ABSTRACT

Flavor is one of the most important attributes of any fermented dairy product. Dairy consumers are known to be willing to experiment with different flavors; thus, many companies producing fermented dairy products have looked at culture manipulation as a tool for flavor diversification. The development of flavor is a complex process, originating from a combination of microbiological, biochemical, and technological aspects. A key driver of flavor is the enzymatic activities of the deliberately inoculated starter cultures, in addition to the environmental or “nonstarter” microbiota. The contribution of microbial metabolism to flavor development in fermented dairy products has been exploited for thousands of years, but the availability of the whole genome sequences of the bacteria and yeasts involved in the fermentation process and the possibilities now offered by next-generation sequencing and downstream “omics” technologies is stimulating a more knowledge-based approach to the selection of desirable cultures for flavor development. By linking genomic traits to phenotypic outputs, it is now possible to mine the metabolic diversity of starter cultures, analyze the metabolic routes to flavor compound formation, identify those strains with flavor-forming potential, and select them for possible commercial application. This approach also allows for the identification of species and strains not previously considered as potential flavor-formers, the blending of strains with complementary metabolic pathways, and the potential improvement of key technological characteristics in existing strains, strains that are at the core of the dairy industry. An in-depth knowledge of the metabolic pathways of individual strains and their interactions in mixed culture fermentations can allow starter blends to be custom-made to suit industry needs. Applying this knowledge to starter culture research programs is enabling research and development scientists to develop superior starters, expand flavor profiles, and potentially develop new products for future market expansion.

Key words: flavor, fermentation, genomics, whole-genome sequencing, metabolomics

INTRODUCTION

Flavor is primarily a combination of a product’s taste and aroma, influenced by visual, aural, and textural sensations, our expectations based on previous experiences, and our genetic ability to perceive flavor (Taylor and Roberts, 2004). The flavor of dairy products varies based on the complexity of the product—from raw milk, with a basic mild sweet flavor, to aged cheeses, with strong sensory characteristics derived principally from the activities of diverse microbial populations. The flavor of milk is influenced by a range of factors at the farm level and during processing. Milk composition and quality are critically important and it is now well established that ruminant diet and breed, among other factors, also directly and indirectly affect sensory perception (Kilcawley et al., 2018). Thus, any dairy products derived from milk are also potentially affected by these same factors. However, the effect of such factors likely diminishes in fermented dairy products because of the proliferation of added microbes that create taste and aroma compounds through their enzymatic and metabolic activity over time. In products where the flavor is primarily derived from externally added ingredients, such as in fruit-flavored yogurts, the sensory impact of the base dairy substrate is diminished. In addition, heat treatments and other intrusive processing techniques can result in the creation, reduction, and modification of both nonvolatile and volatile flavor components that alter sensory perception (Kilcawley et al., 2018). Flavor is a critical purchasing factor in all products and intrinsically linked to product quality (Velasco et al., 2015). Constancy of production in relation to flavor is vitally important in gaining—and maintaining—market share and can be more difficult to
attain in some fermented dairy products because of the dynamic nature of the product. Therefore, advances in the understanding of flavor pathways of the microbes used in the production of fermented dairy foods have greatly improved flavor consistency and quality and have also supported new product development.

Although many of the metabolic pathways associated with flavor compound generation in dairy product-associated microbes were studied comprehensively in the pre-genomics era, the availability of extensive numbers of genome sequences of these and other flavor-contributing species and their analysis through novel genomic technologies (Figure 1) has allowed the identification of novel pathways contributing to flavor generation, the examination of the genomic basis for the strain diversity observed between members of individual species, the elucidation of microbial metabolic networks that contribute to flavor compound development, and the possibility of functional genomic studies, which give an insight into the genomic basis of flavor and other industrially important features. Recent studies combining metagenomic and metabolomic approaches represent a novel approach to associating bacteria in a particular dairy product with the presence of the principal flavor-contributing compounds (Walsh et al., 2016; Bertuzzi et al., 2018).

In this article, we review developments in genomic research in relation to the development of flavor, specifi-
cally in relation to the starter and nonstarter bacteria and fungi that drive flavor development in fermented foods. We review new information on the metabolic potential of flavor-forming species afforded by genomic investigations and the possibilities for exploitation of this potential in industrial applications. We also discuss system-level approaches, integrating genomics and metabolomics, which set the stage for knowledge-based manipulation of the dairy microbiota for generation of products with desirable flavor characteristics.

THE ROLE OF MICROBES IN FLAVOR DEVELOPMENT

The generation of flavor in fermented dairy products is primarily driven by the starter and nonstarter microbiota of individual products and their metabolic activities (Ganesan and Weimer, 2016). These organisms arise either through natural contamination of the product raw materials or manufacturing environment or through deliberate inoculation of the raw materials, either through back-slopping (inoculation of the raw material with a sample from a previous batch) or by addition of specific culture blends (Thierry et al., 2016). Today, most large-scale, industrially manufactured products rely on commercially produced defined culture blends in which the strains have been specifically selected, cultivated, and blended under tightly controlled conditions to ensure optimized fermentation and quality attributes. In contrast, many fermented foods produced and consumed in developing countries are derived from natural, or spontaneous, fermentation (e.g., kimchi, kefir) and many artisanal products are produced using back-slopping methods without addition of specific starters (e.g., raw milk cheeses; Thierry et al., 2016).

