

Next-generation sequencing as an approach to dairy starter selection

Philip Kelleher¹ · James Murphy¹ ·
Jennifer Mahony¹ · Douwe van Sinderen^{1,2}

Received: 26 January 2015 / Revised: 25 March 2015 / Accepted: 2 April 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Lactococcal and streptococcal starter strains are crucial ingredients to manufacture fermented dairy products. As commercial starter culture suppliers and dairy producers attempt to overcome issues of phage sensitivity and develop new product ranges, there is an ever increasing need to improve technologies for the rational selection of novel starter culture blends. Whole genome sequencing, spurred on by recent advances in next-generation sequencing platforms, is a promising approach to facilitate rapid identification and selection of such strains based on gene-trait matching. This review provides a comprehensive overview of the available methodologies to analyse the technological potential of candidate starter strains and highlights recent advances in the area of dairy starter genomics.

Keywords LAB · *Lactococcus* · *Streptococcus* · Dairy fermentation · Phage · Genomics

1 Introduction

The first origins of fermented dairy products can be traced back over 7000 years to the pottery sieves used by some of the first farmers in Europe (Curry 2013). While the dairy industry has evolved since then, the same two basic ingredients remain its cornerstone: milk and lactic acid bacteria (LAB). LAB produce lactic acid from the degradation of hexose sugars and have a long history of safe use in food fermentations, and consequently, many LAB have been granted the so-called generally regarded as

This paper is part of the Special issue dedicated to the 9th International Cheese Symposium held in Cork, Ireland and organized by Teagasc in collaboration with University College Cork and INRA, 12th & 13th November 2014.

✉ Douwe van Sinderen
d.vansinderen@ucc.ie

¹ School of Microbiology, University College Cork, Cork, Ireland

² Alimentary Pharmabiotic Centre, Biosciences Institute, University College Cork, Cork, Ireland

safe (“GRAS”) status (Wegmann et al. 2007). LAB encompass an array of bacterial genera including *Abiotrophia*, *Aerococcus*, *Carnobacterium*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Oenococcus*, *Pediococcus*, *Streptococcus*, *Tetragenococcus*, *Vagococcus* and *Weissella* (Makarova et al. 2006). The most frequently employed starter cultures in dairy fermentations are strains of *Lactococcus lactis* and *Streptococcus thermophilus*, while strains of *Lactobacillus* spp. are primarily employed as adjunct cultures for flavour development (Leroy and De Vuyst 2004).

The earliest dairy fermentations occurred through the activity of the autochthonous organisms present in milk, thus causing spontaneous fermentation. The success of such natural fermentations is variable and depends on the particular natural blend of microorganisms present in each batch. The first attempt to control these fermentations was borne out by the “back-slopping” technique which involves inoculating milk with a small amount of a successful fermentate thus seeding the milk with favourable bacteria. This practice is still commonly applied to produce certain artisanal cheese products, and in such cases, the starter culture mixes are termed undefined starters. In industrial dairy fermentations, starter cultures may be defined or undefined. Defined starter cultures are those which consist of a specific number of known strains of bacteria which had been isolated from undefined starter cultures, while undefined are those of which the constituent strains are unknown (Smid et al. 2014). In Italian artisanal cheeses where back-slopping is employed, starter cultures are said to be undefined starters, while most cheddar-type cheeses are produced with a defined set that consist of two to four strains. The application of defined starter culture blends and rotations aims to ensure consistency of the fermentation process and particularly the quality of the final product in terms of flavour, appearance, aroma and safety.

This review will focus on *L. lactis* and *S. thermophilus* as the dominant starter cultures employed by the dairy industry (Beresford et al. 2001). In particular, it will focus on the methods used to characterise these strains and on the role of next-generation sequencing (NGS) technologies in facilitating rational starter culture selections and rotations.

L. lactis is a Gram-positive, catalase-negative, non-motile and coccoid bacterium (Schleifer and Kilpper-Bälz 1987). Genetically, a typical *L. lactis* chromosome ranges in size from ~2.2 to 2.5 Mb and is often accompanied by a rich plasmid complement (Ainsworth et al. 2014c). *L. lactis* species can be further defined as subsp. *cremoris*, subsp. *lactis* or subsp. *lactis* biovar. *diacetyllactis*, the latter capable of metabolising citrate. Both citrate metabolism and lactose utilisation are plasmid-encoded traits in *L. lactis*. Notably in dairy isolates of *L. lactis*, genome decay and redundancy have been widely reported (Makarova et al. 2006; Goh et al. 2011; Ainsworth et al. 2013), along with the presence of prophages and a host of transposable elements (Chopin et al. 2001).

S. thermophilus strains are Gram-positive, catalase-negative and non-motile cocci (Schleifer and Kilpper-Bälz 1987), which contain chromosomes that are comparably smaller than their *L. lactis* counterparts (1.8 Mb compared to 2.5 Mb), and normally, about 59% of strains carry one or two small plasmids (<10 kb) (Turgeon and Moineau 2001).

2 Traditional methods for screening/selection of starter strains

There is a wide variety of microbiological techniques that have been used in the dairy industry to screen and define starter cultures. Traditionally, the majority of these

methods have, unsurprisingly, placed significant emphasis on the technological phenotypes of the cultures and include growth performance or activity testing, phage robustness, flavour testing and matrix formation analysis.

2.1 Strain differentiation

For dairy lactococcal strains, an early characterisation step is to assign a subspecies, *lactis* or *cremoris*. The major phenotypic differentiator between *L. lactis* ssp. *lactis* and *cremoris* is their adaptive stress response. There are a number of protocols to differentiate between the two sub-species based on temperature, salt tolerance, pH tolerance, maltose fermentation and arginine deamination. One such protocol, based on arginine deamination, is the arginine broth assay described by Harrigan (1998) in which the sub-species identity is determined based on a pH-dependent colour change. *L. lactis* ssp. *lactis* strains are typically capable of metabolising arginine, a process that causes the release of ammonia and concomitant increase in pH, thereby producing red coloured colonies on arginine broth plates (due to a pH indicator). In contrast, *L. lactis* ssp. *cremoris* strains, which normally do not metabolise arginine, produce yellow coloured colonies on this medium because, compared to *L. lactis* ssp. *lactis* strains, their growth results in a lower pH (Villani et al. 2001; Murphy et al. 2013). A number of studies have focused on the response of strains of either sub-species to salt, temperature and pH, where *L. lactis* ssp. *lactis* strains can typically tolerate 4% salt, pH 9.2 and temperatures of up to 40 °C, while *L. lactis* ssp. *cremoris* growth is typically inhibited by these challenging conditions (Schleifer et al. 1985).

2.2 Performance testing

An important aspect of selecting strains for use as starter cultures is performance testing, in which the growth rate, acid production and responses to temperature and salt are assessed. These criteria are an early indicator of how the strains will behave in a fermentation process. The “Pearce activity test” is commonly used in the dairy industry as an indicator for growth and temperature-induced autolysis of starter strains (Feirtag and McKay 1987). The test replicates the temperature cycles of the relevant cheese production process, while monitoring cell counts and pH. The change in pH indicates the acidification activity of the strain, while the point of temperature-induced autolysis is determined via a decrease in cell count after a temperature shift. This test can be employed to assess the level of autolysis, as was done for two dairy starter strains, *L. lactis* subsp. *cremoris* HP and *L. lactis* subsp. *cremoris* AM2, and to determine the resulting proteolytic enzymes released during such autolysis (Wilkinson et al. 1994). Furthermore, the intracellular enzyme lactate dehydrogenase (LDH) can be used as an indicator of autolysis in dairy starter strains (Horvath and Barrangou 2010; Karginov and Hannon 2010), using a method proposed by Wittenberger and Angelo (1970), where LDH is measured by a decrease in absorbance at 340 nm as a result of pyruvate-dependent oxidation of NADH.

2.3 Flavour capabilities

The contribution of lactococcal and streptococcal starter strains to cheese flavour development is predominantly through the metabolism of lactose, lactate, citrate, lipids,

proteins and free amino acids (McSweeney 2004). A number of approaches have been employed to assess the contribution of particular starter strains to the relevant pathways, ranging from simple culture techniques to biochemical profiling of key enzymes, and quantitative analysis of sample products using various experimental approaches and/or sensory panels.

Proteolysis is the most complex and possibly the most important of these processes in terms of primary flavour development in cheese. Additionally, it is responsible for the liberation of peptides and (then) amino acids, thereby supplying substrates for various secondary pathways of amino acid catabolism (McSweeney 2004). To assess the proteolytic capabilities of starter cultures, a frequently used simple, direct test is to determine the degree of hydrolysis (DH) of peptide bonds, which provides a measure for amino acid availability and corresponding catabolism in milk (Mosher et al. 2013). Three commonly used methods for the determination of the degree of hydrolysis involve the application of trinitrobenzenesulphonic acid (TNBS) and o-phthalaldehyde (OPA), which both react with the amino groups released by the hydrolysis of peptides, and pH stat, a titration method where the protons released during peptide bond hydrolysis are used to quantify the DH (Quail et al. 2012). These methods give an overall evaluation of the proteolytic capability of the strains. An alternative method to determine the activity of specific intracellular peptidases was developed by Kato et al. (1978). The method utilises substrates labelled with 7-amino-4-methyl coumarin (AMC), which when cleaved fluoresces at excitation and emission wavelengths of 370 and 440 nm, respectively. The intensity of the fluorescence is used to determine the peptidase concentration. This method has been employed in a number of studies involving lactococci and lactic streptococci with particular emphasis on X-propyl and post-proline dipeptidyl aminopeptidase activity (Mills et al. 2010; Millen et al. 2012; Schloss et al. 2015).

LAB produce aroma compounds through amino acid catabolism, which contribute to cheese flavour, and the activity of glutamate dehydrogenase (GDH) in these strains is closely related to their ability to produce such compounds (Tanous et al. 2002). GDH produces α -ketoglutarate, a compound required for amino acid transaminations by LAB, which is an important step for the synthesis of various amino acids (McSweeney 2004). GDH activity in dairy LAB is usually assessed using a colourimetric assay which measures the glutamate-dependent reduction of NAD^+ and NADP^+ in a coupled reaction with diaphorase (Kieronczyk et al. 2003). Amino acid transferases catalyse the transamination of amino acids to α -ketoacids using α -ketoglutarate as an α -ketoacid acceptor (Tanous et al. 2002). These analyses applied in-parallel can provide a detailed biochemical profile of the flavour capabilities of particular strains, though they are limited in providing information that explain the (genetic and/or biochemical) reasons for strain-specific flavour properties.

