UNIVERSITÀ DI PARMA

UNIVERSITA' DEGLI STUDI DI PARMA

DOTTORATO DI RICERCA IN
SCIENZE DEGLI ALIMENTI

CICLO XXXI

Evolutionary study over space and time of lactic acid bacteria used as starter in cheeses manufacturing

Coordinatore:
Chiar.mo Prof. Furio Brighenti

Tutore:
Chiar.mo Prof. Erasmo Neviani

Co-Tutor:
Chiar.ma Prof. Monica Gatti

Dottoranda: Gaia Bertani

Anni 2015/2018

“It is not the strongest of the species that survives, nor the most intelligent that survives. It is the one that is most adaptable to change.”

Charles Darwin

“Non è la più forte delle specie che sopravvive, né la più intelligente, ma quella più reattiva ai cambiamenti.”

Charles Darwin

To Marco

Table of contents

Abstract

Microorganisms have been mutating and evolving on Earth for billions of years. Undoubtedly, microorganisms in general and bacteria in particular are extremely successful in inhabiting and adapting to diverse ecological niches available in the biosphere. That's because bacteria are characterized by a short generation times and large population sizes, but also thanks to transposons, plasmids and bacteriophages, that play a key role, moving and rearranging the presence and the order of genes in the chromosome. Food matrices can be viewed as different ecosystems, characterized by diverse stressful and selective conditions, leading to bacteria to adapt rapidly, acquiring or maintaining genes useful for their survival or losing genes no more helpful for the cell in those conditions. Lactic acid bacteria (LAB) are the most important bacteria traditionally associated with food fermentation due to their long and safe history of application. LAB have a spread ecological distribution; they are to be found in different environments, from milk and dairy products, vegetable and plants until cereals and meat. Besides, their presence in different ecological niches, LAB are commonly associated to dairy environment, in particular to starter cultures, which 'start' the fermentative process, producing lactic acid; their primary purpose in cheese manufacture. Many studies have demonstrated that these starter cultures have a large variability at strain level, rather than at species level. Even if, LAB biodiversity was extensively assessed in different works, a depth exploration of technological traits, lost or acquired, over the years, and in different ecological niches, have been very poorly investigated.

Based on these assumptions, the aim of this PhD thesis is to investigate how selective pressures, due to environmental or technological drivers, shape the bacterial genome of LAB and microbial communities over the time and/or through different ecological niches (space) (Figure 1).

The main object of this doctoral project was to lay the foundations for understanding the factors that contribute to the selection, in order to direct the technology to favour a biotype rather than another

in a specific environment. This PhD thesis was organized in four main sections. The first section was devoted to study a new protocol to perform in a quickly, cheaply and simply way, Amplified Fragment Length Polymorphism (AFLP) analysis. The second section was focused on *Streptococcus thermophilus* strains isolated from Pecorino Toscano ecosystem over time. It has been investigated if the strains isolated in different years from this niche, were able to produce histamine and tyramine, in order to understand if there has been an evolution over time of genetic traits involved in biogenic amines production. The third section has concerned the genome study of *Lactobacillus helveticus* strains isolated from fully different ecosystems, to gain an insight into the role of selective pressures in the genome evolution and to identify highly desirable traits with technological potential. Throughout the last section the microbial community of natural starter culture, used in Parmigiano Reggiano manufacturing, was investigated to understand how selective and environmental forces (like the seasonality and different temperatures in the technological process) shape the structure of this bacterial consortium.

The results obtained highlighted the importance to study how technological and environmental pressures, over time and space, change the ecosystem scenario, leading bacteria to acquire, maintain or lose genetic traits. Overall, the biodiversity, found in different ecological niches, should be perceived as a great richness; the leak of these biotypes could mean losing technological traits that may reoccur useful for various applications.

Sommario

Per miliardi di anni i microrganismi si sono trasformati ed evoluti sulla Terra. Indubbiamente, i microrganismi in generale e i batteri in particolare hanno un enorme successo nel risiedere e adattarsi nelle diverse nicchie ecologiche disponibili nella biosfera. Questo perché i batteri sono caratterizzati da tempi di generazione brevi e grandi dimensioni della popolazione, ma anche grazie a trasposoni, plasmidi e batteriofagi, che giocano un ruolo chiave, spostando e riorganizzando la presenza e l'ordine dei geni nel cromosoma. Le matrici alimentari possono essere considerate come diversi ecosistemi, caratterizzate da diverse condizioni stressanti e selettive, che portano i batteri a adattarsi rapidamente, acquisendo o mantenendo i geni utili per la loro sopravvivenza o perdendo geni non più utili per la cellula in quelle condizioni. I batteri lattici sono batteri tradizionalmente associati alla fermentazione degli alimenti a causa della loro lunga e sicura storia di applicazione. I batteri lattici hanno una distribuzione ecologica diffusa; possono risiedere in diversi ambienti, dal latte a prodotti lattiero-caseari, dai vegetali alle piante, fino a cereali e carne. I batteri lattici sono comunemente associati all'ambiente lattiero-caseario, in particolare alle colture starter, che "avviano" il processo fermentativo, producendo acido lattico; il loro scopo principale nella produzione di formaggio. Molti studi hanno dimostrato che queste colture starter hanno una grande variabilità a livello di ceppo, piuttosto che a livello di specie. Anche se la biodiversità dei batteri lattici, è stata ampiamente studiata in diversi lavori, una profonda esplorazione dei tratti tecnologici, persi o acquisiti, nel corso degli anni, e in diverse nicchie ecologiche è stata studiata molto poco.

Sulla base di questi presupposti, lo scopo di questa tesi di dottorato è quello di indagare come le pressioni selettive, dovute a fattori ambientali o tecnologici, modellano il genoma batterico dei batteri lattici e delle comunità microbiche nel tempo e/o attraverso diversi ecosistemi (spazio) (Figure 1).

Il principale obbiettivo di questo progetto di dottorato era porre le fondamenta per comprendere i fattori che contribuiscono alla selezione, al fine di indirizzare la tecnologia nel favorire un biotipo piuttosto che un altro in un ambiente specifico. Questa tesi di dottorato è stata organizzata in quattro sezioni principali. La prima sezione è stata dedicata allo studio di un nuovo protocollo per eseguire in modo rapido, economico e semplice, l'analisi genomica Amplified Fragment Lenght Polymorphism (AFLP). La seconda sezione si è focalizzata sui ceppi di *Streptococcus thermophilus* isolati nel tempo dall'ecosistema del Pecorino Toscano. È stato studiato se i ceppi isolati in anni diversi da questa nicchia, fossero in grado di produrre istamina e tiramina al fine di comprendere se nel tempo vi sia stata un'evoluzione dei tratti genetici coinvolti nella produzione di ammine biogene. La terza sezione ha riguardato lo studio del genoma di diversi ceppi di *Lactobacillus helveticus* isolati da ecosistemi completamente diversi, per avere una panoramica sul ruolo delle pressioni selettive nell'evoluzione del genoma e identificare tratti genetici altamente interessanti dal punto di vista tecnologico. Nell'ultima sezione la comunità microbica di colture starter, utilizzate nella produzione del Parmigiano Reggiano, è stata studiata al fine di comprendere come forze selettive e ambientali (come la stagionalità, le diverse temperature nel processo tecnologico) modellano la struttura di questo consorzio batterico.

I risultati ottenuti hanno evidenziato l’importanza di studiare come le pressioni tecnologiche e ambientali, nel tempo e nello spazio, cambiano lo scenario dell’ecosistema, portando i batteri a modificare il proprio genoma, acquisendo, mantenendo o perdendo tratti genetici. Complessivamente, la biodiversità, che si trova in diverse nicchie ecologiche, dovrebbe essere percepita come una grande ricchezza; la perdita di questi biotipi potrebbe significare perdere tratti tecnologici che potrebbero ripresentarsi utili per varie applicazioni future.

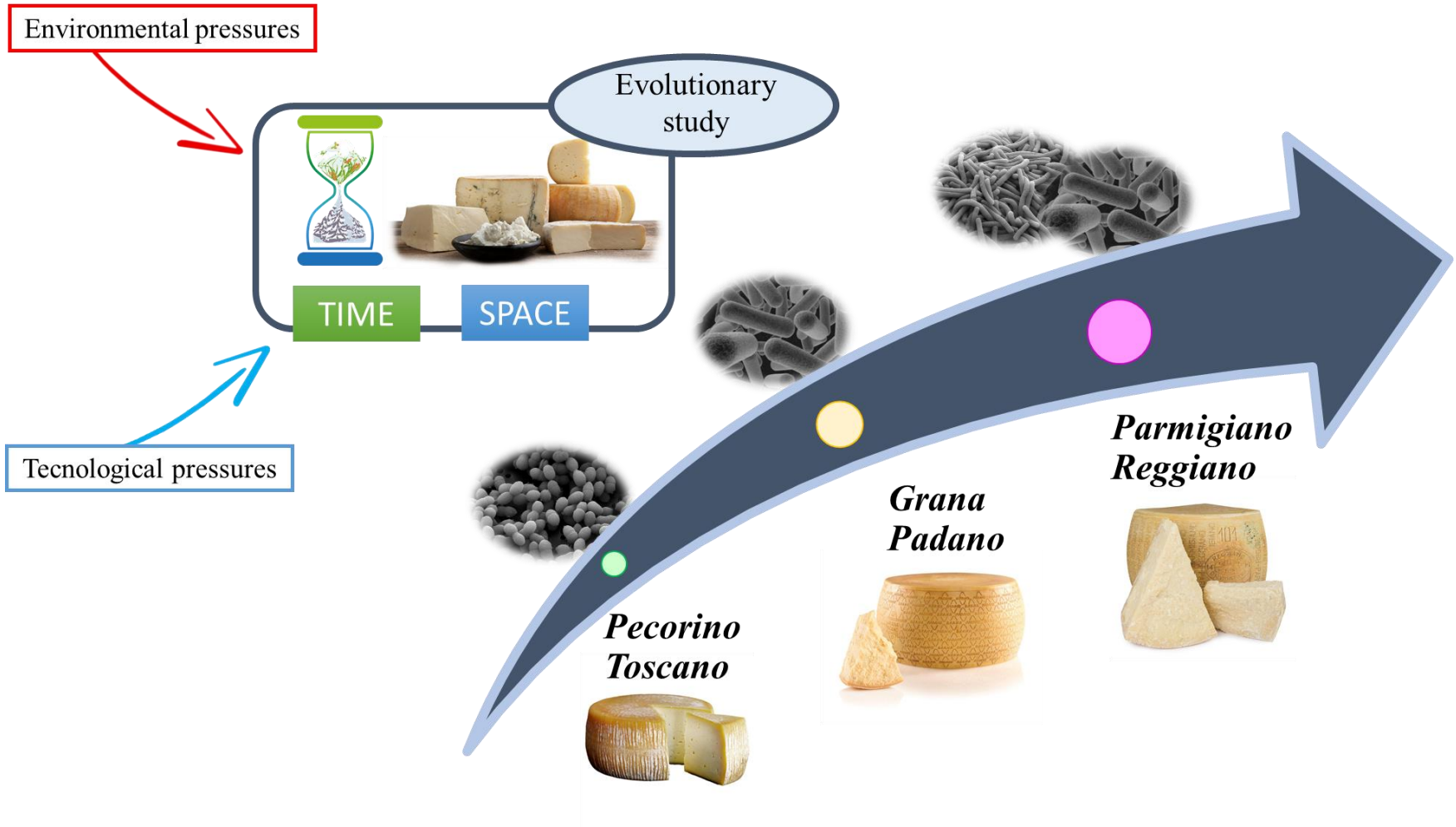


Figure 1 Graphical abstract about the evolutionary study over space and time performed in this PhD thesis.

Introduction

Key 'ingredients' for evolutionary bacteria success

Bacteria exist for a long time in the history of life on Earth. Over years, microorganisms have been mutating and evolving to adapt in an every-changing ecosystem (Bleuven et al., 2016). Evolution is deceptively simple as Darwinian concept, relying on genetic variation and selection of fitness (Brookifield et al., 2009). Given their short generation times and large population sizes, bacteria can evolve rapidly (Bell et al., 2016). Microbial populations adapt rapidly when they are introduced in a new environment, but at the same time microbial populations might continue to improve indefinitely, albeit slowly, even in a constant environment thanks to the contributions of individual mutations to fitness improvement (Elena et al., 2003).

Mutations in DNA sequence occurs constantly in all organisms. Even if these mutations, most of the times, have not any consequence for the fitness of the cell, it can happen that these DNA changes can enhance or retard the growth of the cell, leading to a acquire or lose a genetic trait (Loewe et al., 2010). No doubt, the great ability to transfer genetic information, at strain and species level, a distinctive feature in bacteria, is the major player in adaptation to changes in environment. Genetic exchanges occur frequently in bacterial world, and even if the amount of DNA switched is small, this movement can occur between distantly related microorganisms. Plasmids carry genes that could take part into bacterial chromosome, becoming a stable trait of the microorganisms' heritage. Furthermore, also transposons and bacteriophages, could play a key role, moving and rearranging the presence and the order of genes in the chromosome (Wiedenbeck et al., 2011). Recently, it has been also discovered that, under stressful conditions, bacterial cells may give arise to other cells, named isogenic variants, that despite they share the same DNA, tend to display intercellular differences, that can result in a differential behaviour. Altogether these considerations

suggest that these concepts, were essential ‘ingredients’ for evolutionary bacteria success (Ryall et al., 2012).

Food ecosystems

Food products can be viewed as different ecosystems, that offering diverse stressful and selective conditions, leading to bacteria undergo of diverse environmental and technological pressures. In this way, bacteria could develop distinct genetic and metabolic features among different habitats, acquiring or maintaining useful genetic traits for their survival or losing genes not any more helpful for the cell in that conditions.

Lactic acid bacteria (LAB) are the most important bacteria used in food fermentations (Teusink et al., 2017). LAB encompass a heterogeneous group of Gram-positive microorganisms, which have as common metabolic feature the production of lactic acid starting from the fermentation of carbohydrates (Fijian et al., 2014). Traditionally these microorganisms have been associated with food fermentation due to their long and safe history of application. The industrial importance of LAB is further evidenced by their Generally Regarded as Safe (GRAS) status, due to their ubiquitous appearance in food and their contribution to the healthy microflora of human mucosal surfaces. LAB have a wide ecological distribution; they are to be found in different environments, including milk and dairy products, vegetable and plants, cereals and meat (Mayo et al., 2008). A great phenotypic and genotypic biodiversity among the strains belonging to LAB has been observed (Lombardi et al., 2002; Gatti et al., 2004). This diversity represents an essential component of bacterial evolutionary success. Biodiversity, over years, has been seen, as a stabilizing force, representing a form of biological potentiality to different environmental responses (Ryall et al., 2012). LAB biodiversity was extensively assessed in different works, but a depth exploration of technological traits, lost or acquired, over the years, and in different ecological niches, has been very poorly studied.

Cheese as a reservoir of lactic acid bacteria

Cheese is a microbiologically and biochemically dynamic ecosystem (Gobbetti et al., 2018). LAB are the key players, albeit not always exclusive, of the biochemicals events that govern the production of most out the overall cheese varieties (Kongo et al., 2013).

Even if, less complex than natural ecosystems, the cheese niche is characterized by a variable core microbiota, whose evolution depends by many and different drivers. The stability and the synchronization of microbiological events were necessary to obtain a specific cheese (Neviani et al., 2013).

Intriguingly, the diverse bacteria capability to grow, survive, adapt to different unforeseen stresses, induced by technological processes, are key mechanisms to obtain different cheeses. The long-ripened cheeses produced from raw milk using NWS are characterized by complex interactions between starter lactic acid bacteria (SLAB) and non-starter lactic acid bacteria (NSLAB), which have diverse abilities to grow in a changing substrate (Gatti et al., 2014). The technological process, induce to heat-related, acidic, osmotic and oxidative stresses on microorganisms, leading to change in heat load, pH, water activity (aw), and redox potential gradients in the matrix (Beresford et al., 2001). As a result, the viability of SLAB is decreased within a few hours or a few days after the cheesemaking start, and suddenly a considerable fraction of SLAB, undergo autolysis, releasing their enzymes into the matrix. Conversely, NSLAB, have the capability to grow using different energy sources than lactose, are more resistant to environmental stresses, grow slowly becoming in ripened cheese the dominant microbiota (Gatti et al., 2014).

A little glimpse of starter cultures

Starter culture is a small, invisible ‘ingredient’, which disappears from the cheese once used. Despite its fugacity, starter culture, can be considered one of the main actors in the production of fermented foods (Bassi et al., 2015). A Starter culture can be defined as a culture of

microorganisms, for instance LAB, natural or selected, which is used to begin a fermentation process and acidify milk and curd.

Normally, acidification is reached by *in situ* production of lactic acid. Originally, and until relatively recently, the acidification of cheese is related to the autochthonous microbiota of the milk, particularly LAB. This biota is highly variable and unpredictable. As regard artisanal cheeses, the acidification still depends on the indigenous microbiota. Anyway, for consistency, a starter culture of LAB is inoculated in the milk for all factory-produced cheese (Gobbetti et al., 2018).

In several industrial cheesemaking, nowadays, it is almost common practice to add a starter culture of selected LAB to raw or pasteurized cheese milk, in order to reach a uniform and predictable acidification rate. Natural whey starter cultures are largely used for some traditional Italian cheeses and usually, the microbial composition of these natural cultures is closely related to the temperature used for their manufacture, which selects thermophilic or mesophilic lactic acid bacteria (Gobbetti et al., 2018). Indeed, the microorganisms are selected, by man in selected cultures and by technological process in natural cultures, for their capacity to reach fast acidification, produce lactic acid for curd production and a low pH in order to avoid spoilage, and to provide the enzymes involved in proteolysis and the conversion of amino acids into flavour compounds (Savijoki et al., 2006; Moser et al., 2018). Moreover, another important role of SLAB is to ensure a suitable environment for redox potential, pH and moisture content in the cheese, allowing the enzyme activity of the calf rennet and starter cultures and also promoting the growth of NSLAB (Beresford et al., 2001).

In Pecorino Toscano cheese the whole ewe's milk is generally pasteurized and added with mixed cultures, natural or selected strains belonging to *Streptococcus thermophilus* and *Lactococcus lactis* species, usually obtained from autochthonous strains isolated from milk and cheeses produced in the PDO area (Gobbetti et al., 2018). Conversely, the starter cultures used in Grana Padano and

Parmigiano Reggiano cheesemaking are exclusively natural, artisanal cultures based on technique of back-slopping. Natural whey starter (NWS) is obtained from the previous day's cheesemaking whey incubated at a decreasing temperature (Gatti et al., 2014). The selective pressures, such as chemical and physical drivers, that characterized this ecosystem, lead to the selection of a microflora, composed by thermophilic, acidic, and moderately heat-resistant LAB (Rossetti et al. 2009, Gatti et al. 2011). The different biotypes found in this environment are necessary and at the same time ensure the development of the ecosystem itself (Gatti et al., 2004). Many studies have demonstrated that these starter cultures have a large variability at strain level, rather than at species level (Gatti et al., 2014) The generally dominant species recovered in these starter cultures is *Lactobacillus helveticus* (Bottari et al., 2010). Many studies have revealed a great biodiversity, as regard phenotypic and genotypic characterization (Giraffa et al. 1998, Gatti et al. 1999, Giraffa et al. 2000, Lombardi et al. 2002, Rossetti et al. 2008; Gatti et al. 2004). Interestingly, this wide biodiversity, concerns also strains isolated from the same niche (Gatti et al., 2003; Lombardi et al., 2002; Andrighetto et al., 2004), suggesting once again that the great richness found at strain level, is essential for the ecosystem itself.

References

Andrighetto, C., Marcazzan, G., Lombardi, A. (2004). Use of RAPD-PCR and TTGE for the evaluation of biodiversity of whey cultures for Grana Padano cheese. *Letters in Applied Microbiology, 38,* 400–405.

Bas Teusink, B., Molenaar, D. (2017). Systems biology of lactic acid bacteria: For food and thought. *Current opinion in Systems Biology, 6,* 7-13.

Bassi, D., Puglisi, E., Cocconcelli, P.S. (2015). Comparing natural and selected starter cultures in meat and cheese fermentations. *Current Opinion in Food Science, 2,* 118-122.

Bell, G. (2016). Experimental macroevolution. *Proceedings Biological Sciences, 13,* 283.

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

Bleuven, C., Landry, C.R. (2016). Molecular and cellular bases of adaptation to a changing environment in microorganisms. *Proceedings Biological Sciences,* 283.

Bottari, B., Santarelli, M., Neviani, E., Gatti, M. (2010). Natural whey starter for Parmigiano Reggiano: culture-independent approach. *Journal of Applied Microbiology, 108,* 1676-84.

Brookfield, J.F.Y. (2009). Evolution and evolvability: celebrating Darwin 200. *Biology Letters, 5,* 44-46.

Elena, S.F., Lenski, R.E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics, 4,* 457-469.

Fijian, S. (2014). Microorganisms with claimed probiotic properties: an overview of recent literature. *International Journal of Environmental Research and Public Health, 11,* 4745-4767.

Gatti, M., Bottari, B., Lazzi, C., Neviani, E., Mucchetti, G. (2014). Invited review: microbial evolution in raw milk, long-ripened cheeses produced using undefined natural whey starters. *Journal of Dairy Science 97,* 573–591.

Gatti, M., Bottari, B., Santarelli, M., Neviani, E. (2011). Comparison of natural whey starters for Grana Padano cheese using sunray plots. *Annals of Microbiology, 61,* 475-481.

Gatti, M., Contarini, G., Neviani, E. (1999). Effectiveness of chemometrical techniques in discriminating *Lactobacillus helveticus* biotypes from natural dairy starter cultures by phenotypic characteristics. *Applied and Environmental Microbiology, 65,* 1450-1455.

Gatti, M., Lazzi, C., Rossetti, C., Mucchetti, G., Neviani, E. (2003). Biodiversity in *Lactobacillus helveticus* strains in natural whey starter used for Parmigiano Reggiano cheese. *Journal of Applied Microbiology, 95,* 463-470.

Gatti, M., Trevisano, C., Fabrizi, E., Neviani, E., Gardini, F. (2004). Biodiversity within *Lactobacillus helveticus* isolated from different natural whey starter cultures as revealed by classification trees. *Applied and Environmental Microbiology, 70,* 182-190.

Giraffa, G., De Vecchi, P., Rossi, P., Nicastro, G., Fortina. M.G. (1998). Genotypic heterogeneity among *Lactobacillus helveticus* strains isolated from natural cheese starters. *Journal of Applied Microbiology, 85,* 411–416.

Giraffa, G., Gatti, M., Rossetti, L., Senini, L., Neviani, E. (2000). Molecular diversity within *Lactobacillus helveticus* as revealed by genotypic characterization. *Applied and Environmental Microbiology, 66,* 1259-1265.

Gobbetti, M., Neviani, E., Calasso, M., De Angelis, M., Fox, F. P., Di Cagno, R. (2018). Drivers that establish and assembly the lactic acid bacteria biota in cheeses, *Trends in Food Science & Technology, 78,* 244–254.

Gobbetti, M., Neviani, E., Fox, P. (2018). The Cheeses of Italy: Science and Technology. (1st ed.) *Ed. Springer International,* (Chapter 6).

Kongo, M. (2013). Lactic Acid Bacteria as Starter-Cultures for Cheese Processing: Past, Present and *Future Developments*.

Loewe, L., Hill, W.G. (2010). The population genetics of mutations: good, bad and indifferent. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *27,* 1153-1167.

Lombardi, A., Dal Maistro, L., De Dea, P., Gatti, M., Giraffa, G., Neviani, E. (2002). A poliphasic approach to highlight genotypic and phenotypic diversities of *Lactobacillus helveticus* strains isolated from dairy starter cultures and cheeses. *Journal of Dairy Research, 69,* 139-149.

Mayo, B., van Sinderen, D., Ventura, M. (2008). Genome analysis of food grade lactic Acid-producing bacteria: from basics to applications. *Current Genomics, 9,* 169-183.

Moser, A., Schafroth, K., Meile, L., Egger, L., Badertscher, R., Irmler, S. (2018). Population Dynamics of *Lactobacillus helveticus* in Swiss Gruyère-Type Cheese Manufactured with Natural Whey Cultures. *Frontiers in Microbiology, 9,* 637.

Neviani, E., Bottari, B., Lazzi, C., Gatti, M. (2013). New developments in the study of the microbiota of raw-milk, long-ripened cheeses by molecular methods: the case of Grana Padano and Parmigiano Reggiano. *Frontiers in Microbiology, 4,* 36.

Rossetti, L., Fornasari, M.E., Gatti, M., Lazzi, C., Neviani, E., Giraffa G. (2008). Grana Padano cheese whey starters: microbial composition and strain distribution. *International Journal of Food Microbiology, 127,* 168-171.

Ryall, B., Eydallin, G., Ferenci, T. (2012). Culture history and population heterogeneity as determinants of bacterial adaptation: the adaptomics of a single environmental transition. *Microbiology and Molecular Biology Reviews, 76,* 597-625.

Savijoki, K., Ingmer, H., Varmanen, P. (2006). Proteolytic systems of lactic acid bacteria. *Applied Microbiology Biotechnology, 71,* 394-406.

Wiedenbeck, J., Cohan, F.M. (2011). Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiology Reviews, 35,* 957-976.

The aim

Overall the aim of this PhD thesis is to investigate how selective pressures, due to environmental or technological drivers, shape the bacterial genome and microbial communities over the time and/or through different ecological niches (space). To reach this purpose, different dairy ecosystems, characterized by diverse stringency, and different species belonging to LAB, were deepen evaluated.

Bearing that in mind, the present thesis reports the main results, summarized as follows, regarding one of the first evolutionary study over space and time of LAB used as starter in some DOP Italian cheeses manufacturing.

Chapter 1

In this chapter is reported an improvement of the Amplified Fragment Length Polymorphism (AFLP) protocol, a power genomic technique, largely applied for the classification of bacteria at the species and strain level with high discriminatory power and good reproducibility. In the last years, the Applied Biosystems AFLP Microbial Fingerprint Kit protocol was widely used in various studies to carry out AFLP characterization of selected bacteria strains. Considering that this kit is not anymore available, we have developed a new protocol to perform in a quickly, cheaply and simply way this analysis. This advanced protocol, takes into consideration the increasing of the alleles amplification efficiency and resolution, time-saving and cost effectiveness, if compare to the results obtained using the commercial AFLP microbial kit.

Chapter 2

In this chapter is reported the study of *Streptococcus thermophilus* strains isolated from Pecorino Toscano ecosystem over time. Pecorino Toscano cheese, is a specific, selective, environmental ecosystem modelling based on production of ewe's milk system. This peculiar production system,

characterized by ewe's milk, non-intensive farming, manually milk collection and some artisanal traits of cheesemaking, it means that the milk microbiota could be considered a 'mirror' of the ecosystem strictly related to the area of production. It is known that the microbiota of Pecorino Toscano cheese plays a key role in the biochemical and sensory characteristics of the product but, on the other hand, it could also be responsible for the production of biogenic amines (BAs). The aim of this chapter is to investigate if the strains isolated from different steps of Pecorino Toscano cheesemaking in different years, were able to produce histamine and tyramine, understanding if there has been an evolution over the years of genetic traits involved in biogenic amines production.

Chapter 3

In this chapter is reported the genome study of *Lactobacillus helveticus* strains isolated from natural whey starter used in Grana Padano (GP) and Parmigiano Reggiano (PR) manufacturing. Three strains of *L. helveticus* (Lh 12, Lh 23 and UC1156) were sequenced. Lh 12 and Lh 23 were isolated from GP NWS, while UC1156 from PR NWS. The aim of this chapter is understanding the genome evolution of *L. helveticus* under different selective pressures and identifying highly desirable traits with technological potential. To reach this purpose, these three strains were compared to other eleven strains isolated from fully different ecosystems through a bioinformatic approach.

Chapter 4

In this chapter is reported the evolutionary study of natural whey starter ecosystem used in the production of PR cheese, analysing over time two different production lines ("Conventional" and "Organic"). The aim of this work has two main objectives: (i) to characterise the microbial community of the two NWS collected to develop traditional knowledge and (ii) to investigate how selective and environmental forces (like the seasonality and different temperatures in the technological process) shape the microbial community of NWS.

Chapter 1

Advanced AFLP protocol for microbial diversity fingerprinting

Gaia Bertani, Maria Luisa Savo Sardaro, Erasmo Neviani and Camilla Lazzi

Journal of Applied Genetics (Submitted)

Abstract

Over the last decade, for the identification and genotyping of prokaryotic and eukaryotic organisms, several methods at the DNA level have been developed. These genomic methods differ as regard taxonomic range, discriminatory power, reproducibility, and ease of interpretation and standardization. The Amplified Fragment Length Polymorphism (AFLP) technique is a very powerful DNA fingerprinting technique for DNAs of any source or complexity, varying in both size and base composition. In addition, this method shows high discriminatory power and good reproducibility showing to be efficient for discriminating at the species but also at the strain level. The development and application of AFLP have allowed to significant progress in the study of biodiversity and taxonomy of microorganisms. In the last years, the Applied Biosystems AFLP Microbial Fingerprint Kit protocol was widely used in various studies to perform AFLP characterization of selected bacteria strains, previously described by Vos et al. (1995), although including several modifications. This study aims to propose an alternative protocol to replace the commercial kit out of production, giving so the possibility to use the method for bacteria genetic fingerprinting analysis in biodiversity study. In particular previous results on different species (*Listeria monocytogenes, Lactobacillus plantarum* and *Streptococcus thermophilus*) obtained with the commercial kit were compared with new AFLP procedure to validate the protocol. In comparison with the AFLP Microbial Fingerprint Kit, the new designed protocol shows high reproducibility, resolution and it is a faster method with lower costs.

Keywords: AFLP protocol, Bacterial diversity, Genome Polymorphisms, Phylogenetic analysis

Introduction

In the last decades, several PCR-based fingerprinting methods have been developed for bacteria genomic fingerprinting and genome screening purposes. It is possible to differentiate high-throughput approaches based on DNA or RNA analyses based on several next-generation sequencing techniques (Ercolini et al. 2015; De Filippis et al. 2017; Garofalo et al. 2017) and the low-throughput methodologies that are more largely used. Among the last, can be considered the randomly amplified polymorphic DNA (RAPD) (Cocconcelli et al. 1995; Franklin et al. 1999; Rossetti and Giraffa, 2005; Perin et al. 2017), the primers based on repetitive elements in the genome (rep-PCR) (Versalovic et al. 1991; De Bruijn, 1992; Anderson et., 2010), the amplified ribosomal DNA restriction analysis (ARDRA) (Vaneechoute et al. 1992; Gulitz et al. 2013), the automated ribosomal intergenic spacer analysis (ARISA) (Kovacs et al. 2010) and length heterogeneity-PCR (LH-PCR) (Lazzi et al. 2004, Gatti et al. 2008, Savo Sardaro et al. 2018). Besides these, Amplified fragment length polymorphism (AFLP) (Zabeau and Vos, 1993; Vos et al. 1995; Janssen et al. 1996) has been largely used for genomic fingerprinting of DNA from a variety of sources. As widely reported in literature, AFLP is a valuable technique for the classification of bacteria at the species and strain level with high discriminatory power and good reproducibility (Janssen et al. 1996; Blears et al. 1998; Savelkoul et al. 1999; Jarraud et al. 2002). The AFLP fingerprinting provides several advantages over other techniques (Curtin et al. 2007). First of all, prior knowledge of a microorganism's genome sequence is not necessary. Moreover, AFLP alleles can be fluorescently labelled, allowing a parallel characterization of several samples in automatic genome analysers. Once the technique is fine-tuned, it is possible to obtain accurate information fast enough to allow for an efficient identification and differentiation of species and strains. The development and application of AFLP as a fingerprinting method has led to significant progress in

the study of the genetic diversity and taxonomy of bacteria (Heir et al. 2000; Giraffa et al. 2001; Cocolin and Ercolini, 2008; Cappello et al. 2008; Lazzi et al. 2009; Di Cagno et al. 2010; Lévesque et al. 2012; Nabhan et al. 2012; Hamza et al. 2012; Bernini et al. 2013; Jérôme et al. 2016). Most of the bacteria strain characterization, in the last years, was developed following the Applied Biosystems AFLP Microbial Fingerprint Kit protocol, according to the manufacturer's instructions. In this paper, is reported an improvement of the AFLP protocol taking into consideration the increasing of the alleles amplification efficiency and resolution, time-saving and cost effectiveness, comparing the advanced protocol with the results obtained using the commercial AFLP microbial kit.

Materials and methods

Bacteria and growth conditions

Twenty-one strains isolated from different food matrices were used in this study (Table 1). They include 7 *Streptococcus thermophilus*, 7 *Listeria monocytogenes* and 7 *Lactobacillus plantarum* strains. Bacterial strains were maintained as frozen stocks (−80 °C) in M17 (*S. thermophilus*), TSB broth (*L. monocytogenes*) and MRS (*L. plantarum*) (Oxoid, Milan, Italy) supplemented with 15% glycerol (w/v). Before use, the cultures were propagated twice with a 3% (v/v) inoculum into the appropriate media and incubated at 42°C (*S. thermophilus*), 37°C (*L. monocytogenes*) and 30°C (*L. plantarum*) for 24 h in optimal growth conditions. All *S. thermophilus* and *L. monocytogenes* strains belong to the collection of Food and Drug Department, University of Parma, Italy; 7 *L. plantarum* strains (POM1, POM31, POM43, POM40, POM8, C6, POM38) were kindly given by the Department of Soil, Plant and Food Science, University of Bari, Italy.

Table 1 Strains used in this study.

Species	***Strains***	***Source***
S. thermophilus	100	Pecorino Toscano cheese
S. thermophilus	145	Pecorino Toscano cheese
S. thermophilus	159	Pecorino Toscano cheese
S. thermophilus	418	Pecorino Toscano cheese
S. thermophilus	4027	Pecorino Toscano cheese
S. thermophilus	4028	Pecorino Toscano cheese
S. thermophilus	4042	Pecorino Toscano cheese
L. plantarum	POM1	Tomato
L. plantarum	POM8	Tomato
L. plantarum	POM31	Tomato
L. plantarum	POM38	Tomato
L. plantarum	POM40	Tomato
L. plantarum	POM43	Tomato
L. plantarum	C6	Carrot
L. monocytogenes	Lm6	Gorgonzola cheese
L. monocytogenes	Lm9	Gorgonzola cheese
L. monocytogenes	Lm34	Gorgonzola cheese
L. monocytogenes	Lm35	Gorgonzola cheese
L. monocytogenes	Lm40	Gorgonzola cheese
L. monocytogenes	Lm41	Gorgonzola cheese
L. monocytogenes	Lm44	Gorgonzola cheese

AFLP analysis

Preparation of primary template for AFLP analysis

The AFLP procedure was performed according to the method of Vos et al. (1995) with the modifications described below. Restriction-Ligation reactions were performed in a final volume of 50 µl containing 1×T4 DNA ligase buffer with 1 mM ATP, 250 ng/µl BSA, 10 mM ATP, 20 U/µl of *Eco*RI, 10 U/µl of *Mse*I and 500 ng of genomic DNA. Two different adapters (sequences shown in Table 2), one for the *Eco*RI sticky ends and one for the *Mse*I sticky ends, were ligated to the DNA by adding to the reaction of a mix containing 5 pmol/µL of *Eco*RI adaptor, 50 pmol/µl of *Mse*I adaptor, and 200 U/µl of T4 DNA ligase (New England Biolabs). The reaction was incubated for 4h at 37°C. Two replication for each sample were performed.

The digested-ligated DNA product to be used as templates for the first amplification reaction was diluted 10-fold with RNAse and DNAse free water and stored at -20 °C.

Pre-amplification

The "non-selective" primers *Eco*RI-0 and *Mse*I-0 (Table 2) were used for pre-amplification of digested-ligated DNA. Each pre-amplification contained 5 µl of digested-ligated DNA previously described, 1.5 µl of unlabelled *Mse*I-0 primer (10 µM) and 1.5 µl of labelled *Eco*RI-0 primer (10 µM), 25 µl of GoTaq® Colorless Master Mix (PROMEGA, Madison, Wisconsin, USA).

The reaction was subjected to the following PCR conditions: 3 min at 94 °C, 14 cycles (45 s at 94 °C, 30 s at 65 °C and in each cycle the annealing temperature decrease 1°C, 1 min at 72°C), 19 cycles (45 s at 94°C, 30 s at 56°C and 1 min at 72°C), 5 min extension at 72°C and a final step for 15 min at 30 °C. All amplifications were performed in a GeneAmp® PCR System 2700 (Applied Biosystem). Subsequently, the pre-amplification product was diluted 10-fold with RNAse and DNAse free water and stored at -20 °C.

