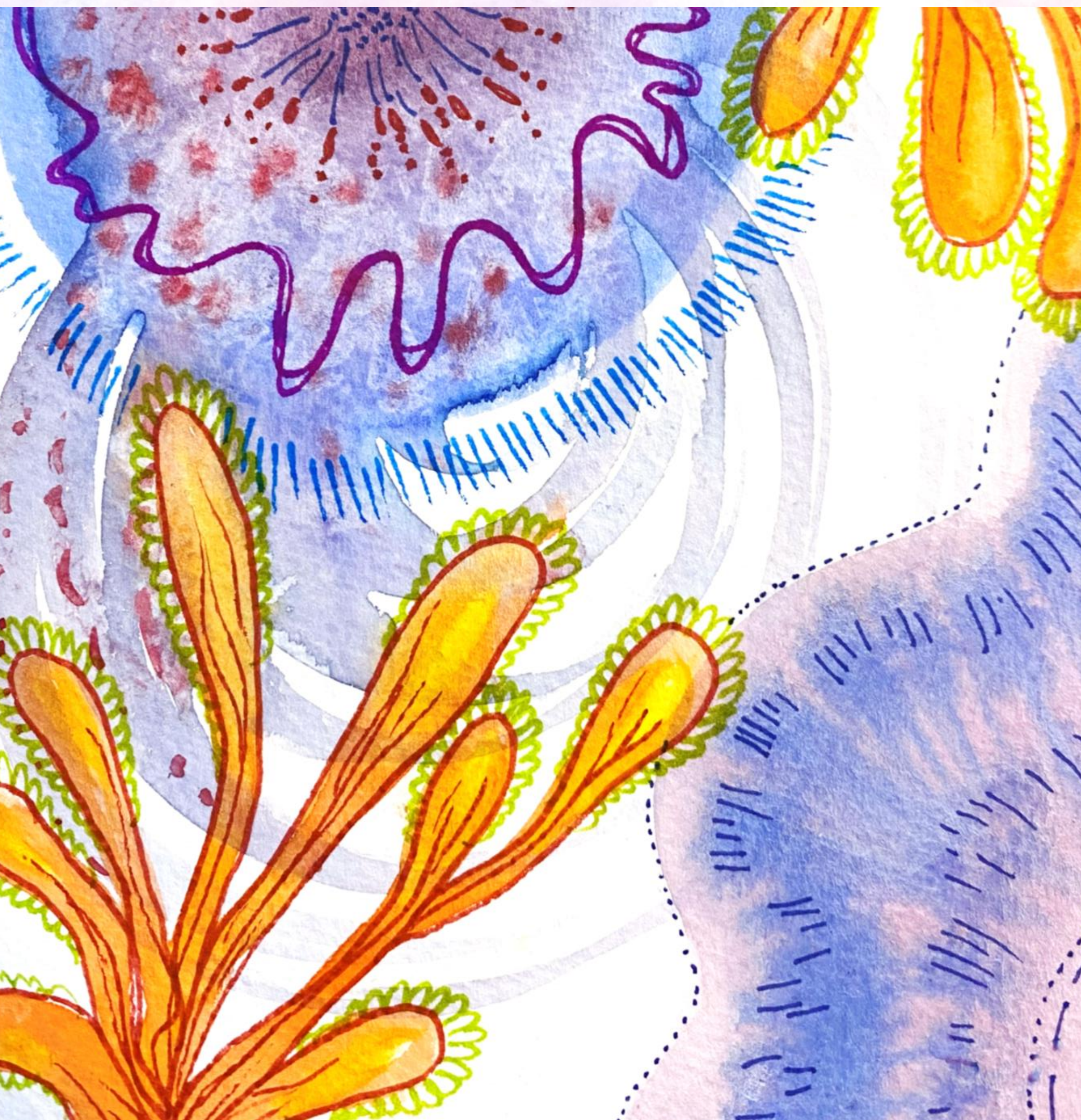
Ecology & evolution of microbial community assembly



Alanna Molly Leale

Propositions:

1. Stability is a main contributor to community function.
   (this thesis)

2. 16S rRNA sequencing is not suitable for studying microbial community functions.
   (this thesis)

3. Handwritten notes provide researchers a tangible fulfilment that computers cannot.

4. Being physically cold by choice is invigorating, whereas by circumstance it is miserable.

5. Universities are businesses.

6. Processes of creativity required for scientific advances are overlooked in academia today.

Propositions belonging to the thesis, entitled

Ecology and evolution of microbial community assembly

Alanna Molly Leale
Wageningen, 13 November 2024

Ecology and evolution of microbial community assembly

Alanna Molly Leale

Thesis committee

Promotors

Prof. Dr Bas Zwaan
Professor of the Laboratory of Genetics
Wageningen University & Research

Dr Sijmen E. Schoustra
Associate Professor at the Laboratory of Genetics
Wageningen University & Research

Prof. Dr Eddy J. Smid
Personal Chair at the Laboratory of Food Microbiology
Wageningen University & Research

Other members

Prof. Dr Hauke Smidt, Wageningen University & Research
Prof. Dr Joana Falcao Salles, University of Groningen
Dr Matti Gralka, Vrije Universiteit Amsterdam
Dr Danna Gifford, University of Manchester, UK

This research was conducted under the auspices of the C. T. de Wit Graduate School for Production Ecology and Resource Conservation

Ecology and evolution of microbial community assembly

Alanna Molly Leale

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr Caroline Kroeze,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 13 November 2024
at 3.30p.m. in the Omnia Auditorium

Alanna Molly Leale
Ecology and evolution of microbial community assembly
196 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2024)
With references, with summary in English

DOI: https://doi.org/10.18174/673035

Table of Contents

Chapter 1	General Introduction	7
Chapter 2	Influence of metabolic guilds on a temporal scale in an experimental fermented food derived microbial community	35
Chapter 3	Shifts and rebound in microbial community function following repeated introduction of a novel species	75
Chapter 4	*S. cerevisiae* serves as keystone species for spoilage resistance in experimental synthetic wine yeast communities	107
Chapter 5	General Discussion	145
Appendix	Summary	181
	Acknowledgements	185
	About the author	190
	List of publications	192
	PE&RC Training and Education Statement	193

1

Chapter 1

General Introduction

1. Community composition-function relationships:

Regardless of biology expertise, there is general appreciation about the critical role biodiversity plays in ecosystem functions across our planet. An ecosystem is considered to be a "complex of living organisms, their physical environment, and all their interrelationships in a particular unit of space", whereas an ecological community is more narrowly defined as an "interacting group of various species in a common location" (Britannica Encyclopaedia 2020, 2024). A community is structured by many factors including which species are present (identity), the number and relative abundances of unique species (diversity), the absolute number of individuals (abundance or population size), and how quickly the community changes over time and after disturbances (stability and ecological succession). Diversity is a component of community composition that is highlighted to contribute to community function. Diversity encompasses both the richness (number of unique species) and the evenness (relative abundances), as well as genetic variation within species.

The link between diversity and community function has long been a central focus of ecology (Cardinale *et al.* 2012; Gonzalez *et al.* 2020). Community function has most commonly been estimated as community productivity or total biomass (Cardinale *et al.* 2007; Tilman, Reich and Isbell 2012; van Moorsel *et al.* 2021), as well as metabolic output (Louca *et al.* 2016, 2018), however measures of stability also contribute to function (Aubree *et al.* 2020). Furthermore, it is becoming recognised that the best suited measure of community function is system dependent. Studies generally demonstrate that higher diversity is associated with increases in measures of function (Balvanera *et al.* 2006); the positive relationship is typically attributed as an artifact of functional complementarity by collections of species with different niches (Cardinale *et al.* 2012). Sometimes described as a "portfolio" or "insurance effect", temporal and functional complementarity between diverse species further contributes to community function through enhanced stability against abiotic or biotic perturbations (Loreau 2010; Loreau and De Mazancourt 2013). Despite strong past attention, there are components of composition-function relationships still requiring further research and consideration. Notably, the time scales of experiments have thus far been limited and thereby overlook the role of evolution and long-term ecological processes (Donohue *et al.* 2016; Gonzalez *et al.* 2020). As scientists look forward, it is meaningful to investigate how the influence of community composition on community function changes over many generations.

Hence, the overarching objective of this PhD was to gain insight on the interconnected impacts that community composition, particularly diversity, and timescale have on community function. Inspired by concepts of traditional community ecology and

population genetics, I applied multi-species microbial communities of fermented foods to investigate community composition-function relationships. In this introduction, I first describe key ecological concepts of community ecology (Table 2, part A) and place them in the context of microbial communities. I then highlight relevant open questions in microbial community ecology and explain an inspiring framework that applies population genetics theory (genetically diverse individuals within a one species population) to multi-species communities (Vellend 2016) (Table 1). I close by proposing the suitability of Mabisi – a traditional fermented milk from Zambia - and wine yeasts, as model experimental systems and translate key ecological concepts to fermented foods systems (Table 2, part B). Lastly, I outline the specific objectives of my research chapters.

2. Key ecological concepts:

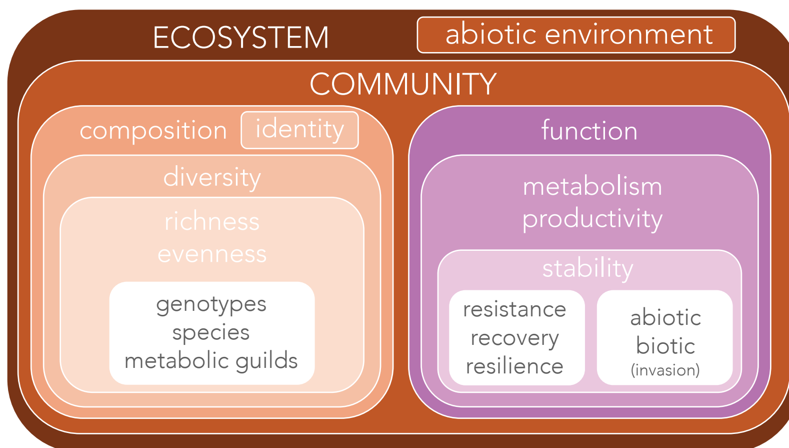


Figure 1: the "Russian doll" of community ecology concepts relevant to this thesis.

2a) Community function:

What is considered a well-functioning community in nature? Defining community function has not been consistent in past literature and it is complicated by combining ecological and humankind perspectives. In ecological terms, community function refers to internal processes of efficient nutrient conversion contributing to growth, with close associations often made to measurable human benefits – commonly food production (De Groot, Wilson and Boumans 2002). A bulk of past diversity-function studies have been performed in plant communities and other primary producers (Balvanera *et al.* 2006), hence total biomass is a common measure of community

function (Cleland 2011). Stability of these properties also contributes to function, which I detail later in this introduction under "Community stability". General ecological theories and associated definitions, which were mostly formulated with multicellular eukaryotes in mind, also apply to microbes (Mallon, Elsas and Salles 2015; Prosser and Martiny 2020).

Microbial communities & metabolic functions:
Although unseen to our naked eyes, microbial communities offer us intricate yet manipulatable ecosystems to investigate core questions about community composition and function. One may easily imagine rainforests, savannahs, and ocean reefs when thinking of diverse ecological communities, but similarly complex interactions including competition for resources, cooperation, and predation, shape the composition and function of microbial communities. The great value of microbial communities for both natural ecosystems and human-desired functions is rapidly being recognised. The power of natural microbial communities has long been utilised by humankind, and while traditional uses of naturally occurring microbial communities continues (i.e., food fermentation, composting), engineering approaches are also advancing to develop selected community compositions for specified functions (Sivasubramaniam and Franks 2016; Jones, Marken and Silver 2024).

Defining and measuring the function of microbial communities most commonly is highly connected to metabolic properties of the entire microbial community, particularly since exchange of metabolites is regarded as a main driver of microbe-microbe interactions. For example, the diverse metabolic properties of natural microbial communities are also the source of biogeochemical cycles for decomposition (Osburn *et al.* 2021; Ramond *et al.* 2022), drivers of food transformation via fermentation (Smid and Lacroix 2013; Gänzle *et al.* 2023), and important for digestion or infection protection in gastrointestinal-tract microbiota (Leser and Mølbak 2009; Rowland *et al.* 2018). Furthermore, the integral role of microbial communities in our daily lives is linked to their ability to persist in wide ranges of environments, supported by metabolic flexibility or plasticity (Voolstra and Ziegler 2020; Varahan and Laxman 2021; Marasco *et al.* 2023). Natural microbial communities however are often inherently complex - with diverse community compositions and resultant potential interactions, combined with metabolic plasticity and redundancy - which can pose challenges for predictability of desired functions and dynamics, especially over the long term. Microbial communities exist across nature, regardless of their functions of interest to humans; the reality is that their evolutionary histories may conflict with human aspirations of what we want them to do. By using natural communities as an inspiration or baseline, their limits can be potentially extended by reducing complexity to a limited number of selected strains (i.e., a synthetic microbial community).

Harnessing the potential for engineering synthetic communities, with selected species or strains, is a thriving sector of applied research in the domains of environment, nutrition, and agriculture (Sivasubramaniam and Franks 2016; Jones, Marken and Silver 2024). Please note, I refer throughout this thesis to "synthetic communities" as a consortium of selected species and/or strain compositions (Grosskopf and Soyer 2014). I do not refer to synthetic communities to mean that they have been genetically engineered with desired novel genes and consequent functions (Jones, Marken and Silver 2024).

Community stability:
Stability is one contributor to community function (Kinzig, Pacala and Tilman 2013; Aubree *et al.* 2020), and it holds particular attention as our ecosystems face quickening global changes. Ecosystem stability is frequently referred to in environmental policies (WWF 2020), yet a clear and collective definition is hard come by. As an initial thought, a dictionary-like description of being stable comes to mind - "not changing or fluctuating" (Mirriam-Webster 2024). In this thesis I focused on community stability to refer to an array of measurements. In fact, Ives and Carpenter 2007 list thirteen definitions of stability for a community and emphasise choosing the correct one should be guided by knowing the dynamics of the system and what perturbations are relevant (Ives and Carpenter 2007). Defining and measuring community stability is complicated by its multi-dimensional nature and plenitude of potential defining features (Savageau 1983; Ives and Carpenter 2007; van Elsas *et al.* 2007; Donohue *et al.* 2016), such as biotic or abiotic perturbation, immediate or long term responses, plus if stability is measured at the level of community composition or function. Trade-offs also exist between different features of community stability, so measuring one disturbance type against one measure of stability is inadequate to form generalities about community composition-stability relationships (Ives and Carpenter 2007; Donohue *et al.* 2016).

There are several unique components of stability, only some of which I assessed in this thesis. Firstly, the core components of stability are resistance, resilience, and recovery; resistance is the degree of change in a variable, recovery is if and the degree that the variable returns to its pre-disturbance state, and resilience is the rate at which it returns (Pimm 1984; Allison and Martiny 2008; Shade *et al.* 2012) (Table 2). Secondly, stability can be measured in different forms; in microbial communities, measurements of stability are often in the form of species composition and/or a function of interest (biomass production, compound degradation). Thirdly, there are also multiple forms of biotic or abiotic disturbances that can be applied to test stability. Common abiotic stresses to apply to microbial communities include temperature, salinity, acidity, antibiotic, and osmotic pressures (sugar concentrations) (Sjöstedt *et al.* 2018; Cairns

et al. 2020; Renes *et al.* 2020; Pourcelot *et al.* 2024), whereas biotic stressors can be applied in the form of a predator, phage, or competing species (van Elsas *et al.* 2007; Mallon *et al.* 2018; Mickalide and Kuehn 2019; Amor, Ratzke and Gore 2020; Spus *et al.* 2023). Lastly, the disturbance can be either a long term "press" disturbance, or a short-term "pulse" disturbance (Lake 2000; Shade *et al.* 2012), and the frequency of the disturbance may also vary. An extensive number of approaches consequently exist to investigate community stability due to the many possible combinations of measurement types, biotic *versus* abiotic, pulse *versus* press, and disturbance frequency. Evidently, I and others recognise that defining and measuring community stability is inconsistent across studies and therefore emphasise the importance to carefully specify your study's approach to stability. For example, clearly state 1) the component of stability assessed (such as resistance), 2) the type of perturbation (such as invasion by a novel species), 3) the measurement (such as change in community composition), and 4) your definition of the stability component (such as "resistance is the degree of change in community composition after versus before a perturbation").

2b) Diversity-invasion hypothesis:

Perturbation with a novel species introduction is one disturbance type to investigate resistance and recovery that is particularly relevant for studying diversity-stability relationships. The introduction and potential establishment of a novel species, or invader, is a common type of disturbance. Just as in macroecological communities, microbial communities also experience plentiful potential invasions, which involves the processes of introduction, establishment, spread, and impact by the introduced species (Mallon, Elsas and Salles 2015; Kinnunen *et al.* 2016; Vila *et al.* 2019). Species diversity is a commonly stated contributor to the fate of an invader, and thus community resistance, in both microbial (Jousset *et al.* 2011; van Elsas *et al.* 2012; Mallon, Elsas and Salles 2015; Kinnunen *et al.* 2016) and non-microbial communities (Kennedy *et al.* 2002; Tilman 2004; Petruzzella *et al.* 2020; Ernst *et al.* 2022). Termed the "diversity-invasion hypothesis", a positive relationship between diversity and resistance to invasion is observed when an introduced species exhibits lower levels of survival in resident communities with higher species richness (Kennedy *et al.* 2002; Mallon, Elsas and Salles 2015). Similarly, "biotic resistance theory" predicts that more phylogenetically diverse communities will be more resistant to invasion due to denser filled niche space, limiting colonisation by introduced species (Elton 1958; D'Antonio and Thomsen 2004; Petruzzella *et al.* 2020).

2c) Niche exclusion principle (NEP):

Understanding the link between community diversity and resistance against invasion can be interpreted through the classic "niche exclusion principle" (NEP) of ecology. The NEP states that only one ecological niche can be filled by one species at a time

(Gause 1934; Hardin 1960). An ecological niche of an animal was defined as "its place in the biotic environment, [and] its relations to food and enemies" (Elton 1927), and this can extend to the microbial world. So then, how to define a microbial niche? As discussed previously for community functions, for microbes, the niche refers strongly to their metabolic activities. Metabolic properties of a microbe are influenced by external abiotic and biotic factors, as well as their intrinsic genetic capacities (Malard and Guisan 2023). Overall, the traits of interest defining a microbe's niche are 1) which substrates are consumed? 2) which substrates are produced? and 3) under which abiotic and biotic conditions?

It is now realised that in its original simplistic form, the NEP is removed from the complexities of natural communities where multitudes of species may coexist. However, the NEP provides an influential foundation for understanding the role that community composition plays for function, including stability to perturbations, such as resistance to invasion. Based upon the niche exclusion principle, it is valid to state that establishment of an introduced species in a community can indicate to available niche space (Eisenhauer *et al.* 2013). The composition of a resident community plays a critical role in its resistance to invasion. Community composition can entail several components, including the number or richness of species (i.e., how many?), the evenness of species (i.e., dominant or rare types?), the phylogenetic diversity of species (i.e., how different?), the population density (i.e., how abundant?), and the identity of species (i.e., who's there?). As a community increases in species diversity, a wider range of resources are consumed, thereby limiting opportunities for a novel species to establish, and hence builds the basis of the diversity-invasion effect (Mallon, Elsas and Salles 2015; Petruzzella *et al.* 2020). Rather than species diversity, it is argued to instead view microbial communities at the level of distinct functional types, or metabolic guild diversity. Metabolic guilds could provide better classification than the current standard of species or genus level, which is decoupled from many functions of interest due to high variation between genotype lineages not captured in the DNA encoding 16S rRNA gene (Louca *et al.* 2016, 2018; Gralka *et al.* 2020; Allen *et al.* 2023; Reynolds *et al.* 2023). Metabolic guilds or functional types can be considered equivalent to the classic ecological classification of non-microbial species into guilds, lifeforms, or strategies (Fauth *et al.* 1996; Stroud *et al.* 2015). This is the basis of mine and others' terminology of metabolic guilds to refer to groups of microbes that can perform the same ecological functions (Louca *et al.* 2018; Allen *et al.* 2023; Leale *et al.* 2023; Reynolds *et al.* 2023).

3. Ongoing research directions:

3a) Why is community diversity not always a good predictor of invasion resistance?

The diversity-invasion hypothesis is an attractive perspective with convincing theory and examples, yet why does it not always hold true? Alternatively, overall diversity may play a negligible role due to influences of keystone or dominant species (Emery and Gross 2007; Ernst *et al.* 2023), the presence of a species that is phylogenetically closely related to the invader (Petruzzella *et al.* 2020), complex food webs (Thébault and Loreau 2005), scale or resource heterogeneity (Jiang and Morin 2004; Fridley *et al.* 2007), and interaction strength (Mallon *et al.* 2015; Ratzke, Barrere and Gore 2020). An "invasion paradox" of contrasting results means that a positive role of diversity against invasion is still not a certainty and under debate (Fridley *et al.* 2007), probably because a combination of factors predicts the success of invasion. Although much insight has been gained using plant communities (Kennedy *et al.* 2002; van Ruijven, De Deyn and Berendse 2003; Petruzzella *et al.* 2020; Ernst *et al.* 2022), as well as diverse environmental microbial communities (van Elsas *et al.* 2012; Eisenhauer *et al.* 2013; Mallon *et al.* 2015; Xing *et al.* 2021), disentangling the multiple factors contributing to invasion resistance remains difficult in such complex systems. As future engineering of microbial communities aims to achieve stability in community composition and function, it is necessary to investigate how and when species richness *versus* identity impacts resistance to invasion. Furthermore, most studies do not assess the entire process of invasion to a microbial community and only consider snapshots at one or a few time points (van Elsas *et al.* 2012; Mallon *et al.* 2015; Amor, Ratzke and Gore 2020).

3b) How are changes in microbial community composition linked to community function?

Community species composition is found to be decoupled from community metabolic function in microbial communities (Louca *et al.* 2016, 2018), which means that responses to perturbation in one (composition) may not align with the other (function). Natural microbial communities are found to have functional redundancies – meaning that different species perform overlapping functions, therefore the loss of one species can be compensated by another (Fetzer et al, 2015). Functional redundancies (Louca *et al.* 2018) combined with high metabolic flexibility of microbes (Carbonero, Oakley and Purdy 2014) can contribute to stability in function despite considerable changes in community composition. While beneficial to persistence of a community facing disturbances, metabolic redundancies, and metabolic plasticity in microbes pose challenges for predicting functions of interest from community composition information alone. It is of general interest to know

how community responses to perturbation in their composition *versus* function contrast, including whether shifts to alternative states persist once perturbation stops. To grasp the complexity of feedback loops between ecological (interactions between species and their environment), evolutionary (genotypes within a species), and physiological (metabolism, gene expression) dynamics, research needs to look beyond explaining "changes to changes" and consider when null expectations are found; for example, investigating when shifts in composition (species, strains) are observed while community functions remain stable (i.e., productivity, metabolic output) (Figure 2). The reverse is equally insightful – if community function changes yet composition goes unaltered suggests plasticity in gene expression or genetic changes below the level of detection (i.e., species, sequenced gene) (Fig. 2). Understanding the link, or lack of link, between community composition and function would also benefit from studies done over extended time frames of multiple generations.

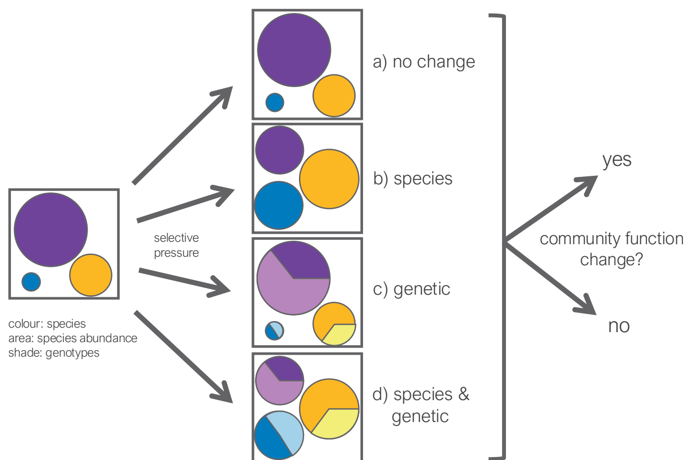


Figure 2: potential dynamics of community composition response to a selective pressure.

3c) How does the relationship between community composition and function change over time?

The interconnection of community composition with function has long been a central topic in ecology, yet it is still unclear how this relationship evolves over time. When a diverse community experiences a new and consistent environment, how does it respond – both in community composition and function? Species sorting – where differential fitness or adaption between species induce changes in community composition – may alter community composition and diversity. For instance, when a particular species or type (such as a metabolic guild) is better adapted to the new selective condition, it

may increase in relative abundance (Vellend 2010). In this way, species sorting may affect overall community functioning, when species or types that increase or decrease in abundance are key to overall community function. While bodies of theory exist on the theme of species sorting in communities (Loeuille and Leibold 2008), experiments explicitly testing species sorting effects on community function output are not common (Fiegna *et al.* 2015; Sjöstedt *et al.* 2018; Groenenboom *et al.* 2022). Since most experimental studies on the influence of community diversity on community function focus on short-term ecological timescales of one or few generations (Wagg *et al.* 2014; Sierocinski *et al.* 2018; Aubree *et al.* 2020; Gonzalez *et al.* 2020), relatively little is known on how this influence changes over multiple generations (Fiegna *et al.* 2015; Louca *et al.* 2020).

In an applied engineering context, such as consortia of microbes in bioreactors, the ability to predict and control microbial community functions overtime is also highly sought, if not a strict requirement. Yet, the dynamic complexity between functions from multiple species, genotypes, and environment poses challenges, especially over extended timescales of multiple generations and propagation batches. Similar as for engineering purposes, predicting community level functions is a primary goal for fundamental ecology research since it provides insight to the consequences of ecological and evolutionary processes. Whether utilising natural microbial communities or composing novel synthetic ones, researchers must understand the ecological and evolutionary mechanisms contributing to stable, or at least predictable, community function. Investigating community composition-function relationships on longer time scales has been advocated (Hendry 2019; Gonzalez *et al.* 2020) and such research can be facilitated by the short generation times of microbial communities. The rapid generational turnover of microbes creates a tight dynamic between processes of ecology and evolution. Ecological *versus* evolutionary processes have traditionally been studied in isolation, but it is now recognised they should be considered simultaneously, despite the challenges poses by such an overlap. Understanding how changes across hierarchical levels - from genes, to species, to communities - influence one another is undoubtedly complicated but merits initial steps to elucidate, which can be supported by a suitable framework and model system. Approaches from population genetics can be useful for studying species sorting and associated community function under novel conditions.

4. Multi-species community framework:

4a) Translating population genetics to community ecology:

Understanding the dynamic feedbacks between genetic changes, shifts in community species composition, and environmental alterations has been a core tenet of the rapidly developing field of evolutionary ecology (Hendry 2017). It is challenging to find a cohesive framework that unites several fields, including but not limited to population genetics, community ecology, and ecosystem ecology; however, resemblances between research concentrations do exist and should be built upon (Vellend 2016). First consider population genetics' conceptualisation of evolutionary processes, which concentrates on tracing genetic changes within a species (Gillespie 2004). For adaptive evolution, rises and declines in genotypes are then linked to selection forces (biotic or abiotic) and associated fitness of each genotype. Now compare community ecology; the focus here is on shifts in species composition of a community in response to an environmental pressure. Further impacts on community function are also often considered to understand species' roles.

A core inspiration for this thesis were concepts from community ecologist Mark Vellend, who proposed: can unique genotypes within a species be paralleled to different species within a community (Vellend 2016)? He thoroughly argues to extend the framework of population genetics to interpret community ecology. Comparable to how a single species' population is composed of various genotypes, communities and ecosystems contain various species. Furthermore, species diversity and genetic diversity have parallel processes and patterns, but they just act at different hierarchical levels (Vellend 2016) (Table 1). Of further consideration is how community function compares to genotype fitness (Sanchez *et al.* 2023), such that selection could act on community members for functional traits just as genotypes with differing fitness are selected (Goodnight 1990; Williams and Lenton 2007; Doolittle and Inkpen 2018). A community level selection or breeding process could then arise (Arias-Sánchez, Vessman and Mitri 2019; Doulcier *et al.* 2020), which I explore more in Chapter 5's general discussion. The conceptual framework outlined by Mark Vellend inspired me to design experiments of multi-species microbial communities to investigate effects of species sorting on community functions.

Table 1: Paralleled processes and components in population genetics and community ecology (Vellend 2016). Italics indicate personal extrapolations.

Evolution (population genetics)	**Community Ecology**
selection	species sorting
mutation	speciation
drift (genotypes)	drift (species)
gene flow	dispersal *(introduction/invasion)*
genotype	single species
single species	community of species
species fitness	*community function*
species persistence	*compositional / functional stability*

4b) Experimental evolution of multi-species microbial communities:

One general motivation for my thesis was to investigate multi-species microbial communities in the style of classic microbial experimental experiments – where traditionally a single isogenic strain is repeatedly grown and propagated to fresh media for an indefinite period. The longest and most well-known exemplar microbial experimental evolution study is the "Lenski *E. coli* experiment" at Michigan State University, which reached over 73 000 generations in 2020. Until now, microbial studies of evolutionary ecology dynamics have largely focused on specific interactions with a limited number of species and/or laboratory strains. Microbial experimental evolution is well established for several single isolate study systems, where studies have revealed the emergence and maintenance of diversity through cross feeding interactions, resource use specialisation, and environmental niche specialisation (Rosenzweig *et al.* 1994; Rainey and Travisano 1998; Treves, Manning and Adams 1998; Kassen 2009). While studies investigating diversity within a species have provided essential understanding about diversity-function relationships in microbial communities (Jousset *et al.* 2011; Eisenhauer *et al.* 2013), it is time to extend beyond single species and recent microbial studies are now including more diverse and natural communities collected from the environment (Gravel *et al.* 2011; Castledine, Padfield and Buckling 2020; Gorter, Manhart and Ackermann 2020; Piccardi *et al.* 2022; Ruiz *et al.* 2023). The simplicity yet power of experimentally evolving multi-species communities is undoubtably appealing, but study systems must also not be *too* complex, since directed hypotheses should still be elegantly investigated.

The value of bridging experimental evolution studies closer to nature by using complex communities is debatable, yet often still considered desirable; natural communities can pose considerable difficulties for creating controlled and traceable systems but offer

stronger translation of findings to the real world. In this context, pairwise studies of two species are relatively common (D'Souza *et al.* 2018; Ratzke and Gore 2018; McClean *et al.* 2019; Giri *et al.* 2021), and a growing number of studies have used isolates from varying environments for laboratory propagation of multi-species microbial communities (Cairns *et al.* 2020; Scheuerl *et al.* 2020; Piccardi *et al.* 2022; Chang *et al.* 2023). Constructing experimental multi-species microbial systems can take two approaches; either a community is sampled from the environment then diversity best maintained (or manipulated, often through dilution) in laboratory conditions (van Elsas *et al.* 2012; Mallon *et al.* 2015; Castledine, Padfield and Buckling 2020; Xing *et al.* 2021; Groenenboom *et al.* 2022; Leale *et al.* 2023), or synthetic communities are assembled from a collection of single isolates that represent the community (Scheuerl *et al.* 2020; Piccardi *et al.* 2022; Chang *et al.* 2023; Pourcelot *et al.* 2023; Ruiz *et al.* 2023; Weiss *et al.* 2023). Building upon such previous studies of complex microbial communities over multiple generations, I aimed to study the relationship between community composition and function, including stability, over time. An ideal experimental study system combines the complexity of natural communities and environments with the control and quick generation times of laboratory evolution experiments. Guided by core community ecology concepts, I applied both natural and synthetic community approaches in my thesis to demonstrate the suitability of microbial communities of fermented foods to achieve such a balance between natural environments and laboratory control.

5. Fermented foods as model systems:

5a) Fermented foods in general:

There is a growing movement of using fermented foods as a model system in ecology and evolutionary biology (Wolfe and Dutton 2015; Carbonetto *et al.* 2018; May *et al.* 2019; Alekseeva *et al.* 2021; Conacher *et al.* 2021; Melkonian *et al.* 2023; Pourcelot *et al.* 2024). Contemplating the strengths and pitfalls of laboratory studies *versus* natural communities is a continuous feat. Laboratory studies provide ease of manipulation and control but come with the downfall of removal from the real world and hence debatable relevance for more complex ecosystems. I believe that fermented foods provide an innovative approach to strike a balance between the controlled, quick generation times of laboratory microbial studies, with the complexity, yet resemblance to diverse natural communities and their natural environment. Fermented foods provide better tractable communities to study compared to other environments (Wolfe and Dutton 2015; Alekseeva *et al.* 2021). Studying natural microbial communities from soils or water systems, for example, can pose many challenges or pitfalls for laboratory experiments; notably, successfully culturing species to maintain high

diversity and mimicking their natural conditions *in vitro* can be difficult (Bodor *et al.* 2020). Comparatively, the substrates of fermented foods (i.e., milk, grape juice, tea) provide a complex yet easily reproduced environment in the laboratory with less bottlenecking of diversity due to high culturing success of microbial community members in fermented foods. Also advantageous to fermented foods is that they have defined measurements that describe functions of their microbial community function (i.e., aroma, acidity, thickness, pathogen resistance). The known functional properties of fermented foods facilitate investigating the influence between diversity and community function. Exploring microbial dynamics in fermented foods also advances knowledge beyond evolutionary ecology by providing valuable insight for improving food product safety, quality, and nutrition. With consideration, knowledge can contribute to multidisciplinary objectives. I sought to achieve this in my thesis by using Mabisi - a traditionally fermented milk beverage from Zambia (Chapters 2, 3), and wine yeast communities (Chapter 4) as model systems for evolutionary ecology questions with real-world applications in mind.

5b) Mabisi:

Mabisi is a traditionally fermented milk beverage from Zambia, in which processing methods differ regionally across Zambia. In the "active back-slopping" method, a small volume of the final product re-inoculates another batch of raw milk over successive cycles (Moonga *et al.* 2019), therefore, some Mabisi microbial communities have been potentially co-evolving for several years. The "back-slop" process is equivalent to experimental evolution studies used in evolutionary biology research (Jessup *et al.* 2004; Schoustra *et al.* 2013). Additionally, unlike mass produced yogurts which contain a few bacterial species, Mabisi hosts a diverse microbial community composed of six to ten dominant lactic and acetic acid bacteria, as well as yeasts and viruses (Schoustra *et al.* 2013; Moonga *et al.* 2020). In line with diversity-invasion hypothesis and NEP, the long shelf life of Mabisi and other traditional fermented foods against pathogens could reasonably be an outcome of their diverse microbial communities (Groenenboom 2019). The manageable yet moderate level of diversity in Mabisi that is well-maintained under laboratory conditions (Groenenboom *et al.* 2022), makes Mabisi a suitable object for laboratory studies of community composition-function relationships overtime and in the face of perturbation by introduction of novel species (i.e., pathogens).

Figure 3: Local sale of Mabisi in Mumbwa and Monze, Zambia. Mabisi is prepared at home in large plastic buckets, then transferred to small plastic bottles for sale. Pictures by Robert Nhlane.

5c) Wine:

Wine yeast communities serve as a useful model system to test ecological questions due to their relatively manageable diversity, the wealth of knowledge on their biochemistry or system function, and the presence of established synthetic media for laboratory experiments (Bagheri et al. 2020; Conacher et al. 2021; Pourcelot et al. 2023; Ruiz et al. 2023). It is already well known that extracted grape juice, termed wine must, initially harbours a diversity of yeast and bacteria, but this diversity drops as the fermentation progresses. Overtime, there is a rise and final dominance of *Saccharomyces cerevisiae* as it withstands fermentation conditions of high alcohol concentration and low oxygen (Albergaria and Arneborg 2016; Conacher et al. 2021). Despite being at negligible or even undetectable levels in wine must (Drumonde-Neves *et al.* 2021; Pourcelot 2023), *S. cerevisiae* is considered necessary for complete fermentation "to dryness" (<2 g/L sugar) and is commonly added as a starter culture by wine producers (Albergaria and Arneborg 2016; Ciani et al. 2016; Binati et al. 2020). The diversity and composition of wine microbial communities is gaining importance as appeal grows from consumers and producers for natural wines made through spontaneous fermentation or wines made with starter cultures of naturally occurring, and often non-*S. cerevisiae* yeasts (Galati *et al.* 2019; Roudil *et al.* 2020). Although naturally occurring non-*S. cerevisiae* yeasts can contribute favourably to aroma and sensory measure (Binati *et al.* 2020; Roudil *et al.* 2020), they can also pose problems with slow or stuck fermentations (Ciani, Beco and Comitini 2006; Medina *et al.* 2012; Taillandier *et al.* 2014) and consequent risk of spoilage by undesirable microorganisms

(Albergaria and Arneborg 2016). Ecologically, the common problem of spoilage during wine production and during storage is an example of community instability against invasion. *S. cerevisiae* is known to prevent spoilage by achieving a quick, full fermentation which consumes available resources and leaves little opportunity for other species to proliferate (Albergaria and Arneborg 2016). However, the dominance of *S. cerevisiae* has been shown to be influenced by ecological interactions with non-*S. cerevisiae* species (Boynton and Greig 2016; Bagheri *et al.* 2020; Conacher *et al.* 2021; Ruiz *et al.* 2023). Thus, in ecological terms, *S. cerevisiae* demonstrates a keystone species effect and helps make wine yeast communities a good model system to test diversity versus composition identity effects to invasion resistance.

5d) Ecological definitions in fermented foods:

So, do definitions of well-functioning ecosystems in nature translate to fermented food model systems? In microbial communities of fermented foods, it is suitable to focus on community level metabolic output and product qualities relevant for consumption as measures of community function. Metabolic profiles measured through gas chromatography mass spectrometry (GC-MS) quantify volatile organic compounds (VOCs) contributing to aroma and taste, while inferring to active pathways and gene expression across the community (Krömer *et al.* 2004; Smid *et al.* 2005; De Filippis *et al.* 2016). Another especially important function is acidification by growth of dominant lactic and acetic acid bacteria, since acidity impacts both consumer preferences (Ott *et al.* 2000; Sikombe *et al.* 2023) and prevention of food borne pathogens (Mpofu *et al.* 2016; Schoustra *et al.* 2022). As mentioned previously, stability is also an element of community function and perturbation by introduction of a novel species is one approach to test community stability. In fermented foods, contamination by food borne pathogens or spoilage microorganisms are, in ecological terms, considered an invasion of the resident community. Abiotic disturbances of temperature or sugar concentrations are also relevant for fermented foods but were not a focus in my thesis. Well-defined measurements of community function are a considerable advantage in using fermented foods as model systems for studying ecological and evolutionary processes.

Table 2: Community ecology concepts defined in general (part A) and in the context of microbial communities of fermented foods (part B).

Term	A) traditional (field) community ecology	B) microbial communities of fermented foods
Function	Biomass, total productivity, stability in these properties	Metabolic profiles, acidity, thickness, and stability in these properties
Composition	Identity and relative abundances of species present	a) unique clusters from 16S rRNA sequencing, b) a known consortium of species or genotypes
Niche	"An [organisms'] place in the biotic environment, [and] its relations to food and enemies" (Gause 1934)	Which substrates are consumed and produced, and under which conditions (i.e., metabolism)
Functional guild	"A group of species that exploit the same class of environmental resources in a similar way" (Fauth *et al.* 1996)	Capable of performing the same metabolic processes (example: lactic acid versus acetic acid bacteria)
Diversity	Number of species and evenness of species' relative abundances	Number and evenness of a) unique clusters from 16S rRNA sequencing (i.e., species), b) functional guilds
Stability: *resistance*	Change from pre-disturbance state following disturbance. Little or no change = high resistance.	Change in metabolic profiles and community composition
Stability: *recovery*	Return to a pre-disturbance state once disturbance is stopped.	Return in metabolic profiles and community composition once disturbance stopped.
Invasion	The introduction, establishment, and spread of species outside their native range.	Contamination or spoilage by microbial food borne pathogens or contaminants (i.e., spoilage)
Species sorting	Changes in community composition of relative species abundances	Changes in relative abundances of a) unique clusters from 16S rRNA sequencing (i.e., species), b) functional guilds

6. Thesis outline:

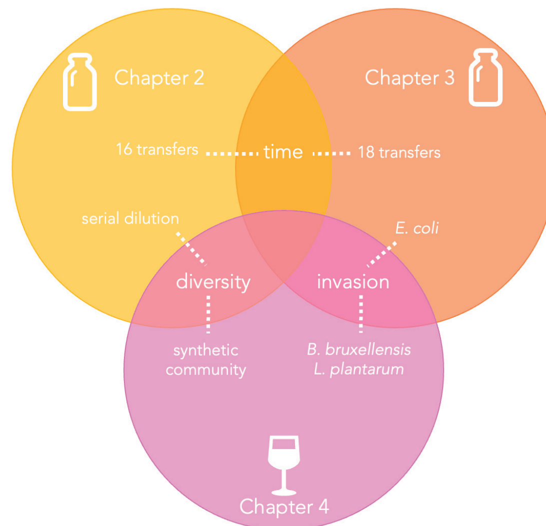


Figure 4: visual outline of thesis chapters.

This PhD presents microbial communities of Mabisi and wine as model systems to investigate the interacting roles community composition, particularly diversity, and timescale have on community function, including stability to an introduced species. The three data chapters have overlaps in their objectives and approaches (Fig. 4), with a commonality of being components of community assembly (the processes that shape the identity and abundance of species within ecological communities) (Chase 2003).

Chapter 2

In Chapter 2 I investigated the influence of metabolic guild diversity on associated community function by propagating Mabisi communities diluted to various levels for approximately 100 generations. I specifically asked:

1. Does altering initial metabolic guild diversity, with the associated loss of low abundance microbial types, influence community function?
2. And if so, are such shifts in community function stable over repeated propagation?
3. Does removing metabolic guilds influence bacterial species sorting trajectories over repeated cycles of propagation?

I found that the starting richness of metabolic guilds had a repeatable effect on bacterial community compositions, metabolic profiles, and acidity. The influence of the yeast guild played a dramatic role on function, but interestingly not on long-term species sorting trajectories of the remaining bacterial community. This chapter's findings were interpreted to suggest an unexpected niche division between yeast and bacterial communities, and to be evidence of ecological selection on the bacterial communities.

Chapter 3

I continued using Mabisi in Chapter 3 as a model system to investigate the persistent impacts by failed invaders. I conducted an experimental test of community stability over multiple generations against repeated novel species introduction by propagating a Mabisi community for approximately 100 generations, with or without, repeated introduction of *Escherichia coli* at each transfer. I asked the following:

1. Does repeated introduction of a novel species (*E. coli*) shift resident community (Mabisi) composition and function over time?
2. If yes, does the resident community rebound once introduction is stopped?

Community function was determined through metabolic profiles, which demonstrated alterations immediately after *E. coli* introduction, followed by recovery, or rebound once ceased. In contrast to this proxy of community function, I never detected changes in the bacterial community composition. This chapter evidenced that community composition and function do not necessarily respond in parallel to an introduced species, potentially due to genotypic changes below species level detection or metabolic plasticity.

Chapter 4

In Chapter 4 I explored a second model system – microbial communities of wine fermentation. Using synthetic communities comprised of combinations of four wine yeasts (*S. cerevisiae, L. thermotolerans, T. delbrueckii, S. bacillaris*), I tracked over 21 days the presence of introduced *B. bruxellensis* spoilage yeast and *L. plantarum* lactic acid bacteria to ask the following:

1. Does yeast community species richness impact the establishment of *B. bruxellensis* yeast and *L. plantarum* bacteria during wine fermentation?
2. How does yeast species identity (i.e., absence/presence) influence such establishment?

I found that species identity rather than richness drove the prevention of establishment of *B. bruxellensis* and *L. plantarum*, with *S. cerevisiae* playing a critical keystone

species role. Aside from spoilage prevention by *S. cerevisiae*, the four resident yeast species demonstrated a strict dominance ranking of competitive exclusion regardless of background community composition. This chapter lends evidence against the commonly predicted positive relationship between species richness and resistance to invasion.

Chapter 5

As with all research, just as many unknowns remain than were solved during this thesis. There are several research questions which inspired me during my research journey but were not yet realised in this thesis. In my general discussion I highlight a few future perspectives, including the exciting potential for community level selection in microbial communities, predicting microbial community trajectories overtime, and next steps for profiling microbial community compositions. I also briefly outline existing synthetic microbial community model systems and then propose a Mabisi-inspired system for tracking ecological and evolutionary processes. I conclude by considering practical applications of my research for traditional fermentation practices.

References:

Albergaria H, Arneborg N. Dominance of Saccharomyces cerevisiae in alcoholic fermentation processes: role of physiological fitness and microbial interactions. Appl Microbiol Biotechnol 2016;100:2035–46.

Alekseeva AY, Groenenboom AE, Smid EJ et al. Eco-Evolutionary Dynamics in Microbial Communities from Spontaneous Fermented Foods. Int J Environ Res Public Health 2021;18:10093.

Allen B, Gonzalez-Cabaleiro R, Ofiteru ID et al. Diversity and metabolic energy in bacteria. FEMS Microbiol Lett 2023;370:fnad043.

Allison SD, Martiny JBH. Resistance, resilience, and redundancy in microbial communities. Proc Natl Acad Sci 2008;105:11512–9.

Amor DR, Ratzke C, Gore J. Transient invaders can induce shifts between alternative stable states of microbial communities. Sci Adv 2020;6:eaay8676.

Arias-Sánchez FI, Vessman B, Mitri S. Artificially selecting microbial communities: If we can breed dogs, why not microbiomes? PLoS Biol 2019;17:e3000356.

Aubree F, David P, Jarne P et al. How community adaptation affects biodiversity–ecosystem functioning relationships. Ecol Lett 2020;23:1263–75.

Bagheri B, Bauer FF, Cardinali G et al. Ecological interactions are a primary driver of population dynamics in wine yeast microbiota during fermentation. Sci Rep 2020;10:4911.

Balvanera P, Pfisterer AB, Buchmann N et al. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. Ecol Lett 2006;9:1146–56.

Binati RL, Lemos Junior WJF, Luzzini G et al. Contribution of non-Saccharomyces yeasts to wine volatile and sensory diversity: A study on Lachancea thermotolerans, Metschnikowia spp. and Starmerella bacillaris strains isolated in Italy. Int J Food Microbiol 2020;318:108470.

Bodor A, Bounedjoum N, Vincze GE et al. Challenges of unculturable bacteria: environmental perspectives. Rev Environ Sci Biotechnol 2020;19:1–22.

Boynton PJ, Greig D. Species richness influences wine ecosystem function through a dominant species. Fungal Ecol 2016;22:61–72.

Britannica Encyclopaedia. Community. 2020.

Britannica Encyclopaedia. Ecosystem. 2024.

Cairns J, Jokela R, Becks L et al. Repeatable ecological dynamics govern the response of experimental communities to antibiotic pulse perturbation. Nat Ecol Evol 2020;4:1385–94.

Carbonero F, Oakley BB, Purdy KJ. Metabolic Flexibility as a Major Predictor of Spatial Distribution in Microbial Communities. PLOS ONE 2014;9:e85105.

Carbonetto B, Ramsayer J, Nidelet T et al. Bakery yeasts, a new model for studies in ecology and evolution. Yeast 2018;35:591–603.

Cardinale BJ, Duffy JE, Gonzalez A et al. Biodiversity loss and its impact on humanity. Nature 2012;486:59–67.

Cardinale BJ, Wright JP, Cadotte MW et al. Impacts of plant diversity on biomass production increase through time because of species complementarity. Proc Natl Acad Sci 2007;104:18123–8.

Castledine M, Padfield D, Buckling A. Experimental (co) evolution in a multi-species microbial community results in local maladaptation. Ecol Lett 2020;23:1673–81.

Chang C-Y, Bajić D, Vila JCC et al. Emergent coexistence in multispecies microbial communities. Science 2023;381:343–8.

Chase JM. Community assembly: when should history matter? Oecologia 2003;136:489–98.

Ciani M, Beco L, Comitini F. Fermentation behaviour and metabolic interactions of multistarter wine yeast fermentations. Int J Food Microbiol 2006;108:239–45.

Ciani M, Capece A, Comitini F et al. Yeast Interactions in Inoculated Wine Fermentation. Front Microbiol 2016;7.

Cleland EE. Biodiversity and Ecosystem Stability. Nat Educ 2011.

Conacher C, Luyt N, Naidoo-Blassoples R et al. The ecology of wine fermentation: a model for the study of complex microbial ecosystems. Appl Microbiol Biotechnol 2021;105:3027–43.

D'Antonio CM, Thomsen M. Ecological Resistance in Theory and Practice. Weed Technol 2004;18:1572–7.

De Groot RS, Wilson MA, Boumans RM. A typology for the classification, description and valuation of ecosystem functions, goods and services. Ecol Econ 2002;41:393–408.

Donohue I, Hillebrand H, Montoya JM et al. Navigating the complexity of ecological stability. Ecol Lett 2016;19:1172–85.

Doolittle WF, Inkpen SA. Processes and patterns of interaction as units of selection: An introduction to ITSNTS thinking. Proc Natl Acad Sci 2018;115:4006–14.

Doulcier G, Lambert A, De Monte S et al. Eco-evolutionary dynamics of nested Darwinian populations and the emergence of community-level heredity. Elife 2020;9:e53433.

Drumonde-Neves J, Fernandes T, Lima T et al. Learning from 80 years of studies: a comprehensive catalogue of non-Saccharomyces yeasts associated with viticulture and winemaking. FEMS Yeast Res 2021;21:foab017.

D'Souza G, Shitut S, Preussger D et al. Ecology and evolution of metabolic cross-feeding interactions in bacteria. Nat Prod Rep 2018;35:455–88.

Eisenhauer N, Schulz W, Scheu S et al. Niche dimensionality links biodiversity and invasibility of microbial communities. Funct Ecol 2013;27:282–8.

van Elsas JD, Chiurazzi M, Mallon CA et al. Microbial diversity determines the invasion of soil by a bacterial pathogen. Proc Natl Acad Sci 2012;109:1159–64.

van Elsas JD, Hill P, Chroňáková A et al. Survival of genetically marked Escherichia coli O157:H7 in soil as affected by soil microbial community shifts. ISME J 2007;1:204–14.

Elton CS. Animal Ecology. Sidgwick & Jackson, London, 1927.

Elton CS. The Ecology of Invasions by Animals and Plants. Boston, MA: Springer US, 1958.

Emery SM, Gross KL. Dominant Species Identity, Not Community Evenness, Regulates Invasion in Experimental Grassland Plant Communities. Ecology 2007;88:954–64.

Ernst AR, Barak RS, Glasenhardt M-C et al. Dominant species establishment may influence invasion resistance more than phylogenetic or functional diversity. J Appl Ecol 2023;60:2652–64.

Ernst AR, Barak RS, Hipp AL et al. The invasion paradox dissolves when using phylogenetic and temporal perspectives. J Ecol 2022;110:443–56.

Fauth JE, Bernardo J, Camara M et al. Simplifying the Jargon of Community Ecology: A Conceptual Approach. Am Nat 1996;147:282–6.

Fiegna F, Moreno-Letelier A, Bell T et al. Evolution of species interactions determines microbial community productivity in new environments. ISME J 2015;9:1235–45.

De Filippis F, Genovese A, Ferranti P et al. Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. Sci Rep 2016, DOI: 10.1038/srep21871.

Fridley JD, Stachowicz JJ, Naeem S et al. The Invasion Paradox: Reconciling Pattern and Process in Species Invasions. Ecology 2007;88:3–17.

Galati A, Schifani G, Crescimanno M et al. "Natural wine" consumers and interest in label information: An analysis of willingness to pay in a new Italian wine market segment. J Clean Prod 2019;227:405–13.

Gänzle MG, Monnin L, Zheng J et al. Starter Culture Development and Innovation for Novel Fermented Foods. 2023, DOI: 10.1146/annurev-food-072023-034207.

