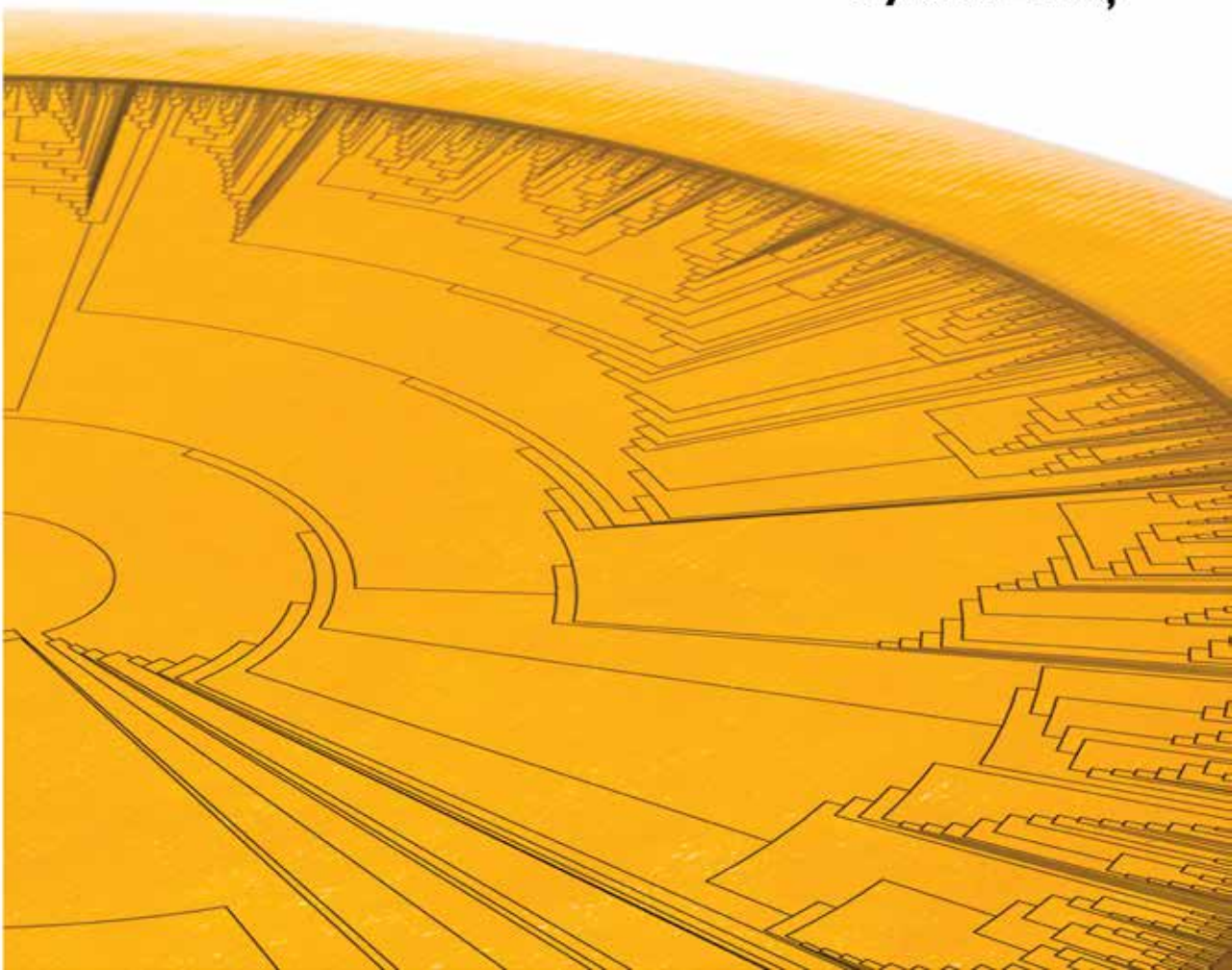
Community Dynamics of Complex Starter Cultures for Gouda-type Cheeses and its Functional Consequences

Oylum Erkuş



Community Dynamics of Complex Starter Cultures for Gouda-type Cheeses and its Functional Consequences

Oylum Erkuş

Thesis committee

Promotor

Prof. Dr. M. Kleerebezem
Personal chair in the Host Microbe Interactomics Group
Wageningen University

Co-promotor

Dr Eddy. J. Smid
Associate professor, Laboratory of Microbiology
Wageningen University

Other members

Prof. Dr. A.C.M. van Hooijdonk, Professor of Dairy Science, Wageningen University
Prof. Dr. P.R. Jensen, Professor of Systems Biotechnology, Technical University of Denmark
Prof. Dr. R.J. Siezen, Professor of Bacterial Genomics, Radboud University Medical Centre
Dr. A. Nauta, Principal Scientist, FrieslandCampina Innovation

This research was conducted under the auspices of the Graduate School of VLAG

Community Dynamics of Complex Starter Cultures for Gouda-type Cheeses and its Functional Consequences

Oylum Erkuş

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 11 June 2014
at 4 p.m. in the Aula.

Oylum Erkuş
Community Dynamics of Complex Starter Cultures for Gouda-type Cheeses and its Functional Consequences, 216 pages.

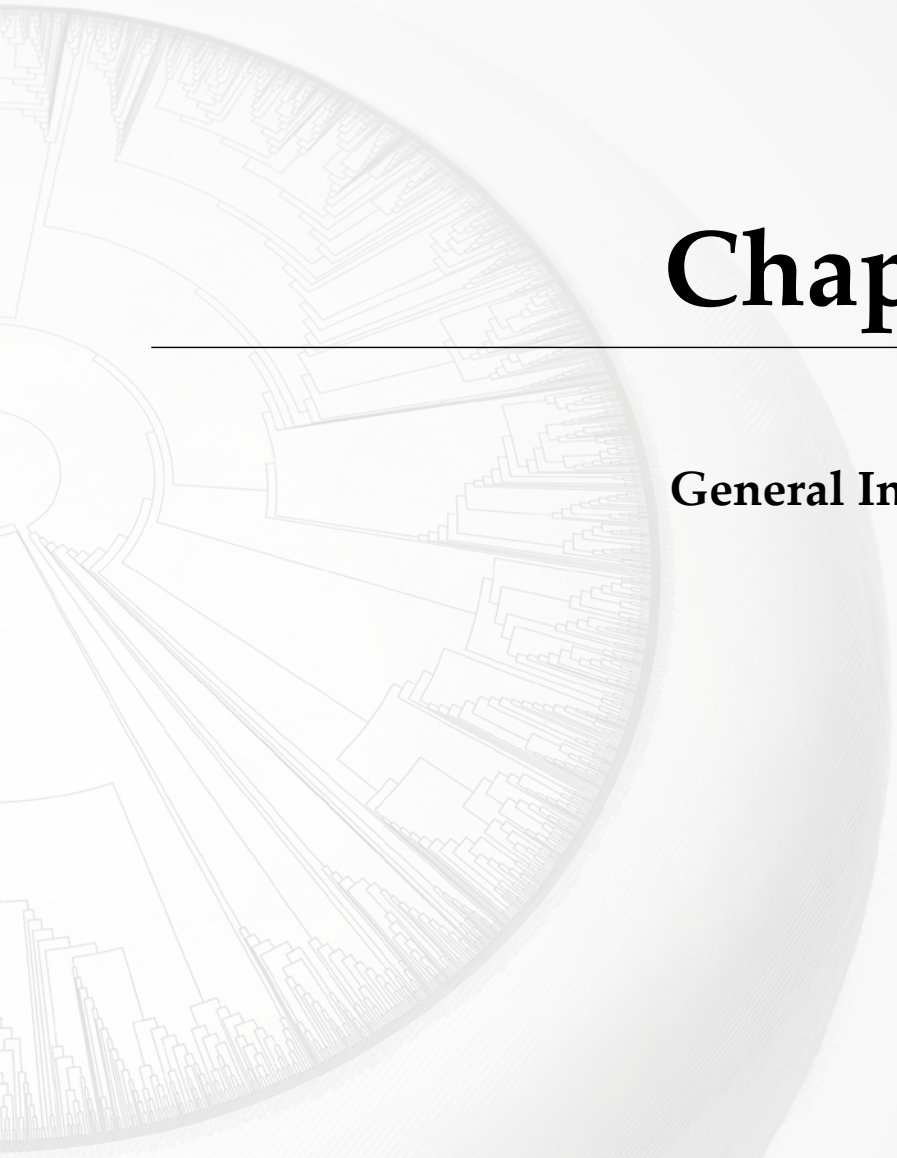
PhD thesis, Wageningen University, Wageningen, NL (2014)

With references, with summaries in Dutch and English

ISBN 978-94-6257-010-8

For the memory of my lovely grandfather

Table of contents



Chapter 1

General Introduction

Introduction

1

Ecosystems in nature have a substantial difference in diversity of the microbial inhabitants. The microbial communities inhabiting the complex ecosystems such as soil and intestine can accommodate up to 2000 species (1). On the contrary, some other ecosystem such as bacterial biofilms that inhabit acid mine drainage (2) and cheese fermentations accommodate a few species but can contain a high clonal diversity. The clonal diversification has been underestimated in the studies focused on species-rich ecosystem so far, but the influences of this level of clonal diversity is considerable in industrial dairy fermentations, and is the main topic of this thesis.

In artisanal cheese manufacturing tradition, the endogenous microbiota present in the previous batch of cheese or fermented milk has long been employed to initiate a new fermentation process in milk. In industrial production of today, this artisanal procedure of back-slopping has been replaced with the use of starter cultures with a lower number of lactic acid bacteria species to achieve more consistent process conditions, and standard final cheese attributes. The starter cultures used for Gouda cheese production accommodate strains of the species *Lactococcus lactis* and *Leuconostocs mesenteroides* and they can be of several types. Being more specific, the undefined and mixed type of starter cultures encompass several strains from each subspecies, which may show considerable functionality variation that contributes to industrial relevance during cheese manufacturing. Therefore, this thesis aims to describe the clonal diversity present in an undefined and mixed Gouda cheese starter culture and to understand better the importance of this clonal diversity in terms of their composite functionalities that are relevant for cheese manufacturing. Describing the strain level diversity and studying the functions of these distinct populations present in these so called “simple microbial communities” requires different approaches than those used for the microbial communities with high species diversity. As a result, the development of tools that are required to discriminate, monitor, and quantify the clonal populations is included in the thesis as a prerequisite to achieve these main targets.

To introduce the main subjects covered in the thesis, first of all, the evolutionary and ecological basis of the speciation and the relevance of the recent species definition will be introduced with a focus on the ecology of *Lactococcus lactis*. Then, several identification and fingerprinting techniques that are available for culture-dependent profiling of the microbial communities with a variable level of phylogenetic diversity

will be addressed. Following the culture-dependent analysis, the sequencing technologies applicable for the culture-independent community analyses will be discussed as a complementary approach to culturing. After introducing these basics of clonal diversification, and the tools available to work with this level of diversity, Gouda cheese starter cultures will be described as the main microbial community studied in this thesis. In the last part, the heterogeneity present in the Gouda cheese starter cultures will be explained and the influence of this heterogeneity on the functionalities that are relevant during cheese manufacturing will be emphasized. Finally, the outline of this thesis will be provided.

The Evolutionary and Ecological Basis of Speciation and Phylogenetic Analysis in Bacteria

The speciation in prokaryotes is driven by spontaneous mutations and horizontal gene transfer events. The divergence within and between the species can be variable depending on the spontaneous mutation rate, the time period since the divergence from the common ancestor, the number of ecological niches that the species can inhabit, the natural competence for transformation, as well as the availability of mobile genetic elements such as phages and plasmids for horizontal genes transfer (3). The definition of species is the basis for the phylogenetic classification of a bacterium and to evaluate the phylogenetic relatedness among the bacteria. At present, prokaryotic species is defined as a group of strains that have pairwise DNA re-association values over 70% in DNA-DNA hybridization (4). In addition, all the members of the species should share a certain degree of phenotypic characteristics. However, it is hard to incorporate the ecology of the micro-organism into this definition of species since the boundaries of the species are not determined by parameters derived from ecological theories (5). With the intention to integrate evolution and ecology considerations, various theoretical concepts described the definition of species (6-8), among which the ecotype concept (9, 10) is frequently considered to best explain the speciation in bacteria. According to the ecotype concept, all the related members of an ecological niche belong to a single, stable ecotype (a prokaryotic species), and the genetic diversity within the ecotype is restricted by cohesive forces such as periodic selection and genetic drift. Divergence of a member of an ecotype is only possible when a mutation or a recombination event results in the foundation of a new ecotype with a new ecological niche, such as sympatric speciation. The ecotypes are distinct in such a way that periodic selection or genetic drifts in one ecotype cannot prevent or induce the divergence in the other one.

These genetically congruent and ecologically distinct ecotypes are very promising units to take the evolutionary and ecological driving forces into account for the definition and delineation of the prokaryotic species. For example, *L. lactis* is a very diverse species, including *L. lactis* ssp. *lactis*, ssp. *cremoris* and ssp. *hordniae* as the main sub-clusters. The first two subspecies are encountered mostly in milk and dairy products, and less dominantly in various other ecological niches such as kimchi, orange juice, soil, and a variety of plant materials like, peas, grass, radish, onion, soya and bean sprouts (11). Historically, the subspecies names were assigned on basis of phenotypic traits. For example, *L. lactis* strains that were able to grow at 40°C or in the presence of more than 4% NaCl, that were able to utilize arginine as an energy source and that possessed active glutamate decarboxylase were designated as subspecies *lactis*. The genotypic characterization of these environmental isolates sometimes does not match with their subspecies designation determined based on the phenotypes observed (12, 13). For example, the isolates having ssp. *cremoris* -like genotypes are frequently encountered to display the phenotypes indicated (see above) for ssp. *lactis*. Methods like 16S rRNA gene sequencing or multi-locus sequence typing (MLST) are not powerful enough to reveal such niche-specifying ecologically distinct groups that can be encompassed within the same genotype clusters since they delineate the strains based on the single nucleotide polymorphisms within a small set of conserved genes (11). On the other hand, the comparative genome hybridization study, that clusters *L. lactis* strains on basis of the presence and absence of individual genes across the complete genome content, allowed the distinction of this ecologically different group showing ssp. *lactis* phenotypic characteristics from the rest of the strains showing ssp. *cremoris* phenotypic characteristics in the same ssp. *cremoris* type-strain-like genotype cluster (14). Therefore, the clustering based on the niche-specifying genetic information, which is analogous to the CGH analysis, provides biologically more meaningful discrimination and delineations of the members of a species.

A rapidly increasing number of complete bacterial genome sequences becomes available, which has been accelerating by the advances on high-throughput next generation sequencing technologies. To date, 32181 bacterial isolates of 7236 species have been sequenced (Genome online database, 25.01.2014). The availability of multiple complete genomes for single species also allows genome comparison studies, which improves our understanding of the main genetic polymorphisms that underlies the diversification in bacteria (3). Genome comparisons revealed unambiguously that the diversity in prokaryotes is much more characterised by

genetic content, e.g. strains carrying different sets of genes, than by allelic variation (point mutation) of conserved genes. This is in clear contrast to what has been concluded for eukaryotic genomes (15). This led to the use of the pan-genome concept to define the intra-species diversity observed in prokaryotes. The pan-genome of a species is defined as the combination of core genome and the accessory genes present in the members of that particular species. The core genome consists of genes shared by all the strains of the species (16) and in general includes the essential genes encoding for housekeeping and information processing functions. In contrast, the accessory genes (also called dispensable genome) are either strain-specific or present in some but not in all strains of the species (16). The accessory genes generally encode products involved in the interaction with the external environment, additional biochemical pathways, or associated with phages or mobile genetic elements (3). They frequently provide strain specific selective advantages during adaptation. A recent genome comparison of plant associated *L. lactis* ssp. *lactis* and the dairy associated *L. lactis* ssp. *lactis* indicated that the plant isolates carry many accessory gene sets that are associated with the adaptation to the plant niche and are not found on dairy isolates (17). These were mostly gene sets for the degradation of plant polymers such as xylan, arabinan, glucans, and fructans, genes for the uptake and conversion of plant cell wall degradation products such as α-galactosides, β-glucosides, arabinose, xylose, galacturonate, glucuronate, and gluconate, as well as genes involved in environmental competition such as nisin biosynthesis. Notably, the observation that most of the accessory genes are related to niche adaptation and that they are the main source of strain variations fits very well to the ecotype concept of speciation (see above). For example, the CGH analysis of 39 plant and dairy associated *L. lactis* isolates was employed to assess the presence and absence of genes using multi-strain microarrays, and enabled the distinct sub-clustering of dairy associated strains relative to the rest of the isolates within the ssp. *lactis* genotype cluster. This indicates that the clustering of accessory genes allowed the discrimination of niche (dairy) adaptation gene repertoires of *L. lactis* strains (14).

Biodiversity Beyond the Species Level and Fingerprinting Tools

The evolution rate of an organism is the most important factor when examining the intra-species diversity. For rapidly evolving species, the analysis of the slowly evolving core genome is probably a more solid basis for the identification of species and sub-species, while it to some extent may also enable the detection of clonal diversification. In general, small-subunit 16S rRNA gene sequencing is the standard

technique for the identification of the species, and multi-locus sequence typing (MLST) was designed to delineate the major clades and/or clonal complexes within a species and describe their phylogenetic relationship, by targeting the slowly evolving core genome elements (18). MLST exploits sequence polymorphisms located within several highly conserved housekeeping loci selected from the core genome of a species, and the degree of delineation depends on the number of loci used and the number of polymorphisms accumulated in the clade and/or clonal complexes. Since the unit of comparison is based on alleles of highly conserved genes, the confounding noise coming from horizontal or vertical gene transfer events is minimal (19). However, since MLST uses a very limited part of the core genome, it does not provide information about the divergence of accessory genome and gene content. Therefore, this technology provides a very useful approach for the reliable identification and diversification of species with a high evolution rates, and is frequently employed for the pathogenic species (20), but is inappropriate for the determination of strain level diversity that depend much more on their discriminative pan-genomic content (see above).

Strain delineation at a resolution that goes beyond the major clade and clonal complexes, requires profiling approaches that can detect polymorphisms at a genome-wide scale. Among the commonly used fingerprinting techniques for this purpose are, repetitive PCR (rep-PCR) (21), randomly amplified polymorphic DNA (RAPD) (22), amplified fragment length polymorphism (AFLP) (23), and microarray based comparative genome hybridization studies (14). However, with the consistently increasing efficacy of whole genome sequencing methodologies, the full-genome sequencing approach for delineation of clades and clonal clusters is becoming more realistic, and obviously provides the highest resolution. The common basis of Rep-PCR, RAPD, and AFLP is that they make use of a range of amplified fragments of the genome. The fragments differ in size depending on the presence of single nucleotide polymorphisms (SNP) within the genomic DNA that change relative position of restriction sites or determine the annealing position of primers within the genome, whereas the presence of insertions- or deletions also contribute to the length variations of amplified fragments (24). Therefore, the markers generated in PCRs are a cumulative measure of the SNPs and structural variations in the genome, and their size distribution patterns are used for clustering based strain delineation. The strain-specific markers generated with these techniques can also be used for tracing or quantification purposes in population dynamics studies (25). These (high resolution) genetic fingerprinting techniques are especially useful to describe the

diversity within relatively simple microbial communities with limited species but high clade and clonal diversity, such as those found in dairy fermentations. Comparative genomics through direct comparison of full genome sequences of individual isolates, or through indirect comparison of isolates with comparative genome hybridization, can provide higher resolution and a more comprehensive strain delineation. It provides information about the ecology of the strains by generating insight in the presence and absence of genes within each genome that can be employed for ecotype distinction (26).

Use of Sequencing Technologies for Community Composition Analyses

The description of the community compositions by molecular identification and typing of the community members requires the isolation of community members with the use of selective and/or enrichment media. However, the micro-organisms present in the communities might have variable growth requirements that might not be accommodated in the growth media and/or growth conditions applied during culturing, which leads to the differential culturability of community members and eventually to biased community composition information. The use of sequencing technologies additional to the culture based analyses can improve the accuracy of the community profiling studies. In order to determine the resolution of the sequencing technology that is required for the adequate profiling of the diversity within the community, the complexity of the community and its species richness should be considered carefully (26). For complex communities with high species diversity, 16S rRNA gene pyrosequencing and deep-sequencing based metagenomics are suitable approaches to obtain information about the genera or the species present in the community. 16S rRNA gene pyrosequencing has been applied on very complex environmental samples such as soil (27) but also on more simple communities involved in food fermentations, including soybean (28) and rice bran (29) fermentations. Metagenomics is also widely used to understand the structure of complex environmental samples such as those obtained from the mammalian GI-tract that can contain up to 2000 microbial species (1). However, when a lower complexity is expected in the microbial communities such as those encountered in industrial food fermentations, the resolution provided by 16S rRNA approaches is insufficient to characterize the actual community richness. For example, the metagenomic analysis of Kimchi fermentation revealed that the microbial community involved was dominated by *Leuconostoc, Lactobacillus* and *Weissella* species, but failed to detect the clades and strains present in the system

(30). Similarly, the 16 S profiling and metagenomic analyses of small intestinal tract community revealed a relatively simple community structure that was dominated in many individuals by Streptococcal species (31) but failed to identify the variety among the Streptococcal clades and lineages that was discovered through cultivation approaches (32). Increasing the metagenomic sequencing depth combined with mapping of the metagenome information on whole genome sequences of isolates recovered from the microbial community can provide the resolution that enables the monitoring of the dynamics of clades and/or individual strains during the fermentation process (26).

The microbial ecology studies in the literature are frequently focused on the analysis of microbial communities with high species diversity, describing the ecology of the microbial communities residing in the respective niches by studying the interactions at the species level. However, the clonal diversification within these communities was hardly investigated. This may in part be due to the lack of proper high resolution tools to apply in clonal diversity research within a phylogenetically diverse microbial community, whereas various tools are available for the genetic and functional analysis of communities at the species level. Studying the 'simple' microbial communities holds great potential for the improvement of, and innovation in the industrial and scientific applications of these 'simple' communities. A very relevant example for such an industrial application with a high added value in economy is the starter cultures that are used in the production of Gouda cheeses. In order to discover the influences of this level of clonal diversity in the cheese manufacturing process, this thesis focuses on Gouda cheese starter cultures, which will be introduced in further sections.

Gouda Cheese Starter Cultures and Cheese Manufacturing

Archaeological evidence indicates that the use of agricultural techniques and production of fermented milk such as yoghurt appeared in the same time frame in several civilizations such as Sumerians, Babylonians, Indians, and the Pharoes in north-east Africa (33), and dates back to approximately 3000 BC (34). Milk sours rapidly by the exposure to high air temperatures and lack of cooling, and when microbial contamination occurs from hand-milking, open contact with air and animals. This is how nomadic people are believed to have discovered the milk fermentations. Subsequently, further steps such as, heating of milk that lowers the endogenous load of microorganisms, concentration of yoghurt or whey separation and salting, were

coupled to the fermentation to extend the preservation period of the coagulated milk (33). The traditional method of back-slopping, using part of the previous batch of a fermented product to inoculate a new batch of milk, has been used for centuries in dairy fermentations. However, the entire process of traditional cheese making was not standardized and the final product attributes were inconsistent. With the increasing industrial need for consistency, the cheese manufacturing process was standardized and more controlled lactic acid bacteria cultures, known as starter cultures, were developed to initiate the dairy fermentations. There are more than 1000 cheese varieties produced worldwide (35), which differ due to the differences in ambient climate, microbial flora of the region, moisture and salt content, as well as other differences in processing. Most of the well-known cheese varieties are manufactured now at industrial scale with the use of starter cultures and sometimes with the addition of secondary starter cultures containing different species of lactic acid bacteria, fungi or yeast depending on the variety of cheese produced. These cheese varieties have been reviewed elsewhere (36), and here, focus will be given to Gouda type cheese starter cultures.

The Gouda cheese is a semi-hard, Netherlands-originated cheese variety with a fat content in dry matter between 48 and 51% (w/w), and it has a mild creamy taste. The starter cultures used for Gouda cheese production are composed of mesophilic bacteria harbouring several strains from the subspecies' of *Lactococcus lactis,* such as *L. lactis* ssp. *lactis*, *L. lactis* ssp. *cremoris*, and *L. lactis* ssp. *lactis* biovar. *diacetylactis* and from *Leuconostocs mesenteroides* in different combinations. The industrial starter cultures are classified in several different ways based on how well the composition of the culture is known, the type of the subspecies present in the culture, and their propagation environment. Based on the characterization of the composition, starters are classified into two main categories: defined and undefined mixed cultures (Beresford, 2001). Defined cultures contain a known number of strains (single or multiple) usually with known functional traits, whereas undefined cultures contain an unknown number of strains and species. Based on species presence, starter cultures can be classified into 4 types; those that contain only *L. lactis* ssp. *lactis* and ssp. *cremoris* are called O-type; those that contain *L. lactis* ssp. *lactis*, ssp. *cremoris*, and additionally the citrate utilizing ssp. *lactis* biovar. *diacetylactis* are called D-type; those that contain *L. lactis* ssp. *lactis*, ssp. *cremoris* and additionally the citrate utilizing *Leuconostoc mesenteroides* are called L- or B- type; and finally those that contain *L. lactis* ssp. *lactis*, ssp. *cremoris* and additionally both of the citrate utilizers *L. lactis* ssp. *lactis* biovar. *diacetylactis* and *Leuconostoc mesenteroides* are called LD-

or BD-type cultures (37). Finally, the cultures can be further classified on basis of their propagation environment, where starter cultures that are propagated under sterile conditions are called L- cultures (L stands for laboratory) and those that are propagated under non-aseptic conditions in dairy plants and that cannot exclude the presence of bacteriophages are called P-cultures (P stands for practice) (38).

Gouda cheese manufacturing is schematically represented in Figure 1. Very briefly, the process starts with the milk coagulation, which is initiated with the addition of chymosin and the starter culture. Milk is composed of casein miscelles that are stabilized with the dissolved salt ions such as calcium and phosphate around the pH of 6.6. K-casein surrounds the casein micelles, preventing the contact between α- and β-casein. Chymosin breaks down K-casein into para-K-casein, resulting in the precipitation of α- and β-casein with calcium upon interaction, which is known as milk curdling. Additionally, the lactic acid accumulated by the activity of the starter culture also facilitates the destabilization of the casein miscelles by lowering the pH. Following the curdling, the curd is cut into small pieces to allow the separation of whey, followed by washing to remove excess lactose, cooking to improve the synersis, pressing for the complete removal of the whey, brining to limit the microbial activity during the ripening, and ends with the actual ripening, which takes a period of 6 weeks to 2 years (39, 40).

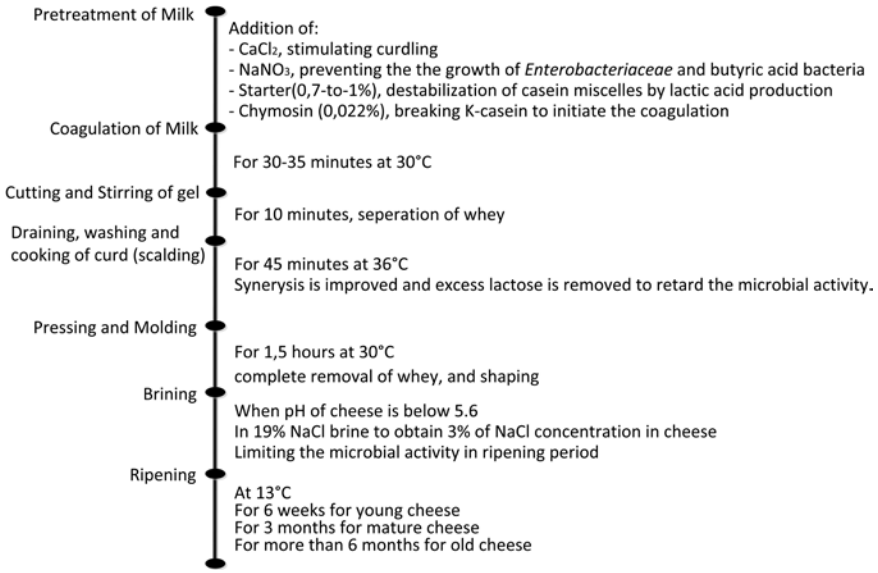


Figure 1. Gouda Type Cheese Manufacturing Process

The Importance of Microbial Heterogeneity in Starter Cultures for Cheese Manufacturing

The microbial community of the Gouda type cheese starter culture can be considered as a simple community harbouring only two species: *Lactococcus lactis* and *Leuconostocs mesenteroides*. *L. lactis* is a facultative heterofermentative species, but under most conditions executes a typical homofermentative metabolism, by which lactose is fermented almost exclusively into lactic acid. *Leuconostoc* species are generally considered as heterofermentative, and produce lactate, ethanol acetate and CO_2. They can ferment lactose only at a very low rate. Therefore, the rapid lactic acid accumulation during the coagulation step, which drives the curd formation together with the activity of chymosin, is mostly dependent on the *L. lactis* cells within the starter culture. All three *L. lactis* subspecies, which were given the subspecies names as ssp. *cremoris*, ssp. *lactis* and ssp. *lactis* biovar. *diacetylactis* based on their phenotypes, can be present in the starter culture, and they provide quite distinctive metabolic functions to the overall starter culture functionality. The strains with the subspecies *cremoris* phenotype are the most fastidious; are unable to grow at 40°C or in the presence of more than 4% NaCl; can't utilize arginine as an energy source; lack glutamate decarboxylase activity, which are all discriminating the strains with ssp. *cremoris* phenotypes from the strains with subspecies *lactis* phenotypes (11). These differential phenotypes may contribute to differential population dynamics of ssp. *lactis* and ssp. *cremoris* cells during cheese manufacturing, especially after the scalding challenge at 46°C, and after the brining step that proceeds to the ripening period. Furthermore, the third subspecies *lactis* biovar. *diacetylactis* is characterised by its capacity to ferment citrate into diacetyl and acetoin, which plays a significant role in flavour formation during cheese ripening. Finally, *Leuconostoc mesenteroides* represents only a small fraction (approximately 1%) of the overall starter culture compositions in general, which is due to the non-optimal growth conditions used during the starter production. Nevertheless, this species is considered to be very important for the production of diacetyl during cheese ripening next to the biovar. *diacetylactis*. Diacetyl generation during cheese ripening is considered as one of the key-examples that involve metabolic cooperation between *Lactococcus* and *Leuconostocs*, in which lactose fermentation by *Lactococcus* rapidly generates lactic acid that lowers the pH, which favours the uptake of citrate by the citrate permease of *Leuconostoc* as well as *L. lactis* ssp. *lactis* biovar. *diacetylactis*.

Even though the starter cultures are considered simple in terms of species diversity, the undefined and mixed starter culture communities that have been propagated under non-aseptic conditions and unprotected from bacteriophages show a large diversity beyond the subspecies level in terms of plasmid content, phage sensitivity, and for many other functionalities, such as proteolytic activity, acid production, autolysis and flavour related metabolite formation. For example, the analysis of commonly used undefined and mixed starter cultures showed the presence of strains with different plasmid profiles, acidification rates as well as different range of phage sensitivities (41). The plasmid complement of *Lactococcus* strains can make up around 10% of their genetic repertoire and it is one of the major sources of the genetic diversity among strains (42). *Lactococcus* cells can harbour up to 7 plasmids (42), and importantly many of the functions that are involved in dairy fermentations such as lactic acid production, proteolysis, citrate metabolism, bacteriophage resistance, bacteriocin production and antibiotic resistance are encoded on plasmids rather than the chromosome (43). Next to these well-established and industrially relevant functionalities, plasmids also encompass numerous hypothetical genes, which may provide various selective advantages during cheese fermentation to strains that harbour them.

Similar to plasmids, phages are directly involved in dairy fermentations as well. Lysis of the starter culture populations due to phage predation at early steps of cheese manufacturing may slow down or eliminate the acid production, leading to elevated pH of the curd, and substantial amounts of residual lactose in cheese. These factors impact on synersis, and lead to a higher moisture content of the cheese, which may allow the growth of pathogens during the ripening period (44). These phage predation problems early during the manufacturing process, are especially known for defined starter cultures (single strain or combination of strains), in which phage predation can eliminate a substantial proportion of the overall starter culture used. This is in clear contrast to the higher phage resistance that is commonly observed in undefined mixed cultures, especially those that were propagated with back-slopping regimes executed in non-aseptic dairy factory environments and are thus not protected against phages (practice [P-] cultures) (38). This is probably due to the presence of subpopulations in these undefined cultures with differential phage resistance profiles.

Besides their predator activity, phages can also contribute significantly to horizontal gene transfer events in bacteria. Especially prophages are believed to have an

important role in the adaptation of bacterial strains to their ecological niche (45). The comparison of closely related bacterial genomes indicates that the prophage related sequences account for a big fraction of the differences observed (46) among the genomes. Commonly, these accessory genes in bacterial genomes are clustered as genomic islands (e.g. prophage regions), which can in some cases be recognized by their difference in G+C content relative to the overall genomic average (15). Notably, *L. lactis* strains may also contain many phage-related sequences in their genomes, such as the completely sequenced *L. lactis* strain IL1403 that encodes six inducible and non-inducible prophage regions within its genome (47).

In addition to the continuous genomic and functional variability contributed by the plasmids and phages, *L. lactis* is very diverse in terms of gene content and structural variations like single nucleotide polymorphisms, insertions, deletions, inversions and other genomic rearrangements. Thus, the functions that are encoded by the chromosomal DNA, such as the formation of metabolites that are important for flavour formation through amino acid catabolism, can be substantially different between strains. *L .lactis* cells can persist in a wide range of environments, and were shown to be evolved to ecotypes that more effectively acquire the nutrient resources available in particular environments (17). For example, plant associated *L. lactis* isolates are generally not proteolytic since the protein concentration is very low in plant niches. They encode complete amino acid biosynthesis pathways, and in general have no (or only few) amino acid auxotrophies (48). In contrast, the dairy associated *L. lactis* isolates are characterised by a highly common auxotrophy for branched chain amino acids (BCAA) and histidine (49, 50), which are the most abundant amino acids generated through degradation of casein in milk. The BCAA auxotrophies were shown to be either due to point mutations in the genes of the corresponding synthesis pathways (e.g. in *L. lactis* IL1403 (49)) or due to complete loss of gene(s) involved in the pathway, which is clearly illustrated by the CGH analysis of dairy and non-dairy associated *L. lactis* strains revealing the absence or incompleteness of the *leu*ABCD–*ilv*DBHCA genes in several dairy isolates (14). Moreover, a recently performed experimental evolution experiment in which a plant associated *L. lactis* strain was adapted to effective growth in milk, also confirmed that adaptation to the protein-rich milk environment involves suppression of functions and loss of functions that are not relevant for persistence and survival in that niche (51). The genetic repertoire of branched chain amino acid metabolism was shown to be enriched in specific mutations during adaptation to milk, which also led to substantial changes at gene-transcription level. In addition, comparative

transcriptome analysis of four closely related *L. lactis* ssp. *cremoris* strains revealed that the expression of oligopeptide transporters and branched chain aminotransferase gene *bcaT* were strain specific under the cheddar cheese manufacturing conditions (52). Degradation products from branched-chain amino acids are among the most important flavour compounds in cheese (53), and such genome or transcriptome variations can be anticipated to affect the flavour profile of the cheese produced.

As demonstrated, the mixed Gouda cheese starter cultures harbour very large strain variation at many different levels: in plasmid content, phage sensitivity and chromosomal genome complement, which eventually have a considerable effect on the metabolic functions required in cheese manufacturing. The studies that are covered in this thesis focus on the comprehensive analysis of the Gouda-type mixed cheese starter culture with the aim to better understand, improve and innovate the cheese manufacturing processes, and the outline of the performed studies are given below.

Outline of This Thesis

The studies presented in this thesis aim to decipher the level of diversity within a complex starter culture, identify the ecological basis of this diversity, investigate the biological source that generates and maintains this diversity and finally what the industrial relevance of this diversity is in terms of functionalities during cheese manufacturing and ripening. These approaches are intended to eventually enable accurate prediction and control of the cheese manufacturing process using (un) defined starter cultures, but may also allow rational design and development of new starter cultures.

Chapter 2 describes the development of a high throughput and high resolution amplified fragment length polymorphism (AFLP) based fingerprinting technique, which allows the discrimination of closely related *L. lactis* strains. The resolution depth of the optimized procedure was confirmed through robust and reliable diversity analysis on 82 *L. lactis* strains, including both closely and distantly related strains of dairy and non-dairy origin. Additionally, the AFLP fingerprints also allowed the identification of a genetic marker for the discrimination of the strains with subs. *cremoris* and subs. *lactis* specific phenotypes within the subsp. *cremoris* genotype cluster.

Chapter 3 describes the analysis of the microbial community of a mixed Gouda cheese starter culture that has a long history of use in cheese manufacturing. The analysis was performed using culturing and metagenomic approaches. The community diversity was investigated for a large number of cultured isolates, using the AFLP technique optimized in Chapter 2. Moreover, the plasmid content was also profiled in these isolates using an adapted AFLP-based plasmid typing approach. These genetic typing approaches identified seven distinct genomic lineages of the species *Lactococcus lactis* with variable plasmid content, and a single lineage of *Leuconostoc mesenteroides*. Phenotypic assessments of the isolates included determination of phage sensitivities using the plaque assays, revealing substantial variations of phage sensitivities among the strains encompassed within the lactococcal lineages. The community dynamics throughout the cheese manufacturing process was monitored at the genomic lineage level. Moreover, the role of phage predation as a determinant of starter culture diversity during starter-culture propagation was investigated.

Chapter 4 describes the development of a technique for the selective amplification of DNA extracted from live fraction of the microbial community in cheese using propidium monoazide (PMA). In general, environmental samples contain significant numbers of dead cells with compromised membranes and /or free DNA, and the community composition data deduced from (meta)genetic analyses, may fail to reflect the live fraction. In this chapter, the use of PMA was tested for the inhibition of PCR amplification on DNA derived from dead or compromised cells in different cheese samples. The selective genetic detection of viable populations during cheese manufacturing was compared to community dynamics known from culture dependent methods. These analyses show that PMA facilitates the selective amplification of live starter culture populations in cheese, which is also enabling selective metagenomic analysis of the viable population using whole genome amplification procedures.

Chapter 5 describes the molecular basis of the selective recovery of *L. lactis* cells that belong to genetic lineage 7 on Reddy's agar, which was observed during cultivation of the microbial community present in Gouda cheese starter culture on complementary growth media in chapter 2. In this chapter, arginine was identified as the medium component that determined the higher recovery of *L. lactis* cells on Reddy's agar. Furthermore, the molecular response of the lineage 7 representative strain (*L. lactis* TIFN7) to arginine was investigated by whole genome gene expression profiling in the absence and presence of excess arginine. This study pinpoints the arginine

mediated growth effects to a deficiency in C1-metabolism of this *L. lactis* lineage, which can be compensated by environmental supplementation of excess arginine, but also by other environmental C1-sources like carbon dioxide or folate.

Chapter 6 describes the effects of the variation in propagation regime on the community composition of a mixed starter culture and connects the composition change to the functionalities that impact on flavour development during cheese manufacturing. In this chapter, the mixed starter culture was serially propagated from different growth phases with the aim to modulate and steer the community composition, and to evaluate whether back-slopping regimes can be used as a novel approach towards the development of new mixed starter cultures. The final community compositions, acidification potential and proteolytic activity of the serially propagated cultures were compared to evaluate the modulation of the community compositions by the back-slopping regimes. Moreover, the flavour formation in 12-weeks ripened mini-cheeses manufactured with distinctly propagated starter cultures were compared to evaluate the effect of community composition change on flavour formation capacity.

Chapter 7 summarizes the main results obtained in this thesis, and gives some concluding remarks and future perspectives.

References

1. **Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Jian M, Zhou Y, Li Y, Zhang X, Qin N, Yang H, Wang J, Brunak S, Dore J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, Ehrlich SD.** 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature **464:**59-65.
2. **Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF.** 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature **428:**37-43.
3. **Feil EJ.** 2004. Small change: keeping pace with microevolution. Nature reviews. Microbiology **2:**483-495.
4. **Achtman M, Wagner M.** 2008. Microbial diversity and the genetic nature of microbial species. Nature reviews. Microbiology **6:**431-440.
5. **Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, Stackebrandt E, Van de Peer Y, Vandamme P, Thompson FL, Swings J.** 2005. Opinion: Re-evaluating prokaryotic species. Nature reviews. Microbiology **3:**733-739.
6. **de Queiroz K.** 2005. Ernst Mayr and the modern concept of species. Proceedings of the National Academy of Sciences of the United States of America **102 Suppl 1:**6600-6607.
7. **Rossello-Mora R, Amann R.** 2001. The species concept for prokaryotes. FEMS microbiology reviews **25:**39-67.
8. **Staley JT.** 2006. The bacterial species dilemma and the genomic-phylogenetic species concept. Philosophical transactions of the Royal Society of London. Series B, Biological sciences **361:**1899-1909.

9. **Cohan FM, Perry EB.** 2007. A systematics for discovering the fundamental units of bacterial diversity. Current biology : CB **17:**R373-386.
10. **Koeppel A, Perry EB, Sikorski J, Krizanc D, Warner A, Ward DM, Rooney AP, Brambilla E, Connor N, Ratcliff RM, Nevo E, Cohan FM.** 2008. Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. Proceedings of the National Academy of Sciences of the United States of America **105:**2504-2509.
11. **Rademaker JL, Herbet H, Starrenburg MJ, Naser SM, Gevers D, Kelly WJ, Hugenholtz J, Swings J, van Hylckama Vlieg JE.** 2007. Diversity analysis of dairy and nondairy *Lactococcus lactis* isolates, using a novel multilocus sequence analysis scheme and (GTG)5-PCR fingerprinting. Applied and environmental microbiology **73:**7128-7137.
12. **Kelly W, Ward L.** 2002. Genotypic vs. phenotypic biodiversity in *Lactococcus lactis*. Microbiology **148:**3332-3333.
13. **Salama MS, Sandine WE, Giovannoni SJ.** 1993. Isolation of *Lactococcus lactis* subsp. *cremoris* from nature by colony hybridization with rRNA probes. Applied and environmental microbiology **59:**3941-3945.
14. **Siezen RJ, Bayjanov JR, Felis GE, van der Sijde MR, Starrenburg M, Molenaar D, Wels M, van Hijum SA, van Hylckama Vlieg JE.** 2011. Genome-scale diversity and niche adaptation analysis of *Lactococcus lactis* by comparative genome hybridization using multi-strain arrays. Microbial biotechnology **4:**383-402.
15. **Rodriguez-Valera F, Martin-Cuadrado AB, Rodriguez-Brito B, Pasic L, Thingstad TF, Rohwer F, Mira A.** 2009. Explaining microbial population genomics through phage predation. Nature reviews. Microbiology **7:**828-836.
16. **Tettelin H, Riley D, Cattuto C, Medini D.** 2008. Comparative genomics: the bacterial pan-genome. Current opinion in microbiology **11:**472-477.
17. **Siezen RJ, Starrenburg MJ, Boekhorst J, Renckens B, Molenaar D, van Hylckama Vlieg JE.** 2008. Genome-scale genotype-phenotype matching of two *Lactococcus lactis* isolates from plants identifies mechanisms of adaptation to the plant niche. Applied and environmental microbiology **74:**424-436.
18. **Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG.** 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America **95:**3140-3145.
19. **Maiden MC, van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA, McCarthy ND.** 2013. MLST revisited: the gene-by-gene approach to bacterial genomics. Nature reviews. Microbiology **11:**728-736.
20. **Perez-Losada M, Cabezas P, Castro-Nallar E, Crandall KA.** 2013. Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases **16:**38-53.
21. **Ventura M, Zink R.** 2002. Specific identification and molecular typing analysis of L*actobacillus johnsonii* by using PCR-based methods and pulsed-field gel electrophoresis. FEMS microbiology letters **217:**141-154.
22. **Foschino R, Nucera D, Volponi G, Picozzi C, Ortoffi M, Bottero MT.** 2008. Comparison of *Lactococcus garvieae* strains isolated in northern Italy from dairy products and fishes through molecular typing. Journal of applied microbiology **105:**652-662.
23. **Thompson FL, Hoste B, Vandemeulebroecke K, Swings J.** 2001. Genomic diversity amongst Vibrio isolates from different sources determined by fluorescent amplified fragment length polymorphism. Systematic and applied microbiology **24:**520-538.
24. **Savelkoul PH, Aarts HJ, de Haas J, Dijkshoorn L, Duim B, Otsen M, Rademaker JL, Schouls L, Lenstra JA.** 1999. Amplified-fragment length polymorphism analysis: the state of an art. Journal of clinical microbiology **37:**3083-3091.
25. **Meudt HM, Bayly MJ.** 2008. Phylogeographic patterns in the Australasian genus Chionohebe (Veronica s.l., Plantaginaceae) based on AFLP and chloroplast DNA sequences. Molecular phylogenetics and evolution **47:**319-338.
26. **van Hijum SA, Vaughan EE, Vogel RF.** 2013. Application of state-of-art sequencing technologies to indigenous food fermentations. Current opinion in biotechnology **24:**178-186.

27. **Will C, Thurmer A, Wollherr A, Nacke H, Herold N, Schrumpf M, Gutknecht J, Wubet T, Buscot F, Daniel R.** 2010. Horizon-specific bacterial community composition of German grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes. Applied and environmental microbiology **76:**6751-6759.

28. **Nam YD, Lee SY, Lim SI.** 2012. Microbial community analysis of Korean soybean pastes by next-generation sequencing. International journal of food microbiology **155:**36-42.

29. **Sakamoto N, Tanaka S, Sonomoto K, Nakayama J.** 2011. 16S rRNA pyrosequencing-based investigation of the bacterial community in nukadoko, a pickling bed of fermented rice bran. International journal of food microbiology **144:**352-359.

30. **Jung JY, Lee SH, Kim JM, Park MS, Bae JW, Hahn Y, Madsen EL, Jeon CO.** 2011. Metagenomic analysis of kimchi, a traditional Korean fermented food. Applied and environmental microbiology **77:**2264-2274.

31. **Zoetendal EG, Raes J, van den Bogert B, Arumugam M, Booijink CC, Troost FJ, Bork P, Wels M, de Vos WM, Kleerebezem M.** 2012. The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. The ISME journal **6:**1415-1426.

32. **van den Bogert B, Erkus O, Boekhorst J, de Goffau M, Smid EJ, Zoetendal EG, Kleerebezem M.** 2013. Diversity of human small intestinal *Streptococcus* and *Veillonella* populations. FEMS microbiology ecology **85:**376-388.33. **Tamime AY, Robinson, R.K.** 1985. Biochemistry of Fermentation, p. 295-327. *In* Tamime AY (ed.), Yoghurt: Science and Technology. Pergamon Press, Great Britain.

34. **Battcock M, Azam-Ali, S.** 1998. Fermented fruits and vegetables. A global perpective, FAO Agricultural Services Bulletin, vol. 134. Food and Agriculture Organization of the United Nations, Rome, Italy.

35. **Sandine WE, Elliker, P. R.** 1970. Microbially induced flavours and fermented foods flavour in fermented dairy products. Journal of Agriculture and Food Chemistry **18:** 557-562.

36. **Beresford TP, Fitzsimons, N.A., Brennan, N.L., Cogan,T.M. .** 2001. Recent advances in cheese microbiology. International Dairy Journal **11** 259–274.

37. **Vedamuthu ER.** 2006. Starter Cultures for Yogurt and Fermented Milks, p. 89-116. *In* Chandan RC (ed.), Manufacturing Yogurt and Fermented Milks. Blackwell Publishing Ltd, UK.

38. **Stadhouders J, Leenders, G.J.M.** 1984. Spontaneously developed mixed-strain cheese starters. Their behaviourstowards phages, and theri use in Dutch cheese industry. Netherlands Milk and Dairy Journal **38:**25.

39. **Walstra P.** 1993. The Synersis of curd, p. 141-191. *In* Fox PF (ed.), Cheese:Chemistry, physics and microbiology, vol. Vol. 1. Elsevier Applied Science, London.

40. **van den Berg G, Meijer, W.C. , Düsterhöft, E.-M. , Smit, G. .** 2004. Gouda and related cheeses, p. 103-140. *In* Fox PF (ed.), Cheese: Chemistry, physics and microbiology, vol. Vol. 2. Elsevier Applied Science, London.

41. **Lodics T, Steenson, L. .** 1993. Phage–host interactions in commercial mixed-strain cultures: Practical significance, a review. Journal of Dairy Science **76:**2380–2391.

42. **Mills S, McAuliffe OE, Coffey A, Fitzgerald GF, Ross RP.** 2006. Plasmids of lactococci - genetic accessories or genetic necessities? FEMS microbiology reviews **30:**243-273.

43. **Siezen RJ, Renckens B, van Swam I, Peters S, van Kranenburg R, Kleerebezem M, de Vos WM.** 2005. Complete sequences of four plasmids of *Lactococcus lactis* subsp. *cremoris* SK11 reveal extensive adaptation to the dairy environment. Applied and environmental microbiology **71:**8371-8382.

44. **Marco MB, Moineau S, Quiberoni A.** 2012. Bacteriophages and dairy fermentations. Bacteriophage **2:**149-158.

45. **Brussow H, Hendrix RW.** 2002. Phage genomics: small is beautiful. Cell **108:**13-16.

46. **Glaser P, Frangeul L, Buchrieser C, Rusniok C, Amend A, Baquero F, Berche P, Bloecker H, Brandt P, Chakraborty T, Charbit A, Chetouani F, Couve E, de Daruvar A, Dehoux P, Domann E, Dominguez-Bernal G, Duchaud E, Durant L, Dussurget O, Entian KD, Fsihi H, Garcia-del Portillo F, Garrido P, Gautier L, Goebel W, Gomez-Lopez N, Hain T, Hauf J, Jackson D, Jones LM, Kaerst U, Kreft J, Kuhn M, Kunst F, Kurapkat G, Madueno E, Maitournam A, Vicente JM, Ng E, Nedjari H, Nordsiek G, Novella S, de Pablos B, Perez-Diaz JC, Purcell R, Remmel B, Rose M, Schlueter T, Simoes N, Tierrez A, Vazquez-Boland JA, Voss H, Wehland J, Cossart P.** 2001. Comparative genomics of *Listeria* species. Science **294:**849-852.

47. **Chopin A, Bolotin A, Sorokin A, Ehrlich SD, Chopin M.** 2001. Analysis of six prophages in *Lactococcus lactis* IL1403: different genetic structure of temperate and virulent phage populations. Nucleic acids research **29:**644-651.

48. **Ayad EH, A. Verheul, C. De Jong, J. T. Wouters, and G. Smit.** 1999. Flavour forming abilities and amino acid requirements of *Lactococcus lactis* strains isolated from artisanal and non-dairy origin. . International Dairy Journal **9:**725-735.

49. **Godon JJ, Delorme C, Bardowski J, Chopin MC, Ehrlich SD, Renault P.** 1993. Gene inactivation in *Lactococcus lactis*: branched-chain amino acid biosynthesis. Journal of bacteriology **175:**4383-4390.

50. **Delorme C, Godon JJ, Ehrlich SD, Renault P.** 1993. Gene inactivation in *Lactococcus lactis:* histidine biosynthesis. Journal of bacteriology **175:**4391-4399.

51. **Bachmann H, Starrenburg MJ, Molenaar D, Kleerebezem M, van Hylckama Vlieg JE.** 2012. Microbial domestication signatures of *Lactococcus lactis* can be reproduced by experimental evolution. Genome research **22:**115-124.

52. **Taibi A, Dabour N, Lamoureux M, Roy D, LaPointe G.** 2011. Comparative transcriptome analysis of *Lactococcus lactis* subsp. *cremoris* strains under conditions simulating Cheddar cheese manufacture. International journal of food microbiology **146:**263-275.

53. **Smit G, Smit BA, Engels WJ.** 2005. Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. FEMS microbiology reviews **29:**591-610.

Chapter 2

High-resolution AFLP Typing of *Lactococcus lactis* Strains Enables Identification of Genetic Markers for Subspecies Related Phenotypes

Oylum Erkuş Kütahya, Marjo J. C. Starrenburg, Jan L. W. Rademaker, Corné H.W. Klaassen, Johan E. T. van Hylckama Vlieg, Eddy J. Smid, Michiel Kleerebezem

This chapter has been published as Appl. Environ. Microbiol; 2011 August; 77(15):5192-8

Abstract

A high-resolution amplified fragment length polymorphism (AFLP) methodology was developed to achieve the delineation of closely related *Lactococcus lactis* strains. The differentiation depth of 24 enzyme/primer/nucleotide combinations was experimentally evaluated to maximize the number of polymorphisms. The resolution depth was confirmed by performing diversity analysis on 82 *L. lactis* strains including both closely and distantly related strains with dairy and nondairy origins. Strains clustered into two main genomic lineages of *L. lactis* subsp. *lactis* and *L. lactis* subsp. *cremoris* type-strain-like-genotypes and a third novel genomic lineage rooted from *L. lactis* subsp. *lactis* genomic lineage. Cluster differentiation was highly correlated with small subunit rRNA homology and multi-locus sequence analysis studies (MLSA). Additionally, the selected enzyme/primer combination generated *cremoris*-phenotype specific fragments irrespective of their genotype. These phenotype-specific markers allowed the differentiation of *lactis*-phenotype from *cremoris*-phenotype strains within the same *L. lactis* subsp. *cremoris* type-strain-like genomic lineage, illustrating the potential of AFLP for the generation of phenotype-linked genetic markers.

Introduction

Lactococcus lactis is commonly found in defined and artisan cheese starter cultures as blends of strains with differential functionality during cheese production process. *L. lactis* accommodates three subspecies: *L. lactis* subsp. *lactis, L. lactis* subsp. *cremoris and L. lactis* subsp. *hordniae* with the first two having significant economic importance to dairy fermentations. The subspecies differentiation and nomenclature are traditionally based on phenotypes according to the taxonomic reclassification published by Schleifer et al.(1). The so-called *lactis* phenotype is differentiated from the *cremoris* phenotype by its growth at 40°C, resistance to 4% salt and utilization of arginine and maltose (1, 2). *L. lactis* subsp. *lactis* also includes the biovariety *diacteylactis* that is a citrate fermenting variant of the *lactis* phenotype and the corresponding metabolic pathway is generally encoded on a plasmid (3, 4).

Phenotypes are of primary importance for screening of potential starter cultures because of their postulated linkage to the performance in the fermentation. However, molecular characterization of field isolates often does not match with their subspecies designation based on the phenotypes (5-8). Fingerprinting studies reveal two main genomic lineages: *L. lactis* subsp. *cremoris* genotype (a genotype similar to that of the *L. lactis* subsp. *cremoris* type strain SK11) and *L. lactis* subsp. *lactis* genotype (a genotype similar to that of the *L. lactis* subsp. *lactis* type strain IL1403). The *lactis*-genotype lineage comprises mainly *lactis*-phenotype strains, but the *cremoris*-genotype lineage often harbors both *lactis*- and *cremoris*-phenotypes, hampering the development of genetic markers for a certain phenotype or trait. The discrepant strains have a *cremoris*-genotype while they display a *lactis*-phenotype (9-13) and MG1363 *L. lactis* subsp. *cremoris* lab strain is the well-known representative for these frequently encountered diverging strains.

Current genomic fingerprinting studies are far from the resolution level required to reveal a phenotype-specific genetic signature. High-resolution fingerprinting, optimized for the species of interest, may deliver phenotype-specific genetic markers that could be employed as an effective alternative for the labor-intensive phenotypic screens. Additionally, the markers generated for a particular strain can be used as a signature to determine their presence and abundance in the commercial mixed culture fermentations (14-16) and/or for source tracking (17). AFLP is a fingerprinting technique that offers high reproducibility, superior resolution and the highest correlation to DNA-DNA hybridizations, although Rep-PCR and randomly amplified polymorphic DNA analysis (RAPD) are more frequently applied (10, 18-21).

In this chapter, we aimed to develop an optimized, high-throughput and high-resolution AFLP methodology for *L. lactis* strains derived from different ecological niches around the world. High-throughput was implemented by the use of fluorophore labels in combination with capillary electrophoresis. The large genetic diversity within the 82-strain collection used in this study allowed true validation of the resolution power and the species-wide applicability of the method. Moreover, AFLP fingerprinting of strains with discrepant genotype-phenotype characteristics enabled the identification of *cremoris* phenotype specific genetic markers, for which the robustness was verified for many representative strains next to the well-known representative *L. lactis* subsp. *cremoris* reference strain MG1363.

Materials and Methods

Bacterial Strains

A collection of 82 *L. lactis* strains of diverse geographical origin and isolation source was used in this study. The designations of the strains, their origin, subspecies and genotype information are provided in supplementary Table S1. Strains were routinely grown in M17 broth (Oxoid Ltd.) supplemented with 0.5 % (wt/vol) glucose (GM17).

Chromosomal DNA Isolation

Bacterial chromosomal DNA was isolated with QIAgen Tissue and Culture DNA Isolation Kit (QIAgen GmbH) according to the manufacturer`s instruction. Chromosomal DNA quality and concentrations were determined by Nanodrop (Coleman technologies Inc.) and 0.8% (wt/vol) agarose gels.

AFLP Prescreening Using Radioactive Detection

Enzyme-primer combinations likely to allow high-resolution fingerprinting were selected using the web-based Simulation of Double Digestion Fingerprinting Techniques tool (22). The published *L. lactis* genomes of strains SK11(23), IL1403(24), MG1363(15) were double-digested using combinations of HindIII *and* EcoRI as rare cutters, and TaqI and MseI as frequent cutters. Selective PCR was simulated adding single 3'-end selective nucleotides for each primer. Enzyme-primer combinations producing 60 to 100 well distributed fragments *in silico* were applied during

experimental prescreening on *L. lactis* subsp. *cremoris* NIZO B32, *L. lactis* subsp. *lactis* NIZO B644 and NIZO B2211 and *L. lactis* subsp. *lactis* biovar. *diacetylactis* NIZO B1592 according to the protocol of Vos *et al* (25) (Figure S1 in supplementary material).

AFLP Methodology Using Fluorophore Detection and Capillary Electrophoresis

Adaptor annealing, restriction and ligation reactions were based on the AFLP protocol by Meudt *et al* (15). 250ng of chromosomal DNA was used for the template preparations. Adaptors and selective primers (Table 1) were synthesized by Sigma-Aldrich (Zwijndrecht, The Netherlands) and HindIII selective primer was end-labeled with carboxyfluorescein fluorophore (FAM). AFLP reactions were optimized based on the protocol of Vos *et al* (25) and Meudt *et al* (15) with the following modifications. Pre-selective PCR was completely omitted from the protocol. Selective PCR was performed using 1 unit of Taq DNA Polymerase (Roche, Mannheim, Germany) in 25 µl volume containing 1 µM of TaqI and HindIII selective primers, 200 µM dNTP (of each nucleotide), and 5 µl ligation products (15 times diluted). PCR amplification using a Westburg Biometria T1 cycler was initiated by 2 min of nick-filling at 72°C, followed by 30 cycles of 30 s denaturation at 94°C, 60 s annealing at 50°C and 60 s extension at 72°C and was completed with a 10 min final extension step at 72°C. Ramping speed was fixed at 5 °C/s. 1 µl PCR products (3-times diluted) was mixed with 9 µl ET550-R size standard (36 times diluted; GE Healthcare, Little Chalfont Buckinghamshire, UK) and analyzed on a MegaBACE 500 automated DNA platform (GE Healthcare, Diegem, Belgium) according to the manufacturer's instructions.

Table 1. Adaptors and primers

Primer Name	Sequence	Label
HindIII Linker 1	5`-CTCGTAGACTGCGTACC- 3`	No label
HindIII Linker 2	5`-AGCTGGTACGCAGTC- 3`	No label
HindIII Core Primer	5`- GACTGCGTACCAGCTT-3`	No label
HindIII Selective Primer **(H1)**	5`-GACTGCGTACCAGCTT**A**- 3	FAM
195bp HindIII Selective Primer	5`-GACTGCGTACCAGCTT**ATCAA**-3`	FAM
TaqI Linker 1	5`-CGGTCAGGACTCAT- 3`	No label
TaqI Linker 2	5`-GACGATGAGTCCTGAC -3`	No label
TaqI Core Primer	5`-CGATGAGTCCTGACCGA- 3`	No label
TaqI Selective Primer **(T1)**	5`-GATGAGTCCTGACCGA**A**- 3`	No label
195 bp TaqI Selective Primer	5`-GATGAGTCCTGACCGA**ATATC**-3`	No label
195bp normal PCR forward primer	5`-CCGCCAGATTCGAATATC- 3`	No label
195bp normal PCR reverse primer	5`-CAAATGGAAGCTTATCAA- 3`	No label

AFLP Profile Analysis

Data were analyzed using the BioNumerics 4.5 software suite (Applied Maths, Sint Martens-Latem, Belgium). Pairwise similarities were calculated using the Pearson correlation coefficient and cluster analyses were performed applying the unweighted-pair group method with arithmetic averages algorithm (UPGMA) for the range of 80-400 bps within the AFLP profiles.

Identification of Subspecies Phenotype Specific Markers on Fingerprints

The candidate phenotype-markers unique for *cremoris*- or *lactis*-phenotypes were searched within the AFLP patterns and traced back on the reference genomes with web-based Simulation of Double Digestion Fingerprinting Techniques tool (22). The original length of the digests was calculated by subtracting the number of nucleotides coming from adaptor sequences. The sequences of the fragments were exported from the sequence link available in web-based Simulation of Double Digestion Fingerprinting Techniques tool (22). The position of the candidate markers and probable mutations on reference genomes (strains MG1363, SK11 and IL1403) were checked by BLAST (http://www.ncbi.nlm.nih.gov/blast/) analysis in NCBI nr-nucleotide database. The gene specific PCR and AFLP PCR primers were designed according to sequences obtained from the BLAST hits. The AFLP primers were extended with 3 more selective nucleotides from each end to increase their selectivity for the marker fragment sequence (Table 1) and the presence of marker fragments was confirmed by gene specific PCR and refined AFLP amplifications.

Phenotypic Diversity Analysis

Phenotypic differentiation of lactococcal subspecies and bio-varieties was achieved using the previously established phenotypic analysis, including 4% (wt/vol) NaCl resistance, citrate utilization, arginine hydrolysis and maltose utilization (11). In addition to these standard tests for subspecies differentiation, several other phenotypic datasets including enzyme activities, sugar fermentation capacity profiles, polysaccharide utilization profiles, antibiotics, metal ion and nisin sensitivity have been used. The details of the phenotypic testing protocols are given in the supplementary material. The determination of phenotypic diversity among strains was facilitated by UPGMA cluster analysis based on Canberra metric coefficient, separately for each dataset. The congruence of groupings based on separate phenotypic data sets and the AFLP data was calculated using Pearson correlation.

Results

Methodology Implementation and Reproducibility Analysis

24 enzyme-primer combinations were selected based on *in silico* analysis and verified experimentally using the classical AFLP approach that makes use of radioactive labeled primers and slab-gel analyses (25) (Supplementary material, Figure S1). The HindIII(A)-TaqI(A) combination revealed the highest number of polymorphisms, including 19 AFLP markers discriminating between the reference strains tested (see M&M for details). This combination was subsequently used for genomic diversity analysis of the natural isolates using fluorophore labeled primers (see Supplementary material for details). Following optimization, the eventual standard-procedure involved a relatively straightforward protocol that does not require pre-amplification, touchdown PCR or slow-ramping PCR procedures. The discriminatory power of AFLP fingerprinting is not only determined by the number of polymorphisms generated but also by the level of reproducibility. To evaluate reproducibility of the AFLP protocol, independent replicates were generated for the 4 reference *L. lactis* strains, 3 of which were used for *in silico* simulations. The variation as a consequence of genome preparations was not significant within a single experiment, but the reproducibility was predominantly affected by the variations between independent AFLP experiments, resulting in distinct clusters for strain MG1363 and strain DRA4 (Figure 1). Based on this analysis, 90% similarity was taken as a threshold for the separation of individual strains.

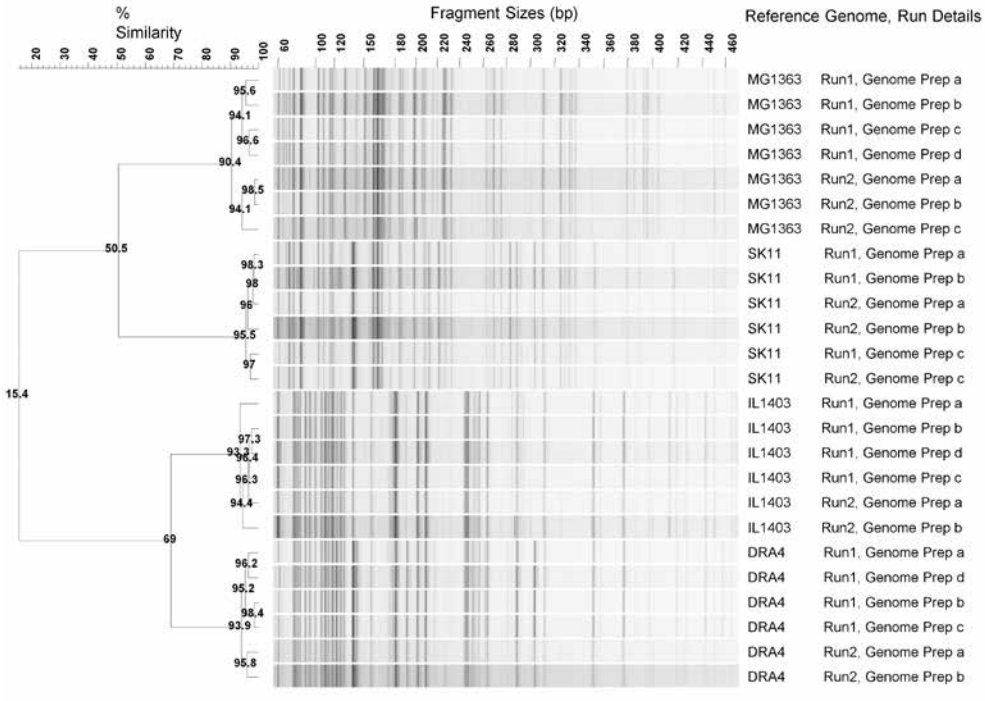


Figure 1. Reproducibility evaluation for the reference strains of *Lactococcus lactis.* Run1 and 2 correspond to independent AFLP experiments and genome preparations a, b, c, and d correspond to independent genome preparations of the same strain. UPGMA clustering was based on Pearson Correlation Coefficient on the range of (80-400) bps.

