Microb Ecol (2010) 59:14–24
DOI 10.1007/s00248-009-9601-1

GENES AND GENOMES

Resolution of Phenotypically Distinct Strains of *Enterococcus* spp. in a Complex Microbial Community Using cpn60 Universal Target Sequencing

Catherine J. Vermette · Amanda H. Russell · Atul R. Desai · Janet E. Hill

Received: 7 August 2009 / Accepted: 5 October 2009 / Published online: 21 October 2009

Abstract Characterization of complex microbial communities is frequently based on the examination of polymerase chain reaction amplified sequences from a single phylogenetic marker, usually the 16S rRNA gene. However, this commonly used target often does not offer robust resolution of species or sub-species and is thus not a sufficiently informative target for understanding microbial population dynamics occurring at the strain level. We have used the cpn60 universal target sequence to characterize *Enterococcus* isolates from feces of growing pigs and have shown that sub-species groups, not detected using 16S rRNA sequences, can be resolved. Furthermore, groups resolved by cpn60-based phylogenetic analysis have distinct phenotypes. We report changes in the structure and function of *Enterococcus* communities in pig feces sampled from individual animals at three times, from suckling through to maturity. *Enterococcus faecalis* was largely replaced by *Enterococcus hirae* between suckling and 9 weeks of age, and a shift from one sub-species group of *E. hirae* to another was observed in all animals between 9 and 15 weeks. Conversely, *E. faecalis* strains remained consistent throughout the study period. Our results demonstrate that cpn60 sequences can be used to detect strain level changes in *Enterococcus* populations during succession in the fecal microbiota of growing pigs.

This manuscript is dedicated to Dr. Roland Brousseau who passed away too soon in May 2008.

Catherine J. Vermette and Amanda H. Russell contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00248-009-9601-1) contains supplementary material, which is available to authorized users.

C. J. Vermette · A. H. Russell · A. R. Desai · J. E. Hill (✉)
Department of Veterinary Microbiology,
Western College of Veterinary Medicine,
University of Saskatchewan,
52 Campus Drive,
Saskatoon, SK S7N 5B4, Canada
e-mail: Janet.Hill@usask.ca

Introduction

Recent advances in whole genome sequencing of prokaryotes have led to an increased recognition of the phenomenon of genome diversity within traditionally defined taxa [26]. Genomic diversity is observed among strains of well-characterized pathogens [4, 31] and in organisms isolated from natural microbial communities [23]. Twenty percent to 40% of open reading frames in prokaryotic genomes encode proteins of unknown function, and these uncharacterized genes are over-represented among unique, strain-specific genes [40]. Strain-specific genes are likely related to ecological niche specialization [40], and selection acting on these genes drives evolutionary diversification between strains [22]. It has been proposed that "natural taxa" are comprised of groups of strains or ecotypes that emerge in populations through periodic selection [6]. In gene sequence-based studies, these ecotypes or strains are detected as clusters of sequences where the average distance between ecotypes is greater than the average distances within the cluster. Our understanding of micro-diversity and ecotype dynamics is limited by difficulty in detecting phylogenetically discrete strains in complex samples using common, 16S rRNA sequence-based methods. Also, despite growing sequence-based evidence of sub-species level diversity within natural communities and an increasing appreciation of its potential biological significance, studies of the phenotypic relationships of sub-species groups defined by sequence relationships are rare. Without this information, it is difficult to evaluate the

significance of micro-diversity to the function of complex microbial communities.

