DEVELOPMENT OF MACROARRAY TECHNOLOGY TO PROFILE BACTERIAL COMPOSITION OF INTESTINAL COMMUNITIES

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

in the Department of Animal and Poultry Science

University of Saskatchewan

Saskatoon

By Angela Dawn Goldfinch

PERMISSION TO USE

In presenting this thesis in partial fulfillment to the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Animal and Poultry Science

University of Saskatchewan

Saskatoon, Saskatchewan, S7N 5A8

ABSTRACT

The gastrointestinal tract is colonized by an abundant and diverse community of microorganisms which has a profound impact on the health of the host. The profiling of these microbial communities with traditional culture-based methods identifies only a fraction of microbes present with limited specificity, high labour costs and limited sample throughput. To overcome these limitations, a molecular hybridization assay was developed and characterized using the target gene chaperonin 60 (*cpn*60). The interspecies discriminatory ability of the hybridization assay was determined by hybridizing *cpn*60 gene fragments from a known species to a series of *cpn*60 gene fragments derived from related species with distinct but similar *cpn*60 sequences. Species with less than 85% *cpn*60 sequence identity to the probe DNA were effectively distinguished using the hybridization approach. To characterize complex microbial communities, universal PCR primers were used to amplify a fragment of 549-567 nucleotides from *cpn*60 (the *cpn*60 universal target (UT)) using template DNA extracted from the ileal contents of pigs fed diets based on corn (C), barley (B), or wheat (W), or from plasmids containing the *cpn*60 UT selected from a clone library generated from these contents. The intensity of hybridization signals generated using labelled probes prepared from library clones designated B1 (Bacillales-related), S1 (*Streptococcus*-related), C1 (Clostridiales-related), and L10 (*Lactobacillales*-related) and targets prepared from ileal contents of C, W, or B-fed pigs correlated closely with the number of genomes of each bacterial group as determined by quantitative PCR. Universal PCR primers were also used to amplify genomic DNA extracted from jejeunal contents of pre- and post-weaning piglets. Labelled probe DNA was prepared

from S1, L10, LV (*Lactobacillus vaginalis*-related) and EC (*E.coli*) library clones. The resulting signal intensities correlated with quantitative polymerase chain reaction (qPCR) data for L10 and LV, but minimal correlation was observed for the S1 and EC groups. A *cpn*60-based macroarray has potential as a tool for identification and semi-quantification of shifts in colonization abundance of bacteria in complex communities, providing a similar amount of data as techniques such as denaturation gradient gel electrophoresis or terminal restriction fragment length polymorphism analysis.

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge my co-supervisors, Dr. Andrew van Kessel and Dr. Sean Hemmingsen, for their endless support and advice throughout the duration of this project. I would also like to extend sincere thanks to the other members of my committee – Dr. Tim Dumonceaux and Dr. Fiona Buchanan for their support and suggestions during both the research and writing phases of this project. I would also like to thank Dr. van Kessel for his patience during all the times I had to show him results that didn't work. Thank you to everyone who has made my working environment so enjoyable over the past few years – Ben, Laura, Seth, Darryl, Gita, JP, Logan, Kelsey, Lisa, Jason, Charlotte, and Blair. Thanks for the memories and for being my friends as well as my colleagues over the past few years.

I would also like to extend heartfelt gratitude to my family – my Mom and Dad, Brenda and Wayne, brothers, Chris and Kelly, and sister-in-law, Cindy. Without their constant love and support this work would never have been completed. Thanks to the three felines in my life – Kyla, Ivy, and Duchess – for all the helpful input provided by walking and/or sitting on the keyboard as I wrote. Thank you to my good friend, Julie, who has been a pillar of support for me since first-year. And thank you to Ian – who makes me laugh even at times when it feels like there is nothing to laugh about.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AP-PCR	Arbitrary-primed polymerase chain reaction
B1	Bacillales-related
BCIP	5-Bromo-4-Chloro-3'-Indolyphosphate p-Toluidine Salt
CWB	Corn, wheat, barley
*cpn*60	Chaperonin 60
CSPD	disodium 3-(4-metho xyspiro {1,2-dioxetane-3,2-(5-chloro)tricyclo [3.3.1.1^{3,7}]decan}-4-yl)phenyl phosphate
DGGE	Denaturation gradient gel electrophoresis
DIG-dUTP	digoxigenin-11-2'-deoxyuridine-5'-triphosphate
DNA	Deoxyribonucleic acid
dTTP	2'-deoxythymidine 5'-triphosphate
EC	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
FISH	Fluorescent *in situ* hybridization
ITS	Internal transcribed spacer
kDa	kilo Dalton
L10	*Lactobacillus amylovorous*-related
LV	*Lactobacillus vaginalis*-related
lytA	Autolysin encoding gene
ml	Milliliter
µl	Microliter
NBT	Nitro-Blue Tetrazolium Chloride
NCBI	National Center for Biotechnology Information
ng	Nanogram
nifH	dinitrogenase reductase gene
NIH	National Institute of Health
NLM	National Library of Medicine
nM	Nanomolar
nt	Nucleotide
PCR	Polymerase chain reaction
pg	Picogram
PVPP	Polyvinylpolypyrrolidone
qPCR	Quantitative polymerase chain reaction
recA	Recombinase A
rpoB	RNA polymerase beta subunit
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
S1	*Streptococcus*-related
SSC	Sodium chloride-sodium citrate

SDS	Sodium dodecyl sulfate
SA	*Streptococcus alactolyticus*-related
TGGE	Temperature gradient gel electrophoresis
T-RFLP	Terminal-Restriction fragment length polymorphism
Tris	trishydroxymethylaminomethane
UT	Universal target

1.0 Introduction

Throughout life, animals are colonized by dynamic and complex arrays of microorganisms (Hooper & Gordon, 2001). While we are in contact with diverse microbial populations life long, exceptionally little is known about the influence of microbiota on normal development and survival (Hooper *et al.*, 2001). Remarkably, even the taxonomic composition of rich microbial communities is poorly defined. Ecological studies are difficult as there is a shortage of appropriate methods that adequately characterize community components (Hill *et al.*, 2002). In the past, the majority of information available regarding microbial community composition has been derived from culture-based techniques such as selective enrichments, species isolation, and number estimates (Raskin *et al.*, 1997). However, only a small proportion of the microorganisms present in complex communities can be observed using these methods in a laboratory setting. Estimates of the exact fraction of culturable organisms in microbial communities varies from as little as 0.001 to 40% depending on the community in question (Amann *et al.*, 1995; Hill *et al.*, 2002; Staley & Konopka, 1985).

Development of molecular microbial characterization methods has allowed new insight into microbial ecology while alleviating the limits surrounding culture based analysis. Essentially, molecular characterization methods are based on the comparison of the nucleic acid sequences of specific genetic targets within the microbial genome (Hill *et al.*, 2002; O'Sullivan, 2000) often relying on polymerase chain reaction (PCR) for gene isolation and amplification. Traditionally, the most popular gene target for molecular characterization has been the 16S

rRNA gene (Olsen *et al.*, 1986; Pace *et al.*, 1986). An alternative gene target for microbial characterization is the chaperonin 60 gene (*cpn*60, also known as *hsp*60 and *groEL*) which codes for a 60kDa protein found in all eubacteria and in the mitochondria and chloroplasts of eukaryotic cells (Gupta, 1995). Universal PCR primers have been designed to amplify an approximately 600 base pair (bp) region of *cpn*60, termed the universal target (UT) which contains conserved regions flanking a highly variable region. The *cpn*60 UT has been established as a molecular diagnostic target for the identification and classification of isolated bacterial species (Brousseau *et al.*, 2001; Goh *et al.*, 1996; Goh *et al.*, 1997; Goh *et al.*, 1998; Goh *et al.*, 2000). For profiling of complex microbial community composition, high throughput sequencing has been employed to generate large *cpn*60 UT sequence libraries prepared by universal primer amplification of extracted genomic DNA (Hill *et al.*, 2002). While techniques such as sequence library construction used alone or in combination with quantitative PCR (qPCR) (Berndt *et al.*, 1995; Countway & Caron, 2006; Ritalahti *et al.*, 2006) allow in depth investigation into the components of microbial communities, they are costly and time consuming. The approach is therefore not suitable for community analysis of large numbers of environmental samples, which would be required for assessment of community dynamics. qPCR provides an inclusive look at complex communities, but is limited to analysis of specific targets that must be determined *a priori*. A method to provide an overview of community components on an individual basis in a timely and cost-effective fashion is required.

2.0 Literature Review

2.1 Methods for Characterization of Microbial Community Composition

2.1.1 Culture-dependent techniques

One of the most basic requirements for any ecological study is the ability to identify accurately the members of the community of interest. For the microbial ecologist, this has proven to be a difficult hurdle to overcome. Historically, most of the knowledge relating to microbial communities has been obtained via culture based methods (Raskin *et al.*, 1997). These methods typically include culturing organisms, followed by enumeration and characterization with various biochemical and physiological tests (Fingegold & Sutter, 1978; Holdeman & Moore, 1972; Moore & Holdeman, 1974; Moore *et al.*, 1978; Tannock, 2002). Non-selective culture methods involve plating fresh samples of microbial communities, obtained under appropriate atmospheric conditions to preserve viability, onto a non-selective media such as brain-heart infusion agar in order to estimate total numbers of aerobic and anaerobic organisms present when cultured under the appropriate atmospheric condition (O'Sullivan, 2000). Although non-selective media are designed to permit growth of most bacteria (example; luria broth), they are known to select against some species present in many microbial communities. These species may be selected against as a result of competition with more abundant species, or the requirement for specific growth conditions that are not met. Conversely, numerous types of selective growth media are available which vary in nutrient composition and/or additives

such as antibiotics and are able to select for specific species under appropriate temperature and atmospheric conditions, facilitating their enumeration (e.g., Beerens agar is reported as selective for growth of *Bifidobacteria*) (O'Sullivan, 2000). However, total toxicity to other species as well as total selectivity for the species in question may render counts inaccurate (Apajalahti *et al.*, 2003; Hartemink & Rombouts, 1999; Rada *et al.*, 1999).

These methods have provided the vast majority of our knowledge regarding microbial communities to date and their importance should not be underestimated. Techniques such as culture-based enumeration still contribute important data regarding the proportions of the components of complex microflora and estimations of bacterial viability. Although the contributions of culture-based techniques are great, the techniques are laborious and time consuming (O'Sullivan, 2000; Tannock, 2002). In addition, only a small proportion of the microbes in complex communities are culturable due to unknown growth requirements or uncharacterized interactions with the host or other microbes, thereby imposing a serious limitation on the scope of these traditional methods (Amann *et al.*, 1990; Giovannoni *et al.*, 1988; O'Sullivan, 2000; Tannock, 2002; Zoetendal *et al.*, 1998). The exact fraction of culturable organisms in microbial communities is estimated to vary from as little as 0.001% to 40% (Amann *et al.*, 1995; Hill *et al.*, 2002; Staley & Konopka, 1985). In addition, characterization of microorganisms accomplished with traditional methods is often not sufficient for definitive classification (Woese, 1987). Similar biochemical properties found in organisms of different families may disguise both genetic similarities and diversity, thereby making species and strains difficult to distinguish (Stahl *et al.*, 1988). In order to avoid all of the complications inherent to culture-based methods, a number of culture-independent methods have been developed.

2.1.2 Culture-independent characterization

The development of culture-independent methods for microbial characterization has addressed many of the gaps left by characterization with culture-dependent methods. Classic culture-independent methods garner valuable information about the true number of bacteria present in a population, although they may be of less value when attempting to characterize specific organisms (O'Sullivan, 2000). For example, direct microscopic analysis of faecal samples using a light microscope will provide an estimate of total number of microbial cells. However, this technique relies on heat fixation and cell staining, during which many cells may be lost resulting in an inaccurate estimate of cells present.

A second classic culture-independent method is the analysis of enzymes and metabolites produced by a microbial community as an indication of the presence of specific microflora. As an example, measurement of short chain fatty acids (SCFA) – specifically propionate and butyrate – has been correlated to the presence of certain bacterial metabolism occurring in the intestinal tract (Apajalahti *et al.*, 2003; Blank *et al.*, 2001; Cabrita *et al.*, 2003; Mathew *et al.*, 1998; Siiger *et al.*, 1996). As increased SCFA has been associated with beneficial bacterial colonization in the gastrointestinal tract (Liong & Shah, 2005; Sakata *et al.*, 2003), changing levels are often used as an indicator of probiotic colonization in monogastric animals. While this is not a direct enumeration technique, metabolite analysis can be used as an indirect method for identifying microbial groups present in a complex community.

While classic culture-independent techniques have provided information regarding microbial communities that is inaccessible using traditional culture-dependent methods, the advent of molecular characterization methods has produced an even greater understanding. Molecular characterization methods generally consist of detection or comparison of the nucleic

acid sequence of specific target genes. When dealing with environments where identification of an individual known species is important, such as pathogen identification in clinical applications, the nucleic acid sequence for a target gene unique to the bacterium is often desirable. An example of this is the identification of *Streptococcus pneumoniae* as the causative microbe in a central nervous system infection via the presence of the *S. pneumoniae* autolysin-encoding gene (*lytA*) (Taha & Olcen, 2004). When bacterial identification is required where there is limited prior knowledge of species identity, analysis of the nucleic acid sequence of universally present target genes can be used. Universal target genes containing conserved regions flanking regions of high sequence variability are ideal (Tannock, 2002), as the comparison of regions of high variability allows the identification of different phylogenetic groups or even species, while the areas of low variability help to facilitate ease of analysis. The selection of a proper target gene is crucial to the efficiency and overall success of molecular characterization methods.

2.2 Target Gene Selection for Molecular-based Community Characterization

Perhaps the most important component of molecular microbial analysis is the selection of a proper target gene. If identification of a known individual bacterial species is of interest, the target gene must be unique to the species of interest (Taha & Olcen, 2004). If the identification of all constituents in a microbial community is desired, or if the identification of unidentified community isolates is desired, the target gene in question must be present in all bacterial groups or at least in a subgroup for which there is specific interest (e.g.) the dinitrogenase reductase gene (*nifH*) – a gene present in the nitrogen fixation metabolic pathway. In addition, the potential target will ideally contain conserved regions interspersed with regions of hypervariability (Tannock, 2002). The conserved regions allow for the development of universal PCR primers, an important asset in the evaluation of a potential target gene by allowing one set

of PCR primers to amplify the gene in question from all species. The regions of variability allow for genus or species-specific identification of bacteria via the comparison of sequences obtained with universal PCR primers. The level of variability is important as it is the key factor in determining the species of organism according to target gene sequence. A lack of variability will result in a lack of specific identification, while too much variability means the identified organisms will be unable to be grouped according to phylogeny. The size of the target gene is also important – large target genes are more difficult to sequence completely in a high-throughput manner and thus are less useful than genes that consist of shorter nucleotide sequence. However, if a target gene has a nucleotide sequence that is too short, it will not yield enough discriminatory power. Finally, the ideal target gene is present in only a single copy in the bacterial genome – an important attribute for quantification purposes. That is, a single copy allows for more accurate quantification as it represents only a single organism. A gene that is present in multiple copies may not allow accurate quantification as each organism contains a variable number of genetic copies. In addition, mutations in multiple copy genes can affect some but not all copies present in a single organism, thereby affecting quantification results.

2.2.1 16S rRNA

There are three ribosomal RNA (rRNA) genes present in bacteria which are identified by their relative sizes – 5S, 16S, and 23S. As ribosomes are important components of the cell responsible for protein synthesis, the rRNAs have been preserved in all organisms. Historically, both the 5S and 16S rRNA genes were popular targets for molecular microbial profiling (Olsen *et al.*, 1986). The small size of 5S rRNA allowed for complete sequence analysis to be done by the late 1960's (Olsen *et al.*, 1986), however a lack of sequence variability limited its usefulness. 16S rRNA was also used for microbial identification studies, but its larger size limited its

usefulness until the advent of more modern DNA cloning and sequencing techniques. Today, 16S rRNA is the most popular target for molecular profiling methods (Hopkins *et al.*, 2001; La *et al.*, 2003; Matsuki *et al.*, 2002; Olsen *et al.*, 1986; Pace *et al.*, 1986; Salzman *et al.*, 2002; Suau *et al.*, 1999). As it meets the majority of criteria presented above, the 16S rRNA gene is a good target for molecular identification methods (Olsen *et al.*, 1986; Pace *et al.*, 1986). 16S rRNA is universally present and functionally constant in all bacterial species (Coenye & Vandamme, 2003; Olsen *et al.*, 1986; Pace *et al.*, 1986; Vandamme *et al.*, 1996), as are all of the rRNA genes. The 16S rRNA gene contains regions of both conservation and variability, however the presence of several often evolutionarily diverged copies of the 16S rRNA gene in many microbes has proven to be a major drawback of using 16S rRNA as a target (Goh *et al.*, 1996). Multiple copies of a gene potentially provide inaccuracy when quantification is desired in addition to identification, as it may be difficult to determine how many copies of a gene are present per organism (Bach *et al.*, 2002). Finally, 16S rRNA genes may not contain sufficient sequence variation to provide identification to the species level (Clayton *et al.*, 1995; Coenye & Vandamme, 2003; Fox *et al.*, 1992; Goh *et al.*, 1996). It has also been shown that strains of *Brevundimonas alba* have identical 16S rRNA sequences, even though DNA-DNA hybridization analysis and various PCR fingerprinting analysis indicated they were physiologically very different (Jaspers & Overmann, 2004).

2.2.2 Internal transcribed spacer

Because the 16S rRNA does not contain sufficient sequence variation to allow for species-specific identification for some genera, alternative targets potentially offering more variation for species identification have been investigated. One alternative is the analysis of a second molecular target in addition to analysis of the 16S rRNA gene. An example of this

methodology is the analysis of the genetic region located between the 16S and 23S rRNA genes in the bacterial genome (O'Sullivan, 2000). Known as the internal transcribed spacer (ITS), sequence data from this region has been used to identify specific *Lactobacillus* species found in various microbial communities (Tannock *et al.*, 1999). The length of the ITS in various species differs, but is normally a size that is manageable with conventional PCR (Christensen *et al.*, 2000; Tannock *et al.*, 1999). While the ITS is universally conserved in bacteria, heterogeneity within the same bacterial strain has been observed (Christensen *et al.*, 2000), making it a useful supplementary tool for microbial profiling but an unlikely candidate for a sole target gene.

2.2.3 *recA*

Another target gene that may be used as an alternative to16S rRNA is the recombinase A (*recA*) gene. The RecA protein appears to have many functions, being involved in DNA repair and the SOS response, as well as homologous recombination (Yang *et al.*, 2001). *recA* is considered to be universally present in both prokaryotic and eukaryotic cells with a high degree of sequence conservation. Past studies have used *recA* sequence analysis to identify bacterial species in the *Erwinia* genus via PCR-RFLP (Waleron *et al.*, 2002). Ramette *et al.* (2005) used *recA* sequence data in conjunction with 16S rRNA analysis to identify and enumerate the species *Burkholderia cepacia* in the general environment. Also, Kullen *et al.* (1997) used *recA* sequence analysis to identify six species from the genus *Bifidobacterium* from isolates taken from the human intestinal tract. Phylogenetic analysis using *recA* compared favourably with phylogenetic analysis using 16S rRNA from the same samples.

2.2.4 *rpoB*

A third alternative to 16S rRNA as a target gene is the gene for RNA polymerase beta subunit (*rpoB*) (Dahllof *et al.*, 2000). Like both 16S rRNA and *recA*, *rpoB* is a good target gene as it is universally present and contains regions that are conserved in some areas and hypervariable in others. Also, *rpoB* is present in a single copy in the bacterial genome, increasing its utility for enumeration techniques (Dahllof *et al.*, 2000). Previous work by Dahllof *et al*. (2000) has shown that molecular profiling of the microbial community found on the surface of a marine rock using *rpoB* as a target gene was more accurate than analysis using 16S rRNA – predominately due to the presence of multiple copies of 16S rRNA in the bacterial genome. *rpoB* has also been used to profile microbial communities found in tropical soils (Peixoto *et al.*, 2002) and on the surface of cold-smoked salmon (Giacomazzi *et al.*, 2004).

2.2.5 Nitrogen reductase

In the selection of a target gene, it is sometimes desirable to focus in on a select group of microorganisms. An example of this is the use of nitrogen reductase genes as a microbial profiling target for nitrogen-fixing organisms. These organisms are central players in the nitrogen cycle, which is responsible for the reduction of nitrate to important gaseous nitrogen compounds used in several sectors such as agriculture, waste treatment, and environmental preservation (Cheneby *et al.*, 2003). Previous work regarding molecular profiling for nitrogen-reducing organisms has focused on the *narGHJI* operon (Cheneby *et al.*, 2003) or the *nifH* gene (Jenkins *et al.*, 2004; Tan *et al.*, 2003). These gene targets provide a high amount of phylogenetic sequence diversity; however this has resulted in highly degenerate universal primers. While the use of these target genes does exclusively select for the profile of nitrogen-

fixing organisms, the possibility exists that unknown organisms have an alternative nitrogen-reducing enzyme and are thus being excluded from study.

2.2.6 *cpn*60

A fifth alternative target for molecular profiling methods is the chaperonin 60 gene (*cpn*60*, hsp60, groEL*), which encodes a 60 kDa protein present in essentially all eubacteria and in the mitochondria and chloroplasts of eukaryotes (Brousseau *et al.*, 2001; Goh *et al.*, 1996; Gupta, 1995). *cpn*60 has been shown to be essential for cellular growth in several types of microorganisms, as well as playing an essential role in post-translational folding and assembly and transport of the resulting polypeptide chains (Ellis & Van der Vies, 1991; Hallberg, 1990; Langer & Neupert, 1991; Rusanganwa & Grupta, 1993; Zeilstra-Ryalls *et al.*, 1991). In most bacteria, *cpn*60 exists only in a single copy in the genome (Gupta, 1995; Kong *et al.*, 1993; Picketts *et al.*, 1989; Rusanganwa & Grupta, 1993), negating ambiguity associated with multiple and potentially divergent copies as reported for 16S rRNA gene target. In the rare instances where *cpn*60 occurs in multiple copies within a single organism, there is normally a large amount of sequence divergence between the copies (Kong *et al.*, 1993; Martel *et al.*, 1990; Rusanganwa & Grupta, 1993), allowing more accurate microbial quantification. Also, analysis of the protein-encoding nucleic acid sequence of *cpn*60 appears to yield more phylogenetic information than does analysis of the structural 16S rRNA gene sequence, allowing for species specific identification of most organisms (Brousseau *et al.*, 2001; Goh *et al.*, 1996; Hill *et al.*, 2002; Hill *et al.*, 2004).

Early work with *cpn*60 as a molecular microbial identification tool revealed that *cpn*60 can be used to identify various *Staphylococcus* species and subspecies (Goh *et al.*, 1997), as well as being able to distinguish between *Enterococcus* species and phenotypically similar

Lactococcus and *Vagococcus* species (Goh *et al.*, 2000). *cpn*60 has also been used as a clinical diagnostic tool to identify *Streptococcus iniae*, a human and animal pathogen (Goh *et al.*, 1998). More recent work has revealed that *cpn*60 can be used as a target gene for high-throughput sequencing for extensive profiling of the intricate microbial community located in the pig intestine (Hill *et al.*, 2002). Further research revealed that *cpn*60 can be used to compare the microbial communities present in pig ileum according to diet treatment using a combination of high-throughput sequencing and qPCR (Dumonceaux *et al.*, 2006; Hill *et al.*, 2005). This work has demonstrated that *cpn*60 is a useful clinical diagnostic tool for bacterial isolates, and is an excellent target for molecular profiling of complex microbial communities.

2.3 Molecular profiling methods for microbial communities

2.3.1 FISH

Fluorescent in situ hybridization (FISH) is a detection method that involves the hybridization of fluorescently labelled rRNA-targeted oligonucleotide probes directly to cells that are attached to a solid glass support (O'Sullivan, 2000; Satokari *et al.*, 2003; Zoetendal *et al.*, 2004). Specialized treatment of the cell membrane renders it permeable to the small, labelled oligonucleotides. The cells that contain hybridized oligonucleotides can be observed via fluorescent microscopy. While FISH has previously been used for the identification of bacterial groups found in human feces, its use in the characterization of the gastrointestinal tract is increasing (Zoetendal *et al.*, 2004). FISH allows for the determination of several issues surrounding microbial profiling with a single method. FISH has the ability to identify individual populations and their habitats within a more complex community. The use of several probes allows for the detection of multiple populations at one time without the use of culturing methods. In addition, FISH allows for the enumeration of population constituents although it often

involves manual counting with a microscope (Wagner *et al.*, 2003; Zoetendal *et al.*, 2004). Although FISH is able to provide the answer to many questions at one time, the entire method is dependent on previous knowledge of the species in question for probe design. In addition, the permeability of the cell determines the effectiveness of the method – the membranes of some species are not easily disrupted, rendering identification and quantification values inaccurate (Zoetendal *et al.*, 2004). FISH has been coupled with flow cytometry in an attempt to increase quantification efficiency (Wagner *et al.*, 2003; Wallner *et al.*, 1997), increasing the attractiveness of this method as a profiling technique.

2.3.2 Sequence analysis of community isolates

Although limited to culturable bacteria, one method of community profiling is to non-selectively culture community samples followed by identification of selected colonies. One of the most widely accepted methods for identification of unknown bacterial isolates is through sequence analysis of a selected universal target gene, most commonly the 16S rRNA gene (Clarridge, 2004; O'Sullivan, 2001). Sequence data is obtained for all target genes from isolated colonies in essentially the same way – the basis of the technology lies in the PCR. Essentially, PCR primers are designed to anneal to regions of high sequence conservation flanking and therefore amplifying a region of variable nucleic acid sequence between the two primer sites (Clarridge, 2004; Goh *et al.*, 1996; O'Sullivan, 2001; Olsen *et al.*, 1986). Since the primers anneal at regions conserved among bacteria, they may be considered universal in that they will amplify the target from any bacterial isolate. The starting material for the PCR can be whole cells from bacterial colonies, or bacterial genomic DNA extracted from the cells using variety of methods many of which are available as commercial kits (Clarridge, 2004; O'Sullivan, 2001).

