

Application of molecular biological methods for studying probiotics and the gut flora

A. L. McCartney*

Food Microbial Sciences Unit, School of Food Biosciences, University of Reading, Whiteknights, Reading RG6 6AP, UK

Increasingly, the microbiological scientific community is relying on molecular biology to define the complexity of the gut flora and to distinguish one organism from the next. This is particularly pertinent in the field of probiotics, and probiotic therapy, where identifying probiotics from the commensal flora is often warranted. Current techniques, including genetic fingerprinting, gene sequencing, oligonucleotide probes and specific primer selection, discriminate closely related bacteria with varying degrees of success. Additional molecular methods being employed to determine the constituents of complex microbiota in this area of research are community analysis, denaturing gradient gel electrophoresis (DGGE)/temperature gradient gel electrophoresis (TGGE), fluorescent *in situ* hybridisation (FISH) and probe grids. Certain approaches enable specific aetiological agents to be monitored, whereas others allow the effects of dietary intervention on bacterial populations to be studied. Other approaches demonstrate diversity, but may not always enable quantification of the population. At the heart of current molecular methods is sequence information gathered from culturable organisms. However, the diversity and novelty identified when applying these methods to the gut microflora demonstrates how little is known about this ecosystem. Of greater concern is the inherent bias associated with some molecular methods. As we understand more of the complexity and dynamics of this diverse microbiota we will be in a position to develop more robust molecular-based technologies to examine it. In addition to identification of the microbiota and discrimination of probiotic strains from commensal organisms, the future of molecular biology in the field of probiotics and the gut flora will, no doubt, stretch to investigations of functionality and activity of the microflora, and/or specific fractions. The quest will be to demonstrate the roles of probiotic strains *in vivo* and not simply their presence or absence.

Molecular methods: Probiotics: Gut flora

Introduction

In recent years it has become increasingly evident that the food industry and gastrointestinal microbiologists require sensitive and reliable methods to identify and characterise the microbial content of foods and the host's gastrointestinal (GI) tract. Particular interest is afforded to the lactic acid bacteria (LAB; especially *Lactobacillus* spp. and *Bifidobacterium* spp.) due largely to: (i) the association of these organisms with health-promoting properties; (ii) their inclusion in numerous food products as 'probiotics'; and (iii) the requirements of legislative and

industrial bodies, as well as the consumer, with respect to safety, labelling and strain integrity (Charteris *et al.* 1997; Prasad *et al.* 1998; Hozapfel *et al.* 2001). Additional areas of interest revolve around contamination, food-borne pathogens, and any underlying microbiological basis to GI disorders, or susceptibility to such illnesses. This review will concentrate on the application of molecular methods in probiotic research, their advantages and limitations, including the identification and characterisation of LAB, techniques available to differentiate probiotic strains from one another and from the indigenous LAB population, and the potential of

Abbreviations: AFLP, amplified fragment length polymorphism; ARDRA, amplified ribosomal DNA restriction analysis; DGGE, denaturing gradient gel electrophoresis; FAME, fatty acid methyl ester; FISH, fluorescent *in situ* hybridisation; F6PPK, fructose 6-phosphate phosphoketolase; GI, gastrointestinal; ITS, intergenic spacer region; LAB, lactic acid bacteria; MLEE, multilocus enzyme electrophoresis; MLST, multilocus sequence typing; PCR, polymerase chain reaction; PFGE, pulsed-field gel electrophoresis; RAPD, randomly amplified polymorphic DNA; REA, restriction enzyme analysis; (T)RFLP, (terminal) restriction fragment length polymorphism; TAP, triplet arbitrary primed PCR; TGGE, temperature gradient gel electrophoresis.

* **Corresponding author:** Dr A. L. McCartney, fax +44 118 9357222, email a.l.mccartney@reading.ac.uk

molecular biological advances to elucidate functionality and activity.

Classical identification methods rely heavily on phenotypic characterisation, including morphology (both cellular and clonal), growth requirements and characteristics, fermentation profiles, cell-wall protein analysis, serology and, more recently, fatty acid methyl ester (FAME) analysis (Klein *et al.* 1998; Giraffa & Neviani, 2000). Although the application of some of these techniques has proven useful for certain LAB (e.g. cell-wall protein profiling for thermophilic lactobacilli; cell-wall peptidoglycan analysis for identifying *Bifidobacterium longum*, *Bifidobacterium infantis* and *Bifidobacterium suis* (Hozapfel *et al.* 2001); and whole-cell protein profiling for certain lactobacilli at both the species and subspecies levels (Giraffa & Neviani, 2000)), there is a general awareness that some aspects of phenotypic characterisation are principally flawed, i.e. the observation of a similar phenotype does not always equate to a similar, or closely-related, genotype. There has therefore been a shift towards molecular-based investigations of phenotypic characteristics and/or employment of genotypic characterisation in order to provide more robust classification and differentiation (Charteris *et al.* 1997; O'Sullivan, 1999; Giraffa & Neviani, 2000; Hozapfel *et al.* 2001). Additional weaknesses of phenotypic methods include poor reproducibility, ambiguity of some techniques (largely resulting from the plasticity of bacterial growth), extensive logistics for large-scale investigations and poor discriminatory power. However, genetic characterisation techniques are not without limitations and thus a polyphasic, or combined approach is often necessary.

Differential plating methodologies are very useful for isolating and enumerating probiotic organisms, especially from a mixed environment (Charteris *et al.* 1997). However, such techniques provide a limited view of the diversity and dynamics of the GI microbiota, as they rely on the cultivation of organisms. At best, such investigations provide an insight into the predominant culturable population, but can not fully elucidate the microbiological community. Estimates range from as little as 10% up to 40–50% of the GI population being accounted for by such methodologies (Zoetendal *et al.* 1998). Recent genetic-based investigations confirm the enormity of the 'previously unaccounted for' fraction of this extensive and diverse microbiota (the so-called 'unculturables'; Zoetendal *et al.* 1998; Suau *et al.* 1999).

