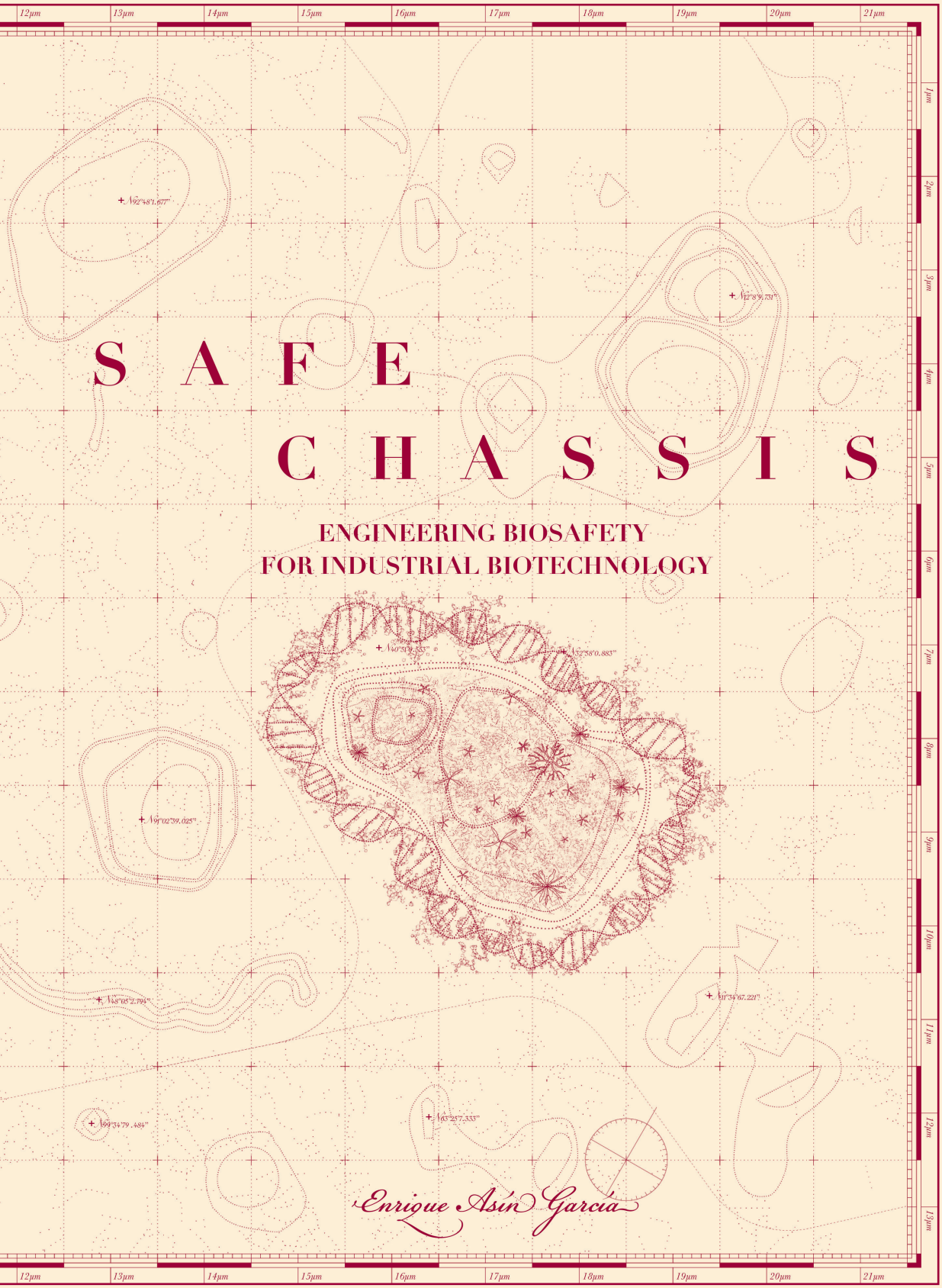
SAFE CHASSIS

ENGINEERING BIOSAFETY FOR INDUSTRIAL BIOTECHNOLOGY



Enrique Asín García

Propositions

1. In absence of experimental validation, any claim of orthogonality is mere speculation.
   (this thesis)

2. Genetic safeguards are yet another object to distract us from the fact that we scientists have no influence on the GMO discussion.
   (this thesis)

3. Automation and artificial intelligence are not only facilitating but also humanizing disciplines.

4. Despite appearances, multidisciplinary research combining life and social sciences remains out of scope for most scientific journals.

5. Transferring information without transferring knowledge is irresponsible.

6. Science is, for better or worse, based on metaphors.

7. In our society both the fool and the wise concur that the latter knows nothing.

8. Both personal and scientific quality and excellence owe much to emotional intelligence.

Propositions belonging to the thesis entitled,
"SafeChassis: Engineering biosafety for industrial biotechnology"

Enrique Asín García
Wageningen, 21 April 2022

SAFECHASSIS: ENGINEERING BIOSAFETY FOR INDUSTRIAL BIOTECHNOLOGY

Enrique Asín García

Thesis committee

Promotor

Prof. Dr Vitor A. P. Martins dos Santos
Professor of Biomanufacturing and Digital Twins, Bioprocess Engineering
Wageningen University & Research

Co-promotor

Dr Zoë Robaey
Assistant professor, Philosophy Group
Wageningen University & Research

Other members

Prof. Dr J. van der Oost, Wageningen University & Research
Prof. Dr V. de Lorenzo, Centro Nacional de Biotecnología, CSIC, Madrid, Spain
Dr L. Asveld, TU Delft, Delft, The Netherlands
Dr M. Guarnieri, National Renewable Energy Laboratory, Golden, Colorado, United States of America

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

SAFECHASSIS: ENGINEERING BIOSAFETY FOR INDUSTRIAL BIOTECHNOLOGY

Enrique Asín García

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A. P. J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Thursday 21 April 2022
at 11 a.m. in the Aula.

Enrique Asín García
SAFECHASSIS: ENGINEERING BIOSAFETY FOR INDUSTRIAL BIOTECHNOLOGY,
386 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2022)
With references, with summary in English

ISBN: 978-94-6447-081-9
DOI: 10.18174/561881

To those who leave their homeland
seeking an opportunity

Contents

1 General introduction and thesis outline 1

2 ReScribe: an unrestrained tool combining multiplex recombineering and minimal-PAM ScCas9 for genome recoding *Pseudomonas putida* 33

3 Metagenomics harvested genus-specific single-stranded DNA annealing proteins improve and expand recombineering in *Pseudomonas* species 73

4 GenoMine: a CRISPR-Cas9-based kill-switch for biocontainment of *Pseudomonas putida* 113

5 Phosphite synthetic auxotrophy as an effective biocontainment strategy for the industrial chassis *Pseudomonas putida* 149

6 Genetic safeguards for Safety-by-Design: so close yet so far 187

7 Context matters: on the road to responsible biosafety technologies in synthetic biology 199

8 Exploring the impact of tensions in stakeholder norms on designing for value change: the case of biosafety in industrial biotechnology 217

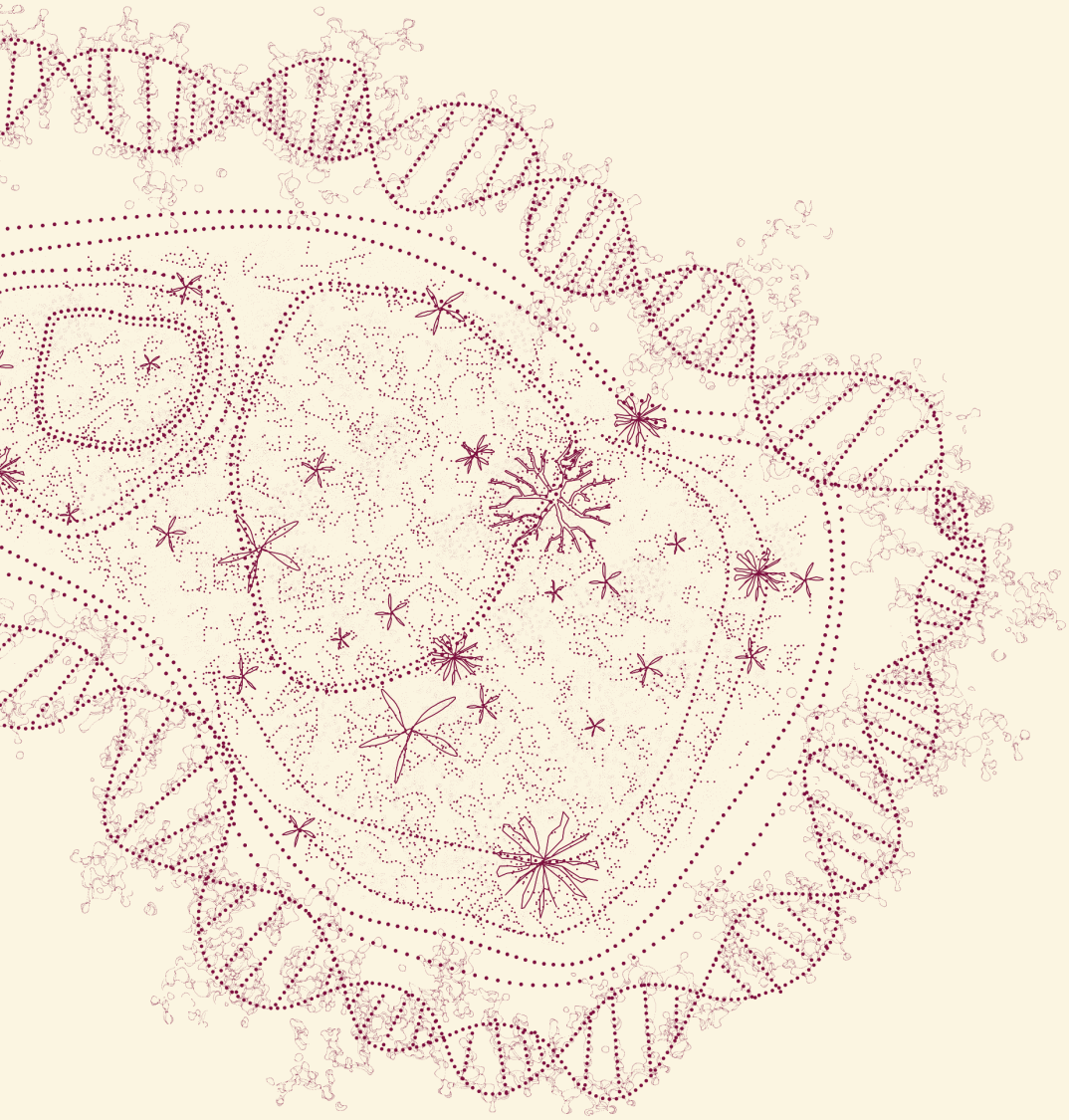
9 Towards an attitude for responsibility in safety education 249

10 Navigating the Valley of Death: perceptions of industry and academia on production platforms and opportunities in biotechnology 277

11 A snapshot of biomanufacturing and the need for enabling research infrastructure 307

12 Thesis summary, general discussion and future perspectives 335

Appendices 365



CHAPTER 1

General introduction and thesis outline

Enrique Asin-Garcia

Preface

This thesis was conceived within the framework of the research project "SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis" (funded by the Dutch Research Organization, NWO). Hereby I address technical and scientific objectives, as well as broader societal ones. The research carried out innovated in three main areas of investigation: genetic safeguards, biosafety and biotechnology and biomanufacturing. In this introduction, I sketch the context in which this investigation took place by guiding the reader through important concepts and historical considerations that have shaped the *status quo* of the research fields encompassed within this dissertation. First, genetically modified organisms, generated with metabolic engineering and synthetic biology, are presented as valuable assets in the transition towards the bioeconomy. Later, I argue that to fulfil the promise of their potential, engineered organisms need to be, among other things, safe in all its forms. Here, I focus on mitigating biosafety risks mainly related to undesired propagation of microbes and horizontal gene transfer, which might be achieved by means of biosafety measures and biocontainment strategies. Different types of genetic safeguards are then described, and ultimately, *Pseudomonas putida* is introduced as a microbial platform that could certainly benefit from the implementation of these safe-by-design approaches on its way to becoming a relevant synthetic biology chassis. The introduction culminates with a list of objectives and an outline, where I describe the motivations behind this thesis and its corresponding chapters.

Genetically modified organisms and their role towards a sustainable bioeconomy

A brief history of the bioeconomy

Life on Earth has always been full of uncertainties and concerns. Globalisation and interconnectivity, however, have exacerbated these to a higher range of systemic risks, such as pandemics, economic shockwaves, underlying disparities, societal fragmentation, and failure in climate change mitigation [1]. Last century's technological innovations motivated the economic growth that raised the living standards and erased the boundaries amongst societies. Alas, they also triggered a chain reaction of negative side effects on ecosystems, biodiversity and climate [2]. Our current economic growth model is directly proportional to the levels of energy consumption, which in turn is associated to different kinds of environmental burdens including waste production, water consumption and greenhouse gas emissions [3]. Thus, to change this dispiriting equation, a new model that decouples economic development from resource scarcity and environmental issues is desperately required.

In an attempt to cope with the ecological, social and economic sustainability challenges (as well as to maximize profits by valorising waste material [4]), the concept of bioeconomy was developed towards the end of the 20th century based on the use of bio-based resources, recirculated products and processes, and renewable energy [5–8]. The European Commission defines it as the knowledge based "production of renewable biological resources and the conversion of these resources and waste streams into value added products, such as food, feed, bio-based products and bioenergy" [9, 10]. By avoiding fossil carbon, this model intends to transition into a more sustainable and resource-efficient world with a low-carbon footprint [11]. By developing in concert all sectors of the bioeconomy, nature, economy and society, prominent global objectives could be satisfied including food security, improved nutrition and health, cleaner and more efficient industrial processes and a significant contribution to the effort of mitigating climate change. Thus, we should aspire to a situation where all three sectors are equally strong, properly linked and inter-connected, with the objective of creating

a prosperous bioeconomy web, able to yield the maximum benefit [12].

The growth resulted from the bioeconomy goes hand in hand with the evolution of the biotechnology sector [12–14]. In turn, biotech-bioeconomy development is rooted in the extension of "traditional biotechnologies" towards the engineering of biological systems [15, 16]. The ability to modify organisms has been possible due to scientific breakthroughs that led to the inception of disciplines such as synthetic biology and metabolic engineering.

Metabolic engineering, synthetic biology and biomanufacturing

Manipulating organisms for human benefit is no recent matter. Our ancestors selected and crossed the best and most productive plant and animal species. Later, with the advent of microbiology, microorganisms that naturally produced desired molecules were identified and then improved through classical strain engineering [17]. Later, in the second half of the 20th century, a flourishing collection of molecular biology tools encompassing restriction enzymes, plasmids and markers opened the door to bioengineering and the reprogramming of gene networks and whole organisms [18, 19]. The first applications of this technology were aimed at producing chemicals in an efficient manner using microorganisms, which resulted in the coined term "cell factories" [20, 21]. The design and construction of these cell factories brought together molecular biologists and chemical engineers to set the foundations of the metabolic engineering field [22–24]. This research field can be defined as "the development of methods and concepts for analysis of metabolic networks, typically with the objective of finding targets for engineering of cell factories" and involves strategies that aim at improving *e.g.*, the range of substrates and products, titre, rate and yield (TRY) of a biotransformation, cellular robustness, or tolerance towards toxic compounds or fluctuating conditions [17].

The search for a tighter control and steering of chemical production, motivated the generation of the first genetic circuits inspired on electronic microcircuit design, which in turn were one of the first applications of what we know today as synthetic biology [25, 26]. Useful for metabolic engineering purposes and for many more biotechnological endeavours, this field encompasses the "application of science, technology and engineering to facilitate and accelerate the design, manufac-

ture and/or modification of genetic materials in living organisms" [27]. Since its emergence, synthetic biology has evolved as an umbrella concept that covers the rational design of biological systems by realistically incorporating engineering principles while stepping away of the previous *ad hoc* and serendipitous practices common in traditional biotechnology [28]. In this way, this approach opts for robustness and reliability, and evolves towards simplification, automation and scalability of biological processes [29]. An increasingly expanding landscape of companies, organizations and funding agencies has turned synthetic biology into a global enterprise [30]. Based on the promise of disruptive technological advances, the field enjoys the support from governmental and funding bodies to thrivingly access the market [29, 31], where it is expected to reach a stake valued at US$2-4 trillion of annual direct economic potential globally in 2030-40 [32]. In the same vein, associated enabling technologies such as DNA sequencing, DNA synthesis, genome editing or mathematical modelling have become incredible profitable both in economic and knowledge-generation terms [33, 34].

Adhering to engineering principles as indicated above, the core premise of synthetic biology encompasses foundational concepts such as standardization of biological constructs and functions, modularity of parts, and abstraction of elements into individual devices [35, 36]. However, the engineering approach of assembling a fully functional system by simply "plugging and playing" standard modular parts is not always feasible due to the many uncontrolled variables and interrelationships occurring within a living system [37]. As a consequence, a more thorough engineering concept has been adopted during the last years to discover, design and optimise bioengineered systems: the design-build-test-learn (DBTL) cycle [38–40], which had also been incorporated by other non-biological manufacturing fields before. Iterations of the DBTL cycle accelerate and reduce the cost of discovery and optimization of desired target solutions [29], being the most immediately recognized and advanced application the improvement of the TRY credentials of a microbial strain [41] (Figure 1.1).

At this stage, and next to other research disciplines including the aforementioned metabolic engineering, systems biology, cascade biocatalysis, etc., synthetic biology can already claim a contribution to a new stage of biomanufacturing [42–

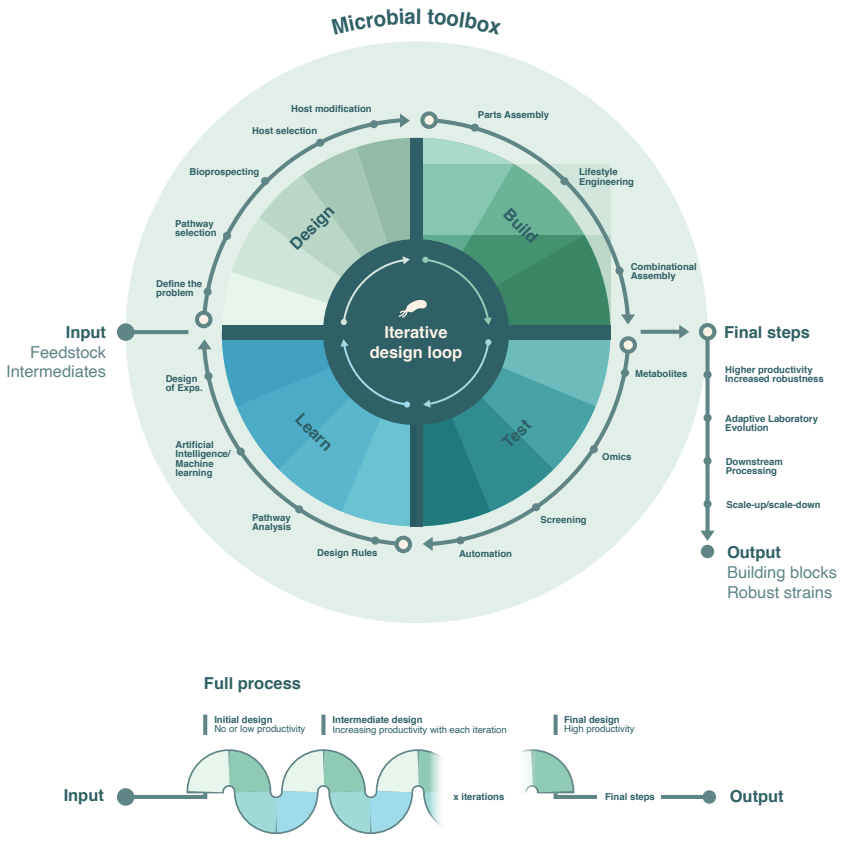


Figure 1.1: Schematic representation of the deployment of the design-build-test-learn (DBTL) cycle for an exemplary re-design of microbial catalysts for the microbial conversion of feedstocks and intermediates into building blocks (from the NWA-ORC project proposal of HARNESS – Hybrid Autotrophic Platform for Sustainable Polymers for a Clean, Circular Economy).

46]. The current challenges, together with the flourishing set of innovation tools have stimulated a revolution in the biomanufacturing field that has now the opportunity to generate not only new products, but also existing ones in a more effective fashion in terms of TRY, productivity, robustness, scale-up feasibility and sustainability [40, 47, 48]. Beyond biomanufacturing, synthetic biology holds the promise for addressing other global needs related to health, material and data science [49], encompassing applications such as therapeutic genome editing [50], diagnostics [51], living biotherapeutics [52], biocomputing [53], electronic interfacing [54], living materials [55] and cellular recording [56].

Beyond the promise

Despite all these advances, a series of challenges slow down the progress of the field, including technology transfer and scale-up considerations. In this regard, efforts must be made towards designing increasingly complex cell functions in robust microbial chassis with more precisely predictable functionalities [57]. This relates to the idea that most developments are currently not translatable to "outside-the-lab" applications, where the conditions are not as well-controlled as they are in laboratory settings, and long-term storage stability as well as resource-limited and off-the-grid scenarios end up hindering the operation [49]. Another main issue is the lack of sufficiently abundant and accurate biological information at the levels of genome, transcriptome and metabolome, beyond the few main model organisms. New proteins, pathways and even biological phenomena are still out there, unknown to the community, with an unexplored vast potential for new processes and innovations [58]. Closely related to this is the necessity of having genetic tools and devices to obtain information. Genetic manipulation of non-model organisms remains challenging as a consequence of organism-specific nuances which hinder universal tools and translatable biochemical and regulatory knowledge [59]. Model or non-model, there is a strong need for quantitative data from biological systems [60]. High expectations are placed on further characterization of biological components and genetic networks, alongside mathematical and computational models, which will hopefully provide us with a better understanding of the systems and phenomena, as well as their dynamics [61].

Just as the overarching biotechnology field, synthetic biology also faces conceptual, societal and ethical challenges [62, 63]. Conceptually speaking, as a relatively considered emerging technology, the field holds a transformational potential because of its capacity to produce changes in social, commercial and physical environments [64]. As a consequence, it involves uncertainty regarding the possible outcomes [65, 66] and ambiguity in regard of legal regimes, codes of conduct and embedded beliefs and values associated to the corresponding practices, products and outcomes [67]. Moreover, the research conducted in synthetic biology has implications in concepts as fundamental as the ones of life and nature [68–70], both from the deterministic [71] and the allegorical [72] points of view.

Underlying or in addition to all those concerns stand the scientific, legal and ethical risks and uncertainties related to the safety and security of synthetic biology, which are associated with its impact on the environment, biological diversity as well as human health [73, 74]. With the advancement of synthetic biology, more and more questions have been raised about the adverse effect that synthetic microbes might have if more widely used or applied under non-contained conditions [75–78]. Could synthetic or genetically engineered microorganisms (GEMs) proliferate outside their optimal conditions? Could they outcompete native species and disrupt ecosystems? Could their synthetic genetic material be transmitted to other hosts? What would be the resulting effects of these situations and what can we do to prevent them? The following sections of this introduction focus on biosafety aspects of synthetic microorganisms and how to address these and other issues from a synthetic biology-oriented perspective.

Utility of genetically engineered microorganisms (GEMs) is dependent on the development of secure biological designs

Risk history of GEMs

Biosafety concerns of synthetic biology echo old issues that come back to the first days of recombinant DNA [79]. Scientists at the 1975 Asilomar conference established a series of guidelines based on caution and founded on the incorporation of both physical and biological containment into the experimental designs of these technologies [80]. These first agreements were aimed at minimizing environmental risks of the use of cisgenics and transgenics (*i.e.*, alteration of sequences within a native host or transference to another species, respectively) and included suggestions such as the implementation of metabolic auxotrophies. Almost five decades later, this foundation still endures and, consequently or not, we have not witnessed any significant disaster [77, 81]. Two main reasons have ensured the success of the Asilomar approach. In the first place, there are generally stringent regulations that confines the use of GEMs to contained settings, limiting the opportunities for derived risks [77]. In the second place, we typically find a rather poor fitness and effectiveness of GEMs outside their targeted conditions that prevents them from

establishing themselves in a meaningful way in wild habitats [82–85].

Synthetic biology, nonetheless, is poised to take a step forward. From synthetic genomes [86, 87] to non-canonical components and alternative biochemistries [88–91], and passing by a growing ability to edit and synthesize *de novo* DNA [92], the field is yielding game-changing GEMs that have little to do with those of half century ago. On top of that, and as mentioned above, robustness has become one of the top-priorities in engineering microorganisms, so that they are better prepared to perform under industrial conditions [93–95]. However, this increased robustness may increase the chances of generating microbes capable of proliferating under non-targeted conditions (*e.g.*, an accidental release). According to these facts, uncontrolled propagation of more robust and radically engineered synthetic organisms, in addition to the possibility of transferring transgenic alien DNA through horizontal gene transfer (HGT) could, in principle, lead to an expanded list of risks including, for example, the generation of novel pathogens [96–99] (Figure 1.2). As a consequence, many groups, including governmental bodies, non-profit research organizations, industry associations and third-party advocacy groups, have examined, during the last years, the implications for matters of risk, uncertainty, safety, and governance, and incorporating the field's distinctive aspects into their oversight [100–103]. In their comprehensive review, Hewett *et al.* identified 44 discrete risks of synthetic biology. Out of the 44, 18 pertained to human health and were related to allergies, antibiotic resistance, carcinogens and pathogenicity; and 26 applied to the environment, including those linked to changes of the environments, competition with native species, HGT and toxicity [102]. Yet, the full risks associated with the spread of this type of GEMs cannot be completely ascertained. It is precisely this very uncertainty that makes it important to implement control and regulation over synthetic microorganisms [99].

Biosafety and biocontainment

Biosafety, in general, can be defined as the set of "containment principles, facility design, practices and procedures to prevent occupational problems or the release of the engineered organisms to non-permissive environments" [105–107]. Detached from the malicious connotations of biosecurity, other definitions of biosafety fo-

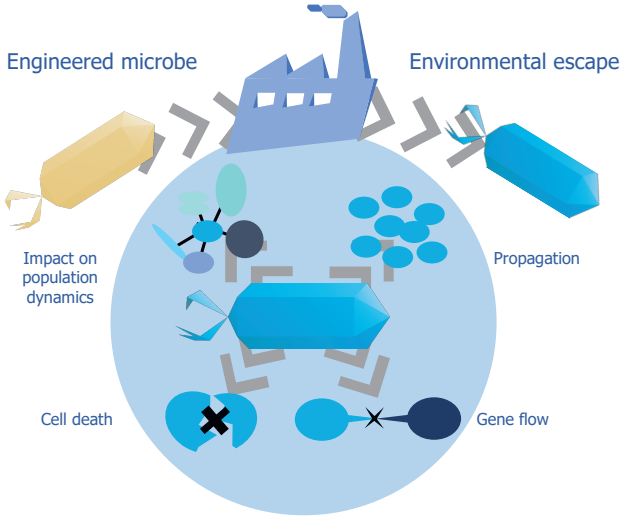


Figure 1.2: Potential environmental fates upon escape of GEMS in native ecosystems (adapted from Arnolds et al., 2021 [104])

cus directly on the risks and the capabilities to cause disease of greater or lesser severity in other organisms or disruption of ecosystems [108, 109].

Legislation regarding biosafety and handling of biological materials has been established by different organizations and conventions both on the international and the regional level [109, 110]. Despite this, regulatory differences among countries have led to new ways of thinking about biosafety [111]. One of such new ways is the Safe-by-Design (SbD) concept, a preemptive approach to risk management (stemming from construction and other traditional engineering disciplines [112]) that aims at minimizing risks by making safer choices from the early stages of an innovation, preferably at the R&D and design phases [113]. On top of the physical containment barriers, in synthetic biology, SbD can also materialize at different biological levels: from choosing the right host organism to designing built-in-safety biocontainment strategies or genetic safeguards [114]. Since the times of Asilomar, new biocontainment methods have been harnessed by emerging technologies in order to keep up with the levels of safety required by the new practices [115].

As a general rule, genetic safeguards aim to target maximal containment with minimal impact on fitness [104]. In order to achieve this, a series of key require-

ments must be met [115]. First, the system must be robust to prevent the escape of even a small fraction of the population from the permissive conditions. This is assessed by means of the escape frequency. In addition, the system needs to be stable in the long term, given the many generations that GEMs are expected to go through during their application. To maintain this stability, attention must be paid to several intrinsic cell factors including DNA recombination, mutagenesis and metabolic burden of the biocontainment system itself, which might disrupt the safeguard apparatus resulting in its inactivation. Furthermore, a biocontainment strategy should be ideally customizable in an easy and rapid fashion, to be able to implement it on demand in a variety of strains, applications and environments [77, 78, 115].

Types of genetic safeguards

A considerable collection of genetic safeguards has been developed with two objectives in mind: biological and genetic isolation. Biological containment aims at conditionally restricting the host cell viability or its ability to proliferate to a permissive environment. In turn, genetic containment intends to prevent the distribution of engineered or synthetic genetic material between organisms and across species, which typically occurs *via* HGT [78, 99, 115–117].

One of the most classic and common strategies is to engineer auxotrophies in GEMs, so these become dependent on externally provided molecules. By removing their capacity to synthesize or incorporate essential metabolites, these organisms will only be able to survive and proliferate under those permissive conditions where these molecules are supplied. The history of metabolic auxotrophies comes back to the times of Asilomar [99, 118]. Since then, strategies have evolved from dependencies on nucleosides [119, 120] and amino acids [121] to reliance on non-canonical [122], synthetic [123] and naturally almost unavailable molecules [124]. While the former strategies were easily overcome as a consequence of environmental availability, cross-feeding and metabolite complementation, the latter have managed to improve considerably the strength of auxotrophic-based biocontainment reaching some of the lowest levels of escape frequency ever reported. Alas, this type of strategies does not consider genetic isolation. In the first place, engineered and

synthetic genes necessary to implement the safeguard apparatus could potentially be transmitted to natural species *via* HGT. By the same token, the strategy could be inactivated by reacquisition of genes that eliminate the metabolic dependency [115, 117].

Another option is the use of genetic circuits for biocontainment. In this case, the viability or the proliferation of a GEM is subjected to specific physicochemical or nutritional conditions which control synthetic regulatory gene modules [125]. Under permissive conditions, these modules are able to detect the presence of environmental signals and translate them into a survival response which can be mediated by toxin-antitoxin balance [126], transcriptional or translational regulation of either toxic or essential genes [127, 128], or an equilibrated mixture of signals [129]. A myriad of mechanisms has been presented during the last years adopting many different complexities and responding to a great diversity of inputs, including natural inducers [130], synthetic molecules [131], temperature [132], pH [133], etc. This flexibility represents the biggest advantage of this type of safeguards, which can be tailor-made or easily adapted for each specific application [115]. As a downside, most genetic circuits have shown a rather poor long-term stability as a result of deactivation caused by recombination and mutagenesis [134]. Thus, their escape frequency generally tends to go up after a few generations.

Furthermore, genetic safeguards can be based on semantic containment. In this revolutionary approach, the conserved genetic code of the cells is altered to prevent the correct expression of genes if they are transmitted *via* HTG [116, 135, 136]. The new genetic information encompasses either non-canonical DNA [137–140] or non-canonical amino acids encoded by recoded codons [141]. The machinery to synthesize these unnatural building blocks can be engineered into the cell [142]. Otherwise, they need to be provided externally when incorporated in essential genes [143], which in turn constitutes and additional layer of auxotrophic containment. Altered or recoded genomes, achieved either *via* rational mutagenesis [136] or *de novo* synthesis [144], can impart new properties beyond biocontainment such as virus resistance, new or improved functions achieved through alternative biochemistries, the means to study fundamental biological concepts related to global transcription or translation, etc [145]. Unfortunately, the implementation of this

type of strategy requires either an extensive genome editing effort supported by extremely efficient genetic tools or a significant economic investment, which considerably limit its availability [115].

Biocontainment today

Biosafety and biocontainment strategies feature an outstanding place in the synthetic biology literature. These publications, however, differ not only in the way containment is assessed, but also in the way it is reported. Currently, the risk assessment of strains equipped with genetic safeguards is mostly performed qualitatively, where the probability of an adverse outcome is considered as more or less likely than a comparative scenario, which is typically the wild type strain [146]. At this point, it is very important to stress that there is a great need for standardized metrics beyond the mere calculation of fitness and escape frequencies [147]. Although informative, these two factors are limited for several reasons. First, fitness is typically inferred from growth curves obtained from optical density (OD) measurements. These do not distinguish living from dead cells, which might end up corrupting the true survival ratio. In addition, experiments to assess escape frequency are commonly performed in standard-size Petri dishes, which limits the sensitivity of the assay. This restriction originates from the number of colony forming units (CFU) that are cultured onto non-permissive plates which should not exceed 10^{11}-10^{12} CFU. Beyond these numbers, cells are expected to clump impeding an accurate evaluation of the plates and hampering the calculation of the ratio between CFU under permissive and non-permissive conditions [99]. As an alternative, lower degrees of escape frequency could be detected by means of a morbidostat, where a population over 10^{16} can easily be exposed to the non-permissive conditions [148]. Still, new techniques and metrics to evaluate other parameters such as stability in permissive conditions [149], or the degree of HGT probability [150] would also be informative and useful.

Either way, assessment data is disparate in the reporting conditions: different media, different contextual circumstances, etc [73]. Hence the need for standard operating procedures (SOPs) that incorporate benchmarks and best practices to enable genuine comparisons among strategies [147]. These inconsistencies might

be one of the major reasons why genetic safeguards are seldom applied in real world applications. What else is preventing these technologies from moving forward beyond the research laboratories? Is it only due to technical reasons? Or are there other underlying motives holding them back?

At this point, it is worth mentioning that genetic safeguards cannot be designed to fulfil their purpose under every conceivable condition, meaning that the available risk assessment data is not representative for most scenarios. In addition, they have been mainly devised in specific chassis, most of them in the model bacterium *Escherichia coli*, which is not always the most suitable workhorse [151–154]. In this context, the utopian idea of a single and perfect synthetic biology chassis for every possible application has been replaced by the more realistic concept of a defined repertoire of characterized and standardized chassis for their corresponding best suited scenarios [147, 155]. In this way, specific hosts equipped with concrete genetic safeguards and intended for a series of fitting applications will need to meet distinct and less ambiguous specifications resulting in a more authentic and pragmatic risk assessment. Particularly, the next section highlights *Pseudomonas putida* as an example of synthetic biology chassis in which genetic safeguards could be implemented to maximize its value as an industrial and environmental platform.

Endowing *P. putida* with genetic safeguards: a SafeChassis for industrial biotechnology and biomanufacturing

P. putida as a chassis

The soil and water Gram-negative bacterium *Pseudomonas putida* KT2440 has gained popularity as a synthetic biology chassis for metabolic engineering purposes with potential for industrial and environmental applications. This is typically justified with a series of credentials that include fast growth in the laboratory, high tolerance to solvents, endurance to stress and genetic accessibility [156, 157]. But how do these characteristics actually benefit *P. putida* KT2440 in synthetic biology applications? A key feature of an ideal synthetic biology chassis is its ability to host recombinant constructs in a stable and durable way. Genetic devices often compete with the host's cellular machinery for the cell's resources, and sometimes they

even trigger a stress response caused by a high cellular burden or toxic by-products (*e.g.*, reactive oxygen species, ROS) [158–160]. As a consequence, these introduced devices end up being mutated or lost. In this regard, *P. putida* KT2440 presents two major advantages for the implantation of genetic devices, intrinsically connected to its core metabolic background [147]. In the first place, the performance of the SOS scheme is poor in this bacterium as a consequence of a weak RecA-based DNA repair system [161]. This prevents a general mutagenesis response when the DNA is damaged and the removal of foreign DNA *via* recombination. In the second place, this organism presents a high rate of endogenous NAD(P)H regeneration due to the EDEMP overflow metabolism on hexoses [162], which allows an effective quenching of the ROS impact on the DNA through misincorporation of 8-oxoguanidine alleviating their toxic effect [163]. Moreover, *P. putida* is very flexible and adapts quickly to steady changing conditions due to its powerful regulatory apparatus [164, 165], which results tremendously advantageous in large-scale fermentations where the conditions are not always strictly homogeneous [166]. Lastly, the microbe naturally presents a versatile catabolism of carbon sources, which has been synthetically expanded even further [167]. All these attributes foster the implementation of not only routes for biodegradation and bioproduction, but also other exogenous genetic devices such as genetic safeguards.

The acquisition of knowledge about the genetics, biochemistry and physiology of *P. putida* has been relentlessly advancing during the last half century culminating in major efforts such as the complete annotation of its genome [168, 169] or the construction of genome-scale metabolic models [170, 171]. In parallel, most wet synthetic biology tools have also been implemented in *P. putida* [155]. Together with its intrinsic capabilities, these expanding knowledge and technology set the stage for using *P. putida* as a chassis in industrial applications. Currently, this microbe has already proven to be an excellent bacterial host for the production of polymers, bulk chemicals, drugs and high added-value compounds (Figure 1.3) [167], with a range of applications that increases day by day [172].

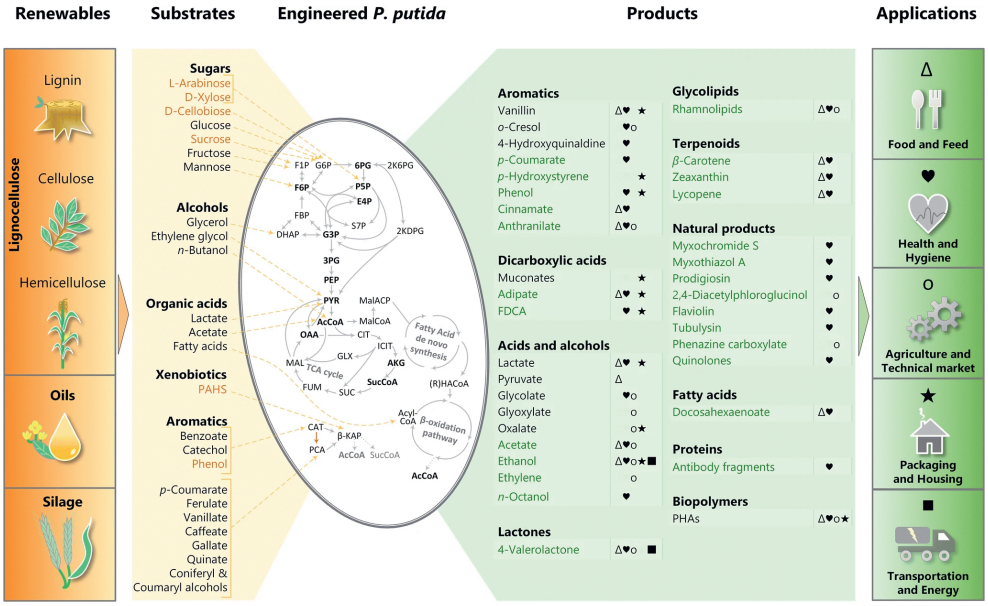


Figure 1.3: Bioproduction using engineered P. putida strains. Substrates generated from renewable feedstocks, such as lignocellulose, oils and silage can be used to produce value added products for application in the food and feed (triangle), health and hygiene (heart), packaging and housing (star), transportation and energy sector (square), and for agriculture and technical application (circle). Entry points of the respective substrates in the metabolism (dashed orange line), new to P. putida substrates (orange) and products (green) (from Weimer et al., 2020 [167])

P. putida as a safe chassis: building in increased safety

For all the reasons mentioned in the previous section, another important feature when considering a synthetic biology chassis for real world applications is its biosafety level. *P. putida* KT2440 is classified by the FDA as host-vector 1 (HV1) certified, indicating that it is safe to use in a P1, BSL1 or ML1 environment with no special containment measures [173]. While different agencies and institutions accredit safety in different ways and through different certifications [174], in most scenarios, this status is granted based on a historical absence of safety concerns, which is supported by years of study, comprehensive knowledge and lack of pathogenic properties [147]. The aforementioned synthetic biology concepts and technologies, however, have the power of enabling to modify the *P. putida* KT2440's nature in a game-changing profound way, providing the means for alternative biochemistries, metabolisms and properties in general. As a consequence, the reductionistic con-

Table 1.1: Previous biocontainment strategies to prevent escape in P. putida.

Study	Year	Toxic gene	In response to	Escape frequency
(Bej *et al.*, 1992) [175]	1992	*hok / gef*	IPTG	10^{-5}
(Munthali *et al.*, 1996) [176]	1996	*colE3*	3-methyl benzoate	N/A
(Szafranski *et al.*, 1997) [177]	1997	streptavidin	3-methyl benzoate	10^{-7}
(Molina *et al.*, 1998) [178]	1998	*gef*	3-methyl benzoate	10^{-8}
(Ronchel and Ramos, 2001) [121]	2001	*gef*	3-methyl benzoate	$<10^{-9}$

cept of assigning the ancestor's biosafety status to a radically engineered strain should be approached with caution.

Both to enhance the biosafety levels in current applications, and to provide a layer of containment in those potential applications lacking any means of physical containment, genetic safeguards represent a valuable addition to strengthen the significance of *P. putida* KT2440 as a chassis. Back in the 1990s, several biocontainment strategies were already developed in this organism with promising results (Table 1.1). In all those cases, population control was based upon the expression of a toxic gene in response to the common inducers IPTG or 3-methylbenzoate. This type of systems is susceptible to escape due to evolutionary mutations in the toxin's open reading frame (ORF), regulatory mechanism or target, or due to acquisition of resistance genes. Moreover, any trace of leakiness results in system's instability under non-induced conditions, an obstacle that nonetheless can be tackled by adding the corresponding antitoxin to counteract those undesired levels of leaked toxin [127, 132].

With the advent of synthetic biology, robust genetic circuits with more complex and programmable functionalities emerged overcoming some of the limitations of prior biocontainment strategies [128]. Unfortunately, the toolbox of *P. putida* KT2440 is still poor in this regard, missing some key elements such as riboswitches and riboregulators that would enable tight control and dynamic ranges of gene expression in response to ligand concentrations [155]. Even the most advance circuits are however still unable to prevent the issue of evolution, which we can only try to avoid by engineering quick and highly toxic responses under non-permissive conditions minimizing in this way the chances for evolution that arise under pressure [179], as well as the problem of HGT. The safest choice for biocontainment of *P. putida* would be in this sense one that grants both biological and genetic iso-

lation, such as genome recoding [147]. While the ability to rewrite genomes has revolutionized the field, it necessitates of the well-functioning of genome editing technologies for an optimal performance [145]. Ideally, these techniques would be able to perform efficient genome engineering in a multiplex manner, given the high number of mutations required for an endeavour of this magnitude [136]. Due to this complexity, the development of additional and customized *P. putida* strains equipped with different genetic safeguards remains also an important matter for achieving high biosafety levels should these be required for safer applications of today and tomorrow.

Aims of the dissertation and outline

The overarching goals of my research project are to investigate, develop, implement and assess genetic safeguards that would render starkly engineered microbial chassis intrinsically safe(-by-design) for industrial biotechnology applications. To this end, I aim to address both technical and non-technical aspects underlying these goals. Hence, I formulated the following specific objectives: a) to develop multiplex genome engineering technologies into *P. putida* to enable whole-genome recoding, b) to construct a proof-of-concept minimally recoded genome of *P. putida* by engineering a TAG— strain, c) to develop additional robust intrinsic biocontainment strategies in *P. putida*, d) to explore the different reasonings and understandings behind the concept of genetic safeguards for SbD and risk assessment, e) to examine biosafety as a value from a perspective of responsibility, and f) to ascertain the current opportunities and future prospects in the sectors of industrial biotechnology and biomanufacturing. With these objectives in mind, the strategy depicted in Figure 1.4 was developed around three main pillars: genetic safeguards, biosafety and biotechnology and biomanufacturing.

Every pillar of the project consists of a series of self-contained studies addressing different aspects of the corresponding theme. The first pillar after this general introduction (chapter 1) includes the implementation of different types of genetic safeguards into *P. putida* KT2440 (chapters 2 to 5). In chapter 2, I developed (jointly with my collaborators) the ReScribe genome editing method, a highly optimized re-

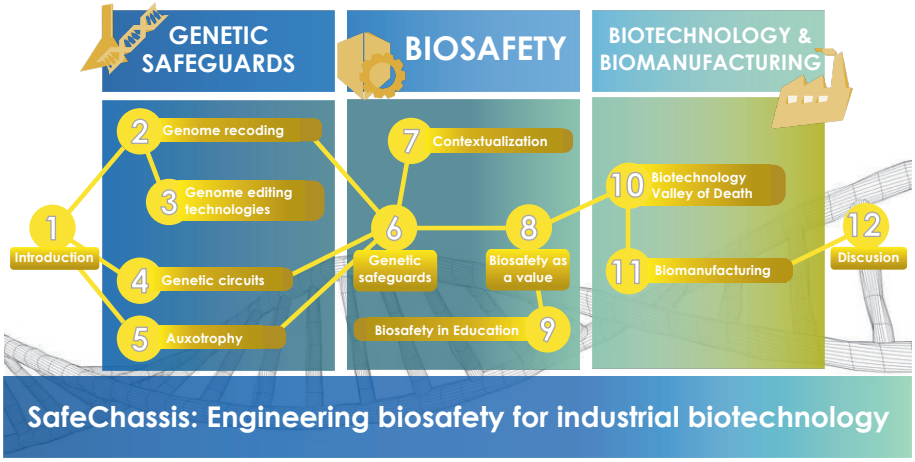


Figure 1.4: Structural contents of this dissertation, based on three pillars: Genetic Safeguards, Biosafety and Biotechnology & Biomanufacturing.

combineering tool enhanced by CRISPR-Cas9-mediated counterselection to generate a minimally recoded strain, setting the basis for the establishment of semantic containment in this bacterium. In an attempt to find even more efficient recombineering systems not only for *P. putida* and but also for other relevant *Pseudomonas* species, chapter 3 describes a high-throughput workflow for screening protein libraries, which allowed us to identify new effective recombinases for enhanced mutagenesis in several species. In chapter 4, I implemented different genetic circuits in *P. putida* that react to exogenous signals with the objective of controlling a CRISPR-Cas9-based kill switch that activates a highly genotoxic response. Lastly, in chapter 5, I generated an industrially appealing synthetic phosphite auxotroph attained *via* metabolic rewiring.

The following pillar (chapter 6 to 9) encompasses investigations of various concepts surrounding the topic of biosafety. Chapter 6 questions the actual role that genetic safeguards play in the biosafety ecosystem and points out the practical bottlenecks the impede their integration in practice, as well as potential solutions. Chapter 7 delves specifically into one of these recommendations: an explicit strategy of contextualization. Here, I propose that an early emphasis on potential ap-

plications can assist the development of genetic safeguards. In chapter 8, I turn to a more philosophical vision of biosafety as a value and analyse how the tensions in stakeholder norms impact the way of designing for biosafety in the industrial biotechnology context. Furthermore, chapter 9 contains an investigation on what aspects of safety are taught and how these are imparted in life sciences programmes of higher education institutions.

In the next pillar (chapters 10 and 11), the focus is shifted to the fields of biotechnology and biomanufacturing. Chapter 10 collects perceptions of industrial partners and academics on production platforms and opportunities in the biotechnology field, which were surveyed by means of a series of in-depth interviews. In addition, chapter 11 provides a snapshot of the state of biomanufacturing and includes a proposition of the idea of meta-workflows as unifying ecosystems amongst research infrastructures for a successful transition to global manufacturing that utilizes biobased and safe production platforms. Finally, the knowledge gathered up to this point was subsequently used for discussing the state of affairs and the future perspectives in chapter 12.

Bibliography

1. World Academic Forum. *The Global Risks Report 2021 16th Edition* 2021.
2. Lopes, M. S. G. Engineering biological systems toward a sustainable bioeconomy. *Journal of Industrial Microbiology and Biotechnology* **42,** 813–838 (2015).
3. Kasman, A. & Duman, Y. S. CO2 emissions, economic growth, energy consumption, trade and urbanization in new EU member and candidate countries: A panel data analysis. *Economic Modelling* **44,** 97–103 (2015).
4. Vivien, F. D., Nieddu, M., Befort, N., *et al.* The Hijacking of the Bioeconomy. *Ecological Economics* **159,** 189–197 (2019).
5. Enriquez, J. *The Biotechonomy (1.0): A Rough Map of Bio-data Flow* (2002).
6. OECD. The Bioeconomy to 2030: Designing a Policy Agenda: Scoping Paper (2006).
7. Geoghegan-Quinn, M. *Innovating for sustainable growth: A Bioeconomy for Europe* in. Press conference, Brussels. **13** (2012).
8. Bell, J., Paula, L., Dodd, T., *et al.* EU ambition to build the world's leading bioeconomy—Uncertain times demand innovative and sustainable solutions. *New biotechnology* **40,** 25–30 (2018).
9. Commission, E. *Innovating for Sustainable Growth: A Bioeconomy for Europe: Communication from the Commission to the European Parliament, the Council and the European Economic and Social Committee and the Committee of the Regions* (2012).
10. Patermann, C. & Aguilar, A. The origins of the bioeconomy in the European Union. *New biotechnology* **40,** 20–24 (2018).
11. Carus, M. & Dammer, L. The Circular Bioeconomy—Concepts, Opportunities, and Limitations. *Industrial Biotechnology* **14,** 83–91 (2018).
12. Lokko, Y., Heijde, M., Schebesta, K., *et al.* Biotechnology and the bioeconomy—Towards inclusive and sustainable industrial development. *New Biotechnology. Bioeconomy* **40,** 5–10 (2018).
13. Oborne, M. The bioeconomy to 2030: designing a policy agenda. *Organisation for Economic Cooperation and Development. The OECD Observer,* 35 (2010).
14. Philp, J. The bioeconomy, the challenge of the century for policy makers. *New biotechnology* **40,** 11–19 (2018).
15. Bud, R. Biotechnology in the twentieth century. *Social studies of science* **21,** 415–457 (1991).
16. Befort, N. Going beyond definitions to understand tensions within the bioeconomy: The contribution of sociotechnical regimes to contested fields. *Technological Forecasting and Social Change* **153,** 119923 (2020).
17. Nielsen, J., Fussenegger, M., Keasling, J., *et al.* Engineering synergy in biotechnology. *Nature Chemical Biology* **10,** 319–322 (2014).
18. Beckwith, J. R. Regulation of the Lac Operon: Recent studies on the regulation of lactose metabolism in Escherichia coli support the operon model. *Science* **156,** 597–604 (1967).
19. Westerhoff, H. V. & Palsson, B. O. The evolution of molecular biology into systems biology. *Nature Biotechnology* **22,** 1249–1252 (2004).
20. Bailey, J. E. Toward a science of metabolic engineering. *Science* **252,** 1668–1675 (1991).
21. Stephanopoulos, G. & Vallino, J. J. Network rigidity and metabolic engineering in metabolite overproduction. *Science* **252,** 1675–1681 (1991).
22. Nielsen, J. Metabolic engineering. *Applied microbiology and biotechnology* **55,** 263–283 (2001).
23. Stephanopoulos, G., Alper, H. & Moxley, J. Exploiting biological complexity for strain improvement through systems biology. *Nature biotechnology* **22,** 1261–1267 (2004).

24. Keasling, J. D. Manufacturing molecules through metabolic engineering. *Science* **330,** 1355–1358 (2010).

25. Chen, W., Kallio, P. T. & Bailey, J. E. Construction and characterization of a novel cross-regulation system for regulating cloned gene expression in Escherichia coli. *Gene* **130,** 15–22 (1993).

26. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403,** 339–342 (2000).

27. European Commission. Directorate General for Health and Consumers. *Opinion on synthetic biology I: definition.* (LU, 2014).

28. Flores Bueso, Y. & Tangney, M. Synthetic Biology in the Driving Seat of the Bioeconomy. *Trends in Biotechnology* **35,** 373–378 (2017).

29. Clarke, L. J. & Kitney, R. I. Synthetic biology in the UK–an outline of plans and progress. *Synthetic and systems biotechnology* **1,** 243–257 (2016).

30. Oldham, P., Hall, S. & Burton, G. Synthetic Biology: Mapping the Scientific Landscape. *PLOS ONE* **7,** e34368 (2012).

31. Council, N. R. *Positioning synthetic biology to meet the challenges of the 21st century: summary report of a six academies symposium series* (2013).

32. Chui, M., Evers, M., Manyika, J., *et al. The Bio Revolution Innovations transforming economies, societies, and our lives* (McKinsey Global Institute, 2020).

33. Manyika, J., Chui, M., Bughin, J., *et al. Disruptive technologies: Advances that will transform life, business, and the global economy* (2013).

34. Kosuri, S. & Church, G. M. Large-scale de novo DNA synthesis: technologies and applications. *Nature methods* **11,** 499–507 (2014).

35. Endy, D. Foundations for engineering biology. *Nature* **438,** 449–453 (2005).

36. Gardner, T. S. & Hawkins, K. Synthetic Biology: evolution or revolution? A co-founder's perspective. *Current Opinion in Chemical Biology. Synthetic biology • Synthetic biomolecules* **17,** 871–877 (2013).

37. Kitney, R. & Freemont, P. Synthetic biology – the state of play. *FEBS Letters* **586,** 2029–2036 (2012).

38. Nielsen, J. & Keasling, J. D. Engineering Cellular Metabolism. *Cell* **164,** 1185–1197 (2016).

39. Clarke, L. J. Synthetic biology UK: progress, paradigms and prospects. *Engineering Biology* **1,** 66–70 (2017).

40. Carbonell, P., Jervis, A. J., Robinson, C. J., *et al.* An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Communications Biology* **1,** 1–10 (2018).

41. Orsi, E., Claassens, N. J., Nikel, P. I., *et al.* Growth-coupled selection of synthetic modules to accelerate cell factory development. *Nature Communications* **12,** 5295 (2021).

42. Ro, D.-K., Paradise, E. M., Ouellet, M., *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440,** 940–943 (2006).

43. Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451,** 86–89 (2008).

44. Paddon, C. J. & Keasling, J. D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nature Reviews Microbiology* **12,** 355–367 (2014).

45. Lee, S. Y., Kim, H. U., Chae, T. U., *et al.* A comprehensive metabolic map for production of bio-based chemicals. *Nature Catalysis* **2,** 18–33 (2019).

46. Chen, F. Y. H., Jung, H.-W., Tsuei, C.-Y., *et al.* Converting Escherichia coli to a Synthetic Methylotroph Growing Solely on Methanol. *Cell* **182,** 933–946.e14 (2020).

47. Reuß, D. R., Altenbuchner, J., Mäder, U., *et al.* Large-scale reduction of the Bacillus subtilis genome: consequences for the transcriptional network, resource allocation, and metabolism. *Genome Research* **27,** 289–299 (2017).

48. Zhang, Y.-H. P., Sun, J. & Ma, Y. Biomanufacturing: history and perspective. *Journal of Industrial Microbiology and Biotechnology* **44,** 773–784 (2017).

49. Brooks, S. M. & Alper, H. S. Applications, challenges, and needs for employing synthetic biology beyond the lab. *Nature Communications* **12,** 1390 (2021).

50. Doudna, J. A. The promise and challenge of therapeutic genome editing. *Nature* **578,** 229–236 (2020).

51. Steens, J. A., Zhu, Y., Taylor, D. W., *et al.* SCOPE enables type III CRISPR-Cas diagnostics using flexible targeting and stringent CARF ribonuclease activation. *Nature Communications* **12,** 5033 (2021).

52. Charbonneau, M. R., Isabella, V. M., Li, N., *et al.* Developing a new class of engineered live bacterial therapeutics to treat human diseases. *Nature Communications* **11,** 1738 (2020).

53. Grozinger, L., Amos, M., Gorochowski, T. E., *et al.* Pathways to cellular supremacy in biocomputing. *Nature Communications* **10,** 5250 (2019).

54. Kumar, A., Hsu, L. H.-H., Kavanagh, P., *et al.* The ins and outs of microorganism–electrode electron transfer reactions. *Nature Reviews Chemistry* **1,** 1–13 (2017).

55. Gilbert, C. & Ellis, T. Biological Engineered Living Materials: Growing Functional Materials with Genetically Programmable Properties. *ACS synthetic biology* **8,** 1–15 (2019).

56. McCarty, N. S., Graham, A. E., Studená, L., *et al.* Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nature Communications* **11,** 1281 (2020).

57. Clarke, L. & Kitney, R. Developing synthetic biology for industrial biotechnology applications. *Biochemical Society Transactions* **48,** 113–122 (2020).

58. García-Granados, R., Lerma-Escalera, J. A. & Morones-Ramírez, J. R. Metabolic Engineering and Synthetic Biology: Synergies, Future, and Challenges. *Frontiers in Bioengineering and Biotechnology* **7,** 36 (2019).

59. Yan, Q. & Fong, S. S. Challenges and Advances for Genetic Engineering of Non-model Bacteria and Uses in Consolidated Bioprocessing. *Frontiers in Microbiology* **8,** 2060 (2017).

60. Le Novère, N. Quantitative and logic modelling of molecular and gene networks. *Nature Reviews Genetics* **16,** 146–158 (2015).

61. Hwang, M., Garbey, M., Berceli, S. A., *et al.* Rule-Based Simulation of Multi-Cellular Biological Systems—A Review of Modeling Techniques. *Cellular and Molecular Bioengineering* **2,** 285–294 (2009).

62. Asveld, L., Ganzevles, J. & Osseweijer, P. Trustworthiness and Responsible Research and Innovation: The Case of the Bio-Economy. *Journal of Agricultural and Environmental Ethics* **28,** 571–588 (2015).

63. Braun, M., Fernau, S. & Dabrock, P. (Re-)Designing Nature? An Overview and Outlook on the Ethical and Societal Challenges in Synthetic Biology. *Advanced Biosystems* **3,** 1800326 (2019).

64. *Technik und Kultur: Bedingungs- und Beeinflussungsverhältnisse* (eds Banse, G., Grundwald, A. & (TH), U. K.) *Karlsruher Studien Technik und Kultur* **Bd. 1** (Karlsruhe, 2010).

65. Kaebnick, G. E., Heitman, E., Collins, J. P., *et al.* Precaution and governance of emerging technologies. *Science* **354,** 710–711 (2016).

66. Stirling, A. & Coburn, J. From CBA to Precautionary Appraisal: Practical Responses to Intractable Problems. *The Hastings Center Report* **48 Suppl 1,** S78–S87 (2018).

67. Moran, M., Calvert, J., Cook, T., *et al.* Emerging biotechnologies: technology, choice and the public good, a guide to the report (2012).

68. Deplazes-Zemp, A. & Biller-Andorno, N. Explaining life. Synthetic biology and non-scientific understandings of life. *EMBO reports* **13,** 959–963 (2012).

69. Deplazes-Zemp, A. The conception of life in synthetic biology. *Science and Engineering Ethics* **18,** 757–774 (2012).

70. Deplazes-Zemp, A., Gregorowius, D. & Biller-Andorno, N. Different Understandings of Life as an Opportunity to Enrich the Debate About Synthetic Biology. *NanoEthics* **9,** 179–188 (2015).

71. Porcar, M. & Peretó, J. Nature versus design: synthetic biology or how to build a biological non-machine. *Integrative Biology* **8,** 451–455 (2016).

72. Lakoff, G. & Johnson, M. *Metaphors we live by* (Chicago, 2003).

73. European Commission. Directorate General for Health and Consumers. *Final opinion on synthetic biology III: risks to the environment and biodiversity related to synthetic biology and research priorities in the field of synthetic biology.* (LU, 2015).

74. Scientific Committee on Health and Environmental Risks (SCHER), Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR) & Scientific Committee on Consumer Safety (SCCS). *Opinion on Synthetic Biology II Risk assessment methodologies and safety aspects. European Comission* 2015.

75. Dana, G. V., Kuiken, T., Rejeski, D., *et al.* Synthetic biology: Four steps to avoid a synthetic-biology disaster. *Nature* **483,** 29 (2012).

76. Moe-Behrens, G. H. G., Davis, R. & Haynes, K. A. Preparing synthetic biology for the world. *Front. Microbio.* **4** (2013).

77. Wright, O., Stan, G.-B. & Ellis, T. 2. Building-in biosafety for synthetic biology. *Microbiology* **159,** 1221–1235 (Pt_7 2013).

78. Schmidt, M. & de Lorenzo, V. Synthetic bugs on the loose: containment options for deeply engineered (micro)organisms. *Current Opinion in Biotechnology* **38,** 90–96 (2016).

79. Berg, P. & Singer, M. F. The recombinant DNA controversy: twenty years later. *Proceedings of the National Academy of Sciences of the United States of America* **92,** 9011 (1995).

80. Berg, P., Baltimore, D., Brenner, S., *et al.* Summary statement of the Asilomar conference on recombinant DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America* **72,** 1981 (1975).

81. Benner, S. A. & Sismour, A. M. Synthetic biology. *Nature Reviews Genetics* **6,** 533–543 (2005).

82. Cases, I. & Lorenzo, V. d. Genetically modified organisms for the environment: stories of success and failure and what we have learned from them (2005).

83. Lorenzo, V. d. Environmental biosafety in the age of synthetic biology: do we really need a radical new approach? *BioEssays* **32,** 926–931 (2010).

84. Singh, J. S., Abhilash, P. C., Singh, H. B., *et al.* Genetically engineered bacteria: an emerging tool for environmental remediation and future research perspectives. *Gene* **480,** 1–9 (2011).

85. Dvořák, P., Nikel, P. I., Damborský, J., *et al.* Bioremediation 3.0: Engineering pollutant-removing bacteria in the times of systemic biology. *Biotechnology Advances* **35,** 845–866 (2017).

86. Gibson, D. G., Glass, J. I., Lartigue, C., *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *science* **329,** 52–56 (2010).

87. Zhang, W., Mitchell, L. A., Bader, J. S., *et al.* Synthetic Genomes. *Annual Review of Biochemistry* **89,** 77–101 (2020).

88. Mirts, E. N., Bhagi-Damodaran, A. & Lu, Y. Understanding and Modulating Metalloenzymes with Unnatural Amino Acids, Non-Native Metal Ions, and Non-Native Metallocofactors. *Accounts of Chemical Research* **52,** 935–944 (2019).

89. Bowie, J. U., Sherkhanov, S., Korman, T. P., *et al.* Synthetic Biochemistry: The Bio-inspired Cell-Free Approach to Commodity Chemical Production. *Trends in Biotechnology. Special Issue: Metabolic Engineering* **38,** 766–778 (2020).

90. Drienovská, I. & Roelfes, G. Expanding the enzyme universe with genetically encoded unnatural amino acids. *Nature Catalysis* **3,** 193–202 (2020).

91. Dunkelmann, D. L., Oehm, S. B., Beattie, A. T., *et al.* A 68-codon genetic code to incorporate four distinct non-canonical amino acids enabled by automated orthogonal mRNA design. *Nature Chemistry* **13,** 1110–1117 (2021).

92. Carroll, D. Genome Editing: Past, Present, and Future. *The Yale Journal of Biology and Medicine* **90,** 653–659 (2017).

93. Gong, T., Xu, X., Dang, Y., *et al.* An engineered Pseudomonas putida can simultaneously degrade organophosphates, pyrethroids and carbamates. *Science of The Total Environment* **628-629,** 1258–1265 (2018).

94. Wehrs, M., Tanjore, D., Eng, T., *et al.* Engineering Robust Production Microbes for Large-Scale Cultivation. *Trends in Microbiology* **27,** 524–537 (2019).

95. Jiang, T., Li, C., Teng, Y., *et al.* Recent advances in improving metabolic robustness of microbial cell factories. *Current Opinion in Biotechnology. Tissue, Cell and Pathway Engineering* **66,** 69–77 (2020).

96. SCoEaNIH, R. *Opinion on Synthetic Biology II. Risk Assessment Methodologies and Safety Aspects* (2015).

97. El Karoui, M., Hoyos-Flight, M. & Fletcher, L. Future Trends in Synthetic Biology—A Report. *Frontiers in Bioengineering and Biotechnology* **7,** 175 (2019).

98. EFSA Scientific Committee. Evaluation of existing guidelines for their adequacy for the microbial characterisation and environmental risk assessment of microorganisms obtained through synthetic biology. *EFSA Journal* **18,** 6263 (2020).

99. Stirling, F. & Silver, P. A. Controlling the Implementation of Transgenic Microbes: Are We Ready for What Synthetic Biology Has to Offer? *Molecular Cell* **78,** 614–623 (2020).

100. Rodemeyer, M. *New life, old bottles: Regulating first-generation products of synthetic biology* (2009).

101. Carter, S. R., Rodemeyer, M., Garfinkel, M. S., *et al. Synthetic biology and the US biotechnology regulatory system: Challenges and options* (J. Craig Venter Institute, Rockville, MD (United States), 2014).

102. Hewett, J. P., Wolfe, A. K., Bergmann, R. A., *et al.* Human Health and Environmental Risks Posed by Synthetic Biology R&D for Energy Applications: A Literature Analysis. *Applied Biosafety* **21,** 177–184 (2016).

103. National Academies of Sciences, E. *Genetically engineered crops: Experiences and prospects* (2016).

104. Arnolds, K. L., Dahlin, L. R., Ding, L., *et al.* Biotechnology for secure biocontainment designs in an emerging bioeconomy. *Current Opinion in Biotechnology. Analytical Biotechnology* **71,** 25–31 (2021).

105. World Health Organization. *Laboratory biosafety manual* (2004).

106. World Health Organization. *Biorisk management: Laboratory biosecurity guidance* (2006).

107. Nordmann, B. D. Issues in biosecurity and biosafety. *International Journal of Antimicrobial Agents. Global Trends of Emerging Zoonotic Diseases* **36,** S66–S69 (2010).

108. Kelle, A. Synthetic biology and biosecurity: From low levels of awareness to a comprehensive strategy. *EMBO reports* **10** (S1 2009).

109. Beeckman, D. S. A. & Rüdelsheim, P. Biosafety and Biosecurity in Containment: A Regulatory Overview. *Frontiers in Bioengineering and Biotechnology* **8,** 650 (2020).

110. *Cartagena protocol on biosafety to the Convention on Biological Diversity: text and annexes* (ed of the Convention on Biological Diversity, S.) (Montreal, 2000).

111. Trevan, T. Biological research: Rethink biosafety. *Nature* **527,** 155–158 (2015).

112. Van Gelder, P., Klaassen, P., Taebi, B., *et al.* Safe-by-Design in Engineering: An Overview and Comparative Analysis of Engineering Disciplines. *International Journal of Environmental Research and Public Health* **18,** 6329 (2021).

113. Van de Poel, I. & Robaey, Z. Safe-by-Design: from Safety to Responsibility. *NanoEthics* **11,** 297–306 (2017).

114. Robaey, Z. *Dealing with risks of biotechnology: understanding the potential of Safe-by-Design.* Report commissioned by the Dutch Ministry of Infrastructure and Water Management (2018).

115. Lee, J. W., Chan, C. T. Y., Slomovic, S., *et al.* Next-generation biocontainment systems for engineered organisms. *Nature Chemical Biology* **14,** 530–537 (2018).

116. Torres, L., Krüger, A., Csibra, E., *et al.* Synthetic biology approaches to biological containment: pre-emptively tackling potential risks. *Essays in Biochemistry* **60** (ed Pinheiro, V. B.) 393–410 (2016).

117. Whitford, C. M., Dymek, S., Kerkhoff, D., *et al.* Auxotrophy to Xeno-DNA: an exploration of combinatorial mechanisms for a high-fidelity biosafety system for synthetic biology applications. *J Biol Eng* **12,** 13 (2018).

118. Berg, P., Baltimore, D., Brenner, S., *et al.* Asilomar Conference on Recombinant DNA Molecules. *Science* **188,** 991–994 (1975).

119. Steidler, L., Neirynck, S., Huyghebaert, N., *et al.* Biological containment of genetically modified Lactococcus lactis for intestinal delivery of human interleukin 10. *Nature biotechnology* **21,** 785–789 (2003).

120. Bahey-El-Din, M., Casey, P. G., Griffin, B. T., *et al.* Efficacy of a Lactococcus lactis Δ pyrG vaccine delivery platform expressing chromosomally integrated hly from Listeria monocytogenes. *Bioengineered bugs* **1,** 66–74 (2010).

121. Ronchel, M. C. & Ramos, J. L. Dual system to reinforce biological containment of recombinant bacteria designed for rhizoremediation. *Applied and environmental microbiology* **67,** 2649–2656 (2001).

122. Rovner, A. J., Haimovich, A. D., Katz, S. R., *et al.* Recoded organisms engineered to depend on synthetic amino acids. *Nature* **518,** 89–93 (2015).

123. Lopez, G. & Anderson, J. C. Synthetic Auxotrophs with Ligand-Dependent Essential Genes for a BL21(DE3) Biosafety Strain. *ACS synthetic biology* **4,** 1279–1286 (2015).

124. Hirota, R., Abe, K., Katsuura, Z.-i., *et al.* A Novel Biocontainment Strategy Makes Bacterial Growth and Survival Dependent on Phosphite. *Scientific Reports* **7,** 44748 (2017).

125. Callura, J. M., Dwyer, D. J., Isaacs, F. J., *et al.* Tracking, tuning, and terminating microbial physiology using synthetic riboregulators. *Proceedings of the National Academy of Sciences* **107,** 15898–15903 (2010).

126. Torres, B., Jaenecke, S., Timmis, K. N., *et al.* A dual lethal system to enhance containment of recombinant microorganisms. *Microbiology* **149,** 3595–3601 (2003).

127. Gallagher, R. R., Patel, J. R., Interiano, A. L., *et al.* Multilayered genetic safeguards limit growth of microorganisms to defined environments. *Nucleic acids research* **43,** 1945–1954 (2015).

128. Chan, C. T. Y., Lee, J. W., Cameron, D. E., *et al.* 'Deadman' and 'Passcode' microbial kill switches for bacterial containment. *Nature Chemical Biology* **12,** 82–86 (2016).

129. Xia, P.-F., Ling, H., Foo, J. L., *et al.* Synthetic genetic circuits for programmable biological functionalities. *Biotechnology Advances. Microbial Engineering Biotechnologies* **37,** 107393 (2019).

130. Green, A. A., Silver, P. A., Collins, J. J., *et al.* Toehold Switches: De-Novo-Designed Regulators of Gene Expression. *Cell* **159,** 925–939 (2014).

131. Chen, Y., Ho, J. M. L., Shis, D. L., *et al.* Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. *Nature Communications* **9,** 64 (2018).

132. Stirling, F., Bitzan, L., O'Keefe, S., *et al.* Rational design of evolutionarily stable microbial kill switches. *Molecular cell* **68,** 686–697. e3 (2017).

133. Stirling, F., Naydich, A., Bramante, J., *et al.* Synthetic Cassettes for pH-Mediated Sensing, Counting, and Containment. *Cell Reports* **30,** 3139–3148.e4 (2020).

134. Yang, S., Sleight, S. C. & Sauro, H. M. Rationally designed bidirectional promoter improves the evolutionary stability of synthetic genetic circuits. *Nucleic Acids Research* **41,** e33 (2013).

135. Wang, L., Brock, A., Herberich, B., *et al.* Expanding the genetic code of Escherichia coli. *Science* **292,** 498–500 (2001).

136. Isaacs, F. J., Carr, P. A., Wang, H. H., *et al.* Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement. *Science* **333,** 348–353 (2011).

137. Schmidt, M. Xenobiology: A new form of life as the ultimate biosafety tool. *Bioessays* **32,** 322–331 (2010).

138. Marlière, P., Patrouix, J., Döring, V., *et al.* Chemical Evolution of a Bacterium's Genome. *Angewandte Chemie International Edition* **50,** 7109–7114 (2011).

139. Zhang, Y., Ptacin, J. L., Fischer, E. C., *et al.* A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* **551,** 644–647 (2017).

140. Handal-Marquez, P., Anupama, A., Pezo, V., *et al.* Beneath the XNA world: Tools and targets to build novel biology. *Current Opinion in Systems Biology. Systems immunology & host-pathogen interaction (2020)* **24,** 142–152 (2020).

141. Lajoie, M. J., Rovner, A. J., Goodman, D. B., *et al.* Genomically recoded organisms expand biological functions. *science* **342,** 357–360 (2013).

142. Chen, Y., Tang, J., Wang, L., *et al.* Creation of Bacterial Cells with 5-Hydroxytryptophan as a 21st Amino Acid Building Block. *Chem* **6,** 2717–2727 (2020).

143. Vargas-Rodriguez, O., Sevostyanova, A., Söll, D., *et al.* Upgrading aminoacyl-tRNA synthetases for genetic code expansion. *Current Opinion in Chemical Biology. Synthetic Biology / Synthetic Biomolecules* **46,** 115–122 (2018).

144. Fredens, J., Wang, K., de la Torre, D., *et al.* Total synthesis of Escherichia coli with a recoded genome. *Nature* **569,** 514–518 (2019).

145. Kuo, J., Stirling, F., Lau, Y. H., *et al.* Synthetic genome recoding: new genetic codes for new features. *Current Genetics* **64,** 327–333 (2018).

146. Committee on Future Biotechnology Products and Opportunities to Enhance Capabilities of the Biotechnology Regulatory System, Board on Life Sciences, Board on Agriculture and Natural Resources, *et al. Preparing for Future Products of Biotechnology* (Washington, D.C., 2017).

147. De Lorenzo, V., Krasnogor, N. & Schmidt, M. For the sake of the Bioeconomy: define what a Synthetic Biology Chassis is! *New Biotechnology* **60,** 44–51 (2021).

148. Toprak, E., Veres, A., Yildiz, S., *et al.* Building a morbidostat: an automated continuous-culture device for studying bacterial drug resistance under dynamically sustained drug inhibition. *Nature protocols* **8,** 555–567 (2013).

149. Maddamsetti, R., Lenski, R. E. & Barrick, J. E. Adaptation, clonal interference, and frequency-dependent interactions in a long-term evolution experiment with Escherichia coli. *Genetics* **200,** 619–631 (2015).

150. Schmidt, M. A metric space for semantic containment: Towards the implementation of genetic firewalls. *Bio Systems* **185,** 104015 (2019).

151. Becker, J. & Wittmann, C. Bio-based production of chemicals, materials and fuels–Corynebacterium glutamicum as versatile cell factory. *Current opinion in biotechnology* **23,** 631–640 (2012).

152. Hutchison, C. A., Chuang, R.-Y., Noskov, V. N., *et al.* Design and synthesis of a minimal bacterial genome. *Science* **351** (2016).

153. Niehus, X., Crutz-Le Coq, A.-M., Sandoval, G., *et al.* Engineering Yarrowia lipolytica to enhance lipid production from lignocellulosic materials. *Biotechnology for Biofuels* **11,** 1–10 (2018).

154. Chi, H., Wang, X., Shao, Y., *et al.* Engineering and modification of microbial chassis for systems and synthetic biology. *Synthetic and systems biotechnology* **4,** 25–33 (2019).

155. Martin-Pascual, M., Batianis, C., Bruinsma, L., *et al.* A navigation guide of synthetic biology tools for Pseudomonas putida. *Biotechnology Advances* **49,** 107732 (2021).

156. Poblete-Castro, I., Becker, J., Dohnt, K., *et al.* Industrial biotechnology of Pseudomonas putida and related species. *Applied Microbiology and Biotechnology* **93,** 2279–2290 (2012).

157. Kivisaar, M. 2. Narrative of a versatile and adept species Pseudomonas putida. *Journal of Medical Microbiology* **69,** 324–338 (2020).

158. Del Vecchio, D., Ninfa, A. J. & Sontag, E. D. Modular cell biology: retroactivity and insulation. *Molecular systems biology* **4,** 161 (2008).

159. Slusarczyk, A. L., Lin, A. & Weiss, R. Foundations for the design and implementation of synthetic genetic circuits. *Nature Reviews Genetics* **13,** 406–420 (2012).

160. Akkaya, Ö., Nikel, P. I., Pérez-Pantoja, D., *et al.* Evolving metabolism of 2, 4-dinitrotoluene triggers SOS-independent diversification of host cells. *Environmental microbiology* **21,** 314–326 (2019).

161. Martínez-García, E., Jatsenko, T., Kivisaar, M., *et al.* Freeing P. seudomonas putida KT2440 of its proviral load strengthens endurance to environmental stresses. *Environmental microbiology* **17,** 76–90 (2015).

162. Nikel, P. I., Chavarría, M., Fuhrer, T., *et al.* Pseudomonas putida KT2440 strain metabolizes glucose through a cycle formed by enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and pentose phosphate pathways. *Journal of Biological Chemistry* **290,** 25920–25932 (2015).

163. Akkaya, Ö., Pérez-Pantoja, D. R., Calles, B., *et al.* The metabolic redox regime of Pseudomonas putida tunes its evolvability toward novel xenobiotic substrates. *MBio* **9,** e01512–18 (2018).

164. Sudarsan, S., Dethlefsen, S., Blank, L. M., *et al.* The functional structure of central carbon metabolism in Pseudomonas putida KT2440. *Applied and environmental microbiology* **80,** 5292–5303 (2014).

165. Kukurugya, M. A., Mendonca, C. M., Solhtalab, M., *et al.* Multi-omics analysis unravels a segregated metabolic flux network that tunes co-utilization of sugar and aromatic carbons in Pseudomonas putida. *Journal of Biological Chemistry* **294,** 8464–8479 (2019).

166. Ankenbauer, A., Schäfer, R. A., Viegas, S. C., *et al. Pseudomonas putida* KT2440 is naturally endowed to withstand industrial-scale stress conditions. *Microbial Biotechnology* **13,** 1145–1161 (2020).

167. Weimer, A., Kohlstedt, M., Volke, D. C., *et al.* Industrial biotechnology of Pseudomonas putida: advances and prospects. *Applied Microbiology and Biotechnology* **104,** 7745–7766 (2020).

168. Nelson, K. E., Weinel, C., Paulsen, I. T., *et al.* Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440. *Environmental microbiology* **4,** 799–808 (2002).

169. Belda, E., Van Heck, R. G., José Lopez-Sanchez, M., *et al.* The revisited genome of Pseudomonas putida KT2440 enlightens its value as a robust metabolic chassis. *Environmental microbiology* **18,** 3403–3424 (2016).

170. Puchałka, J., Oberhardt, M. A., Godinho, M., *et al.* Genome-Scale Reconstruction and Analysis of the Pseudomonas putid a KT2440 Metabolic Network Facilitates Applications in Biotechnology. *PLoS computational biology* **4,** e1000210 (2008).

171. Nogales, J., Mueller, J., Gudmundsson, S., *et al.* High-quality genome-scale metabolic modelling of Pseudomonas putida highlights its broad metabolic capabilities. *Environmental microbiology* **22,** 255–269 (2020).

172. Nieto-Domínguez, M. & Nikel, P. I. Intersecting Xenobiology and Neometabolism To Bring Novel Chemistries to Life. *ChemBioChem* **21,** 2551–2571 (2020).

173. Kampers, L. F., Volkers, R. J. & Martins dos Santos, V. A. Pseudomonas putida KT 2440 is HV 1 certified, not GRAS. *Microbial biotechnology* **12,** 845–848 (2019).

174. Herman, L., Chemaly, M., Cocconcelli, P. S., *et al.* The qualified presumption of safety assessment and its role in EFSA risk evaluations: 15 years past. *FEMS microbiology letters* **366,** 260 (2019).

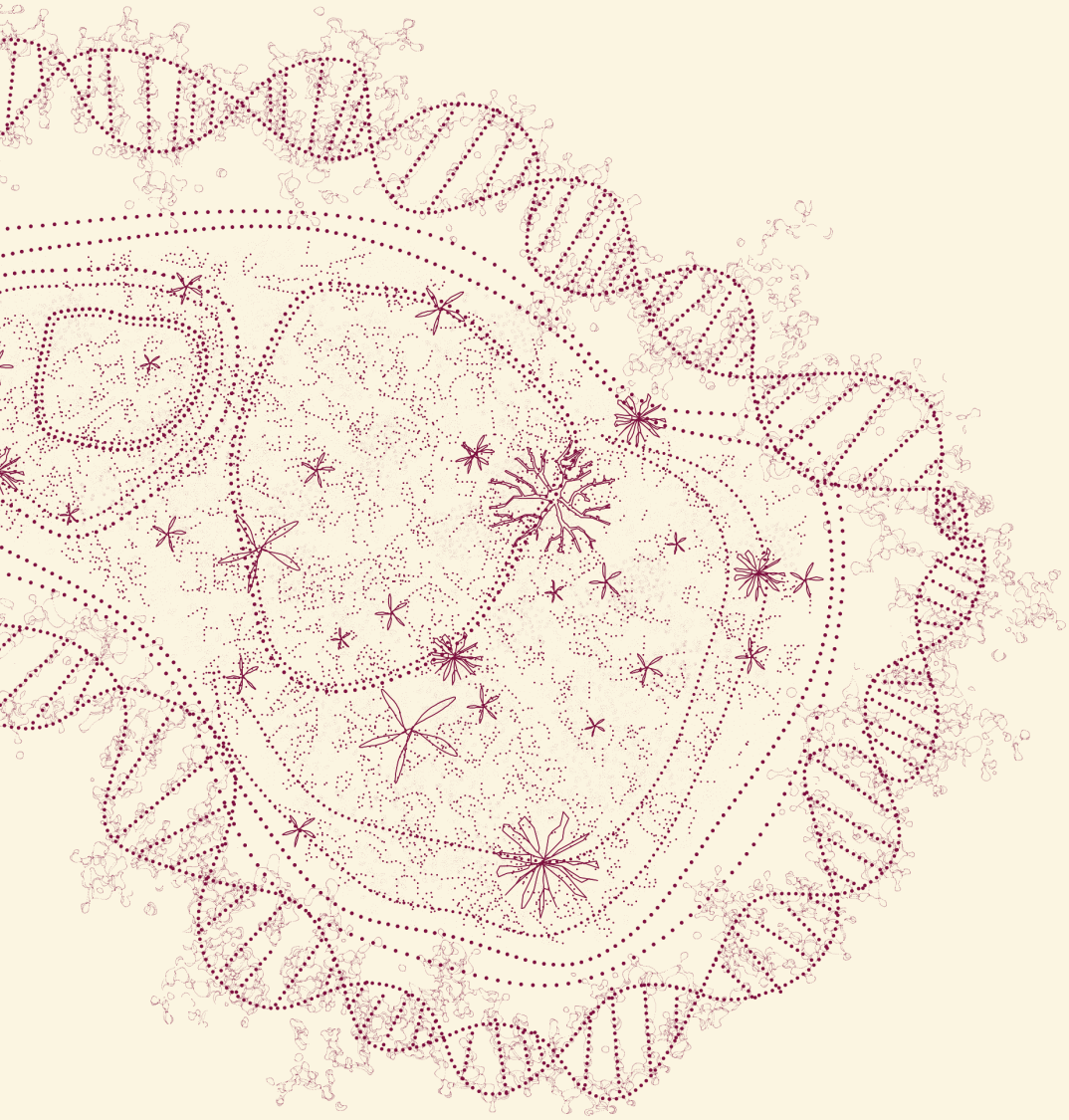
175. Bej, A. K., Molin, S., Perlin, M., *et al.* Maintenance and killing efficiency of conditional lethal constructs in Pseudomonas putida. *Journal of industrial microbiology and biotechnology* **10,** 79–85 (1992).

176. Munthali, M. T., Timmis, K. N. & Diaz, E. Use of colicin E3 for biological containment of microorganisms. *Applied and Environmental Microbiology* **62,** 1805–1807 (1996).

177. Szafranski, P., Mello, C. M., Sano, T., *et al.* A new approach for containment of microorganisms: dual control of streptavidin expression by antisense RNA and the T7 transcription system. *Proceedings of the National Academy of Sciences* **94,** 1059–1063 (1997).

178. Molina, L., Ramos, C., Ronchel, M.-C., *et al.* Construction of an efficient biologically contained Pseudomonas putida strain and its survival in outdoor assays. *Applied and Environmental Microbiology* **64,** 2072–2078 (1998).

179. Castanon, O., Smith, C. J., Khoshakhlagh, P., *et al.* CRISPR-mediated biocontainment. *biorxiv* (2020).



CHAPTER 2

ReScribe: an unrestrained tool combining multiplex recombineering and minimal-PAM ScCas9 for genome recoding *Pseudomonas putida*

Enrique Asin-Garcia*, Maria Martin-Pascual*,
Luis Garcia-Morales, Richard van Kranenburg, Vitor A. P. Martins dos Santos
* Contributed equally

Adapted for publication as:
ReScribe: An Unrestrained Tool Combining Multiplex Recombineering and Minimal-PAM ScCas9 for Genome Recoding Pseudomonas putida,
2021, ACS Synthetic Biology

Abstract

Genome recoding enables incorporating new functions into the DNA of microorganisms. By reassigning codons to non-canonical amino acids, the generation of new-to-nature proteins offers countless opportunities for bioproduction and biocontainment in industrial chassis. A key bottleneck in genome recoding efforts, however, is the low efficiency of recombineering, which hinders large-scale applications at acceptable speed and cost. To relieve this bottleneck, we developed ReScribe, a highly optimized recombineering tool enhanced by CRISPR-Cas9-mediated counterselection built upon the minimal PAM 5'-NNG-3' of the *Streptococcus canis* Cas9 (ScCas9). As a proof of concept, we used ReScribe to generate a minimally recoded strain of the industrial chassis *Pseudomonas putida* by replacing TAG stop codons (functioning as PAMs) of essential metabolic genes with the synonymous TAA. We showed that ReScribe enables nearly 100 % engineering efficiency of multiple *loci* in *P. putida*, opening promising avenues for genome editing and applications thereof in this bacterium and beyond.

Introduction

Due to its physiological robustness, stress resistance, metabolic versatility, and fast growth, *Pseudomonas putida* KT2440 has become a platform for metabolic engineering aimed at industrial and environmental applications [1–3]. On top of its intrinsic features, the toolbox for genetic programming of this bacterium has significantly improved over the past years, propelling it to the front ranks of the synthetic biology platforms [4]. However, its full potential is still held back by limitations on genome-scale editing that hinder ambitious bioengineering projects such as genome recoding. The ability to rewrite genomes provides the opportunity to incorporate new properties to the DNA of industrially relevant microorganisms, thereby increasing their value as biotechnological platforms [5]. By reassigning natural codons to non-canonical amino acids, this approach allows the synthesis of new-to-function proteins and peptides, leading to a significantly expanded space for bespoke biocatalysis [6, 7]. Moreover, alterations in the translation machinery derived from the repurposing of codons can also be part of their biosafety assurance by impairing the ability of recoded microbes to express foreign DNA or by producing proteins that cannot be functionally expressed in other organisms [8, 9]. This impact on horizontal gene transfer and viral infection enhances the biosafety and stability of the engineered strains constituting a powerful biocontainment strategy [10].

Over the past decade, a few recoding efforts have been published using different approaches [5, 8, 11–15]. The first strategy used recombineering-based multiplex automated genome engineering (MAGE) to create a genome-wide recoded *Escherichia coli* [8]. In that pioneering approach, all TAG stop codons were replaced with synonymous TAA codons, allowing the deletion of release factor 1 and the reassignment of UAG translation function to non-canonical amino acids (ncAA) [16]. This alteration allowed the incorporation of the recoded codon in essential genes conferring metabolic dependence on the ncAA for cell viability [10, 17] and hampering the dispersal of functional DNA from the synthetic chassis cell to natural microbes [15]. Other strategies have been based on chemical synthesis of recoded DNA and its incorporation into the target microbe either by substituting genome segments of different sizes [5, 13, 14, 18] or by substituting the entire genome [19].

Despite their potential, these strategies keep requiring a costly DNA synthesis investment and a very laborious assembly process. Hence, there is a need for an efficient site-directed editing genome engineering tool to enable recoding.

Recombineering is a powerful genome editing technique based on a recombinase protein that promotes the incorporation of single-stranded DNA (ssDNA) molecules mimicking Okazaki fragments in the replication fork during DNA replication producing the intended mutation [20, 21]. Over the past years, different recombinases have been tested in *P. putida* including the redβ-recombinase and the RecET system from the *E. coli* Lambda phage and Rac prophage, the activity of which is relatively low in *Pseudomonas* species in absence of selection [4] Other *P. putida*-borne recombinases identified through genus-specific bioinformatic mining, Ssr and Rec2, have been experimentally validated with promising results [22, 23] Recently, Rec2-mediated ssDNA recombineering has been merged with transitional inhibition of the native mismatch machinery repair (MMR) system [24] by coexpressing a dominant-negative allele of *mutL*, and further improved by iterating the recombineering protocol. Yet, efficiencies dropped dramatically when multiplexing, [25] which prevents pursuing a full genome recoding and other high-throughput enterprises of mutations at genome-scale. The authors pointed to the core recombinase as the key limitation suggesting that a different or an optimized protein may work better in the proposed pipeline. Alternatives have already arisen in the way of (i) new recombinases like PapRecT, which has enabled efficient recombineering in the related species *Pseudomonas aeruginosa*; and (ii) optimization methods, such as RBS strengthening [26].

In addition, CRISPR-Cas9 can be used as counterselection in recombineering by eliminating non-edited cells. To such end, the endonuclease Cas9, guided by the spacer, induces double-stranded DNA breaks (DSB) in the target site when the cell has not been mutated. In most bacteria, DSB can be prevented only *via* homology-directed repair (HDR) if a dsDNA template is provided. Therefore, in the absence of such template, cells will die as they typically lack a functional non-homologous end joining (NHEJ) repair system [27]. The targeting and cleavage specificity of Cas9 proteins requires two RNA elements, the precursor CRISPR RNA (crRNA), and the trans-activating RNA (tracrRNA). Each mature crRNA:tracrRNA:Cas9 ribonucleopro-

tein complex comprises a single transcribed spacer, a part of the neighboring repeat sequence, the tracrRNA and the Cas9. The design of the spacer allows directing the Cas9 protein to the desired protospacer (complementary spacer sequence present in the genome). The only requirement of the target site is the presence of a protospacer adjacent motif (PAM), a short (3–8 nucleotides) sequence, commonly found at the 3' end of the protospacer that varies among Cas proteins.

CRISPR-Cas9-mediated counterselection has already been used for increasing the efficiency of recombineering in *P. putida* [28] employing the paradigm *Streptococcus pyogenes* Cas9 (SpyCas9), which needs the PAM sequence 5'-NGG-3'. This general version of the method presents limitations for high-throughput applications like bacterial genome recoding, in which the difference between the wild type and mutated genotypes can be a single nucleotide. First, this single change might be insufficient for preventing the Cas9 targeting when positioned in the protospacer sequence, since single mismatches across the spacer sequence can be tolerated [29] and recombinant cells would not be properly discriminated by the Cas9 cleavage. Furthermore, although short and abundant in high-GC content genomes such as the one of *P. putida*, the 5'-NGG-3' PAM is not available in every desired target site required for genome recoding. In this scenario, the deployment of SpCas9 as counterselection method becomes a serious challenge, if not an impossible enterprise.

Here, we develop ReScribe (Recombineering + ScCas9-mediated counterselection), a highly efficient tool for genome recoding *P. putida* making use of the TAG stop codon itself as PAM for high on-target efficiencies. We boosted recombineering efficiencies by using the *Streptococcus canis* Cas9 ortholog, ScCas9. Previous *in silico* analysis showed a ScCas9 PAM specificity of 5'-NNGTT-3', which was later refined to 5'-NNG-3' in *in vivo* studies [30]. Additionally, an engineered version of the ScCas9 in which the loop D367–376 had been removed showed a concomitant change in the specificity of the PAM from 5'-NNG-3' to 5'-NAG-3' [30]. Both PAM sequences would allow to edit all possible TAG stop codons of the genome of *P. putida* KT2440, permitting an efficient counterselection after their recoding to TAA.

We thus propose ReScribe as a key solution to other multiple approaches that failed to surpass the reference recombineering efficiencies in *P. putida* KT2440 en-

compassing: (i) oligonucleotide design, (ii) RBS strengthening, and (iii) the alternative PapRecT recombinase. By using ReScribe, we showed near to 100 % cleavage efficiencies with both ScCas9 and ScCas9Δloop using a wide range of spacers in different *P. putida* and *E. coli* strains, revealing their virtually unrestrained applicability in these two different bacteria. Moreover, ReScribe reached allelic replacement efficiencies higher than 90 % after a single round of recombineering in single and multiple *loci* simultaneously. Ultimately, such impressive efficiencies allowed us to build a minimally recoded *P. putida* KT2440 strain, in which the TAG stop codons of essential metabolic genes were replaced by the synonymous TAA in a highly efficient manner. This first milestone evidences the power of our technology, not only as a mean for whole genome recoding, but also as an unprecedented tool for precise and specific targets, removing the PAM boundaries of other recombineering and CRISPRCas9mediated counterselection methods.

Results

Analysis of the *P. putida* genome for recoding

The 6.18 Mb genome of *P. putida* KT2440 contains a total of 5671 open reading frames [31] (NCBI accession number NC_002947.4). On the basis of this annotation and the aforementioned UAG recoding strategy (2011) [16], 654 genes that contain the least frequent TAG stop codon were computationally identified (Supplementary Table S6), representing ~11.5 % of the total, which is significantly higher than that of *E. coli*, probably related to the high GC-content of *P. putida*'s genome [32, 33]. A comprehensive analysis of these genes displayed features such as genomic coordinates, orientation, and size. Moreover, other characteristics that might complicate codon conversion were considered, namely, overlapping reading frames and essentiality (Supplementary Table S6). The list included 67 genes in which the TAG stop codon overlaps a different reading frame. Out of these 67, in 35 instances the recoding of TAG to TAA would result in a non-synonymous amino acid change in the product of the second reading frame, which could have an impact in the ultimate recoded phenotype and thus is not desired. Essentiality is a conditional feature subjected to the physiological context as a function of several factors. In absence

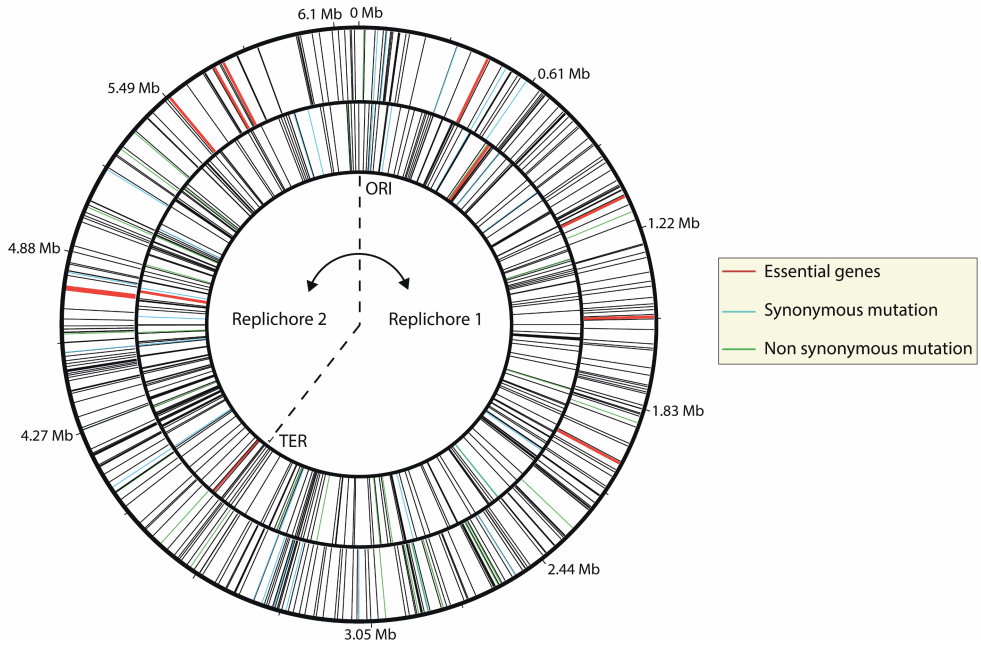


Figure 2.1: P. putida KT2440 genome's representation illustrating with lines the coordinates and orientation of all 654 genes terminated by TAG stop codons. TAG stop codons that are clockwise transcribed on the + DNA strand are depicted in the outer ring while those counterclockwise transcribed on the – DNA strand are represented in the inner ring. Essential genes terminated by TAG stop codons are indicated by the thick red lines. Genes finished in TAG stop codons whose mutation would result in a synonymous or non-synonymous amino acidic change on an overlapping gene are shown in blue or green, respectively. Genomic coordinates are represented around the circle, whereas origin of replication (ORI), terminus (TER), and replichores 1 and 2 are plotted in the inner circle.

of studies providing a comprehensive record of all essential genes of *P. putida* (*e.g.*, high-density TnSeq [34, 35] or CRISPRi-based screening libraries), we defined the essential metabolic genes based on the predictions made through an experimentally validated genome-scale, constraint-based metabolic model [36–38]. In order to identify the essential genes containing a TAG stop codon, a list of 270 conditionally essential genes in glucose minimal media was predicted using flux balance analysis (FBA) and the *P. putida* genome scale metabolic model iJP962 [36]. Of these, 12 were terminated by a TAG stop codon (Figure 2.1 and Supplementary Table S7).

Due to either poor efficiency or unsuitability of their selection methods, current genome editing tools did not seem suitable for conducting 654 mutations in *P. putida*. Therefore, for a task of this magnitude, it became imperative to develop a

powerful new technique that relied on efficient multiplex recombineering.

RBS optimization strategy does not increase ssDNA recombineering efficiency in *P. putida*

With an efficiency in the range of 10 % of single target replacements after 10 iterative cycles, the recombineering protocol with recombinase Rec2 [25] served as the baseline for the optimization study of our method. According to the results of our oligonucleotide optimization study (Figure 2.2), replacement efficiency was highest when mediated by 60-mer oligonucleotides. As a consequence, the oligos used in subsequent experiments were designed of 60 nucleotides in length, with the desired mutations included in the middle of the sequence. In addition, while phosphorothioate bonds located at the terminal bases of recombineering oligonucleotides had been reported to increase the replacement efficiency in *E. coli* by evading nuclease degradation [39, 40], our findings depicted in Figure 2.2 indicated that oligonucleotides with phosphorothioate bonds do not result in higher recombineering efficiency in *P. putida*. Therefore, phosphorothioate bonds were not included in the oligos used in this study.

In an attempt to enhance the editing efficiency of the reference plasmid pSEVA2514-rec2-mutLE36KPP by increasing the gene expression of its elements, the strongest predicted ribosomal binding sites (RBS) were designed and cloned upstream of both *rec2* and *mutLE36KPP* gene sequences. To evaluate the effect of the optimized variants, we used two different readouts to test the frequencies of mutant appearance: streptomycin resistance conferred by the K43T mutation in the *rpsL* gene in a WT *P. putida* KT2440 strain, and green fluorescence granted by the restoration of the genomic *gfp* gene in *P. putida* Tn7GFPstop.

The results of applying 10 iterative recombineering cycles with oligonucleotides RO rpsL 60 and RO gfp stop, respectively, are depicted in Figure 2.3 A and 2.3 B. The percentage of mutated cells increased from 2.40 % (Cycle 1) to 11.87 % (Cycle 10) for *rpsL*, and from 0.13 % (Cycle 1) to 5.05 % (Cycle 10) for *gfp*, when using the pSEVA2514-rec2-mutLE36KPP, which is in line with previously reported results [25]. Yet, the RBS optimized variants (Figure 2.3 C) carrying pSEVA2514-rec2RBSopt-mutLE36KPPRBSopt did not render significant increases of editing efficiency nor did

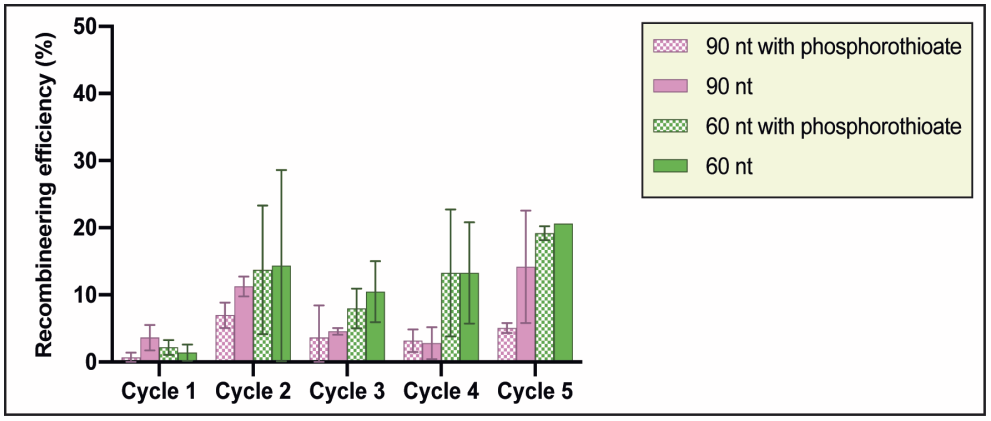


Figure 2.2: Assessment of oligonucleotide length and phosphorothioate modifications as oligonucleotide features that affect the allelic replacement frequency in P. putida KT2440. Evaluation is performed by screening of streptomycin resistant CFU after mutation K43T of the rpsL gene during 5 iterations of the recombineering cycle using Rec2. Recombineering efficiency was calculated in appropriate dilutions as the ratio between streptomycin resistant CFU growing in LB-sm plates and total CFU growing in LB plates. Replacement efficiency as a function of oligonucleotide length is depicted by colors, purple for 90-mer and green for 60-mer oligos, respectively. The effect in replacement efficiency of terminal phosphorothioate bonds positioned at both the 3' and 5' termini can be distinguished by dashed or full-colored bars, representing presence or absence of such backbone modifications, respectively. (Mean ± s.d., n = 4 biological).

they seem to result in a burden for the cells. In this case, results indicated percentages in the range of 3.10 % (Cycle 1) and 9.82 % (Cycle 10) for *rpsL*, and 2.26 % (Cycle 1) and 6.82 % (Cycle 10) for *gfp*. Even though RBS optimization *via* RBS strengthening has been reported to increase recombineering levels with other recombinases [26], our results are in consistency with those previously obtained with Rec2 in the related species *P. aeruginosa* [26], concluding that most likely Rec2 activity is not limited by the levels of expression of the recombinase gene but by different intrinsic factors.