2.4 Matrix formation

Exopolysaccharides (EPS) produced by LAB are polysaccharides that are deposited outside the cell wall. EPS may remain attached to the cell as capsular EPS or secreted in an unattached form (Hassan et al. 2003). Dairy strains are generally described as “ropy” or “non-ropy”, a term which describes the threads drawn with a needle from the surface of the colonies or fermented liquid (Hassan 2008). The EPS produced by some dairy

LAB can impact on the protein matrix of fermented dairy products by affecting the casein gel structure and acting as a filler (Hassan et al. 1995). A common assay for the differentiation of ropy and non-ropy colonies utilises ruthenium red stain in milk agar plates. Ruthenium red stains the cell wall, thereby producing red colonies for non-ropy, non-EPS producing cells, yet is less well capable of staining cell walls of ropy, EPS producers, which thus remain predominantly white (Borucki et al. 2003).

2.5 Phage robustness

While the technological attributes of dairy starter cultures are essential to achieve the desired flavours and characteristics in the final product, the phage robustness of these strains is also an important consideration. Since the discovery of lactococcal (bacterio)phages (i.e. viruses that infect bacterial cells) by Whitehead and Cox in 1935, phage infection has been recognized as the main cause of fermentation problems within the dairy industry with concomitant economic problems (Whitehead and Cox 1935). Selecting the correct starter culture traditionally involved assessing their susceptibility to phage infection, but with the advent of modern sequencing technologies, strain suppliers can now also screen strains to determine the presence of prophages as well as the arsenal of plasmid- and/or chromosomally encoded phage resistance mechanisms (Allison and Klaenhammer 1998; Ainsworth et al. 2014a, b).

The phage life cycle may take one of two possible routes, i.e. the lytic or the temperate/lysogenic life cycle, depending on the phage and the environmental circumstances. Ultimately, all phage species must enter the lytic cycle to harness the host machinery, replicate and release progeny phages by lysing the host cell. However, conditions may not favour lytic replication, and as a result, some phages may engage in a lysogenic life cycle by incorporating their genomes within that of the host, allowing phage replication *in situ* with that of the host's genome. This process allows the phage to replicate 'silently' each time the bacterium undergoes cell division by binary fission. Under certain conditions, the lysogenic phage will excise from the host genome and enter the lytic cycle. When selecting appropriate starter cultures for the production of various dairy products, it is important to determine if a particular strain harbours any prophages as these pose the risk of becoming active during the fermentation process leading partial or complete culture lysis. The presence of prophages can be determined by phage induction assays whereby the bacterium is placed under particular stressful conditions (chemical treatment or exposure to UV-light) in order to stimulate excision of the integrated phage which will subsequently turn on its lytic life cycle and lyse the host cell (Chopin et al. 2001).

2.6 Bacteriocin production

Bacteriocins are a diverse group of ribosomally synthesized peptides, produced by some bacteria and archaea, which convey a bactericidal or bacteriostatic effect on other bacteria when secreted (Dobson et al. 2012). Bacteriocin production is a double-edged sword, and consideration should be given in selecting starter cultures, as producing strains may inhibit other desirable strains in mixed starter cultures or adjunct cultures added later in the fermentation process. However, they also offer the benefit of inhibiting the growth of spoilage bacteria in food products. Traditionally, a range of

culture-based methods have been used in screening for bacteriocin producers, most commonly based on the principles of diffusion in agar plates and cell-free supernatants (Kékessy and Piguet 1970; Barefoot and Klaenhammer 1983; Yang et al. 1992).

3 Modern methods for screening/selection of starter strains

Culture-based techniques are an excellent foundation for the study of LAB strains but are both labour- and time-intensive. While these methods have limitations, they can still be applied on selected strains, yet are impractical for screening large strain collections. Current approaches to identify, classify and characterise industrial starter cultures for application in the dairy industry is increasingly reliant on molecular techniques because phenotypic traits alone are not sufficiently discriminative since some strains of *L. lactis* ssp. *cremoris* display the characteristics of *L. lactis* ssp. *lactis* (Urbach et al. 1997; Tailliez et al. 1998). Molecular tools possess several advantages over classical approaches including a shorter assay period, higher throughput and a greater ability to discriminate at the species level. Up until recently, the gold standard of molecular typing of bacterial isolates was pulsed field gel electrophoresis (PFGE). Other methodologies rely on a polymerase chain reaction (PCR) which employs oligonucleotide primers to amplify DNA fragments. Such techniques include PCR-restriction fragment length polymorphism (PCR-RFLP), repetitive sequence-based PCR (REP-PCR), PCR-ribotyping, amplified fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD) PCR and multilocus sequence typing (MLST). As such methods have been discussed extensively, we will only provide a brief overview of the advantages and/or disadvantages of each technique in the context of selecting industrially important dairy starter cultures.

3.1 Restriction endonuclease-based methods

The use of PFGE allows for the separation of large DNA fragments, produced by one or more suitable restriction endonucleases (REases), employing alternating electric currents resulting in the generation of a strain-specific genomic fingerprint (Cantor et al. 1988; Mathew et al. 1988; Tanskanen et al. 1990). PFGE has been used extensively to characterise members of both *L. lactis* and *S. thermophilus* (Boutrou et al. 1995; Ward et al. 2004; Kelly et al. 2010; Zinno et al. 2010). PFGE is highly discriminatory in comparison to other methods, but the procedure is time-consuming, technically challenging, and the lack of fingerprint databases limits its widespread use in the dairy industry. Similar to PFGE, AFLP involves the digestion of chromosomal DNA with restriction endonucleases. Following digestion, oligonucleotide adapters containing a core sequence and a sequence homologous to the sticky ends produced by the endonuclease are added to the DNA fragments creating a target site for the AFLP-specific primers. These primers are designed to match the core sequence and the restriction enzyme-specific sequence (Vos et al. 1995; Vuylsteke et al. 2007). Two to three additional non-specific bases may be added to the 3'-end of the primer to increase the selectivity of the protocol. The AFLP method can be automated by the use of fluorophore-labelled primers, coupled to fragment separation by denaturing gel electrophoresis and subsequent detection of AFLP fragments using an automated sequencer (Gancheva et al. 1999). While initially labour-intensive, the development of a particular

protocol can be used to characterise hundreds of bacterial strains using only small amounts of template DNA as well as allowing for faster processing of samples and a high-throughput analysis of strains without any prior knowledge of the genome (Paun and Schönswetter 2012). However, such approaches require an automated DNA sequencer which can be expensive to purchase, run and maintain. To date, there has been no reported use of AFLP in characterising *L. lactis* strains used as dairy starter cultures. Analysis of dairy *S. thermophilus* strains using AFLP indicated that this approach provides a higher degree of differentiation at both the species and strain level compared to other methods such as RAPD PCR (Lazzi et al. 2009).

3.2 PCR-based methods

RAPD PCR is a fingerprinting technique which uses ten nucleotide (10-mer) primers to amplify random segments of the bacterial genome without the requirement of having previous sequence knowledge of the genome (Williams et al. 1990). Diversification or discrimination at the intraspecies and interspecies level as well as the identification of mutations within a genome is based on whether or not the primer will amplify a segment of DNA resulting in the generation of unique DNA banding patterns. The use of RAPD PCR has been extensively applied to the characterisation of members of the LAB, as well as lactococcal and streptococcal phages (Rossetti and Giraffa 2005; Rodríguez et al. 2008; Chouayekh et al. 2009). RAPD PCR has been shown to differentiate the lactococcal sub-species and also members of the same subspecies, particularly where members of a particular subspecies display different phenotypic characteristics (Tailliez et al. 1998; Samarzija et al. 2002). PCR-RFLP, which utilises a combination of PCR and a REase-based approach, is a genotyping technique that relies on the presence of single nucleotide polymorphisms (SNPs) or multi-nucleotide polymorphisms (MNPs) within a specific region of targeted homologous DNA (Saiki et al. 1985). The presence of SNPs or MNPs within the amplified region of interest can alter the recognition site of the REase resulting in different banding patterns, allowing the discrimination of different (sub)species within a bacterial genus/species, or members of the same species. PCR-RFLP has been employed to discriminate between *L. lactis* ssp. *lactis* and *L. lactis* ssp. *cremoris* by targeting different genes such as those specifying N-acetylmuramidase (*acmA*), the 16S rRNA and glutamate decarboxylase (*gadB*) (Buist et al. 1995; Ward et al. 1998; Nomura et al. 2002). Recently, two additional gene targets, the genes encoding a serine protease (*htrA*) and a non-proteolytic protein peptidase family M16 member (*yueF*) were found to be more suitable for differentiating between the two subspecies *cremoris* and *lactis* when used in combination with the REases TaqI and AluI, respectively (Khemariya et al. 2013). Both RAPD and PCR-RFLP target specific DNA regions, whereas rep-PCR employs sets of primers that target non-coding repetitive short polynucleotide sequence tracts that can be found scattered across the bacterial genome (Stern et al. 1984; Versalovic et al. 1991). Amplification of these repeating sequences results in the generation of fragments of various lengths depending on the length of the DNA located between these repetitive elements generating strain-specific patterns when separated by gel electrophoresis. Several short repetitive elements can be used as targets for amplification including the repetitive extragenic palindromic (REP) elements; enterobacterial repetitive intergenic consensus (ERIC) sequences (127-bp imperfect palindromes) and BOX

elements (modular DNA segments encompassing three subunits, A-C, which can exist in various combinations) (Koeuth et al. 1995; Tobes and Ramos 2005; Wilson and Sharp 2006). While it has been demonstrated that rep-PCR is comparable to RFLP and RAPD, several limitations such as poor resolution and band separation have been observed for rep-PCR (Olive and Bean 1999). To overcome these limitations, fluorophore-labelled oligonucleotides can be used where the amplified sequences are separated using a DNA sequencer similar to the AFLP approach (Versalovic et al. 1995), and this approach has been used to characterise *S. thermophilus* strains and establish links between *S. thermophilus* strains and their geographical origin (Brusetti et al. 2008). There have been mixed reports on the differentiating ability of rep-PCR. While Urbach et al. (1998) applied rep-PCR for the characterisation of *L. lactis* strains and reported it as a useful and reproducible technique at differentiating strains, subsequent studies have shown that rep-PCR cannot distinguish between ssp. *lactis* and ssp. *cremoris* (Prodělalová et al. 2005).

