Feasibility Study into the Use of DNA Sequencing for the Identification of Probiotic Bacteria

Claire Bushell and Malcolm Burns[(a)], Molecular and Cell Biology, LGC, Queens Road, Teddington, Middlesex, TW11 0LY, UK.

(a) To whom all correspondence should be sent. E-mail: Malcolm.Burns@lgcgroup.com, Telephone: (+44) (0)208 943 7000, Fax: (+44) (0)208 943 2767

Summary

In recent years the European Union and UK markets have witnessed a large increase in the availability of probiotic foods and food supplements. The current study investigated the use of the molecular biology techniques of PCR and DNA sequencing for the detection and identification of probiotic bacteria. The aim was to provide "proof of principle" of the applicability of these methods in the area of food analysis. These methods have the potential to improve efficacy of identification compared to more traditional culture approaches, as well as providing a confirmatory role. The work involved amplifying specific regions of DNA from a selection of probiotic bacteria commonly found in foods, followed by sequencing and species identification using sequence databases. Initial results showed clear potential for successful species identification, however issues relating to the use of generic assays, background contamination of reagents with bacterial species and interpretation of database results were also highlighted.

Keywords

DNA sequencing, PCR, Probiotic bacteria, 16S, *Bifidobacterium*, *Lactobacillus*, *Streptococcus*

Introduction

Probiotic foods are those containing health-promoting probiotic bacteria. Probiotics are defined as "live microorganisms which when administered in adequate amounts confer a health benefit on the host"[1]. Probiotic foods range from fermented milks to lyophilised (freeze-dried) preparations containing both single and multiple bacterial strains. In order to be effective, the bacteria need to be viable and present in significant numbers. Traditional species of interest include specific strains of *Lactobacillus* and *Bifidobacterium*. However, the taxonomy of probiotic species has changed over the past 15 years and new species and strains are likely to be further introduced by the food industry in the future. Doubts have been raised about whether some of the bacteria added to food products actually survive food processing, distribution and digestion, and whether the correct strains have been added. Recent Food Standards Agency (FSA) sponsored research stated that out of 35 bacterial strains in 12 commercial products tested only *Lactobacillus* was sufficiently robust to survive the whole digestive process[2]. A number of product surveys have found low numbers of viable probiotic bacteria[3] and also misidentification and mislabelling of strains[4].

The market for probiotics is experiencing rapid growth globally, including recent expansion within the UK[5]. While there are many reputable products on the market previous work by Public Analysts, the FSA and a

Government Chemist referee case have demonstrated problems including a) quantities of viable probiotic bacteria very much lower than the label declarations and b) the use of some inappropriate analytical methods for detection [6]. Media reports have suggested that half of the products available in the UK, many on the World Wide Web, had incorrectly labelled bacteria present which were of no benefit, prompting continuing consumer confusion. Additionally, recent legislation will only permit health and nutrition claims if they are based on independently assessed science[7]. Regulators and the legitimate trade wish to address these issues but may be constrained currently by lack of confidence in test methods and low availability of credible laboratory testing services in respect of probiotics.

Traditional approaches for detection and identification of probiotic bacteria include, typically, an overnight culturing stage followed by microscopical examination. In this study, the aim was to provide "proof of principle" that the technology of DNA sequencing could be applied for the detection and identification of probiotic bacteria. The potential benefits of this approach would permit more objective identification of probiotic bacterial strains, allow the speed and efficacy of identification to be improved, as well as providing an alternative and confirmatory approach for probiotic identification.

The work involved targeting regions of the 16S rRNA complex found in bacteria. In addition to highly conserved primer binding sites, 16S rRNA gene sequences contain hypervariable regions which can provide species-specific signature sequences useful for bacterial identification. There is limited published literature on the use of DNA sequencing for probiotic bacteria identification, so this approach is considered of value to the scientific community.

