Functional genomics of lactic acid bacteria: from food to health

François P Douillard1, Willem M de Vos1,2,3*

From 11th International Symposium on Lactic Acid Bacteria
Egmond aan Zee, the Netherlands. 31 August - 4 September 2014

Abstract
Genome analysis using next generation sequencing technologies has revolutionized the characterization of lactic acid bacteria and complete genomes of all major groups are now available. Comparative genomics has provided new insights into the natural and laboratory evolution of lactic acid bacteria and their environmental interactions. Moreover, functional genomics approaches have been used to understand the response of lactic acid bacteria to their environment. The results have been instrumental in understanding the adaptation of lactic acid bacteria in artisanal and industrial food fermentations as well as their interactions with the human host. Collectively, this has led to a detailed analysis of genes involved in colonization, persistence, interaction and signaling towards the human host and its health. Finally, massive parallel genome re-sequencing has provided new opportunities in applied genomics, specifically in the characterization of novel non-GMO strains that have potential to be used in the food industry. Here, we provide an overview of the state of the art of these functional genomics approaches and their impact in understanding, applying and designing lactic acid bacteria for food and health.

Introduction & outline
Lactic acid bacteria (LAB) and humans share a long and intricate history. Well known are the first food fermentations reported in ancient times that contributed to the preservation and quality improvement of raw plant, meat and milk substrates. Most likely the transition from hunter-gatherers to an agricultural lifestyle, some 10,000 years ago, contributed to the further development of these food fermentations that are now practiced worldwide on an industrial scale. However, our interactions with LAB are more intimate and have a much longer history than the food fermentations that were initiated by the LAB present at that time (Figure 1). In addition to many plants and animals, the human body is also colonized by LAB and early culturing studies already documented the presence of LAB at different locations, e.g. the gastro-intestinal tract or the oral cavity [1]. However, many microbes cannot yet be cultured and this also holds for LAB [2]. Until recently, technological limitations precluded the global characterization of human microbiota in terms of composition, diversity and dynamics. Massive parallel sequencing and other high throughput approaches have offered novel ways to explore and examine the microbiota from different human body cavities [3-5]. Much attention has been given to the human gastro-intestinal (GI) tract but the number of endogenous (autochtonous) LAB in the human system is rather low (Douillard and De Vos, in press; see also below). This contrasts with many animals where the GI-tract is a well-established habitat for high numbers of endogenous LAB, such as the fore-stomach of mice and other rodents, as well as the crop of chicken and other birds [6,7]. Hence, these animal systems, similar to many plants that are colonized with LAB in the phyllosphere, may constitute reservoirs for LAB found in food fermentations or even the human body (Figure 1).

In retrospect, it may be argued that the low level of endogenous LAB in human explains the impact of passenger (allochtonous) LAB on the human host, as exemplified with LAB that are marketed as probiotics and after consumption have shown to provide health benefits [8,9]. The continuing consumer interest in these and
other LAB-containing functional foods may be a reason for the special fondness for these bacteria that go beyond any personal affection [10]. This interest has a long history as the first association of LAB with traditional fermentations, naturalness and long life, has been described over 100 years ago for what is now known as Lactobacillus delbrueckii subsp. bulgaricus [11]. Moreover, it is widely known that LAB are highly versatile and include phylogenetically related bacterial taxa that are essentially non-pathogenic.

The early days of the genome sequencing era witnessed a strong focus on pathogens, starting with Haemophilus influenza in 1995 [12]. In hindsight, this medical focus explains why the first LAB genomes were only deciphered some years later, in the early 2000s with Lactococcus lactis subsp. lactis [13] and Lactobacillus plantarum [14]. Ever since, the number of sequenced LAB genomes has grown exponentially and currently genomic data from over 100 LAB species and strains are available in various public databases. These offer a wealth of information, to further understand LAB with respect to their gene content, their properties, and their ecological role in human health as well as in food fermentations [15].

The present review aims at discussing and describing the latest functional genomic advances in LAB species that are associated with food and health (Figure 1). As prototype functional genomics studies rely on a complete genome sequence, we focus here on the LAB that comply with this criterion and these include rod-shaped LAB (Lactobacillus) and a dozen coccoid LAB, including Lactococcus, Streptococcus, Enterococcus, Pediococcus spp. and Oenococcus spp. (Table 1). Remarkably, this morphological distinction is reflected in a dichotomy in the genome-based phylogenetic tree (Figure 2). We will specifically focus on food-related fermentations where much basic progress has been on the global expression control using transcriptome and proteome approaches that are facilitated by the fact that these systems are easily accessible or can be mimicked in the laboratory. In contrast, the human associated LAB are more difficult to access and most studies that will be discussed relate to LAB with clear health benefits to the human host. Finally, we will address the evolutionary impact of the genomic adaptations (Figure 1) and describe some of the latest genomics approaches applied to LAB for improved food fermentations or health benefits.

Figure 1 Overview of LAB associations with plants and animals, human and foods. The estimated time frames of the evolutionary events relating to the emergence of human (top) and domestication (bottom) are indicated - please note their different dimensions. For a further explanation, see text.
Table 1 Genomic features of a selected number of lactic acid bacteria related to human lifestyle and health.

| Bacterial species | Example of Sequenced Strain | Isolation Source | Genome Size (Mbp) | Number of Plasmids | %GC | Number of Proteins | References |
|-------------------|-----------------------------|------------------|-------------------|--------------------|-----|-------------------|------------|
| **Lactobacilli**  |                             |                  |                   |                    |     |                   |            |
| Lactobacillus acidophilus | NCFM | GI tract (Feces) | 1.99              | 0                  | 34.7 | 1,832            | [112]     |
| Lactobacillus amylovorus | GRL1112 | Pig GI Tract (Feces)* | 2.13              | 2                  | 38.1 | 2,121            | [195]     |
| Lactobacillus brevis | ATCC 367 | Unknown          | 2.34              | 2                  | 46.0 | 2,218            | [19]      |
| Lactobacillus buchneri | ATCC 11577 | Oral Cavity | 2.86 | n.d. | 39.5 | 3,002 | DS | |
| Lactobacillus casei | BL23 | Food (Cheese) | 3.08 | 0      | 46.3 | 2,997 | [196] |
| Lactobacillus crispatus | EM-LC1 | GI Tract (Feces) | 1.83 | n.d. | 37.0 | 1,751 | DS | |
| Lactobacillus delbrueckii subsp. bulgaricus | ATCC 11842 | Food (Dairy product) | 1.87 | 0      | 49.7 | 1,529 | [19] |
| Lactobacillus fermentum | IFO 3956 | Food (Plant) | 2.1 | 0      | 51.5 | 1,843 | [197] |
| Lactobacillus gasseri | ATCC 33523 | Human origin | 1.89 | 0      | 35.3 | 1,755 | [19] |
| Lactobacillus helveticus | DPC 4571 | Food (Cheese) | 2.08 | 0      | 37.1 | 1,610 | [113] |
| Lactobacillus iners | AB-1 | Vaginal Cavity | 1.29 | 0      | 32.7 | 1,209 | [154] |
| Lactobacillus jensenii | 269-3 | Vaginal Cavity | 1.69 | n.d. | 34.4 | 1,575 | DS | |
| Lactobacillus johnsonii | NCC 533 | GI Tract (Intestine) | 1.99 | 0      | 34.6 | 1,821 | [198] |
| Lactobacillus kefiranofaciens subsp. kefiranofaciens | ZW3 | Food (Kefir) | 2.35 | 2      | 37.4 | 2,162 | [199] |
| Lactobacillus paracasei subsp. casei | N1115 | Food (Dairy products) | 3.06 | 4      | 46.5 | 2,985 | [200] |
| Lactobacillus plantarum | WCF51 | Oral Cavity (Saliva) | 3.35 | 3      | 44.4 | 3,063 | [201] |
| Lactobacillus reuteri | DSM 20016 | GI Tract (Intestine) | 2.0 | 0      | 38.9 | 1,900 | DS | |
| Lactobacillus rhamnosus | GG | GI Tract (Intestine) | 3.01 | 0      | 46.7 | 2,913 | [69] |
| Lactobacillus ruminis | ATCC 25544 | GI Tract (Intestine) | 2.07 | 0      | 43.7 | 2,153 | [111] |
| Lactobacillus salivarius subsp. salivarius | UCC118 | GI Tract (Intestine) | 2.13 | 3      | 33.0 | 2,013 | [106] |
| Lactobacillus sakei subsp. sakei | 23K | Food (Meat) | 1.88 | 0      | 41.3 | 1,871 | [202] |
| **Lactococci**    |                             |                  |                   |                    |     |                   |            |
| Lactococcus lactis subsp. lactis | IL1403 | Food (Cheese) | 2.37 | 0      | 35.3 | 2,277 | [13] |
| Lactococcus lactis subsp. cremoris | MG1363 | Food (Dairy Products) | 2.53 | 0      | 35.7 | 2,434 | [203] |
| **Streptococci**  |                             |                  |                   |                    |     |                   |            |
| Streptococcus salivarius | CCHSS3 | Oral Cavity | 2.22 | 0      | 39.9 | 2,027 | DS | |
| Streptococcus thermophilus | CNRZ1066 | Food (Yoghurt) | 1.8 | 0      | 39.1 | 1,914 | [181] |
| **Enterococci**   |                             |                  |                   |                    |     |                   |            |
| Enterococcus faecalis | V583 | Clinical Sample (Blood) | 3.36 | 3      | 37.4 | 3,264 | [204] |
| Enterococcus faecium | DO | Clinical Sample | 3.05 | 3      | 37.9 | 3,114 | [205] |
| **Oenococci**     |                             |                  |                   |                    |     |                   |            |
| Oenococcus oeni | PSU-1 | Food (Plant) | 1.78 | 0      | 37.9 | 1,691 | [19] |
| **Pediococcus**   |                             |                  |                   |                    |     |                   |            |
| Pediococcus pentosaceus | ATCC 25745 | Food (Plant) | 1.83 | 0      | 37.4 | 1,752 | [19] |
| Pediococcus clausnng | ATCC BAA-344 | Food (Beer) | 1.98 | 8      | 37.0 | 1,881 | [182] |
| **Leuconostoc**   |                             |                  |                   |                    |     |                   |            |
| Leuconostoc mesenteroides | ATCC 8293 | Food (Olives) | 2.08 | 1      | 37.7 | 2,003 | [19] |
| Leuconostoc citreum | KM20 | Food (Khimch) | 1.9 | 4      | 38.9 | 1,820 | [136] |
| Leuconostoc gelidum | JB7 | Food (Khimch) | 1.89 | 0      | 36.7 | 1,796 | [206] |
| Leuconostoc camosus | JB16 | Food (Khimch) | 1.77 | 4      | 37.1 | 1,691 | [207] |
| Leuconostoc kimchi | IMSNU 11154 | Food (Khimch) | 2.1 | 5      | 37.9 | 2,129 | [208] |
| Leuconostoc gasmoritatum | LMG 188117 | Food (Spoilage) | 1.95 | 0      | 36.7 | 1,912 | [209] |

Legend: DS, Direct Submission to sequence databases; n.d, not defined; *, No human isolates have been sequenced yet. Due to some discrepancies between the original references and the sequence databases, the data shown in the table were exclusively retrieved from NCBI databases as on 4th of April 2014.
widely used as a host with the NICE system [17,18], only 4 other complete genomes of this taxon have been reported. These genomes include their plasmid complement, which is of crucial importance as it harbors many important dairy functions [16]. The first strain was SK11, a well-studied good flavor-producing strain used as model in earlier genetic studies [19,20]. More recent examples include strain A76, isolated from a cheese production system and strain UC 509.9, an Irish starter with the smallest genome [21]. Moreover, the complete genome of *Lactococcus lactis* subsp. *cremoris* KW2, derived from a corn-fermentation, has been elucidated [22]. This and another plant isolate,
**Lactococcus lactis** subsp. *lactis* KF147, isolated from mung bean sprouts, with one of the largest lactococcal genomes [23], serve as models for domestication studies (Figure 1) and will be discussed below.