The resulting amplicon from the PCR can be sequenced and the sequence information compared with sequence data for biochemically typed bacteria to allow phylogenetic classification.

2.3.3 Sequence analysis from whole community samples

In addition to simple sequence comparison for individual isolates, techniques have also been developed for the profiling of microbial community members without culture and isolation. Construction of PCR-derived plasmid-based libraries containing the sequence of the target gene of interest can be made in order to produce an overview of the community components. Hill *et al.* (2002) have published a large study utilizing a molecular *cpn*60-based diagnostic method in order to produce libraries of *cpn*60 nucleotide sequences from a diverse range of microorganisms. Briefly, bacterial DNA was extracted from the microbial community present in fecal material obtained from pigs fed various diets. The resulting DNA was amplified with *cpn*60 universal primers and the amplicons were ligated into plasmids and cloned using an appropriate strain of *E. coli*. The resulting clones were sequenced and high quality sequence was obtained for more than 1100 of the resulting clones. The DNA sequences obtained were compared both to each other and to reference *cpn*60 sequences obtained from isolated strains of bacteria. Almost 400 unique nucleotide sequences were identified, encoding 280 unique peptide sequences (Hill *et al.*, 2002). Pair-wise comparison and phylogenetic analysis was able to determine the major taxa present in the various communities, as well as the level of sequence identity between typed strains and previously uncharacterized strains. Widmer *et al.* (1999) used a similar analysis technique targeting the *nifH* gene to characterize the diverse microbial population located in the soil and litter in a Douglas fir forest site. Bourne and Munn (2005) used a similar technique for the construction of clone libraries targeting both *rpoB* and 16S rRNA to profile diverse bacteria present in coral located on the Great Barrier Reef.

While the development of plasmid-based libraries targeting specific target genes have allowed a greater understanding of the components of complex microbial communities, aspects of the technique may inaccurately overemphasize the proportion of certain bacterial species within the communities (Green & Keller, 2006). The root of this problem lies in the use of universal PCR primers designed to anneal to a conserved area of the target gene in order to amplify a selected portion of the target gene from every species present in a complex microbial community. In studies in which 16S rRNA is the target gene, the presence of several conserved regions of DNA sequence make chimeric PCR products a possibility. A study by Acinas *et al*. (2004) focused on the production of two clone libraries analyzing a coastal bacterial community using 16S rRNA as the target gene. One clone library was constructed using universal PCR primers in a regular PCR, while the second library was constructed using an altered PCR specially designed to reduce the amount of chimeric molecules and *Taq* polymerase errors that occurred. Comparison of the number of unique sequences detected in the two libraries showed a decrease from 76% in the regular PCR library to 61% in the modified PCR library, indicating a potentially significant impact of PCR error in the estimation of microbial diversity when using 16S rRNA as a target gene.

Another contributing factor to the inaccurate representation of the diversity complex microbial communities is the inability of so-called universal primers to amplify a specific group of microorganisms. Some species of bacteria containing a high proportion of G+C residues in their genomic sequence are known to be present in microbial communities through culture-based analysis (Hill *et al.*, 2006). However, these high G+C organisms are conspicuously absent from PCR-based clone library production (Hill *et al.,* 2002). One proposed explanation for the lack of amplification is the inability of *Taq* polymerase to amplify these templates because of the higher

denaturation temperature required for a G+C rich template (Varadarajb & Skinnera, 1994). A second explanation is found in the makeup of the primers themselves. Primers designed against target genes such as *cpn*60 are often degenerate in order to maximize the number of species amplified (Goh *et al.*, 1996). As a consequence, the primers may contain inosine residues at specific locations to minimize degeneracy. However, the inclusion of inosine in universal PCR primers alters the thermodynamic stability, stacking interactions, and overall structure of the primers, potentially creating a bias against the amplification of G+C rich organisms (Hill *et al.*, 2006). Development of primers that amplify G+C rich organisms efficiently and including these primers in the PCR with regular universal primers for clone library construction may result in a more accurate representation of the diversity present in microbial communities (Hill *et al.*, 2006).

2.3.4 Reference databases and use in organism identification

Nucleotide sequence comparison is used to determine the identity and phylogenetic relationships of unidentified members of the microbial community. A picture of bacterial evolutionary relationships can be produced via comparison of recovered nucleotide sequences with the sequences of typed reference isolates maintained in computer databases (Raskin *et al.*, 1997). Typically, 97% sequence similarity of recovered nucleotide sequences of the 16S rRNA gene allows an unknown organism to be assigned to an individual species. This criterion is based on the observation that organisms with greater than 97.5% sequence similarity will most likely be identified as the same species (Leser *et al.*, 2000). However, this criterion is true for full-length target gene sequences only, as less than full length sequences may overestimate the proportional changes in both conserved or hypervariable regions (Hugenholtz *et al.*, 1998; Leser *et al.*, 2000). Bacteria that are determined to be nearly identical to typed strains are often designated as closely related typed strain-like clusters until more closely examined. Clusters of

bacteria that are highly related to each other, but relatively unrelated to known typed organisms are often denoted as being 'like' their nearest neighbour, but are treated as a separate unidentified cluster (Hill *et al.*, 2002).

A database for sequence comparison is GenBank (http://www.ncbi.nlm.nih.gov/). Maintained by the National Center for Biotechnology Information (NCBI), in conjunction with the National Institute of Health (NIH) and the National Library of Medicine (NLM), in 2005 GenBank contained sequence information for over 165 000 named organisms (Benson *et al.*, 2005). The majority of submissions received are from individual authors, while collaborations with institutes such as the US Office of Patents and Trademarks also contribute. However, investigation into the accuracy of GenBank sequences have indicated that up to 20% of species with the same name contain at least 2% sequence variability (Clarridge, 2004). These discrepancies are most likely due to inadequate or faulty phenotypic testing, or simply due to a lack of complete phenotypic understanding regarding the species in question. In order to avoid the problems produced by such a broad database, smaller databases have been developed to deal with individual target genes. For example, the Ribosomal Database Project II (http://rdp.cme.msu.edu/) has been developed for the analysis of 16S rRNA gene sequences. Because the 16S rRNA gene has traditionally been the most widely used gene target, there exists sequence data from a multitude of organisms for this gene. cpnDB (http://cpndb.cbr.nrc.ca/) is another database developed for sequence analysis, this time for the target gene *cpn*60. cpnDB now contains more than 4200 records for more than 240 identified species (Hill *et al.*, 2004) and is the largest collection of sequences for *cpn*60 in the world.

2.3.5 Quantitative Real-Time PCR

The development of real-time PCR has allowed for the development of detection and quantification assays for various microbial species. Real-time PCR is based on the periodic monitoring of the change in fluorescence as an indicator of product generation during the exponential amplification phase of the PCR (Heid *et al.*, 1996; Malinen *et al.*, 2003; Suzuki & Giovannoni, 1996). Several chemistries have been developed to permit fluorescence detection of PCR product accumulation including for example, SYBRGreen® (Molecular Probes Inc. Oregon, USA), Taqman® (Heid *et al.*, 1996) and Molecular Beacons (Tyagi & Kramer, 1996). The Taqman® and molecular beacon chemistries are advantaged over SYBRGreen® in that for both methods fluorescence emission is PCR product specific, avoiding signal detection associated with non specific double stranded product or even PCR primer dimers. SYBRGreen® is an intercalating dye that produces a fluorescent signal when bound to double stranded DNA (Malinen *et al.*, 2003). As the PCR proceeds, an increasing amount of double stranded DNA is present and the fluorescence of SYBRGreen® is also proportionally increased. While SYBRGreen® lacks the specificity of the *Taq*man system; SYBRGreen® requires only primers for the target of interest and therefore provides more flexibility in its application. Developed by Tyagi and Kramer (1996), molecular beacon detection relies on hairpin-shaped oligonucleotide probes that undergo a conformational change upon hybridization with an intended target. The hairpin probe consists of the intended oligonucleotide target sequence in the center of the hairpin, flanked by two 'arm' sequences complementary to each other. A fluorophore is covalently linked to one end of the oligonucleotide sequence, while a quenching molecule is bound to the other end. The hairpin structure of the unbound probe allows the fluorophore and quencher to remain in close proximity to one another, thus quenching possible fluorescence.

Following hybridization to the intended target, the hairpin structure unfolds, removing the quenching molecule from the vicinity of the fluorophore and fluorescence occurs. For the Taqman® method a detection probe is designed to anneal to the PCR product just downstream of either the forward or reverse PCR primer. A fluorescent dye (e.g. 6-carboxy-fluorescein) is chemically coupled to the 5' nucleotide and an appropriately matched florescence quencher (6-carboxy-tetramethyl-rhodamine) is attached to the 3' nucleotide (Ludwig & Schliefer, 2000; Malinen *et al.*, 2003a). The proximity of fluorescence emitter and quencher when the probe is intact prevents fluorescence emission. However, relying on the 5' to 3' exonuclease activity of *Taq* polymerase the Taqman® probe is degraded during primer extension since it is designed to anneal to the template just 3' to either primer. Design of probes with a lower annealing temperature than the primers ensures the probe anneals prior to primer extension. Probe degradation results in the separation of the fluorescent dye and quencher during the amplification phase of the PCR allowing the dye to fluoresce under appropriate excitation. Thus, the more specific amplified product that is produced, the greater the amount of fluorescence that is detected.

The use of primers specific for microbial species or groups in real-time PCRs has proven to be a sensitive method for quantifying specific groups of bacteria in various communities (Dumonceaux *et al.*, 2006; Fortin *et al.*, 2001; Huijsdens *et al.*, 2002; Koesters *et al.*, 2001; Nadkarni *et al.*, 2002). Known as quantitative PCR (qPCR), the technique involves the use of standardized samples that contain known numbers of genetic copies. The basis of standard curves can come from several sources. Wise and Siragusa (2005) developed a 16S rRNA-based qPCR protocol to enumerate the presence of *Clostridium perfingens* in the avian gastrointestinal tract. By spiking samples known to be lacking *C. perfingens* with known amounts of *C.*

perfingens as determined by culture enumeration in a serial dilution, a logarithmic-linear relationship with cycle threshold number was observed when analyzed by qPCR. Castillo *et al*. (2006) used PCR product of the 16S rRNA gene of *Echerishia coli* and *Lactobacillus acidophilus* to construct a qPCR standard curve to quantify bacterial populations in the porcine jejunum. Dumonceaux *et al*. (2006) examined several methods to enumerate various bacterial populations via a *cpn*60-based qPCR protocol and found that the most efficient standard curve was created with a serial dilution of plasmids containing a fragment of *cpn*60 of the species in question. Amplification of these defined standards results in the production of a reliable standard curve, allowing for extrapolation of the quantity of the bacteria in question in unknown samples.

The use of qPCR in microbial characterization has many advantages. Many samples can be analyzed at one time and results are obtained in a timely fashion with relatively low labour costs once an established protocol has been developed and targets to be analyzed have been established. However, the equipment required and reagents used are expensive. The presence of PCR inhibitors and fluorescence quenchers extracted in environmental samples may also affect accuracy (Dumonceaux *et al.*, 2006). In addition, previous knowledge of the nucleotide sequence of the bacterial species in question is required to develop functional primer and probe sets. Thus, the target organism must be selected before analysis can begin. qPCR is an excellent method when one is interested in a small number of specific groups; however, both extensive cost and the requirement for previous knowledge of target organism sequence prohibit this method from being a widely used profiling technique.

2.3.6 Multiplex and arbitrary primed PCR

When the PCR technique was first developed, it was originally intended that only a single target gene would be amplified during each reaction. However, it has been discovered that the inclusion of more than one primer pair in the PCR may result in the production of more than one PCR product from the same reaction (Markoulatos, 2002). First proposed in 1988, multiplex PCR was used to screen for causative deletion mutations in the Duchenne muscular dystrophy gene locus (Chamberlain *et al.*, 1988). Multiplex PCR has since been used as a valuable tool for the identification and profiling of several parasites, bacteria, and viruses (Chamberlain *et al.*, 1988; McIver *et al.*, 2005; Markoulatos, 2002), as well as a tool for gene screening and molecular diagnostics (Garcia *et al.*, 2005). In addition, inclusion of variable probe color and analysis of melting temperature results in the expansion of the multiplex PCR technique (Wittwer *et al.*, 2001). Combining multiplex PCR technology with qPCR could result in a higher throughput qPCR application.

The inclusion of multiple primer sets increases the probability of the amplification of non-specific amplified product, likely because of the formation of primer-dimers (Markoulatos, 2002). Variation in number of copies of target gene, secondary, genome structure, and variation in G/C content may also contribute to preferential amplification of one target gene over another during the reaction (Markoulatos, 2002).

While the inclusion of multiple primer sets requires careful monitoring of PCR conditions, arbitrary primed PCR (AP-PCR) requires much less stringent conditions. In AP-PCR, a single short primer representing an arbitrarily selected sequence is used (Hernandez *et al.*, 1997). Under non-stringent annealing temperatures and PCR conditions, the arbitrary primer creates a set of PCR amplification products unique to each species. Thus a unique fingerprint

may be obtained following electrophoresis (Hernandez *et al.*, 1997). While subtle changes in reaction conditions may result in changes in banding patterns, making reproducibility difficult (O'Sullivan, 2000), the major advantage of this method is that it can be used to characterize species for which no sequence information is known. While AP-PCR is not commonly used as a profiling tool, it has been used to identify various microbial species (Hernandez *et al.*, 1997; van Belkum, 1994).

2.3.7 DGGE and TGGE

As opposed to highly specific, quantitative molecular methods such as qPCR, a number of molecular fingerprinting techniques have been developed in order to give an overview of the components of microbial communities. A frequently used technique for this purpose is denaturing gradient gel electrophoresis (DGGE). First introduced by Muyzer *et al*. (1993), DGGE is based on the separation of PCR amplicons of the 16S rRNA in response to an increasing denaturation environment (Ercolini, 2004). Briefly, a microbial community is subjected to a DNA extraction. PCR amplicons from the variable region of the 16S rRNA gene are generated using universal primers that anneal to conserved regions of the gene. While one of the primers is normal, the other primer has a GC-rich 5' tail known as the 'GC clamp' in order to prevent complete denaturation of the amplicons (Tannock, 2000). In addition, primers specific for the 16S rRNA gene of a particular bacterial taxon may be used to examine shifts in individual populations of interest (e.g. Konstantinov *et al*. 2004) examined the response of a *Lactobacillus amylovorous*-like species in the porcine gastrointestinal tract according to prebiotic treatment). The resulting amplicons are subjected to electrophoresis with a denaturation gradient running parallel to the voltage-induced direction of electrophoresis. The differential denaturation gradient separates the PCR amplicons depending on their melting point, which is dictated by the

differential sequence amplified within the variable region of the 16S rRNA gene (Tannock, 2000). A similar technique to DGGE is temperature gradient gel electrophoresis (TGGE). While the technical aspects are virtually identical, TGGE relies on a differential temperature gradient to separate the PCR amplicons according to their sequence variation (Muyzer, 1999).

It has been shown that PCR coupled with TGGE was a successful method for comparative analysis of gastrointestinal and fecal microflora (Felske *et al.*, 2001; McCracken *et al.*, 2001; Zoetendal *et al.*, 1998). DGGE has also proven to be an effective tool for molecular profiling of microbial communities from various other communities (Favier *et al.*, 2002; Heilig *et al.*, 2002; Satokari *et al.*, 2001; Walter *et al.*, 2001). While both methods have found widespread use, they have inherent advantages and disadvantages. Both TGGE and DGGE are reproducible, rapid, and inexpensive, while allowing for the simultaneous generation of results from many samples (Muyzer, 1999). In addition, bands of interest can easily be extracted and identified via sequencing. Unlike qPCR, analysis is not dependent on a specific target organism; instead all changes in the microbiological population can be observed at once. However, generation of genetic fingerprints can be difficult with TGGE and DGGE due to a lack of separation of small DNA fragments, and the co-migration and separation of fragments of different sequences (Vallaeys *et al.*, 1997). In addition, DGGE and TGGE have a relatively high limit of detection of around $10^7 – 10^8$ colony forming units (cfu) (Collier *et al*. 2003). Thus DGGE and TGGE are rarely used as the sole method for profiling microbial communities. Instead, limitations in sensitivity are reduced by the use of real time PCR or hybridization methods in conjunction with DGGE and TGGE.

2.3.8 T-RFLP

Terminal restriction fragment length polymorphism (T-RFLP) is a molecular profiling tool based on the measurement of the size of restriction fragments from a PCR amplified marker (Marsh, 1999). Briefly, universal or taxon-specific PCR primers are designed for the amplification of the gene of interest. This target gene has traditionally been 16S rRNA due to the wealth of sequence information available, however, any suitable target gene may be used (Marsh, 1999). One of the PCR primers has a fluorescent molecule coupled to the 5' nucleotide resulting in PCR products fluorescently labelled at one terminus. The labelled PCR amplicons are cleaved with selected restriction endonucleases, resulting in the production of fragments of various lengths depending on product sequence and enzyme specificity. The fragments are separated on a polyacrylamide gel and visualized by ultraviolet excitation or an automated DNA sequencer can be used to provide a fluorescence pattern equivalent to fragment nucleotide length (Saikaly *et al.*, 2005).

The advent of T-RFLP emerged after the innovation of DGGE and TGGE. However, the use of T-RFLP is becoming more common, perhaps due to several key advantages of the technology (Marsh, 1999). Primarily, sequence data can be generated for the unique terminal restriction fragments (T-RFs), allowing reference to sequence databases. In addition, sizes of unique T-RFs are maintained in a database (http://www.trefid.net/ - Rosch and Bothe (2005). Thus, the unique T-RFs obtained from a digestion can be directly compared with the database in order to obtain phylogenetic information. Secondly, T-RFLP has greater resolution ability than DGGE and TGGE (Marsh, 1999). Lastly, the analysis of T-RFLP is immediate during gel analysis and the resulting data is digital.

T-RFLP has been used to assess bacterial diversity in several different microbial communities. Saikaly *et al*. (2005) assessed the impact of solid retention time on the bacterial diversity of activated sludge, while Hayashi *et al.* (2005) characterized the microbiota found in the human intestinal tract using T-RFLP technology.

2.3.9 G + C profiling

Not all profiling methods require PCR as a part of their technique. Total microbial communities may be analyzed using technology based on the comparison of the percent guanine (G) and cytosine (C) found within the bacterial chromosomal DNA extract (Apajalahti *et al.*, 1998). As the proportion of G and C differ between bacterial genera, analysis of the percentage can be used to create a bacterial profile.

DNA is extracted from the community of interest using a method that involves several rounds of high speed centrifugation and washing to extract maximum amounts of bacterial cells (Apajalahti *et al.*, 1998). Following lysis, the DNA is purified with several rounds of caesium cholide-ethidium bromide equilibrium-gradient centrifugation (Apajalahti *et al.*, 1998). The resulting highly purified DNA is subjected to a caesium chloride-bisbenzimidazole gradient. Bisbenzimidazole is a DNA-binding dye that specifically binds adenine (A) and thymidine (T) (Apajalahti *et al.*, 1998). Exposure to a bisbenzimidine gradient allows the fractionation of DNA dependent on its AT/GC content. When compared to gradients of known GC content, the GC content of the sample community can be estimated.

G + C analysis has been used to determine shifts in bacterial communities according to species of inoculation and intestinal location (Apajalahti *et al.*, 1998), as well as differences in diet composition (Apajalahti *et al.*, 2001).

While G + C analysis is an effective method for determining a bacterial profile from an unknown community, it provides only an overall snapshot of the bacterial community. No sequence data can be easily obtained from this analysis, so if more detailed information is required regarding the community of interest, more specific techniques must be used.

2.4 Hybridization Techniques

2.4.1 Macroarray

It has long been known that single stranded DNA will naturally find and bind to its complementary strand in a process known as hybridization. Observation of this phenomenon has resulted in the invention of DNA-based hybridization macroarrays that can be used for a variety of analyses, including gene expression assays, genetic mutation analysis, bacterial species identification, and profiling techniques for complex communities. Originally developed by Southern (1975) and Thomas (1980), macroarrays are based on the immobilization of single stranded DNA on a solid support. Denatured DNA in a liquid phase is applied to the immobilized DNA and hybridization between complimentary strands results in double stranded DNA being immobilized on the solid support. A variety of detection methods (chemiluminescence, radioactivity) can be used to detect the presence of probe DNA. Microarrays are based on the same basic principles, although they are performed on a much smaller scale and typically use fluorescence intensity as a detection method (Schena *et al.* 1995). While both are based on the same technology, hybridization macroarrays have the potential to be a powerful identification and profiling method for microbes within a microbial community without the excessive cost of microarray technology.

There are two DNA sources that must be identified for macroarray technology to work. The first is the identification of target DNA; i.e., the DNA that is immobilized on the array

membrane. Target DNA can either be PCR products or synthesized oligonucleotides. PCR product-based target DNA typically consists of longer DNA fragments than oligonucleotide-based arrays (Gabig & Wegrzyn, 2001). The increased length of the PCR product-based target DNA allows for improved species discrimination when profiling complex microbial communities as longer sequence is available for the detection of mismatched DNA. A second advantage of PCR product-based target DNA is cost of production – in-house PCR synthesis of target DNA is much less expensive than the commercial synthesis of oligonucleotide-based target DNA (Gabig & Wegrzyn, 2001; Ye *et al.*, 2001). However, the production of PCR product-based target DNA can be time consuming, especially as purification steps are required to remove contaminants such as remaining PCR components and non-specific amplification products (Ye *et al.*, 2001). In addition, it is difficult to obtain reliable PCR amplification for some bacterial species, specifically those with a high G+C content. In contrast, oligonucleotide-based target DNA does not require an amplification step in addition to having a much lower risk of external contamination resulting from the fewer steps needed for production. However, it remains very expensive to synthesize large arrays using oligonucleotide as target DNA.

The second DNA sequence required for macroarray technology is labelled probe DNA created from whole DNA extracts from a complex community of interest or individual bacterial species. Probe DNA can be labelled in a variety of ways with the most popular occurring either during the PCR or immediately following its completion (Ye *et al.*, 2001). In addition, there are a variety of substances that can be used as the labelling agent. Early experimentation with hybridization-based arrays used radioactive isotopes as the labelling agent for probe DNA (e.g. ^{32}P or ^{33}P) (Mansfield *et al.*, 1995; Ye *et al.*, 2001). However, in recent years there has been a movement towards the increased use of non-radioactive labelling methods based on marker

enzymes; due to the inconvenience of radioactive work, the majority of detection methods used in this technology today are based on non-radioactive methods. The purpose of these methods is to label the probe DNA for direct or indirect detection using colourmetric, chemiluminescent, bioluminescent, or fluorescent molecules (Mansfield *et al.*, 1995). While indirect detection is often more sensitive than direct, it is more time consuming and expensive due to the extra steps, reagents, and equipment required (Mansfield *et al.*, 1995). Colourmetric detection is the least sensitive of the non-radioactive detection methods. The method is based on the formation of coloured insoluble products that precipitate at the site of hybridization, allowing for simple visual detection. While the formation of the precipitate results in the development of a permanent record of that assay, it is difficult to remove it if the membrane is to be used again (Mansfield *et al.*, 1995). Fluorescent detection is based on the presence of molecules that absorb light at specific wavelengths and emit light at a different wavelength. Numerous assays have been developed around fluorescent detection to date and it is considered the most versatile of the non-radioactive detection techniques (Mansfield *et al.*, 1995). Chemiluminescence detection is based on the production of light as a detection reaction. One of the most sensitive and efficient methods of probe labelling and detection is accomplished with the incorporation of digoxigenin-11-2'-deoxyuridine-5'-triphosphate (DIG-dUTP) into the PCR used to produce labelled target DNA (Goh *et al.*, 1996). DIG-dUTP is included in the PCR mixture and thereby incorporated into the product in order to produce a labelled target. Detection of the presence of bound labelled targets is accomplished with chemiluminescent detection following stringency washes and the exposure of the membrane to anti-DIG antibody complexed with alkaline phosphatase in combination with the light-producing substrate disodium 3-(4-metho xyspiro {1,2-dioxetane-3,2-(5-chloro)tricyclo [3.3.1.1^{3,7}]decan}-4-yl)phenyl phosphate (CSPD). Hybridization is detected

by exposing the hybridized membrane to X-ray film. Various detection methods are available for hybridization macroarrays, including colorimetric, radioisotope, and chemiluminescent detection. A major disadvantage of chemiluminescence detection is the lack of linearity with regards to the signal emission and decay. That is, the chemilumenescent signal does not directly correlate to the amount of probe DNA bound to the membrane, making the determination of relative abundance more difficult.

The physical reaction during which labelling of the probe DNA occurs can involve many methods. For some probes (typically shorter than 400 nt), post-PCR labelling of the 3' end of the sequence with the use of a terminal transferase enzyme can be efficient (Mansfield *et al.*, 1995). While this method is efficient, it is specific for short probe DNA sequences and is more time consuming due to the post-PCR modification. Probe DNA can also be labelled during the actual PCR by the inclusion of a 5' labelled primer or the incorporation of labelled nucleotides in the reaction cocktail. However, the incorporation efficiencies of the modified nucleotides may be low and can differ between different labels thereby altering detection efficiency (Ye *et al.*, 2001).

Following the successful labelling reaction of probe DNA from the community of interest, the membrane containing the immobilized target DNA is exposed to the labelled probe. Resulting hybridization indicates the presence and relative abundance of the panel of species present on the membrane. Actual hybridization can be detected with either direct or indirect methods. In direct methods, the molecule (enzyme or radioisotope) used for the generation of the positive reaction is directly linked to the probe DNA (Mansfield *et al.*, 1995). In indirect methods, a small molecule (e.g. DIG, fluorescein, biotin) is incorporated into the probe DNA. Following hybridization, the array is exposed to the conjugates appropriate for the substance

incorporated in the probe DNA and colourmetric, chemiluminescent, or fluorescent detection ensues.