3.3 Multilocus sequencing typing (MLST)

MLST is a technique that targets the internal region (400–500 bp) of several housekeeping genes and is now by many considered to represent the gold standard for species differentiation (Maiden et al. 1998). MLST has been used to characterise both *L. lactis* and *S. thermophilus* strains, including wild-type and industrial starters. MLST has been shown to effectively discriminate collections of lactococcal strains into subgroups or clusters, as well as differentiating the sub-species based on housekeeping genes such as peptidase N and X-prolyl dipeptidyl aminopeptidase (Fernández et al. 2011). MLST can be used to separate lineages among the sub-species of *L. lactis*, although some lineages were found to include representatives of both subsp. *lactis* and *cremoris*, suggesting that MLST is not completely discriminatory at the sub-species level (Rademaker et al. 2007). In a recent study, the MLST scheme for *L. lactis* was revised based on additional genes that were not necessarily housekeeping genes, particularly genes that act as indicators of species divergence (Passerini et al. 2010). MLST has also been applied to differentiate isolates of *S. thermophilus* using primers that target eight housekeeping genes, previously applied for MLST purposes on strains of *Streptococcus salivarius* (Delorme et al. 2007, 2010).

3.4 Cell wall polysaccharide (CWPS) operon typing

The (Gram-positive) lactic acid bacterial cell envelope represents a complex structure composed of a thick peptidoglycan layer, teichoic acids, cell wall polysaccharides (CWPS) and various surface carbohydrates (Chapot-Chartier 2014). Recently, it has been demonstrated by a multiplex PCR approach that, based on the genetic composition of their CWPS-encoding gene cluster, (most) lactococcal dairy strains can be assigned to one of three types (types A, B and C) (Mahony et al. 2013). Strains possessing the CWPS C type can be further divided into five subtypes (designated C₁ through to C₅), which can also be identified by multiplex PCR (Ainsworth et al. 2014b). These PCR-based approaches are important and rapid in classifying the CWPS biosynthesis cluster, in particular from phage sensitivity prediction and strain blend derivation perspectives.

4 Next-generation sequencing (NGS) approaches

Sanger sequencing was first described in the 1970s by Frederick Sanger and colleagues (Sanger et al. 1977) and became the dominant method of sequencing DNA for the remainder of the 20th century. It was used to sequence the complete human genome (Lander et al. 2001; Venter et al. 2001) and is a highly accurate method of sequencing. However, Sanger sequencing, referred to as a “first generation” sequencing method, is expensive and impractical for large sequencing projects, and in recent times, this method has partly been replaced by “next-generation” sequencing (NGS) methods (Metzker 2010).

NGS methods are high-throughput DNA sequencing technologies permitting the sequencing of millions of DNA strands in parallel generating large volumes of sequence data in a relatively short period of time (Pettersson et al. 2009). There are currently a number of methods in use (Table 1), but for the purposes of this review, we will briefly compare four of the most commercially viable methods, namely Roche 454-pyrosequencing, Illumina-Solexa, Life Sciences Ion-Torrent and Pacific Biosciences Single-molecule real-time sequencing (SMRT). For more in-depth information on each of these technologies, see Loman et al. (2012a, b).

4.1 Comparison of NGS approaches

There are a number of next-generation techniques available with associated advantages and disadvantages to each technique depending on the desired application (Table 1). For the study of lactococcal or streptococcal starter cultures, any of the aforementioned techniques may be applied to obtain finished genome sequences due to the small genome size of these species: strains of *L. lactis* typically possess a ~2.5 Mb chromosome, whereas *S. thermophilus* strains have a smaller genome of around 1.8 Mb.

454-Pyrosequencing is a next-generation, high-throughput sequencing methodology based on the sequence by synthesis approach and is useful due to its longer read length compared to read lengths generated by the current Illumina or Ion-torrent platforms: 700 bp compared to 300 and 400 bp, respectively. While 454-pyrosequencing has been used extensively over the last 10 years, it was announced in 2013 that Roche will phase out this sequencing platform by mid-2016. The reasons for this discontinuation of 454-pyrosequencing include the advent of lower cost, high(er) throughput sequencing technologies, along with increasing read lengths of the alternative NGS technologies (Chaisson et al. 2009). Errors in homopolymer sequence tracts have also been reported with the 454-pyrosequencing method (Gilles et al. 2011) and ion torrent technology (Loman et al. 2012b).

The Ion-Torrent PGM represents a low-cost and rapid sequencing methodology generating around 80 million sequence reads in a single run of approximately 90 min. The Illumina approach has been one of the most widely used sequencing approaches in recent years and can generate a large volume of sequencing data (Metzker 2010), although the average read length is relatively low, in particular when compared to the newer PacBio SMRT platform. Current Illumina sequencing-by-synthesis (SBS) instruments are capable of generating over 1 terabase of data in a single run and are capable of sequencing bacterial genomes in a matter of hours. These properties,

Table 1 Comparison of next-generation sequencing technologies adapted from Metzker (2010)

Platform	Library preparation	Chemistry	Consensus accuracy	Average read length (bp)	Reads per run	Run time	Pros	Cons	Application	References
Roche 454 GS FLX Titanium+	Fragment, Mate Pair/ emPCR	Pyrosequencing	99.997%	~700, maximum 1000	~1,000,000	23 h	Long read length	High reagent cost, homopolymer repeat errors	De novo assemblies, metagenomics	http://454.com/products/gs-flx-system/index.asp
Illumina- Solexa MiSeq	Fragment, Mate Pair/ Solid phase	Reversible terminator	98%	2 × 300	~2.5,000,000	~55 h 300 bp reads	Widely used platform	Short read length	Small genomes, 10x amplicon, improving coverage	http://systems.illumina.com/systems/miseq.html
Life Technologies SOLiD 5500 Series	Fragment, Mate Pair/ emPCR	Sequencing by ligation (Cleavable probe)	99.99%	Mate-paired 2 × 60, Paired-end 75 × 35 Fragment 75	1.2–1.4 billion	1–2 weeks	Low-cost	Slow, issues with palindromic sequences reported	Whole genome re-sequencing, variant analysis	http://www.lifetechnologies.com/ (Huang et al. 2012)
Life Technologies Ion Torrent	Fragment, Mate Pair/ emPCR	Sequential ion detection	98%	35–400	80,000,000	90 min	Fast and inexpensive	Reported homopolymer errors	Small genomes, gene expression, ChIP-SEQ	http://www.lifetechnologies.com/
Pacific Biosciences SMRT RS II	Fragment only/ Single molecule	Real-time	99.999%	N50 14,000, maximum >40,000	~50,000	30 min–4 h	Longest read length, detects base modifications, fast	Low single sequence accuracy 87%	De novo assemblies, Base modification detection, Transcriptionome sequencing	(Murray et al. 2012; Chin et al. 2013) http://www.pacificbiosciences.com/

combined with low sequencing costs, have in recent years made Illumina become the dominant sequencing technology.

The PacBio SMRT approach has the advantage of the longest read lengths of any sequencing technology currently in use (Chin et al. 2013), with Pacific Biosciences reporting N50 read lengths of >14,000 bp and maximum read lengths of >40,000 bp, which is extremely useful for covering repetitive regions of genomes, particularly so in lactococcal genomes where a large number of insertion sequence (IS) elements cause problems during sequence read assembly (Daveran-Mingot et al. 1998; Chopin et al. 2001; Kok et al. 2005). The SMRT sequencing approach also moves beyond traditional detection of the four DNA bases as it is the first high-throughput approach to directly detect DNA base modifications (Flusberg et al. 2010). This allows SMRT sequencing to differentiate between unmodified bases and those with m6A, m4C or m5C base modifications (Clark et al. 2012). One drawback of the PacBio SMRT platform which should be considered is the reports of higher error rates compared to other NGS platforms. Since launching the SMRT platform, Pacific Biosciences have addressed this issue by incorporating circular consensus sequencing (CCS), which has led to greatly reduced error rates (Hodkinson and Grice 2015) and achieving a higher consensus accuracy, currently reported at 99.999% by Pacific Biosciences. In comparison to early studies which reported a sequencing inaccuracy rate of ~13–18% (Quail et al. 2012; Mosher et al. 2013), more recent studies have reported a large reduction in these rates (Schloss et al. 2015).

As discussed, the PacBio SMRT platform currently possesses a number of unique advantages over other NGS methods. However, a noteworthy new single molecule sequencing method currently in development is nanopore sequencing which may challenge PacBio's dominance in this area. Nanopore sequencing is predicted to deliver long read lengths and base modification data, while the simple sample preparation and possibility of label-free DNA sequencing are expected to reduce sequencing costs dramatically (Clarke et al. 2009).

Genotypes of lactococcal and streptococcal strains derived from genomics (Table 2) can provide a myriad of information about industrially important traits. There is an impressive array of tools available for post-sequencing and comparative genome analyses, and readers requiring more information should refer to Edwards and Holt (2013). Here, we discuss some of the key genetic markers derived from genomic analysis which can be used for strain selection with particular emphasis on phage resistance and flavour development.

4.2 Practical implications of genomics-driven characterisation of dairy cultures

4.2.1 Metabolic capabilities

Dairy LAB encode a number of key metabolic pathways that are necessary for the production of cheese and other fermented dairy products. These include genes required for lactose utilisation, degradation of peptides and amino acids, citrate metabolism and lipolysis, all of which can be characterised at a genetic level given the availability of sequencing data.

Table 2 Sequencing methods used for finished quality lactococcal and streptococcal genomes available from the NCBI (correct January 2014)

Organism	GenBank accession	Sequencing technology	Year	Reference
<i>L. lactis</i>				
ssp. <i>lactis</i> IL1403	GCA_000006865.1	Sanger	2001	Bolotin et al. 2001
ssp. <i>cremoris</i> SK11	GCA_000014545.1	Sanger	2006	Makarova et al. 2006
ssp. <i>cremoris</i> MG1363	GCA_000009425.1	Sanger	2007	Wegmann et al. 2007
ssp. <i>lactis</i> KF147	GCA_000025045.1	Combined 454-pyrosequencing & Illumina	2009	Siezen et al. 2010
ssp. <i>cremoris</i> NZ9000	GCA_000143205.1	Illumina	2010	Linares et al. 2010
ssp. <i>cremoris</i> A76	GCA_000236475.1	Sanger	2011	Bolotin et al. 2012
ssp. <i>lactis</i> CV56	GCA_000192705.1	454-pyrosequencing	2011	Gao et al. 2011
ssp. <i>lactis</i> IO-1	GCA_000344575.1	Sanger	2012	Kato et al. 2012
ssp. <i>cremoris</i> UC509.9	GCA_000312685.1	Combined 454-pyrosequencing & Illumina	2012	Ainsworth et al. 2013
ssp. <i>cremoris</i> KW2	GCA_000468955.1	454-pyrosequencing	2013	Kelly et al. 2013
ssp. <i>lactis</i> KLDS 4.0325	GCA_000479375.1	Illumina	2013	Yang et al. 2013
ssp. <i>lactis</i> NCDO 2118	GCA_000478255.2	SOLiD, Ion PGM & Ion Torrent PGM	2014	Oliveira et al. 2014
ssp. <i>lactis</i> SO	GCA_000807375.1	Ion Torrent PGM	2014	McCulloch et al. 2014
ssp. <i>lactis</i> AI06	GCA_000761115.1	454-pyrosequencing	2014	Unpublished
<i>S. thermophilus</i>				
CNRZ1066	GCA_000011845.1	Sanger	2004	Bolotin et al. 2004
LMG 18311	GCA_000011825.1	Sanger	2004	Bolotin et al. 2004
LMD-9	GCA_000014485.1	Sanger	2006	Makarova et al. 2006
ND03	GCA_000182875.1	Combined 454-pyrosequencing & Illumina	2010	Sun et al. 2011
JIM 8232	GCA_000253395.1	Sanger & SOLiD	2011	Delorme et al. 2011
MN-ZLW-002	GCA_000262675.1	Combined 454-pyrosequencing & Illumina	2012	Kang et al. 2012
ASCC 1275	GCA_000698885.1	454-pyrosequencing	2014	Wu et al. 2014