In recent years genomic interest has developed into the so called Non-Starter LAB (NSLAB) that are naturally present in dairy fermentations and in some cases have been developed into adjunct starters that contribute to flavor development or quality improvement of fermented foods [24,25]. An example is the recent genomic characterization of *Lactobacillus helveticus* strain CNRZ 32, used as an adjunct starter to reduce bitterness and found to encode 4 different cell-envelope proteinases, in contrast to other Lactobacilli that have one or none [26].

A variety of functional genomics approaches have been reported in the last decade that relate to LAB found in food fermentations. Most have focused on the dairy LAB and here we will discuss the salient features of the common elements that relate to the control of gene expression and serve as models for other LABs. Moreover, functional studies have targeted a variety of foods where attention has been focused on starter LAB, NSLAB and spoilage LAB. Finally, in these studies a series of discoveries have been described that affect the lifestyle of LAB and these are briefly summarized.

**Growth & global regulation**

LAB are known to be rather fastidious bacteria that compete based on rapid growth and lactic acid production in a selected number of habitats (see Figure 2) Genomic-based metabolic reconstructions and modeling have confirmed the dependence on external sources of sugar and protein that are found in complex media such as milk, meat and some plant materials. So much attention has been focused on starter LAB, NSLAB and spoilage LAB. Finally, in these studies a series of discoveries have been described that affect the lifestyle of LAB and these are briefly summarized.

**Growth & global regulation**

LAB are known to be rather fastidious bacteria that compete based on rapid growth and lactic acid production in a selected number of habitats (see Figure 2) Genomic-based metabolic reconstructions and modeling have confirmed the dependence on external sources of sugar and protein that are found in complex media such as milk, meat and some plant materials. So much attention has been focused on starter LAB, NSLAB and spoilage LAB. Finally, in these studies a series of discoveries have been described that affect the lifestyle of LAB and these are briefly summarized.

Several nitrogen control systems are present in LAB and the most studied include GlnR and CodY. While GlnR is present in all LAB genomes, CodY is only present in *Lactococcus*, *Streptococcus* and *Enterococcus* spp. [40]. A comparative genomic study of GlnR regulon, revealed its target site to be present in all LAB genomes and, supported by published transcriptome analyses, predicted GlnR to be involved in controlling the import of nitrogen-containing compounds and the synthesis of intracellular ammonia under conditions of high nitrogen availability [40]. In *Lactococcus lactis* MG1363 GlnR was found to be rather specific but CodY appeared to be a much more global control system [41]. This appeared to be the cases for all other coccoid LAB where it is present. Similar to the identification of the CcpA regulon, a comparative transcriptome approach using an isogenic *codY* mutant was followed to identify the CodY regulon in *Lactococcus lactis* MG1363 [42]. Over 30 genes mainly involved in amino acid metabolism were identified to be under control of CodY in strain MG1363 and in later study in strain IL1403 some more were predicted based on the CodY target (CodY box) in the genome of this strain [43]. The CodY box is present in the promoter of the *codY* gene, explaining that *codY* regulates its own synthesis and does so in response to branched chain amino acids [42]. Importantly, CodY controls the proteolytic system of *Lactococcus lactis* and notably the cell-wall proteinase (PrtP), the key enzyme in milk degradation that prior to the genomic era was shown to be controlled at the transcriptional level by milk-derived
peptides [44]. During growth of strain MG1363 in milk, CodY also acts as a regulator of a major network and detailed transcriptional studies identified a second CodY box in the intergenic regions of 3 operons but the function of this element remains enigmatic in absence of further experimental work [39]. An integrated transcriptomic and proteomic analysis of the adaptation of strain IL1403 to isoleucine starvation showed that CodY was specifically dedicated to the control of the supply of this branched chain amino acid [45]. In *Streptococcus thermophilus* CodY was found to be also involved in the control of the proteolytic system but the study failed to identify a conserved CodY box, indicative of a species-specific *cis*-acting control elements [46]. Remarkably, CodY in pathogenic *Streptococci* was shown to provide a link between amino acid and carbon metabolism as well as virulence factors such as nasopharynx colonization and the synthesis of exoproteins [47,48]. It would be of interest to determine whether CodY of *Streptococcus salivarius* has a similar role in the colonization of this and related species in the oral or other human related cavities (see below). The absence of a *codY* gene in the genomes of *Oenococcus* and *Pediococcus* suggests that these bacteria have a life style where they do not need such an intricate protein control [40]. Alternatively, these bacterial species may employ different regulatory mechanisms, possibly involving unrecognized regulators.

Apart from the above-mentioned CcpA, GlnR and CodY, many other specific and global regulators have been described and functionally studied. In many cases new links may be observed as the control systems all seem to be interlinked. With the development of high throughput transcriptome and RNAseq approaches, new avenues to identify and map these are emerging. The recent analysis of the global regulatory networks, identified during growth of *Lactococcus lactis* subsp. *cremoris* MG1363 in milk, is such an example [39]. This is expected to be followed by other studies that will provide insights into the global control, the *cis*-acting elements, and their nodes. The challenge is to relate these transcriptional networks to the metabolic networks that are now well-developed to increase the predictability of LAB in the model systems, in food products and in association with human [49].

**Expression in foods**

To improve the understanding of growth and function of LAB in fermented foods, numerous global transcriptional, proteomic and recently also metabolomic studies have been performed. Model and starter strains of *Lactococcus lactis* have been the first to be studied. A lactose-proficient derivative of the model strain MG1363 was used in an artificial cheese system using an expression technology approach [50]. While a series of genes involved in amino acid transport and metabolism were identified, the approach suffered from the fact that the strain used was plasmid-free and did not contain the PrtP-encoded system and hence was not proteolytic. This caveat also applies to the elegant study of strain MG1363 in milk elucidating the global networks [39]. However, several other studies have addressed the expression in cheese of starter lactococci that are capable of rapid growth in milk. Using cheeses made from milk concentrated by ultrafiltration (UF-Cheese) and the starter *Lactococcus lactis* subsp. *lactis* biovar diacetylactis LD6, a detailed study was made of the in situ global gene expression [51]. Genes of the proteolytic system were increased due to down-regulation of CodY repression, while acid and oxidative stress-related genes were increased. Moreover, carbon limitation was apparent and involved release of CcpA-mediated control. In similar UF-Cheeses made with strain LD6, recently the metabolites were determined using an unsupervised mass-spectrometry approach, illustrating the power of other non-targeted functional approaches [52]. In an unrelated study, four *Lactococcus lactis* subsp. *cremoris* starter strains (SK11, and proteolytic variants of HP, Wg2 and E8) were used in parallel cheese vats and analyzed for their transcriptomic response [53]. This resulted in the definition of a core transcriptome with almost 200 genes, mainly encoding for house-keeping functions but also those involved in cysteine metabolism. Several of these were found to be under control of the CodY regulator, reiterating the common theme discussed above. As indicated below, correlations between CcpA, CodY and the stringent response exist and it is expected that these regulatory circuits are all operating during these complex fermentations in cheese making. As often mixtures of LAB strains are used as cheese starter cultures, various approaches have been developed to differentiate between the components of the starter. Various metagenomics and quantitative PCR approaches have been tested and shown to have potential for strain differentiation or expression [54,55]. Sequence analysis of 16S rRNA transcripts has recently been used to identify the microbial composition and activity of Cheddar cheese batches, identifying both LAB and NSLAB. These and similar investigations can be coupled to RNAseq studies to analyze the expression in real time of the different components.

Only few other global gene expression studies have been performed in food products other than those derived from fermented milk. The global transcriptome of *Lactococcus garviae*, a fish and opportunistic human pathogen was analyzed and revealed a heme-dependent and cold-induced respiration system [56]. This had already been described some years ago in another strain of *Lactococcus garviae* [57]. Such a respiration system was also identified in a transcriptomic approach of
**Leuconostoc gasicomitatum**, an emerging food spoilage organism, when grown in meat [58]. The endogeneous heme present in meat allowed respiration and this increased growth rate and yield. Interestingly, this had no impact on the transcriptional response of *Leuconostoc gasicomitatum*, similar to what has been observed in *Lactococcus lactis* [59]. However, it has been described that the meat-grown *Leuconostoc gasicomitatum* respiration activity was increased 1000-fold and was paralleled by the production of different metabolites, suggesting that its control is at the metabolic rather than the transcriptional level [58].

**Novel insight and functions**

While providing a molecular understanding of the adaptation of LAB to the food environment, the genomics studies discussed here also present insight in novel functions. An example is the identification of a novel stress regulon under the control of the protein Ldb0677 in *Lactobacillus delbrueckii* subsp. *bulgaricus* by using a proteomic approach and its characterization by molecular techniques [60]. Moreover, studies in other model systems may shed new light on the findings in LAB. One such new insight derives from findings in *Bacillus subtilis*, which reportedly shares a common ancestor with the LAB [19]. It has recently been shown that CcpA forms complexes with CodY in *Bacillus subtilis* and there is no reason to assume this would not be possible in LAB [61]. This strongly suggests that the carbon and nitrogen control in LAB are intimately connected. Similarly, structural analysis of the *Bacillus subtilis* CodY indicated that GTP is a ligand for this conserved regulator and hence CodY reacts to (p)ppGpp levels formed in the stringent response [62-64]. The stringent response of the (p)ppGpp alarmone may well be one of the general triggers that operate in LAB during cheese fermentation.

The discovery of aerobic respiration in LAB and its genetic elucidation has been well documented together with its biotechnological application [59]. This heme-dependent property has now been found to be operating in many LAB, including several *Lactobacillus*, *Leuconostoc* and *Enterococcus* spp. [57,65,66]. Strictly speaking respiration is the coupling of a membrane potential to the reduction of oxygen and this only has been shown to operate in *Lactococcus lactis* subsp. *cremonis* MG1363 when grown on heme [67]. It is of interest to note that this respiration is so widely spread and appears to occur in food fermentations when there is a supply of heme-containing media. Remarkably, also the genome of *Oenococcus oeni* contains the genes for aerobic respiration but its functionality has not yet been tested [67].

By an elegant combination of genomics and expression studies, it has been shown that the *Lactococcus lactis* model strains IL1403 contains the genes for pili production that can be expressed and are involved in biofilm production [68]. Prior to this discovery such proteinaceous pili had only been described in the GI tract isolate *Lactobacillus rhamnosus* GG where they bind human mucus as well as have a set of other functions, e.g. immunogenicity [69,70]. The presence of these functional pili genes in strain IL1403 prompted comparative genomics studies that revealed their presence in various *Lactococcus lactis* strains, including the other model strain MG1363, the plant isolate KF147 (see above), and various other plant and human isolates [68]. The presence of pili production genes in dairy and plant strains suggests that this property is multifunctional and provides competitive advantage in various environments. Interestingly, by using a combination of proteomics and genomics, a functional pili cluster that enables mucus binding was also detected in another plant isolate, strain TIL448 [71]. Here, the genes for the pili production are located on a plasmid, suggesting horizontal gene transfer and proving a possible mechanism for the apparently wide spread of this novel function in dairy and plant lactococci.