The majority of investigation and use of macroarray technology has been in clinical diagnostic area for identification of bacterial isolates. For these arrays, a panel of bacterial species is selected for application to a positively charged membrane – these species applied to the membrane are known as target DNA. Typically, universal PCR primers are used to amplify a specific target gene with the application of the resulting amplicons onto the membrane (Goh *et al.*, 1996). However, several genes or entire bacterial genomes may also be used as target DNA (Trad *et al.*, 2004). Probe DNA is constructed by labelling either amplified or total DNA from the unidentified species of interest with a radioactive (e.g.^{32}P) or non-radioactive (e.g. DIG-dUTP) labelling agent. Probe DNA is exposed to the immobilized target DNA and hybridization is allowed to occur. Under stringent conditions, the probe DNA will bind to its complimentary target DNA. The identification of the unknown probe DNA can be determined according to the identification of the target DNA at the membrane location with the positive hybridization reaction, thus providing an identification tool for the unknown species. This technology was used to identify *Streptococcus iniae* (Goh *et al.*, 1998), various *Staphylococcus* subspecies (Goh *et al.*, 1997), as well as distinguishing *Enterococcus* species from phenotypically similar *Lactococcus* and *Vagococcus* species (Goh *et al.*, 2000). Lubeck *et al.*, (2004) were able to distinguish between three subspecies of *Trichoderma*, while Trad *et al.*, (2004) identified several isolates of *Staphylococcus aureus* using this approach. Masson *et al*. (2006) combined macroarray technology with *cpn*60 and 16S rRNA target genes to design a plastic DNA array for the detection of pathogenic *Helicobacter* species in human waste sludge.

When using macroarray technology for profiling complex microbial communities, there are two major approaches. The first approach involves applying target DNA from individual bacterial species of interest, which together form a panel of species of interest immobilized in individual spots in predetermined locations on a membrane. Labelled probe DNA extracted from complex microbial communities of interest (e.g. human intestinal tract, etc.) is applied to the immobilized panel DNA present on the membrane. The location and intensity of the hybridization signal indicates the presence of those bacteria in the complex community under evaluation. This hybridization technology is quite similar to the methods used in the identification of clinical isolates used above. This technology was also used to detect the presence of several intestinal bacteria in human feces (Wang *et al.*, 2002a). The 16S rRNA sequences of 20 different bacterial species commonly found in the human intestinal tract were amplified with universal 16S rRNA primers and were labelled during the amplification reaction with DIG-dUTP. Oligonucleotide sequences specific for each of these bacterial species were synthesized and applied to a nitrocellulose membrane in pre-determined locations. Hybridization was determined using Nitro-Blue Tetrazolium Chloride/5-Bromo-4-Chloro-3'-Indolyphosphate p-Toluidine Salt (NBT/BCIP) colourmetric detection. Wang *et al.* (2002b) successfully showed that this type of macroarray development is a useful method for the profiling of common bacterial species found in the human intestinal tract, although results were recorded strictly on a comparative basis and no attempt was made to quantify the bacterial groups present in different subjects. This method of hybridization macroarray offers the advantage of providing a method by which numerous bacterial species can be detected at the same time. In contrast, PCR-based methods used for detection and quantification of various bacterial species must test for individual

bacterial species separately (Greiner *et al.*, 2001; Pahl *et al.*, 1999; Tajima *et al.*, 2001; Wang *et al.*, 2002b; Wang *et al.*, 1997) thereby posing a serious limitation to this specific methods.

Alternatively, total DNA extracts from complex microbial communities can be amplified with universal PCR primers, with the resulting universal PCR product spotted onto the membrane at predetermined locations. Exposure of these complex PCR product spots to labelled DNA probes consisting of either oligonucleotides or PCR products representing specific bacterial species can indicate the presence and level (by intentisty of hybridization signal) of each individual organism in the complex microbial communities under evaluation. Macroarrays developed in this fashion were used to identify and partially quantify microbial communities found in the waters of Chesapeake Bay (Jenkins *et al.*, 2004; Steward *et al.*, 2004). This macroarray technology focused on the target gene encoding the dinitrogenase reductase (*nifH*) enzyme. Various *nifH* fragments from complex clone libraries were amplified were spotted in predetermined locations on a nylon membrane (Steward *et al.*, 2004). Individual clones representing single bacterial species were amplified and labelled via biotinylation and applied to a nylon membrane. Detection of hybridization was accomplished using chemiluminescent detection (Jenkins *et al.*, 2004).

There are inherent problems with hybridization macroarray techniques. One commonly observed property in macroarray technology is the phenomenon of cross hybridization. Cross hybridization occurs when the labelled target binds to a probe of similar, but not identical, nucleotide sequence resulting in a false positive detection. It should be noted that when developing a molecular profiling tool, cross hybridization is likely to occur to some level due to the large amount of genetic overlap seen in complex communities. Parameters such as hybridization temperature, and temperature of stringency washes may be altered to ensure the

least amount of cross hybridization occurs (Zheng *et al.*, 1996). For example, Steward *et al*. (2004) determined that the amount of cross hybridization observed was dependent on both formamide concentration and percent similarity between the probe and target DNA sequence. High concentrations of formamide (40%) showed a decrease in cross hybridization; however, probes with ≥ 86% sequence identity to the targets indicated hybridization when using a portion of the *nifH* gene as the genetic target. This finding is backed up by Jenkins *et al*. (2004), who indicate a specificity level of 85% sequence identity over the same portion of the *nifH* gene. Chiang *et al*. (2006) also found that sequences with >85% homology between probe and target DNA may show measurable cross hybridization after developing a 16S rRNA-based DNA macroarray technique. However, Goh *et al.* (1997) used a *cpn*60-based DNA macroarray technique to show cross reaction is only noted at sequence identities ≥94%. Goh *et al*. (2000) used the same technique to identify several *Enterococcus*, *Lactococcus*, and *Vagococcus* species with a sequence identity ranging from 69 to 88%. A simple probe study by Miller *et al*. (2002) designed to investigate the occurrence of cross hybridization in DNA macroarrays indicated that the level of cross hybridization is dependent on the concentration of the labelled probe DNA. That is, if a probe DNA is present in high concentration in relation to all other DNA, it is more likely cross hybridization will be observed when the sequence identity between the probe and target DNA is ≥ 90%. Another factor affecting potential cross hybridization is the location and distribution of the mismatched nucleotide sequence over the target gene sequence. Kane *et al*. (2000) determined a level of cross hybridization of 75% sequence identity when greater than 15 bp stretches of identical DNA sequence were present over a 50 bp range between the probe and target DNA.

Another important issue surrounding DNA macroarray technology is the detection limit (sensitivity). In other words, what is the smallest amount of target DNA that can be detected with this technology. Sensitivity levels of macroarray technologies can be difficult to determine as they depend on somewhat subjective conditions. Reliable quantification of the original source of the target DNA is of absolute importance, as this is the basis for the determination of sensitivity levels. Sensitivity levels also depend on the detection method being employed, as methods such colourmetric detection have lower sensitivities than methods such as chemilumenescence detection (Mansfield *et al.*, 1995). Using radioisotope-labelled probe DNA, Malinen *et al.* (2003a) determined a detection limit of approximately 10^7 genomes of target DNA. This correlates to genomic DNA extracted from approximately 30g of faecal contents. Using a colourmetric detection method, Chiang *et al*. (2006) successfully detected PCR products amplified from 1 x 10^3 cfu/ml. Using chemiluminescent detection, Socransky *et al*. (2004) determined a sensitivity limit of 10^4 cells using probe DNA based on whole genomic DNA. Undoubtedly, there are many issues surrounding macroarray technology that must be addressed so the technique will produce valuable results.

2.4.2 Microarray

The recent development of microarray technology has created a revolution in the molecular analysis of biological systems. Originally developed for gene expression studies, the use of microarrays has expanded to detect genetic mutations in specific genes, as well as characterizing the components of complex microbial communities (Zhou, 2003). Based on the same basic principles as macroarrays, microarrays have the advantage of being able to process thousands of potential targets in a much more efficient manner.

There are two major types of microarray – oligonucleotide-based arrays and PCR product-based arrays (Ye *et al.*, 2001). PCR product-based arrays consist of previously amplified PCR amplicons usually representing a gene of interest from various bacterial species immobilized onto a predetermined area of a glass slide in a process commonly known as spotting. Common difficulties arising from this method are variation in spot morphology and high background, possibly related to batch variability observed in the production of PCR product-based slides (Ye *et al.*, 2001). Conversely, an oligonucleotide-based array consists of oligonucleotides representing a target gene of interest are constructed by spotting the oligonucleotides in predetermined locations on a glass slide. Oligonucleotide-based arrays can also be generated via the Affymetrix system, which involves the oligonucleotides representing a target gene of interest being synthesized *in situ* directly on a glass slide. Issues such as high background and cross hybridization commonly seen in PCR product-based arrays are not seen in oligonucleotide-based arrays. However, the synthesis and attachment of oligonucleotides – particularly long ones – is very expensive (Ye *et al.*, 2001).

Following construction of the array slides, probes are constructed from the community of interest. Following total DNA extraction, the probes are labelled with fluorescent markers – typically Cy5 or Cy3 – and are hybridized to the slides. Analysis of resulting signal intensities indicate the level of members of microbial samples as expressed as relative abundance. Wang *et al.* (2002a) used this technology for the detection of numerous intestinal bacterial species found in human feces.

2.5 Microbial Ecology of the Porcine Gastrointestinal Tract

Although sterile at birth, the porcine gastrointestinal tract becomes colonized with a rapid succession of various microbiota (Leser *et al.*, 2000; Robinson *et al.*, 1981; Tannock *et al.*,

1990). Tannock *et al.* (1990) used a plasmid profiling method to identify a succession of *Lactobacillus* species that colonize the pig gut during the first few weeks of life. The strains identified in pig gut samples were also detected in the feces of the sows in contact with the piglets. This indicates that maternal contact is an important method of colonization for the piglet in early life. Pedersen and Tannock (1989) suggest that the presence of *Lactobacillus* species in the digestive tract of the pig may be beneficial in early life. *Lactobacillus* species appear to contribute to the high acidity levels in the piglet gastrointestinal tract during the first days of life. This may act as an acid barrier through which potential pathogens must pass (Pedersen & Tannock, 1989).

One of the most stressful times of a piglet's life comes at weaning. This is not only important from a physiological standpoint, but from a microbiological standpoint as well. The switch from a diet made up solely of milk to one consisting of solid feed is thought to cause a dramatic shift in the microbiota population in the pig gut. Swords *et al*. (1993) published a study attempting to characterize the changes in select bacterial groups in the pig colon within the first 120 days of life. Using traditional culture-based methods, it was determined that very low counts of bacteria are present at birth. Within 12 hours, however, the colon was rapidly colonized with bacteria. These early colonizing bacteria were predominately aerobic bacteria, but by 48 hours after birth more than 90% of the bacteria present were anaerobes. During the suckling stage of life, the predominate organism identified were *Clostridium* species particularly in the early stages. During the latter stages of the pre-weaning period, *Clostridium* species began to decrease in proportional abundance, indicating the increase in other anaerobic species. Post-weaning, *Clostridium* species drastically decreased to be replaced by members of the genus *Bacteroides*.

There is relatively little detailed information available on the normal microbial ecology of the normal post-weaning pig gut. This lack of detailed information is not due to a lack of studies done on the microbial communities, but is instead a result of the shortcomings of traditional, culture-based analysis (Leser *et al.*, 2002). However, these early culture-based studies are not without value and have indicated that the majority of the bacteria present in the porcine gastrointestinal tract consists mainly of gram-positive, anaerobic *streptococci*, *lactobacilli*, *eubacteria*, and *clostridia*, while the gram-negative portion consists mainly of *Bacteroides* species (Leser *et al.*, 2002; Pedersen & Tannock, 1989; Robinson *et al.*, 1981; Tannock *et al.*, 1990).

Culture-independent methods have been used to attempt to analyze the complex microbial community found in the normal pig gut. Simpson *et al*. (1999) used DGGE as a tool to examine intestinal and fecal microflora of pigs at various stages of development. While DGGE proved to be a useful tool for monitoring changes in bacterial populations, sequencing is required to identify the bacterial groups that are changing. Detailed studies on the microbial ecology of the pig gut have been done by Pryde *et al*. (1999), Hill *et al*. (2005), Leser *et al*. (2002). Both Leser *et al*. (2002) and Pryde *et al*. (1999) used 16S rRNA-based culture-independent techniques to characterize the microflora found in the ileum, cecum, and colon of pigs from a variety of ages, diets, and overall health. Hill *et al*. (2005) used *cpn*60-based culture-independent methods to thoroughly characterize the ileal microflora of pigs fed corn, wheat, or barley. These studies indicated that majority of the flora present in the pig intestinal tract consists of low G+C gram-positive organisms, however, this may be due to biased primer selectivity during the PCR step of the cloning process. Hill *et al*. (2005) found that the ileal flora of pigs was dominated by *Lactobacillales*-like species regardless of the diet fed. *Clostridiales*-like sequences were found

to be more abundant in the corn-fed ileum, while *Streptococcus*-like sequences were more abundant in the barley-based diet. These observations were in agreement with observations made by Leser *et al*. (2002) and Pryde *et al*. (1999), although the microbiological environment observed in the pig ilea was not as complex as the one observed in colonic and fecal contents (Hill *et al.*, 2002). Molecular-based studies on environments such as the porcine gastrointestinal tract agree with culture-based studies as far as the identity of the major components, however, molecular-based methods allow for a true appreciation of the complexity that exists here.

2.5.1 Microbial contributions to pig health and performance

Fermentation within the pig gut as dictated by the activity of the microbial community there is known to be important to pig health and performance (Lalles *et al.* 2007). The main products of microbial fermentation are volatile fatty acids (VFAs), which are known to have an important role in water absorption, regulation of pH within the gut, and pathogen inhibition. VFAs are also important in stimulation of gut motility, improvement of energy yield and vitamin production, and stimulation of gut immunity (Lalles et al. 2007).

Establishment of a complex and stable microbial community within the pig gut depends largely on diet (Lalles *et al.* 2007; Knudsen 2001). The development of a healthy commensal population is imperative and during highly stressful times such as weaning, the microbial population can easily be swayed towards enteric disease (Lalles *et al.* 2007). The use of antimicrobial growth promoters (AGP) is common in the pig production industry, as one of the potential benefits is a reduction in pathogenic bacteria in the gut (Knudsen 2001). However, the continued use of AGP in swine production has come under scrutiny with increasing prevalence of antimicrobial-resistant strains of bacteria. With the use of AGP on the decline, investigation into other means of controlling microbial establishment in the pig gut has increased.

The roles through which the microbial population in the pig gastrointestinal tract affect pig health and development are not clearly defined (Knudsen 2001); however they appear to be closely linked. Investigation into the components of this complex microbial community will allow a greater understanding of the exact influence of gut microbial activity on swine health and development and may reveal alternative avenues to increasing the efficiency of swine production.

3.0 Hypothesis and Objectives

The purpose of this project was to develop and evaluate a hybridization method based on macroarray technology as a profiling tool for complex microbial communities. It is hypothesized that hybridization macroarray technology based on *cpn*60 target gene technology will be an effective profiling tool for complex microbial communities, specifically those of the gastrointestinal tract. The primary objectives of this project are to develop macroarray hybridization methodology with sufficient sensitivity (defined as the smallest amount of DNA that can be detected with this technology) and specificity (defined as the level of sequence identity that can be distinguished with this technology) to characterize complex microbial communities found in the gastrointestinal tract of animals. A method with potentially suitable characteristics will initially be established based on analysis of single species hybridization to multiple single species targets with variable sequence identity. Subsequently, suitability for complex community analysis will be assessed using samples of gastrointestinal contents. In one case, the structures of the gastrointestinal communities were previously characterized by extensive *cpn*60 UT sequence libraries and qPCR. Finally, macroarray technology will be used to characterize the complex microbiota found in the gastrointestinal tracts of pre and post-weaning piglets – an environment that has not previously been investigated with molecular profiling techniques.

4.0 Materials and Methods

4.1 Assay development and characterization

4.1.1 Preparation of target DNA from plasmid templates

Target DNA was defined as DNA that was applied and crossed-linked to the nylon membrane. Target DNA was prepared by PCR using plasmid DNA templates (pGEM-T Easy; Promega, Madison, Wisconsin) containing *cpn*60 universal target (UT) inserts corresponding to various selected bacterial species listed in Table 4-1. Selected *cpn*60 UT inserts were amplified in a PCR consisting of 20 mM Tris pH 8.4, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTP, 1 unit Taq Polymerase (Invitrogen, Burlington, ON), 500 nM of each universal *cpn*60 primer (H279 and H280 Goh *et al*. 1996) or plasmid specific primers (T7 and SP6) and ddH_2O to 50 µl volume. PCR conditions were 1 cycle of 3 minutes at 95^{o}C, then 40 cycles of 95^{o}C for 1 minute, 46^{o}C for 1 minute, and 72^{o}C for 1 minute. Target DNA PCR product concentration was determined using Quant-iT^{TM} PicoGreen$^{®}$ dsDNA reagent (Invitrogen, Burlington, ON) with a Fluroskan Ascent FL fluorometer (Thermo Labsystems, Franklin, Massachusetts). Standard curve was generated using commercial lambda DNA (Invitrogen).

Table 4-1. Species identification of plasmid-based *cpn*60 UT used for synthesis of target DNA.

Species identification	cpnDB ID[a]	Type Strain	Pairwise identity to type strain
Prevotella zoogleoformans	B6851	Yes	100%
Prevotella heparinolytica	B6844	Yes	100%
Bacteroides stercoris	B6774	Yes	100%
Bacteroides eggerthii	B6771	Yes	100%
Bacteroides caccae	B6770	Yes	100%
Bacteroides helcogenes	B6773	Yes	100%
Bacteroides vulgatus	B1030	Yes	100%
Bacteroides fragilis	B6393	No	99.8%
Citrobacter freundii	B3375	Yes	100%
Escherichia coli	B7482	No	Not available[b]
Salmonella enterica	B4831	No	Not available[b]
Enterobacter cloacae	B7411	Yes	100%

[a] http://cpndb.cbr.nrc.ca/
[b] No type strain available in cpnDB

4.1.2 Preparation of target DNA from mixed community DNA

4.1.2.1 DNA extraction

For DNA extraction from mixed microbial communities found in the gastrointestinal tract of pre- and post-weaning piglets, frozen intestinal contents from the pre-weaning jejunum were thawed and 3 aliquots of 333 mg were weighed into sterile bead beating tubes (# 12811-100-DBT, Bio/Can Scientific, Etobicoke, Ontario, Canada). For post-weaning jejunal samples, which were high in moisture content, 1g of contents was weighed and centrifuged at 19 000 x g for 10 minutes. The resulting supernatant was removed and the remaining solid fraction was transferred to a bead-beating tube. To extract total bacterial DNA, 365 µl buffer B1 (50mM Tris-HCl, 50 mM EDTA with 0.5% Tween 20 and 0.5% Triton X-100), 7.5 µl 100 mg/ml lysozyme, 20 µl 20 mg/ml proteinase K were added to the bead beating tube and samples were incubated at 37^{o}C for 45 minutes. A further 135 µl of buffer B2 (3M guadnidine HCl with 20% Tween 20) was added and samples were incubated at 50^{o}C for 30 minutes. Samples were then frozen at -70^{o}C. After thawing at room temperature, 700 µl of 25:24:1 phenol/chloroform/isoamyl alcohol was added and the samples were placed on ice. The samples were then placed in a Fast Prep bead-beating unit (Bio101 Thermo Savant) and were agitated for 1 minute. The samples were then centrifuged at 19 000 x g for 15 minutes. The aqueous phase was removed to a new tube and an equal volume of chloroform was added. Following centrifugation for 10 minutes at 19 000 x g as above, the aqueous phase was removed to a new tube and 0.1 volume 3 M sodium acetate pH 5.2 and 1.0 volume of isopropanol was added. Samples were centrifuged for 15 minutes at 19 000 x g. The resulting pellet was washed with 1

ml of 70% ethanol, dried, and redissolved in 100 µl 10mM Tris-Cl/1 mM EDTA pH 8.0 (TE buffer).

Following extraction, the DNA was purified using a polyvinylpolypyrrolidone (PVPP) purification technique (Holben *et al*. 1988). Briefly, the PVPP was prepared by suspending insoluble PVPP in 3M HCl for 12 hours at room temperature. The PVPP was then strained through Miracloth (Calbiochem Cat # 475855), suspended in 20 mM potassium phosphate (pH 7.4) and stirred at room temperature for 2 hours. This process was repeated until the pH reached 7.0, at which point the PVPP was stored in potassium phosphate. For purification of DNA, a 1 cc syringe was packed with glass wool to approximately the 1cc mark. 0.5ml of PVPP was added and the syringes were centrifuged at 250 x g for 5 minutes. The flow-through fraction was discarded and 100 µl of the extracted DNA was added to the packed syringes. Purified DNA was collected by centrifuging the syringes for 5 minutes using a 0.2 ml microfuge tube to collect the elutate.

4.1.2.2 Preparation of target DNA from complex community extracts

Target DNA was prepared using template DNA extracted and purified as described in section 3.1.2.1 in a PCR consisting of 20 mM Tris pH 8.4, 50 mM KCl, 1.5 mM MgCl, 0.2 mM dNTP, 1 unit *Taq* Polymerase (Invitrogen, Burlington, ON), 500 nM of each universal *cpn*60 primer (H279 and H280 - Goh *et al*. 1996) and ddH_2O to 50 µl volume. PCR conditions were 1 cycle of 3 minutes at 95^{o}C, then 40 cycles of 95^{o}C for 1 minute, 46^{o}C for 1 minute, and 72^{o}C for 1 minute. Target DNA PCR product concentration was determined as described in section 3.1.1.

4.1.3 Application of target DNA to nylon membrane

Degenerate, universal PCR primers (H279, H280) were used to amplify 549-567 nt *cpn*60 UT from genomic DNA in a PCR as described in section 3.1.1. The resulting amplicons were quantified using Picogreen® (Invitrogen, Burlington, ON) and were applied to the membrane in duplicate spots without further purification. Target DNA was denatured and applied to positively charged nylon membranes (Roche) either by hand or with a Minifold I Microsample Filtration Manifold (Schleicher and Schuell). When applied by hand, amplified DNA from each strain was diluted to the desired concentration, denatured with 45 µl 0.4M NaOH for 10 minutes at room temperature, then neutralized with 50 µl 0.5 M Tris-HCl (pH 7.5). Duplicate 10 µl spots of this mixture were applied to the membrane using a micropipettor. When applied using the filtration manifold, undiluted, amplified DNA from each strain was denatured with 100 µl of 0.4 M NaOH and neutralized with a variable amount of 0.5 M Tris-HCl (pH 7.5) to bring the solution to a concentration of 10, 5, or 1 ng/µl in a total volume of 250 µl. Using vacuum pressure applied to the filtration manifold, 100 µl spots were applied to the membrane in duplicate. Irrespective of the method of application of target DNA, membranes were dried at 80^{o}C for 10 minutes and the DNA was UV crosslinked (UVC 500 UV Crosslinker, GE Healthcare) for 3 minutes to ensure attachment of DNA to the membrane.

4.1.4 Synthesis of probe

Probes were labelled with digoxigenin-dUTP (DIG-dUTP) (Roche) using PCR. Templates consisted of plasmid vectors (pGEM-T Easy (Promega) or pCR2.1-TOPO (Invitrogen)) containing *cpn*60 inserts from the species of interest. A typical labelling reaction consisted of 20 mM Tris pH 8.4, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dCTP, dATP, and dGTP, 0.13 mM dTTP, 1 unit Taq Polymerase (Invitrogen, Burlington, ON), 80 pM DIG-dUTP, 500

nM of universal (H279 and H280 primers) or plasmid-specific primers (T7 and M13-RP primers for pCR2.1-TOPO or T7 and SP6 primers for pGEM-T easy), and 500 pg of template DNA in a final volume of 25 µl. PCR conditions were 3 minutes at 95°C (1 cycle), then 95°C for 1 minute, 50°C for 1 minute, and 72°C for 1 minute (35 cycles). Amounts of labelled probe produced were determined as described in section 3.1.1.

4.1.4.1 Modification of probe specific activity

To alter the specific activity of DIG labelling, unlabelled DNA was amplified using the same template used to create the labelled probe in a PCR containing 20 mM Tris pH 8.4, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM dNTP, 1 unit Taq Polymerase , 500 nM of plasmid-specific primers (T7 and M13-RP or T7 and SP6), and 1 ng of template DNA in a final volume of 50 µl. PCR conditions were identical to those used to amplify the labelled probe. Amount of unlabelled probe DNA produced was determined using Picogreen® (Invitrogen, Burlington, ON) as described in section 3.1.1.

4.1.5 Hybridization

Nylon membranes were prehybridized at 42°C for at least 1 hour prior to exposure to labelled probe. Hybridization buffer consisted of 5x sodium chloride, sodium chloride sodium citrate (5x SSC), 1% blocking solution (Roche), 0.1% N-laurylsarcosine, 0.02% SDS, 50% formamide, and sterile water. In certain cases, labelled probe was not quantified and the entire labelling reaction was diluted in 20 ml of hybridization buffer; alternatively, 10 ng of labelled probe and 100 ng of unlabeled probe (competitor) DNA were diluted in 10 ml of hybridization buffer. Following addition of the probe and unlabelled DNA, fresh hybridization mixture was

denatured at 100°C for 10 minutes and was then placed in an ice water bath for 5 minutes. Hybridization proceeded overnight at 42°C.