4.2.2 Lactose and lactate metabolism

The gene products of the *lac* operon facilitate and govern lactose utilisation in LAB and provide dairy strains with the ability to rapidly ferment lactose required for growth in milk. In *L. lactis*, the plasmid-borne *lac* operon consists of the genes *lacABCDEFGHIX* and is regulated by a repressor, encoded by the adjacent *lacR* gene (van Rooijen and de Vos 1990; van Rooijen et al. 1992). Loss of the *lac* operon has been reported due to the

instability of the large extra-chromosomal element on which it is encoded (McKay et al. 1972; Ainsworth et al. 2014c), resulting in spontaneous mutants that are incapable of growth in milk. Interestingly, the plasmid-cured laboratory strain *L. lactis* MG1363, which does not harbour the *lac* operon, is capable of growth on lactose-supplemented media following prolonged adaptation due to the activity of a cellobiose-specific phosphotransferase system (PTS), which can act as an alternative lactose utilisation pathway (Solopova et al. 2012). Another example of an alternative lactose metabolic pathway is found in the slow lactose fermenter *L. lactis* NCDO2054 which metabolises lactose via the Leloir pathway (Bissett and Anderson 1974). This occurs as a result of *lacA*, which encodes a galactoside acetyltransferase, and *lacZ*, which encodes a β -galactosidase, being integrated into the *gal* (galactose) operon (Vaughan et al. 1998). Such data suggests that phenotypic growth on lactose may not be an absolutely reliable indicator for the presence of the *lac* operon within lactococcal strains. Further studies have suggested that certain PCR-based techniques may also be somewhat unreliable in indicating the true lactose genotype.

A recent study by Ferrario et al. (2012) reported on the screening for isolates of *L. garvieae* in the dairy environment using primers targeting the *lacG* gene. They found that *lacG* is variably present among *L. garvieae* isolates from the meat environment and is not limited to dairy isolates, demonstrating the need for complete genome sequences for the correct identification of dairy isolates.

Lactose metabolism by *S. thermophilus* differs from that observed in *Lactococcus* spp.: lactose is transported into the cell by the secondary transport system LacS in antiport with galactose (Delcour et al. 2000). The *lacS* gene is organised in an operon with the β -galactosidase-encoding gene, *lacZ*, similar to the system in *Lactobacillus delbrueckii* (Schroeder et al. 1991). The *lacZ* gene in *S. thermophilus* is well conserved and has been used for PCR-based identification purposes of *S. thermophilus* from fermented dairy products (Lick et al. 1996) and for the isolation of folate-producing *S. thermophilus* strains from Indian fermented milk products (Iyer et al. 2011).

4.2.3 Citrate metabolism

Citrate metabolism in dairy fermentations conducted by citrate-positive (Cit⁺) lactococci and *Leuconostoc* spp. is important as it leads to the production of a number of volatile flavour compounds (McSweeney and Sousa 2000). Citrate uptake and subsequent diacetyl production are governed by the plasmid-encoded *citQRP* operon in lactococcal species (Drider et al. 2004). It has been demonstrated that the *citP* gene is well conserved amongst LAB with approximately 98% amino acid identity making it a useful screening target for Cit⁺ starters (Drider et al. 2004). Lactococci capable of metabolising citrate are classified as *L. lactis* subsp. *lactis* biovar diacetyllactis (Kelly et al. 2010), a classification that has led to confusion since plasmid-encoded characteristics such as citrate and arginine metabolism can be transferred to subsp. *cremoris* strains leading to incorrect characterisation based on phenotype (Kelly et al. 2010). It is also noteworthy that recent studies have indicated potential adverse health effects associated with diacetyl production, which may lead to the future removal of diacetyl-producing LAB from starter cultures (Shibamoto 2014).

4.2.4 Proteolysis

Proteolysis and the degradation of casein from milk are one of the most important contributors to flavour development in cheese (McSweeney 2004). Lactococcal strains contribute to proteolysis through the hydrolysis of casein by peptidases and proteases and the catabolism of peptides and amino acids from casein breakdown (Steele et al. 2013). There are a number of genes which contribute to this function, such as various and mostly chromosomally specified peptidase-encoding genes (e.g. *pepC*, *pepN*, *pepX*, *pepP*, *pepA*, *pepF2*, *pepDA1*, *pepDA2*, *pepQ*, *pepT*, *pepM* and *pepO1*), the plasmid-encoded *opp* operon, which specifies an oligopeptide-uptake system, and the plasmid-borne gene that specifies the *L. lactis* cell wall-associated protease PrtP, required for the proteolytic phenotype (Yu et al. 1996). The majority of the genes mentioned above are monocistronic (e.g. *pepC*, *pepN* and *prtP*) or co-transcribed, such as *opp* and *pepO1*, while *pepF2*, *pepM* and *pepT* are transcribed with genes that are (apparently) unrelated to proteolysis (Guédon et al. 2001). There are also a number of uncharacterised proteins which contain peptidase-associated domains, many of which are strain-specific and their roles may become clearer as more genome sequences become available (Siezen et al. 2005).

As discussed, proteolysis contributes greatly to cheese flavour development; however, high levels of proteolysis can also cause bitterness in cheese (Broadbent et al. 2002). The *L. lactis* extracellular cell wall proteinase (lactocepin) has been shown to be directly involved in the bitter flavour defect in cheddar cheese varieties, specifically starters which produce lactocepin of the so-called group a, e or h (Broadbent et al. 2002). Broadbent et al. (2002) concluded that the bitterness defect in cheese could be altered through gene exchange or replacement in the starter culture. These findings highlight the benefits of subsp. *cremoris* strains in lactococcal starter cultures in comparison to subsp. *lactis*.

A recent study by Liu et al. (2010) indicates that our knowledge of the proteolytic system in LAB can be enhanced by systematic genome-wide studies of the regions encoding proteins involved in proteolysis. These authors indicated that comparative genomics can be used to distinguish various sub-groups within protein super-families involved in proteolysis where the generated information predicts the proteolytic ability of LAB strains. A major finding from this study was the confirmation of proteolytic diversity among ssp. *lactis* and ssp. *cremoris* strains and the provision of a genetic basis for this diversity, linked to distinct patterns in the presence or absence of genes encoding proteolytic functions (Liu et al. 2010).

4.2.5 Lipolysis

Lipolysis involves the breakdown of milk fats and hydrolysis of triglycerides into lipids and fatty acids, activities that are considered to be crucial for flavour development in cheese production, particularly in the production of cheddar varieties (McSweeney and Sousa 2000). In LAB, lipolytic enzymes involved in lipolysis are mainly esterases and lipases belonging to a class of enzymes called the carboxylic ester hydrolases (Verger 1997). Apparently, *estA* is the only esterase-encoding gene in *L. lactis*, being capable of hydrolysing short chain fatty acid esters (Nardi et al. 2002). However, this research area

of cheese flavour development remains considerably under-represented in lactococcal studies compared to those related to proteolysis (McSweeney and Sousa 2000).

Therefore, a genomics approach may be beneficial in broadening our scope of knowledge on lipolysis in lactococcal and streptococcal strains as demonstrated in other LAB.

4.3 Phage robustness

4.3.1 Prophages

As mentioned previously, it is possible to determine the presence of prophages by phage induction assays (Chopin et al. 2001). However, such approaches are time-consuming and require the assessment of large collections of strains. In addition, ‘true’ prophage induction can only be determined using additional methods such as confirmation of the presence of prophages by performing phage sensitivity assays upon identification of a sensitive host strain, PCR or flow cytometry (Sozhamannan et al. 2006). Whole genome sequencing can readily identify the presence of temperate phages within the host genome. Furthermore, the availability of programmes such as Phage_Finder as well as gene annotation tools aid in the determination of the presence of intact or cryptic prophage elements (Fouts 2006). The presence of prophages is more common in *L. lactis* than *S. thermophilus* with some of the former harbouring six prophages (Chopin et al. 2001; Wegmann et al. 2007). Lysogenic *S. thermophilus* strains have been identified and have been correlated with strain autolysis (Husson-Kao et al. 2000; Neve et al. 2003). While the presence of prophages in commercial strains has generally been considered an undesirable trait due to the risk of phage excision, some prophage elements have been found to encode systems which are beneficial for phage defence namely, superinfection exclusion systems (Sie) in lactococcal strains and Lipoprotein (Ltp) in streptococcal strains (Gasson and Davies 1980; McGrath et al. 2002; Mahony et al. 2008).

4.3.2 Restriction-modification (R-M) systems and abortive infection (Abi) systems

Genes encoding R-M systems are present on approximately 90% of currently available bacterial and archaeal genome sequences (Roberts et al. 2003). These systems can be plasmid or chromosomally encoded, and their general role is to recognize and target invading foreign DNA with restriction enzymes, while simultaneously protecting the host’s DNA by methyltransferase (MTase) activity. Four types of R-M systems (I, II, III and IV) are currently recognized and have been extensively reviewed previously (Roberts et al. 2003; Loenen et al. 2013; Pingoud et al. 2014; Rao et al. 2014). The presence of various R-Ms in industrial starter cultures plays an important role in phage defence, as invading phage DNA, if unmethylated (except in the case of type IV R-M systems), will be subject to endonuclease activity. Traditional and laborious approaches to the identification of R-Ms such as the use of crude cell extracts, extract fractionation and restriction endonuclease assays are time-consuming and often only suitable for type II R-Ms such as the lactococcal ScrFI, LlaBAI and LlaBI (Fitzgerald et al. 1982; Nyengaard et al. 1993; Mruk et al. 2003) and *S. thermophilus* Sth4551 (Guimont et al. 1993). These methods have been replaced with the emergence of more accessible

sequencing technologies allowing for the prediction of chromosomal- and plasmid-encoded R-Ms through similarity searches, e.g. the *L. lactis* systems LlaII, LldI and LlaI (Hill et al. 1989; Deng et al. 2000; O'driscoll et al. 2004) and the *S. thermophilus* R-M system Sth3681 (Burrus et al. 2001). In recent years, the emergence of SMRT sequencing technology (as discussed above) has revolutionised the identification of whole genome modification and the function of R-Ms. Combining whole genome sequencing and MTase motif analysis, the functions of one or more bacterially encoded R-Ms can be predicted which can then be confirmed using heterologous gene expression coupled with restriction endonuclease assays. This approach has been applied to both bacteria and bacteriophages alike (Murphy et al. 2014; O'Connell-Motherway et al. 2014), though it has not yet been applied to lactococcal or streptococcal strains.