Diversity Analysis of Field Isolates

AFLP profiles were generated for 82 dairy and non-dairy isolates of *L. lactis*. Cluster analysis revealed two main AFLP-clusters with the resolution of 15% similarity for *lactis* and *cremoris* genotype discrimination (Figure 2). The AFLP profiles of *lactis*-genotype cluster were further separated into two distinct clusters. The cluster 1B contained the majority of the subspecies *lactis* strains and all of the biovar *diacetylactis* strains.

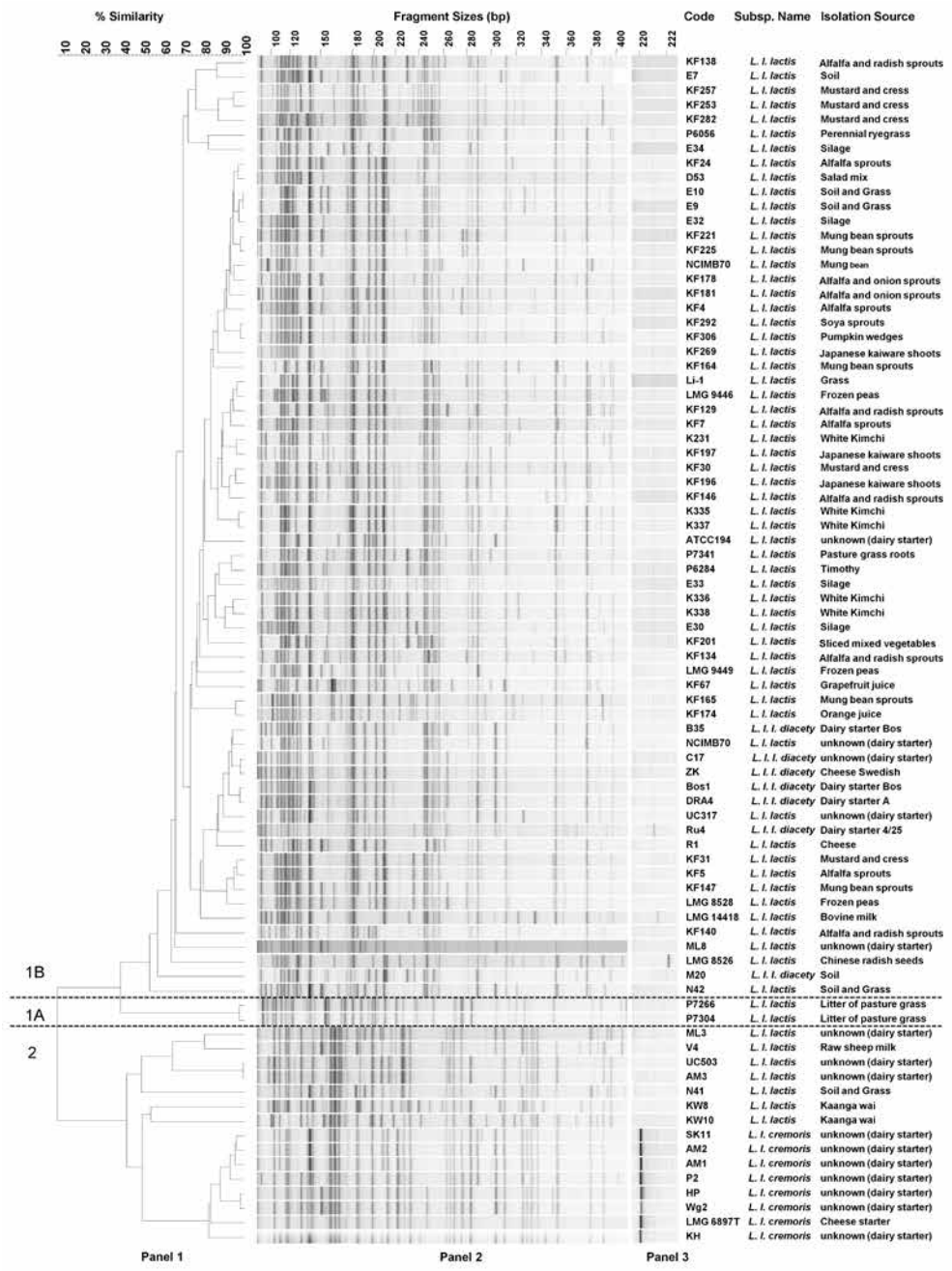


Figure 2. AFLP on *L. lactis* natural isolates 1st Panel: UPGMA Clustering of AFLP Profiles with Pearson Correlation; 2nd Panel: AFLP profiles of isolates, 80-400 bps; 3rd Panel: 221 bp marker (corresponding to 195bp digest) selectively isolated by the addition of 3 more nucleotides to the end of original HindIII (A)-TaqI (A) primers.

Notably, all dairy associated biovar. *diacetylactis* strains clustered together, but were distinguished from strain M20 that was isolated from soil. The smaller cluster 1A contained indistinguishable AFLP profiles of two strains (*L. lactis* subsp. *lactis* NIZO B2207 and NIZO B2208) isolated from litter of pasture grass (Figure 2). This cluster 1A isolates were previously proposed to be members of a third *L. lactis* genomic lineage (11). It is characterized by a *lactis*-phenotype but displays neither a *lactis*- nor a *cremoris*-genotype based on 16S rRNA gene sequence and MLSA analysis. The AFLP analysis enables their distinction as a unique genotype at a high resolution. The *cremoris*-genotype cluster 2 included all strains displaying typical *cremoris*-phenotype but also encompassed a subset of strains with a *lactis*-phenotype. The *cremoris*-phenotype strains grouped together with similarities exceeding 80% and formed a distinct sub-cluster, while the *lactis*-phenotype strains were too diverse to form a distinct sub-cluster.

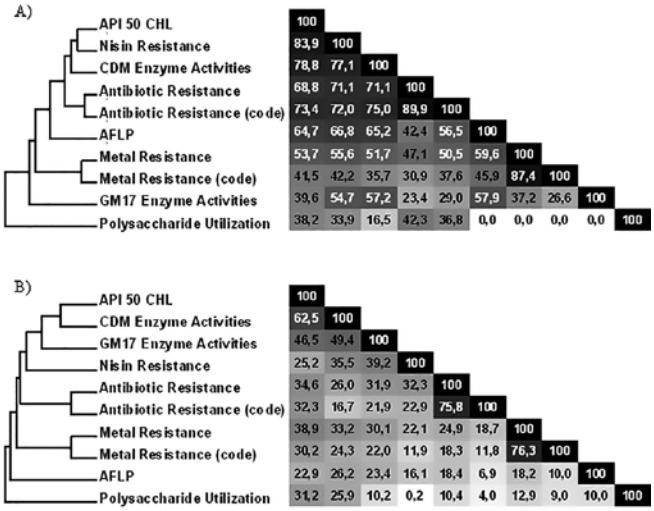


Figure 3. Correlations between phenotype variations and AFLP clusterings for *cremoris* (a) and *lactis* (b) genotype clusters.

The possible correlations between specific phenotypic traits and AFLP clustering were further investigated for *lactis*- and *cremoris*- genotype clusters (cluster 1 and 2) to evaluate whether strain discriminating phenotypes were reflected in the AFLP profiles. Within *cremoris*-genotype cluster, API 50 carbohydrate fermentations, nisin sensitivity and enzyme activity levels measured in CDM, correlated relatively well with the AFLP based clustering (Figure 3, panel A; Product Moment Correlations ranging from 60-70%). In contrast, other phenotypic traits like polysaccharide

utilization, enzyme activity levels in GM17, antibiotic and metal resistances, didn`t show significant correlation with the AFLP clustering. Notably, none of the phenotypic differences between strains within the *lactis*-genotype clusters (1A and 1B) appeared to correspond to AFLP profile differences (Figure 3, panel B).

AFLP Markers Discriminating *cremoris*- and *lactis*-phenotype

In addition to strain specific discrimination, the AFLP patterns generated in this study enabled the identification of marker fragments for the *cremoris*-phenotype. The pair of fragments between 212-214 bps and the triplet fragments between 220-223 bps were consistently observed in the AFLP profiles of *cremoris*-phenotype strains but were absent in the profiles generated for *lactis*-phenotype strains, irrespective of their genotype. The 221 bp (195bp when the number of nucleotides coming from adaptor sequences are subtracted) *cremoris*-phenotype specific fragment was selected for further validation. BLAST analysis indicated that the corresponding gene encodes for the LACR_2087, a hypothetical membrane protein (genebank accession number: YP_811677), that is present both in SK11 and MG1363 genomes, but absent in the IL1403 genome. The corresponding gene on MG1363 genome has a single nucleotide substitution that abolishes the Hind III restriction site, as well as some other substitutions affecting the selective nucleotides adjacent to this site, explaining the lack of this marker fragment in the MG1363 AFLP-fingerprint. The genomic region in which LACR_2087 is encoded appeared to be hyper-variable, with many gene deletions in MG1363 compared to SK11, such as the deletion of LACR 2089, a phage related hypothetical protein, LACR_2088, a cell wall associated hydrolase, and LACR_2085, a hypothetical protein that may be linked to the *cremoris*-phenotype. Moreover, the 82 isolates were screened both with gene specific PCR and with more stringent AFLP targeting of the 195 bp fragment. The gene specific PCR generated the anticipated product in all *cremoris*-phenotype strains, while no products were obtained for strains exhibiting a *lactis*-phenotype. High stringency AFLP PCR confirmed that the 221 bp fragment could consistently be generated in *cremoris*-phenotype strains, thereby proving that the 195 bp fragment indeed is responsible for the generation of the observed 221 bp marker in the AFLP analyses (Figure 3). Additional reference strains as well as novel isolates displaying either the *cremoris*- or *lactis*-phenotype were also tested for the robustness of these marker fragments. All novel *cremoris*-phenotype isolates produced the marker-fragments, while they were consistently absent for the *lactis*-phenotype strains (data not shown). These experiments establish the marker fragment identified here as an

appropriate and generic marker to distinguish *cremoris*- and *lactis*-phenotype strains at the genetic level.

Discussion

The present study delivers a high-resolution AFLP protocol for *Lactococcus lactis* that was adapted to a fluorophore-labeled procedure for high-throughput screenings. In contrast to many AFLP protocols described (25), the optimized procedure does not include pre-selection and touchdown-PCR procedure and employs fast-ramping cycles. The validity of the AFLP methodology developed was established on 82 *L. lactis* strains that has previously served as a benchmark strain collection in other strain-typing efforts, including 16S rRNA analysis, 5-loci MLSA and REP-PCR genomic fingerprint analysis (11). AFLP-based fingerprinting with a cut-off value of 15% similarity for *cremoris*- and *lactis*-genotype discrimination vastly exceeds the resolution power of REP-PCR, which has 40 % similarity cut-off for subspecies discrimination. Published AFLP studies for other species usually offer 60-90 % and 40-60 % for intra and interspecies variation, respectively, based on UPGMA clustering with Pearson correlation coefficient (14, 16, 18). The AFLP dendrogram also shared its clustering topology with those obtained from MLSA and 16S rRNA homology analysis (16). Recently, a subset of the strain collection used in this study was also subjected to a genome-scale diversity analysis by comparative genome hybridization (CGH) (26). The comparison of AFLP and CGH dendrograms showed similar clustering of the strains with a *lactis*-phenotype within the *cremoris*-genotype cluster and an analogous compact clustering for *lactis*-genotype strains, with outlying members of the third genomic lineage (the strains P7304 and P7266). These comparisons demonstrate the resolution capacity of the AFLP methodology, which enables strain specific typing with high reliability and biological relevance.

The commonly observed phenotype-genotype mismatching has directed many researchers to generate genetic markers linked to these so-called *cremoris*- and *lactis*-phenotypes. A method based on probing EcoRI digested chromosomal DNA with a *his* operon gene-probe enabled differentiation of *lactis*-phenotype strains with different genomic make-up (27). An elegant RFLP approach targeting the *gadB* gene was also able to determine both phenotype and genotype at the same time (28). However, their robustness remains to be confirmed since they were validated with only one discrepant strain (MG1363) (2, 4). *Cremoris*-phenotype specific marker

fragments generated in this study provide subspecies phenotype information within the fingerprint. Elaborating on this concept, markers that are representative for a particular phenotypic trait can be developed rationally. For example, a phenotypic trait related to a specific gene can be screened in the simulation tool with different enzyme-primer combinations for the generation of AFLP markers, which might be amplified for the strains carrying the functional gene of interest, and absent for the others. Such a strategy potentially provides effective genetic markers for screening of large strain collections for the phenotype of interest additional to the genomic diversity information.

The AFLP procedure presented in this chapter also provides a tool for the analysis of population dynamics in industrial fermentation processes that employ microbial consortia of closely related strains such as a starter culture (e.g. Gouda Cheese production). Identification of strain-specific markers analogous to the approach presented for the *cremoris*-phenotype may allow the design of strain-specific primers for tracing and quantification of the strains without re-isolation or viable count enumeration at different time points during fermentation, or under different processing conditions. It is a valuable tool to determine the number of genomic lineages present in isolated collections of strains which in turn justifies selection of representative strains for whole genome sequencing.

In conclusion, the AFLP protocol developed for *L. lactis* enables the delineation of closely related strains and its reproducibility and resolution exceed that of other currently available methods. Advanced analysis of the AFLP profiles in combination with *in silico* genome analysis can facilitate the recognition of genetic markers responsible for specific phenotypic traits. Furthermore, the high-throughput strain typing method using a genetic marker, which accurately predicts the phenotype, allows fast and predictive screening of culture collections.

Acknowledgements

We acknowledge Hanneke Witsenboer from Keygene B.V. for her support with the radio-labeling based AFLP screenings. Furthermore, we acknowledge Jumna Murat Bayjanov, Prof. Dr. Roland Siezen and Dr. Sacha van Hijum for kindly providing dendrograms produced for Comparative Genome Hybridization Data and also for their valuable comments. We also would like to acknowledge Dr. Herwig Bachmann

for sharing phenotype data and Jan Hoolwerf for technical assistance. Finally we would like to thank Patrick Janssen, Ingrid van Alen and Dr. Lucie Hazelwood for all the useful discussions and technical supports.

This work was funded by Top Institute Food and Nutrition (TIFN) and is part of Kluyver Centre for Genomics of Industrial Fermentation, a Centre of Excellence of The Netherlands Genomics Initiative.

References

1. **Schleifer KH, J. Kraus, C. Dvorak, R. Kilpper-Balz, M. D. Collins, and W. Fisher.** 1985. Transfer of *Streptococcus lactis* and related streptococci to the genus Lactococcus gen. nov. . Syst. Appl. Microbiol **6:**183-195.
2. **Garvie EI, Farrow JAE.** 1982. NOTES:*Streptococcus lactis* subsp. *cremoris* (Orla-Jensen) comb. nov. and *Streptococcus lactis* subsp. *diacetilactis* (Matuszewski et al.) nom. rev., comb. nov. Int J Syst Bacteriol **32:**453-455.
3. **Stiles ME, Holzapfel WH.** 1997. Lactic acid bacteria of foods and their current taxonomy. Int J Food Microbiol **36:**1-29.
4. **Hugenholtz J.** 1993. Citrate metabolism in lactic acid bacteria. FEMS Microbiology Reviews **12:**165-178.
5. **Urbach E, Daniels B, Salama MS, Sandine WE, Giovannoni SJ.** 1997. The ldh phylogeny for environmental isolates of *Lactococcus lactis* is consistent with rRNA genotypes but not with phenotypes. Appl Environ Microbiol **63:**694-702.
6. **de la Plaza M, Rodriguez A, Fernandez de Palencia P, Martinez-Cuesta MC, Pelaez C, Requena T.** 2006. Discrepancies between the phenotypic and genotypic characterization of *Lactococcus lactis* cheese isolates. Lett Appl Microbiol **43:**637-644.
7. **Pu ZY, Dobos M, Limsowtin GK, Powell IB.** 2002. Integrated polymerase chain reaction-based procedures for the detection and identification of species and subspecies of the Gram-positive bacterial genus Lactococcus. J Appl Microbiol **93:**353-361.
8. **Tanigawa K, Kawabata H, Watanabe K.** 2010. Identification and typing of *Lactococcus lactis* by matrix-assisted laser desorption ionization-time of flight mass spectrometry. Appl Environ Microbiol **76:**4055-4062.
9. **Kelly W, Ward L.** 2002. Genotypic vs. phenotypic biodiversity in *Lactococcus lactis*. Microbiology **148:**3332-3333.
10. **Nomura M, Kobayashi M, Narita T, Kimoto-Nira H, Okamoto T.** 2006. Phenotypic and molecular characterization of *Lactococcus lactis* from milk and plants. J Appl Microbiol **101:**396-405.
11. **Rademaker JL, Herbet H, Starrenburg MJ, Naser SM, Gevers D, Kelly WJ, Hugenholtz J, Swings J, van Hylckama Vlieg JE.** 2007. Diversity analysis of dairy and nondairy *Lactococcus lactis* isolates, using a novel multilocus sequence analysis scheme and (GTG)5-PCR fingerprinting. Appl Environ Microbiol **73:**7128-7137.
12. **Salama MS, Sandine WE, Giovannoni SJ.** 1993. Isolation of *Lactococcus lactis* subsp. *cremoris* from nature by colony hybridization with rRNA probes. Appl Environ Microbiol **59:**3941-3945.
13. **van Hylckama Vlieg JE, Rademaker JL, Bachmann H, Molenaar D, Kelly WJ, Siezen RJ.** 2006. Natural diversity and adaptive responses of *Lactococcus lactis*. Curr Opin Biotechnol **17:**183-190.
14. **Janssen P, Maquelin K, Coopman R, Tjernberg I, Bouvet P, Kersters K, Dijkshoorn L.** 1997. Discrimination of Acinetobacter genomic species by AFLP fingerprinting. Int J Syst Bacteriol **47:**1179-1187.
15. **Meudt HM, Bayly MJ.** 2008. Phylogeographic patterns in the Australasian genus Chionohebe (Veronica s.l., Plantaginaceae) based on AFLP and chloroplast DNA sequences. Mol Phylogenet Evol **47:**319-338.

16. **Savelkoul PH, Aarts HJ, de Haas J, Dijkshoorn L, Duim B, Otsen M, Rademaker JL, Schouls L, Lenstra JA.** 1999. Amplified-fragment length polymorphism analysis: the state of an art. J Clin Microbiol **37:**3083-3091.

17. **Scheirlinck I, Van der Meulen R, De Vuyst L, Vandamme P, Huys G.** 2009. Molecular source tracking of predominant lactic acid bacteria in traditional Belgian sourdoughs and their production environments. J Appl Microbiol **106:**1081-1092.

18. **Rademaker JL, Hoste B, Louws FJ, Kersters K, Swings J, Vauterin L, Vauterin P, de Bruijn FJ.** 2000. Comparison of AFLP and rep-PCR genomic fingerprinting with DNA-DNA homology studies: Xanthomonas as a model system. Int J Syst Evol Microbiol **50 Pt 2:**665-677.

19. **Foschino R, Nucera D, Volponi G, Picozzi C, Ortoffi M, Bottero MT.** 2008. Comparison of *Lactococcus garvieae* strains isolated in northern Italy from dairy products and fishes through molecular typing. J Appl Microbiol **105:**652-662.

20. **Ventura M, Zink R.** 2002. Specific identification and molecular typing analysis of *Lactobacillus johnsonii* by using PCR-based methods and pulsed-field gel electrophoresis. FEMS Microbiol Lett **217:**141-154.

21. **Thompson FL, Hoste B, Vandemeulebroecke K, Swings J.** 2001. Genomic diversity amongst *Vibrio* isolates from different sources determined by fluorescent amplified fragment length polymorphism. Syst Appl Microbiol **24:**520-538.

22. **San Millan R, Garaizar J, Bikandi J.** 2005. In silico simulation of fingerprinting techniques based on double endonuclease digestion of genomic DNA. In Silico Biol **5:**341-346.

23. **Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, Shakhova V, Grigoriev I, Lou Y, Rohksar D, Lucas S, Huang K, Goodstein DM, Hawkins T, Plengvidhya V, Welker D, Hughes J, Goh Y, Benson A, Baldwin K, Lee JH, Diaz-Muniz I, Dosti B, Smeianov V, Wechter W, Barabote R, Lorca G, Altermann E, Barrangou R, Ganesan B, Xie Y, Rawsthorne H, Tamir D, Parker C, Breidt F, Broadbent J, Hutkins R, O'Sullivan D, Steele J, Unlu G, Saier M, Klaenhammer T, Richardson P, Kozyavkin S, Weimer B, Mills D.** 2006. Comparative genomics of the lactic acid bacteria. Proc Natl Acad Sci U S A **103:**15611-15616.

24. **Clarke AC, Meudt, H. M.** 2005, posting date. AFLP (Amplified Fragment Length Polymorphism) for multi-locus genomic fingerprinting. [Online.]

25. **Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al.** 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res **23:**4407-4414.

26. **Siezen RJ, Bayjanov JR, Felis GE, van der Sijde MR, Starrenburg M, Molenaar D, Wels M, van Hijum SA, van Hylckama Vlieg JE.** 2011. Genome-scale diversity and niche adaptation analysis of *Lactococcus lactis* by comparative genome hybridization using multi-strain arrays. Microb Biotechnol **4:**383-402.

27. **Godon JJ, Delorme C, Ehrlich SD, Renault P.** 1992. Divergence of Genomic Sequences between *Lactococcus lactis* subsp. *lactis* and *Lactococcus lactis* subsp. *cremoris*. Appl Environ Microbiol **58:**4045-4047.

28. **Nomura M, Kobayashi M, Okamoto T.** 2002. Rapid PCR-based method which can determine both phenotype and genotype of *Lactococcus lactis* subspecies. Appl Environ Microbiol **68:**2209-2213.

Supplementary Information for Chapter 2

Optimizations for Selective Amplification

Optimizations aimed for the most appropriate PCR conditions for a flourophore labeled HindIII (A)-TaqI (A) primer combination on SK11, IL1403, MG1363 *L. lactis* reference genome preparations. Ligation reactions were diluted 25 times with nuclease free water and 5 μL of diluent was used as a template. Optimization parameters included touchdown and normal PCR cycles operated at two different ramping speeds (1 °C/s, 5 °C/s). In touchdown procedures, annealing times (30s, 15s), touchdown temperature decreases (0,5**°C,** 1°C) and cycle numbers (10, 20) were investigated. Normal PCR cycles included trials with different annealing temperatures over a range of (65°C - 45°C) and annealing times of 1min-30s by gradient PCR. Reactions were also optimized in terms of Mg^{2+}, polymerase and primer concentrations. The optimized PCR protocol was used in further AFLP analysis of *L. lactis* isolates.

AFLP is based on selective amplification of restriction products coupled to adaptors and selectivity is achieved by the addition of selective nucleotides to the end of primers which are complementary to adaptor sequences and restriction enzyme sites. Preselection by core primers without any selective base is commonly applied to increase the number of template fragments in the pool before the selective amplification step. This prevents the domination of hindering fragments generated by frequent cutters in complex genomes. Therefore, the first point studied in optimization was the added value of the preselection step. In our trials, addition of preselective PCR before selective PCR as successive steps led to the enrichment of certain fragments in the pool, while poorly amplified fragments in the first PCR were almost vanished in the second one. (Figure S2 in supplemantary material, lane a, b, c, d). High intensities coming from those fragments may interfere with data analysis since Pearson product moment correlation coefficient is not tolerant to outliers. However direct selective PCR resulted in even amplification of fragments. Hence, enrichment of certain fragments was prevented by applying direct selectice PCR on ligation products without preselection in further optimizations. Secondly, in the selective amplification step, touchdown PCR schemes are usually the method of choice for an efficient coverage of fragments in AFLP studies. However, too dense profiles were obtained with the touchdown protocol when compared to normal PCR cycles (Figure S2, lane e). Decreasing the annealing times from 30s to 15s in

touchdowns led to more clear profiles but with reduced reproducibility (Figure S2, lane f,g). Normal PCR cycles were preferred in further experiments. Another effect which had high impact on amplification was ramping speed of thermal cyclers. AFLP protocols in literature apply the reactions strickly with slow ramping, however fast ramping gave better amplifications for our FAM labelled primers (Figure S2, lane h). Finally, 50°C was found as an optimum annealing temperature and no significant difference was observed between 30s and 1min annealing times in normal PCR trials (data not shown).

Phenotypic Testing

Strain-specific activity levels of several enzymes including branched chain aminotransferase (BcaT), hydroxyl-isocaproic-acid dehydrogenases (HicDH), aminopeptidase N (PepN), x-prolyl-dipeptidyl aminopeptidase (PepXP) and esterase have been reported previously (Bachmann et al., Appl. Environ. Microbiol. **75**:5687-5694, 2009) and determined in both glucose supplemented M17 broth or chemically defined medium (CDM) (Bachmann et al., J. Dairy Sci. **92:**5868-5882, 2009). Additional phenotyping of individual strains included API 50 CHL assays (BioMérieux, Marcy l`Etoile, France) to determine sugar fermentation capacity profiles, while polysaccharide utilization capacities of the strains were evaluated using M17 medium supplemented with 0.5 or 1% (wt/vol) of the polysaccharide of interest. Antibiotic, metal ion and nisin sensitivity levels of strains were determined by growth response in GM17, supplemented either with metal ions, antibiotics, or nisin. The level of stress was varied within a concentration range determined depending on the antibiotics, nisin or metal salt used. The antibiotics evaluated included chloramphenicol, erythromycin, tetracycline, vancomycin, eciprofloxacin, neomycine, penicillin G, trimethoprim, while metal ions analyzed included Cd^{2+}, Hg^{2+}, Cr^{3+}, Cu^{2+}, Mn^{2+}, Pb^{2+} as Cl^{-1} salts. Wells were inoculated 1% (vol/vol) with GM17 active cultures and growth was assessed by turbidity measurement at 600nm after overnight incubation at 30°C. Growth was scored as negative (0), weak (1) and positive (2) when it was lower than 25%, between 25% -75% and over 75% of the turbidity obtained from growth without any antibiotic, metal or nisin stress, respectively. Results for all phenotypic tests that were performed are given in Supplementary tables S2 to S7 (see below).

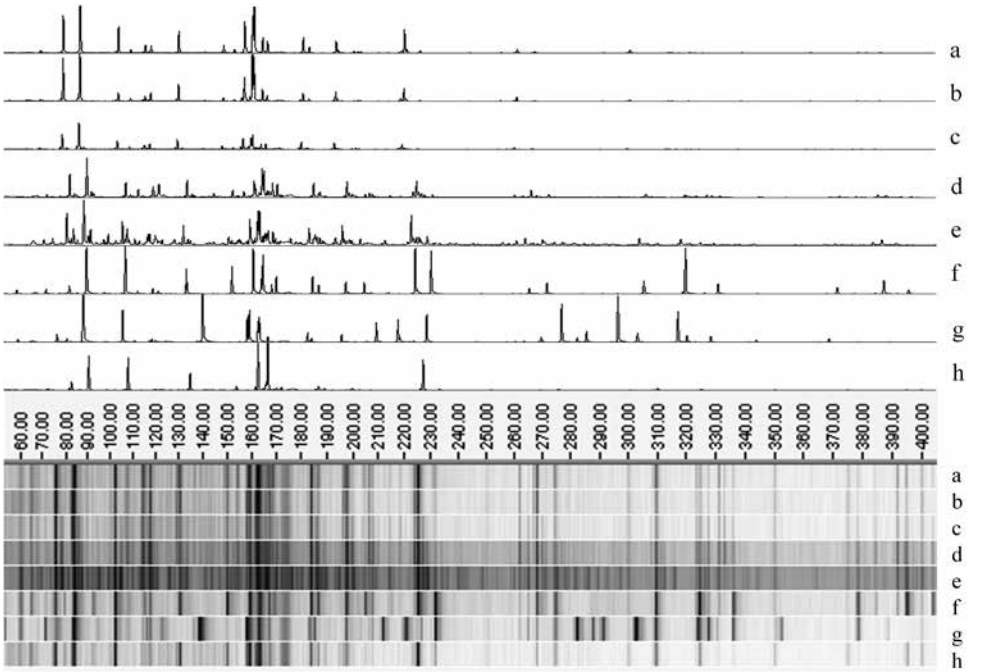


Figure S2. Optimization of Selective PCR Conditions on MG1363 *L. lactis* ssp. *cremoris* genome. Profiles are represented both as chromatograms and bands. Lane a, Normal PCR on preselection products with fast ramping Replicate 1; Lane b, Normal PCR on preselection products with fast ramping Replicate 2; Lane c, Normal PCR on ligation products with fast ramping Replicate 1; Lane d, Normal PCR on ligation products with fast ramping Replicate 2; Lane e, Touchdown PCR, 30 s annealing with fast ramping; Lane f, Touchdown PCR, 15 s annealing with fast ramping Replicate 1; Lane g, Touchdown PCR, 15 s annealing with fast ramping Replicate 2; Lane h, Touchdown PCR, 30s annealing with slow ramping.

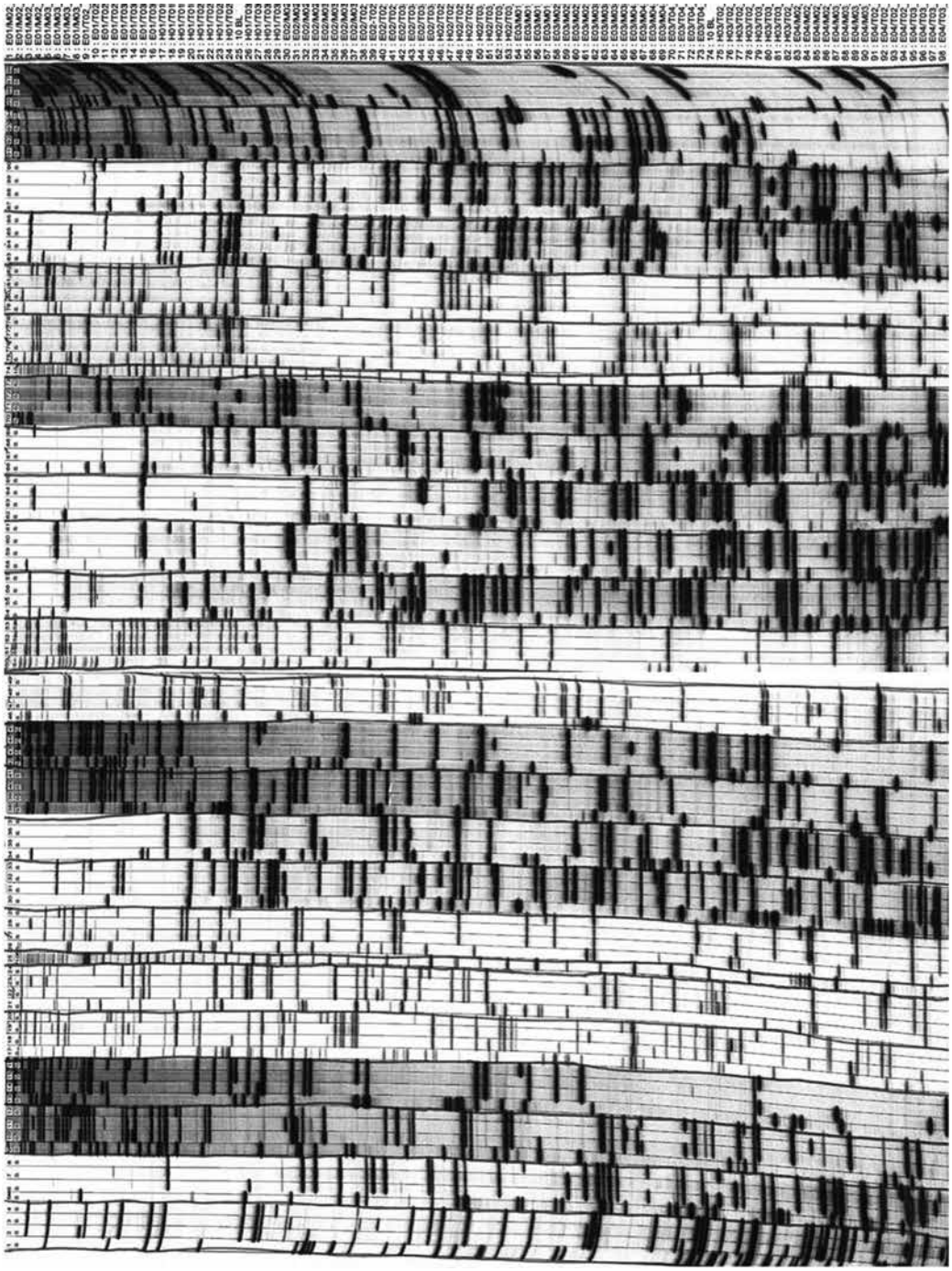


Figure S1: Presecreening on *L. lactis* strains with the strain order of NIZO B32 *L. lactis* ssp. *cremoris*, NIZO B644 *L. lactis* ssp. *lactis,* NIZO B2211 *L. lactis* ssp. *lactis*, and NIZO B1592 *L. lactis* ssp. *lactis* biovar. *diacetylactis*.. The column on the right side shows the order of the lanes. The abbreviations E, M, H, T indicate EcoRI, MseI, HindIII, and TaqI enzymes respectively. The numbers 1,2,3,4 correspond to selective nucleotides A, C, G, T respectively. Marker lanes are indicated with BL.

Table S1. Data for Strains Used in the Study

No	NIZO Number	Other Designations	Origin	Subspecies	Genotype
1	NIZO34	P2	unknown (dairy starter)	*L. l. cremoris*	*L. l. cremoris*
2	NIZO42	HP	unknown (dairy starter)	*L. l. cremoris*	*L. l. cremoris*
3	NIZO49	KH	unknown (dairy starter)	*L. l. cremoris*	*L. l. cremoris*
4	NIZO32	SK11	unknown (dairy starter)	*L. l. cremoris*	*L. l. cremoris*
5	NIZO33	AM2	unknown (dairy starter)	*L. l. cremoris*	*L. l. cremoris*
6	NIZO65	Wg2	unknown (dairy starter)	*L. l. cremoris*	*L. l. cremoris*
7	NIZO2418T	LMG 6897T	Cheese starter	*L. l. cremoris*	*L. l. cremoris*
8	NIZO48	AM1	unknown (dairy starter)	*L. l. cremoris*	*L. l. cremoris*
9	NIZO643	ML3	unknown (dairy starter)	*L. l. lactis*	*L. l. cremoris*
10	NIZO1157	V4	Raw sheep milk	*L. l. lactis*	*L. l. cremoris*
11	NIZO2250	UC503	unknown (dairy starter)	*L. l. lactis*	*L. l. cremoris*
12	NIZO2251	AM3	unknown (dairy starter)	*L. l. lactis*	*L. l. cremoris*
13	NIZO1175	N41	Soil and Grass	*L. l. lactis*	*L. l. cremoris*
14	NIZO2248	KW8	Kaanga wai	*L. l. lactis*	*L. l. cremoris*
15	NIZO2249	KW10	Kaanga wai	*L. l. lactis*	*L. l. cremoris*
16	NIZO2192	P7341	Pasture grass roots	*L. l. lactis*	*L. l. lactis*
17	NIZO2219	KF7	Alfalfa sprouts	*L. l. lactis*	*L. l. lactis*
18	NIZO2225	KF129	Alfalfa,radish sprouts	*L. l. lactis*	*L. l. lactis*
19	NIZO5	R1	Cheese	*L. l. lactis*	*L. l. lactis*
20	NIZO2124	LMG 9449	Frozen peas	*L. l. lactis*	*L. l. lactis*
21	NIZO2220	KF24	Alfalfa sprouts	*L. l. lactis*	*L. l. lactis*
22	NIZO2226	KF134	Alfalfa,radish sprouts	*L. l. lactis*	*L. l. lactis*
23	NIZO20	ML8	unknown (dairy starter)	*L. l. lactis*	*L. l. lactis*
24	NIZO2142	NCIMB702727	Mung bean	*L. l. lactis*	*L. l. lactis*
25	NIZO2206	P7266	Litter of pasture grass	*L. l. lactis*	*L. l. lactis*
26	NIZO2227	KF138	Alfalfa,radish sprouts	*L. l. lactis*	*L. l. lactis*
27	NIZO29T	ATCC19435T	unknown (dairy starter)	*L. l. lactis*	*L. l. lactis*
28	NIZO2211	NCIMB700895	unknown (dairy starter)	*L. l. lactis*	*L. l. lactis*
29	NIZO2207	P7304	Litter of pasture grass	*L. l. lactis*	*L. l. lactis*
30	NIZO2221	KF30	Mustard and cress	*L. l. lactis*	*L. l. lactis*
31	NIZO2228	KF140	Alfalfa,radish sprouts	*L. l. lactis*	*L. l. lactis*
32	NIZO644	UC317	unknown (dairy starter)	*L. l. lactis*	*L. l. lactis*
33	NIZO2190	P6056	Perennial ryegrass	*L. l. lactis*	*L. l. lactis*
34	NIZO2217	KF4	Alfalfa sprouts	*L. l. lactis*	*L. l. lactis*
35	NIZO2222	KF31	Mustard and cress	*L. l. lactis*	*L. l. lactis*
36	NIZO2229	KF146	Alfalfa,radish sprouts	*L. l. lactis*	*L. l. lactis*
37	NIZO2191	P6284	Timothy	*L. l. lactis*	*L. l. lactis*
38	NIZO2218	KF5	Alfalfa sprouts	*L. l. lactis*	*L. l. lactis*
39	NIZO2223	KF67	Grapefruit juice	*L. l. lactis*	*L. l. lactis*
40	NIZO2230	KF147	Mung bean sprouts	*L. l. lactis*	*L. l. lactis*
41	NIZO1171	E30	Silage	*L. l. lactis*	*L. l. lactis*
42	NIZO1236	E10	Soil and Grass	*L. l. lactis*	*L. l. lactis*
43	NIZO2198	D53	Salad mix	*L. l. lactis*	*L. l. lactis*
44	NIZO2236	KF196	Japanese kaiware shoots	*L. l. lactis*	*L. l. lactis*
45	NIZO2242	KF257	Mustard and cress	*L. l. lactis*	*L. l. lactis*
46	NIZO26	LMG 8526	Chinese radish seeds	*L. l. lactis*	*L. l. lactis*
47	NIZO1172	E33	Silage	*L. l. lactis*	*L. l. lactis*
48	NIZO2199	K231	White Kimchi	*L. l. lactis*	*L. l. lactis*
49	NIZO2231	KF164	Mung bean sprouts	*L. l. lactis*	*L. l. lactis*
50	NIZO2237	KF197	Japanese kaiware shoots	*L. l. lactis*	*L. l. lactis*

51	NIZO2243	KF269	Japanese kaiware shoots	*L. l. lactis*	*L. l. lactis*
52	NIZO1173	E34	Silage	*L. l. lactis*	*L. l. lactis*
53	NIZO2200	K335	White Kimchi	*L. l. lactis*	*L. l. lactis*
54	NIZO2232	KF165	Mung bean sprouts	*L. l. lactis*	*L. l. lactis*
55	NIZO2238	KF201	Sliced mixed vegetables	*L. l. lactis*	*L. l. lactis*
56	NIZO1174	E32	Silage	*L. l. lactis*	*L. l. lactis*
57	NIZO1240	E9	Soil and Grass	*L. l. lactis*	*L. l. lactis*
58	NIZO2201	K336	White Kimchi	*L. l. lactis*	*L. l. lactis*
59	NIZO2233	KF174	Orange juice	*L. l. lactis*	*L. l. lactis*
60	NIZO2239	KF221	Mung bean sprouts	*L. l. lactis*	*L. l. lactis*
61	NIZO2244	KF282	Mustard and cress	*L. l. lactis*	*L. l. lactis*
62	NIZO1154	E7	Soil	*L. l. lactis*	*L. l. lactis*
63	NIZO2122	LMG 8528	Frozen peas	*L. l. lactis*	*L. l. lactis*
64	NIZO2202	K337	White Kimchi	*L. l. lactis*	*L. l. lactis*
65	NIZO2234	KF178	Alfalfa and onion sprouts	*L. l. lactis*	*L. l. lactis*
66	NIZO2240	KF225	Mung bean sprouts	*L. l. lactis*	*L. l. lactis*
67	NIZO2245	KF292	Soya sprouts	*L. l. lactis*	*L. l. lactis*
68	NIZO1156	Li-1	Grass	*L. l. lactis*	*L. l. lactis*
69	NIZO1230	N42	Soil and Grass	*L. l. lactis*	*L. l. lactis*
70	NIZO2123	LMG 9446	Frozen peas	*L. l. lactis*	*L. l. lactis*
71	NIZO2203	K338	White Kimchi	*L. l. lactis*	*L. l. lactis*
72	NIZO2235	KF181	Alfalfa and onion sprouts	*L. l. lactis*	*L. l. lactis*
73	NIZO2241	KF253	Mustard and cress	*L. l. lactis*	*L. l. lactis*
74	NIZO2246	KF306	Pumpkin wedges	*L. l. lactis*	*L. l. lactis*
75	NIZO2424	LMG 14418	Bovine milk	*L. l. lactis*	*L. l. lactis*
76	NIZO82	Bos1	Dairy starter Bos	*L. l. lactis diacetylactis*	*L. l. lactis*
77	NIZO86	Ru4	Dairy starter 4/25	*L. l. lactis diacetylactis*	*L. l. lactis*
78	NIZO1594	C17	unknown (dairy starter)	*L. l. lactis diacetylactis*	*L. l. lactis*
79	NIZO1929	B35	Dairy starter Bos	*L. l. lactis diacetylactis*	*L. l. lactis*
80	NIZO87	ZK	Cheese Swedish	*L. l. lactis diacetylactis*	*L. l. lactis*
81	NIZO844	M20	Soil	*L. l. lactis diacetylactis*	*L. l. lactis*
82	NIZO1592	DRA4	Dairy starter A	*L. l. lactis diacetylactis*	*L. l. lactis*

The tables: S2, S3, S4, S5, S6 and S7 that encompass the data for carbohydrate fermentation profiles, antibiotic resistance profiles, metal resistance, GM17 and CDM enzyme activities, polysaccharide utilization and nisin sensitivity of the *L. lactis* isolates respectively are not provided in the thesis due to extensive information, which is not readily presented in a concise table format. These supplementary tables can be obtained on the open access link (http://aem.asm.org/content/77/15/5192/suppl/DC1), provided by the Journal of Applied and Environmental Microbiology for the supplementary material of the publication.

Chapter 3

Multifactorial Diversity Sustains Microbial Community Stability

Oylum Erkuş, Victor C.L. de Jager, Maciej Spus, Ingrid J. van Alen-Boerrigter, Irma Van Rijswijck, Lucie Hazelwood, Patrick W.M. Janssen, Sacha A.F.T. van Hijum, Michiel Kleerebezem & Eddy J. Smid

This chapter has been published as The ISME Journal; 2013 November; 7:2126-36.

Abstract

Maintenance of a high degree of biodiversity in homogeneous environments is poorly understood. A complex cheese starter culture with a long history of use was characterized as a model system to study simple microbial communities in this chapter. Eight distinct genetic lineages were identified, encompassing two species: *Lactococcus lactis* and *Leuconostoc mesenteroides.* The genetic lineages were found to be collections of strains with variable plasmid content and phage sensitivities. Kill-the-winner hypothesis explaining the suppression of the fittest strains by density dependent phage predation was operational at the strain level. This prevents the eradication of entire genetic lineages from the community during propagation regimes (back-slopping), stabilizing the genetic heterogeneity in the starter culture against environmental uncertainty.

Introduction

Complex microbial communities exist everywhere in nature. Nutrient rich and dynamic environments such as the mammalian gastro-intestinal tract support complex microbial communities in terms of species diversity (1). On the other hand, nutrient poor or extreme environments like high salinity ponds contain microbial communities with a lower species complexity (2). Interestingly, the latter microbial communities still display a high degree of intraspecies diversity (micro-diversity). Population heterogeneity is thought to be linked with resilience against environmental uncertainty (3), and mobile genetic elements are considered to be the primary sources of heterogeneity (4). Intraspecies diversity among closely related strains is commonly linked to functionally adaptive traits encoded on genomic islands (5) that are acquired by horizontal gene transfer. The generation of subpopulations with varying plasmid content in natural communities brings selective advantages in the face of environmental uncertainty (6, 7). Analogously, bacteriophages play a regulatory role in population dynamics through density dependent predation (8). "Kill the winner" model predicts that an increase in a particular strain of a microbial community leads to an increase of a viral population that prays on that specific strain so that the population size of rapidly growing strain will be reduced and eventually it will secure the heterogeneity of the total community (9).

The inherent resilience (10) of simple microbial communities towards environmental fluctuations makes such adaptive clonal populations also interesting for biotechnological applications such as waste water treatment, beer, wine, and dairy fermentations. Many currently used dairy starter cultures are undefined and originate from farms that employed artisanal back-slopping practices (11). Starter culture communities may face several (a)biotic selective pressures during back-slopping regimes. Despite this environmental uncertainty, mixed-strain undefined cultures are more resilient and display a more robust performance as compared to defined low-strain-diversity cultures in dairy processes (12).

This study presents an in-depth analysis of the genetic and functional diversity of an undefined complex cheese starter culture. Community heterogeneity was investigated at high-resolution, including the enumeration of individual genetic lineages, profiling of in-lineage heterogeneity for plasmid content and phage sensitivity, as well as community metagenomics and genome sequencing. The heterogeneity in plasmid content and phage predation were identified as the basis for the resilience of the culture as a whole.

Materials and Methods

The protocols for lab scale mini-cheese making, single colony isolations during cheese manufacturing and fingerprinting, plasmid isolations and profiling, phage isolations and sensitivity testing are provided in the supplementary information.

Sequencing and assembly of representative (TIFN) strains for each genetic lineage

One representative of each genetic lineage was selected randomly from the starter point isolates. Chromosomal DNA was isolated with the DNeasy Tissue and Culture DNA Isolation Kit (QIAgen, Hilden, Germany) according to manufacturer`s protocol and sequenced with Roche 454 Titanium (13) and Illumina mate-pair (14) sequencing technologies (insert size of 5 kb). 454 Titanium reads were assembled with Celera WGS version 6.1 assembler (15). The structural variations (SVs) in the form of homopolymer streches were determined by aligning the Illumina mate-pair reads against the pre-assembled contigs using RoVar (http://trac.nbic.nl/rovar), and the structural variations in the form of single nucleotide polymorphisms (SNPs) or small indels were corrected using a local Perl script only allowing the SVs with a maximum of one read mismatch. These pre-assembled contigs were further assembled (scaffolding) using the Illumina mate-pair data by SSPACE v1.1 (16), and the assemblies were subjected to gene function annotation using RAST (17). The genomes have been deposited at the NCBI database with the following bioproject ID's. Strain TIFN1: PRJNA175624; strain TIFN2: PRJNA175670; strain TIFN3: PRJNA175671; strain TIFN4; PRJNA175672; strain TIFN5: PRJNA175673; strain TIFN6: PRJNA175674; strain TIFN7: PRJNA175675; strain TIFN8: PRJNA175676. The statistics of TIFN genome sequences are given in Supplementary Table 1.

Core- and pan-genome analysis

RAST annotated protein coding genes of TIFN strains and protein coding gene sequences of publically available four reference strains MG1363 *L. lactis* ssp. *cremoris* (18), IL1403 *L. lactis* ssp. *lactis* (19), SK11 *L. lactis* ssp. *cremoris* (20), and *KF147 L. lactis* ssp. *lactis* (21) were subjected to orthology prediction using the orthoMCL algorithm V2.0.2 (22). Orthologous groups (OGs) were subsequently used for core- and pan-genome analyses. The core-genome was computed first for the set of TIFN1-7 *L. lactis* strains by stepwise increase of the number of genomes included to the computation. In the second part of the analysis, OGS from four *L. lactis* reference strains were

added to the initial TIFN1-7 set in the same way. At each step, the intersection of OGS were computed from all possible combinations of genomes for the number of genomes included at that step, resulting in a distribution of shared OGs. The mean, minimum, maximum and standard deviation of the obtained distributions at each step was depicted on the core-genome graph in order to see the core gene set reduction. In pan-genome analysis, the *L. lactis* reference strains and TIFN *L. lactis* strains were analyzed in a single set by stepwise increase of the number of genomes included to the computation. The cumulative numbers of non-redundant OGs were computed from all possible combinations of genomes for the number of genomes included at that step, and the mean, minimum, maximum and standard deviation of the resulting distribution was depicted on the pan-genome graph. Additionally, the pan-genome analysis was performed for the separate set of *L. lactis* reference strains for the comparison of TIFN pan-genome and reference pan-genome sets in order to find the contribution of TIFN strains to *L. lactis* pan-genome. The COG functional annotations of the OGs in core- and pan-genomes were assigned as described (23).

Sub-species assignments of TIFN strains were determined by PhyML (Guindon et al 2010) clustering of the core OGs that are present only as a single copy on the genome sequences of TIFN1-8 *L. lactis*, *Lc. mesenteroides*, as well as reference *L. lactis* ssp. *cremoris* (SK11, MG1363), and reference *L. lactis* ssp. *lactis* (IL1403, KF147) strains. The core OGs that are present as single copy on all the indicated genome sequences were aligned using MUSCLE (Edgar 2004) with the standard settings. The varying bases in the alignment were concatenated as a pseudo sequence for each strain, which were finally aligned and clustered by PhyML using standard settings.

The gene coverage of *Lactococcus* and *Leuconostoc* strains on the global metabolic pathway was determined by projecting the COG classes of OGs from TIFN1-7 *L. lactis* pan-genome and the TIFN8 *Lc. mesenteroides* genome separately on iPath v.2 (24) (http://pathways.embl.de/iPath2.cgi). The constructed metabolic maps were superimposed with Gimp (http://www.gimp.org) to visualize the complementing reactions of two species.

Strain-specific OGs Analysis and QPCR Primer Designs

Lineage-specific genetic markers were identified as follows. The protein sequences of all members in each strain-specific OGs were aligned to the reads of all TIFN genomes with TBLASTN v.2.2.26+ (25) using default parameters. Reads with

alignments of more than 99% identity over 90% of the read length within at least 66 nucleotides (The reads with a length shorter than 64 nucleotides were filtered out before the genome assembly.) were used for further analysis. The OGs of which the members only give hits to the reads of the expected genome were considered as candidate genetic marker loci. The OGs annotated with potentially phage or plasmid originated functions were excluded, and the remaining OGs were used for primer design with Primer Express software (Applied Biosystem, Warrington, UK), aiming for consistent annealing temperatures (60°C) and predicted amplicon lengths (between 60-70 bps) to obtain similar PCR conditions for each detection reactions. 6 primer sets were designed targeting genetic markers specific for TIFN(1&5), TIFN(2&4), TIFN3, TIFN6, TIFN7, and TIFN8, and biologically validated on chromosomal and plasmid DNA preparations of the TIFN strains. Additionally, the conservation of genetic markers for strains clustered within a genetic lineage was validated with QPCR for all starter culture isolates.

Metagenome Sequencing and Analysis

Metagenome analysis was performed on total DNA isolated from the starter culture (see supplementary methods for details), using Roche 454 titanium sequencing technology (13). The reads were filtered by the sff2CA tool of Celera WGS assembler (version 6.1) with the following command (sffToCA -libraryname <LIBNAME> -clear 454 -clear discard-n -trim hard -linker titanium -output <DATADIR>). The obtained fragment files (.frg) and log files provided information on the quality of the reads with the former converted into fastq or fasta files for further analyses. The statistics of the metagenome sequences are given in Supplementary Table 2. In order to determine the taxonomic composition, 16S rRNA derived reads were classified by mapping all the metagenome sequence reads to 16S rRNA sequences of the Ribosomal Data Project database (RDP 10.28) using the RDP Classifier v2.2. The metagenome sequence reads were compared to all TIFN genome scaffolds using BLASTN with a 95% sequence identity, 32bp word size, and 90% length cut-off as filtering parameters. The percentage coverage of the metagenome with the TIFN strains was determined by dividing the number of reads mapped to all TIFN strains by the total number of reads in the metagenome. The taxonomic origin of the reads that could not be mapped on the genomes of the TIFN strains was predicted by comparing the unmapped metagenome reads to the NCBI non-redundant protein sequence database (March, 2012) with BLASTP v2.2.26+ using the default settings, and classified using MEGAN v4.63.1 (26). The percentage of phage related sequences in the unmapped metagenome reads was

defined as the ratio of the number of reads classified into the phage related sequences in MEGAN output divided by the total number of unmapped metagenome reads.

Relative Quantification of Strain-specific OGs in the Metagenome

In order to see if metagenome datasets can be used for the quantification of the community composition to enable the comparison with culture dependent analyses, TIFN genome prevalence in the metagenome was quantified based on the strain-specific OGs. The strain specific OGs, of which the members only give hits to the reads of the expected genome at the nucleotide level were aligned with BLAST (27) to the starter metagenome reads to count the number of hits. The matching criteria were as follows: 1) more than 99% identity over 90% of the read length was considered as cut-off when the length of the read is smaller than the length of the target protein, 2) more than 99% identity over 50% of the read length was considered as cut-off when the read matches to the part of the target protein 3) more than 99% identity over 99% of the read length was considered as cut-off when the length of the read is larger than the length of the target protein. The numbers of metagenome reads aligning to the strain specific OGs of TIFN1-8 were summed to calculate the total hit count for each TIFN strain. Additionally, the cumulative hit count for all strains was calculated by adding up the total hit counts of each TIFN strain. Finally, the contribution of each TIFN genome in the metagenome was calculated by dividing the total hit count for that particular TIFN strain by the cumulative hit count of all TIFN strains.

Results

Community of Starter Culture

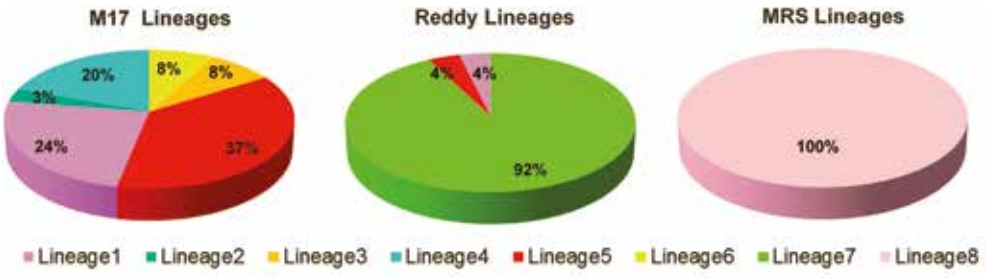
For this study, an undefined Gouda-type cheese starter culture (named UR)(28) that has a long history of use in cheese manufacturing was selected. Therefore, the composition of the microbial community is assumed to be shaped in adaptation to conditions applied in the propagation of the culture. The structure of the community was analyzed using culture dependent and independent methods. The metagenome of the starter culture that was propagated in milk was determined, and relative abundance analyses of classified 16S rRNA sequences revealed the dominance of *Lactococcus lactis* (99%) with a minor population of *Leuconostoc mesenteroides* (1%) (Supplementary Fig. 1).

The metagenome does not allow the extraction of information concerning strain-level diversity since the resolution beyond the species level is lost. For that reason, 140 single colony isolates (68, 56, and 14 colony isolates from LM17, Reddy's, and MRS media, respectively) were obtained from the starter culture with three complementary plating media (Fig. 1), and fingerprinted using an improved amplified fragment length polymorphism (AFLP) procedure with superior resolution compared to all other fingerprinting techniques that was described in chapter 3 . In total, five genetic lineages of *L. lactis* ssp. *cremoris*, two genetic lineages of *L. lactis* ssp. *lactis* and a single *Lc. mesenteroides* lineage could be discriminated (Supplementary Fig. 2) among the isolates. 1.4% of the total community was classified as *Lc. mesenteroides* subsp. *cremoris, and* the remaining 98.6% of the community consisted of the 7 genetic lineages of *L. lactis*, in line with the metagenome analysis. The cumulative enumeration of genetic lineages quantitatively matched with the total cell counts obtained with fluorescence activated cell sorting (FACS) (Fig.1).

Genome Sequences of Representative Strains

To reveal the genetic content of the lineages, genome sequences of single colony isolates from each *L. lactis* (designated strains TIFN1 to TIFN7) and *Lc. mesenteroides* (designated TIFN8) genetic lineage were determined by a combination of 454 and Illumina sequencing. The analysis of *L. lactis* TIFN1-7 core-genome resulted in detection of 1541 orthologous gene clusters (OGs) (Fig. 2A) that is very similar to the core genome definition of *L. lactis* reference genomes (1584 OGs, data not shown). Combined analysis of TIFN1-7 and public *L. lactis* genomes reduced the core-genome to 1337 OGs, indicating the presence of 204 OGs that are conserved among the 7 TIFN *L. lactis* strains but absent at least in one of the reference genomes (Supplementary Table 3). Only 8 of these 204 OGs were absent from all reference strains, and therefore might be niche specific genes for cheese. They were annotated to encode transposases (2 OGs), ComEC late competence protein, transcriptional regulator EpsR, DNA directed DNA polymerase, and hypothetical proteins (3 OGs). The remaining 196 OGs were clustered into COG categories. Carbohydrate and amino acid transport and metabolism were the most abundant, while the rest of the categories contributed less than 10% among all the categories determined (Supplementary Fig. 3). The pan-genome of reference *L. lactis* strains consisted of 3472 OGs (data not shown). Addition of TIFN1- 7 *L. lactis* strains increased the set by 2860 OGs to a total of 6332 OGs (Fig. 2B). The COG distributions of OGs contributed by TIFN strains were dominated by replication and recombination related functions of mostly transposases (43%, Supplementary Fig. 4).

The core OGs that are present as a single copy per genome (Supplementary Table 4) were additionally employed to cluster TIFN 1-to-8 *L. lactis*, *Lc. mesenteroides* and the 4 reference *L. lactis* strains in order to see the sub-species diversification of TIFN strains (Supplementary Fig. 5). As expected, TIFN1, 3, 5, 6, 7 genomes were clustered with SK11 and MG1363 *L. lactis* ssp. *cremoris* reference genomes, and TIFN2 and 4 were clustered together with IL1403 and KF147 *L. lactis* ssp. *lactis* reference genomes. TIFN8, as the only member from *Lc. mesenteroides*, was placed as a distant outlier. Hence, clustering of the single copy core gene set validated the subspecies assignment of TIFN strains that were previously determined by phenotypic tests and AFLP analysis.



Selective Medium	Avg. Cell Count	Std Dev.	Lineage 1	Lineage 5	Lineage 2	Lineage 4	Lineage 3	Lineage 6	Lineage 7	Lineage 8
FACS Total Cell Counts	1,2E+09	1,2E+08	NA	NA	NA	NA	NA	NA	NA	NA
LM17 Plate Counts	2,1E+08	1,2E+07	4,9E+07	7,4E+07	6,5E+06	3,9E+07	1,6E+07	1,6E+07	ND	ND
Reddy Plate Counts	1,0E+09	8,5E+07	3,6E+07	3,6E+07	ND	ND	ND	ND	8,6E+08	ND
MRSV Plate Counts	1,6E+07	1,3E+06	ND	ND	ND	ND	ND	ND	ND	1,5E+07
Estimated Counts	1,1E+09	NA	4,9E+07	7,4E+07	6,5E+06	3,9E+07	1,6E+07	1,6E+07	8,6E+08	1,5E+07
Pecentages based on Plate Counts	NA	NA	11.21		4.22		1,5	1,5	78,2	1,4
Percentages based on Metagenomics	NA	NA	17.81		4.22		7.1	2.9	65	3.6

Figure 1. Genetic lineage coverage of propagated starter culture. MRS supplemented with vancomycine (MRSV), M17-lactose (LM17), Reddy's Agar plate counts and FACS total cell counts are the mean of three replicates. The highest viable count for a particular genetic lineage among all selective media was taken as the closest approximation for the contribution of that genetic lineage to the community and given as estimated count. ND and NA are the abbreviations for not-detected and not-applicable, respectively. Genetic lineage 7 cells that belong to *L. lactis* subsp. *cremoris* accounted for 78.2% of the total starter community and was exclusively recovered from Reddy Agar Medium. The majority of the remaining *L. lactis* isolates belonged to two genetic lineages of subspecies *cremoris* (genetic lineage 1 and 5), whereas the remaining isolates constituted only a minor fraction and encompassed two *L. lactis* subsp. *cremoris* (genetic lineage 3 and 6) and two *L. lactis* subsp. *lactis* biovar. *diacetylactis* (genetic lineage 2 and 4) lineages. These quantitatively minor lineages were recovered only on M17-lactose agar media, because their small population size relative to the dominating lineage 7 prevented their detection on Reddy agar. 1 and 2 represents the cumulative contribution of genetic lineage 1 and 5, and the cumulative contribution of genetic lineage 2 and 4, respectively.

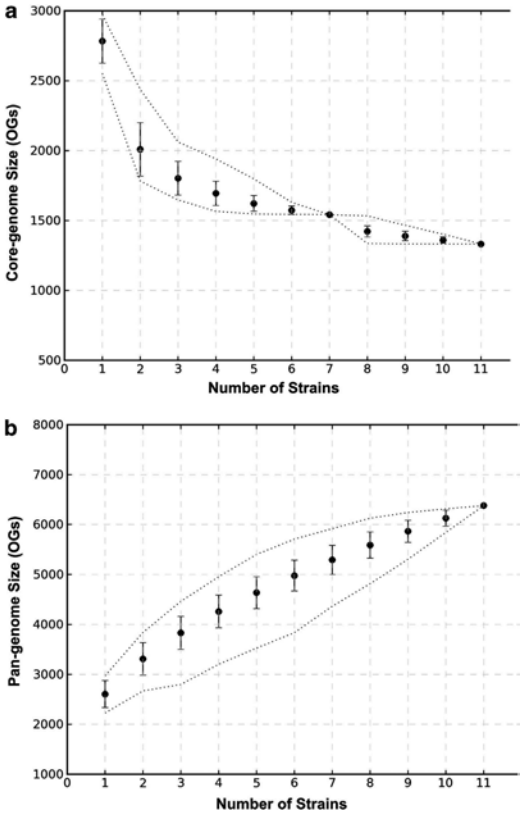


Figure 2. The core- and pan-genome analysis of TIFN *L. lactis* strains in combination with the reference *L. lactis* strains. Panel A: Core-genome analysis. Core-genome sizes were defined as number of shared OGs among all the members of the strain-set at each genome addition. Panel B: Pan-genome analysis. Pan-genome sizes are defined as the number of all non-redundant OGs among all the members of the strain-set at genome addition. The mean, minimum, maximum and standard deviation of the obtained distributions at each genome addition was depicted on the graphs.

Based on the presence and absence of OGs in the individual lineages, TIFN8 was clearly distinct from any of the *L. lactis* strains (772 TIFN8-specific OGs, Supplementary Table 5). Among TIFN *L. lactis* genomes, TIFN6 was identified as the most distant lineage (157 TIFN6-specific OGs), followed by TIFN3 and TIFN7 (91, and 66 strain-specific OGs, respectively). The most closely related strains for which hardly any strain-specific OGs were identified were TIFN2 and TIFN4, (0 and 9 strain-specific OGs, respectively), as well as TIFN1 and 5 (6 and 21 strain-specific OGs, respectively). Additionally, strain-specific OG analyses were employed to identify lineage-specific marker genes for quantification in QPCR assays. No genetic marker was identified for the differentiation of genetic lineages 1 from 5 and 2 from 4, therefore these lineages were detected as lineage-couples. In total 6 markers (Supplementary Table 6) were identified for the specific detection of genetic lineages 1&5, 2&4, 3, 6, 7, 8.

The comparison of the metagenome data with TIFN1-8 genomes indicated that the vast majority of the starter genetic potential is covered by the 8 isolates' genomes

(96%). 1.2% of total metagenome reads in the remaining 4% corresponded to lactococcal sequences, mostly phage encoded functions (Supplementary Fig. 6), whereas the remainder did not match any sequence in NCBI NR protein sequence database (March, 2012). This analysis implies that the metagenome datasets can be used for strain-specific OGs enumeration for the comparison of culture dependent and independent community analyses. Analogous to the culture dependent analysis, TIFN7-specific OGs dominated the metagenome (65%), and TIFN1&5 was identified as the second most abundant group at 17.8%. TIFN3, 6, 8, 2&4 specific OGs provided the smallest contributions to the metagenome (7.1%, 2.9%, 3.6%, and 3.5%, respectively). The contributions of lineage 1&5 and lineage 3 to the community was calculated to be 5% more, while the contribution of lineage 7 was 13% less when compared to the culture based approach (Fig. 1). In general, the distribution of the lineage contributions was found to be in good agreement with the results obtained in the culture dependent approach.

Metabolic Complementation of *Lactococcus* and *Leuconostoc*

To evaluate to what extent *Lactococcus* and *Leuconostoc* strains can metabolically complement each other, the COG categories of the *L. lactis* TIFN1- 7 pan-genome and TIFN8 *Leuconostoc* genome were mapped on the Interactive Pathways Explorer (iPath) (24) for the general metabolism separately and the maps were superimposed (Supplementary Fig. 7). The COGs related to a variety of metabolic pathways were either detected in *L. lactis* pan-genome or in *Leuconostoc* TIFN8 genome, suggesting a complementation for each other (Supplementary Table 7). The *gad* operon, involved in acid stress response of *L. lactis* ssp. *Lactis* (29), provides a specific example for this putative cross-feeding potential (Fig. 3). Cytoplasmic decarboxylation of glutamate is coupled to an expenditure of a proton from the cytoplasm, thereby increasing the intracellular pH, and the reaction product 4-aminobutanoate (GABA) is excreted. The *gad* operon was shown to be induced by chloride and the decarboxylase enzyme has an optimum pH of 4.0 to 5.0 (30), suggesting a role during cheese ripening. The gadR was shown to be induced in cheese by recombinant *in vivo* expression technology (31) as well. The superimposition of the metabolic maps suggests that GABA excreted by *L. lactis* may serve as substrate in the formation of succinate by *L. mesenteroides* since TIFN8 genome encodes a predicted GABA permease, a GABA aminotransferase, and succinate-semialdehyde dehydrogenase (Fig. 3). Notably, GABA utilization has been studied in a few species involved in smear-ripened cheeses (32, 33), but has not been documented for *L. mesenteroides* to date. Metabolic

complementation between the two genera in the catabolism of specific amino acids and possible flavor-formation pathways clearly bears relevance for the performance of the overall starter culture in the cheese production process.