4.1.6 Chemiluminescent development and detection

Membranes were removed from the oven and washed in 25 ml 0.1x SSC/0.1% SDS at room temperature for 1 minute, and then washed twice in 50 ml of 0.1x SSC/0.1% SDS at 68°C. Development of the chemiluminescent signal proceeded according to the manufacturer's recommendations (Roche). Chemiluminescent signals were detected by exposing the membrane to BioMax MR-1 film (Kodak) for 1 to 35 minutes. Densitometric analysis of the X-ray film exposures was performed using Scion imager software (Scion Corporation, Frederick, MD). Briefly, the drawing tool was used to define an area encircling the approximate area of a strong hybridization reaction. The same defined area was read from each positive hybridization signal. The mean pixel density within the circular area was reported for each spot. Measurements from each spot were recorded and pixel densities from duplicate spots were averaged.

4.2 Profiling microbial communities in ileal contents of pigs fed corn, wheat, or barley

Genomic microbial DNA extracted from the ileal contents of pigs fed corn, wheat, or barley as described in Hill *et al.* (2005) were examined using macroarray technology. Based on differences in clone frequencies, phylogenetic analysis, and qPCR quantification as determined by Hill *et al.* (2005), four bacterial groups were selected for hybridization analysis: B1 - Bacillales-related, S1 – *Streptococcus*-related, C1 – Clostridiales-related, and L10 – *Lactobacillus amylovorus* (Table 4-2). Labelled probe was prepared as described in section 3.1.3 using plasmids containing the *cpn*60 UT from the selected groups. Target DNA was amplified

Table 4-2. Plasmid-based *cpn*60 UT used for synthesis of target DNA for corn, wheat, and barley characterization (top half) and characterization of pre and post-weaning jejunal contents of piglets (bottom half).

Species identification	cpnDB ID	Nearest neighbour	Cloning vector
Bacillales-related – B1	B6076	*Bacillus circulans* (68.4%)	pCR2.1-TOPO
Streptococcus-related – S1	B5883	*Streptococcus intestinalis* (96.9%)	pCR2.1-TOPO
Clostridiales-related – C1	B6042	*Clostridium disporicum* (84.7%)	pCR2.1-TOPO
Lactobacillus amylovorous – L10	B7447	*Lactobacillus curvatus* (99.6%)	pCR2.1-TOPO
Lactobacillus amylovorous – L10	B7447	*Lactobacillus curvatus* (99.6%)	pCR2.1-TOPO
Escherichia coli – EC	B7482	*Shigella sonnei* (100%)[a]	pGEM-T Easy
Lactobacillus vaginalis-like – LV	V7588	*Lactobacillus vaginalis* (88.9%)	PGEM-T Easy
Streptococcus-related – S1	B5883	*Streptococcus intestinalis* (96.9%)	pCR2.1-TOPO

[a]*E. coli* and *S. sonnei* have identical *cpn*60 UT sequences

from genomic DNA extractions from pools of ileal contents of pigs fed each of the diets as described in section 3.1.1. In addition, universal primers were used to amplify the same plasmid used to create the labelled probe. Amplicons from genomic DNA were applied to the membrane using a dot-blot apparatus as described in section 3.1.3 in duplicate spots consisting of 100, 50, and 10 ng. Amplicons generated from the plasmid representing the labelled probe were applied to the membrane in duplicate spots consisting of 500, 300, 100, 50, 30, 10, 5 and 3 pg. Labelled probe was hybridized to the membranes and the chemiluminescent signal was detected and analyzed by densitometric analysis as described in sections 3.1.5 and 3.1.6. The resulting pixel densities were correlated with previously produced qPCR data (Hill *et al*. 2005).

4.3 Characterization of microbiota found in jejunal contents of pre and post weaning piglets

4.3.1 Selection of animals

Four piglets from each of four litters from Prairie Swine Center, Inc. (Floral, SK), were selected on day 20, prior to weaning (n=16). Two piglets per litter were euthanized on day 20 by submersion in CO_2 and exsanguination (pre-weaning treatment). The remaining eight piglets were weaned on day 20 onto a diet based on wheat and euthanized on day 25 (post-weaning treatment). Intestinal contents were collected from a 100 cm intestinal segment located at 50% of the length of the small intestine and were frozen at -20^{o}C.

4.3.2 Hybridization Analysis

DNA was extracted from the samples as described in section 3.1.2.1. Four different labelled probes were prepared for hybridization analysis of the pre and post weaning DNA samples. Labelled probes were created as described in section 3.1.4 from plasmids containing

*cpn*60 inserts representing *Lactobacillus amylovorous* (L10), *Escherichia coli* (EC), and a *Lactobacillus vaginalis*-like (LV) and a *Streptococcus alactolyticus*-like species (S1) isolated from the ileal contents of a mature pig gut (Table 4-2). DNA samples from pre and post-weaning piglets were used as template in a PCR using *cpn*60 universal primers as described in section 3.1.1. Amplicons were quantified and applied to a membrane as described in section 3.1.2 in duplicate spots of 100, 50, and 10 ng. In addition, target DNA was prepared from plasmids containing the relevant *cpn*60 sequence and applied to membranes in duplicate spots of 500, 300, 100, 10, 30, 10, 5, and 3 pg. Hybridization and densitometric analysis were carried out as described in sections 3.1.5 and 3.1.6.

4.3.3 qPCR

Quantitative PCR assays were used to enumerate the four bacterial groups examined using hybridization analysis in the digesta samples from pre- and post-weaning piglets, as well as the species *Streptococcus alactolyticus* (SA). The forward(f) and reverse(r) primer sequences for each target were S1f 5'-TTGACGTGGTTGAAGG-3'; S1r 5'-GTTTTCAAGACTTCTTCAAGCAA-3'; ECf 5'-GGCTATCATCACTGAAGGTCTG-3'; ECr 5'-TTCTTCAACTGCAGCGGTAAC-3'; LAf 5'-CATCTAAAGAAGTTGGTGACT-3'; LAr 5'-CAATGATTAACAAAGCCTTAC-3'; LVf 5'-TGGGATCGACAAGGCTACGC-3'; LVr 5'-CGTCGGCAATCAGCTTACCAA-3'; SAf 5'- CGAATCTGCGGTTTCAGTAGC-3' and SAr 5'- GATGATACGGCAGCAACTTGG-3'. A TaqMan® fluorescent probe was used for the EC assay (5'-TGTTGCTGCGGGCATGAACC-3'), which was labelled with 6-carboxyfluroecein at the 5'end and Black Hole Quencher 1 at the 3' end (Integrated DNA Technologies, Inc.)

For all assays, qPCR was performed using an iCycler iQ real-time PCR detection system (Bio-Rad Laboratories). Calibration standards consisted of a 10-fold dilution series of the same

plasmid used for construction of labelled target during hybridization analysis. For the L10, LV,S1 and SA assays, each reaction mixture contained iQ SYBRGreen Supermix, 500 nM each primer, 2 µl of each plasmid calibration standard or genomic DNA extraction, and H_2O to a final volume of 25 µl. Amplification conditions were 50°C for 2 minutes, 95°C for 3 minutes, followed by 40 cycles of 30 seconds at 95°C, 30 seconds at 51°C (S1), 55°C (L10 and SA), or 58°C (LV), and 30 seconds at 72°C. Fluorescence data was collected during the 72°C step. Immediately following all assays except the EC assay, the melting temperature of the PCR product was determined under the following conditions: 30 second dwell at 55°C, followed by increasing temperature at 0.5°C increments for 80 cycles of 10 seconds each. For the EC assay, the reaction contained Platinum Quantitative PCR Supermix-UDG (Invitrogen), 500 nM each primer, 200 nM probe, and 2 µl plasmid calibration standard or genomic DNA extraction in a total reaction volume of 25 µl. PCR conditions were 50°C for 2 minutes, 94°C for 3 minutes, followed by 40 cycles of 94°C for 30 seconds, and 61°C for 30 seconds. The number of genomes present in the digesta samples was determined as described in Dumonceaux *et al*. (2005).

4.4 Statistical Analysis

Densitometric analysis from both the CWB library and pre and post-weaning treatments were correlated with corresponding qPCR results. Correlation was done using Microsoft Excel with correlation values reported.

Statistical analysis of pre and post-weaning treatments was done using SPSS. A one-way ANOVA was used to analyze the difference between means. A level of significance was set at $p = 0.05$.

5.0 Results

5.1 Assay development and characterization

5.1.1 Target generation with *cpn*60 UT primers

To study the degree of cross hybridization of labelled probe with target DNA spotted on the nylon membrane, universal *cpn*60 PCR primers were used to prepare target DNA by amplification of plasmids containing *cpn*60 UT inserts from *Prevotella zoogleoformans* and from several bacterial species with *cpn*60 UT ranging from 81-89% identity with *P. zoogleoformans*. When membranes were hybridized with DIG-labelled *P. zoogleoformans cpn*60 UT that had also been prepared by universal *cpn*60 primer amplification of plasmid DNA containing the relevant *cpn*60 UT insert, significant cross-hybridization was observed even at a sequence identity of 81% (Figure 5-1). The intensity of the hybridization signal was not related to the degree of sequence identity.

5.1.2 Target generation with plasmid-specific primers

It was hypothesized that the used of *cpn*60 universal primers to generate both the target and probe DNA was a contributing factor to the level of cross hybridization observed. In an effort to decrease background signals and improve the specificity of the hybridization signal, plasmid-specific primers (SP6, T7, and M13) were used to replace universal *cpn*60 primers for the generation of target DNA.

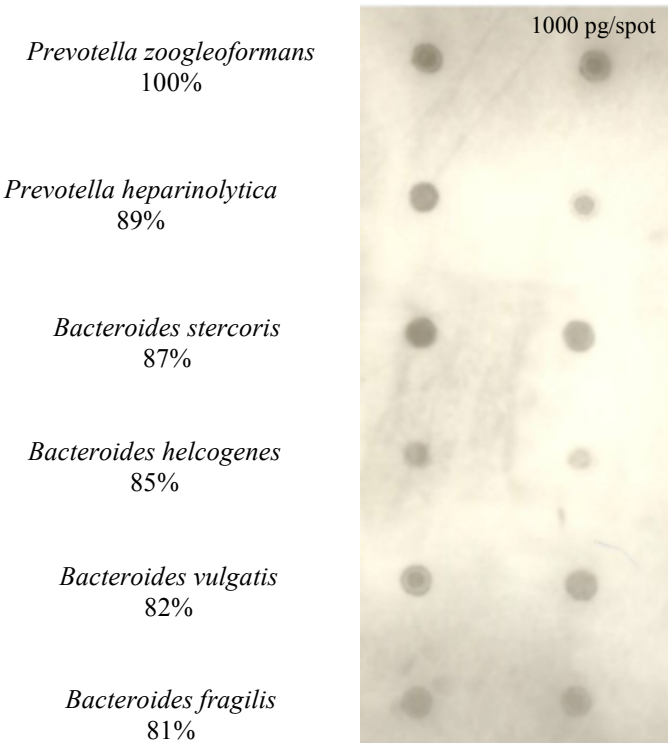

Figure 5-1. Macroarray demonstrating the level of cross-hybridization observed with universal primers used to amplify probe and target DNA. Duplicate *cpn*60 UT target DNA spots (1000 pg/spot) were hybridized a with DIG-labelled *P. zoogleoformans cpn*60 UT probe. The identities of target DNA spots and pair-wise identities to probe are labelled at left. The membrane was exposed to X-ray film for 30 minutes. This figure is representative of several replicates of this experiment.

Target DNA was amplified from several plasmids containing *cpn*60 UT with both universal *cpn*60 primers and plasmid-specific primers and applied to the membrane in duplicate in a dilution series ranging from 50-500 pg/spot. Following hybridization with DIG-labelled *P. zoogleoformans cpn*60 UT amplified from plasmid, the level of cross-hybridization produced by each primer set was observed by visual comparison of the intensity of the chemiluminescent signal in the corresponding spots. Signal intensities obtained using the primers SP6/T7-generated target DNA was comparable to the signal for the universal primer target DNA for 100% sequence identity (Figure 5-2). At 85% sequence identity, SP6/T7 primer-generated target DNA showed markedly less apparent cross-hybridization than target DNA generated using *cpn*60 universal primers for all dilutions. At 81% sequence identity, little apparent cross-hybridization was observed using either primer set.

To determine the level of cross-hybridization above 85% pairwise identity, another series of related probe and target DNA was prepared. Two membranes cross linked (500 pg/spot) with a different set of target DNA were prepared. For the first membrane, target DNA was prepared by universal *cpn*60 primer amplification of *cpn*60 UT with a sequence identity range of 95 to 92% with the selected probe. Universal PCR primers were chosen to amplify target DNA because the *cpn*60 UT sequences were in different plasmid vectors, therefore the use of multiple plasmid-specific primers would have been required. This change was not expected to impact the results. The selected probe was prepared from a plasmid containing *cpn*60 UT from *Citrobacter freundii* using plasmid-specific SP6 and T7 primers. Following hybridization, all spots with exception of the 64% negative control showed a similar high chemiluminescent signal intensity indicating strong cross-hybridization (Figure 5-3A). For the second membrane, universal *cpn*60 primers were used to amplify target DNA from several species with pairwise identities ranging

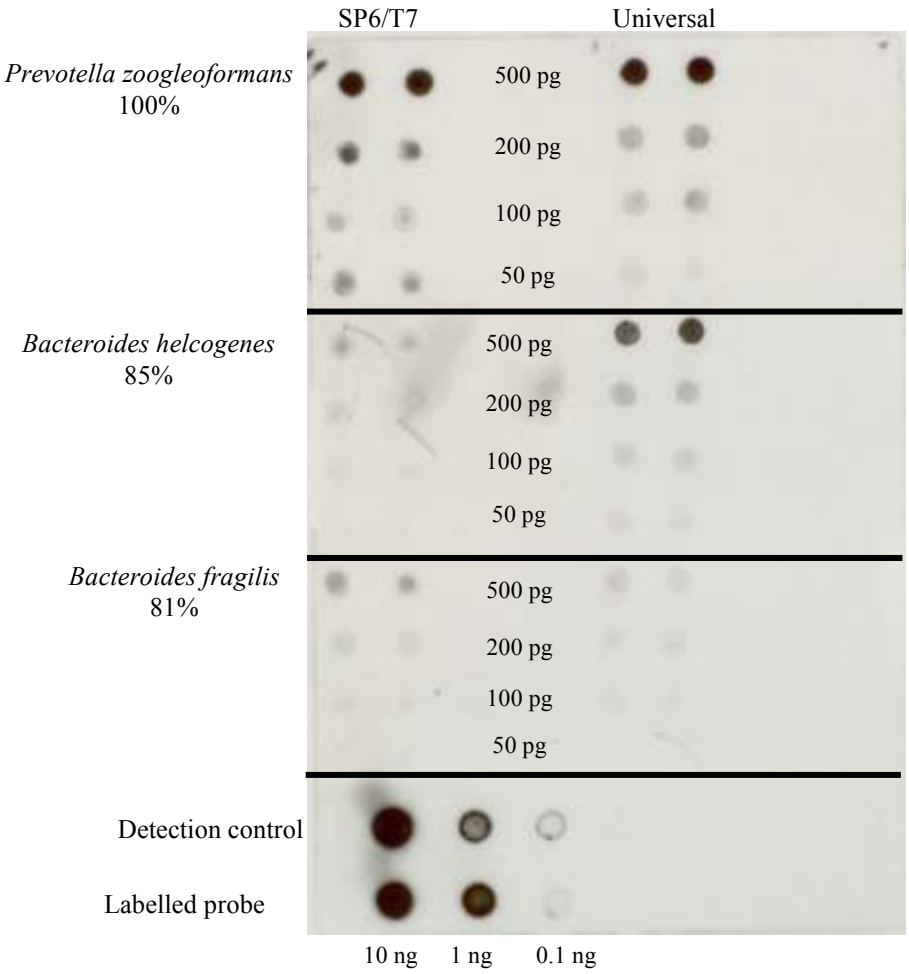

Figure 5-2. A macroarray demonstrating the amount of cross-hybridization observed when target DNA amplified with either universal *cpn*60 (H279/H280) or plasmid-specific (SP6/T7) primers was hybridized with a DIG-labelled *P. zoogleoformans* *cpn*60 UT probe amplified with universal *cpn*60 (H279/H280) primers. The identities of target DNA spots and pair-wise identities to probe are labelled at left. Amount of target DNA in each duplicate spot is indicated in the centre column. The detection control was DIG-labelled λ-DNA cross linked directly to the membrane. Probe labelling efficiency (labelled probe control) was assessed by crosslinking labelled probe (10-0.1 ng) directly to membrane. The membrane was exposed to X-ray film for 15 minutes.

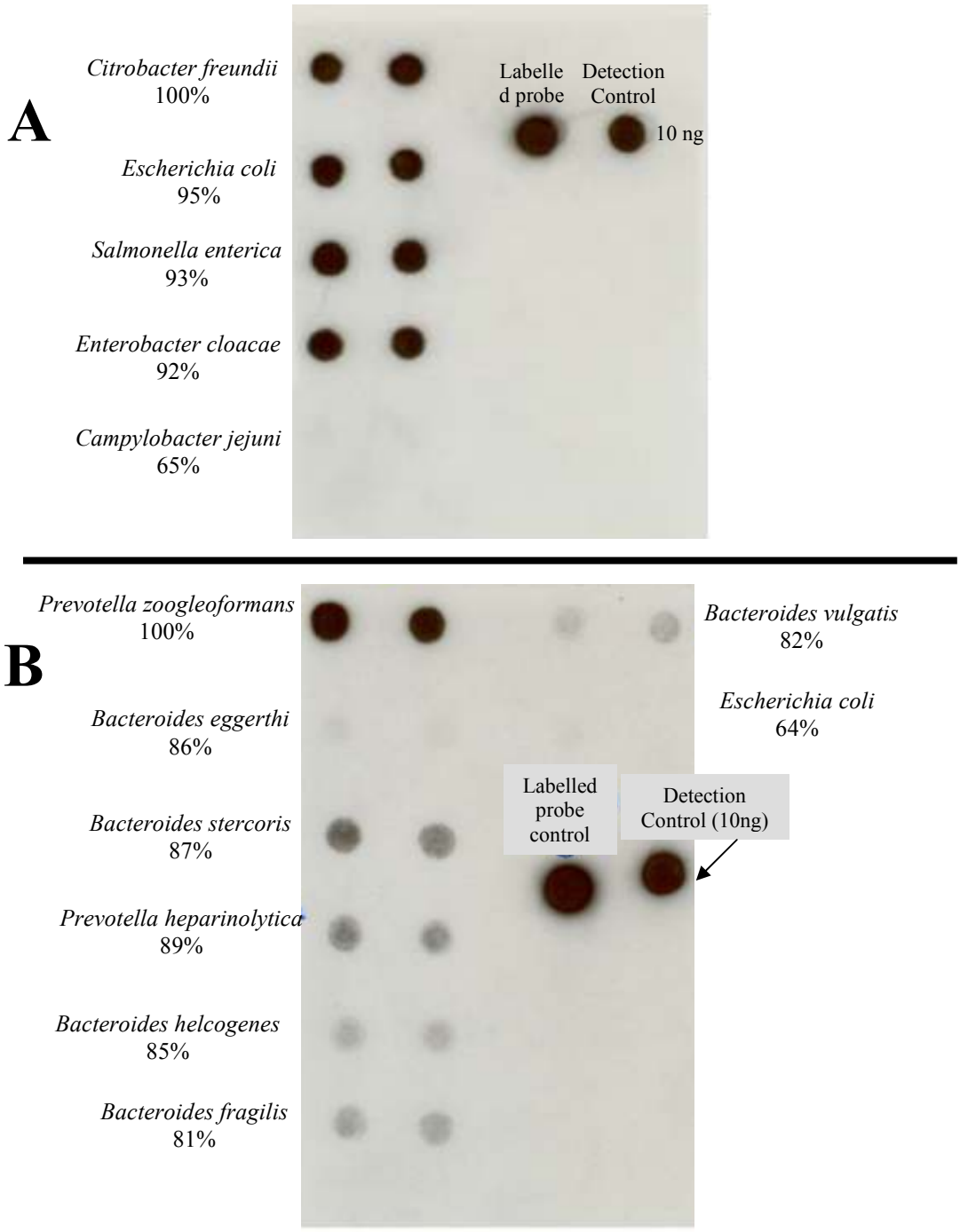

Figure 5-3. Macroarrays demonstrating the level of cross-hybridization observed using target DNA amplified with universal *cpn*60 primers and DIG-labelled probe produced from *C. freundii* (panel A) and *P. zoogleoformans* (panel B) with plasmid specific (SP6/T7) primers. The identities of target DNA spots and pair-wise identities to probe are labelled at left. Amount of target DNA in each duplicate spot is 500 pg. The detection control is DIG-labelled λ-DNA. Probe labelling efficiency (Labelled probe control) was assessed by crosslinking labelled probe (10ng) directly to membrane. The membranes were exposed to X-ray film for 15 minutes for panel A and 5 minutes for panel B.

from 89-81% to the selected probe, *P. zoogleoformans,* prepared as previously using SP6 and T7 primers. In addition, a negative control target DNA was also prepared with 64% identity to *P. zoogleoformans*. Following hybridization, a strong signal was observed for the DNA target spots with 100% identity to probe. Target DNA with 89% and 87% sequence identity showed a weak hybridization signal that was clearly greater than background. All remaining spots showed low but variable level of hybridization intensity that was not directly related to pairwise identity with the probe (Figure 5-3B).

In an effort to decrease cross-hybridization further, we attempted to modify the specific activity of the labelled probe by dilution of DIG-labelled probe with unlabeled probe. Unlabeled probe DNA was prepared using the same template as the labelled probe under identical amplification conditions; however no DIG-dUTP was included in the PCR. Unlabeled probe DNA was included in the hybridization reaction at 1x and 10x the amount of labelled probe and the experiment in Figure 5-3B was repeated (Figure 5-4). A considerable and reproducible reduction in cross-hybridization was observed with the inclusion of 10x unlabeled DNA in the hybridization reaction. All subsequent experiments included the use of 10x unlabeled DNA in the hybridization in an effort to decrease cross-hybridization.

To improve the consistency of spot size and cross hybridization a membrane dot blotting apparatus was used to apply target DNA to the nylon membrane. The cross hybridization experiments described above were essentially repeated using a dot blotting apparatus as shown in Figure 5-5. Universal *cpn*60 PCR primers were used to amplify target DNA and plasmid specific PCR primers (SP6/T7) were used to amplify labelled probe DNA, as well as the unlabeled DNA included in the hybridization reaction at 10 fold the amount of labelled probe. Cross- hybridization was observed at approximately 85% sequence identity between target DNA

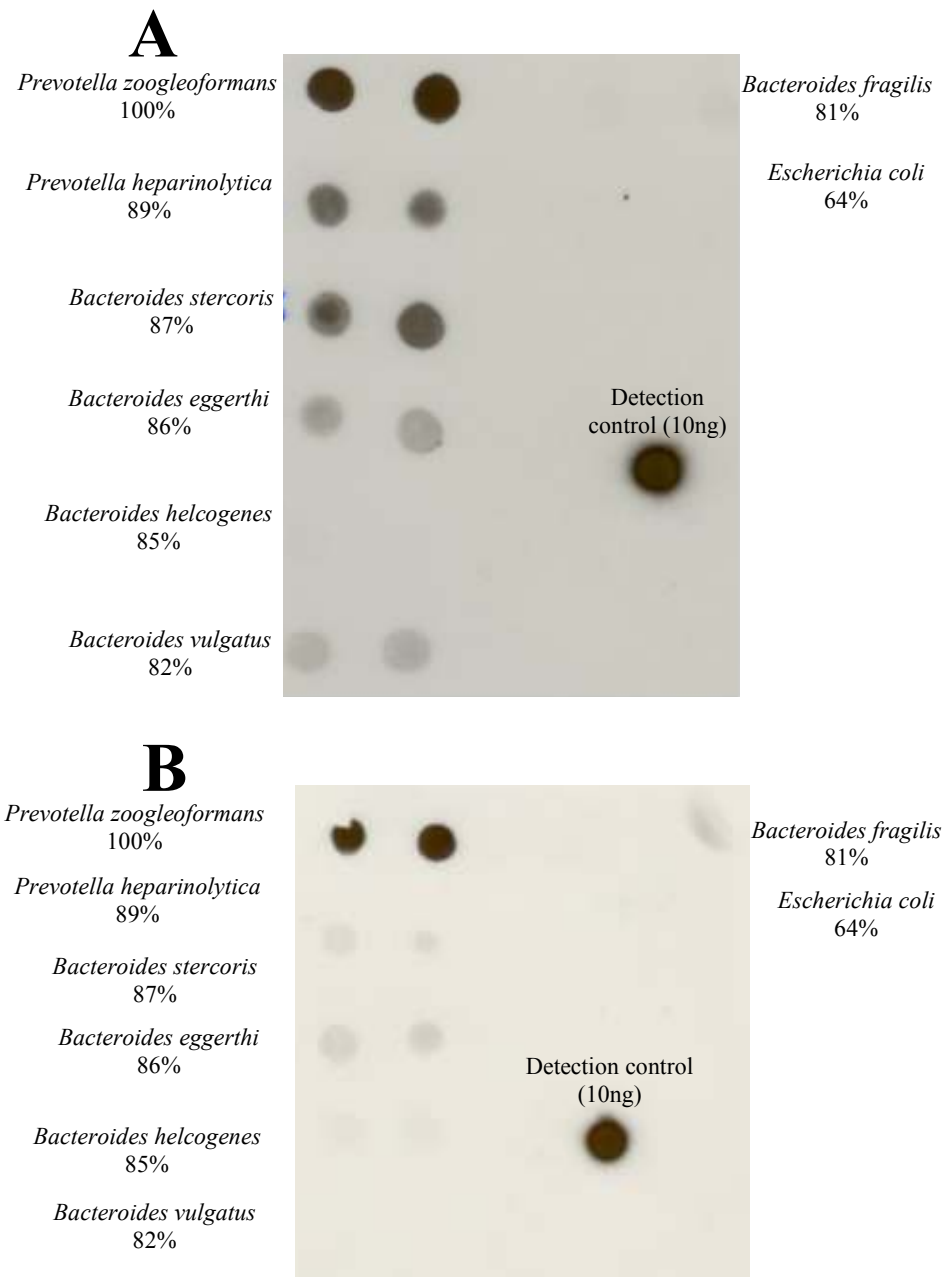

Figure 5-4. Macroarrays demonstrating the level of cross hybridization observed when specific activity of DIG-labelled probe DNA produced from *P. zoogleoformans* was reduced by mixing with unlabelled probe DNA. The identities of target DNA spots and pair-wise identities to probe are labelled at left. Amount of target DNA cross-linked at each duplicate spot is 500pg. The detection control is DIG-labelled λ-DNA. Unlabeled and labelled probe DNA were combined in equal (1:1 ratio) amounts (Panel A) or at a 10:1 ratio of unlabelled to labelled probe (Panel B). X-ray film exposure time was 15 minutes. Amount of target DNA was 500pg/spot.

