FEMS Microbiology Reviews, 2022, **46**, 1–15
DOI: 10.1093/femsre/fuac014
Advance access publication date: 7 March 2022
Review Article

Expanding natural transformation to improve beneficial lactic acid bacteria

Stefano Di Giacomo [1], Frédéric Toussaint [1], Laura Ledesma-García [1], Adrien Knoops [1], Florence Vande Capelle [1], Christophe Fremaux [2], Philippe Horvath [2], Jean-Marc Ladrière [2], Hassina Ait-Abderrahim [3], Pascal Hols [1,*,†], Johann Mignolet [1,†,$]

[1]Biochemistry and Genetics of Microorganisms (BGM), Louvain Institute of Biomolecular Science and Technology, Université catholique de Louvain, Croix du Sud 4-5, (box L7.07.06), B-1348 Louvain-la-Neuve, Belgium
[2]Health and Biosciences, IFF Danisco France SAS, CS 10010, F-86220 Dangé-Saint-Romain, France
[3]Lesaffre, 137 Rue Gabriel Péri, 59700 Marcq-en-Barœul, France
*__Corresponding author:__ Biochemistry and Genetics of Microorganisms, Louvain Institute of Biomolecular Science and Technology, Université catholique de Louvain, Croix du Sud 4-5 (box L7.07.06), B-1348 Louvain-La-Neuve, Belgium. Tel: +3210478896; Fax: +3210472825; E-mail: pascal.hols@uclouvain.be
One sentence summary: The authors review the advantages of using natural DNA transformation to genetically improve lactic acid bacteria for human benefits by providing a step-by-step strategy to predict and trigger this natural gene exchange process.
†These authors contributed equally.
$Current address: Department of Fundamental Microbiology, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland.
Editor: Aimee Shen

Abstract

Nowadays, the growing human population exacerbates the need for sustainable resources. Inspiration and achievements in nutrient production or human/animal health might emanate from microorganisms and their adaptive strategies. Here, we exemplify the benefits of lactic acid bacteria (LAB) for numerous biotechnological applications and showcase their natural transformability as a fast and robust method to hereditarily influence their phenotype/traits in fundamental and applied research contexts. We described the biogenesis of the transformation machinery and we analyzed the genome of hundreds of LAB strains exploitable for human needs to predict their transformation capabilities. Finally, we provide a stepwise rational path to stimulate and optimize natural transformation with standard and synthetic biology techniques. A comprehensive understanding of the molecular mechanisms driving natural transformation will facilitate and accelerate the improvement of bacteria with properties that serve broad societal interests.

Keywords: competence, natural transformation, lactic acid bacteria, industrial strains, *Lactococcus*, *Lactobacillus*, *Streptococcus*, bioengineering

Introduction

The constant rise of human population puts under pressure our agricultural systems, while, at the medical level, emerging diseases (e.g. due to antibiotic resistant bacteria) and modification of our environment impinge on human health (de Kraker *et al.* 2016, van Duin, Paterson 2016). Therefore, new complementary strategies and optimization of current bioproduction methods will be required in a near future. In this scope, the huge diversity of beneficial lactic acid bacteria (LAB) turns out to be a promising pool of metabolic factories that we can tap into.

For decades, LAB have assumed a tremendous and growing economic importance in the agro-food sector, thanks to their fermentative lifestyle that converts various sugars into (isomers of) lactic acid (De Filippis *et al.* 2020). They have a key role in beverages and food maturation/transformation of dairy products and various fishes and vegetables (e.g. cheese, cabbage, yogurt, kefir, sourdough bread, kimchi, sauerkraut, and pickles), and have been used as natural preservatives for millenniums. Currently, LAB participate in a myriad of processes, including improvement of nutrient value, organoleptic diversity, and product texture (Johansen 2018). In addition, their epithelium colonization capabilities and potentially their secretion of antimicrobial compounds such as bacteriocins make several LAB valuable and safe probiotics (Hols *et al.* 2019, Mathur *et al.* 2020). Other applications based on LAB are emerging in the industry and could replace the traditional and well-known *Escherichia coli* and *Saccharomyces cerevisiae* bio-based factories (Hugenholtz, Smid 2002, Papagianni 2012). The elaboration of fermentation molecules (aroma), biopolymers, vitamins, cofactors, plastic building blocks, solvents, pharmaceuticals, and cosmetics might be expanded to LAB as they neatly adapt to food-related environments. First, LAB could be maintained in a bioreactor at low cost considering that the nutritional media used to grow these LAB cell factories could emanate from raw and nontransformed agricultural products. Second, LAB might be exploited with a recycling scope to valorize biomaterial waste such as whey for bioproduction purposes. In addition, LAB are part of the phyto-microbiome and could be used for biocontrol of agricultural lands as a safe substitute for pesticides (Lamont *et al.* 2017). Prospectively, LAB gain attention in regard to drug and vaccine delivery, diagnostic, and therapeutics (Van Braeckel-Budimir *et al.* 2013, Bron, Kleerebezem 2018, Chowdhury *et al.* 2014) such as fecal transplants and curing oral vectors for specific gastro-intestinal disorders.

Received: November 24, 2021. **Revised:** February 14, 2022. **Accepted:** March 1, 2022

Sometimes, beneficial properties required for a specific purpose are borne by different microorganisms. They have to be blended to operationally optimize functionalities. However, some of them could be mutually exclusive or impede growth in industrially relevant conditions. Therefore, these genetically encoded capabilities could be transferred into a unique microorganism adapted to human constraints (e.g. fermentation reactors or food safety issues). Selection of upgraded LAB, thus represents a promising scope to optimize their probiotic or phytosanitary properties, control their antimicrobial behavior, favor their protection against bacteriophages, combine multiple interesting genetic traits in a unique cell factory, and/or enhance the yield of value-added products. From a scientific viewpoint, genetic improvement might also be used to decipher LAB metabolism, stress response, food adaptation, and molecular interactions with the host (plant and animals) immune system (Borner *et al.* 2019, Cho *et al.* 2020, Hols *et al.* 2019).

A comprehensive understanding of how beneficial LAB thrive in relevant environments will be a prerequisite to strain improvement. Interestingly, bacterial natural transformation involving homologous DNA is a highly efficient process for the modification of LAB genomes for research and applied purposes. This intrinsic capacity enables recipient cells to capture extracellular fragments of DNA available in the surrounding medium and promote foreign DNA insertion into their chromosome (Johnsborg *et al.* 2007). Working with pathogenic LAB strains (i.e. *Streptococcus pneumoniae*), Frederick Griffith discovered natural transformation in 1928 and directly realized the tremendous potential to transfer genetic traits in between bacteria (Griffith 1928). Nowadays, it is a simple and fast technique to routinely make genetic changes to a significant number of bacteria in laboratory environments.

In this review, we showcase beneficial LAB for their intrinsic property to take up DNA and to modify their genome. We describe the physiological process, the regulation mechanisms, and the protein factors required to generate a productive natural transformation. We use bacterial paradigms, including *Bacillus subtilis* (non-LAB species) and *S. pneumoniae* (pathogenic LAB species), to extrapolate some aspects for beneficial LAB strains. We also show how widespread these factors are within industrially relevant LAB. We next elaborate on various practical strategies to augment transformability of LAB strains. Finally, we discuss the benefits of natural transformation to decipher bacterial physiology with concrete examples.

Biogenesis of the competence nanomachine

To be naturally transformable, bacteria have to enter a transient physiological state known as competence, during which they execute a dedicated sequential developmental program, i.e. intricately wired to fundamental processes of bacterial life cycle (Chen, Dubnau 2004, Salvadori *et al.* 2019). During this stage, they synthesize about 15 proteins that assemble into a sophisticated helical nanomachine at the cellular surface to internalize extracellular naked DNA (Claverys *et al.* 2009, Dubnau, Blokesch 2019; Fig. 1A). All these proteins that form the so-called transformasome are essential for an effective transformation process and they are typically encoded in half a dozen of operons, including the so-called late competence (*com*) genes (as opposed to early *com* genes; see next section; Fig. 1B). They can be clustered into three specific phases.

DNA docking

The first transformation step requires free stretches of double-stranded DNA floating in the environment. To initiate electrostatic contacts with DNA, bacteria assemble a pseudo-pilus macrostructure reminiscent of type IV pili and the type II secretion system (Laurenceau *et al.* 2013, Muschiol *et al.* 2019). They are formed by an ordered heteromer of major (ComGC) and minor (ComGD-GE-GF-GG) pilin subunits that span the cell wall to be exposed at the cellular surface (Chen *et al.* 2006, Muschiol *et al.* 2017; Fig. 1A). Synthesized as precursors, the pilin subunits are processed *via* ComC (pre-pilin peptidase; Chen *et al.* 2006) to cleave the hydrophilic class III signal sequence prior to incorporation of mature pilins at the base of the filament anchored by the transmembrane platform protein ComGB. Inside the fiber, both major and minor pilins are linked through noncovalent interactions, while major pilins are also bound *via* covalent disulfide bridges (Chen *et al.* 2006). The pilus extension is driven by pilin incorporation and requires energy, drawn in the proton-motive force, which is generated by the AAA^+ (ATPases Associated with various cellular Activities) ATPase ComGA (Laurenceau *et al.* 2013). The ComG proteins are required to bind DNA even if no such capacity was reported for the pilin polymer itself (Chung, Dubnau 1998, Briley *et al.* 2011, Laurenceau *et al.* 2013). Instead, the pseudopilus is supposed to facilitate DNA access to the DNA receptor protein ComEA *via* the modification of the cell wall and envelope, presumably with the creation of a pore (Provvedi, Dubnau 1999). Nonexclusively, free DNA strands might "slide" on the pilus, or the pilus retraction (depolymerization of pilin subunits) might escort DNA to bring it in intimate contact with ComEA (Lam *et al.* 2021). Its C-terminus encompasses helix-hairpin-helix motifs responsible for the nonspecific sequence recognition of DNA (Provvedi, Dubnau 1999). It is noteworthy that a *comEA* deletion mutant still sustains a residual DNA-binding activity, arguing for the presence of other DNA receptors on the cell envelope (Inamine, Dubnau 1995).

DNA translocation

Due to a membrane anchor at its N-terminus, ComEA is in close proximity with the outer leaflet of the cellular membrane. Once ComEA is loaded with DNA, it interacts with and delivers DNA to the channel ComEC (Kramer *et al.* 2007) and the nuclease EndA (Berge *et al.* 2013). This large permease ComEC features seven transmembrane helices and homodimerizes to allow the passage of DNA (Draskovic, Dubnau 2005). Simultaneously, the constitutively produced EndA converts double-stranded DNA into single-stranded DNA that translocates in the 3'–5' orientation (Mejean, Claverys 1988) through ComEC with the concomitant release of soluble nucleotides into the surrounding environment (Midon *et al.* 2011, Moon *et al.* 2011). The energy required for DNA transport across the membrane is powered by the inner leaflet-associated ComFA that hydrolyses ATP through its ATP-binding cassette domain (Diallo *et al.* 2017). Its function could be to merely energize the single-stranded DNA passage by itself. However, ComFA, in association with co-factors, might alternatively consume ATP to gate the permease ComEC or to unwind double-stranded DNA.

