REVIEW ARTICLE

The Evolution of gene regulation research in Lactococcus lactis

Jan Kok*, Lieke A. van Gijtenbeek, Anne de Jong, Sjoerd B. van der Meulen, Ana Solopova and Oscar P. Kuipers

Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, the Netherlands

*Corresponding author: Department of Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, the Netherlands. Tel: + 50 363 2111; E-mail: jan.kok@rug.nl

One sentence summary: An overview is given of the big strides that have been made over the past 50 years in the genetics and genomics analyses of Lactococcus lactis, with an emphasis on gene regulation in this important industrial microorganism.

Editor: Eddy Smid

ABSTRACT

Lactococcus lactis is a major microbe. This lactic acid bacterium (LAB) is used worldwide in the production of safe, healthy, tasteful and nutritious milk fermentation products. Its huge industrial importance has led to an explosion of research on the organism, particularly since the early 1970s. The upsurge in the research on L. lactis coincided not accidentally with the advent of recombinant DNA technology in these years. The development of methods to take out and re-introduce DNA in L. lactis, to clone genes and to mutate the chromosome in a targeted way, to control (over)expression of proteins and, ultimately, the availability of the nucleotide sequence of its genome and the use of that information in transcriptomics and proteomics research have enabled to peek deep into the functioning of the organism. Among many other things, this has provided an unprecedented view of the major gene regulatory pathways involved in nitrogen and carbon metabolism and their overlap, and has led to the blossoming of the field of L. lactis systems biology. All of these advances have made L. lactis the paradigm of the LAB. This review will deal with the exciting path along which the research on the genetics of and gene regulation in L. lactis has trodden.

Keywords: Lactococcus lactis; gene regulation; -omics; gene technology; systems biology; single-cell analysis

INTRODUCTION

A lot of research has been devoted in the last half century to the architype of lactic acid bacterium (LAB), Lactococcus lactis. This should not surprise us, as it is an economically highly important microorganism that is used throughout the world for the production of nutritious and healthy foods and is used as such already since ancient times. Man has been using L. lactis unknowingly for many thousands of years in food (milk) fermentations. Fermentations were started by ‘back-slopping’: the inoculation of fresh milk with a portion of a previous fermentation. Cheese making has changed in half a century from a small scale rather local activity to a genuine industrial process with almost complete automation and little intervention by man. The demands on the starter culture, the mixtures of various strains of L. lactis and, sometimes, other species of LAB that are added to the cheese milk have enormously increased. Most importantly, the bacteria are supposed to do their job in a predictable and stable manner, quickly acidifying the milk and, more slowly, providing flavor, taste and texture to the end products. In other words, L. lactis is expected to ‘run like a machine’, smoothly and without any restraint. In fact, it does so under proper and stable...
growth conditions while keeping bacteriophages at bay or, even better, away.

*Lactococcus lactis* is not only one of the oldest ‘domesticated’ bacterial species, but was also one of the first to be isolated in pure form. To be able to follow and understand the literature, it is important to realize that *Lactococcus lactis* has changed names several times over the almost 150 years that is being studied, mostly to emphasize its food-related and healthy role (Teuber 1995). At the end of the 19th century, after Pasteur had examined lactic acid fermentation, Lister produced the first pure culture of ‘Bacterium lactis’ (Lister 1873). It was later called ‘*Streptococcus lactis*’ (Orla-Jensen 1919, 1942; Teuber 1995) and part of the genus *Streptococcus* Group N (Lancefield 1993; Wicken and Knox 1975). Until 1985, two species were identified: *S. lactis* (including *S. lactis* biovar. *diacetylactis*) and *S. cremoris*. Schleifer et al. (1985) proposed to transfer both to a new genus, ‘*Lactococcus*’, as two subspecies, namely *L. lactis* ssp. *cremoris* and *L. lactis* ssp. *lactis*. Several other species of Lactococcus have been identified and described, which are not employed in milk fermentation and of which a number are in fact associated with food spoilage, bovine mastitis or fish lactococcosis (Teuber 1995; Rahkila et al. 2012; Pothakos et al. 2014).

As indicated, the main importance of *L. lactis* lies in the dairy industry. Consequently, a plethora of research papers deal with the important industrial characteristics of the organism, such as its sugar fermentation capacities, its ability to break down milk protein, to form aroma and texturizing as well as antimicrobial compounds, its sensitivity to bacteriophages and resistance to a variety of (industrial) stressors. The Symposium on Lactic Acid Bacteria, held triannually in the Netherlands since 1983, have carefully documented these and many more advances in the field with comprehensive review papers in publicly available conference proceedings. In-depth knowledge has been gained over the last five decades on these attributes of *L. lactis*. For a large part, this has been made possible by the development, early on, of tools and techniques to genetically access, modify and, thus, study the organism. The implementation of genetic engineering methodologies, keeping pace with the advances made in this field in the other bacterial work horses, *Escherichia coli* and *Bacillus subtilis*, has allowed targeting single genes and operons, unraveling their functions, understanding their regulation, and revealing (metabolic) pathways in the cell and their interconnections. A major step forward could subsequently be made around the turn of the millennium, with the advances made in genome sequencing. These developments ultimately enabled genome-wide studies and represented a switch from genetics research, studying one or only a few genes at a time, to genomics research in which, in principle (the expression of) all genes of an organism as possible by combining the ‘big data’ obtained from different -omics studies. It also meant that a change was possible to applying discovery rather than hypothesis-driven approaches in the research of *L. lactis*.

Novel uses of *L. lactis* have emerged as a spin-off of the studies performed of its ‘traditional’ features. For example, it has been modified into a carrier, deliverer and/or displacer of oral vaccines, and as an expression host for proteins and antimicrobial peptides such as a variety of lantibiotics (Mierau and Kleerebezem 2005; Wyszyńska et al. 2015). In fact, it was the first genetically modified microorganism that was used as a live vehicle for the delivery of a therapeutic protein (interleukin 10) in the treatment of a human disease, Crohn’s disease (Braat et al. 2006). *Lactococcus lactis* is a credible alternative for *E. coli* and *B. subtilis* for the production of recombinant membrane and secreted proteins (Kunji et al. 2005; Morello et al. 2008). These new applications of *L. lactis* would not have been possible without the incredible depth of know-how gained on the molecular biology of the organism; they represent major advances that could have far-fetching implications for the food and health sector and offer excellent possibilities for successful startup companies.

In this review, an overview is given of the big strides that have been made over the past 50 years or so in the genetics and genomics analyses of *L. lactis*. The field is enormous and that is why an emphasis will be placed on the progress in the research on gene regulation and gene regulatory networks (GRN) that are operative in *L. lactis*.

**THE EARLY YEARS: DEVELOPMENT OF GENETIC ENGINEERING IN LACTOCoccus Lactis**

Many years of research have gone into the traits of *L. lactis* that are important for its use in industrial milk fermentations. Physiology, (bio)chemistry and enzymology all have been applied to describe and understand carbon metabolism and proteolysis and their role in successful fermentation and in flavor and texturizing attributes of this important dairy organism. Biochemical knowledge of lactic acid fermentation really boomed in the 1950s and has led to the first detailed descriptions of the various industrially important pathways present in *L. lactis* and, for that matter, in other LAB species (Kandler 1983; Law and Kolstad 1983). One aspect of some of the major pathways, lactose and protein utilization by *L. lactis*, has intrigued researchers for many years namely the often strain-dependent unpredictability of their activity. The variability of lactose fermenting capacity was described as early as in the 1930s by a number of different researchers (McKay 1983 and references therein). A major breakthrough came with the discovery of plasmids in the organism, finally explaining the instability of not only lactose utilization but quite some of the other industrial phenotypes. This observation energized the research on the genetics of *L. lactis* and has ultimately led to the successful implementation of gene technology.

**Plasmids and traits**

Cords, McKay and Guerry (1974), in a publication on extrachromosomal elements in group-N streptococci, reported for the first time on the occurrence of plasmids in *L. lactis*. A number of critically important traits of *L. lactis*, such as lactose metabolism, proteolytic activity, exopolysaccharide production, phage resistance and the fermentation of citrate, were subsequently shown to be encoded by plasmids of various sizes in many industrially used strains of *L. lactis* (Davies and Gasson 1981; Ainsworth et al. 2014). McKay (1983) offered one of the earliest reviews on plasmid functionalities in lactococci, on the first Symposium [on] Lactic Acid Bacteria in 1983. The advent of genetic engineering techniques, most notably the development of the fast and reliable technique of agarose gel electrophoresis, simplified, accelerated and thereby revolutionized the analysis of plasmids in *L. lactis*. We now know that strains of *L. lactis* typically contain from nil to up to more than 10 different plasmids ranging in size from a little over 2000 bp to over 100 kbp (Klaenhammer, McKay and Baldwin 1978).
Certain plasmids were shown to be transferable by conjugation (for a recent overview, see Ainsworth et al. 2014), which was one of the ways used in the early days to obtain evidence for the location of structural genes for industrial functions on plasmids (Kempler and McKay 1979; Gasson and Davies 1980; Kondo and McKay 1985; Gasson and Fitzgerald 1994). To unequivocally prove this point required the cloning and molecular analysis of these genes, which was initially done by introducing them in heterologous hosts (Bacillus subtilis and Escherichia coli). The Holy Grail, of course, was to be able to clone and analyze the genes in L. lactis itself. To do this, it was essential to develop gene-cloning strategies for L. lactis, which entailed the construction of gene cloning vectors as well as a way to introduce (plasmid) DNA into the organism. In 1982, two groups simultaneously reported the uptake of DNA by protoplasts of L. lactis. Geis (1982) introduced bacteriophage DNA in the protoplasts, which ultimately led to the liberation of phage particles from the ‘transfected’ cells. Kondo and McKay (1982) used protoplasts of a lactose-negative mutant of L. lactis and incubated these with a 32–37 kb lactose plasmid. By plating the mixture on selective, lactose-containing plates and allowing time for regeneration of cell walls on the protoplasts and proper cell division, lactose-fermenting ‘transformants’ were obtained. Several other groups reported on protoplast transformation, but the frequencies with which transformants were obtained were low, not in the least because of the very strain-dependent regeneration frequencies of the protoplasts into properly walled cells (Kondo and McKay 1982, 1984; Simon, Rouault and Chopin 1986).

Two parallel developments ultimately led to the establishment of advanced genetic engineering in L. lactis (Fig. 1). First, a number of the plasmids in L. lactis, notably the smaller ones, have been used to construct gene-cloning vectors. An important driver herein was the finding that the smallest plasmid from L. lactis subsp. cremoris Wg2, pWV01, could also replicate in B. subtilis and, most importantly, the laboratory cloning horse, E. coli (Fig. 1A; Vosman and Venema 1983; Kok, van der Vossen and Venema 1984). This directly coupled the nursling genetics of L. lactis to the advanced genetic tools already available for E. coli. This connection is still a great accelerator of genetics and, more general, molecular biology research in L. lactis and the LAB alike, as the many vectors that have been constructed over the years from this plasmid (and from its almost identical counterpart from L. lactis NCDO712, pSH71) (Fig. 1B; De Vos 1987) also replicate and function in a wide variety of LAB species (and, for that matter, in many other Gram-positive bacterial species) (Kok 1991). Second, the discovery of the technique of electroporation as a way to introduce (plasmid) DNA in bacteria has really made all the difference (Harlander 1987). From initial frequencies of only a few to tens of transformants per microgram of DNA, transformation frequencies of 105–107 are now routinely attainable (but, again, this is quite strain dependent and even within the lactococci success cannot be guaranteed (Fig. 1B; Holo and Nes 1989; Papagianni, Avramidis and Filoussis 2007)). Electroporation has allowed efficiently and quite successfully employing all the tools and tricks from the ‘genetic engineering toolbox’ to the study of L. lactis, and to use them in both fundamental and applied research.