Gause GF. The Struggle for Existence. Williams and Wilkins, Baltimore, 1934.
Gillespie JH. Population Genetics: A Conside Guide. JHU Press, 2004.
Giri S, Oña L, Waschina S et al. Metabolic dissimilarity determines the establishment of cross-feeding interactions in bacteria. Curr Biol 2021;31:5547-5557.e6.
Gonzalez A, Germain RM, Srivastava DS et al. Scaling-up biodiversity-ecosystem functioning research. Ecol Lett 2020;23:757–76.
Goodnight CJ. Experimental studies of community evolution I: The response to selection at the community level. Evolution 1990;44:1614–24.
Gorter FA, Manhart M, Ackermann M. Understanding the evolution of interspecies interactions in microbial communities. Philos Trans R Soc B Biol Sci 2020;375:20190256.
Gralka M, Szabo R, Stocker R et al. Trophic interactions and the drivers of microbial community assembly. Curr Biol 2020;30:R1176–88.
Gravel D, Bell T, Barbera C et al. Experimental niche evolution alters the strength of the diversity–productivity relationship. Nature 2011;469:89–92.
Groenenboom AE. Microbial community dynamics in traditionally fermented milk. Unpubl PhD Thesis WUR 2019.
Groenenboom AE, van den Heuvel J, Zwaan BJ et al. Species dynamics in natural bacterial communities over multiple rounds of propagation. Evol Appl 2022;15:1766–75.
Grosskopf T, Soyer OS. Synthetic microbial communities. 2014;18:72–7.
Hardin G. The Competitive Exclusion Principle: An idea that took a century to be born has implications in ecology, economics, and genetics. science 1960;131:1292–7.
Hendry AP. Eco-Evolutionary Dynamics., 2017.
Hendry AP. A critique for eco-evolutionary dynamics. Funct Ecol 2019;33:84–94.
Ives AR, Carpenter SR. Stability and diversity of ecosystems. science 2007;317:58–62.
Jessup CM, Kassen R, Forde SE et al. Big questions, small worlds: Microbial model systems in ecology. Trends Ecol Evol 2004, DOI: 10.1016/j.tree.2004.01.008.
Jiang L, Morin PJ. Productivity gradients cause positive diversity–invasibility relationships in microbial communities. Ecol Lett 2004;7:1047–57.
Jones EM, Marken JP, Silver PA. Synthetic microbiology in sustainability applications. Nat Rev Microbiol 2024:1–15.
Jousset A, Schmid B, Scheu S et al. Genotypic richness and dissimilarity opposingly affect ecosystem functioning. Ecol Lett 2011;14:537–45.
Kassen R. Toward a General Theory of Adaptive Radiation. Ann N Y Acad Sci 2009;1168:3–22.
Kennedy TA, Naeem S, Howe KM et al. Biodiversity as a barrier to ecological invasion. Nature 2002;417:636–8.
Kinnunen M, Dechesne A, Proctor C et al. A conceptual framework for invasion in microbial communities. ISME J 2016;10:2773–9.
Kinzig AP, Pacala S, Tilman D. The Functional Consequences of Biodiversity: Empirical Progress and Theoretical Extensions (Mpb-33). Princeton University Press, 2013.
Krömer JO, Sorgenfrei O, Klopprogge K et al. In-Depth Profiling of Lysine-Producing Corynebacterium glutamicum by Combined Analysis of the Transcriptome, Metabolome, and Fluxome. J Bacteriol 2004, DOI: 10.1128/JB.186.6.1769-1784.2004.
Lake PS. Disturbance, patchiness, and diversity in streams. J North Am Benthol Soc 2000;19:573–92.
Leale A, Auxier B, Smid EJ et al. Influence of metabolic guilds on a temporal scale in an experimental fermented food derived microbial community. FEMS Microbiol Ecol 2023;99:fiad112.
Leser TD, Mølbak L. Better living through microbial action: the benefits of the mammalian gastrointestinal microbiota on the host. Environ Microbiol 2009;11:2194–206.

Loeuille N, Leibold MA. Evolution in Metacommunities: On the Relative Importance of Species Sorting and Monopolization in Structuring Communities. Am Nat 2008;171:788–99.
Loreau M. Linking biodiversity and ecosystems: towards a unifying ecological theory. Philos Trans R Soc B Biol Sci 2010;365:49–60.
Loreau M, De Mazancourt C. Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. Ecol Lett 2013;16:106–15.
Louca S, Jacques SMS, Pires APF et al. High taxonomic variability despite stable functional structure across microbial communities. Nat Ecol Evol 2016;1:1–12.
Louca S, Polz MF, Mazel F et al. Function and functional redundancy in microbial systems. Nat Ecol Evol 2018;2:936–43.
Louca S, Rubin IN, Madilao LL et al. Effects of forced taxonomic transitions on metabolic composition and function in microbial microcosms. Environ Microbiol Rep 2020;12:514–24.
Malard LA, Guisan A. Into the microbial niche. Trends Ecol Evol 2023;38:936–45.
Mallon CA, Elsas JDV, Salles JF. Microbial Invasions: The Process, Patterns, and Mechanisms. Trends Microbiol 2015;23:719–29.
Mallon CA, Le Roux X, van Doorn GS et al. The impact of failure: unsuccessful bacterial invasions steer the soil microbial community away from the invader's niche. ISME J 2018;12:728–41.
Mallon CA, Poly F, Le Roux X et al. Resource pulses can alleviate the biodiversity–invasion relationship in soil microbial communities. Ecology 2015;96:915–26.
Marasco R, Fusi M, Coscolín C et al. Enzyme adaptation to habitat thermal legacy shapes the thermal plasticity of marine microbiomes. Nat Commun 2023;14:1045.
May A, Narayanan S, Alcock J et al. Kombucha: a novel model system for cooperation and conflict in a complex multi-species microbial ecosystem. PeerJ 2019;7:e7565.
McClean D, Friman V-P, Finn A et al. Coping with multiple enemies: pairwise interactions do not predict evolutionary change in complex multitrophic communities. Oikos 2019;128:1588–99.
Medina K, Boido E, Dellacassa E et al. Growth of non-Saccharomyces yeasts affects nutrient availability for Saccharomyces cerevisiae during wine fermentation. Int J Food Microbiol 2012;157:245–50.
Melkonian C, Zorrilla F, Kjærbølling I et al. Microbial interactions shape cheese flavour formation. Nat Commun 2023;14:8348.
Mickalide H, Kuehn S. Higher-Order Interaction between Species Inhibits Bacterial Invasion of a Phototroph-Predator Microbial Community. Cell Syst 2019;9:521-533.e10.
Mirriam-Webster. Definition of STABLE. 2024.
Moonga HB, Schoustra SE, van den Heuvel J et al. Composition and Diversity of Natural Bacterial Communities in Mabisi, a Traditionally Fermented Milk. Front Microbiol 2020;11:1–12.
Moonga HB, Schoustra SE, Linnemann AR et al. The art of mabisi production: A traditional fermented milk. PLoS ONE 2019;14:1–19.
van Moorsel SJ, Hahl T, Petchey OL et al. Co-occurrence history increases ecosystem stability and resilience in experimental plant communities. Ecology 2021;102:e03205.
Mpofu A, Linnemann AR, Nout MJR et al. Inactivation of bacterial pathogens in yoba mutandabota, a dairy product fermented with the probiotic Lactobacillus rhamnosus yoba. Int J Food Microbiol 2016, DOI: 10.1016/j.ijfoodmicro.2015.09.016.
Osburn ED, Badgley BD, Strahm BD et al. Emergent properties of microbial communities drive accelerated biogeochemical cycling in disturbed temperate forests. Ecology 2021;102:e03553.
Ott A, Hugi A, Baumgartner M et al. Sensory investigation of yogurt flavor perception: Mutual influence of volatiles and acidity. J Agric Food Chem 2000, DOI: 10.1021/jf990432x.
Petruzzella A, da S. S. R. Rodrigues TA, van Leeuwen CHA et al. Species identity and diversity effects on invasion resistance of tropical freshwater plant communities. Sci Rep 2020;10:5626.

Piccardi P, Alberti G, Alexander JM et al. Microbial invasion of a toxic medium is facilitated by a resident community but inhibited as the community co-evolves. ISME J 2022;16:2644–52.
Pimm SL. The complexity and stability of ecosystems. Nature 1984;307:321–6.
Pourcelot E. Développement d'une communauté modèle de levures oenologiques. 2023.
Pourcelot E, Conacher C, Marlin T et al. Comparing the hierarchy of inter- and intra-species interactions with population dynamics of wine yeast cocultures. FEMS Yeast Res 2023;23:foad039.
Pourcelot E, Vigna A, Marlin T et al. Design of a new model yeast consortium for ecological studies of enological fermentation. 2024:2024.05.06.592697.
Prosser JI, Martiny JBH. Conceptual challenges in microbial community ecology. Philos Trans R Soc B Biol Sci 2020;375:20190241.
Rainey PB, Travisano M. Adaptive radiation in a heterogeneous environment. Nature 1998;394:69–72.
Ramond J-B, Jordaan K, Díez B et al. Microbial Biogeochemical Cycling of Nitrogen in Arid Ecosystems. Microbiol Mol Biol Rev 2022;86:e00109-21.
Ratzke C, Barrere J, Gore J. Strength of species interactions determines biodiversity and stability in microbial communities. Nat Ecol Evol 2020;4:376–83.
Ratzke C, Gore J. Modifying and reacting to the environmental pH can drive bacterial interactions. PLOS Biol 2018;16:e2004248.
Renes SE, Sjöstedt J, Fetzer I et al. Disturbance history can increase functional stability in the face of both repeated disturbances of the same type and novel disturbances. Sci Rep 2020;10:11333.
Reynolds R, Hyun S, Tully B et al. Identification of microbial metabolic functional guilds from large genomic datasets. Front Microbiol 2023;14.
Rosenzweig RF, Sharp RR, Treves DS et al. Microbial evolution in a simple unstructured environment: genetic differentiation in Escherichia coli. Genetics 1994;137:903–17.
Roudil L, Russo P, Berbegal C et al. Non-Saccharomyces Commercial Starter Cultures: Scientific Trends, Recent Patents and Innovation in the Wine Sector. 2020, DOI: 10.2174/2212798410666190131103713.
Rowland I, Gibson G, Heinken A et al. Gut microbiota functions: metabolism of nutrients and other food components. Eur J Nutr 2018;57:1–24.
van Ruijven J, De Deyn GB, Berendse F. Diversity reduces invasibility in experimental plant communities: the role of plant species. Ecol Lett 2003;6:910–8.
Ruiz J, de Celis M, Diaz-Colunga J et al. Predictability of the community-function landscape in wine yeast ecosystems. Mol Syst Biol 2023;19:e11613.
Sanchez A, Bajic D, Diaz-Colunga J et al. The community-function landscape of microbial consortia. Cell Syst 2023;14:122–34.
Savageau MA. Escherichia coli Habitats, Cell Types, and Molecular Mechanisms of Gene Control. Am Nat 1983;122:732–44.
Scheuerl T, Hopkins M, Nowell RW et al. Bacterial adaptation is constrained in complex communities. Nat Commun 2020;11:1–8.
Schoustra S, van der Zon C, Groenenboom A et al. Microbiological safety of traditionally processed fermented foods based on raw milk, the case of Mabisi from Zambia. LWT 2022;171:113997.
Schoustra SE, Kasase C, Toarta C et al. Microbial Community Structure of Three Traditional Zambian Fermented Products: Mabisi, Chibwantu and Munkoyo. PLoS ONE 2013;8, DOI: 10.1371/journal.pone.0063948.
Shade A, Peter H, Allison S et al. Fundamentals of Microbial Community Resistance and Resilience. Front Microbiol 2012;3.
Sierocinski P, Bayer F, Yvon-Durocher G et al. Biodiversity–function relationships in methanogenic communities. Mol Ecol 2018;27:4641–51.

Sikombe TW, Moonga HB, Schoustra SE et al. Sensory characteristics and consumer acceptability of four variants of mabisi, a traditionally fermented Zambian dairy product. LWT 2023;188:115410.
Sivasubramaniam D, Franks AE. Bioengineering microbial communities: Their potential to help, hinder and disgust. Bioengineered 2016;7:137–44.
Sjöstedt J, Langenheder S, Kritzberg E et al. Repeated disturbances affect functional but not compositional resistance and resilience in an aquatic bacterioplankton community. Environ Microbiol Rep 2018;10:493–500.
Smid EJ, Lacroix C. Microbe–microbe interactions in mixed culture food fermentations. Curr Opin Biotechnol 2013;24:148–54.
Smid EJ, Molenaar D, Hugenholtz J et al. Functional ingredient production: Application of global metabolic models. Curr Opin Biotechnol 2005, DOI: 10.1016/j.copbio.2005.03.001.
Spus M, Wardhana Y, Wolkers-Rooijackers J et al. Lytic bacteriophages affect the population dynamics of multi-strain microbial communities. Microbiome Res Rep 2023;2:null-null.
Stroud JT, Bush MR, Ladd MC et al. Is a community still a community? Reviewing definitions of key terms in community ecology. Ecol Evol 2015;5:4757–65.
Taillandier P, Lai QP, Julien-Ortiz A et al. Interactions between Torulaspora delbrueckii and Saccharomyces cerevisiae in wine fermentation: influence of inoculation and nitrogen content. World J Microbiol Biotechnol 2014;30:1959–67.
Thébault E, Loreau M. Trophic Interactions and the Relationship between Species Diversity and Ecosystem Stability. Am Nat 2005;166:E95–114.
Tilman D. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. Proc Natl Acad Sci 2004;101:10854–61.
Tilman D, Reich PB, Isbell F. Biodiversity impacts ecosystem productivity as much as resources, disturbance, or herbivory. Proc Natl Acad Sci 2012;109:10394–7.
Treves DS, Manning S, Adams J. Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of Escherichia coli. Mol Biol Evol 1998;15:789–97.
Varahan S, Laxman S. Bend or break: how biochemically versatile molecules enable metabolic division of labor in clonal microbial communities. Genetics 2021;219:iyab109.
Vellend M. Conceptual synthesis in community ecology. Q Rev Biol 2010;85:183–206.
Vellend M. The theory of ecological communities. The Theory of Ecological Communities. Princeton University Press, 2016.
Vila JC, Jones ML, Patel M et al. Uncovering the rules of microbial community invasions. Nat Ecol Evol 2019;3:1162–71.
Voolstra CR, Ziegler M. Adapting with Microbial Help: Microbiome Flexibility Facilitates Rapid Responses to Environmental Change. BioEssays 2020;42:2000004.
Wagg C, Bender SF, Widmer F et al. Soil biodiversity and soil community composition determine ecosystem multifunctionality. Proc Natl Acad Sci 2014;111:5266–70.
Weiss AS, Niedermeier LS, von Strempel A et al. Nutritional and host environments determine community ecology and keystone species in a synthetic gut bacterial community. Nat Commun 2023;14:4780.
Williams HT, Lenton TM. Artificial selection of simulated microbial ecosystems. Proc Natl Acad Sci 2007;104:8918–23.
Wolfe BE, Dutton RJ. Fermented foods as experimentally tractable microbial ecosystems. Cell 2015;161:49–55.
WWF. Living Planet Report 2020 - Bending the Curve of Biodiversity Loss., 2020.
Xing J, Jia X, Wang H et al. The legacy of bacterial invasions on soil native communities. Environ Microbiol 2021;23:669–81.

2

Chapter 2

Influence of metabolic guilds on a temporal scale in an experimental fermented food derived microbial community

Alanna M. Leale[1], Ben Auxier[1], Eddy J. Smid[2], Sijmen Schoustra[1,3]

[1] Laboratory of Genetics, Wageningen University and Research, 6700 HB Wageningen, The Netherlands
[2] Food Microbiology, Wageningen University and Research, 6700 HB Wageningen, The Netherlands
[3] Department of Food Science and Nutrition, School of Agricultural Sciences, University of Zambia, Lusaka 10101, Zambia

Published in FEMS Microbiology Ecology: https://doi.org/10.1093/femsec/fiad112

ABSTRACT:

The influence of community diversity, which can be measured at the level of metabolic guilds, on community function is a central question in ecology. Particularly, the long-term temporal dynamic between a community's function and its diversity remains unclear. We investigated the influence of metabolic guild diversity on associated community function by propagating natural microbial communities from a traditionally fermented milk beverage diluted to various levels. Specifically, we assessed the influence of less abundant microbial types, such as yeast, on community functionality and bacterial community compositions over repeated propagation cycles amounting to approximately 100 generations. The starting richness of metabolic guilds had a repeatable effect on bacterial community compositions, metabolic profiles, and acidity. The influence of a single metabolic guild, yeast in our study, played a dramatic role on function, but interestingly not on long-term species sorting trajectories of the remaining bacterial community. Our results together suggest an unexpected niche division between yeast and bacterial communities and evidence ecological selection on the microbial communities in our system.

INTRODUCTION:

What factors affect dynamics of diversity in natural communities and how this links to community function has long been a central question in ecology (Cardinale *et al.* 2012; Gonzalez *et al.* 2020). Defining community function and community diversity at the level of species can take many forms. Community function has been estimated by quantifying parameters such as community productivity, overall metabolic output, and stability or resistance against invasion (Aubree *et al.* 2020). Community diversity captures the living organisms present and their relative abundances. This diversity can be expressed at several levels of taxonomic richness; at the broader level of functional types or metabolic guilds (i.e., capable of performing the same metabolic processes (Allen *et al.* 2023; Reynolds *et al.* 2023)), then genus or species, or more narrowly at intraspecific diversity of genotypes within a species. A general positive association between higher diversity and increased measures of community function has been shown (Balvanera *et al.* 2006), and it is typically interpreted to result from functional complementarity (Tilman 1999; Cardinale *et al.* 2007) of collections of species with different niches (Cardinale *et al.* 2012).

Ecological processes such as species sorting – the sorting of variation at the level of species along an ecological or evolutionary timescale – may alter community diversity. For instance, when a particular species or type is better adapted to the new selective condition, they may increase in relative abundance (Vellend 2010). The response to selection may depend on initial species diversity — both in the number of species present and their relative abundance – or diversity at the level of metabolic guilds. In this way, species sorting may affect overall community functioning, when species or metabolic guilds that increase or decrease in abundance are key to overall community function. While bodies of theory exist on the theme of species sorting in communities (Loeuille and Leibold 2008), experimental tests are relatively few (Langenheder and Székely 2011; Cairns *et al.* 2018, 2020; Groenenboom *et al.* 2022). Since most experimental studies on the influence of community or metabolic guild diversity on community function focus on short-term ecological timescales of one or few generations (Wagg *et al.* 2014; Sierocinski *et al.* 2018; Aubree *et al.* 2020; Gonzalez *et al.* 2020), relatively little is known on how this influence changes over multiple generations (Fiegna *et al.* 2015).

As an experimental model system, the microbial communities of fermented foods provide a powerful method to study effects of selection on species or metabolic guild diversity and on community function (Wolfe and Dutton 2015; Wolfe 2018; Alekseeva *et al.* 2021; Conacher *et al.* 2021). In these communities, function is often quantified through changes in pH and through metabolic output measured as

volatile compound production, contributing to aroma and taste, which can be further connected to known biochemical pathways (Krömer *et al.* 2004; Smid *et al.* 2005; De Filippis *et al.* 2016). Metabolic guild diversity can be traced to the level of microbial groups known to be responsible for fermentation, such as lactic acid bacteria, acetic acid bacteria, and alcohol producing yeast. Metabolic guilds or functional types can be considered equivalent to the classic ecological classification of species into guilds, lifeforms, or strategies (Louca *et al.* 2018). This is the basis of ours and others' terminology of metabolic guilds to refer to groups of microbes that can perform the same ecological functions (Allen *et al.* 2023; Reynolds *et al.* 2023). The well-studied microbial metabolic guilds, or fermentative types of fermented foods are therefore relevant communities for studying fundamental questions in ecology.

Specifically, a traditionally fermented milk beverage from Zambia, Mabisi (Moonga et al., 2020; Schoustra et al., 2013), provides a model system to study the relationship between metabolic guild diversity and functioning of the community, and how this may change over repeated cycles of sequential propagation. The moderate diversity of Mabisi, composed of six to ten dominant lactic and acetic acid bacterial species, plus numerous other low abundance types (other bacterial species, yeasts, and viruses), further facilitates experimental design and analysis (Moonga et al., 2020; Schoustra et al., 2013). The defined and measurable functional properties of Mabisi, including metabolite profiles and acidity, frame investigations of the influence between metabolic guild diversity and community function. Four fermentative types are naturally present at various levels of relative abundance. These include: 1. Alcohol producers, 2. Alcohol consumers producing acetic acid, 3. Homofermentative lactic acid producers (only lactic acid produced), and 4. Heterofermentative lactic acid producers (lactic acid, ethanol, acetic acid, and carbon dioxide produced) (Gänzle 2015). The presence and ratios of these four fermentative types, or metabolic guilds, is expected to affect community metabolic profiles due to the distinctive metabolic capabilities each community member possesses.

Here, we present an experimental test of predictions on community functioning and species sorting using microbial communities differing in diversity of metabolic guilds, species, and genotypes, which were repeatedly propagated for 16 cycles (approximately 100 generations) in milk in a laboratory environment. Our design was inspired by a prediction of altered community function at various levels of species diversity, and consequent metabolic guild diversity. Upon diluting a natural, "full" community (10^0) to medium (10^{-4}) and low (10^{-9}) diversity levels, then by using a synthetic community of five isolates from the community (Fig. 1), we progressively eliminated rare types. Specifically, yeast was a priori determined to be eliminated in low and synthetic communities as visualised by eye and calculated from previous abundance

estimates (Moonga et al., 2019; Schoustra et al., 2013). At transfers 1, 5, and 17 we then measured metabolic output as a proxy for community function, and bacterial diversity by full 16S rRNA gene amplicon sequencing. At every second transfer, pH was also measured as indication of community function since acidity strongly impacts both consumer preferences (Ott *et al.* 2000) and stability against pathogen invasion (Mpofu *et al.* 2016). With this we aimed to answer three main research questions: 1. Does altering initial metabolic guild diversity, with the associated loss of low abundance microbial types, influence community function? 2. And if so, are such shifts in community function stable over repeated propagation? 3. Does removing metabolic guilds influence bacterial species sorting trajectories over repeated cycles of propagation?

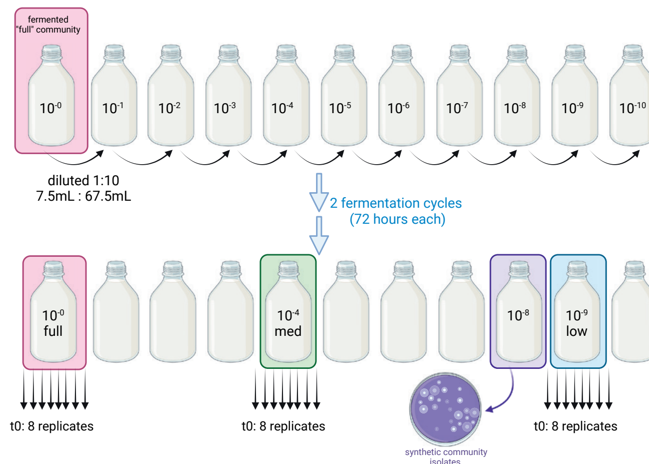


Figure 1: *Experimental design.* Dilution approach to create initial communities (t0), progressively removing rare types, arriving at four levels of initial species diversity (full, medium, and low diversity, and one synthetic community). The synthetic community was created of five single isolates, thus creating the lowest diversity community. From each diversity level eight replicates were then propagated for 16 transfers of 1% to sterile milk. Image created with BioRender.com. Abbreviations: med = medium, t0 = transfer 0 (i.e., starting inoculum).

METHODS:

Preparing communities (t0):

A fermented Mabisi sample containing its full microbial community was serially diluted up to a factor of 10^{-10} using 7.5 mL culture with 67.5 mL UHT full fat milk (Jumbo brand, Houdbare Volle Melk). Only one replicate dilution series was performed, creating one culture bottle for each dilution level. The entire volumes of these dilutions were then fermented for 72 hours at 28 °C, then 0.75 mL of final culture was transferred to 75 mL of fresh milk for a second 72-hour fermentation at 28 °C. Cultures were swirled well to mix, especially at the air-liquid phase, before transferring. After the second fermentation, the Mabisi products created from each serially diluted community were compared. A thickened, fermented product was observed up to the 10^{-9} diluted community and yeast were visibly present at the air interface up until the 10^{-4} diluted community. Visual observations informed the choice of three levels of diversity to use in the main evolution experiment: full community (10^{0} dilution), medium (10^{-4} dilution), and low (10^{-9} dilution). Initial T_0 communities were archived from the final product after two fermentation cycles without glycerol at -20 °C and with glycerol at -80 °C (1 mL culture + 0.5 mL 85% glycerol). Yeast and whey production was therefore seen only in medium and full communities.

A synthetic community of five individual isolates was also created as a fourth diversity level (Table S1). A Mabisi community that had been diluted to 10^{-8} and undergone two 72 hours fermentation cycles at 28 °C was diluted and grown aerobically at 28 °C on MRS (de Man, Rogosa, Sharpe) agar. Initially eight unique appearing morphotypes (labelled A-H) were selected from the agar plate for investigation. The community was also plated on PCA (plate count agar) and M17 agars, but the MRS provided the clearest and most diverse collection of colony types. Colonies were grown for five days in MRS broth and archived at -80 °C (1 mL culture + 0.5 mL 85% glycerol). The eight colonies were assessed for their growth in MRS broth, morphotype on MRS agar, API metabolism test (BioMérieux), and ability to acidify milk in isolation. Using the mentioned assessments, a final five colonies (A, B, C, F, G) were chosen based upon being the most dissimilar (confident at least three were unique). Each colony was streaked twice on MRS agar and inoculated five days growth at 28 °C in MRS broth (1.5 mL broth in 24 well plate). The cultures' optical density was measured (600nm wavelength reading) and then combined in volumes containing equal cell densities of each. This mixed culture formed the synthetic T_0 diversity community and was archived at -80 °C (1 mL culture + 0.5 mL 85% glycerol). Later Sanger sequencing of the full 16S rRNA gene (27F, 1492R primers) with Blast search in the NCBI database was unable to decipher the colonies to the species level (Table 1), as top identifications

were all a >99% match. However, the 16S rRNA gene sequences indicated that at least three unique types were present – two *Acetobacter* and one *Lactobacillus*.

Transferring and archiving:

For each diversity level (full, medium, low, synthetic) eight replicate lines were inoculated using their respective T_0 frozen cultures (1.5 mL total, with glycerol) in 75 mL of milk (total 32 evolution lines, plus one uninoculated milk as a negative control). Samples underwent repeated 72-hour fermentation cycles at 28 °C for 16 transfers (approximately 100 generations). For logistical purposes, after every second transfer final cultures were inoculated into fresh milk and stored at 4 °C for 24 hours before being moved to 28 °C. Therefore, a "true" cycle was completed after every second transfer (i.e., 7 days). Fresh samples of final products were archived at transfers 1, 2, 3, 5, 11, and 17, with the following archived: with glycerol at -80 °C (1.27 mL culture + 0.63 mL 85% glycerol), and without glycerol at -20 °C for DNA analysis and GC-MS analysis (~9 mL). Please note that archive labelling was done as follows: t5 = final product from t4 (i.e., the Mabisi culture used to inoculate at transfer 5), t17 = final product from t16. Due to an error in transferring, the M8 line (medium diversity, replicate 8) was lost early in the experiment, thus it is excluded from all analyses.

GC-MS Analysis:

Larger tubes of samples without glycerol at -20 °C were defrosted at 4 °C, thoroughly mixed, 1.8 mL pipetted into headspace vials, then stored at -20 °C until analysis. A sample of 1.8 mL Jumbo Brand Volle Melk and 1.8 mL Jumbo Kefir Naturel were included with every time point as controls. After incubating for 20 minutes at 60 °C, a SPME fibre (Car/DVB/PDMS, Suppelco) extracted volatiles for 20 minutes at 60 °C. Volatiles were desorbed from the fibre under the following conditions: Stabilwax- DA-Crossbond-Carbowax-polyethylene-glycol column (2 min), PTV split mode at a ratio of 1:25 (heated to 250 °C), helium carrier gas at 1.2 mL/min, GC over temperature at 35 °C (2 min) raised to 240 °C (10 C/min), kept at 240 °C (5min). Mass spectral data was collected over a range of 33-250 m/z in full scan mode with 3.0030 scans/seconds. Results were analysed with Chromeleon 7.2 CDS Software (ThermoFisher) where the following signal peaks were identified as volatile metabolites according to their elution time and mass spectral data: acetaldehyde; acetone; ethanol; hexanal; 2-heptanone; 1-butanol, 3-methyl; hexanoic acid, ethyl ester; 2-butanone, 3-hydroxy; 2-heptanol; 5-hydroxy-4-octanone; 2-butanone, 4-hydroxy; 2-nonanone; acetic acid; propanoic acid, 2-methyl; 2-undecanone; butanoic acid, 3-methyl; propanedioic acid, propyl; octanoic acid; and n-decanoic acid. MS quantification peak counts were exported to Excel.

Data was first normalised by compound using the calculation , where x is the quantification peak count for a given compound in a given sample (i.e., area under compound peak), and the median ion count across all samples for that compound. Due to the viscous nature of Mabisi, especially the synthetic community samples at later time points, accuracy and reliability of volumes were questionable. For example, pipetting the highly viscous and thick synthetic diversity samples was very difficult, creating probable variation in sampling volumes. To be conservative, data was therefore further standardised by sample using the equation , where is the mean quantification peak count and the standard deviation (following normalisation by compound) across compounds for a given sample. We realised that this removed all quantitative information about total aroma intensity between samples or diversity levels and left information only about relative compound peak heights. However, it was the justifiable approach considering the likely inaccuracies of sample volumes.

DNA extraction:

Adapted from Groenenboom et al. 2020 and Schoustra et al. 2013 (Groenenboom et al., 2020; Schoustra et al., 2013). DNA extraction was completed for all 8 replicates, of all four communities at transfers t01, t05, and t17. A sample of 1.8 mL of fermented milk was spun down (2 minutes, 12000 RPM), then the supernatant and curd removed with a sterile scoopula. Cells were re-suspended in a mix of 64 µL EDTA (0.5 M, pH 8), 160 µL Nucleic Lysis Solution, 5 µL RNAse, 120 µL lysozyme [10 mg/mL] and 40 µL pronase E [10 mg/mL] and incubated for 60 minutes at 37 °C with agitation of 350 RPM. Cells were dislodged by manually flicking the tube occasionally during incubation. This was to improve mixing with suspension mixture. Bead beading was then performed for 3 minutes (1 minute, 5 minute rest, repeated three times) with sand sized beads, then 400 µL ice-cold ammonium acetate (5 M) added, and the mixture immediately cooled on ice for 15 minutes. The mixture was spun down (13000 rpm x g, 4 min) and 700 µL of supernatant transferred to a new 1.5 mL tube. Equal volume of phenol (700 uL) was added, the tube vortexed, and then spun down (6 minutes, 12000 RPM, 4 °C). 300 µL of supernatant was transferred to a new tube, 300 µL chloroform added, then was vortexed and its content spun down (2 minutes, 12000 RPM). From here, 300 µL of supernatant was transferred to another new tube, 400uL of 2-isopropanol added and vortexed. This mixture was left at -20 C overnight to precipitate.

Following overnight precipitation, the tube was spun down (13000 rpm, 4 °C, 15 minutes), then the supernatant carefully poured out so that the DNA pellet remained. 1mL of 70% cold ethanol was added to the tube and then spun down (10 min, 12000 rpm, 4 °C). The supernatant was carefully poured out and the DNA pellet washed again with 1mL 70% cold ethanol, spun down, and supernatant poured out. The

DNA pellet was left to dry at 37 °C for 5 min, then dissolved in 20 µL of TE buffer (pH 8.0). A brief incubation (<30 min) at 37 °C improved dissolving the DNA pellet. Extracted DNA samples were stored at -20 °C. Extracted DNA was measured on Qubit using dsDNA High Sensitivity kit, and a new diluted sample made with DNA concentration 0.5 ng/µL.

Nanopore MinIon protocol:

Adapted from Beekman et al. 2022 (Beekman *et al.* 2022).

DNA concentrations in these steps were always measured on Qubit 2.0 fluorometer using dsDNA High Sensitivity Assay kit (Thermo Fisher).

Determination of number of PCR cycles:

An approximate appropriate number of PCR cycles was first determined on just three samples (1-F1, 1-L1, 1-S1), using 13, 16, 18, and 25 cycles. Primers, amounts of reagents, and PCR settings were as described below in "Tailed PCR reaction". End products were visualised on 1% agarose gel and the minimal number of cycles decided based upon the fewest number of cycles where a sufficient band was seen (determined to be 23-30 cycles).

Tailed PCR reaction:

Nanopore tailed forward: 5' TTTCTGTTGGTGCTGATATTGC-[27F] 3'
Nanopore tailed reverse: 5' ACTTGCCTGTCGCTCTATCTTC-[1492R] 3'
27F: 5' AGA GTT TGA TCC TGG CTC AG 3'
1492R: 5' TAC GGY TAC CTT GTT ACG ACT T 3'

The first step in Nanopore sequencing was a PCR reaction using Nanopore specific tailed primers. The specific number of cycles used for each sample is seen in Table S2. Two positive controls were included – the ZymoBIOMICS Microbial Community DNA Standard D6305, and a previously sequenced Mabisi sample (Groenenboom *et al.* 2020). The tailed primer PCR reaction was as follows:

Tailed reaction reagents:

1 uL – DNA [0.5ng/µL]
12.5 uL – Phusion High Fidelity PCR 2X master mix (ThermoFisher)
1.25 uL – forward tailed primer [10uM]
1.25 uL – reverse tailed primer [10uM]
9 uL – MilliQ water

Tailed cycle conditions:
98 °C 10 sec
98 °C 5 sec (~ 25X, see Table S2 for specifics)
57 °C 5 sec (~ 25X)
72 °C 30 sec (~ 25X)
72 °C 1 min
12 °C infinity

The tailed PCR reaction was performed another two times, resulting in three separate tailed primer PCR products per sample. Placement of PCR tubes in machine was adjusted for each PCR to avoid edge effects. Each amplified DNA sample was all visualised on 1% agarose gel to confirm successful amplification, then 8 µL of each PCR reaction were combined. A total of 24 µL of amplified DNA per samples was used for the clean-up.

PCR Clean-up:
For each sample, the 24 µL of amplified DNA was cleaned with 24 µL of homemade SPRI beads (i.e., 1:1 ratio) (1 ml Sera-Mag SpeedBeads (Cytiva, Marlborough, MA, USA) cleaned and dissolved in 50 ml end volume containing 2.5 M NaCL, 20 mM PEG, 10mM Tris-HCL and 1 mM EDTA) and eluted into 20 µL of MilliQ water. DNA concentration of cleaned amplicons was measured using Qubit 2.0 Fluorometer. A new dilution of 15 µL of the cleaned, amplified 16S rRNA gene PCR product was made into a new diluted sample with DNA concentration 0.5 nM.

Barcoding:
The PCR for each sample was barcoded to enable pooling using the PCR Barcoding Expansion 1-96 Kit (Oxford Nanopore Technologies). Reaction volumes were adapted from the Nanopore barcoding protocol to save in reagents used. The reaction was as follows with a unique barcode per sample:

Barcoding PCR (per reaction/sample):
0.3 µL barcode (Oxford Nanopore Technologies)
7.2 µL [0.5nM] cleaned PCR
7.5 µL LongAmp Taq 2x Master Mix (New England Biolabs)

Cycle conditions:
95 °C – 3 min (x1)
95 °C – 15 sec (x16)
62 °C – 15 sec (x16)
65 °C – 1.5 min (x16)
65 °C – 2 min (x1)
4 °C – infinity

Before pooling of the PCR barcoded products, each was visualised on 1% agarose gel. All samples were combined with 2 µL, apart from samples with fainter bands that were subjectively determined to require 3 µL (1-F1, 1-S1, 5-L1), 4 µL (5-S1, 1-M5), or 5 µL (1-S2, 1-F5, negative control). The pooled sample (total volume = 208 µL) was cleaned using homemade SPRI beads in 1:1 volumetric ratio and eluted in 200 µL of milliQ water. The DNA concentration of the cleaned, pooled sample was measured using Qubit fluorometer.

In 47 µL of milliQ water, 1 µg of the barcoded, pooled, cleaned library was prepared. From here, the library was repaired, end-prepped and adaptor ligated according to the Oxford Nanopore Technologies PCR barcoding (96) amplicons (SQK-LSK109) protocol, version "PBAC96_9069_v109_revO_14Aug2019". Reagents used were NEBNext FFPE DNA Repair Buffer (E7181A), NEBNext FFPE DNA End Repair Mix (E7182A), NEBNext Ultra II End Prep Reaction Buffer (E7183A), and NEBNext Ultra II End Prep Enzyme Mix (E7184A) (New England Biolabs).

The prepared library was loaded on a SpotON Flow Cell (FLO-MIN106D) on a MinION MK111775 sequencing device (Oxford Nanopore Technologies). Basecalling was performed shortly after using Guppy software version 6.2.4+a11ce76.

Barcodes removed due to too few reads were bc16 (1-S2), bc43 (5-L4), bc49 (1-F5), bc73 (1-F7). Also removed from final analyses were bc90 (neg), bc94 (zymo), bc86 (positive control, DNA isolated from Groenenboom et al. 2022 (Groenenboom *et al.* 2022)). Barcodes with fewer reads, but still sufficient (~7000 reads) were bc8 (5-S1), 50 (1-M5), 61 (1-F6), 64 (1-S6).

Bioinformatics:

To assign 16S rRNA gene sequences to taxonomic identity, we first downloaded the SILVA reference database (Quast *et al.* 2013). As we were interested in variation above the species level, a custom database was produced using vsearch to cluster 16S rRNA gene sequences at a 95% similarity threshold (Rognes *et al.* 2016). Reads were aligned to this database using minimap2 (Li 2018), and the best matching cluster was assigned

as the taxonomic identity for each read. Details of sequence compositions for main clusters are found in Table S3 Excel file. Clusters to which less than five reads matched were assumed to be trace contamination and discarded.

The overall frequency of clusters with <1% abundance is similar across diversity levels, as evidenced by the amount of "missing reads" to complete 1.00 abundance on bar plot (i.e., blank space at top of bar plot). Since full communities and synthetic communities have similar frequency of the rarest clusters, they are likely present due to general contamination and not true diversity being removed through our analyses. Unexpected "external" reads were identified (purple and green colouring) but are at reasonable levels as expected from low level laboratory contamination amplified by repeated PCRs during Nanopore sequencing.

We used the SILVA database reference sequences to identify types of bacteria. Our bioinformatic analyses were able to identify *Lactobacillus* and *Limosilactobacillus* types to the species level, with all clusters containing nearly 100% of sequences with same species identification (Table S3 Excel file). This was not true for *Acetobacter* types, where clusters included sequences matching to several *Acetobacter* species and these species matched to multiple clusters. For example, sequences labelled as *A. lovaniensis* are included in *Acetobacter A, B,* and *C* clusters. Hence, for the purpose of our study we only classified into general genus level groups for *Lactobacillus*, *Limosilactobacillus, and Acetobacters.* Our bioinformatic analysis could classify *Lactobacillus* and *Limosilactobacillus* clusters as homo or heterofermentative, respectively. The very rare *Limosilactobacillus* type detected in only some full and medium communities, is the sole identified heterofermentative type - *Limosilactobacillus fermentum*. The yeast present in the Mabisi communities used in this study was isolated and identified as *Geotrichum candidum* using amplification and sequencing of the ITS (internal transcribed spacer) region (top match GenBank: MK967716.1). Yeast was visibly seen growing at the air interface only in the full diversity and medium diversity communities.

Statistical analyses:

All analyses were performed in R version 4.3.1 (R Core Team 2023). Significance was defined for all analyses as p-value < 0.05. Results outputs and details of statistical analysis are found in Supplementary Material.

Measures of bacterial diversity:

The "vegan" R package functions "specnumber" and "diversity" were used to calculate bacterial species richness and Shannon diversity, respectively. A two-way ANOVA of community*transfer was performed separately for richness and Shannon diversity values using the "lm" function from "lme4" R package, type 3 partial sum of squares

(Bates *et al.* 2015). Post-hoc Tukey's pairwise comparisons with adjusted p-values were then performed using the "emmeans" function from "vegan" R package at 95% confidence intervals (Oksanen *et al.* 2022)

Acidity:

A two-way ANOVA of community*transfer was performed on pH values using the "lm" function from "lme4" R package was used, type 3 partial sum of squares (Bates *et al.* 2015). Post-hoc Tukey's pairwise comparisons with adjusted p-values were then performed using the "emmeans" function from "vegan" R package at 95% confidence intervals (Oksanen *et al.* 2022).

Community compositions:

A PERMANOVA was performed on the full data set of community compositions to evaluate the effect of community (i.e., diversity level), time point, and their interaction on community composition. The "adonis" function from "vegan" R package was used, with "bray" method (Oksanen *et al.* 2022). Post-hoc comparisons between communities were made separately for each transfer subset. The "pairwise.perm.manova" function with "Pillai test" was used from the "RVAideMemoire" R package (Herve 2023), followed by Bonferroni correction of p-values using "p.adjust" function.

Differences in dispersion of communities was tested to confirm that the significant effects from PERMANOVA were because group centroids differed and not because group variances differed. This was performed using the "betadisper" function of "vegan" R package (Oksanen *et al.* 2022)

RESULTS & DISCUSSION:

Bacterial species sorting trajectories are similar following dilution of diversity:

Figure 2 shows relative abundance of bacterial species in each of the communities at transfer 1, 5, and 17. Species profiles of the bacterial communities at transfer 1 show a clear signature of the progressive serial dilution of diversity into full diversity, medium diversity, low diversity, and synthetic communities: *Lactobacillus A* type (light turquoise) has a higher relative abundance in the medium diversity communities (mean = 0.21) compared to the full diversity communities (mean = 0.08), and is the dominant type in low diversity communities (mean = 0.52). The compositions of bacterial types at transfer 1 significantly differ (Table S13: $p < 0.05$ for all pairwise comparisons of community compositions with Bonferroni correction), yet the bacterial species richness is similar across full, medium, and low diversity communities (Figure S5: mean bacterial species richness transfer 1, full = 4.5, medium = 4.43, low = 4, synthetic = 3). Thus, while our dilution approach significantly changes initial bacterial community composition, it is not evident with the standard diversity measure of species richness. However, our dilution approach did eliminate metabolic guilds (Table 1, based from Fig. 2), and communities did differ in their initial metabolic guild diversity due to serial dilution of the full community. Most notably is the loss of yeast in the low and synthetic communities.

Table 1: Presence/absence of functional types (metabolic guilds) between diversity level treatments.

		Diversity level			
Microbial type	**Functional guild**	**full**	**medium**	**low**	**synthetic**
Geotrichum (yeast)	alcohol producer	+	+	-	-
Lactobacillus	homo fermentative	+	+	+	+
Limosilactobacillus	hetero fermentative	+	+	-	-
Acetobacter	acetic acid producer	+	+	+	+
others & genotypes		+	+/-	+/-	-

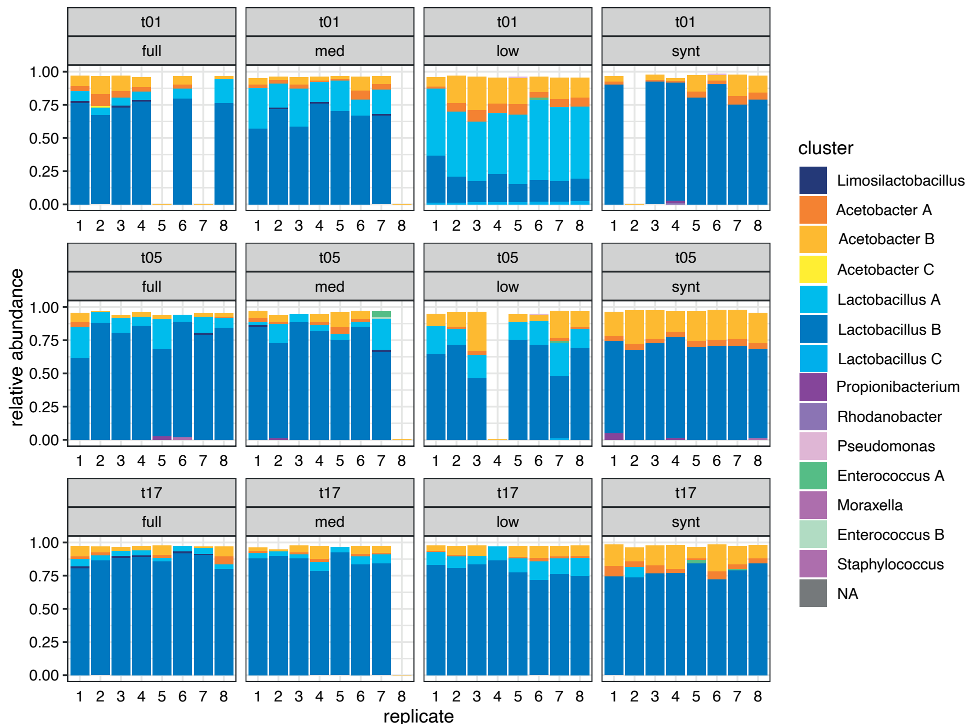


Figure 2: *Starting diversity transiently alters bacterial community composition.* Bacterial community compositions at transfers 1, 5, and 17. *Lactobacillus* clusters are shown in blue colour shades, *Acetobacter* in yellow-orange, *Limosilactobacillus* in dark navy blue, and low abundance contaminant types in green or purple. Clusters with <1% abundance were not identified nor plotted. Some replicate populations are missing due to insufficient DNA concentrations, resulting either from poor DNA extraction or library preparation. Medium community, replicate 8 was lost early in propagation and removed from all analyses. Abbreviations: med = medium, synt = synthetic; t01 = transfer 1, t05 = transfer 5, and t17 = transfer 17.

The increase in relative abundance seen at transfer 1 of *Lactobacillus A* type diminishes over time, with *Lactobacillus B* type (medium blue) dominating in all replicates and all initial dilution treatments by transfer 17 (mean relative abundance *Lactobacillus B*: full = 0.87, medium = 0.86, low = 0.79, synthetic = 0.70). However, the abundance of *Lactobacillus A* cluster remains comparatively slightly higher in low diversity communities at the end of propagation (transfer 17 mean relative abundance *Lactobacillus A*: full = 0.04, medium = 0.05, low = 0.11, synthetic = 0.00). This *Lactobacillus A* cluster was not included in the synthetic community but surprisingly appears in one synthetic community replicate at transfer 17, likely due to cross contamination during the serial propagation.

Bacterial community composition significantly differed per diversity level and time point (Table S12: PERMANOVA on community compositions, Bray method used: effect of initial diversity ($p = 0.001$, $F = 70.9$, $DF = 3$), transfer ($p = 0.001$, $F = 55.1$, $DF = 2$), diversity x transfer interaction ($p = 0.001$, $F = 27.8$, $DF = 6$)). We further conclude that significant results are due to true differences in group centroids and not dispersion within groups (Table S11: $P > 0.05$ for ANOVA community effect of distances to group centroid at all time points). Derived from the same data presented in Figure 4, NMDS plots of bacterial community compositions for transfer 1, 5, and 17 per initial diversity level (Fig. 3) demonstrate convergence of bacterial community composition across full, medium, and low diversity communities over time. Furthermore, final relative ratios of *Acetobacter A, Acetobacter B, Lactobacillus A, and Lactobacillus B* are similar across all communities at transfer 17, apart from synthetic diversity ones (Fig. S3, Table S13: $P < 0.05$ for pairwise comparisons at transfer 17 of community compositions for synthetic versus full, medium, and low).

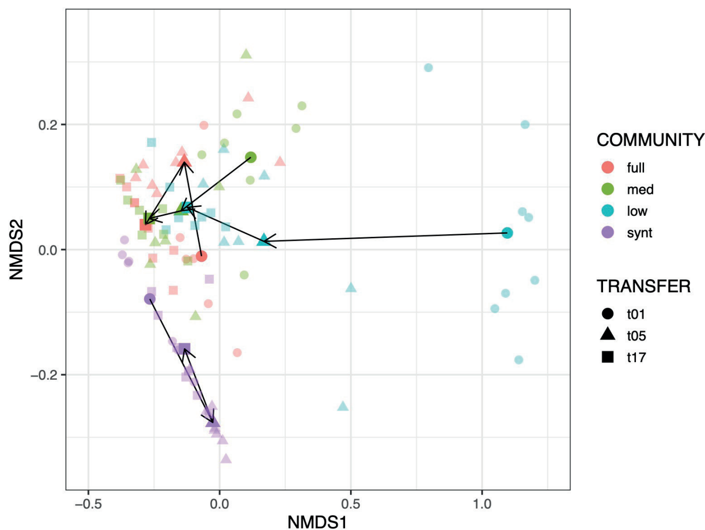


Figure 3: *Dilution of diversity largely only transiently alters bacterial community composition.* NMDS depiction of bacterial community compositions of diversity treatments at transfer 1, 5, and 17. Solid coloured points are centroids of data points for community*transfer grouping. Arrows connect centroids over time. Abbreviations: med = medium, synt = synthetic; t01 = transfer 1, t05 = transfer 5, and t17 = transfer t17.

Acetobacters are at highest relative abundance in the low diversity communities early in propagation (mean relative abundance *Acetobacters* transfer 1, full = 0.12 (sd = 0.07), medium = 0.08 (sd = 0.05), low = 0.23 (sd = 0.08), synthetic = 0.11 (sd = 0.08)), but

synthetic communities demonstrate the highest proportions of *Acetobacter* types by transfer 17 (mean relative abundance *Acetobacters* transfer 17, full = 0.06 (sd =0.05), medium = 0.05 (sd =0.04), low = 0.08 (sd =0.04), synthetic = 0.18 (sd =0.06)). However, the relative abundances of *Acetobacter* versus *Lactobacillus* types do converge across full, medium, and low communities by transfer 17 (Fig. S2).

Although removal of metabolic guilds across our starting communities affected initial and final metabolic function (see Results section below), it interestingly did not have a large effect on the outcome of species sorting trajectories of the remaining bacterial community over 16 cycles of propagation. This is especially true when focusing on the relative ratios of the major bacterial metabolic guilds - acetic acid versus homofermentative lactic acid bacteria, which were surprisingly uniform across full diversity, medium diversity, and low diversity communities after 16 cycles of propagation. If replicate communities of higher diversity diverged more so in their compositions overtime, then our results would have aligned with the theory that "diversity begets diversity" via greater niche construction (Madi *et al.* 2020). Reversely, our results could have exhibited replicate communities of lower diversity diverging from one another, which would support the concept that greater diversity restricts possible trajectories due to less available niche space (van Moorsel *et al.* 2021). We observe neither outcome. The minimal effect of diversity on sorting trajectories in our experiment could be explained by high functional redundancies (Allison and Martiny 2008) in the major bacterial types (i.e., acetic acid and lactic acid bacteria).

Yeast drives division of function shown in metabolic profiles:

Metabolic profiles of replicate communities at four levels of diversity in metabolic guilds show that two groupings persist across the three time points analysed, with metabolic profiles of full and medium diversity communities clustering together, versus low and synthetic (Fig. 4). The divide exists along principal component 1 (PC1) but not the second component (PC2). PC1 explains between 53.1% (transfer 5) to 65.5% (transfer 1) of the variation, while PC2 explains between 11.9% (transfer 1) to 19.6% (transfer 17). A division in metabolic profiles between full and medium communities is observed along PC2 at transfer 1 (Fig. 4a), disappearing at transfer 5 (Fig. 4b). The metabolic profiles of communities at the four diversity levels inconsistently exhibit relatively larger or smaller spread in data points across the time points but replicates of the synthetic community (purple ellipses) arguably show closer clustering in metabolic profiles throughout the experiment.

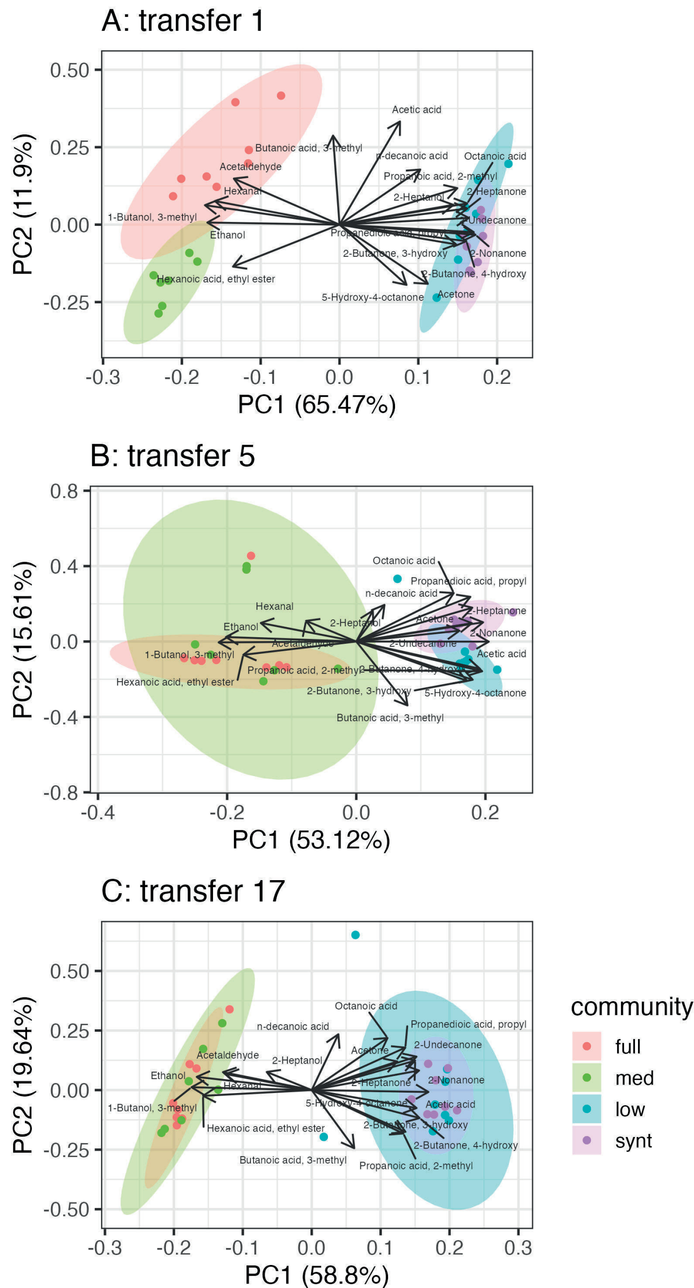


Figure 4: *Diversity of metabolic profiles persist through propagation. Compounds associated with yeast metabolism primarily impact divisions in metabolic profile.* PCA analysis of metabolic profiles compiled from 19 compounds using GC-MS analyses at transfer 1 (A), 5 (B), and 17 (C). Dots represent individual replicates. Ellipses are 95% confidence intervals. Direction of vectors indicate the compound's contribution to either principal component, whereas the vector length indicates the amount of variation explained by the two plotted principal components. Yeast associated compounds: ethanol, acetaldehyde, hexanoic acid-ethyl ester, 1-butanol 3-methyl. Abbreviations: med = medium, synt = synthetic.