Molecular biological applications in probiotics and indigenous lactic acid bacteria

The basic taxonomic unit is the species; classically, species delineation is based on $\geq 70\%$ homology by DNA–DNA hybridisation (Wayne *et al.* 1987). With the advent of 16S rRNA gene sequencing it is now generally accepted that species delineation may be based on 3% sequence divergence between a novel species and its closest relative provided corresponding phenotypic data allows differentiation from already recognised species of the genera in question. Where it is clear that an isolate represents a new species, but shows less than 3% divergence with its nearest relative, DNA–DNA hybridisation data is required

along with phenotypic data. There are currently no guidelines with respect to subspecies delineation, only that there are recognisable phenotypic differences. Various genetic fingerprinting techniques, as well as genetic marking, have proven useful in subspecies discrimination or strain differentiation. Alternatively, monoclonal antibody assays or antibiotic resistance markers have been employed for identifying or tracking specific strains in mixed samples.

To date, molecular techniques have been employed in studies of probiotics and human GI microflora for four main aims: (i) characterisation of bacterial diversity within samples; (ii) enumeration of phylogenetically related groups of bacteria; (iii) tracking or monitoring of specific organisms or populations, both quantitatively and qualitatively; and (iv) definitive identification of isolates, especially probiotics. Probiotics are 'live microbial supplements' that are culturable and therefore the full range of classification protocols is suitable, including the genetic techniques discussed below.

Genetic probing strategies

Probing techniques are based on the hybridisation of synthetically prepared oligonucleotides to specific target sequences on bacterial DNA. The specificity of the probe is largely dependent on the target sequence, although the stringency of the hybridisation conditions and washings are also critical (Charteris *et al.* 1997; O'Sullivan, 1999). To date, probing methods have concentrated on different regions of the bacterial ribosome. Highly conserved regions are targeted for universal, or domain-based, oligonucleotide probes, whereas different variable regions are used for genus-specific probes; highly variable regions are targeted for species-specific probes. Though ribosomal oligonucleotides are frequently used (predominantly 16S rDNA probes), due to the extensive databases of known sequences, other genetic targets may be useful in some instances. For example, the fructose 6-phosphate phosphoketolase enzyme (F6PPK; particular to *Bifidobacterium* and *Gardnerella* species) gene–gene complex may contain regions that provide greater inter- and intra-species discriminatory power for bifidobacteria (O'Sullivan, 1999). Alternatively, random selection of DNA fragments from restriction digests of plasmid, or total DNA, has previously been used to isolate specific probes for some species (including *Lactobacillus* species; Giraffa & Neviani, 2000). A vast list of oligonucleotide probes is now available, including ones specific to most food-related bacteria, LAB (see Table 1) and the predominant members of the indigenous microflora of the human GI tract. The major limitation of genetic probes is the level of specificity. To date, species-specificity is the optimal discriminatory power of such probes. As the genetic information available in databases continues to grow, and with the increasing use of whole-genome sequencing (as opposed to gene sequencing) of organisms, subspecies target sequences suitable for probing strategies may be identified. Such advances in molecular biology may provide future approaches with both increased specificity for genetic probing and probing assays that demonstrate activity or function.

Labelled oligonucleotide probes are employed in a

Table 1. Summary of oligonucleotide probes available in the literature for *Lactobacillus* and *Bifidobacterium* species

Bacterial species/group	Probe types
<i>Bifidobacterium adolescentis</i>	16S rDNA, DNA fragment*
<i>Bifidobacterium adolescentis/coryneforme</i>	16S rDNA
<i>Bifidobacterium animalis</i>	DNA fragment*
<i>Bifidobacterium breve</i>	16S rDNA
<i>Bifidobacterium bifidum</i>	16S rDNA, DNA fragment*
<i>Bifidobacterium infantis</i>	16S rDNA, DNA fragment*
<i>Bifidobacterium infantis/longum/suis</i>	16S rDNA
<i>Bifidobacterium longum</i>	16S rDNA, DNA fragment*
<i>Bifidobacterium pseudocatenulatum</i>	16S rDNA
<i>Lactobacillus acidophilus</i>	16S rDNA, 23S rDNA, rRNA
<i>Lactobacillus brevis</i>	Genomic DNA
<i>Lactobacillus casei</i>	Genomic DNA
<i>Lactobacillus casei/rhamnosus</i>	23S rDNA
<i>Lactobacillus casei/paracasei/rhamnosus</i>	23S rDNA
<i>Lactobacillus crispatus</i>	23S rDNA
<i>Lactobacillus delbrueckii</i>	DNA fragment*
<i>Lactobacillus fermentum</i>	16S rDNA
<i>Lactobacillus gasseri</i>	23S rDNA
<i>Lactobacillus helveticus</i>	DNA fragment*
<i>Lactobacillus johnsonii</i>	23S rDNA
<i>Lactobacillus paracasei</i>	23S rDNA
<i>Lactobacillus plantarum</i>	16S rDNA, plasmid DNA fragment, genomic DNA
<i>Lactobacillus plantarum/pentosus</i>	23S rDNA

* Identified from randomly cloned DNA fragments from this species in *Escherichia coli*.

number of assays including colony, dot-blot (including probe grids) and *in situ* hybridisations (Charteris *et al.* 1997; O'Sullivan, 1999; Giraffa & Neviani, 2000). Improved technologies for fluorescent labelling, and other non-radioactive labelling techniques, have reduced the need to employ radioactive labels and also increased the application of genetic probes in gut microbiology, seen with the recent abundance of fluorescent *in situ* hybridisation (FISH) studies.

As the name suggests, colony hybridisation involves probing of bacterial colonies that have been transferred onto membranes (e.g. nitrocellulose membranes). Such methods were formerly used for rapid identification of bacterial colonies and enumeration of specific bacterial groups or species prior to *in situ* assays that do not require cultivation. Additionally, the automation of gene sequencing has negated the use of colony hybridisation, as improved discriminatory power is available for identifying and differentiating bacterial isolates using this approach.