Approach

Use of 16S rRNA

Ribonucleic acid (RNA) is a biologically-important type of molecule that consists of a long chain of single stranded nucleotide units. Ribosomal RNA (rRNA) is a type of non-coding RNA that is involved in translating messenger RNA (mRNA) into proteins, and acts as the catalytic component of the ribosome. The 16S rRNA is a part of the ribosomal RNA − a 1542 nucleotide long component of the small prokaryotic ribosomal subunit (30S). The 16S rRNA sequence is used for phylogenetic studies as it is highly conserved between different species of bacteria. DNA primers, for use in the Polymerase Chain Reaction (PCR), can be designed to the highly-conserved sequences on the 16S DNA, which often flank hypervariable regions that can provide species-specific signature sequences useful for identification. As a result, 16S DNA gene sequencing has become prevalent in medical microbiology as a rapid, accurate alternative to phenotypic methods of bacterial identification.

Selection of a Subset of Probiotic Bacteria

A review of the current literature for probiotic bacteria was performed using NCBI PubMed searches[8] to identify probiotic strains of importance to the food industry. Of particular interest was the frequency with which a specific bacterium was used as a target for detection in published literature, as well as lists of the most common supplement probiotic species[9]. This information was combined with data gathered from consultation with internal and external experts (personal communications from Health Food Manufacturers' Association) on the prevalence of important probiotic species, in order to make an informed decision on which subset of probiotic bacteria to study.

Based on these findings, and the availability of DNA from ATCC reference materials (LGC standards, Teddington, UK), five probiotic bacteria were selected as templates for PCR detection assays (Table 1). Sequence information for each of the chosen bacteria was obtained from the GenBank database[10] (where available) with particular emphasis on the 16S gene region.

Table 1 Five Probiotic Bacterial DNA Templates Selected for the Study

Probiotic Bacteria	ATCC DNA #	Source*
Lactobacillus acidophilus	4357D-5	Human
Lactobacillus bulgaricus	11842D-5	Dairy products (Bulgarian yogurt)
Lactobacillus casei	334D-5	Dairy products (Emmental cheese)
Streptococcus thermophilus	BAA-250D-5	Dairy products ([111718] commercially-prepared yogurt)
Bifidobacterium breve	15700D-5	Instestine of infant

*Source refers to the source of the reference material as stated on the ATCC reference standards provision form

Selection of Primer Sets for PCR Amplification of DNA from Probiotic Bacteria

Several PCR primer sets were chosen from published literature (Table 2) to include both generic and group (genus) specific detection of the particular probiotic strains selected. All PCR assays targeted the 16S gene except for the *Streptococcus* group set which was targeted towards the *tuf* gene (elongation factor *Tu*). This region had been identified in the literature as an alternative target area for potentially improved bacterial species specific detection[11 12].

Table 2 Four Assays used to Amplify Probiotic Bacteria

Target	Publication ref	Primers/Probe sequences (5'-3')	TM °C	Amplicon length (bp)	Design region
Streptococcus group	12	Tuf-Strep-1: GAAGAATTGCTTGAATTGGTTGAA Tuf-Strep-R: GGACGGTAGTTGTTGAAGAATGG	62	560	tuf gene
All bacteria ('All bac 1')	14	F_eub TCC TAC GGG AGG CAG CAG T R_eub GGA CTA CCA GGG TAT CTA ATC CTG TT	59 58	466	16S 16S
Bifidobacterium group	15	g-Bifid-F: CTCCTGGAAACGGGTGG g-Bifid-R: GGTGTTCTTCCCGATATCTACA	50	549-563	16S 16S
Lactobacillus group	16	Lab 159aF: GGAAACAGATGCTAATACCG Lab 677R: CACCGCTACACATGGAG	61	545	16S 16S

The primer sequences were aligned to GenBank sequences[10] for the chosen ATCC bacterial strains, to check for the presence of the correct priming sites. The expected amplicon sizes were all approximately 500bp (± about 60bp), with some variation expected between different species and strains. A check of theoretical primer cross-reactivity with other bacteria and non-bacterial organisms was conducted by performing a BLAST analysis[13] of each oligonucleotide. Of most interest were any hits on species common in foodstuffs e.g. wheat, soya and rice, as well as human or animal sequences and bacteria belonging to other groups (for the group-specific assays). None of the results suggested any potential specificity problems. All primers were synthesised by Sigma Aldrich (St. Louis, MO, USA) with HPLC purification.

