Journal of Forensic Research
ISSN: 2157-7145

Masson et al., J Forensic Res 2011, S2
DOI: 10.4172/2157-7145.S2-002

Special Issue Open Access

Molecular Forensic Approaches to the Taxonomic Assessment of Bacteria in a Commercial Consortia

Luke Masson[1]*, Thomas A. Edge[2], Christine Maynard[1], Roland Brousseau[1], Nathalie Fortin[1], Charles W. Greer[1] and Jack T. Trevors[3]

[1]*Environmental Sector, Biotechnology Research Institute, NRC, 6100 Royalmount Avenue, Montreal, Quebec, H4P 2R2, Canada*
[2]*National Water Research Institute, Environment Canada, 867 Lakeshore Rd., Burlington, Ontario, L7R 4A6, Canada*
[3]*School of Environmental Sciences, University of Guelph, Guelph, Ontario, N1G 2W1, Canada*

Abstract

The characterization of low complexity (only a few species present) bacterial communities or commercial consortia products in terms of microbial composition can require a variety of molecular techniques for supporting forensic investigations. We examined a low complexity commercial consortium productfor water treatment application as a model for a tiered molecular approach to studying microbial communities. PCR amplification of 16S rDNA and cpn60 genes were performed on total genomic DNA extracted from the consortium. First, the PCR amplicons were cloned, sequenced and subjected to both DGGE and RFLP analysis, or they were fluorescently labeled and hybridized to a dual backbone taxonomic DNA microarray. Secondly, total genomic DNA from the commercial consortium was subjected to quantitative PCR to determine the concentration of the different components.

The data showed that the dual backbone DNA microarray is extremely useful as a first step to identify the major members of the consortium, including lot-to-lot variation of the commercial product, as validated by independent analyses. More importantly, the DNA microarray proved to be a useful screening tool to detect unexpected and potentially pathogenic microbes in the commercial product. This tiered approach using a DNA microarray screen can be a useful guide for application of more rapid and targeted molecular tools in forensic investigations of microbial communities.

Introduction

A serious difficulty facing both federal regulatory agencies and those involved in environmental risk assessments is the lack of information concerning the specific composition of commercial microbial bioproducts or consortia. This is partially due to the lack of standardized methodologies for characterizing microbial communities or commercial microbial consortia even though there are numerous biochemical, microbiological and molecular biology techniques [1]. Knowledge of the compositionof commercial microbial consortia is critical to address quality control issues related to product efficacy as it relates to variations in lot -to -lot production and accidental introduction of unwanted microbial contaminants including pathogens. Applications of molecular techniques need to be investigated for better characterizing microbial communities, whether for supporting forensic enforcement investigations to characterize commercial consortia products, or for better understanding the composition of complex microbial communities in nature.

Although numerous bacterial identification methodologies exist, both classical and molecular based, most have inherent limitations and are used to examine either single species (cultivating) or several species using immunological (e.g. flow cytometry, ELISA), nucleic acid (e.g. PCR) or biosensor-based approaches. For bacterial communities of lower complexity, molecular techniques like DGGE or T-RFLP on taxonomic amplicons can provide a rapid partial index of biodiversity using total extracted DNA [2]. Moreover, subsequent sequencing of 16S rRNA gene amplicons from gels can also provide a certain measure of bacterial identification. Although laborious and time consuming, small subunit rRNAgene or metagenomic sequencing remains one of most accurate means of determining microbial community content and diversity.

Microarray technology offers another means to assess complex communities due to its parallel processing power of thousands of immobilized taxonomic probes. Although numerous publications have appeared over the last decade utilizing the 16S rRNA small subunit or other regions for taxonomic identification[3,4], it has been demonstrated that DNA microarrays typically show a broader diversity of species than 16S rRNA cloning analysis [5]. Due to its conserved nature and large existing databases, array probes designed using 16S rRNA alone are ideal to identify the microbial composition of a complex community at higher taxonomic levels. However, when used with DNA microarrays, these probes lose resolution when applied to lower taxonomic levels (i.e., species/strains). This inherent limitation also affects the rapid identification of specific pathogens that could contaminate production lots of commercial microbial consortia. To improve this resolution, we have developed a taxonomic microarray that uses a dual-backbone approach for microbial identification. The backbone terminology stems from the phylogenetic backbone tree produced by the Ribosomal Database Project. Release 8, which produced a tree of SSU sequences and created 217 branches [6], allowed us to design representative probes from the type species from each branch to create a chip providing maximum coverage of prokaryotic sequence diversity. To increase the resolution of the chip, the majority of the representative type strains also had a cpn60 probe included, hence the dual-backbone name. Inclusion of this heat shock gene permits the extraction of a greater amount of phylogenetic

***Corresponding author:** Luke Masson, Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Ave, Montreal, Quebec, Canada, H4P 2R2, Tel: 1-514-496-3123; E-mail: Luke.Masson@nrc-cnrc.gc.ca

Received July 07, 2011; **Accepted** July 22, 2011; **Published** July 27, 2011

Citation: Masson L, Edge TA, Maynard C, Brousseau R, Fortin N, et al. (2011) Molecular Forensic Approaches to the Taxonomic Assessment of Bacteria in a Commercial Consortia. J Forensic Res S2:001. doi:10.4172/2157-7145.S2-001

information when compared to 16S rRNA particularly when applied to a group of highly similar species [7,8]. Our current dual backbone DNA microarray uses both 16S rRNA and cpn60 to identify over 600 bacterial species among which can be found representative members of all the different 16SrRNA backbone branches as well as a number of key bacterial pathogens.