In fermented dairy products, members of the lactic acid bacteria (LAB) are the most prevalent organisms responsible for fermentation (Smid and Kleerebezem, 2014). This is a diverse group of gram-positive microorganisms that produce lactic acid as the main product of sugar degradation. The LAB are classified in the order Lactobacillales and comprise species of the genera Lactococcus, Lactobacillus, Streptococcus, Leuconostoc, Oenococcus, and Pediococcus, among others (McAuliffe, 2018). Many of the members of the LAB group are afforded GRAS (generally regarded as safe) or QPS (qualified presumption of safety) status (Bourdichon et al., 2012). Their main function in dairy fermentations is to rapidly acidify the milk, thus improving the microbial quality and safety of the final product. Along with other non-LAB species, and certain yeasts and molds, they also have a significant effect on end-product flavor through key metabolic pathways and enzyme activities (Smid and Kleerebezem, 2014).

KEY METABOLIC PATHWAYS FOR FLAVOR FORMATION

From a microbial perspective, the metabolic pathways providing the key enzymes and driving the generation of flavor compounds in dairy fermentations are sugar metabolism (glycolysis), protein degradation (proteolysis), and fat breakdown (lipolysis) (Figure 2; McAuliffe, 2017). The LAB metabolize sugars or carbohydrates through fermentation, producing ATP by substrate-level phosphorylation without the requirement for a membrane or electron transport system. The conversion of lactose (the principal milk sugar) to glucose and galactose, and the subsequent generation of lactic acid from glucose through the glycolytic pathway is the foundation upon which all milk-based fermentations are based. Lactose can undergo either homofermentation or heterofermentation, depending on the LAB species, the substrate, and the environmental conditions. Homofermentation results in lactic acid as the main end product, the flavor of which contributes to the pleasant acidic taste of many fermented foods. Homofermentative LAB species include Lactococcus lactis, Streptococcus thermophilus, Pediococcus, Enterococcus, and certain species of Lactobacillus (e.g., Lactobacillus casei, Lactobacillus curvatus; Thierry et al., 2016). In addition to lactic acid, heterofermentative species produce a wide variety of fermentation products, including carbon dioxide, ethanol, acetic acid, and formic acid, the latter of which can act either as aroma compounds or precursors to aroma compounds. These microorganisms include Leuconostoc, Oenococcus, and Lactobacillus species such Lactobacillus brevis, Lactobacillus fermentum, and Lactobacillus reuteri. Sugars enter the pathways of glycolysis through either the Embden–Meyerhof–Parnas (EMP) pathway, in the case of homofermentative LAB, or the phosphoketolase pathway (PKP), in the case of heterofermentative species. A key intermediate in the generation of lactic acid through glycolysis is pyruvate, a fraction of which can be converted to various flavor compounds such as diacetyl and acetoin, 2 compounds that give butter its characteristic flavor and that are found in high concentrations in sour (cultured) cream, cultured buttermilk, and cultured butter. Metabolism of citrate by certain citrate-metabolizing (Cit+) dairy species, such as Lactococcus lactis ssp. lactis biovar diacetylactis and Leuconostoc mesenteroides, can also lead to the formation of diacetyl (Starrenburg and Hugenholtz, 1991). Other flavor compounds arising from pyruvate metabolism include acetaldehyde, the green
apple or nutty flavor responsible for the typical taste and aroma of yogurt (Chen et al., 2017), and acetic acid, a key flavor component in certain dairy products that can lead to off-flavors at higher concentrations.

Lipolysis, or the degradation of milk fats by lipases or esterases, results in the formation of free fatty acids, which act as precursors for certain flavor compounds such as methylketones, secondary alcohols, esters, and lactones. In general, LAB contribute little by way of lipolytic activity (Smit et al., 2005) but recently, certain “wild” strains of Lc. lactis have shown significantly higher levels of lipase activity than their domesticated counterparts (McCarthy, 2017). In mold-ripened cheeses, such as Camembert and Roquefort, the organisms used for ripening often have significant lipid-degradation capabilities and produce compounds such as n-

Figure 2. The primary microbial metabolic pathways leading to the generation of flavor in fermented dairy foods. EMP = Embden–Meyerhof–Parnas pathway; PP = pentose phosphate pathway; T6P = tagatose 6 phosphate pathway.
methyl-ketones, impact compounds in blue-veined and other cheese varieties (Martin and Coton, 2017). The impact compound for a particular flavor or aroma can be defined as the chemical substance that provides the principal sensory identity (McGorrin, 2012).

It is undoubtedly the proteolytic cascade—along with the addition of rennet—that is central to generating the key volatile and nonvolatile flavor compounds associated with most hard and semi-hard cheese types (Engels et al., 1997; Smit et al., 2005). Because of their predominantly auxotrophic nature, many LAB species harbor a potent proteolytic system that fulfills their requirements for AA (Liu et al., 2010). Short peptides are released from casein through the action of added rennet in conjunction with the cell envelope proteinases (CEP) of inoculated starter cultures. These act as substrates for starter-derived peptidases, which generate AA that are further converted to (1) aldehydes, derived from the decarboxylation of various α-keto acids; (2) alcohols, derived from aldehydes by dehydrogenation; (3) esters, derived from alcohols by esterases, and (4) sulfur compounds, derived from methionine through the action of various enzymes. A general overview of the main metabolic pathways leading to the formation of some impact flavor compounds in milk-based products is shown in Figure 2.