Selective-amplification

Different primer combinations were used, based on different species analysed: *Eco*RI-A/*Mse*I-C for *L. monocytogenes* and *Eco*RI-A/*Mse*I-A for *S. thermophilus* and *L. plantarum* (Table 2). Each selective-amplification contained 5 µl of the diluted pre-amplification product described previously, 1.5 µl of unlabelled *Mse*I-A primer (10 µM) and 1.5 µl of labelled *Eco*RI-A primer (10 µM), 25 µl of GoTaq® Colorless Master Mix (PROMEGA).

The thermocycler program consisted in 2 min at 72 °C, 33 cycles (30 s at 94 °C, 1 min at 56 °C, 1 min at 72 °C), 2 min extension at 72°C and a final step for 30 min at 60 °C. All amplifications were performed in a GeneAmp® PCR System 2700 (Applied Biosystem). Ten microliters of each selective amplification product were separated by electrophoresis on a 1.2% agarose gel at 90 V/cm for 20 min to check the amplifications.

Table 2 Primers used for AFLP analysis.

Primer Name	***Sequence (5'-3')***
*Eco*RI-0	GACTGCGTACCAATTC (labelled FAM 5')
*Mse*I-0	GATGAGTCCTGAGTAA
*Eco*RI-A	GACTGCGTACCAATTCA (labelled FAM 5')
*Mse*I-A	GATGAGTCCTGAGTAAA
*Mse*I-C	GATGAGTCCTGAGTAAC

Fragments analysis

Eight microliters of each amplified products from selective amplification were added to 1.5 μl of GeneScan-500 [LIZ] size standard (Applied Biosystem-Pe Corporation) and 27 μl of deionized formamide. The mixture was heated 5 min at 95°C and cooled on ice. Samples were loaded and run on the ABI Prism 310 (Applied Biosystem-Pe Corporation, Waltham, Massachusetts, USA) and analysed using GeneMapper Analysis Software (Applied Biosystem-Pe Corporation). The data for each run were saved as an individual GeneScan file and displayed as an electropherogram. A threshold, used in scoring to consider only sharp and easily distinguishable peaks, of 50 RFU (Relative Fluorescent Unit) was considered for results obtained by new protocol, while a threshold of 80 RFU for results obtained with AFLP Microbial Fingerprinting Kit; all the signals under this value were treated as background and not scored. Peaks representing AFLP fragments from 50 to 500 bp were reported as binary format with "1" for the presence of a band and "0" for its absence.

Results and discussion

This study focuses on the develop an advanced AFLP procedure for the detection of polymorphisms in bacterial genomes (Zabeau & Vos, 1993; Vos et al.1995; Janssen et al. 1996). The new advanced procedure was mainly compared with AFLP Microbial Fingerprinting Kit, considering its wide employment, in the last years, for bacteria AFLP fingerprinting analysis. The main aim has been to develop an AFLP protocol, to replace the AFLP Microbial Fingerprinting Kit, able to improve the

profile quality of the assay, increasing sensitivity and precision and decreasing scoring time and errors.

The first modification in the protocol, in comparison to AFLP Microbial Fingerprinting Kit, is related to performing a combined digestion and adaptor ligation at the same time without affecting the number of bands and the end results. The experiment was performed on the 21 strains showed in Table 1 (Appendix, Supplementary Figure 1, 2, 3) and an example of the result is presented in Figure 1 where the comparison between the advanced protocol and the commercial kit on the strain Lm44 showed the same profiles.

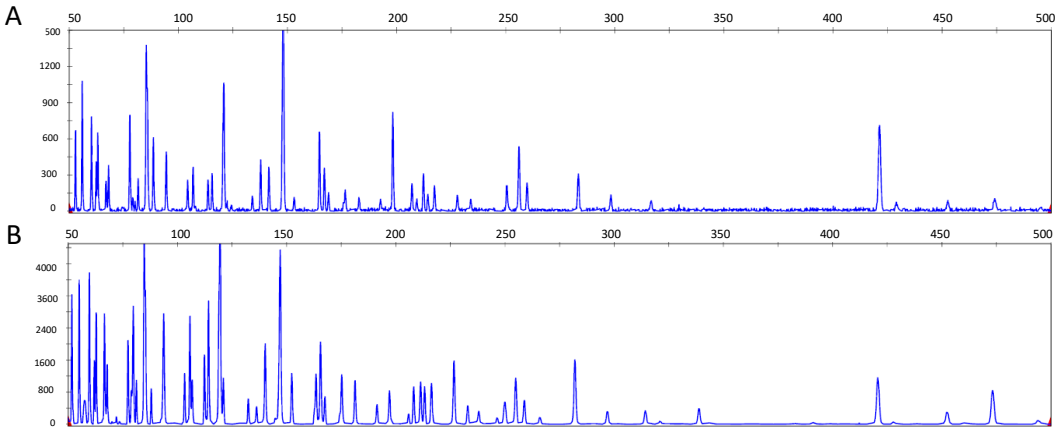


Figure 1 Electropherograms of *L. monocytogenes* strain Lm44 AFLP profiles. Comparison of AFLP profiles obtained with combined and not-combined digestion-ligation procedure. A) AFLP profile strain Lm44 obtained with not-combined digestion-ligation procedure using AFLP Microbial Fingerprinting kit; B) AFLP profile strain Lm44 performed with combined digestion-ligation procedure in the advanced AFLP protocol.

This modification gives the opportunity to save time and to reduce the laboratory costs for analysis and this was also considered positively by Curtin et al. (2007). The second condition that has been also evaluated for modification, is the dilution of the digested-ligated DNA fragments to be used as templates in the first amplification reaction. The dilution in AFLP Microbial Fingerprinting Kit

procedure is 20-fold while in the optimized AFLP protocol this dilution of digested-ligated DNA fragments has a negative influence in the electropherograms resolution, whereas the optimal dilution of digested-ligated DNA fragments is obtained with 10-fold. Figure 2 shows the electropherograms obtained with the two dilutions using the strains *S. thermophilus* 100 and 4042.

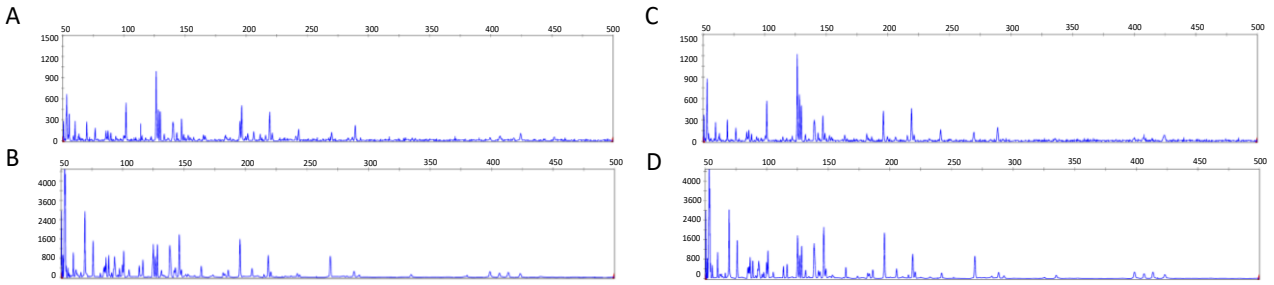


Figure 2 Electropherograms of *S. thermophilus* strains 100 and 4042 AFLP profiles obtained using 20-fold and 10-fold dilution of digested-ligated DNA product. A) AFLP profile of strain 100 related to 20-fold dilution of digested-ligated DNA product; B) AFLP profile of strain 100 related to 10-fold dilution of digested-ligated DNA product; C) AFLP profile of strain 4042 related to 20-fold dilution of digested-ligated DNA product; D) AFLP profile of strain 4042 related to 10-fold dilution of digested-ligated DNA product.

The third parameter that was consider is the one related with the dilution of the pre-selective PCR products used as templates for the subsequent selective PCR. Two different conditions were evaluated one without any dilution in compare with samples diluted 10-fold. The amplification results analysed by capillary electrophoresis were very similar one each other (data not shown) so for this parameter the condition reported by AFLP Microbial Fingerprinting Kit and other authors (Vos et al. 1995; Janssen et al. 1996) of 10-fold dilution, was maintained. Finally, the PCR conditions have been modified with the touchdown PCR applied in the pre-amplification reaction and not in the selective one as in the kit. In addition, the number of PCR cycles in the selective amplification has been taken in consideration with the aim to increase the peak intensity without

introduce high level of *Taq* polymerase errors, that can give also differences in peak base pair size. The number of selective PCR cycles has been increased to 33 comparing with Vos et al. (1995), Janssen et al. (1996) and the commercial kit where this parameter was 24 and 30 respectively. The modification introduced allows us to have electropherograms with peaks higher in the intensity of fluorescence and more defined for the next step of the data elaboration (Figure 3).

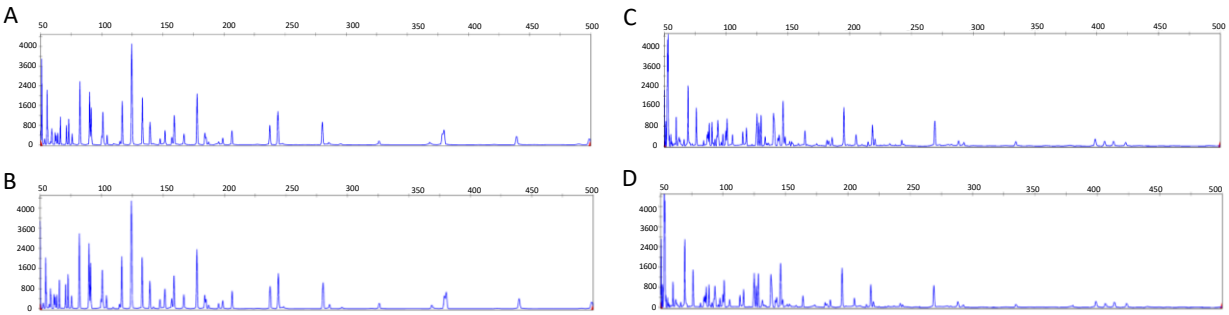


Figure 3 Electropherograms of *S. thermophilus* strains 201 and 100 AFLP profiles of the PCR product obtained using 33 and 30 PCR cycles in selective PCR. A) AFLP profile of strain 201 related to 30 cycles in selective PCR; B) AFLP profile of strain 201 related to 33 cycles; C) AFLP profile of strain 100 related to 30 cycles in selective PCR; D) AFLP profile of strain 100 related to 33 cycles in selective PCR.

In addition, the new optimized protocol provides an improvement in the signal-to-background ratio in the electropherograms and increase the intensity of the peaks' profiles obtained (Figure 4 and Appendix, Supplementary Figure 1, 2, 3). Also, the possibility to maintain and compare previous data obtained with the AFLP microbial kit is showed in Figure 4 where two strains previously analysed, by Lazzi et al. (2009) using the commercial kit, have the same peaks profiles with the advanced AFLP protocol and it is also shown in supplementary material on the other 19 strains.

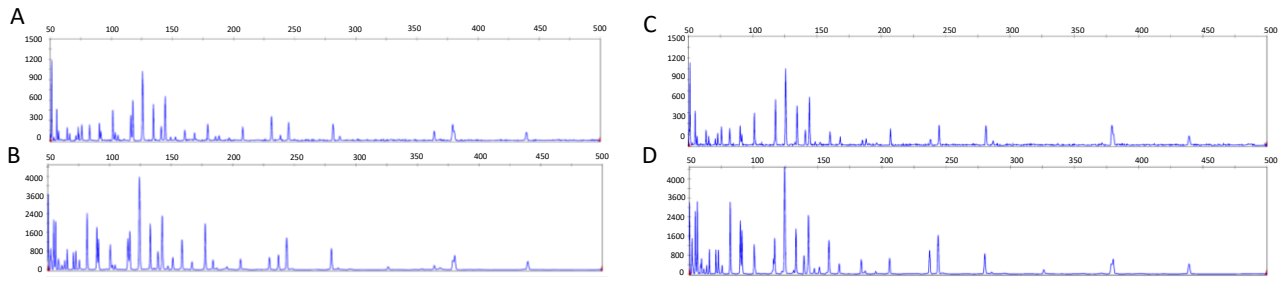


Figure 4 Comparison of the AFLP profiles of *S. thermophilus* strains 4027 and 4028 obtained with the AFLP microbial Kit and the advanced AFLP protocol. A) AFLP profile of strain 4027 obtained with the AFLP microbial Kit; B) AFLP profile of strain 4027 obtained with the advanced AFLP protocol; C) AFLP profile of strain 4028 obtained with the AFLP microbial Kit; D) AFLP profile of strain 4028 obtained with the advanced AFLP protocol.

This opens the possibility to all laboratory to continue still their phylogenetic study using data previously obtained with the commercial kit.

Conclusions

AFLP is an excellent technique to differentiate strains or very closely related species and as a good phylogenetic tool. In the last years, the use of several restriction enzymes and many fluorescence molecules at the same time has given the opportunities to achieve a very extensive screening of the bacteria genomes. The modification of the digested-ligated step reduces drastically the time needed for the sample's analysis. Moreover, the different condition of dilution of digested-ligated DNA fragments and the possibility to increase the number of PCR cycles allow to obtain comparable and better results in terms of distinctiveness and intensity of the band's peaks in comparison with the commercial AFLP Microbial kit. In addition, considering that this kit is not available anymore and the resulted electropherograms are the same with both protocols, thanks to this method we can compare AFLP profiles with previous database without repeating the analysis of all the samples.

Overall, the modified protocol gives the opportunities to reduce the time-consuming and labour-intensive, and costs-effective given the possibilities to continue to use the AFLP technique as an excellent tool to analyse many samples.

Acknowledgements

Authors acknowledge the Department of Soil, Plant and Food Science, University of Bari, Italy for providing *L. plantarum* strains (POM1, POM31, POM43, POM40, POM8, C6, POM38). This research was partially supported by Consorzio per la tutela Pecorino Toscano DOP and by Consorzio per la tutela del formaggio Gorgonzola.

References

Anderson, P.N., Hume, M.E., Byrd, J.A., Hernandez, C., Stevens, S.M., Stringfellow, K., Caldwell, D.J. (2010). Evaluation of repetitive extragenic palindromic-polymerase chain reaction and denatured gradient gel electrophoresis in identifying *Salmonella* serotypes isolated from processed turkeys. *Poultry Science, 89,* 1293-1300.

Bernini, V., Bottari, B., Dalzini, E., Sgarbi, E., Lazzi, C., Neviani, E., Gatti, M. (2013). The presence, genetic diversity and behaviour of *Listeria monocytogenes* in blue-veined cheese rinds during the shelf life. *Food Control, 34,* 323-330.

Blears, M., De Grandis, S., Lee, H., Trevors, J.T. (1998). Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. *Journal of Industrial Microbiology & Biotechnology, 21,* 99–114.

Cappello, M.S., Stefani, D., Grieco, F., Logrieco, A., Zapparoli, G. (2008). Genotyping by Amplified Fragment Length Polymorphism and malate metabolism performances of indigenous *Oenococcus oeni* strains isolated from Primitivo wine. *International Journal of Food Microbiology, 127,* 241-245.

Cocconcelli, P.S., Porro, D., Galandini, S., Senini, L. (1995). Development of RAPD protocol for typing of strains of lactic acid bacteria and enterococci. *Letters in Applied Microbiology, 21,* 376-379.

Cocolin, L., Ercolini, D. (2008). Molecular Techniques in the Microbial Ecology of Fermented Foods. *Food Microbiology and Food Safety,* Springer, New York, pp. 1–30.

Curtin, C.D., Bellon, J.R., Henschke, P.A., Godden, P.W., de Barros Lopes, M.A. (2007). Genetic diversity of *Dekkera bruxellensis* yeasts isolated from Australian wineries. *FEMS Yeast Research, 7,* 471–481.

De Bruijn, F.J. (1992). Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. *Applied and Environmental Microbiology, 58,* 2180-2187.

De Filippis, F., Parente, E., Ercolini, D. (2017). Metagenomics insights into food fermentations. *Microbial Biotechnology, 10,* 91-102.

Di Cagno, R., Minervini, G., Sgarbi, E., Lazzi, C., Bernini, V., Neviani, E., Gobbetti, M. (2010). Comparison of phenotypic (Biolog System) and genotypic (random amplified polymorphic DNA-polymerase chain reaction, RAPD-PCR, and amplified fragment length polymorphism, AFLP) methods for typing *Lactobacillus plantarum* isolates from raw vegetables and fruits. *International Journal of Food Microbiology, 143,* 246-253.

Ercolini, D., Francavilla, R., Vannini, L., De Filippis, F., Capriati, T., Di Cagno, R., Iacono, G., De Angelis, M., Gobbetti, M. (2015). From an imbalance to a new imbalance: Italian-style gluten-free diet alters the salivary microbiota and metabolome of African celiac children. *Scientific Reports, 5,* 18571.

Franklin, R.B., Taylor, D.R., Mills, A.L. (1999). Characterization of microbial communities using randomly amplified polymorphic DNA (RAPD). *Journal of Microbiological Methods, 35,* 225-235.

Garofalo, C., Osimani, A., Milanović, V., Taccari, M., Cardinali, F., Aquilanti, L., Riolo, P., Ruschioni, S., Isidoro, N., Clementi, F. (2017). The microbiota of marketed processed edible insects as revealed by high-throughput sequencing. *Food Microbiology, 62,* 15-22.

Gatti, M., De Dea Lindner, J., De Lorentiis, A., Bottari, B., Santarelli, M., Bernini, V., Neviani, E. (2008). Dynamics of whole and lysed bacterial cells during Parmigiano-Reggiano cheese production and ripening. *Applied and Environmental Microbiology, 74,* 6161-6167.

Giraffa, G., Paris, A., Valcavi, L., Gatti, M., Neviani, E. (2001). Genotypic and phenotypic heterogeneity of *Streptococcus thermophilus* strains isolated from dairy products. *Journal of Applied Microbiology, 91,* 937-943.

Gulitz, A., Stadie, J., Ehrmann, M.A., Ludwig, W., Vogel, R.F. (2013). Comparative phylobiomic analysis of the bacterial community of water kefir by 16S rRNA gene amplicon sequencing and ARDRA analysis. *Journal of Applied Microbiology, 114,* 1082-1091.

Hamza, A.A., Robene-Soustrade, I., Jouen, E., Lefeuvre, P., Chiroleu, F., Fisher-Le Saux, M., Gagnevin, L., Pruvost, O. (2012). MultiLocus Sequence Analysis and Amplified Fragment Length Polymorphism-based characterization of xanthomonads associated with bacterial spot of tomato and pepper and their relatedness to *Xanthomonas* species. *Systematic and Applied Microbiology, 35,* 183-190.

Heir, E., Lindstedt, B.A., Vardund, T, Wasteson, Y., Kapperud, G., (2000). Genomic fingerprinting of shigatoxin-producing *Escherichia coli* (STEC) strains: comparison of pulsed-field gel electrophoresis (PFGE) and fluorescent amplified-fragment-length polymorphism (FAFLP). *Epidemiology & Infection, 125,* 537–548.

Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., Zabeau, M., Kersters, K. (1996). Evaluation of the DNA Fingerprinting Method AFLP as a New Tool in Bacterial Taxonomy. *Microbiology, 142,* 1881-1893.

Jarraud, S., Mougel, C., Thioulouse, J., Lina, G., Meugnier, H., Forey, F., Vandenesch, F. (2002). Relationships between *Staphylococcus aureus* Genetic Background, Virulence Factors, agr Groups (Alleles), and Human Disease. *Infection and Immunity, 70,* 631–641.

Jérôme, M., Macé, S., Dousset, X., Pot, B., Joffraud, J.J. (2016). Genetic diversity analysis of isolates belonging to the *Photobacterium phosphoreum* species group collected from salmon products using AFLP fingerprinting. *International Journal of Food Microbiology, 217,* 101-109.

Kovacs, A., Yacoby, K., Gophna, U. (2010). A systematic assessment of automated ribosomal intergenic spacer analysis (ARISA) as a tool for estimating bacterial richness. *Research in Microbiology, 161,* 192-197.

Lazzi, C., Rossetti, L., Zago, M., Neviani, E., Giraffa, G. (2004). Evaluation of bacterial communities belonging to natural whey starters for Grana Padano cheese by length heterogeneity-PCR. *Journal of Applied Microbiology,* 96, 481–490.

Lazzi, C., Bove, C.G., Sgarbi, E., Gatti, M., La Gioia, F., Torriani, S., Neviani, E. (2009). Application of AFLP fingerprint analysis for studying the biodiversity of *Streptococcus thermophilus*. *Journal of Microbiological Methods, 79,* 48-54.

Lévesque, S., St-Pierre, K., Frost, E., Arbeit, R.D., Michaud, S. (2012). Use of amplified-fragment length polymorphism to study the ecology of *Campylobacter jejuni* in environmental water and to predict multilocus sequence typing clonal complexes. *Applied and Environmental Microbiology, 78,* 2470-2473.

Nabhan, S., Wydra, K., Linde, M., Debener, T. (2012). The use of two complementary DNA assays, AFLP and MLSA, for epidemic and phylogenetic studies of pectolytic enterobacterial strains with focus on the heterogeneous species *Pectobacterium carotovorum*. *Plant Pathology, 61,* 498-508.

Perin, L.M., Savo Sardaro, M.L., Nero, L.A., Neviani, E., Gatti, M. (2017). Bacterial ecology of artisanal Minas cheeses assessed by culture-dependent and-independent methods. *Food Microbiology, 65,* 160-169.

Rossetti, L., Giraffa, G. (2005). Rapid identification of dairy lactic acid bacteria by M13-generated, RAPD-PCR fingerprint databases. *Journal of Microbiological Methods, 63,* 135-144.

Savelkoul, P.H., Aarts, H.J., De Haas, J., Dijkshoorn, L., Duim, B., Otsen, M., Rademaker, J.L., Schouls, L., Lenstra, J.A. (1999). Amplified-fragment length polymorphism analysis: the state of an art. *Journal of Clinical Microbiology, 37,* 3083-3091.

Savo Sardaro, M.L., Perin, L.M., Bancalari, E., Neviani, E., Gatti, M. (2018). Advancement in LH-PCR methodology for multiple microbial species detections in fermented foods. *Food Microbiology, 74,* 113-119.

Vaneechoutte, M., Rossau, R., De Vos, P., Gillis, M., Janssen, D., Paepe, N., De Rouck, A., Fiers, T., Claeys, G., Kersters, K. (1992). Rapid identification of bacteria of the *Comamonadaceae* with amplified ribosomal DNA-restriction analysis (ARDRA) *FEMS Microbiological Letters, 93,* 227-234

Versalovic, J., Koeuth, T., Lupski, J.R. (1991). Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Research, 19,* 6823–6831.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research, 23,* 4407-4414.

Zabeau, M., Vos, P. (1993). Selective restriction fragment amplification: a general method for DNA fingerprinting. *European Patent Office,* Publication 0 534 858 Al.

Chapter 2

Pecorino Toscano cheese: genetic diversity and histamine production of Streptococcus thermophilus strains over time

Annalisa Ricci[$], Gaia Bertani[$], Luca Calani, Valentina Bernini, Camilla Lazzi and Erasmo Neviani

[$] Both authors contributed equally to this manuscript

Food Research International (Submitted)

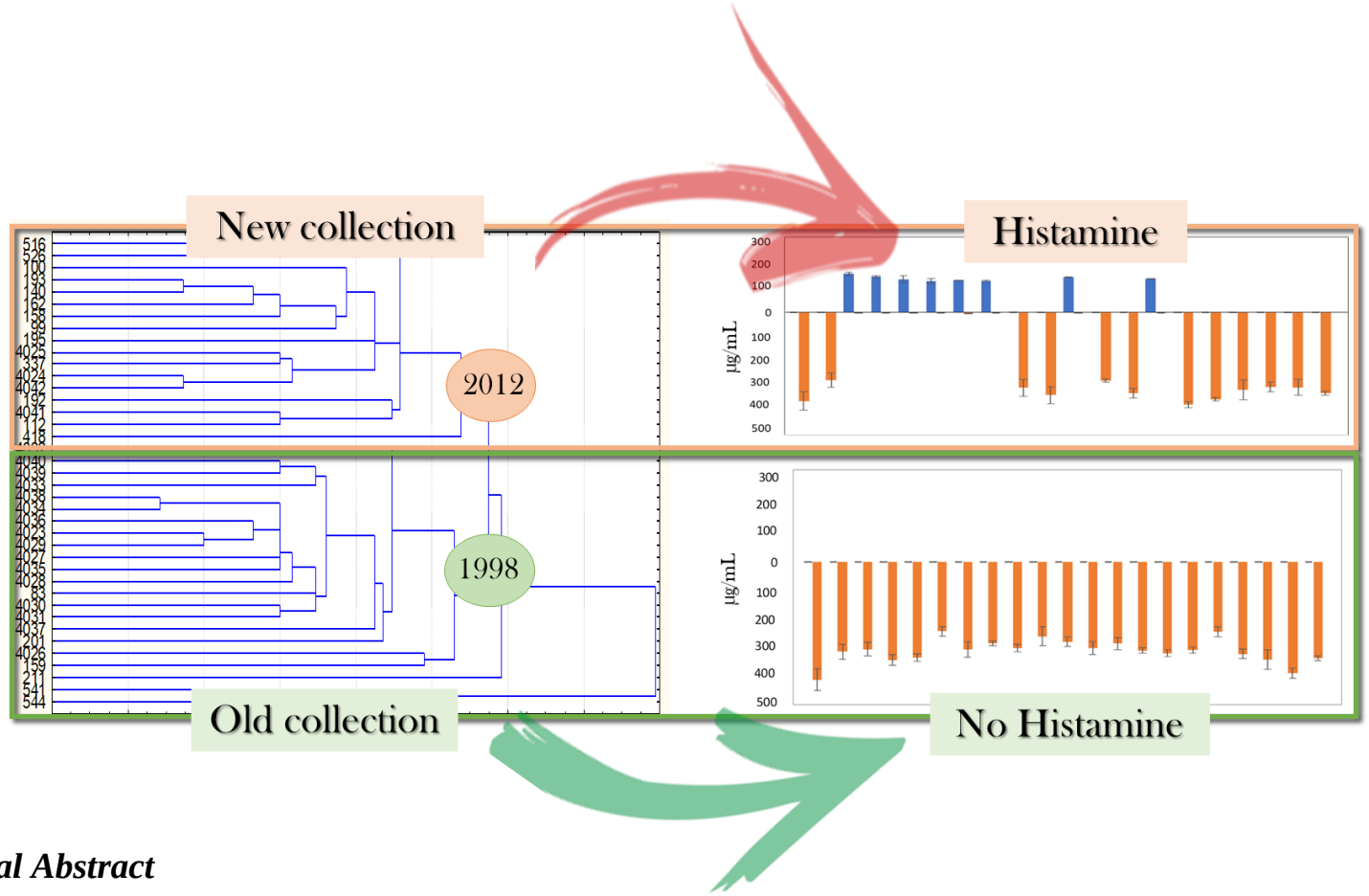
Prevalence of tyramine producing strains in Streptococcus thermophilus from Pecorino Toscano cheese

Gaia Bertani[$], Annalisa Ricci[$], Luca Calani, Camilla Lazzi, Valentina Bernini and Erasmo Neviani

[$] Both authors contributed equally to this manuscript

Food Control (Submitted)

Pecorino Toscano cheese: genetic diversity and histamine production of Streptococcus thermophilus strains over time



Graphical Abstract

Abstract

In this work the genetic diversity and the ability to produce histamine in *in vitro* conditions of 39 *Streptococcus thermophilus* strains used as starter in Pecorino Toscano PDO cheesemaking were investigated. To explore differences over time, 19 isolates in 2012 and 20 isolates in 1998 were analyzed. Amplified Fragment Length Polymorphism (AFLP) analysis revealed a wide heterogeneity and biodiversity among strains, even if the groping observed could be related to a change in genotypic profile over time. The 42% of the strains isolated in 2012 were able to metabolize histidine, converting this amino acid in histamine. Overall, the histidine was metabolized completely with a ratio of approximately 1:1. When high histidine concentrations were added, the histidine-histamine ratio decreased resulting in the conversion of less than half of the precursor, suggesting the existence of a different metabolic pathway. Interestingly, no histamine was synthesized by the strains isolates in 1998, suggesting a possible modification of Pecorino Toscano niche during the last two decades leading to the histamine producer strains. Our results suggest the relevance to take under control the biogenic ammine producer strains in starter cultures, but, at the same time, the need to supervise the genetic possible changes over time possible related to technological changes or evolutionary drivers.

Keywords: Pecorino Toscano, *Streptococcus thermophilus*, AFLP, UHPLC-ESI-MS/MS, Histamine

Introduction

Pecorino is the name given to all Italian cheeses made from sheep's milk. The composition and the microbiological traits of this product are related to the areas of production (Schirone et al., 2013), principally located in middle and southern Italy. Pecorino Toscano is a Protected Denomination of Origin (PDO) Italian cheese manufactured in the Tuscany region. Two varieties of the cheese are produced: the soft one, ripened at least 20 days, and the semi-hard one, ripened at least 120 days. The whole ewe's milk is generally pasteurized and added with mixed starter cultures, natural or selected strains belonging to *Streptococcus thermophilus* and *Lactococcus lactis* species, usually obtained from autochthonous strains isolated from milk and cheeses produced in the PDO area (Neviani et al., 1998). The microflora of Pecorino Toscano cheese is therefore predominantly characterized by starter lactic acid bacteria (SLAB) and by non-starter lactic acid bacteria (NSLAB) resistant to milk heat treatment (Neviani et al., 1998; Gobbetti et al., 2018). Furthermore, it could be underlined that sheep's raw milk usually presents many microorganisms consequently a not negligible number of microorganisms survive to pasteurization. This surviving microbiome, becoming part of the cheese ecosystem, plays an essential role during the ripening process.

The addition of selected starter is useful to reach a fast acidification in cheese. Because of this technological feature a selection of strains to be used as starter culture was suggested by PDO Protection Consortium. To maintain a link between cheese and production area the use of selected *S. thermophilus* starter strains was suggested among those naturally present in milk and curd.

This microbiota plays an important role in the biochemical and sensory characteristics of the product. The proteolytic and the decarboxylation activities can also be responsible for the production of biogenic amines (BAs), such as histamine, tyramine, putrescine, cadaverine and 2-phenylethylamine (Benkerroum et al., 2016; Schirone et al., 2012). The ingestion of BAs could cause several problems depending on the efficiency of the detoxification system of consumer,

which may be different among individuals and can be influenced by specific drugs, smoke or alcohol consumption (Torracca et al., 2015). In general, the most common symptoms associated with the ingestion of BAs are nausea, respiratory distress, hot flushes, sweating, heart palpitation, headache, bright red rush, oral burning hyper or hypo-tension (Schirone et al., 2013). Among BAs, histamine represents one of the most serious food safety concerns: its poisoning is often associated with contaminated foods consumption and particularly cheese is the most common food implicated (EFSA, 2011). In foods histamine is mainly produced by bacteria able to decarboxylate histidine (Tabanelli et al., 2012), in relation to the specificity of the strains. Consequently, it is of great importance to study the production of BAs in bacteria involved in cheese fermentation, not only to estimate the risk of BAs content but also to prevent their accumulation (Schirone et al., 2013).

The present study aimed to investigate the potential in BA production of *S. thermophilus* strains characteristic of Pecorino Toscano cheese production. In particular, the genetic diversity and the ability to produce histamine in *in vitro* conditions of different strains belonging to the microbial collection of Pecorino Toscano PDO Protection Consortium was explored in order to evaluate the characteristics of the strains over time comparing old isolates (1998) and recent ones (2012).

Materials and methods

Chemicals

L-histidine, histamine dihydrochloride as well as HPLC-grade solvents and reagents (acetonitrile, methanol and ammonium formate) were purchased from Sigma-Aldrich (St. Louis, MO, USA), while HPLC grade water and formic acid were obtained from VWR International (Milan, Italy).

Microbiological analysis

The samples, collected in different Pecorino Toscano dairy farms, were 10-fold diluted in Ringer's solution (Oxoid Ltd., Basingstoke, UK) for bacterial count on M17 agar (Oxoid Ltd., Basingstoke, UK). In the case of solid samples, 10 g were blended 1:10 in Ringer solution (Oxoid) before

dilution. After anaerobic incubation of plates at 42°C for 48h, the colonies belonging to cocci were microscopically identified (Olympus BX51TF, Olympus, Tokyo, Japan) and counted. The colony counts were performed in duplicates and the average values were calculated. Cocci were isolated through three purification steps on M17 agar medium (Oxoid) under anaerobic conditions at 42°C and, after a purity check by microscopic examination, 19 isolates were selected and included in the collection of Pecorino Toscano PDO Protection Consortium as new collection (NC) strains (isolated in 2012).

Bacterial strains

Thirty-nine strains were overall considered in this study. Nineteen strains representing the new collection (NC) and were isolated as previously reported. Twenty *S. thermophilus* strains, isolated during different steps of Pecorino Toscano PDO cheesemaking in 1998 (Neviani et al. 1998), representing the old microbial collection (OC) of the Pecorino Toscano PDO Protection Consortium. All the isolates were maintained as frozen stock cultures at −80 °C in M17 (Oxoid Ltd., Basingstoke, UK), with 15% (w/v) of glycerol. Before use, each frozen stock culture was transferred twice (2% inoculum, v/v) in M17 broth medium and incubated at 42° C for 18 hours under anaerobic conditions (Oxoid Ltd., Basingstoke, UK).

Genomic DNA extraction and identification of the new collection (NC) strains

The total genomic DNA of the 39 strains, belonging to both OC and NC, was extracted and purified from 2 ml of a fresh overnight culture. The cells were centrifuged at 12,000×g for 5 min and the pellets were resuspended in TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0). Genomic DNA was extracted using the phenol/chloroform/isoamyl alcohol (25:24:1) method described by de Los Reyes-Gavilan et al. (1992). The 19 isolates belonging to the NC were then identified by $tRNA^{Ala}$–23S rDNA-RFLP according to Mancini et al. (2012).

Genotypic characterization of *S. thermophilus* old collection (OC) and new collection (NC) strains by Amplified Fragment Length Polymorphism (AFLP) analysis

All 39 strains of *S. thermophilus* belonging to the OC and NC were genotypically characterized by AFLP analysis.

The AFLP protocol was based on the method published by Lazzi et al. (2009) using an AFLP Microbial Fingerprinting Kit (Applied Biosystem-Pe Corporation, Foster City, California, USA) with modification of the primers. The primers *EcoRI*-0/*MseI*-0 and *EcoRI*-A/*MseI*-A were used for pre-selective and selective PCR respectively, as described by Lazzi et al. (2009). Peaks representing AFLP fragments from 50 to 500 bp were reported in binary format ("1" presence and "0" absence) elaborated by GeneMapper® Software Version 4.0 (Applied Biosystem-Pe Corporation, Foster City, California, USA). Hierarchical cluster analysis was performed with Statistica 6.0 software (Statsoft Italia, Padova, Italia) using the Euclidean distance and unweighted pair group method with arithmetic mean (UPGMA) setting.

Evaluation of histamine production by *S. thermophilus* strains

The ability of the 39 *S. thermophilus* strains to metabolize histidine and to accumulate histamine was determined in M17 broth added with 0.5% (w/v) lactose (LM17) (Oxoid Ltd., Basingstoke, UK) with and without the addition of 400 µg/mL of L-histidine. Each strain was inoculated at a concentration of 6 Log CFU/mL and incubated at 42°C for 72 h in anaerobic condition. After growth the microbial cultures were centrifuged at 8000 g for 10 min at 10°C. The histidine and histamine concentration in the supernatants were measured through Ultra high performance liquid chromatography-electrospray ionization-tandem mass spectrometry (UHPLC-ESI-MS/MS) analysis. All the analyses were performed in triplicate.