The objective of this research was to assess the utility of this dual backbone array to rapidly screen a microbial consortium to address two key issues: 1) the identification of the genus/species composition in a commercial bioproduct; and 2) to assess lot to lot variability of microbial composition which includes the potential introduction of potentially pathogenic bacteria. The data generated by the array was compared to and confirmed by a variety of conventional molecular techniques including DGGE and RFLP mapping, quantitative–PCR and the cloning and sequencing of 16S rRNA and cpn60 amplicons.

Our results showed that the taxonomic dual-backbone DNA microarray provides a good initial pre-screening of microbial consortia of low to medium complexity by providing a rapid assessment of the major microbial components in the mixture and the detection of the presence of potential pathogens. The approach described here should find broad applicability in microbial forensic investigations to characterize commercial consortia products, as well as investigations of a range of other potential microbial communities.

Materials and Methods

Microbial consortia origin

The low complexity consortium utilized in this study was a dry packaged commercial water treatment product used to treat wastewater, and not to be used with potable water. The product was purchased as a powdered microbial formulation in sealed plastic containers. For anonymity purposes, this material was labelled as Product A. Two lots were examined that possessed different expiry dates about one year apart. All DNA extractions were completed on both products one month before the first expiry date came into effect (this older lot was subsequently named 'expired'). The newer lot of Product A was called 'new'. The product label and manufacturer's website indicated that the microbial contents were various Bacillus species.

DNA extraction

Total consortium DNA was recovered using the methodology developed by Fortin et al. [9]. In this method, each EDTA wash was performed once and any proteins and/or humic materials were precipitated with ammonium acetate. An RNAse A treatment and phenol/chloroform/isoamyl alcohol precipitation were also incorporated. To prevent external contamination of the container's contents, caution was taken to open and remove samples in a laminar flow biosafety cabinet using aseptic techniques. To ensure uniformity of sampling, the contents of the container were stirred with a sterile spatula.

Denaturing Gradient Gel Electrophoresis (DGGE) and DNA banding pattern analysis

DGGE was performed as described elsewhere [10] using the U341/U803 primer set [11,12,13] complementary to conserved regions of the 16S rRNA gene to amplify a 464 base pair (bp) fragment corresponding to positions 341 to 803 in the E. coli sequence (Table 1).About 550 ng of 16S rRNA PCR product was applied and electrophoresed on an 8% polyacrylamide gel containing a gradient of 40%–65% denaturant (7 mol·l^{-1} urea and 40% deionized formamide were 100% denaturant). DGGE was performed with a DCode Universal Mutation Detection System (Bio-Rad, Mississauga, Ontario, Canada). Electrophoresis was performed at a constant voltage of 80 V for 16 h at 60°C in 1x TAE running buffer. Nucleic acids were visualized by staining with SYBR Gold Nucleic Acid Gel Stain (Molecular Probes) and photographed with the FluoroImager System Model 595 (Molecular Dynamics, Sunnyvale, California).The total number of bands in each lane and the percentage of the gel gradient that it took to cover all the bands were analyzed with Gel Compar II version 4.6 (Applied Maths, Sint-Martens-Latem, Belgium). A band was defined as "present" if its peak intensity was at least 3% of the most intense bandin the lane. Also, the relative band intensity was estimated using a ChemiImager (Alpha Innotech, San Leandro, California).

DNA microarrays

The custom designed taxonomic DNA microarray, called GHI 10, contained 700 16S rRNA probes and 669 cpn60probes printed on Corning GAPSII slides in duplicate. Consequently, in this 'dual backbone' chip, there are 567 species or strains for which both 16S rRNA and cpn60 are represented. The list of bacterial species/strains and their associated designed 16S rRNA or cpn60probes for the dual-backbone chip is listed in Table S1 in supplementary data.A large number of the surface immobilized probes have been validated in other studies [3,14].

Amplicon generation and DNA labelling

To generate fluorescently labeled 16S rRNA and cpn60 amplicons from different lots of Product A, 100ng of total extracted consortium DNA was added in the master mix as template. Amplification of the 16SrRNA or cpn60 regions was performed using the 16S rRNA (Dorschand Stackebrandt, 1992) or cpn60 PCR primers [15] listed in Table 1. The PCR reaction mixture included 5 µl of 10 X PCR buffer (100 mM Tris-HCl, pH 9.0, 15 mM $MgCl_2$ and 500 mM KCl), 0.5 µl of 20 mM dNTPs, 1 µl of each of the forward and reverse primers (stock concentration: 25 µM), 0.5 µl (2.5 units) of Taq DNA polymerase (GE Healthcare), with sterile distilled water added to a 50 µl final volume. For cpn60 amplification, 1 µl of 100 mM $MgCl_2$ solution was also added to the master mixture.

The annealing temperatures used during amplification were 52°C for 16S rRNA and 60°C for cpn60 primers. Amplifications were performed in a Mastercycle EPgradient System (Eppendorf) according to the following scheme: Hot start: 5 min at 94°C, 40 cycles of amplification: 30 sec at 94°C, annealing temperature 30 sec at 65°C, 45 sec at 72°C, final extension: 7 min at 72°C. The length of the generated amplicons was 528 bp for 16S rRNA and 555 bp for cpn60. Amplicons were purified with the QIA quick PCR purification Kit (QIAGEN Inc., Ontario, Canada) according to the manufacturer's instructions before being labeled.