Genome remodeling

On the cytoplasmic face of the membrane, ComFA forms a complex with ComFC, a competence-essential protein of elusive function, and DprA, that protects the incoming single-stranded DNA from intracellular housekeeping nuclease attack (Diallo *et al.* 2017). DprA subsequently facilitates homologous recombination by loading RecA onto the foreign DNA (Mortier-Barriere *et al.* 2007, Quevillon-Cheruel *et al.* 2012). RecA forms a filament that surrounds DNA and scouts the bacterial chromosome for sequence homology (Berge *et al.* 2003). In combination with DprA, SsbB (and SsbA), and CoiA, it stabilizes single-stranded DNA in the cyto-

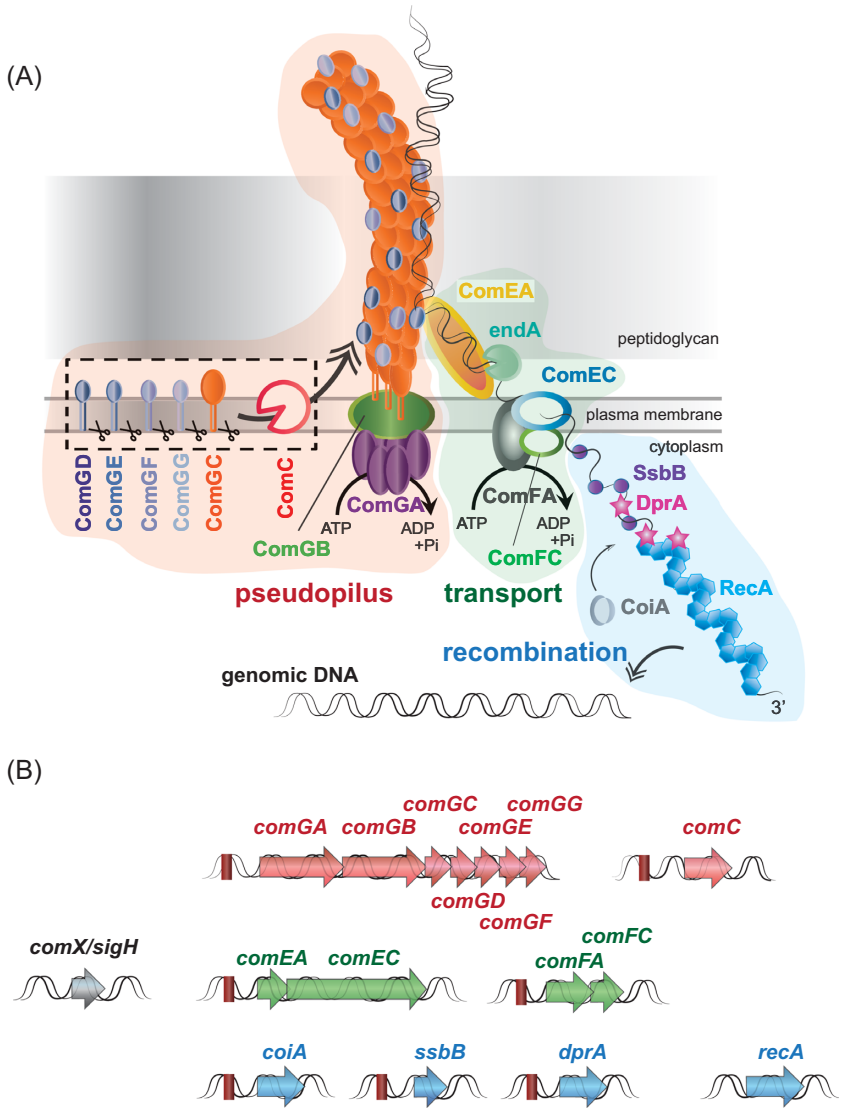

Figure 1. The transformasome biogenesis. **(A)** Schematic representation of the transformasome architecture. The genes involved in the three steps of DNA transformation, namely the DNA docking, transport, and processing, are color-coded in red, green, and blue, respectively. The proteins composing the competence machinery are shaded accordingly. ComC processes the major and ancillary pilin subunits to promote their incorporation into the pseudopilus filament. The two ATPases ComGA and ComFA generate energy *via* ATP hydrolysis. **(B)** Operon structure of the late *com* genes. The red boxes in the promoter region of *com* genes depict the ComX binding sites (*cin*-boxes).

plasm, increases its survival time, and maximizes the probability of genetic exchange (Desai, Morrison 2007, Attaiech *et al.* 2011, Morrison *et al.* 2007).

The central node for competence activation

Competence imposes a dramatic burden on cell physiology (Nester, Stocker 1963, Haijema *et al.* 2001, Mignolet *et al.* 2018). First, the production of transformasome subunits and the dynamics of the pseudo-pilus consume abundant building blocks (amino acids) and energy (ATP and PMF). Additionally, competent state engages a series of DNA-related processes (e.g. recombination) barely compatible with a vegetative growth (Berge *et al.* 2017). This explains why competence is widely regulated in time and space according to internal and environmental conditions (Johnston *et al.* 2014, Fontaine *et al.* 2015). In contrast, model LAB strains such as *S. pneumoniae* R6, *Streptococcus mutans* UA159, and *Streptococcus thermophilus* LMD-9, were extensively studied because they tolerate to develop competence in laboratory conditions, sometimes for extended periods of time, and display an intrinsic high rate of transformation (Gardan *et al.* 2009, Fontaine *et al.* 2010a, Johnston *et al.* 2020, Underhill *et al.* 2019). Owing to disadvantageous collateral effects, competence requires a robust "no-return" commitment at appropriate moments. Therefore, a master competence regulator, part of the early phase of competence, serves as a cellular checkpoint that integrates external and internal cues at the transcriptional and post-translational level to implement the competence developmental program (Fig. 2). In LAB, this bottleneck regulator is an alternative sigma factor named ComX (or SigX; σ^X) in streptococci, lactococci, and *Leuconostoc* sp. (Fontaine *et al.* 2015), or SigH in lactobacilli and pediococci (Schmid *et al.* 2012). Of note, the taxonomy of the *Lactobacillus* genus was recently revised by its splitting into numerous new genera (Zheng *et al.* 2020). For clarity and simplicity, we will hereunder refer to them as *Lactobacillus*, occasionally providing the new name into square brackets. Both sigma factors are phylogenetically related,

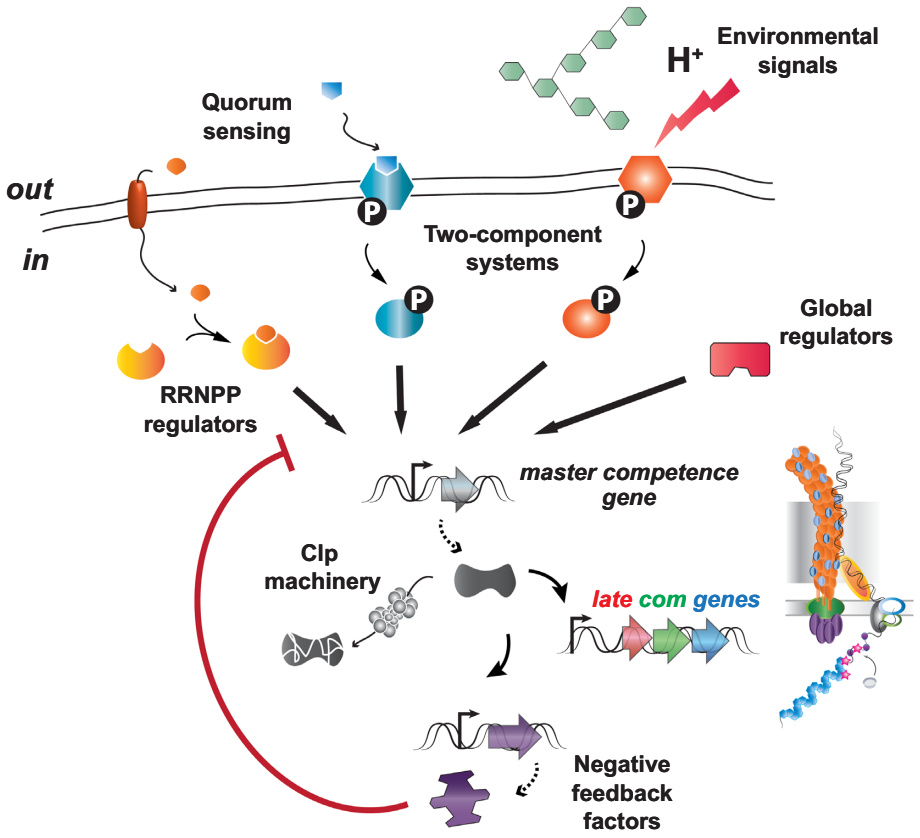

Figure 2. Regulation circuitry of master competence activator. Scheme of the global network regulating the abundance of the master competence activator. Pheromones-based (Quorum sensing) or environmental signals activates membrane (two-component systems relaying phosphate moiety) or cytoplasmic (RRNPP) sensors that in turn influence the master competence gene expression. Nonexclusively, specific global regulators can bind the master competence gene promoter to modulate its expression. The Clp bacterial proteasome controls the steady-state level of the master competence activator through a directed proteolysis. Some factors under the master competence activator control ensure a negative feedback loop that inhibits the competence activation and is responsible for the shut-off mechanism.

however, they are genus-specific given that there has been no reshuffling/horizontal transfer in between the two clades (Fig. 3). This presumably means that the sequence-responsive elements present in the promoter of ComX- and SigH-dependent genes are too distant and that the swapping of one sigma factor by the other one would lead to the incapacity to coordinate the activation of the *com* regulon. In some species, more than one master regulator coexists and can act synergistically to induce competence. For instance, several strains of *S. pneumoniae*, *Streptococcus gordonii*, and *Streptococcus pyogenes* encode two *comX* genes (*comX1* and *comX2*; Martin *et al.* 2006), or even three copies in the case of *Streptococcus anginosus* (Olson *et al.* 2013), while *Staphylococcus aureus* (a non-LAB species) combines a sigma factor (SigH) to one or two transcriptional regulators (ComK1 and ComK2) in order to execute competence activation (Fagerlund *et al.* 2014). Fairly conserved across clades, an asymmetric *cin*-box (TTNCGAATAA) occurs in each streptococcal promoter region of essential late *com* genes (at position -10) and is bound by ComX that contacts the RNA polymerase for a coordinated wave of transactivation (Opdyke *et al.* 2001, Luo, Morrison 2003, Mignolet *et al.* 2018). Additionally, the ComX regulon includes accessory *com* genes that are usually species-specific. For example, the competence activation is usually synchronized with the secretion of antimicrobial molecules, including the ribosomally synthesized bacteriocins (Shanker, Federle 2017). In *S. mutans*, ComX has a direct role on BlpRH, the two-component system that regulates the bacteriocin production and positively influences bacterial predation (Reck *et al.* 2015).