Dot-blot hybridisation techniques involve probing DNA extracts, from either bacterial isolates or environmental samples, and enable multiple probes and multiple samples to be analysed concurrently, provided that the different probes have a shared hybridisation stringency. Indeed probe grids, or 'checkerboard' hybridisations, employ such a strategy, with the DNA extracts loaded along one axis and labelled probes sequentially loaded along the other axis, resulting in a composition profile for each sample (O'Sullivan, 1999). Such an approach is also feasible using well-assays, where the labelled probes are bound to the wells and the samples are applied. The ability to extract total DNA or RNA from environmental samples, or mixed populations, means that such techniques are not reliant on cultivation.

FISH enables the direct enumeration of whole bacterial

cells in environmental samples, or other mixed populations, using either fluorescent microscopy or flow cytometry (Langendijk *et al.* 1995). An abundance of such studies has been seen in gut microbiology in recent years, largely concentrating on the predominant bacterial groups or genera (Franks *et al.* 1998). Although such methods would also be applicable to numerically dominant bacterial species, the detection threshold of this technique limits its use for examining subdominant populations.

Genetic fingerprinting

Genetic fingerprinting is very useful for differentiating bacterial isolates and has been applied to tracking bacteria of interest and elucidating the complexity, dynamics and diversity of bacterial populations, including the GI microflora. Subspecies discrimination is achieved in most cases, although there are also examples of fingerprinting strategies that afford speciation. The primary limitation of most fingerprinting methods is the need to first isolate the organisms. However, recent developments in genetic fingerprinting combine polymerase chain reaction (PCR) amplification with profiling strategies, sometimes termed PCR-typing, thus overcoming the reliance on cultivation.

The majority of genetic fingerprinting techniques are based on restriction fragment length polymorphism (RFLP) analysis, essentially a profiling tool based on the banding patterns obtained from DNA restriction digests (Charteris *et al.* 1997; Hozapfel *et al.* 2001). The crucial element of this technique is the selection of the restriction enzymes. Use of rare-cutting enzymes reduces the number of DNA fragments, but usually requires more sophisticated techniques, such as pulsed-field gel electrophoresis (PFGE), to separate the large DNA fragments generated. Along with PFGE, a number of modifications to restriction

enzyme analysis (REA) are available to simplify the resultant banding profiles. Examples include ribotyping and amplified ribosomal DNA restriction analysis (ARDRA; Roy *et al.* 2001).

PFGE employs an alternating field of electrophoresis to allow separation of the large DNA fragments obtained from restriction digests with rare-cutting enzymes, with increasing pulse times throughout the run (O'Sullivan, 1999; Hozapfel *et al.* 2001). Crucial to PFGE is the extraction of intact chromosomal DNA, which is usually achieved by embedding the bacteria in low-melting-point agarose prior to cell lysis. As such, the technique can be more time consuming than other fingerprinting strategies. However, since the PFGE profile generated represents the whole genome this technique has superior discriminatory power. Indeed, excellent subspecies differentiation has been shown using PFGE for a number of organisms, including lactobacilli and bifidobacteria (McCartney *et al.* 1996; Kimura *et al.* 1997; O'Sullivan & Kullen, 1998). In some cases PFGE has enabled the grouping of bacterial strains within a species, and there are also examples of the potential of this technique to speciate bacterial isolates. Genome integrity may also be checked using PFGE, especially when multiple enzyme profiles are used, as DNA insertions, deletions and rearrangements alter the fingerprint (O'Sullivan, 1999).

A number of studies have demonstrated the usefulness of PFGE to monitor changes in the predominant bifidobacterial and lactobacilli populations of humans, both within individuals over time and between individuals (McCartney *et al.* 1996; Kimura *et al.* 1997). Other workers have shown the ability of PFGE to differentiate probiotic strains. Strain typing has been successfully achieved by PFGE for the *Lactobacillus acidophilus* complex, *Lactobacillus casei*, *Lactobacillus delbrueckii* and its three subspecies (*bulgaricus*, *delbrueckii* and *lactis*), *Lactobacillus fermentum*, *Lactobacillus helveticus*, *Lactobacillus plantarum*, *Lactobacillus rhamnosus* and *Lactobacillus sakei* (Klein *et al.* 1998; Giraffa & Neviani, 2000). Overall, PFGE has been shown to differentiate strains belonging to the same LAB species, group strains within a species, distinguish between strains of different LAB species, and even to place isolates in specific *Lactobacillus* species. In a recent comparison, PFGE was shown to be more discriminatory in typing closely related *Lactobacillus casei* and *Lactobacillus rhamnosus* strains than either ribotyping or randomly amplified polymorphic DNA (RAPD) analysis (Tynkkynen *et al.* 1999). However, it should be noted that in this case PFGE differentiation was performed combining the results from two separate enzymes (*Sfi*I and *Not*I), whereas a single restriction enzyme was employed in ribotyping (*Eco*RI). Use of multiple ribopatterns to determine the overall ribotype of isolates has previously been shown to increase the discriminatory power of this technique (McCartney *et al.* 1996). Additionally, *Eco*RI ribotyping was able to differentiate the *Lactobacillus rhamnosus* strains VS 1020 and VS 1021, which neither PFGE nor RAPD analyses separated, thus demonstrating the improved discriminatory power of the technique in some instances. PFGE is generally considered to afford the greatest differentiation of conventional fingerprinting techniques.

Ribotyping is a modification of conventional RFLP, whereby the REA is simplified by highlighting those DNA fragments that contain rDNA through Southern blotting and hybridisation with an rDNA probe. Ribotyping affords good subspecies differentiation and has been successfully used for both bifidobacteria and lactobacilli, and for studying the diversity of strains within these two groups (McCartney *et al.* 1996). Greater discrimination is afforded when multiple ribopatterns are used to define each ribotype, especially when restriction enzymes with distinctly different recognition sequences are used. Amplified rDNA restriction analysis (ARDRA) is essentially the reverse of ribotyping, i.e. the REA of 16S rDNA PCR amplicons. Ribotyping generally affords greater discriminatory power than ARDRA, due to the inclusion of the flanking regions of the 16S rRNA genes in the fingerprint. Overall, the conservative nature of 16S rRNA genes limits the discriminatory power of these techniques compared to methods which utilise the whole genome, such as PFGE and RAPD (O'Sullivan, 1999). Nevertheless, ribotyping has successfully differentiated various species or strains within the *Lactobacillus acidophilus* complex, *Lactobacillus casei*, *Lactobacillus delbrueckii*, *Lactobacillus fermentum*, *Lactobacillus helveticus*, *Lactobacillus plantarum*, *Lactobacillus reuteri*, *Lactobacillus rhamnosus* and *Lactobacillus sakei* (Giraffa & Neviani, 2000; Hozapfel *et al.* 2001). Other studies have demonstrated species recognition of *Lactobacillus plantarum* isolates using ribotyping, while further investigations identified species-specific ribotypes for *Lactobacillus crispatus* and *Lactobacillus gasseri* (Charteris *et al.* 1997). On the other hand, ARDRA has been used to differentiate a variety of lactobacilli at species level, including *Lactobacillus delbrueckii* and its three subspecies (*bulgaricus*, *delbrueckii* and *lactis*), *Lactobacillus acidophilus* and *Lactobacillus helveticus* (Roy *et al.* 2001).