One µg of purified amplicon was chemically labeled with a Mirus Cy5 Label IT® Nucleic Acid Labeling kit (Mirus, Madison, USA) according to the manufacturer's instructions. The unreacted reagents were removed using a QIA quick PCR purification kit (QIAGEN Inc., Ontario, Canada). Quantification of fluorescent cyanine dye incorporation was done by scanning the DNA sample from 200 to 700 nm and subsequently inputting the data into the Internet-based Percent Incorporation Calculator found at http://www.pangloss.com/seidel/Protocols/percent_inc.html

Hybridization of labeled DNA

Microarrays were prehybridized at 50°C for 1 hour under a

Lifterslip (25 x 60 mm; Erie Scientific Company, Portsmouth, NH, USA) using a SlideBooster hybridization workstation (model SB800; Advalytix, Germany), with 50 µl of pre warmed (37°C) digoxigenin (DIG) Easy Hyb Buffer (Roche Diagnostics, Laval, Quebec, Canada) supplemented with 5% (vol/vol) bovine serum albumin (1 mg/ml; New England Biolabs Inc., Beverly, MA). After pre-hybridization, the lifter slip was removed by dipping the slides in 0.1X SSC (saline-sodium citrate) and were air-dried. Before hybridization, the samples were dried and resuspended in 15 µl of hybridization buffer DIG and denatured for five minutes at 95°C. One microgram of labeled genomic DNA was hybridized on the taxonomic microarray under a lifter slip (18 x 18 mm). The hybridization was carried out overnight at 50°C in a Slide Booster hybridization workstation. After hybridization, lifter slips were removed by dipping the slides in a 0.1X SSC and 0.1% SDS (sodium dodecyl sulfate) solution. Post-hybridization washes were performed at 37°C: two washes with 0.1X SSC and 0.1%,SDS for ten and five minutes, respectively and one last wash in 0.1X SSC for five minutes. The microarrays were then air-dried.

Signal acquisition and analysis

The dried microarrays were scanned with a Scan Array Lite fluorescent microarray analysis system (Perkin-Elmer, Mississauga, Ontario, Canada). Acquisition of fluorescent spots was performed using the Scan Array Express software (Perkin-Elmer, Foster City, CA).All the microarrays were normalized using the same method. Fluorescent spot intensities were quantified using Quant Array software version 3.0 (Canberra-Packard) after normalizing the data by subtracting local background from the recorded spot intensities from arrays on the same slide. The median value of each set of duplicate spotted probes was compared to the median of the buffer spotsand probes that had a signal-to-noise fluorescence ratio greater than greater than the established threshold (log2 ratio = 4 in this study), were considered positive.

Cloning and sequencing of 16S rRNA and cpn60 genes

For 16S rRNAamplification, universal primers F1 and R2 (Table 1) were used to generate a PCR fragment of ~520 bp containing the V1,V2, and V6 variable regions of the 16S rRNA gene. The final concentration of components in the PCR reaction mix was as follows:250 µM dNTP's, 0.4 µM of each primer, 2.5 Units of Taq polymerase, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 10 ng of purified consortium DNA extract in a total volume of 100µl. PCR reactions were done in an ABI thermocycler for 35 cycles (94°C for 30 sec; 54°C for 30 sec; and 72°C for 30 sec), followed by a single elongation at 70°C for 7 minutes.For cpn60 amplification, PCR conditions were the same as for 16SrRNA, except for the different primers and an annealing temperature of 50°C. The universal cpn60 primers H279 and H280 were combined with 10 ng (1.0 µl) of template DNA to generate a PCR fragment of about 600 bp. Sequencing reactions were resolved on a 3730XL DNA Analyzer systems (Genome Québec Innovation Centre Sequencing Platform). 16S rRNAand cpn60 sequences were analyzed by BLAST against Genbank and cpn60 sequences were also compared to a cpn60 database [8] using FASTA and BLASTp. Sequences were also checked for the presence of putative 16S rRNA gene chimeras using the Bellerophon server [16]. Amplicon cloning was performed by ligating overnight at 16°C in a final volume of 10 µl containing 50 ng of linearized pDrive cloning vector (QIAGEN) and a 10x molar ratio of insert (200ng). After electroporation, amp-resistance colonies that were lac+ were selected. The pDrive vector contains a T7 and SP6 promoter on either side of the cloning site, allowing sequencing using standard sequencing primers.

RFLP mapping

Restriction fragment length polymorphism (RFLP) analyses were done on isolated 16S rRNA recombinant colonies cloned from both Product A lots. The 16S rRNA gene was amplified with the F1-R2 oligonucleotide primers listed in Table 1 to produce a 520 bp amplicon [17] using PCR conditions identical to those described above for cpn60 cloning with the exception of the primers and template DNA. Several PCR reactions from each lot were pooled and precipitated overnight at -20°C with 1/10 volume of 3M sodium acetate pH 5.2, and 2.5 volumes of 100% ethanol. The DNA was centrifuged then washed with 70 % ethanol. The dry pellet was resuspended in 100 µl of TE (10 mM Tris-HCl, 0.1 mM EDTA pH 8). The excess primers, dNTPs were removed using Montage PCR centrifugal filters (Millipore, Bedford, MA, USA). The DNA was eluted with 20 µl of distilled deionized water. All recombinant plasmids were digested with the restriction enzymes AluI and HaeIII (new lot) or AluIand Sau3A (expired lot).The reactions were incubated for 3 hours at 37°C after which the restricted fragments were resolved on a 4% (w/v) Nusieve (FMC, Rockland, ME, USA) agarose gel in TBE buffer. The DNA fragments were visualized by staining with ethidium bromide.