PapRecT as an alternative to Rec2 for ssDNA recombineering in *P. putida*

Next, our efforts focused on the replacement of Rec2. The nature of the core recombinase and its source have been shown to play an important role in the efficiency of recombineering [41] suggesting that promising alternatives might be found among *Pseudomonas* species genomes and phages. Given its reported activity in other bacterial species, especially in the related *P. aeruginosa*, we selected PapRecT (orig-

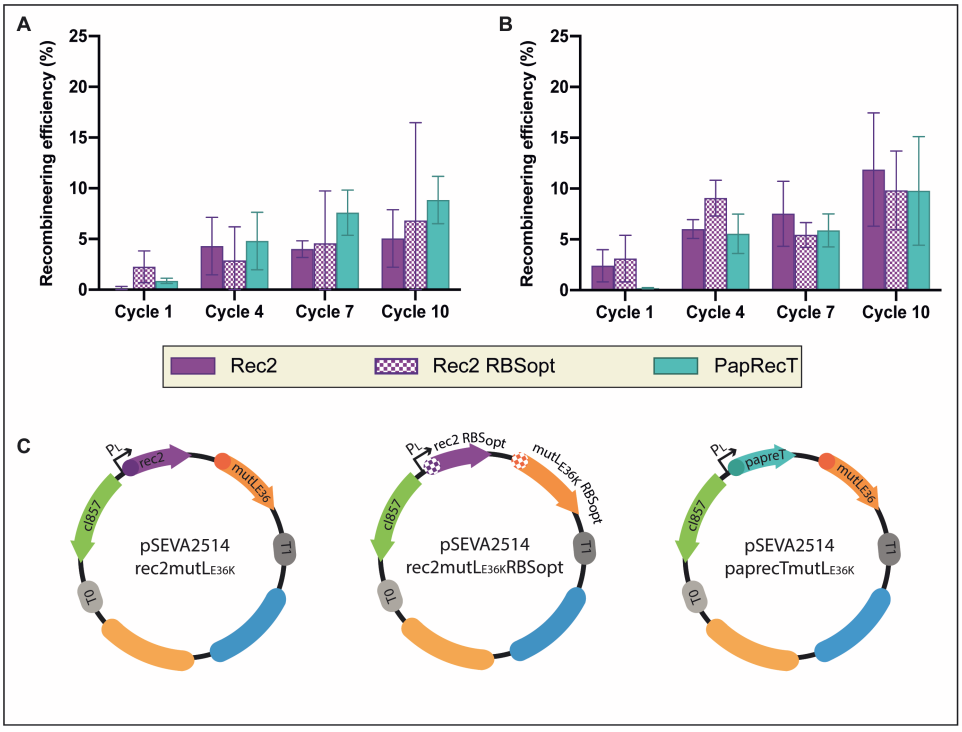


Figure 2.3: Comparison of allelic replacement efficiency of Rec2 with its RBS optimized version and with the alternative recombinase PapRecT. Recombineering was applied during 10 iterative cycles and samples were monitored after recovery steps of cycles 1, 4, 7, and 10. Graphs depict recombineering efficiency of Rec2, RBS optimized Rec2 and PapRecT mediating: (A) rpsL K43T mutation and (B) gfp stop66Y restoration mutation. Recombineering efficiency was calculated: (i) for the rpsL K43T mutation readout as the ratio between streptomycin resistant CFU growing in LB-sm plates and total CFU growing in LB plates; and (ii) for the gfp stop66Y restoration mutation as the ratio between fluorescent (mutated) and total CFU. Non-induced recombinase samples were included as controls and were subtracted from the absolute values. (Mean ± s.d., n = 4 biological). (C) The pSEVA2514-rec2-mutLE36KPP plasmid, harboring the thermolabile cI857 repressor (green), the rec2 recombinase gene (purple) and mutLE36KPP (orange). The pSEVA2514-rec2-mutLE36KPPRBSopt plasmid, harboring the thermolabile cI857 repressor (green), the RBS optimized rec2 recombinase gene (purple squared pattern) and the RBS optimized mutLE36KPP (orange squared pattern). The pSEVA2514-paprecT-mutLE36KPP plasmid, harboring the thermolabile cI857 repressor (green), the paprecT recombinase gene (turquoise), and mutLE36KPP (orange).

inating from a *P. aeruginosa* phage) [26] for our setup in *P. putida* KT2440 (Figure 2.3 C). The PapRecT recombinase resulted in recombineering efficiencies in the same range as those obtained with Rec2: 0.19 % (Cycle 1) and 9.77 % (Cycle 10) for the *rpsL* readout; and 0.87 % (Cycle 1) and 8.85 % (Cycle 10) for *gfp* (Figure 2.3 A and 2.3 B). Although no significant differences could be seen between the results from both

recombinases, PapRecT can be used as an equally valid alternative choice to the reference Rec2 in *P. putida* KT2440 due to their similar efficiency levels. In contrast to Rec2, the PapRecT RBS optimized version failed to provide any form of allelic replacement (data not shown).

ScCas9 efficiently cleaves *E. coli* and *P. putida* genome

Unable to boost the recombineering efficiency by optimizing or replacing the core recombinase, we therefore aimed at developing a recombineering tool enhanced by CRISPR-Cas9-mediated counterselection that would reach the efficacy level required for multiscale engineering purposes. To this end, we needed a Cas protein with PAM specifications compatible with the TAG stop codon, which we found in the ScCas9 having 5'-NNG-3' as PAM.

First, we codon optimized the ScCas9 gene for *P. putida* KT2440 and analyzed its functionality by performing fluorescent loss assays (Supplementary Figure S2.1). We targeted different regions of a *sfgfp* gene placed on a pSEVAb44 plasmid by using different spacers that were located next to a variety of 5'-NNG-3' PAMs. These preliminary results provided initial evidence that the codon optimized ScCas9 was functional in *P. putida* KT2440 and could target sequences adjacent to this minimal PAM. The spacer located next to a 5'-GAG-3' resulted in the most pronounced fluorescence loss, whereas the rest of the spacers showed different degrees of fluorescence depletion. While this preliminary experiment did not include a comprehensive number of spacers including all types of possible 5'-NNG-3' PAMs, it laid the groundwork for a more exhaustive analysis that included variations of a large array of factors such as characteristics of the minimal PAM, target genes, ScCas9 versions and bacterial systems. As very little was known about the use of this Cas9 variant with minimal PAM in bacteria [30], we included the bacterial model *E. coli* BL21 in our experiments. In addition, we also included *P. putida* EM383 and *E. coli* DH5α, both lacking the *recA* gene, as a certain tolerance for weak spacers has been found in *E. coli* when the HDR system is activated [42–44], with *recA* being the main element of the HDR pathway.

We transformed a total of 12 pSEVAb23-crRNA_sp plasmids with different targeting spacers in *P. putida* KT2440, *P. putida* EM383, *E. coli* BL21 and *E. coli* DH5α, har-

boring the pSEVAb62-ScCas9 plasmid (Figure 2.4 A) or the pSEVAb62-ScCas9Δloop plasmid, in which the loop D367–376 from the ScCas9 had been removed. We included both ScCas9 and ScCas9Δloop in our study since both their putative corresponding PAMs suit our purpose of using the TAG stop codon as PAM (5'-NNG-3' and 5'-NAG-3', respectively).

The 12 different spacers targeted 3 non-essential genes, *aceEF*, *rpsL*, and *speA* to avoid that the absence of colonies was due to the interference of Cas9 instead of its cleavage activity [45, 46]. Additionally, the latter two *loci* had previously been targeted by SpCas9 in *E. coli* with high efficiency [42]. For each *locus*, we designed 4 spacers with different PAMs that target different positions. The different PAMs: 5'-NBGTT-3', 5'-NBGVV-3', 5'-NAGTT-3', and 5'-NAGVV-3', were selected to analyze the PAM specificity of the ScCas9 and ScCas9Δloop and more specifically whether the nucleotides at positions 4 and 5 had higher specificity for T rather than A, G, and C [30]. We showed targeting efficiencies near to 100 % with all the spacers in *P. putida* KT2440 (Figure 2.4 B) and *P. putida* EM383 (Figure 2.4 C) with both the ScCas9 and the ScCas9Δloop. Additionally, we also proved that these codon-optimized ScCas9 variants are highly efficient in *E. coli* BL21 (Figure 2.4 D) and *E. coli* DH5α (Figure 2.4 E). In contrast to *P. putida* strains, in which all the spacers led to cell death at near 100 % efficiency, *E. coli* strains survived with some spacers which were less efficient or had efficiencies similar to the non-targeting spacer.

In contrast to previous results in which the removal of the loop resulted in a concomitant change in the specificity of the PAM from the minimal 5'-NNG-3' to 5'-NAG-3' [30], here we show that ScCas9Δloop is equally able to cleave targets positioned next to the 5'-NNG-3' PAMs as efficiently as the intact ScCas9 variant, both in *P. putida* and *E. coli* strains. Therefore, we decided to proceed with only the ScCas9 variant. With the perspective of combining both technologies: CRISPR (ScCas9) and recombineering (Rec2), we aimed at simplifying the CRISPR-ScCas9 design by combining the 2 plasmids, the pSEVAb23-crRNA_sp and the pSEVAb62-ScCas9, into one: the pSEVAb62-ScCas9-crRNA_sp plasmid. We cloned the same 12 spacers used in the 2-plasmid system into the pSEVAb62-ScCas9-crRNA_sp plasmid and transformed those plasmids in *P. putida* KT2440 and *P. putida* EM383 (Figure 2.4 F). We showed efficiencies between 76 % and 100 % in *P. putida* KT2440 (Figure 2.4

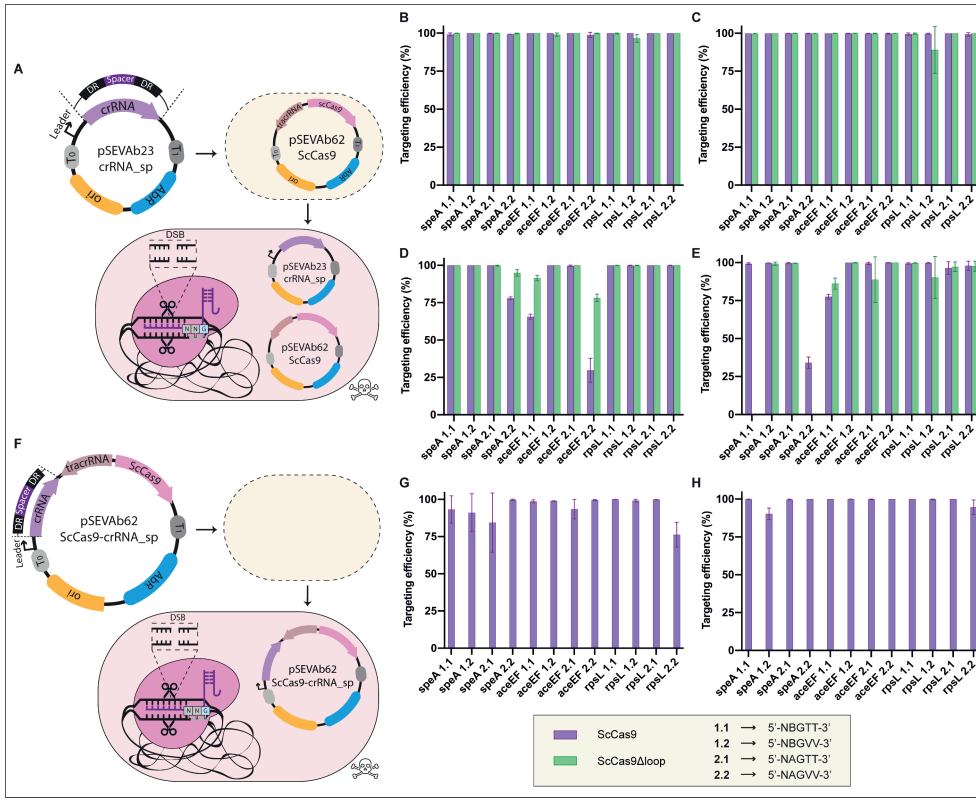


Figure 2.4: Cleavage assays with different ScCas9-based systems in P. putida and E. coli strains. (A) The two plasmids system is based on the pSEVAb23-crRNA and the pSEVAb62-ScCas9 plasmids. The pSEVAb23-crRNA plasmid harbors the crRNA comprised by the spacer (dark purple) interspersed by two direct repeats (black). The crRNA is expressed constitutively from the leader sequence. The pSEVAb23-crRNA plasmid is transformed into bacterial cells already harboring the pSEVAb62-ScCas9 plasmid. The pSEVAb62-ScCas9 plasmid expresses the ScCas9 (pink) and the tracrRNA (light brown) constitutively. The ScCas9:crRNA:tracrRNA complexes, directed by the spacer sequence, bind and unwind the target DNA, inducing a double strand break (DSB), causing bacterial cell death. The targeting efficiency is reported with different spacers targeting the genome of P. putida KT2440 (recA$^+$) (B), P. putida EM383 (recA$^-$) (C), E. coli BL21 (recA$^+$) (D), and E. coli DH5α (recA$^-$) (E), expressing sccas9 (purple bars) or sccas9Δloop (green bars). (F) The one plasmid system is based on the pSEVAb62-ScCas-crRNA_sp plasmid, which harbors the crRNA (light purple), the tracrRNA (light brown), and the ScCas9 (pink). All elements are expressed constitutively. After transformation the pSEVAb62-ScCas-crRNA_sp plasmid into bacterial cells, ScCas9:crRNA:tracrRNA complexes are formed, eliciting bacterial cell death. The targeting efficiency is reported with different spacers cleaving the genome of P. putida KT2440 (recA$^+$) (G) and P. putida EM383 (recA$^-$) (H). The average targeting efficiency (%) was calculated by normalizing the CFU numbers obtained with targeting spacers, with the CFU numbers obtained with nontargeting spacer (control) (mean ± s.d., n = 3 biological). Targeting spacers 1.1 have a PAM specificity of 5'-NBGTT-3', 1.2 of 5'-NBGVV-3', 2.1 of 5'-NAGTT-3' PAM, and 2.2 of 5'-NAGVV-3'.

G) and 90–100 % in *P. putida* EM383 (Figure 2.4 H).

ScCas9-mediated counterselection boosts the efficiency of ssDNA recombineering in *P. putida*

By converging recombineering and ScCas9-mediated counterselection we developed ReScribe, a method applying the unique features of ScCas9 that held the potential to enhance effectively the net efficiency of recombineering. In this study, Rec2 was maintained as core recombinase given its reported efficiency in a larger variety of *loci*. After one complete cycle of Rec2 recombineering (including recovery and segregation of the mutation), the heterogeneous cell population was subsequently transformed with a pSEVAb62-ScCas9-crRNA_sp targeting the wild type population, and therefore sifting for the engineered cells (Figure 2.5). To verify our hypothesis and in line with our initial recoding objective, we tested the unified ReScribe protocol with three essential genes of the genome of *P. putida* KT2440 that ended in TAG: *mraY*, *pvdJ*, and *bioD*. Individually, these three genes, were effectively mutated with efficiencies higher than 94 % after only one recombineering iteration as demonstrated by MASC-PCR [16, 40] and HRM [48, 49] (Figure 2.6 A, C, and D).

The high efficiency of the system for achieving individual single point mutations prompted us to test if ReScribe would enable the simultaneous mutagenesis of multiple *loci*, which in *P. putida* typically results in very low frequencies when relying on recombineering alone [25]. Two or three oligonucleotides were cotransformed during the recombineering protocol and a single plasmid containing the ScCas9, tracrRNA, and a CRISPR array with the respective two or three spacers was used for selection. To increase our chances of generating a significant population of cells containing all combined mutations, counterselection was applied to different samples that had experienced 1, 2, and 3 cycles of recombineering. As a result of ScCas9's cleavage and the modest efficiencies of Rec2, wild type cells were wiped out from the population leading to plates with a significantly reduced number of colonies (Supplementary Table S8). This deficit was nonetheless outweighed by the high ratios of edited cells granted by ReScribe. After one single iteration, our selection system was already able to easily single out colonies containing two and three simultaneous mutations. Moreover, virtually every colony (97.6 % and 95.2 %

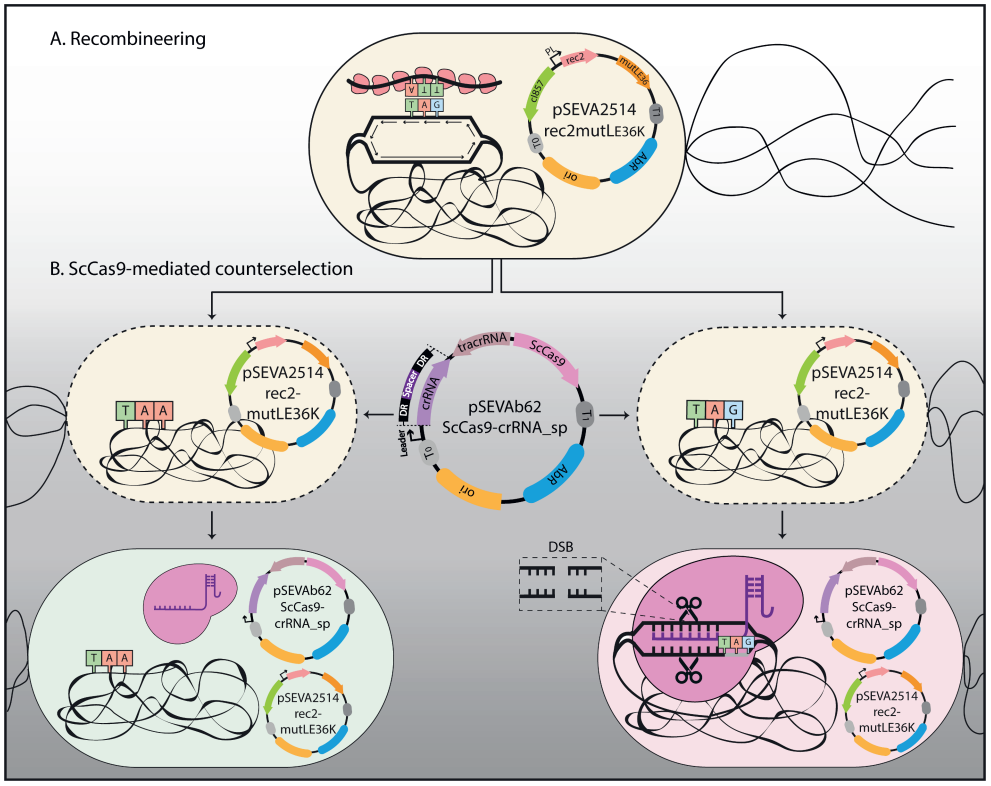


Figure 2.5: Enhanced genome-scale editing in P. putida with ReScribe. (A) The ssDNA oligonucleotide carrying TTA mutation and the pSEVA2514-rec2-mutLE36KPP plasmid, harboring the thermolabile cI857 repressor (green), the rec2 recombinase gene (light red), and mutLE36KPP (light orange) are transformed to P. putida. The expression of both rec2 and mutLE36KPP is controlled by the thermoinducible cI857/PL system. With the increase of temperature from 30 to 42°C, the cI857 repressor is degraded and rec2 and mutLE36KPP are expressed [47]. The expression of these two elements contributes to the incorporation of single nucleotide mutations in the genome of P. putida mediated via oligonucleotides that have been transformed. Consequently, a mix population of wild type and edited cells is generated. (B) The pSEVAb62-ScCas9-crRNA_sp plasmid, constitutively expressing all the CRISPR components (crRNA in light purple, tracrRNA in light brown and ScCas9 in pink), is transformed to the mixed population of P. putida wild type cells (genome with the TAG stop codon) and edited cells (genome with the TAA stop codon). The ScCas9:crRNA:tracrRNA ribonucleoprotein complex with a PAM specificity of 5'-NNG-3', recognizes the TAG as PAM and ScCas9 cleaves both strands. The double strand break (DSB) is lethal for P. putida wild type (cell with the light red background). In contrast, the edited cells have no PAM to be recognized in the site complementary to the spacer and escape ScCas9 activity (cell with light green background).

for two and three simultaneous mutations, respectively) presented all the intended alterations after three recombineering cycles (Figure 2.6 B).

Though most cells without the desired mutations died from the DSB in the chro-

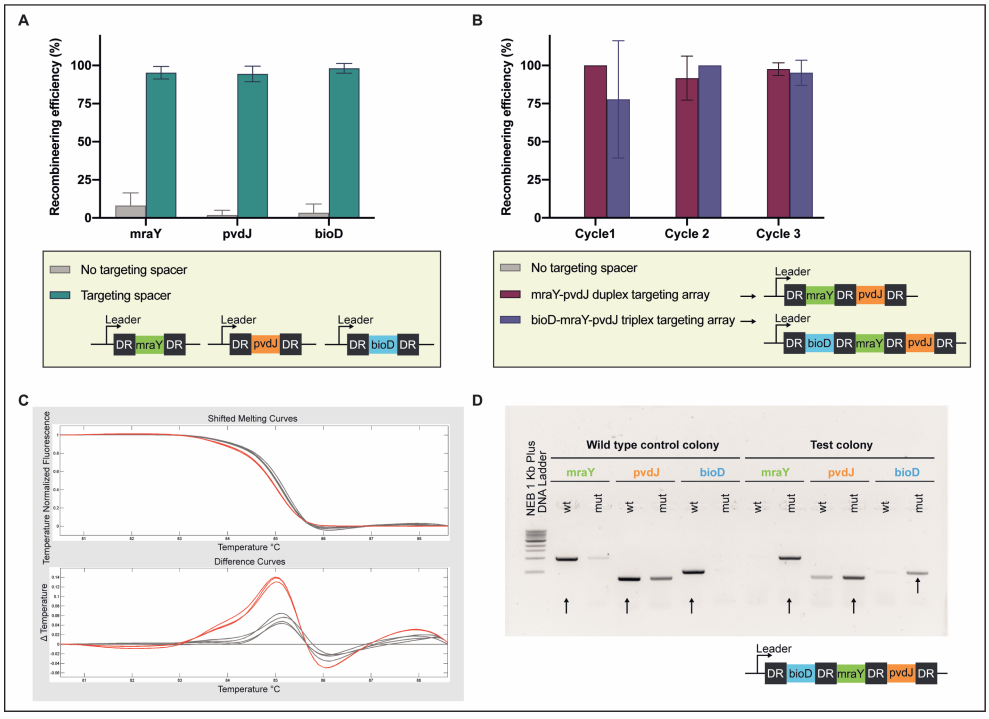


Figure 2.6: Allelic replacement efficiency for single and multiple targets with ReScribe in P. putida KT2440. (A) Single targets. Recombineering was applied during 1 cycle, after which the pSEVAb62-ScCas-crRNA_sp plasmid harboring the targeting spacers was electroporated for counterselection. A non-targeting spacer was used as control. Samples were monitored after recovery steps of cycle 1 and efficiency was calculated for samples with the non-targeting spacer (gray bars) and targeting spacer (turquoise bars) as the ratio between edited and non-edited colonies (mean ± s.d., n = 3 biological). (B) Multiple targets. Recombineering was applied during 1, 2, and 3 cycles, after which the pSEVAb62-ScCas-crRNA_sp plasmid harboring duplex and triplex arrays was electroporated for counterselection. A non-targeting spacer was used as control. Samples were monitored after recovery steps of cycle 1, 2, and 3 and efficiency was calculated for samples with the non-targeting spacer (gray bars), mraY-pvdJ duplex array (dark red bars), and mraY-pvdJ-bioD triplex array (dark purple bars) as the ratio between edited and non-edited colonies (mean ± s.d., n ≥ 2 biological). (C) Single-nucleotide polymorphism TAG to TAA assessment by HRM. Interrogation by HRM analysis of the genotypes of a series of test colonies after mutation of mraY. HRM analysis is performed on PCR amplicons supplemented with a fluorescent dye by monitoring the separation of the two strands of DNA in real-time. Single-nucleotide mutations are observed as two different melt curves, WT control curve (dark red) and test curve (gray), due to the process high resolution. (D) Single-nucleotide polymorphism TAG → TAA assessment by MASC-PCR. Comparison of MASC-PCR binary results between a WT and a test colony with three targeted loci (mraY, pvdJ, bioD). Screening of each mutation is performed in two reactions: one with a FW_{WT} and RV pair of primers (wt) and another one with a FW_{mut} and RV pair (mut). FW_{WT} and FW_{mut} primers are identical differing only in the 3'-terminal base which can be either a G or an A, consequently annealing to the WT or the mutant genotype, respectively.

mosome caused by the pSEVAb62-ScCas9_crRNA_sp targeting plasmid, a small percentage was able to escape this lethal cleavage (Figure 2.5 B). In addition, smaller colonies often appeared on selection plates upon prolonged incubation longer than the standard 24 h, which turned-out to be false positives. These colonies are easily differentiated by visual inspection from the edited ones given their small size and late apparition.

To get a full understanding of these results, we further analyzed the aforementioned escapers by reculturing them on selection plates and sequencing all the CRISPR elements of the pSEVAb62-ScCas9_crRNA_ mraY-pvdJ-(bioD) targeting plasmids. From the reculturing experiments, only few colonies were able to fully grow again in selection plates. Those were subsequently grown in liquid cultures and their plasmids were isolated and sequenced. Sequencing results showed miscellaneous cases of mutations and reorganizations, including recombination between the direct repeats of the CRISPR array or complete deletion of CRISPR-Cas9 machinery elements such as the tracrRNA, that would inactivate the pressure of the counterselection plasmid.

Recombineering *vs.* ReScribe to construct a minimally recoded *P. putida* strain

Given the efficient and multisite recombineering possibilities granted by ReScribe, our next goal was to compare the standard Rec2-mediated recombineering and ReScribe in terms of efficiency and time, and to construct a minimally recoded *P. putida* strain. For such end, we used the informatic analysis previously described in this study and we aimed at recoding all predicted TAG codons (12 in total) that reside in conditionally essential genes in glucose minimal media.

As a proof of concept, we completed the recoding of all TAG codons by editing them to the synonymous TAA codon, generating the minimally recoded *P. putida* KT2440Rc12 strain by using both standard Rec2-mediated recombineering and ReScribe. Thereby, we highlighted the benefits of ReScribe in speed and efficiency when compared to the previous standard technique. The first six mutations of *P. putida* KT2440Rc12 (*dxs*, *pvdT*, *vdh*, *bioA*, *cobK*, *murA*) were introduced individually and consecutively *via* standard Rec2-mediated recombineering. This typically required 6 working days per mutation with an average efficiency of 8.3 ± 2.8 %. Next,

three mutations (*ubiB*, *wbpL*, *ompQ*) were performed making use of single-targeting ReScribe resulting in a decrease in working time from 6 to 3 days per mutation and a considerable increase of average efficiency to 90.5 ± 9.9 %. Possibly due to a high plasmid burden imposed by the different targeting pSEVAb62-ScCas9_crRNA_(spacer) plasmids, easy isolation of plasmid-cured colonies with an efficiency of 100 % within 24 h was possible after two rounds of antibiotic-pressure-free media passaging. Finally, multiplex ReScribe was utilized for the simultaneous recoding of the three remaining genes (*bioD*, *mraY*, *pvdJ*) which was achieved in 3 days with an efficiency of 77.8 ± 38.5 % (Figure 2.7 A). In comparison, ReScribe reduced the working time of standard Rec2-mediated recombineering to the half in its single-targeting version, and 6-fold when multiplexing, while it increased ~10-fold the efficiency levels.

Once completed (Figure 2.7 B), the minimally recoded strain showed an unaffected fitness with equal doubling time to its ancestor, both in LB and M9-glucose (Figure 2.7 C). Beyond the rapid screening performed with MASC-PCR (Figure 2.7 D) and HRM, whole-genome sequencing confirmed the presence of all 12 mutations in the final KT2440Rc12 strain (Supplementary Table S9). This analysis revealed as well the presence of 40 off-target mutations when compared with the reference *P. putida* KT2440 genome (accession no GCF_000007565.2) and the sequenced genome of an in-house reference strain from our laboratory (Supplementary Table S9). Considering that KT2440Rc12 had undergone at least 34 recombineering cycles (6 mutations with 5 cycles of standard ssDNA recombineering + 3 mutations with 1 cycle of single ReScribe + 3 mutations with 1 cycles of multiplex ReScribe; Figure 2.7 A), the average number of off-target mutations would be 1.17 per recombineering cycle.

Discussion

We developed ReScribe as a highly efficient method for multiplex recombineering of *P. putida*. The key element of ReScribe is the deployment of the minimal-PAM CRISPR-ScCas9 system, which provides single base-pair resolution and therefore permits the counterselection against wild type genotypes after introducing single-nucleotide polymorphisms. Taking advantage of the here validated minimal PAM

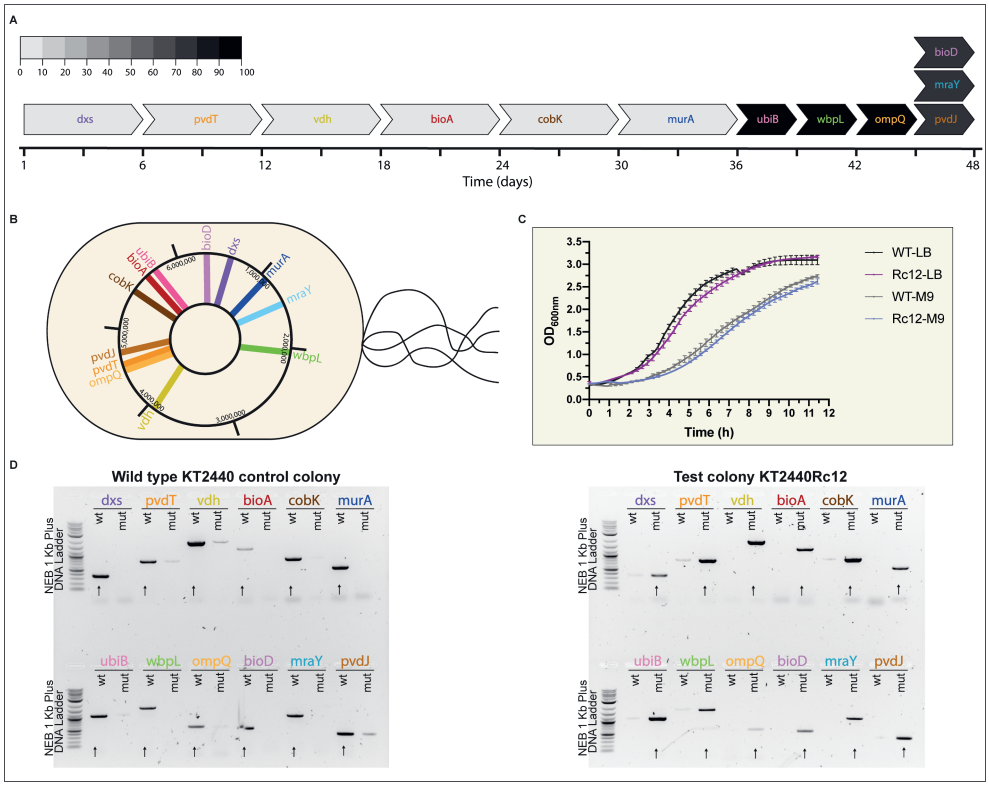


Figure 2.7: Minimally recoded P. putida KT2440Rc12 strain. (A) Timeline of the sequential mutation of the TAG stop codons in metabolic essential genes. Required time for performing each mutation is represented in days in the X axis while recombineering efficiency is depicted in percentage according to the gray scale. (B) Illustration of P. putida KT2440Rc12 genome portraying metabolic essential genes in glucose minimal medium terminated by TAG stop codons (dxs, pvdT, vdh, bioA, cobK, murA, ubiB, wbpL, ompQ, bioD, mraY, and pvdJ). (C) Fitness comparison between KT2440 (WT-LB) and KT2440Rc12 (Rc12-LB) in LB and M9-glucose media (WT-M9 and Rc12-M9). Data represent OD_{600} over 12 h (mean ± s.d., $n \geq 3$ biological replicates). (D) Comparison of MASC-PCR binary results between a wild type KT2440 control colony and a test colony KT2440Rc12. The wild type control colony showed stronger bands in the PCR reactions with the wt set of primers than in the PCR reactions with the mut set of primers, indicating that the genotype is wild type for all 12 genes. The test colony showed stronger bands in the PCR reactions with the mut set of primers than in the PCR reactions with the wt set of primers, indicating that the TAG stop codons of all 12 genes have been mutated to TAA.

5'-NNG-3', very precise and specific *loci* can be targeted in a highly efficient manner, which is otherwise impossible with CRISPR systems with more restricting PAMs. By applying ReScribe, we edited the genome of *P. putida* by substituting native TAG stop codons with the synonymous TAA stop codon with unprecedented efficiencies, 90–100 %, for both single and multiplex genome engineering. As a result, we built a

minimally recoded *P. putida* KT2440 strain of essential metabolic genes, establishing the first step toward a whole-genome recoding process. The need for ReScribe was the result of failed attempts to increase the recombineering efficiency of the current reference Rec2 recombinase in *P. putida* (10 % efficiency for single targets and 2×10^{-4} % and 6×10^{-6} % efficiency forfour and five targets, after 10 recombineering cycles) [25]. In this study, different factors were tested: (i) strandedness, structure, length, and backbone modifications of the recombineering oligonucleotides (Figure 2.2), (ii) expression of the Rec2 recombinase with strong synthetic RBSs (Figure 2.3), and (iii) PapRecT as an alternative recombinase to the baseline Rec2 (Figure 2.3). However, none of those changes led to significantly increased recombineering efficiencies, which are imperative for genome-scale applications such as our intended genome recoding.

The power of recombineering relies on the ability to rapidly edit the genome of organisms with high accuracy on a scale that is not feasible with previous traditional tools. For such an end, exceptionally high editing levels are crucial [50]. Our findings demonstrate the great utility of the codon optimized ScCas9 variant to boost recombineering efficiencies in our system, which could be potentially reproduced in bacteria beyond *P. putida*. What makes ScCas9 and ReScribe especial, as compared to other CRISPR-Cas9-mediated counterselection systems, is the minimal PAM 5'-NNG-3'. It is worth mentioning that the ScCas9Δloop variant tested in this study was equally able to cleave targets adjacent to 5'-NNG-3' PAMs despite previous *in silico* predictions that assigned it the more restricted 5'-NAG-3' PAM specificity [30]. However, these results align with the *in vitro* assessment of the authors in human cells in which ScCas9Δloop was able to efficiently cleave at 5'-NGG-3', 5'-NNGA-3' and 5'-NNGN-3' targets [30]. While in this study the two variants were tested due to the suitability of both 5'-NAG-3' and 5'-NNG-3' PAMs for the aimed objective, the latter PAM represents the most convenient choice for a broader applicability of the tool. This less restrained requirement expands dramatically the number of targetable sites and becomes critical when there is no flexibility in selecting protospacer sequences (which must be followed by the PAM sequence). ReScribe therefore represents an auspicious opportunity for those bacteria that, like *P. putida*, show limited recombineering activity with the currently available recom-

binases [51–55]. Moreover, the potential niches of application of ReScribe go even further: given the high efficiency of the ScCas9 cleavage, the tool could be deployed for recalcitrant targets [56, 57] or for facilitating the process in bacteria with fully established recombineering systems such as *E. coli*. On top of that, ReScribe has demonstrated the feasibility of efficient multisite genome editing, which, in turn, enables further genome-scale engineering applications.

The major drawbacks of ReScribe are intrinsically connected to those of the parental techniques: Rec2-mediated recombineering and CRISPR-Cas9 technologies. In the first place, the efficiency of recombineering varies with the relative location of the gene. In this context, cold and hotspots for recombineering have been identified in other studies [58]. Additionally, the position of the target *loci* with respect to the two replichores may have an effect too. Thus, genes closer to the origin of replication will be edited at higher levels than those located farther away [59, 60]. Moreover, the recombineering efficiency can also be affected by the nucleotide composition of the mutagenic oligonucleotide. Besides, in our particular case, the specific efficacy of Rec2 might be an extra limitation for the recombineering part of the workflow, especially for multiplexing experiments. Nonetheless, if any better alternative should exist, the core recombinase could easily be replaced, either for a better performance in *P. putida* or for the application of the pipeline in other organisms. The rapid emergence of high-throughput methods for surveying complex libraries of recombinases [26] holds the potential of finding more suitable candidates, especially among Pseudomonas-borne counterparts for this particular case. At this point, account has to be taken of the intrinsic limitations of *P. putida* KT2440 as a receptor of exogenous DNA. While a better recombinase could improve the overall efficiency of the protocol, the poor ability of this bacterium to capture synthetic ssDNA could be an even greater constraint for recombineering in this species [25]. Besides recombinase expression levels, recent studies highlighted additional interactions with single-stranded DNA-binding proteins (SSBs) within the replication fork as a way to improve recombineering efficiency in a given host [61, 62]. On the basis of results reported in other *Pseudomonas* species, coexpression of these SSBs together with recombinases might be an option for enhancing the allelic replacement efficiencies even further [63].

Ultimately, increasing the efficiency of recombineering would be desired, not only as a way for eliminating the need of counterselection, but also for exploring higher-order selections than those showed in this work. Considering the number of required mutations, these orders of multiplexing (*i.e.*, penta- to decaplex) would become useful for genome-scale editing applications. While coexpression of a high number of crRNAs might be a challenge, it would not be impossible from multiple expression units. However, the generation of a cell population containing all the mutations, high enough to be set apart after the counterselection step, remains the most important obstacle, given the reported number of colonies that were obtained in our experiments with triplex selection.

Furthermore, ReScribe still has some of the limitations of the CRISPR technology, including (i) the need for constructing CRISPR plasmids directed at each modification *locus*; (ii) variations in the efficiency of the spacer, probably caused by differences in the secondary structure of the crRNA, which depends on the nucleotide composition of the spacer [58]; and (iii) the loss of functionality of the CRISPR plasmid caused by homologous recombination between the direct repeats or by mutating one of the CRISPR elements [28, 29, 64]. Regarding possible substitutions in the CRISPR-Cas9 counterpart, the recently engineered Sc2+ and HiFi-Sc2+ optimized ScCas9 variants could enhance the performance and robustness of ReScribe [65].

Lastly, the significant number of 40 off-target mutations was found in the genome sequence of KT2440Rc12. While the overall count could be considered high, it is important to have into account the elevated number of recombineering cycles to which this strain was subjected. If we would consider off-targets per cycle, the average number of instances would be 1.17, which is only slightly higher than recent reports about *E. coli* cells expressing a redβ-recombineering system (1.0 ± 0.7 off-targets per recombineering cycle), and significantly lower than *E. coli* cells containing the recombinase PapRecT (3.3 ± 0.6) [26]. While the accumulation of off-target mutations remains one of the main limitations of highly efficient recombineering systems [16, 66], the number of off-targets observed in this study aligns with those obtained by other systems with considered low off-target mutagenesis, such as pORTMAGE2, 3, or 4 [24, 66]. Nevertheless, this negative effect deserves further investigation and needs to be dealt with special prudence when applying ReScribe

to large-scale applications such as genome recoding.

In the context of recoding, the aforementioned limitations could be eventually overcome by following a strategy of high-fidelity, total genomic synthesis. Contemporary DNA synthesis and assembly methods have enabled the generation of entire or largely synthetic genomes of *Mycoplasma genitalium* and *Mycoplasma mycoides* [67, 68], *Saccharomyces cerevisiae* [12, 69], and *E. coli* [14]. Although those works provide a blueprint for future efforts, costs of whole-genome synthesis remain prohibitively expensive, accuracy of such lengthy sequences keep not being guaranteed, and assembly methods are still limited or insufficient in some particular organisms, thus hindering such approaches in most recoding undertakings [15].

The tolerated stop codon changes already performed in the strain for essential genes support the feasibility of a whole-genome recoded project in *P. putida*. With ReScribe, this herculean enterprise could be carried out in separate strains in order to accelerate the process. Each strain can be used to recode a section of the genome (*e.g.*, 100 kb) which can then be assembled together in a single chromosome. For this convergence, the large edited genomic fragments can be captured in BAC (or YAC) plasmids and transformed into the recipient cell to replace the corresponding non-edited fragment using a CRISPR/Cas9-based strategy [14] or a recombinase-mediated cassette exchange [70].

Overall, fully recoding *P. putida* would be a major step toward a new chemical landscape by enabling it to maximize and expand its attractive metabolic possibilities for bioproduction. The ability of reprogramming codons to encode alternative amino acids will allow the exploration of neo-transmetabolisms, with the incorporation of elements and biochemistries beyond the cell's customary repertoire, *e.g.*, silicon or halogens such as fluorine [7, 71]. These ncAA confer at the same time the opportunity of implementing a powerful genetic safeguard addressing both (i) biological isolation in defined environments with a supply of the ncAA, and (ii) genetic isolation by preventing horizontal gene transfer (HGT) of the neo-transgenes between organisms and across species. Such biosafety credentials would contribute to expand the possibilities for risk management of the strain and therefore could serve as a prelude for a more suitable and realistic consideration of *P. putida* for non-contained environmental applications [72].

While this work was conceived as a means of developing an efficient tool for recoding *P. putida*, we just had a glimpse of the capabilities of ReScribe. Our results support the hypothesis that counterselection can enhance the efficiencies of recombineering to nearly absolute levels in a multiplex manner and in an increased targetable space. ReScribe is therefore not limited by the size of the edit, the necessity of targeting gene by gene, or the location of a complex PAM. In addition, the proposed pipeline is neither restricted to *P. putida* nor hampered by the native mismatch repair machinery, which makes it a conveniently unrestrained tool for highly efficient engineering of arduous targets and endeavors.

Materials and methods

Bacterial strains and media

All bacterial strains with their respective characteristics used in the present study are listed in Supplementary Table S1. *E. coli* DH5α and BL21 cells were made chemically competent as previously described [73]. While the first were used for cloning purposes, fluorescence loss assays, and cleavage assays, the latter were only utilized for cleavage assays. Subsequently, electrocompetent *P. putida* strains were prepared as previously described [74] and used for cleavage assays and recombineering experiments. Unless otherwise stated, *P. putida* and *E. coli* were cultured on LB (10 g/L NaCl, 10 g/L tryptone, and 5 g/L yeast extract) medium at 30 and 37°C, respectively. Antibiotics were added when required, at the following concentrations: kanamycin, 50 mg/L; gentamicin, 10 mg/L; chloramphenicol, 20 mg/L; streptomycin, 50 mg/L and 100 mg/L for *E. coli* and *P. putida*, respectively. Fluorescence loss assays were performed on M9 minimal medium (1.63 g/L NaH_2PO_4, 3.88 g/L K_2HPO_4, 2 g/L $(NH_4)_2SO_4$, 10 mg/L EDTA, 100 mg/L $MgCl_2 \cdot 6\ H_2O$, 2 mg/L $ZnSO_4 \cdot 7\ H_2O$, 1 mg/L $CaCl_2 \cdot 2\ H_2O$, 5 mg/L $FeSO_4 \cdot 7\ H_2O$, 0.2 mg/L $Na_2MoO_4 \cdot 2\ H_2O$, 0.2 mg/L $CuSO_4 \cdot 5\ H_2O$, 0.4 mg/L $CoCl_2 \cdot 6\ H_2O$, and 1 mg/L $MnCl_2 \cdot 2\ H_2O$) supplemented with 70 mM of glucose. Recombineering experiments were performed on TB medium (12 g/L tryptone, 24 g/L yeast extract, 0.4 % (v/v) glycerol, and 10 % (v/v) phosphate buffer (23.12 g/L KH_2PO_4 and 125.4 g/L K_2HPO_4)).

Construction of *P. putida* Tn7GFPstop strain

To assess the efficiency levels of allelic replacement, easily selectable mutations had to be selected for our protocol. These mutations should have a visual phenotypic readout for easy screening, *e.g.*, antibiotic resistance, change of color, or fluorescence emission. With this objective in mind, we generated *P. putida* Tn7GFPstop. This strain was created by introducing a *gfp* gene cassette with a *gfp* ORF disrupted by a TAG stop codon replacing Tyr66 (Supplementary Table S5), in the attTn7 landing site of the *P. putida* KT2440 genome. By mutating back the introduced TAG stop codon into the original sequence, the *gfp* coding sequence would be restored and thus the strain would become fluorescent. The cassette was integrated following a previously described protocol for I-SceI-mediated homologous recombination [75]. In brief, the cassette was first amplified with primers M46 and M47 (Supplementary Table S3) from pSB1C3 disrupted *gfp* and cloned into the pGNW suicide vector (amplified with M40 and M41) between 500-bp upstream and downstream regions (amplified with M42-M43 and M44-M45, respectively) of the attTn7 landing site. *P. putida* KT2440 cells were transformed with pGNW *via* electroporation and resultant fluorescent colonies (cointegrates) were grown for pSEVA628 I-SceI vector transformation. Expression of I-SceI meganuclease was induced with 3-methylbenzoate mediating the excision of pGNW from the genome, leading to nonfluorescent colonies. Final clones were tested for revertant (*i.e.*, wild type) or mutant (*i.e.*, knock in) genotype.

Plasmids

Plasmids used in the present study are fully described in Supplementary Table S2. All PCR reactions for cloning purposes were performed with the NEB Q5 High-Fidelity DNA polymerase, according to manufacturer's instructions (M0491). PCR fragments were subjected to 1% w/v agarose gel electrophoresis, and isolated using Nucleospin Gel and PCR Clean-up (BIOKÉ) kit. Plasmids were built using the SevaBrick Assembly method [76], unless otherwise stated, and introduced by heat-shock in chemically competent *E. coli* DH5α cells. Plasmids were isolated using the GeneJET Plasmid Miniprep Kit (Thermo Scientific) and colony PCR was performed to verify the right assembly of the different fragments. Plasmid sequence was con-

firmed by Sanger sequencing from Macrogen (MACROGEN Inc. DNA Sequencing Service; Amsterdam, The Netherlands).

Recombineering experiments were performed using pSEVA2514-rec2-mutLE36KPP, pSEVA2514-rec2RBSopt-mutLE36KPPRBSopt, pSEVA2514-paprecT-mutLE36KPP, and pSEVA2514-paprecTRBSopt-mutLE36KPPRBSopt. The pSEVA2514-rec2-mutLE36KPP plasmid was a kind gift from the Molecular Environmental Microbiology Laboratory (CNB-CSIC) of Madrid (GenBank #MN180222) and was used as reference vector. The pSEVA2514-paprecT-mutLE36KPP plasmid was built *via* Gibson Assembly using NEBuilder HiFi DNA Assembly Master Mix by substituting *rec2* with *paprecT*. The backbone was amplified with M70–M71 primers from the reference vector, and *paprecT* was amplified with M72–M73 primers from pORTMAGEE502B, which was purchased from Addgene (#128971). Strengthening of RBSs for *rec2*, *paprecT*, and *mutLE36KPP* genes was done by predicting the strongest RBS upstream of the ORFs with the automated design tool De Novo DNA [77] and can be found in Supplementary Table S5. Putatively optimal RBSs were incorporated in primers M76, M78, and M80, which were used together with M77, M79, and M81, to amplify *rec2*, *paprecT*, and *mutLE36KPP*, respectively. RBSopt amplicons were cloned in a three-part ligation (recombinase + *mutLE36KPP*) into the linear backbone amplified with M74-M75 primers from the reference vector, for the construction of pSEVA2514-rec2RBSopt-mutLE36KPPRBSopt and pSEVA2514-paprecTRBSopt-mutLE36KPPRBSopt.

Fluorescence loss assays were performed using pSEVAb62-ScCas9, pSEVAb23-RhaBAD-crRNA_sp and pSEVAb44-sfGFP. The gene encoding the ScCas9 was codon optimized for *P. putida* using the Jcat codon optimization tool (www.jcat.de) (Supplementary Table S5). A pCCI-4k plasmid with the optimized *sccas9* gene was synthesized and delivered by GenScript. The pSEVAb62-ScCas9 plasmid was built by PCR amplifying the ScCas9 gene, tracrRNA and their respective promoters (cargo of the pCCI-4K plasmid) with Ep-Pp primers and cloning the fragment into a linearized pSEVAb62 backbone with Ev-Pv primers. The pSEVAb23-RhaBAD-crRNA_eforRed was built by using the pSB1C3-RhaBAD and pSEVAb23-crRNA_eforRed in-house plasmids. The pSB1C3-RhaBAD plasmid was used to amplify the rhamnose inducible promoter together with the activators rhaS and rhaR (rhaSR-PrhaBAD) with

Ep-Sp primers. The pSEVAb23-crRNA_eforRed plasmid was used to amplify the leader sequence and crRNA array with Xp-Pp primers. The crRNA is composed of two directed repeats interspaced by the transcriptional unit P_{J23100}-RBS_{BBa_B0034}-$eforRed_{BBa_K592012}$, which, in turn, is flanked by two BsaI sites [78]. The two PCR amplified fragments were cloned into a linearized pSEVAb23 backbone with Ev-Pv primers using the previously mentioned SevaBrick Assembly method with some modifications. The customary enzyme deactivation step at 80°C was replaced by a process halting at 16°C. The desired spacers were introduced in pSEVAb23-RhaBAD-crRNA_eforRed using the previously described protocol called One-step Golden Gate-based cloning for the assembly of single and multiple spacers into the crRNA cassette [78], by replacing the eforRed chromoprotein (pSEVAb23-RhaBAD-crRNA_ sp). The pSEVAb44-sfGFP plasmid was built by cloning the transcriptional unit, PJ23106-RBS-sfGFP (amplified from pSB1C3-sfGFP in-house plasmid with Ep-Pp primers) into pSEVAb44 backbone (amplified with Ev-Pv primers).

Cleavage assays were performed using (i) two plasmid system: pSEVAb62-ScCas9/pSEVAb62-ScCas9Δloop and pSEVAb23-crRNA_sp and (ii) one plasmid system: pSEVAb62-ScCas9-crRNA_sp. The pSEVAb62-ScCas9Δloop plasmid was built by removing the loop D367–376 from the ScCas9 using the pSEVAb62-ScCas9 plasmid as template. Two PCR fragments were created using 597-M160 and 594-M161 primers and ligated *via* SevaBrick Assembly.

The pSEVAb23-crRNA_amilCP plasmid was built using pSEVA231-CRISPR as template [28]. The leader sequence and crRNA array were amplified with crRNA-F-crRNA-R primers. The translational unit, comprised by the BBa_J23100 Anderson promoter (P_{J23100}), the BBa_B0034 RBS (RBS_{BBa_B003}) and the amilCP, blue chromoprotein (BBa_K592009), was amplified with 804–940-primers from pSB1C3-amilCP in-house plasmid. The two PCR amplified fragments were cloned into a linearized pSEVAb23 backbone with 475–476 primers using the previously mentioned SevaBrick Assembly method with some modifications. The customary enzyme deactivation step at 80°C was replaced by a process halting at 16°C. As result, the crRNA array is comprised by two directed repeats interspaced by the transcriptional unit P_{J23100}-RBS_{BBa_B0034}-$amilCP_{BBa_K592009}$, which, in turn, is flanked by two BsaI sites [78]. The pSEVAb62-ScCas9-crRNA_eforRed plasmid was built using the NEBuilder

HiFi DNA Assembly Master Mix. A constitutive version of the pGCRi-R [78] was used to amplify the crRNA array (P_{J23100}-RBS_{BBa_B0034}-$eforRed_{BBa_K592012}$) with M94–M95 primers. The PCR amplified crRNA array was cloned into pSEVAb62-ScCas9, linearized with M92–M93 primers. The desired spacers are introduced in pSEVAb23-crRNA_amilCP and pSEVAb62-ScCas9-crRNA_eforRed as previously described for the pSEVAb23-RhaBAD-crRNA_eforRed plasmid. All the spacers used in the present study can be found in Supplementary Table S4.

Oligonucleotides

Single-stranded (ss) DNA oligonucleotides employed in this study (Supplementary Table S3) were ordered from Integrated DNA technologies (IDT) as salt-free without further purification, resuspended in milli-Q at 100 μM and long-term stored at -20°C.

Recombineering oligos (Supplementary Table S3) were designed to be complementary to the lagging strand of replicating DNA and according to the optimized design criteria shown in the Supporting Information and Supplementary Figure S2.1. In sum, they were 60 nt long and carried mutation changes at the middle positions of their DNA sequence; predicted folding energies were higher than ≥ 16 kcal/mol [22], and no phosphorothioate bonds were included in the sequences.

Design of optimized recombineering oligonucleotides

In order to optimize the design of recombineering oligonucleotides, different parameters were considered.

(i) Strandedness: Recombineering oligos were designed to anneal to the lagging strand of the replication fork since hybridization of ssDNA is supposed to occur there according to the principles of recombineering [21].

(ii) Structure: ssDNA with higher predicted ΔG score is suggested to recombine at higher frequencies [39]. While the optimal folding energy for *E. coli* has been reported to be ~12.5 kcal/mol [40], the higher GC content and lower optimal growth temperature (30°C) suggest that the optimal range could be different for *P. putida* [22]. According to the DNA folding predictor tool mfold-UNAFold [79], all the recombineering oligos used in this study had a folding energy ≥ 16 kcal/mol.

(iii) Length: Oligonucleotides of 90, 60, and 40 nt were tested by using recombineering with the K43T mutation in the *rpsL* gene that confers resistance to streptomycin. After the optimization study, the oligos were designed of 60 bp in length, with the desired mutations included in the middle of the sequence.

(iv) Backbone modifications: For the oligonucleotide optimization study, oligos both with and without phosphorothioate bonds were tested with the K43T mutation in the *rpsL* gene. Accordingly, subsequent oligonucleotides did not include phosphorothioate bonds.

Recombineering cycling protocol

Recombineering experiments were performed according to the previously described standard protocol [80] based on the coexpression of a recombinase and a mismatch repair machinery disruptor, both under the control of the thermoinducible cI857/PL expression system. An overnight culture of *P. putida* KT2440 harboring pSEVA2514-recombinase*-mutLE36KPP (* indicates any of the recombinases used in this work: Rec2, PapRecT and their respective RBS variations) was grown in 20 mL of LB supplemented with kanamycin. The next day, bacterial cultures were diluted to an OD_{600} of 0.1 in LB-kan and incubated at 30°C, 200 rpm until an OD_{600} of 0.5–0.7 (mid log phase). Once the appropriate OD_{600} was reached, recombinase and *mutLE36KPP* transcription was thermoinduced by 10 min incubation at 42°C in a shaking water bath. *P. putida* cultures were chilled on ice for 5 min and harvested by centrifugation at room temperature and 4700 g for 10 min. Subsequently, cells were made electrocompetent by consecutive washing steps of 10, 2, and 1 mL of sucrose 300 mM. The washed cultures were finally resuspended in 200 μL of sucrose 300 mM. 100 μL of electrocompetent cells were transformed with 1 μL of recombineering oligo (100 μM). Samples in which CRISPR-ScCas9 counterselection was applied were transformed additionally with 100 ng of pSEVAb62-ScCas9-crRNA_sp carrying the appropriate spacer, in addition to the corresponding recombineering oligonucleotide. Electroporation was performed in 2 mm gap Bio-Rad electroporation cuvettes. A single exponential decay pulse was applied using a Gene Pulser X-Cell (Bio-Rad) set at 2.5 kV, 200 Ω and 25 μF. Cells were first resuspended in 5 mL of terrific broth (TB) with kanamycin and recovered for 1 h. Afterward, 15 mL of LB-kan

was added to the transformed cells. Cultures were grown to an OD_{600} of ~0.4, and stored at 4°C until the next day. In consecutive cycles, cultures stored at 4°C were reactivated by ~30 min incubation at 30°C and 200 rpm until an OD_{600} of 0.5–0.7 before continuing with subsequent recombinase and *mutLE36KPP* induction.

Before storing the cultures at 4°C, 1 mL of the bacterial cells with OD_{600} of ~0.4 from cycles 1, 4, 7, and 10 was inoculated into 2 mL LB-kan and grown at 30°C and 200 rpm, overnight. Appropriate dilutions from the overnight cultures were plated for screening and subsequent efficiency calculation. Cultures without thermoinduction were included as negative controls. Editing efficiencies of non-thermoinduced controls were subtracted from those of the thermoinduced samples to calculate final recombineering efficiencies as it is assumed that such background levels are not directly derived from the action of the tested recombinases [25].

Screening of cells edited with recombineering

On the basis of the phenotypic outcome of the introduced mutations, three different readouts and screening methods were analyzed in this study.

First, *P. putida* KT2440 clones edited in the *rpsL* gene (PP_0449) were engineered with oligo RO rpsL 60 for the change of the AAA codon (Lys43) by ACA (Thr43) and were screened by the resistant to streptomycin conferred by this change. Recombineering efficiency was calculated as the ratio between streptomycin resistant and total CFU. Screening was performed using both LB and LB-sm plates [23].

Second, *P. putida* Tn7GFPstop clones edited in a heterologous and disrupted *gfp* gene were engineered with oligo RO gfp stop in order to revert the functional expression of the green fluorescence reporter (Stop66Tyr). Recombineering efficiency was calculated as the ratio between fluorescent (mutated) and total CFU. Screening was performed in LB-kan agar plates after ~48 h, allowing GFP maturation.

Lastly, efficiencies of recombinant cells without screenable or selectable phenotypes (TAG → TAA mutants) were determined by multiplex allele-specific colony PCR (MASC-PCR) [16, 40] or high-resolution melt analysis (HRM) [48, 49]. For MASC-PCR, three primers were designed for each targeted *locus*: (i) FW primer specific to the wild type genotype, (ii) FW primers specific for the mutant genotype, and (iii) RV primer common to both. The two FW primers only differed at their 3'-terminal

bases allowing discrimination of single nucleotide changes. Two MASC-PCR reactions were required to screen each colony: one to test the wild type genotype and one to test the mutant genotype. Colony genotype was therefore revealed by the binary result yielded by the two reactions. When possible, several *loci* were interrogated in a single reaction by designing different primer sets with the same melting temperature but different amplicon length. In this case, primer pools were used for the PCR reactions. For HRM, specific primers for 100 bp amplicons were used for colony PCR supplemented with LCGreen Plus+ Melting Dye. Since exact melting temperatures of DNA molecules are determined by their nucleic acid sequence, differences between amplicon samples even with only one single nucleotide variation result in melting profiles that are unique to these particular genotypes, allowing for differentiation between amplicons containing the TAG WT and the TAA mutant genotypes. After amplification, samples were transferred to a LightScanner Instrument (BIOKÉ) for melting and acquisition of melt curves, and subsequent data analysis was performed by the LightScanner software. Both types, MASC and HRM colony PCRs, were performed with Phire Hot start II DNA Polymerase (Thermo Fisher Scientific) according to manufacturer's guidelines. Recombineering efficiency was ultimately calculated as the ratio between mutants and total CFU.

Fluorescence loss assays

P. putida KT2440 (*recA*$^+$), harboring pSEVAb62-ScCas9, pSEVAb23-RhaBAD-crRNA_sp with the desired spacer and pSEVAb44-sfGFP, were grown at 30°C and 200 rpm, overnight in 10 mL LB media supplemented with kanamycin, gentamycin, and streptomycin. Overnight cells were harvested at 4700g for 10 min and washed with minimal M9 medium in order to eliminate LB traces. Cells were resuspended to an OD_{600} of 0.3 and grown aerobically at 30°C in fresh minimal M9 medium supplemented with 70 mM of glucose and the appropriated antibiotics (kanamycin and gentamicin at 50 and 10 μg/mL, respectively) on 96-well black wall and transparent round-bottom plate in a total volume of 200 μL per well. Additionally, 5 mM of l-rhamnose was added to the media under induced conditions. Optical density (OD_{600}) and green fluorescence (excitation 467 nm, emission 508 nm) readings were monitored in a BioTek Synergy Mx Multi-Mode Microplate reader over 24 h. Fluores-

cence values were normalized to OD_{600} values. Biological and technical triplicates were included.

Cleavage assay

Strains *P. putida* KT2440 (*recA*$^+$), *P. putida* EM383 (*recA*$^-$), *E. coli* BL21 (*recA*$^+$) and *E. coli* DH5α (*recA*$^-$), harboring pSEVAb62-ScCas9 or pSEVAb62-ScCas9Δloop were used in the cleavage assays based on two-plasmid system. Electrocompetent *P. putida* and chemical-competent *E. coli* cells were transformed with 100 ng of pSEVAb23-crRNA_sp (with different targeting spacers) and plated in LB-Kan-Gen solid agar media. The pSEVAb23-crRNA_nt plasmid harbors a non-targeting spacer that does not target any region in the genome and was used as control. Strains *P. putida* KT2440 (*recA*$^+$) and *P. putida* EM383 (*recA*$^-$) were used in the cleavage assays based on one-plasmid system. pSEVAb62-ScCas9-crRNA_sp with non-targeting spacer (control) and targeting spacers were transformed in electrocompetent *P. putida* strains. Targeting efficiency for both strategies was calculated by as the percentage of surviving CFU present in plates transformed with targeting spacers divided by the number of CFU present in plates transformed with the non-targeting spacer. All transformations were repeated at least two times.

Fitness and toxicity assays

To measure the fitness of the strains and the toxicity of ScCas9 for *P. putida*, growth assays were conducted in LB and M9-glucose media in an Elx808 Absorbance Microplate Reader (BioTek Instruments, Inc., VT, U.S.). Optical density at 600 nm was monitored for 24 h after seeding with 200 μL cultures at OD_{600} = 0.1.

Whole-genome sequencing

In order to confirm the mutated *loci* of *P. putida* KT2440Rc12 and to measure off-target mutagenesis, gDNA of the strain was isolated by using the GenElute Bacterial Genomic DNA kit (Sigma-Aldrich). Extracted gDNA was sent for sequencing to Novogene Co. Ltd. (Beijing, China) for Illumina sequencing. Raw lllumina reads were trimmed for low quality and adapters with fastp (v0.20.0). Mutations were found using breseq (v0.35.5) using the reference genome and annotation of *Pseu-*

domonas putida KT2440 (GCF_000007565.2). To calculate the number of off-target mutations, the total number of non-intended mutations was divided by the number of recombineering cycles performed in the KT2440Rc12 strain, after removing all those mutations also present in the sequence of an in-house reference strain that was subjected to the same whole-genome sequencing and analysis.

Acknowledgements

We would like to thank Víctor de Lorenzo, Tomás Aparicio, and Esteban Martínez-García for providing the reference vector pSEVA2514-rec2-mutLE36KPP and, most importantly, for their valuable advice and comments throughout the different stages of this study. We wish to convey our gratitude as well to Farren Isaacs for the feedback and suggestions within the project; to Christos Batianis for his help with the *P. putida* genome scale, constraint-based metabolic model iJP962, and the useful conversations about this work; and to Lorenzo Olivi for his guidance with Python and Artemis. We also thank Daniel Moñino-López for discussing some ingenious ideas that have materialized on some of the key points of this paper. This work is part of the research programs SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis and Putida for Plastics, both supported by the Dutch Research Council (NWO) [15814, GSGT.2019.028]. Funding for open access charge: Dutch Research Council.

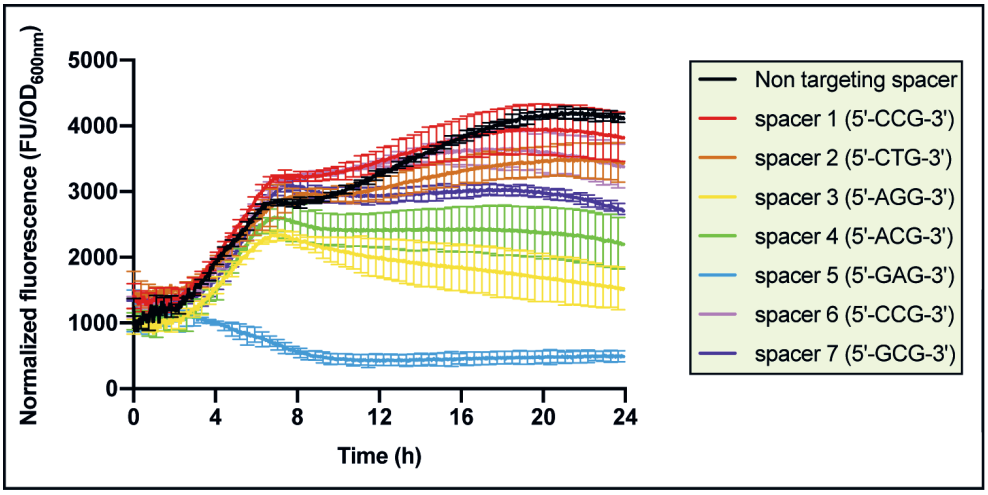
Conflict of interest

The authors declare there are no conflicting interests.

Author contributions

Conceptualization: EAG/MMP/VAPMdS; Data curation: EAG/MMP; Formal analysis: EAG/MMP; Funding acquisition: MMP/RvK/VAPMdS; Investigation: EAG/MMP; Methodology: EAG/MMP; Resources: VAPMdS; Software: EAG/MMP; Supervision: LGM/RvK/VAPMdS; Validation: EAG/MMP; Visualization: EAG/MMP; Writing – original draft: EAG/MMP; Writing – review & editing: EAG/MMP/LGM/RvK/VAPMdS

Supplementary material



Supplementary Figure S2.1: Fluorescent loss essays. Fluorescence curves of *P. putida* KT2440, harboring the pSEVAb44-sfGFP, pSEVAb62-ScCas9 and pSEVAb23-RhaBAD-crRNA_sp with the different spacers. Data represent normalized fluorescence levels, expressed as a ratio with the OD_{600} nm over 24 h under induced conditions (mean ± s.d., n ≥ 3 biological). The non-targeting spacer was used as control. The spacers 1 to 7 targeted the pSEVAb44-sfGFP plasmid. The spacers 1, 2, 3, 4, 5, 6 and 7 were, respectively, in close proximity with the PAMs 5'-CCG-3', 5'-CTG-3', 5'-AGG-3', 5'- ACG-3', 5'-GAG-3', 5'-CCG-3' and 5'-GCG-3'

The rest of the supplementary material of this chapter, including Supplementary Tables S1 to S9, can be accessed *via*:

https://pubs.acs.org/doi/10.1021/acssynbio.1c00297

Bibliography

1. Poblete-Castro, I., Becker, J., Dohnt, K., *et al.* Industrial biotechnology of Pseudomonas putida and related species. *Applied Microbiology and Biotechnology* **93,** 2279–2290 (2012).
2. Nikel, P. I. & de Lorenzo, V. Pseudomonas putida as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism. *Metabolic Engineering. Metabolic Engineering Host Organism Special Issue* **50,** 142–155 (2018).
3. Calero, P. & Nikel, P. I. Chasing bacterial *chassis* for metabolic engineering: a perspective review from classical to non-traditional microorganisms. *Microb. Biotechnol.* **12,** 98–124 (2019).
4. Martin-Pascual, M., Batianis, C., Bruinsma, L., *et al.* A navigation guide of synthetic biology tools for Pseudomonas putida. *Biotechnology Advances* **49,** 107732 (2021).
5. Lau, Y. H., Stirling, F., Kuo, J., *et al.* Large-scale recoding of a bacterial genome by iterative recombineering of synthetic DNA. *Nucleic Acids Research* **45,** 6971–6980 (2017).
6. Chin, J. W. Reprogramming the genetic code. *EMBO J* **30,** 2312–2324 (2011).
7. Nikel, P. I. Synthesis of Recoded Bacterial Genomes toward Bespoke Biocatalysis. *Trends in Biotechnology* **37,** 1036–1038 (2019).
8. Lajoie, M. J., Rovner, A. J., Goodman, D. B., *et al.* Genomically recoded organisms expand biological functions. *science* **342,** 357–360 (2013).
9. Ma, N. J. & Isaacs, F. J. Genomic Recoding Broadly Obstructs the Propagation of Horizontally Transferred Genetic Elements. *Cell Syst* **3,** 199–207 (2016).
10. Rovner, A. J., Haimovich, A. D., Katz, S. R., *et al.* Recoded organisms engineered to depend on synthetic amino acids. *Nature* **518,** 89–93 (2015).
11. Martínez, M. A., Jordan-Paiz, A., Franco, S., *et al.* Synonymous Virus Genome Recoding as a Tool to Impact Viral Fitness. *Trends in Microbiology* **24,** 134–147 (2016).
12. Richardson, S. M., Mitchell, L. A., Stracquadanio, G., *et al.* Design of a synthetic yeast genome. *Science* **355,** 1040–1044 (2017).
13. Wang, K., Fredens, J., Brunner, S. F., *et al.* Defining synonymous codon compression schemes by genome recoding. *Nature* **539,** 59–64 (2016).
14. Fredens, J., Wang, K., de la Torre, D., *et al.* Total synthesis of Escherichia coli with a recoded genome. *Nature* **569,** 514–518 (2019).
15. Kuo, J., Stirling, F., Lau, Y. H., *et al.* Synthetic genome recoding: new genetic codes for new features. *Current Genetics* **64,** 327–333 (2018).
16. Isaacs, F. J., Carr, P. A., Wang, H. H., *et al.* Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement. *Science* **333,** 348–353 (2011).
17. Mandell, D. J., Lajoie, M. J., Mee, M. T., *et al.* Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518,** 55–60 (2015).
18. Ostrov, N., Landon, M., Guell, M., *et al.* Design, synthesis, and testing toward a 57-codon genome. *Science* **353,** 819–822 (2016).
19. Hutchison, C. A., Chuang, R.-Y., Noskov, V. N., *et al.* Design and synthesis of a minimal bacterial genome. *Science* **351** (2016).
20. Murphy, K. C. λ Recombination and Recombineering. *EcoSal Plus* **7** (2016).

21. Ellis, H. M., Yu, D., DiTizio, T., *et al.* High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proceedings of the National Academy of Sciences* **98,** 6742–6746 (2001).

22. Aparicio, T., Jensen, S. I., Nielsen, A. T., *et al.* The Ssr protein (T1E_1405) from *Pseudomonas putida* DOT-T1E enables oligonucleotide-based recombineering in platform strain *P. putida* EM42. *Biotechnology Journal* **11,** 1309–1319 (2016).

23. Ricaurte, D. E., Martínez-García, E., Nyerges, Á., *et al.* A standardized workflow for surveying recombinases expands bacterial genome-editing capabilities. *Microb. Biotechnol.* **11,** 176–188 (2018).

24. Nyerges, Á., Csörgő, B., Nagy, I., *et al.* A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *Proc Natl Acad Sci U S A* **113,** 2502–2507 (2016).

25. Aparicio, T., Nyerges, A., Martínez-García, E., *et al.* High-efficiency multi-site genomic editing of Pseudomonas putida through thermoinducible ssDNA recombineering. *Iscience* **23,** 100946 (2020).

26. Wannier, T. M., Nyerges, A., Kuchwara, H. M., *et al.* Improved bacterial recombineering by parallelized protein discovery. *PNAS* **117,** 13689–13698 (2020).

27. Selle, K. & Barrangou, R. CRISPR-Based Technologies and the Future of Food Science. *J Food Sci* **80,** R2367–2372 (2015).

28. Aparicio, T., Lorenzo, V. d. & Martínez-García, E. CRISPR/Cas9-Based Counterselection Boosts Recombineering Efficiency in Pseudomonas putida. en. *Biotechnology Journal* **13,** 1700161 (2018).

29. Jiang, W., Bikard, D., Cox, D., *et al.* RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature biotechnology* **31,** 233–239 (2013).

30. Chatterjee, P., Jakimo, N. & Jacobson, J. M. Minimal PAM specificity of a highly similar SpCas9 ortholog. *Science Advances* **4,** eaau0766 (2018).

31. Nelson, K. E., Weinel, C., Paulsen, I. T., *et al.* Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440. *Environmental microbiology* **4,** 799–808 (2002).

32. Povolotskaya, I. S., Kondrashov, F. A., Ledda, A., *et al.* Stop codons in bacteria are not selectively equivalent. *Biol Direct* **7,** 30 (2012).

33. Korkmaz, G., Holm, M., Wiens, T., *et al.* Comprehensive Analysis of Stop Codon Usage in Bacteria and Its Correlation with Release Factor Abundance*. *Journal of Biological Chemistry* **289,** 30334–30342 (2014).

34. Poulsen, B. E., Yang, R., Clatworthy, A. E., *et al.* Defining the core essential genome of Pseudomonas aeruginosa. *PNAS* **116,** 10072–10080 (2019).

35. Wetmore, K. M., Price, M. N., Waters, R. J., *et al.* Rapid Quantification of Mutant Fitness in Diverse Bacteria by Sequencing Randomly Bar-Coded Transposons. *mBio* **6** (2015).

36. Belda, E., Van Heck, R. G., José Lopez-Sanchez, M., *et al.* The revisited genome of Pseudomonas putida KT2440 enlightens its value as a robust metabolic chassis. *Environmental microbiology* **18,** 3403–3424 (2016).

37. Puchałka, J., Oberhardt, M. A., Godinho, M., *et al.* Genome-Scale Reconstruction and Analysis of the Pseudomonas putid a KT2440 Metabolic Network Facilitates Applications in Biotechnology. *PLoS computational biology* **4,** e1000210 (2008).

38. Molina-Henares, M. A., de la Torre, J., García-Salamanca, A., *et al.* Identification of conditionally essential genes for growth of Pseudomonas putida KT2440 on minimal medium through the screening of a genome-wide mutant library. *Environ Microbiol* **12,** 1468–1485 (2010).

39. Wang, H. H., Isaacs, F. J., Carr, P. A., *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460,** 894–898 (2009).

40. Gallagher, R. R., Li, Z., Lewis, A. O., *et al.* Rapid editing and evolution of bacterial genomes using libraries of synthetic DNA. *Nat Protoc* **9,** 2301–2316 (2014).

41. Chang, Y., Wang, Q., Su, T., *et al.* The efficiency for recombineering is dependent on the source of the phage recombinase function unit. *bioRxiv,* 745448 (2019).

42. Cui, L. & Bikard, D. Consequences of Cas9 cleavage in the chromosome of Escherichia coli. *Nucleic acids research* **44,** 4243–4251 (2016).

43. Bikard, D., Hatoum-Aslan, A., Mucida, D., *et al.* CRISPR interference can prevent natural transformation and virulence acquisition during in vivo bacterial infection. *Cell Host Microbe* **12,** 177–186 (2012).

44. Komor, A. C., Badran, A. H. & Liu, D. R. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell* **168,** 20–36 (2017).

45. Qi, L. S., Larson, M. H., Gilbert, L. A., *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152,** 1173–1183 (2013).

46. Peters, J. M., Colavin, A., Shi, H., *et al.* A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. *Cell* **165,** 1493–1506 (2016).

47. Aparicio, T., Lorenzo, V. d. & Martínez-García, E. Improved Thermotolerance of Genome-Reduced Pseudomonas putida EM42 Enables Effective Functioning of the PL/cI857 System. *Biotechnology Journal* **14,** 1800483 (2019).

48. Wittwer, C. T. High-resolution DNA melting analysis: advancements and limitations. *Human Mutation* **30,** 857–859 (2009).

49. Song, M., Li, J., Xiong, C., *et al.* Applying high-resolution melting (HRM) technology to identify five commonly used Artemisia species. *Scientific Reports* **6,** 34133 (2016).

50. Abbasi, M. N., Fu, J., Bian, X., *et al.* Recombineering for Genetic Engineering of Natural Product Biosynthetic Pathways. *Trends in Biotechnology. Special Issue: Metabolic Engineering* **38,** 715–728 (2020).

51. Sun, Z., Deng, A., Hu, T., *et al.* A high-efficiency recombineering system with PCR-based ssDNA in Bacillus subtilis mediated by the native phage recombinase GP35. *Appl Microbiol Biotechnol* **99,** 5151–5162 (2015).

52. Binder, S., Siedler, S., Marienhagen, J., *et al.* Recombineering in C. glutamicum combined with optical nanosensors: a general strategy for fast producer strain generation. *Nucleic Acids Res* **41,** 6360–6369 (2013).

53. Guo, T., Xin, Y., Zhang, Y., *et al.* A rapid and versatile tool for genomic engineering in Lactococcus lactis. *Microb Cell Fact* **18,** 22 (2019).

54. Van Pijkeren, J.-P. & Britton, R. A. High efficiency recombineering in lactic acid bacteria. *Nucleic Acids Res* **40,** e76 (2012).

55. Penewit, K., Holmes, E. A., McLean, K., *et al.* Efficient and Scalable Precision Genome Editing in Staphylococcus aureus through Conditional Recombineering and CRISPR/Cas9-Mediated Counterselection. *mBio* **9** (2018).

56. Zhao, D., Feng, X., Zhu, X., *et al.* CRISPR/Cas9-assisted gRNA-free one-step genome editing with no sequence limitations and improved targeting efficiency. *Scientific Reports* **7,** 16624 (2017).

57. Torkelson, J., Harris, R. S., Lombardo, M. J., *et al.* Genome-wide hypermutation in a subpopulation of stationary-phase cells underlies recombination-dependent adaptive mutation. *EMBO J* **16,** 3303–3311 (1997).

58. Creutzburg, S. C. A., Wu, W. Y., Mohanraju, P., *et al.* Good guide, bad guide: spacer sequence-dependent cleavage efficiency of Cas12a. *Nucleic Acids Res* **48,** 3228–3243 (2020).

59. Sobetzko, P., Travers, A. & Muskhelishvili, G. Gene order and chromosome dynamics coordinate spatiotemporal gene expression during the bacterial growth cycle. *PNAS* **109,** 355–356 (2012).

60. Pines, G., Freed, E. F., Winkler, J. D., *et al.* Bacterial Recombineering: Genome Engineering via Phage-Based Homologous Recombination. *ACS Synth. Biol.* **4,** 1176–1185 (2015).

61. Filsinger, G. T., Wannier, T. M., Pedersen, F. B., *et al.* Characterizing the portability of phage-encoded homologous recombination proteins. *Nature Chemical Biology,* 1–9 (2021).

62. Caldwell, B. J., Zakharova, E., Filsinger, G. T., *et al.* Crystal structure of the Red β C-terminal domain in complex with λ Exonuclease reveals an unexpected homology with λ Orf and an interaction with Escherichia coli single stranded DNA binding protein. *Nucleic Acids Research* **47,** 1950–1963 (2019).

63. Yin, J., Zheng, W., Gao, Y., *et al.* Single-Stranded DNA-Binding Protein and Exogenous RecBCD Inhibitors Enhance Phage-Derived Homologous Recombination in Pseudomonas. *iScience* **14,** 1–14 (2019).

64. Azpiroz, M. F. & Laviña, M. Analysis of RecA-independent recombination events between short direct repeats related to a genomic island and to a plasmid in Escherichia coli K12. *PeerJ* **5,** e3293 (2017).

65. Chatterjee, P., Jakimo, N., Lee, J., *et al.* An engineered ScCas9 with broad PAM range and high specificity and activity. *Nat Biotechnol* **38,** 1154–1158 (2020).

66. Wannier, T. M., Ciaccia, P. N., Ellington, A. D., *et al.* Recombineering and MAGE. *Nat Rev Methods Primers* **1,** 1–24 (2021).

67. Gibson, D. G., Benders, G. A., Andrews-Pfannkoch, C., *et al.* Complete Chemical Synthesis, Assembly, and Cloning of a Mycoplasma genitalium Genome. *Science* **319,** 1215–1220 (2008).

68. Gibson, D. G., Glass, J. I., Lartigue, C., *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *science* **329,** 52–56 (2010).

69. Annaluru, N., Muller, H., Mitchell, L. A., *et al.* Total Synthesis of a Functional Designer Eukaryotic Chromosome. *Science* **344,** 55–58 (2014).

70. Garcia-Morales, L., Ruiz, E., Gourgues, G., *et al.* A RAGE Based Strategy for the Genome Engineering of the Human Respiratory Pathogen Mycoplasma pneumoniae. *ACS Synth. Biol.* **9,** 2737–2748 (2020).

71. Calero, P., Volke, D. C., Lowe, P. T., *et al.* A fluoride-responsive genetic circuit enables in vivo biofluorination in engineered Pseudomonas putida. *Nature Communications* **11,** 5045 (2020).

72. Asin-Garcia, E., Kallergi, A., Landeweerd, L., *et al.* Genetic Safeguards for Safety-by-design: So Close Yet So Far. *Trends in Biotechnology* **38,** 1308–1312 (2020).

73. Green, R. & Rogers, E. J. Transformation of Chemically Competent E. coli. **529,** 329–336 (2013).

74. Choi, K.-H., Kumar, A. & Schweizer, H. P. A 10-min method for preparation of highly electrocompetent Pseudomonas aeruginosa cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J Microbiol Methods* **64,** 391–397 (2006).

75. Wirth, N. T., Kozaeva, E. & Nikel, P. I. Accelerated genome engineering of *Pseudomonas putida* by I- *Sce* I—mediated recombination and CRISPR-Cas9 counterselection. *Microb. Biotechnol.* **13,** 233–249 (2020).

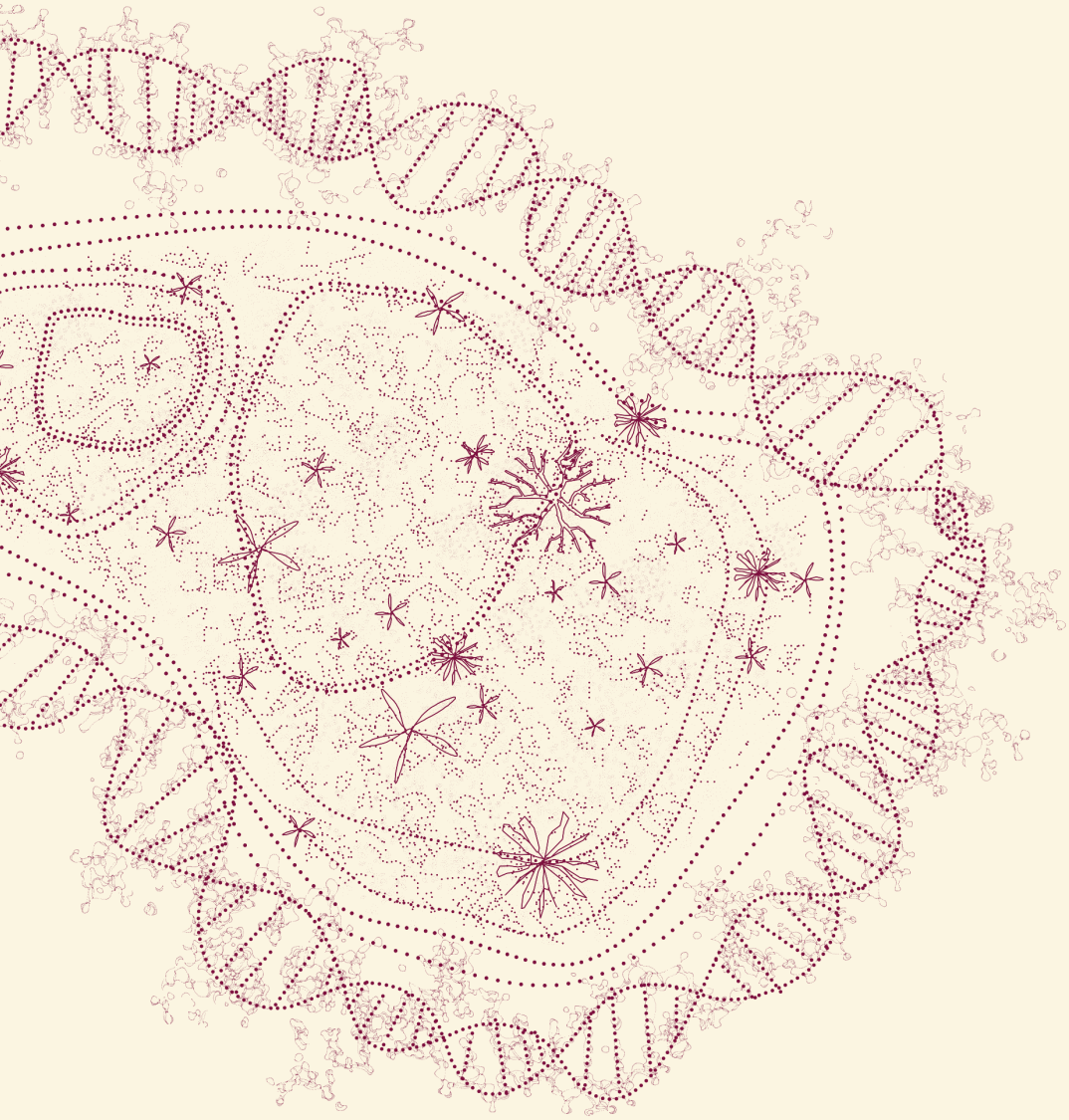
76. Damalas, S. G., Batianis, C., Martin-Pascual, M., *et al.* SEVA 3.1: enabling interoperability of DNA assembly among the SEVA, BioBricks and Type IIS restriction enzyme standards. *Microbial biotechnology* **13,** 1793–1806 (2020).

77. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* **27,** 946–950 (2009).

78. Batianis, C., Kozaeva, E., Damalas, S. G., *et al.* An expanded CRISPRi toolbox for tunable control of gene expression in Pseudomonas putida. *Microbial biotechnology* **13,** 368–385 (2020).

79. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* **31,** 3406–3415 (2003).

80. Aparicio, T., Nyerges, A., Nagy, I., *et al.* Mismatch repair hierarchy of Pseudomonas putida revealed by mutagenic ssDNA recombineering of the pyrF gene. *Environmental Microbiology* **22,** 45–58 (2020).



CHAPTER 3

Metagenomics harvested genus-specific single-stranded DNA annealing proteins improve and expand recombineering in *Pseudomonas* species

Enrique Asin-Garcia, Luis Garcia-Morales, Tessa Bartholet, Zhuobin Liang, Farren J. Isaacs, Vitor A. P. Martins dos Santos

Abstract

The widespread *Pseudomonas* genus embraces a series of relevant species with remarkable abilities to degrade and produce a whole range of compounds. While their unusual characteristics of tolerance and stress resistance already position them as valuable hosts for industrial and environmental biotechnology, easier and more efficient DNA manipulation tools would propel them to the front ranks of the metabolic engineering platforms. To improve their genome editing capabilities, we generated a genus-specific library of potential single-stranded DNA annealing proteins, also known as recombinases, that were later screened in a high-throughput manner using a serial enriching workflow followed by Oxford Nanopore NGS analysis. Among different active candidates with variable levels of allelic replacement frequency, efficient recombinases were found and characterized for mediating recombineering in four different species. New variants yielded higher allelic replacement frequencies than previous standards in *P. putida* and *P. aeruginosa*, and set a precedent for efficient recombineering in *P. taiwanensis* and *P. fluorescens*. These findings will enhance the mutagenesis capabilities of these members of the *Pseudomonas* genus, increasing the possibilities for biotransformation and boosting their potential for synthetic biology applications.

Introduction

In recent years, the number of studies exploiting *Pseudomonas* species as bacterial chassis for synthetic biology applications has increased exponentially [1]. Among the reasons why some of the species of this genus have become significantly relevant to industrial and environmental biotechnology are their metabolic versatility and genetic plasticity [2]. Their distinct biochemical network allows them to efficiently tolerate solvents, counteract high levels of oxidative stress and carry out strong redox reactions [3, 4]. These traits, remarkably valuable for the harsh conditions of industrial bioreactors, fostered a flourishing research interest in *Pseudomonas* genetics that has resulted in an extensive repertoire of molecular tools, studies and biotechnological applications including biological control, bioremediation, biosensing and biotransformations [4].

Amongst the *Pseudomonas*, we encounter several species that stand out as biological chassis due to their distinctive capacities. In this article, we focus on four of them. In the first place, the species *P. putida*, in particular strain KT2440, naturally displays a remarkable capability to degrade toxic compounds [5]. This feature is motivated by a distinct carbon metabolism that produces high levels of reductive power to help counteract oxidative stress [6] and a high tolerance to solvents and antibiotics as a consequence of its numerous export pumps [7]. As a result, *P. putida* KT2440 has gained notoriety as a chassis in synthetic biology for metabolic engineering purposes, especially towards applications involving complex redox reactions that other bacteria could not sustain [4, 8, 9]. Second, *P. taiwanensis* VLB120 has emerged in the recent years as an excellent solvent-tolerant strain with a unique set of resistance pumps and with the ability of growing in the presence of a second phase aliphatic alcohols and hydrophobic aromatics [10–13]. Furthermore, this bacterium is able to use a wide range of organic molecules as carbon source [14], most prominent among them being xylose, what makes it one of the few *Pseudomonas* that is able to utilize this substrate without any genetic modification [15]. Third, the rhizobacterium *P. fluorescens* SBW25 is popular as a biocontrol agent for its production of secondary metabolites that help suppressing soilborne plant pathogens and therefore promoting plant growth [16, 17]. Able as well to use disparate carbon sources, this microbe tolerates wide ranges of temper-

ature and pH [18] in addition to a plethora of other adverse conditions [19]. These properties position *P. fluorescens* as a popular workhorse for bioremediation, next to its pathogen control and biofertilizer applications, and make it a great candidate to research metabolic reconfigurations in response to abiotic stress [20, 21]. Lastly, we find *P. aeruginosa*, an opportunistic pathogen ubiquitous in the environment, that flourishes in individuals with weakened immune systems causing devastating effects due to its ability to rapidly acquire antibiotic resistances [22]. Nevertheless, this bacterium, especially the common strain PA01, has proved extremely useful in medicine, industries, and environment, due to its ability to produce secondary metabolites with antimicrobial properties and other compounds such as alginate, rhamnolipids or biosurfactants [23–25]. *P. aeruginosa* has therefore beneficial uses not only in medicine but in various other sectors that range from waste degradation and agriculture to textile and cleaning products [26].

To a greater or lesser extent, the role that these *Pseudomonas* strains play in the biotechnology field is beyond question and their potential is only limited by our ability to genetically engineer them and make use of their special traits. While improving intrinsic lifestyle characteristics such as tolerance, handling, growth and substrate consumption rates remains challenging, the incredible progress of genetic engineering during the last decades ensures fast and efficient tailoring of microbial metabolism for a large array of bioconversions [27, 28]. Currently, the level of development of synthetic biology tools varies amongst the aforementioned species. Ideally though, a standardized and orthogonal tool set that could be easily applicable and transferable for engineering different *Pseudomonas* would propel their value as promising hosts for the bio-based industry [29, 30].

Thanks to the recent advances in gene editing technologies, precise manipulation of genomes is becoming a reality for more and more organisms. Challenges of the field such as genome-scale reverse genetics or the generation of combinatorial genetic diversity at base-pair resolution could be addressed *via* recombineering (recombination-mediated genetic engineering) [31]. This efficient bioengineering technique harnesses phage-derived proteins for *in vivo* scarless manipulation of genomes thanks to the integration of single-stranded DNA (ssDNA) [32] or double stranded DNA (dsDNA) [33] (through an ssDNA intermediate) [34, 35] into a repli-

cating chromosome. Recombinase proteins, also known as single-stranded DNA-annealing proteins (SSAPs), promote the incorporation of the ssDNA segments during DNA replication making use of their short homology arms. In this way, such DNA segments bearing designed modifications can be inherited and spread across the cell population allowing the generation of genomic deletions, insertions and single nucleotide substitutions [36].

In *Escherichia coli*, the redβ protein of the λ Red system has long been the standard recombinase for oligo-mediated recombineering [32]. Cycles of recombinase expression during subsequent rounds of DNA segregation can induce multiple random or specific mutations on discrete sites [37, 38]. Thus, improvement of the recombineering process utilising phage-derived recombinases fostered genome-scale editing in *E. coli* and set the basis for tools like Multiplex Automated Genome Engineering (MAGE) [37], Directed Evolution with Random Genomic mutations (DIvERGE) [39] or retron-recombineering [40, 41]. The huge impact of these recombineering-based technologies in *E. coli* and other bacteria has been continuously demonstrated in a growing number of projects including single gene evolution [42], metabolism rewiring *via* pathway diversification [37] or whole-genome recoding [43, 44].

Unfortunately, the recombineering power of redβ is not always readily imparted to other bacterial species including the *Pseudomonas*, which implies specificity of the bacteria-phage interaction [45]. This apparent host tropism results from the distinct recognition of the bacterial single stranded DNA-binding protein (SSB) at the carboxy-terminal site by the recombinase [46]. Attempts of redβ-analogue discovery in *Pseudomonas* species have taken place during the last years for *P. putida* [47, 48] and *P. aeruginosa* [29, 49] with promising results. Yet, the efficiency of the best candidates has not always been consistent across species or has dropped dramatically when multiplexing resulting in recombineering levels far from those obtained in *E. coli* with redβ [50]. While the *P. putida*-borne Rec2 brought about the High-Efficiency Multi-site Genomic Editing system (HEMSE) in its native species [50], it quickly became overshadowed by PapRecT in *P. aeruginosa* [49]. Conversely, this recombinase that originates from a *P. aeruginosa* phage and was identified by Serial Enrichment for Efficient Recombineering (SEER) showed improved recombineering

frequency in *P. aeruginosa* [49] but equal or lower activity than Rec2 in *P. putida* [51].

In order to find a suitable recombineering toolkit, in this study we seek to expand the bioprospecting of SSAPs and the genetic tractability of biotechnologically relevant *Pseudomonas* species. Through computational identification of SSAP homologues, we ended up with a library of the most promising genus-related recombinases. We used DNA synthesis to obtain the SSAP library and express it into the aforementioned strains. Successive recombineering cycles introducing selectable mutations were used to enrich the population containing the most efficient SSAPs. This allowed us to efficiently scan for the best performers in a high-throughput manner in multiple hosts using the long read sequencing devices of Oxford Nanopore Technologies [52, 53]. Putatively active recombinases were later characterized individually and, ultimately, multiple active variants yielded significant results in the four tested *Pseudomonas*. These findings not only demonstrate the value of the utilized workflow for protein identification, but also expand the recombineering toolboxes of the most popular *Pseudomonas* species and usher in a new era of genome editing for those ones lacking any previous functional tool and for the *Pseudomonas* genus as a whole.

Results

Metagenomic discovery of genus specific SSAPs for high-throughput screening in *Pseudomonas* species

We first performed a comprehensive search of potential SSAP functional homologues in the metagenomes of *Pseudomonas* relevant bacteria species and bacteriophages. To maximize our chance to uncover active SSAPs from metagenomic databases, we manually curated a list of seven SSAPs (Supplementary Table S1) from three diverse SSAP superfamilies (RecT, Erf and Sak) with relativity high activity in prior studies to build the initial hidden Markov model (HMM) for functional homologue searching (PMID: 20194117, 18230724, 20451472). After 15 rounds of jackHMMER searches, we identified more than 10000 hits, among which 287 come from diverse species in the *Pseudomonas* genus and 18 come from their bacteriophages

(Supplementary Table S2). Protein sequence clustering of these 305 *Pseudomonas* hits resulted in 48 separated groups (Supplementary Table S3) from which we surveyed a total of 49 *Pseudomonas* SSAP candidates (Supplementary Table S4) for gene synthesis and subsequent testing in four *Pseudomonas* species (P. putida, *P. taiwanensis*, *P. fluorescens* and *P. aeruginosa*).

The predicted 49 SSAP candidates (Supplementary Table S5) were synthesized and cloned into pSEVA2514 vectors upstream a *mutLE36KPP* gene and under the control of the cI857/PL inducible expression system [50]. In addition, Rec2 [48], PapRecT [49], and the λ Red recombineering system [54] were generated in the same

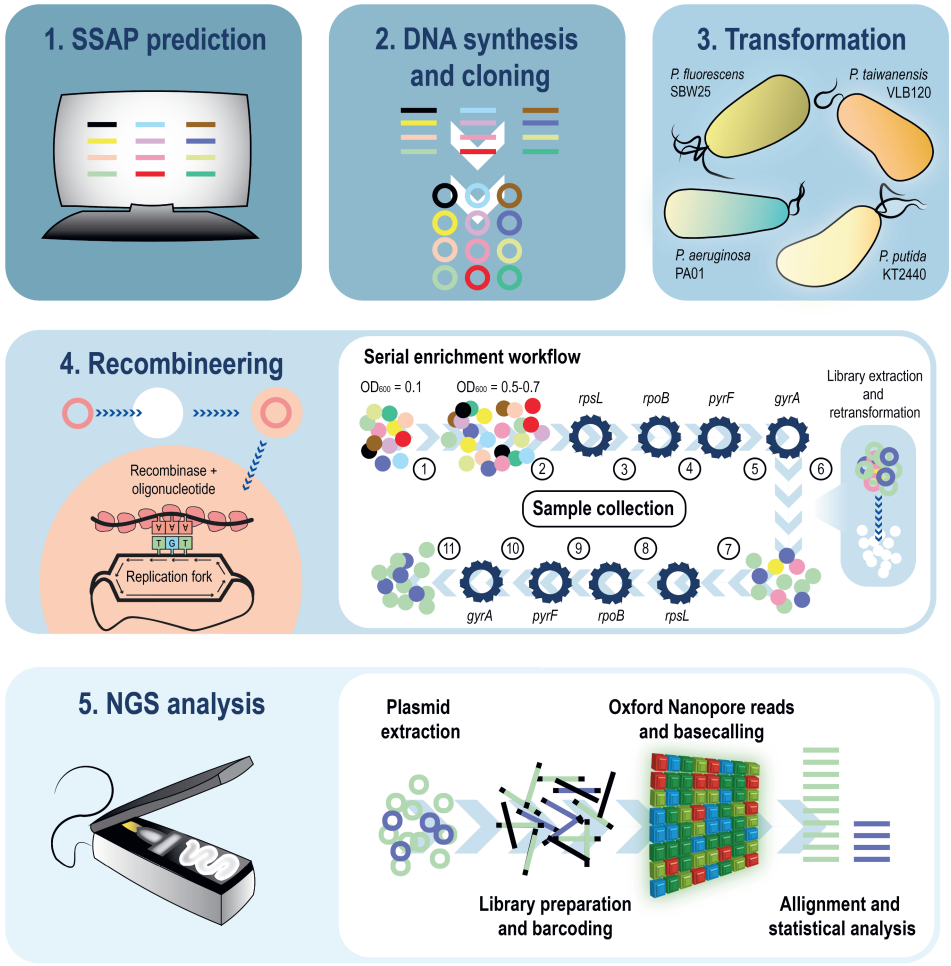


Figure 3.1: Step-by-step workflow to find active recombinases in Pseudomonas species. (1) Bioinformatics-metagenomic approach for SSAP prediction. (2) DNA synthesis and cloning of all the predicted SSAP candidates into pSEVA2514-mutLE36KPP vectors via Golden Gate assembly. (3) Plasmid transformation of all the SSAP candidates into the different Pseudomonas species included in this study (P. putida KT2440, P. taiwanensis VLB120, P. fluorescens SBW25 and P. aeruginosa PAO1). (4) Serial enrichment recombineering workflow. All the strains from the same species containing the different candidates are combined in an equimolar pool of cells and grown together until reaching mid-log phase. At that point, each library of strains is exposed to four iterative recombineering cycles that mediate four specific mutations that confer different antibiotic resistances. By selecting with the appropriate antibiotics, only those cells in which the mutation has been adequately introduced will survive the pressure, enriching the pool in those cells containing active recombinases. After the fourth mutation, all the plasmid content is extracted from the cell population for transformation of a new naïve parental strain for a second round of enrichment using the same selective mutations. Eventually, and after 8 recombineering iterations, the final population contains only those cells carrying the most efficient recombinases. Samples were collected at the 11 different time points indicated for NGS analysis. (5) After extracting the plasmid content, each of the collected samples was barcoded accordingly. Recombinase abundances were determined by Nanopore sequencing after basecalling and aligning the long reads to the sequence of the candidate recombinase genes.

fashion as controls. Naïve libraries of *P. putida* KT2440, *P. taiwanensis* VLB120 and *P. fluorescens* SBW25 were then constructed to assess the activity of the recombinases in the different hosts *via* enriching cycles of recombineering [49]. For this, we made use of compatible oligonucleotides that introduce small mutations in the different genomes conferring specific antibiotic resistances. In this way, four readouts were employed: streptomycin resistance caused by a mutation in the *rpsL* gene; rifampicin resistance conferred by mutating the *rpoB* gene; 5-fluoroorotic acid (5-FOA) resistance due to the introduction of a stop codon in the *pyrF* gene; and nalidixic acid or ciprofloxacin resistance granted after mutating the *gyrA* gene [50]. After 4 consecutive iterations of recombineering introducing the aforementioned mutations and their corresponding selections, plasmids were extracted from the surviving cell populations and reintroduced in naïve parental strains for a second 4-readout round. Samples were collected prior the serial enrichment process and after each of the recombineering cycles for plasmid extraction. Barcoding and library preparation using these samples was performed using the Rapid Barcoding Kit (Oxford Nanopore Technologies), which does not require amplification of the target sequence. Nanopore sequencing was employed to track the presence and evolution of the different recombinase candidates during the whole workflow (Fig-

ure 3.1).