Amplified fragment length polymorphism (AFLP) analysis is based on the selective amplification of restriction fragments from total digests of genomic DNA; the DNA fragments are then separated by polyacrylamide gel electrophoresis (Vos *et al.* 1995). Originally developed for plant systematics, AFLP has been found to be a very useful fingerprinting technique for bacteria, affording both species resolution and strain differentiation. To date, AFLP has mostly been employed in epidemiological studies and in investigations aiming to distinguish virulence markers in food-borne pathogens (such as *Listeria* and *Salmonella*). However, species-level discrimination has also been shown for the phylogenetically closely related species *Lactobacillus pentosus*, *Lactobacillus plantarum* and *Lactobacillus pseudopantarum* using this method (Giraffa & Neviani, 2000).

Polymerase chain reaction-based methodologies

PCR-based techniques, including multiplex PCR using specific primers and RAPD, have been used to detect probiotic lactobacilli in faecal samples. The list of specific PCR primers continues to grow, with species-specific primer combinations and even strain-specific PCR primer sets being developed (Charteris *et al.* 1997; Giraffa &

Neviani, 2000). Such primers are useful in direct PCR analysis of samples, demonstrating the presence and semi-quantification of bacterial groups and/or species and identification work. More recently, genus- and group-specific primer sets have been utilised in studies concerned with the composition and/or diversity within specific bacterial groups, including PCR-typing and community analysis studies. Overall, the development of PCR-based methodologies has provided simple and rapid tools for identifying bacteria and investigating the diversity of mixed populations without the need for cultivation (Wang *et al.* 1996). Additionally, such advances in molecular methodology provide great potential for rapid and reliable quality control of probiotic products.

The major limitation of PCR-based methodologies is the inherent biasing of the technique (Wintzingerode *et al.* 1997). Every step (biological, chemical or physical) in the genetic analysis of environmental samples may be a source of bias, potentially distorting the picture obtained. The greatest hazards in studying the composition of mixed populations include insufficient or preferential cell lysis, PCR inhibition, differential amplification and formation of chimeric or artefactual PCR products. A further complication in PCR-based diversity studies encompasses differences in the *rrn* operons of different organisms, both with respect to copy number and heterogeneity (Wintzingerode *et al.* 1997). A number of strategies are available to reduce and/or alleviate such biasing, e.g. incorporating PCR facilitators in the reaction, performing multiple reactions with different 'universal' primer sets, or controlling the PCR conditions, such as number of cycles and elongation time. Overall, PCR-based techniques provide a powerful tool for examining the microbial diversity of mixed populations. The general recommendation for 'a good PCR-mediated analysis of 16S rRNA from environmental samples' is comparative studies incorporating different DNA extraction, PCR amplification and cloning strategies, with replication and controlled PCR conditions (Wintzingerode *et al.* 1997). Although each step in the procedure from sample collection to analysis may cause biasing, the PCR amplification and cloning stages are the most likely source. Wintzingerode *et al.* (1997) suggest performing multiple PCR amplifications and cloning experiments on a mixture of nucleic acids obtained from a number of extraction methods, thus obtaining a more accurate picture of the diversity of the mixed population of each sample. The advent of capillary sequencing facilities and robotic automation may indeed allow such analyses to be carried out. However, the costs would still be prohibitive, especially as a tool for investigating the diversity of the gut microflora and the effects of dietary modification or probiotic administration on such an ecosystem. It would be more plausible to produce composite profiles for each sample from PCR-typing following multiple amplifications with a mixture of nucleic acids obtained from a number of extraction procedures.

RAPD analysis utilises short arbitrary primers and low-stringency to randomly amplify DNA fragments, which are separated to give a fingerprint (Tilsala-Timisjarvi & Alatosava, 1998). Though a simple and rapid method, RAPD is prone to poor reproducibility, largely because

small changes in reaction conditions can alter the profiles. RAPD must therefore be performed under controlled conditions. Alternatively, refinement of the protocol to perform multiple RAPD PCR can improve reproducibility of the technique, e.g. triplet arbitrary primed PCR (TAP-PCR) or tri-RAPD, which involves aliquoting the PCR reaction mix into three tubes and performing amplification at three annealing temperatures (each differing by 2°C) (O'Sullivan & Kullen, 1998; O'Sullivan, 1999). The RAPD profile is then identified as the profile consisting of the shared bands across the three patterns. Such a method was tested by O'Sullivan & Kullen (1998) for fingerprinting bifidobacteria and found to be 'highly strain-specific'.