UHPLC-ESI-MS/MS analysis

For method development a Xbridge BEH HILIC XP column 100x2.1 mm, 2.5 µm particle size and Acquity UPLC HSS T3 column 100x2.1 mm, 1.8 µm particle size were purchased from Waters (Milford, MA, USA), while Ultra Aqueous C18 column 100x2.1 mm, 3 µm particle size was from Restek (Bellefonte, PA, USA). The Xbridge BEH HILIC XP was used in the preliminary analyses because of high hydrophilic characteristics of histidine and histamine. The mobile phase pumped at 0.4 mL/min. with a constant 20% ammonium formate 20 mM in 1% formic acid, and water and acetonitrile both acidified with 0.2% formic acid. Unfortunately, histidine and histamine showed a very poor chromatographic efficiency with the HILIC column, with a chromatographic width close to 3 min. Any effect on peak width was obtained by increasing acetonitrile as a sample diluent in order to reduce the strength of the water as diluent in a HILIC system. Thus, a reverse phase Ultra Aqueous C18 column from Restek was tested with a binary solvent gradient of water and acetonitrile both acidified with 0.2% formic acid. This column was previously used to retain polar compounds as amino acids (Sardella et al., 2013). The starting gradient used in the reverse phase system was kept at 99% water and 1% acetonitrile ensuring the chromatographic retention of both histidine and histamine. Ultra Aqueous C18 column was compared with a HSS T3 reverse phase column, since the latter was previously used for analyses of very polar compounds (Anal Bioanal Chem (2017) 409:295–305). A similar chromatographic behaviour was obtained with Acquity UPLC HSS T3 column (Waters) with the same mobile phase used with Ultra Aqueous C18 column (Restek). However, the Acquity UPLC HSS T3 column was preferred than Ultra Aqueous C18 column because the lower particle size (1.8 µm vs 3 µm) ensured a higher chromatographic efficiency. All the samples are diluted 20-fold with 0.2% aqueous formic acid because this step improved the peak shape of histamine rather than the 5- and 10-fold dilution step (data not shown).

Histidine and histamine were analysed through a DIONEX Ultimate 3000 UHPLC coupled to a TSQ Vantage triple quadrupole (Thermo Fisher Scientific Inc., San Josè, CA, USA) fitted with a heated-electrospray ionisation (H-ESI II) probe (Thermo Fisher Scientific Inc., San Josè, CA, USA). For UHPLC, mobile phase A was 0.2% aqueous formic acid and mobile phase B was acetonitrile containing 0.2% formic acid. Separations were performed using an Acquity UPLC HSS T3 column (100 × 2.1 mm, 1.8 µm) equipped with its VanGuard pre-column (5 x 2.1 mm, 1.8 µm) (Waters, Milford, MA, USA). The gradient started with 1% B and maintaining isocratic conditions for 2 min. After 4 min B turned up to 80% keeping these conditions for 2 min before to restablish the starting gradient maintaining it 7 min to re-equilibrate the column. The flow rate was 0.25 mL/min, the injection volume was 5 µL, the autosampler temperature was set at 10°C and the column oven was set at 40 °C.

The ESI source of mass spectrometer worked in positive ionization mode. The H-ESI vaporizer temperature was 250°C, while the capillary temperature was set to 270°C. The spray voltage was 3 kV and the sheath and auxiliary gases (both nitrogen) were set to 40 and 10 (arbitrary units), respectively. The S-Lens value was 49 V. Ultra-high purity argon was used for collision energy (CE) of both histidine and histamine during the Selected Reaction Monitoring (SRM) analyses. Histidine had a protonated molecule ([M+H]+) at m/z 156.0, which produced product ions at m/z 82.9 (CE 24), m/z 92.9 (CE 22) and m/z 110.0 (CE 14). The latter product ion was the tandem mass base peak ion used as quantifier. Histamine had a [M+H]+ equal to 112.1, which produced product ions at m/z 40.9 (CE 28), m/z 67.9 (CE 21), m/z 83.0 (CE 13) and m/z 95.0 (CE 13). The latter product ion was the tandem mass base peak ion used as quantifier. The scan width was 0.2 Da while the scan time was 0.1 s. Data processing was performed using Xcalibur 2.2 software (Thermo Scientific Inc. San Josè, CA, USA). Both histidine and histamine have been quantified using

external calibration curves of their corresponding standard compounds, at concentrations ranging from 0.1 to 50 µg/mL (n= 9).

Statistical analysis

To evaluate the normally distribution for every group of independent samples Shapiro-Wilk test was used. One-way ANOVA was applied for discriminate the significant differences among the samples using Tukey test. Moreover, Pearson correlation analysis was performed to measure the strength of association and the direction of the relationship between histamine and histidine. All the analyses were conducted with SPSS Statistics 21.0 software (SPSS Inc., Chicago, IL) and the results were considered statistically different for values of $p < 0.05$ and highly different for values of $p < 0.001$.

Results and discussion

Quantification and identification of *S. thermophilus*

The samples, collected in different Pecorino Toscano dairy farms were analysed in order to quantify the thermophilic cocci by plate count on M17 agar (Oxoid) at 42°C for 48 h. Bacterial counts ranged between 9.30-9.53 Log CFU/g, in line with data reported by Schirone et al. (2012).

The 19 cocci strains isolated from plates counts were purified and the genomic DNA was extracted for their identification by tRNAAla–23S rDNA-RFLP (Mancini et al., 2012). All the strains isolated belonged to the species *S. thermophilus*.

AFLP analysis

The characterization of 39 strains of *S. thermophilus*, 20 belonging to OC and 19 to NC, was carried out by AFLP analysis. Using the best combination of fluorescently selective primers (*Eco*RIA/*Mse*IA), previously described in Lazzi et al. (2009), different fragments (ranging in size from 50-500bp) were visualized as peaks in the electropherograms. A total of 171 peaks

representing AFLP fragments was obtained and the data acquired were statistically analysed by UPGMA cluster analysis. The dendrogram revealed two clusters (A and B) while three strains (211, 541 and 544) remained unclustered showing a different genotypic profile (Figure 1). Cluster A consisted of 17 strains: 13 of them were recently isolated (2012) instead only 4 strains belonged to OC. On the contrary the cluster B was mainly composed by strains isolated in 1998 (16 out of 19). Overall a marked DNA variability was found even if the groping observed could be related to a change in genotypic profile over time.

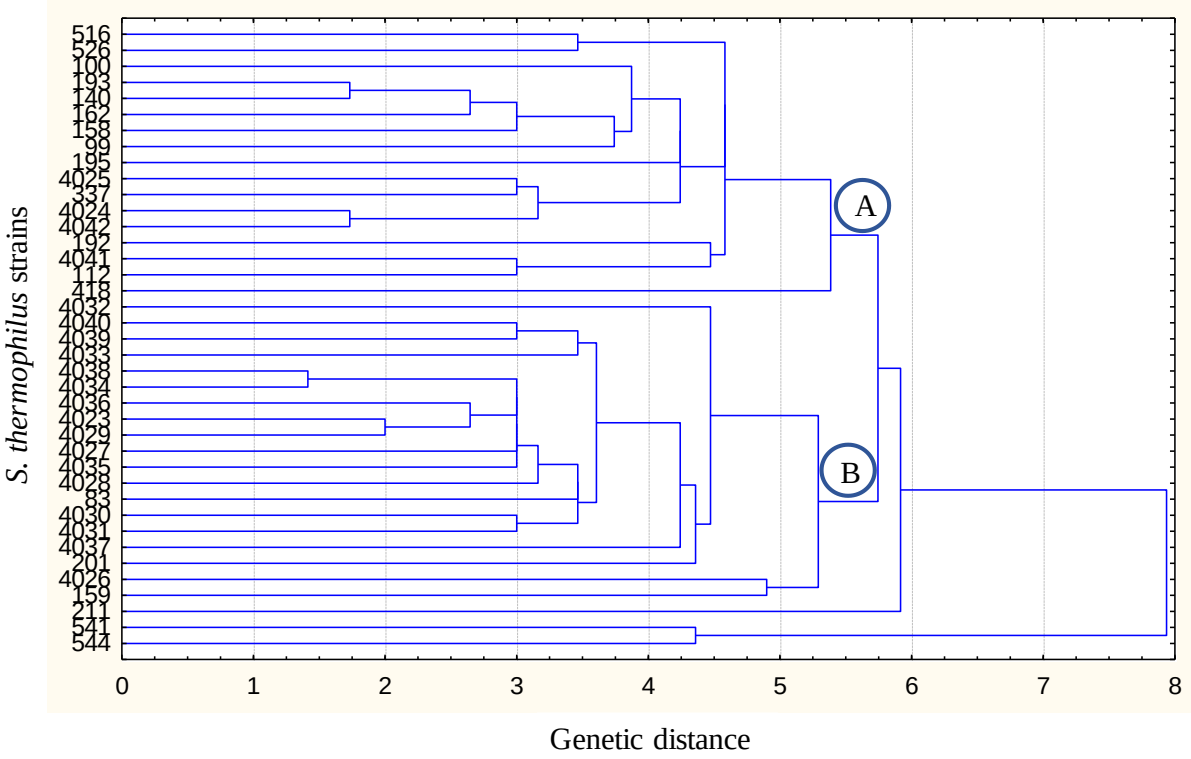


Figure 1 UPGMA dendrogram derived from AFLP analysis of the 39 *S. thermophilus* strains using the primer combinations *Eco*RI0/*Mse*I0–*Eco*RIA/*Mse*IA.

Evaluation of histamine production and histidine consumption by *S. thermophilus* strains

The ability of the 39 *S. thermophilus* strains to metabolize histidine accumulating histamine was evaluated in LM17 broth with or without the addition of 400 µg/mL of L-histidine. The bacterial growth conditions and the medium used were chosen in accordance with the best parameters found

for the highest histamine production by Tabanelli et al. (2012). In order to measure the strengths of association and the direction of the relationship between histidine and histamine Pearson's correlation analysis was performed. A negative highly significant correlation was observed (r = -0.366, p-value = 0.001), proving that histidine was consumed by *S. thermophilus* strains during fermentation while histamine was produced.

In Figure 2 (A and B) is reported the concentration of histamine and histidine in LM17 medium. The 42% of strains belonging to NC collection (140, 158, 159 162, 192, 193, 211, 516) were able to metabolize the histidine, detected in the LM17 broth (control), converting this amino acid in histamine (28.8 ± 4.8 μg/mL) (Figure 2B). Instead, the remaining 11 NC strains couldn't produce this compound. Interestingly, no histamine was produced by the OC strains (Figure 2A). Regarding histidine concentrations, in histamine producer strains generally all the histidine detected in the medium was metabolized completely with a ratio of approximately 1:1. On the other hand, some non-producer strains showed a different histidine amount (statistically significant, p value < 0.05) (Figure 2A and 2B) that could be related to the metabolic behaviour of the different strains (Goh et al., 2011; Fernàndez et al., 2006).

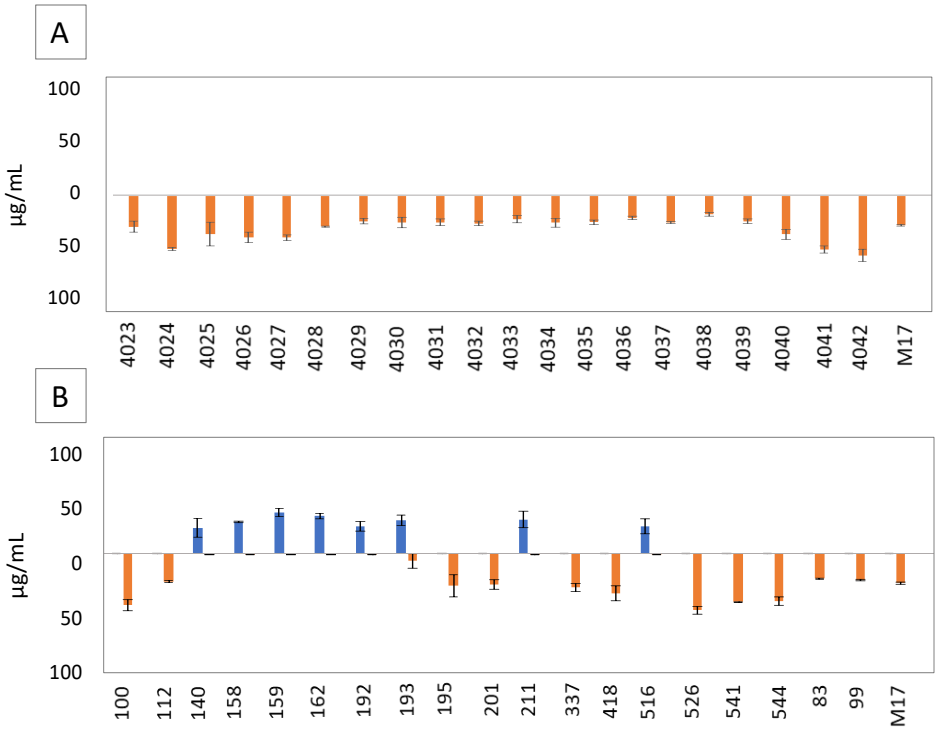


Figure 2 Histamine (blue bars) and histidine (orange bars) concentrations (µg/mL) detected in cultures of *S. thermophilus* strains belonging to the old microbial collection (OC), isolated in 1998 (A), and to the new collection (NC), isolated in 2012 (B), Pecorino Toscano PDO Protection Consortium. All the strains were cultivated in LM17 medium without L-histidine addition.

Despite the addition of histidine to LM17 medium the most of analysed strains were not able to produce histamine (Figure 3A and 3B). However, in the eight producer strains, previously mentioned, the production of histamine reached a mean value of 147.0 ± 11.4 µg/mL. It is interesting to note that high histidine concentrations led to a histidine-histamine ratio decrease resulting in the conversion of less than 50% of the precursor detected in the medium, suggesting the existence of a different metabolic pathway.

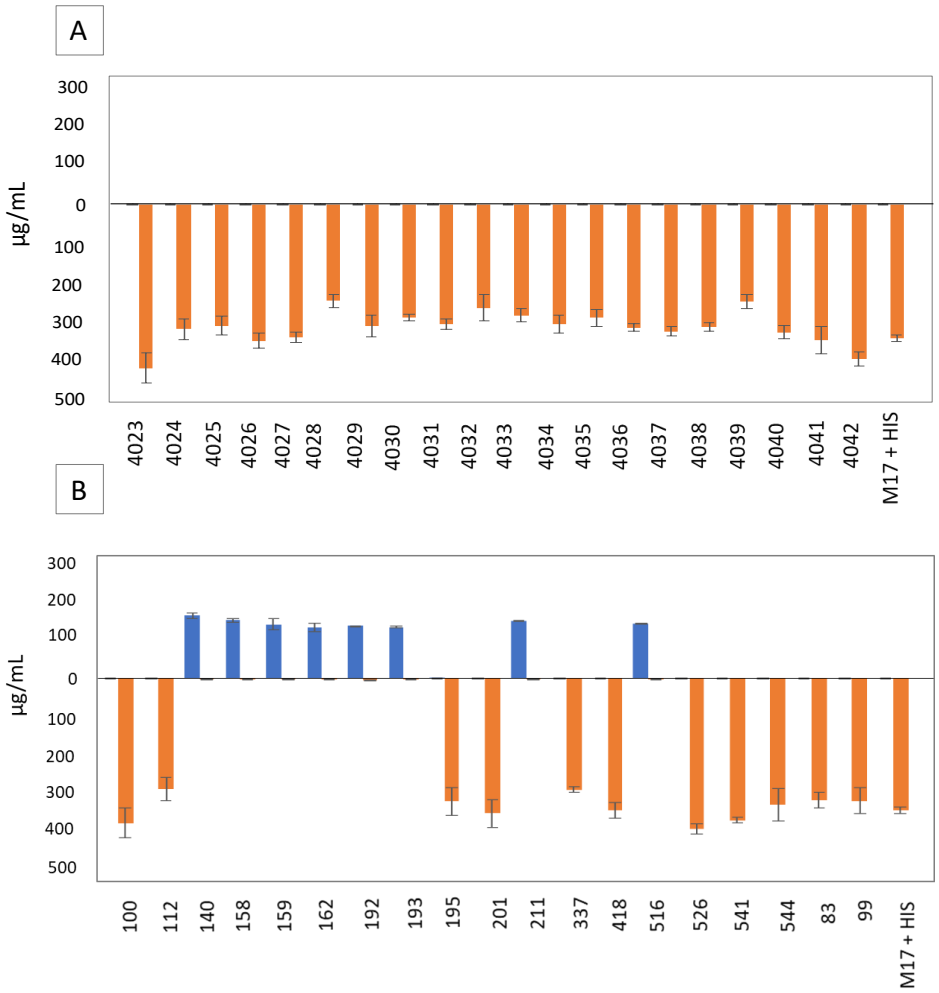


Figure 3 Histamine (blue bars) and histidine (orange bars) concentrations (µg/mL) detected in cultures of *S. thermophilus* strains belonging to the old microbial collection (OC), isolated in 1998 (A), and to the new collection, isolated in 2012 (B), Pecorino Toscano PDO Protection Consortium All the strains were cultivated in LM17 with L-histidine addition.

Indeed, previously genes *hutH* and *hutU* encoding for histidine ammonia-lyase (EC 4.3.1.3) and urocanate hydratase (EC 4.2.1.49) were found in *S. thermophilus* species, suggesting an alternative histidine catabolic pathway (Fernàndez et al., 2006), as also reported in Kyoto Encyclopedia of genes and genomes (KEGG) (Kanehisa et al., 2000). Moreover, histidine can be incorporated in proteins during their synthesis.

Eight of 39 *S. thermophilus* strains considered in this work revealed a stronger ability to synthesize histamine compared to the findings of Rossi et al. (2011), showing an increase of 41% in LM17 medium and 71% in LM17 medium added with histidine, confirming the data of Tabanelli et al. (2012). As an overall assessment, a trend, over the last twenty years, to acquire the ability to synthesize histamine could be supposed considering the tested strains. A possible explanation of this phenomena could be the supplementary energy obtained through BA production, thanks to the protonmotive force associated to the membrane, resulting in a cell's benefit (Gardini et al., 2016). The *S. thermophilus* strains in this work were all isolated from different steps of Pecorino Toscano PDO cheesemaking, disclosing that some of them could accumulate histamine. Nevertheless, this hazard is not only associated with species and matrix just mentioned, but is widely spread in fermented foods, especially in dairy products (Tabanelli et al., 2012, Gardini et al., 2016, Gezginc et al., 2013; Trip et al., 2012; Calles-Enríquez et al., 2010; Spizzirri et al., 2013; Russo et al., 2012). Among the same species, not all the strains revealed genes involved in histidine decarboxylation, suggesting strains-dependant implication in food safety. Starting from this observation the employment of selected starter cultures intends to restrict BAs accumulation in fermented products, with a particular focus on dairy products (Linares et al., 2012).

To this aim different studies have reported successful results using autochthonous strains (Renes et al., 2014; Novella-Rodríguez et al., 2002). In accordance with these conclusions the European Food Safety Agency (EFSA), approves the use of autochthonous selected starter with suitable technological profile with a low tendency to produce BA. In Pecorino Toscano cheese the presence of BAs observed in previous studies is ascribed to endogenous mesophilic microflora which survive after thermal treatment (Torracca et al., 2015; Schirone et al., 2012). Our results suggest taking under control the inclusion of BAs producer strains in starter cultures, according to Linares et al. (2012) but, above all, the need to supervise the genetic possible changes over time.

Conclusions

One of the main risks associated with fermented foods is represented by the poisoning due to histamine accumulation, gathers from bacterial histidine decarboxylase activity. For this reason, the present study was aimed to evaluate over time the potential ability of *S. thermophilus* strains, arising from Pecorino Toscano cheesemaking, to convert histidine into histamine. Even if most of the strains analysed didn't show this aptitude, the recently isolated population, characterized by different genetic profile, showed a greater potential for biogenic amines conversion. The appearance of new producer strains highlights the risk related to BAs production in the Pecorino Toscano cheese. This evidence could be due to the use of selected strains coming from other dairy niches. Moreover, it could also be related to the modification of autochthonous microflora, as a consequence of horizontal gene transfer and/or technological pressures. Considering that *S. thermophilus* selected starter have to be used for the property acidification rate during Pecorino Toscano cheesemaking, the selection of autochthonous strains should base on technological properties but also on the BAs production over time.

Acknowledgements

The authors are grateful to "Consorzio Tutela del Pecorino Toscano DOP" for samples providing.

References

Benkerroum, N. (2016). Biogenic amines in dairy products: origin, incidence, and control means. *Comprehensive reviews of Food Science and Food Safety, 15,* 801–826.

Calles-Enríquez, M., Eriksen, B.H., Andersen, P.S., Rattray, F.P., Johansen, A.H., Fernández, M., Ladero, V., Alvarez, M.A. (2010). Sequencing and transcriptional analysis of the *Streptococcus thermophilus* histamine biosynthesis gene cluster: factors that affect differential *hdcA* expression. *Applied and Environmental Microbiology, 76,* 6231–6238.

De Los Reyes-Gavilán, C.G., Limsowtin, G.K., Tailliez. P., Séchaud. L., Accolas. J.P. (1992). A *Lactobacillus helveticus*-Specific DNA Probe Detects Restriction Fragment Length Polymorphisms in This Species. *Applied and Environmental Microbiology, 58,* 3429-3432.

EFSA (2011). Scientific opinion on risk-based control of biogenic amine formation in fermented foods. *EFSA Journal, 9,* 2393–2486.

Fernández, M., del Río, B., Linares, D.M., Martín, M.C., Alvarez, M.A. (2006). Real-time polymerase chain reaction for quantitative detection of histamine-producing bacteria: use in cheese production. *Journal of Dairy Science, 89,* 3763-3769.

Gardini, F., Özogul, Y., Suzzi, G., Tabanelli, G., Özogul, F. (2016). Technological Factors Affecting Biogenic Amine Content in Foods: A Review. *Frontiers in Microbiology, 7,* 1218.

Gezginc, Y., Akyol, I., Kuley, E., Özogul, F. (2013). Biogenic amines formation in *Streptococcus thermophilus* isolated from home-made natural yogurt. *Food Chemistry, 138,* 655-662.

Gobbetti, M., Neviani, E., Fox, P. (2018). The Cheeses of Italy: Science and Technology. (1st ed.) *Ed. Springer International*, (Chapter 6).

Goh, Y.J., Goin, C., O'Flaherty, S., Altermann, E., Hutkins, R. (2011). Specialized adaptation of a lactic acid bacterium to the milk environment: the comparative genomics of *Streptococcus thermophilus LMD-9. Microbial Cell Factories, 10,* Suppl 1: S22.

Kanehisa, M., Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research, 28,* 27-30.

Lazzi, C., Bove, C. G., Sgarbi, E., Gatti, M., La Gioia, F., Torriani, S., Neviani, E. (2009). Application of AFLP fingerprint analysis for studying the biodiversity of *Streptococcus thermophilus. Journal of Microbiological Methods, 79,* 48-54.

Linares, D.M., del Río, B., Ladero, V., Martínez, N., Fernández, M., Martín, M.C., Alvarez, M.A. (2012). Factors influencing biogenic amines accumulation in dairy products. *Frontiers in Microbiology, 3,* 180.

Mancini, A., Lazzi, C., Bernini, V., Neviani, E., Gatti, M. (2012). Identification of dairy lactic acid bacteria by tRNAAla-23S rDNA-RFLP. *Journal of Microbiological Methods, 91,* 380-390.

Neviani, E., Bizzarro, R., Righini, A., Mucchetti, G. (1998). Pecorino toscano DOP: tecniche di produzione e caratteristiche microbiologiche. (1st ed.) *Tecniche nuove,* (Chapter 6).

Novella-Rodríguez, S., Veciana-Nogués, M., Roig-Sagués, A., Trujillo-Mesa, A., Vidal-Carou, M. (2002). Influence of starter and non-starter on the formation of biogenic amine in goat cheese during ripening. *Journal of Dairy Science, 85,* 2471-2478.

Renes, E., Diezhandino, I., Fernandez, D., Ferrazza, R. E., Tornadijo, M.E., and Fresno, J.M. (2014). Effect of autochthonous starter cultures on the biogenic amine content of ewe's milk cheese throughout ripening. *Food Microbiology, 44,* 271–277.

Rossi, F., Gardini, F., Rizzotti, L., La Gioia, F., Tabanelli, G., Torriani S. (2011). Quantitative analysis of histidine decarboxylase gene (*HdcA*) transcription and histamine production by *Streptococcus thermophilus* PRI60 under conditions relevant to cheese making. *Journal of Applied Environmental Microbiology, 77,* 2817-2822.

Russo, P., de Palencia, P. F., Romano, A., Fernández, M., Lucas, P., Spano, G., López, P. (2012). Biogenic amine production by the wine *Lactobacillus brevis* IOEB 9809 in systems that partially mimic the gastrointestinal tract stress. *BMC Microbiology, 12,* 247.

Sardella, R., Lisanti, A., Marinozzi, M., Ianni, F., Natalini, B., Blanch, G.P., del Castillo, M.L.R. (2013). Combined monodimensional chromatographic approaches to monitor the presence of D-amino acids in cheese. *Food Control, 34,* 478-487.

Schirone, M., Tofalo, R., Fasoli, G., Perpetuini, G., Corsetti, A., Manetta, A. C., Ciarrocchi, A., Suzzi, G. (2013). High content of biogenic amines in Pecorino cheeses. *Food Microbiology, 34,* 137-144.

Schirone, M., Tofalo, R., Visciano, P., Corsetti, A., Suzzi, G. (2012). Biogenic amines in Italian Pecorino cheese. *Frontiers in Microbiology, 3,* 171.

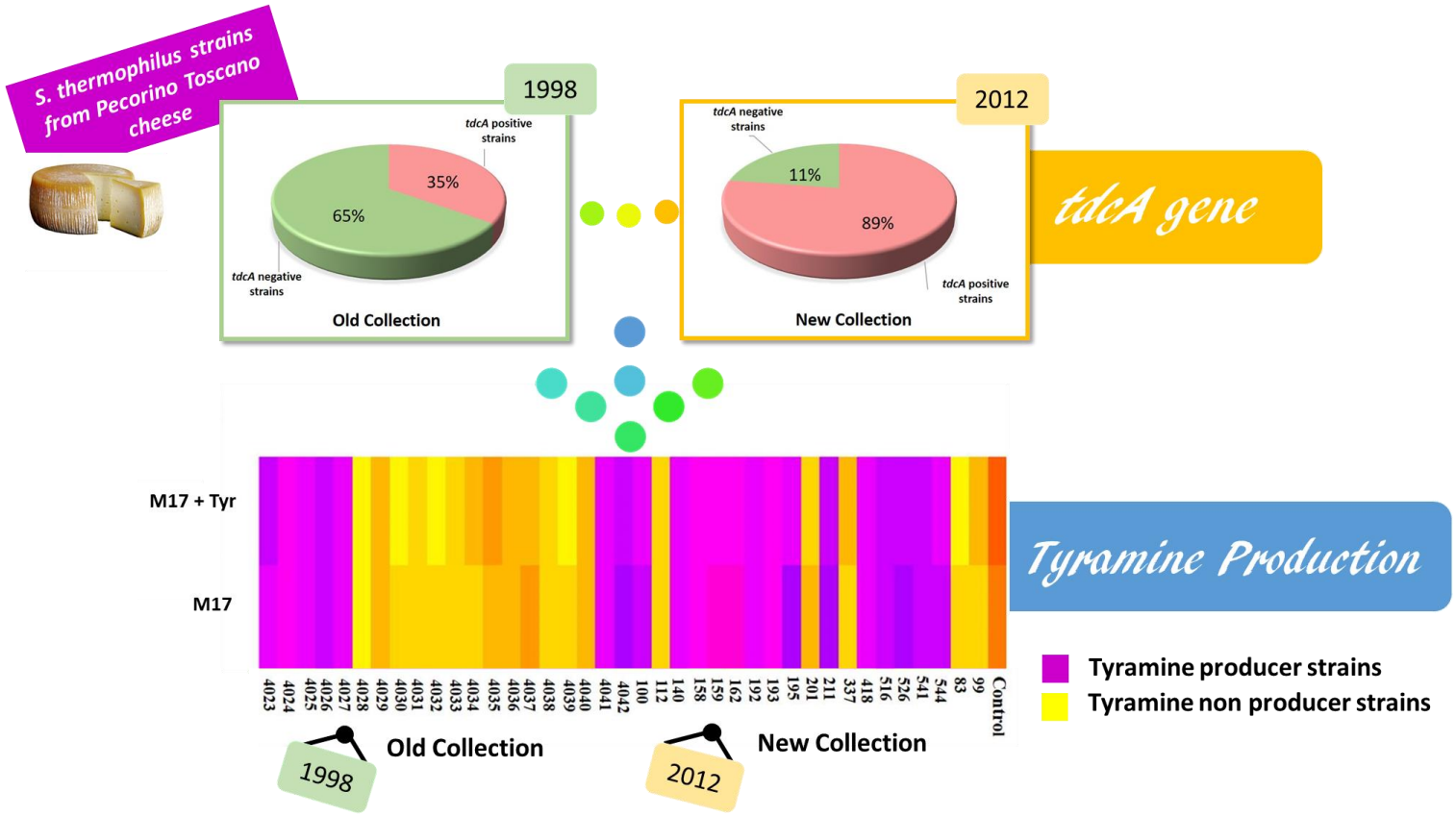
Spizzirri, U.G., Restuccia, D., Curcio, M., Parisi, O. I., Iemma, F., and Picci, N. (2013). Determination of biogenic amines in different cheese samples by LC with evaporative light scattering detector. *Journal of Food Composition and Analysis, 29,* 43–51.

Tabanelli, G., Torriani, S., Rossi, F., Rizzotti, L., Gardini, F. (2012). Effect of chemico-physical parameters on the histidine decarboxylase (*hdcA*) enzymatic activity in *Streptococcus thermophilus* PRI60. *Journal of Food Science, 77,* 231–237.

Torracca, B., Nuvoloni, R., Pedonese, F., Bacci, C., Ducci, M. (2015). Biogenic Amines Content of Four Types of "Pecorino" Cheese Manufactured in Tuscany. *International Journal of Food Properties, 18,* 999-1005.

Trip, H., Mulder, N.L., and Lolkema, J.S. (2012). Improved acid stress survival of *Lactococcus lactis* expressing the histidine decarboxylation pathway of *Streptococcus thermophilus* CHCC1524. *The Journal of Biological Chemistry, 287,* 11195–11204.

Prevalence of tyramine producing strains in *Streptococcus thermophilus* from Pecorino Toscano cheese



Graphical Abstract

Abstract

Several microorganisms belonging to lactic acid bacteria (LAB) group are responsible for tyramine accumulation in dairy foods. To understand the *in vitro* production of tyramine by *S. thermophilus*, a species widely employed as starter, thirty-nine strains isolated from Pecorino Toscano cheese in different years were evaluated. The old collection (1998) and the new collection (2012) were considered to investigate the presence of *tdcA* gene and the tyramine accumulation in M17 medium. Many strains revealed the presence of the *tdcA* gene, but for some of them the production of tyramine was not detected by UHPLC-ESI-MS/MS. Overall, the 35% of old strains and the 74% of new strains were able to produce tyramine. The tyramine production observed after fermentation was higher than the amount of precursor in the control suggesting an additional source of tyrosine probably released by the proteolytic system activity.

Keywords: Pecorino Toscano cheese, *Streptococcus thermophilus*, *tdcA* gene, UHPLC-ESI-MS/MS, Tyramine

Introduction

Biogenic amines (BAs) are low-molecular-weight organic compounds, which play important roles as neurotransmitters in the nervous system or as vasoactive substances in the vascular system (Benkerroum, 2016). The most important BAs in dairy foods are histamine, tyramine, putrescine, cadaverine, *β*-phenylethylamine and tryptamine (Gardini et al., 2016). Tyramine, formed by the decarboxylation of tyrosine, acts in the central nervous system as a releasing agent of cathecolamines including adrenaline, noradrenaline and dopamine (Benkerroum, 2016). When the amount of tyramine increases, this compound causes adverse effects on human health like migraine and hypertensive crisis (EFSA, 2011). Among the microorganisms involved in BAs production several lactic acid bacteria (LAB) can be responsible for tyramine accumulation in dairy foods (Gatto et al., 2016). Two main favourable conditions make cheese an ideal substrate for the production and the accumulation of BAs by LAB. First of all, the high developed proteolytic system of LAB able to hydrolyse caseins in free amino acid (Schirone et al., 2012) and, in addition, the low pH generated by LAB in the matrix supporting the decarboxylase gene transcription and enzyme activity (Ladero et al., 2015). Over recent years several studies have investigated different dairy products and their microbiota responsible for the production of biogenic amines (Ladero et al., 2015; Guarcello et al., 2016; Ylmaz et al., 2017). Regarding Italian products, Pecorino cheeses are widely spread: they are made from sheep's milk and many varieties can be recognized in relation to the areas of production. Among these, Pecorino Toscano is a Protected Denomination of Origin (PDO) Italian cheese manufactured in the Tuscany region. During the manufacturing process whole milk is generally pasteurized and autochthonous strains belonging to *Streptococcus thermophilus* and *Lactococcus lactis*, isolated from milk and cheeses produced in the PDO area, are used as starter cultures (Neviani et al., 1998, Gobbetti et al., 2018). The cheese microbiota is composed by

mesophilic and thermophilic strains, coming from sheep raw milk, which survive to heat treatment and starter strains.

The starter cultures play a key role in cheese acidification and flavoring, so a steered selection of strains, among those usually naturally present in milk and curd, was suggested by PDO Protection Consortium.

Beyond the technological traits the starter strains could also generate toxic metabolites, such as BAs, which can be accumulated in the final product constituting a risk to consumer health.

Therefore, this study aimed to shed light on the production of tyramine by *S. thermophilus* strains belonging to the microbial collection of Pecorino Toscano Consortium. Differently from other works, the aptitude to convert tyrosine into tyramine *in vitro* condition was evaluated over time, considering strains isolated in 1998 and in 2012. In particular, 39 autochthonous strains of *S. thermophilus* were first screened by PCR in order to verify the presence of *tdcA* gene encoding for tyrosine decarboxylase and the amount of tyramine produced was detected by UHPLC-ESI-MS/MS.

Materials and methods

Chemicals

Tyramine, L-tyrosine as well as HPLC-grade solvents and reagents (acetonitrile, methanol and ammonium formate) were purchased from Sigma-Aldrich (St. Louis, MO, USA), while HPLC-grade water and formic acid were obtained from VWR International (Milan, Italy).

Bacterial strains

Thirty-nine *S. thermophilus* strains were analysed. Twenty of these, previously identified by 16S rRNA gene sequencing, belong to the original microbial collection (OC) of the "Pecorino Toscano PDO Protection Consortium" and have been isolated during different Pecorino Toscano PDO cheesemaking in 1998 (Neviani et al., 1998). The remaining 19 strains belong to the new collection

(NC). These strains have been isolated in different Pecorino Toscano dairy farms from different manufacturing steps by the Department of Food Science (University of Parma, Parma, Italy) in 2012. All isolates were maintained as frozen stock cultures at −80 °C in M17 (Oxoid Ltd., Basingstoke, UK), with 15% (v/v) of glycerol. Before use, each frozen stock culture was transferred twice (2% inoculum, v/v) in the culture medium and incubated at 42° C for 18 hours under anaerobic conditions (Oxoid Ltd., Basingstoke, UK).

Detection of *tdcA* gene

Genomic DNA extraction was carried out using DNeasy Blood & Tissue Kit (Qiagen, Milan, Italy). Primers TDC1 (5'AACTATCGTATGGATATCAACG3') and TDC2 (5'TAGTCAACCATATTGAAATCTGG3') were used for the amplification of *tdcA* gene as described by Guarcello et al. (2016). Amplification was carried out in a 20 µl final volume 1X GoTaq® Colorless Master Mix (PROMEGA Madison, Wisconsin, US), 0.5 µM of each primer and 1 µl of template DNA. PCR amplifications were performed, in accordance with Guarcello et al. (2016), in a conventional thermocycler (GeneAmp® PCR System 2700 Applied Biosystem) with the following thermal cycling program: 2 min at 94 °C, 35 cycles (30 s at 94 °C, 40 s at 52 °C, 30 s at 72 °C), 5 min extension at 72°C. The PCR products were separated on a 1% agarose gel in 1×TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.2) and visualized by UV transillumination after staining with SYBR™ Safe DNA Gel Stain (Invitrogen Life Technologies, Carlsbad, California, USA). A GeneRuler 1Kb plus DNA ladder (Invitrogen Life Technologies, Carlsbad, California, USA) was used as a molecular size standard.

Evaluation of tyramine production by UHPLC-ESI-MS/MS analyses

The ability of the 39 *S. thermophilus* strains to metabolize tyrosine and to accumulate tyramine was determined in M17 broth added with 0.5% (w/v) lactose (LM17) (Oxoid Ltd., Basingstoke, UK) with or without the addition of 400 µg/mL of L-tyrosine. All the analyses were performed in

triplicate and each strain was inoculated at a concentration of 6 Log CFU/mL and incubated at 42°C for 72 h in anaerobic condition. After growth the microbial cultures were centrifuged at 8000 g for 10 min at 10°C. The supernatants were stored at – 80°C and then diluted 10-fold with acetonitrile before UHPLC-MS/MS analyses.