Application of our workflow in *P. putida* KT2440 revealed the presence of four active recombinases: R8, R12, R35 and R47 (Figure 3.2 A, 3.2 B and 3.2 C). Their activity surpassed the one of the positive control Rec2, especially after the second recombineering cycle (barcode 4) and, while their abundance varied during the course of the enrichment, these four clearly stood out over the other candidates which did not report any sign of recombineering activity. Furthermore, the positive control PapRecT was the most abundant recombinase of the whole library presenting an absolute number of reads significantly higher than that of the other positive control Rec2. However, the change in percentage abundance throughout the cycles did not differ much from other recombinases (Supplementary Figure S3.1). Taking into account the percentages of the cell population containing PapRecT compared to those of previous cycles, relative growth of PapRecT was quite subtle after the first introduction and the subsequent reintroduction of the enriched library into the naïve strain KT2440 (barcodes 3 and 8). In addition, analysis of barcodes 1 and 2, previous to the recombineering selection pressure showed no sign of strains containing R4, R5, R11 and R46 in the initial naïve library after the pooling of samples meaning these SSAPs were finally not assessed in *P. putida* (Figure 3.2 A).

In *P. taiwanensis* VLB120, two active recombinases were uncovered after the library enrichment: R9 and R47 (Figure 3.2 D, 3.2 E and 3.2 F). R47 was the most abundant recombinase after every cycle, while R9 was mostly represented after the mutations on the *rpoB* gene (barcodes 4 and 9), likely indicating that this recombinase is able to mediate that specific mutation more efficiently. Rec2 and PapRecT completely disappeared from the assay after the first antibiotic pressure (barcode 3) suggesting that these two recombinases might be inefficient in this *Pseudomonas* species, whereas λ Red showed low levels of activity. The current absence of a recombineering methodology in *P. taiwanensis* explains the lack of positive controls and the role of Rec2 and PapRecT as simple SSAP candidates during the assay in this species. Analysis of the first two barcodes revealed that the strain containing R46 was again not present during the enrichment process (Figure 3.2 D).

In the third place, five active recombinases were identified in *P. fluorescens* SBW25. R9 was the most represented candidate during all the cycles, but R26, R27,

R35 and R47 reported activity as well with variable numbers of reads throughout the enrichment process (Figure 3.2 G, 3.2 H and 3.2 I). Once again, the controls Rec2 and λ Red showed very low levels of activity, while PapRecT was not even detected before the selection pressure. Unfortunately, a large amount of SSAP candidates were not detected in barcodes 1 and 2, with 21 strains missing the assessment in *P. fluorescens* (R1, R3, R10, R12, R14, R17, R18, R20, R21, R24, R25, R28, R30, R32, R36, R37, R41, R44, R46, R49 and PapRecT) (Figure 3.2 G). This species' ability to readily acquire resistance to antibiotics might have caused the strains to lose their recombinase plasmids as they became resistant to kanamycin themselves and the presence of their plasmids obsolete for survival [55].

New recombinases prove to efficiently mediate recombineering

Once that the enrichment workflow yielded active recombinases in the different species, these were assessed individually to obtain more accurate data of their efficiency. Since editing efficiency is typically dependent on the target *locus* [50, 51], we used two different readouts per strain to test the frequencies of mutant appearance. For an easy screening, we made use of oligonucleotides that confer different antibiotic resistances through small mutations as we had previously done during the serial enrichment.

In *P. putida*, the results from the controls Rec2, PapRecT and λ Red were in line with those obtained in previous research [48, 51]. While the number of reads of PapRecT in the pooled population of the enrichment workflow was significantly higher than those of Rec2 or of any of the candidates, its individual assessment (0.17 ± 0.05 % for *rpsL* and 0.26 ± 0.01 % for *pyrF*) resulted in recombineering efficiencies in the same range as those obtained with Rec2 (0.31 ± 0.21 % for *rpsL* and 0.20 ± 0.04 % for *pyrF*). On the other hand, λ Red showed low recombination efficiency (0.00 ± 0.00 % for both *rpsL* and *pyrF*), which is in agreement with the results of the workflow as well as with previous literature. Regarding the selected SSAP candidates, R8 yielded allelic replacement frequencies (ARF) of 0.33 ± 0.10 % for *rpsL* and 0.92 ± 0.04 % for *pyrF*. R12 scored 0.48 ± 0.05 % for *pyrF* and an inconsistent 0.69 % with a standard deviation of 0.64 for *rpsL*. R35 showed the lowest efficiencies of the individual assessment with 0.21 ± 0.07 % for *rpsL* and 0.11 ± 0.03 % for *pyrF*. Lastly,

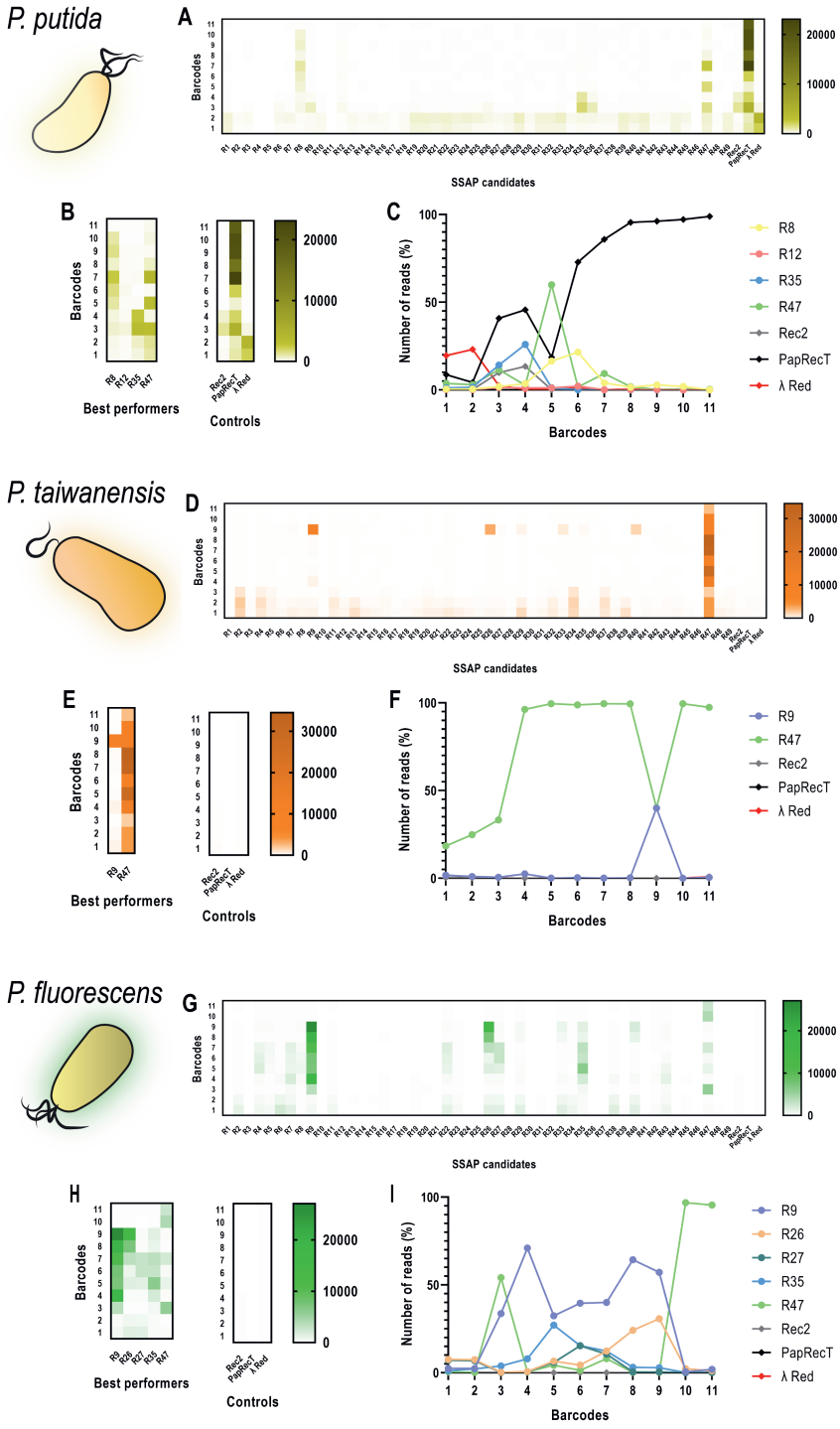

P. putida
A
Barcodes
SSAP candidates
B
Best performers
Controls
C
Number of reads (%)
Barcodes
R8
R12
R35
R47
Rec2
PapRecT
λ Red
P. taiwanensis
D
E
F
R9
R47
P. fluorescens
G
H
I
R26
R27

Figure 3.2: Abundance of NGS reads per recombinase over the selective enrichment cycles in P. putida KT2440 (yellow), P. taiwanensis VLB120 (orange) and P. fluorescens SBW25 (green). (A, D and G) Heatmaps depicting abundance of all 49 SSAP candidates plus controls for each of the 11 barcoded samples in P. putida, P. taiwanensis and P. fluorescens, respectively. (Mean ± s.d., n = 2 technical) The number of reads associated to the abundance within the cell population is represented in a colour scale mapping according to the depicted corresponding legends. (B, E and H) Heatmaps zooming into the abundance of the best performers and the three controls per species (P. putida, P. taiwanensis and P. fluorescens, respectively). (C, F and I) Percentage of reads per best performing recombinase (circles) and per control (diamonds) over the enrichment cycles in P. putida, P. taiwanensis and P. fluorescens, respectively.

R47 reported the highest levels for *pyrF* with a 1.22 ± 0.07 %, and an ARF of 0.30 ± 0.12 % for *rpsL* (Figure 3.3 A). While the individual performance of these four candidates was found within the same range than the ones of Rec2 and PapRecT for the *rpsL* mutation (not significative differences), in the case of *pyrF*, the four of them performed significantly better than both positive controls.

Rec2 and PapRecT showed the best individual results in *P. taiwanensis* despite their scarce presence in the enrichment workflow of this bacterium (2.15 ± 0.90 % and 1.63 ± 0.62 % for *rpsL*, and 0.22 ± 0.06 % and 0.48 ± 0.04 % for *rpoB*, respectively) (Figure 3.3 B). By the contrary, R47, which was the most represented candidate of the pooled library over the enrichment cycles, yielded lower efficiencies (0.30 ± 0.09 % for *rpsL* and 0.16 ± 0.02 % for *rpoB*). However, significative differences were only observed between the results of R47 and PapRecT for the *rpoB* mutation, which can most likely be explained due to the variance of the Rec2 and PapRecT replicates. Lastly, R9 showed very low levels of activity (0.11 ± 0.12 % for *rpsL* and 0.01 ± 0.004 % for *rpoB*) while the λ Red system did not seem to mediate any recombineering in this *Pseudomonas* species either.

The individual assessment of SSAPs in *P. fluorescens* resulted in R9 as the best recombinase with an ARF of 0.43 ± 0.17 % for *rpsL*, which was significantly higher than any other recombinase, and an ARF of 0.05 ± 0.02 % for *pyrF* (Figure 3.3 C). R26, R27 and R35 yielded similarly low results for the *pyrF* mutation (0.02 ± 0.02 %, 0.05 ± 0.02 % and 0.09 ± 0.02 %, respectively). However, of this three, only R35 could efficiently mediate the *rpsL* mutation (0.05 ± 0.04 %) while R26 and R27 showed insignificant efficiency levels lower than 0.002 %. This was also the case for R47, PapRecT and the λ Red system, which did not score efficiencies higher than 0.002

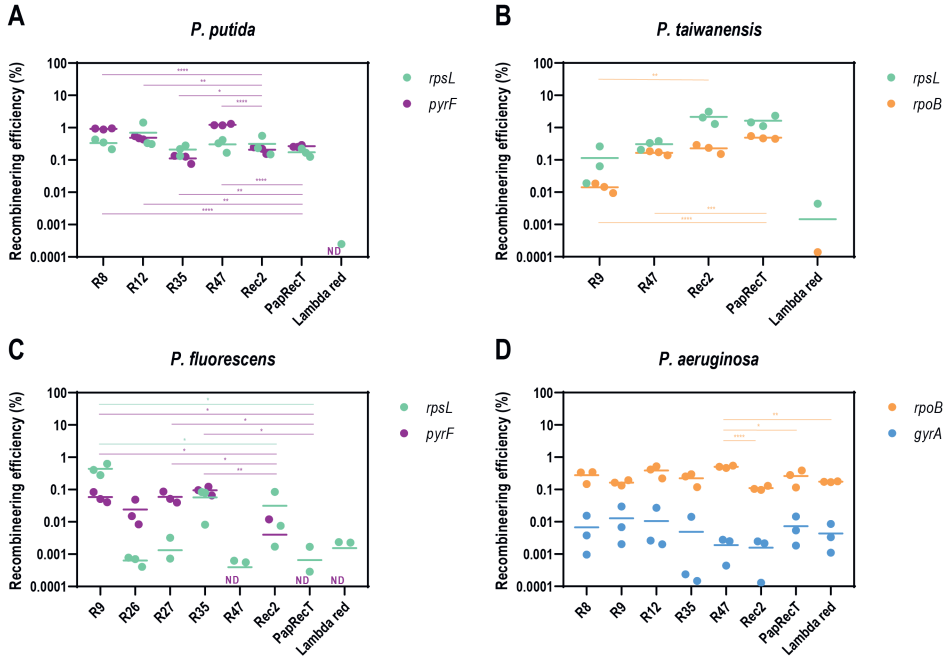


*Figure 3.3: Editing efficiencies of selected SSAPs in individual strains in percentage. Graphs display recombineering efficiency mediating: (A) rpsL K43T (green) and pyrF E50stop (purple) mutations in P. putida KT2440, (B) rpsL K43T (green) and rpoB Q518L (orange) mutations in P. taiwanensis VLB120, (C) rpsL K43T (green) and pyrF G50stop (purple) mutations in P. fluorescens SBW25, and (D) rpoB D521V (orange) and gyrA T83I (blue) mutations in P. aeruginosa PAO1. Recombineering efficiency was calculated as the ratio between antibiotic resistant CFU growing in pressure plates of LB with the corresponding antibiotic and total CFU growing in LB plates. ND labels indicate that no mutants were detected under those specific circumstances of mutation and recombinase. Non-induced recombinase samples were included as controls and were subtracted from the absolute values. Only significant values are indicated for a parametric two-tailed t test between two groups, where *P ≤ 0.05; **P < 0.01; ***P < 0.001; and ****P < 0.0001; non-significant values were not depicted. (Mean ± s.d., n = 3 biological)*

% for neither of the mutations. Rec2 did nonetheless show some low activity levels with efficiencies of 0.03 ± 0.04 % for *rpsL* and 0.004 ± 0.006 % for *pyrF*.

Finally, we evaluated the most efficient recombinases from the three previous *Pseudomonas* in the related species *P. aeruginosa* PAO1 (Figure 3.3 D). Given the natural resistance of this species to Kanamycin [56], selected recombinases were cloned into pSEVA6514 plasmids that confer resistance to gentamicin. For the test, we also utilized two different antibiotic resistance readouts, the *rpoB* D521V and the *gyrA* T83I mutations, which had been previously characterized in this bacterium

[49]. In general, all our candidates, namely R8, R9, R12, R35, R47, Rec2, PapRecT and the λ Red system, showed recombineering activity in *P. aeruginosa*, with higher efficiency levels for the *rpoB* than for the *gyrA* mutation, which were all equal or lower than 0.01 %. Our *rpoB* results of PapRecT and Rec2 (0.26 ± 0.13 % and 0.11 ± 0.01 %, respectively) were in line with those obtained in previous research [49] in the sense than Rec2 seems to have some activity in this species but it is less suitable for editing in *P. aeruginosa* than the efficient PapRecT. Contrarily, when we tested the λ Red system, we obtained a relatively modest efficiency of 0.17 ± 0.005 %, which was higher than the reported by other previous studies [49, 57, 58]. Furthermore, all R8, R9, R12 and R35 yielded efficiencies in the same range than PapRecT (0.27 ± 0.11 %, 0.16 ± 0.03 %, 0.38 ± 0.15 % and 0.22 ± 0.09 %), while R47 significantly surpassed all the other candidates and controls with a notable efficiency of 0.50 ± 0.04 %.

Allelic replacement rate differs across recombinases

Motivated by the differences between the behaviour of the recombinases within the pooled library and in the individual tests, we performed an assay to determine if other intrinsic characteristics beyond the ARF could have influenced the results of the enrichment workflow. Our hypothesis was that recombinases could have different action rates and that faster candidates would take over naïve cell populations that were quickly subjected to the pressure thanks to a faster incorporation of the resistance mutation. During our enrichment workflow, we applied the selection pressure to our samples *via* antibiotic addition around 4 h after the end of the recombineering cycle (when OD_{600} reached 0.3-0.5). At that moment, the naïve cell population likely contained more resistant cells mutated with the faster recombinases causing an increase on the percentage of those cells, which would not be derived from a higher efficiency but from a higher speed. To test our hypothesis, we compared the recombineering efficiency of PapRecT and R12 in *P. putida* over time using the readout of the *rpsL* K43T mutation. Previously, PapRecT had shown a significantly higher number of reads than R12 during the enrichment workflow (98.96 % *vs.* 2.24 %, respectively, at their highest moment). However, the latter surpassed the efficiency of the former when tested separately.

In this experiment, we subjected to pressure two different cell populations con-

taining the aforementioned recombinases at different times after the recombineering cycle. R12 yielded again a higher final average efficiency than PapRecT (1.68 ± 0.50 % *vs.* 1.53 ± 0.24 %) even though their differences were not significant after 24 h (Figure 3.4). Nevertheless, it was at the early stages of the assay when we could observe some differences. Interestingly, PapRecT reached a 71.79 % of its final efficiency (1.10 ± 0.30 %) only 1 h after the end of the recombineering cycle, while R12 needed more than 7 h to reach that percentage of its own final levels. Protocols of cyclic recombineering usually recommend an incubation of at least 3 h before proceeding with a consecutive cycle [59] or reaching an OD_{600} of mid-log phase [50, 59] to ensure that the introduced mutation gets spread across the population. In the case of PapRecT, only 1 h sufficed to almost achieve the target population of mutants, shedding light on the overwhelming cell population percentages containing this recombinase during the enrichment process (Figure 3.2 A, 3.2 B and 3.2 C).

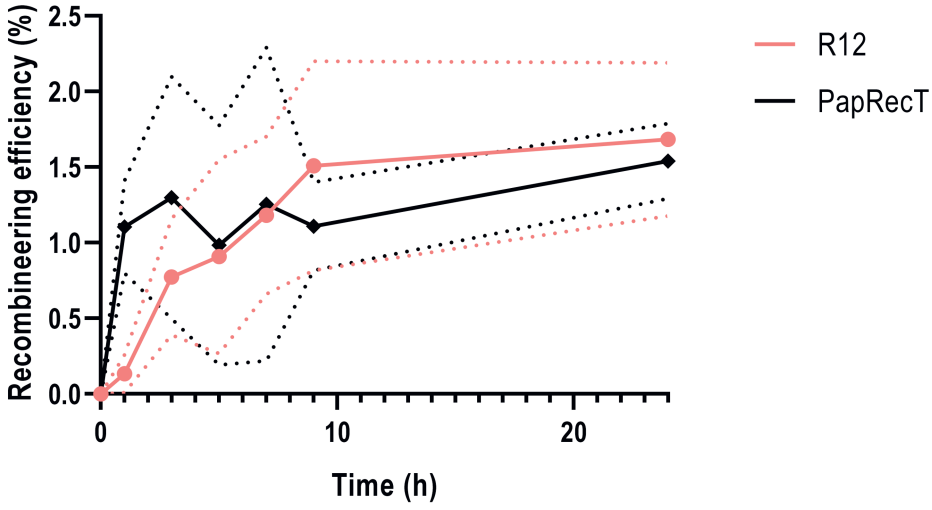


Figure 3.4: Comparative of the editing efficiencies of R12 and PapRecT when subjected to pressure at different time points after recombineering during 24 h. Full lines represent average values while dashed lines represent the standard deviation field. Graph displays recombineering efficiency mediating the rpsL K43T mutation in P. putida KT2440 which was calculated as the ratio between streptomycin resistant CFU growing in LB-sm and total CFU growing in LB plates. Samples before the recombineering cycle were included as controls and were subtracted from the absolute values. (Mean± s.d., n = 3 biological).

Combining several recombinases does not result in a synergetic effect

In an attempt to enhance recombineering efficiencies even further, we combined five of our best recombinases (R9, R12, R47, Rec2 and PapRecT) on a single chimeric plasmid. Previous studies had demonstrated that increasing the gene expression of the recombinase by using strong ribosomal binding sites (RBS) could result in a significant improvement of editing efficiency [49]. Along this line, we hypothesized that multiple SSAP variants expressed at the same time could favour allelic replacement as a consequence of their synergetic effect. To assess the efficiency of the chimeric plasmid in the different *Pseudomonas*, we mutated again the *rpsL* (in *P. putida*, *P. taiwanensis* and *P. fluorescens*) and the *rpoB* (in *P. aeruginosa*) genes as readouts to test the frequency of mutant appearance.

The results after applying one recombineering cycle are depicted in Figure 3.5. The percentage of mutated cells was 0.48 ± 0.30 % in *P. putida*, 1.18 ± 0.77 % in *P. taiwanensis*, 0.02 ± 0.006 % in *P. fluorescens*, and of 0.29 ± 0.17 % in *P. aeruginosa*. In none of the cases, the efficiency achieved by the chimeric plasmid reached the average obtained with the best species-specific recombinase individually. These differences, however, were not significative suggesting that the best candidates (R12 for *P. putida*, Rec2 for *P. taiwanensis* and R47 for *P. aeruginosa*) still performed equally well as part of the recombinase cassette, except in the case of *P. fluorescens*, where R9 achieved a 0.43 ± 0.17 % efficiency for *rpsL* when expressed individually but only a 0.02 ± 0.006 % when expressed in the chimeric plasmid.

Off-target effect of active recombinases

To obtain an even more comprehensive view of our recombinases, we examined the accumulation of off-target mutations using whole genome sequencing data from different strains that had gone through one successful recombineering cycle. We studied the off-target effect of the most efficient variants (R8, R9, R12, R35 and R47), and, additionally, we studied that of R47 in the four different *Pseudomonas* hosts (Table 3.1). *P. putida* cells expressing *r8* and *r12* showed an elevated number of off-targets mutations, whereas *P. putida* cells expressing *r35* and *P. fluorescens* cells expressing *r9* resulted in a lower off-target effect. When looking at the results of cells expressing *r47*, we observed inconsistencies across species. *P. putida* and

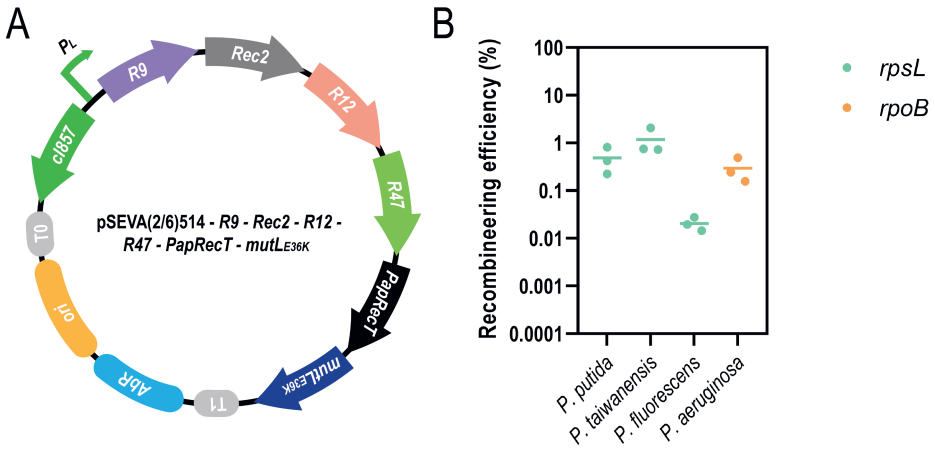


Figure 3.5: The chimeric pSEVA(2/6)514-R9-Rec2-R12-R47-PapRecT-mutLE36KPP plasmid, harbouring the thermolabile cI857 repressor (dark green) controlling a synthetic operon of recombinase genes. Antibiotic resistance gene (blue) varied between kanamycin in pSEVA2514, which was used for P. putida, P. taiwanensis and P. fluorescens; and gentamicin in pSEVA6514, which was used for P. aeruginosa. (B) Editing efficiencies of the chimeric pSEVA(2/6)514-R9-Rec2-R12-R47-PapRecT-mutLE36KPP plasmid in individual strains in percentage. Graphs display recombineering efficiency mediating: rpsL K43T mutation (green) in P. putida KT2440, rpsL K43T mutation (green) in P. taiwanensis VLB120, rpsL K43T mutation (green) in P. fluorescens SBW25, and rpoB D521V (orange) mutation in P. aeruginosa PAO1. Recombineering efficiency was calculated as the ratio between antibiotic resistant CFU growing in pressure plates of LB with the corresponding antibiotic and total CFU growing in LB plates. (Mean± s.d., n = 3 biological).

P. taiwanensis cells expressing this variant contained high numbers of mutations, which were within the same range than those shown by *r8*- and *r12*-expressing *P. putida* cells. However, *P. fluorescens* and *P. aeruginosa* cells carrying the *r47* gene exhibited lower numbers, this time comparable to those obtained for R9 and R35 in *P. fluorescens* and *P. putida*, respectively.

Discussion

In this research, we screened a library of recombinase candidates for the genus *Pseudomonas* in four representative species utilizing a user-friendly, streamlined and cost-effective high-throughput enrichment workflow. The key element of our method was the application of Oxford Nanopore sequencing with the minION Mk1B and the Rapid Barcoding Sequencing kit by ONT. As a result, we identified a num-

Table 3.1: Off-target mutations per colony after one recombineering cycle (rpsL gene mutation in P. putida, P. taiwanensis and P. fluorescens, and rpoB gene mutation in P. aeruginosa).

Plasmid	Species	N	Off-target mutations
pSEVA2514-R8-mutLE36KPP	*P. putida*	2	14 ± 5.66
pSEVA2514-R12-mutLE36KPP	*P. putida*	2	11.5 ± 3.54
pSEVA2514-R35-mutLE36KPP	*P. putida*	2	2 ± 1.41
pSEVA2514-R47-mutLE36KPP	*P. putida*	2	13.5 ± 2.12
pSEVA2514-R47-mutLE36KPP	*P. taiwanensis*	2	15.5 ± 7.78
pSEVA2514-R9-mutLE36KPP	*P. fluorescens*	2	1 ± 0.00
pSEVA2514-R47-mutLE36KPP	*P. fluorescens*	2	1 ± 0.00
pSEVA6514-R47-mutLE36KPP	*P. aeruginosa*	2	3.5 ± 3.54

ber of active recombinases for the studied species which were later characterized individually. Four active recombinases, R8, R12, R35 and R47, were uncovered for *P. putida* KT2440. While their efficiency was, in general, in the same range than previous standards Rec2 and PapRecT, R8, R12 and specially R47 surpassed the controls' levels when targeting specific *loci*. In *P. taiwanensis* VLB120, we identified two new active recombinases, R9 and R47, and confirmed that *P. putida*'s controls Rec2 and PapRecT have recombineering activity in this species too. The latter two showed outstanding efficiencies in *P. taiwanensis*, almost one order of magnitude higher than those obtained in *P. putida* postulating them as preferred for recombineering in this species. Six recombinases showed some activity in *P. fluorescens* SBW25, R9, R26, R27, R35, R47 and Rec2, but only R9 yielded significant efficiency levels when tested individually. While previous *P. putida* standards did not work in *P. fluorescens* as efficiently as they did in *P. taiwanensis*, newly characterized R9 represents a promising option for kicking off recombineering applications in this species. Lastly, a set of the most active candidates (R8, R9, R12, R35, R47, Rec2, PapRecT and λ Red) was tested in *P. aeruginosa* PAO1. All of them showed recombineering activity in this species, but R47 was highlighted as the best performer, even surpassing the paradigm PapRecT.

The strategy that we followed to find efficient SSAP variants encompassed five stages: a bioinformatic-metagenomic approach for candidate prediction; DNA sequencing and cloning of the most relevant candidates into a library of recombinase vectors; transformation into the *Pseudomonas* systems; a recombineering workflow

to enrich the best candidates within the library; and a final NGS analysis (Figure 3.1).

To ensure that we take into account the natural diversity of SSAPs in candidate prediction, we carefully selected seven representative SSAPs based on their unique occupying positions in the three SSAP superfamilies and their relatively well-characterized recombineering activity. Due to the cost limitation of gene synthesis and capability of simultaneous testing of many candidates, we devised a set of simple yet efficient filtering steps to help narrowing down the initial library to a modest-size collection. The fact that we successfully harvested half a dozen new recombinases with significant recombineering activity in four different *Pseudomonas* species from a relatively small number of candidates strongly indicates that our choice of this set of broad-range "founder" sequences coupling with subsequent genus-specific filtering is efficient for enriching functional relevant SSAP homologues. This implies that such metagenomic bioinformatics workflow may be applicable for predicting active recombinases in other bacterial genus.

Our choice for the NGS analysis landed upon the minion Mk1B by Oxford Nanopore Technologies, as using this platform was more time- and cost-efficient for our purpose and allowed us to overcome some technical challenges that could arise from using other platforms as Illumina, which could derive on biased results [60] given our primary setup. We used full recombinase sequences instead of short, standardized barcodes as the basis for sequencing libraries. This posed an issue due to their different lengths since time-consuming qPCR would have been necessary to normalize the sequences amplification and these results tend to be biased towards easier amplifications (usually shorter amplifications). Shorter products are amplified more efficiently than longer ones in the sequencing process in Illumina devices [61], which could have ultimately yielded the aforementioned bias towards the shorter recombinase sequences in this research. Furthermore, conventional library preparation for Illumina sequencing encompasses several steps (DNA fragmentation, end-repair, adapter ligation and PCR amplification) [62]. As a result, library preparation time for Illumina is generally longer, and even though it depends on the kits used, it ranges from 3 (with kits making use of tagmentation) to 11 h. In contrast, Nanopore sequencing library preparation using the transposome-based Rapid Barcoding kit by ONT consists only of a two-step process that includes

tagmentation/barcoding and concentration of the library, taking around 30 min per library [63]. The resulting library is hence prepared from the actual plasmid molecules without the need of amplifying the target region. Finally, Nanopore sequencing is more cost-effective than other sequencing methods for a number of reasons including cheaper costs of flow cells, kits and appliances [53], and ultimately the possibility of an in-house real-time sequencing. Thus, when enough data is acquired, the experiment can be halted and another library loaded onto the flow cell, increasing its usability and contributing to the cost-efficiency of the method. Nanopore sequencing is however still less accurate than other sequencing platforms as Illumina [64], but its purpose in this study was to help us assessing the active recombinases over the enriching workflow, not to accurately determine their sequences, which made it the best option for our objectives. We would like to highlight nonetheless the importance of exploring multiple options when considering NGS. Every method presents some benefits but also some disadvantages and selecting the appropriate sequencing alternative can be crucial for the optimal development of the experiment and should be selected from the inception of the experiment.

Many different tools have been developed by both manufacturers and third parties for the analysis of sequencing data generated by NGS devices. The application for which the specific NGS has been deployed determines the workflow for data analysis and the tools required for every step [65]. As our focus was mainly on quantifying the data rather than an application that requires high quality and accuracy, like *e.g.*, *de novo* assembly, our workflow was very simple but effective in generating results. When looking at the results of the enrichment of the naïve library in *P. putida*, it is notable that the control PapRecT seemed to be significantly more present than any other candidate or control (Figures 3.2 A, 3.2 B and 3.2 C). This recombinase's efficiency had previously been shown to be within the same range than the control Rec2 [51], which was later confirmed by our individual assessment of recombinases. Despite this inconsistency, the abundance of PapRecT remained relatively stable over the selective cycles (Supplementary Figure S3.1), which could indicate that once that an efficient recombinase is unintentionally the most prominent member of the cluster, it will likely take over the whole popula-

tion. This dominance effect could have masked the performance of other recombinases, highlighting an important drawback of this kind of high-throughput screening methods, where the different elements are competing for their place within the cell population. Individual testing of the active recombinases and the controls revealed their efficiencies in a non-competitive environment. As mentioned, PapRecT behaved then coherently to previous studies (Figure 3.3 A), but this was not the only inconsistency that was observed between the enrichment workflow results and the individual assessment. R12, despite of being the one with the lowest numbers out of the four active candidates in the selective enrichment of *P. putida* (Figure 3.2 B), ended up showing similar levels of recombineering to the other active candidates and positive controls (Figure 3.3 A), an effect that was probably caused by PapRecT taking over the population. Similarly, Rec2 and PapRecT reached very high levels of recombineering activity during their individual evaluation in *P. taiwanensis* (Figure 3.3 B) whereas they had been completely overshadowed by R47 earlier during the enrichment workflow (Figure 3.2 E).

Differences between the pooled and the individual assessment made us wonder if there would be other factors influencing the performance of a recombinase beyond its intrinsic recombineering efficiency, such as the action rate. We confirmed this hypothesis by comparing the ARF reached by PapRecT and R12 at different time points after recombineering. After only 1 h, the former had reached more than 70 % of the efficiency that it would reach after one whole day, while R12 would need 6 h more to achieve these levels, demonstrating that different recombinases mediate mutations at different speed (Figure 3.4). On the one hand, specific protein toxicity could affect the growth rate of cells carrying particular recombinases. This phenomenon has previously been described for Rec2, whose toxicity might have led to a longer lag phase or higher death rate of Rec2-producing cells after transformation [48]. On the other hand, alongside toxicity affecting growth speed, other aspects such as the process of homology detection and annealing by the recombinase might determine the action rate of the recombinases. For the λ Red system, it has been described that a complex needs to be formed between two redβ proteins bound to the ssDNA (ΔGdimerisation) as well as hybridisation between the matching strands (ΔGhybridisation) in order to overcome the energetic barrier which in-

Table 3.2: Overview of the origin of the best performing recombinases in the enrichment workflow of the different Pseudomonas species.

Recombinase	GenBank description	Original species
R8	DNA recombinase	*Pseudomonas sp. strain M1*
R9	Phage recombination protein β	*Pseudomonas sp. 91RF*
R12	Phage recombination protein β	*Pseudomonas putida*
R26	Uncharacterized	*Pseudomonas sp. AU12215*
R27	Recombination protein recT	*Pseudomonas psychrotolerans*
R35	DNA recombinase	*Pseudomonas sp. ICMP 8385*
R47	Phage recombination protein β	*Pseudomonas duriflava*

hibits interaction of the ssDNA bound to redβ with the host DNA (ΔG‡). The energy required for this interaction, or clamping, which is larger than the thermal energy (Boltzmann constant (kB) × absolute temperature (T)) can only be generated when the sequence homology is adequate [66]. The rate at which these processes occur can differ and might explain some of the inconsistencies between the two assays.

Annealing of phage RecT proteins has been linked to specific interactions with bacterial single stranded binding proteins (SSBs), which mainly relies on the recognition of seven amino acids at the SSB's C-terminal. This interaction appears to be host-specific, hence the host tropism displayed by λ redβ for disparate bacterial species. However, some RecT proteins, like PapRecT, can function more broadly by interacting with multiple SSBs [46], whose portability seems to rely on the conservation of the C-terminal sequence of the SSBs. From the newly found recombinases, R47 postulates as a broadly acting recombinase given its activity in all four tested species. This candidate derives from a phage of *Pseudomonas duriflava*, a close relative of the tested *Pseudomonas* species (Table 2). While the association between recombinase efficiency and taxonomic proximity has often been questioned in previous literature [48, 49, 67], seeking recombinases in taxonomically related species has also been successful in numerous occasions [45, 68–70]. In our study, out of our best performers, only R12 was derived from one of the tested species (Table 3.2). Further investigation on host specificity in different bacterial species might provide further insights into this question.

In addition to the limitation of the study related to "slow" potential active SSAPs masked by faster candidates, we encountered the problem of missing variants after the pooling of samples during the enrichment processes (Figures 3.2 A, 3.2 D

and 3.2 G). Especially for *P. fluorescens*, a lot of candidates (21) were not present during the enrichment workflow after pooling all strains. This species' ability to readily acquire resistance to antibiotics might have caused the strains to lose their recombinase plasmids as they became resistant to kanamycin themselves and the presence of their plasmids obsolete for survival [55]. R46 was found to be absent in all three experiments of enrichment, which likely indicates that this recombinase presents a high toxicity for all the strains, similar to what was described for Rec2 [48] but in much higher level of lethality. Another major drawback is that ARF appears to be affected by the target *locus* and consequently the oligonucleotide sequence, which results for some recombinases in showing different outcomes depending on the targeted gene. The design of the oligonucleotides does therefore influence recombineering efficiency as secondary structures can occur for oligonucleotides that target regions containing high GC content or repetitive sequences, hence decreasing efficiency [37]. In this way, we see that *i.e.*, R9 seems to be more efficient when targeting *rpoB* (barcodes 4 and 9) both in *P. taiwanensis* and *P. fluorescens* (Figures 3.2 F and 3.2 I).

As efficiency appears to be dependent on the targeted genomic region and also the action rate of the recombinase, we attempted to synergize the activity of several of the best recombinases by combining them in a single plasmid. On top of providing higher ARF levels, this chimeric plasmid could be introduced in any *Pseudomonas* host regardless which mutation needed to be incorporated, skipping the selection of the recombinase best suited for the situation. Unfortunately, this synergetic effect was not observed in any of the species, concluding that most likely the levels of recombinase protein are not the limiting factors of the allelic replacement process as it had been previously suggested [51]. The chimeric plasmid however was able to report similar efficiencies to those obtained by the best performers individually in the different *Pseudomonas*, except *P. fluorescens*, which makes it a still useful tool to use indistinctively for different species.

Lastly, we assessed the accumulation of off-target mutations while using the newly characterized SSAPs (Table 3.1). Cells expressing *r9* and *r35* showed off-target results within the same range than data previously reported for redβ, PapRecT and Rec2 (1.0 ± 0.7, 3.3 ± 0.6, and 1.7, respectively) [49, 51]. This was also the case for

r47-expressing *P. fluorescens* and *P. aeruginosa* cells. Unfortunately, *P. putida* and *P. taiwanensis* cells expressing the same recombinase genes accumulated elevated numbers of mutations. These differences on the off-target effect of R47 in the four *Pseudomonas* species could be due to the fact that the mismatch repair (MMR) machinery disruptor employed in all the plasmids of this study, MutLE36KPP, is the dominant-negative mutant of *P. putida* KT2440 MutL. While the MutL proteins of *P. putida* and *P. taiwanensis* share a 98.6 % of identity, the one of *P. putida* shares with the ones of *P. fluorescens* and *P. aeruginosa* only 85.5 and 81.7 %, respectively. This could indicate that the transient expression of MutLE36KPP could have not disabled MMR completely or nothing at all in the latter two species, which would have resulted in the observed reduced amount of off-target mutations. Even though previous studies have demonstrated that this transient overexpression of dominant-negative MutL mutants greatly reduces the off-target effect when compared to complete inactivation of the MMR machinery [38, 71], we observed an elevated accumulation of this type of mutations only in those species in which MutLE36KPP had most likely taken effect. *P. putida* cells carrying the other two recombinases, R8 and R12, also showed higher background mutation rates, which would support this hypothesis. R35, conversely, could have not yielded the same results due to its lower overall efficiency in this host. Given that the off-target effects are one of the biggest constrains of recombineering [31], this phenomenon deserves further investigation in terms of recombinases and species-specificy, as well as MMR inhibition.

Overall, this study provides new means for efficient recombineering in four *Pseudomonas* species, two of them previously missing any tool for this kind of genome editing. In two cases, the barrier of the 1 % ARF was overcome [31], with R47 in *P. putida* KT2440 and both with Rec2 and PapRecT in *P. taiwanensis* VLB120, while values higher than 0.4 % ARF were achieved for *P. fluorescens* and *P. aeruginosa* with R9 and R47, respectively. These findings offer the promise of radically expanding the scope and throughput of genetic manipulation of these organisms increasing their relevance as biotechnological chassis. Furthermore, as important as identifying efficient recombinases was expanding the understanding of key processes that accompany recombineering and improving SSAP screening methods. The fact that out of an extensive number of potential recombinases, the final mod-

est selection of 49 yielded several efficient candidates provides confidence in the proposed workflow to find effective recombinases not only for the here researched *Pseudomonas* sp. but also for other bacterial species in the future.

Materials and methods

Identification of *Pseudomonas* SSAP candidates

To comprehensively survey publicly available metagenomics databases for potential SSAP candidates, we first assembled seven SSAPs which represent the diverse nature of bacteriophage-derived recombinases with single-stranded annealing activity for multiple sequence alignment using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). These SSAPs include β from *Escherichia coli*/ λ phage (GeneBank: NP_040617.1), S065 from *Vibrio cholerae* (GeneBank: AAL59710.1), Plu2935 from *Photorhabdus laumondii* (GeneBank: WP_011147155.1), ERF from *Salmonella typhimurium*/P22 phage (GeneBank: NP_059596.1), Sak from *Lactococcus lactis*/ul36 phage (GeneBank: NP_663647.1), ERF from *Pseudomonas aeruginosa*/D3 phage (GeneBank: NP_061548.2) and DdRB from *Deinococcus radiodurans* (GeneBank: DAA06535.1) (Supplementary Table S1). Subsequently, we used jackHMMER (https://www.ebi.ac.uk/Tools/hmmer/search/jackhmmer) to search for protein homologous in UniProtKB within the taxonomy of Eubacteria and Viruses based on the HMM-profile built from multiple sequence alignment of the above-mentioned SSAPs. jackHMMER hits after 15 rounds were filtered for the ones relevant to *Pseudomonas* (Supplementary Table S2). Next, clustering of protein sequencing of *Pseudomonas* relevant hits was done using CD-HIT (PMID: 16731699) (http://weizhong-lab.ucsd.edu/cdhit-web-server/cgi-bin/index.cgi) with a setting of similarity = 0.7, which resulted in 48 clusters (Supplementary Table S3). By prioritizing members of each cluster based on the HMMER e-value (the smaller the e-value, the higher the homology to the SSAP HMM-profile) and the relevance to a phage source, we selected a total of 49 *Pseudomonas* SSAP candidates (Supplementary Tables S4 and S5) to screen for their recombineering activity.

Table 3.3: List of antibiotics used in this study with their corresponding concentrations for each species.

Species	Kanamycin	Gentamicin	Streptomycin	Rifampicin	5-Fluorooritic acid	Nalidixic acid	Ciproflaxin
P. putida	50 g/mL	-	100 g/mL	100 g/mL	250 g/mL	50 g/mL	-
P. taiwanensis	50 g/mL	-	100 g/mL	100 g/mL	250 g/mL	50 g/mL	-
P. fluorescens	100 g/mL	-	100 g/mL	100 g/mL	250 g/mL	50 g/mL	-
P. aeruginosa	-	100 g/mL	-	100 g/mL	-	-	0.5 g/mL

Bacterial strains and media

All strains used in this research are listed in Supplementary Table S7 and were grown in Lysogeny Broth (LB) (10 g/L NaCl, 10 g/L tryptone, and 5 g/L yeast extract) supplemented with the proper antibiotics when needed, unless stated otherwise. *E. coli* DH5α cells were made chemically competent as previously described [72] and were used for cloning purposes. Electrocompetent *Pseudomonas* strains were prepared as previously described [73] and used for recombineering experiments. After standard transformation, cells were allowed to recover in super optimal broth with catabolite repression (SOC) (31 g/L SOC powder), while Terrific Broth (TB) (12 g/L tryptone, 24 g/L yeast extract, 0.4 % (v/v) glycerol, and 10 % (v/v) phosphate buffer (23.12 g/L KH_2PO_4 and 125.4 g/L K_2HPO_4)) was used for the same purpose after electroporation during recombineering cycles. *E. coli* cells were grown at 37°C and all *Pseudomonas* strains at 30°C, even *P. aeruginosa*, whose optimal temperature is 37°C, to prevent the activation of the thermo-inducible cIpL promoter controlling the recombinase genes and *mutLE36KPP*. Medium was supplemented with antibiotics and supplements following the concentrations shown in Table 3.3. *P. putida* KT2440, *P. taiwanensis* VLB120 and *P. fluorescens* SBW25 were managed in an ML-1 laboratory, whereas *P. aeruginosa* PAO1 was managed in an ML-2 under more secured conditions due to its potential pathogencity.

Plasmids

All plasmids employed during this study, as well as the oligonucleotides used to construct them can be found in Supplementary Tables S8 and S9, respectively. Putative recombinase sequences were manufactured by IDT as eBlocks of maximal 900 bps, with BsaI compatible sticky ends. All of them were cloned into a pSEVA2514-mutLE36KPP backbone previously amplified with primers C1 and C2 by means of Golden Gate or the Golden Gate-based SevaBrick Assembly [74]. Addition-

ally, recombinases R8, R9, R12, R35 and R47 were equally cloned into a pSEVA6514-mutLE36KPP backbone for the experiments in *P. aeruginosa* PAO1. Gibson Assembly using the NEBuilder HiFi DNA Assembly Master Mix was utilised following manufacturer's instructions when Golden Gate did not suffice due to difficulties during the assembly process. This was the case for R26, R30, R37 and R49, which had to be amplified using the sets of primers GA R26 FW-RV, GA R30 FW-RV, GA R37 FW-RV and GA R49 FW-RV in order to be cloned into a pSEVA2514-mutLE36KPP backbone generated with primers GA BB FW-RV. Construction of the chimeric pSEVA(2/6)514-R9-Rec2-R12-R47-PapRecT-mutLE36KPP plasmids was done *via* Golden Gate using aforementioned backbones pSEVA2514-mutLE36KPP and pSEVA6514-mutLE36KPP and DNA fragments of R9, Rec2, R12, R47 and PapRecT, amplified with primer pairs C3-C4, C5-C6, C7-C8, C9-C10, C11-C12, respectively. All resulting plasmids were transformed *via* heat shock into chemically competent *E. coli* DH5α cells for plasmid amplification and subsequently isolated using GenJET Plasmid Miniprep Kit® (Thermo Scientific). Plasmid presence was confirmed by colony PCR using Phire Hot start II DNA polymerase (Thermo Fisher Scientific) and primers V1 and V2 according to manufacturer's guidelines. Plasmid sequence was confirmed using Sanger sequencing and primers V1 and V2 by Macrogen (MACROGEN Inc. DNA Sequencing Service; Amsterdam, The Netherlands).

Oligonucleotides

Single-stranded (ss) DNA oligonucleotides employed in this study (Supplementary Table S9) were purchased from Integrated DNA technologies (IDT) as salt-free without further purification, resuspended in milli-Q at 100 μM and long-term stored at 20°C.

Recombineering oligonucleotides were designed according to previously optimized parameters [47, 51]. In brief, they were 60-mer oligonucleotides in which the mutation change was incorporated at the middle position and without phosphorothioate bonds in their sequence. They were designed complementary to the lagging strand of replicating DNA and with predicted folding energies higher or equal than 16 kcal/mol. Recombineering oligonucleotides used in this study were design to confer different antibiotic resistances by means of small mutations (up to 3 nu-

cleotides) in key genes of the different hosts: (A) *rpsL* K43T mutation in *P. putida*, *P. taiwanensis* and *P. fluorescens* conferred streptomycin resistance, (B) *rpoB* Q518L mutation in *P. putida*, *P. taiwanensis* and *P. fluorescens* and *rpoB* D521V mutation in *P. aeruginosa* conferred rifampicin resistance, (C) *pyrF* E50stop mutation in *P. putida*, *pyrF* Q50stop mutation in *P. taiwanensis* and *pyrF* G50stop mutation in *P. fluorescens* conferred 5-fluorooritic acid resistance, (D) *gyrA* D87N mutation in *P. putida*, *P. taiwanensis* and *P. fluorescens* conferred nalidixic acid resistance, and (E) *gyrA* T83I mutation in *P. aeruginosa* conferred ciproflaxin resistance (Supplementary Table S9).

Serial enrichment recombineering

Recombinase library plasmids were transformed into electrocompetent *P. putida* KT2440, *P. taiwanensis* VLB120 and *P. fluorescens* SBW25 cells. Presence of the plasmids was confirmed again by using cPCR using Phire Hot start II DNA polymerase (Thermo Fisher Scientific) and primers V1 and V2 according to manufacturer's guidelines. Overnight cultures of all strains containing each of the putative recombinases and controls were diluted to OD_{600} = 0.1 and pooled into two replicas containing the control strains and one without them. Once the cultures grew to mid-log phase ($OD600$ = 0.5-0.7), samples were induced for 10 min at 42°C, inactivated on ice for 5 min, washed twice with 10 and 1 mL 300 mM sucrose, respectively, and finally resuspended in 200 μL of 300 mM sucrose. A 100 μL aliquot of competent cells was mixed with 1 μL of an appropriate 100 μM oligonucleotide for subsequent electroporation (2500 V, 200 Ω and 25 F). Cells were allowed to recover in TB until $OD600$ = 0.3-0.5 before applying selective pressure *via* LB with antibiotics and overnight growth. The steps were repeated with all 4 oligonucleotides and then the entire 4-cycle protocol was repeated by miniprepping the samples after the first round and then reintroducing the enriched library into naive host strains.

Oxford Nanopore NGS Workflow

Samples for sequencing were collected after every recombineering step and stored as 25 % glycerol stocks until all cycles were completed. Samples were later grown, and plasmid content was subsequently extracted and diluted to get the optimal

amount of DNA (400 ng) for the sequencing library preparation. Library preparation was done using the ONT Rapid Barcoding Sequencing kit SQK-RBK004, which fragments the DNA and adds a barcode in one step, and the ONT Flow Cell FLO-MIN106D R9 was primed using the ONT Flow Cell Priming kit EXP-FLP002, which prepares the cell for the library and provides fuel for the experiment. The Rapid Barcoding kit provided 12 barcodes, which were distributed for each one of the sample collection points (11) (Supplementary Table S10) and for the λ control DNA (Control Expansion kit EXP-CTL001) provided by ONT, to ensure the libraries were of sufficient quality. Next, sequencing was initiated and monitored by the accompanying software minKNOW by ONT. Fast basecalling was chosen as our interest lies in the quantification of the library, not high accurate sequencing. Bias Voltage was adjusted when necessary, according to the guidelines on the ONT website. In between runs, the flow cell was washed (wash kit EXP-WSH004 by ONT) to ensure no traces of the previous sample remained, as we made use of the same barcodes in every library. A total of three libraries were run for the *P. putida* experiment, two with controls and one without, and only the two samples with controls were run for *P. fluorescens* and *P. taiwanensis*.

Next Generation Sequencing data analysis

The raw sequencing data from the minION stored in Fast5 files was basecalled using the Guppy Basecaller (ONT) yielding FastQ files. The basecaller software allowed for demutiplexing the files based on their barcodes (1-12). Files were then merged into one FastQ file per barcode per sample. The Guppy Aligner (ONT) was used for the alignment of the merged files and generated the alignment files in BAM-format. BAM (and SAM) files were manipulated with readily available tools like Samtools, which was used in this instance to quantify the recombinase abundance over the cycles of the selective enrichment (Samtools idxstats). This data was finally used for visualising the enrichment. The software used, minKNOW GUI and Guppy, are available on the Oxford Nanopore Technologies website and the Samtools code is freely available on Github or https://www.htslib.org/. The relevant recombinase genes of this study R8, R9, R12, R26, R27, R35, R47 and the controls Rec2, PapRecT and λ Red were *de novo* assembled from the raw data from barcode01 using the

flye software, and polished with the Medaka tool to create consensus sequences and ensure the integrity of these genes in the final experiment before enrichment.

Individual Assessment of selected SSAPs

Strains containing recombinases that had shown activity in the serial enrichment workflow were subjected to individual assessment *via* recombineering. Recombineering cycles were performed according to the previously described standard protocol [51, 71], in a similar manner than it had been done during the serial enrichment workflow but without pooling the different samples and only for one cycle. After transformation with the corresponding oligonucleotides, cells were left to recover in 5 mL of TB at 30°C and 250 rpm. Afterwards, 15 mL of LB was added, and samples were incubated overnight under the same conditions. Next day, appropriate dilutions were cultured onto LB plates both containing and lacking the corresponding antibiotic whose resistance would have been conferred by the mutation incorporated during the recombineering cycle for subsequent efficiency calculation. Cultures without thermo-induction were included as negative controls during one of the replicas to discard any possible native or spontaneous resistance to the antibiotics. Recombineering efficiency was calculated as the ratio between antibiotic resistant and total colony forming units (CFU).

During the assay to evaluate the different recombineering efficiency at different times, cells were plated and therefore subjected to the antibiotic pressure at different times after the transformation of the oligonucleotides. Specifically, cells were plated after 1, 3, 5, 7, 9 and 24 h after the addition of the TB that initiates the phase of recovery.

Off-target analysis

In order to investigate the off-target mutations that occur in the cells during the process of recombineering when using each individual SSAP, we used the GenElute Bacterial Genomic DNA kit (Sigma-Aldrich) to isolate gDNA from two colonies of the next strains: (i) *P. putida* expressing pSEVA2514-R8-mutLE36KPP, (ii) *P. putida* expressing pSEVA2514-R12-mutLE36KPP, (iii) *P. putida* expressing pSEVA2514-R35-mutLE36KPP, (iv) *P. putida* expressing pSEVA2514-R47-mutLE36KPP, (v) *P. taiwanen-*

sis expressing pSEVA2514-R47-mutLE36KPP, (vi) *P. fluorescens* expressing pSEVA2514-R9-mutLE36KPP, (vii) *P. fluorescens* expressing pSEVA2514-R47-mutLE36KPP, (viii) *P. aeruginosa* expressing pSEVA6514-R8-mutLE36KPP. Successful incorporation of an oligonucleotide mediating either the *rpsL* gene mutation in *P. putida*, *P. taiwanensis* and *P. fluorescens*, or the *rpoB* gene mutation in *P. aeruginosa* had been performed using one recombineering cycle in the aforementioned strains prior gDNA isolation. In the same way, but without undertaking any recombineering, we extracted the gDNA from one colony of each four wild type parental strains. Isolated gDNA was sent for whole genome sequencing to Novogene Co. Ltd. (Beijing, China) for Illumina sequencing. Fastp (v0.20.0) was used to trim the raw Illumina reads for low quality and Illumina adapters. Subsequently, breseq (v0.35.5) was utilized to find the mutations using the reference genomes and annotations of *P. putida* KT2440 (NC_002947.4), *P. taiwanensis* VLB120 (NC_022738), *P. fluorescens* SBW25 (NC_012660.1), and *P. aeruginosa* PAO1 (NC_002516.2). Final number of off-target mutations was calculated as the average of the total number of non-intended mutations in the two colonies after the removal of those mutations also present in the sequencing results of the parental strains.

Acknowledgements

We would like to thank Lars Blank, Xu Cheng and Janneke Elzinga for providing the reference strains *P. taiwanensis* VLB120, *P. fluorescens* SBW25 and *P. aeruginosa* PAO1, respectively. We also thank Peter N. Ciaccia for the useful discussions about this work. This work is part of the research programs SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis, which is supported by the Dutch Research Council (NWO) [15814]. Funding for open access charge: Dutch Research Council.

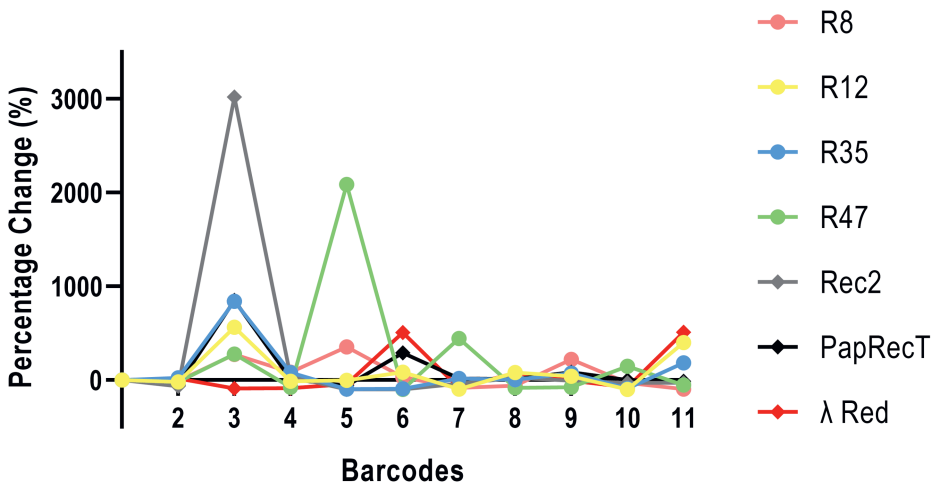
Conflict of interest

The authors declare there are no conflicting interests.

Author contributions

Conceptualization: EAG/ZL/FJI/VAPMds; Data curation: EAG/LGM/TB/ZL; Formal analysis: EAG/LGM/TB/ZL; Funding acquisition: FJI/VAPMdS; Investigation: EAG/LGM/TB/ZL; Methodology: EAG/LGM/TB/ZL; Resources: FJI/VAPMdS; Software: LGM/TB/ZL; Supervision: FJI/VAPMdS; Visualization: EAG; Writing – original draft: EAG; Writing – review & editing: EAG/LGM/TB/ZL/VAPMdS

Supplementary material



Supplementary Figure S3.1: Graph displaying the percentage change (increase and decrease) of the number of reads per recombinase compared to the previous cycle during the serial enrichment recombineering of *P. putida*. To calculate percentage change, the difference (increase or decrease) between the percentages of a given recombinase within the whole population in two consecutive cycles was calculated. Then, the difference was divided by the percentage of that recombinases in the first of the two cycles (original number) and multiplied by 100.

The rest of the supplementary material of this chapter, including Supplementary Tables S1 to S10, can be accessed *via*:

https://figshare.com/s/4e5c392861d598297f59

Bibliography

1. Poblete-Castro, I., Wittmann, C. & Nikel, P. I. Biochemistry, genetics and biotechnology of glycerol utilization in Pseudomonas species. *Microbial Biotechnology* **13,** 32–53 (2020).
2. Kumar, A., Verma, H., Singh, V. K., *et al.* in *Agriculturally Important Microbes for Sustainable Agriculture: Volume 2: Applications in Crop Production and Protection* (eds Meena, V. S., Mishra, P. K., Bisht, J. K., *et al.*) 195–215 (Singapore, 2017).
3. Chavarría, M., Nikel, P. I., Pérez-Pantoja, D., *et al.* The Entner-Doudoroff pathway empowers Pseudomonas putidaKT2440 with a high tolerance to oxidative stress. *Environmental Microbiology* **15,** 1772–1785 (2013).
4. Martínez-García, E. & de Lorenzo, V. Molecular tools and emerging strategies for deep genetic/genomic refactoring of Pseudomonas. *Current opinion in biotechnology* **47,** 120–132 (2017).
5. Nikel, P. I., Martínez-García, E. & De Lorenzo, V. Biotechnological domestication of pseudomonads using synthetic biology. *Nature Reviews Microbiology* **12,** 368–379 (2014).
6. Nikel, P. I., Chavarría, M., Fuhrer, T., *et al.* Pseudomonas putida KT2440 strain metabolizes glucose through a cycle formed by enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and pentose phosphate pathways. *Journal of Biological Chemistry* **290,** 25920–25932 (2015).
7. Dos Santos, V. A. P. M., Heim, S., Moore, E. R. B., *et al.* Insights into the genomic basis of niche specificity of Pseudomonas putida KT2440. *Environmental Microbiology* **6,** 1264–1286 (2004).
8. Poblete-Castro, I., Becker, J., Dohnt, K., *et al.* Industrial biotechnology of Pseudomonas putida and related species. *Applied Microbiology and Biotechnology* **93,** 2279–2290 (2012).
9. Loeschcke, A. & Thies, S. Pseudomonas putida-a versatile host for the production of natural products. *Applied Microbiology and Biotechnology* **99,** 6197–6214 (2015).
10. Panke, S., Witholt, B., Schmid, A., *et al.* Towards a biocatalyst for (S)-styrene oxide production: characterization of the styrene degradation pathway of Pseudomonas sp. strain VLB120. *Applied and Environmental Microbiology* **64,** 2032–2043 (1998).
11. Park, J.-B., Bühler, B., Panke, S., *et al.* Carbon metabolism and product inhibition determine the epoxidation efficiency of solvent-tolerant Pseudomonas sp. strain VLB120DeltaC. *Biotechnology and Bioengineering* **98,** 1219–1229 (2007).
12. Volmer, J., Schmid, A. & Bühler, B. The application of constitutively solvent-tolerant P. taiwanensis VLB120ΔCΔttgV for stereospecific epoxidation of toxic styrene alleviates carrier solvent use. *Biotechnology Journal* **12** (2017).
13. Wynands, B., Otto, M., Runge, N., *et al.* Streamlined Pseudomonas taiwanensis VLB120 Chassis Strains with Improved Bioprocess Features. *ACS Synthetic Biology* **8,** 2036–2050 (2019).
14. Wordofa, G. G. & Kristensen, M. Tolerance and metabolic response of Pseudomonas taiwanensis VLB120 towards biomass hydrolysate-derived inhibitors. *Biotechnology for Biofuels* **11,** 199 (2018).
15. Köhler, K. A. K., Blank, L. M., Frick, O., *et al.* D-Xylose assimilation via the Weimberg pathway by solvent-tolerant Pseudomonas taiwanensisVLB120. *Environmental Microbiology* **17,** 156–170 (2015).
16. Paulsen, I. T., Press, C. M., Ravel, J., *et al.* Complete genome sequence of the plant commensal Pseudomonas fluorescens Pf-5. *Nature Biotechnology* **23,** 873–878 (2005).
17. David, B. V., Chandrasehar, G. & Selvam, P. N. *Pseudomonas fluorescens: a plant-growth-promoting rhizobacterium (PGPR) with potential role in biocontrol of pests of crops* 221–243 (2018).
18. Gonçalves, L. D. d. A., Piccoli, R. H., Peres, A. d. P., *et al.* Predictive modeling of Pseudomonas fluorescens growth under different temperature and pH values. *Brazilian Journal of Microbiology* **48,** 352–358 (2017).

19. MacLean, A., Bley, A. M., Appanna, V. P., *et al.* Metabolic manipulation by Pseudomonas fluorescens: a powerful stratagem against oxidative and metal stress. *Journal of Medical Microbiology* **69,** 339–346 (2020).

20. Wasi, S., Tabrez, S. & Ahmad, M. Use of Pseudomonas spp. for the bioremediation of environmental pollutants: a review. *Environmental Monitoring and Assessment* **185,** 8147–8155 (2013).

21. Olanrewaju, O. S., Glick, B. R. & Babalola, O. O. Mechanisms of action of plant growth promoting bacteria. *World Journal of Microbiology & Biotechnology* **33,** 197 (2017).

22. Sriramulu, D. *Pseudomonas Aeruginosa: An Armory Within* (2019).

23. Huertas, M. J. & Matilla, M. A. Training bacteria to produce environmentally friendly polymers of industrial and medical relevance. *Microbial Biotechnology* **13,** 14–16 (2019).

24. Qi, L. & Christopher, G. F. Role of Flagella, Type IV Pili, Biosurfactants, and Extracellular Polymeric Substance Polysaccharides on the Formation of Pellicles by Pseudomonas aeruginosa. *Langmuir* **35,** 5294–5304 (2019).

25. Soberón-Chávez, G., González-Valdez, A., Soto-Aceves, M. P., *et al.* Rhamnolipids produced by Pseudomonas: from molecular genetics to the market. *Microbial Biotechnology* **14,** 136–146 (2021).

26. Wu, W., Jin, Y., Bai, F., *et al. Chapter 41 - Pseudomonas aeruginosa* (eds Tang, Y.-W., Sussman, M., Liu, D., *et al.*) 753–767 (Boston, 2015).

27. Lee, S. Y. & Kim, H. U. Systems strategies for developing industrial microbial strains. *Nature Biotechnology* **33,** 1061–1072 (2015).

28. Straathof, A. J. J., Wahl, S. A., Benjamin, K. R., *et al.* Grand Research Challenges for Sustainable Industrial Biotechnology. *Trends in Biotechnology* **37,** 1042–1050 (2019).

29. Yin, J., Zheng, W., Gao, Y., *et al.* Single-Stranded DNA-Binding Protein and Exogenous RecBCD Inhibitors Enhance Phage-Derived Homologous Recombination in Pseudomonas. *iScience* **14,** 1–14 (2019).

30. Blombach, B., Grünberger, A., Centler, F., *et al.* Exploiting unconventional prokaryotic hosts for industrial biotechnology. *Trends in Biotechnology* (2021).

31. Wannier, T. M., Ciaccia, P. N., Ellington, A. D., *et al.* Recombineering and MAGE. *Nat Rev Methods Primers* **1,** 1–24 (2021).

32. Ellis, H. M., Yu, D., DiTizio, T., *et al.* High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proceedings of the National Academy of Sciences* **98,** 6742–6746 (2001).

33. Yu, D., Ellis, H. M., Lee, E. C., *et al.* An efficient recombination system for chromosome engineering in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **97,** 5978–5983 (2000).

34. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences* **97,** 6640–6645 (2000).

35. Mosberg, J. A., Lajoie, M. J. & Church, G. M. Lambda red recombineering in Escherichia coli occurs through a fully single-stranded intermediate. *Genetics* **186,** 791–799 (2010).

36. Marinelli, L. J., Hatfull, G. F. & Piuri, M. Recombineering: A powerful tool for modification of bacteriophage genomes. *Bacteriophage* **2,** 5–14 (2012).

37. Wang, H. H., Isaacs, F. J., Carr, P. A., *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460,** 894–898 (2009).

38. Nyerges, Á., Csörgő, B., Nagy, I., *et al.* A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *Proc Natl Acad Sci USA* **113,** 2502–2507 (2016).

39. Nyerges, Á., Csörgő, B., Draskovits, G., *et al.* Directed evolution of multiple genomic loci allows the prediction of antibiotic resistance. *Proc Natl Acad Sci USA* **115,** E5726–E5735 (2018).

40. Farzadfard, F. & Lu, T. K. Genomically encoded analog memory with precise in vivo DNA writing in living cell populations. *Science* **346** (2014).

41. Schubert, M. G., Goodman, D. B., Wannier, T. M., *et al.* High-throughput functional variant screens via in vivo production of single-stranded DNA. *Proceedings of the National Academy of Sciences* **118** (2021).

42. Amiram, M., Haimovich, A. D., Fan, C., *et al.* Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nature Biotechnology* **33,** 1272–1279 (2015).

43. Isaacs, F. J., Carr, P. A., Wang, H. H., *et al.* Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement. *Science* **333,** 348–353 (2011).

44. Lajoie, M. J., Rovner, A. J., Goodman, D. B., *et al.* Genomically recoded organisms expand biological functions. *science* **342,** 357–360 (2013).

45. Sun, Z., Deng, A., Hu, T., *et al.* A high-efficiency recombineering system with PCR-based ssDNA in Bacillus subtilis mediated by the native phage recombinase GP35. *Appl Microbiol Biotechnol* **99,** 5151–5162 (2015).

46. Filsinger, G. T., Wannier, T. M., Pedersen, F. B., *et al.* Characterizing the portability of phage-encoded homologous recombination proteins. *Nature Chemical Biology,* 1–9 (2021).

47. Aparicio, T., Jensen, S. I., Nielsen, A. T., *et al.* The Ssr protein (T1E_1405) from *Pseudomonas putida* DOT-T1E enables oligonucleotide-based recombineering in platform strain *P. putida* EM42. *Biotechnology Journal* **11,** 1309–1319 (2016).

48. Ricaurte, D. E., Martínez-García, E., Nyerges, Á., *et al.* A standardized workflow for surveying recombinases expands bacterial genome-editing capabilities. *Microb. Biotechnol.* **11,** 176–188 (2018).

49. Wannier, T. M., Nyerges, A., Kuchwara, H. M., *et al.* Improved bacterial recombineering by parallelized protein discovery. *PNAS* **117,** 13689–13698 (2020).

50. Aparicio, T., Nyerges, A., Martínez-García, E., *et al.* High-efficiency multi-site genomic editing of Pseudomonas putida through thermoinducible ssDNA recombineering. *Iscience* **23,** 100946 (2020).

51. Asin-Garcia, E., Martin-Pascual, M., Garcia-Morales, L., *et al.* ReScribe: An Unrestrained Tool Combining Multiplex Recombineering and Minimal-PAM ScCas9 for Genome Recoding Pseudomonas putida. *ACS Synthetic Biology* (2021).

52. Deamer, D., Akeson, M. & Branton, D. Three decades of nanopore sequencing. *Nature Biotechnology* **34,** 518–524 (2016).

53. Li, R., Xie, M., Dong, N., *et al.* Efficient generation of complete sequences of MDR-encoding plasmids by rapid assembly of MinION barcoding sequencing data. *GigaScience* **7,** 1–9 (2018).

54. Murphy, K. C. λ Recombination and Recombineering. *EcoSal Plus* **7** (2016).

55. Trivedi, M., Patil, S., Shettigar, H., *et al.* Antimicrobial sensitivity pattern of Pseudomonas fluorescens after biofield treatment. *Infectious Diseases and Therapy* **3** (2015).

56. Poole, K. Pseudomonas Aeruginosa: Resistance to the Max. *Frontiers in Microbiology* **2,** 65 (2011).

57. Lesic, B. & Rahme, L. G. Use of the lambda Red recombinase system to rapidly generate mutants in Pseudomonas aeruginosa. *BMC Molecular Biol* **9,** 20 (2008).

58. Liang, R. & Liu, J. Scarless and sequential gene modification in Pseudomonas using PCR product flanked by short homology regions. *BMC microbiology* **10,** 1–9 (2010).

59. Gallagher, R. R., Li, Z., Lewis, A. O., *et al.* Rapid editing and evolution of bacterial genomes using libraries of synthetic DNA. *Nat Protoc* **9,** 2301–2316 (2014).

60. Pennisi, E. Search for Pore-fection. *Science* **336,** 534–537 (2012).

61. Head, S. R., Komori, H. K., LaMere, S. A., *et al.* Library construction for next-generation sequencing: overviews and challenges. *BioTechniques* **56,** 61–64, 66, 68, passim (2014).

62. Bruinsma, S., Burgess, J., Schlingman, D., *et al.* Bead-linked transposomes enable a normalization-free workflow for NGS library preparation. *BMC Genomics* **19,** 722 (2018).

63. Hess, J., Kohl, T., Kotrová, M., *et al.* Library preparation for next generation sequencing: A review of automation strategies. *Biotechnology Advances* **41,** 107537 (2020).

64. Lu, H., Giordano, F. & Ning, Z. Oxford Nanopore MinION Sequencing and Genome Assembly. *Genomics, Proteomics & Bioinformatics* **14,** 265–279 (2016).

65. Amarasinghe, S. L., Su, S., Dong, X., *et al.* Opportunities and challenges in long-read sequencing data analysis. *Genome Biology* **21,** 30 (2020).

66. Ander, M., Subramaniam, S., Fahmy, K., *et al.* A Single-Strand Annealing Protein Clamps DNA to Detect and Secure Homology. *PLOS Biology* **13,** e1002213 (2015).

67. Datta, S., Costantino, N., Zhou, X., *et al.* Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proceedings of the National Academy of Sciences of the United States of America* **105,** 1626–1631 (2008).

68. Van Kessel, J. C. & Hatfull, G. F. Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering: characterization of antimycobacterial drug targets. *Molecular Microbiology* **67,** 1094–1107 (2008).

69. Swingle, B., Bao, Z., Markel, E., *et al.* Recombineering Using RecTE from Pseudomonas syringae. *Applied and Environmental Microbiology* (2010).

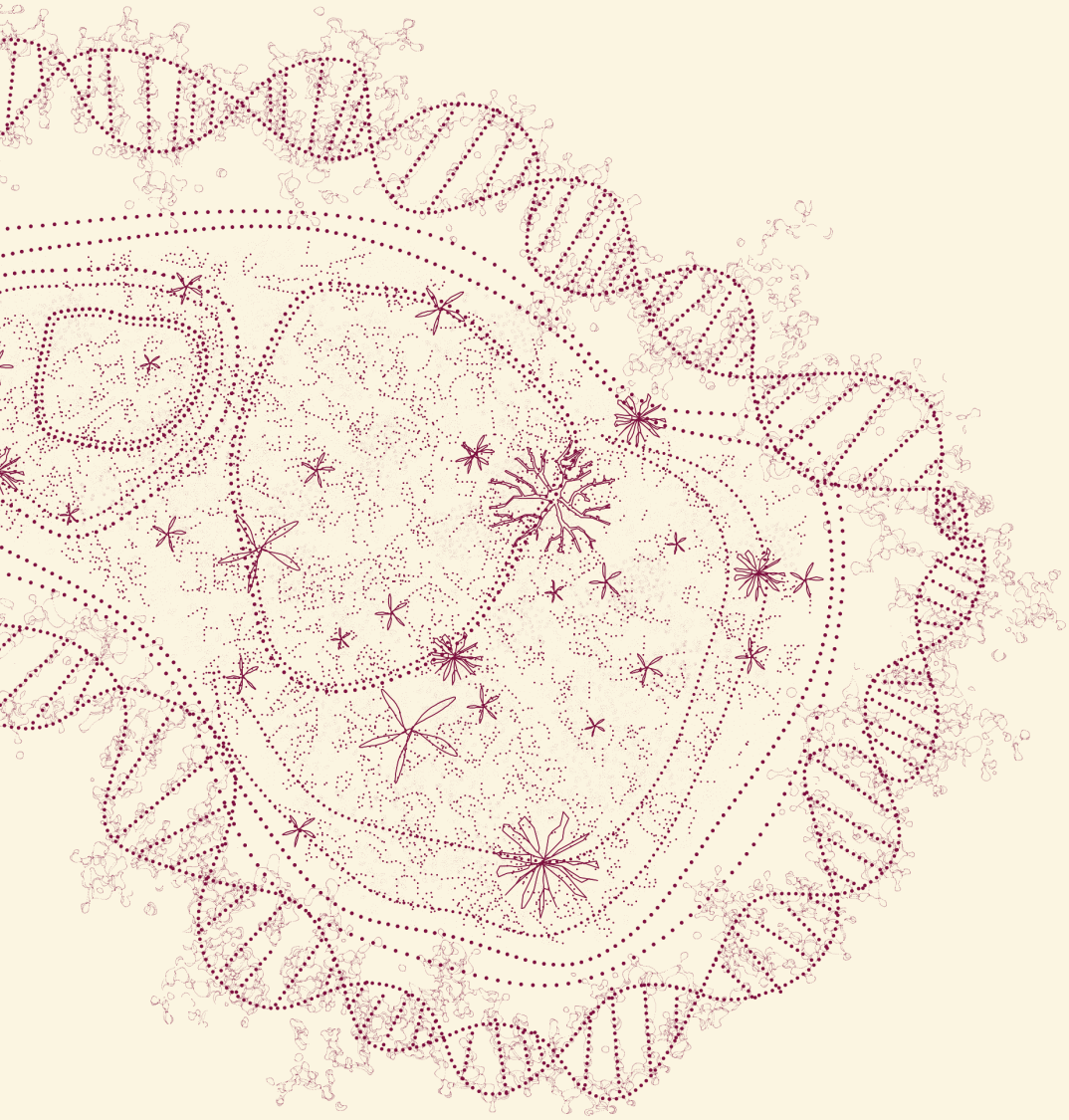
70. Van Pijkeren, J.-P. & Britton, R. A. High efficiency recombineering in lactic acid bacteria. *Nucleic Acids Res* **40,** e76 (2012).

71. Aparicio, T., Nyerges, A., Nagy, I., *et al.* Mismatch repair hierarchy of Pseudomonas putida revealed by mutagenic ssDNA recombineering of the pyrF gene. *Environmental Microbiology* **22,** 45–58 (2020).

72. Green, R. & Rogers, E. J. Transformation of Chemically Competent E. coli. **529,** 329–336 (2013).

73. Choi, K.-H., Kumar, A. & Schweizer, H. P. A 10-min method for preparation of highly electrocompetent Pseudomonas aeruginosa cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J Microbiol Methods* **64,** 391–397 (2006).

74. Damalas, S. G., Batianis, C., Martin-Pascual, M., *et al.* SEVA 3.1: enabling interoperability of DNA assembly among the SEVA, BioBricks and Type IIS restriction enzyme standards. *Microbial biotechnology* **13,** 1793–1806 (2020).



CHAPTER 4

GenoMine: a CRISPR-Cas9-based kill-switch for biocontainment of *Pseudomonas putida*

Enrique Asin-Garcia, Maria Martin-Pascual, Claudia de Buck, Max Allewijn, Alexandra Müller, Vitor A. P. Martins dos Santos

Abstract

Synthetic genetic circuits have revolutionized our capacity to control cell viability by conferring microorganisms with programmable functionalities to limit survival to specific environmental conditions. Here we present the GenoMine safeguard, a CRISPR-Cas9-based kill switch for the biotechnological workhorse *Pseudomonas putida* that employs repetitive genomic elements as cleavage targets to unleash a highly genotoxic response. To regulate the system's activation, we tested various circuit-based mechanisms including the digitalized version of an inducible expression system that operates at transcriptional level and different options of post-transcriptional riboregulators. All of them were applied not only to directly control Cas9 and its lethal effects, but also to modulate the expression of two of its inhibitors: the AcrIIA4 anti-CRISPR protein and the transcriptional repressor TetR. Either upon direct induction of the endonuclease or under non-induced conditions of its inhibitors, the presence of Cas9 suppressed cell survival which could be exploited beyond biocontainment in situations where further CRISPR genome editing is undesirable.

Introduction

In recent years, *Pseudomonas putida* KT2440 has come to occupy a prominent place on the list of synthetic biology chassis [1]. Its unique metabolic and physiological properties that provide high levels of solvent tolerance and stress resistance made it first appealing for degradation studies and harsh biotransformations [2]. However, the expanding plethora of molecular and *in silico* tools for targeted genetic manipulations is behind the seemingly unstoppable rise of this bacterium as one of the preferred flexible and engineerable hosts for general metabolic engineering and bioproduction [3, 4]. While the toolbox of methods and genetic parts grows day by day, many of them remain unapplied and overshadowed by traditionally reliable techniques and other established elements. Specifically, numerous native and heterologous inducible expression systems [5–7] alongside different synthetic libraries of genetic regulatory elements like promoters, ribosome binding sites (RBS) and terminators have been characterized in *P. putida* [8–12]. However, and by the same token, the majority of these has never been exploited in an application for gene expression modulation yet, which is fundamental for a proper control of engineered metabolic pathways. At the same time, more complex elements such as riboswitches and riboregulators to fine-tune gene expression are scarce in this bacterium hindering the development of synthetic genetic circuits to tightly control cellular functions [4].

Programming microorganisms through circuit-oriented innovations has led to novel applications and functionalities that range from diagnostics and therapeutics [13] to advanced bioproduction [14, 15]. Moreover, genetic circuits offer an excellent means for controlling cell viability and consequently for biological containment. These engineered cells would sense the environment and respond accordingly by being functional or inactivated, thus maintaining their performance only under permissive conditions [16]. Specifically, kill switches are circuit systems that can result in cell death by activating lethal genes under certain conditions [17]. Addressing challenges such as high degrees of efficiency and robustness becomes imperative when building such circuits in order to eliminate all the targeted population. Additionally, the kill switch's response needs to be quick in order to prevent instabilities and other detrimental consequences of a delayed action [18]. Popular among

the lethal genes used in kill switches we encounter toxins [16, 17, 19], proteases [16] and CRISPR nucleases, such as Cas9 [20]. The endonuclease Cas9 creates double strand breaks (DSB) in the bacterial chromosome triggering cell death [21, 22]. However, some bacterial cells can survive such DNA damage since some positions in the genome are targeted more efficiently than others [23], a limitation that could be overcome by targeting more than one *locus* simultaneously. To avoid the design and utilization of different CRISPR independent targets, the genotoxicity of Cas9 could be improved by targeting repetitive regions of the genome [24].

Repetitive extragenic palindromic (REP) elements are highly conserved inverted repeats found in the genomes of some bacterial species [25], including *P. putida* KT2440 [26]. This strain has > 900 REPs with a highly conserved 35 bp sequence typically occurring as single units or pairs. REPs are also the target for some insertion sequences (IS) which are small, mobile genetic elements carrying information for their own transposition [27]. Particularly, the transposase ISPpu9 is encoded seven times also in repetitive extragenic palindromic regions of the *P. putida* KT2440 genome [28]. As a consequence, by using only the REP and ISPpu9 sequences, hundreds of parallel DSBs could be generated by Cas9 as long as those sequences were positioned next to the required protospacer adjacent motif (PAM).

With the aim of developing a quick and effective CRISPR-Cas9-based kill switch, we took advantage of the minimal PAM (5'-NNG-3') of the ScCas9 variant [29]. This allowed the generation of the *P. putida* GenoMine strain by installing into its genome a CRISPR cassette containing only two spacers to target the repetitive REP and ISPpu9 regions. To tightly control the lethal cleavage of the GenoMine strain's chromosome, we subjected Cas9 to the control of robust genetic circuits that had previously shown outstanding performance in bacteria. These circuits include AND gates such as conventional riboregulators [30, 31] and toehold riboregulators [32], and an ON/OFF digitalizer module (a YES gate) [33]. The basic design principle of a kill switch, however, should be a NOT, a NOR or a NAND gate to activate the lethal genes when the external signal(s) commonly administered under permissive and controlled conditions stop(s) being perceived. Consequently, the Cas9 controlled by the circuits was replaced by two repressing elements, AcrIIA4 and TetR, whereas an independent Cas9 was expressed on a different module. Circuits were used first

to control the expression of the anti-CRISPR protein AcrIIA4 that inhibits the enzymatic activity of type II Cas9 proteins [34, 35]; and secondly, to control the expression of TetR, a transcriptional repressor of the pLtetO promoter [36] which was in turn accommodated to control the independent Cas9.

Results

ScCas9 effectively targets repetitive genomic regions in *P. putida* starting a chain reaction of chromosome cleavages that result in cell death

The genome of *P. putida* KT2440 contains more than 900 REPs and 7 ISPpu9 sites [28]. Based on their highly conserved sequences, we designed two corresponding spacers of 30 nucleotides (Supplementary Table S4). Considering that ScCas9 requires an 5'-NNG-3' PAM located downstream of the protospacer or targeted region, these two spacers could theoretically target the genome at 27 different *loci* that shared 100 % sequence identity. However, it is estimated that the cRNA:tracrRNA:Cas9 might be able to target a much higher number of *loci*, due to the fact that the Cas9 complex can target sequences that differ by one or two nucleotides from the spacer, as long as these occur out of the seed sequence (the 8-10 nucleotides immediately adjacent to the PAM) [37] (Figure 4.1 A). Subsequently, we generated the *P. putida* GenoMine strain by introducing a constitutive CRISPR cassette containing both the REP and ISPpu9 spacers and the sequence for the tracrRNA (Figure 4.1 B), downstream the *PP_5322* gene to achieve high levels of expression [38].

To test the genotoxicity of ScCas9 in the GenoMine strain (Figure 4.1 C), we transformed it with the plasmid pSEVAb62-ScCas9, in which the endonuclease was constitutively expressed. As a result, a targeting efficiency of 99.95 ± 0.03 % was achieved when compared with the wild type strain transformed with the same plasmid (Figure 4.1 D). As a negative control, an empty pSEVAb62 plasmid was used to transform both the GenoMine and the wild type. In this case, growth of the GenoMine strain was not limited compared to the wild type strain, indicating that the former without Cas9 activity does not show growth limitations compared to the latter.

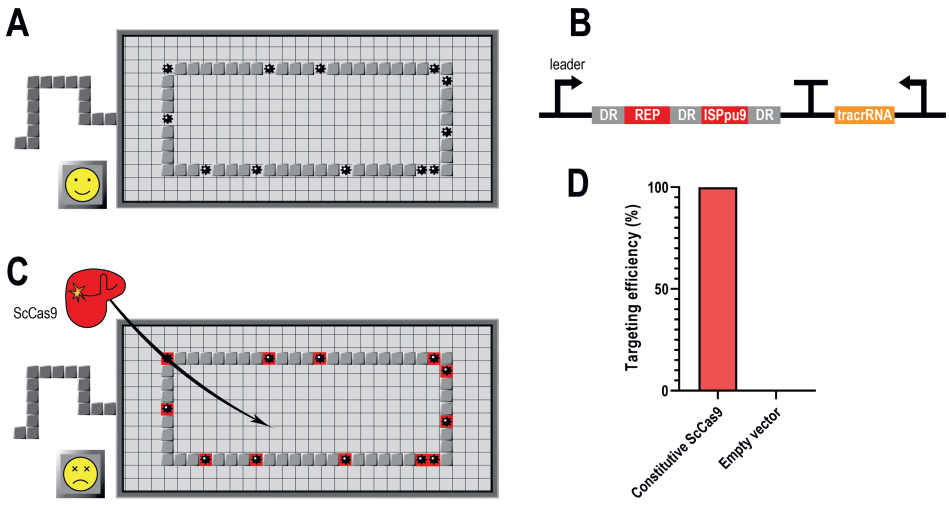


Figure 4.1: Scheme of GenoMine, a CRISPR-Cas9-based kill switch in P. putida. (A) The genome of P. putida KT2440 contains abundant repetitive sequences. These repetitive regions can easily be used as CRISPR-Cas9 targets leading to a chain reaction of cleavages in the bacterial chromosome. (B) GenoMine cassette introduced downstream the PP_5322 locus of the genome of P. putida KT2440 containing a CRISPR array with two constitutively expressed spacers (REP and ISPpu9) that target repetitive regions of the genome itself. DR stands for direct repeat. (C) When ScCas9 is transformed into cells containing the GenoMine cassette, a chain reaction of cleavages starts resulting in cell death. (D) Experimental validation of the hypothesis depicting the targeting efficiency associated with cell death of P. putida GenoMine relative to that obtained in the wild type strain when transformed with a pSEVAb62 plasmid constitutively expressing sccas9 or with an empty plasmid (mean ± s.d., n = 3 biological).

Utilization of different genetic circuits to control gene expression in *P. putida*

In order to control the triggering of the lethal cleavages, our efforts focused on finding and testing systems that would react to external signals to activate gene expression in *P. putida*. Our first approach involved the use of engineered riboregulators (RR) which operate at post-transcriptional level through small synthetic RNAs that enable gene silencing and activation. Specifically, we selected the riboregulators #12 and #10 (RR12 and RR10) from Isaacs *et al.* [30] due to their effectiveness in *E. coli* (Supplementary Table S5). Their structure consists of (i) a complementary cis-sequence which binds to the ribosomal binding site (RBS) at the 5'- untranslated region (UTR) of a given gene, and consequently forms a stem-loop structure, silencing gene expression by blocking the RBS, and (ii) a small non-coding trans-activating RNA (taRNA) that acts as counterpart and targets the cis-

repressing RNA (crRNA) with high affinity, enabling a conformation change which frees the RBS for the binding of the translation machinery (Figure 4.2 A) [30]. These two molecules were cloned into a pSEVAb23 plasmid under the control of the xylS/pM and rhaRS/pRham expression systems, induced by 3-methylbenzoate (3MB) and L-rhamnose, respectively, to regulate the expression of a reporter *gfp* gene. In the first version of our system (V1), xylS/pM was controlling the crRNA-gfp molecule whereas rhaRS/pRham was controlling the taRNA. In the second version (V2), the inducible expression systems were interchanged to evaluate the effect and potential leakiness of the promoters.

Fluorescence assays revealed the correct silencing and activation of *gfp* under the corresponding non-induced and double induction (3MB + rhamnose) conditions using the two versions of both riboregulators (Figure 4.2 C). Regarding RR12, very low levels of fluorescence were observed in both versions under non-induced conditions. Conversely, when both inducers were supplied, GFP expression levels increased dramatically, with version 2 showing a more stable and longer expression than version 1. When using RR10, fluorescence levels were insignificant under non-induced conditions. Upon double induction, RR10 showed similar profiles in both versions, albeit expression levels of version 1 were higher. When compared to RR12, both versions showed higher expression with RR10.

With respect to the individual induction with 3MB or rhamnose, the outcomes were similar when using either RR12 or RR10. Results showed no expression (version 2) or very low GFP levels (version 1) when only the taRNA was induced. Conversely, the inducer of the crRNA molecule alone was able to yield significant levels of GFP expression (3MB-induction for version 1 and rhamnose-induction for version 2) indicating either leaky expression of the promoters [5] or improper functioning of the crRNA molecule (Figure 4.2 C). The first hypothesis was quickly discarded based on the results produced by the different configurations, as it would be unlikely that both the pRham and the pM promoters would exhibit the same level of leakiness. The second hypothesis regarding the functioning of the crRNA molecule was experimentally investigated with pSEVAb23 plasmids containing the crRNA-gfp module but lacking the taRNA part.

GFP expression was observed with the crRNA part of both RR12 and RR10 and

with their corresponding two versions despite the absence of taRNA within the system. In this case, relative fluorescence levels with the crRNA inducer only and with both inducers were very similar and consistent with those of the previous assays when only the inducer for the crRNA was provided (Figure 4.2 C). These results showed that *gfp* translation was still possible even without taRNA, indicating that the cis-repressor sequence did not manage to block completely the access of the ribosome to the *gfp*'s RBS.

In an attempt to solve this issue and move away from relying on the RBS sequence, we tested the toehold riboregulator approach. The principle of a toehold switch is based on the design of two RNA sequences flanking the start codon that match and form a hairpin with another complementary RNA sequence situated before the RBS, which is left completely unpaired in this configuration (Figure 4.2 B). Because neither the RBS nor the start codon are necessary for hairpin formation, the crRNA can have any arbitrary sequence instead of a limited amount of them, which allows greater riboregulator variability. In this case, the reconfigured crRNA sequence is called switch RNA, whereas the equivalent of the former taRNA receives the name of trigger RNA [32]. For our study, we selected the toehold switch #1 of the second- or forward engineering- generation (#2.1) of Green *et al.* [32] which had previously shown the highest ON/OFF GFP fluorescence ratio (665 ± 135) in their study using *E. coli* (Supplementary Table S5). Toehold riboregulator #2.1 (Toehold 2.1) was cloned into a pSEVAb23 vector controlling a *gfp* reporter gene and, this time, only version 1, meaning switchRNA-gfp under the control of xylS/pM and triggerRNA controlled by rhaRS/pRham, was generated. In this case, the fluorescence assay using the Toehold 2.1 reported a correct functioning of the circuit under non-induced and double induction conditions. Unfortunately, undesired GFP expression levels were observed when only the switchRNA inducer was provided, as it had previously happened with RR12 and RR10 (Figure 4.2 C). However, these one-inducer activation levels were considerably lower than those obtained with the conventional riboregulators in the same conditions (Figure 4.2 C).

With the objective of minimizing the levels of gene expression under non-induced conditions, our third approach was the use of an ON/OFF digitalizer module (DM) that supresses the basal level of promoters entirely, impeding gene expression

in the absence of induction. The rationale of this circuit is based on the interplay of a translation-inhibitory small RNA controlled by pLac with the translational coupling of the gene of interest to the repressor LacI (Figure 4.2 D) [33]. Instead of

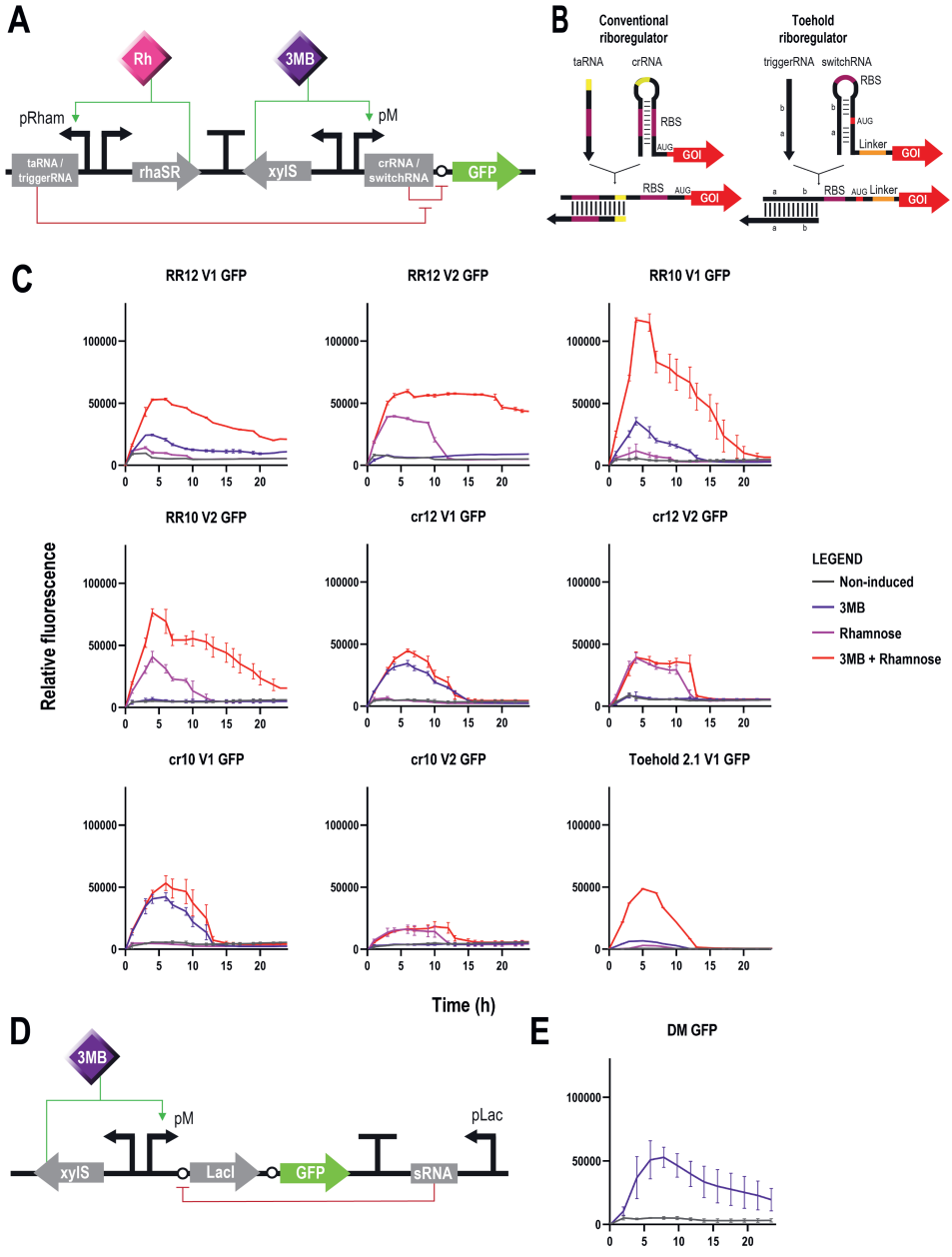


Figure 4.2: Different circuits controlling GFP expression in P. putida. (A) Schematic representation of a riboregulator circuit. Two different modules are cloned in opposite directions and separated by a transcriptional terminator: on the left, the trans-activating RNA (conventional riboregulator) or trigger RNA (toehold) molecule is controlled by the rhaRS/pRham expression system; and on the right, the cis-repressing RNA (conventional riboregulator) or switch RNA (toehold) is transcriptionally coupled to a gfp gene under the control of the xylS/pM expression system (Promoter version 1). The transcription of the aforementioned modules is induced by rhamnose and 3-methylbenzoate, respectively. Translation of gfp is blocked by the crRNA/switchRNA in absence of the taRNA/triggerRNA, but in its presence, the binding of the two RNA molecules results in a conformational change that allows the ribosome to access the RBS for gfp translation. (B) Structural differences between conventional and toehold riboregulators. While the former represses translation by base pairing directly to the RBS region (pink), the latter does it through base pairs programmed before and after the AUG start codon (red), leaving this and the RBS (pink) completely unpaired. (C) Fluorescence assays of the different riboregulator variants controlling the expression of GFP depicting relative fluorescence over time for 24 h. From top left to bottom right: RR12 version 1, RR12 version 2, RR10 version 1, RR10 version 2, crRNA 12 version 1, crRNA 12 version 2, crRNA 10 version 1, crRNA 10 version 2 and Toehold 2.1 version 1. Grey lines represent non-induced conditions, purple lines induction only with 3MB, pink lines only with rhamnose and red lines double induction with both 3MB and rhamnose (mean ± s.d., $n \geq 2$ biological and 3 technical replicates). (D) Schematic representation of the ON/OFF digitalizer module (DM). Expression of gfp is controlled by the xylS/pM expression system and an additional mutual inhibition circuit that regulates the translation step. This switch-like inhibition circuit includes a small RNA that inhibits the translation of the transcriptionally coupled LacI, a repressor of the pLac promoter that in turn controls the small RNA production. Activation of the system occurs only under 3MB induction, whereas in its absence, the small RNA interferes with any possible basal translation levels. (E) Fluorescence assay of a digitalized version of the xylS/pM system controlling the expression of GFP depicting relative fluorescence over time for 24 h (mean ± s.d., $n \geq 2$ biological and 3 technical replicates).

an AND gate like the riboregulators that theoretically require the presence of both inducers for the expression of a gene, the ON/OFF DM needs only a single input to accomplish a tight control of the system (YES gate). Even though it had already been tested in *P. putida* using the xylS/pM expression system and GFP as reporter [33], we repeated the experiment and obtained results consistent with previous research (Figure 4.2 E). Basal fluorescence levels were minimal while expression under 3MB-induced conditions were similar to those achieved with double induced RR12, version 1 (Figure 4.2 C).

Induced conditions activate ScCas9 resulting in cell death

To determine whether controlled induction of the different systems could result in cell death in the GenoMine strain, we replaced in all four circuits the reporter *gfp* with the *sccas9* gene which had previously been codon optimized for *P. putida*. For

convenience and given the scarce behavioural differences between the two configurations, only version 1 of the riboregulators was used in subsequent experiments (Figures 4.3 A). Plasmids pSEVAb23 containing these constructs (Figures 4.3 A and 4.3 B), alongside an empty pSEVAb23 vector were transformed in both wild type and GenoMine *P. putida* strains, which were subsequently plated on media supplemented with the different inducers to quantify the cell population survival to the CRISPR-Cas9 lethal response. We used for each strain- and plasmid-transformation the number of colony forming units (CFU) present in the non-induced plates as the

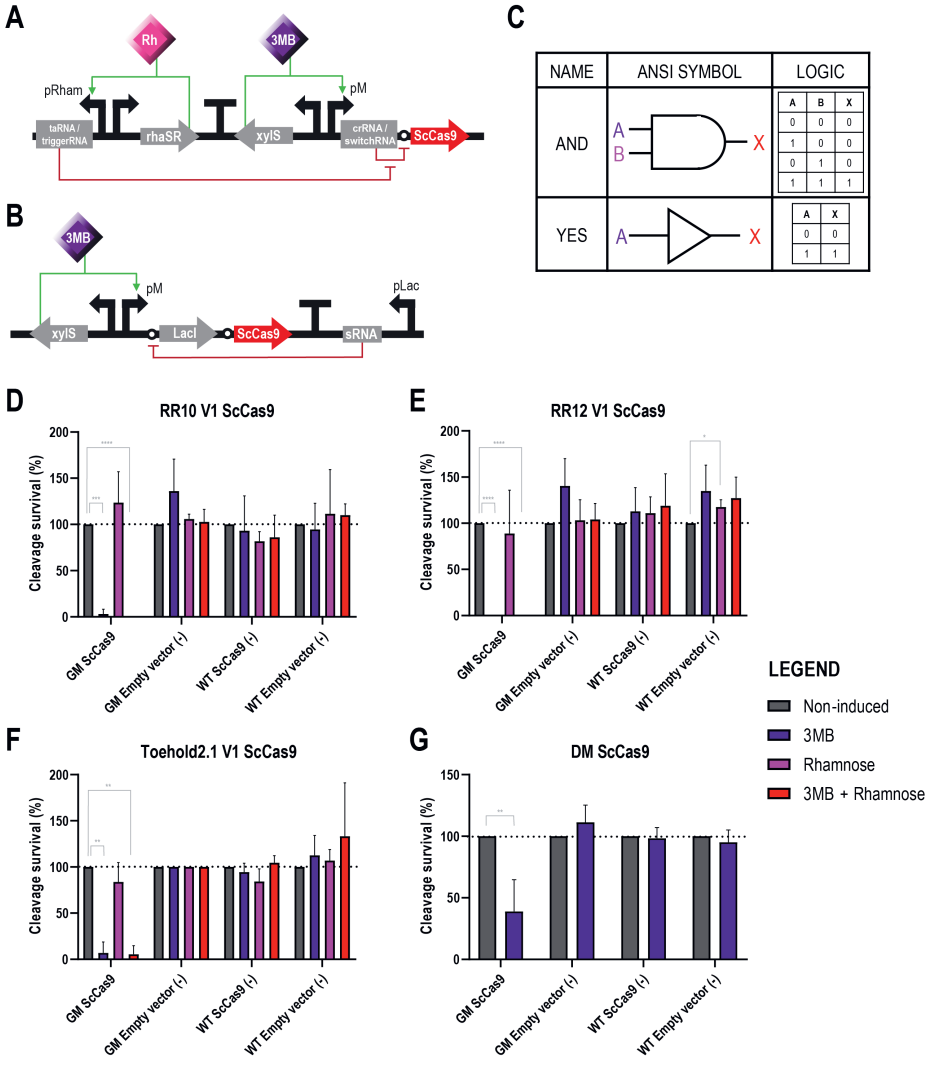


*Figure 4.3: Different circuits controlling the expression of ScCas9 in P. putida GenoMine. (A) Schematic representation of the riboregulator circuits (conventional or toehold) controlling the expression of ScCas9. (B) Schematic representation of the xylS/pM expression system endowed with the ON/OFF digitalizer module and controlling the expression of ScCas9. (C) Designs of AND and YES gates according to the American National Standards Institute (ANSI), in this case corresponding to the riboregulators and the ON/OFF DM, respectively. Following binary logic, there are only two states allowed, 1 and 0 or 'on and off'. A and B represent the inputs 3MB and rhamnose, and X represents the output of the circuit, in this case the activation of Cas9. (D) Cleavage survival after transformation of ScCas9 under the control of conventional RR10. (E) Cleavage survival after transformation of ScCas9 under the control of conventional RR12. (F) Cleavage survival after transformation of ScCas9 under the control of Toehold 2.1. (G) Cleavage survival after transformation of ScCas9 under the control of a digitalized version of the xylS/pM system. Relative cleavage survival was calculated separately for each strain in each figure as the ratio between CFU growing upon different induction conditions (purple bars represent single induction with 3MB, pink bars single induction with rhamnose and red bars double 3MB + rhamnose induction) and total CFU growing on plates without the addition of inducers (grey bars) and expressed in percentage. Only significant values are indicated for a parametric two-tailed t test between two groups, where *P < 0.05; **P < 0.01; ***P < 0.001; and ****P < 0.0001; non-significant values were not depicted. (Mean ± s.d., n ≥ 3 biological)*

reference of 100 % survival.