In culture-independent, 16S rRNA sequence-based microbial ecology studies, polymerase chain reaction (PCR) with universal bacterial primers is conducted on total community DNA and products are sequenced, either via production of clone libraries or direct pyrosequencing. Usually, only a portion of the gene is amplified (approximately 500 bp) to simplify target amplification and sequencing. Taxonomic and phylogenetic assignment of sequences is facilitated by comparison to reference sequence databases, particularly the Ribosomal Database Project [25]. Traditionally, population data is lumped into "operational taxonomic units (OTU)", which appear as discrete clusters of nearly identical sequences. The reasons for the observed sequence micro-diversity within OTUs identified by 16S rRNA sequence, or any PCR-amplified gene target, include PCR artifacts arising from non-proofreading polymerases [3]. However, fine-scale population structure, including discrete groups of sequences with <1% sequence divergence, is also apparent in 16S rRNA libraries where artifacts have been minimized through technical adjustments [2].

The gene encoding the universal 60 kDa chaperonin (cpn60) has been established as a target for identification of bacterial species and has the advantage of offering superior discrimination between closely related organisms: at the species [17, 21, 28] and sub-species levels [5]. Among protein-coding alternatives to 16S rRNA, cpn60 is the most developed, and a large, curated database of sequence data, cpnDB (http://cpndb.cbr.nrc.ca), is available [18]. Since cpn60 is a protein-coding gene, the codon position of differences between nearly identical sequences generated in microbial ecology studies can be informative. In a study of human vaginal microbial communities where cpn60 was used as a target rather than the 16S rRNA gene [15], the distribution of differences between members of a cluster of sequences with >90% identity to each other was skewed toward synonymous substitutions at the third codon position, rather than the random distribution expected from polymerase error. Similar observations have been made of cpn60 sequence data from microbial populations in animal intestines [10, 16, 19] and an industrial environment [11]. This observation suggests that at least some sequence micro-diversity in microbial populations may be indicative of real strain diversity, rather than technical artifacts. Furthermore, it suggests that cpn60 sequences may be used to characterize sub-species level diversity within natural microbial communities.

The objectives of the current study were to compare sub-species resolution of the cpn60 and 16S rRNA gene targets and to determine if the cpn60 target could be used to detect species and sub-species level changes in a model microbial community: *Enterococcus* in the feces of pigs from weaning to maturity. We have also used carbon source utilization profiling to investigate the phenotypic diversity among strains within the fecal enterococci and to determine if strains identified based on cpn60 sequence have distinct phenotypes. We chose *Enterococcus* spp. as a model population since they are easily isolated from the intestinal microbiota of pigs and have a broad repertoire of metabolic capabilities.

Materials and Methods

Swine Feces Collection

Rectal fecal samples were collected from pigs in a biosecure, high health production facility. Collections were made at 3, 9, and 15 weeks of age. Fourteen pigs were included in the study: two to four pigs from each of four litters, and a mixture of male and female piglets were sampled from each litter. Piglets at 3 weeks were suckling. Diet composition at 9 weeks consisted of 55% wheat, 26% soymeal, and 15% barley. At 15 weeks, the diet consisted of 30% corn, 23% barley, 14% soymeal, 8% canola meal, and 23% wheat.

Bacterial Culture

Fecal samples were diluted in peptone water and plated on mEnterococcus agar (Difco) using a spiral plater to facilitate colony counting. Plates were incubated at 35°C for 48 h. The average number of colonies from at least three plates was calculated for each sample to determine colony-forming units per gram. Isolated colonies were picked into 150 μL of Luria-Bertani broth with 12% (*v*/*v*) glycerol and incubated overnight at 35°C prior to freezing at −80°C for storage.

Type strains of *Enterococcus faecalis* (ATCC 19433), *Enterococcus hirae* (ATCC 8043), *Enterococcus faecium* (ATCC 19434), *Enterococcus avium* (ATCC 14025), and *Enterococcus mundtii* (ATCC 43186) were obtained from the American Type Culture Collection.