**Functional genomics of LAB in human health**

The colonization of LAB in and on the human body has been well established and 16S rRNA-based phylogenetic studies have identified LAB at different body sites, such as the skin, oral cavity, GI tract, and vaginal cavity [72-77]. Further comprehensive phylogenetic and metagenomic characterizations of the human-associated microbiota using massive parallel sequencing, have extended this notion and identified the presence, level and genetic content of the various LAB in the microbial communities in the human body [4,5]. Based on these data it can be concluded that the number of total microbes varies considerably in the various body sites, as does the fraction of LAB (Figure 3).

The recent genome-based molecular inventories have shown that the fraction of LAB in the GI-tract is low and barely reaches over 1 % in only few persons (Figure 3). It is assumed that many of these LAB are passengers rather than endogenous inhabitants. Still, a detailed phenotypic and genomic characterization of strains from each LAB species is needed to clarify their role within the GI tract, since some LAB have a high intraspecies diversity and include both endoenous and passenger strains. This has been confirmed in human feeding studies with marked *Lactococcus lactis*, showing unexpected survival of viable cells [78]. Moreover, a recent high fat feeding trial where fecal DNA was analysed using massive parallel sequencing, revealed the transit of *Lactococcus lactis*, *Streptococcus thermophilus* and *Pediococcus acidilacti*, which are components of dairy and meat starters [79]. However, based on genomic or
sequence characteristics various LAB strains have found to be endogenous in human [56,73,75,80]. By far the highest fractions of LAB are found in the oral and vaginal cavities since the environment of these relatively open systems is more accessible than that of the human GI tract (Figure 3).

While our mouth as the port d’entrée of the GI tract is receiving a rather variable microbial load of mainly passengers, the vaginal cavity has a rather stable microbiota. This explains why the endogeneous vaginal LAB were found to be specifically associated with health [81]. This contrasts with the GI tract where most specific associations with health have been described for other members of the complex human-associated communities than LAB [82]. An exception is a recent metagenome study, where *Lactobacillus gasseri* was associated with the incidence of type 2 diabetes in a Swedish cohort [83]. However, this was not reproduced in another large type 2 diabetes cohort and the observed genes may have derived from passenger LAB [84].

As many of the genomes of human-derived LAB have been determined (Figure 2), we summarize the recent functional genomics studies of these strains below.

### Oral Cavity

The mouth constitutes the first cavity from which food is introduced into the digestive tract. As an ecological habitat, it hosts hundreds of different bacterial species, including LAB, that are colonizing the teeth, the gum, the saliva and various locations on the tongue [4]. Teeth, as hard tissues, form an excellent surface for biofilm formation [85]. A dozen Lactobacilli are found to be the most prevalent LAB detected in the oral cavity (Figure 2) [86,87]. Metaproteomic analysis also confirmed the presence of Lactobacilli in the human saliva [88]. Some LAB have been used to restore healthy oral microbiota and the well-known probiotic *Lactobacillus rhamnosus* GG was shown to reduce the population of *Streptococcus mutans*, the common cause of caries [89]. Genomic and phenotypic characterization of oral isolates of *Lactobacillus rhamnosus* indicated that these were closely related to cheese isolates, suggesting that they may originate from food products [80]. However, genomic characterization of *Lactobacillus rhamnosus* strains isolated from dental pulp showed that these were unique and contained an additional set of approximately 250 unique genes [90]. These genes included those

---

**Figure 3** Overview of the level of LAB in the different body sites. The estimated LAB fraction is based on several complete and comprehensive phylogenetic and metagenomic datasets and the total number of bacteria per gram of homogenized tissue or fluid or square centimeter of skin [4,94,95,220,221].
coding for the biosynthesis of exopolysaccharides that could be involved in biofilm formation, while others encoded transcriptional regulators and ferric iron ABC transporters. In the oral isolates of both studies, the spaCBA-srtC1 plius gene cluster was lacking, suggesting that such trait is not essential for persisting in the oral cavity [80,90].

The gastro-intestinal tract
Isolated or detected throughout the whole digestive tract, LAB only represent a minor proportion of gastro-intestinal microbial communities [73,91]. Typically, representatives of the Lactobacillus/Enterococcus group constitute 0.01-1.8% of the overall fecal microbiota, as shown by qPCR techniques [92]. Their abundance in the GI tract significantly ranges from less than 10^5 CFU/ml (small intestine) to 10^6 CFU/g (faeces) (Figure 3) [73,74,93-95]. The human small intestine was shown to harbour a diverse population of Streptococci [96]. However, sequence analysis of the rRNA gene does not allow determining whether these detected LAB strains are endogenous or transient. Up to date, more than 20 LAB species have been detected in the digestive tract (Figure 2). Some of these are consumed as probiotics, such as Lactobacillus plantarum, Lactobacillus casei or Lactobacillus rhamnosus [8,10,97]. Others are present in the mouth where they may be derived from food or be endogenous (see above). This suggests that some of the LAB isolated from the GI tract may in fact originate from food or the oral cavity [96,98,99]. Detailed comparative and functional genomic characterization of human LAB isolates may provide answers whether they are endogenous or transient, as well as generate a better understanding of their ecological fitness, their adaptation, and their role in their dedicated niche. The first of these studies related to Lactobacillus johnsonii and Lactobacillus gasseri, which were genetically characterized ten years ago (Table 1). Genomic data complemented with experimental work provide evidence for the ecological adaptation and fitness of Lactobacillus gasseri to the GI tract, as recently reviewed [100]. Transcriptomic analysis of Lactobacillus johnsonii NCC533 identified a number of genes that could relate to its persistence within the intestinal tract [101]. The isolation and sequencing of intestinal LAB along with LAB from other sources has allowed us to compare strains and to determine the diversity of each species from an ecological but also evolutionary perspective. In a recent comparative genomic study, the examination of 100 Lactobacillus rhamnosus isolates showed possible correlations between ecological fitness, phenotypic traits and genomic modifications [80]. The intraspecies diversity in Lactobacillus rhamnosus was mostly concentrated in 17 lifestyle isolates. Compared to Lactobacillus rhamnosus food isolates, a subset of GI tract isolates harbored more prevalently genes associated with specific carbohydrate pathways (fucose metabolic genes), host adhesion (mucus-binding SpaCBA pilus gene cluster), defence and immunity system (CRISPR system) and biofilm formation (exopolysaccharide cluster). These are likely to provide an improved capacity to colonize and persist in the GI tract [80]. Intestinal Lactobacillus rhamnosus isolates were shown to be resistant to bile, whereas isolates from dairy niches for example were generally less bile-resistant [80]. Two other closely related species Lactobacillus casei and Lactobacillus paracasei shared some lifestyle islands with Lactobacillus rhamnosus that were syntenous [102,103]. Using hybridization arrays and multilocus sequence typing, the genomic diversity of Lactobacillus salivarius was studied [104]. In line with findings in other LAB, the intraspecies diversity was found to be concentrated on 18 chromosomal regions that included gene clusters encoding for the production of exopolysaccharides [104]. An important fitness factor with applied potential is the capacity to produce a broad host-range bacteriocin that allowed Lactobacillus salivarius to outcompete Listeria monocytogenes [105]. In addition to chromosomal variations, the presence of plasmids and other mobile elements are playing an important role. One remarkable example contributing to intraspecies diversity is the presence of megaplasmids in some Lactobacillus salivarius strains. Lactobacillus salivarius subsp. salivarius UCC118 harbors the megaplasmid pMP118 (242 kb in size) [106]. Further analysis of two other subspecies identified other megaplasmids with a different size, suggesting a possible role in ecological adaptation [106]. Some species such as Lactobacillus reuteri are specialized to one particular host. Lactobacillus reuteri is also commonly in different human body sites, i.e. breast milk, GI tract, vagina but it is also found in other vertebrates [107,108]. Work on the Lactobacillus reuteri species revealed that strains have distinctly evolved between different hosts. Gut isolates from different mammals, i.e. rodents and humans have distinct genetic signatures. This may be explained by the fact that the anatomical differences between human and rodent gut resulted in different colonization strategies [109]. The host specialization observed among Lactobacillus reuteri strains results from similar genetic mechanisms as in other symbiotic bacteria [109]. The role played by transposases in the genome dynamics between rodent and human isolates differs. The genomes of Lactobacillus reuteri human gut isolates tend to be smaller with higher number of pseudogenes [109], as previously reported in other host-dependent bacteria [110]. In contrast with the Lactobacillus reuteri strains, where it was shown that strain differ according to the host, comparative genomic analysis showed that the human gut strain Lactobacillus ruminis
ATCC 25644 is highly similar to the bovine isolate *Lactobacillus ruminis* ATCC 27782 [111]. They, however, significantly differ from the closely related *Lactobacillus salivarius* (Figure 3). *Lactobacillus acidophilus* and *Lactobacillus helveticus* are closely related (Figure 2). However, *Lactobacillus helveticus* is typically more specialized to the dairy environment compared to the gut-adapted *Lactobacillus acidophilus*, which has conserved more biological functions. In the *Lactobacillus helveticus* genome, adhesion factors, such as mucus-binding proteins, are absent along with a narrower gene repertoire encoding for PTS transporters [112,113].

Genome sequences of LAB provided a basis to identify the secretome and interactome of LAB found in the human GI tract. Within the *Lactobacillus casei* group, the respective LPXTG protein-encoding gene repertoires of *Lactobacillus rhamnosus*, *Lactobacillus casei* and *Lactobacillus paracasei* shared several similarities [102]. Among others, pilus gene clusters were identified. However, only in *Lactobacillus rhamnosus*, the functionality and expression of one of the gene cluster encoding mucus-binding pili (spaCBA-srtC1) has been so far demonstrated [69,114]. This single and outstanding trait contributes to the highly efficient adhesion of *Lactobacillus rhamnosus* GG to the intestinal mucosa [69]. Within the *Lactobacillus rhamnosus* species, pilus-associated genes were significantly more present in intestinal isolates (56 %) compared to dairy isolates (13 %) [80].

Genome-wide analysis of *Lactobacillus salivarius* UCC118 identified 108 predicted secreted proteins, including 10 sortase-anchored proteins. Gene deletion of sortase and one sortase-anchored protein significantly reduced the epithelium-binding ability of the strain UCC118 [115]. A recent review discussed the central role of sortases and LPXTG proteins for LAB, especially for the ones found in the GI tract [116]. Interestingly, some *Lactobacillus ruminis* strains, *i.e.* ATCC 27782, also possess a set of genes encoding for a complete and functional flagellar apparatus, *i.e.* 45 flagellar genes, providing motility [117]. The discovery of motile commensal LAB suggests unique and uncovered impact on the gut ecology in terms of host signaling and colonization. In the intestinal *Lactobacillus gasseri* ATCC 33323, among the 271 predicted cell surface proteins, at least 14 mucus-binding proteins were identified [118], suggesting a potential role in adherence with the intestinal mucosa. In *Lactobacillus acidophilus* L-92, the attachment to epithelial cell lines altered the expression of 78 genes, *i.e.* membrane proteins, transporters and regulators [119]. Comparative proteomic analysis led to the identification of 18 proteins with potential adhesive properties, including surface-layer protein A. Further work showed that the latter protein has a central role in the adherence of *Lactobacillus acidophilus* L-92 to epithelium [120]. Moreover, one of the well-characterized surface-layer proteins, SdpA of *Lactobacillus acidophilus* NCFM, was found to bind to the DC-SIGN receptor of dendritic cells, indicative of a role in intestinal signaling [121,122].