Upon full induction of the genetic circuits, which is 3MB + rhamnose for the riboregulators (AND gate) and 3MB only for the ON/OFF DM (YES gate) (Figure 4.3 C), the endonuclease should be expressed leading to the chromosome's cleavage of the strain carrying the GenoMine CRISPR cassette. Conversely, in the three control cases lacking any of the CRISPR-Cas9 machinery elements, namely the two transformations of the wild type and the transformation of the empty vector into the GenoMine strain, cleavages were not expected. Accordingly, the obtained number of CFU in the three negative controls was generally not affected by the different inducer conditions, apart from a single exception (WT strain transformed with an empty RR12 vector upon rhamnose induction, Figure 4.3 E) (Figures 4.3 D, 4.3 E, 4.3 F and 4.3 G).

On the other side, we had our test situation when the GenoMine strain was transformed with vectors carrying ScCas9. Upon fully induced conditions, there was a strong decrease in the percentage of surviving cells with all four circuits. When the endonuclease was under the control of the conventional riboregulators, the levels of lethality under fully induced conditions were absolute, with no CFU present in the plates. When using the toehold, 5.39 ± 9.33 % survived, whereas 38.97 ± 25.66 % of

the CFU managed to endure the cleavage with the DM (Figures 4.3 D, 4.3 E, 4.3 F and 4.3 G). In absence of 3MB, upon single induction of the taRNA/triggerRNA riboregulator counterparts with rhamnose, the cleavage survival was restored to levels that did not significantly differ from those of the non-induced conditions (Figures 4.3 D, 4.3 E and 4.3 F). However, as we had previously seen with the GFP, activation of ScCas9 still occurred upon single 3MB induction of the crRNA/switchRNA molecules. Percentages of cleavage survival dropped to 3.03 ± 5.24 %, 0 % and 6.86 ± 11.88 % when ScCas9 was controlled by RR10, RR12 or toehold 2.1, respectively (Figures 4.3 D, 4.3 E and 4.3 F).

While Figures 4.3 D, 4.3 E, 4.3 F and 4.3 G depict relative data calculated separately for each strain and transformation, we must remark that the absolute numbers in the obtained plates were variable (Supplementary Table S6). It should be noted that the absolute CFU count of non-induced GenoMine and wild type strains transformed with ScCas9 under the control of the different circuits was heavily impacted compared to the CFU count of GenoMine strains transformed with an empty vector (Supplementary Table S6). This indicates that there is either: (i) ScCas9-activity

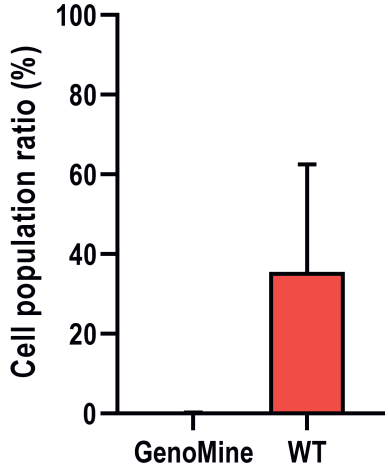


Figure 4.4: Effect of ScCas9 on P. putida's cell populations. Percentage of cell population was calculated as the ratio between CFU obtained when transformed with either pSEVAb62-ScCas9 (targeting conditions) or an empty pSEVAb62 (non-targeting conditions) for both the GenoMine and the wild type strains (mean ± s.d., $n \geq 3$ biological).

under non-induced conditions, which could be attributed to leakiness of the circuits that are in control of ScCas9 expression, or (ii) a significant burden or toxicity derived from hosting plasmids containing the endonuclease, even when this does not result in any chromosome's cleavage. To quantify this effect, we calculated the CFU difference between the wild type strain transformed with either the *sccas9*-expressing or the empty plasmids (Figure 4.4). As we had already seen in Figure 4.1 D, the lethal effect of ScCas9 was almost absolute in the GenoMine strain. However, in the wild type strain, the presence of ScCas9 also places a substantial burden that results in reduced cell populations, even without mediating any cleavage.

Controlling *P. putida* GenoMine through ScCas9 DNA-binding inhibition and transcriptional repression

In order to develop a kill-switch that is activated under non-induced conditions (*i.e.*, non-permissive conditions) instead of induced conditions, we attempted two different strategies: ScCas9's DNA-binding inhibition using the anti-CRISPR (Acr) protein AcrIIA4, and ScCas9's transcriptional repression using the repressing TetR/-pLtetO expression system. These new elements would work as inverter devices to reverse the circuits' outcome and logic, from AND and YES gates, in which ScCas9 expression was only possible in case of a full induction (Figure 4.3 C), to NAND and NOT gates, in which ScCas9 expression should happen under any other scenario that is not full induction (Figure 4.6 D).

In our first strategy, we tested if AcrIIA4 could effectively inhibit ScCas9, as had been previously reported for SpyCas9 [34, 35]. For this, we first transformed the GenoMine strain with pSEVAb23-AcrIIA4, in which the anti-CRISPR protein was constitutively expressed. Secondly, pSEVAb62-ScCas9 was also introduced to evaluate AcrIIA4's efficacy to prevent the action of ScCas9. The outcome of this experiment showed however no effect of AcrIIA4. Either in the presence or in the absence of this inhibitor, ScCas9 was equally able to target 99.73 ± 0.25 and 99.74 ± 0.25 % of the respective cell populations, when compared with those transformed with an empty pSEVAb62 vector (Figure 4.5). As a consequence, we decided not to continue with this approach and attempted an alternative way to repress ScCas9's activity.

In our second strategy, we substituted the *sccas9* gene for the one encoding

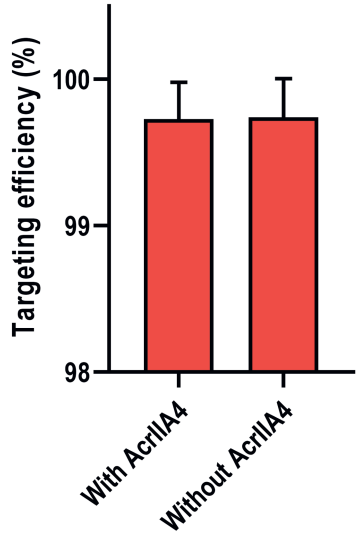


Figure 4.5: AcrIIA4 does not inhibit ScCas9's cleavage ability in the GenoMine strain. Targeting efficiency was calculated in P. putida GenoMine using the ratio of CFU obtained when transformed with either pSEVAb62-ScCas9 (targeting conditions) or an empty pSEVAb62 (non-targeting conditions), both in the presence and in the absence of a constitutively expressed AcrIIA4 (mean ± s.d., n = 2 biological).

the transcriptional repressor TetR in all four circuits (Figures 4.6 A and 4.6 B). In turn, *sccas9* was expressed on a separate plasmid under the control of the pLtetO promoter, whose activity can be repressed by binding of TetR [36] (Figure 4.6 C). Plasmids pSEVAb23 carrying the TetR circuit constructs were first transformed into the wild type and GenoMine strains as we had done with the ScCas9 circuit constructs in the previous assay. This time, however, a second transformation using either pSEVAb62-pLtetO-ScCas9 or an empty pSEVAb62 vector was performed, followed by an evaluation of the cleavage survival under the different inducer combinations. The number of CFU accounted under non-induced conditions represented a 100 % survival for each strain- and plasmid-transformation. It is important to consider that this survival, however, is the result of an activated ScCas9 with the power of killing the GenoMine cells. Upon induced conditions, TetR should be activated repressing ScCas9's transcription, resulting in a larger number of CFU and therefore cleavage survival levels higher than 100 %.

Once again, in the negative controls with an incomplete CRISPR-Cas9 system (wild type strains and GenoMine transformed with the empty vector), the addition of the inducers did not significantly affect the cleavage survival except in a few cases (Figures 4.6 E, 4.6 F, 4.6 G and 4.6 H). This time, however, inducers had no significant effect on the cleavage survival of the test strain either (GenoMine transformed with ScCas9).

In this way, the GenoMine strain carrying ScCas9 did not show a significantly higher cleavage survival when the transcriptional repressor TetR was expressed under fully induced conditions of any of the circuits. Even though the highest increases in average cleavage survival were observed upon induction of the riboregulators in this strain, the large standard deviations made the results not significantly different from the non-induced conditions (Figures 4.6 E, 4.6 F and 4.6 G). Single induction of the crRNA/switchRNA molecules resulted again in activation of the riboregulator systems, also boosting the average cleavage survival although in a non-significant manner (Figures 4.6 E, 4.6 F and 4.6 G). In addition, single induction of the taRNA when using RR12 reported a significant cleavage survival that reached 212.9 ± 34.04 % (Figure 4.6 E). Lastly, when assessing TetR under the control of the digitalized version of the xylS/pM expression system, induced conditions did not contribute to increases in cleavage survival at all (Figure 4.6 H).

Relative data illustrated in Figure 4.6 was calculated separately for each strain and transformation. As aforementioned for the cleavage assays depicted in Figure 4.3, the absolute CFU numbers in the obtained plates were variable (Supplementary Table S6), with a dramatic reduction of CFU in strains transformed with the *sccas9*-expressing plasmid.

Discussion

This study describes the generation of a CRISPR-Cas9-mediated biocontainment strategy in *P. putida* and details direct and indirect mechanisms of transcriptional, post-transcriptional and post-translational control over the system. After knocking in a genotoxic CRISPR array that targets repetitive genomic elements, cell viability becomes challenged on the presence of a Cas9 protein. In this strain, that we named

GenoMine, the powerful lethal effect arises from the large collection of *loci* spread across the chromosome susceptible to the targeted cleavage (Figure 4.1). To trigger this chain reaction, Cas9 was configurated by means of kill switches that included some genetic regulatory elements never used before in *P. putida*.

When using engineered riboregulators, we found that conformational changes in RNA structures and corresponding intermolecular crRNA-taRNA and switchRNA-triggerRNA interactions resulted in the expected RBS exposition that unleashes translation (Figure 4.2 B and 4.2 C) [30, 32]. However, flawed or incomplete repressive

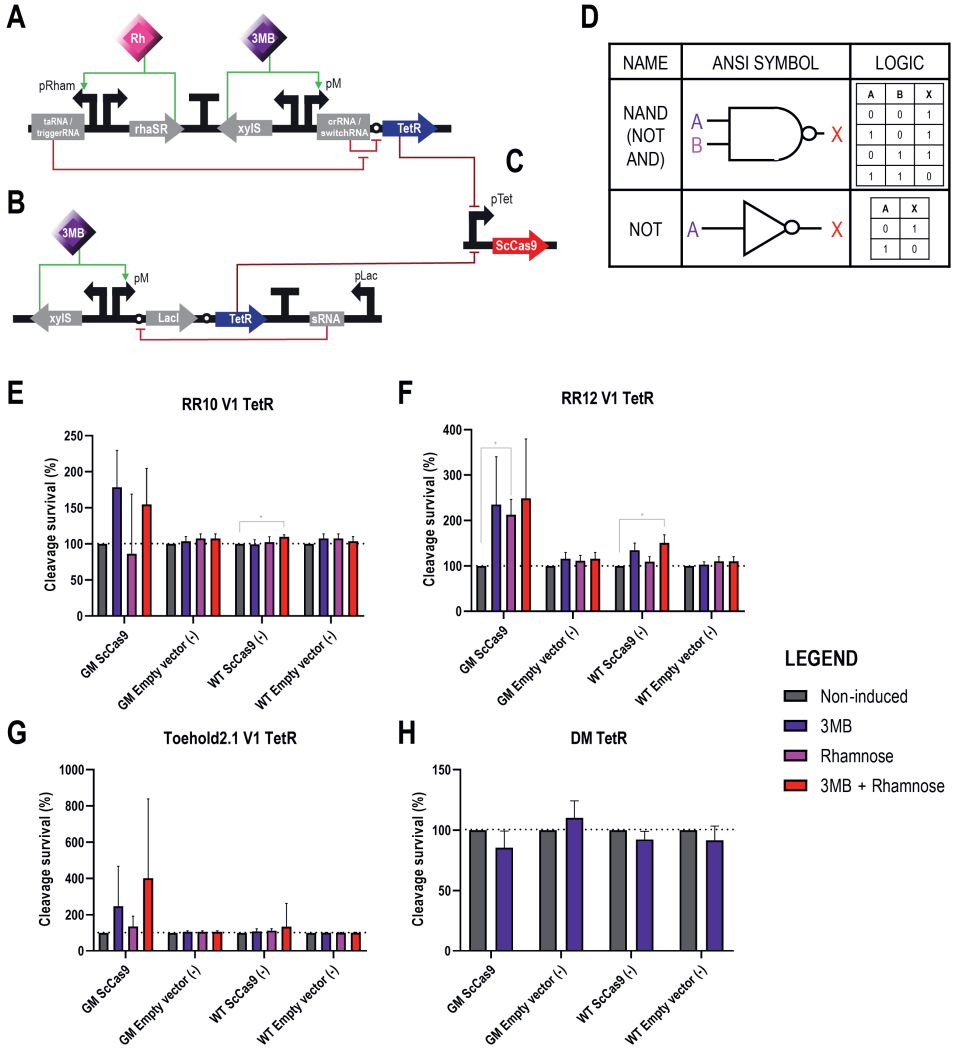


*Figure 4.6: Different circuits controlling the expression of TetR, which works as a transcriptional repressor of ScCas9 in P. putida GenoMine. (A) Schematic representation of the riboregulator circuits (conventional or toehold) controlling the expression of TetR. (B) Schematic representation of the xylS/pM expression system endowed with the ON/OFF digitalizer module and controlling the expression of TetR. (C) ScCas9 located on a separated plasmid under the control of the TetR/pLtetO expression system. ScCas9 can be repressed in the presence of TetR, which in turn gets expressed upon full induction of the genetic circuits depicted in either A or B. (D) Designs of NAND (NOT + AND) and NOT gates according to the American National Standards Institute (ANSI), in this case corresponding to the riboregulators and the ON/OFF DM, respectively. Following binary logic, there are only two states allowed, 1 and 0 or 'on and off'. A and B represent the inputs 3MB and rhamnose, and X represents the output of the circuit, in this case the activation of TetR and, correspondingly, the repression of ScCas9. (E) Cleavage survival of cells bearing TetR under the control of RR10 after transformation with pLtetO-ScCas9. (F) Cleavage survival of cells bearing TetR under the control of conventional RR12 after transformation with pLtetO-ScCas9. (G) Cleavage survival of cells bearing TetR under the control of Toehold 2.1 after transformation with pLtetO-ScCas9. (H) Cleavage survival of cells bearing TetR under the control of a digitalized version of the xylS/pM expression system after transformation with pLtetO-ScCas9. Relative cleavage survival was calculated separately for each strain in each figure as the ratio between CFU growing upon different induction conditions (purple bars represent single induction with 3MB, pink bars single induction with rhamnose and red bars double 3MB + rhamnose induction) and total CFU growing on plates without the addition of inducers (grey bars) and expressed in percentage. Only significant values are indicated for a parametric two-tailed t test between two groups, where *P < 0.05; **P < 0.01; ***P < 0.001; and ****P < 0.0001; non-significant values were not depicted. (Mean ± s.d., n = 3 biological)*

ability of the 5'-UTR cis elements of the crRNA and switchRNA molecules allowed undesired access of the ribosome to the RBS and subsequent post-transcriptional gene expression (Figure 4.2 C). The cis RNA molecules of the different riboregulators used in this study (RR10, RR12 and Toehold 2.1) had previously shown silencing levels > 98 % derived from translational repression [30, 32]. These experiments were performed *in vivo* at intermediate transcription rate in *E. coli*. While some silencing effect was exhibited, the leaky levels of gene expression that could be observed with the fluorescence data (Figure 4.2 C) were high enough to obtain significantly similar outcomes when inducing only the cis RNA molecule, or both the cis and trans modules of the kill switches in the GenoMine strain (Figures 4.3 and 4.6).

This behavioural difference of the cis elements of both conventional and toehold riboregulators between *E. coli* and *P. putida* and the common outcome reached by all three choices applied in this study (GFP, ScCas9 and TetR) suggest two things: (i) some degree of independence from the specific nucleotide sequences, and conversely, (ii) an influence of host-specific factors on crRNA and switchRNA loop con-

formation and stability as the cause of this phenomenon. Because the intermolecular RNA interactions rely on specific RNA structures, the utilized cis RNA sequences were initially generated based on appropriate secondary structure predictions [30, 32]. While temperature is a function of the equilibrium constant ΔGcrR-crR and the optimal growth temperature is different for the two organisms in question, our Mfold analysis [39] revealed an even lower value of ΔGcrR-crR at 30°C than at 37°C, indicating that these structures should be even more stable at *P. putida*'s growth conditions than at *E. coli*'s. On another note, host specific small RNAs [40, 41] or even RNA binding proteins that typically cause translational inhibition such as Crc [42, 43] and Hfq [44] might interfere with the hybridization loop of the crRNA and switchRNA molecules. Specifically, these proteins could have bound to the A-rich sequence upstream the cis elements of the riboregulators in the used vectors, since it has been shown that they bind to this type of motifs in *Pseudomonas* [45, 46]. Even though the issues seem to originate from other intrinsic factors, further sequence- and structure-based efforts to characterize novel small RNAs [47–49] in *P. putida* could help identifying cis repressing sequences that provide a robust and complete suppression of post-transcriptional expression.

Without the proper functioning of the cis repressing elements, the logic of the AND and NAND gates of the engineered riboregulators evolved into an ON/OFF mechanism (YES and NOT gates, respectively), similar to that of conventional inducible promoters. From the GenoMine strategy standpoint, this would not need to negatively affect their mission, as long as they would be able to offer a close to ON/OFF response when transitioning from non-induced to induced conditions. Key considerations in this regard would be: (i) high induced expression levels, and (ii) very low basal expression when no inducer is provided. These two properties were also accredited to the digitalizer module, which, unlike the riboregulators, had previously been evaluated in *P. putida* [33].

When controlling ScCas9, riboregulators showed high induced expression levels resulting in the desired complete (RR10 and RR12) or almost complete (Toehold 2.1) elimination of the cell population (Figures 4.3 D, 4.3 E and 4.3 F). Conversely, while the GFP expression produced under the DM's control (Figure 4.3 E) seemed in accordance with prior results, the ScCas9 levels yielded under induced conditions of

the digitalized xylS/pM were only able to kill off 61 % of the GenoMine cell population (Figure 4.3 G). Things get more complicated, however, when considering the basal expression when no inducer is provided. Even though the OFF state of the DM had previously shown zero transcription and the GFP assays depicted in Figure 4.2 C reported very low expression levels when the riboregulators were not induced, ScCas9 had a strong effect on the population in absence of inducers. Since this influence was also observed in the strain lacking the CRISPR module (Figure 4.4), we hypothesized that it occurs as a result of plasmid burden and ScCas9's own toxicity [50, 51]. Additionally, this also assures than any level of leakiness no matter how small, can be very detrimental to the cell population.

Moving on to the control over TetR, the basal expression in absence of inducers stopped being an element of concern since leaky expression of the repressor would not significantly harm the functioning of the kill switches. Instead, the high induced expression levels became even more important given the crucial task that the repressors had of blocking the action of every ScCas9 protein. Unfortunately, the induced levels of TetR by means of any of the circuits were not sufficient to inhibit the pLtetO-mediated transcription of ScCas9 (Figure 4.6 E, 4.6 F, 4.6 G and 4.6 H).

After the unsuccessful attempt to control ScCas9's response by inhibiting its DNA-binding with AcrIIA4 and by repressing its transcription with TetR, we are in need of more powerful repressing systems. This concerns a set of elements that is quite scarce in *P. putida* [4], but that is strongly needed for the well-functioning of the GenoMine strategy. When we attempted the use of anti-CRISPR, we hypothesized that that AcrIIA4 would effectively inhibit ScCas9, based on the effective inhibition of SpyCas9 by AcrIIA4 [34, 35] and the fact that ScCas9 and SpCas9 share a sequence similarity of 89.2 %. However, the fact that the PAM-specificity of ScCas9 is different from SpCas9, and that AcrIIA4 acts preventing PAM-recognition, might be the factor that negatively influenced interactions between ScCas9 and AcrIIA4 (Figure 4.5). Regarding the transcriptional repression, we believe that the levels of TetR that were generated upon induction were simply not enough to completely inhibit ScCas9's transcription, which, as we have explained, is able to significantly reduce cell population at very low expression levels. Consequently, alternative re-

pressors with capabilities to work as genetic inverters by totally inhibiting ScCas9 need to be investigated to be placed under the control of genetic devices.

While heterologous circuits are typically designed and assumed to be orthogonal (no direct genetic crosstalk interaction) to the host cell genetic background [52, 53], the performance of the different genetic networks used in this study shows otherwise. Most of the presumed orthogonal components and circuits have not been experimentally tested for their effects on different hosts' genetic machinery [54]. The original configuration of the riboregulators included the lacI/pLaciQ and tetR/pLtetO inducible expression systems for the transcriptional control of the RNA molecules [31]. However, these two systems had shown erratic behaviour in previous *P. putida* studies [4]: an excess of leakiness in the case of the lacI/pLaciQ system, and an inconsistent ability of anhydrotetracycline (aTc) to inhibit TetR repression over pLtetO. The latter fact prevented us from properly using the whole (aTc)/tetR/pLtetO system as an inducible promoter, but still allowed us to use TetR as a repressor. Overall, these reasons motivated us to substitute them by the well-established xylS/pM and rhaRS/pRham. In addition, the xylS/pM system endowed with the ON/OFF digitalizer device had performed effectively when controlling the toxic antibacterial colicin E3 in *P. putida* [33]. These circumstances evince, as many other investigations before [52, 55–57], that new configurations and bacterial hosts in the former case, and new controlled genes in the latter, can affect the final outcome of heterologous gene networks. Additionally, previous research has shown that circuit plasmid copy number plays an important role in host gene expression and dramatically affects orthogonality, burden and functionality of heterologous circuits [54]. This key factor did not receive special attention during the course of this study since we consistently used medium copy number plasmids for the expression of the circuits. In the future, more attention should be paid to achieving a balanced system behaviour by considering multiple compatible plasmids with different copy numbers to address the requirements of a given circuit in the *P. putida* GenoMine strain [52, 58].

Apart from the ones contemplated in this study, other regulatory elements and genetic circuits could be considered to modulate the genotoxic response of the GenoMine strategy. Especially attractive would be the aforementioned exploration

of gene expression repressing elements, and new mechanisms for inducible control over CRISPR-Cas systems [59–61], given that the key element for the well-functioning of GenoMine is precisely the tight control of the cleavage performed by Cas9.

Once its elements are properly optimized, the GenoMine kill switch would pose an excellent strategy for biocontainment studies in *P. putida* since it brings together an effective killing mechanism and a modular sensing part that allows the characterization of new genetic regulatory elements and circuits in this bacterium. While the general challenges of efficiency, stability and robustness typically associated to genetic circuits [18, 62] keep being present, we aimed at accelerating the circuit's response by using two different spacers with the ability of targeting multiple genomic *loci*. This is expected to reduce the time window for mutations and evolution that might end up affecting the targeting efficiency of Cas9. In addition, the GenoMine strain could be considered safe in a broader sense since, once optimized, it would have the ability to prevent any further genomic alteration made with CRISPR-Cas9-based editing technologies. Thus, this host might eventually be used as an intellectual property strain to prevent industrial espionage [63, 64], to biologically store sensible information, and to use in any situation in which further modifications are undesired since it offers protection against any editing that involves Cas9. These industrially attractive features built upon the value of biosafety would contribute to the realization of a more thorough, and ultimately more legitimized, safe-by-design approach in which the properties of the biocontainment strategy itself can also be profitable in other aspects of the industrial process or final application context.

Methods

Bacterial strains and media

All bacterial strains utilized in this study with their respective characteristics can be found in Supplementary Table S1. Standard cultivation was done at the strains' optimal temperature, 37°C for *E. coli* and 30°C for *P. putida*, in Lysogeny Broth (LB) (10 g/L NaCl, 10 g/L tryptone, and 5 g/L yeast extract) supplemented with appro-

priate antibiotics according to the following concentrations: kanamycin, 50 mg/L; gentamicin, 10 mg/L; chloramphenicol, 20 mg/L. When required, relevant inducers were as well supplemented into the medium: 3-methylbenzoate at 1 mM and rhamnose at 3.75 mM, if not stated otherwise. *E. coli* cells were made chemically competent as previously described [65] and used for cloning purposes. Electrocompetent *P. putida* strains were prepared following previously described standard protocols [66] and used for cleavage assays and circuit experiments.

Fluorescence assays were performed on M9 minimal medium (1.63 g/L NaH_2PO_4, 3.88 g/L K_2HPO_4, 2 g/L $(NH_4)_2SO_4$, 10 mg/L EDTA, 100 mg/L $MgCl_2 \cdot 6\ H_2O$, 2 mg/L $ZnSO_4 \cdot 7\ H_2O$, 1 mg/L $CaCl_2 \cdot 2\ H_2O$, 5 mg/L $FeSO_4 \cdot 7\ H_2O$, 0.2 mg/L $Na_2MoO_4 \cdot 2\ H_2O$, 0.2 mg/L $CuSO_4 \cdot 5\ H_2O$, 0.4 mg/L $CoCl_2 \cdot 6\ H_2O$, and 1 mg/L $MnCl_2 \cdot 2\ H_2O$) supplemented with 50 mM of glucose.

Plasmids

Plasmids used in the present study are fully described in Supplementary Table S2. Primers to amplify the genetic parts (Supplementary Table S5) and to verify the products can be found in more detail in Supplementary Table S3. All genetic parts were amplified by PCR using the NEB Q5® High-Fidelity DNA polymerase, according to the company's instruction. PCR fragments were subjected to a 1 % (w/v) agarose gel electrophoresis and isolated using Nucleospin Gel and PCR Clean-up (BIOKÉ) kit. Plasmids were constructed using Golden Gate and the Golden Gate-based SevaBrick Assembly method [12], unless otherwise stated, and transformed by heat-shock in chemically competent *E. coli* DH5α cells. Isolation of the plasmids was done with the GeneJET Plasmid Miniprep Kit® (Thermo Scientific) and colony PCR was performed to verify the right assembly of the different fragments. Plasmid sequence was confirmed by Sanger sequencing from Macrogen (MACROGEN Inc. DNA Sequencing Service; Amsterdam, The Netherlands).

First, we generated a CRISPR-Cas9 vector pSEVAb62-ScCas9-tracrRNA-crRNA_REP_ISPpu9 for the construction of the appropriate GenoMine CRISPR array. Spacers REP and ISPpu9 (Supplementary Table S4) were introduced in pSEVAb62-ScCas9-crRNA_eforRed using the previously described protocol called "One-step Golden Gate-based cloning for the assembly of single and multiple

spacers into the crRNA cassette" [7] by replacing the eforRed chromoprotein. Once the array was built, it was PCR amplified with primers G3-G4 and cloned into a G1-G2 amplified pGNW backbone to knock it in the *PP_5322 locus* of the *P. putida* KT2440 genome. Plasmids containing the reporter *gfp* were used to evaluate how the different circuits control gene expression in *P. putida*. Plasmids pZE21 Y12 a12C GFP containing conventional riboregulator #12 and pSEVA238 D.M. GFP were kind gifts from the Isaacs Lab (Yale University) and the Molecular Environmental Microbiology Laboratory (CNB-CSIC), respectively. Riboregulator #12's cassette was amplified using primers R1-R2 and cloned into a R3-R4 amplified pSEVAb23 backbone. Primers R5-R6 were then used to substitute the original promoters pLaciQ and pLtetO by the xylS/pM (R7-R8) and rhaRS/pRham (R9-R10) expression systems, which were amplified from in-house vectors, to create pSEVAb23 RR12-GFP pM-pRha. Subsequently, linearized fragments lacking both the xylS/pM and the rhaRS/pRham expression systems were created using R15-R16 for riboregulator #12 and R17-R18 for riboregulator #10, which was generated by introducing the corresponding crRNA and taRNA differences with #12 in the amplification primers. Using P7 and P10, xylS/pM and rhaRS/pRham were cloned to create pSEVAb23 RR10-GFP pM-pRha, while the use of R19-R20 allowed the construction of both pSEVAb23 RR10-GFP pRha-pM and pSEVAb23 RR10-GFP pRha-pM. To build all the versions without taRNA, R22 and R21 were used as forward primers when the crRNA promoters were either pM or pRham, respectively, combined with the reverse primer R14, that eliminated the taRNA part. To obtain the last circuit, the toehold riboregulator #2.1, the pSEVAb GFP backbone was amplified with T2 and T4, whereas the part containing the xylS/pM and rhaRS/pRham expression systems of the riboregulator #12 version 1 was obtained using T1 and T3, yielding pSEVAb23 Toehold2.1-GFP pM-pRha. For the replacement of *gfp* with the genes encoding ScCas9, TetR and AcrIIA4, the backbones of pSEVAb23 RR12/10-GFP pM-pRha, pSEVAb23 Toehold2.1-GFP pM-pRha and pSEVA238 D.M. GFP were amplified with R24-R25, R24-T5 and D1-D2, respectively. ScCas9 was amplified from pSEVAb62-ScCas9-tracrRNA-crRNA_REP_ISPpu9 using primers S1-S2 (for the DM), S3-S4 (for both conventional riboregulators) and S5-S4 (for the Toehold2.1). The TetR fragment was obtained from an in-house pSEVAb65 pTet-GFP with primers

TR1-TR2 (for the DM), TR3-TR4 (for both conventional riboregulators) and TR5-TR4 (for the toehold2.1). Lastly, AcrIIA4 was amplified from an in-house pSB1C3-AcrIIA4 vector making use of A1-A2 (for the DM), A3-A4 (for both conventional riboregulators) and A5-A4 (for the Toehold2.1). The constitutive pSEVAb23-AcrIIA4 was generated by amplifying the backbone pSEVAb23 with the Anderson 100 promoter attached using primers R4 and R26 and by cloning into it the AcrIIA4 sequence amplified with R27-R28 from the pSB1C3-AcrIIA4. Finally, pSEVAb62-pTet-ScCas9 was made using primers S6-S7 which changed the native ScCas9 promoter present in pSEVAb62-ScCas9 for the pLtetO expression system.

Construction of *P. putida* GenoMine

P. putida GenoMine was generated by introducing downstream the *PP_5322* genomic *locus* [38] a CRISPR cassette containing two spacers that target highly repetitive regions of the *P. putida* KT2440 genome. The CRISPR cassette was introduced *via* I-SceI mediated homologous recombination [67]. In brief, the fluorescent suicide vector pGNW crRNA_REP_ISPpu9 PP_5322 KI, which carried the CRISPR cassette between two 500-bp homology regions upstream and downstream the exact introduction location, was first transformed by electroporation into *P. putida* KT2440 cells. Cointegrates, easily identified by the green fluorescence, were subjected to a second transformation with the standard pQURE6 which contains the inducible I-SceI meganuclease. Upon 3-methylbenzoate induction, a second crossover resulted in the excision of the cointegrated pGNW followed by non-fluorescent cells which needed to be scanned for revertant (*i.e.*, wild type) or mutant (*i.e.*, knock in) genotype.

Cleavage assays

Strains *P. putida* KT2440 and *P. putida* GenoMine, either harbouring or not circuit plasmids, were used in different types of cleavage assays with ScCas9 plasmids. First, to test the efficiency of the GenoMine strategy, electro-competent *P. putida* cells were transformed with 100 ng of either the constitutive pSEVAb62-ScCas9 or the empty pSEVAb62, and subsequently plated onto LB-Gen solid agar media after 1.5 h of antibiotic-free recovery. In this first case, targeting efficiency was calculated

by comparing the number of CFU present in the GenoMine plates (targeting) and the number of CFU present in the corresponding wild type plates (non-targeting).

Second, to measure the cleavage survival when ScCas9 was controlled by the different genetic circuits, electro-competent wild type and GenoMine cells were transformed with 100 ng of either pSEVAb23 RR10-ScCas9, pSEVAb23 RR12-ScCas9, pSEVAb23 Toehold2.1-ScCas9 or pSEVA238 D.M. ScCas9, alongside their non-targeting counterparts carrying GFP instead of ScCas9. After 1.5 h of recovery in absence of antibiotics and inducers, 100 μL of each transformed sample was plated on LB-Kan, LB-Kan-3MB, LB-Kan-Rham and LB-Kan-3MB-Rham in the case of the riboregulators, and onto LB-Kan and LB-Kan-3MB, in the case of the DM. Cleavage survival was calculated after 48 h as the ratio between CFU growing upon different induction conditions and total CFU growing on plates without the addition of inducers and expressed in percentage.

Third, to measure cleavage survival when the circuits were controlling the expression of AcrIIA4, strains KT2440 and GenoMine were first transformed with the AcrIIA4-containing plasmid pSEVAb23-AcrIIA4. For the subsequent transformation, cultures were grown overnight under all the different inducer-combinations and then OD600 corrected. Then, they were transformed with 100 ng of either the constitutive pSEVAb62-ScCas9 or the empty pSEVAb62 and, in this case, recovery of each sample was performed for 1.5 h in the presence of the same inducer in which it had been previously incubated. Subsequently, 100 μL of the samples was correspondingly plated onto LB-Kan-Gen.

Fourth, to also measure cleavage survival, but this time when the circuits were controlling the expression of the TetR repressor, both strains KT2440 and GenoMine had to be transformed in advance with the TetR-containing plasmids pSEVAb23 RR10-TetR, pSEVAb23 RR12-TetR, pSEVAb23 Toehold2.1-TetR and pSEVA238 D.M. TetR. All these different strains were subsequently transformed with 100 ng of either pSEVAb62-pTet-ScCas9 or the empty pSEVAb62 and incubated for 1.5 h of recovery in the absence of antibiotics but in presence of all the inducer combinations. 100 μL of the samples was then correspondingly plated onto LB-Kan-Gen, LB-Kan-Gen-3MB, LB-Kan-Gen-Rham or LB-Kan-Gen-3MB-Rham in the case of the riboregulators, and onto LB-Kan-Gen or LB-Kan-Gen-3MB in the case of the digitalizer module.

Cleavage survival was calculated after 48 h as the ratio between CFU growing upon different induction conditions and total CFU growing on plates without the addition of inducers and expressed in percentage. Every cleavage assay was performed at least in biological triplicates.

Fluorescence assays

P. putida KT2440 harboring either pSEVAb23 RR12-GFP pM-pRham, pSEVAb23 RR12-GFP pRham-pM, pSEVAb23 RR10-GFP pM-pRham, pSEVAb23 RR10-GFP pRham-pM, pSEVAb23 cr12-GFP pM-pRham, pSEVAb23 cr12-GFP pRham-pM, pSEVAb23 cr10-GFP pM-pRham, pSEVAb23 cr10-GFP pRham-pM, pSEVAb23 Toehold2.1-GFP pM-pRham or pSEVA238 D.M. GFP were grown at 30°C and 200 rpm, overnight in 10 mL LB-Kan. Overnight cells were harvested at 4700 g for 5 min and washed twice with minimal M9 medium to eliminate LB traces. Cells were then resuspended to an OD_{600} = 0.3 in fresh M9 medium supplemented with 50 mM of glucose and 50 mg/L kanamycin on appropriate 96-well plates for the measurement of both absorbance and fluorescence in a total volume of 200 μL per well. In addition, inducers were also provided when necessary, including L-rhamnose at 3.75 mM and 3-methylbenzoate at 2 mM. Optical density (OD_{600}) and fluorescence (excitation 467 nm, emission 508 nm) were monitored over 24 h using a BioTek Synergy Mx Multi-Mode Microplate reader. Relative fluorescence values were calculated by normalizing the values of fluorescence to the OD_{600} values. Two biological and three technical replicates were included.

Acknowledgements

We would like to thank Farren Isaacs and Belén Calles for providing the vectors pZE21 Y12 a12C GFP and pSEVA238 D.M. GFP. In addition, we wish to convey our gratitude to Karsten Pawellek, Ivar de Jong, Kesi Kurnia and Sanne van Hamond for their help with other parts related to this work. This work is part of the research programme SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis, which is supported by the Dutch Research Council (NWO) [15814].

Conflict of interest

The authors declare there are no conflicting interests.

Author contributions

Conceptualization: EAG/MMP; Data curation: EAG/CdB/MA/AM; Formal analysis: EAG/CdB/MA/AM; Funding acquisition: VAPMdS; Investigation: EAG/MMP/CdB/-MA/AM; Methodology: EAG/MMP/CdB/MA/AM; Resources: VAPMdS; Supervision: VAPMdS; Visualization: EAG; Writing – original draft: EAG; Writing – review & editing: EAG/MMP/CdB/AM/VAPMdS

Supplementary material

The supplementary material of this chapter, including Supplementary Tables S1 to S6, can be accessed *via*:

https://figshare.com/s/c817200d1d06bf2a6398

Bibliography

1. De Lorenzo, V., Krasnogor, N. & Schmidt, M. For the sake of the Bioeconomy: define what a Synthetic Biology Chassis is! *New Biotechnology* **60,** 44–51 (2021).
2. Nikel, P. I., Chavarría, M., Danchin, A., *et al.* From dirt to industrial applications: Pseudomonas putida as a Synthetic Biology chassis for hosting harsh biochemical reactions. *Current Opinion in Chemical Biology. Synthetic Biology * Synthetic Biomolecules* **34,** 20–29 (2016).
3. Nikel, P. I. & de Lorenzo, V. Pseudomonas putida as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism. *Metabolic Engineering. Metabolic Engineering Host Organism Special Issue* **50,** 142–155 (2018).
4. Martin-Pascual, M., Batianis, C., Bruinsma, L., *et al.* A navigation guide of synthetic biology tools for Pseudomonas putida. *Biotechnology Advances* **49,** 107732 (2021).
5. Calero, P., Jensen, S. I. & Nielsen, A. T. Broad-host-range ProUSER vectors enable fast characterization of inducible promoters and optimization of p-coumaric acid production in Pseudomonas putida KT2440. *ACS synthetic biology* **5,** 741–753 (2016).
6. Martínez-García, E. & de Lorenzo, V. Molecular tools and emerging strategies for deep genetic/genomic refactoring of Pseudomonas. *Current opinion in biotechnology* **47,** 120–132 (2017).
7. Batianis, C., Kozaeva, E., Damalas, S. G., *et al.* An expanded CRISPRi toolbox for tunable control of gene expression in Pseudomonas putida. *Microbial biotechnology* **13,** 368–385 (2020).
8. Zobel, S., Benedetti, I., Eisenbach, L., *et al.* Tn7-based device for calibrated heterologous gene expression in Pseudomonas putida. *ACS synthetic biology* **4,** 1341–1351 (2015).
9. Elmore, J. R., Furches, A., Wolff, G. N., *et al.* Development of a high efficiency integration system and promoter library for rapid modification of Pseudomonas putida KT2440. *Metabolic engineering communications* **5,** 1–8 (2017).
10. Amarelle, V., Sanches-Medeiros, A., Silva-Rocha, R., *et al.* Expanding the toolbox of broad host-range transcriptional terminators for Proteobacteria through metagenomics. *ACS synthetic biology* **8,** 647–654 (2019).
11. Aparicio, T., Nyerges, A., Martínez-García, E., *et al.* High-efficiency multi-site genomic editing of Pseudomonas putida through thermoinducible ssDNA recombineering. *Iscience* **23,** 100946 (2020).
12. Damalas, S. G., Batianis, C., Martin-Pascual, M., *et al.* SEVA 3.1: enabling interoperability of DNA assembly among the SEVA, BioBricks and Type IIS restriction enzyme standards. *Microbial biotechnology* **13,** 1793–1806 (2020).
13. Riglar, D. T. & Silver, P. A. Engineering bacteria for diagnostic and therapeutic applications. *Nature Reviews Microbiology* **16,** 214–225 (2018).
14. Xu, P., Li, L., Zhang, F., *et al.* Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proceedings of the National Academy of Sciences* **111,** 11299–11304 (2014).
15. Lo, T.-M., Chng, S. H., Teo, W. S., *et al.* A two-layer gene circuit for decoupling cell growth from metabolite production. *Cell systems* **3,** 133–143 (2016).
16. Chan, C. T. Y., Lee, J. W., Cameron, D. E., *et al.* 'Deadman' and 'Passcode' microbial kill switches for bacterial containment. *Nature Chemical Biology* **12,** 82–86 (2016).
17. Stirling, F., Bitzan, L., O'Keefe, S., *et al.* Rational design of evolutionarily stable microbial kill switches. *Molecular cell* **68,** 686–697. e3 (2017).
18. Xia, P.-F., Ling, H., Foo, J. L., *et al.* Synthetic genetic circuits for programmable biological functionalities. *Biotechnology Advances. Microbial Engineering Biotechnologies* **37,** 107393 (2019).

19. Piraner, D. I., Abedi, M. H., Moser, B. A., *et al.* Tunable thermal bioswitches for in vivo control of microbial therapeutics. *Nature chemical biology* **13,** 75–80 (2017).

20. Bikard, D. & Barrangou, R. Using CRISPR-Cas systems as antimicrobials. *Current opinion in Microbiology* **37,** 155–160 (2017).

21. Anders, C., Niewoehner, O., Duerst, A., *et al.* Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513,** 569–573 (2014).

22. Jinek, M., Jiang, F., Taylor, D. W., *et al.* Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science (New York, N.Y.)* **343,** 1247997 (2014).

23. Cui, L. & Bikard, D. Consequences of Cas9 cleavage in the chromosome of Escherichia coli. *Nucleic acids research* **44,** 4243–4251 (2016).

24. Castanon, O., Smith, C. J., Khoshakhlagh, P., *et al.* CRISPR-mediated biocontainment. *biorxiv* (2020).

25. Stern, M. J., Ames, G. F.-L., Smith, N. H., *et al.* Repetitive extragenic palindromic sequences: A major component of the bacterial genome. *Cell* **37,** 1015–1026 (1984).

26. Aranda-Olmedo, I., Tobes, R., Manzanera, M., *et al.* Species-specific repetitive extragenic palindromic (REP) sequences in Pseudomonas putida. *Nucleic Acids Research* **30,** 1826–1833 (2002).

27. Siguier, P., Gourbeyre, E., Varani, A., *et al.* Everyman's Guide to Bacterial Insertion Sequences. *Microbiology Spectrum* **3,** MDNA3–0030–2014 (2015).

28. Gómez-García, G., Ruiz-Enamorado, A., Yuste, L., *et al.* Expression of the ISPpu9 transposase of Pseudomonas putida KT2440 is regulated by two small RNAs and the secondary structure of the mRNA 5-untranslated region. *Nucleic Acids Research* **49,** 9211–9228 (2021).

29. Chatterjee, P., Jakimo, N. & Jacobson, J. M. Minimal PAM specificity of a highly similar SpCas9 ortholog. *Science Advances* **4,** eaau0766 (2018).

30. Isaacs, F. J., Dwyer, D. J., Ding, C., *et al.* Engineered riboregulators enable post-transcriptional control of gene expression. *Nature biotechnology* **22,** 841–847 (2004).

31. Gallagher, R. R., Patel, J. R., Interiano, A. L., *et al.* Multilayered genetic safeguards limit growth of microorganisms to defined environments. *Nucleic acids research* **43,** 1945–1954 (2015).

32. Green, A. A., Silver, P. A., Collins, J. J., *et al.* Toehold Switches: De-Novo-Designed Regulators of Gene Expression. *Cell* **159,** 925–939 (2014).

33. Calles, B., Goñi-Moreno, Á. & de Lorenzo, V. Digitalizing heterologous gene expression in Gram-negative bacteria with a portable ON/OFF module. *Molecular Systems Biology* **15,** e8777 (2019).

34. Dong, D., Guo, M., Wang, S., *et al.* Structural basis of CRISPR-SpyCas9 inhibition by an anti-CRISPR protein. *Nature* **546,** 436–439 (2017).

35. Yang, H. & Patel, D. J. Inhibition Mechanism of an Anti-CRISPR Suppressor AcrIIA4 Targeting SpyCas9. *Molecular cell* **67,** 117–127.e5 (2017).

36. Saenger, W., Orth, P., Kisker, C., *et al.* The Tetracycline Repressor—A Paradigm for a Biological Switch. *Angewandte Chemie International Edition* **39,** 2042–2052 (2000).

37. Jiang, W., Bikard, D., Cox, D., *et al.* RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature biotechnology* **31,** 233–239 (2013).

38. Chaves, J. E., Wilton, R., Gao, Y., *et al.* Evaluation of chromosomal insertion loci in the Pseudomonas putida KT2440 genome for predictable biosystems design. *Metabolic Engineering Communications* **11,** e00139 (2020).

39. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research* **31,** 3406–3415 (2003).

40. Prévost, K., Salvail, H., Desnoyers, G., *et al.* The small RNA RyhB activates the translation of shiA mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. *Molecular microbiology* **64,** 1260–1273 (2007).

41. Sonnleitner, E., Gonzalez, N., Sorger-Domenigg, T., *et al.* The small RNA PhrS stimulates synthesis of the Pseudomonas aeruginosa quinolone signal. *Molecular microbiology* **80,** 868–885 (2011).

42. Moreno, R., Ruiz-Manzano, A., Yuste, L., *et al.* The Pseudomonas putida Crc global regulator is an RNA binding protein that inhibits translation of the AlkS transcriptional regulator. *Molecular Microbiology* **64,** 665–675 (2007).

43. Moreno, R., Marzi, S., Romby, P., *et al.* The Crc global regulator binds to an unpaired A-rich motif at the Pseudomonas putida alkS mRNA coding sequence and inhibits translation initiation. *Nucleic acids research* **37,** 7678–7690 (2009).

44. Sonnleitner, E. & Bläsi, U. Regulation of Hfq by the RNA CrcZ in Pseudomonas aeruginosa Carbon Catabolite Repression. *PLoS Genetics* **10** (ed Garsin, D. A.) e1004440 (2014).

45. Moreno, R., Hernández-Arranz, S., La Rosa, R., *et al.* The Crc and Hfq proteins of Pseudomonas putida cooperate in catabolite repression and formation of ribonucleic acid complexes with specific target motifs. *Environmental Microbiology* **17,** 105–118 (2015).

46. Hernández-Arranz, S., Sánchez-Hevia, D., Rojo, F., *et al.* Effect of Crc and Hfq proteins on the transcription, processing, and stability of the Pseudomonas putida CrcZ sRNA. *RNA* **22,** 1902–1917 (2016).

47. Zadeh, J. N., Steenberg, C. D., Bois, J. S., *et al.* NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry* **32,** 170–173 (2011).

48. Apura, P., Saramago, M., Peregrina, A., *et al.* Tailor-made sRNAs: a plasmid tool to control the expression of target mRNAs in Pseudomonas putida. *Plasmid* **109,** 102503 (2020).

49. Pobre, V., Graça-Lopes, G., Saramago, M., *et al.* Prediction of novel non-coding RNAs relevant for the growth of Pseudomonas putida in a bioreactor. *Microbiology* **166,** 149–156 (2020).

50. Jakočiūnas, T., Jensen, M. K. & Keasling, J. D. CRISPR/Cas9 advances engineering of microbial cell factories. *Metabolic Engineering* **34,** 44–59 (2016).

51. Wu, G., Yan, Q., Jones, J. A., *et al.* Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends in Biotechnology* **34,** 652–664 (2016).

52. Wang, B., Kitney, R. I., Joly, N., *et al.* Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nature communications* **2,** 1–9 (2011).

53. Moon, T. S., Lou, C., Tamsir, A., *et al.* Genetic programs constructed from layered logic gates in single cells. *Nature* **491,** 249–253 (2012).

54. Liu, Q., Schumacher, J., Wan, X., *et al.* Orthogonality and Burdens of Heterologous AND Gate Gene Circuits in E. coli. *ACS Synthetic Biology* **7,** 553–564 (2018).

55. Cardinale, S., Joachimiak, M. P. & Arkin, A. P. Effects of genetic variation on the E. coli host-circuit interface. *Cell reports* **4,** 231–237 (2013).

56. Bradley, R. W. & Wang, B. Designer cell signal processing circuits for biotechnology. *New biotechnology* **32,** 635–643 (2015).

57. Borkowski, O., Ceroni, F., Stan, G.-B., *et al.* Overloaded and stressed: whole-cell considerations for bacterial synthetic biology. *Current Opinion in Microbiology. Antimicrobials • Microbial systems biology* **33,** 123–130 (2016).

58. Nielsen, A. A. K., Der, B. S., Shin, J., *et al.* Genetic circuit design automation. *Science* **352** (2016).

59. Cañadas, I. C., Groothuis, D., Zygouropoulou, M., *et al.* RiboCas: a universal CRISPR-based editing tool for Clostridium. *ACS synthetic biology* **8,** 1379–1390 (2019).

60. Patinios, C., Creutzburg, S. C. A., Arifah, A. Q., *et al.* Streamlined CRISPR genome engineering in wild-type bacteria using SIBR-Cas. *Nucleic Acids Research* (gkab893 2021).

61. Zhao, J., Inomata, R., Kato, Y., *et al.* Development of aptamer-based inhibitors for CRISPR/Cas system. *Nucleic acids research* **49,** 1330–1344 (2021).

62. Arnolds, K. L., Dahlin, L. R., Ding, L., *et al.* Biotechnology for secure biocontainment designs in an emerging bioeconomy. *Current Opinion in Biotechnology. Analytical Biotechnology* **71,** 25–31 (2021).

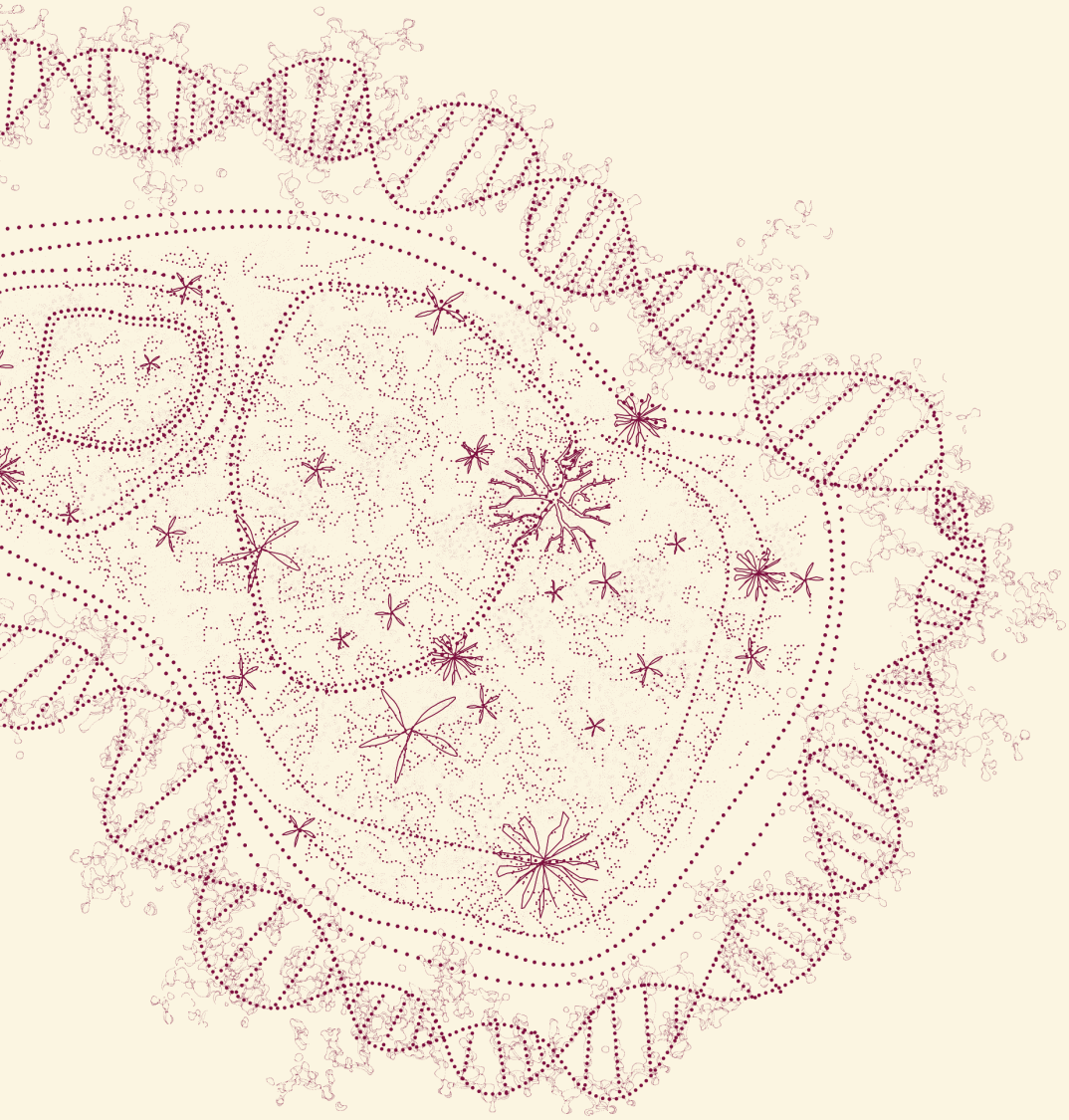
63. Cai, Y., Agmon, N., Choi, W. J., *et al.* Intrinsic biocontainment: Multiplex genome safeguards combine transcriptional and recombinational control of essential yeast genes. *Proceedings of the National Academy of Sciences* **112,** 1803–1808 (2015).

64. Kim, D. & Lee, J. W. Genetic Biocontainment Systems for the Safe Use of Engineered Microorganisms. *Biotechnology and Bioprocess Engineering* **25,** 974–984 (2020).

65. Green, R. & Rogers, E. J. Transformation of Chemically Competent E. coli. **529,** 329–336 (2013).

66. Choi, K.-H., Kumar, A. & Schweizer, H. P. A 10-min method for preparation of highly electrocompetent Pseudomonas aeruginosa cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J Microbiol Methods* **64,** 391–397 (2006).

67. Wirth, N. T., Kozaeva, E. & Nikel, P. I. Accelerated genome engineering of *Pseudomonas putida* by I- *Sce* I—mediated recombination and CRISPR-Cas9 counterselection. *Microb. Biotechnol.* **13,** 233–249 (2020).



CHAPTER 5

PHOSPHITE SYNTHETIC AUXOTROPHY AS AN EFFECTIVE BIOCONTAINMENT STRATEGY FOR THE INDUSTRIAL CHASSIS *PSEUDOMONAS PUTIDA*

Enrique Asin-Garcia, Christos Batianis, Yunsong Li,
James D. Fawcett, Ivar de Jong, Vitor A. P. Martins dos Santos

Abstract

Inclusion of biosafety strategies into strain engineering pipelines is crucial for the safe-by-design of biobased processes. This in turn might enable a more rapid regulatory acceptance of bioengineered organisms in both industrial processes and environmental applications. For this reason, we equipped the industrially relevant microbial chassis *Pseudomonas putida* KT2440 with an effective biocontainment strategy based on a synthetic dependency on phosphite. The produced PSAG-9 strain was first engineered to assimilate phosphite through the genome-integration of a phosphite dehydrogenase and a phosphite-specific transport complex. Subsequently, to deter the strain from growing on the naturally assimilated phosphate, all native genes related to its transport were identified and deleted generating a strain unable to grow on media containing any other phosphorous source than phosphite. Characterization of PSAG-9 reported fitness levels with phosphite similar to those of the wild type with phosphate, and low levels of escape frequency. In addition, this strategy conferred other industrially beneficial properties to *P. putida* such as the capacity to be cultured under non-sterile conditions using phosphite as the sole phosphorous source with a reduced risk of contamination by other microbes, and additional NADH regenerative capacity. These features complement the metabolic strengths of this species promoting it as a synthetic biology chassis with potential uses in industry with a suitability towards environmental release.

Introduction

Safe-by-design industrial chassis

The KT2440 strain of the gram-negative soil bacterium *Pseudomonas putida* shows considerable potential as a proper host for the development of enabling synthetic biology technologies and as a metabolically versatile chassis with a proven safety record and a tolerance towards genome reduction, toxic compounds and redox stressors [1–4]. With an increasingly accurate annotated genome [5] and a diverse collection of *in silico* and genome modification tools [6], the strain is well suited for high-throughput and automated engineering through the performance of iterative DBTL cycles in the context of biofoundries, biomanufacturing and industrial biotechnology in general [7, 8]. Thus, the generation of engineered *P. putida* strains with defined and beneficial industrial characteristics could play an enabling role in standardizing this microbe as a chassis for future generation of synthetic biology derived products and for its use in potential environmental applications [9].

As a consequence of the increasing demand for *P. putida* in research and industry, more profound genetic and metabolic engineering has been applied to this bacterium [10, 11]. This has led to a growing focus on improving the biosafety credentials of the engineered strains in order to prevent their escape from closed settings and their uncontrolled transmission in open or semi-open environments [12–14]. To this end, advocates of biosafety believe that implementation of a safe-by-design approach would be necessary to facilitate commonplace use of heavily engineered strains for applications involving environmental release [15, 16]. However, genetic safeguard technologies, still in their infancy for real-world applications, are not widely used within industry [17]. With biomanufacturing becoming the driving force behind the development and regulation of engineered organisms, industries will likely be required to consider biosafety upfront and incorporate biocontainment strategies in the early phases of strain development. In keeping with the synthetic biology tenet of standardisation, it would be beneficial if the biocontainment strategies selected for this were versatile and portable between different synthetic biology chassis [18]. “Plug-in” biocontainment modules would be indispensable in the rapid construction of strains with built-in biosafety, which could

help in streamlining the process of regulation [19].

Phosphite synthetic auxotrophy

With a non-complex technical implementation and a current record for the lowest detectable escape frequency of any genetic safeguard, synthetic auxotrophy is a good candidate to examine the potential portability of biocontainment amongst chassis [13, 20–22]. One of the most recent and successful strategies is the alteration of the phosphorous (P) metabolism to engineer a synthetic auxotrophy as an effective means of biocontainment. This auxotrophy has so far been successfully incorporated in both *Escherichia coli* and the cyanobacteria *Synechococcus elongatus* PCC 7942, two gram-negative model organisms that are representative of effective synthetic biology chassis [23, 24]. Phosphate (Pi, PO_4^{3-}), the biologically available form of P, is utilised in several key biological molecules, including nucleic acids, phospholipids and ATP. Most bacteria possess multiple Pi transporter genes [25] to provide transporter redundancy, in the off chance that one acquires a deleterious mutation, and also allows for the fine tuning of Pi import through differential gene regulation. The genome of *P. putida* KT2440 [5] indicates the presence of two low-affinity inorganic Pi transporters, PitB and PitA, and, unlike in *E. coli*, two copies of the high-affinity inorganic Pi transport complex, PstSCAB and PstSCAB-II. Additionally, orthologs of other genes that have previously been related to phosphate transport in *E. coli* can also be found in the KT2440 genome, including the phosphonate transporter PhnCEptxBC and organic Pi transporters.

Phosphite (Pt, HPO_3^{2-}), a reduced form of Pi, is an uncommon form of P in the environment and occurs naturally at only low concentrations [26–28]. Nevertheless, several species of bacteria have developed methods to metabolise Pt and utilise it as a source of P. The genes that comprise these metabolic processes can be isolated and used to endow a Pt-assimilation behaviour in a recombinant organism. In previous studies, the researchers inserted the Pt specific ABC transporter HtxBCDE, derived from the *Pseudomonas stutzeri* strain WM88, and the Pt dehydrogenase PtxD, taken from *Ralstonia* sp. 4506, to convert the newly imported Pt into Pi [23, 24, 29, 30]. Once the Pt utilization pathway was constructed, the ability of the engineered strain to incorporate the metabolizable Pi needed to be eliminated to

achieve a full auxotrophy.

Here, we established a Pt auxotrophy in *P. putida* KT2440 by integrating a functional Pt assimilation pathway and by deleting all the relevant native Pi transporters in order to produce a strain capable of utilising Pt, but not Pi, as the sole P source. In addition, the characteristics of this strategy equipped the strain with beneficial properties for industrial applications beyond biosafety, including a competitive growth advantage over contaminating organisms on media containing Pt as the sole P source, and an increased supply of NADH generated during the obligatory conversion of Pt into Pi. Thus, the development of this biocontainment strategy positions *P. putida* as a safer biological chassis with potential uses in research and industrial biotechnology, especially towards environmental applications.

Results

The *htxBCDE* and *ptxD* genes permit *P. putida* to utilise Pt as a P source

To confirm that *P. putida* KT2440 cannot naturally utilise and grow on Pt as a P source, a wild type *P. putida* strain (Figure 5.1 A) was first seeded into MOPS minimal medium containing 50 mM Glucose and relevant P sources (Figure 5.2 A). As expected, this strain was able to grow in media supplemented with inorganic Pi at a concentration of 1 mM. This data also suggests that *P. putida* KT2440 is unable to enter the log phase of growth without any P source, with the strain appearing to enter a stationary growth phase for the whole duration of the assay, likely maintaining a state of latency supported by internally accumulated Pi [31]. Likewise, the strain decreased below the seeding OD_{600} of 0.1 in both 1 mM and 2 mM Pt conditions, suggesting cell starvation and death, and confirming that Pt cannot be utilized as a P source.

As expected from the results of previous studies [23, 24], the introduction of the hypophosphite transporter HtxBCDE, derived from *P. stutzeri* WM88, and the phosphate dehydrogenase PtxD from *Ralstonia* sp. 4506 (in this study collectively referred to as the phosphite synthetic auxotrophy genes, abbreviated as PSAG) into the wild type KT2440 resulted in the strain gaining the ability to grow on Pt (Figure 5.2 A). After integrating both the *htxBCDE* operon and *ptxD* gene under the control

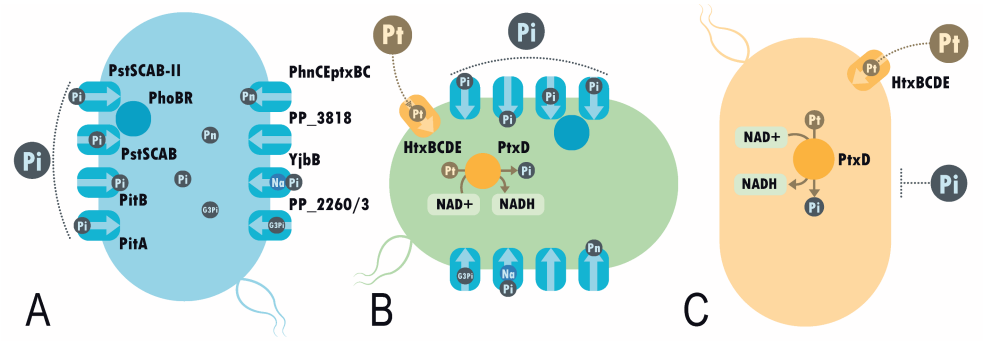


Figure 5.1: Graphical representation of the different stages of the generation of a phosphite auxotrophic P. putida strain. (A) In blue, P. putida KT2440 with its native Pi-transport-related genes. (B) In green, P. putida PSAG with its corresponding native genes, plus the hypophosphite transporter HtxBCDE from P. stutzeri WM88 and the phosphate dehydrogenase PtxD from Ralstonia sp. 4506. (C) In yellow, P. putida PSAG-9 equipped only with the phosphite assimilation genes htxBCDE and ptxD and deprived of its native Pi-transport related genes.

of the lac and tac promoters, respectively, into the *att*Tn7 site of KT2440 genome, the growth of the strain on Pt provided evidence that the PSAG were functionally expressed and metabolically active. *P. putida* PSAG (Figure 5.1 B) was able to grow in Pt to an optical density close to that of the wild type in Pi containing media, although the latter showed a higher fitness, achieved a higher final OD_{600} after 48 h and entered the acceleration and log phase of growth earlier than the PSAG strain. *P. putida* PSAG also appeared to grow on Pt at both the tested concentrations of 1 and 2 mM. Growth was significantly higher in the 2 mM Pt condition, indicating a growth-rate dependency on ambient Pt concentration. However, growth in 2 mM Pt was still significantly lower than growth in Pi at 1 mM.

Although there appears to be no native specific Pt transporters in *P. putida*, it was expected that the cells were capable of importing Pt through the native Pi transporters due to the structural similarities between the two forms of P [23]. To confirm this, a 24-h growth assay was prepared to examine if a *P. putida* strain containing *ptxD* but no *htxBCDE* in its genome was able to grow on a Pt source (Supplementary Figure S5.1). The PtxD-containing strain could indeed grow on Pt, indicating that Pt is able to enter the cell most likely though the native Pi transporters. This strain showed a significantly lower fitness than the *P. putida* PSAG on Pt, suggesting that the HtxBCDE transporter does function in importing Pt, and is more efficient

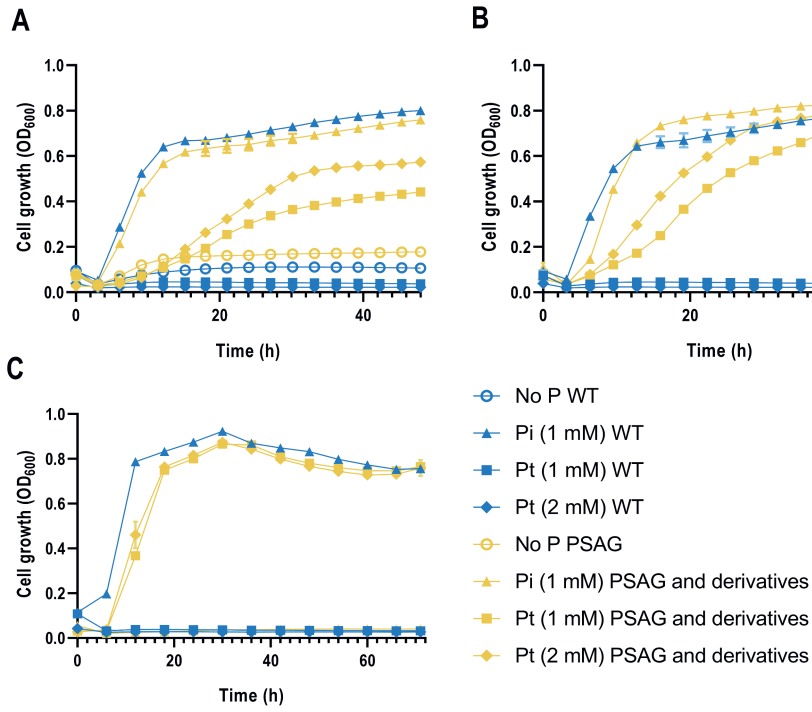


Figure 5.2: Differences in the growth capabilities of P. putida in different P sources after addition of the PSAG assimilation genes and the removal of its Pi native transporters. (A) Growth of P. putida wild type (blue) and PSAG (yellow) in MOPS with different P sources. The circle represents MOPS-0, the triangle MOPS-Pi (1 mM), the square MOPS-Pt (1 mM), and the diamond MOPS-Pt (2 mM). (B) Growth of P. putida wild type (blue) and PSAG-2 (yellow) in MOPS with different P sources. The triangle represents MOPS-Pi (1 mM), the square MOPS-Pt (1 mM), and the diamond MOPS-Pt (2 mM). (C) Growth of P. putida wild type (blue) and PSAG-4 (yellow) in MOPS with different P sources. The triangle represents MOPS-Pi (1 mM), the square MOPS-Pt (1 mM), and the diamond MOPS-Pt (2 mM). Growth was monitored by measuring OD_{600} in an ELx808 Absorbance Microplate Reader (BioTek Instruments, Inc., VT, U.S.). Error bars represent the standard deviation among biological duplicates and technical triplicates for each condition.

at Pt transport than any native Pi transporter. On a different note, the fact that the native Pi transporters can incorporate Pt while the wild type KT2440 is unable to grow with Pt as sole P source suggests that no native *P. putida* enzyme can convert Pt to Pi in an efficient enough manner as to provide sufficient levels of Pi for the cell to grow. This is relevant due to the existence of a native PtxD from *P. putida* encoded by *PP_3376* which was inferred from homology. Since the native PtxD does not seem to mediate our function of interest, we do not consider it in this study

and refer to the *Ralstonia* sp. 4506's protein when we talk of PtxD, unless stated otherwise.

Deletion of *P. putida*'s native Pi transporters prevents growth on Pi and enhances growth on Pt

Once it was verified that introduction of *htxBCDE* and *ptxD* allowed *P. putida* PSAG to grow on Pt as sole P source, the strain was engineered so as to be unable to grow in Pi. This process was initiated by deleting the low affinity inorganic Pi transporters PitB and PitA. Their removal appeared to show no major fitness deficit compared to the wild type strain when growing on Pi (Figure 5.2 B and Supplementary Figure S5.2 A), likely because of the high affinity transporters entirely taking over the Pi transport. When grown on Pt, the knock-out strains showed a considerable increase in fitness (Figure 5.2 B) compared to *P. putida* PSAG (Figure 5.2 A). This could be due to a lack of competition for membrane space that would be present in strains producing both HtxBCDE and the Pit transporters, allowing more Pt-specific transporters to be present at the cell surface. Membrane space competition can be expected from high transporter expression, which is here granted by the strong lac promoter controlling *htxBCDE*. However, this high expression could most likely cause a growth defect, which in turn, is not apparent in Figure 5.2 B. On 2 mM of Pt, *P. putida* PSAG-2 meets the final OD_{600} of the same strain grown on 1 mM of Pi after 48 h, establishing the suitability of this Pt concentration in the culture of the auxotroph and suggesting that the Pt concentration used is not yet limiting.

Following the removal of the low affinity Pi transporters PitB and PitA, the high-affinity PstSCAB and PstSCAB-II Pi transporters were removed from the strain. As both gene clusters are upregulated during Pi starvation [32], their removal would be likely to significantly hinder the ability of the cell to import inorganic Pi. As predicted, the removal of both transporters was sufficient to completely suppress growth in minimal Pi media for 72 h (Figure 5.2 C). When grown both in 1 or 2 mM of Pt, *P. putida* PSAG-4 reached an OD_{600} comparable to that of the wild type in 1 mM Pi after approximately 30 h of culture (Figure 5.2 C). This 4-knock-out strain in Pt has the closest growth profile to the wild type grown in Pi, with an elongated lag phase but comparable growth rate during the logarithmic phase.

While *P. putida* PSAG-4 was not able to grow in MOPS-Pi for 72 h, that was not the case when growing in rich media such as LB, probably due to the presence of other P sources or the cotransport of Pi when associated to other molecules missing in the minimal medium. To prevent our PSAG-4 strain from importing any other form of P that was not Pt, we knocked out a series of genes that were involved in P transport. First, we deleted the operon *phnCEptxBC*, whose product is mainly involved in metabolism and transport of phosphonates, though it has also been reported to work as an additional Pi transporter in other species [33]. Secondly, the *PP_3818* gene, which had been predicted to encode a periplasmic phosphate-binding protein [34], was subsequently deleted. Thirdly, the putative glycerol-phosphate transporter encoded by *PP_2260/3* was removed to avoid the incorporation of trace amounts of Pi *via* organic Pi transporters. Lastly, the Na^+/Pi-symporter YjbB was eliminated to exclude the possibility of Pi transport into the cell in Na^+-rich conditions, even though previous studies indicated that YjbB has none or little Pi-uptake activity in *E. coli* [35]. These four deletions resulted in the strains PSAG-5, PSAG-6, PSAG-7 and PSAG-8, respectively. Additional information concerning the fitness of all these intermediate strains is depicted in Supplementary Figure S5.2.

Furthermore, *P. putida* contains a complete set of *xcp* genes (type II secretion system) that participates in the secretion of enzymes involved in phosphate acquisition under low Pi conditions [36]. Overexpression of the corresponding genes is controlled by a two-component regulatory system that consists of PhoB as the response regulator and PhoR as the histidine kinase [37]. Given the high number of phosphatases and other enzymes involved in this Pi scavenging process, we opted for the removal of the *phoBR* regulon to avoid the activation of the system under Pi-limiting growth conditions, which resulted in *P. putida* PSAG-9 (Figure 5.1 C).

The growth and fitness of PSAG-9 were examined to verify the auxotrophic nature of the strain in both minimal and rich media using the wild type KT2440 as a control. As expected, PSAG-9 was not capable of growing in MOPS-Pi or LB lacking Pt (Figure 5.3). Although the results showed that the growth of PSAG-9 in either MOPS-Pt or LB-Pt was similar to the growth of the wild type in either MOPS-Pi or LB (Figure 5.3), the final cell density of PSAG-9 was lower compared with the wild

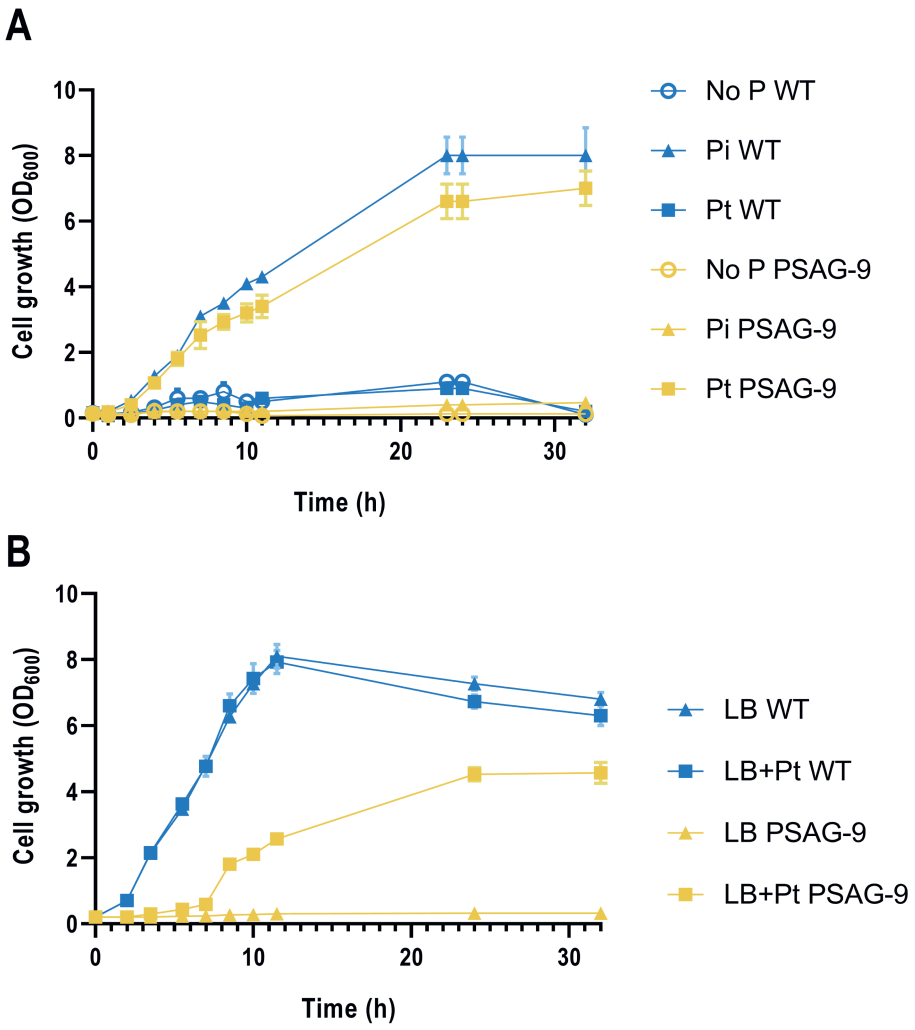


Figure 5.3: PSAG-9 strain only grows on Pt-supplemented media. (A) Growth of P. putida wild type (blue) and PSAG-9 (yellow) in MOPS with different P sources. The circle represents MOPS-o, the triangle MOPS-Pi, and the square MOPS-Pt. (B) Growth of P. putida wild type (blue) and PSAG-9 (yellow) in LB with and without Pt. The triangle represents LB and the square LB-Pt. Growth was monitored by measuring OD_{600} in 500 mL flasks. Error bars represent the standard deviation between biological triplicates in each condition.

type in either MOPS or LB, especially in the latter. To further investigate the growth of PSAG-9 in a Pt environment compared with the wild type in Pi, the growth rate of the two strains was calculated in different media. PSAG-9 on MOPS-Pt and LB-Pt was 0.51 ± 0.1 and 0.44 ± 0.02 (n = 3, mean ± sd) which was about 100 % of the growth

rate of the wild type on MOPS-Pi and 57 % of that on LB, respectively. Regarding final growth, at 32 h PSAG-9 had reached on MOPS-Pt an 87.5 % of the final OD_{600} of the wild type on MOPS-Pi. In LB-Pt however, the final OD600 of PSAG-9 was only 67.1 % of that reached by the wild type strain in LB. These results revealed that PSAG-9 is not able to make use of Pi in either inorganic or organic forms, but can consume Pt in a similarly effective manner to the wild type when making use of Pi, this being more evident in minimal media.

The auxotrophic PSAG-9 strain strictly uses Pt as the only P source

After producing a complete auxotrophic strain in liquid MOPS and LB media, a spot assay was carried out for a prolonged period of time to investigate if PSAG-9 was ultimately able to utilize other P sources present in different rich media typically used in laboratory settings (Figure 5.4). Results showed that the growth of the wild type strains on the rich media LB, Columbia agar, and terrific broth (TB) was significantly better than on the MOPS and soil extract plates as a consequence of their abundance in nutrients. However, the P sources from those media could not support the growth of PSAG-9 after 21 days. In light of the native habitat of *P. putida*, soil extract agar was also included in this experiment alongside the aforementioned rich and MOPS media. Soil extract has a complex composition and typically contains plenty of nutrients required for the growth of soil bacterium [38, 39]. Nevertheless, PSAG-9 was also not able to grow on the soil extract agar either. In contrast, growth of PSAG-9 was only observed on MOPS-Pt. These results further demonstrated that PSAG-9 could no longer use any organic or inorganic P source apart from Pt corroborating that it was a fully Pt-dependent auxotrophic strain.

PSAG-9 has a sufficiently low escape frequency

Subsequently, we studied the escape frequency of PSAG-9 under non-permissive conditions. To do so, two independent experiments were performed at different volume scales. First, we started with a cell culture of small volume (50 mL) containing approximately 2.7 × 10^8 cells, which was plated onto both MOPS-Pt- and MOPS-Pi-agar plates. CFU were counted after 21 days of incubation resulting in an escape frequency of 4.8 x 10^{-8} ± 1.81 x 10^{-8} (n = 3, mean ± sd). Next, in an at-

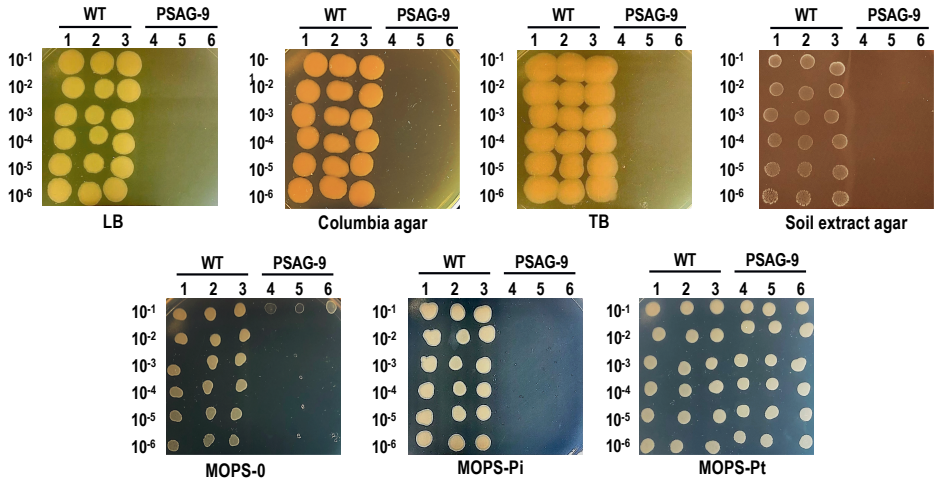


Figure 5.4: PSAG-9 could not use other various P sources except Pt. A spot assay was performed on seven different types of agar-containing media: LB, Columbia, TB, Soil extract, MOPS-0, MOPS-Pi, and MOPS-Pt. Three biological replicates were plated in different dilutions for the wild type (1, 2 and 3) and PSAG-9 (4, 5 and 6) respectively. For the spots, dilutions from an overnight culture ranging from 10^{-1} to 10^{-6} were used. Pictures were taken after 21 days of incubation at 30°C.

tempt to get a more accurate number for industrial production, we repeated the experiment on a bigger scale (1 L) with a larger cell culture that contained an initial number of approximately 3.0 × 10^{10} cells. After three weeks, the escape frequency was then estimated as 3.0 x 10^{-8} ± 0.91 x 10^{-8} (n = 3, mean ± sd). In both experiments, we demonstrated that PSAG-9 has an escape frequency in the order of 10^{-8}. For Host-Vector 2 systems, the recommended standards set by the NIH for GMOs it is precisely <10^{-8} [40]. While *P. putida* is Host-Vector 1 certified [4] and these NIH guidelines are not an official standard for biological containment, we used them as an estimation of having reached acceptable levels of biosafety in regard of the escape frequency metric.

PSAG-9 allows fermentations under non-sterile conditions

Given the industrial relevance of *P. putida* as an industrial chassis, we used PSAG-9 to test other beneficial traits that could be attractive for industrial applications besides biosafety. Strains that offer the possibility to operate under non-sterile conditions have become one of the most commercially appealing options when it

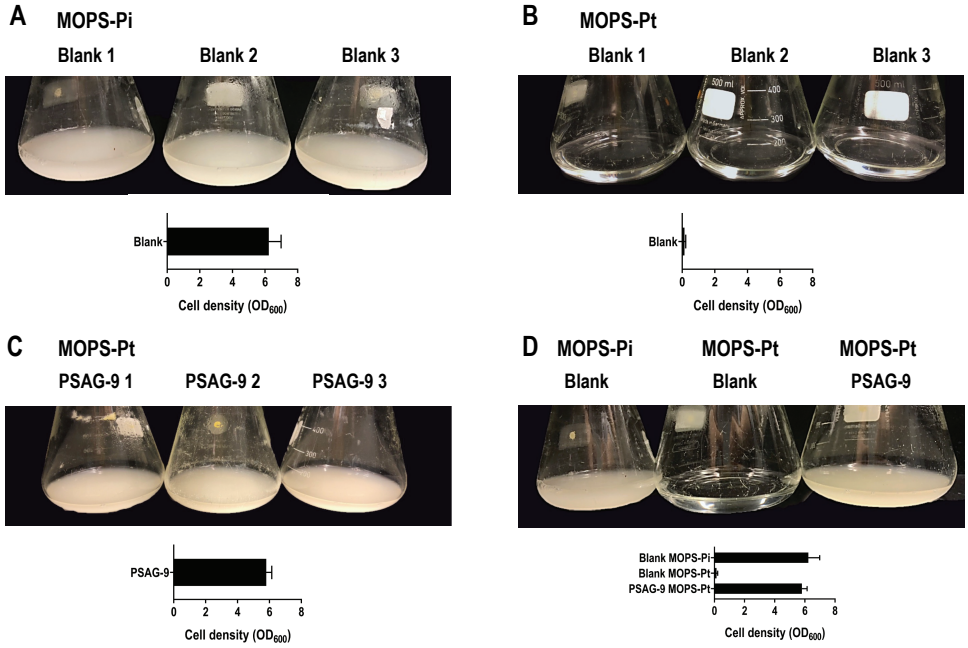


Figure 5.5: Pt as the only P source effectively inhibited the growth of other biological contaminants. PSAG-9 were cultured under non-sterile conditions in 250 mL flasks. Blank cultures of MOPS-Pi and MOPS-Pt were included as controls and each group was incubated in triplicates. Pictures were taken after 5 days of incubation. (A) The blanks with only MOPS-Pi medium; (B) the blanks with only MOPS-Pt medium; (C) the growth of PSAG-9 at 5 days post-inoculation in MOPS-Pt; and, (D) comparison of the three groups.

comes to industrial production [41, 42]. It is already known that Pt is barely available in the environment and can hardly be used by other organisms [24]. As a consequence, the Pt auxotrophy was expected to grant the PSAG-9 strain the possibility of performing non-sterile fermentation without the risk of biological contamination. To verify this, a non-sterility test was performed by using non-sterilized flasks and media. All steps were done under non-sterile conditions including inoculation. We observed that the blank of MOPS-Pt remained clear after 5 days of incubation (Figure 5.5 B) while the blank of MOPS-Pi showed contamination since the second day of the experiment (Figure 5.5 A). Meanwhile, the turbidity of the flasks inoculated with PSAG-9 was significantly increased after 5 days of incubation (Figure 5.5 C). The results indicated that the media with Pt as the only P source was effectively not suitable for the growth of other microorganisms.

To further investigate the bacterial growth from the two turbid groups (initially, blank MOPS-Pi and MOPS-Pt inoculated with PSAG-9), samples were diluted and spread onto plates and bacterial communities were distinguished by microbial morphology. On the one hand, the colonies that grew from the blank MOPS-Pi flask presented bigger and more yellowish morphologies when compared to PSAG-9 (Supplementary Figure S5.3, left). On the other hand, the morphology of the colonies derived from MOPS-Pt-PSAG-9 was more uniform and resembled the typical characteristics of the strain (Supplementary Figure S5.3, right) suggesting that the group MOPS-Pt-PSAG-9 only contained the PSAG-9 that was inoculated.

The phosphite assimilation pathway provides reducing power

The Pt dehydrogenase PtxD requires NAD^+ as a cofactor and creates NADH during the conversion of Pt to Pi [43]. The high redox potential of the NADH/NAD^+ couple suggests that growing *P. putida* PSAG or its derivatives on Pt could also have the additional advantage of providing a source of metabolic currency to carry out NADH-dependent metabolic processes. The production of NADH on a source of Pt would also give an indication that PtxD functionally converts Pt to Pi, and detecting NADH production would provide a suitable measure of the activity of the enzyme. To examine this, cell extracts were prepared from both *P. putida* KT2440 and *P. putida* PSAG cell cultures to follow their NADH production over time (Supplementary Figure S5.4). The rate of increasing NADH concentration per mg/mL of protein was calculated as 0.87 mM/min/mg/mL for the PtxD-containing cultures. Interestingly, extracts of cells deficient in PtxD also showed a NADH increase of around 0.326 mM/min/mg/mL of protein, suggesting that the wild type could have a native Pt dehydrogenase activity, probably due to the presence of the native PtxD (*PP_3376*), which appeared to be significantly less efficient than the PtxD from *Ralstonia* sp. 4506 (Figure 5.6).

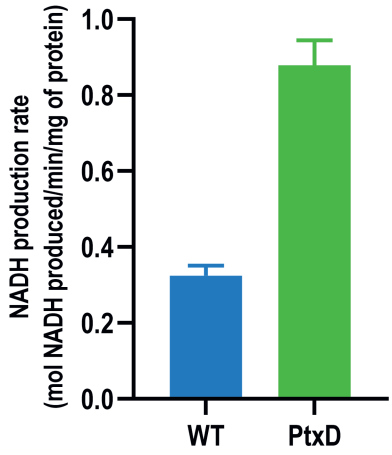


Figure 5.6: NAD^+/NADH conversion in wild type (blue) and PtxD-containing (green) cell extracts. Wells were seeded with 0.3-0.45 mg/ml of total protein from clarified cell extracts. Values were normalised for total protein concentration in each extract. OD_{340} was measured for 90 min in 30-sec intervals and the depicted bars represent the increasing NADH concentration rate throughout that period. Error bars represent the standard deviation between biological triplicates and technical duplicates in each condition.

Discussion

Assimilation of P in *P. putida*

We successfully engineered a phosphite synthetic auxotrophy in *P. putida* KT2440. Our research showed that the auxotrophic strain PSAG-9, which contains in its genome the *ptxD* and *htxBCDE* genes, can effectively uptake Pt as a P source for supporting growth (Figure 5.2 A). Furthermore, the deletion of all native Pi transporters including the inorganic Pi transporters, Pn transporter, organic Pi transporters, and the PhoBR regulatory system suppressed the ability of the strain to consume Pi (Figure 5.3). Ultimately, no growth of PSAG-9 was observed in any medium lacking Pt (Figures 5.3 and 5.4).Prior to this research, references to the native Pi transport system of *P. putida* KT2440 were scarce [44, 45], and most transporters' annotations were based on their homology with their *E. coli*'s counterparts. During this study, we found that deletions in what we identified by homology as PitA (*PP_4103*) did not dramatically affect the growth of *P. putida* in Pi (Supplementary Figure S5.2 A). In *Mycobacterium smegmatis*, deletions of PitA similarly did not alter phosphate

uptake and cell growth, even though PitA is the only low-affinity Pi transporter annotated in this organism [46]. However, the deletion of PitA in *M. smegmatis* was found to cause the upregulation of PstSCAB. This further supports the notion that Pi assimilation relies on a redundancy of regulatable transporters, and that PstSCAB alone is likely sufficient to support growth. This was similarly indicated in our results with *P. putida*, which suggested that the strain could still grow efficiently after the deletion of both PitA and PitB (Figure 5.2 B). Furthermore, only the PstB component of the second PstSCAB complex was annotated as *pstB-II* at the beginning of our research. First, we hypothesized the possibility of PstB-II being a single isoenzyme of PstB. However, we later concluded based on the homology of the neighbouring genes that the operon *PP_5326-PP_5329* could be a completely separate PstSCAB complex. It should be noted that only when both *pstSCAB* and *pstSCAB-II* were completely deleted the strain was unable to grow on Pi in minimal medium, confirming our latter hypothesis.

In the environment, many species are able to fulfil their P requirements through import of organic Pi-esters, including phosphonates. Around 10 % of the dissolved P in the ocean is in a phosphonate form, and some marine life has adapted to assimilate it [47]. Phosphonates are also known to be utilised by soil bacteria. Specifically, *P. putida* has been demonstrated to utilise the phosphonate 2-aminoethylphosphonic acid as its sole source of carbon, nitrogen and P, further demonstrating the robustness of this organism's metabolism [48]. Interestingly, the presence of Pt specifically inhibited the activity of the enzymes required to convert the phosphonate into Pi. Regardless, we still removed *P. putida*'s ability to transport phosphonates by deleting the *phnCEptxBC* genes to ensure that the organism could not grow on any other environmental form of P other than Pt.

Regarding the exact function of the deleted *PP_3818*, little is known beyond its annotation as a periplasmic phosphate transport system substrate-binding protein. While it contains an outer membrane protein A (OmpA) domain, much like the phosphate scavenging protein 3 (psp3) found in comparative genomic analysis of *Pseudomonas* isolates [49], there is no evidence of the product of *PP_3818* having this scavenger function. Nevertheless, exoproteins psp, secreted under Pi-limited conditions to scavenge extracellular Pi before being transported back into the cells

through the PstSCAB transporters [50], which had been previously knocked-out of our strain's genome. On a different note, homologs of the glucose-3-phosphate transporter genes are widely found in many living organisms and their function had been previously characterized in *E. coli* [51, 52], revealing that on top of glycerol-3-phosphate, this transporter may participate as well in Pi uptake [53]. In contrast, the existing literature on the Na^+/Pi symporter YjbB is limited but the few studies on *E. coli*'s YjbB suggest that it might also be involved in the transport of Pi and the regulation of the *pho* regulatory genes [35]. The PHO Pi starvation regulatory system has recently been studied in depth in *P. putida* [54]. The mechanism involves the response regulator PhoB and the histidine kinase PhoR, which were known in *P. putida* KT2440 for regulating the expression of a series of genes related to Pi transport or assimilation under inorganic Pi starvation [36].

The main effect of Pho induction is the upregulation of *pstSCAB* to compensate for the low internal Pi concentration, which may explain why the deletion of *pitB* and *pitA*, and the likely initial condition of Pi starvation that the deletions caused, resulted in an increased ability to grow on Pi (Figure 5.2 B). Similarly, the lack of growth defect in Pi upon *pstSCAB* deletion (Supplementary Figure S5.2 B) may have been a result of the upregulation of the second *pstSCAB-II* variant as a way of compensation. As a result of this Pi concentration-dependant feedback system, it would be difficult to calculate the relative effect that each of the sequential deletions actually has on the wild type KT2440. Although the goal of this project was to ablate growth in Pi, necessitating the removal of all Pi transporters, it would be useful to measure the effect of each transporter deletion on the resulting upregulation and expression of the other transporters to verify our initial assumption, which could be done through a comparative RNASeq analysis to establish the relative abundance of mRNA for the transporters in each knockout configuration [55].

PSAG-9 characteristics and biosafety credentials

Notably, as it can be seen in Figures 5.3 and 5.4, the wild type *P. putida* could grow slightly in both liquid and solid media without P source and with only Pt as the only P source. Studies have found that *P. putida* has the ability to survive under phosphate starvation [56]. There are probably two reasons to explain this phenomenon.

Firstly, the active aforementioned PHO starvation regulatory system, which results in the strong Pi storage capacity of the strain that accumulates it in the form of polyP. There is evidence that these polyPs play an important role in the cells under stress conditions [31]. Second, the phosphate-deficient environment puts strong selection pressure on *P. putida*, which can lead to the development of mutants that have a growth advantage in the environment and are able to grow with nutrients from dead cells [56]. However, the same phenomenon was not found in PSAG-9, suggesting that the phosphite auxotrophy strategy weakens the tolerance of PSAG-9 to P starvation to a great extent, ensuring that the viability of PSAG-9 will be rapidly reduced even when PSAG-9 is released to the open environment. On the other hand, the fitness assay (Figure 5.3) reflected that although the growth rate of PSAG-9 in the presence of Pt was similar to the wild type in Pi during the growth phase, its stationary phase value was lower than that of KT2440. The difference between the two strains in the stationary phase was more pronounced in LB. This indicated that the engineering of Pt auxotrophy has a certain effect on the fitness of bacteria. The stationary phase usually occurs when nutrients are not enough to support cell division and growth [57]. As mentioned above, the wild type is more tolerant to P starvation than PSAG-9, therefore, the former is always able to survive better than PSAG-9 when there is no supplementation of the P source. Considering that LB is a rich medium, we speculated that LB contains a higher concentration of P source to support the growth and proliferation of the wild type than MOPS medium does when supplied only with an inorganic P source. Since the P source of PSAG-9 is still Pt, the difference between the two strains in LB at the stationary stage should be expected to be more significant.

In addition, our study showed that PSAG-9 cannot survive under any non-permissible conditions (Figures 5.3 and 5.4). In the absence of Pt, the probability of occurrence of escape mutants surpassed the NIH standard of 10^{-8}. For the practical application of *P. putida* as a synthetic biology chassis, the lower the escape frequency, the higher the guarantee for biosafety. As the phosphite auxotrophy applied in *P. putida* yielded a significantly higher escape frequency compared to that in *E. coli* [23], this may mean that *P. putida* may be better able to cope with P starvation [56]. From the biosafety standpoint, this might be a liability for the stability

of this specific strategy in this species.

PSAG-9 advantages and limitations

Nevertheless, as a biocontainment strategy, Pt synthetic auxotrophy provides many advantages. Pt is cheap, non-toxic and simple to produce, and as such would make an acceptable supplement for large quantities of culture media [58]. Pt is a stable compound and is not likely to be depleted by the majority of organisms that might share an environment with the auxotrophic *P. putida*. In addition, its centrality to the maintenance of the auxotrophic strain, in being required for nucleic acid and ATP synthesis, suggests that a circumvention of biocontainment is less likely than in strategies that encode, for example, toxin: antitoxin pairs. The PSAG-9 strain constructed in this project not only guarantees its biosafety as a biochassis, but also endows it with the advantage of reducing the application cost. Compared to the vitamins and amino acids commonly used in auxotrophy strategies, the costs of feeding Pt are extremely low [23]. Due to the cultivation characteristics of PSAG-9, it has the effect of controlling contamination from other undesired bacteria (Figure 5.5), which provides a stable guarantee for its outdoor application. In addition, this feature also provides the potential for large-scale non-sterilized fermentation, saving the cost and time of sterilization of equipment and media in industrial production.

Moreover, as a by-product of the Pt conversion to Pi by the dehydrogenase PtxD, NAD^+ is reduced to NADH [43]. The high reducing power of NADH makes it a valuable metabolic currency with uses in many catabolic processes [59]. *P. putida* already has a proven high rate of NADH regeneration [60]. The high availability of NADH is vital to the strategy of *P. putida* in removing toxic compounds, as oxidation of the cofactor supplies the energy required to drive the organism's efflux pumps [61]. The metabolism of polycyclic aromatic hydrocarbons (PAH), catalysed by PAH dioxygenases, is another process dependent on NADH availability [62]. In this study, we verified that *ptxD*-encoding *P. putida* strains produce higher levels of NADH than the wild type organism when assimilating Pt (Figure 5.6). This would be a great benefit to *P. putida* strains encoding this auxotrophy that are designed to have applications in environmental bioremediation, as the Pt conversion provides additional

energy to drive the catabolism of toxic compounds in the form of more available NADH.

Although PSAG-9 has numerous benefits and application potential, there is still room for further research and optimization given some of the strain's limitations. First, only a single biocontainment strategy was applied to this strain. The study has shown that horizontal gene transfer might be a risk for this strategy because it can enable Pt-dependent strains to regain the ability to transport Pi if any Pt transporter is acquired horizontally [23]. Consequently, it would be convenient to combine this strategy with other safeguards for PSAG-9 to make up for the inherent shortcomings of the individual protection measures. For example, the Pt auxotrophy established in *P. putida* could be complemented with a methanol auxotrophy, meaning two entire auxotrophy systems would have to be circumvented to allow environmental escape [63]. More extreme changes, such as the establishment of sematic containment, could also be considered in future combinatorial biocontainment strategies for *P. putida*, in an effort to demonstrate biosafety to a level acceptable to regulators [64, 65]. Secondly, Pt is still widely used in agriculture as a fungicide [66]. If PSAG-9 is to be used for bioremediation caused by pesticides pollution, the fields should be screened for whether they contained Pt or not, which could increase the workload of the overall cleaning procedure. Finally, research has suggested that the microbial community in soil and water is usually capable of rapidly absorbing and oxidizing Pt [23]. Whether the extensive use of Pt in the open environment supplied for the use of PSAG-9 will damage the balance of the original microbial community and cause the enrichment of the community remains to be verified.

Orthogonality and concluding remarks

As a final remark, we would like to reflect on the orthogonality of genetic safeguards. While we may have ambitions towards having a repertoire of synthetic biology chassis that can be equipped on-demand with different standard modules and genetic safeguards, each of these features will still require accounting for their corresponding cellular and application contexts [19]. In this specific case, although it was fair to assume that the PSAG components described in the previous studies

in *E. coli* and *S. elongatus* would function sufficiently in our chosen organism, it quickly became evident that bacterial species have distinct variations concerning their methods of Pi transport. In *Pseudomonas* alone, several species encode different combinations of Pi transporters [49]. Additionally, our initial search for transporters suggested that Pi transport in *P. putida* had not been detailed exhaustively, with many Pi-interacting proteins remaining unannotated and uncharacterised in the species. Similarly, the native Pi transporters of *E. coli* were disrupted by P1 transduction [23]. Our chosen method of gene deletion in *P. putida*, based on the integrative plasmid pGNW, required more technical steps and periods of overnight incubation to complete than P1 transduction, making the enactment of this biocontainment strategy in our chosen organism more time and resource intensive [67, 68].