Numerous research groups have demonstrated the reliability of RAPD-typing for inter- and intra-specific differentiation of most *Lactobacillus* species. Indeed, it is considered, along with 16S rRNA gene sequencing, to be the most widely used technique for characterising food-associated lactobacilli, and has been shown to accurately distinguish between strains of *Lactobacillus pentosus* and *Lactobacillus plantarum* (Charteris *et al.* 1997). Additionally, multiplex RAPD (RAPD with multiple arbitrary primers) has been employed successfully for the differentiation of LAB isolated from the GI tract (Lucchinin *et al.* 1998). More recently, reports include the use of RAPD to

Table 2. Summary of *Lactobacillus* and *Bifidobacterium* species and strains for which specific polymerase chain reaction (PCR) primers are currently available

Species-specific PCR primers
<i>Lactobacillus acidophilus</i>
<i>Lactobacillus casei</i>
<i>Lactobacillus crispatus</i>
<i>Lactobacillus curvatus</i>
<i>Lactobacillus delbrueckii</i> subsp. <i>bulgaricus</i>
<i>Lactobacillus delbrueckii</i> subsp. <i>lactis</i>
<i>Lactobacillus fermentum</i>
<i>Lactobacillus gasseri</i>
<i>Lactobacillus helveticus</i>
<i>Lactobacillus johnsonii</i>
<i>Lactobacillus paracasei</i> subsp. <i>paracasei</i>
<i>Lactobacillus paraplantarum</i>
<i>Lactobacillus pentosus</i>
<i>Lactobacillus plantarum</i>
<i>Lactobacillus reuteri</i>
<i>Lactobacillus rhamnosus</i>
<i>Lactobacillus sanfranciscensis</i>
<i>Lactobacillus sharpeae</i>
<i>Lactobacillus zeae</i>
<i>Bifidobacterium adolescentis</i>
<i>Bifidobacterium angulatum</i>
<i>Bifidobacterium bifidum</i>
<i>Bifidobacterium breve</i>
<i>Bifidobacterium catenulatum</i> group
<i>Bifidobacterium dentium</i>
<i>Bifidobacterium gallicum</i>
<i>Bifidobacterium infantis</i>
<i>Bifidobacterium longum</i>
Strain-specific PCR primers
<i>Bifidobacterium breve</i> Y8
<i>Bifidobacterium infantis</i> Y1
<i>Bifidobacterium longum</i> Y10
<i>Lactobacillus gasseri</i> 4B2

identify species- and strain-specific probes or PCR primers for some lactobacilli including *Lactobacillus gasseri* and *Lactobacillus rhamnosus* (Lucchinin *et al.* 1998; Tilsala-Timisjarvi & Alatossava, 1998).

Direct PCR methodologies are also available for the differentiation of LAB. These include *in situ* PCR, specific PCR and community analysis (cloning libraries or denaturing gradient gel electrophoresis/temperature gradient gel electrophoresis, DGGE/TGGE). As with genetic probes, the list of species-specific PCR primer sets is extensive for LAB and is increasing (see Table 2). Indeed, the application of PCR-based methods for differentiating LAB surpasses that of oligonucleotide-probing strategies. Along with 16S rDNA primers and RAPD-derived PCR primers, there is a growing interest in utilising intergenic sequences (ITS; most notably the 16S–23S rRNA spacer region); (Lebond-Bourget *et al.* 1996; Tilsala-Timisjarvi & Alatossava, 1997; Torriani *et al.* 1999), as well as protein- and enzyme-encoding genes, such as heat-shock protein genes (*hsp*; Jian *et al.* 2001), the *recA* gene (Kullen *et al.* 1997) and the *ldh* gene (Roy & Sirois, 2000). Recent work by Tannock *et al.* (1999) demonstrated the usefulness of sequencing the 16S–23S ITS region for strain identification of lactobacilli. This strategy proved to have greater characterisation powers than sequencing the 16S V2–V3 region, discriminating *Lactobacillus casei* and *Lactobacillus rhamnosus* strains. The 16S–23S ITS has also demonstrated good intraspecific analysis for bifidobacteria

(Lebond-Bourget *et al.* 1996), with strain-specific primers developed for three *Bifidobacterium* strains following sequence comparison of this spacer region (Brigidi *et al.* 2000).

A number of PCR techniques are available for the identification and/or differentiation of LAB. Some of these rely principally on the amplification of a PCR product or on the appropriately sized amplification product, e.g. using species-specific primers, whereas others depend on sequence comparison of the resulting amplicons. Work investigating the detection limit of such techniques, using *Bifidobacterium* species-specific primers, has calculated the threshold of PCR amplification as 100 cells per PCR reaction (equivalent to 10^6 cells per gram of faeces; Matsuki *et al.* 1999).

Multiplex PCR amplifies multiple DNA targets simultaneously by performing PCR amplification with multiple primer sets. In general, the more target sequences amplified, the greater the accuracy and reliability. The major limitation with multiplex PCR concerns optimisation of the PCR reaction conditions. This normally necessitates selection of primer sets with similar reaction specifications, especially annealing temperature. With the appropriate sequence information available it is feasible to develop or identify potential multiplex-PCR protocols. Workers have demonstrated the usefulness of this technique in detecting *Lactobacillus gasseri* strain 4B2 amongst colonies grown on *Lactobacillus*-selective agar (Lucchinin *et al.* 1998).