The tyrosine and tyramine concentration have been evaluated through a DIONEX Ultimate 3000 UHPLC coupled to a TSQ Vantage triple quadrupole (Thermo Fisher Scientific Inc., San Josè, CA, USA) fitted with a heated-electrospray ionization (H-ESI II) probe (Thermo Fisher Scientific Inc., San Josè, CA, USA). For UHPLC, mobile phase A was methanol/water 50:50 (v:v), mobile phase B was acetonitrile containing 0.2% formic acid, while mobile phase C was ammonium formate 20 mM containing 1% of formic acid. Separations were performed using an XBridge® BEH HILIC column (2.1 x 100mm, 2.5 μm particle size) (Waters, Milford, MA, USA). The mobile phase C containing the buffer eluent was maintained constant during the whole analysis at 16%. The gradient started with 1% A and maintaining isocratic conditions for 2 min, and then A turned up to reach 63% at 4 min before to re-establish the starting gradient maintaining it 5 min to re-equilibrate the column. The flow rate was 0.4 mL/min, the injection volume was 5 μL, the autosampler temperature was set at 10°C and the column oven was set at 35 °C.

The ESI source of mass spectrometer worked in positive ionization mode. The H-ESI vaporizer temperature was 250°C, while the capillary temperature was set to 270°C. The spray voltage was 3.2 kV while the sheath and auxiliary gases (both nitrogen) were set to 50 and 15 (arbitrary units), respectively. The S-Lens value was 48 V. Ultra-high purity argon was used for collision energy (CE) of both tyrosine and tyramine during the Selected Reaction Monitoring (SRM) analyses. Tyrosine had a positively charged molecular ion ([M+H]+) at m/z 182.0, which produced product ions at m/z 77.0 (CE 38), m/z 91.0 (CE 32), m/z 119.0 (CE 19), m/z 122.9 (CE 19) and m/z 136.0 (CE 19). The product ion at m/z 122.9 was the tandem mass base peak ion used as quantifier.

Tyramine had a [M+H]+ equal to 138.1, which produced product ions at m/z 77.0 (CE 30), m/z 91.0 (CE 12) and m/z 121.0 (CE 12). The latter product ion was the tandem mass base peak ion used as quantifier. The scan width was 0.2 Da while the scan time was 0.1 s. Data processing was performed using Xcalibur 2.2 software (Thermo Scientific Inc. San Josè, CA, USA). Both tyrosine and tyramine were quantified using external calibration curves of their corresponding standard compounds.

Statistical analyses

To evaluate the normal distribution for every group of independent samples Shapiro-Wilk test was used. One-way ANOVA was applied to discriminate the significant differences among the samples using Tukey test. Moreover, Pearson correlation analysis was performed to measure the strength of association and the direction of the relationship between tyrosine and tyramine. All the analyses were conducted with SPSS Statistics 21.0 software (SPSS Inc., Chicago, IL) and the results were considered statistically different for values of $p < 0.05$.

Results and discussion

Detection of *tdcA* gene

Thirty-nine *S. thermophilus* strains were screened by PCR using the primers pairs TDC1 and TDC2 targeting the gene *tdcA* in order to evaluate the potential ability to metabolize tyrosine into tyramine. Twenty-four strains (62% of the strains) presented the *tdcA* gene and, in particular, 7 of them were isolated in 1998 (35% of the OC) instead 17 (89% of the NC) were recently isolated (Table 1 and Figure 1).

Table 1 Strains of *S. thermophilus* analysed, year of isolation and presence (+) / absence (-) of the gene *tdcA* encoding for tyrosine decarboxylase.

[1]Abbreviations: OC (old collection 1998), NC (new collection 2012).

Strains	***Collection[1]***	***tdcA gene***
4023	OC	+
4024	OC	+
4025	OC	+
4026	OC	+
4027	OC	+
4028	OC	-
4029	OC	-
4030	OC	-
4031	OC	-
4032	OC	-
4033	OC	-
4034	OC	-
4035	OC	-
4036	OC	-
4037	OC	-
4038	OC	-
4039	OC	-
4040	OC	-
4041	OC	+
4042	OC	+
100	NC	+
112	NC	+
140	NC	+
158	NC	+
159	NC	+
162	NC	+
192	NC	+
193	NC	+
195	NC	+
201	NC	-
211	NC	+
337	NC	+
418	NC	+
516	NC	+
526	NC	+

The prevalence of *tdc*A gene inside *S. thermophilus* species is rather relevant compared to what have been previously observed. La Gioia et al. (2011) have targeted the same gene with a different primer pair, observing the presence of only 1% among *S. thermophilus* dairy strains tested. Instead, focusing on Pecorino cheeses, Guarcello et al. (2016) has reported the presence of *tdcA* gene in 29% of the strains. Our study is the first about the tyramine production in *S. thermophilus* strains used as starter microflora in Pecorino Toscano cheese and compared with literature our results highlighted the highest prevalence of this gene in this niche.

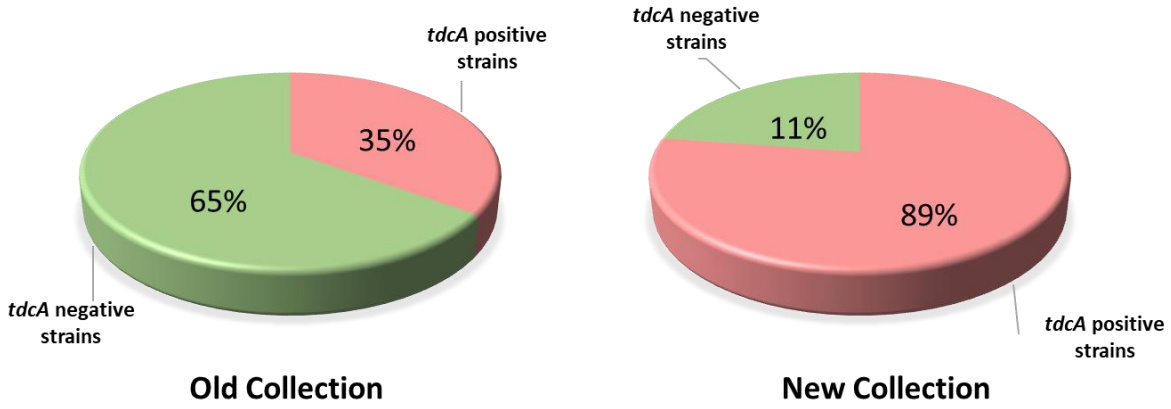


Figure 1 Pie chart showing prevalence of *tdcA* gene in *S. thermophilus* strains, belonging to old collection (1998) and new collection (2012).

TdcA gene, encoding for tyrosine decarboxylase, belong to *tdc* cluster in which take part also *tyrP*, a putative tyrosine-tyramine antiporter, and *tyrS*, a tyrosyl-tRNA synthetase gene (Fernández et al., 2004; Cruz Martín et al., 2011). The *tdc* cluster has been characterized in different species of LAB such as *Lactobacillus brevis* (Lucas et al., 2003), *Enterococcus faecalis* (Connil et al., 2002), *S. thermophilus* and *Lactobacillus curvatus* (La Gioia et al., 2011). Different studies have suggested that *tdcA* could be involved in horizontal gene transfer, at both intraspecies and interspecies level, such as between *L. curvatus* and *S. thermophilus* (La Gioia et al., 2011; Coton & Coton, 2009).

Therefore, the tyramine production risk could be extended to LAB involved in dairy manufacturing sharing the same ecosystem.

Evaluation of tyramine production and tyrosine consumption *in vitro*

The *in vitro* production of tyramine *via* decarboxylation of tyrosine was evaluated through UHPLC-ESI-MS/MS analyses. Thirty-nine *S. thermophilus* strains were analysed in LM17 broth with or without the addition of 400 µg/mL of L-tyrosine to study their ability to convert tyrosine into tyramine. These two conditions were tested i) to quantify the conversion of the precursor into the corresponding BA under optimal growth condition ii) to verify the maximum concentration of tyramine reached in non-limiting precursor conditions. Results from this study clearly indicated that *S. thermophilus* strains isolated from Pecorino Toscano cheese, to our knowledge never studied previously, have a great potential to produce tyramine and this phenotype is most widespread comparing to the past. Monitoring by UHPLC-ESI-MS/MS analyses the precursor, tyrosine, and the end-product, tyramine, we confirmed the effective conversion that has been reinforced by Pearson's correlation analysis ($r = -0.503$, $p<0.001$).

Overall, a total of 21 strains (54%) resulted to be tyramine producers. Seven strains belonging to OC (35%) were able to metabolize tyrosine detected in the LM17 broth (control). In the samples where these strains were grown for 72 h at 42°C, the concentration of tyramine observed was 49.62 ± 5.07 µg/mL (mean value), as a result of the total consume of the amino acid (Table 2). The remaining 13 OC strains did not produce tyramine. Among the NC, 14 strains (74%) were able to produce tyramine reaching a concentration of 58.14 ± 5.79 µg/mL (mean value). These results indicated that the decarboxylation of tyrosine can occur also in media without precursor addition, as previously reported by Bargossi et al. (2015). As expected, LM17 medium supplemented with L-tyrosine led to an increase of the tyramine concentration, reaching 248.39 ± 9.39 µg/mL (mean value) in OC strains and 257.85 ± 7.15 µg/mL (mean value) in NC (Table 2).

Table 2 Tyramine and tyrosine concentrations (μg/mL) observed in *S. thermophilus* strains belonging to original microbial collection (OC) of the "Consorzio per la Tutela del Pecorino Toscano" isolated in 1998 and to the new collection (NC) isolated in 2012. All the strains were cultivated in LM17 medium with or without tyrosine addition. Mean values within the column with different letters are statistically significant different ($P < 0.05$).

	Strains	*LM17 medium*		*LM17+Tyr. medium*	
		Tyrosine	*Tyramine*	*Tyrosine*	*Tyramine*
OLD COLLECTION	**4023**	1.0533 ± 0.0607^{a}	48.7393 ± 1.6293^{cde}	0.5819 ± 0.0703^{a}	228.9797 ± 5.7477^{de}
	4024	0.9240 ± 0.1009^{a}	60.2815 ± 5.4572^{fg}	1.3212 ± 0.0656^{a}	291.5566 ± 10.1381^{il}
	4025	0.9262 ± 0.7921^{a}	52.7223 ± 8.0634^{ef}	0.7695 ± 0.0656^{a}	262.9108 ± 2.5528^{gh}
	4026	0.5955 ± 0.4345^{a}	44.2261 ± 2.0237^{bcde}	0.7073 ± 0.0361^{a}	205.2189 ± 8.6407^{bcd}
	4027	1.2237 ± 0.0140^{a}	51.5032 ± 3.2320^{ef}	0.7031 ± 0.0493^{a}	270.5144 ± 9.2540^{hi}
	4028	32.0857 ± 0.7261^{h}	0.1054 ± 0.0024^{a}	152.9933 ± 13.1017^{hi}	0.4240 ± 0.0177^{a}
	4029	21.6016 ± 1.6956^{cd}	0.0727 ± 0.0057^{a}	125.6365 ± 7.6646^{bcdef}	0.2945 ± 0.0114^{a}
	4030	27.7974 ± 3.9167^{fgh}	0.1006 ± 0.0147^{a}	140.4854 ± 14.9170^{fgh}	0.4121 ± 0.0111^{a}
	4031	22.3318 ± 1.1942^{cde}	0.0935 ± 0.0050^{a}	129.6132 ± 10.2855^{def}	0.3746 ± 0.0172^{a}
	4032	26.7569 ± 6.1103^{efg}	0.0920 ± 0.0228^{a}	138.4196 ± 14.5734^{fgh}	0.4229 ± 0.0227^{a}
	4033	25.4599 ± 5.7908^{defg}	0.0973 ± 0.0263^{a}	133.9147 ± 8.1784^{efg}	0.3462 ± 0.0167^{a}
	4034	23.2413 ± 0.5613^{def}	0.0878 ± 0.0117^{a}	135.8711 ± 7.2411^{fgh}	0.3365 ± 0.0223^{a}
	4035	25.6641 ± 1.8379^{defg}	0.0847 ± 0.0183^{a}	148.5204 ± 12.9736^{ghi}	0.2707 ± 0.0268^{a}
	4036	18.0332 ± 3.3816^{bc}	0.0723 ± 0.0128^{a}	116.5265 ± 7.3408^{bcde}	0.3193 ± 0.0256^{a}
	4037	15.5021 ± 2.3269^{b}	0.0636 ± 0.0154^{a}	107.2296 ± 3.4848^{b}	0.3130 ± 0.0096^{a}
	4038	23.6199 ± 0.3789^{def}	0.1030 ± 0.0066^{a}	127.9435 ± 9.4143^{cdef}	0.3897 ± 0.0194^{a}
	4039	28.9708 ± 2.2685^{gh}	0.1022 ± 0.0076^{a}	159.4654 ± 14.6247^{il}	0.4333 ± 0.0326^{a}
	4040	25.0865 ± 0.1688^{defg}	0.0802 ± 0.0020^{a}	97.2229 ± 2.5722^{bcd}	0.3176 ± 0.0078^{a}
	4041	1.2254 ± 0.1931^{a}	49.9926 ± 2.7290^{de}	0.8511 ± 0.1102^{a}	237.1920 ± 3.1030^{ef}
	4042	0.9841 ± 0.2581^{a}	38.9494 ± 4.6924^{bc}	0.4299 ± 0.0266^{a}	212.0716 ± 1.1889^{bcd}
NEW COLLECTION	**100**	1.0205 ± 0.0339^{a}	44.4680 ± 4.5395^{bcde}	0.7713 ± 0.0461^{a}	247.0991 ± 2.7890^{efgh}
	112	25.1315 ± 1.0986^{defg}	0.0944 ± 0.0026^{a}	135.1257 ± 10.2994^{efgh}	0.3607 ± 0.0284^{a}
	140	0.4982±0.0007^{a}	49.7829 ± 5.2753^{de}	1.1565 ± 0.0988^{a}	251.5352 ± 6.6446^{efgh}
	158	1.5962 ± 0.9477^{a}	68.3653 ± 1.8439^{gh}	1.1082 ± 0.1615^{a}	288.2487 ± 8.9858^{i}
	159	1.6901 ± 0.9287^{a}	81.3898 ± 9.4400^{i}	1.4703 ± 0.0716^{a}	314.6611 ± 12.2465^{l}
	162	0.8163 ± 0.1232^{a}	74.5478 ± 2.3436^{hi}	1.1316 ± 0.0883^{a}	293.0830 ± 9.1633^{il}
	192	1.2497 ± 0.5913^{a}	53.3178 ± 10.8321^{ef}	0.8262 ± 0.1036^{a}	240.8252 ± 8.7757^{efg}
	193	1.1136 ± 0.1736^{a}	67.5999 ± 5.0718^{gh}	1.0590 ± 0.1576a	289.7161 ± 9.9014^{i}
	195	0.5793 ± 0.1077^{a}	34.8956 ± 2.1386^{b}	1.2796 ± 0.0625^{a}	250.1136 ± 1.8388^{efgh}
	201	27.7005 ± 1.6199^{fgh}	0.0822 ± 0.0080^{a}	143.8138 ± 14.9102^{fghi}	0.3614 ± 0.0134^{a}
	211	0.6627 ± 0.4694^{a}	36.8250 ± 1.7614^{b}	0.5949 ± 0.0624^{a}	201.0379 ± 2.0554^{b}
	337	21.8077 ± 0.4523^{cd}	0.0874 ± 0.0087^{a}	109.4997 ± 7.6573^{bc}	0.2939 ± 0.0068^{a}
	418	1.3124 ± 0.4364^{a}	49.7025 ± 16.7351^{de}	0.8226 ± 0.0345^{a}	249.9254 ± 7.0848^{efgh}
	516	0.4106 ± 0.0597^{a}	39.8931 ± 1.6717^{bcd}	0.7315 ± 0.0711^{a}	203.1071 ± 3.1285^{bc}
	526	0.7749 ± 0.1934^{a}	38.6988 ± 0.1422^{bc}	0.5466 ± 0.0601^{a}	227.1924 ± 1.5393^{cde}
	541	1.4486 ± 0.3395^{a}	43.3249 ± 0.2786^{bcde}	0.6482 ± 0.0603^{a}	237.0207 ± 8.2406^{ef}
	544	1.3868 ± 0.1118^{a}	40.4282 ± 6.2069^{bcd}	0.6572 ± 0.0553^{a}	260.5293 ± 5.7209^{fgh}
	83	23.7546 ± 4.9882^{def}	0.0911 ± 0.0159^{a}	149.5493 ± 13.9495^{ghi}	0.4662 ± 0.0392^{a}
	99	23.4937 ± 1.6226^{def}	0.0962 ± 0.0136^{a}	125.9478 ± 5.8591^{bcdef}	0.3286 ± 0.0390^{a}
	M17	25.5426 ± 0.8365^{defg}	0.0548 ± 0.0066^{a}	-	-
	M17+Tyr.	-	-	176.7941 ± 11.1396^{l}	0.1681 ± 0.0054^{a}

To note, in LM17 added with tyrosine (control), after sterilisation, a reduction of tyrosine was observed (approximately 223 µg/mL), probably due to the Strecker degradation and/or the oxidation of the same amino acid (Bergès et al., 2011). Since the tyramine production observed after fermentation was higher than the amount of precursor in the control, we can assume that an additional source of tyrosine was consumed. This because *S. thermophilus* proteolytic system can increase the source of tyrosine in the substrate that could be subsequently metabolized (Hafez et al., 2015). The great differences observed in the proteolytic activity of *S. thermophilus* strains (Cui et al., 2016) can also explain the diverse concentrations, resulted statistically different ($p < 0.05$), of tyramine that we observed for our producer strains (Figure 2).

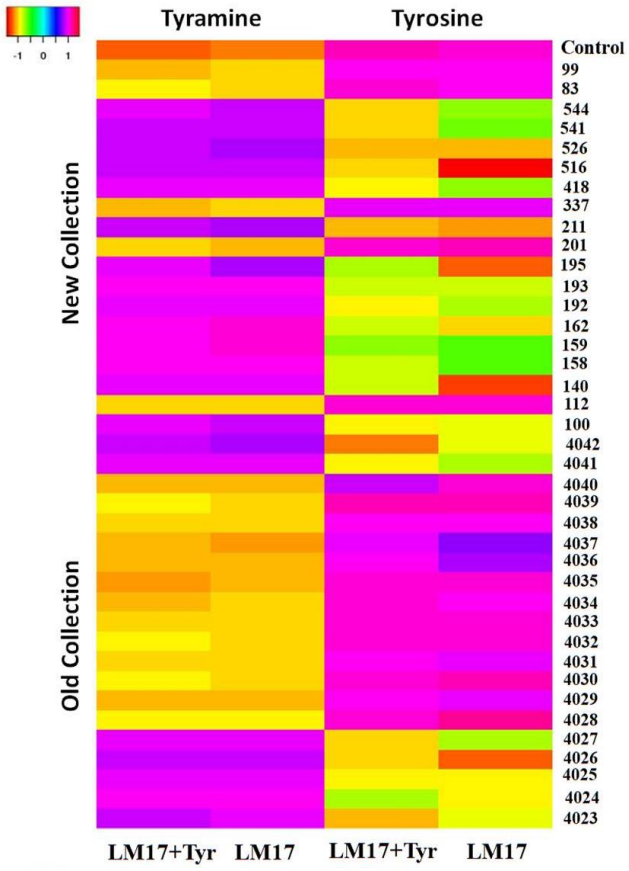


Figure 2 Hierarchical clustering and heat map based on tyrosine and tyramine concentration observed after *S. thermophilus* fermentation in LM17 with or without the addition of L-tyrosine.

It should be noted that, in addition to the high prevalence, the amount of amine produced is greater than what is reported in other studies; indeed, while La Gioia et al. (2011) found one among 83 *S. thermophilus* strains able to produce less than half of the tyrosine added in the medium, in this study we observed 21 strains able to produce more tyramine than the tyrosine detected. Among the fourteen non-producers, different concentrations of precursor were detected ($p< 0.05$), pointing out that tyrosine can be catabolized through an alternative pathway, like the conversion by an aminotransferase into the corresponding α-keto acid (4-Hydroxyphenylpyruvate) (Cui et al., 2016). Combining the PCR and UHPLC-ESI-MS/MS analyses, all the 21 producer strains had the *tdcA* gene, but three of them (99, 112 and 337), despite the presence of the gene, were not able to produce tyramine. The lack of BA production might be correlated with a non-functionality of *tdc* cluster or because of non-optimal condition tested. Even if the presence of the gene in the genome is not a *sine qua non* condition for its expression and further study will be necessary to explore the tyramine amount directly in Pecorino Toscano cheese, a genetic screening could represent a suitable method to identify the potential BA producing strains, as suggested by Ladero et al. (2015). This analysis could be introduced as an additional criterion in the steered selection of strains suggested by PDO Protection Consortium and could be useful to control the starter cultures over time, considering the changes of the prevalence observed during the years. Indeed, the higher numbers of producer strains belonging to the NC, respect to the OC ones, suggest that this metabolic trait can be acquired over time, probably as a result of horizontal gene transfer or technological pressure.

Conclusions

Various species used as starter in fermented foods are able to accumulate tyramine causing adverse effects on human health like migraine and hypertensive crisis when BA amount increases. Pecorino Toscano cheese represents an ecological niche little investigated for this trait. Therefore, this study aimed to focus on the aptitude to convert tyrosine into tyramine in *in vitro* condition by *S. thermophilus* strains, belonging to the microbial collection of Pecorino Toscano Consortium, over time. The appearance of new producer strains in the last years, observed in this work, could be a consequence of the modification of autochthonous microflora, related to horizontal gene transfer and/or technological pressures. It is noteworthy that *S. thermophilus* starter play an essential role in cheese acidification, ripening and for flavour accumulation. Anyway, these first evidences on the high prevalence of BA producer strains highlighted the need to introduce an adjunct criterion in the selection of the starter, such as a genetic screening of the *tdcA* gene, to exclude undesired strains from the collection of autochthonous strains of Pecorino Toscano cheese.

Acknowledgements

The authors are grateful to “Consorzio Tutela del Pecorino Toscano DOP” for samples providing.

References

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, A., Wishart, D. S. (2016). Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Research, 44,* 147-153.

Bargossi, E., Gardini, F., Gatto, V., Montanari, C., Torriani, S., Tabanelli, G. (2015). The capability of tyramine production and correlation between phenotypic and genetic characteristics of *Enterococcus faecium* and *Enterococcus faecalis* strains. *Frontiers in Microbiology, 6,* 1371.

Benkerroum, N. (2016). Biogenic amines in dairy products: origin, incidence, and control means. *Comprehensive reviews of Food Science and Food Safety, 15,* 801–826.

Bergès, J., Trouillas, P., Houée-Levin, C. (2011). Oxidation of protein tyrosine or methionine residues: From the amino acid to the peptide. *Journal of Physics: Conference series, 261,* 012003.

Connil, N., Breton, Y.L., Dousset, X., Auffray, Y., Rincé, A., and Prévost, H. (2002). Identification of the *Enterococcus faecalis* tyrosine decarboxylase operon involved in tyramine production. *Applied and Environmental Microbiology, 68,* 3537–3544.

Coton, E. & Coton, M. (2009). Evidence of horizontal transfer as origin of strain to strain variation of the tyramine production trait in *Lactobacillus brevis*. *Food Microbiology 26,* 52–57.

Cui, Y., Xu, T., Qu, X., Hu, T., Jiang, X., Zhao, C. (2016). New Insights into Various Production Characteristics of *Streptococcus thermophilus* Strains. *International Journal of Molecular Sciences. 17,* pii: E1701.

EFSA (2011). Scientific opinion on risk-based control of biogenic amine formation in fermented foods. EFSA J. 9 2393–2486.

Fernández, M., Linares, D. M., Álvarez, M. A. (2004). Sequencing of the tyrosine decarboxylase cluster of *Lactococcus lactis* IPLA 655 and the development of a PCR method for detecting tyrosine decarboxylating lactic acid bacteria. *Journal of Food Protection, 67,* 2521–2529.

Gardini, F., Özogul, Y., Suzzi, G., Tabanelli, G., Özogul, F. (2016). Technological Factors Affecting Biogenic Amine Content in Foods: A Review. *Frontiers in Microbiology*, 7, 1218.

Gatto, V., Tabanelli, G., Montanari, C., Prodomi, V., Bargossi, E., Torriani, S., and Gardini, F. (2016). Tyrosine decarboxylase activity of *Enterococcus mundtii*: new insights into phenotypic and genetic aspects. *Microbial Biotechnology, 9,* 801-813.

Gobbetti, M., Neviani, E., Fox, P. (2018). The Cheeses of Italy: Science and Technology. (1st ed.) *Ed. Springer International*, (Chapter 6).

Guarcello, R., De Angelis, M., Settanni, L., Formiglio, S., Gaglio, R., Minervini, F., Moschetti, G., Gobbetti, M. (2016). Selection of amine-oxidizing dairy lactic acid bacteria and identification of the enzyme and gene involved in the decrease of biogenic amines. *Applied and Environmental Microbiology, 82,* 6870–6880.

Hafeez, Z., Cakir-Kiefer, C., Girardet, J. M., Lecomte, X., Paris, C., Galia, W., Dary, A., Miclo, L. (2015). New Insights into the Proteolytic System of *Streptococcus thermophilus*: Use of Isracidin To Characterize Cell-Associated Extracellular Peptidase Activities. *Journal of Agricultural and Food Chemistry, 63,* 7522-7531.

La Gioia, F., Rizzotti, L., Rossi, F., Gardini, F., Tabanelli, G., Torriani, S. (2011). Identification of a tyrosine decarboxylase gene (tdcA) in *Streptococcus thermophilus* 1TT45 and analysis of its expression and tyramine production in milk. *Applied and Environmental Microbiology, 77,* 1140–1144.

Ladero, V., Martín, M., Redruello, B., Mayo, B., Flórez, A., Fernández, M., Alvarez, M. (2015). Genetic and functional analysis of biogenic amine production capacity among starter and non-starter lactic acid bacteria isolated from artisanal cheeses. *European Food Research and Technology, 241,* 377–383.

Linares, D. M., Martín, M. C., Ladero, V., Álvarez, M. A., Fernández, M. (2011). Biogenic amines in dairy products. *Critical Reviews in Food Science and Nutrition, 51,* 691–703.

Lucas, P., Landete, J., Coton, M., Coton, E. & Lonvaud-Funel, A. (2003). The tyrosine decarboxylase operon of *Lactobacillus brevis* IOEB 9809: characterization and conservation in tyramine-producing bacteria. *FEMS Microbiology Letters, 229,* 65–71.

Molenaar, D., Bosscher, J. S., Ten Brink, B., Driessen, A. J. M., and Konings, W. N. (1993). Generation of a proton motive force by histidine decarboxylation and electrogenic histidine/histamine antiport in *Lactobacillus buchneri*. *Journal of Bacteriology, 175,* 2864–2870.

Neviani, E., Bizzarro, R., Righini, A., Mucchetti, G. (1998). Pecorino toscano DOP: tecniche di produzione e caratteristiche microbiologiche. (1st ed.) *Tecniche nuove,* (Chapter 6).

Rhee, J.E., Rhee, J.H., Ryu, P.Y., and Choi, S.H. (2002). Identification of the cadBA operon from *Vibrio vulnificus* and its influence on survival to acidstress. *FEMS Microbiol. Letters, 208,* 245–251.

Schirone, M., Tofalo, R., Visciano, P., Corsetti, A., Suzzi, G. (2012). Biogenic amines in Italian Pecorino cheese. *Frontiers in Microbiology, 3*, 171.

Van de Guchte, M., Serror, P., Chervaux, C., Smokvina, T., Ehrlich, S.D., and Maguin, E. (2002). Stress responses in lactic acid bacteria. *Antoine van Leeuwenhoek, 82,* 187–216.

Vido, K., Le Bars, D., Mistou, M.Y., Anglade, P., Gruss, A., and Gaudu, P. (2004). Proteome analyses of heme-dependent respiration in *Lactococcus lactis*: Involvement of the proteolytic system. *Journal of Bacteriology, 186,* 1648–1657.

Yılmaz, C., & Gökmen, V. (2017). Comparative evaluation of the formations of gamma-aminobutyric acid and other bioactive amines during unhopped wort fermentation. *Journal of Food Processing and Preservation, 42*, e13405.

Chapter 3

Draft genome sequence of Lactobacillus helveticus strain Lh 12 isolated from natural whey starter

Gaia Bertani, Daniela Bassi, Monica Gatti, Pier Sandro Cocconcelli, Erasmo Neviani

Published in Genome Announcement

Genetic traits of technological interest in the draft genome of Lactobacillus helveticus Lh 23 isolated from natural whey starter

Gaia Bertani, Daniela Bassi, Monica Gatti, Pier Sandro Cocconcelli, Erasmo Neviani

Manuscript in preparation

The dynamics and genetic traits of Lactobacillus helveticus in different ecological niches

Gaia Bertani, Daniela Bassi, Claudia Cortimiglia, Monica Gatti, Pier Sandro Cocconcelli and Erasmo Neviani

Manuscript in preparation

Draft genome sequence of Lactobacillus helveticus strain Lh 12 isolated from natural whey starter

Abstract

Lactobacillus helveticus is a lactic acid bacterium widely used in cheese-making and for the production of bioactive peptides from milk proteins. Here, we described the draft genome sequence and annotation of *L. helveticus* Lh 12 strain isolated from natural whey starter used in the production of Grana Padano cheese.

Genome Announcement

Lactobacillus helveticus is a homofermentative, thermophilic lactic acid bacterium widely used as a starter culture in the manufacture of Swiss-type and long-ripened Italian hard-cooked cheeses and as a flavor-enhancing adjunct culture for other types of cheese (Broadbent et al., 2011). This microorganism is furthermore able to produce peptides with a biological function, such as inhibitory activity on the angiotensin-converting enzyme (Giraffa, 2014).

L. helveticus Lh 12 was originally isolated from natural whey starter (NWS), the undefined starter culture used to produce Protected Designation of Origin Grana Padano cheese. Based on technique of back-slopping, NWS is obtained from the previous day's cheese-making whey incubated at a decreasing temperature, selecting, in this way, the growth of a large number of thermophilic lactic acid bacteria that, together with the raw milk microbiota, preserve from damage the final product and determine final cheese attributes (Gatti et al., 2014; Bassi et al., 2015).

Differences in the biotypes' composition of NWS, that distinguishes natural from selected cultures, is modulated by the incubation conditions of the drained whey, in strict synergy with small variations in the applied dairy technology (Gatti et al., 2014; Rossetti et al., 2008).

Nowadays the interest in these undefined complex starter cultures is growing for their low vulnerability to bacteriophage attack and because they represent a source for the isolation of new dairy strains with interesting functional characteristics (Smid et al., 2014).

In a previous study, different strains of *L. helveticus* were analyzed for genes involved in the proteolytic system and amino acid catabolism (Broadbent et al., 2011). Lh 12 (renamed in the study UPR2) was genetically different from the other strains, therefore sequencing of this genome may reveal alternative and interesting metabolic pathways.

The genome was sequenced with a 917.3-fold overall genome coverage using an Illumina HiSeq 1000 platform from Parco Tecnologico Padano (PTP) Genomic Platform, Lodi, Italy. The reads set were *de novo* assembled using SPAdes software (version 3.1.0). This strategy resulted in 334 contigs with a calculated genome size of 2.13 Mb and a GC content of 36.6%. A total of 2,327 genes were predicted by annotating the genome with both the NCBI Prokaryotic Genome Automatic Annotation Pipeline (PGAP) and the RAST Annotation Server (Aziz et al., 2008; Brettin et al., 2015). The draft genome contained 2,327 genes in total, including 1820 coding genes, 5 rRNA and 61 tRNAs and 502 hypothetical proteins. The functional annotation was performed using RAST version 2.0.

The proteolytic nature of this strain was reflected in the genome by the presence of genes involved in the amino acid metabolism. Lh 12 also harbors genes involved in sugar metabolism, transport and uptake, like genes participating in lactose, galactose, fructose, chitin and N-acetylglucosamine utilization. The genome analysis revealed also genes involved in stress response (heat shock response, osmotic and oxidative stress), and in DNA metabolism (coding for clustered regularly interspaced short palindromic repeat [CRISPR]-associated proteins and restriction-modification systems). Finally, two intact and three partial phage regions were identified by using the PHAge Search Tool (PHAST) (Zhou et al., 2011).

Accession number(s). This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under accession number LSVI00000000. The version described in this paper is the first version, LSVI01000000.

References

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., 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.

Bassi, D., Puglisi, E., Cocconcelli, P.S. (2015). Understanding the bacterial communities of hard cheese with blowing defect. *Food Microbiology, 52,* 106-118.

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Thomason, J.A., Stevens, R., Vonstein, V., Wattam, A.R., Xia, F. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports, 5,* 8365.

Broadbent, J.R., Cai, H., Larsen, R.L., Hughes, J.E., Welker, D.L., De Carvalho, V.G., Tompkins, T.A., Ardö, Y., Vogensen, F., De Lorentiis, A., Gatti, M., Neviani, E., Steele, J.L. (2011). Genetic diversity in proteolytic enzymes and amino acid metabolism among *Lactobacillus helveticus* strains. *Journal of Dairy Science, 94,* 4313-4328.

Gatti, M., Bottari, B., Lazzi, C., Neviani, E., Mucchetti, G. (2014). Invited Review: Microbial evolution in raw milk, long-ripened cheeses produced using undefined natural whey starters. *Journal of Dairy Science 97,* 573-591.

Giraffa, G. (2014). *Lactobacillus helveticus*: importance in food and health. *Frontiers in Microbiology, 5,* 338.

Rossetti, L., Fornasari, M.E., Gatti, M., Lazzi, C., Neviani, E., Giraffa, G. (2008). Grana Padano cheese whey starters: microbial composition and strain distribution. *International Journal of Food Microbiology, 127,* 168-171.

Smid, E.J., Erkus, O., Spus, M., Wolkers-Rooijackers, J.C., Alexeeva, S., Kleerebezem, M. (2014). Functional implications of the microbial community structure of undefined mesophilic starter cultures. *Microbial Cell Factories, 13,* (suppl. 1).

Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., Wishart, D.S. (2011). PHAST: A Fast Phage Search Tool. *Nucleic Acids Research, 39,* 347-352.

Genetic traits of technological interest in the draft genome of Lactobacillus helveticus Lh 23 isolated from natural whey starter

Abstract

Lactobacillus helveticus is a microorganism largely used as starter culture thanks to several genes, recovered in its genome, with potential industrial application. Here, we present the shotgun genome of *L. helveticus* strain Lh 23, isolated from natural whey starter used in the production of Grana Padano PDO cheese. Draft genome consisted in a circular chromosome of 2,100,230 bp with a GC content of 36.5%. Lh 23 harboured genes involved in the biosynthesis of folate and exopolysaccharides and twenty-four genes related to stress response to face several changes due to specific technological processes of the dairy environment. Several components of a strong proteolytic system were found, moreover an enzyme, L-serine ammonia lyase, involved in serine degradation was identified, allowing the strain to survive longer in a sugarless environment such as cheese during ripening. The sequenced genome was characterized also by the presence of four intact prophages elements integrated in the chromosome.

Keywords: *Lactobacillus helveticus*, Genome, Folate biosynthesis, EPS biosynthesis, Bacteriophages

Introduction

Lactic acid bacteria (LAB) are widespread microorganisms present in nature and characterized by a great diversity. These microorganisms play a key role in improving the nutritional value, the flavor and favoring preservation of different fermented products (Liu et al., 2015).

Among LAB, *Lactobacillus helveticus* is a homofermentative, thermophilic bacterium, widely distributed in different ecological niches, commonly found in the dairy environment, but also in other completely different contests like whisky fermentations, human reproductive or gastrointestinal tracts. Thanks to several genes in its chromosome that confer technological properties, *L. helveticus* is commonly found in natural starter used for the production of long-ripened Italian hard-cooked cheeses and Swiss-type cheeses (Giraffa et al., 2014). Particularly, undefined natural whey starter culture (NWS) used to produce Protected Designation of Origin (PDO) Grana Padano (GP), a long ripened Italian hard cheese, made from partially skimmed cow's milk, is composed mostly by *L. helveticus* (Neviani et al., 2013; Santarelli et al., 2013; Gatti et al., 2014). It is well known that the main technological role of *L. helveticus* in GP production is firstly the lactose depletion during the curd acidification and later the proteolytic activity over the long ripening. Proteolytic activity of this LAB species is well known (Savijoki et al., 2006; Moser et al., 2018) but, interestingly, *L. helveticus* Lh 23 has been previously considered in a comparative study where it was denominated UPR16 (Broadbent et al., 2011). Comparing by mean of microarray the proteolytic system of 38 different strains, Lh 23 was genetically different from the other strains isolated from dairy environment (Broadbent et al., 2011; Bertani et al., 2018), highlighting a great phenotypic and genotypic biodiversity observed in this species as a result of different selective pressures due to environmental and technological changes (Lombardi et al., 2002; Fortina et al., 1998; Gatti et al., 1999; Gatti et al., 2003; Gatti et al., 2004). In this paper we present the genome sequencing and study of *L. helveticus* Lh 23. The genetic information obtained from Lh 23

sequencing could be exploited to gain an insight into its technological traits and alternative metabolic pathways that could be useful to discover a natural ecological niche, where technological pressures play an important role in the manufacturing and furthermore in the adaptation of the strain to this particular environment.