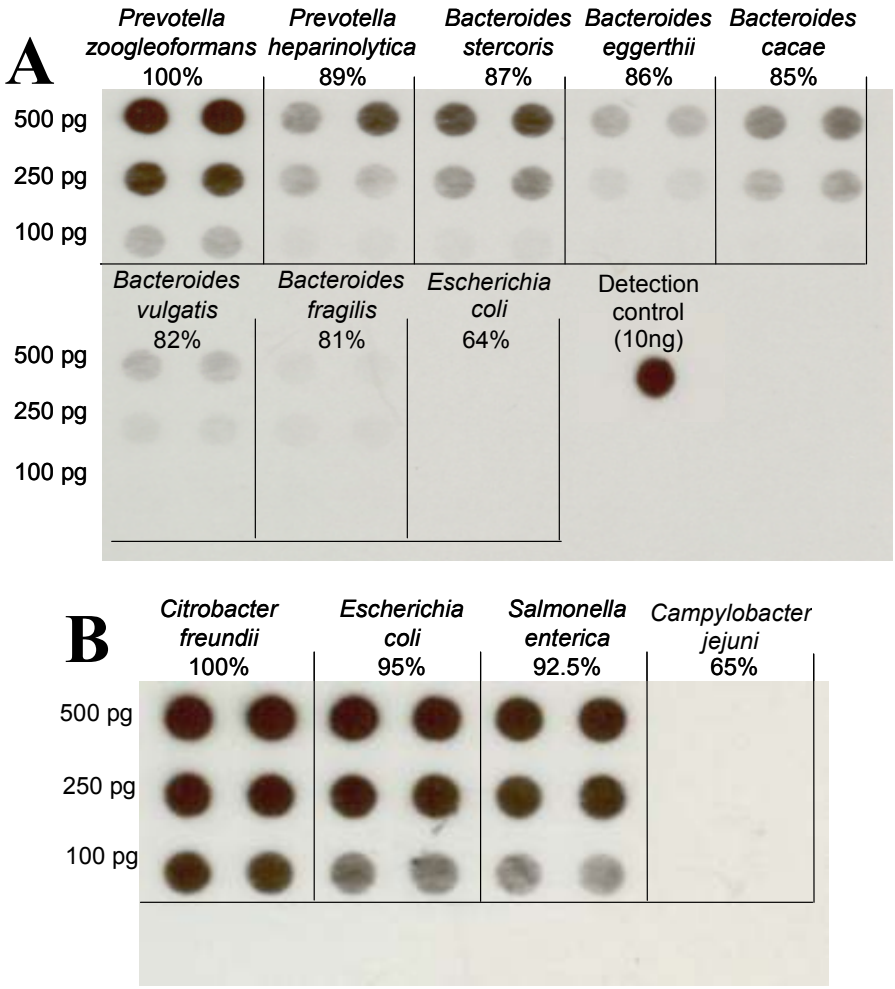

Figure 5-5. Macroarrays demonstrating the level of cross-hybridization observed between target and probe DNA for pairwise sequence identities ranging between 81% and 100% (panel A) and 92.5% and 100% (panel B). DIG-labelled probe DNA was produced from *P.zoogleoformans* (panel A) and *C. freundii* (panel B). Target DNA was applied to the membrane using a Minifold I Microsample Filtration Manifold. The identities of target DNA spots and pair-wise identity to probe are labelled at top. Amount of target DNA in each duplicate spot are indicated at left. The detection control is DIG-labelled λ-DNA. Unlabeled competitor DNA was included at 10x the probe amount. The membrane was exposed to X-ray film for 15 minutes for both panels.

and the labelled probe when using the dot blotting apparatus. In other words, any bacterial species that has 85% or greater *cpn*60 UT sequence identity will cause a positive reaction using this method.

It should be noted that very little cross hybridization was noted in the macroarray displayed in Figure 5-4B, when the array was applied to the membrane by hand. More cross hybridization was observed when using a filtration manifold to apply DNA to the membrane. The filtration manifold was used because it allowed for more accurate and uniform application of the target DNA to the membrane, resulting in a more uniform distribution of signal intensity.

As evidenced by Figures 5-4 and 5-5, hybridization signal intensity decreases with lower sequence identities. This decrease in signal intensity is not always proportional to the sequence identity of each species. In Figure 5-5A, both *Bacteroides stercoris* and *Bacteroides cacae* showed slightly higher signal intensity than other species with a higher sequence identity to the probe species. However, both signal intensities are noticeably less than the signal intensity produced by the identical match to the probe species.

5.2 Profiling microbial communities found in ileal contents of pigs fed corn, wheat, or barley

Following assay development and partial characterization, we further evaluated this technology by its application to assess microbial composition in previously well characterized samples of intestinal contents. A study conducted by Hill *et al.* (2005) characterized the microbial flora found in the ileal contents of pigs fed diets based on corn, wheat, or barley (CWB). Briefly, ileal contents were obtained from pigs fed diets based on corn, wheat, or barley. From these contents, genomic DNA was isolated and *cpn*60-based PCR product libraries were created and sequenced. The resulting sequences were compared with cpnDB and the

identity of the nearest neighbour sequence was determined. Sequences obtained from each diet treatment were compared based on the genus or taxonomic group of the nearest identified neighbour. Based on this phylogenetic profile of unique *cpn*60 UT sequences from the pig ileum, four target groups (S1, C1, B1, and L10 – see Table 4-2) were selected for analysis with the macroarray methodology established above. These target groups were selected because they showed major differences in prevalence among the different treatment groups.

For macroarray analysis of each of the four target groups, target DNA was amplified using universal *cpn*60 primers from pooled total DNA extracted from ileal contents of the pigs fed the corn-based diet. This was repeated for both the barley and wheat-based diets. A dilution series of each of these 3 target DNAs was applied to the membrane in duplicate using the dot blot apparatus. DNA was also amplified from a plasmid containing a representative *cpn*60 UT of the target group under examination using universal *cpn*60 primers. A serial dilution of this DNA was applied to the membrane using a dot blot apparatus. The membrane was then hybridized with a DIG-labelled probe prepared by amplification of the plasmid containing the representative *cpn*60 UT of the target group being examined using SP6 and T7 primers (10:1 mixture of unlabelled and DIG labelled amplicons).

Figure 5-6 shows an example of the macroarray design used to analyze each of the four groups. The results shown indicate the hybridization results obtained for the B1 group. Decreasing target DNA concentration affected the intensity of the hybridization signal in a dose dependant manner indicating a relationship between signal intensity on target amount and potential for target quantification. A very strong signal was observed in the spots corresponding to pooled ileal contents from pigs fed corn, with almost no signal observed for the spots corresponding to pooled ileal contents from pigs fed wheat or barley.

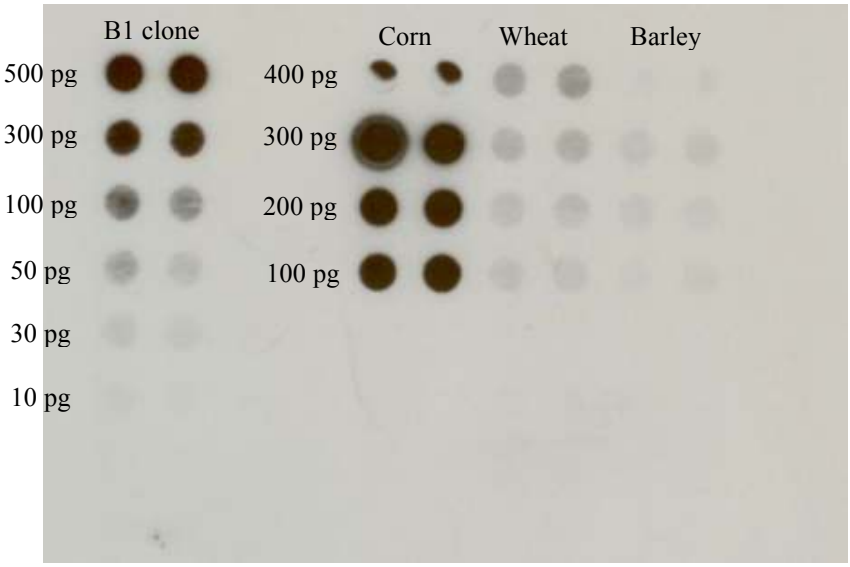

Figure 5-6. A macroarray demonstrating results observed when a DIG-labelled probe DNA created from representative B1group probe was hybridized to target DNA from prepared from ileal contents of pigs fed corn, wheat, or barley-based diets. Target DNA identical to the representative B1 group probe is on the left side of the membrane in a duplicate dilution series from 10-500 pg probe per duplicate spot. Target DNA prepared from pooled ileal contents for each diet is at right in a duplicate dilution series from 100-400 pg. The membrane was exposed to the X-ray film for 15 minutes. Note that the lack of signal for the 400 pg spot for the corn diet is likely the result of lost target DNA due to a seal leak in the filtration manifold.

Table 5-1 shows the correlation between frequency of sequence recovery for the four groups examined from the plasmid libraries (from Hill *et al.* 2005), qPCR data, signal intensity determined from hybridization results, and the correlation between qPCR and signal intensities. Figure 5-7 depicts densitometric data obtained using the S1 probe and the B1 probe. The first 3 blots (A,B,C) depict densitometric data observed with the S1 probe at various X-ray film exposure times. The greatest difference among the hybridization signals for each of the pooled ileal contents was observed using low amounts of target DNA (10-50ng) and short X-ray film exposure times (5 minutes). Larger amounts of target DNA and longer exposure times resulted in overexposure of the membrane such that signal intensity was at maximum value and in the ability to distinguish quantitative differences between the target DNA populations was lost (Figure 5-7B, 5-7C). As indicated in Figure 4-6, densitometric data depicted in Figure 5-7D shows a high amount of B1 present in the corn library, but very low levels in the wheat and barley libraries.

5.3 Profiling of microbiota in jejunal contents of pre and post-weaning piglets

Macroarray technology was used to characterize the microbial communities present in the jejunum of eight pre-weaning and eight post-weaning piglets. Four bacterial targets were selected for analysis – *Lactobacillus amylovorous* (L10), *Lactobacillus vaginalis* (LV), S1 (*Streptococcus*-like), and *E. coli* (EC). Target DNA was applied to one of two membranes that were prepared for each probe species being analyzed. As an interassay control, DNA from two randomly chosen targets were mixed in equal proportions and applied to both membranes. No data was obtained from a pre-weaning piglet from litter 3 and a post-weaning piglet from litter 1, as not enough contents were obtained for successful DNA extraction to occur.

Table 5-1 Correlation of densitometric analysis and qPCR for groups B1, S1, C1 and L10.

Probe	Library	Signal Intensity[a]	No. of genomes[b] (log/g of contents)	Correlation Coefficient[c]	Plasmid Frequency
C1	Corn	172.3	9.1 ± 0.2	0.916	88
	Wheat	150.0	8.8 ± 0.1		69
	Barley	134.7	8.2 ± 0.3		18
S1	Corn	116.4	5.3 ± 0.1	0.887	1
	Wheat	129.3	5.9 ± 0.2		4
	Barley	140.6	6.0 ± 0.2		48
B1	Corn	139.0	8.5 ± 0.0	0.952	16
	Wheat	23.7	6.4 ± 0.0		7
	Barley	9.39	5.4 ± 0.0		0
L10	Corn	85.44	10.0 ± 0.0	0.249	506
	Wheat	131.5	10.2 ± 0.0		616
	Barley	108.5	9.8 ± 0.0		686

[a] Signal intensities for C1 and S1 probes are the mean of 2 determinations, taken from a 10 ng spot after 15 minutes of exposure; signal intensities for B1 probe is the mean of 2 determinations, taken from a 100 ng spot after 25 minutes of exposure; signal intensities for L10 probe is the mean of 2 determinations, taken from a 5 ng spot after 5 minutes of exposure

[b] The mean of triplicate determinations as determined by Hill *et al* 2005

[c] Correlation between No. of genomes and signal intensity

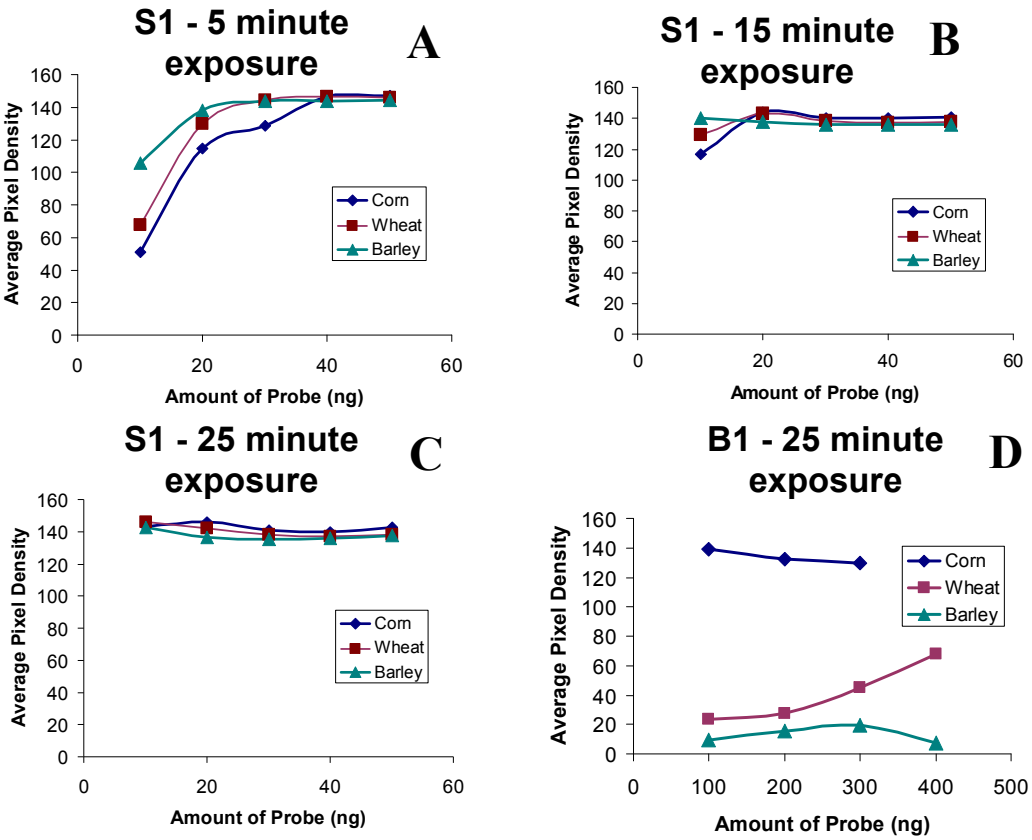

Figure 5-7. Densitometric data from S1 and B1 hybridization analysis of CWB libraries. Length of X-ray film exposure times and bacterial group used are indicated above the graphs.

Figure 5-8 depicts a membrane blot used to determine the levels of L10 DNA in the jejunal contents of pre and post-weaning piglets. Target DNA was amplified from DNA extracts of jejunal contents using universal PCR primers and applied to the membrane in duplicate spots ranging from 100 – 10 ng. Differences in the amounts of L10 DNA were apparent among piglets particularly when lower amounts of target DNA prepared from jejunal contents was applied to the membrane. Densitometric analysis of all hybridization arrays was performed. The pixel densities observed for L10 hybridization are shown in Figure 5-9A and B. Panel A depicts data observed from the 50 ng spot of target DNA after 5 minutes of exposure. The majority of the spots for both pre and post-weaning piglets appear to be nearing saturation at this amount of target DNA. The averages for pre and post-weaning levels of L10 were not significantly different ($p = 0.072$). Panel B depicts pixel density from the 10ng spot after 5 minutes of X-ray film exposure. With the exception of one post-weaning piglet from litter 2, the pre-weaning piglets show notably more signal than do the post-weaning piglets. A statistically significant difference was observed between the average of the pre and post-weaning levels of L10 ($p < 0.01$). Panel C represents the number of L10 genomes enumerated in each of the jejunal sample DNA extracts using L10 specific primers and qPCR. No data for one pig in litter 2 could be reported as the total DNA extract from this piglet appeared to be contaminated with a substance that fluoresced at the same wavelength as SYBR® Green. As in panel B, a significant difference was observed between the average number of L10 genomes in the pre and post-weaning jejunum with qPCR. A positive correlation ($r2 = 0.7473$) was observed for L10 pixel density determined by hybridization (10ng of target DNA and 5 minutes of exposure) and L10 genomes determined by qPCR.

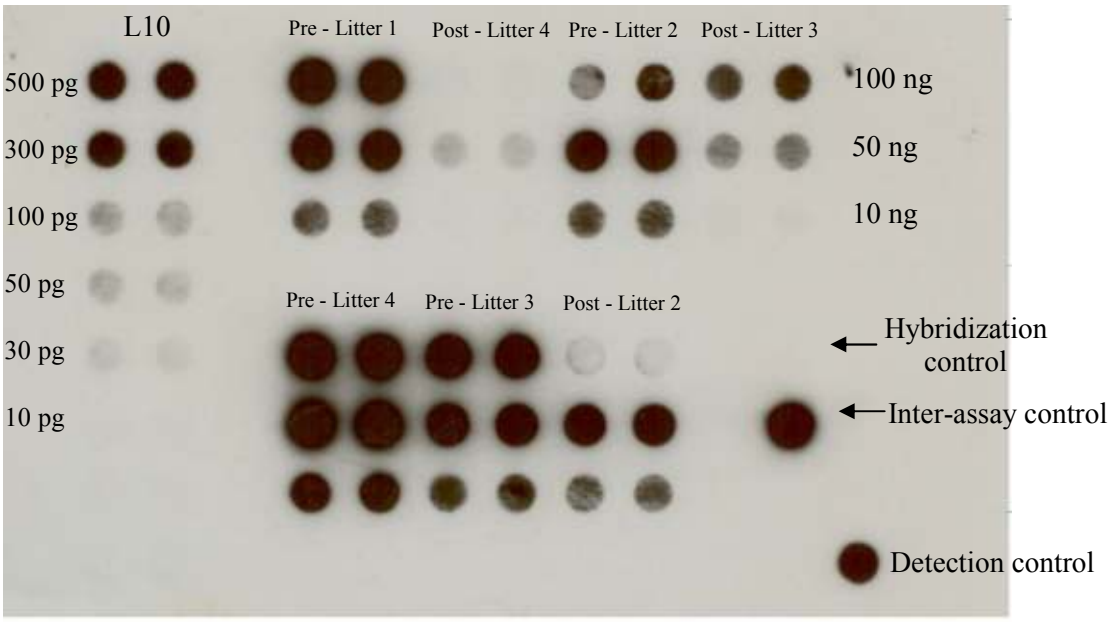

Figure 5-8. A representative macroarray demonstrating hybridization of L10 DIG-labelled probe to L10 target DNA and target DNA prepared from the jejunal contents of pre and post-weaning piglets. Target DNA identical to the L10 group probe is on the left side of the membrane in a duplicate dilution series from 10-500 pg probe per duplicate spot. Target DNA prepared from jejunal contents for each diet and litter is at right in a duplicate dilution series from 10-100 pg. The hybridization control is 500pg of *E. coli* DNA, which is approximately 65% related to the L10 probe. The interassay control is 50 pg of target DNA from Pre – litter 4 and 50 pg of target DNA from Pre – litter 2 (present on the macroarray developed in conjunction with macroarray seen above – not shown). The detection control is 10 ng of DIG-labelled lambda DNA. The membrane was exposed to X-ray film for 5 minutes.

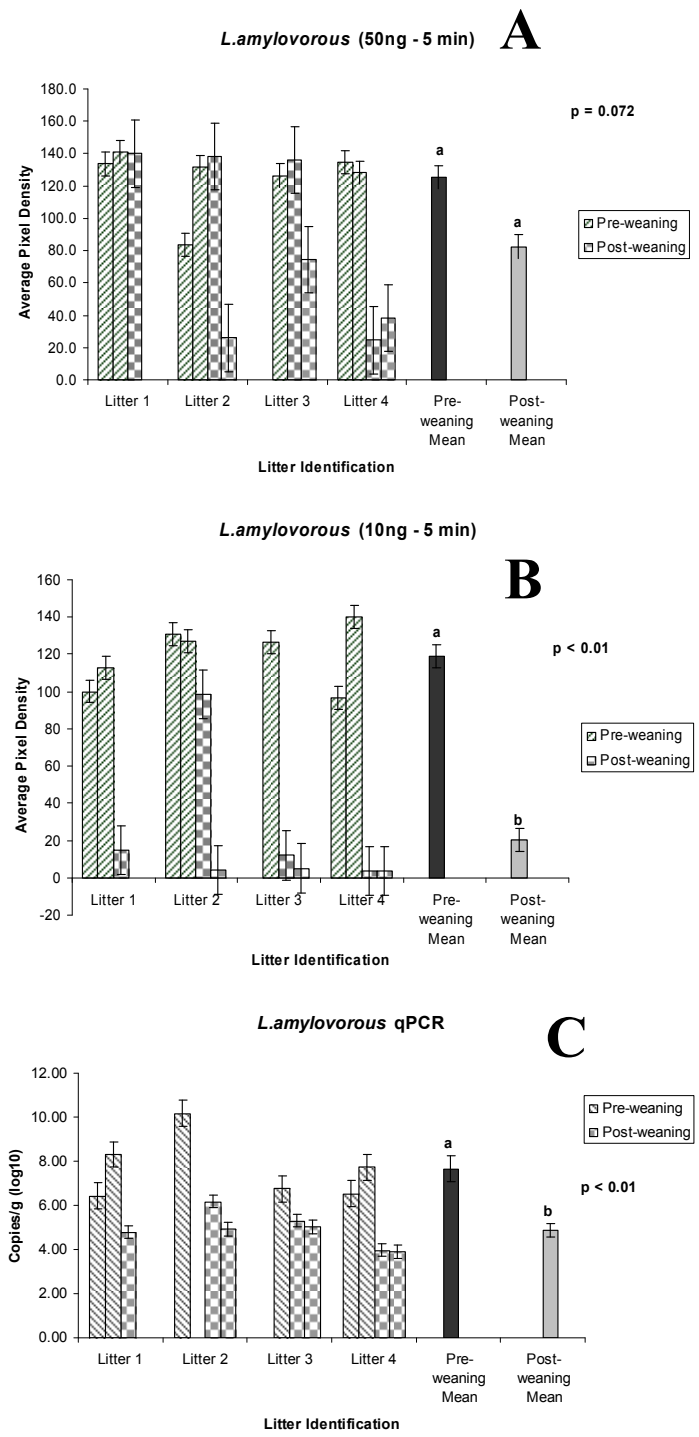

Figure 5-9. Mean (±SE) *L. amylovorous* pixel density determined by hybridization assay or *L. amylovorous* genome copy number determined in DNA extracted from jejunal contents of pre and postweaning pigs. Panel A shows pixel density for 50 ng target DNA after 5 minutes X-ray film exposure. Panel B shows pixel density for 10 ng target DNA after 5 minutes X-ray film exposure. Panel C show genome copy number determined by qPCR. Data for duplicate values for individual pigs within each litter is presented at left. The mean value pre and post weaning is presented at right.

Figure 5-10 represents the hybridization and qPCR data for pre and post-weaning levels of *Lactobacillus vaginalis* (LV). Both methods indicated a statistically significant decrease LV bacteria in jejunum of the post-weaning pig. However, LV abundance determined by hybridization analysis showed a higher degree of variation than did abundance observed with qPCR. Correlation between hybridization and qPCR data was calculated to be moderately positive ($r^2 = 0.3862$). Levels of S1 (*Streptococcus*-related) in pre and post-weaning piglets are depicted in Figure 5-11. Pixel density for S1 DNA was significantly greater for pre-weaning than postweaning in jejunal contents when either 50 ng ($p < 0.01$) or 10 ng ($p < 0.05$) of target DNA was spotted on the membrane (Figure 5-11 A,B). However, analysis of S1 genome number with qPCR showed a much different trend, with no significant difference observed between pre and post-weaning levels (Figure 5-11 C). Poor correlation was observed between hybridization data and qPCR ($r^2 = 0.0208$).

The high level of cross reaction observed for the hybridization based assay, relative to the high specificity qPCR primers used, could explain the discordance between hybridization and qPCR results for S1 abundance. Thus the higher level of S1 in pre-weaning jejunum as measured by densitometric analysis could be explained by the presence of a bacterium possessing a *cpn*60 UT sequence with > 85% identity to S1. This species would be detected by the S1 probe, but not by S1 qPCR primers. To test this hypothesis, *cpn*60 UT sequences libraries prepared from pig intestinal contents (Hill *et al*. 2002, Hill *et al* .2005) were examined to determine bacterial species that may cross hybridize with the S1 probe. The *cpn*60 UT for *Streptococcus alactolyticus* was recovered from pig intestinal libraries and is 97% identical to the S1 *cpn*60 UT suggesting cross reaction with S1 during hybridization.