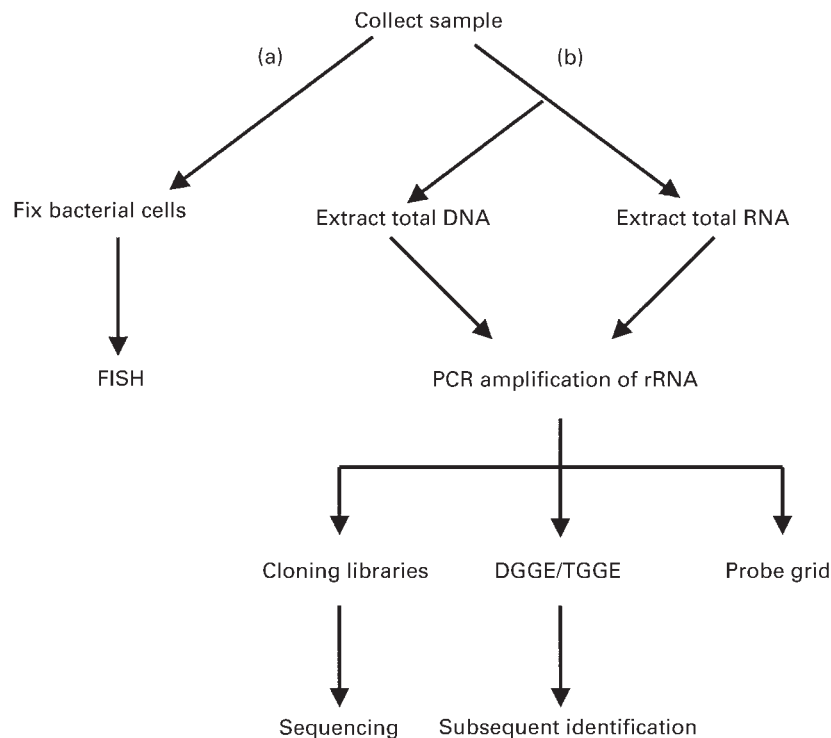


Fig. 1. Strategies for direct community analysis of mixed microbial populations by (a) detection of individual cells or (b) investigating the composition of 16S rRNA community. DGGE/TGGE, denaturing gradient gel electrophoresis/temperature gradient gel electrophoresis; FISH, fluorescent *in situ* hybridisation; PCR, polymerase chain reaction.

Community analysis

Two main strategies for the direct investigation of the diversity of mixed microbial populations (community analysis) are currently available. Both rely on 16S rRNA (see Fig. 1). Firstly, direct detection of individual bacterial cells within the population utilising fluorescently labelled 16S rDNA probes, FISH (Franks *et al.* 1998), and secondly, studying the composition of the 16S rRNA population, which is amplified using universal PCR primers. There are currently three techniques employed for such investigations: sequencing individual rRNA genes using cloning libraries (Suau *et al.* 1999), or profiling the 16S rRNA population by: (i) separation of individual rRNA genes using particular gel electrophoresis methods (such as DGGE and TGGE; Muyzer, 1999); or (ii) probing techniques, such as probe grids or checkerboard hybridisation (O'Sullivan, 1999).

Probing techniques, including FISH and checkerboard hybridisation, have been discussed previously in this article. Both of these methods have been applied successfully to studies of the human colonic microflora (Franks *et al.* 1998; O'Sullivan, 1999). Although these techniques are reliable and relatively easy to use, probe design and detection limits are the main disadvantages. Such protocols are useful for following bacterial groups or predominant genera; however, an extensive list of probes would be necessary for diversity studies. Conversely, probing strategies are excellent for following a limited number of bacterial groups of interest in mixed populations, provided that they are present at levels of 10^6 /ml and the appropriate probes are available.

Gene sequencing of 16S rRNA cloning libraries also has limited application in studies of the human GI microbiota (Suau *et al.* 1999), largely due to the number of clones produced and the fact that this method only allows qualitative analysis of the bacterial composition. Along with the general limitations of PCR-based methodologies, a further consideration is the biasing incurred due to the cloning system employed. Although useful for identification of novel or 'unknown' organisms, such methods are too laborious for clinical trials and/or extensive human studies. Advances in molecular machinery, such as capillary sequencers, go some way to addressing this; however, until the overall costs and analysis packages improve, the usefulness of this technique in probiotics and human microbiota research is minimal. Indeed, perhaps the primary use of this method is the identification of previously unaccounted for members of the microflora and the design of appropriate probes to investigate their predominance in larger population studies. Gene sequencing itself, however, has been shown to be a very effective method for identification and taxonomic purposes (Lebond-Bourget *et al.* 1996; Wintzingerode *et al.* 1997; O'Sullivan, 1999; Giraffa & Neviani, 2000). The majority of work so far has focused on the ribosomal genes, especially the 16S gene and the 16S–23S ITS region, but other genetic markers have also been identified. Comparisons of sequences of a number of genes and/or partial genes have been studied for both phylogenetic analysis and for species or strain differentiation.

Profiling the 16S rRNA population by DGGE or TGGE enables both the rapid estimation of diversity and the simultaneous analysis of multiple samples (Muyzer *et al.* 1993; Zoetendal *et al.* 1998; Muyzer, 1999). The general principle of DGGE/TGGE is the separation of individual rRNA genes based on differences in chemical stability or melting temperature of these genes. Polyacrylamide gels consisting of a linear denaturing gradient, formed by urea and formamide are employed for DGGE, whereas a linear temperature gradient is used during TGGE. Most DGGE/TGGE studies to date have concentrated on the microbial diversity of different samples by comparing the number and patterns of bands generated. However, as with all PCR-based techniques, semi-quantification is also possible based on the intensity of the bands. Competitive reverse transcription PCR of community rRNA and known standards enables absolute quantification of band intensities in the profile. However, one limitation of DGGE/TGGE is that heterologous sequences may migrate similarly, and thus bands at the same position in the gel are not necessarily phylogenetically related (Muyzer *et al.* 1993). The resolution may be improved by narrowing the gradient. Additionally, subsequent identification work with DGGE profiles, such as hybridisation analysis, or cloning and sequencing of excised bands improves the information obtained on the complexity of the bacterial community with respect to diversity and identity (Muyzer, 1999).

The most important limitations specific to DGGE/TGGE are heteroduplex detection, detection of heterologous rRNA operons from the same organism and poor detection sensitivity of rare members. Indications are that DGGE/TGGE profiles represent 90–99% of the bacterial community (Zoetendal *et al.* 1998). Given the assumption that human faecal samples contain 10^{11} cells per gram, this suggests a detection limit of 10^9 . However, the presence of heteroduplexes can result in overestimation of bacterial diversity. Recent work has shown that heteroduplex frequency is a function of PCR product production, increased cycle number, increased template concentration and species diversity (Qiu *et al.* 2001). Heteroduplexes can be eliminated, by cutting the bubble in the heteroduplex with T7 endonuclease I, for example or by using polyacrylamide gel electrophoresis. However, the experimental conditions for both of these systems are critical, especially for separating heteroduplexes formed between highly related strains or those with very similar conformations to homoduplex molecules (Muyzer *et al.* 1993).