Quantitative real time Polymerase Chain Reaction (Q-PCR)

With the exception of Enterococcus faecium, specific cpn60 primers were designed against the cpn60 genes sequenced from the different isolates in this study namely B. licheniformis, B. cereus/thuringiensis/anthracis group, B. subtilis, and Acinetobacterbaumannii.The primers used for E. faecium were designed by Cheng et al. [18]. All primer sets amplified a100 to 150 bp PCR product. All Q-PCR primers had a melting temperature between 48-50°C and were initially tested on known genomic DNA to test their specificity. Primer optimization was performed using four primer concentrations (0.2 to 0.8 µM) for both forward and reverse primers with only one primer concentration (0.6 µM) for the lower threshold cycle value (Ct) selected for all experiments. DNA master SYBR Green I Mix (including Taq DNA polymerase, dNTP, $MgCl_2$ and SYBR Green I dye, Roche Molecular Biometricals) was used to quantify the cpn60 relative concentration which was calculated as a percentage of total extracted consortium DNA. Each 20 µl reaction contained 10 µl of master mix, 0.6 µM of each specific primer (forward and reverse), 2.6 µl of RNase-free water and one of three different concentrations of consortium DNA (i.e. 5.0, 0.5 or 0.05 ng in 5 µl of water).

Q-PCR was performed with the Rotor-Gene™ 3000-A system (Corbett Life Science) and a final reaction mix containing 1X QuantiTect SYBR green PCR mastermix (QIAgen), 0.6 µM of each specific primer (forward and reverse), 5 µl of diluted plasmid standards or consortia DNAto a final volume of 20 µl. Amplification conditions were 95°C for 15 min, followed by 40 cycles of 10 s at 95°C, 15 s at 64°C (B. licheniformis), 45°C (B. cereus/thuringiensis/anthracis group and B. subtilis), 52°C (A.baumannii), or 50°C (E. faecium), and 15 sec at 72°C or 68°C (B. cereus/thuringiensis/anthracis group).

Standard curves for the assays were developed using a 10-fold dilution series of plasmids containing cpn60 sequence representatives of each target clones. Plasmid copy number was calculated using plasmid molecular weight and plasmid concentration measured using NanoDrop® ND-1000 UV-Vis Spectrophotometer (NanoDrop Technologies, Inc.). The data was expressed by number of target bacteria genomes or ng of target bacteria DNA per ng of total extracted Product A DNA. A negative control (no DNA) was performed for each run to check the quality and specificity of the amplification. In addition, melting curves analyses were done after each run to determine the

annealing temperature of every product. There were no nonspecific products or primer-dimer formation after 45 cycles of amplification and efficiency of the reaction for each specific product was calculated using the Rotor-Gene 3000 software.

Results

Microarray assessment of consortium DNA

By using a taxonomic DNA microarray based primarily on representatives of the different Eubacterial rRNA phylogenic branches, an assessment or fingerprint of the spectrum of potential microbial groups within different bacterial communities can be made. To determine potential lot to lot variation of a commercial microbial consortium or the introduction of potentially pathogenic species, total DNA was extracted from two different lots of Product A, as shown in Figure 1. In both lots, the universal bacterial probe (EUB338) confirmed that both products were bacterial in nature. At the 16S rRNA level, both products appeared similar as the general Bacillus sp probe was positive indicating the presence of one or more Bacillus species. Unexpectedly however, two Enterococcal 16S rRNA (E. saccharolyticus and E. faecium) probes were positive for the expired lot of Product A Figure 1A whereas only the E. saccharolyticus probe was positive in the new lot of Product A (Figure 1B). (just slightly above the threshold cut-off value of 4).