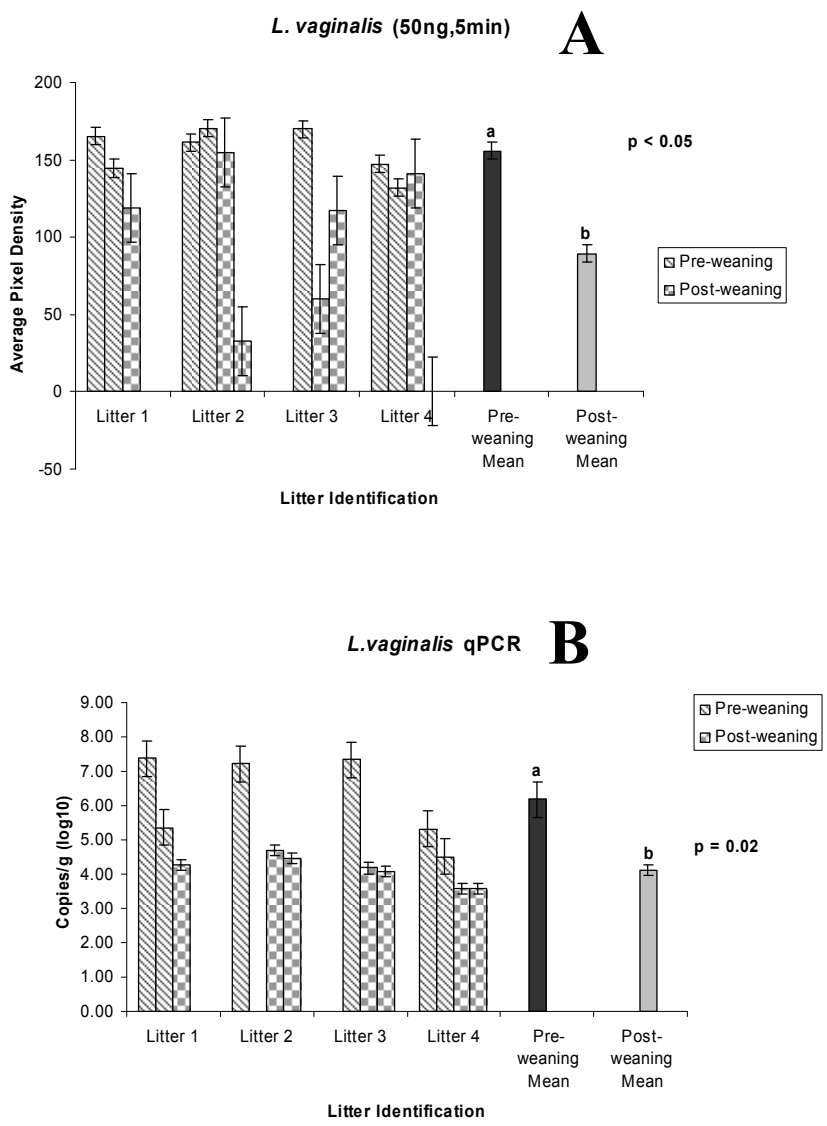

Figure 5-10. Mean (±SE) LV pixel density determined by hybridization assay or LV genome copy number determined in DNA extracted from jejunal contents of pre and postweaning pigs. Panel A shows pixel density for 50 ng target DNA after 5 minutes X-ray film exposure. Panel B shows genome copy number determined by qPCR. Data for duplicate values for individual pigs within each litter is presented at left. The mean value pre and post weaning is presented at right.

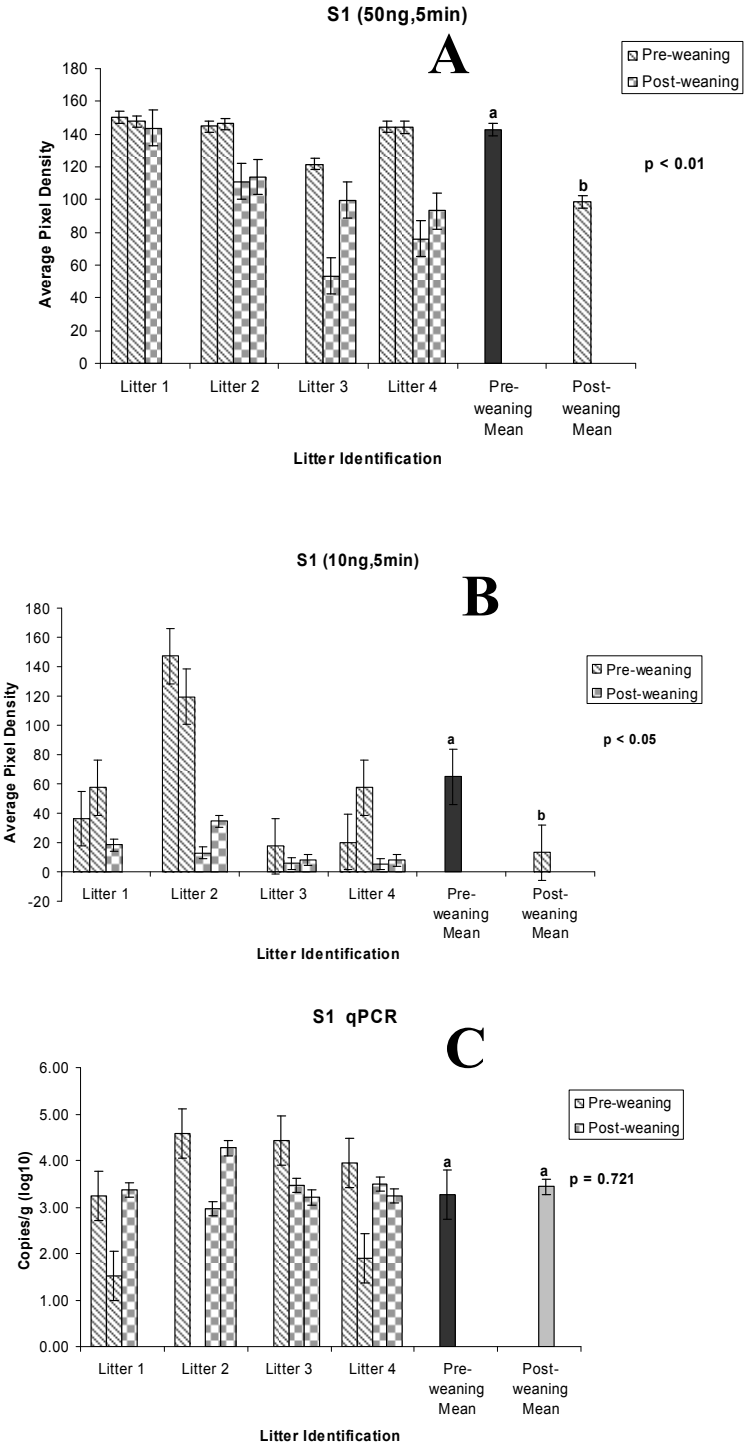

Figure 5-11. Mean (±SE) S1 pixel density determined by hybridization assay or S1 genome copy number determined in DNA extracted from jejunal contents of pre and postweaning pigs. Panel A shows pixel density for 50 ng target DNA after 5 minutes X-ray film exposure. Panel B shows pixel density for 10 ng target DNA after 5 minutes X-ray film exposure. Panel C shows genome copy number determined by qPCR. Data for duplicate values for individual pigs within each litter is presented at left. The mean value pre and post weaning is presented at right.

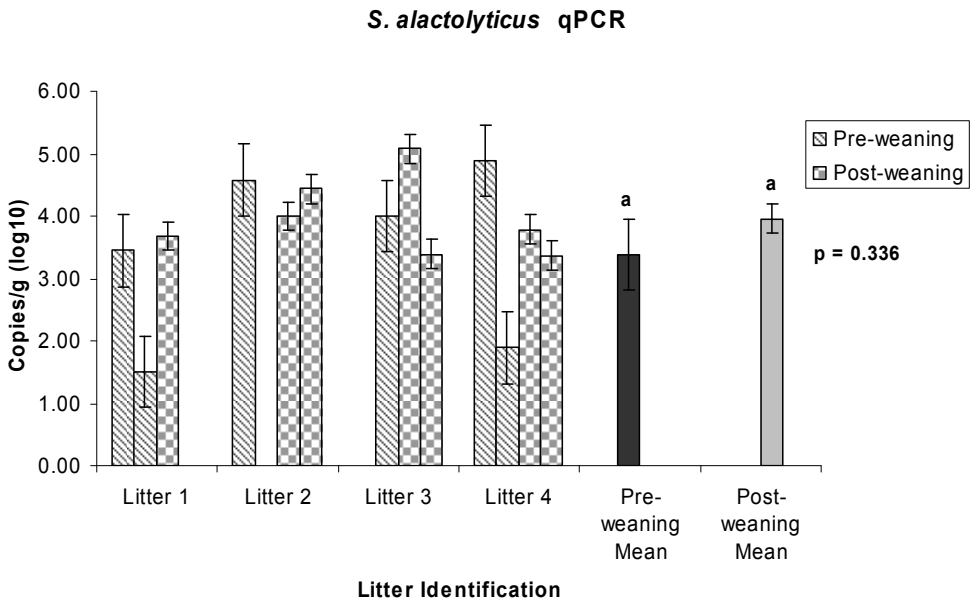

Figure 5-12. Genome copy number of *Steptococcus alactolyticus* determined by qPCR. Data for duplicate values for individual pigs within each litter is presented at left. The mean value pre and post weaning is presented at right.

qPCR using *S. alactolyticus* specific primers is provided in Figure 5-12. qPCR results showed that *S. alactolyticus* was present at lower levels than the S1 group,but a nearly identical trend to S1 genome copy number was observed (i.e. similar number pre and post weaning). Thus high colonization by *S. alactolyticus* pre-weaning, resulting in cross hybridization with the S1 probe does not explain the discordant results for S1 abundance as determined by hybridization assay and qPCR.

The pre and post-weaning abundance of *E.coli* was also determined as shown in Figure 5-13. No significant differences were observed between pre and post-weaning samples, although a large amount of individual variation was observed (Figure 5-13A). When compared to the levels of other species evaluated with macroarray technology, levels of *E. coli* were lower as greater amounts of target DNA and longer exposure times were required for adequate signal to be obtained. qPCR results agreed with this observation (Figure 5-13B), where lower copy numbers were observed than any other target organism analyzed here. The correlation between hybridization data and qPCR results was low ($r^2 = 0.0129$), however, both qPCR and hybridization showed slightly higher levels of *E.coli* in the post-weaning group though neither method showed values that were significantly different.

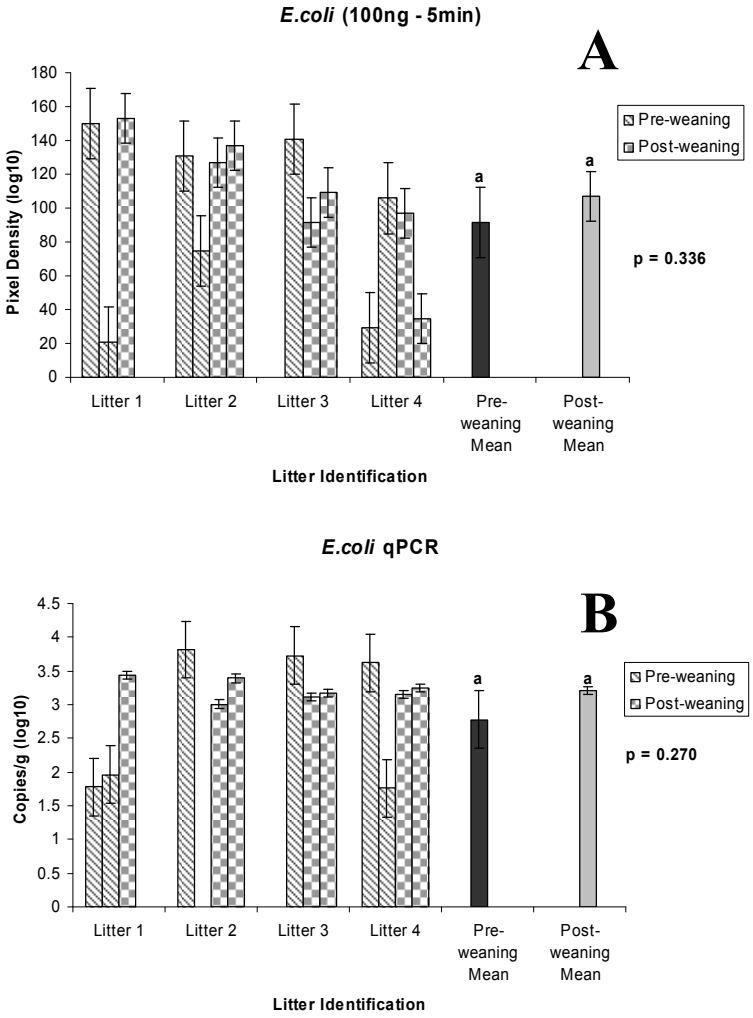

Figure 5-13. Mean (±SE) EC pixel density determined by hybridization assay or EC genome copy number determined in DNA extracted from jejunal contents of pre and postweaning pigs. Panel A shows pixel density for 100 ng target DNA after 5 minutes X-ray film exposure. Panel B shows genome copy number determined by qPCR. Data for duplicate values for individual pigs within each litter is presented at left. The mean value pre and post weaning is presented at right.

6.0 Discussion

DNA hybridization arrays have been developed for a multitude of purposes including, as common examples; gene expression assays, gene mutation analysis, and bacterial species identification. The purpose of this project was to develop and characterize a hybridization-based macroarray method for characterization of complex microbial communities using the target gene *cpn*60. Universal and plasmid specific primers targeting *cpn*60 were used to develop a macroarray technique able to discriminate between microorganisms that have *cpn*60 sequences that have less than 85% sequence identity. This method was used to analyze previously characterized communities found in the porcine ileum when fed diets based on several different cereal ingredients, and to examine the shifts in related microbiota in the pre and post-weaning porcine intestinal tract. qPCR was used to validate the qualitative and quantitative results obtained from the macroarray analysis method.

The vast majority of molecular phylogenetic data available for microbial communities is based on analysis of 16S rRNA as a target gene (Hill *et al.*, 2002). Universally present and functionally constant in all bacterial species, as well as containing regions of conservation and hypervariability,16S rRNA has proven to be an extremely useful target gene. However, multiple copies of 16S rRNA can be present in the same organism (Goh *et al.*, 1996), which can complicate molecular enumeration.

The majority of work using macroarray technology involves the identification of clinical diagnostic bacterial isolates (Goh *et al.*, 2000). For these arrays, a panel of individual bacterial species was applied in predetermined locations

to a positively charged membrane. Probe DNA prepared by labelling the *cpn*60 UT from unidentified bacterial isolate genomic DNA was applied to the membrane. Using stringent hybridization conditions, Goh *et al*. (2000) were able to demonstrate sequence specific hybridization of the unidentified probe *cpn*60 UT DNA with complementary target *cpn*60 UT for 17 different *Enterococcus* species and phenotypically similar *Lactococcus* and *Vagococcus* species. The nucleotide sequence identity of the *Enterococcus*, *Lactococcus* and *Vagococcus* species ranged from 69 to 88% in this study with virtually no cross hybridization observed. The *cpn*60 hybridization analysis method of Goh *et al*. (2000) was also capable of distinguishing between two *Enterococcus* species previously misidentified by traditional phenotypic identification practices.

Macroarray technology has also been used to identify microbial species present in complex communities (Jenkins *et al.*, 2004; Steward *et al.*, 2004; Wang, 2002b; Wu *et al.*, 2001); however this method can also be applied to identifying all components of complex microbial communities. DNA extracted from complex microbial communities can be labelled as the probe DNA and applied to the membrane containing known target DNA. The resulting hybridization signals on the membrane can be analyzed to identify not only the species present in the complex communities, but the intensity of the hybridization signal should give an indication of the prevalence of each species in the complex community.

The potential advantages of using an array approach to characterize complex bacterial communities include rapid identification of many components of a complex community with one array, as well as obtaining some information about the prevalence of each of those species within that community. This screening of community composition allows one to determine the identities of many of the species present within the community, rather than focusing on the

presence or absence of a single species. Disadvantages of this setup include the potential for missing a major bacterial group if it is not present on the membrane. In order to obtain a complete analysis of a complex community, some previous knowledge of the community components is required. This proves particularly difficult for unculturable organisms, which are unable to be qualified with traditional culture-based techniques. In such instances, phylogenetic data stored in reference databases may be the most useful method of identification. Labelled probes of highly related sequences may give incorrect indication of their presence due to cross hybridization. In addition, variation in hybridization kinetics make comparisons between arrays difficult, thus comparisons of hybridization intensities for individual species between several arrays may not be valid (Steward *et al.*, 2004; Wu *et al.*, 2001). The resulting hybridization intensities can be analyzed via comparison to the hybridization intensities of internal controls present on each membrane, but results can only be expressed as relative, not absolute, values.

At the beginning of this project, the intent was to use the macroarray setup described above (apply labelled probed prepared from complex community genomic DNA to membrane spotted with DNA such that each spot represented a unique potential community member). However, several difficulties regarding labelling efficiency of complex microbial PCR product and variation in hybridization signal intensity between arrays were encountered. When complex microbial communities were used as a template for PCR amplification for the labelling reaction of probe production, the robustness of the PCR appeared to be severely compromised. Very low quantities of labelled probe were produced in PCR using complex communities as a template for amplification. The application of these probes to target DNA present on the membrane typically resulted in low hybridization signal intensities. It was hypothesized that the labelling PCR was not robust enough to amplify and label probe DNA from a complex community. After several

attempts to improve the labelling reaction by both increasing and decreasing the amount of complex community DNA present in the reaction, an alternative array experimental design was established. In this project, *cpn*60 UT from complex microbial communities was immobilized on the nylon membrane as the target DNA. Labelled probe *cpn*60 UT from individual bacterial species was created for hybridization. By exchanging the DNA sources for the probe and target DNA, the robustness of the probe labelling reaction and the variation in hybridization signal intensities improved dramatically.

This experimental setup allows the examination of the prevalence of a specific bacterial group in several complex community samples with the use of only one array. However, a potential disadvantage of this method includes being restricted to the analysis of a single bacterial group on each array. To combat this, the labelled probe can by stripped from previously hybridized membranes by exposure to a denaturing buffer, which allows for the membrane to be rehybridized with another probe species. Also, there is the potential for missing an important bacterial group that is affecting the complex community due to either the lack of DNA probe available or simply neglecting to include it in the analysis. As above, some previous knowledge of the complex community under investigation is required in order for an unbiased analysis to occur. This is different than techniques such as DGGE or TGGE, where no prior information is required to analyze the complex community. Highly related sequences may be present in drastically different concentrations, but cross hybridization effects may indicate false results. Hybridization conditions remain the same as the first experimental setup outlined above – kinetic variation is high and direct comparison of hybridization intensities between different membranes is difficult.

One of the major challenges involved in developing macroarray hybridization arrays is the persistent presence of cross hybridization. In the techniques presented here, we were able to distinguish species at a level of approximately 85% sequence identity between the probe and target DNA. Common to all hybridization assays, cross hybridization can decrease the specificity of the array and thus threshold levels must be clearly defined. Little detail is known about the aspects of hybridization on a solid state at a molecular level, and this discrepancy has resulted in a wide range of cross-hybridization thresholds being defined for different hybridization techniques (Wu *et al.*, 2001). The threshold level of cross hybridization appears to be primarily dependent on the degree and location of sequence divergence between the target and probe sequences. Wu *et al*. (2005) used analysis of a free energy model to determine that cross hybridization is often caused by runs of 10-16 complementary nucleotides in a sequence, with a greater amount of cross hybridization occurring at the 5' end of the probe sequence. This observation agrees with Lu *et al*. (2000), who also observed that stretches of 15 complementary bases or more had increased cross-hybridization levels in 50mer oligonucleotide sequences during microarray analysis.

In the present study, universal *cpn*60 primers were initially used to amplify the target sequences for both probe and target DNA. Universal primers are designed based on the presence of a conserved region present in the DNA sequence. According to Wu *et al*. (2005), sequence similarity at the end regions of a hybridization target – especially similarity at the 5' end - will increase cross hybridization. The use of universal primers for amplification guarantees the presence of conserved regions at the ends of the amplicon sequence, therefore potentially contributing to cross hybridization. The use of plasmid specific primers for the generation of target DNA improved cross hybridization substantially, even though the conserved region for the

*cpn*60 universal primer sites were still present in target DNA although no longer be at the extreme 5' and 3'ends of the sequence. In agreement with Wu *et al*. (2005), it is possible that the movement of the highly homologous regions of sequence away from the end of the strand results in less cross hybridization.

It is likely that the major contribution to reduced cross hybridization when plasmid based primers were used for target DNA amplification was associated with *Escherichia coli* genomic DNA contamination of plasmid templates prepared from transformed *E. coli* cultures. Sufficient amounts of contaminating *E. coli* genomic DNA may have been present to serve as template for amplification when universal *cpn60* primers where used to amplify both the labelled probe (whether a plasmid or genomic DNA extracted from digesta was used as template) and target DNA (Figure 5-1). Thus both target and probe DNA could contain a considerable amount *E.coli cpn*60 UT amplicons resulting in significant cross hybridization.

Although cross hybridization was reduced by the use of plasmid specific primers, the level of cross hybridization remained unacceptable. We suspected that the high specific activity of the labelled probe would result in a considerable chemiluminescent signal detection (hybridization detection) even though only a very small number of probes may have hybridized to a particular spot as a result of cross hybridization. It is commonplace to include unlabeled, non-specific DNA in the hybridization reaction to decrease non-specific signal by occupying areas where non-specific binding may occur (example salmon sperm DNA) (Wren *et al.*, 2002). Less common, although still accepted, is the inclusion of unlabeled DNA identical or closely related to the labelled probe DNA in the hybridization reaction in an effort to decrease cross hybridization (Anamthawat-Jonsson & Heslop-Harrison, 1996; Socransky *et al.*, 2004). We employed this approach here to reduce probe specific activity, and a marked decrease in cross-hybridization

was observed (Figure 5-3). Rather than decreasing non-specific binding in general, the primary effect of this modification was likely a shift in the specific signal to noise ratio. For target DNA with low sequence identity compared to the labelled probe, the cross-hybridization probably occurred at the same level, however, as a result of the reduction in specific activity, the threshold level of DIG required for hybridization detection was not met. For target DNA with high sequence identity compared to the labelled probe, the amount of DIG associated with abundant hybridization would also be reduced by the reduction in probe specific activity. However, in this case the signal remains above the threshold level of detection with any signal dilution compensated for with longer X-ray film exposure times. This also implies that prior to reducing probe specific activity, chemiluminescent signal detection was at or near the saturation for low level cross hybridization and at or well above saturation for high level hybridization such that even with short radiographic exposure differences in spot intensity were small. This is consistent with a narrow dynamic range of signal, common in hybridization assays (Wu *et al.*, 2001).

Cross hybridization can be influenced by several additional factors, including hybridization and wash temperatures, formamide concentration, ionic strength, and sequence length. Ye *et al*. (2001) observed that different levels of specificity can be obtained by altering the hybridization temperature such that low stringency was observed at low temperatures (45^{o}C) while higher temperatures garnered better specificity (70-75^{o}C). Hybridization temperatures are typically determined based on several factors, including G + C content, length of probe and target sequence, and cation concentration within the hybridization buffer (Stahl & Amann, 1991). Steward *et al*. (2004) found that a higher concentration (40%) of formamide in the hybridization buffer also produced the most stringent conditions. The presence of formamide in the hybridization buffer helps to improve hybridization specificity by maintaining dissociation of

DNA, thereby decreasing the temperature required for hybridization to occur. The methodology reported here was initially based on previously determined hybridization protocols designed for the target gene *cpn*60 (Goh *et al.*, 1996). The hybridization buffer contained 50% formamide and the hybridization temperature was 42^{o}C. It is unusual for hybridization reactions to occur at such high temperatures when a high proportion of formamide is present, as a common problem is loss of target DNA from the membrane. DNA loss from the membrane was avoided here by UV crosslinking DNA to the membrane prior to dehybridization. An ideal solution to the issue of choosing a temperature would involve the calculation of model hybridization temperatures for each individual blot based on the sequence of the individual labelled probe being used (Stahl & Amann, 1991), however it is much more commonplace to adopt a standard set of hybridization parameters as seen here. The parameters used here were used to identify bacterial species with up to 95% sequence identity (Goh *et al.*, 1997) and are comparable to other hybridization techniques found in the literature (Steward *et al.*, 2004; Wang *et al.*, 2002b).

As depicted in Figure 5-5A, the relationship between signal intensity and sequence identity between target and probe DNA was not linear. This variation in target DNA behaviour has also been observed by Steward *et al*. (2004) and is considered to be indicative of considerable variability in the behaviour of target DNA from different bacterial sources. Possible explanations include variation in G + C content and the distribution of mismatches throughout the target gene sequence. While no major differences were observed in the G + C content of the organisms examined here (data not shown), it is possible that even slight differences in G + C concentration may have been a contributing factor. In addition, Bodrossy and Sessitsch (2004) suggested that hybridization is a more efficient method of detection when the mismatches are distributed more among the middle of the sequence, rather than the ends. Although some of the

sequences used had less sequence variation present in the central area than on the ends, substantial variation was present throughout the entire sequence (data not shown). However, even slight discrepancies in the distribution of sequence variation between different target DNA species may contribute to the variation in target DNA hybridization intensities observed.