Amplification of Probiotic DNA and Cross-Reactivity Studies

All of the PCR assays were performed using standard PCR conditions on an Applied Biosystems® (Life Technologies, Carlsbad, California, USA) 9700 thermal cycler (95°C for 10min followed by 40 cycles of 95°C, 30sec; 60°C, 30sec; 72°C, 30sec and a final extension of 72°C for 7 minutes). The PCR products were visualised using an Agilent 2100 Bioanalyser (Agilent Technologies, Santa Clara, California, USA) with the DNA 1000 series chips.

Provision of PCR Products and Primers for Sanger Dideoxy Sequencing

DNA sequencing was performed by LGC Genomics (Berlin, Germany) using their barcode sequencing service for all reactions yielding a positive PCR result. PCR products were prepared according to the provider's specifications: prior to sequence analysis the amplicons were cleaned using a QIAquick PCR purification kit (Qiagen, Hilden, Germany) to remove any primers and PCR buffer remaining within the samples. 10µL of the template DNA solution was required per sequencing read at a concentration of 10ng/µL for 200-500bp products; and 20ng/µL for 500-1000bp products. 4µL of the custom primers were added to each sample at a concentration of 5µM as the sequencing primer. Two forward and two reverse reads were performed on each PCR product.

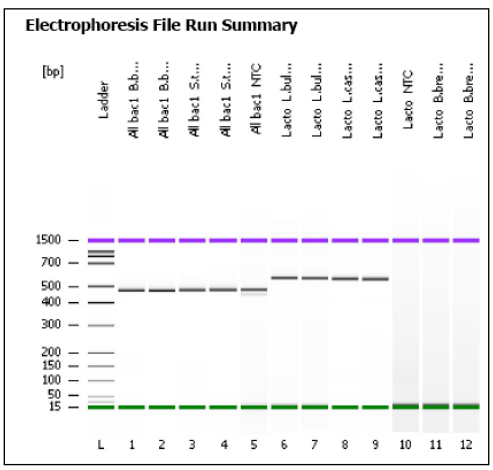
Verification of Sequencing Results against known Probiotic Sequences

The sequencing results from the PCR products analysed by LGC Genomics were compared to sequences held in the GenBank database using the BLAST sequence alignment program[13]. The sequences of the amplicons, generated from each of the four test assays, were aligned against publicly available sequences within this database, in order to assign candidate sequences with a genus and where possible a species name to each sample (16S rDNA or *tuf* gene region). The results with the highest "Max Identity" (%) were taken as the closest matches for each query sequence and subsequently used to assign the most likely candidate species to the DNA sequence query. The Max Identity is defined as the percentage sequence information that is identical between the query and candidate sequence[13]. The sequencing results were also aligned to the closest matched reference sequences for the specific ATCC templates used in each assay (where available), held within the NCBI database. This enabled a percentage alignment score to be established for each sample.

Results and Discussion

Results from the PCR which were visualised on the Agilent Bioanalyser, demonstrated that all of the assays produced DNA amplicons in the expected size range. An example electropherogram is shown in Figure 1.

Figure 1 Agilent Electropherogram displaying Typical PCR Amplicons



A molecular ladder is shown in lane "L" ranging from 15 to 1500 base pairs of DNA. Lanes 1-5 show products from the generic "All bacteria" assay with *Bifidobacterium breve* template (lanes 1-2); *Streptococcus thermophilus* template (lanes 3-4) and a positive amplification in the No Template Control (lane 5). Lanes 6-12 show products from the *Lactobacillus* group assay with *Lactobacillus bulgaricus* template (lanes 6-7);

Lactobacillus casei (lanes 8-9); a PCR No Template Control (lane 10); and no cross reactivity with *Bifidobacterium breve* (lanes 11-12).