A number of similarities in terms of response to the GI environment have been observed among gut-isolated LAB species and relate among others to metabolic re-routing, cell wall modifications or activation of resistance/stress mechanisms. The mechanisms by which these genes are induced when LAB are in the human gut are not fully comprehended. Specific attention has been given to the exposure to bile salts and acids as during the transit (and eventual colonization) in the GI tract, LAB are exposed to these environmental stimuli. Recent proteomic and transcriptomic analysis of the intestinal *Lactobacillus rhamnosus* strain GG under bile stress revealed the activation of numerous genes related to cell wall functions and possibly operate as a stimulus for adherence in the intestinal tract [123]. *Lactobacillus rhamnosus* strain GG also generated a specific response towards acid environments, as examined by proteomic analysis [124]. Similarly, in *Lactobacillus casei* BL23, 52 proteins showed an altered expression under bile stress, and these were predicted to be involved in general stress response, cell wall functions and also carbohydrate metabolism [125]. Remarkably, in *Lactobacillus acidophilus*, glycogen metabolism was found to be associated with bile resistance [126]. Apart from these laboratory studies also a series of model animal and human studies have been reported. An in vivo expression technology (IVET) study in *Lactobacillus plantarum* WCFS1 identified a set of 72 genes that were induced when transiting the GI tract of mice [127]. These mainly include genes associated with carbohydrate metabolism, biosynthetic pathways and transport and also four genes potentially relating to host interactions, *i.e.* cell wall anchor proteins [127]. Reciprocally, *Lactobacillus plantarum* WCFS1 cells triggered the expression of over 400 genes in the mucosa of the human small intestine [128,129]. A mouse study further addressed the transcriptional responses of *Lactobacillus plantarum* to different dietary regimes [130]. Finally, the transcriptional responses to *Lactobacillus plantarum* WCFS1 in mice and human were described in a detailed comparative study that revealed high level similarities between those systems [131]. The transcriptomic profile of *Lactobacillus plantarum* WCFS1 was also found to be modified upon exposure to p-coumaric acid, a component present in vegetables or fruits, possibly signaling *Lactobacillus plantarum* to its entry to the digestive tract [132]. Similarly, the transcriptional response of *Lactobacillus plantarum* to bile was also investigated, revealing a set of genes whose expression is bile-inducible [133]. Within the
Lactobacillus plantarum species, strains have different bile sensitivity, i.e. showing either resistance (strain 299V) or sensitivity (strain LC56) [134]. Comparative proteomic analysis of three different strains led to the identification of 13 proteins related to bile resistance mechanisms [134]. In addition, alteration of genes associated with cell surface proteins and metabolism suggests that Lactobacillus plantarum underwent adaptation when exposed to the murine tract [135]. In intestinal isolates of Lactobacillus reuteri, a total of 28 genes were shown to be induced under bile salt exposure and proteomic analysis indicated that the encoded proteins were associated with metabolic pathways, stress-induced response and also pH homeostasis, which possibly relate to resistance mechanisms of Lactobacillus reuteri to bile salt stress [136]. A similar mechanistic response was observed when exposed to acids [137]. Mice studies showed that the transcriptome of Lactobacillus johnsonii NCC533 is changing throughout the GI tract, suggesting specific responses to each of the GI sites [138]. Using a mouse model, it was found that 174 Lactobacillus johnsonii NCC533 genes were expressed in vivo, including EPS-associated glycosyltransferase genes and PTS transporters [101].

In conclusion, LAB when present in the GI-tract express a number of common characteristics that relates their adaptation. These could be summarized as follows: i. a large repertoire of genes encoding transporters (ABC, PTS or permeases) to optimally utilize nutrients available in the gut niche, ii. the presence of genes associated with acid and bile resistance, iii. a wide range of genes promoting interactions and signaling with the host, such as pili that contain mucus-binding proteins.

The vaginal cavity
LAB members constitute a dominant proportion (~80%) of bacteria inhabiting the vaginal cavity of healthy women [139] and are consistently detected in healthy vaginal microbiota from patients of different ethnic groups and/or living in different geographical locations [139-143]. Four main bacterial species were typically identified: Lactobacillus crispatus, Lactobacillus iners, Lactobacillus jensenii and Lactobacillus gasseri along with, at lesser extent, some other lactobacilli, such as Lactobacillus acidophilus, Lactobacillus ruminis, Lactobacillus rhamnosus or Lactobacillus vaginalis [139,144-146]. The high abundance of LAB is strongly associated with healthy vagina, whereas a low abundance of LAB, i.e. alteration of the vaginal microbiota, was more prevalent in women with a medical condition, i.e. bacterial vaginosis (BV) [140,145,147]. The beneficial roles of LAB in preserving a healthy vagina include the maintenance of acidic vaginal pH [148], the prevention of infections by producing bacteriocins, hydrogen peroxide and acids, but also by signaling to the host [148-150]. The understanding of the vaginal microbiota composition not only contributes to the comprehension of the ecology of this habitat in health and disease but also offers avenues towards the development of better diagnostic and therapeutic solutions [147,151,152].

Four LAB species are predominantly detected in human vagina (Lactobacillus crispatus, Lactobacillus gasseri, Lactobacillus iners and Lactobacillus jensenii) but co-dominance between LAB species is seldom [142]. This indicates that each vaginal species may harbor genes that relate to (unique) adaptation signatures and allow the non-symbiotic persistence and colonization regardless of the presence of other LAB members [153]. Interestingly, these LAB genomes also showed to be significantly smaller and contained a lower GC content than other LAB genomes, suggesting a loss of non-essential genes towards a vaginal adaptation [153].

One of the most studied vaginal LAB is the Lactobacillus iners. Remarkably, strains from the Lactobacillus iners species have a relatively small genome compared to the LAB, i.e. ~1.3 Mb for Lactobacillus iners AB-1 genome [154] and its intraspecies diversity is peculiarly low [143]. In line with its genome size, Lactobacillus iners is not able to biosynthesize many vitamins, cofactors and amino acids, while compensating these metabolic limitations by the presence of numerous genes encoding transporters [154]. When compared to Lactobacillus crispatus, Lactobacillus gasseri and Lactobacillus jensenii, Lactobacillus iners carries a variety of unique genes encoding ABC transporters [153]. The poor metabolic and biosynthetic capabilities illustrate its strong dependency to the host niche, from where Lactobacillus iners acquires most of its nutrients. This may also explain why this species is rarely detected in other ecological niches that are more demanding in terms of metabolic capabilities [155,156]. Lactobacillus iners is lacking numerous transcriptional regulators or integral membrane proteins [153]. The detailed mechanisms involved the persistence of Lactobacillus iners in the vagina remain unclear. However, a number of genes encoding potential adhesins (a total of 11 LPXTG proteins) were identified in Lactobacillus iners AB-1 [154], along with genes encoding fibronectin-binding type adhesins [157], indicating that interactions occur between the bacterial cells and the vaginal tissues. Such association (lactobacilli-epithelium) promotes exclusion of pathogens [158], as shown with the displacement of biofilms formed by Gardnerella vaginalis [159]. In addition, Lactobacillus iners AB-1 is able to use mucin as a carbon source, which is clearly beneficial for persisting in a mucosal niche (vagina) [154]. Interestingly, the genome of Lactobacillus iners AB-1 contains a gene (LINAB_0216) that encodes a cytolysin [154]. This gene is also found in other Lactobacillus iners isolates and its
product is similar to cholesterol-dependent cytolysins produced in species such as *Streptococcus* or *Gardnerella*, [160]. However, its function in *L. iners* is unclear, i.e. attachment to host tissues, antimicrobial activity or pathogenesis [143,160]. A recent meta-RNA-seq based study showed that during a BV episode *Lactobacillus iners* AB-1 modified the expression of genes encoding the CRISPR-cas system, the cholesterol-dependent cytolysin and the mucin and glycerol transporters [81]. This underlines adaptive mechanisms towards the persistence of *Lactobacillus iners* in changing vaginal microbiota, i.e. change of nutrient use (mucin and glycogen) and protection against bacteriophages [81]. The overexpression of the cholesterol-dependent cytolysin by *Lactobacillus iners* during BV appeared to have a detrimental role towards the host [152]. Based on genomic and transcriptomic data, *Lactobacillus iners* was found to be specifically adapted the vaginal niche under different conditions, i.e. healthy or non-healthy vaginal microbiota. This remarkable adaptation suggests a strong association of *Lactobacillus iners* with the host, possibly contributing to maintaining a healthy microbiota, though its role in BV needs to be further examined.

In contrast with the *Lactobacillus iners* species, strains of all three other vaginal LAB, *Lactobacillus crispatus*, *Lactobacillus gasseri* and *Lactobacillus jensenii* are also found in other ecological niches than the vagina (Figure 2). Intestinal *Lactobacillus gasseri* isolates have genotypic traits beneficial for persistence and colonization in the gut (see above) [118]. Comparative genomic analysis identified a series of species- and/or niche-specific gene sets mostly consisting of different ABC transporters and regulators and in some cases toxin-antitoxin systems or cell envelope proteins [153]. However, no clear vaginal gene sets were defined in *Lactobacillus crispatus*, *Lactobacillus gasseri* and *Lactobacillus jensenii*. Vaginal strains of *Lactobacillus crispatus* have a larger genome than other strains of this species, possibly resulting from an abundance of IS-encoded transposases [153].

Apart from the four dominant LAB species that are recurrently detected in healthy vaginal microbiota, also other *Lactobacillus* spp., can be found and show, in some cases, unique patterns in both phenotypes and genomes (Figure 2). In a recent study, vaginal *Lactobacillus rhamnosus* isolates were compared with the *Lactobacillus rhamnosus* strain GG at both genomic and phenotypic level [80]. Four main genotypic/phenotypic traits were highlighted: the lack of mucus-binding pili, their bile resistance (100% of all isolates), an altered or deficient CRISPR-cas system compared to strain GG and some metabolic capabilities similar to food isolates. It was hypothesized that vaginal LAB may have originated from food environments or the oral cavity and survived through the gastro-intestinal tract (bile resistant, antimicrobial activity), before colonizing the vaginal cavity [80]. The loss of the pilus gene cluster indicates that it is not beneficial for *Lactobacillus rhamnosus* in the vaginal cavity. This is consistent with genomic data on other vaginal LAB, such as *Lactobacillus iners*, *Lactobacillus gasseri* or *Lactobacillus crispatus*, with genomes that does not contain such cluster. Recent work on other LAB, i.e. *Lactobacillus plantarum*, showed that the vaginal adhesion of the bacterial cells is sortase-dependent and therefore relies on LPXTG anchor proteins that likely do not form pili [161]. Similar mechanisms may occur as well in other LAB, such as *Lactobacillus rhamnosus*. No other studies on vaginal *Lactobacillus rhamnosus* genomics have been reported but it seems that only a subset of the *Lactobacillus rhamnosus* species may be able to colonize the vaginal cavity. Most clinical trials using *Lactobacillus rhamnosus* strains showed promising results [162,163]. However, each strain within the species appear to have a distinct ecological fitness and intestinal *Lactobacillus rhamnosus* strain GG with a pheno-genotype different from vaginal isolates, was poorly colonizing the vagina cavity, indicating that it lacks a number of genes promoting its ecological fitness to the vaginal cavity [164].