In microarray analysis developed for the characterization of Arabidopsis seeds gene sequences, cross-hybridization was observed in genes with 70 – 80% sequence identity (Girke *et al.*, 2000). For macroarray development, cross-hybridization of target and probe at approximately 94% sequence identity was observed for analysis of closely related genes (Miller *et al.*, 2002). Steward *et al*. (2004) observed cross-hybridization at 78-88% sequence identity when developing an array to evaluate nitrogenase gene diversity. Our experimentation showed a cross-hybridization threshold of approximately 85% sequence identity, which is comparable to most other studies using hybridization technology. We also observed that signal intensity was markedly lower at up to 90% sequence identity between probe and target DNA. However, in previous work utilizing *cpn*60 as a target gene for bacterial isolate identification in a "cross checkerboard" nylon membrane hybridization scheme, no cross hybridization was observed for sequences less than 94% identical (Goh *et al.*, 1996; Goh *et al.*, 1997; Goh *et al.*, 1998). This value is significantly higher than the cross-hybridization threshold observed in this project. While similar hybridization and washing protocols were used, a potential explanation for the discrepancy in cross-hybridization threshold levels is found in the templates used to create both targets and probes. Previous work (Goh *et al.*, 1997) used *cpn*60 universal primers to amplify both target and probe DNA using genomic DNA as template. In contrast, in this project plasmid-based primers were used generate probe DNA using plasmids containing *cpn*60 inserts as templates. A second difference noted between previous work and the work presented here is the

amount of target DNA present on the membrane and X-ray film exposure time used to determine the specificity of this methodology. While Goh *et al*. (1997) applied approximately 1.2 ng of target DNA per spot and an X-ray film exposure time of 20-30 minutes, we used 50 – 500 pg of target DNA per spot and X-ray film exposure times of approximately 15 minutes because we found target DNA amounts greater than 500 pg became saturated too quickley. This indicates we obtained a potentially higher labelling efficiency as we were able to detect lower amounts of target DNA in a shorter time period. A greater amount of labelled probe DNA may have been used in the methodology presented here, however the amount of probe DNA used by Goh *et al*. (1997) was not presented. These variations in methodology appear to have resulted in a lower level of specificity.

While hybridization kinetics have been studied intensively (Wetmur, 1976), it is a highly complex area and little is known about the phenomenon at a molecular level when one phase is attached to a solid support. One of the major difficulties of using hybridization as both a detection tool and a profiling tool is the limited dynamic quantification range of the assay. Typically, a linear relationship would be expected between the concentration of target DNA present on the membrane that is identical to the labelled probe and the intensity of the signal generated, thus rendering hybridization an accurate quantitative technique. As depicted by the signal intensities observed for the dilution series for each bacterial species, signal intensity is greatly dependent on target DNA concentration (example: Figure 5-6). However, studies indicate that while a linear relationship is predictable for a narrow range of target DNA, a linear relationship is not observed for a wide range of concentrations of target DNA (Wu *et al.*, 2001). In addition, we were faced with the challenge of detecting various bacterial groups that were present in unknown abundance. As more abundant groups required less target DNA to avoid

signal saturation, while less abundant groups required more target DNA to produce a detectable signal it was difficult to determine a fixed amount of target DNA to be applied to the membrane. In order to improve the quantitative range of the assay, a dilution series of each target DNA was placed on the membrane. This ensured a variable amount of probe-specific target DNA at each spot and enhanced the opportunity to obtain a quantitative chemiluminescent signal that was below saturation but above background (i.e. within dynamic range). In agreement with other studies (Steward *et al.*, 2004; Wu *et al.*, 2001), we observed that quantitative data was difficult to obtain as higher amounts of target DNA from both simple and complex communities quickly reach saturation of signal intensity. Exposing X-ray film for a variable amount of time helped to acquire useful signal intensity data for quantitative analysis, as shorter exposure times allowed signal intensity data to be obtained before saturation was reached for most highly abundant species. However, another issue surrounding the lack of linearity between target DNA concentration and signal intensity of dilution series of simple communities may possibly be attributed to the non-linear response of light detecting films in reaction to light emission (Mansfield *et al.*, 1995). Due to the kinetics of chemiluminescent signal decay, light-detecting films used for chemiluminescent detection do not respond in a linear fashion, thereby making both comparisons of signal intensity between different arrays and quantitative measurements from individual arrays difficult at best.

Other issues surrounding generation of quantifiable macroarray data concern the indirect detection method used here. In this method, DIG-dUTP is incorporated into probe DNA during the PCR reaction. Anti-DIG antibodies that are conjugated with alkaline phosphatase are exposed to the hybridized probe. The alkaline phosphatase catalytically dephosphorylates CSPD, which results in the production of light. Indirect detection methods such as this provide

higher sensitivity than direct detection methods (Mansfield *et al.* 1995). While this high sensitivity may be highly desirable in applications where the presence of an individual species is being detected (i.e. pathogen detection), it may contribute to the quantitative measurements able to be generated using this technology. In addition, increased probe density bound to an area of target DNA can potentially affect the linearity of signal detected due to steric hindrance. That is, an extremely high density of probe DNA hybridized to a small surface area may not result in a single anti-DIG-dUTP antibody being bound to every molecule of DIG-dUTP present due to physical space limitations. This limits the amount of signal generated for that area of target DNA. It is possible that an area of less concentrated target DNA can generate the same amount of signal intensity as the antibodies bound to the hybridized probe DNA may not be affected by steric hindrance, thus affecting the linearity of signal intensity generated.

There are several options that could potentially be explored to increase the quantitative range of the methodology used here. As chemiluminescent detection is highly sensitive, the application of a tighter dilution series on each membrane may expand the narrow linear range observed. In addition, the application of more dilutions and smaller amounts of target DNA to the membrane may increase the generation of quantifiable data, as a larger range of target DNA amounts is available for analysis. In addition, a less sensitive detection method such as colourmetric detection or a direct detection method may generate a broader linear range more suitable for quantification of data.

While variations in cross-hybridization and signal intensity as related to target amount are noted, this method may still be used to characterize complex microbial communities. However, it is difficult to distinguish differences in signal intensity due to population abundance from those observed due to cross-hybridization from variation in sequence divergence. Ideally, complex

communities show differences in hybridization signal intensities as a result of variation in the concentration of probe species present in that community. However, cross hybridization between closely related species will also result in variation in signal intensities dependent upon their abundance. It becomes an issue of determining the amount of signal variation attributed to probe species abundance as opposed to variation attributed to cross hybridization. In order to determine the amount of signal intensity that can be attributed to cross hybridization, We *et al*. (2001) suggests performing hybridization under a range of stringency levels, as increasing stringency would decrease cross-hybridization thereby giving an estimation of the proportion of signal intensity due to cross-hybridization. While cross-hybridization may slightly influence the detection of some bacterial targets in some environmental communities, this technology appears to produce reliable estimations for the proportions of various bacterial groups even without compensating for variations in signal intensity due to cross hybridization. In addition, small amounts of cross hybridization may be observed as an advantage if this method is to be used as a wide identification method for broad ranges of bacterial species or taxa in complex microbial communities. Lower specificity of hybridization could ensure the detection of many bacterial species in a closely related group where broad-based community compositional changes are of interest. This could possibly serve as an initial screening method followed by more detailed analysis of the composition of the taxa identified as affected. In comparison, highly stringent hybridization methodology will be highly accurate for individual species identification (for example, pathogen identification); however shifts in broad taxa will not be detected. Using the technology developed here, shifts in broad taxa can be analyzed up to a sequence specificity of approximately 85%. In order to obtain higher specificity for individual species identification,

more stringent hybridization parameters, a different detection system, and potentially a different target gene could all be explored.

While quantitative data from macroarray technology is difficult to acquire at best, qPCR is a highly specific enumeration approach for determining the components of complex microbial communities. Comparison of signal intensity data from macroarray methodology with qPCR results for the CWB analysis showed high correlation for all bacterial groups examined except for the *Lactobacillus amylovorous* group (Table 5-3). The discrepancy observed for *Lactobacillus amylovorous* group may be attributed to the extreme abundance of the *Lactobacillus* species within the pig ileum (Hill *et al.*, 2005). Because of the high abundance, the *Lactobacillus amylovorous* group appeared to saturate chemiluminescent detection even with the lowest amounts of target DNA applied to the membrane and very short X-ray film exposure times. Further dilution of the sample target DNA for the *Lactobacillus amylovorous* group may provide a more accurate quantitative analysis of community composition. It should be noted that qPCR results determined for the presence of *Lactobacillus amylovorous* and related species also did not correlate with library clone frequency results (Hill *et al.*, 2005), also probably due to high abundance.

We chose to assess further the utility of the macroarray methodology described here by characterizing the complex microbial community found in the pre and post-weaning jejunum of neonatal piglets. While little knowledge exists about the communities located in the jejunum of the pig, previous work (Swords *et al.*, 1993) has indicated that there is a considerable shift in the components of the microbial communities of the pig gut when weaning occurs. As with the CWB analysis, a dilution series was included on the arrays used to analyze the microbial community found in the jejunum of pre and post-weaning piglets. However, a lack of linearity in

this dilution series hindered efforts to quantify hybridization results. That is, the densitometric analysis of the signal intensity generated by the dilution series did not produce suitable results to generate a linear standard curve to provide an estimate of the amount probe species present in the jejunal samples. While qPCR results correlated well with signal intensities from the *Lactobacillus amylovorous* group, a slightly lower correlation was observed for the *Lactobacillus vaginalis* group. qPCR results for the *Streptococcus*-like group and *E. coli* group showed low correlation with hybridization signal intensities. As macroarray technology presented here has a cross-hybridization threshold of approximately 85%, it is possible that one related species present in the jejunum of pre and post-weaning piglets contributed to the signal intensity observed for the *Streptococcus*-like probe. Use of the more specific method of qPCR diagnostic for *Streptococcus*-like would not have detected this closely related organism due to the higher specificity attained with qPCR. In an effort to examine this discrepancy further, qPCR analysis was conducted for the target *Streptococcus alactolyticus*, a species with 97% sequence identity to the *Streptococcus*-like organism. However, qPCR results indicated the same profile as the *Streptococcus*-like qPCR assay indicating that variation in colonization by *S. alactolyticus* does not account for the poor correlation between array and qPCR quantification of *Streptococcus*-like group. It is quite possible that another closely related species present in the sample could account for quantitative discrepancy between a *Streptococcus*-like array and PCR. *S. alactolyticus* was identified as a candidate organisms based on the abundance of this organism in previous investigations of pig ileal populations using the sequence library approach (Hill *et al.*, 2002; Hill *et al.*, 2005; Smith, 2006). Based on these studies, other relatively abundant *Streptococcus* species that could have contributed to *Streptococcus*-like signal intensity using the array include *S. thermophilus* and *S. orisratti*. The microbiota residing in the porcine jejunum

has not been significantly studied, thus there may be several other bacterial species present in sufficient amounts that are closely related to all of the groups examined that may alter hybridization results. These species may be identified as further molecular-based investigations of the pig jejunal microbiota continue.

While the macroarray methodology presented here does not provide completely quantitative data for individual bacterial species, it does provide a more cost-efficient method to obtain semi-quantitative analysis of the abundance of various species in microbial communities when compared to costly, labour intensive methods such as clone library construction. It may be used either as an alternative or as a supplement to other broad characterization methods such as DGGE, TGGE, library construction, and qPCR. Our approach provides the advantage of analyzing the abundance of a single target group in many environmental samples at once in a more broad and inexpensive manner than highly specific techniques such as qPCR. Macroarray hybridization could offer a more efficient method to evaluate shifts in broad microbial groups, as methods such as DGGE and TGGE require the excision and subsequent sequencing of genomic bands of interest. While the method does not provide the wealth of sequence and phylogenetic data acquired by plasmid library construction of a complex microbial community, macroarray hybridization is much more efficient, less labour intensive, and a more cost effective means of obtaining information on the microbial shifts in different environments. The reliable, semi-quantitative results of the macroarray methodology presented here gives a general overview of the abundance of species under examination, however the presence of significant cross hybridization in species greater than 85% sequence identity, narrow dynamic range of the technique, and semi-quantitative parameters are not functional at times when highly specific and accurate measurements of individual organisms are required.

7.0 Conclusions

The objective of this project was to develop and characterize macroarray hybridization methods for the qualitative and quantitative characterization of complex microbial communities. With *cpn*60 as the target gene selected for profiling analysis, we defined the cross-hybridization threshold for this technology using target DNA amplified from plasmids containing *cpn*60 inserts with *cpn*60 universal primers for immobilization on a positively charged membrane. DIG-dUTP-labelled probe was created using plasmid-specific primers. The inclusion of unlabeled DNA identical to the labelled probe in the hybridization mixture acted to lower probe specific activity and thus improved the cross-hybridization threshold to approximately 85% sequence identity with low amounts of cross-hybridization observed to 90% sequence identity. Signal intensities generally correlated with the amount of target DNA present on the membrane, but densitometric analysis of target DNA of known sequence indicated some variation between the relationship of signal intensity and target DNA amounts. While a linear relationship between target DNA amount and signal intensity is predictable for a narrow range of loading, a linear relationship is not observed over a wide range of loading. Thus, the results produced by macroarray technology are semi-quantitative at best. For future work, linear range could potentially be expanded by spotting a tighter dilution series of target DNA onto the membrane. Direct detection methods or colormetric detection, which are generally regarded as being less sensitive detection methods, could also be examined as the lower sensitivity may increase the linear range of detectable response. This technology was used to determine the presence of several bacterial groups in ileal contents extracted from pigs fed diets based on corn, wheat, or

barley, and in the jejunum of pre- and post-weaning piglets. The macroarray results were correlated with qPCR results. While high correlation hybridization and qPCR data was observed for most bacterial groups analyzed, groups present in extremely high abundance showed poor correlation. In addition, some bacterial groups analyzed in the pre and post-weaning pig jejunum also showed poor correlation, possibly due to the presence of an unidentified species that was detected by macroarray technology, but not detected with the more specific qPCR technique.

In this project, macroarray technology was able to provide some information regarding the shift of gastrointestinal microflora in the piglet intestine following weaning. Macroarray technology detected a notable decrease in both *Lactobacillus amylovorous*-related species and *Lactobacillus vaginalis*-related species in post-weaning piglets, both of which correlated with qPCR results. These changes may be indicative of increased microbial diversity in the gastrointestinal environment following the transition onto solid feed. In addition, *E. coli* levels detected by macroarray technology and qPCR were considerably lower in both pre and post-weaning piglets than other bacterial groups characterized. Unlike the levels of *L. amylovorous*, the levels of *E. coli* did not appear to decrease following weaning.

Macroarray technology gives semi-quantitative results about proportional presence of several bacterial groups found in several microbial communities. Though not as stringent and detailed as techniques such as qPCR, macroarray technology provides a relatively inexpensive, though still reliable method for the characterization of complex microbial communities. The broad profiling ability of macroarray technology may be used to recognize changes in microbial communities in response to numerous stimuli, thereby identifying taxa where further analysis may be warranted.

8.0 References

Acinas, S., Klepac-Ceraj, V., Hunt, D., Pharino, C., Ceraj, I., Distel, D. & Polz, M. (2004). Fine-scale phlygenetic architecture of a complex bacterial community. *Nature* **430,** 551-554.

Amann, R. I., L. Krumholz & Stahl, D. A. (1990). Flourescent-oligonucleotide probing of whole cells for deteremmentive, phylogenetic, and environmental studies in microbiology. *Journal of Bacteriology* **172**, 762-770.

Amann, R. I., Ludwig, W. & Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of microbial cells without cultivation. *Microbiological Reviews* **59**, 143-169.

Anamthawat-Jonsson, K. & Heslop-Harrison, J. (1996). Establishing relationships between genomic DNA as a probe. *Methods Mol Bio* **50**, 209-225.

Apajalahti, J. H. A., Sarkilahti, L. K., Maki, B. R. E., Heikkinen, J. P., Nurminen, P. H. & Holben, W. E. (1998). Effective Recovery of Bacterial DNA and Percent-Guanine-Plus-Cytosine-Based Analysis of Community Structure in the Gastrointestinal Tract of Broiler Chickens. *Appl Environ Microbiol* **64**, 4084-4088.

Apajalahti, J. H. A., Kettunen, A., Bedford, M. R. & Holben, W. E. (2001). Percent G+C Profiling Accurately Reveals Diet-Related Differences in the Gastrointestinal Microbial Community of Broiler Chickens. *Appl Environ Microbiol* **67**, 5656-5667.

Apajalahti, J. H. A., Kettunen, A., Nurminen, P. H., Jatila, H. & Holben, W. E. (2003). Selective Plating Underestimates Abundance and Shows Differential Recovery of Bifidobacterial Species from Human Feces. *Appl Environ Microbiol* **69**, 5731-5735.

Bach, H. J., Tomanova, J., Schloter, M. & Munch, J. C. (2002). Enumeration of total bacteria and bacteria with genes for proteolytic activity in pure cultures and in environmental samples by quantitative PCR mediated amplification. *J Microbiol Methods* **49**, 235-245.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. (2005). GenBank. *Nucl Acids Res* **33**, D34-38.

Berndt, C., Bebenroth, M., Oehlschlegel, K., Hiepe, F. & Schossler, W. (1995). Quantitative Polymerase Chain Reaction Using a DNA Hybridization Assay Based on Surface-Activated Microplates. *Analytical Biochemistry* **225**, 252.

Blank, R., Sauer, W. C., Mosenthin, R., Zentek, J., Huang, S. & Roth, S. (2001). Effect of Fumic Acid Supplementation and Dietary Buffering Capacity on the Concentration of Microbial Metabolites in Ileal Digesta of Young Pigs. *Can J Anim Sci* **81**, 345-353.

Bodrossy, L. & Sessitsch, A. (2004). Oligonucleotide microarrays in microbial diagnostics. *Current Opinion in Microbiology* **7**, 245.

Bourne, D. & Munn, C. (2005). Diversity of bacteria associated with the coral *Pocillopora damiconis* from the Great Barrier Reef. *Enivron Micro* **7**, 1162-1174.

Brousseau, R., Hill, J. E., Prefontaine, G., Goh, S.-H., Harel, J. & Hemmingsen, S. M. (2001). Streptococcus suis Serotypes Characterized by Analysis of Chaperonin 60 Gene Sequences. *Appl Environ Microbiol* **67**, 4828-4833.

Cabrita, A. R. J., Fonseca, A. J. M., Dewhurst, R. J. & Gomes, E. (2003). Nitrogen Supplementation of Corn Silages. 2. Assessing Rumen Function Using Fatty Acid Profiles of Bovine Milk. *J Dairy Sci* **86**, 4020-4032.

Castillo, M., Martin-Orue, S. M., Manzanilla, E. G., Badiola, I., Martin, M. & Gasa, J. (2006). Quantification of total bacteria, enterobacteria and lactobacilli populations in pig digesta by real-time PCR. *Veterinary Microbiology* **114**, 165-175.

Chamberlain, J., Gibbs, R., Ranier, J., Nguyen, P. & Caskey, C. (1988). Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Research* **16**, 11141-11156.

Cheneby, D., Hallet, M., Mondon, M., Martin-Laurent, F., Germon, J. & Philippot, L. (2003). Gentic Characterization of the Nitrate Reducing Community Based on *narG* Nucleotide Sequence Analysis. *Microbial Ecology* **46,** 113-121.

Chiang, Y. C., Yang, C.Y., Li, C., Ho, Y.C., Lin, C.K. & Tsen, H.Y. (2006). Identification of Bacillus spp., Escherichia coli, Salmonella spp., Staphylococcus spp. and Vibrio spp. with 16S ribosomal DNA-based oligonucleotide array hybridization. *International Journal of Food Microbiology* **107,** 131.

Christensen, H., Moller, P. L., Vogensen, F. K. & Olsen, J. E. (2000). Sequence variation of the 16S to 23S rRNA spacer region in Salmonella enterica. *Research in Microbiology* **151,** 37.

Clarridge, J. E., III (2004). Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases. *Clin Microbiol Rev* **17**, 840-862.

Clayton, R. A., Sutton, G., Hinkle, P. S., Bult, C. & Fields, C. (1995). Intraspecific variation in small-subunit rRNA sequences in GenBank: why single sequences may not adequately represent prokaryotic taxa. *International Journal of Systematic Bacteriology* **45**, 595-599.

Coenye, T. & Vandamme, P. (2003). Intragenomic heterogeneity between multiple 16S ribosomal RNA operons in sequenced bacterial genomes. *FEMS Microbiology Letters* **228**, 45-49.

Collier, C. T., Smiricky-Tjardes, M. R., Albin, D. M., Wubben, J. E., Gabert, V. M., Deplancke, B., Bane, D., Anderson, D. B. & Gaskins, H. R. (2003). Molecular ecological analysis of porcine ileal microbiota responses to antimicrobial growth promoters. *J Anim Sci* **81**, 3035-3045.

Countway, P. D. & Caron, D. A. (2006). Abundance and Distribution of Ostreococcus sp. in the San Pedro Channel, California, as Revealed by Quantitative PCR. *Appl Environ Microbiol* **72**, 2496-2506.

Dahllof, I., Baillie, H. & Kjelleberg, S. (2000). rpoB-Based Microbial Community Analysis Avoids Limitations Inherent in 16S rRNA Gene Intraspecies Heterogeneity. *Appl Environ Microbiol* **66**, 3376-3380.

Dumonceaux, T. J., Hill, J. E., Briggs, S. A., Amoako, K. K., Hemmingsen, S. M. & Van Kessel, A. G. (2006). Enumeration of specific bacterial populations in complex intestinal communities using quantitative PCR based on the chaperonin-60 target. *Journal of Microbiological Methods* **64**, 46.

Ellis, R. J. & Van der Vies, S. M. (1991). Molecular chaperones. *Annual Review of Biochemistry* **60**, 321-347.

Ercolini, D. (2004). PCR-DGGE fingerprinting: novel strategies for detection of microbes in food. *Journal of Microbiological Methods* **56**, 297.

Favier, C. F., Vaughan, E. E., De Vos, W. M. & Akkermans, A. D. L. (2002). Molecular Monitoring of Succession of Bacterial Communities in Human Neonates. *Appl Environ Microbiol* **68**, 219-226.

Felske, A., Pauling, B. V., von Canstein, H. F., Li, Y., Lauber, J., Buer, J. & Wagner, D. I. (2001). Detection of small sequence differences using competitive PCR: Molecular monitoring of genetically improved, mercury-reducing bacteria. *Biotechniques* **30**, 142-148.

Fingegold, S. M. & Sutter, V. L. (1978). Fecal flora in different populations, with special reference to diet. *American Journal of Clinical Nutrition* **31**, S116-S122.

Fortin, N., Mulchandani, A. & Chen, W. (2001). Use of real-time polymerase chain reaction and molecular beacons for the detection of Escherichia coli O157: H7. *Analytical Biochemistry* **289**, 281-288.

Fox, G. E., Wisotzkey, J. D. & Jurtshuk, P. (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *International Journal of Systematic Bacteriology* **42**, 166-170.

Gabig, M. & Wegrzyn, G. (2001). An introduction to DNA chips: principles, technology, applications, and analysis. *Acta Biochimica Polonica* **48**, 615-622.

Garcia, E. P., Dowding, L. A., Stanton, L. W. & Slepnev, V. I. (2005). Scalable Transcriptional Analysis Routine--Multiplexed Quantitative Real-Time Polymerase Chain Reaction Platform for Gene Expression Analysis and Molecular Diagnostics. *J Mol Diagn* **7**, 444-454.

Giacomazzi, S., Leroi, F., L'Henaff, C. & Joffraud, J. J. (2004). rpoB-PCR amplified gene and temporal temperature gradient gel electrophoresis: a rapid tool to analyse bacterial strains representative of cold-smoked salmon microflora. *Letters in Applied Microbiology* **38**, 130-134.

Giovannoni, S. J., DeLong, E. F., Olsen, G. J. & Pace, N. R. (1988). Phylogenetic group-specific oligo-deoxynucleotide probes for identification of single microbial cells. *Journal of Bacteriology* **170**, 720-726.

Girke, T., Todd, J., Ruuska, S., White, J., Benning, C. & Ohlrogge, J. (2000). Microarray Analysis of Developing Arabidopsis Seeds. *Plant Physiol* **124**, 1570-1581.

Goh, S., Potter, S., Wood, J., Hemmingsen, S., Reynolds, R. & Chow, A. (1996). HSP60 gene sequences as universal targets for microbial species identification: studies with coagulase-negative staphylococci. *J Clin Microbiol* **34**, 818-823.

Goh, S. H., Santucci, Z., Kloos, W. E., Faltyn, M., George, C. G., Driedger, D. & Hemmingsen, S. M. (1997). Identification of Staphylococcus species and subspecies using the Chaperonin-60 gene identification method and reverse checkerboard hybridization. **35**, 3116-3121.

Goh, S. H., Driedger, D., Gillett, S. & other authors (1998). Streptococcus iniae, a human and animal pathogen: specific identification by the Chaperonin-60 gene identification method. **36**, 2164-2166.

Goh, S. H., Facklam, R. R., Chang, M. & other authors (2000). Identification of *Enterococcus* species and phenotypically similar *Lactococcus* and *Vagococcus* species by reverse checkerboard hybridization to chaperonin 60 gene sequences. *Journal of Clinical Microbiology* **38**, 3953–3959.

Green, B. D. & Keller, M. (2006). Capturing the uncultivated majority. *Current Opinion in Biotechnology* **17**, 236.

Greiner, O., Day, P., Bosshard, P., Imeri, F., Altwegg, M. & Nadal, D. (2001). Quantitative detection of *Streptococcus pneumoniae* in nasopharyngeal secretions by real-time PCR. *Journal of Clinical Microbiology* **39**, 3129-3134.

Gupta, R. S. (1995). Evolution of the chaperonin families (Hsp60, Hsp10, and Tcp-1) of the proteins and the origin of eukaryotic cells. *Molecular Microbiology* **15**, 1-11.

Hallberg, R. L. (1990). A mitochondrial chaperonin: genetic, biochemical and molecular characteristics. *Seminars in Cell Biology* **1**, 37-46.

Hartemink, R. & Rombouts, F. M. (1999). Comparison of media for the detection of bifidobacteria, lactobacilli and total anaerobes from faecal samples. *Journal of Microbiological Methods* **36**, 181.

Hayashi, H., Takahashi, R., Nishi, T., Sakamoto, M. & Benno, Y. (2005). Molecular analysis of jejunal, ileal, caecal and recto-sigmoidal human colonic microbiota using 16S rRNA gene libraries and terminal restriction fragment length polymorphism. *J Med Microbiol* **54**, 1093-1101.

Heid, C., Stevens, J., Livak, K. J. & Williams, P. M. (1996). Real Time quantitative PCR. *Genome Res* **6**, 986-994.