**GENOMICS OF FLAVOR FORMATION IN CORE DAIRY SPECIES**

*Lactobacillus Species*

The *Lactobacillus* genus is one of the largest and most well researched groups of LAB and comprises more than 200 species and subspecies isolated from different habitats, including plant material, food, and the gastrointestinal tracts of humans and animals (Duar et al., 2017). Because of their fermentation capacity, lactobacilli have found numerous applications in the production of fermented dairy, meat, and plant products (Sun et al., 2015). In fermented dairy products, strains of the species *Lactobacillus delbrueckii* ssp. *bulgaricus* and *Lactobacillus helveticus* are mainly used as starter or adjunct cultures, with a considerable effect on flavor generation (El Kafsi et al., 2014). Additionally, non-starter lactobacilli such as *Lb. casei*, *Lb. paracasei*, *Lb. rhamnosus*, and less often *Lb. plantarum*, previously considered as adventitious dairy microflora, have been shown to contribute significantly to flavor development in certain dairy products (Milesi et al., 2010; Sgarbi et al., 2013; Gobbetti et al., 2015). Numerous comparative studies on the available genomes of lactobacilli showed that the common features of dairy isolates are genome minimization and metabolic simplification, with the accumulation of pseudogenes, repeats, and insertion sequences, leading to adaptation to the nutritionally rich environment of milk (Cai et al., 2009; O’Sullivan et al., 2009). Additionally, dairy isolates are characterized by increased numbers of proteolytic genes and AA and sugar transportation genes, which support efficient nutrient uptake by dairy isolates (Makarova et al., 2006).

In lactobacilli, as in other species of LAB, proteolysis is considered the main pathway for the development of flavor compounds. This cascade begins with the degradation of casein, the main milk protein, by the CEP. A comparative genomic study performed on 213 *Lactobacillus* species and associated genera, showed that CEP genes were present in 60 of the analyzed genomes (Sun et al., 2015). The distribution of the CEP genes corresponded with phylogenetic clades. Apart from different anchoring mechanisms to the cell wall, a combination of domains with different specificities for substrates can result in different ranges of metabolic products, eventually improving and diversifying product flavor (Sun et al., 2015). Unlike the majority of lactobacilli that usually possess a single CEP gene, strains of *Lb. helveticus*, such as CNRZ32, can carry up to 4 prt genes (Broadbent et al., 2013). Differences in numbers and substrate specificities of proteinases could be responsible for efficient proteolytic capacity of *Lb. helveticus*, which can be a determining element of flavor generation by these strains (Broadbent et al., 2011). In a recent *Lb. helveticus* comparative genome study, it was shown that *prtH* encoded the predominant CEP responsible for proteolysis, and strains with frameshift mutations in this gene did not show proteolytic activity (Schmid et al., 2018). Although CEP-negative strains are also used as flavor-enhancing strains, they are used in conjunction with CEP-positive strains that can provide the substrates for the downstream cascade responsible for the production of volatile flavor compounds.

Following transportation into the cells, peptides are metabolized further to free amino acids by general (i.e., PepN, PepC) and specific (i.e., PepA, PepS, PepL) aminopeptidases. Genome analysis of lactobacilli revealed that aminopeptidases are a common feature of the genus, including dairy-related species (Cai et al., 2009; Liu et al., 2010). Interestingly, although PepN is usually encoded by single genes, Pep C/E and PepO are present as multiple homologs, potentially contributing to adaptation of dairy strains to habitats rich in proteins, such as milk (Cai et al., 2009). Variation in the number and types of peptidases is evident even at the species level, as seen in the comparison of 4 complete *Lb. delbrueckii* genomes. In strain ND02, the highest number of genes encoding proteinases, pepti-
dases, and amino acid transporter systems was identified. The same strain also possessed 3 strain-specific intracellular peptidases. On the other hand, 2 surface peptidases present exclusively in strain 2038 imply a higher proteolytic capacity with a substantial effect on flavor compound generation (Liu et al., 2012).

Low-molecular-weight organic molecules, such as alcohols, aldehydes, ketones, and acids are some of main flavor compounds in dairy products (Yvon, 2006). They are generated by metabolism of free AA, which starts with the activity of aminotransferases that transfer amino groups from AA to α-keto acids. Genome analysis of AA metabolism enzymes in LAB, including 12 Lactobacillus species, showed that a single branched-chain aminotransferase (bcaT) gene was present in all dairy-related Lactobacillus strains, in contrast to the aromatic aminotransferase (araT) gene, which was usually present in a higher number of homologs (Liu et al., 2008). To support the activity of aminotransferase, glutamate dehydrogenase (GDH) recycles α-ketoglutarate, the main receptor of the amino group during transamination, and is thus indirectly linked to flavor compound generation. The majority of nonstarter LAB possess the gdh gene (Gobbetti et al., 2015). Conversely, sequenced Lb. delbrueckii strains were lacking this gene, but probably compensated by the production of glutamate from 2-oxoglutarate and L-aspartate by the activity of 2 enzymes homologous to aspartate aminotransferase (Liu et al., 2012).