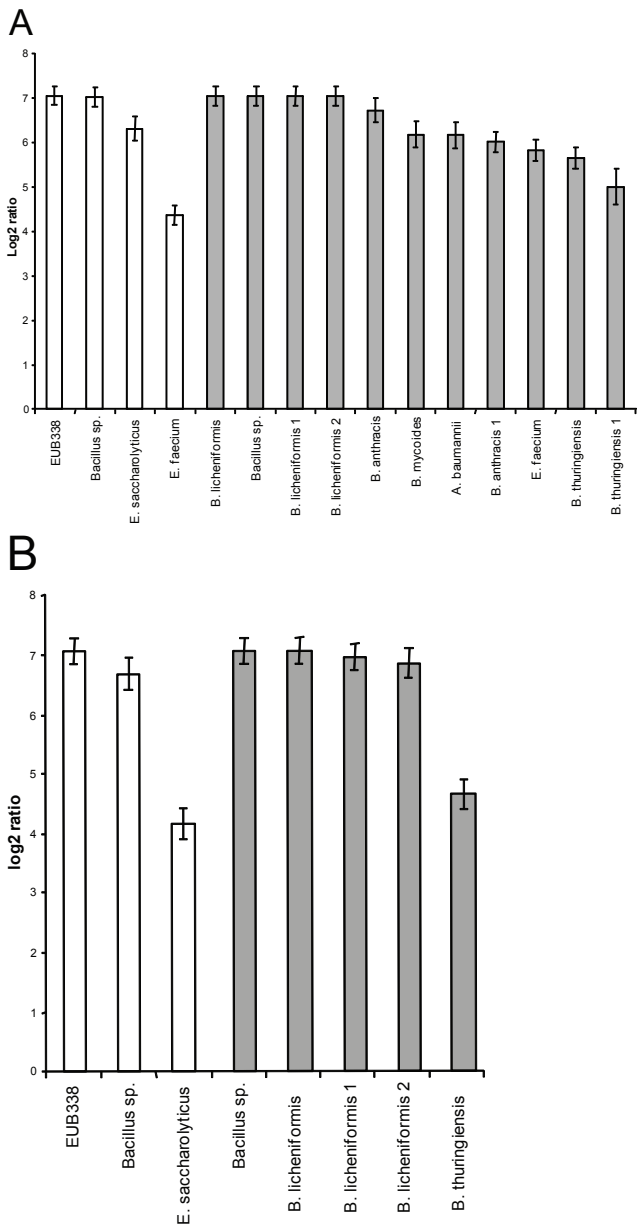

Figure 1: Taxonomic microarray analyses of different Product A lots. Fluorescent intensity ratios were plotted for both the 16S rRNA (white bars) and the *cpn60* (gray bars) hybridizations. The results represent the average of three independent experiments with each probe present on each array as two separate spots.The standard errors bars are shown. Panel A represents amplicons generated from the expired Product A and Panel B represents amplicons generated from the new Product A. Numbers at the end of the species name represents an independent species-specific probe.

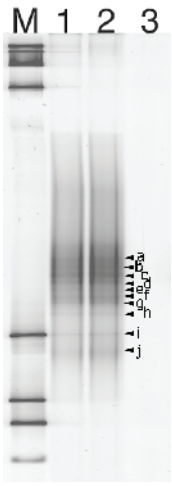

Figure 2: DGGE analysis of Product A. M, marker 100 bp; lane 1, Product A (new) (400 ng); lane 2, Product A (new) (600 ng); lane 3, empty; M, marker 100bp.Bands that were extracted for sequencing are labelled a to j.

Examination of the cpn60 positive probes, which presents a higher taxonomic resolution than the genus-level 16S rRNA probes, also showed variation between the expired product and a new lot of the same consortium. The expired lot showed the presence of B. licheniformis and a member or members from the B. cereus family. It is noteworthy that members of this highly similar family, in this case B. anthracis, B. thuringiensis and B. mycoides, cannot be distinguished sufficiently to produce discriminating microarray probes. It was surprising that no enterococcal cpn60 probe was positive in the new lot. In contrast however, the Enterococcus faecium, and Acinetobacter baumannii cpn60 probes were positive in the expired lot. These two organisms are potential pathogens and known causes of nosocomial infections in hospitals [19,20]. It is noteworthy that the Acinetobacter 16S rRNA probe was not positive in either lot, however, our 16S rRNA probe was designed using the Acinetobacter radio resistans 16S rRNA sequence which introduced a 1 base mismatch against the A. baumannii 16S rRNA probe. Overall, the microarray hybridization data showedlot to lot variation in the Bacillus spp present in the consortium, as well ascontamination with potential pathogens in one of the consortium lots.

DGGE analyses

When faced with an unknown mixture, one can obtain information on the level of microbial complexity by using DGGE to assess the

Primer	Primer sequences 5'-3'	Target	Reference
16S F1 16S R2	GAGTTTGATCCTGGCTCAG GWATTACCGCGGCKGCTG	16S rRNA gene	Dorsch and Stackebrandt 1992
H279 H280	GAIIIIGCIGGIGAYGGIACIACIAC YKIYKITCICCRAAICCIGGIGCYTT	*cpn60*	Goh et al. 2000
U341 U803	[a]GC clamp-CCTACGGGAGGCAGCAG CTACCAGGGTATCTAATCC	16S rRNA gene	Muyzer et al, 2003 Baker et al, 2003 Van de Peer et al, 1996
Ef-F Ef-R	GAGGCAGACCAGATTGACG ACTAGGTGTGGAACGGATG	*E. faecium* specific DNA	Cheng et al. 1997
Bl-f Bl-R	CGTTGCTTCTGATTGCTGAAGACG CGTTGAATGTTCCGCGAAGCTTGT	*Bacillus licheniformis* *cpn60*	This study
Bc-f BcR	CAAGAAATCTTACCAGTATTA TTCACTACTAATGTAGCTAACG	*Bacillus cereus cpn60*	This study
Bs-f Bs-R	GACAATCCTTACATCTTAATCAC AACAAGTGTTGCAAGTGCTT	*Bacillus subtilis cpn60*	This study
Brev-f Brev-R	CTGGAGCAAGTTGTACAAAGC CAGCAACAGCTGTGAAGGTA	*Brevibacillus choshinensis* *cpn60*	This study
Ab-F Ab-R	CATTCGTGAATTGATTTCTGT CATGTTGTTTACAACAAGAGT	*Acinetobacter baumannii cpn60*	This study

[a]GC clamp sequence: GCGGGCGGGGCGGGGGCACGGGGGGCGCGGCGGGCGGGGCGGGGG

Table 1: Primer sequences.

Gel sample	Fragment	Similarity (%)	0rganism	Accession #
Product A (new)	a,b,c,d,e,f,	100	*B. licheniformis*	CP000002
	g	99	*B. licheniformis*	CP000002
	h	99	*B. licheniformis*	CP000002
	i, j	99	*B. subtilis*	AB188212

Table 2: Taxonomic identification by 16S rRNA gene sequences from extracted DGGE gel bands.

RFLP [a] Pattern	% of total # of clones	% Similarity	Organism	Accession number
1	9	99.5	*Acinetobacter baumannii*	CP001172
2	18	98	*Bacillus cereus*	CP000227
3	6	99	*Bacillus licheniformis*	EF685205
4	18	96.6	*Enterococcus faecium*	GQ405286

[a]A total of 40 clones were subjected to RFLP analysis

Table 3: RFLP patterns generated with the *AluI* and *Sau3A* restriction enzymes from the expired Product A clone library and percentage of recombinant plasmids in each category.