To protect cells against undesired persisting competence activation, several molecular shut-off mechanisms have evolved (Fig. 2). At the post-transcriptional level, ComX undergoes a rapid turnover under the control of the nonspecific prokaryote proteasome ClpEP or ClpCP (Biornstad, Havarstein 2011, Fontaine *et al.* 2015). In LAB, this targeted degradation is promoted by an adaptor protein, MecA, that sequesters ComX and plugs it onto the Clp machinery (Boutry *et al.* 2012, Tian *et al.* 2013, Wahl *et al.* 2014). Although the exact role of MecA on the shut-off phase is still under debate, it definitely ensures a hindrance on the ComX steady-state level. This system acts as a locking device to avoid stochastic competence activation under inappropriate environmental conditions. Superimposed, RNA-seq analyses in *Streptococcus salivarius* and *S. thermophilus* upon competence induction revealed that the deletion of *comX* prolongs the activity of early competence phase regulator (ComR, see hereunder; Haustenne *et al.* 2015, Mignolet *et al.* 2018). This suggests that genes under ComX control negatively impact upstream control systems governing ComX. Strikingly in *S. pneumoniae*, ComX targets the polar accumulation of DprA. This essential member of the transformasome plays a dual role and interferes with the regulation of ComX abundance by promoting the dissociation of ComE dimers, an activator of the early competence phase (see hereunder; Mirouze *et al.* 2013, Johnston *et al.* 2020). In *S. mutans*, the regulator of the early competence phase (ComR, see hereunder) is analogously the target of a negative feedback loop mechanism. Indeed, XrpA (*comX* Regulatory Protein A) interacts with and inhibits ComR to switch off *comX* expression. However,

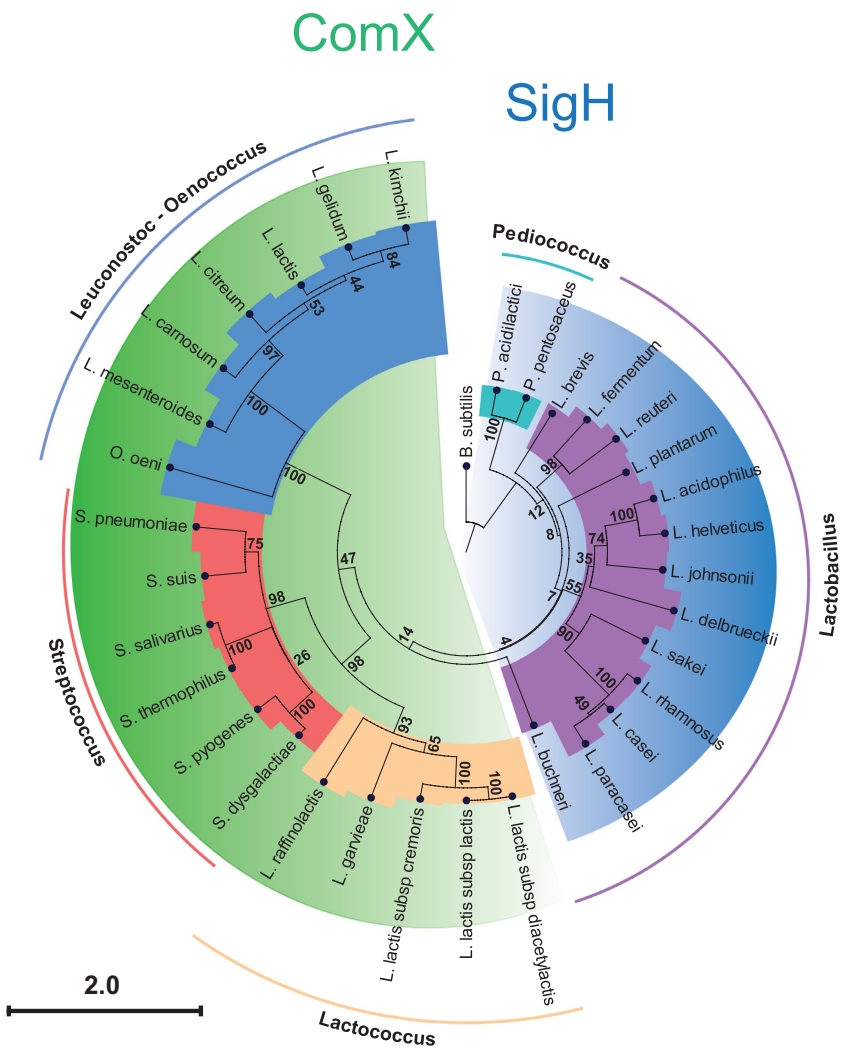

Figure 3. Evolutionary relationship between ComX and SigH. Phylogeny of the master competence activator ComX (green) and SigH (blue) in relevant LAB. A total of five LAB genera are color-assigned: *Pediococcus* (green), *Lactobacillus* (purple), *Lactococcus* (beige), *Streptococcus* (red), and *Leuconostoc-Oenococcus* (blue). The evolutionary history was inferred by the Neighbor joining algorithm and the distance assessed with the Kimura protein method. A total of 1000 replicates were performed to build the tree and bootstrap values are indicated for each phylogenetic branching. The tree with the highest log-likelihood is shown. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site (scale bar, 2 substitution/site). SigH from *B. subtilis* was used as the outgroup to root the tree. The multiple alignments, evolutionary analyses and the tree layout were performed with the CLC Main WorkBench software (http://www.clcbio.com/products/clc-main-workbench/).

the *xrpA* gene is not under the ComX control, but is encoded inside *comX* and expressed from an alternative reading frame (Kaspar *et al.* 2018).

Prediction of LAB transformation potential

Microbiologists face a recurring question in bioengineering: without any preconception on transformation-favorable conditions, what is the successful path to demonstrate that a strain has the potentiality to uptake naked DNA? Impairment of the central regulator or one of the late *com* genes results in a drastic reduction or even a complete abrogation of DNA acquisition (Peterson *et al.* 2004, Claverys *et al.* 2009, Mignolet *et al.* 2018). Therefore, mutation, indel, transposon-insertion, truncation, or absence of at least one *com* gene in a genome is definitely a tangible hint that a strain cannot be naturally transformed, especially when the first gene of an operon is inactivated. Conversely, an intact set of essential *com* genes does not provide an absolute certitude. Several bacteria not known to be transformable possess homologs for *comX*/*sigH* and all late *com* genes. Nevertheless, it is certainly a prerequisite. Thence, a careful biocomputational inspection of the genome is the prime rational step to predict natural transformability.

We performed a systematic study (see Supplementary Appendix 1) of industrial LAB genomes to quantitate the proportion of potential naturally transformable strains. We comprehensively browsed the genome of Generally Recognized As Safe strains in streptococci, lactococci, lactobacilli, pedicococci, oenococci, and *Leuconostoc* sp. and sought for the presence of intact essential competence proteins. In order to avoid ambiguities and false-negative cases (nonsequenced DNA fragments including *com* genes), we only considered strains for which a complete and closed sequenced genome was publicly available, i.e. 482 strains in this analysis. In each genus, we defined a reference set of valid competence proteins (Claverys *et al.* 2009, Bron *et al.* 2019; see Tables S1–S5, Supporting Information) that was used to examine the other strain genomes inside the same genus, estimated the sequence coverage and homology, and reported pseudogenes, disrupted genes (due to small indel, or insertion of prophage, transposon,

or any other mobile genetic element), or absent genes in the *com* gene set. Considering the minor pilin protein (ComGD/GE/GF and GG) that can exhibit a high sequence variability, we looked at the global architecture of the coding operon (gene number and size) to decide whether they are present or absent.

Across LAB species and strains, we noticed a huge variation in the conservation of the *com* gene content and the presumed capability to perform natural transformation (Fig. 4A). On one side, more than 92% of the *S. thermophilus* and *S. salivarius* strains possess a full set of *com* genes. This high number correlates with their transformability, as natural transformation was empirically demonstrated for a representative handful of strains (Table 1; Fontaine *et al.* 2010a). This reflects a high selective pressure to maintain such a capacity. In contrast, no more than 25% of lactococci genomes (mainly plant isolates) contain a complete *com* set. Some *Lactococcus lactis* strains have been reported to be transformable in the absence of the essential protein DprA (Mulder *et al.* 2017). Nevertheless, this low percentage is particularly indicative considering that several strains display a massive *com* gene decay. With the accumulation of up to eight inactive genes, it is unlikely that strains are still proficient for transformation. Lactobacilli, pediococci and *Leuconostoc* sp. are in an intermediate situation, ranging from 40% to 65% of strains with a potentially intact transformasome. This is highly encouraging to envision natural transformation as a rapid genome-editing tool, even if species-to-species profiles display a strong heterogeneity, with 95% for *Lactobacillus [Latilactobacillus] sakei* and 0% for *Leuconostoc lactis* and *Oenococcus oeni*. However, none of these strains so far have been shown to perform natural transformation, making it impossible to relate the *com* gene content to the transformability.

The exact reason why some strains maintain an intact set of *com* genes while some other tolerate a fragmentary function remains elusive in most cases. Obviously, the selective pressure imposed by the biotope and the microbial partners represent a key factor. In lactococci from the dairy industry, some strains have been domesticated through a process called back-slopping (iteration of strain re-inoculation in new batches of fresh dairy products from a previous fermented product). Such stable environments favor a reductive evolution, where costly functions are rapidly discarded by gene mutation if they cannot confer a substantial advantage for growth (Bolotin *et al.* 2004, Kelleher *et al.* 2017, Makarova *et al.* 2006). This is typically the case for the flagellar operons when bacterial mobility is dispensable for survival (e.g. rich laboratory media under constant shaking; Dragosits, Mattanovich 2013, Knoppel *et al.* 2018). Akin to the motorized, polysubunited flagellum, the competence pseudopilus necessitates the ATP-dependent production and polymerization of thousands of major and ancillary pilin subunits, while the transport protein ComEC hydrolyses ATP molecules to transport the DNA inward. Moreover, commitment in the competence phase triggers a global metabolic and physiologic switch that disturbs the cell cycle and slows down the growth. Therefore, we can speculate that wild-type lactococcal strains initially encoded a functional transformasome before they underwent a massive *com* gene decay to adjust to and thrive in their new industrial niches. By contrast, *S. salivarius* strains had to cope in fiercely competitive ecological niches (e.g. gut and oral cavity) with tremendous and recurrent variations at the physicochemical and biological levels. Maintaining an adequate ability to be naturally transformed can be seen as a strategy to regularly acquire new genetic traits beneficial to outcompete microbial adversaries, and reinforce interactions with symbiotic bacterial consortia and host epithelial cells.

The domestication conditions are pretty similar for *L. lactis* and *S. thermophilus* strains, and *S. thermophilus* exhibits on average a higher number of pseudogenes, presumably due to the cultivation in a dairy niche (Makarova *et al.* 2006). However, the output is totally conflicting because *S. thermophilus* strains apparently face a huge selective pressure to keep an intact *com* gene set. How to explain this discrepancy? Considering the proportion of pseudogenes, *S. thermophilus* is prone to accelerated genome decay with a high mutation rate. One reason might be that *S. thermophilus* (in contrast to *L. lactis*) misses *recQ* that codes for a helicase important in genome maintenance and DNA repair from UV, alkylation, and oxidative stresses or mere polymerase errors (Bolotin *et al.* 2004). According to the Muller ratchet theory, asexual reproduction drives the incremental accumulation of mutations to a critical level. Uptake of DNA from kin origin might be, therefore, an opportunity for strains to correct and rescue from detrimental indels and nucleotide polymorphisms. However, *S. thermophilus* possesses an immune system, known as CRISPR-Cas, that prevents the invasion of double-stranded DNA from bacteriophage and plasmid origin (Garneau *et al.* 2010). In a collateral manner, this system drastically reduces the transfer of DNA material from transduction or conjugation that could explain why competence and CRISPR systems coexist and coevolves (Jorth, Whiteley 2012, Gonzalez-Torres *et al.* 2019). Evolvability lies in a delicate trade-off between too low and too high genome plasticity/instability and replication fidelity. While evolutionary trajectories of *L. lactis* mainly relied on plasmid conjugation and phage transduction, *S. thermophilus* used natural transformation to maintain sustainable genome flexibility in presence of stringent foreign DNA acceptor (CRISPR) systems. Possibly, competence development in *S. thermophilus* might be used as an essential nucleotide-based nutrition support that strengthens the energy metabolism during the yogurt acidification process in multispecies environment (e.g. *S. thermophilus* and *Lactobacillus delbrueckii* subsp. *bulgaricus*).