Recently, a comparative genome analysis was performed on 3 Lb. paracasei strains isolated as members of the nonstarter microbiota of Cheddar cheese (Stefanovic and McAuliffe, 2018). In earlier phenotypic studies, these strains showed significant differences in their proteolytic enzyme activities and generated variable volatile profiles in different model systems (Stefanovic et al., 2017a,b). Although genomic comparison revealed substantial genetic variation between the strains, no major differences were observed in regard to the components of the proteolytic cascade. This suggests that any strain-specific differences observed in terms of the variation of the flavor compounds produced most likely result from enzyme activity, regulation, or both. However, an important difference was found in genes encoding enzymes involved in the development of volatile sulfur compounds (VSC), including methanethiol (MTL), dimethyl disulfide (DMDS), dimethyl trisulfide, and methional, which are generated by catabolism of the sulfur-containing AA methionine and cysteine and considered important cheese flavor contributors. Volatile sulfur compounds are impact flavor compounds in a variety of ripened cheeses. They include thiols such as MTL, a characteristic flavor note of French cheeses such as Camembert, often described as “cooked cabbage.” This flavor note is also found in Cheddar, Limburger, and Parmesan cheeses. Others include DMDS, a typical “garlic” odor associated with certain ripened cheeses, and S-methylthioesters such as S-methylthioacetate associated with French cheeses Vacherin and Pont-l’Évêque. Each of the 3 Lb. paracasei strains differed in the presence and number of homologs of cystathionine-β and cystathionine-γ lyases (CBL, CGL) and cystathionine-β synthase (CBS; Stefanovic and McAuliffe, 2018). Although their main activity is in methionine anabolism, these enzymes have been reported to contribute to flavor compound production (Fernandez and Zumiga, 2006; Bustos et al., 2011). The highest numbers of homologs encoding CBL, CGL, and CBS were detected in strain DPC4206, and this could be the determining factor for more efficient VSC production and a rationale for selection of this strain as an adjunct to support desirable flavor development (Stefanovic and McAuliffe, 2018; Stefanovic et al., 2018).

Apart from the proteolytic pathway, other metabolic reactions in Lactobacillus also contribute to development of flavor. For example, diacetyl and acetoin, low-molecular-weight ketones, contribute to buttery notes in dairy products (Curioni and Bosset, 2002). They arise from pyruvate metabolism, which is generated in the glycolytic/citrate pathway (Marilley and Casey, 2004). One of the key enzymes involved in the synthesis of these flavor compounds is acetolactate synthase, which converts pyruvate to acetolactate, a precursor of acetoin/diacetyl. Whole-genome sequencing (WGS) and comparison of 4 Lb. rhamnosus strains demonstrated that the genes that encode the enzymes of citrate metabolism are conserved. However, one strain was unable to produce either diacetyl or acetoin in milk. Further genetic analysis revealed that a frameshift mutation in the gene encoding acetolactate synthase (als) resulted in the inability to produce the diacetyl/acetoin precursor compound acetolactate. After the introduction of the wild-type als gene in this deficient strain, the production of the 2 flavor compounds was established (Lo et al., 2018).

The availability of whole-genome sequences of various species of Lactobacillus has enabled the establishment of genome-scale metabolic models (GSMM), mathematical models used to predict metabolic processes under various conditions based on the presence or absence of genes encoding enzymes involved in metabolic pathways. These GSMM have been used to predict the production of diacetyl and acetoin in L. casei LC2W in the presence of oxygen. It was also shown that acetoin accumulates in both aerobic and anaerobic conditions of cultivation, whereas diacetyl production is enhanced.
Lactococcus lactis

*Lactococcus lactis* is used as a starter organism in the production of many fermented dairy products such as soft and hard cheeses, buttermilk, and sour cream. Genomic analysis of “domesticated” strains of this species, strains evolved to the dairy niche, has revealed that the ancestor of these strains occupied a plant niche and that they have minimized their genome, with the loss of genetic systems not required for growth in milk (McAuliffe, 2018). Although its principal role in dairy fermentations is the production of lactate for acidification and coagulation of the milk, lactic acid also contributes to flavor, conferring a pleasant acidic taste often associated with many fermented dairy foods (Thierry et al., 2016). The principal route to lactic acid production in *Lc. lactis* is via the lactose-specific phosphotransferase system (lac-PTS) encoded by *lacEF*. This generates lactose phosphate, which is acted on by LacABCD of the tagatose-6-phosphate pathway and LacG, or phospho-β-galactosidase, to yield galactose-6-phosphate and glucose (Callanan and Ross, 2004). This metabolic capability is generally encoded by plasmids and was well characterized in the pre-genome era. However, WGS of various *Lc. lactis* strains has led to the discovery of alternative pathways for lactose catabolism in lactococci. Certain lactococcal strains utilize a lactose permease-β-galactosidase system that is chromosomally encoded, depending on enzymes encoded by the Leloir pathway operon, including *lacS*, *lacA*, and *lacZ*, the latter of which encodes β-galactosidase (Aleksandrzak-Piekarczyk, 2013). Lactose is transported in an unphosphorylated form and subsequently cleaved by β-galactosidase. Another lactose uptake system, based on the chromosomally encoded cellobiose-specific PTS system (PTSCel-Lac), has been identified in the plasmid-free strain *Lc. lactis* sspp. *lactis* IL1403 (Aleksandrzak-Piekarczyk et al., 2005, 2011). The presence of the *ptcB* and *ptcA* genes, along with *celB* and *bgIS*, which encodes an enzyme with both phospho-β-glucosidase and phospho-β-galactosidase activity, enables this strain to transport lactose. This system is also present in certain lactococcal strains that were previously described as being unable to utilize lactose, such as MG1363 (Solopova et al., 2012).