![Figure 1](image-url)
**Tools and tricks**

A full set of technologies required to properly study *L. lactis* and to ultimately modify it for application purposes includes, apart from the plasmids and plasmid transfer systems mentioned above, methods to express (foreign) genes, and to mutate the genome at will, by mutating or deleting genes, or introducing others. All of these basic technologies have been pioneered in *L. lactis* and implemented with quite some success.

**Vectors**

While in the 1970s a lot of research was conducted on discovering and describing (traits specified by) plasmids, the 1980s saw the development of the first gene cloning vectors. Over the years, several different vectors with a variety of antibiotic selection markers have been developed for gene cloning and expression in *L. lactis*. Also, many plasmids have been described in the other species of LAB, and some have been employed for cloning vector construction (Cui et al. 2015). A very versatile tool was built from pWV01 by randomly mutating a derivative of that plasmid, pGK12 (Kok, van der Vossen and Venema 1984), and selecting for a temperature-sensitive replication phenotype. One variant contained four mutations in the plasmid replication protein gene, repA. One or a combination of the nucleotide changes, all of which led to amino acid alterations, rendered the RepA protein temperature sensitive (Maguin et al. 1992; Biswas et al. 1993). In *L. lactis*, this so-called pG7-host vector replicates at 28°C but not above 37°C. A line of pG7-host plasmids have been constructed that have been used for random and targeted mutagenesis in *L. lactis* and other (lactic acid) bacteria (see below).

**Gene expression**

Plasmid pWV01 was also used to construct and employ the first promoter- and terminator screening vectors in an early effort to understand what genes in *L. lactis* looked like and how they were driven, that is, by which type(s) of promoters and regulatory elements. The first promoters cloned and analyzed in this way revealed that they conform to the canonical eubacterial promoter, with well conserved –10 and –35 consensus sequences separated by 17 nucleotides (van der Vossen, van der Lelie and Venema 1987). Indeed, sequencing of the entire genomes of several lactococcal strains has confirmed the conclusions from this initial work on transcription initiation signals in lactococci: they have only one vegetative sigma factor (σ70) that, together with RNA polymerase, uses the canonical promoter sequence (De Vos 1987; van de Guchte, Kok and Venema 1992). One immediate advance was to use the isolated promoters to construct gene expression vectors for *L. lactis*, to express genes that would not be expressed from their native promoter. Streptococcus equisimilis streptokinase, bovine prochymosine, B. subtilis neutral protease and hen egg white lysozyme were among the first heterologous enzymes that were successfully expressed in *L. lactis* (De Vos 1987; Laplace et al. 1989; van de Guchte et al. 1989, 1990). Currently, numerous proteins/enzymes have been expressed in *L. lactis*, and a great choice of gene expression tools is available. Lactococcus lactis constitutive promoters have been identified and studied in great depth, and, for instance, a collection of promoters from the *L. lactis* ssp cremoris LM0230 chromosome has been established that together cover a wide range of protein expression levels (Jeong et al. 2006). In another approach, synthetic promoter libraries were developed to tune gene expression in *L. lactis* (Jensen and Hammer 1998) and in Lactobacillus plantarum (Rud et al. 2006). In both cases, the consensus promoter sequence was kept constant while the non-consensus sequences were fully randomized. A synthetic promoter library from one promoter is thus created that in the case of the *L. lactis* library regulates gene expression over a range of 1–400 (Fig. 1C). As this range of promoter activities is covered in small increments of activity increase, the promoter library becomes very useful for fine-tuning of gene expression in *L. lactis*.

**Inducible gene expression**

An important next driver of molecular biology research in *L. lactis* was the discovery and use of inducible gene expression systems in *L. lactis* (Fig. 1D and E). Several promoters have been uncovered that respond to changes in the environment and have been used to build inducible gene expression vectors (Fig. 1E). These include the promoter P170, which is upregulated at a low pH during the transition to stationary phase; the dnp/promoter; induced by heat shock; as well as purine- or sugar-regulated promoters and bacteriophage promoters (see references in Morello et al. 2008). The best characterized and most widely used of these is the nisin-inducible promoter *P* _nisA_ (Fig. 1D). It is employed in the *L. lactis* NICE (Nlacin-Controlled Expression) system. NICE can under certain conditions and with careful amendments to the system components and induction protocol also be used in other species of LAB and in other Gram-positive bacteria (de Ruyter, Kuipers and de Vos 1996; Kleerebezem et al. 1997). The system is based on the quorum sensor two-component systems NisR and NisK that responds to the presence of the inducer nisin, an antimicrobial peptide produced by certain strains of *L. lactis* (de Ruyter, Kuipers and de Vos 1996; Kuipers et al. 1998). The nisin sensor NisK is located in the cytoplasmic membrane, interacts with nisin and in response phosphorylates the transcription regulator NisR. Phosphorylated NisR is the active molecule driving the transcription of its target genes, among which those for the biosynthesis of nisin. The nisin/NisR–P responsive promoter *P* _nisA_ is placed in the gene expression vector, while the nisRK genes, in the implementation of the system in *L. lactis*, are inserted in the chromosome. For use in the other bacteria, the two modules are present on either two plasmids or combined on one. Increasing the nisRK copy number, however, can lead to nisin-independent, leaky expression. A multitude of proteins have been (over)expressed using the NICE system in *L. lactis*. The nisin promoter has been modified to obtain differences in induction strength (Guo, Hu and Kong 2013).

To date, two additional systems for inducible gene expression have been reported that are both responsive to zinc availability. An expression system was constructed from the *P* _znC_ promoter of the *L. lactis* zit operon, which is involved in high-affinity uptake of Zn2+, and its repressor ZirR (Lull and Poquet 2004). The *P* _znC-ZirR_ system is highly inducible by starving the cells for divalent cations by growing them in a chemically defined medium, or by the addition of the divalent metal chelator EDTA. It is strongly repressed in the presence of excess Zn ions. Recently, a heterologous system has also been shown to work in *L. lactis* (Mu et al. 2013). This zinc-inducible expression system, called Zirex for Zinc-Regulated Expression, consists of the promoter *P* _znC_ from the non-food LAB *S. pneumoniae* and its regulator protein, the pneumococcal repressor SczA (Kloosterman et al. 2007). Overexpression of proteins is achieved by adding non-toxic amounts of zinc to the growth medium. The system holds promise, especially since it could be combined with the NICE system, for the expression of different proteins at different moments in the growth cycle of *L. lactis*. The recently developed agmatine-controlled expression system displays an excellent dose-response similar to that of the NICE system, and thus can be used instead of or in combination with the latter.
It employs the L. lactis-derived AguR protein that senses agmatine and activates transcription from the aguB promoter (Linares et al. 2015a).

**Fluorescent proteins**

Bright fluorescent proteins (FPs) are key components in the molecular biology toolbox. The fusion of FP genes to promoters or to genes of endogenous proteins can reveal gene and protein expression levels and (dynamics of) subcellular protein localization, all of which can help to better understand cellular processes (Fig. 1F). FPs have enabled population-wide and single-cell expression profiling through transcriptional fusions and high-throughput screening methods. The set of available FPs and fluorescence microscopy techniques has expanded considerably (Day and Davidson 2009; Yao and Carballido-López 2014). The last decade has been specifically enriched by the development of various superresolution microscopy techniques that allow researchers to peak beyond the diffraction limit of light. Many FPs were not readily available for use in every organism as major determinants for FP functionality such as pH, temperature, oxygen availability and codon usage are parameters that differ between organisms (Remington 2006). This may be illustrated by the large number of green FP variants, each of which functions optimally in a specific host (Overkamp et al. 2013). Lactococcus lactis grows best as standing cultures at around 30°C, but is not strictly anaerobic. As a consequence, it accepts mild shaking conditions during growth to increase otherwise limiting levels of oxygen, which is needed for FP maturation. Recently, improved fluorescent reporter proteins for use in B. subtilis, S. pneumoniae and L. lactis have been developed, which upgraded the detection of fluorescently tagged proteins in these organisms (Overkamp et al. 2013; Beilharz et al. 2015) and allowed for the first time to perform single-cell studies in L. lactis (see below). Employing a chemically defined medium without riboflavin, which was critical in reducing background fluorescence, enabled time-lapse microscopy as well as the ability to directly monitor fluorescence development in growing cultures (Solopova et al. 2014; van Gijtenbeek et al. 2016).

**Chromosomal knock-out and knock-in approaches**

Access to the chromosome was another issue to tackle early on, as that would allow both inserting genes and changing/mutating the chromosome for functional studies. Both random and targeted mutagenesis techniques have been developed (Mills 2001). An efficient native system of random mutation by transposon mutagenesis has been realized in L. lactis. Maguin et al. (1996) placed the lactococcal insertion sequence ISS1 on one of the conditionally replicating pG⁺ host plasmids to effectively uncouple the two more or less infrequent steps in the mutagenesis procedure, transformation and subsequent transposition. The strategy allows to first introduce the mutagenic plasmid in the cells via electroporation, and to subsequently let the ISS1 element transpose. In this way, a high frequency of random single insertions was achieved in L. lactis (1%) allowing for efficient mutagenesis. As ISS1 undergoes replicative transposition, the vector will end up in the chromosome between duplicated copies of ISS1. Excision of pG⁺ host will leave a copy of the ISS1 sequence and preserves the mutation. The system was also shown to work in other species of LAB: Enterococcus faecalis, S. thermophilus and thermophilic Lactobacilli (Russell and Klaenhammer 2001).

Various other systems have been developed over the years, using transposons from various origins (Mills 2001). In vitro transposon mutagenesis strategies have been developed on the basis of the small (~ 1.3 kb) mariner (-like) transposable element that is widely spread in animals. It encodes a single protein (the mariner transposase) and is flanked by short inverted terminal repeats. The transposase of one of these elements, Himar1, is the only protein factor required for in vitro transposition (Lampe, Churchill and Robertson 1996). Insertion takes place at TA dinucleotides and is, thus, random. A derivative of the Himar1 minitranposon, Magellan6, was used to insert randomly in the genome of the pathogenic LAB S. pneumoniae (Van Opijnen, Bodi and Camilli 2009). After the in vitro transposition, the natural competence for DNA uptake of S. pneumoniae was employed to integrate the transposed DNA in the genome of the organism, resulting in a library of mutant strains with single transposon insertions. However, rather low transformation efficiencies are a serious impediment for the implementation of this in vitro and the in vivo transposon systems in several of the other LAB species. A system for in vivo random mutagenesis overcoming at least some of the issues was developed on the basis of IS1223 from Lb. johnsonii (Licandro-Seraut et al. 2012). It comprises two plasmids, one carrying the transposase gene and the other is a suicide transposon plasmid that carries the substrate for the transposase and integrates at random in the genome. The system efficiently operates in other LAB (among which L. lactis) and has recently been used in a signature-tagged mutagenesis approach to identify genes that are necessary for establishment of Lb. casei in the rabbit gut (Licandro-Seraut et al. 2014).