Although at minority abundances in typical Mabisi communities (Schoustra et al., 2013), yeasts contribute a unique metabolism since they are the sole alcohol fermentative type and are known to produce other distinct compounds such as methyl-esters. We thus predicted that presence of yeast in the full and medium diversity communities or, conversely, its absence in the low diversity and synthetic communities, would significantly influence community function. Initial metabolic profiles across all time points provide support for this expectation, where alcohol and ester compounds typically linked to yeast metabolism are found. The presence of yeast in only the full and medium diversity communities is reflected by presence of ethanol, acetaldehyde, and hexanoic acid-ethyl ester among the main metabolites detected. Acetaldehydes and esters are a known products of yeast metabolism (Liu and Pilone 2000; Dzialo *et al.* 2017). An additional compound with vector contributions towards the full and medium community grouping is 1-butanol-3-methyl, which is a breakdown product of leucine via the Ehrlich pathway found in various yeast species (Szudera-Kończal *et al.* 2020). The heterofermentative type *Limosilactobacillus* is also absent in low diversity and synthetic communities, whose presence is expected to alter metabolite profile; however, since this overlaps with the presence/absence of yeast in the communities, we are unable to disentangle which metabolite changes are specific to the presence or absence of heterofermentative types.

All diversity levels contained the two dominant fermentative types – homofermentative lactic acid bacteria (*Lactobacillus* A, B) and acetic acid bacteria *(Acetobacter* A, B) – which our data suggest are the core bacterial members for community function at the level of metabolite production, in addition to yeast. While yeast appear as a core microbial community member in our study's Mabisi sample, interestingly, not all natural Mabisi products contain yeast (Moonga et al., 2020; Schoustra et al., 2013). Microbial community profiles of Mabisi samples vary across regions and processors in Zambia. There is past and ongoing research to link microbial community compositions to processing methods (Moonga *et al.* 2020) and consumer preferences.

Furthermore, metabolic profiles overlap between low diversity and synthetic communities (Fig. 4), yet their bacterial compositions differ with loss of intraspecies genotype diversity and *Lactobacillus A* in synthetic communities (Fig. 2, Table 1). Hence, in Mabisi communities, intraspecies bacterial genotype compositions do not appear to strongly influence metabolic profiles, suggesting functional redundancies in bacterial metabolic capacities. Others have found low functional redundancy in microbial communities for particular functions, for example methane production (Sierocinski *et al.* 2018), where there is considerable impact by loss of any species. Our results align regarding yeast but not for species or genotypes of the abundant lactic acid and acetic acid bacteria. However, we measured broader, more generalist

functions of overall aroma profiles and acidity. High functional redundancy of Mabisi microbial communities compared to environmental systems such as feedstock fermenters (Sierocinski *et al.* 2018; Wagg *et al.* 2019) or soil (Wagg *et al.* 2014, 2019) is likely due to immense microbial diversity in these other systems and their specific measures of function.

Functional redundancy exhibited for acidity:

Figure 5 shows pH of cultures after a full cycle of growth over the course of the experiment. The overall pH differences across diversity treatments ranged only between approximately pH 3.4 to 3.8, yet with a significant effect of community, transfer, and their interaction on acidity (Table S8: $p < 2.2e-16$ for all effects). Full diversity communities and medium diversity communities maintained similar pH throughout propagation (Table S9: Tukey's pairwise comparison full vs. medium $p>0.05$ for all time points). Low diversity communities maintained a higher pH until a drop between transfer 15 and 17 where they converge to a comparable value as the full and medium communities (Table S9: Tukey's pairwise comparison low vs. full and low vs. medium both $p<0.05$ for time points 1 through 15, then both $p>0.05$ for transfer 17). Whereas the pH in synthetic communities was significantly higher than all others by transfer 17 (Table S9: $p<0.0001$ for Tukey's pairwise comparisons of synthetic vs. full, medium, and low at transfer 17). The pH dropped significantly between transfer 1 and transfer 17 for full, medium, and low diversity treatments (Table S10: Tukey's pairwise comparison transfer 1 vs. transfer t17 $p<0.05$ for full, medium, and low communities).

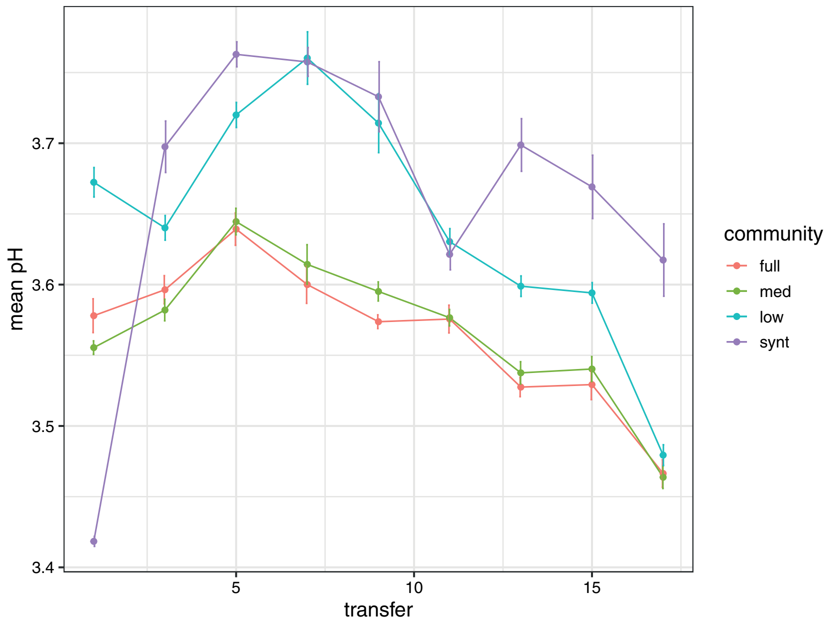


Figure 5: *Loss of diversity initially increases pH, which decreases over continual propagation*. Measures of pH of full diversity, medium diversity, low diversity and synthetic communities over 16 rounds of serial propagation. Mean pH across replicates of the same community diversity level measured at every second transfer. Points show mean value for the 8 replicate communities (except medium diversity treatment with 7 replicates), vertical bars show standard error of the mean. Abbreviations: med = medium, synt = synthetic.

High functional redundancy is observed in the bacterial part of the microbial community (Wagg *et al.* 2019; Gralka *et al.* 2020), thus we predicted that general acidification properties of lactic acid and acetic acid bacteria would be maintained in diluted communities, regardless of the loss of yeast or rare genotypes (Wagg *et al.* 2019; White *et al.* 2020). We found support for this prediction; however, a very slightly higher pH was observed for the synthetic, hence lowest diversity, community. This modest increase suggests that the presence of yeast and rare bacterial genotypes does not impact community acidification. A dynamic jump in pH in synthetic diversity communities between transfers 1 and 3 is not surprising since this community was more naïve in its member structure and substrate (these isolates had been pre-cultured in MRS broth). The range in pH between diversity levels is minimal (~pH 3.4 to 3.85) but may still represent a selective force on the microbial communities. Mabisi acidity increases quickly during fermentation, with the most notable pH drop between 24 to 48 hours (Groenenboom *et al.* 2020). Our observed trend of continually decreasing pH may be explained by the preferential propagation of, and thus selection for, persisting abundant types in the acidic conditions created after 72 hours of fermentation. Increased acidity is a common outcome of bacterial community domestication

and repeated back-slop propagation of fermentation starter cultures (Bachmann *et al.* 2011; Spuś 2016; van Kerrebroeck *et al.* 2016). It would be interesting to test how much lower the pH would reduce with further propagation and when or if pH stabilisation would be reached. Evolving communities could be improving resource use, consequently excreting more metabolites due to larger supported population sizes, further lowering the pH over additional repeated transfers.

CONCLUSIONS:

In this study we assessed how progressively diluting a natural microbial community altered community function and how this function, as well as bacterial community composition, would subsequently change on an ecological time-scale due to selection upon repeated cycles of propagation. By exploring the relationship between community function and metabolic guild diversity over time, we observed repeatable changes of replicate lineages for metabolic profiles and acidity related to starting diversity levels of metabolic guilds. These results show a clear division in metabolic profiles with full diversity and medium diversity communities on one hand, and low diversity and synthetic communities together on the other hand. This division was sustained throughout repeated rounds of propagation. Further, we found that changes in bacterial community composition (i.e., bacterial species sorting trajectories) over repeated cycles of propagation generally resulted in a convergence of bacterial communities to the same composition, irrespective of initial metabolic guild diversity.

Most surprising was the seeming lack of influence of yeast presence or absence on bacterial compositions after 16 cycles of propagation. The convergence of bacterial community compositions, regardless of the presence of yeast communities, suggests yeast and bacteria exist in sufficiently unique niches of resource use. Evidence for strong division in resources and hence lack of influence of yeast on bacterial community composition in a fermented food was surprising and contrasts against previous findings showing metabolic associations between the two (Mendes *et al.* 2013; Suharja, Henriksson and Liu 2014; Ponomarova *et al.* 2017; Blasche *et al.* 2021; Xu *et al.* 2021). However, these investigations mostly focus on *S. cerevisiae* and specific *Lactobacilli* species or strains; the yeast in our system is identified as *Geotrichum candidum*. If *Lactobacilli* versus *Acetobacters* had differing metabolic associations with yeast, we would expect shifts in their relative abundances following the removal of *G. candidum,* but this was not observed. We can thus hypothesise that in the Mabisi system, yeast do not affect bacterial growth by consuming end products of bacterial metabolism such as lactate, thus are simply secondary in the metabolic route. In addition to more cross-feeding like interactions, yeast and bacteria could be

in resource competition, in which case the removal of yeast would open niche space for certain metabolic guilds or species. Those whose niche space previously overlapped more with yeast, or were poor competitors, would expectedly establish at higher abundances in the absence of yeast; this is not what we observe. Overall, our results support that amongst bacterial metabolic types in our community, they are all in similar, at most very weak, resource competition with yeast since bacterial community compositions converge across all community diversities regardless of yeast's presence. A next step to further elucidate the influence of yeast on community function in Mabisi is to add isolated *G. candidum* to the low and synthetic communities, then compare metabolic profiles and acidity. The reverse direction could also be taken, by eliminating yeasts from the full and medium communities with fungicide, which would maintain rare bacterial types.

We interpret the observed repeatability during the propagation cycles between replicate lineages and convergence in bacterial communities to be evidence of ecological selection (Vellend 2010, 2016) in our study system. In evolutionary biology research, repeated or convergent ratios of genotypes under a common environment is interpreted as evidence for adaptive selection (Hughes, 1999); we take the same interpretation for our results but at the ecological level of ratios of bacterial metabolic guilds. Currently, we can only hypothesise about the sources of ecological selection in our study, with possibilities including our chosen laboratory conditions (temperature, fermentation time and associated acidity, supermarket milk content) and unspecified biotic species interactions. While prokaryotes and bacteriophage are known members of the microbial community of our natural experimental system (Schoustra et al., 2013), we have focussed our analyses on species sorting in the bacterial part of the microbial community since bacteria are present in all communities regardless of starting diversity (i.e., dilution) treatment. There is growing evidence of highly repeatable species sorting trajectories in microbial community assembly, both in varied environmental samples (Goldford *et al.* 2018; Diaz-Colunga *et al.* 2022) and communities of isolated strains (Cairns *et al.* 2020). Our research similarly finds convergence overtime in community compositions but instead after disturbance of initial diversity.

Studying functionality responses on longer time scales after an erosion of metabolic guilds, species, and genotypic diversity remains an underexplored topic that we made initial steps here to explore. The ecological processes of species sorting in a community can be paralleled to evolutionary dynamics of genotypic changes within a species (Vellend 2010, 2016). Understanding how changes across hierarchical levels — from genes, to communities, to ecosystems — influence one another is undoubtedly complicated and it merits further steps to elucidate. In this study we focused on ecological processes without investigating genotypic changes. However, we foresee

exciting future avenues of research to unravel the role of ecology versus evolution in altering communities and their functionality. Creating a synthetic community comprised of four fermentative types (i.e., metabolic guilds) that grow in a defined media, combined with whole genome sequencing and/or metagenomics could allow identification of novel genotypes, and is an exciting avenue to better explore ecological-evolutionary dynamics.

AUTHOR CONTRIBUTIONS:

Experimental design and conceptual ideas were developed by AML and SS. Experiments and data collection were performed by AML with help of Dr. Francisca Reyes Marquez. Data analysis was completed by AML and BA, with BA creating the bioinformatic analysis pipeline. Manuscript written by AML and SS with editing support of BA and EJS.

FUNDING:

This work was supported by a Wageningen University and Research Interdisciplinary Research and Education Fund (INREF).

ACKNOWLEDGMENTS:

We greatly thank Dr. Francisca Reyes Marquez for her laboratory expertise and support with DNA extractions and Nanopore MinIon sequencing protocols. We also thank Judith Wolkers-Rooijackers for running the GC-MS machine and helping with its associated data analysis. Thank you as well to Mariska Beekman for sharing her experience with Nanopore sequencing and its bioinformatic analysis.

DATA AVAILABILITY:

Source data files and codes available on GitHub (amleale/diversity_function_mabisi): https://github.com/amleale/diversity_function_mabisi.git . Raw amplicon sequence data for the communities are available under NCBI BioProject PRJNA937001, with each barcode available separately under SAMN35661351-SAMN35661446. Isolate 16S rRNA gene sequences of the five community members in the synthetic community are available under NCBI GenBank accessions OQ284045-OQ284049.

CONFLICT OF INTEREST:

The authors declare they have no conflict of interest.

REFERENCES:

Alekseeva AY, Groenenboom AE, Smid EJ et al. Eco-Evolutionary Dynamics in Microbial Communities from Spontaneous Fermented Foods. Int J Environ Res Public Health 2021;18:10093.

Allen B, Gonzalez-Cabaleiro R, Ofiteru ID et al. Diversity and metabolic energy in bacteria. FEMS Microbiol Lett 2023;370:fnad043.

Allison SD, Martiny JBH. Resistance, resilience, and redundancy in microbial communities. Proc Natl Acad Sci 2008;105:11512–9.

Aubree F, David P, Jarne P et al. How community adaptation affects biodiversity–ecosystem functioning relationships. Ecol Lett 2020;23:1263–75.

Bachmann H, Molenaar D, Kleerebezem M et al. High local substrate availability stabilizes a cooperative trait. ISME J 2011;5:929–32.

Balvanera P, Pfisterer AB, Buchmann N et al. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. Ecol Lett 2006;9:1146–56.

Bates D, Mächler M, Bolker B et al. Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw 2015;67:1–48.

Beekman MM, Donner SH, Litjens JJ et al. Do aphids in Dutch sweet pepper greenhouses carry heritable elements that protect them against biocontrol parasitoids? Evol Appl 2022;15:1580–93.

Blasche S, Kim Y, Mars RA et al. Metabolic cooperation and spatiotemporal niche partitioning in a kefir microbial community. Nat Microbiol 2021;6:196–208.

Cairns J, Jokela R, Becks L et al. Repeatable ecological dynamics govern the response of experimental communities to antibiotic pulse perturbation. Nat Ecol Evol 2020;4:1385–94.

Cairns J, Jokela R, Hultman J et al. Construction and characterization of synthetic bacterial community for experimental ecology and evolution. Front Genet 2018;9:312.

Cardinale BJ, Duffy JE, Gonzalez A et al. Biodiversity loss and its impact on humanity. Nature 2012;486:59–67.

Cardinale BJ, Wright JP, Cadotte MW et al. Impacts of plant diversity on biomass production increase through time because of species complementarity. Proc Natl Acad Sci 2007;104:18123–8.

Conacher C, Luyt N, Naidoo-Blassoples R et al. The ecology of wine fermentation: a model for the study of complex microbial ecosystems. Appl Microbiol Biotechnol 2021;105:3027–43.

De Filippis F, Genovese A, Ferranti P et al. Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. Sci Rep 2016;6:1–11.

Diaz-Colunga J, Lu N, Sanchez-Gorostiaga A et al. Top-down and bottom-up cohesiveness in microbial community coalescence. Proc Natl Acad Sci 2022;119.

Dzialo MC, Park R, Steensels J et al. Physiology, ecology and industrial applications of aroma formation in yeast. FEMS Microbiol Rev 2017;41:S95–128.

Fiegna F, Moreno-Letelier A, Bell T et al. Evolution of species interactions determines microbial community productivity in new environments. ISME J 2015;9:1235–45.

Gänzle MG. Lactic metabolism revisited: metabolism of lactic acid bacteria in food fermentations and food spoilage. Curr Opin Food Sci 2015;2:106–17.

Goldford JE, Lu N, Bajić D et al. Emergent simplicity in microbial community assembly. Science 2018;361:469–74.

Gonzalez A, Germain RM, Srivastava DS et al. Scaling-up biodiversity-ecosystem functioning research. Ecol Lett 2020;23:757–76.

Gralka M, Szabo R, Stocker R et al. Trophic interactions and the drivers of microbial community assembly. Curr Biol 2020;30:R1176–88.

Groenenboom AE, van den Heuvel J, Zwaan BJ et al. Species dynamics in natural bacterial communities over multiple rounds of propagation. Evol Appl 2022;15:1766–75.

Groenenboom AE, Shindano J, Cheepa N et al. Microbial population dynamics during traditional production of Mabisi, a spontaneous fermented milk product from Zambia: a field trial. World J Microbiol Biotechnol 2020;36:184.

Herve M. RVAideMemoire: Testing and Plotting Procedures for Biostatistics. 2023.

Hughes AL, others. Adaptive Evolution of Genes and Genomes. Oxford University Press, USA, 1999.

van Kerrebroeck S, Bastos FCC, Harth H et al. A low pH does not determine the community dynamics of spontaneously developed backslopped liquid wheat sourdoughs but does influence their metabolite kinetics. Int J Food Microbiol 2016;239:54–64.

Krömer JO, Sorgenfrei O, Klopprogge K et al. In-Depth Profiling of Lysine-Producing Corynebacterium glutamicum by Combined Analysis of the Transcriptome, Metabolome, and Fluxome. J Bacteriol 2004;186:1769–84.

Langenheder S, Székely AJ. Species sorting and neutral processes are both important during the initial assembly of bacterial communities. ISME J 2011;5:1086–94.

Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 2018;34:3094–100.

Liu S-Q, Pilone GJ. An overview of formation and roles of acetaldehyde in winemaking with emphasis on microbiological implications. Int J Food Sci Technol 2000;35:49–61.

Loeuille N, Leibold MA. Evolution in Metacommunities: On the Relative Importance of Species Sorting and Monopolization in Structuring Communities. Am Nat 2008;171:788–99.

Louca S, Polz MF, Mazel F et al. Function and functional redundancy in microbial systems. Nat Ecol Evol 2018;2:936–43.

Madi N, Vos M, Murall CL et al. Does diversity beget diversity in microbiomes? Weigel D, Kemen E, Wolfe BE (eds.). eLife 2020;9:e58999.

Mendes F, Sieuwerts S, de Hulster E et al. Transcriptome-based characterization of interactions between Saccharomyces cerevisiae and Lactobacillus delbrueckii subsp. bulgaricus in lactose-grown chemostat cocultures. Appl Environ Microbiol 2013;79:5949–61.

Moonga HB, Schoustra SE, Linnemann AR et al. The art of mabisi production: A traditional fermented milk. PloS One 2019;14:e0213541.

Moonga HB, Schoustra SE, Van den Heuvel J et al. Composition and diversity of natural bacterial communities in mabisi, a traditionally fermented milk. Front Microbiol 2020;11:1816.

van Moorsel SJ, Hahl T, Petchey OL et al. Co-occurrence history increases ecosystem stability and resilience in experimental plant communities. Ecology 2021;102:e03205.

Oksanen J, Blanchet FG, Friendly M et al. vegan: Community Ecology Package. R package version 2.5-7. 2020. 2022.

Ponomarova O, Gabrielli N, Sévin DC et al. Yeast creates a niche for symbiotic lactic acid bacteria through nitrogen overflow. Cell Syst 2017;5:345–57.

Quast C, Pruesse E, Yilmaz P et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 2013;41:D590–6.

R Core Team. R: A language and environment for 2023.

Reynolds R, Hyun S, Tully B et al. Identification of microbial metabolic functional guilds from large genomic datasets. Front Microbiol 2023;14.

Rognes T, Flouri T, Nichols B et al. VSEARCH: a versatile open source tool for metagenomics. PeerJ 2016;4:e2584.

Schoustra SE, Kasase C, Toarta C et al. Microbial community structure of three traditional zambian fermented products: mabisi, chibwantu and munkoyo. PloS One 2013a;8:e63948.

Sierocinski P, Bayer F, Yvon-Durocher G et al. Biodiversity–function relationships in methanogenic communities. Mol Ecol 2018;27:4641–51.

Smid EJ, Molenaar D, Hugenholtz J et al. Functional ingredient production: application of global metabolic models. Curr Opin Biotechnol 2005;16:190–7.

Spuś M. Mixed culture engineering for steering starter functionality. 2016.

Suharja AA, Henriksson A, Liu S-Q. Cerevisiae on Viability of Probiotic Rhamnosus in Fermented Milk under Ambient Conditions. J Food Process Preserv 2014;38:326–37.

Szudera-Kończal K, Myszka K, Kubiak P et al. The use of sour and sweet whey in producing compositions with pleasant aromas using the mold galactomyces geotrichum: identification of key odorants. J Agric Food Chem 2020;68:10799–807.

Tilman D. The Ecological Consequences of Changes in Biodiversity: A Search for General Principles. Ecology 1999;80:1455–74.

Vellend M. Conceptual synthesis in community ecology. Q Rev Biol 2010;85:183–206.

Vellend M. The theory of ecological communities. The Theory of Ecological Communities. Princeton University Press, 2016.

Wagg C, Bender SF, Widmer F et al. Soil biodiversity and soil community composition determine ecosystem multifunctionality. Proc Natl Acad Sci 2014;111:5266–70.

Wagg C, Schlaeppi K, Banerjee S et al. Fungal-bacterial diversity and microbiome complexity predict ecosystem functioning. Nat Commun 2019;10:1–10.

White L, O'Connor NE, Yang Q et al. Individual species provide multifaceted contributions to the stability of ecosystems. Nat Ecol Evol 2020;4:1594–601.

Wolfe BE. Using cultivated microbial communities to dissect microbiome assembly: challenges, limitations, and the path ahead. Msystems 2018;3:e00161-17.

Wolfe BE, Dutton RJ. Fermented foods as experimentally tractable microbial ecosystems. Cell 2015;161:49–55.

Xu Z, Lu Z, Soteyome T et al. Polymicrobial interaction between Lactobacillus and Saccharomyces cerevisiae: Coexistence-relevant mechanisms. Crit Rev Microbiol 2021;47:386–96.

SUPPLEMENTARY MATERIAL (Chapter 2):

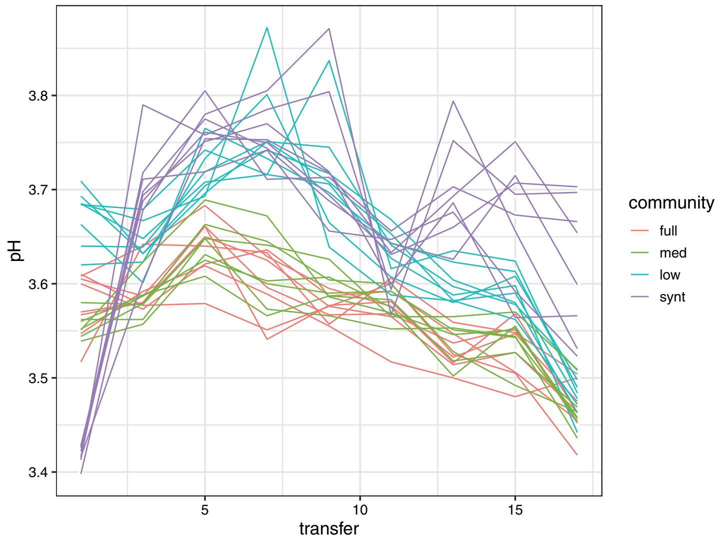


Figure S1*:* Measured pH at every second transfer for individual replicate communities over 16 rounds of propagation. Abbreviations: med = medium, synt = synthetic.

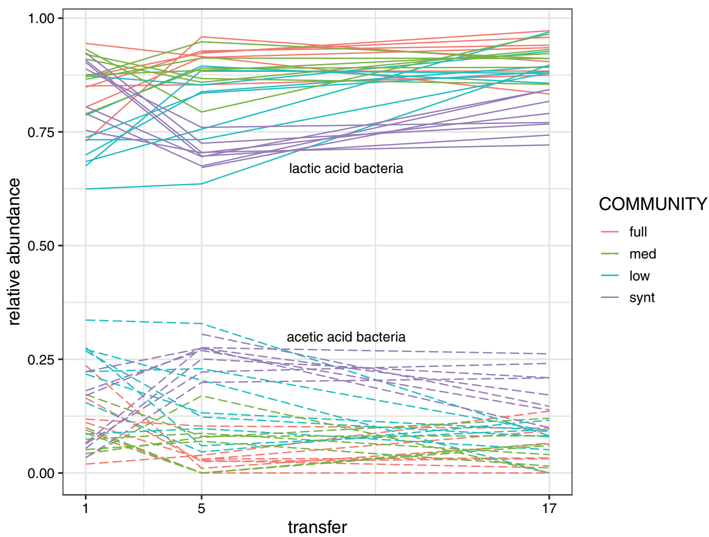


Figure S2: *Community compositions become more similar over time, with the exception of synthetic community with more acetic acid bacteria.* Relative abundance of lactic acid and acetic acid bacteria overtime. Lines follow each replicate over time. Solid lines: lactic acid bacteria, dashed lines: acetic acid bacteria. Lactic acid bacteria grouping: *Limosilactobacillus, Lactobacillus A, B, and C*. Acetic acid bacteria grouping: *Acetobacter A, B, and C*. Four replicate populations missing as in Figure 4. Abbreviations: med = medium, synt = synthetic

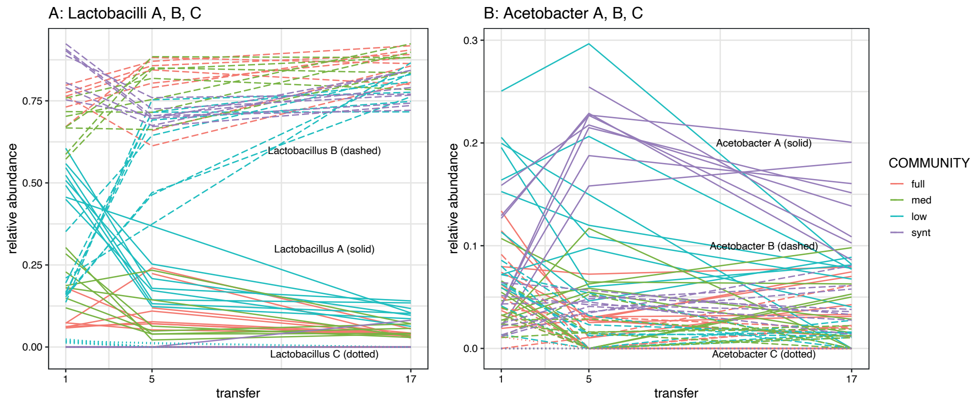


Figure S3: *Relative abundances of lactic acid and acetic acid bacterial types converge overtime.* Lines follow each replicate overtime. Four replicate populations missing as in Figure 4. Abbreviations: med = medium, synt = synthetic

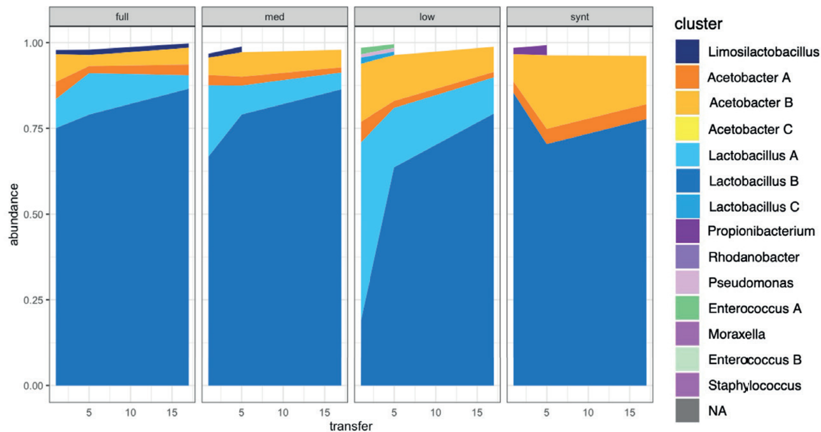


Figure S4: *Mean relative abundances of bacterial clusters overtime in each diversity treatment. Alternative visualization of data from Fig. 2 in manuscript. Lactobacillus* clusters are shown in blue colour shades, *Acetobacter* in yellow-orange, *Limosilactobacillus* in dark navy blue, and low abundance contaminant types in green or purple. Means calculated from eight replicate lineages per diversity treatment. Some replicate populations are missing due to insufficient DNA concentrations, resulting either from poor DNA extraction or library preparation. Abbreviations: med = medium, synt = synthetic.

Table S1: Identity of the five isolates used in synthetic diversity communities with percent identity probabilities in brackets.

Isolate ID	NCBI Blast top hits (27F – 1492R 16S region)
A	*Acetobacter fabarum (100), A. lovaniensis (99.9), A. ghanensis (99.8)*
B	*Acetobacter orientalis (100), A. cibinongensis (99.5), A. cerevisiae (99.4)*
C	*Acetobacter orientalis (99.9), A. cibinongensis (99.5), A. cerevisiae (99.4)*
F	*Lactobacillus gallinarum (99.4), L. helveticus (99.4), L. acidophilus (99.0)*
G	*Lactobacillus gallinarum (99.6), L. helveticus (99.6), L. acidophilus (99.0)*

Table S2: Number of cycles used for PCR amplification in Nanopore sequencing.

Transfer-community	PCR cycles used
1-full	25
1-med	25
1-low	23
1-syn	25
5-full	25
5-med	25
5-low	24
5-synt	30
17-full	24
17-med	24
17-low	24
17-synt	24

Table S3: Separate excel file with details of bioinformatic clustering.

Table S4: Mean relative abundances of bacterial clusters per diversity treatment and transfer. Means calculated from eight replicate lineages.

COMMUNITY	full	full	full	low	low	low	med	med	med	synt	synt	synt
TRANSFER	**t01**	**t05**	**t17**	**t01**	**t05**	**t17**	**t01**	**t05**	**t17**	**t17**	**t01**	**t05**
Limosilactobacillus	0.01	0.02	0.01	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00
Lactobacillus A	0.08	0.12	0.04	0.52	0.17	0.11	0.21	0.08	0.05	0.08	0.00	0.00
Lactobacillus B	0.75	0.79	0.87	0.19	0.64	0.79	0.67	0.79	0.86	0.78	0.85	0.70
Lactobacillus C	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Acetobacter A	0.05	0.02	0.03	0.06	0.02	0.02	0.03	0.03	0.02	0.04	0.03	0.04
Acetobacter B	0.08	0.03	0.05	0.17	0.13	0.07	0.05	0.07	0.05	0.14	0.08	0.21
Acetobacter C	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Propionibacterium	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.03
Rhodanobacter	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Pseudomonas	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Staphylococcus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Moraxella	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Enterococcus A	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.04	0.00	0.02	0.00	0.00
Enterococcus B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00

Statistical Analyses (Chapter 2):

Significant, or non-significant, p-values that are previously referred to in the main text are highlighted in yellow.

1. BACTERIAL SPECIES RICHNESS

Stats were performed on same data set used for barplot Figure 2 and NMDS Figure 3. (ie., clusters of <1% abundance removed).

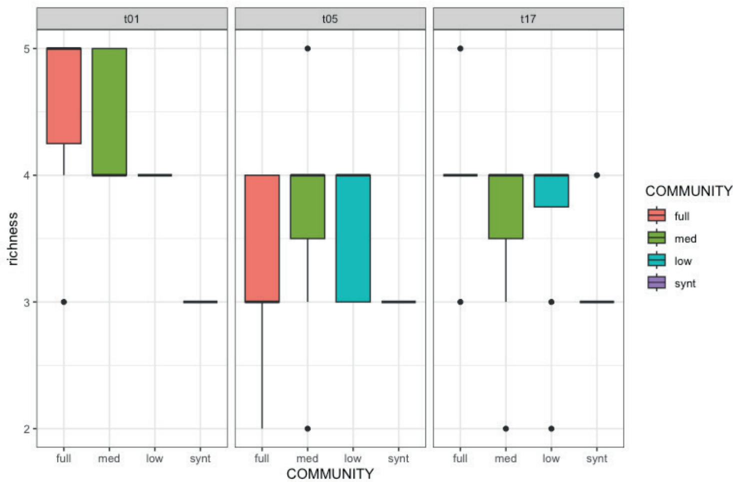


Figure S5: *calculated species richness.* Abbreviations: med = medium, synt = synthetic, t01 = transfer 1, t05 = transfer 5, t17 = transfer 17.

1a)

```
> rich_lm <- lm(richness ~ COMMUNITY*TRANSFER, data = pivoted)
> emmeans(rich_lm, pairwise ~ COMMUNITY|TRANSFER)
P value adjustment: tukey method for comparing a family of 4 estimates
```

Table S5: Output of Tukey's pairwise comparisons between communities for species richness.

transfer	contrast	estimate	SE	df	t.ratio	p.value
t01	full-med	0.0714	0.325	77	0.220	0.9962
t01	full-low	0.5000	0.316	77	1.583	0.3940
t01	full-synt	15.000	0.325	77	4.611	**0.0001**
t01	med-low	0.4286	0.303	77	1.416	0.4932
t01	med-synt	14.286	0.313	77	4.571	**0.0001**
t01	low-synt	10.000	0.303	77	3.305	**0.0077**
t05	full-med	-0.4643	0.303	77	-1.534	0.4223
t05	full-low	-0.3214	0.303	77	-1.062	0.7134
t05	full-synt	0.2500	0.292	77	0.855	0.8277
t05	med-low	0.1429	0.313	77	0.457	0.9680
t05	med-synt	0.7143	0.303	77	2.361	0.0936
t05	low-synt	0.5714	0.303	77	1.888	0.2415
t17	full-med	0.4286	0.303	77	1.416	0.4932
t17	full-low	0.3750	0.292	77	1.283	0.5766
t17	full-synt	0.8750	0.292	77	2.993	0.0190
t17	med-low	-0.0536	0.303	77	-0.177	0.9980
t17	med-synt	0.4464	0.303	77	1.475	0.4572
t17	low-synt	0.5000	0.292	77	1.710	0.3254

2. BACTERIAL SHANNON DIVERSITY

Stats were performed on same data set used for barplot Figure 2 and NMDS Figure 3. (ie., clusters of <1% abundance removed).

2a)

```
> shan_lm <- lm(shannon ~ COMMUNITY*TRANSFER, data = pivoted)
> emmeans(shan_lm, pairwise ~ COMMUNITY|TRANSFER)
```

P value adjustment: tukey method for comparing a family of 4 estimates

Table S6: Output of Tukey's pairwise comparisons between communities for calculated Shannon diversity.

transfer	contrast	estimate	SE	df	t.ratio	p.value
t01	full-med	-0.0892	0.0873	77	-1.021	**0.7378**
t01	full-low	-0.3533	0.0848	77	-4.166	0.0005
t01	full-synt	0.3523	0.0873	77	4.034	0.0007
t01	med-low	-0.2641	0.0813	77	-3.250	0.0091
t01	med-synt	0.4415	0.0839	77	5.261	<.0001
t01	low-synt	0.7056	0.0813	77	8.684	<.0001
t05	full-med	-0.0474	0.0813	77	-0.584	0.9367
t05	full-low	-0.3465	0.0813	77	-4.265	0.0003
t05	full-synt	-0.1884	0.0785	77	-2.399	0.0857
t05	med-low	-0.2991	0.0839	77	-3.564	0.0035
t05	med-synt	-0.1409	0.0813	77	-1.734	0.3133
t05	low-synt	0.1582	0.0813	77	1.947	0.2175
t17	full-med	0.0247	0.0813	77	0.303	0.9902
t17	full-low	-0.1780	0.0785	77	-2.268	0.1147
t17	full-synt	-0.1671	0.0785	77	-2.129	0.1532
t17	med-low	-0.2027	0.0813	77	-2.494	0.0688
t17	med-synt	-0.1918	0.0813	77	-2.360	0.0937
t17	low-synt	0.0109	0.0785	77	0.139	0.9990

Chapter 2

2b)

```
> shan_lm <- lm(shannon ~ COMMUNITY*TRANSFER, data = pivoted)
> emmeans(shan_lm, pairwise ~ TRANSFER|COMMUNITY)
P value adjustment: tukey method for comparing a family of 3 estimates
```

Table S7: Output of Tukey's pairwise comparisons between timepoints for calculated Shannon diversity.

community	contrast	estimate	SE	df	t.ratio	p.value
full	t01-t05	0.2427	0.0848	77	2.862	0.0148
full	t01-t17	0.3112	0.0848	77	3.670	**0.0013**
full	t05-t17	0.0685	0.0785	77	0.872	0.6592
medium	t01-t05	0.2844	0.0839	77	3.389	0.0031
medium	t01-t17	0.4250	0.0839	77	5.064	**<.0001**
medium	t05-t17	0.1406	0.0839	77	1.675	0.2212
low	t01-t05	0.2494	0.0813	77	3.070	0.0082
low	t01-t17	0.4865	0.0785	77	6.197	**<.0001**
low	t05-t17	0.2370	0.0813	77	2.917	0.0127
synthetic	t01-t05	-0.2980	0.0813	77	-3.667	0.0013
synthetic	t01-t17	-0.2082	0.0813	77	-2.563	**0.0327**
synthetic	t05-t17	0.0897	0.0785	77	1.143	0.4907

3. pH / ACIDITY

```
> ph_data2$transfer<- as.factor(ph_data2$transfer)
> lm_ph <- lm(pH ~ community*transfer, data = ph_data2)
```

Anova Table (Type III tests)

Response: pH

Table S8: Output of ANOVA analysis for measured pH values.

	SumSq	Df	F value	Pr(>F)
(Intercept)	102.417	1	80309.661	< 2.2e-16
community	0.264	3	68.938	< 2.2e-16
transfer	0.164	8	16.069	< 2.2e-16
community:transfer	0.579	24	18.924	< 2.2e-16
Residuals	0.310	243		

3a)

```
> emmeans(lm_ph, pairwise ~ community|transfer)
```

P value adjustment: tukey method for comparing a family of 4 estimates

Table S9: Output of Tukey's pairwise comparisons between communities for measured pH values.

transfer	contrast	estimate	SE	df	t.ratio	p.value
t01	full-med	0.022571	0.0185	243	1.221	**0.6140**
t01	full-low	-0.094375	0.0179	243	-5.285	<.0001
t01	full-synt	0.159625	0.0179	243	8.940	<.0001
t01	med-low	-0.116946	0.0185	243	-6.328	<.0001
t01	med-synt	0.137054	0.0185	243	7.415	<.0001
t01	low-synt	0.254000	0.0179	243	14.225	<.0001
t03	full-med	0.014375	0.0185	243	0.778	**0.8645**
t03	full-low	-0.043750	0.0179	243	-2.450	0.0706
t03	full-synt	-0.101125	0.0179	243	-5.664	<.0001
t03	med-low	-0.058125	0.0185	243	-3.145	0.0100
t03	med-synt	-0.115500	0.0185	243	-6.249	<.0001
t03	low-synt	-0.057375	0.0179	243	-3.213	0.0081
t05	full-med	-0.005321	0.0185	243	-0.288	**0.9917**
t05	full-low	-0.080750	0.0179	243	-4.522	0.0001
t05	full-synt	-0.123625	0.0179	243	-6.924	<.0001
t05	med-low	-0.075429	0.0185	243	-4.081	0.0004
t05	med-synt	-0.118304	0.0185	243	-6.401	<.0001
t05	low-synt	-0.042875	0.0179	243	-2.401	0.0795
t07	full-med	-0.014286	0.0185	243	-0.773	**0.8666**
t07	full-low	-0.160250	0.0179	243	-8.975	<.0001
t07	full-synt	-0.157500	0.0179	243	-8.821	<.0001
t07	med-low	-0.145964	0.0185	243	-7.898	<.0001
t07	med-synt	-0.143214	0.0185	243	-7.749	<.0001
t07	low-synt	0.002750	0.0179	243	0.154	0.9987

Table S9: Continued

transfer	contrast	estimate	SE	df	t.ratio	p.value
t09	full-med	-0.021393	0.0185	243	-1.157	**0.6542**
t09	full-low	-0.140500	0.0179	243	-7.869	<.0001
t09	full-synt	-0.159125	0.0179	243	-8.912	<.0001
t09	med-low	-0.119107	0.0185	243	-6.444	<.0001
t09	med-synt	-0.137732	0.0185	243	-7.452	<.0001
t09	low-synt	-0.018625	0.0179	243	-1.043	0.7243
t11	full-med	-0.000946	0.0185	243	-0.051	**1.0000**
t11	full-low	-0.054750	0.0179	243	-3.066	0.0128
t11	full-synt	-0.045750	0.0179	243	-2.562	0.0533
t11	med-low	-0.053804	0.0185	243	-2.911	0.0204
t11	med-synt	-0.044804	0.0185	243	-2.424	0.0752
t11	low-synt	0.009000	0.0179	243	0.504	0.9581
t13	full-med	-0.010071	0.0185	243	-0.545	**0.9478**
t13	full-low	-0.071375	0.0179	243	-3.997	0.0005
t13	full-synt	-0.171250	0.0179	243	-9.591	<.0001
t13	med-low	-0.061304	0.0185	243	-3.317	0.0058
t13	med-synt	-0.161179	0.0185	243	-8.721	<.0001
t13	low-synt	-0.099875	0.0179	243	-5.594	<.0001
t15	full-med	-0.011036	0.0185	243	-0.597	**0.9329**
t15	full-low	-0.064875	0.0179	243	-3.633	0.0019
t15	full-synt	-0.139875	0.0179	243	-7.834	<.0001
t15	med-low	-0.053839	0.0185	243	-2.913	0.0203
t15	med-synt	-0.128839	0.0185	243	-6.971	<.0001
t15	low-synt	-0.075000	0.0179	243	-4.200	0.0002
t17	full-med	0.002661	0.0185	243	0.144	**0.9989**
t17	full-low	-0.013000	0.0179	243	-0.728	**0.8858**
t17	full-synt	-0.151000	0.0179	243	-8.457	<.0001
t17	med-low	-0.015661	0.0185	243	-0.847	**0.8317**
t17	med-synt	-0.153661	0.0185	243	-8.314	<.0001
t17	low-synt	-0.138000	0.0179	243	-7.729	<.0001

3b)

> emmeans(lm_ph, pairwise ~ transfer|community)

P value adjustment: tukey method for comparing a family of 9 estimates

Table S10: Output of Tukey's pairwise comparisons between timepoints for measured pH values. Full list of comparisons (beyond transfer3-transfer5) not shown for brevity.

community	contrast	estimate	SE	df	t.ratio	p.value
full	transfer1-transfer3	-0.01837	0.0179	243	-1.029	0.9827
full	transfer1-transfer5	-0.06125	0.0179	243	-3.430	0.0200
full	transfer1-transfer7	-0.02200	0.0179	243	-1.232	0.9487
full	transfer1-transfer9	0.00425	0.0179	243	0.238	10.000
full	transfer1-transfer11	0.00237	0.0179	243	0.133	10.000
full	transfer1-transfer13	0.05050	0.0179	243	2.828	0.1126
full	transfer1-transfer15	0.04875	0.0179	243	2.730	0.1428
full	transfer1-transfer17	0.11163	0.0179	243	6.252	**<.0001**
full	transfer3-transfer5	-0.04288	0.0179	243	-2.401	0.2880
medium	transfer1-transfer3	-0.02657	0.0191	243	-1.392	0.9000
medium	transfer1-transfer5	-0.08914	0.0191	243	-4.670	0.0002
medium	transfer1-transfer7	-0.05886	0.0191	243	-3.083	0.0571
medium	transfer1-transfer9	-0.03971	0.0191	243	-2.081	0.4891
medium	transfer1-transfer11	-0.02114	0.0191	243	-1.108	0.9727
medium	transfer1-transfer13	0.01786	0.0191	243	0.936	0.9907
medium	transfer1-transfer15	0.01514	0.0191	243	0.793	0.9970
medium	transfer1-transfer17	0.09171	0.0191	243	4.805	**0.0001**
medium	transfer3-transfer5	-0.06257	0.0191	243	-3.278	0.0322
low	transfer1-transfer3	0.03225	0.0179	243	1.806	0.6781
low	transfer1-transfer5	-0.04763	0.0179	243	-2.667	0.1653
low	transfer1-transfer7	-0.08787	0.0179	243	-4.921	0.0001
low	transfer1-transfer9	-0.04188	0.0179	243	-2.345	0.3195
low	transfer1-transfer11	0.04200	0.0179	243	2.352	0.3155
low	transfer1-transfer13	0.07350	0.0179	243	4.116	0.0017
low	transfer1-transfer15	0.07825	0.0179	243	4.382	0.0006
low	transfer1-transfer17	0.19300	0.0179	243	10.809	**<.0001**
low	transfer3-transfer5	-0.07988	0.0179	243	-4.473	0.0004
synthetic	transfer1-transfer3	-0.27913	0.0179	243	-15.632	<.0001
synthetic	transfer1-transfer5	-0.34450	0.0179	243	-19.294	<.0001
synthetic	transfer1-transfer7	-0.33913	0.0179	243	-18.993	<.0001
synthetic	transfer1-transfer9	-0.31450	0.0179	243	-17.614	<.0001
synthetic	transfer1-transfer11	-0.20300	0.0179	243	-11.369	<.0001
synthetic	transfer1-transfer13	-0.28037	0.0179	243	-15.702	<.0001
synthetic	transfer1-transfer15	-0.25075	0.0179	243	-14.043	<.0001
synthetic	transfer1-transfer17	-0.19900	0.0179	243	-11.145	**<.0001**
synthetic	transfer3-transfer5	-0.06538	0.0179	243	-3.661	0.0092
synthetic	transfer3-transfer7	-0.06000	0.0179	243	-3.360	0.0250
synthetic	transfer3-transfer9	-0.03537	0.0179	243	-1.981	0.5580
synthetic	transfer3-transfer11	0.07612	0.0179	243	4.263	0.0010
synthetic	transfer3-transfer13	-0.00125	0.0179	243	-0.070	10.000
synthetic	transfer3-transfer15	0.02838	0.0179	243	1.589	0.8100
synthetic	transfer3-transfer17	0.08013	0.0179	243	4.487	**0.0004**

4. PERMANOVA: checking group dispersions

Completed with betadisper in vegan. Community profile data for each transfer was analysed separately. We conclude that distance to centroid for each community (i.e., dispersion) does not significantly differ (anova on distances p>0.05).

> beta <- wide %>% filter(**TRANSFER** == "tXX")

> dst <- dist(beta[,5:18])

> wide.bd <- betadisper(dst, beta$**COMMUNITY**)

> anova(wide.bd)

Analysis of Variance Table

Response: Distances

Table S11: Output of PERMANOVA test for differences in dispersion of community compositions between communities (i.e., full, medium, low, synthetic).

transfer		Df	SumSq	Mean Sq	F value	Pr (>F)
t01	Groups	3	0.002787	0.0009289	0.2725	0.8446
t01	Residuals	24	0.081823	0.0034093		
t05	Groups	3	0.030427	0.0101422	15.495	0.2255
t05	Residuals	26	0.170186	0.0065456		
t17	Groups	3	0.0008032	0.00026774	0.2481	0.862
t17	Residuals	27	0.0291365	0.00107913		

5. PERMANOVA and post-hoc pairwise comparisons differences in community compositions.

> adonis2(wide[,5:18] ~ COMMUNITY*TRANSFER, data = wide[,1:4], method = "bray")

Permutation test for adonis under reduced model

Terms added sequentially (first to last)

Permutation: free

Number of permutations: 999

Table S12: Output of PERMANOVA tests of 16S rRNA community compositions.

	Df	SumOfSqs	R2	F	Pr(>F)
COMMUNITY	3	15.159	0.37527	70.882	0.001
TRANSFER	2	0.7855	0.19445	55.093	0.001
COMMUNITY:TRANSFER	6	11.892	0.29439	27.802	0.001
Residual	77	0.5489	0.13589		
Total	88	40.396	100.000		

For post hoc comparisons, PERMANOVA of each transfer analysed separately, then pairwise comparisons made of distances using Bonferroni correction.

```
> wide_subset <- wide %>% filter(TRANSFER == "tXX")
> dist_dml <- vegan::vegdist(x=as.matrix(wide_subset[,5:18]), method="bray", binary=FALSE, diag=TRUE, upper=TRUE, na.rm=FALSE)
> y_permanova <- vegan::adonis2(dist_dml ~ COMMUNITY, data=wide_subset, permutations=999, method="euclidean", parallel=4)
> permtst <- RVAideMemoire::pairwise.perm.manova(resp = dist_dml, fact = wide_subset$COMMUNITY, test = "Pillai", nperm = 999, progress = TRUE, p.method = "none")
> df <- reshape2::melt(permtst$p.value)
> colnames(df) <- c("comm1", "comm2", "pvalue")
> df <- df[-which(is.na(df$pvalue)), ]
> df$pvalue.adj <- p.adjust(p = df$pvalue, method = "bonferroni", n = length(df$pvalue))
```

Table S13: Output of pairwise comparisons between communities of their 16S rRNA community compositions.

transfer	comm1	comm2	pvalue	pvalue.adj
t01	med	full	0.007	0.042
t01	low	full	0.001	0.006
t01	synt	full	0.005	0.030
t01	low	med	0.001	0.006
t01	synt	med	0.001	0.006
t01	synt	low	0.002	0.012
t05	med	full	0.634	**1.000**
t05	low	full	0.007	0.042
t05	synt	full	0.001	0.006
t05	low	med	0.010	**0.060**
t05	synt	med	0.001	0.006
t05	synt	low	0.001	0.006
t17	med	full	0.506	**1.000**
t17	low	full	0.003	0.018
t17	synt	full	0.001	0.006
t17	low	med	0.009	**0.054**
t17	synt	med	0.002	0.012
t17	synt	low	0.001	0.006

3

Chapter 3

Shifts and rebound in microbial community function following repeated introduction of a novel species

Alanna M. Leale[1], Francisca Reyes Marquez[1], Bas Zwaan[1], Eddy J. Smid[2], Sijmen Schoustra[1,3]

[1] Laboratory of Genetics, Wageningen University and Research, 6700 HB Wageningen, The Netherlands
[2] Food Microbiology, Wageningen University and Research, 6700 HB Wageningen, The Netherlands
[3] Department of Food Science and Nutrition, School of Agricultural Sciences, University of Zambia, Lusaka 10101, Zambia

In revision at Oikos. Pre-print online: https://doi.org/10.1101/2024.04.05.588252

ABSTRACT:

Natural microbial communities continually encounter novel species that may successfully establish or simply be transient, yet both outcomes can alter the resident community composition and function. Preserving natural microbial communities and innovating synthetic ones requires insight on the immediate and long-term impact of species introductions on both composition and function. For instance, it remains unclear whether there are gradual and long-term impacts from repeated invasions where the introduced species fails to establish – so-called failed invaders. To investigate the persistent impacts by failed invaders, we present an experimental test of community stability over multiple generations against repeated novel species introduction. We propagated a natural microbial community from a traditional fermented milk beverage for approximately 100 generations, with or without, repeated introduction of *Escherichia coli* at each transfer. Community function was determined by metabolic profiling, and we observed alterations therein immediately after *E. coli* introduction, followed by recovery, or rebound once ceased. In contrast to this proxy of community function, changes in the bacterial community composition were never detected. Our results evidence that community composition and function do not necessarily respond in parallel to an introduced species, potentially due to genotypic changes below species level detection or metabolic plasticity. Our work shows an ability for functional recovery in microbial communities and contributes insight on long-term community stability to sustained disturbances.

INTRODUCTION:

Natural ecosystems and the species communities they harbour are commonly faced with perturbations by invading novel species, whether gradual and continuous or in singular events. Thinking of invasive species typically recalls non-native plants, insects, or animals proliferating in a habitat. For instance, zebra mussels in North American fresh lakes (Karatayev and Burlakova 2022), the common water hyacinth plant in tropical ponds globally (Datta *et al.* 2021), or the emerald ash borer beetle in European and North American forests (McCullough 2020), are just a few well-known cases when introduced species successfully, or devastatingly *per se*, established in a novel environment. The same processes occurring at a macroecological level also apply to microbial ecosystems that harbour communities of co-existing microbes. The spoilage of food (Oro *et al.* 2019), some harmful cyanobacteria blooms (Bolius, Wiedner and Weithoff 2019), biofertilizers (Gu *et al.* 2019), probiotics (Albright *et al.* 2020), or infection of human microbiota (Libertucci and Young 2019), can all be considered instances where the resident microbial community is confronted by novel species (Kinnunen *et al.* 2016). But even if unsuccessful at establishing, an introduced species can change the structure of interactions between members of a microbial community (Padrón *et al.* 2009) and furthermore influence overall community function (Mallon *et al.* 2018), even shifting the community to an alternative stable compositional state (Amor, Ratzke and Gore 2020). Further, community species composition is found to be decoupled from community metabolic function (Louca *et al.* 2016, 2018) meaning that responses to perturbation at one level may not align with the other. As the great value of microbial communities for both natural ecosystems and human-desired functions is rapidly being recognised (Kinnunen *et al.* 2016; Sivasubramaniam and Franks 2016), it is of interest to know how they may respond to perturbations and how responses in community composition *versus* function contrast, including the persistence of shifts to alternative states once perturbation stops.