The generic primer set (All bacteria) successfully amplified all three of the ATCC bacterial groups (*Lactobacillus, Bifidobacterium and Streptococcus*) as expected. The No Template Controls (NTC's) also produced a detectable product with this assay, which was likely to be due to background amplification associated with the bacterial 16S target region (discussed below). The group specific assays were all able to amplify their matched target bacterial strain, although there was some potential weak cross reactivity of *S. thermophilus* with the *Lactobacillus* assay and some relatively strong cross reactivity of the *B. breve* template with the *Streptococcus* assay, as shown in Figure 1. "Weak positive" results were identified visually as very feint bands on the gel image and quantified as less than 1ng/µl per reaction using the Agilent Bioanalyser. NTC's in the group specific assays were clear in the majority of cases with the one exception being one replicate of the *Streptococcus* assay which had a weakly positive result (Table 3).

Table 3 Assay Specificity Determined Experimentally

Assay / Template	All bac1	Lacto group	Bifido group	Strepto group
L. acidophilus				
L. bulgaricus				
L. casei				
B. breve				
S. thermophilus				
NTC				

Key	
	Strong positive
	Weak positive
	Negative
	Not tested

A summary of the BLAST identification results based on sequencing all of the PCR products for all positive PCR samples is displayed in Table 4. Experiment #1 shows that the generic 16S assay ("All bac1") enabled identification of the *Lactobacillus* sample to the correct genus, and the top candidate sequences at 99% maximal identity also included the correct species of *L. acidophilus.* Hence the correct species of *L. acidophilus* was listed alongside other *Lactobacillus* species but not uniquely identified as the only candidate species. Experiment #2 shows that there were also issues using the generic "All bac1" assay in general with bacterial detection in the NTCs (discussed below).

Table 4 Summary of Sequencing Results based on BLAST Identification using "Max Identity" Result Criterion to show most-likely Candidate Sequence Matches using NCBI Database

Experiment	Probiotic Template	Assay	BLAST Identification	Maximal Identity	Correct Genus?	Correct Species?
#1	*L.acidophilus*	All bac1	Various *Lactobacillus* species including *L.acidophilus*	99%	Yes	Not uniquely
#2	NTC	All bac1	Predominantly *Escherichia* species but also *Shigella*, *Cronobacter* and "Uncultured bacterial clones"	99 to 100%	N/A	N/A
#3	*L.acidophilus*	*Lactobacillus*	*Lactobacillus acidophilus* (FR683087) and other "Uncultured bacterial clones"	100%	Yes	Likely[1]
#4	*L.bulgaricus*	*Lactobacillus*	*Lactobacillus casei* and. *paracasei* strains	100%	Yes	No
#5	*L.casei*	*Lactobacillus*	Numerous *Lactobacillus casei* and occasional *L. paracasei*	99%	Yes	Likely[2]
#6	*B.breve*	All bac1	Predominantly *Bifidobacterium breve* with occasional *Bifidobacterium sp.* and "Uncultured bacterial clones"	100%	Yes	Likely[1]
#7	*B.breve*	*Bifidobacterium*	Predominantly *Bifidobacterium breve* with occasional *Bifidobacterium sp.* and "Uncultured bacterial clones"	99%	Yes	Likely[1]
#8	*S.thermophilus*	All bac1	*Streptococcus thermophilus* and "Uncultured bacterial clones"	99%	Yes	Likely[1]
#9	*S.thermophilus*	*Streptococcus*	Single 100% hit for *Streptococcus thermophilus* (CP000023)	100%	Yes	Yes
#10	*B.breve*	*Streptococcus*	Single 100% hit for *Streptococcus agalactiae* (AE009948)	100%	No	No
#11	*L.acidophilus*	*Streptococcus*	*Streptococcus pneumoniae*	100%	No	No

Likely[1]: likely that specific species has been identified if the "Uncultured bacterial clones" are shown to be the same species as the query sequence
Likely[2]: likely that specific species has been identified if the taxonomic classification of *L. paracasei* is the same as *L. casei*