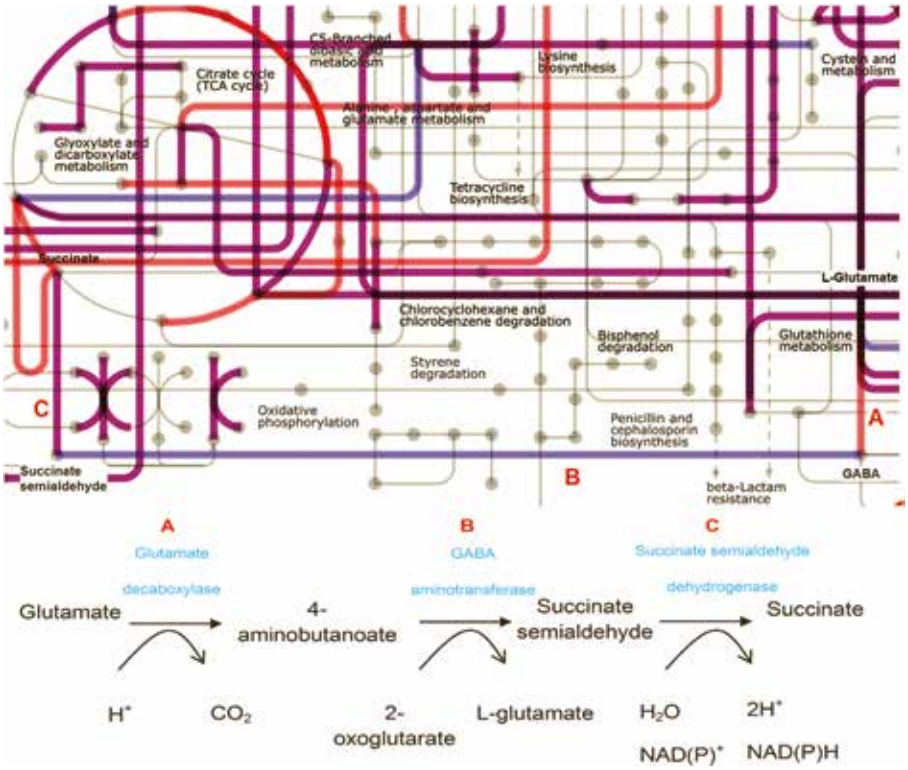


Figure 3. The complementing reactions of *L. lactis* pan-genome and TIFN8 *Lc. mesenteroides* genome on glutamate degradation pathway. The global metabolic maps of *L. lactis* pan-genome (red lines) and TIFN8 *Lc. mesenteroides* genome (blue lines) were superimposed, and the shared genes were represented with purple lines.

Plasmid Content Diversity of Starter Culture

Dairy associated lactococci are known to harbor high number of plasmids (4 to 7) (34) that provides a crucial complement to starter culture functionality and appear to reflect the strain's adaptation to the dairy environment (35). To determine whether plasmid diversity among isolates is independent of their genetic lineage assignment, the plasmid profiles of the starter culture isolates including 2-to-20 representatives of each genetic lineage (68 *L. lactis* and 17 *Lc. mesenteroides*) were determined and 8 distinct plasmid profiles (P1 to P8, Fig. 4) were revealed. Plasmid content heterogeneity was apparent for the isolates of genetic lineages 1 and 5, that

both contained plasmid profiles P2 and P4, while also P6 was identified among the isolates of lineage 1 (Fig. 4). The isolates of the other genetic lineages harbored a single unique plasmid profile for their corresponding lineage (Fig. 4). Analogous to the previous reports (36, 37), functions that are relevant for growth in a dairy environment were found to be plasmid associated, including the citrate transport encoding gene (*citP*), extracellular protease (*prtP*), and lactose utilization (*lac*-operon; *lacG*) (Supplementary Fig. 8). Based on this analysis, cross-lineage plasmid transfer appears to be operational but not all host-plasmid combinations were encountered, suggesting certain incompatibilities.

Role of Bacteriophages in Genetic Diversity of the Starter Culture

Metagenome sequencing of the starter culture revealed the presence of 1.15% phage related DNA sequences in the metagenome. OGs analysis also indicated the presence of prophage DNA in *L. lactis* (average of 1.75%) and *Lc. mesenteroides* (0.65%) genome scaffolds. These observations illustrate the presence of mobile (pro)phages that interact with the strains. To demonstrate the presence of free phages in the starter culture supernatant, plaque formation was monitored using the 7 sequenced *L. lactis* strains as indicator strains. Approximately 10^6 plaque forming unit (pfu)/ml, 10^4 pfu/ml, 10^2 pfu/ml were detected in the supernatant with TIFN1, TIFN5, and TIFN7 as indicators, respectively (data not shown). The phages were isolated and designated as ΦTIFN1, ΦTIFN5 and ΦTIFN7, for which the host range and plaque forming efficacy were determined. Preliminary sequence analysis indicated that ΦTIFN1 and ΦTIF7 belong to the p335 (38), while ΦTIFN5 belongs to p936 subgroup of lactococcal bacteriophages (data not shown).

To find out whether the phage resistance profile is conserved within genetic lineages, 2-to-4 single colony isolates from each genetic lineage were challenged with ΦTIFN1, ΦTIFN5 and ΦTIFN7. The host strains TIFN1, TIFN5, and TIFN7 displayed the highest sensitivity for their corresponding phages (Fig. 4). All strains that were tested and belonged to genetic lineage 1 were sensitive for ΦTIFN1, albeit with remarkably different magnitudes of sensitivity (4 to 7 orders of magnitude in pfu). A similar sensitivity profile for ΦTIFN5 was observed for the isolates belonging to genetic lineage 5. In contrast, TIFN7 was found to be the only sensitive strain to the phage ΦTIFN7 within genetic lineage 7. The phages were also active on other genetic lineages, with the exception of lineage 2 and 3. These results demonstrate a considerable phage sensitivity variation within and between the genetic lineages.

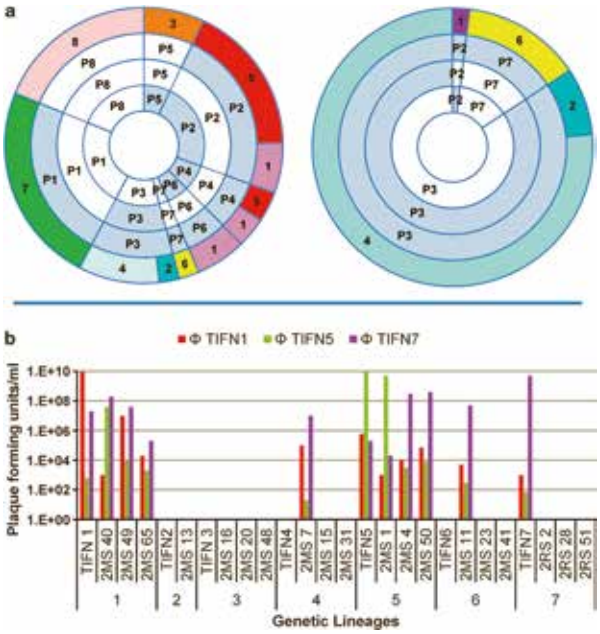


Figure 4. Plasmid profiles and phage sensitivities of starter culture isolates. Panel A: Plasmid profile distribution of the isolates from propagated starter culture (left chart) and from 2-Weeks ripened cheese (right chart). 1st circle represents the AFLP derived genetic lineage classifications. Inner circles represent the corresponding plasmid profiles. Hybridization of lac-operon, citrate permease, and protease permease probes to the plasmid profiles are indicated on 2nd, 3rd, and 4th circles, and hybridized, not-hybridized, and not-determined are indicated with blue, grey and white filling colors, respectively. Panel B: Phage sensitivity levels of isolates from propagated starter culture within and between genetic lineages.

Community stability of the starter culture under a back-slopping propagation regime

Back-slopping is the traditional way of propagating starter cultures. Experience learns that undefined, mixed cultures with a history of back-slopping are more stable against environmental fluctuations as compared to defined starters. One reason for the observed resilience may be the continuous diversification and the resulting heterogeneity of the community members, which can prevent the loss of genetic potential from the population as a whole under environmental fluctuation. In line with this hypothesis, heterogeneous phage sensitivity of strains within and between the genetic lineages supports a dynamic process where phage sweeps eradicate sensitive cells but fail to eradicate an entire lineage.

To demonstrate that this process is operational, we propagated the starter for several weeks in milk by daily sub-culturing and monitored the titer of phages using *L. lactis* TIFN1 as an indicator organism (Fig. 5, Panel A). The initial phage titer of 10^7

pfu/ml increased in both parallel cultures and followed a sharp decline after 5 days, indicating the disappearance of the host strain population from the community. After 11 days, phage counts showed a sharp increase reaching the levels of almost 10^9 pfu/ml, followed by a second decline. Interestingly, genetic lineage quantification by QPCR didn't indicate steep fluctuations in the community composition (Fig. 5, Panel B). Even though the phage predation pressure on TIFN1 strain showed dramatic changes over time, the genetic lineages 1&5 were maintained with only small changes in their relative abundance. This experiment illustrates that the individual strains, which are sensitive to a particular phage may be eradicated at certain stages of propagation, but this only marginally impacts on the relative abundance of the lineage they belong to within community.

Community dynamics in cheese

Fluctuations in the community members during cheese manufacturing have an effect on the functionality of the starter such as acidification or flavor formation (39). Therefore, relative abundance of each genetic lineage during cheese manufacturing was quantified by AFLP typing of randomly picked single colony isolates to reveal the community dynamics (Fig. 6).

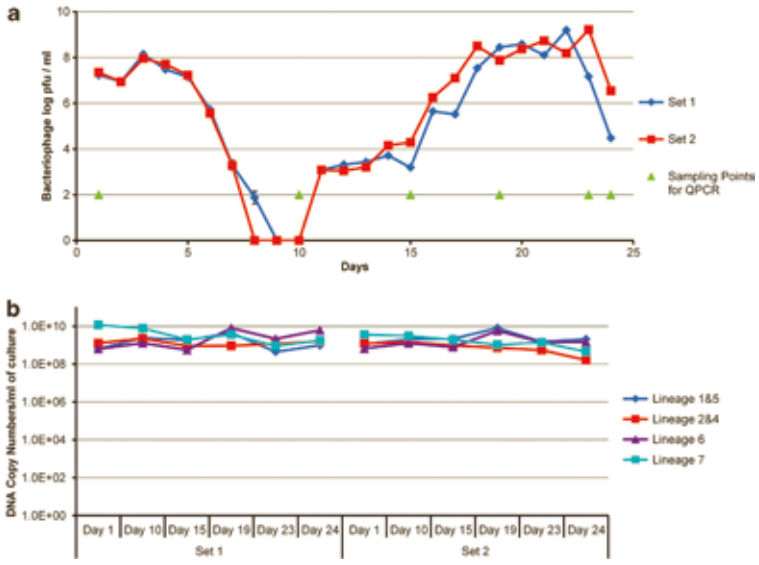


Figure 5. Phage and community dynamics during a back-slopping propagation regime. Panel A: Dynamics in the titer of phages using *L. lactis* TIFN1 as an indicator organism for 25 days of sub-culturing in milk. The Sampling points for QPCR are from where the community dynamics data shown in Panel B were generated. Panel B: Absolute abundance of the genetic lineages 1 to 8 during propagations as determined by genetic lineage specific QPCR. The absolute abundances of genetic lineages 6 and 8 were below the detection level during sub-culturing, and excluded from the graph.

The protease-negative genetic lineage populations had the highest relative abundance in cheese until the ripening period (0.1% - 35%) as consistent with the previous observations (40), indicating a peptide cross feeding for the acquisition of essential amino acids during growth in milk (41). The protease positive cells were stably maintained in the starter community, which is generally not observed in single strain populations (42). The relative abundance of genetic lineage 2&4 (both *L. lactis* subsp. *lactis* biovar *diacetylactis*), lineage 6 (*L. lactis* subsp. *cremoris*), and lineage 8 (*Lc. mesenteroides* subsp. *cremoris)* increased after the brining step. The enrichment of certain populations during cheese ripening was followed at the plasmid content level as well. The profiles P2, P3 and P7 were enriched among isolates of 2-weeks ripened cheese, and the impressive dominance of P3 (85%) carrying the citrate permease gene was in agreement with the relatively good survival of *L. lactis* subsp. *diacetylactis* lineages (Fig. 4A).

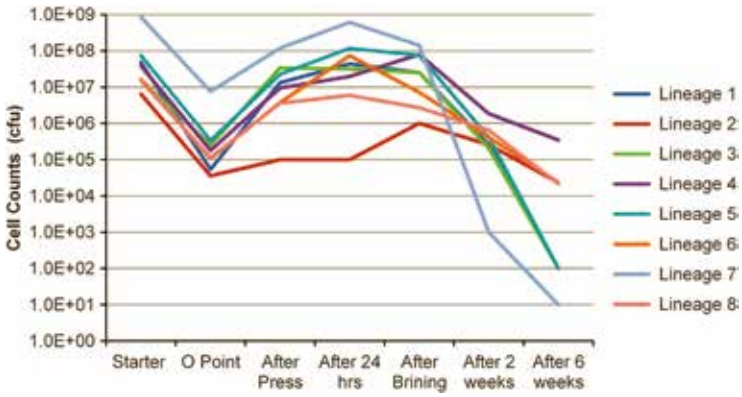


Figure 6. Relative abundance of genetic lineages during cheese manufacturing as determined by AFLP typing of randomly picked isolates plated on complementary media. Community dynamics have an effect on the functionality during cheese manufacturing. The key flavor compounds in Gouda are the metabolic side products of glycolysis (mostly lactose and citrated derived), lipolysis and amino acid metabolism. Therefore, the better survival of citrate utilizing *L. lactis* subsp. *lactis* biovar *diacetylactis* and *Lc. mesenteroides* populations (genetic lineage 2 and 4 and 8) after brining is expected to enhance the flavor formation via citrate metabolism. Amino acid derived flavor development, on the other hand, is enhanced with the concerted activity of intact and lysed cells, and the intracellular peptidase complement of lysed *L. cremoris* strains after brining may control the free amino acid pool during ripening.

Discussion

Environments without stringent selective pressures can accommodate many species, however strongly selective conditions reduce the species diversity. For example, the increase in salinity of hypersaline ponds was shown to lead to an increase in

the abundance of Archeal species, while fresh water was occupied with a variety of prokaryote species (43). Similarly, milk can accommodate numerous microbial species (44). However, strong selective pressures like heat treatment (45), dairy processing conditions (46), intrinsic properties of lactic acid fermentation and antagonistic interactions (47), enrich the microbiota towards a few lactic acid bacteria species.

Metagenomics has been widely used to understand the structure of microbial communities for complex systems such as mammalian GI-tract, which can contain up to 2000 species (48). However its use for simpler communities is not straightforward since the resolution beyond the species level is lost. The metagenomic analysis of a simple community that is involved in Kimchi fermentation revealed the domination of community by members of *Leuconostocs*, *Lactobacillus* and *Weissella*, however the resolution at the strain level was not achieved (49). In the approach presented in this study, the resolution was enhanced with the quantification of lineage-specific OGs that were determined from the genome sequences of distinct lineages. This allowed monitoring of the community dynamics at the genetic lineage level using a metagenome approach.

Our results revealed a high degree of genetic heterogeneity at the strain level rather than the species level. The persistence of closely related strains within each genetic lineage was in agreement with the density dependent phage predation model (8). Genetic lineages appeared stably present despite severe phage predation, exemplifying that "kill-the-winner principles" (9) are operational at the strain level, and not at the genetic lineage level. The heterogeneous phage sensitivity within genetic lineages is the most likely basis for community resilience. Phage predation ensures diversity by suppression of the more abundant strains, thereby stabilizing the overall community functionality. Similar community dynamics were also observed in aquatic environments, where the relative abundance of species was stable while the abundance of the encompassed strains fluctuated dramatically (43). These findings show that this generic phenomenon is operational in a variety of (simple) community systems and may represent a generic driver of community dynamics that has remained underestimated to date.

In conclusion, the multilevel analyses performed here illustrated that a relatively simple bacterial community represent a large genetic landscape and the main phylogenetic groups (or genetic lineages) are stabilized within the community by kill-the-winner phage dynamics that impact at the strain level.

Acknowledgements

We acknowledge Corné H.W. Klaassen (Canisius Wilhelmina Hospital) for his technical support in AFLP, Marja Twigt for the set-up of the mini-scale cheese manufacturing procedure, Herwig Bachmann (NIZO Food Research) for the set-up of mini-scale cheese manufacturing and the critical reading of the manuscript. This work was part of Kluyver Centre for Genomics of Industrial Fermentation.

References

1. **Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM, Nelson KE.** 2006. Metagenomic analysis of the human distal gut microbiome. Science **312:**1355-1359.
2. **Benlloch S, Lopez-Lopez A, Casamayor EO, Ovreas L, Goddard V, Daae FL, Smerdon G, Massana R, Joint I, Thingstad F, Pedros-Alio C, Rodriguez-Valera F.** 2002. Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern. Environ Microbiol **4:**349-360.
3. **Heuer H, Abdo Z, Smalla K.** 2008. Patchy distribution of flexible genetic elements in bacterial populations mediates robustness to environmental uncertainty. FEMS Microbiol Ecol **65:**361-371.
4. **Dobrindt U, Hochhut B, Hentschel U, Hacker J.** 2004. Genomic islands in pathogenic and environmental microorganisms. Nat Rev Microbiol **2:**414-424.
5. **Penn K, Jenkins C, Nett M, Udwary DW, Gontang EA, McGlinchey RP, Foster B, Lapidus A, Podell S, Allen EE, Moore BS, Jensen PR.** 2009. Genomic islands link secondary metabolism to functional adaptation in marine Actinobacteria. ISME J **3:**1193-1203.
6. **Lilley AK, Bailey MJ.** 1997. The acquisition of indigenous plasmids by a genetically marked pseudomonad population colonizing the sugar beet phytosphere is related to local environmental conditions. Appl Environ Microbiol **63:**1577-1583.
7. **Lilley AK, Bailey MJ.** 1997. Impact of Plasmid pQBR103 Acquisition and Carriage on the Phytosphere Fitness of *Pseudomonas fluorescens* SBW25: Burden and Benefit. Appl Environ Microbiol **63:**1584-1587.
8. **Rodriguez-Valera F, Martin-Cuadrado AB, Rodriguez-Brito B, Pasic L, Thingstad TF, Rohwer F, Mira A.** 2009. Explaining microbial population genomics through phage predation. Nat Rev Microbiol **7:**828-836.
9. **Thingstad TF.** 2000. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnol Oceanogr **45:**1320-1328.
10. **Jousset A, Schulz W, Scheu S, Eisenhauer N.** 2011. Intraspecific genotypic richness and relatedness predict the invasibility of microbial communities. ISME J **5:**1108-1114.
11. **Buckenhuskes HJ.** 1993. Selection Criteria for Lactic-Acid Bacteria to Be Used as Starter Cultures for Various Food Commodities. Fems Microbiology Reviews **12:**253-272.
12. **de Vos WM.** 2011. Systems solutions by lactic acid bacteria: from paradigms to practice. Microb Cell Fact **10**.
13. **Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376-380.
14. **Bennett S.** 2004. Solexa Ltd. Pharmacogenomics **5:**433-438.

15. **Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter JC.** 2000. A whole-genome assembly of Drosophila. Science **287:**2196-2204.
16. **Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W.** 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics **27:**578-579.
17. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.** 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics **9:**75.
18. **Wegmann U, O'Connell-Motherway M, Zomer A, Buist G, Shearman C, Canchaya C, Ventura M, Goesmann A, Gasson MJ, Kuipers OP, van Sinderen D, Kok J.** 2007. Complete genome sequence of the prototype lactic acid bacterium *Lactococcus lactis* subsp. *cremoris* MG1363. Journal of bacteriology **189:**3256-3270.
19. **Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, Ehrlich SD, Sorokin A.** 2001. The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. Genome Res **11:**731-753.
20. **Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, Shakhova V, Grigoriev I, Lou Y, Rohksar D, Lucas S, Huang K, Goodstein DM, Hawkins T, Plengvidhya V, Welker D, Hughes J, Goh Y, Benson A, Baldwin K, Lee JH, Diaz-Muniz I, Dosti B, Smeianov V, Wechter W, Barabote R, Lorca G, Altermann E, Barrangou R, Ganesan B, Xie Y, Rawsthorne H, Tamir D, Parker C, Breidt F, Broadbent J, Hutkins R, O'Sullivan D, Steele J, Unlu G, Saier M, Klaenhammer T, Richardson P, Kozyavkin S, Weimer B, Mills D.** 2006. Comparative genomics of the lactic acid bacteria. Proceedings of the National Academy of Sciences of the United States of America **103:**15611-15616.
21. **Siezen RJ, Bayjanov J, Renckens B, Wels M, van Hijum SA, Molenaar D, van Hylckama Vlieg JE.** 2010. Complete genome sequence of *Lactococcus lactis* subsp. *lactis* KF147, a plant-associated lactic acid bacterium. Journal of bacteriology **192:**2649-2650.
22. **Li L, Stoeckert CJ, Jr., Roos DS.** 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res **13:**2178-2189.
23. **Snel B, Bork P, Huynen MA.** 2002. The identification of functional modules from the genomic association of genes. Proceedings of the National Academy of Sciences of the United States of America **99:**5890-5895.
24. **Letunic I, Yamada T, Kanehisa M, Bork P.** 2008. iPath: interactive exploration of biochemical pathways and networks. Trends Biochem Sci **33:**101-103.
25. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research **25:**3389-3402.
26. **Huson DH, Auch AF, Qi J, Schuster SC.** 2007. MEGAN analysis of metagenomic data. Genome Res **17:**377-386.
27. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. Journal of molecular biology **215:**403-410.
28. **Stadhouders J, Leenders, G.J.M.** 1984. Spontaneously developed mixed-strain cheese starters. Their behaviourstowards phages, and theri use in Dutch cheese industry. Netherlands Milk and Dairy Journal **38:**25.
29. **Cotter PD, Hill C.** 2003. Surviving the acid test: responses of gram-positive bacteria to low pH. Microbiology and molecular biology reviews : MMBR **67:**429-453, table of contents.
30. **Sanders JW, Leenhouts K, Burghoorn J, Brands JR, Venema G, Kok J.** 1998. A chloride-inducible acid resistance mechanism in *Lactococcus lactis* and its regulation. Molecular microbiology **27:**299-310.
31. **Bachmann H, de Wilt L, Kleerebezem M, van Hylckama Vlieg JE.** 2010. Time-resolved genetic responses of *Lactococcus lactis* to a dairy environment. Environ Microbiol **12:**1260-1270.
32. **Schroder J, Maus I, Trost E, Tauch A.** 2011. Complete genome sequence of Corynebacterium variabile DSM 44702 isolated from the surface of smear-ripened cheeses and insights into cheese ripening and flavor generation. BMC Genomics **12:**545.

33. **Monnet C, Loux V, Gibrat JF, Spinnler E, Barbe V, Vacherie B, Gavory F, Gourbeyre E, Siguier P, Chandler M, Elleuch R, Irlinger F, Vallaeys T.** 2010. The arthrobacter arilaitensis Re117 genome sequence reveals its genetic adaptation to the surface of cheese. PloS one **5:**e15489.

34. **Mills S, McAuliffe OE, Coffey A, Fitzgerald GF, Ross RP.** 2006. Plasmids of lactococci - genetic accessories or genetic necessities? FEMS Microbiol Rev **30:**243-273.

35. **Siezen RJ, Starrenburg MJ, Boekhorst J, Renckens B, Molenaar D, van Hylckama Vlieg JE.** 2008. Genome-scale genotype-phenotype matching of two *Lactococcus lactis* isolates from plants identifies mechanisms of adaptation to the plant niche. Appl Environ Microbiol **74:**424-436.

36. **Feirtag JM, Petzel JP, Pasalodos E, Baldwin KA, McKay LL.** 1991. Thermosensitive plasmid replication, temperature-sensitive host growth, and chromosomal plasmid integration conferred by Lactococcus lactis subsp. cremoris lactose plasmids in *Lactococcus lactis* subsp. *lactis*. Appl Environ Microbiol **57:**539-548.

37. **Siezen RJ, Renckens B, van Swam I, Peters S, van Kranenburg R, Kleerebezem M, de Vos WM.** 2005. Complete sequences of four plasmids of *Lactococcus lactis* subsp. *cremoris* SK11 reveal extensive adaptation to the dairy environment. Appl Environ Microbiol **71:**8371-8382.

38. **Samson JE, Moineau S.** 2010. Characterization of *Lactococcus lactis* phage 949 and comparison with other lactococcal phages. Appl Environ Microbiol **76:**6843-6852.

39. **Smit G, Smit BA, Engels WJ.** 2005. Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. FEMS Microbiol Rev **29:**591-610.

40. **Hugenholtz J, Splint R, Konings WN, Veldkamp H.** 1987. Selection of Protease-Positive and Protease-Negative Variants of *Streptococcus cremoris*. Appl Environ Microbiol **53:**309-314.

41. **Kunji ER, Mierau I, Hagting A, Poolman B, Konings WN.** 1996. The proteolytic systems of lactic acid bacteria. Antonie Van Leeuwenhoek **70:**187-221.

42. **Stadhouders J.** 1986. The control of cheese starter activity. Netherlands Milk and Dairy Journal **40:**18.

43. **Rodriguez-Brito B, Li L, Wegley L, Furlan M, Angly F, Breitbart M, Buchanan J, Desnues C, Dinsdale E, Edwards R, Felts B, Haynes M, Liu H, Lipson D, Mahaffy J, Martin-Cuadrado AB, Mira A, Nulton J, Pasic L, Rayhawk S, Rodriguez-Mueller J, Rodriguez-Valera F, Salamon P, Srinagesh S, Thingstad TF, Tran T, Thurber RV, Willner D, Youle M, Rohwer F.** 2010. Viral and microbial community dynamics in four aquatic environments. ISME J **4:**739-751.

44. **Delbes C, Ali-Mandjee L, Montel MC.** 2007. Monitoring bacterial communities in raw milk and cheese by culture-dependent and -independent 16S rRNA gene-based analyses. Appl Environ Microbiol **73:**1882-1891.

45. **Van Hoorde K, Verstraete T, Vandamme P, Huys G.** 2008. Diversity of lactic acid bacteria in two Flemish artisan raw milk Gouda-type cheeses. Food Microbiol **25:**929-935.

46. **Randazzo CL, Torriani S, Akkermans AD, de Vos WM, Vaughan EE.** 2002. Diversity, dynamics, and activity of bacterial communities during production of an artisanal Sicilian cheese as evaluated by 16S rRNA analysis. Appl Environ Microbiol **68:**1882-1892.

47. **Lindgren SE, Dobrogosz WJ.** 1990. Antagonistic activities of lactic acid bacteria in food and feed fermentations. FEMS Microbiol Rev **7:**149-163.

48. **Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Jian M, Zhou Y, Li Y, Zhang X, Qin N, Yang H, Wang J, Brunak S, Dore J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, Ehrlich SD.** 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature **464:**59-65.

49. **Jung JY, Lee SH, Kim JM, Park MS, Bae JW, Hahn Y, Madsen EL, Jeon CO.** 2011. Metagenomic analysis of kimchi, a traditional Korean fermented food. Appl Environ Microbiol **77:**2264-2274.

Supplementary Information for Chapter 3

Supplementary Information for Methods

Lab Scale Mini-cheese Making

2L of bovine milk was standardized for fat (3.5 %), protein (3.4 %) and lactose (4.5 %) content, and heat treated at 72.5 °C for 9 s. The starter culture was propagated at 20 °C for 20 h in sterile skimmed milk two times prior to cheese making. The cheese milk was warmed to 32 °C, and supplemented with 230 μL of renneting enzyme (Kalase, 150 IMCU, CSK Food Enrichment, Ede, The Netherlands), 400 μL of a 33 % (wt/vol) $CaCl_2$ solution, 400 μL of $NaNO_3$ solution, and 1 % (v/v) of propagated starter culture per liter of milk. After curdling at 30.5 °C for 35-45 min, the curd was cut for 10 min using a custom-made stirring device. 800 mL of whey was removed and replaced with 700 mL of sterile deionized water at 45 °C. The curd was incubated at 36 °C for 45 min with gentle stirring. Whey was removed from curd under 1kg weight pressure for 1,5 h at 30°C. The cheese was incubated at 30 °C until the pH was reduced to the range of 5.4-5.6. Subsequently, it was brined with a sterile 19 % (wt/vol) NaCl solution for 4 h at 13 °C to obtain 3 % salt in dry matter in the center of the cheese. The cheese with the final moisture content in the range of (42-44) % and pH in the range of (5.4-5.2) was vacuum-packed and ripened at 13 °C for 6 weeks.

Total Genomic DNA Isolation From the Starter Culture

2,5 gr of propagated starter culture was mixed with 2 % (wt/vol) sodium citrate solution at 45 °C in 25 ml final volume, and homogenized for 5 min with stomacher. The homogenate was centrifuged at 13.750 g for 10 min at room temperature (Beckman Avanti J-20, CA, USA). The supernatant and fat layers were removed and the cells were re-suspended in 1 ml of the same solution at 45 °C for washing. The cell suspension was centrifuged at 10.750 g for 5 min (Eppendorf 5417R Microcentrifuge, Hamburg, Germany), and the washing step was repeated for 2-3 times until all fat was removed. The DNA from the washed cell pellet was further isolated with Quiagen DNeasy Blood and Tissue Isolation Kit with the following modifications. The cells were lysed in 1 ml of lysis buffer for 1 h at 37 °C. All kit ingredients were adjusted to 1 ml lysis volume. The lysed cell suspension was protease treated for 1 h at 56 °C and RNase treated for 30 min at room temperature subsequently. Final DNA solution was applied to a single spin column, and washing steps were repeated 2

times to get rid of rest of the milk culture impurities. DNA was eluted with 300 µl of nuclease free water in three steps (100 µl in each step), the concentrations were determined by Nanodrop (Coleman technologies Inc., Glen Mills, PA, USA) and 0.8 % (wt/vol) agarose gels.

Quantitative PCR

QPCR reactions were performed in 20 µl of final volume, combining 10 µl of SYBR® Green PCR master mix (Applied Biosystem, Warrington, UK) with 0,2 µl of each primer (10 µM), and 2 µl of DNA template. PCR amplification (7500 Fast System, Applied Biosystem, Warrington, UK) was initiated with 10 min of initial denaturation at 95 °C, followed with 40 cycles of 15 s of denaturation at 95 °C and 1 min of annealing and extension at 60 °C.

Determination of Community Dynamics During Cheese Manufacturing and Ripening

For the time series analysis, 10 g of samples were collected at 8 points during cheese manufacturing: from initial starter culture just before milk inoculation, starter culture inoculated milk (zero point), pressed cheese, 24-h-incubated cheese, brined cheese, 2-weeks-ripened cheese and 6-weeks-ripened cheese. The samples were homogenized with 2 % (wt/vol) sodium citrate solution at 45 °C in 100 ml final volume, diluted from 10^{-1} to 10^{-8} with peptone water, and plated on selective media. M17 agar(1) supplemented with 0.5 % (wt/vol) lactose and Reddy agar(2) were used for the isolation of lactococci at 30 °C and MRS agar (Merck, Darmstadt, Germany) supplemented with vancomycine (20 µg/ml) for the isolation of *Leuconostoc* at 25 °C. The dilutions plates that has between 10 and 150 colonies were considered relevant for isolation, and all the colonies present on the relevant dilution plate were picked to obtain an unbiased lineage distribution. The isolated colonies were activated in the liquid culture of the same isolation medium, and stored in glycerol stocks (30 % glycerol, 70 % liquid culture) at -80 °C.

Phenotypic differentiation of *Lactococcus* subspecies and bio-varieties was achieved using previously established phenotypic analysis(3). Protease activities were determined with glycerophosphate milk agar(4) screening, and fingerprinted with the AFLP protocol optimized for *L. lactis* in Chapter 2. For the quantification of genetic lineages during cheese manufacturing, the highest viable count for a particular

genetic lineage among all selective media was taken as the closest approximation of the contribution for that genetic lineage to the community.

Plasmid isolation, Southern Blot Hybridizations, and High-Throughput Plasmid Profiling

The overnight grown cultures in M17 were diluted 10 times with the same medium, incubated further to OD_{600nm} of 0,5. 5 ml of this culture was centrifuged at 1560 g for 10 min (Heraeus Megafuge 1R, Heraeus Instruments, Hanau, Germany), washed with 1 ml of nuclease free water and centrifuged again at 20.800 g for 2 min (Eppendorf 5417R Microcentrifuge, Hamburg, Germany). The pellet was re-suspended in 380 µl of lysis buffer (6.7 % sucrose, 500 mM Tris.HCl, 1 mM EDTA pH 8.0, 10 µg/ml RNase) and incubated for 2 min at 37 °C. Subsequently, 96 µl of lysozyme (10 mg/ml in 25mM Tris.HCl pH 8.0) was added to lysis solution and incubated further for 7 min at 37 °C. 250 µl of Tris-EDTA solution (50 mM EDTA, 50 mM Tris pH 8.0) and 28 µl of SDS solution (20 % SDS in 50 mM Tris-HCl, 20 mM EDTA pH 8.0) was added sequentially with gentle mixing in the end of lysozyme treatment. The lysis was completed with 10 min incubation at 37 °C, and 10 s full speed vortexing. Before the DNA extraction, the lysate was treated with NaOH (addition of 28 µl of NaOH (3 M), 10 min incubation on ice), with Tris.HCl (addition of 50 µl of Tris.HCl (2 M, pH 7.0), 5 min incubation at 25 °C), and with NaCl (addition of 72 µl of 5 M NaCl, 30 min incubation on ice) sequentially. The cell debris was separated by 10 min centrifugation at 20.800 g (Eppendorf 5417R Microcentrifuge, Hamburg, Germany), and the plasmid containing supernatant was taken for phenol-chloroform extraction (addition of 700 µl cold phenol saturated with 3 % NaCl, centrifugation at 20.800 g (Eppendorf 5417R Microcentrifuge, Hamburg, Germany) for 5 min, transferring the water phase into 700 µl chloroform, centrifugation at 20.800 g (Eppendorf 5417R Microcentrifuge, Hamburg, Germany) for 2 min. The DNA in the water phase was precipitated with an equal volume of isopropanol, and by centrifugation at 20.800 g (Eppendorf 5417R Microcentrifuge, Hamburg, Germany) for 5 min. The pellet was washed with 500 µl of ethanol (70 %), dried for 20 min at room temperature and dissolved in 50 µl TE buffer. Finally 10 µl was loaded on agarose gel (0.7 % agarose, 0.5 µg/ml EtBr) and run at 70 mA for 4 h for profiling. Southern Blot Hybridizations were performed as described before(5).

High-throughput Plasmid Profiling was based on fluorophore labeling of digested plasmid DNA fragments. FAM6 fluorophore labeled adapters were prepared from

linker I ([6FAM] GAC GAT GAG TCC TGA G) and linker II (TAC TCA GGA CTC AT) in adapter mix (50 μM linker I, 50 μM linker II, 4.5 mM Tris, 0.45 mM EDTA) by heating for 5 min at 95 °C and cooling down to 5 °C in 30 min. 10 μl of plasmid DNA was digested with 0.2 Unit *MseI* restriction enzyme (New England Biolabs, Beverly, MA, USA) in 10 μl of restriction buffer for 2 h at 37 °C. 5 μl of digested plasmid DNA was ligated with adapters in ligation mix (1 X ligation buffer, 2.5 μM adapter, 0.2 Unit T4 DNA ligase) at 37 °C for 3 h. Ligated fragments were purified with MSB HTS PCRapace DNA fragment purification kit (Invitek GmbH, Berlin, Germany) according to the manufacturer's instruction. For capillary electrophoresis, 8 volumes of ligation product were mixed with 1.75 volume of PCR grade water and 0.25 volume of MegaBACE™ ET550-R size standard (GE Healthcare, Little Chalfont Buckinghamshire, UK). All samples were analyzed with MegaBACE 500 48-capillary electrophoresis system (GE Healthcare, Diegem, Belgium) using FAM6 sensitive detector according to manufacturer`s protocol.

Phage Isolations and Sensitivity Testing

Starter culture was propagated (1 % inoculated) two times in sterile skimmed milk for 20 h at 20 °C. 10 ml of culture was centrifuged at 3000 g for 10 min (Heraeus Megafuge 1R, Heraeus Instruments, Hanau, Germany), and the supernatant was used as a phage solution. TIFN1-7 strains were used as host to test the presence of phages in the supernatant. 100 μl of a full grown strain was mixed with 100 μl of phage solution (in dilution series) and incubated for 10 min at 37 °C. 3 ml LM17 soft agar with 10 mM $CaCl_2$ (46 °C) was added to phage-bacteria mix and the mixture was plated on M17 plates (1 % lactose and 10 mM $CaCl_2$). The plates were incubated at 30 °C, and the phage titer was determined after 20 h. Single plaques were isolated with a Pasteur pipette from agar plates with the host strains TIFN1, TIFN5 and TIFN7, and kept in 1 ml of 50 mM Tris.HCl phage buffer (pH 7.5, 100 mM NaCl, 5 mM $CalCl_2$, 1 mM $MgSO_4$, 0.01 % gelatin). The phage solution was stored at 4 °C.

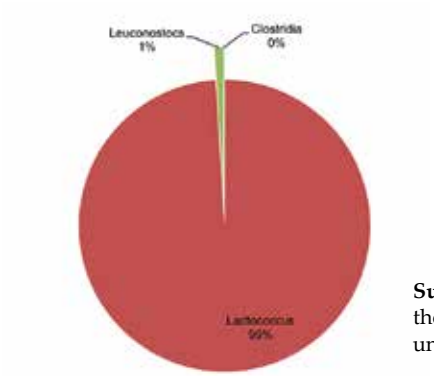
High density phage solutions were prepared by cell culture lysis for sensitivity testing. Phage solutions from previous isolations were mixed with over-night grown cells in 5:1 volume ratio, followed by 10 min adsorption time at 37 °C and diluted 10 times in M17 broth (1 % lactose, 10 mM $CalCl_2$). Lysis of the culture occurred after 6 h of incubation at 30 °C. Cell debris was removed by centrifugation, the supernatant was filter sterilized and the phage titer was determined by plating as described before. If the phage titer was low, the phage isolation was repeated either by another

plate lysis or cell culture lysis. Four isolates of each genetic lineage (including the TIFN strains) were tested for their phage sensitivity, using the high density phage preparations of ϕtifn1, ϕtifn5, ϕtifn7 and Ur-supernatant with dilutions from 10^{10} to 10^{2} pfu/ml.

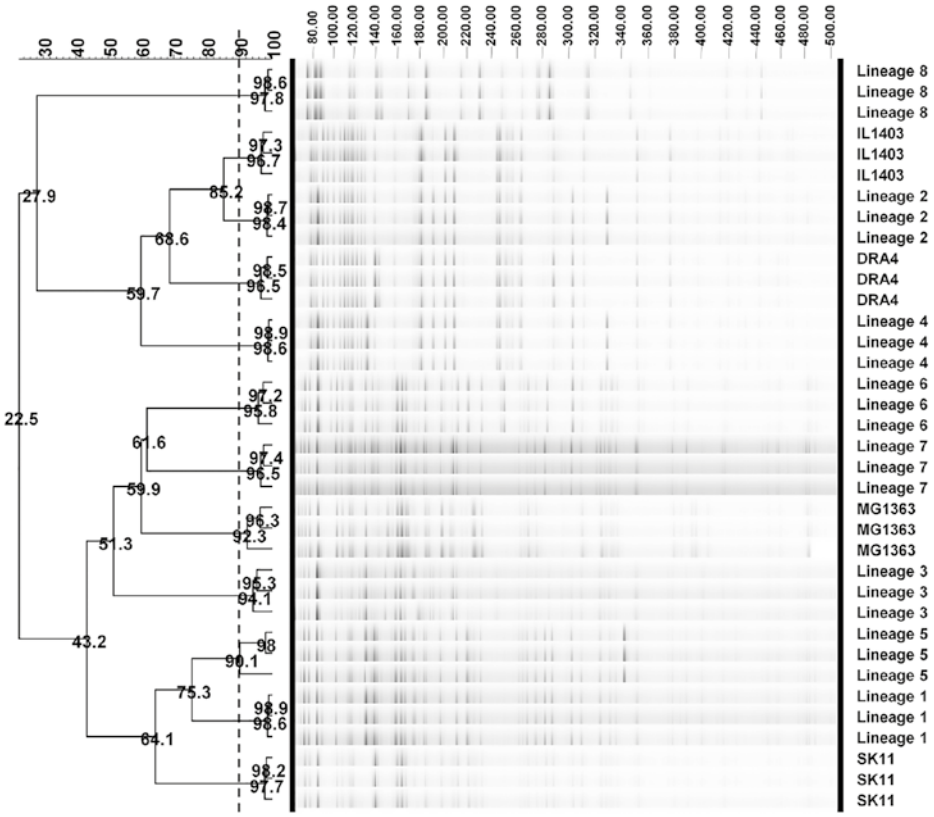
UR Back-Slopping and Phage Titer Determinations

The starter culture was cultivated in skimmed milk (0 % fat, UHT, Friesland Campina, Ede, The Netherlands) at 30 °C overnight. In back-slopping regime, the culture was further sequentially propagated by inoculation of the 5 ml skimmed milk with 50 µl (1 %) of the previous day culture and incubated at 30 °C for 24 hours. The back-slopping was performed in duplicate. Indicator strain TIFN1 was grown in M17 broth media (Oxoid, Hampshire, UK) with 1 % of lactose (w/v) at 30 °C overnight. Bacteriophage titers were determined as follow; 10 ml of culture was centrifuged at 3000 g for 10 min (Heraeus Megafuge 1R, Heraeus Instruments, Hanau, Germany) and the supernatant was used as phage solution. Serial dilutions of phage solution were prepared with 10 times dilutions in phage buffer (pH 7.5, 50 mM Tris.HCl, 100 mM NaCl, 5 mM $CalCl_2$, 1 mM $MgSO_4$, 0.01 % gelatin). 100 µl of Ur culture supernatant (or phage dilutions) were added to 100 µl of the overnight grown indicator strain, kept at 37 °C for 10 min, and subsequently mixed with 3 ml of M17 agar (1 % lactose, 10 mM of $CalCl_2$, 0.75 % agar) to be used as top agar. Finally, the top agar preparations were poured on M17 bottom agar plates (1 % lactose, 10 mM of $CaCl_2$, 1.5 % agar) and incubated at 30°C for 24 h in duplicate. Plaque forming units were counted and expressed as log pfu/ml.

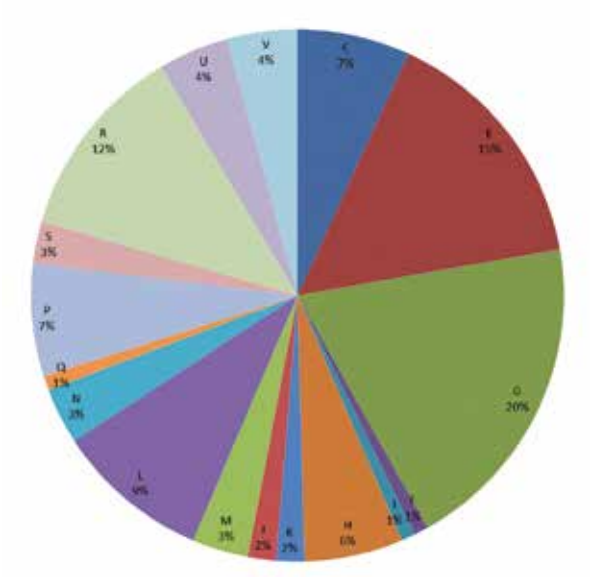
Supplementary Figures



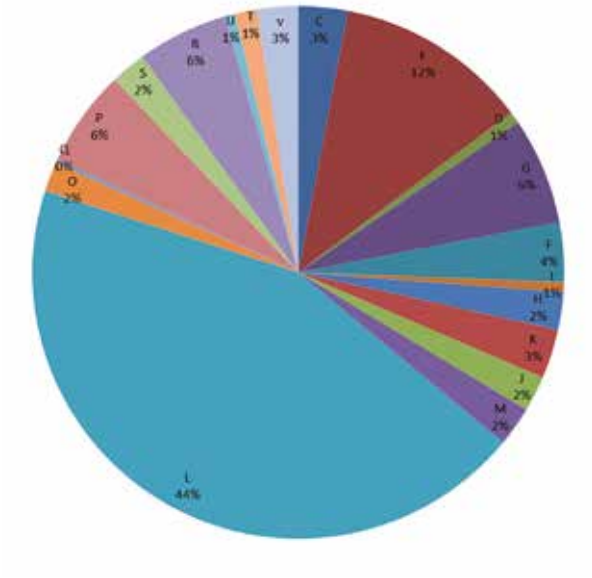
Supplementary Figure 1. Taxonomic composition of the classified 16S rRNA reads in the metagenome of the undefined cheese starter culture.



Supplementary Figure 2. Clustering of AFLP profiles of 3 independently picked representatives for the genetic lineages found in starter culture and independent triplicate profiles of the reference *L. Lactis* strains IL1403, MG1363, SK11, and DRA4. Clustering was based on Pearson correlation coefficient on the range of (80-500) base pairs. Similarity of the genetic profiles of less than 90% similarity (dotted line) was considered as cut-off for defining seperate genetic lineages.

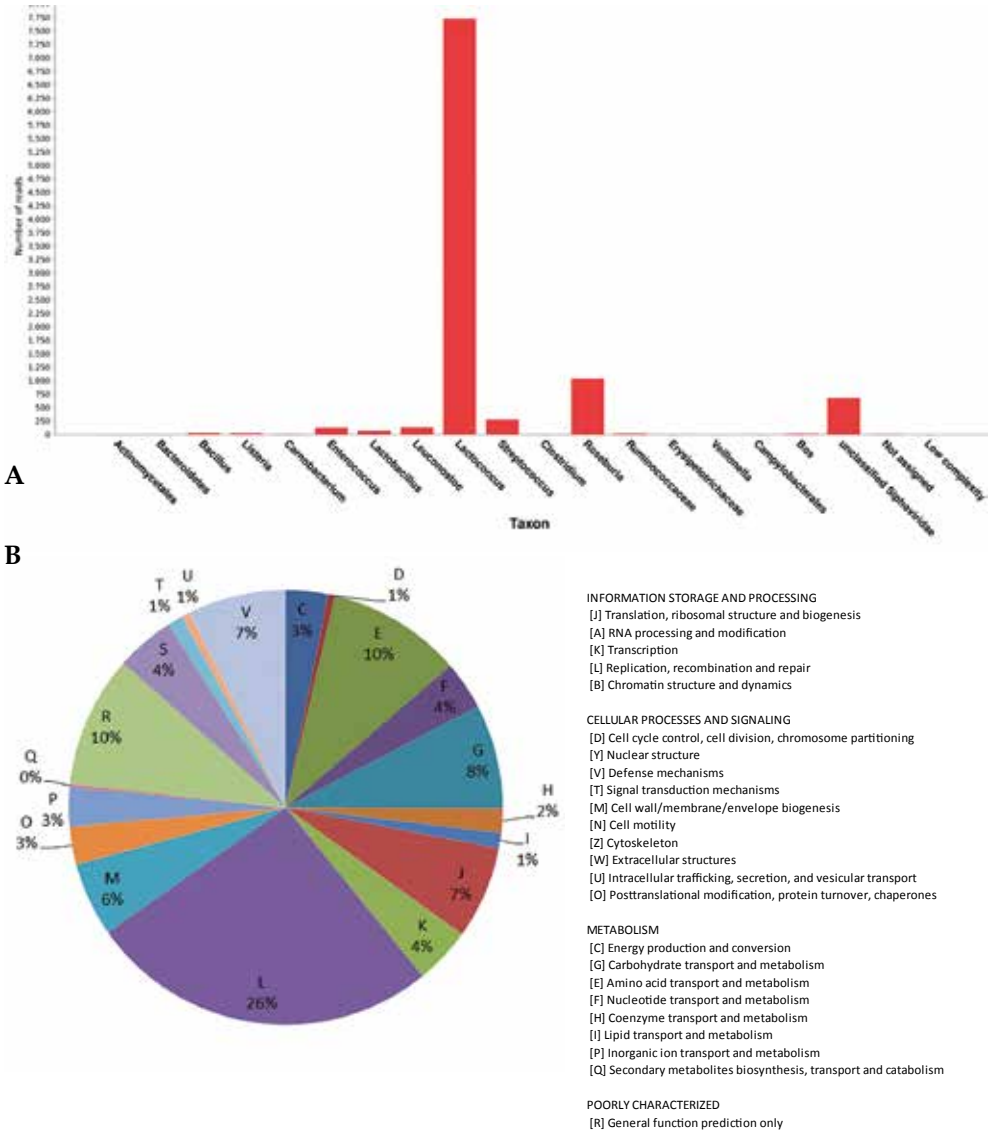


Supplementary Figure 3. Distribution of COG Categories for the 204 OGs conserved exclusively on TIFN1-to-7 *L. lactis* core-genome, but not conserved for the reference *L. lactis* strains.

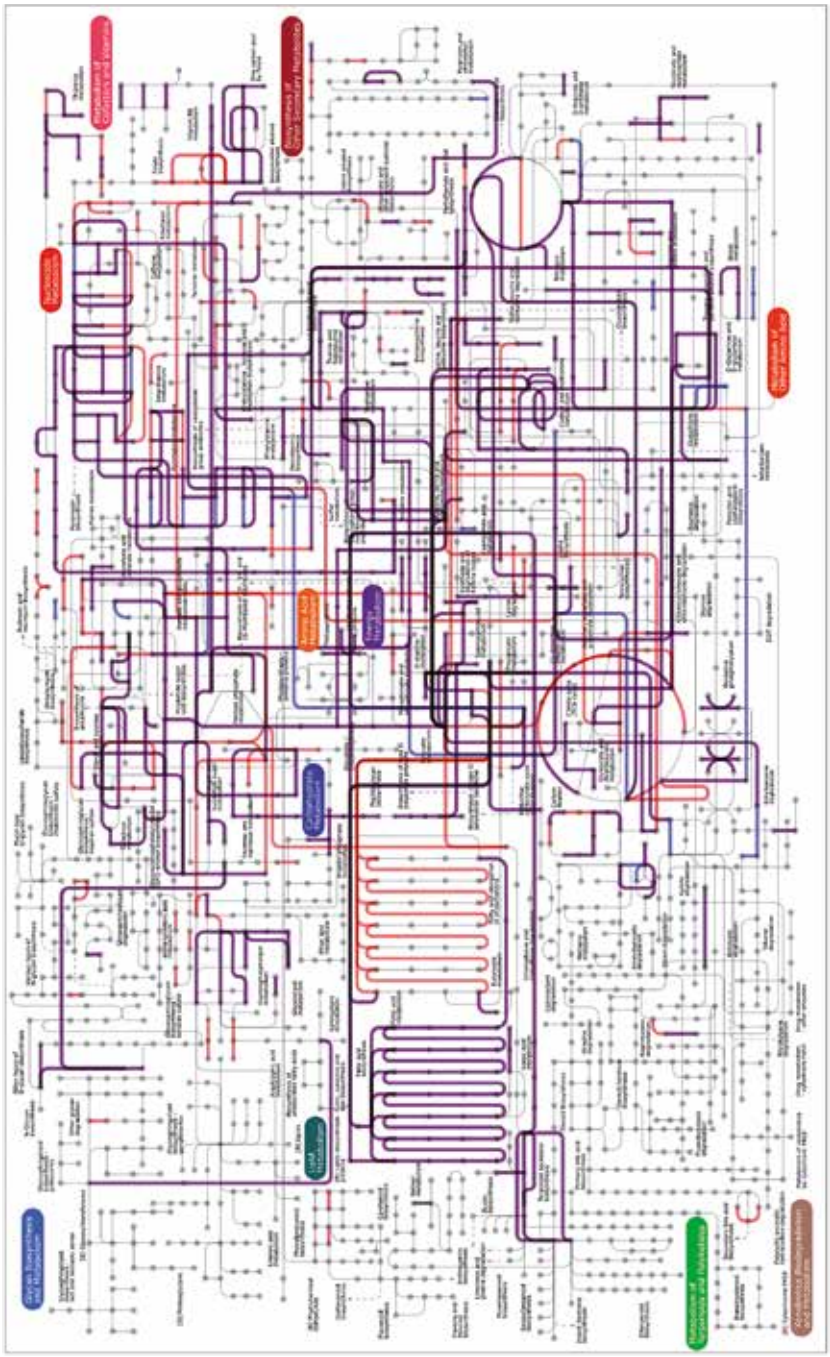


Supplementary Figure 4. Distribution of COG Categories for the 2860 OGs contributed by TIFN1-to-7 *L. lactis* genomes to the pan-genome of *L. lactis* reference strains.

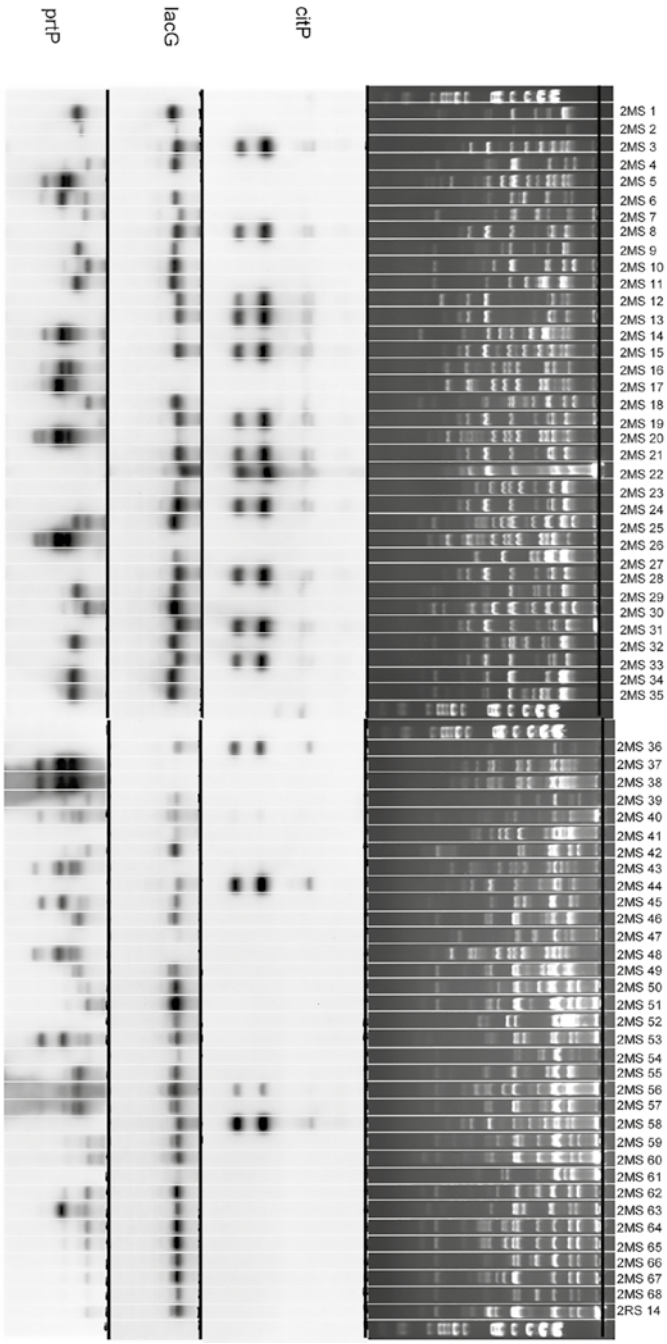
Supplementary Figure 5. Clustering of TIFN 1-to-8 *L. lactis, Lc. mesenteroides* and the 4 reference *L. lactis* strains. The dendrogram was constructed based on the concatenated variable nucleotides of the core OGs that are present as a single copy on all the genomes included to the clustering.



Supplementary Figure 6. Taxonomy-based binning of metagenome reads that do not match to TIFN 1-to-8 genomes. Panel A: Taxa identified for the sequences of the metagenome that do not match to TIFN genomes. Panel B: Distribution of COG Categories for the reads in Panel A.



Supplementary Figure 7. Global metabolic maps of *L. lactis* pangenome (red lines) and TIFN8 *Lc. mesenteroides* genome (blue lines). The shared pathways are represented with purple lines.



Supplementary Figure 8: Agarose gel electrophoresis of plasmid preparations, and southern blot hybridization patterns for *citP(6), lacG(7),* and *prtP(8)* genes. The codes of the isolates that the plasmid profiles belong to were given at the top of the figure

Supplementary Table 1. Statistics of TIFN Genome Sequences

Strain	n>200	n>N50	coverage	scaffolds with 'N' stretches	min	N80	N50	N20	max	sum	RAST CDS	prodigal CDS complete	5' clip	3' clip	fragment	Prodigal CDS sum	Type
TIFN1	123	2	34	25	623	127788	525167	1024008	1024008	2679966	2961	2596	149	154	7	2906	Lactococcus lactis subsp cremoris
TIFN2	57	2	22	13	702	223560	580079	1208704	1208704	2505042	2640	2435	80	66	6	2587	Lactococcus lactis subsp. lactis biovar diacetylactis
TIFN3	235	3	22	28	728	45334	399043	500333	500862	2725178	3069	2600	186	204	20	3010	Lactococcus lactis subsp cremoris
TIFN4	70	2	29	14	692	136786	422206	1104390	1104390	2550393	2655	2478	73	76	7	2634	Lactococcus lactis subsp. lactis biovar diacetylactis
TIFN5	159	4	18	34	677	34625	155364	442104	487864	2541509	3023	2401	364	322	36	3123	Lactococcus lactis subsp cremoris
TIFN6	271	8	17	80	636	16152	74666	231059	296060	2591499	3166	2389	456	425	51	3321	Lactococcus lactis subsp cremoris
TIFN7	156	5	33	24	692	60154	195833	309889	462454	2634063	3242	3102	385	349	40	3876	Lactococcus lactis subsp cremoris
TIFN8	55	1	39	7	536	241353	1131021	1131021	1131021	1710071	2023	1708	68	89	4	1869	Leuconostoc mesenteroides subsp. cremoris

N>200	Number of scaffolds larger than 200 nucleotides
n>N50	Number of scaffolds larger than N50 length
coverage	average coverage of the scaffolds
scaffolds	Number of scaffolds that contain stretches of undefined nucleotides as a result of the scaffolding procedure.
min	Length of the shortest scaffold
max	Length of the largest scaffold
sum	Total length of the assembly
N80	N<percentage> represents the length N such that the given percentage of the entire assembly is contained in scaffolds equal to or larger than N. Example: and N50 of 231059 nt means that 50% of the total assembly size is represented by contigs equal or lar
N50	N<percentage> represents the length N such that the given percentage of the entire assembly is contained in scaffolds equal to or larger than N. Example: and N50 of 231059 nt means that 50% of the total assembly size is represented by contigs equal or lar
N20	N<percentage> represents the length N such that the given percentage of the entire assembly is contained in scaffolds equal to or larger than N. Example: and N50 of 231059 nt means that 50% of the total assembly size is represented by contigs equal or lar
RAST	Number of ORFs predicted by RAST
prodigal	Number of complete ORF's predicted by Prodigal 2.60
5'clip	Number of ORFs predicted by Prodigal 2.60 without 5' start
3'clip	Number of ORFs predicted by Prodigal 2.60 without 3'end
fragmen	Number of ORFs predicted by Prodigal 2.60 without proper 5'start and 3'end.
Prodigal	Total number of CDSs predicted by Prodigal
Type	Lactococcus subtype as indicated by AFLP and confirmed by phylogenetic analysis based on predicted protein sequences as compared to the reference strains.

Supplementary Table 2. Statistics of Stater Metagenome Sequences

Library	Reads	Short (<64bp)	Long	Trimmed	OK	Filtered	Duplicate	Avg length
Nizo_run_1.1	607576	20	0	347735	259821	553450	16	596
Nizo_run_1.2	531399	26	0	404486	126887	458739	8	599

Library	Name of the sequence library
Reads	Number of reads in the library
Short	number of reads shorter than 64 bp (excluded from the final set)
Long	number of reads longer than 1000nt (These are considered artifacts in Roche 454 sequencing)
Trimmed	number of reads trimmed by sffToCA at default settings
OK	number of reads that did not need trimming
Filtered	number of reads left after filtering by sffToCA
Duplicate	number of duplicate reads
Avg Length	Average Length of the reads (expected target length was 500)

Supplementary Table 4. The core gene set that are present as single copy on the genome sequences of TIFN1-8 and four reference *L. lactis* Strains

OG	Annotation
OG_00027	pyrroline-5-carboxylate reductase [Lactococcus lactis subsp. lactis KF147]
OG_00028	anthranilate synthase component I [Lactococcus lactis subsp. lactis KF147]
OG_00029	amidophosphoribosyltransferase [Lactococcus lactis subsp. lactis KF147]
OG_00030	tRNA (Guanine-N(1)-)-methyltransferase [Lactococcus lactis subsp. lactis KF147]
OG_00031	glutamine synthetase [Lactococcus lactis subsp. lactis KF147]
OG_00032	universal stress protein A [Lactococcus lactis subsp. lactis KF147]
OG_00033	ATP-dependent endopeptidase Clp proteolytic subunit ClpP [Lactococcus lactis subsp. lactis KF147]
OG_00034	hypothetical protein LLKF_0598 [Lactococcus lactis subsp. lactis KF147]
OG_00035	protein-(glutamine-N5) methyltransferase [Lactococcus lactis subsp. lactis KF147]
OG_00036	general stress protein/S1-type RNA-binding domain-containing protein [Lactococcus lactis subsp. lactis KF147]
OG_00037	hypothetical protein LLKF_0005 [Lactococcus lactis subsp. lactis KF147]
OG_00038	lipoate-protein ligase [Lactococcus lactis subsp. lactis KF147]
OG_00039	GTP pyrophosphokinase/guanosine-3,5-bis(diphosphate) 3-pyrophosphohydrolase [Lactococcus lactis subsp. lactis KF147]
OG_00040	hypothetical protein LLKF_0117 [Lactococcus lactis subsp. lactis KF147]
OG_00041	deoxyuridine 5′-triphosphate nucleotidohydrolase [Lactococcus lactis subsp. lactis KF147]
OG_00043	hypothetical protein LLKF_0226 [Lactococcus lactis subsp. lactis KF147]
OG_00044	cobalt ABC transporter ATP-binding protein [Lactococcus lactis subsp. lactis KF147]
OG_00045	O-sialoglycoprotein endopeptidase [Lactococcus lactis subsp. lactis KF147]
OG_00046	cobalt ABC transporter permease [Lactococcus lactis subsp. lactis KF147]
OG_00047	peptide ABC transporter ATP-binding protein [Lactococcus lactis subsp. lactis KF147]
OG_00048	NADPH-dependent FMN reductase [Lactococcus lactis subsp. lactis KF147]
OG_00049	two-component response regulator [Lactococcus lactis subsp. lactis KF147]
OG_00050	hypothetical protein LLKF_0475 [Lactococcus lactis subsp. lactis KF147]
OG_00051	DegV family fatty acid-binding protein [Lactococcus lactis subsp. lactis KF147]
OG_00052	Myo-inositol-1(or 4)-monophosphatase [Lactococcus lactis subsp. lactis KF147]
OG_00053	O-succinylbenzoic acid--CoA ligase [Lactococcus lactis subsp. lactis KF147]
OG_00054	tRNA (m(7)G46) methyltransferase [Lactococcus lactis subsp. lactis KF147]
OG_00055	mannose-6-phosphate isomerase [Lactococcus lactis subsp. lactis KF147]
OG_00056	acetyl-CoA carboxylase,carboxyl transferase subunit alpha [Lactococcus lactis subsp. lactis KF147]
OG_00057	alkaline phosphatase superfamily protein [Lactococcus lactis subsp. lactis KF147]
OG_00058	dipeptidase [Lactococcus lactis subsp. lactis KF147]
OG_00059	gamma-D-glutamyl-meso-diaminopimelate peptidase [Lactococcus lactis subsp. lactis KF147]
OG_00060	hypothetical protein LLKF_0991 [Lactococcus lactis subsp. lactis KF147]
OG_00061	ribosomal large subunit pseudouridine synthase D [Lactococcus lactis subsp. lactis KF147]
OG_00062	50S ribosomal protein L27 [Lactococcus lactis subsp. lactis KF147]
OG_00063	UDP-N-acetylmuramyl tripeptide synthase [Lactococcus lactis subsp. lactis KF147]
OG_00064	xanthine phosphoribosyltransferase [Lactococcus lactis subsp. lactis KF147]
OG_00065	spermidine/putrescine ABC transporter permease [Lactococcus lactis subsp. lactis KF147]
OG_00066	imidazole glycerol phosphate synthase cyclase subunit [Lactococcus lactis subsp. lactis KF147]
OG_00067	acyltransferase, MBOAT family [Lactococcus lactis subsp. lactis KF147]
OG_00068	nucleoside ABC transporter ATP-binding protein [Lactococcus lactis subsp. lactis KF147]
OG_00070	GntR family transcriptional regulator [Lactococcus lactis subsp. lactis KF147]
OG_00071	hypothetical protein LLKF_1745 [Lactococcus lactis subsp. lactis KF147]
OG_00072	3-dehydroquinate dehydratase [Lactococcus lactis subsp. lactis KF147]
OG_00073	two-component system sensor histidine kinase [Lactococcus lactis subsp. lactis KF147]
OG_00074	phosphate ABC transporter permease [Lactococcus lactis subsp. lactis KF147]
OG_00076	tributyrin esterase [Lactococcus lactis subsp. lactis KF147]
OG_00077	SUF system FeS cluster assembly protein ATP-dependent transporter SufC [Lactococcus lactis subsp. lactis KF147]
OG_00078	excinuclease ABC subunit A [Lactococcus lactis subsp. lactis KF147]
OG_00079	cysteinyl-tRNA synthetase [Lactococcus lactis subsp. lactis KF147]

OG_00080	aminopeptidase C [Lactococcus lactis subsp. lactis KF147]
OG_00082	hsp33 family chaperonin [Lactococcus lactis subsp. lactis KF147]
OG_00083	galactokinase [Lactococcus lactis subsp. lactis KF147]
OG_00084	uridylate kinase [Lactococcus lactis subsp. lactis KF147]
OG_00085	Xaa-Pro dipeptidyl-peptidase [Lactococcus lactis subsp. lactis KF147]
OG_00086	adenylate kinase [Lactococcus lactis subsp. lactis KF147]
OG_00087	50S ribosomal protein L14 [Lactococcus lactis subsp. lactis KF147]
OG_00088	30S ribosomal protein S10 [Lactococcus lactis subsp. lactis KF147]
OG_00089	prolyl-tRNA synthetase [Lactococcus lactis subsp. lactis KF147]
OG_00090	methyltransferase [Lactococcus lactis subsp. lactis KF147]
OG_00091	valyl-tRNA synthetase [Lactococcus lactis subsp. lactis KF147]
OG_00092	D-alanyl-D-alanine serine-type carboxypeptidase [Lactococcus lactis subsp. lactis KF147]

Supplementary Table 7. The Metabolic genes coded only by Lactococcus pan-genome or only by TIFN8 Leuconostoc Strain

COG	Enzyme	Pathways involved	Lactococcus	Leuconostocs
COG0183	Acetyl-CoA acetyltransferase	Fatty acid metabolism.Valine, leucine, isoleucine degradation. Benzoate degradation. Ethylene benzene degradation	**X**	
COG 2235	Arginine deiminase	Arginine and proline metabolism Benzoate degradation	**X**	
COG1053	Succinate dehydrogenase/fumarate reductase, flavoprotein subunit	2,4 Dichloro benzoate degradation. Butanoate metabolism, TCA Cycle	**X**	
COG0075	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase	Alanine, aspartate, glutamate metabolism. Glycine, serine, threonine metabolism.	**X**	
COG0191	Fructose/tagatose bisphosphate aldolase	Glycolysis. Pentose phophate pathway. Fructose and mannose metabolism.	**X**	
COG0371	Glycerol dehydrogenase and related enzymes	Glycerolipid metabolism	**X**	
COG0297	Glycogen synthase	Starch and Sucrose Metabolism	**X**	
COG1554	Trehalose and maltose hydrolases (possible phosphorylases)	Starch and sucrose metabolism	**X**	
COG0800	2-keto-3-deoxy-6-phosphogluconate aldolase	Pentose phosphate pathway. Arginine and proline matabolism	**X**	
COG0076	Glutamate decarboxylase and related PLP-dependent proteins	Alanine, aspartate, glutamate metabolism. Glycine, serine, threonine metabolism. Butanoate metabolism. Panthatonate Coa biosynthesis	**X**	
COG0108	3,4-dihydroxy-2-butanone 4-phosphate synthase	Riboflavin metabolism	**X**	
COG0294	Dihydropteroate synthase and related enzymes	Folate Biosynthesis	**X**	
COG0160	4-aminobutyrate aminotransferase and related aminotransferases	Alanine, aspartate, glutamate metabolism. Glycine, serine, threonine metabolism. Butanoate metabolism. Propanoate metabolims,		**X**
COG0161	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase	Alanine, aspartate, glutamate metabolism. Glycine, serine, threonine metabolism. Butanoate metabolism. Propanoate metabolims,		**X**
COG1027	Aspartate ammonia-lyase	Alanine, aspartate, glutamate metabolism. Nitrogen metabolism		**X**
COG0105	Nucleoside diphosphate kinase	Purine, pyrimidine metabolism		**X**

Supplementary Table 6. Strain-specific qPCR primers

Primer	Annotation of target gene	Primer Sequences
TIFN1-5	Hypothetical protein	5`-AAAAGAAAAAGGCTGGCTGAAA- 3`
		5`-GCACCGAGGTCAGACCAAGA- 3`
TIFN2-4	Ferrochelatase	5`-GCTTGTGCTGCCTCTTTAATTTC-3`
		5`-TTTACCCGAGCGTCTATTAGCAA-3`
TIFN3	Conserved hyphotetical protein LacR E2 of SK11	5`-CACGTCGTCAAACTGGTTACTCA-3`
		5`-TGAACAGGCTTGCTCTTATCTATGA-3`
TIFN6	Polysaccharide biosynthesis protein	5`-TCAGGCAAGGCGAGCAA-3`
		5`-GTATGCGGAAGAAAATTCATGGA-3`
TIFN7	Hyphotetical protein	5`-GGGAGCAAGCCTATCCTCACT-3`
		5`-ACTGCCATCTTTGGTGATTTCTAAA-3`
TIFN8	Accessory secretory protein Asp1	5`-AATGAGCGACATCAACAACAAGTAA-3`
		5`-ATGGCGGAACGATAGCAACT-3`

The supplementary table 3 and 5 that encompass the data for the COG Categories and the annotations of OGs conserved exclusively for TIFN 1-to-7 *L. lactis* strains and the data for strain-specific OGs of TIFN 1-to-7 *L. lactis* strains respectively are not provided in the thesis due to extensive information, which is not readily presented in a concise table format.