Gas-producing mesophilic cultures include *Lc. lactis* sspp. *lactis* biovar *diacetylactis*, a biovariant that is capable of catalyzing citrate to produce CO₂ along with the flavor compounds diacetyl and acetoin, and is used in the production of Edam, Gouda, and soft mold-ripened cheeses such as Camembert and Brie. Although these compounds can be generated from the pyruvate generated during glycolysis by the *lactis* and *cremoris* subspecies, the biovar *diacetylactis* produces these compounds via carbohydrate/citrate-co-metabolism (Pudlik and Lolkema, 2011). This organism contains the *citP* plasmid in association with the chromosomally encoded *cit* operon, *citM-citI-citCDEFXG*. The *citP* plasmid encodes CitP, a citrate permease which transports citrate into the cell where it is cleaved by citrate lyase (CitDEF) and its auxiliary proteins (CitC, CitX and CitG) into acetate and oxaloacetate (McAuliffe, 2017). Subsequently, oxaloacetate is decarboxylated to pyruvate by CitM, the oxaloacetate decarboxylase. Two molecules of pyruvate can be condensed into acetylacetate, which is converted by spontaneous oxidative decarboxylation to diacetyl or acetaldehyde. Genomic analysis of the *diacetylactis* biovar demonstrated that most lactococci, both domesticated and environmental, can produce diacetyl, acetoin, or both (Passerini et al., 2013). The presence of the *cit* operon is what delineates the *diacetylactis* biovar from the subspecies, and it is related to rapid accumulation of aroma. The *lactis* and *cremoris* species can also produce these compounds but through a much slower metabolic route via pyruvate. Indeed, it has been suggested that “citrate” might be a more appropriate name for the *diacetylactis* biovar, given its true citrate fermentation (Passerini et al., 2013). In addition, the authors noted that genomic features do not fully account for the flavor or aroma potential of a strain and that a combination of genomic and phenotypic analysis should be performed (Passerini et al., 2013). This has also been a lesson from other genomic investigations of flavor formation.

Proteolysis is also a key biochemical pathway in flavor formation by *Lc. lactis*. Through genome-wide comparisons, the diverse nature of the proteolytic system in *Lc. lactis* was investigated. Enzyme components of the proteolytic system; that is, proteinases, peptidases, and peptide transporters, were found to be unevenly distributed among lactococcal strains, typically correlating with the presence or absence of particular plasmids (Liu et al., 2010). The subspecies *cremoris* normally harbors the CEP PrtP/PrtM pair,
whereas they are generally absent from strains of the \textit{lactis} subspecies. There are exceptions, of course, with the dairy \textit{Lc. lactis} ssp. \textit{cremoris} HP lacking the gene pair, and the \textit{L. lactis} ssp. \textit{lactis} strains ML8, from a dairy source, and KF24, from a plant source, being PrtP/PrtM-positive. This is an interesting feature of KF24 and other non-dairy strains. The presence of large sets of proteolytic enzymes in non-dairy strains has prompted investigation of both genomic and flavor development potential of strains from diverse sources. Whole-genome sequencing of \textit{Lc. lactis} strains from grass, rumen, and vegetable sources by Cavanagh et al. (2015) revealed strains that could be genotypically classified as subspecies \textit{cremoris} but that behaved phenotypically like subspecies \textit{lactis}, with growth at 4°C and in 4% NaCl and with the ability to hydrolyze arginine. The WGS analysis also revealed sufficient phylogenetic distance between the strains to suggest that subspecies \textit{lactis} and \textit{cremoris} should be defined as separate species, rather than separate subspecies (Cavanagh et al., 2015). From a flavor perspective, the presence of certain key enzymes indicated the propensity of these strains to produce volatile flavor compounds. The non-dairy strains tested produced higher levels of branched-chain aldehydes and their corresponding alcohols, ethanol, acetoin, 2,3 butanediol, and diacetyl. As discussed previously, these latter compounds are generally produced in dairy strains from citrate fermentation and associated with desirable natural flavor in certain cheeses. This was application-tested in a mini-Gouda cheese model, where the non-dairy lactococcal strains were linked to nutty flavor notes and certain textural attributes, but with bitterness and astringency in the case of certain non-dairy strains (Cavanagh et al., 2014).

Based on available genome sequences, several GSMM have been generated for strains of \textit{Lc. lactis} (Oliveira et al., 2005; Flahaut et al., 2013), most notably the metabolism of nitrogen and its associated flavor-forming pathways (Flahaut et al., 2013). Interestingly, the GSMM predicted that \textit{Lc. lactis} is prototrophic for all AA except leucine, and that the broad substrate ranges of aminotransferases in \textit{Lc. lactis} allow the synthesis of more AA than previously predicted. Of the 754 reactions predicted, almost 60 reactions were identified to be directly or indirectly involved in flavor formation, including those involved in the production of VSC from methionine and cysteine. The model predicted the production of MTL and hydroxy acid from leucine and valine with an increase in their production coupled with an increase in growth rate. Thus, the model represents a valuable approach to prediction of flavor compound generation by providing an in silico mechanistic-based hypothesis to explain flavor compound production in relation to the growth of \textit{Lc. lactis} (Flahaut et al., 2013).

**Propionibacterium Species**

Certain species of \textit{Propionibacterium} play a particular role in the flavor development of Swiss-type cheeses. These species produce large amounts of propionate and acetate and include \textit{Propionibacterium freudenreichii} and \textit{Propionibacterium acidipropionici}. This so-called dairy group of \textit{Propionibacterium} are generally associated with dairy environments but also silage environments and they have a long history of safe use in food production (Rabah et al., 2017). From a flavor development perspective, \textit{P. freudenreichii} is the best studied as an integral component of the starter flora of Swiss-type cheeses. The characteristic flavor associated with Swiss cheese is a result of the conversion of L-lactate, produced by the starter LAB, into the flavor-forming propionate, acetate, and CO$_2$ that result in the characteristic eye formation. There is also a significant contribution by the products of milk fat hydrolysis to the characteristic flavor, such as free fatty acids released via lipolysis of milk glycerides (Abeijón Mukdsi et al., 2014).