Experiments #3 to #5 show that the genus-specific 16S assays for *Lactobacillus* were able to successfully identify samples to the correct genus. Specificity was improved with these genus-specific assays as expected, and the NTC's did not yield positive results. This was an issue previously observed with the generic primer set, so the genus-specific assays had the benefit of lower false positive rates. In Experiment #3 the genus-specific *Lactobacillus* assay with the *L. acidophilus* template identified the correct species alongside other un-classified and un-named "uncultured bacterial clones" If these clones are shown to also be *L. acidophilus* then it is likely the assay is useful for identifying the correct species as well. In Experiment #4 the *Lactobacillus* assay with the *L. bulgaricus* template was able to identify the correct genera but not the correct species. In Experiment #5 the *Lactobacillus* assay with the *L. casei* template identified the top candidate sequences as being *L. casei* and *L. paracasei*. Under certain taxonomic schemes the latter may be classified as *L. casei* in which case this assay has shown correct species identification as well.

In Experiment #6 the generic 16S assay ("All bac1") enabled identification of the *Bifidobacterium* sample to the correct genus, and the top candidate sequences also included the correct species of *B. breve* as well as other uncultured bacterial clones. Again, if the latter are shown to actually be sequences originating from *B. breve* then the generic "All bac1" assay may also facilitate species specific identification.

The genus specific 16S assay for *Bifidobacterium* in Experiment #7 was able to show genus-specific identification. Additionally this assay may also show species-specific identification if the *Bifidobacterium sp*. and uncultured bacterial clones are shown to be *B. breve* in origin.

Using the "All bac1" assay the *S. thermophilus* template in Experiment #8 was identified correctly as a candidate DNA sequence alongside other uncultured bacterial clones The *Streptococcus* assay targeting the *tuf* gene appeared to have good specificity for its target bacterial genus, and in the case of Experiment #9 the *S. thermophilus* template showed complete specificity for the correct species at 100% maximal identity. However the specificity and sensitivity of the *Streptococcus* assay may be too high to use in a screening type approach as the assay also produced a positive result for its specific target even when other bacteria were present as the main template. This is shown in Experiment #10 where the *Streptococcus* assay suggested *S. agalactiae* was present even when the DNA template was *B. breve*, and in Experiment #11 where the assay suggested *S. pneumoniae* was present when the DNA template was *L. acidophilus*. This suggests that the assay does not necessarily demonstrate cross-reactivity with other groups, but is able to amplify environmental or low levels of *Streptococcus* species even in the presence of other bacteria at high concentrations. Additionally note that the reference materials provided are based on the presence of DNA from a particular bacterial species, and this does not preclude inclusion of other bacterial DNA in the reference standard. This could lead to mis-identification of the constituent probiotic bacteria within a test sample.

Whilst it was useful to show the closest match between the query sequence and likely candidate sequences on NCBI based on the Max Identity criterion (Table 4), it was also useful to examine the percentage identity match between the query sequence and available sequence information for the true reference/template samples where available (Table 5). The results indicate close matches between the query sequence and the template/reference sequence, with percentage identity varying between 99 and 100% in all except two cases. The first exception is for 89% identity between the query sequence for the *L. bulgaricus* template and the ATCC standard when using the *Lactobacillus* assay. This may be in part due to the specificity of the general "*Lactobacillus*" primers that appear selective for the genus *Lactobacillus* but may not allow differentiation between closely related species. The second exception was the *Streptococcus* assay, which produced a *Streptococcus* specific amplicon when in the presence of the *Bifidobacterium breve* template. As mentioned, this latter example may be because of low level environmental contamination as opposed to actual amplification from the *Bifidobacterium* template itself.