Abortive infection (Abi) systems are host-encoded resistance mechanisms that disrupt critical stages in the lytic phage cycle such as phage transcription, translation, DNA replication or phage DNA packaging and have been extensively studied in *L. lactis* (Ainsworth et al. 2014a, c). Abi-mediated resistance typically culminates in the death of the infected host cell in order to limit the release of progeny particles, thus protecting the neighbouring bacterial population. Currently, 23 Abi systems (AbiA-AbiZ) are known for *L. lactis*, which, with the exception of AbiN and AbiV, are all plasmid-encoded (Prevots and Ritzenthaler 1998; Chopin et al. 2005; Ainsworth et al. 2014a, c). The presence of Abi systems was first identified due to the protective effect that certain lactococcal plasmids have against phage infection, by causing a decreased burst size and an altered phage plaque morphology as observed for plasmids pTR2030 and pIL105 (Sing and Klaenhammer 1986; Gautier and Chopin 1987). Subsequently, plasmids that conferred such resistance to infecting phages were digested with restriction endonucleases and the fragments cloned into suitable shuttle vectors. The various recombinant derivatives were then screened to determine if a particular fragment provided phage resistance as observed for AbiE and AbiF encoded on the lactococcal plasmid pNP40 (Garvey et al. 1995). Such experimental approaches are time-consuming, particularly if the approach was to be employed in order to assess a collection of dairy starters. As with R-Ms, whole genome sequencing can readily identify both chromosomally and plasmid-encoded Abi systems, in particular in the case of SMRT sequencing technology, which in many cases can produce completed, single contig bacterial genome and plasmid sequence information for use in comparative sequence analyses (Sistla and Rao 2004).

4.3.3 CRISPR/Cas systems

Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) genes form an acquired adaptive immunity system against foreign genetic elements in prokaryotes (Horvath et al. 2008; van der Oost et al. 2009; Deveau et al. 2010). CRISPR systems are composed of a series of conserved repeats which are separated by protospacers, variable sequences involved in target recognition, an A-T-rich leader region located at the 5' end of the CRISPR locus and Cas genes (Deveau et al. 2008). CRISPR systems play an important role in phage resistance in dairy starter strains (Mills et al. 2010), and furthermore, CRISPR systems can be used as a tool for the typing and comparative analyses of strains of *S. thermophilus* (Horvath et al. 2008). CRISPR typing of *S. thermophilus* performed by Horvath et al. (2008), based on a

combination of primers targeting conserved regions and Sanger sequencing resulted in the identification of CRISPR3 and demonstrated the diversity of CRISPR systems across 124 *S. thermophilus* strains. To date, there have been four distinct CRISPR loci identified in *S. thermophilus*, designated as CRISPR1 through to CRISPR4 (Makarova et al. 2011). In *L. lactis*, only one plasmid-encoded CRISPR/Cas locus has been characterised (Millen et al. 2012).

5 Future perspectives and conclusions

While it is likely that “omics”-based technologies will never completely replace traditional culture-based methods, there is a vast array of knowledge to be gained from integrating these disciplines. Small-scale trial fermentations will continue to be the only genuine test of the performance of starter cultures within an industrial setting; yet, it is an impractical approach for screening large culture banks for suitable strains. The recent advances in NGS technologies have ensured that sequencing has become a more accessible avenue that may permit a rational approach to reduce the number of potential candidates for such trials, and to minimize screening times and labour-intensive culture techniques. Further combinations of sequencing-based approaches with other “omics”-based technologies, such as transcriptomics and proteomics in cheese, may help to moderate the genotype-phenotype link in the future.

Genome decay and redundancy, as highlighted in dairy lactococcal isolates (Makarova et al. 2006; Goh et al. 2011; Ainsworth et al. 2013), coupled to *cremoris* type strains which are believed to be descended from a few closely related lineages (Kelly et al. 2010), are factors likely to limit the selection of novel starter strains in the future. This is perpetuated by the likelihood of large redundancies in culture collections from different institutes globally and the differentiation of many of these strains. Additionally, the possibility of incorrect phenotype/genotype association, such as the plasmid-encoded citrate metabolism trait, is likely to only be resolved by complete genome sequencing.

Acknowledgments Philip Kelleher is funded by the Department of Agriculture, Food and the Marine under the Food Institutional Research Measure (FIRM) (Ref: 10/RD/TMFRC/704–‘CheeseBoard 2015’ project). Douwe van Sinderen is supported by a Principal Investigator award (Ref. No. 13/IA/1953) through Science Foundation Ireland (SFI). Jennifer Mahony is in receipt of a Technology Innovation Development Award (TIDA) (Ref. No. 14/TIDA/2287) funded by Science Foundation Ireland (SFI).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Ainsworth S, Zomer A, de Jager V, Bottacini F, van Hijum SA, Mahony J, van Sinderen D (2013) Complete genome of *Lactococcus lactis* subsp. *cremoris* UC509. 9, host for a model lactococcal P335 bacteriophage. *Genome Announc* 1(1):e00119–00112
- Ainsworth S, Mahony J, van Sinderen D (2014) "The plasmid complement of *Lactococcus lactis* UC509. 9 encodes multiple bacteriophage resistance systems." *Appl. Environ. Microb.*: AEM. 01070–01014

- Ainsworth S, Sadovskaya I, Vinogradov E, Courtin P, Guerardel Y, Mahony J, Grard T, Cambillau C, Chapot-Chartier M-P, van Sinderen D (2014b) Differences in lactococcal cell wall polysaccharide structure are major determining factors in bacteriophage sensitivity. *mBio* 5(3):e00880–00814
- Ainsworth S, Stockdale S, Bottacini F, Mahony J, Sinderen D (2014c) "The *Lactococcus lactis* plasmidome: much learnt, yet still lots to discover.". *FEMS Microbiol Rev* 38(2014):1066–1088
- Allison GE, Klaenhammer TR (1998) Phage resistance mechanisms in lactic acid bacteria. *Int Dairy J* 8(3): 207–226
- Barefoot SF, Klaenhammer TR (1983) Detection and activity of lactacin B, a bacteriocin produced by *Lactobacillus acidophilus*. *Appl Environ Microbiol* 45(6):1808–1815
- Beresford TP, Fitzsimons NA, Brennan NL, Cogan TM (2001) Recent advances in cheese microbiology. *Int Dairy J* 11(4–7):259–274
- Bissett DL, Anderson RL (1974) Lactose and d-Galactose Metabolism in Group N *Streptococci*: Presence of Enzymes for Both the d-Galactose 1-Phosphate and d-Tagatose 6-Phosphate Pathways1. *J Bacteriol* 117(1):318
- Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, Ehrlich SD, Sorokin A (2001) The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* IL1403. *Genome Res* 11(5):731–753
- Bolotin A, Quinquis B, Renault P, Sorokin A, Ehrlich SD, Kulakauskas S, Lapidus A, Goltsman E, Mazur M, Pusch GD (2004) Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat Biotechnol* 22(12):1554–1558
- Bolotin A, Quinquis B, Ehrlich SD, Sorokin A (2012) "Complete genome sequence of *Lactococcus lactis* subsp. *cremoris* A76.". *J Bacteriol* 194(5):1241–1242
- Borucki MK, Peppin JD, White D, Loge F, Call DR (2003) Variation in biofilm formation among strains of *Listeria monocytogenes*. *Appl Environ Microbiol* 69(12):7336–7342
- Boutrou R, Thuault D, Bourgeois C (1995) Identification and characterization of *Streptococcus thermophilus* strains by pulsed-field gel electrophoresis. *J Appl Bacteriol* 79(4):454–458
- Broadbent JR, Barnes M, Brennan C, Strickland M, Houck K, Johnson ME, Steele JL (2002) Contribution of *Lactococcus lactis* cell envelope proteinase specificity to peptide accumulation and bitterness in reduced-fat Cheddar cheese. *Appl Environ Microbiol* 68(4):1778–1785
- Bruseti L, Malkhazova I, Gtari M, Tamagnini I, Borin S, Merabishvili M, Chanishvili N, Mora D, Cappitelli F, Daffonchio D (2008) Fluorescent-PCR for resolving bacterial genetic diversity, endemism and biogeography. *BMC Microbiol* 8(1):220
- Buist G, Kok J, Leenhouts KJ, Dabrowska M, Venema G, Haandrikman AJ (1995) Molecular cloning and nucleotide sequence of the gene encoding the major peptidoglycan hydrolase of *Lactococcus lactis*, a muramidase needed for cell separation. *J Bacteriol* 177(6):1554–1563
- Burrus V, Bontemps C, Decaris B, Guédon G (2001) Characterization of a Novel Type II Restriction-Modification System, Sth3681, Encoded by the Integrative Element ICES1 of *Streptococcus thermophilus* CNRZ368. *Appl Environ Microbiol* 67(4):1522–1528
- Cantor CR, Smith CL, Mathew MK (1988) Pulsed-field gel electrophoresis of very large DNA molecules. *Annu Rev Biophys Biophys* 17(1):287–304
- Chaisson MJ, Brinza D, Pevzner PA (2009) De novo fragment assembly with short mate-paired reads: Does the read length matter? *Genome Res* 19(2):336–346
- Chapot-Chartier M.-P (2014) "Interactions of the cell-wall glycopolymers of lactic acid bacteria with their bacteriophages." *Front Microbiol* 5
- Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10(6):563–569
- Chopin A, Bolotin A, Sorokin A, Ehrlich SD, Chopin M-C (2001) Analysis of six prophages in *Lactococcus lactis* IL1403: different genetic structure of temperate and virulent phage populations. *Nucleic Acids Res* 29(3):644–651
- Chopin M-C, Chopin A, Bidnenko E (2005) Phage abortive infection in lactococci: variations on a theme. *Curr Opin Microbiol* 8(4):473–479
- Chouayekh H, Serror P, Boudebouze S, Maguin E (2009) "Highly efficient production of the staphylococcal nuclease reporter in *Lactobacillus bulgaricus* governed by the promoter of the *hlyA* gene.". *FEMS Microbiol Lett* 293(2):232–239
- Clark TA, Murray IA, Morgan RD, Kislyuk AO, Spittle KE, Boitano M, Fomenkov A, Roberts RJ, Korlach J (2012) Characterization of DNA methyltransferase specificities using single-molecule, real-time DNA sequencing. *Nucleic Acids Res* 40(4):e29