Other body sites and clinical cases

In general, LAB are considered to be safe and many species are on the list of Qualified Presumed Safety (QPS) of the European Food Safety Authority [165]. This does not apply to *Enterococcus faecalis* and *Enterococcus faecium*, two species of enterococci that have been and are used as starters in various food fermentations as well as marketed as probiotics (Figure 2) [166]. These enterococci emerged as the leading causes of antibiotic-resistant infection of bloodstream, urinary tract and surgical wounds [167]. However, most if not all human are carrying these *Enterococcus* spp. in their GI tract and it has been suggested that enterococci may have been ubiquitous colonizers of the gut since the early Devonian period, i.e. 400 million years ago [168]. Comparative genomic studies have now shed light on how such normal colonizing species may have developed into a major group of pathogens. It appeared that the genomes of hospital adapted enterococcal strains consist of over 25 % of mobile elements, have lost CRISPR-cas systems that limit horizontal gene transfer, and have accumulated multiple antibiotic resistance and virulence traits [168]. It has been proposed that the introduction of antibiotics approximately 75 years ago and their widespread use in both human and veterinary medicine promoted the rapid evolution of the present epidemic hospital-adapted lineage not from human commensals but from a population that included animal strains [168]. There is some apparent disagreement about the moment of divergence between the
remarkable similarities apart from the large 21-gene biome, which is rapidly developing and the dairy cow. Recent culture-independent analysis of the foliar microbiome shows that Enterococcus spp. and pathogenic Streptococci shared more gene families than did the genomes from non-pathogens, such as other LAB [170].

Inspection of the present QPS listing reveals that some LAB have incidental cases where they are implicated in non-nosocomial and other clinical infections. This has been described previously for Lactobacillus rhamnosus and has been recently reviewed [171]. However, the increased intake of Lactobacillus rhamnosus GG did not lead to an increase in bacteremia cases [172]. Hence, EFSA concluded that clinical infections especially of Lactobacillus rhamnosus, should be closely monitored [165]. This also relates to an increasing number of reports that imply LAB in other body sites than the canonical caveats (Figure 3). These include strains of Lactococcus lactis, Leuconostoc lactis, Lactobacillus casei, Lactobacillus paracasei and Pediococcus sp. [165].

The number of reports linking Lactococcus lactis, often the subsp. cremoris, to clinical cases is increasing. Recent studies include the isolation of Lactococcus lactis from human brain or neck abscesses or bovine mastitis [173-175]. It should be remembered that Lactococcus lactis (then appropriately termed Bacterium lactis) was the first bacterium grown as a pure culture by Joseph Lister in 1878. Ironically, Lister compared the fermentation process with an infection process in his attempts to illustrate the cause of infectious disease in humans [176]. It can be expected that further comparative and functional genomic studies of clinical, food and other LAB isolates will be instrumental in understanding the adaptations to the human body as well as assessing the safety of LAB used in the food or pharmacy industry.

**Evolutionary LAB genomics**

**Adaptation and horizontal gene transfer**

It is generally believed that plant material is the archetype source of the dairy LAB, though some inoculation from the dairy cow and its milk is also possible (Figure 1). Recent culture-independent analysis of the foliar microbiome, which is rapidly developing and the dairy cow’s teat showed LAB to be present in both environments [177,178]. Hence, detailed genomic analysis is needed to distinguish between the sources of the dairy LAB. Comparing the genome of the plant isolate Lactococcus lactis subsp cremoris KW2 with the dairy strains showed remarkable similarities apart from the large 21-gene cluster coding for the biosynthesis of wall teichoic acids that is partially absent or truncated in the model strain MG1363 or the dairy starters SK11, UC509.9 or A76. In contrast to the dairy starters, the plant strain KW2 does not contain any plasmids or IS sequences. This substantiates the earlier suggestions that these mobile elements are recent acquisitions by horizontal gene transfer. Moreover, the presence of the gene cluster for the wall teichoic acid production seems to be a plant adaptation as it is also found in Lactococcus lactis subsp. lactis KF147 isolated from mung bean sprouts that has been studied extensively as a non-dairy model for lactococci [23]. This strain KF147 has one of the largest genomes, shows high identity and synteny to the genome of Lactococcus lactis subsp. lactis IL1403 but contains a variety of plant adaptations that have been lost in the dairy starter of this taxon [23,179]. Hence, for Lactococci there is ample evidence that plants are the sources of the dairy strains (Figure 2).

The genome Lactobacillus iners AB-1 is the smallest among the LAB (Table 1) suggesting that important gene loss occurred in that species towards the specialization to one unique ecological habitat, i.e. vaginal cavity. The genome size reduction possibly reflects the dependency of vaginal LAB to their host, as previously reported in other symbiotic bacteria, such as Candidatus Tremblaya princes (genome size of 139 kb) [180]. The limited coding capacities of Lactobacillus iners do not only reflect a remarkable ecological-driven specialization to the vaginal host but also a strong dependency to this habitat. The high number of genes associated with DNA repair, RNA modification and the alteration of a number of metabolic pathways clearly underline how most of these vaginal lactobacilli rely on the host for surviving and persisting. There is a potential mutualistic relationship between the host and the vaginal LAB. The host provides a stable environment, from where vaginal LAB can utilize nutrients (mucin, glycogen) or by-products from other inhabitants. In return, vaginal lactobacilli are warrant of the maintenance of a healthy vaginal microbiota. Although Lactobacillus iners has been reported in rare clinical cases [155], these may constitute evolutionary dead-ends that are usually not associated with any adaptation traits.

As detailed in the first large scale comparative genomic study, most LAB are phylogenetically closely related (Figure 2) but mainly differ by the gain of novel genes or the loss/decay of ancestral genes [19]. In addition, the number of pseudogenes is highly variable among LAB, i.e. S. thermophilus CNRZ1066 (182 pseudogenes) [181] or Pediococcus pentosaceus ATCC 25745 (19 pseudogenes) [19]. The presence of plasmids or megaplasmids in some strains are also of interest, since they may carry additional genes involved in metabolic pathways, production of bacteriocins and bile salt hydrolase. Two
striking examples are: the co-existence of 8 plasmids in *Pediococcus clausenii* ATCC BAA-344 [182] and the presence of a 242-kb megaplasmid pMP118 in *Lactobacillus salivarius* UCC118 [106]. In addition, horizontal gene transfer further contribute to genus and species diversification, as previously reported in *Lactobacillus acidophilus, Lactobacillus casei, Lactobacillus delbrueckii* subsp. *bulgaricus* and *Lactobacillus johnsonii* [103,183-185]. Significant differences observed in LAB genomic features give primary evidence for possible ecological adaptation and specialization: genome size (coding capacities), pseudogenes or plasmids (Table 1). Only a further detailed examination of these genomes may highlight gained, duplicated, decayed or lost gene sets that are encoding biological functions relating to one particular ecological context. The role played by transposases in the genome dynamics between rodent and human isolates differs. The genomes of *Lactobacillus reuteri* human gut isolates tends to be smaller with higher number of pseudogenes [109], as previously reported in other host-dependent bacteria [110].

**Applied LAB genomics**

The use of functional and comparative genomics has greatly enhanced a variety of applications. First, there is the issue of strain identity and protection. Many manufacturers of LAB starters or producers that market LAB as probiotics, have started to characterize their strains by complete genomic analysis. While supporting rapid strain characterization, this is also instrumental in strain mining and speedily selecting specific properties. Moreover, safety, administrative and legal processes can be supported by genome sequences and LAB strains of competitors can be benchmarked. With respect to safety, one should realize that knowledge of a genome sequence does not make a strain safe or not. However, lessons learned from the adaptation of notably *Enterococcus* strains discussed above could be helpful in further predicting safety of LAB.

The rapid implementation of next generation sequencing technologies for comparative genome analysis has allowed for several well-known commercial strains to be made public. It was recently shown that *Lactobacillus casei* strains marketed in Yakult and Actimel products were found to contain only a few dozen single nucleotide polymorphisms (SNPs) and a prophage [186]. This approach also showed that *Lactobacillus rhamnosus* GG isolated from several products was highly stable [114]. A new genomics approach that is only possible by the rapid advances in sequencing technology is capitalizing on genomic resequencing approaches. In a first published example *Lactococcus lactis* NZ9000, containing the *nisRK* two-component system genes that are used in conjunction with the nisin-controlled expression system, was mutated to increase expression of a variety of membrane proteins [17,187]. The genomes of the resulting 3 strains were compared and found to carry notably SNPs in the sensor NisK gene [17]. This coupling of adaptive evolution and high throughput sequencing has been used in many other studies with LAB, e.g. experimental evolution of *Lactobacillus plantarum* when exposed to the murine digestive tract [135]. A recent report describes an elegant study with the plant isolate *Lactococcus lactis* KFI47 (see above) that propagated for 1000 generations in milk resulting in faster growth and biomass yields [188]. Three of the resulting strains were resequenced and found in two of the cases to have lost the conjugative transposon needed for growth in plants (see above). In the rest of the genome only few (6-28) mutations were detected in various genes, including those involved in amino acid production and transport. Remarkably, the strain with most mutations also contained a mutated mutL gene involved in mismatch repair and believed to increase the mutation frequency [188]. This example illustrates not only the power of experimental evolution and the used sequencing technology but also highlights the domestication process of a plant strain to the dairy environment.

A final but appealing approach where applied genomics has been used is the in the selection for *Lactococcus lactis* strains [189]. Cells of the strain MG1363 were mutagenized and serially propagated in water-in-oil emulsions to allow for selection of strains with increased biomass yield. One of the resulting strains coupled an increased biomass to slightly different growth kinetics and the conversion from homolactic into a mixed acid fermentation. Genomic ressequencing revealed a SNP mutation in the *ptnC* gene, encoding a component of the glucose PTS transport system. The phenotype of this mutant is explained by decreased glucose uptake rates, resulting in less acidification and higher yields without pH control. A series of revertants were also isolated that upon genomic ressequencing were found to contain an IS905 copy inserted in front of the *ptnABCD* operon, resulting in upregulation of the glucose PTS transport [189]. While these experiments generated further insight in fundamental aspects of the adaptation processes they also represent the proof of concept on how to use high throughput screening and sequencing allowing rapid analysis of the results. The examples of applied genomics described here are only a few of the possibilities that can be envisaged. Notably, strain optimization in combination with genomic re-sequencing will be a highly useful tool for improving starter strains or LAB marketed as probiotics. As natural or induced mutations do not lead to genetically modified organisms, the generated and improved strains can be used immediately for food or pharmaceutical applications.
Concluding remarks

Benefiting from the rapid development of next generation sequencing techniques, multiple genome sequencing projects on LAB were initiated since the beginning of the millennium. The data available up to now provide a comprehensive view on the complexity of the heterogeneous LAB group (Figure 2). Detailed comparative analysis of these genomic data emphasized the remarkable diversity within the LAB group at numerous taxonomic levels, *i.e.* order, family, group, genus and even species. This diversity results from the interactions between genome and environment as is schematically depicted (Figure 4). The abundance and variety of nutrients available in a habitat has a direct impact of the catabolic and biosynthetic properties of LAB. In many LAB species, the loss of metabolic genes is compensated by genome enrichment in genes encoding for transporters (ABC or PTS systems), allowing LAB to use nutrients and by-products from their niche. This specialization is evident from genome size reduction, presence of pseudogenes, and genome decay. Still, other LAB species or strains maintain a broad ecological flexibility, which may cause a high resilience to drastic environmental changes.