Table 5 Summary of the BLAST Identification Results for all Positive PCR Samples

Probiotic Template	Assay	Percentage "Identity" Alignment to Reference Sequences
L. acidophilus	All bac1	99% to *L. acidophilus* partial 16S rRNA gene, strain VPI 6032 (FR683087)
L. acidophilus	*Lactobacillus*	100% to *L. acidophilus* partial 16S rRNA gene, strain VPI 6032 (FR683087)
L. acidophilus	*Streptococcus*	No significant alignment found to *L. acidophilus* partial 16S rRNA gene, strain VPI 6032 (FR683087)
L. bulgaricus	*Lactobacillus*	89% to ATCC 11842 (NC_008054.1) 89% to *L.delbrueckii* subsp. *bulgaricus* 2038, complete genome (CP000156.1)
L. casei	*Lactobacillus*	99% to *L. casei* strain NQ2-2 16S ribosomal RNA gene, partial sequence (GU299083.1) 99% to *L. casei* ATCC 334, complete genome (NC_008526)
B. breve	All bac1	100% to ATCC 15700 (AB006658.1) 100% to *B. breve* DSM 20213 B_breve-1.0.1_Cont1.3, whole genome shotgun sequence (NZ_ACCG02000012.1)
B. breve	*Bifidobacterium*	99% to ATCC 15700 (AB006658.1) 99% to *B. breve* DSM 20213 B_breve-1.0.1_Cont1.3, whole genome (NZ_ACCG02000012.1)
B. breve	*Streptococcus*	No significant alignment found to AB006658.1 *Bifidobacterium breve* gene for 16S rRNA, partial sequence, strain: ATCC 15700
S. thermophilus	All bac1	99% to *S.thermophilus* LMG 18311 chromosome, complete genome (NC_006448.1)
S. thermophilus	*Streptococcus*	100% to *Streptococcus thermophilus* LMG 18311 chromosome, complete genome (NC_006448.1)
NTC	All bac1	No template

The percentage alignments based on the percentage identity between the query sequence and the ATCC template candidate sequence in the database, taking into account any gaps in sequences if present.

Background Bacterial Contamination

Bacterial contamination of NTC's, as seen in the generic 16S assays, is a well-documented issue arising from background DNA present in reagents, particularly polymerases[17]. Researchers have tested methods to overcome the problem using UV treatment of reagents and shorter PCR cycling[18], but success is limited as this can compromise sensitivity of the assays. The generic nature of using broad-range primers that amplify the 16S ribosomal DNA sequence (which is often present as multiple copies) has meant that low levels of background bacterial contamination are commonly observed and reported in PCR experiments[16]. The issue is compounded by the fact that the enzymes used to mediate the PCR reaction (e.g. *Taq* polymerase) are often derived from bacteria themselves, and unless additional purification steps have been conducted to isolate a cleaner version of the enzyme, it is unlikely that the enzyme will be completely free of bacteria DNA. It is known that *Taq* polymerase has a high affinity for DNA and this DNA is co-purified during the production of the enzyme, resulting in commercially available *Taq* polymerases often containing some residual bacterial DNA. The same issue is also relevant to other PCR reagents, buffers and distilled water, where additional precautions may be necessary to ensure that these components are pure and free from background bacterial DNA.

The potential misidentification of bacterial species extends to the NTC's as well. The identification of *Shigella* species in some of the NTC's is more likely to be the presence of *E. coli* species. These are commonly found and known to be present in untreated Taq polymerase and often share over 99% sequence identity with *Shigella* species for the target region on the 16S gene target.

There are a variety of reasons for being unable to correctly identify with 100% certainty bacteria at the species level using sequencing based approaches. Firstly, the choice of assay primers may not be suitable for speciation purposes if, for example, the selected region does not have enough variation between different species to enable differentiation. Secondly, the degree of completeness and accuracy of the database being searched for sequence alignment will influence the possible outcomes. Issues surrounding database entries and the naming of bacterial species have previously been highlighted[19]. The prevalence of entries marked as uncultured bacterial clone without any further classification information does not promote accurate species identification. The lack of standardised criteria for 16s rDNA sequence information on public databases derived from different isolates of bacteria from the same species or genera has also been commented upon[18]. It has been suggested that for members of the same species 98-100% homology should be expected, whilst for members of the same genus this figure could show 97-99% homology. Finally, the individual users' interpretation of BLAST results e.g. how many results to consider, what percentage identity is used as a threshold for a positive identification etc., will also affect the outcome.