Protecting natural microbial communities or engineering novel synthetic ones both require understanding immediate and long-term responses of community composition and function to perturbation by novel introduced species. The stability of microbial communities against perturbation by a novel species has been investigated with multiple experimental and modelling approaches across study systems (Mallon, Elsas and Salles 2015; Vila *et al.* 2019; Kruk *et al.* 2021; Kurkjian, Akbari and Momeni 2021; Philippot, Griffiths and Langenheder 2021). Much research has focused on a single introduction of a novel species to microbial communities over a few generations of a single growth cycle (Allison and Martiny 2008; Shade *et al.* 2012; van Elsas *et al.* 2012; Mallon *et al.* 2018; Amor, Ratzke and Gore 2020), with few studies assessing multiple rounds of propagation and assessments at multiple time

points (Xing *et al.* 2021; Albright *et al.* 2022). Single "flash" introduction events are undoubtably important to understand as they occur in natural ecosystems; however, it is also common that novel species are repeatedly introduced into a community with compound effects (Blackburn, Lockwood and Cassey 2015; Philippot, Griffiths and Langenheder 2021). Specifically, a key finding motivating our study is that even a transient novel species, or a failed invader, can induce a lasting alteration to a bacterial community's composition and resource use breadth (Yao *et al.* 2014; Mallon *et al.* 2018; Amor, Ratzke and Gore 2020; Xing *et al.* 2021). The permanence of induced changes is relevant to investigate – do microbial communities rebound in metabolic function and composition, or remain altered after repeated exposure to a failed invader? To test this, observations on community stability against novel species over extended timescales of multiple generations are needed (Philippot, Griffiths and Langenheder 2021; Albright *et al.* 2022), and would help elucidate unanswered questions about the persistent impacts by failed invaders.

We used a microbial community of a traditionally fermented milk beverage from Zambia – Mabisi – as a model system to investigate the impact of repeated introduction of a novel species on community function. Microbial communities of Mabisi sampled in the field typically contain a diverse yet manageable composition of approximately 6 to 12 species of bacteria, but samples can vary depending on location and processing methods (Schoustra *et al.* 2013; Moonga *et al.* 2020). As an introduced species, we chose a non-pathogenic *E. coli* strain for its relevance to food safety in spontaneously fermented foods (Capozzi *et al.* 2017) and ease of laboratory use for selective plating at a low biohazard level. *E. coli* is also used in other microbial community stability research (van Elsas *et al.* 2012; Mallon *et al.* 2018; Xing *et al.* 2021) due to its ability to grow, or at least persist, in secondary environments from its primary environment of the vertebrate gut (Savageau 1983; van Elsas *et al.* 2007). Mabisi is traditionally made by spontaneous fermentation of raw, unpasteurised milk, thus its microbiological safety is called into question and its sale restricted to household level in Zambia (Materia *et al.* 2021). However, past research exhibited that Mabisi communities are highly resistant against food borne pathogens including *Enterobacteriaceae* (Schoustra *et al.* 2022); we therefore did not predict *E. coli* establishment. Expected failed *E. coli* invasion in Mabisi therefore provided a suitable approach to evaluate how the resident community species composition and function would be impacted over repeated failed invasion events, as well as community recovery afterwards.

The varying macro-ecological and microbial systems for researching ecological stability has led to inconsistent terminology (Donohue *et al.* 2016; Kinnunen *et al.* 2016; Philippot, Griffiths and Langenheder 2021), so we here stablish our definitions. We refer to the resident community as the members and their relative abundances

stably present in a natural setting, whereas the novel species as one that is absent or at negligible abundances in the initial resident community, hence considered introduced. We do not use the term "invader" due to the term's implication of domination, proliferation, or at least establishment, whereas simply the alteration of community structure, regardless of relative abundance (i.e., establishment) of the novel species, is of focus in our work. In line with other microbial ecology studies on stability, our study assessed both resistance and recovery in a microbial community (Allison and Martiny 2008; Shade *et al.* 2012; Philippot, Griffiths and Langenheder 2021), doing so in terms of metabolic function and community composition (Table 1). Resistance is the degree of change in a variable, for example, how much does the community composition or function shift following disturbance. Recovery is if, and the degree that, the community returns to its pre-disturbance state (Shade *et al.* 2012). Metabolic profiles are an especially suitable proxy for community function in fermented foods since they indicate compounds associated with aroma and taste (Alekseeva *et al.* 2021), while inferring to active pathways and gene expression across the community (Smid *et al.* 2005; De Filippis *et al.* 2016). Acidity was additionally measured since it influences consumer preference (Ott *et al.* 2000; Sikombe *et al.* 2023) but more importantly, the prevention of establishment of food borne pathogens in fermented foods (Mpofu *et al.* 2016).

Table 1: community stability related definitions and measures, as used in this paper.

Term	Definition related to this study's results
resident community	composition of community members stably present in a natural setting
introduced/novel species	species that is absent or at negligible abundances in the initial resident community
resistance	change in metabolic function and community composition following disturbance. Little or no change = high resistance.
recovery	return to a pre-disturbance state of metabolic function and community composition once disturbance is stopped.
community composition	full-length 16S rRNA gene sequences clustered at 95% similarity
community function	metabolic profiles (GC-MS analyses) and acidity

In this study, our general research aim was to investigate repeated introduction of a novel species to a microbial community over many generations. Our experimental approach had two parts: we first tested if, when repeatedly introduced, *E. coli* is unable to establish, but still can shift the Mabisi resident community composition and function? Secondly, we were interested in the Mabisi resident community's recovery to rebound in composition and function once introduction was stopped. To do so, we propagated replicate natural Mabisi communities for 18 growth cycles (approximately 110 mitotic generations) in the laboratory under two treatments; a "Control" treatment without the addition of non-pathogenic *E. coli*, and an "Introduction"

treatment where *E. coli* was added at every transfer (Fig. 1). After 14 transfers, the "Introduction" treatment replicates were then split to generate a third "Recovery" treatment where *E. coli* was no longer added for the remaining transfers. Community composition (species identification through full 16S rRNA gene) and metabolic profiles (GC-MS) were measured throughout to assess the community resistance and recovery to repeated novel species introduction.

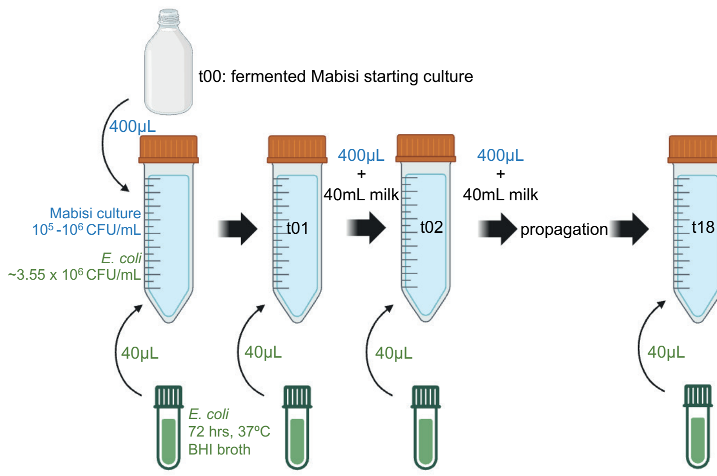


Figure 1: *experimental set up of "Introduction" treatment*. "Control" treatment was identical except without addition of *E. coli*. Both treatments were performed in replicates of six. At transfer 14, the "Introduction" replicates were used to create a "Recovery" treatment where *E. coli* was no longer added at each transfer. Please note that the same *E. coli* stock was used across transfers (i.e., *E. coli* was not evolving). Figure created with BioRender.com.

METHODS:

Preparation for transfer 0:

A fermented Mabisi sample stored at -80 C in 50%v/v glycerol was defrosted and 800uL inoculated into 40mL of UHT full fat milk (Campina Langlekker volle melk) then incubated unshaken at 28ºC for 72 hours. On the same day, a glycerol stock of *E. coli* (strain DSM498) was inoculated into 2mL of BHI broth (VWR 84626.0500) and grown for 72 hours shaken at 37ºC. The choice of 72 hours for *E. coli* growth culture was done instead of 24 or 48 hours because of feasibility with laboratory working hours and weekends. *E. coli* cultures were therefore always at stationary phase when inoculating into milk with Mabisi. Cell density of *E. coli* after 72-hours in BHI broth was estimated by plate counts to be 3.55 x 10^9 CFU/mL (Fig. S3).

Transfer 0 inoculations:

From the 72-hour fermented Mabisi culture and *E. coli* culture, two treatments were created with six replicate lines each. Following inversion and mixing of the Mabisi culture, 400 uL of culture was added to 40 mL of fresh full milk for both the Control (CON) and Introduction (INT) replicates. The Introduction treatment additionally received 40 uL of *E. coli* culture, producing an estimated 3.55 x 10^6 CFU/mL of *E. coli* at Transfer 0 in 40 mL of milk; this concentration is greater than field study estimates of enterobacteria found in raw milk used for Mabisi production in Zambia (Schoustra *et al.* 2022). Samples then underwent repeated 72-hour fermentation cycles at 28 ºC for 18 transfers (approximately 110 generations); fermented cultures were mixed then 400 uL transferred to 40 mL of new milk, with 40 uL of new 72-hour *E. coli* culture also added to Introduction treatment replicates at every transfer. For logistical purposes, after every second transfer final cultures were inoculated into fresh milk and stored at 4 ºC for 24 hours before being moved to 28 ºC. Therefore, a "true" cycle was completed after every second transfer (i.e., 7 days). The 72-hour *E. coli* culture was repeatedly grown fresh from the same glycerol stock for each inoculation in the Introduction treatment. At transfer 14, cultures from each replicate of the Introduction treatment were used to create a third "Recovery" treatment, where fresh *E. coli* culture was not added at the remaining transfers. Fresh samples of final products were archived at transfers 1, 5, 9, 12, and 19, with the following archived: with glycerol at -80 ºC (1.27 mL culture + 0.63 mL 85% glycerol), and without glycerol at -20 ºC for DNA analysis and GC-MS analysis (remaining volume in tube). Please note that archive labelling was done as follows: t9 = final product from transfer 8 (i.e., the Mabisi culture used to inoculate at transfer 9), t19 = final product from transfer 18. The pH was also measured at every second transfer.

GC-MS Analysis of VOCs:

Larger tubes of samples without glycerol at -20 ºC were defrosted at 4 ºC, thoroughly mixed, 1.8 mL pipetted into headspace vials, then stored at -20 ºC until analysis. A sample of 1.8 mL Jumbo Brand Volle Melk and 1.8 mL Jumbo Kefir Naturel were included with every time point as controls. After incubating for 20 minutes at 60 ºC, a SPME fibre (Car/DVB/PDMS, Suppelco) extracted volatiles for 20 minutes at 60 ºC. Volatiles were desorbed from the fibre under the following conditions: Stabilwax- DA-Crossbond-Carbowax-polyethylene-glycol column (2 min), PTV split mode at a ratio of 1:25 (heated to 250 ºC), helium carrier gas at 1.2 mL/min, GC over temperature at 35 ºC (2 min) raised to 240 ºC (10 C/min), kept at 240 ºC (5min). Mass spectral data was collected over a range of 33-250 m/z in full scan mode with 3.0030 scans/seconds. Results were analysed with Chromeleon 7.2 CDS Software (ThermoFisher) where the following signal peaks were identified as volatile metabolites according to their elution time and mass spectral data: acetaldehyde; Ethyl Acetate; Ethanol; 2-Heptanone; 1-Butanol, 3-methyl-; 2-Butanone, 3-hydroxy-; Propanoic

acid, 2-hydroxy-, ethyl ester, (S)-; Acetic anhydride; Pentanal, 2-methyl-; Acetic acid; Propanoic acid, 2-methyl-; Butanoic acid, 3-methyl-; Hexanoic acid; Phenylethyl Alcohol; Octanoic Acid; Benzoic acid. MS quantification peak counts were exported to Excel. Data was first normalised by compound using the calculation , where x is the quantification peak count for a given compound in each sample (i.e., area under compound peak), and the median ion count across all samples for that compound.

DNA extraction:

Larger tubes of samples without glycerol at –20 °C were defrosted at 4°C, thoroughly mixed, and ~1.75 mL pipetted into centrifuge tube and stored at -20 C. On the day of extraction, the sample was defrosted and spun down (2 min, 13000 g), then the supernatant and curd removed with a sterile scoopula. Cells were re-suspended in a mix of 64 µL EDTA (0.5 M, pH 8), 160 µL Nucleic Lysis Solution, 5 µL RNAse, 120 µL lysozyme [10 mg/mL] and 40 µL pronase E [10 mg/mL] and incubated for 60 minutes at 37 °C with agitation of 350 rpm. Cells were dislodged by manually flicking the tube occasionally during incubation, to improve mixing with suspension mixture. Bead beading (MP Biomedical FastPrep-25 5G device) was then performed for 3 minutes (1 minute, 5-minutes rest, repeated three times) with sand sized beads, then 400 µL ice-cold ammonium acetate (5 M) added, and the mixture immediately cooled on ice for 15 minutes. The mixture was spun down (4 min, 13000 g) and 650 µL transferred to a 2 mL flat bottom 96-well plate (Greiner #789271). The plate was sealed tightly with a rubber cover mat, inverted to mix, incubated for 10min on ice, then centrifuged (5 min, 3000 g). The six replicates per treatment were split so that there were three replicates per 96 well plate; this organisation remained for all future steps, including Oxford Nanopore sequencing.

In a new 1 mL 96-well round bottom plate (Greiner #780201), 400 µL of supernatant from the previous plate and 400 µL of homemade SPRI beads were added (1 mL Sera-Mag SpeedBeads (Cytiva, Marlborough, MA, USA) cleaned and dissolved in 50mL end volume containing 2.5 M NaCl, 20 mM PEG, 10mM Tris-HCl and 1mM EDTA). The plate was sealed tightly with a rubber cover mat, placed on its side and shaken at 50 rpm for 1 hour. The plate was then left on the 96 well magnet (Thermo Fisher #AM10027) for 5-10 minutes, and the supernatant was poured off (plate inverted at sink while held on magnet). A wash was performed three times by adding each time 500 µL of 80% ethanol, sealing the plate tightly with the rubber cover mat, vortexing, returning plate to magnet, then pouring off ethanol. After the third rinse, beads were left to dry for 10-30 min and 50 µL of sterile milliQ water added. The plate was shaken horizontally to resuspend the beads, left for 10min at room temperature, then returned to the magnet for 5-15min. While remaining on the magnet, 45 µL of the eluate containing genomic DNA was transferred to a regular 96 well plate and

stored at -20 °C. DNA concentrations were then measured by Invitrogen Quant-iT PicoGreen according to manufacturer instructions (ThermoFisher #P7589).

Nanopore MinIon protocol:

This protocol was adapted from Beekman et al. 2022 and Leale et al. 2023 (Beekman *et al.* 2022; Leale *et al.* 2023). DNA concentrations in these steps, with the exception of the final pooled sample, were measured by Invitrogen Quant-iT PicoGreen.

Tailed PCR reaction:

The first step in Nanopore sequencing was a PCR reaction of 19 cycles using Nanopore specific tailed primers. Three positive controls were included – the ZymoBIOMICS Microbial Community DNA Standard D6305, a previously sequenced Mabisi sample (community "F1.1" from Leale et al. 2024) (Leale *et al.* 2023), and a personally created sample titled "M+". The M+ control contained cleaned colony PCR products of the full 16S rRNA gene (27F, 1492R primers) of five species, combined in equal proportions by measured DNA concentration – *Limosilactobacillus fermentum* (DSM20052), *Lactobacillus helveticus* (DSM20075), *Lactobacillus delbrueckii* (DSM20072), *Lactococcus lactis* (DSM20481), *Acetobacter orientalis* (DSM15550), *Acetobacter lovaniensis* (DSM4491). The tailed primer PCR reaction was as follows:

Primers:
Nanopore tailed forward: 5' TTTCTGTTGGTGCTGATATTGC-[27F] 3'
Nanopore tailed reverse: 5' ACTTGCCTGTCGCTCTATCTTC-[1492R] 3'
27F: 5' AGA GTT TGA TCC TGG CTC AG 3'
1492R: 5' TAC GGY TAC CTT GTT ACG ACT T 3'

Tailed reaction reagents:
1 μL – DNA [0.5ng/μL]
12.5 μL – Phusion High Fidelity PCR 2X master mix (ThermoFisher)
1.25 μL – forward tailed primer [10uM]
1.25 μL – reverse tailed primer [10uM]
9 μL – MilliQ water

Tailed cycle conditions:
98 °C 10 sec
98 °C 5 sec (19X)
57 °C 5 sec (19X)
72 °C 30 sec (19X)
72 °C 1 min
12 °C infinity

The tailed PCR reaction was performed another two times, resulting in three separate tailed primer PCR products per sample. Placement of PCR tubes in machine was adjusted for each PCR to avoid edge effects. Each amplified DNA sample was all visualized on 1% agarose gel to confirm successful amplification, then 8 µL of each PCR reaction were combined. A total of 24 µL of amplified DNA per samples was used for the clean-up.

PCR Clean-up:
For each sample, the 24 µL of amplified DNA was cleaned with 24 µL of homemade SPRI beads (i.e., 1:1 ratio) (1 ml Sera-Mag SpeedBeads (Cytiva, Marlborough, MA, USA) cleaned and dissolved in 50 ml end volume containing 2.5 M NaCl 20 mM PEG, 10mM Tris-HCl and 1 mM EDTA) and eluted into 20 µL of MilliQ water. DNA concentration of cleaned amplicons was measured using Invitrogen Quant-iT PicoGreen. A new dilution of 15 µL of the cleaned, amplified 16S PCR product was made into a new diluted sample with DNA concentration 0.5 nM.

Barcoding:
The PCR for each sample was barcoded to enable pooling using the PCR Barcoding Expansion 1-96 Kit (Oxford Nanopore Technologies). Reaction volumes were adapted from the Nanopore barcoding protocol to save in reagents used. The reaction was as follows with a unique barcode per sample:

Barcoding PCR (per reaction/sample):
0.3 µL barcode (Oxford Nanopore Technologies)
7.2 µL [0.5nM] cleaned PCR
7.5 µL LongAmp Taq 2x Master Mix (New England Biolabs)

Cycle conditions:
95 °C – 3 min (x1)
95 °C – 15 sec (x16)
62 °C – 15 sec (x16)
65 °C – 1.5 min (x16)
65 °C – 2 min (x1)
4 °C – infinity

Before pooling of the PCR barcoded products, each was visualised on 1% agarose gel. All samples were combined with 2 µL, apart from samples with fainter bands that were subjectively determined to require 3 µL (1-F1, 1-S1, 5-L1), 4 µL (5-S1, 1-M5), or 5 µL (1-S2, 1-F5, negative control). The pooled sample (total volume = 208 µL) was cleaned using homemade SPRI beads in 1:1 volumetric ratio and eluted

in 200 µL of milliQ water. The DNA concentration of the cleaned, pooled sample was measured on the Qubit 2.0 fluorometer using dsDNA High Sensitivity Assay kit (Thermo Fisher).

In 47 µL of milliQ water, 1 ug of the barcoded, pooled, cleaned library was prepared. From here, the library was repaired, end-prepped and adaptor ligated according to the Oxford Nanopore Technologies ligation sequencing kit V14 (SQK-LSK114) , following the protocol: ligation-sequencing-amplicons-sqk-lsk114-ACDE_9163_v114_revJ_29Jun2022-minion". Reagents used were NEBNext FFPE DNA Repair Buffer (E7181A), NEBNext FFPE DNA End Repair Mix (E7182A), NEBNext Ultra II End Prep Reaction Buffer (E7183A), and NEBNext Ultra II End Prep Enzyme Mix (E7184A) (New England Biolabs).

The prepared library was loaded on a R10.4.01 SpotON Flow Cell (FLO-MIN114) on a MinION MK1c sequencing device (Oxford Nanopore Technologies) at 5 kHz sampling rate and 400 bases/second reading rate. Base-calling was performed shortly after using the "fast base-calling" setting.

Barcodes removed from final analyses were plate 1 and 2, bc 93 = "F1.1" natural Mabisi community DNA (Leale *et al.* 2023), bc94 = "M+" mock synthetic Mabisi community created from DNA of DSMZ strains, bc95 = ZymoBIOMICS Microbial Community DNA Standard, bc96 = negative control.

Bioinformatics:

To assign 16S rRNA sequences to taxonomic identity, we first downloaded the SILVA reference database (Quast *et al.* 2013). Since we were interested in variation above the species level, a custom database was produced using vsearch to cluster 16S rRNA sequences at a 95% similarity threshold (Rognes *et al.* 2016). Reads were aligned to this database using minimap2 (Li 2018) and the best matching cluster was assigned as the taxonomic identity for each read. Details of sequence compositions for main clusters are found in Table S3 from Leale et al. 2023 (Leale *et al.* 2023). Clusters to which less than five reads matched were assumed to be trace contamination and discarded. The overall frequency of clusters with <1% abundance is similar across diversity levels, as evidenced by the amount of "missing reads" to complete 1.00 abundance on bar plot (i.e., blank space at top of bar plot).

Due to a complication between the newer MinIon flowcell and older machine versions, the base-calling and then bioinformatic analysis for the first 96-well plate posed substantial complications. This was however solved for the second batch of samples (replicates 4, 5, 6). Since there was minimal species diversity in, nor variation

between, replicate populations 4, 5, and 6, we decided to only consider these three. The presence of only two lactic acid bacterium types in the in our starting Transfer 0 community was not expected, which we yet to have a clear explanation for. Our DNA extraction method differed from Leale et al. 2024, but a preliminary test confirmed that the abundances of types did not significantly differ between extraction techniques, reassuring our approach.

The yeast present in the Mabisi communities used in this study was isolated and identified as *Geotrichum candidum* using amplification and sequencing of the ITS (internal transcribed spacer) region (top match GenBank: MK967716.1). Yeast was visibly seen growing at the air interface of both Control and Introduction treatments.

Statistical analyses:

All analyses were performed in R version 4.3.2 (R Core Team 2023). Significance was defined for all analyses as p-value < 0.05. Results outputs and details of statistical analysis are found in Supplementary Material.

Community compositions:

A PERMANOVA was performed on the full data set of community compositions to evaluate the effect of treatment, transfer, and their interaction on community composition. The “adonis” function from “vegan” R package was used, with “bray” method. Community richness was also calculated using the “vegan” package’s “specnumber” function (Oksanen *et al.* 2022).

Acidity:

Excluding Recovery treatment populations, a two-way ANOVA of community*transfer was performed on pH values using the “lm” function from “lme4” R package, type 3 partial sum of squares (Bates *et al.* 2015). Post-hoc Tukey’s pairwise comparisons with adjusted p-values were then performed using the “emmeans” function from “vegan” R package at 95% confidence intervals (Oksanen *et al.* 2022).

RESULTS & DISCUSSION:

E. coli introduction does not detectably shift bacterial community composition:

Using full-length 16S rRNA gene Oxford Nanopore sequencing at four time points, we assessed whether repeated *E. coli* introduction shifted resident bacterial community compositions and whether recovery was observed after introduction stopped. We were interested in the ecological process of species sorting – the sorting of variation at the

level of species over time – in response to novel species introduction. We did not detect shifts in bacterial community composition between treatments of repeated *E. coli* introduction, at least at the full-length 16S rRNA sequencing level of bacterial type detection (Fig. 2). The compositions of bacterial types did not differ between treatments but did change overtime (Table S1: PERMANOVA on community compositions, Bray method used: effect of treatment ($p = 0.085$, $F = 2.78$, $DF = 2$), transfer ($p = 0.001$, $F = 16.83$, $DF = 3$), diversity x transfer interaction ($p = 0.034$, $F = 3.60$, $DF = 3$)). Most notably, our starting Mabisi community contained lower bacterial type richness than expected (Table S2: mean species richness transfer 1 of Control treatment = 2.67, Introduction treatment = 3), based upon prior experiments using the same sample (Leale *et al.* 2023) we assumed to find approximately five types (three lactic acid bacterium types and two acetic acid bacterium types).

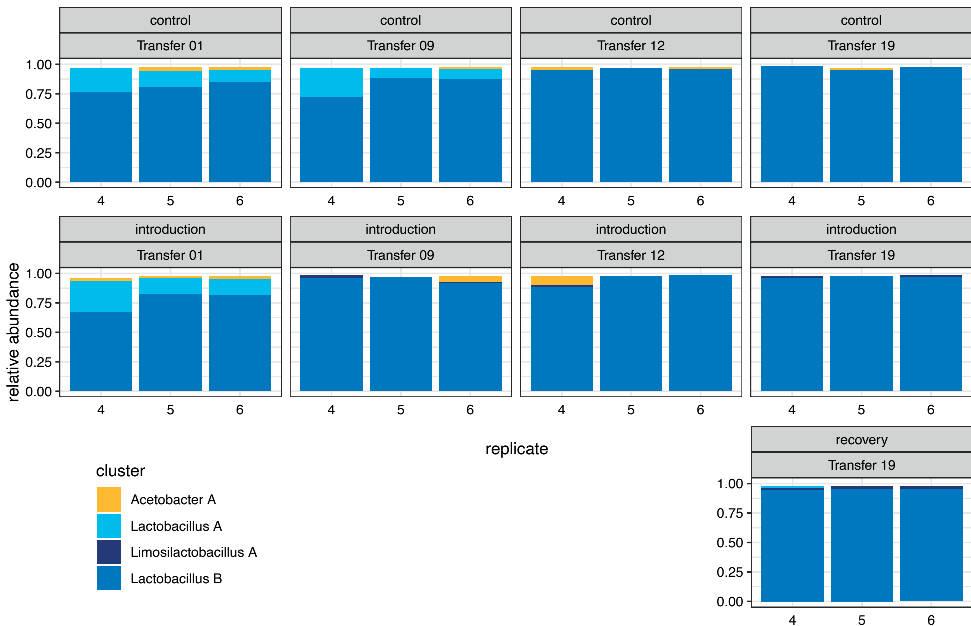


Figure 2: *bacterial community profiles show no shift in species composition from E. coli introduction.* Bacterial community compositions from full-length 16S rRNA sequencing at transfers 1, 9, 12, and 19. Only three of six replicates shown due to technical errors with sequencing in the first batch. Empty bars are clusters of <1% abundance which were not identified nor plotted. Transfer number refers to the end products of previous propagation (i.e., "transfer 9" are community compositions of the end products from transfer 8).

Although we did not detect changes in community composition at the resolution of the full-length 16S rRNA gene in the Introduction treatment (Fig. 2), it should not necessarily be inferred that underlying evolutionary processes are similarly static. So

called "cryptic dynamics" has been proposed to describe the potential role of evolution in achieving ecological stability (Kinnison, Hairston Jr and Hendry 2015; Hendry 2019). In our experiment for example, there may be shifts in genotypes that are undetected since we can only detect types in the full-length 16S rRNA gene region, thus suggesting the community remains stable at the ecological level of bacterium type composition. Metagenomic sequencing would be required to investigate potential eco-evolutionary feedback mechanisms by detecting changes at the genotype, rather than species level. Although our null expectation of no change in community composition was observed, future research should look deeper into whether there are hidden evolutionary shifts in metabolism or gene expression.

The full resistance of our Mabisi communities to *E. coli* establishment, at least to detectable levels after 72 hours of fermentation, was expected considering *E. coli's* optimal pH of 6.5 to 7.5 (Davey 1994). Our results differ from others' who observed *E. coli* successfully establish in new environments over a long period, such as soil water sheds over an entire year (Ishii *et al.* 2006). Although *E. coli* is known for impressive persistence and adaptation to new environments, where it has survived for up to seven days in soil pH of 4.57-5.14 (van Elsas *et al.* 2007; Zhang *et al.* 2013), this is still however above Mabisi's final pH of approximately 4.0 or lower (Groenenboom *et al.* 2020; Moonga *et al.* 2020; Leale *et al.* 2023). The inability of *E. coli* to establish in our experiment even after 18 rounds of repeated introduction, strengthens the argument for Mabisi's biological safety and upscaling beyond household production in Zambia (Schoustra *et al.* 2022).

Overall, we observe high resistance of Mabisi communities against *E. coli* in terms of resident bacterial community species composition and establishment. Unfortunately, the lack of species diversity in starting communities may have restricted our ability to detect potential changes in the composition of resident species. Furthermore, in the present study we focussed on potential shifts in the bacterial community and this approach overlooks potential changes in yeast abundances which are also present in the populations used here and our previous study (Leale *et al.* 2023). Regardless, we can say that the introduction of *E. coli* did not exert great enough pressure to induce bacterial species sorting, yet this was not true when considering community function responses, which we elaborate in the next section.

Community function shifts with *E. coli* introduction then rebounds once stopped:

Resistance and recovery in community function in response to repeated novel species introduction was tested using the production of volatile organic compounds (VOCs) as a proxy (Smid *et al.* 2005; De Filippis *et al.* 2016; Alekseeva *et al.* 2021). By measuring

metabolic profiles of VOCs overtime, we observe changes in community function in response to *E. coli* introduction immediately at transfer 01 and remaining until transfer 19 (Fig. 3). The communities exhibit a rebound, or recovery, in function after *E. coli* introduction was stopped at transfer 14 (comparing transfer 12 to 19, Fig. 3 C and E); at transfer 19 the PCA ellipses of metabolic profiles of the Recovery and the Control treatments overlap, while the Introduction treatment remains separated. The divide in community function between Control and Introduction treatments exists along principal component 1 (PC1) but not the second component (PC2) across transfers. However, the contribution of PC1 reduces overtime from 80% (transfer 1) to 53% (transfer 19) of the variation, while PC2 explains between 13% (transfer 1) to 31% (transfer 19); thus, strikingly, while for all transfers and treatments PC1 and PC2 explain more than 80% of the variation, the relative contribution of PC1 decreases and PC2 increases overtime. Moreover, the loadings on the metabolic compounds changed for PC1 from being all in the same direction at transfer 1 to contrasting directions at transfer 19, while the reverse was true for PC2, suggesting a change of metabolism over time.

Whether or not initially failed invaders may later establish in future introductions can be influenced by shifts in resource use of the resident community away (permissive) or towards overlapping (preventing) with the introduced species (Mallon *et al.* 2018). The same would apply for shifts in community function, which we measured as metabolic profiles of VOCs. We did not assess resource use in our experiment, and instead observed metabolic profiles to monitor changes in community function. The persistent division in metabolic profiles between Introduction *versus* Control or Recovery communities are characterised by a few VOCs which are repeatedly found at higher levels in the Introduction communities. These compounds include alcohols and esters potentially associated with Ehrlich pathway activity of the yeast *Geotrichum candidum* (ethanol, acetaldehyde, hexanoic acid-ethyl ester, 1-butanol 3-methyl) (Szudera-Kończal *et al.* 2020). It is possible that *E. coli* interacts with yeast and via changes in yeast abundances or metabolism, the community metabolic profiles are altered. Our previous work showed that the bacterial community compositions of Mabisi were not influenced by the presence-absence of *Geotrichum candidum* yeast (Leale *et al.* 2023); thus, as observed in this experiment, we would not expect to see shifts in *bacterial* community compositions even if yeast abundances were altered by *E. coli* introduction. Future work would provide insight on possible *E. coli* – yeast interactions in Mabisi.

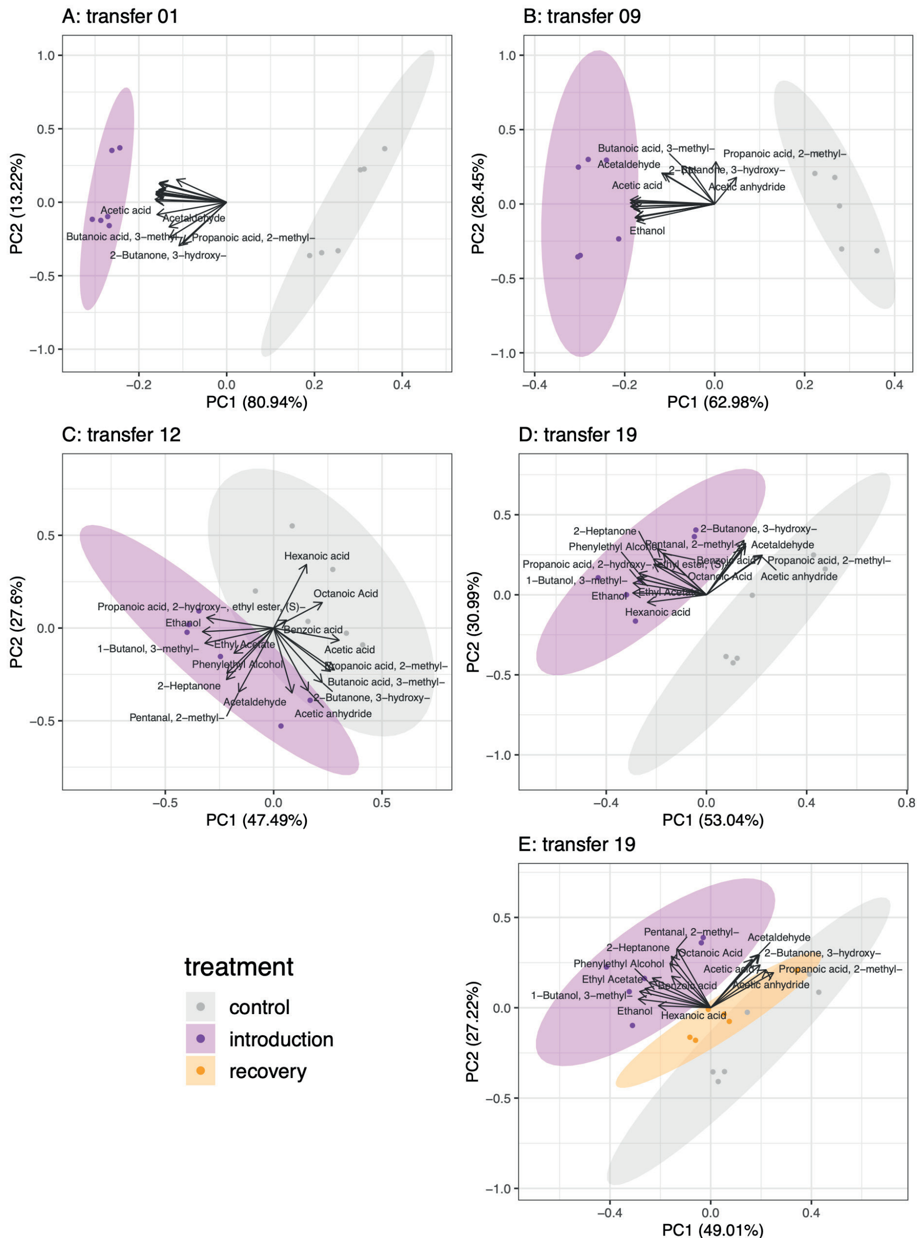


Figure 3: *Community function is immediately impacted by E. coli introduction but rebounds following cessation.* PCA analysis of metabolic profiles compiled from 16 volatile organic compounds (VOCs) using GC-MS analyses at transfer 1 (A), 9 (B), 12, (C) and 19 (D, E). Dots represent individual replicates. Ellipses are 95% confidence intervals. Direction of vectors indicate the compound's contribution to either principal component, whereas the vector length indicates the amount of variation explained by the two plotted principal components. Some labels are not shown due to text overlap, but all compounds are listed in methods section. Transfer number refers to the end products of previous propagation (i.e., "transfer 9" are metabolic profiles of the end products from transfer 8). Details of compound loadings are found in supplementary Table S3.

We are unable to disentangle whether the observed shift and recovery in metabolic profiles are due to metabolism of resident community itself, or if the shifts are simply metabolic contributions of any transient *E. coli* growth. Sampling was done after 72 hours of fermentation, when the pH was already at minimal levels for approximately 24 hours (the largest pH drop in Mabisi fermentation occurs in first 48 hours (Groenenboom *et al.* 2020)). Although *E. coli* was not detected by full-length 16S rRNA gene sequencing at 72 hours (Fig. 3), it is possible that metabolism from growth earlier during fermentation remains evident in final metabolic profiles. Few VOCs can be linked to metabolic activity of *E. coli*. For instance, 2-heptanone has been found in *E. coli* cultures (Maddula *et al.* 2009), but this compound may as well originate from a lactic acid bacterium (Albright *et al.* 2020). The generality of metabolism of the species in our community in combination with increased intensity of most, or all compounds (i.e., Fig. 4 vectors similar in length and direction) means we cannot make specific associations between detected VOCs and certain community members such as *E. coli*. In hindsight, metabolic profiles of the Recovery treatment at transfer 15 would have shown whether community function immediately rebounded to similarity of Control lines, therefore lending evidence of *E. coli*'s metabolic contributions. Detecting the metabolic input of *E. coli* and yeast through transcriptomic data across transfers and within a cycle (i.e., 6, 12, 24 hours), could help us understand our observed metabolic plasticity.

Low pH likely explains *E. coli* exclusion:

We assessed pH over time, which is a higher order yet practical component of community function. A fast fermentation with a corresponding drop in pH is a contributing factor to the high biological safety of fermented food products (Mpofu *et al.* 2016) and consumer preference (Ott *et al.* 2000), including Mabisi (Sikombe *et al.* 2023). Low pH and high acidity are key determinants for the survival of *E. coli* in soil environments (Xing *et al.* 2021). In our experiment, the fully grown Mabisi culture had a pH of around 3.25 (Fig. 4) which is well below the optimal pH of *E. coli* of 6.5 to 7.5 (Davey 1994), lending explanation to the inability of *E. coli* to establish (Fig. 3). Interestingly, the Introduction treatment had a significant lower pH as compared to the Control treatment (Table S4: ANOVA excluding Recovery treatment, effect of treatment ($p < 0.0001$, $F = 23.51$, $DF = 1$), transfer ($p < 0.0001$, $F = 34.40$, $DF = 7$), and treatment x transfer interaction ($p = 0.06052$, $F = 2.04$, $DF = 7$)), with the exception at transfer 4 and until later time points when Control and Introduction treatments were similar (Table S5: Tukey's pairwise comparison Control versus Introduction $p > 0.05$ at transfers 4, 16, and 18). *E. coli* does grow alone in milk and drops the pH because of metabolic products (Fig. S2 and S3), so the small drop in pH in the Introduction treatment is likely simply the result of an overall higher number of cells (i.e., CFU Mabisi community + CFU *E. coli*), and not

any interaction of the resident community with *E. coli*. We consider no biological significance in interpreting such a minor reduction of pH (< 0.05 pH units) resulting from *E. coli* introduction.

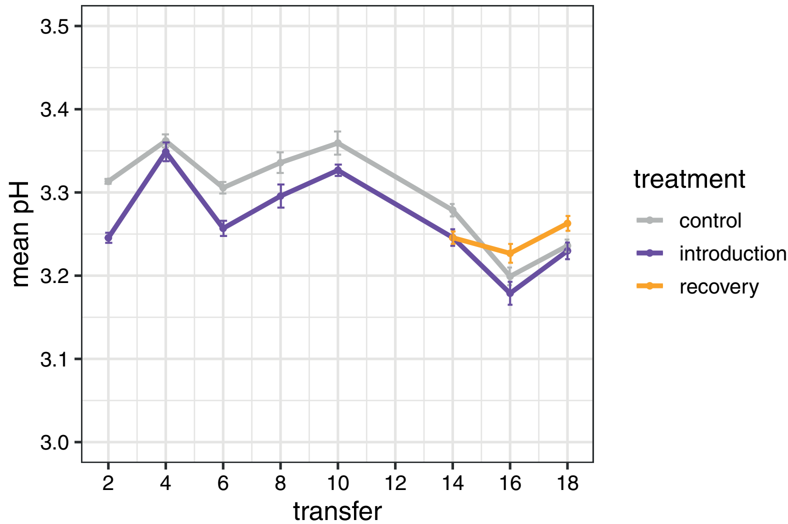


Figure 4: *E. coli introduction slightly reduces pH.* Measures of pH of Control, Introduction, and Recovery communities over 18 rounds of serial propagation. Points show mean value for the 6 replicate communities, vertical bars show standard error of the mean. Transfer number refers to the end products of previous propagation (i.e., pH "transfer 10" is the pH of end products from transfer 9).

GENERAL DISCUSSION:

The main aim of this study was to investigate repeated introduction of a novel species to a microbial community over many generations. We were interested in whether a novel species not predicted to establish may still alter the resident community composition and function over time when repeatedly introduced, and if there would be recovery once introduction of the novel species was stopped. We immediately observed alteration of community function (i.e., metabolic profiles) followed by recovery once *E. coli* introduction ceased. The observed shifts in community function in response to novel species introduction was not similarly reflected in bacterial community species composition. No change in the bacterial community composition was detected; however, the likelihood of assessing possible community composition shifts was reduced due to unexpectedly low starting species diversity in our ancestor community. Overall, we observed a shift and rebound of metabolic profiles when *E. coli* was introduced without changes in species composition.

We cannot yet realise the several potential existing mechanisms by which the shift and rebound in community function occurs. Our first explanation of the observed metabolic shift is an interaction between *E. coli* and *G. candidum* yeast found in our Mabisi samples (Leale et al. 2023). We secondly speculate that the metabolism of the resident community was changed by *E. coli*'s presence, yet the species compositions themselves remained unaltered. Thirdly, there may be changes at the genotypic level which are influencing community function, but they were not detected in our analysis level of full-length 16S rRNA gene sequencing. The low diversity we found, and its consequent limitations, emphasises the importance of the choice of resident community and novel introduced species in future research. Next steps to investigate these three possible mechanisms could involve a multitude of approaches; first estimating yeast abundances through selective agar plating, meta-transcriptomics could then further provide detailed insight into the metabolic contributions of community members, including *E. coli* and yeast, and lastly, metagenomics could show hidden diversity and dynamics of bacterial sub-species genotypes not detected here. Microorganisms can change their metabolism in response to their biotic environment and community member interactions (van Rijswijck *et al.* 2017). To investigate possible "metabolic plasticity"– a cell's ability to change their nutrient conversion (Fendt, Frezza and Erez 2020) – to *E. coli* introduction, we should also demonstrate stability in the community genotypes. There are many exciting future directions, yet what we can say for now from our experiment is that *E. coli* introduction produced community level metabolic changes without sufficient pressure to alter bacterial species sorting.

A community's resistance against novel species may be strengthened or weakened with repeated exposure overtime. Previous research on unsuccessful invasions (Mallon *et al.* 2018; Amor, Ratzke and Gore 2020; Xing *et al.* 2021) brings to question whether potential compositional and functional shifts in the resident community are beneficial to the resident community, or to the introduced species? We pose three hypotheses; either the shift makes the novel species more successful at establishing later, the reverse - that the resident community is now more stable against future introductions, or thirdly the shift could have no effect. The resident community may adapt ecologically and genetically in response to selection by continued introduction of a novel species over many generations. It is suggested that evolutionary processes can contribute to stability, such as to invasion, at the community level (Jousset *et al.* 2011; Kinnison, Hairston Jr and Hendry 2015; Hendry 2019; Vila *et al.* 2019). Specialization, generalization, and diversification are evolutionary trajectories that can create complementarity in resource use or overall niche space and enhance exclusion of a novel species (Gravel *et al.* 2011; van Moorsel *et al.* 2021). Simply ecological processes of species sorting, without genotypic changes, could also improve total resource use and leave less available niche space through similar processes of complementarity. If true, we would predict

that communities initially unable to resist establishment by an introduced species will increase resistance after experimental evolution due to reducing remaining resources. Our results do not support either prediction about the fate of establishment overtime by an introduced species – *E. coli* establishment was neither lower nor greater after 19 propagation cycles (approximately 110 generations). In general, introduced species seldomly persist, or at least do not induce lasting changes to community function. It is proposed that biotic interactions between the introduced species and resident community are more important for invader persistence than the dosage pressure of introduction (Albright *et al.* 2020, 2022). Our results support such a system-specific fitness perspective, as the fitness of *E. coli* in the Mabisi environment appears to be too low even with its high initial inoculation levels (approximately 50% *E. coli* CFU/mL).

Controlled laboratory microbial evolution experiments tracking establishment overtime of a repeatedly introduced novel species could identify critical tipping points in resistance of the resident community. Our experiments with an *E. coli* invader in the highly acidic Mabisi environment did not result in long-term establishment of an initially unsuccessful invader. However, an exciting future avenue to pursue is using an introduced species that is ecologically alike to, but not found in, our Mabisi community, such as *a Lactococcus*. Introducing *Lactococcus* is also relevant since natural Mabisi communities contain either dominant levels of lactococci or lactobacilli, or they coexist at similar abundances (Schoustra *et al.* 2013; Moonga *et al.* 2021). It is not yet entirely clear what selection factors determine the dominance of lactobacilli or lactococci in a Mabisi community, whether frequency dependence contributes, if it is a stochastic drift process, or if temperature is a main factor (Moonga *et al.* 2021). We could begin investigating their coexistence in Mabisi by following the fate of an introduced *Lactococcus* strain in our communities where it is not initially present (Fig. 2 this paper and Fig. 2 (Leale *et al.* 2023)) under varied temperatures. Similar work traced the fate of an introduced probiotic lactic acid bacteria in a traditional fermented milk from Senegal. The probiotic strain was detected at minimal levels (<1%) after only the first fermentation cycle, then was quickly lost following the second back-slop propagation (Groenenboom *et al.* 2019). The remaining question then is if stable establishment would eventually result from repeated introduction of the probiotic strain at every fermentation cycle?

Microbial communities play critical roles to environmental, industrial, and human ecosystem functions, all of which face disturbance by species introductions (Kinnunen *et al.* 2016; Bolius, Wiedner and Weithoff 2019; Gu *et al.* 2019; Libertucci and Young 2019; Oro *et al.* 2019; Albright *et al.* 2020). An introduced species may not successfully establish yet still impact the dynamics of species interactions and have a potentially lasting influence on the overall community composition and function (Mallon *et al.*

2018; Amor, Ratzke and Gore 2020; Xing *et al.* 2021). Community composition and function may not respond in parallel to an introduced species, therefore the engineering of new microbial ecosystems or the preservation of existing ones requires concurrently investigating both components. Ours and others' observed change in function yet not composition in response to disturbance (Sjöstedt *et al.* 2018) is an important consideration for industrial applications; for example, a bioreactor or food product may not be detectably contaminated by an unwanted organism, yet the function of interest is still impacted. There is additionally a critical emphasis to expand the time scales of measuring microbial community responses to perturbations and testing the effect of multiple, rather than single incidents (Vila *et al.* 2019; Philippot, Griffiths and Langenheder 2021; Xing *et al.* 2021; Albright *et al.* 2022). Soil systems heavily represent existing research on microbial community resistance to invasion (van Elsas *et al.* 2007; Allison and Martiny 2008; Shade *et al.* 2012; Yao *et al.* 2014; Mallon *et al.* 2018; Amor, Ratzke and Gore 2020; Xing *et al.* 2021), thus our work expands upon the field by using a unique fermented food model system. Our study's effort to measure microbial community composition and function over a hundred generations, including following cessation of repeated novel species introduction, contributes to understanding long-term community stability to sustained disturbances.

AUTHOR CONTRIBUTIONS:

Experimental design and conceptual ideas were developed by AML and SS. Experiments and data collection were performed by AML and FRM. Data analysis was completed by AML. Manuscript written and edited by AML, SS, EJS, BZ.

FUNDING:

This work was supported by a Wageningen University and Research Interdisciplinary Research and Education Fund (INREF).

ACKNOWLEDGMENTS:

A big thanks to Dr. Ben Auxier for creating the bioinformatic analysis pipeline and his support with data analysis and visualisation. We additionally thank Judith Wolkers-Rooijackers for running the GC-MS machine.

DATA AVAILABILITY:

Source data files and codes available on GitHub (amleale/ecoli_mabisi_2022): https://github.com/amleale/ecoli_mabisi_2022. Raw amplicon sequence data for the communities are available under NCBI BioProject PRJNA1096163 with each barcode available separately under SAMN40628846 - SAMN40629034.

CONFLICT OF INTEREST:

The authors declare they have no conflict of interest.

REFERENCES:

Albright MBN, Louca S, Winkler DE et al. Solutions in microbiome engineering: prioritizing barriers to organism establishment. ISME J 2022;16:331–8.
Albright MBN, Sevanto S, Gallegos-Graves LV et al. Biotic Interactions Are More Important than Propagule Pressure in Microbial Community Invasions. mBio 2020;11:10.1128/mbio.02089-20.
Alekseeva AY, Groenenboom AE, Smid EJ et al. Eco-Evolutionary Dynamics in Microbial Communities from Spontaneous Fermented Foods. Int J Environ Res Public Health 2021;18:10093.
Allison SD, Martiny JBH. Resistance, resilience, and redundancy in microbial communities. Proc Natl Acad Sci 2008;105:11512–9.
Amor DR, Ratzke C, Gore J. Transient invaders can induce shifts between alternative stable states of microbial communities. Sci Adv 2020;6:eaay8676.
Beekman MM, Donner SH, Litjens JJ et al. Do aphids in Dutch sweet pepper greenhouses carry heritable elements that protect them against biocontrol parasitoids? Evol Appl 2022;15:1580–93.
Blackburn TM, Lockwood JL, Cassey P. The influence of numbers on invasion success. Mol Ecol 2015;24:1942–53.
Bolius S, Wiedner C, Weithoff G. Low invasion success of an invasive cyanobacterium in a chlorophyte dominated lake. Sci Rep 2019;9:8297.
Capozzi V, Fragasso M, Romaniello R et al. Spontaneous Food Fermentations and Potential Risks for Human Health. Fermentation 2017;3:49.
Datta A, Maharaj S, Prabhu GN et al. Monitoring the Spread of Water Hyacinth (Pontederia crassipes): Challenges and Future Developments. Front Ecol Evol 2021;9.
Davey KR. Modelling the combined effect of temperature and pH on the rate coefficient for bacterial growth. Int J Food Microbiol 1994;23:295–303.
De Filippis F, Genovese A, Ferranti P et al. Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. Sci Rep 2016;6:1–11.
Donohue I, Hillebrand H, Montoya JM et al. Navigating the complexity of ecological stability. Ecol Lett 2016;19:1172–85.
van Elsas JD, Chiurazzi M, Mallon CA et al. Microbial diversity determines the invasion of soil by a bacterial pathogen. Proc Natl Acad Sci 2012;109:1159–64.
van Elsas JD, Hill P, Chroňáková A et al. Survival of genetically marked Escherichia coli O157:H7 in soil as affected by soil microbial community shifts. ISME J 2007;1:204–14.
Fendt S-M, Frezza C, Erez A. Targeting metabolic plasticity and flexibility dynamics for cancer therapy. Cancer Discov 2020;10:1797–807.
Gravel D, Bell T, Barbera C et al. Experimental niche evolution alters the strength of the diversity–productivity relationship. Nature 2011;469:89–92.
Groenenboom AE, Parker ME, De Vries A et al. Bacterial community dynamics in lait caillé, a traditional product of spontaneous fermentation from Senegal. PloS One 2019;14:e0215658.
Groenenboom AE, Shindano J, Cheepa N et al. Microbial population dynamics during traditional production of Mabisi, a spontaneous fermented milk product from Zambia: a field trial. World J Microbiol Biotechnol 2020;36:184.
Gu Y, Meng D, Yang S et al. Invader-resident community similarity contribute to the invasion process and regulate biofertilizer effectiveness. J Clean Prod 2019;241:118278.
Hendry AP. A critique for eco-evolutionary dynamics. Funct Ecol 2019;33:84–94.
Ishii S, Ksoll WB, Hicks RE et al. Presence and Growth of Naturalized Escherichia coli in Temperate Soils from Lake Superior Watersheds. Appl Environ Microbiol 2006;72:612–21.

Jousset A, Schmid B, Scheu S et al. Genotypic richness and dissimilarity opposingly affect ecosystem functioning. Ecol Lett 2011;14:537–45.

Karatayev AY, Burlakova LE. What we know and don't know about the invasive zebra (Dreissena polymorpha) and quagga (Dreissena rostriformis bugensis) mussels. Hydrobiologia 2022, DOI: 10.1007/s10750-022-04950-5.

Kinnison MT, Hairston Jr NG, Hendry AP. Cryptic eco-evolutionary dynamics. Ann N Y Acad Sci 2015;1360:120–44.

Kinnunen M, Dechesne A, Proctor C et al. A conceptual framework for invasion in microbial communities. ISME J 2016;10:2773–9.

Kruk C, Piccini C, Devercelli M et al. A trait-based approach predicting community assembly and dominance of microbial invasive species. Oikos 2021;130:571–86.

Kurkjian HM, Akbari MJ, Momeni B. The impact of interactions on invasion and colonization resistance in microbial communities. PLOS Comput Biol 2021;17:e1008643.

Leale A, Auxier B, Smid EJ et al. Influence of metabolic guilds on a temporal scale in an experimental fermented food derived microbial community. FEMS Microbiol Ecol 2023;99:fiad112.

Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 2018;34:3094–100.

Libertucci J, Young VB. The role of the microbiota in infectious diseases. Nat Microbiol 2019;4:35–45.

Louca S, Jacques SMS, Pires APF et al. High taxonomic variability despite stable functional structure across microbial communities. Nat Ecol Evol 2016;1:1–12.

Louca S, Polz MF, Mazel F et al. Function and functional redundancy in microbial systems. Nat Ecol Evol 2018;2:936–43.

Maddula S, Blank LM, Schmid A et al. Detection of volatile metabolites of Escherichia coli by multi capillary column coupled ion mobility spectrometry. Anal Bioanal Chem 2009;394:791–800.

Mallon CA, Elsas JDV, Salles JF. Microbial Invasions: The Process, Patterns, and Mechanisms. Trends Microbiol 2015;23:719–29.

Mallon CA, Le Roux X, van Doorn GS et al. The impact of failure: unsuccessful bacterial invasions steer the soil microbial community away from the invader's niche. ISME J 2018;12:728–41.

Materia VC, Linnemann AR, Smid EJ et al. Contribution of traditional fermented foods to food systems transformation: value addition and inclusive entrepreneurship. Food Secur 2021;13:1163–77.

McCullough DG. Challenges, tactics and integrated management of emerald ash borer in North America. For Int J For Res 2020;93:197–211.

Moonga HB, Schoustra SE, Linnemann AR et al. Influence of fermentation temperature on microbial community composition and physicochemical properties of mabisi, a traditionally fermented milk. LWT 2021;136:110350.