In summary, by developing a biocontainment strategy based on synthetic auxotrophy for *P. putida*, this work attempted to enable its deployment as a synthetic biology chassis with potential uses in industry with enhanced, built-in biosafety features. Additionally, this may also provide a basis for its possible use in non-contained environmental applications.

Materials and methods

Strains and growth conditions

All bacterial strains used in this study are listed in Supplementary Table S1. Culturing of bacterial strains was conducted in Lysogeny Broth. Strains carrying antibiotic selection markers were cultured with their respective antibiotics in the following concentrations: kanamycin 50 μg/mL, gentamicin 10 μg/mL and 15 μg/mL (for *E. coli* and *P. putida* respectively), apramycin 50 μg/mL and chloramphenicol 20 μg/mL. Morpholinepropansesulfonicacid-glucose (MOPS) synthetic medium supplemented with 50 mM Glucose and variable P sources and concentrations was used as a minimal medium [69] for Phosphite Synthetic Auxotrophy experiments. Pt stock solutions were prepared dissolving phosphorous acid in distilled water followed by pH neutralization, filtration and storage at -20°C. MOPS media containing Pi, Pt are designated as MOPS-Pi, MOPS-Pt respectively. P-free MOPS medium is designated

as MOPS-0. *E. coli* and *P. putida* strains were grown at optimal temperatures, 37°C and 30°C respectively.

Plasmids

A list of all plasmids utilized in this study with their respective features is provided in Supplementary Table S2. Primers used for the amplification of genetic parts, genome editing verification and sequencing purposes (Supplementary Table S3) were synthesized by Integrated DNA Technologies (IDT) as salt-free without further purification, resuspended in milli-Q at 100 μM and long-term stored at -20°C. Q5® High-Fidelity DNA Polymerase (New England Biolabs) and Platinum SuperFi II DNA Polymerase (Thermo Fisher Scientific) were used for PCR amplification of genetic parts according to manufacturers' instructions. The resulting PCR products were run on 1 % (v/v) agarose gel to verify fragments size. PCR fragments were isolated from PCR product or from agarose gel using NucleoSpin® Gel and PCR Clean-up (BIOKÉ).

All plasmids were assembled using Golden Gate or the Golden Gate-based SevaBrick Assembly [70]; therefore, BsaI recognition sites were included in the primers to produce 4-nucleotide overhangs on the 5'- and 3'-ends of each PCR amplicon used for plasmid construction. Assembled plasmids were transformed into chemically competent *E. coli* DH5α using heat shock transformation [71]. Transformant *E. coli* colonies were subsequently selected and correct assembly of the constructs was screened *via* Phire Hot Start II DNA Polymerase (Thermo Fisher Scientific) colony PCR according to manufacturer's instructions. Corresponding clones were inoculated in LB liquid cultures and plasmids were isolated using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) following manufacturer's protocol. Plasmid insert sequences were ultimately confirmed with standard DNA sequencing (MACROGEN Europe).

Engineering knock-out and knock-in *P. putida* mutants

Phosphite assimilation genes were introduced in the genome of *P. putida* KT2440 *via* site-specific recombination. First, *ptxD* from *Ralstonia* sp. 4506 under the control of the tac promoter [72] and *htxBCDE* from *P. stutzeri* WM88 (Supplementary Table S4)

under the control of the lac promoter were amplified from plasmids pSTV28 with primers C3 and C4, and C5 and C6, respectively, and cloned together in pSEVAb84 ptxD + htxBCDE, using the pSEVAb84 backbone amplified with C1 and C2 primers. Either *ptxD* only, by using primers C11 and C12, or the whole cassette, using primers C11 and C13, were amplified for construction of pSEVAb22 cre + ptxD and pSEVAb22 cre + ptxD + htxBCDE, using the in-house amplicon pSEVAb22 cre which was previously amplified with C9 and C10 and works as a standard backbone for cre lox integration in a lox71-lox66/2m landing pad located in the *att*Tn7 *locus* of the *P. putida* genome. Transformation of pSEVAb22 cre plasmids was performed *via* electroporation in competent *P. putida* KT2440 [73] cells containing the mentioned landing pad. Site-specific recombination and subsequent cassette integration was confirmed with V1 and V2 primers and resulted in the creation of the *P. putida* PSAG strain.

P. putida PSAG-derived knock-out and knock-in strains were generated as previously described using tri-parental conjugation for chromosomal integration of suicide pGNW vectors followed by the action of meganuclease I-SceI, which was located in a pQURE6.H plasmid [68, 74]. Construction of pGNW plasmids for knock-outs was based on the cloning of an upstream and a downstream homologous regions of 500 bp required for homologous recombination into the pGNW backbone amplified with P1 and P2 primers. Individual *P. putida* clones were tested for wild type or mutant genotype *via* Phire Hot Start II DNA Polymerase (Thermo Fisher Scientific) colony PCR according to manufacturer's instructions with corresponding verification primers [75].

Following the aforementioned knock-out protocol, all endogenous phosphate transport encoding genes were sequentially knocked-out of the *P. putida* PSAG genome including those that encode: (a) the low affinity inorganic phosphate transporters PitA and PitB (*PP_4103* and *PP_1373*); (b) the high affinity inorganic phosphate transporters PstSCAB and PstSCAB-II (*PP_2656-PP_2659* and *PP_5326-PP_5329*); (c) the phosphate and phosphonate transporter complex PhnCE-PtxBC (*PP_0824-PP_0827*); (d) the periplasmic phosphate-binding protein encoded by *PP_3818*; (e) the glycerol-phosphate transporter encoded by *PP_2260-PP_2063*; (f) the Na+/phosphate symporter YjbB (*PP_0145*); and (g) the two-component regula-

tory system PhoBR (*PP_5320-PP_5321*).

Growth and fitness assays

To determine the growth and fitness of the different *P. putida* PSAG strains in different P sources, assays were performed on a smaller and a larger scale. Precultures of *P. putida* KT2440 and PSAG strains were prepared 18-24 h prior the harvesting which was done by centrifugation at 4700 g for 5 min at room temperature. Pellets were washed twice with an equal volume of MOPS-0, and finally resuspended in MOPS-0 to a final volume of 5 mL. The growth and fitness of the samples were monitored by measuring OD_{600}. For the small-scale experiments, monitoring was performed for 24-72 h at 30°C under gentle agitation in an ELx808 Absorbance Microplate Reader (BioTek Instruments, Inc., VT, U.S.). Cultures were seeded in biological duplicates and technical triplicates to an initial OD_{600} of 0.3 in either MOPS-0, MOPS-Pi, MOPS-Pt, LB, or LB-Pt media, to a final volume of 200 μL per well in a 96-well plate. Large scale experiments were perfomed in 500 mL flasks. Each sample was inoculated into MOPS-0, MOPS-Pi, and MOPS-Pt to an initial OD_{600} of 0.2 in a final volume of 50 mL. Cultures were incubated at 200 rpm and 30°C for 24 h. Growth was monitored by measuring OD_{600} every 1 to 1.5 h. The growth rate was calculated as a function of the variation in the number of cells during the exponential phase period.

Assessment of different P source viability

For assaying whether the *P. putida* PSAG strain strictly uses Pt as the only P source, the growth of the strain on different types of media containing a variety of P sources was examined. Three biological replicates of *P. putida* PSAG-9 and *P. putida* KT2440 were grown overnight. Precultures were pelleted down and washed with an equal volume of MOPS-0. Cells were resuspended in 5 mL MOPS-0 and bacterial resuspensions were seeded into a 96-well plate to a final volume of 200 μL with an equal OD_{600}. 2 μL of corresponding serial dilutions were spotted onto different agar plates including Lysogenic Broth (10 g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract and 15 g/L bacteriological agar), Columbia medium (10 g/L pancreatic digest of casein, 5 g/L meat peptic digest,3 g/L heart pancreatic digest, 5 g/L yeast extract, 1 g/L maize starch, 5 g/L sodium chloride, 12 g/L bacteriological agar), Terrific Broth (12

g/L tryptone, 24 g/L yeast extract, 0.4 % (v/v) glycerol, 10 % (v/v) phosphate buffer (23.12 g/L KH_2PO_4 and 125.4 g/L K_2HPO_4) and 15 g/L bacteriological agar), Soil extract agar (500 mL/l soil extract and 15 g/l agar), MOPS-Pi, MOPS-Pt, and MOPS-0 [69]. The plates were incubated at 30°C and growth of the spots was monitored for 21 days.

Escape frequency assay

To examine the reliability of the Pt-dependent containment strategy on *P. putida*, escape frequency of auxotrophic strains was assessed in permissive and non-permissive conditions. Three biological replicates of the auxotrophic *P. putida* PSAG-9 strain were incubated at 30°C overnight. Precultures were then used for inoculation of 50 mL and 1 L-flask cultures of MOPS-Pt to an initial OD_{600} of 0.2. Next day, grown cultures were centrifuged at 4700 g for 5 min to collect the pellet. Cells were washed once with an equal volume of MOPS-0 and resuspended in 1 mL MOPS-0 for subsequent preparation of serial dilutions down to 10-20. 50 μL from relevant dilutions was plated on MOPS-Pi and MOPS-Pt. Plates were grown for 21 days before colonies were counted. Escape frequency was calculated as the ratio between the number of colonies that escaped and could grow in MOPS-Pi and the total number of colonies that grew in MOPS-Pt.

Non-sterile fermentation ability test

For the non-sterile fermentation test, cultures of 10 mL of *P. putida* PSAG-9 were grown in triplicates to an OD_{600} of 5 and were used to inoculate MOPS-Pt in 0.5 L flasks to a final volume of 50 mL with a starting OD_{600} of 0.1. 0.5 L flasks containing 50 mL of MOPS-Pt and MOPS-Pi were used in triplicates as blank during the experiment. The flasks and medium utilized in this experiment were not sterilized and the inoculation was performed under non-sterile conditions. Flasks were incubated at 30°C and 200 rpm for 5 days. Samples of the cultures were collected every 24 h for plating and visual analysis of the grown microbial community.

NAD+/NADH conversion assay

P. putida PSAG was grown to mid-exponential phase, alongside *P. putida* KT2400. Cells were harvested by centrifugation at 4700 g for 10 min at 4°C and resuspended in cold PBS buffer containing 10 mM 2-mercaptoethanol. The washed cells were centrifuged again under the same conditions, resuspended in 750 μL of the same PBS buffer, and transferred to bead-beater tubes containing 0.1 mm silica beads. Cells were subsequently lysed by pulsing with a cell disruptor for 30 sec, and then cooled for 5 min on ice before being pulsed for a second time. The crude lysates were clarified by centrifugation at 7500 g for 30 min at 4°C before being transferred to a fresh 1.5 mL tube. The protein concentrations of the cell extracts were calculated by a 30-45-minute incubation with Bradford reagent and comparing absorbance at OD_{595} with a BSA standard curve using a Synergy™ Mx microplate reader. To measure NAD^+/ NADH conversion rates, a buffer solution containing 50 mM K_2HPO_4, 5 mM $MgSO_4$ and 0.5 mM NAD^+ were mixed with 40 μL of cell lysate solution from the wild type and *ptxD*-containing cultures in a 96-well plate in biological triplicates and technical duplicates. 0.5 M Pt was then added to the samples to a final volume of 200 μL, with water otherwise used as a negative control. Immediately after mixing, the plate was monitored for NADH production by measuring absorbance at 340 nm over 90 min. The OD_{340} values were then converted into NADH concentration by comparison with a prepared NADH standard curve and normalised for protein concentration.

Acknowledgements

We would like to thank Ryuichi Hirota for providing the initial vectors containing the *htxBCDE* and *ptxD* genes. In addition, we wish to convey our utmost gratitude to Lyon Bruinsma and Maria Martin-Pascual for their feedback, help and suggestions during all the stages of the project. This work was funded by the research program SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis, which is supported by the Dutch Research Council (NWO) [15814]. Funding for open access charge: Dutch Research Council.

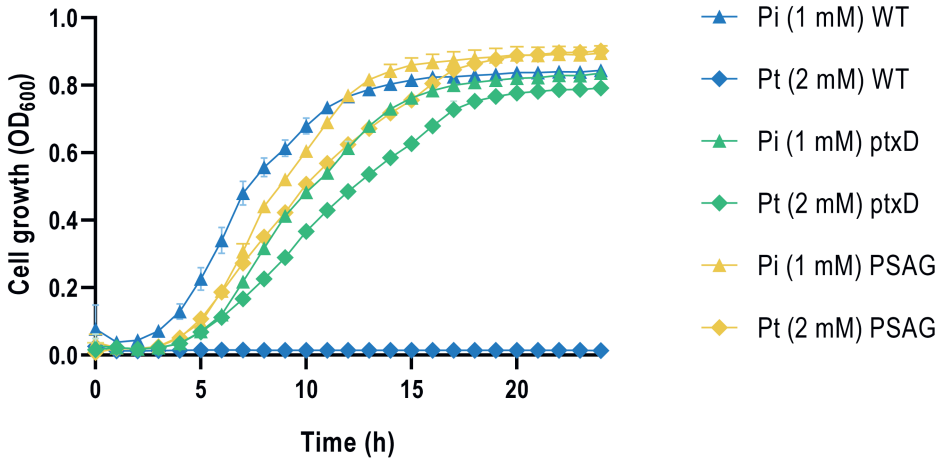
Conflict of interest

The authors declare there are no conflicting interests.

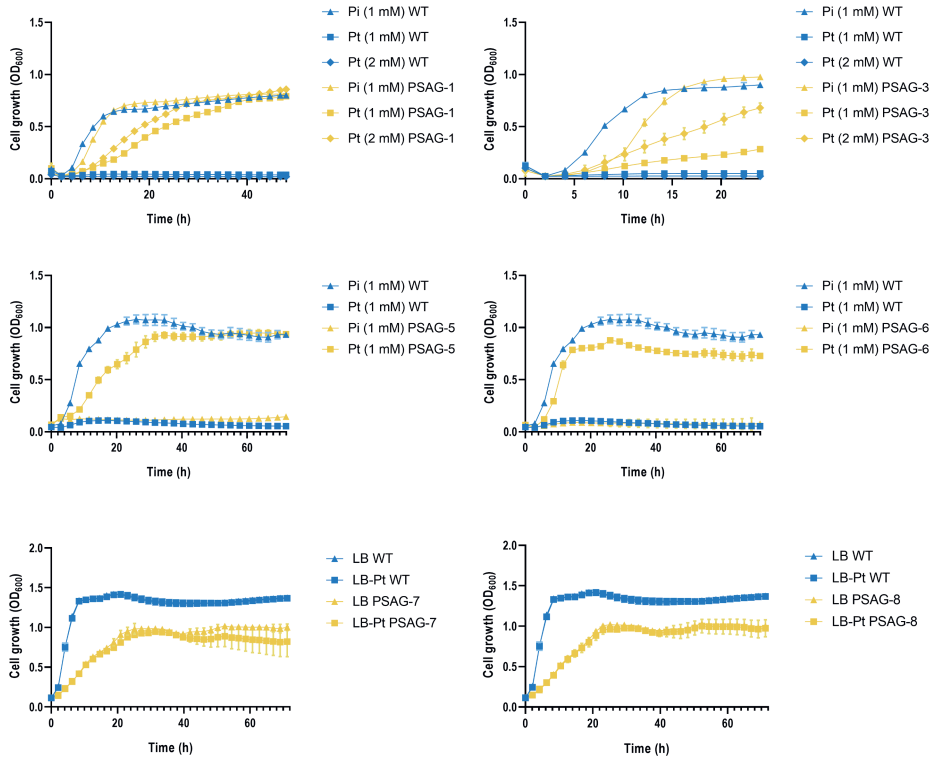
Author contributions

Conceptualization: EAG/CB/IdJ; Data curation: EAG/YL/JDF/IdJ; Formal analysis: EAG/YL/JDF/IdJ; Funding acquisition: VAPMdS; Investigation: EAG/CB/YL/JDF/IdJ; Methodology: EAG/CB/YL/JDF/IdJ; Resources: VAPMdS; Supervision: VAPMdS; Validation: EAG/CB; Visualization: EAG; Writing – original draft: EAG/YL/JDF; Writing – review & editing: EAG/CB/YL/JDF/IdJ/VAPMdS

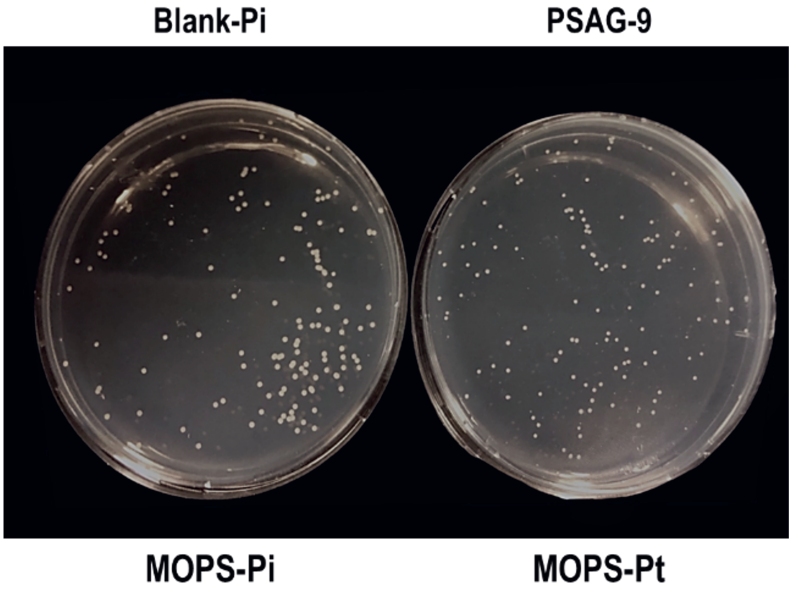
Supplementary material



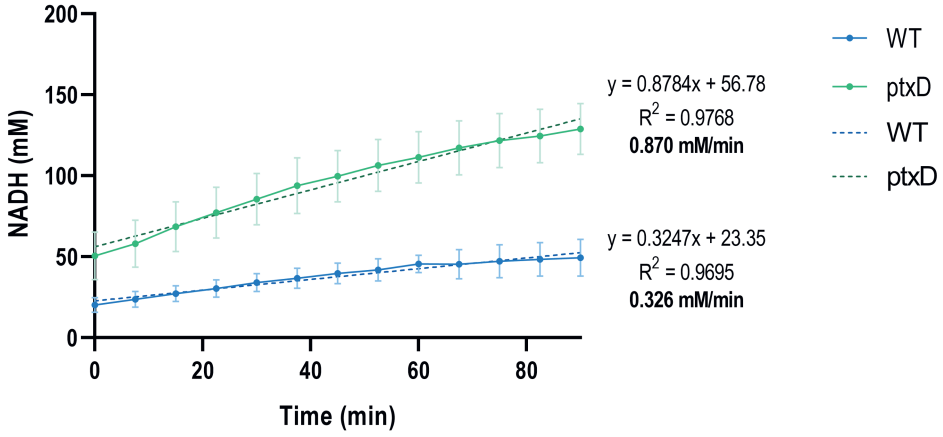
Supplementary Figure S5.1: PtxD only is required to grow on Pt. Growth of *P. putida* wild type (blue), PtxD-containing strain (green) and PSAG (yellow) in MOPS with different P sources. The triangle represents MOPS-Pi (1 mM), and the diamond MOPS-Pt (2 mM). Growth was monitored by measuring OD_{600} in an ELx808 Absorbance Microplate Reader (BioTek Instruments, Inc., VT, U.S.). Error bars represent the standard deviation among biological duplicates and technical triplicates for each condition.



Supplementary Figure S5.2: Evolution of the growth of *P. putida* in different P sources after addition of the PSAG assimilation genes and the removal of its Pi native transporters. (A) Growth of *P. putida* wild type (blue) and PSAG-1 (yellow) in MOPS with different P sources for 48 h. The triangle represents MOPS-Pi (1 mM), the square MOPS-Pt (1 mM), and the diamond MOPS-Pt (2 mM). (B) Growth of *P. putida* wild type (blue) and PSAG-3 (yellow) in MOPS with different P sources for 24 h. The triangle represents MOPS-Pi (1 mM), the square MOPS-Pt (1 mM), and the diamond MOPS-Pt (2 mM). (C) Growth of *P. putida* wild type (blue) and PSAG-5 (yellow) in MOPS with different P sources for 72 h. The triangle represents MOPS-Pi (1 mM) and the square MOPS-Pt (1 mM). (D) Growth of *P. putida* wild type (blue) and PSAG-6 (yellow) in MOPS with different P sources for 72 h. The triangle represents MOPS-Pi (1 mM) and the square MOPS-Pt (1 mM). (E) Growth of *P. putida* wild type (blue) and PSAG-7 (yellow) in LB with and without Pt for 72 h. The triangle represents LB without Pt and the square LB with Pt. (F) Growth of *P. putida* wild type (blue) and PSAG-8 (yellow) in LB with and without Pt for 72 h. The triangle represents LB without Pt and the square LB with Pt. Growth was monitored by measuring OD_{600} in an ELx808 Absorbance Microplate Reader (BioTek Instruments, Inc., VT, U.S.). Error bars represent the standard deviation among biological duplicates and technical triplicates for each condition.



Supplementary Figure S5.3: The cultures were collected from the initially blank MOPS-Pi (left) and MOPS-Pt inoculated with PSAG-9 (right) and were plated onto MOPS-Gluc-Pi and MOPS-Gluc-Pt plates, respectively. The bacterial communities were distinguished by colony morphology. Picture was taken after 24 h.



Supplementary Figure S5.4: NADH production is detected in both wild type and PtxD-containing cell extracts. Wells were seeded with 0.3-0.45 mg/mL of total protein from clarified cell extracts. Values were normalized for total protein concentration in each extract. OD_{340} was measured for 90 min in 30-sec intervals. Error bars represent the standard deviation between biological triplicates and technical duplicates in each condition.

The rest of the supplementary material of this chapter, including Supplementary Tables S1 to S4, can be accessed *via*:

https://figshare.com/s/074a416fb06fbd6e081d

Bibliography

1. Fernández, M., Duque, E., Pizarro-Tobías, P., *et al.* Microbial responses to xenobiotic compounds. Identification of genes that allow Pseudomonas putida KT2440 to cope with 2,4,6-trinitrotoluene. *Microbial biotechnology* **2,** 287–294 (2009).
2. Martínez-García, E., Nikel, P. I., Aparicio, T., *et al.* Pseudomonas 2.0: genetic upgrading of P. putida KT2440 as an enhanced host for heterologous gene expression. *Microbial Cell Factories* **13,** 159 (2014).
3. Nikel, P. I., Martínez-García, E. & De Lorenzo, V. Biotechnological domestication of pseudomonads using synthetic biology. *Nature Reviews Microbiology* **12,** 368–379 (2014).
4. Kampers, L. F., Volkers, R. J. & Martins dos Santos, V. A. Pseudomonas putida KT 2440 is HV 1 certified, not GRAS. *Microbial biotechnology* **12,** 845–848 (2019).
5. Belda, E., Van Heck, R. G., José Lopez-Sanchez, M., *et al.* The revisited genome of Pseudomonas putida KT2440 enlightens its value as a robust metabolic chassis. *Environmental microbiology* **18,** 3403–3424 (2016).
6. Martin-Pascual, M., Batianis, C., Bruinsma, L., *et al.* A navigation guide of synthetic biology tools for Pseudomonas putida. *Biotechnology Advances* **49,** 107732 (2021).
7. Kampers, L. F. C., van Heck, R. G. A., Donati, S., *et al.* In silico-guided engineering of Pseudomonas putida towards growth under micro-oxic conditions. *Microbial Cell Factories* **18,** 179 (2019).
8. Mezzina, M. P., Manoli, M. T., Prieto, M. A., *et al.* Engineering Native and Synthetic Pathways in Pseudomonas putida for the Production of Tailored Polyhydroxyalkanoates. *Biotechnology Journal* **16,** 2000165 (2021).
9. Adams, B. L. The Next Generation of Synthetic Biology Chassis: Moving Synthetic Biology from the Laboratory to the Field. *ACS synthetic biology* **5,** 1328–1330 (2016).
10. Poblete-Castro, I., Becker, J., Dohnt, K., *et al.* Industrial biotechnology of Pseudomonas putida and related species. *Applied Microbiology and Biotechnology* **93,** 2279–2290 (2012).
11. Nikel, P. I. & de Lorenzo, V. Pseudomonas putida as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism. *Metabolic Engineering. Metabolic Engineering Host Organism Special Issue* **50,** 142–155 (2018).
12. Guan, Z.-j., Schmidt, M., Pei, L., *et al.* Biosafety Considerations of Synthetic Biology in the International Genetically Engineered Machine (iGEM) Competition. *BioScience* **63,** 25–34 (2013).
13. Whitford, C. M., Dymek, S., Kerkhoff, D., *et al.* Auxotrophy to Xeno-DNA: an exploration of combinatorial mechanisms for a high-fidelity biosafety system for synthetic biology applications. *J Biol Eng* **12,** 13 (2018).
14. Li, J., Zhao, H., Zheng, L., *et al.* Advances in Synthetic Biology and Biosafety Governance. *Frontiers in Bioengineering and Biotechnology* **9,** 173 (2021).
15. Schmidt, M. & de Lorenzo, V. Synthetic bugs on the loose: containment options for deeply engineered (micro)organisms. *Current Opinion in Biotechnology* **38,** 90–96 (2016).
16. Van de Poel, I. & Robaey, Z. Safe-by-Design: from Safety to Responsibility. *NanoEthics* **11,** 297–306 (2017).
17. Asin-Garcia, E., Kallergi, A., Landeweerd, L., *et al.* Genetic Safeguards for Safety-by-design: So Close Yet So Far. *Trends in Biotechnology* **38,** 1308–1312 (2020).
18. De Lorenzo, V., Krasnogor, N. & Schmidt, M. For the sake of the Bioeconomy: define what a Synthetic Biology Chassis is! *New Biotechnology* **60,** 44–51 (2021).
19. Kallergi, A., Asin-Garcia, E., Martins dos Santos, V. A., *et al.* Context matters: On the road to responsible biosafety technologies in synthetic biology. *EMBO Reports* **22,** e51227 (2021).
20. Kato, Y. An engineered bacterium auxotrophic for an unnatural amino acid: a novel biological containment system. *PeerJ* **3,** e1247 (2015).

21. Lopez, G. & Anderson, J. C. Synthetic Auxotrophs with Ligand-Dependent Essential Genes for a BL21(DE3) Biosafety Strain. *ACS synthetic biology* **4,** 1279–1286 (2015).

22. Gómez-Tatay, L. & Hernández-Andreu, J. M. Biosafety and biosecurity in Synthetic Biology: A review. *Critical Reviews in Environmental Science and Technology* **49,** 1587–1621 (2019).

23. Hirota, R., Abe, K., Katsuura, Z.-i., *et al.* A Novel Biocontainment Strategy Makes Bacterial Growth and Survival Dependent on Phosphite. *Scientific Reports* **7,** 44748 (2017).

24. Motomura, K., Sano, K., Watanabe, S., *et al.* Synthetic Phosphorus Metabolic Pathway for Biosafety and Contamination Management of Cyanobacterial Cultivation. *ACS synthetic biology* **7,** 2189–2198 (2018).

25. Jansson, M. Phosphate uptake and utilization by bacteria and algae. *Hydrobiologia* **170,** 177–189 (1988).

26. Morton, S. C. & Edwards, M. Reduced Phosphorus Compounds in the Environment. *Critical Reviews in Environmental Science and Technology* **35,** 333–364 (2005).

27. White, A. K. & Metcalf, W. W. Microbial Metabolism of Reduced Phosphorus Compounds. *Annual Review of Microbiology* **61,** 379–400 (2007).

28. Bisson, C., Adams, N. B. P., Stevenson, B., *et al.* The molecular basis of phosphite and hypophosphite recognition by ABC-transporters. *Nature Communications* **8,** 1746 (2017).

29. Guo, Z.-W., Ou, X.-Y., Liang, S., *et al.* Recruiting a Phosphite Dehydrogenase/Formamidase-Driven Antimicrobial Contamination System in Bacillus subtilis for Nonsterilized Fermentation of Acetoin. *ACS Synthetic Biology* **9,** 2537–2545 (2020).

30. Lei, M., Peng, X., Sun, W., *et al.* Nonsterile l-Lysine Fermentation Using Engineered Phosphite-Grown Corynebacterium glutamicum. *ACS Omega* **6,** 10160–10167 (2021).

31. Nikel, P. I., Chavarría, M., Martínez-García, E., *et al.* Accumulation of inorganic polyphosphate enables stress endurance and catalytic vigour in Pseudomonas putida KT2440. *Microbial Cell Factories* **12,** 50 (2013).

32. Harris, R. M., Webb, D. C., Howitt, S. M., *et al.* Characterization of PitA and PitB from Escherichia coli. *Journal of Bacteriology* **183,** 5008–5014 (2001).

33. Metcalf, W. W. & Wanner, B. L. Mutational analysis of an Escherichia coli fourteen-gene operon for phosphonate degradation, using TnphoA'elements. *Journal of bacteriology* **175,** 3430–3442 (1993).

34. Nelson, K. E., Weinel, C., Paulsen, I. T., *et al.* Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440. *Environmental microbiology* **4,** 799–808 (2002).

35. Motomura, K., Hirota, R., Ohnaka, N., *et al.* Overproduction of YjbB reduces the level of polyphosphate in Escherichia coli: a hypothetical role of YjbB in phosphate export and polyphosphate accumulation. *FEMS Microbiology Letters* **320,** 25–32 (2011).

36. Putker, F., Boxtel, R. T.-v., Stork, M., *et al.* The type II secretion system (Xcp) of Pseudomonas putida is active and involved in the secretion of phosphatases. *Environmental Microbiology* **15,** 2658–2671 (2013).

37. Wanner, B. L. Gene regulation by phosphate in enteric bacteria. *Journal of Cellular Biochemistry* **51,** 47–54 (1993).

38. Taylor, C. B. Nature of the Factor in Soil-extract Responsible for Bacterial Growth-stimulation. *Nature* **168,** 115–116 (1951).

39. Rao, N. S. S. *Soil microorganisms and plant growth.* (1995).

40. NIH Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules (NIH Guidelines) - April 2019, 149 (2019).

41. Shaw, A. J., Lam, F. H., Hamilton, M., *et al.* Metabolic engineering of microbial competitive advantage for industrial fermentation processes. *Science (New York, N.Y.)* **353,** 583–586 (2016).

42. Pan, S., Chen, G., Wu, R., *et al.* Non-sterile Submerged Fermentation of Fibrinolytic Enzyme by Marine Bacillus subtilis Harboring Antibacterial Activity With Starvation Strategy. *Frontiers in Microbiology* **10,** 1025 (2019).

43. Metcalf, W. W. & Wolfe, R. S. Molecular Genetic Analysis of Phosphite and Hypophosphite Oxidation by Pseudomonas stutzeriWM88. *Journal of Bacteriology* **180,** 5547–5558 (1998).

44. Wu, H., Kosaka, H., Kato, J., *et al.* Cloning and characterization of Pseudomonas putida genes encoding the phosphate-specific transport system. *Journal of Bioscience and Bioengineering* **87,** 273–279 (1999).

45. Montenegro, R., Vieto, S., Wicki-Emmenegger, D., *et al.* The putative phosphate transporter PitB (PP1373) is involved in tellurite uptake in Pseudomonas putida KT2440. *Microbiology (Reading, England)* **167** (2021).

46. Gebhard, S., Ekanayaka, N. & Cook, G. M. The low-affinity phosphate transporter PitA is dispensable for in vitro growth of Mycobacterium smegmatis. *BMC microbiology* **9,** 254 (2009).

47. Whitney, L. P. & Lomas, M. W. Phosphonate utilization by eukaryotic phytoplankton. *Limnology and Oceanography Letters* **4,** 18–24 (2019).

48. Cook, A. M., Daughton, C. G. & Alexander, M. Phosphonate utilization by bacteria. *Journal of Bacteriology* **133,** 85–90 (1978).

49. Lidbury, I. D. E. A., Murphy, A. R. J., Scanlan, D. J., *et al.* Comparative genomic, proteomic and exoproteomic analyses of three Pseudomonas strains reveals novel insights into the phosphorus scavenging capabilities of soil bacteria. *Environmental Microbiology* **18,** 3535–3549 (2016).

50. Zhang, X.-X., Scott, K., Meffin, R., *et al.* Genetic characterization of psp encoding the DING protein in Pseudomonas fluorescens SBW25. *BMC Microbiology* **7,** 114 (2007).

51. Fann, M.-C., Busch, A. & Maloney, P. C. Functional characterization of cysteine residues in GlpT, the glycerol 3-phosphate transporter of Escherichia coli. *Journal of Bacteriology* **185,** 3863–3870 (2003).

52. Huang, Y., Lemieux, M. J., Song, J., *et al.* Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli. *Science (New York, N.Y.)* **301,** 616–620 (2003).

53. Lemieux, M. J., Huang, Y. & Wang, D.-N. Glycerol-3-phosphate transporter of Escherichia coli: structure, function and regulation. *Research in Microbiology* **155,** 623–629 (2004).

54. Torres-Bacete, J., Luís García, J. & Nogales, J. A portable library of phosphate-depletion based synthetic promoters for customable and automata control of gene expression in bacteria. *Microbial Biotechnology* (2021).

55. Cui, J. Y., Gunewardena, S. S., Yoo, B., *et al.* RNA-Seq reveals different mRNA abundance of transporters and their alternative transcript isoforms during liver development. *Toxicological Sciences: An Official Journal of the Society of Toxicology* **127,** 592–608 (2012).

56. Eberl, L., Givskov, M., Sternberg, C., *et al.* Physiological responses of Pseudomonas putida KT2442 to phosphate starvation. *Microbiology (Reading, England)* **142,** 155–163 (1996).

57. Navarro Llorens, J. M., Tormo, A. & Martínez-García, E. Stationary phase in gram-negative bacteria. *FEMS microbiology reviews* **34,** 476–495 (2010).

58. Thao, H. T. B. & Yamakawa, T. Phosphite (phosphorous acid): Fungicide, fertilizer or bio-stimulator? *Soil Science & Plant Nutrition* **55,** 228–234 (2009).

59. Spaans, S., Weusthuis, R., Van Der Oost, J., *et al.* NADPH-generating systems in bacteria and archaea. *Frontiers in Microbiology* **6,** 742 (2015).

60. Blank, L. M., Ionidis, G., Ebert, B. E., *et al.* Metabolic response of Pseudomonas putida during redox biocatalysis in the presence of a second octanol phase. *The FEBS journal* **275,** 5173–5190 (2008).

61. Ramos, J. L., Duque, E., Gallegos, M.-T., *et al.* Mechanisms of solvent tolerance in gram-negative bacteria. *Annual Review of Microbiology* **56,** 743–768 (2002).

62. Johnsen, A. R., Wick, L. Y. & Harms, H. Principles of microbial PAH-degradation in soil. *Environmental Pollution (Barking, Essex: 1987)* **133,** 71–84 (2005).

63. Chen, C.-T., Chen, F. Y. .-.-H., Bogorad, I. W., *et al.* Synthetic methanol auxotrophy of Escherichia coli for methanol-dependent growth and production. *Metabolic Engineering* **49,** 257–266 (2018).

64. Schmidt, M. A metric space for semantic containment: Towards the implementation of genetic firewalls. *Bio Systems* **185,** 104015 (2019).

65. Asin-Garcia, E., Martin-Pascual, M., Garcia-Morales, L., *et al.* ReScribe: An Unrestrained Tool Combining Multiplex Recombineering and Minimal-PAM ScCas9 for Genome Recoding Pseudomonas putida. *ACS Synthetic Biology* (2021).

66. McDonald, A. E., Grant, B. R. & Plaxton, W. C. Phosphite (phosphorous Acid): Its Relevance in the Environment and Agriculture and Influence on Plant Phosphate Starvation Response. *Journal of Plant Nutrition* **24,** 1505–1519 (2001).

67. Thomason, L. C., Costantino, N. & Court, D. L. E. coli genome manipulation by P1 transduction. *Current Protocols in Molecular Biology* **Chapter 1,** Unit 1.17 (2007).

68. Wirth, N. T., Kozaeva, E. & Nikel, P. I. Accelerated genome engineering of *Pseudomonas putida* by I- *Sce* I—mediated recombination and CRISPR-Cas9 counterselection. *Microb. Biotechnol.* **13,** 233–249 (2020).

69. Neidhardt, F. C., Bloch, P. L. & Smith, D. F. Culture Medium for Enterobacteria. *Journal of Bacteriology* **119,** 736–747 (1974).

70. Damalas, S. G., Batianis, C., Martin-Pascual, M., *et al.* SEVA 3.1: enabling interoperability of DNA assembly among the SEVA, BioBricks and Type IIS restriction enzyme standards. *Microbial biotechnology* **13,** 1793–1806 (2020).

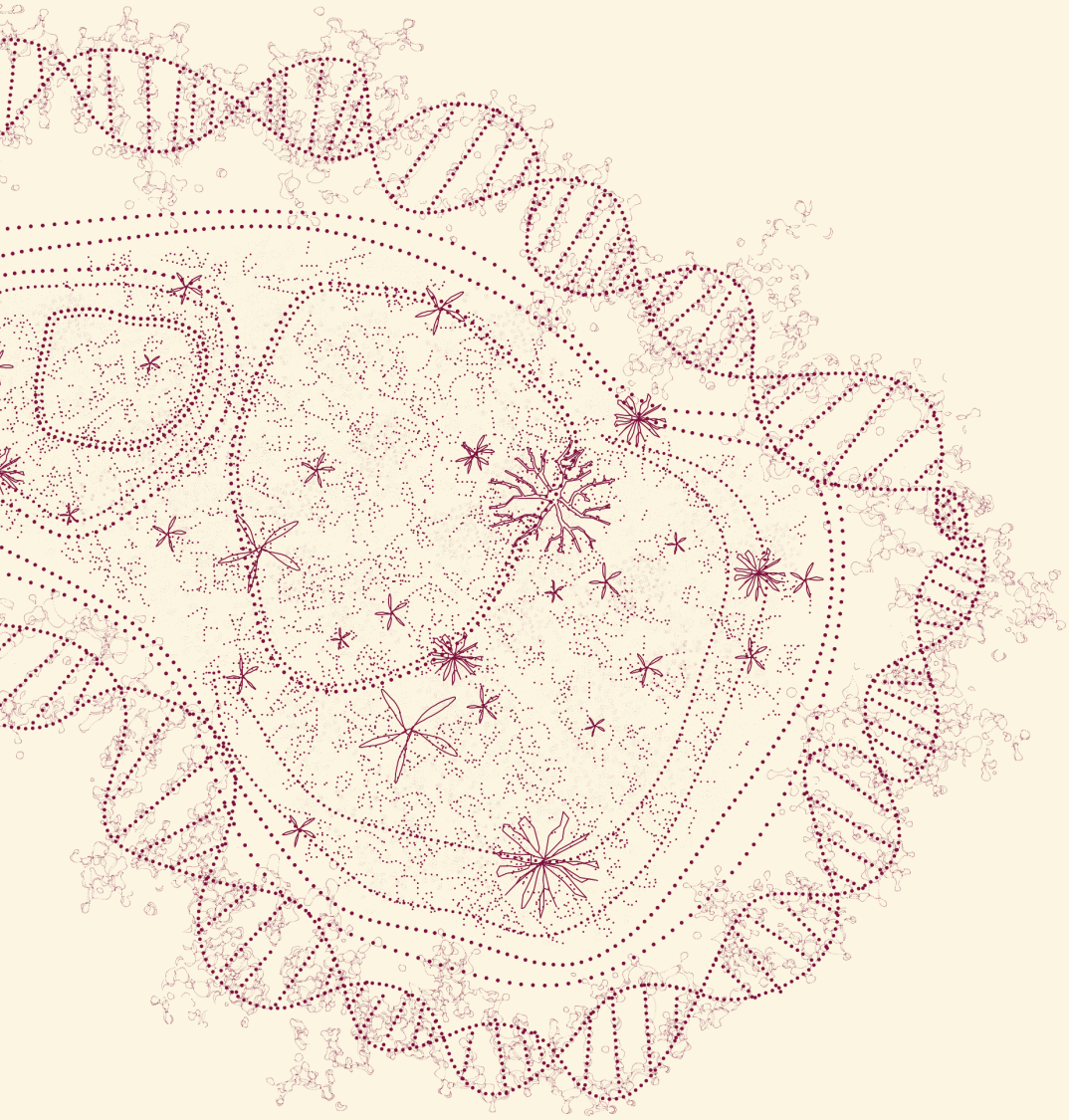
71. Green, R. & Rogers, E. J. Transformation of Chemically Competent E. coli. **529,** 329–336 (2013).

72. Hirota, R., Yamane, S.-t., Fujibuchi, T., *et al.* Isolation and characterization of a soluble and thermostable phosphite dehydrogenase from Ralstonia sp. strain 4506. *Journal of Bioscience and Bioengineering* **113,** 445–450 (2012).

73. Choi, K.-H., Kumar, A. & Schweizer, H. P. A 10-min method for preparation of highly electrocompetent Pseudomonas aeruginosa cells: application for DNA fragment transfer between chromosomes and plasmid transformation. *J Microbiol Methods* **64,** 391–397 (2006).

74. Volke, D. C., Friis, L., Wirth, N. T., *et al.* Synthetic control of plasmid replication enables target- and self-curing of vectors and expedites genome engineering of Pseudomonas putida. *Metabolic Engineering Communications* **10,** e00126 (2020).

75. Gallagher, R. R., Li, Z., Lewis, A. O., *et al.* Rapid editing and evolution of bacterial genomes using libraries of synthetic DNA. *Nat Protoc* **9,** 2301–2316 (2014).



CHAPTER 6

Genetic safeguards for Safety-by-Design: so close yet so far

Enrique Asin-Garcia*, Amalia Kallergi*,
Laurens Landeweerd, Vitor A. P. Martins dos Santos
* Contributed equally

Adapted for publication as:
Genetic Safeguards for Safety-by-design: so close yet so far,
2020, Trends in Biotechnology, 38, 1308–1312

Teaser

Safety-by-design (SbD) is paramount for risk management in synthetic biology, with genetic safeguards being a key technology for its implementation. While attractive in theory, the integration of genetic safeguards into SbD strategies is rarely exercised in practice, despite technological advances. Here we question why and what might be done about it.

Genetic safeguards: a tool for Safety-by-design in synthetic biology

Synthetic biology (synthetic biology) may yield breakthroughs of societal and economic importance for medicine, industry, agriculture, and the environment. However, its unique features considerably complicate the assessment and management of its risks. From the countless possibilities granted by advances in genome editing and DNA synthesis, to the access of amateur users to this technology, these critical aspects suggest a myriad of application scenarios and deployment contexts that should be approached with caution [1]. Accordingly, policymakers and regulatory bodies face demands for renewed regulation that adequately addresses innovation in synthetic biology [2]. As noted in the corresponding EU scientificopinion [3], current risk assessment expertise may be applicable to synthetic biology at present, but it will not suffice in the long term.

For an innovation to be deployed safely and responsibly, risks should be anticipated, assessed, and minimized as early as possible, preferably at the R&D and design phases. In emerging technologies, where uncertainty is high, this objective has been formalized under the concept of Safety-by-design (SbD). SbD is a preemptive approach to risk management that aspires to minimize risks by making safer design choices upfront. In the case of synthetic biology, this often materializes as built-in safety mechanisms implemented at the genetic level, during the design of the organism carrying out the operations envisaged [4]. Thus, in addition to new-to-nature biological entities, synthetic biologists are also building safety locks to genetically contain them. Biocontainment strategies (henceforth, genetic safeguards) were conceived in the 1980s [5, 6],but the onset of synthetic biology renders them increasingly relevant for both technical and societal reasons: more sophisticated designs for biocontainment are achievable [7] while the risks and benefits of synthetic biology increase the demand for built-in safety [1, 8]. This new momentum has resulted in a prolific collection of strategies (Box 6.1).

Box 6.1: A small synthetic biology library of genetic safeguards

An extensive repertoire of strategies has been produced, either to conditionally restrict the cell host viability to defined environments (biological isolation), or to prevent horizontal or vertical gene transfer between organisms and across species (genetic isolation) [8]:

Auxotrophy

The engineered host is not able to synthesize or incorporate an essential molecule, which therefore must be provided externally in a restrained environment. While the concept has been applied for more than 40 years, the list of new auxotrophs keeps ever growing for both natural and synthetic compounds. A remarkable recent example is the dependency on phosphite. By engineering the cell to transport only phosphite as the sole phosphorous source and by facilitating its conversion into the metabolizable phosphate, an extremely tight auxotroph was achieved [9]. To date, this strategy has resulted in the lowest escape frequency ever reported (1.9×10^{-13}).

Genetic circuits and switches

Synthetic regulatory modules are incorporated to limit the survival of the microbes to specific physicochemical or nutritional conditions. These modules can adopt different complexities and are typically based on toxinantitoxin systems, transcriptional factors, translational regulation, or signal balance. Two of the most intricate designs are 'DeadMan' and'PassCode' [10]. While the first plays with the unbalance between reciprocal repressors, the second is based on chimeric regulators. In both, inducer molecules perform as specific input signals that are coupled to cell survival.

Semantic containment

The genetic code of the cells is altered in order to avoid horizontal gene transfer. The new information might include non-canonical DNA or recoded codons that encode non-canonical amino acids, which, at the same time, must be provided externally constituting, in this way, an additional auxotrophic layer of biosafety. Projects of massive mutagenesis for rewriting genomes have amazed the scientific community during the last decade, but the total synthesis of a recoded one set a historical landmark [7], opening the way to any conceivable genome-wide design [11].

Combinatorial systems

Combinatorial systems include multiple strategies. While multi-layered systems might diminish escape frequencies, they may as well result in decreased fitness and survival levels. The strategy proposed by Gallagher and colleagues [12] combines essential gene riboregulation and cleavage of the bacterial genome in the absence of synthetic small molecules. By overlapping two different genetic circuits, they were able to reach significantly low levels of escape frequency (1.3×10^{-12}).

Technical solutions such as genetic safeguards should not be the sole SbD measure but do have the potential to facilitate its implementation. They are implicitly

part of SbD but their deployment beyond academic research remains uncertain. As SbD is slowly but steadily institutionalized, there is a discouraging lack of deliberation among stakeholders on the utility of genetic safeguards. Such interactive processes must occur at the early stages of research and innovation. As we will discuss, a pragmatic consideration of what safeguards can or cannot do for SbD is impeded by unproductive assumptions, mismatches in stakeholder perceptions, and practical bottlenecks.

From wishful thinking to reality check

Genetic safeguards are at an experimental stage and not yet widely used commercially. Nevertheless, in academic publications, policy drafts, and informal talk, an idealized image of genetic safeguards as drivers of innovation emerges (Figure 6.1). Many describe them as a precondition for a wider adoption of synthetic biology, especially non-contained applications in the environment or the human body [1, 6]. The EU scientific opinion acknowledges their potential [3] while the iGEM competition recommends them to its participating teams [8]. Others see them as key to unlocking the benefits (and markets) of synthetic biology. Built-in safety, the reasoning goes, can satisfy regulatory demands and influence public opinion positively. While tempting, such expectations are far from straightforward. Are we, the scientific community, overestimating the capacity of safeguards to provide a required level of safety; and perhaps confusing the absence of barriers (*i.e.*, risks) with positive drivers for innovation? Disconcertingly, such claims also reduce public concerns to safety concerns only, ignoring broader concerns for, among others, patenting, ownership, just distribution, naturalness, and the integrity of life. Finally, none of the factors that safeguards are supposed to affect positively (*i.e.*, safety, innovation, or public attitudes) can ever be assessed, evaluated, or agreed upon without some view on concrete application contexts.

Genetic safeguards are situated among a complex network of stakeholders, all of whom approach this technology from their particular standpoint. Synthetic biologists understand them as proofs of concept for solutions that do what is required (*i.e.*, limit growth of an engineered strain or prevent horizontal gene trans-

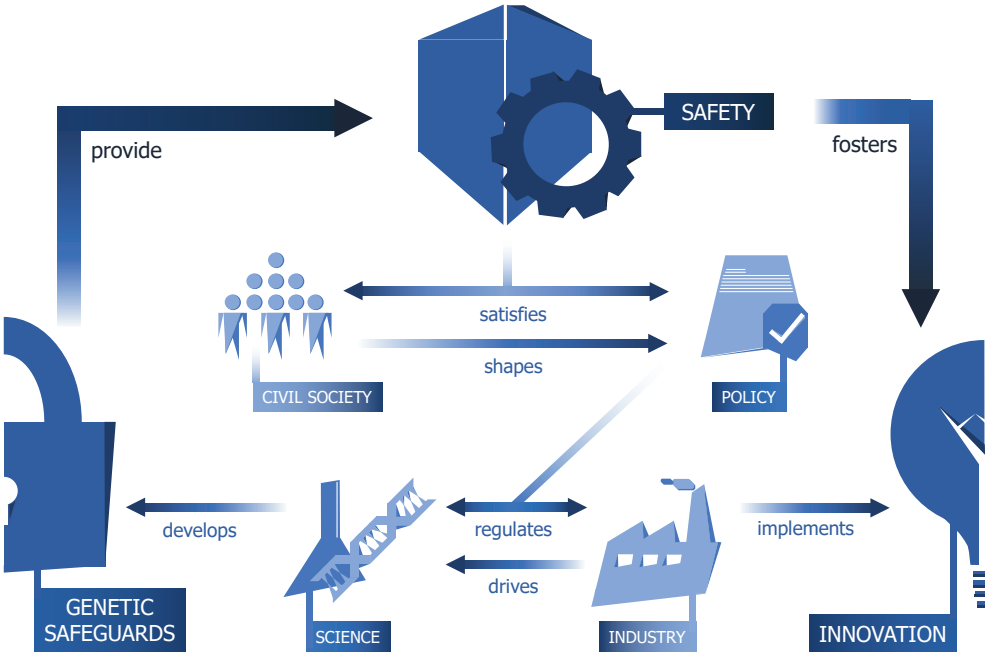


Figure 6.1: Idealized representation of genetic safeguards for Safety-by-design (SbD) and their assumed role in synthetic biology (synthetic biology) innovation. Genetic safeguards as mechanisms for safety are assumed to foster innovation by affecting the perceptions and actions of stakeholders. By increasing the safety of synthetic biology applications, genetic safeguards are expected to satisfy regulatory demands and to lead to positive public attitudes. Public attitudes and risk perceptions of emerging technologies are complex phenomena influenced by a range of individual and social factors; nevertheless, it is conceivable that genetic safeguards carry some persuasive power in communications about risk. Note that policy is represented here as a governing body attentive and responsive to public opinion. Also, note that this figure only captures the assumptions frequent in the dominant discourse surrounding genetic safeguards, not the reality or actual workings of innovation.

fer). Regulators and policymakers often consider safeguard research as a source of data for risk assessment and of metrics for policy; the integration of genetic safeguards in regulation remains distant and is likely to be context specific. Industry representatives consider them redundant for physically contained applications, which are sufficiently covered by current infrastructure and regulation; novel applications remain theoretical, as they do not yet qualify for risk assessment. Civil society is hardly represented while both developers and users express concerns about emerging public perceptions of safeguards.

This dissonance in stakeholder perceptions reveals a troubling lack of deliberation about needs and expectations. Rather than maintaining an idealist view, academic researchers must engage stakeholders in an early dialogue about the relevance and utility of safeguard technologies. Frameworks such as Responsible Research and Innovation [13] may contribute to research that is better attuned to its broader socioeconomic context, and the calibration of societal and economic forces required for innovation.

Practical bottlenecks

The much-needed deliberation over the utility of genetic safeguards is further hindered by two practical bottlenecks: (i) a knowledge gap in their assessment and; (ii) an evasion of responsibility regarding their implementation.

Knowledge gap

Safeguards are typically assessed under laboratory conditions according to a limited set of metrics, namely escape frequency and fitness. Different strategies have reported levels of escape frequency under the limit of detection with fitness equal to the wild type strains in laboratory settings [8, 9]. While encouraging, these results are far from sufficient. The detection of escapers is limited by the feasibility of the methods while fitness is only tested in a series of standard media, which makes these metrics a rather simplistic assessment of reality. More importantly, current practice is unprepared to evaluate the effects of evolution and the effects of the environment on the proposed strategies. Both factors are of paramount relevance, if genetic safeguards are to become a relevant component of a SbD strategy.

The intrinsic capacity of synthetic biology agents to evolve has obvious effects on the fitness of the produced strains and the functioning of safeguards. Existing systems tend to be bypassed by evolutionary forces and, except for the use of combinatorial strategies, the field has few ideas on how to overcome this obstacle. Research efforts focus on the development of evolutionary more stable systems with few alternative proposals suggesting the use of evolution itself as a driver for new biosafety rationales [14]. In any case, the lack of experimental methods to

monitor the development of the safeguard system remains the greatest challenge. An equally relevant matter is the evaluation of synthetic biology agents equipped with safeguards in their final environment. When this environment contains other organisms, possible consequences of every possible interaction should be considered and assessed, including competition, pathogenicity, predation, or susceptibility [1]. Specific applications of synthetic biology (*e.g.*, health, food/feed, and cosmetics) would require additional studies such as toxicity and allergenicity tests. Horizontal gene transfer (HGT) is another pressing issue exclusively considered in semantic containment strategies (Box 1), with only one recent framework for its risk assessment based on similarities among genetic codes [15].

Evasion of responsibility

Assuming that suitable metrics and methods are available, a second bottleneck lies in the transition from a proof of concept to a mature technology. This transition would require the commission of longitudinal field studies which directly raises questions of responsibility: who is responsible for conducting such studies and who is responsible for funding them? Within academic synthetic biology research, the virtual absence of long-term safety studies may be explained by pragmatic reasons: they require a multidisciplinary approach and are time-consuming, underfunded, and potentially non-prestigious. They may also be genuinely uninteresting academically. Underlying beliefs about the function of scientists in technology development may also play a role. Finally, it is plausible that prevailing mechanistic worldviews regarding the nature of life result in a neglect to seriously take into account risks associated with the complexity of life. Outside academia, interest in validating safeguard technologies remains low, presumably due to the lack of immediate benefits and the lack of imaginable applications. As a result, safeguard research is reaching a stalemate: all actors acknowledge the need for more data but none assumes responsibility for the required studies.

A way forward

Given the discussed bottlenecks, a convincing assessment of safeguards remains out of reach. In addressing our need for improved, further-reaching analyses and metrics, new technologies remain vital. High-throughput methodologies postulate as a promising avenue for higher resolution screening of conditions and escapers, by providing the possibility of testing safeguards in innumerable settings and by expanding the limits of detection. Yet, the field is unlikely to surpass its practical bottlenecks without stakeholder collaboration and investments, which, in turn, require a view on concrete application contexts. Finally, resolving practical bottlenecks will not automatically render safeguards relevant unless relevance and utility are co-created upfront, in a context-aware and socially responsive manner. Early stakeholder engagement, proper contextualization of genetic safeguards in concrete innovation scenarios and a discussion on the distribution of responsibility should guide safeguard research towards directions that make practical and societal sense.

Acknowledgements

This work is part of the research program SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis with project number 15814, which is financed by the Dutch Research Council (NWO). We would like to thank Dr Esther Kok and Tiny van der Werff for their valuable comments in the manuscript.

Conflict of interest

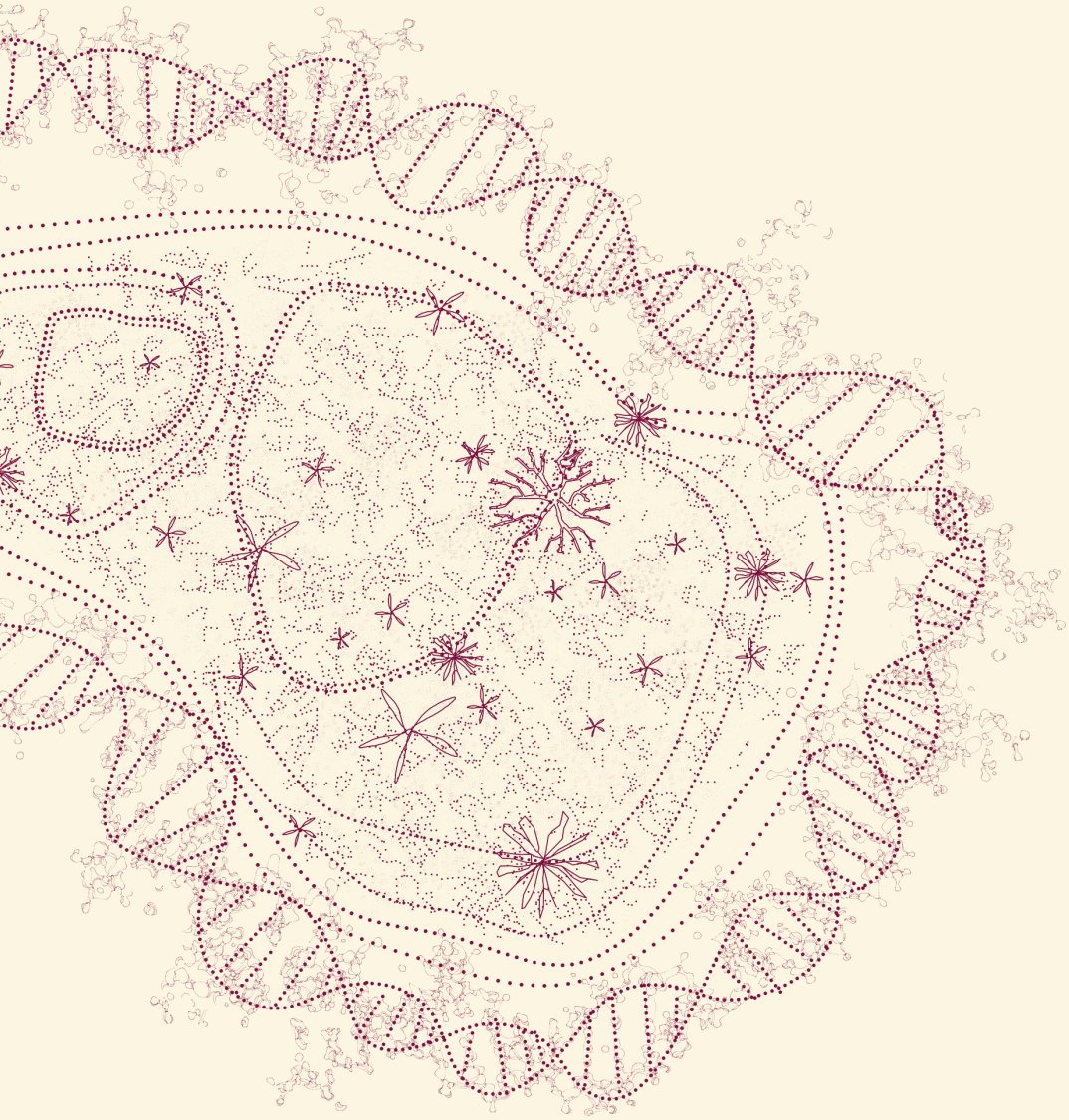
The authors declare there are no conflicting interests.

Author contributions

Conceptualization: EAG/AK; Formal analysis: EAG/AK; Funding acquisition: LL/VAPMdS; Investigation: EAG/AK; Supervision: LL/VAPMdS; Visualization: EAG; Writing - original draft: EAG/AK; Writing - review & editing: EAG/AK/LL/VAPMdS

Bibliography

1. Schmidt, M. & de Lorenzo, V. Synthetic bugs on the loose: containment options for deeply engineered (micro)organisms. *Current Opinion in Biotechnology* **38,** 90–96 (2016).
2. Trump, B. D., Cummings, C. L., Galaitsi, S. E., *et al. Synthetic biology: perspectives on risk analysis, governance, communication, and ELSI* (eds Trump, B. D., Cummings, C. L., Kuzma, J., *et al.*) 1–18 (2020).
3. Scientific Committee on Health and Environmental Risks (SCHER), Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR) & Scientific Committee on Consumer Safety (SCCS). *Opinion on Synthetic Biology II Risk assessment methodologies and safety aspects. European Comission* 2015.
4. Van de Poel, I. & Robaey, Z. Safe-by-Design: from Safety to Responsibility. *NanoEthics* **11,** 297–306 (2017).
5. Lee, J. W., Chan, C. T. Y., Slomovic, S., *et al.* Next-generation biocontainment systems for engineered organisms. *Nature Chemical Biology* **14,** 530–537 (2018).
6. Moe-Behrens, G. H. G., Davis, R. & Haynes, K. A. Preparing synthetic biology for the world. *Front. Microbio.* **4** (2013).
7. Fredens, J., Wang, K., de la Torre, D., *et al.* Total synthesis of Escherichia coli with a recoded genome. *Nature* **569,** 514–518 (2019).
8. Whitford, C. M., Dymek, S., Kerkhoff, D., *et al.* Auxotrophy to Xeno-DNA: an exploration of combinatorial mechanisms for a high-fidelity biosafety system for synthetic biology applications. *J Biol Eng* **12,** 13 (2018).
9. Hirota, R., Abe, K., Katsuura, Z.-i., *et al.* A Novel Biocontainment Strategy Makes Bacterial Growth and Survival Dependent on Phosphite. *Scientific Reports* **7,** 44748 (2017).
10. Chan, C. T. Y., Lee, J. W., Cameron, D. E., *et al.* 'Deadman' and 'Passcode' microbial kill switches for bacterial containment. *Nature Chemical Biology* **12,** 82–86 (2016).
11. Kuo, J., Stirling, F., Lau, Y. H., *et al.* Synthetic genome recoding: new genetic codes for new features. *Current Genetics* **64,** 327–333 (2018).
12. Gallagher, R. R., Patel, J. R., Interiano, A. L., *et al.* Multilayered genetic safeguards limit growth of microorganisms to defined environments. *Nucleic acids research* **43,** 1945–1954 (2015).
13. Schomberg, R. v. *A Vision of Responsible Research and Innovation* (eds Owen, R., Bessant, J. & Heintz, M.) 51–74 (2013).
14. Robaey, Z. *Dealing with risks of biotechnology: understanding the potential of Safe-by-Design.* Report commissioned by the Dutch Ministry of Infrastructure and Water Management (2018).
15. Schmidt, M. A metric space for semantic containment: Towards the implementation of genetic firewalls. *Bio Systems* **185,** 104015 (2019).



CHAPTER 7

CONTEXT MATTERS: ON THE ROAD TO RESPONSIBLE BIOSAFETY TECHNOLOGIES IN SYNTHETIC BIOLOGY

Amalia Kallergi*, Enrique Asin-Garcia*,
Vitor A. P. Martins dos Santos, Laurens Landeweerd,
* Contributed equally

Adapted for publication as:
Context matters: On the road to responsible biosafety technologies in synthetic biology,
2021, EMBO Reports, 22, e51227

Teaser

For a responsible development of genetic safeguards, a strategy of contextualization could confer practical benefits in regard of representation, assessment of safety and evaluation of social desirability. Here, we explore these benefits but also the potential tensions and potential caveats that such concept could bring along.

Preface

Synthetic biology abounds with engineering metaphors, from living machines to genetic code. Another key metaphor is the car: when broken down to its essential parts, the humble car may have something to teach us about safety in synthetic biology and about the significance of context for the development of biosafety technologies.

A microbial chassis is an essential organic structure that provides the frame for the design of a synthetic organism. In a direct analogy with a car chassis, a chassis cell can be augmented with desirable features and functionality. Synthetic biologists and bioengineers work on developing robust and appropriate chassis cells for real-life applications. Part of their effort is dedicated to developing *safer* chassis by implementing safety measures in their genome. Biocontainment strategies or genetic safeguards could reduce some of the risks of synthetic biology *e.g.*, by limiting the proliferation of the synthetic organism to controlled conditions or by limiting horizontal gene transfer. Much like seat belts in cars, genetic safeguards are built-in safety measures that could inherently increase the safety of the chassis and derived synthetic organisms.

Virtually every car manufactured today has seat belts installed by default. Yet, seat belts were not always a standard and required component of cars; in fact, they were not designed for cars to begin with. Originally conceived for aviation purposes, 2-point seat belts were somewhat unconvincingly introduced in motorized vehicles. It was not until the original aviation design was reconsidered in the context of driving that 3-point seat belts would become the life-saving equipment that they are today. A proper understanding of the context of application was, thus, one of the defining factors to their success and uptake. This is equally valid for safety measures at the genetic level of synthetic organisms.

The significance of context for technological innovation is evident in yet another car part, namely its wheels. A rotating object that was initially used for pottery led to a revolution in transportation and human civilization. Still, this revolution was far from self-explanatory. Consider for example the complete absence of wheeled vehicles in the pre-columbian Americas – presumably due to the absence of domesticated animals fit for carriages – or the dominion of camels over wheeled carriages

in Northern Africa. Wheeled transportation may appear obvious and inevitable today but it is the result of a combination of contextual factors, many of which would appear irrelevant to the function of the technological object.

Innovation processes in contemporary times are much different from the circumstances of prehistoric societies. Similarly, awareness of technological risks is much higher today than in the early days of car manufacture. Still, the histories of these now omnipresent objects remind us that no innovation is inevitable. Context matters, provides meaning and defines what a technological object will be best suited for and what its desirable features should be. As we are to discuss, what was true for wheels and seat belts may also be true for biosafety technologies such as genetic safeguards.

Genetic safeguards, responsibly

Scientific progress renders technological risks both more visible and more urgent. Seen from this perspective, genetic safeguards are a response to demands for safer innovations in synthetic biology [1]. Related to this objective is the Safety-by-design paradigm, a risk management approach that emphasizes minimization of risks at the early stages of research and development [2]. Inspired by the principles of inherent safety, where less hazardous options are favored over hazardous ones, Safety-by-design in synthetic biology encourages synthetic organisms that are designed with risk minimization in mind. Built-in safety as achieved by one or multiple genetic safeguards could, thus, contribute to inherently safer synthetic organisms and safer-by-design biotechnology innovations of economic and societal relevance. Yet, although academic research continues to produce promising results, the integration of genetic safeguards in real-life innovation scenarios remains unarticulated.

Safety-by-design places the value of safety at the core of technology development. Undoubtedly, this is an essential and praiseworthy requirement. However, to design for safety also implies to assess safety. This is practically impossible without a view on context of application. Importantly, risk assessment alone is no longer deemed sufficient for science policy. Contemporary societies increasingly

Table 7.1: RRI process dimensions as defined by the RRI-tools project (https://www.rri-tools.eu/about-rri). Other conceptualizations of RRI have been proposed but they all share a comparable emphasis on mutual responsiveness and shared responsibility.

Dimension	Description
Diverse & inclusive	Involve early a wide range of actors and publics in R&I practice, deliberation, and decision-making to yield more useful and higher quality knowledge. This strengths democracy and broadens sources of expertise, disciplines and perspectives.
Anticipative & reflective	Envision impacts and reflects on the underlying assumptions, values, and purposes to better understand how R&I shapes the future. This yields to valuable insights and increase our capacity to act on what we know.
Open & transparent	Communicate in a balanced, meaningful way methods, results, conclusions, and implications to enable public scrutiny and dialogue. This benefits the visibility and understanding of R&I.
Responsive & adaptive to change	To be able to modify modes of thought and behaviour, overarching organizational structures, in response to changing circumstances, knowledge, and perspectives. This aligns action with the needs expressed by stakeholders and publics.

demand that research and innovation are conducted in a socially attentive manner, *i.e.*, a manner that takes societal needs and values into account. During the past decades, ongoing developments in science governance have tried to capture and address this shift [3]. The latest example is the notion of Responsible Research and Innovation (RRI) [4], a broad framework to orient science governance that has become prominent in European science policy. In a nutshell, RRI calls for research and innovation that is conducted *for* and *with* society [5]. This means that both its processes and its products are aligned with societal needs, expectations and values and that this alignment happens in consultation with stakeholders. Note that safety may be one of these societal values but it is certainly not the only one that matters. Methodologically, RRI requires innovation trajectories that are diverse and inclusive, anticipative and reflective, open and transparent and responsive and adaptive (Table 7.1).

The rise of frameworks such as RRI reminds us that synthetic biology must anticipate and respond not only to its risks (or risks/benefits balance), but also to corresponding societal views that frame its evaluation. RRI can practically aid in coordinating such evaluation, provided there is a view on contexts of application. Consider as an example the so-called soft aspects of innovation: next to quantitative risk assessment, attention should be paid to qualitative aspects such as views on naturalness, individual autonomy, community rights etc. These RRI aspects can

only be assessed when concrete contexts of application become apparent.

Safeguard research may facilitate innovation through synthetic biology and may strengthen a responsible approach to innovation. Nonetheless, technical assessments of risk will need to be complemented by a view on possible beneficial applications and their socioeconomic context. This is relevant both for synthetic biology applications and for biosafety technologies. We suggest that an explicit strategy of contextualization, *i.e.*, the early emphasis on potential application contexts, can assist the development of genetic safeguards in a manner that complies with the priorities of RRI. Unfortunately, such an emphasis is noticeably lacking in current safeguard research: biocontainment strategies are assumed to increase the safety of synthetic organisms, but discussions about their use and applicability hardly ever become more specific than that. Proofs of concept are being designed, developed and evaluated irrespectively of their final context of application, with contextual considerations being postponed for some later stage. This lack of contextualization poses considerable obstacles to a responsible development of genetic safeguards or similar biosafety technologies.

To remain inclusive of and responsive to the values, needs and interests of stakeholders, the field needs to move from its abstract, one-size-fits-all vision to visions that are considerably more attentive to context. And whilst it is notoriously difficult to predict future applications, some view on the potential applications of synthetic organisms can facilitate present and future uptake in industry and SME. It is not until one has assessed the utility and desirability of potential applications that questions about safety can be defined and addressed.

Which context?

Synthetic biology agents are generally organisms designed to perform specific, dedicated actions and for a deployment site in mind. Genetic safeguards, on the other hand, are a family of biocontainment and control mechanisms that would theoretically be applicable to a variety of host cells, irrespectively of their application. As such, they are indicative of what Mampuys & Brom [6] describe as “horizontal integration” of biotechnology, namely the development of techniques

that are increasingly more versatile and less application-specific. To date, this vision has dominated laboratory research on genetic safeguards which remains largely context-agnostic: the body of related academic literature suggests that these strategies are designed, implemented and tested without a reference to a final application context. Henceforth, we define "context" as the final context of application of a "safelocked" synthetic organism. For the time being, let us agree that this context may be as broad as its corresponding branch of biotechnology (*e.g.*, green, red or white) or as specific as a deployment in a specified moment and location by a specified entity.

Note that our definition of context does not include the biological context (*i.e.*, the physical cell) in which a safeguard strategy is implemented. Genetic safeguards may be envisioned as "plug-ins" to be implemented in different chassis but, in reality, they are far less disembodied than this vision would suggest. To begin with, current laboratory practice still develops genetic safeguards in and for specific organisms; the effort required to implement an existing strategy in a new species is not neglectable. Secondly, different chassis are not interchangeable: the particularities of the organism a chassis originates from are not to be quickly dismissed with different microorganisms being praised for different properties. Eventually, it is reasonable to expect that different application contexts will dictate the use of different biological chassis [7]. While we acknowledge that biological context is relevant, it is admittedly a low-level, technical concept that may not be directly related to our present discussion.

While research on genetic safeguards is presently carried out without an explicit focus on applications, the academic discourse that surrounds it hints at a wide range of potential applications. In fact, any of the popularly envisioned application domains of synthetic biology, from medical to industrial to agricultural and environmental, could presumably benefit from the incorporation of genetic safeguards. Nonetheless, most gains are expected in novel and pervasive applications that are currently considered too risky, such as medical or environmental applications in uncontained settings [8, 9]. Clearly, the referenced domains vary greatly in their characteristics, from their deployment site and ecological scale of intervention to their perceived benefits and the human practices they will directly affect.

It should also be noted that our definition of context as the final context of application includes but is not limited to the final context of use. While multi-faceted, dynamic and relational, a context of use remains a bounded setting and one that involves those directly or indirectly interacting with the technological product. A context of application, on the other hand, is accessible to a broader set of actors. Moreover, it is unavoidably associated with the practices, beliefs, imaginaries and values that these actors maintain. We hypothesize that explicit mentions to application context will activate the corresponding frames of reference and associated prior knowledge of involved stakeholders, including an understanding of the broader societal context in which this application is deployed. Naturally, this aim is directly motivated by the tenets of science and technology studies (STS), where technological developments exist and can only be understood *in context*.

Given the diversity of scenarios in which genetic safeguards may be applicable, it is striking that they are expected to be designed, evaluated and regulated independently of their context. What is more, it is highly unlikely that (the same) strategies (or combinations thereof) will be equally useful, relevant and desirable across such grossly diverse contexts. After all, seat belts in passenger cars are distinctively different from seat belts in racing cars, a ski lift or a rollercoaster. As we are about to see, contextualization, *i.e.*, the early emphasis on potential application contexts, offers tangible benefits for our RRI efforts to capture and respond to the particularities of each context and the corresponding needs and interests of stakeholders.

Genetic safeguards, practically and realistically

The development of biosafety technologies, such as genetic safeguards, involves several scientific, ethical, legal and societal questions that are pending research and deliberation. We suggest that the efforts needed to resolve our pending questions will be both technically easier and qualitatively richer if they are organized around a specific context. Specifically, contextualization introduces concrete practical benefits both for laboratory research and for multi-stakeholder dialogue (Figure 7.1).

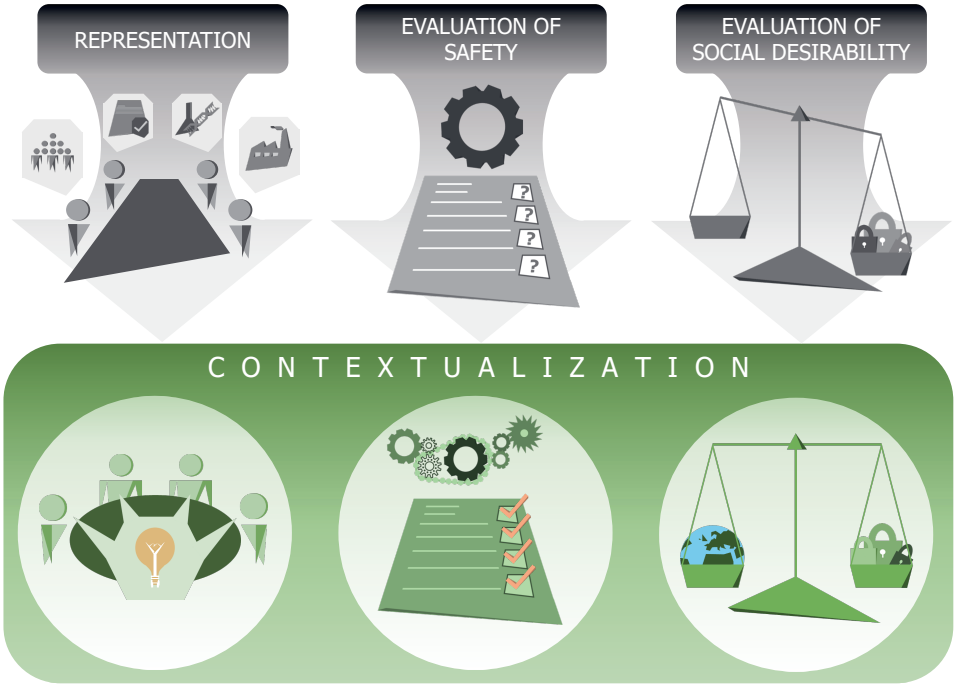


Figure 7.1: Practical benefits of contextualization for a responsible development of genetic safeguards.

Representation requires context

RRI emphasizes the value of early stakeholder engagement but a multi-stakeholder dialogue is as successful as its representative power. A sufficiently broad and yet manageable and representative participation is a challenging task that only becomes harder if context is absent. Furthermore, whose voices are being heard and whose stakes are being considered needs to go beyond the stereotypical categories of “science”, “industry” and “civil society”. Rather, we need more pluralistic and nuanced representations of these societal groups. Obviously, different application domains will demand interactions with different primary stakeholders. For example, the presence of medical practitioners and patient groups will be obligatory when negotiating the applicability of genetic safeguards to medical applications, whereas these groups may or may not have an interest in a deliberation process over agricultural applications of synthetic biology. To take this further, medical

applications themselves span across a diverse set of cases and practices that will require further delineation. For example, the experiences, insights and needs of medical professionals and patients will be noticeably different in the case of synthetic biology treatments for gut health as compared to ones for cancer.

An early focus on context should thus allow project leaders and researchers to identify and involve relevant stakeholders early and efficiently. It provides a manageable and practical way to perform stakeholder analysis, to effectively distribute our (often scarce) resources for dialogue and to go beyond stereotypical generalizations. It may also allow interested societal groups to express themselves from a position of engaged interest rather than as representatives of an amorphous body of *e.g.*, "interested citizens" or "industry representatives". In other words, contextualization should help establish a common ground for a multi-stakeholder dialogue that is a joint investigation on a shared problem rather than an articulation of abstract views.

Evaluating safety requires context

Research on genetic safeguards must provide satisfactory data about the safety and efficiency of the proposed strategies. This is not a straightforward task with laboratory practice being impeded by a lack of adequate metrics, methods and resources [10]. Firstly, it is practically impossible to validate safeguard strategies in every conceivable condition or application setting. An early focus on context should thus allow developers to specify finite and appropriate test conditions for a first evaluation of genetic safeguards in a given context or condensed cluster of contexts. Secondly, as laboratory test conditions and corresponding metrics are often insufficient for real-life applications, contextualization may provide a way to better organize research efforts on this front: it may enable us to better define necessary steps for future research and better divide responsibilities during the transition from development to application. Finally, the problem of funding (longitudinal) biosafety-related studies is not to be neglected. A clearer view on applications may allow safeguard research to engage relevant stakeholders with the power to fund or conduct these studies in their respective domains of expertise. It may also allow researchers to gather support and traction from societal groups with high stakes in

the given context.

At present, the role of genetic safeguards in the regulation of synthetic biology remains to be determined. Synthetic biology developments continue to test the limits of existing risk assessment protocols and risk management practices, with many voicing a need for a more flexible and adaptive risk governance [11]. That said, the paradigm of Safety-by-design, to which genetic safeguards directly contribute, opts for a mode of risk management that emphasizes prevention and risk minimization *via* appropriate design choices. In this case, the question of what constitutes a low enough risk or, equivalently, a safe enough alternative remains relevant. Obviously, this question is a highly contextual one. The practice of Safety-by-design as an up-front risk management tool (and which includes and builds upon risk assessment) begins with an accurate and complete formulation of the problem domain which requires contextual information about the organism, its application and its site of deployment. A "safelocked" application is likely to be evaluated in a similar manner, unless genetic safeguards evolve into a set of standardized and previously certified parts which will partially reduce the need for a case-by-case assessment of the synthetic biology agents equipped with them. Even so, it is highly unlikely that such a toolkit can be validated for every conceivable setting. Moreover, if genetic safeguards are to satisfy a generally agreed level of safety, we must decide whether this level should cater for the most demanding settings, which is bound to cause unnecessary overhead in less demanding ones, or for a minimum set of commonly shared risks, which will inevitably require additional measures per application.

Evaluating social desirability requires context

A technical assessment of safety as understood in risk assessment can inform governance by providing estimates of the probable risks of a "safelocked" synthetic biology application. However, it cannot reveal much about its social desirability. Between two equally "safe" (or equally unsafe) technologies, a society may choose to utilize only the one that strives to address sustainable development goals. Or it may choose to utilize the one that does so without offending core values such as justice, privacy or equality. Early stakeholder engagement is expected to inform such decision-making but, despite the best of our intentions, we might be asking

our stakeholders an impossible question: to formulate an opinion and make an assessment about the desirability of synthetic biology by and large, *i.e.*, without a view on what this technology is actually used for. Furthermore, with absolute safety guarantees being unattainable and with deep uncertainty being the distinguishing factor of emerging technologies, an assessment of synthetic biology must also consider the tension between uncertainty and a need to act or the risk of not acting in the face of a recognized danger.

In light of these observations, contextualization should allow us to organize dialogue and deliberation in a way that enables societal actors to take the particularities of each context into account. It should allow us to tap into our existing knowledge and make assessments about the ethical, social, economic and cultural impacts of a technology. It may also allow societal actors to agree on a conditional "go" for science and innovation, *i.e.*, one that ties a technology to its desired applications. More importantly, however, contextualization appears to be the only realistic way to an assessment that is not technology-oriented but problem-oriented. Risk assessors, ethicists and public participation experts all agree that our decision (or license) to act (or not) should be based on the evaluation of alternative solutions rather than the evaluation of the risks and benefits of an isolated technological option. Contextualization is, thus, the only strategy that can allow us to compare "safelocked" synthetic biology solutions to alternative solutions that are available in the same context.

Tensions and potential caveats

While contextualization comes with obvious practical benefits, one must also question whether these benefits are conditional and whether contextualization could, in fact, introduce new hurdles to the responsible development of genetic safeguards. What may be the risks of an early emphasis on context? And what tensions can we already anticipate regarding this approach?

We discussed contextualization as part of an RRI trajectory that values multistakeholder dialogue, clear distribution of responsibilities and knowledge co-creation. That said, an early emphasis on context implies, by definition, some form

of selection and prioritization. As such, it unavoidably raises questions about the motivations behind the choices made. For example, it is conceivable that contextualization becomes a strategy for persuasion or for the legitimization of genetic safeguards or synthetic biology. An early emphasis on application contexts could be easily misunderstood as yet another instance of overpromising or hyping the capacities of synthetic biology to generate funding or societal support. Urgent applications scenarios, in particular, while noble from the point of view of societal well-being, could also be used as a means to pressurize societal actors or to speed up regulatory reform. To counter these risks to responsible development, contextualization needs to be accompanied by a reflective mindset, a scientific agenda that is mutually shaped and a governance that favors resilience in the form of multiple, even redundant, solutions as opposed to silver bullets.

Contextualization offers useful constraints to organize research and dialogue but what may be missed by getting too specific too early? We will refrain here from revisiting the discussion on whether synthetic biology research should be conducted with a primarily fundamental or applied orientation. Rather, we are curious about any restraining effects when envisioning desirable futures and usages for genetic safeguards. It is possible that an early emphasis on context will favor known and accessible innovation scenarios at the expense of bolder or hypothetical ones. To counter this risk, contextualization may need to be complemented by speculative exercises with stakeholders, such as the use of future scenarios in technology assessment, or even by collaborations with creative forces such as artists and speculative designers. To prevent a preliminary technology lock-in, contextualization should once again be understood as part of an effort to generate and evaluate multiple alternative solutions for a given context.

As safeguards research embraces contextualization, it will inevitably face the very practical challenge of delineating an appropriate context. How specific should a context be for it to be useful? Or, alternatively, what is a meaningful clustering of application settings? Based on which factors should different contexts be aggregated? Developing genetic safeguards for a dedicated, strictly defined context should result in an optimal solution for that specific context but may also reduce the reusability of the developed solution. Besides, developing for too specific a

context will be practically similar to a case-by-case risk assessment of a synthetic organism, which diminishes the benefits of a standardized "safelocked" chassis. This problem of delineation points at a fundamental tension between contextualization and the basic pillars of the field, namely standardization, modularity and abstraction. Synthetic biology promotes standardization in the form of a hypothetical repertoire of standardized chassis applicable to multiple contexts. This further implies that any of the interchangeable modules or "plug-ins" created for these standardized cells should also be developed in a context-independent manner so as to work optimally everywhere, without worrying about every aspect of the module-chassis combination. In the case of genetic safeguards, this is a rather unrealistic expectation so some middle way is needed between this vision of universal "plug-ins" and the customization associated with contextualization. It remains to be seen what the appropriate balance is between an informative level of detail and a useful level of abstraction and aggregation. To complicate things more, the same context can be modeled in multiple ways as different stakeholders will be interested in different factors. Multi-stakeholder dialogue should, thus, prevent any blind spots during this process of modeling and aggregation.

Conclusion

Contextualization as the early emphasis on potential application contexts enables us to better manage a range of technical and societal questions relevant for a responsible development of genetic safeguards. Some of these questions are simply too broad and will require too many resources to be addressed in all conceivable settings and contexts. Contextualization, thus, provides a pragmatic way to effectively distribute our limited resources as well as to raise the investment of affiliated stakeholders, both monetary and in terms of engagement. Next, contextualization offers a means to improve the quality and relevance of our generated observations, *e.g.*, by requiring (ethical) assessments on concrete issues as opposed to abstract and de-personalized terms and by tapping into the local knowledge of participants. As such, it emerges as an essential strategy to meet the principles of RRI in the field of biosafety.

Our interest in contextualization corresponds with an interest in developing genetic safeguards from proofs of concept into biosafety tools that are usable and desirable in their respective contexts. Our proposed strategy is primarily a response to very practical considerations that impede this process. While contextualization does not provide all the answers on how to move forward, it could support us in devising appropriate and actionable roadmaps for our next steps per agreed application context. At the same time, contextualization comes with obvious implications about the future *use* of genetic safeguards as part of the risk governance of synthetic biology or of a Safety-by-design strategy. It implies that a one-size-fits-all level of biosafety is inappropriate, even irrelevant, and that built-in (multiple) safety measures will need to be customized to different contexts. The way these biosafety technologies will be deployed in practice remains open, but contextualization should bring us one step closer to co-creating their future in a manner that makes technical and societal sense.

Acknowledgements

This work is part of the research program SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis with project number 15814, which is financed by the Dutch Research Council (NWO).

Conflict of interest

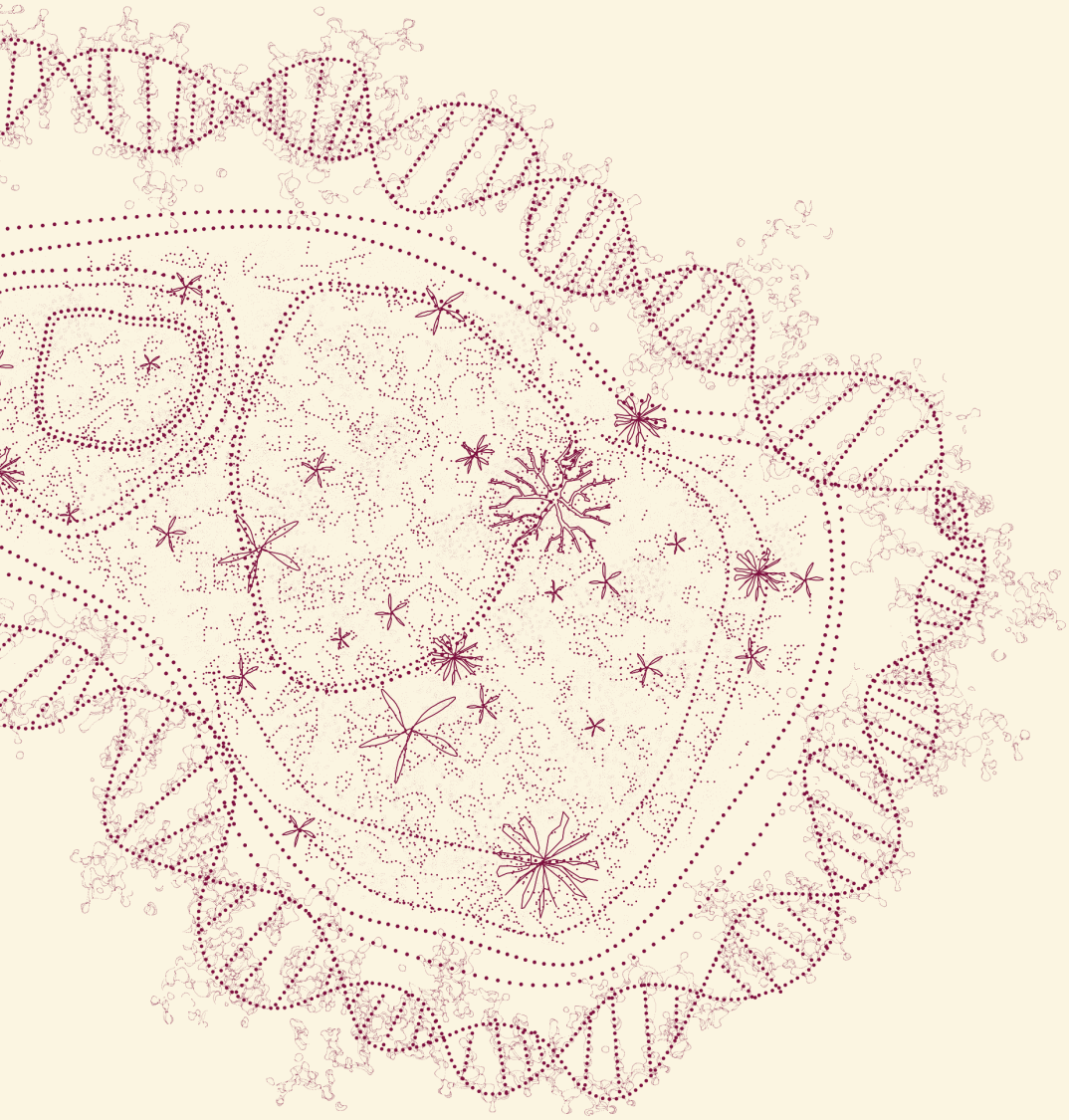
The authors declare there are no conflicting interests.

Author contributions

Conceptualization: AK/EAG; Formal analysis: AK/EAG; Funding acquisition: VAPMdS/LL; Investigation: AK/EAG; Supervision: VAPMdS/LL; Visualization: EAG; Writing - original draft: AK/EAG; Writing - review & editing: AK/EAG/VAPMdS/LL

Bibliography

1. EFSA Scientific Committee. Evaluation of existing guidelines for their adequacy for the microbial characterisation and environmental risk assessment of microorganisms obtained through synthetic biology. *EFSA Journal* **18,** 6263 (2020).
2. Robaey, Z. *Dealing with risks of biotechnology: understanding the potential of Safe-by-Design.* Report commissioned by the Dutch Ministry of Infrastructure and Water Management (2018).
3. Zwart, H., Landeweerd, L. & van Rooij, A. Adapt or perish? Assessing the recent shift in the European research funding arena from 'ELSA' to 'RRI'. *Life Sciences, Society and Policy* **10,** 11 (2014).
4. Schomberg, R. v. *A Vision of Responsible Research and Innovation* (eds Owen, R., Bessant, J. & Heintz, M.) 51–74 (2013).
5. Owen, R., Macnaghten, P. & Stilgoe, J. Responsible research and innovation: From science in society to science for society, with society. *Science and public policy* **39,** 751–760 (2012).
6. Mampuys, R. & Brom, F. Emerging crossover technologies: How to organize a biotechnology that becomes mainstream? *Environ Syst Decis* **38,** 163–169 (2018).
7. De Lorenzo, V., Krasnogor, N. & Schmidt, M. For the sake of the Bioeconomy: define what a Synthetic Biology Chassis is! *New Biotechnology* **60,** 44–51 (2021).
8. Moe-Behrens, G. H. G., Davis, R. & Haynes, K. A. Preparing synthetic biology for the world. *Front. Microbio.* **4** (2013).
9. Schmidt, M. & de Lorenzo, V. Synthetic bugs on the loose: containment options for deeply engineered (micro)organisms. *Current Opinion in Biotechnology* **38,** 90–96 (2016).
10. Asin-Garcia, E., Kallergi, A., Landeweerd, L., *et al.* Genetic Safeguards for Safety-by-design: So Close Yet So Far. *Trends in Biotechnology* **38,** 1308–1312 (2020).
11. Oye, K. A. Proactive and adaptive governance of emerging risks: The case of DNA synthesis and synthetic biology. *International Risk Governance Council, Geneva* (2012).



CHAPTER 8

Exploring the impact of tensions in stakeholder norms on designing for value change: the case of biosafety in industrial biotechnology

Enrique Asin-Garcia, Zoë Robaey
Linde F. C. Kampers, Vitor A. P. Martins dos Santos

Abstract

Synthetic biologists design and engineer organisms for a better and more sustainable future. While the manifold prospects are encouraging, concerns about the uncertain risks of genome editing affect public opinion as well as regulation. As a consequence, biosafety and associated concepts, such as the Safe-by-design framework and genetic safeguards technologies, have gained notoriety and occupy a central position in the GMO conversation. Yet, as regulatory interest and academic research in genetic safeguards technologies advance, the implementation in industrial biotechnology, a sector that is already employing engineered microorganisms, lags behind.

In this article, we explore the case of using genetic safeguards technologies for designing biosafety in industrial biotechnology. Based on our results, we posit that biosafety is a case of a changing value, by means of further specification of how to realize biosafety. Our investigation is inspired by the Value Sensitive Design framework, to investigate scientific and technological choices in their appropriate social context. Our findings discuss stakeholder norms for biosafety, reasonings about genetic safeguards, and how these impact the practice of designing for biosafety. We show that tensions between stakeholders occur at the level of norms, and that prior stakeholder alignment is crucial for value specification to happen in practice. Finally, we elaborate in different reasonings about genetic safeguards for biosafety building upon a mid-level ethical theory and conclude that, in absence of a common multi-stakeholder effort, the differences in informal biosafety norms and the disparity in biosafety thinking could end up leading to design requirements for compliance instead of for safety.

Introduction

Biotechnology, synthetic biology and biosafety

Biotechnology, and the field of synthetic biology (SynBio) in particular, have been described as promising technologies that at the same time raise concerns of uncertainty and risks to health and the environment [1–6]. In this paper, we turn our attention to engineered microorganisms in industrial biotechnology (IB), a mature sector with promises of decreasing environmental impacts, positive socioeconomic effects, competitive processes and products, and development in rural areas [2]. Developments in the last decade have superseded the initial hype brought by the onset of synthetic biology with impressive technological achievements [7] like the synthesis of working bacterial genomes [8], the development of computer-aided systems for logic circuit construction in bacteria [9], the CRISPR-Cas9 genome editing revolution [10] and the use of non-natural building blocks and non-canonical chemistry that gave rise to the subfield of xenobiology [11]. The development of these technologies has brought forth the concepts of trophic and semantic containment of engineered microorganisms, in other words, possibilities to build safety as an inherent feature of a microorganism. The idea of genetic safeguards technologies for biosafety and related possible implementations therefore emerged once again, as it had previously done in the eighties linked to the idea of uncontained genetically engineered organisms.

Biosafety then

Biosafety is a complex concept traditionally defined as the set of containment principles, facility design, practices and procedures to prevent occupational problems or the release of engineered organisms to a non-permissive environment [12]. Other definitions of biosafety focus directly on the risks and the capability of the engineered biological agents to cause disease, of greater or lesser severity, in humans, animals and plants [13, 14].

Almost 50 years ago, the Asilomar Conference set the foundations for the way recombinant DNA is dealt with by delivering a series of principles to handle potential biohazards. These guidelines mainly focused on making containment an

essential consideration in the experimental design and matching the effectiveness of such containment to the estimated risk. Moreover, it was indicated that given the difficulty of risk estimation, this would be intuitive at first and improved later as additional knowledge and technologies came into the picture. Thus, technology progress and new research developments would bring along the means to assess and balance risks with reexamined appropriate levels of containment [15].

The proposals of Asilomar were meant as a guide that included recommendations regarding, amongst others the development of safer vectors and hosts, laboratory procedures and education and reassessment, thereby providing the means for a structured and standard approach. As a consequence, the value of biosafety today builds largely on foundations and discussions that took place in the seventies, and slowly crystallized into policies in the nineties and early two-thousands. We observe however that the principles, designs and practices are changing, and therefore that biosafety can present a case of value change, where scientific knowledge about new ways to realize biosafety could afford new uses of engineered microorganisms and therefore the fulfillment of the benefits of new biotechnologies.

Biosafety now

In recent times biosafety is attracting more attention because it may be a limiting factor in the development of advanced technologies. Taking into account the pace and progress of biotechnology and synthetic biology, a lack of international norms is apparent, as well as a dearth of guidance at some of the value levels, mostly related to the risks that have emerged as a consequence of the progress of technology [16, 17]. As indicated, Asilomar principles resulted in norms that were established decades ago, but now more information related to biosafety seems to be necessary to set policy for the recent developments.

In a recent comprehensive review, Hewett *et al.* identified 44 discrete risks in synthetic biology: 18 of those related to human health, which were subsequently clustered into 4 main categories: allergies, antibiotic resistance, carcinogens, and pathogenicity or toxicity; and 26 related to the environment, later categorized into: change or depletion of the environment, competition with native species, horizontal gene transfer, and pathogenicity or toxicity [18, 19]. This points to the need for

further specification of issues when speaking of the value of biosafety.

Despite international agreements such as the Cartagena Protocol on Biosafety, legal regulation differs from country to country, the EU being known for its more stringent regulation. This has led to recurrent calls to reconsider how to do biosafety [20] (and in the related field of biosecurity [21]), suggesting that the specification of biosafety as a value is not only limited to understanding risks, but also extends to how biosafety is understood. One of the recent promising trends to deal with biosafety is centered around the concept of Safe-by-design, deriving from other technological fields [22]. Applying Safe-by-design (SbD) principles to biotechnology and synthetic biology offers a pre-emptive approach to risk management and aspires to minimize the risks of these technologies by making safer design choices during the early stages of the innovation trajectory, preferably at the R&D and design phases [23].

Genetic safeguard strategies, previously popular in the eighties and nineties, have experienced a new golden age this past decade resulting in a prolific repertoire of diverse approaches [24–26]. These technological investigations have been justified as inherent safety mechanisms to control the aforementioned risks of engineered microorganisms [27], a claim that has been underlined by policy in numerous occasions [28–30]. Biosafety is at the epicenter of discussions about governance and ethics of technology [6, 31, 32] and remains an essential component of the synthetic biology paradigm, as reflected by its importance in the International Competition on the Genetically Engineered Machine (the iGEM competition), the cradle of the synthetic biologists of the future [33, 34].

Nowadays, one could argue that approaches to safety and responsibility keep partly shaping the research agenda of the biotechnology and synthetic biology laboratories [35]. While this might be true for certain academic research projects, the situation is probably different for an already established sector like IB [2, 36–38], where biosafety might be simply understood as technical compliance to the current regulations about management of genetically modified organisms (GMO). References to strategies for increasing the safety of engineered or synthetic organisms are widespread, but discussions about their use and applicability hardly ever get more specific than that. Consequently, while the appreciation of genetic safeguard

technologies by policy and regulatory bodies continues developing, the formulation of such tools for real-life scenarios beyond academic research remains an unfulfilled ideal [39, 40].

Investigating a changing value

This paper explores this tension between the on-going scientific development of genetic safeguard technologies and their potential implementation in real world IB applications. This tension can also indicate a case of changing value. The type of value change we consider here is value specification according to van de Poel's taxonomy [41]. This paper seeks to further explicate how biosafety presents a case of value specification. In addition, because the first author of this paper is a bioengineer working on the design of said genetic safeguards in IB, we organize our investigation following a framework that lends itself to understanding scientific and technological choices in their social context. Value Sensitive Design (VSD) invites considering the role of the designer [42, 43] so this one can formulate design requirements that can answer to given norms to fulfill stakeholder values [44]. Norms can be formal and informal and capture the dominant way of doing things in a community, so the designer's role is embedded in this context of norms. While we do not follow a VSD approach to identify stakeholder values, we structure our findings in a VSD manner in order to investigate the changing nature of the value of biosafety. We ask: what are the stakeholder norms at play? How do these norms influence the understanding of the value of biosafety and stakeholder participation? What kind of limitations do these norms impose on design choices? How does this analysis inform the concept of value change?

Herein, we show that tensions between stakeholders occur at the level of norms. Formal and informal norms present a crucial component of the value hierarchy that needs to be addressed when designing for value change. We discuss this by first reviewing important aspects of VSD research, then we present results of qualitative interviews with practitioners in biosafety and IB, and finally we discuss stakeholder norms for biosafety, reasoning about genetic safeguards, how these impact the practice of designing for biosafety, and what this entails for designing for value specification.

Methods and findings

Designing for biosafety is akin to safety engineering [45] as it also delivers design options for safety like different types of barriers, of which genetic safeguards are the most recent and illustrative exemplar. In order to investigate the changing value of biosafety in IB and the existing design options, we structure our desk research and qualitative interviews following the steps of the VSD methodology as described by Friedman and colleagues [46] with conceptual, empirical and finally technical investigations. We do not do a VSD study as we concentrate on one value, but rather use the framework to understand dynamics of value hierarchy [44] in the context of a changing values, by looking at translation to norms and design requirements.

One recent contribution to the field of VSD is of great relevance to our case, namely design for value change where technical features of adaptability, flexibility and robustness help dealing with change [41]. When considering value specification as a case of value change, how do these features play a role? We consider stakeholders' reasoning about biosafety in order to further understand how the ethics of risk are understood in biosafety and identify norms and design requirements of biosafety. This also informs a discussion on ethical commitments [42, 43], not only of designers, but also of users in industrial biotechnology.

This section successively presents methods and results for conceptual, empirical and finally technical investigations.