Conclusion

The UK market for so-called "healthy bacteria" in the form of probiotic food and food supplements has shown a large growth in recent years[5]. As with all packaged foods however, the sale of these products are subject to strict European Union food labelling legislation. There is a need to unequivocally detect and identify particular probiotic bacteria that are present in these food and food supplements in order to verify the labelling. Traditional approaches to identification of probiotics are based on culturing techniques followed by examination by eye, which could be criticised in terms of length of time and subjectivity. The aim of the current study was to provide "proof of principle" that the molecular approach of DNA sequencing could be used for identification of probiotic bacteria.

The results showed that the use of generic primers that target the 16S rDNA region has the potential to correctly identify DNA template from specific bacterial species, although some cross reactivity or background amplification does occur. Additionally, in common with results in the reported literature, blank PCR controls are likely to give background contamination with bacterial DNA present in the PCR enzyme. Greater specificity was observed using group (genera) specific primers and there were no incidences of background contamination in the blank controls used as part of these group specific assays.

Considering all the limitations of this sequencing-based method (issues with naming of bacteria, control of database entries, primer region used and background and low-level amplification) DNA sequencing shows promise for being used as an additional tool for confirmation of species-identification to complement pre-existing methods, as opposed to becoming a definitive stand-alone method itself at the current stage. These conclusions are based on a small set of data meant as a first step towards assessing the potential of this technique for probiotic bacterial speciation. Further investigation of different primer regions and a wider range of bacterial species could yield a better understanding of the potential of a DNA sequencing approach in this field.

There are a number of caveats associated with the findings in this study. DNA extraction procedures have not been examined, and it is likely that extraction of DNA from processed food materials may provide some challenges. Additionally, the approach of Sanger dideoxy sequencing[20] directly from uncultured bacterial samples may not provide clear results when the DNA target is in a mixed population, where more than one PCR product has been amplified. This can be overcome for Sanger dideoxy sequencing by using a cloning selection process or by culturing individual bacteria prior to sequencing, although this needs additional time and effort. The aspects of DNA extraction and mixed populations of probiotic bacteria are likely to cause additional challenges but will not be insurmountable given current knowledge and technologies e.g. the application of

advanced DNA extraction approaches and the use of Next Generation sequencing. Hence the results described in this article provide the foundation for further work into the usefulness and applicability of DNA sequencing for probiotic bacteria identification.

Acknowledgements

The work described in this paper was funded by the UK National Measurement System under the Measurement for Innovators Programme, as part of the Government Chemist Programme 2008-2011. The authors thank Dr. Carole Foy for her invaluable advice and assistance associated with the project.

References

1 World Health Organisation, 2001, Report of a Joint FAO/WHO Expert Consultation on "Evaluation of Health and Nutritional Properties of Probiotics in Food Including Powder Milk with Live Lactic Acid Bacteria" http://www.who.int/foodsafety/publications/fs_management/en/probiotics.pdf

2 Food Standards Agency, 2011, Final Technical report for FSA project ref G01022 "An Evaluation of Probiotic Effects in the Human Gut: Microbial Aspects." http://www.food.gov.uk/multimedia/pdfs/probioticreport.pdf

3 N.P. Shah, W.E.V. Lankaputhra, M. Britz, and W.S.A. Kyle, Survival of *L. acidophilus* and *Bifidobacterium bifidum* in Commercial Yogurt during Refrigerated Storage. International Dairy Journal, 1995, **5(5)**, 515-521

4 R. Temmerman, B. Pot, G. Huys, and J Swings, Identification and Antibiotic Susceptibility of Bacterial Isolates from Probiotic Products. International Journal of Food Microbiology, 2003, **81(1),** 1-10

5 C.Stanton, G.Gardiner, H. Meehan, K. Collins, G. Fitzgerald, P B. Lynch, and R P Ross, Market Potential for Probiotics. American Journal of Clinical Nutrition, 2001, **73(suppl)**, 476S-483S