Interestingly, we observed a mutation bias toward genes and phases in competence-deficient strains. Several genes, such as *comX/sigH*, *comGC*, *comGD*, *comC*, *comEA*, or *coiA*, appear as preferential targets of inactivation (Fig. 4B, right panel). Presumably, these genes would be much less favorable for growth in a context where the transformability would be (partially) impaired. For instance, genes could interfere with physiological processes (such as *coiA* for DNA replication or maintenance) and generate a certain toxicity that would be counterbalanced by adaptability benefits in competence-proficient strains. Alternatively, some mutations merely economize energy and amino acids building block, such as mutations in highly expressed genes (e.g. pilin genes) or mutations in one of the first genes of an operon that usually affects negatively the expression of downstream genes (e.g. *comEA*), the extreme case being the loss of the regulatory central node (i.e. *comX* or *sigH*) that totally annihilates the inducible production of all competence proteins. Nevertheless, compared to the DNA docking and translocation genes (Fig. 4B, right panel), the DNA recombination genes (through *coiA* and *dprA*, mainly) proportionally accumulate more mutations. In addition to the colliding effects, these genes could have an impact on the genome integrity; this could also suggest that the two other phases could retain/acquire satellite roles, related or not to competence. As examples, the pseudopilus could be exploited as an adhesive appendage to stick to substrate or capture molecules (DNA and/or other biological constituents), while the DNA translocation system would ensure a nutrition role *via* the assimilation of nucleotides and phosphate moieties. Interestingly, once the competence gene decay is

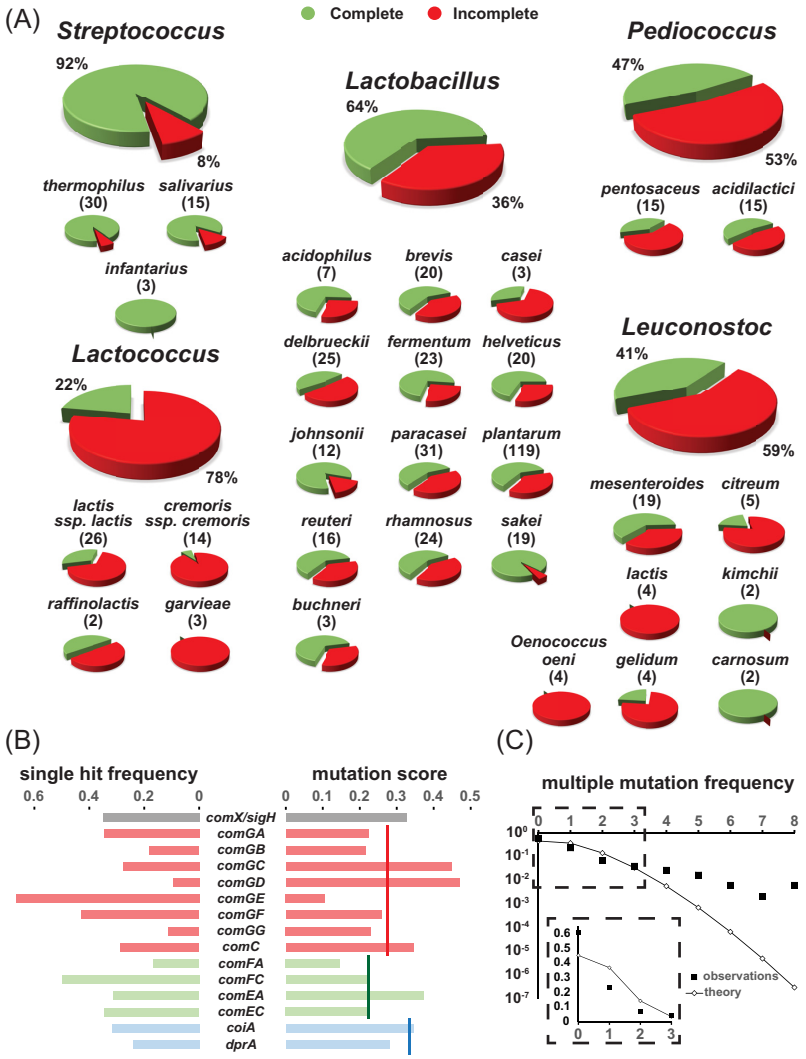

Figure 4. Conservation of competence genes across safe LAB species. **(A)** Pie charts highlighting the percentage of strains that encode an intact (green) vs. incomplete (red) set of *com* genes. Small pie charts quantify it for all the strains of a species (number indicated in brackets), while big pie charts summarize the data collected for the five relevant LAB genera. **(B)** Mutation score and frequency for each *com* gene. The mutation score (right panel) is the occurrence frequency of mutation for each gene normalized to the gene size (in kb). The single hit frequency (left panel) is defined as the frequency to have two or more mutated *com* genes for each strain bearing a mutated *com* gene. Bars are color-coded in agreement to the gene function in the DNA transformation process shown in Fig. 1B (pseudopilus, red; DNA transport, green; and DNA processing, blue). The color-coded vertical bars depict the mean value for the gene set involved in DNA docking, transport, or processing. The *recA* and *ssbB* genes have been dismissed from the analysis because *recA* is not specific for natural transformation and is required for the vegetative growth, while *ssbB* has a paralog (*ssbA*) that could bias the analysis. **(C)** Theoretical and empirical frequency of multiple inactivating mutations in a strain genome (semi-log scale). The observation curve shows the actual frequency for a strain to gather no, a single, or multiple gene mutations (up to eight genes hit) in the *com* gene set. The theoretical curve is considered as if the mutation events were independent. The inset provides a zoom on values ranging from zero to three gene disruptions/mutations in the *com* gene set (linear scale).

initiated, there is a selective pressure to massively inactivate *com* genes (Fig. 4C). Indeed, the actual frequency of mutations for one or two *com* genes is lower than expected with a theoretical model, while the occurrence of strains possessing four or more mutations is high above predictions where the mutation events would be fully independent (up to a 10 000-fold increase for an octuple inactivation). In this perspective, some genes like *comGB*, *comGD*, *comGG*, or *comFA* (but not *comX/sigH*) are rarely inactivated alone, suggesting that they accelerate the *com* gene decay or compensate toxicity due to other mutations (Fig. 4B, left panel). This reflects their important and conflicting role on competence and cell fitness.

To conclude, the comparative analyses of genomes represent a first hint that indicates whether a strain might be naturally transformable. Nevertheless, even in the scenario of an intact set of *com* genes, there is a necessity to validate that they do not harbor loss-of-function single-point mutations, are all fully functional, and can sustain an efficient transformation process.

Governing the master competence regulator

The next step to highlight the transformability of a specific strain requires to demonstrate whether the competence machinery might be coordinately produced under the (inducible) overexpression of an extra copy of the master regulator gene (ComX or SigH; Fig. 5, strategy I). Typically, this supplementary gene is borne onto a (multicopy) plasmid, i.e. transferred into the bacterium *via* an artificial (chemical or electric shocks) transformation (Blomqvist *et*

Table 1. Avowed transformable LAB.

Strain	Source	Set of late *com* gene	Activation of competence[a]			Reference
			Spontaneous	ComX^{++}	Pheromone	
Streptococcus thermophilus[b]						
LMD-9	Dairy	Complete	Moderate	High	High	Gardan *et al.* (2009), Fontaine *et al.* (2010a), Fontaine *et al.* (2010b)
LMG18311	Dairy	Complete	Low	High	High	Blomqvist *et al.* (2006), Fontaine *et al.* (2010a), Fontaine *et al.* (2010b), Gardan *et al.* (2009)
CNRZ1066	Dairy	Complete	No	High	High	Gardan *et al.* (2009), Fontaine *et al.* (2010a), Fontaine *et al.* (2010b)
DGCC7710	Dairy	Complete	Low	High	High	Fontaine *et al.* (2010b)
Streptococcus salivarius[b]						
ATCC25975	Human oral cavity	Complete	-	NT	High	Fontaine *et al.* (2010a)
JIM8777	Human oral cavity	Complete	-	NT	High	Fontaine *et al.* (2010a)
JIM8780	Human oral cavity	Complete	Low	NT	High	Fontaine *et al.* (2010a)
HSISS4	Human GI tract	Complete	-	NT	High	Mignolet *et al.* (2018)
Streptococcus macedonicus						
DSM15879T	Dairy	Unknown	Low	NT	High	Morrison *et al.* (2013)
139C	Dairy	Unknown	No	NT	High	Morrison *et al.* (2013)
Streptococcus infantarius[b]						
ATCC BAA-102T	Human GI tract	Complete	-	NT	High	Morrison *et al.* (2013)
11FA-1	Dairy	Complete	-	NT	High	Morrison *et al.* (2013)
AB2VB1-2	Dairy	Unknown	-	NT	High	Morrison *et al.* (2013)
AV2A(1)	Dairy	Unknown	-	NT	High	Morrison *et al.* (2013)
Lactococcus lactis* subsp. *lactis						
IL1403	Dairy	*comX**, *dprA**	-	Low	-P	Mulder *et al.* (2017)
KF147	Plant	Complete	-	High	-P	Mulder *et al.* (2017)
Lactococcus cremoris* subsp. *cremoris						
KW2	Plant	Complete	-	High	-P	David *et al.* (2017), Mulder *et al.* (2017)
Lactococcus raffinolactis						
LMG13098	Plant	Unknown	Low	NT	-P	Toussaint (unpublished work)
LMG14164	Bird	Unknown	Low	NT	-P	Toussaint (unpublished work)

[a]High $>10^{-4}$ transformation rate; moderate from 10^{-4} to 10^{-6} ; low $<10^{-6}$; - = not identified; NT = not tested; and -P = no identified pheromone.
[b]Only a subset of strains is listed in this table (mainly close genome-strains).

al. 2006, David *et al.* 2017, Fontaine *et al.* 2010b, Mulder *et al.* 2017). In parallel, the strain can be engineered to carry a reporter fusion (e.g. luciferase or GFP encoding genes) under the control of a late *com* promoter (e.g. P_{comGA} or P_{dprA}). This strategy first demonstrates whether the alloexpression of the competence master regulator impacts late *com* gene transcription. Moreover, it can also disclose the timing of competence extinction (early/mid/late exponential or stationary phase) that can diverge from species to species (see hereunder). Beside *S. thermophilus* (Blomqvist *et al.* 2006, Fontaine *et al.* 2010b), three recent examples lie in *L. lactis/cremoris* strains (Table 1) that sometimes encode a full set of *com* genes, but were non-naturally transformable in human hands until 2017. In contrast to closely related strains partly deficient for *com* genes (point mutations or indels), David *et al.* (2017) showed that *comX* overexpression in strain KW2 activates late *com* genes, and they were able to induce natural transformation, providing a streptomycin-resistant *rpsL* or miscellaneous deletion alleles as donor DNA. In parallel, Mulder *et al.* (2017) showed that two other *L. lactis* subsp. *lactis* strains (KF147 and IL1403) were also transformable with similar plasmid-borne inducible extra copy of *comX*. Additionally, deep sequencing of the transcriptional response in *comX/sigH*-overexpressing backgrounds or conditions gives a wider picture of the direct and indirect regulon under the master competence regulator control. This strategy was used in *L. sakei*, where the *sigH*-overexpression was reported to enhance transcription of the full set of late *com* genes (Schmid *et al.* 2012). Unfortunately, *L. sakei* was non-naturally transformable in this synthetic condition.