Materials and methods

Bacterial growth and DNA extraction

L. helveticus Lh 23 was grown in MRS broth (Oxoid Ltd., Basingstoke, UK) at 42°C, under anaerobic conditions, for 24 hours. High-quality genomic DNA of *L. helveticus* Lh 23 strain was extracted with DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germania) according to manufacturer's instructions.

High-density sequencing and sequence assembly of the genome

Complete genomic sequencing was conducted using a second-generation sequencing platform, Illumina HiSeq 1000, from Parco Tecnologico Padano (PTP) Genomic Platform, Lodi, Italy, resulting in 2.38 Mb data (11,924,037 reads with 488bp average insert size and 993.0-fold coverage). The genome was directly assembled into 312 contings by SPAdes software v3.1.0.

Genome annotation and in silico analysis

Gene functional annotation was performed using BLASTP with KEGG, Swiss-Prot and RAST databases (Kanehisa et al., 2000; Aziz et al., 2008; Brettin et al., 2015). rRNA and tRNA genes were identified using RNAmmer and tRNAscan-SE (Lagesen et al., 2007; Lowe et al., 1997; Lowe et al., 2016). PHAge Search Tool (PHAST) was used to identify the presence of prophages in Lh 23 genome (Zhou et al., 2011). As regard the characterization of exopolysaccharide (EPS) biosynthesis proteins, Search Tool for the Retrieval (STRING) web tool was performed in order to explore protein–protein interactions (Von Mering et al., 2003).

Nucleotide sequence accession number

The complete genome sequence of *Lactobacillus helveticus* Lh 23 has been deposited in GenBank under accession number LSVJ00000000. The strain belongs to the collection of the Laboratory of Food Microbiology of the Department of Food and Drug of the University of Parma

Results and discussion

Genome features

The draft genome of *L. helveticus* Lh 23 was composed of 2,100,230 bp with a GC content of 36.5% (Figure 1).

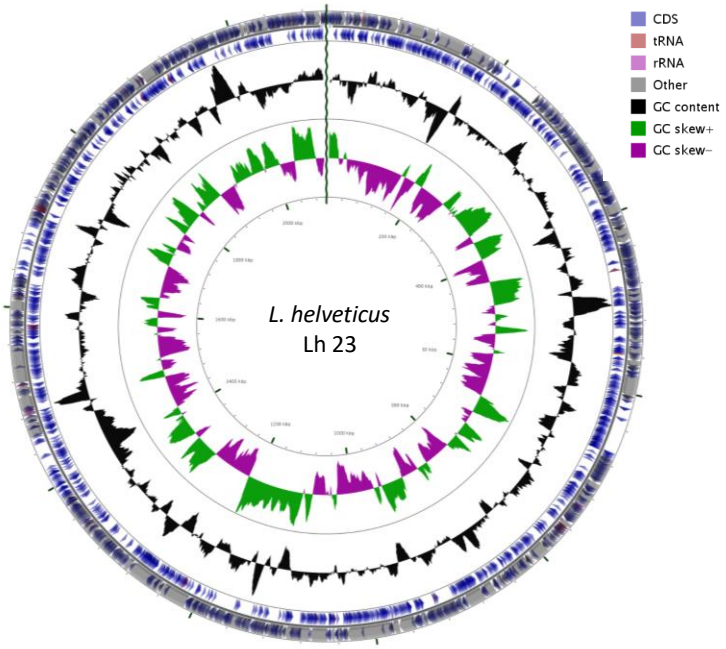


Figure 1 Circular chromosome graph of *Lactobacillus helveticus* Lh 23. The 1st and the 2nd circles represent the GC skew (G-C) / (G+C). The 3rd circle displays the G + C content of the genome. The 4th (forward strand) and 5th (reverse strand) indicate the sites of CDSs/rRNA/tRNA on the genome.

The chromosome of Lh 23 totally contained 2210 genes, which include 1751 protein coding genes, 420 pseudo genes, 5 rRNA and 63 tRNA for transfer of all 20 amino acids. Biological functions

were identified for 1407 (80.35%) genes, and 344 (19.65%) genes encoded hypothetical proteins (Table 1).

Table 1 Genome features of *L. helveticus* Lh 23.

Features	Chromosome Lh 23
Length (bp)	2,100,230bp
G + C content (%)	36.5%
Total genes	2210
Protein coding genes	1751
tRNA genes	63
rRNA genes	5
Pseudo genes	420
Hypothetical protein with unknown function	344

Biosynthesis and degradation

RAST annotation has shown 128 genes involved in the amino acid metabolism. Sixty-three genes are involved in lysine, threonine, methionine and cysteine metabolism. *L. helveticus* species was characterized by a high level of auxotrophy regarding amino acid biosynthesis (Slattery et al., 2010). Having a strong proteolytic system, *L. helveticus* can produce short peptides and liberate amino acids from the casein matrix, allowing its easy growth in milk (Griffiths et al., 2013). The proteolytic system of *L. helveticus* is characterized by the presence of a cell wall-associated serine proteinase, transport systems specific for di-, tri-, and oligopeptides, and a multitude of intracellular peptidases (Cremonesi et al., 2012). Four proteinases, PrtH (gLh23_seq109), PrtH2 (AY471_RS02465), PrtH3 (AY471_RS06440) and PrtH4 (AY471_RS01515) were found in the draft genome of Lh 23 (Table 2), in line with the results previously obtained in Broadbent et al. (2011). Regarding the transport systems for di-, tri-, and oligopeptides, Lh 23 chromosome contains five ABC oligopeptide transporters (AY471_RS01185, AY471_RS01190, AY471_RS01710,

AY471_RS01715, AY471_RS06020). Thirty genes encoding intracellular peptidases with different specificities were found in Lh 23 chromosome (Table 2).

Table 2 Proteinases and peptidases related proteins of *L. helveticus* Lh 23.

Peptidase/Proteinase	Peptidase/Proteinase related proteins	Locus tag
Proteinases	Metalloproteinase	AY471_RS00540
	PrtH2	AY471_RS02465
	PrtH3	AY471_RS06440
	PrtH4	AY471_RS01515
Peptidases	M13 peptidase	AY471_RS00400, AY471_RS02530, AY471_RS04455, AY471_RS05330
	Aminopeptidase	AY471_RS01175, AY471_RS02620
	Dipeptidase	AY471_RS01690, AY471_RS08365
	Peptidase T	AY471_RS02720, AY471_RS10520
	IV prepilin peptidase	AY471_RS03230
	S26 family signal peptidase	AY471_RS03470
	Peptidase M42	AY471_RS03610
	Signal peptidase II	AY471_RS03990
	Signal peptidase I	AY471_RS04110
	Dipeptidase A	AY471_RS04805
	Xaa-Pro dipeptidase	AY471_RS05970
	Oligoendopeptidase F	AY471_RS06325
	Xaa-Pro dipeptidyl-peptidase	AY471_RS06785
	X-Pro dipeptidase	AY471_RS06850
	D-alanyl-D-alanine carboxypeptidase	AY471_RS06960
	Pyroglutamyl-peptidase I	AY471_RS07180
	Methionyl aminopeptidase	AY471_RS08300
	HslU--HslV peptidase ATPase subunit	AY471_RS10085
	Peptidase M23	AY471_RS10545
	Dipeptidase PepV	AY471_RS11235
	Murein L,D-transpeptidase	AY471_RS06950
	Aminopeptidase E	AY471_RS01220
	Aminopeptidase N	AY471_RS02315
	Aminopeptidase C	AY471_RS07750

About amino acid degradation in the chromosome of *L. helveticus* Lh 23, L-serine ammonia lyase, an enzyme involved in serine degradation was found. L-serine ammonia lyase (SDH) or serine dehydratase, belongs to the β-family of pyridoxal phosphate-dependent enzymes, and catalyses the deamination of L-serine into pyruvate, with the release of ammonia. SDH could be a useful gene, because in LAB glycolysis terminates when glucose residues arising from lactose end. Serine is a very frequent amino acid in milk casein and it is released during cheese ripening. Thus, strains possessing this enzyme can survive longer in a sugarless environment such as ripened cheese, because when the lactose ends, the proteolytic system can release serine residues that are therefore converted to pyruvate.

Adaptation to stress

Twenty-four genes involved in the stress response system were found in Lh 23. During cheese production LAB microflora undergoes several changes such as heat shock, adverse pH, reduction of redox potential, water activity and nutrient content (Bottari et al., 2010). Five intracellular proteases Clp (AY471_RS02825, AY471_RS03465, AY471_RS04935, AY471_RS07230, AY471_RS09350) and two peptidases HslV (AY471_RS10085, AY471_RS10090) were found which could degrade aberrant and nonfunctional proteins. To deal with heat shock, several heat shock proteins were found in the sequenced genome; in addition, also one cold-shock protein, CspA (AY471_RS07375), was also identified. Regarding the oxidative stress response, five thioredoxin (AY471_RS01045, AY471_RS04685, AY471_RS06990, AY471_RS08055, AY471_RS10340), ferroxidase (gLH23_seq185), iron-binding ferritin-like antioxidant (gLH23_seq185), protein DNA-binding protein Dps (gLH23_seq185), redox-sensitive transcriptional regulator (gLH23_seq224) were also observed. F0F1-ATPase is the main regulator of intracellular pH, and seven proteins belonging to this complex were found in Lh 23 genome (AY471_RS07025, AY471_RS07030, AY471_RS07035, AY471_RS07040, AY471_RS07045, AY471_RS07050, AY471_RS07060).

Moreover, regarding pH regulation and Na^+ homeostasis, two sodium-proton (Na^+/H^+) antiporters and an alkaline shock protein (AY471_RS01440), which improve the acid tolerance of the strain, were identified in the Lh 23 chromosome.

Carbohydrate metabolism

RAST annotation has shown 187 genes involved in sugar metabolism, transport and uptake, such as genes participating in lactose, galactose, maltose, fructose, chitin and N-acetylglucosamine utilization. Many genes participating in the Embden-Meyerhof-Parnas (EMP) pathway were found in the genome of Lh 23. Similarly to other sequenced genomes available in databases, *L. helveticus* Lh 23, did not encode an intact citrate acid cycle (TCA), but several enzymes involved in this pathway were found, such as class II fumarate hydratase (AY471_RS09515), phosphoenolpyruvate carboxylase (AY471_RS11225), citrate lyase subunit alpha (AY471_RS08430), citrate lyase subunit beta (AY471_RS08435), citrate lyase ACP (AY471_RS08440), citrate (pro-3S)-lyase ligase (AY471_RS08445), citrate lyase holo (AY471_RS09255). Through glycolysis pathway, glucose is converted to pyruvate, which is subsequently transformed into lactate by the activity of lactate dehydrogenase enzyme. Lactic acid produced by *L. helveticus* is an equal mixture of L(+) and D(−)-isomers (Kylä-Nikkilä et al., 2000). In the Lh 23 chromosome, three L-lactate dehydrogenase genes (AY471_RS02600, AY471_RS02905, AY471_RS07800, AY471_RS09705) and one D-lactate dehydrogenase gene (AY471_RS07800) were identified.

Bacteriophages

The Lh 23 genome contained six prophage elements (four intact, two incomplete) (Figure 2A). One of the four intact prophage regions resembled to Entero_lato (5.6 Kb, region 1), the second prophage region looks like Paenib_Fern (40.2Kb, region 2), the third prophage identified is like Pseudo_Andromeda (35.8 Kb, region 3) and the fourth phage resembles to Lactob_phiAQ113 (43.2 Kb, region 4) (Figure 2B). The prophage Lactob_phiAQ113 was already sequenced and found in a

strain of *L. helveticus* isolated from NWS of Grana Padano cheese (Zago et al., 2013). Integrases and transposases are markers for mobile DNA elements. Four integrases (PP_00633, PP_00655, PP_00705, PP_00800) were identified in region 2, 3 and 4. All these regions contained also attL and attR sites, which identified the extent of the prophage.

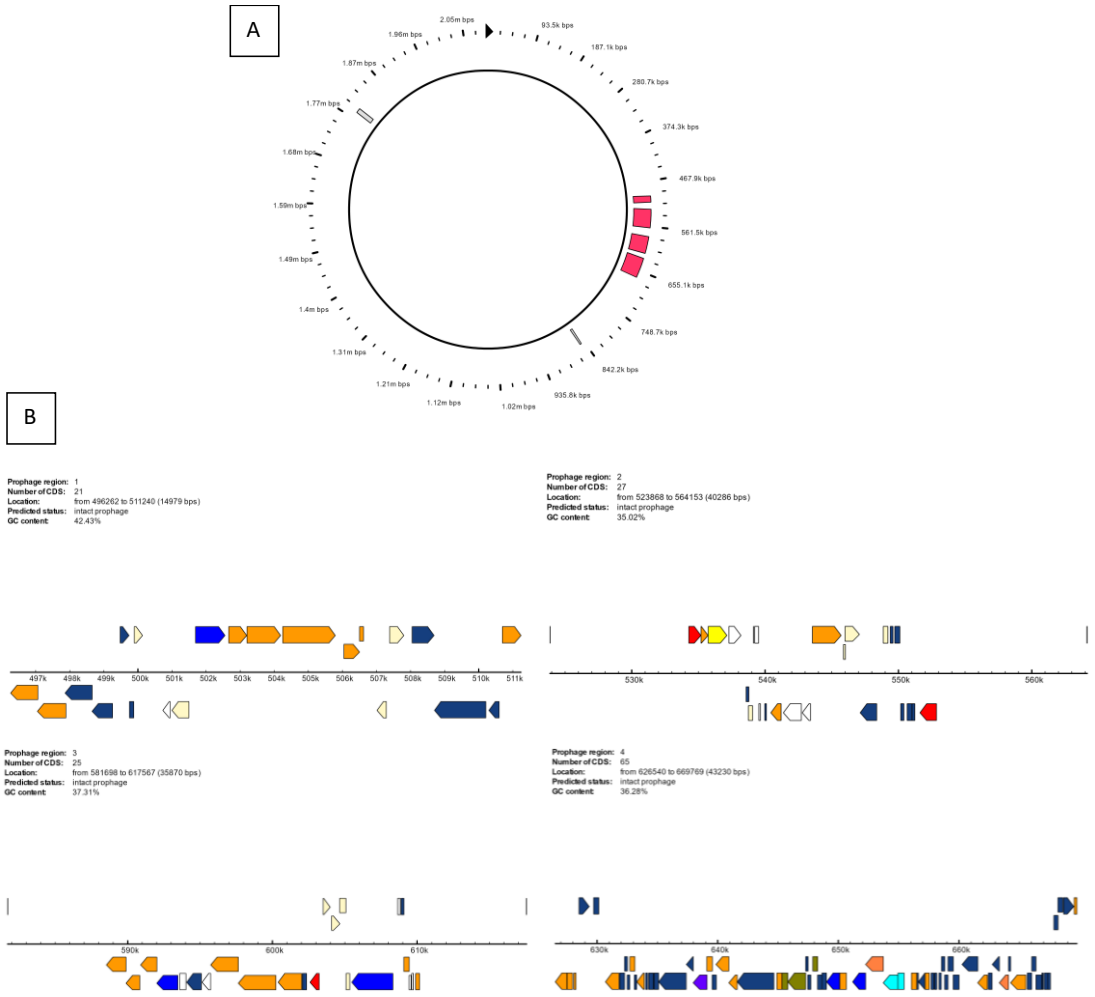


Figure 2 Prophage regions and predicted elements in the Lh 23 genome. Part A represents the prophage regions predicted in Lh 23 genome, that include 4 intact regions and two incomplete regions. Parts B represents the four intact prophages integrated in the chromosome.

A greatest advantage of NWS was the great resistance to phage attacks, while in different dairy fermentative processes, bacteriophage infections represent a serious problem (Moineau et al., 2005).

Indeed, NWS in Grana Padano production is expected to be highly tolerant to phage infections because strains grow in the presence of numerous phages, thus leading to the dominance of resistant or tolerant ones (Marcó et al., 2012).

Technological Properties

The draft genome of Lh 23 included all genes required for folate biosynthesis. The folate biosynthesis pathway has been deepen investigated in *Lactobacillus plantarum* WCFS1 (Wegcamp et al., 2009). KEGG analysis revealed a total of seven genes involved in the complete metabolic pathway for the biosynthesis of folate, GTP cyclohydrolase I FolE (AY471_RS11175), dihydroneopterin aldolase (AY471_RS11180), 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase (AY471_RS11175), dihydropteroate synthase (AY471_RS11165), bifunctional folylpolyglutamate synthase/dihydrofolate synthase (AY471_RS09800 , AY471_RS11170), dihydrofolate reductase (AY471_RS05115). Folates have a key role in the biosynthesis of purines and pyrimidines and regeneration of methionine. In addition, humans have an auxotrophic requirement for vitamin B12 and folate, so this technological property could be useful to create functional and probiotic products.

Moreover, two genes (AY471_RS06125, AY471_RS06130) involved in EPS biosynthesis were identified through RAST annotation. STRING web tool was carried out for exploring protein-protein interactions and to determine the relationship of these proteins with other proteins located in a network hub. The two encoded proteins (WP_061313842.1, WP_020828979.1) were involved in the functional association of 11 nodes (Figure 3). All proteins identified in this network association were mainly related to EPS biosynthesis pathway, such as phospho-glucosyltransferase (lhv_1803), glycosyltransferase (lhv_1797, lhv_1800, lhv_1801) galactosyl transferase (lhv_1802). The ability to synthetize polysaccharides, well known in *L. helveticus* (Torino et al., 2005), could be important

in the dairy industry, to improve the texture and the rheology of fermented products (De Vuyst et al., 1999; Duboc et al., 2001).

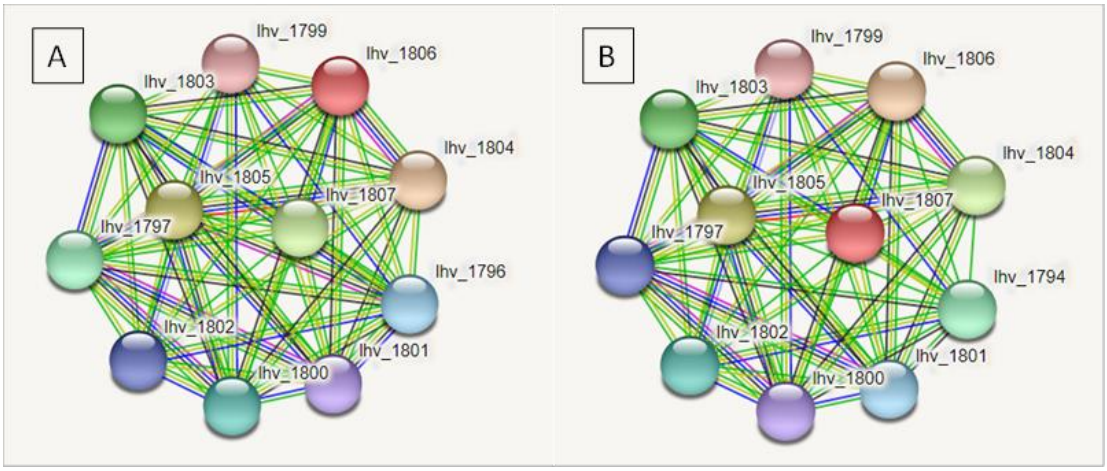


Figure 3 Protein–protein interactions analysis of exopolysaccharide biosynthesis proteins using STRING network. A) STRING network analysis between exopolysaccharide biosynthesis protein (lhv_1805) and other associated proteins. B) STRING network analysis of exopolysaccharide biosynthesis protein (lhv_1807).

In summary, the genome sequence and the analysis performed highlight interesting genetic traits of technological interest of *L. helveticus* Lh 23. Moreover, metabolic pathways for folate and exopolysaccharides biosynthesis was found. These findings could be helpful for the use of this strain as targeted application in functional products.

Acknowledgments

This research was partially supported by Consorzio per la Tutela del Formaggio Grana Padano.

References

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., 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.

Bertani, G., Bassi, D., Gatti, M., Cocconcelli, P.S., Neviani, E. (2018). Draft Genome Sequence of *Lactobacillus helveticus* Strain Lh 12 Isolated from Natural Whey Starter. *Genome Announcements, 6,* pii: e00139-18.

Bottari, B., Santarelli, M., Neviani, E., Gatti, M. (2010). Natural whey starter for Parmigiano Reggiano: culture-independent approach. *Journal of Applied Microbiology, 108,* 1676-1684.

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Thomason, J.A., Stevens, R., Vonstein, V., Wattam, A.R., Xia, F. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports, 5,* 8365.

Broadbent, J.R., Cai, H., Larsen, R.L., Hughes, J.E., Welker, D.L., De Carvalho, V.G., Tompkins, T.A., Ardö, Y., Vogensen, F., De Lorentiis, A., Gatti, M., Neviani, E., Steele, J.L. (2011). Genetic diversity in proteolytic enzymes and amino acid metabolism among *Lactobacillus helveticus* strains. *Journal of Dairy Science, 94,* 4313–4328.

Cremonesi, P., Chessa, S., Castiglioni, B. (2012). Genome sequence and analysis of *Lactobacillus helveticus*. *Frontiers in Microbiology, 3,* 435.

De Vuyst, L., Degeest, B. (1999). Heteropolysaccharides from lactic acid bacteria. *FEMS Microbiology Reviews, 23,* 153-77.

Duboc, P., Mollet, B. (2001). Applications of exopolysaccharides in the dairy industry. *International Dairy Journal, 11,* 759-768.

Fortina, M.G., Nicastro, G., Carminati, D., Neviani, E., Manachini, P.L. (1998). *Lactobacillus helveticus* heterogeneity in natural cheese starters: the diversity in phenotypic characteristics. *Journal of Applied Microbiology, 84,* 72-80.

Gatti, M., Bottari, B., Lazzi, C., Neviani, E., Mucchetti, G. (2014). Invited review: Microbial evolution in raw-milk, long-ripened cheeses produced using undefined natural whey starters. *Journal of Dairy Science, 97,* 573-591.

Gatti, M., Contarini, G., Neviani, E. (1999). Effectiveness of chemometric techniques in discrimination of *Lactobacillus helveticus* biotypes from natural dairy starter cultures on the basis of phenotypic characteristics. *Applied and Environmental Microbiology, 65,* 1450.

Gatti, M., Lazzi, C., Rossetti, L., Mucchetti, G., Neviani, E. (2003). Biodiversity in *Lactobacillus helveticus* strains in natural whey starter used for Parmigiano Reggiano Cheese. *Journal of Applied Microbiology, 95,* 463-480.

Gatti, M., Trivisano, C., Fabrizi, E., Neviani, E., Gardini, F. (2004). Biodiversity within *Lactobacillus helveticus* isolated from different natural whey starter cultures as revealed by classification trees. *Applied and Enviromental Microbiology, 70,* 182-190.

Giraffa, G. (2014). *Lactobacillus helveticus*: importance in food and health. *Frontiers in Microbiology, 5,* 338.

Griffiths, M.W., Tellez, A.M. (2013). *Lactobacillus helveticus*: the proteolytic system. *Frontiers in Microbiology, 4,* 30.

Kanehisa, M., Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research, 28,* 27-30.

Kylä-Nikkilä, K., Hujanen, M., Leisola, M., Palva, A. (2000). Metabolic Engineering of *Lactobacillus helveticus* CNRZ32 for Production of Pure L-(+)-Lactic Acid. *Journal of Applied and Environmental Microbiology, 66,* 3835-3841.

Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.H., Rognes, T., Ussery, D.W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research, 35,* 3100-3108.

Liu, J.J., Liu, X.P., Ren, J.W., Zhao, H.Y., Yuan, X.F., Wang, X.F., Wang, X.F., Salem, A.Z.M., Cui, Z.J. (2015). The effects of fermentation and adsorption using lactic acid bacteria culture broth on the feed quality of rice straw. *Journal of Integrative Agriculture, 14,* 503-513.

Lombardi, A., Dal Maistro, L., De Dea, P., Gatti, M., Giraffa, G., and Neviani, E. (2002). A poliphasic approach to highlight genotypic and phenotypic diversities of *Lactobacillus helveticus* strains isolated from dairy starter cultures and cheeses. *Journal of Dairy Research, 69,* 139-149.

Lowe, T.M., Chan, P.P (2016). tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research, 44,* 54-57.

Lowe, T.M., Eddy, S.R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research, 25,* 955-964.

Marcó, M.B., Moineau, S., Quiberoni, A. (2012). Bacteriophages and dairy fermentations. *Bacteriophage, 2,* 149-158.

Moineau, S., Lévesque. C. (2005). Control of bacteriophages in industrial fermentations, eds. *Bacteriophages: biology and applications*. 285-296.

Moser, A., Schafroth, K., Meile, L., Egger, L., Badertscher, R., Irmler, S. (2018). Population Dynamics of *Lactobacillus helveticus* in Swiss Gruyère-Type Cheese Manufactured with Natural Whey Cultures. *Frontiers in Microbiology, 9,* 637.

Neviani, E., Bottari, B., Lazzi, C., Gatti, M. (2013). New developments in the study of the microbiota of raw-milk, long-ripened cheeses by molecular methods: the case of Grana Padano and Parmigiano Reggiano. *Frontiers in Microbiology, 4,* 36.

Santarelli, M., Bottari, B., Lazzi, C., Neviani, E., and Gatti, M. (2013). Survey on community and dynamics of lactic acid bacteria in Grana Padano cheese. *Systematic and Applied Microbiology, 36,* 593–600.

Savijoki, K., Ingmer, H., Varmanen, P. (2006). Proteolytic systems of lactic acid bacteria. *Applied Microbiology Biotechnology, 71,* 394-406.

Slattery, L., O'Callaghan, J., Fitzgerald, G.F., Beresford, T., Ross, R.P. (2010). Invited review: *Lactobacillus helveticus*: a thermophilic dairy starter related to gut bacteria. *Journal of Dairy Science, 93,* 4435-4454.

Torino, M.I., Hébert, E.M., Mozzi, F., Font de Valdez, G. (2005). Growth and exopolysaccharide production by *Lactobacillus helveticus* ATCC 15807 in an adenine-supplemented chemically defined medium. *Journal of Applied Microbiology, 99,* 1123-1129.

Von Mering, C., Huynen, M., Jaeggi, D., Schmidt, S., Bork, P., Snel, B. (2003). STRING: a database of predicted functional associations between proteins. *Nucleic Acids Research, 31,* 258-261.

Wegkamp, A., de Vos, W.M., Smid. E.J. (2009). Folate overproduction in *Lactobacillus plantarum* WCFS1 causes methotrexate resistance. *FEMS Microbiology Letters, 297,* 261-265.

Zago, M., Scaltriti, E., Rossetti, L., Guffanti, A., Armiento, A., Fornasari, M.E., Grolli, S., Carminati, D., Brini, E., Pavan, P., Felsani, A., D'Urzo, A., Moles, A., Claude, J.B., Grandori, R., Ramoni, R., and Giraffa, G. (2013). Characterization of the Genome of the Dairy *Lactobacillus helveticus* Bacteriophage ΦAQ113. *Journal of Applied and Environmental Microbiology, 79,* 4712-4718.

Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., Wishart, D.S. (2011). PHAST: A Fast Phage Search Tool. *Nucleic Acids Research, 39,* 347–352.

The dynamics and genetic traits of Lactobacillus helveticus in different ecological niches

Abstract

Lactobacillus helveticus is a lactic acid bacterium, widely used in the production of fermented milk beverages and different types of hard cheeses. This microorganism can be also found in other diverse ecological niches fully different from dairy ecosystems, such as oral cavity, intestinal or vaginal tract. Considering that, genome sequencing is a power tool to investigate strain-specific genes, providing an optimal basis for functional genomics studies, in this study a bioinformatic approach was applied. Three strains of *L. helveticus* (Lh 12, Lh 23 and UC1156) isolated from natural whey starter used in the production of Parmigiano Reggiano and Grana Padano cheeses were sequenced. Here, we present the first comparative genomics study of strains isolated from this undefined, artisanal, 'cyclical' ecosystem, compared to other eleven strains of *L. helveticus* isolated from diverse sources, in order to gain an insight into genome evolution of *L. helveticus* species and to discover highly desirable traits with technological potential. This comparison has identified a great strain heterogeneity in amino acid and sugar metabolism, showing no correlation to the source of isolation. Genes involved in autolytic and lipolytic mechanism, were differently found in all strains analyzed, demonstrating a great variability especially due to a frameshift mutation. A great biodiversity was also discovered in cell-envelope proteinases distribution, and those differences could be very important for strain functionality in cheese and in production of bioactive peptides in fermented milks. Furthermore, 4 sequenced genomes were characterized also by the presence of intact prophages elements integrated in the chromosome, that could play a key role in the evolution, moving and rearranging the presence and the order of genes in the chromosome.

Keywords: *Lactobacillus helveticus*, Genome, Proteolytic system, Bacteriophages

Introduction

Lactic acid bacteria (LAB) encompass a heterogeneous group of Gram-positive microorganisms, which are able to produce lactic acid starting from the fermentation of carbohydrates (Douillard et al., 2014). LAB are commonly distributed in milk and dairy products, but also in vegetable and plants, cereals and meat (Mayo et al., 2008). Thanks to their long and safe history of application LAB have always been associated with food fermentation. Furthermore, the importance of LAB is highlighted by their Generally Regarded as Safe (GRAS) status (Mattu et al., 2013). *Lactobacillus helveticus* is widely used as a starter culture or adjunct culture in Swiss-type cheeses and long-ripened Italian cheeses manufacturing (Gobbetti et al., 2018). It has been demonstrated, in recent years, that this species has an important role in human health. This microorganism can be also found in diverse ecological niches such as dairy, oral, vaginal and gastrointestinal tract (Dellaglio and Felis, 2005). Different niches offer diverse stressful and selective conditions (Senan et al., 2014). In a dairy-based environment, lactobacilli strains are characterized to resist adverse conditions encountered in technological processes, like several changes such as heat shock, adverse pH, reduction of redox potential, water activity and nutrient content (Gobbetti et al., 2018). While in a gut environment, strains should be able to survive into the digestive tract and colonize the digestive mucosa against perturbations and competition (Senan et al., 2014).

It is well recognized that in food, bacteria underwent of environmental pressures and thus developed distinct genetic and metabolic features among different habitats. Given their short generation times and large population sizes, bacteria can evolve rapidly. Adaptation play key role in the evolution (Elena et al., 2003), and the adaptation responses could be similar or widely differ between species and strains. These different responses could be due to gene loss or decay, horizontal gene transfer, gene up regulation or mutation (Senan et al., 2014).

Complete genome sequencing is an excellent instrument to define and to describe many genes and regulatory features as well as numerous other elements, acquired lost or maintained in different environments. The decreasing costs and increasing speed of complete genomes sequencing have led to an enormous amount of data (Inglin et al., 2018). Recently, the genome sequences of fifty *L. helveticus* strains isolated from different environments, such as Swiss cheese, industrial dairy starter cultures, fermented milk, vaginal and gastrointestinal tract, have been deposited in GenBank. Seen the widespread use of *L. helveticus* in cheese technology there is a great interest in analysing the genome of this species. Indeed, the first three genomes of *L. helveticus* species (Lh 12, Lh 23 and UC1156), isolated from natural whey starter culture (NWS) used in manufacturing of DOP Italian cheeses (Grana Padano and Parmigiano Reggiano) were sequenced (Bertani et al., 2018). NWS, for Grana Padano and Parmigiano Reggiano cheeses production, is obtained from the previous day's cheese-making whey incubated at a decreasing temperature, selecting, in this way, the growth of a large number of thermophilic lactic acid bacteria that, together with the raw milk microbiota, preserve from damage the final product and determine final cheese attributes (Gatti et al., 2014, Bassi et al., 2015).

In order to understand the genome evolution of *L. helveticus* in different environments, and first of all to identify highly desirable traits with technological potential for the production of cheese and other milk derivatives, the three strains recently sequenced were compared to other eleven strains isolated from fully different niches (swiss cheese, fermented milk, vaginal and gastrointestinal tract).

Materials and methods

Bacterial strains and genome sequencing

Genomic DNA of *L. helveticus* Lh 23 and Lh 12 strains, isolated from Grana Padano natural whey starter (NWS) belonging to the Laboratory of Food Microbiology of the Department of Food and

Drugs of the University of Parma, and UC1156 strain, isolated from Parmigiano Reggiano NWS, belonging to the University Cattolica del Sacro Cuore of Piacenza were extracted by DNeasy Blood & Tissue Kit (Qiagen). Briefly, all strains were grown in MRS broth (Merck, Darmstadt, Germany) at 37°C, under anaerobic conditions, for 24-48 hours. The active cultures were pelleted by centrifuging 3 ml of culture at 11000 rpm. Supernatants were discarded, and pellets were resuspended in 180 µl enzymatic lysis buffer, following all the steps of the manufacturer's protocol. The DNA was visualized on agarose gel and quantified using Quant-iT™ HS ds-DNA assay kit (Invitrogen, Paisley, UK) in combination with the QuBit™ fluorometer. The genomes of the 3 strains were sequenced using Illumina HiSeq 1000 platform with a 2x100bp paired-end (PE) protocol. *De novo* assembly of all reads was performed by SPAdes (Bankevich et all., 2012).

The following quality filtering was applied: reads were removed when they are duplicated or if contained more than 50 bases with a wrong call probability higher than 20%; reads shorter than 20 bases were discarded. The assembly generated 334 contigs for Lh 12, 312 contigs for Lh 23 and 392 contigs for UC1156. These draft-genome shotgun projects were submitted in NCBI under the accession numbers LSVJ00000000, LSVI01000000 and NRQX00000000.

Genome annotation and *in silico* analysis

Gene functional annotation was performed on fourteen sequenced strains, isolated from diverse sources (Table 1), using BLASTP with KEGG, Swiss-Prot and RAST databases (Kanehisa et al., 2000; Aziz et al., 2008; Brettin et al., 2015). PHAge Search Tool (PHAST) was used to identify the presence of prophages in all the fourteen strains analyzed in this study (Zhou et al., 2011). All the general features of *L. helveticus* were evaluated using RAST. A screening of genes involved in amino acid biosynthesis and degradation, proteolytic system, lipolytic and autolytic system, sugar metabolism was performed.

Table 1 Genome sequenced of *L. helveticus* strains used in this study and their respective origins.

Strain	Accession	Source
Lh 12	LSVI00000000	Isolated from natural whey starter used in Grana Padano manufacturing (Parma, Italy)
Lh 23	LSVJ00000000	Isolated from natural whey starter used in Grana Padano manufacturing (Parma, Italy)
UC1156	NRQX00000000	Isolated from natural whey starter used in Parmigiano Reggiano manufacturing (Piacenza, Italy)
CAUH18	NZ_CP012381	Isolated from Koumiss (Xinjiang, China)
CNRZ 32	NC_021744	Used as industrial cheese starter and cheese flavour adjunct
D76	NZ_CP016827	Isolated from intestine from healthy child (Leningrad, Russia)
DSM 20075	ACLM00000000	Isolated from Emmental cheese
DPC4571	NC_010080	Cheese starter and cheese flavour adjunct isolated from Swiss cheese
H10	NC_017467	Isolated from traditional fermented milk (Tibet, China)
H9	NZ_CP002427	Isolated from kurut (Tibet, China)
KLDS1.8701	NZ_CP009907	Isolated from sour milk (Sinkiang, China)
MB2-1	NZ_CP011386	Isolated from fermented milk (Southern Xinjiang, China)
MTCC5463	AEYL00000000	Vaginal tract
R0052	NC_018528	Probiotic strain isolated from sweet acidophilus milk (France)

Results and discussion

Genome sequencing, assembly and annotation of UC1156, Lh 12 and Lh 23 and genome features of *L. helveticus* strains

The *L. helveticus* strains UC1156, Lh 12 and Lh 23 have been sequenced through Illumina HiSeq 1000 platform with a 2x100bp paired-end (PE) protocol. Three programs have been tested on *Lactococcus lactis* sp. 36a dataset: Velvet[3], SOAPdenovo2[4] and SPAdes[5], in order to identify the best software for assembling *de novo*. Among these software tried, the results showed that SPAdes[5] gave the best results. Compared to SOAPdenovo2[4] and Velvet[3] software, SPAdes[5] showed the less redundant assembly, with the minimum number of sequences covering the same genome length of the other software, 1.5 times less and ~3 times less than the tested software respectively. Regarding the fragmentation of the assembly the longest scaffold is produced by SOAPdenovo2[4], but this software included the highest number of erroneous gaps (10 out of 15), followed by Velvet[3] in order to identify juxtaposition of contiguous regions. SPAdes[5] on the contrary was able to obtain the less fragmented assembly with the highest N50, N90 and average sequence length for both the assembly

and the contiguous regions. Another important aspect to evaluate was the continuity of the assembled sequences. SPAdes[5], compared to other software, was able to produce a gap-free assembly, suggesting this as the best tool to use for the assembly of these genomes. The genome of UC1156, Lh 12 and Lh 23 were annotated at NCBI, under the accession number NRQX00000000, LSVJ00000000 and LSVI01000000 respectively. An overview of the three genomes features was shown in Table 2 and Figure 1.