The genomes of \textit{P. freudenreichii} CIRM-BIA1 and \textit{P. freudenreichii} ssp. \textit{shermanii} JS have been elucidated, and these sequences have been analyzed in comparative genomic studies to investigate the genomic basis of some of the commercially important features of these strains (Falentin et al., 2010; Ojala et al., 2017). What is interesting to note about this species, based on these comparative analyses, is the apparent lack of any specific adaptations of these organisms to the milk environment. Unlike many species of LAB, \textit{P. freudenreichii} has the capacity for de novo biosynthesis of most AA and vitamins. Although both strain CIRM-BIA1 and JS are able to ferment lactose due to the presence of a β-galactosidase gene most likely acquired by horizontal gene transfer, these organisms do not harbor any protease capable of casein hydrolysis. But the primary role of this species in dairy fermentations is to utilize the lactate produced by the starter LAB and convert it to flavor-forming compounds. Analysis of the genome sequences of the above-mentioned strains revealed pathways central to the formation of flavor compounds. In silico reconstruction of the metabolic pathways of \textit{P. freudenreichii} ssp. \textit{shermanii} strain JS revealed the genes encoding the 2 key metabolic cycles involved in propionate production; namely, the Wood-Werkman cycle and the citric acid or tricarboxylic acid cycle. \textit{Propionibacterium freudenreichii} CIRM-BIA1 appears to exclusively ferment L-lactate, as L-lactate permease
is present in the genome, whereas no D-lactate permease gene was identified. The sequenced CIRM-BIA1 strain also had numerous AA catabolism pathways, which can result in production of VFA. The organism has the potential to produce 2-methylbutanoic acid and 3-methylbutanoic acid (short branched-chain fatty acids) from the metabolism of isoleucine and leucine, respectively. These compounds are known to impart the "cheesy/sweaty" notes in certain cheese types. In addition, an extensive network of esterases was identified, potentially involved in the formation of free fatty acids and esters. Several secreted, surface-bound, and intracellular esterases were identified, predicted to be involved in the synthesis of the volatile esters associated with the fruity flavor of cheese (Falentin et al., 2010).

The 2.68-Mbp genome sequence of *P. freudenreichii* ssp. *shermanii* JS has shed further light on the capabilities of this species for flavor formation (Ojala et al., 2017). Similar to strain CIRM-BIA1, strain JS possesses all genes to convert lactate to propionate via the Wood-Werkman and tricarboxylic acid cycles, in addition to capacity to produce the flavor compound acetate. What is interesting is the apparent ability of strain JS to produce diacetyl from pyruvate via acetolactate by a pathway similar to Cit+ lactococci and *Leuconostoc* species. However, because no diacetyl reductase- or butanediol-encoding genes are present in the JS genome, it is unlikely that the subsequent conversion of diacetyl to acetoine occurs. It has been suggested for lactococci that rerouting of pyruvate to the diacetyl/acetoine pathway can be physiologically advantageous during lactose fermentation at low pH, as it may prevent accumulation of excess pyruvate inside the cell, resulting in intracellular acidification (Passerini et al., 2013; Zuljan et al., 2014). The authors of the study on *P. freudenreichii* ssp. *shermanii* JS propose a similar system operating on the co-metabolism of aspartate and lactate (Ojala et al., 2017).

Transcriptome profiling has been used to determine the activity of *P. freudenreichii* ssp. *shermanii* JS during warm-room (20°C) and cold-room (5°C) ripening of cheese (Ojala et al., 2017); RNA sequencing was performed on both warm- and cold-room samples and differentially expressed genes were identified. Both the central carbon and propionate production pathways of *P. freudenreichii* were more highly expressed during warm-room ripening than during cold-room ripening. This was also the case for genes encoding proteins involved in AA transport and metabolism. In contrast, enrichment analysis found that genes induced in cold-room conditions were primarily concerned with replication, recombination, and repair. These observations suggest that *P. freudenreichii* grows faster and is more active metabolically during warm-room ripening, where a strong increase in population numbers occurs. It is also evidence that the activity of the organism during warm-room ripening is the main driver of flavor development, although some flavor-contributing enzymes remain active during cold-room ripening, such as branched-chain amino acid aminotransferases and certain esterases (Ojala et al., 2017).