RFLP[a] Pattern	% of total # of clones	% similarity	Organism	Accession number
1	16	99	*Bacillus licheniformis*	AB188216
2	27	99	*Bacillus licheniformis*	AB188216
3	2	99	*Bacillus licheniformis*	AB188216
4	17	99	*Bacillus subtilis*	AB188212
5	30	99	*Bacillus subtilis*	AB188212
6	1	99	*Bacillus subtilis*	AB188212
7	1	99	*Bacillus subtilis*	AB188212

[a]A total of 118 clones were subjected to RFLP analysis.

Table 4: RFLP patterns generated with the *AluI* and *HaeIII* restriction enzymes from the new Product A clone library and percentage of recombinant plasmids in each category.

Organism	# clones sequenced (expired lot)[a]	PCR (expired lot)[b]	# clones sequenced (new lot)	PCR (new lot)
Acinetobacter baumannii CP001172	5	+	0	-
Bacillus cereus CP000227	3	+	2	+
Bacillus licheniformis EF685205	15	+	58	+
Enterococcus faecium GQ405286	94	+	0	-
Brevibacillus choshinensis AB038650	0	+	2	+
Bacillus subtilis	0	-	38	+

[a]117 cpn60 isolates from the expired Product A and 100 *cpn60* isolates from the new lot were sequenced. Bacterial sequences were determined by BLAST analysis using NCBI or against the cpn60 database (http://www.cpndb.ca/cpnDB/home.php) (Hill et al, 2004)
[b]PCR represents a positive signal against total extracted consortia DNA using cpn60 primers specific for that species

Table 5: Consortia composition by *cpn60* isolate sequencing.

Bacterial species[a]	Weight of one genome (g)	Product A (expired)		Product A (new)	
		Avg. number genome copies/ng total DNA	% total DNA	Avg. number genome copies/ng total DNA	% total DNA
Bacillus cereus	5.93E-15	1243.1 (57.8)	0.65	179.7 (8.3)	0.12
Bacillus licheniformis	4.63E-15	184965.2 (8688.9)	97.17	106260.2 (5390.8)	72.98
Bacillus subtilis	4.62E-15	16.8 (2.4)	0.008	38840.9 (1672.8)	26.68
Brevibacillus choshinensis	5.00E-15	186.7 (7.3)	0.10	58.5 (3.6)	0.04
Acinetobacter baumannii	3.94E-15	1251.4 (79.5)	0.66	5.5 (0.8)	0.003
Enterococcus faecium	3.53E-15	2686.5 (238.2)	1.41	248.1 (22.7)	0.17

[a]The extracts were tested by PCR using common eukaryotic and yeast probes but were negative

Table 6: Relative number of bacteria in extracted consortia DNA.

number of different DNA amplicons having same size by separating them on the basis of their nucleotide sequence differences. Asshown in Figure 2, DGGE products from two different concentrations of Product A (new) were amplified and electrophoresed. A second higher concentration was used to ensure no weak yet specific bands would appear. About 10 bands were observed showing a low to medium level of complexity. The DNA bands were extracted from the gel, sequenced, and the results presented in Table 2 showed a very low level of species complexity. Most of the bands were variants of B. licheniformis while two other bands represented B. subtilis thus providing independent confirmation of both species occurring in the new lot of Product A.

Microbial screening by cloning and RFLP analyses

To assess the consortium microbial content using the 16SrRNA taxonomic marker, we amplified and cloned the 16S rRNA amplicons and subjected them to RFLP analysis. As shown in Table 3, the expired lot of Product A showed the presence of four 16S rRNA RFLP groups (after the removal of chimeric amplicons) corresponding to the four different bacterial species found by the DNA microarray with the exception of the Acinetobacter isolate.

A set of cloned 16S rRNA amplicons derived from the new lot of Product A was also subjected to RFLP analysis. As shown in Table 4, seven different RFLP patterns were determined. Surprisingly, these seven patterns only corresponded to two species, B. licheniformis and B. subtilis, which were the only two species determined by DGGE analysis. The multiple patterns assigned to each species may have been due to a better discriminatory capacity of the restriction enzyme combination AluI/HaeIII over AluI/Sau3A.

Microbial screening by cpn60cloning and sequencing

As the DNA microarray and RFLP data verified lot to lot variation, a cloning and sequencing approach was adopted to assess the microbial content using cpn60as a higher resolution taxonomic marker. About 100 isolates from cpn60 clone libraries obtained from either Product A lot were sequenced and the results presented in Table 5. The presence of A. baumannii and E. faecium was confirmed in the expired lot but not in the new lot. The surprisingly high number of E. faecium isolates in the expired lot is presumably explained by PCR biased amplification of the DNA extracted from this lot. B. subtilis was only detected in the new lot and not the expired lot confirming the 16S rRNA sequence data that there is a different composition of Bacillus species between the two lots. In addition, a new species, Brevibacillus choshinensis was found only in the new lot of Product A. Using the cpn60 sequences, specific PCR primers were designed for each of the six bacteria listed in Table 5. With the exception of B. choshinensis, the PCR amplification patterns followed the presence or absence of the cloned cpn60 isolates. In the case of B. choshinensis, a positive PCR was found in the expired lot of Product A suggesting that, although it was not cloned, it is present in this lot.