These supplementary tables can be found on the online access link of the publication where the supplementary tables were published, provided by the ISME Journal. (http://www.nature.com/ismej/journal/v7/n11/full/ismej2013108a.html)

References for Supplementary Figures and Note

1. **Terzaghi BE, Sandine WE.** 1975. Improved medium for lactic streptococci and their bacteriophages. Appl Microbiol **29:**807-813.
2. **Reddy MS, Vedamuthu ER, Washam CJ, Reinbold GW.** 1972. Agar medium for differential enumeration of lactic streptococci. Appl Microbiol **24:**947-952.
3. **Rademaker JL, Herbet H, Starrenburg MJ, Naser SM, Gevers D, Kelly WJ, Hugenholtz J, Swings J, van Hylckama Vlieg JE.** 2007. Diversity analysis of dairy and nondairy *Lactococcus lactis* isolates, using a novel multilocus sequence analysis scheme and (GTG)5-PCR fingerprinting. Appl Environ Microbiol **73:**7128-7137.
4. **Limsowtin GKY, Terzaghi, B.E.** 1976. Agar medium for the differentiation of "fast" and "slow" coagulating cells in lactic streptococcal cultures. N.Z.J. Dairy Science and Technology **11:**2.
5. **Siezen RJ, Bayjanov JR, Felis GE, van der Sijde MR, Starrenburg M, Molenaar D, Wels M, van Hijum SA, van Hylckama Vlieg JE.** 2011. Genome-scale diversity and niche adaptation analysis of *Lactococcus lactis* by comparative genome hybridization using multi-strain arrays. Microb Biotechnol **4:**383-402.
6. **Stadhouders J, Leenders, G.J.M.** 1984. Spontaneously developed mixed-strain cheese starters. Their behaviourstowards phages, and theri use in Dutch cheese industry. Netherlands Milk and Dairy Journal **38:**25.
7. **Simons G, Nijhuis M, de Vos WM.** 1993. Integration and gene replacement in the *Lactococcus lactis* lac operon: induction of a cryptic phospho-beta-glucosidase in LacG-deficient strains. Journal of bacteriology **175:**5168-5175.
8. **Stadhouders J.** 1986. The control of cheese starter activity. Netherlands Milk and Dairy Journal **40:**18.

Chapter 4

Use of Propidium Monoazide for the Selective Profiling of Live Microbial Populations during Cheese Ripening

Oylum Erkuş, Victor C.L. de Jager, RenskeT.C.M. Geene , Ingrid van Alen-Boerrigter, Lucie Hazelwood, Sacha A.F.T. van Hijum, Michiel Kleerebezem, Eddy J. Smid

This chapter is submitted for publication.

Abstract

Microbial community profiling on the basis of DNA extracted from environmental samples can be confounded by the presence of DNA derived from dead cells. Selective amplification of DNA extracted from the live fraction of the microbial community by propidium monoazide (PMA) treatment could circumvent this problem. In this chapter, we evaluate the use of PMA for community profiling of complex environmental samples with many membrane-compromised cells in the sample matrix using a well-defined Gouda cheese manufacturing and ripening model. The community dynamics of mixed-strain Gouda cheese starter cultures were determined with a genetic lineage-specific quantitative PCR (qPCR) and metagenomics approaches with and without the application of PMA. PMA effectively inhibited the amplification of DNA derived from membrane-compromised cells and enhanced the selective analysis of the viable population of the cheese starter, leading to the community profiles that are more accurately reflecting community composition compared to untreated samples. Additionally, following PMA treatment, a two-step whole genome amplification based sequencing methodology was evaluated for selectively enhanced metagenome sequencing of the viable microbial communities present in cheese.

Introduction

Profiling tools to reveal the composition of microbial communities, such as denaturing gradient gel electrophoresis (DGGE), 16S rRNA gene profiling, 16S rRNA targeted quantitative PCR approaches (qPCR) and metagenomics are based on deducing information from the total DNA pool of the community. However, the community composition information commonly fails to reflect the presence of subpopulations with different viability state (1) and/or the persistence of DNA in the sample matrix although the cells might have (long) lost their integrity (2, 3). A recent study comparing the microbial community profiles of human faecal sample fractions that were sorted based on the bacterial RNA concentration revealed that the profiles of the active fraction in all samples differed highly from the whole community profiles (4). It has previously been reported that up to 60% of faecal microbiota may be dead or damaged, and sequencing of the dominant DGGE bands that were separately obtained from the sorted viable, dead, and damaged cells revealed remarkable phylogenetic diversity differences among the sorted sub-populations (1).

Propidium monoazide (PMA) treatment is increasingly used for the discrimination of live and dead cells in microbial communities (5), and offers a promising approach to overcome the limitations of DNA-based community profiling methods. PMA is a modified version of propidium iodide (PI), a fluorescent dye that is routinely used for staining cells with compromised membranes and therefore considered as damaged or dead cells. The iodide group of PI was replaced with an azide group that covalently links to DNA (and other organic moieties) upon photo activation and inhibits DNA amplification by PCR (6). The dye is membrane impermeable, modifying only the DNA freely present in the sample matrix or in permeable cells, but not of the intact (viable) cells that are surrounded by PI-impermeable membranes. Therefore, pretreatment of the (environmental) sample with PMA prior to DNA extraction and subsequent PCR allows the selective amplification of the DNA derived from viable cells in the community (7).

PMA was successfully applied to selectively determine the viable microbial community using different profiling techniques on environmental samples such as 16S rRNA gene pyrosequencing of sea and canal water (8), 16S rRNA based DGGE profiling of marine sediment (7) and qPCR based quantification of probiotic strains in human feces (9). The approach using PMA for selective detection of live cells could also be useful for the study of community dynamics in food fermentations, where

there is substantial variation in cell viability and where community composition may change during the fermentation process due to succession or differential loss of viability. Gouda cheese manufacturing employs mixed-strain starter cultures that are blends of closely related strains of the species *Lactococcus lactis* and *Leuconostoc mesenteroides* (This thesis, Chapter 3). The composition of the community is expected to be dynamic under the cheese manufacturing regime applied, and we have previously described that the major community changes occur after the brining step (This thesis, Chaper 3). A substantial decrease in viable cell counts determined by colony forming unit enumeration was observed, in combination with an increase in relative abundance of salt and acid resistant strains during the cheese ripening period. The balance between viable (live) and membrane-compromised (dead and damaged) cells in cheese is crucial for the final product texture and flavor formation (10). The dying cells, that have increasingly permeable cell membranes, are considered essential for the release of intracellular peptidases into the cheese matrix, which convert the casein derived peptides into amino acids (11). The comparison of *L. lactis* strains with varying autolytic capacities revealed that the decrease in cell viability was highly correlated with the release of intracellular peptidases like an X-prolyl dipeptidyl aminopeptidase and an aminopeptidase with broad substrate specificity (12). Promoting *in situ* cell lysis has been considered a strategy to increase the free amino acid pools that are available for flavor formation (13, 14), which would accelerate cheese ripening. However, also intact (live) cells that contain the catabolic enzymes of complete metabolic pathways are important for the formation of key aroma compounds derived from degradation of amino acids, carbohydrates and/or lipids (15). Notably, aroma compounds are produced as side- or intermediate-products when intact cells metabolize the amino acids for biosynthetic and catabolic purposes, for the regeneration of co-substrates, or for survival under environmental stresses i.e. internal pH regulation (16). Therefore, understanding the pivotal effect of intact versus membrane-compromised cells on the aroma formation requires the discrimination of live populations from the dead and damaged populations of the starter culture community during cheese ripening.

In this chapter, PMA treatment was incorporated into the time resolved analysis of the microbial community during Gouda cheese production, using a mixed-strain starter culture. PMA treatment procedures were optimized on pure DNA and on *in vitro* cultured *L. lactis* cultures. The community dynamics of the different microbial lineages in a complex Gouda cheese starter were studied using lineage-specific qPCRs during the initial manufacturing and subsequent ripening periods in samples that

were treated or not-treated with PMA. Finally, the PMA treated total DNA isolated from 6-weeks ripened cheese was subjected to whole-genome amplification based metagenome sequencing to evaluate the use of PMA for the selective analysis of viable microbial communities in difficult-to-process environmental samples (cheese), which are known to contain significant numbers of dead cells with compromised membranes and /or free DNA in the matrix.

Materials and Methods

Bacterial Strains, Starter Cultures

The Gouda type starter culture named "UR" (17) used in this study was selected because of its long history of use in cheese manufacturing. Therefore, the diversity present in this undefined mixed culture was assumed to be shaped in adaptation to propagation regime applied. The microbial diversity present in this starter culture has been studied in chapter 3, and determined to be consist of 7 genetic lineages of *L. lactis* and 1 genetic lineage of *Ln. mesenteroides*(This thesis). Representative strains for each genetic lineage (designated as TIFN 1 to 8) have been subjected to whole genome sequencing, comprising 5 *L. lactis* ssp.*cremoris* strains (TIFN1, 3, 5, 6, and 7), *2 L. lactis* ssp. *lactis* biovar. *diacetylactis* strains (TIFN 2 and 4), and 1 *Ln. mesenteroides* strain (TIFN8)(This thesis, chapter 3). The starter culture was propagated two times at 20°C for 20 hours in sterile skimmed milk prior to cheese making. The propagated culture and the genomic DNA of *L. lactis* ssp. *cremoris* TIFN7 were used in PMA treatment optimizations, and the strain was propagated in lactose-M17 broth (Merck, Germany) at 30°C overnight prior to experiments.

Quantitative PCR

16S rRNA targeting methods do not provide sufficient discriminative resolution to detect and quantify different genetic lineages of the same species. Therefore, the lineage-specific genetic markers to be employed for lineage-specific quantitative detection using qPCR approach, were identified by orthologous group analysis of the whole genome sequences of representative TIFN strains as described in chapter 3 (This thesis). Lineage-specific primer sets (6 sets) were designed targeting genetic markers specific for genetic lineages 3, 6, 7, and 8 separately, and for the lineage 1&5, and the lineage 2&4 as groups of lineages. The genetic lineage specificity of

the primer sets were experimentally validated with qPCR for the conservation of genetic markers in the complete set of strains within each genetic lineage. The qPCR amplifications were performed in 20 µl of final reaction volume, combining 10 µl of SYBRgreen master mix (Applied Biosystem, UK) with 0,2 µl of each primer (10 µM), and 2 µl of DNA template. PCR amplification (7500 Fast Real-Time PCR System, Applied Biosystem, Warrington, UK) was initiated with 10 min of initial denaturation at 95°C, followed with 40 cycles of 15 s of denaturation at 95°C and 1 min of annealing and extension at 60°C. The Ct values were calculated with the default settings.

PMA Treatment of Genomic DNA

Chromosomal DNA of the *L. lactis* TIFN7 (representative strain of genetic lineage 7) was isolated with DNeasy Tissue and Culture DNA Isolation Kit (QIAGEN, Hilden, Germany) according to manufacturer`s protocol and diluted with water to a final concentration of 1 ng/µl with DNAse free water. Propidium monoazide (Biotium Inc, Hayward, California) was dissolved in 20% of DMSO (Merck, Germany) to a 20 mM stock concentration. Genomic DNA solution (1 ng/µl) was aliquoted to 1 ml portions in microcentrifuge tubes and PMA was added to the aliquots to the final concentrations of 3, 30, 50, and 100 µM in duplicate. The samples were mixed and incubated in the dark and at room temperature for 5 minutes and exposure to light at 20 cm distance while cooled on ice for increasing exposure times (0, 1, 2, 3, 4, 5 min), using a 650 W halogen light bulb as a light source with the following specifications (120 V, 650 W, FCW, 3200 K, General Electric Co., OH). Aliquots of 2 µl of PMA treated samples used as template DNA in quantitative PCR (7500 Fast Real-Time PCR System, Applied Biosystem, UK) using the primers targeting the genetic lineage 7. DNA (2 µl with concentration of 1 ng/µl) that was not treated with PMA was taken as a negative, normalization control during qPCR quantification, by subtracting the Ct value of PMA-treated sample from the Ct value of the untreated sample.

PMA Treatment of *Lactococcus lactis* Cells

L. lactis TIFN7 was cultured overnight in lactose-M17 (L-M17) liquid broth at 30ºC. The overnight culture was subcultured in L-M17 until the culture had reached an optical density of 1 at 600 nm. This exponential phase culture was used for harvesting the live cells. Additionally, 500 µl of exponential phase culture was mixed with 25

μl of n-butanol (analytical grade, J.T. Baker Chemicals, Deventer, The Netherlands) followed by incubation at 25ºC for 20 min for the preparation of damaged (permeable) cells, and 500 μl of exponential phase culture was heat treated at 80ºC for 5 min for the preparation of dead cells. All types of cell preparations were centrifuged at 6800*g for 5 min at 4°C (Eppendorf Centrifuge 5417R, NY), re-suspended in 500 μl of Na-citrate buffer (2% (w/v), pH 8.0), labelled with LIVE/DEAD *Bac*Light™ Bacterial Viability and Counting Kit according to manufacturer`s protocol (Invitrogen) and finally sorted by BD FACSAria III Cell Sorter (BD Biosciences, NJ) to further purify the preparations for the live, damaged and dead cell content. 10^8 cells were sorted from each live, damaged and dead cell fractions, harvested by centrifugation at 6800*g for 10 min at 4°C (Eppendorf Centrifuge 5417R, NY), and re-suspended in 1 ml of Na-citrate buffer (2 % w/v), pH 8.0). PMA was added to a final concentration of 50 μM, followed by thorough mixing (vortex) and incubation in the dark for 5 min at room temperature, and subsequently placed on ice and exposed to light for 5 min (see above for light exposure specifications). Subsequently, the cells were harvested by centrifugation at 6800*g for 5 min at 4°C (Eppendorf Centrifuge 5417R, NY). The supernatant was discarded to remove excess PMA, and the cell pellets were re-suspended in 380 μl of Na-citrate buffer (2% (w/v), pH 8.0). 10^8 cells sorted from live, dead and damaged cell preparations without any PMA treatment were used as negative controls. Total DNA was extracted from all cell preparations using the DNeasy Tissue and Culture DNA Isolation Kit (QIAGEN, Hilden, Germany), according to the manufacturer`s protocol. The DNA preparations were diluted 10-fold in Na-citrate buffer (2% w/v), pH 8.0) before it was used in the qPCR assays. All samples were prepared in triplicate.

Lab-scale Cheese Manufacturing and Samplings

Mini and micro scale cheese models are useful tools mimicking the complex cheese environment for high throughput screenings of starter cultures (18). Time course of the microbial community dynamics of Gouda cheese manufacturing and ripening was studied in a mini scale cheese model. 3L of thermized (13.7 s at 68°C), bactofugated, and pasteurized (13.7 s at 73°C) bovine milk was standardized for fat/protein ratio of 1.085 with a final fat, protein and lactose content of 3.56% , 3.28% and 4.41% (w/w), respectively. The starter culture was propagated two times at 20°C for 20 hours in sterile skimmed milk prior to cheese making. The mini cheeses were manufactured from 3 L of milk and with 1% of starter culture addition as described by Bachmann *et al.* (19). At the end of the pressing, the mini cheeses were incubated

at 30°C until the pH was reduced to the range of 5.4-5.6. Subsequently, they were brined in 19% (wt/vol) sodium chloride solution for 4 hours at 13°C, flushed with 0.8 atmospheric pressure of nitrogen, vacuum-sealed in a plastic foil and ripened at 13°C for 6 weeks. The cheeses manufactured had a final moisture contents in the range of 42-44% and pH values ranging between 5.4 and 5.2. The cheeses were manufactured in duplicate and sampled throughout the manufacturing process at 7 time points: (i) from the propagated starter culture just before cheese milk inoculation, (ii) after pressing of the cheeses, (iii) after 24 hours just before brining, (iv) after brining, and from (v) 2-weeks, (vi) 4-weeks, and (viii) 6-weeks ripened cheeses during the ripening period.

Fluorescence Activated Cell Counting during Cheese Manufacturing and Ripening

2.5 gr of propagated starter culture or cheese sample was mixed with 2% (wt/vol) Na-citrate buffer (pH 8.0) at 45°C in 25 ml final volume, and homogenized for 5 min with the stomacher. The homogenate was centrifuged at 8000*g for 10 min at 4°C (Beckman-Coulter, Avanti JE Centrifuge, JA20 rotor, Fullerton, CA). The supernatant and fat layers were removed and bacterial cell pellets were washed three times and resuspended in 1 ml of the same buffer. 10 µl of cell suspensions were differentially stained using the LIVE/DEAD® *Bac*Light™ Bacterial Viability and Counting Kit according to manufacturer`s protocol (Invitrogen, The Netherlands) and the cells in live, dead, damaged fractions were quantified using the BD FACSAria III Cell Sorter (BD Biosciences, NJ) according to manufacturer's instructions. Analysis was performed in duplicate.

Total DNA Isolations during Cheese manufacturing and Ripening

For total DNA isolation, cells were recovered from milk culture or cheese matrix by Na-citrate buffer washings as described in the paragraph dealing with FACS analyses. Total DNA was isolated in duplicate from cheese derived microbial suspensions with and without PMA treatment. For the treated samples, PMA was added to 1 ml of cell suspension in 2% (wt/vol) Na-citrate solution (pH 8.0) to a final concentration of 50 µM. The cell suspension was thoroughly mixed (vortex), incubated in the dark for 5 min at room temperature, followed by placing the suspension on ice and exposure to light for 5 min (for details about light exposure, see above). The cells were harvested by centrifugation at 6800*g and 4°C for 5 min (Eppendorf Centrifuge 5417R, NY),

and the supernatant was discarded to remove excess PMA. Subsequently, the cells were washed once more in 1 ml of Na-citrate (pH 8.0) buffer, and the total DNA was isolated from the final, fat-free cell pellets using QIAGEN DNeasy Blood and Tissue Isolation Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocols with the following modifications. The cells were lysed in 1 ml instead of 380 µl of lysis buffer for 1h at 37°C. Accordingly, all subsequently described volumes in the manufacturers' protocols were increased 2.5-fold to accommodate this larger starting volume. Subsequently, the lysed cell suspension was treated with 50 µl of Proteinase K (600 maul/ ml, QIAGEN, Hilden, Germany) for 1 hour at 56°C and with 10 µl of RNase (100 mg/ ml, QIAGEN, Hilden, Germany) for 30 min at room temperature. The resulting DNA solution was purified on a spin column (QIAGEN, Hilden, Germany), including two extra on-column washing steps to remove residual milk derived impurities. DNA was eluted with 300 µl of nuclease free water in three steps (100 µl in each step), and the DNA concentration was determined using a Nanodrop Spectrophotometer (Coleman technologies Inc., FL). The samples containing total DNA were 10-fold diluted prior to their use as template in the qPCR.

Whole Genome Amplification of Total DNA and Sequencing

The total DNA preparation of PMA treated 6-weeks ripened cheese (the replicate B) was subjected to whole genome amplification (WGA) to enrich the total amount of DNA from live cell fraction in cheese prior to sequencing. Whole genome amplifications were performed in triplicate using 10 ng/µl of total DNA isolated from cheese according to manufacturer's protocol (GenomePlex Whole Genome Amplification (WGA) Kit, Sigma-Aldrich, MO). DNA concentrations were determined using Nanodrop Spectrophotometer (Coleman technologies Inc., FL), and size distributions of amplification products were evaluated with 0.8% agarose gel electrophoresis. WGA products were further purified with PCR purification kit according to manufacturer's protocol (MSB Spin PCRapace, STRATEC Molecular GmbH, Berlin, Germany). The initial total DNA preparation and the purified WGA products were subjected to lineage-specific qPCR in order to determine the community compositions before and after the WGA reaction to check for the amplification biases. Finally, purified WGA products were pooled and subjected to paired-end whole genome sequencing using Illumina HiSeq technology (GATC, Konstanz, Germany).

The genetic lineage prevalence in the metagenome of the cheese samples was quantified based on the genetic lineage-specific orthologous group analysis as described in chapter 3with the following modifications. The lineage-specific orthologous groups were aligned to the metagenome reads using Bowtie2 Version 2.0.6 (20)om/office/word/2010/wordprocessingInk to enumerate the coverages within the metagenome dataset, followed by the conversion of alignment information in to variant call format (VCF) files (21) using Samtools (22). The coverage for each lineage-specific orthologous group was calculated by taking the median coverage value from the central 50% of the open reading frame (ORF) sequence as reported in the combined depth across samples (DP) field of the VCF file in order to correct for reads that aligned at the edge of a given ORF. For each lineage, the average of the median coverages of the lineage-specific orthologous groups was calculated, signifying the read coverage of lineage-specific ORFs. For all lineages, these average lineage coverages were added to a cumulative coverage value: the total average read coverage over all lineages. Finally, the lineage-specific coverages normalized through dividing by the cumulative coverage were used as measure of the relative abundance of a given genetic lineage in the microbial community.

Results and Discussion

Effect of PMA on the PCR Amplification of *L. lactis* DNA

A DNA preparation of *L. lactis* TIFN7 was used as reference material to reflect DNA derived from dead cells, to optimize the PMA concentration and the light exposure time for complete inhibition of DNA amplification. The DNA was treated with 4 different PMA concentrations (3, 30, 50, and 100 μM), and exposed to a fixed light source using different exposure times (0, 1, 2, 3, 4, 5 minutes). PMA treated and untreated DNA material was used as a qPCR template with the primer pair specific for the genetic lineage 7, to which *L. lactis* TIFN7 belongs, to determine the difference of threshold cycle (Ct) number required for detection (Fig. 1A). The Ct differences between PMA treated and untreated samples were 8 to 10 for 3 or 30 μM of PMA. A further increase of PMA concentration to 50 or 100 μM led to a Ct difference of 18 or a complete loss of detection, respectively. The latter lack of detection was due to direct inhibition of the amplification reaction independent of the PMA reaction with the DNA template, because light exposure was not required to elicit this inhibition (Fig. 1A). Direct PCR inhibition by PMA did not occur on the samples derived from

the cheese matrix or laboratory cultures because the excess PMA was removed with the supernatant in the washing step prior to DNA isolation (data not shown). With respect to light exposure time, Ct values increased as the exposure time increased up to 5 minutes for the samples treated with 3 μM of PMA concentration, but the increase in the range of 2-to-5 minutes exposure time for the samples treated with 30, 50, and 100 μM of PMA concentration did not result in a significant change in the Ct values obtained. Therefore, all subsequent experiments employed a PMA concentration of 50 μM, combined with a 5 minutes light exposure time. In a previous study, Nocker et al. (6) showed that Ct values increased by 12 cycles during the amplification *E.coli* DNA when 30 μM of PMA was applied in 2 minutes exposure time and this value compared well with the Ct difference observed for *L. lactis* DNA under the same conditions. On the other hand, the inverse correlation between Ct difference and PMA concentration observed by Nocker et al. (6) was not seen on *L. lactis* DNA in our study.

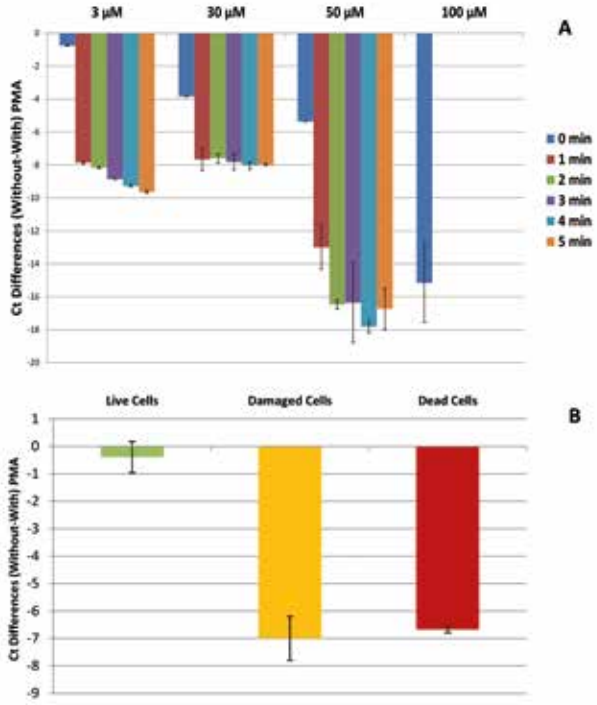


Figure 1. PMA optimizations on genomic DNA and *L. lactis* cell cultures. Panel A: Effect of different PMA concentrations and increasing light exposure times on the inhibition of *L. lactis* TIFN7 DNA amplification. Error bars represent the standard deviations from 2 independent replicates. Panel B: The effect of cell viability state on PMA inhibition of qPCR amplification of *L. lactis* DNA. PMA treatment of FACS sorted cells was performed with 50μM of PMA with 5 min. of light exposure. Error bars represent the standard deviations from 3 independent replicates.

Validation of PMA Treatment on *L. lactis* Cells

The penetration of PMA into cells with different viability states (dead, damaged, or live) has not yet been studied in detail. Therefore, the influence of PMA treatment on live, damaged and dead cell fractions of a *L. lactis* TIFN7 culture was investigated. To this end, bacterial cultures were sorted into live, dead and damaged cell-fractions by FACS, and DNA was isolated from sorted cell fractions with and without PMA treatment. The DNA preparations were used as template in qPCR targeting the genetic marker specific for the lineage that *L. lactis* strain TIFN7 belongs to, resulting in a clear PMA-treatment-induced difference in Ct values obtained for live (within one Ct interval; no significant difference) and dead/ damaged (high Ct difference) cell populations. No significant difference was observed between the detection efficiency of dead and damaged cell populations, implying an equal rate of PMA penetration and DNA modification in these cell populations. The maximum difference in Ct-value induced by PMA treatment was around 6.6, corresponding to an approximate 100-fold reduction in detected template concentration (Fig. 1B). The inhibition of the qPCR amplification due to PMA treatment of live cells was within one Ct cycle interval.

The relative quantification limit of 6.6 PCR cycles for the discrimination of dead/ damaged cells upon PMA treatment as compared to the relative Ct difference obtained with PMA treatment of pure DNA (18 cycles) indicates a reduced efficiency of PMA penetration into these cells. One reason for this reduced quantification range of PMA-qPCR observed on PMA-treated sorted cells may be the type of cell treatment. For example, Lee et al. (23) showed that PMA-qPCR quantification was not well correlated with the plate counts when mild heat treatments (incubation at 50°C-to-60°C for 5 minutes) were applied during the analysis of mixed bacterial cultures of fish origin. Additionally, the PMA-qPCR quantification of cells that were incubated at 80°C for 5 minutes still showed 1-5% of cell viability even though considerably fewer cells were recovered with plating. The 100 fold less PMA-qPCR detection of *L. lactis* cells that were heat killed at 80°C for 5 minutes in our study compared well with the findings of Lee et al. (23) and this is most probably due to insufficient PMA diffusion or physical hindering of damaged membrane in photo activation of PMA binding. In another study by Desfossés-Foucault et al.(24), the qPCR quantified cell counts of *L. lactis* pure cultures, which had dead-to-viable ratio of 10^3, was reduced only by two orders of magnitude (logarithmic scale) when the culture was PMA treated, indicative of similar limitations that we have been observed with pure *L. lactis* cultures in our study.

Analysis of Microbial Community in Cheese Using PMA Treatment

The Gouda cheese starter culture UR, whichwas used in this study is composed of 8 co-existing genetic lineages (5 *L. lactis* ssp. *cremoris*, 2 *L. lactis* ssp. *lactis* biovar. *diacetylactis* and 1 *Lc. mesenteroides* ssp. *cremoris*) (This thesis, chapter 3). Community dynamics of the viable fraction of this starter culture during cheese production has been studied using amplified fragment length polymorphism (AFLP) fingerprinting of single colonies isolated at different time points during the cheese making process (This thesis, chapter 3). Viable plate counting indicated a gradual increase of colony forming units (CFU) representing all genetic lineages up to the brining point, after which the total viable CFU count dropped 2-3 orders of magnitude. The relative contribution of each genetic lineage to the community of survivors changed after brining due to differential inactivation rates of the members of respective genetic lineages. The relative abundance of cells belonging to the genetic lineage 2 and 4 (both *L. lactis* subsp. *lactis* biovar *diacetylactis*), lineage 6 (*L. lactis* subsp. *cremoris*), and lineage 8 (*Lc. mesenteroides* subsp. *cremoris)* increased during the ripening period, while other lineages within the community virtually disappeared. The metagenomic analysis of the starter community during cheese manufacturing and ripening failed to show these community changes, which was most likely due to the fact that membrane compromised cells (dead and damaged) did not form a colony but still contributed to the total DNA isolated from the ecosystem (Fig. 2). Thereby, the metagenome sequencing approach using total DNA preparations from the cheese matrix failed to discriminate between viable and non-viable communities and provided a skewed view of the cheese community dynamics during production.

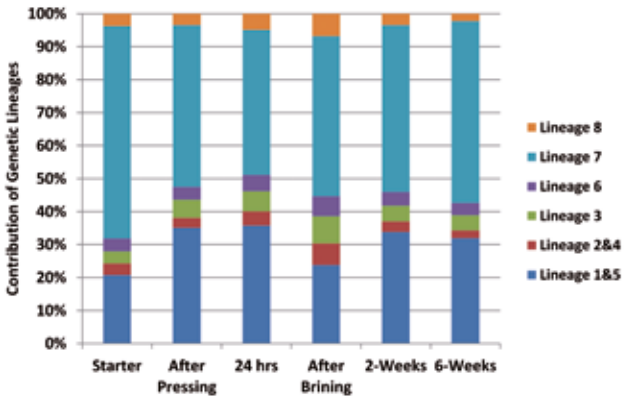


Figure 2. Community profiles deduced from the metagenome analysis of total DNAs isolated from PMA untreated cheese samples during cheese manufacturing and ripening period. The contributions of genetic lineages in the metagenome of the cheese samples were quantified based on the genetic lineage-specific orthologous group analysis.

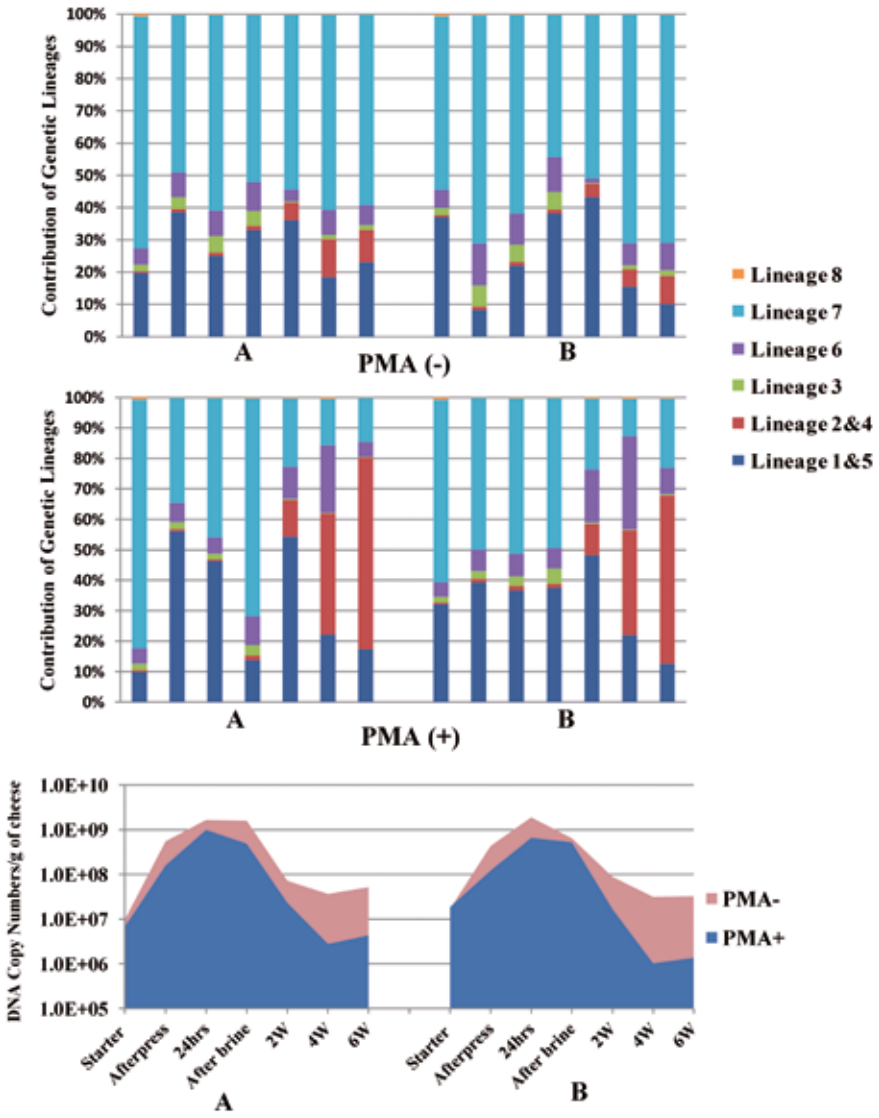


Figure 3. Community dynamics of genetic lineages during cheese manufacturing and ripening. Quantifications were performed in independent cheese duplicates (A & B) with lineage-specific qPCR on the total DNA samples that were not treated with PMA (PMA (-), 1st graph), and treated with PMA (PMA (+), 2nd graph). Total cell counts were calculated by adding up the lineage-specific qPCR counts for both PMA-treated and PMA-untreated samples (3rd graph).

In order to evaluate the added value of PMA to discriminate between viable and non-viable cells, total DNA isolations were performed with and without PMA treatment during cheese production and qPCRs targeting genetic markers of

each genetic lineage were performed using the PMA-treated and –untreated total DNA samples as template. The qPCR analyses of untreated total DNA samples indicated initial domination of cheese by *L. lactis* ssp. *cremoris* strains belonging to lineage 7 (55-75%) and lineage 1&5 (20-40%) (Fig. 3, Panel A). The rest 5-10% of the community consisted of lineage 3 and lineage 6 strains (*L. lactis* ssp. *cremoris*) strains with very little contribution of lineage 2 & 4 strains (*L. lactis* ssp. *lactis* biovar. *diacetylactis*). Minor community composition changes were observed during the ripening including a 5-10% increase in relative abundance of the lineage 3 (*L. lactis* ssp. *cremoris*), and lineage 2&4 (*L. lactis* ssp. *lactis* biovar. *diacetylactis*) populations. The community dynamics observed by qPCR analyses of untreated total DNA matched well with the previous metagenome analyses of the total DNA isolated without PMA treatment (Fig. 2), and did not match at all to the recovered viable community compositions established with AFLP profiling in chapter 3 (This thesis). Furthermore, the qPCR analyses of PMA-treated total DNA samples delivered quantitatively and qualitatively the same dynamics of genetic lineages (Fig. 3), as observed in AFLP profiling of the recovered viable community (This thesis, chapter 3). The community compositions up to the brining time-point were very similar to the ones obtained from untreated samples (Fig. 3, Panel B). However, the community compositions detected at subsequent time points during the ripening were strongly enriched for the lineage 2&4 cells *(L. lactis* ssp. *lactis* biovar. *diacetylactis)* with a gradual increase over time of their relative abundance from 2% to 75% of the overall community detected.

In addition to the relative community composition over time, the qPCR data were also used to estimate the total cell counts during cheese manufacturing, by conversion of the Ct values to DNA copy numbers per gram of cheese for each of the lineages analysed. The DNA copy numbers calculated from cumulative qPCR data were not significantly different in PMA-treated and -untreated samples that were taken up to the brining stage (Fig.3, Panel C). However, the impact of PMA treatment on the cumulative DNA copy numbers detected at later stages confirmed the decrease of overall viability of the microbial community. The decrease in viability of the overall community and observed viability changes during cheese ripening was supported with flow cytometer analysis (Fig. 4). The initial cell count in the cheese milk was found to be 10^7 cells/gr of cheese, and increased to 10^9 at the start of the brining procedure. During this initial phase of cheese production, the contribution of live cells to the cheese community was between 90% and 98%. However, during brining and the subsequent ripening period, the total number of cells decreased approximately

10-fold and the relative contribution of the viable fraction to the overall community decreased gradually from 90 to 5% (Fig. 4). In addition, the contribution of the dead cells in the population declined during extended ripening. The damaged fraction of cells represented the majority in the bacterial community recovered from the cheese matrix during ripening.

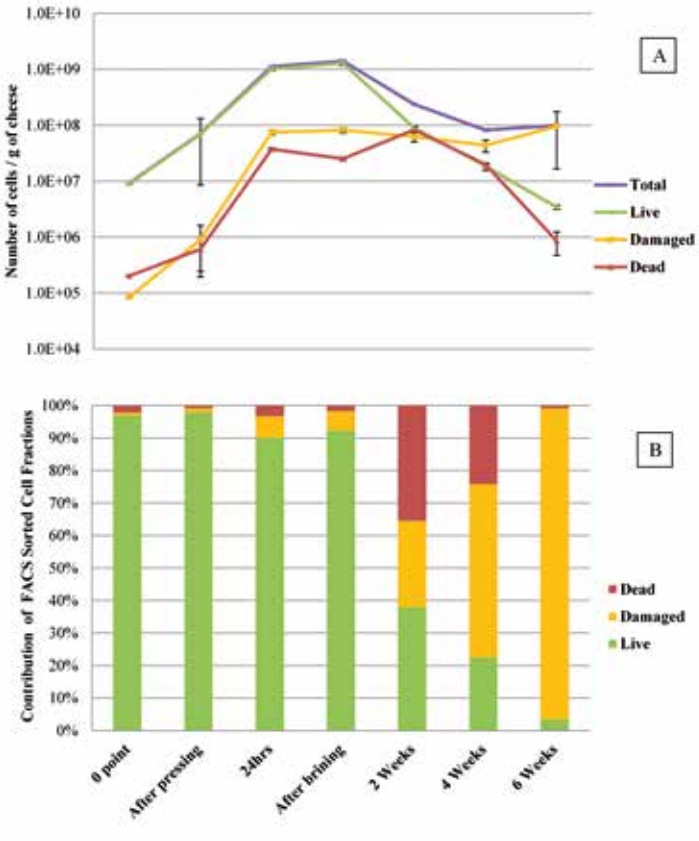


Figure 4. FACS counts of cheese samples taken during the manufacturing and ripening process. Panel A: dynamics of live, dead, and damaged fractions of sorted cells (by FACS) at different time point (absolute counts). Error bars represent the standard deviations from 2 independent replicates. Panel B: relative contribution of live, dead, and damaged fractions to the cheese community at different time points. Profiles were calculated from 2 independent replicates.

Both the qPCR estimation of cell counts of PMA-treated samples and the FACS based enumeration of live cells indicated a 2.5 orders of magnitude decrease of viability between the brining stage and the 6-weeks ripening time point, supporting the selective detection of live population in cheese when treated with PMA. However, the absolute difference in copy-number counts between PMA-treated and –untreated

bacteria was almost 2 orders of magnitude in the ripening period (decrease in logarithmic cell count of maximally 1.7), implying that further reduction in cell count would not be detected with PMA treatment. Nevertheless, these analyses indicate that the application of PMA enables a significant enhancement (up to 2 orders of magnitude) of the selective quantification of viable communities within a complex ecosystem obtained from an environmental matrix. A similar PMA-qPCR strategy was previously published by Desfossés-Foucault et al. (24) for monitoring of the viability of probiotics and starter culture in Cheddar cheese manufacturing and ripening. The authors found that, PMA-qPCR counts did not decrease as much as viable plate counts did during the ripening time and only 2-to-3 log cfu/g of cheese cell count decrease was detected with PMA-qPCR quantification, indicating similar limitations that we have observed in Gouda cheese ripening.

To our knowledge, this is the first successful attempt to employ the PMA strategy for enhanced focus on viable populations using a well-studied model system of Gouda cheese manufacturing. PMA can directly be applied on the cheese sample prior to total DNA isolation to be used in qPCR and metagenome analysis. This presents an efficient alternative to labour intensive sorting based techniques like FACS that require intense clean-up procedures prior to analysis, and is restricted to only a small fraction of the entire community that can be sorted.

Metagenomics of PMA-Treated Microbial Community Recovered from 6-Weeks Ripened Cheeses

The viable community analyses demonstrated substantial community changes over time during cheese making process. However, these community dynamics were poorly reflected in the total DNA pool derived from the matrix implying that DNA derived from dead or membrane compromised cells in the cheese matrix is persisting during extended cheese ripening periods. Sequencing of total DNA derived from PMA-treated bacterial communities of environmental samples has the potential to selectively determine the metagenome of the viable members in the microbial community. Therefore, total DNA of PMA-treated bacterial community from 6-weeks ripened cheese was sequenced to evaluate the added value of PMA treatment in metagenomic studies. The initial DNA was prepared for sequencing in two subsequent steps. In the first step, the total DNA was subjected to whole genome amplification (WGA) to enrich the DNA that was not modified with PMA. In the second step, the whole genome amplification product was used for library

construction and subjected to high-throughput random sequencing using Illumina technology.

In order to evaluate whether WGA introduces a bias in the sequence datasets, the community composition represented in the original DNA sample as well as the WGA product (in triplicate) derived from it, were determined using lineage-specific qPCR amplifications. The community composition profiles before and after WGA were similar, although an apparent 5-10% relative increase of lineage 7 and lineage 1&5 were observed after WGA (Fig. 5). With the WGA

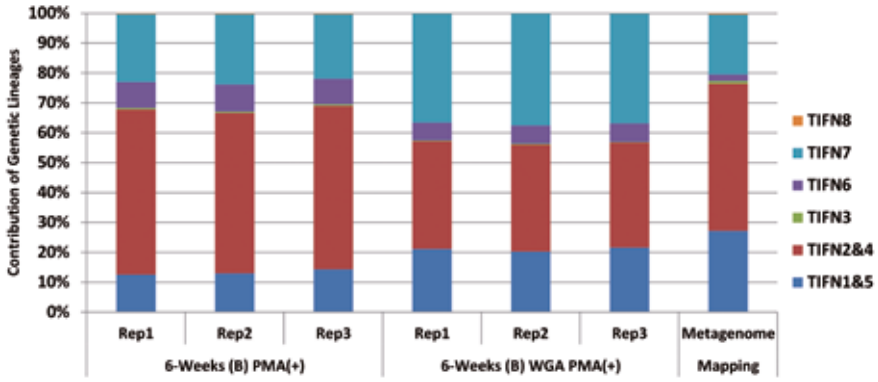


Figure 5. Community profiles of the PMA-treated total DNA sample (the replicate B) isolated from 6-weeks ripened cheese based on genetic lineage-specific qPCR and metagenomic analyses. The community profiles that were deduced from the qPCR analyses of the total DNA sample and WGA product were represented with three replicates.

procedure, the amount of DNA that was not modified with PMA and accessible for sequencing was enriched 100-fold in the PMA-treated total DNA sample of 6-weeks ripened cheese (data not shown).

Metagenome analysis of the WGA product of PMA-treated total DNA sample matched well with the community composition information deduced from qPCR analyses of WGA products, and appropriately reflects the active population at that time point in the cheese making process. PMA has been used once in a pyrosequencing study for the selective amplification of 16S rRNA genes of live cells in canal water (8). However, there is no protocol developed yet for the direct metagenome sequencing of PMA-modified total DNA isolated from environmental samples. The outcome of the metagenomic analysis presented in our work indicated that the proposed two-

step total DNA amplification and sequencing approach can effectively reveal the active (membrane intact) members of the microbial communities in PMA treated environmental samples.

One limitation that should be taken into account for the PMA treatment of environmental samples is that the effect of PMA may vary among different bacterial species (5). Especially the differences in the cell wall structure of Gram-negative and Gram-positive species may lead to differential PMA diffusion, which eventually may create community composition biases if they co-occur in the same environmental sample. For example, the PMA application on Gram-positive *Listeria innocua* was shown to lead to an overestimation of live cells during quantification (25), while Gram negative *E.coli* O157:H7 was quantified accurately (26). With the same approach, all environmental conditions that can modify the cell structure such as stress or growth conditions (lag or log phase) may in principle have an effect on PMA quantification. Therefore, the PMA studies on complex environmental samples should take into account the differential permeability of cell membranes of different species and investigators should be aware of the quantification limits under different environmental or growth conditions.

In conclusion, PMA was shown to enhance the selective analysis of the viable cell population during cheese manufacturing and ripening in this study. Dead or damaged cells with compromised membranes are permeable to PMA and thereby were selectively excluded from DNA amplification, enabling the discriminative analysis of live population in the microbial community of cheese. PMA treatment enabled an approximate 100-fold enhanced focus on the viable microbial community of cheese, which represents a difficult-to-process matrix with substantial proportion of membrane-compromised cells. PMA treatment was shown to be not a limitation for the direct sequencing of total DNA from environmental samples using the two step WGA-sequencing approach presented in this chapter. Our findings exemplify the importance of the selective analysis of the viable fraction within a microbial community and illustrate the usefulness of PMA treatment for this purpose in complex ecosystems. The approach described in this study may well be applied to artisanal and industrial mixed culture fermentation processes, host-associated microbiota in animals or humans (e.g. intestinal tract microbiota), or microbial communities from other environmental samples like the plant rhizosphere and bioremediation sites.

Acknowledgements

We would like to acknowledge Marja Twigt and Herwig Bachmann for the set-up of the mini-scale cheese manufacturing procedure. This work was funded by Top Institute Food and Nutrition (TIFN) and is part of Kluyver Centre for Genomics of Industrial Fermentation, a Centre of Excellence of The Netherlands Genomics Initiative.

References

1. **Ben-Amor K, Heilig H, Smidt H, Vaughan EE, Abee T, de Vos WM.** 2005. Genetic diversity of viable, injured, and dead fecal bacteria assessed by fluorescence-activated cell sorting and 16S rRNA gene analysis. Applied and environmental microbiology **71:**4679-4689.
2. **Masters CI, Shallcross JA, Mackey BM.** 1994. Effect of stress treatments on the detection of *Listeria monocytogenes* and enterotoxigenic *Escherichia coli* by the polymerase chain reaction. The Journal of applied bacteriology **77:**73-79.
3. **Novitsky JA.** 1986. Degradation of dead microbial biomass in a marine sediment. Applied and environmental microbiology **52:**504-509.
4. **Peris-Bondia F, Latorre A, Artacho A, Moya A, D'Auria G.** 2011. The active human gut microbiota differs from the total microbiota. PloS one **6:**e22448.
5. **Elizaquivel P, Aznar R, Sanchez G.** 2014. Recent developments in the use of viability dyes and quantitative PCR in the food microbiology field. Journal of applied microbiology **116:**1-13.
6. **Nocker A, Cheung CY, Camper AK.** 2006. Comparison of propidium monoazide with ethidium monoazide for differentiation of live vs. dead bacteria by selective removal of DNA from dead cells. Journal of microbiological methods **67:**310-320.
7. **Nocker A, Sossa-Fernandez P, Burr MD, Camper AK.** 2007. Use of propidium monoazide for live/dead distinction in microbial ecology. Applied and environmental microbiology **73:**5111-5117.
8. **Nocker A, Richter-Heitmann T, Montijn R, Schuren F, Kort R.** 2010. Discrimination between live and dead cellsin bacterial communities from environmental water samples analyzed by 454 pyrosequencing. International microbiology : the official journal of the Spanish Society for Microbiology **13:**59-65.
9. **Fujimoto J, Tanigawa K, Kudo Y, Makino H, Watanabe K.** 2011. Identification and quantification of viable *Bifidobacterium breve* strain Yakult in human faeces by using strain-specific primers and propidium monoazide. Journal of applied microbiology **110:**209-217.
10. **Crow VL, Martley FG, Coolbear T, Roundhill SJ.** 1995. The Influence of Phage-Assisted Lysis of Lactococcus *lactis* subsp *lactis* Ml8 on Cheddar Cheese Ripening. Int Dairy J **5:**451-472.
11. **ODonovan CM, Wilkinson MG, Guinee TP, Fox PF.** 1996. An investigation of the autolytic properties of three lactococcal strains during cheese ripening. Int Dairy J **6:**1149-1165.
12. **Chapot-Chartier MP, Deniel C, Rousseau M, Vassal L, Gripon JC.** 1994. Autolysis of Two Strains of *Lactococcus lactis* during Cheese Ripening. Int Dairy J **4:**251-269.
13. **de Ruyter PG, Kuipers OP, Meijer WC, de Vos WM.** 1997. Food-grade controlled lysis of *Lactococcus lactis* for accelerated cheese ripening. Nature biotechnology **15:**976-979.
14. **Fox PF, Wallace JM, Morgan S, Lynch CM, Niland EJ, Tobin J.** 1996. Acceleration of cheese ripening. Antonie van Leeuwenhoek **70:**271-297.
15. **Smit G, Smit BA, Engels WJ.** 2005. Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. FEMS microbiology reviews **29:**591-610.

16. **Ardo Y.** 2006. Flavour formation by amino acid catabolism. Biotechnology advances **24:**238-242.
17. **Stadhoouders J, Leenders, G.J.M.** 1984. Spontaneously developed mixed-strain cheese starters. Their behaviourstowards phages, and theri use in Dutch cheese industry. Netherlands Milk and Dairy Journal **38:**25.
18. **Bachmann H, Kruijswijk Z, Molenaar D, Kleerebezem M, van Hylckama Vlieg JE.** 2009. A high-throughput cheese manufacturing model for effective cheese starter culture screening. Journal of dairy science **92:**5868-5882.
19. **Bachmann H, de Wilt L, Kleerebezem M, van Hylckama Vlieg JE.** 2010. Time-resolved genetic responses of *Lactococcus lactis* to a dairy environment. Environmental microbiology **12:**1260-1270.
20. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. Nature methods **9:**357-359.
21. **Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, Genomes Project Analysis G.** 2011. The variant call format and VCFtools. Bioinformatics **27:**2156-2158.
22. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S.** 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25:**2078-2079.
23. **Lee JL, Levin RE.** 2009. A comparative study of the ability of EMA and PMA to distinguish viable from heat killed mixed bacterial flora from fish fillets. Journal of microbiological methods **76:**93-96.
24. **Desfosses-Foucault E, Dussault-Lepage V, Le Boucher C, Savard P, Lapointe G, Roy D.** 2012. Assessment of Probiotic Viability during Cheddar Cheese Manufacture and Ripening Using Propidium Monoazide-PCR Quantification. Frontiers in microbiology **3:**350.
25. **Lovdal T, Hovda MB, Bjorkblom B, Moller SG.** 2011. Propidium monoazide combined with real-time quantitative PCR underestimates heat-killed Listeria innocua. Journal of microbiological methods **85:**164-169.
26. **Elizaquivel P, Sanchez G, Aznar R.** 2012. Application of propidium monoazide quantitative PCR for selective detection of live *Escherichia coli* O157:H7 in vegetables after inactivation by essential oils. International journal of food microbiology **159:**115-121.

Chapter 5

Arginine Stimulated Culturability of *Lactococcus lactis* Cells Limited for C1-Resourses

Oylum Erkuş, P.W.M. Janssen, I. van Alen-Boerigter, M. Wels, R. Bongers, Michiel Kleerebezem and Eddy J. Smid

Manuscript in preparation

Abstract

Culture-based microbial community analyses are inevitably biased due to differences in the metabolism and viability states of the community members. In chapter 3, the microbial community present in complex Gouda cheese starter culture was investigated, and full recovery of *L. lactis* populations in the community was achieved only with the complementary plating on LM17 and Reddy's agar media (This thesis, Chapter 3). The differentially recovered *Lactococcus* cells on Reddy's agar (5 times more) were determined to belong to one specific genetic lineage of the 7 different lineages recognized within this complex starter culture. In the present study, arginine was found to be the growth stimulating medium component for this genetic lineage on Reddy's agar. The molecular basis of this stimulatory effect was investigated in the lineage-representing strain *L. lactis* TIFN7 by transcriptome analysis before and after the addition of arginine in arginine-limited chemostat cultures. For comparative analyses, the *L. lactis* TIFN7 transcriptome of the steady state culture was compared to that of a steady state culture of an arginine unresponsive strain (*L. lactis* TIFN5). Prior to arginine addition, *L. lactis* TIFN7 expressed the genes coding for the arginine, pyrimidine and purine biosynthesis pathways at a very low level, and displayed a high level expression of genes related to single carbon (C1) metabolism and reactions involving $C0_2$ production, compared to the expression levels of genes coding for the same pathways in *L. lactis* TIFN5. $C0_2$ and glutamine are the main substrates of the metabolic pathways leading to the production of arginine and pyrimidines. Arginine addition relieved both the arginine and $C0_2$ limitation in strain TIFN7, and redirected these metabolic resources towards the pyrimidine and purine biosynthesis. The stringent C-1 limitation in strain TIFN7 when grown on LM17 agar medium was confirmed by enhanced growth on the same medium in 5% CO_2 enriched air atmosphere and on the same medium enriched with uracil, folate, aspartate, and arginine.

Introduction

Microbial ecology studies make use of culture dependent or independent approaches to determine the community compositions of environmental samples. Culture based methods are still required in studies of microbial communities with a high degree of diversity below the species level (This thesis, Chapter 3) or when community members are intended to be studied in more detail for example through genomic sequencing or transcriptomic analyses. Commonly, microbes from environmental samples are recovered with the use of selective and/or enrichment media. Thereby, culture based community analyses are intrinsically biased by culturability/viability related differences of the community members. Different viability and culturability states may occur among bacteria depending on the cell membrane integrity, metabolic activity, and growth (1). The cells that succeed to form colonies on solid growth media meet all the criteria, and are considered as viable and culturable. However, if the cells are intact and metabolically active but not able to replicate on the selected growth media, they are defined as "viable but non-culturable" (VBNC) (2). VBNC is a conflicting term (3) and has been used both for bacterial cells that appeared conditionally non-culturable and for those that have never been cultured (4). A conditionally non-culturable state is often observed when the cells have been exposed to sublethal stress conditions (5). Alternatively, bacteria in a community may be considered as VBNC when they display heterogeneity in their growth requirements, have unknown auxotrophies or when trophic interactions are present among the community members that are not complemented in the growth media and/or growth conditions applied during culturing. Metagenomic studies indicate that the majority of the microbes present in environmental samples have not been cultured yet, but the actual percentage of (un)cultured members differs very much between different environmental habitats (6, 7).

Many microbial species, including *Lactococcus lactis* (2), have been observed to enter the VBNC state under environmental stress conditions. *L. lactis* is commonly found in defined and undefined cheese starter cultures either as single strain, or as blends of strains with differential functionality during cheese production process (This thesis, Chapter 3). Stuart et al (1999) demonstrated that *L. lactis* cells become completely non-culturable in a strain-dependent manner between 1 week and 3 months, under carbohydrate starvation even though they can sustain their cellular integrity and a low level of metabolic activity up to 3.5 years (8). Non-culturability might also occur due to unfulfilled growth requirements of specific *Lactococcus* strains under the

growth conditions employed. Commonly, *L. lactis* strains have multiple amino acid auxotrophies. For example, the dairy associated *L. lactis* ssp. *lactis* strains are mostly auxotrophic for branched-chain amino acids and histidine, while the plant derived *L. lactis* strains are frequently prototrophic for these amino acids (9). This was proposed to be a reflection of the evolutionary adaptation to long term growth in the amino acid- and protein-rich environment of milk. Adaptive evolution of a plant-derived *Lactococcus* strains to growth in milk led to the accumulation of mutations related to enhance efficacy of peptide and amino acid uptake systems from the protein rich environment of milk, supporting the reduced dependency for endogenous amino acid synthesis in this habitat (10). The accumulation of mutations and deletions on the amino acid biosynthesis pathways in the genomes of dairy associated *L. lactis* strains (11, 12) support this hypothesis as well.

In this study, we investigate the molecular basis of non-culturability of *L. lactis* observed during the enumeration of the microbial community present in a complex Gouda cheese starter culture, which was reported in chapter 3. Full recovery of the cells present in the microbial community was achieved only when three complementary growth media were employed. With the media used for the recovery of all *Lactococcus* cells (Reddy's and LM17 agar), 5 times more cells were recovered on Reddy's agar compared to LM17 agar plates. Interestingly, the differentially recovered cells on Reddy's agar consistently belonged to a single genetic lineage (designated genetic lineage 7) in the community (This thesis, Chapter 3). In the present study, arginine that is present in higher amounts in Reddy's agar was identified as the medium component that determined the higher recovery of *L. lactis* cells on Reddy's agar. Furthermore, the molecular response to arginine was investigated in a lineage 7 representative strain (*L. lactis* TIFN7) by whole genome gene expression profiling before and after providing an arginine switch to arginine-limited chemostat cultures, and comparing the transcriptome of arginine limited *L. lactis* TIFN7 to the transcriptome of an arginine-non-responsive strain (*L. lactis* TIFN5) cultured under the same conditions. These analyses revealed that *L. lactis* TIFN7 had lower activity of genes encoding the arginine, pyrimidine and purine biosynthesis pathways compared to *L. lactis* TIFN5. Conversely, *L. lactis* TIFN7 displayed very high expression of genes encoding CO_2 or single carbon (C1)-folate intermediate producing pathways, suggesting that the availability of C1 compounds was the major limitation for arginine biosynthesis in TIFN7. The relief in arginine limitation after the arginine supplementation allowed *L. lactis* TIFN7 cells to re-direct the resources such as C1 compounds, glutamine, and aspartate towards

the biosynthesis of pyrimidines and purines, which elevated the activity of these pathways in *L. lactis* TIFN7 to the level observed in *L. lactis* TIFN5 under steady state condition, and thereby sustained *L. lactis* TIFN7 growth. Furthermore, the growth of *L. lactis* TIFN7 on LM17 agar medium in 5% CO2 enriched air atmosphere or on LM17 agar medium supplemented with uracil, folate, aspartate, and arginine compensated well for C1 compound limitation of *L. lactis* TIFN7.

Materials and Methods

Starter Culture, Bacterial Strains and Growth Conditions

The undefined and mixed-strain Gouda type cheese starter culture that was designated 'Ur' (13) was used in this study. For *L. lactis* bacterial isolations, the starter culture was propagated in sterile 10% (w/v) skimmed milk two times at 20°C for 20 hours, and plated on two selective media: lactose supplemented M17 agar (LM17; Merck, Darmstadt, Germany) and Reddy's agar (14), and the colony forming units were enumerated and isolated after 2 days of incubation at 30^{0}C. M17 and Reddy`s agar isolates as well as *L. lactis* ssp. *cremoris* TIFN7 and TIFN5 strains were propagated in LM17 broth at 30°C overnight prior to experiments unless stated differently.

Mixed Culture Growth on Modified-LM17 Media

The modified LM17 media were prepared by single supplementation of LM17 agar medium with components that are present in higher concentration in Reddy medium as compared to the basic LM17 recipe (Table 1). The supplements and the final concentrations at which they were added to LM17 agar medium are as follows: yeast extract (5g/L), beef extract (5g/L), papaic digest of soybean (5g/L), polypeptone (5g/L) and L-arginine (1.5g/L). The influence of the supplements on starter culturability was evaluated by plating the complete starter culture on modified-LM17 as well as Reddy's, and LM17 agar media as described above. Two-tailed t-test with two groups was applied for the pairwise mean comparisons of colony forming units (CFU) recovered on the LM17 agar media with different supplements. The CFU recovered on two different agar media were considered significantly different if the P value was lower than 0.05.

Determination of Arginine Responses of Single Strains

For the preparation of a collection of single strain isolates, the complex starter culture was plated on arginine supplemented and non-supplemented LM17 agar plates as described above and 40 single colony isolates were picked from each agar medium. The recovered colonies were propagated overnight in LM17 broth at 30^{0}C, plated on arginine supplemented and non-supplemented LM17 agar plates with high throughput plating technique (15) in duplicate for the enumeration of CFU.

To evaluate the arginine dose response, LM17 agar plates were supplemented with arginine to the final arginine concentrations of 2, 3, 4, 5, 6, 7, 8, 9 or 10 mM. Two arginine-responsive and two arginine-non-responsive isolates, which were picked from Reddy's and LM17 agar plates during the mixed-culture growth experiments described above, were propagated overnight in LM17 broth at 30^{0}C, and plated on LM17 agar plates with increasing arginine concentration for the enumeration of CFU.

Genetic Fingerprinting of Arginine Responsive Cells

The starter culture Ur was plated on arginine supplemented LM17, Reddy's and LM17 agar media and the CFU were enumerated as described above. Prior to DNA isolation, the single colonies were grown overnight and the genomic DNA of individual cultures was isolated with QIAgen Tissue and Culture DNA Isolation Kit (QIAgen GmbH, Hilden, Germany) according to the manufacturer's instruction. DNA quality and concentrations were determined by Nanodrop (Coleman technologies Inc., Orlando, FL) and 0.8% (wt/vol) agarose gel electrophoresis. The genetic lineage corresponding to each of the isolates was determined using the previously described high throughput AFLP typing methodology that employs protocols that are specifically optimized for the effective and reliable genetic fingerprinting of *Lactococcus lactis* isolates (This thesis, Chapter 2)

Growth Stimulating Medium Supplements

In order to evaluate whether any other supplements beside arginine can compensate the lack of growth of *L. lactis* TIFN7, LM17 agar medium was supplemented with compounds other than arginine. The supplements and the final concentrations at which they were added to LM17 agar medium are as follows: glutamate (5g/L),

aspartate (4.2g/L), arginine (1.25g/L), formate (10mg/L) and uracil (100mg/L). The influence of these additional supplements on the recovery rate of *L. lactis* TIFN7 was evaluated by CFU enumeration of *L. lactis* TIFN7 on supplemented and non-supplemented LM17 agar media in normal air and 5% CO_2 enriched atmosphere in triplicate as described above.