**Streptococcus thermophilus**

In addition to using WGS as a means to uncovering new pathways for flavor generation, the technology can also be used as a screening tool to monitor changes in known pathways in the generation of spontaneous mutants with enhanced flavor capabilities. This approach was taken by Sørensen et al. (2016) in an elegant study where WGS analysis of spontaneous mutants derived from strains of *S. thermophilus* and *Lb. delbrueckii* ssp. *bulgaricus* led to the identification of food-grade mutants that consume lactose and galactose, which are considered nonsweet sugars, and secrete glucose, considered much sweeter. This resulted in yogurt with enhanced sweetness without the need for added sucrose and additional calorie content (Sørensen et al., 2016). Typically, strains of *S. thermophilus* and *Lb. delbrueckii* ssp. *bulgaricus* are not able to ferment the galactose produced from the breakdown of lactose into glucose and galactose (de Vin et al., 2005). It is secreted into the milk; thus, dairy products produced using these cultures contain significant amounts of galactose as well as high levels (30 to 40 mg/mL) of residual unfermented lactose (Sørensen et al., 2016). However, because glucose is perceived as being far sweeter than galactose or lactose, selection of mutants secreting glucose into the growth medium would be advantageous. Galactose-fermenting mutants of *S. thermophilus* were obtained through selection on M17 galactose plates. Subsequent WGS analysis revealed mutations in genes encoding elements of the glucose/mannose phosphotransferase system (PTS), the *galKTEM* promoter and the glucokinase gene. These strains were capable of metabolizing galactose but unable to phosphorylate glucose, either internally or via the PTS. In a similar food-grade manner, *Lb. delbrueckii* ssp. *bulgaricus* mutants were generated, disrupting the glucose/mannose phosphotransferase system and creating strains that were unable to take up glucose from the milk. These strains cannot grow on exogenous glucose but can metabolize internal glucose released from lactose by β-galactosidase. A combination of these mutant strains produces yogurt with no detectable lactose, moderate levels of galactose, and high levels of glucose, thus perceptibly enhancing the sweetness of yogurt without the need for added
sucrose and additional calorie content (Sorensen et al., 2016). Although the acidification rates of both strains were affected, it was found that pretreatment of cells before inoculation could improve these rates. However, whether this is economically feasible remains to be demonstrated. Of note, the metabolic rerouting of the strains in this study did not involve the use of non-food-grade DNA methods and provided evidence that classical methods of strain manipulation can result in food-grade strains with valuable characteristics. In addition, the use of WGS to screen for other mutations that may alter the behavior of the mutants and prohibit their use in foods is an approach that can be adopted to any screening project.

**Fungal Species**

Flavor development in surface-ripened cheeses, such as Camembert, and blue cheeses, such as Roquefort, is driven mainly by the fungal microflora of the product. Species of critical importance in the ripening of these cheeses include the fungi *Penicillium camemberti*, *Geotrichum candidum*, and *Penicillium roqueforti* (Irlinger et al., 2015). Via their proteolytic and, in particular, lipolytic enzyme activities, these organisms are major contributors to the organoleptic quality and typical characteristics of these mold-ripened cheese types. Although limited WGS information is available for these fungal species, a recent metatranscriptome study investigated the genetic determinants for these specific enzyme activities in *P. camemberti* and *G. candidum* and their expression over the 77-d ripening process of Camembert cheese (Lessard et al., 2014). As with other cheese varieties, proteolysis is a major contributor to cheese flavor in Camembert (Lessard et al., 2012). Metallopeptidases and serine peptidases were the most abundant peptidase families expressed, with protease and peptidase transcripts detected early in the ripening period, mainly in the first 21 d (Lessard et al., 2014). Ripening yeasts such as *Geotrichum* species are known for their VSC production, and indeed, the *cgl* gene, encoding cystathionine γ-lyase (key in the generation of VSC from methionine), was identified in the top 5% of expressed contigs in *G. candidum* (Lessard et al., 2014). In *G. candidum*, expression of *cgl* is linked to cabbage and sulfur aroma development through MTL production (Martinez-Cuesta et al., 2013).

Because lipid metabolism plays such a major role in the aroma and texture of Camembert-type cheeses, generating flavor compounds such as methylketones, esters, alcohols, and lactones, the pathways of lipolysis were also examined. It was found that genes of the lipolytic pathway (e.g., genes for lipases, which hydrolyze triglycerides into di- and monoglycerides, free fatty acids, and glycerol) were expressed at the beginning of the ripening period (observed at d 5) and continued to be expressed throughout the ripening period, albeit at a very low rate (Lessard et al., 2014). In addition, enzymes of the peroxisomal and mitochondrial pathways of β-oxidation were found to be expressed in *P. camemberti*, whereas only enzymes of the peroxisomal β-oxidation pathway were present in *G. candidum*. β-Oxidation of fatty acids produces ketone bodies, which are aroma precursors (Urbach, 1997). The peroxisomal multifunctional enzyme (MFE), which is also responsible for the β-oxidation of fatty acids in fungi, was expressed in both species, with 69% of all transcripts related to the β-oxidation cycle coding for the MFE (Lessard et al., 2014). This suggests that MFE could have a central biological role and act as a potential biomarker for Camembert-type cheese ripening, given its expression over time in both microorganisms. Overall, the study concluded that *P. camemberti* has a higher lipolytic potential than *G. candidum* and therefore is more likely to have a greater influence on the production of volatile flavor compounds in Camembert cheese (Lessard et al., 2014).