The basis for targeted integration strategies that have been developed over the years for L. lactis are plasmids that cannot replicate in L. lactis. Therefore, they require another host for construction purposes and propagation prior to their application in L. lactis. Several of the vectors should, thus, be applicable in the other LAB species as well. This depends on the selection markers used on the plasmids and whether transformation efficiencies are high enough to allow for two independent processes to take place in quick succession: the uptake of the integration vector by the cell and the subsequent integration in the chromosome via homologous recombination. Various E. coli vectors have been tested in the past but they have met with various issues of stability and amplification in the chromosome, thwarting easy analysis of the obtained integrants (Leenhouts, Kok and Venema 1989). Again, pWV01 was used to circumvent these hurdles: it was bisected and the fragment containing the gene for plasmid replication (repA) was inserted in the chromosome of E. coli, B. subtilis or L. lactis (Leenhouts, Kok and Venema 1991; Leenhouts et al. 1996) (see Fig. 2A). In these so-called Rep⁺ host strains, the other part of pWV01, the fragment with the plasmid’s origin of replication (ori), can replicate as a circular DNA fragment (pORI). Endowed with an insert of L. lactis chromosomal DNA, obtained in one of the Rep⁺ hosts, these small circles of DNA cannot replicate when introduced in L. lactis. The pORI plasmids can be rescued by homologous recombination between the inserts on pORI and the chromosome. Selection is based on antibiotic resistance specified by pORI (Fig. 2A). Several pORI versions have been made employing erythromycin, chloramphenicol, spectinomycin or tetracycline resistance genes. The plasmid pG⁺ host can be used to uncouple the two steps in chromosomal integration (Fig. 2B). By first introducing the pORI vector plus its chromosomal insert in the to-be-mutated strain carrying pG⁺ host, the former will replicate by virtue of the RepA⁺ replication protein specified by pG⁺ host. After pORI has established, the temperature is raised to above what pG⁺ host can handle. By relieving the selective pressure on pG⁺ host but keeping the antibiotic that selects for pORI, the only cells that will survive are those in which the integration vector has inserted in the host chromosome. This would be the end product of single crossover (SCO).
recombination with insertion of the entire pORI plasmid. The genetic construct is stably maintained in the culture by simply selecting for the pORI-specified antibiotic resistance. In another version of the system using the couple pORI/pG+ host, double crossover (DCO) can be enforced between the chromosome and two homologous regions cloned in pORI. The E. coli lacZ gene in the integration vector is used as an easy marker to select for excisants. Entire genes or even larger operons or regions can thus be deleted. As all intervening DNA between the two targeted regions on the chromosome is removed, no foreign (plasmid) DNA is left after the DCO recombination event. Importantly, this deletion strain can be used in successive rounds of DCO deletion, allowing stacking mutations in a single strain for functional studies (Leenhouts et al. 1996). In this way, for instance, multiple peptidase mutations have been made in L. lactis, allowing studying the importance of each of them alone or in various combinations for their role in proteolysis by L. lactis and its growth in milk (Mierau et al. 1996). The system has also been used to implement a random mutagenesis protocol for L. lactis (Law et al. 1995).

A rapid and very convenient integration system was developed on the basis of L. lactis orOP, a plasmid-derived gene
specifying a dedicated orotate transporter (Defoor, Kryger and Martinussen 2007) (Fig. 2C). The transporter allows the cells to use orotate as a sole pyrimidine source, but at the same time renders the cells sensitive to the toxic analog 5-fluoro-orotate. Thus, the orof gene is employed as a counterselection marker. The original promoter driving orof led to lethal amounts of OroP in E. coli while it was not strong enough to provide toxicity when it was integrated in single copy in the chromosome of L. lactis. Both problems were solved by selecting from the synthetic promoter library described above a promoter that was applicable in E. coli and was strong enough to allow for counterselection in L. lactis. The promoter-orof cassette is present in the integration plasmid pCS1966 (Solem et al. 2008). This vector can only replicate in E. coli and can be used for integration in a bacteriophage attachment site in the lactococcal chromosome as well as for homologous recombination-driven sequence-specific SCO or DCO integration.

Recombineering

Novel genome engineering technologies that form a welcome addition to the strategies described above have recently been pioneered in LAB. The so-called recombination-mediated genetic engineering methods (or short, recombineering) normally employ bacteriophage-derived recombinase enzymes to specifically make subtle mutations in the genome or even to insert fragments containing up to thousands of basepairs (van Pijkeren and Britton 2014). The mutation(s) is present in DNA oligonucleotide that is otherwise identical to the location in the chromosome to be mutated. After introduction of the mutagenic oligonucleotide in the cell by electroporation, it is protected from exonucleolytic breakdown by the single-strand DNA(ssDNA)-binding recombinase, which is provided in trans by inducing its expression from an inducible promoter. As mutation frequencies can be very high, identification of recombinants can be done by a diversity of (colony) PCR strategies in combination with nucleotide sequencing (van Pijkeren and Britton 2014; Fig. 2D). To reach such high frequencies is not trivial, though. It requires (strain-dependent) optimization in order to avoid the mismatch repair system from reverting the mutation, and careful considerations with respect to the sequence of the oligonucleotide. The synthetic oligonucleotide should ideally withstand intracellular exonucleases, by chemical modifications, and be identical to the lagging strand (in other words, resemble an Okazaki fragment in DNA synthesis). A bacterial recombinase, RecT from a strain of L. reuteri, was shown to have a high level of activity in ssDNA recombineering. It was employed to pioneer the system in Lactococcus and L. lactis. The paper by van Pijkeren and Britton (2014) at the 11th Symposium on Lactic Acid Bacteria offers an excellent overview of the topics to be considered when trying to implement the system in the LAB of choice. Although selection is not needed to find the proper mutant strain if efficiencies are high enough, this may not (initially) be the case in any organism in which recombineering has to be set up for the first time. A smart combination with the CRISP-Cas9 system greatly enhances the chances of success (Fig. 2F). Cas9 is DNA endonuclease from S. pyogenes that utilizes so-called guide RNA to recognize, via complementarity, and cleave foreign (bacteriophage or plasmid) DNA (Lander 2016). Using this RNA-guided DNase, up to 90%–100% mutation efficiencies by recombineering could be obtained in certain cases. CRISPR-Cas9 is introduced in the cell after the actual oligonucleotide-directed recombineering mutation has taken place to effectively kill those cells that did not incorporate the mutagenic primer. This is effectuated by directing trans-activating RNA (tracrRNA)-activated Cas9 endonuclease to the site of mutation where additional (otherwise silent) mutation(s) in the oligonucleotide have removed a short 2–5 bp PAM sequence (see below) that is critical for recognition by CRISPR-Cas9 and thus its activity. Therefore, the mutant strain is protected against the endonucleaseytic attack while the chromosome in wild-type cells is cut and degraded, effectively removing them from the population (Fig. 2E). The selection through CRISPR–Cas9 holds great promise for genome editing in bacteria with low recombineering efficiencies. It would also obviate the need for time-consuming ssDNA recombineering optimization procedures. While originally the guide RNA was composed of two separate RNA molecules, the CRISPR RNA (crRNA) and tracrRNA, these could be combined into a chimeric single-guide RNA (sgRNA), which greatly simplifies Cas9 targeting (Jinek et al. 2012).

CRISPRi

Cas9 makes double-strand breaks upon base pairing between the sgRNA and target DNA. The binding specificity is determined by both sgRNA-DNA base pairing and a short DNA motif, the protospacer adjacent motif or PAM sequence (NGG), immediately next to the DNA complementary region (Marraffini and Sontheimer 2010). Variants of the Cas9 endonuclease have been made that bind to but do not cleave DNA by mutating two crucial catalytic residues in the enzyme. This ‘dead’ Cas9 (dCas9) protein can be directed to a gene of interest through the simultaneous expression of a proper sgRNA targeting the non-template strand of the gene. The resulting dCas9-sgRNA-DNA complex acts as a roadblock for RNA polymerase, repressing transcription of the target gene, either at the promoter region or during the elongation phase of transcription (Fig. 2F). By allowing mismatches in the sgRNA that weaken the interaction of dCas9 with its programmed target sequence, the extent of gene silencing can be adjusted (Bikard et al. 2013). The system can be used to target and, reversibly, shut down multiple genes at the same time.

It cannot be overemphasized how important these gene knock-out, knock-in and knock-down tools have been and still are in the entire research field that is under scrutiny in this review. Any new gene or small RNA or other genetic element of which one would like to know the biological function will have to be deleted, mutated, overexpressed or studied in situ or ectopically either on its own or in fusion with a (fluorescent) tag; all of that would not be possible without this genetic surgery technology.

SEQUENCING—A COMING OF AGE

The turn of the millennium marked a big change in the way we can look at and understand biology. This enormous leap in how and to what enormous depth we are now able to tackle biological questions has been made possible by the unprecedented speed and accuracy with which the nucleotide sequences of genomes can be determined. Although an entire chromosome had already been sequenced as early as 1977 (Smith et al. 1977), it was in 1995 that the first nucleotide sequence of the genome (chromosome) of a free-living organism was published in its entirety (Fleischmann et al. 1995). Importantly, a so-called shotgun sequencing approach was employed in which no prior knowledge of the relative positions of the fragments in the chromosome was used. This strategy, which could not have been executed without a parallel and huge increase in computing power as well as the simultaneous development of bioinformatics tools to deal with the massive amount of generated data, is now common practice.
The progress was accompanied by a dramatic decrease in the price per sequenced nucleotide of more than 100,000-fold in the last decade (https://www.genome.gov/sequencingcostsd ata/), allowing well-resourced molecular biology laboratories to now sequence their pet bacteria at affordable costs.

**Genome sequencing**

The first sequences of LAB genomes became available in the early years of the new millennium. A draft sequence of the genome of *Lactococcus lactis* IL1403 was published at the Sixth Symposium of Lactic Acid Bacteria in 1999 (Bolotin et al. 1999). The full genome was disclosed in 2001 (Bolotin et al. 2001; Makarova et al. 2006). In 2005, comprehensive reviews with comparative analyses of the genomes of lactobacilli and lactococci have been published by FEMS Microbiology Reviews (Vol. 29, No. 3) in the context of the Eighth Symposium of Lactic Acid Bacteria. Currently, the genomes of over thousands of species and strains of LAB have been sequenced, among which at least 82 (of which 15 are complete) lactococcal genomes are present in public databases. As was detailed above, *L. lactis* strains characteristically carry several different plasmid species per cell and the entire genomes (chromosome plus plasmids) of several strains have been sequenced. The nucleotide sequences of the *L. lactis* genomes allow extensive (phylogenetic) comparisons to be made and provide a wealth of information on their makeup, such as the presence of genes, their organization in operons, the incidence and localization of mobile genetic elements (IS elements, remnants of prophages and transposons) and more. They are a lucky dip for geneticists and molecular biologists alike and a lot of presents are still there to be unwrapped. Proper bioinformatics is essential in order to help revealing all the treats.

Dairy starter cultures are usually defined or complex undefined mixtures of LAB species and strains. As these microbial communities are of major importance for the quality of the final product, their composition and performance are being extensively examined by various strategies of high-throughput sequencing of all the genetic material, the metagenome, of food fermentation samples. Spatio-temporal resolution of the distribution of the microbial population during the fermentation process can thus be obtained (De Filippis, Parente and Ercolini 2016). The actual activity of the microbial community, in the form of the in vivo gene expression, can also be studied by RNA sequencing (RNA-seq), in which the so-called metatranscriptome is being uncovered (Lessard et al. 2014). By combining the genome information of a number of representative dairy strains with the metagenome and metatranscriptome data, the metabolic potential of the microbiome can be assessed (Almeida et al. 2014).