Chemostat Cultivation and Arginine Supplementation

For both L. lactis ssp. cremoris TIFN5 and TIFN7 chemostat cultivations, the bioreactors were flushed with nitrogen at a flow of 50 ml/min and with the agitation speed of 125 rpm. Chemostat reactors were operated at 30°C with a dilution rate of 0.1 h-1, and pH was maintained at 6.5 by automated titration of 2.5 M NaOH.

The arginine-responsive *L. lactis* ssp. *cremoris* TIFN7 was cultured under chemostat conditions (500 ml working volume) in duplicate, established by arginine limitation in chemically defined medium (Supplementary Table 1, CDM with 100 mM glucose; 0,05 mM Arginine) at 30°C. Steady-state was considered to be reached when the optical density of the culture at 600 nm remained constant for at least 5 bioreactor volume changes, and the arginine limitation was confirmed by determination of the residual glucose concentration in the spent medium using the Diabur test glucose dipsticks (Accu-Chek, Roche, Almere, Netherlands) according to the manufacturer's protocol. The arginine concentration of steady-state chemostat cultures were increased to 50mM by arginine supplementation to the fermenter and by simultaneous switching the feed medium of the fermenter (arginine-limited CDM) to the arginine supplemented (50 mM) CDM medium. Chemostat mode was continued using this medium until the end of the cultivation. 50 ml of samples were collected in duplicate before the medium switch, and 15, 30 and 120 minutes after the switch to arginine supplemented media, for the transcriptome analysis. The cells were harvested by centrifugation at 13000 g for 5 minutes at 4°C, and resuspended in 1 ml of cold TE buffer.

For comparative transcriptome analysis, the arginine non-responsive *L. lactis* TIFN5 strain was grown in duplicate under the same condition (CDM; 100 mM glucose; 0,05 mM Arginine) that *L. lactis* TIFN7 was grown prior to arginine supplementation. None of the replicates of *L. lactis* TIFN5 chemostat cultures could be maintained for 5 volumes of change to reach the steady-state conditions (data not shown). The TIFN5 chemostat cultures were glucose limited as confirmed by the lack of detectable concentrations of residual glucose in the spent medium using the Diabur test glucose

dipsticks (Accu-Chek, Roche, Almere, Netherlands), according to the manufacturer's protocol. The transcriptomes of *L. lactis* TIFN5 chemostat cultures were determined under these glucose limited condition to compare with the transcriptomes of the arginine limited *L. lactis* TIFN7 chemostat cultures. Approximately 50 ml of *L. lactis* TIFN5 chemostat culture was sampled in duplicate for transcriptome analysis and samples were processed as described above.

RNA Isolations, cDNA Synthesis and Labelling

Aliquots of 500 **µl** cell suspensions in TE buffer were added to 500 µl phenol/chloroform (5:1), 30 µl of SDS (10 %), 30 µl of sodium acetate buffer (3 M, pH 5,2), and 500 mg of glass beads (75-150 µm), frozen in liquid nitrogen, and kept at -80°C. Prior to RNA isolation, the cells were disrupted by bead beating (Savant FastPrep FP120, Rome, Italy), 3 times for 40 seconds at speed 4, and lysates were centrifuged for 1 minute at 13.000 rpm and at 4°C (Eppendorf Centrifuge, Hamburg, Germany). The aqueous phase was removed and extracted with 400 **µl of cold chloroform (4°**C) to remove residual phenol. RNA was further purified using the High Pure RNA isolation kit (Roche Diagnostics, Mannheim, Germany) according to manufacturer's protocol. RNA concentrations were determined with Nanodrop spectrophotometer (Model ND-100, Coleman technologies Inc., Orlando, FL) and the quality was checked using the Bioanalyzer (Model 2100, Agilent Technologies, Santa Clara, CA). cDNA synthesis and labelling were carried out using 10 µg of total RNA as described previously (16). The concentration of cDNA and the dye incorporation were checked using the Nanodrop spectrophotometer (Model ND-100, Coleman technologies Inc., Orlando, FL). The cDNA preparations all had 260/280nm absorption ratios in the range of 1,6-to-1,8 and 0,3 **µg of** this cDNA in 25 **µl total volume** were used for hybridization.

Microarray Design, Hybridizations, and Data Analysis

DNA microarrays specific for *L. lactis* TIFN5 and TIFN7 were designed independently using a custom probe design procedure that uses the strain's genome sequences. Single slides with 8 microarrays (Agilent 8x15K) were used for each strain. All probes were designed as 60-mers, minimizing the cross hybridization and targeting 100 % match at 65^0C hybridization and 37^0C washing temperatures. In total 11088 and 10811 probes were designed for *L. lactis* TIFN5 and TIFN7 respectively, and the annotated genes were represented by up to 5 probes with an average of 4.5 probes. The hybridization schemes were designed as closed loop model connecting all the

time points in a minimal loop and allowing the comparison of gene expression at any time point to each other (Supplementary Figure 1).

cDNAs labelled with cyanine 3 and cyanine 5 (0,3 μg of each) were hybridized to the slides using the Agilent *in situ* Hybridization Kit Plus (Agilent, Palo Alto, CA), according to the manufacturer's protocol for 17 hours at 65°C and 10 rpm. Slides were washed with wash solution 1 at room temperature for 1 minute and subsequently with wash solution 2 (at 37°C) for 1 minute. The slides were dried with nitrogen at 1 atm pressure, and kept in dark until scanning. Slides were scanned using an Agilent DNA microarray scanner (Agilent, Palo Alto, CA) as described by Meijerink et al (17). Image analysis was performed using the ImaGene Version 7.5 software (BioDiscovery Inc., Marina Del Rey, CA, USA). Each microarray was scanned at different intensities to select for the best image with optimal signal distribution. The data were normalized with Lowess normalization using MicroPrep (18), and corrected for differences between slides on the basis of total signal intensity per slide using Postprep (18)`. The median intensity of all the probes available per gene was accepted as the gene expression intensity. CyberT was used to compare the different transcriptomes, taking into account the duplicates of each condition (19)`, which resulted in a gene expression ratio and false discovery rate (FDR) for each gene. Gene expressions with a false discovery rate (FDR) values lower than 0.05 were considered statistically significant. Additionally, the median intensity of all the probes available for each gene was used for the comparison of absolute gene expressions between strains (*L. lactis* TIFN5 and TIFN7). Differentially expressed genes on the general metabolic pathways were visualized by mapping their corresponding clustered orthologous group (COG) categories on the Interactive Pathways Explorer (iPath) (20) to be used as a template for manual drawing of up and down-regulated pathways in Figure 4.

Results and Discussion

Plate Count Anomaly and Arginine Triggered Recovery of Bacterial Cells

The Gouda cheese starter culture (designated 'UR',(13)) that was used in this study was previously reported to encompass 8 co-existing genetic lineages based on amplified fragment length polymorphism (AFLP) fingerprinting of single colonies obtained by using different selective growth media (This thesis, Chapter 3) and the representative strains of each genetic lineage have previously been subjected to

whole genome sequencing, comprising 5 *L. lactis* ssp.*cremoris* strains (TIFN1, 3, 5, 6, and 7), 2 L. *lactis* ssp. *lactis* biovar. *diacetylactis* strains (TIFN 2 and 4), and 1 *Leuconostoc mesenteroides* strain (TIFN8) (This thesis, Chapter 3). In chapter 3, Reddy's and LM17 agar were used for the recovery of *Lactococcus lactis* isolates from the starter culture. Notably, 5 times more lactococcal isolates were recovered using Reddy's agar as compared to LM17 agar.

In order to understand the molecular basis for the better recovery of lactococcal cells on Reddy's agar plates, LM17 medium was complemented with individual components that are exclusively present, or present at a higher concentration, in the Reddy's medium relative to LM17 medium (Table 1). Neither yeast extract, beef extract, papaic digest of soybean, nor the polypeptone supplementation could significantly enhance the plating efficacy of the starter culture UR on LM7 agar plates (P values > 0.05). In contrast, arginine supplementation of LM17 agar medium, to the final concentration of 8.6mM that is also present in Reddy's agar, led to a plating efficacy that was comparable to that observed on Reddy's medium (P values > 0.05 between the cfus recovered on Reddy's and on arginine supplemented LM17, but <0.05 for the other supplemented media, Figure 1). Therefore, the low arginine concentration in LM17 medium (which in was determined to be 1.28 mM) was identified as the major limiting factor for the plating efficacy of the starter culture on LM17 agar plates in comparison with the Reddy's agar plates.

Fingerprinting of Arginine Responsive Cells

7 distinct lactococcal genetic lineages were identified in the starter culture in chapter 3. The representatives of *L. lactis* subsp. *cremoris* lineage 7 appeared to be dominantly present (> 75 % of the total lactococcal community) in this complex starter and were exclusively recovered from Reddy's agar plates (This thesis, Chapter 3). These observations in combination with the findings presented above would imply that the enhanced plating efficacy on arginine supplemented LM17 agar plates is due to the enhanced recovery of the lineage 7 representatives. To verify, to which genetic lineage the individual colonies recovered on the arginine supplemented LM17 agar plates belong, AFLP analysis was performed. AFLP fingerprinting revealed that 78% of all colonies recovered from arginine supplemented LM17 plates belonged to genetic lineage 7, which is similar to the relative dominance level of this lineage that was found previously using Reddy's agar plates in the undefined starter culture (This thesis, Chapter 3). Moreover, in agreement with previous observations in chapter 3,

none of the colonies isolated from non-supplemented LM17 agar plates could be assigned to genetic lineage 7. These findings clearly establish that increased arginine concentrations promote the recovery of colonies that belong to genetic lineage 7. The recovery of colonies at increased arginine concentration in LM17 agar plates was also determined for strains that are the representatives of the other genetic lineages in the starter culture. None of the strains representing the other lineages were concluded to display arginine-supplementation dependent growth.

Table 1. Comparison of Reddy's Agar and LM17 Agar Medium

Ingredients (g/L)	Reddy's Agar	LM17 Agar
Beef extract[a]	5	0
"Lab Lemco" powder	0	5
Papaic digest of soybean[a]	5	0
Soya peptone	0	5
Polypeptone[a]	5	0
Yeast extract[a]	5	2,5
Tryptone	0	5
L-Arginine.HCl[a]	1,5	0
Lactose	1,5	5
Sodium carboxymethylcellulose	10	0
Calcium citrate	10	0
Agar	15	12
Magnesium sulphate	0	0,25
Ascorbic Acid	0	0,5
Di-sodium glycerphosphate	0	19

[a] The highlighted ingredients were added to the modified-M17 agar media at the final concentrations that they were present in Reddy's Agar Medium

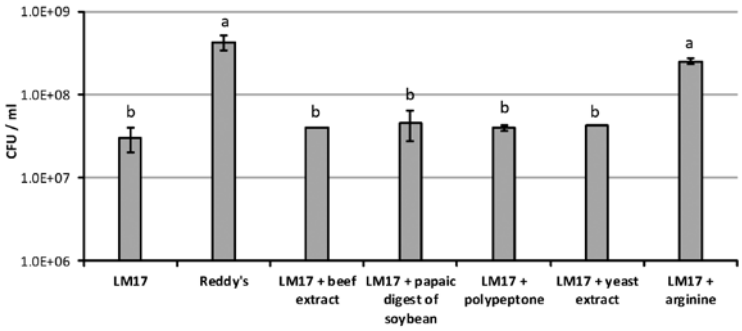


Figure 1. CFU recoveries on modified-LM17 agar media plated with Gouda cheese starter culture. Bars represent the CFU recovered on LM17 agar media with different supplements (the average of the two measurements). Significance testing was based on the pairwise comparison of CFUs recovered on different media using two-tailed t-test. The groups; a and b were determined to be significantly different from each other with $P < 0.05$.

Arginine Responsive Growth of Single Strains Derived from UR

One interesting observation for the arginine responsive cells of *L. lactis* genetic lineage 7 was that they don't have growth limitation in LM17 liquid medium. The colonies picked from arginine supplemented LM17 agar plates had very comparable growth and acidification rates to those of the isolates picked from non-supplemented LM17 plates in liquid LM17 growth during propagations before plating experiments (data not shown). This observation is indicative of several possibilities for the reduced growth on agar medium, such as problems related to inefficient transport of nutrients due to reduced diffusion rate in agar medium or increased oxygen sensitivity. A certain population of genetic lineage 7 cells may have a higher demand for the limiting nutrients and have an inability to acquire them from agar medium due to low rate of diffusion or they may have the same diffusion limitation in acquiring certain metabolites from adjacent colonies on agar medium that were otherwise easily exchanged in liquid medium. Another possibility could also be the higher oxygen sensitivity of a certain population of genetic lineage 7 cells on agar plates with direct exposure to oxygen in air although they may tolerate it better in liquid medium.

In order to eliminate the possibility of cross feeding effect due to the proximity of the colonies in the liquid medium, the arginine stimulated recovery was also tested using single strains. The starter culture UR was plated on arginine supplemented and non-supplemented LM17 agar plates and 40 single colonies picked from each agar medium were cultivated in LM17 liquid_medium. The CFU in these overnight cultures were subsequently enumerated on arginine supplemented and non-supplemented LM17 agar plates. The CFU enumeration obtained from overnight cultures of single colony isolates from LM17 plates were very similar (only 1- to 2-fold difference) when plated on arginine supplemented and non-supplemented LM17 agar plates. In contrast, overnight cultures of single colony isolates obtained from arginine supplemented LM17 agar plates displayed 2- to 33-fold higher CFU recovery on arginine supplemented LM17 agar plates compared to the non-supplemented LM17 agar plates (Figure 2). This observation confirms that the higher recovery of *L. lactis* cells by plating on LM17 agar medium supplemented with arginine is also achieved with single strain representative isolates recovered from the complex starter culture UR.

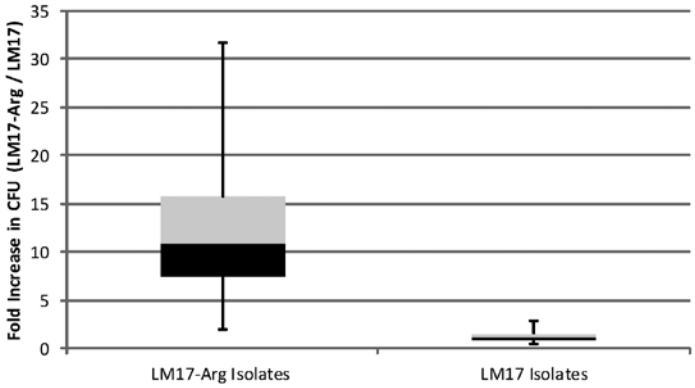


Figure 2. Relative CFU recovery increase of isolates on arginine supplemented LM17 agar plates compared to LM17 agar plates. LM17-Arg Isolates (n=40) and LM17 isolates (n=40) represent the individual isolates from Gouda cheese starter culture that were recovered on arginine supplemented LM17 and nonsupplemented LM17 agar plates, respectively. The whiskers show the total range of measured values and the boxes show the median, the 25th and 75th percentiles.

To determine the minimum level of arginine required for the complete recovery of arginine-responsive cells in the starter culture, we analysed the relationship between the arginine concentration in the growth medium and the recovery rate of arginine-responsive *L. lactis* isolates on the plates. Two arginine-responsive and two arginine-nonresponsive isolates that were picked from Reddy's and LM17 agar plates, respectively, were grown LM17 agar plates supplemented with increasing arginine concentrations, ranging from 2 to 10 mM, using 1mM increments. Cultures of the arginine-nonresponsive isolates revealed consistent CFU recovery, irrespective of the arginine concentration provided in the plates (Figure 3). However, a clear dose-response relationship between CFU enumeration and the supplemented arginine concentration was observed for the arginine-responsive cultures, which displayed a gradual increase from 2-10 * 10^6 cfu/ml on LM17 plates containing 2 mM final arginine concentration to 8 * 10^8 cfu/ml on LM17 plates with 6 mM of final arginine concentration (Figure 3). Arginine concentrations above 6 mM did not further stimulate the recovery of colonies, indicating that 6 mM arginine supplementation in LM17 leads to complete recovery of the arginine-responsive *L. lactis* isolates.

Responses of *L. lactis* TIFN7 and TIFN5 Transcriptomes in Chemostat Cultivations

L. lactis TIFN7 is a representative strain of the arginine-responsive genetic lineage 7 in the undefined starter culture. In order to understand how arginine supplementation affects the metabolism of *L. lactis* cells belonging to genetic lineage 7, the global

gene expression response of the representative strain *L. lactis* TIFN7 to arginine supplementation was determined. In media with low-arginine concentration, steady state chemostat effluents of *L. lactis* TIFN7 cultures contained 60 mM of residual glucose, indicating that the chemostats were not glucose limited (Supplementary Figure 2). After switching the chemostat cultures to a medium with a high arginine concentration (50 mM), the optical density (OD_{600}) of the *L. lactis* TIFN7 chemostat culture increased from 1 to 3 demonstrating the arginine limitation. The transcriptome of an arginine-limited (0.5 mM) chemostat culture of *L. lactis* TIFN7 (in duplicate) was compared to transcriptomes obtained after switching the *L. lactis* TIFN7 chemostat culture to an arginine-excess (50 mM) medium (15 min, 30 min. and 120 min. after the switch).

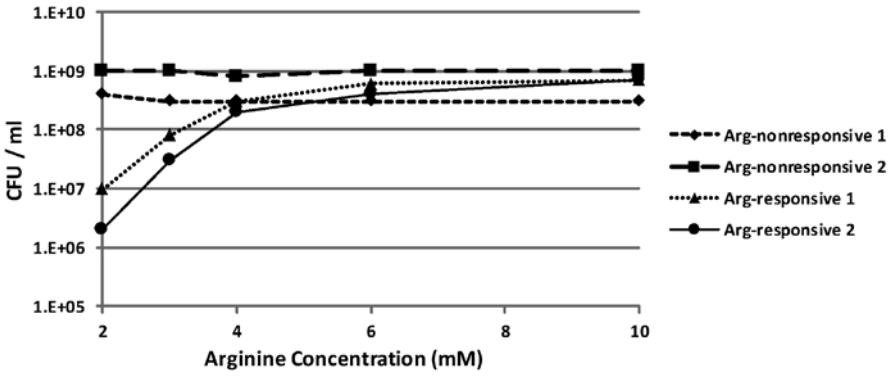
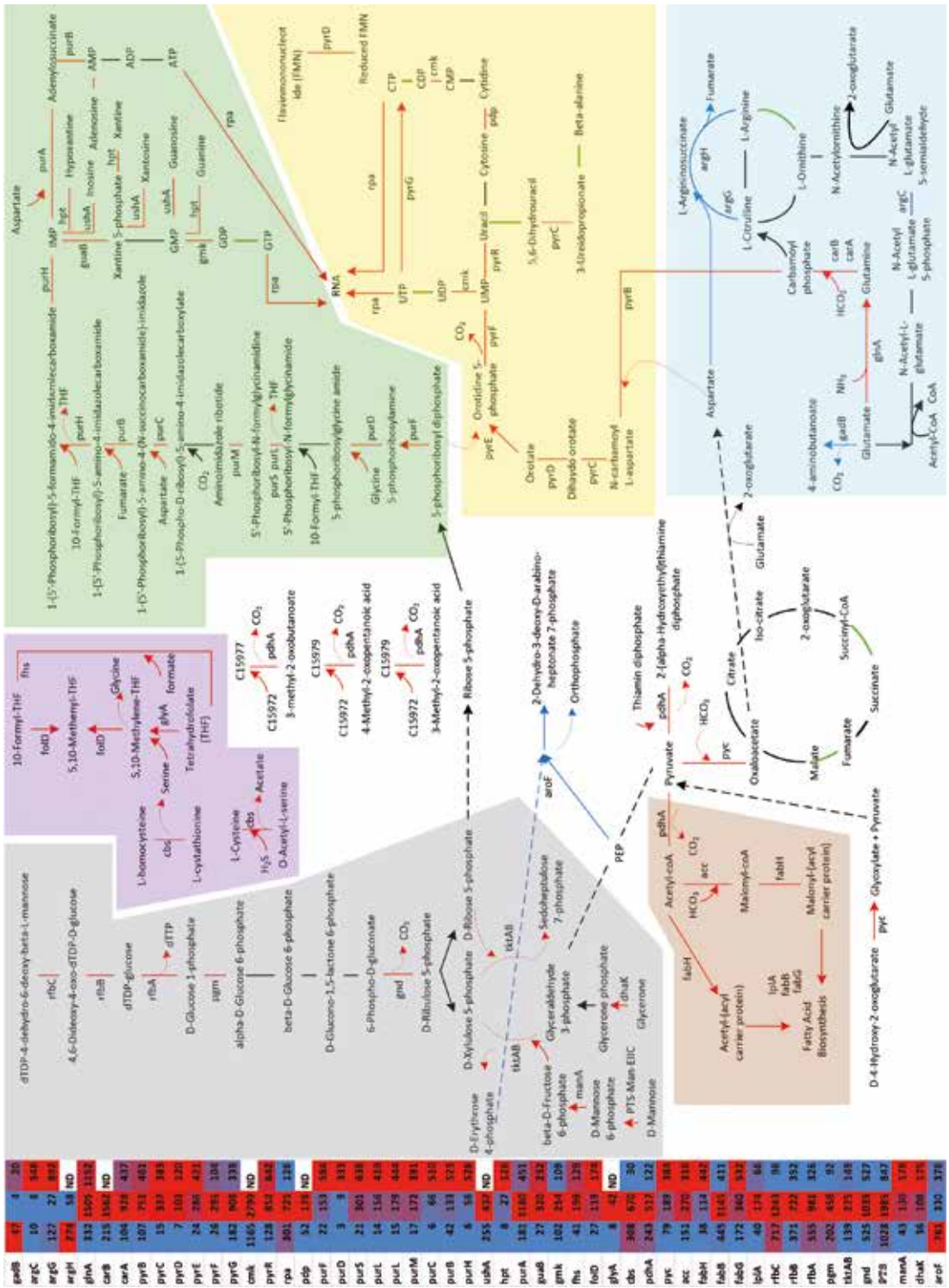


Figure 3. The recovery of arginine-responsive and -nonresponsive isolates (CFU per ml of culture) with increasing arginine concentrations on LM17 agar medium. The CFUs are the average of the two measurements.

Figure 4. The genes that were differentially expressed after 15 minutes of growth in arginine excess chemostat cultures of *L. lactis* TIFN7. The genes represented with blue colour were differentially downregulated and the genes represented with red colour were differentially upregulated compared to the steady state condition before the arginine supplementation. The genes represented with black colour are present on *L. lactis* TIFN7 genome, but not significantly detected in *L. lactis* TIFN7 transcriptome, and the genes represented by green colour are not present on *L. lactis* TIFN7 genome. The absolute expression levels of the genes depicted in the metabolic map are given in the left panel for steady state arginine limited cultivation of *L. lactis* TIFN7 (2nd column), 15 minutes after the culture was switched to arginine excess medium (3rd column), and glucose limited growth of *L. lactis* TIFN5 (4th column), respectively. The gene names are given on the 1st column.


Aspartate
purA
Adenylosuccinate
purB
AMP
Adenosine
Inosine
hpt
Hypoxanthine
Xanthine 5-phosphate
ushA
Xanthosine
ADP
GMP
ushA
Guanosine
ATP
gmk
GDP
hpt
Guanine
GTP
rpa
1-(5'-Phosphoribosyl)-5-formamido-4-imidazolecarboxamide
purH
IMP
10-Formyl-THF
THF
purH
1-(5'-Phosphoribosyl)-5-amino-4-imidazolecarboxamide
purB
Fumarate
1-(5'-Phosphoribosyl)-5-amino-4-(N-succinocarboxamide)-imidazole
purC
Aspartate
1-(5-Phospho-D-ribosyl)-5-amino-4-imidazolecarboxylate
purM
Aminoimidazole ribotide
5'-Phosphoribosyl-N-formylglycinamidine
purS purL
5'-Phosphoribosyl-N-formylglycinamide
THF
10-Formyl-THF
5-phosphoribosylglycine amide
purD
Glycine
5-phosphoribosylamine
purF
5-phosphoribosyl diphosphate
Flavinmononucleotide (FMN)
pyrD
Reduced FMN
RNA
rpa
CTP
pyrG
CDP
cmk
CMP
UTP
UDP
cmk
UMP
pyrH
Uracil
Cytosine
pdp
Cytidine
CO2
Orotidine 5-phosphate
pyrF
Orotate
pyrD
5,6-Dihydrouracil
Dihydro-orotate
pyrC
3-Ureidopropionate
Beta-alanine
N-carbamoyl L-aspartate
pyrB
L-Argininosuccinate
Fumarate
argH
L-Citruline
L-Arginine
argG
Carbamoyl phosphate
L-Ornithine
carB
carA
HCO3
N-Acetylornithine
2-oxoglutarate
Glutamine
4-aminobutanoate
gabB
NH3
Glutamate
gltA
N-Acetyl-L-glutamate
argC
N-Acetyl L-glutamate 5-phosphate
Glutamate
N-Acetyl Glutamate
Acetyl-CoA
CoA
5-semialdehyde
Aspartate
2-oxoglutarate
Glutamate
dTDP-4-dehydro-6-deoxy-beta-L-mannose
rfbC
4,6-Dideoxy-4-oxo-dTDP-D-glucose
rfbB
dTDP-glucose
rfbA
dTTP
D-Glucose 1-phosphate
pgm
alpha-D-Glucose 6-phosphate
beta-D-Glucose 6-phosphate
10-Formyl-THF
fhs
folD
5,10-Methenyl-THF
folD
Glycine
5,10-Methylene-THF
glyA
formate
Tetrahydrofolate (THF)
L-homocysteine
cbs
Serine
L-cystathionine
metB
L-Cysteine
cbs
Acetate
H2S
O-Acetyl-L-serine
C15977
CO2
C15972
pdhA
3-methyl-2-oxobutanoate
C15979
CO2
C15972
pdhA
4-Methyl-2-oxopentanoic acid
C15979
CO2
C15972
pdhA
3-Methyl-2-oxopentanoic acid
D-Gluconic-1,5-lactone 6-phosphate
6-Phospho-D-gluconate
gnd
CO2
D-Ribulose 5-phosphate
D-Xylulose 5-phosphate
D-Ribose 5-phosphate
D-Erythrose 4-phosphate
tktAB
tktAB
Sedoheptulose 7 phosphate
beta-D-Fructose 6-phosphate
Glyceraldehyde 3-phosphate
manA
D-Mannose 6-phosphate
Glycerone phosphate
dhaK
PTS-Man-EIIC
Glycerone
D-Mannose
2-Dehydro-3-deoxy-D-arabino-heptonate 7-phosphate
aroF
Orthophosphate
PEP
Thiamin diphosphate
pdhA
2-(alpha-Hydroxyethyl)thiamine diphosphate
CO2
Pyruvate
HCO3
PYC
Citrate
Iso-citrate
2-oxoglutarate
Glutamate
Succinyl-CoA
Oxaloacetate
Malate
Fumarate
Succinate
Acetyl-coA
pdhA
CO2
HCO3
acc
Malonyl-coA
fabH
fabI
Acetyl-[acyl carrier protein]
Malonyl-[acyl carrier protein]
lplA
fabB
fabG
Fatty Acid Biosynthesis
D-4-Hydroxy-2-oxoglutarate
pyc
Glyoxylate + Pyruvate
gadB
argC
rfbC
rfbB
rfbA
pgm
tktAB
gnd
pgi
manA
dhaK
aroF

Fifteen minutes after switching to arginine excess medium, the *L. lactis* TIFN7 culture differentially expressed 175 genes compared to the arginine limited chemostat culture (Supplementary Table 2). These initial responses after the arginine supplementation encompassed the downregulation of genes encoding arginine biosynthesis, and the upregulation of pyrimidine, purine and fatty acid biosynthesis, single-carbon (C1-) metabolism, pyruvate metabolism, and pentose phosphate pathways (Figure 4). A large subset of these initial transcriptional responses (81 genes) that are associated with pyrimidine, purine and fatty acid biosynthesis, single-carbon (C1-) metabolism as well as arginine biosynthesis continued to be differentially expressed compared to the arginine limited culture even after 120 minutes of arginine excess growth (Supplementary Table 3). Interestingly, 52 genes, which were not affected after 15 minutes of arginine excess growth, were differentially regulated after 30 minutes of arginine excess growth compared to the arginine limited growth (Supplementary Table 2). Of these genes, 12 were upregulated and related to stress response (cold shock and other general stress proteins) as well as to secondary metabolism (such as o-succinylbenzoate-CoA synthase, aconitate hydratase) while the residual 40 genes were downregulated and associated with several functions related to amino acid metabolism (threonine dehydrotase, aspartokinase, 5-enolpyruvylshikimate-3-phosphate synthase), stress response (spermidine and putrescine ABC transporter), and some other transporters with unknown specificity, transposases and hypothetical proteins (Supplementary Table 3). Prolongation of growth under arginine excess for 120 minutes led to a further modulation of 54 genes that were unaffected at earlier time points (15 and 30 minutes). The 35 upregulated genes in this gene-set have functions related to purine metabolism, replication and recombination, fatty acid metabolism and biosynthesis of secondary metabolites. The 19 remaining, down regulated, genes corresponded to phages, dNTP metabolism, and competence (Supplementary Table 3).

For comparative analysis, the arginine non-responsive strain *L. lactis* TIFN5 was grown in chemostat in duplicate at low-arginine (0.05 mM) concentration. Both replicates of the low-arginine (0.05mM) culture of *L. lactis* TIFN5 were established to be glucose limited by the observation that residual glucose was not detectable (data not shown). The transcriptomes of these *L. lactis* TIFN5 chemostat cultures were determined and compared with the transcriptome profiles of the *L. lactis* TIFN7 chemostat cultures. To this end, the absolute gene expression values obtained for *L. lactis* TIFN5 grown under low-arginine chemostat conditions (See Figure 4 and Supplementary Table 4) were compared with the gene expression values of *L.*

lactis TIFN7 cultures. The comparative analysis of the low-arginine concentration growth gene expression profiles revealed that *L. lactis* TIFN7 displayed very low transcriptional activity of arginine, pyrimidine and purine biosynthesis pathways relative to the same pathways in *L. lactis* TIFN5 grown under the same conditions. After arginine supplementation, the expression levels of the genes encoding the enzymes of the arginine biosynthesis pathway were further decreased in *L. lactis* TIFN7 in the presence of excess arginine, which is expected. On the other hand, the expression levels of purine and pyrimidine biosynthesis pathways of *L. lactis* TIFN7 were elevated to levels that corresponded more closely to the levels observed in the arginine non-responsive *L. lactis* TIFN5 (Figure 4). Therefore, the initial transcriptional responses of *L. lactis* TIFN7 (after 15 minutes) appeared to largely compensate for the difference in predicted activity of the purine and pyrimidine biosynthesis pathways relative to that observed in *L. lactis* TIFN5. This finding indicates that conditions of arginine excess (50 mM) elicit metabolic responses in *L. lactis* TIFN7 that are already operational in the non-responsive *L. lactis* TIFN5 strain under low-arginine conditions (0.05 mM). Notably, this clearly contrasts with the expression levels of most of the genes corresponding to the pentose phosphate pathway, and part of the single carbon, fatty acid biosynthesis, and pyruvate metabolism, which already displayed similar (or higher) expression values in *L. lactis* TIFN7 as compared to *L. lactis* TIFN5 under low-arginine conditions, and were further elevated in *L. lactis* TIFN7 after switching to the arginine supplemented medium (Figure 4).

The implication of the above described regulatory changes in expression of genes encoding arginine, pyrimidine, purine biosynthesis pathways as well as single carbon metabolism of *L. lactis* TIFN7 are further discussed below.

Expression of Genes Involved in Arginine Metabolism

Arginine biosynthesis in *L. lactis* uses glutamate as a primary substrate, which is initially converted into ornithine by several enzymatically catalysed steps. Ornithine is subsequently converted into citrulline and arginino-succinate by the enzymes of the urea cycle, where carbamoyl phosphate and aspartate are used as co-substrates (21, 22). The subsequent break down of arginino-succinate leads to the formation of arginine and fumarate. The corresponding genes encoding the enzymes involved in this biosynthetic pathway are found on three separate operons *argCJDBF*, *argGH*, and *gltS-argE* in *L. lactis* (23). Arginine biosynthesis and catabolic operons were shown to be tightly regulated by the availability of arginine in the environment

in *L. lactis* MG1363 (24). In the absence of arginine, the regulator ArgR has higher affinity for the binding sites (*arc* boxes) located upstream of the arginine catabolic (*arc*) operon. However, in presence of arginine, ArgR forms a complex with AhrC, which has high affinity for alternative *arc* boxes that are located upstream of the arginine biosynthesis operons, thereby downregulating arginine biosynthesis capacity. Additionally, ArgR is shown to repress its own expression in the presence of arginine (25). Analogous to the transcriptional control of genes encoding arginine biosynthesis in *L. lactis* MG1363, the genes encoding arginine biosynthesis pathway and its regulator *argR* in *L. lactis* TIFN7 were repressed after 15 minutes of exposure to excess arginine and repression was sustained at further time points (30 minutes, and 2 hours after the arginine switch, Supplementary Tables 3 and 4). Remarkably, the expression of genes encoding N-acetyl-gamma-glutamyl-phosphate reductase (*argC*) and argininosuccinate synthase (*argG*) were 54- and 7-fold lower in the *L. lactis* TIFN7 relative to the expression of the same functions in *L. lactis* TIFN5 (Figure 4) under low-arginine (0.05 mM) growth condition. This finding implies that growth of *L. lactis* TIFN7 under low arginine availability may be limited by the low expression level of arginine biosynthesis. Interestingly, none of the arginine catabolic pathways such as the agmatine, D-arginine, urea and gamma aminobutyric acid (GABA) production pathways, or the arginine deiminase (ADI) pathway are completely encoded in the genome sequences of either of the strains used in this study. Only the gene encoding the ornithine aminotransferase (encoded by *rocD*) that converts L-ornithine to spermidine appeared to be differentially regulated in *L. lactis* TIFN5 after the arginine supplementation.

The main arginine transport system described in *L. lactis* is an antiporter that exchanges arginine and ornithine (26) and is encoded by the genes *arcD1* and *arcD2* (23). These genes encoding the arginine/ornithine antiporter ArcD was strongly downregulated 15 minutes after the arginine supplementation in *L. lactis* TIFN7 (Supplementary Table 4), which is in agreement with its previously reported repression by increasing arginine and decreasing ornithine intracellular concentrations in *L. lactis* (27). Notably, besides the *arcD1*/*D2*, also some other amino acid transporter encoding genes were downregulated by the arginine supplementation, such as those encoding the glutamate and lysine permeases, as well as several amino acid transport systems of unknown substrate specificity. The downregulation of these additional transporters may imply that TIFN7 required importing several other amino acids under low-arginine (0.05 mM) growth conditions as well.

Expression of Genes Involved in the Biosynthesis of Pyrimidine, Purine and C1 Intermediates

Arginine and pyrimidine biosynthesis pathways are tightly linked through the common intermediate carbamoyl-phosphate (22) and are both subjected to complex transcriptional regulation (28). Carbomoyl-phosphate is synthesized from NH_3 and CO_2 by carbamoyl-phosphate synthase (CPS) with the expenditure of 2 ATP, in which NH_3 is produced from hydrolysis of glutamine by glutamine synthase (encoded by *glnA*). Many species of Gram positive bacteria, such as *Lactobacillus plantarum* and *Bacillus subtilus*, have two distinct CPS enzymes that are regulated either by arginine or pyrimidine attenuation mechanism at the transcription level (29). However, the genome of *L. lactis* appears to encode a single CPS enzyme complex composed of a small and large subunit, respectively encoded by *carA* and *carB* (29). The lactococcal *carA* gene is located within the pyrimidine biosynthesis operon *pyrRPB-carA*, whereas its *carB* gene is transcribed monocistronically (30). The transcription of both the *carA* and *carB* genes of *L. lactis* is subject to feed-back regulation controlled by the intracellular pyrimidine level (31). Therefore, the transcription of lactococcal CPS is only known to be regulated through a pyrimidine mediated attenuation mechanism. The initial transcriptional responses observed in *L. lactis* TIFN7 after switching to arginine excess medium (15 min time point) included the upregulation of the *carA* and *carB* as well as the glutamine synthase encoding *glnA* gene (4.5, 8.9 and 8.5 times, respectively). Notably, the *glnA* and *carA* expression levels of *L. lactis* TIFN7 grown under arginine limited conditions (0.05 mM) were substantially lower (almost 4 fold) than those observed in *L. lactis* TIFN5 grown under the same conditions, illustrating that arginine supplementation compensates for the low expression of these functions in *L. lactis* TIFN7 (Figure 4). These findings imply that *L. lactis* TIFN7 experiences a higher demand for pyrimidines after elevating the external arginine concentration to 50 mM and adapts its gene expressions to elevate pyrimidine biosynthesis capacity. This is also supported by the upregulation of the pyrimidine biosynthesis genes, including aspartate carbamoyltransferase (*pyrB*), dihydroorotase (*pyrC*), dihydroorotate dehydrogenase (*pyrD*), orotate phosphoribosyltransferase (*pyrE*), orotidine-5'-phosphate decarboxylase (*pyrF*), CTP synthase (*pyrG*) and uracil phosphoribosyltransferase *(pyrR),* which were among the genes with the highest fold-change of upregulation upon switching to arginine excess conditions (up-regulated 6 and 22-fold, Supplementary Table 4). Moreover, the absolute expression value of the *cmk* gene encoding cytidylate kinase, which converts UMP and CMP to UTP and CTP respectively, was among the highest expressed genes in *L. lactis* TIFN7 in either

low- or high-arginine medium conditions (Figure 4), whereas this gene appeared not to be significantly expressed in *L. lactis* TIFN5 in low-arginine growth conditions. In addition, also the genes encoding the xanthine/uracil/thiamine transporters were upregulated after the switch of *L. lactis* TIFN7 to a medium with excess arginine (7-fold, Supplementary Table 4). This response elevated the absolute expression level of these genes to a level that is comparable to that observed for *L. lactis* TIFN5 when grown in low-arginine medium. The upregulation of *L. lactis* TIFN7 genes encoding the enzymes of the pyrimidine biosynthesis pathway, as well as *carA*, *carB*, *glnA* and pyrimidine transport genes was at its peak 15 minutes after switching to the medium with excess arginine. The corresponding expression levels after 30 minutes of growth in the presence of excess arginine were lower, but were still 2-to-4 fold higher as compared to the initial gene expression levels observed in arginine limited medium (Supplementary Table 4). This initial transient and strong upregulation of gene expression values may be considered as an adaptation to arginine excess. Overall, the expression of genes linked to pyrimidine biosynthesis and import at any time point in the presence of excess arginine was higher compared to the expression levels in arginine limited growth, which supports the higher need of nucleotides for the elevated growth rate of *L. lactis* TIFN7 after the arginine limitation was relieved.

Purine biosynthesis starts with the formation of 5-phosphoribosyl diphosphate (PRPP) from ribose-5-phosphate, which is derived from the pentose phosphate pathway. PRPP is an precursor in both purine and pyrimidine biosynthesis and is converted into imidazole monophosphate (IMP) in 10 enzymatic steps for the synthesis of purine nucleotides AMP and GMP (32). The genes coding for the enzymes located on the purine biosynthesis pathway are clustered into four operons: *purCSQLF*, *purMN*, *purDEK* and *hprT-purH* for *L. lactis* (33). PRPP is a known inducer for the expression of the *pur* operons in *L. lactis* (34), and the activation mechanism involves the formation of a PRPP–PurR complex, which has a higher affinity for RNA polymerase (35). A comparative analysis of *L. lactis* ssp. *cremoris* proteomes obtained under standard and purine-starved conditions revealed that the latter condition led to upregulation of purine biosynthesis, which was synchronized with the upregulation of the genes *glyA* and *fhs* that are involved in the formation of single carbon-intermediates, as well as the upregulation of some glycolytic enzymes such as fructose biphosphate aldolase and pyruvate kinase (36). Our transcriptome analyses indicated that the absolute expression levels of the purine biosynthesis operons was extremely low in *L. lactis* TIFN7 under arginine limited conditions, relative to the expression levels of the same genes in *L. lactis* TIFN5 grown under the same conditions (between

12- and 100-fold difference). Upon switching the *L. lactis* TIFN7 to arginine excess conditions, the expression of the *pur* operons increased 3- to 15-fold already after 15 minutes, illustrating the generic nucleotide demand exerted by increased growth rates, but still retaining the *pur*-operon expression values below those observed in *L. lactis* TIFN5 under low-arginine growth conditions. Moreover, in accordance with the proteomics work conducted on *L. lactis* under purine starvation, upon arginine excess the *L. lactis* TIFN7 also upregulated the *glyA* and *fhs* genes to enhance the supply of single carbon-formate intermediates, and the genes corresponding to the pentose-phosphate pathway that supplies the ribose-5-phosphate substrate required for the biosynthesis of purines.

Single carbon (C1) metabolism is very closely connected to purine, pyrimidine and arginine biosynthesis pathways. Several C1-folate intermediates are required for the formation of purines, and carbon dioxide is used as the main substrate for arginine, pyrimidine and purine biosynthesis (Figure 4). Apart from substrate requirements, aerobic or anaerobic growth conditions were reported to influence the efficiency of these biosynthetic pathways. For example, the activity of the enzyme PurF of *Bacillus subtilus* requires anaerobiosis, and was shown to be inactivated in the presence of oxygen during the stationary phase of growth (37). In another study, the arginine responsive CPS of *L. plantarum* was found to be active only in CO_2 enriched air and this was proposed to be due to the low affinity of CPS for CO_2/HCO_3^- as a substrate (38). Even though *L. lactis* harbours only a pyrimidine responsive CPS, this enzyme may be affected by similar limitations. Thereby, the availability of $C0_2$ and single carbon intermediates seem to be critical for the efficiency of these biosynthetic pathways. Strikingly, the only absolute gene expression values that were higher in arginine-limited *L. lactis* TIFN7 cultures compared to *L. lactis* TIFN5 grown under low-arginine conditions, corresponded to functions related to the metabolism of single carbon-folate intermediates and CO_2 producing pathways (Figure 4). These functions included the *aroF* gene product that is involved in folate biosynthesis, the cysteine synthase (*cbs*) that plays a role in C1-tetrahydrofolate conversions, acetoin dehydrogenase (*pdhA*) that produces $C0_2$, and the genes *rfbA*, *rfbB*, *rfbC*, *PTS-Man-EIIC*, the products of which are all related to the pentose phosphate pathway that generates ribose-5-phosphate. The very high expression values of genes involved in single carbon metabolism in the arginine-limited culture of *L. lactis* TIFN7 relative to their expression in *L. lactis* TIFN5 growing in the same medium, indicates that *L. lactis* TIFN7 displayed high activity of C1 production pathways prior to the relief of the arginine limitation, but failed to overcome the overall limitations in arginine, purine and pyrimidine biosynthesis pathways.

Complementation of C1-Limitations in Arginine and Pyrimidine Biosynthesis Pathways

In order to further understand the limitations observed in arginine, purine and pyrimidine biosynthesis as well as single carbon metabolism of *L. lactis* TIFN7 under arginine-limiting conditions, growth of this strain was investigated on LM17 agar media supplemented with aspartate, arginine, uracil and folate, in both normal air and CO_2 enriched (5%) air conditions. The LM17 propagated culture of *L. lactis* TIFN7 was plated on each of the supplemented LM17 agar media and the recovery of colony forming units was determined. Supplementation of LM17 agar with aspartate, arginine, uracil and folate increased the CFU recovery rate to a level similar to that observed when LM17 agar plates were incubated in 5% CO_2 enriched air (Supplementary Figure 3), demonstrating that C1- compounds are the limiting factor for *L. lactis* TIFN7 cell recovery on regular LM17 medium. Additionally, incubation of the aspartate, arginine, uracil, or folate supplemented LM17 agar media in 5% CO_2 enriched atmosphere did not increase the CFU recovery further, suggesting that all these supplementations fulfil the same growth requirement in this *L. lactis* TIFN7. This is in agreement with the observed high expression of several C1-producing pathways in arginine-limited cultures of *L. lactis* TIFN7 (see above) and also with the further upregulation of these pathways after switching the culture to arginine excess medium. The requirement for elevated CO_2 availability for growth of *L. lactis* TIFN7 on LM17 agar could be related to a relatively low affinity of the carbomoyl phosphate synthatase for CO_2 and thereby limiting the formation of carbamoyl phosphate, which has also been documented for *L. plantarum* (38).

Overall, our data suggest that C1-availability (e.g. compensated by $C0_2$ enrichment) is the major limitation in *L. lactis* TIFN7 cultivated on non-supplemented LM17 agar and the relief of the growth limitation by supplementing the LM17 medium with aspartate, arginine, folate and uracil implies that these resources either directly lead to the elevated C1-sources, or they decrease the requirement for C1-sources. For example, folate or CO_2 directly elevates the amount of C1 compounds, and aspartate may facilitate CO_2 production via decarboxylation reactions or by increasing the availability of aspartate for aspartate carbamoyltransferase in pyrimidine biosynthesis pathways. Additionally, supplementation with uracil or arginine eliminates the requirement of the endogenous biosynthesis of a metabolite that requires a C1-source in its synthetic pathway. In conclusion, any of the supplements (folate, aspartate, arginine, and uracil) or CO_2 enriched growth itself

in the complementation experiments increased the cell recovery probably either by directly relieving the C1-source limitation or by relieving the arginine or pyrimidine end product limitation in which the C1-sources were utilized.

Concluding Remarks

The regulation of arginine and pyrimidine biosynthesis in *L. lactis* is controlled through the pyrimidine responsive carbamoyl-phosphate synthase genes *carA* and *carB* in *L. lactis* (31). The initial low expression levels of genes encoding the pyrimidine, purine and arginine biosynthesis pathways indicated that *L. lactis* TIFN7 had limited activity in these pathways as compared to the activity of *L. lactis* TIFN5 under arginine-limiting conditions. Interestingly, only the genes with the functions related to the metabolism of single carbon-folate intermediates and CO_2 producing pathways were very active in *L. lactis* TIFN7 under arginine-limiting conditions. The *L. lactis* TIFN7 cells did not show any pyrimidine limitation response under this limited activity in steady state condition. Instead, the pyrimidine limitation response started only after the arginine supplementation as evidenced with the enhanced expression levels of *carA, carB* and all other genes located in the pyrimidine biosynthesis pathway (Figure 4). This observation suggests that under arginine-limiting condition, *L. lactis* TIFN7 was limited for C-1 resources, which restricted the amount of arginine that can be synthesized for growth, and this arginine-restricted growth rate limits the requirement for nucleotide synthesis capacity. However, when the arginine supplementation relieved the cells from the C1- demand of arginine biosynthesis, a higher growth rate became possible by the increased availability of arginine and C1-sources, which were not required anymore for arginine biosynthesis. Therefore, the expression of genes belonging to purine and pyrimidine biosynthesis pathways were elevated to sustain higher growth rates as were seen after arginine supplementations to the *L. lactis* TIFN7 chemostat cultures. Furthermore, the relative degree of growth enhancement of TIFN7 on LM17 agar plates that were supplemented with C1-sources (folate and growth in 5% CO2 enriched air), pyrimidine (uracil) and arginine were found to be very similar, indicating that they liberate the *L. lactis* TIFN7 cells from growth arrest either by directly relieving the C1-source limitation or by relieving the arginine or pyrimidine end product limitation.

Acknowledgements

We would like to acknowledge Dr. Sacha van Hijum and V.C.L. de Jager for the *L. lactis* TIFN7 and *L. lactis* TIFN5 genome assembly and gene annotations. This work was funded by Top Institute Food and Nutrition (TIFN) and is part of Kluyver Centre for Genomics of Industrial Fermentation, a Centre of Excellence of The Netherlands Genomics Initiative.

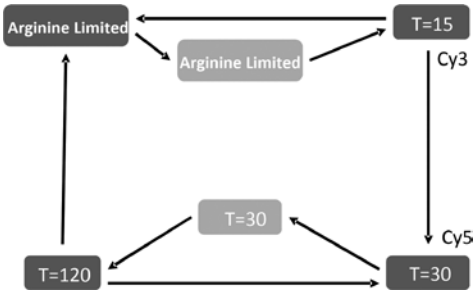
References

1. **Nocker A, Camper AK.** 2009. Novel approaches toward preferential detection of viable cells using nucleic acid amplification techniques. Fems Microbiol Lett **291:**137-142.
2. **Oliver JD.** 2005. The viable but nonculturable state in bacteria. Journal of microbiology **43 Spec No:**93-100.
3. **Kell DB, Kaprelyants AS, Weichart DH, Harwood CR, Barer MR.** 1998. Viability and activity in readily culturable bacteria: a review and discussion of the practical issues. Antonie van Leeuwenhoek **73:**169-187.
4. **Barer MR, Harwood CR.** 1999. Bacterial viability and culturability. Advances in microbial physiology **41:**93-137.
5. **Desnues B, Cuny C, Gregori G, Dukan S, Aguilaniu H, Nystrom T.** 2003. Differential oxidative damage and expression of stress defence regulons in culturable and non-culturable *Escherichia coli* cells. EMBO reports **4:**400-404.
6. **G. K. Joshi JJ, J. P. Bhatt.** 2014. Metagenomics: The Exploration of Unculturable Microbial World, p. 105-115. *In* Indu Ravi MB, Jyoti Saxena (ed.), Advances in Biotechnology.
7. **Streit WR, Schmitz RA.** 2004. Metagenomics--the key to the uncultured microbes. Current opinion in microbiology **7:**492-498.
8. **Stuart MR, Chou LS, Weimer BC.** 1999. Influence of carbohydrate starvation and arginine on culturability and amino acid utilization of *Lactococcus lactis* subsp. *lactis.* Appl Environ Microbiol **65:**665-673.
9. **Ayad EHE, Verheul, A., de Jong, C.,Wouters, J.T.M., Smit, G. .** 1999. Flavour forming abilities and amino acid requirements of *Lactococcus lactis* strains isolated from artisanal and non-dairy origin. International Dairy Journal **9:**11.
10. **Bachmann H, Starrenburg MJ, Molenaar D, Kleerebezem M, van Hylckama Vlieg JE.** 2012. Microbial domestication signatures of *Lactococcus lactis* can be reproduced by experimental evolution. Genome research **22:**115-124.
11. **Godon JJ, Delorme C, Bardowski J, Chopin MC, Ehrlich SD, Renault P.** 1993. Gene inactivation in *Lactococcus lactis*: branched-chain amino acid biosynthesis. Journal of bacteriology **175:**4383-4390.
12. **Delorme C, Godon JJ, Ehrlich SD, Renault P.** 1993. Gene inactivation in *Lactococcus lactis:* histidine biosynthesis. Journal of bacteriology **175:**4391-4399.
13. **Stadhoouders J, Leenders, G.J.M.** 1984. Spontaneously developed mixed-strain cheese starters. Their behaviours towards phages, and theri use in Dutch cheese industry. Netherlands Milk and Dairy Journal **38:**25.
14. **Reddy MS, Vedamuthu ER, Washam CJ, Reinbold GW.** 1972. Agar medium for differential enumeration of lactic streptococci. Appl Microbiol **24:**947-952.
15. **Sieuwerts S, de Bok FA, Mols E, de vos WM, Vlieg JE.** 2008. A simple and fast method for determining colony forming units. Letters in applied microbiology **47:**275-278.
16. **Stevens MJ, Wiersma A, de Vos WM, Kuipers OP, Smid EJ, Molenaar D, Kleerebezem M.** 2008. Improvement of *Lactobacillus plantarum* aerobic growth as directed by comprehensive transcriptome analysis. Appl Environ Microbiol **74:**4776-4778.
17. **Meijerink M, van Hemert S, Taverne N, Wels M, de Vos P, Bron PA, Savelkoul HF, van Bilsen J, Kleerebezem M, Wells JM.** 2010. Identification of genetic loci in *Lactobacillus plantarum* that modulate the immune response of dendritic cells using comparative genome hybridization. PloS one **5:**e10632.

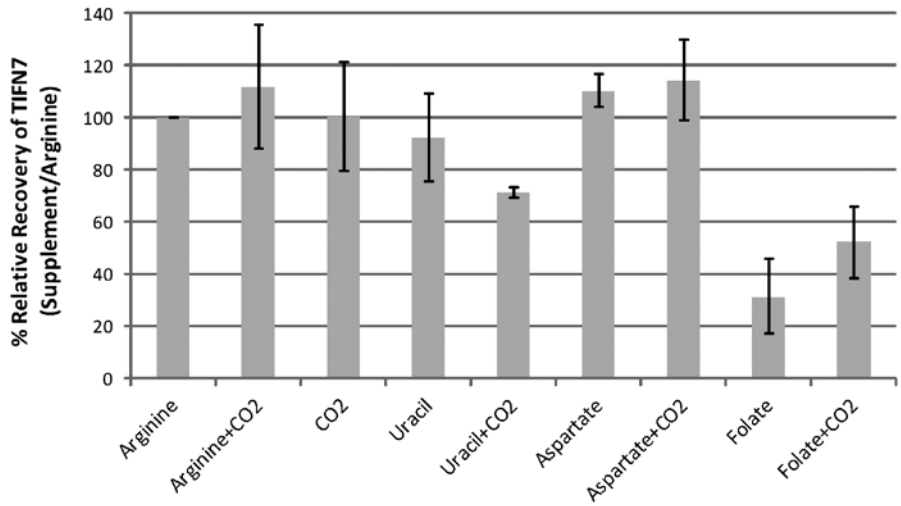
18. **van Hijum SA, Garcia de la Nava J, Trelles O, Kok J, Kuipers OP.** 2003. MicroPreP: a cDNA microarray data pre-processing framework. Applied bioinformatics **2:**241-244.
19. **Baldi P, Long AD.** 2001. A Bayesian framework for the analysis of microarray expression data: regularized t -test and statistical inferences of gene changes. Bioinformatics **17:**509-519.
20. **Letunic I, Yamada T, Kanehisa M, Bork P.** 2008. iPath: interactive exploration of biochemical pathways and networks. Trends Biochem Sci **33:**101-103.
21. **Larsen R, Buist G, Kuipers OP, Kok J.** 2004. ArgR and AhrC are both required for regulation of arginine metabolism in *Lactococcus lactis*. Journal of bacteriology **186:**1147-1157.
22. **Cunin R, Glansdorff N, Pierard A, Stalon V.** 1986. Biosynthesis and metabolism of arginine in bacteria. Microbiological reviews **50:**314-352.
23. **Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, Ehrlich SD, Sorokin A.** 2001. The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. Genome research **11:**731-753.
24. **Larsen R, Kok J, Kuipers OP.** 2005. Interaction between ArgR and AhrC controls regulation of arginine metabolism in *Lactococcus lactis.* The Journal of biological chemistry **280:**19319-19330.
25. **Lim DB, Oppenheim JD, Eckhardt T, Maas WK.** 1987. Nucleotide sequence of the argR gene of *Escherichia coli* K-12 and isolation of its product, the arginine repressor. Proceedings of the National Academy of Sciences of the United States of America **84:**6697-6701.
26. **Driessen AJ, Poolman B, Kiewiet R, Konings W.** 1987. Arginine transport in *Streptococcus lactis i*s catalyzed by a cationic exchanger. Proceedings of the National Academy of Sciences of the United States of America **84:**6093-6097.
27. **Poolman B, Driessen AJ, Konings WN.** 1987. Regulation of arginine-ornithine exchange and the arginine deiminase pathway in *Streptococcus lactis*. Journal of bacteriology **169:**5597-5604.
28. **Caldara M, Charlier D, Cunin R.** 2006. The arginine regulon of *Escherichia coli:* whole-system transcriptome analysis discovers new genes and provides an integrated view of arginine regulation. Microbiology **152:**3343-3354.
29. **Nikoloff Herve HJ-C, Bringel Francoise.** 2001. Carbomoyl-phosphate synthetases in lactic acid bacteria and other Gram-positive bacteria. Lait **81:**9.
30. **Martinussen J, Schallert J, Andersen B, Hammer K.** 2001. The pyrimidine operon pyrRPB-carA from *Lactococcus lactis*. Journal of bacteriology **183:**2785-2794.
31. **Martinussen J, Hammer K.** 1998. The carB gene encoding the large subunit of carbamoylphosphate synthetase from *Lactococcus lactis* is transcribed monocistronically. Journal of bacteriology **180:**4380-4386.
32. **Kilstrup M, Hammer K, Ruhdal Jensen P, Martinussen J.** 2005. Nucleotide metabolism and its control in lactic acid bacteria. FEMS microbiology reviews **29:**555-590.
33. **Nilsson D, Kilstrup M.** 1998. Cloning and expression of the *Lactococcus lactis* purDEK genes, required for growth in milk. Appl Environ Microbiol **64:**4321-4327.
34. **Kilstrup M, Martinussen J.** 1998. A transcriptional activator, homologous to the Bacillus subtilis PurR repressor, is required for expression of purine biosynthetic genes in *Lactococcus lactis.* Journal of bacteriology **180:**3907-3916.
35. **Kilstrup M, Jessing SG, Wichmand-Jorgensen SB, Madsen M, Nilsson D.** 1998. Activation control of pur gene expression in *Lactococcus lactis*: proposal for a consensus activator binding sequence based on deletion analysis and site-directed mutagenesis of purC and purD promoter regions. Journal of bacteriology **180:**3900-3906.
36. **Beyer NH, Roepstorff P, Hammer K, Kilstrup M.** 2003. Proteome analysis of the purine stimulon from *Lactococcus lactis.* Proteomics **3:**786-797.
37. **Turnbough CL, Jr., Switzer RL.** 1975. Oxygen-dependent inactivation of glutamine phosphoribosylpyrophosphate amidotransferase in stationary-phase cultures of *Bacillus subtilis.* Journal of bacteriology **121:**108-114.
38. **Nicoloff H, Hubert JC, Bringel F.** 2000. In *Lactobacillus plantarum*, carbamoyl phosphate is synthesized by two carbamoyl-phosphate synthetases (CPS): carbon dioxide differentiates the arginine-repressed from the pyrimidine-regulated CPS. Journal of bacteriology **182:**3416-3422.

Supplementary Information for Chapter 4

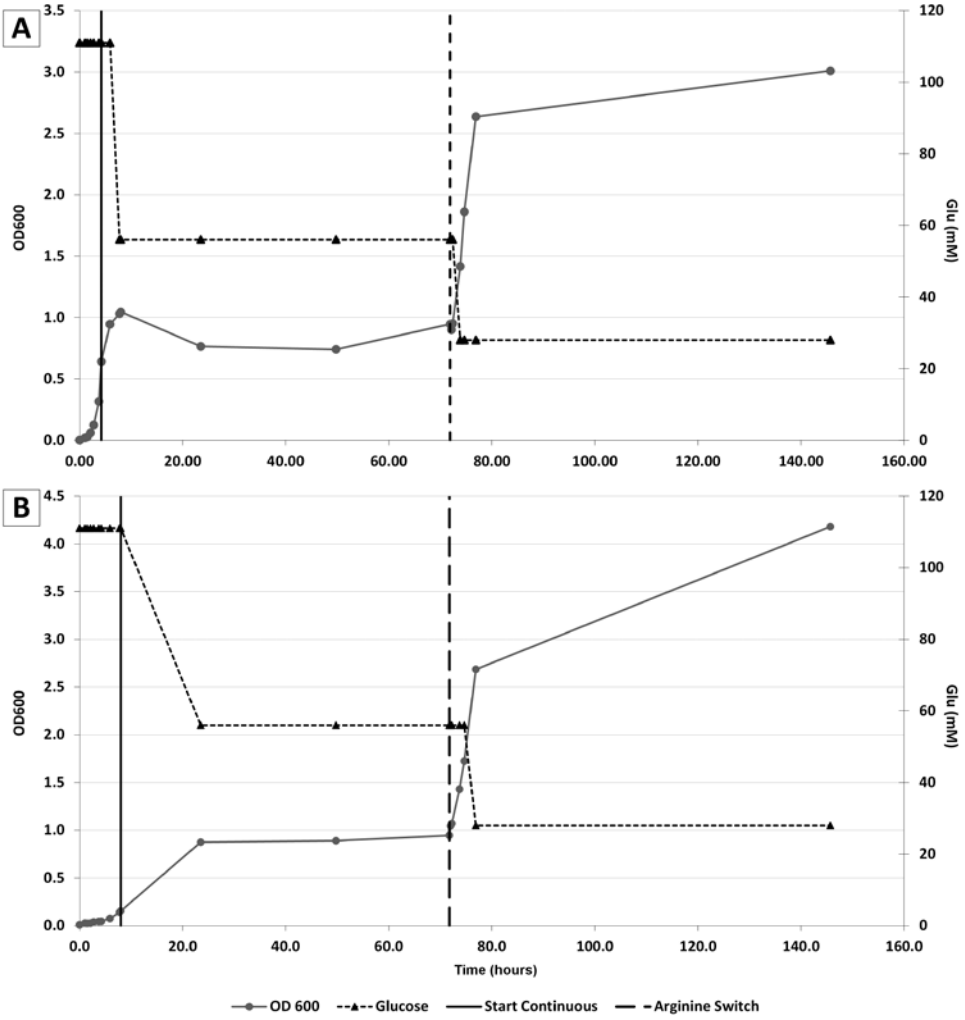
Supplementary Figures



Supplementary Figure 1. The DNA Microarray hybridization scheme of *L. lactis* TIFN7. Hybridizations were designed in closed-loop model. Arginine limited, T=15, T=30, and T=120 are representatives for the time points just before the arginine supplementation and 15 minutes, 30 minutes, and 120 minutes after the arginine supplementation, respectively. The dark grey and the light grey rectangles represent the duplicates for mRNA samples of each time point. The arrows indicate the hybridization orientation. The bottom of the arrow is representative for Cy3 labelled sample, and the top of the arrow is representative for the Cy5 labelled sample.



Supplementary Figure 3. The CFU recovery of *L. lactis* TIFN7 on LM17 agar plates supplemented with arginine, uracil, aspartate and folate under normal air and 5% CO_2 enriched air conditions compared to the CFU recovery obtained on arginine supplemented LM17 agar plates. The CFUs are the average.



Supplementary Figure 2. Chemostat cultivations of *L. lactis* ssp. *cremoris* TIFN7. The chemostats were performed in duplicate (Graphs A&B). The grey circles and the black triangles represent the optical density and the residual glucose concentration in the chemostat culture, respectively. The continuous black line indicates start of the continuous fermentation, and the dotted black line indicates the arginine supplementation. The first part of the cultivation is under arginine limitation (0.05mM) and the arginine concentration was increased to 50mM after the black dotted line.

Supplementary Table 1.The Recipe of the arginine limited fermentation medium with 0,05 mM arginine and 100mM glucose concentration (pH 6.8)

Components of Chemically Defined Medium*	Amount	Units
Basal Medium	8.2	liter
100mM C-bron in medium	500	ml
10x Amino acid solution	1000	ml
100x DNA precursor mix	100	ml
100x Vitamin solution	100	ml
100x Metal solution	100	ml
Total Volume	10	liter

* The ingredients of each component is given below.

Basal Medium(in 8.2 Liter volume)	
K2HPO4	10 gr
KH2PO4	50 gr
Na-acetate	10 gr
(NH4)3-citrate	6 gr
tyrosine	2.5 gr
β-glycerophosphate	19 gr
100mM C-bron in medium (in 500 ml volume)	
Glucose monohydrate	198.17 gr
10x Amino acid solution(in 1 Liter volume)	
alanine	2.4 gr
arginine	0.0883 gr
aspartic acid	4.2 gr
cysteïne-HCl	1.3 gr
glutamic acid	5 gr
histidine	1.5 gr
isoleucine	2.1 gr
leucine	4.75 gr
lysine	4.4 gr
phenylalanine	2.75 gr
proline	6.75 gr
serine	3.4 gr
threonine	2.25 gr
tryptophane	0.5 gr
valine	3.25 gr
glycine	1.75 gr
methionine	1.25 gr
L(+)Ascorbic acid	5 gr

100x DNA precursor mix (in 100 ml volume)	
adenine	100 mg
guanine	100 mg
xanthine	100 mg
uracil	100 mg
100x Vitamin solution (in 1 Liter volume)	
pyridoxamine-HCl	500 mg
D-biotin	250 mg
6,8-thioctic acid	250 mg
pyridoxine-HCl	200 mg
nicotinic acid	100 mg
Ca-(D+)pantothenate	100 mg
Riboflavin	100 mg
thiamin-HCl	100 mg
vitamin B12	100 mg
p-aminobenzoëic acid	1000 mg
orotic acid	500 mg
thymidine	500 mg
inosine	500 mg
100x Metal solution (in 1 Liter volume)	
$MgCl_2 \times 6 H_2O$	20 gr
$CaCl_2 \times 2 H_2O$	5 gr
$MnCl_4 \times H_2O$	1.6 gr
$FeCl_3 \times 6 H_2O$	0.3 gr
$FeCl_2 \times 4 H_2O$	0.5 gr
$ZnSO_4 \times 7 H_2O$	0.5 gr
$CoSO_4 \times 7 H_2O$	0.25 gr
$CuSO_4 \times 5 H_2O$	0.25 gr
$(NH_4)_6Mo_7O_{24} \times 4 H_2O$	0.25 gr

Supplementary Table 2. The number of differentially expressed genes in time series transcriptome analysis of TIFN7 chemostat cultures (FDR< 0,05)

	Time Point	No of genes		Shared with		
				15 minutes/SS	30 Minutes/SS	2 Hours/SS
Genes that were differentially expressed for the first time at	15 Minutes/SS	175	DE:	145	82	72
			DEL:	30	17	9
			ND:	0	76	94
	30 Minutes/SS	52	DE:	0	40	9
			DEL:	0	12	1
			ND:	52	0	42
	2 hours/SS	54	DE:	0	0	40
			DEL:	0	0	14
			ND:	54	54	0
Total number of differentialy expressed genes at each time point				175	151	145

DE: Differentially expressed more than 2 folds
DEL: Differentially expressed less than 2 folds
ND: Not detected
SS: Compared to steady state arginine limiting condition

Supplementary Table 3. Differentially expressed genes (FDR< 0,05) in time series transcriptome analysis of TIFN7 chemostat cultures. Genes that are differentially upregulated more than two folds are highlighted with dark grey colour and the genes that were differentially downregulated are highlighted with light grey colour. ND and SS are used as abbreviations for ``not significantly detected`` and ``steady state condition just before arginine switch`` respectively.