Conceptual investigations: identifying peripheral issues to genetic safeguards

Methods

For this section, an exploratory qualitative approach was adopted to uncover the different stakeholders that interact with genetic safeguards and to investigate their particular stakes, roles and positions towards the technology. For the purpose of this study, stakeholder mapping was performed using literature review through a key word analysis of related scientific articles in combination with previous experience derived from workshops within the SafeChassis research project. To further enrich the mapping, stakeholders were categorized into direct or indirect according to those who make decisions and those who are affected by the decisions made.

Results

Our research finds a complex network of direct and indirect stakeholders is positioned around the concept of designing for biosafety. Direct stakeholders, namely researchers, regulators, risk assessors, policy makers and industry, perceive the technology of genetic safeguards in different ways, which do not always align with each other or with the mainstream academic discourse found in the literature (as argued more generally for SbD in [22]). First, biotechnologists and synthetic biologists envision and design genetic safeguards as creative ideas that might or might not be applied but are deemed capable of effectively addressing a given biosafety concern in the laboratory setting (*i.e.*, limit growth of an engineered strain or prevent horizontal gene transfer). An extensive collection of proof-of-concepts has been produced by researchers [26] but, despite technological advances, the integration of these tools into SbD strategies is rarely executed [31]. Second, regulators, risk assessors and policy makers can use technological designs, safeguard research and lab and field studies to collect data and information for risk assessment, establishment of metrics and policy making. Meanwhile, the integration of these designs and tools in regulation remains hypothetical and most probably will ultimately be context-specific and not universal [40]. Lastly, industry tackles physically contained applications with currently sufficient infrastructure and regulation, which makes genetic and biological isolation appear redundant for these settings. Non- or semi-contained applications of genetically engineered agents remain theoretical and, therefore, they are still not contemplated for industrial risk assessment [24] (Table 8.1).

While the aforementioned stakeholders have an influence or an interest in the design, development and application of the technology; there are two broad categories of indirect stakeholders: the public and the environment. As part of the public, we encounter civil society. Despite little evidence of its position on genetic safeguards for biosafety, there are concerns expressed about the use of genetic modification in IB [47]. In some cases, there is stark opposition against an IB innovation [48, 49], while in other cases, the social debate around biotechnology and synthetic biology concentrates precisely on the balance between the technology's benefits and the biosafety [22, 50]. Taking into account biases and factors of cog-

Table 8.1: Overview of stakeholders surrounding the biosafety technologies based on the conceptual investigations.

Stakeholders		Relation with biosafety technologies	Challenge
Direct	Researchers (biotechnologists, synthetic biologists)	Development, proof-of-concepts to verify hypotheses	Research does not continue beyond laboratory settings
	Regulation and policy organisms (regulators, risk assessors, policy makers)	Source of data for risk assessment and, ultimately, policy making	Integration in real applications is not executed
	Industry	Compliance with regulation	Current regulation is already covered by current infrastructure
Indirect	Public (civil society, workers)	Balance between technology's benefits and harms (potential health problems)	Difficulties to assess the risks given the underdevelopment of the technology
	Environment	Balance between technology's benefits and harms (HGT, hazardous "superbugs", etc.)	

nitive or even sociocultural nature, the public is still the group that bears the potential biosafety risks, specifically for workers, and for human health in general [51, 52]. Yet, the impacts of genetic safeguards and other potential biosafety designs are difficult to assess because the early technical designs generally do not resemble the version that will ultimately be adopted in the final application context [6] (Table 8.1).

In IB, genetically engineered organisms are meant to be physically contained in bioreactors, semi-contained or non-contained at all. The impacts of intentional and the risks of unintentional release position the environment as an indirect stakeholder. Upon release to non-permissive environments, engineered microorganisms are historically not expected to survive in the long run due to their laboratory domestication which likely results in a lower ecological fitness and higher vulnerability to competitors and predators [53]. Nonetheless, these engineered microorganisms might live long enough to transmit engineered DNA to other microbes altering the genetic structure of the ecosystem. This horizontal gene transfer phenomenon represents not only an ecological hazard but could result in the generation of "superbugs" carrying antibiotic resistance genes in nature [19].

The risk assessment of potential environmental impacts of engineered microbes will depend on the particular engineered function, the equipped genetic safeguard and the context of application. The complexity of biotechnology and synthetic biology together with our limited understanding of natural microbial communities and ecosystems bring about an enormous degree of uncertainty, which in itself also constitutes a critical challenge [54].

This first step of conceptual investigation into the use of genetic safeguards for biosafety in IB underlines that all stakeholders are engaging with the value of biosafety, or are concerned by its absence. In addition, while genetic safeguards are a technology that mostly remains within the confines of academic research, they do seem to provide solutions for potential risks for indirect stakeholders like the public and the environment.

Empirical investigations: one value of biosafety, several meanings

Methods

In this empirical investigation, we investigate stakeholders' views about biosafety in IB, with a focus on practitioners in the biotechnology and synthetic biology fields. We interviewed representatives of industry (9), academia (4), regulatory bodies (2) and technology transfer experts (2) from May 2019 to June 2021 with two researchers present. Interviewees were recruited considering their experience (senior position within the company/institution) and professional domain. The final list of participants (Ntot = 17) included companies based in Europe or America from a variety of sectors (pharmaceutical, food, chemical or production organism development industries), academics from Europe whose expertise lies in microbial biotechnology and who are engaged in collaborations with multiple industrial sectors, and regulatory bodies and offices from the Netherlands and Denmark (Table 8.2). After having interviewed approximately half of our participants, the content of the provided answers started to show overlap with previous interviews indicating data saturation. In addition, we found consistency and complementarity of our results with recent research in stakeholder perception of SbD in biotechnology [22], which led us to consider our list of participants a representative sample for our exercise. In this phase of the study, we focused on the direct stakeholders with a relation to the industrial context.

The interviews followed a semi-structured approach with open-ended and general questions to prevent bias and to allow for a more in-depth discussion and included questions focusing on the concepts of genetic safeguards' implementation and utility, SbD and perceptions of risks and uncertainty. In addition, the use of appreciative inquiry [55] in the questions encouraged and inspired the partic-

Table 8.2: Overview of stakeholder groups with their expertise and location.

Sector	Area	Location		Total
		Europe	USA	
Industry	Pharmaceuticals	1	1	2
	Food	1		1
	Chemicals	1		1
	Development of production platforms	1	1	2
	Combination	2	1	3
Academia	IB and SynBio	4		4
Public sector	Regulation	1		1
	Policy making	1		1
	Technology transfer	2		2
				17

ipants to answer according to their own perspectives, ideas and experiences, as opposed to following a strict interview structure. With the prior verbal and written consent of the participants, all interviews were audio-recorded for subsequent transcription (intelligent verbatim style). In order to eliminate bias, transcripts were pseudonymized and encoded with a letter (I = industry, A = academic, R = regulation, policy or technology transfer representative) and a number, before data analysis which was later performed using QCA map [56] and following the six steps for qualitative data analysis [57]. Based on predefined top-down codes and bottom-up codes derived during the transcription of the interviews, three main themes were formulated: A) meanings and norms of biosafety; B) reasoning about genetic safeguards; and, C) implementation of genetic safeguards.

Results

A) Meanings and norms of biosafety

The value of biosafety has shown many different meanings throughout the course of our investigation. This plurality soon became apparent from the norms ascribed by our participants to biosafety. In Table 8.3 we present our findings of four main norms of biosafety: I) Compliance with regulation; II) Evaluation of microbes; III) Responsibility; and, IV) Assessment of risk and uncertainty.

The norms collected in Table 8.3 reflect the plurality of the different understandings of biosafety encompassing categories that sometimes do not even vary along the same underlying dimensions (some of them are rules, some are individual prac-

tices, and some are general assumptions). While most of them could be considered complementary and particular facets of the value, some of them appear contradicting, which originates the possibility of value tensions (*e.g.*, historical argument of biosafety *vs*. scientific uncertainty).

Genetic safeguards are not found in our empirical investigation. However, one could envision genetic safeguards as part of the biological containment (norm of compliance), or as part of multi-actor responsibility by giving deliberate attention to biosafety since the conception of a strain's engineering (norms of responsibility).

B) Reasoning about genetic safeguards

Stakeholders reason differently about the need of genetic safeguards as design requirements for biosafety. Table 8.4 collects the main arguments presented under the following aspects identified from our analysis: I) Influence of stakeholders on industrial GMOs; II) Incentives of industry; and, III) Missing elements for safer IB.

C) Implementing genetic safeguards

To determine the position of genetic safeguards in the stakeholders' ecosystem, participants were asked about the suitability and feasibility of implementing genetic safeguards in IB (Figure 8.1 A), and about the new opportunities that they could bring (Figure 8.1 B). When it comes to implementation, we find overall mixed positions on those issues, ranging from positive, neutral and negative.

Overall, our empirical findings highlight heterogeneous ascriptions to different norms of biosafety, to reasoning about biosafety and to implementing genetic safeguards. This is striking because when it comes to biotechnology, certain stakeholder groups are typically associated with certain views on safety. This is, for example, the case of the stereotypically concerned and scared public, the very strict regulators or the careless corporations blinded by the money.

Table 8.3: Main norm groups representing the different stakeholders' views of the value of biosafety together with the corresponding design approaches. Each group is accompanied by a representative quote and sources supporting or stating each of the design approaches are shown in the last column with a numerical code (I = industry, A = academic, R = regulation, policy or technology transfer representative).

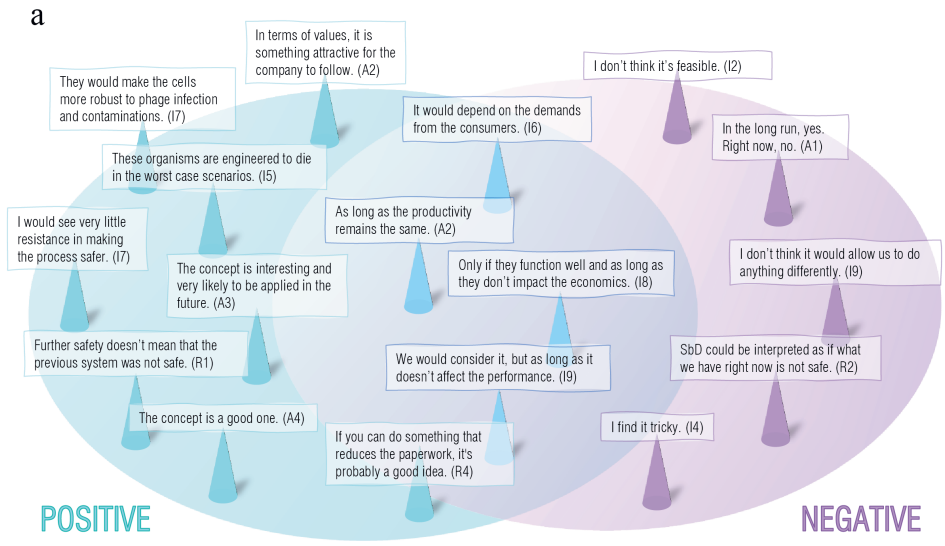
Norms	Design approach	Description	Source
Compliance with regulation ***"The physical containment is sufficient to comply with regulations, and that is the key thing, we need to comply with regulations."*** **(I2)**	Biological containment	Use of hosts microorganisms with a reduce host range, with natural or genetically modified characteristics that diminish their invading capacity or virulence, self-inactivating vectors, etc.	I3, I4, I6, R1, A4
	Physical containment	All the physical barriers that prevent or minimize the escape of the microorganisms from the controlled settings.	I2, I3, I6, R2, A4, I8
	GMO-free products	Separation of producer and products and inactivation of the biomass.	I3, I4, I5, A2, R1, I8
	Historical argument of biosafety	Engineered strains retain the biosafety category granted to their ancestors.	I2, I7, A3, I8, I9
	Regulation	Biosafety committees that take care of specific controls and standards. Additional approvals and bigger dossiers than other bioprocesses.	I2, I5, I6, I7, A3, R1, R2, I8, I9, R4
Evaluation of microbes ***"The whole thing is that you should really study your microorganisms carefully and monitor things and be aware of what could happen."*** **(A4)**	Study of introduced genetic elements	Monitorization of stability and mobility of introduced genetic elements.	I3, I9
	Sequencing	Sequence check of plasmids and full genome.	I4, I5, I6, I7, I9
	Other assays	Growth, productivity and fitness assays.	I5, I7, I8, I9

Responsibility ***"That's why the Safe-by-design concept try to promote a proactive approach by actors so that the government doesn't need to solve problems afterwards, but that the actors who develop something, who innovate, who develop a new technology or new application, think about the safety aspects during that process."*** **(R2)**	Multi-actor responsibility	Proactive responsibility at all stages of the process. Safe-by-design framework.	I4, R2, R3
	Cellular barcoding	Accountability through identification of labelled cells through space, time and even cell division which allows the instant access to all the information associated to a particular construct including its origin, its nature, if it is sensitive to antibiotics, what countermeasures one could take, etc.	A3, I9
Assessment of risk and uncertainty ***"If we work in a reasonably well-established organism like E. coli or Saccharomyces, I would say that we do not do any test to see whether it's safe or not. We basically go off the literature where others have done those tests already."*** **(I8)**	Domestication	Human selection of strains to obtain cultivated variants that thrive in artificial niches and meet specific requirements. During this process, microbes become more efficient in consuming particular nutrients, coping with research- or industry-specific stress factors, and producing the target compounds, but this usually comes at the cost of a dramatic decrease of fitness in their natural environment.	I4, I6, I9, R1
	Scientific uncertainty	Uncertain risks beyond the imposed norms and extra measures.	I2, I4, I6, A3, A4, R2, I8, R3
	Non-fitting assessment	Current regulation does not cover all the aspects of the technology.	A2, R2

Table 8.4: Subthemes affecting the establishment of stakeholders' norms with their respective supporting arguments. Each subtheme is accompanied by a representative quote and sources supporting or stating each of the arguments are shown in the last column with a numerical code (I = industry, A = academic, R = regulation, policy or technology transfer representative).

Subthemes	Arguments concerning	Arguments	Source
Stakeholders' influence on industrial GMOs ***"The applications of industrial biotech are so broad, and they're increasing. They're coming out different areas, different sectors, etc. that there are different situations. And the rules are changing. You know, as public perception changes, and people become more comfortable sometimes with the guidelines."*** **(I5)**	Public	Uses industrial GMO products daily.	I4, I5
		Perceives risks in GMO products.	I2, I4, I5, R1, I9, R3
		Demands GMO-free products.	I3, A2, R1, R3
		Responds to marketing campaigns.	I6, R2, R3
		Thinks that the technology should be more understandable and simpler, as well as transparent, communicative and honest.	I4, A3, R1, R2
		Argues that more biosafety might imply that the former products were not safe.	I7
	Environmental	Naturalness argument (natural is more desired than engineered).	I3, I8
		GMO products are more sustainable.	I3, I8, R3
	Specificity	Territorial differences: GMOs are better seen and less regulated in USA than in Europe.	I5, I6, A2, I8, R3, R4
		Industrial field differences: GMOS are better seen in chemical industry than in health and food industries.	A1, A2, R3
	Regulation	Makes a balance between harms and benefits.	A1
		Dictates what products reach the market.	A3, R2, R3
		Needs to track responsibility.	A3, R2
		Encounters secrecy and a non-continuous dialogue with industry.	R2
	Other companies	IP prevents the use of some techniques.	I4

Incentives of industry ***"Their difficulties are 'what do I need to deliver in order to get the approval of being undertaking that activity in my plant?' 'How much time and effort does it cost in the Netherlands, and is it worthwhile setting up a plant here, or should I do it abroad where perhaps it costs me less information and less money and the permits are given earlier?' because, in R&D, time is of the essence because the competition is everywhere."*** **(R2)**	Marketing	New markets	I2, R2
		Good image	A2, R2, R3, R4
		Communication, proactivity and honesty	I4, R2
	Production	Cheaper costs	I2, I3, R1, R2
		Minimal alteration of the production chain	I2, I7, A4, I8, R3
		Better (or at least the same) titers, rates and yields	I4, I6, A4, I8
		More robust strains and with better lifestyle	I7, A3
		Sustainable bioprocesses	I4, I6, R2
	Safety	Intrinsic safety	I4, A3, R2
		Broader view in risk assessment	A2, R1
		Tracing responsibility (barcoding) to prevent industrial espionage	A3, I9
Missing elements for safer IB ***"I think that safety for release of environmental bacteria have to first overcome the problem number one: how to put them in an environment that works. And when it works, we can discuss about safety issues. But let's face it, the limitation has never been the prohibitions or the regulations for their release, the limitation has always been the efficacy."*** **(A3)**	Technical part	High-throughput and automation	I2, A3
	Communication	Education of costumers	I3, R2, R3
		Rebranding of GMOs	I3, R3
	Regulation	New, more fitting regulation	A1, R2
		Holistic approach	R2
	Safety	Genetic containment	I5, I7, R1, R2
		Further research on genetic containment and gap analysis	I5, I7, R1, R2, A4, R3
		Biosafety metrics	A3, R2
		Successful stories	A3, R3, R4
		Barcodes and elements of responsibility	I6, A3, R2, I9
		Prudence	R2



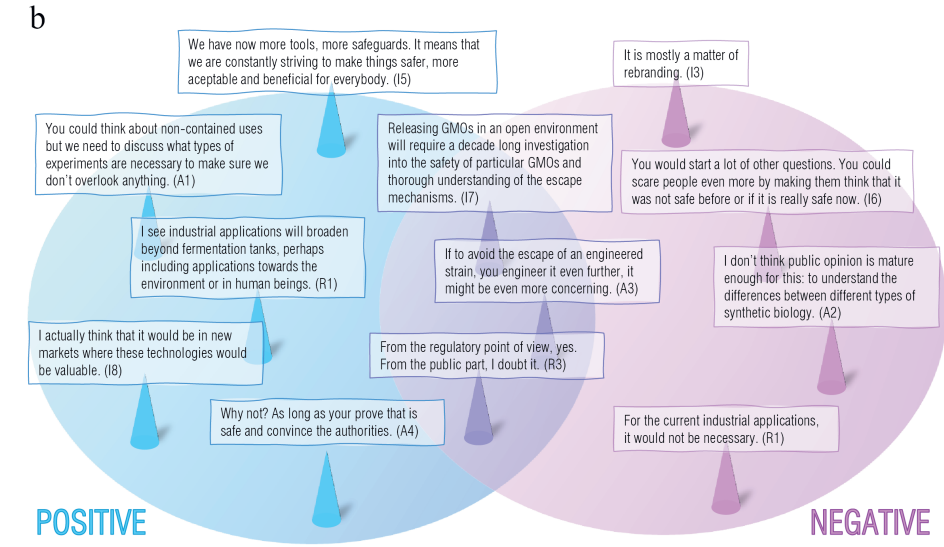


Figure 8.1: Venn diagrams containing participants' views on suitability of feasibility of genetic safeguards (A) and the opportunities of this technology (B) in the field of IB. The schematic overview represents the positive opinions in the left circles and the negative considerations in the right circles, while mixed aspects are found in the overlapping regions. Each statement has been directly obtained or simplified from a representative quote whose source is indicated with a numerical code (I = industry, A = academic, R = regulation, policy or technology transfer representative).

Technical investigations: biosafety, a value ready to change

Methods

In this section we analyze the properties of the genetic safeguards as technical measures to address the value of biosafety. For this purpose, we reviewed the scientific literature on the existing technology and connected the known technical features with the value. Moreover, we relied on our background knowledge in addition to the gathered information from literature to pinpoint what the limitations to support the committed value are and how these could be resolved though new biosafety design requirements.

Results

Since the second half of the XIX century, biosafety practices have been implemented in response to potential risks associated with the exposure to microorganisms cultured in laboratories. Over the years, more and more protective measures against biological risks have been developed and adopted, typically combining physical containment, working practices and personal protective equipment, and focusing mainly on occupational safety [13]. With the advent of recombinant DNA and cloning, experts started recommending an extra layer of biological containment on top of the previous barriers [15]. Shortly after, the U. S. National Institutes of Health (NIH) published the first "Guidelines for Research Involving Recombinant DNA Molecules" [58] which worked as a baseline for many of the present regulations on contained use.

In this paper, the focus of the technical investigations lies on the scientific properties of those biological and genetic layers of containment and their underlying mechanisms related to the value of biosafety. As we have mentioned before, the widespread view is that these built-in safety genetic safeguards can satisfy regulatory demands by reducing the risks and uncertainty of biotechnology and synthetic biology. This concept however remains mostly speculative given the limited experimental data to support or refute the notion [59].

Risk assessment and reduction of risks should be considered as quantitatively as possible. This would be achieved not only by using the appropriate containment tools but also with a series of well-supplied quantitative metrics and robust assess-

ment methodologies. However, these remain scarce. At present, risks are typically considered qualitatively, where the probability of an adverse outcome is expressed as more or less likely than a comparative scenario (*e.g.*, wild type comparators) [60]. Special mention should be made of the evaluation of the escape frequency for a given genetic safeguard. Escape frequency is, together with the strain fitness, one of the few quantitative metrics available and the most popular across the biocontainment literature and expresses the probability that a microbe equipped with a genetic safeguard has of escaping the permissive conditions. Nonetheless, the detection limit to assess escape frequencies is currently not low enough and it would need to be significantly lower for assessing genetically engineered organisms that are not intended to be physically contained (*e.g.*, intentional environmental release or human body).

Added to this, metrics should be international and standardized in order to provide sufficient incentives to commit the resources required to achieve high levels of biosafety in laboratories, companies and institutions. [16, 17, 30]. Data cannot be compared if they are obtained in different media or different contextual circumstances, which means that beyond further metrics, standard protocols to implement them and conduct them are also necessary [61]. Furthermore, it is practically impossible to test and validate beforehand safeguard strategies for every conceivable condition or application setting, which implies that assessment of their safety must be, in a higher or lower degree, contextual [40]. This early focus on context should reduce the infinite test conditions to an attainable number, which would likely be still large, given that any ecosystem is, by nature, dynamic. The key to address this point is developing technology to test genetic safeguards in a very large number of environmental conditions. Lorenzo and Schmidt proposed as a solution the generation of thousands of microenvironments in small droplets. By making use of micro and milli-fluidics devices, one could create a myriad of scenarios with specific and fluctuating parameters (oxygen concentration, temperature, pH, humidity, etc.) to test engineered organisms equipped with genetic safeguards in a high-throughput manner. Even though complete ecosystems cannot be completely mimicked in a small drop, this first approach could be the source of copious amounts of valuable information [61].

Table 8.5: Technical issues for biosafety risk assessment with genetic safeguards.

Issues	Current limitations	New biosafety design requirements
Risks are mostly considered qualitatively	Not enough metrics and data for risk assessment	Quantitative or semiquantitative new metrics beyond escape frequency
Data is obtained in different contextual circumstances	No consensus on how to measure biosafety and which are the expected levels	Standardized protocols for assessment
Standard validation should be performed in infinite scenarios	It practical terms, it is impossible to test in infinite application settings	Contextualization (reduction of the number of test conditions)

These developments indicate that biosafety has the potential to be redefined at a technical level, indicating a case of value specification. This section also underscores the limitations of using these technical measures, which also require more research attention (Table 8.5).

Discussion

The importance of norms in value change

In this research, we identified a tension between the renewed interest in available design options procured by the synthetic biology field for biosafety presented in the introduction and the apparent potential for implementation (Figure 8.1). By considering and applying the notion of value hierarchy [44] to biosafety, we honed in on the sources of this tension, which we found located at the level of norms. While the conceptual investigation underlined the interest for genetic safeguards, the empirical investigations, uncovered a heterogeneous understanding of various norms of biosafety independent of membership of specific stakeholder groups where the role of genetic safeguard was yet to be defined. We clustered these norms in four categories, namely compliance with regulation, evaluation of microbes, norms that attend to responsibility, and assessment of risk and uncertainty (Table 8.3). We can further analyze these norms in terms of formal and informal norms. For instance, “Compliance with regulation” encompasses formal norms codified in laws and regulations. The other three groups of norms identified consist of informal norms that are sometimes used, but not consistently across stakeholders, and without being formally codified. This results in disparities. For example, the monitoring of the uncertain aspects of engineered organisms collides with the idea of domestication. While the former aims at considering unknown risks, the latter reasons that industrial strains are not fit for proliferation out of the laboratory setting and therefore

not immediately concerning. In the same line than domestication, we find the historical argument of safety that allows engineered strains to retain the biosafety category of their ancestors despite the many unknowns associated to this practice. Later in the technical investigation and the results depicted in Figure 8.1, it became apparent that for value specification to happen in practice, there must be a prior stakeholder alignment.

The tension we identify can therefore be explained by formal norms that may stifle the possibilities for novel design requirements that stem from informal norms of biosafety practice. This tension at the level of norms brings technical innovations at odds with societal needs. Indeed, formal norms contribute mainly to current regulatory compliance, instead of making space for other stakeholder norms found in practice. This affects the diversity of views that can get as fundamental as the own orientation of the value. For most, the strains would be generally perceived as the risk actor of the discussion, whereas some stakeholders switch the direction and interpret them as the sensitive element subjected to the risks (industrial espionage, secret and property), what would certainly derive in very different design requirements. Moreover, elements of responsibility will differ as well depending to whom this one is attributed, whether it falls to the developer or to all actors involved.

When designing for value change, our findings indicate that inclusion and alignment of stakeholders should be considered as an essential part of the flexibility that van de Poel presented as one of the technical features that allow to better deal with value change, next to adaptability and robustness [41]. Flexibility is understood in terms of different possibilities for using a design. In an absolutely predictable situation, a designer would not need or want the design to be flexible as its function and the values that it is to meet are known. However, when dealing with non-predictable scenarios and changing values like biosafety, flexibility is imperative as it is the involvement of relevant stakeholders to help envisioning how to better meet certain values in the new circumstances [41]. Complementary to the necessary but inflexible formal norms, other stakeholder norms become indispensable to refine the use of the technology for a changing value.

Designing for biosafety is an exemplar for value specification as understanding of risks expands and adjusts according to the increasing demands for biotechno-

logical innovations. Current trends in the field such as cultures of microbial communities [62], accelerated evolution and genetic diversification [63], and a shift towards robust non-traditional chassis [64] step away from the classical definition of domesticated microbes and call for a more holistic approach regarding safety. Additionally, most of the norms depicted in Table 8.3 refer to an IB that is carried out in physically contained settings. However, the future prospects of synthetic biology and IB incorporate applications beyond the bioreactor, including non-controlled environments such as the open field or the human body. One can expect that as scientific knowledge on GMOs increases, awareness and judgement of how to reach safe scenarios will also change accordingly.

Reasoning about genetic safeguards for biosafety: building a mid-level ethical theory for biosafety

This study emerges as an initiative of synthetic biology designers to reflect on the goal of the biosafety research carried out within the Dutch TTW funded project SafeChassis. The VSD literature calls for ethical commitment of designers [42, 43]; since, in practice, designers often find themselves not only having to design for competing values [65], but also dealing with competing stakeholder norms as highlighted in the previous section. Using what Jacobs and Huldtgren [42] call explanatory power, justificatory power, and simplicity and practicability, we find that such categories can streamline a discussion regarding a changing value where informal norms can be made further explicit, and contribute to designing for a changing value with the general goal of safer IB.

First, the explanatory power constitutes of insights on the purpose and status of morality, and how the principles, rules and rights ultimately relate to obligations and norms [66]. As we have indicated, these norms respond to some of the nuances of the biosafety spectrum and are often guided by stakeholders' principles and morals. In our particular case, public opinion plays a fundamental role, not only affecting the norms, but also shaping some of the design requirements of the operations. Public perception is the main pillar of the marketing campaigns and hence prevention of a negative image ends up being one of the most important drivers during the design of a business model in IB (Table 8.4, Figure 8.1). Previous

studies on opinion towards GMOs showcase that certain events or scandals might impact public opinion in the long term and affect later policy decisions [4, 67, 68]. Arguments concerning the public on Table 8.4 also illustrate the general belief that the public perceives risks in GMO products and consequently demands GMO-free products. However, some more recent studies report that public views do not always align with experts' opinions and that editing technologies are sometimes portrayed as positive [69, 70]. This, of course, depends on different factors including the cultural context and the type of industry and product. While the stigma around GMOs is greater in Europe than it is in the United States, the use of engineered microorganisms for bioproduction in the field of IB is equal in the two regions, despite the *a priori* different legislations and norms. Nonetheless, more considerations besides the territory arise as one gets closer to anything that approaches consumer products. The use of GMOs for the generation of chemicals is better seen than their use for medicines and drugs, and this one would yet be better ranked than their employment for food production, which correspondingly derives too in different obligations and norms (Table 8.4, Stakeholders' influence on industrial GMOs, Specificity). Secondly, the justificatory power is based on the justification and argumentation regarding moral claims to judge or legitimize value prioritizations. Arguments such as naturalness, risk and uncertainty have historically been used to justify rejection towards GMOs [71], whereas sustainability, innovation and an overall positive balance of benefits versus harms have played the opposite role. In the matter of genetic safeguards, the negative and neutral quotes presented in Figure 1 showcase how the application of this technology can be simply deemed unfeasible or even superfluous when the process is already considered safe enough. In these grounds, prioritizations of other values such as productivity or economic interests over technologies for biosafety will always be legitimized. Furthermore, and even if these other major interests would remain unaffected, investing in genetic safeguards could be interpreted from two opposite angles. One could either see them as a way to make the process safer, or as a way to make it merely safe, implying that previous processes were not safe in the first place. This argument constitutes the main opposition towards the idea of genetic safeguards. If biosafety would be positioned not as a standalone, but as a complementary value that would benefit other

interests such as economic growth and broader applications (Figure 8.1 B), judgement of the genetic safeguards' technology would be adjusted. Genetic safeguards could then cease to appear superfluous and represent the only necessary design requirement for biosafety in non-contained application contexts [24]. In addition, some of these strategies' features hold the promise of being profitable for other aspects of the industrial process by, for example, decreasing some of the costs [72, 73], or increasing the production yields [74].

This leads to the third desiderata, simplicity and practicability. As we present in this research, genetic safeguards are not yet widely used in real-world applications despite being a powerful innovation for biosafety. To legitimize their use, genetic safeguards first need to be ready to deliver. The arguments concerning missing elements presented in Table 8.4 and the technical issues described in Table 8.5 underline the need for further efforts and research for genetic safeguards to be able to yield the expected outcomes towards the value of biosafety. Discussions on the changing value of biosafety therefore need to take design issues as seriously as stakeholders' norms.

Regarding stakeholders' norms, we find there are tensions at the level of norms accompanied by a variety of justifications and explanations for their need. Our research underlines that unless this is a common effort with multiple stakeholders, the differences in informal biosafety norms and reasoning about biosafety will inevitably lead to design requirements that seek compliance [39], a formal norm, instead of innovation in design for safety.

Conclusions

A possible limitation of our research was to look at genetic safeguards broadly in IB, since our results suggest that biosafety design requirements depend on the application setting. Further research should consider designing for biosafety in a context-specific application [40] that would lead to design changes [75] and would consider not only the development of technologies but rather their entire lifecycle [76]. Despite being largely an educational project, this point is exemplified by the iGEM competition teams throughout thorough studies of genetic safeguards in their

very specific innovations [31].

Due to the identified tensions at the level of stakeholder norms and arguments *e.g.*, compliance *vs.* innovation, sustainability *vs.* naturalness, uncertainty *vs.* domestication, responsibility from or towards engineered strains, etc., we suggest that efforts recommended in RRI to achieve stakeholder alignment [77] can also benefit a designing for changing values. Since the value of biosafety has significant societal impacts, we consider crucial stakeholder participation during the process to discuss concerns and to formulate alternative solutions [78]. Our research suggests that these insights should also be extended to discussions about biosafety norms in order to address tensions stemming from formal and informal norms, and design for the value, in our case, of biosafety.

Acknowledgements

This work is part of the research programme SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis with project number 15814, which is financed by the Dutch Research Council (NWO). This work is also part of the research programme "Virtues for Innovation in Practice (VIPs): A Virtue Ethics Account of Responsibility for Biotechnology" with project number VI. Veni.191F.010, which is financed by the Dutch Research Council (NWO). Additionally, the project received funding (VMdS) by the EU-2020 programme Industrial Biotechnology Innovation and Synthetic Biology Accelerator (IBISBA, under grant agreements No 730976 and 871118). We would like to thank Amalia Kallergi (TU Delft) and Henk van den Belt (Wageningen University & Research) for their valuable comments in the manuscript, as well as to express our gratitude to all our interviewees for their participation in this research, for their time and for sharing their views on the subject.

Conflict of interest

The authors declare there are no conflicting interests.

Author's contributions

Conceptualization: EAG/ZR/LK/VAPMdS; Data curation: EAG/LK; Formal analysis: EAG; Funding acquisition: ZR/VAPMdS; Investigation: EA/ZR; Methodology: EAG/ZR; Resources: VAPMdS; Supervision: ZR/VAPMdS; Visualization: EAG; Writing – original draft: EAG/ZR; Writing – review & editing: EAG/ZR/LK/VAPMdS

Bibliography

1. Andrianantoandro, E., Basu, S., Karig, D. K., *et al.* Synthetic biology: new engineering rules for an emerging discipline. *Molecular Systems Biology* **2,** 2006.0028 (2006).
2. Fröhling, M. & Hiete, M. *Sustainability and Life Cycle Assessment in Industrial Biotechnology: A Review of Current Approaches and Future Needs* (eds Fröhling, M. & Hiete, M.) 143–203 (2020).
3. Linkov, I., Trump, B. D., Anklam, E., *et al.* Comparative, collaborative, and integrative risk governance for emerging technologies. *Environment Systems and Decisions* **38,** 170–176 (2018).
4. More, S., Bampidis, V., Benford, D., *et al.* Evaluation of existing guidelines for their adequacy for the microbial characterisation and environmental risk assessment of microorganisms obtained through synthetic biology. *EFSA Journal* **18,** e06263 (2020).
5. Remer, S., Ang, S. H. & Baden-Fuller, C. *Dealing with Uncertainties in the Biotechnology Industry: The Use of Real Options Reasoning* SSRN Scholarly Paper ID 3371923 (Social Science Research Network, Rochester, NY, 2001).
6. Trump, B. D., Cummings, C. L., Galaitsi, S. E., *et al. Synthetic biology: perspectives on risk analysis, governance, communication, and ELSI* (eds Trump, B. D., Cummings, C. L., Kuzma, J., *et al.*) 1–18 (2020).
7. Meng, F. & Ellis, T. The second decade of synthetic biology: 2010-2020. *Nat Commun* **11,** 5174 (2020).
8. Gibson, D. G., Glass, J. I., Lartigue, C., *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *science* **329,** 52–56 (2010).
9. Nielsen, A. A. K., Der, B. S., Shin, J., *et al.* Genetic circuit design automation. *Science* **352** (2016).
10. Jinek, M., Chylinski, K., Fonfara, I., *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337,** 816–821 (2012).
11. Schmidt, M., Pei, L. & Budisa, N. *Xenobiology: State-of-the-Art, Ethics, and Philosophy of New-to-Nature Organisms* (eds Zhao, H. & Zeng, A.-P.) 301–315 (2018).
12. Nordmann, B. D. Issues in biosecurity and biosafety. *International Journal of Antimicrobial Agents. Global Trends of Emerging Zoonotic Diseases* **36,** S66–S69 (2010).
13. Beeckman, D. S. A. & Rüdelsheim, P. Biosafety and Biosecurity in Containment: A Regulatory Overview. *Frontiers in Bioengineering and Biotechnology* **8,** 650 (2020).
14. Kelle, A. Synthetic biology and biosecurity: From low levels of awareness to a comprehensive strategy. *EMBO reports* **10** (S1 2009).
15. Berg, P., Baltimore, D., Brenner, S., *et al.* Asilomar Conference on Recombinant DNA Molecules. *Science* **188,** 991–994 (1975).
16. Kwik Gronvall, G. A Biosafety Agenda to Spur Biotechnology Development and Prevent Accidents. *Health security* (2017).
17. Kwik Gronvall, G. & Rozo, M. Addressing the Gap in International Norms for Biosafety. *Trends in Microbiology* **23,** 743–744 (2015).
18. Hewett, J. P., Wolfe, A. K., Bergmann, R. A., *et al.* Human Health and Environmental Risks Posed by Synthetic Biology R&D for Energy Applications: A Literature Analysis. *Applied Biosafety* **21,** 177–184 (2016).
19. Wang, F. & Zhang, W. Synthetic biology: Recent progress, biosafety and biosecurity concerns, and possible solutions. *Journal of Biosafety and Biosecurity* **1,** 22–30 (2019).
20. Trevan, T. Biological research: Rethink biosafety. *Nature* **527,** 155–158 (2015).
21. Evans, S. W., Beal, J., Berger, K., *et al.* Embrace experimentation in biosecurity governance. *Science* **368,** 138–140 (2020).

22. Bouchaut, B. & Asveld, L. Safe-by-Design: Stakeholders' Perceptions and Expectations of How to Deal with Uncertain Risks of Emerging Biotechnologies in the Netherlands. *Risk Analysis* **40,** 1632–1644 (2020).
23. Van de Poel, I. & Robaey, Z. Safe-by-Design: from Safety to Responsibility. *NanoEthics* **11,** 297–306 (2017).
24. Asin-Garcia, E., Kallergi, A., Landeweerd, L., *et al.* Genetic Safeguards for Safety-by-design: So Close Yet So Far. *Trends in Biotechnology* **38,** 1308–1312 (2020).
25. Schmidt, M. & de Lorenzo, V. Synthetic bugs on the loose: containment options for deeply engineered (micro)organisms. *Current Opinion in Biotechnology* **38,** 90–96 (2016).
26. Whitford, C. M., Dymek, S., Kerkhoff, D., *et al.* Auxotrophy to Xeno-DNA: an exploration of combinatorial mechanisms for a high-fidelity biosafety system for synthetic biology applications. *J Biol Eng* **12,** 13 (2018).
27. Schmidt, M. Xenobiology: A new form of life as the ultimate biosafety tool. *Bioessays* **32,** 322–331 (2010).
28. Gutmann, A., Wagner, J., Allen, A. L., *et al.* Presidential commission for the study of bioethical issues. *Moral science: protecting participants in human subjects research. Washington DC* (2011).
29. Ministerie van Volksgezondheid, W. e. S. *Trendanalyse biotechnologie 2016 - Advies - Gezondheidsraad* https://www.gezondheidsraad.nl/documenten/adviezen/2016/06/14/trendanalyse-biotechnologie-2016 (2021).
30. Scientific Committee on Health and Environmental Risks (SCHER), Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR) & Scientific Committee on Consumer Safety (SCCS). *Opinion on Synthetic Biology II Risk assessment methodologies and safety aspects. European Comission* 2015.
31. Robaey, Z. *Dealing with risks of biotechnology: understanding the potential of Safe-by-Design.* Report commissioned by the Dutch Ministry of Infrastructure and Water Management (2018).
32. Stemerding, D., Betten, W., Rerimassie, V., *et al.* Future making and responsible governance of innovation in synthetic biology. *Futures* **109,** 213–226 (2019).
33. Guan, Z.-j., Schmidt, M., Pei, L., *et al.* Biosafety Considerations of Synthetic Biology in the International Genetically Engineered Machine (iGEM) Competition. *BioScience* **63,** 25–34 (2013).
34. Millett, P., Binz, T., Evans, S. W., *et al.* Developing a Comprehensive, Adaptive, and International Biosafety and Biosecurity Program for Advanced Biotechnology: The iGEM Experience. *Applied Biosafety* **24,** 64–71 (2019).
35. Aparicio, A. 'That would break the containment': the co-production of responsibility and safety-by-design in xenobiology. *Journal of Responsible Innovation* **0,** 1–22 (2021).
36. Nuzzo, A., Puccio, S., Martina, C., *et al.* Containment of a genetically modified microorganism by an activated sludge system. *New Biotechnology* **55,** 58–64 (2020).
37. Straathof, A. J. J., Wahl, S. A., Benjamin, K. R., *et al.* Grand Research Challenges for Sustainable Industrial Biotechnology. *Trends in Biotechnology* **37,** 1042–1050 (2019).
38. Asveld, L., Osseweijer, P. & Posada, J. A. Societal and Ethical Issues in Industrial Biotechnology. *Advances in Biochemical Engineering/Biotechnology* **173,** 121–141 (2020).
39. Bouchaut, B. & Asveld, L. Responsible Learning About Risks Arising from Emerging Biotechnologies. *Science and Engineering Ethics* **27,** 22 (2021).
40. Kallergi, A., Asin-Garcia, E., Martins dos Santos, V. A., *et al.* Context matters: On the road to responsible biosafety technologies in synthetic biology. *EMBO Reports* **22,** e51227 (2021).
41. Van de Poel, I. Design for value change. *Ethics and Information Technology* (2018).
42. Jacobs, N. & Huldtgren, A. Why value sensitive design needs ethical commitments. *Ethics and Information Technology* (2018).

43. Van den Hoven, J. *ICT and Value Sensitive Design* in *The Information Society: Innovation, Legitimacy, Ethics and Democracy In honor of Professor Jacques Berleur s.j.* (eds Goujon, P., Lavelle, S., Duquenoy, P., *et al.*) (Boston, MA, 2007), 67–72.

44. Van de Poel, I. *Translating Values into Design Requirements* (eds Michelfelder, D. P., McCarthy, N. & Goldberg, D. E.) 253–266 (2013).

45. Doorn, N. & Hansson, S. O. Design for the value of safety. *Handbook of ethics and values in technological design,* 491–509 (2015).

46. Friedman, B., Kahn, P. H., Borning, A., *et al. Value Sensitive Design and Information Systems* (eds Doorn, N., Schuurbiers, D., van de Poel, I., *et al.*) 55–95 (2013).

47. Wouters, M. & Rerimassie, V. *DE BURGER AAN HET WOORD: PUBLIEKSOPVATTINGEN OVER MODERNE BIOTECHNOLOGIE. Onderzoeksrapport InSites Consulting.* https://docplayer.nl/65535566-De-burger-aan-het-woord-publieksopvattingen-over-moderne-biotechnologie-onderzoeksrapport-insites-consulting-november-in-opdracht-van-ienw.html (2021).

48. Asveld, L. & Stemerding, D. Social learning in the bioeconomy: the Ecover case. *New perspectives on technology in society: experimentation beyond the laboratory,* 103–124 (2017).

49. ETC Group. *17 Groups call on Ecover and Method to drop extreme genetic engineering plans* ETC Group. https://www.etcgroup.org/content/groups-call-ecover-and-method-drop-extreme-genetic-engineering-plans (2021).

50. Schmidt, M. Diffusion of synthetic biology: a challenge to biosafety. *Systems and Synthetic Biology* **2,** 1–6 (2008).

51. Merad, M. *Highlights on the Risk Governance for Key Enabling Technologies: From Risk Denial to Ethics* (eds Trump, B. D., Cummings, C. L., Kuzma, J., *et al.*) 399–408 (2020).

52. Trump, B. D., Foran, C., Rycroft, T., *et al.* Development of community of practice to support quantitative risk assessment for synthetic biology products: contaminant bioremediation and invasive carp control as cases. *Environment Systems and Decisions* **38,** 517–527 (2018).

53. Cases, I. & Lorenzo, V. d. Genetically modified organisms for the environment: stories of success and failure and what we have learned from them (2005).

54. Warner, C. M., Carter, S. R., Lance, R. F., *et al. Synthetic Biology: Research Needs for Assessing Environmental Impacts* (eds Trump, B. D., Cummings, C. L., Kuzma, J., *et al.*) 19–50 (2020).

55. Cooperrider, D. L., Whitney, D. K. & Stavros, J. M. *Appreciative inquiry handbook* (2003).

56. Fenzl, T. & Mayring, P. QCAmap: Eine interaktive Webapplikation für Qualitative Inhaltsanalyse (2017).

57. Braun, V. & Clarke, V. Using thematic analysis in psychology. *Qualitative research in psychology* **3,** 77–101 (2006).

58. U.S. National Institutes of Health. Recombinant DNA Research Guidelines. *Federal Register* **41,** 27902–27943 (1976).

59. Rycroft, T., Hamilton, K., Haas, C. N., *et al.* A quantitative risk assessment method for synthetic biology products in the environment. *Science of The Total Environment* **696,** 133940 (2019).

60. Committee on Future Biotechnology Products and Opportunities to Enhance Capabilities of the Biotechnology Regulatory System, Board on Life Sciences, Board on Agriculture and Natural Resources, *et al. Preparing for Future Products of Biotechnology* (Washington, D.C., 2017).

61. De Lorenzo, V., Krasnogor, N. & Schmidt, M. For the sake of the Bioeconomy: define what a Synthetic Biology Chassis is! *New Biotechnology* **60,** 44–51 (2021).

62. Peng, X.-Y., Wu, J.-T., Shao, C.-L., *et al.* Co-culture: stimulate the metabolic potential and explore the molecular diversity of natural products from microorganisms. *Marine Life Science & Technology* (2021).

63. Simon, A. J., d'Oelsnitz, S. & Ellington, A. D. Synthetic evolution. *Nature Biotechnology* **37,** 730–743 (2019).

64. Martin-Pascual, M., Batianis, C., Bruinsma, L., *et al.* A navigation guide of synthetic biology tools for Pseudomonas putida. *Biotechnology Advances* **49,** 107732 (2021).

65. Van de Kaa, G., Rezaei, J., Taebi, B., *et al.* How to Weigh Values in Value Sensitive Design: A Best Worst Method Approach for the Case of Smart Metering. *Science and Engineering Ethics* **26,** 475–494 (2020).

66. Beauchamp, T. L. & Childress, J. F. *Principles of biomedical ethics* (2001).

67. Ferretti, M. P. Why Public Participation in Risk Regulation? The Case of Authorizing GMO Products in the European Union. *Science as Culture* **16,** 377–395 (2007).

68. Jr, J. S. L. & Durant, R. F. Public Opinion, Risk Assessment, and Biotechnology: Lessons from Attitudes toward Genetically Modified Foods in the European Union. *Review of Policy Research* **27,** 59–76 (2010).

69. Marcon, A., Master, Z., Ravitsky, V., *et al.* CRISPR in the North American popular press. *Genetics in Medicine* **21,** 2184–2189 (2019).

70. McCaughey, T., Budden, D. M., Sanfilippo, P. G., *et al.* A Need for Better Understanding Is the Major Determinant for Public Perceptions of Human Gene Editing. *Human Gene Therapy* **30,** 36–43 (2018).

71. Sandin, P. & Robaey, Z. *Modern Biotechnology, Agriculture, and Ethics* (eds Poff, D. C. & Michalos, A. C.) 1–4 (2020).

72. Ma, N. J. & Isaacs, F. J. Genomic Recoding Broadly Obstructs the Propagation of Horizontally Transferred Genetic Elements. *Cell Systems* **3,** 199–207 (2016).

73. Selão, T. T., Włodarczyk, A., Nixon, P. J., *et al.* Growth and selection of the cyanobacterium Synechococcus sp. PCC 7002 using alternative nitrogen and phosphorus sources. *Metabolic Engineering* **54,** 255–263 (2019).

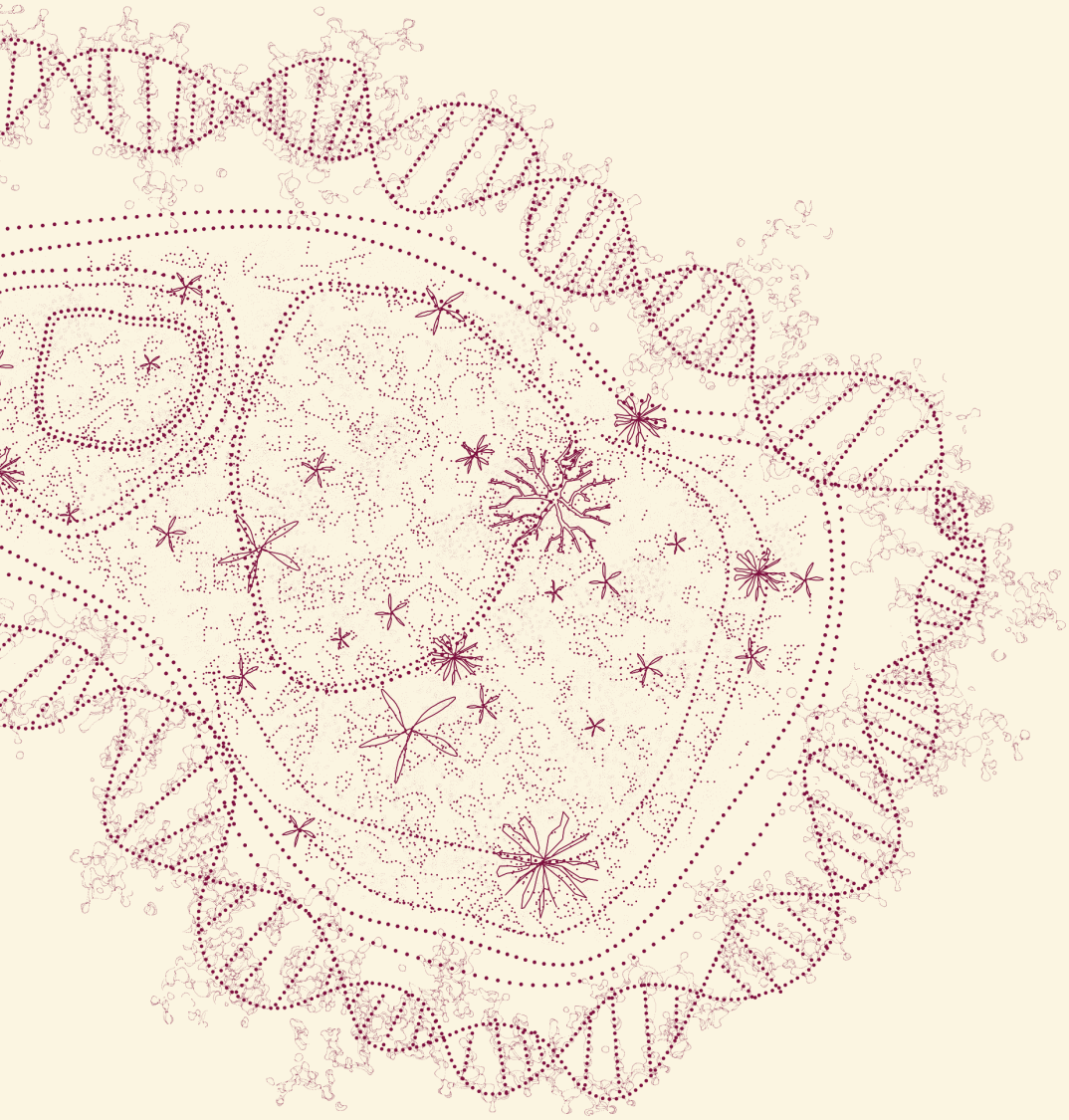
74. Aslan, S., Noor, E. & Bar-Even, A. Holistic bioengineering: rewiring central metabolism for enhanced bioproduction. *Biochemical Journal* **474,** 3935–3950 (2017).

75. Winkler, T. & Spiekermann, S. Twenty years of value sensitive design: a review of methodological practices in VSD projects. *Ethics and Information Technology* (2018).

76. De Reuver, M., van Wynsberghe, A., Janssen, M., *et al.* Digital platforms and responsible innovation: expanding value sensitive design to overcome ontological uncertainty. *Ethics and Information Technology* **22,** 257–267 (2020).

77. Kuzma, J. & Roberts, P. Cataloguing the barriers facing RRI in innovation pathways: a response to the dilemma of societal alignment. *Journal of Responsible Innovation* **5,** 338–346 (2018).

78. Mok, L. & Hyysalo, S. Designing for energy transition through Value Sensitive Design. *Design Studies* **54,** 162–183 (2018).



CHAPTER 9

Towards an attitude for responsibility in safety education

Hetty Huijs*, Enrique Asin-Garcia*,
Zoë Robaey, Vitor A. P. Martins dos Santos
* Contributed equally

Abstract

The fast pace of innovation in the life sciences leads to new questions regarding the safety and security of new technologies and processes. As a consequence, it is important to investigate what is taught about safety and how safety is taught during the early stages of a scientific career in higher education programmes. In this study, we performed a gap analysis by means of an inventory that delves into the learning contents on different dimensions and elements of safety, including understandings and reasonings around the concept of Safe-by-design. We used Wageningen University as a case study for our research, where we conducted a series of qualitative interviews with programme directors and teachers from different life science divisions. Our results show that the technical aspect of safety receives the most attention, especially when teaching about the development of an innovation. At the same time, other skills more related to anticipating and managing risks as well as other elements of responsibility need to be strengthened in the curricula. Interestingly, the frequency at which a safety element is taught was shown to be always higher than the amount of skills instructed on the same topic. This illustrates that more attention should be paid to aiding students by teaching them skills to improve their own judgements on safety in their future careers.

Introduction

The last two decades have witnessed an increasing urgency of pursuing solutions to serious global problems such as climate change, food security, health-associated issues, industrial restructuring and energy shortage [1]. To address these grand challenges, the concept of bioeconomy has been proposed as a key element towards solving these issues by means of a transition from fossil-based to bio-based products and energy [2]. The term bioeconomy is explicitly connected to an increasing use of biology and biotechnology across sectors [3] including amongst others food, plant and animal sciences. In general, all those scientific branches that involve the study of life receive the name of life sciences and, in a broad sense, most of them connect *via* three ideal visions to the different dimensions of what the notion of bioeconomy constitutes [4, 5]. We encounter in the first place the biotechnology vision, focused on the potential applicability of bioscience. Next, we have the bio-resource vision that emphasizes the upgrading and conversion of biological raw materials. Lastly, the bio-ecology vision highlights the potential for regionally concentrated circular and integrated processes and systems [1]. While the first two visions are significantly influenced by a global technical perspective and by engineering science, the last is more related to sustainability and environmental sciences. Hence, all of these interconnected disciplines belonging to the life sciences field are involved in the knowledge production and innovation process underpinning the progress and success of the bioeconomy.

In this context, and as technological innovations within life sciences develop, new questions arise regarding their safety which often include complex ethical, social, and environmental issues. On top of the previous risks intrinsically associated with handling biological agents, especially microbes such as bacteria, viruses, parasites, fungi, prions and other related agents [6], recent technologies have added a list of new concerns typically associated to genetic modification [7]. Risk is a meaningful concept in life science, but also for daily life in general. Every day, we are exposed and need to confront different types of risks and, while our responses to them are often straight forward, sometimes their complexity requires from us specific knowledge to properly assess and manage the situation [8]. While the basic safety skillset to undertake common daily risks is taught to most citizens when

they are young both at home and in primary and secondary school (think of traffic rules, sex education, etc.), the more complex the risks become, the more specialized information and skills are needed to apply an adequate decision-making process to confront them. Not only life science risks, but also derived uncertainty, are therefore critical elements to be taken into account when assessing safety across these disciplines. To give an example, the life sciences are contributing to the speedy progress of medicine with advance strategies to treat and monitor disease, including gene therapy and personalised medicine. The former involves the introduction of genetic material into the cells of an individual's body for its subsequent expression to treat an illness. The latter is based on the use of specific treatments tailored to the genetic profile, lifestyle and environment of individual patients. Although these two trends offer clear benefits, they present a highly diverse collection of challenges and associated risks, whose nature is not only related to the scientific methods, side-effects and consequences to the patients' health, but also to the functioning and affordability of healthcare, the licensing of the treatments and numerous ethical issues [9]. Since this type of high-profile issues will become a part of life scientists' careers, an appropriate training and literacy on safety becomes of paramount importance in the higher education programmes [10]. While this might be perceived as a natural part of socio-scientific issues in life science education, it is rarely acknowledged as a learning content in its own right [8]. The following paragraphs describe different ways and motivations from which the concept of safety could be approached for educational purposes.

On the one hand, one could learn to adhere to the safety rules and checks simply because they are compulsory. This means the only aim for incorporating safety would be compliance with regulations. On the other hand, a more effective way of looking and managing safety would be with a deeper focus on responsibility in addition to compliance [11], where researchers take a more responsible and active role in the safety of the technologies and products they develop. This relates to the framework of Responsible Research and Innovation (RRI) [12] and is particularly important in instances where there are risks for which no compliance mechanisms have been established. In a nutshell, RRI requires that innovative products and processes align with societal needs, expectations and values in consultation

with stakeholders [13], making the researchers take more responsibility for their work. Furthermore, RRI encourages scientists to anticipate on future applications and to include varying opinions [14]. One can conceptualize responsibility as being more concerned with results, whereas duty or compliance is more about a known course of actions [15]. Due to the large number of possibilities that have arisen in life sciences, the safety of some of the actions cannot always be controlled and, therefore, a shift towards responsibility beyond compliance represents a more pre-emptive approach. As a consequence the concept of Safe-by-design (SbD) [11, 16, 17] emerged originally as a Dutch policymaking initiative for dealing with safety issues by combining different methods of assessment and ethics of technology. Based on responsibility and aiming at including the researcher from an early stage, SbD challenges research, industry and education to pay attention to the wide scope of safety issues from the design and development stages up to the final application phase. While students seek higher education programmes that provide them with the best knowledge and tools to perform adequately in their future careers [18, 19], universities need to integrate and balance safety topics in their life science curricula in light of the new technologies and advances that expeditiously drive the shift towards the bioeconomy [20]. In this research, we explore which gaps exist in higher education curricula regarding safety education by questioning what is taught about safety, how safety is taught and make recommendations on what can be improved. For this, we use Wageningen University & Research (WUR), an institution specialized in life sciences in The Netherlands, as a case study. In our gap analysis, seven directions of life sciences are included in order to obtain advice on what elements each study programme needs to improve or introduce as part of its curriculum.

Concepts and framework: understanding elements of Safety

Safe-by-design: technical safety, risk assessment and RRI

SbD is a concept in which safety is meant to be considered in every step of the design process in order to procure an inherently safer product [11]. Generally, SbD is a broad term, widely applicable to many different fields that range from biotechnology [17] to nanotechnology [21, 22]. In particular, this research aims to evaluate

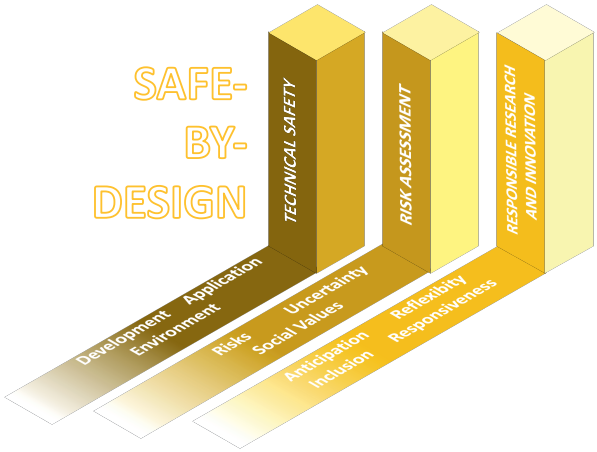


Figure 9.1: Safe-by-design dimensions together with their respective elements.

the extent to which aspects of the SbD paradigm are present in the life sciences study programmes at WUR. To be able to perform this investigation, it is therefore necessary to first explore the different dimensions and the elements they entail. Three key dimensions of the framework had been previously defined in an internal collaborative document between the WUR and the Dutch Ministry of Infrastructure and Water Management (https://www.safe-by-design-nl.nl/default.aspx): technical safety, risk assessment, and responsible research and innovation (Figure 9.1). We specify their elements below.

Technical safety

We focus first on the examination of a technological innovation's safety from a technical viewpoint. Whether a product is safe can be assessed by looking at its development, its application and its effect on the environment [16]. The development of a product should be safe, with the necessary protection measures for the researcher. This includes a series of considerations such as looking at what products are used during the developmental stages and how they are disposed of. Also, application of the product should be safe, making sure it contains no toxic substances for the consumers or the people who are ultimately in contact with the final version of the product. Lastly, during the whole process, safety for the environment must

be taken into account.

Risk assessment

SbD has also been defined as a specific risk assessment strategy that addresses safety by means of design measures [11]. Risk analysis is an important step to increase the safety of a product and, looking for risks throughout the entire process helps identifying more of them. Classical risk assessment looks at risks like calculable entities combining the probability of an event happening with its estimated effect (risk = probability × effect) [23]. This way of looking at risks is the first element of the risk assessment dimension. However, risk assessment for technological innovations in live sciences also contains an increasing number of uncertainties since either the probability of an accident or the associated effect is unknown. Hence, the importance of looking at these uncertain aspects as the second element of this aspect. In the third place, risks are perceived differently depending on a person's social values. What might be acceptable to someone, could be unacceptable to someone else. Thus, public opinion and social values about risks need to be taken into account representing the third aspect of the risk assessment dimension.

Responsible Research and Innovation

Like SbD, RRI also focuses on a proactive role of the scientist in finding responsible and safe solutions. Stilgoe *et al.* [14] developed a framework for RRI consisting of four elements that work as the base for our study of SbD in education. The first element is anticipation, prompting researchers to ask themselves "What if....?". The anticipation of researchers on how their research will be used can lead to new safer opportunities for innovation. The second element is reflexivity, reflecting on one's research, activities, and assumptions. Inclusion is the third element, meaning the consideration of new opinions and engaging the wider public to reach responsible, safe, and socially accepted solutions, which is also related to the attention to social values in risk assessment. The last element is responsiveness, which implies responding to social values and circumstances and allowing flexibility in research to change according to these factors.

What is taught about safety and how?

Education in life sciences at WUR offers both theoretical and practical sessions. Within these classes, students examine in depth momentous issues such as the transition to a more circular economy, health-related pandemics and epidemics, and climate change, which propitiates an increasing awareness of risks within their future career functions [19]. After their education, students realize, situations where they have to make important decisions on risks will be frequent [24]. Therefore, it is important not only to acknowledge the issues but to teach students skills on how to deal with them and how to develop an attitude towards responsibility and safety. Understanding of this requires skill-building tasks with a gradual increase of responsibility [24]. Concepts of ethics help understand that taking responsibility implies one needs to actively self-supervise in order to achieve a desired end, whereas duties demand following specific instructions. These fundamental concepts have been translated in applied ethics in the context of risk. Specifically, responsibility has been specifically translated to the ability to deal with uncertain risks whereas compliance represents a duty where known actions are dictated to get to the desired end, for example through regulation [25]. In other words, responsibility for safety in education should mean more than following the rules.

In addition, when evaluating learning contents in education one should not only look at what is taught but also how is taught. Thus, the combination of these two approaches is required for a proper assessment of the situation [26]. The first approach is the quantitative, also called standards-based or criterial approach, which looks at descriptive variables [26]. The other approach is the qualitative or responsive approach that relies on interpretive thinking and is often more based on experiences and narratives [26]. Related to quantitative and qualitative approaches, when evaluating educational contents, it is important not only to look at the inputs but also at the outputs [27]. Input can be seen as the amount of time spend on a certain subject, whereas the output would be the knowledge and skills acquired through it.

To investigate the gaps in education on SbD, two scales were developed for assessing the attention paid to each of the aforementioned elements and dimensions (Table 9.1). First, we have the descriptive scale that indicates how often the element

in question is discussed in different lectures or practical sessions, revealing how much time and attention is paid to the subject. This scale measures the input given to an element by looking at simple criteria that need to be met. It can therefore be considered criterial and descriptive. The second is the analytical scale that gives an indication of the depth of the education: is the element merely brought up in the lecture? Do students learn skills on how to handle the element? Or is the element a key-value and are they encouraged to develop an attitude to pay attention to this element? This scale is more interpretive and comes from an analysis of the experiences and narratives of the interviewees. This scale focuses on the output of the education learning obtained on a given dimension's element.

Table 9.1: Descriptive and analytical scale for the gap analysis. Levels from the analytical scale were inspired by Laudonia et al. [28].

	Descriptive scale (what is taught and how often)	**Analytical scale (how it is taught)**
Level 1	This element is not present or is only discussed in 1 lecture	Cognition: students are told about this element
Level 2	This element Is discussed 2-7 times in lectures or practical sessions	Skills: students learn skills to deal with this element
Level 3	This element is discussed more than 7 times throughout the study programme	Attitude: students develop an attitude towards this element to always consider it

Methods

To gain information on what SbD dimensions and their elements are present in education on life sciences at WUR and how they are taught, a series of in-depth qualitative interviews were performed with people involved in different educational programmes of the university.

Recruitment and procedure

In order to cover a broad scope of study programmes, education on life sciences at WUR was divided into seven study directions corresponding to the offered bachelor programmes. Within each of these directions, at least two key programme coordinators and lecturers were selected for our study (except in the case of the Nutrition & Health direction, in which only one interview was finally performed) (Table 9.2). For most directions, the programme director was first approached given the good overview of the respective sector. These initial conversations were performed with

Table 9.2: Overview of the different study directions and the number of associated people queried.

Direction	Number of interviews*
Animal Science	2
Biology	2
Biotechnology	2
Food Science	2
Molecular Life Sciences	2
Nutrition & Health	1
Plant Science	2
General	3

*Some interviews covered multiple study directions. However, for this table only the main one was considered.

the objective of gaining information in the following matters: (i) courses that deal with SbD elements; (ii) teachers that could be approached to get a better understanding of how their courses deal with the different notions of safety; (iii) give information on key courses of the programme they are involved in.

Potential interviewees were initially approached *via* email to arrange a virtual meeting and informed in advance about the objective and methodology of our research. Interviewees were approached for their professional insights into the curricula, and not their personal views. In addition, prior each interview, participants gave oral consent to use the outcome of their interviews as source material for this study. The whole set of interviews (Ntot = 16) was conducted by one interviewer over a period of 4 months, from April to July 2020.

Interviews

The interviews followed a semi-structured approach including general and open-ended questions in order to avoid study direction-bias and stimulate in-depth discussions.

A first section was built around three key aspects: technical safety, risk assessment and RRI. Second, a more open approach was used to explore the respondents' understanding of the terms SbD and RRI and to explain and discuss these concepts with them. The third final section worked as a more conversational conclusion and included a request for further interviewees and references.

Recording and data analysis and protection

Interviews were consensually audio-taped and subsequently transcribed (intelligent verbatim style) by the GoTranscripts service. Transcripts were then further analysed using Atlas.ti and organized in codes by study direction and according to the elements and scales described in the Concepts and Framework section. To support recurrent themes, direct quotes representing participants' opinions and views are provided in the results and discussion sections. Analysed data, interview transcripts and identity of the interviewees remain protected according to the WUR protocol for safe data management.

Results

Descriptive and analytical learning

From the interviews, a good understanding could be gained on how the different elements of SbD are considered in the different study programmes (Figure 9.2). From this analysis, all study directions were scored on two different three-tier scales which included the aforementioned quantity of education and the depth of education on a subject (Table 9.1). In general, we observed that technical safety was the part to which most attention was paid, mostly encompassing dark coloured tiles in all the study directions. Risk assessment and RRI showed more balanced results, but dark coloured tiles were scarce here whereas light ones were abundant, meaning less attention was paid to these dimensions. In addition, it is worth mentioning that when the two scales did not yield same-level results for a specific element, the descriptive scale always scored higher than the analytical scale. Particularly from the analytical point of view, the risk assessment and RRI dimensions were exclusively focused on cognition and skill learning, whereas students were taught to develop an attitude only towards elements of the technical safety dimension.

Regarding this technical dimension specifically, safety during the development phase received the most attention from both the descriptive and analytical points of view in all the divisions, as indicated by all the dark coloured tiles of the first row. In turn, safety towards the environment was discussed more than 7 times throughout the programme according to the descriptive scale, but only reached

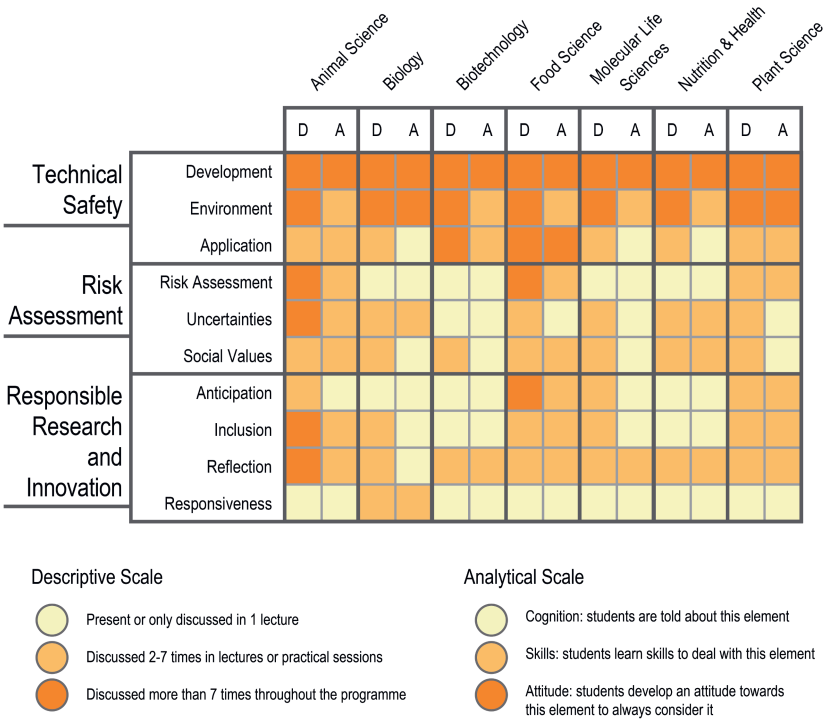


Figure 9.2: Overview of interviews' results per element and per study direction. Descriptive scale on the frequency and occurrence and analytical scale on the depth of learning are indicated with D and A, respectively.

the level of learning skills in the analytical scale. Exceptions of this were the Biology and Plant Science directions where this element seemed to be very important from both perspectives. Regarding safety for life science applications, only in Food Science the element was discussed more than 7 times and at the attitude learning level, whereas it received less attention in the other study divisions. In Molecular Life Sciences, Nutrition & Health and Biology, this element was discussed between 2 and 7 times throughout the whole programme and only taught at the cognition level. This is consistent with the fact that not all study programmes aim to yield technological applications.

Moving on to the risk assessment dimension, we appreciate that according to the descriptive scale, only the elements of risk assessment and uncertainties were discussed more than 7 times in Animal Science, next to the risk assessment element in

the Food Science. The risk assessment element was however only discussed once in Biology, Biotechnology, Molecular Life Sciences and Nutrition & Health. Moreover, the element of uncertainties was also treated only once in Biotechnology. In all the other situations, the elements were discussed between 2 and 7 times. Regarding the analytical scale, none of the elements were taught towards developing an attitude as indicated above. In the Animal Science direction, students learnt skills to deal with the three elements, risk assessment, uncertainties and social values. This was also the case for risk assessment in Food and Plant Sciences, for uncertainties in Biology and Nutrition & Health, and for social values in Food Science and Nutrition & Health. In every other situation, elements were only taught at the cognition level. Especially remarkable was the situation of Biotechnology and Molecular Life Sciences in which students were simply told about the three elements that constitute this dimension (cognition level).

Regarding the RRI dimension from the descriptive perspective, only the elements of inclusion and reflection in Animal Science, and the element of anticipation in Food Science were discussed more than 7 times. Anticipation was treated only once in Biology, Biotechnology and Nutrition & Health, and inclusion, only once in Biotechnology and Nutrition & Health. Discussed only during one lecture was also the element of responsiveness in every study direction, except in Biology where it was discussed at least twice. In every other occasion, elements were discussed between 2 and 7 times. The element of reflection was an exemplary of this since it came up at least twice in all the study programmes. According to the analytical scale, none of the elements of this dimension were taught at the attitude level either. Students of Food and Plant Sciences learnt skills to deal with anticipation and inclusion, while their counterparts of Animal Science only did for inclusion. Reflection was taught at the skill level in every study direction, except in Biology where only the cognition level was reached. By contrary, responsiveness was simply told about (cognition level) in all the programmes aside from Biology, where students learnt skills to deal with this element.

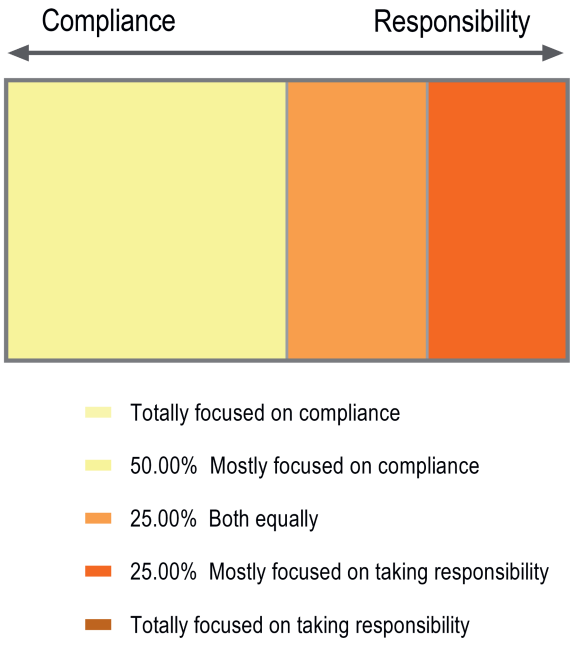


Figure 9.3: Bar graph containing the range of interviewees' answers to the question "Do you think education at WUR is currently more focused on compliance or on responsibility?" (Ntot = 12).

Responsibility *vs.* compliance

As stated in the introduction, SbD is an approach encouraging people in taking responsibility for safety. After discussing with the participants throughout the interviews this concept and its dimensions and elements in addition to the understanding of responsibility as earlier described in the Introduction and Concepts and framework sections, the question was asked on whether current education at WUR was more focused on compliance with rules or towards taking responsibility for safety. The answers from the interviews were translated to fall into the categories "mostly focused on compliance" (50 %), "both equally" (25 %) or "mostly focused on responsibility" (25 %) (Figure 9.3), suggesting a stronger focus on compliance in education.

Discussion and conclusions

This research aimed at finding gaps regarding higher education on safety from a SbD perspective using Wageningen University & Research as a case study. To do this, different elements that shape the SbD paradigm were proposed in the Concepts and framework section of this work. Next, a series of exploratory qualitative interviews were conducted with teachers and programme directors to gain a better understanding of the status of each of these elements in the different study programmes. Collected material was then analysed on the basis of the SbD elements and structured in two different scales that responded to the frequency and the depth of the learning contents. Based on the outcomes of our investigation, two main conclusions could be drawn. In the first place, elements always received either an equal or a higher score in the descriptive than in the analytical scale; and secondly, compliance-oriented elements such as those related to technical safety are much more present than responsibility-oriented ones.

Descriptive over analytical

A noticeable trend from the overview (Figure 9.2) is that the descriptive scale is always positioned on an equal or higher tier than the analytical scale, which suggests that the input is higher than the output. Generally speaking, this means that specific elements are brought up in lectures, but further contents on how to address them are lacking. While a developed capacity for critical thinking and transferable skills are concepts that most designers of programmes in higher education would be likely to endorse, it is complex to determine how such capacities might be optimally established in an institutional setting [29]. Strong on the provision of the relevant multidisciplinary knowledge base, higher education has not factually been so strong on the development of the competences essential for its effective use in real world applications [30, 31].

Despite the dramatic changes that university learning environments have undergone in the recent decades, with more constructivist and learning-focused approaches taking over the more traditional content-focus ones [32–34], our findings clearly indicate that the latter keep being dominant when educating about safety. From our interviews, we find that while lecturers often have a good theoretical and

technical background, they lack knowledge related to the social perspectives of safety. Previous research has shown the importance of developing teacher's approaches to teaching towards being learning-focused and the positive impact of teacher's pedagogical development programmes for this purpose. Such a training would bring about opportunities to reflect upon how teaching and assessment practices affect the learning process of student [35]. In addition, meaningful interactions with other lecturers from the social department experienced and trained in a different and complementary skillset, related for instance to the concepts of risks and uncertainty, could help supplementing the current curricula by means of interdisciplinary thinking [36]. Amongst the most representative elements in which the descriptive scale receives a clearly higher attention than the analytical scale, we find the environmental safety (Figure 9.2). With the rise in concern for environmental sustainability that propitiated the transition towards the bioeconomy, education is called to play a critical and multifaceted role in the transition to more sustainable forms of development [37]. Its relevance makes of the topic a recurrent theme during the lectures, perhaps including direct problem-based learning [38]. However, deeper forms of environmental safety application should be considered including instrumental approaches that address specific environmental challenges aiming at promoting certain responsive behaviours, and emancipatory education to build key competencies such as reflecting and thinking critically, working collaboratively and engaging with different viewpoints, with the objective of identifying solutions [37]. This situation is intrinsically related to the element of anticipation but also to the other factors of the RRI perspective. Paying attention to the environment's safety during the developmental stages of an innovation requires a strong component of anticipation, inclusion and reflection, which, according to our study, remain as well more strongly rooted in the descriptive level. Opportunities arise to teach students better skills on the anticipation dimension by letting them understand, engage and appreciate concepts, methods and applications for exploring and managing possible futures [39].

Compliance over responsibility

Investigating the issue of compliance over responsibility also relates to going deeper in learning, preferably reaching levels of attitude for responsibility. For this, it is important to highlight the broader context of practice in engineering in the life sciences. Furthermore, the concept of responsibility is best teased apart building on literature in engineering ethics. As a broader context, The Royal Netherlands Society of Engineers code of ethics states as first principle:

> "We shall take into account how our technical decisions influence the health and safety of people and their surroundings. We will not hide any factors that influence the safety of society and the environment."
> *https://www.kivi.nl/uploads/media/5a587110c2160/2018-01%20Code%20of%20Ethics.pdf*

Looking at Figure 9.2, one realizes that the dark coloured tiles associated to higher levels of attention concentrate at the part of the heatmap related to technical safety, especially during the development stages. This fact was justified during the interviews by the importance of following laboratory safety rules for the personal safety of students, which is paramount during these sessions. In general, laboratory safety in the academic setting receives a great deal of attention by committing resources that include training, safety plans and access to online information [40–42]. This is still unquestionably necessary as research underlines that academic environments still suffer from a less well-established safety culture than industrial and governmental laboratories, which is mainly due to the students' lack of knowledge and experience regarding the risks associated to their experiments [43]. However, this keen focus on compliance to rules to avoid accidents, push precisely to the background aspects intrinsically related to risk assessment and responsibility. This idea was confirmed by the results depicted on Figure 9.3.

There are two related problems with understanding safety education as mere compliance. First, codes, and compliance-based education can leave room for interpretation [44], where students might not then be equipped with the right skills or attitude to further act in a desirable manner. Second, the ability to comply and carry out technical assessments does not guarantee skills to understanding com-

plex contextual safety issues like risk assessment and RRI would. This could present students with difficulties when making decisions about risky situations [45]. The overall low scores on the RRI dimension are especially worrisome, as

> "[RRI] intends to stimulate responsible behaviour and encourage individuals to assume responsibility beyond the mere compliance with positive legal frameworks and traditional role responsibilities such as those of engineers, scientists and policy-makers" [46].

Particularly, it is worth mentioning the lack of attention paid to responsiveness, a vital element for the students to develop a capacity to react to the situation taking into account the stakeholder's input [47]. The integration of this element in the RRI framework works as a strong link between cognitive and attitude-oriented education and enables the required two-way dialogue for being responsive to stakeholder's perspectives and anticipated scenarios [48]. The issue of 'mere compliance' is an often-voiced critique in engineering ethics when it comes to speaking about responsibility of engineers [49]. To remedy this, engineering ethics scholarship has been calling for further integration of virtue ethics in education.

However, there are encouraging trends in our analysis: differences between technical safety and the dimensions of risk assessment and RRI were less pronounced for study directions such as food and plant sciences, probably due to their close proximity to societal applications. Moreover, based on the interviews, WUR strives towards more responsibility in education on safety. This is visible from a series of actions and implementations that came up during the interviews including a responsibility learning line that is being implemented within the programme of Biology, and a collection of safety modules that *"are really suited for education and drift away from the rule-based modules towards a more mindset modules"* (Anonymous Interviewee) which will be obligatory for all life science programmes since the start of the academic year 2021-2022. However, our interviewees acknowledged that *"producing modules like this is still a challenge that takes time and multiple developments and that, while for now they are still more focused on compliance, they will be improved in the coming years"* (Anonymous Interviewee). In order to support students in learning an attitude for responsibility for important issues likes safety,

more integration with recent work in virtues for engineering ethics might help fill the gaps.

Learnings and perspectives: the way forward

Studies on education have consistently highlighted the importance of developing practical knowledge and student-centred programmes in higher education institutions to tackle future issues [50–52]. As we have mentioned, the underlying idea is that giving students the tools and the responsibility for their own learning will activate and motivate them, helping them to better prepare for their professional careers [53, 54]. Alas, evidence from our case study suggests that safety learning in life sciences is more focused on compliance than in responsibility, with a dominance of descriptive cognitive learning over a more analytical attitude-oriented education.

Coming back to our initial example, these life science students learn in lectures and practical sessions contents encompassing novel topics such as gene therapy and personalized medicine. Issues and challenges of these technologies are presented, mostly at the technical dimension and related to development safety: safety laboratory practices and rules, levels of drug toxicity, vector dose and potency, etc. [55, 56]. While pragmatically useful, this information exclusively conveys a message of compliance to standards and regulations. The way forward should be focused on developing relevant skills and attitudes towards the different dimensions and elements of safety. Despite the importance of technical safety in medicine and health, the risk assessment and RRI dimensions would help providing a more comprehensive overview of the current and future risks, uncertainties and scenarios [9] as well as building up a personal ethical profile to question issues derived from new technologies. Thus, when studying gene therapy, one should learn: how to deal with the high expectations embedded in this type of promising and powerful techniques [57], how to analyse the risks associated with, for example, genetically modified viruses and deviant cell behaviours [58], how to understand the licensing required for clinical trials [59], and how to judge the costs of these therapies and their affordability in different points of the world [60]. In turn, when discussing personalized medicine, students should consider the high prices of the

limited and tailored drugs [61], the difficulties within the clinical trials with a reduced amount of individuals [62], and the reliability on statistics to establish a safe connection between the results expected according epidemiological research and the actual outcome in particular patients [63]. This is not an easy task. Furthermore, and related to social values and the elements of RRI, it is important to become aware that the aforementioned developments and issues concern various stakeholder groups. In particular, gene therapy and personalized medicine may have implications for patients and consumers, medical professionals, scientific researchers, pharmaceutical companies and even insurers and governmental bodies [9].

A major barrier to a less compliance-oriented safety education is the fact that teachers are often specialized in their own didactic systems or basic units of teaching. Thus, a teacher mastering a new paradigm needs support through collective pedagogical action [64]. We argue that substantial changes and collaboration with social sciences teachers are required in both curricula and pedagogical practice to challenge the current epistemologies and to unsettle the current ways of acting and thinking about safety [65]. In this way, inter- and multidisciplinary education will boost student capacities for critical and reflective thinking. At the same time, safety contents should be provided at the adequate phases of the education programmes, by adjusting the depth of teaching to the degree of receptibility. One will not be able to properly comprehend the risks of gene therapy after a basic Cell Biology course. It will not be until the students attend more advance courses such as Gene Technology or Molecular Biology Toolbox that they will have the resources to develop skills and attitudes towards certain safety elements. Here, we also find that closer connection with recent development in virtues for engineering ethics can come to support educating for an attitude for responsibility for safety.

Our case study focussed on bachelor and master level programs because increasingly, students can join the workforce after a bachelor's or master's degree. This is however a limitation in that we have not considered what kind of learning would fit best at which stage of education. This would call for an ordering of which levels of attitude for responsibility for safety would be satisfactory for which level of education. In addition, this could be expanded to industry, considering that educa-

tion for safety could become a more prominent cultural aspect of various practices in the life science [66].

While a core of safety material could be taught *via* a standalone subject [67, 68], we advocate for covering safety topics throughout the different courses of the programme by means of carefully selected and relevant case studies. These tools that present students with real and potential situations that could result in issues, controversies or accidents are an excellent way to engage them in learning [69, 70]. Case studies can be presented in lectures taking time to highlight key points and the chain of events that led to undesired outcomes, or through presentations carried out by the students themselves after working in groups in concrete cases. The latter approach would not only help developing students' presentation and general communication skills, but would also foster students' analytical abilities and receptibility [69]. While these practices might increase the relevance of safety, they might also seem difficult to fit within the contents of the curricula. As a consequence, anticipatory skill- and attitude-oriented SbD methods should always be prioritized over mere cognitive compliance information. In a world filled with complex problems that require even more complex solutions and replete of unexpected and uncertain non-documented issues, it is naïve to discipline students only with theoretical contents and routines, and non-transferable skills [71]. Instead, safety education must provide the tools and the resources to engage learners and empower them so that they take responsibility for their own life-wide learning and have the drive to make their own safe choices in their future life science careers.

Acknowledgements

We would like to thank all our interviewees that contributed to this research for sharing with us their time, their views and their experiences. This work is part of the research programme SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis with project number 15814, which is financed by the Dutch Research Council (NWO). This work is also part of the research programme "Virtues for Innovation in Practice (VIPs): A Virtue Ethics Account of Responsibility for Biotechnology" with project number VI. Veni.191F.010,

which is financed by the Dutch Research Council (NWO).

Conflict of interest

The authors declare there are no conflicting interests.

Author contributions

Conceptualization: HH/EAG/ZR/VAPMdS; Literature research: HH/EAG/ZR; Data curation: HH; Formal analysis: HH; Funding acquisition: ZR/VAPMdS; Investigation: HH/EAG; Methodology: HH; Supervision: EAG/ZR/VAPMdS; Visualization: EAG; Writing - original draft: EAG/ZR; Writing - review & editing: HH/EAG/ZR/VAPMdS

Bibliography

1. Bugge, M. M., Hansen, T. & Klitkou, A. What Is the Bioeconomy? A Review of the Literature. *Sustainability* **8,** 691 (2016).
2. Pülzl, H., Kleinschmit, D. & Arts, B. Bioeconomy–an emerging meta-discourse affecting forest discourses? *Scandinavian Journal of Forest Research* **29,** 386–393 (2014).
3. Wield, D. Bioeconomy and the global economy: Industrial policies and bio-innovation. *Technology Analysis & Strategic Management* **25,** 1209–1221 (2013).
4. Levidow, L., Birch, K. & Papaioannou, T. Divergent paradigms of European agro-food innovation: The knowledge-based bio-economy (KBBE) as an R&D agenda. *Science, Technology, & Human Values* **38,** 94–125 (2013).
5. Staffas, L., Gustavsson, M. & McCormick, K. Strategies and policies for the bioeconomy and bio-based economy: An analysis of official national approaches. *Sustainability* **5,** 2751–2769 (2013).
6. Bayot, M. L. & Limaiem, F. Biosafety guidelines (2019).
7. Knight, A. J. Perceptions, Knowledge and ethical concerns with GM foods and the GM process. *Public Understanding of Science* **18,** 177–188 (2009).
8. Schenk, L., Hamza, K., Arvanitis, L., *et al.* Socioscientific Issues in Science Education: An opportunity to Incorporate Education about Risk and Risk Analysis? *Risk Analysis* (2021).
9. COGEM & Gezondheidsraad. Trend analysis biotechnology 2016: A regulatory disconnect. CGM/160614-01. *The Netherlands Commission on Genetic Modification, Bilthoven* (2016).
10. Millar, R. Twenty First Century Science: Insights from the Design and Implementation of a Scientific Literacy Approach in School Science. *International Journal of Science Education* **28,** 1499–1521 (2006).
11. Van de Poel, I. & Robaey, Z. Safe-by-Design: from Safety to Responsibility. *NanoEthics* **11,** 297–306 (2017).
12. Schomberg, R. v. *A Vision of Responsible Research and Innovation* (eds Owen, R., Bessant, J. & Heintz, M.) 51–74 (2013).
13. Kallergi, A., Asin-Garcia, E., Martins dos Santos, V. A., *et al.* Context matters: On the road to responsible biosafety technologies in synthetic biology. *EMBO Reports* **22,** e51227 (2021).
14. Stilgoe, J., Owen, R. & Macnaghten, P. Developing a framework for responsible innovation. *Research Policy* **42,** 1568–1580 (2013).
15. Goodin, R. E. Responsibilities. *Philosophical Quarterly* **36,** 50–56 (1986).
16. Noorlander, C., Sips, A., Höck, J., *et al.* NANoREG Safe-by-Design (SbD) Concept. *Report RIVM and TEMAS AG* (2016).
17. Robaey, Z. *Dealing with risks of biotechnology: understanding the potential of Safe-by-Design.* Report commissioned by the Dutch Ministry of Infrastructure and Water Management (2018).
18. Barnett, R. Knowing and becoming in the higher education curriculum. *Studies in Higher Education* **34,** 429–440 (2009).
19. Thekdi, S. A. & Aven, T. Risk Science in Higher Education: The Current and Future Role of Risk Science in the University Curriculum. *Risk Analysis* (2021).
20. Hill, R. H. Undergraduates Need a Safety Education! *Journal of Chemical Education* **93,** 1495–1498 (2016).
21. Kraegeloh, A., Suarez-Merino, B., Sluijters, T., *et al.* Implementation of Safe-by-Design for Nanomaterial Development and Safe Innovation: Why We Need a Comprehensive Approach. *Nanomaterials* **8,** 239 (2018).
22. Yan, L., Zhao, F., Wang, J., *et al.* A Safe-by-Design Strategy towards Safer Nanomaterials in Nanomedicines. *Advanced Materials* **31,** 1805391 (2019).

23. Van Est, R., Walhout, B. & Brom, F. *Risk and Technology Assessment* (eds Roeser, S., Hillerbrand, R., Sandin, P., *et al.*) 1067–1091 (2012).
24. Ross, A. & Athanassoulis, N. The Social Nature of Engineering and its Implications for Risk Taking. *Science and Engineering Ethics* **16,** 147–168 (2010).
25. Robaey, Z. Looking for Moral Responsibility in Ownership: A Way to Deal with Hazards of GMOs. *Journal of Agricultural and Environmental Ethics* **28,** 43–56 (2015).
26. Stake, R. E. *Standards-based and responsive evaluation* (2003).
27. Gray, P. J., Patil, A. & Codner, G. *The background of quality assurance in higher education and engineering education* 3–25 (2009).
28. Laudonia, I., Mamlok-Naaman, R., Abels, S., *et al.* Action research in science education – an analytical review of the literature. *Educational Action Research* **26,** 480–495 (2018).
29. Walker, P. & Finney, N. Skill Development and Critical Thinking in Higher Education. *Teaching in Higher Education* **4,** 531–547 (1999).
30. Harvey, L. Quality Assessment in Higher Education: The collected papers of the QHE Project (Birmingham, Quality in Higher Education, The University of Central England in Birmingham) (1993).
31. Leckey, J. F. & McGuigan, M. A. Right Tracks—Wrong Rails: The Development of Generic Skills in Higher Education. *Research in Higher Education* **38,** 365–378 (1997).
32. Prosser, M. & Trigwell, K. Qualitative variation in approaches to university teaching and learning in large first-year classes. *Higher Education* **67,** 783–795 (2014).
33. Tynjälä, P. Towards expert knowledge? A comparison between a constructivist and a traditional learning environment in the university. *International journal of educational research* **31,** 357–442 (1999).
34. Uiboleht, K., Karm, M. & Postareff, L. The interplay between teachers' approaches to teaching, students' approaches to learning and learning outcomes: a qualitative multi-case study. *Learning Environments Research* **21,** 321–347 (2018).
35. Postareff, L., Lindblom-Ylänne, S. & Nevgi, A. The effect of pedagogical training on teaching in higher education. *Teaching and teacher education* **23,** 557–571 (2007).
36. Spelt, E. J. H., Biemans, H. J. A., Tobi, H., *et al.* Teaching and Learning in Interdisciplinary Higher Education: A Systematic Review. *Educational Psychology Review* **21,** 365 (2009).
37. Wals, A. E. J. & Benavot, A. Can we meet the sustainability challenges? The role of education and lifelong learning. *European Journal of Education* **52,** 404–413 (2017).
38. Leal Filho, W., Pallant, E., Enete, A., *et al.* Planning and implementing sustainability in higher education institutions: an overview of the difficulties and potentials. *International Journal of Sustainable Development & World Ecology* **25,** 713–721 (2018).
39. Tassone, V. C., O'Mahony, C., McKenna, E., *et al.* (Re-)designing higher education curricula in times of systemic dysfunction: a responsible research and innovation perspective. *Higher Education* **76,** 337–352 (2018).
40. Hill, D. J., Williams, O. F., Mizzy, D. P., *et al.* Introduction to Laboratory Safety for Graduate Students: An Active-Learning Endeavor. *Journal of Chemical Education* **96,** 652–659 (2019).
41. Johnson, J. & Kemsley, J. Academic lab safety under exam. *Chemical & Engineering News* **89,** 25–27 (2011).
42. Salazar-Escoboza, M. A., Laborin-Alvarez, J. F., Alvarez-Chavez, C. R., *et al.* Safety climate perceived by users of academic laboratories in higher education institutes. *Safety Science* **121,** 93–99 (2020).
43. Schröder, I., Huang, D. Y. Q., Ellis, O., *et al.* Laboratory safety attitudes and practices: A comparison of academic, government, and industry researchers. *Journal of Chemical Health & Safety* **23,** 12–23 (2016).

44. Pritchard, M. S. Responsible engineering: The importance of character and imagination. *Science and Engineering Ethics* **7,** 391–402 (2001).

45. Athanassoulis, N. & Ross, A. A virtue ethical account of making decisions about risk. *Journal of Risk Research* **13,** 217–230 (2010).

46. Steen, M., Sand, M., Van de Poel, I., *et al.* Virtue Ethics for Responsible Innovation. *Business and Professional Ethics Journal* **40,** 243–268 (2021).

47. Van Atteveldt, N., Tijsma, G., Janssen, T., *et al.* Responsible Research and Innovation as a Novel Approach to Guide Educational Impact of Mind, Brain, and Education Research. *Mind, Brain, and Education* **13,** 279–287 (2019).

48. Ansari, D., Coch, D. & De Smedt, B. Connecting education and cognitive neuroscience: Where will the journey take us? *Educational philosophy and theory* **43,** 37–42 (2011).

49. Frigo, G., Marthaler, F., Albers, A., *et al.* Training responsible engineers. Phronesis and the role of virtues in teaching engineering ethics. *Australasian Journal of Engineering Education* **26,** 25–37 (2021).

50. Graham, R. The global state of the art in engineering education. *Massachusetts Institute of Technology (MIT) Report, Massachusetts, USA* (2018).

51. Katajavuori, N., Lindblom-Ylänne, S. & Hirvonen, J. The significance of practical training in linking theoretical studies with practice. *Higher Education* **51,** 439–464 (2006).

52. Voon, X. P., Wong, L.-H., Chen, W., *et al.* Principled practical knowledge in bridging practical and reflective experiential learning: case studies of teachers' professional development. *Asia Pacific Education Review* **20,** 641–656 (2019).

53. Thurlings, M. & van Diggelen, M. Perceptions of practical knowledge of learning and feedback among academic teachers. *European Journal of Engineering Education* **46,** 139–160 (2021).

54. Wallin, P. & Adawi, T. The reflective diary as a method for the formative assessment of self-regulated learning. *European Journal of Engineering Education* **43,** 507–521 (2018).

55. Douali, N. & Jaulent, M. C. Integrate personalized medicine into clinical practice to improve patient safety. *IRBM. Digital Technologies for Healthcare* **34,** 53–55 (2013).

56. Gonin, P., Buchholz, C. J., Pallardy, M., *et al.* Gene therapy bio-safety: scientific and regulatory issues. *Gene Therapy* **12,** S146–S152 (2005).

57. Horst, M. Public Expectations of Gene Therapy: Scientific Futures and Their Performative Effects on Scientific Citizenship. *Science, Technology, & Human Values* **32,** 150–171 (2007).

58. Bystrykh, L. V., de Haan, G. & Verovskaya, E. Barcoded vector libraries and retroviral or lentiviral barcoding of hematopoietic stem cells. *Methods in Molecular Biology (Clifton, N.J.)* **1185,** 345–360 (2014).

59. Shahryari, A., Saghaeian Jazi, M., Mohammadi, S., *et al.* Development and Clinical Translation of Approved Gene Therapy Products for Genetic Disorders. *Frontiers in Genetics* **10,** 868 (2019).

60. What price affordable access? *Nature Biotechnology* **31,** 467–467 (2013).

61. Grosse, S. D. Economic analyses of genetic tests in personalized medicine: clinical utility first, then cost utility. *Genetics in Medicine* **16,** 225–227 (2014).

62. Rubin, D. B. & van der Laan, M. J. Statistical Issues and Limitations in Personalized Medicine Research with Clinical Trials. *The international journal of biostatistics* **8,** 18 (2012).

63. Hollands, G. J., French, D. P., Griffin, S. J., *et al.* The impact of communicating genetic risks of disease on risk-reducing health behaviour: systematic review with meta-analysis. *bmj* **352** (2016).

64. Lvovskiy, V., Agapov, A. & Mysina, T. Training of educational technicians specialized in activity-based learning at institutions of higher education. *SHS Web of Conferences* **98** (2021).

65. Howlett, C., Ferreira, J.-A. & Blomfield, J. Teaching sustainable development in higher education: Building critical, reflective thinkers through an interdisciplinary approach. *International Journal of Sustainability in Higher Education* **17,** 305–321 (2016).

66. Trevan, T. Biological research: Rethink biosafety. *Nature* **527,** 155–158 (2015).

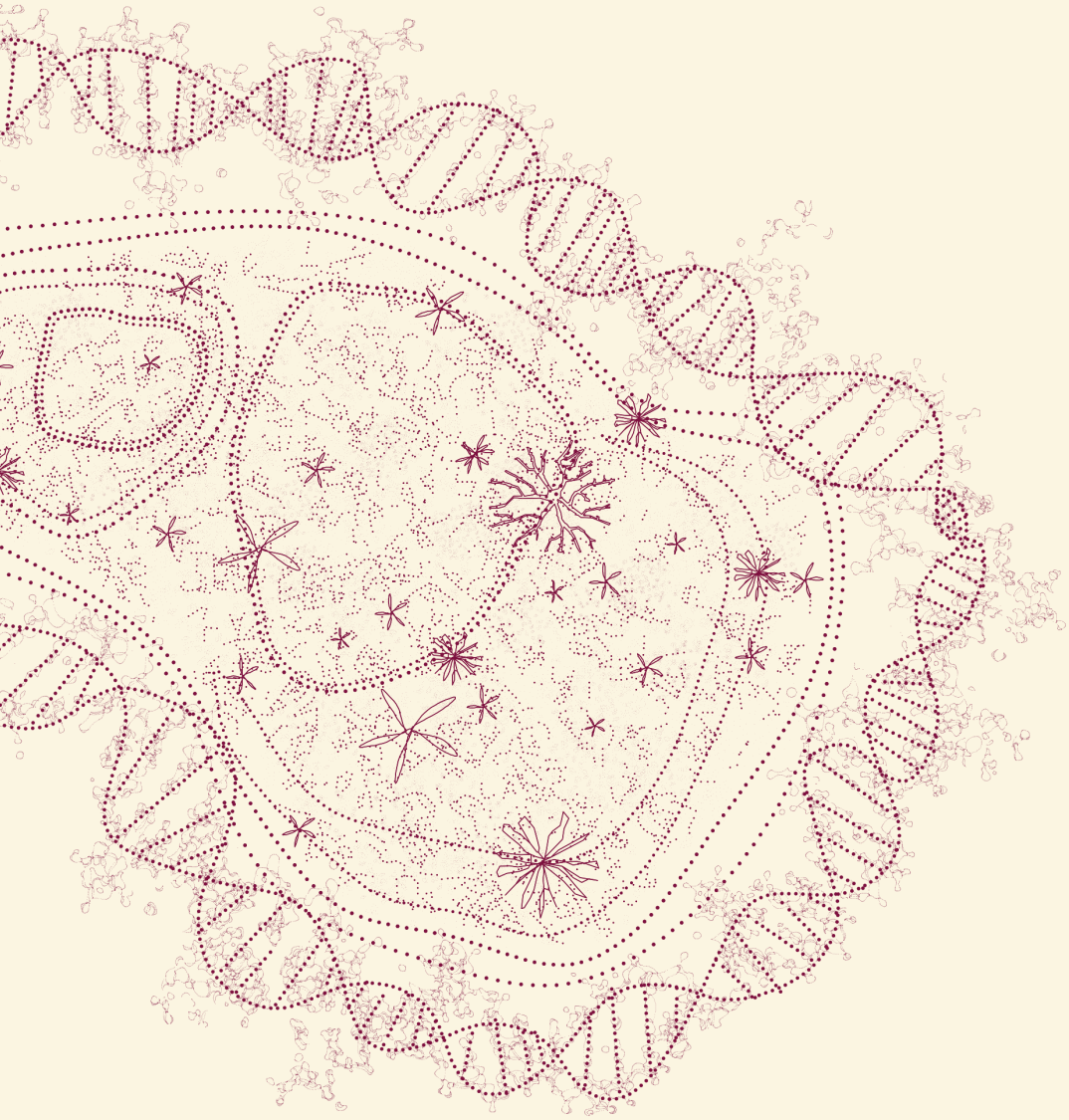
67. Cortés, J. M., Pellicer, E. & Catala, J. Integration of occupational risk prevention courses in engineering degrees: Delphi study. *Journal of Professional Issues in Engineering Education & Practice* **138,** 31–36 (2012).

68. Saleh, J. H. & Pendley, C. C. From learning from accidents to teaching about accident causation and prevention: Multidisciplinary education and safety literacy for all engineering students. *Reliability Engineering & System Safety* **99,** 105–113 (2012).