DGGE/TGGE can be employed for whole community analysis, or for investigations of specific populations or groups within the sample. The selection of PCR primers determines the level at which the profiling is aimed either global, genus or species. Additionally, recent work has demonstrated the ability of DGGE profiling to group *Lactobacillus* isolates which were then subjected to more rigorous classification techniques for definitive species identification (Walter *et al.* 2000). This work demonstrated the potential of polyphasic studies to reduce the time, cost and extensive testing required while also maintaining specificity. Both TGGE and

DGGE have been successfully used in studying the complexity and dynamics of the human faecal flora. These techniques are reliable, rapid, comparatively inexpensive and show good reproducibility. More recent studies have demonstrated the application of such methods in examining the diversity of the bifidobacteria (Satokari *et al.* 2001) and lactobacilli (Tannock *et al.* 1999) populations of mixed communities.

Other molecular techniques

As the knowledge and information increases so does the advancement of molecular technologies. There are always new or improved methods being described. Two techniques worthy of mention here are terminal restriction fragment length polymorphism (T-RFLP) and multilocus sequence typing (MLST).

Terminal restriction fragment length polymorphism analysis can determine subtle differences in genotypes (Marsh, 1999). Fluorescently tagged amplicons of specific sequences are produced during PCR by using fluorescently tagged primers. Subsequent restriction digests and visualisation of the fluorescent-labelled terminal fragments on high-resolution sequencing gels produces the T-RFLP patterns. Such analyses have been used in environmental studies, including studies of bacterial soil communities, and are an additional tool to community diversity studies. The main advantages of T-RFLP over other community analysis techniques, such as DGGE/TGGE, is the improved resolution afforded by nucleic acid sequencing, compared to electrophoretic systems, and the digital output. The method also shows good sensitivity and high-throughput, and has been employed for strain identification as well as comparative community analysis. To date, most T-RFLP has concentrated on 16S rRNA sequences. However, other genetic markers, such as *hsp* genes, and markers for glutamine synthetase, ATPases and topoisomerases, may be useful in estimating biodiversity (Marsh, 1999).

MLST is a modification of multilocus enzyme electrophoresis (MLEE), using nucleic acid sequencing of house-keeping loci rather than electrophoretic mobilities of their products (Spratt, 1999). Primarily, MLST is an identification tool for differentiating isolates with highly related genotypes. Because of the increased sensitivity of nucleotide sequencing, compared to electrophoretic mobilities, MLST based on seven loci can provide excellent discrimination. Bacterial isolates are defined by a string of integers representing their allelic profile. This digital output is an important advantage of the technique, as it allows easy comparison within and between laboratories via a central MLST website (<http://www.mlst.zoo.ox.ac.uk/>; Spratt, 1999). To date MLST applications have largely involved clinical isolates and known pathogens, although it is ineffective in the case of *Mycobacterium tuberculosis* isolates, as they contain virtually no sequence variation in their house-keeping genes.

Summary

Genetic characterisation techniques are suitable for monitoring complex communities. DNA fingerprinting enables

differentiation of bacterial strains, but requires cultivation in most instances. A number of *in situ* methods based on 16S rRNA gene sequences are available, some of which are limited in discriminating closely related bacterial species or strains within a species. Taken together, a polyphasic strategy principally based on genetic technologies is necessary in order to achieve an accurate interpretation of the bacteriology of the GI tract and to elucidate the 'true' microbiological impact of probiotics. Employment of such an approach will overcome the inherent limitations of each system. With the current plethora of genetic techniques for identification, differentiation and community characterisation, the stage is set for an inundation of information from thorough investigations of the normal colonic microbiota, its overall complexity, diversity and dynamics, and the effects of probiotics and other dietary modulation at the bacterial level. Additionally, the tools are available for definitive analysis of the bacterial components of probiotic products, quality control and strain integrity. Future developments in molecular biology should be aimed at identifying and developing methods to study the status of functional elements and biomarkers suitable for unravelling the activity of the bacterial population, specific organisms and/or genes. The quest will be to demonstrate the roles of the microflora and of probiotic strains *in vivo* and not simply to characterise the population or demonstrate the presence or absence of particular organisms or groups.

References

- Brigidi P, Vitali B, Swennen E, Altomare L, Rossi M & Matteuzzi D (2000) Specific detection of *Bifidobacterium* strains in a pharmaceutical probiotic product and in human faeces by PCR. *Systematic and Applied Microbiology* **23**, 391–399.
- Charteris WP, Kelly PM, Morelli L & Collins JK (1997) Review article: selective detection, enumeration and identification of potential probiotic *Lactobacillus* and *Bifidobacterium* species in mixed bacterial populations. *International Journal of Food Microbiology* **35**, 1–27.
- Franks AH, Harmsen HJM, Raangs GC, Jansen GJ, Schut F & Welling GW (1998) Variations of bacterial populations in human faeces measured by fluorescent *in situ* hybridisation with group-specific 16S rRNA-targeted oligonucleotide probes. *Applied and Environmental Microbiology* **64**, 3336–3345.
- Giraffa G & Neviani E (2000) Molecular identification and characterisation of food-associated lactobacilli. *Italian Journal of Food Science* **4**, 403–423.
- Hozapfel WH, Haberer P, Geisen R, Bjorkroth J & Schillinger U (2001) Taxonomy and important features of probiotic microorganisms in food and nutrition. *American Journal of Clinical Nutrition* **73**(S), 365S–373S.
- Jian W, Zhu L & Dong X (2001) New approach to phylogenetic analysis of the genus *Bifidobacterium* based on partial HSP60 gene sequences. *International Journal of Systematic and Evolutionary Microbiology* **51**, 1622–1638.
- Kimura K, McCartney AL, McConnell MA & Tannock GW (1997) Analysis of fecal populations of bifidobacteria and lactobacilli and investigation of the immunological responses of