RAST Identifier	Annotation	COG/EC	15 minutes/SS		30 minutes/SS		2 hours/SS	
			Exp.	FDR	Exp.	FDR	Exp.	FDR
Genes that were differentially expressed 15 minutes after the arginine supplementation								
820	Abortive phage resistance protein abiGi		5.003	3.03E-02	5.344	8.06E-04	5.746	4.28E-02
2320	Phosphoribosylformylglycinamidine synthase, PurS subunit (EC 6.3.5.3)	E6.3.5.3	3.830	4.50E-05	3.003	1.65E-06	5.479	3.28E-07
2319	Phosphoribosylformylglycinamidine synthase (EC 6.3.5.3)	COG0047	3.580	3.88E-06	3.100	4.03E-07	5.346	5.88E-09
3172	GMP reductase (EC 1.7.1.7)	COG0516	3.555	7.53E-06	3.515	9.99E-10	2.442	7.62E-04
2497	Nitrogen regulatory protein P-II	COG0347	3.487	1.05E-04	2.336	2.81E-03	3.849	3.52E-05
2321	Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)	COG0152	3.466	5.96E-05	2.457	6.56E-04	4.760	6.27E-07
2318	Phosphoribosylformylglycinamidine synthase, synthetase subunit (EC 6.3.5.3)	COG0046	3.443	1.41E-04	3.575	2.32E-07	6.050	8.78E-08
2786	Lysine-specific permease	COG0833	3.361	7.05E-05	2.832	1.62E-05	2.050	1.23E-02
1975	IMP cyclohydrolase (EC 3.5.4.10)	COG0138	3.332	7.24E-04	2.224	1.53E-03	3.559	1.41E-04
2115	Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)	COG0150	3.305	8.79E-05	2.295	8.29E-05	3.566	3.98E-05
504	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)	COG0458	3.099	9.37E-05	3.265	2.36E-02	3.769	2.11E-02
2145	Xanthine/uracil/thiamine/ascorbate permease family protein	COG2252	3.015	3.16E-07	2.271	7.42E-07	3.014	3.17E-07
502	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)	COG0458	2.861	5.67E-08	1.313	6.87E-05	1.149	3.54E-03
2317	Amidophosphoribosyltransferase (EC 2.4.2.14)	COG0034	2.797	6.18E-05	1.806	2.00E-03	4.243	1.34E-07
907	Adenylosuccinate synthetase (EC 6.3.4.4)	COG0104	2.702	3.20E-07	1.721	1.51E-05	1.840	3.97E-05
1502	Phosphonate ABC transporter phosphate-binding periplasmic component		2.632	6.44E-03	1.873	1.72E-02	3.190	9.78E-04
1503	Phosphonate ABC transporter ATP-binding protein (TC 3.A.1.9.1)	COG3638	2.629	3.58E-02	ND	ND	3.334	4.18E-03
3092	Nucleoside-binding protein	COG1744	2.481	4.01E-07	1.514	1.15E-05	1.545	1.37E-04
1903	Serine hydroxymethyltransferase (EC 2.1.2.1)	COG0112	2.390	3.55E-03	2.307	2.45E-03	2.882	5.58E-04
1336	CTP synthase (EC 6.3.4.2)	COG0504	2.320	6.20E-06	1.378	1.29E-04	1.355	3.34E-03
397	Formate--tetrahydrofolate ligase (EC 6.3.4.3)	COG2759	2.269	1.84E-04	1.759	1.48E-03	2.589	5.25E-05
2817	Glutamine synthetase type I (EC 6.3.1.2)	COG0174	2.180	1.84E-06	1.234	3.31E-04	2.320	4.39E-07
597	Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)	COG0190	2.138	2.10E-03	ND	ND	1.458	4.86E-02
1157	Mevalonate kinase (EC 2.7.1.36)	COG1577	2.106	3.07E-03	ND	ND	1.945	6.99E-03
1969	Phosphoribosylamine--glycine ligase (EC 6.3.4.13)	COG0151	2.072	2.12E-02	ND	ND	3.271	4.22E-04
1974	IMP cyclohydrolase (EC 3.5.4.10) / (EC 2.1.2.3)	COG0138	2.026	2.44E-02	ND	ND	3.265	3.56E-04
2818	glutamine synthetase repressor	COG0789	1.857	1.03E-04	1.465	6.88E-05	1.862	2.87E-05
TIFN7_02880	Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2)	COG0213	1.769	2.61E-03	1.716	1.16E-03	1.458	1.80E-02
1271	Methionine ABC transporter substrate-binding protein	COG1464	1.748	2.86E-04	2.041	1.04E-06	1.493	2.31E-03

141	Heptaprenyl diphosphate synthase component II (EC 2.5.1.30)	E2.5.1.30	1.702	3.99E-04	1.516	1.31E-04	1.664	5.45E-04
3148	Substrate-specific component NiaX of predicted niacin ECF transporter		1.697	5.89E-03	ND	ND	1.383	2.66E-02
751	hypothetical protein (Fe-S Oxidoreductase)	ncbi-gi:116513115	1.686	4.98E-03	1.384	3.04E-02	2.196	4.77E-04
2033	Adenylosuccinate lyase (EC 4.3.2.2)	COG0015	1.663	9.76E-03	1.321	2.08E-02	1.611	1.61E-02
62	Transcriptional regulatory protein	COG0718	1.638	8.36E-04	1.387	1.26E-03	1.362	5.99E-03
1270	Methionine ABC transporter substrate-binding protein	COG1464	1.598	2.27E-03	1.186	1.03E-02	1.100	4.12E-02
1754	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	COG0331	1.596	8.65E-03	ND	ND	1.426	2.52E-02
406	preprotein translocase subunit SecG	ncbi-gi:116511758	1.438	1.71E-02	1.181	4.60E-02	1.269	4.14E-02
2556	tRNA dihydrouridine synthase B (EC 1.-.-.-)	COG0042	1.428	2.46E-02	1.398	1.16E-02	1.369	4.17E-02
140	Heptaprenyl diphosphate synthase component II (EC 2.5.1.30)	E2.5.1.30	1.387	4.81E-03	1.389	3.07E-04	1.257	1.23E-02
1756	3-oxoacyl-[acyl-carrier-protein] synthase, KASII (EC 2.3.1.41)	COG0304	1.364	4.70E-04	1.126	4.19E-04	0.911	1.29E-02
1149	Cell division trigger factor (EC 5.2.1.8)	COG0544	1.333	7.16E-04	0.848	1.59E-02	1.112	3.74E-03
2796	hypothetical phage protein		1.261	2.02E-03	1.133	5.89E-04	1.192	4.08E-03
2567	UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8)	COG0773	1.242	4.74E-03	0.962	6.94E-03	1.030	2.20E-02
1915	Flavodoxin	COG0716	1.181	4.22E-02	1.285	1.16E-02	1.517	1.16E-02
1267	Methionine ABC transporter substrate-binding protein	COG1464	1.174	2.15E-02	1.477	7.53E-04	1.106	4.46E-02
2636	hyphotetical protein	ncbi-gi:116513040	1.163	4.18E-02	1.067	4.03E-02	1.273	2.59E-02
1751	Transcriptional regulator of fatty acid biosynthesis FabT	ncbi-gi:116511579	1.152	1.46E-02	1.145	6.75E-03	1.208	1.29E-02
1758	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)	COG0764	1.144	3.29E-03	0.761	2.82E-02	0.917	1.73E-02
1187	FMN-dependent NADH-azoreductase	COG1182	1.128	8.69E-03	0.901	2.09E-02	1.373	2.27E-03
637	aromatic ring hydroxylating enzyme, PaaD-like protein (DUF59)	COG2151	1.094	2.48E-02	1.101	1.21E-02	1.363	6.02E-03
2300	Tripeptide aminopeptidase (EC 3.4.11.4)	COG2195	1.092	3.04E-03	1.068	7.64E-04	1.147	9.38E-04
TIFN7_01964	Hypothetical protein		1.089	8.68E-03	1.332	8.01E-05	1.524	3.70E-04
2872	hypothetical protein	ncbi-gi:116512933	1.067	3.87E-03	0.888	5.53E-03	1.126	2.28E-03
1753	Acyl carrier protein	COG0236	1.041	5.79E-03	1.227	8.16E-04	1.329	2.26E-04
53	cationic amino acid transporter	COG0531	0.993	8.10E-03	0.974	5.88E-03	1.470	2.22E-04
1994	Transketolase (EC 2.2.1.1)	COG0021	0.987	1.61E-02	ND	ND	0.875	3.99E-02
1931	6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)	COG0362	0.986	3.74E-03	ND	ND	0.734	4.02E-02
197	hypothetical protein	COG0217	0.892	1.31E-02	0.778	2.83E-02	0.774	3.95E-02
1866	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)	COG0764	0.882	4.17E-02	0.980	5.66E-03	1.389	1.58E-03
77	Inner membrane protein translocase component YidC, short form OxaI-like	COG0706	0.862	1.47E-02	0.828	1.12E-02	0.973	7.81E-03
2179	Glutamine ABC transporter, glutamine-binding protein/permease protein	COG0765	0.772	3.70E-02	ND	ND	0.910	1.41E-02
955	FIG00756687: hypothetical protein		-0.853	1.83E-02	ND	ND	0.845	1.67E-02
954	FIG00755954: hypothetical protein		-0.860	2.81E-02	ND	ND	0.940	1.17E-02
2586	Cysteine ABC transporter, substrate-binding protein	ncbi-gi:116513001	-0.980	7.20E-03	-0.792	1.74E-02	-1.115	3.04E-03
2565	Arginyl-tRNA synthetase (EC 6.1.1.19)	COG0018	-1.255	7.07E-03	-1.893	6.51E-05	-1.704	6.31E-04
139	Oxygen-insensitive NAD(P)H nitroreductase / Dihydropteridine reductase	E1.5.1.34	-1.289	1.19E-04	-1.572	3.75E-07	-1.128	6.56E-04

996	Lactose phosphotransferase system repressor	ncbi-gi:116326616	-1.474	1.23E-04	-1.819	6.00E-07	-1.213	1.33E-03
2566	Arginine pathway regulatory protein ArgR, repressor of arg regulon	COG1438	-1.525	9.98E-05	-1.995	1.42E-06	-2.144	4.79E-06
2742	Arginine/ornithine antiporter ArcD	COG0531	-1.717	3.73E-02	-2.300	1.53E-03	-2.109	1.41E-02
2280	glutamate ABC transporter permease protein	COG0765	-2.073	1.30E-02	-2.906	3.72E-05	-3.710	2.33E-05
73	Argininosuccinate lyase (EC 4.3.2.1)	COG0165	-2.232	4.11E-06	-3.412	4.03E-09	-3.417	6.49E-08
72	Argininosuccinate synthase (EC 6.3.4.5)	COG0137	-2.237	5.37E-05	-2.746	7.92E-07	-2.828	3.03E-05
1212	Lysine-specific permease	COG0833	-2.377	7.25E-05	-2.942	7.24E-07	-2.723	4.81E-06
1786	N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38)	COG0002	-2.467	4.12E-06	-4.383	1.12E-09	-4.258	3.19E-07
2278	Amino acid ABC transporter, ATP-binding protein	COG1126	-2.876	1.76E-04	-3.538	3.11E-07	-2.958	5.51E-05
1862	Amino acid ABC transporter, amino acid-binding/permease protein	ncbi-gi:116511362	-3.547	2.11E-08	-6.869	0.00E+00	-6.450	4.95E-09
887	transposase		-3.700	1.46E-04	ND	ND	-1.956	4.28E-02
1865	Acetylornithine deacetylase (EC 3.5.1.16)	EC 3.5.1.16	-3.836	1.02E-09	-6.745	0.00E+00	-8.898	0.00E+00
1864	Acetylornithine deacetylase (EC 3.5.1.16)	EC 3.5.1.16	-3.848	1.20E-07	-8.606	0.00E+00	-8.474	2.51E-12
1863	glutamate or arginine ABC transporter substrate binding protein	ncbi-gi:116511362	-4.148	2.85E-08	-7.449	0.00E+00	-8.617	4.71E-13
2279	Amino acid ABC transporter, ATP-binding protein	COG1126	-5.638	1.56E-05	-5.844	4.75E-10	-7.000	2.54E-06
731	OrfA; hypothetical protein; putative DNA recombinase	ncbi-gi:281491175	5.560	8.60E-04	5.512	1.88E-03	ND	ND
489	transposase	ncbi-gi:125623538	3.056	1.75E-02	3.332	4.13E-04	ND	ND
540	Dihydroorotate dehydrogenase electron transfer subunit (EC 1.3.3.1)	COG0543	3.821	2.61E-05	1.718	1.73E-02	ND	ND
3081	Orotate phosphoribosyltransferase (EC 2.4.2.10)	COG0461	3.592	3.75E-06	1.839	1.55E-03	ND	ND
542	Orotidine 5′-phosphate decarboxylase (EC 4.1.1.23)	COG0284	3.480	4.16E-06	1.248	3.41E-02	ND	ND
2512	Carbamoyl-phosphate synthase small chain (EC 6.3.5.5)	COG0505	3.156	5.64E-08	1.333	2.41E-04	ND	ND
2515	Uracil permease	COG2233	2.921	4.85E-05	1.279	7.18E-03	ND	ND
2513	Aspartate carbamoyltransferase (EC 2.1.3.2)	COG0540	2.816	7.43E-08	1.191	1.02E-03	ND	ND
2516	Uracil phosphoribosyltransferase / Pyrimidine operon regulatory protein PyrR	COG2065	2.738	7.29E-08	1.135	1.86E-03	ND	ND
309	Deoxyribose-phosphate aldolase (EC 4.1.2.4)	COG0274	2.140	5.41E-04	1.429	7.19E-03	ND	ND
3041	Glycyl-tRNA synthetase alpha chain (EC 6.1.1.14)	COG0752	1.777	3.71E-02	1.637	3.05E-02	ND	ND
912	Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)	COG0016	1.501	3.31E-04	1.270	2.63E-04	ND	ND
2428	DNA-directed RNA polymerase omega subunit (EC 2.7.7.6)	COG1758	1.270	6.46E-04	0.838	9.52E-03	ND	ND
1869	alpha-phosphoglucomutase	E5.4.2.2	1.182	2.78E-03	0.733	4.52E-02	ND	ND
TIFN7_01512	hyphotetical protein		1.120	8.74E-03	0.953	5.08E-03	ND	ND
1283	Cysteine synthase (EC 2.5.1.47)	COG0031	1.120	2.26E-03	0.942	2.93E-03	ND	ND
940	Threonyl-tRNA synthetase (EC 6.1.1.3)	COG0441	1.102	4.46E-02	1.090	1.98E-02	ND	ND
2127	Phosphate transport system permease protein PstA (TC 3.A.1.7.1)	COG0581	0.963	1.46E-02	0.733	4.21E-02	ND	ND
152	LSU ribosomal protein L28p	COG0227	0.876	3.64E-02	1.011	1.89E-02	ND	ND
2385	Cell division initiation protein DivIVA	ncbi-gi:116512742	-0.885	4.02E-02	-1.079	3.12E-03	ND	ND
980	Universal stress protein family	ncbi-gi:116326637	-0.901	5.96E-03	-1.062	4.11E-04	ND	ND
2122	Phosphoglycerate mutase family 2	ncbi-gi:116512567	-0.966	1.96E-02	-1.630	1.35E-05	ND	ND

979	Manganese transport protein MntH	COG1914	-0.973	6.05E-03	-1.126	3.21E-04	ND	ND
1766	Gls24 family general stress protein	ncbi-gi:281491279	-1.021	2.01E-03	-0.654	4.21E-02	ND	ND
2060	transglycosylase	ncbi-gi:116512524	-1.030	4.81E-03	-0.858	7.95E-03	ND	ND
58	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha	COG0722	-1.206	4.48E-03	-1.027	2.93E-03	ND	ND
738	FIG00755240: hypothetical protein	COG3382	-1.314	6.50E-03	-1.408	1.65E-03	ND	ND
1375	FIG00755211: hypothetical protein		-1.662	1.92E-02	-1.578	2.14E-02	ND	ND
2116	ClpB protein	COG0542	-1.908	4.20E-02	-2.022	8.11E-04	ND	ND
468	L-proline glycine betaine ABC transport system permease protein ProV	COG1125	-2.774	8.68E-03	-2.595	7.44E-04	ND	ND
2936	Glutamate decarboxylase (EC 4.1.1.15)	COG0076	-3.599	7.43E-04	-1.749	2.35E-02	ND	ND
3080	Dihydroorotase (EC 3.5.2.3)	COG0044	4.469	3.37E-06	ND	ND	ND	ND
541	Dihydroorotate dehydrogenase, catalytic subunit (EC 1.3.3.1)	COG0167	3.834	5.73E-05	ND	ND	ND	ND
1372	PTS system, cellobiose-specific IIA component (EC 2.7.1.69)	COG1447	3.183	2.36E-03	ND	ND	ND	ND
1809	regulator of polyketide synthase expression	ncbi-gi:116511630	2.829	1.79E-02	ND	ND	ND	ND
1364	Multiple sugar ABC transporter, ATP-binding protein	COG3839	2.268	1.84E-02	ND	ND	ND	ND
1617	cation-transporting ATPase, E1-E2 family	COG0474	2.227	3.66E-02	ND	ND	ND	ND
3098	Ribosomal RNA small subunit methyltransferase C (EC 2.1.1.52)	COG2813	2.149	4.75E-03	ND	ND	ND	ND
2947	Poly(glycerophosphate chain) D-alanine transfer protein DltD		2.134	3.60E-02	ND	ND	ND	ND
1799	YibE/F superfamily transporter	UNIPROT:F2HJA2	2.027	4.24E-02	ND	ND	ND	ND
2090	Maltose/maltodextrin ABC transporter, substrate binding protein MalE	COG2182	2.026	5.85E-03	ND	ND	ND	ND
1355	PTS system, trehalose-specific IIA component (EC 2.7.1.69)	COG2190	1.981	1.91E-02	ND	ND	ND	ND
1373	PTS system, cellobiose-specific IIB component (EC 2.7.1.69)	COG1440	1.894	3.61E-04	ND	ND	ND	ND
1354	PTS system, trehalose-specific IIB component (EC 2.7.1.69)	COG1264	1.894	2.67E-02	ND	ND	ND	ND
2438	N-acetylmuramoyl-L-alanine amidase (involved in cell division)	E3.5.1.28	1.858	3.66E-02	ND	ND	ND	ND
1977	Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8)	COG0634	1.735	4.67E-02	ND	ND	ND	ND
2089	Maltose/maltodextrin ABC transporter, substrate binding protein MalE	COG2182	1.714	3.42E-02	ND	ND	ND	ND
2582	Glycerol uptake facilitator protein	COG0580	1.702	8.39E-04	ND	ND	ND	ND
1284	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA	COG4166	1.676	1.84E-02	ND	ND	ND	ND
398	NLP/P60 family protein	COG0791	1.635	4.65E-02	ND	ND	ND	ND
3202	FIG00756303: hypothetical protein	ncbi-gi:116511972	1.602	1.19E-04	ND	ND	ND	ND
237	Phosphoenolpyruvate-dihydroxyacetone phosphotransferase (EC 2.7.1.121)	COG2376	1.589	4.87E-03	ND	ND	ND	ND
1750	Mannose-6-phosphate isomerase (EC 5.3.1.8)	COG1482	1.587	4.97E-03	ND	ND	ND	ND
3227	Probable transport protein	COG0531	1.512	2.87E-02	ND	ND	ND	ND
TIFN7_01499	Hypothetical protein		1.455	4.20E-02	ND	ND	ND	ND
3146	UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.1.1.158)	COG0812	1.412	3.29E-02	ND	ND	ND	ND
1171	Heat shock protein 60 family co-chaperone GroES	COG0234	1.399	5.76E-03	ND	ND	ND	ND
1868	Inner membrane protein translocase component YidC, OxaA protein	COG0706	1.358	4.41E-02	ND	ND	ND	ND
2429	Guanylate kinase (EC 2.7.4.8)	COG0194	1.313	6.01E-03	ND	ND	ND	ND

2072	6-phospho-beta-glucosidase (EC 3.2.1.86)	COG2723	1.296	2.37E-03	ND	ND	ND	ND
422	Chaperone protein DnaK	COG0443	1.273	1.20E-03	ND	ND	ND	ND
TIFN7_00581	Cytidylate kinase (EC 2.7.4.14)	COG0283	1.264	7.65E-04	ND	ND	ND	ND
1757	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	COG0511	1.262	6.15E-03	ND	ND	ND	ND
2130	Phosphate ABC transporter, periplasmic phosphate-binding protein PstS	COG0226	1.219	1.20E-02	ND	ND	ND	ND
2274	Iron-sulfur cluster assembly ATPase protein SufC	COG0396	1.190	3.86E-02	ND	ND	ND	ND
2208	cation transporter		1.190	3.31E-02	ND	ND	ND	ND
328	DedA Protein		1.185	1.37E-02	ND	ND	ND	ND
421	Heat shock protein GrpE	COG0576	1.153	3.51E-03	ND	ND	ND	ND
TIFN7_01389	ESAT-6/Esx family secreted protein EsxA/YukE		1.145	4.92E-03	ND	ND	ND	ND
2653	Large-conductance mechanosensitive channel	COG1970	1.135	4.35E-02	ND	ND	ND	ND
308	Acetoin dehydrogenase E1 component alpha-subunit (EC 1.2.4.-)	COG1071	1.089	4.98E-03	ND	ND	ND	ND
2129	Phosphate ABC transporter, periplasmic phosphate-binding protein PstS	COG0226	1.084	5.88E-03	ND	ND	ND	ND
1755	3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)	COG1028	1.066	7.16E-03	ND	ND	ND	ND
913	Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)	COG0073	0.994	2.53E-02	ND	ND	ND	ND
158	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)	COG1088	0.961	1.45E-02	ND	ND	ND	ND
2105	PTS system, mannose-specific IIC component (EC 2.7.1.69)	COG3715	0.949	9.29E-03	ND	ND	ND	ND
706	Preprotein translocase subunit YajC (TC 3.A.5.1.1)	ncbi-gi;116513085	0.948	5.15E-03	ND	ND	ND	ND
1167	Two-component response regulator SA14-24	COG0745	0.935	1.38E-02	ND	ND	ND	ND
1482	FIG00756147: hypothetical protein	ncbi-gi:116511112	0.903	8.06E-03	ND	ND	ND	ND
76	Ribonuclease P protein component (EC 3.1.26.5)	COG0594	0.849	2.18E-02	ND	ND	ND	ND
1760	Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)	COG0777	0.842	3.69E-02	ND	ND	ND	ND
154	Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)	COG1209	0.822	1.84E-02	ND	ND	ND	ND
989	6-phospho-beta-galactosidase (EC 3.2.1.85)	COG2723	0.809	2.54E-02	ND	ND	ND	ND
156	dTDP-4-dehydrorhamnose 3,5-epimerase (EC 5.1.3.13)	COG1898	0.794	2.16E-02	ND	ND	ND	ND
134	5′-nucleotidase family protein in cluster with NagD-like phosphatase		0.780	4.56E-02	ND	ND	ND	ND
2353	LSU ribosomal protein L20p	COG0292	-0.825	4.00E-02	ND	ND	ND	ND
1749	Ribosome-binding factor A		-0.897	4.18E-02	ND	ND	ND	ND
1087	Ketol-acid reductoisomerase (EC 1.1.1.86)	COG0059	-1.138	1.20E-02	ND	ND	ND	ND
56	FMN-dependent NADH-azoreductase 1 (EC 1.7.-.-)	ncbi-gi:281490577	-1.278	4.30E-02	ND	ND	ND	ND
1024	FIG00755147: hypothetical protein		-2.204	8.76E-03	ND	ND	ND	ND
2897	IS1191, transposase, IS256 family		-2.820	2.17E-02	ND	ND	ND	ND
TIFN7_02719	Transposase	COG2801	-2.906	2.19E-02	ND	ND	ND	ND
1746	Translation initiation factor 2		-3.311	4.31E-02	ND	ND	ND	ND
TIFN7_01059	SSU ribosomal protein S9p (S16e)	COG0103	-3.597	6.40E-03	ND	ND	ND	ND

Genes that were differentially expressed for the first time at 30 minutes after the arginine supplementation								
817	metal-dependent membrane protease	gi116511499	ND	ND	1.282	3.03E-02	2.812	1.14E-02
1605	Aconitate hydratase (EC 4.2.1.3)	COG1048	ND	ND	2.200	3.13E-02	2.139	3.81E-02
1675	Cold shock protein	COG1278	ND	ND	1.361	5.55E-03	1.641	2.29E-03
1692	O-succinylbenzoate-CoA synthase (EC 4.2.1.-)	ncbi-gi:116511533	ND	ND	1.259	3.15E-02	1.593	2.74E-02
2868	Alkylphosphonate utilization operon protein PhnA	ncbi-gi:116513188	ND	ND	1.124	4.02E-02	1.565	1.26E-02
868	cold shock protein D	COG1278	ND	ND	1.146	1.81E-02	1.264	3.34E-03
118	Cold shock protein CspC	COG1278	ND	ND	0.923	1.48E-02	0.939	1.88E-02
623	peptide ABC transporter ATPase	COG1136	ND	ND	-2.718	3.99E-05	1.146	4.03E-02
851	RepB family protein	ncbi-gi:116326564	ND	ND	-1.330	2.92E-03	-1.010	3.92E-02
1929	General stress protein	COG4768	ND	ND	-1.912	1.99E-03	-1.886	1.27E-02
1657	FMN reductase superfamily flavodoxin	ncbi-gi:116511510	ND	ND	2.633	2.19E-02	ND	ND
1996	Transport protein SgaT, putative		ND	ND	2.511	1.27E-04	ND	ND
859	hypothetical protein		ND	ND	2.483	3.39E-02	ND	ND
2535	FIG00755147: hypothetical protein		ND	ND	2.293	1.07E-06	ND	ND
614	rRNA small subunit methyltransferase H	COG0275	ND	ND	1.760	1.86E-04	ND	ND
2349	Oligopeptide transport system permease protein OppC (TC 3.A.1.5.1)	ncbi-gi:116512714	ND	ND	1.221	1.21E-02	ND	ND
867	Cold shock protein CspF	ncbi-gi:116326594	ND	ND	0.983	3.88E-02	ND	ND
81	LSU ribosomal protein L34p	COG0230	ND	ND	0.901	1.01E-02	ND	ND
146	hypothetical protein	ncbi-gi:116511016	ND	ND	0.866	2.04E-02	ND	ND
184	Inosine-5′-monophosphate dehydrogenase (EC 1.1.1.205)	E1.1.1.205	ND	ND	-0.694	3.38E-02	ND	ND
2584	Glutamine amidotransferase, class I	COG2071	ND	ND	-0.851	7.15E-03	ND	ND
882	Transposase and inactivated derivatives-like protein	COG3316	ND	ND	-0.872	4.80E-02	ND	ND
1665	ABC transporter permease/ATP-binding protein	COG1136	ND	ND	-0.907	3.85E-02	ND	ND
498	cation-transporting ATPase	COG0474	ND	ND	-0.932	4.21E-02	ND	ND
2150	integral membrane protein, interacts with FtsH	COG0670	ND	ND	-0.956	2.42E-02	ND	ND
2699	ABC transporter, ATP-binding protein EcsA	COG1131	ND	ND	-0.986	3.89E-02	ND	ND
2120	Microcin C7 self-immunity protein mccF	COG1619	ND	ND	-0.994	3.03E-02	ND	ND
TIFN7_00040	Hypothetical protein		ND	ND	-1.017	6.07E-04	ND	ND
863	ImpB/MucB/SamB family protein	COG0389	ND	ND	-1.022	2.37E-02	ND	ND
2390	proline synthase co-transcribed bacterial homolog PROSC	COG0325	ND	ND	-1.062	7.12E-03	ND	ND
1086	Threonine dehydratase (EC 4.3.1.19)	COG1171	ND	ND	-1.068	2.33E-02	ND	ND
1666	FIG00755159: hypothetical protein		ND	ND	-1.071	3.96E-03	ND	ND
951	Transposase and inactivated derivatives-like protein	COG3316	ND	ND	-1.075	1.80E-02	ND	ND
2187	ATP synthase gamma chain (EC 3.6.3.14)	COG0224	ND	ND	-1.076	1.23E-02	ND	ND
2435	stress-responsive transcriptional regulator, PspC family	ncbi-gi:281492518	ND	ND	-1.085	2.46E-03	ND	ND

3002	Transposase and inactivated derivatives-like protein	COG3316	ND	ND	-1.085	1.64E-02	ND	ND
241	Serine/threonine protein kinase PrkC, regulator of stationary phase		ND	ND	-1.119	1.25E-02	ND	ND
144	CAAX amino terminal protease family protein	ncbi-gi:281490674	ND	ND	-1.138	1.99E-02	ND	ND
1570	Putative stomatin/prohibitin-family membrane protease subunit YbbK	COG0330	ND	ND	-1.240	4.20E-02	ND	ND
1709	Aspartokinase (EC 2.7.2.4)	COG0527	ND	ND	-1.265	3.89E-03	ND	ND
37	hypothetical protein	ncbi-gi:116510917	ND	ND	-1.351	3.15E-02	ND	ND
708	LSU m3Psi1915 methyltransferase RlmH ## ybeA	COG1576	ND	ND	-1.353	3.87E-02	ND	ND
2710	Peptide methionine sulfoxide reductase MsrA (EC 1.8.4.11)	COG0225	ND	ND	-1.359	2.24E-02	ND	ND
TIFN7_01007	Hypothetical protein		ND	ND	-1.477	2.97E-02	ND	ND
2160	5-Enolpyruvylshikimate-3-phosphate synthase (EC 2.5.1.19)	COG0128	ND	ND	-1.599	4.21E-02	ND	ND
3144	Spermidine Putrescine ABC transporter permease component PotB	COG1176	ND	ND	-1.602	2.83E-02	ND	ND
2740	similarity to aminoacyl-tRNA editing enzymes YbaK, ProX	COG3760	ND	ND	-1.721	2.65E-02	ND	ND
1857	NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	COG0057	ND	ND	-1.915	1.75E-03	ND	ND
2076	Two component system histidine kinase (EC 2.7.3.-)		ND	ND	-1.968	4.60E-02	ND	ND
622	ABC transporter permease protein		ND	ND	-2.092	5.40E-04	ND	ND
1519	ABC transporter, ATP-binding/permease protein	COG1132	ND	ND	-2.774	1.03E-03	ND	ND
1491	MerR family transcriptional regulator	ncbi-gi:125623173	ND	ND	-2.898	6.64E-04	ND	ND
Genes that were differentially expressed for the first time at 2 hours after the arginine supplementation								
3165	Xanthine phosphoribosyltransferase (EC 2.4.2.22)	COG0503	ND	ND	ND	ND	4.411	8.89E-07
3164	Xanthine permease	COG2233	ND	ND	ND	ND	4.007	2.26E-05
1715	Replicative DNA helicase (EC 3.6.1.-) [SA14-24]	COG0305	ND	ND	ND	ND	3.124	1.06E-02
1552	Transcriptional regulator, MarR family / nitroreductase family protein	ncbi-gi:125623862	ND	ND	ND	ND	2.721	1.23E-02
838	hypothetical protein	ncbi-gi:15672064	ND	ND	ND	ND	2.460	2.17E-03
1504	Phosphonate ABC transporter permease protein phnE2 (TC 3.A.1.9.1)	COG3639	ND	ND	ND	ND	2.425	3.29E-02
459	Topoisomerase IV subunit A (EC 5.99.1.-)	COG0188	ND	ND	ND	ND	2.359	7.78E-03
793	transport and binding protein	ncbi-gi:15672068	ND	ND	ND	ND	2.009	4.03E-03
TIFN7_01126	Integrase		ND	ND	ND	ND	1.960	4.86E-02
TIFN7_02572	Hypothetical protein		ND	ND	ND	ND	1.904	1.26E-02
3135	Putative NAD(P)H nitroreductase(EC:1.-)	COG0778	ND	ND	ND	ND	1.688	4.87E-02
711	Glycosyltransferase LafA, responsible for the formation of Glc-DAG	COG0438	ND	ND	ND	ND	1.670	6.46E-03
1451	transporter	ncbi-gi:116513196	ND	ND	ND	ND	1.628	3.29E-02
1452	metal-dependent membrane protease	ncbi-gi:116513195	ND	ND	ND	ND	1.540	2.19E-04
792	hypothetical protein		ND	ND	ND	ND	1.305	4.11E-03
143	1,4-dihydroxy-2-naphthoate octaprenyltransferase (EC 2.5.1.-)	ncbi-gi:116511015	ND	ND	ND	ND	1.300	4.49E-02
262	oxidoreductase of aldo/keto reductase family, subgroup 1	COG0656	ND	ND	ND	ND	1.271	3.24E-02
1188	FMN-dependent NADH-azoreductase		ND	ND	ND	ND	1.224	4.81E-02
444	Peptide chain release factor 2		ND	ND	ND	ND	1.050	1.07E-02

368	Malolactic enzyme (EC 1.-.-.-)	COG0281	ND	ND	ND	ND	1.039	4.05E-02
2301	FIG00756323: hypothetical protein		ND	ND	ND	ND	1.030	1.67E-02
2043	Proline dipeptidase (EC 3.4.13.9)	E3.4.13.9	ND	ND	ND	ND	1.022	3.46E-02
2600	Pili retraction protein pilT		ND	ND	ND	ND	0.995	1.80E-02
1867	Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)	COG0623	ND	ND	ND	ND	0.988	1.62E-02
330	Shikimate/quinate 5-dehydrogenase I beta (EC 1.1.1.282)	COG0169	ND	ND	ND	ND	0.964	2.68E-02
1486	Transcription accessory protein (S1 RNA-binding domain)	COG2183	ND	ND	ND	ND	0.938	3.73E-02
1239	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein OppA	COG4166	ND	ND	ND	ND	0.914	2.53E-02
2795	hypothetical protein	ncbi-gi:116326585	ND	ND	ND	ND	0.895	4.25E-02
1512	Lysyl aminopeptidase (EC 3.4.11.15)	COG0308	ND	ND	ND	ND	0.890	3.97E-02
1727	GTP-binding protein EngA	COG1160	ND	ND	ND	ND	0.880	2.58E-02
136	Ribosomal RNA large subunit methyltransferase N (EC 2.1.1.-)	COG0820	ND	ND	ND	ND	0.865	3.21E-02
1695	Naphthoate synthase (EC 4.1.3.36)	COG0447	ND	ND	ND	ND	0.865	3.79E-02
1014	Oligopeptide transport ATP-binding protein OppD (TC 3.A.1.5.1)	COG0444	ND	ND	ND	ND	0.847	1.61E-02
1613	hypothetical protein	ncbi-gi:125623501	ND	ND	ND	ND	0.801	2.37E-02
78	Inner membrane protein translocase component YidC, short form OxaI-like	COG0706	ND	ND	ND	ND	0.749	3.82E-02
1587	Ribosomal subunit interface protein	COG1544	ND	ND	ND	ND	-0.840	1.25E-02
688	Phage antirepressor protein	ncbi-gi:116511842	ND	ND	ND	ND	-1.289	3.58E-03
2262	orf10	ncbi-gi:30023991	ND	ND	ND	ND	-1.297	1.30E-02
685	prophage pi2 protein 09	ncbi-gi:116512480	ND	ND	ND	ND	-1.416	1.81E-02
666	hypothetical protein		ND	ND	ND	ND	-1.460	6.67E-04
671	Deoxyuridine 5′-triphosphate nucleotidohydrolase (EC 3.6.1.23)	COG0756	ND	ND	ND	ND	-1.469	6.35E-03
2248	unknown		ND	ND	ND	ND	-1.504	1.06E-02
2259	hypothetical protein	ncbi-gi:125623654	ND	ND	ND	ND	-1.523	1.25E-02
670	Deoxyuridine 5′-triphosphate nucleotidohydrolase (EC 3.6.1.23)	COG0756	ND	ND	ND	ND	-1.552	1.45E-02
2250	hypothetical protein	ncbi-gi:116512491	ND	ND	ND	ND	-1.563	6.05E-03
TIFN7_00363	FIG00755207: hypothetical protein		ND	ND	ND	ND	-1.611	5.86E-03
2245	Deoxyuridine 5′-triphosphate nucleotidohydrolase (EC 3.6.1.23)	COG0756	ND	ND	ND	ND	-1.632	5.82E-03
2247	hypothetical protein	ncbi-gi:116512801	ND	ND	ND	ND	-1.663	6.16E-03
1811	Leucyl-tRNA synthetase (EC 6.1.1.4)	COG0495	ND	ND	ND	ND	-3.010	3.69E-03
2326	hypothetical protein	ncbi-gi:116512330	ND	ND	ND	ND	-3.658	4.96E-02
2159	Shikimate kinase I (EC 2.7.1.71)	E2.7.1.71	ND	ND	ND	ND	-3.816	4.93E-02
1463	Ribonucleotide reductase of class III (anaerobic), activating protein	E1.97.1.4	ND	ND	ND	ND	-4.146	1.29E-02
696	Late competence protein ComGC, access of DNA to ComEA, FIG007487	ncbi-gi:116513076	ND	ND	ND	ND	-4.889	1.66E-02
2358	hypothetical protein	ncbi-gi:125624806	ND	ND	ND	ND	-8.551	4.29E-04

Supplementary Table 4. The absolute expression values of the genes that were differentially regulated 15 minutes after the arginine supplementation in *L. lactis* TIFN7. Genes that are differentially upregulated more than two folds are highlighted with dark grey and the genes that were differentially downregulated are highlighted with light grey. ND and is used as abbreviation for ``not detected`` and the absolute expression values are given as 10^6 times.

Differential Exp. (log2)	Annotation	Gene Names	COG/EC Number	TIFN7 Steady State	TIFN7 15 minutes	TIFN7 30 minutes	TIFN7 2 hours	TIFN5 Steady State
3.83	OrfA; hypothetical protein; putative DNA recombinase			2.0	95.4	92.3	1.5	10.9
3.83	Abortive phage resistance protein abiGi			0.0	0.7	0.9	1.2	80.2
3.82	Dihydroorotase (EC 3.5.2.3)	pyrC	COG0044	15.2	337.2	51.1	49.4	382.6
3.59	Dihydroorotate dehydrogenase, catalytic subunit (EC 1.3.3.1)	pyrD	COG0167	7.2	103.4	21.1	16.0	119.6
3.58	Phosphoribosylformylglycinamidine synthase, PurS subunit	purS	E6.3.5.3	21.2	301.3	169.8	944.8	637.9
3.56	Dihydroorotate dehydrogenase electron transfer subunit	pyrD	COG0543	7.0	98.5	22.9	15.1	262.4
3.49	Orotate phosphoribosyltransferase (EC 2.4.2.10)	pyrE	COG0461	23.8	286.5	85.0	56.7	421.2
3.48	Phosphoribosylformylglycinamidine synthase	purL	COG0047	14.9	178.5	128.0	607.1	443.7
3.47	GMP reductase (EC 1.7.1.7)	guaB	COG0516	27.2	320.1	311.4	147.9	232.2
3.44	Nitrogen regulatory protein P-II		COG0347	5.4	60.9	27.4	78.3	39.3
3.36	Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23)	pyrF	COG0284	26.4	294.6	62.7	71.9	103.8
3.33	Phosphoribosylaminoimidazole-succinocarboxamide synthase	purC	COG0152	6.0	66.0	32.8	161.9	510.3
3.31	Phosphoribosylformylglycinamidine synthase, synthetase subunit	purL	COG0046	14.3	155.8	170.8	949.3	419.3
3.18	Lysine-specific permease		COG0833	225.4	43.4	29.3	34.1	94.3
3.16	IMP cyclohydrolase (EC 3.5.4.10)	purH	COG0138	5.6	56.0	26.0	65.5	527.8
3.10	Phosphoribosylformylglycinamidine cyclo-ligase (EC 6.3.3.1)	purM	COG0150	17.4	172.0	85.4	206.1	391.4
3.06	PTS system, cellobiose-specific IIA component (EC 2.7.1.69)		COG1447	17.0	154.2	47.7	27.8	142.2
3.01	Carbamoyl-phosphate synthase small chain (EC 6.3.5.5)	carA	COG0505	104.2	928.3	262.4	189.9	437.2
2.86	transposase			1.2	9.7	11.8	0.4	ND
2.83	Xanthine/uracil/thiamine/ascorbate permease family protein		COG2252	89.4	722.7	431.5	722.3	444.8
2.82	Uracil permease		COG2233	54.8	414.9	132.9	85.0	615.6
2.80	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)	carB	COG0458	215.1	1562.3	534.3	477.1	ND
2.74	regulator of polyketide synthase expression		ncbi-gi:116511630	99.1	120.7	81.8	93.7	244.8
2.70	Aspartate carbamoyltransferase (EC 2.1.3.2)	pyrB	COG0540	106.6	750.8	243.4	167.1	461.1
2.63	Amidophosphoribosyltransferase (EC 2.4.2.14)	purF	COG0034	22.0	152.7	76.8	416.0	584.1
2.63	Uracil phosphoribosyltransferase (EC 2.4.2.9) / PyrR	pyrR	COG2065	127.7	852.1	280.5	196.1	642.0
2.48	Adenylosuccinate synthetase (EC 6.3.4.4)	purA	COG0104	181.3	1180.0	597.7	649.4	451.0
2.39	Phosphonate ABC transporter phosphate			4.4	27.6	16.3	40.6	596.1
2.32	Phosphonate ABC transporter ATP-binding protein (TC 3.A.1.9.1)		COG3638	2.1	13.0	6.6	21.1	133.3

2.27	Nucleoside-binding protein		COG1744	237.2	1324.2	677.3	691.9	635.8
2.27	Serine hydroxymethyltransferase (EC 2.1.2.1)	glyA	COG0112	8.0	42.1	39.7	59.2	ND
2.23	CTP synthase (EC 6.3.4.2)	pyrG	COG0504	181.8	907.8	472.7	465.3	338.8
2.18	Formate--tetrahydrofolate ligase (EC 6.3.4.3)	fhs	COG2759	41.1	198.2	139.2	247.5	129.4
2.15	Multiple sugar ABC transporter, ATP-binding protein		COG3839	3.1	15.0	4.6	5.0	70.5
2.14	cation-transporting ATPase, E1-E2 family		COG0474	1.5	2.3	2.8	2.5	59.1
2.14	Glutamine synthetase type I (EC 6.3.1.2)	glnA	COG0174	332.1	1504.7	781.3	1658.0	1151.8
2.13	Ribosomal RNA small subunit methyltransferase C (EC 2.1.1.52)		COG2813	17.8	78.9	35.3	55.8	106.4
2.11	Lipoate-protein ligase A	lplA	COG0095	39.5	174.2	106.4	22.7	65.8
2.07	Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)	folD	COG0190	27.1	119.3	48.8	74.4	173.9
2.03	Poly(glycerophosphate chain) D-alanine transfer protein DltD			4.8	21.1	10.2	24.1	281.8
2.03	Mevalonate kinase (EC 2.7.1.36)		COG1577	20.5	88.3	55.1	78.9	192.4
2.03	Phosphoribosylamine--glycine ligase (EC 6.3.4.13)	purD	COG0151	3.2	9.4	5.5	6.6	332.5
1.98	YibE/F superfamily transporter		UNIPROT:F2HJA2	7.2	29.5	10.4	18.3	43.1
1.89	IMP cyclohydrolase (EC 3.5.4.10)		COG0138	5.9	24.1	14.4	56.8	249.5
1.89	Maltose/maltodextrin ABC transporter, MalE		COG2182	14.7	60.0	24.9	24.0	ND
1.86	PTS system, trehalose-specific IIA component (EC 2.7.1.69)		COG2190	8.1	32.0	14.9	10.4	22.1
1.86	PTS system, cellobiose-specific IIB component (EC 2.7.1.69)		COG1440	71.0	264.1	98.4	57.1	143.9
1.78	PTS system, trehalose-specific IIB component (EC 2.7.1.69)		COG1264	8.3	30.9	17.2	11.0	17.1
1.77	N-acetylmuramoyl-L-alanine amidase (involved in cell division)		E3.5.1.28	13.9	50.4	32.4	25.0	422.8
1.75	glutamine synthetase repressor		COG0789	963.6	3490.4	2659.4	3501.6	1434.5
1.73	Glycyl-tRNA synthetase alpha chain (EC 6.1.1.14)		COG0752	7.8	26.7	24.3	20.0	80.4
1.71	Pyrimidine-nucleoside phosphorylase (EC 2.4.2.2)	pdp	COG0213	51.6	175.9	169.5	141.8	ND
1.70	Methionine ABC transporter substrate-binding protein		COG1464	117.2	393.8	482.4	330.0	64.9
1.70	Hypoxanthine-guanine phosphoribosyltransferase (EC 2.4.2.8)	hpt	COG0634	8.2	27.4	14.2	15.3	125.7
1.70	Maltose/maltodextrin ABC transporter, MalE		COG2182	13.1	42.9	22.7	23.9	ND
1.69	Glycerol uptake facilitator protein		COG0580	60.6	197.2	123.6	121.3	126.1
1.68	Heptaprenyl diphosphate synthase component II (EC 2.5.1.30)		E2.5.1.30	86.4	280.9	247.0	273.6	0.0
1.66	Substrate-specific component NiaX of niacin ECF transporter			156.4	507.1	315.8	407.9	340.5
1.64	hypothetical protein (Fe-S Oxidoreductase)		ncbi-gi:116513115	33.0	106.1	86.1	151.2	184.8
1.64	Oligopeptide ABC transporter, OppA		COG4166	22.1	70.5	44.1	40.7	ND
1.60	Adenylosuccinate lyase (EC 4.3.2.2)	purB	COG0015	42.0	133.1	105.0	128.3	522.9
1.60	Transcriptional regulatory protein		COG0718	93.3	290.3	243.9	239.8	73.5
1.60	NLP/P60 family protein		COG0791	15.1	46.9	28.2	22.7	267.2
1.59	FIG00756303: hypothetical protein		ncbi-gi:116511972	135.0	409.6	85.4	155.9	394.5
1.59	Methionine ABC transporter substrate-binding protein		COG1464	84.8	256.7	192.9	181.8	69.3
1.51	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	fabH	COG0331	37.9	114.4	81.0	101.7	441.8

1.50	Phosphoenolpyruvate-dihydroxyacetone phosphotransferase	dhaK	COG2376	36.1	108.5	60.6	69.9	175.0
1.46	Mannose-6-phosphate isomerase (EC 5.3.1.8)	manA	COG1482	43.3	130.1	85.8	73.1	178.4
1.44	Probable transport protein		COG0531	25.0	71.3	36.5	40.3	38.9
1.43	Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)		COG0016	111.9	316.8	269.8	161.3	145.5
1.41	Hypothetical protein			133.5	366.0	116.5	194.5	ND
1.40	Preprotein translocase subunit SecG (TC 3.A.5.1.1)		UNIPPROT:Q02ZS0	58.9	159.6	133.5	141.9	102.6
1.39	tRNA dihydrouridine synthase B (EC 1.-.-.-)		COG0042	26.9	72.5	70.9	69.6	79.7
1.36	UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.1.1.158)		COG0812	35.6	94.9	68.0	57.6	600.5
1.36	Heat shock protein 60 family co-chaperone GroES		COG0234	68.2	179.9	91.8	76.8	180.1
1.33	Heptaprenyl diphosphate synthase component II (EC 2.5.1.30)		E2.5.1.30	97.8	255.8	256.2	233.8	190.2
1.31	3-oxoacyl-[acyl-carrier-protein] synthase, KASII (EC 2.3.1.41)	fabB	COG0304	445.1	1145.4	971.3	836.8	410.9
1.30	Inner membrane protein translocase component YidC, OxaA		COG0706	45.0	115.4	65.0	87.9	216.4
1.27	Cell division trigger factor (EC 5.2.1.8)		COG0544	463.7	1168.4	834.8	1002.3	714.7
1.27	Guanylate kinase (EC 2.7.4.8)	gmk	COG0194	102.3	254.1	173.9	156.7	109.5
1.26	6-phospho-beta-glucosidase (EC 3.2.1.86)		COG2723	109.3	268.4	168.2	104.3	ND
1.26	Chaperone protein DnaK		COG0443	559.4	1352.1	607.5	494.7	ND
1.26	DNA-directed RNA polymerase omega subunit (EC 2.7.7.6)	rpa	E2.7.7.6	300.6	724.9	537.3	449.5	126.1
1.24	Cytidylate kinase (EC 2.7.4.14)	cmk	COG0283	1165.3	2798.9	1503.6	1235.5	0.0
1.22	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	pyc	COG0511	79.0	189.4	124.5	130.9	384.3
1.19	hypothetical phage protein			215.2	515.8	471.8	491.6	ND
1.19	UDP-N-acetylmuramate--alanine ligase (EC 6.3.2.8)		COG0773	180.4	426.7	351.6	368.3	469.9
1.18	Phosphate ABC transporter, PstS		COG0226	64.5	150.2	96.9	100.9	68.8
1.18	Iron-sulfur cluster assembly ATPase protein SufC		COG0396	69.5	158.5	119.0	104.4	182.8
1.18	cation transporter			63.6	145.0	104.7	117.5	ND
1.17	DedA Protein			89.7	203.9	136.3	151.1	ND
1.16	alpha-phosphoglucomutase	pgm	E5.4.2.2	201.9	458.1	335.6	338.3	92.4
1.15	Flavodoxin		COG0716	49.7	112.6	121.1	142.2	75.4
1.15	Methionine ABC transporter substrate-binding protein		COG1464	61.7	139.1	171.7	132.7	54.4
1.14	hyphotetical protein		ncbi-gi:116513040	52.9	118.5	110.9	127.9	25.1
1.14	Heat shock protein GrpE		COG0576	201.5	448.2	174.8	138.7	512.3
1.13	Transcriptional regulator of fatty acid biosynthesis FabT		ncbi-gi:116511579	84.2	187.2	186.3	194.6	207.8
1.13	ESAT-6/Esx family secreted protein EsxA/YukE			117.4	259.7	200.6	143.8	ND
1.12	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase	fabG	COG0764	456.2	1008.5	773.3	861.4	248.5
1.12	Large-conductance mechanosensitive channel		COG1970	83.5	183.4	105.6	114.2	384.5
1.10	FMN-dependent NADH-azoreductase		COG1182	189.1	413.3	353.0	489.9	172.9
1.09	Hypothetical protein			258.0	560.7	499.7	326.3	ND
1.09	Cysteine synthase (EC 2.5.1.47)	cbs	COG0031	308.3	669.9	592.4	347.9	30.0

1.09	Threonyl-tRNA synthetase (EC 6.1.1.3)		COG0441	78.2	167.8	166.4	137.8	135.4
1.09	probably aromatic ring hydroxylating enzyme		COG2151	84.5	180.4	181.2	217.4	67.0
1.08	Tripeptide aminopeptidase (EC 3.4.11.4)		COG2195	733.6	1564.0	1537.8	1624.5	613.7
1.07	Acetoin dehydrogenase E1 component alpha-subunit (EC 1.2.4.-)	pdhA	COG1071	243.2	517.4	330.2	243.1	121.6
1.07	Hypothetical protein			212.1	451.0	533.9	610.1	ND
1.04	Phosphate ABC transporter, PstS		COG0226	161.7	342.7	235.0	220.7	135.8
0.99	hypothetical phage protein			286.3	599.6	529.6	624.7	ND
0.99	3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)	fabG	COG1028	171.9	359.9	228.5	231.0	531.7
0.99	Acyl carrier protein		COG0236	963.8	1982.9	2256.5	2421.4	645.1
0.99	Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)		COG0073	110.5	220.0	179.4	159.7	198.8
0.96	cationic amino acid transporter		COG0531	291.6	580.5	572.8	807.7	103.8
0.96	Transketolase (EC 2.2.1.1)	tktA	COG0021	138.8	275.1	215.2	254.5	149.1
0.95	6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)	gnd	COG0362	524.7	1039.0	769.6	872.9	527.4
0.95	Phosphate transport system permease protein PstA (TC 3.A.1.7.1)		COG0581	179.1	349.2	297.5	307.1	260.3
0.94	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)	rfbB	COG1088	370.9	722.0	519.8	552.8	351.9
0.90	PTS system, mannose-specific IIC component (EC 2.7.1.69)	PTS	COG3715	1028.1	1985.5	1140.5	1151.0	847.1
0.89	Preprotein translocase subunit YajC (TC 3.A.5.1.1)		ncbi-gi;116513085	376.6	726.5	605.2	464.5	142.8
0.88	Two-component response regulator SA14-24		COG0745	496.5	949.5	588.9	750.0	769.4
0.88	FIG00756147: hypothetical protein		ncbi-gi:116511112	482.2	902.0	666.6	730.0	ND
0.86	hypothetical protein		ncbi-gi:116511061	305.6	567.0	524.1	522.8	ND
0.85	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase		COG0764	192.2	354.3	379.2	503.4	ND
0.84	LSU ribosomal protein L28p		COG0227	866.0	1589.0	1745.2	1438.4	87.0
0.82	Inner membrane protein translocase component YidC		COG0706	328.7	597.6	583.4	645.0	ND
0.81	Ribonuclease P protein component (EC 3.1.26.5)		COG0594	419.6	755.8	572.5	642.8	500.3
0.79	Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)	acc	COG0777	150.8	270.4	194.9	233.1	315.7
0.78	Glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24)	rfbA	COG1209	554.9	981.1	757.6	811.3	326.5
0.77	6-phospho-beta-galactosidase (EC 3.2.1.85)		COG2723	862.3	1511.1	575.5	661.1	ND
0.00	Acetylornithine deacetylase (EC 3.5.1.16)		EC 3.5.1.16	1703.8	119.3	15.9	3.6	1145.8
0.00	glutamate or arginine ABC transporter substrate binding protein		ncbi-gi:116511362	3324.4	187.5	19.0	8.5	645.2
0.00	Amino acid ABC transporter, ATP-binding protein		COG1126	190.1	3.8	3.3	1.5	79.9
-0.83	dTDP-4-dehydrorhamnose 3,5-epimerase (EC 5.1.3.13)	rfbC	COG1898	716.8	1242.7	1008.0	1055.8	97.8
-0.85	5′-nucleotidase family protein with NagD-like phosphatase	ushA	COG0737	254.6	437.3	317.3	352.6	ND
-0.86	Glutamine ABC transporter, glutamine-binding protein		COG0765	274.9	469.4	316.4	516.4	362.7
-0.89	LSU ribosomal protein L20p		COG0292	2174.9	1227.5	1738.6	1816.8	501.8
-0.90	FIG00756687: hypothetical protein			652.3	361.0	611.6	1171.8	96.8
-0.90	FIG00755954: hypothetical protein			324.7	179.0	311.0	622.8	62.6
-0.97	Cell division initiation protein DivIVA		ncbi-gi:116512742	349.0	188.9	165.2	232.0	688.7

-0.97	Ribosome-binding factor A			215.8	115.9	187.4	194.9	ND
-0.98	Universal stress protein family		ncbi-gi:116326637	1019.6	545.9	488.2	727.5	175.8
-1.02	Phosphoglycerate mutase family 2		ncbi-gi:116512567	456.6	233.7	147.6	264.3	29.7
-1.03	Manganese transport protein MntH		COG1914	820.4	417.8	375.9	507.3	220.2
-1.14	Cysteine ABC transporter, substrate-binding protein		ncbi-gi:116513001	997.0	505.6	576.0	460.3	231.7
-1.21	Gls24 family general stress protein		ncbi-gi:281491279	908.6	447.8	577.4	612.0	267.9
-1.26	transglycosylase		ncbi-gi:116512524	1715.5	840.0	946.3	1207.0	942.7
-1.28	Ketol-acid reductoisomerase (EC 1.1.1.86)		COG0059	134.4	61.1	86.5	87.3	50.1
-1.29	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I	aroF	COG0722	760.6	329.7	373.3	510.3	376.2
-1.31	Arginyl-tRNA synthetase (EC 6.1.1.19)		COG0018	242.9	101.8	65.4	74.5	792.1
-1.47	FMN-dependent NADH-azoreductase 1 (EC 1.7.-.-)		ncbi-gi:281490577	48.8	20.1	21.8	30.7	42.6
-1.52	Dihydropteridine reductase (EC 1.5.1.34)		E1.5.1.34	923.9	378.1	310.7	422.7	438.6
-1.66	FIG00755240: hypothetical protein		COG3382	315.8	127.0	119.0	283.4	196.3
-1.72	Lactose phosphotransferase system repressor		ncbi-gi:116326616	507.4	182.7	143.8	218.9	171.8
-1.91	Arginine pathway regulatory protein ArgR		COG1438	277.2	96.3	69.5	62.7	285.7
-2.07	FIG00755211: hypothetical protein			40.9	12.9	13.7	24.1	ND
-2.20	Arginine/ornithine antiporter ArcD		COG0531	43.8	13.3	8.9	10.1	42.0
-2.23	ClpB protein		COG0542	622.6	165.9	153.3	201.3	660.0
-2.24	glutamate ABC transporter permease protein		COG0765	23.1	5.5	3.1	1.8	60.6
-2.38	FIG00755147: hypothetical protein			1.3	0.0	1.5	3.3	ND
-2.47	Argininosuccinate lyase (EC 4.3.2.1)	argH	COG0165	274.1	58.3	25.8	25.7	ND
-2.77	Argininosuccinate synthase (EC 6.3.4.5)	argG	COG0137	126.8	26.9	18.9	17.8	892.4
-2.82	Lysine-specific permease		COG0833	15.6	160.1	110.9	64.5	27.0
-2.88	N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38)	argC	COG0002	357.8	64.7	17.2	18.7	191.8
-2.91	L-proline glycine betaine ABC transport permease protein ProV		COG1125	36.6	5.3	6.1	17.1	103.1
-3.31	IS1191, transposase, IS256 family			2.0	0.0	1.2	3.7	ND
-3.55	Amino acid ABC transporter, ATP-binding protein		COG1126	60.9	8.3	5.2	7.8	ND
-3.60	Transposase		COG2801	2.2	0.0	1.1	1.4	ND
-3.60	Translation initiation factor 2			2.9	0.0	3.3	0.7	ND
-3.70	Amino acid ABC transporter, amino acid-binding/permease protein		ncbi-gi:116511362	667.0	57.1	5.7	7.6	645.2
-3.84	SSU ribosomal protein S9p (S16e)		COG0103	3.5	0.0	1.5	1.2	ND
-3.85	Glutamate decarboxylase (EC 4.1.1.15)	gadB	COG0076	46.8	3.9	13.9	14.1	20.0
-4.15	transposase			3.1	0.2	0.5	0.8	ND
-5.64	Acetylornithine deacetylase (EC 3.5.1.16)		EC 3.5.1.16	1097.5	76.2	2.8	3.1	737.1

Chapter 6

Starter Culture Composition and Its Flavor Formation Capacity as a Function of Propagation Regimes

Oylum Erkuş, R.T.C.M. Geene, Patrick W.M. Janssen, Michiel Kleerebezem & Eddy J. Smid

Manuscript in preparation

Abstract

Dairy industry employs both defined and undefined starter cultures for cheese manufacturing. Defined blends of strains with known phenotypic features provide a certain degree of control on the starter culture functionality but the relatively simple composition makes them vulnerable for phage attack. On the other hand, undefined starters function very robustly and are less vulnerable to phage attacks or other environmental uncertainties due to their complex nature. In contrast to defined cultures, modulating the community composition to alter certain functionality could be relatively difficult in mixed undefined starters due to their complexity. In this study, different propagation regimes were used to modulate the microbial community composition of an undefined starter culture with the objective to change the starter culture functionality in Gouda cheese production. The undefined starter culture was serially propagated from different growth phases in milk. The final community compositions and functionalities, such as acidification potential and proteolytic activity of serially propagated cultures were found to depend on the back-slopping regime employed. For example, cultures propagated at the stationary growth phase shifted towards a higher relative abundance of proteolytic strains as compared to the cultures propagated at the exponential growth phase. The flavour formation in 12-weeks ripened mini-cheeses that were manufactured with serially propagated starter cultures were specific for the propagation regime used. These results imply that alternative back-slopping can be employed to steer the microbial community composition and thereby modulate starter culture functionality.

Introduction

Dairy starter cultures are lactic acid bacteria cultures that are used to initiate or accelerate the fermentation of milk products (1). In Gouda cheese manufacturing, the starter cultures harbour different strains of *Lactococcus lactis* and *Leuconostoc mesenteroides* species, which acidify the milk by lactic acid production and thereby induce, in concert with added chymosin, the formation of curd. Certain activities associated with the lactic acid bacteria in the starter culture (i.e., proteolysis, lipolysis, citrate utilization, amino acid degradation, and exopolysaccharide production) contribute to the flavour and texture development in cheese (2, 3). Starter cultures may constitute single or multiple strains from several species and in general are classified into two categories: defined and undefined mixed cultures (4).

Defined starter cultures contain a known number of strains (single or multiple) usually with known functional traits. For example, *L. lactis* strains with a high level of peptidase activity and high autolytic abilities were shown to increase the release of flavour precursors (amino acids) in the cheese matrix (5), and they are used in starter culture formulations for rich flavour development. Tailor-made defined starter cultures can also be composed from multiple strains that complement each other for particular metabolic activity, for example, to stimulate the production of specific aroma compounds in the product. To illustrate this, Ayad et al. demonstrated that the formation of 3-methyl-butanal in cheese is limited in single cultures of *L. lactis* B1157 and *L. lactis* SK110, whereas 3-methyl-butanal formation was strongly enhanced when they were grown together (6). Detailed enzymatic and genomic analysis showed that *L. lactis* SK110 was able to release branched chain amino acids from casein but was limited in decarboxylase activity. The non-proteolytic *L. lactis* B1157 was found to complement *L. lactis* SK110 by decarboxylation of the released leucine into 3-methyl butanal (6). This example not only illustrates microbial cooperation but also supports the value of genomic and biochemical knowledge on the rational design of defined starter culture combinations with predictable functionalities.

Undefined or artisanal starter cultures contain unknown number of strains and species. Back-slopping, which is the inoculation of a new batch of unfermented food material by transfer from the previous batch (7), is the traditional way of propagating undefined starters. These spontaneously developed mixed-strain cultures were propagated without any protection against contamination with phages (also known as Practice cultures) (8), and thereby the microbial diversity present in them can be regarded as the outcome of a long-term adaptive evolution process.

Undefined mixed starter cultures are still commonly used at industrial scale, for instance for the production of Gouda cheese. The microbial community composition of one of these undefined mixed Gouda cheese starter (named Ur) has been characterized in Chapter 3 of this thesis, and was shown to encompass a total 7 genetic lineages of *Lactococcus lactis* and a single genetic lineage of *Leuconostocs mesenteroides*. Genetic lineages were shown to accommodate strains with variable phage sensitivities and harbouring a certain degree of plasmid content variability, which generates a level of diversity that significantly exceeds that of the discriminated genetic lineages. This sub-lineage diversity generates a high degree of stability for each genetic lineage and eventually the overall starter culture, which makes the undefined culture highly robust under the environmental challenges imposed by variable physicochemical conditions and/or phage predation (This thesis, Chapter 3). Engineering mixed undefined starter cultures to stimulate specific functionality on demand can be considered more difficult due to microbial complexity and interactions present within these microbial communities compared to the defined cultures encompassing single or limited number or strains with known functionalities. However, (adaptation of) back-slopping regimes can potentially provide a credible strategy to steer the composition of the starter cultures into a certain direction and to modulate final end-product attributes such as flavour formation. For example, in a recent study, *S. thermophilus* and *L. bulgaricus* were shown to adapt to each other in sequential yoghurt fermentations (back-slopping). The evolved co-culture was shown to produce higher quantities of yoghurt specific flavour compounds as compared to the initial, unadapted co-cultures (9). In a similar approach, modulation of growth conditions and propagation regimes in back-slopping of undefined mixed-strain cultures may force the system to adapt to these alternative propagation conditions, which may result in the emergence of an alternative community composition.

In order to test this hypothesis, we studied the effect of propagation regime modulations on the undefined mixed-strain Gouda cheese starter culture UR (8) that has long been maintained using back-slopping propagations. The transfer time point from old batch of culture to a subsequent subculture was varied based on the growth status of the starter culture in the propagation regime. Starter culture UR was serially propagated in three parallel lines from exponential and stationary phases of growth, and the final community compositions and functionalities (acidification and proteolysis) of the newly adapted cultures were determined. Finally, the original and the adapted starter cultures were applied in cheese manufacturing, and the variations in flavour profiles of 12-weeks ripened cheeses were correlated to the community composition changes.

Materials and Methods

Starter Culture and Propagation Regimes

The starter culture UR used in cheese manufacturing is a mixed-strain undefined Gouda type cheese starter culture (8). The starter culture consists of 7 genetic lineages of *Lactococcus lactis* and 1 genetic lineage of *Leuconostoc mesenteroides* (This thesis, Chapter 3). Representative strains for each genetic lineage (designated as TIFN1 to 8) have previously been subjected to whole genome sequencing, comprising *5 L. lactis* ssp. *cremoris* strains (TIFN1, 3, 5, 6, and 7), 2 *L. lactis* ssp. *lactis* biovar. *diacetylactis* strains (TIFN 2 and 4), and 1 *Ln. mesenteroides* ssp. *cremoris* strain (TIFN8) (This thesis, Chapter 3). The stocks of the starter culture were stored at -80°C.

During back-slopping experiments, the transfer time point of the culture from old batch to new one in the propagation regime was varied based on the growth curve of the starter culture. Previous analysis of the growth curve of the original starter culture in 10% skimmed milk (1% (w/v) inoculation) at 20°C indicated that the end of exponential phase was observed approximately 20 hour after inoculation (data not shown). Therefore, the propagation regimes employed 12 hours and 48 hours for the propagations from logarithmic phase of growth and stationary phase of growth, respectively. In total 7 replicates of the original starter culture were included in the experimental design. Three of the starter cultures were sub-cultured every 12 hours from the exponential phase for 500-700 generations (cultures ExpA, ExpB, ExpC). Three of the remaining cultures were sub-cultured every 48 hours from the stationary phase for 100-150 generations (cultures StaA, StaB, StaC). The remaining starter culture was stored to be used as original starter culture control (Org).