Quantification of microbial composition

A Q-PCR assay was set up based on the cpn60 gene from all six bacterial species found among the various molecular approaches in this study. Quantification of the individual species was performed on DNA extracted from both lots of Product A and the results are presented in Table 6. Q-PCR analysis confirmed that all six bacterial species were present in both lots, however it is clear that B. licheniformis is a major component of both lots while B. subtilis is a major component only in the new lot comprising ~27% of the total extracted DNA. In the expired lot, < 0.7% of the total extracted DNA was composed of B. subtilis. E. faecium was present at a slightly higher level than B. subtilis (1.4%) in the same lot. The Q-PCR results also confirmed that E. faecium and A. baumanii were present in both lots however, the low number of genome copies in the extracted DNA explains the lack of detection in the new lot of Product A. However, their presence in both lots and their high levels in the expired lot strongly suggest that they do not represent a minor contaminant introduced after fermentation or mixing.

Discussion

This study developed a tiered approach using several molecular forensic techniques to characterize the composition of a commercial microbial consortium product. This approach enabled a screening level assessment of the composition of the microbial product, followed by more detailed assessments using molecular techniques like Q-PCR to enumerate specific microorganisms. We hypothesized that multiple approaches and the data obtained may reveal significant differences between methods, including detection limits. The approach was applied to a relatively low species-complexity commercial product as a model for investigating the composition of microbial communities. The initial screening level exploited the parallel processing power of DNA microarrays. Although larger taxonomic chips have been developed and used commercially, the custom chip in this study served to show that a smaller, less expensive, targeted chip could suffice especially if the investigator was targeting a limited number of organisms. Moreover, the data also illustrates the utility of using a second marker in the form of a protein-encoding housekeeping gene like cpn60 to increase the taxonomic resolution of the chip data beyond the capabilities of the standard 16S rDNA gene as well as providing an added level of certainty when both taxonomic markers are positive.

The positive cpn60 signals for B. anthracis was not alarming as this simply meant that the associated isolate(s) are members of the B. cereus groupsensu latto. This means that the isolate could be B. anthracis, B. cereus, B. mycoides, B. pseudomycoides, B. thuringiensis or B. weihenstephanensis due to the high genetic similarities between

these organisms. For example, BLAST analysis of the cpn60 sequence of one of the sensu latto group (we used B. pseudomycoides) brings up all of the members of the sensu latto group within about 6% distance. For example, the cpn60 sequence for B. pseudomycoides DSM12442 is only 0.2% distant from B. mycoides ATCC31101 and 0.4% distant form B. thuringiensis ATCC10206. Sequence alignments clearly indicate that DNA microarray hybridizations using immobilized 50-mer oligonucleotides are unable to discriminate between these amplicons at this level of similarity. This did not apply to the unexpected appearance of a Brevibacillus species as its lack of detection by the DNA microarray could have been that a specific cpn60 probe was not present on the array. However, since this species was also not observed by cpn60 cloning in the different lots, it suggests that it was a minor component in the consortium which was later confirmed by Q-PCR.

One observation arising from our approach is that due to the vast biodiversity inherent in the microbial world and the vastly differential concentrations of specific microorganisms within a mixture, there is no single all-encompassing methodology to determine the make-up of a microbial community. For example, we did not see B. subtilis in the DNA microarray in either the new or expired Product A yet it was observed in the DGGE, Q-PCR and the amplicon cloning and sequencing assays. In fact, B. subtilis was observed by its specific cpn60 probe in the new Product A lot, however, the signal intensity was slightly below the cutoff value and was ignored (data not shown). Clearly, a better B. subtilis probe should be designed for the chip in order to increase the signal. Alternatively, since the Q-PCR results showed a substantial amount of B. subtilis in this lot, it is possible that PCR bias occurred in the original amplification of the cpn60 signal and that the level of B. subtilis cpn60 amplification remained low [21].

The RFLP approach was the least helpful characterization methodology. Like DGGE, it is a useful tool for determining the level of biodiversity/complexity of a microbial population and perhaps assessing gross changes in the population, however, specific information from these approaches requires the sequencing of the amplicons. Nonetheless it did provide support for the compositions determined by the DNA microarray.

In addition to being able to identity the major components of the commercial consortium at the genus level on our microarray through the use of 16S rDNA probes and the species level by using cpn60 probes, we also found two foreign species, E.faecium and A.baumannii in the expired lot. Although Q-PCR detected both of them in the new lot of Product A, their numbers were too low for the other methods to detect. Their presence was of particular concern sincevarious reports of multi-drug resistant nosocomial infections have been caused by either A. baumannii or E. faecium [22,19]. Both pathogens seemed to be present in surprisingly high concentrations in the expired bacterial consortium suggesting that their presence was probably not due to external contamination by poor aseptic technique by personnel and perhaps more due to contamination during the cell culturing or subsequent processing (formulation).

Although our tiered approach was applied to a lower complexity microbial community, the initial screening by DNA microarrays should still be appropriate for larger, more diverse communities thus narrowing down the number of targets to quantify by either conventional microbiology or by a molecular technique like Q-PCR.

1) **Conclusions and future perspective**Our validated taxonomic, dual backbone microarray provides an excellent first level screening of an unknown bacterial consortium to provide crucial information for the deployment of subsequent quantitative methods like Q-PCR.
2) The Q-PCR approach provides an excellent means of quantifying the levels of bacteria in consortia in order to identify the major bacterial species.
3) Lottolot variation in the Product A consortium was observed by both DNA microarray and cpn60analysis (sequencing and Q-PCR).
4) The presence of potentially pathogenic contaminants in one consortium lot was detected by the DNA microarray and confirmed by Q-PCR.
5) 16S rRNA gene analysis alone or combined with RFLP mapping does not possess sufficient discriminatory powers to accurately identify consortium bacteria down to the species level, however, it provides a useful complement to cpn60analysis due to its larger sequence databank.
6) The identity of the B. cereus group member should be explored further as B. cereusis considered a human pathogen. EitherB. anthracis or B. thuringiensis can be rapidly verified by PCR analysis for the presence of toxin genes.