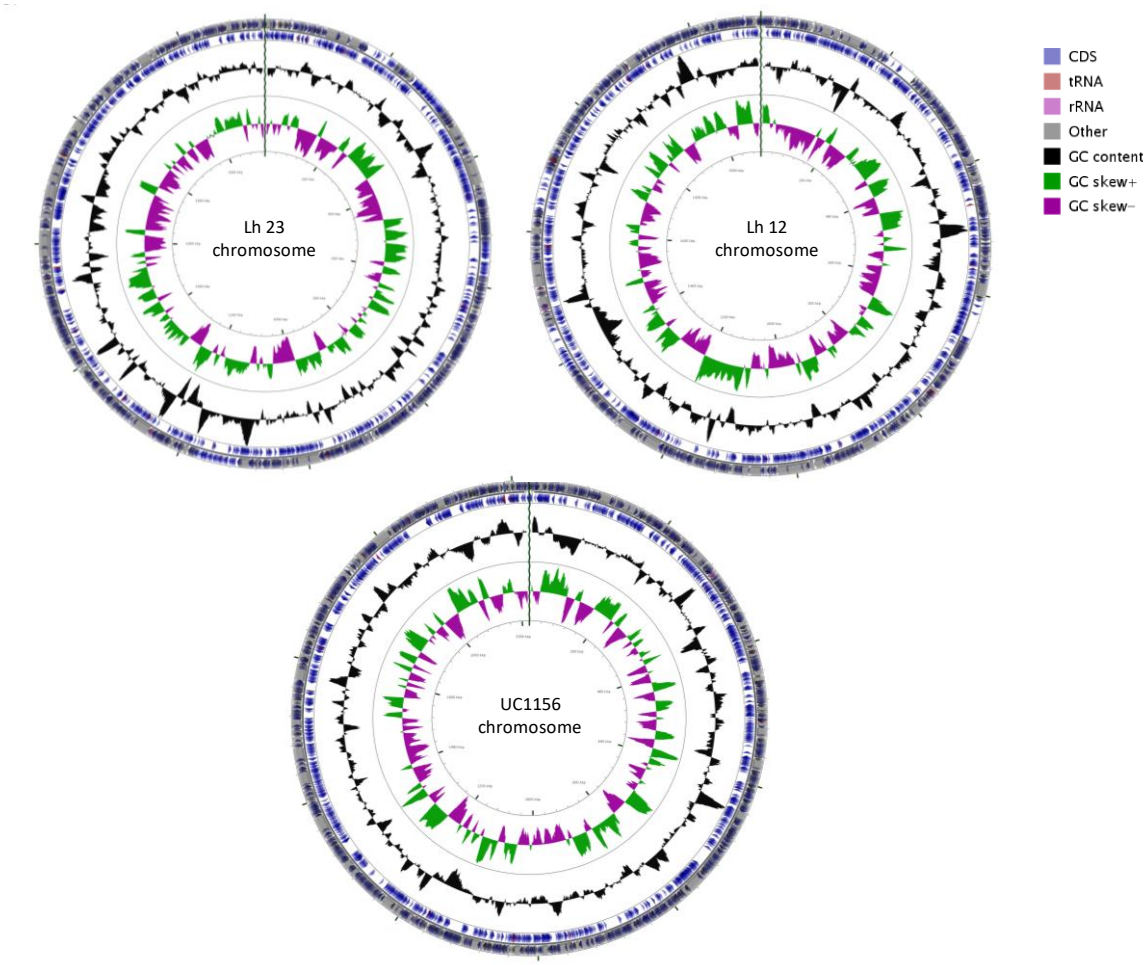


Figure 1 Circular genome map for Lh 23, Lh 12 and UC1156, generated using CGview. For every CircularMap, the following features are shown: the 1st and the 2nd circles represents the GC skew (G-C) / (G+C). The 3rd circle displays the G + C content of the genome. The 4th (forward strand) and 5th (reverse strand) indicate the sites of CDSs/rRNA/tRNA on the genome.

Table 2 Genome features of *L. helveticus* Lh 12, *L. helveticus* Lh 23 and *L. helveticus* UC1156.

Features	Lh 12	Lh 23	UC1156
Length (bp)	2,133,405 bp	2,100,230 bp	2,023,650 bp
G+C content (%)	36.6%	36.5%	36.7%
Total genes	2242	2270	2454
Protein coding genes	1783	1780	1872
tRNA genes	61	63	60
rRNA genes	5	5	6
Pseudo genes	439	420	513
Hypothetical protein with unknown function	408	478	333
GenBank accession	LSVJ00000000	LSVI01000000	NRQX00000000

The *L. helveticus* strains selected in this study represent a wide insight of genetic and ecological diversity in this species. Among the fourteen strains, five were draft genomes (Lh 12, Lh 23, UC1156, MTCC5463 and DSM20075) and had a coverage ranging from 13.61X and 992X (mean 503X) and total contig number ranged from 49 and 553. The total number of predicted CDS features in the fourteen genomes analyzed extended from 1,871 to 2,226 Mb, with a GC content of 36.5-37.0% (data not shown).

Genes relevant in a cheese ecosystem

The versatile nature of this species, to adapt easily in different niches, could be underlined by major biochemical processes, such as protein and amino acid degradation. About the proteolytic system, several studies have demonstrated that LAB, without any functional cell-envelope proteinases (CEPs), were not able to survive in milk environment (Mayo et al., 2010, Griffiths et al., 2013), indeed most of the LAB possess only one proteinase. Through a strong proteolytic system, it has been demonstrated that *L. helveticus,* in a cheese environment is able to release intracellular and bioactive peptides, which remain active for several weeks in a curd (Valence et al., 2000; Deutsch et al., 2003). Among the fourteen *L. helveticus* analysed, only two strains (CNRZ 32 and Lh 23),

encoded for 4 intact CEPs named PrtH1, PrtH2, PrtH3 and PrtH4, in accordance with previous study of Broadbent et al. (2011).

On the other hand, the remaining 12 strains possessed at least only one CEP-encoding gene in the chromosome suggesting a great intraspecific diversity of genes encoding CEPs. Indeed, previously studies have revealed a rich diversity, regarding CEPs, among several strains of *L. helveticus*, and this biodiversity observed could have wide implications for the application of these strains in cheese manufacturing and bioactive foods (Griffiths et al., 2013, Broadbent et al., 2011).

Autolytic and lipolytic ability associated to the proteolytic system are important traits involved in improving flavour and aroma in the final product, consequently, the genes involved in these processes, were studied in all the strains previously mentioned.

About the autolytic system, in Jebava et al. (2011) Vinogradeov et al. (2013), it has been reported that in *L. helveticus* species differences in peptidoglycan hydrolases caused variations in autolytic ability, and through PCR analysis have demonstrated that nine peptidoglycan hydrolases were part of core-genome of *L. helveticus*. Our results, in accordance with Schmid et al., (2018), identified, conversely to studies previously mentioned, only five (three N-acetylmuramidase, one amidase and one lysozyme) of these nine hydrolases, presented in all genomes analysed (Table 3).

During cheese ripening, lipolysis was the main biochemical event that occurs which contributes significantly to flavour development. Lipolytic enzymes as esterase and lipase, cleaving the ester linkage in a triglyceride, releasing free fatty acid (FFA), responsible to and di- and mono-acyglycerols (Holland et al., 2005; Hickley et al., 2007).

Microbial lipolytic enzymes, could improve flavour notes, or accelerate the maturation of various fermented products. If compared to other species like *Pseudomonas*, *Bacillus*, and *Achromobacter*, LAB had a weak lipolytic activity.

Table 3 Overview of peptidoglycan hydrolases (PGH) genes in fourteen *L. helveticus* genomes.

(PS): Pseudogene with premature stop codon (FS): Pseudogene with frameshift

Strains	Amidase 363aa	Lysin 406aa	Lysin 373aa	Lysozyme 234aa	M23 family peptidase 143aa	M23 family peptidase 218aa	N-acetylmuramidase 183aa	N-acetylmuramidase 215aa	N-acetylmuramidase 407aa
Lh 12	✓	✓	✓	✓	FS	✓	✓	✓	✓
Lh 23	✓	✓	✓	✓	✓	✓	✓	✓	✓
UC1156	✓	✓	FS	✓	✓	✓	✓	✓	✓
CAUH18	✓	✓	✓	✓	PS	✓	✓	✓	✓
CNRZ 32	✓	FS	FS	✓	✓	✓	✓	✓	✓
D76	✓	✓	FS	✓	✓	✓	✓	✓	✓
DPC4571	✓	✓	FS	✓	✓	✓	✓	✓	✓
DSM 20075	✓	✓	X	✓	FS	✓	✓	✓	✓
H10	✓	✓	✓	✓	PS	✓	✓	✓	✓
H9	✓	✓	FS	✓	FS	FS	✓	✓	✓
KLDS 1.8701	✓	✓	FS	✓	✓	✓	✓	✓	✓
MB2-1	✓	✓	FS	✓	PS	✓	✓	✓	✓
MTCC 5463	✓	✓	X	✓	PS	✓	✓	✓	✓
R0052	✓	✓	✓	✓	✓	✓	✓	✓	✓

Based on RAST analysis, six intact and highly conserved esterase were identified (> 98% average pairwise identity aa) in the functional core genome. The remaining esterase found in all the strains analysed were inactivated in one or several strains (Table 4).

Table 4 Overview of genes involved in lipolytic system in fourteen *L. helveticus* genomes.

Strains	Acyl-ACP thioesterase 244aa	Arylesterase 186aa	DHH family phosphoesterase 673aa	Esterase 190aa	Esterase 337aa	Glycerophosphodiester phosphodiesterase 365aa	Glycerophosphodiester phosphodiesterase 456aa	Glycerophosphodiester phosphodiesterase 596aa	Glycerophosphodiester phosphodiesterase 231aa
Lh 12	✓	✓	✓	✓	✓	FS	✓	FS	✓
Lh 23	✓	✓	✓	✓	✓	✓	✓	FS	✓
UC1156	✓	✓	✓	✓	✓	✓	✓	FS	✓
CAUH18	✓	✓	✓	✓	✓	✓	✓	✓	✓
CNRZ 32	✓	✓	✓	✓	✓	FS	✓	FS	✓
D76	✓	✓	✓	✓	✓	FS	✓	FS	✓
DPC4571	✓	✓	✓	✓	✓	FS	✓	FS	✓
DSM 20075	✓	✓	✓	✓	✓	FS	✓	FS	✓
H10	✓	✓	✓	✓	✓	✓	FS	FS	✓
H9	✓	✓	✓	✓	✓	FS	✓	FS	✓
KLDS 1.8701	✓	✓	✓	✓	✓	X	✓	FS	✓
MB2-1	✓	✓	✓	✓	✓	FS	FS	FS	✓
MTCC 5463	✓	✓	✓	✓	✓	FS	FS	FS	✓
R0052	✓	✓	✓	✓	✓	✓	✓	✓	✓

Like the majority of LAB, *L. helveticus* species is auxotrophic for many amino acids (Kandler and Weiss, 1986). Previous analysis on *L. helveticus* CNRZ 32, demonstrated that this strain was auxotrophic for 14 amino acids (Christensen and Steele, 2003; Christiansen et al., 2008). RAST annotation showed the ability of all fourteen strains to synthesize *de novo* four amino acids (Gln, Ser, Cys and Gly), and the ability to interconvert Asn and Asp (hence, only one of these amino acids was required), suggesting a high dependency on external supplies for the other 15 amino acids. About serine biosynthesis, the absence of *serC*, which encodes phosphoserine transaminase, was observed in 8 strains analysed. In CNRZ 32 chromosome this enzyme is adjacent of *serA* (3-phosphoglycerate dehydrogenase) and *pgm* (phosphoglycerate mutase), and it has been demonstrated that all these three genes were upregulated during microbial growth in milk environment (Broadbent et al., 2011). Interestingly, *serC* gene was present also in strain D76, isolated from human niche. These data contrasted with Broadbent et al., (2011) where the absence of this gene among all the whisky isolates, had suggested that this gene play a key role in the specialization of certain *L. helveticus* strains associated to a milk environment.

A second observation involved the absence of SDH, which encodes L-serine ammonia lyase, in 5 strains. L-serine ammonia lyase or serine dehydratase, belongs to the β-family of pyridoxal phosphate-dependent enzymes, and catalyses the deamination of L-serine into pyruvate, with the release of ammonia. SDH could be a useful gene, in milk environment. When glucose residues arising from lactose end, LAB glycolysis terminates. Milk casein are rich in serine residues, therefore the strains possessing this enzyme can survive longer in a sugarless environment such as ripened cheese, because when the lactose ends, the proteolytic system can release serine residues that are therefore converted to pyruvate.

Carbohydrate transport and metabolism

Comparative genome analysis of *L. helveticus* strains revealed a great variability regarding genes encoding various carbohydrates, involved in sugar metabolism, transport and uptake. The ability to exploit simple or complex carbohydrates was investigated in the fourteen strains to understand the diversity in sugar metabolism in relation to different sources of isolation. All the *L. helveticus* strains analysed had the capability to metabolize lactose and galactose, while regarding the other pathways involved in sugar metabolism, several differences were observed in the genomes analysed.

RAST analysis showed that 9 strains have the ability to metabolize maltose. As reported by Breton et al. (2005), in LAB the uptake and metabolism of maltose occurred through four different mechanisms, in the strains analysed were found two different metabolic pathways involved in the uptake and metabolism of this disaccharide. Nine strains (Lh 23, CNRZ 32, DPC4571, CAUH18, H10, H9, R0052, DSM20075 and MTCC5463) of the fourteen *L. helveticus* studied, had the capability to uptake the maltose into the cell through a permease, and then the sugar was cleaved by a maltose phosphorylase producing glucose and glucose-1-phosphate. Another possible system involved in maltose uptake into the cell is a phosphotransferase (PTS) system, the maltose-6-phosphate glucosidase is an enzyme linked to this system, converting maltose to glucose and glucose-6-phosphate. The proteins involved in PTS system were found in 8 strains (CNRZ 32, DPC4571, Lh 12, Lh 23, MB2-1, KLDS 1.8701, D76 and CAUH18) while the maltose-6-phosphate glucosidase was identified in 6 strains (CAUH18, H10, Lh 12, UC1156, R0052 and DSM20075). The majority of the fourteen *L. helveticus* possessed a complete pathway for the uptake and the metabolism of sucrose, with the exception of Lh 12 and DPC4571. In lactobacilli three pathways for sucrose metabolism were found (Gänzle and Follador, 2012). In ten strains analysed, sucrose

was transported into the cell through a PTS system, followed by the action of a sucrose phosphorylase, an enzyme that converts this disaccharide into glucose-6-phosphate and fructose.

As reported by Crowe et al. (2001), lactobacilli were also able, at the strain or species level, to metabolize several disaccharides like cellobiose, trehalose, turanose, maltulose, leucrose and palatinose. Genes involved in trehalose biosynthesis were largely distributed in nature, and it has been demonstrated that this disaccharide has been produced in response to an osmotic stress or to survive extreme conditions like too heat or cold, dehydration or entry into stationary phase (Ruhal et al. 2013). Seven *L. helveticus* strains analysed in this study had a complete pathway for the uptake and metabolism of this compound; trehalose entered into the cell through a PTS system, then two enzymes trehalose-6-phosphate hydrolase and trehalose phosphorylase, are responsible for the conversion of this compound into glucose and glucose 1-phosphate or glucose and glucose 6-phosphate respectively.

Bacteriophages of *L. helveticus*

Bacteriophages were widely distributed in nature and could be a serious issue in fermented products, in particular in cheese manufacturing, where the use of starter cultures in an intensive production system produced ideal conditions for the multiplication and survival of phage. NWS biological ecosystem have shown a great resistance to lytic system, because strains grown in the presence of phages, which leading to dominance of resistant or tolerant ones (Marcó et al., 2012).

Phages and phage remnants were found in several species belong to LAB and play a key role at species and strains level variability. Through PHAST analysis, the presence of bacteriophages in the fourteen strains analysed was investigated, identifying the presence of intact phages in four strains analysed (Lh 23, Lh 12, CNRZ 32 and MB2-1) (Figure 2). Interestingly, not only the strains isolated from dairy environment have interacted with bacteriophages, but also one strain isolated from fermented milk. These results were confirmed also by RAST analysis, showing in these strains

phage-related proteins, like proteins involved in phage replication, in packaging machinery and in the composition of capsid.

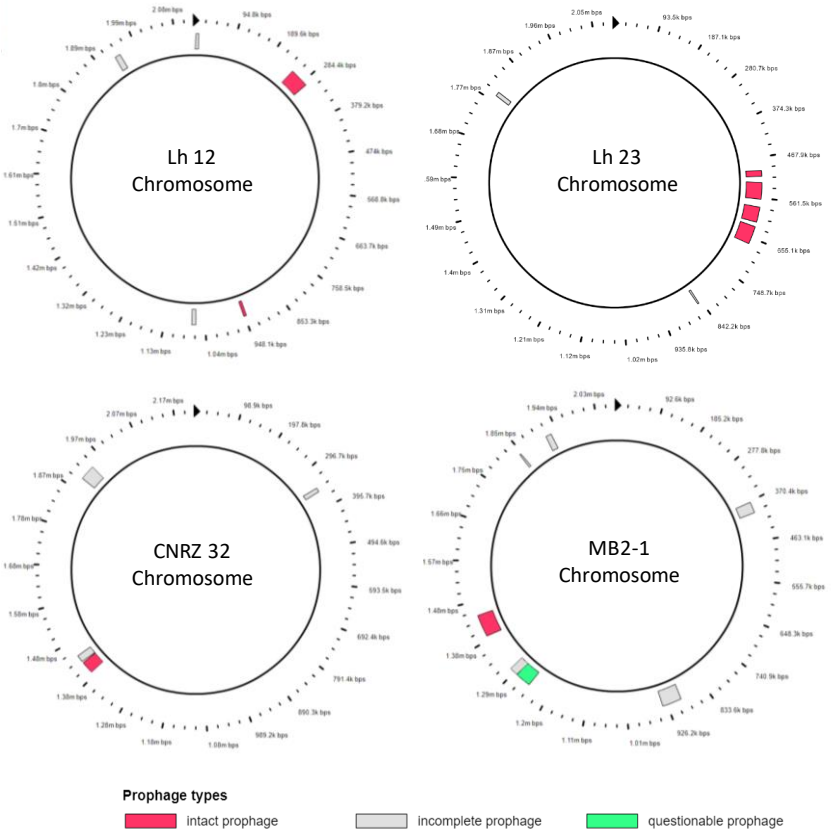


Figure 2 Prophage regions predicted in Lh 12, Lh 23, CNRZ 32 and MB2-1 genomes, through PHAST analysis.

Comparing through a multiple alignment the bacteriophages' sequences found, it has been observed that one phage integrated in the chromosome of Lh 23 was highly conserved (>99% average pairwise identity aa) in the chromosome of CNRZ 32. These results highlighted that, probably, these strains, both isolated from NWS, have interacted with the same bacteriophage, and have integrated this sequence in their chromosome.

Conclusions

Usually, in bacteria the concept of species is traditionally related to features encoded by the core genome, but which often are not exhaustive to adequately describe the biodiversity found at strain level (Broadbent et al., 2012). Thus, the comparison among fourteen strains, isolated from diverse ecological niches, could gain an accurate understanding on genome evolution of *L. helveticus* identifying also highly desirable traits with technological potential. To our knowledge, this is the first study that investigate the natural whey starter cultures used in DOP Italian cheeses manufacturing, comparing them against other ecosystems through a bioinformatic approach. Results from this study, suggest that probably the genetic traits investigated, have not been sufficient to capture the adaptation of *L. helveticus* species to different ecosystems. Nevertheless, this comparison has identified a great strain heterogeneity in amino acid and sugar metabolism, showing no correlation to the source of isolation. Regarding autolytic and lipolytic mechanism, the data obtained showed that the differences in gene content, are more likely due to a frameshift mutation, that affect the activity of the enzymes involved in these reactions. Marked genetic differences were discovered also in the distribution of CEPs, and those differences are very important traits of strain functionality in cheese and in the production of bioactive peptides in fermented milks.

References

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., 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.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology, 19,* 455-477.

Bassi, D., Puglisi, E., Cocconcelli, P.S., (2015). Understanding the bacterial communities of hard cheese with blowing defect. *Food Microbiology, 52,* 106-118.

Bertani, G., Bassi, D., Gatti, M., Cocconcelli, P.S., Neviani, E. (2018). Draft Genome Sequence of *Lactobacillus helveticus* Strain Lh 12 Isolated from Natural Whey Starter. *Genome Announcements, 6,* e00139-18.

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Thomason, J.A., Stevens, R., Vonstein, V., Wattam, A.R., Xia, F. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports, 5,* 8365.

Broadbent, J.R., Cai, H., Larsen, R.L., Hughes, J.E., Welker, D.L., De Carvalho, V.G., Tompkins, T.A., Ardö, Y., Vogensen, F., De Lorentiis, A., Gatti, M., Neviani, E., Steele, J.L. (2011). Genetic diversity in proteolytic enzymes and amino acid metabolism among *Lactobacillus helveticus* strains. *Journal of Dairy Science, 94,* 4313–4328.

Christensen, J.E., Steele, J.L. (2003). Impaired growth rates in milk of *Lactobacillus helveticus* peptidase mutants can be overcome by use of amino acid supplements. *Journal of Bacteriology, 185,* 3297-3306.

Christiansen, J.K., Hughes, J.E., Welker, D.L., Rodríguez, B.T., Steele, J.L., and Broadbent. J.R. (2008). Phenotypic and genotypic analysis of amino acid auxotrophy in *Lactobacillus helveticus* CNRZ 32. *Applied and Environmental Microbiology, 74,* 416–423.

Crowe, J.H., Crowe, L.M., Oliver, A.E., Tsvetkova, N., Wolkers, W., Tablin, F. (2001). The trehalose myth revisited: introduction to a symposium on stabilization of cells in the dry state. *Cryobiology, 43,* 89-105.

Dellaglio, F., and G. E. Felis. (2005). Taxonomy of Lactobacilli and Bifidobacteria., p. 48-73.

Deutsch, S.M., Neveu, A., Guezenec, S., Ritzenthaler, P., Lortal, S. (2003). Early lysis of *Lactobacillus helveticus* CNRZ 303 in Swiss cheese is not prophage-related. *International Journal of Food Microbiology, 81,* 147–157.

Douillard, F. P., & de Vos, W. M. (2014). Functional genomics of lactic acid bacteria: from food to health. *Microbial cell factories, 13,* (Suppl 1), S8.

Elena, S.F., Lenski, R.E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics, 4,* 457-469.

Gänzle, M.G., & Follador, R. (2012). Metabolism of oligosaccharides and starch in lactobacilli: a review. *Frontiers in Microbiology, 3,* 340.

Gatti, M., Bottari, B., Lazzi, C., Neviani, E., Mucchetti, G. (2014). Invited review: Microbial evolution in raw-milk, long-ripened cheeses produced using undefined natural whey starters. *Journal of Dairy Science, 97,* 573-591.

Gobbetti, M., Neviani, E., Fox, P. (2018). The Cheeses of Italy: Science and Technology. (1st ed.) *Ed. Springer International,* (Chapter 6).

Griffiths, M.W., Tellez, A.M. (2013). *Lactobacillus helveticus*: the proteolytic system. *Frontiers in Microbiology, 4,* 30.

Hickey, D.K., Kilcawley, K.N., Beresford, T.P., Wilkinson, M.G. (2007). Lipolysis in cheddar cheese made from raw, thermized, and pasteurized milks. *Journal of Dairy Science, 90,* 47-56.

Holland, R., Liu, S.Q., Crow, V.L., Delabre, M.L., Lubbers, M., Bennett, M., Norris, G. (2005). Esterases of lactic acid bacteria and cheese flavour: Milk fat hydrolysis, alcoholysis and esterification. *International Dairy Journal, 15,* 6-9.

Inglin, R.C., Meile, L., & Stevens, M. (2018). Clustering of Pan- and Core-genome of *Lactobacillus* provides Novel Evolutionary Insights for Differentiation. *BMC genomics, 19,* 284.

Jebava, I., Plockova, M., Lortal, S., and Valence, F. (2011). The nine peptidoglycan hydrolases genes in *Lactobacillus helveticus* are ubiquitous and early transcribed. *International Journal of Food Microbiology, 148,* 1–7.

Kandler, O. and Weiss, N. (1986). Regular nonsporing Gram-positive rods. *In Bergey's Manual of Systematic Bacteriology, Vol. 2,* pp. 1208–1234.

Kanehisa, M., Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research, 28,* 27-30.

Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.H., Rognes, T., Ussery, D.W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research, 35,* 3100-3108.

Le Breton, Y., Pichereau, V., Sauvageot, N., Auffray, Y., Rincé, A. (2005). Maltose utilization in *Enterococcus faecalis*. *Journal of Applied Microbiology, 98,* 806-813.

Marcó, M.B., Moineau, S., Quiberoni, A. (2012). Bacteriophages and dairy fermentations. *Bacteriophage, 2,* 149-158.

Mattu, B., Chauhan, A. (2013). Lactic Acid Bacteria and Its Use in Probiotics. *Journal of Bioremediation & Biodegradation, 4,* e140.

Mayo, B., Ammor, M. S., Delgado, S., Alegría, A. (2010). Fermented milk products. *Fermented foods and beverages of the World,* pp. 263–288.

Mayo, B., van Sinderen, D., & Ventura, M. (2008). Genome analysis of food grade lactic Acid-producing bacteria: from basics to applications. *Current genomics, 9,* 169-183.

Ruhal, R., Kataria, R., & Choudhury, B. (2013). Trends in bacterial trehalose metabolism and significant nodes of metabolic pathway in the direction of trehalose accumulation. *Microbial Biotechnology, 6,* 493-502.

Schmid, M., Muri, J., Melidis, D., Varadarajan, A.R., Somerville, V., Wicki, A., Moser, A., Bourqui, M., Wenzel, C., Eugster-Meier, E., Frey, J.E., Irmler, S., Ahrens, C.H. (2018). Comparative Genomics of Completely Sequenced *Lactobacillus helveticus* Genomes Provides Insights into Strain-Specific Genes and Resolves Metagenomics Data Down to the Strain Level. *Frontiers in Microbiology, 9,* 63.

Senan, S., Prajapati, J.B., Joshi, C.G. (2014). Comparative genome-scale analysis of niche-based stress-responsive genes in *Lactobacillus helveticus* strains. *Genome, 57,* 185-192.

Valence, F., Deutsch, S.M., Richoux, R., Gagnaire, V., Lortal, S. (2000). Autolysis and related proteolysis in Swiss cheese for two *Lactobacillus helveticus* strains. *Journal of Dairy Research, 67,* 261–271.

Vinogradov, E., Valence, F., Maes, E., Jebava, I., Chuat, V., Lortal, S., et al. (2013). Structural studies of the cell wall polysaccharides from three strains of *Lactobacillus helveticus* with different autolytic properties: DPC4571, BROI, and LH1. *Carbohydrate Research, 379,* 7–12.

Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., Wishart, D.S. (2011). PHAST: A Fast Phage Search Tool. *Nucleic Acids Research, 39,* 347–352.

Chapter 4

Do technological and environmental pressures shape the microbial ecosystem? The case of natural whey starter of Parmigiano Reggiano cheesemaking

Gaia Bertani[$], Alessia Levante[$], Camilla Lazzi, Monica Gatti and Erasmo Neviani

[$]Both authors contributed equally to this manuscript

Manuscript in preparation

Abstract

Natural starter cultures are complex multiple-strains cultures community of thermophilic lactic acid bacteria, whose structure could be affected by small changes in the parameters of the process and differences in the acidity and pH. This study aims to obtain a deeper insight into the modification of microbial composition of Parmigiano Reggiano (PR) natural whey starters under environmental and technological pressures. To reach this purpose, two different production lines were monitored ("Conventional" and "Organic"), from August to October 2017, sampling 91 whey starter during different steps of PR cheesemaking. RealTime qPCR and High Throughput Sequencing showed that application of technological steps affect the microbial composition of the NWS, which is dominated by the species *L. helveticus* and *L. delbrueckii* ssp. *lactis*. Intriguingly, the cooking step that is performed during PR manufacturing leads to an increase of the abundance of *L. delbrueckii* ssp. *lactis* in both the production lines, with a concomitant decrease of the species *L. helveticus*. Nevertheless, overnight incubation of the NWS, that is used in the following day production, allows an increase of the population of the latter species, re-establishing the microbial composition. Our results suggest a different behaviour of these two thermophilic species during the production phases, probably due to their different ability to resist under acidic or thermal selective pressures.

Keywords: Natural starter culture, Parmigiano Reggiano cheese, Microbial dynamics, RealTime qPCR, High Throughput Sequencing

Introduction

Parmigiano Reggiano (PR) is a Protected Designated Origin (PDO) cooked, long-ripened and hard-textured cheese exclusively produced in a defined geographical area (Gatti et al., 2014).

PR is manufactured employing raw cow's milk, using mixture of partially skimmed and whole milk. In detail, the evening milk is partially skimmed after overnight spontaneous creaming, and subsequently mixed in a 1:1 ratio with whole milk from the following morning milking. The vat milk is added of natural whey starter culture (NWS) and then addition of calf rennet powder, at 32–34°C, is used to achieve milk coagulation. As soon as curd reaches the proper firmness, it is broken down with the use of 'spino', a dedicated tool to obtain small-sized granules (approximately the size of rice grains), with a size distribution that is not too large (Iezzi et al., 2012). Afterwards the curd is heated up to 54°C (53–56°C) for 5-15 min under stirring. After cooking the granules previously obtained aggregate together at the bottom of the vat. Curd is removed from vat, manually cut in two parts and stored in a mold for 2 days, entering the subsequent manufacturing steps that will lead to the final product, as carefully reviewed elsewhere (Gatti et al., 2014; Gobbetti, Neviani and Fox, 2018). After curd cooking the resulting sweet whey is recovered and incubated at a decreasing temperature and used the next day for cheesemaking, therefore maintaining a microbiological *liaison* among batches of production. During the first few hours of cheesemaking, the thermophilic microorganisms present in NWS grow in the curd and play a key role in the acidification step (Johnson et al., 2013). Thus, the thermophilic flora able to resist to a cooking temperature is crucial for the cheese ripening (De Dea et al., 2008; Gatti et al., 2008). NWS cultures are complex multiple-strain cultures community of lactic acid bacteria (LAB) composed by a mixture of thermophilic microorganisms (Giraffa et al., 1999; Giraffa et al., 2000; Lombardi et al., 2002; Giraffa et al., 2004). In a previous work (Bottari et al, 2010), with the use of techniques such as Length Heterogeneity (LH)-PCR and Fluorescence In Situ Hybridization (FISH) it has been

demonstrated that *Lactobacillus helveticus* was found in the highest percentage in all NWS used for PR cheese analysed. *Lactobacillus delbrueckii* ssp. *lactis* was also commonly present in varying amounts, while the species *Lactobacillus fermentum* and *Streptococcus thermophilus* were occurred in low concentrations (Bottari et al., 2010). Small changes in the parameters of the process like temperature of curd cooking treatment, temperature and modality of whey cooling and differences in the acidity and pH, could affect the bacterial consortium present in NWS (Gatti et al., 2014; Rossetti et al., 2008). In this study 91 samples of NWS were collected from one dairy located in the province of Parma, belonging to the Parmigiano Reggiano PDO cheese area of production, in a period ranging from August to October 2017. Fifty-seven whey cultures were collected just before addition to the vat milk, while seventeen samples were collected after the use of 'spino' and the curd cooking treatment. In detail, the NWS belonged to two different production lines: a "Conventional" NWS and a "Organic" NWS (NWS). The first production line uses milk that respects the disciplinary regulations provided for PR manufacturing (Single document of the PDO Parmigiano Reggiano), while the "Organic" one uses milk from breeders certifying that cows only eat organic feed and are treated with homeopathic substances, excluding the use of antibiotics.

This study was performed in order to understand the dynamics of microbial structure of NWS over the seasons, in response to the technological steps and to different technological parameters. The aim of this work was: (i) to characterise the microbial community of the two NWS collected to develop traditional knowledge (ii) to investigate how selective and environmental forces (like the seasonality and different temperatures in the technological process) shape the microbial community typical of NWS.

Materials and methods

Whey starter samples: recruitment and collection

Ninety-one natural whey starter used in PR manufacturing were sampled from one dairy located in the province of Parma belonging to the Parmigiano Reggiano PDO cheese area of production, selected to carry out the study. Two different production lines were monitored; a "Conventional" production line that respects the disciplinary regulations provided for PR manufacturing and a "Organic" line that provides the use of milk, for which breeders certifying that cows, treated with homeopathic substances, eat only organic feed. The samples, for each production line, were collected during different steps of PR cheesemaking: 57 NWS samples were collected just before the addition to the vat milk, 17 samples after the use of 'spino', i.e. right after the curdling step, and 17 samples after the curd cooking step, often referred to as sweet whey. The 57 NWS samples were collected three times a week in a period from August to October, while the whey after the use of 'spino' and the sweet whey were sampled once a week in the same period. Samples were shipped to the laboratory under refrigerated conditions, and immediately stored at -20°C.

DNA extraction and quantification

Microbial DNA extraction was performed using DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) modified as follow. The following steps were performed to remove fat and milk impurities: 1 mL of natural whey starter culture was mixed with 2 % (w/v) of tri-sodium citrate solution at 50°C in 10 mL final volume for 30'. Regarding samples of whey starter after the use of 'spino' and sweet whey, 10 mL were mixed with 2 % (w/v) of sodium citrate solution at 50°C in 20 mL final volume for 30'. The mixture was centrifugated at 10000 rpm for 10 min at 4°C. The supernatant and fat layers were removed, and the cells were re-suspended in 1 mL of the same solution, repeating the washing step for 2 times until all impurities were removed. The cell suspensions were centrifuged at 10000 rpm for 10 min at 4°C. Subsequently, the manufacturer's

protocol for DNA extraction from Gram-positive bacteria was followed, by doubling the reagents volumes. Briefly, the cells were lysed in 360 µl of lysis buffer containing 25 mg/mL of lysozyme for 30' at 37 °C. The lysed cell suspension was protease treated for 30' at 56 °C. At the end of the spin-column protocol, the DNA was eluted with 60 µl of nuclease-free water, and the concentration and purity of the extracted nucleic acids were determined by Nanodrop (NanoDrop™ 2000, ThermoFisher Scientific, Waltham, Massachusetts, USA).

RealTime qPCR assay

The absolute quantification of *L. helveticus* and *L. delbrueckii* ssp. *lactis* species, was performed on the 91 samples collected, using specific primers (Table 1) designed on *pheS* gene sequences (Bottari et al., 2013). A master mix for each primer pairs contained: 1× PowerUp SYBR Green Master Mix (ThermoFisher Scientific, Milan, Italy), forward and reverse primers at concentration of 250 nM and nuclease-free water to a total of 20 µl per well. All the reactions were performed in duplicate, and no template controls were included in each experiment. The total DNA, previously extracted, was diluted 10-fold with nuclease free water and added to the reaction in a 5 µl volume. The plate, after a short centrifugation, was placed in the QuantStudio® 3 instrument (Thermo Fisher Scientific), the thermic cycle was as follows: a first hold stage of 2 minutes at 50 °C followed by 10 minutes at 95°C, 40 cycles of 15 s at 95°C and 1 minute at 60°C, during which fluorescence acquisition took place, and a final melting curve stage from 60° to 95°C with a temperature gradient of 0.1°C/s.