Moonga HB, Schoustra SE, Van den Heuvel J et al. Composition and diversity of natural bacterial communities in mabisi, a traditionally fermented milk. Front Microbiol 2020;11:1816.

van Moorsel SJ, Hahl T, Petchey OL et al. Co-occurrence history increases ecosystem stability and resilience in experimental plant communities. Ecology 2021;102:e03205.

Mpofu A, Linnemann AR, Nout MJR et al. Inactivation of bacterial pathogens in yoba mutandabota, a dairy product fermented with the probiotic Lactobacillus rhamnosus yoba. Int J Food Microbiol 2016;217:42–8.

Oro L, Canonico L, Marinelli V et al. Occurrence of Brettanomyces bruxellensis on Grape Berries and in Related Winemaking Cellar. Front Microbiol 2019;10.

Ott A, Hugi A, Baumgartner M et al. Sensory investigation of yogurt flavor perception: mutual influence of volatiles and acidity. J Agric Food Chem 2000;48:441–50.

Padrón B, Traveset A, Biedenweg T et al. Impact of Alien Plant Invaders on Pollination Networks in Two Archipelagos. PLOS ONE 2009;4:e6275.

Philippot L, Griffiths BS, Langenheder S. Microbial Community Resilience across Ecosystems and Multiple Disturbances. Microbiol Mol Biol Rev 2021;85:10.1128/mmbr.00026-20.

Quast C, Pruesse E, Yilmaz P et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 2013;41:D590–6.

van Rijswijck IMH, Wolkers – Rooijackers JCM, Abee T et al. Performance of non-conventional yeasts in co-culture with brewers' yeast for steering ethanol and aroma production. Microb Biotechnol 2017;10:1591–602.

Rognes T, Flouri T, Nichols B et al. VSEARCH: a versatile open source tool for metagenomics. PeerJ 2016;4:e2584.

Savageau MA. Escherichia coli Habitats, Cell Types, and Molecular Mechanisms of Gene Control. Am Nat 1983;122:732–44.

Schoustra S, Kasase C, Toarta C et al. Microbial Community Structure of Three Traditional Zambian Fermented Products: Mabisi, Chibwantu and Munkoyo. PLOS ONE 2013;8:e63948.

Schoustra S, van der Zon C, Groenenboom A et al. Microbiological safety of traditionally processed fermented foods based on raw milk, the case of Mabisi from Zambia. LWT 2022;171:113997.

Shade A, Peter H, Allison S et al. Fundamentals of Microbial Community Resistance and Resilience. Front Microbiol 2012;3.

Sikombe TW, Moonga HB, Schoustra SE et al. Sensory characteristics and consumer acceptability of four variants of mabisi, a traditionally fermented Zambian dairy product. LWT 2023;188:115410.

Sivasubramaniam D, Franks AE. Bioengineering microbial communities: Their potential to help, hinder and disgust. Bioengineered 2016;7:137–44.

Sjöstedt J, Langenheder S, Kritzberg E et al. Repeated disturbances affect functional but not compositional resistance and resilience in an aquatic bacterioplankton community. Environ Microbiol Rep 2018;10:493–500.

Smid EJ, Molenaar D, Hugenholtz J et al. Functional ingredient production: application of global metabolic models. Curr Opin Biotechnol 2005;16:190–7.

Szudera-Kończal K, Myszka K, Kubiak P et al. The use of sour and sweet whey in producing compositions with pleasant aromas using the mold galactomyces geotrichum: identification of key odorants. J Agric Food Chem 2020;68:10799–807.

Vila JC, Jones ML, Patel M et al. Uncovering the rules of microbial community invasions. Nat Ecol Evol 2019;3:1162–71.

Xing J, Jia X, Wang H et al. The legacy of bacterial invasions on soil native communities. Environ Microbiol 2021;23:669–81.

Yao Z, Wang H, Wu L et al. Interaction between the Microbial Community and Invading Escherichia coli O157:H7 in Soils from Vegetable Fields. Appl Environ Microbiol 2014;80:70–6.

Zhang T, Wang H, Wu L et al. Survival of Escherichia coli O157:H7 in Soils from Jiangsu Province, China. PLOS ONE 2013;8:e81178.

SUPPLEMENTARY (Chapter 3):

Statistical Analyses – supplementary details:

1. COMMUNITY COMPOSITIONS

"COMMUNITY" refers to the treatment of "Control", "Introduction", and "Recovery".

*> adonis2(wide[,9:12] ~ COMMUNITY*TRANSFER, data = wide[,2:3], method = "bray")*

Permutation test for adonis under reduced model
Terms added sequentially (first to last)
Permutation: free
Number of permutations: 999

Table S1: Output of PERMANOVA tests of full 16S rRNA gene community compositions.

	Df	SumOfSqs	R2	F	Pr(>F)
COMMUNITY	2	0.015199	0.06560	2.7825	0.085
TRANSFER	3	0.137874	0.59504	16.8272	0.001
COMMUNITY:TRANSFER	3	0.029472	0.12720	3.5970	0.034
Residual	18	0.049161	0.21217		
Total	26	0.231706	1.00000		

Table S2: Mean species richness from full 16S rRNA gene community compositions

	Control	Invasion	Recovery
t01	2.67	3	n/a
t09	2.33	2	n/a
t12	1.67	1.67	n/a
t19	1.33	1.67	2.33

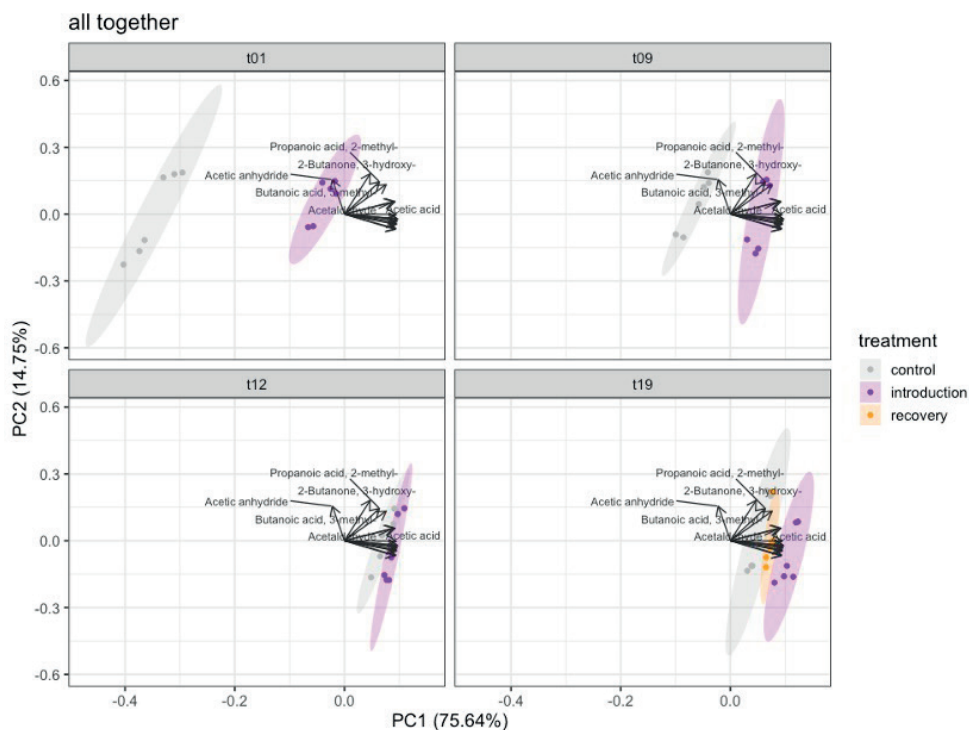


Figure S1: PCA of metabolic profiles calculated for all data points combined and plotted per transfer, rather than a separate PCA analysed for each transfer. T01 = transfer 1, t09 = transfer 9, t12 = transfer 12, t19 = transfer 19.

2. METABOLIC PROFILES

Table S3: PCA loadings of aroma compounds in Figure 3.

	transfer 1 (fig3.a)			transfer 9 (fig3.b)			transfer 12 (fig3.c)			transfer 19 (fig3d)		
	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3
Acetaldehyde	-0,22	-0,29	-0,06	-0,19	0,34	0,22	0,09	-0,38	0,20	0,17	0,36	-0,01
Ethyl Acetate	-0,26	0,15	-0,19	-0,29	-0,14	-0,02	-0,20	-0,15	0,31	-0,30	0,11	-0,24
Ethanol	-0,26	0,23	-0,15	-0,29	-0,19	-0,04	-0,36	-0,02	0,13	-0,33	0,01	-0,16
2-Heptanone	-0,27	0,10	-0,07	-0,31	-0,03	-0,09	-0,24	-0,30	0,12	-0,24	0,25	0,07
1-Butanol, 3-methyl-	-0,26	0,21	-0,11	-0,29	-0,15	0,09	-0,34	-0,09	-0,03	-0,33	0,08	-0,22
2-Butanone, 3-hydroxy-	-0,18	-0,50	-0,01	-0,09	0,41	0,41	0,24	-0,32	-0,08	0,17	0,38	-0,02
Propanoic acid, 2-hydroxy-, ethyl ester, (S)-	-0,27	0,12	0,00	-0,29	-0,15	-0,02	-0,33	0,06	0,15	-0,31	0,12	-0,02
Acetic anhydride	-0,19	0,26	0,91	0,08	0,30	-0,73	0,17	-0,37	0,14	0,25	0,27	-0,06
Pentanal, 2-methyl-	-0,28	0,03	-0,02	-0,31	0,04	-0,09	-0,17	-0,37	0,19	-0,22	0,32	-0,01
Acetic acid	-0,27	-0,14	-0,08	-0,18	0,33	-0,23	0,33	-0,07	0,14	0,16	0,32	0,41
Propanoic acid, 2-methyl-	-0,17	-0,48	0,26	0,00	0,47	-0,06	0,30	-0,24	-0,09	0,25	0,28	-0,05
Butanoic acid, 3-methyl-	-0,22	-0,41	0,08	-0,12	0,43	0,21	0,28	-0,25	-0,13	0,18	0,34	-0,23
Hexanoic acid	-0,27	0,11	-0,05	-0,30	-0,03	-0,21	0,16	0,37	0,29	-0,26	-0,05	0,57
Phenylethyl Alcohol	-0,27	0,12	-0,07	-0,30	-0,06	0,17	-0,23	-0,26	0,01	-0,30	0,15	-0,26
Octanoic Acid	-0,27	0,09	-0,06	-0,30	0,03	-0,21	0,24	0,15	0,47	-0,24	0,21	0,49
Benzoic acid	-0,27	0,02	-0,02	-0,31	0,01	0,02	0,06	0,05	0,63	-0,19	0,29	-0,04

3. ACIDITY

Recovery treatment excluded.

```
lm_ph <- lm(ph ~ treatment*transfer, data = invade)
> Anova(lm_ph, type = 3)
Anova Table (Type III tests)
```

Table S4: Output of ANOVA analysis for measured pH values.

	SumSq	Df	F value	Pr(>F)
(Intercept)	65.869	1	1.1219e+05	< 2.2e-16
treatment	0.014	1	2.3512e+01	6.004e-06
transfer	0.141	7	3.4402e+01	< 2.2e-16
treatment:transfer	0.008	7	2.0358e+00	0.06052
Residuals	0.047	80		

emmeans(lm_ph, pairwise ~ treatment|transfer)

Table S5: Output of Tukey's pairwise comparisons between communities for measured pH values.

transfer	contrast	estimate	SE	df	t.ratio	p.value
t02	con-inv	0.06783	0.014	80	4.849	<.0001
t04	con-inv	0.01317	0.014	243	0.941	0.3494
t06	con-inv	0.04883	0.014	80	3.491	0.0008
t08	con-inv	0.04017	0.014	80	2.871	0.0052
t10	con-inv	0.03267	0.014	80	2.335	0.0220
t14	con-inv	0.03283	0.014	80	2.347	0.0214
t16	con-inv	0.02033	0.014	80	1.453	0.1500
t18	con-inv	0.00617	0.014	80	0.441	0.6605

4. *E.COLI* ALONE IN MILK

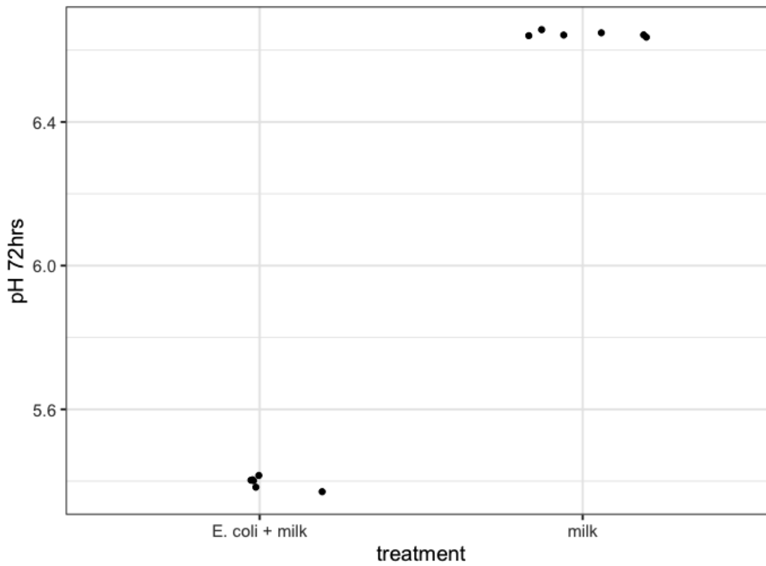


Figure S2: pH of milk inoculated or not with *E. coli* after 72 hours of fermentation at 28°C. Replication = 6, p-value < 2.2e-16.

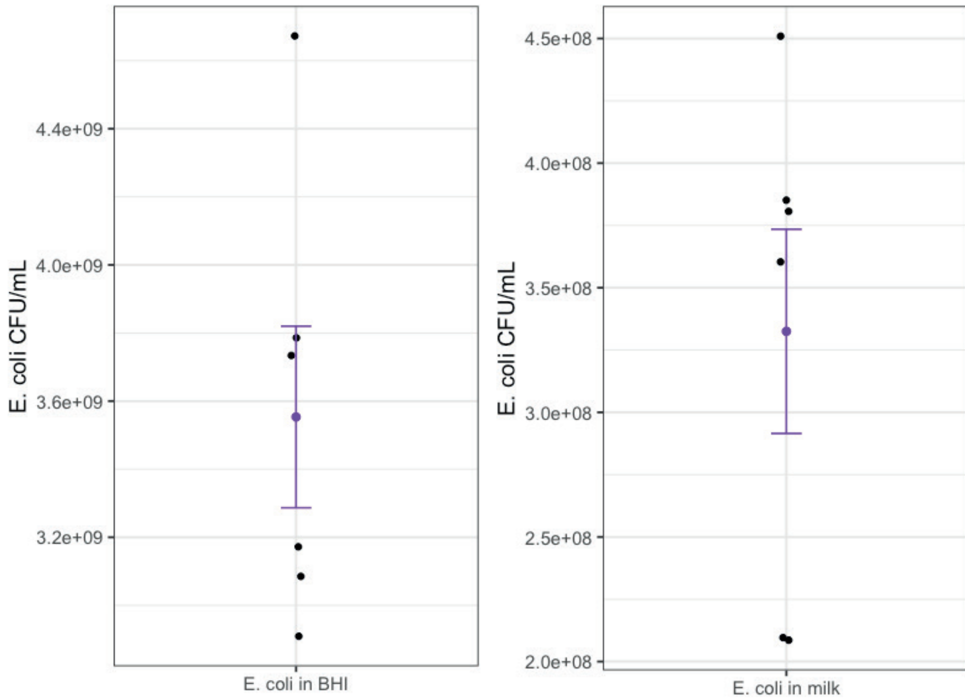


Figure S3: CFU/mL of *E. coli* growing alone for 72 hours in Brain Heart Infusion broth (37°C) and milk (28°C). Estimates calculated from colony counts on violet-red-bile (VRB) agar (replication = 6 biological replicates).

5. DNA controls for Oxford Nanopore sequencing

F1.1 control: DNA isolated and previously sequenced in Leale et al. 2023

M+ control: mock synthetic community composed of DNA isolated from DSMZ strains and combined in attempted equal by DNA concentrations. *Lactobacillus delbreuckii (*DSM20072*), Lactobacillus helveticus (*DSM20075*), Acetobacter orientalis (*DSM15550*), Acetobacter lovaniensis (*DSM4491*), Limosilactobacillus fermentum (*DSM20052*), Lactococcus lactis (*DSM20481*).*

Zymo control: ZymoBIOMICS Microbial Community DNA Standard. *Listeria monocytogenes* - 12%, *Pseudomonas aeruginosa* - 12%, *Bacillus subtilis* - 12%, *Escherichia coli* - 12%, *Salmonella enterica* - 12%, *Lactobacillus fermentum* - 12%, *Enterococcus faecalis* - 12%, *Staphylococcus aureus* - 12%, *Saccharomyces cerevisiae* - 2%, and *Cryptococcus neoformans* - 2%.

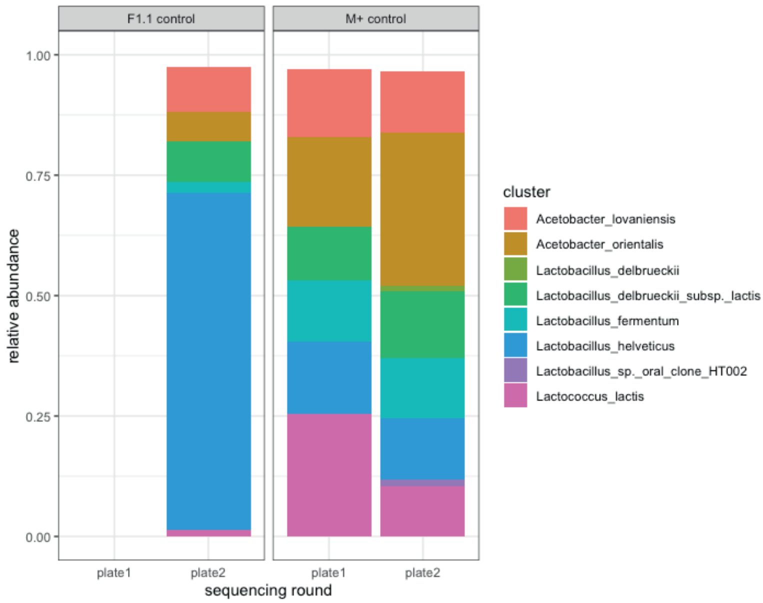


Figure S4: *our sequencing and bioinformatic pipeline suitably distinguished bacterial types and their proportions predicted in Mabisi.* Community compositions of Mabisi related positive control samples.

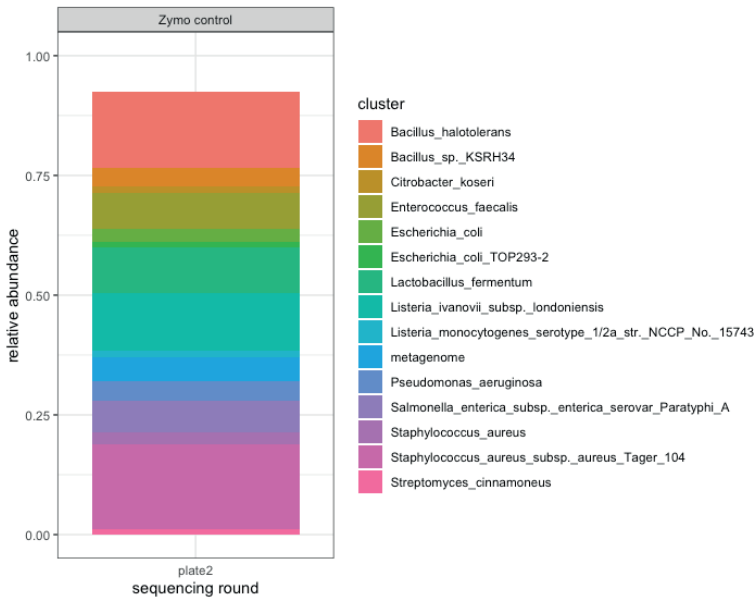


Figure S5: *our sequencing and bioinformatic pipeline suitably distinguished bacterial types and their proportions predicted in Zymo postive control.* Community compositions of ZymoBIOMICS Microbial Community DNA Standard

4

Chapter 4

S. cerevisiae serves as keystone species for spoilage resistance in experimental synthetic wine yeast communities

Alanna M. Leale[1], Eléonore Pourcelot[2], Stéphane Guezenec[2], Delphine Sicard[2], Thibault Nidelet[2]

[1] Laboratory of Genetics, Wageningen University and Research, 6700 HB Wageningen, The Netherlands
[2] SPO, Université de Montpellier, INRAE, Institut Agro, 34060 Montpellier, France

Pre-print online: https://doi.org/10.1101/2024.07.04.602080

ABSTRACT:

Species diversity is a commonly stated contributor to the fate of an invader, and thus community resistance, in both microbial and non-microbial communities. Termed the "diversity-invasion hypothesis", a positive relationship between diversity and resistance to invasion is observed when an introduced species exhibits lower levels of survival in resident communities with higher species richness. The diversity-invasion hypothesis is an attractive perspective with convincing theory and examples, yet an "invasion paradox" of contrasting results means that a positive role of diversity against invasion is still not a certainty and under debate. In this study we investigated the relationship between resistance to invasion and resident community species richness versus species identity (i.e., keystone species). Using synthetic communities comprised of combinations of four wine yeasts (*Saccharomyces cerevisiae, Lachancea thermotolerans, Torulaspora delbrueckii, Starmerella bacillaris*), we tracked over 21 days the presence of introduced *Brettanomyces bruxellensis* spoilage yeast and *Lactiplantibacillus plantarum* lactic acid bacteria to ask the following: 1. Does yeast community species richness impact the establishment of *B. bruxellensis* yeast and *L. plantarum* bacteria during wine fermentation? 2. How does yeast species identity influence such establishment? We found that species identity rather than richness drove the prevention of establishment of *B. bruxellensis* and *L. plantarum*, with *S. cerevisiae* playing a critical keystone species role. Aside from spoilage prevention by *S. cerevisiae*, the four resident yeast species demonstrated a strict dominance ranking of competitive exclusion regardless of background community composition. Our research lends evidence against the commonly predicted positive relationship between species richness and resistance to invasion. Furthermore, as spontaneously fermented natural wines and diverse starter cultures gain popularity, our findings support a remaining importance of *S. cerevisiae* in preventing *B. bruxellensis* spoilage.

INTRODUCTION:

Ecological communities are collections of species or individuals that interact in a particular environment to contribute both individual and collective community functions. Communities may exist stably without changing species composition or function, or they may experience shifts in response to disturbances. The introduction and potential establishment of a novel species, or invader, is a common type of disturbance. Just as in macroecological communities, microbial communities also experience plentiful potential invasions, which involves the processes of introduction, establishment, spread, and impact by the introduced species (Mallon, Elsas and Salles 2015; Kinnunen *et al.* 2016; Vila *et al.* 2019). Moreover, species diversity is a commonly stated contributor to the fate of an invader, and thus resistance of both microbial (Jousset *et al.* 2011; van Elsas *et al.* 2012; Mallon, Elsas and Salles 2015; Kinnunen *et al.* 2016) and non-microbial communities (Kennedy *et al.* 2002; Tilman 2004; Petruzzella *et al.* 2020; Ernst *et al.* 2022). Termed the "diversity-invasion effect", a positive relationship between diversity and resistance to invasion is observed when an introduced species exhibits lower levels of survival in resident communities with higher species richness (Kennedy *et al.* 2002; Mallon, Elsas and Salles 2015). Similarly, "biotic resistance theory" predicts that more phylogenetically diverse communities will be more resistant to invasion due to denser filled niche space, limiting colonisation by introduced species (Elton 1958; D'Antonio and Thomsen 2004; Petruzzella *et al.* 2020).

Resistance is considered the degree of change in a variable after versus before a perturbation (Allison and Martiny 2008; Donohue *et al.* 2016; Philippot, Griffiths and Langenheder 2021), thus in the context of invasion success, resistance is measured as the abundance of a novel species. As a community increases in species diversity, a wider range of resources are consumed, thereby limiting opportunities for a novel species to establish, and hence builds the basis of the diversity invasion effect and biotic resistance theory (Mallon, Elsas and Salles 2015; Petruzzella *et al.* 2020). Alternatively, overall diversity may play a negligible role due to influences of keystone species (Emery and Gross 2007; Ernst *et al.* 2023), the presence of a phylogenetically closely related species to the invader (Petruzzella *et al.* 2020), complex food webs (Thébault and Loreau 2005), resource heterogeneity (Jiang and Morin 2004; Fridley *et al.* 2007), and interaction strength (Mallon *et al.* 2015; Ratzke, Barrere and Gore 2020). The positive role of diversity against invasion is still not a certainty and under debate, probably because a combination of opposing factors. Although much insight has been gained using plant communities (Kennedy *et al.* 2002; van Ruijven, De Deyn and Berendse 2003; Petruzzella *et al.* 2020; Ernst *et al.* 2022), as well as diverse environmental microbial communities (van Elsas *et al.* 2012; Eisenhauer *et al.* 2013;

Mallon *et al.* 2015; Xing *et al.* 2021), disentangling the multiple factors contributing to invasion resistance remains difficult in such complex systems. Therefore, defined synthetic microbial communities could help test single factors of invasion resistance factors at a time.

The composition of a resident community plays a critical role in its resistance to invasion. Community composition can entail several components, including the number or richness of species (i.e., How many?), the evenness of species (i.e., Dominant or rare types?), the phylogenetic diversity of species (i.e., How different?), the population density (i.e., How abundant?), and the identity of species (i.e., Who's there?). We were specifically interested in exploring how a community's resistance to invasion is influenced by identity effects of keystone species (Emery and Gross 2007; Ernst *et al.* 2023) compared to species richness. Similar to other recent research using synthetic microbial communities (Weiss *et al.* 2023), we *do not* use a restricted definition of a keystone species since we do not require a keystone species to have a low relative abundance compared to its influence on community function (Mills, Soulé and Doak 1993; Mouquet *et al.* 2013). We instead focus on microbial communities in which we consider a keystone species more generally as a species that, when absent, significantly alters community composition and/or function (Berry and Widder 2014; Weiss *et al.* 2023). In the context of microbial community invasion resistance, simply the presence versus absence of a keystone species can therefore be the influential factor determining whether an invader successfully establishes.

To explore the diversity-invasion resistance versus identity-invasion effect, we used a model synthetic community composed of wine yeasts. Wine yeast communities serve as a useful model system to test ecological questions due to their relatively manageable diversity, wealth of knowledge on their biochemistry or system function, and established synthetic media for laboratory experiments (Bagheri *et al.* 2020; Conacher *et al.* 2021; Pourcelot *et al.* 2023; Ruiz *et al.* 2023). It is already well known that starting extracted grape juice, termed grape must, initially harbours high diversity of yeast and bacteria, but this diversity drops as the fermentation progresses. Overtime, there is a rise and final dominance of *S. cerevisiae* as it withstands fermentation conditions of high alcohol concentration and low oxygen (Albergaria and Arneborg 2016; Conacher *et al.* 2021). Despite being at negligible or even undetectable levels in grape must (Fleet 2003; Zott *et al.* 2008; Drumonde-Neves *et al.* 2021), *S. cerevisiae* is considered necessary for complete fermentation "to dryness" (<2 g/L sugar) and is commonly added as a starter culture by wine producers (Albergaria and Arneborg 2016; Ciani *et al.* 2016; Binati *et al.* 2020). Furthermore, *S. cerevisiae* is known to prevent spoilage by achieving a quick, full fermentation which consumes available resources and leaves little opportunity for other species to proliferate (Ivey, Massel

and Phister 2013; Williams, Liu and Fay 2015; Albergaria and Arneborg 2016). Additionally associated to *S. cerevisiae*'s strong fermentative capacities is its production of ethanol – an antimicrobial compound inhibiting growth of many other microbes (Du Toit and Pretorius 2000; Pereira, Freitas and Paschoalin 2021). The combination of strong resource competition and creation of a harsh environment with high ethanol concentrations by *S. cerevisiae*, together prevent spoilage. In ecological terms, the common problem of spoilage during wine production and during storage is an example of community instability against invasion. Thus, *S. cerevisiae* demonstrates a keystone species effect and helps make wine yeast communities a good model system to test diversity versus identity effects to invasion resistance.

Oenological fermentations are vulnerable to invasion by so-called spoilage species, such as the common problematic yeast *Brettanomyces bruxellensis*. To potentially combat spoilage and improve other product characteristics, more importance is being placed on promoting endemic yeast diversity and "non-*Saccharomyces*" yeasts in wine fermentations (Galati *et al.* 2019; Roudil *et al.* 2020). "Non-*Saccharomyces*" is a common term in wine science literature, but we avoid its use because the term is not ecologically based and actually typically refers specifically to the *S. cerevisiae* species, not the entire *Saccharomyces* genus (Jolly, Varela and Pretorius 2014). For the fundamental community ecology focus of our research, we here instead prefer referring to yeasts more specifically as non-*S. cerevisiae,* or communities being "*S. cerevisiae*-free". Although naturally occurring non-*S. cerevisiae* yeasts can contribute favourably to aroma and sensory measures (Binati *et al.* 2020; Roudil *et al.* 2020), they can also pose problems with slow or stuck fermentations (Ciani, Beco and Comitini 2006; Medina *et al.* 2012; Taillandier *et al.* 2014). The dominance of *S. cerevisiae* has been shown to be influenced by ecological interactions with non-*S. cerevisiae* species (Boynton and Greig 2016; Bagheri *et al.* 2020; Conacher *et al.* 2021; Ruiz *et al.* 2023). If *S. cerevisiae* growth contributes to quick resource consumption and ethanol production, yet its growth also depends on its community, that brings to question - does species richness and consequent interactions also influence invasion by spoilage microorganisms?

In this work, we investigated the relationship between resistance to invasion and resident community species richness versus species identity (i.e., keystone species) using as model system synthetic wine yeast communities composed of four yeast species naturally found in grape must: *Saccharomyces cerevisiae*, *Lachancea thermotolerans*, *Starmerella bacillaris,* and *Torulaspora delbrueckii* (Pourcelot *et al.* 2023, 2024). The yeast *Brettanomyces bruxellensis* and the lactic acid bacteria *Lactiplantibacillus plantarum* were introduced as "invaders". *S. cerevisiae* was a clear choice as an exemplary keystone species due to its strong fermentative performance and necessity

in full fermentations. *L. thermotolerans* is a moderate fermenter and a popular non-*S. cerevisiae* yeast (Urbina, Calderón and Benito 2021). *S. bacillaris* is abundant in grape must (Csoma and Sipiczki 2008; Urso *et al.* 2008) and a weak-moderate fermenter, yet it was an interesting member to include due to its unique sugar preference for fructose rather than glucose metabolism (Englezos *et al.* 2018). Lastly, *T. delbrueckii* is a moderate fermenter that is phylogenetically and phenotypically (nitrogen source use) more similar to *S. cerevisiae* (Kemsawasd *et al.* 2015; Ramírez and Velázquez 2018). The high alcohol tolerance and ability to grow on limited resources of *B. bruxellensis* makes it a common spoilage organism in wine fermentation, including after bottling, where it produces undesirable "barnyard" or "sweaty-leather" aromas (Wedral, Shewfelt and Frank 2010; Malfeito-Ferreira and Silva 2019; Harrouard *et al.* 2023). *L. plantarum* was chosen due to its growing appeal for malolactic fermentation - a critical and desirable process in red wine fermentation (du Toit *et al.* 2011; Bravo-Ferrada *et al.* 2013; Urbina, Calderón and Benito 2021). Malolactic fermentation by lactic acid bacteria converts tart malic acid to the softer lactic acid and reduces the wine's acidity (Lonvaud-Funel 1999; Virdis *et al.* 2021). However, a fine balance must be achieved because overgrowth of *L. plantarum* and other lactic acid bacteria also leads to undesirable levels of acetic acid (du Toit *et al.* 2011; Bartowsky, Costello and Chambers 2015; Urbina, Calderón and Benito 2021); interestingly then, it can be considered either a desired addition, or a spoilage invader. Furthermore, the *L. plantarum* bacteria was predicted to have less resource overlap with the resident yeast community, which may therefore contribute to its successful establishment or not. Using different communities varying for their species richness, we aimed to address two main questions: a) Does yeast community species richness impact the invasion of *B. bruxellensis* yeast and *L. plantarum* bacteria during wine fermentation? b) How does species identity (i.e., absence/presence) influence such invasion?

METHODS:

Table 1: community compositions of 17 synthetic communities. Associated colours and icons for yeast species are used consistently across all figures.

	Lactiplantibacillus plantarum	*Brettanomyces bruxellensis*	*Saccharomyces cerevisiae*	*Lachancea thermotolerans*	*Starmerella bacillaris*	*Torulaspora delbrueckii*	richness
CFU/mL day 00	**10^3**	**10^3**	**combined total 10^6**				
C01 (control)			●				n/a
C02	●	●	●				1
C03	●	●	●	●			2
C04	●	●	●		●		2
C05	●	●	●			●	2
C06	●	●	●	●	●		3
C07	●	●	●	●		●	3
C08	●	●	●		●	●	3
C09	●	●	●	●	●	●	4
C10 (control)	●	●					n/a
C11	●	●		●			1
C12	●	●			●		1
C13	●	●				●	1
C14	●	●		●	●		2
C15	●	●		●		●	2
C16	●	●			●	●	2
C17	●	●		●	●	●	3

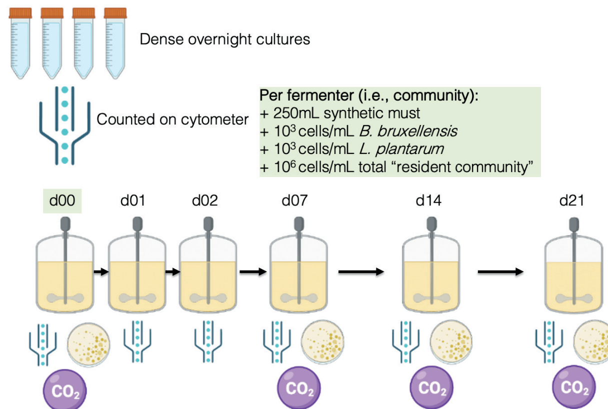


Figure 1: experimental setup. Figure created with BioRender.com.

Preparation of synthetic grape must (SGM) media:

A synthetic grape must (SGM) media was used for all fermentations (prepared as in Bely et al. 1990) and adjusted to pH 3.3. Containing: 100 g/L glucose, 100 g/L fructose, 0.18g/L malic acid, 0.18g/L citric acid, and 425 mg/L of yeast assimilable nitrogen (mixture of ammonium chloride and amino acids). One large batch was prepared and kept frozen in smaller volumes until defrosted for each round of fermentations. Four separate rounds of fermentations were completed, thus providing four replicate data points. The day before inoculations, 250 mL of SGM was pasteurised in each fermenter, aerated for 20min, and stored overnight at 4°C. On the day of inoculations (d00), 850µL of β-sitosterol (final concentration of 5 mg/L) solution was added to each fermenter.

Strains and fluorescent marking:

Four yeast species comprised the "resident community" in our study: *Saccharomyces cerevisiae*, *Lachancea thermotolerans, Starmerella bacillaris,* and *Torulaspora delbrueckii* (details in Table 2). The strains used for *S. cerevisiae, L. thermotolerans, T. delbrueckii,* and *S. bacillaris* were each uniquely fluorescently tagged with an integrated fluorescent protein gene. This was done previously as in Pourcelot et al. 2023.

Two other microbes of interest were also included: the spoilage yeast *Brettanomyces bruxellensis* and the lactic acid bacteria *Lactiplantibacillus plantarum. L. plantarum* and *B. bruxellensis* strains were not fluorescently tagged. *L. plantarum* was chosen due to easier lab culturing compared to the better-known lactic acid bacteria *Oenococcus oeni*.

Table 2: strains used and culture conditions.

species	Tag used	Strain name	Broth conditions
Lactiplantibacillus plantarum	none	Lp VPR1-30	28°C, 24 hrs, shaken, MRS
Brettanomyces bruxellensis	none	CLIB3747	28°C, 48 hrs, shaken, YEPD
Saccharomyces cerevisiae	mCherry + EGFP	CLIB3930	28°C, 24 hrs, shaken, YEPD
Lachancea thermotolerans	mCherry	CLIB3932	28°C, 24 hrs, shaken, YEPD
Starmerella bacillaris	EGFP	CLIB3933	28°C, 24 hrs, shaken, YEPD
Torulaspora delbrueckii	mCitrine	CLIB3931	28°C, 24 hrs, shaken, YEPD

Overnight cultures and day 0 (d00):

Overnight cultures of each species were prepared from a single colony (plates made less than a week prior from -80C° cell stock in YPD + 20% glycerol) in growth conditions described in Table 1. On day 00, the cell density of these overnight cultures was measured by flow cytometry. From the estimated cells/mL by cytometer, the correct volume of each species' culture was added to a 50mL tube to reach a final 10^3 cells/mL of *B. bruxellensis*, 10^3 cells/mL of *L. plantarum*, and a 10^6 cells/mL total of the remaining community members later in 250mL of SGM media (see Table 2). For example, in community C17 this was approximately 3.33 x 10^5 *L. thermotolerans*, 3.33 x 10^5 *S. bacillaris*, 3.33 x 10^5 *T. delbrueckii* for a total of 10^6 cells/mL. The mixed cultures were centrifuged (4500rpm, 5min, 4°C) then supernatant poured off, washed with 5mL of 9g/L NaCl, centrifuged again then supernatant poured off, and lastly 5mL of sterile SGM media added. The cells were resuspended in the SGM media by vortex and each community added to a separate pasteurised fermenter containing 250mL of SGM media. This created the d00 communities. An approximately 5mL sample was removed for d00 flow cytometry and plate counts. Bells were added to fermenters and their weights recorded in the Fermini computer software (internally developed), then placed at 28°C on 280 rpm spinners.

CO_2 production/fermentation kinetics:

At day d00, d01, d02, d07, d14, and d21, approximately 5mL was sampled from each fermenter for plate counts and/or flow cytometry. The weight was recorded before and after sampling in Fermini software. Estimated cumulative CO_2 production (g/L) was calculated by the software at each time point and data exported to Excel.

Plate counts:

At day d00, d07, d14, and d21 the sample from each fermenter (I.e., community) was plated onto three selective medias to enumerate total yeast (excluding *B. bruxellensis*), *L. plantarum*, and *B. bruxellensis* (Table 3). A *B. bruxellensis* selective agar media was possible because of its resistance to cycloheximide and sensitivity of the other four yeast species (Morneau, Zuehlke and Edwards 2011). A conservative concentration of 50 µg/mL cycloheximide was previously confirmed from pilot experiments to stop growth of all yeast species except *B. bruxellensis*. The antibiotic chloramphenicol was also added to yeast selective medias to ensure full suppression of *L. plantarum* growth. Plates were counted on an automatic colony counter (Interscience Scan 300) and the calculated CFU values exported to Excel. A value of zero was only entered for empty plates at 10^0 dilutions, otherwise the data point was considered missing.

Table 3: selective plating conditions. **Excludes *B. bruxellensis*

	Base Media	Additives	Incubation
total viable yeast**	YEPD	50 µg/mL chloramphenicol	28°C, aerobic, 48 hrs
B. bruxellensis	YEPD	50 µg/mL chloramphenicol 50 µg/mL cycloheximide	28°C, aerobic, 12 days
L. plantarum	MRS	50 µg/mL cycloheximide	37°C, aerobic, 48 hrs

As planned, the negative control community C10, containing only *B. bruxellensis and L. plantarum* exhibited no viable yeast colonies at day 00 (Fig. S1). Our negative control community C01 of *S. cerevisiae* alone, without *B. bruxellensis*, showed no growth throughout the experiment, lending confirmation that cycloheximide was selective for *B. bruxellensis* (Fig. S2*).* Our negative control community C01 of *S. cerevisiae* alone, without *L. plantarum*, showed no growth throughout the experiment, lending confirmation that our chloramphenicol media was selective for *L. plantarum*. (Fig. S3).

To note, we struggled finding correct dilution factors, leading to some missing data or lawns with closest CFUs estimated. Most notably, we believe a human error in Round 1 showed zero viable cells at day 21, which inaccurately guided our dilution factor choices at later rounds and resulted in lawns that were uncountable. Lawns were conservatively entered as the minimum possible colonies (i.e., 100) and the dilution factor (i.e., a lawn at 10^{-1} = 10^{4} CFU/mL). Also to note, we assume that there was early contamination by *S. cerevisiae* of the C10 community "negative" control (containing only slow growing *B. bruxellensi*s and the poor growing bacteria *L. plantarum)* in replicate round 3 and 4 (Fig. S1). The contamination is not surprising at all since *S. cerevisiae* is ubiquitous in the lab environment and is the strongest competitor in our synthetic must environment.

Flow cytometry population dynamics:

The fermenting must was sampled aseptically at day d00, d01, d02, d07, d14, d21 to determine number cells of each species with flow cytometry. Population cell number was monitored by flow cytometry with the Attune NxT™ Thermofisher® Flow Cytometer (Life Technologies, Singapore) equipped with an AttuneNXT Autosampler. Briefly, cells were washed in PBS (130 mM NaCl, 2.6 mM KCl, 7 mM Na2HPO4, 1.2 mM KH2PO4, pH 7.4; Sigma) and diluted to obtain cell concentrations about 1 to 5.10^{5} cells/mL. Numeration of live cells (i.e., viability) was assessed after staining cells with 1 µg/mL propidium iodide (PI, stored at 4°C protected from light; Calbiochem). Each population was previously tagged with one or two fluorescent proteins (Pourcelot et

al. 2023) and then detected with a specific set of channels described in Table X. Gating was done with the AttuneNxT software.

Table 4: Description of the set of channels and description of filters used for flow cytometry detection of the different yeast populations.

Population	Fluorochrome	Channel	Voltage	Fluorescence
Dead cells	PI	BL3 (488 – 695-40)	340	+
Live cells	PI	BL3 (488 – 695-40)	340	-
S. cerevisiae	mCherry + GFP	YL2 (600DLP-620/15)	320	+
		BL1 (495DLP-530/30)	260	+
L. thermotolerans	mCherry	YL2 (600DLP-620/15)	320	+
		BL1 (495DLP-530/30)	260	-
S. bacillaris	mCitrine	BL1 (495DLP-530/30)	260	+
		VL2 (495DLP-512/25)	400	-
T. delbrueckii	EGFP	BL1 (495DLP-530/30)	260	+
		VL2 (495DLP-512/25)	400	-

Our gatings proved to be appropriate, since unexpected species were not found in corresponding communities (i.e., only *L. thermotolerans* was found in C11 where only *L. thermotolerans* was inoculated). There is probable underestimation of *T. delbrueckii* cell counts by flow cytometry due to its flocculating nature. We put our best efforts to thoroughly resuspend those *T. delbrueckii* containing samples by vortex and manual pipetting.

Statistical analyses:

All analyses were performed in R version 4.3.2 (R Core Team 2023). Significance was defined for all analyses as p-value < 0.05. Results outputs and more details of statistical analysis are found in Supplementary Material.

Cumulative CO_2:

Separately for data points at day 2 and then day 21, a linear model testing the effect by community richness and each species' presence on cumulative CO_2 was fitted using the "lm" function from "lme4" R package. Cumulative CO_2 was the continuous response variable with absence versus presence of each species (α, β, δ, ε) and richness (θ) as additive categorical fixed effects (Bates *et al.* 2015). $\gamma = \alpha + \beta + \delta + \varepsilon + \theta$, where γ = cumulative CO_2, α = presence of *S. cerevisiae* (0, 1), β = presence of *L. thermotolerans* (0, 1), δ = presence of T. *delbrueckii* (0, 1), ε = presence of *S. bacillaris* (0, 1), θ = richness level (1, 2, 3, 4).

A PERMANOVA was performed to evaluate the variation in cumulative CO_2 of *S. cerevisiae*-free versus *S. cerevisiae*-containing communities. The "vegdist" function was

Chapter 4

used to calculate Euclidean distance of variation in cumulative CO_2, followed by the "adonis" function from the "vegan" R package, with "bray" method. Presence of *S. cerevisiae* was the categorical effect variable (0, 1) with Euclidean distance as the continuous response variable (Oksanen *et al.* 2022).

Total viable yeast, B. bruxellensis, L. plantarum plate counts:
Inexactness of cell counts across time points (i.e., agar plate lawns, empty plates) led us to perform all statistical tests not as quantitative cell counts but instead as binary data (absence versus presence, presence being any colony growth detected). Fisher's Exact Tests were completed using the "fisher.test" function of base R, to evaluate the effect of community richness and each species' presence on the detection of viable yeast, *B. bruxellensis,* and *L. plantarum*. The Fisher's exact test is recommended for smaller data sets below 1000 points, which is suitable for our study. A separate Fisher's Exact test was completed separately for each output, day, and variable combination. A Bonferroni correction of the p-value significance threshold was performed to account for multiple tests (i.e., adjusted p-value threshold = 0.05 / # tests).

Flow cytometry maximum total CFU/mL:
The same as for *Cumulative CO_2,* a linear model testing the effect on maximum CFU per mL of each species' presence and community richness was fitted using the "lm" function from "lme4" R package. Maximum CFU reached was the continuous response variable with absence versus presence of each species and richness as additive categorical fixed effects (Bates *et al.* 2015).

RESULTS & DISCUSSION:

S. cerevisiae-free communities ferment slower and more variably:

In this study, fermentations with combinations of four resident wine yeast species (*S. cerevisiae*, *L. thermotolerans*, *S. bacillaris, T. delbrueckii*) were performed to investigate the impact of species richness and community composition on the invasion of two potential spoilage micro-organisms (*B. bruxellensis*, *L. plantarum)*. The communities varied in resident species richness (1, 2, 3, or 4 species) and composition, where for example, half the communities included *S. cerevisiae* (*S. cerevisiae*-containing) and the other half did not (*S. cerevisiae*-free) (community compositions listed in Table 1). Firstly, the fermentation performance of the 15 communities in synthetic grape must was tracked over 21 days through weight measurements that is proportional to CO_2 production and calculated cumulative CO_2 (Fig. 2). A higher rate of fermentation was observed in communities containing *S. cerevisiae* (C02-C09, Fig. 2 dashed lines). As expected, *S. cerevisiae*-free communities in comparison showed

overall slower fermentation with a significant lower cumulative CO_2 production at day 2 (C11-C17, Fig. 2 solid lines; $F_{6, 52}$ = 82.79, p-value <2e-16, Table S4) and most did not reach comparable cumulative CO_2 levels (i.e., fermentation level) as those containing *S. cerevisiae,* even by day 21. However, we saw that some *S. cerevisiae*-free communities (C13, C15, C16, C17) reached cumulative CO_2 levels comparable to *S. cerevisiae*-containing communities at day 21 (using the cumulative CO_2 plateau of C02-C09 as a reference). The common component across these four communities was the inclusion of *T. delbrueckii* in their starting compositions. They reached in average at day 21 a CO_2 production of 95.40±6.45 g/L compared to 90.76±9.10 g/L for the communities without *T. delbrueckii* ($F_{6, 53}$ = 4.75, p-value = 0.03251, Table S5). Additionally, *S. cerevisiae*-free communities (C11-C17) had greater variation compared to when *S. cerevisiae* was present (C02-C09) (p-value = 0.002, Table S6). The minimal variation in fermentation across *S. cerevisiae*-containing communities reflects the known standardisation effect that *S. cerevisiae* starter cultures have on wine fermentation (Ciani *et al.* 2010). When analysing across all timepoints, *S. cerevisiae* was the only species whose presence had a significant effect on cumulative CO_2 ($F_{7, 352}$ = 92.98, p-value = 3.88e-07, Table S7).

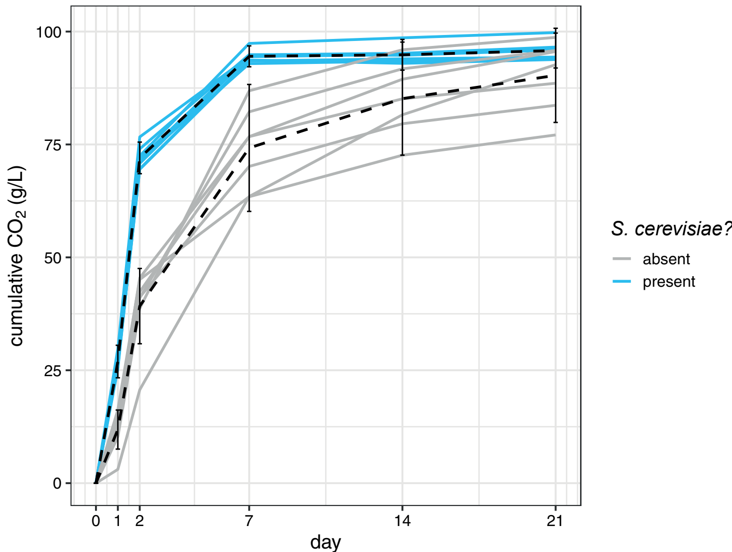


Figure 2: *fermentation is stunted in the absence of S. cerevisiae.* Cumulative CO_2 produced over time demonstrates fermentation progression in all 17 synthetic communities. Solid lines are the mean of four replicates for each community. Dotted lines show mean across all *S. cerevisiae*-free (grey) or *S. cerevisiae*-containing (blue) communities, with error bars showing one standard deviation.

A recent large-scale study found that *S. cerevisiae* growth and consequent final sugar consumption was reduced when co-cultured with other yeast species (Ruiz *et al.* 2023). They suggested that the reduction in *S. cerevisiae* growth was due to heightened negative species interactions resulting from greater species richness. While we did not directly measure consumed sugars, we do still see in our cumulative CO_2 data that all *S. cerevisiae*-containing communities reach fermentation plateau by day 7, regardless of community richness ($F_{1, 30} = 0.0053$, richness level p-values > 0.05, Table S8). The discrepancy may be because some strains used by Ruiz et al. 2023 were more strongly predicted to have negative interactions on *S. cerevisiae* (mainly *P. kudriavzevii, H. opuntiae,* and *M. pulcherrima),* whereas these species were not included in our study. Another possibility is that slight composition differences in their synthetic grape must media compared to ours (i.e., yeast assimilable nitrogen, trace elements, vitamins) influences competition dynamics since species interactions depend on the environment and available substrates. Overall, fermentation in *S. cerevisiae*-containing communities plateaued earlier and progressed more consistently than *S. cerevisiae*-free communities, which implies slower, more variable, and incomplete resource consumption across *S. cerevisiae*-free communities.

Viable yeasts persist after 7 days in *S. cerevisiae*-free communities:

Using colony counts on YEPD media agar plates, we estimated total viable yeast (excluding *B. bruxellensis*) in the 15 resident communities at day 00, day 07, day 14, and day 21 (i.e., total combined *S. cerevisiae, S. bacillaris, L. thermotolerans,* and *T. delbrueckii*) (Fig. 3). We confirmed that we successfully inoculated approximately 10^6 cells/mL total resident community yeast cells at day 00 (Fig. S1). To assess the growth of resident yeasts over time, we scored replicate plates as presence (calculated CFU/mL > 0) or absence (calculated CFU/mL = 0) of yeast and calculated the frequency of replicates where viable yeast was detected for each community and time point (Fig. 3A). Viable yeast cells were not detected by day 21 in communities containing *S. cerevisiae* (C02-C09), which aligns well with established knowledge of wine fermentation dynamics demonstrating rapid growth and consumption of resources by *S. cerevisiae* (Albergaria and Arneborg 2016). Comparatively, in *S. cerevisiae*-free communities (C11-C17), viable yeast cells were still observed at day 14 and even day 21 (non-empty plate proportion: day 14 = 0.75, day 21 = 0.50), which we interpret as a slower consumption of resources and delayed approach to carrying capacity. Furthermore, our detection of total viable yeast cells supports the rapid fermentation by *S. cerevisiae* containing communities observed in CO_2 fermentation performance data (Fig. 2). Early and fast consumption of resources by the resident yeast community is predicted to strengthen resistance against *B. bruxellensis* and *L. plantarum* because fewer remaining resources restricts opportunities for spoilage microorganisms to establish. Quick fermentation by *S. cerevisiae* will also contribute

to high levels of ethanol earlier; however, this is likely not a main factor preventing *B. bruxellensis* growth since the spoilage yeast exhibits ethanol tolerances as strong as *S. cerevisiae* (Renouf *et al.* 2006; Childs, Bohlscheid and Edwards 2015; Smith and Divol 2016). Although strain dependent, ethanol stress tolerance can also be relatively high for *L. plantarum* or other lactic acid bacteria, with growth observed at ethanol concentrations of 10 to 13% (G-Alegrıía *et al.* 2004; Ngwenya, Nkambule and Kidane 2023).