Because LAB are heterotrophs they have developed intimate interactions with plants and, most likely later, with animals and humans (Figure 1). Host-associated LAB contain a large and diverse repertoire of interaction proteins to adhere and signal to the host. It is tempting to speculate that the GI tract, as the site where plants enter the animal body, has played an important role in this evolutionary process. LAB adapted to the food environment may not require interaction with any host and therefore would generally possess a distinct repertoire of cell surface proteins. Thus, alternative surface proteins may be involved in the interactions between LAB and food constituents as compared to the interplay with the host mucosa [190]. Horizontal gene transfer appears a major driver of the genomic diversity and plasticity, affecting genome size and the acquisition of new genes. Plasmids of different sizes (up to mega-plasmids) and conjugative transposons have found to be involved in gene gain and loss.

Surviving in a niche also means to compete with other microbes and to defend against other inhabitants, including bacteriophages. The controlled production of organic acids and antimicrobials is a highly effective strategy in this microbiological warfare. Moreover, LAB harbor
CRISPR-cas systems to protect from bacteriophages and other foreign DNA. It seems that the loss of these defense systems may promote the promiscuous transfer of various traits, including antibiotic resistance or virulence factors. Finally, tolerance and resistance systems to endure physico-chemical properties, such as temperature, acid, salt or bile salts, are essential for LAB living in foods, the GI tract or other harsh environments.

The area of host-microbe, microbe-microbe and microbe-molecule interaction is a highly relevant and timely theme, notably in view of the rapidly expanding interest in the human GI tract [191]. It may be expected that the insight worked out for LAB may serve as model for other microbes. Moreover, as many LAB have immediate application potential, these systems also may result in improved or novel strains or processes, as seen for the discovery of peptide-based quorum sensing in Lactococcus lactis [192]. Some of the models with impact at various levels include the CRISPR-cas system discovered in Streptococcus thermophilus [193], the communication of Lactobacillus plantarum with the human host [129], the production of host-interacting pili in Lactobacillus rhamnosus [69], the evolution of metabolic strategies in Lactococcus lactis [189] or the finding of a novel metal-depend ing lactate racemase in Lactobacillus plantarum that is widely distributed [194]. The discovery of these models has relied for a large part on functional genomics, stressing the importance of this approach in LAB. This provides a promising outlook for the future where soon all LAB species will be characterized at the genomic level, many strains will have been genetically-modified organisms; LAB, lactic acid bacteria; NSLAB, Non-Starter LAB.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We are grateful to Prof. Todd R. Klaenhammer for his excellent contribution to the research on lactic acid bacteria for food and health during the last 40 years, his inspiration to many young and not-so young scientists, and his approval to us to contribute to this special issue. WmDvS was sponsored by the European Research Council (ERC, grant 250712 - Microbes Inside), the Academy of Finland (grant number 141130) and the Gravity and Spinoza Awards of the Netherlands Organization for Scientific Research. FPW was supported by the Academy of Finland (postdoctoral research fellowship, grant number 252123). We also thank Dr. Jarmo Ritari for generating the phylogenetic tree. We apologize that we could not cite all relevant research work on LAB functional genomics due to space limitations and time constraints.

Declarations

This article has been published as part of Microbial Cell Factories Volume 13 Supplement 1, 2014: Proceedings of the 11th International Symposium on Lactic Acid Bacteria. The full contents of the supplement are available online at http://www.microbialcellfactories.com/supplements/13/S1. Publication charges for this supplement were funded by the Stichting Symposium on Lactic Acid Bacteria. The articles have undergone the journal’s standard peer review process for supplements. EJ is employed by Chr Hansen A/S, is a shareholder in Chr Hansen A/S, and has applied for patents relating to the content of an article which he did not edit but which forms part of this supplement. MK is an employee of Nizo. The other Supplement Editors declare that they have no competing interests.

Authors’ details

1Department of Veterinary Biosciences, University of Helsinki, Agnes Sjöberginkatu 2, 00790 Helsinki, Finland. Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands. 2Department of Bacteriology & Immunology, University of Helsinki, Haartmaninkatu 3, 00014 Helsinki, Finland.

Published: 29 August 2014

References

1. Tannock GW: Normal microflora: an introduction to microbes inhabiting the human body. Springers; 1995.
2. Vaughan EE, Heilig HGH, Ben-Amor K, de Vos WM: Diversity, vitality and activities of intestinal lactic acid bacteria and bifidobacteria assessed by molecular approaches. FEMS Microbiol Rev 2005, 29:477-490.
3. Zoetendal EG, Rajilé-Stojanová M, de Vos WM: High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. Gut 2008, 57:1605-1615.
4. Human Microbiome Project Consortium: Structure, function and diversity of the healthy human microbiome. Nature 2012, 486:207-214.
5. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichench C, Nielsen T, Pons N, Levenez F, Yamada T, et al: A human gut microbiota gene catalogue established by metagenomic sequencing. Nature 2010, 464:59-65.
6. Sekelj M, Rud A, Knutzen SH, Denstadli V, Westereng B, Naes T, Rudi K: Abrupt temporal fluctuations in the chicken fecal microbiota are explained by its gastrointestinal origin. Appl Environ Microbiol 2012, 78:2941-2948.
7. Frese SA, Mackenze DA, Peterson DA, Schmaltz R, Fangman T, Zhou Y, Zhang C, Benson AK, Cody LA, Mulholland F, et al: Molecular characterization of host-specific biofilm formation in a vertebrate gut symbiont. PLoS Genet 2013, 9:e1004057.
8. Saxelin M, Tynkkynen S, Mattila-Sandholm T, de Vos WM: Probiotic and other functional microbes: from markets to mechanisms. Curr Opin Biotechnol 2003, 16:204-211.
9. Perschow R, Dorre J, Hörber P, Dinan T, Reid G, Blaser M, Cani PD, Degnan FH, Foster J, Gibson G, et al: Probiotics, prebiotics, and the host microbiome: the science of translation. Ann N Y Acad Sci 2013, 1296:1-17.
10. Tannock GW: A special fondness for lactobacilli. Appl Environ Microbiol 2004, 70:1389-1394.
11. Douglas LM: The bacillus of long life. London: TC & EC Jack; 1911.
12. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Mirecki JM, et al: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science 1995, 269:496-512.
13. Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, Ehrlich SD, Sorokin A: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403. Genome Res 2001, 11:731-753.
14. Keereezem M, Roekhorst J, van Kranenburg R, Molenaar D, Kuipers OP, Leer R, Tarchini R, Peters SA, Sandbrink HM, Fiers WMEJ, et al: Complete genome sequence of Lactococcus lactis plantarum WCFS1. Proc Natl Acad Sci USA 2003, 100:1990-1995.
15. Johnson BR, Klaenhammer TR: Impact of genomics on the field of probiotic research: historical perspectives to modern paradigms. Antonie Van Leeuwenhoek 2014.
16. de Vos WM: Systems solutions by lactic acid bacteria: from paradigms to practice. Microb Cell Fact 2011, 10(Suppl 1):S2.
17. Linares DM, Kok J, Poolman B. Genome sequences of Lactococcus lactis MG1363 (revised) and NZ9000 and comparative physiological studies. J Bacteriol 2010, 192:5806-5812.
18. de Ruyter PG, Kuipers OP, de Vos WM. Controlled gene expression systems for Lactococcus lactis with the food-grade inducer nisin. Appl Environ Microbiol 1996, 62:3662-3667.
19. Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, et al. Comparative genomics of the lactic acid bacteria. Proc Natl Acad Sci USA 2003, 103:15611-15616.
20. van den Bogaard PTC, Kleerebezem M, Kuipers OP, de Vos WM, van Hijum SAFT, Molenaar D, Siezen RJ, Bayjanov J, Renckens B, Wels M, et al. Complete genome of Lactobacillus lactis subsp. cremoris UC5099, host for a model lactococcal P335 bacteriophage. Genome Announc 2013, 1:.
21. Kelly WJ, Altemann E, Lambeke SC, Leboy SC. Interaction between the genomes of Lactobacillus lactis and phages of the P335 species. Front Microbiol 2013, 4:.
22. Siezen RJ, Bayjanov J, Renckens B, Wels M, van Hijum SAFT, Molenaar D, Siezen RJ, Bayjanov J, Renckens B, Wels M, et al. Complete genome of Lactobacillus lactis subsp. cremoris UC5099, host for a model lactococcal P335 bacteriophage. Genome Announc 2013, 1:.
23. L. Muscariello, C. Marino, C. Capri, U. Vastano, V. Marasco, R. Sacco, M. Mazzeo, M. Cacace, G. Peluso, A. Zotta, L. Muscariello, V. Vastano, E. Parente, et al. Overall control of nitrogen metabolism in Lactococcus lactis by CodY, and possible models for CodY regulation in Firmicutes. Microbiology 2005, 151:3895-3909.
24. K. Makarova, A. Slesarev, Y. Wolf, A. Mirkin, B. Koonin, E. Pavlov, A. Pavlova, N. Karamychev, V. Polouchine, N. Sorokin, A., et al. Comparative genomics of the lactic acid bacteria. Proc Natl Acad Sci USA 2003, 103:15611-15616.
25. Z. Papanikolaou, M. Hatzikamari, P. Georgakopoulos, M. Yiangou, M. Litopoulou-Tzaneti, E. Tzaneti, N. Zomer, A. de Jong, H. Estevao, T. Hoogenboezem, T. de Jong, A., et al. Complete genome sequence of Lactobacillus helveticus CNRZ 32, an industrial cheese starter and flavor adjunct. Genome Announc 2013, 1:.
26. J. Barrangou, R. Azcarate-Peril, M. A. D. Kelly, R. M. W. de Vos, P. Stouthalet, H. Estevao, T. Hoogenboezem, T. de Jong, A., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
27. A. Berks, B. C. De Wilt, L. Kleerebezem, M. de Jong, A., et al. Global transcriptome analysis of Lactococcus lactis cremoris KF147, a plant-associated lactic acid bacterium. Proc Natl Acad Sci USA 2006, 103:15615-15620.
28. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. cremoris SK11. Journal of Bacteriology 1996, 178:1525-1531.
29. D. J. Broadbent, J. E. Hughes, J. E. Welker, D. L. Tompkins, T. A. Tompkins, T. A., et al. Complete genome sequence of Lactobacillus sakei B2649-3550. Journal of Bacteriology 1998, 180:3769-3779.
30. Z. Papanikolaou, Z. Hatzikamari, M. Georgakopoulos, P. Yiangou, M. Litopoulou-Tzaneti, E. Tzaneti, N. Zomer, A. de Jong, H. Estevao, T. Hoogenboezem, T. de Jong, A., et al. Complete genome sequence of Lactobacillus helveticus CNRZ 32, an industrial cheese starter and flavor adjunct. Genome Announc 2013, 1:.
31. J. Barrangou, R. Azcarate-Peril, M. A. D. Kelly, R. M. W. de Vos, P. Stouthalet, H. Estevao, T. Hoogenboezem, T. de Jong, A., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
32. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactobacillus helveticus CNRZ 32, an industrial cheese starter and flavor adjunct. Genome Announc 2013, 1:.
33. J. Barrangou, R. Azcarate-Peril, M. A. D. Kelly, R. M. W. de Vos, P. Stouthalet, H. Estevao, T. Hoogenboezem, T. de Jong, A., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
34. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactobacillus helveticus CNRZ 32, an industrial cheese starter and flavor adjunct. Genome Announc 2013, 1:.
35. J. Barrangou, R. Azcarate-Peril, M. A. D. Kelly, R. M. W. de Vos, P. Stouthalet, H. Estevao, T. Hoogenboezem, T. de Jong, A., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
36. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
37. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
38. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
39. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
40. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
41. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
42. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
43. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
44. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
45. E. D. van den Bogaard, P. T. C., Kleerebezem, M., Kuipers, O. P., de Vos, W. M., et al. Complete genome sequence of Lactococcus lactis subsp. lactis IL1403. Journal of Dairy Research 2011, 78:247-257.
Ecological role of lactobacilli in the gastrointestinal tract:

58. Kankainen M, Paulin L, Tynkkynen S, von Ossowski I, Reunanen J,
59. Pedersen NR, Gauft P, Lechardeur D, Pett M, Strass A. Aerobic
respiration metabolism in lactic acid bacteria and uses in biotechnology.
60. Ann Rev Food Sci Technol 2012, 3:37-58.