- Clarke J, Wu H-C, Jayasinghe L, Patel A, Reid S, Bayley H (2009) Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol* 4(4):265–270
- Curry A (2013) "The milk revolution." *Nature* 500
- Daveran-Mingot M-L, Campo N, Ritzenthaler P, Le Bourgeois P (1998) A Natural Large Chromosomal Inversion in *Lactococcus lactis* Is Mediated by Homologous Recombination between Two Insertion Sequences. *J Bacteriol* 180(18):4834–4842
- Delcour J, Ferain T, Hols P (2000) Advances in the genetics of thermophilic lactic acid bacteria. *Curr Opin Biotechnol* 11(5):497–504
- Delorme C, Poyart C, Ehrlich SD, Renault P (2007) Extent of horizontal gene transfer in evolution of *streptococci* of the *salivarius* group. *J Bacteriol* 189(4):1330–1341
- Delorme C, Bartholini C, Bolotine A, Ehrlich SD, Renault P (2010) Emergence of a cell wall protease in the *Streptococcus thermophilus* population. *Appl Environ Microbiol* 76(2):451–460
- Delorme C, Bartholini C, Luraschi M, Pons N, Loux V, Almeida M, Guédon E, Gibrat J-F, Renault P (2011) Complete genome sequence of the pigmented *Streptococcus thermophilus* strain JIM8232. *J Bacteriol* 193(19):5581–5582
- Deng Y-M, Liu C-Q, Dunn NW (2000) Lld I, a Plasmid-Encoded Type I Restriction and Modification System in *Lactococcus lactis*. *Mitochondr DNA* 11(3–4):239–245
- Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, Romero DA, Horvath P, Moineau S (2008) Phage Response to CRISPR-Encoded Resistance in *Streptococcus thermophilus*. *J Bacteriol* 190(4):1390–1400
- Deveau H, Garneau JE, Moineau S (2010) CRISPR/Cas System and Its Role in Phage-Bacteria Interactions. *Annu Rev Microbiol* 64(1):475–493
- Dobson A, Cotter PD, Ross RP, Hill C (2012) Bacteriocin Production: a Probiotic Trait? *Appl Environ Microbiol* 78(1):1–6
- Drider D, Bekal S, Prévost H (2004) Genetic organization and expression of citrate permease in lactic acid bacteria. *Genet Mol Res* 3(2):271–281
- Edwards DJ, Holt KE (2013) Beginner's guide to comparative bacterial genome analysis using next-generation sequence data. *Microb Inform Exp* 3(1):2
- Feirtag JM, McKay LL (1987) Thermoinducible Lysis of Temperature-Sensitive *Streptococcus cremoris* Strains. *J Dairy Sci* 70(9):1779–1784
- Fernández E, Á Alegría, S Delgado, MC Martín, B Mayo (2011) "Comparative Phenotypic and Molecular Genetic Profiling of Wild *Lactococcus lactis* subsp. *lactis* Strains of the *lactis* and *cremoris* Genotypes Isolated from Starter-Free Cheeses Made of Raw Milk." *Appl. and Environmental Microb.: AEM*. 02991–02910
- Ferrario C, Ricci G, Borgo F, Rollando A, Fortina MG (2012) Genetic investigation within *Lactococcus garvieae* revealed two genomic lineages. *FEMS Microbiol Lett* 332(2):153–161
- Fitzgerald GF, Daly C, Brown LR, Gingeras TR (1982) ScrFI: a new sequence-specific endonuclease from *Streptococcus cremoris*. *Nucleic Acids Res* 10(24):8171–8179
- Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korlach J, Turner SW (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods* 7(6):461–465
- Fouts DE (2006) Phage_Finder: automated identification and classification of prophage regions in complete bacterial genome sequences. *Nucleic Acids Res* 34(20):5839–5851
- Gancheva A, Pot B, Vanhonacker K, Hoste B, Kersters K (1999) A Polyphasic Approach towards the Identification of Strains Belonging to *Lactobacillus acidophilus* and Related Species. *Syst Appl Microbiol* 22(4):573–585
- Gao Y, Lu Y, Teng K-L, Chen M-L, Zheng H-J, Zhu Y-Q, Zhong J (2011) Complete genome sequence of *Lactococcus lactis* subsp. *lactis* CV56, a probiotic strain isolated from the vaginas of healthy women. *J Bacteriol* 193(11):2886–2887
- Garvey P, Fitzgerald G, Hill C (1995) Cloning and DNA sequence analysis of two abortive infection phage resistance determinants from the lactococcal plasmid pNP40. *Appl Environ Microbiol* 61(12):4321–4328
- Gasson MJ, Davies FL (1980) Prophage-Cured Derivatives of *Streptococcus lactis* and *Streptococcus cremoris*. *Appl Environ Microbiol* 40(5):964–966
- Gautier M, Chopin M-C (1987) Plasmid-determined systems for restriction and modification activity and abortive infection in *Streptococcus cremoris*. *Appl Environ Microbiol* 53(5):923–927
- Gilles A, Meglec E, Pech N, Ferreira S, Malausa T, Martin J-F (2011) Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC Genomics* 12(1):245
- Goh YJ, 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. *Microbiol Cell Fact* 10(Suppl 1):S22

- Guédon E, Renault P, Ehrlich SD, Delorme C (2001) Transcriptional Pattern of Genes Coding for the Proteolytic System of *Lactococcus lactis* and Evidence for Coordinated Regulation of Key Enzymes by Peptide Supply. *J Bacteriol* 183(12):3614–3622
- Guimont C, Henry P, Linden G (1993) Restriction/modification in *Streptococcus thermophilus*: isolation and characterization of a type II restriction endonuclease Sth455I. *Appl Microbiol Biotechnol* 39(2):216–220
- Harrigan WF (1998) Laboratory methods in food microbiology (3rd ed). Academic Press, San Diego
- Hassan AN (2008) ADSA Foundation Scholar Award: Possibilities and Challenges of Exopolysaccharide-Producing Lactic Cultures in Dairy Foods. *J Dairy Sci* 91(4):1282–1298
- Hassan AN, Frank JF, Farmer MA, Schmidt KA, Shalabi SI (1995) Formation of Yogurt Microstructure and Three-Dimensional Visualization as Determined by Confocal Scanning Laser Microscopy. *J Dairy Sci* 78(12):2629–2636
- Hassan AN, Ipsen R, Janzen T, Qvist KB (2003) Microstructure and Rheology of Yogurt Made with Cultures Differing Only in Their Ability to Produce Exopolysaccharides. *J Dairy Sci* 86(5):1632–1638
- Hill C, Pierce K, Klaenhammer T (1989) The conjugative plasmid pTR2030 encodes two bacteriophage defense mechanisms in lactococci, restriction modification (R+/M+) and abortive infection (Hsp+). *Appl Environ Microbiol* 55(9):2416–2419
- Hodkinson BP, EA Grice (2015) "Next-Generation Sequencing: A Review of Technologies and Tools for Wound Microbiome Research." *Advances in wound care*
- Horvath P, Barrangou R (2010) CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science* 327(5962):167–170
- Horvath P, Romero DA, Coûté-Monvoisin A-C, Richards M, Deveau H, Moineau S, Boyaval P, Fremaux C, Barrangou R (2008) Diversity, Activity, and Evolution of CRISPR Loci in *Streptococcus thermophilus*. *J Bacteriol* 190(4):1401–1412
- Huang Y-F, Chen S-C, Chiang Y-S, Chen T-H, Chiu K-P (2012) Palindromic sequence impedes sequencing-by-ligation mechanism. *BMC Syst Biol* 6(Suppl 2):S10
- Husson-Kao C, Mengaud J, Cesselin B, van Sinderen D, Benbadis L, Chapot-Chartier M-P (2000) The *Streptococcus thermophilus* autolytic phenotype results from a leaky prophage. *Appl Environ Microbiol* 66(2):558–565
- Iyer R, Tomar S, Mohanty A, Singh P, Singh R (2011) Bioprospecting of strains of *Streptococcus thermophilus* from Indian fermented milk products for folate production. *Dairy Sci Technol* 91(2):237–246
- Kang X, Ling N, Sun G, Zhou Q, Zhang L, Sheng Q (2012) Complete genome sequence of *Streptococcus thermophilus* strain MN-ZLW-002. *J Bacteriol* 194(16):4428–4429
- Karginov FV, Hannon GJ (2010) The CRISPR System: Small RNA-Guided Defense in Bacteria and Archaea. *Mol Cell* 37(1):7–19
- Kato T, Nagatsu T, Kimura T, Sakakibara S (1978) Fluorescence assay of x-prolyldipeptidyl-aminopeptidase activity with a new fluorogenic substrate. *Biochem Med* 19(3):351–359
- Kato H, Shiwa Y, Oshima K, Machii M, Araya-Kojima T, Zendo T, Shimizu-Kadota M, Hattori M, Sonomoto K, Yoshikawa H (2012) Complete genome sequence of *Lactococcus lactis* IO-1, a lactic acid bacterium that utilizes xylose and produces high levels of L-lactic acid. *J Bacteriol* 194(8):2102–2103
- Kékessy DA, Piguet JD (1970) New Method for Detecting Bacteriocin Production. *Appl Microbiol* 20(2):282–283
- Kelly WJ, Ward LJ, Leahy SC (2010) Chromosomal diversity in *Lactococcus lactis* and the origin of dairy starter cultures. *Genome Biol Evol* 2:729–744
- Kelly WJ, E Altermann, SC Lambie, SC Leahy (2013) "Interaction between the genomes of *Lactococcus lactis* and phages of the P335 species." *Front Microbiol* 4
- Khemariya P, Singh S, Nath G, Gulati AK (2013) Development of a PCR-RFLP assay for the identification of *Lactococcus lactis* ssp. *lactis* and *cremoris*. *Ann Microbiol* 63(1):109–115
- Kieronczyk A, Skeie S, Langsrud T, Yvon M (2003) Cooperation between *Lactococcus lactis* and nonstarter lactobacilli in the formation of cheese aroma from amino acids. *Appl Environ Microbiol* 69(2):734–739
- Ko euth T, Versalovic J, Lupski JR (1995) Differential subsequence conservation of interspersed repetitive *Streptococcus pneumoniae* BOX elements in diverse bacteria. *Genome Res* 5(4):408–418
- Kok J, Buist G, Zomer AL, Van Hijum SAFT, Kuipers OP (2005) Comparative and functional genomics of lactococci. *FEMS Microbiol Rev* 29(3):411–433
- Lander ES, Linton LM, Birren B, Nussbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822):860–921
- Lazzi C, Bove CG, Sgarbi E, Monica G, La Gioia F, Sandra T, Neviani E (2009) Application of AFLP fingerprint analysis for studying the biodiversity of *Streptococcus thermophilus*. *J Microbiol Methods* 79(1):48–54
- Leroy F, De Vuyst L (2004) Lactic acid bacteria as functional starter cultures for the food fermentation industry. *Trends Food Sci Technol* 15(2):67–78