**Plasmid sequencing**

Apart from the chromosomes of lactococci, a lot of effort has gone into the sequencing of the plasmid content of these organisms, because they specify some of the industrially most important functions. Well over 80 completely sequenced *L. lactis* plasmids are available in the public domain. Van Sinderen and coworkers have recently presented an extensive survey of these plasmids, representing the first glimpse of the lactococcal ‘plasmidome’, the overall plasmid content in a given species or environment (Walker 2012). Using bioinformatics approaches, it was demonstrated that there is still a lot to be discovered as this plasmidome is actively evolving. Conjugation and transduction, the transfer of DNA by bacteriophages through the error-neous packaging of DNA other than the phage genome, could be the cause of constant influx of new genetic material in this species. Conjugation has been shown to disregard LAB species borders. In fact, the first paper on conjugation in LAB was the transfer of pAMβ1, a broad host range antibiotic resistance plasmid from *Enterococcus faecalis* into, among others, *Lactobacillus casei* (Gibson et al. 1979). In addition, plasmids like pAMβ1 have been used to mobilize and transfer other plasmids, some carrying important industrial traits, between different strains of *L. lactis* (Gasson and Fitzgerald 1994). Interestingly, it has recently been shown that transduction can also bypass the species border in Gram-positive bacteria. Ammann et al. (2008) described the transfer of a 9.2-kb plasmid from *Streptococcus thermophilus* to *L. lactis* by three different plages from the former species. The high similarity between genes of *L. lactis* and *S. thermophilus* bacteriophages supports the notion that horizontal gene transfer does occur between these species and, also, that recombination between their phages apparently takes place, either between an infecting phage and a resident prophage (*L. lactis* strains contain many (in)complete prophages in their chromosomes) or between two simultaneously infecting phages.

**Transcriptomics and more**

The genome sequences have enabled setting up transcriptomics, a technology used to uncover expression profiles of all genes in a specific cell type or in (sub)populations of cells at a given moment in time. Thus, gene regulation could be studied much more comprehensively. The availability of the genome sequences has also allowed successfully applying proteomics and metabolomics approaches. Whole-genome transcription analysis was made possible again by advances in technology (miniarization, robotization) and proper bioinformatics to deal with the large data sets obtained from transcriptomics studies (eg, http://genome2d.molgenrug.nl). DNA macro- (filter) and micro- (glass slides) arrays carrying many or all known genes of the *L. lactis* strain under study have been and still are being used. The technology is within reach of most well-equipped molecular biology groups now, since the costs of commercial DNA microarrays have gone down considerably. A leap in transcriptomics research possibilities is offered by RNA-seq, which is fast and reliable, and in addition offers single nucleotide resolution, a high dynamic range, less noisy signals and does not suffer from cross hybridization. Therefore, it quickly replaces DNA microarrays as the preferred method for transcriptome analysis in research (Wang, Gerstein and Snyder 2009).

The responses of all the genes in the genome of *L. lactis* on (industrial) stressors such as acidification, temperature up- and down shifts, oxygen, salt, antimicrobials and antimicrobial peptides and many more have been extensively scrutinized over the past decade (for recent extensive reviews, see Papadimitriou et al. 2016). Also, many studies have examined the adaptation of the organism to shorter or longer duration of the stress conditions; have adopted more elaborate schemes of chronotranscriptomics, in which a situation was sampled and analyzed over a course of time; or have examined strain performance in a spatio-temporal way in actual fermentation processes. These studies have led to a wealth of novel information on the functioning of genes, on gene regulatory pathways and on responses of *L. lactis* to a variety of growth conditions and stimuli. A full account of all original literature dealing with employing transcriptomics to the study of the life and sufferings of *L. lactis* is beyond the scope of this review. Below, only carbon and nitrogen metabolism, the two major pathways from the
perspective of milk fermentation, their main regulators, their interplay and the stringent response (SR) will be detailed.

**GENE REGULATION RESEARCH V2.0(00)**

By automated annotation and manual curation, the accuracy of the lactococcal genome sequences has improved over the years. Among others, this has allowed obtaining quite a complete view of the transcription regulation potential in the organism. It has to be noted that the NCBI bacterial database has changed to the RefSeq system (O’Leary et al. 2016) at the end of 2014. Locus tags have been changed, removed or added, while many gene names have been deleted from the original annotation, seriously hampering the bioinformatic analysis of –omics data. A tool is available at http://genome2d.molgenrug.nl/to interconvert new and old locus tags of all complete bacterial genomes for ‘90% of the locus tags.

*Lactococcus lactis* subsp. cremoris MG1363 contains 154 genes annotated as (putatively) specifying transcription regulators (Wegmann et al. 2007). Forty-one have been studied to a more or lesser extent in strain MG1363 and/or in *L. lactis* subsp. lactis strain IL1403. More than one-third of these regulators show high homology to transcription regulators with known function in other bacteria while 68 are supposed transcription regulators, as they carry conserved protein motifs that could be involved in DNA binding. Seven genes are part of an operon specifying a two-component system (for an overview, see Table 1).

The identity and functionality of transcription regulators has been studied in the past by employing several of the tools and techniques mentioned above, that is, by knocking out the gene of interest and studying the effect on gene expression of the target gene(s), if known. Some of the regulators work at a close distance from where their genes are situated in the genome and control the expression of the gene(s) ‘next door’. Quite detailed knowledge has been achieved in doing so, with various footprinting techniques permitting pinpointing the actual regulator binding sites and understanding the molecular details of protein–DNA interaction and the mechanism of gene activation or repression. Quickly after its introduction, DNA microarray technology was used to study gene regulation in *L. lactis* more deeply, and to probe interactions between regulators. The summary in Table 1 lists all the regulators that have been studied to date in the two

<table>
<thead>
<tr>
<th>TF</th>
<th>Strain</th>
<th>Description/target</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>AguR</td>
<td>IL1403</td>
<td>Activator of the putrescine biosynthesis operon</td>
<td>Linares et al. (2015b)</td>
</tr>
<tr>
<td>AhrC</td>
<td>MG1363</td>
<td>Arginine sensor and ArgR co-regulator</td>
<td>Larsen et al. (2004); Larsen, Kok and Kuipers (2005)</td>
</tr>
<tr>
<td>ArgR</td>
<td>MG1363</td>
<td>Regulator of arginine metabolism</td>
<td>Larsen et al. (2004); Larsen, Kok and Kuipers (2005)</td>
</tr>
<tr>
<td>BglR</td>
<td>IL1403</td>
<td>Beta-glucoside utilization, transcriptional antiterminator</td>
<td>Bardowski, Ehrlich and Chopin (1994)</td>
</tr>
<tr>
<td>BusR</td>
<td>IL1403</td>
<td>Repressor of the osmoprotectant uptake system</td>
<td>Romeo et al. (2003); Romeo, Bouvier and Gutierrez (2007)</td>
</tr>
<tr>
<td>CcpA</td>
<td>MG1363</td>
<td>Catabolite control protein A</td>
<td>Zomer et al. (2007)</td>
</tr>
<tr>
<td>ClaR</td>
<td>IL1403</td>
<td>Cellobiose and lactose metabolism</td>
<td>Aleksandrzak-Piekarczyk et al. (2015)</td>
</tr>
<tr>
<td>CodY</td>
<td>MG1363</td>
<td>Global nitrogen regulator</td>
<td>den Hengst et al. (2005b); den Hengst et al. (2006)</td>
</tr>
<tr>
<td>ComX</td>
<td>IL1403</td>
<td>Competence regulator</td>
<td>Wydau et al. (2006)</td>
</tr>
<tr>
<td>CopR</td>
<td>IL1403</td>
<td>Copper metabolism and resistance</td>
<td>Magnani et al. (2008)</td>
</tr>
<tr>
<td>CtsR</td>
<td>MG1363</td>
<td>Protein thermosensor, master regulator of protein quality control</td>
<td>Varmanen, Ingmer and Vogensen (2000); Varmanen et al. (2003)</td>
</tr>
<tr>
<td>FabT</td>
<td>MG1363</td>
<td>Fatty acid biosynthesis</td>
<td>Eckhardt et al. (2013)</td>
</tr>
<tr>
<td>FhuR</td>
<td>IL1403</td>
<td>Activator of metC-cysK operon</td>
<td>Fernandez et al. (2002)</td>
</tr>
<tr>
<td>FlpA</td>
<td>MG1363</td>
<td>FNR-like protein A</td>
<td>Akyol and Shearman (2008)</td>
</tr>
<tr>
<td>FlpB</td>
<td>MG1363</td>
<td>FNR-like protein B</td>
<td>Akyol and Shearman (2008)</td>
</tr>
<tr>
<td>FruR</td>
<td>IL1403</td>
<td>Repressor of fructose operon</td>
<td>Barriere et al. (2005)</td>
</tr>
<tr>
<td>GadR</td>
<td>IL1403</td>
<td>Activator of gadCB; acid stress response</td>
<td>Sanders, Venema and Kok (1997); Sanders et al. (1998)</td>
</tr>
<tr>
<td>HdiR</td>
<td>MG1363</td>
<td>DNA damage and heat-stress</td>
<td>Savijoki et al. (2003)</td>
</tr>
<tr>
<td>HtrR</td>
<td>IL1403</td>
<td>Repressor of htrRRA, specifying a heme efflux pump</td>
<td>Lechardeur et al. (2012)</td>
</tr>
<tr>
<td>LacI</td>
<td>MG1363</td>
<td>Cellobiose metabolism</td>
<td>Solopova, in preparation</td>
</tr>
<tr>
<td>LlrA</td>
<td>MG1363</td>
<td>Two-component system; arginine; acid stress</td>
<td>O’Connell-Motherway et al. (2000)</td>
</tr>
<tr>
<td>LlrB</td>
<td>MG1363</td>
<td>Two-component system</td>
<td>O’Connell-Motherway et al. (2000)</td>
</tr>
<tr>
<td>LlrC</td>
<td>MG1363</td>
<td>Two-component system; acid stress resistance</td>
<td>O’Connell-Motherway et al. (2000)</td>
</tr>
<tr>
<td>LlrE</td>
<td>MG1363</td>
<td>Two-component system; phosphatase expression</td>
<td>O’Connell-Motherway et al. (2000)</td>
</tr>
<tr>
<td>LlrF</td>
<td>MG1363</td>
<td>Two-component system; oxidative stress</td>
<td>O’Connell-Motherway et al. (2000)</td>
</tr>
<tr>
<td>LmrR</td>
<td>MG1363</td>
<td>Repressor of lmrCD, specifying a multidrug transporter</td>
<td>Agustiandari et al. (2008)</td>
</tr>
<tr>
<td>MalR</td>
<td>IL1403</td>
<td>Maltose operon repressor</td>
<td>Andersson and Radstrom (2002)</td>
</tr>
<tr>
<td>PhoU</td>
<td>IL1403</td>
<td>Phosphate uptake</td>
<td>Cesselin et al. (2009)</td>
</tr>
<tr>
<td>PurR</td>
<td>IL1403</td>
<td>pur operon repressor</td>
<td>Kilstrup and Martinussen (1998)</td>
</tr>
<tr>
<td>PyrR</td>
<td>MG1363</td>
<td>pyr operon attenuation</td>
<td>Martinussen et al. (2001)</td>
</tr>
<tr>
<td>RcfB</td>
<td>IL1403</td>
<td>Acid adaptation</td>
<td>Madsen et al. (2005)</td>
</tr>
<tr>
<td>SpxB</td>
<td>IL1403</td>
<td>Control of O-acetylation of peptidoglycan</td>
<td>Veiga et al. (2007)</td>
</tr>
<tr>
<td>ZitR</td>
<td>IL1403</td>
<td>Repressor of zitRSQP operon</td>
<td>Lull and Poquet (2004)</td>
</tr>
</tbody>
</table>
laboratory strains of *L. lactis*, MG1363 and IL1403. Some of the regulators, such as CodY and CcpA, are pleiotropic and therefore have large and partly overlapping regulons, while others are more dedicated.