61. Zhai Z, Douillard FP, An H, Wang G, Guo X, Luo Y, Hao Y. Protonic
characterization of the acid tolerance response in Lactobacillus
delbrueckii subsp. bulgaricus CAUH1 and functional identification of a
62. novel acid stress-related transcriptional regulator Ldb0677. Environ
Microbiol 2013, 16:1524-1537.

63. Wunschke A, Hammer E, Bartholomei M, Volker U, Burkovski A, Seidel G,
64. Hillen W. CcppA forms complexes with CodY and RpoA in Bacillus subtilis.
FEBS J 2012, 279:2201-2214.

65. Winstedt L, Frankenberg L, Hederstedt L, von Wachenfeldt C:
66. The synthesis and function of the alarmone (p)pGpp in firmicutes.
67. J Bacteriol 2007, 189:5203-5209.

68. Woll C, Geiger T, Goerke C. The structure of CodY, a GTP- and
69. isocitruic-acid-responsive regulator of stationary phase and virulence in gram-positive bacteria.
70. J Biol Chem 2006, 281:11366-11137.

71. Geiger T, Wolz C. Intersection of the stringent response and the CodY
72. regulon in low GC Gram-positive bacteria. Int J Med Microbiol 2014,
73. 304:150-155.

74. Wolz C, Geiger T, Goerke C. The synthesis and function of the alarmone
75. (p)pGpp in firmicutes. Int J Med Microbiol 2010, 300:142-147.

76. Wolk C, Paulin L, Tynkkynen S, von Ossowski I, Reunanen J,
77. Partela CAF, Smart KF, Tumanov S, Cook GM, Villas-Boas SG. The global
78. metabolic response of Enteroctococcus faecalis to oxygen. J Bacteriol 2014.

79. Brooomans RJW, Poolman B, Schuurman-Wolters GK, de Vos WM,
80. Hugenholtz J. Generation of a membrane potential by Lactococcus
81. lactis through aerobic electron transport. J Bacteriol 2007,
82. 189:5203-5209.

83. Oxarai V, Lede-Cler F, Dieye H, Henry JM, Pechoux C, Meythrec U,
84. Brandert R, Juillard V, Pard J C. Plus biogenesis in Lactococcus lactis:
molecular characterization and role in aggregation and biofilm
85. formation. PLoS One 2012, 7:e03589.

86. Kankainen M, Paulin L, Tynkkynen S, von Ossowski I, Reunanen J,
87. Partanen S, Takatalo V, Rushton M, Tkalec S, Lehtola J. Comparative genomic analysis of Lactobacillus rhamnosus GG reveals pil
88. containing a human- mucus binding protein. Proc Natl Acad Sci U S A
89. 2009, 106:17193-17198.

90. Segers ME, Lebeer S. Lactobacillus rhamnosus GG - host interactions.
91. Microb Cell Fact 2014.

92. Meyrand M, Guillet A, Gou M, Furlan S, Armalyte J, Kulakauskas S, Cortes-
93. Perez NG, Thomas G, Chat S, Pechoux C, et al. Surface proteome analysis of a natural isolate of Lactococcus lactis reveals the presence of pil
94. able to bind human intestinal epithelial cells. Mol Cell Proteomics 2013,
95. 12:3935-3947.

96. Vaughn EE, de Vries MC, Zoetendal EG, Ben-Amor K, Akkermans ADL, de
97. Vos WM. The intestinal LABs. Antonie Van Leeuwenhoek 2002,
98. 82:341-352.

99. Walter J. Ecological role of lactobacilli in the gastrointestinal tract:
implications for fundamental and biomedical research. Appl Environ
Microbiol 2008, 74:4545-4549.

100. Bello FD, Walter J, Hammes WP, Hertel C. Increased complexity of the system of processes of lactic acid bacteria in human feces revealed by
alternative incubation condition. Microb Ecol 2003, 45:455-463.

101. Heilig HGHU, Zoetendaal EG, Vaughn EE, Manteau P, Akkermans ADL, de
102. Vos WM. Molecular diversity of Lactobacillus spp. and other lactic acid
bacteria in the human intestine as determined by specific amplification of 165 ribosomal DNA. Appl Environ Microbiol 2002, 68:114-123.

103. Ahnne S, Nobaek S, Jeppsson B, Adlerberth I, Wold AE, Molin G. The normal
104. Lactobacillus flora of healthy human rectal and oral mucosa. J Appl
Microbiol 1996, 85:98-94.

105. Molin G, Jeppsson B, Johansson ML, Ahnne S, Nobaek S, Stahl M, Bengmark S. Numerical taxonomy of Lactobacillus spp. associated with
healthy and diseased mucosa of the human intestines. J Appl Bacteriol
1993, 74:314-323.

106. Klijn N, Weerkamp AH, de Vos WM. Genetic marking of Lactococcus lactis shows its survival in the human gastrointestinal tract. Appl Environ
Microbiol 1995, 61:2771-2774.
microbiota, Lactobacillus rhamnosus GG and serum lipids isolated by integrated analysis of high-throughput profiling data. PeerJ 2013, 26.

100. Selle K, Kleihammer TR: Genomic and phenotypic evidence for probiotic influences of Lactobacillus gasseri on human health. FEMS Microbiol Rev 2013, 37:915-935.

101. Denou E, Priemore RD, Berger B, Panoff JM, Arigoni F, Brussow H: Identification of genes associated with the long-persistence phenotype of the probiotic Lactobacillus johnsonii strain NCC533 using a combination of genomics and transcriptional analysis. J Bacteriol 2008, 190:3161-3168.

102. Toh H, Oshima K, Nakano A, Takahata M, Murakami M, Takaki T, Nishiyama H, Igimi S, Haraki M, Monta H: Genomic Adaptation of the Lactobacillus casei group. PLoS ONE 2013, 8:e75073.

103. Cai H, Rodríguez BT, Zhang W, Broadbent JR, Steele JL: Genotypic and phenotypic characterization of Lactobacillus casei strains isolated from different ecological niches suggests frequent recombination and niche specificity. Microbiol 2007, 153:2655-2665.

104. Rahfs EJ, Salvetti E, Tornani S, Felix GE, OTtoe PW: Genomic diversity of Lactobacillus salivarius. Appl Environ Microbiol 2011, 77:954-965.

105. Corr SC, Li Y, Riedel CU, OTtoe PW, Hill C, Gahan CG: Bacteriocin production as a mechanism for the antiinfective activity of Lactobacillus salivarius UCC118. Proc Natl Acad Sci USA 2007, 104:7617-7621.

106. Claesson MJ, Li Y, Leahy S, Canchaya C, van Pijkeren JP, Cerdeño-Tárraga AM, Parkhill J, Flynn S, O’Sullivan GV, Collins JK, et al: Multireplicon genome architecture of Lactobacillus acidophilus. Proc Natl Acad Sci USA 2006, 103:6183-6187.

107. Naito S, Hayashidani H, Kano K, Ogawa M, Benno Y: Development of intestinal lactobacilli in normal piglets. J Appl Bacteriol 1995, 79:230-236.

108. Molin G, Johansson ML, Ståhl M, Ahmed S, Andersson R, Jeppsson B, Bengmark S: Systematics of the Lactobacillus population on rat intestinal mucosa with special reference to Lactobacillus reuteri. Antonie Van Leeuwenhoek 1992, 61:175-183.

109. Frese SA, Benson AK, Takahata M, Murakami M, Takaki T, Heng NCK, Patil PB, Juge N, Bengmark S: Phenotypic characterization of different ecological niches suggests frequent recombination and niche specificity. Microbiol 2007, 153:2655-2665.

110. Moran NA, Plague GR: Characterization of pro-inflammatory flagellin proteins of Lactobacillus plantarum W65. Front Microbiol 2013, 3:1357-1374.

111. De Vos WM, Tynkkynen S, Kalkkinen N, Yamanen P: Effect of acid stress on protein expression and phosphorylation in Lactobacillus rhamnosus GG. J Proteomics 2012, 75:1357-1374.

112. Cai H, Rodríguez BT, Zhang W, Broadbent JR, Steele JL: Genotypic and phenotypic characterization of Lactobacillus casei strains isolated from different ecological niches suggests frequent recombination and niche specificity. Microbiol 2007, 153:2655-2665.

113. Alcántara C, Zúñiga M: Proteomic and transcriptomic analysis of the response to bile stress of Lactobacillus casei BL23. Microbiology 2012, 158:1206-1218.

114. Moen NA, Plague GR: Transformation of host restriction in Lactobacillus salivarius UCC118. PLoS Genet 2011, 7:e1001314.

115. Forde BM, McEntee E, O’Donnell MM, Riberon-Bissot E, Claesson MJ, Coghlan A, Ross RP, OTtoe PW: Genome sequences and comparative genomics of two Lactobacillus rhamnosus strains isolated from the gastrointestinal tract of mice. Genomics 2008, 93:217-225.

116. Forde BM, McEntee E, O’Donnell MM, Riberon-Bissot E, Claesson MJ, Coghlan A, Ross RP, OTtoe PW: Genome sequences and comparative genomics of two Lactobacillus rhamnosus strains isolated from the gastrointestinal tract of mice. Genomics 2008, 93:217-225.

117. Moen NA, Plague GR: Transformation of host restriction in Lactobacillus salivarius UCC118. PLoS Genet 2011, 7:e1001314.

118. Alcántara C, Zúñiga M: Proteomic and transcriptomic analysis of the response to bile stress of Lactobacillus casei BL23. Microbiology 2012, 158:1206-1218.

119. Goh YJ, Kleihammer TR: A functional glycan biosynthesis pathway in Lactobacillus acidophilus: expression and analysis of the gly operon. Mol Microbiol 2013, 89:1187-1200.

120. Bron PA, Grangette C, Mercier J, Van de WM, Kleerebezem M: Identification of Lactobacillus plantarum genes that are induced in the gastrointestinal tract of mice. J Bacteriol 2004, 186:5721-5729.

121. Otten J, van Baarlen P, Troost FJ, Molenaar D, Ahrné S, de Vos WM: Comparative proteomic analysis of Lactobacillus plantarum in the duodenum of healthy humans correlating with immune tolerance. Proc Natl Acad Sci USA 2009, 106:2371-2376.

122. Marco ML, Peters THF, Bongers RS, Molenaar D, Van Hemert S, van der Meer C, de Vos WM, van Baarlen P, Troost FJ, Molenaar D, Ahrné S, de Vos WM: Comparative proteomic analysis of Lactobacillus plantarum in the duodenum of healthy humans correlating with immune tolerance. Proc Natl Acad Sci USA 2009, 106:2371-2376.