69. Shallcross, D. C. Safety education through case study presentations. *Education for Chemical Engineers* **8,** e12–e30 (2013).

70. Swuste, P., Galera, A., Van Wassenhove, W., *et al.* Quality assessment of postgraduate safety education programs, current developments with examples of ten (post)graduate safety courses in Europe. *Safety Science* **141,** 105338 (2021).

71. Liu, W. C., Wang, J. C. K. & Ryan, R. M. *Understanding Motivation in Education: Theoretical and Practical Considerations* (eds Liu, W. C., Wang, J. C. K. & Ryan, R. M.) 1–7 (2016).



CHAPTER 10

Navigating the Valley of Death: Perceptions of Industry and Academia on Production Platforms and Opportunities in Biotechnology

Linde F. C. Kampers, Enrique Asin-Garcia,
Peter J. Schaap, Annemarie Wagemakers, Vitor A. P. Martins dos Santos

Adapted for publication as:
From Innovation to Application: Bridging the Valley of Death in Industrial Biotechnology,
2021, Trends in Biotechnology
and
Navigating the Valley of Death: Perceptions of Industry
and Academia on Production Platforms and Opportunities,
2020, bioRxiv

Abstract

Rational lifestyle engineering through computational methods and synthetic biology has made it possible to substantially improve industrial performance of microbial cell factories which produce a range of bio-based chemicals in Industrial Biotechnology. However, only an estimated 1 in 5000 to 10000 innovations make it through the Valley of Death to market implementation. To gain in-depth insights into the views of industry and academia on key bottlenecks and opportunities to reach market implementation in Biotechnology, an exploratory study was performed by conducting fourteen in-depth interviews with eight industrial and four academic participants before information saturation occurred. The characteristics that any cell factory and bioprocess must have were schematically listed, and commonly recognised opportunities were identified. The results obtained were validated with in-depth interviews with two Technology Transfer experts. We found that academics are limited mostly by technical factors in their research, whereas industry is restricted in their research choices and flexibility by technical, sector-dependent and societal factors. This leads to a misalignment of interest between academics and funding industrial partners, often resulting in miscommunication and missed cooperation opportunities. Although both are of the opinion that academia must perform curiosity-driven research to find innovative solutions, there does exist pressure to aim for short-term industrial applications. These factors add up to widen the Valley of Death in Biotechnology. A third party, in the form of start-up companies, could be the answer to traverse this valley more effectively, in particular when embedded in adequate innovation ecosystems.

Introduction

The demand for industrial production by microorganisms is increasing rapidly. In 1990, the global market value of industrial enzymes was close to US$1 billion, crossed the US$2 billion mark in 2005 [1], was valued at over US$4.5 billion in 2016 and is expected to reach over US$6 billion in 2022 [2]. Nowadays, bacteria, yeasts, fungi, and microalgae are used in the industrial production processes of food, enzymes, vitamins, pharmaceuticals, biofuels, bioplastics, bio-insecticides, nanocomposites for electronic devices and a large variety of chemicals and enzymes with industrial value [3–5]. The reasons to widely apply a diverse set of microorganisms are manifold; they represent a broad biochemical diversity, increase the feasibility of large scale production, can produce sustainable products or reduce processing time, require low energy input, increase cost effectiveness and can be selected for non-toxicity [6–15]. However, as heavily as the current biotechnology industry depends on production by microbes, production methods are by far not as stable as desired due to the biological variability of microorganisms [16]. Many external factors affect the production capacity of microorganisms, including temperature, pH, oxygen, ion and carbon source availability, cell density, and biofilm formation [17–26]. All factors combined, this often results in a high variability in yield, product profile or selectivity. For industrial production purposes and within a given range of desired products, it might be ideal to select only one microorganism as the general production platform, tailoring it for specific needs. By focusing research and resources on a single microbe, the microbial toolkit would increase, production would stabilize, and production costs would decrease as more universal bioreactor set-ups could be used. An ideal industrial microbial workhorse must be safe to work with (preferably generally regarded as safe, or GRAS), genomically accessible, metabolically and environmentally flexible, resistant to external industrial stress factors, grow on cheap medium, and capable to produce a high variety of products with high and stable yields *à la carte* (upon appropriate tailoring) [27]. The field of systems and synthetic biology potentially offers rational solutions to these requirements. By restructuring the genome and the metabolism of microorganisms, microbes found in nature can be perfected to produce any *à la carte* product of interest in a safe way. Whereas random evolution and selection are common meth-

ods to increase productivity, yield or performance of a species, *in silico* systems and *in vivo* synthetic biology offers the tools to not only adapt microorganisms to gain novel characteristics in a directed way, but to create them. This goes far beyond the engineering of metabolic pathways for production of compounds and extends to re-programming the lifestyle of microbes to operate outside of their natural boundaries. An example of research conducted in academia to improve microbial workhorses for robust biocatalysis is one reported by Sandberg and colleagues, who aimed to render the industrially relevant *Escherichia coli* more resilient to temperature fluctuations [28]. In this study, *E. coli* K-12 MG1655, which has an optimal growth temperature at 37°C, was subjected to adaptive laboratory evolution to improve strain performance at 42°C, its upper critical growth temperature. Through rational engineering procedures, Kampers *et al.* recently reported a robust, rational process to increase tolerance to temperature shifts in both prokaryotes and eukaryotes microbes up to 42°C (patent WO2020094828A1). Another example is the lifestyle adaptation of the industrially applied *Pseudomonas putida*, which was enriched with three heterologous genes to better survive micro-oxic conditions through oxygen gradients [18]. So far, however, these examples are not applied in industry. The large gap for academic research to advance from the technological prototyping at laboratory scale to actual commercialization in industry is generally referred to as the "Valley of Death" [29–31]. In biotechnology, including pharmaceuticals and therapeutic medical device development, it has been estimated that only one in 5000 to 10000 innovations survive the long route from the initial findings to product commercialization [32–34]. This reflects a large break between results from academic research and industrial product development. The different levels of product and process development are known as technology readiness levels, or TRLs (Table 10.1) [35, 36]. Academic research generally ranges between TRL 1 to 3, while industry generally covers TRL 8 and 9. At TRL 4 to 7, the discovery process is often considered too applied for academic funding, whereas the industrial sector still considers it too risky to fund for market implementation [37]. As academics introduce scientific innovations, they often limit their research to proof-of-principle. Industry, on the other hand, needs marketable production rates, yields and titres that allow for a competitive business model. Even though this seems

Table 10.1: Technology Readiness Levels (TRLs) as instated by NASA [36].

TRL	Definition
1	Basic principles observed and reported
2	Technology concept and/or application formulated
3	Analytical and experimental critical function and/or characteristic proof-of-concept
4	Component validation in laboratory environment
5	Component validation in relevant environment
6	System/subsystem model or prototype demonstration in relevant environment
7	System prototype demonstration in relevant environment
8	Actual system completed and qualified through test and demonstration
9	Actual system proven though successful operation

rather straightforward, the implications are far-reaching.

Some reasons that new technology does not survive this Valley of Death include cumbersome contracting or procurement of technology requirements, lack of exposure, lack of entrepreneurial management, lack of adequate funding for further development, and the lack of a strong link between technology development efforts and industrial deployment [32–34, 38]. To decrease the amount of research that does not bridge the Valley of Death, funding agencies now often demand a direct collaboration between academia with industrial partners. Still, the Valley of Death remains as large as ever. Because of a perceived lack of systemic exploration of the factors that influence the Valley of Death, we aim to gain insight into this pressing matter by exploring the distance between academia and industry by conducting in-depth interviews amongst practitioners. Specifically, we aimed to ascertain I) what factors influence the choice for an industrial microorganism and II) opportunities to close the distance between academia and industry.

Methods

A qualitative and exploratory study was performed to gain a comprehensive insight into the perception of industry and academia by conducting in-depth interviews amongst practitioners. To this end, we approached multiple industrial and academic experts to assess their perceptions on how to increase chances of research surviving the Valley of Death. Results were validated in consultation with representatives of two Technology Transfer Offices (TTO).

Recruitment and procedure

Recruitment took place by gathering a convenience sample based on our direct network and visited conferences. A list was devised of possible participants, from which 15 companies were selected based on their location, Europe or America, and field of research, pharmaceuticals, food, industrial chemicals, or production organism development. Some companies selected cover a combination of these fields. A variety of large and small and new and established companies were selected carefully from each continent, to determine the effect of a different regulation in industrial applications of bacteria (GMOs in particular) and over different companies from different sectors. Chief executive officers or chief technical officers were mainly approached to ensure the most overhead view on proceedings and the clearest perspective on the general aim, flexibility and workings of the company. In three cases, we were redirected to someone else within the company. Of the fifteen companies approached, nine agreed to participate, five did not answer, and one declined participation citing confidentiality over proprietary information. Of the nine companies that agreed, one decided to withdraw from participating as they could not obtain consent from their legal department after receiving the transcripts. The distribution of companies over continent and field are shown in Table 10.2. Of the convenience sample, 4 European universities were selected based on their diverting expertise, similarly to companies. All universities approached agreed to participate. Three of the academic participants had present or past experience in working at a company. All participants were engaged in collaborations with multiple industrial sectors. The distribution of all the participants over continent and field are shown in Table 10.2.

Table 10.2: Participants in the interviews: the field occupied by participating companies and academics, and their location.

		Pharmaceuticals	Food	Chemicals	Development of production platforms	Combination	Total
Industry	Europe	1	1	1	1	1	5
	USA	1			1	1	3
Academia	Europe			1	1	2	4

Potential participants were approached up to three times *via* email to make an appointment. If possible, interviews were conducted in person, otherwise face-to-

face through Skype, WebEx or Bluejeans, according to the participants preference. During one interview, the camera of the participant failed, and the interview was done by voice only. Both through e-mail and at the start of each interview the aim and methodology of the research was explained to the participants. Before the start of each interview, participants were presented with written information and contact details, and were asked to sign a declaration of consent. Interviews were conducted over a period of 8 months, from April to November 2019. After 8 interviews, the content of answers started to show overlap with previous interviews indicating data saturation.

Interviews

Based on literature and experience [18, 39–43], a topic list was constructed to guide the interviews:

1. **Choice of production platform** was addressed by questions such as which production organism is used now, what factors may influence the choice for a production organism, including limitations or possible restrictions, company and set-up flexibility and what a production organism would need for the participating companies to switch.

2. **Opportunities between industry and academia** was addressed by questions including common ground, differences, challenges, opportunities, interesting fields of research, impact of science in industry, and collaboration. To this end, we included the presentation of a simplified overview of industrial production from the Design-Build-Test-Learn (DBTL) cycle [44] *via* laboratory scale to industrial production scale, where participants could indicate the main bottlenecks which indirectly represent the main short-term research opportunities for academia (Figure 10.1). Figure 10.1 was used to stimulate in-depth discussion on challenges and opportunities, allowing us to construct a plan to navigate the Valley of Death.

Prior to data collection, the interview questions were discussed with an expert working in academia and working closely with industry, from which adjustments were made. Adjustments were focused on making the questions more open-ended

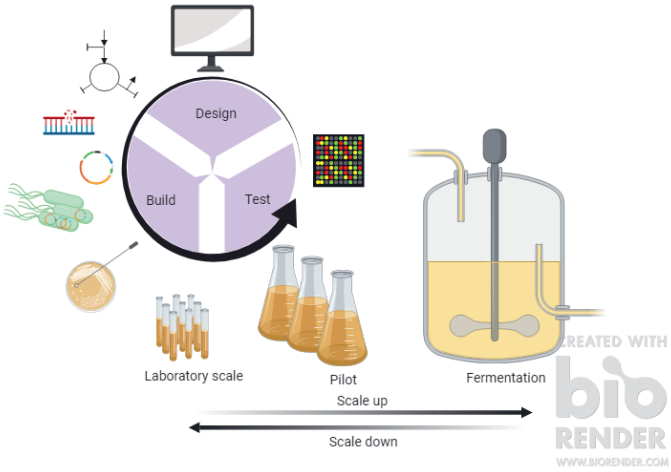


Figure 10.1: Overview of industrial development and production process. The schematic covers the main steps between the DBT-cycle and the industrial production scale.

and general to prevent bias and allow for a more in-depth discussion by leaving room for follow up questions. The use of appreciative inquiry in the questions [45] encouraged and inspired the participants to answer according to their own perspectives, ideas and experiences, as opposed to following a strict interview structure. In practice, this means that questions are asked with a positive spin, to focus more on solutions rather than on problems. A pilot interview with an expert working both in industry and academia was then held, indicating that the qualitative face-to-face interviews offered the expected in-depth information and clarity for both the participant and the interviewer, as it allowed for elaboration on specific topics, insight into technical differences or company aims. Special care was taken during the interviews to make sure there is a consensus of participants on definitions and terms used, by asking follow-up questions. Each interview was conducted by two members of the research team. The 13 interviews lasted on average 41 minutes, ranging from 26 to 54 minutes.

Recording and analysis of data

With the participants' verbal and written consent, all interviews were audio-taped and transcribed (intelligent verbatim style). The data was pseudonymized. When necessary, audio files were used to confirm transcripts and listen to excerpts within

their original context. Transcripts and specific quotes used from each transcript were shared with the participants to consent upon prior to analysis. All participants responded, and minor editorial changes were made. Data was processed using the six steps for qualitative data analysis [46]. All analyses were done by the two interviewers, LK and EAG. In the first step, transcripts were made and carefully read. In the second step, codes were formulated manually using the research questions. Thirdly, the transcripts were coded using the formulated top-down codes, and where necessary deriving new codes from the data. The processing of the transcripts was conducted *via* inductive category formation using QCAmap [47], according to a thematic content analysis [46, 48]. The interviewers discussed their respective individual top-down and bottom-up codes to define a final set of codes. In the fourth step, all codes were clustered into the research questions: A) Choice of Production Platform, and B) Opportunities between Industry and Academia. In step five and six, recurring themes found to affect multiple codes were derived from the data and defined carefully. The nine themes found pertain either one or both research questions. The themes also led to the establishment of an overarching topic: all themes found relate to a difference in aims of industry and academia. To support the themes found, direct quotes are provided in the result section, selected to represent participants' opinions, views and experiences. Table 10.3 presents an overview of the clusters, codes and derived themes.

Consultation with TTO representatives

Technology transfer programs aim to assist and empower researchers to take their ideas to the next levels by stimulating, supporting and accelerating (sometimes through brokering seed- and soft-funding) the application of promising research outcomes from an early stage. Technology transfer offices and other innovation-oriented vehicles may provide a powerful impulse to potential entrepreneurs who want to evaluate, estimate and capitalize curiosity-driven and innovative ideas on the market. To validate the data, representatives of a Dutch and a Danish TTO were consulted to discuss the biotechnological innovation process and the Valley of Death.

Table 10.3: Themes, research questions and codes of transcriptions. Top-down codes were formed before coding, bottom-up codes were derived during transcription.

Themes	Top-down codes	Bottom-up codes
Research Question 1: Choice of production platform		
• I) Characteristics of microbial production platform • II) Industrial process and set-up flexibility • III) Regulations • IV) Public perception	• What production platform is used • Characteristics of used production platform • Flexibility on industrial set-up	• Limitations on choice of production platform • External influences on choice of production platform
Research Question 2: Opportunities between industry and academia		
• V) Genome engineering • VI) Scale-up • VII) Increase strain robustness	• Opportunities and developments • Distance between industry and academia • Impact of industry on academia	• Personal or company aim • Influence of industrial experiences
The aims of industry and academia		
• VIII) Curiosity-driven versus solution-based research • IX) The rise of small companies	• What production platform is used • Characteristics of used production platform • Flexibility on industrial set-up • Opportunities and developments • Distance between industry and academia • Impact of industry on academia	• Limitations on choice of production platform • External influences on choice of production platform • Personal or company aim • Influence of industrial experiences • Communication

Data protection and transparency

Identities of participants and interview data remain protected, stored in secure university locations according to the local protocols for safe data management. The aims and methods of the study were explained to the participants beforehand. All participants provided informed consent and agreed to publication of the anonymized data.

Results

The main results are presented in line with each of the greater themes (Table 10.3).

(A) Choice of production platform

Four main themes pertain the choice for a certain production organism: I) Characteristics of microbial production platform, II) Industrial process and set-up flexibility, III) Regulations, and IV) Public perception (Table 10.4).

Table 10.4: Factors affecting the choice of production platform with representational quotes from participants from industry (I) or academia (A), indicated by their numerical identifier.

Factor	Representative Quote	Source
I. Characteristics of microbial production platform		
Titre-rate-yield	"The microorganism itself is not so relevant, the whole picture should be okay. Generally speaking, you can say that for any fermentation process you need to have a good titre, a good rate and a good yield, so the famous try-numbers, titre-rate-yield."	I.2
Toolbox availability	"I think trying to predict and overcome toxicity problems of given molecules to given organisms is very difficult. Trying to adapt one organism to secrete something where another does it more naturally is very difficult. I don't see why you would do it, I think it is a better strategy to develop the broadest range of tools for different organisms including extremophiles."	I.5
	"Genetic accessibility, wealth of available data, free available data."	A.6
Genomic accessibility	"If a host is genetically accessible we can start to metabolically engineer it from day one, that is preferred to do six or nine months of method development to even get DNA inside the cell."	I.12
	"[The main advantage] is that we don't have to worry about the genetic tools."	A.11
Product of interest	"The higher the cell complexity is, the more advanced the product should be. This statement somehow mimics necessary R&D expenses that need to be covered by cost calculations."	A.9
	"These microbes have not been evolved to make a 100 grams per litre of compounds A or B, so we typically start with a selection where we screen different hosts for tolerance towards the products."	I.12
Expertise available	"Everybody needs to be trained on the new host, and please don't forget that as soon as you have transferred something to a production line, there will be lots of questions coming up in the aftermath. For instance, downstream processing matters need to be checked and must fulfil given constraints, too. That requires experts who care for those questions."	I.9
Production of unwanted side-products	"We produce flavour and fragrance ingredients (...) they should smell and taste in the right way afterwards, after purification, so what we have to avoid is off odours."	I.3
Fundamental interests	"There is this fundamental interest in understanding how new insights, new activities in new networks emerged under natural conditions because if we are able to domesticate the corresponding mechanisms then we should be able to reproduce them in the laboratory a nd make them work much faster and much more efficiently. The basic questions that we are interested in are not limited to [microorganism], it just happens that [microorganism] is a very good experimental system to address them."	A.10
II. Industrial process and set-up flexibility		
Investment requirements	"It depends on the numbers, and on the cost, but changes in bioreactor setup of course mean that you have to either build or find another bioreactor."	I.3

Choice of industrial setup	"[The producer] might want a semi-continuous process. The cell may have high productivity but it dies after a week; they will swap that out to something that has a lower productivity but can last for three weeks in a continuous or semi-continuous culture. It's all about space-time yield for low value products."	I.5
Downstream processing	"There is also the downstream processing part. In essence, you need to separate your product equally simple from the broth as you can do with established hosts."	I.9
Complete industrial process	"If your fermentation is going too fast, faster than the downstream can handle, you're creating a bottleneck."	I.2
Stage of development	"[Producers] will like small changes where they don't have to invest in new capital (...) but if it's a process that's not been implemented at pilot scale then they're more likely to make a radical change as long as it's consistent with their engineers' understanding of what the eventual process will look like."	I.5
III. Regulations		
GMO regulation	"If you're heading for products of food, feed, flavour or fragrance markets, you may face severe opposition regarding the use of recombinant strains, per se."	I.9
Safety regulations	"All proteins that are produced for food and feed, they have to apply to EFSA or FDA regulations, and the more similar you can do a process, the more easy it is to get permission."	I.4
IV. Public perception		
Social acceptability	"Sometimes the social acceptability of it as well. I mean you might find that Klebsiella or Salmonella would be good hosts to make something and you could use perfectly viable lab strains. But they might prefer Lactobacillus because it's a food grade organism and public perception ultimately might be very different."	I.5
Halal or Kosher	"It's not just GMO regulations, it's not just FDA and what have you, it's also these kinds of considerations [like halal or kosher] that must be taken into account. So you really have to think through all the consequences if you want to introduce a new microbe in your process and make other products."	I.2

The factors that influence the choice of production platform in industry and academia can be divided into technical factors, such as the characteristics of the production platform and industrial process and set-up flexibility; sector-specific factors like regulations; and social factors, including public perception. Many of these technical, sector-based and social factors restrict a company when choosing a production host. As they must keep the end-user in mind throughout the entire development and production process, products must be made as cost-effective, fast, environmentally-friendly and safe as possible [49]. For academia, mostly technical factors play a role when deciding to work with a specific microorganism. These boil down to availability of a suitable genetic and biochemical toolbox, genomic accessibility, product of interest and fundamental interests. Only the latter is specific to academia, the rest is shared with industry. This means that academics are freer to work with any microorganism, even with microbes that cannot be applied *yet* to industrial processes.

(B) Opportunities between academia and industry

To determine what the main future opportunities are, participants were asked for future research of interest, the main bottlenecks in their production process and possibilities, and what scientific progress they expected to have most future impact (Figure 10.2).

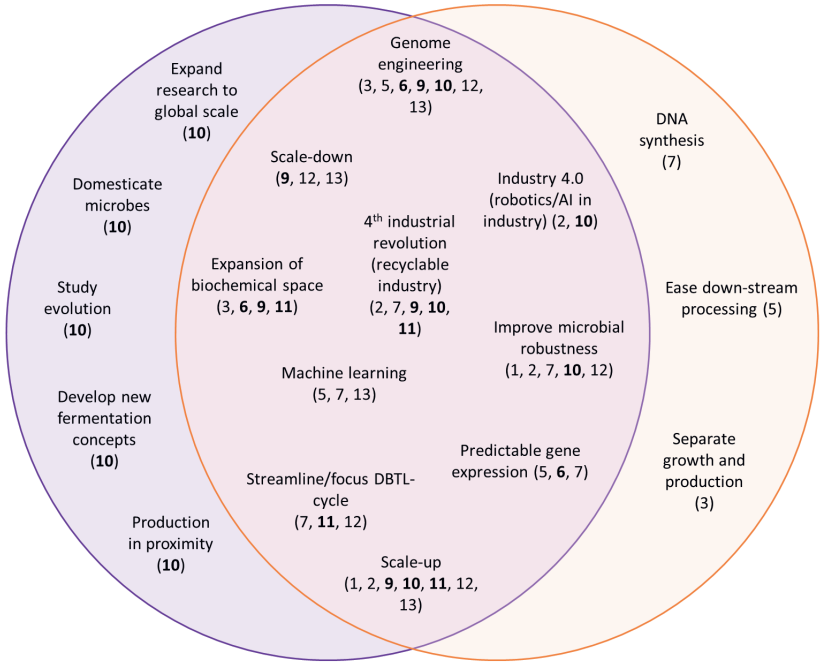


Figure 10.2: Venn diagram of research areas of interest per field. A schematic overview of the research areas indicated as interesting by academia (left) and industry (right). Overlapping research of interest is in the middle. Per subject, the identifier of each participant is indicated. Identifiers of academic collaborators are printed in bold.

From this overview, three main technical themes could be deduced: V) Genome engineering, VI) Scale-up, and VII) Improve robustness. These will be discussed in more detail.

V) Genome engineering

When presented with Figure 10.1, most industrial participants identified (parts of) the DBTL-cycle as the main opportunity for improvement.

> "You test [a new design], you find something out and then you have to design, you get a lot of data, you get some results, you don't understand them, and you have to analyse. The learn part is the bottleneck because all the other things are automatable or scalable to some extent, so the learning bit is not." *Interviewee 3*

For industry, the possibility to affordably construct an entire genome combined with the rise of novel tools such as CRISPR-Cas for genome engineering significantly ease this bottleneck.

> "On the build side [of the DBTL-cycle] there's a lot of change. Synthetic DNA is becoming cheaper and cheaper by the day, sequencing is becoming cheaper by the day, CRISPR is like 6 years old and [recently] there was a new version released that really put the whole town plus all of the genetic engineering upside down. [...] It's just changing the timelines and the ability so much." *Interviewee 12*

For academia, such tools pose promising ways to improve upon fundamental research, which is required for understanding strain design.

> "I would hope that it will accelerate the way at which we can generate industrially meaningful strains, and it will expand the extent of manipulation that we can achieve. So rather than looking at one gene or one pathway, we should hopefully be able to look at one entire section of metabolism in one go, and this would hopefully have a very good influence on the strain design that we want to achieve." *Interviewee 6*

VI) Scale-up

To reliably scale-up results obtained under lab-conditions to industrial scale is challenging at best. Scaling-up was recognised as the second main bottleneck.

Even in academia, the importance of reliable scale-up to bridge the Valley of Death was recognised:

> "You can develop the most sophisticated CRISPR system, the most sophisticated recombineering, the most sophisticated reactions, you can do wonderful things in the lab: if you cannot scale them up, then industry will not be interested, period." *Interviewee 10*

From an industrial point of view, there are few different reasons why scale-up can be a bottleneck, besides technical difficulties:

> "The availability of equipment is actually the whole bottleneck. There is not a lot of pilot facilities in the US, so either you build your own or you go to Europe for piloting. That's non-trivial. I think there's a lot of SynBio companies that are mostly comprised of genetic engineers or metabolic engineers that have never seen a large-scale fermenter and a large-scale production process including purification. So, if you start a scale-up trajectory without people that know what they're doing then it's going to be a bottleneck. But if you have a team, [...] then you should be able to do that in a pretty smooth trajectory. There's always going to be surprises and hiccups. What I've also seen is that lots of SynBio companies have very aggressive timelines, probably due to their investors pushing really hard, and then they are going to take short-cuts and that's going to blow up in your face. There is not a technical bottleneck, but it's more of a managerial bottleneck at that point." *Interviewee 12*

VII) Improve strain robustness

When asked directly, all participants indicated that the robustness of their used or preferred production platform could be improved. Reasons for improving robustness circle back to scale-up without fail, since industrially used production organisms are grown in conditions that are strictly controlled.

> "Robustness in an industrial sense is that it is operationally robust: meaning it is not very sensitive to infections, it is not very sensitive to

> minor variations that you always have in the conditions of the process, *e.g.*, due to scale. So, there should be some tolerance in the microbes to slight variations in the process conditions, which of course we try to control that as much as possible." *Interviewee 2*

In the few cases in which robustness was not recognised as an issue, it was solely because the production platform of interest already performed optimally at large scale:

> "We experience our microorganism is very robust under the conditions that are relevant to us, that is industrial fermenters, so we don't see a need for further improvement there." *Interviewee 3*

Robustness is thus only deemed important to aid production platforms to perform reliably on industrial scale. This link was made again by academics:

> "Fluctuations on large scale are much more prominent then on small scale, and constant performance through larger fluctuations in large scale most probably could be improved." *Interviewee 6*

Overall, the long- and short-term innovation interests of academia and industry align very well. However, there exist differences in interest. In industry, research interests are often very practical, focused on an increase in titre-rate-yield. Subjects included the use of a different C-source, separated growth- and production cycles, and easing downstream processing. Amongst academics, research interests are of a more fundamental or futuristic nature. These included for example the development of new fermentation concepts, expanding of research to global scale and studying evolution.

The aims of industry and academia

All nine themes together lead to the overarching main topic: a difference in aim between industry and academia. Two themes proved of vital importance as they influence both research questions: VIII) Curiosity-driven versus solution-based research, and IX) the rise of small companies.

VIII) Curiosity-driven versus solution-based research

The precarious balance between curiosity-driven and solution-based research, the latter mainly with industrial application in mind or on stakeholders' demand, can be defined most clearly by looking at the collaborations between academia and industry. All participants, both from industry and academia, indicated current involvement in collaborative projects. However, these collaborations are not necessarily initiated by industry nor based on industrial interests. By academics, it was more described as follows:

> "It's rarely the case that somebody [from industry] comes to me and says "*This is something that you should work on*", but of course you talk to people at conferences, at seminars, and you try to understand what the problems are, and you try to find out if you can solve some of these problems with the tools that you have available. But it's a bit of a balance, I try to make sure that a substantial part of our activities does not have a direct line. In a way, it's curiosity driven." *Interviewee 6*

Collaborations between academia and industry are not based on a demand-supply model, but rather on academics trying to fill existing or predicted gaps in industry. All academic partners did indicate a strong influence on their work from industry, even if not directly. However, a way to deal with this potential struggle was explained:

> "We are driven by curiosity and obviously we have to frame our curiosity and our personal interest in the interest of our sponsors, and in the interest of society, and this is perfectly fine. Just remember that the big artist in the Renaissance were sponsored by the popes, by the emperors and so, and [the artists] had to produce things that [the sponsors] liked. I think this is not completely shocking that we are constrained by the ideas of our sponsors. I think that our challenge [...] has a point even of enjoyment. It is how you connect your own personal interest and our own curiosity with these big demands that your sponsors have for your research." *Interviewee 10*

Industry thus does not approach academia with questions, but rather shops around for interesting opportunities they can easily adapt and apply. This indicates that they rely heavily on research performed in academia to make large strides in innovation. Industry indicates this when finding new investment opportunities:

> "The biggest opportunities lie in the combination for a company like us, or a hospital research group for example, that have developed something novel, something proprietary, but [they, academia] have no understanding of the hurdles they'll face or the means to demonstrate its viability to be a competitive product in the market. Because that's what's going to be necessary to attract investment. To give it a chance of survival." *Interviewee 5*

The clear distinction between technical, sector-driven and social factors influencing main research choices and definition of opportunities can be derived from the difference in the main aim of each field. Where industry looks for applications and adaptable solutions to day-to-day problems, academics aim to discover answers to fundamental questions without being limited by end-user demands. Communication to align these different aims towards one goal is vital:

> "First of all, there has to be good communication, because academia works in a completely different way to industry. Industry is all about timelines, deadlines, deliverables, project management, progress towards measurable output. Academia is about testing, discovery, interesting products, interesting processes, new methods, means of discovering where science can go." *Interviewee 5*

IX) The rise of small companies

Small companies that arise from academia to commercialise promising academic findings are called start-ups. Start-ups do not only focus on the academically recognised technical factors; they include the sector-based and social factors that might influence the successful industrial introduction of their product, focusing specifically on TRL 4-7. Some see start-ups as a natural way to close the gap between

academia and industry, by presenting interesting research ready for commercialisation on a silver platter:

> "It's those partnerships in between, it's the networking and the identification of viable product opportunities that are emerging from either academic groups, hospital groups, virtual companies and then the partnerships that takes that feasibility demonstration onto viability. That's where the opportunity is. Bridging that gap from feasibility to viability. Because that's what derisks it, and brings in funding to take it forward. I think otherwise so many things can sit in academia, very interesting, very exciting, and probably more so in Europe than in the United States, and they just don't see the light of day because they haven't taken the next step and that's where certainly a company like ours, but other small companies can come in and make that difference." *Interviewee 5*

Other interviewees indicate that the rise of start-ups widens the communicative and collaborative distance between academia and industry. However, they do understand why companies prefer dealing with start-up companies over academics:

> "What we observe is that all these new developments, new technologies and everything, is mostly the business of small companies. The big companies just become very conservative. [...] Big companies prefer to deal with smaller companies than dealing directly with scientists, and the reason [...] is that scientists have too big a mouth. That means that the moment you have discovered something interesting you immediately start spreading the word, and this is something that the companies don't like at all." *Interviewee 10*

Consultation with TTOs

The consultations with TTOs confirmed the results described and contributed to contextualising our data interpretation and resulting discussion. In particular, the TTOs provided a richer view of the transitions and processes involved. Of common note was the observation that such transitions may be more likely to succeed when embedded in or supported by adequate innovation ecosystems, which can entail

innovation incubators, accelerators, or non- and for-profit commercial foundations. Unsurprisingly, availability of seed- and soft-funding was seen as a key factor, along with entrepreneurial know-how and exposure.

Discussion

The importance of technical, sector-based and societal factors

Three groups of factors were recognised when considering the choice of microorganisms currently applied in industrial biotechnology: technical, sector-based and social. Current host systems widely applied include bacteria, yeasts, filamentous fungi and unicellular algae. Industrially applied strains all come with their own strengths and drawbacks [6, 50, 51]. Technical factors that influence the choice for production platform comprise, among others, the ability to produce complex chemicals, growth rate, the ability to perform post-translational modifications and the use of cheap medium components. Sector-based factors include the used microorganism to be a natural producer of the product of interest, the cultivation and production process to fit within the complete industrial process, or familiarity with working with a specific strain from out the company history. Societal factors for the wide application of these production hosts include public familiarity, an environmentally sustainable production process and the safety status of a species. Academic participants only cited technical reasons for the selection of a microorganism such as ease to work with, genomic accessibility and wealth of publicly available data. In industry, sector-based factors and social influences derived from demands of the end-user or regulations applied by the government also strongly influence choices made. Despite this, the main innovation opportunities identified are of a technical nature: genome engineering, predictability, stability, scale-up, and increased strain robustness. Any innovation not making it through the Valley of Death thus occurs when the societal and sector-based aspects come in play. Changing the production organism used in an established industrial process is considered a radical industrial shift, which heavily depends on research. Adapting an existing production process is expensive, requires specific scientific knowledge and takes time to develop and test. As changing a production platform is generally

bound to affect the industrial setup, the gain in yield and characteristics has to be substantial to make up for the investment costs. Additionally, microbes with more interesting traits which better produce specific substances are being discovered every day. It is thus not surprising that there is a lot of research focused not on finding the perfect production host for specific purposes, but on the *de novo* design and construction of one, either by making a cell including only traits specifically selected, or by rewiring and streamlining an existing production microbial platform. Computational methods are paving the way towards big data processing, enabling the on-demand production host design in any genomically available species. Despite this, implementation and testing of many different designs through the DBTL cycle and unpredictability of gene expression were identified as main bottlenecks by both industrial and academic participants.

The rise of start-ups

The distance between academia and industry arises from the difference in their perspectives. Where academia focuses on understanding a microorganism on a fundamental level, industry always looks at the entire production process. Industry works on solving the problems of today, academia focuses on the problems of tomorrow. Simultaneously, both parties need each other for funding, direction, creativity, innovation and exploring boundaries. Both parties also agree that better communication would improve cooperation and increase opportunities for process or product development towards actual market introduction. As earlier indicated, academia works mostly at TRL 1-3 while industry tends to focus on TRL 8-9. To bridge this gap, a solution may have naturally emerged: start-up companies. Rather than direct translation of academic research and technology transfer to industry, companies have emerged directly from academia, which when successful are often acquired by larger companies of investment groups. Start-up companies are closely linked to both industry and academia. They are set up to operate in TRL 4-7, more willing to take risks, and use a versatile production platform. The continuous process of using novel advanced technologies to build a firm has moved to the heart of innovation strategy [52]. By presenting the best, most lucrative results achieved in academia on a silver platter, start-ups attract initial funding and stir commercial

interest. This trend currently courses through many different sectors. A well-known example of the once-start-up turned multinational IT company Facebook acquiring multiple start-ups [53]. In the early twentieth century, government-funded science fostered a first academic revolution. Entrepreneurial science, where collaborations were formed with industry, chimed in the second academic revolution [54]. Nowadays, government funded research often demands collaborations with industrial partners, as seen in European projects such as EmPowerPutida, P4SB or IBISBA [55, 56]. Such collaborations limit publication opportunities: scientific excellence is often measured by their publications, industry can only allow publications results concerning TRL 1-3 due to patenting potential, creating a natural stop of research performed by academics (Box 10.1). The swift rise of start-ups has chimed in a third academic revolution, where securing the IP of novel discoveries will grow in importance. This impacts the role of academia, which is pushed towards faster development of more applied research.

Box 10.1: Other limitations in technology transfer

Although the main focus here is on problems, limitations and miscommunications occurring during the attempt to traverse the Valley of Death, there are many reasons for academics to not even attempt to bring their innovation to the market. Being driven by the need to teach, conduct research, write papers, and finance their efforts, there is simply no incentive for marketing. Not only does the trajectory cost time and resources better spent elsewhere, but the rewards are also highly limited. Risk of failure is simply too high [29]. In addition, academics must decide themselves which innovations are worth pursuing, and which they think will not make it across the Valley of Death. Communication in these early phases is highly limited: there is a constant pressure in academia to publish or risk their research being scooped. Many academic institutions are now providing aid in the form of patenting offices, start-up support and technology transfer experts [57], but still a strong incentive is required to pursue innovation application.

All academic and some industrial participants have indicated a growing distance between academia and industry, affecting the communication and collaborations. Some have indicated this growing distance to be worrisome: as start-up companies present commercialised techniques that can be quickly integrated in any company, they mask the years of university-driven research and development that occurred before forming the start-up. Industry then much rather obtains the techniques

or products from the start-up company directly than investing in the necessary academic research.

TTOs and other innovation agents (such as non- and for-profit foundations) support the creation of spin-offs and start-ups, which are positioned in a higher level of circularity in their businesses when compared with bigger corporations, rendering them a logical bridge between discovery and application. TTOs and other innovation agents contribute to educate in entrepreneurial skills including finances and social abilities to provide the means for the development and often implementation of a solid business model, fostering entrepreneurial exposure and assisting with funding opportunities (*e.g.*, through seed funding or investors). Alternatively, new spin-off companies can emerge from larger corporations that acquire academic IP to capitalise company creation with a new team of professionals. In any case, traversing the Valley of Death is a long journey that needs a multistakeholder effort and support from many associated organisations and institutions. In terms of circular economy, for example, policy makers can be among the biggest innovation drivers *e.g.*, through regulation or measures that stimulate the development of products and processes for more sustainable solutions. The natural rise of start-ups and spin-offs should be seen as an opportunity. This type of companies can ease the transition of academic research to actual industrial application, closing the Valley of Death by forming an independent bridge between academia and industry [58–60]. It also opens up room for academics to focus on what they do best: education and research. It was earlier recognised [61] that having opportunities in academia to translate their research effectively into concrete products benefits academic institutions, faculty members, industry, and society [54, 62]. A case could be made for collaborations to include of the possible development of start-ups from their inception, which might lead to more attention to promising results and how to grow them into market applications from an early point on.

Strengths and limitations

After approximately eight interviews, data saturation occurred: no new information was obtained. This indicates a good reflection of the target community. The large spread of different companies over different biotechnology sectors, of different

sizes and spread over different locations around the globe allowed to reasonably generalise the findings so that the universal bottlenecks, challenges and opportunities could be easily recognised. In addition, data was validated and confirmed by two TTO representatives. An alternative interesting approach to this research could be to assess the subtle differences from companies in different continents, expanding on not only sectorial but also cultural and market-related differences. A larger sample would be more reflective of the continental community, including academic instances from other continents than Europe.

Conclusions

Within industrial biotechnology, research on suitable production platforms is paramount. Existing production platforms are improved, or new ones are developed. From an academic point of view, accessibility, and ease to work with, along with scientific interest, are the main reasons to choose one specific microorganism. For industry, technical factors such as the intrinsic characteristics of the production platform, industrial process and set-up flexibility, sector-based factors such as regulations, and social factors such as public perception are important influences. To ensure that research innovation makes it through the Valley of Death to market commercialisation, these factors must be considered early on. Universally recognised research topics of interests are genome engineering, scale-up, and improved strain robustness. The main cause for lack of market introduction are the vastly different TRLs that academia and industry operate on. Industry mostly focuses on solution-based research *e.g.*, to improve the titre-rate-yield, whereas academia generally focuses on curiosity-driven, fundamental research, vast steps removed from industrial application. Neither party really wants this to change, while both indicate a need for improved communication to avoid lost opportunities [57]. As confirmed by the technology transfer offices, start-ups, in particular when embedded in advanced innovation ecosystems, can serve as the bridge over the Valley of Death, connecting feasibility to viability, acting as an active broker between academia and industry to set-up, sharpen, develop and fund innovation to implementation.

Acknowledgements

The authors gratefully acknowledge financial support from the Wageningen University & Research IP/OP project, the TTW-NWO project SafeChassis (grant agreement No. 15814) and the European Horizon 2020 projects EmPowerPutida (grant agreement No 635536) and IBISBA (grant agreements No. 730976 and 871118). The funders had no role in preparation of the manuscript. We are very thankful to all our interviewees for participating in our research, for their time and for sharing their views. We thank Ruud A. Weusthuis as well for his valuable advice on various aspects of the interview.

Conflict of interest

The authors declare there are no conflicting interests.

Author contributions

Conceptualization: VAPMdS/LFCK; Literature research: LFCK/EAG; Data curation: LFCK/EAG; Formal analysis: LFCK/EAG/AW; Funding acquisition: VAPMdS; Investigation: LFCK/EAG/VAPMdS; Methodology: LFCK/EAG; Supervision: AW/VAPMdS; Writing - original draft: LFCK/EAG/PS/AW/VAPMdS; Writing - review & editing: AW/PJS/-VAPMdS

Bibliography

1. Binod, P., Palkhiwala, P., Gaikaiwari, R., *et al.* Industrial Enzymes - Present status and future perspectives for India. *JSIR Vol.72(05) [May 2013]* (2013).
2. MarketsandMarkets. *Industrial Enzymes Market by Type & Source - Global Forecast 2022 | MarketsandMarkets* 2016.
3. Arora, N. K., Mishra, J. & Mishra, V. *Microbial enzymes: roles and applications in industries* (2020).
4. Danielson, N., McKay, S., Bloom, P., *et al.* Industrial Biotechnology—An Industry at an Inflection Point. *Industrial Biotechnology* **16,** 321–332 (2020).
5. Rastegari, A. A., Yadav, A. N. & Yadav, N. *New and future developments in microbial biotechnology and bioengineering: Trends of microbial biotechnology for sustainable agriculture and biomedicine systems: perspectives for human health* (2020).
6. Adrio, J.-L. & Demain, A. L. Recombinant organisms for production of industrial products. eng. *Bioengineered Bugs* **1,** 116–131 (2010).
7. Dien, B. S., Cotta, M. A. & Jeffries, T. W. Bacteria engineered for fuel ethanol production: current status. en. *Applied Microbiology and Biotechnology* **63,** 258–266 (2003).
8. Diggins, F. W. E. The true history of the discovery of penicillin, with refutation of the misinformation in the literature. English. *British Journal of Biomedical Science; London* **56,** 83–93 (1999).
9. Gupta, R., Gupta, N. & Rathi, P. Bacterial lipases: an overview of production, purification and biochemical properties. en. *Applied Microbiology and Biotechnology* **64,** 763–781 (2004).
10. Mielenz, J. R. Ethanol production from biomass: technology and commercialization status. *Current Opinion in Microbiology* **4,** 324–329 (2001).
11. Nogi, M. & Yano, H. Transparent Nanocomposites Based on Cellulose Produced by Bacteria Offer Potential Innovation in the Electronics Device Industry. *Advanced Materials* **20,** 1849–1852 (2008).
12. Povolo, S. & Casella, S. Bacterial production of PHA from lactose and cheese whey permeate. en. *Macromolecular Symposia* **197,** 1–10 (2003).
13. Rogers, K. & Kadner, R. J. *Bacteria - Bacteria in industry* en. Chicago, IL, United States, 2019.
14. Singh, R., Kumar, M., Mittal, A., *et al.* Microbial enzymes: industrial progress in 21st century. *3 Biotech* **6** (2016).
15. Vuyst, L. D. & Leroy, F. Bacteriocins from Lactic Acid Bacteria: Production, Purification, and Food Applications. *Journal of Molecular Microbiology and Biotechnology* **13,** 194–199 (2007).
16. Pilyugin, S. S. & Waltman, P. Multiple limit cycles in the chemostat with variable yield. *Mathematical Biosciences* **182,** 151–166 (2003).
17. DiCesare, C., Yu, M., Yin, J., *et al.* Development, qualification, and application of a bioreactor scale-down process: modeling large-scale microcarrier perfusion cell culture. *Bioprocess Int.* **14,** 18–29 (2016).
18. Kampers, L. F. C., van Heck, R. G. A., Donati, S., *et al.* In silico-guided engineering of Pseudomonas putida towards growth under micro-oxic conditions. *Microbial Cell Factories* **18,** 179 (2019).
19. Lai, B., Nguyen, A. V. & Krömer, J. O. Characterizing the Anoxic Phenotype of Pseudomonas putida Using a Bioelectrochemical System. *Methods and protocols* **2,** 26 (2019).
20. Lorenzo, V. d., Prather, K. L., Chen, G.-Q., *et al.* The power of synthetic biology for bioproduction, remediation and pollution control: The UN's Sustainable Development Goals will inevitably require the application of molecular biology and biotechnology on a global scale. en. *EMBO reports,* e45658 (2018).

21. Niess, A., Löffler, M., Simen, J. D., *et al.* Repetitive short-term stimuli imposed in poor mixing zones induce long-term adaptation of E. coli cultures in large-scale bioreactors: experimental evidence and mathematical model. *Frontiers in microbiology* **8,** 1195 (2017).

22. Nikel, P. I. A Brief Guide to Pseudomonas Putida as a Microbial Cell Factory. *Bioessays* (2013).

23. Olughu, W., Deepika, G., Hewitt, C., *et al.* Insight into the large-scale upstream fermentation environment using scaled-down models. *Journal of Chemical Technology & Biotechnology* **94,** 647–657 (2019).

24. Reniere, M. L. Reduce, Induce, Thrive: Bacterial Redox Sensing during Pathogenesis. en. *Journal of Bacteriology* **200,** e00128–18 (2018).

25. Sporer, A. J., Kahl, L. J., Price-Whelan, A., *et al.* Redox-Based Regulation of Bacterial Development and Behavior. *Annual Review of Biochemistry* **86,** 777–797 (2017).

26. Steen, A., Utkur, F. O., Borrero-de Acuna, J. M., *et al.* Construction and characterization of nitrate and nitrite respiring Pseudomonas putida KT2440 strains for anoxic biotechnical applications. *Journal of Biotechnology. From Gene to ProductFrom Gene to Product* **163,** 155–165 (2013).

27. De Lorenzo, V., Krasnogor, N. & Schmidt, M. For the sake of the Bioeconomy: define what a Synthetic Biology Chassis is! *New Biotechnology* **60,** 44–51 (2021).

28. Sauser, B., Verma, D., Ramirez-Marquez, J., *et al. From TRL to SRL: The concept of systems readiness levels* in *Conference on Systems Engineering Research, Los Angeles, CA* (2006), 1–10.

29. De Lorenzo, V. & Couto, J. The important versus the exciting: reining contradictions in contemporary biotechnology. *Microb Biotechnol* **12,** 32–34 (2018).

30. Frank, C., Sink, C., Mynatt, L., *et al.* Surviving the "valley of death": A comparative analysis. *The Journal of Technology Transfer* **21,** 61–69 (1996).

31. Linton, J. D. & Xu, W. Understanding and Managing the Biotechnology Valley of Death. *Trends Biotechnol* **39,** 107–110 (2021).

32. Ding, M. & Eliashberg, J. Structuring the new product development pipeline. *Management Science* **48,** 343–363 (2002).

33. Schoonmaker, M., Carayannis, E. & Rau, P. The role of marketing activities in the fuzzy front end of innovation: a study of the biotech industry. *The Journal of Technology Transfer* **38,** 850–872 (2013).

34. Stremersch, S. & Van Dyck, W. Marketing of the life sciences: A new framework and research agenda for a nascent field. *Journal of Marketing* **73,** 4–30 (2009).

35. Fruehauf, H. M., Enzmann, F., Harnisch, F., *et al.* Microbial Electrosynthesis—An Inventory on Technology Readiness Level and Performance of Different Process Variants. *Biotechnology Journal* **15,** 2000066 (2020).

36. Sandberg, T. E., Pedersen, M., LaCroix, R. A., *et al.* Evolution of Escherichia coli to 42 °C and Subsequent Genetic Engineering Reveals Adaptive Mechanisms and Novel Mutations. *Molecular Biology and Evolution* **31,** 2647–2662 (2014).

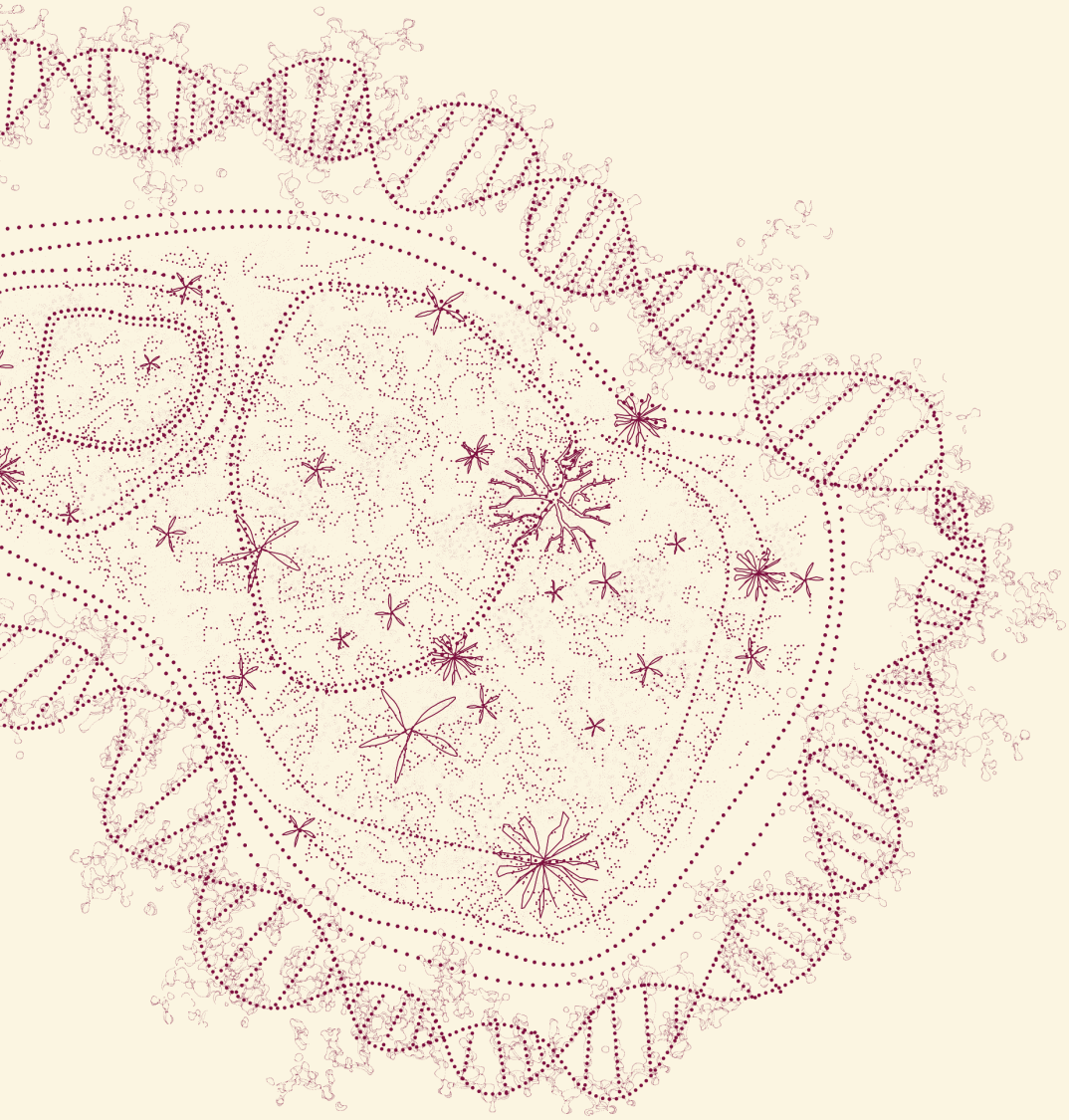
37. Baglieri, D., Baldi, F. & Tucci, C. L. University technology transfer office business models: One size does not fit all. *Technovation. DNA of the Triple Helix* **76-77,** 51–63 (2018).

38. Fröhling, M. & Hiete, M. *Sustainability and Life Cycle Assessment in Industrial Biotechnology: A Review of Current Approaches and Future Needs* (eds Fröhling, M. & Hiete, M.) 143–203 (2020).

39. Rosemann, A. & Molyneux-Hodgson, S. Industrial Biotechnology: To What Extent Is Responsible Innovation on the Agenda? *Trends in biotechnology* (2019).

40. Chen, X., Wei, P., Fan, L., *et al.* Generation of high-yield rapamycin-producing strains through protoplasts-related techniques. *Applied microbiology and biotechnology* **83,** 507–512 (2009).

41. Gong, T., Xu, X., Dang, Y., *et al.* An engineered Pseudomonas putida can simultaneously degrade organophosphates, pyrethroids and carbamates. *Science of The Total Environment* **628-629,** 1258–1265 (2018).
42. Franden, M. A., Jayakody, L. N., Li, W.-J., *et al.* Engineering Pseudomonas putida KT2440 for efficient ethylene glycol utilization. *Metabolic engineering* **48,** 197–207 (2018).
43. Kampers, L. F., Volkers, R. J. & Martins dos Santos, V. A. Pseudomonas putida KT 2440 is HV 1 certified, not GRAS. *Microbial biotechnology* **12,** 845–848 (2019).
44. Carbonell, P., Jervis, A. J., Robinson, C. J., *et al.* An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Communications Biology* **1,** 1–10 (2018).
45. Cooperrider, D. L., Whitney, D. K. & Stavros, J. M. *Appreciative inquiry handbook* (2003).
46. Braun, V. & Clarke, V. Using thematic analysis in psychology. *Qualitative research in psychology* **3,** 77–101 (2006).
47. Fenzl, T. & Mayring, P. QCAmap: Eine interaktive Webapplikation für Qualitative Inhaltsanalyse (2017).
48. Burnard, P. A method of analysing interview transcripts in qualitative research. *Nurse education today* **11,** 461–466 (1991).
49. Hallagan, J. B., Hall, R. L. & Drake, J. The GRAS provision - The FEMA GRAS program and the safety and regulation of flavors in the United States. *Food and Chemical Toxicology* **138,** 111236 (2020).
50. Demain, A. L. & Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. eng. *Biotechnology Advances* **27,** 297–306 (2009).
51. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in Microbiology* **5** (2014).
52. Etzkowitz, H. & Zhou, C. *The triple helix: University–industry–government innovation and entrepreneurship* (2017).
53. Brueller, N. N., Ellis, S., Segev, E., *et al.* Knowing when to acquire: The case of multinational technology firms. *International Business Review* **24,** 1–10 (2015).
54. Etzkowitz, H., Webster, A. & Healey, P. *Capitalizing knowledge: New intersections of industry and academia* (1998).
55. Calzolari, A., Valerio, A., Capone, F., *et al.* The European research infrastructures of the ESFRI Roadmap in biological and medical sciences: status and perspectives. *Annali dell'Istituto superiore di sanità* **50,** 178–185 (2014).
56. Renda, A. & Massimiliano, G. Innovation Policy in the European Union: towards Horizon2020 (2012).
57. Ellwood, P., Williams, C. & Egan, J. Crossing the valley of death: Five underlying innovation processes. *Technovation,* 102162 (2020).
58. Gatto, F. & Re, I. Circular Bioeconomy Business Models to Overcome the Valley of Death. A Systematic Statistical Analysis of Studies and Projects in Emerging Bio-Based Technologies and Trends Linked to the SME Instrument Support. *Sustainability* **13,** 1899 (2021).
59. Gilding, M., Brennecke, J., Bunton, V., *et al.* Network failure: Biotechnology firms, clusters and collaborations far from the world superclusters. *Research Policy* **49,** 103902 (2020).
60. Takata, M., Nakagawa, K., Yoshida, M., *et al.* Nurturing entrepreneurs: How do technology transfer professionals bridge the Valley of Death in Japan? *Technovation,* 102161 (2020).
61. Bauer, E. A. & Cohen, D. E. The changing roles of industry and academia. *Journal of Investigative Dermatology* **132,** 1033–1036 (2012).
62. Patermann, C. & Aguilar, A. The origins of the bioeconomy in the European Union. *New biotechnology* **40,** 20–24 (2018).



CHAPTER 11

A SNAPSHOT OF BIOMANUFACTURING AND THE NEED FOR ENABLING RESEARCH INFRASTRUCTURE

James D. Fawcett*, Enrique Asin-Garcia*,

Vitor A. P. Martins dos Santos

* Contributed equally

Abstract

In the search for sustainable alternatives to conventional petrochemical-based manufacturing, the industrial and commercial sectors are increasingly anticipating a transition towards a global economy grounded in biomanufacturing. The past two decades have seen the establishment of numerous diverse biotech companies, as well as some clear shifts in governmental policy that stress the importance placed on biomanufacturing for their future economies. To help bridge the gap between the conception and translation of these technologies, interdisciplinary research infrastructures have begun to address many of the outstanding issues that are currently limiting progress in bioprocess development and its competitiveness as compared to established petrochemical-based processes. We propose that these efforts could be taken even further through the adoption of an ecosystem of unified workflows that function within and between research infrastructures. These meta-workflows could act as a framework that emphasises the importance of interoperability, harmonisation and sustainability in making a successful transition to global manufacturing that utilises bio-based production platforms.

Biomanufacturing as cornerstone of the bio-based economy

The growth of the circular bioeconomy, or the commercial and industrial application of sustainable bio-based products and services, has long been an ambition for many governments and private stakeholders. Transitioning towards a system of production firmly rooted in biotechnological innovation, and away from traditional petroleum-based manufacturing practices, is dependent on the enabling research from fields such as synthetic biology (SynBio), information technology (IT) and industrial biotechnology (IB). Many pioneering technologies that facilitate bio-based production are being explored by start-ups, SMEs and larger companies alike, and span the entire range of biomanufacturing. There is also a growing support for these innovations from the public sector; public funding is finding its way into bio-based markets, and the desire to improve capabilities in biomanufacturing is becoming increasingly represented in national economic development strategies and roadmaps.

Despite the rising enthusiasm and investment in cutting-edge biomanufacturing processes, there are considerable barriers to the viability, commercialisation and acceptance of such technologies. Due to the swift and sudden development of biological engineering paradigms, as well as the use of potentially misleading analogies and simplifications when pitching the science, there is significant hype around the promise of many of these technologies in the context of industry. There is currently a divide between what is theoretically possible and what is both technically and economically feasible when pitching the adoption of many biomanufacturing processes into the industrial domain.

In this paper, we outline the current state of play of biomanufacturing in the private sector, and describe the efforts made to enable the development of new biomanufacturing practices through both publicly and privately funded research infrastructures (RIs). Finally, we propose the development of “meta-workflows" as a strategy to facilitate cooperation and interplay within the ecosystem of RIs in an effort to cultivate sustainability within the vision of a growing bioeconomy.

Setting the stage for private biomanufacturing

In terms of interest and involvement in state-of-the-art biomanufacturing principles and enterprises, the private sector needs little assistance. The past two decades have seen the establishment of numerous business ventures that feature SynBio and IB components that are core to their operation (Table 11.1). These are in fields long serviced by biotechnology, including agriculture, cosmetics and healthcare, but companies are increasingly branching out. More recent research avenues, including the use of DNA in data-storage and the expanding diversity of chemical and material production from biological sources, have enabled biomanufacturing to become relevant to many other areas of industry.

Healthcare makes up a large proportion of industry allied with biomanufacturing. In 2006, the biotechnology company Amyris was an integral player in the production of antimalarial drug artemisinin in yeast, which is now recognised as one of biggest success stories in synthetic biology [1]. Many SMEs now involve themselves with key proprietary technologies for the production of probiotics, small molecule and protein-based drugs, as well as cell-scale therapeutics like CAR-T cell therapy [2]. Other biomanufacturing enterprises focus on solutions to climate change and explore the production of alternative energies. Companies such as LanzaTech leverage engineered bacteria to produce biofuels from waste-feedstocks, while many including Photanol, Kiverdi and Algae for Future use them to actively sequester CO2 from the atmosphere [3].

In addition to explicit biomanufacturers, several other service providers have emerged that aim to facilitate the projects of the former. Companies such as Ginkgo Bioworks, the “Organism Company”, offer their infrastructure and expertise in metabolic engineering to design efficient bioprocesses for other enterprises, and have attracted many partners as a result. Ginkgo Bioworks achieves this in part through its biofoundry capability, in which bioprocesses can be designed and screened through iterative rounds of optimisation in a high-throughput manner [4]. Zymergen takes a different approach by placing emphasis on the design of novel products that both address unmet market needs and can be feasibly derived from biology. The company then designs a biomanufacturing process around the product by optimising selected microbes towards a platform for synthesis at scale.

Table 11.1: Biomanufacturing companies

Company	Headquarters	Established	Market & Specialties (if listed)	Self-description	Link
AgriMetis	Lutherville, Maryland, US	2014	Synthetic biology, Chemistry, and Agrichemicals	AgriMetis is a developer of biological products for crop protection. It offers fungicides, herbicides, insecticides, and nematicides which act against biotic stresses on crops including prevention against attacks caused by specific types of fungus, bacteria, and pests.	ttps://www.agrimetis.com/
Algenuity	Stewartby, UK	2009	Algal Strain Modification, Photobioreactors, Algae Biotechnology	Algenuity researches, develops, manufactures and provides labscale photobioreactors to optimize growth parameters and algal genome modification services to increase desired traits and to harness microalgal strains as SynBio chassis organisms.	ttps://www.algenuity.com/
Aromyx	Mountain View, California, US	2014	Biosensor, Scent and taste analytics, Artificial Intelligence	Aromyx uses the biological receptors from the human nose to create quantitative representations of sensory data for multiple applications.	ttp://www.aromyx.com
Biomason	Triangle, Virginia, US	2012		BioMASON employs natural microorganisms and chemical processes to manufacture biological cement-based masonry building materials.	ttps://biomason.com/
Biomilq	Durham, North Carolina, US	2019		BIOMILQ's goal is to mimic human breast milk in a lab setting by culturing mammary epithelial cells.	ttps://www.biomilq.com/
Bluegene Technologies	London, UK	2015	Textile manufacturing and dyeing industries, Cosmetics/food colourants, Fine chemicals	BlueGene's speciality is developing unique microbial strains and technologies for bioconversions of low-cost starting compounds to high-value chemicals and the production of novel chiral precursors for use in organic syntheses.	ttp://www.synbicite.com/collaboration/Partners/bluegene-technologies-ltd/
Bluepha	Beijing, China.	2016	Metabolic Engineering, PHA, Bioplastics	(Bluepha) uses SynBio techniques to make a biodegradable polymer called polyhydroxyalkanoate (PHA) out of seawater, and carbon containing industrial and domestic waste products. Its biodegradable polymer products are sold to companies operating in a variety of industries such as agriculture, packaging, industrial moulding, restoration, 3D printing, and textiles.	ttp://en.bluepha.com/
BOLT Threads	Emeryville, California, US	2009		Bolt is a materials company that invents and scales materials that put the planet on a path towards a better future. Biofabricated silk fibers Microsilk (is) based on proteins found in nature, produced using Bolt Threads proprietary technology and revolutionary bioengineering concepts.	ttps://boltthreads.com/
C16 Biosciences	New York, US	2017		C16 Biosciences is a biotechnology firm that uses microbiology to brew sustainable alternatives to palm oil.	ttps://www.c16bio.com/
CHAIN Biotechnology	Nottingham, UK	2014	Metabolic engineering, Chiral chemical synthesis, microbiome	CHAIN develops oral vaccines and immuno-therapies targeting the lower gastrointestinal tract with several therapeutic candidates in pre-clinical development.	ttps://chainbiotech.com/about/
Chr. Hansen	Hørsholm, Denmark	1874		(Chr Hansen) develop and produce cultures, enzymes, probiotics and natural colours for a rich variety of foods, confectionery, beverages, dietary supplements as well as even animal feed and plant protection.	ttps://www.chr-hansen.com/en

c-LEcta GmbH	Leipzig, Germany	2004	Tailored enzymes, bioprocess development, strain engineering	c-LEcta is a leading innovator and supplier of biotech products for regulated markets like food and pharma. Best-in-class technologies are applied to provide customized enzymes and microbial strains to novel and cost-efficient industrial manufacturing processes.	ttps://www.c-lecta.com/
Codexis	Redwood City, California, US	2002	Biocatalysis, Directed Evolution, Enzyme Optimization	Codexis is a biosciences company applying synthetic chemistry technologies to create improved versions of small molecule therapeutics.	ttps://www.codexis.com/
Colorifix	Norwich, UK	2016	Textile manufacturing and dyes, DNA Sequencing	Colorifix is developing a revolutionary dyeing process to help the textile industry dramatically reduce its environmental impact in a cost-effective manner using a synthetic biology-based approach.	ttps://colorifix.com/
Cysbio	Hørsholm, Denmark	2019	Biotechnology, Metabolic Engineering, and Synthetic Biology	Cysbio uses advanced metabolic engineering and synthetic biology approaches to construct bacterial cell factories for transforming the production of biochemicals from renewable feedstocks.	https://cysbio.com/
Deinove	Grabels, France	2006	Bioprocesses, antibiotics, antimicrobials	Deinove is a green technology company specialized in development and commercial exploitation of innovative, environmentally friendly, high-performance processes for production of biofuels and other compounds of industrial or pharmaceutical value.	ttps://www.deinove.com/en
Demuris	Newcastle upon Tyne, UK	2007		Demuris is focused on the discovery of new antibiotics to combat the threat from emerging diseases and the rise of antibiotic resistance. Demuris also provides the opportunity to discover and develop small molecule inhibitors against a range of interesting drug targets.	ttps://demuris.co.uk/
Desktop Genetics	London, UK	2012	Functional genomics, Cell Line Engineering, Drug Discovery, Machine Learning	(Desktop Genetics) have supported thousands of labs around the world with best-in-class CRISPR genome editing design and analysis technology, across basic research, drug discovery, translational research and IND-enabling studies.	ttp://www.desktopgenetics.com/
Eligo Bioscience	Paris, France	2014		Eligo Bioscience is a biotechnology company that develops a new class of biotherapeutics for microbiome precision-editing.	ttps://eligo.bio/
FUJIFILM Diosynth Biotechnologies	Billingham, UK	2011	Biologics, Process Development, Biologics cGMP Manufacturing	(FUJIFILM Diosynth) combine(s) technical leadership in cell culture, microbial fermentation and gene therapies with world class cGMP manufacturing facilities to advance tomorrow's medicines	ttps://fujifilmdiosynth.com/
Geltor	San Leandro, California, US	2015	Analytical Chemistry, Protein Design, Biological Manufacturing	Geltor is a bio-design company that makes proteins used in beauty, food, & beverage products like collagen and elastin, but without animals.	ttps://geltor.com/
Ginkgo Bioworks	Boston, Massachusettes, US	2008	Cultured ingredients, biological engineering, metabolic engineering	(Ginkgo Biowork's) cell programming platform is enabling the growth of biotechnology across diverse markets, from food to fragrance to pharmaceuticals. Ginkgo is also actively supporting a number of COVID-19 response efforts, including community testing, epidemiological tracing, vaccine development and therapeutics discovery.	ttps://www.ginkgobioworks.com/

GOURMEY	Paris, France	2019		GOURMEY is France's first cultivated meat company. Its mission is to accelerate the world's transition toward humane, sustainable and healthy meat by harvesting it from animal cells.	https://gourmey.com/en/
Green Biologics	Abingdon, UK	2003	Fermentation chemistry, Renewable Chemicals, Biofuels	(Green Biologic's) bio acetone, n-butanol and IPA provide a platform for a variety of bio-based solutions for derivatives and consumer product development.	ttp://www.greenbiologics.com
Ingenza	Roslin, UK	2002	Enzyme Discovery, Enzyme Evolution, Synthetic Chemistry	Ingenza is a biotechnology company specialising in the design, development and manufacture of diverse, high-value industrial products and therapeutic-proteins. (Ingenza) exploit synergies between synthetic and evolutionary biology, providing next generation microbial and mammalian manufacturing platforms	ttp://www.ingenza.com
LanzaTech	Skokie, Illinois, US	2005	Green chemistry, carbon recycling, fuels and chemicals	LanzaTech's unique microbial gas fermentation process provides a sustainable pathway to produce platform chemicals that serve as building blocks to products that have become indispensable in our lives such as rubber, plastics, synthetic fibers and fuels.	ttps://www.lanzatech.com/
LifeMine Therapeutics	Cambridge, Massachusetts, US	2017		Discovering powerful new drugs from eukaryotic microbes by combining Genomics with AI and SynBio.	ttps://lifeminetx.com/
Lonza Pharma & Biotech	Basel, Switzerland	1897		(Lonza) have extensive experience with microbial processes using advanced engineering and process development capabilities.	ttps://pharma.lonza.com/offerings/microbial
Meatable	Delft, The Netherlands	2018		Meatable is an innovative, Dutch food company, aiming to deliver, at scale, the new natural, cultivated meat that looks like, tastes like, and has the nutritional profile of traditional meat. Its proprietary opti-ox technology enables Meatable to produce meat rapidly, sustainably, and without harming animals.	ttps://www.meatable.com/
MicroGen Biotech	Carlow, Ireland	2012	Biotechnology and bioremediation	(MicroGen) utilise patented isolation and high-throughput screening methods to isolate functional, high-performance microbiomes for application in agricultural crop production and environmental remediation.	ttp://www.microgenbiotech.com/
Micro SynBiotiX	Cork, Ireland	2016	Oral vaccine production, research and development, and biotechnology	MicroSynBiotiX is developing cost-effective oral vaccines to combat these infections and improve global food security. We are developing a novel, patent-pending method of producing oral vaccines using transgenic microalgae	ttp://www.microsynbiotix.com/
Nextbiotics	San Francisco, California, US	2017	Biotech, CRISPR, and Digital Biology	(Nextbiotics) are using SynBio to drastically enhance bacteriophage stability and efficacy, allowing us to efficiently destroy pathogens, enhance animal nutrition and promote animal health in an antibiotic-free way.	ttps://www.next-biotics.com/
Oxford Biotrans	Oxford, UK	2013		(Oxford Biotrans) will develop and commercialise enzymatic process technologies that yield high-value chemical compounds.	ttps://oxfordbiotrans.com/

Oxitech	Abingdon, UK	2002	Vector Control, agriculture, sustainability	Oxitec has developed the world's leading insect-based biological control system to safely and sustainably control insects that transmit disease and destroy crops.	ttps://www.oxitec.com/
Pembient	San Francisco, California, US	2015		Pembient is leveraging advances in biotechnology to fabricate wildlife products, such as rhino horn and elephant ivory, at prices below the levels that induce poaching.	ttps://www.pembient.com/
Perfect Day	Emeryville, California, US	2014	Dairy, Food/Beverages, and Biotechnology	(Perfect Day) make milk protein that is nutritionally identical to protein from cow's milk, but made using flora fermentation instead of animals.	ttps://perfectdayfoods.com/
Photanol	Amsterdam, The Netherlands	2012		Photanol develops a breakthrough technology to convert CO_2 and sunlight into valuable organic compounds. We are applying this technology in a number of markets, ranging from food ingredients to chemical bio-blocks and biofuels.	ttps://photanol.com/
PILI Bio	Toulouse, France	2015	Bioproduction, Biomaterials, Dyes, Textile, and Inks	PILI is focused on the biofabrication of a wide range of colors produced by microorganisms as an alternative to the petrochemical versions.	ttps://www.pili.bio/
Pivot Bio	Berkley, California, US	2011		Pivot Bio is improving and using a deep understanding of the microbiome to advance crop nutrition. It currently focuses on enabling microbes to fix and supply more nitrogen to corn.	ttps://www.pivotbio.com/
Prokarium	London, UK	2012	Vaccines, Oncology, Microbial Immunotherapy	Prokarium's mission is to use its expertise in engineering biology and immunology to improve the lives of patients with solid tumours, with its lead oncology program expected to enter the clinic in 2022.	ttps://www.prokarium.com/
Puraffinity	London, UK	2015	Water Treatment, Micropollutants, Adsorbents, Sustainability	(Puraffinity) specialises in the design of advanced materials, and has developed a bio-based novel adsorbent capable of removing highly challenging pollutants from contaminated water and wastewater.	ttps://www.puraffinity.com/
Swiftscale Biologics	San Francisco, California, US	2019		Swiftscale provides cell-free protein synthesis for fast expression, high-throughput synthesis, and complex protein production.	ttps://www.swiftscalebio.com/
Synthace	London, UK	2011	Synthetic biology, engineering biology, and programming	Synthace's goal is to raise universal bioscience productivity, enabling people to better engineer biology for health, food, energy and manufacturing. Central to its technology is Antha, an operating system that enables ease of designing and optimizing biological unit operations that are linked into executable workflows that are reliable, shareable and scalable.	ttps://www.synthace.com/
Zymergen	Emeryville, California, US	2013		Zymergen has specialized in engineering microbes for new materials, including electronics and insect repellent, or for manufacturing chemicals and fuels.	ttps://www.zymergen.com/
Zymvol Biomodeling	Barcelona, Spain	2017	Protein engineering and Molecular modeling	Zymvol Biomodeling SL is a privately held biotech company dedicated to amplifying biocatalytic solutions in the market through cloud-based computer simulations.	ttps://zymvol.com/

In addition to those providing physical biodesign infrastructure, software developers such as Synthace and Riffyn provide tools that enable computer-aided biology, bioprocess design and consistent data analysis, all of which further support those directly involved in the manufacturing sector. As contemporary biomanufacturing practices continue to mature, the most recent developments in these technologies are being made progressively more comprehensible and approachable for potential investors. Networks such as SynBioBeta (https://synbiobeta.com/) document the minutia of the funding, acquisition, deals and successes of businesses that employ SynBio and IB in their platform, and provide a much-needed context and accessibility for non-experts in the increasingly market-facing discipline.

The combination of wider public exposure alongside deliberate attempts to push SynBio and IB innovation in the direction of meaningful industrial translation has led to a considerable amount of investment in the integration of novel biotechnologies within manufacturing and production domains. This steady trend potentially reflects changes in attitudes and a growing confidence from investors regarding the commercial possibility of biodesign and biomanufacturing technologies, especially in a sector known to entail a considerable amount of risk. The novel and untested nature of many of the technologies underpinning a biomanufacturer's business strategy often requires vast amounts of seed money, as well as time and specialist knowledge for any significant payoff, and many promising ideas are likely to eventually peter out with little or no profit. The prevailing sentiment from private investors, however, appears to be that the substantial potential rewards make the expenses worth the risk.

Living up to the hype

In the 2010 publication "The Five Hard Truths about Synthetic Biology", biotechnologist Martin Fussenegger opined that the field "has had its hype phase", and that it "now needs to deliver" [5]. Originally attributed to SynBio research, it has become apparent that elements of this hype have been transferred into the biomanufacturing investment sector, where seemingly endless promise surrounds several enabling technologies in cases where their true effectiveness still remains to be seen.

While many commercial initiatives will likely succeed in advancing the state of the biomanufacturing field, it is probable that many will fail in some capacity due to deficient technologies.

The aforementioned publication also highlighted a flawed narrative used when describing SynBio technologies, in that they are often represented in an abstract form analogous to "LEGO blocks". The author reiterated that "hacking biology" would, in reality, involve much more complexity and variability than any engineering analogies would suggest. These same misrepresentations about the readiness of new SynBio and IB techniques may have similarly been pitched to investors with a limited understanding of the underlying technological reality; this form of upselling, whether intended or not, could lead to financial backlash when biotechnology companies find their products cannot deliver on their promise. Nevertheless, multiple innovations thought to be hype in 2010 have since been translated into robust and applicable technologies, and fulfil the requirements of the companies that adopt them [6].

Although many diverse and well-funded companies can be found in each subsection of biomanufacturing, it is difficult to quantify the number of successful commercial ventures that have significantly contributed to the bioeconomy thus far. As expected, the prosperity, weaknesses and profits of these private entities are obscured from all those except the members of the board and others who are in-the-know. What is evident is that most externally funded start-ups and SMEs take a long time to provide any meaningful return on investment, however substantial that return may eventually be. It is, in most cases, too early to make any accurate estimates on the success or failure of many burgeoning technologies aimed at facilitating modern biomanufacturing and the bioeconomy at large; this is a question that should be re-visited periodically over the span of several years. At this stage, one should measure the health of the modern biomanufacturing sector by its high, and continually growing, levels of investment, as well as through the developmental support offered to academic partners from optimistic commercial and industrial stakeholders.

The pitfalls of biomanufacturing

Despite the promising activity in the private sector concerning biomanufacturing, several key issues in bioprocess design make commercialisation of these technologies difficult. The landmark creation of artemisinin-producing yeast took an estimated 150 person-years of work to achieve and required significant subsidies [5]. Long bioprocess development times have been echoed in the production of other small chemicals through metabolic engineering by pharmaceutical giants such as DuPont [7]. These upfront expenditures are enough to give any investor pause, especially when considering financing the same activities being attempted by burgeoning, under-resourced SMEs.

Although rapid improvements in biomanufacturing facilitation are ongoing, many areas of the industrial sector are currently better served by existing and well-established petroleum-based manufacturing methods. This is especially evident in the stalled transition from fossil fuels to biofuels; with only a few exceptions, production of the latter to the required scales is still the less economically favourable option when compared against the price and availability of conventional fuels. Similarly, while several interesting candidates exist, the promise of bioplastics that are both cheap and readily producible is still some time away. Whilst the healthcare and agricultural sectors are already seeing the fruits of bio-based production methods, the more classical manufacturing fields are not presently able to be properly integrated into any contemporary version of the bioeconomy.

> "Biotechnology takes too long, costs too much and fails too often to effectively address the global challenges that must be solved", *James Philp, workshop OECD/ICL, Sept 1, 2018*

Because of the often unpredictable and irreproducible nature of new biomanufacturing technologies, most designed bioprocesses require considerable testing and optimisation, which is necessarily time and resource consuming. The economy of scale for manufactured products is vital for businesses to recuperate their investments in the costly initial design of these processes, and so economically viable biomanufacturing needs to operate over large production scales. In reality, most SME and start-ups tasked with bringing these processes to market lack the ac-

cess to pilot plants where they can test at the volumes required. It is prohibitively expensive for these smaller companies to justify the construction of a bioprocessing plant to test their products at production-level scales, which is a well-known barrier-to-entry for smaller biomanufacturers [8]. For this reason, many promising biomanufacturing innovations are inevitably lost to the "Valley of Death", a scenario in which a project is abandoned due to the owner's lack of resources and inability to test their process at realistic scales and optimise the technology accordingly [9]. To navigate this issue, investment in accessible research infrastructure that allows for economical testing at scale is increasingly warranted.

> "The big technology barrier (to commercialisation) is how you move from a laboratory-based strain that produces a product towards a 3000 L bioreactor producing the same product at high titre. That's always been the challenge in industrial biotechnology, and I think it still is the challenge.", *Paul Freemont, Global Alliance Biofoundries, Imperial College London*

The need for research infrastructure

In addition to private investment, many governments are acknowledging the need to invest in biomanufacturing. Clarke and Kitney describe several economic drivers for the development of the bioeconomy, which include the desire to "actively mitigate climate change, to improve food security" and to "establish more sustainable manufacturing supply chains" [10]. The incentive for a national capability to deal with these issues is apparent, as is the want to foster a competitive industrial base; manufacturing is a major component in most national economies and lagging behind their neighbours can have serious ramifications. These aspects are often surmised through roadmaps, which outline intentions of publicly funded research focus in areas such as sustainable agriculture, renewable fuels and carbon-neutral waste processing [11–15].

Kitney *et al.* also suggest that biomanufacturing capabilities can be further supported through more public-private partnerships, as well as policy initiatives that increase the confidence of the private sector in the forthcoming transition to

bio-based manufacturing practices [16]. Presently, however, the majority of public resources support the development of the biomanufacturing sector through the funding and establishment of research infrastructures (RIs).

In the context of biomanufacturing, RIs form an architecture that provides access to dedicated research facilities, such as biorefineries and biofoundries, and resources for research and innovation to those within their network [17]. They commonly have an overarching research theme and often function at the interface of academia and industry, with the explicit goal of translating innovation from publicly funded institutions into realised commercial applications. A major strength of RIs is in their collaborative spirit; data generated by members of the group can be shared openly and contributors are able to dedicate their focus towards individual and specialist aspects of the research goal. It is therefore necessary for RIs to be cross-disciplinary, which allows coordination and integration of wet-lab science, computational design and analysis, and the practical considerations necessary for realistic industrial translation and commercialisation. The key component of an RI is its ability to manage, consolidate and distribute data between its members and collaborators, permitting research to be achieved on striking scales.

> "In other words, there is no one technology at fault, but it is the failure to integrate several technologies that defeats commercialization.", *Kitney et al., 2019 [16]*

At present, there are multiple RIs and similar consortia that focus on different specific aspects to enable biomanufacturing and the bioeconomy at large (Table 11.2). Launched in 2018, the Global Biofoundry Alliance (GBA) is a network of public biofoundries and associated biofoundry facilitators that aims to both mature their platform and accelerate the commercial translation of SynBio research [18]. With 29 international participants, a main tenant of the organisation is the open sharing of data and materials between GBA members for the benefit of improving biofoundry capability as a whole. With the aims to increase research democratisation, a future focus of the GBA will be in promoting the mobilisation of biofoundries, with plans to establish ties with facilities in the African and South American continents [19]. With access to high-throughput biological testing facilities, countries

with limited research capabilities and resources will be able to develop healthcare and agricultural bioprocesses that best suit their individual needs and lessen their dependence on the research activities from other areas of the world.

> "I would love to see the mobilisation concept of the biofoundries developed further. This pandemic has illustrated that there are roles for biofoundries in acting quickly… vaccine development and optimisation, antigen and synthetic virus-like particle production… there are many things that can be done in the vaccine arena, as well as in low-cost diagnostics", *Paul Freemont, Global Alliance Biofoundries, Imperial College London*

The Future Biomanufacturing Research Hub (FBRH) is leading the biomanufacturing strategy of the UK, with specific priority placed in the production of pharmaceuticals, value-added chemicals, engineering materials and advanced biofuels. The FBRH is made up of several major UK universities, as well as publicly funded innovation centres specialising in industrial scale up and commercialisation. Although many UK institutions are involved with other biomanufacturing RIs, including the GBA, the FBRH focuses on the low research levels of TLR 1-3, and aims to be unique in its service to the UK's vision of the bioeconomy. Members of the EU are serviced by multiple biomanufacturing RIs, comprising of both national and union-wide initiatives. The Fraunhofer Society, Germany's applied research organisation, is well placed to pivot their considerable manufacturing infrastructure and research capability towards bio-based production lines [20]. The German government also has plans to construct demonstration plants that promote the scale-up of bioprocesses to industrial capacities [21].

The Industrial Biotechnology Innovation and Synthetic Biology Accelerator (IBISBA, ibisba.eu), a French led, pan-European initiative, involves close cooperation from research infrastructure operators in multiple EU member states. Behind the IBISBA concept is the ambition to link research infrastructure within a coordinated network capable of producing services that support biotechnology project pipelines (TRL 2-5), including biofoundry and scale-up operations. The IBISBA infrastructure network is also developing the requisite business practices to smooth

the hand-on/hand-off transitions between individual RI services, thus streamlining overall support to projects. Primarily, the initiative aims to make research infrastructure in biotechnology more readily available, and to actively support the translation of cutting-edge knowledge in the field of bioengineering into innovation and commercially mature products and services. As a pan-European initiative, IBISBA is well placed to grow European biomanufacturing capability, in part to keep pace with the developments made in China and the United States.

The Bio-based Industries Joint Undertaking (BBI JU) is another public-private European group which has seen the establishment of biorefineries across the continent, in an advancement towards the European Green Deal climate initiative [22]. The ambitious programming and funding agency aims to invest €3.7 billion in bio-based industries by 2024, with a focus on de-risking investment and creating new and internationally competitive value chains.

In the US, the National Institute of Standards and Technology (NIST) program primarily works on developing suitable metrologies for use in biomanufacturing, and also offers the use of their advanced imaging facilities to their academic and industrial partners. Through the distributed Agile Biofoundry platform, the US Department of Energy has financed numerous projects across the country aimed at enabling biomanufacturing, encompassing both specific bioprocesses such as antimicrobial peptide, fatty acid and biodegradable plastic synthesis to more process design-oriented optimisations in host cell engineering [23]. The Department of Defence has allocated US$87.5 million to the newly established Bioindustrial Manufacturing And Design Ecosystem (BioMADE, biomade.org), bringing total investment in the RI up to over US$260 million [24]. Over 90 public and private entities contribute to BioMADE's goal of translating lab-scale innovations towards industrial capacities, which emerged from the broader national Engineering Biology Research Consortium (EBRC, ebrc.org) to focus specifically on deployment of biomanufacturing technologies.

Table 11.2: Biomanufacturing research infrastructures

Facility	Headquarters	Establishe	Key Focus Areas
A*STAR	Singapore	1991	Prokaryotic / Eukaryotic therapeutic production Product recovery and analysis Omics technologies

ACIB (Austrian Center of Industrial Biotechnology)	Austria	2010	Bionanoparticle production Microbial CO2 Sequestration Green bioprocess technology development
AFOB (Asian Federation of Biotechnology)	Korea	2008	Agricultural biotechnology Bioprocess and Biorefinery Environmental technology
Agile Biofoundry	US	2016	DBTL process optimisation Host onboarding Scale-up and downstream process development
AlgaePARC (Wageningen UR)	The Nether-lands	2011	Microalgae strain development Sustainable microalgae production platforms Microalgae feedstocks and biorefinery
Algae-UK	UK	2019	Biomass production and processing Strain domestication and platform development Exploration of algae metabolic diversity
BBI JU (Bio-based Industries Joint Under-taking)	EU (distributed)	2014	Biomass feedstock production Biorefinery process scale-up and optimisation Bio-based product marketing and policy
BioCRF (Biosciences Central Research Fa-cility)	Hong Kong	2010	Liquid chromatography and mass spectrometry Tissue processing and metabolomics Sequencing services
BioMADE (Bioindustrial Manufacturing and Design Ecosystem)	US	2020	Scale-up and downstream processing R&D commercial translation Sustainable and reliable biomanufacturing technology
Bioplatforms Australia	Australia	2007	Genomics, proteomics and metabolomics platform
BIRAC (Biotechnology Industry Research Assistance Council)	India	2012	Biotech innovation incubators Research to industry translation Public-private partnership formation
BPF (Bioprocess Pilot Facility)	The Nether-lands	2012	Biomass valorisation Fermentation scale-up and processing Pilot-scale biorefinery demos
CPI (Center for Process Innovation)	UK	2003	Data-driven rational bioprocess design Agritech ingredient production platforms Pharmaceutical/therapeutic biomanufacturing
CSIRO BioFoundry	Australia	2021	Application domains: Foundation Technologies, Envi-ronment & Biocontrol, Chemicals & Fibres, Health & Medicine and Maximising Impact
DIPLOMICS (Distributed Platform in Omics)	South Africa	2017	Genomics/Proteomics infrastructure Downstream data bioinformatic analysis LIMS development
DTU Biosustain	Denmark	2011	Petrochemical to biological production transition Microbial genomics and strain design Biosustainability
Euro-BioImaging ERIC	Finland	2019	Multi-modal molecular and biomedical imaging Instrument training and data management
FBRH (Future Biomanufacturing Research Hub)	UK	2019	Sustainable bio-based manufacturing in pharmaceu-ticals, value-added chemicals, engineering materials and advanced synthetic fuels
GIFS (Global Institute for Food Security)	Canada	2012	Resilient Agriculture Root-Soil-Microbiome Interactions Plant Improvement & International Partnerships

Global Biofoundries Alliance (GBA)	Multinational	2018	Enhance international biofoundry collaboration DBTL cycle technology optimisation Accessible pre-competitive infrastructure
iBioIC (Industrial Biotechnology Innovation Center)	Scotland	2014	Bioprocessing Scale-Up Centres Talent development and funding provision
IBISBA (Industrial Biotechnology Innovation and Synthetic Biology Accelerator)	EU (distributed)	2018	Promote circular bioeconomy transition End-to-end bioprocess development Biocatalysts and microbial cell factories
ISBE (Infrastructure Systems Biology Europe)	EU (distributed)	2009	Data, tool and protocol standardisation Data modelling Systems biology workgroup integration
MIRRI (Microbial Research Resource Infrastructure)	EU (distributed)	2012	Valorisation of antimicrobials and bioactives Microbe-based bioremediation Biological management of soil and crops
NRC-HRT (Human Health Therapeutics Research Centre)	Canada	2019	Animal and microbial cell pilot plants Biologics and vaccine production platform Monoclonal antibody production
NUS SynCTI (Synthetic Biology Foundry)	Singapore	2015	Synthetic genomics Enzyme and biocatalyst engineering Microbiome and microbial ecosystems
SynBioChem	UK	2015	Microbial engineering for specialty chemicals DBTL pipeline prototyping High-throughput laboratory robotics

The additional benefits of RIs

Aside from enabling the development of biomanufacturing technologies, the integration of hardware, software and expertise from both public and private research entities into a focused RI has numerous other advantages within biomanufacturing and beyond. As a public effort with both societal and economic involvement, well-established RIs can be deployed in multiple ways to bolster the needs of their respective governments. This became apparent during the COVID pandemic, during which several biofoundries refocused their efforts towards high-throughput research, testing and vaccine production [25]. Their flexibility in use highlights the benefits of public investment in high-throughput research infrastructure; once built, RI capability can be refactored and refit to perform various automated functions as needed. In many areas, RIs could be set up to serve immediate public needs as responsive centres for disease testing and therapeutic development, rather than as a permanent platform for efficient bioprocess design.

Much importance has been placed on the biofoundry for the enabling of new biomanufacturing techniques [26, 27]. However, biofoundries could be simply con-

sidered bioscience tools powered by digital technology, whereas building a scaled bioprocess also requires a good understanding of the chemical engineering involved in this endeavour. Most current biofoundries, for example, are ill-equipped to handle process scale-up and therefore benefit from collaboration with pilot plants and biorefineries. In addition to physical centres for experimentally driven research, data processing, machine learning and AI tools are vital not only for designing robust bioprocesses, but also ones that are economically feasible, sustainable and attractive to investors. Taken alone, biofoundries struggle to create fully validated end-to-end processes that are suitable and ready for industrial adoption; much of the interdisciplinary work required for effective translation of biomanufacturing can only be provided in the context of an integrated and multifaceted RI that includes members from each of the relevant disciplines.

By prioritising data sharing, albeit to differing degrees of openness to the public, RIs contribute to the global improvement of generalised biomanufacturing processes. As pre-competitive enabling biotechnologies mature, we can expect to see individual stakeholders become more private and protective over their unique IP; we have seen as much already in the private biofoundry community, which enables a select few biomanufacturers to flourish but does little to advance the overall platform. It is difficult to imagine the establishment of truly accessible infrastructure and open research frameworks without funding and oversight from the public sector, which will likely be required for the enrichment of the whole. Additionally, as each member within an RI will have interests and activities independent of the goals of the collective, their required proximity and working relationships with collaborators from other disciplines provides ample opportunities for innovative spin offs and separate lines of research into new directions.

Areas for growth

As integrative groups, RIs have many components; in addition to wet-lab facilities operating at both low and high scales, they also need to coordinate workflows and protocols, computational and data management tools, and commercial interests and requisites. Harmonisation of both working practices and objectives within the

RI is therefore of a high priority, and one which becomes increasingly more difficult as partners separated both geographically and by discipline are required to cooperate when designing biomanufacturing processes. Biofoundry members already find difficulty in harmonising workflows and protocols, as well as agreeing on which standards, metrology and reference materials should be used. These issues are likely to be exacerbated within RIs, wherein stakeholders from divergent fields are required to find common ground regarding an appropriate procedure or method of communication. Conflicting viewpoints also become apparent when formulating the primary goal and objectives of an RI, aspects which are often left general and wide in scope. Despite this, the larger objective focused RIs inherently lack flexibility, and may struggle to branch out in promising research directions that fall outside of their mandate.

While standardised protocols and metrology can be established by RIs from the top down, this would require many collaborators to completely refactor their working practices, leading to a less productive and cohesive working group. To address this, a major area of focus in facilitating effective RIs is data interoperability. Common denominators to define the hand-on/hand-off points and translation of the data output from disparate members into a format that can be easily understood by all disciplines and stakeholders within an RI enable each player to govern their own practices, whilst maintaining efficient collaboration. Misrepresentation and misunderstanding of results or large data deposits that lack a detailed context could lead to inaccurate key assumptions, which may see significant effects not detected until later in the project timeline.

RIs can also be limited by a lack of accessibility. Researchers in the SynBio and IB fields may already be well-aware of the services that RIs offer, as will closely allied companies. However, it would be advantageous if these services were better described and advertised to people outside of these close communities. Theoretically, any individual working with a compound of unexplored interest, or looking to optimise the manufacture of both simple and complex products, should be able to consult an RI regarding the potential for a biomanufacturing solution which they may not have otherwise explored on their own. Increased interaction with members of different research communities could result in multiple opportunities for

collaborative innovation, either for individual RI members or for the group as a whole.

Although RIs can encompass all that is required for biomanufacturing process design, individuals from outside the group should be able to engage with relevant group members to differing degrees; for instance, many labs and SMEs already possess a limited capacity for small-scale in-house automation and may only need to outsource certain aspects to conduct testing and development at higher scales. Other labs may only require access to the computational tools or data management systems generated within the RI, and not make use of any physical infrastructure. User-friendliness should be prioritised to increase client engagement with RI activities and services, which has the benefit of adding additional venue streams that the RI can reinvest into their specific focus enabling biomanufacturing technologies.

The more sophisticated an infrastructure gets, the more specialized your staff need to be o support that development. And that makes it difficult to be open.

Developing a Meta-workflow

RIs work to address several key areas of bioeconomic growth by bringing together collaborators from different backgrounds. However, as goal-oriented organisations, their focus is not explicitly on how research can be performed and effectively translated within the same space. The initial transition into bio-based manufacturing, on both a national and global level, is only part of the challenge; the links between research and positive commercial outcome have to be sustained and continually updated as the enabling technologies mature. Refactoring and setting up new RIs *ad hoc* as their specific objectives evolve would be a waste of resources. At this stage, it makes sense to step back and consider how the substantial brainpower available to an RI could be better leveraged to support and encourage the evolution of a project and its core technologies over a long period of time.

To this end, we suggest a focus on developing a “meta-workflow” for how the research performed within RIs can be made more streamlined, accessible and resilient to rapidly evolving disciplines and technological developments, with an emphasis being placed on how the specific research objectives of an RI can fit into the

bigger picture of an expanding bioeconomy (Figure 11.1). Developing meta-workflow practices should prioritise data interoperability; although we are capable of generating huge amounts of sequencing and modelling data, this often lacks the proper context and can be easily misrepresented by focusing on certain aspects. In concert, promoting harmonisation of the procedures performed by RI members should help produce data that is simultaneously more interoperable and more repeatable. Excessive standardisation of protocols and biological components could stifle the possibilities that their use may uniquely contribute to a specific objective, and so some diversity in this regard should be encouraged. Attention can, however, be dedicated to how the resulting data is outputted, and in ensuring that it is relevant alongside and compatible with other data forms used throughout an RI. In practice, this could involve requiring defined common standards only at points where data is exchanged between different members.

Another argument for a meta-workflow is to facilitate effective communication between distinct RIs that share similar areas of application. Some problems may be best addressed through multiple or mutually exclusive strategies, and the potential for lucrative commercialisation can foster a competitive nature between contrasting approaches and viewpoints. Within the framework of a meta-workflow, RIs could further enable their own formation, interplay and competition, and together support the testing of comparable, yet unique approaches in many areas of bioeconomic interest. Comprehensive, seamless and interdisciplinary work practices and communication styles should be inherent to this ecosystem, which when adopted could aid in defining the most efficient and cost-effective biomanufacturing practices across a number of different industries.

> "An infrastructure is most powerful when it allows to derisk investment of taking multiple ideas to products and thereby enable generating value.", *Markus Herrgård, BioInnovation Institute, Copenhagen*

Looking forward, an ecosystem of RIs governed by an efficient meta-workflow could encourage contributors to distil how their work can best contribute towards the overall vision of a universal bioeconomy. In addition to creating bio-based manufacturing platforms, we also need to conceive of ways to support their later growth

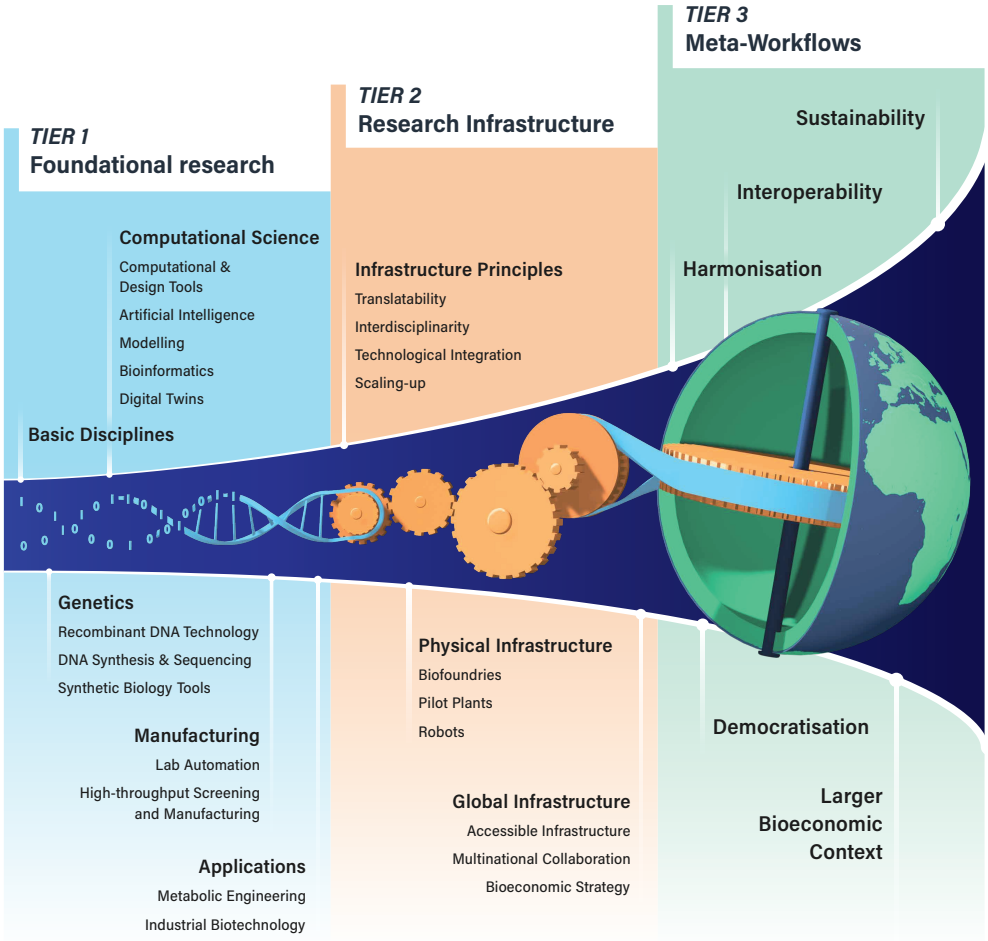


Figure 11.1: Tiers and areas of focus in enabling the bioeconomy.

and maintenance that are both sensible and sustainable. This will require both continued exploratory research, as well as the consolidation and optimisation of the foundational enabling technologies that we already possess. The larger bioeconomy will forever be inextricably linked to the nascent biotechnologies developed within the research sphere, and so both will need to be evolved reciprocally without the development of one significantly outpacing the other. Bioeconomic growth will need to advance organically and satisfy specific concerns and objectives within manufacturing if the transition from petrochemical to bio-based production is to be economically favourable and sustainable indefinitely. An ecosystem of RIs gov-

erned by a meta-workflow that encompasses all aspects of this growth, with attention paid to how the sector may look in the near and far future, could help ensure that the research output from RIs has a greater potential for translation and gives the bioeconomy its greatest chance of universal success.

> "... biofoundries [that] will make it easier for the process of designing new synthetic lifeforms to be scaled up from the bespoke boutique business it is now to something more like a global industry. That day is not yet here. But if there is demand, then biofoundries will surely play their part in the next phase of the Industrial Revolution.", *The Economist, March 1, 2018*

A vision for meta-workflows in the bioeconomy

Investment in the facilitation of biomanufacturing is a priority for those in the public sector who desire rapid translation, and has the added benefits of providing a robust and flexible platform for the production of other materials, including vital healthcare and agricultural products. Due to the economy of scale, it is difficult for economical bioprocesses to be developed effectively; much money is wasted on establishing piloting plants and in the outsourcing of expertise required to optimise the bioprocess. As biofoundries have proliferated to enable bioprocess design, several RIs have been set up to support other areas within biomanufacturing facilitation. Distributed, and importantly, accessible RIs facilitate the translation of ideas rooted in biological research into being commercially viable by lowering the barrier-for-entry for SMEs into high-throughput and high-scale testing and optimisation.

Whilst few private ventures of note can be pointed to as a prime example of a bioeconomic success as of yet, progress can be measured by the steps taken by RIs and private investors to enable the biomanufacturing process itself. RIs result in high cooperation in the development of these processes; not only will this likely lead to more translational success stories, but improvement of the foundations that burgeoning biomanufacturers can depend upon for growth. We now propose the idea of meta-workflows for RIs, which would act as a framework to further expedite

the translation of technologies and would help define their applicability within the larger bioeconomic context. An ecosystem of connected RIs that places emphasis on data interpretability, harmonisation, democratisation and sustainability could play a crucial role in enabling the development of the next-generation biomanufacturing principles necessary for a global transition towards a sustainable bioeconomy.