**Regulation of nitrogen metabolism**

*Lactococcus lactis*, being an amino acid auxotroph, has evolved an intricate system to degrade and utilize milk proteins, consisting of an extracellular proteinase, various amino acid and peptide uptake systems and an army of intracellular peptidases. Several regulators are involved in the regulation of nitrogen metabolism in *L. lactis* (see Table 1) of which CodY is the most studied and has the largest regulon (Guédon et al. 2001a,b; Petranovic et al. 2004; den Hengst et al. 2005a,b). Other regulators involved are more specific, with smaller regulons, such as GlnR and the couple ArgR/AhrC. Over 30 genes of *L. lactis* are under control of the pleiotropic regulator CodY (Guedon et al. 2005; den Hengst et al. 2005b). CodY acts as a transcriptional repressor of these genes when the cells are growing under nitrogen-rich conditions, an activity that is relieved when nitrogen becomes limiting. Using a clean transcriptomics/bioinformatics approach on a clean knock-out *L. lactis* codY mutant, a CodY-binding box could be identified upstream of the CodY target genes (den Hengst et al. 2005b). CodY regulates its own production as a CodY box is also present upstream of its own gene. Branched-chain amino acids act as CodY co-repressors and stimulate binding of CodY to its binding site (Guédon et al. 2001b; Petranovic et al. 2004; den Hengst et al. 2005a). The homologous nutritional repressor CodY from *B. subtilis*, in addition to interacting with branched-chain amino acids, also interacts with GTP, a feature not seen with *L. lactis* CodY (Ratnayake-Lecamwasam et al. 2001; Petranovic et al. 2004). The *L. lactis* CodY regulon contains a number of genes from the Krebs cycle, indicating that a link exists between nitrogen and carbon metabolism. Conversely, the global regulator of carbon metabolism in *L. lactis*, CcpA regulates a number of genes involved in nitrogen metabolism (see below).

DNA microarray analyses of cells growing under nitrogen-poor conditions revealed that the genes coding the glutamine synthetase GlnA and the putative ammonium transporter and sensor AmtB-GlnK were highly derepressed in an *L. lactis* mutant in which the glnR gene had been removed by DCO recombinatation. Also, the expression of the gene encoding the glutamine synthetase GlnA and the putative ammonium transporter and sensor AmtB-GlnK were highly derepressed in an *L. lactis* glnR mutant. The fact that the catalytic genes were upregulated in the *argRahrC* double mutant showed that AhrC is not necessary for the activation of *arc* and, at the same time, that ArgR does play a role in arc regulation. A so-called ARG box with high similarity to the ARG boxes in other organisms is present in the promoter regions of the three arginine biosynthesis operons, while ARG box half-sites are located upstream of the catalobic genes cluster. Electric mobility shift assays with purified ArgR and AhrC proteins showed that the latter interacts with a T-loop in AmtB that then inserts into the exit of the channel and probably blocks transport (Conroy et al. 2007). *Escherichia coli* GlnK is uridylylated at its Tyr53 residue under N-limitation (low intracellular glutamine levels). As a consequence, it cannot bind to AmtB and transport can take place (Javelle et al. 2004). It has been proposed that GlnK fine-tunes the active transport of ammonium by AmtB (Boogerd et al. 2011) and prevents a futile cycle of (NH₄)₄ being transported into the cell and a passive outward-directed flow of NH₃. Although it is not known how AmtB-GlnK operates in *L. lactis*, the composition and responses to nitrogen availability of the GlnK regulon in this organism are in agreement with a role of GlnR in tightly controlling the incorporation of ammonium into central metabolism via glutamine synthetase (GlnRA) and through glutamine/glutamate (GlnPQ) and ammonium transport (AmtB-GlnK).

The amino acids arginine, glutamine and glutamate are closely connected. The *de novo* biosynthesis of arginine starts with glutamate while glutamine is the precursor for the energy-rich compound carbamoylphosphate required to make arginine. Carbamoylphosphate is also produced in the arginine catabolic route, the arginine deiminase (ADI) pathway, and is a precursor in pyrimidine metabolism. The *L. lactis* ADI pathway is encoded by one large gene cluster, *arcABD1C2T2D2yvaD*, from which various transcripts are being produced (Budin-Verneuil et al. 2006; Larsen et al. 2008; van der Meulen, et al. 2016). ArcA, ArcB and ArcC perform the three enzymatic steps; ArcD1 is the main transporter of the pathway, exchanging L-arginine for L-ornithine, while ArcD2 has been proposed to function as an L-arginine/L-alanine exchanger in a pathway together with ArcT (Noens et al. 2015). Two dedicated regulators are required for the arginine-dependent repression of arginine metabolism in *L. lactis* and *Lactobacillus plantarum* (Larsen et al. 2004; Nicoloff et al. 2004). A greatly increased expression of the three arginine biosynthetic operons *argCDFB*, *argGH* and *gltS-argE* was observed when either one or both of the regulators in *L. lactis*, ArgR and AhrC, were deleted. The two regulators have different functions in the regulation of arginine breakdown: deletion of AhrC resulted in a strong downregulation of *arc*, while the operon was not affected in an *L. lactis* argR mutant. The fact that the catalobic genes were upregulated in the *argRahrC* double mutant showed that AhrC is not necessary for the activation of *arc* and, at the same time, that ArgR does play a role in arc regulation. A so-called ARG box with high similarity to the ARG boxes in other organisms is present in the promoter regions of the three arginine biosynthesis operons, while ARG box half-sites are located upstream of the catalobic genes cluster. Electric mobility shift assays with purified ArgR and AhrC proteins showed that the latter had no DNA-binding activity (Larsen et al. 2006). ArgR bound to a probe carrying the ArgR boxes from the *argC* promoter and to a ParaC promoter fragment. The binding of ArgR was not arginine sensitive: arginine-dependent binding to both probes was only obtained when both regulators were present.

A model has been presented in which AhrC is considered to be the main arginine sensor, while ArgR acts as the DNA-binding protein in the arginine-dependent regulation of arginine metabolism in *L. lactis* (Larsen, Kok and Kuipers 2005). It is based on the data obtained in *L. lactis* and on structural knowledge on ArgR(AhrC)-type arginine regulators from other bacterial species.
The regulatory control of carbohydrate metabolism in lactococci is primarily governed by the bifunctional transcriptional regulator CcpA. CcpA plays a central role in the carbon catabolic repression (CCR) system. CCR is a global regulatory mechanism that helps bacteria adapt to changing carbon availability by repressing genes that would be unnecessary under certain growth conditions. This is essential for bacteria to maintain optimal fitness in diverse environments.

The CcpA regulon comprises a large number of genes involved in the expression of metabolic pathways that utilize sugars other than glucose. CcpA activation is triggered by the phosphorylation state of the bifunctional protein HPr, which is a central regulator of sugar utilization. CcpA activates the expression of genes involved in the utilization of carbohydrates through the phosphorylation of HPr, leading to the derepression of target genes.

The CcpA regulon is induced under conditions of carbon starvation, particularly when bacteria are grown on a carbon source that is less preferred than glucose. This regulatory mechanism allows bacteria to efficiently respond to changes in their environment, ensuring optimal resource utilization and growth.

In summary, the carbohydrate metabolism in lactococci is intricately regulated by CcpA, enabling bacteria to adapt to varying carbon conditions and maintain their fitness in diverse environments.
but also carbon source dependent (Crow and Thomas 1982; Poolman, Driessen and Konings 1987). Indeed, CcpA was shown to repress the ADI pathway genes during the exponential phase of growth, while arc was transcribed in the transition phase. Multiple copies of the cre element were detectable in the promoter region of the arc operon (Zomer et al. 2007).

In L. lactis, CCR does not only play an important role in sugar uptake but also in regulation of central carbon metabolism. The glycolytic las operon (pgk-pyk-ldh), encoding the enzymes phosphofructokinase, pyruvate kinase and lactate dehydrogenase, is activated by binding of CcpA-HPr Ser46-P. By contrast, transcription of the genes for acetic kinase and pyruvate dehydrogenase, which belong to the heteroacetification fermentation branch (see below), is repressed by the complex (Zomer et al. 2007).

One of the L. lactis genes recently shown to be under CCR is that of the abundant non-coding RNA 6S (van der Meulen et al. 2016). The gene for RNA 6S, LLMGnc_004, is highly conserved in prokaryotes, located immediately downstream of the CcpA-controlled mtlARFD operon mentioned above and contains a cre element just upstream of the −35 region of its promoter. The 6S gene is approximately 3-fold upregulated in the L. lactis ccpA mutant. In Escherichia coli, 6S RNA accumulates maximally in the late stages of stationary phase, plays a role in balancing the nutrient usage during extended stationary phase and helps saving energy for long-term survival. RNA 6S binds σ70- RNA polymerase, inhibiting transcription from σ70- driven promoters. Upon nutrient reintroduction, 6S RNA dissociates from RNA polymerase and is degraded. Operons that are repressed during the stationary phase can then be transcribed again (Wasserman 2007). Lactococcus lactis 6S RNA becomes abundant in glucose-grown cells only during the stationary phase. 6S RNA is expressed already in the exponential growth phase when L. lactis is growing on cellobiose or galactose, but not on fructose (van der Meulen et al. 2016). Thus, 6S might fine-tune the CcpA regulon, coming into play when CcpA repression is relieved eg, during stationary phase and/or growth on alternative carbon sources. Although the full extent of the regulon of 6S RNA and the functions of the regulon members in L. lactis remain to be elucidated, its expression should help increase cellular fitness under suboptimal conditions of growth.

Besides being subject to global regulation, catabolic operons are usually under control of a specific transcriptional protein regulator (see Table 2), ensuring that costly transporters and enzymes are not synthesized unnecessarily. A regulatory non-coding RNA was recently shown to be involved in the control of the L. lactis cryptic sugar utilization operon llmg_0957(rpe2)-llmg_0963 (see below) (van der Meulen et al. 2016). LLMGnc_147 is an sRNA of 102 nucleotides that is located immediately downstream of the transcriptional activator gene tenA. Short-term overexpression of LLMGnc_147 led to a highly increased expression (23 to 60-fold) of the llmg_0957(rpe2)-llmg_0963 cluster, which specifies the putative PTS IIC component Llmg_0963, enabling the cell to import galactose. LLMGnc_147 likely stabilizes the gene cluster transcript(s). The expression of the sRNA itself is controlled by galactose and cellobiose, possibly via a cellobiose-specific transcriptional activator. In fact, a candidate gene for such a regulator is present in the operon itself: lllmg_0962 specifies a putative AraC transcriptional regulator. Both the LLMGnc_147 sRNA gene and the lllmg_0957-llmg_0963 cluster seem to be also under control of CCR as their promoter regions possess credible cre sites.

Curiously, the lllmg_0957(rpe2)-llmg_0963 gene cluster, which on the basis of homology presumably was originally dedicated to the transport and metabolism of the plant sugars xylose, ribulose and/or cellobiose, is only present in a few dairy L. lactis subsp. cremoris strains. Lactococcus lactis can import and utilize a limited number of carbohydrates and sugar alcohols as a result of reductive evolution and adaptation to the milk environment (Siezen et al. 2011; Price et al. 2012). Although L. lactis has lost many plant niche-specific genes during its adaptation to milk, the ability of Lllmg_0963 to import galactose has apparently saved this cluster from complete silencing or loss. When L. lactis is exposed to a plant sugar-rich environment, new hybrid transporters and metabolic pathways can become operative. In L. lactis II1403, proteins encoded by two cellobiose-induced operons form a hybrid transporter CelB-PtcAB which can

Table 2. Specific regulators of sugar import and catabolism in L. lactis.