Beside gene overexpression, an alternative way to modulate the ComX/SigH steady-state level is to hijack proteins involved in the shut-off/repression mechanisms or the turnover rate (Fig. 5, strategy II). For instance, alteration of the core and accessory proteolysis machinery (i.e. Clp subunits and MecA, respectively) repre-

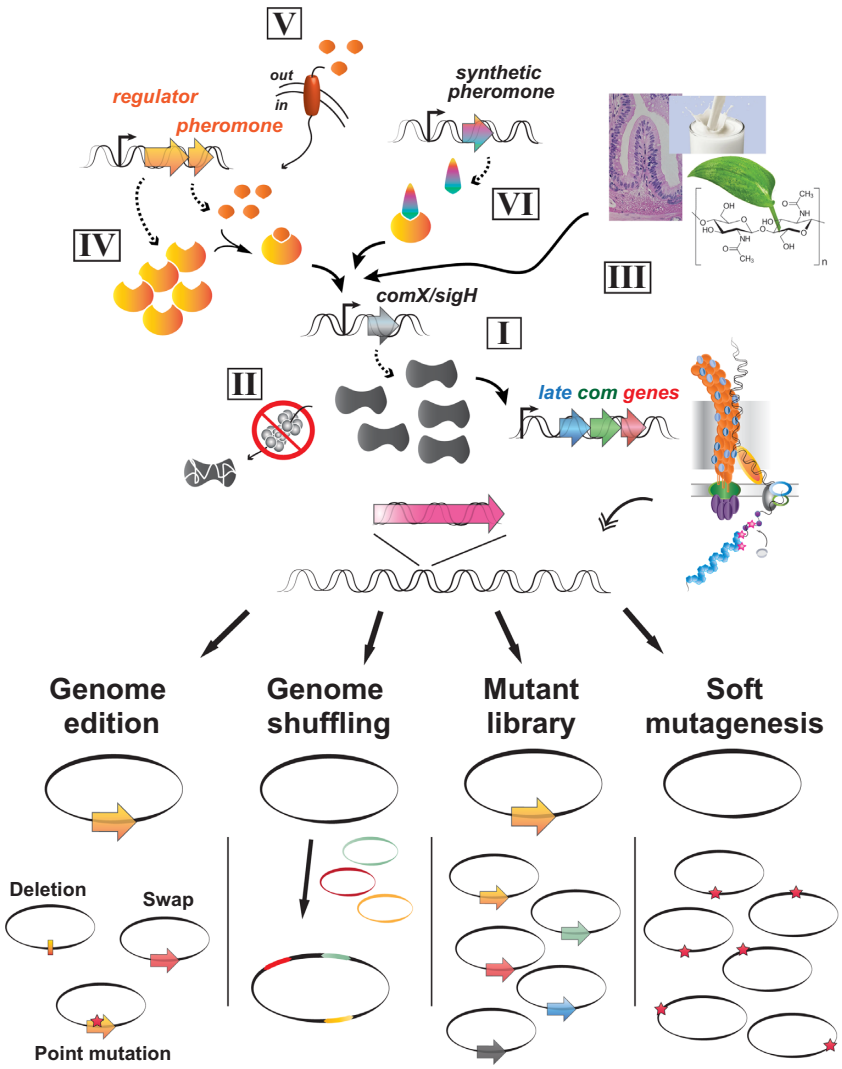

Figure 5. Competence activation strategies and bioediting outputs. Top panel: scheme depicting the six different levers to unlock competence activation and promote natural transformation. (I) Overexpression of *comX/sigH*. (II) Inhibition of factors that proteolyse or decrease the abundance of the master competence activator. (III) Natural or laboratory conditions that activate *comX/sigH* expression. (IV) Overexpression of *comX/sigH* activators (pheromones or sensors). (V) Exogenous supply of natural or synthetic pheromones that activate the *comX/sigH* expression. (VI) Synthetic screen to identify peptide-based molecules produced intracellularly and able to stimulate the *comX/sigH* expression. Bottom panel: potentialities of bioediting *via* natural transformation. Single genes can be swapped (with point-mutated alleles or nonrelated genes) or deleted. Large genomic reorganization or multiple gene exchange can be performed. Libraries of mutant for specific loci/genes can be easily generated. Competence activation increases the rate of spontaneous mutations (soft mutagenesis) that can be selected for a desired phenotype.

sents an extra control at the post-transcriptional level to increase the intracellular concentration of the master competence regulator. In *S. thermophilus*, it was demonstrated that the proteolysis-mediated stabilization of ComX can trigger competence in subpermissive conditions (Biornstad, Havarstein 2011, Boutry *et al.* 2012). This phenomenon is also partially true in *L. cremoris* subsp. *cremoris* KW2 (previously known as *L. lactis* subsp. *cremoris* KW2; David *et al.* 2017).

These strategies are highly informative and essential to accelerate/simplify genomic modifications in naturally transformable strains. Nonetheless, the alloproduction of ComX/SigH that drives a permanent or spread competence state is usually toxic and mutagenic, and could lead to undesired mutations or loci rearrangements all over the genome (David *et al.* 2017, Mulder *et al.* 2017). This is likely to generate an undesired genome instability or drift that could alter the nature of the microorganism. Thus, extra investigations are required to unveil the "natural" exosignals (e.g. molecules and physicochemical parameters) that stimulate competence activation.

Discovery of the master regulator guardians

Regulatory circuitries that control competence are intricately fine-tuned and differ between species, while the accurate timing of activation is somehow strain-specific. For instance, several bacterial species naturally turn on competence early during the exponential phase (e.g. *S. thermophilus*; Fontaine *et al.* 2013), while other clades pace competence according to the stationary phase (e.g. *B. subtilis*; Hoch 1993). A first glance at *comX/sigH* promoter might shed light on regulatory sequences, and therefore, the known transcriptional regulators that bind them. Competence is typically triggered in specific conditions encountered in the natural ecosystem of the bacterium. Usually, they deviate from laboratory environments or are difficult to recreate. The signals could range

from basic carbohydrates, such as monosaccharides, to complex (branched) biomolecules and media. *S. thermophilus* LMD-9 becomes naturally transformable in milk, i.e. the starting substrate used in industries for yogurt and cheese production (Gardan *et al.* 2013). Concerning *Vibrio cholerae*, Meibom *et al.* (2005) discovered that competence is induced by chitin, the main biopolymer forming the exoskeleton of crustacean that coexist with *V. cholera* in aquatic habitats. One-at-a-time testing of conditions that could arise in the natural biotope of a bug is laborious with uncertain chances of success. Instead, phenotypical microarrays allow to screen thousands of defined media elaborated with a range of pH, or media supplemented with nutrients (amino acids, sugars, nucleic acids, lipids, and vitamins), energy sources (carbon, nitrogen, phosphate, sulfur, and glycolysis/Krebs cycle intermediates), oligo-elements, biomolecules, osmolytes, antibiotics, chemicals, or xenobiotics (Fig. 5, strategy III). Growing reporter strains (e.g. harboring P_{comX}-*luxAB* constructs) in these diversified conditions is a powerful way to identify inducing molecules (and thereby, the cognate protein sensor or two-component system) and time-monitor competence activation at a high-throughput level.

In Gram-positive bacteria, competence regulation is intimately linked to the metabolic state and relies on a contribution from overlapping layers of quorum sensing, nutritional signals, and other stress responses (Nester, Stocker 1963, Dubnau 1991, Fontaine *et al.* 2015). Specifically in streptococci, ComX is under the exclusive direct control of either ComE (in mitis and anginosus groups, e.g. *S. pneumoniae and S. anginosus*, respectively) or ComR (in mutans, bovis, pyogenes, suis, and salivarius groups, e.g. *S. mutans*, *Streptococcus infantarius*, *S. pyogenes*, *Streptococcus suis*, and *S. salivarius*, respectively), two pheromone-responsive systems (Johnston *et al.* 2014). In both cases, the pheromone is a peptide produced, processed/matured and secreted outside of the cell envelope. However, ComE is a response regulator part of the two-component system that extracellularly detects the presence of the pheromone CSP (Competence-Stimulating Peptide; mature form of ComC) through the direct interaction with and the autophosphorylation of the membrane histidine kinase ComD, and the transfer of the phosphate moiety onto ComE (Pestova *et al.* 1996, Martin *et al.* 2013). In contrast, XIP (*comX*-Inducing Peptide, mature form of ComS), the cognate peptide pheromone of ComR, penetrates into the cytoplasmic space to accommodate the peptide-binding pocket of the cytosolic ComR, a protein of the RRNPP (Rgg-Rap-NprR-PrgX-PlcR) family (Fontaine *et al.* 2010a, Mashburn-Warren *et al.* 2010). Subsequent subtle conformational changes liberate the ComR DNA-binding domain and expose key residues involved in the selective recognition of specific nucleotides of the *comX* promoter (Shanker *et al.* 2016, Ledesma-Garcia *et al.* 2020, Talagas *et al.* 2016). Interestingly, this system was discovered in *S. thermophilus* while performing a transcriptomic profiling of a *comX* null-mutant. Following a study that highlighted the role of an unidentified pheromone-sensitive system controlling competence activation (Gardan *et al.* 2009), Fontaine *et al.* (2010a) realized that the inactivation of *comX* in conditions favoring competence (i.e. growth in chemically defined medium) bolstered the expression of the ComR direct regulon, namely the bacteriocin regulatory genes as well as the inducing pheromone gene *comS* itself (positive feedback loop). This data implies that a gene under ComX control negatively influences *comX* expression through the modulation of the ComR activation level (retro-control). Analogous feedback circuitries were observed for the close relative *S. pneumoniae* (see DprA bifunctionality, hereinabove) and *S. salivarius* (Mirouze *et al.* 2013, Mignolet *et al.* 2018). Therefore, in such a network configuration, upstream early *com* genes might be identified *via* ComX-associated phenotypes and unbiased high-throughput profiling techniques (e.g. deep sequencing or microarrays). In this respect, pheromone-responsive systems, either two-component or RRNPP proteins, and to a lesser extent transcriptional regulators, can be considered as promising candidates that would influence the steady-state level of master competence regulators. Mutations, bioengineering, or environmental conditions that drive the overproduction of the pheromone or the sensor might, therefore, increase the transformation rate and illuminate the process of competence development for fundamental research (Fig. 5, strategy IV). For instance, CovR is an environmental sensor directly repressing *comR* expression in the salivarius group of streptococci. In *S. thermophilus* LMD-9 strain, a natural single substitution of the conserved Asp98 (for a Glu) mimics a CovR unphosphorylated state that partially alleviates *comR* repression. This translates into higher ComR concentration into the cytoplasm and higher rate of transformation compared to other *S. thermophilus* or *S. salivarius* strains (Knoops *et al.* 2022). The identification of external signals modulating the phosphorylation state or activity of CovR will presumably magnify the transformation capabilities of some refractory strains.