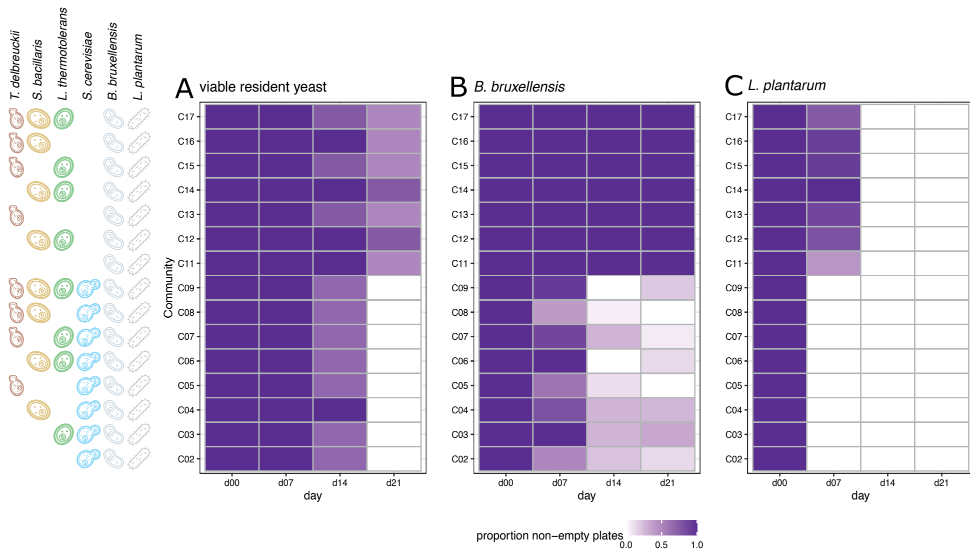


Figure 3: **A)** *viable yeast remain in S. cerevisiae-free communities.* Detected presence of total viable yeast (excluding *B. bruxellensis*) over time from YEPD agar plating and counting after 48 hours. **B)** *B. bruxellensis establishes in the absence of S. cerevisiae.* Detected presence of *B. bruxellensis* over time from selective agar plating on YEPD supplemented with cycloheximide and counted after approximately 14 days. **C)** elimination of *L. plantarum* is delayed in the absence of *S. cerevisiae.* Detected presence of *L. plantarum* over time from selective agar plating on MRS supplemented with chloramphenicol and counted after 48 hours. Species compositions of each community indicated as icons on left side.

The presence of a viable resident yeast community was influenced by a negative identity effect of *S. cerevisiae* presence at days 14 and 21 (Fisher's Exact Tests: Table S1). Species richness of the community had a negative effect of viable yeast of the resident community at day 21 (Fisher's Exact Tests: Table S1). However, when only *S. cerevisiae*-free communities and only *S. cerevisiae*-containing communities were separately analysed, the negative effect on viable yeast by species richness disappeared (Fisher's Exact Tests: Tables S2 and S3), lending evidence that richness effects are due to a confounded greater probability of *S. cerevisiae* inclusion, i.e., identity effect, rather than richness alone. Having observed that viable yeast remained in *S. cerevisiae*-free

Chapter 4

communities until day 21, we next looked at the presence of *B. bruxellensis* overtime across our 15 communities of varied species richness and composition.

B. bruxellensis proliferates in S. cerevisiae-free communities:

We then estimated the abundance of the spoilage yeast *Brettanomyces bruxellensis* overtime. The analysis of *B. bruxellensis* population density at day 00 showed that each community was successfully inoculated with approximately 10^3 cells/mL of *B. bruxellensis* (Fig. S2). As for total viable yeast, we then calculated the frequency of replicates where *B. bruxellensis* was detected as present (calculated CFU/mL > 0) (Fig. 3B). After seven days in communities containing *S. cerevisiae* (C02-C09), *B. bruxellensis* was already starting to decrease (Fig. 3B). The gradual elimination of *B. bruxellensis* in *S. cerevisiae*-containing communities continued further by day 14 and day 21 (non-empty plate proportion: day 07 = 0.40 to 1.00, day 14 = 0.00 to 0.28, day 21 = 0.00 to 0.33). In comparison, *B. bruxellensis* populations persisted in *S. cerevisiae*-free communities until day 21 (minimum non-empty plate proportion = 1.00). From day 7 onwards, *B. bruxellensis* demonstrated cell counts remaining above target inoculation values (Fig. S2). There was a significant negative effect by species richness and the presence of *S. cerevisiae* on *B. bruxellensis* detection at days 07, 14, and 21 (Fisher's Exact Tests: Table S1). Yet, the negative effect by species richness on *B. bruxellensis* disappeared when *S. cerevisiae*-free and *S. cerevisiae*-containing communities were separately analysed (Fisher's Exact Tests: Tables S2 and S3). Greater species richness is confounded with a greater probability of *S. cerevisiae* being present, thus our results support a strong keystone species role of *S. cerevisiae* that causes an overall observed effect of species richness. The persistence or even growth of *B. bruxellensis* in *S. cerevisiae*-free communities aligns with the concept of breached biotic resistance of the resident community and successful establishment of the invader (Mallon, Elsas and Salles 2015). Entering the growth and spread phase exhibits *B. bruxellensis*' ability to access local resources in communities lacking *S. cerevisiae*, which aligns with our results of fermentation progression (Fig. 2) and presence of viable yeast in the resident community (Fig. 3A).

Our results suggest a critical role for *S. cerevisiae* in the prevention of spoilage by *B. bruxellensis* in wine fermentation, however we must consider limitations from the simplicity of our synthetic grape must media (SGM). In real wine production settings, *B. bruxellensis* is still a problematic spoilage organism, even if *S. cerevisiae* is present (Wedral, Shewfelt and Frank 2010; Malfeito-Ferreira and Silva 2019; Harrouard *et al.* 2023). We cannot perfectly mimic the high resource complexity found in natural grape must from vineyards, which may provide opportunities for *B. bruxellensis* to persist since the species is known to have minimal nutrient requirements (Childs, Bohlscheid and Edwards 2015; Smith and Divol 2016). None-the-less, as enthusiasm

rises for natural wine fermentation and the beneficial roles of naturally occurring non-*S. cerevisiae* yeasts for wine characteristics (Galati *et al.* 2019; Roudil *et al.* 2020), wine producers should not underestimate the importance of *S. cerevisiae*. Creative strategies to permit the contributions of both *S. cerevisiae* and other present yeast species are being sought, such as delayed inoculation of starter cultures containing *S. cerevisiae* (Taillandier *et al.* 2014; Conacher *et al.* 2021), reduced temperatures (Alonso-del-Real *et al.* 2017), or oxygenation early in fermentation (Shekhawat, Bauer and Setati 2017) to promote growth of less competitive species (Englezos *et al.* 2022). In addition to *S. cerevisiae*, other overlooked *Saccharomyces* species with varying fermentation capabilities should be studied for their potential role in spoilage prevention. Our model wine yeast community model system could replace *S. cerevisiae* with various alternative high or low fermenter *Saccharomyces* species and see if keystone species effects are also observed generally for the *Saccharomyces* genus, or if they are specific to *S. cerevisiae*. Such an approach would enable testing theories of predicted niche overlap as a function of phylogenetic distance (Elton 1958; D'Antonio and Thomsen 2004; Petruzzella *et al.* 2020) versus phenotypic metabolic traits (i.e., sugar consumption efficiency, alcohol tolerance, nitrogen availability limitations).

L. plantarum is eliminated slower in *S. cerevisiae*-free communities:

We also explored the effect of species richness versus species identity on the persistence of lactic acid bacteria during wine fermentation by tracking *Lactiplantibacillus plantarum* in our synthetic wine yeast communities. We estimated abundances of *L. plantarum* overtime using colony counts on agar media selective for lactic acid bacteria, and then calculated the frequency of replicates where *L. plantarum* was detected as present (calculated CFU/mL > 0) (Fig. 3). We first confirmed that the initial *L. plantarum* population was successfully inoculated at approximately 10^3 cells/mL (Fig. S3). *L. plantarum* was rapidly eliminated by day 7 from communities containing *S. cerevisiae* (C02 to C09) with a significant negative effect of *S. cerevisiae* and species richness on the presence of *L. plantarum* (Fisher's Exact Tests: Table S1). In *S. cerevisiae*-free communities (C11-C17), *L. plantarum* abundances remained detected (Fig. 3C) (minimum non-empty plate proportion = 0.42) but mostly below target starting levels (Fig. S3), thus it did not proliferate in these communities. By day 14, *L. plantarum* was no longer detected in any of the fifteen communities, regardless of resident community composition. Establishment of an invader progresses either because it possesses unique metabolic traits that permit it fill an empty niche, or more commonly, because it outcompetes at least one member of the resident community (Kinnunen *et al.* 2016). We had speculated that unique bacterial metabolism of *L. plantarum*, compared to resident yeast species, could potentially enable it to fill an unutilised niche. Contrary to our expectation, it does not appear that *L. plantarum* was able to breach biotic resistance of the resident community. We interpret

its inability to enter a growth phase to demonstrate that local resources were not available to the bacteria and that the resident yeasts, regardless of species identity and richness, eventually outcompete *L. plantarum*. We did not measure gas or liquid phase metabolic profiles in this study, but it would be useful for future studies to assess if there are metabolic traces of *L. plantarum* in communities where it persisted to day 7. The aromatic and flavour contributions of *L. plantarum* is worthy of investigation since lactic acid bacteria can produce both desirable and undesirable characteristics to wine. Overgrowth of lactic acid bacteria can produce excess acetic acid (du Toit *et al.* 2011; Bartowsky, Costello and Chambers 2015; Urbina, Calderón and Benito 2021), yet malolactic fermentation is also essential to reduce the acidity of wine by converting malic acid to lactic acid (Lonvaud-Funel 1999; Virdis *et al.* 2021). An interest for diverse lactic acid bacteria starters is forming since typically just *Oenococcus oeni* is used (Bartowsky, Costello and Chambers 2015; Cappello *et al.* 2017), thus it is of value to find other suitable species and strains for starter culture communities. Finding suitable strains requires experiments that also measure organic acids (malic acid, lactic acid, acetic acid, etc.) and pH to assess the effectiveness of various lactic acid bacteria (including *O. oeni* and *L. plantarum*) for converting malic acid, versus their production of undesirable compounds. *L. plantarum* is often isolated together with *S. cerevisiae* in various fermented foods (Liu *et al.* 2022), yet our results suggest that *L. plantarum* is not a strong competitor with our synthetic wine yeast communities and must environment, especially when *S. cerevisiae* is present. Contrasting results between our synthetic experiment versus natural wine fermentation could be an outcome of the simplicity of our SGM media and our inoculation of *S. cerevisiae* at higher ratios than typically found in natural grape must (Albergaria and Arneborg 2016; Conacher *et al.* 2021; Drumonde-Neves *et al.* 2021).

Hierarchical dominance across yeast species observed by flow cytometry:
Our four resident species were each uniquely fluorescently marked, enabling us to estimate their abundances over time using flow cytometry. As expected, flow cytometry data demonstrated a quick dominance of *S. cerevisiae* and a nearly complete exclusion of the other yeast species (Fig. 4, blue lines, C02-C09). The only slight exception of exclusion by *S. cerevisiae* was in communities where *L. thermotolerans* briefly coexisted from day 01 to day 03, albeit at low abundance (C03, C06, C07, C09 - green lines). *T. delbrueckii* (red lines) and *S. bacillaris* (gold lines) were entirely outcompeted by *S. cerevisiae* within the first days of fermentation. The stronger competitive ability of *L. thermotolerans* over *T. delbrueckii* and *S. bacillaris* is further demonstrated by its dominance in *S. cerevisiae*-free communities C14, C15, and C17. *S. bacillaris* is evidently the poorest competitor, as it only reached detectable levels when in monoculture (community C12) and was outcompeted when paired with *T. delbrueckii* (community C16). Furthermore, in monocultures, *L. thermotolerans* reached the

greatest maximum population size after *S. cerevisiae* (Fig. S5). The dominance of a species was not affected by the presence of others in the community and followed the ranking order from strongest to weakest competitor of: *S. cerevisiae, L. thermotolerans, T. delbrueckii, S. bacillaris*. Importantly, this dominance ranking held regardless of the community composition; for example, *L. thermotolerans* always outcompeted *T. delbrueckii,* regardless of which other yeast species were present.

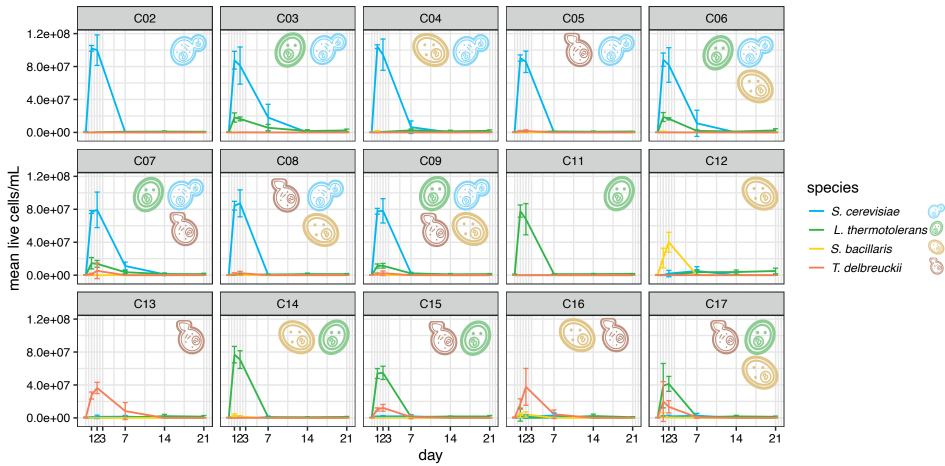


Figure 4: *Flow cytometry reveals species dominance ranking: S. cerevisiae > L. thermotolerans > T. delbrueckii > S. bacillaris.* Fluorescent flow cytometry estimates of resident community yeast species overtime. The mean live cells/mL of four replicate rounds of fermentation is plotted, with error bars showing one standard deviation. Icons indicate species composition of each community.

It was not surprising that *S. cerevisiae* quickly outcompeted the other species, given its strong fermentative ability and high inoculation concentration (equivalent to fellow resident species) compared to in natural grape must where it is a minority (Albergaria and Arneborg 2016; Conacher *et al.* 2021; Drumonde-Neves *et al.* 2021). We can compare our dominance rankings of the other three species to another study of pairwise interactions of *L. thermotolerans, S. bacillaris,* and *S. cerevisiae* (as well as other species); they similarly saw that *L. thermotolerans* could persist well through fermentation, whether inoculated or naturally present as indigenous strains (Bagheri *et al.* 2020). Our results demonstrated *S. bacillaris* as the poorest competitor even when in monoculture, whereas their research found that it could dominate if *S. cerevisiae* was absent. Our results additionally differ slightly from their finding that population dynamics between wine yeast species change in *S. cerevisiae*-free communities (Bagheri *et al.* 2020); we observed the same dominance of *L. thermotolerans* > *T. delbrueckii* > *S. bacillaris* when *S. cerevisiae* was present or not. Bagheri et al. 2020 used more complex

and diverse natural communities, whereas our study's simplicity of a four species synthetic community maybe misses complex ecological interactions that influence wine yeast species dynamics. Most importantly, our results align with a closely related study to ours that studied pairs of the same fluorescently marked strains; although they did detect interactions between the species in co-culture, there were consistent dominance rankings (Pourcelot *et al.* 2023).

Interestingly, despite *L. thermotolerans* reaching considerable maximum population sizes in communities C11 and C14 (Fig. 4, Fig. 5A), *B. bruxellensis* still proliferated (Fig. 3B). However, the maximum number of live cells in C11 and C13 is still below that of *S. cerevisiae*-containing communities (Fig. 5A). The persistence of *B. bruxellensis* indicates that resources were not exhausted, or at least could not be used quick enough, by the resident community. We therefore hypothesise that compared to *S. cerevisiae, L. thermotolerans* is less efficient in consumption of carbon and nitrogen sources, consequently leaving a greater opportunity for *B. bruxellensis* to establish. In other words, we could interpret that the resource niche overlap appears greater between *B. bruxellensis* and *S. cerevisiae,* compared to *B. bruxellensis* and *L. thermotolerans*. Resource use efficiency could be tested by growing *B. bruxellensis* in spent media of *L. thermotolerans* versus *S. cerevisiae* at multiple time points of fermentation. Successful growth of the spoilage yeast in spent media could reveal its window of opportunity for establishment and demonstrate niche overlap with alternative resident yeast species.

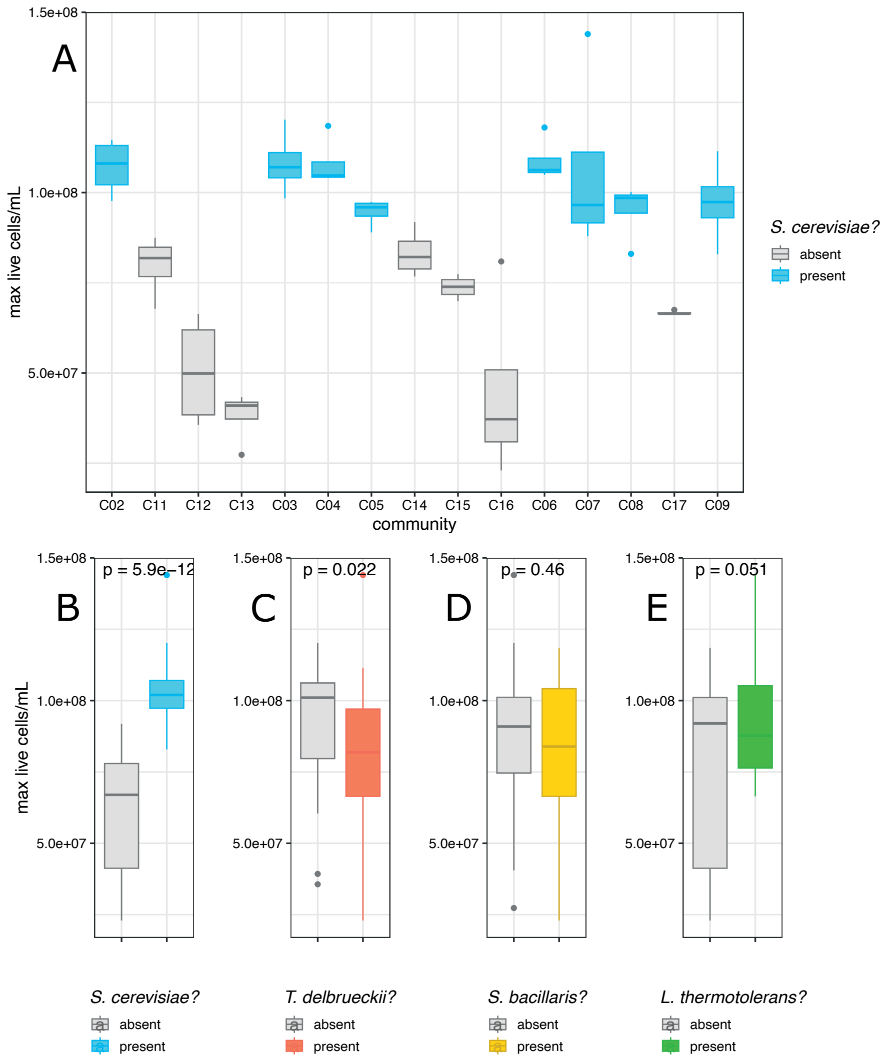


Figure 5: *Maximum CFU is increased only by the presence of S. cerevisiae*. Fluorescent flow cytometry estimates of maximum live cells/mL reached during fermentation of the resident community yeasts (i.e., excludes *B. bruxellensis*). Replication = 4 fermentation rounds per community. **A)** each community separately, ordered on the X-axis by community species richness (richness 1 = C02, C11, C12, C13; richness 2 = C03, C04, C05, C14, C15, C16; richness 3 = C06, C07, C08, C17; richness 4 = C09, **B) to E)** Data points pooled together based on absence versus presence of each resident yeast species. P-value is of t-test comparing the two groups.

To also note is the discrepancy in our data between estimated alive cells from the cytometer versus plate counts for total yeast. When observing cytometer readings on a log10 scale, it is revealed that all communities have up to 10^6 cells/mL at day 14 and day 21 (Fig. S4), yet total viable yeast plate counts for communities C02-C09 show no live cells (Fig. 3A). The proliferation of *B. bruxellensis* in *S. cerevisiae*-free communities cannot explain the differences since both cytometer and viable yeast plate count measurements exclude the spoilage species; *B. bruxellensis* was not fluorescently tagged for cytometer detection, and yeast plate counts were performed at 48 hours, long before the growth of *B. bruxellensis* colonies. The best explanation we have for the differences is because the CFU/mL calculation from cytometer data involves large multiplication factors. As an artefact, the calculation of just a few mislabelled "live" cells from cytometer data consequently produces falsely high estimated cell counts. For example, only one mislabelled live cell at a 100-fold dilution in the cytometer estimates 10^5 cells/mL (1 cell/µL * 100-fold dilution * 1000µL/mL = 10^5 cells/mL). Viewing the data on a log10 scale (Fig. S4) and inspection of the data in R reveals that this is especially a problem for *L. thermotolerans,* but also slightly for *S. cerevisiae*. Similar problems were considered and discussed by Pourcelot 2023, who concluded that cytometer counts below 10^4 cells/mL could not be trusted (Pourcelot 2023). We therefore chose to focus only on the maximum CFU reached, which are above 10^4 cells/mL, to minimise or eliminate the effect on our overall conclusions by incorrectly labelled "live" cells. The potential overestimation of live cells must be considered in other future experiments using fluorescent flow cytometry, especially when the bias differs between fluorescent markings (as seen here for *L. thermotolerans*).

Overall, we saw a determinable hierarchy of dominance across our four resident yeast species. Like in our communities, strongly hierarchical competitive exclusion has also been observed as a common outcome in experimental *bacterial* communities (Chang *et al.* 2023). If such competitive exclusion is found in bacterial communities with common complex metabolic interdependencies (Zelezniak *et al.* 2015; Zengler and Zaramela 2018), it makes sense that the same occurs in yeast communities; weak, or even no, niche breadth versus growth rate trade-offs in yeasts (Opulente *et al.* 2024) would predict that simply the strongest fermenting yeast species should dominate, which was seen in our experiment's cytometer data. Our research contributes to an underrepresented yet growing research field of synthetic yeast communities (Walker and Pretorius 2022b; Pourcelot *et al.* 2023; Ruiz *et al.* 2023). Future engineering of microbial communities - yeast, bacterial, or mixed - must consider when there is a tendency for hierarchical dominance and disproportionate impacts by keystone species. If strong identity effects override the influence of species richness, it becomes logical to consider simpler synthetic communities of fewer species.

CONCLUSIONS:

We built upon previous research of species interactions in diverse wine yeast communities (Boynton and Greig 2016; Walker and Pretorius 2022a; Pourcelot *et al.* 2023, 2023; Ruiz *et al.* 2023) with the research aim to test the impact of species richness versus identity on resistance to invasion by introduced species. Using synthetic communities comprised of all combinations of four wine yeasts (*S. cerevisiae, L. thermotolerans, T. delbrueckii, S. bacillaris*), we included and tracked the presence of *B. bruxellensis* yeast and *L. plantarum* lactic acid bacteria over 21 days. In line with our predictions and previous work on the functional impact of *S. cerevisiae* (Boynton and Greig 2016), we found that species identity rather than richness drove the prevention of establishment of *B. bruxellensis* and *L. plantarum*, with *S. cerevisiae* playing a critical keystone species role. Aside from *S. cerevisiae*, our results suggest a most promising role for *L. thermotolerans* as a non-*S. cerevisiae* yeast to include in wine fermentation starter cultures. In addition to spoilage prevention by *S. cerevisiae*, the four resident yeast species demonstrated a strict dominance ranking of competitive exclusion regardless of background community composition.

A unique aspect of our study is its use of synthetic yeast communities, whereas thus far, most microbial community ecology research focuses on bacterial communities (Fiegna *et al.* 2015; Piccardi, Vessman and Mitri 2019; Hu *et al.* 2022; Chang *et al.* 2023). If the metabolic overlap across yeasts versus across bacteria differs, this shifts our predictions about the influence of species richness and niche exclusion on community resistance to invasion. Yeast species arguably consume and produce similar metabolites but just at different efficiencies (Opulente *et al.* 2024), whereas in comparison, bacteria have more unique metabolic capacities since different strains consume and produce specific metabolites, leading to cross-feeding interdependencies (Zelezniak *et al.* 2015; D'Souza *et al.* 2018; Zengler and Zaramela 2018; Kost *et al.* 2023). The association of species richness to stability and function is therefore predictably weaker in yeast communities than bacteria. Our conclusion of a keystone species rather than species richness effect on community invasion makes sense in the context of our yeast synthetic communities. We also realise that the absent effect of richness could be a result of the small number of yeast species tested here (maximum richness = 4) and that the outcome could differ if more species were included. Furthermore, most studies do not assess the entire process of invasion to a microbial community and only consider snapshots at one or few time points (van Elsas *et al.* 2012; Mallon *et al.* 2015; Amor, Ratzke and Gore 2020). We assessed at three time points, which revealed potentially missed dynamics, such as the eventual exclusion of *L. plantarum* regardless of community species richness and identities. Our assessment of the sustained persistence or exclusion of *B. bruxellensis* and *L. plantarum* after an

extended time scale of 21 days is applicable knowledge for wine production, where spoilage can often occur even after bottling (Malfeito-Ferreira and Silva 2019).

As spontaneously fermented natural wines and diverse starter cultures gain popularity (Galati *et al.* 2019; Roudil *et al.* 2020), our research emphasises a remaining importance of *S. cerevisiae* during fermentation for resistance against *B. bruxellensis* spoilage. However, if new non-*S. cerevisiae* yeasts are also desired for their aromatic or sensory contributions, it is desirable to better promote their persistence. Here, we inoculated all species concurrently and at equal abundances, which is not realistic to real wine fermentations where *S. cerevisiae* is initially at very low abundances in natural grape must (Albergaria and Arneborg 2016; Conacher *et al.* 2021; Drumonde-Neves *et al.* 2021). It would be interesting to further test varied initial abundance ratios of *S. cerevisiae* to the other three yeasts and see the impact on both resistance against *B. bruxellensis* and desired metabolic characteristics, including those produced by lactic acid bacteria such as *L. plantarum*.

Our work contributes to a growing movement of using fermented foods as a model system in ecology and evolutionary biology (Wolfe and Dutton 2015; Alekseeva *et al.* 2021; Conacher *et al.* 2021). Community microbial ecology research is being done in diverse natural and synthetic communities with prospects for long term evolution of multispecies communities (Fiegna *et al.* 2015; Piccardi, Vessman and Mitri 2019; Hu *et al.* 2022; Chang *et al.* 2023). An exciting example in fermented foods is the evolution of wine yeast with bacteria for over 100 generations, where the emergence of mutualistic benefits was observed (du Toit *et al.* 2020). Future research could work towards promoting coexistence in our wine yeast model system, such as propagating at 24 hours before the dominance of particular species, lowering the temperature, oxygenation, or reducing SO_2 to promote less competitive species compared to *S. cerevisiae* (Jolly, Varela and Pretorius 2014; Alonso-del-Real *et al.* 2017; Bagheri *et al.* 2020). Further tests would also need to test the stability of our unique fluorescent tags over multiple generations. Experimental evolution and tracing of species sorting in synthetic yeast communities is an especially exciting prospect since they are underrepresented compared to synthetic bacterial communities.

Our research lends evidence against the commonly predicted positive relationship between species diversity and resistance to invasion. We instead observed a strong identity effect by keystone species, which could be predicted by consideration of high resource overlap in our synthetic yeast community. We recognise that the effect of species identity is not independent from species richness, since increasing richness brings greater probability that a keystone species is included. Despite the entanglement of species richness and identity effects, we feel our conclusions demonstrate the

importance to test ecological theories across varied model systems and to consider both species identity and species richness when engineering microbial communities.

AUTHOR CONTRIBUTIONS:

Experimental design and conceptual ideas were developed by AML, EP, DS, TN. Experiments and data collection were performed by AML and SG, with assistance by EP, DS, and TN. Data analysis was completed by AML and TD. Original manuscript draft written by AML, with editing by AML, EP, DS, and TN.

ACKNOWLEDGEMENTS:

We thank Diego Segond for his help with data collection and cytometry in the final replicate fermentation round.

FUNDING:

This work was supported by the Wageningen University Interdisciplinary Research and Education Fund (INREF) and an Erasmus+ grant to AML, alongside additional support by INRAE.

DATA AVAILABILITY:

Source data files and codes available on GitHub (amleale/2023_montpellier): https://github.com/amleale/2023_montpellier.

CONFLICT OF INTEREST:

The authors declare they have no conflict of interest.

REFERENCES:

Albergaria H, Arneborg N. Dominance of Saccharomyces cerevisiae in alcoholic fermentation processes: role of physiological fitness and microbial interactions. Appl Microbiol Biotechnol 2016;100:2035–46.

Alekseeva AY, Groenenboom AE, Smid EJ et al. Eco-Evolutionary Dynamics in Microbial Communities from Spontaneous Fermented Foods. Int J Environ Res Public Health 2021;18:10093.

Allison SD, Martiny JBH. Resistance, resilience, and redundancy in microbial communities. Proc Natl Acad Sci 2008;105:11512–9.

Alonso-del-Real J, Lairón-Peris M, Barrio E et al. Effect of Temperature on the Prevalence of Saccharomyces Non cerevisiae Species against a S. cerevisiae Wine Strain in Wine Fermentation: Competition, Physiological Fitness, and Influence in Final Wine Composition. Front Microbiol 2017;8.

Amor DR, Ratzke C, Gore J. Transient invaders can induce shifts between alternative stable states of microbial communities. Sci Adv 2020;6:eaay8676.

Bagheri B, Bauer FF, Cardinali G et al. Ecological interactions are a primary driver of population dynamics in wine yeast microbiota during fermentation. Sci Rep 2020;10:4911.

Bartowsky E j., Costello P j., Chambers P j. Emerging trends in the application of malolactic fermentation. Aust J Grape Wine Res 2015;21:663–9.

Bates D, Mächler M, Bolker B et al. Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw 2015;67:1–48.

Berry D, Widder S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. Front Microbiol 2014;5.

Binati RL, Lemos Junior WJF, Luzzini G et al. Contribution of non-Saccharomyces yeasts to wine volatile and sensory diversity: A study on Lachancea thermotolerans, Metschnikowia spp. and Starmerella bacillaris strains isolated in Italy. Int J Food Microbiol 2020;318:108470.

Boynton PJ, Greig D. Species richness influences wine ecosystem function through a dominant species. Fungal Ecol 2016;22:61–72.

Bravo-Ferrada BM, Hollmann A, Delfederico L et al. Patagonian red wines: selection of Lactobacillus plantarum isolates as potential starter cultures for malolactic fermentation. World J Microbiol Biotechnol 2013;29:1537–49.

Cappello MS, Zapparoli G, Logrieco A et al. Linking wine lactic acid bacteria diversity with wine aroma and flavour. Int J Food Microbiol 2017;243:16–27.

Chang C-Y, Bajić D, Vila JCC et al. Emergent coexistence in multispecies microbial communities. Science 2023;381:343–8.

Childs BC, Bohlscheid JC, Edwards CG. Impact of available nitrogen and sugar concentration in musts on alcoholic fermentation and subsequent wine spoilage by Brettanomyces bruxellensis. Food Microbiol 2015;46:604–9.

Ciani M, Beco L, Comitini F. Fermentation behaviour and metabolic interactions of multistarter wine yeast fermentations. Int J Food Microbiol 2006;108:239–45.

Ciani M, Capece A, Comitini F et al. Yeast Interactions in Inoculated Wine Fermentation. Front Microbiol 2016;7.

Ciani M, Comitini F, Mannazzu I et al. Controlled mixed culture fermentation: a new perspective on the use of non-Saccharomyces yeasts in winemaking. FEMS Yeast Res 2010;10:123–33.

Conacher C, Luyt N, Naidoo-Blassoples R et al. The ecology of wine fermentation: a model for the study of complex microbial ecosystems. Appl Microbiol Biotechnol 2021;105:3027–43.

Csoma H, Sipiczki M. Taxonomic reclassification of Candida stellata strains reveals frequent occurrence of Candida zemplinina in wine fermentation. FEMS Yeast Res 2008;8:328–36.

D'Antonio CM, Thomsen M. Ecological Resistance in Theory and Practice. Weed Technol 2004;18:1572–7.
Donohue I, Hillebrand H, Montoya JM et al. Navigating the complexity of ecological stability. Ecol Lett 2016;19:1172–85.
Drumonde-Neves J, Fernandes T, Lima T et al. Learning from 80 years of studies: a comprehensive catalogue of non-Saccharomyces yeasts associated with viticulture and winemaking. FEMS Yeast Res 2021;21:foab017.
D'Souza G, Shitut S, Preussger D et al. Ecology and evolution of metabolic cross-feeding interactions in bacteria. Nat Prod Rep 2018;35:455–88.
Du Toit M, Pretorius IS. Microbial spoilage and preservation of wine : using weapons for nature's own arsenal. 2000.
Eisenhauer N, Schulz W, Scheu S et al. Niche dimensionality links biodiversity and invasibility of microbial communities. Funct Ecol 2013;27:282–8.
van Elsas JD, Chiurazzi M, Mallon CA et al. Microbial diversity determines the invasion of soil by a bacterial pathogen. Proc Natl Acad Sci 2012;109:1159–64.
Elton CS. The Ecology of Invasions by Animals and Plants. Boston, MA: Springer US, 1958.
Emery SM, Gross KL. Dominant Species Identity, Not Community Evenness, Regulates Invasion in Experimental Grassland Plant Communities. Ecology 2007;88:954–64.
Englezos V, Cocolin L, Rantsiou K et al. Specific Phenotypic Traits of Starmerella bacillaris Related to Nitrogen Source Consumption and Central Carbon Metabolite Production during Wine Fermentation. Appl Environ Microbiol 2018;84:e00797-18.
Englezos V, Jolly NP, Di Gianvito P et al. Microbial interactions in winemaking: Ecological aspects and effect on wine quality. Trends Food Sci Technol 2022;127:99–113.
Ernst AR, Barak RS, Glasenhardt M-C et al. Dominant species establishment may influence invasion resistance more than phylogenetic or functional diversity. J Appl Ecol 2023;60:2652–64.
Ernst AR, Barak RS, Hipp AL et al. The invasion paradox dissolves when using phylogenetic and temporal perspectives. J Ecol 2022;110:443–56.
Fiegna F, Moreno-Letelier A, Bell T et al. Evolution of species interactions determines microbial community productivity in new environments. ISME J 2015;9:1235–45.
Fleet GH. Yeast interactions and wine flavour. Int J Food Microbiol 2003;86:11–22.
Fridley JD, Stachowicz JJ, Naeem S et al. The Invasion Paradox: Reconciling Pattern and Process in Species Invasions. Ecology 2007;88:3–17.
Galati A, Schifani G, Crescimanno M et al. "Natural wine" consumers and interest in label information: An analysis of willingness to pay in a new Italian wine market segment. J Clean Prod 2019;227:405–13.
G-Alegrıía E, López I, Ruiz JI et al. High tolerance of wild Lactobacillus plantarum and Oenococcus oeni strains to lyophilisation and stress environmental conditions of acid pH and ethanol. FEMS Microbiol Lett 2004;230:53–61.
Harrouard J, Eberlein C, Ballestra P et al. Brettanomyces bruxellensis: Overview of the genetic and phenotypic diversity of an anthropized yeast. Mol Ecol 2023;32:2374–95.
Hu J, Amor DR, Barbier M et al. Emergent phases of ecological diversity and dynamics mapped in microcosms. Science 2022;378:85–9.
Ivey M, Massel M, Phister TG. Microbial Interactions in Food Fermentations. Annu Rev Food Sci Technol 2013;4:141–62.
Jiang L, Morin PJ. Productivity gradients cause positive diversity–invasibility relationships in microbial communities. Ecol Lett 2004;7:1047–57.

Jolly NP, Varela C, Pretorius IS. Not your ordinary yeast: non-Saccharomyces yeasts in wine production uncovered. FEMS Yeast Res 2014;14:215–37.

Jousset A, Schmid B, Scheu S et al. Genotypic richness and dissimilarity opposingly affect ecosystem functioning. Ecol Lett 2011;14:537–45.

Kemsawasd V, Viana T, Ardö Y et al. Influence of nitrogen sources on growth and fermentation performance of different wine yeast species during alcoholic fermentation. Appl Microbiol Biotechnol 2015;99:10191–207.

Kennedy TA, Naeem S, Howe KM et al. Biodiversity as a barrier to ecological invasion. Nature 2002;417:636–8.

Kinnunen M, Dechesne A, Proctor C et al. A conceptual framework for invasion in microbial communities. ISME J 2016;10:2773–9.

Kost C, Patil KR, Friedman J et al. Metabolic exchanges are ubiquitous in natural microbial communities. Nat Microbiol 2023;8:2244–52.

Liu J, Huang T-Y, Liu G et al. Microbial Interaction between Lactiplantibacillus plantarum and Saccharomyces cerevisiae: Transcriptome Level Mechanism of Cell-Cell Antagonism. Microbiol Spectr 2022;10:e01433-22.

Lonvaud-Funel A. Lactic acid bacteria in the quality improvement and depreciation of wine. In: Konings WN, Kuipers OP, In 't Veld JHJH (eds.). Lactic Acid Bacteria: Genetics, Metabolism and Applications: Proceedings of the Sixth Symposium on Lactic Acid Bacteria: Genetics, Metabolism and Applications, 19–23 September 1999, Veldhoven, The Netherlands. Dordrecht: Springer Netherlands, 1999, 317–31.

Malfeito-Ferreira M, Silva AC. Spoilage Yeasts in Wine Production. In: Romano P, Ciani M, Fleet GH (eds.). Yeasts in the Production of Wine. New York, NY: Springer, 2019, 375–94.

Mallon CA, Elsas JDV, Salles JF. Microbial Invasions: The Process, Patterns, and Mechanisms. Trends Microbiol 2015;23:719–29.

Mallon CA, Poly F, Le Roux X et al. Resource pulses can alleviate the biodiversity–invasion relationship in soil microbial communities. Ecology 2015;96:915–26.

Medina K, Boido E, Dellacassa E et al. Growth of non-Saccharomyces yeasts affects nutrient availability for Saccharomyces cerevisiae during wine fermentation. Int J Food Microbiol 2012;157:245–50.

Mills LS, Soulé ME, Doak DF. The Keystone-Species Concept in Ecology and Conservation. BioScience 1993;43:219–24.

Morneau AD, Zuehlke JM, Edwards CG. Comparison of media formulations used to selectively cultivate Dekkera/Brettanomyces. Lett Appl Microbiol 2011;53:460–5.

Mouquet N, Gravel D, Massol F et al. Extending the concept of keystone species to communities and ecosystems. Ecol Lett 2013;16:1–8.

Ngwenya MP, Nkambule TP, Kidane SW. Physicochemical attributes and acceptability of marula wine fermented with natural Lactiplantibacillus plantarum and Saccharomyces cerevisiae. Heliyon 2023;9:e21613.

Opulente DA, LaBella AL, Harrison M-C et al. Genomic factors shape carbon and nitrogen metabolic niche breadth across Saccharomycotina yeasts. Science 2024;384:eadj4503.

Pereira PR, Freitas CS, Paschoalin VMF. Saccharomyces cerevisiae biomass as a source of next-generation food preservatives: Evaluating potential proteins as a source of antimicrobial peptides. Compr Rev Food Sci Food Saf 2021;20:4450–79.

Petruzzella A, da S. S. R. Rodrigues TA, van Leeuwen CHA et al. Species identity and diversity effects on invasion resistance of tropical freshwater plant communities. Sci Rep 2020;10:5626.

Philippot L, Griffiths BS, Langenheder S. Microbial Community Resilience across Ecosystems and Multiple Disturbances. Microbiol Mol Biol Rev 2021;85:10.1128/mmbr.00026-20.

Piccardi P, Vessman B, Mitri S. Toxicity drives facilitation between 4 bacterial species. Proc Natl Acad Sci 2019;116:15979–84.

Pourcelot E. Développement d'une communauté modèle de levures oenologiques. 2023.

Pourcelot E, Conacher C, Marlin T et al. Comparing the hierarchy of inter- and intra-species interactions with population dynamics of wine yeast cocultures. FEMS Yeast Res 2023;23:foad039.

Pourcelot E, Vigna A, Marlin T et al. Design of a new model yeast consortium for ecological studies of enological fermentation. 2024:2024.05.06.592697.

Ramírez M, Velázquez R. The Yeast Torulaspora delbrueckii: An Interesting But Difficult-To-Use Tool for Winemaking. Fermentation 2018;4:94.

Ratzke C, Barrere J, Gore J. Strength of species interactions determines biodiversity and stability in microbial communities. Nat Ecol Evol 2020;4:376–83.

Renouf V, Falcou M, Miot-Sertier C et al. Interactions between Brettanomyces bruxellensis and other yeast species during the initial stages of winemaking. J Appl Microbiol 2006;100:1208–19.

Roudil L, Russo P, Berbegal C et al. Non-Saccharomyces Commercial Starter Cultures: Scientific Trends, Recent Patents and Innovation in the Wine Sector. 2020, DOI: 10.2174/2212798410666190131103713.

van Ruijven J, De Deyn GB, Berendse F. Diversity reduces invasibility in experimental plant communities: the role of plant species. Ecol Lett 2003;6:910–8.

Ruiz J, de Celis M, Diaz-Colunga J et al. Predictability of the community-function landscape in wine yeast ecosystems. Mol Syst Biol 2023;19:e11613.

Shekhawat K, Bauer FF, Setati ME. Impact of oxygenation on the performance of three non-Saccharomyces yeasts in co-fermentation with Saccharomyces cerevisiae. Appl Microbiol Biotechnol 2017;101:2479–91.

Smith BD, Divol B. Brettanomyces bruxellensis, a survivalist prepared for the wine apocalypse and other beverages. Food Microbiol 2016;59:161–75.

Taillandier P, Lai QP, Julien-Ortiz A et al. Interactions between Torulaspora delbrueckii and Saccharomyces cerevisiae in wine fermentation: influence of inoculation and nitrogen content. World J Microbiol Biotechnol 2014;30:1959–67.

Thébault E, Loreau M. Trophic Interactions and the Relationship between Species Diversity and Ecosystem Stability. Am Nat 2005;166:E95–114.

Tilman D. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. Proc Natl Acad Sci 2004;101:10854–61.

du Toit M, Engelbrecht L, Lerm E et al. Lactobacillus: the Next Generation of Malolactic Fermentation Starter Cultures—an Overview. Food Bioprocess Technol 2011;4:876–906.

du Toit SC, Rossouw D, du Toit M et al. Enforced Mutualism Leads to Improved Cooperative Behavior between Saccharomyces cerevisiae and Lactobacillus plantarum. Microorganisms 2020;8:1109.

Urbina Á, Calderón F, Benito S. The Combined Use of Lachancea thermotolerans and Lactiplantibacillus plantarum (former Lactobacillus plantarum) in Wine Technology. Foods 2021;10:1356.

Urso R, Rantsiou K, Dolci P et al. Yeast biodiversity and dynamics during sweet wine production as determined by molecular methods. FEMS Yeast Res 2008;8:1053–62.

Vila JC, Jones ML, Patel M et al. Uncovering the rules of microbial community invasions. Nat Ecol Evol 2019;3:1162–71.

Virdis C, Sumby K, Bartowsky E et al. Lactic Acid Bacteria in Wine: Technological Advances and Evaluation of Their Functional Role. Front Microbiol 2021;11.

Walker RS, Pretorius IS. Synthetic biology for the engineering of complex wine yeast communities. Nat Food 2022a;3:249–54.

Walker RSK, Pretorius IS. Synthetic biology for the engineering of complex wine yeast communities. Nat Food 2022b;3:249–54.

Wedral D, Shewfelt R, Frank J. The challenge of Brettanomyces in wine. LWT - Food Sci Technol 2010;43:1474–9.

Weiss AS, Niedermeier LS, von Strempel A et al. Nutritional and host environments determine community ecology and keystone species in a synthetic gut bacterial community. Nat Commun 2023;14:4780.

Williams KM, Liu P, Fay JC. Evolution of ecological dominance of yeast species in high-sugar environments. Evolution 2015;69:2079–93.

Wolfe BE, Dutton RJ. Fermented foods as experimentally tractable microbial ecosystems. Cell 2015;161:49–55.

Xing J, Jia X, Wang H et al. The legacy of bacterial invasions on soil native communities. Environ Microbiol 2021;23:669–81.

Zelezniak A, Andrejev S, Ponomarova O et al. Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proc Natl Acad Sci 2015;112:6449–54.

Zengler K, Zaramela LS. The social network of microorganisms — how auxotrophies shape complex communities. Nat Rev Microbiol 2018;16:383–90.

Zott K, Miot-Sertier C, Claisse O et al. Dynamics and diversity of non-Saccharomyces yeasts during the early stages in winemaking. Int J Food Microbiol 2008;125:197–203.

SUPPLEMENTARY (Chapter 4):

Table S1: p-value of Fisher's Exact tests for all communities of detected colony growth on agar media (i.e., confounded effect of richness and *S. cerevisiae* presence). Tests completed separately for each output, day, and variable combination. Bonferroni corrected p-value significance threshold = 0.0017 (i.e., 0.05 / 30 tests). Example analysis for day 14, total viable yeast, *S. bacillaris* shown below in "Supplementary (statistical analyses), 2. Binomial plate count data".

	Day	richness	*S. cerevisiae*	*L. thermotolerans*	*T. delbrueckii*	*S. bacillaris*
Viable yeast	**d07**	n/a	n/a	n/a	n/a	n/a
	d14	0.073	0.004	0.546	0.044	0.106
	d21	0.001	< 2.2e-16	0.736	0.616	0.177
B. bruxellensis	**d07**	0.774	1.407e-09	2.642e-06	0.392	0.196
	d14	1.05e-14	< 2.2e-16	0.274	1.00	0.398
	d21	3.425e-12	< 2.2e-16	0.386	1.00	0.385
L. plantarum	**d07**	2.709e-06	< 2.2e-16	0.883	0.185	0.140
	d14	n/a	n/a	n/a	n/a	n/a
	d21	n/a	n/a	n/a	n/a	n/a

Table S2: Fisher's Exact tests for *S. cerevisiae*-free communities only (C11-C17). Tests completed separately for each output, day, and variable combination. Bonferroni corrected p-value significance threshold = 0.0042 (i.e., 0.05 / 12 tests).

	Day	richness	*L. thermotolerans*	*T. delbrueckii*	*S. bacillaris*
Viable yeast	**d07**	n/a	n/a	n/a	n/a
	d14	0.194	0.729	0.004	0.165
	d21	0.910	1.00	0.181	0.274
B. bruxellensis	**d07**	n/a	n/a	n/a	n/a
	d14	n/a	n/a	n/a	n/a
	d21	n/a	n/a	n/a	n/a
L. plantarum	**d07**	0.008	0.591	0.192	0.181
	d14	n/a	n/a	n/a	n/a
	d21	n/a	n/a	n/a	n/a

Table S3: Fisher's Exact tests for *S. cerevisiae*-containing communities only (C02-C09). Tests completed separately for each output, day, and variable combination. Bonferroni corrected p-value significance threshold = 0.0031 (i.e., 0.05 / 16 tests).

	Day	richness	*L. thermotolerans*	*T. delbrueckii*	*S. bacillaris*
Viable yeast	**d07**	n/a	n/a	n/a	n/a
	d14	0.825	0.605	0.605	0.605
	d21	n/a	n/a	n/a	n/a
B. bruxellensis	**d07**	0.065	4.851e-06	0.1698	0.4923
	d14	0.067	0.817	0.247	0.035
	d21	0.231	0.295	0.034	1.00
L. plantarum	**d07**	n/a	n/a	n/a	n/a
	d14	n/a	n/a	n/a	n/a
	d21	n/a	n/a	n/a	n/a

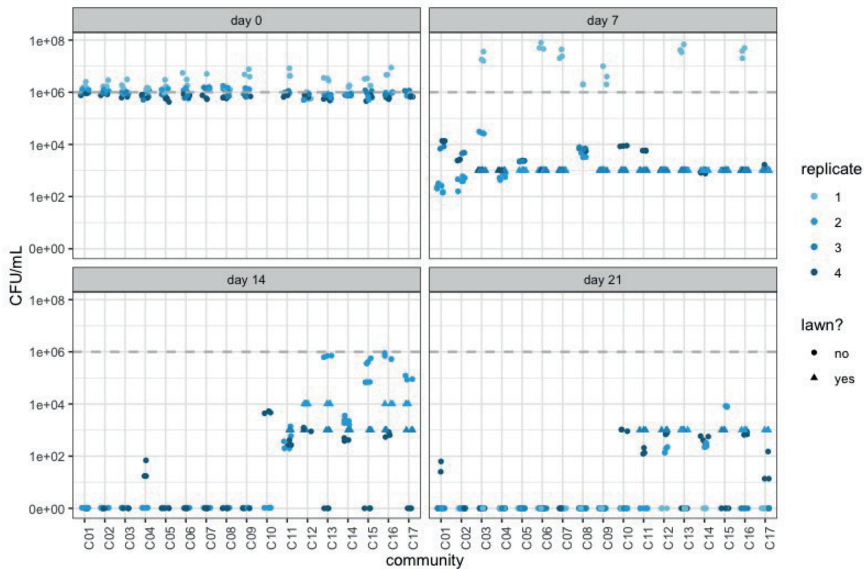


Figure S1: *viable yeasts remain in S. cerevisiae-free communities.* Estimated cell density of total viable yeast (excluding *B. bruxellensis*) over time from YEPD agar plating and counting after 48 hours. Dashed grey line indicates projected initial inoculated cell density at day 0 (i.e., 10^6 cells/mL). Lawns were conservatively estimated as the minimum possible colonies (i.e., 100) and the dilution factor (i.e., a lawn at 10^{-1} dilution = 10^4 CFU).

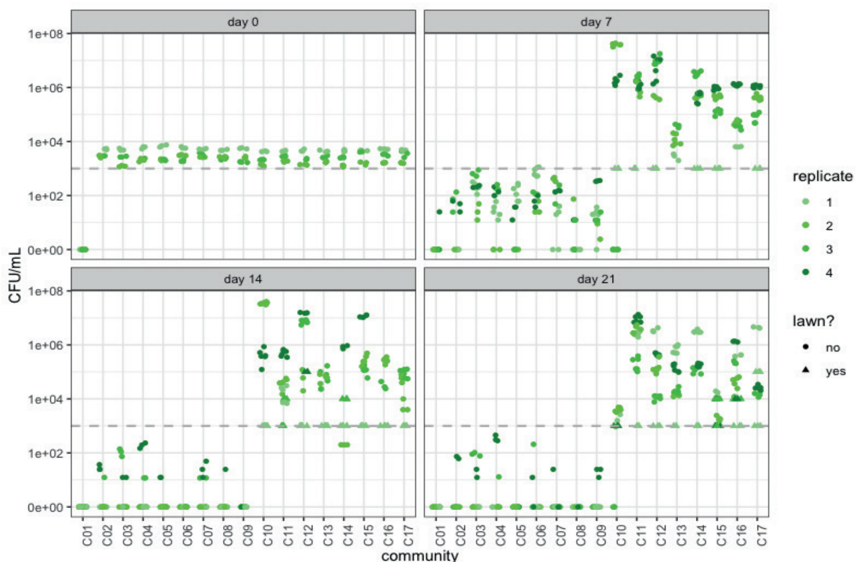


Figure S2: *B. bruxellensis establishes in the absence of S. cerevisiae.* Estimated cell density of *B. bruxellensis* over time from selective agar plating on YEPD supplemented with cycloheximide and counted after approximately 14 days. Dashed grey line indicates projected initial inoculated cell density of at day 0 (i.e., 10^3 cells/mL). Lawns were conservatively estimated as the minimum possible colonies (i.e., 100) and the dilution factor (i.e., a lawn at 10^{-1} = 10^4 CFU).

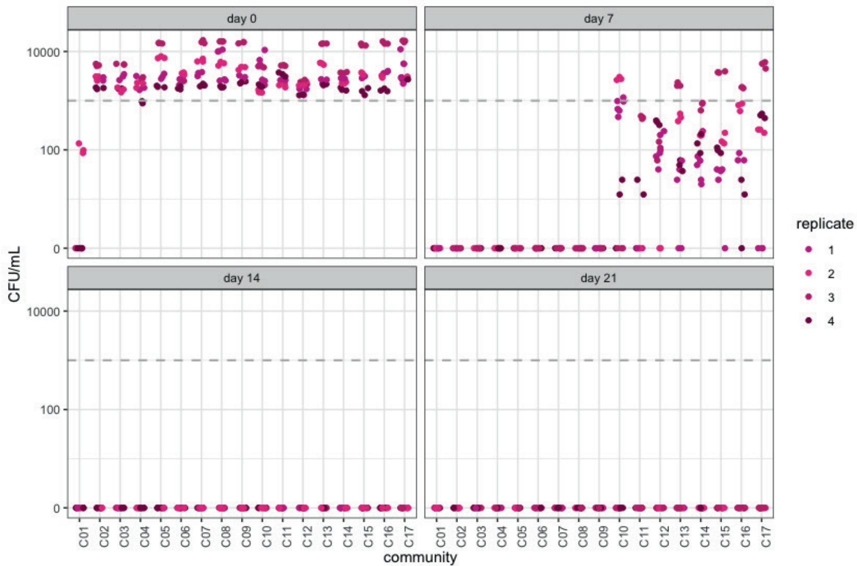


Figure S3: *elimination of L. plantarum is delayed in the absence of S. cerevisiae.* Estimated cell density of *L. plantarum* over time from selective agar plating on MRS supplemented with chlorophenicol and counted after 48 hours. Dashed grey line indicates projected initial inoculated cell density of at day 0 (i.e., 10^3 cells/mL). Estimated CFU of lawns were entered depending on the dilution factor used.