<table>
<thead>
<tr>
<th>Regulator</th>
<th>Family</th>
<th>Carbohydrate-specific regulon</th>
<th>Imported sugar (inducer molecule)</th>
<th>L. lactis strain</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BglIR</td>
<td>Bgl/GSacY antiterminator</td>
<td>pthA</td>
<td>Arbutin, esculin, salicin</td>
<td>IL1403</td>
<td>Bardowski, Ehrlich and Chopin (1994)</td>
</tr>
<tr>
<td>ClaR</td>
<td>RpiR</td>
<td>bgIS, celB</td>
<td>Cellobiose, lactose</td>
<td>IL1403</td>
<td>Aleksandrzak-Piekarczyk et al. (2015)</td>
</tr>
<tr>
<td>LacR</td>
<td>DeoR</td>
<td>lacABCDFEGX</td>
<td>Lactose (tagatose-6P)</td>
<td>NCD0712 (on lactose-protease plasmid)</td>
<td>van Rooijen, Gasson and de Vos (1992)</td>
</tr>
<tr>
<td>SacR</td>
<td>Lacl/GalR</td>
<td>sacABK</td>
<td>Sucrose (sucrose-6P)</td>
<td>N29800 (on transposon Tn5276)</td>
<td>Luesink et al. (1999)</td>
</tr>
<tr>
<td>FruR</td>
<td>DeoR</td>
<td>fruAC</td>
<td>Fructose (fructose-1P) Xylose</td>
<td>IL1403 IO-1 B-4449</td>
<td>Barriere et al. (2005) Erlandson et al. (2000)</td>
</tr>
<tr>
<td>MalIR</td>
<td>Lacl/GalR</td>
<td>malEFG</td>
<td>Maltose</td>
<td>19435</td>
<td>van der Meulen, et al. (2016)</td>
</tr>
<tr>
<td>Llmg_1239</td>
<td>Lacl/sRNA</td>
<td>llmg_1240-1244</td>
<td>Cellobiose Galactose, Cellobiose (?)</td>
<td>MG1563 MG1563</td>
<td></td>
</tr>
</tbody>
</table>
also import lactose (Aleksandrzak-Piekarczyk et al. 2011). A promoter-up mutation upstream of a homologous but silent cell cluster in *L. lactis* MG1363 allowed the mutant strain to utilize lactose by virtue of the expression of the cellbiose-specific PTS IIC component, CelB (Solopova et al. 2012). The mutant utilizes a novel pathway for lactose degradation, involving the hybrid transporter CelB-PtCAB, phospho-β-glucosidases BglS and AscB, lactose 6-phosphate dephosphorylase and the Leolir pathway enzymes. A different way by which silent sugar utilization operons can be awakened was shown for a cellbiose, beta-glucoside- or lichenan-catabolizing population containing yet another PTS IIC component gene. Insertion of an IS element at different positions in a gene coding for a LacI-type transcriptional repressor located upstream of the cluster apparently led to the activation of the otherwise silent operon (Solopova, manuscript in preparation).

**CCR from the perspective of single cells**

For more than 75 years diauxie, the fact that a bacterial population growing on a mixture of two sugars will first consume the preferred sugar and only then switch to the less-preferred one (Monod 1949) is the essential textbook example of strict CCR and uniform cell behavior. A lag phase of no growth separates the two growth phases and is needed for the cells to synthesize the enzymes required for the metabolism of the second sugar. However, microorganisms live in constantly changing environments, which often creates a selective pressure that does not favor strict CCR (Siegal 2015). Populations of microorganism often exhibit phenotypic heterogeneity, whereby the various phenotypes allow fine-tuning of adaptation at the population level, while the regulatory mechanisms fail to do so at the single-cell level. Several recent studies revisited diauxic growth at the single-cell level (Siegal 2015). They show that there are many different ways to respond to a shift in available carbon sources and that a lot of heterogeneity exists among and within populations during this process.

An isogenic population of *L. lactis* differentiates into two metabolic phenotypes at the switch point during the diauxic shift from glucose to cellbiose or lactose utilization (Fig. 3A) (Solopova et al. 2014). The lag phase is not a result of a temporal growth arrest of the whole population, but is observed because only a subpopulation is fit enough to partake in the second outgrowth. The non-participating (non-growing) cells are viable, probably because of induction of the SR (see below), and are able to cope better when a third carbon source is administered. Diauxic shift was not observed in an *L. lactis* ccpA culture: all cells immediately also start consuming cellbiose. The observed phenotypic heterogeneity was proposed to result from differences in the capability of cells to deal with the time constraint between CCR relief and the activation of the SR. The metabolic state of individual cells determines whether SR is induced or whether they can make the switch to cellbiose consumption. The failure to fine-tune two global regulatory mechanisms, CCR and SR, at the single-cell level seems very curious from an evolutionary perspective. The two subpopulations were not equally fit when a third sugar became available. Thus, the phenotypic heterogeneity could be the result of natural selection and represent a bet-hedging strategy, as was underpinned by an evolutionary model based on the growth rates of both cell types. Apparently, when future conditions are unpredictable, a genotype generating a set of phenotypes is better adapted to multiple conditions at the same time (Solopova et al. 2014; Grimbergen et al. 2015).

This and other examples illustrate that only long-term cultivation in a constant environment favors strict regulation, while growing under continuously varying conditions, a relaxed CCR and bet hedging allow the cells to adapt faster to the environmental changes (New et al. 2014; Siegal 2015).

**Stringent response**

The SR is a protective mechanism that inhibits major energy-consuming processes and stimulates certain anabolic pathways. It is induced as soon as bacterial cells encounter a nutrient limitation or several other stresses (Mechold and Malke 1997). The SR factors (the RelA family) produce the phosphorylated purine-derived alarmones (p)ppGpp in response to the presence of uncharged tRNA molecules (Potrykus and Cashel 2008). In *E. coli*, the alarmone and its cofactor DksA directly interact with RNA polymerase, changing its specificity for certain promoters (Ross et al. 2013). In other bacteria, (p)ppGpp allosterically inhibit enzymes or compete with GTP for binding to proteins (Kanjee, Ogata and Houry 2012). *Bacillus subtilis* RNA polymerase does not bind (p)ppGpp. In this organism, the alarmones inhibit the first enzyme of the GTP biosynthesis pathway, inosine monophosphate (IMP) dehydrogenase (GuaB). IMP then accumulates in the cytosol and stimulates ATP synthesis. Hypoxantine phosphorybosyltransferase Hpt and guanylate kinase Gmk are

---

**Figure 3.** Illuminating *L. lactis* at the single-cell level. (A) Fluorescence microscopy analysis of diauxie of *L. lactis* MG1363 carrying a Pcel-gfp transcriptional fusion. The strain is growing in chemically defined medium with 0.1% glucose and 1% cellbiose. After glucose depletion, only a fraction of the cells start using cellbiose and become fluorescent. An overlay of green fluorescence and phase-contrast images is shown. (B) Deconvolved image of fluorescence microscopy of *L. lactis* cells in which transcripts of well expressed (left cell) or poorly produced (cells on the right) membrane proteins have been visualized (in green). The membrane is stained with Nile red (red) and the chromosomal DNA with DAPI (blue).
also inhibited by (p)ppGpp. Thus, the level of GTP decreases during SR while that of ATP increases. The identity of first base of the transcript (either a G or an A) plays a crucial role in the activity of the promoter during SR in B. subtilis (Krásný et al. 2008). Genes downregulated during SR such as those for rRNA, translation initiation and transcription termination factors, and ribosomal proteins, often possess a G as the +1 nucleotide. The transcripts of the upregulated genes, namely those for amino acid synthesis enzymes and sporulation proteins, usually start with an A (Krásný et al. 2008). It was shown that decrease of GTP, GDP or GMP concentration in the cytosol is a strong signal for multistress-resistant phenotype induction in L. lactis (Ryssel et al. 2014).

Lactococcus lactis RelA exhibits both (p)ppGpp synthesis and degradation activities (Rallu et al. 2000). Studies with a relA mutant, which accumulated (p)ppGpp under acidic conditions, revealed that the alarmones activate transcription of the important glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase GAPDH, although the overall glycolytic flux was strongly reduced (Rallu et al. 2000; Mercade, Coaigin-Bousquet and Loubière 2006). Lactococcus lactis relA exhibited a 1000-fold increased survival rate (Mercade, Coaigin-Bousquet and Loubière 2006) compared to the wild-type strain under acidic conditions. Thus, also in L. lactis growth rate and survival are opposite parameters (Rallu et al. 2000; Mercade, Coaigin-Bousquet and Loubière 2006; Ryssel et al. 2014). Since activation of SR induces branched-chain amino acid synthesis and survival during the stationary phase, interaction exists between SR and CodY (Eymann et al. 2002; Krásný et al. 2008). As induction of SR instantly reduces the pool of GTP, the co-repressor of B. subtilis CodY, repression of genes by CodY is consequently relieved. Lactococcus lactis CodY is unresponsive to GTP (Petrnovic et al. 2004; den Hengst et al. 2005b). However, it was shown in S. mutans that regulation of the ilv-leu pathway depends on a basal level of (p)ppGpp and CodY (Lemos et al. 2008). Thus, although a link exists between SR induction and a relief of CodY repression in LAB, the exact mechanism of how this works still needs to be elucidated. Overall, the interplay between the CcpA, LMGnc, CodY and other regulatory factors results in a dynamic, well-organized balancing of cellular processes between growth and survival.

Regulation with RNAs in lactic acid bacteria

Although conventional DNA microarrays can give a lot of information on genome-wide gene transcription, one major drawback is that it builds on available know-how, with respect to the identity and location of genes in the genome. Unknown genes or other regions that produce RNA that are not represented on the microarray slides are ‘invisible’ to the experimenter. High-density tiling arrays can be used to get a rough indication of transcript boundaries and identify unknown RNA species such as antisense RNAs (asRNAs) and small non-coding regulatory RNAs (sRNAs) (Nicolas et al. 2012). RNA-seq is quickly becoming the preferred technology for transcriptome analysis. Specific enrichment of 5′- or 3′-ends of total RNA can be achieved, allowing predicting TSS detection, identification of novel RNAs and operon structures. Novel technologies are being developed such as, for instance, ribosomal profiling by which actively translated mRNAs can be recognized (Ingolia et al. 2009).

Genome-wide RNA-seq analyses have allowed adding numerous novel elements to previously annotated genomes, such as sRNAs, asRNAs and riboswitches. It is now clear that bacteria use a wide diversity of RNA molecules to regulate the expression of their genes (Wagner and Romby 2015). Particularly in the last decade, hundreds of sRNA genes have been identified in a variety of bacterial genomes. sRNAs now greatly outnumber the protein regulators specified by bacterial genomes, not only in absolute number, but also in diversity of functions (Narbonhaus and Vogel 2009). sRNAs are small transcripts that are very heterogeneous in size, ranging from around 50 to up to ~350 nucleotides, and, generally, do not encode a protein. sRNA genes are typically located in intergenic regions, are controlled by orphan promoters and have Rho-independent terminators. sRNAs act post-transcriptionally in the regulation of mRNAs by base pairing. They can distinguish their mRNA targets by an astounding accuracy, as is shown by the Salmonella Typhymurium sRNA SgrS, which can discriminate a target mRNA based on a single hydrogen bond (Papenfort et al. 2012). This surgical precision is a reflection of the extreme fine-tuning that can take place in bacterial gene regulation. It also poses a huge challenge to in silico sRNA target prediction (Wright et al. 2013).