Acknowledgements

We would like to thank SynbiCITE co-director Professor Paul Freemont for our very useful discussions and for sharing his insights concerning issues in contemporary biomanufacturing. We wish to convey our gratitude as well to Dr. Michael O'Donohue (INRAE, www.ibisba.eu) and Dr. Markus Herrgård (CTO BioInnovation Institute, bii.dk) for their valuable comments in the manuscript and discussions on infrastructure. We thank Dr. James Philp (OECD) as well for fruitful discussions. We are also very thankful to a large number of www.IBISBA.eu team members for many discussions over time on aspects related to the reflections depicted here (in particular Prof. Merja Penttilä, VTT, Finland, and Dr. Heleen de Wever, VITO, Belgium). Additionally, we gratefully acknowledge the financial support from the research program SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis which is financed by the Dutch Research Council (NWO) (project number 15814), and the European Horizon 2020 project IBISBA (grant agreements No 730976 and 871118).

Conflict of interest

The authors declare there are no conflicting interests.

Authors' contributions

Conceptualization: JDF/EAG/VAPMdS; Funding acquisition: VAPMdS; Investigation: JDF/EAG; Methodology: JDF; Supervision: EAG/VAPMdS; Visualization: JDF/EAG/VAPMdS; Writing – original draft: JDF; Writing – review & editing: JDF/EAG/VAPMdS

Bibliography

1. Ro, D.-K., Paradise, E. M., Ouellet, M., *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440,** 940–943 (2006).
2. Tew, D. Synthetic biology and healthcare. *Emerg Top Life Sci* **3,** 659–667 (2019).
3. *LanzaTech eyes two more spin-off companies* TechCrunch. https://social.techcrunch.com/2020/09/15/lanzatech-eyes-two-more-spinoff-companies/ (2021).
4. Chao, R., Mishra, S., Si, T., *et al.* Engineering biological systems using automated biofoundries. *Metab Eng* **42,** 98–108 (2017).
5. Kwok, R. Five hard truths for synthetic biology. *Nature* **463,** 288–290 (2010).
6. Meng, F. & Ellis, T. The second decade of synthetic biology: 2010-2020. *Nat Commun* **11,** 5174 (2020).
7. Hodgman, C. E. & Jewett, M. C. Cell-free synthetic biology: thinking outside the cell. *Metab Eng* **14,** 261–269 (2012).
8. Shears, J. Is there a role for synthetic biology in addressing the transition to a new low-carbon energy system? *Microb Biotechnol* **12,** 824–827 (2019).
9. Linton, J. D. & Xu, W. Understanding and Managing the Biotechnology Valley of Death. *Trends in Biotechnology* (2020).
10. Clarke, L. & Kitney, R. Developing synthetic biology for industrial biotechnology applications. *Biochemical Society Transactions* **48,** 113–122 (2020).
11. Clarke, L., Adams, J., Sutton, P., *et al. A Synthetic Biology Roadmap for the UK* https://connect.innovateuk.org/web/synthetic-biology-special-interest-group/roadmap-for-synthetic-biology (2021).
12. Zaken, M. v. A. *The position of the bioeconomy in the Netherlands - Leaflet - Government.nl* https://www.government.nl/documents/leaflets/2018/04/01/the-position-of-the-bioeconomy-in-the-netherlands (2021).
13. *Biofutures | State Development, Infrastructure, Local Government and Planning* https://www.statedevelopment.qld.gov.au/industry/priority-industries/biofutures (2021).
14. *Bioeconomy strategy: 2018 to 2030* GOV.UK. https://www.gov.uk/government/publications/bioeconomy-strategy-2018-to-2030 (2021).
15. Aurand, E., Keasling, J., Friedman, D., *et al.* Engineering Biology: A Research Roadmap for the Next-Generation Bioeconomy (2019).
16. Kitney, R., Adeogun, M., Fujishima, Y., *et al.* Enabling the Advanced Bioeconomy through Public Policy Supporting Biofoundries and Engineering Biology. *Trends Biotechnol* **37,** 917–920 (2019).
17. *A snapshot of UK research infrastructures | Royal Society* https://royalsociety.org/topics-policy/publications/2018/research-infrastructures-uk-snapshot/ (2021).
18. Hillson, N., Caddick, M., Cai, Y., *et al.* Building a global alliance of biofoundries. *Nat Commun* **10,** 2040 (2019).
19. Zhao, W. A forum on synthetic biology: meet the great challenges with new technology. *National Science Review* **8** (2021).
20. Miehe, R., Bauernhansl, T., Beckett, M., *et al.* The biological transformation of industrial manufacturing – Technologies, status and scenarios for a sustainable future of the German manufacturing industry. *Journal of Manufacturing Systems* **54,** 50–61 (2020).
21. Energie, B. f. W. u. *BMWi startet Förderung der industriellen Bioökonomie – Start-ups, KMU und Leuchtturmprojekte im Fokus* https://www.bmwi.de/Redaktion/DE/Pressemitteilungen/2020/12/20201228-bmwi-startet-foerderung-der-industriellen-biooekonomie-startups-kmu-und-leuchtturmprojekte-im-fokus.html (2021).

22. Ruiz Sierra, A., Zika, E., Lange, L., *et al.* The bio-based industries joint undertaking: A high impact initiative that is transforming the bio-based industries in Europe. *N Biotechnol* **60,** 105–112 (2021).

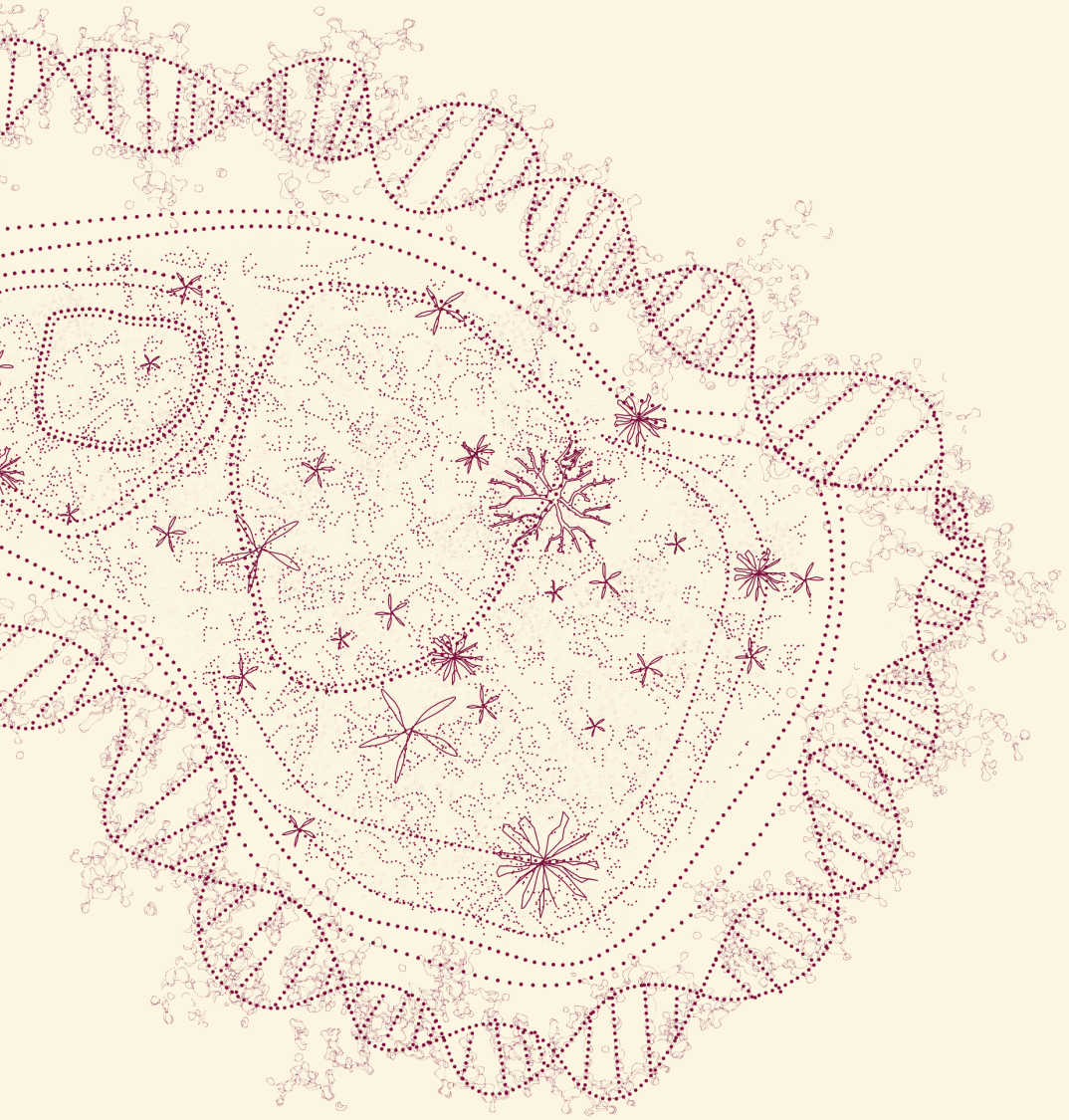
23. *Department of Energy Selects New Projects to Accelerate Innovation and Growth in the Biomanufacturing Sector* Energy.gov. https://www.energy.gov/eere/articles/department-energy-selects-new-projects-accelerate-innovation-and-growth (2021).

24. *U.S. Department of Defense Awards $87.5 Million to EBRC-led BioMADE Establishing the Bioindustrial Manufacturing Innovation Institute* https://www.businesswire.com/news/home/20201020006281/en/U.S.-Department-of-Defense-Awards-87.5-Million-to-EBRC-led-BioMADE-Establishing-the-Bioindustrial-Manufacturing-Innovation-Institute (2021).

25. *Engineering a solution to high throughput COVID-19 virus testing in just nine days | SynbiCITE* http://synbicite.com/news-events/synbicite-blog/engineering-solution-high-throughput-covid-19-anti/ (2021).

26. Holowko, M. B., Frow, E. K., Reid, J. C., *et al.* Building a biofoundry. *Synthetic Biology* **6,** ysaa026 (2021).

27. Farzaneh, T. & Freemont, P. S. Biofoundries are a nucleating hub for industrial translation. *Synthetic Biology (Oxford, England)* **6,** ysab013 (2021).



CHAPTER 12

Thesis summary, general discussion and future perspectives

Enrique Asin-Garcia

Preface

Biotechnological innovations have the ability to stir a range of often conflicting societal responses, as different stakeholder groups perceive risks, benefits, harms, desirability, and necessity of new technologies in contrasting ways [1]. The work presented in this thesis is part of the SafeChassis project, which in turn is included within the NWO Biotechnology and Safety programme (https://www.nwo.nl/en/projects/15814) and aims at implementing and assessing genetic safeguards for lifestyle engineering of the versatile industrial chassis *Pseudomonas putida*. Owing to the multidisciplinary nature of this project, I seized the opportunity to engage with a broad array of stakeholders and the Safety-by-Design approach to address biosafety in the field of synthetic biology applied to biomanufacturing and industrial biotechnology. Thereby, I aimed to contribute to the development of the biosafety field by taking advantage of the immense potential of synthetic biology and by exploring cross-cutting biosafety issues in the laboratory, the classroom, the regulatory and public sectors, and business.

Thesis summary

Biosafety is critical when considering the use of a microorganism as a robust and reliable chassis for industrial applications. In this thesis I explore what makes a synthetic biology chassis a safe chassis and how this can be achieved. To address these questions, I adapted and developed bespoke genetic safeguards using *Pseudomonas putida* as a model organism and positioned biocontainment strategies into the general context of synthetic biology. In addition, I examined different reasonings and understandings related to biosafety as a value, and the *status quo* and prospects of innovation in the transition towards industrial stages of the biotechnology ecosystem.

In chapter 1, I introduce and contextualize the different topics addressed in this thesis. I start by narrating the emergence of metabolic engineering, synthetic biology and biomanufacturing within the context of the bioeconomy, to later continue highlighting why biosafety and biocontainment are valuable features of genetically engineered microorganisms. Semantic containment, genetic circuits and auxotrophy are then introduced as the main types of genetic safeguards and *Pseudomonas putida* is proposed as a relevant chassis that could benefit from the implementation of these biocontainment strategies. Subsequently, I present an outline of the chapters of this thesis clustered in three sections, namely "Genetic Safeguards in *P. putida*" (chapters 2 to 5), "Biosafety" (chapter 6 to 9) and "Biotechnology & Biomanufacturing" (chapters 10 and 11).

Arguing that genome recoding could enhance bioproduction and biocontainment by means of novel functions incorporated into the DNA of *P. putida*, chapter 2 relieves the low recombineering efficiency bottleneck that was hindering this type of genome-scale mutagenesis endeavours. Based on the single-nucleotide precision granted by the minimal PAM ScCas9 variant, the counterselection mediated by this CRISPR-Cas9 system had the power to boost our recombineering capabilities to efficiency levels higher than 90 %, for both single and multiplex genome engineering. The method, which we called ReScribe, allowed us to generate as a proof-of-concept the minimally recoded strain KT2440Rc12, in which all the TAG stop codons from essential metabolic genes were substituted by the synonymous TAA. Thus, ReScribe was shown to be a powerful and convenient editing tool which

represents an advancement for recombineering of non-model organisms. Additionally, its expanded targetable space makes CRISPR-Cas9 counterselection possible in applications where it could not be used before.

To find efficient recombineering systems that would not require counterselection, chapter 3 describes the high-throughput screening of a library of potential recombinases, which was generated using a bioinformatics-metagenomic approach. Candidates were surveyed using a serial enrichment recombineering workflow whose progress could be followed by means of a Next-Generation Sequencing (NGS) analysis. Unlike previous pipelines of parallelized protein discovery, our method makes use of Oxford Nanopore sequencing for a more user-friendly, streamlined and cost-effective approach that ultimately allowed us to uncover recombinases with efficient genome editing capabilities. The efficacy of the system to yield active candidates in *P. putida* prompted us to test the library in other related *Pseudomonas* species with relevance in the biotechnology field. In this way, we reported the first recombineering systems available for *Pseudomonas taiwanensis* VLB120 and *Pseudomonas fluorescens* SBW25, and new alternative recombinases for *Pseudomonas aeruginosa* PAO1.

Since a fully recoded strain was beyond the scope of this thesis, we pursued biocontainment solutions rooted in other types of robust genetic safeguards with additional features of industrial interest. Taking advantage of the high cleavage efficacy and the minimal PAM requirements of ScCas9 shown in *P. putida* in chapter 2, we made use of this endonuclease to implement a highly genotoxic genomic device for biocontainment in chapter 4. To achieve a quick and efficient lethal response upon Cas9 activation, we exploited *P. putida*'s repetitive genomic regions as cleavage targets of the CRISPR-Cas9 mechanism creating *P. putida* GenoMine. This strategy could be exploited beyond biocontainment in situations where further CRISPR genome editing is not desired. In addition, and with the objective of limiting the viability of this strain to controlled environments, we attempted to block the activation of Cas9 with repressors controlled by different genetic circuits based on riboregulators or a digitalized version of an inducible promoter. Unfortunately, repression levels were insufficient to completely prevent Cas9 cleavage upon circuit induction, which set us on the quest for strongest repressors for *P. putida*.

The implementation of a synthetic phosphite auxotrophy in *P. putida* for effective biocontainment is described in chapter 5. To generate the auxotrophic strain PSAG-9, we first implemented in the genome a phosphite assimilation pathway consisting of a phosphite-specific transport complex and a phosphite dehydrogenase. In this way, *P. putida* could incorporate phosphite that could be then converted to metabolizable phosphate. Second, all the genes related to the native transport of phosphate and its derivatives were knocked out to prevent the strain from growing on media containing other phosphorous source than phosphite. After this objective was achieved, the strain showed competitive fitness levels when growing in phosphite and low levels of escape frequency when this compound was not present in the media. Beyond biocontainment, this dependency provided additional NADH regenerative capacity due to the NADH produced during the conversion from phosphite to phosphate, and allowed us to grow *P. putida* under non-sterile conditions without the risk of biological contamination given that phosphite is generally not metabolizable by other microorganisms.

The more technical "Genetic Safeguards in *P. putida*" first section of this thesis transitions into the more philosophical "Biosafety" section in chapter 6, in which genetic safeguards are critically analyzed as the materialization of the Safe-by-Design paradigm in synthetic biology. While theoretically attractive, integrating genetic safeguards in real-world applications is rarely exercised in practice, in spite of technological progress and regulatory claims about their utility as risk assessment tools. In this chapter, we point at a knowledge gap in their assessment and an evasion of responsibility regarding their ultimate implementation as the main bottlenecks to this end, and conclude providing some advice in the matter.

In chapter 7, we highlight the importance of application contexts for a responsible and pragmatic development of genetic safeguards. We argue that some view on contexts of application for these technologies is required to facilitate responsible innovations in synthetic biology. Consequently, we propose that an explicit strategy of contextualization is needed meaning that potential applications contexts should be emphasised at an early stage. This leads to a socially responsible development of genetic safeguards that takes into account different values and interests. Additionally, we discuss concrete practical benefits of contextualization in

terms of representation, assessment of safety and evaluation of safety desirability, and anticipate potential tensions derived from embracing this approach as a responsible research and innovation strategy.

Continuing in the line of values, chapter 8 gets inspired by the Value Sensitive Design framework to investigate scientific and technological design choices in their appropriate social context. Here, we came back to primary research approaches to explore biosafety norms, reasonings about genetic safeguards and how these two impact the practice of designing for the value of biosafety in industrial biotechnology. Our findings show that tensions among stakeholders are found at the level of norms, and that, a practical value specification of how to realize biosafety necessitates of an early stakeholder alignment. Without this common effort, we conclude that the differences in informal biosafety norms and the disarray in biosafety thinking lead to design requirements for compliance instead of for safety.

The investigation about the value of safety continues in chapter 9, but this time on the field of life sciences education. Here, we explored what is taught about safety and how safety is taught in higher education programmes using Wageningen University as a case study. A gap analysis on safety education was created based on qualitative interviews with teachers, which ultimately yielded two main gaps. In the first place, most attention is paid to technical safety indicating a bias towards compliance, and in the second place, safety elements' cognition receives more attention than the skills to address them. We propose that this can be addressed early in scientific and engineering training by shifting the focus to higher levels of cognition focusing on skills and attitudes, and setting up education for responsibility instead of for compliance.

Chapter 10 goes beyond biosafety and initiates the third section of this thesis which treats more general matters related to "Biotechnology & Biomanufacturing". Specifically, this chapter presents our quest to ascertain the factors underlying the traditionally called Valley of Death in Biotechnology and possible opportunities to transverse it. To this end, we carried out an exploratory study consisting of in-depth interviews with industrial and academic participants across sectors and national boundaries, and dealt with a number of specific technical, business and societal aspects. The inquiries indicated that academics are limited mostly by technical

factors, whereas industry is restricted by technical, sector-dependent and societal ones, which leads to a misalignment of interest, miscommunication and missed cooperation opportunities. Although both are of the opinion that academia must perform curiosity-driven research to find innovative solutions, there is pressure to aim for short-term industrial applications. These factors add up to widen the Valley of Death. A third party, in the form of start-ups, could be the answer to traverse this valley more effectively, particularly when embedded in adequate innovation ecosystems.

The message of chapter 11 is conveyed by means of a revision, a request and a personal recommendation. First, we review the biomanufacturing field including *status quo*, expectations and pitfalls of a sector that plays a fundamental role in the bioeconomy. Next, we claim that to accelerate bioproduction progress, there is an urgent need for interdisciplinary research infrastructure. Lastly, we introduce the concept of meta-workflows as frameworks that connect ecosystems of research infrastructures, highlighting the importance of interoperability, harmonisation and sustainability.

Finally, this chapter, 12, recapitulates the contents of this thesis with this summary and discusses future directions by reflecting on three main themes: *P. putida* as a safe chassis for industrial biotechnology, the prospects of biocontainment, and the bridges between innovation and application in biotechnology.

General discussion and future perspectives

P. putida as a safe chassis for industrial biotechnology

Not there yet

Long gone are the days when *Pseudomonas putida* was simply considered for its potential to biodegrade xenobiotics [2]. Today, after more than half a century of study, streamlining and technology development, the industrialization of this bacterium has shifted onto a whole different level. Beyond bioremediation studies [3–6], efficient bioproduction of a wide range of products using this microbe has also been demonstrated [7].

At present, several companies are making use of this microbe (*e.g.*, Corbion (The Netherlands) produces FDCA from hydroxymethyl furfuraldehyde (HMF) [8], different companies such as Kaneka (Japan), Telles (USA), Jiangsu Nantian (China) or Biomer Biotechnology (Germany) produce PHA polymers, etc. [7, 9–11]). Yet, the road towards actual industrial applications has not always been straightforward for *P. putida* given the many factors at play that influence the choice of the production platform, as we found in chapter 10, both at technical and societal levels. Despite meeting the desired criteria, *P. putida* is often not selected for an application because of the historical use of model organisms, typically *Escherichia coli* or *Saccharomyces cerevisiae* [12]. We must always bear in mind, that the primordial objective of a private company is profit for its shareholders, and that, according to this principle, non-cost-effective actions are seldom taken. In this context, either the substitution of the previous platform by *P. putida* does not affect the production process, or its titre-rate-yield (TRY) credentials are higher enough than those of the previous chassis for the revenue to exceed the replacement investment and adjusted operational conditions. The sad part is that, at times, not even this is enough, since industry must always keep the end-users in mind, who typically demand products that are sustainable and above all safe. These claims have far-reaching implications when considering the use of *P. putida* as an industrial chassis. For example, as we saw in chapter 8, some companies, especially those in which the end-product comes into contact with the costumers (food, fragrances, etc.), prefer to employ native producers (or, illogically from the scientific point of view, even

strains enhanced *via* classical random mutagenesis) rather than genetically engineered microorganisms (GEMs) to avoid the negative connotations associated to GMO labeling, and for other marketing purposes [13]. In these cases, *P. putida* bioproduction would therefore be limited to products that can be generated naturally. By the same token, biodegradation applications are also restricted to the natural capabilities of *P. putida* if these need to happen out of physically contained settings as a consequence of the stringent GMO regulations.

In the end, it seems to me that the potential of *P. putida* as a chassis for industrial and environmental applications needs more than just proofs of feasibility and productivity. With production of PHA, FDCA and rhamnolipids as rare exceptions, large-scale industrial applications of this microbe remain scarce [7]. For this situation to change, our focus needs to be placed on improving strains on industry-scale conditions and on further optimizing processes [14]. Moreover, I believe that only by dedicating attention to other key aspects of the process such as safety, robustness or sustainability, more business cases featuring *P. putida* will become viable and successful.

On the first part of this dissertation, I focused on the implementation of different genetic safeguards for this bacterium's biocontainment in an attempt to enhance its safety credentials. As described in chapter 1, the main strategies to achieve this goal are generally semantic containment (*e.g.*, recoding), genetic circuits and auxotrophy. The three of them were pursued in *P. putida* KT2440, generating the strains KT2440Rc12 (chapter 2), GenoMine (chapter 4) and PSAG-9 (chapter 5) (Figure 12.1), but since only the first of these strategies has the ability to attain isolation at both biological and genetic level, we first aimed at altering the microbe's genetic code by constructing a recoded genome without a TAG stop codon.

Recoded *P. putida*

Replacing a specific codon globally with a synonymous one in the degenerate 64-codon code [15], followed by the corresponding tRNA or releasing factor removal, allows the repurposing of this free codon [16, 17], which in turn grants numerous valuable features [18–21]. These include biological isolation [22], genetic isolation [20, 23], resistance to bacteriophages [24, 25], a tool to address fundamental bio-

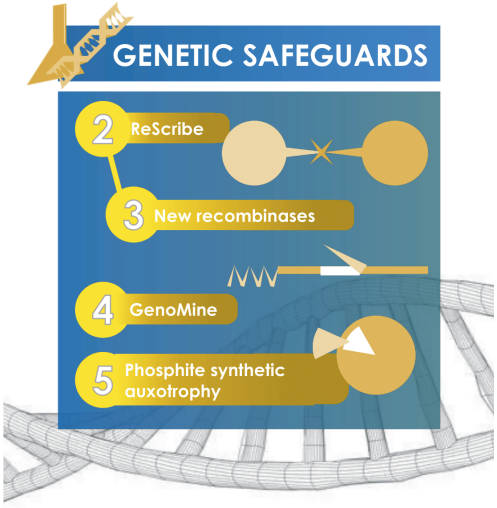


Figure 12.1: Strains KT2440Rc12 (minimally recoded strain of essential metabolic genes), GenoMine (strain equipped with a kill-switch device), and PSAG-9 (phosphite auxotroph) were generated as part of this thesis to illustrate the different types of genetic safeguards in P. putida.

logical questions [26], and the power to expand the properties of proteins towards novel functions never found before in nature [27]. Unfortunately, to achieve that, the recoding process needs to be done throughout the whole genome.

KT2440Rc12 is simply a minimally recoded strain because TAG stop codons were substituted by synonymous TAA only in essential metabolic genes (chapter 2). When I started this project, I held the view that the effort, time and resources necessary for generating a fully recoded *P. putida* strain were monumental and required more efficient genome editing methodologies than the ones that were available at that moment [28, 29]. In that respect, it is worth mentioning that at present, even after 10 years of its first report, no other organism has been fully recoded beyond *E. coli* [16], except *Caulobacter ethensis*, which has a 0.78 Mb genome and was ultimately non-viable [30, 31]. Large recoding enterprises beyond that one have only taken place in *Salmonella typhimurium* [32] and yeast [33]. In addition, the *E. coli*'s recoding effort relied on 321 TAG substitutions, whereas *P. putida*'s genome contains 654 genes finished in the TAG stop codon. As a consequence, the starting point of our work was the development of technology and the search for new systems that would

allow large-scale manipulation of *P. putida*'s genome. Ultimately, this motivated the findings of chapter 2 and 3, which in turn built upon key technologies such as the minimal PAM ScCas9 variant, new recombineering systems and Oxford Nanopore next-generation sequencing (NGS).

Cas enzymes with relaxed PAM have been receiving more and more attention given the access they grant to a much more expanded fraction of the genome [34–37]. Today these developments already allow the edition of many previously inaccessible regions improving a variety of editing applications [38]. However, I believe that further investigation of these variants will ultimately enhance the precision editing prospects towards even higher-resolution genome perturbations. Useful as it is, CRISPR-ScCas9-mediated counterselection is, in this particular method, also a source of limitations for the recombineering process as discussed in chapter 2, which encouraged us to seek alternative recombineering strategies in chapter 3 using a high-throughput method involving NGS. Unlike previous pipelines of parallelized protein discovery [39], our NGS analytical method makes use of Oxford Nanopore sequencing for a more user-friendly, streamlined and cost-effective approach that ultimately allowed us to uncover recombinases or single-stranded annealing proteins (SSAPs) with high genome editing capabilities in four different species of *Pseudomonas*.

Despite of this, none of the recombinases was able to yield results within the category of high allelic replacement frequencies (ARF 10-25 %), a threshold that has only been surpassed in a handful of microorganisms [39–41]. To tackle potential intrinsic factors that might be limiting the allelic replacement process in *P. putida*, several approaches have been proposed during the last years. These include the addition of mismatch repair inhibitors [42], which has already been adopted by the *P. putida* community [43–45], or the portability across species of compatible SSAP-SSB pairs [46, 47]. Furthermore, special mention needs to be made of game-changing strategies such as retron library recombineering (RLR), which uses the targeted reverse-transcription activity of retrons to produce ssDNA *in vivo* without the need of exogenously provided oligonucleotides [48]. While the system is able to yield edited cell population fractions >90 %, its continuous editing nature unavoidably derives in significant off-target effects.

Despite the progress in recombineering and other genome editing techniques, the relentless pace at which DNA synthesis and assembly methods advance make me wonder to what extent whole-genome mutagenesis endeavors are still worth our time and resources. In 2019, the world witnessed the first total synthesis of an *E. coli*'s recoded genome by Fredens and colleagues [49]. While this type of projects remains currently unaffordable for most research groups, the increasing demand will surely drive down DNA synthesis costs in the not-too-distant future [20]. At this point, we need to consider if our eagerness for recoding organisms and their advantageous features outweighs the enormous efforts, or if instead it would be a more efficient option to wait for the expenses of DNA synthesis to become less prohibitive even though this will probably entail a certain trivialization of the process. At present, moving forward from this scenario seems to me only possible by researching efficient genome editing tools with parallelized potential applications beyond recoding organisms to amortize any possible investment. In the meantime, I dare say that the exploitation of the appealing features of recoding in *P. putida* will only be possible by developing alternative strategies with the potential to confer some of these benefits to the microbe.

Other sources of biocontainment for *P. putida*

Recognizing that a fully recoded strain was beyond the scope of this thesis, I pursued biocontainment solutions rooted in genetic circuits and synthetic auxotrophy with additional features of industrial interest. Strains GenoMine and PSAG-9 were thus generated in chapters 4 and 5, respectively. With these developments, I aimed to connect with the idea of building a suite of *P. putida* strains harboring a subset of individual safeguards with high performing metrics and potentially applicable to customized contexts (chapter 7). Nonetheless, there is room left for improvement as regards enhancing the robustness of existing strains and expanding the repertoire.

Concerning our incapability to tightly control Cas9 in *P. putida* GenoMine, future efforts will need to be concentrated on finding optimal repressing molecules, for this application and for the toolbox of this bacterium in general, which remains scarce in this kind of genetic elements [50]. A possible option would be the use of

RNA binding proteins (RBPs) from the coat of bacteriophages with the capacity to repress translation in bacteria by inhibiting or directly competing for the binding site of the ribosome [51]. To date, several RBPs have been characterized originating from different bacteriophages [52–55] with different strength levels of repression and the capacity to bind natural and altered coat protein binding sites (CPBS) in *E. coli* [51]. Still, these molecules have not yet been harnessed as elements of complex genetic circuits, which most likely will require the appropriate fine-tuning for specific species and applications.

Furthermore, to control these repressing molecules, other regulatory elements could be considered for circuit construction given the limitations of the employed systems. Recently, a theophylline-dependent [56] and a fluoride-responsive [57] riboswitches have been successfully applied in *P. putida*. In addition, a complete and robust CRISPRi toolbox to efficiently control this bacterium's transcriptional levels has been published featuring a modular and standard architecture [58]. Cell's dynamics can then be regulated by linking the CRISPRi system to inducible promoters or biosensors based, for example, on temperature [59] or the quorum sensing machinery [60].

Regarding auxotrophies for biocontainment, a few significant contributions have been made during the last years reporting strains dependent on ligands [61] and non-canonical proteinogenic amino acids [62]. Still, no better strategy than the one we used to create *P. putida* PSAG-9 has emerged in the last years in terms of escape frequency [63]. Unfortunately, and despite reaching acceptable safety levels ($<10^{-8}$), our phosphite-dependent *P. putida* was still far from the impressive escape frequency in the order of 10^{-13} that was achieved in *E. coli* by Hirota *et al.* [64]. With the potential reasons for this difference detailed in chapter 5, I would like to use this opportunity to turn briefly to the issue of orthogonality, a metaphor I consider abused by synthetic biologists.

Orthogonality is a concept borrowed from mathematics that originally referred to the notion of perpendicularity to the linear algebra of bilinear forms. The use that we give to the term is, however, related to a factual independence between otherwise co-existing systems. In biological terms, this becomes even less accurate and simply implies a lesser dependence of the host's entity [65]. While very few

biological elements would deserve a genuine orthogonal qualification (*e.g.*, some bacteriophage polymerases), the idea behind the metaphor is constantly attributed to functions, devices, systems and even whole cells in the synthetic biology literature. The subfield of biocontainment does not lag behind and endures the notion of orthogonality applied to parts, circuits and genetic safeguard strategies [66–70]. In an ideal situation, an engineered regulatory element, such as, for example, riboregulator #12 (chapter 4), whose functioning is based on the basic biological principles of RNA base-pairing, would yield the same output in two bacteria as related as *E. coli* and *P. putida*. However, we have seen that this was not the case. Orthogonality is often claimed after exclusively experimenting in *E. coli*, which by no means proves that the behavior will be the same on a different organism. As much as synthetic biology strives to simplify living systems, there is so much that we do not fully understand yet. The independence that an orthogonal element requires should, however, be possible only by fully comprehending the framework within which it operates. In this context, catalytic vesicular "cells" deprived of their genome [71–73], and bottom-up synthetic cells and biomolecular systems [74, 75] represent a step forward towards predictability and device orthogonality, due to their reduced amount of unknowns. However, taking all this into account, I wonder until which point we should rely on data and indications obtained for a specific device when installed in model organisms. Will those biocontainment strategies with lower escape frequencies in *E. coli* end up being the safest in other organisms and in the field? Which credit should we give to the currently available metrics when these are performed in a setting that has nothing to do with the final application context? How orthogonal a biological system can be if biological systems themselves are fully dependent on their environment?

With this disclaimer, my intentions were far from playing down the great work that my synthetic biology peers do every day with model organisms in controlled laboratory settings. Instead, I wanted to stimulate reflection about the irrelevance of decontextualized data and numbers as well as the damage that metaphors can inflict when we lose sight of the big picture.

The prospects of biocontainment

Multilayering and computational integration

Multilayered approaches, in which multiple safeguards are combined into a single strain, rely on different modes of action to address shortcomings inherent to each individual safeguard [76–78]. Importantly, each safeguard mechanism would be capable of scaling and thus provide solutions with many combinations that can be implemented across strains for varied use. While the idea of adding multiple layers appears reasonable to multiply the safety levels, it can also evoke those medieval knights equipped with several layers of heavy armor consisting of tabards, mail colfs, leather brigandines and metal plackart, who eventually could barely move. Hence, one might need to take into account that the strain's performance might be affected by the heavy burden of several combinatorial devices controlling cell's viability, as well as unpredictable results derived from the incomplete knowledge base associated therewith [79].

By aiming for predictive design and control capacity, computational models represent powerful technologies for enhancing biocontainment strategies [80]. In the first place, genome-scale, constraint-based metabolic models could be used to identify (multiple) gene targets like essential genes for genetic circuits or metabolic targets to engineer an auxotrophy. Moreover, integrated metabolic and gene expression models, as well as digital twins, with the capacity to compute cellular processes beyond metabolism (synthesis of proteins, translation effectors, energetic costs, regulation, etc.) can provide an even more comprehensive framework [81–83]. In addition, co-culture models can be exploited for the prediction of biocontainment within microbial networks facilitating predictions related to interactions, perturbations and metabolite exchange [84–86]. Lastly, metabolic robustness analysis postulates as a valuable tool to predict and assess the fitness of a dynamic biological system, which would be particularly relevant when this one is expected to behave differently in the laboratory and the final application settings. Under permissive conditions, a strain equipped with a biocontainment circuit will show steady-state metabolism provided by the availability of the key effector that controls the function of a key enzyme. However, under non-permissive conditions, in the absence of such an effector, cell decay will commence exhibiting a series of

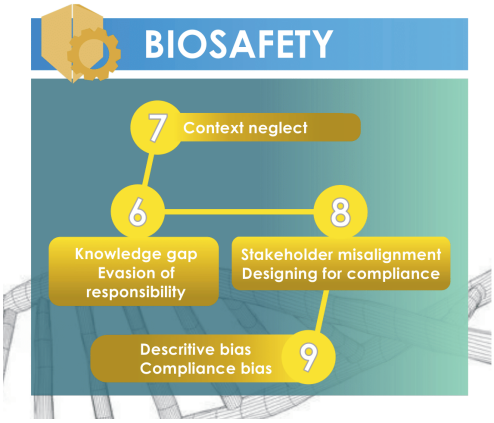


Figure 12.2: Main limitations related to the implementation of genetic safeguards and preemptive approaches identified within distinct exploratory studies of this thesis which analysed biosafety from different angles.

phenomena including the programmed change in the key enzyme, a decrease of vital metabolites and an increase of stress-associated factors. Metabolic robustness, and therefore the probability of a system failing due to a perturbation, could be assessed by modeling different kinetic data obtained from playing with the parameters and perturbations that affect the levels of the key enzyme and the corresponding maximum rate. In this way, we could study the viability of our safe chassis under a variety of environmental conditions [79, 87, 88].

Lessons for the future

Despite technical opportunities, the future of biocontainment is fraught with obstacles. In the second part of this thesis, I navigated through the subject of biosafety highlighting genetic safeguards as a design basis for achieving this value. In this way, I identified practical, axiological and societal bottlenecks directly related to the implementation of approaches that: (i) consist of genetic safeguards (chapter 6), (ii) are deployed in real and contextual applications (chapter 7), and (iii) aimed at considering different safety aspects of a biotechnological process (chapter 8). In addition, chapter 9 recognized safety limitations that occur at the education stages, which have the power to shape the mindset of future scientists (Figure 12.2).

On top of all the recommendations provided in each of the chapters, my own

opinion on the subject is that to move forward, we all need to be on the same page. During my trip exploring biosafety, I spoke to many people, heard even more and read extensively on the topic. While there are a few basic thoughts that almost all of us share ("biosafety deserves some attention", "there is a certain degree of uncertainty in genetically engineering organisms", etc.), a profound disarray of opinions, understandings and reasonings dominates the biosafety field, dramatically complicating any effort towards concerted progress. As I mentioned in chapter 8, these beliefs are not even stakeholder group-specific. Rather, they seem to be simply individual-specific, which in turn makes the matter anything but simple. Based on the same set of observations, these stakeholders perceive different problems, deduct different factors that cause them, and propose different recommendations to address them, all of it most likely determined by their individual worldviews [89]. These are constituted by ontologies, which in turn encompass particular anthropological, epistemological and axiological visions [90]. Joachim Spangenberg quoted, "The world we see shapes the world we create" [91], which aligns remarkably well with what I have observed when exploring technical designs for biosafety.

Despite the absence of biosafety incidents reported so far, the difference in speed between innovation and regulation demands not only common guidelines and governance to help us prevent future crises [92], but also more flexible frameworks that allow us to consider new biotechnologies from overarching perspectives. The Safe-by-design (SbD) approach, referred to so many times throughout this thesis, is an exemplary of this vision, but so are educating in skills and attitudes and linking biosafety to other values to increase its own relevance. We need to stop thinking about biosafety as another ticked box during the process, and instead, we should start harnessing it as a driver for innovation. By developing attitudes and conducts of responsibility and consideration about future implications of their own research (even when these are not yet legislated), scientists will have the capacity of embedding these safety aspects within their designs, hopefully contributing to a conscientious maturation of their research fields. This rationale is not completely exempt from controversy since it requires anticipatory knowledge that might collide with current regulations, and trusts a concrete group of people to judge universal social values. However, this could be solved by means of adaptive actions

and measures, as well as comprehensive and equitable innovation ecosystems.

Additionally, value co-creation can contribute to adopting a safety culture in biotechnology. Not only is safety intrinsically associated with other broadly positive values related to responsibility, society and the environment, but it also has the capacity to result in industrially appealing ones. Such is the case of the phosphite auxotrophy biocontainment (chapter 5), which on top of boosting the biosafety credentials, is able to prevent economic losses associated with biological contamination in the bioreactors, and confer additional reducing power for enhanced bioproduction. Personally, I believe this value co-creation is the key to set in motion the incorporation by companies of biosafety technologies in real applications, at least during the first stages of the transition towards more responsible and less compliance-oriented processes. Sooner or later, many responsibility choices, not only related to safety, but also to matters of wealth distribution, resources and sustainability, will have to be made if we want our economic system to overcome the great environmental and (therefore) financial challenges that lie ahead.

A biotechnology of bridges

Increasing awareness of the significant global challenges derived from our fossil-based economic system has positioned the bio-based economic model in the spotlight of industries and governments [93]. As a consequence, much hope has been placed on biotechnology and its power to replace former technologies with more sustainable alternatives. After treating these considerations in the introductory chapter, I dedicated the third part of this dissertation to deepening into general concepts related to biotechnology and biomanufacturing. The two chapters comprising this section focused on ways to help bridging the gap between the inception of biotechnologies and their translation to industrial applications (Figure 12.3).

Thomas Harrison said that a bridge is “not just a symbol, but the symbol of symbols, a means to a region holding hidden significance” [94]. While curiosity and innovation are the first indispensable dwelling of science, current expectations need biotechnology to reach that other shore of significant gains by means of genuinely profitable applications. To ensure that international commitments to mitigate the climate crisis are fulfilled, changes of established industrial operations

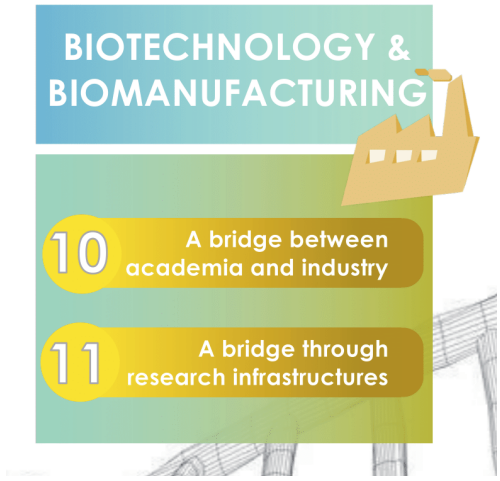


Figure 12.3: The two chapters included in the third section of this thesis consider broader topics related to general biotechnology and biomanufacturing and discuss "bridges" to cross the biotechnological Valley of Death, opportunities between industry and academia, and the potential of interdisciplinary research infrastructures to address issues of current innovation processes.

need to be radical and immediate [95, 96]. At this critical point, biotechnology must leave behind its clichéd "future potential" and embrace now its responsibility for the bioeconomy.

In this way, biotechnology's responsibility has become dual. On the one hand, we have the responsibility derived from the field's uncertain nature, which, as it has been discussed throughout this thesis, has to do mainly with the safety of the bioproducts, even though there are other non-risk-related concerns that have also been recognized for some time now [97]. On the other hand, we encounter a responsibility founded on the quality, efficacy and thus the ability of the biological alternatives to replace the previous, less sustainable products and processes. These two perspectives often feel distant because while the latter is encouraged by urgency, the former questions every step of the innovation process. However, I believe their clashing to be unlikely. Ethical debates on biotechnological issues are fertile ground for the democratization of technology, and the establishment of adequate innovation ecosystems and roadmaps. In absence of multidisciplinary information and independent experts, biotechnology would be vulnerable to the

full force of industrial and political lobbies seeking to control their own scientific choices. Heraclitus' sense of unity was that "by being at variance with itself, it is in agreement, like the harmony of bow or lyre" [98]. This is yet another bridge, that like a bow or a lyre, can only fulfil its purpose if there is a tension that prevents it from collapsing. Harrison beautifully describes this dilemma as a lesson to learn on "how to respect the eternal apartness of things while making them dance together" [94].

Broader multi-stakeholder discussions, as well as other participatory and deliberative activities are great contributors for the identification of perceptions and values that can be considered in policy and decision making. However, despite the efforts towards a more inclusive debate on biotechnology, the general voting behaviour on the safety assessment for market applications goes ahead slowly. This is most likely rooted not only on the resilience of normative frames, but also on heterogeneous expectations and underlying visions of the world, future and nature [99]. In the context of the aforementioned dual responsibility, we biotechnologists need, perhaps even more than other actors, to tune our appreciative system, that is, our judgement of reality, interests and accountability [100], at the professional level to understand these disparities. The 1989 edition of On Being a Scientist stated that "scientists must always be aware that their work ultimately may have a great impact of society [101]. However, biotechnologists often remain secluded to the technicalities of a laboratory, instead of becoming the bridge among stakeholders that current innovation ecosystems demand. For our research to be successful and as a key for bioeconomic growth, we need to incorporate new principles to our mindsets (sustainability, interoperability, harmonisation, democratisation and larger bioeconomic context were, for example, the pillars on which the meta-workflows of chapter 11 were sustained) and reach out to others' conceptions of our work. Like it would happen in ancient port cities like Istanbul, we need to bridge cultures, pluralistic ideas and understandings, if we want to exploit the benefits of progress.

The situation of the *isolated* biotechnologist is one that I see exacerbated at early scientific stages in academic environments. Under pressure to quickly become productive members of the laboratory in which he/she is part of and the scientific

Figure 12.4: A biotechnology of bridges. Puente de los Franceses (Bridge of the French) over the Águeda river located on the Natural Park of Arribes del Duero, San Felices de los Gallegos and Puertoseguro, Salamanca, Spain.

community in the particular field of research, many PhD programmes have reduced the load of opportunities for research contextualization, which ultimately ends up training specialists instead of thinkers [102]. It goes without saying that I completely disagree with such an approach. Scientific productivity in our educational systems (at different academic levels) tends to depend more on rote knowledge than on competences in critical thinking, which leaves no room for uncertainty, information not yet available and emerging situations. This type of education is on a collision course with the necessary development of biotechnology, which will only come from self-monitoring its contributions to the world through the lens of social commitment and responsibility.

During the course of this PhD candidacy, I had the opportunity to connect with many people, from different stakeholder groups (academic researchers, industrial workers, teachers, study programme coordinators, regulators, risk assessors, policy makers, politicians, philosophers, ethicists, artists, technology transfer officers, serious-game developers, etc.). These connections did not cancel out our dispar-

ities, cement ever-lasting unions or merge me with other professions, but worked once again as bridges "binding territories that remain disparate, albeit joined" (Figure 12.4) [94]. After this dissertation, I am still *just* a biotechnologist working on synthetic biology. I believe, however, that the broad education that I received (and sought) during these last years will help me doing science more thoughtfully and creatively. In this way, and as part of the scientific quest for a better world, I only hope this work contributes to the field of biotechnology and therefore society at large.

Concluding remarks

This project represents a meaningful advancement in the area of biosafety in the fields of biotechnology and synthetic biology. It does so by developing and implementing technical genetic safeguards, as well as by addressing cross-cutting biosafety issues.

On the technical side, the first three objectives of this thesis were focused on the implementation of sophisticated measures using *P. putida* as a case study aiming to mitigate environmental and health-related safety effects of synthetic biology. In this way, different safeguarded *P. putida* strains were generated equipped with biocontainment strategies that additionally granted other industrially appealing co-created values to the microbial chassis and rendered them as valuable material for risk assessment.

Furthermore, this thesis combined the previous technical aspects with in-depth philosophical reconnaissance, analysis and assessment of risk perception. Different explorations, including those of reasonings and understandings behind the concept of genetic safeguards, biosafety as a value, and opportunities and prospects of industrial biotechnology and biomanufacturing, are as well presented as outcomes of this thesis.

This research has posed questions that needed to be addressed sooner than later in the fields of biocontainment, biosafety and biotechnology and has filled gaps in existing knowledge. In addition, it will hopefully contribute to pave the way to a type of science that strives on self-improvement by encouraging an engagement with bigger-picture problems and with social responsibility.

Bibliography

1. De Witt, A., Osseweijer, P. & Pierce, R. Understanding public perceptions of biotechnology through the "Integrative Worldview Framework". *Public Understanding of Science* **26,** 70–88 (2017).
2. Nakazawa, T. Travels of a Pseudomonas, from Japan around the world. en. *Environmental Microbiology* **4,** 782–786 (2002).
3. Shim, H. & Yang, S.-T. Biodegradation of benzene, toluene, ethylbenzene, and o-xylene by a coculture of Pseudomonas putida and Pseudomonas fluorescens immobilized in a fibrous-bed bioreactor. *Journal of Biotechnology* **67,** 99–112 (1999).
4. Jimenez, J. I., Minambres, B., Garcia, J. L., *et al.* Genomic analysis of the aromatic catabolic pathways from Pseudomonas putida KT2440. *Environmental Microbiology* **4,** 824–841 (2002).
5. De Lorenzo, V. Systems biology approaches to bioremediation. *Current Opinion in Biotechnology. Chemical biotechnology / Pharmaceutical biotechnology* **19,** 579–589 (2008).
6. Udiković-Kolić, N., Scott, C. & Martin-Laurent, F. Evolution of atrazine-degrading capabilities in the environment. *Applied Microbiology and Biotechnology* **96,** 1175–1189 (2012).
7. Weimer, A., Kohlstedt, M., Volke, D. C., *et al.* Industrial biotechnology of Pseudomonas putida: advances and prospects. *Applied Microbiology and Biotechnology* **104,** 7745–7766 (2020).
8. Corbion. *The future of plastics is biobased* https://www.corbion.com/bioplastics/fdca (2021).
9. Poblete-Castro, I., Becker, J., Dohnt, K., *et al.* Industrial biotechnology of Pseudomonas putida and related species. *Applied Microbiology and Biotechnology* **93,** 2279–2290 (2012).
10. Poltronieri, P. & Kumar, P. Polyhydroxyalcanoates (PHAs) in Industrial Applications. *Handbook of Ecomaterials,* 1–30 (2017).
11. Van de Graaf, A. *FDCA-race naar de markt* C2W. https://www.sciencelink.net/nieuws-and-verdieping/fdca-race-naar-de-markt/1027.article (2022).
12. Calero, P. & Nikel, P. I. Chasing bacterial *chassis* for metabolic engineering: a perspective review from classical to non-traditional microorganisms. *Microb. Biotechnol.* **12,** 98–124 (2019).
13. Wunderlich, S. & Gatto, K. A. Consumer Perception of Genetically Modified Organisms and Sources of Information. *Advances in Nutrition* **6,** 842–851 (2015).
14. Yan, Q. & Fong, S. S. Challenges and Advances for Genetic Engineering of Non-model Bacteria and Uses in Consolidated Bioprocessing. *Frontiers in Microbiology* **8,** 2060 (2017).
15. Plotkin, J. B. & Kudla, G. Synonymous but not the same: the causes and consequences of codon bias. *Nature Reviews Genetics* **12,** 32–42 (2011).
16. Isaacs, F. J., Carr, P. A., Wang, H. H., *et al.* Precise Manipulation of Chromosomes in Vivo Enables Genome-Wide Codon Replacement. *Science* **333,** 348–353 (2011).
17. Lajoie, M. J., Rovner, A. J., Goodman, D. B., *et al.* Genomically recoded organisms expand biological functions. *science* **342,** 357–360 (2013).
18. Lajoie, M. J., Söll, D. & Church, G. M. Overcoming Challenges in Engineering the Genetic Code. *Journal of Molecular Biology* **428,** 1004–1021 (2016).
19. Mukai, T., Lajoie, M. J., Englert, M., *et al.* Rewriting the Genetic Code. *Annual Review of Microbiology* **71,** 557–577 (2017).
20. Kuo, J., Stirling, F., Lau, Y. H., *et al.* Synthetic genome recoding: new genetic codes for new features. *Current Genetics* **64,** 327–333 (2018).

21. Santos-Moreno, J. & Schaerli, Y. Changing the biological Rosetta stone: the (commercial) potential of recoded microbes. *Microbial Biotechnology* **13,** 11–13 (2019).

22. Rovner, A. J., Haimovich, A. D., Katz, S. R., *et al.* Recoded organisms engineered to depend on synthetic amino acids. *Nature* **518,** 89–93 (2015).

23. Mandell, D. J., Lajoie, M. J., Mee, M. T., *et al.* Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518,** 55–60 (2015).

24. Ma, N. J. & Isaacs, F. J. Genomic Recoding Broadly Obstructs the Propagation of Horizontally Transferred Genetic Elements. *Cell Systems* **3,** 199–207 (2016).

25. Wang, L., Jiang, S., Chen, C., *et al.* Synthetic Genomics: From DNA Synthesis to Genome Design. *Angewandte Chemie (International Ed. in English)* **57,** 1748–1756 (2018).

26. Martínez, M. A., Jordan-Paiz, A., Franco, S., *et al.* Synonymous Virus Genome Recoding as a Tool to Impact Viral Fitness. *Trends in Microbiology* **24,** 134–147 (2016).

27. Agostini, F., Völler, J.-S., Koksch, B., *et al.* Biocatalysis with Unnatural Amino Acids: Enzymology Meets Xenobiology. *Angewandte Chemie International Edition* **56,** 9680–9703 (2017).

28. Aparicio, T., Jensen, S. I., Nielsen, A. T., *et al.* The Ssr protein (T1E_1405) from *Pseudomonas putida* DOT-T1E enables oligonucleotide-based recombineering in platform strain *P. putida* EM42. *Biotechnology Journal* **11,** 1309–1319 (2016).

29. Ricaurte, D. E., Martínez-García, E., Nyerges, Á., *et al.* A standardized workflow for surveying recombinases expands bacterial genome-editing capabilities. *Microb. Biotechnol.* **11,** 176–188 (2018).

30. Venetz, J. E., Medico, L. D., Wölfle, A., *et al.* Chemical synthesis rewriting of a bacterial genome to achieve design flexibility and biological functionality. *Proceedings of the National Academy of Sciences* **116,** 8070–8079 (2019).

31. Singh, T., Yadav, S. K., Vainstein, A., *et al.* Genome recoding strategies to improve cellular properties: mechanisms and advances. *aBIOTECH* **2,** 79–95 (2021).

32. Lau, Y. H., Stirling, F., Kuo, J., *et al.* Large-scale recoding of a bacterial genome by iterative recombineering of synthetic DNA. *Nucleic Acids Research* **45,** 6971–6980 (2017).

33. Richardson, S. M., Mitchell, L. A., Stracquadanio, G., *et al.* Design of a synthetic yeast genome. *Science* **355,** 1040–1044 (2017).

34. Chatterjee, P., Jakimo, N. & Jacobson, J. M. Minimal PAM specificity of a highly similar SpCas9 ortholog. *Science Advances* **4,** eaau0766 (2018).

35. Nishimasu, H., Shi, X., Ishiguro, S., *et al.* Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* **361,** 1259–1262 (2018).

36. Miller, S. M., Wang, T., Randolph, P. B., *et al.* Continuous evolution of SpCas9 variants compatible with non-G PAMs. *Nature Biotechnology* **38,** 471–481 (2020).

37. Walton, R. T., Christie, K. A., Whittaker, M. N., *et al.* Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368,** 290–296 (2020).

38. Christie, K. A. & Kleinstiver, B. P. Making the cut with PAMless CRISPR-Cas enzymes. *Trends in Genetics* **37,** 1053–1055 (2021).

39. Wannier, T. M., Nyerges, A., Kuchwara, H. M., *et al.* Improved bacterial recombineering by parallelized protein discovery. *PNAS* **117,** 13689–13698 (2020).

40. Van Pijkeren, J.-P., Neoh, K. M., Sirias, D., *et al.* Exploring optimization parameters to increase ssDNA recombineering in Lactococcus lactis and Lactobacillus reuteri. *Bioengineered* **3,** 209–217 (2012).

41. Wannier, T. M., Ciaccia, P. N., Ellington, A. D., *et al.* Recombineering and MAGE. *Nat Rev Methods Primers* **1,** 1–24 (2021).

42. Nyerges, Á., Csorgő, B., Nagy, I., *et al.* Conditional DNA repair mutants enable highly precise genome engineering. *Nucleic Acids Research* **42,** e62 (2014).

43. Aparicio, T., Nyerges, A., Nagy, I., *et al.* Mismatch repair hierarchy of Pseudomonas putida revealed by mutagenic ssDNA recombineering of the pyrF gene. *Environmental Microbiology* **22,** 45–58 (2020).

44. Aparicio, T., Nyerges, A., Martínez-García, E., *et al.* High-efficiency multi-site genomic editing of Pseudomonas putida through thermoinducible ssDNA recombineering. *Iscience* **23,** 100946 (2020).

45. Asin-Garcia, E., Martin-Pascual, M., Garcia-Morales, L., *et al.* ReScribe: An Unrestrained Tool Combining Multiplex Recombineering and Minimal-PAM ScCas9 for Genome Recoding Pseudomonas putida. *ACS Synthetic Biology* (2021).

46. Yin, J., Zheng, W., Gao, Y., *et al.* Single-Stranded DNA-Binding Protein and Exogenous RecBCD Inhibitors Enhance Phage-Derived Homologous Recombination in Pseudomonas. *iScience* **14,** 1–14 (2019).

47. Filsinger, G. T., Wannier, T. M., Pedersen, F. B., *et al.* Characterizing the portability of phage-encoded homologous recombination proteins. *Nature Chemical Biology,* 1–9 (2021).

48. Schubert, M. G., Goodman, D. B., Wannier, T. M., *et al.* High-throughput functional variant screens via in vivo production of single-stranded DNA. *Proceedings of the National Academy of Sciences* **118** (2021).

49. Fredens, J., Wang, K., de la Torre, D., *et al.* Total synthesis of Escherichia coli with a recoded genome. *Nature* **569,** 514–518 (2019).

50. Martin-Pascual, M., Batianis, C., Bruinsma, L., *et al.* A navigation guide of synthetic biology tools for Pseudomonas putida. *Biotechnology Advances* **49,** 107732 (2021).

51. Katz, N., Cohen, R., Solomon, O., *et al.* An in Vivo Binding Assay for RNA-Binding Proteins Based on Repression of a Reporter Gene. *ACS Synthetic Biology* **7,** 2765–2774 (2018).

52. Gott, J. M., Wilhelm, L. J. & Uhlenbeck, O. C. RNA binding properties of the coat protein from bacteriophage GA. *Nucleic Acids Research* **19,** 6499–6503 (1991).

53. Peabody, D. S. The RNA binding site of bacteriophage MS2 coat protein. *The EMBO journal* **12,** 595–600 (1993).

54. Lim, F., Spingola, M. & Peabody, D. S. The RNA-binding Site of Bacteriophage Qβ Coat Protein *. *Journal of Biological Chemistry* **271,** 31839–31845 (1996).

55. Lim, F. & Peabody, D. S. RNA recognition site of PP7 coat protein. *Nucleic Acids Research* **30,** 4138–4144 (2002).

56. Patinios, C., Creutzburg, S. C. A., Arifah, A. Q., *et al.* Streamlined CRISPR genome engineering in wild-type bacteria using SIBR-Cas. *Nucleic Acids Research* (gkab893 2021).

57. Calero, P., Volke, D. C., Lowe, P. T., *et al.* A fluoride-responsive genetic circuit enables in vivo biofluorination in engineered Pseudomonas putida. *Nature Communications* **11,** 5045 (2020).

58. Batianis, C., Kozaeva, E., Damalas, S. G., *et al.* An expanded CRISPRi toolbox for tunable control of gene expression in Pseudomonas putida. *Microbial biotechnology* **13,** 368–385 (2020).

59. Noll, P., Treinen, C., Müller, S., *et al.* Evaluating temperature-induced regulation of a ROSE-like RNA-thermometer for heterologous rhamnolipid production in Pseudomonas putida KT2440. *AMB Express* **9,** 154 (2019).

60. Liu, Y., Chen, J., Crisante, D., *et al.* Dynamic Cell Programming with Quorum Sensing-Controlled CRISPRi Circuit. *ACS Synthetic Biology* **9,** 1284–1291 (2020).

61. Lopez, G. & Anderson, J. C. Synthetic Auxotrophs with Ligand-Dependent Essential Genes for a BL21(DE3) Biosafety Strain. *ACS synthetic biology* **4,** 1279–1286 (2015).

62. Tack, D. S., Ellefson, J. W., Thyer, R., *et al.* Addicting diverse bacteria to a noncanonical amino acid. *Nature Chemical Biology* **12,** 138–140 (2016).

63. Whitford, C. M., Dymek, S., Kerkhoff, D., *et al.* Auxotrophy to Xeno-DNA: an exploration of combinatorial mechanisms for a high-fidelity biosafety system for synthetic biology applications. *J Biol Eng* **12,** 13 (2018).

64. Hirota, R., Abe, K., Katsuura, Z.-i., *et al.* A Novel Biocontainment Strategy Makes Bacterial Growth and Survival Dependent on Phosphite. *Scientific Reports* **7,** 44748 (2017).

65. De Lorenzo, V. Beware of metaphors: Chasses and orthogonality in synthetic biology. *Bioengineered Bugs* **2,** 3–7 (2011).

66. Kwon, I., Kirshenbaum, K. & Tirrell, D. A. Breaking the degeneracy of the genetic code. *Journal of the American Chemical Society* **125,** 7512–7513 (2003).

67. Rackham, O. & Chin, J. W. A network of orthogonal ribosome x mRNA pairs. *Nature Chemical Biology* **1,** 159–166 (2005).

68. Wang, K., Neumann, H., Peak-Chew, S. Y., *et al.* Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nature Biotechnology* **25,** 770–777 (2007).

69. Stanton, B. C., Nielsen, A. A. K., Tamsir, A., *et al.* Genomic mining of prokaryotic repressors for orthogonal logic gates. *Nature Chemical Biology* **10,** 99–105 (2014).

70. Jia, B., Qi, H., Li, B.-Z., *et al.* Orthogonal Ribosome Biofirewall. *ACS Synthetic Biology* **6,** 2108–2117 (2017).

71. Noireaux, V. & Libchaber, A. A vesicle bioreactor as a step toward an artificial cell assembly. *Proceedings of the National Academy of Sciences of the United States of America* **101,** 17669–17674 (2004).

72. Hindley, J. W., Elani, Y., McGilvery, C. M., *et al.* Light-triggered enzymatic reactions in nested vesicle reactors. *Nature Communications* **9,** 1093 (2018).

73. Song, J., Baeg, Y., Jeong, H., *et al.* Bacterial Outer Membrane Vesicles as Nano-Scale Bioreactors: A Fatty Acid Conversion Case Study. *ChemCatChem* **13,** 4080–4086 (2021).

74. Laohakunakorn, N., Grasemann, L., Lavickova, B., *et al.* Bottom-Up Construction of Complex Biomolecular Systems With Cell-Free Synthetic Biology. *Frontiers in Bioengineering and Biotechnology* **8,** 213 (2020).

75. Olivi, L., Berger, M., Creyghton, R. N. P., *et al.* Towards a synthetic cell cycle. *Nature Communications* **12,** 4531 (2021).

76. Cai, Y., Agmon, N., Choi, W. J., *et al.* Intrinsic biocontainment: Multiplex genome safeguards combine transcriptional and recombinational control of essential yeast genes. *Proceedings of the National Academy of Sciences* **112,** 1803–1808 (2015).

77. Gallagher, R. R., Patel, J. R., Interiano, A. L., *et al.* Multilayered genetic safeguards limit growth of microorganisms to defined environments. *Nucleic acids research* **43,** 1945–1954 (2015).

78. Wright, O., Delmans, M., Stan, G.-B., *et al.* GeneGuard: A Modular Plasmid System Designed for Biosafety. *ACS Synthetic Biology* **4,** 307–316 (2015).

79. Arnolds, K. L., Dahlin, L. R., Ding, L., *et al.* Biotechnology for secure biocontainment designs in an emerging bioeconomy. *Current Opinion in Biotechnology. Analytical Biotechnology* **71,** 25–31 (2021).

80. Zuñiga, C., Li, T., Guarnieri, M. T., *et al.* Synthetic microbial communities of heterotrophs and phototrophs facilitate sustainable growth. *Nature Communications* **11,** 3803 (2020).

81. Chao, R., Mishra, S., Si, T., *et al.* Engineering biological systems using automated biofoundries. *Metab Eng* **42,** 98–108 (2017).

82. Al-Bassam, M. M., Kim, J.-N., Zaramela, L. S., *et al.* Optimization of carbon and energy utilization through differential translational efficiency. *Nature Communications* **9,** 4474 (2018).

83. Tellechea-Luzardo, J., Winterhalter, C., Widera, P., *et al.* Linking Engineered Cells to Their Digital Twins: A Version Control System for Strain Engineering. *ACS Synthetic Biology* **9,** 536–545 (2020).

84. Zuñiga, C., Zaramela, L. & Zengler, K. Elucidation of complexity and prediction of interactions in microbial communities. *Microbial Biotechnology* **10,** 1500–1522 (2017).

85. Zuñiga, C., Li, C.-T., Yu, G., *et al.* Environmental stimuli drive a transition from cooperation to competition in synthetic phototrophic communities. *Nature Microbiology* **4,** 2184–2191 (2019).

86. Zengler, K. & Zaramela, L. S. The social network of microorganisms — how auxotrophies shape complex communities. *Nature Reviews Microbiology* **16,** 383–390 (2018).

87. Lee, Y., Lafontaine Rivera, J. G. & Liao, J. C. Ensemble Modeling for Robustness Analysis in engineering non-native metabolic pathways. *Metabolic Engineering* **25,** 63–71 (2014).

88. Wu, C., Jiang, H., Kalra, I., *et al.* A generalized computational framework to streamline thermodynamics and kinetics analysis of metabolic pathways. *Metabolic Engineering* **57,** 140–150 (2020).

89. Daly, H. E. *Beyond growth: the economics of sustainable development* Nachdr. (Boston, Mass, 1996).

90. Hedlund-de Witt, A. Exploring worldviews and their relationships to sustainable lifestyles: Towards a new conceptual and methodological approach. *Ecological Economics. The Economics of Degrowth* **84,** 74–83 (2012).

91. Spangenberg, J. H. The world we see shapes the world we create: how the underlying worldviews lead to different recommendations from environmental and ecological economics - the green economy example. *International Journal of Sustainable Development* **19,** 127–146 (2016).

92. Wang, F. & Zhang, W. Synthetic biology: Recent progress, biosafety and biosecurity concerns, and possible solutions. *Journal of Biosafety and Biosecurity* **1,** 22–30 (2019).

93. From the Commission to the European Parliament, C., the Council, the European Economic and Social Committee, *et al. The European Green Deal* (2019).

94. Harrison, T. J. *Of bridges: a poetic and philosophical account* (Chicago, 2021).

95. IPCC. *Special Report: Global Warming of 1.5°C* https://www.ipcc.ch/sr15/ (2021).

96. UN Climate Change Conference UK 2021. *COP26 The Glasgow Climate Pact* https://ukcop26.org/the-conference/cop26-outcomes/ (2021).

97. Devos, Y., Sanvido, O., Tait, J., *et al.* Towards a more open debate about values in decision-making on agricultural biotechnology. *Transgenic Research* **23,** 933–943 (2014).

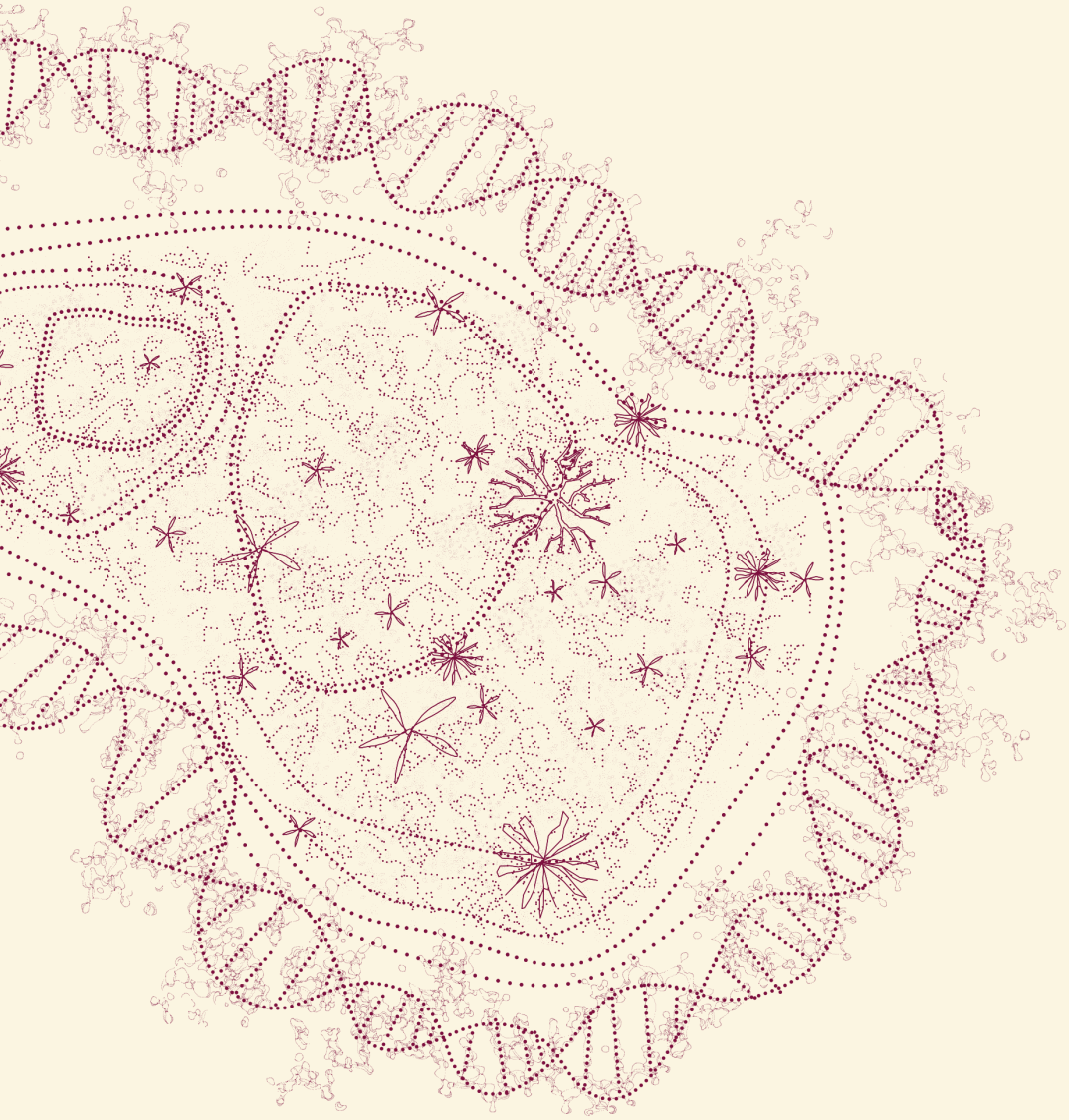
98. Wardy, R. Heraclitus' Symposium. *Revue de philosophie ancienne,* 97–139 (2016).

99. Mampuys, R. & Brom, F. Emerging crossover technologies: How to organize a biotechnology that becomes mainstream? *Environ Syst Decis* **38,** 163–169 (2018).

100. Vickers, G. *The art of judgment: A study of policy making* (1965).

101. National Academy of Sciences (US). Committee on the Conduct of Science. *On being a scientist* (1989).

102. Bosch, G. Train PhD students to be thinkers not just specialists. *Nature* **554,** 277–277 (2018).



Appendices

Enrique Asin-Garcia

Proverbios y cantares (XXIX)

Caminante, son tus huellas
el camino y nada más;
caminante, no hay camino,
se hace camino al andar.
Al andar se hace el camino,
y al volver la vista atrás
se ve la senda que nunca
se ha de volver a pisar.
Caminante no hay camino
sino estelas en la mar.

Antonio Machado

About the author

Enrique Asín García was born on the 23rd of January 1993 in Leganés (Madrid, Spain), but spent his childhood in the small medieval village of San Felices de los Gallegos (Salamanca, Spain). Among the western meadows, the rough hills and the pastureland, he fell in love with nature at a very young age. Always curious and non-conformist, he received a microscope as a present from his grandfather when he was only 9, which set him out on the quest of looking beyond what the bare eye can see.

Fascinated by the harmonious complexity hidden in natural processes and the possibility of altering those principles to solve global problems, Enrique studied Biotechnology at the University of Salamanca (Salamanca, Spain) from which he obtained his BSc in 2015. During this time, he learnt his way into the laboratory during his BSc thesis in vaccine development at the Laboratory of Parasitic and Molecular Immunology (Research Centre for Tropical Diseases of University of Salamanca, Salamanca, Spain) and his internship in biofuel generation at the Laboratory of Adaptive Mechanisms for Plant Biotechnology and Energy Crops (Institute of Molecular and Cellular Biology of Plants, Valencia, Spain). During these years, he acted as student representative of the Bachelor Biotechnology and became treasurer and part of the Executive Board of the Association of Biotechnology of Salamanca. In 2015, he left his home country to expand his horizons studying Cellular and Molecular Biotechnology at Wageningen University & Research (Wageningen, The Netherlands), from which he obtained his MSc degree. As part of his masters, he conducted two MSc theses, one in yeast and filamentous fungi bioengineering at the Laboratory of Systems and Synthetic Biology and one in reverse genetics of virus in insect cells at the Laboratory of Virology, both part of Wageningen University.

Deeply interested in a research career, Enrique started his PhD research project in September 2017 back at the Laboratory of Systems and Synthetic Biology of Wageningen University under the supervision and mentorship of Professor Vitor Martins dos Santos. By means of synthetic biology tools, genome recoding and metabolic engineering, his work has been the main foundation of the NWO-funded project "SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engi-

neering of a Versatile Industrial Chassis", focused on biosafety tools for safe-by-design industrial bioprocesses. Throughout his PhD candidacy, he also engaged in broader discussions related to social, ethical and philosophical aspects of his project, which brought him close to the research of Dr. Zoë Robaey, who eventually became his co-promotor and a fundamental piece of the journey.

At the beginning of 2022, Enrique founded SynBioNL, the Synthetic Biology Association of The Netherlands, with the position of treasurer as part of the Executive Board. Currently, he is employed as a PostDoc researcher at the Laboratory of Systems and Synthetic Biology of Wageningen University where he is focused on high-throughput genome and metabolic engineering approaches for biomanufacturing and safety-by-design.

Co-author affiliations

✉ Laboratory of Systems and Synthetic Biology - Wageningen University & Research

Wageningen, 6708 WE, The Netherlands

Maria Martin-Pascual, Luis Garcia-Morales, Vitor A. P. Martins dos Santos[a], Tessa Bartholet, Claudia de Buck, Max Allewijn, Alexandra Müller[a], Christos Batianis, Yunsong Li, James D. Fawcett[b], Ivar de Jong[c], Hetty Huijs, Linde F. C. Kampers, Peter J. Schaap

✉ LifeGlimmer GmbH

Berlin, 12163, Germany

Vitor A. P. Martins dos Santos

✉ Laboratory of Microbiology - Wageningen University & Research

Wageningen, 6703 HB, The Netherlands

Richard van Kranenburg

✉ Department of Molecular, Cellular, and Developmental Biology - Yale University

New Haven, CT 06520, USA

Zhuobin Liang[d], Farren Isaacs

✉ Systems Biology Institute - Yale University

West Haven, CT 06516, USA

Zhuobin Liang[d], Farren Isaacs

✉ Institute for Science in Society - Radboud University

Nijmegen, 6500 GL, The Netherlands

Amalia Kallergi[e], Laurens Landeweerd

✉ Department of Social Sciences - Wageningen University & Research

Wageningen, 6708 WE, The Netherlands

Hetty Huijs, Zoë Robaey, Annemarie Wagemakers

Current addresses

a. Bioprocess Engineering Group - Wageningen University & Research

Wageningen, 6700 AA, The Netherlands

b. Department of Life Sciences - Imperial College London

South Kensington, London, SW72BX, United Kingdom

c. The Novo Nordisk Foundation Center for Biosustainability - Technical University of Denmark

Kgs. Lyngby, 2800, Denmark

d. Institute of Molecular Physiology - Shenzhen Bay Laboratory

Shenzhen, 518132, China

e. Department of Biotechnology - Delft University of Technology

Delft, 2629 HZ, The Netherlands

List of publications

- ✎ T. Laothanachareon, **E. Asin-Garcia**, R. J. M. Volkers, J. A. Tamayo-Ramos, V. A. P. Martins dos Santos and P. J. Schaap. Identification of *Aspergillus niger* aquaporins involved in hydrogen peroxide signaling. *(Manuscript submitted for publication)*

- ✎ **E. Asin-Garcia**, L. Garcia-Morales, T. Bartholet, Z. Liang, F. Isaacs and V. A. P. Martins dos Santos. Metagenomics harvested genus-specific single-stranded DNA annealing proteins improve and expand recombineering in *Pseudomonas* species. *(Manuscript submitted for publication)*

- ✎ H. Huijs*, **E. Asin-Garcia***, Z. Robaey and V. A. P. Martins dos Santos. Towards an attitude for responsibility in safety education. *(Manuscript submitted for publication)*

- ✎ **E. Asin-Garcia**, C. Batianis, Y. Li, J. D. Fawcett, I. de Jong and V. A. P. Martins dos Santos. Phosphite synthetic auxotrophy as an effective biocontainment strategy for the industrial chassis *Pseudomonas putida*. *(Manuscript submitted for publication)*

- ✎ J. D. Fawcett*, **E. Asin-Garcia*** and V. A. P. Martins dos Santos. A snapshot of biomanufacturing and the need for enabling research infrastructure. *(Manuscript submitted for publication)*

- ✎ **E. Asin-Garcia**, L. F. C. Kampers, Z. Robaey and V. A. P. Martins dos Santos. Exploring the impact of tensions in stakeholder norms on designing for value change: the case of biosafety in industrial biotechnology. *(Manuscript submitted for publication)*

- ✎ L. F. C. Kampers, **E. Asin-Garcia**, P. J. Schaap, A. Wagemakers and V. A. P. Martins dos Santos. Navigating the Valley of Death: Perceptions of Industry and Academia on Production Platforms and Opportunities. *(Manuscript submitted for publication)*

- ✎ **E. Asin-Garcia***, M. Martin-Pascual*, L. Garcia-Morales, R. van Kranenburg and V. A. P. Martins dos Santos. ReScribe: an unrestrained tool com-

bining multiplex recombineering and minimal-PAM ScCas9 for genome recoding *Pseudomonas putida*. ACS Synthetic Biology 10, 2672–2688 (2021) doi.org/10.1021/acssynbio.1c00297

✎ L. F. C. Kampers*, **E. Asin-Garcia***, P. J. Schaap, A. Wagemakers and V. A. P. Martins dos Santos. Bridging Academia and Industry in the Biotechnology Valley of Death. Trends in Biotechnology, Science & Society, 39, 1240-1242 (2021) doi.org/10.1016/j.tibtech.2021.04.010.

✎ M. Martin-Pascual*, C. Batianis*, L. Bruinsma*, **E. Asin-Garcia**, L. Garcia-Morales, R. A. Weusthuis, R. van Kranenburg and V. A. P. Martins dos Santos. A navigation guide of synthetic biology tools for *Pseudomonas putida*. Biotechnology Advances 49, 107732 (2021) doi.org/10.1016/j.biotechadv.2021.107732

✎ A. Kallergi*, **E. Asin-Garcia***, V. A. P. Martins dos Santos and L. Landeweerd. Context matters: On the road to responsible biosafety technologies in synthetic biology. EMBO reports, Science & Society, 22, e51227 (2021) doi.org/10.15252/embr.202051227.

✎ **E. Asin-Garcia***, A. Kallergi*, L. Landeweerd and V. A. P. Martins dos Santos. Genetic safeguards for Safety-by-design: so close yet so far. Trends in Biotechnology, Science & Society, 38, 1308-1312 (2020) doi.org/10.1016/j.tibtech.2020.04.005.

* denotes equal contribution

Overview of completed training activities

Discipline specific activities	Organised by	Year
5th EmPowerPutida GA	EPP consortium	2017
6th EmpowerPutida GA	EPP and P4SB consortia	2018
7th EmPowerPutida GA	EPP consortium	2018
8th EmPowerPutida GA	EPP consortium	2019
Biotechnology and Safety Programme Kick Off	NWO	2018
Biotechnology and Safety Programme Impact Pathways	NWO	2019
Biotechnology and Safety Programme Day	NWO	2021
Safety-by-design as part of RRI in practice Workshop	WUR and RIKILT	2018
Lost in Tranlation - Analysing SynBio Workshop	WUR and Radboud University	2018
The Lively Sciences Workshop	iArts Maastricht and Radboud University	2019
Autonomous Bio-Machines! BioArt Programme	Creative Innovation WUR	2018
ynBio @ WUR Impulse Seminar	WUR	2018
T-TRIPP Kick off	Rathenau Institute and TU Delft	2018
T-TRIPP Stakeholder Workshop	Rathenau Institute and TU Delft	2018
T-TRIPP MachiaCELLi+ Workshop	Rathenau Institute, TU Delft and WUR	2019
Engineered safeguards to limit environmental effects of applications of synthetic biology Workshop	SafeChassis consortium	2019
Seminar for the Molecular Environmental Microbiology Laboratory - CNB	CNB	2020
T-TRIPP Engineering plants and the rhizosphere Workshop	TU Delft	2021
Conversations about Metabolic Engineering Seminar	SynBio4Flav	2021
Genome Engineering and Synthetic Biology Conference (3rd Edition)	VIB	2018
Pseudomonas Grassroots Symposium (1st Edition)	Dechema-Haus, iAMB, RWTH, Aachen University	2018
CRISPRcon Conference	Keystone Policy Center	2019
Applied Synthetic Biology in Europe Conference V	European Federation of Biotechnology	2020
German Conference of Synthetic Biology2021	German Association of Synthetic Biology	2021
4TU.Ethics Bi-annual Conference on Ethics and Technology IT'S ALIVE!	4tu Centre for Ethics and Technology	2021
MB7.0 (Microbial Biotechnology Meeting)	Board of Microbial Biotechnology	2021
General courses	**Organised by**	**Year**
VLAG PhD week (40th Edition)	VLAG	2018
Brain-friendly Working and Writing	WGS	2018
Philosophy and ethics of Food Science and Technology	WGS	2019
Making Impact: Increase the relevance of your Research through science-society interaction	WGS	2019
The Choice: Un-box your PhD process and take charge of your performance	WGS	2019
Career Orientation	WGS	2020
Assisting in teaching and supervision activities		**Year**
Courses SSB-31312 Toolbox in Systems and Synthetic Biology, SSB-20806 iBiosystems, and SSB-50806 Systems and Synthetic Biology	WUR - SSB	2018-2020
Supervision of BSc and MSc students	WUR - SSB	2018-2021
Other activities	**Organised by**	**Year**
Preparation of research proposal	SSB	2017

PhD study tour to Boston and New York	SSB and MIB	2019
Organization of PhD study tour to Boston and New York	SSB and MIB	2018-2019
Weekly and montly SSB group meetings	SSB	2017-2021
SSB Journal club	SSB	2021
iGEM meetings and supervision	SSB and MIB	2018-2019
Co-founder and treasurer of SynBioNL (Synthetic Biology Association from The Netherlands)	Independent	2021
Interview for Daniele Maccari for "Metaphors we know by: An investigation in Synthetic Biology"	Radboud University	2019
Interview for Daan Schuurbiers for "Over biotechnologie en veiligheid: Verslag van een serie interviews met onderzoekers binnen het NWO-TTW programma Biotechnologie en Veiligheid"	Ministerie van Infrastructuur en Waterstaat (IenW)	2020
Interview for Mario Juárez Rodríguez for the SEVA platform blog	D3veloperSCS_SEVA, SEVA platform	2021

Acknowledgements

UBUNTU

which in the Xhosa culture means "I am because we are"

Independence seems to be one of the most appreciated and valuable personal qualities of our modern western societies. However, the reality is that humans are interdependent beings: all of us need support. This is true for everyone, even though the degree of interdependence might vary during the different stages of our lives and according to personal circumstances. During the last four and a half years, I have depended on many people to do the work collected in this book, not only in terms of scientific counsel and resources, but also because many times I have been fragile and vulnerable. All things considered, these feelings make me only human, but the fact that I could always count on someone makes me a very lucky one.

I am afraid that these acknowledgements will inevitably fall short of the actual amounts of gratitude that I have inside, but nothing gives me more joy than making you all part of this story by giving you a glimpse of how important your support has been throughout this process.

First and foremost, my gratitude goes to **Vitor**, my promotor. You have always encouraged me to become my best self by procuring me with opportunities, connections and advice to develop, to grow and to set my own path. You believe in me more than I do sometimes, and this reassurance and your characteristic optimism have undoubtedly been one of the main engines of this PhD adventure. I would not be the professional that I am today without the countless conversations we have shared, not only about projects, papers and collaborations, but also about how to confront and overcome the obstacles of the scientific career and life in general. You have been a truly valuable mentor and I am very thankful to you for always being there for me.

Zoë, you entered my life coincidentally, and how beautiful and enriching coincidence that was. Eventually, you would become my co-promotor, but we clicked instantly and built together a fruitful collaboration around the concept of biosafety ethics that expanded my horizons to a whole new level. You have taught me a lot,

professionally and personally, and your sensibility and empathy have been fundamental for my well-being during the last stages of this process. Thank you, above all, for caring for the person behind the researcher and for giving me energy to continue pursuing my life goals.

I would also like to take this opportunity to thank the **members of the thesis committee** for your time and critical assessment of my thesis.

Beside my supervisors, this thesis could have never been achieved without three people, my work pillars, who have worked relentlessly for the success of our team and have backed me up throughout this entire journey: **María**, **Christos** and **Lyon**.

María, no one without the experience can really fathom the perks of working alongside a long lasting friend. Sometimes, I think not even we realized how fortunate we were for spending our days together and having a loved one near. But that was only sometimes. At this point, I am completely aware of how valuable is to have the possibility of working and collaborating in an environment of complete trust, reliance, support, encouragement and mutual learning. Our friendship is the result of a relationship cultivated over the course of many years, in hundreds of conversations and experiences that go beyond the contents of this book. You will always be a nostalgic memento of an old life and the spark of my present days. You are like a sister to me and you will always feel like home.

Christos, my PhD brother. What a crazy adventure we have lived together! I have witnessed many changes during my PhD candidacy, and the only constant I can recall is you. You have always been there. You have acted as a daily supervisor, as a reviewer, as a consultant, as a counselor... and even though the road has not always been easy for us, I feel extremely proud of what we have achieved together. All this time you were simply doing what a true friend would have done, and I am grateful for all the mutual feedback, the side-by-side development, all the fun and all the kindness you have shown to me. Undoubtedly, you have been my rock during all these years and my future scientific career is already missing that brilliant mind of yours.

Lyon, I feel so fortunate for having you in my daily life. Your special personality is a treasure with the power of cheering anyone up in the darkest moments. Your optimism has helped me putting problems in perspective many days and has

brightened many early mornings at the lab with music, Spanish greetings and the widest of smiles. Your goofiness and craziness make of you the best company in gatherings, parties and long working days, and even though I often wonder what really goes on in that mind of yours full with enzymes, metabolic pathways, useless trivia and weird dreams, I always enjoy our conversations. They always end up turning into unpredictable, wonderful directions, with huge doses of laughter insured.

My gratitude goes also to all the colleagues of the Systems and Synthetic Biology chair group. Throughout all these years, I have met many SSB members, both staff and students, and I would like to express my acknowledgements for making this time enjoyable and for helping me whenever I needed it. **Luis**, thank you for the time you have dedicated to support me, my students and the lab in general during this last couple of years. Your help with the third chapter of this thesis, especially with the NGS experiments, as well as with the use of cre-lox site-specific recombineering has been invaluable. **Tom**, we owe much of the smooth way things run at the lab to you. It has been a treat to do all my experiments in such an organized and tidy environment, and to have had you there by my side whenever I needed technical advice. In this sense, I would also like to extend my gratitude to **Willemijn** and **Bart** for making my PhD journey a little bit easier. I really appreciate all the help, bureaucratic and computational, that I have received from both you. I am also thankful to the staff members of the group: to **Peter**, for being the one who first saw something in me and brought me back to SSB to start this PhD; to **María**, for all the valuable thesis and Latex tips and every other piece of advice; to **Edoardo**, for letting me instruct my dear Toolbox course with you for a couple of years; to **Jasper**, for everything I learnt during the adventure into your project about domains of unknown function; to **Cristina**, for your support with the analytics and the students at the lab and thesis ring; and to **Rob**, also for your support with the students in various aspects (thesis ring, iGEM, etc.), but also for your help with the ongoing riboregulator project.

Sara, **Rik**, **Marco** and **Chunzhe**, one of my favourite parts of the project has definitely been the opportunity to share together our ideas, results and advice. I am very proud of our *P. putida* team and I have learnt tons about our favourite microbe

from each one of you. To the rest of my SSB colleagues, old and new, **Anna**, **Anna**, **Bastian**, **Benoit**, **Dorett**, **Efsun**, **Emma**, **Erika**, **Henk**, **Marta**, **Melanie**, **Nhung**, **Niels**, **Nicolas**, **Niru**, **Rita**, **Ruben**, **Sabine**, **Sanjee**, **Sara**, **Stamatios**, and **Wasin**, thank you for the lessons together and for all the moments out of the lab and the office, the Christmas events, the PhD trips, the lab trips, the barbeques, the cakes on the sixth floor, the drinks after the seminars and all the other activities. I am happy to have met all of you!

But this round through SSB could not finish without two people I have shared much with: **Linde** and **Nong**. You were very important for me during the first years of the PhD and I want to let you know that I treasure the long conversations at the office, the common projects, the road trips with Linde to conduct interviews, the organization of the PhD trip with Nong, the dinners, the coffee breaks and the thousand moments more that we experienced together. I am truly grateful for all of them!

The work of this book could have not been possible without the help of many talented BSc, MSc, internship and Toolbox students. I would like to express my thanks to: **Karsten**, for all the things I learnt about supervising a student with you and for your work with the riboregulators; to **Katrin**, **Nafisha**, **Rosyida** and **Timo**, for making such a great team and setting up with me the MASC-PCR technique that so important would become for the success of chapter 2; to **Alexandra**, for juggling three projects at once well, for our learning process together and for staying in my life all this time; to **Ivar**, for your great efforts and for laying the basis for the work collected in the fifth chapter of this thesis; to **Kesi** and **Sanne**, for researching the genome integration techniques that so important would become in the next years of research in our lab; to **Alba** and **Ben**, for making even greater the iGEM experience, for your resilience and for trusting your supervisors at all times; to **Carmen**, **Kimberly** and **Lorenzo**, for your work with ScCas9 and PapRecT, which brought María and me to the findings of chapter 2; to **James**, for your dedication, your motivation and two fantastic pieces of work that have become major contributions to chapters 5 and 11 of this thesis; to **Max**, for dealing so well with a bachelor thesis during the uncertain times of a starting pandemic and for your work on chapter 4; to **Hetty**, for your strength working on your thesis during difficult times, for your resourcefulness

and for the work that has materialized in chapter 9; to **Tessa**, for your hard work and determination, for the new ideas and for the protocols that we set up together to achieve the results collected in chapter 3; to **Yunsong**, for your cheerfulness, your enthusiasm and your eagerness to learn new things every single day, which led you to the completion of chapter 5; and to **Claudia**, for your independence, for your reasoning and critical-thinking and for supporting me during the writing of this thesis contributing to chapter 4 and expanding its scope to exciting new avenues.

As part of the SafeChassis project, I had the honour to work with various experts from diverse research fields. **Farren** and **Zhuobin**, thank you for providing me with the bioinformatics-metagenomics material to pursue the project collected in chapter 3. Your knowledge and insights in recombineering, MAGE and genetic safeguards have been essential for the fulfillment of this thesis. **Ken**, thank you for the lessons about biosafety considerations and for setting the stage for multi-stakeholder discussions on the topic. **Laurens**, thank you very much for bringing me closer to philosophy and art through your stimulating workshops. I have enjoyed that part of SafeChassis very much and I am very thankful for our collaboration together. **Amalia**, I cannot thank you enough for all the things you taught me about writing and thinking. I am so proud for the beautiful things we created together (chapters 6 and 7 of this thesis) and I will always be grateful for the enriching process we shared together. In addition, I would like to thank **Cécile**, **Kyra**, **Lilian**, **Marie-Louise**, **Markus**, **Michelle**, **Titia**, **Virgil**, plus the rest of colleagues from the T-TRIPP project, for the discussions about my work. Especial mention goes to **Richard**, who has always been helpful beyond the user committee meetings, and played a fundamental work in the development of chapter 2.

In the same way, I feel extremely lucky for having been part of the EmPower-Putida project. **Andrea**, **Bernhard**, **Cecília**, **Ian**, **Mechthild**, **Ralf**, **Sven** and all the PhD candidates and PostDocs that were part of the consortium, thank you very much for all the inputs and the amazing time we had during the general assembly meetings. Prominent amongst these contributions have been the support, generosity and encouragement from **Víctor** and his team, especially **Tomás** and **Esteban**, who have shed light more than once on the intrepid world of *P. putida*.

And talking about organizations, I would like to thank my co-organizers of the

PhD trip 2019 to the Boston and New York areas, **Caifang**, **Catarina**, **Costas**, **Giannis**, **Ivette**, **Lot**, **Nong** and **Ran**, for setting up one of the best experiences of this PhD. It was enriching, it was productive and it was crazy fun! Thank you also to **Diana** and **Raymond** for being there for and with us!

Luckily, there have been many more moments of fun during these last years. I started this journey together with many of the friends I made during my masters, and we have had a wonderful time surviving the PhD. **Belén**, **Carrie**, **Janneke**, **Jannie**, **Joep**, **Ivette**, **Max**, **Patrick**, **Sharon** and **Thijs**, thank you for all the good memories of those beginnings with the Thursdays Friesdays, the lunches at the canteen, the dinners, the drinks and all the lessons we learnt together during that time. **Carrie** and **Jannie**, thank you for all the necessary talks over coffee or dinner. With you two I could always talk about my feelings, fears and life prospects, and the way you have always supported has shown me how caring and especial you are for me, and how vital that has been for my mental health. **Ivette**, you have been there since the moment I set foot in this university, and I am not sure if it was that or our matching personalities, but something makes our connection one-of-a-kind. I am always happy when I am with you, either dancing in the IC, chatting at Helix or getting lost in NYC. I will always feel fortunate for that serendipitous day when I sat next to you in that lecture of Microbial Ecology. **Belén**, I also met you in the early days at Wageningen, but our friendship intensified during the last years. You are a good and caring soul, always kind and willing to help. Thank you for listening to me in the good and in the bad moments, and for knowing what to say to make me feel better. Later on, I got to know **Despoina**, **Fotis**, **Jeroen**, **Lorenzo** and **Pilar**. Thank you, guys, for bringing me out of the routine with nice food and fun summer trips. I thoroughly enjoy our time together and look forward to the travels and plans that the future is holding for us.

My personal development (and sanity) during these last years has been also influenced by a bunch of friends, old and not that old, in The Netherlands and in Spain.

Anton and **Judit**, the beers, picnics, dinners and calçots we share are never enough, but they have been sufficient for me to consider you great friends. Thanks for the good vibes I feel when I spend time with you, which fill me with energy and

positivity for the whole week. **Jeroen**, amigo mío, you are always a reminder of the old Wageningen: Haarweg, the IC and those times when we used to know everyone in town. I am thankful that you remain in my life. I love our catch-ups and reliving the old stories over some beers. In the same vein, I want to thank **Ana Belén**, **Fausto** and **Virginia**. Even though it has been many years and even though we have created a lot of new memories, you will always be my little piece of Salamanca in this country. Somehow you keep me connected to a part of my life I do not ever want to leave behind, and I am very glad we undertook this literal journey together and continued building this beautiful friendship in a new country.

Alejandro, **Andrea**, **Curro** and **Javi**, ours has always been a special friendship, built on the distance but on a strong foundation. Time has proven that no matter how far we are, we will always be there for each other in the moments that count. Thank you for being, against all odds, such a wonderful and meaningful part of my life. Especial appreciation goes, of course, to **Alejandro**, the designer of this thesis' cover, for sharing your reflections and your art with me during all this process.

Adrián, **Adrián**, **África**, **Alberto**, **Ángel**, **Cristina**, **Daniel**, **David**, **Diego**, **Gerardo**, **Gloria**, **Héctor**, **Hermes**, **Iker**, **Irene**, **Javier**, **José Ángel**, **Juan Carlos**, **Laura**, **Leticia**, **Miriam**, **Rocío**, **Sara** and **Vega**, you have been there for me since I can remember. We have gone through everything in this life together and I cannot conceive a purest bond than the one we share. Without a doubt, the most difficult part of this PhD has been to endure a life so far away from you. Fortunately, you are daily on the other side of the screen and always present in the happiest of my adventures. **Raúl**, you taught me the hardest of lessons and, at the same time, the most beautiful: life might finish at any point, so live and love always to the fullest*. Nunca dejaré de agradeceros la vida que me dais.

Daniel, you are my greatest friend from Wageningen and your contribution to the completion of this book has been invaluable. The list of things I need to thank you for is endless, but I believe that everything is well summarized by simply referring to your enormous heart. You are a generous, emphatic human being, with the capacity of taking my frustration and sorrows and baking them away into a delicious cake to cheer me up. I am very happy that you stayed here all those years ago, since I am very proud of how much we have grown together. You and **María** are our family

in this country, and that comforting feeling has been critical to fight the nostalgia and the loneliness one feels when is so far away from his beloved ones. Because of you, this place started long ago to feel like home.

A mi familia, mis abuelos, mis tíos y primos. Millones de gracias por creer siempre en mí. Vuestro apoyo, vuestra confianza y vuestro cariño siempre me ha motivado a perseguir mis metas. Valoro mucho cada momento que pasamos en familia, quizá porque no son tantos como me gustaría. Sé de sobra cuánto celebráis mis éxitos y cuánto me animáis cuando las cosas no van tan bien. Solo deciros que cuando estoy rodeado por todos vosotros, los problemas parecen menos problemas, y las victorias saben doblemente mejor. **Abuela** y **abuelo**, gracias por cuidarme tanto, por facilitarme el camino, y por todos esos consejos que siempre llevaré en el corazón.

Mamá, **papá**, vosotros me habéis enseñado que no importa cuántas veces uno tropiece por el camino, que lo que importa es levantarse las veces que haga falta para seguir adelante. La nuestra siempre ha sido una historia de superación, de trabajo duro y de constancia. Pero el hecho de que hoy esté yo aquí escribiendo estas líneas es el fruto de vuestro enorme esfuerzo y de muchos sacrificios, no solo por haberme brindado todas las posibilidades en este mundo, sino por haberme apoyado en todas esas decisiones que inexorablemente me alejaban de casa. De vosotros. Gracias por vuestro apoyo incondicional, por la vida que me habéis dado y por el amor que me habéis transmitido por la naturaleza, por los libros, por los amigos y por las cosas bien hechas. Sabéis bien que nunca me conformo con nada, pero si tuviera que conformarme con algo, sería con volver siempre a nuestro rincón de San Felices, donde el aire huele a chimenea, a jara y a olivo, y la vida pasa a ritmo de poema.

Yolanda, gracias por ser corazón, y por ser coraza. Por ser refugio, y a la vez ser selva. Gracias por ser aquello de lo que más orgulloso estoy en mi vida. Te agradezco el hacerme fuerte cada día, por todas esas veces que tengo que serlo por los dos. Te agradezco el hacerme vulnerable, por todas esas veces en las que la fuerte has tenido que ser tú. Oír tu voz a través del móvil en todas esas llamadas inesperadas me ha sanado en momentos de amargura y frustración, y esas risas que nos echamos, eso, no lo cambio por nada. Gracias por tus palabras en nuestro

lenguaje secreto y por todas esas que se dicen sin hablar. Por ser mi roca, y por ser mi luz, gracias.

This PhD has been a beautiful but hard journey. I have been stressed out most of the time, hurrying to the next experiment, to the next meeting, to the next article, etc. It was then when **Blue** came into my life. One of the early days after she arrived, walking with her after work, but still absent, hurrying off to get back home to continue, and listing in my mind all the tasks I still had left to do, I felt she would not continue walking. When I looked back, I found her still, smelling a little flower. It was then when I understood what life was really about. Thank you, caracol, for teaching me that nothing is as important and deserves more attention than those things that are dear to us and make us happy.

And the one that makes me the happiest is **Yahya**; the happiest, the craziest and definitely the strongest. When the weight of the PhD was too much, you were there, unable to lift the PhD burden, but willing to lift me up instead. You are the spark of my life, filling my days with excitement, laughter and joy. You believe in me much more than I often do myself, and this faith has been the fuel of many of the beautiful things collected in this book. I found in you a kindred spirit, the purest of souls and the most heroic support, given all the ups and downs that this process has meant to you too. Thank you for being the music of this adventure, sometimes literally and sometimes figuratively. It has made it seem like one of those great stories, the ones that really matter, in which the bad moments pass and the new day's sun shines out the brightest. I am glad and thankful for you to be with me, here, at the end of this chapter and at the beginning of the rest of things.

About the cover

The cover of this thesis invites the reader to get immersed in a micro-geography generated by the superposition of two scales of perception, one microscopic and one territorial; a cartography that aims to transfer, through graphic representation, an environment of permissive and non-permissive living conditions, a genetic firework and an infinity of potential possibilities within it.

Starting from a scientific premise, the limits between technology and philosophy, between a tool and its multiple uses, are blurred by issuing a series of more or less precise regions where everything could happen, or where nothing should happen. Almost as in a portolan map or in a nautical chart, the place where the different crossings converge is the main object of the surface: a specific region, delimited by a genetic safeguard, and distinguished by few granted conditions that allow the development of certain events, places or intentions.

The way this book presents itself arises not only from interpretation, but also from the suggestiveness of and endless number of horizons for biotechnology and the knowledge it contains, such as the endless trajectories that a traveller could draw on a map.

Cover art by
Alejandro Infantes

Colophon

The research described in this thesis was financially supported by the NWO project SafeChassis: Implementing and Assessing Safeguards for Lifestyle Engineering of a Versatile Industrial Chassis (Project reference: 15814).

Printed by: ProefschriftMaken.nl

1µm
2µm
3µm
4µm
5µm
6µm
7µm
8µm
9µm
10µm
11µm
12µm
13µm
23°887.336"
72°55'2.947"
62°98'17.374"
26°997.233"
87°12'9.142"
60°11'4.375"
83°26'47.099"
69°18.154"