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

Hemmingsen, S. M., Woolford, C., Van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W. & Ellis, R. J. (1988). Homologous plant and bacterial proteins chaperone oligometric protein assembly. *Nature* **333**, 330-334.

Hernandez, J., Ferrus, M. A., Hernandez, M. & Owen, R. J. (1997). Arbitrary primed PCR fingerprinting and serotyping of clinical Pseudomonas aeruginosa strains. *FEMS Immunology and Medical Microbiology* **17**, 37.

Hill, J. E., Seipp, R. P., Betts, M., Hawkins, L., Kessel, A. G. V., Crosby, W. L. & Hemmingsen, S. M. (2002). Extensive profiling of a complex microbial community by high-throughput sequencing. *Applied and Environmental Microbiology* **68**, 3055-3066.

Hill, J. E., Penny, S. L., Crowell, K. G., Goh, S. H. & Hemmingsen, S. M. (2004). cpnDB: A Chaperonin Sequence Database. *Genome Res* **14**, 1669-1675.

Hill, J. E., Hemmingsen, S. M., Goldade, B. G., Dumonceaux, T. J., Klassen, J., Zijlstra, R. T., Goh, S. H. & Van Kessel, A. G. (2005). Comparison of Ileum Microflora of Pigs Fed Corn-, Wheat-, or Barley-Based Diets by Chaperonin-60 Sequencing and Quantitative PCR. *Appl Environ Microbiol* **71**, 867-875.

Hill, J. E., Town, J. R. & Hemmingsen, S. M. (2006). Improved template represetation in *cpn*60 polymerase chain reaction (PCR) product libraries generated from complex templates by application of a specific mixture of PCR primers. *Enivron Micro* **8**, 741-746.

Holben, W.E., Jansson, J.K., Chelm, B.C. & Tiedje, J.M. (1988) DNA probe method for the detection of specific microorganisms in the soil bacterial community. *Appl Environ Microbiol* **54(3)**, 708-711.

Holdeman, L. V. & Moore, W. E. (1972). Roll-tube techniques for anaerobic bacteria. *American Clinical Nutrition* **25**, 1314-1317.

Hooper, L. V. & Gordon, J. I. (2001). Commensal Host-Bacterial Relationships in the Gut. *Science* **292**, 1115-1118.

Hooper, L. V., Wong, M. H., Thelin, A., Hansson, L., Falk, P. G. & Gordon, J. I. (2001). Molecular analysis of commensal host-microbial relationships in the intestine. *Science* **291**, 881-884.

Hopkins, M. J., Sharp, R. & Macfarlane, G. T. (2001). Age and disease related changes in intestinal bacterial populations assessed by cell culture, 16S rRNA abundance, and community cellular fatty acid profiles. *Gut* **48**, 198-205.

Hugenholtz, P., Goebel, B. M. & Pace, N. R. (1998). Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. *J Bacteriol* **180**, 4765-4774.

Huijsdens, X., Linskens, R., Mak, M., Meuwissen, S., Vandenbroucke, G. C. & Savelkoul, P. (2002). Quantification of bacteria adherent to gastrointestinal mucosa by real-time PCR. *Journal of Clinical Microbiology* **40**, 4423-4427.

Jaspers, E. & Overmann, J. (2004). Ecological Significance of Microdiversity: Identical 16S rRNA Gene Sequences Can Be Found in Bacteria with Highly Divergent Genomes and Ecophysiologies. *Appl Environ Microbiol* **70**, 4831-4839.

Jenkins, B. D., Steward, G. F., Short, S. M., Ward, B. B. & Zehr, J. P. (2004). Fingerprinting Diazotroph Communities in the Chesapeake Bay by Using a DNA Macroarray. *Appl Environ Microbiol* **70**, 1767-1776.

Kane, M. D., Jatkoe, C., Stumpf, J., Lu, J., Thomas, J. D. & Madore, S. J. (2000). Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucl Acids Res* **28**, 4552-4557.

Knudsen, K.E. (2001). Development of antibiotic resistance and options to replace antimicrobials in animal diets. *Proc. Nutr. Soc.* **60**, 291-299.

Koesters, K., Riffelmann, M. & von, K. C. H. W. (2001). Evaluation of a real-time PCR assay for detection of Bordetella pertussis and B. parapertussis in clinical samples. *Journal of Medical Microbiology* **50,** 436-440.

Kong, T., Coates, A., Butcher, P., Hickman, C. & Shinnick, T. (1993). Mycobacterium tuberculosis Expresses Two Chaperonin-60 Homologs. *PNAS* **90**, 2608-2612.

Konstantinov, S. R., Awati, A., Smidt, H., Williams, B. A., Akkermans, A. D. L. & de Vos, W. M. (2004). Specific Response of a Novel and Abundant Lactobacillus amylovorus-Like Phylotype to Dietary Prebiotics in the Guts of Weaning Piglets. *Appl Environ Microbiol* **70**, 3821-3830.

Kullen, M. J., Brady, L. J. & O'Sullivan, D. J. (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.

La, T., Phillips, N. D. & Hampson, D. J. (2003). Development of a duplex PCR assay for detection of Brachyspira hyodysenteriae and Brachyspira pilosicoli in pig feces. *J Clin Microbiol* **41**, 3372-3375.

Lalles , J.P., Bosi, P., Smidt, H., and Stokes, C.R. (2007) Nutritional management of gut health in pigs around weaning. *Proc Nutr Soc* **66**, 260-268.

Langer, T. & Neupert, W. (1991). Heat-shock proteins hsp60 and hsp70: their roles in folding, assembly and membrane translocations of proteins. *Current Topics in Microbiology and Immunology* **167**, 3-29.

Leser, T. D., Lindecrona, R. H., Jensen, T. K., Jensen, B. B. & Moller, K. (2000). Changes in Bacterial Community Structure in the Colon of Pigs Fed Different Experimental Diets and after Infection with Brachyspira hyodysenteriae. *Appl Environ Microbiol* **66**, 3290-3296.

Leser, T. D., Amenuvor, J. Z., Jensen, T. K., Lindecrona, R. H., Boye, M. & Moller, K. (2002). Culture-Independent Analysis of Gut Bacteria: the Pig Gastrointestinal Tract Microbiota Revisited. *Appl Environ Microbiol* **68,** 673-690.

Liong, M. T. & Shah, N. P. (2005). Optimization of cholesterol removal, growth and fermentation patterns of Lactobacillus acidophilus ATCC 4962 in the presence of mannitol, fructo-oligosaccharide and inulin: a response surface methodology approach. *Journal of Applied Microbiology* **98**, 1115-1126.

Lu, J., Kane, M. D., Jatkoe, T. A., Stumpf, C. R., Thomas, J. D. & Madore, S. J. (2000). Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Research* **28**, 4552.

Lubeck, M., Bulat, S., Alekhina, I. & Lieckfeldt, E. (2004). Delineation of species within the Trichoderma viride/atroviride/koningii complex by UP-PCR cross-blot hybridization. *FEMS Microbiology Letters* **237**, 255.

Ludwig, W. & Schliefer, K. (2000). How quantitative is quantitative PCR with respect to cell counts. *Systematic and Applied Microbiology* **23**, 556-562.

Malinen, E., Kassinen, A., Rinttila, T. & Palva, A. (2003). Comparison of real-time PCR with SYBR Green I or 5'-nuclease assays and dot-blot hybridization with rDNA-targeted oligonucleotide probes in quantification of selected faecal bacteria. *Microbiology* **149**, 269-277.

Mansfield, E. S., Worley, J. M., McKenzie, S. E., Surrey, S., Rappaport, E. & Fortina, P. (1995). Nucleic acid detection using non-radioactive labelling methods. *Molecular and Cellular Probes* **9**, 145.

Markoulatos, P. (2002). Multiplex polymerase chain reaction: A practical approach. *Journal of Clinical Laboratory Analysis* **16**, 47-51.

Marsh, T. L. (1999). Terminal restriction fragment length polymorphism (T-RFLP): An emerging method for characterizing diversity among homologous populations of amplification products. *Current Opinion in Microbiology* **2**, 323.

Martel, R., Cloney, L. P., L.E., P. & Hemmingsen, S. M. (1990). Unique composition of plastid chaperonin-60: alpha and beta polypeptide-encoding genes and highly divergent. *Gene* **94**, 181-187.

Masson, L., Maynard, C., Brousseau, R., Goh, S.-H., Hemmingsen, S. M., Hill, J. E., Paccagnella, A., Oda, R. & Kimura, N. (2006). Identification of pathogenic Helicobacter species by chaperonin-60 differentiation on plastic DNA arrays. *Genomics* **87**, 104.

Mathew, A. G., Chattin, S. E., Robbins, C. M. & Golden, D. A. (1998). Effects of a direct-fed yeast culture on enteric microbial populations, fermentation acids, and performance of weanling pigs. *J Anim Sci* **76**, 2138-2145.

Matsuki, T., Watanabe, K., Fujimoto, J., Miyamoto, Y., Takada, T., Matsumoto, K., Oyaizu, H. & Tanaka, R. (2002). Development of 16S rRNA-gene-targeted group-specific primers for the detection and identification of predominant bacteria in human feces. *Applied and Environmental Microbiology* **68,** 5445-5451.

McCracken, V. J., Simpson, J. M., Mackie, R. I. & Gaskins, H. R. (2001). Molecular Ecological Analysis of Dietary and Antibiotic-Induced Alterations of the Mouse Intestinal Microbiota. *J Nutr* **131**, 1862-1870.

McIver, C. J., Jacques, C. F. H., Chow, S. S. W., Munro, S. C., Scott, G. M., Roberts, J. A., Craig, M. E. & Rawlinson, W. D. (2005). Development of Multiplex PCRs for Detection of Common Viral Pathogens and Agents of Congenital Infections. *J Clin Microbiol* **43**, 5102-5110.

Miller, N. A., Bryan, Q. G. R., Ruvolo, M., Turner, L. A. & LaBrie, S. T. (2002). Cross-hybridization of closely related genes on high-denisty macroarrays. *Biotechniques* **32**, 620-625.

Moore, W. E. & Holdeman, L. V. (1974). Special problems associated with the isolation and identifcation of intestinal bacteria in fecal flora studies. *American Journal of Clinical Nutrition* **27**, 1450-1455.

Moore, W. E., Cato, E. P. & Holdeman, L. E. (1978). Some current topics in intestinal bacteria. *American Journal of Clinical Nutrition* **31**, S33-S42.

Muyzer, G., de Waal, E. & Uitterlinden, A. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes encoding for 16S rRNA. *Applied and Environmental Microbiology* **59**, 695-700.

Muyzer, G. (1999). DGGE/TGGE a method for identifying genes from natural ecosystems. *Current Opinion in Microbiology* **2**, 317.

Nadkarni, M. A., Martin, F. E., Jacques, N. A. & Hunter, N. (2002). Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology* **148**, 257-266.

O'Sullivan, D. J. (2000). Methods for analysis of the intestinal microflora. *Current Issues in Intestinal Microbiology* **1**, 39-50.

O'Sullivan, D. J. (2001). Screening of intestinal microflora for effective probiotic bacteria. *Journal of Agricultural and Food Chemistry* **49**, 1751-1760.

Olsen, G. J., Lane, D. J., Giovannoni, S. J. & Pace, N. R. (1986). Microbial ecology and evolution: A ribosomal RNA approach. *Annual Review of Microbiology* **40**, 337-365.

Pace, N. R., Stahl, D. A., Lane, D. J. & Olsen, G. J. (1986). The analysis of natural microbial populations by ribosomal RNA sequences. *Advances in Microbial Evolution* **9**, 1-55.

Pahl, A., Kuehlbrandt, U., Brune, K., Roellinghoff, M. & Gessner, A. (1999). Quantitative detection of Borrelia burgdorferi by real-time PCR. *Journal of Clinical Microbiology* **37**, 1958-1963.

Pedersen, K. & Tannock, G. W. (1989). Colonization of the porcine gastrointestinal tract by lactobacilli. *Appl Environ Microbiol* **55**, 279-283.

Peixoto, R. S., da Costa Coutinho, H. L., Rumjanek, N. G., Macrae, A. & Rosado, A. S. (2002). Use of rpoB and 16S rRNA genes to analyse bacterial diversity of a tropical soil using PCR and DGGE. *Letters in Applied Microbiology* **35**, 316-320.

Picketts, D. J., Mayanil, C. S. K. & Grupta, R. S. (1989). Molecular cloning of a Chinese hamster mitochondrial protein related to the 'chaperonin' family of bacterial and plant proteins. *Journal of Biological Chemistry* **264**, 12001-12008.

Pryde, S. E., Richardson, A. J., Stewart, C. S. & Flint, H. J. (1999). Molecular Analysis of the Microbial Diversity Present in the Colonic Wall, Colonic Lumen, and Cecal Lumen of a Pig. *Appl Environ Microbiol* **65**, 5372-5377.

Rada, V., Sirotek, K. & Peter, J. (1999). Evaluation of Selective Media for Bifidobacteria in Poultry and Rabbit Caecal Samples. *J Vet Med* **46**, 369-373.

Ramette, A., LiPuma, J. J. & Tiedje, J. M. (2005). Species Abundance and Diversity of Burkholderia cepacia Complex in the Environment. *Appl Environ Microbiol* **71**, 1193-1201.

Raskin, L., Capman, W. C., Sharp, R., Poulsen, L. K. & Stahl, D. A. (1997). Molecular Ecology of Gastrointestinal Ecosystems. In *Gastrointestinal Microbes and Host Interactions*, pp. 243-298. Edited by R. I. Mackie, B. A. White & R. E. Issacson. New York: International Thomson Publishing.

Ritalahti, K. M., Amos, B. K., Sung, Y., Wu, Q., Koenigsberg, S. S. & Loffler, F. E. (2006). Quantitative PCR Targeting 16S rRNA and Reductive Dehalogenase Genes Simultaneously Monitors Multiple Dehalococcoides Strains. *Appl Environ Microbiol* **72**, 2765-2774.

Robinson, I. M., Allison, M. J. & Bucklin, J. A. (1981). Characterization of the cecal bacteria of normal pigs. *Appl Environ Microbiol* **41**, 950-955.

Rosch, C. & Bothe, H. (2005). Improved Assessment of Denitrifying, N2-Fixing, and Total-Community Bacteria by Terminal Restriction Fragment Length Polymorphism Analysis Using Multiple Restriction Enzymes. *Appl Environ Microbiol* **71**, 2026-2035.

Rusanganwa, E. & Grupta, S. (1993). Cloning and characterization of mutiple *groEL* chaperonin-encoding genes in Rhizobium meliloti. *Gene* **126**, 67-75.

Saikaly, P. E., Stroot, P. G. & Oerther, D. B. (2005). Use of 16S rRNA Gene Terminal Restriction Fragment Analysis To Assess the Impact of Solids Retention Time on the Bacterial Diversity of Activated Sludge. *Appl Environ Microbiol* **71,** 5814-5822.

Sakata, T., Kojima, T., Fijieda, M., Takahashi, M. & Michibata, T. (2003). Influences of probiotic bacteria on organic acid production by pig caecal bacteria *in vitro*. *Proceedings of the Nutrition Society* **62**, 73-80.

Salzman, N. H., de Jong, H., Paterson, Y., Harmsen, H. J., Welling, G. W. & Bos, N. A. (2002). Analysis of 16S libraries of mouse gastrointestinal microflora reveals a large new group of mouse intestinal bacteria. *Microbiology* **148**, 3651-3660.

Satokari, R. M., Vaughan, E. E., Akkermans, A. D. L., Saarela, M. & de Vos, W. M. (2001). Bifidobacterial Diversity in Human Feces Detected by Genus-Specific PCR and Denaturing Gradient Gel Electrophoresis. *Appl Environ Microbiol* **67**, 504-513.

Satokari, R. M., Vaughan, E. E., Smidt, H., Saarela, M., Matto, J. & de Vos, W. M. (2003). Molecular Approaches for the Detection and Identification of Bifidobacteria and Lactobacilli in the Human Gastrointestinal Tract. ***Systematic and Applied Microbiology*** 26, 572.

Schena, M., Shalon, D., Davis, R. & Brown, P. (1995). Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. ***Science*** **270,** 467-470.

Siiger, U., Tamm, E., Torm, S., Lustar, I., Salminen, S. & Midtvedt, T. (1996). Effect of bacterial infection and administration of a probiotic on faecal short-chain fatty acids. *Microbial Ecology in Health and Disease*, 271-277.

Simpson, J. M., McCracken, V. J., White, B. A., Gaskins, H. R. & Mackie, R. I. (1999). Application of denaturant gradient gel electrophoresis for the analysis of the porcine gastrointestinal microbiota. *Journal of Microbiological Methods* **36,** 167.

Smith, L. F. (2006). Impact of tylosine phosphate, flaxseed, and flaxseed fractions on small intestinal microbial profiles in pigs. *Thesis - Masters of Science*.

Socransky, S., Haffajee, A., Smith, C., Martin, L., Haffajee, J., Uzel, N. & Goodson, J. (2004). Use of checkerboard DNA-DNA hybridization to study complex microbial ecosystems. *Oral Microbiology and Immunology* **19,** 352-362.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503.

Stahl, D. A., Flesher, B., Mansfield, H. R. & Montgomery, L. (1988). Use of phylogenetically based hybridization probes for studies of ruminal microbial ecology. *Applied and Environmental Microbiology* **54,** 1079-1084.

Stahl, D. A. & Amann, R. I. (1991). Development and Application of Nucleic Acid Probes. *Nucleic Acid Techniques in Bacterial Systematics*, 205-248.

Staley, J. T. & Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annual Review of Microbiology* **39**, 321-346.

Steward, G. F., Jenkins, B. D., Ward, B. B. & Zehr, J. P. (2004). Development and Testing of a DNA Macroarray To Assess Nitrogenase (nifH) Gene Diversity. *Appl Environ Microbiol* **70**, 1455-1465.

Suau, A., Bonnet, R., Sutren, M., Godon, J. J., Gibson, G. R., Collins, M. D. & 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.

Suzuki, M. & Giovannoni, S. (1996). Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl Environ Microbiol* **62,** 625-630.

Swords, W., Wu, C., Champlin, F. & Buddington, R. (1993). Postnatal changes in selected bacterial groups of the pig colonic microflora. *Biol Neotate* **63**, 191-200.

Taha, M.-K. & Olcen, P. E. R. (2004). Molecular genetic methods in diagnosis and direct characterization of acute bacterial central nervous system infections. *APMIS* **112**, 753-770.

Tajima, K., Aminov, R., Nagamine, T., Matsui, H., Nakamura, M. & Benno, Y. (2001). Diet-dependent shifts in the bacterial population of the rumen revealed with real-time PCR. *Applied and Environmental Microbiology* **67**, 2766-2774.

Tan, Z., Hurek, T. & Reinhold-Hurek, B. (2003). Effect of N-fertlilization, Plant Genotype, and Environmental Conditions on *nifH* Gene Pools in Roots of Rice. *Enivron Micro* **5**, 1009-1015.

Tannock, G. W., Fuller, R. & Pedersen, K. (1990). Lactobacillus succession in the piglet digestive tract demonstrated by plasmid profiling. *Appl Environ Microbiol* **56,** 1310-1316.

Tannock, G. W., 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. *Appl Environ Microbiol* **65**, 4264-4267.

Tannock, G. W. (2002). Analysis of the intestinal microflora using molecular methods. *European journal of clinical nutrition* **56**, S44-S49.

Tannock, G. W., K. Munro, H.J.M. Harmsen, G.W. Welling, J. Smart, and P.K. Gopal (2000). Analysis of the fecal microflora of human subjects consuming a probiotic product containing Lactobacillus rhamnosus DR20. *Applied and Environmental Microbiology* **66**, 2578-2588.

Thomas, P. S. (1980). Hybridization of Denatured RNA and Small DNA Fragments Transferred to Nitrocellulose. *PNAS* **77**, 5201-5205.

Trad, S., Allignet, J., Frangeul, L. & other authors (2004). DNA Macroarray for Identification and Typing of Staphylococcus aureus Isolates. *J Clin Microbiol* **42,** 2054-2064.

Tyagi, S. & Kramer, F. (1996). Molecular beacons: probes that fluoresce upon hybridization. *Nature Biotechnology* **14**, 303-308.

Vallaeys, T., Topp, E., Muyzer, G., Macheret, V., Laguerre, G., Rigaud, A. & Soulas, G. (1997). Evaluation of denaturing gradient gel electrophoresis in the detection of 16S rDNA sequence variation in rhizobia and methanotrophs. *FEMS Microbiology Ecology* **24**, 279.

van Belkum, A. (1994). DNA fingerprinting of medically important microorganisms by use of PCR. *Clin Microbiol Rev* **7**, 174-184.

Vandamme, P., Pot, B., Gillis, M., De vos, P., Kersters, K. & Swings, J. (1996). Polyplasic taxonomy, a consensus approach to bacterial systems. *Microbiological Reviews* **60**, 407-438.

Varadarajb, K. & Skinnera, D. (1994). Denaturants or cosolvents improve the specificity of PCR amplification of a G+C rich DNA using genetically engineered DNA polymerase. *Gene* **140**, 1-5.

Wagner, M., Horn, M. & Daims, H. (2003). Fluorescence in situ hybridisation for the identification and characterisation of prokaryotes. *Current Opinion in Microbiology* **6**, 302.

Waleron, M., Waleron, K., Podhajska, A. J. & Lojkowska, E. (2002). Genotyping of bacteria belonging to the former Erwinia genus by PCR-RFLP analysis of a recA gene fragment. *Microbiology* **148**, 583-595.

Wallner, G., Fuchs, B., Spring, S., Beisker, W. & Amann, R. (1997). Flow sorting of microorganisms for molecular analysis. *Appl Environ Microbiol* **63**, 4223-4231.

Walter, J., Hertel, C., Tannock, G. W., Lis, C. M., Munro, K. & Hammes, W. P. (2001). Detection of Lactobacillus, Pediococcus, Leuconostoc, and Weissella Species in Human Feces by Using Group-Specific PCR Primers and Denaturing Gradient Gel Electrophoresis. *Appl Environ Microbiol* **67**, 2578-2585.

Wang, R.F., Beggs, M. L., Robertson, L. H. & Cerniglia, C. E. (2002a). Design and evaluation of oligonucleotide-microarray method for the detection of human intestinal bacteria in fecal samples. *FEMS Microbiology Letters* **213**, 175-182.

Wang, R.F., L.H. Robertson, and C.E. Cerniglia (2002b). Development of a membrane-array method for the detection of human intestinal bacteria in fecal samples. *Molecular and Cellular Probes* **16**, 341-350.

Wang, R.F., Cao, W. W. & Cerniglia, C. E. (1997). PCR detection of *Ruminococcus* spp. in human and animal faecal samples. *Molecular and Cellular Probes* **11**, 259-265.

Wetmur, J. G. (1976). Hybridization and renaturation kinetics of nucleic acids. *Annu Rev Biophys Bioeng* **5**, 337-361.

Widmer, F., Shaffer, B. T., Porteous, L. A. & Seidler, R. J. (1999). Analysis of nifH Gene Pool Complexity in Soil and Litter at a Douglas Fir Forest Site in the Oregon Cascade Mountain Range. *Appl Environ Microbiol* **65**, 374-380.

Wise, M. G. & Siragusa, G. R. (2005). Quantitative Detection of Clostridium perfringens in the Broiler Fowl Gastrointestinal Tract by Real-Time PCR. *Appl Environ Microbiol* **71**, 3911-3916.

Wittwer, C. T., Herrmann, M. G., Gundry, C. N. & Elenitoba-Johnson, K. S. J. (2001). Real-Time Multiplex PCR Assays. *Methods* **25**, 430.

Woese, C. R. (1987). Bacterial Evolution. *Microbiological Reviews* **51**, 221-271.

Wren, J. D., Kulkarni, A., Joslin, J., Butow, R. A. & Harold, R. G. (2002). Cross-hybridization on PCR spotted arrays. *Engineering in Medicine and Biology Magazine IEEE* **21**, 71-75.

Wu, C., Carta, R. & Zhang, L. (2005). Sequence dependence of cross-hybridization on short oligo microarrays. *Nucl Acids Res* **33,** e84-91.

Wu, L., Thompson, D. K., Li, G., Hurt, R. A., Tiedje, J. M. & Zhou, J. (2001). Development and Evaluation of Functional Gene Arrays for Detection of Selected Genes in the Environment. *Appl Environ Microbiol* **67**, 5780-5790.

Yang, M.-K., Chou, M.-E. & Yen-Chun, Y. (2001). Molecular Characterization and Expression of the recX Gene of Xanthomonas campestris pv. citri. *Current Microbiology* **42**, 257.

Ye, R. W., Wang, T., Bedzyk, L. & Croker, K. M. (2001). Applications of DNA microarrays in microbial systems. *Journal of Microbiological Methods* **47**, 257.

Zeilstra-Ryalls, Fayet, J. O. & Georgopoulos, C. P. (1991). The universally conserved GroE (Hsp60) chaperonins. *Annual Review of Microbiology* **45**, 301-325.

Zheng, D., Alm, E. W., Stahl, D. A. & Raskin, L. (1996). Characterization of universal small-subunit rRNA hybridization probes for quantitative molecular microbial ecology studies. *Appl Environ Microbiol* **62**, 4504-4513.

Zhou, J. (2003). Microarrays for bacterial detection and microbial community analysis. *Current Opinion in Microbiology* **6**, 288.

Zoetendal, E. G., Akkermans, A. D. L. & De Vos, W. M. (1998). Temperature Gradient Gel Electrophoresis Analysis of 16S rRNA from Human Fecal Samples Reveals Stable and Host-Specific Communities of Active Bacteria. *Appl Environ Microbiol* **64**, 3854-3859.

Zoetendal, E. G., Collier, C. T., Koike, S., Mackie, R. I. & Gaskins, H. R. (2004). Molecular Ecological Analysis of the Gastrointestinal Microbiota: A Review. *J Nutr* **134**, 465-472.