Bacterial sRNAs mostly target mRNAs by positively or negatively influencing mRNA stability or translation. sRNAs can regulate a range of different biological processes such as stress response (Hoe et al. 2013), virulence (Grijenland et al. 2010) and metabolism (Sonleitner and Haas 2011). Some sRNAs can regulate multiple mRNAs, while one mRNA can be regulated by various sRNAs (Storz, Vogel and Wassarman 2011). In some cases, sRNAs perform a crucial role in overcoming a potentially lethal stress, while in other cases fine-tuning by sRNAs enables homeostasis (Masse and Gottesman 2002). A number of sRNAs influence the function of specific proteins (Barrick et al. 2005). asRNAs derive from the DNA strand complementary to that of protein coding genes and therefore can make perfect base pairing with their target mRNA. In addition to affecting RNA stability and translation, certain asRNAs operate by transcription interference and influencing transcription termination (Georg and Hess 2011). Despite the obvious targets of asRNAs, it cannot be excluded that they also regulate other mRNAs in trans. To date, only a relatively small number of chromosomally encoded asRNAs have been characterized and assigned a specific function (Wagner and Romby 2015). They control plasmid replication and plasmid copy numbers and some are expressed from bacteriophage genomes and transposons (Thomason and Storz 2010). Hundreds of sRNAs have been studied, particularly in Gram-negative (pathogenic) bacteria from the genus Enterobacteriaceae. Most of the work on regulatory RNAs has been performed in pathogenic LAB, in Streptococcal species such as S. pyogenes and S. pneumoniae (Brantl and Brückner 2014). Strand-specific RNA-seq in Lb. delbrueckii revealed one sRNA gene, located between the genes LBU_0613 and LBU_0612 but a functional study (Hosoda et al. 2016) revealed one sRNA gene, located between the genes LBU_0613 and LBU_0612 but a functional study (Hosoda et al. 2016) revealed one sRNA gene, located between the genes LBU_0613 and LBU_0612 but a functional study (Hosoda et al. 2016). Several of the newly discovered sRNAs have been studied to some extent, e.g., the above-described S6 RNA, LLMGnc_172 (ArgX) and LLMGnc_147. An analysis of the 5-min stress response of L. lactis NCDO712 by RNA-seq...
revealed many uncharacterized small regulatory RNAs (Van der Meulen et al. submitted).

The sRNAs and asRNAs add a whole new layer to the complexity of control of gene regulation, and considering their omnipresence and influence represent a nearly untapped source of new regulatory mechanisms in LAB.

**SYSTEMS BIOLOGY**

Large amounts of –omics data are currently generated and used to describe the functions of and mutual interactions between genes, proteins, (regulator) RNAs and metabolites. In systems biology, this extensive experimental data are combined with engineering approaches, physics, mathematics and computer science with the aim of developing a quantitative and conceptual understanding of how biological systems ‘work’. By an iterative approach, it tries to predict and accurately simulate complex biological behavior. As for the topic of this review: systems biology has huge potential in biotechnology, in trying to comprehend the functioning of LAB and to subsequently use that knowledge to optimize these organisms e.g. for food fermentation purposes (Teusink, Bachmann and Molenaar 2011). One way to use the genomics data is to construct genome-scale models containing all the above-mentioned interactions and all available experimental and existing (literature) data of an organism.

Several genome-scale metabolic models have been presented for a number of different LAB species, while on the other hand in-depth analyses of subparts of the cell’s machinery are also still undertaken (Oliveira, Nielsen and Forster 2005; Teusink et al. 2006; Voit et al. 2006; Pastink et al. 2009; Flahaut et al. 2013). For comprehensive overviews, see Smid et al. (2005); Smid and Hugenholtz (2010); de Vos (2011); dos Santos, de Vos and Teusink (2013). These metabolic models can be used a.o., to analyze fluxes in metabolic networks, the physiology of growth in complex media or microbial ecosystems, or understand genotype–phenotype relationships (Teusink, Bachmann and Molenaar 2011; Goel et al. 2015). A genome-scale metabolic model of Lactococcus lactis was developed that contained the information of 518 genes, 650 metabolites and 754 reactions of which 59 were (in)directly or involved in flavor formation (Flahaut et al. 2013). It could be used to predict and, most importantly, subsequently validate the formation of volatile sulfur compounds during fermentation. In another study, the transcriptome data obtained from aerobically grown cultures that consistently showed a temporary stagnation in growth during the early logarithmic phase were plotted onto a metabolic map using a metabolic model for Lactobacillus plantarum WCFS1 (Stevens et al. 2008). The analysis revealed that in the cells that resumed growth a number of CO2-producing pathways had been activated. The supposition that growth had stopped because of a limitation in CO2 availability proved correct and could be alleviated by increasing the CO2 gas partial pressure during aerobic fermentation. Genome-scale models and experimental data have also been used in a comparative way to examine the metabolic differences between Streptococcus thermophilus and L. lactis and Lb. plantarum (Pastink et al. 2009). After construction of a genome-scale metabolic model of S. thermophilus, these differences could be visualized by direct projection on the metabolic map. The comparative analysis highlighted the limited amino acid dependency of the S. thermophilus strain used, its broad potency of producing a variety of volatiles from amino acids as well as a number of industrially relevant assets, such as the unique pathway for acetaldehyde (yogurt flavor) production.

These models, in combination with extensive experimentation, have also been used to examine the shift from homolactic to heterolactic fermentation in L. lactis. When L. lactis is growing fast, in the presence of a preferred carbon source such as glucose, the main fermentation product is lactate. This so-called homolactic fermentation yields two molecules of ATP per glucose molecule. When a less favorable carbon source is present or during aerobic growth, it performs heterolactic fermentation generating, besides lactate, acetate, formate, ethanol, CO2, aconitase and other end products. This growth mode yields three ATPs per glucose. The shift from heterolactic to homolactic fermentation occurs as the growth rate of bacteria increases (Thomas, Ellwood and Longyear 1979; Neves et al. 2005; Goel et al. 2015). Several molecular mechanisms have been entertained over the years to explain these observations but none of them gave a satisfactory explanation for the phenomenon. Metabolic shifts are usually viewed as a trade-offs between catabolic rate and ATP yield: the free energy of the substrate of a pathway can be used either to produce high free-energy intermediates or to drive the pathway quickly (Goel et al. 2012; Bachmann et al. 2013). At low substrate concentrations, efficient metabolism leads to higher growth rates, while inefficient metabolism gives higher growth rates at high substrate concentrations (Molenaar et al. 2009; Goel et al. 2012). It is generally assumed that cells adjust their components—transcripts, ribosomes and enzymes—to changing environmental conditions and to adapt in an ‘ecomomical’ way, to guarantee survival. This is, however, not the case for L. lactis adapting to glucose availability (Goel et al. 2015). Although higher growth rates lead to increasing glycolytic fluxes, the amount of (the majority of) transcripts, ribosomes and proteins changes only modestly and not in proportion to the growth rate. The enzymes for both metabolic strategies—homolactic and heterolactic fermentation—are always present in the cell. Thus, L. lactis is always prepared for fast growth once environmental conditions become optimal (Goel et al. 2015). In order to keep an almost constant amount of ribosomes in a cell at different growth rates, L. lactis dimerizes its ribosomes. When not in use, ribosomes are stored as 100S complexes. The ribosome dimerization factor YfA is highly expressed during slow growth periods such as stationary phase (Puri et al. 2014; Breuner et al. 2016).

By using the associations between mRNA levels, the structures of operons and regulons and data on transcription factor binding sites, functional classes of proteins, metabolic pathways and sRNAs, GRN can be built. Extensive GRNs on manually curated data are available for Escherichia coli and Bacillus subtilis (Salgado et al. 2013; Michna et al. 2014). On the basis of a time-series transcriptome analysis of L. lactis growing in milk, in combination with literature data of known L. lactis regulons, a GRN was constructed (de Jong et al. 2013). A more extended GRN based on a large number of publically available transcriptome data was recently reconstructed (Omony et al. submitted). The structure of this L. lactis MG1363 GRN showed high similarity to those of the gold standard networks of E. coli K-12 and B. subtilis 168. An example of a gene regulatory subnetwork of the N- and C-regulators discussed in this review is presented in Fig. 4.

**THE IMPORTANCE OF STUDYING SINGLE LACTOCOCCUS LACTIS CELLS**

Industrial dairy fermentations are by no means homogeneous and as soon as the starter bacteria are added to the milk, differences will inevitably occur even if the bacteria/milk mixture
Figure 4. GRN of the five pleiotropic regulators described in this review. Cytoscape model depicting the non-scaled interactions of five selected transcriptional regulators (green diamonds) with genes (gray circles) and other regulators (yellow diamonds), on a blurred background of the total GRN of L. lactis MG1363.

would be constantly and thoroughly stirred (which one normally doesn’t!). In such an industrial setting, variations with respect to the local microenvironment of the cells of the starter can arise shortly after spatial separation as a consequence of the coagulation of milk during the fermentation process. Thus, L. lactis cells are usually quickly spatially scattered and faced with local nutrient and other gradients that pose different challenges to the bacteria with respect to growth and survival. Isogenic subpopulations with different phenotypes and activities will inevitably occur. In fact, as we have seen above, even growth in fully liquid media can lead to the occurrence of subpopulations of otherwise isogenic cells, displaying different phenotypes (Solopova et al. 2014). As is clear from this study of glucose to cellobiose/lactose diauxie, observations at the population level might differ quite considerably with what is actually happening at the level of the single cells that make up the population. This work is one of many that underscore the importance of single-cell studies to expose temporal diversifications emerging in monoclonal populations, to resolve macroscopic processes such as fermentation of complex food matrices or (mixed-species) biofilm formation (Grimbergen et al. 2015 and references therein).

To study major aspects of bacterial life not as an average of all cells in a population but, instead, at the level of the single cell or of ensembles of single cells is a rapidly emerging field. Again, big strides have recently been made possible by major advances in technology, in this case of high- and superresolution (microscopy) imaging, allowing examining cellular processes beyond the diffraction limit of light (Schneider and Basler 2016). Several key processes in bacterial cells appear to be well orchestrated and organized; they transiently or statically occur at specific sites in the cell (Govindarajan, Nevo-Dinur and Amster-Choder 2012). Thus, highly dynamic and important macromolecules and macromolecular assemblies, such as lipids, proteins, ribosomes, plasmids and RNA species move in varying diffusive states and with different velocities in a crowded intracellular environment (Mika and Poolman 2011). Fluorescent reporter fusions have been employed to monitor transcriptional activity of a promoter and to study the timing and level of gene expression in individual bacterial cells. In this way, the appearance of stochasticity and bistability in monoclonal bacterial populations of promoters as well as stress responses or changes in transcriptional activity could be elucidated at the single-cell level (reviewed in Norman et al. 2015).