6 Selvarani Elahi, Peter Farnell, Kevin Thurlow, Claudio Scotti, Alan Varnam, Referee Analysis of Probiotic Food Supplements. Food Control, 2008, **19**, 925-929

7 Food Standards Agency, "UK List of Health Claims" http://www.food.gov.uk/scotland/regsscotland/ull/claims/

8 U.S. National Library of Medicine (NLM), National Center for Biotechnology Information (NCBI) at the National Institutes of Health (NIH), 2010, http://www.ncbi.nlm.nih.gov/pubmed/

9 "Methods for the Official Control of Probiotics used as Feed Additives" published under the European Commission, Community Research project as Report EUR 20873/1 (/2 /3) EN - ISBN 92-894-6249-3 (Set) by the Office for Official Publications of the European Communities, 2002

10 U.S. National Library of Medicine (NLM), National Center for Biotechnology Information (NCBI) at the National Institutes of Health (NIH), 2010 http://www.ncbi.nlm.nih.gov/genbank/

11 S.Sheu, W.Hwang, H. Chen, Y. Chlang, and H. Tsen, Development and Use of tuf Gene–Based Primers for the Multiplex PCR Detection of *Lactobacillus acidophilus*, *Lactobacillus casei* Group, *Lactobacillus delbrueckii*, and *Bifidobacterium longum* in Commercial Dairy Products. Journal of Food Protection, 2009, **72 (1)**, 93-100

12 M.C. Collado, S. Delgado, A. Maldonado and J.M. Rodrı´guez, Assessment of the Bacterial Diversity of Breast Milk of Healthy Women by Quantitative Real-Time PCR. Letters in Applied Microbiology, 2009, **48**, 523-528

13 Basic Local Alignment Search Tool (BLAST), U.S. National Library of Medicine (NLM), National Center for Biotechnology Information (NCBI) at the National Institutes of Health (NIH), 2010, http://blast.ncbi.nlm.nih.gov/Blast.cgi

14 M.A. Nadkarni, F. E. Martin, N. A. Jacques and N.Hunter, Determination of Bacterial Load by Real-Time PCR using a Broad-range (universal) Probe and Primers Set. Microbiology 2002, **148**, 257-266

15 T. Matsuki, K. Watanabe, J.Fujimoto, Y. Miyamoto, T. Takada, K.Matsumoto, H.Oyaizu, and R. Tanaka, Development of 16S rRNA-Gene-Targeted Group-Specific Primers for the Detection and Identification of Predominant Bacteria in Human Faeces. Applied and Environmental Microbiology, 2002, **68 (11)** 5445-5451.

16 H. G. H. J. Heilig, E. G. Zoetendal, E. E. Vaughan, P. Marteau, A. D. L. Akkermans and W. M. de Vos, Molecular Diversity of *Lactobacillus* spp. and Other Lactic Acid Bacteria in the Human Intestine as Determined by Specific Amplification of 16S Ribosomal DNA. Applied and Environmental Microbiology, 2002, **68 (1)** 114–123

17 N. Witt, G. Rodger, J. Vandesompele, V. Benes, A. Zumla, G. A. Rook and J. F. Huggett, An Assessment of Air as a Source of DNA Contamination Encountered When Performing PCR. Journal of Biomolecular Techniques, 2009, **20**, 236–240

18 K. A. Harris and J. C. Hartley, Development of Broad-Range 16S rDNA PCR for use in the Routine Diagnostic Clinical Microbiology Service. Journal of Medical Microbiology, 2003, **52**, 685-691

19 P. S. M. Yeung, M. E. Sanders, C. L. Kitts, R. Cano and P. S. Tong, Species-Specific Identification of Commercial Probiotic Strains. Journal of Dairy Science, 2002, **85**, 1039-1051

20 F. Sanger, S. Nicklen, and A. R. Coulson, DNA Sequencing with Chain-Terminating Inhibitors. Proceedings of the National Academy of Sciences USA, 1977, **74(12)**, 5463-5467