In addition to testing microbial formulations, molecular-based microbial forensics can be used for bioterrorism analysis and unknown suspected contaminated samples such as dried formulations, liquid formulations, water, soil, sediment, biosolids and food samples. Until agreed upon, standardized molecular methods are incorporated into guidelines and protocols used by regulatory agencies we believe that the tiered approach used in this study is logical and efficient.

Acknowledgements

This article is dedicated to Dr. Roland Brousseau who during the course of our collaborative research was taken from his loved ones and colleagues by a sudden illness. Roland was foremost our trusted friend, and through his knowledge and passion for science was a scholar, educator and honorable gentleman to all of us. We will always be grateful for his friendship, knowledge and all he has taught us. This research was supported by Environment Canada's CRSB program for biotechnology. Research by JTT is also supported by NSERC (Canada).

References

1. Spiegelman D, Whissell G, Greer CW (2005) A survey of the methods for the characterization of microbial consortia and communities. Can J Microbiol 51: 355-386.
2. Nowak E, Brousseau R, Garrett J, Masson L, Maynard C, et al. (2008) Characterization of formulated microbial products by DGGE, total cellular fatty acid analysis and DNA-microarray analysis. Can J Microbiol 54: 380-390.
3. Dubois JW, Hill S, England LS, Edge T, Masson L, et al. (2004) The development of a DNA microarray-based assay for the characterization of commercially formulated microbial products. J Microbiol Methods 58: 251-262.
4. Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, et al. (2010) Deep-Sea oil plume enriches indigenous oil-degrading bacteria. Science 330: 204-208.
5. DeSantis TZ, Brodie EL, Moberg JP, Zubieta IX, Piceno YM, et al. (2007) High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. Micro Ecol 53: 371–383.
6. Maidak BL, Cole JR, Lilburn TG, Parker CT Jr, Saxman PR, et al. (2001) The RDP-II (Ribosomal Database Project). Nucl Acids Res 29:173-174.
7. Goh SH, Facklam RR, Chang M, Hill JE, Tyrrell GJ, et al. (2000) Identification of *Enterococcus* species and phenotypically similar *Lactococcus* and *Vagococcus* species by reverse checkerboard hybridization to chaperonin 60 gene sequences. J Clin Microbiol 38: 3953-3959.
8. Hill JE, Penny SL, Crowell KG, Goh SH, Hemmingsen SM (2004) cpnDB: A Chaperonin Sequence Database. Genome Res 14: 1669-1675.
9. Fortin N, Beaumier D, Lee K, Greer CW (2004) Soil washing improves the recovery of total community DNA from polluted and high organic content sediments. J Microbiol Methods 56: 181-191.

10. Yeung CW, Woo M, Lee K, Greer CW (2011)Characterization of the bacterial communitystructure of Sydney Tar Ponds sediment Can J Microbiol 57: 493–503.

11. Muyzer G, de Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol 59: 695-700.

12. Van de Peer Y, Chapelle S, De Wachter R (1996) A quantitative map of nucleotide substitution rates in bacterial rRNA. Nucl Acids Res 24: 3381-3391.

13. Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-specific 16S primers. J Microbiol Methods 55: 541-555.

14. Maynard C, Berthiaume F, Lemarchand K, Harel J, Payment P, et al. (2005) Water-borne pathogen detection using oligonucleotide-based microarrays. Appl Environ Microbiol 71: 8548-8557.

15. Hill JE, Seipp RP, Betts M, Hawkins L, Van Kessel AG, et al. (2002) Extensive profiling of a complex microbial community by high-throughput sequencing. Appl Environ Microbiol 68: 3055-3066.

16. Huber T, Faulkner G, Hugenholtz P (2004) Bellerophon; a program to detect chimeric sequences in multiple sequence alignments. Bioinformatics 20: 2317-2319.

17. Dorsch M, Stackebrandt E (1992) Some modifications in the procedure of direct sequencing of PCR amplified 16S rDNA. J Microbiol Methods 16: 271-279.

18. Cheng S, McCleskey FK, Gress MJ, Petroziello JM, Liu R, et al. (1997) A PCR assay for identification of *Enterococcus faecium*. J Clin Microbiol 35: 1248-1250.

19. Hsueh PR, Teng LJ, Chen CY, Chen WH, Yu CJ, et al. (2002) Pandrug-Resistant *Acinetobacter baumannii* Causing Nosocomial Infections in a University Hospital, Taiwan. Emerg Infect Dis 8: 827-832.

20. Giridhara Upadhyaya PM, Ravikumar KL, Umapathy BL (2009) Review of virulence factors of *Enterococcus*: an emerging nosocomial pathogen. Indian J Med Microbiol 27: 301–305.

21. Ishii K, Fukui M (2001) Optimization of annealing temperature to reduce bias caused by a primer mismatch in multitemplate PCR.Appl Environ Microbiol 67: 3753–3755.

22. Afzal-Shah M, Livermore DM (1998) Worldwide emergence of carbapenem-resistant *Acinetobacter* spp. J Antimicrob Chemother 41: 576–577.