- their human hosts to the predominant strains. *Applied and Environmental Microbiology* **63**, 3394–3398.
- Klein G, Pack A, Bonaparte C & Reuter G (1998) Taxonomy and physiology of probiotic lactic acid bacteria. *International Journal of Food Microbiology* **41**, 103–125.
- Kullen MJ, Brady LJ & O'Sullivan DJ (1997) Evaluation of using a short region of the *recA* gene for rapid and sensitive speciation of dominant bifidobacteria in the human large intestine. *FEMS Microbiology Letters* **154**, 377–383.
- Langendijk PS, Schut F, Jansen GJ, Raangs GC, Kamphuis GR, Wilkinson MHF & Welling GW (1995) Quantitative fluorescence *in situ* hybridisation of *Bifidobacterium* spp. with genus-specific 16S rRNA-targeted probes and its application in fecal samples. *Applied and Environmental Microbiology* **61**, 3069–3075.
- Lebond-Bourget N, Philippe H, Mangin I & Decaris B (1996) 16S rRNA and 16S to 23S internal transcribed spacer sequence analyses reveal inter- and intraspecific *Bifidobacterium* phylogeny. *International Journal of Systematic Bacteriology* **46**, 102–111.
- Lucchinin F, Kmet V, Cesena C, Coppi L, Bottazzi V & Morelli L (1998) Specific detection of a probiotic *Lactobacillus* strain in faecal samples by using multiplex PCR. *FEMS Microbiology Letters* **158**, 273–278.
- Marsh TL (1999) Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterising diversity among homologous populations of amplification products. *Current Opinions in Microbiology* **2**, 323–324.
- Matsuki T, Watanabe K, Tanaka R, Fukuda M & Oyaizu H (1999) Distribution of bifidobacterial species in human intestinal microflora examined with 16S rRNA-gene-targeted species-specific primers. *Applied and Environmental Microbiology* **65**, 4506–4512.
- McCartney AL, Wang W & Tannock GW (1996) Molecular analysis of the composition of the bifidobacterial and lactobacillus microflora of humans. *Applied and Environmental Microbiology* **62**, 4608–4613.
- Muyzer G (1999) DGGE/TGGE a method for identifying genes from natural ecosystems. *Current Opinions in Microbiology* **2**, 317–322.
- Muyzer G, de Waal EC & Uitterlinden AG (1993) Profiling of complex microbial populations by DGGE analysis of PCR-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology* **59**, 695–700.
- O'Sullivan DJ (1999) Methods for analysis of the intestinal microflora. In *Probiotics. A Critical Review*, pp. 23–44 [GW Tannock, editor]. Norfolk, UK: Horizon Scientific Press.
- O'Sullivan DJ & Kullen MJ (1998) Tracking of probiotic bifidobacteria in the intestine. *International Dairy Journal* **8**, 513–525.
- Prasad J, Gill H, Smart J & Gopal PK (1998) Selection and characterisation of *Lactobacillus* and *Bifidobacterium* strains for use as probiotics. *International Dairy Journal* **8**, 993–1002.
- Qiu X, Wu L, Huang H, McDonel PE, Palumbo AV, Tiedje JM & Zhou J (2001) Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. *Applied and Environmental Microbiology* **67**, 880–887.
- Roy D & Sirois S (2000) Molecular differentiation of *Bifidobacterium* species with amplified ribosomal DNA restriction analysis and alignment of short regions of the *ldh* gene. *FEMS Microbiology Letters* **191**, 17–24.
- Roy D, Sirois S & Vincent D (2001) Molecular discrimination of lactobacilli used as starter and probiotic cultures by amplified ribosomal DNA restriction analysis. *Current Microbiology* **42**, 282–289.
- Satokari RM, Vaughan EE, Akkermans ADL, Saarela M & de Vos WM (2001) Bifidobacterial diversity in human faeces detected by genus-specific PCR and denaturing gradient gel electrophoresis. *Applied and Environmental Microbiology* **67**, 504–513.
- Spratt BG (1999) Multilocus sequence typing: molecular typing of bacterial pathogens in an era of rapid DNA sequencing and the Internet. *Current Opinions in Microbiology* **2**, 312–316.
- Suau A, Bonnet R, Sutren M, Godon J-J, Gibson GR, Collins MD & Dore J (1999) Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Applied and Environmental Microbiology* **65**, 4799–4807.
- Tannock GW, Tilsala-Timisjarvi A, Rodtong S, Ng J, Munro K & Alatossava T (1999) Identification of *Lactobacillus* isolates from the gastrointestinal tract, silage, and yoghurt by 16S–23S rRNA gene intergenic spacer region sequence comparisons. *Applied and Environmental Microbiology* **65**, 4264–4267.
- Tilsala-Timisjarvi A & Alatossava T (1997) Development of oligonucleotide primers from the 16S–23S rRNA intergenic sequences for identifying different dairy and probiotic lactic acid bacteria by PCR. *International Journal of Food Microbiology* **35**, 49–56.
- Tilsala-Timisjarvi A & Alatossava T (1998) Strain-specific identification of probiotic *Lactobacillus rhamnosus* with randomly amplified polymorphic DNA-derived PCR primers. *Applied and Environmental Microbiology* **64**, 4816–4819.
- Torriani S, Zapparoli G & Dellaglio F (1999) Use of PCR-based methods for rapid differentiation of *L. delbrueckii* subsp. *bulgaricus* and *L. delbrueckii* subsp. *lactis*. *Applied and Environmental Microbiology* **65**, 4351–4356.
- Tynkkynen S, Satokari R, Saarela M, Mattila-Sandholm M & Saxelin M (1999) Comparison of ribotyping, randomly amplified polymorphic DNA analysis, and pulsed-field gel electrophoresis in typing of *Lactobacillus rhamnosus* and *L. casei* strains. *Applied and Environmental Microbiology* **65**, 3908–3914.
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M & Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**, 4407–4414.
- Walter J, Tannock GW, Tilsala-Timisjarvi A, Rodtong S, Loach DM, Munro K & Alatossava T (2000) Detection and identification of gastrointestinal *Lactobacillus* species by DGGE and species-specific PCR primers. *Applied and Environmental Microbiology* **66**, 297–303.
- Wang R-F, Cao W-W & Cerniglia CE (1996) PCR detection and quantification of predominant anaerobic bacteria in human and animal faecal samples. *Applied and Environmental Microbiology* **62**, 1242–1247.
- Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP & Truper HG (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *International Journal of Systematic Bacteriology* **37**, 463–464.
- Wintzingerode FV, Gobel UB & Stackebrandt E (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiology Reviews* **21**, 213–229.
- Zoetendal EG, Akkermans ADL & de Vos WM (1998) Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria. *Applied and Environmental Microbiology* **64**, 3854–3859.