- Lick S, Keller M, Bockelmann W, Heller J (1996) Rapid Identification of *Streptococcus thermophilus* by Primer-specific PCR Amplification Based on its lacZ Gene. *Syst Appl Microbiol* 19(1):74–77
- Linares DM, Kok J, Poolman B (2010) Genome sequences of *Lactococcus lactis* MG1363 (revised) and NZ9000 and comparative physiological studies. *J Bacteriol* 192(21):5806–5812
- Liu M, Bayjanov J, Renckens B, Nauta A, Siezen R (2010) The proteolytic system of lactic acid bacteria revisited: a genomic comparison. *BMC Genomics* 11(1):36
- Loenen WA, DT Dryden, EA Raleigh, GG Wilson (2013) "Type I restriction enzymes and their relatives." *Nucleic Acids Res.*: gkt847
- Loman NJ, Constantinidou C, Chan JZM, Halachev M, Sergeant M, Penn CW, Robinson ER, Pallen MJ (2012a) High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nat Rev Microbiol* 10(9):599–606
- Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ (2012b) Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* 30(5):434–439
- Mahony J, McGrath S, Fitzgerald GF, van Sinderen D (2008) Identification and characterization of lactococcal-prophage-carried superinfection exclusion genes. *Appl Environ Microbiol* 74(20):6206–6215
- Mahony J, Kot W, Murphy J, Ainsworth S, Neve H, Hansen LH, Heller KJ, Sørensen SJ, Hammer K, Cambillau C (2013) Investigation of the relationship between lactococcal host cell wall polysaccharide genotype and 936 phage receptor binding protein phylogeny. *Appl Environ Microbiol* 79(14):4385–4392
- Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA (1998) Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A* 95(6):3140–3145
- Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, Shakhova V, Grigoriev I, Lou Y, Rohksar D, Lucas S, Huang K, Goodstein DM, Hawkins T, Plengvidhya V, Welker D, Hughes J, Goh Y, Benson A, Baldwin K, Lee JH, Diaz-Muniz I, Dosti B, Smeianov V, Wechter W, Barabote R, Lorca G, Altermann E, Barrangou R, Ganesan B, Xie Y, Rawsthorne H, Tamir D, Parker C, Breidt F, Broadbent J, Hutkins R, O'Sullivan D, Steele J, Unlu G, Saier M, Klaenhammer T, Richardson P, Kozyavkin S, Weimer B, Mills D (2006b) Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A* 103(42):15611–15616
- Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, Moineau S, Mojica FJM, Wolf YI, Yakunin AF, van der Oost J, Koonin EV (2011) Evolution and classification of the CRISPR–Cas systems. *Nat Rev Microbiol* 9(6):467–477
- Mathew MK, Smith CL, Cantor CR (1988) High-resolution separation and accurate size determination in pulsed-field gel electrophoresis of DNA. 1. DNA size standards and the effect of agarose and temperature. *Biochem* 27(26):9204–9210
- McCulloch JA, de Oliveira VM, de Almeida Pina AV, Pérez-Chaparro PJ, de Almeida LM, de Vasconcelos JM, de Oliveira LF, da Silva DEA, Rogez HLG, Cretenet M (2014) Complete genome sequence of *Lactococcus lactis* strain AI06, an endophyte of the Amazonian açai palm. *Genome Announc* 2(6):e01225–01214
- McGrath S, Fitzgerald GF, van Sinderen D (2002) Identification and characterization of phage-resistance genes in temperate lactococcal bacteriophages. *Mol Microbiol* 43(2):509–520
- McKay LL, Baldwin KA, Zottola EA (1972) Loss of Lactose Metabolism in Lactic *Streptococci*. *Appl Microbiol* 23(6):1090–1096
- McSweeney PLH (2004) Biochemistry of cheese ripening. *Int J Dairy Technol* 57(2–3):127–144
- McSweeney PLH, Sousa MJ (2000) Biochemical pathways for the production of flavour compounds in cheeses during ripening: A review. *Lait* 80(3):293–324
- Metzker ML (2010) Sequencing technologies –the next generation. *Nat Rev Genet* 11(1):31–46
- Millen AM, Horvath P, Boyaval P, Romero DA (2012) Mobile CRISPR/Cas-mediated bacteriophage resistance in *Lactococcus lactis*. *PLoS ONE* 7(12):e51663
- Mills S, Griffin C, Coffey A, Meijer W, Hafkamp B, Ross R (2010) CRISPR analysis of bacteriophage-insensitive mutants (BIMs) of industrial *Streptococcus thermophilus*—implications for starter design. *J Appl Microbiol* 108(3):945–955
- Mosher JJ, Bernberg EL, Shevchenko O, Kan J, Kaplan LA (2013) Efficacy of a 3rd generation high-throughput sequencing platform for analyses of 16S rRNA genes from environmental samples. *J Microbiol Methods* 95(2):175–181

- Mruk I, Cichowicz M, Kaczorowski T (2003) Characterization of the LlaCI methyltransferase from *Lactococcus lactis* subsp. *cremoris* W15 provides new insights into the biology of type II restriction–modification systems. *Microbiol* 149(11):3331–3341
- Murphy J, Royer B, Mahony J, Hoyles L, Heller K, Neve H, Bonestroo M, Nauta A, van Sinderen D (2013) Biodiversity of lactococcal bacteriophages isolated from 3 Gouda-type cheese-producing plants. *J Dairy Sci* 96(8):4945–4957
- Murphy J, Klumpp J, Mahony J, Mary O, Nauta A, van Sinderen D (2014) Methyltransferases acquired by lactococcal 936-type phage provide protection against restriction endonuclease activity. *BMC Genomics* 15(1):831
- Murray IA, Clark TA, Morgan RD, Boitano M, Anton BP, Luong K, Fomenkov A, Turner SW, Korlach J, Roberts RJ (2012) The methylomes of six bacteria. *Nucleic Acids Res* 40(22):11450–11462
- Nardi M, Fiez-Vandal C, Tailliez P, Monnet V (2002) The EstA esterase is responsible for the main capacity of *Lactococcus lactis* to synthesize short chain fatty acid esters in vitro. *J Appl Microbiol* 93(6):994–1002
- Neve H, Freudenberg W, Diestel-Feddersen F, Ehlert R, Heller KJ (2003) Biology of the temperate *Streptococcus thermophilus* bacteriophage TP-J34 and physical characterization of the phage genome. *Virology* 315(1):184–194
- Nomura M, Kobayashi M, Okamoto T (2002) Rapid PCR-based method which can determine both phenotype and genotype of *Lactococcus lactis* subspecies. *Appl Environ Microbiol* 68(5):2209–2213
- Nyengaard N, Vogensen FK, Josephsen J (1993) LlaAI and LlaBI, two type-II restriction endonucleases from *Lactococcus lactis* subsp. *cremoris* W9 and W56 recognizing, respectively, 5'-GATC-3' and 5'-C/TRYA G-3'. *Gene* 136(1):371–372
- O'Connell-Motherway M, Watson D, Bottacini F, Clark TA, Roberts RJ, Korlach J, Garault P, Chervaux C, van Hylckama Vlieg JET (2014) Identification of Restriction-Modification Systems of *Bifidobacterium animalis* subsp. *lactis* CNCM I-2494 by SMRT Sequencing and Associated Methylome Analysis. *PLoS ONE* 9(4):e94875
- O'driscoll J, Glynn F, Cahalane O, O'Connell-Motherway M, Fitzgerald GF, Van Sinderen D (2004) Lactococcal plasmid pNP40 encodes a novel, temperature-sensitive restriction-modification system. *Appl Environ Microbiol* 70(9):5546–5556
- Olive DM, Bean P (1999) Principles and applications of methods for DNA-based typing of microbial organisms. *J Clin Microbiol* 37(6):1661–1669
- Oliveira LC, Saraiva TD, Soares SC, Ramos RT, Sá PH, Carneiro AR, Miranda F, Freire M, Renan W, Júnior AF (2014) Genome Sequence of *Lactococcus lactis* subsp. *lactis* NCDO 2118, a GABA-Producing Strain. *Genome Announc* 2(5):e00980–00914
- Passerini D, Beltramo C, Coddeville M, Quentin Y, Ritzenthaler P, Daveran-Mingot M-L, Le Bourgeois P (2010) Genes but not genomes reveal bacterial domestication of *Lactococcus lactis*. *PLoS ONE* 5(12):e15306
- Paun O, P Schönswetter (2012) Amplified fragment length polymorphism: an invaluable fingerprinting technique for genomic, transcriptomic, and epigenetic studies. *Plant DNA Fingerprinting and Barcoding*, Springer: 75–87
- Pettersson E, Lundeberg J, Ahmadian A (2009) Generations of sequencing technologies. *Genomics* 93(2): 105–111
- Pingoud A, GG Wilson, W Wende (2014) "Type II restriction endonucleases—a historical perspective and more." *Nucleic Acids Res.*: gku447
- Prevots F, Ritzenthaler P (1998) Complete Sequence of the new Lactococcal Abortive Phage Resistance Gene *abiO*. *J Dairy Sci* 81(6):1483–1485
- Prodělalová J, Španová A, Rittich B (2005) Application of PCR, rep-PCR and RAPD techniques for typing of *Lactococcus lactis* strains. *Folia Microbiol* 50(2):150–154
- Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13(1):341
- Rademaker JL, Herbet H, Starrenburg MJ, Naser SM, Gevers D, Kelly WJ, Hugenholtz J, Swings J, van Hylckama Vlieg JE (2007) Diversity analysis of dairy and nondairy *Lactococcus lactis* isolates, using a novel multilocus sequence analysis scheme and (GTG) 5-PCR fingerprinting. *Appl Environ Microbiol* 73(22):7128–7137
- Rao DN, Dryden DT, Bheemanaik S (2014) Type III restriction-modification enzymes: a historical perspective. *Nucleic Acids Res* 42(1):45–55
- Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, Blumenthal RM, Degtyarev SK, Dryden DT, Dybvig K (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res* 31(7):1805–1812