With regard to absolute quantification, standard curves have been constructed using purified genomic DNA of type strains of *L. helveticus* and *L. delbrueckii* ssp. *lactis*, *pheS* copy number was calculated as described in Bottari et al. (2013). The standard curves were constructed from serially 10-fold diluted reference strains DNA at known copy number, covering a dilution range of 6 orders of magnitude, and plotting the resulting threshold cycles (Ct), against the logarithm of the *pheS*

copy number. The copy number of *pheS* gene of each species was calculated for the 91 natural whey starter samples by comparing the Ct of the sample with that of the respective standard curve.

Table 1 Primer pairs used in this study, the sequences are the same as reported in Bottari et al., 2013.

Primer	Primer sequence (5'→3')	Lenght (bp)	Size	GC%
LlpheSF	ACGTTGACGCTGACCACC	18	51	60
LlpheSR	GGCTTGAACTGGTGAAGTCTG	21		
LhpheSF	TTGATGGTGAAGACTTGCTTAGAA	24		
LhpheSR	CTCTGGCTTGGTCACCTGAA	20	51	45

Microbial community profiling

Microbial diversity was evaluated in 34 samples (17 NWS and 17 sweet whey), by sequencing the amplified V3-V4 region of the 16S rRNA gene using primers and PCR conditions previously described in Takahashi et al. (2014). Library preparation and sequencing were carried out by BMR Genomics (Padova, Italy), on a MiSeq platform (Illumina Italy s.r.l., Milan, Italy), leading to 300bp, paired-ends approach. The quality of the 16S rRNA amplicon raw reads was checked through FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Using PrinSeq (http://prinseq.sourceforge.net/) raw reads shorter than 380 bp were discarded. Through FLASH (Magoč and Salzberg, 2011), paired-ends reads, without primers and barcodes, were joined. Sequences merged were then analysed using QIIME 1.9.1 software (Caporaso et al., 2010). The OTUs were picked at 99% of similarity and the most representative sequences of each cluster were selected in order to assign taxonomy, using the database Greengenes. All samples were rarified to 6,250 reads per sample and singleton OTUs were discarded for subsequent analysis, to prevent any biases, due to different depths of sequencing. Alpha diversity (Good's coverage, Chao1 richness and Shannon diversity) was carried out using QIIME.

Results

Absolute quantification

The absolute abundances of *L. helveticus* and *L. delbrueckii* ssp. *lactis* used as starter for PR manufacturing were evaluated through a culture-independent approach, RealTime qPCR assay. To estimate the absolute quantification of these species, specific primers designed on *pheS* gene sequences were used (Bottari et al., 2010). Standard calibration curves from a purified genomic DNAs of type strains of *L. helveticus* and *L. delbrueckii* ssp. *lactis*, with known concentration was performed. The resulting curves were linear, with R^2 in the range 0.95–1.00 for both species. Calibration curves were exploited to assess for these species the absolute abundances in 91 whey starter samples, belonging to two different production lines ("Conventional" and "Organic"), and collected in different steps of PR cheesemaking. RealTime qPCR data, shown in Figure 1, represented the mean of "Conventional" and "Organic" samples collected, during PR cheesemaking. In detail, all the NWS collected showed a predominance of *L. delbrueckii* ssp. *lactis* (statistically significant p value < 0.05) in the two production lines, showing a mean concentration of about 7.35 ± 0.53 log copy n/mL in the conventional line, and a mean value of 7.39 ± 0.44 log copy n/mL in the organic one. On the other hand, *L. helveticus* species showed a different absolute amount (statistically significant, p value < 0.05) in "Conventional" and "Organic" production lines, reaching a mean value of 7.20 ± 0.18 log copy n/mL and 7.32 ± 0.16 log copy n/mL respectively. When the PR cheesemaking starts, it can be observed a decrease of the abundances of *L. helveticus* and *L. delbrueckii* ssp. *lactis*, reaching in 'spino' samples a value of about 6.41 ± 0.20 and 6.98 ± 0.42 log copy n/mL. This decrease is only apparent, since it's due to the dilution of the NWS into the vat milk, with a ratio of about 1:100, showing that, indeed, there has been a quick adaptation of the NWS to the vat environment, and an overall increase of the two species of about one order of magnitude. After curd cooking a mean value of 6.51 ± 0.15 and 7.76 ± 0.28 log copy n/mL is

observed for *L. helveticus* and *L. delbrueckii* ssp. *lactis* respectively, showing that the latter species has a better adaptation to the vat milk substrate and an increased thermal stress tolerance. Intriguingly, any statistical difference among the "Conventional" and "Organic" lines, found initially in NWS samples, was lost during the cheesemaking process.

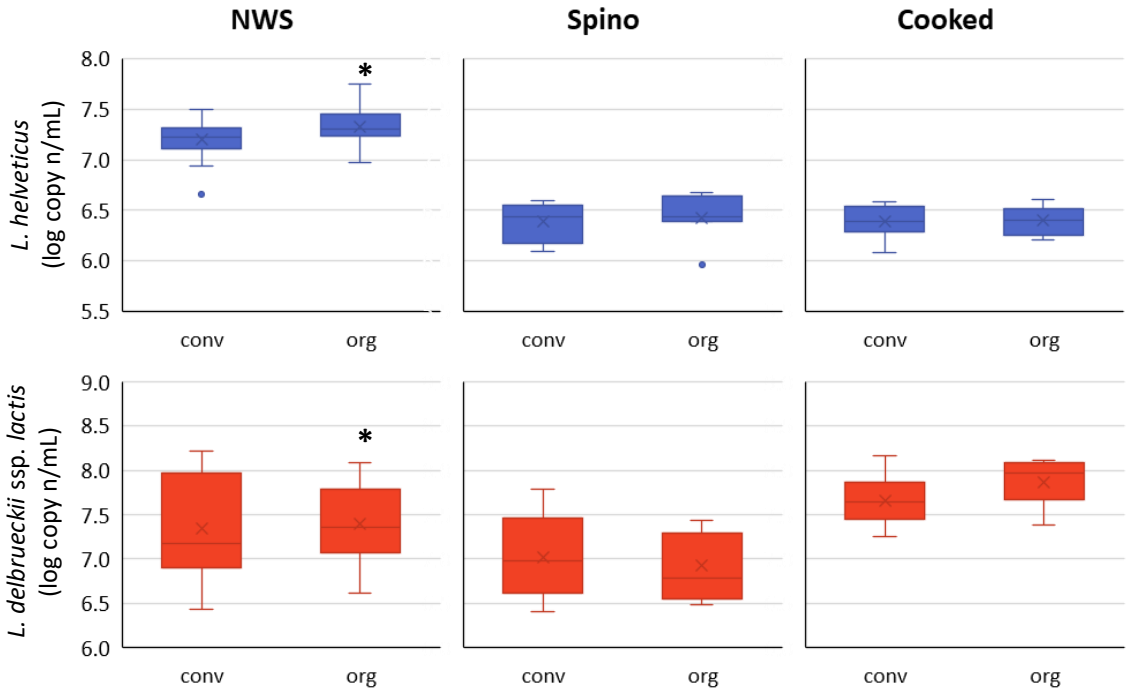


Figure 1 Boxplot of the absolute quantification of the species *L. helveticus* and *L. delbrueckii* ssp. *lactis* in the different samples collected. NWS plot described the samples collected after overnight acidification of the whey, 'spino' samples are sampled after addition to the vat milk, and cooked samples after exposure to the cooking step. * Statistically significant difference, $p < 0.005$.

Figure 2 shows the relative abundance, expressed as percentage, of *L. helveticus* and *L. delbrueckii* ssp. *lactis* in all samples collected in the two different production lines, over ten weeks, from August to October. The results obtained, show the prevalence of *L. delbrueckii* ssp. *lactis* compared to *L. helveticus*. During the ten weeks of sampling, a certain degree of variability can be observed in the data, more evidently in the "Organic" production line compared to the "Conventional" one.

Moreover, it's interesting to note that in all samples collected from "Conventional" line, *L. delbrueckii* ssp. *lactis* was the predominant species over the ten weeks (Figure 2). On the other hand, as regards "Organic" production line, during the first few weeks, *L. delbrueckii* ssp. *lactis* represented the highest percentage of all samples collected, while in the last few weeks, *L. helveticus* became the predominant species in whey cultures (Figure 2).

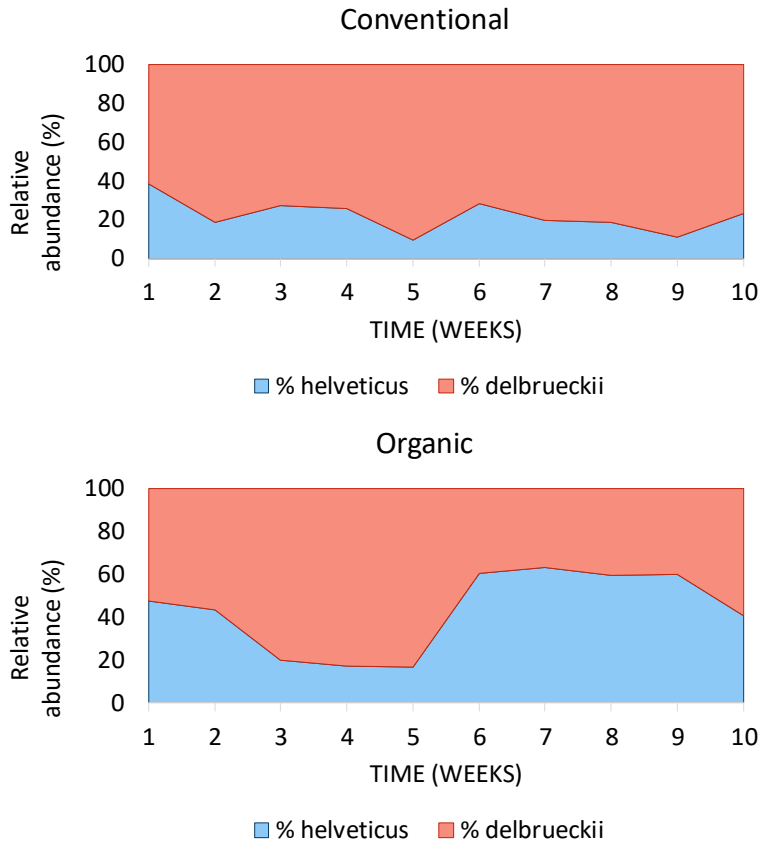


Figure 2 Average distribution of the relative abundance of the species *L. helveticus* and *L. delbrueckii* ssp. *lactis* during the 10 weeks sampling.

16s rRNA Gene High Throughput Sequencing

A total of 3,727,122 raw reads were obtained after the sequencing step, of which 884,189 passed the filtering steps performed using QIIME software pipeline, with an average value of 26,006 reads per sample, and an average read length of 468 bp. For each sample, indices describing the alpha-diversity of the microbial community were obtained and are reported in Table 2.

Table 2 Alpha diversity metrics for microbial communities in natural starter cultures, collected in "Conventional" and "Organic" production line.

Abbreviations: n: Samples collected before the addition to the vat milk

y: Samples collected after the curd cooking

Type	Sample	Cooking (y/n)	obs. OTUs	chao1	shannon	ESC
Conventional	NWS-7/9	n	1232 ± 18.44	5421.27 ± 416.38	5.73 ± 0.05	0.85 ± 0
	NWS-27/9	n	1312.9 ± 28.29	5307.19 ± 343.92	6.29 ± 0.04	0.84 ± 0
	NWS-18/9	n	1326.4 ± 24.73	5283.88 ± 361.09	6.03 ± 0.05	0.84 ± 0
	NWS-31/8	n	1268.8 ± 16.03	4941.34 ± 228.95	6.33 ± 0.03	0.85 ± 0
	NWS-4/8	n	1356.1 ± 28.38	5684.44 ± 331.57	6.04 ± 0.06	0.83 ± 0
	NWS-7/8	n	1387.5 ± 25.47	5725.98 ± 466.4	6.03 ± 0.04	0.83 ± 0
	NWS-16/8	n	1334.4 ± 23.06	5545.74 ± 425.93	6 ± 0.06	0.84 ± 0
	NWS-23/8	n	1241.3 ± 16.05	4768.54 ± 160.23	6.37 ± 0.03	0.85 ± 0
	NWS-13/9	n	1284.3 ± 19.64	4890.46 ± 447.88	6.2 ± 0.04	0.85 ± 0
	NWS-6/10	n	1230.4 ± 20	4458.67 ± 311.28	6.3 ± 0.04	0.85 ± 0
	NWS-c-7/9	y	1253.7 ± 18.86	4943.3 ± 439.22	5.99 ± 0.05	0.85 ± 0
	NWS-c-27/9	y	1220.8 ± 11.72	4949.73 ± 427.56	6.21 ± 0.04	0.85 ± 0
	NWS-c-18/9	y	1149.6 ± 21.21	4733.55 ± 314.33	5.94 ± 0.06	0.86 ± 0
	NWS-c-31/8	y	1256.9 ± 23.64	4675.01 ± 258.64	6.38 ± 0.03	0.85 ± 0
	NWS-c-4/8	y	1361.6 ± 29.86	4852.38 ± 232.65	6.51 ± 0.05	0.84 ± 0
	NWS-c-7/8	y	1360.1 ± 33.06	5369.89 ± 335.92	6.27 ± 0.07	0.84 ± 0.01
	NWS-c-16/8	y	1293.7 ± 24.41	5394.88 ± 372.51	6.25 ± 0.05	0.84 ± 0
	NWS-c-23/8	y	1209.8 ± 18.24	4759.47 ± 294.9	6.2 ± 0.06	0.85 ± 0
	NWS-c-13/9	y	1264.7 ± 22.3	5029.55 ± 372.28	6.06 ± 0.05	0.84 ± 0
	NWS-c-6/10	y	1325.9 ± 20.1	5079.62 ± 461.23	6.46 ± 0.04	0.84 ± 0
Organic	ORG-7/9	n	1240.3 ± 20.73	5432.77 ± 360.7	4.78 ± 0.05	0.84 ± 0
	ORG-29/9	n	1151.4 ± 14.24	4985.08 ± 438	4.43 ± 0.04	0.86 ± 0
	ORG-1/9	n	1394.4 ± 18	5566.55 ± 260.25	6.62 ± 0.05	0.83 ± 0
	ORG-8/8	n	1330.4 ± 25.33	5247.08 ± 394.63	5.76 ± 0.07	0.84 ± 0
	ORG-16/8	n	1330.7 ± 30.99	5567.09 ± 612.54	6.54 ± 0.04	0.84 ± 0.01
	ORG-23/8	n	1263 ± 0	5748.1 ± 0	6.33 ± 0	0.85 ± 0
	ORG-10/10	n	1042.5 ± 17.19	4741.83 ± 435.55	3.9 ± 0.07	0.87 ± 0
	ORG-c-7/9	y	1220.7 ± 11.18	5460.18 ± 295.31	5.61 ± 0.05	0.85 ± 0
	ORG-c-29/9	y	1124.9 ± 24.39	4784.23 ± 391.92	5.23 ± 0.05	0.86 ± 0
	ORG-c-1/9	y	1173.9 ± 23.89	4743.01 ± 443.4	6.09 ± 0.04	0.86 ± 0
	ORG-c-8/8	y	1344.9 ± 19.01	5566.02 ± 279.15	6.3 ± 0.04	0.84 ± 0
	ORG-c-16/8	y	1245.9 ± 21.89	5174.25 ± 302.43	6.37 ± 0.03	0.85 ± 0
	ORG-c-23/8	y	1214.9 ± 24.42	5047.92 ± 441.72	6.23 ± 0.06	0.85 ± 0
	ORG-c-10/10	y	1093.6 ± 21.9	4985 ± 468.59	4.94 ± 0.05	0.86 ± 0

Bioinformatic analysis allowed to identify a total of 102 OTUs, of which only eight have a relative abundance higher than 0.1 % in at least five samples, as reported in Figure 3 for "Conventional" and "Organic" farming samples. *L. helveticus* and *L. delbrueckii* were the major OTUs found in all analysed samples, with mean values of 52% and 34%, respectively. Regarding "Conventional" samples the relative abundance of the species *L. delbrueckii* ranged from a minimum of 16.5% to a maximum of 60.1%, while for the species *L. helveticus* the minimum value is 27.9% and the maximum value is 63.5% (Figure 3). On the other hand, the results measured in "Organic" samples are of a minimum of 4% and a maximum of 57.1% for *L. delbrueckii*, and values ranging from 28.9% to 94.9% for *L. helveticus*, which is significantly higher than the values found in conventional samples (T-test, $p < 0.05$). Interestingly, in all the samples except for one there is an increase in the relative abundance of the species *L. delbrueckii* after the cooking step, that varies from a minimum of 4.1% to a maximum of 39.5%, both measured in "Conventional" samples, that show a mean value of 11.7 ± 12.7 % (mean ± standard deviation). Conversely, the percentage increase of the species *L. delbrueckii* is less variable in "Organic" samples, with a mean value of 20.8% ± 5%. Interestingly, the genus *Streptococcus* spp. is found in all samples, with mean values corresponding to 8,1 ± 6.4% for organic farming and 11.5 ± 3,6% for conventional farming. As for the other minor species, unidentified *Lactobacillus* spp. was only found in organic farming samples, at less than 3%, while unidentified *Streptococcus* spp. and *Enterobacter* spp. were present below 1%.

Regarding the effect of cooking, this step significantly increased the relative abundance of *L. delbrueckii* in samples belonging to the "Organic" production line, where the average shifts from 18.47 ± 15.9 % to 39.3 ± 15.7% after the cooking step (T-test, $p < 0.005$). The same effect is observed in "Conventional" samples, too, where the relative abundance passes from 31 ± 12.9% in the starting NWS sample to 42.8 ± 12.1% in the cooked whey sample. Interestingly, the species that

is significantly affected by the cooking step in the "Conventional" line samples is *L. helveticus*, that decreases in the relative abundance from 52.7 ± 11.5% to 40 ± 12.1% (T-test, $p < 0.05$). Finally, "Conventional" samples after the cooking step show a significant increase of *Enterobacter* spp. and *Streptococcus* spp., despite retaining values that are less than 1% of the relative abundance (T-test, $p < 0.05$).

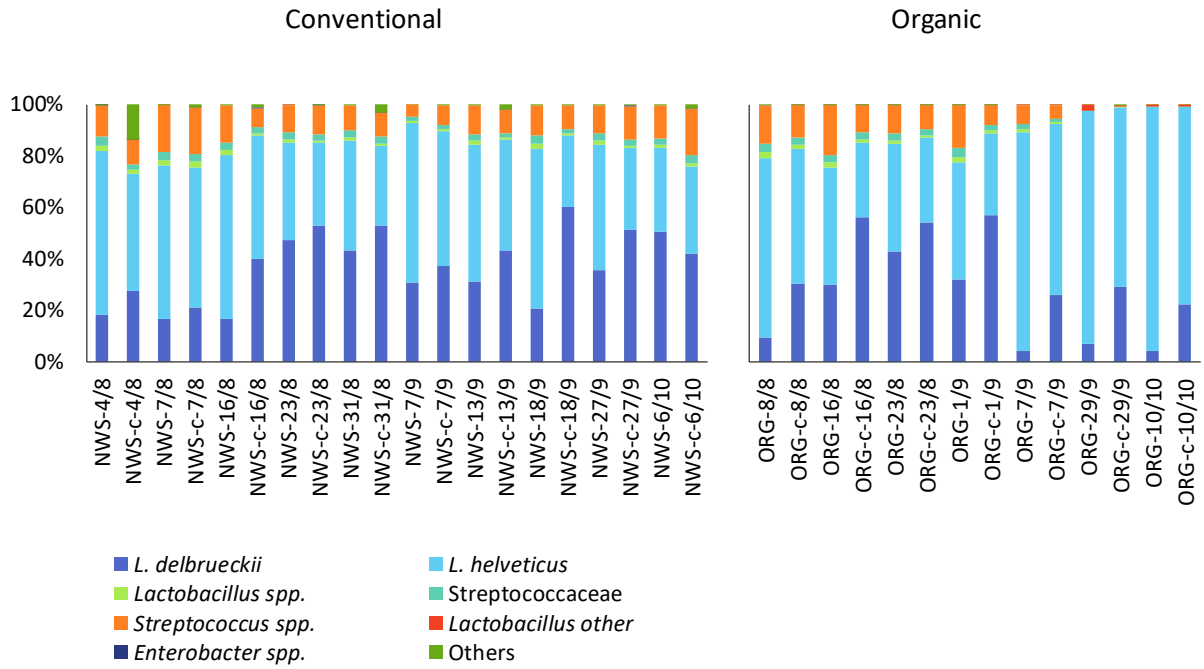


Figure 3 Incidence of OTUs based on 16S rRNA gene pyrosequencing at different sampling times. The graphs are divided according to the production line ("Conventional" or "Organic"), the NWS and cooked samples from the same date are shown next to each other. The bar graph describes the relative abundance of each species at different sampling date from August to October. Only the OTUs whose abundance is higher than 0.1% in at least 90% of the samples are reported, the remaining OTUs are grouped in the category "Others".

Abbreviations: NWS-c-: Samples collected after the curd cooking

ORG-c-: Samples collected after the curd cooking

Discussion

Nowadays it is increasingly common to study microbial community on high-throughput sequencing technology which, giving the relative abundance at taxon level, serves to achieve an overview of the community composition (Widder et al., 2016). Still, in order to estimate the community dynamics in complex bacterial cultures under technological and selective pressures, more detailed information is necessary, such as the absolute quantification to evaluate how community structure changes in different conditions (Philippot et al., 2009; Berry and Widder, 2014). This study aimed to investigate the complexity of natural starter cultures, used in PR cheesemaking, focusing on the modification of this microbial community, under environmental and technological pressures. To reach this purpose, two different production lines were monitored, from August to October 2017, sampling 91 whey starter during different steps of PR cheesemaking. The data obtained through RealTime qPCR assay showed different community dynamics in the collected samples. The NWS of this dairy plant shows a dominance of the species *L. delbrueckii* ssp. *lactis* over *L. helveticus*, differently from what reported in the prevalent literature about the NWS of Parmigiano Reggiano cheese, that consistently find an opposite distribution of the two thermophilic species (Cocconcelli et al., 1997; Coppola et al., 2000; Gatti et al., 2003; Gatti et al., 2004); nevertheless, this "inverted ratio" could be a peculiarity of the chosen dairy plant, since the intrinsically artisanal nature of the NWS leads to a great variability among the samples of different PR manufacturer. Intriguingly, it can be observed that the absolute abundances of *L. helveticus* and *L. delbrueckii* ssp. *lactis* in the NWS, prior to the start of the cheesemaking process, showed a significant difference between samples belonging to "Conventional" or "Organic" farming (p-value <0.05). This difference is subsequently abolished in response to the technological treatment, that is mainly represented by the curdling process, followed by the breaking step ('spino'), and the cooking step, consisting in a heating stress. These data could suggest that NWS samples, regardless of their initial different

composition, respond stably following the treatment. The technological process, in particular the addition of NWS to milk mixture at approximately 2.7 to 3.5%, the curdling formation and subsequent breaking and the cooking step at 54.7-55°C, probably are the selective pressures that shape the stability of the starter culture. It is well known from literature, that weak changes due to seasonal variations or technological parameters modifications, such as differences in the acidity and pH, could affect the bacterial consortium present in NWS (Gatti et al., 2014; Rossetti et al., 2008). Our results, in accordance to what reported in the literature, showed a great degree of variability as regard *L. helveticus* and *L. delbrueckii* ssp. *lactis* in both production lines, over the ten weeks.

Regarding the composition of the microbial community, the calculated alpha indices show that there are not significant differences in the microbial composition of the samples coming from both the production lines, despite the unusually high number of identified OTUs, the final composition that is still dominated by the main acidifying thermophilic species, that is in good agreement with findings observed from other studies (De Filippis et al., 2014; Giraffa et al., 1996; Bottari et al., 2010). Interestingly, the variations observed in terms of numerical variation of the amount of *L. helveticus* and *L. delbrueckii* ssp. *lactis* in response to the technological treatments agree with the data from high-throughput sequencing, that lead to similar observations. Nevertheless, the sole analysis of metagenomics data fails to depict if the observed change in the relative abundance of *L. helveticus* is due to an increase of the number of *L. delbrueckii* ssp. *lactis*, or to a progressive loss of viability of the strain itself. The analysis of quantitative data obtained through RealTime qPCR clearly shows that, taken into account the dilution effect due to the inoculation of NWS into the vat milk, roughly a 1:100 dilution ratio, *L. helveticus* quantity remains mostly unchanged after the curdling and cooking step, while the population belonging to *L. delbrueckii* ssp. *lactis* increases, resulting in a NWS composition that is different from the one observed before the technological step. Still, it has to be kept in mind the ‘cyclical’ nature of the NWS, that is slowly cooled down at

the end of the production, overnight, to be ready for the following day's production. To this respect, it is interesting to underline that the NWS undergoes two different thermal gradients: after inoculation into the milk vat, the temperature increases, exposing the bacteria to a thermal stress, while after separation of the whey from the curd, the temperature goes down, releasing the stringent thermal conditions. The data obtained, in agreement with other works (Neviani et al., 1995; Cocconcelli et al., 1997; Giraffa et al., 1998; Chamba, 2000), suggest that the two thermophilic species present in the whey appear to respond differently to the gradients and to the different composition of the environment during the production phases. Indeed, *L. helveticus* is regarded as a more acid-tolerant strain, and this might explain why this strain might increase in number after the overnight incubation of the NWS, reaching lower pH values. Conversely, the inoculated vat milk has higher pH values that, in combination with a possibly higher tolerance to the thermal stress of *L. delbrueckii* ssp. *lactis*, might explain the numerical increase observed after the phases of 'spino' and cooking.

Conclusions

This study provides an in-depth description of the natural starter cultures microbiota involved in Parmigiano Reggiano DOP Italian cheesemaking over a time series and in response to different technological steps, monitoring two different production lines ("Conventional" and "Organic"). In the early stage of the cheesemaking process, the results showed diverse absolute abundances of *L. helveticus* and *L. delbrueckii* ssp. *lactis*, in the two production lines. Once the technological process starts, the differences previously observed among these species disappeared, suggesting that the technological pressures, represented by the curdling process, followed by the breaking step ('spino'), and the cooking step, shape the structure of natural starter cultures. Over the ten weeks, from August to October, a great degree of variability was observed in both production lines, with regard to *L. helveticus* and *L. delbrueckii* ssp. *lactis*. Moreover, metagenomic analysis highlighted a different behaviour of these two thermophilic species during the production phases, in agreement with quantitative results obtained with RT-qPCR. *L. helveticus* is regarded as a more acid-tolerant strain and increases in number after the overnight incubation of the NWS reaching lower pH values, conversely, the higher pH values in combination with a higher tolerance to the thermal stress applied during cooking step of the cheesemaking lead to an increase of *L. delbrueckii* ssp. *lactis*. Overall, the data obtained highlighted that environmental and technological pressures shape the microbial structure of NWS, revealing also the different aptitude to resist under acidic or heat stress of the thermophilic moiety of the starter, that differently adapt to the technological pressure that characterizes this 'cyclical' ecosystem.

References

Berry, D., & Widder, S. (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontiers in Microbiology, 5,* 219.

Bottari, B., Agrimonti, C., Gatti, M., Neviani, E., Marmiroli, N. (2013). Development of a multiplex real time PCR to detect thermophilic lactic acid bacteria in natural whey starters. *International Journal of Food Microbiology, 160,* 290-297.

Bottari, B., Santarelli, M, Neviani, E., Gatti, M. (2010). Natural whey starter for Parmigiano Reggiano: culture independent approach. *Journal of Applied Microbiology, 108,* 1676–1684.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods, 7,* 335-336.

Chamba, J.F. (2000). L'emmental, an écosystème complex. Conséquences sur la sélection et l'utilisation des ferments. *Sciences des Aliments, 20,* 37-54.

Cocconcelli, P.S., Parisi, M.G., Senini, L., Bottazzi, V. (1997). Use of RAPD and 16S rDNA sequencing for the study of *Lactobacillus* population dynamics in natural whey culture. *Letters in Applied Microbiology, 25,* 8-12.

Coppola, R., Nanni, M., Iorizzo, M., Sorrentino, A., Sorrentino, E., Chiavari, C., et al. (2000). Microbiological characteristics of Parmigiano Reggiano cheese during the cheesemaking and the first months of the ripening. *Lait, 80,* 479–490.

De Dea Lindner, J., Bernini, V., De Lorentiis, A., Pecorari, A., Neviani, E., Gatti, M. (2008). Parmigiano Reggiano cheese: evolution of cultivable and total lactic microflora and peptidase activities during manufacture and ripening. *Dairy Science Technology, 88,* 511–523.

De Filippis, F., La Storia, A., Stellato, G., Gatti, M., and Ercolini, D. (2014). A selected core microbiome drives the early stages of three popular Italian cheese manufactures. *PLoS ONE, 9,* e89680.

Gatti, M., Bottari, B., Lazzi, C., Neviani, E., Mucchetti, G. (2014). Invited review: microbial evolution in raw milk, long-ripened cheeses produced using undefined natural whey starters. *Journal of Dairy Science 97,* 573–591.

Gatti, M., De Dea Lindner, J., De Lorentiis, A., Bottari, B., Santarelli, M., Bernini, V., Neviani, E. (2008). Dynamics of whole and lysed bacterial cells during Parmigiano-Reggiano cheese production and ripening. *Applied and Environmental Microbiology, 74,* 6161–6167.

Gatti, M., Lazzi, C., Rossetti, L., Mucchetti, G., and Neviani, E. (2003). Biodiversity in *Lactobacillus helveticus* strains present in natural whey starter used for Parmigiano Reggiano cheese. *Journal of Applied Microbiology, 95,* 463–470.

Gatti, M., Trivisano, C., Fabrizi, E., Neviani, E., Gardini, F. (2004). Biodiversity within *Lactobacillus helveticus* isolated from different natural whey starter cultures as revealed by classification trees. *Applied and Enviromental Microbiology, 70,* 182-190.

Giraffa, G., Andrighetto, C., Antonello, C., Gatti, M., Lazzi, C., Marcazzan, G., Lombardi, A., Neviani, E. (2004). Genotypic and phenotypic diversity of *Lactobacillus delbrueckii subsp. lactis* strains of dairy origin. *International Journal of Food Microbiology, 91,* 129-139.

Giraffa, G., Gatti, M., Rossetti, L., Senini, L., Neviani, E. (2000). Molecular diversity within *Lactobacillus helveticus* as revealed by genotypic characterization. *Applied and Environmental Microbiology, 66,* 1259.

Giraffa, G., Rossetti, L., Mucchetti, G., Addeo, F., Neviani, E. (1998). Influence of the temperature gradient on the growth of thermophilic lactobacilli used as natural starters in Grana cheese. *Journal of Dairy Science, 81,* 31-36.

Giraffa, G., Mucchetti, G., Neviani, E. (1996). Interactions among thermophilic lactobacilli during growth in cheese whey. *Journal of Applied Bacteriology, 80,* 199.

Giraffa, G., Neviani, E. (1999). Different *Lactobacillus helveticus* strain populations dominate during Grana Padano Cheese-making. *Food Microbiology, 16,* 205.

Gobbetti, M., Neviani, E., Fox, P. (2018). The Cheeses of Italy: Science and Technology. (1st ed.) *Ed. Springer International,* (Chapter 6).

Johnson, M.E. (2013). Mesophilic and Thermophilic Cultures Used in Traditional Cheesemaking. *Microbiology Spectrum, 1,* no. 1.

Iezzi, R., F. Locci, R. Ghiglietti, C. Belingheri, S. Francolino, and G. Mucchetti. (2012). Parmigiano Reggiano and Grana Padano cheese curd grains size and distribution by image analysis. *Lebenson. Wiss. Technol, 47,* 380–385.

Lombardi, A., Dal Maistro, L., De Dea, P., Gatti, M., Giraffa, G., Neviani, E. (2002). A poliphasic approach to highlight genotypic and phenotypic diversities of *Lactobacillus helveticus* strains isolated from dairy starter cultures and cheeses. *Journal of Dairy Research, 69,* 139-149.

Magoč, T., Salzberg, S.L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics, 27,* 2957-2963.

Neviani, E., Divizia, R., Abbiati, E., Gatti, M. (1995). Acidification activity of thermophilic lactobacilli under the temperature gradient of Grana cheese making. *Journal of Dairy Science, 78,* 1248-1252.

Philippot, L., Cuhel, J., Saby, N.P.A., Chèneby, D., Chronakova, A., Bru, D., et al. (2009). Mapping field-scale spatial distribution patterns of size and activity of the denitrifier community. *Environmental Microbiology, 11,* 1518–1526.

Rossetti, L., Fornasari, M.E., Gatti, M., Lazzi, C., Neviani, E., Giraffa G. (2008). Grana Padano cheese whey starters: microbial composition and strain distribution. *International Journal of Food Microbiology, 127,* 168-171

Widder, S., Allen, R.J., Pfeiffer, T., Curtis, T.P., Wiuf, C., Sloan, W.T., Cordero, O.X., Brown, S.P., Momeni, B., Shou, W., Kettle, H., Flint, H.J., Haas, A.F., Laroche, B., Kreft, J.U., Rainey, P.B., Freilich, S., Schuster, S., Milferstedt, K., van der Meer, J.R., Grobkopf, T., Huisman, J., Free, A., Picioreanu, C., Quince, C., Klapper, I., Labarthe, S., Smets, B.F., Wang, H. (2016). Challenges in microbial ecology: building predictive understanding of community function and dynamics. *ISME Journal, 10,* 2557-2568.

Summary and conclusions

Summary and conclusions

Bacteria have been mutating and evolving on Earth for billions of years (Bleuven et al., 2016).

'Survival of the fittest' is the key concept behind the Darwinian natural selection that drives the evolution of microorganisms (Brookifield et al., 2009).

Thanks to their short generation times and large population sizes, bacteria can evolve rapidly (Elena et al., 2003). Bacterial evolution concerns to the heritable genetic changes that a microorganism accumulates during its life time, which can arise principally from adaptations in response to environmental or technological changes (Loewe et al., 2010).

Food ecosystems can be viewed as different environments, characterized by diverse selective pressures (Senan et al., 2014), which lead bacteria to adapt rapidly, acquiring genes essential for their survival as well losing genetic traits not helpful for the cell.

The work presented in this thesis has focused on the evolution of LAB, used as starter in DOP Italian cheeses manufacturing over time and space. The research question has been performed through different approaches, analysing different dairy ecosystems, characterized by diverse stringency.

Previous studies carried out on dairy isolates of LAB had highlighted a great genotypic and phenotypic biodiversity in diverse ecological niches (Giraffa et al. 1998, Gatti et al. 1999, Giraffa et al. 2000, Lombardi et al. 2002, Rossetti et al. 2009 Gatti et al. 2004), nevertheless a depth exploration regards how genetic traits and microbial communities mutate over the time, and in different ecological niches, has been poorly investigated.

Our results represent one of the first evolutionary study of LAB over time and space.

To reach this fascinating purpose, at first, we have developed a new protocol to perform in a quickly, cheaply and simply way, Amplified Fragment Length Polymorphism (AFLP) analysis. AFLP is an excellent technique to differentiate strains or very closely related species and as a good

phylogenetic tool (Janssen et al. 1996; Blears et al. 1998; Savelkoul et al. 1999; Jarraud et al. 2002). The advanced protocol fine-tuned, gives the opportunities to reduce the time-consuming and labour-intensive, and costs effective given the possibilities to perform genotypic characterization of many samples, by being an excellent and easier tool to study genome evolution (Chapter 1).

Different dairy ecosystems stability, characterized by different stringency, were investigated. Environmental and technological pressures impose the modification of genome and structure of the cheese microbiota.

Pecorino Toscano cheese, is a specific, selective, environmental ecosystem modelling based on production of ewe's milk system (Gobbetti et al., 2018). The appearance of *S. thermophilus* tyramine-histamine producer strains, over the years, highlights the risk related to BAs production in the Pecorino Toscano cheese, suggesting a possible modification of Pecorino Toscano niche during the last two decades. This acquired aptitude, could be due to the use of selected strains coming from other dairy niches. Moreover, it could also be related to the modification of autochthonous microflora, as a consequence of horizontal gene transfer and/or technological pressures (Chapter 2).

Natural whey starter culture ecosystem, used for Parmigiano Reggiano and Grana Padano PDO cheeses, is obtained from the previous day's cheese-making whey incubated at a decreasing temperature, selecting, in this way, the growth of a large number of thermophilic LAB that, together with the raw milk microbiota, preserve from damage the final product and determine final cheese attributes (Gatti et al., 2014; Bassi et al., 2015). Therefore, this ‘cyclical’ ecological niche, is a highly selective ecosystem, which every day, is subjected to the same selective pressures such as acidifying stress and heat treatment.