The main importance of L. lactis lies in the dairy industry, but L. lactis has also emerged as a potential candidate in oral vaccination strategies, and as an expression host for a host of different biologically active peptides and proteins (Mierau and Kleerebezem 2005; Robert and Steidler 2014; Wyszyńska et al. 2015). Specifically, L. lactis has gained attention as a valid and valuable alternative for E. coli and B. subtilis for the production of recombinant membrane and secreted proteins (Kunji et al. 2005; Morello et al. 2008). It is easy to handle, fast-growing, and
well-defined molecular toolkits are available (see above), while protein inclusion bodies like those in E. coli have never been detected in L. lactis (Kunjii, Slotboon and Poolman 2003; van Gijtenbeek et al. 2016). Notwithstanding this, expression of heterologous membrane proteins is still suboptimal and recent single-cell studies have shed light on the bottlenecks involved (van Gijtenbeek et al. 2016). The transcripts of a well-expressed homologous membrane protein (BcaP) localized to the cytoplasmic membrane (Fig. 3B). Disruption of the translation process led to loss of the membrane-proximal mRNA foci, suggestive of a process of transertion, the coupling of transcription, translation and protein insertion in the membrane. It was shown that mRNAs encoding poorly produced membrane proteins accumulate in mRNA-dense immobile bodies at the cell poles (Fig. 3B). Cells carrying such polar mRNA foci stopped growing; growth only resumed when these mRNA foci had disappeared. The heat-shock response was activated in the severely stressed cells and a lowering of the pool of active ribosomes is also observed. It is therefore highly likely that the SR is involved due to nutrient limitation or expression of the heterologous membrane protein (van Gijtenbeek et al. 2016; Marreddy et al. 2011). It was observed that, like in E. coli (Prilusky and Bibi 2009), the occurrence of uracil residues in transcripts of membrane proteins is higher than that in mRNAs of cytoplasmic proteins. The cold-shock proteins CspE and CspC interact with the uracil-rich transcripts as well as the inner membrane of E. coli and might therefore not only stabilize but also carry them to the membrane (Benhalley et al. 2015). Interestingly, a similar mechanism might operate in L. lactis, as the expression of most CSP genes was downregulated when a heterologous, but not a homologous, membrane protein was overexpressed (Marreddy et al. 2011; Pinto et al. 2011). Both the in-depth population (Pinto et al. 2011) and single-cell analyses (van Gijtenbeek et al. 2016) allowed improving L. lactis as a membrane protein production host.

CONCLUSIONS AND FUTURE PROSPECTS

The evolution of gene regulatory research in Lactococcus lactis: it has rather been a revolution as may be clear from this review. In some 30 to 40 years, the know-how on genes, their functions and mutual interactions in L. lactis have exploded. This has for a large part been made possible by the development of improved or entirely novel technologies, both in hardware and software, in automation and miniaturization. The ‘biotechnology’, using life’s biomolecules such as the great diversity of DNA modifying enzymes, plasmids and functional DNA ‘parts’, has been fundamental and instrumental in this respect. In this day and age, a young researcher entering the field for the first time has at his/her disposal a large array of tools, methods and machinery. The expanding extension of the genetics toolbox, such as the recent advances in CRISP-Cas technology and the synthesis of large DNA molecules and their assembly into genome-sized molecules by in vitro enzymatic methods and in vivo recombination in the yeast Saccharomyces cerevisiae (Gibson et al. 2010), offers entirely new possibilities and enables looking at fundamental as well as more applied (industrial) issues from a completely different angle. Two research avenues have emerged as a result of the latest boosts in technology. Both will be important for further fundamental research on L. lactis as well as for future applications of L. lactis in the ‘old’ but certainly also in new biotechnology.

On the one hand, the ever-increasing span and speed at which DNA and RNA can be sequenced allows for entirely new ways of looking at scientific issues. Any new strain that is sequenced or isolated will now first have its entire genome sequence determined. Not one at a time, but if need be with many strains in a single sequence run, allowing to probe the enormous diversity that potentially exists among L. lactis strains. Integration of these data with high-throughput phenotypic analyses such as ultrafast GC-time of flight MS technology (Smit et al. 2004; de Bok et al. 2011) should allow screening of large strain collections for L. lactis starter culture strains with characteristics that are desired in dairy fermentation (Tan-a-ram et al. 2011; Dhaisne et al. 2013). In addition to single strains or mixes of pure strains, the whole complex of LAB strains and species in fermentation products is being heavily scrutinized by metagenome sequencing approaches (Quigley et al. 2012; Erkus et al. 2013; O’Sullivan et al. 2015). Apart from these metagenomes, the metatranscriptomes are also accessible and can be routinely queried to answer major (industrial) questions. In this way, not only the presence, distribution and relative abundance of the microbes are recorded but, most importantly, their mutual interactions in the form of genome activity (van Hijum, Vaughan and Vogel 2013; Dugat-Bony et al. 2015) Thus, the aim would be to understand, reproduce and ultimately influence the sequential development of the complex metabolic patterns that are active in milk fermentation, or during cheese ripening.

High-throughput transcriptomic and, to a lesser extent, proteomic analyses have aided in uncovering the intricacies of GRNs and their interconnections that are operational in L. lactis. Some of the inconsistencies occasionally observed between transcriptome, proteome and metabolic data can now be addressed by deep RNA sequencing, in the realization that post-transcriptional regulation by sRNAs is very important and omnipresent in bacteria and eukaryots alike. The fact that some sRNAs are regulated by transcription factors while, conversely, sRNAs can also control transcription factors (Mandin and Guiller 2013) shows that GRNs can be very complex, and that a whole new layer of control is waiting to be described for L. lactis (van der Meulen et al. 2016). Determining the functions of sRNAs and their interaction partners will certainly fill some of the gaps that exist in our understanding of gene expression behavior. One of the issues to be solved in L. lactis is the lack of an obvious sRNA chaperone such as Hfq in Gram-negative bacteria (van der Meulen, unpublished). Hfq acts as an RNA chaperone that can bind both sRNA and mRNA and is as such crucial for sRNA functioning, at least for those that interact with Hfq. Deletion of Hfq in E. coli leads to pleiotropic effects and increased sensitivity to some stresses (Tsui, Leung and Winkler 1994). Staphylococcus aureus Hfq seems not to be crucial for stress tolerance (Bohn, Rigoulay and Boulouc 2007). Hfq is present in nearly 50% of the bacteria that have been sequenced, especially in those with a high GC content (Jousselin, Metzinger and Felden 2009). Recently, ProQ has been identified as another important RNA-binding protein that can bind to and stabilize sRNAs on a large scale (Smirnov et al. 2016). In L. lactis, neither Hfq nor ProQ is present, suggesting that its sRNAs, specifically those that are trans-encoded, either do not require a chaperone or use another yet to be identified RNA-binding protein(s). From the perspective of application, a deeper mechanistic insight in gene regulation by non-coding RNAs could also facilitate more effective gene silencing by asRNAs. The asRNA targeting methodology has been applied, with limited success, already several decades ago in LAB (Kim and Batt 1991). A recent example in Oenococcus oeni shows that asRNAs can be useful in the study of gene function when genetic tools are scant (Darsonval et al. 2015). RNA can not only be used for control purposes, the CRISPR-Cas9 technology offers an exciting novel set of applications of RNA-directed genome
engineering that will undoubtedly have far-reaching consequences also for research in LAB.

Next-generation sequencing in combination with experimental evolution approaches offers completely new possibilities of obtaining strains with novel phenotypes. The underlying genetic make-up of the strains can be easily and quickly determined by sequencing their entire genomes. Several examples exist of clever experimental evolution approaches and evolved LAB strains but, as the topic of experimental evolution in LAB is dealt with separately in this issue of FEMS Microbiology Reviews, it was not discussed in this paper.

On the other hand, the recently emerging field of high-resolution spatio-temporal imaging of processes in single bacterial cells will offer entirely new views on academic as well application-oriented issues at single cell of even single-molecule resolution. How bacteria respond to nutrient starvation, or to a diversity of other (industrial) stressors, during the various stages of growth in laboratory media or throughout milk fermentation, can now be assessed at the single-cell level. Many microorganisms growing in rapidly changing, unpredictable environment produce siblings with diverse phenotypes, a response that increases the chance of survival in an altered milieu of at least a subpopulation of the cells. This so-called phenotypic heterogeneity can be generated by intracellular noise such as stochastic variations in gene expression and/or by differences in local concentrations of extracellular nutrients (extracellular noise; Elowitz et al. 2002; Fraser and Kaern 2009). The first insights in the ‘choices’ that L. lactis makes upon entering a lag phase as a consequence of carbon source depletion have already shown that even in a liquid culture of isogenic cells subpopulations with different phenotypes exist (Solopova et al. 2014). Superresolution microscopy now also allows peering deep into bacterial cells and obtaining molecular mechanistic understanding of important intracellular processes, their dynamics and cellular location, by examining the spatial distribution and behavior of proteins, protein complexes and other biomacromolecules (Yao and Carballido-López 2014; Schneider and Basler 2016). From the fundamental point of view, exciting discoveries are to be made in the near future.

These single-cell responses and adaptations of L. lactis are also relevant for more application-oriented goals, considering that dairy fermentation creates the perfect setting for the starter bacteria to make individual choices. Milk is an emulsified colloidal suspension, while the matrix of milk fermentation products is a subpopulation of the cells. This so-called phenotypic heterogeneity can be generated by intracellular noise such as stochastic variations in gene expression and/or by differences in local concentrations of extracellular nutrients (extracellular noise; Elowitz et al. 2002; Fraser and Kaern 2009). The first insights in the ‘choices’ that L. lactis makes upon entering a lag phase as a consequence of carbon source depletion have already shown that even in a liquid culture of isogenic cells subpopulations with different phenotypes exist (Solopova et al. 2014). Superresolution microscopy now also allows peering deep into bacterial cells and obtaining molecular mechanistic understanding of important intracellular processes, their dynamics and cellular location, by examining the spatial distribution and behavior of proteins, protein complexes and other biomacromolecules (Yao and Carballido-López 2014; Schneider and Basler 2016). From the fundamental point of view, exciting discoveries are to be made in the near future.

These single-cell responses and adaptations of L. lactis are also relevant for more application-oriented goals, considering that dairy fermentation creates the perfect setting for the starter bacteria to make individual choices. Milk is an emulsified colloidal suspension, while the matrix of milk fermentation products is a colloidal structure with diverse niches of casein micelle clusters, fat globules and void spaces filled with whey. Together, they constitute a blend of microenvironments with quite different characteristics and nutrient conditions. Interactions among starter culture cells and between these cells and matrix components are important for the ultimate location of the bacteria and, thus, their activity in the fermented dairy foods. In cheese, for instance, the bacteria can form microcolonies in the curd (Hannon et al. 2006; Jeanson et al. 2011). Partitioning of the bacteria is dependent on many physico-chemical interactions in which the cell wall, a composite structure of glyco-polymers and proteins, is an important player (Sheehan et al. 2009; Jeanson et al. 2011; Burgain et al. 2013; Chapot-Chartier and Kulakauskas 2014). Cell wall charge and composition are highly flexible and provide ample opportunity for strain variation and possible subsequent changes in location and performance in dairy products. It is evident that dairying provides highly complex ecosystems in which, depending on the fermentation product, a multitude of bacterial and fungal strains and species are expected to each perform their required task, in cooperation or in competition, and often under pressure of invading bacteriophages and other stressors. A lot still needs to be learnt in order to understand and ultimately fully control and steer fermentation by L. lactis, be it for (improvement of) traditional foods, for entirely novel food- or medicine-related products, or even for biorefinery purposes using bulk materials other than milk (Gaspar et al. 2013). Ideally, L. lactis should perform in these processes with maximum predictability, like a smoothly operating machine or, rather, a whole factory. With the tools at hand to modify the organism and others to precisely measure the resultant effects at population and single-cell levels, and with undoubtedly much more technology emanating, this should ultimately be a realistic goal.

ACKNOWLEDGEMENTS

The authors thank all past and present collaborators. The authors apologize to those researchers whose original research work could not be cited due to space limitations.

FUNDING

Part of the work presented here was associated with the Top Institute Food and Nutrition. SBvdM, OPK and JK were supported by BeBasic FS10, AS by an Stichting Technische Wetenschappen (STW) grant in the scope of Project 10619, and the EU FP7 project BachBerry, and LAVG by a TOP-GO program granted by the Netherlands Organization for Scientific Research (NWO), project 700.10.353.

Conflict of interest. None declared.

REFERENCES


Bolotin A, Mauger S, Malarme K et al. The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403. Genome Res 2001;11:731–53.