Lab-Scale Cheese Manufacturing and Sampling

The time-course analyses of the microbial communities present in mini-cheeses that were manufactured with the adapted cultures and the original starter culture were studied with a mini scale cheese model (10). 3L of thermally treated (13.7 s at 68°C), bactofugated, and pasteurized (13.7 s at 73°C) bovine milk was standardized for fat/protein ratio of 1.085 with a final fat, protein and lactose content of 3.56% , 3.28% and 4.41% (w/w), respectively. Prior to cheese making, the original starter culture controls (2 cultures) and the adapted cultures (6 cultures) were propagated in sterile skimmed milk (10%, w/v) with 1% (w/v) inoculation followed by incubation at 20°C

for 20 hours, and the propagations were repeated two times. The mini cheeses were manufactured in duplicate for each starter culture.

Each mini cheese was manufactured from 3 L of milk and with 1% (w/v) of starter culture addition as described by Bachmann et al (10). At the end of the pressing, the mini cheeses were incubated at 30°C until the pH was reduced to the range of 5.4-5.6. Subsequently, they were brined in 19% (w/v) sodium chloride solution for 4 hours at 13°C, flushed with 0.8 atmospheric pressure of nitrogen, vacuum-sealed in a plastic foil and ripened at 13°C for 6 weeks. The manufactured cheeses had a final moisture contents in the range of 42-44% and pH values ranging between 5.4 and 5.2. The cheeses were sampled throughout the manufacturing process at 8 time points: (i) from the propagated starter culture just before cheese milk inoculation, (ii) after pressing of the cheeses, (iii) after 24 hours just before brining, (iv) after brining, and from (v) 2-weeks, (vi) 4-weeks, (viii) 6-weeks and (iX) 6-weeks ripened cheeses during the ripening period.

Total DNA Isolations during Cheese Manufacturing and Ripening

For total DNA isolation, 2.5 gr of propagated starter culture or cheese sample was mixed with 2% (wt/vol) Na-citrate buffer (pH 8.0) at 45°C in 25 ml final volume, and homogenized for 5 min using the stomacher. The homogenate was centrifuged at 8000*g for 10 min at 4°C (Beckman-Coulter, Avanti JE Centrifuge, JA20 rotor, Fullerton, CA). The supernatant and fat layers were removed and bacterial cell pellets were washed three times and re-suspended in 1 ml of the same buffer.

Prior to total DNA isolation, the cell suspension was treated with propidium monoazide (PMA) in order to prevent the amplification of the DNA originated from dead or membrane compromised cells during quantitative PCR analysis. PMA was added into 1 ml of cell suspension in 2% (wt/vol) Na-citrate solution (pH 8.0) to a final concentration of 50 μM. The cell suspension was thoroughly mixed (vortex), incubated in the dark for 5 min at room temperature, followed by placing the suspension on ice and exposure to light for 5 min as described (This thesis, Chapter 4). The cells were harvested by centrifugation at 6800*g and 4°C for 5 min (Eppendorf Centrifuge 5417R, NY), and the supernatant was discarded to remove excess PMA. Subsequently, the cells were washed once more in 1 ml of Na-citrate (pH 8.0) buffer, and the total DNA was isolated from the final, fat-free cell pellets using QIAgen DNeasy Blood and Tissue Isolation Kit (QIAgen, Hilden, Germany)

as described (This thesis, Chapter 4). The total DNAs were isolated in duplicate from each sample.

Quantitative PCR

The community composition of the cheese and starter culture samples were determined with genomic lineage specific quantitative PCR. The lineage specific primer sets that were employed in qPCRs were identified by orthologous group analysis of the whole genome sequences of representative TIFN strains as described in our previous study (This thesis, Chapter 3). Lineage specific primer sets (6 sets) were designed targeting genetic markers specific for genetic lineages 3, 6, 7, and 8 separately, and for the lineage 1&5, and the lineage 2&4 as groups of lineages (This thesis, Chapter 3). The qPCR reactions were performed in 20 µl of final reaction volume, combining 10 µl of SYBRgreen master mix (Applied Biosystem, UK) with 0,2 µl of each primer (10 µM), and 2 µl of DNA template. PCR amplification (7500 Fast Real-Time PCR System, Applied Biosystem, Warrington, UK)) was initiated with 10 min of initial denaturation at 95°C, followed with 40 cycles of 15 s of denaturation at 95°C and 1 min of annealing and extension at 60°C. The Ct values were calculated with the default settings.

Flavour Compound Profiling

For the relative quantification of key flavour compounds for Gouda cheese varieties in 6-weeks ripened mini-cheese samples, a headspace solid phase micro extraction gas chromatography (HS-SPME-GC) coupled to mass spectrometry (HS-SPME-GC/MS) was applied. 1 gram of grinded cheese sample was transferred into a 10 ml amber screw cap vial. A grey SPME fibre (CAR/PDMS/DVB) was exposed to the sample. The fibre was then desorbed in a hot split-less injector and the analytes were refocused in a cold trap. After trapping, the analytes were separated on an Rxi-1ms 30 m*0.25 µm column with a film thickness of 1 µm using a helium flow of 1.5 ml/min. The GC column was initially held at 20 °C for 0.4 minutes, subsequently raised to 250 °C with a rate of 200 °C/min and held for 0.5 minutes. Mass spectra were determined using Thermo plus Time of Flight mass spectrometer (Thermo Fisher Scientific Inc., Waltham, USA). Mass spectra were recorded over a range of m/z 30-200. The peaks were identified using NIST MS version 2.0 and the peak areas were quantified using XCalibur 1.4SR1 software (Thermo Fisher Scientific Inc., Waltham, USA). The analyses were performed in duplicate. The Pairwise similarities of flavour

profiles were calculated using the Pearson correlation coefficient, clustering was performed applying the unweighted-pair group method with arithmetic averages algorithm (UPGMA) using GeneMath software (Applied Maths, Sint Martens-Latem, Belgium).

Results

Effect of Back-Slopping on the Starter Culture Community Composition

The effect of the variations of back-slopping regime on final community composition of undefined cheese starter cultures was investigated with the model starter culture Ur. Ur is an undefined mixed-strain starter culture originating from artisanal cheese making practices. The culture, most likely has a long history of exposure to bacteriophages present in the dairy farm environment (8). Being classified as a practice (P) starter culture, it is highly resistant against phage attacks from the environment (11). We have previously reported on the composition of Ur that was determined using metagenomics and AFLP typing of community members (This thesis, Chapter 3). Approximately 1.4% of the community was classified as *Lc. mesenteroides* subsp. *cremoris* and the remaining 98.6% of the community consisted of 7 genetic lineages of *L. lactis*. The vast majority of *L. lactis* subsp. *cremoris* cells belonged to genetic lineage 7, comprising approximately 78 % of the total starter community. The remaining isolates belonged to four genetic lineages of *L. lactis* subsp. *cremoris* (genetic lineage 1, 5, 3 and 6) and two genetic lineages of *L. lactis* subsp. *lactis* biovar. *diacetylactis* (genetic lineage 2 and 4).

Variations in the back-slopping regime were based on the growth curve of the starter culture. Two propagation transfer time points were selected: (i) from the exponential phase of growth and (ii) from the stationary phase of growth. The starter cultures were sequentially sub-cultured every 12 hours from the exponential phase for 500-700 generations (in triplicate, coded with ExpA, ExpB, ExpC) and every 48 hours from the stationary phase for 100-150 generations (in triplicate, coded with StaA, StaB, StaC) in 10% sterile skimmed milk at 20 °C. In order to determine the changes that occurred within the community compositions of propagated cultures, the initial and the final community compositions of cultures were analysed by with quantitative PCR (qPCR) using primer pairs that amplify genetic loci that are specific for each genetic lineage.

Community analyses revealed that the cultures that were subcultured from stationary-phase (StaA, StaB, StaC) were dominated by *L. lactis* ssp. *cremoris* cells that belong to genetic lineages 1&5 together with a smaller contribution of *L. lactis* ssp. *lactis* biovar. *diacetylactis* strains that belong to the lineages 2&4 (Figure 1, Panel A). The anticipated selective pressure that the starter culture was exposed to in the stationary phase is the high acidity during this stage of growth due to lactic acid accumulation. Improved persistence of *diacetylactis* strains under acid stress may be correlated with their capacity to execute citrate conversion, which generates additional metabolic energy to cope with cell wall damage and contributes to the pH homeostasis (12) as well as glutamate decarboxylase activity (13). However, there is no obvious explanation for the better survival of protease positive *cremoris* cells that belong to lineage 1&5. Interestingly, in a previous study, we demonstrated that the cells that belong to the lineages 1&5, as well as 2&4 were relatively better survivors under cheese ripening conditions (This thesis, Chapter 3). Lineages 2&4 were also found to exhibit the highest resistance towards NaCl exposure (data not shown), corroborating the broad spectrum robustness of these lineages under stress conditions associated with cheese manufacturing conditions.

The cultures that were propagated from exponential phase (ExpA, ExpB, ExpC) were found to be dominated with the protease negative *L. lactis* ssp. *cremoris* cells that belong to genetic lineage 7 (Figure 1, Panel A). The genetic lineage 7 cells were previously shown to dominate the original starter culture Ur and the microbial community present during the initial acidification stage of the cheese making process, prior to the brining step (This thesis, Chapter 3). These findings support the competitive advantage of genetic lineage 7 cells during the exponential phase of the growth, which most likely relates to the relatively high growth rate of this protease negative lineage that does not have the burden to protease expression in a community that also encompasses the protease positive lineages 1&5, and 3, which is in agreement with previous observations (14, 15).

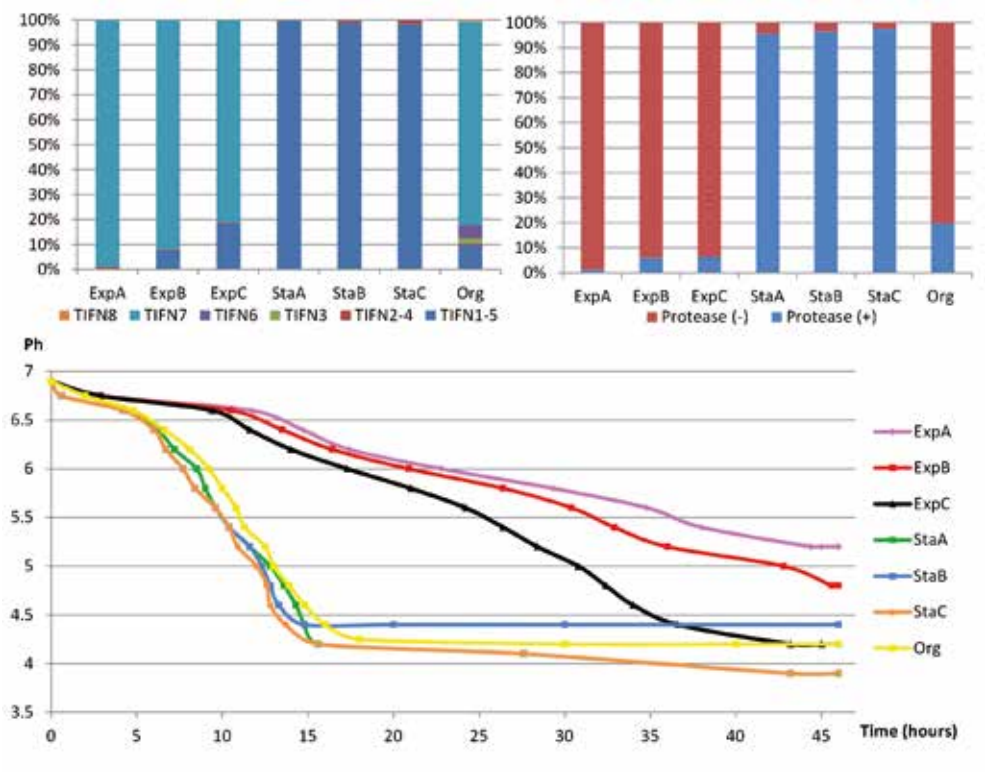


Figure 1. Composition and Functional Properties of Adapted Cultures and the Original Starter Culture. The compositions of the starter cultures (Panel A) were determined based on qPCR data. The contribution of protease positive/negative cells (Panel B) were determined based on Glucose-Milk-Agar (GMA) Plating. The acidification rates of the starter cultures (Panel C) were determined in sterile 10% skimmed milk at 20°C. All analyses were performed on three parallel lines of exponential phase propagated cultures (ExpA, ExpB, ExpC), stationary phase propagated cultures (StaA, StaB, StaC) and the original starter culture (Org).

Characterisation of Exponential and Stationary Phase Propagated Cultures

Acidification and proteolytic activity are among the major functionalities that are considered to be essential during cheese production. Therefore, these two features were evaluated in the newly developed starter cultures. The original starter culture and the cultures that were serially propagated from exponential and stationary phases were transferred (1% (v/v) inoculum) to skimmed milk and incubated at 20°C. The time required by the cultures to reach pH 4.0 was taken as a measure for their acidification rates. The stationary phase propagated cultures (StaA, StaB, StaC) displayed the highest acidification rates and required 15 hours to reach pH 4.0, whereas the original starter culture UR (Org) required 20 hours to reach this

pH and the exponential phase propagated cultures (ExpA, ExpB, ExpC) required more than 48 hours (Figure 1, Panel C). The relative abundance of protease-positive cells in the cultures were determined by plating the cultures on glycerophosphate milk agar plates (16), revealing a high level of dominance of protease positive cells in stationary phase propagated cultures (> 90 %), whereas the protease-positive population was estimated to be only 10% of the overall population in the cultures propagated from the exponential phase of growth (Figure 1, Panel B). An intermediate level of protease-positive cells was encountered in the original starter culture (approximately 20 % of the overall culture). These values imply that acidification rates are not directly related to the relative abundance of the protease positive population, for which similar values were found for the original culture and the stationary phase propagated culture (20 and 15 hours, respectively).

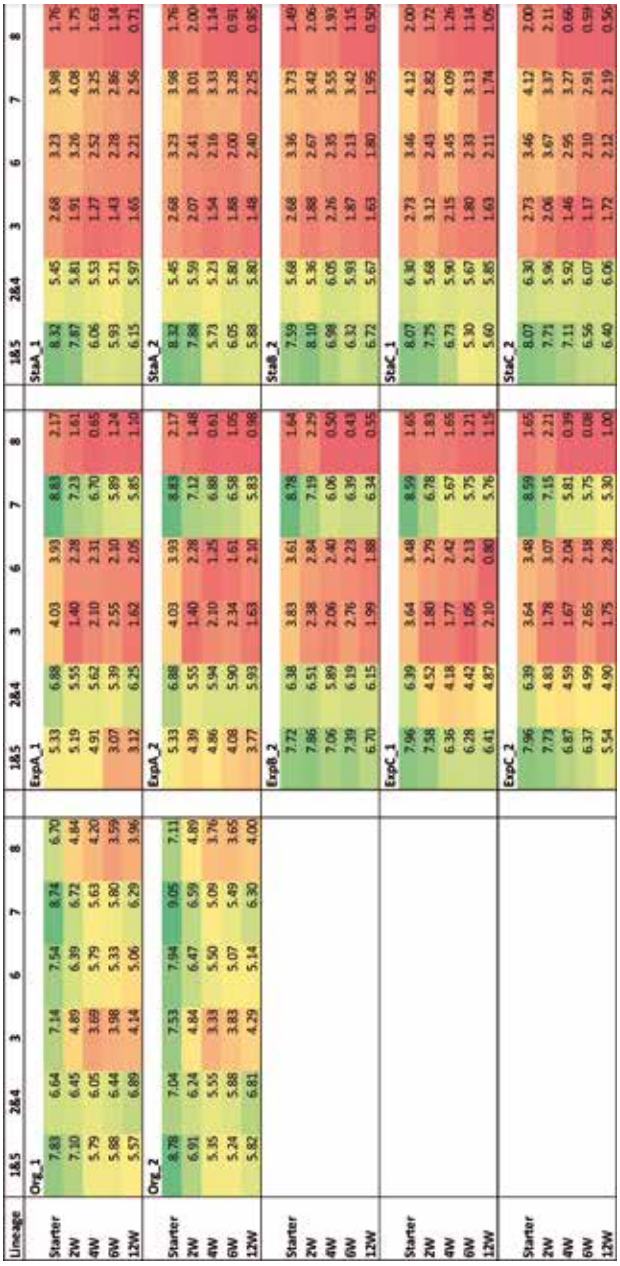
Community Dynamics during Mini-cheese Manufacturing and Ripening

Next, the exponential and stationary phase propagated starter cultures as well as the original starter culture were applied in cheese making using the mini-cheese model (10). The mini-cheeses were manufactured in three parallel production lines (A, B, and C) of exponential and stationary phase propagated starter cultures, encompassing two biological replicates (1 and 2) and the community compositions of the viable cells in the cheeses were monitored throughout the 12 week ripening period. Additionally, the final volatile metabolite profiles in 12-weeks ripened cheeses were determined to describe the impact of the community composition (and its dynamics) on flavour formation.

The initial community of the stationary phase propagated starter cultures during mini-cheese making was dominated with representatives of the protease positive genetic lineages 1&5 (protease positive *L. lactis* subsp. *cremoris*, 99%), and only had a minor contribution from the representatives of genetic lineages 2&4 (*L. lactis* subsp. *diacetylactis*) (Figure 2). The relative abundance of the latter genetic lineages (2&4) gradually increased during the mini-cheese ripening period. This is most likely due to a better survival of lineages 2&4 as compared to the lineages 1&5 representatives during the ripening period, and is not a consequence of a substantial amount of growth of either of these lineages. This is supported by the finding that the qPCR based estimate of the lineages 2&4 abundance was stable throughout the entire process of cheese production and ripening, whereas the qPCR estimated abundance of lineages 1&5 representatives decreased 10 to 1000 fold. The other genetic lineages

remained below the detection limit (10E3/g of cheese) at all stages except for genetic lineage 7, which was detected in the initial starter cultures, but subsequently vanished from the mini-cheese derived communities during cheese ripening.

The initial community composition of exponential phase propagated starter cultures that were employed for mini-cheeses manufacturing were consistently dominated by protease-negative cells that belong to genetic lineage 7 (90-to 99%), and contained minor populations of genetic lineages 2&4 (1%) (Figure 2). The representatives of protease-positive lineage 1&5 represented almost 10% of the total community in the starters used in ExpB, and ExpC, whereas, the abundance of these lineages was estimated to be approximately 100-fold lower in the ExpA starter cultures (0.1% of the total community). The remaining genetic lineages were below the detection level in these initial starter cultures. The high abundance of representatives of lineage 7 observed in initial starter cultures appeared to gradually decrease during cheese production, finally ending approximately a 1000-fold lower than the initial abundance in the cheeses manufactured with any of the exponential phase propagated cultures (Exp A, ExpB, ExpC). Interestingly, lineages 2&4 representatives appeared to sustain a stable abundance during cheese production and ripening, with the only exception of ExpC starter cultures. The abundance of lineage 2&4 representatives in the ExpC starter cultures appeared to decrease approximately 50-fold during cheese ripening. Analogous to the community dynamics observed for genetic lineage 7, all of the exponential phase propagated cultures displayed a gradually decreasing abundance of lineage 1&5 during cheese ripening. The relative abundance of genetic lineage 1&5 decreased around 10 fold for the starter cultures ExpB and ExpC in ripening period. Unfortunately, the relative abundance decrease of genetic lineage 1&5 in the ripening



Lineage	1&5	2&4	3	6	7	8		1&5	2&4	3	6	7	8		1&5	2&4	3	6	7	8
	Org_1							ExpA_1							StaA_1					
Starter	7.83	6.64	7.14	7.54	8.74	6.70		5.33	6.88	4.03	3.93	8.83	2.17		8.32	5.45	2.68	3.23	3.98	1.76
2W	7.10	6.45	4.89	6.39	6.72	4.84		5.19	5.55	1.40	2.28	7.23	1.61		7.87	5.81	1.91	3.26	4.08	1.75
4W	5.79	6.05	3.69	5.79	5.63	4.20		4.91	5.62	2.10	2.31	6.70	0.65		6.06	5.53	1.27	2.52	3.25	1.63
6W	5.88	6.44	3.98	5.33	5.80	3.59		3.07	5.39	2.55	2.10	5.89	1.24		5.93	5.21	1.43	2.28	2.86	1.14
12W	5.57	6.89	4.14	5.06	6.29	3.96		3.12	6.25	1.62	2.05	5.85	1.10		6.15	5.97	1.65	2.21	2.56	0.71
	Org_2							ExpA_2							StaA_2					
Starter	8.78	7.04	7.53	7.94	9.05	7.11		5.33	6.88	4.03	3.93	8.83	2.17		8.32	5.45	2.68	3.23	3.98	1.76
2W	6.91	6.24	4.84	6.47	6.59	4.89		4.39	5.55	1.40	2.28	7.12	1.48		7.88	5.59	2.07	2.41	3.01	2.00
4W	5.35	5.55	3.33	5.50	5.09	3.76		4.86	5.94	2.10	1.25	6.88	0.61		5.73	5.23	1.54	2.16	3.33	1.14
6W	5.24	5.88	3.83	5.07	5.49	3.65		4.08	5.90	2.34	1.61	6.58	1.05		6.05	5.80	1.88	2.00	3.28	0.91
12W	5.82	6.81	4.29	5.14	6.30	4.00		3.77	5.93	1.63	2.10	5.83	0.98		5.88	5.80	1.48	2.40	2.25	0.85
								ExpB_2							StaB_2					
Starter								7.72	6.38	3.83	3.61	8.78	1.64		7.59	5.68	2.68	3.36	3.73	1.49
2W								7.86	6.51	2.38	2.84	7.19	2.29		8.10	5.36	1.88	2.67	3.42	2.06
4W								7.06	5.89	2.06	2.40	6.06	0.50		6.98	6.05	2.26	2.35	3.55	1.93
6W								7.39	6.19	2.76	2.23	6.39	0.43		6.32	5.93	1.87	2.13	3.42	1.15
12W								6.70	6.15	1.99	1.88	6.34	0.55		6.72	5.67	1.63	1.80	1.95	0.50
								ExpC_1							StaC_1					
Starter								7.96	6.39	3.64	3.48	8.59	1.65		8.07	6.30	2.73	3.46	4.12	2.00
2W								7.58	4.52	1.80	2.79	6.78	1.83		7.75	5.68	3.12	2.43	2.82	1.72
4W								6.36	4.18	1.77	2.42	5.67	1.66		6.73	5.90	2.15	3.45	4.09	1.26
6W								6.28	4.42	1.05	2.13	5.75	1.21		5.30	5.67	1.80	2.33	3.13	1.14
12W								6.41	4.87	2.10	0.80	5.76	1.15		5.60	5.85	1.63	2.11	1.74	1.05
								ExpC_2							StaC_2					
Starter								7.96	6.39	3.64	3.48	8.59	1.65		8.07	6.30	2.73	3.46	4.12	2.00
2W								7.73	4.83	1.78	3.07	7.15	2.21		7.71	5.96	2.06	3.67	3.37	2.11
4W								6.87	4.59	1.67	2.04	5.81	0.39		7.11	5.92	1.46	2.95	3.27	0.66
6W								6.37	4.99	2.65	2.13	5.75	0.08		6.56	6.07	1.17	2.10	2.91	0.59
12W								5.54	4.90	1.75	2.28	5.30	1.00		6.40	6.06	1.72	2.12	2.19	0.56

Figure 2. Community Dynamics of mini-cheeses that were manufactured using exponential phase propagated cultures (ExpA, ExpB, ExpC), stationary phase propagated cultures (StaA, StaB, StaC) and the original starter culture (Org) in the time course of 12 weeks ripening period. 2w, 4w, 6w and 12w represents the time points 2 weeks, 4 weeks, 6 weeks and 12 weeks, in cheese ripening period respectively. The cheeses were manufactured in biological duplicates (indicated as 1 and 2) and ExpB_1 and StaB_1 were excluded from the study due to contamination. The abundances of genetic lineage reported are the absolute abundances of the template DNA, quantified with qPCR in samples treated with PMA and the abundances are given in logarithmic scale.

period were not possible to calculate for ExpA cultures since the abundance of this lineage was already below the detection limit due to its very low abundance in the initial ExpA starter cultures, which had a clear impact on the acidification rates observed in the cheese production process as well (slow acidification, data not shown).

The community dynamics of the original starter culture resembled very much to the dynamics previously reported for the original UR starter culture (This thesis, Chapter 3), which is characterised by a lineage 7 dominated initial starter cultures (60-to-80% of the overall community), and a clearly detectable level of each of the other genetic lineages. The abundance of viable cells representing each of the genetic lineages decreased between 100-to-1000 fold during the 12-weeks ripening period, except the representatives of lineages 2&4 that appeared to comprise a relatively consistent size in terms of viable population.

Flavour Formation in Mini-Cheeses

The key flavour compounds that give the characteristics flavour to Gouda type cheese such as aromatic and branched chain aldehydes, acids, alcohols and sulphur compounds are predominantly generated in cheese through the degradation of casein and casein-derived peptides, fat, lactose and citrate in milk by the enzyme repertoire provided by starter culture bacteria (17). To evaluate how the community composition changes elicited by the alternative propagation regimes impacted on the flavour component profiles after cheese ripening, we determined and compared the volatile flavour compound profiles of the mini-cheeses (12-weeks ripened) manufactured with the three different starter cultures described above.

The most abundant flavour compounds derived from amino acid catabolism were 3-methylbutanal, 2-methylbutanol, phenyl-ethyl alcohol and 2-methylpropanoic acid (Fig. 3). No sulphur compounds derived from methionine and cysteine degradation could be detected in any of the mini-cheeses. In general, the flavour compounds originating from amino acid degradation were more abundant in the cheeses manufactured with starters obtained by propagations from exponential and stationary phase as compared to the cheeses manufactured with the original starter culture.

Gouda type cheeses have a high fat content (more than 48%), and enzymatic fat

hydrolysis (lipases and esterases produced by the starter culture bacteria) and chemical oxidation alters the chemical composition of the fat in cheese during the ripening period (18). Fatty acids that are produced through the hydrolysis of triglyceraldehydes, such as butanoic and hexanoic acids were the most abundant fat-derived flavour compounds detected in all mini-cheeses. These compounds were most dominantly present in mini-cheeses manufactured with the original starter culture and the exponential phase propagated cultures. The esterification product of alcohols and fatty acids such as aliphatic esters and the oxidation product of unsaturated fatty acids such as ketones and alcohols were also detected at lower levels, but aldehydes were consistently absent.

Besides protein and fat degradation, lactose and citrate metabolism is very important for flavour formation and texture development of cheese. Lactic acid bacteria ferment lactose into lactic acid, which reduces the pH and determines the Ca^{2+} and phosphate loss during whey drainage, thereby determining the elasticity characteristics of the cheese (19). The breakdown of the citrate that is present in milk leads to the production of diacetyl, which gives the highly appreciated buttery flavour in Gouda cheeses. In addition, citrate metabolism also contributes to the formation of several other metabolic products such as acetate, $C0_2$, acetoin, butanediol and formate, in a manner that depends on carbohydrate limitation, the presence of additional substrates, the degree of aeration and the pH of the fermentation (20). Citrate utilization is only executed by *L. lactis* ssp. *diacetylactis* and/or *Lc. mesenteroides* in Gouda type cheeses because the other *L. lactis* species in the Gouda-type starter cultures lack the citrate permease as well as the citrate lyase (20). Most of the diacetyl produced in cheese is converted into acetoin and 2,3-butandiol during ripening. The highest amounts of diacetyl and acetoin were produced in the mini-cheeses that were manufactured using the original starter cultures, whereas the highest levels of acetic acid were detected in cheeses produced with the exponential phase propagated starter cultures, which coincided with the lowest levels of diacetyl and acetoin.

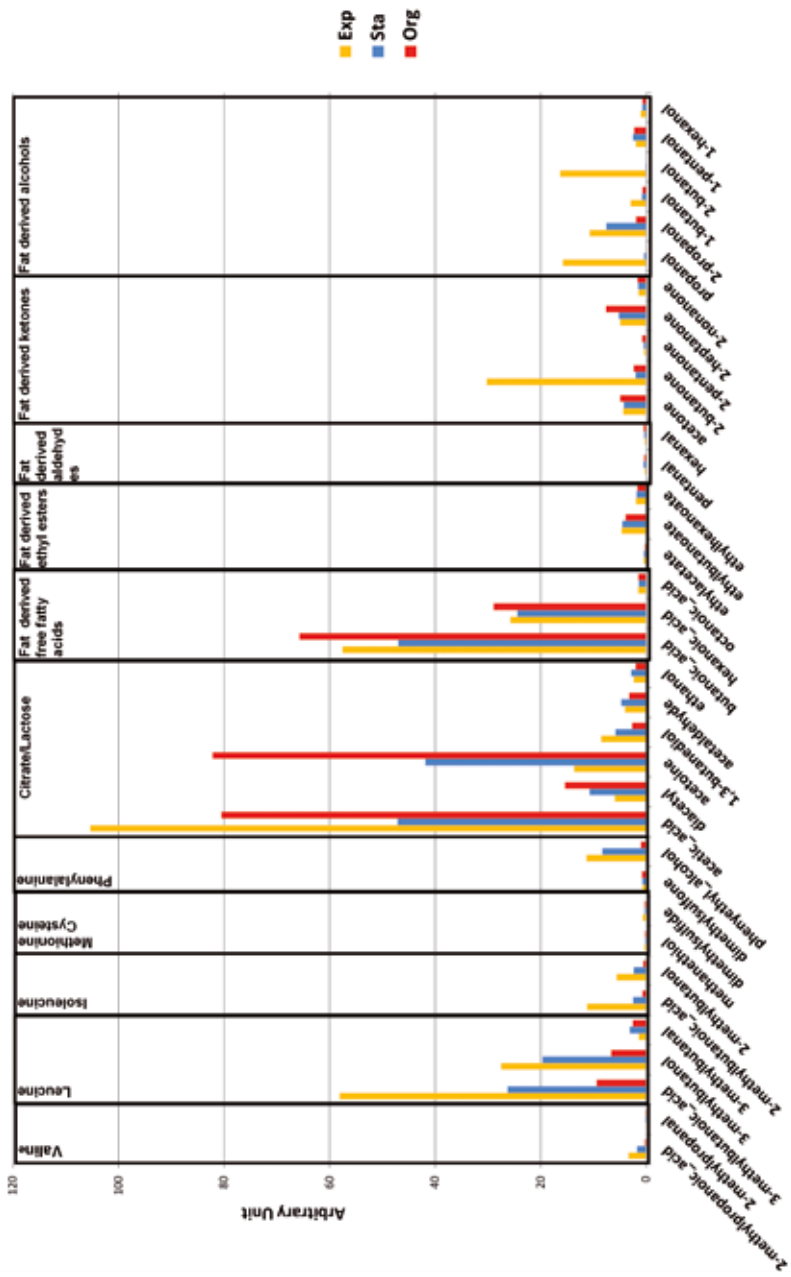


Figure 3. Quantities (arbitrary units) of flavour compounds measured with GC/MS as number of counts. The abundances of flavour compound given in the graph as Exp, Sta, and Org are the averages of the abundances of flavour compounds measured in the cheeses that were produced using exponential phase propagated cultures (ExpA, ExpB, ExpC, in duplicate), stationary phase propagated cultures (StaA, StaB, StaC, in duplicate) and the original starter culture (Org_1, Org_2, in duplicate) respectively. The quantities are additionally sub-grouped on the x-axis based on the sources of flavour compounds.

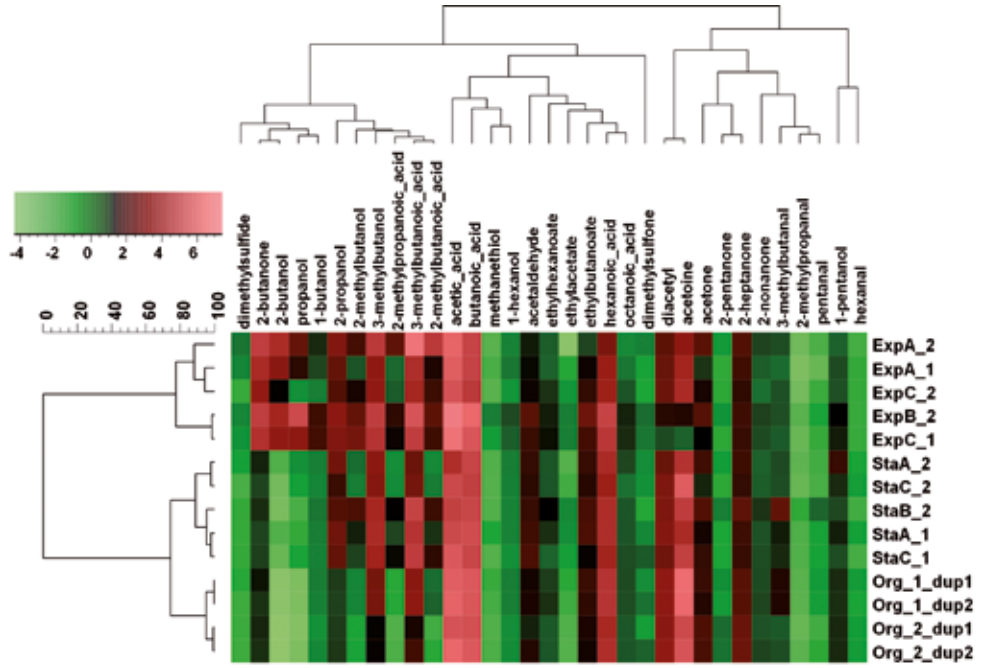


Figure 4. Comparison of flavour profiles obtained from 12 weeks-ripened cheeses manufactured using exponential phase propagated cultures (ExpA, ExpB, ExpC), stationary phase propagated cultures (StaA, StaB, StaC) and the original starter culture (Org). The cheeses were manufactured in biological duplicates (indicated as 1 and 2). The flavour profiling of cheeses Org_1 and Org_2 were performed in duplicate (indicated as dup1 and dup2). ExpB_1 and StaB_1 were excluded from the study due to contamination. The quantities of flavour compounds were log2 transformed prior to analysis, and clustering was performed using Pearson correlation coefficient with UPGMA for both starter cultures and flavour compounds.

To determine possible correlations between the variations of flavour compounds detected in the different mini-cheeses and the community composition of the starter culture with which they were produced, the complete flavour profiles of mini-cheeses were clustered using Pearson correlation coefficients with UPGMA. Importantly, the flavour profiles of the different mini-cheeses could be separated into distinct sub-clusters, that were grouped according to the type of starter culture employed for their production. The cheeses produced with the exponential phase propagated starter cultures clustered together and were most distinct from rest of the mini-cheeses in the analysis. Although more closely related, also the cheeses produced with the stationary phase propagated and original starter cultures could be separated into sub-clusters. These results illustrate that the modulations of the starter culture composition introduced by the different back-slopping regimes has a profound impact on the flavour profile of the cheeses produced with these starters. Typical characteristics of the three distinct flavour profile clusters illustrate that the

cheeses produced with exponential and stationary phase propagated cultures are relatively rich in branched chain amino acid derived flavour compounds, including 2-methylbutanol, 3-methylbutanol, 2-methylpropionic acid, 2-methylbutanoic acid, 3-methylbutanoic acid, and 2-butanol (see also above). Notably, the most distinguishing flavour compounds that enable the separation of these two groups of cheeses, were among the fat derived alcohols like 2-butanone, 2-butanol, propanol, butanol, and the sulphur containing flavour-compound dimethylsulfide. These metabolites are most probably produced via the enzyme repertoire provided by the representatives of genetic lineage 7, which is the only strong discriminator in the starter composition between the exponential and stationary phase propagated cultures. The clustering also exemplifies that the amino acid and fat derived flavour compounds were present at substantially lower levels in the cheeses produced with the original starter culture Ur (see also above). The cheeses made with the original Ur culture uniquely contained detectable levels of the representatives of the *L. lactis* genetic lineages 3, 6 and the *Lc. mesenteroides* lineage 8, which may imply that the amino acid and fat derived compounds could be consumed, or converted into other metabolites, by catabolic pathways present in these genetic lineages (3, 6 and 8). Notably, the citrate degradation derived metabolites diacetyl and acetoin were detected in all cheeses, but they were more abundant in the cheeses produced with the original starter Ur. This may be explained by the presence of the citrate degrading *Lc. mesenteroides* genetic lineage 8 at a clearly detectable level that was uniquely observed in this starter culture, in addition to *L. lactis* ssp. *lactis diacetylactis* genetic lineages 2&4.

Discussion

The metabolic enzymes of *Lactococcus* and *Leuconostoc* strains – either present in intact cells or released into the cheese matrix through autolytic activity- during cheese manufacturing are considered the key players in the process of flavour formation during Gouda cheese ripening (17). Metabolic conversion of lactose and citrate (20), degradation of casein through the activity of proteases, peptidases, and amino acid catabolic pathways (21), and the enzymatic and chemical conversion of milk lipids (22) contribute to the final flavour-compound profile of cheese. Therefore, the strain specific differences among the members of the starter culture in terms of genomic content, transcriptional activity and metabolic regulation of these particular pathways, including the specific activities and substrate preferences of the enzymes

involved play a major role in the final flavour profile generated. In a recent study, the comparative transcriptome analysis of four closely related *L. lactis* ssp. *cremoris* strains revealed that the expression of genes related to the osmotic stress and amino acid catabolism, such as activation of oligopeptide transporters and branched chain aminotransferase *bcaT* were strain specific under the cheddar cheese manufacturing conditions (23). Strain specific variation of osmotic stress tolerance infers viability differences among the members within a starter culture during cheese production, in particular after the brining step, while differential activity of Opp and BcaT could infer differential favour formation capacities linked to amino acid catabolism. Moreover, an elaborate activity profiling study for several flavour forming enzymes in different strains of *L. lactis* exemplified the regulatory diversity of these particular enzymes (24), implying that flavour formation capacities in situ may be highly strain specific. Therefore, steering the community composition of an undefined starter culture is a fruitful approach to alter the final flavour profiles of the cheeses that can be produced with these starter cultures, and the present study clearly illustrates the strength of this approach.

The relatively simple modulation of the propagation regime employed during back-slopping induces drastic changes in the relative and absolute abundances of the various members of the undefined starter community, including the protease-positive versus protease-negative population ratio, and the relative abundance of citrate metabolizing organisms (either *L. lactis* subsp. *diacetylactis* or *Lc. mesenteroides*). Some propagation regimes may generate relatively unreliable community composition outcomes, as to some extent was seen with the exponential phase propagation regime employed in this study, which generated quite different community compositions at the end of the propagation procedure. These variations were reflected in the flavour profiles obtained in the cheeses produced with these cultures (Figure 4), strengthening the conceptual notion that the initial starter community composition is of pivotal importance for the characteristics of the cheese that can be produced with them.

The protease positive members of the starter culture liberate peptides and amino acids from milk proteins, benefiting not only themselves but also supplying the protease negative members of the community with these essential nutrients, thereby playing a prominent role in the overall starter culture performance. Nevertheless, there appears to be no direct relationship between the relative protease-positive population abundance and the rate of acidification, suggesting that very limited amounts of protease positive cells suffice to supply peptides for the overall

community, which is in good agreement with the observation that protease activity may be relatively rapidly lost in *L. lactis* (14, 15). However, the relative abundance appears to have a lower limit, which is illustrated by one of the cultures obtained by the exponential phase propagation regime, which contained less than 1 % of protease positive cells, and failed to efficiently acidify milk.

The classical back-slopping procedure for maintenance and propagation of undefined starter cultures appears to be highly amenable for modulations to create novel starter cultures, with distinct and possibly attractive flavour profile production characteristics. Importantly, the flavour profiles generated in the cheeses produced appear to be strongly correlated to the composition of the starter culture used, implying that on basis of population composition analysis it can be predicted what trends in the alterations of the flavour profile characteristics may be expected in the eventual product that can be produced with these cultures. Based on these findings, it is of interest to evaluate other perturbations that can be implemented within the back-sloping regimes to change the composition of the undefined starter culture communities and thereby create a variety of derivative starter culture compositions with alternative (and predictable) flavour forming capacities during cheese making. For example, one could consider propagating the original undefined starters in conditions that are kept at a constant and relatively high pH values that may drive the community towards domination of fast growing and pH sensitive *L. lactis* ssp. *cremoris* strains, which may be more autolytic during cheese ripening conditions and thus release more of the intracellular peptidolytic enzymes that may enhance amino acid catabolism derived flavour formation. Alternatively, modulation of the propagation temperature employed (in this study standard at 20°C) might improve the relative fitness of *Leuconostoc* compared to *Lactococcus* strains, which may lead to *Leuconostoc* rich starters that could display enhanced CO_2 production (eye-formation) and diacetyl formation during cheese ripening.

During the starter culture propagations reported in this study, serial transfers were performed under aseptic conditions in the laboratory, which may have the drawback that this could lead to reduced phage resistance of the starter culture. The undefined cultures (so-called P cultures) were reported to lose their broad range phage resistance if they were transferred from dairy practices to the laboratory for subculturing (8). Therefore, in order to take this aspect into account, the alternative back-slopping experiments could be performed in not strictly aseptic conditions, e.g., a dairy factory environment.

Acknowledgements

We would like to acknowledge Marja Twigt and Herwig Bachmann for the set-up of the mini-scale cheese manufacturing procedure. This work was funded by Top Institute Food and Nutrition (TIFN) and is part of Kluyver Centre for Genomics of Industrial Fermentation, a Centre of Excellence of The Netherlands Genomics Initiative.

References

1. **Leroy F, De Vuyst, F.** 2004. Lactic acid bacteria as functional starter cultures for the food fermentation industry. Food Science &Technology **15:**67-78.
2. **Smid EJ, Kleerebezem, M. .** 2014. Production of Aroma Compounds in Lactic Fermentations. Annual Review of Food Science and Technology **5**.
3. **Ross RP, Stanton, C., Hill, C., Fitzgerald, G.F., Coffey, A.** 2000. Novel cultures for cheese improvement. Food Science &Technology **11:**96-104.
4. **Beresford TP, Fitzsimons, N.A., Brennan, N.L., Cogan, T.M.** 2001. Recent advances in cheese microbiology. International Dairy Journal **11:**259-274.
5. **Hannona JA, Kilcawley, K.N., Wilkinson, M.G., Delahunty, C.M. , Beresford, T.P. .** 2007. Flavour precursor development in Cheddar cheese due to lactococcal starters and the presence and lysis of *Lactobacillus helveticus*. International Dairy Journal **17:**316-327.
6. **Ayad EHE, Verheul, A., Engels, W.J.M., Wouters, J.T.M., Smit, G.** 2001. Enhanced flavour formation by combination of selected lactococci from industrial and artisanal origin with focus on completion of a metabolic pathway. Journal of Applied Microbiology **90:**59-67.
7. **Holzapfel WH.** 2002. Appropriate starter culture technologies for small-scale fermentation in developing countries. International journal of food microbiology **75:**197-212.
8. **Stadhouders J, Leenders, G.J.M.** 1984. Spontaneously developed mixed-strain cheese starters. Their behaviourstowards phages, and theri use in Dutch cheese industry. Netherlands Milk and Dairy Journal **38:**25.
9. **Sieuwerts S, Molenaar D, van Hijum SA, Beerthuyzen M, Stevens MJ, Janssen PW, Ingham CJ, de Bok FA, de Vos WM, van Hylckama Vlieg JE.** 2010. Mixed-culture transcriptome analysis reveals the molecular basis of mixed-culture growth in *Streptococcus thermophilus* and *Lactobacillus bulgaricus.* Applied and environmental microbiology **76:**7775-7784.
10. **Bachmann H, de Wilt L, Kleerebezem M, van Hylckama Vlieg JE.** 2010. Time-resolved genetic responses of *Lactococcus lactis* to a dairy environment. Environmental microbiology **12:**1260-1270.
11. **Stadhouders J.** 1986. The control of cheese starter activity. Netherlands Milk and Dairy Journal **40:**18.
12. **Magni C, de Mendoza D, Konings WN, Lolkema JS.** 1999. Mechanism of citrate metabolism in *Lactococcus lactis*: resistance against lactate toxicity at low pH. Journal of bacteriology **181:**1451-1457.
13. **Nomura M, Nakajima I, Fujita Y, Kobayashi M, Kimoto H, Suzuki I, Aso H.** 1999. *Lactococcus lactis* contains only one glutamate decarboxylase gene. Microbiology **145 (Pt 6):**1375-1380.
14. **Hugenholtz J, Splint R, Konings WN, Veldkamp H.** 1987. Selection of Protease-Positive and Protease-Negative Variants of *Streptococcus cremoris*. Applied and environmental microbiology **53:**309-314.
15. **Bachmann H, Molenaar D, Kleerebezem M, van Hylckama Vlieg JE.** 2011. High local substrate availability stabilizes a cooperative trait. The ISME journal **5:**929-932.

16. **Limsowtin GKY, Terzaghi, B.E.** 1976. Agar medium for the differentiation of "fast" and "slow" coagulating cells in lactic streptococcal cultures. N.Z.J. Dairy Science and Technology **11:**2.

17. **Steele J, Broadbent J, Kok J.** 2013. Perspectives on the contribution of lactic acid bacteria to cheese flavor development. Current opinion in biotechnology **24:**135-141.

18. **Alewijn M, Sliwinski, E.L., Wouters, J.T.M. .** 2005. Production of fat-derived (flavour) compounds during the ripening of Gouda cheese. International Dairy Journal **15:**733–740.

19. **Singh TK, Drake, M.A., Cadwallader, K.R. .** 2006. Flavor of Cheddar Cheese: A Chemical and Sensory Perspective. Comprehensive Reviews in Food Science and Food Safety **2:**139-162.

20. **Hugenholtz J.** 1993. Citrate metabolism in lactic acid bacteria. FEMS microbiology reviews **12:**165-178.

21. **Yvon M, Rijnen, L.** 2001. Cheese flavour formation by amino acid catabolism. International Dairy Journal **11:**185-201.

22. **Molimard P, Spinnler, H. E. .** 1996. Compounds involved in the flavour of surface mold-ripened cheeses: Origins and properties. Journal of Dairy Science **79:**169–184.

23. **Taibi A, Dabour N, Lamoureux M, Roy D, LaPointe G.** 2011. Comparative transcriptome analysis of *Lactococcus lactis* subsp. *cremoris* strains under conditions simulating Cheddar cheese manufacture. International journal of food microbiology **146:**263-275.

24. **Bachmann H, Starrenburg MJ, Dijkstra A, Molenaar D, Kleerebezem M, Rademaker JL, van Hylckama Vlieg JE.** 2009. Regulatory phenotyping reveals important diversity within the species *Lactococcus lactis.* Applied and environmental microbiology **75:**5687-5694.

Chapter 7

General Discussion and Future Perspectives

General Discussion

The heterogeneity of microbial communities in nature differs widely. Compared to the large and open ecosystems, such as soil and sea, the food associated microbial communities, such as those involved in cheese and sourdough fermentation, have a smaller range of species diversity. For example, a recent 16S rRNA gene pyrosequencing of raw milk microbial community indicated the presence of 61 operational taxonomic units (with 97% sequence similarity cut-off) in milk, including the species from *Firmicutes*, *Proteobacteria*, *Actinobacteria* and *Bacteriodetes* (1), and this initial raw milk diversity was reduced to 21 species in the artisanal cheese fermentation process, favouring the dominance of lactic acid bacteria species. In industrial cheese manufacturing processes, the microbial load of raw milk is even further reduced or completely diminished with the increased demand for control and consistency of the fermentation process and the resulting product. As a consequence cheese starter cultures are tending to contain lower diversity of lactic acid bacteria species to replace the artisanal milk microbiota. For instance, the Gouda cheese starter culture UR studied in this thesis consists of only the species *Lactococcus lactis* and *Leuconostocs mesenteroides*. Interestingly, the members of these species show considerable variability at the strain level in the undefined industrial starter cultures, artisanal cheese microbiota or in their other ecological niches. For example, the *L. lactis* isolates that were recovered from commonly used industrial cheese starter cultures showed large heterogeneity in terms of phage susceptibilities, plasmid contents, and other functionalities (This thesis, chapter 3) (2). A recent comparative genome hybridization analysis of 37 *L. lactis* strains on multi-strain arrays indicated that only 1121 of the 3877 orthologous groups of genes that form the pan-genome of four fully sequenced strains IL1403, KF147, SK11 and MG1363, appeared to be universally conserved in the 37 *L. lactis* strains tested (3). This finding indicates the large size of the *L. lactis* accessory genome. So far, 31 *L. lactis* genomes (11 complete and published) and 7 *Leuc. mesenteroides* genomes (3 complete and published) have been sequenced (Genome online database, 09.02.2014), improving the knowledge on the heterogeneity of the complete genome complement among different strains of these species. Importantly, the variation among *L. lactis* strains has been studied at the level of enzyme activities and flavour metabolite formation. Screening of 84 *L. lactis* strains for their specific activities of the enzymes branched chain aminotransferase (BcaT), aminopeptidase (PepN), X-prolyl dipeptidyl peptidase (PepX), alpha-hydroxyisocaproic acid dehydrogenase (HicDH) and esterase that are considered to be of importance for flavour development in cheese,

revealed significantly different regulatory phenotypes among the strains in different media (4), indicating that even the enzyme levels of *L. lactis* is perturbed in variable environmental conditions and only in situ screening in cheese is of real relevance to evaluate individual strains in cheese manufacturing. Moreover, some of these *L. lactis* strains of dairy and non-dairy origin were also shown to produce variable flavour volatiles in milk and in a cheese model system (5).

All of these functionality studies demonstrate that large strain-to-strain variation is present at the level of metabolic functions that play a role during cheese manufacturing and ripening. Understanding and improving the cheese manufacturing processes, or identifying new starter cultures for different dairy applications requires underpinning of this level of lineage diversity among the strains of the canonically used species, as well as the functional responses of strains and their interactions during the process of cheese manufacturing and ripening. Therefore, in this chapter, the main focus is given to the discussion of the set of tools partly described in this thesis, which enable to deciphering of the microbial diversity in starter cultures using culture dependent and independent approaches, and allow monitoring of lineage-level community dynamics and functional responses. Finally, some future perspectives are discussed that may help to discover the interactions of strains within the cheese-associated microbiota. Overall, the community analyses approaches employed to decipher the cheese starter culture composition and function, may serve as a template for the analysis of other 'simple' microbial communities associated with food production.

High Resolution Fingerprinting of Clonal Sub-Populations in Starter Culture Communities

One of the major targets in this thesis was to discover the diversity within the mixed Gouda type cheese starter culture UR, which has a long history of use in dairy industry, with the ambition to understand how the community structure is shaped under the influences of the stress conditions encountered during cheese manufacturing. Therefore, techniques were required that allow discrimination and monitoring of the clonal sub-populations (genetic lineages) of *L. lactis* and *Leuc. mesenteroides* in cheese. Chapter 1 briefly discussed the evolutionary and ecological basis of the microbial speciation, the level of microbial diversity expected in the mixed cheese starter cultures and the techniques available for the phylogenetic clustering and delineation of the bacterial strains depending on the resolution levels aimed for. The culturing based analysis of the starter culture community, which is

described in Chapter 3, employed amplified fragment length polymorphism (AFLP) on basis of its high resolution and reproducibility to discriminate genetic lineages, and its compatibility with high-throughput fingerprinting of a large set of isolates. For the optimization of its resolution power to be used in starter culture community analysis, the original AFLP technique described in the literature (6) was optimized to generate maximum number of strain-specific markers for the closely related members of *L. lactis* in Chapter 2.

To develop a high resolution AFLP typing, firstly the restriction enzyme/selective nucleotide combinations that are generating the maximum number of strain-specific fragments for *L. lactis* were determined using in silico analysis and slab-gel validations. Use of carboxyfluorescein fluorophore (FAM) labelled primers in combination with capillary electrophoresis upgraded the procedure to a high-throughput level and the resolution power of the fingerprinting technique was validated using the 84 *L. lactis* isolates of dairy and non-dairy origin that were previously used for specific enzyme screening (4). AFLP provided much higher resolution than repetitive PCR (rep-PCR) and multi locus sequence typing (MLST) in the delineation of this set of strains. The subspecies of *L. lactis* are determined based on their phenotypes, in which *L. lactis* ssp. *lactis* is discriminated from ssp. *cremoris* by its capacity to grow at 40°C and in the presence of more than 4% NaCl, and its capacity to utilize arginine and its expression of an active glutamate decarboxylase. In genetic fingerprinting studies, the genotypes of some of these strains did not match with their subspecies designation assigned on basis of their phenotypes (7). For example, the isolates having ssp. *cremoris* type-strain-like genotypes are in some cases encountered to display the phenotypes indicated for ssp. *lactis*, and therefore named as ssp. *lactis*. AFLP analysis was found very useful to discriminate this group of strains with phenotype-genotype discrepancy as well. The topology of the phylogenetic clustering on basis of AFLP patterns was found to be very similar to the topology of comparative genome hybridization clustering, enabling the discrimination of strains with a typical *lactis*-phenotype from a compact cluster of *cremoris*-phenotype strains within the *cremoris*-genotype cluster.

The developed AFLP technique was applied for the culture-based community analysis of a mixed cheese starter culture as well as to monitor the community dynamics in the course of cheese manufacturing in Chapter 3. In principal, the AFLP methodology was implemented for the species *L. lactis*. However, in silico AFLP analysis of the available *Leuc. mesenteroides* complete genome sequences indicated

that the enzyme/primer combination applied for *L. lactis* also is predicted to generate a high number of strain-specific markers for *Leuc. mesenteroides*. Therefore, the high throughput AFLP approach could be used for the effective fingerprinting of the isolates from both species. In total, 8 genetic lineages were discriminated among the isolates recovered from the mixed starter culture. In addition, the AFLP fingerprinting approach also allowed high resolution monitoring of the population dynamics among the genetic lineages during cheese manufacturing.

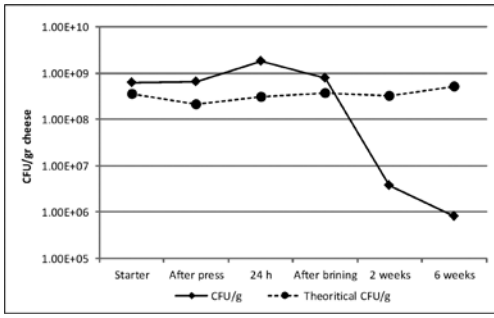
Beside the cheese microbial community, the AFLP method developed was also successfully applied to decipher the richness of *Streptococcus* and *Veillonella* isolates that were recovered from the human small intestine and oral cavity (8, 9). The AFLP fingerprinting of 160 *Streptococcus* and 37 *Veillonella* isolates were differentiated into 8 *Streptococcus* and 1 *Veillonella* genetic lineages. This finding exemplified the presence of the clonal diversity that in many cases remains unexplored in complex ecosystems with high species diversity like the human intestinal tract. The resolution power of AFLP was found to be better than the randomly amplified polymorphic DNA (RAPD) analysis for the *Streptococcus* isolates, and the AFLP-based phylogenetic clustering displayed congruency with the 16S rRNA gene phylotyping, indicating the generic nature and suitability of the technique for the species (or genus) that are closely related to the genus *Lactococcus*.

Cheese Starter Population Dynamics using Molecular Methods; Dead-or-Alive Makes a Difference

The AFLP approach has proven to provide high-resolution delineation of 8 genetic lineages in the cheese starter culture community. During cheese manufacturing, there is a chance that some of the genetic lineages, which remained undetected in the starter culture due to the low abundance, may have better fitness and become more abundant under the conditions encountered later in the process. In order to make sure that all genetic lineages are covered that are present in the starter culture, AFLP analysis was performed on the isolates recovered from several time points during cheese manufacturing and ripening, which revealed the population dynamics of the genetic lineages throughout the process but did not reveal any lineages that remained undetected in the original starter culture. Nevertheless, monitoring the population dynamics of the genetic lineages by AFLP typing of isolates required a substantial effort, implying that a faster high-resolution tool could accelerate community composition analyses in further experiments. To this end, the complete

genome sequencing of the representative strains (TIFN 1-to-TIFN8) selected for each genetic lineage in chapter 3 allowed the identification of a strain-specific orthologous gene pool, which could be exploited for the development of quantitative PCR assays. The major challenge in the development of strain-specific qPCR assays for the TIFN strains was the limited number of strain-specific OGs, of which most were annotated as phage or plasmid encoded functions (approximately 90 % of them). Therefore, only the strain-specific OGs that were predicted to be of chromosomal origin were taken as targets for the primer design, which enabled the detection and quantification of TIFN1&5 and TIFN2&4 as pairs of strains, and TIFN3, TIFN6, TIFN7 and TIFN8 as separate strains. Evaluation of the qPCR primer sets on separate chromosomal and plasmid DNA preparations of the TIFN strains enabled the validation of the chromosomal location of the marker genes targeted. Moreover, in order to investigate if the unique markers employed for representative TIFN strains could also be employed for the discrimination of the genetic lineages, the conservation of the targeted markers were verified for the complete set of isolates clustered within each genetic lineage by performing qPCR for all the starter culture isolates, which supported the reliability of the selected genetic markers for the discrimination of the members of separate lineages. The developed qPCR assays enabled the detection and quantification of genetic lineages in cheese in chapter 4 and 6.

Another difficulty encountered in developing a qPCR detection assay applicable for cheese samples was the discrepancy between the total DNA isolated from these samples and the live bacterial populations present in them. The total DNA isolations at several time points during cheese manufacturing indicated that the amount of DNA isolated at various stages of cheese ripening was more or less equal, whereas the colony forming units (cfu/gr of cheese) recovered from cheese were markedly decreasing during the ripening period (Figure 1). This finding implied that the total amount of DNA recovered from the cheese matrix was stable and included a major amount of DNA derived from cells that were no longer able to form a colony, indicating that the DNA of bacteria was not degraded in the cheese matrix even though the bacteria were losing their culturability.



Calculation Factors	Symbol	Number	Units
Genome size	NT	2.50E+06	nucleotides
Multiplication factor for double strand	D	2	
Correction factor for slow growing cultures	S	1.8	
Atomic mass unit	AU	1.66E-27	kg/Da
Molecular weight nucleotide	MW	660	Da/bp
The amount of DNA recovered from g of cheese	m		µg/g
Calculation			
Total amount of DNA per cell	NT*D*S*MW*AU	9.86E-18	in kg/cell
		9.86E-09	µg DNA/cell
Theoritical CFU/g of cheese	m/9.86E-09		CFU/g

Figure 1. Left Panel: The comparison of the theoretical cfu/ g of cheese calculated from the amount of DNA (m) recovered from cheese with the actual cfu/g of cheese recovered during cheese manufacturing. Right Panel: Calculation of the theoretical cfu/ g of cheese.

In order to eliminate the qPCR confounding DNA of the disintegrated / non-culturable cells in cheese matrix from the total DNA obtained, propidium monoazide (PMA) treatment was incorporated into the procedure of total DNA isolation from milk cultures and cheese. PMA covalently links to DNA upon photo activation and inhibits the DNA amplification during polymerase chain reaction (10). When samples are treated with PMA, the compound is expected to modify only the free DNA that is present in the sample matrix or the DNA residing in cells with compromised membranes, while not modifying the DNA in cells with intact and impermeable membrane. Therefore, the pre-treatment of the cheese sample with PMA prior to DNA extraction and subsequent qPCR would ideally results in the exclusive amplification of the DNA recovered from intact cells in cheese. The implementation of the PMA for qPCR and metagenome based community profiling of the cheese microbiota was described in Chapter 4, and allowed a 100-fold enhanced selective monitoring of the viable microbial community in cheese. This 100-fold enhancement appeared to be relatively low when compared to the other applications of PMA for community analyses that have been reported in the literature (11). Similar levels of selective amplification of viable cells at the expense of damaged or dead cells were also observed in pure LM17 cultures of *L. lactis*, suggesting that the relatively low selective enhancement of amplification was not due to the hindering effect of the complex cheese matrix, but rather due to a species-specific limitation, which is in agreement with the observation that the effect of PMA may vary among different bacterial species (12). Therefore, if PMA is intended to be used in the analysis of the complex communities with high species diversity, the differential PMA efficacy to penetrate in different species with different cell wall structures, such as Gram positive and negative bacteria, should be taken into consideration.

Discriminating the live or disintegrated members of the cheese microbial community is essential to allow the linkage of specific functions, such as production of flavour compounds, to specific community members. Previous studies reported that the decrease in cell viability was highly correlated with the release of the intracellular peptidases X-prolyl dipeptidyl aminopeptidase and aminopeptidase with broad substrate specificity (13), which is proposed to accelerate the release of the amino acids from matrix embedded peptides and maximize the contribution of these enzymes during ripening period. Conversely, the formation of key aroma compounds derived from the degradation of amino acids as well as carbohydrate and/or lipid components may require intact and viable cells that harbour the catabolic enzymes that collectively constitute the required metabolic pathway involved in the metabolic conversion (14). As a consequence, the viability state of bacterial cells is expected to have a profound influence on the final flavour profile of the cheese.

We have also investigated the effects of the viability state on the flavour profiles of 6-weeks ripened micro-cheeses that were prepared with permeabilized and/or viable cultures of the different TIFN strains. Briefly, the extracellular proteases produced by *L. lactis* TIFN 1, 3, 5 cultures were released from the cell wall by incubating the pellet in 5 ml of 50 mM Tris-HCl (pH 6.5) at 25°C for 30 minutes using the protocol described by Juillard and co-workers (15), and the cheese milk was treated with this freshly isolated protease mixture (combined in equal volumes) prior to micro-cheese production, in order to facilitate the peptide release from casein that is required for the growth of protease negative TIFN strains (TIFN 2, 4, 6, 7, and 8). The TIFN strains were cultured in protease treated milk at 20°C for 20 hours in duplicate and one set of the cultures was permeabilized using 5 % butanol treatment, according to the protocol described by Extercate and co-workers (16). Following permeabilization, cells were washed 5 times in prior to their use in cheese manufacturing. The micro-cheeses were manufactured in duplicate using the protease treated cheese milk with permeabilized or intact (viable) cultures of the TIFN strains using the micro-cheese protocol developed by Bachmann and co-workers (17). Micro-cheese manufacturing using the permeabilized bacterial preparations included a pH adjustment to pH 5.4 by the addition of lactic acid during the milk coagulation step to compensate for the incapacity of the permeabilized bacteria to acidify the cheese milk, like their intact and viable counterparts. Finally, the flavour profiles of 6-weeks ripened micro-cheeses were determined as described in chapter 6. The flavour profiles obtained displayed a substantial difference between the micro-cheeses produced with intact (viable) or permeabilized TIFN

culture preparations. In general, the flavour compounds 3-methyl butanal, 2-methyl butanal, 2-nonanone, phenyl ethyl alcohol, isoamyl propanoate, methanethiol were more abundant in micro-cheeses produced with permeabilized cell preparations. In contrast, the flavour compounds diacetyl (see below), hexanoic acid, octanoic acid, and acetaldehyde were more dominant in the micro-cheeses produced with intact and viable bacterial preparations. The variation in flavour formation among the micro-cheeses produced with TIFN 1-to-7 *L. lactis* strains could not be differentiated at high resolution in this set of experiments, although they were all very different compared to the micro-cheeses manufactured with TIFN8 *Leuc. mesenteroides*. The flavour profiles of the micro-cheeses produced with *Leuc. mesenteroides* TIFN8 were distinguished from those manufactured with *L. lactis* strains on basis of their lack of the hexanoic acid, octanoic acid and methanethiol formation. The major difference observed among the flavour profiles of *Leuc. mesenteroides* TIFN8 micro-cheeses was the higher abundance of 3-methylbutanal in the micro- cheeses manufactured with permeabilized cells compared to the ones manufactured with intact and viable cells. Notably, 3-methylbutanal, 2-methylbutanal and 2-methylpropanal are known to be responsible for the sensation of malty or chocolate flavour in cheese (18) and produced via sequential transamination and decarboxylation of leucine, isoleucine and valine, respectively (19, 20). The high abundance of 3-methylbutanal and 2-methylbutanal in the micro-cheeses manufactured with permeabilized cells, suggested that permeabilization facilitated the release of the branched chain aminotransferase and decarboxylases into the cheese matrix, leading to the elevated flavour compound formation. Another interesting finding was the higher abundance of diacetyl in the micro-cheeses produced with intact and viable cell preparations, which was especially detected in micro-cheeses manufactured with the intact and viable cultures of TIFN2, 4, and 8, and to a lower extent in the micro-cheeses manufactured with the intact and viable cultures of TIFN 1,3,5,6 and 7. Importantly, diacetyl formation was not detected at all in the micro-cheeses manufactured with any of the permeabilized cell preparations, indicating that intact and metabolically active cells are required for the formation of diacetyl in cheese.

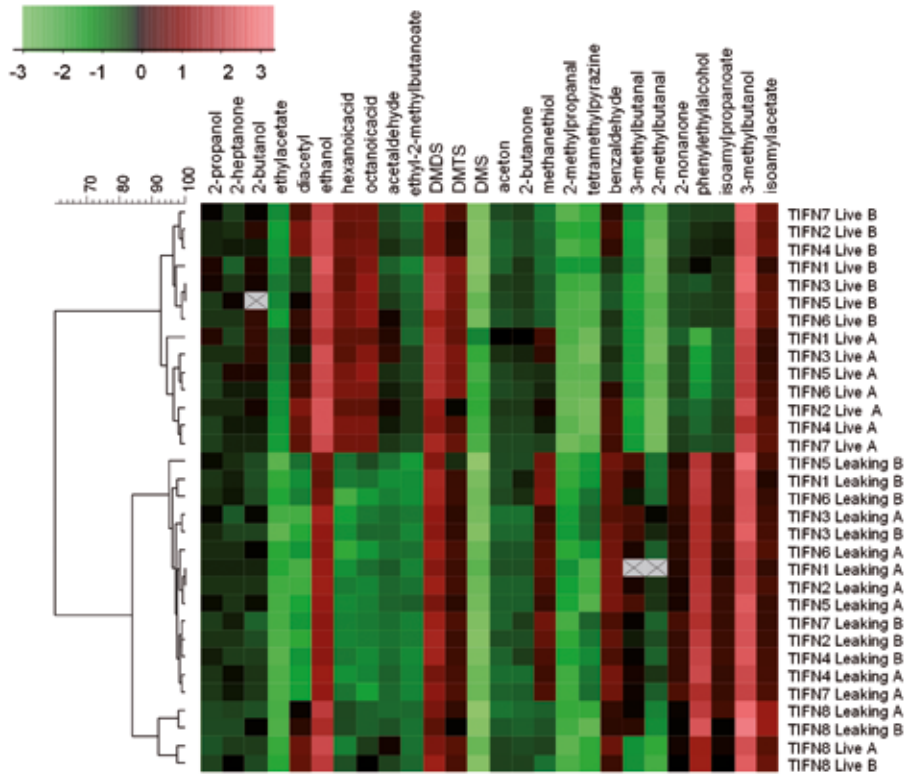


Figure 2. Clustering of the flavour profiles of the micro-cheeses that were manufactured with the permeabilized and intact (viable) cell preparations of TIFN1 to TIFN8 strains. The flavour profiles were log2 transformed and normalized with mean centring prior to clustering, and they were clustered with the UPGMA algorithm using Pearson correlation coefficient.