**SYSTEM-LEVEL TECHNOLOGIES TO STUDY FLAVOR FORMATION**

To date, much of the study of flavor in dairy products has dealt with single strains. However, with few exceptions, food fermentations rely on mixed communities of microbes, interacting via various mechanisms (Sieuwerts et al., 2008). The multi-organism contribution of a complex dairy microbiota can have a positive effect on the stability, reproducibility, and productivity of the product flora, with all elements of that flora contributing to product functionality. Advances in next-generation sequencing and the availability of advanced bioinformatics tools has enabled high-level analysis of bacterial communities in dairy products and the active biochemical pathways during the fermentation process. Recent studies integrating metabolomic and metagenomic analyses have begun to address the relationship between the microbiota and flavor development. A metagenomic study that aimed to correlate genomic information and flavor compounds in kefir revealed that *Lb. kefiranofaciens* was the most prevalent in the bacterial community in early ripening, whereas *Leu. mesenteroides* dominated later (Walsh et al., 2016). These changes corresponded to variations in the levels of volatile compounds throughout ripening. For instance, the presence of *Lb. kefiranofaciens* was correlated with the production of carboxylic acids, ketones, and esters, which contribute to cheesy and fruity flavors, and *Leu. mesenteroides* was linked with an
abundance of 2,3-butanedione (buttery notes) and acetic acid (vapor flavor). These data were also in agreement with sensory analysis. A study on smear-ripened cheeses sought to investigate the succession of microbial populations on the rind through metagenome shotgun sequencing and to correlate this with volatile analysis (Bertuzzi et al., 2018). Yeasts such as Debaryomyces hansenii and G. candidum were found to dominate in the early stages of ripening, with bacterial species such as Brevibacterium linens and Glutamicibacter arilaitensis coming to the fore later in the ripening process. The authors correlated the presence of individual species to volatile flavor compounds detected by GC-MS at the cheese surface. Debaryomyces hansenii was linked with the production of alcohols and carboxylic acids, B. linens and G. candidum with sulfur compounds, and Glutamicibacter arilaitensis with alcohols, carboxylic acids, and ketones. The association of the production of volatile flavor compounds and the species present in the microbiota allows researchers to investigate the metabolic potential of the microbial community and can demonstrate how variations in the microbial populations influence important aspects of cheese flavor development. Such integrated approaches represent a novel means for studying flavor development in cheese and other dairy products.

SUMMARY

With the advances in genomic technologies made in the past decades, the potential of dairy microbes for flavor generation through fermentation has been investigated using various complementary approaches, including the in silico mining of genomic information and the assessment of culture performance at the transcriptome and proteome levels. Genomic investigations have shed further light on the metabolic potential of these organisms, allowing for a more knowledge-based approach for the selection of desirable cultures for flavor generation. In addition, the generation of whole-genome sequences and their analysis has led to exploitation of the metabolic diversity of organisms that might not normally be associated with dairy products. The power afforded by WGS allows evaluation of flavor production potential in tandem with assessment of the suitability of the organism for use in food; for example, the absence of antibiotic resistance genes, biogenic amine production capacity, and so on, which opens a range of possibilities in terms of culture selection for diversification of product flavor. However, caution is advised, because many reports conclude that revelations from genomic studies cannot fully account for the flavor or aroma potential of a strain, and that integrated methods combining genomic- and metabolomic-based approaches are required.

The beginnings of the application of systems-level analysis to tackle the significant challenge of linking knowledge gained from studies on pure cultures to studies of the microbial community within a product is a welcome development. Investigation of the synergistic interactions that occur between key organisms in mixed-culture fermentations is an understudied area of research. To date, much of the work on the generation of flavor through microbial fermentation has focused on the activities of pure cultures, which have been well characterized and exploited. Mixed-culture fermentations present an interesting case, because the flavor compounds produced are strongly influenced by the microbial activity of each active strain during and after fermentation and directly relate to the strains involved and their dynamics within the population. As such, small changes in the population dynamics can have a dramatic effect on metabolism and, in turn, flavor production. Recently, integrated approaches have examined flavor production in a variety of mixed-culture systems or microbial communities, including the combination of metagenomics and flavor analyses in the study of a complex multispecies solid-state fermentation as occurs in acetic acid fermentation (Wang et al., 2016), and the successful integration of metabolomic and transcriptomic data for the analysis of flavor profiles generated by multiple yeasts during wine fermentation (Mendes et al., 2017), as well as the dairy-based systems discussed previously (Bertuzzi et al., 2018).

Significant advances in metabolomic methods are also enhancing our understanding of flavor development by dairy microbes. Most of the flavor metabolites in dairy products are aromatic, and thus volatile, and therefore the optimal approach is GC-MS. Key developments relate to improved automated volatile extraction and concentration techniques, comprehensive or 2-dimensional GC, and improved sensitivity of quadrupole, time-of-flight and orbitrap mass spectrometers. However, significant advances in data processing to aid identification and quantification using deconvolution software are aiding rapid metabolite identification in complex chromatograms. Advanced GC-MS coupled with next-generation genomic technologies is enabling correlations to be made between microbial species and volatile compounds using multivariate statistical analysis. Mass spectrometric profiling is also enabling the study of the “sensometabolome,” or the population of the key flavor compounds in our food. This type of approach can be used to define a “molecular blueprint” of the chemosensory properties of a food and allow monitoring of any changes in odor and flavor profiles that
occur along the processing chain, from the raw materials through manufacturing process intermediates, and all the way to the consumer. From the perspective of fermented dairy foods, knowledge-based selection of strains would be a key step in such a “molecular food engineering” approach—the tailoring of processing parameters to better flavor improvement.

ACKNOWLEDGMENTS

The authors acknowledge our many Teagasc colleagues who have contributed to the Cultures, Fermentation and Biotransformation and Flavor Chemistry research programmes at Teagasc, particularly Daniel Cavanagh, Clara Roces Rodriguez, Shahnelle Mazhar, Mary Rea, Diarmuid Sheehan, and David Mannion, in addition to our valued collaborators Anne Thierry (INRA, France), Eric Altermann (AgResearch, New Zealand), and Maurice O’Sullivan (University College Cork, Ireland). We also acknowledge the invaluable assistance of Lidia Anna Soszynska (Belgrade, Serbia) in the generation of the figures in this manuscript. Our research team has been funded by Teagasc and the Teagasc Walsh Fellowship programme (Carlow, Ireland), Dairy Research Ireland (Cork, Ireland), IRCSET (Dublin, Ireland), and the EU Marie Curie Actions Clarin Co-Fund (Asturias, Spain).

REFERENCES