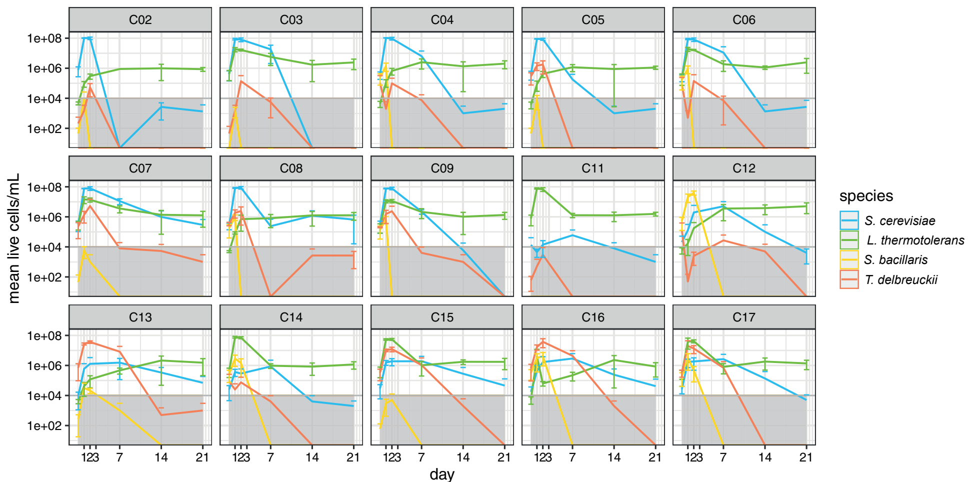


Figure S4: viewing main text's Fig. 3 on log10 scale reveals biased heightened live cell counts at days 14 and 21, especially for *L. thermotolerans* and *S. cerevisiae*. Fluorescent flow cytometry estimates of resident community yeast species overtime. The mean live cells/mL of four replicate rounds of fermentation is plotted, with error bars showing one standard deviation. Grey area indicates below 10^4 cells/mL, our proposed minimum reliable count (i.e., values below are unreliable).

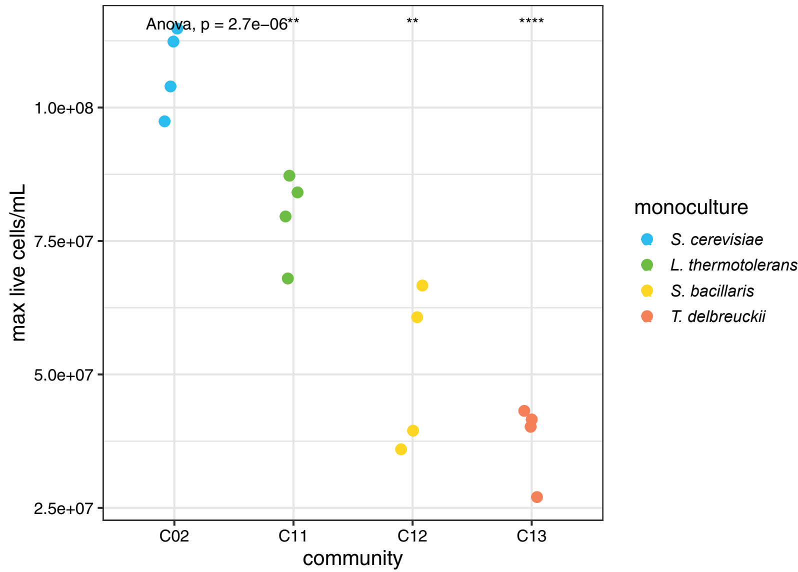


Figure S5: *L. thermotolerans' monoculture reaches the greatest maximum CFU/mL of the non-S. cerevisiae species.* Fluorescent flow cytometry estimates of maximum live cells/mL reached during fermentation for the monoculture of each resident yeast species. Replication = 4 fermentation rounds per community.

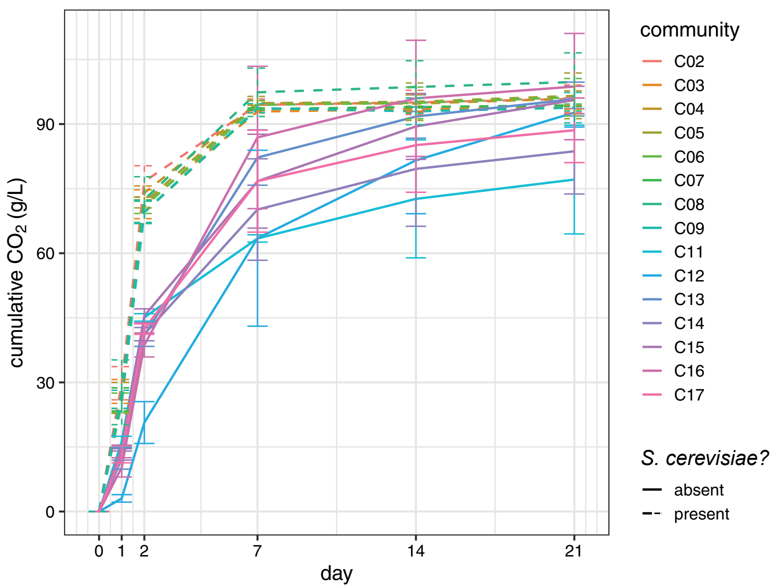


Figure S6: *fermentation is stunted in the absence of S. cerevisiae.* Cumulative CO_2 produced over time demonstrates fermentation progression in all 17 synthetic communities. Each line is the mean of four replicates for each community, with error bars showing one standard deviation. Dashed lines indicate communities containing *S. cerevisiae*, solid lines are *S. cerevisiae*-free communities.

Statistical analyses (Chapter 4):

1. Cumulative CO2:

Control communities C01 and C10 were excluded throughout. We did not have sufficient replication and degrees of freedom to test species richness 4.

Day 2:

lm(formula = CO2cum ~ Sc + Lt + Td + Sb + as.factor(richness))

Table S4: Output of linear model analysis of cumulative CO_2 at day 02.

	Estimate	SE	t-value	p-value
intercept	37.937	2.152	17.626	<2e-16
S. cerevisiae	32.449	1.650	19.669	<2e-16
L. thermotolerans	2.241	1.650	1.359	0.1800
T. delbrueckii	1.115	1.650	0.676	0.5022
S. bacillaris	-4.150	1.650	-2.516	0.0149
Richness 2	3.328	1.801	1.848	0.0702
Richness 3	2.106	2.454	0.858	0.3947
Richness 4	NA	NA	NA	NA

Day 21:

lm(formula = CO2cum ~ Sc + Lt + Td + Sb + as.factor(richness))

Table S5: Output of linear model analysis of cumulative CO_2 at day 21.

	Estimate	SE	t-value	p-value
intercept	89.2160	2.5601	34.848	< 2e-16
S. cerevisiae	5.1128	1.9623	2.606	0.01188
L. thermotolerans	-5.6880	1.9623	-2.899	0.00544
T. delbrueckii	4.3086	1.9623	2.196	0.03251
S. bacillaris	0.8643	1.9623	0.440	0.66140
Richness 2	2.5382	2.1420	1.185	0.24131
Richness 3	2.0704	2.9190	0.709	0.48126
Richness 4	NA	NA	NA	NA

Coefficient of variation:

```
## Remove rows with missing values in either 'CO2cum' or 'Sc'
df_no_missing <- na.omit(df2[c("CO2cum", "Sc")])
## Calculate beta diversity using Euclidean distance
variation <- vegdist(df_no_missing$CO2cum, method = "euclidean")
## Run PERMANOVA
result <- adonis2(variation ~ Sc, data = df_no_missing)
print(result)
```


Chapter 4

Permutation test for adonis under reduced model
Terms added sequentially (first to last)
Permutation: free
Number of permutations: 999
adonis2(formula = variation ~ Sc, data = df_no_missing)

Table S6: Output of PERMANOVA test demonstrates variation in cumulative CO_2 differs between *S. cerevisiae*-free versus *S. cerevisiae*-containing communities.

	DF	SumSq	R2	F	p-value
S. cerevisiae	1	17299	0.03364	12.464	0.002
Residual	358	496879	0.96636		

All time points:
lm(formula = CO2cum ~ Sc + Lt + Td + Sb + as.factor(richness) + day)

Table S7: Output of linear model analysis of cumulative CO_2 for all communities and time points.

	Estimate	SE	t-value	p-value
intercept	19.6969	3.6370	5.416	1.13e-07
S. cerevisiae	13.6592	2.6405	5.173	3.88e-07
L. thermotolerans	-1.6775	2.6405	-0.635	0.526
T. delbrueckii	2.9111	2.6405	1.102	0.271
S. bacillaris	-0.7964	2.6405	-0.302	0.763
Richness 2	2.4290	2.8823	0.843	0.400
Richness 3	1.7110	3.9279	0.436	0.663
Richness 4	NA	NA	NA	NA
day	3.8529	0.1555	24.781	< 2e-16

Day 07 - S. cerevisiae-containing communities only:
lm_sacc <- lm(CO2cum ~ as.factor(richness), data = sacc)

Table S8: Output of linear model analysis of cumulative CO_2 at day 07 for *S. cerevisiae*-containing communities only (C02-C09).

	Estimate	SE	t-value	p-value
Intercept	94.4639	1.1887	79.470	<2e-16
Richness 2	-0.2490	1.3726	-0.181	0.857
Richness 3	0.6506	1.3726	0.474	0.639
Richness 4	-0.7602	1.6810	-0.452	0.655

2. Binomial plate count data: used to create supplementary Tables S1, S2, S3.

Example of Fisher's Test: (day 14, total viable yeast, S. bacillaris)

```
binary <- cfus2 %>%
mutate(present = ifelse(CFU > 0, "yes", "no")) %>%
filter(!is.na(present)) %>%
filter(day == "d14") %>%
filter(media == "yeast") %>%
filter(community != "C01") %>%
filter(community != "C10")

binary2 <- binary %>%
select(present, Sb)
table(binary2)
```

	S. bacillaris (Sb) present?	
Viable yeast present?	No	Yes
No	18	12
Yes	55	73

```
fisher.test(table(binary2))
```

Fisher's Exact Test for Count Data

data: table(binary2)

p-value = 0.1063

alternative hypothesis: true odds ratio is not equal to 1

95 percent confidence interval:

0.8245106 4.9201388

sample estimates:

odds ratio

1.982165

5

Chapter 5

General Discussion

1. Thesis overview:

Understanding the dynamics of communities in response to environmental changes is a core theme of community ecology research and is central to my thesis research. While fundamental questions pertaining to the mechanisms and consequences of species populations changes are interesting, applying this knowledge is equally critical in shaping perspectives as we face rapid anthropogenic global changes. Both short term and long-term dynamics of communities should be considered, or in other words, both ecological and evolutionary timescales, because genetic changes within community members can change interactions with consequent impacts on community function (Govaert *et al.* 2019; Hendry 2019). Experiments addressing both community composition and community function responses, with consequent impacts on ecosystem functions, are particularly needed for understanding adaptation of communities to increasingly extreme environmental events induced by climate changes (Turner *et al.* 2020; Otto 2023) and sudden anthropogenic perturbations (deforestation, aquatic pollution).

Experimental microbial evolution has proven to be a powerful approach in studying ecological and evolutionary processes for single species in a short period of time, with high replication, and control. But microbial evolution studies have only recently been in mixed species communities of two or more species. The general inspiration for this thesis was to perform experimental evolution styled propagation experiments using diverse microbial communities of multiple species. By doing so, I wanted to consider species in a community in a paralleled view to genetic variants, or genotypes, within a single species (Vellend 2016). More specifically, I aimed to test the relationship between community composition and community function over multiple generations. Focus was made on the diversity component of community composition (Chapters 2 and 4), with community function measured as metabolism (Chapters 2 and 3), and both were considered in relation to stability, or resistance, to an introduced species (Chapters 3 and 4). In my thesis I demonstrate the suitability of microbial communities of fermented foods as model experimental systems to investigate three processes of community assembly – timescale, diversity, and invasion (see Box 1).

An exciting research direction motivated during my thesis and briefly experimentally explored but not to completion, was creating a synthetic community of selected microbial isolates representing a Mabisi community. Here in this general discussion, I first present such a proposed synthetic community in detail, emphasising its goal for exploring concepts in ecology and evolutionary biology. The potential for community level selection and artificial selection of microbial communities is later considered, while proposing its established application already in fermented food production.

Questions that inspired my initial research plans but remain future directions to explore are then discussed, including how they could be investigated using microbial communities. Lastly, applied take-aways for traditional food fermentation from my thesis research are mentioned.

Box 1: summary of research chapters

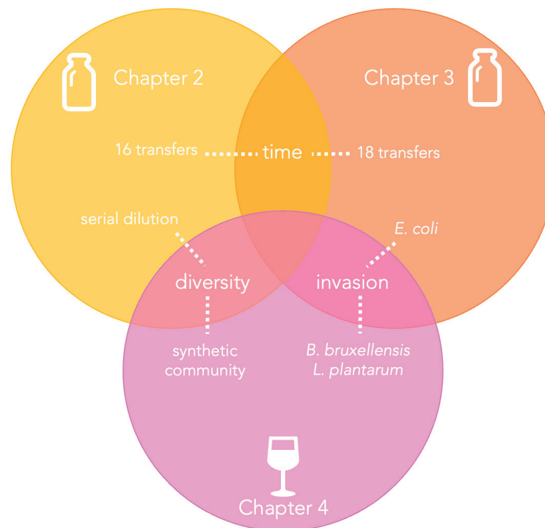


Figure 1: overview of thesis chapter contents

Diversity: Chapters 2 & 4

The effect of diversity on function was evidenced to be underlaid by identity, or inclusion, of certain species in both Chapters 3 and 4. Yeast species *Geotrichum candidum* played a main role in division of metabolic profiles of Mabisi communities in Chapter 2, whereas in Chapter 4 *Saccharomyces cerevisiae* was a keystone species for resistance against spoilage microorganisms in wine yeast communities.

Time: Chapters 2 & 3

Tracing changes in community composition and function over long-term ecological timescales was sought in Chapters 2 and 3, which both demonstrated gradual, not sudden, changes in community composition and function. Species sorting was detected in Chapter 2 where community bacterial compositions converged overtime and showed the importance to track recovery of communities after disturbance. However, the limits of 16S rRNA sequencing to detect meaningful compositional

changes in relation to community function were exposed, since far less could be concluded in Chapter 3 where near monocultures were seen (at least at 16S rRNA detection level were observed). Community function was also tracked overtime though metabolic profiles. Grouping of metabolic profiles based on treatment did not shift overtime in Chapter 2 nor Chapter 3 – any divisions in function observed at initial time points remained throughout propagation. Recovery in metabolic profiles did occur when *Escherichia coli* introduction was ceased in Chapter 3.

Invasion: Chapters 3 & 4

The system dependency - both biotic and abiotic factors of both invader and resident community - of invasions was supported in Chapters 3 and 4 as either failed or successful establishment of an introduced species were observed. In Chapter 2, the failed invasion of *E. coli* in Mabisi despite repeated high propagule pressure of inoculation was likely due to high acidity (abiotic stress) created by resident Mabisi community. Whereas in Chapter 4, the establishment of *Brettanomyces bruxellensis* or *Lactiplantibacillus plantarum* in wine yeast communities was dependent on a combination of biotic factors of both the resident community composition and invader identity (*L. plantarum* was transient in all communities, whereas *B. bruxellensis* persisted but only when *S. cerevisiae* was absent).

2. Controlled laboratory studies of eco-evo concepts using microbial communities:

2a) Advancing microbial experimental evolution:

Continuing from the thriving field of microbial experimental evolution in the 1990-2000s (Turner, Souza and Lenski 1996; Rainey and Travisano 1998; Buckling *et al.* 2000; Kassen 2002; Schoustra *et al.* 2006), investigating the ecology and evolution of multi-species or multi-strain microbial communities is similarly flourishing in recent years (Table 1) (Rivett *et al.* 2016; Cairns *et al.* 2018, 2020; Goldford *et al.* 2018; Piccardi, Vessman and Mitri 2019; Scheuerl *et al.* 2020; Piccardi *et al.* 2022, 2024; Weiss *et al.* 2022, 2023; Chang *et al.* 2023; O'Brien, Culbert and Barraclough 2023; Pourcelot *et al.* 2023, 2024; Ruiz *et al.* 2023). The tools to use diverse microbial communities for experimental evolution are establishing and the approach is sparking excitement; going forward now, we need to identify effective systems and use them to explore hypothesis driven questions. Experimental evolution of multi-species microbial communities can be a powerful tool to investigate fundamental eco-evo processes but also hold impactful biotechnological engineering applications. This thesis has motivated a Mabisi inspired model system, which I outline later in detail alongside another cyanobacteria model system; both would add to the growing

number of research groups using various model systems, a few of which I compare in Table 1. Every model system evidently has pros and cons, but I believe a Mabisi inspired system would be comparable to existing research while contributing novel aspects.

2b) Existing model systems:

I first expand here on two existing model systems - Cairns et al. (2018, 2020) and Piccardi et al. (2019, 2022, 2024) - which are particularly related to my objective to detect both genotype and species sorting processes within a multi-species community (Cairns *et al.* 2018, 2020; Piccardi, Vessman and Mitri 2019; Piccardi *et al.* 2022, 2024). Cairns et al.'s experiments have impressively maintained a highly diverse community of over 30 species in laboratory cultures while detecting novel mutants in response to antibiotic pressure. Their ability to assess intraspecies mutations while tracking species level dynamics is an exciting and inspiring achievement that encourages my aspiration to explicitly trace genotypes within a species to create a multi-species Muller plot (Fig. 2b). However, a simpler community with fewer species, such as four proposed species from Mabisi, would benefit this goal. Neither Cairns et al's consortia choice of community members nor their artificially defined media reflect a specific natural environment, which can raise questions about translating findings to the real-world and applied questions. Furthermore, their time scale of 54 generations is commendable and sufficient under strong selection of antibiotic selection, but I argue longer propagation is desirable for investigation of evolutionary changes in milder novel environments. In comparison, Piccardi et al.'s experiments use a much simpler synthetic community of four species in an undefined metal working fluids media. Their four species used can be selectively plated and isolated from one another, which is advantageous for tracing genotypes to a particular species and double-checking the absence versus presence of types. A minor limitation is that bacterial growth in their metal working fluid is relatively slow at a 7-day transfer time. Despite slower growth, Piccardi et al.'s experiments have impressively reached 300 generations, which is a substantial period for predicted evolutionary dynamics from genotypic changes to arise. The undefined nature of metal working fluids, just as with milk (Chapters 2 and 3 of this thesis) can limit research questions regarding environmental selection pressures. An undefined media means that its resources (carbon, nitrogen, etc.) cannot be easily manipulated to test consequent effects on community traits, whereas an artificial and defined media opens more research directions. Nonetheless, Piccardi et al.'s use of a real-world substrate is valuable with applicable insight for bioremediation of metal working fluids.

This thesis' research, alongside the work listed in Table 1 and many others, has inspired me to conceptualise two potential future model microbial systems. I next elaborate on

these two concrete examples, which both have fundamental and practical implications. The first is a synthetic community inspired by Mabisi, a traditional fermented milk, to address specific questions on the contribution of short-term (ecological) processes and long-term (evolutionary) processes to adaptation in multi-species communities. The second uses environmentally relevant multi-strain populations of cyanobacteria to focus more on a community level function – that of toxin production contributing to harmful cyanobacterial blooms.

Table 1: comparison between some other synthetic microbial community model systems.

Paper(s)	# species	# cycles / generations	Transfer time	Media	Environment resembled	Members
Cairns et al. 2018 Cairns et al. 2020	33-34	12 cycles ~ 54 gen.	4 days (at 10%)	Nutrient rich (PPY), low nutrient (M9 + 1% KB)	Not mentioned	bacteria
Piccardi et al. 2019 Piccardi et al. 2022 Piccardi et al. 2024	4	44 cycles ~300 gen.	7 days	Metal working fluids (undefined)	Bioremediation	bacteria
Goldford et al. 2018 Chang et al. 2023	12	12 cycles ~84 gen.	2 days	M9 / M9-glucose	Plant leaf and soil	bacteria
Fetzer et al. 2015	12	1 cycle	3 days	Mineral salt + benzoate + NaCl	Not mentioned	bacteria
Weiß et al. 2023 Weiß et al. 2022	12	4 cycles	1 day	AF media + modified versions	Gut microbiota	bacteria
a) O'Brien et al. 2023 b) Rivett et a. 2016 c) Scheuerl et al. 2020	a) 2-15 b) 1-16 c) 22	a) 1 cycle b) 7 cycles c) 22 cycles	7 days (at 10%)	a) iron-limited LB b) c) boiled leaf tea media	Water filled beech tree holes	bacteria
a) This thesis Ch. 4 b) Pourcelot et al. 2023 c) Pourcelot et al. 2024	a) 6 b) 5 c) 6	1 cycle	3 to 7 days	Synthetic grape must	Grape must	yeast
Ruiz et al. 2023	60	1 cycle	3 to 7 days	Synthetic grape must	Grape must	yeast
Mabisi inspired (proposed)	4	24 cycles ~ 150 gen.	3 days	synthetic (MRS + M17 + YPD mixture) or milk (undefined)	Fermented milk	bacteria & yeast
Cyanobacteria (proposed)	1 (multiple strains)	6 cycles	~ 5 days	Synthetic "COMBO" lake media	Toxic cyanobacterial blooms	bacteria

3. Tracing ecology and evolution during adaptation in multi-species communities:

3a) Background:

Current climate crises continually remind us of our world's rapidly changing environments and lead us to ask ourselves - how will species communities respond? The classical view of adaptation to changing environments focuses on individual species and their potential to survive novel environments via genetic changes (i.e., adaptive evolution) (Hughes 1999). However, this overlooks the necessity of considering species in the context of an entire co-evolving community, both in the short and longer term. Environments in nature harbour complex species communities. Yet, because of this complexity, eco-evo dynamics of complex communities has thus far been mainly limited to theoretical approaches and model simulations (Govaert *et al.* 2019) but should also be experimentally explored and validated. Building on my experimental studies and research on long-term-ecological time scales of multi-species communities (Chapter 2 & 3) I now outline here a potential approach for investigating further on evolutionary timescales.

A community can respond to an altered environment by adjustment of relative species abundances through the process of species sorting (ecology) and may adapt by spontaneous mutations spreading in individual species (evolution) - with the relative importance gradually shifting from ecology to evolution over time. Both processes are traditionally studied in isolation, meanwhile growing evidence supports that in the context of adaptive responses to change, ecology and evolution should be considered simultaneously (Govaert *et al.* 2019; Hendry 2019). Muller plots are an elegant visualisation commonly used to track the emergence, spread, and fate of genetic variants in evolving monocultures (Maddamsetti, Lenski and Barrick 2015). While Muller plots are currently restricted to single species, I aspire to develop a model system that extends the concept to a multi-species community. Essentially, I have been imagining how to achieve a "Muller-like plot" with multiple species, in which a different colour is a particular species abundance, and different shades of that colour are genotypes within that species (Fig. 2b). I propose establishing an experimental system using a defined microbial community of multiple bacterial species and serially propagate it in a novel environment over 100s of generations. Doing so could establish a proof-of-concept experimental system and perform a proof-of-principle experiment of serial propagation in a novel environment to test how and when ecological and evolutionary changes co-occur.

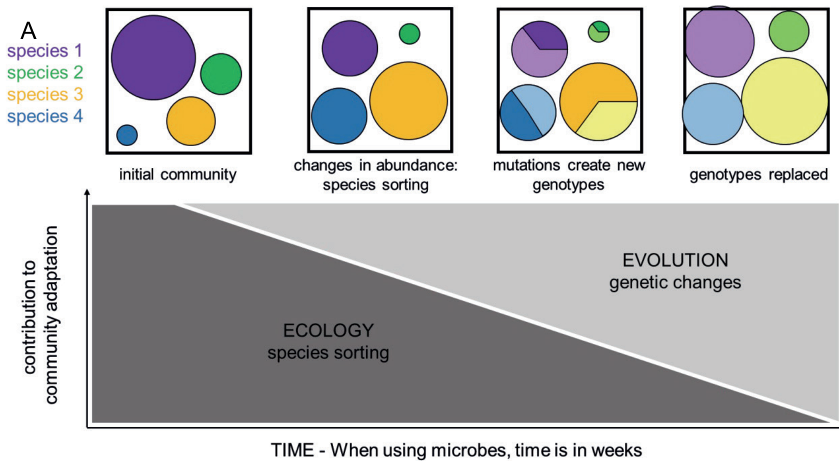
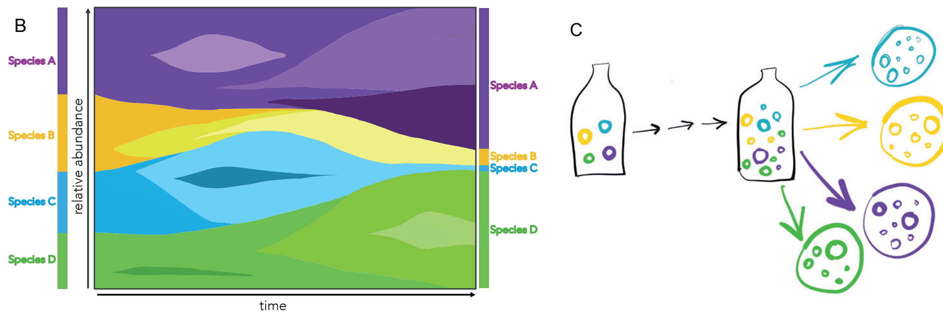


Figure 2: a) A community of four species can stably co-exist in many environments. Upon a change in environmental conditions, the community may initially change in relative abundances of species (species sorting - ecology), where species intrinsically better adapted to the new conditions increase in frequency. Upon prolonged selection, novel adaptive genetic changes (mutations - evolution) may occur, further shaping the communities and changing relative species abundances.



b) Depiction of an imagined Muller-type plot of multiple species (each colour) with sub-genotypes within (shades of colours). **c)** Visual depiction of propagation and selective agar plating of four species, each indicated by a unique colour. The propagation experiment would be initiated by a single isogenic clone of each species.

3b) Potential approach:

As a model system, I suggest a defined set of microbial species taken from natural microbial communities of Mabisi to track contributions of ecology and evolution on species and community level adaptation. For example - *Lactobacillus helveticus, Acetobacter orientalis, Streptococcus thermophilus,* and *Geotrichum candidum*. Classical methods of microbiology (selective strain isolation, colony counts) could be combined with high throughput DNA sequencing to track relative species abundances and identify new mutations in individual strains and species. I would eventually allow for experiments over approximately 500 bacterial generations to now include an

evolutionary timescale. Taken together, this approach would be innovative by combining 1) using four species that represent a natural community, and 2) tracking both species sorting and genetic variant dynamics.

Work plan:
1. *Mark the four co-existing strains with selectable markers.* The markers, for instance an antibiotic resistance or unique ability to use certain growth compounds, allow selective isolation from the community. This allows tracing their frequencies and isolating individual variants to track new mutations. Flow cytometry detecting separate species is a possible alternative, however the ability to isolate species also could allow testing evolved versus ancestor community members in various combinations.

2. *Co-culture the four strains in a serial propagation experiment under two selective environmental conditions.* I would initiate replicate communities in a defined synthetic milk media (40 mL), eight grown at 25°C and eight at 37°C. These temperatures represent two distinct conditions known for their differing selective effect (Moonga *et al.* 2021). Populations would be serially transferred at 1% portions to fresh medium every 3 days for initially 24 rounds of propagation (3 months, ~160 generations). Samples would be regularly frozen for later analysis and continued propagation until approximately 500 generations, or more.

3. *Map changes in relative abundance of the four strains (ecology) and novel mutations (evolution).* Samples would ideally be analysed at regular timepoints for relative species abundances using selective plating colony counts (Fig. 1c) or flow cytometry. Even if selective plating proves troublesome, the profiles of entire communities at initial, middle, and end points would also be completed through metagenomics to reveal population wide novel mutations; this would be validated by and combined with whole genome sequencing of random colony isolates ("individuals" of the community) to assign mutations to certain species.

3c) Innovative approaches & impact:

Contributing to the field of eco-evolutionary biology, this proposed approach extends concepts from classical experimental evolution of single species to study the complexities of community level changes. Successfully creating a microbial community, with traceable members, in a defined and manipulatable environment, would provide an effective opportunity to experimentally study general biological processes. The focus of this proposed work - dynamics between ecological (species sorting) and evolutionary processes (genotypic changes) - is a core, yet not well understood, topic across biological research (Govaert *et al.* 2019; Hendry 2019). Appreciating the capabilities of communities to respond to environments beyond

genetic changes within single species must be understood in the face of rapid global environmental changes. However, most natural communities cannot be easily studied directly, nor be reproduced in laboratory environments; a natural (fermented food) inspired microbial community enables such a combination for experimental evolution. Investigating the dynamics between ecological and evolutionary processes is highly relevant for understanding community responses to change in natural and managed systems on short and longer terms. The fundamental insight gained can also be used for ecology driven "top down" design of starter cultures for small-scale traditional food fermentation (detailed further in the following section "Practical implications").

Successfully setting up a proposed system of four fermentative species in a synthetic milk broth would open numerous experimental and theoretical research directions. This could include 1) investigating repeatability of species sorting and genotypic changes concurrently following selection pressure through manipulating our defined media (i.e., novel sugar), 2) mapping of specific mutations and their roles in overall community performance, 3) linking of community species profiles to community function in their metabolic output and resilience over ecological and evolutionary time, and 4) comparing laboratory experimental evolution to the natural world. The experiments could be repeated with real milk, either in the laboratory or in a field experiment - there are few microbial experimental evolution model systems where such direct comparisons to the natural environment can be made.

Also, the moderate diversity of Mabisi (6-10 major species in environmental samples (Schoustra *et al.* 2013; Moonga *et al.* 2020; Groenenboom *et al.* 2022) means that there is a minimal bottleneck of complexity between natural versus laboratory communities, which is a common pitfall when working with immense diversity in soil or gut microbiomes (Moeller *et al.* 2014; Thompson *et al.* 2017; Jurburg *et al.* 2024; Čaušević *et al.* 2024) where successful laboratory culturing is difficult (Hirano *et al.* 2023; Kapinusova, Lopez Marin and Uhlik 2023). Lastly, most experimental microbial communities use only bacterial members (Table 1), yet eukaryotic microorganisms (fungi, protozoa, algae) are also components of many microbial ecosystems (Pepper and Gentry 2015; Rosenberg 2021). The combination of bacteria and yeast (*G. candidum*) in a Mabisi inspired synthetic community would be a unique aspect.

4. Synthetic cyanobacterial communities to study toxic algal blooms:

Many thanks to David Johnson and Francesco Pomati at Eawag - Dübendorf, CH (Swiss Federal Institute of Aquatic Science and Technology) for their feedback, insight, and enthusiasm when drafting the section below.

4a) Background:

Cyanobacteria are an important pillar of many aquatic food webs and global nutrient cycles. However, cyanobacteria pose environmental and health concerns when flourishing in blooms of high cell densities, where toxins can accumulate to harmful concentrations. Cyanobacterial blooms in aquatic systems have existed for millions of years, but concern is rising as their frequency, intensity, and duration are rapidly increasing globally (Huisman *et al.* 2018). Only some cyanobacterial blooms reach harmful toxin concentrations, which leads to the question - why do some blooms turn toxic while others do not? Harmful cyanobacterial blooms (cyanoHABs) are predicted to further increase in frequency and intensity due to continued global climate and anthropogenic resource use changes (Huisman *et al.* 2018; Hellweger *et al.* 2022), yet the main underlying mechanisms are still debated, which hampers our ability to predict and manage them. Cyanobacterial blooms are complex events encompassing dynamics driven by physiological, ecological, evolutionary, and chemical processes. Laboratory experiments, when grounded in real world scenarios, can provide controlled tests on the processes leading to cyanoHABs and improve monitoring techniques for predicting their overall intensity and toxicity.

The toxicity of cyanoHABs is the result of the production of bioactive secondary metabolites (i.e., those not related to primary metabolism) by certain cyanobacterial strains within and across species. A key aspect to cyanoHABs is that the production of toxic metabolites varies both within and between populations, which poses difficulties for predicting the overall toxin load and the toxicity of blooms based only on taxonomy. CyanoHABs are often simultaneously composed of toxin and non-toxin producing strains (Hellweger *et al.* 2022), but neither the total abundance of cyanobacteria nor the species composition of a bloom is a good predictor of toxin accumulation. While governmental policies typically aim to reduce the total size of cyanoHABs, selection for reduced total cyanobacterial loads may unfortunately contrast with selection for toxic strains, inadvertently increasing bloom toxicity (Hellweger *et al.* 2022).

Recent reviews and synthesis articles have indicated environmental factors implicated in driving the dominance of toxic strains under blooming conditions. Factors highlighted include nitrogen load, phosphorus limitation, atmospheric CO_2,

temperature, and light intensity (Rigosi *et al.* 2014; Huisman *et al.* 2018; Hellweger *et al.* 2022). Clearly, all these factors are aggravated in deep lakes by climate change (Posch *et al.* 2012; Huisman *et al.* 2018; Merz *et al.* 2023); yet empirical proof is lacking on how these factors favour toxic versus non-toxic cyanobacterial isolates and influence toxin production (concentration and diversity) under realistic cyanoHAB conditions. Furthermore, past experiments have been largely based on cyanobacterial responses in monoculture, with rare exceptions (Kurmayer, Deng and Entfellner 2016; Hellweger *et al.* 2022). Studies using mixed populations are therefore needed to investigate the physiological responses and strain interactions driving turnover of toxic and non-toxic cyanobacterial isolates.

4b) Objectives:

I propose here a model laboratory system to identify the underlying mechanisms driving dynamics of diverse synthetic populations of known composition (multiple toxic and non-toxic strains) under conditions relevant to current global changes in lakes, namely variation in temperature and light intensity (Wells *et al.* 2015). These two factors are hypothesised to be among the most important driving changes in genotype composition of cyanobacterial blooms, as they influence the overall cell oxidative stress and associated responses (Hellweger *et al.* 2022). Temperature is a main factor that influences physiological and metabolic processes, but data on relative fitness of closely related toxic and non-toxic cyanobacterial strains under varied temperatures in diverse populations is lacking (Wells *et al.* 2015). Light intensity is predicted to increase due to global climate changes, which is especially relevant since cyanobacterial strains vary in tolerance to UV exposure, ensuing tolerance to oxidative radicals and toxin production responses (Van de Waal *et al.* 2011; Hellweger *et al.* 2022). Like temperature, there is little data about the impacts of light intensity on cyanobacterial population structure in diverse populations.

The experiment would address three specific knowledge gaps:

Q1) How do temperature and light conditions affect the relative fitness of toxic genotypes within a diverse cyanobacterial population?

Q2) How does the diversity of toxic genotypes relate to the diversity and abundance of toxin metabolites produced across a cyanobacterial community?

Q3) Do such genotypic and metabolic shifts change over time during development of a cyanoHAB?

4c) Work Plan:

Synthetic community propagation:

An existing strain collection of *Microcystis aeruginosa* strains sampled across northwestern Europe would be used, which have all been successfully cultured at Eawag and had their genomes sequenced. The *Microcystis* genus is an abundant contributor to cyanoHABs locally in Switzerland and the most important toxic bloom-forming cyanobacterium globally (Huisman *et al.* 2018; Dick *et al.* 2021). Defined synthetic populations would be assembled using multiple *M. aeruginosa* strains determined to be toxin producing or not, with the strain diversity and composition varying. Each synthetic population would be grown with appropriate replication under controlled laboratory conditions in a factorial design of two temperatures and two light intensities. An already established synthetic media (Kilham *et al.* 1998) mimicking temperate lake conditions would be used to transfer 1% of each replicate population to fresh medium every five days for six rounds of transfers, spanning one month - a relevant time frame for cyanobacterial blooms in nature (Huisman *et al.* 2018). Samples would be taken at the beginning, middle, and end points, with the following measurements performed:

1. DNA extraction / whole population community metagenomics (subset of replicates):
   a) Genotype abundances & diversity
   b) Relative genotype fitness, calculated from relative abundances over time
2. Metabolomics/HPLC: secondary metabolites (diversity and abundance of toxins)
3. Flow cytometer: cell counts / total population size (bloom intensity)

The common use of 16S rRNA to track relative abundances at the genus level, or species level if fortunate, does not provide useful information for predicting toxin production in cyanoHABs. Variation in toxicity occurs across genotypes and not just at the genus or species level, so instead a more detailed quantitative assessment of genotypes relevant to toxin production is needed; my proposed experiment aims to achieve this by combining metagenomic with metabolomic data of diverse cyanobacterial populations. It is also important to capture the temporal scale of genetic and metabolic dynamics during the development of toxic blooms under changing environmental conditions to identify causal mechanisms.

4d) Innovative approaches & impact:

Firstly, this proposed approach to *assess multiple strains and stressors*, not just single strains and single environmental pressures, allows possible detection of interactions and better mimics the complexity of natural conditions under global change scenarios, specifically for deep peri-alpine lakes (e.g., Switzerland). Research on the physiology of individual monoculture strains is well established, but multi-strain populations are understudied. Secondly, the proposed work would fill a major knowledge gap

regarding the lack of studies *incorporating metabolomics with genetic data* (Burford *et al.* 2020). Such an assessment of overall community level toxin production is an initial step to detect potential trade-offs and associations between total bloom size versus toxicity (Hellweger *et al.* 2022). Combining metabolomic and metagenomic data can help understand if, and how, diversity in toxic genotypes relates to diversity and abundance of toxic metabolites - a critical connection to know for establishing predictive monitoring techniques. Thirdly, it would contribute to expanding the *power of metagenomics to monitor cyanoHABs* eventually in real lakes and reservoirs by initially testing in a controlled laboratory system. Developing metagenomic tools for monitoring cyanoHABs and seeking assessment for potential toxin production would contribute to improving sustainable management and use of our water systems in the face of global climate change. A fourth exciting potential outcome is that through development of metagenomic monitoring to track toxin-production genes, *lateral gene transfer of toxin gene(s)* between toxic to non-toxic strains may be detected. Lateral gene transfer is known to contribute to the spread of antimicrobial resistance genes (McInnes *et al.* 2020), so the process also possibly occurs in cyanoHABs.

It is critical to investigate the mechanisms behind climate change-induced effects on aquatic ecosystem services, such as concerns related to the management of cyanoHABs in lakes and reservoirs. Research on the adaptation of diverse cyanobacterial populations to climate-change induced blooming conditions would provide insight into why some cyanobacterial blooms turn toxic, which is a critical problem facing the management of Swiss lakes and similar lakes globally. My approach of investigating the dynamics between ecological, evolutionary, and physiological processes of cyanoHABs would strengthen collaboration between disciplines to advance fundamental and applied aquatic ecosystem research. The tractable and controlled data produced from the experiments proposed here would benefit the development of modelling predictions for application in water systems management. Furthermore, comparing laboratory experiment results to field observations of researchers would be a powerful approach to identify the mechanisms contributing to rising cyanoHABs in Switzerland and elsewhere. Combining my laboratory findings with field experiments and computer modelling would provide key considerations for environmental policy recommendations.

5. Outlook on ongoing directions:

5a) Can community level selection, or "breeding", of microbial communities exist?

There is growing enthusiasm in exploring if and how selection can act at levels above individuals comprising a single species in microbial communities (Blouin *et al.* 2015; Arias-Sánchez, Vessman and Mitri 2019; Raynaud *et al.* 2019; Doulcier *et al.* 2020; Xie and Shou 2021; Sanchez *et al.* 2023). Community-level selection refers to the process by which certain traits or characteristics become prevalent in a community due to selective pressures acting on collective traits, rather than on individual traits. Community-level selection considers how the composition and interactions within a community influence its overall fitness and persistence. Philosophising a community of individual species to be one "superorganism" has been debated for nearly a century (Clements 1936; Wilson and Sober 1989; Okasha 2008; Liautaud *et al.* 2019) and the perspective is particularly relevant when considering the evolution of microbial communities, where the fitness of a member of one species is intricately linked to its interactions with the traits of other community members.

Community-level selection requires 1) variation in the community trait of interest, 2) a link between this variation and "fitness" (community function that can be selected), and 3) heritability (Lewontin 1970, Raynaud et al. 2019). Heritability is especially critical but less conclusive at the community level – it means that if parental communities are deconstructed, then reassembled, these offspring communities will again reproduce a similar function to parental communities. Doulcier et al. 2020's modelling paper nicely demonstrates and argues that such heritability can arise from selection on interactions (Doulcier *et al.* 2020). Xie et al. 2019 modelling paper also highlights the importance of interactions but notes that it is incredibly complex to identity interactions involved in functions, never mind modify them for directed evolution purposes. Conceptualising selection at higher hierarchies poses the challenge to translate concepts traditionally found in population genetics to community ecology (Vellend 2016), such as the fitness landscape (Sanchez *et al.* 2023). Several questions with unclear answers arise for theorising fitness landscapes for complex communities, such as what is the "fitness" of a community, and how to measure it?

Does a community level phenotype, or function, compare to genotype fitness (Sanchez et al. 2023), such that selection acts on community members for functional traits just as genotypes with differing fitness are selected (Goodnight 1990; Williams and Lenton 2007; Doolittle and Inkpen 2018)? Like in population genetics' "genotype-fitness" landscapes, Sanchez et al. 2024 describe "community-function" landscapes for understanding the emergence of collective microbial community functions from

interactions. A major complication of developing a function landscape for diverse communities is that there is also an embedded hierarchical network of fitness landscapes for each individual replicating species within the community. A community level fitness function of interest can be applied in experimental conditions, such as degradation of a compound (Swenson, Arendt and Wilson 2000; Wright, Gibson and Christie-Oleza 2019), plant biomass (Swenson, Wilson and Elias 2000), CO_2 emissions (Blouin *et al.* 2015), microbial biomass (Raynaud *et al.* 2019), or colour (Doulcier *et al.* 2020). In these cases, the community fitness is defined and is measured by the unit of interest (i.e., biomass, colour) (Box 2).

It is an exciting direction to reflect if microbial communities can be engineered to improve biotechnology potential, for instance in fields of the environment (degrade pollutants, biofertilizer efficacy), animal health (probiotic efficacy), nutrition (improve vitamin content), and food quality (novel flavours, property combinations) (Tamang 2015; Malusà, Pinzari and Canfora 2016; Sheth *et al.* 2016; van Rijswijck *et al.* 2017; Arias-Sánchez *et al.* 2024). Here, community selection may be powerful because of the potential to select valuable combinations of genes and their functions that do not exist together in an individual species' population. An especially important aspect is the evolutionary stability of a community over time so that its function of interest persists, remaining stable over generations. For example, limiting potential changes in functional output of bioreactors during long time periods, even with environmental perturbations. Similarly for fermented food production, producers want to ensure that starter cultures lead to a reliable product (discussed more in later section "Practical implications"). Using evolutionary selection approaches to optimise communities, perhaps through favouring certain interactions, might be an approach to strengthen stability and persistence.

Exploring community level selection in fermented foods:

Community level selection of microbial communities has effectively been applied for centuries in back-slop production of fermented foods. Back-slop propagation is still used in smear cheeses (rind of aged cheese is washed and transferred to young cheese) (Bockelmann *et al.* 2005; Mounier *et al.* 2017), parmesan cheeses (whey is transferred to initiate the next fermentation) (Bottari *et al.* 2010), or sourdough ("active" back-slop using a small amount of first ferment) (Gänzle, Qiao and Bechtner 2023), and Mabisi ("active" back-slop, or "passive" where buckets or calabashes are not washed with soap) (Moonga *et al.* 2019). I consider community level selection to be dependent on the function of interest not possibly being performed by one strain alone; therefore, characteristics of fermented foods such as aroma, flavour, colour, and texture are community level product properties – that is, no single member can produce them. These mentioned properties of fermented foods are also linked to

community members and their evolution. As community members mutate or change in relative abundances, their contributions to the food characteristic of interest may be altered. Product properties are determined by the entire microbial community and thus can be a proxy measure of the microbial community function; hence another interesting research inquiry regards the fate of species compositions following selection in relation to a community level property (i.e., function). Connecting the final community compositions to their selected characteristics can uncover key species, functional groups, and their associations. Lastly, the traits are heritable since a final community produces a similar aroma, texture, flavour, etc. – this is exactly a reason why humans have used such propagation approaches for centuries.

Box 2: potential community level selection approaches

Thank you to Duur Aanen for his discussions on community level selection and feedback for drafting Figure 3.

There lies a difference between traditional passive “experimental evolution” and active “artificial selection”. All communities are transferred in traditional experimental evolution regardless of measured community function of interest (Fig. 3a), whereas in artificial selection, only the top-performing communities are used to inoculate next generations (Fig. 3c and 5d). Artificial selection for a community level function can furthermore take two approaches of either pooling the top-performing communities together and propagating (Fig. 3c) or keeping individual top-performing communities intact (Fig. 3d); it is unknown whether the effectiveness of artificial selection differs between the two approaches. Beneficial interactions might be broken apart in a community mixing approach, yet it can potentially also introduce new beneficial combinations (Fig. 3c). Lastly, you may have community level selection without artificial selection when top-performing communities (in relation to function of interest) are not preferentially selected to inoculate next generations, yet high cell counts of more productive communities are favoured (Fig. 3b). Multi-level selection (MLS) 1 and 2 refers to definitions in Okasha 2008; MLS1 is defined as selection only on properties of particles (i.e., individuals), whereas MLS2 is selection on both particles and collectives (i.e., communities). In MLS1, the fitness of a collective is the average fitness of its particles (i.e., biomass, colony forming units). Comparatively in MLS2, the fitness of a collective is measured as the number of *offspring collectives*, regardless of their particle fitness.

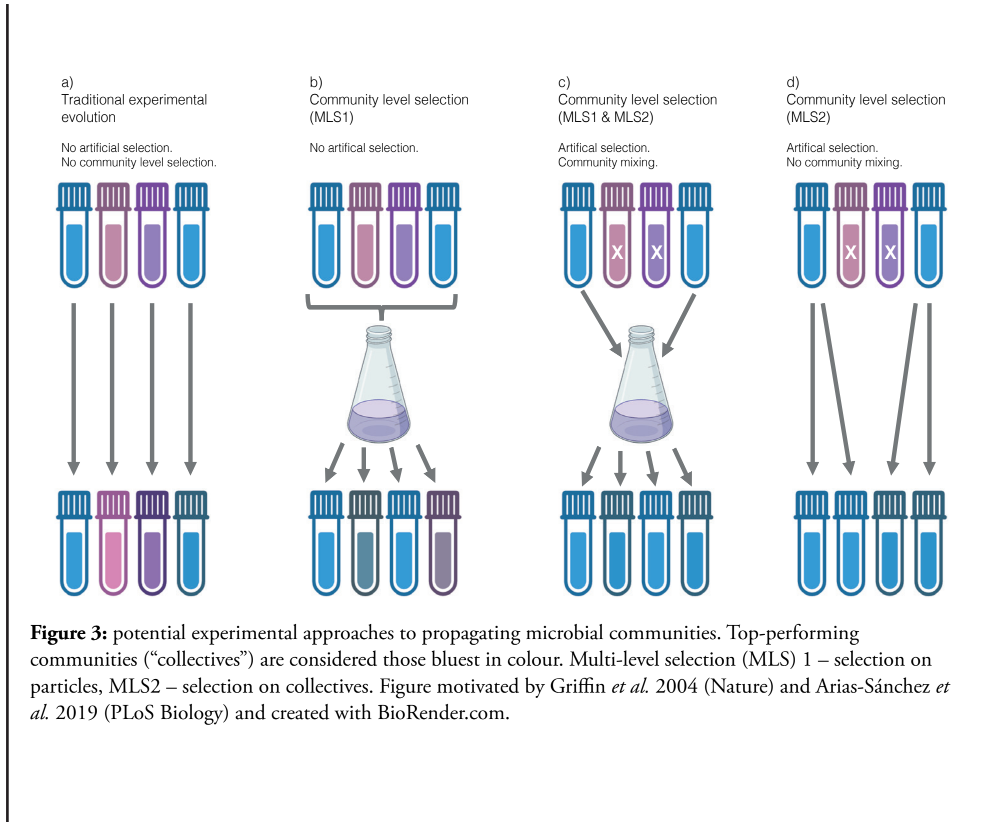


Figure 3: potential experimental approaches to propagating microbial communities. Top-performing communities ("collectives") are considered those bluest in colour. Multi-level selection (MLS) 1 – selection on particles, MLS2 – selection on collectives. Figure motivated by Griffin *et al.* 2004 (Nature) and Arias-Sánchez *et al.* 2019 (PLoS Biology) and created with BioRender.com.

Understanding how artificial selection via selective propagation can alter and optimise fermented food characteristics is important for future product upscaling and protection of traditions. For example, the microbial communities of starter cultures should produce a consistent and reliable product. I brainstorm here a project to help elucidate if selective propagation improves robustness in fermented food product qualities. A general concept could be to perform an experimental evolution experiment where replicate fermented food microbial communities are selected for collective functional properties, such as aroma compound combinations. Inspired by Doulcier *et al.* 2020's modelling research of selection on 50:50 blue to red bacterial cells to achieve a purple colour community trait, in fermented foods, a specific combination of aromas could serve as a suitable comparison. I imagine a distinct area on a PCA plot of aromas would correspond to a particular combination of community members (i.e., proportion acetic acid bacteria, lactic acid bacteria, and yeast) to achieve a particular metabolic profile. The aroma profiles would need to be sampled, monitored, and assessed in real time to select top-performing communities to "reproduce" at each propagation round (Box 2 – Fig. 3c, d). A critical point to consider is that the metabolic compounds present (i.e., proxy for community function) at a particular time point may not directly

correspond to the community composition at that moment. Assembly history can be influential (Sanchez *et al.* 2023) since any metabolites produced prior to the time of sampling will remain, even if the members previously producing them are no longer present in the community. Such a disconnect between community compositions and metabolic profiles was observed in my Chapter 3, where *E. coli* is not present in any community compositions at 72 hours yet its initial presence versus absence is "shadowed" in metabolic profiles.

5b) Do more diverse communities follow more convergent or divergent paths?

There are two predictions of the ecological and evolutionary trajectories of communities in relation to their starting diversity. The first view is that more diverse communities are expected to have a limited number of potential evolutionary trajectories due to reduced available niche space. Less diverse communities in comparison should have more evolutionary space to explore and diverge from their original state, in an adaptive radiation process (van Moorsel *et al.* 2021). If true, I would predict that Mabisi communities in Chapter 2 with higher starting diversity would exhibit stronger parallelism and convergence in final species composition and functional properties after repeated propagation. Conversely, the opposite may be true in the second view of "diversity begets diversity" (Meyer *et al.* 2010; Madi *et al.* 2020); greater standing diversity is less restrictive and provides opportunity for divergence of lineages across replicate populations. I did not detect either prediction in Chapter 2 where the starting diversity of Mabisi microbial communities were manipulated by serial dilution to remove, most notably yeast, as well as genotype variation. However, synthetic communities in Chapter 2 perhaps exhibit more convergence and similarity overtime in their aroma profiles, as estimated by their smaller PCA ellipse size (Leale *et al.* 2023).

When designing Chapter 2's experiment I did consider the secondary question of – would starting diversity influence the repeatability of composition and function across replicate populations overtime? Detecting divergence or convergence in community compositions was limited by assessing only the 16S rRNA gene, thus I wonder if hidden dynamics would be revealed at higher resolution through metagenomic sequencing. Bioinformatic analyses could be performed on metagenomes from archived samples of Chapter 2's laboratory experiment where Mabisi communities of four initial diversities were propagated for approximately 100 generations. The community compositions from metagenome sequencing could be compared between initial and final communities in the selection experiment, then the variation across replicate lines and divergence from initial populations also evaluated. Variation among replicates and divergence from the ancestral population would be compared

between the different starting diversities (full, medium, low, synthetic). Another approach to explore this concept in depth would be to create synthetic communities of isogenic clone combinations in varied levels of species and metabolic guild richness, then similarly track trajectories of community species compositions, genotype compositions, and metabolic profiles over time. See Table 2 for some hypothesised synthetic communities.

Table 2: hypothesised synthetic communities to test community divergence in relation to initial community diversity. Each species would be initially combined as isogenic clones. LAB: lactic acid bacteria, AAB: acetic acid bacteria.

Initial composition				Initial diversity	
Focal species (LAB)	**Other LABs**	**AABs**	**Yeasts**	**Species richness**	**Guild richness**
L. helveticus	*L. delbreuckii* *L. lactis*			3	1
L. helveticus		*A. orientalis* *A. pasteurianus*		3	2
L. helveticus	*L. delbreuckii* *L. lactis*	*A. orientalis* *A. pasteurianus*		5	2
L. helveticus	*L. delbreuckii* *L. lactis*	*A. orientalis* *A. pasteurianus*	*G. candidum*	6	3
L. helveticus	*L. delbreuckii* *L. lactis*		*G. candidum*	4	2
L. helveticus		*A. orientalis* *A. pasteurianus*	*G. candidum*	4	2

5c) Can evolutionary and ecological processes promote community stability over time?

When changes in community function are not observed, it cannot be assumed that underlying ecological or evolutionary processes are similarly static. So called “cryptic dynamics” has been proposed to describe the potential role of evolution in achieving ecological stability (Kinnison, Hairston and Hendry 2015; Hendry 2017). It is suggested that evolutionary processes can contribute to stability at the community level (Kinnison, Hairston and Hendry 2015; Hendry 2017). For example, if true, it is expected that communities initially unable to resist invasion will later exhibit increased stability after undergoing experimental evolution. Specialization, generalization, and diversification are evolutionary trajectories that can create complementarity in resource use or overall niche space to enhance exclusion of an invader (Gravel *et al.* 2011). Simply ecological processes (i.e., species sorting) could also independently increase stability against a novel species establishing since shifts in relative species abundances without genotype changes could improve total resource use and leave less available niche space. I did not observe increased stability overtime in neither Mabisi community composition nor function (metabolic profiles) against *E. coli* introduction

in Chapter 3. It could be argued that there were reduced shifts in function between control versus *E. coli* introduced communities after the 19 transfers, indicated by the distance between PCA ellipses in the two treatments.

A next step to investigate whether evolutionary and ecological processes may promote community stability over time is to perform the same experiment as in Chapter 3 but with a *Lactococcus* species repeatedly introduced. Using an ecologically similar species that is more likely to establish, unlike *E. coli*'s strong limitation by Mabisi's low pH, would better enable possibly detecting if community stability increases overtime. *Lactococcus* is not present in the natural Mabisi communities used in my thesis, however it is found in other Mabisi samples and can be detected by 16S rRNA gene profiles. Ideally, the *Lactococcus* species would also be marked (i.e., antibiotic resistance) for selective plating isolation at multiple time points, allowing 2 by 2 factorial invasion tests of naïve versus evolved *Lactococcus* into naïve versus evolved Mabisi communities. Stability, or resistance, would be measured as the inverse of *Lactococcus* abundance – determined by cell counts on selective agar plates and/or relative abundances by 16S rRNA gene sequencing.