Starter Culture Composition Modulation and its Impact on Flavour Formation

The developed strain-specific qPCR approach in combination with PMA treatment can be used in many applications in cheese and starter culture research. In chapter 6, the undefined mixed starter culture was serially propagated from exponential and stationary growth phases to evaluate whether back-slopping regimes can be used to steer the community compositions as a novel approach to develop new mixed starter cultures. The strain-specific qPCR in combination with PMA treatment was applied to monitor both the final community compositions of the propagated cultures and their dynamics during cheese manufacturing. Eventually, the final community compositions and functionalities of cultures as well as the flavour formation in 12-weeks ripened mini-cheeses were found to depend on the propagation regime applied. In a similar way, several other perturbations in propagation regimes can

readily be implemented to evaluate their impact on the starter culture composition, such as alternative growth and propagation temperatures, or the addition of specific supplements to the propagation medium. Moreover, other parameters that are normally fixed during cheese manufacturing can be modulated to create diverse applications, in which the influence of the change on the initial starter community composition can be evaluated. Among such modifications one could think of varying the degree of acidification, changing the cooking or ripening temperatures, varying the parameters of scalding to change the lactose content left in the curd, varying the salt concentration in the brine or the duration of the brining step, addition of supplements to the cheese milk (e.g. to create specific functional cheeses), addition of other (adjunct) strains or probiotics to the starter culture community, etc.

An interesting example of an alternative cheese production protocol is the manufacture of low-salt-content cheeses that are attractive in the context of a healthier diet. In order to evaluate the impact of lowering the salt content in brine, we monitored the community composition of the starter culture in the ripening period for the micro-cheeses that were brined in the solutions with regular salt content (19%) and reduced salt content (11,4%). The community composition information deduced from total DNA recovered without any PMA treatment revealed a very big contribution of genetic lineage 7 to the community compositions of both set of cheeses brined with regular and low salt content during the ripening period. This was due to the confounding effect of the free DNA of genetic lineage 7 cells that were membrane compromised (or dead) in the cheese matrix. The qPCR on the total DNA recovered with PMA treatment allowed to determine the dominance of viable *diacetylactis* population from genetic lineage 2&4 in the micro-cheeses brined with 19 % salt concentration. The *diacetylactis* cells are known to be more resistant to salt and low pH due to their citrate metabolism (21, 22) and the activity of their glutamate / GABA antiporter (23). Therefore, the better survival of *diacetylactis* strains in ripening was expected. Interestingly, the reduction of salt content in brine reduced the contribution of *diacetylactis* strains in the community composition. This was not due to the poor survival of *diacetylactis* cells, but due to the increased relative survival of genetic lineages 6 and 1&5, which means that salt reduction was favourable for the survival of genetic lineage 1&5 and 6. The community composition changes induced by lower salt concentrations in cheese production as compared to the normal salt levels may results in modified flavour formation in lower-salted cheeses compared to the regular-salted cheeses. The less efficient cell permeabilization of genetic lineage 1&5 and 6 may lower the release of the peptidases into cheese matrix or the

formation of flavour metabolites that require the complete metabolic pathways of viable cells may be improved due to the better survival of genetic lineage 1&5 and 6. In dairy industry, one of the problems encountered in developing cheeses with low salt content is the poor (or different) flavour development in low-salted cheeses compared to the regular-salted cheeses (24). Considering the changes observed on the community composition of cheeses in the ripening period when salt content is lowered, in general the difference observed in flavour development could be due to the changes in the (viable versus deteriorating) community composition. Changing the process parameters or addition of specific nutrients to modulate the community dynamics in ripening period may contribute to the improvement of the flavour of cheeses with a reduced salt content.

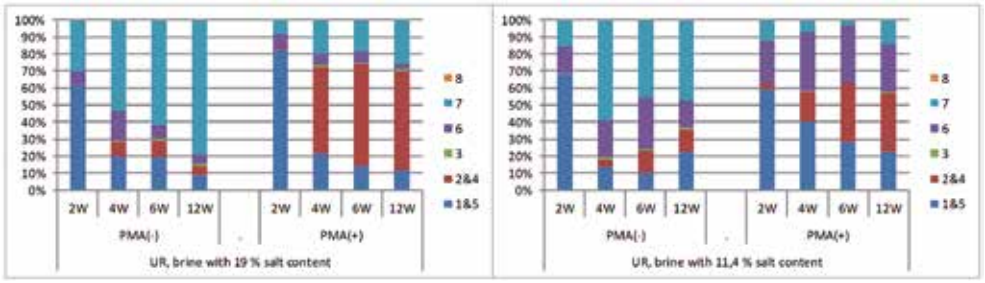


Figure 3. Dynamics of the relative abundances of genetic lineages in micro-cheeses brined with regular salt content (19%, left graph) and reduced salt content (11,4%) during the ripening period. Lineage-specific qPCR quantifications of the total DNA samples treated with PMA (PMA (+)) and not treated with PMA (PMA (-)) were performed in duplicates, and the average of the duplicates is used on the figure.

Metagenomics and Monitoring Genetic Lineages in Metagenome Datasets

Metagenomic analysis is a powerful approach to determine the community composition of environmental samples as well as the metabolic gene complement of the community as a whole in specific environmental niches (25). Because the analysis is based on the total DNA directly recovered from environmental sample, it avoids the culturability biases observed in the classical community analysis approaches. The culture-dependent analysis of microbial communities is hampered by intrinsic biases due to the growth requirements that are not fulfilled by the chosen growth-media or –conditions. For example, in the culture dependent analysis of the cheese starter culture in chapter 3, complementary plating on LM17 and Reddy's agar media were required for the complete recovery of the overall *L. lactis* populations and 5 times more cells were recovered on Reddy's agar medium compared to the LM17 agar medium. The molecular basis of this plate count anomaly was comprehensively investigated

in chapter 5. Interestingly, *L. lactis* cells, which were belonged to genetic lineage 7 and constituted 78% of the starter culture, were found to be recovered exclusively on Reddy's agar medium due to the presence of arginine in medium components compared to the LM17 agar medium. Furthermore, the stimulatory effect of arginine was investigated on *L. lactis* TIFN7 by transcriptome analysis before and after arginine addition in arginine-limited chemostat cultures and the major limitation was determined to be C-1 resources in arginine-limited growth, which resulted in low expression levels on arginine, pyrimidine and purine biosynthesis pathways, probably explaining the limited growth of TIFN7 in LM17 medium. Therefore, the plate count anomaly encountered in culture-based analysis of the starter culture community in this thesis underpins the requirement for multiple complementary media to recover all the members of an undefined community as well as the requirement for culture-independent approach such as metagenomics to complement and/or verify integral recovery of the community members in culturing approaches.

In this context, metagenomics is even more essential for the analyses of the communities with high species diversity, which will otherwise require broad spectrum and demanding culturing procedures to recover all the species present in the community. Therefore, the composition of the communities with high species richness is generally determined by binning the phylogenetic markers that are discriminative for the species, such as 16S rRNA genes, in metagenome dataset. However, this approach is not appropriate for the analysis of more simple microbial communities characterised by low species but high strain diversity, since the discriminative resolution of 16S rRNA gene is insufficient to accurately determine the diversity below the species level. In chapter 3, the relative abundance of 16S rRNA gene sequences that were obtained from metagenome datasets of the starter culture revealed a high level of dominance of *L. lactis* (99%) and a minor population of *Leuc. mesenteroides* (1%), but as expected did not allow further resolution at the genetic lineage level.

In order to circumvent this problem, the strain specific genes (orthologous groups) in the genome sequences of the genetic lineage representative strains (TIFN 1-to-8) were used to determine the contribution of genetic lineages in the metagenome dataset of starter culture community. This approach was successfully used in Chapter 3 for the comparison of culture-independent and culture-dependent analyses of the initial starter culture community, and the genetic lineage contributions deduced from metagenome analyses were in agreement with those obtained by culture

dependent approaches. The strain-specific OGs based genetic lineage quantification in metagenome datasets was also applied in chapter 4 to follow the lineage dynamics during the cheese manufacturing and ripening period. Additionally, in chapter 4, a two-step metagenome sequencing methodology was developed by the inclusion of PMA pre-treatment and selective whole genome amplification to specifically determine the viable population of the cheese community. The quantification of the genetic lineages in the metagenome dataset of the PMA-treated DNA was tested for the 6-weeks ripened cheese sample and it was reflecting the intact and viable population at this time point of cheese making process. Therefore, the approach presented in this chapter appears to hold great promise for the selective metagenome determination of the viable fraction of a complex ecosystem, which may very well be applicable for studying several other food associated or environmental communities.

Metatranscriptomics for the Functional Analysis of Cheese Microbial Communities

L. lactis strains has been subjected to numerous studies to understand the metabolic responses of cells under specific stress conditions associated with cheese manufacturing such as low pH, osmotic change, carbon starvation, temperature change and oxidative stress (26, 27). These comprehensive studies increased our understanding of the metabolism of *L. lactis* under the conditions that might be encountered during cheese manufacturing. However, the responses studied probably do not reflect the real responses of strains when they are part of an starter culture that actually resides in the cheese matrix, since they were almost exclusively performed using pure cultures of single strains in standardized laboratory media, or at best in milk. Therefore, the interactions among the strains as well as the interactions between the strains and continuously changing cheese matrix may not be adequately represented during these analyses. Recently, several techniques were described to unravel the functional responses of *L. lactis* during cheese manufacturing conditions. For example, a recombinase-based *in vivo* expression technology (RIVET) was described to decipher the gene expression analysis in cheese environment, which led to the identification of genes expressed by a *L. lactis* strain during cheese manufacturing in combination with a mixed starter culture (28). This technique effectively revealed the transcriptional responses of a single *L. lactis* strain throughout the manufacturing process, but it is not readily applicable to determine the responses of different microbial community members differentially. Another study reported on the influence of the cheese cooking temperature (38 °C) and brining on the array based transcriptome responses of four closely related *L. lactis* subsp. *cremoris* strains using conditions that simulated cheddar

cheese manufacturing. Comparative analyses of these transcriptomes, revealed strain specific responses to these conditions (29). However, also this approach does not allow the detection of the gene responses to the interactions of different strains within the cheese starter culture. Therefore, the unavailability of high resolution tools to differentiate the transcriptional responses of clonal sub-populations in cheese environment and to link a specific gene activity to a particular sub-population is an important bottleneck in comprehensive functionality studies of cheese microbiota.

Metatranscriptomics, which is studying the total transcriptome response of a microbial community, is a very promising technique to investigate the functional response of the community as a whole under variable environmental conditions or in a particular niche. Just like the transcriptome profiling of single micro-organisms, metatranscriptomes can be obtained using dedicated microarray platforms or through the sequencing of cDNAs that are generated from total mRNAs recovered from environmental samples. The microarray based approach has been used in several food fermentations such as kimchi (30) and sourdough (31) fermentations to determine the active community members and their responses in those particular niches. The major drawback of microarray based analysis is that cross hybridization is a prominent confounder in communities that encompass closely related species and strains. Furthermore, the complete genetic complement of the community should be available to be able to design an array for community. In contrast, metatranscriptome sequencing could enable the high resolution binning of the transcribed sequences to particular species or strains provided that sufficient small differences in the conserved genes are present like single-nucleotide polymorphisms (SNPs), or other structural variations, which can be assessed on by analysis of the genome sequences of the community members. Furthermore, metatranscriptome sequencing may detect the expression of genes expressed in the complete environmental sample, including genes that may not have been annotated on basis of the sequenced strains.

During this thesis project, total RNA isolations were also performed at different time points of cheese manufacturing and ripening with the intention to perform metatranscriptome sequencing. However, the RNA isolation from cheese sample appeared far from trivial, and a variety of modifications on the standard RNA isolation protocols available in kits (Roche Diagnostics, Mannheim, Germany) was evaluated. Prior to cell lysis, cheese samples were homogenized in a stomacher and cleared using a sodium citrate solution (2%), followed by additional washing steps with a Tris-EDTA solution. Cell lysis was subsequently performed by bead beating in the presence of

surfactants (SDS) and denaturing agents (phenol), and RNA was extracted using several chloroform extractions prior to further purification with RNA purification kits (Roche Diagnostics, Mannheim, Germany). RNA isolations during cheese manufacturing indicated that the yield and quality of the material recovered was very much influenced by the changing physicochemical nature of the cheese matrix during the manufacturing process. During the initial stages of cheese making (until brining), the curd was not elastic and the protein impurities of the RNA samples recovered was a major hurdle. In contrast, the cheese matrices at later stages of the manufacturing process were with a higher elasticity and softness, which improved the purity of the RNA recovered. However, these later stage samples suffered from very low RNA recovery, which is likely due to the reduced microbial activity. Overall, RNA isolation during the different stages of cheese manufacturing was not straightforward and provided samples of variable quality and quantity. Nevertheless, the metatranscriptome sequencing was performed on the total RNA samples that were obtained from initial stages of cheese making (until brining) with sufficient level of purity and amount, and the sequence datasets were found to be of reasonable quality. Therefore, metatranscriptome sequencing of cheese microbial community seems feasible to address the differential functionality aspects of the strains or to investigate the interactions between the members of the community during early stages of cheese production. On the other hand, RNA isolation methodology needs further improvement to adequately address the starter culture functionality during ripening stages.

Discovering the Interactions between Starter Culture Community Members

Understanding the evolutionary and ecological drivers that support and create the diversity and the final structure of the starter culture community during its long history of use is not an easy task. However, it may be feasible to describe what maintains the stability of the present complex community structure by studying the interactions between the species, strains or sub-populations in these microbial communities. The interactions among bacteria include variety of molecular and ecological interrelationships, like competition, mutualism, commensalism, ammensalism, and parasitism, depending on how the interactions affect the fitness of the organisms involved (32). The members of the community may respond to each other via trophic interactions in which they exchange metabolites or signalling molecules (33). The most well-known example for such trading in cheese microbial communities is the growth stimulation of protease-negative *L. lactis* cells by peptide and amino acid trading from protease-positive *L. lactis* cells during mixed growth in

milk (34). It was shown that protease-negative strains can invade the community very rapidly, and only a small fraction of protease-positive strains were needed to sustain the growth of entire community with the peptides supplied by protease-positive cells (35). This type of interaction was also described during the community analysis of the starter culture UR in chapter 3 of this thesis study. The starter culture UR was found to be dominated with protease-negative members of the genetic lineage 7 (78.2 %), while the protease positive cells of the genetic lineages 1&5 and 3 were only contributing to the 12,7 % of the community. Other known examples of interactions observed among the dairy micro-organisms are the production of bacteriocin by lactic acid bacteria (36), or the proto-cooperative trading of purines, long chain fatty acids and amino acids between *S. salivarius* subsp. *thermophilus* and *Lb. delbrueckii* subsp. *bulgaricus* in yoghurt fermentations (37). In the latter example, the formic acid, folic acid and long-chain fatty acids were shown to be provided by *S. salivarius* subsp. *thermophilus*, while the amino acids were supplied by the proteolytic activity of *Lb. delbrueckii* subsp. *bulgaricus* in the co-culture growth in milk (37).

Through metabolite trading, micro-organisms can also complement each other to better exploit the resources present in the niche and metatranscriptome sequencing is a powerful tool to discover these potential metabolite exchanges. Genome scale metabolic models can predict which reaction or products of a particular strain can be utilized by other members in the microbial community. Metatranscriptomic analyses of the communities may validate such interactions by mapping the expressed genes on the genomes of the community members. As an example, a potential trophic interaction for trading of GABA between *Leuc. mesenteroides* and *L. Lactis* was already proposed in chapter 3 by superimposing the genomic complements of the separate species. Transcriptome data could verify that GABA producing genes of *Leuc. mesenteroides* are expressed in parallel with GABA utilizing genes of *L. lactis* strains, which thereby could provide experimental support for such predicted metabolic interactions. Such transcriptome analyses could be executed in milk rather than cheese to avoid the complexity of obtaining reliable amounts and quality of RNA from the latter matrix. Moreover, in silico metabolic modelling of the genomes of the genetic lineages may propose additional metabolic dependencies or stimulatory capacities that govern the complex starter culture functioning. The GABA example given is readily extracted manually from the genome scale modelling of metabolic capacities, but refined flux-balance modelling using stoichiometry-based genome scale models may illustrate more refined interactions that remain undisclosed by 'simple' manual analyses of the metabolic capacities predicted on basis of the genomic sequences.

Importantly, metatranscriptome sequencing may also be useful to decipher the contribution of different members of the community to the overall flavour formation in cheese. The key flavour compounds, such as aromatic and branched chain aldehydes, acids, alcohols and sulphur compounds are mostly produced through the degradation of casein and casein-derived peptides, fat, lactose and citrate in milk with the enzymes that are expressed as part of the metabolic pathways of the starter bacteria (38). Mapping metatranscriptome sequences on the genome-predicted flavour metabolite formation pathways for the different community members could allow the assignment of the relative level of activity of particular pathways to specific lineages or strains in the community. Metabolome analysis of flavour metabolites can be combined with the metatranscriptomic analyses to connect the proposed pathway activity levels expressed by community members to the corresponding flavour compound profiles. Thereby, the coupled approaches of genome scale modelling in combination with metatranscriptome and metabolome analyses can improve our understanding of the flavour formation process during cheese production and open new avenues towards innovation strategies that aim to engineer the cheese starter cultures and the cheeses that they produce.

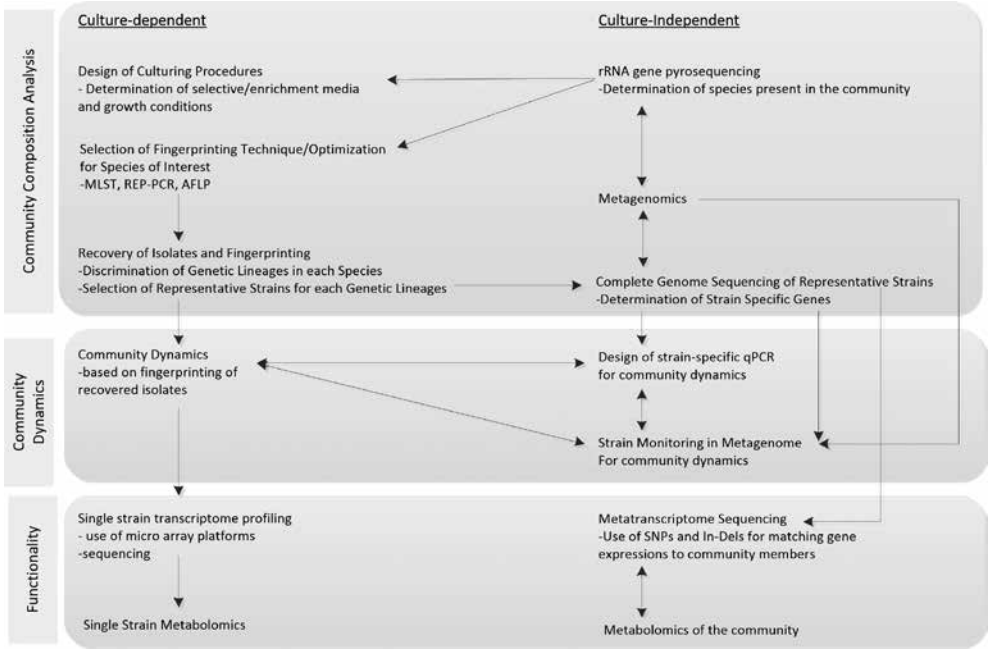


Figure 4. The schematic representation of the tools that can be applicable for the analysis of simple microbial communities.

Concluding Remarks

In this thesis, the community composition of a mixed Gouda cheese starter culture, the strain level dynamics of the community and its reflection on the functionalities during cheese manufacturing were investigated with several different approaches. The underlying molecular mechanism of the observed non-culturability for a specific member of the community was investigated at the molecular level. Furthermore, the back-slopping regimes applied for starter propagations were investigated as an alternative strategy to modulate the starter culture community composition and the corresponding flavour development in cheese. Several of the tools that were originally developed to study the microbial communities with high species diversity were further optimized or re-developed to reach the required resolution for studying the strain-level diversity and function present in the cheese starter culture community. The summary of these high-resolution tools has been compiled in Figure 4. The approaches followed in this thesis may pave the way to study other 'simple' microbial communities encountered in a variety of food fermentations, which can substantially contribute to our understanding of the overall fermentation processes at a molecular level, and can provide novel approaches to improve, control and steer the food-fermentations in desired directions.

References

1. **Delbes C, Ali-Mandjee L, Montel MC.** 2007. Monitoring bacterial communities in raw milk and cheese by culture-dependent and -independent 16S rRNA gene-based analyses. Appl Environ Microbiol **73:**1882-1891.
2. **Lodics T, Steenson, L. .** 1993. Phage–host interactions in commercial mixed-strain cultures: Practical significance, a review. Journal of Dairy Science **76:**2380–2391.
3. **Siezen RJ, Bayjanov JR, Felis GE, van der Sijde MR, Starrenburg M, Molenaar D, Wels M, van Hijum SA, van Hylckama Vlieg JE.** 2011. Genome-scale diversity and niche adaptation analysis of *Lactococcus lactis* by comparative genome hybridization using multi-strain arrays. Microb Biotechnol **4:**383-402.
4. **Bachmann H, Starrenburg MJ, Dijkstra A, Molenaar D, Kleerebezem M, Rademaker JL, van Hylckama Vlieg JE.** 2009. Regulatory phenotyping reveals important diversity within the species *Lactococcus lactis*. Appl Environ Microbiol **75:**5687-5694.
5. **Ayad EH, A. Verheul, C. De Jong, J. T. Wouters, and G. Smit.** 1999. Flavour forming abilities and amino acid requirements of *Lactococcus lactis* strains isolated from artisanal and non-dairy origin.. International Dairy Journal **9:**725-735.
6. **Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al.** 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res **23:**4407-4414.
7. **Kelly W, Ward L.** 2002. Genotypic vs. phenotypic biodiversity in *Lactococcus lactis*. Microbiology **148:**3332-3333.
8. **van den Bogert B, Erkus O, Boekhorst J, de Goffau M, Smid EJ, Zoetendal EG, Kleerebezem M.** 2013. Diversity of human small intestinal *Streptococcus* and *Veillonella* populations. FEMS microbiology ecology **85:**376-388.

9. **van den Bogert B.** 2013. Community and genomic analysis of the human small intestine microbiota. Wageningen University, Wageningne, The Netherlands.
10. **Nocker A, Cheung CY, Camper AK.** 2006. Comparison of propidium monoazide with ethidium monoazide for differentiation of live vs. dead bacteria by selective removal of DNA from dead cells. Journal of microbiological methods **67:**310-320.
11. **Kramer M, Obermajer N, Bogovic Matijasic B, Rogelj I, Kmetec V.** 2009. Quantification of live and dead probiotic bacteria in lyophilised product by real-time PCR and by flow cytometry. Applied microbiology and biotechnology **84:**1137-1147.
12. **Elizaquivel P, Aznar R, Sanchez G.** 2014. Recent developments in the use of viability dyes and quantitative PCR in the food microbiology field. Journal of applied microbiology **116:**1-13.
13. **Chapot-Chartier MP, Deniel C, Rousseau M, Vassal L, Gripon JC.** 1994. Autolysis of Two Strains of *Lactococcus lactis* during Cheese Ripening. International Dairy Journal **4:**251-269.
14. **Smit G, Smit BA, Engels WJ.** 2005. Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. FEMS microbiology reviews **29:**591-610.
15. **Juillard V, Laan H, Kunji ER, Jeronimus-Stratingh CM, Bruins AP, Konings WN.** 1995. The extracellular PI-type proteinase of *Lactococcus lactis* hydrolyzes beta-casein into more than one hundred different oligopeptides. Journal of bacteriology **177:**3472-3478.
16. **Exterkate FA.** 2006. Controlled permeabilization of *Lactococcus lactis* cells as a means to study and influence cheese ripening processes. International Dairy Journal **16:**788-796.
17. **Bachmann H, Kruijswijk Z, Molenaar D, Kleerebezem M, van Hylckama Vlieg JE.** 2009. A high-throughput cheese manufacturing model for effective cheese starter culture screening. J Dairy Sci **92:**5868-5882.
18. **Urbach G.** 1993. Relations between cheese flavour and chemical composition. International Dairy Journal **3:**389-422.
19. **Engels WJM.** 1997. Volatile and non-volatile compounds in ripened cheese: their formation and their contribution to favour. Wageningen Agricultural University, Wageningen, The Netherlands.
20. **Ayad EHE, Verheul, A., Engels, W.J.M., Wouters, J.T.M., Smit, G.** 2001. Enhanced flavour formation by combination of selected lactococci from industrial and artisanal origin with focus on completion of a metabolic pathway. Journal of applied microbiology **90:**59-67.
21. **Sanchez C, Neves AR, Cavalheiro J, dos Santos MM, Garcia-Quintans N, Lopez P, Santos H.** 2008. Contribution of citrate metabolism to the growth of *Lactococcus lactis* CRL264 at low pH. Appl Environ Microbiol **74:**1136-1144.
22. **Sanders JW, Leenhouts K, Burghoorn J, Brands JR, Venema G, Kok J.** 1998. A chloride-inducible acid resistance mechanism in *Lactococcus lactis* and its regulation. Molecular microbiology **27:**299-310.
23. **Small PL, Waterman SR.** 1998. Acid stress, anaerobiosis and gadCB: lessons from *Lactococcus lactis* and *Escherichia coli.* Trends in microbiology **6:**214-216.
24. **Rulikowska A, Kilcawley, K.N., Doolan, I.A., Alonso-Gomez, M., Nongonierma, A.B., Hannon, J.A., Wilkinson, M.G. .** 2013. The impact of reduced sodium chloride content on Cheddar cheese quality. International Dairy Journal **28:**45–55.
25. **Srivastava S, Ghosh, N., Pal, G.** 2013. Metagenomics: Mining Environmental Genomes, p. 161-189, Biotechnology for Environmental Management and Resource Recovery.
26. **Dijkstra AR, Setyawati MC, Bayjanov JR, Alkema W, van Hijum SA, Bron PA, Hugenholtz J.** 2014. Diversity in robustness of *Lactococcus lactis* strains during heat stress, oxidative stress, and spray drying stress. Appl Environ Microbiol **80:**603-611.
27. **van de Guchte M, Serror P, Chervaux C, Smokvina T, Ehrlich SD, Maguin E.** 2002. Stress responses in lactic acid bacteria. Antonie van Leeuwenhoek **82:**187-216.
28. **Bachmann H, de Wilt L, Kleerebezem M, van Hylckama Vlieg JE.** 2010. Time-resolved genetic responses of *Lactococcus lactis* to a dairy environment. Environmental microbiology **12:**1260-1270.
29. **Taibi A, Dabour N, Lamoureux M, Roy D, LaPointe G.** 2011. Comparative transcriptome analysis of *Lactococcus lactis* subsp. *cremoris* strains under conditions simulating Cheddar cheese manufacture. International journal of food microbiology **146:**263-275.

30. **Nam YD, Chang HW, Kim KH, Roh SW, Bae JW.** 2009. Metatranscriptome analysis of lactic acid bacteria during kimchi fermentation with genome-probing microarrays. International journal of food microbiology **130:**140-146.

31. **Weckx S, Allemeersch J, Van der Meulen R, Vrancken G, Huys G, Vandamme P, Van Hummelen P, De Vuyst L.** 2011. Metatranscriptome analysis for insight into whole-ecosystem gene expression during spontaneous wheat and spelt sourdough fermentations. Appl Environ Microbiol **77:**618-626.

32. **Sieuwerts S, de Bok FA, Hugenholtz J, van Hylckama Vlieg JE.** 2008. Unraveling microbial interactions in food fermentations: from classical to genomics approaches. Appl Environ Microbiol **74:**4997-5007.

33. **Smid EJ, Lacroix C.** 2013. Microbe-microbe interactions in mixed culture food fermentations. Current opinion in biotechnology **24:**148-154.

34. **Hugenholtz J, Splint R, Konings WN, Veldkamp H.** 1987. Selection of Protease-Positive and Protease-Negative Variants of *Streptococcus cremoris*. Appl Environ Microbiol **53:**309-314.

35. **Bachmann H, Molenaar D, Kleerebezem M, van Hylckama Vlieg JE.** 2011. High local substrate availability stabilizes a cooperative trait. The ISME journal **5:**929-932.

36. **Leroy F, De Winter, T., Moreno, M.R.F., De Vuyst, L.** 2007. The bacteriocin producer *Lactobacillus amylovorus* DCE 471 is a competitive starter culture for type II sourdough fermentations. . Journal of the Science of Food and Agriculture **87:**1726-1736.

37. **Sieuwerts S, Molenaar D, van Hijum SA, Beerthuyzen M, Stevens MJ, Janssen PW, Ingham CJ, de Bok FA, de Vos WM, van Hylckama Vlieg JE.** 2010. Mixed-culture transcriptome analysis reveals the molecular basis of mixed-culture growth in *Streptococcus thermophilus* and *Lactobacillus bulgaricus*. Appl Environ Microbiol **76:**7775-7784.

38. **Steele J, Broadbent J, Kok J.** 2013. Perspectives on the contribution of lactic acid bacteria to cheese flavor development. Current opinion in biotechnology **24:**135-141.

Summary

SUMMARY

Lactic acid bacteria (LAB) are used as starter and adjunct cultures for the production of many artisanal and industrial fermented foods with plant and animal origin all over the world, such as wine, yoghurt, cheese, kefir, kimchi. Artisanal fermentations have long been propagated with the transfer of an inoculum from old batch of fermented food to the new batch (back-slopping) to initiate the fermentation with the activity of the indigenous microbiota present in the inoculum. In contrast, the industrial production of these traditional fermented foods needs consistent and controlled processes in order to have safe and standard final product quality. Therefore, most of them have been standardized under controlled process parameters in industrial production, and the inocula with indigenous microbiota are replaced with the starter cultures that contain lower numbers of LAB species. Cheese manufacturing is still performed in both artisanal ways, as well as with the use of starter cultures. Gouda cheese starter cultures constitute several strains from the subspecies of *Lactococcus lactis* and *Leuconostocs mesenteroides* in different combinations, and several different types of starter cultures are present depending on how well the composition of the culture is known, the type of the subspecies present in the culture, and the propagation environment used for starter production. The mixed and undefined cheese starter cultures may harbour variable number of strains that contribute unique functionalities to the cheese manufacturing process. Therefore, understanding, controlling and predicting the cheese manufacturing processes require the determination of strain level diversity in the starter culture, their collective and specific metabolic complement, and their activity throughout the cheese manufacturing process, including the interactions between the strains. In this thesis, a mixed and undefined Gouda-type cheese starter culture was studied using several different approaches, aiming to answer some of these research questions that are summarized below.

Mixed cheese starter cultures can be considered as simple microbial communities with limited species, but high strain diversity. There are plenty of tools available in the literature that enable to discriminate the micro-organisms depending on the resolution power required. While 16S rRNA gene sequencing and multi-locus sequence typing (MLST) are considered powerful for the discrimination at the species and subspecies level, the techniques like repetitive polymerase chain reaction (rep-PCR), randomly amplified polymorphic DNA (RAPD), and amplified fragment length polymorphism (AFLP) are more useful for the discrimination of strains since

they use the polymorphisms distributed among whole genome content as the basis of comparison. In this thesis, a high-throughput AFLP technique was developed for the delineation of the closely related strains recovered during the culture-dependent analysis of a mixed starter culture. The number of polymorphisms that can be detected with AFLP was maximized with the proper selection of the method's variables, such as restriction enzyme and selective nucleotides, with in silico and slab-gel analyses of closely related *L. lactis* strains, and the technique was implemented to the high-throughput capillary electrophoresis using the flourophore labelled primers. The developed AFLP technique was validated using a set of *L. lactis* strains with dairy and non-dairy origin, which had already been delineated with several other fingerprinting techniques to verify the improvement of the resolution power. AFLP was found to deliver much higher resolution than the techniques 16 rRNA gene pyrosequencing, MLST, and REP-PCR, and produced clustering topology very similar to the clustering topology obtained by array-based comparative genome hybridization profiles. Furthermore, AFLP fingerprints generated phenotype-specific marker fragments that allowed the discrimination of the strains with *cremoris* genotype into two subsets based on their typical *cremoris*- and *lactis*-phenotypes.

Subsequent to the development of tools for diversity analysis, the community composition of an undefined and mixed Gouda cheese starter was studied with culture dependent and independent approaches. Isolates of *L. lactis* and *Leuconostocs mesenteroides* were recovered from the starter culture and from cheese samples collected at different time points during cheese manufacturing and ripening, using three different complementary cultivation media. The AFLP methodology was applied for fingerprinting of the isolates recovered from the starter culture, and in total 8 genetic lineages were identified. In addition, individual isolates representing each of the genetic lineages were investigated for their plasmid content and phage sensitivity, which indicated a certain extent of plasmid content variation but a very large variation in phage resistance levels among the strains representing each genetic lineage. This finding supports the requirement of a multilevel approach for the composition analysis of simple microbial communities. Additionally, fingerprinting of the isolates recovered from cheese samples collected at different time points during the cheese making process, allowed to monitor the dynamics of the genetic lineages and revealed variable growth and survival patterns for different genetic lineages during cheese manufacturing. The AFLP clustering also formed the basis for the selection of representative strains for each of the genetic lineage for

complete genome sequencing, allowing the comparison of the genome complement of the genetic lineages. In the culture independent approach, the metagenome of the starter culture was determined by total DNA sequencing and revealed the presence of 99% of *Lactococcus lactis* and 1% of *Leuconostoc mesenteroides* strains in the starter culture, which was in good agreement with the relative abundance of the two species deduced from culture dependent analysis. Furthermore, determining the relative abundances of stain-specific orthologous groups in the metagenome sequences allowed the quantification of genetic lineages in metagenome. Moreover, strain-specific orthologous groups allowed the development of genetic lineage-specific quantitative PCR assays to detect and quantify the genetic lineages and their relative abundance using total DNA recovered from the original starter culture or from cheese samples.

Metagenomic analysis of total DNA isolated at different time points of cheese manufacturing nor genetic lineage-specific qPCR revealed significant changes in community composition over time during the initial stages and the ripening period, which was in clear contrast with the community dynamics deduced from culture-based approaches. Moreover, the total DNA quantities isolated at the different time points during the overall cheese manufacturing process were very similar throughout, which implied that DNA from deteriorating, and dying bacterial cells consistently was isolated at each of the time-points even though the cells lost their capacity to form a colony. In order to circumvent this discrepancy between viable cells and their molecular detection, propidium monoazide (PMA) treatment was incorporated into the total DNA extraction protocol, which renders the DNA derived from membrane-compromised or dead cells incompatible with PCR-amplification. The use of PMA was evaluated with genetic-lineage specific qPCR analysis and metagenomics on pure DNA samples of *L. lactis*; on damaged, dead, and live cell cultures; as well as on cheese samples collected at several time points during cheese manufacturing and ripening. PMA inhibited the amplification of DNA derived from membrane-compromised cells approximately 100-fold and allowed the selective analysis of the viable population in cheese. Furthermore, a two-step PMA treatment and whole genome amplification procedure allowed the selective metagenome sequencing of the viable population in cheese.

The classical culture-dependent analysis of microbial communities bears intrinsic biases due to the differences in the growth requirements of the community members, and complementary plating with several media is required to recover the majority

of the community members. In the culture dependent analysis of the cheese starter culture in this thesis, the recovery of *L. lactis* populations were also achieved with complementary plating on LM17 and Reddy's agar media. Interestingly, 5 times more cells, all of which were belonged to genetic lineage 7, were exclusively recovered on Reddy's agar medium due to presence of arginine in Reddy's agar medium when compared to LM17. The molecular basis of the arginine stimulation was investigated on *L. lactis* TIFN7, which is the representative strain for genetic-lineage 7, by transcriptome analysis before and after arginine addition in arginine-limited chemostat cultures. Transcription analysis indicated very low expression levels on arginine, pyrimidine and purine biosynthesis pathways and high expression levels on single carbon C-1 metabolism and $C0_2$ production reactions for TIFN7 under arginine limiting conditions. Following arginine supplementation, the arginine and $C0_2$ limitation in TIFN7 was relieved and growth was enhanced with the redirection of resources towards the pyrimidine and purine biosynthesis, indicating that growth of TIFN7 is mainly limited by sufficient C-1 resources for the production of nucleotides to sustain maximal growth rates and can be enhanced by extracellular C1-supplies. The plate count anomaly observed in this example of distinctive growth requirements of the *L. lactis* strains within a starter culture indicates the necessity to employ complementary plating as well as culture-independent analyses to investigate the complexity of microbial communities.

One of the main reasons why an undefined cheese starter culture was studied in the thesis is that undefined, mixed starter cultures are still frequently preferred in dairy industry due to their robustness under phage attack, which is also exemplified by the findings in this thesis. Compared to the defined starter cultures, which can be engineered by addition or removal of some of the strains from the combination used in the starter, the compositions of the undefined mixed cultures are not easy to steer. They are propagated with strict back-slopping regimes in a traditional way, and modifications in the propagation regimes are the only applicable approach to steer the community compositions without loss of diversity and robustness. In order to test this strategy, and to evaluate its potential as an engineering tool for undefined cultures, the mixed starter culture was serially propagated from different growth phases in milk. In the end of long term serial propagations from exponential and stationary phases, the final community compositions and functionalities of cultures were found to be strongly dependent on the back-slopping regime applied. The distinctly propagated starter cultures were applied in mini-cheese manufacturing and the flavour profiles detected in 12-weeks ripened mini-cheeses were distinct and

very specific for the initial starter composition employed, which was determined by the propagation regime used. Therefore, alternative back-slopping regimes hold great potential for the steering of microbial community compositions, and can generate starter cultures with altered functionality during cheese manufacturing.

In conclusion, this thesis describes the study of a mixed and undefined Gouda cheese starter culture in terms of its community composition, dynamics and functionalities using culture dependent and independent approaches. Many tools have been developed, or optimized to increase the detection resolution to follow the strain-level dynamics and functions in the starter culture during its propagation or during cheese manufacturing. The approaches taken in this thesis for the investigation of the microbial community of a complex Gouda-cheese starter culture can be taken as an example for the development of analogous strategies for other 'simple' microbial communities involved in food fermentations.

Samenvatting

SAMENVATTING

Starter- en adjunct-culturen van melkzuurbacteriën worden wereldwijd gebruikt voor de ambachtelijke of industriële productie van gefermenteerde voedselproducten uit plantaardige of dierlijke grondstoffen, zoals bijvoorbeeld wijn, yoghurt, kaas, kefir of kimchi. Ambachtelijke fermentatie processen worden traditioneel vaak gestart met een ent uit een voorgaande batch gefermenteerd voedsel (back slopping) en de daarin van nature voorkomende micro-organismen. Bij industriële productie van deze traditionele gefermenteerde voedingsmiddelen zijn consistentie van productkwaliteit en productveiligheid essentieel. De gewenste standaardisering van de productie wordt bereikt door zowel nauwkeurige controle op de procesparameters als het vervangen van de natuurlijke ent door startercultures met een beperkte variëteit aan soorten melkzuurbacteriën. Bij het maken van kaas worden zowel startercultures als de traditionele methodes gebruikt. Startercultures voor Goudse kaas zijn mengsels van diverse stammen van de ondersoorten *Lactococcus lactis* en *Leuconostocs mesenteroides*, die onderling niet alleen verschillen in biologische samenstelling maar ook in de methode van propagatie en mate waarin de exacte samenstelling bekend is. Zulke ongedefinieerde startercultuur mengsels bevatten talloze stammen die bijdragen aan de functionaliteit tijdens kaasproductie. Om kaasproductie te kunnen begrijpen, controleren en voorspellen is het belangrijk om de diversiteit van de startercultuur op stamniveau en de collectieve- en individuele bijdrage van de stammen tijdens de volledige duur van het productieproces te bepalen. Hierbij spelen ook de interacties tussen de verschillende melkzuurbacteriestammen een belangrijke rol. In dit proefschrift wordt een ongedefinieerd startercultuurmengsel voor de productie van Goudse kaas bestudeerd aan de hand van verscheidene technieken en benaderingen, met als doel een aantal, hieronder samengevatte, onderzoeksvragen te beantwoorden.

Startercultuurmengsels zijn eenvoudige microbiële systemen met een beperkt aantal soorten, maar met een grote verscheidenheid aan stammen binnen die beperkte hoeveelheid soorten. Er bestaan vele genetische methodes om micro-organismen van elkaar te onderscheiden met een methode-afhankelijk niveau van resolutie. 16S rRNA gen-sequencing of multi-locus sequence typing (MLST) zijn goed in het onderscheiden van micro-organismen op het niveau van soort en ondersoort. Andere technieken, zoals repetitive polymerase chain reaction (rep-PCR), randomly amplified polymorphic DNA (RAPD), en amplified fragment length polymorphism (AFLP) zijn echter beter in het onderscheiden van stammen, omdat daarbij de over

het gehele genoom verspreide genetische verschillen gebruikt kunnen worden. Dit proefschrift beschrijft de ontwikkeling van een high-throughput AFLP methode voor het onderscheiden van de zeer verwante stammen tijdens de cultivatie-afhankelijke analyse van startercultuurmengsels. Om zoveel mogelijk genetische verschillen te identificeren, werd de AFLP methode geoptimaliseerd door middel van de keuze van variabelen, zoals de restrictie enzymen en selectieve nucleotiden, met zowel *in silico* en slab-gel analyse van nauw verwante *L. lactis* stammen. De procedure werd verder geoptimaliseerd door gebruik te maken van fluorofoor-gelabelde primers en high-throughput capillary electrophoresis. Om de verbetering van de resolutie van de ontwikkelde AFLP methode te valideren werd deze getest op een aantal *L. lactis* stammen, waarvan de verwantschap al grotendeels bekend was uit voorgaande studies met andere methoden. Hieruit bleek dat de resolutie van de geoptimaliseerde AFLP methode veel hoger was dan van 16 rRNA gene pyrosequencing, MLST, of REP-PCR. Bovendien hadden de resulterende clusters grote gelijkenis met eerdere op array-gebaseerde vergelijkende genoom hybridisatie profielen. Daarnaast werden met de AFLP methode fenotype-specifieke markers verkregen, waardoor de stammen met een *cremoris* genotype onderverdeeld konden worden in twee groepen gebaseerd op typische *cremoris*- en *lactis*-fenotypes.

De ontwikkelde methode om diversiteit op stamniveau te analyseren werd vervolgens gebruikt om de samenstelling van een ongedefinieerd startercultuurmengsel voor de productie van Goudse kaas te bestuderen met zowel een cultivatie-afhankelijke als een cultivatie-onafhankelijke aanpak. Tijdens de productie en rijping van de kaas werden met behulp van drie complementaire cultivatie methodes stammen van *L. lactis* en *Leuconostocs mesenteroides* gevonden. Met behulp van de AFLP methode werden deze isolaten vervolgens ingedeeld in 8 verschillende genetische lijnen. Daarnaast werd ook de plasmide samenstelling en de gevoeligheid voor fagen van de verschillende isolaten bestudeerd. Hieruit bleek dat naast verschillen in plasmide samenstelling, vooral de resistentie tegen fagen zeer verschillend was voor de stammen binnen de verschillende genetische lijnen. Deze observaties onderstrepen de noodzaak van het analyseren van de samenstelling van de microbiële populatie op de verschillende niveaus. Bovendien maakt deze methode het mogelijk om de tijdsafhankelijke verandering in verschillende genetische lijnen te volgen tijdens de kaasproductie. Hiermee werd aangetoond dat de patronen van groei en overleving verschillen tussen de verschillende genetische lijnen. Op basis van deze op AFLP gebaseerde indeling werd uit iedere genetische lijn een representatieve stam gekozen waarvan vervolgens de gehele genoom sequentie bepaald werd.

De genetische informatie van de 8 genetische lijnen werd vervolgens vergeleken. Daarnaast werd ook het metagenoom van de startercultuur bepaald om op een cultivatie onafhankelijke manier informatie over de samenstelling te verkrijgen. Uit de DNA metagenoomsequentie blijkt dat 99% van de cultuur bestaat uit *Lactococcus lactis* stammen en voor 1% uit *Leuconostoc mesenteroides*. Dit komt overeen met de relatieve frequentie waarmee de twee soorten voorkwamen in de bovenstaande cultivatie afhankelijke methode. Door te kijken naar de stam-specifieke orthologe genen in de data van deze metagenoom analyse kon ook de relatieve fractie van de verschillende genetische lijnen in de startercultuur bepaald worden. Deze stam-specifieke orthologe genen vormen ook de basis voor een lijn-specifiek quantitatieve PCR methode (qPCR) om de verschillende genetische lijnen de detecteren en te kwantificeren op basis van totaal-DNA extracten uit de oorspronkelijke startercultures of uit kaasmonsters.

Hoewel cultivatie afhankelijke methodes duidelijk tijdsafhankelijke verschillen in de populatie micro-organismen lieten zien tijdens de kaasproductie en –rijping, waren deze variaties niet aantoonbaar met zowel de analyse van het metagenoom van de verschillende tijdspunten als de stam specifieke qPCR analyse methode. Daarnaast was de totale hoeveelheid DNA die op ieder tijdpunt geïsoleerd werd nagenoeg gelijk, wat mogelijk betekent dat ook DNA van beschadigde of zelfs dode bacteriën geïsoleerd wordt, terwijl deze in cultivatie afhankelijke methodes niet worden geïsoleerd als kolonie. Om ook met de moleculaire methodes specifiek de levende cellen te kunnen analyseren, werd een propidium monoazide (PMA) behandeling toegevoegd aan het extractie protocol van het DNA. Door deze behandeling is DNA uit cellen waarvan de membranen beschadigd zijn niet meer beschikbaar voor de PCR reacties. Deze methode werd eerst getest met de specifieke qPCR analyse en in de metagenoom bepaling op DNA verkregen van zowel beschadigde, dode als levende cultures van *L. Lactis*. Daarnaast werd de methode getest op monsters tijdens de productie en rijping van kaas. Doordat na behandeling met PMA de amplificatie van DNA uit cellen met een beschadigd membraan ongeveer met een factor 100 was verminderd, kon selectief gekeken worden naar de levende populatie micro-organismen in de kaasmonsters. Ook kon met behulp van een twee-staps PMA behandeling en een whole genoom amplificatie procedure het metagenoom van de levende populatie in kaas bepaald worden.

De uitkomst van traditionele cultivatie-afhankelijke analyses van microbiële populaties wordt beïnvloed door de verschillende benodigde groeifactoren van

de micro-organismen, waardoor uitplaten op verschillende complementaire media nodig is om de meerderheid van de voorkomende micro-organismen te identificeren. Ook tijdens de cultivatie-afhankelijke analyse van de *L. Lactis* stammen van de kaas startercultuur in dit proefschrift is gebruik gemaakt van uitplaten op de complementaire media LM17 en Reddy's agar. In vergelijking met de LM17 platen, werden op de platen met Reddy's agar 5 keer meer kolonies waargenomen. Deze cellen, die uitsluitend op Reddy's agar platen konden groeien door de aanwezigheid van arginine, behoorden allemaal tot de 7^{e} genetische lijn uit de bovenstaande isolaties. Om deze stimulering door de aanwezigheid van arginine verder te bestuderen, werd een vergelijkende transcriptoom studie met stam *L. Lactis* TIFN7 gedaan voor- en na toevoeging van extra arginine aan arginine-gelimiteerde chemostaatcultures. Deze studie liet zien dat de transcript niveaus van de genen die coderen voor de biosynthetische routes van arginine, pyrimidine en purine erg laag waren in TIFN7. Tegelijkertijd waren de mRNA niveaus voor het C1-metabolisme en CO_2 producerende reacties hoog onder de arginine-gelimiteerde condities. Na toevoeging van overmaat arginine nam de groei toe door herverdeling van grondstoffen naar de pyrimidine en purine biosynthese, wat aangeeft dat groei van TIFN7 beperkt wordt door de beschikbaarheid van C1-grondstoffen. Dit voorbeeld van specifieke groeifactor afhankelijkheid van de *L. Lactis* stammen in een startercultuur laat de noodzaak zien van zowel het uitplaten op verschillende complementaire media als van cultivatie-onafhankelijke analyse methodes voor onderzoek aan complexe microbiële gemeenschappen.

Vanwege de robuustheid tegen faaginfecties worden ongedefinieerde, gemengde startercultures nog vaak toegepast in de zuivelindustrie. Dit is ook een van de hoofdredenen waarom deze startercultures in dit proefschrift bestudeerd werden. Echter, in vergelijking met gedefinieerde startercultures, die door middel van toevoeging of verwijdering van specifieke stammen aangepast kunnen worden, is de samenstelling van zulke ongedefinieerde, gemengde startercultures moeilijk te beïnvloeden. Doordat ze alleen vermenigvuldigd worden door het doorzetten van de oorspronkelijke cultures (back slopping), is het aanpassen van de propagatie procedures de enige manier om de samenstelling te beïnvloeden zonder verlies aan diversiteit en robuustheid. Om te onderzoeken of dit inderdaad een goede methode is om de samenstelling van ongedefinieerde cultures te beïnvloeden, werd een startercultuur onderworpen aan seriële propagatie vanuit verschillende groeifases in melk. Het bleek dat de samenstelling van de microbiële gemeenschap, en daarmee ook de functionaliteit van de cultuur, sterk bepaald werd door of doorenten uit

de exponentiele fase of uit de stationaire fase. De op deze verschillende wijze gepropageerde startercultures werden vervolgens getest voor de productie van mini-kazen. De smaakprofielen van de 12-weken gerijpte kaas bleken sterk afhankelijk van de gebruikte propagatie procedure voor de startercultuur. Uit deze resultaten blijkt dat het aanpassen van de propagatie procedures inderdaad gebruikt kan worden om de functionaliteit van ongedefinieerde, gemengde startercultures voor kaasproductie te beïnvloeden.

Dit proefschrift beschrijft het bestuderen van de microbiële samenstelling, de dynamiek en de functionaliteit van een ongedefinieerde, gemengde startercultuur voor de productie van Goudse kaas door middel van zowel cultivatie-afhankelijke als –onafhankelijke methodes. Hierbij zijn diverse gereedschappen ontwikkeld of geoptimaliseerd om de dynamiek op (levensvatbaar) stam- en functieniveau met de gewenste resolutie te kunnen volgen tijdens het kaasproductieproces. De aanpak die in dit proefschrift is toegepast op een startercultuur voor Goudse kaas productie, kan ook gebruikt worden voor de ontwikkeling van analoge strategieën om andere microbiële gemeenschappen in voedselfermentaties te bestuderen.

Acknowledgements

ACKNOWLEDGEMENTS

How I experienced and who accompanied me on this long journey matters as much as much as my work and hereby I would like acknowledge the people who were with me in this period!

First of all, I would like to show my deepest gratitude to my supervisors Michiel Kleerebezem and Eddy Smid. Dear Michiel, your work and publications had been the source of my motivation when seeking the PhD opportunities in Netherlands, the homeland of LAB research. I still remember the cheer I had when I was accepted for being part of Top Institute Food and Nutrition team. Thank you so much for your guidance and help starting from the very first day of my PhD until the end. You have always been motivating, encouraging and path opening, which I definitely acknowledged in the up-and-downs of my PhD period. Dear Eddy, I met you when my previous daily supervisor Johan left the project, and everybody told me how lucky I was to have the opportunity to work with you. Now, after sharing 4 years of research experience with you, I truly understand what they really meant. You have been part of my common sense and I always appreciated your solid and grounded support on my studies. I am very happy to get to know your wise and authentic personality.

During my study, I have been part of several different teams due to the integrated structure of TIFN. I had my research activities at NIZO Food Research, and I was part of Laboratory of Microbiology mostly for my education activities. TIFN had been the consortium that brings these two teams as well as our industrial partners together. I am grateful to each member of my TIFN team, including Patrick, Ingrid, Lucie, Maria, Tom, Milkha, Renske, Raditya, Irma, Maciek, Sacha, Victor, Roland, Michiel Wells, with whom I experienced the strength of teaming up for the excellence. Patrick, you have been my right hand in the lab. It was pleasure to work close to you and to share the office. Now, I am happy to see you on my side as my paranymph on this important day of my life as well. Ingrid, I am always impressed with your high quality technical skills and accuracy, and have learned a lot from you. Thank you very much for your full support in the molecular works. Renske, Raditya, Irma and Maciek, you all have footsteps in the collection of the data. Thanks a lot for all your efforts in different chapters of this thesis work. Sacha, Victor and Michiel Wells, the so called ``Bioinformatics guys``, I am grateful for your support and collaboration on the analysis of whole genome sequences and metagenomes, microarray designs

and statistical works. Tom, it was pleasure to work with you on typing of your Streptococci and to show the world how cheese and gut can find a common ground ;-). Maria, my lovely and the far most authentic friend! I met you in TIFN meetings, but it feels like I knew you from the beginning of my life. Having similar stories in different sides of the Aegean Sea, you became like a sister to me. I will always remember the nice memories of our conversations, tea parties and dinners. I wish to have you around for many more years.

Furthermore, I would like to acknowledge my colleges and friends at NIZO that were not part of my TIFN team; Kerstin, Iris, Roger, Anne, Sven, Onur, Arno and Carol. You were not only a college to me, but also a nice friend that I always enjoyed to have around. Kerstin, my sincere friend, you had been my tennis partner and a life couch ☺, companying me in my joyful and difficult times. Thank you so much. Iris, I always remember you with your cheerful laughs and openness to share. I pity that we did not share the same corridor. Roger, you are a life-saving man in RNA work and I very much enjoyed your sense of humour. Anne, fortunately I don't need to check the back from my shoulder anymore to see if you are in the preparation of a nasty joke ;-). You always helped me out when I run out of a chemical or an enzyme in the middle of experiments, many thanks… Sven, the German superman, he can open a pressurized jar of eppies with his big muscles in one go! You are one of the most interesting men that I ever met. I will always remember you with your cup of coffee in the ice box, milk boxes on the window and sausages in the drawer ☺ and I am very impressed with your artistic gift in photography. Arno, I always enjoyed the questions of your critical mind and talking with you on many interesting topics you can come up with. Carol, I am glad that we met at NIZO. You were a very sweet roommate. It was fun to share the house, to cook, to party and to have long talks on life together. Thank you guys for all these nice memories I collected! Additionally, I acknowledge Herwig Bachmann, Eva Dusseldorf, Wim Engels, Thom Huppertz, Zeger Kruiswijk, Igor Mierau, Bart Smit, Arjen Wagendorp, Wilma Wesselink for their both direct and indirect involvements, discussions, supports and guidance during my thesis study.

My dear friends at Microbiology, it was pleasure to meet you! Janneke, Gerben and Leo, it was fun to share the office and to have inspiring talks on science and life. Rozelin, my pinky and the first friend in Wageningen, it was a lot of fun to spend time together with you in Wageningen and USA. Derya, my crazy mate in practical courses, I wish you keep your big smile forever! Martin, I enjoyed every bit of our

deep philosophical talks on life… Floor, Kal, Nicolas, Thomas, Mauricio, Juan A, Pierpaolo, Teun, Peer, Sebastian, Jing and all the others that I couldn`t mention here… I enjoyed very much to have chats with you during lunch times and around the lab activities. It was a memorable experience to travel China and Japan together.

I am also very thankful to my Turkish friends in Wageningen, who walked with me in this journey far away from home. Dilek, Yunus, Yusuf, Onur, Sami, Mustafa, Sevinc, Morteza, Aslihan, Ceren, Ebru, Basak, Yonter, Robin, Samet ve gelip giden ama gonlumuzde birer parcalarini birakan Erasmus guzelleri… Sizleri tanimak benim icin buyuk bir keyif. Yol arkadasim oldunuz bu uzak diyar dusunde. Raki masalarinda gurbeti ve ozlemi birlikte harmanladik, yeni deneyimlere birlikte yol aldik heyecanla. Dilek, varligini hep yanimda hissettigim, yani basimda beni yureklendirmeye devam eden sevgili paranymphim. ``Umudumuza da acilarimiza da yol ver hayatta yerini bulsun`` demeyi birlikte deneyimledik! Dilerim, daha uzun yillar birlikte buyur ve ozgurlesiriz bu dipsiz kuyuda. Yunus, aklimdaki sorulara, isyanlara ve aydinliga eslik eden, saatlerce sikilmadan konusabildigim, ayni dili paylastigim guzel insan... Varligin benim icin cok degerli! Yusuf, tanidigim en buyuk kalpli dunya vatandasi… Onur, sag duyusuna ve mantigina guvendigim, NIZO daki yol arkadasim.. Sami, acik sozlulugu ve direktligi ile Avrupa standardlarina en cok uyum saglayan guzel kalpli kardesim… Mustafa, izmir meltemini ve nesesini; Sevinc, ODTU anilarimi keyifle paylastigim hemserilerim… Morteza, cografyalar degisse de kulturun insanlari birbirine bagladigina yegane ornek… Aslihan ve Ceren, son yillarda hayatima giren ve muhabbetinden keyif aldigim guzel dostlarim. Ve ismini tek tek anamadigim butun diger arkadaslarim… Aklimda ve gomlumde yer etmis butun bu guzel anilar icin cok tesekkur ederim. Iyi ki varsiniz…

Her zaman yanimda hissettigim sevgili aileme de tesekkurlerimi ifade etmek istiyorum. Sizlerin destegi ve sevgisi, doktorami tamamlayabilmek icin motivasyon kaynagi oldu bana. Annecigim, iyi ve kotu her animda kosulsuz yanimda oldugun ve sevdigin icin, kendiligimi yasamamda ve gerceklestirmemde beni yureklendirdigin icin sana minnettarim. Oz saygim ve sevgim senin eserindir. Abicigim, sevgini ve korumaci gudulerini hep uzerimde hissettim. Hayatima entellektuel beyninle ve naïf ruhunla renk kattigin icin sana cok tesekkur ederim. Biricik kuzenim Isil, ikinci kardesim benim… Zor anlarimda da, mutlu gunlerimde de varliginla bana guc ve umut verdin. Benim icin cok degerlisin. Hayatimin en buyuk kahramani dedecigim… Biyolojiye olan tutkum birlikte karistirdigimiz biyoloji kitaplariyla, ciftlikte seninle gecirdigim guzel anilarla basladi. Nur icinde uyu, seni sevgiyle aniyorum!

Turkiyedeki buyuk ailemin icinde bana uzaklardan destegini hic esirgemeyen en yakin dostlarim da dahil! Burcu, yillara meydan okuyan dostlugumuz dilerim gunden gune daha da guclenir. Cisem ve Zelal, bir fincan kahveninin gercekten 40 yil hatiri varmis kizlar. Bunu da deneyimledik bu surecte. Ergul ve Halil, ODTUde baslayan ve yillandikca daha da guzellesen dostlugumuz beni hep keyiflendirdi. Iyi ki varsiniz. Sevgili hocalarim, Sebnem Harsa ve Haluk Hamamci, laktik asit bakterilerine olan merakim sizlerin bana sagladigi kaynaklar ve projeler icinde sekillendi. Turkiyedeki akademik calismalarim icinde benim icin hem bir yol gosterici oldunuz, hem de ailemden biri haline geldiniz. Guveniniz ve desteginiz icin cok tesekkur ederim.

Hayat arkadasim, birlikte buyudugum yol arkadasim Baris… Hayatimin en ozel anlari sana ait. Duslerimi gerceklestirmemde beni yureklendirdin, yanimda durdun, destegini hic bir zaman esirgemedin. Yollarimiz ayrilmis olsa da bil ki sana olan sevgim icimde bana mutluluk vermeye devam ediyor! Aci ve tatli tum yasanmisliklarimiz, ve birlikte ogrendigimiz hersey onumu aydinlatiyor, bana yol gosteriyor. Varligin ve destegin icin tesekkur ederim…

My dear Ton… You have been my home in Netherlands… I had very difficult times in finding peace with myself in the past periods but you fully embraced me with your whole heart! I wouldn`t make the end of this long journey without your support and love. Thank you so much dear!

Curriculum Vitae

Oylum Erkus was born on 1st of January 1981 in Izmir, Turkey. She attended Middle East Technical University in 1998 (Ankara, Turkey) and graduated with a BSc diploma in Food Engineering in 2003. She obtained her degree of Master of Science in Food Engineering from Izmir Institute of Technology (Izmir, Turkey) in 2007. In her major thesis, she studied on the isolation and molecular typing of yoghurt starter bacteria obtained from artisanal yoghurts of Anatolia region under the project called ``Production of Cheese and Yoghurt Starter Cultures and Lactase Enzyme for the Dairy Industry: Traditional and Modern Solutions against Lactose Intolerance`` financed by Governmental Planning Institute. In addition, she worked in a minor project, investigating the starter functions of Lactobacilli strains to formulate industrial cheese starter cultures.

In June 2008, she was appointed as a PhD student within the project ``Complex Fermentations`` funded by Top Institute Food and Nutrition to make research on the community dynamics of Gouda cheese starter cultures. The PhD project was carried out at NIZO Food Research and Laboratory of Microbiology at Wageningen University under supervision of Prof. Dr. Michiel Kleerebezem and Assoc. Prof. Dr. Eddy J. Smid.

List of Publications

Erkus Kutahya O, Starrenburg MJ, Rademaker JL, Klaassen CH, van Hylckama Vlieg JE, Smid EJ. 2011. High-resolution amplified fragment length polymorphism typing of *Lactococcus lactis* strains enables identification of genetic markers for subspecies-related phenotypes. *Appl Environ Microbiol* **77:** 5192-5198.

Erkus O, de Jager VC, Spus M, van Alen-Boerrigter IJ, van Rijswijck IM, Hazelwood L, Janssen PW, van Hijum SA, Kleerebezem M, Smid EJ. 2013. Multifactorial diversity sustains microbial community stability. *The ISME journal* **7:**2126-2136.

Erkus O, Okuklu B, Yenidunya AF, Harsa S. 2013. High Genetic and Phenotypic Variability of *Streptococcus thermophilus* Strains Isolated from Artisanal Yuruk Yoghurts. *LWT-Food Science and Technology.*Published online ahead (*http://dx.doi.org/10.1016/j.lwt.2013.03.007).*

Bogert B, **Erkus O,** Boekhorst J, Goffau M, Smid EJ, Zoetendal EG, Kleerebezem M. 2013. Diversity of human small intestinal *Streptococcus* and *Veillonella* populations. FEMS microbiology ecology. **85**: 376-388.

Karahan AG, Başyiğit Kılıç G, Kart A, Şanlıdere Aloğlu H, Öner Z, Aydemir S, **Erkuş O,** Harsa Ş. 2010. Genotypic identification of some lactic acid bacteria by amplified fragment length polymorphism analysis and investigation of their potential usage as starter culture combinations in Beyaz cheese manufacture. *Journal of Dairy Science* **93**:1-11.

Erkus O, de Jager VCL, Geene RTCM, Van Alen-Boerigter I, Hazelwood L, Van Hijum SAFT, Kleerebezem M, Smid EJ. Use of Propidium Monoazide for the Selective Profiling of Live Microbial Populations during Cheese Ripening. Submitted for publication.

Smid EJ, **Erkus O**, Spus M, Wolkers-Rooijackers JCM, Alexeeva S, Kleerebezem M. Complex starter cultures for cheese manufacturing. Submitted for publication.

Perdana J, Aguirre Z, **Kutahya O**, Schutyser M, Fox M. Up-scaling of spray drying of *Lactobacillus plantarum* WCFS1 guided by predictive modelling. Submitted for publication.

Erkuş O, Janssen PWM, Van Alen-Boerigter I, Wels M, Bongers R, Kleerebezem M, Smid EJ. Arginine Stimulated Culturability of *Lactococcus lactis* Cells Limited for C1-Resourses. Manuscript in preparation.

Erkuş O, Geene RTCM, Janssen PWM, Kleerebezem M, Smid EJ. Starter Culture Composition and Its Flavor Formation Capacity as a Function of Propagation Regimes. Manuscript in preparation.

Overview of Completed Training Activities

Discipline Specific Activities

Courses

Systems Biology: statistical analysis of ~omics data, 2010, Wageningen, NL
Advanced Course on Microbial Physiology & Fermentation Technology, 2011, Delft, NL
Genetics and Physiology of Food-Associated Microorganisms, 2010, Wageningen, NL
Principles of Ecological Genomics, 2011, Wageningen, NL

Meetings and Conferences

International Society for Microbial Ecology Congress, 2010, Seattle, USA
6th, 7th, 8th and 9th Kluyver Centre Symposium, (2009, 2010, 2011, 2012), Noordwijkerhout, NL
Microbial Networks in the Food Chain, Food Summit, 2009, Wageningen, NL
KNVM & NVMM 2012 Spring Meeting, 2012, Arnhem, NL
10th Symposium on Lactic Acid Bacteria, 2011, Egmond aan Zee, NL
Top Institute Food and Nutrition Mini-Symposium, 2012, Wageningen, NL

General Courses

Teaching and Supervising Thesis Students, 2010, Wageningen, NL
Scientific Writing, 2012, Wageningen, NL

Optional Courses and Activities

Bioinformation Technology Master Course, 2008, Wageningen, NL
PhD trip to China and Japan, 2011
Top Institute Food and Nutrition Project Meetings and WE-DAYS, Wageningen, NL
Meetings at NIZO Food Research, Wageningen, NL
PhD-PostDoc meetings, Wageningen, NL
Preparing PhD research proposal

The research described in this thesis was financially supported by Top Institute Food & Nutrition.

Financial support from Wageningen University and Top Institute Food & Nutrition for printing this thesis is gratefully acknowledged.

Cover Design by Sven Menschel
Printed by GVO drukkers en vormgevers B.V. / Ponsen & Looijen, Ede, NL