Synthetic strategies to hijack regulatory circuitries

Usually, unveiling the "natural" or lab conditions that allow competence development in field strains is costly and often conducted in the long run. Alternatively, hacking the communication systems governing *comX*/*sigH* expression is a tantalizing approach that would facilitate competence induction (at the level of costs and time). This is particularly true for streptococci where both pheromones and sensors are identified, and where pheromones are ribosomally synthesized peptides of relatively small size and with no molecular decorations (Fontaine *et al.* 2015). Strikingly, the exogenous supplementation of synthetic pheromones enables a saturating competence response at extremely low doses (nanomolar range), independently of the sensor cellular location (outside-membrane interface or cytoplasmic compartment; Mignolet *et al.* 2016, Ledesma-Garcia *et al.* 2020, Yang *et al.* 2020; Fig. 5, strategy V). Since the discovery of the ComRS system in *S. thermophilus*, this pheromone-based strategy appears to be highly efficient. It was used several times in the past to expand competence activation to numerous strains of beneficial streptococci belonging to species *S. thermophilus*, *S. salivarius*, *S. infantarius*, and *Streptococcus macedonicus* (Table 1; Fontaine *et al.* 2010a, b, Morrison *et al.* 2013), as well as pathogenic streptococci including *S. mutans* (Mashburn-Warren *et al.* 2010), *S. sobrinus* (Li *et al.* 2021), *S. suis* (Zaccaria *et al.* 2014), and *S. pyogenes* (Mashburn-Warren *et al.* 2012, Marks *et al.* 2014). The current technologies in chemistry allow a fast and somewhat cheap production of conventional peptides (C- to N-terminus synthesis on solid surface). However, rapid advances in this field will definitively give a similar access in the future for "exotic" peptidomimetics that encompass L stereoisoform, unnatural, decorated, or bridged amino acids, and even circular and endocyclic peptides (Bedard, Biron 2018). We specifically think about the competence pheromones of *Bacillus* species, a hexapeptide comprising a modified tryptophan residue (Okada *et al.* 2005), that might also be chemically synthesized to activate competence when desired. However, some microorganisms are insensitive to their own pheromone in inappropriate conditions. A remarkable example is the *S. pyogenes* case that requires both ComS and an organized biofilm on epithelial cells to trigger com-

petence, showing how important is the natural environment to activate the competence regulatory cascade (Mashburn-Warren *et al.* 2012, Marks *et al.* 2014). Alternatively, several screens and rational design works have been successfully performed to identify linear and cyclic molecules of nonpeptide origin able to modulate the competence pathway in *S. pneumoniae* (Yang *et al.* 2020, Milly *et al.* 2021).

Interestingly, peptide-based signaling exhibits a 2-fold benefit. First, an array of variants can be designed and screened for their capacity to improve peptide stability, protease-resistance, diffusion coefficient, affinity/activability toward their cognate sensor, and ultimately the modulation of the transformation rate. Complementarily, random mutations in the pheromone coding gene to seek for optimized properties might generate a higher diversity of sequences compatible with high-throughput genetic screens (Fig. 5, strategy VI). Moreover, the rise of 3D structures available for regulators, in an apoform or in co-crystal with their cognate pheromone, illuminates the recognition at the atom scale between peptides and regulators (Talagas *et al.* 2016). It facilitates the rational design of more potent pheromones or pheromones with a revamped selectivity toward noncognate sensors, as was recently performed for streptococci of the salivarius group (Ledesma-Garcia *et al.* 2020). In a second scope, randomization-based genetic screens can elegantly shine a light on a broad range of peptide sequences, distant from the natural one, and capable of activating signaling regulators. Concerning competence pheromones that impose a high fitness cost and hinder the cell cycle, such strategies can reveal sequences that would be counter-selected in nature because they amplify/prolong the competence state. From another angle, many sensors are only assumed to respond to peptides because their natural peptide has not yet been disclosed. In a recent study, researchers generated a library of short genes for which the last seven codons were randomized, and they managed to unveil a set of peptides that efficiently activate ComR-like regulators unrelated to competence and for which the cognate pheromone stays elusive (Mignolet *et al.* 2019). Sequence comparison extracted common properties (i.e. amino acid positioning) and may give some hint to fish the natural molecular cue. This synthetic pheromone approach is particularly convenient to stimulate a system on demand, even in absence of known signals (Fig. 5, strategy VI). Albeit designed for cytoplasmic sensors, it could be presumably amenable to extracellular sensing (e.g. two-component system).

Competence benefits for LAB optimization

Nowadays, competence for natural transformation is routinely exploited to genetically modify streptococcal strains in the scope of fundamental research. This includes the engineering of deletion, overexpression, or single-point mutants, the design of genetic libraries for synthetic biology screens, the optimization of cell fitness, and the clustering of interesting traits into a unique microbial chassis. The two main handling advantages of competence over other natural transfer mechanisms such as transduction and conjugation or genome-editing tools such as CRISPR-Cas and λ-red-like recombineering are definitely that it does not require steps of cloning, and that the transferable information can be provided as a mere linear fragment of double-stranded DNA (even though entire circular plasmids can be acquired *via* natural transformation). The DNA template can either directly originate from genomic extraction or alternatively be native or recombinant genes amplified by PCR, or chemically synthesized in DNA blocks. The only requirement for chromosomal and extrachromosomal insertions is two DNA homology regions (ideally more than 500 bp) that flank the fragment to be inserted, one on each side. In the context of directed evolution and positive selection, competence is particularly efficient and trendy for DNA shuffling. From a linear DNA library of (active) mutated alleles (or functional homologous genes in the case of "family shuffling"), this technique uses the iteration of PCR-random fragmentation cycles to generate novel combinations (hybrids or swaps) and increase the library diversity. Another substantial added value of competence relies on the fact that multiple pieces of DNA can be transferred simultaneously and drive a global genomic rearrangement (Dalia *et al.* 2017). Therefore, the natural transformation technology is a strain-inherent property that tremendously accelerates and potentiates the genome reshaping, and can produce mutants in a single day.

The transformation rate is definitively strain-specific and relies on intrinsic signaling and regulatory networks. Puzzlingly, it strongly vary across strains of the same species, or even from time to time in a given strain. It can range from 50% in high competence performers such as *S. pneumoniae* R6 (Johnston *et al.* 2020), *Streptococcus mutans* UA159 (Lemme *et al.* 2011), or *Streptococcus thermophilus* LMD-9 to very low value (10^{-4}) in most cases. To select gene acquisition events in poorly transformable strains, a resistance cassette can be linked to the recombinant DNA, and the cognate antibiotic is supplemented in the growth medium to neutralize unmodified cells. On the top of it, counter-selection markers (e.g. *oroP*, *pheS*, or *sacBR*) or site-specific recombinases (Cre) can be combined to introduce or remove alleles in a scarless, or at least a marker-free, output (Fontaine *et al.* 2010b, Dorrazehi *et al.* 2020, Mignolet *et al.* 2018). These methods provide extraordinary opportunities in fundamental research to shed light on bacterial physiological processes. Once robust and convenient stimulating conditions have been developed, natural transformation is the most efficient strategy to select the acquisition of novel phenotypical traits. Particularly convenient for high-throughput screens, hypercompetent LAB can accommodate genetic libraries of a large diversity (up to billion clones) and could be used in the context of bacterial surface display or antigen presentation (Hu *et al.* 2010, Dorrazehi *et al.* 2020, Lecomte *et al.* 2014).

In the context of natural DNA transfers without the use of an antibiotic marker, the major challenges are to use highly transformable strains (Blomqvist *et al.* 2010) or to associate a desired trait with a selectable (growth improvement, color- or texture-marked) phenotype, and to dispose of a natural donor DNA that exhibits a sufficient homology to sustain recombination in the recipient cells. This strategy, reminiscent of self-cloning, was applied to two examples in *S. thermophilus*, the *prtS* locus, and the histidine prototrophy. PrtS is a cell wall-anchored enzyme that fulfills a nutritional role by proteolyzing the milk casein in short importable oligopeptides. Acquisition of the *prtS* genomic island in *prtS*-deficient strains speeds up the cellular growth and subsequently the milk acidification process, leading to a reduction in the industrial process costs (Dandoy *et al.* 2011). The conversion of slow-acidifying strains into fast-acidifying derivatives provides a genuine selective advantage easily visualized on differential agar plate (big colonies surrounded by a yellow halo of acidification). The transfer of prototrophic traits (his^+) into auxotrophic strains (his^-) is even more straightforward. Genomic DNA from the prototrophic donor has to be incubated in competence-inducing conditions with the recipient strains disabled for a biosynthetic pathway. Cells are next plated on selective media depleted for the specific essential amino acids and tested for growth (Fontaine *et al.* 2010b).

As an additional benefit, competence accelerates the evolutionary clock of microorganisms even in the absence of foreign DNA (Grist, Butler 1983, Martin *et al.* 2020). Indeed, competence induction increases the mutation rate, presumably because proteins involved in the recombination process could display an intrinsic replicative error-prone molecular behavior. In comparison with harsh physical (e.g. UVs) or chemical (e.g. ethyl methanesulfonate) mutagenesis that literally riddle the genome with multiple replicative errors, competence activation could, therefore, be exploited as a soft treatment to promote the emergence of spontaneous mutants.

Concluding remarks

The LAB world is a wide reservoir in which we can tap into to elaborate innovative solutions and face current and future major issues (e.g. soil desertification, multidrug resistance, and food/energy availability; Lamont *et al.* 2017, Borner *et al.* 2019, Bron, Kleerebezem 2018, Mathur *et al.* 2020). The synthetic biology and biotechnologies even expand the possibility landscape because they facilitate (temporally and financially) the fine-tuning of microfactories, either through rational design or randomization-based and evolution strategies. So far, competence for natural transformation is definitely the most convenient tool to expand our knowledge about LAB because it permits to convert bacterial genomes, from single-point mutation to large chromosomic rearrangements. A deep understanding of molecular behaviors that govern the whole process in every transformable microorganism is, thence, the keystone to concretize the full potential of LAB for human need.

The conclusive demonstration of strain transformability is usually a long and demanding path. Even for *S. thermophilus* and *S. salivarius*, two species nowadays routinely transformed on lab benches with incredibly simple protocols, more than 10 years of intensive works have been carried out by several research units. With a stepwise strategy based on *in silico* analyses, genetics, synthetic biology, and phenotypic characterization, future researches will be able to rationally minimize the effort to reach this aim. The use of overexpressing/deletion strains or artificially transformed (i.e. *via* enzymes, chemicals, or electroshocks) mutants are excellent premises to finely unravel the mechanistic of competence activation. Nevertheless, the identification of the natural signal (molecule or stress) that dictates the transformability is the final objective to fully unleash the potential of natural transformation.

In the future, the use of competence as a natural genome diversification method in bacteria, analogously to plant or animal breeding, will benefit from the microbial diversity, i.e. present in the giant cryostock libraries established by a plethora of academic, medical, and industrial organizations over the world (Ryan *et al.* 2021) Therefore, there is a flagrant necessity to pursue our effort and widely collect and sequence as possible the highest diversity of microorganisms. They will be the solid basis to combine exciting properties in second generation of improved LAB strains.

Supplementary data

Supplementary data are available at FEMSRE online.

Supplementary Table 1. Conservation of competence genes across the *Streptococcus* genus.

Supplementary Table 2. Conservation of competence genes across the *Lactobacillus* genus.

Supplementary Table 3. Conservation of competence genes across the *Lactococcus* genus.

Supplementary Table 4. Conservation of competence genes across the *Leuconostoc* and *Oenococcus* genera.

Supplementary Table 5. Conservation of competence genes across the *Pediococcus* genus.

Supplementary Appendix. Methods and analysis strategies used to identify the competence gene sets in beneficial LAB genomes, to estimate the mutation frequency in competence genes, and to generate the ComX/SigH phylogenetic tree.