5d) Moving beyond 16S rRNA taxonomic resolution:

The ease and inexpensiveness of sequencing the full length or variable regions of the 16S rRNA gene has made it the standard approach for profiling bacterial communities. Describing bacterial communities from extracted DNA allows revealing unculturable or low abundance species that go undetected with classical agar plating. While useful for descriptive studies of simply "who's there" in an environment, microbial ecologists are struggling to interpret community functions from 16S rRNA taxonomic compositions. The limits of describing bacterial communities through 16S rRNA sequencing has rising scepticism and is bringing forth conversation about better alternatives (Burke *et al.* 2011; Louca *et al.* 2016, 2018; VanInsberghe *et al.* 2020; Reynolds *et al.* 2023) . Our current profiling approaches are lacking in ability to describe community functions, or niches, because we only see the taxonomic level (Burke *et al.* 2011; Malard and Guisan 2023; Reynolds *et al.* 2023). A contributing factor is high functional redundancy across bacterial taxa, so that variation at the 16S rRNA level (or even species level) tells a limited story about the community's capacities (functions) (Burke *et al.* 2011; Louca *et al.* 2016, 2018; Tully *et al.* 2018). Functional redundancy is defined in Louca *et al.* 2018 as "the coexistence of multiple distinct taxa or genomes capable of performing the same focal biochemical function". Metabolic pathways of bacteria are not closely related to taxonomic groupings and are instead found across various clades, meaning that variation in metabolic capacities occurs across genotypes and not just at the genus or species level (Burke *et al.* 2011; Louca *et al.* 2016, 2018; Tully *et al.* 2018). Furthermore, there can be meaningful

functional variation across strains of one species; for example, in *Lactobacillus helveticus*, two protease-type strains exist (Broadbent *et al.* 2011), which can or cannot break down casein and grow in milk. Such differing protease activities between types of *L. helveticus* is just one clear example of significantly different functional abilities that are hidden in 16S rRNA gene sequencing profiles. Bacterial community profiles from 16S rRNA sequences are arguably more insightful for taxonomically diverse communities of soil or gut, where more variation and differentiated functions exist at the genus or phyla level (Moeller *et al.* 2014; Thompson *et al.* 2017; Jurburg *et al.* 2024; Čaušević *et al.* 2024). But alternative sequencing methods are needed to track community dynamics of ecological meaning (Burke *et al.* 2011; VanInsberghe *et al.* 2020), especially in communities where species are genetically similar, such as lactic acid bacteria in Mabisi (Leale *et al.* 2023, 2024). This thesis demonstrates the pitfalls of 16S rRNA sequencing, especially in Chapter 3 where the taxonomic resolution in bacterial communities showed near single species monocultures, despite divisions in metabolic profiles across communities of similar composition. It is hard to imagine that there are not multiple genotypes within the "monocolour blue blocks" of *Lactobacillus* B in Chapter 3 (Ch.3, Fig. 2) (Leale *et al.* 2024), so improved resolution is needed going forward to detect meaningful variation.

The future of microbial community profiling is looking to metagenomics, since the common use of 16S rRNA to track relative abundances at the genus level, or species level if fortunate, does not provide useful information for predicting most functions of interest. As discussed above in Section 3, metagenomics can allow the detection of novel genotypes, or evolution, within multi-species communities; this is an exciting step forward for investigating evolutionary processes and adaptation of complex natural communities, especially because metagenomics can include not only bacteria but also fungi, viruses, and other microbes. Metagenomic profiling of multi-species communities will move the field of evolutionary ecology forward by studying evolving, species-diverse communities over multiple generations. Comparatively, ecological studies of single time points would still benefit, but less from metagenomic sequencing, since it could still show gene compositions that link better to community functions of interest. A first step in testing the power of metagenomics is to initially test a controlled laboratory system of multiple genotypes of a single species, such as proposed in Section 4 addressing toxic cyanobacterial blooms. However, metagenomic compositions still can only indicate *predictions* on community metabolic functions (a community's "fundamental niche"), and not *actual* metabolism ("realised niche") (Malard and Guisan 2023). An additional step would be performing meta-transcriptomics to investigate actual gene transcription and functional activity of the community and its members.

A proposed alternative to 16S rRNA sequencing is to instead classify at the level of functional traits, genes, or metabolic guilds of interest (Burke *et al.* 2011; Goldford *et al.* 2018; Louca *et al.* 2018; Reynolds *et al.* 2023). Binning genes into broad functional families can improve interpreting community function from composition changes – composition now not being taxonomy, but instead groupings based on genes and their functions. The traditional species concept has long been proposed to not fit well for microbes (Rocha 2018; VanInsberghe *et al.* 2020), yet it is still the easiest and standard approach taken via 16S rRNA profiling (including in this thesis). Focusing on genes, or functions, as the unit of classification is an appropriate description of microbial community compositions to explore going forward, while recognising that this alternative view will create hurdles for translating findings to other study systems where taxonomical species will continue to be used. Yet, even beyond microbial community ecology, it has also been proposed to use functional traits instead for macroecological systems (McGill *et al.* 2006). Considering microbial communities to be composed of genes or functions of interest, is a drastic change in view, but its new perspective could provide useful insight for engineering communities with stable or predictable functions. A difficulty with narrowing to functional traits in microbes is that genes or regions of interest are ideally predicted before, but this is not a straightforward task depending on the system and environmental pressures used; for example, genes involved in resistance adaptation to one antibiotic class are a simpler case. But *a priori* knowledge may not be a necessity, as functional guilds have recently emerged from metagenomic data without predictions (Reynolds *et al.* 2023). Clearly, metagenomic sequencing helps overcome hurdles of predicting relevant functional genes prior, but its financial costs remain relatively high, and I argue that researchers still benefit by formulating such predictions. Effective use of metagenomics can be improved through well thought-out experimental design that minimises replicate populations and time points to gain the most insight from sequencing fewer samples. Nonetheless, the costs of metagenomic sequencing are quickly declining and I expect they will soon prove overall superior to the limitations of 16S rRNA sequencing approaches. It is exciting to see how greater resolution will advance our understanding of eco-evo feedbacks with evolving species populations within diverse communities, or the degree in which adaptive evolution is constrained to single microbial species or dispersed across an entire community (i.e., similar genes targeted across community or in only one species). However, while "-omics" approaches are very useful for investigating fundamental eco-evo theory and in large-scale research or production facilities, they are not a feasible future direction for many practical applied uses, such as in household or small-scale fermented food production.

6. Practical implications:

Mabisi upscaling:

Upscaling of Mabisi production in Zambia, mainly by women entrepreneurs, initiated the research directions undertaken by me and others in the Mabisi portion of the INREF Fermented Foods Project [Box 3]. For my research in-particular, the equivalence of back-slopping propagation and classic experimental evolution, inspired linking questions from ecology and evolutionary biology with real world applications for Mabisi production. Understanding a microbial community's ecology and evolution is influential when upscaling through developing starter cultures or controlling processing methods.

Starter culture approach:

Starter cultures are considered a "bottom up" approach to microbial community engineering and they are commonly applied in industrial level fermentations (Taskila 2017). However, using starter cultures poses challenges for rural home Mabisi production since it requires reliance on buying and storing a "seed" source of active microbes. Dependence on starter cultures can also limit opportunities for producers to differentiate their product since if only a few starter cultures are available, it consequently homogenises output across producers. Under current Zambian regulations that require pasteurised milk, a starter culture is initially needed to upscale production for official commercial sale – such as in Parmalat's "LactoMabisi" sold at Zambian supermarkets. Additionally, traditional processing of Mabisi is currently illegal since no formal standards exist – an official code of practice of processing methods could overcome this barrier to entrepreneurship.

Ecology driven processing approach:

The alternative to using starter cultures to achieve a regulated and reliable product is to take a "top down" approach of standardising the methods. Worded more ecologically, standardising methods means selecting environmental factors that shape the community composition and members' physiology over time to produce a desired final function (i.e., product characteristics relevant for consumer consumption). Ecological selection can shape existing natural bacterial communities into predictable composition and functionality, eliminating the need of using preselected strain bacteria and/or raw materials for a particular fermented food. While not termed "ecology driven", the careful selection of environmental conditions during fermentation to achieve a particular product has existed since as long as humans have fermented foods. A traditional yet current example is in natural wines, where starter cultures are not used (Galati et al. 2019; Alonso González and Parga-Dans 2020); winemakers instead closely track and may control variables such as temperature, sugars, acidity,

and oxygen to direct succession of necessary microbial community members while preventing unwanted spoilage during fermentation. If directing natural, spontaneous fermentation is acceptable for wine, the same approach could be applied to upscaling Mabisi production.

Mabisi food safety:
When milk is pasteurised, the natural microbiota is lost and fermentation cannot proceed without new microbial input, as is achieved with adding a starter culture. One argument for using starter cultures, whether with pasteurised or unpasteurised starting material, is that the addition of desirable microbes quickens the fermentation and offers competition against unwanted microorganisms to prevent spoilage (Vinicius De Melo Pereira et al. 2020). Still, the safety of traditional Mabisi (unpasteurised milk, no starter culture) with high resistance to food borne pathogens has been previously tested (Schoustra et al. 2022) and further supported in this thesis' Chapter 3. Regardless of food safety, the alteration of metabolic profiles in Chapter 3 by E. coli addition despite it not establishing, emphasises the need to still minimise any initial contamination. Potential unfavourable alteration of product characteristics can be reduced by promoting sterile and clean conditions during traditional processing.

Box 3: INREF project overview

This thesis was completed in the context of a Wageningen University transdisciplinary INREF project, entitled '*Traditional fermented foods to promote food and nutrition security in Africa; entrepreneurship, value chains, product development and microbial ecology in Zambia, Zimbabwe and Benin*' (acronym: FermFood). The project targeted (a) upgrading food and nutrition security in Africa by (b) ameliorating the quality and use of traditional fermented foods through (c) strengthening the connected local value chains and (d) fostering women's entrepreneurship. Since its start, the project has funded 8 PhD trajectories and 1 postdoc position. The project had three main specific objectives linked to the fourth objective of interdisciplinary integration (see Figure 4).

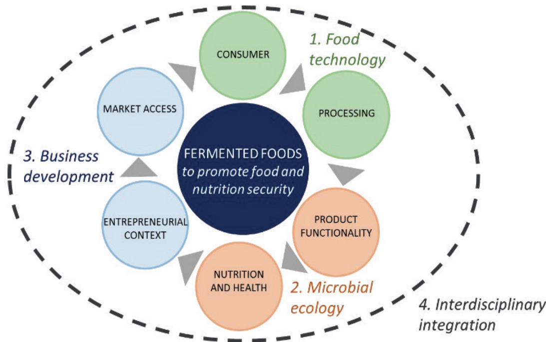


Figure 4: Overview of the research activities and linkages in the FermFood project.

Figure 5: Photo of local Mabisi producer and woman entrepreneur in Monze, Zambia. Photo by Robert Nhlane.

The central objectives involving different disciplines and interdisciplinary integration were as follows:

1 Upgrade traditional processing. Upgrading of artisanal household level production to small-to-medium enterprises at local farmer/producer cooperatives by designing the technological processes needed to upscale the production in current production areas (discipline: food technology);

2. Generalize and exploit functional properties. Define generic ecological properties that underlie successful fermented products to allow generalization and to maximize nutrition and health impact by improving the microbial composition that is key to product quality, targeting enhanced food and nutrition security (disciplines: microbial ecology, evolution and human nutrition);

3. Fostering women's entrepreneurship. Enhance conditions for locally initiated entrepreneurship as a source of income and opportunities for women in local farmer cooperatives and (household level) female producers, and develop value chains embedded in the domestic and local economic, legal and social-cultural environment (discipline: business development);

4. Interdisciplinary integration. Link insights and results for the three specific objectives.

Three traditional fermented foods were selected as representative examples, namely the dairy-based *mabisi* in Zambia, the cereal-based *mahewu* in Zimbabwe and the cereal-based *akpan* in Benin. Together these products cover different aspects, e.g., rural versus urban, and the current levels of standardization and contribution to diets. This range maximizes the relevance of our research and helps to expand findings to other traditional fermented foods and African food systems in general. In essence, scientific output will serve as a blueprint of the current reality and solutions to allow findings to be widely applied in Africa and beyond.

7. Concluding remarks:

There has been a recent call for more prediction focused microbial ecology studies, arguing that there is a worrying trend for descriptive data collection with post hoc conclusions. The growing ease of profiling microbial communities via genomic approaches - financial and practical - has created an overwhelming wealth of taxonomic data that lacks in mechanistic interpretations (Prosser 2020; Prosser and Martiny 2020). I believe the research presented in this thesis meets this call and serves as an example of using microbial communities to test meaningful hypotheses. The link between diversity and community function has long been a central focus of ecology yet the time scales of experiments have thus far been limited and thereby overlook the role of evolution and long-term ecological processes. My chapters were composed of experiments designed to assess thought-out objectives regarding the interconnected impacts that community composition, particularly diversity, and timescale have on community function. Chapter 2's serial dilution and propagation design aimed to answer questions about the impact of metabolic guild diversity on community function and composition over multiple generations. Chapter 3 used the same propagation regime as in Chapter 2 to test resistance and recovery of microbial community function and composition to repeated introduction of a "failed invader" species. Lastly, Chapter 4 manipulated species richness of a synthetic wine-yeast community to investigate resistance to invasion in relation to community richness versus compositional identity (i.e., keystone species). Furthermore, all chapters can be linked to real world applications of fermented foods, namely food spoilage and the role of community composition for aromatic properties.

The overarching objective of this PhD was to explore multi-species microbial communities in the style of classic microbial experimental experiments, while also investigating community diversity-function relationships over many generations. I found that the effect of diversity on function is largely caused by identity effects of keystone species in my model microbial systems of Mabisi (*G. candidum* in Chapter 2) and wine (*S. cerevisiae* in Chapter 4). Tracing changes in community composition and function over long-term ecological timescales in Chapters 2 and 3 highlighted the limits of 16S rRNA sequencing techniques for detecting meaningful community compositional changes; however, at least at the resolution achieved, results showed gradual, not sudden, changes in community composition and function over time. Introduction of novel species in Chapters 3 and 4 demonstrated the role of biotic and abiotic factors of both invader and resident community members for determining successful establishment of a novel species. Guided by core concepts of community ecology and population genetics, I applied both natural and synthetic community approaches in my thesis to demonstrate the suitability of fermented foods as model

microbial community systems. It additionally contributes to the thriving field of microbial community engineering by studying the fate of multi-species microbial communities over multiple generations – both in terms of species composition, function, and stability. In doing so, my research has sparked reflection about future outlooks for experimentally examining the underlying ecological and evolutionary processes of community composition-function relationships.

References:

Alonso González P, Parga-Dans E. Natural wine: do consumers know what it is, and how natural it really is? J Clean Prod 2020;251:119635.

Arias-Sánchez FI, Vessman B, Haym A *et al.* Artificial selection improves pollutant degradation by bacterial communities. *Nat Commun* 2024;**15**:7836.

Arias-Sánchez FI, Vessman B, Mitri S. Artificially selecting microbial communities: If we can breed dogs, why not microbiomes? PLoS Biol 2019;17:e3000356.

Blouin M, Karimi B, Mathieu J et al. Levels and limits in artificial selection of communities. Ecol Lett 2015;18:1040–8.

Bockelmann W, Willems KP, Neve H et al. Cultures for the ripening of smear cheeses. Int Dairy J 2005;15:719–32.

Bottari B, Santarelli M, Neviani E et al. Natural whey starter for Parmigiano Reggiano: culture-independent approach. J Appl Microbiol 2010;108:1676–84.

Broadbent JR, Cai H, Larsen RL et al. Genetic diversity in proteolytic enzymes and amino acid metabolism among Lactobacillus helveticus strains1. J Dairy Sci 2011;94:4313–28.

Buckling A, Kassen R, Bell G et al. Disturbance and diversity in experimental microcosms. Nature 2000;408:961–4.

Burford MA, Carey CC, Hamilton DP et al. Perspective: Advancing the research agenda for improving understanding of cyanobacteria in a future of global change. Harmful Algae 2020;91:101601.

Burke C, Steinberg P, Rusch D et al. Bacterial community assembly based on functional genes rather than species. Proc Natl Acad Sci 2011;108:14288–93.

Cairns J, Jokela R, Becks L et al. Repeatable ecological dynamics govern the response of experimental communities to antibiotic pulse perturbation. Nat Ecol Evol 2020;4:1385–94.

Cairns J, Jokela R, Hultman J et al. Construction and characterization of synthetic bacterial community for experimental ecology and evolution. Front Genet 2018;9:312.

Čaušević S, Dubey M, Morales M et al. Niche availability and competitive loss by facilitation control proliferation of bacterial strains intended for soil microbiome interventions. Nat Commun 2024;15:2557.

Chang C-Y, Bajić D, Vila JCC et al. Emergent coexistence in multispecies microbial communities. Science 2023;381:343–8.

Clements FE. Nature and Structure of the Climax. J Ecol 1936;24:252–84.

Dick GJ, Duhaime MB, Evans JT et al. The genetic and ecophysiological diversity of Microcystis. Environ Microbiol 2021;23:7278–313.

Doolittle FW, Inkpen AS. Processes and patterns of interaction as units of selection: An introduction to ITSNTS thinking. Proc Natl Acad Sci U S A 2018;115:4006–14.

Doulcier G, Lambert A, De Monte S et al. Eco-evolutionary dynamics of nested Darwinian populations and the emergence of community-level heredity. Elife 2020;9:e53433.

Galati A, Schifani G, Crescimanno M et al. "Natural wine" consumers and interest in label information: An analysis of willingness to pay in a new Italian wine market segment. J Clean Prod 2019;227:405–13.

Gänzle MG, Qiao N, Bechtner J. The quest for the perfect loaf of sourdough bread continues: Novel developments for selection of sourdough starter cultures. Int J Food Microbiol 2023;407:110421.

Goldford JE, Lu N, Bajić D et al. Emergent simplicity in microbial community assembly. Science 2018;361:469–74.

Goodnight CJ. Experimental studies of community evolution. I: the response to selection at the community level. Evolution 1990;44:1614–24.

Govaert L, Fronhofer EA, Lion S et al. Eco-evolutionary feedbacks—Theoretical models and perspectives. Funct Ecol 2019;33:13–30.
Gravel D, Bell T, Barbera C et al. Experimental niche evolution alters the strength of the diversity-productivity relationship. Nature 2011;469:89–94.
Groenenboom AE, van den Heuvel J, Zwaan BJ et al. Species dynamics in natural bacterial communities over multiple rounds of propagation. Evol Appl 2022;15:1766–75.
Hellweger FL, Martin RM, Eigemann F et al. Models predict planned phosphorus load reduction will make Lake Erie more toxic. Science 2022;376:1001–5.
Hendry AP. Eco-Evolutionary Dynamics., 2017.
Hendry AP. A critique for eco-evolutionary dynamics. Funct Ecol 2019;33:84–94.
Hirano R, Nishita I, Nakai R et al. Development of culture methods capable of culturing a wide range of predominant species of intestinal bacteria. Front Cell Infect Microbiol 2023;13:1056866.
Hughes AL. Adaptive Evolution of Genes and Genomes. Oxford University Press, USA, 1999.
Huisman J, Codd GA, Paerl HW et al. Cyanobacterial blooms. Nat Rev Microbiol 2018;16:471–83.
Jurburg SD, Blowes SA, Shade A et al. Synthesis of recovery patterns in microbial communities across environments. Microbiome 2024;12:79.
Kapinusova G, Lopez Marin MA, Uhlik O. Reaching unreachables: Obstacles and successes of microbial cultivation and their reasons. Front Microbiol 2023;14:1089630.
Kassen R. The experimental evolution of specialists, generalists, and the maintenance of diversity. J Evol Biol 2002;15:173–90.
Kilham SS, Kreeger DA, Lynn SG et al. COMBO: a defined freshwater culture medium for algae and zooplankton. Hydrobiologia 1998;377:147–59.
Kinnison MT, Hairston NG, Hendry AP. Cryptic eco-evolutionary dynamics. Ann N Y Acad Sci 2015;1360:120–44.
Kurmayer R, Deng L, Entfellner E. Role of toxic and bioactive secondary metabolites in colonization and bloom formation by filamentous cyanobacteria Planktothrix. Harmful Algae 2016;54:69–86.
Leale A, Auxier B, Smid EJ et al. Influence of metabolic guilds on a temporal scale in an experimental fermented food derived microbial community. FEMS Microbiol Ecol 2023;99:fiad112.
Leale AM, Marquez FR, Zwaan B et al. Shifts and rebound in microbial community function following repeated introduction of a novel species. 2024:2024.04.05.588252.
Liautaud K, van Nes EH, Barbier M et al. Superorganisms or loose collections of species? A unifying theory of community patterns along environmental gradients. Ecol Lett 2019;22:1243–52.
Louca S, Jacques SMS, Pires APF et al. High taxonomic variability despite stable functional structure across microbial communities. Nat Ecol Evol 2016;1:1–12.
Louca S, Polz MF, Mazel F et al. Function and functional redundancy in microbial systems. Nat Ecol Evol 2018;2:936–43.
Maddamsetti R, Lenski RE, Barrick JE. Adaptation, Clonal Interference, and Frequency-Dependent Interactions in a Long-Term Evolution Experiment with Escherichia coli. Genetics 2015;200:619–31.
Madi N, Vos M, Murall CL et al. Does diversity beget diversity in microbiomes? Weigel D, Kemen E, Wolfe BE (eds.). eLife 2020;9:e58999.
Malard LA, Guisan A. Into the microbial niche. Trends Ecol Evol 2023;38:936–45.
Maleke M, Doorsamy W, Abrahams AM et al. Influence of Fermentation Conditions (Temperature and Time) on the Physicochemical Properties and Bacteria Microbiota of Amasi. Fermentation 2022;8:57.
Malusà E, Pinzari F, Canfora L. Efficacy of Biofertilizers: Challenges to Improve Crop Production. In: Singh DP, Singh HB, Prabha R (eds.). Microbial Inoculants in Sustainable Agricultural Productivity: Vol. 2: Functional Applications. New Delhi: Springer India, 2016, 17–40.

McGill BJ, Enquist BJ, Weiher E et al. Rebuilding community ecology from functional traits. Trends Ecol Evol 2006;21:178–85.

McInnes RS, McCallum GE, Lamberte LE et al. Horizontal transfer of antibiotic resistance genes in the human gut microbiome. Curr Opin Microbiol 2020;53:35–43.

Merz E, Saberski E, Gilarranz LJ et al. Disruption of ecological networks in lakes by climate change and nutrient fluctuations. Nat Clim Change 2023;13:389–96.

Meyer JR, Schoustra SE, Lachapelle J et al. Overshooting dynamics in a model adaptive radiation. Proc R Soc B Biol Sci 2010;278:392–8.

Moeller AH, Li Y, Mpoudi Ngole E et al. Rapid changes in the gut microbiome during human evolution. Proc Natl Acad Sci 2014;111:16431–5.

Moonga HB, Schoustra SE, Linnemann AR et al. The art of mabisi production: A traditional fermented milk. PloS One 2019;14:e0213541.

Moonga HB, Schoustra SE, Linnemann AR et al. Influence of fermentation temperature on microbial community composition and physicochemical properties of mabisi, a traditionally fermented milk. LWT 2021;136:110350.

Moonga HB, Schoustra SE, Van den Heuvel J et al. Composition and diversity of natural bacterial communities in mabisi, a traditionally fermented milk. Front Microbiol 2020;11:1816.

van Moorsel SJ, Hahl T, Petchey OL et al. Co-occurrence history increases ecosystem stability and resilience in experimental plant communities. Ecology 2021;102:e03205.

Mounier J, Coton M, Irlinger F et al. Chapter 38 - Smear-Ripened Cheeses. In: McSweeney PLH, Fox PF, Cotter PD, et al. (eds.). Cheese (Fourth Edition). San Diego: Academic Press, 2017, 955–96.

O'Brien S, Culbert CT, Barraclough TG. Community composition drives siderophore dynamics in multispecies bacterial communities. BMC Ecol Evol 2023;23:45.

Okasha S. Evolution and the Levels of Selection. Oxford, New York: Oxford University Press, 2008.

Otto FEL. Attribution of Extreme Events to Climate Change. Annu Rev Environ Resour 2023;48:813–28.

Pepper IL, Gentry TJ. Chapter 2 - Microorganisms Found in the Environment. In: Pepper IL, Gerba CP, Gentry TJ (eds.). Environmental Microbiology (Third Edition). San Diego: Academic Press, 2015, 9–36.

Piccardi P, Alberti G, Alexander JM et al. Microbial invasion of a toxic medium is facilitated by a resident community but inhibited as the community co-evolves. ISME J 2022;16:2644–52.

Piccardi P, Ulrich E, Garcia-Garcerà M et al. The evolution of reduced facilitation in a four-species bacterial community. Evol Lett 2024:qrae036.

Piccardi P, Vessman B, Mitri S. Toxicity drives facilitation between 4 bacterial species. Proc Natl Acad Sci 2019;116:15979–84.

Posch T, Köster O, Salcher MM et al. Harmful filamentous cyanobacteria favoured by reduced water turnover with lake warming. Nat Clim Change 2012;2:809–13.

Pourcelot E, Conacher C, Marlin T et al. Comparing the hierarchy of inter- and intra-species interactions with population dynamics of wine yeast cocultures. FEMS Yeast Res 2023;23:foad039.

Pourcelot E, Vigna A, Marlin T et al. Design of a new model yeast consortium for ecological studies of enological fermentation. 2024:2024.05.06.592697.

Prosser JI. Putting science back into microbial ecology: a question of approach. Philos Trans R Soc B Biol Sci 2020;375:20190240.

Prosser JI, Martiny JBH. Conceptual challenges in microbial community ecology. Philos Trans R Soc B Biol Sci 2020;375:20190241.

Rainey PB, Travisano M. Adaptive radiation in a heterogeneous environment. Nature 1998;394:69–72.

Raynaud T, Devers M, Spor A et al. Effect of the reproduction method in an artificial selection experiment at the community level. Front Ecol Evol 2019;7:416.

Reynolds R, Hyun S, Tully B et al. Identification of microbial metabolic functional guilds from large genomic datasets. Front Microbiol 2023;14.
Rigosi A, Carey CC, Ibelings BW et al. The interaction between climate warming and eutrophication to promote cyanobacteria is dependent on trophic state and varies among taxa. Limnol Oceanogr 2014;59:99–114.
van Rijswijck IMH, Wolkers – Rooijackers JCM, Abee T et al. Performance of non-conventional yeasts in co-culture with brewers' yeast for steering ethanol and aroma production. Microb Biotechnol 2017;10:1591–602.
Rivett DW, Scheuerl T, Culbert CT et al. Resource-dependent attenuation of species interactions during bacterial succession. ISME J 2016;10:2259–68.
Rocha EPC. Neutral Theory, Microbial Practice: Challenges in Bacterial Population Genetics. Mol Biol Evol 2018;35:1338–47.
Rosenberg E. Eukaryotic Microorganisms are Part of Holobionts. In: Rosenberg E (ed.). Microbiomes: Current Knowledge and Unanswered Questions. Cham: Springer International Publishing, 2021, 195–229.
Ruiz J, de Celis M, Diaz-Colunga J et al. Predictability of the community-function landscape in wine yeast ecosystems. Mol Syst Biol 2023;19:e11613.
Sanchez A, Bajic D, Diaz-Colunga J et al. The community-function landscape of microbial consortia. Cell Syst 2023;14:122–34.
Scheuerl T, Hopkins M, Nowell RW et al. Bacterial adaptation is constrained in complex communities. Nat Commun 2020;11:1–8.
Schoustra S, Kasase C, Toarta C et al. Microbial Community Structure of Three Traditional Zambian Fermented Products: Mabisi, Chibwantu and Munkoyo. PLOS ONE 2013;8:e63948.
Schoustra S, van der Zon C, Groenenboom A et al. Microbiological safety of traditionally processed fermented foods based on raw milk, the case of Mabisi from Zambia. LWT 2022;171:113997.
Schoustra SE, Debets AJM, Slakhorst M et al. Reducing the cost of resistance; experimental evolution in the filamentous fungus Aspergillus nidulans. J Evol Biol 2006;19:1115–27.
Sheth RU, Cabral V, Chen SP et al. Manipulating Bacterial Communities by in situ Microbiome Engineering. Trends Genet 2016;32:189–200.
Swenson W, Arendt J, Wilson DS. Artificial selection of microbial ecosystems for 3-chloroaniline biodegradation. Environ Microbiol 2000;2:564–71.
Swenson W, Wilson DS, Elias R. Artificial ecosystem selection. Proc Natl Acad Sci 2000;97:9110–4.
Tamang NTAJP. Functionality and Therapeutic Values of Fermented Foods. Health Benefits of Fermented Foods and Beverages. CRC Press, 2015.
Taskila S. Industrial production of starter cultures. Starter Cultures in Food Production. John Wiley & Sons, Ltd, 2017, 79–100.
Thompson LR, Sanders JG, McDonald D et al. A communal catalogue reveals Earth's multiscale microbial diversity. Nature 2017;551:457–63.
Tully BJ, Wheat CG, Glazer BT et al. A dynamic microbial community with high functional redundancy inhabits the cold, oxic subseafloor aquifer. ISME J 2018;12:1–16.
Turner MG, Calder WJ, Cumming GS et al. Climate change, ecosystems and abrupt change: science priorities. Philos Trans R Soc B Biol Sci 2020;375:20190105.
Turner PE, Souza V, Lenski RE. Tests of Ecological Mechanisms Promoting the Stable Coexistence of Two Bacterial Genotypes. Ecology 1996;77:2119–29.
Van de Waal DB, Verspagen JMH, Finke JF et al. Reversal in competitive dominance of a toxic versus non-toxic cyanobacterium in response to rising CO2. ISME J 2011;5:1438–50.

VanInsberghe D, Arevalo P, Chien D et al. How can microbial population genomics inform community ecology? Philos Trans R Soc B Biol Sci 2020;375:20190253.
Vellend M. The theory of ecological communities. The Theory of Ecological Communities. Princeton University Press, 2016.
Vinicius De Melo Pereira G, De Carvalho Neto DP, Junqueira ACDO et al. A Review of Selection Criteria for Starter Culture Development in the Food Fermentation Industry. Food Rev Int 2020;36:135–67.
Weiss AS, Burrichter AG, Durai Raj AC et al. In vitro interaction network of a synthetic gut bacterial community. ISME J 2022;16:1095–109.
Weiss AS, Niedermeier LS, von Strempel A et al. Nutritional and host environments determine community ecology and keystone species in a synthetic gut bacterial community. Nat Commun 2023;14:4780.
Wells ML, Trainer VL, Smayda TJ et al. Harmful algal blooms and climate change: Learning from the past and present to forecast the future. Harmful Algae 2015;49:68–93.
Williams HTP, Lenton TM. Artificial selection of simulated microbial ecosystems. Proc Natl Acad Sci U S A 2007, DOI: 10.1073/pnas.0610038104.
Wilson DS, Sober E. Reviving the superorganism. J Theor Biol 1989;136:337–56.
Wright RJ, Gibson MI, Christie-Oleza JA. Understanding microbial community dynamics to improve optimal microbiome selection. Microbiome 2019;7:85.
Xie L, Shou W. Steering ecological-evolutionary dynamics to improve artificial selection of microbial communities. Nat Commun 2021;12:6799.

Appendix

Summary

Acknowledgements

About the author

List of publications

PE&RC Training and Education Statement

Summary

The relationship beween the composition and function of ecological communties has long been a central topic in ecology, yet there are components of community composition-function relationships still requiring further consideration. Notably, the time scales of experiments have thus far been limited and thereby overlook the role of evolution and long-term ecological processes. When a diverse community experiences a new and consistent environment, how does it respond – both in community composition and function? Investigating community composition-function relationships on longer time scales has been advocated and such research can be facilitated by microbial communities whose rapid generational turnover creates a tight dynamic between processes of ecology and evolution. Understanding how changes across hierarchical levels - from genes, to species, to communities - influence one another is undoubtedly complicated but merits initial steps to elucidate, which can be supported by a suitable framework and model system.

The general inspiration for this thesis was to perform experimental evolution styled propagation experiments using diverse microbial communities of multiple species. By doing so, I wanted to consider species in a community in a paralleled view to genetic variants, or genotypes, within a single species. Inspired by concepts of traditional community ecology and population genetics, I applied multi-species microbial communities of fermented foods to investigate community composition-function relationships. A traditionally fermented milk beverage from Zambia, Mabisi, and wine yeast communities were used as model systems. More specifically, the overarching objective of this PhD was to gain insight on the interconnected impacts that community composition, particularly diversity, and timescale have on community function. Focus was made on the diversity component of community composition (Chapters 2 and 4), community function was measured as metabolism (Chapters 2 and 3), and both were studied in relation to stability, or resistance, to an introduced species (Chapters 3 and 4). Serial dilution and propagation of Mabisi microbial communities in Chapter 2 aimed to answer questions about the impact of metabolic guild diversity on community function and composition over multiple generations. Chapter 3 used the same propagation regime of Mabisi to test resistance and recovery of microbial community function and composition to repeated introduction of the "failed invader" species *Escherichia coli*. Lastly, species richness of a synthetic wine-yeast community was manipulated in Chapter 4 to investigate resistance to invasion in relation to community richness versus compositional identity (i.e., keystone species).

Diversity: The effect of diversity on function was evidenced to be underlaid by identity, or inclusion, of certain species in both Chapters 3 and 4. Yeast species *Geotrichum*

candidum played a main role in division of metabolic profiles of Mabisi communities in Chapter 2, whereas in Chapter 4 *Saccharomyces cerevisiae* was a keystone species for resistance again spoilage microorganisms in wine yeast communities. **Time:** Tracing changes in community composition and function over long-term ecological timescales was sought in Chapters 2 and 3, which both demonstrated gradual, not sudden, changes in community composition and function. Grouping of metabolic profiles based on treatment did not shift overtime in Chapter 2 nor Chapter 3, as any divisions in function observed at initial time points remained throughout propagation. Recovery in metabolic profiles did occur when *E. coli* introduction was stopped in Chapter 3. **Invasion:** The system dependency of invasions – of biotic and abiotic factors of both invader and resident community - was supported in Chapters 3 and 4, as either failed or successful establishment of an introduced species were observed. In Chapter 3, the failed invasion of *E. coli* in Mabisi despite repeated high propagule pressure of inoculation was likely due to high acidity (abiotic stress) created by the resident Mabisi community. Whereas in Chapter 4, the establishment of *Brettanomyces bruxellensis* or *Lactiplantibacillus plantarum* in wine yeast communities was dependent on a combination of biotic factors of both the resident community composition and invader identity (*L. plantarum* was transient in all communities, whereas *B. bruxellensis* persisted only when *S. cerevisiae* was absent).

In Chapters 1 and 5 I conceptualise how microbial communities can adapt to an altered environment both via ecological and evolutionary processes, that is, by adjustment of relative species abundances through the process of species sorting and by spontaneous mutations spreading in individual species. I outline in Chapter 5 possible synthetic microbial community model systems for detecting and interpreting genetic changes within a species amongst a multi-species community. I image a “Muller-like plot” with multiple species, in which a different colour is a particular species abundance, and different shades of that colour are genotypes within that species. Chapter 5 furthermore considers the potential for community level selection and artificial selection of microbial communities, and proposes its established application already in fermented food production.

Overall, this thesis demonstrates the suitability of microbial communities of fermented foods as model experimental systems to investigate three processes of community assembly – timescale, diversity, and invasion. Furthermore, all chapters can be linked to real world applications of fermented foods, namely food spoilage and the role of community composition for aromatic properties.

Acknowledgements

Yes, I recognise this acknowledgment section is fairly lengthy, but I view it as a fortunate outcome of the many amazing people I have crossed paths with.

Firstly, a big thank you to my daily supervisor Sijmen. It was quite overwhelming for me back in 2020 to move across the world for four years, to a country I had never visited, with a supervisor I never met, and in the middle of a pandemic. But it was worth it, and your supervision made a substantial influence on this. I know I wasn't always the easiest with my endless moments of worry and doubt, but you are incredibly understanding, and your patient encouragement allowed me to keep on. The freedom and trust you provided to direct my own research allowed me to feel proud of this thesis and build confidence as an independent scientist. I have learned a lot from your research proposal writing skills Sijmen, and the critical value of not only seeking interdisciplinary science but truly accomplishing it too. And thank you to my co-promotors Bas and Eddy. Bas, the Laboratory of Genetics is very lucky to have you as our department chair who pushes the status-quo of academia – whether it be with bright butterfly DocMartens and corduroy overalls or questioning the politics of university systems. Eddy, I always enjoyed having your smiling, friendly, up-beat personality at meetings, and your expertise of microbe fermentation metabolism was invaluable.

Completing this thesis would not have been possible without my so-called "unofficial supervisors" Francisca and Ben. How to describe Francisca – usually I say something like "she is my lab Mom/lab best friend/lab guardian angel/Santa Francisca de los PCRs, and so much more". The combination of your kindness and intelligence has saved me and many others through difficult times in, or out of the lab. I will continue to look up to you a great amount wherever life leads me next. And Ben, my research would not be what it is without your insight and generosity. I hope I have learned from your critical eye, which I will carry forward to continue improving my research. Thank you for taking the time and effort to help me (and nearly everyone in GEN) with bioinformatic pipelines, debugging an R code for a pretty plot, paper drafts, applications, supervision advice, laboratory techniques, and on and on.

My weekly meetings, social lunches, potlucks, and workshops with the INREF Fermented Foods group (Carole, Blandine, Murambiwa, Mubonda, Shepherd, Thelma, Sakile, Taonga, Robert, Ardinesh, Mundia, Clara) were very well enjoyed, so thank you all for your good humour, excellent cooking, Nshima-kneeding lessons, and support while we made our way through this PhD journey. Thank you, Anna, for your kindness to always ensure myself and the INREF crew settled well into

Wageningen. We all especially appreciated your insight and time to help us understand our bioinformatic analyses.

Merci beaucoup à mes superviseurs à Montpellier – Delphine et Thibault. Without your support to host and supervise me, I would not have been fortunate enough to have my research and life experience at L'Institut Agro. Thank you for your positive attitude and encouragement during my pessimistic panics. Through you I gained the value of optimism and learned to rely on others to complete big experiments, which has made me a better scientist. Merci aussi à Eléonore – I admired your patience to help me throughout, despite finishing your own thesis at the same time. I am greatly impressed by your knowledge of scientific literature and dedication to your own and others' research – you are like a living human bibliography of wine yeast ecology! Et bien sûr un grand merci à Stéphane! Sans toi les manipulations ne seraient pas possibles. Si j'ai bien calculé, nous avons préparé et coulé environ 147L de milieu d'agar, soit 4900 boîtes de petri! De plus, c'est avec toi que j'ai le plus pratiqué mon français, donc, j'apprécie ta patience à essayer de comprendre mon « Franglais ».

Grazie mille Bea for bringing out the dancing, girly, silly, sparkly, scientist in me when we met in Montpellier. As you have said "I am so thankful the world let us meet, but I'm angry it made us live so far apart". You are an incredible friend whose support helped me complete this thesis, while learning some Italian along the way. I can't wait to soon sip cappuccinos with you after a night of dancing (German techno for me, Latin/reggaeton for you)!

Louise, your infectious laugh and ability to know when "something's up" really helped me get through. Your patience to always listen and remind everyone to be easier on oneself is a special character that I have been lucky to have nearby. Thank you for your friendship filled with lunch time wanders, many silly giggles, and shared appreciation for the outdoors.

Big hugs Daniel for being a super supportive friend and showing me one of the only okay-ish lakes in the Netherlands. You were always there to listen through my many teary-eyed highs and lows, fix my bike, water my plants, or pick up parcels that I forgot I ordered just before a vacation. And your heartfelt Latin hugs were appreciated – even if I may not have expressed it back well (sorry!). So, is it finally time to uncork that bottle of "Pixalu" (Pic Saint-Loup) wine?!

And thanks to some new faces in GEN that brought an amazing vibe to the office this past year! Sitting next to you was a non-stop stand-up comedy show Bo and I admire your "work hard, play hard" attitude. You made the dreaded thesis writing phase full

of laughs and terribly trashy, catchy pop music, which I needed – so thank you so much! It has been such a pleasure Maggie-Anne to have you join the department with your top-notch break time conversations on a diversity of subjects (honestly, where do you find all this information) and shared enthusiasm for the cinema. Spyros – I hope that some of your upbeat, positive energy has worn off on me this year with each of your "hello beautiful people" greetings.

Thank you, Duur, for your kind approach and attitude to expressing your interest in others' research and posing critical questions. Your feedback has made my thesis stronger. Eveline, I've greatly appreciated your advice and support as I navigated through PhD ups and downs, with the occasional email in French to keep up our language skills! Gerlinde, I am always inspired by your uplifting energy and enthusiasm for others' research, especially your natural curiosity to ask both the simple and complex questions on your mind. Thanks Judith, for your help running the GC-MS and with all its data analysis after but also importantly for your wonderfully bright energy anytime I saw you at Axis.

I learned a lot of fun (and fundamental!) information from the Evolution Journal Club, so thanks to all the regulars and occasional drop-ins. Thanks Nicky for coordinating it and your always curious questions. Thank you Hilje for bringing a dedicated enthusiasm for fundamental eco-evo concepts (especially for the microbial world – but I guess I'm a bit biased) to journal club, and your useful feedback on many email drafts of mine. Thank you also to the past and present Wageningen Ecology and Evolution Seminars (WEES) members, with whom I am happy to have organised some great invited speakers over the years. A shout out to Lydia, who WEES is very lucky to have around - the world needs more intelligent, authentically nice and caring, organised, adventurous, cool, etc., people like you.

I was lucky to work with excellent students - Stijn, Maria, Merle, and Tiro – thanks for making supervising easy and you're sure to take exciting next steps. Thank you to Marjan for never being bothered by my never-ending MyProjects timesheet questions and corrections. The lab would be a disaster without the lab technicians, so thanks for keeping me from burning the place down and knowing where the extra secret storage for every supply is. Thanks Dave and Francesco at Eawag for your enthusiastic support and feedback on potential postdoc projects, which sparked excitement in me about future research paths.

Then a proost to the longstanding Vrijmibo gang, past and present (René, Tom, Sofia, Alex, Suzette, Murambiwa, Mubonda, Jitpanu, Roel, Jelle (Louise's), Jelle (René's), Lisa, Hylke, Francesco, Alan, Martin, Mirko, Mariana, Jeroen, and many more)!

Surviving a PhD would not have been possible without your friendship and ridiculous conversations. Even if you won't let me paint your tortoises René, you deserve a thank you for your appreciated vrijmibo encouragement and bringing essential humour to the GEN group. GEN and especially its students are lucky to keep your dedicated and caring personality around. Thanks, Suzette, for the wonderful hours climbing with you this last year. Your great company and experienced climbing tips helped stop my busy mind both while hanging from a wall or chatting in-between. I enjoyed our chats Mariska about travel, surfing, yoga, and fermentation projects, so thank you for that, but also your generosity and time to share your MinIon sequencing knowledge.

My brief travel break in 2023 was essential to better enjoying my PhD experience. Merci à Julien du Château Brandeau. My stay at your vineyard during harvest season was unforgettable - an outcome of your fantastic energy and playful joy for winemaking. And to the fellow fun and warm-hearted "Chatty B 2023" grape pickers - I was very fortunate to spend time with you all, so thanks for the good vibes, laughter, wild dance parties, vineyard sing-alongs, jeux-de-boules matches, and more. Thanks to Deirde and family at Cascina Val Liberata for your hospitality and work-relax balance during your vineyard harvest.

There have also been many other faces along the way: To my friend since kindergarten, Michaela– our friendship is a special one that I am incredibly grateful keeps continuing. You always have the right thing to say and are a wonderful listener – from near and far. Jenny – settling into the Netherlands was not always easy for me, so I greatly appreciate your kindness and our many long walks while I made my way. I admire your openness and perseverance to experience new hobbies, people, places, and interests, which I hope has rubbed off on me. Thank you, Tobi, for helping open-up life in Wageningen after COVID and reminding me to pause, laugh, enjoy the present, and stop thinking too far ahead. Fernando – your endless involvement in Wageningen initiatives has contributed to many memorable aspects of life here. Merci Cloé for always caring about how things were going, our shared appreciation of both science and art, and of course your beautiful thesis cover design. Santé à mes amis de Montpellier! - Megan, Tracy, Irene, Sylvia, Lily, et Barbara! You are all lovely ladies whose adventurous ambition I admire, and your friendship was meaningful to me while in Montpellier and I look forward to continuing crossing paths with you all.

And there is *zero* possibility I could have persevered through this PhD without the support of my family back "home-home" in northern Ontario, Canada. Thank you Mom and Dad for putting up with my texts in the middle of your night (sorry, I never did remember how annoying the time difference was for you), the far too many FaceTime calls when I said I just wanted to catch up but was actually mid-meltdown,

always reading my important documents for typos (even if you always deleted my Oxford commas), and filling your suitcase with maple syrup and Kraft Dinner mac 'n cheese for me when visiting the Netherlands. To my sister Carina – thanks for being a strong, beautiful, adventurous, intelligent younger sister who I can look up to and always trust for honest advice. Also, thank you Carina and Nick for always knowing when I needed goofy photos of Lugh to cheer me up. I have missed you all a lot, but I am thankful that you're just a video call away and am always so excited to see you next.

Echt, heel erg bedankt iedereen!

In memory of my friend and a ray of sunshine, Jackie (1992-2024). She will always be remembered for the warm kindness and brightness she shared with everyone she met.

About the author

Alanna Molly Leale was born on 4 June 1993 in Thunder Bay, Ontario, Canada and was raised in the town of Nipigon in northern Ontario at the top of Lake Superior (which for a fun fact, is nearly twice the size of the Netherlands by surface area).

Alanna's upbringing in rural northern Ontario, with parents who were a forester and a fisheries biologist, instilled an early appreciation for nature, including big waters. She has always been eager to experience new places, people, and challenges, which is reflected in her diverse life path. She left her small hometown to do an HBSc at University of Toronto where she was immediately enthusiastic about her introductory ecology and evolutionary biology course, and the subject has motivated her since. She was granted several research opportunities during her bachelor's degree, most notably anthropology dietary isotope analysis in Cape Town, South Africa and managing a *Drosophila melanogaster* laboratory.

Her particular interest in evolutionary medicine led her to apply for and receive the Canadian government NSERC CG-SM graduate scholarship with Dr. Rees Kassen at the University of Ottawa. Her research experimentally tested how varying drug availability in either time or space impacted antibiotic resistance in bacterial populations. After her MSc, Alanna stepped away from academia to reconsider career motivations. During this period, she worked as an English assistant for two school years at French schools in the francophone province of Québec, teaching at both the primary and college level.

She started her PhD at Wageningen University & Research (Netherlands) in August 2020 with Dr. Sijmen Schoustra in a large interdisciplinary project titled "Traditional fermented foods to promote food and nutrition security in Africa". She enjoyed the multidiscipline collaboration required for the project and the diverse insight she gained from her peers. Fundamental principles from community ecology and evolutionary biology inspired her to design experimental evolution styled propagation experiments using microbial communities in fermented foods. Her PhD research

aimed to investigate the interconnected impacts that community composition, particularly diversity, and timescale have on community function. But never staying still in one place for too long, she took the initiative during her PhD to complete a five-month research visit at L'Institut Agro, Montpellier with Dr. Delphine Sicard and Dr. Thibault Nidelet studying the ecology and dynamics of wine yeasts during fermentation. She then left for a two-month leave of absence to travel France and Italy to help organic vineyards with their harvest season (Château Brandeau, FR and Cascina Val Liberata, IT).

Extracurricular academic activities had an influential role throughout her university studies. As a PhD candidate she was a member and chair of the Wageningen Ecology and Evolution Seminars (WEES) and helped organise the 2022 Netherlands Society for Evolutionary Biology (NLSEB) PhD-Postdoc Day. She also was involved in starting a bi-weekly Evolution Journal Club, which is still running.

Alanna has always had an artistic side with an interest in painting, printmaking, film, and music. Some of her work has been presented in Wageningen's arts and culture magazine "Uitwaaien". She keeps busy with outdoor or active pursuits, including jogging, hiking, open water swimming, yoga, and more recently climbing. She is also inspired by her microbiological studies to experiment with fermentation making kombucha, kimchi, fermented vegetables, and sourdough. You could say she brings her work home, as she loves good cheese, chocolate, wine, and beer with friends.

Alanna's varied work experiences make her excited to see where she goes next. A career in research is on the horizon for now (fingers crossed for lakes and mountains nearby), but who knows, maybe she'll still become a cheesemaker in the mountains!

List of publications

Leale, A.M., Pourcelot, E., Guezenec, S., Sicard, D., & Nidelet, T. 2024. *S. cerevisiae* serves as keystone species against spoilage in experimental synthetic wine yeast communities. *bioRxiv.* https://doi.org/10.1101/2024.07.04.602080

Leale, A.M., Reyes-Marquez, F., Smid, E. J., Zwaan, B., & Schoustra, S. 2024. Shifts and rebound in microbial community function following repeated introduction of a novel species. *bioRxiv.* https://doi.org/10.1101/2024.04.05.588252

Leale, A.M., Auxier, B., Smid, E. J., & Schoustra, S. 2023. Influence of metabolic guilds on a temporal scale in an experimental fermented food derived microbial community. *FEMS Microbiology Ecology*, *99*(10), fiad112. https://doi.org/10.1093/femsec/fiad112

Leale, A.M., & Kassen, R. 2018. The emergence, maintenance and demise of diversity in a spatially variable antibiotic regime. *Evolution Letters* 2:134-143. https://doi.org/10.1002/evl3.43

PE&RC Training and Education Statement

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 30 ECTS (= 20 weeks of activities)

Review/project proposal (4.5 ECTS)
- PhD proposal introduction & seminar presentation

Writing of project proposal (5 ECTS)
- PhD proposal: The connected effects of diversity and evolution on microbial community function in Mabisi (2021)
- PE&RC call for institutional collaboration: Track and trace of ecology and evolution during adaptation in multispecies communities (2022)
- NWO XS Grant: Track and trace of ecology and evolution during adaptation in multispecies communities (2022)
- Eawag Postdoctoral Fellowship: How do climate change-induced blooming conditions impact toxicity of cyanobacterial populations? (2024)

Post-graduate courses (8 ECTS)
- Resilience of Living Systems, PE&RC (2021)
- Principles of ecological and evolutionary genomics, PE&RC (2021)
- Microbial timeseries analysis, online, Leuven (2021)
- Tidy data transformation with ggplot2, PE&RC (2020)
- 16s sequencing, online, ZIEL Institute (2021)
- Introduction to LaTex, WGS (2022)
- Intermediate programming in R, PE&RC (2024)

Laboratory training and working visits (4.5 ECTS)
- Lab visit and seminar presentation, RUG (2022)
- Research visit and project, INRAE/Institut Agro, Montpellier, France (2023)

Competence, skills and career-oriented activities (2.7 ECTS)
- Intercultural Communication, online, PE&RC (2021)
- French 4, course level B1, online, WUR Into Languages (2022)
- Knowledge utilisation in evolutionary biology, NLSEB (2022)
- Career Orientation, WGS (2024)
- NWO Insight Out - Academic Mental Health afternoon, online, NWO (2023)
- NWO Insight Out - live event @ Van der Valk Utrecht, NWO (2023)

Scientific Integrity/Ethics in science activities (2.1 ECTS)
- Scientific Integrity, WGS (2020)
- Philosophy and Ethics of Food Science and Technology, VLAG (2021)

PE&RC Annual meetings, seminars and PE&RC weekend/retreat (1.5 ECTS)
- PE&RC First Years Day, online (2020)
- PE&RC Afternoon, online (2020)
- PE&RC Day, online (2021)
- PE&RC INREF Project reflection meeting, online (2021)
- PE&RC Fermentation workshop (2022)
- PE&RC Day, Nijmegen (2022)

National scientific meetings, local seminars, and discussion groups (14.1 ECTS)
- Evolution Journal Club, GEN-WUR (2021-2024)
- WEES organising committee, WUR (2021-2024)
- NLSEB Junior Board member (2021-2022)
- INREF symposia, WUR (2020-2024)
- WUR Seminar: Diversity of crops, food and culture (2023)
- Journée Scientifique: Vigne – Vin, Institut Agro, Montpellier, France (2023)
- Journée Scientifique: Eau et viticulture, Institut Agro, Montpellier, France (2023)

International symposia, workshops and conferences (11 ECTS)
- EMPSEB 26, online (2021)
- NLSEB, Ede, the Netherlands (2022)
- MEE Hubs, Lausanne, Switzerland (2024)
- NLSEB, Ede, the Netherlands (2024)
- EMPSEB 28, Puchberg am Schneeberg, Austria (2024)
- Guarda workshop on evolutionary biology, Guarda, Switzerland (2024)

Societally relevant exposure (0.15 ECTS)
- Illustrator – The Jester newspaper, WUR (2023-2024)

Lecturing/supervision of practicals/tutorials (3.9 ECTS)
- GATC (2023)

BSc/MSc thesis supervision (6.5 ECTS)
- Research topic 1: BSc - Stijn Hanssen
- Research topic 2: Erasmus exchange - Maria Warwony
- Research topic 3: MSc - Tiro van Gestelen

The research described in this thesis was financially supported by the Wageningen University Transdisciplinary Research and Education Fund (INREF).

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover art by Alanna Leale
Cover layout by Cloé Villard
Lay-out and printing by ProefschriftMaken.nl on FSC-certified paper