123. Koskenniemi K, Laasko K, Kankainen M, Savijärvi K, Nyman TA, Surakka A, Salusjärvi T, et al: Proteomics and transcriptomics characterization of bile stress response in probiotic Lactobacillus rhamnosus GG. Mol Cell Proteom 2011, 10.

124. Goh YJ, Kleihammer TR: A functional glycan biosynthesis pathway in Lactobacillus acidophilus: expression and analysis of the gly operon. Mol Microbiol 2013, 89:1187-1200.

125. Bron PA, Grangette C, Mercier J, Van de WM, Kleerebezem M: Identification of Lactobacillus plantarum genes that are induced in the gastrointestinal tract of mice. J Bacteriol 2004, 186:5721-5729.

126. Toole PW: Genomic changes following host restriction in Lactobacillus salivarius. Proc Natl Acad Sci USA 2007, 104:7617-7621.

127. McKendred GG, Harper GM, Nickle J, Cui Y, Neale LJ, O’Toole PW: Identification of the expressed gene products in Lactobacillus salivarius UCC118. FEMS Microbiol Lett 2007, 271:1481-1484.

128. de la Riva B, Muñoz R, López de Felipe F: Genome-wide transcriptomic analysis of a human isolate of Lactobacillus plantarum exposed to p-coumaric acid stress. Molecular Nutrition & Food Research 2012, 56:1848-1859.

129. Bron PA, Marco M, Hoffner SM, van Mullekom E, van de WM, Kleerebezem M: Genetic characterization of the bile salt response of Lactobacillus plantarum and Lactobacillus johnsonii. Microbiology 2011, 157:2047-2056.

130. Marco ML, Peters THF, Bongers RS, Molenaar D, Van Hemert S, van der Meer C, de Vos WM, van Baarlen P, Troost FJ, Molenaar D, Ahrné S, de Vos WM: Comparative proteomic analysis of Lactobacillus plantarum in the duodenum of healthy humans correlating with immune tolerance. Proc Natl Acad Sci USA 2009, 106:2371-2376.

131. Marco ML, Peters THF, Bongers RS, Molenaar D, Van Hemert S, van der Meer C, de Vos WM, van Baarlen P, Troost FJ, Molenaar D, Ahrné S, de Vos WM: Comparative proteomic analysis of Lactobacillus plantarum in the duodenum of healthy humans correlating with immune tolerance. Proc Natl Acad Sci USA 2009, 106:2371-2376.

132. Reverón I, de las Rivas B, Muñoz R, López de Felipe F: Genome-wide transcriptomic analysis of a human isolate of Lactobacillus plantarum exposed to p-coumaric acid stress. Molecular Nutrition & Food Research 2012, 56:1848-1859.

133. de la Riva B, Muñoz R, López de Felipe F: Genome-wide transcriptomic analysis of a human isolate of Lactobacillus plantarum exposed to p-coumaric acid stress. Molecular Nutrition & Food Research 2012, 56:1848-1859.

134. de la Riva B, Muñoz R, López de Felipe F: Genome-wide transcriptomic analysis of a human isolate of Lactobacillus plantarum exposed to p-coumaric acid stress. Molecular Nutrition & Food Research 2012, 56:1848-1859.
Douillard and de Vos Microbial Cell Factories 2014, 13(Suppl 1):S8
http://www.microbialcellfactories.com/content/13/S1/S8

Page 20 of 21

138. Denou E, Berger B, Barretto C, Panoff JM, Arigoni F, Brussow H: Gene expression of commensal Lactobacillus johnsonii strain NCC533 during in vitro growth and in the murine gut. J Bacterial. 2007, 189:8109-8119.

139. Pandharik S, Magopa T, Lanson FG, de Bruyn G, Gray GE, Hammarstrom L, Marcotte H: Identification and characterisation of vaginal lactobacilli from South African women. BMC Infect Dis. 2013, 13:189.

140. Martinez-Peña MD, Castro-Escarpulli G, Aguillera-Areola MG: Lactobacillus species isolated from vaginal secretions of healthy and bacterial vaginosis-intermediate Mexican women: a prospective study. BMC Infect Dis. 2013, 13:189.

141. Zhou X, Hanmann MA, Davis CC, Suzuki H, Brown CJ, Schutte LJ, Pierson JD, Forney LJ. The vaginal bacterial communities of Japanese women resemble those of women in other racial groups. FEMS Immunol Med Microbiol. 2010, 58:169-181.

142. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SSK, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, et al. Vaginal microbiome of reproductive-aged women. Proc Natl Acad Sci USA. 2011, 108(Suppl 1):4680-4687.

143. Zhou X, Brown CJ, Abdo Z, Davis CC, Hanmann MA, Joyce P, Foster JA, Forney LJ. Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. ISME J. 2007, 1:121-133.

144. Drell T, Lillar T, Tummelseth L, Simm J, Aasöll A, Vain E, Saarma I, Dianatpour M, Modarressi MH: Lactobacillus fermentum isolated from the human vagina. BMC Microbiol. 2012, 12:64.

145. Hanlon DE, Moench TR, Cone RA: Vaginal pH and microbicidal lactic acid produced by lactobacilli from South African women. J Dairy Sci. 2013, 96:4967-4968.

146. Hanlon DE, Moench TR, Cone RA: Microbiome-sensitive and specific molecular diagnosis of bacterial vaginosis in women with history of recurrences, undergoing surgical menopause; a prospective pilot study. Eur Rev Med Pharmacol Sci. 2013, 17:1399-1403.

147. Garnier P, Lestrade JL, Brouqui F, Hambali S, Sher M, Hotez PJ, Laxminarayan R, Falagas ME. The highly autoaggregative and adhesive phenotype of the vaginal Lactobacillus plantarum strain CMGP5300 is ciprofloxacin resistant. Appl Environ Microbiol. 2013, 79:4576-4584.

148. Bobbot JM, Cardot JM: Vaginal impact of the oral administration of total freeze-dried culture of LCR 35 in healthy women. Infect Dis Obstet Gynecol. 2012, 2012:30648.

149. Pomplin S, Hentati M, Caputo L, Redaelli A, Quaranta L, Candidini M. The role of vaginal Lactobacillus Rhamnosus (Normobin(R)) in preventing Bacterial Vaginosis in women with history of recurrences, undergoing surgical menopause; a prospective pilot study. Eur Rev Med Pharmacol Sci. 2013, 17:1399-1403.

150. Colodner R, Edelstein H, Chazan B, Raz R: Vaginal colonization by orally administered Lactobacillus rhamnosus GG. Isr Med Assoc J. 2003, 5:767-769.

151. EFSA BIOHAZ Panel (EFSA Panel on Biological Hazards). Scientific Opinion on the maintenance of the list of OPS biological agents intentionally added to food and feed (2013 update). EFSA Journal. 2013, 11:108.

152. Foulquè Moreno MR, Sarantinopoulos P, Taiskalud E, De Vuyst L. The role and application of enterococci in food and health. Int J Food Microbiol. 2006, 106:1-24.

153. Lebreton F, Willems RJL, Gilmore MS: Enterococcus Diversity, Origins in Nature, and Gut Colonization. In Enterococci: from commensals to leading causes of drug resistant infection. Boston;Gilmore MS, Clewell DR; Ise E, Shankar N 2014.

154. Gilmore MS, Lebreton F, van Schaik W: Genomic transition of enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era. Can Open Microbiol. 2013, 16:10-16.

155. Delaplain ME, Ewing L, Sloniger NA, Flavahan NA, Polasek AD, Holzapfel WH, Dewitt JD, Kass EH, Bumgarner RE: Comparative functional analysis of lactic acid bacteria in shrimp intestines. Plos One. 2012, 7:e30187.

156. Lukanceino O, Ussey DW, Wassenaar TM: Comparative genomics of Bifidobacterium, Lactobacillus and related probiotic genera. Microb Ecol 2012, 63:651-673.

157. Gouriet F, Million M, Henri M, Fournier PE, Raoult D: Ratner AJ: Inerolysin, a cholesterol-dependent cytolysin produced by Lactobacillus johnsonii NCC533 during in vitro growth. J Clin Microbiol. 2010, 48:3563-3568.

158. Santer M: Comparative functional genomics of Lactobacillus strains to unravel the cause of infectious disease of humans. Notes Rec R Soc Lond. 2014, 69:614-623.

159. Bouwstra V, van den Hoogen M, de Vries F, van der Zee S, Heuer M, van der Drift C, Paul S: Lactobacillus rhamnosus as a model organism to illustrate the cause of infectious disease. Microb Ecol Health Dis. 2013, 24:290-296.

160. Santer M: Lactobacillus rhamnosus LB5111 and related probiotic genera. Microb Ecol Health Dis. 2014, 25:269-280.

161. Small G, Sneath PHA, van der Zee S, van den Hoogen M, van der Drift C, Paul S: Lactobacillus rhamnosus as a model organism to illustrate the cause of infectious disease of humans. Notes Rec R Soc Lond. 2014, 69:614-623.

162. Dürr J, Hanley DE, Moench TR, Cone RA: Vaginal pH and microbicidal lactic acid when lactobacilli dominate the microbial population. Appl Environ Microbiol. 2013, 79:4680-4690.

163. Sukhov E, Huisman T, de Vos WM, van der Sande AM: Unraveling the microbiota of teat apices of clinically healthy dairy cows. PLoS One. 2012, 7:e38108.

164. Akkermans ADL, de Vos WM, van der Sande AM: Unraveling the microbiota of teat apices of clinically healthy dairy cows. PLoS One. 2012, 7:e38108.

165. Akkermans ADL, de Vos WM, van der Sande AM: Unraveling the microbiota of teat apices of clinically healthy dairy cows. PLoS One. 2012, 7:e38108.

166. Akkermans ADL, de Vos WM, van der Sande AM: Unraveling the microbiota of teat apices of clinically healthy dairy cows. PLoS One. 2012, 7:e38108.
Molecular analysis of the bacterial microbiota in the human stomach. 

Perez G, Blaser MJ, Relman DA: 

Future Microbiol 2009, 4:153.

Bolotin A, Quinquis B, Renault P, Sorokin A, Ehrlich SD, Kulakauskas S, Lapidus A, Goeden M, Mauzerall D, Saier MH Jr, et al: 

Complete resequencing and reannotation of the Lactobacillus plantarum WCFS1 genome. 

Mol Microbiol 2006, 61:1137-1152.

Bolotin A, Quinquis B, Renault P, Sorokin A, Ehrlich SD, Kulakauskas S, Lapidus A, Goeden M, Mauzerall D, Saier MH Jr, et al: 

BLAST-EXPLORER helps you building datasets for phylogenetic analysis. 

BMC Evol Biol 2010, 10:8.

Chevenet F, Brun C, Bañuls AL, Jacq B, Christen R: 

SeqinR 1.0-2: a contributed package to the R project representing biological sequences, and matching algorithms. 

Bioinformatics 2004, 20:172-179.

Perez G, Blaser MJ, Relman DA: 

Future Microbiol 2009, 4:153.

Bolotin A, Quinquis B, Renault P, Sorokin A, Ehrlich SD, Kulakauskas S, Lapidus A, Goeden M, Mauzerall D, Saier MH Jr, et al: 

BLAST-EXPLORER helps you building datasets for phylogenetic analysis. 

BMC Evol Biol 2010, 10:8.

Chevenet F, Brun C, Bañuls AL, Jacq B, Christen R: 

SeqinR 1.0-2: a contributed package to the R project representing biological sequences, and matching algorithms. 

Bioinformatics 2004, 20:172-179.

Perez G, Blaser MJ, Relman DA: 

Future Microbiol 2009, 4:153.