Funding

This work was supported by the Belgian National Fund for Scientific Research (FNRS; PDR T.0110.18 to Pa.H.), the Federation Wallonia-Brussels (ARC 17/22-084 to Pa.H.), and the Belgian Science Policy Office (BELSPO; IAP P17/28 to Pa.H.). J.M. holds a H2020-MSCA-IF-2020 fellowship (grant number 101018461). Pa.H. is Research Director at the FNRS.

Conflict of interest statement. F.T., C.F., Ph.H., and Pa.H., declare that they are listed as inventors on patent(s) or patent application(s) related to transforming bacteria through natural competence.

References

Attaiech L, Olivier A, Mortier-Barriere I *et al.* Role of the single-stranded DNA-binding protein SsbB in pneumococcal transformation: maintenance of a reservoir for genetic plasticity. *PLos Genet* 2011;**7**:e1002156.

Bedard F, Biron E. Recent progress in the chemical synthesis of class II and S-Glycosylated bacteriocins. *Front Microbiol* 2018;**9**:1048.

Berge M, Mortier-Barriere I, Martin B *et al.* Transformation of *Streptococcus pneumoniae* relies on DprA- and RecA-dependent protection of incoming DNA single strands. *Mol Microbiol* 2003;**50**: 527–36.

Berge MJ, Kamgoue A, Martin B *et al.* Midcell recruitment of the DNA uptake and virulence nuclease, EndA, for pneumococcal transformation. *PLoS Pathog* 2013;**9**:e1003596.

Berge MJ, Mercy C, Mortier-Barriere I *et al.* A programmed cell division delay preserves genome integrity during natural genetic transformation in *Streptococcus pneumoniae*. *Nat Commun* 2017;**8**: 1621.

Biornstad TJ, Havarstein LS. ClpC acts as a negative regulator of competence in *Streptococcus thermophilus*. *Microbiology* 2011;**157**: 1676–84.

Blomqvist T, Steinmoen H, Havarstein LS. A food-grade site-directed mutagenesis system for *Streptococcus thermophilus* LMG 18311. *Lett Appl Microbiol* 2010;**50**:314–9.

Blomqvist T, Steinmoen H, Havarstein LS. Natural genetic transformation: a novel tool for efficient genetic engineering of the dairy bacterium *Streptococcus thermophilus*. *Appl Environ Microbiol* 2006;**72**:6751–6.

Bolotin A, Quinquis B, Renault P *et al.*. Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat Biotechnol* 2004;**22**:1554–8.

Borner RA, Kandasamy V, Axelsen AM *et al.* Genome editing of lactic acid bacteria: opportunities for food, feed, pharma and biotech. *FEMS Microbiol Lett* 2019;**366**:fny291.

Boutry C, Wahl A, Delplace B *et al.* Adaptor protein MecA is a negative regulator of the expression of late competence genes in *Streptococcus thermophilus*. *J Bacteriol* 2012;**194**:1777–88.

Briley K, Dorsey-Oresto A, Prepiak P *et al.* The secretion ATPase ComGA is required for the binding and transport of transforming DNA. *Mol Microbiol* 2011;**81**:818–30.

Bron PA, Kleerebezem M. Lactic acid bacteria for delivery of endogenous or engineered therapeutic molecules. *Front Microbiol* 2018;**9**:1821.

Bron PA, Marcelli B, Mulder J *et al.* Renaissance of traditional DNA transfer strategies for improvement of industrial lactic acid bacteria. *Curr Opin Biotechnol* 2019;**56**:61–8.

Chen I, Dubnau D. DNA uptake during bacterial transformation. *Nat Rev Microbiol* 2004;**2**:241–9.

Chen I, Provvedi R, Dubnau D. A macromolecular complex formed by a pilin-like protein in competent *Bacillus subtilis*. *J Biol Chem* 2006;**281**:21720–7.

Cho SW, Yim J, Seo SW. Engineering tools for the development of recombinant lactic acid bacteria. *Biotechnol J* 2020;**15**:e1900344.

Chowdhury MY, Li R, Kim JH *et al.* Mucosal vaccination with recombinant *Lactobacillus casei*-displayed CTA1-conjugated consensus matrix protein-2 (sM2) induces broad protection against divergent influenza subtypes in BALB/c mice. *PLoS ONE* 2014;**9**:e94051.

Chung YS, Dubnau D. All seven comG open reading frames are required for DNA binding during transformation of competent *Bacillus subtilis*. *J Bacteriol* 1998;**180**:41–5.

Claverys JP, Martin B, Polard P. The genetic transformation machinery: composition, localization, and mechanism. *FEMS Microbiol Rev* 2009;**33**:643–56.

Dalia TN, Yoon SH, Galli E *et al.* Enhancing multiplex genome editing by natural transformation (MuGENT) via inactivation of ssDNA exonucleases. *Nucleic Acids Res* 2017;**45**:7527–37.

Dandoy D, Fremaux C, de Frahan MH *et al.* The fast milk acidifying phenotype of *Streptococcus thermophilus* can be acquired by natural transformation of the genomic island encoding the cell-envelope proteinase PrtS. *Microb Cell Fact* 2011;**10**:S21.

David B, Radziejwoski A, Toussaint F *et al.* Natural DNA transformation is functional in *Lactococcus lactis* subsp. *cremoris* KW2. *Appl Environ Microbiol* 2017;**83**:e01074–17.

De Filippis F, Pasolli E, Ercolini D. The food-gut axis: lactic acid bacteria and their link to food, the gut microbiome and human health. *FEMS Microbiol Rev* 2020;**44**:454–89.

de Kraker ME, Stewardson AJ, Harbarth S. Will 10 million people die a year due to antimicrobial resistance by 2050? *PLoS Med* 2016;**13**:e1002184.

Desai BV, Morrison DA. Transformation in *Streptococcus pneumoniae*: formation of eclipse complex in a *coiA* mutant implicates CoiA in genetic recombination. *Mol Microbiol* 2007;**63**:1107–17.

Diallo A, Foster HR, Gromek KA *et al.* Bacterial transformation: ComFA is a DNA-dependent ATPase that forms complexes with ComFC and DprA. *Mol Microbiol* 2017;**105**:741–54.

Dorrazehi GM, Worms S, Chirakadavil JB *et al.* Building scarless gene libraries in the chromosome of bacteria. In: *Peptide and Protein Engineering: From Concepts to Biotechnological Applications*, New York, NY: Springer, 2020, 189–211.

Dragosits M, Mattanovich D. Adaptive laboratory evolution – principles and applications for biotechnology. *Microb Cell Fact* 2013;**12**:64.

Draskovic I, Dubnau D. Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. *Mol Microbiol* 2005;**55**:881–96.

Dubnau D, Blokesch M. Mechanisms of DNA uptake by naturally competent bacteria. *Annu Rev Genet* 2019;**53**:217–37.

Dubnau D. Genetic competence in *Bacillus subtilis*. *Microbiol Rev* 1991;**55**:395–424.

Fagerlund A, Granum PE, Havarstein LS. *Staphylococcus aureus* competence genes: mapping of the SigH, ComK1 and ComK2 regulons by transcriptome sequencing. *Mol Microbiol* 2014;**94**:557–79.

Fontaine L, Boutry C, de Frahan MH *et al.* A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J Bacteriol* 2010a;**192**:1444–54.

Fontaine L, Dandoy D, Boutry C *et al.* Development of a versatile procedure based on natural transformation for marker-free targeted genetic modification in *Streptococcus thermophilus*. *Appl Environ Microbiol* 2010b;**76**:7870–7.

Fontaine L, Goffin P, Dubout H *et al.* Mechanism of competence activation by the ComRS signalling system in streptococci. *Mol Microbiol* 2013;**87**:1113–32.

Fontaine L, Wahl A, Flechard M *et al.* Regulation of competence for natural transformation in streptococci. *Infect Genet Evol* 2015;**33**:343–60.

Gardan R, Besset C, Gitton C *et al.* Extracellular life cycle of ComS, the competence-stimulating peptide of *Streptococcus thermophilus*. *J Bacteriol* 2013;**195**:1845–55.

Gardan R, Besset C, Guillot A *et al.* The oligopeptide transport system is essential for the development of natural competence in *Streptococcus thermophilus* strain LMD-9. *J Bacteriol* 2009;**191**:4647–55.

Garneau JE, Dupuis ME, Villion M *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 2010;**468**:67–71.

Gonzalez-Torres P, Rodriguez-Mateos F, Anton J *et al.* Impact of homologous recombination on the evolution of prokaryotic core genomes. *mBio* 2019;**10**:e02494–18.

Griffith F. The significance of pneumococcal types. *J Hyg* 1928;**27**:113–59.

Grist RW, Butler LO. Effect of transforming DNA on growth and frequency of mutation of *Streptococcus pneumoniae*. *J Bacteriol* 1983;**153**:153–62.

Haijema BJ, Hahn J, Haynes J *et al.* ComGA-dependent checkpoint limits growth during the escape from competence. *Mol Microbiol* 2001;**40**:52–64.

Haustenne L, Bastin G, Hols P *et al.* Modeling of the ComRS Signaling Pathway Reveals the Limiting Factors Controlling Competence in *Streptococcus thermophilus*. *Front Microbiol* 2015;**6**:1413.

Hoch JA. Regulation of the onset of the stationary phase and sporulation in *Bacillus subtilis*. *Adv Microb Physiol* 1993;**35**:111–33.

Hols P, Ledesma-Garcia L, Gabant P *et al.* Mobilization of microbiota commensals and their bacteriocins for therapeutics. *Trends Microbiol* 2019;**27**:690–702.

Hu S, Kong J, Kong W *et al.* Characterization of a novel LysM domain from *Lactobacillus fermentum* bacteriophage endolysin and its use as an anchor to display heterologous proteins on the surfaces of lactic acid bacteria. *Appl Environ Microbiol* 2010;**76**:2410–8.

Hugenholtz J, Smid EJ. Nutraceutical production with food-grade microorganisms. *Curr Opin Biotechnol* 2002;**13**:497–507.

Inamine GS, Dubnau D. ComEA, a *Bacillus subtilis* integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. *J Bacteriol* 1995;**177**:3045–51.

Johansen E. Use of natural selection and evolution to develop new starter cultures for fermented foods. *Ann Rev Food Sci Technol* 2018;**9**:411–28.

Johnsborg O, Eldholm V, Havarstein LS. Natural genetic transformation: prevalence, mechanisms and function. *Res Microbiol* 2007;**158**:767–78.

Johnston C, Martin B, Fichant G *et al.* Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat Rev Microbiol* 2014;**12**:181–96.

Johnston CH, Soulet AL, Berge M *et al.* The alternative sigma factor sigma(X) mediates competence shut-off at the cell pole in *Streptococcus pneumoniae*. *eLife* 2020;**9**:e62907.

Jorth P, Whiteley M. An evolutionary link between natural transformation and CRISPR adaptive immunity. *mBio* 2012;**3**:e00309–12.

Kaspar J, Shields RC, Burne RA. Competence inhibition by the XrpA peptide encoded within the *comX* gene of *Streptococcus mutans*. *Mol Microbiol* 2018;**109**:345–64.

Kelleher P, Bottacini F, Mahony J *et al.* Comparative and functional genomics of the *Lactococcus lactis* taxon; insights into evolution and niche adaptation. *BMC Genomics* 2017;**18**:267.

Knoops A, Vande Capelle F, Fontaine L *et al.* The CovRS environmental sensor directly controls the ComRS signaling system to orchestrate competence bimodality in Salivarius Streptococci. *mBio* 2022;**13**:e03125–03121.