- Rodríguez C, Van der Meulen R, Vaningelgem F, Font de Valdez G, Raya R, De Vuyst L, Mozzi F (2008) Sensitivity of capsular-producing *Streptococcus thermophilus* strains to bacteriophage adsorption. *Let Appl Microbiol* 46(4):462–468
- Rossetti L, Giraffa G (2005) Rapid identification of dairy lactic acid bacteria by M13-generated, RAPD-PCR fingerprint databases. *J Microbiol Methods* 63(2):135–144
- Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230(4732):1350–1354
- Samarzija D, Sikora S, Redzepovic S, Antunac N, Havranek J (2002) Application of RAPD analysis for identification of *Lactococcus lactis* subsp. *cremoris* strains isolated from artisanal cultures. *Microbiol Res* 157(1):13–17
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci* 74(12):5463–5467
- Schleifer KH, Kilpper-Bälz R (1987) Molecular and Chemotaxonomic Approaches to the Classification of *Streptococci*, *Enterococci* and *Lactococci*: A Review. *Syst Appl Microbiol* 10(1):1–19
- Schleifer K, Kraus J, Dvorak C, Kilpper-Bälz R, Collins M, Fischer W (1985) Transfer of *Streptococcus lactis* and Related *Streptococci* to the Genus *Lactococcus* gen. nov. *Syst Appl Microbiol* 6(2):183–195
- Schloss PD, SL Westcott, ML Jenior, SK Highlander (2015) Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system, PeerJ PrePrints
- Schroeder CJ, Robert C, Lenzen G, McKay LL, Mercenier A (1991) Analysis of the lacZ sequences from two *Streptococcus thermophilus* strains: comparison with the *Escherichia coli* and *Lactobacillus bulgaricus* beta-galactosidase sequences. *J Gen Microbiol* 137(2):369–380
- Shibamoto T (2014) Diacetyl: Occurrence, Analysis, and Toxicity. *J Agric Food Chem* 62(18):4048–4053
- Siezen RJ, Renckens B, van Swam I, Peters S, van Kranenburg R, Kleerebezem M, de Vos WM (2005) Complete sequences of four plasmids of *Lactococcus lactis* subsp. *cremoris* SK11 reveal extensive adaptation to the dairy environment. *Appl Environ Microbiol* 71(12):8371–8382
- Siezen RJ, Bayjanov J, Renckens B, Wels M, van Hijum SA, Molenaar D, van Hylckama Vlieg JE (2010) "Complete genome sequence of *Lactococcus lactis* subsp. *lactis* KF147, a plant-associated lactic acid bacterium." *J Bacteriol* 192(10):2649–2650
- Sing WD, Klaenhammer TR (1986) Conjugal transfer of bacteriophage resistance determinants on pTR2030 into *Streptococcus cremoris* strains. *Appl Environ Microbiol* 51(6):1264–1271
- Sistla S, Rao DN (2004) S-Adenosyl-L-methionine-dependent restriction enzymes. *Crit Rev Biochem Mol Biol* 39(1):1–19
- Smid EJ, O Erkus, M. Spus, J. C. Wolkers-Rooijackers, S. Alexeeva and M. Kleerebezem (2014). "Functional implications of the microbial community structure of undefined mesophilic starter cultures." *Microb. Cell Fact.* 13(suppl. 1)
- Solopova A, Bachmann H, Teusink B, Kok J, Neves AR, Kuipers OP (2012) A Specific Mutation in the Promoter Region of the Silent cel Cluster Accounts for the Appearance of Lactose-Utilizing *Lactococcus lactis* MG1363. *Appl Environ Microbiol* 78(16):5612–5621
- Sozhamannan S, Chute MD, McAfee FD, Fouts DE, Akmal A, Galloway DR, Mateczun A, Baillie LW, Read TD (2006) The *Bacillus anthracis* chromosome contains four conserved, excision-proficient, putative prophages. *BMC Microbiol* 6(1):34
- Steele J, Broadbent J, Kok J (2013) Perspectives on the contribution of lactic acid bacteria to cheese flavor development. *Curr Opin Biotechnol* 24(2):135–141
- Stern MJ, Ames GF-L, Smith NH, Clare Robinson E, Higgins CF (1984) Repetitive extragenic palindromic sequences: a major component of the bacterial genome. *Cell* 37(3):1015–1026
- Sun Z, Chen X, Wang J, Zhao W, Shao Y, Wu L, Zhou Z, Sun T, Wang L, Meng H (2011) Complete genome sequence of *Streptococcus thermophilus* strain ND03. *J Bacteriol* 193(3):793–794
- Tailliez P, Tremblay J, Ehrlich SD, Chopin A (1998) Molecular Diversity and Relationship within *Lactococcus lactis*, as revealed by Randomly Amplified Polymorphic DNA (RAPD). *Syst Appl Microbiol* 21(4):530–538
- Tanous C, A Kieronczyk, S Helinck, E Chambellon, M Yvon (2002) Glutamate dehydrogenase activity: a major criterion for the selection of flavour-producing lactic acid bacteria strains. *Lactic Acid Bacteria: Genetics, Metabolism and Applications*, Springer: 271–278
- Tanskanen EI, Tulloch DL, Hillier AJ, Davidson BE (1990) Pulsed-field gel electrophoresis of SmaI digests of lactococcal genomic DNA, a novel method of strain identification. *Appl Environ Microbiol* 56(10):3105–3111
- Tobes R, Ramos JL (2005) REP code: defining bacterial identity in extragenic space. *Environ Microbiol* 7(2): 225–228

- Turgeon N, Moineau S (2001) Isolation and Characterization of a *Streptococcus thermophilus* Plasmid Closely Related to the pMV158 Family. *Plasmid* 45(3):171–183
- Urbach E, Daniels B, Salama MS, Sandine WE, Giovannoni SJ (1997) The *ldh* phylogeny for environmental isolates of *Lactococcus lactis* is consistent with rRNA genotypes but not with phenotypes. *Appl Environ Microbiol* 63(2):694–702
- Urbach E, Schindler C, Giovannoni SJ (1998) A PCR fingerprinting technique to distinguish isolates of *Lactococcus lactis*. *FEMS Microbiol Lett* 162(1):111–115
- van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJJ (2009) CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem Sci* 34(8):401–407
- van Rooijen RJ, de Vos WM (1990) Molecular cloning, transcriptional analysis, and nucleotide sequence of *lacR*, a gene encoding the repressor of the lactose phosphotransferase system of *Lactococcus lactis*. *J Biol Chem* 265(30):18499–18503
- van Rooijen R, Gasson M, De Vos W (1992) Characterization of the *Lactococcus lactis* lactose operon promoter: contribution of flanking sequences and LacR repressor to promoter activity. *J Bacteriol* 174(7):2273–2280
- Vaughan EE, Pridmore RD, Mollet B (1998) Transcriptional Regulation and Evolution of Lactose Genes in the Galactose-Lactose Operon of *Lactococcus lactis* NCDO2054. *J Bacteriol* 180(18):4893–4902
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA (2001) The sequence of the human genome. *Science* 291(5507):1304–1351
- Verger R (1997) ‘Interfacial activation’ of lipases: facts and artifacts. *Trends Biotechnol* 15(1):32–38
- Versalovic J, Koeuth T, Lupski R (1991) Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* 19(24):6823–6831
- Versalovic J, Kapur V, Koeuth T, Mazurek GH, Whittam TS, Musser JM, Lupski JR (1995) DNA fingerprinting of pathogenic bacteria by fluorophore-enhanced repetitive sequence-based polymerase chain reaction. *Arch Pathol Lab Med* 119(1):23–29
- Villani F, Aponte M, Blaiotta G, Mauriello G, Pepe O, Moschetti G (2001) Detection and characterization of a bacteriocin, garviecin L1-5, produced by *Lactococcus garvieae* isolated from raw cow’s milk. *J Appl Microbiol* 90(3):430–439
- Vos P, Hogers R, Bleeker M, Reijns M, Van de Lee T, Hornes M, Friters A, Pot J, Paleman J, Kuiper M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* 23(21):4407–4414
- Vuylsteke M, Peleman JD, van Eijk MJ (2007) AFLP technology for DNA fingerprinting. *Nat Protoc* 2(6):1387–1398
- Ward LJ, Brown J, Davey GP (1998) Two methods for the genetic differentiation of *Lactococcus lactis* ssp. *lactis* and *cremoris* based on differences in the 16S rRNA gene sequence. *FEMS Microbiol Lett* 166(1):15–20
- Ward L, Heap H, Kelly W (2004) Characterization of closely related lactococcal starter strains which show differing patterns of bacteriophage sensitivity. *J Appl Microbiol* 96(1):144–148
- Wegmann U, O’Connell-Motherway M, Zomer A, Buist G, Shearman C, Canchaya C, Ventura M, Goesmann A, Gasson MJ, Kuipers OP (2007) Complete genome sequence of the prototypic lactic acid bacterium *Lactococcus lactis* subsp. *cremoris* MG1363. *J Bacteriol* 189(8):3256–3270
- Whitehead HR, G Cox (1935) "The Occurrence of Bacteriophage in Cultures of Lactic *Streptococci*: A Preliminary Note." *N Z J Dairy Sci Technol*: 313–320
- Wilkinson MG, Guinee TP, O’Callaghan DM, Fox PF (1994) Autolysis and proteolysis in different strains of starter bacteria during Cheddar cheese ripening. *J Dairy Res* 61(02):249–262
- Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18(22):6531–6535
- Wilson LA, Sharp PM (2006) Enterobacterial repetitive intergenic consensus (ERIC) sequences in *Escherichia coli*: Evolution and implications for ERIC-PCR. *Mol Biol Evol* 23(6):1156–1168
- Wittenberger CL, Angelo N (1970) Purification and properties of a fructose-1,6-diphosphate-activated lactate dehydrogenase from *Streptococcus faecalis*. *J Bacteriol* 101:717–734
- Wu, Q., H. M. Tun, F. C.-C. Leung and N. P. Shah (2014). "Genomic insights into high exopolysaccharide-producing dairy starter bacterium *Streptococcus thermophilus* ASCC 1275." *Sci. Rep.* 4
- Yang R, Johnson MC, Ray B (1992) Novel method to extract large amounts of bacteriocins from lactic acid bacteria. *Appl Environ Microbiol* 58(10):3355–3359
- Yang X, Wang Y, Huo G (2013) Complete Genome Sequence of *Lactococcus lactis* subsp. *lactis* KLDS4.0325. *Genome Announc* 1(6):e00962–00913
- Yu W, Gillies K, Kondo JK, Broadbent JR, McKay LL (1996) Loss of Plasmid-Mediated Oligopeptide Transport System in *Lactococci*: Another Reason for Slow Milk Coagulation. *Plasmid* 35(3):145–155
- Zinno P, Janzen T, Bennedsen M, Ercolini D, Mauriello G (2010) Characterization of *Streptococcus thermophilus* lytic bacteriophages from mozzarella cheese plants. *Int J Food Microbiol* 138(1):137–144