The first three strains of *L. helveticus* isolated from starter cultures used in the production of Parmigiano Reggiano and Grana Padano cheese were here sequenced and compared to other eleven strains, through a bioinformatic approach. This comparison has identified highly desirable traits

with technological potential, like genes involved in a strong proteolytic system, genetic traits involved in autolytic and lipolytic mechanism, and finally the co-evolution of bacteriophages integrated into the chromosome (Chapter 3).

Besides, this interesting study, we have also investigated, for the first time, how selective and environmental forces (like the seasonality and different temperatures in the technological process) shape the microbial community typical of NWS. The experiments conducted, from August to October 2017, highlighted that environmental and technological pressures shape the microbial structure of NWS, revealing also the different aptitude to resist under acidic or heat stress of the thermophilic moiety of the starter, that differently adapt to the technological pressure that characterizes this 'cyclical' ecosystem (Chapter 4).

The works presented in this thesis, have demonstrated different approaches to study the evolution of bacteria over space and time. This study lays the foundations for understanding the factors that contribute to the selection, in order to direct the technology to favour a biotype rather than another in a specific environment. The biodiversity, that we found in different ecological niches, should be perceived as a great richness, the leak of these biotypes could mean losing technological traits that may reoccur useful for various applications.

Overall, following the Darwinian concept '*Survival of the fittest',* we suggest that the selection of starter should base on their ability to adapt under different environmental and technological drivers, because these microorganisms turn out to be the most suitable in technological terms. Alongside these considerations, our results suggest also to take under control the inclusion of BAs producer strains in starter cultures, above all, the need to supervise the genetic possible changes over time.

References

Bassi, D., Puglisi, E., Cocconcelli, P.S. (2015). Understanding the bacterial communities of hard cheese with blowing defect. *Food Microbiology, 52,* 106-118.

Blears, M., De Grandis, S., Lee, H., Trevors, J.T. (1998). Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. *Journal of Industrial Microbiology & Biotechnology, 21,* 99–114.

Bleuven, C., Landry, C.R. (2016). Molecular and cellular bases of adaptation to a changing environment in microorganisms. *Proceedings Biological Sciences,* 283.

Brookfield, J.F.Y. (2009). Evolution and evolvability: celebrating Darwin 200. (2008) *Biology Letters, 5,* 44-46.

Elena, S.F., Lenski, R.E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Reviews Genetics, 4,* 457-469.

Gatti, M., Bottari, B., Lazzi, C., Neviani, E., Mucchetti, G. (2014). Invited review: microbial evolution in raw milk, long-ripened cheeses produced using undefined natural whey starters. *Journal of Dairy Science 97,* 573–591.

Gatti, M., Contarini, G., Neviani, E. (1999). Effectiveness of chemometrical techniques in discriminating *Lactobacillus helveticus* biotypes from natural dairy starter cultures by phenotypic characteristics. *Applied and Environmental Microbiology, 65,* 1450-1455.

Giraffa, G., De Vecchi, P., Rossi, P., Nicastro, G., Fortina. M.G. (1998). Genotypic heterogeneity among *Lactobacillus helveticus* strains isolated from natural cheese starters. *Journal of Applied Microbiology, 85,* 411–416.

Giraffa, G., Gatti, M., Rossetti, L., Senini, L., Neviani, E. (2000). Molecular diversity within *Lactobacillus helveticus* as revealed by genotypic characterization. *Applied and Environmental Microbiology, 66,* 1259-1265.

Gobbetti, M., Neviani, E., Fox, P. (2018). The Cheeses of Italy: Science and Technology. (1st ed.) *Ed. Springer International,* (Chapter 6).

Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., Zabeau, M., Kersters, K. (1996). Evaluation of the DNA Fingerprinting Method AFLP as a New Tool in Bacterial Taxonomy. *Microbiology, 142*, 1881-1893.

Jarraud, S., Mougel, C., Thioulouse, J., Lina, G., Meugnier, H., Forey, F., Vandenesch, F. (2002). Relationships between *Staphylococcus aureus* Genetic Background, Virulence Factors, agr Groups (Alleles), and Human Disease. *Infection and Immunity, 70,* 631–641.

Loewe, L., Hill, W.G. (2010). The population genetics of mutations: good, bad and indifferent. *Philosophical Transactions of the Royal Society B: Biological Sciences, 27,* 1153-1167.

Lombardi, A., Dal Maistro, L., De Dea, P., Gatti, M., Giraffa, G., Neviani, E. (2002). A poliphasic approach to highlight genotypic and phenotypic diversities of *Lactobacillus helveticus* strains isolated from dairy starter cultures and cheeses. *Journal of Dairy Research, 69,* 139-149.

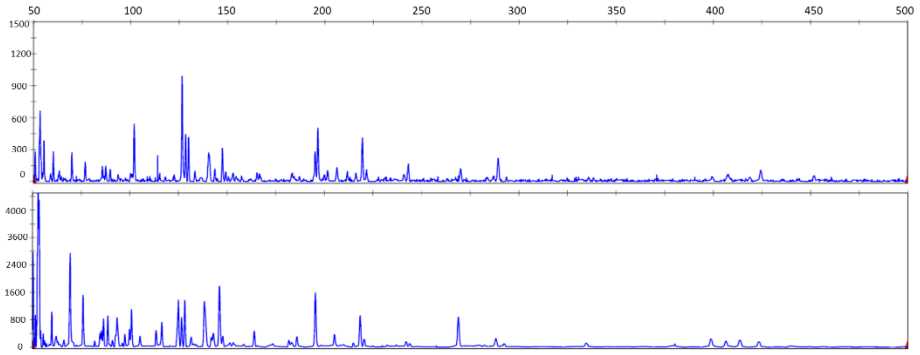
Rossetti, L., Carminati, D., Zago, M., Giraffa, G. (2009). A qualified presumption of safety approach for the safety assessment of Grana Padano whey starters. *International Journal of Food Microbiology, 130,* 70-73.

Savelkoul, P.H., Aarts, H.J., De Haas, J., Dijkshoorn, L., Duim, B., Otsen, M., Rademaker, J.L., Schouls, L., Lenstra, J.A. (1999). Amplified-fragment length polymorphism analysis: the state of an art. *Journal of Clinical Microbiology, 37,* 3083-3091.

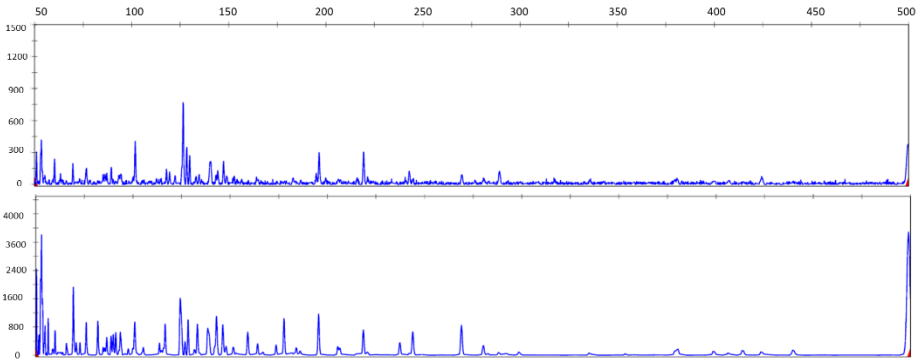
Senan, S., Prajapati, J.B., Joshi, C.G. (2014). Comparative genome-scale analysis of niche-based stress-responsive genes in *Lactobacillus helveticus* strains. *Genome, 57,* 185-192.

Appendix

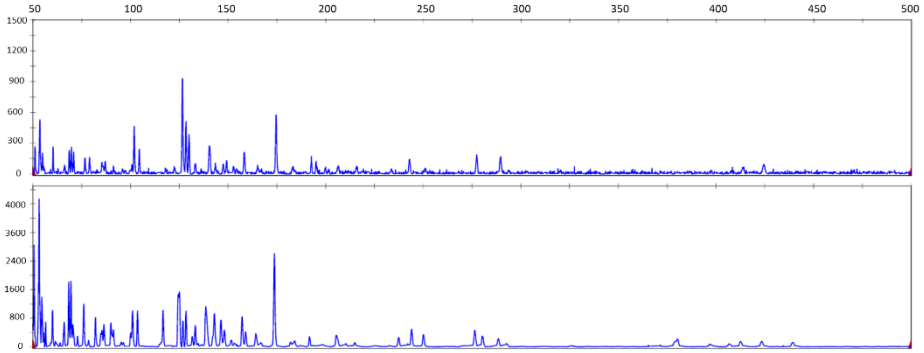
A


50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

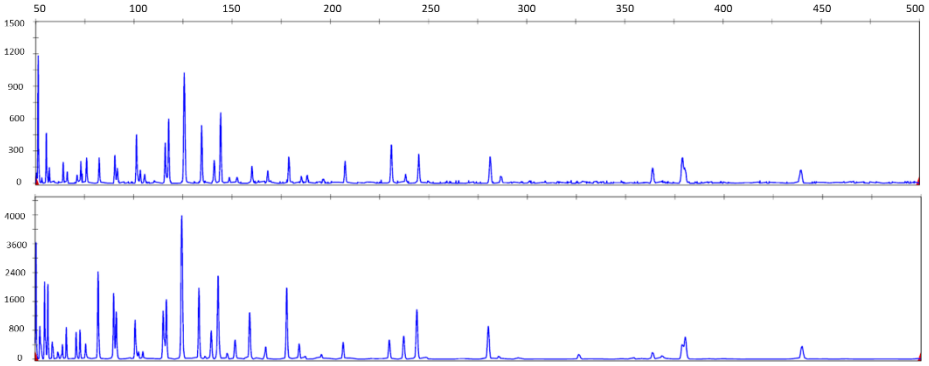
B

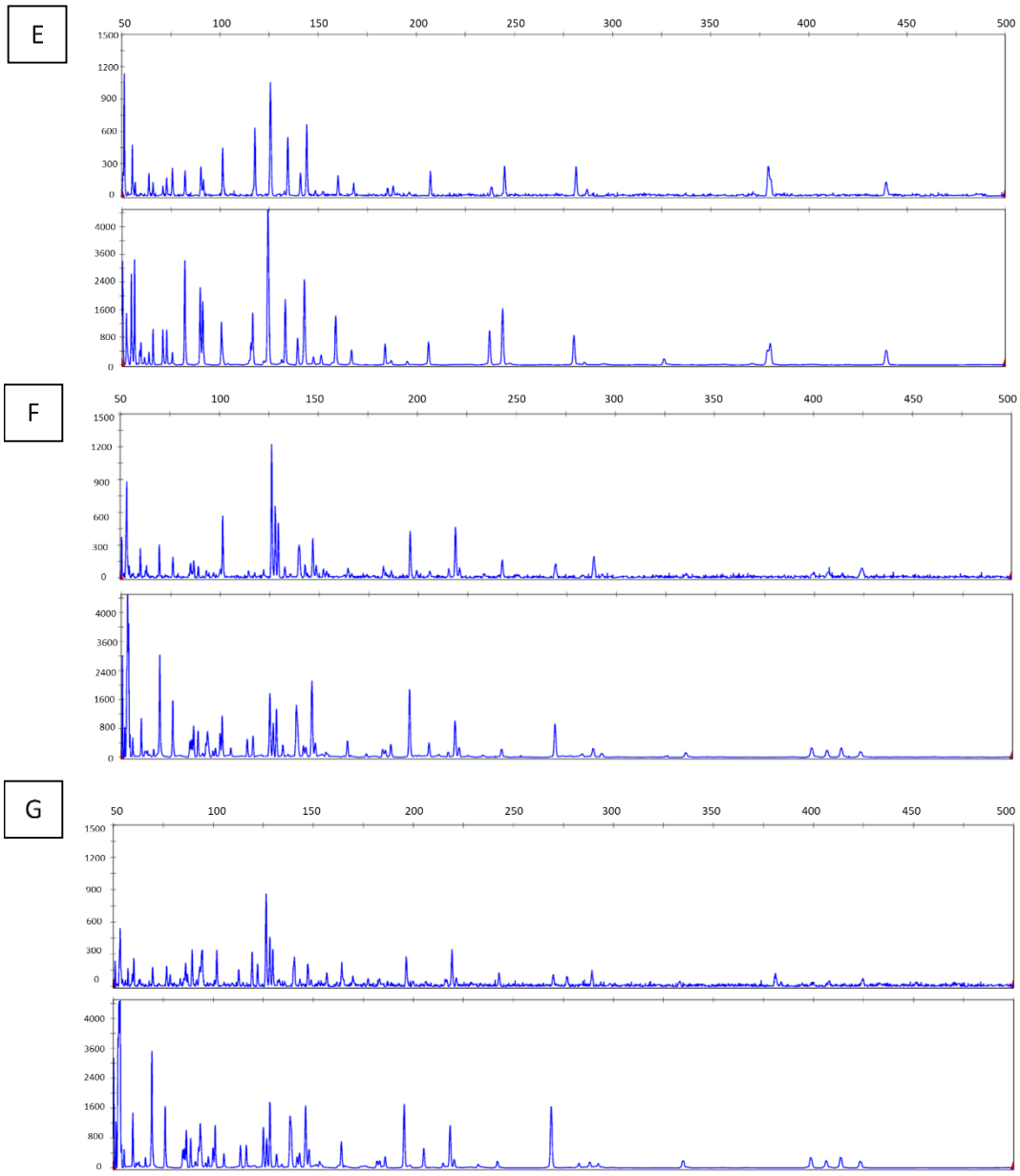

50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

C


50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

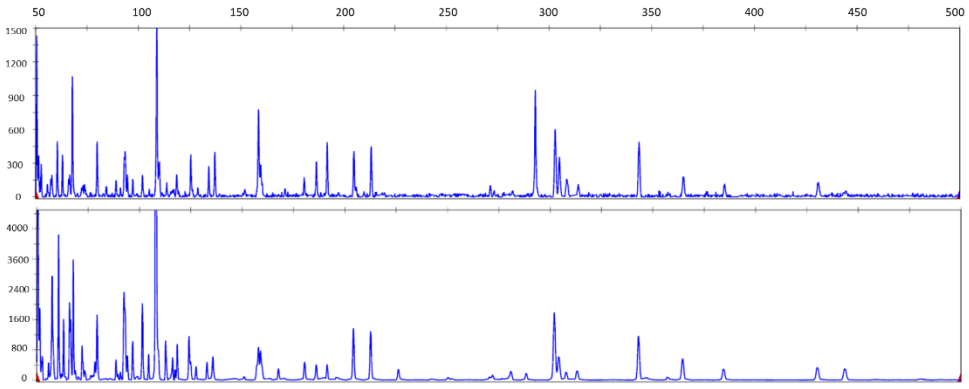
D


50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

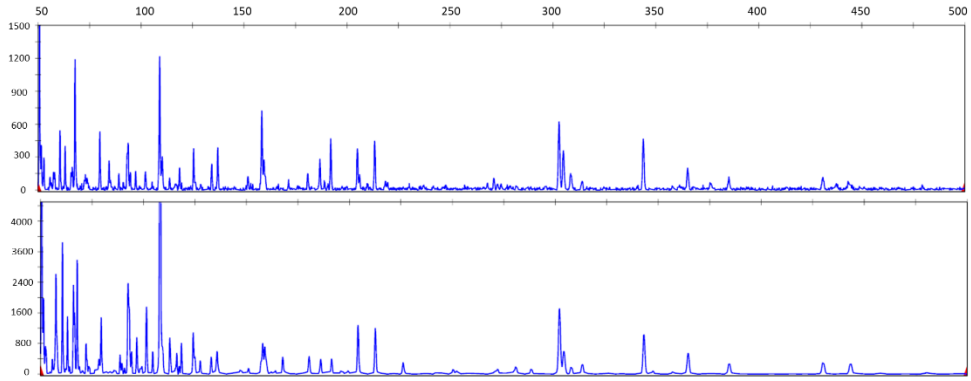


Supplementary Figure 1 Electropherograms of *S. thermophilus* strains AFLP profiles obtained with the AFLP microbial Kit (upward) and with the advanced AFLP protocol (below). A) AFLP profiles of strain 100; B) AFLP profiles of strain 159; C) AFLP profiles of strain 418; D) AFLP profiles of strain 4027; E) AFLP profiles of strain 4028; F) AFLP profiles of strain 4042; G) AFLP profiles of strain 145.

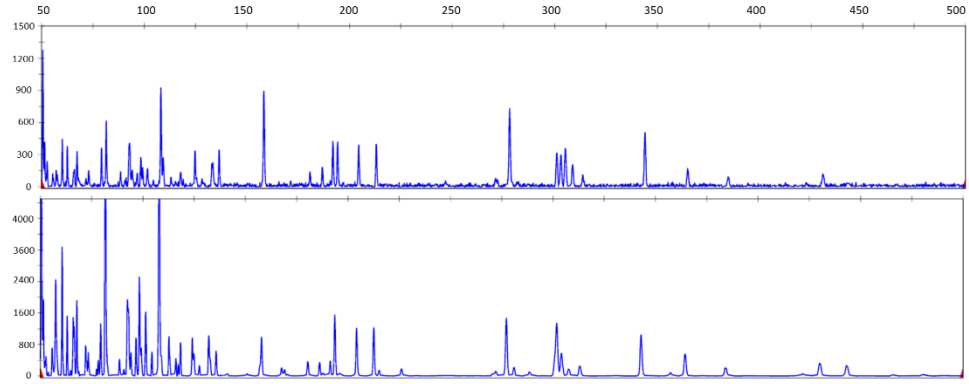
H


50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

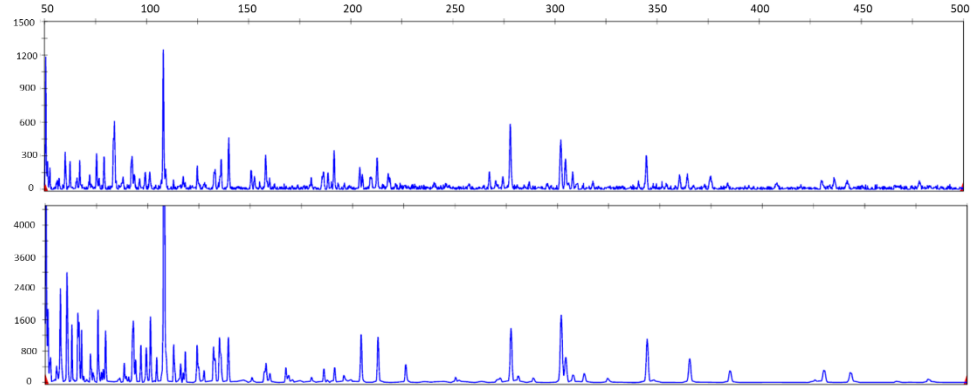
I

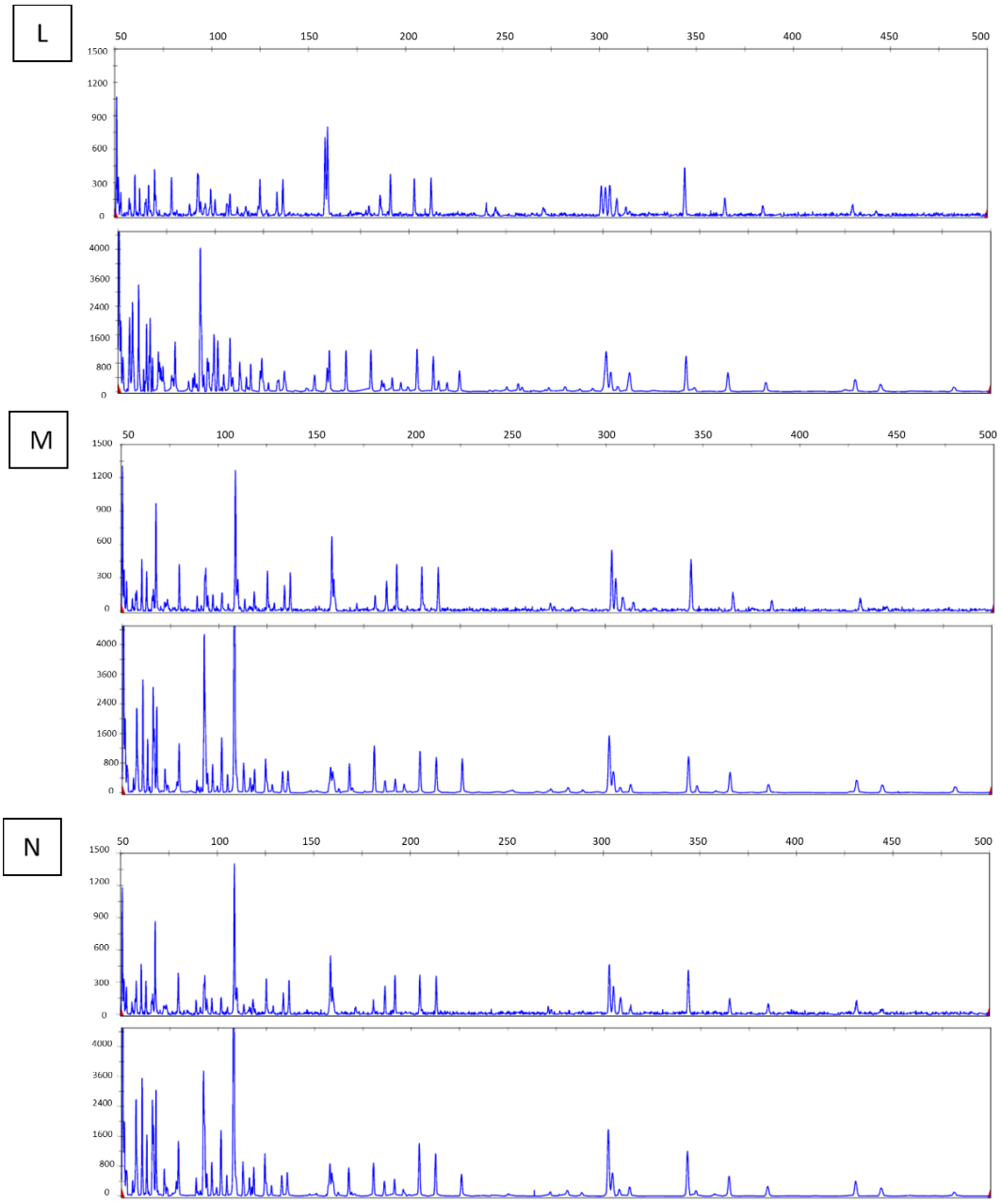

50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

J


50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

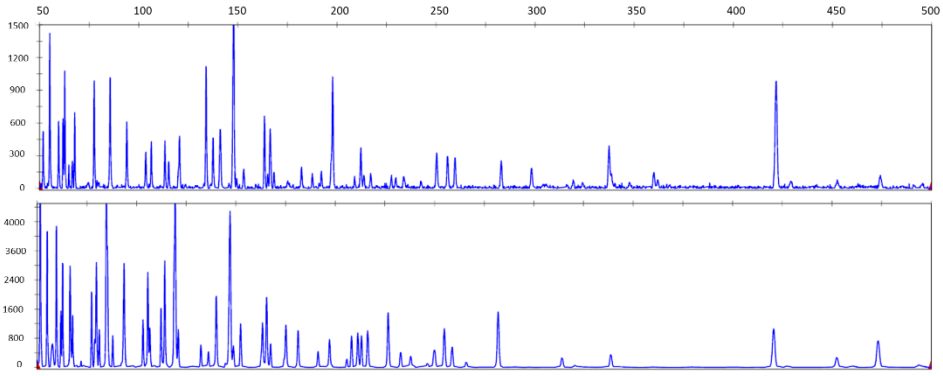
K


50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

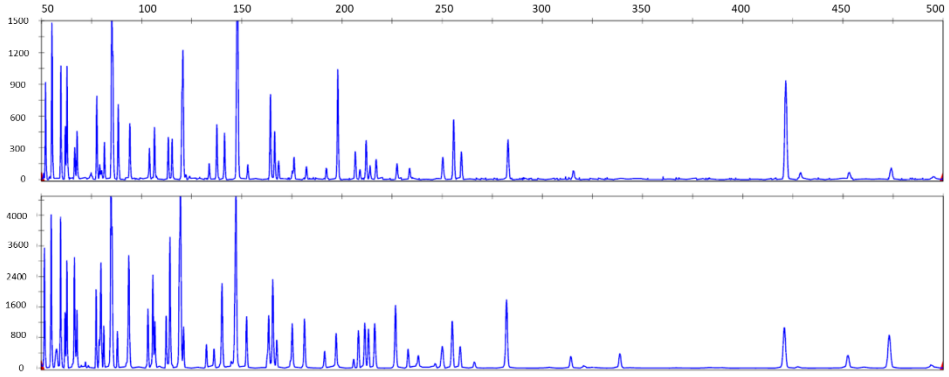


Supplementary Figure 2 Electropherograms of *L. plantarum* strains AFLP profiles obtained with the AFLP microbial Kit (upward) and with the advanced AFLP protocol (below). H) AFLP profiles of strain POM31; I) AFLP profiles of strain POM43; J) AFLP profiles of strain POM8; K) AFLP profiles of strain POM1; L) AFLP profiles of strain C6; M) AFLP profiles of strain POM40; N) AFLP profiles of strain POM38.

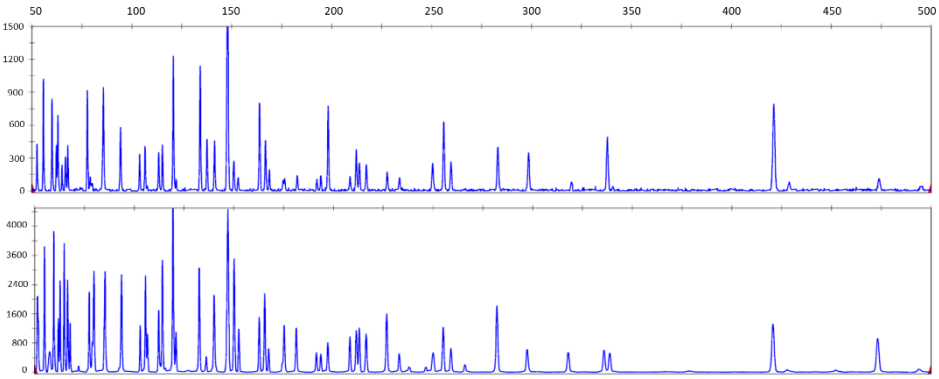
O


50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

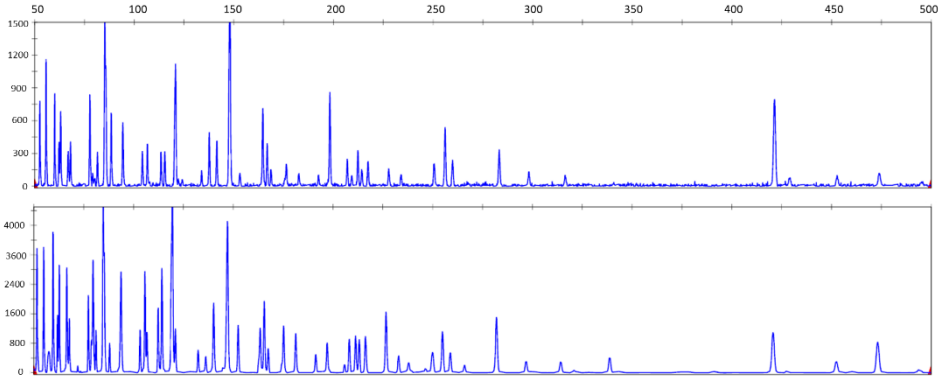
P

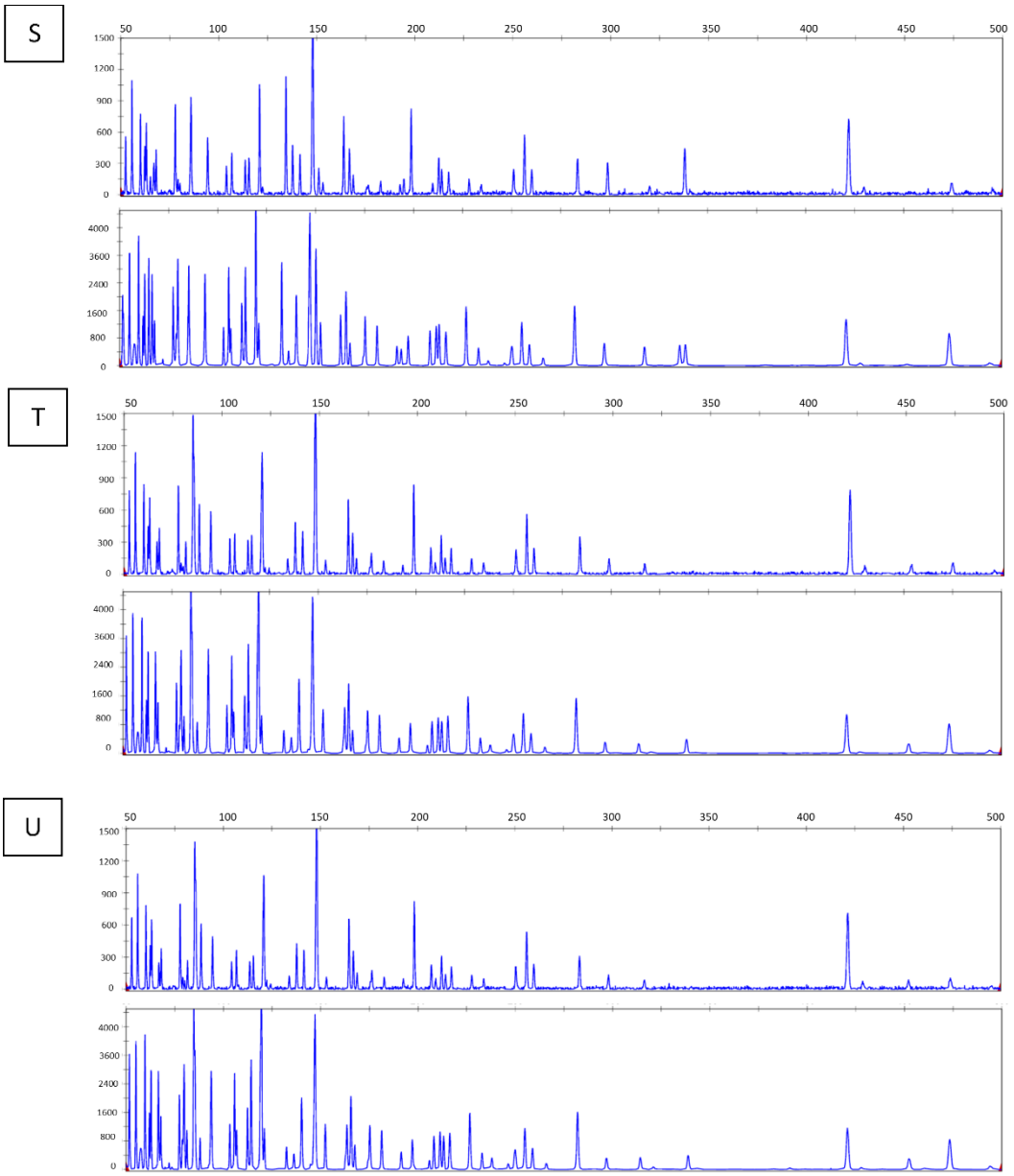

50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

Q


50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0

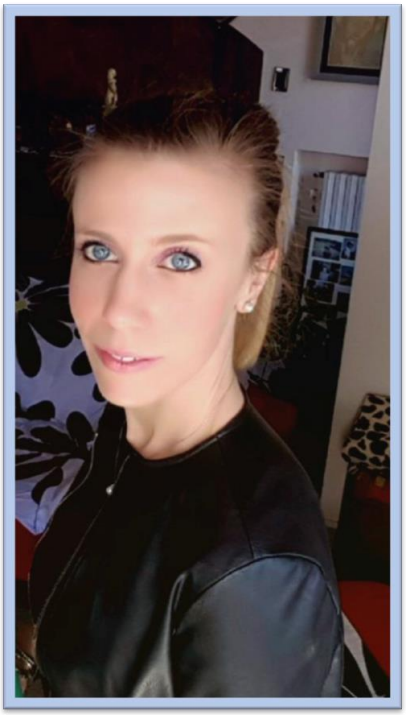
R


50 100 150 200 250 300 350 400 450 500
1500 1200 900 600 300 0
4000 3600 2400 1600 800 0



Supplementary Figure 3 Electropherograms of *L. monocytogenes* strains AFLP profiles obtained with the AFLP microbial Kit (upward) and with the advanced AFLP protocol (below). O) AFLP profiles of strain Lm9; P) AFLP profiles of strain Lm6; Q) AFLP profiles of strain Lm34; R) AFLP profiles of strain Lm40; S) AFLP profiles of strain Lm35; T) AFLP profiles of strain Lm41; U) AFLP profiles of strain Lm44.

About the Author



Gaia Bertani was born on 29th October 1990 in Parma, a province of Emilia Romagna, in Italy. In 2009, was accepted to the degree programme in Biology at the University of Parma. After completion of the bachelor, she got the master's degree in Molecular Biology (University of Parma, Italy) in July 2015, with the project titled "The yeast as a model for the study of neurodegenerative diseases". Subsequently, in November 2015, Gaia passed the selection for the Doctoral School in Food Sciences at the University of Parma, under the supervision of Prof. Erasmo Neviani and Prof. Monica Gatti. During the three years of her PhD she dealt with various aspects concerning the evolution of lactic acid bacteria used as starter in different DOP Italian cheeses manufacturing.

During, this experience, she developed a personal interest towards molecular tools for microbial dynamics investigation, bioinformatics and metagenomics. The results achieved in the three years of the PhD are described in this thesis.

Scientific activity

Original papers

Bertani, G., Bassi, D., Gatti, M., Cocconcelli, P.S., Neviani, E. (2018). Draft Genome Sequence of *Lactobacillus helveticus* Strain Lh 12 Isolated from Natural Whey Starter. Genome Announcements, 6, e00139-18.

Bertani, G., Savo Sardaro, M.L., Neviani, E., Lazzi, C. (2018). Advanced AFLP protocol for microbial diversity fingerprinting. *Journal of Applied Genetics (Submitted)*

Ricci, A.[$], Bertani, G.[$], Calani, L., Bernini, V., Lazzi, C. and Neviani, E. (2018). Pecorino Toscano cheese: genetic diversity and histamine production of *Streptococcus thermophilus* strains over time. *Food Research International (Submitted)*

Bertani, G.[$], Ricci, A.[$], Calani, L., Lazzi, C., Bernini, V. and Neviani, E. (2018). Prevalence of tyramine producing strain in *Streptococcus thermophilus* from Pecorino Toscano cheese. *Food Control (Submitted)*

[$] Both authors contributed equally to this manuscript

Communications

Evolutionary study over space and time of lactic acid bacteria used as starter in cheeses manufacturing. *23rd Workshop on the Developments in the Italian PhD Research on Food Science Technology and Biotechnology* (University of Sassari, Oristano, Italy, September 19th -21st, 2018).

Poster presentation

Gaia Bertani. *Lactobacillus helveticus*: evolutionary study of technological interest traits. *21st Workshop on the Developments in the Italian PhD Research on Food Science Technology and Biotechnology* (University of Naples Federico II, Portici, September 14th-16th, 2016).

Gaia Bertani. The evolutionary study of *Lactobacillus helveticus*. *22nd Workshop on the Developments in the Italian PhD Research on Food Science Technology and Biotechnology* (University of Bolzano, Bolzano, September 20th-22nd, 2017).

Gaia Bertani, Daniela Bassi, Claudia Cortimiglia, Monica Gatti, Pier Sandro Cocconcelli and Erasmo Neviani. Natural whey starter: a *reservoir* of *Lactobacillus helveticus* strains and genetic traits of technological interest. *4th international conference of Microbial Diversity 2017* (Bari, Italy, October 24th-26th, 2017).

Awards and distinctions

FEMS Grant for the participation of '*the 4th International Conference on Microbial Diversity 2017'* (Bari, October 24th-26th, 2017).

SIMTREA Grant for the participation of '*Summer School on Computational Analysis from Genomic Diversity to Ecosystem Structure'* (Firenze, September 3rd-7th, 2018).

Participation to PhD School and workshop

21st Workshop on the Developments in the Italian PhD Research on Food Science Technology and Biotechnology (University of Naples Federico II, Portici, Italy, September 14th -16th, 2016).

22nd Workshop on the Developments in the Italian PhD Research on Food Science Technology and Biotechnology (University of Bolzano, Bolzano, Italy, September 20th -22nd, 2017).

23rd Workshop on the Developments in the Italian PhD Research on Food Science Technology and Biotechnology (University of Sassari, Oristano, Italy, September 19th -21st, 2018).

Summer School on Computational Analysis from Genomic Diversity to Ecosystem Structure (Firenze, Italy, September 3rd -7th, 2018).