Knoppel A, Knopp M, Albrecht LM *et al.* Genetic adaptation to growth under laboratory conditions in *Escherichia coli* and *Salmonella enterica*. *Front Microbiol* 2018;**9**:756.

Kramer N, Hahn J, Dubnau D. Multiple interactions among the competence proteins of *Bacillus subtilis*. *Mol Microbiol* 2007;**65**:454–64.

Lam T, Ellison CK, Eddington DT *et al.* Competence pili in *Streptococcus pneumoniae* are highly dynamic structures that retract to promote DNA uptake. *Mol Microbiol* 2021;**116**:381–396.

Lamont JR, Wilkins O, Bywater-EkegÃrd M. From yogurt to yield: potential applications of lactic acid bacteria in plant production. *Soil Biol Biochem* 2017;**111**:1–9.

Laurenceau R, Pehau-Arnaudet G, Baconnais S *et al.* A type IV pilus mediates DNA binding during natural transformation in *Streptococcus pneumoniae*. *PLoS Pathog* 2013;**9**:e1003473.

Lecomte X, Gagnaire V, Briard-Bion V *et al.* The naturally competent strain *Streptococcus thermophilus* LMD-9 as a new tool to anchor heterologous proteins on the cell surface. *Microb Cell Fact* 2014;**13**:82.

Ledesma-Garcia L, Thuillier J, Guzman-Espinola A *et al.* Molecular dissection of pheromone selectivity in the competence signaling system ComRS of streptococci. *Proc Natl Acad Sci* 2020;**117**:7745–54.

Lemme A, Grobe L, Reck M *et al.* Subpopulation-specific transcriptome analysis of competence-stimulating-peptide-induced *Streptococcus mutans*. *J Bacteriol* 2011;**193**:1863–77.

Li JW, Wyllie RM, Jensen PA. A novel competence pathway in the oral pathogen *Streptococcus sobrinus*. *J Dent Res* 2021;**100**:542–8.

Luo P, Morrison DA. Transient association of an alternative sigma factor, ComX, with RNA polymerase during the period of competence for genetic transformation in *Streptococcus pneumoniae*. *J Bacteriol* 2003;**185**:349–58.

Makarova K, Slesarev A, Wolf Y *et al.* Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci* 2006;**103**:15611–6.

Marks LR, Mashburn-Warren L, Federle MJ *et al.* *Streptococcus pyogenes* biofilm growth *in vitro* and *in vivo* and its role in colonization, virulence, and genetic exchange. *J Infect Dis* 2014;**210**:25–34.

Martin B, Quentin Y, Fichant G *et al.* Independent evolution of competence regulatory cascades in streptococci? *Trends Microbiol* 2006;**14**:339–45.

Martin B, Soulet AL, Mirouze N *et al.* ComE/ComE~P interplay dictates activation or extinction status of pneumococcal X-state (competence). *Mol Microbiol* 2013;**87**:394–411.

Martin HA, Kidman AA, Socea J *et al.* The *Bacillus Subtilis* K-state promotes stationary-phase mutagenesis via oxidative damage. *Genes* 2020;**11**:190.

Mashburn-Warren L, Morrison DA, Federle MJ. A novel double-tryptophan peptide pheromone controls competence in *Streptococcus* spp. via an Rgg regulator. *Mol Microbiol* 2010;**78**:589–606.

Mashburn-Warren L, Morrison DA, Federle MJ. The cryptic competence pathway in *Streptococcus pyogenes* is controlled by a peptide pheromone. *J Bacteriol* 2012;**194**:4589–600.

Mathur H, Beresford TP, Cotter PD. Health benefits of lactic acid bacteria (LAB) fermentates. *Nutrients* 2020;**12**:1679.

Meibom KL, Blokesch M, Dolganov NA *et al.* Chitin induces natural competence in *Vibrio cholerae*. *Science* 2005;**310**:1824–7.

Mejean V, Claverys JP. Polarity of DNA entry in transformation of *Streptococcus pneumoniae*. *Mol Gen Genet* 1988;**213**:444–8.

Midon M, Schafer P, Pingoud A *et al.* Mutational and biochemical analysis of the DNA-entry nuclease EndA from *Streptococcus pneumoniae*. *Nucleic Acids Res* 2011;**39**:623–34.

Mignolet J, Cerckel G, Damoczi J *et al.* Subtle selectivity in a pheromone sensor triumvirate desynchronizes competence and predation in a human gut commensal. *eLife* 2019;**8**:e47139.

Mignolet J, Fontaine L, Kleerebezem M *et al.* Complete genome sequence of *Streptococcus salivarius* HSISS4, a human commensal bacterium highly prevalent in the digestive tract. *Genome Announc* 2016;**4**:e01637–01615.

Mignolet J, Fontaine L, Sass A *et al.* Circuitry rewiring directly couples competence to predation in the gut dweller *Streptococcus salivarius*. *Cell Rep* 2018;**22**:1627–38.

Milly TA, Engler ER, Chichura KS *et al.* Harnessing multiple, non-proteogenic substitutions to optimize CSP:comD hydrophobic interactions in group 1 *Streptococcus pneumoniae*. *ChemBioChem* 2021;**22**:1940–1947.

Mirouze N, Berge MA, Soulet AL *et al.* Direct involvement of DprA, the transformation-dedicated RecA loader, in the shut-off of pneumococcal competence. *Proc Natl Acad Sci* 2013;**110**:E1035–1044.

Moon AF, Midon M, Meiss G *et al.* Structural insights into catalytic and substrate binding mechanisms of the strategic EndA nuclease from *Streptococcus pneumoniae*. *Nucleic Acids Res* 2011;**39**:2943–53.

Morrison DA, Guedon E, Renault P Competence for natural genetic transformation in the *Streptococcus bovis* group streptococci *S. infantarius* and *S. macedonicus*. *J Bacteriol* 2013;**195**:2612–20.

Morrison DA, Mortier-Barriere I, Attaiech L *et al.* Identification of the major protein component of the pneumococcal eclipse complex. *J Bacteriol* 2007;**189**:6497–500.

Mortier-Barriere I, Velten M, Dupaigne P *et al.* A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. *Cell* 2007;**130**:824–36.

Mulder J, Wels M, Kuipers OP *et al.* Unleashing natural competence in *Lactococcus lactis* by induction of the competence regulator ComX. *Appl Environ Microbiol* 2017;**83**:e01320–17.

Muschiol S, Aschtgen MS, Nannapaneni P *et al.* Gram-positive type IV pili and competence. *Microbiol Spectr* 2019;**7**:0011–2018.

Muschiol S, Erlendsson S, Aschtgen MS *et al.* Structure of the competence pilus major pilin ComGC in *Streptococcus pneumoniae*. *J Biol Chem* 2017;**292**:14134–46.

Nester EW, Stocker BA. Biosynthetic latency in early stages of deoxyribonucleic acid transformation in *Bacillus Subtilis*. *J Bacteriol* 1963;**86**:785–96.

Okada M, Sato I, Cho SJ *et al.* Structure of the *Bacillus subtilis* quorum-sensing peptide pheromone ComX. *Nat Chem Biol* 2005;**1**:23–4.

Olson AB, Kent H, Sibley CD *et al.* Phylogenetic relationship and virulence inference of *Streptococcus anginosus* group: curated annotation and whole-genome comparative analysis support distinct species designation. *BMC Genomics* 2013;**14**:895.

Opdyke JA, Scott JR, Moran CP. A secondary RNA polymerase sigma factor from *Streptococcus pyogenes*. *Mol Microbiol* 2001;**42**:495–502.

Papagianni M. Metabolic engineering of lactic acid bacteria for the production of industrially important compounds. *Comput Struct Biotechnol J* 2012;**3**:e201210003.

Pestova EV, Havarstein LS, Morrison DA. Regulation of competence for genetic transformation in *Streptococcus pneumoniae* by an auto-induced peptide pheromone and a two-component regulatory system. *Mol Microbiol* 1996;**21**:853–62.

Peterson SN, Sung CK, Cline R *et al.* Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. *Mol Microbiol* 2004;**51**: 1051–70.

Provvedi R, Dubnau D. ComEA is a DNA receptor for transformation of competent *Bacillus subtilis*. *Mol Microbiol* 1999;**31**:271–80.

Quevillon-Cheruel S, Campo N, Mirouze N *et al.* Structure-function analysis of pneumococcal DprA protein reveals that dimerization is crucial for loading RecA recombinase onto DNA during transformation. *Proc Natl Acad Sci* 2012;**109**:E2466–2475.

Reck M, Tomasch J, Wagner-Dobler I. The alternative sigma factor SigX controls bacteriocin synthesis and competence, the two quorum sensing regulated traits in *Streptococcus mutans*. *PLos Genet* 2015;**11**:e1005353.

Ryan MJ, Schloter M, Berg G *et al.* Development of microbiome biobanks - challenges and opportunities. *Trends Microbiol* 2021;**29**:89–92.

Salvadori G, Junges R, Morrison DA *et al.* Competence in *Streptococcus pneumoniae* and close commensal relatives: mechanisms and implications. *Front Cell Infect Microbiol* 2019;**9**:94.

Schmid S, Bevilacqua C, Crutz-Le Coq AM. Alternative sigma factor sigmaH activates competence gene expression in *Lactobacillus sakei*. *BMC Microbiol* 2012;**12**:32.

Shanker E, Federle MJ.Quorum sensing regulation of competence and bacteriocins in *Streptococcus pneumoniae* and *mutans*. *Genes* 2017;**8**:15.

Shanker E, Morrison DA, Talagas A *et al.* Pheromone recognition and selectivity by ComR proteins among *Streptococcus* Species. *PLoS Pathog* 2016;**12**:e1005979.

Talagas A, Fontaine L, Ledesma-Garcia L *et al.* Structural insights into streptococcal competence regulation by the cell-to-cell communication system ComRS. *PLoS Pathog* 2016;**12**: e1005980.

Tian XL, Dong G, Liu T *et al.* MecA protein acts as a negative regulator of genetic competence in *Streptococcus mutans*. *J Bacteriol* 2013;**195**:5196–206.

Underhill SAM, Shields RC, Burne RA *et al.* Carbohydrate and PepO control bimodality in competence development by *Streptococcus mutans*. *Mol Microbiol* 2019;**112**:1388–402.

Van Braeckel-Budimir N, Haijema BJ, Leenhouts K. Bacterium-like particles for efficient immune stimulation of existing vaccines and new subunit vaccines in mucosal applications. *Front Immunol* 2013;**4**:282.

van Duin D, Paterson DL. Multidrug-resistant bacteria in the community: trends and lessons learned. *Infect Dis Clin North Am* 2016;**30**:377–90.

Wahl A, Servais F, Drucbert AS *et al.* Control of natural transformation in salivarius Streptococci through specific degradation of sigmaX by the MecA-ClpCP protease complex. *J Bacteriol* 2014;**196**:2807–16.

Yang Y, Lin J, Harrington A *et al.* Designing cyclic competence-stimulating peptide (CSP) analogs with pan-group quorum-sensing inhibition activity in *Streptococcus pneumoniae*. *Proc Natl Acad Sci* 2020;**117**:1689–99.

Zaccaria E, van Baarlen P, de Greeff A *et al.* Control of competence for DNA transformation in *Streptococcus suis* by genetically transferable pherotypes. *PLoS ONE* 2014;**9**:e99394.

Zheng J, Wittouck S, Salvetti E *et al.* A taxonomic note on the genus *Lactobacillus*: description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *Int J Syst Evol Microbiol* 2020;**70**: 2782–858.