PCR and Sequencing

All PCR reactions were performed in 96-well plates using 1 μL of glycerol stock of each bacterial isolate as template. The region of the 16S rRNA gene corresponding to nucleotides 11-536 of the *Escherichia coli* 16S rRNA gene (encompassing variable regions V1, V2, and V3) was amplified with universal primers F1 (5′-GAG TTT GAT CCT GGC TCA G-3′) and R2 (5′-GWA TTA CCG CGG CKG CTG-3′) [9]. PCR amplifications were performed in 50 μL reactions consisting of 5 μL of a 10× PCR reaction

buffer, 2 mM $MgCl_2$, 0.2 mM each of a dNTP mixture, 0.4 μM of each primer, and 1 U of HP Taq DNA polymerase (UBI, Calgary, AB, Canada) using an Eppendorf Mastercycler EP. The amplification program was 3 min at 95°C followed by 40 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 30 s, and a final extension of 10 min at 72°C.

For amplification of the cpn60 universal target region corresponding to nucleotides 274-828 of the *E. coli* cpn60 sequence, primers JH0064 (5′-GAT ATY GCW GGW GAY GGW ACA ACD AC-3′) and JH0065 (5′-CGR CGR TCR CCR AAS CCS GGH GCT-3′) were used. JH0064 and JH0065 were designed based on universal cpn60 primers H279 (5′-GAI III GCI GGI GAY GGI ACI ACI AC-3′) and H280 (5′-YKI YKI TCI CCR AAI CCI GGI GCY TT-3′) and modified to match the consensus sequences of *E. hirae*, *E. faecalis*, *Enterococcus cecorum*, and *E. faecium*, which were previously reported to be present in pig feces. These primers were used instead of universal cpn60 primers H279 and H280 for this *Enterococcus*-focused study to avoid the expense of inosine-containing primers. PCR reaction components were identical to those described for 16S rRNA. The amplification program for JH0064/JH0065 was 3 min at 95°C followed by 40 cycles of 95°C for 30 s, 62°C for 60 s, 72°C for 60 s, and a final extension of 10 min at 72°C.

PCR products were purified using the PCR Microarray Purification kit (Arrayit Corporation, Sunnyvale, CA, USA). Purified products were sequenced directly with the amplification primers. Raw sequence data for individual PCR products was manually edited and assembled using the Staden Package with the default settings for quality cutoff in the Pregap4 module [37]. Other sequence manipulation was accomplished using EMBOSS [34]. Phylogenetic trees were created using PHYLIP [12] and viewed with NJplot [32].

Species identification of sequences was made by comparison of isolate sequences to the Greengenes database (16S rRNA sequences; http://greengenes.lbl.gov) or cpnDB (cpn60 sequences; http://cpndb.cbr.nrc.ca) [18].

Carbon Source Utilization Profiling

Carbon source utilization profiles were generated using the BIOLOG GP2 microplate (Biolog Incorporated, Hayward, CA, USA) according to the manufacturer's instructions. Prior to analysis, strains were resurrected from −80°C glycerol stocks on trypticase soy agar plates containing 5% defibrinated sheep blood and incubated for 48 h at 35°C. Results were recorded at 18 h post-inoculation. Statistical analysis of carbon source utilization patterns was accomplished using Statistical Package for the Social Sciences (SPSS) software (SPSS Inc., Chicago, USA).

Results

Enterococcus Population Size Decreases with Age

Table 1 summarizes the total *Enterococcus* colony counts for all pigs over the three time periods. Enormous variation in counts was observed in the weaning pigs (3 weeks of age), ranging from 0 in pig 61 (no growth observed on mEnterococcus agar plates) to $>10^8$ cfu/g in pigs 58, 64, and 82. Observations at the later times were more consistent. By 15 weeks, total *Enterococcus* counts had dropped to 10^3 to 10^4 cfu/g in all pigs for which data was collected, except pig 64 where no growth was observed. All isolates examined were determined to be *Enterococcus* spp.

Identification of *Enterococcus* spp. by 16S rRNA and cpn60 Sequencing

To compare the utilities of the 16S rRNA and cpn60 targets for *Enterococcus* species identification and strain level resolution, we amplified and sequenced both targets from each of 169 isolates representing all three time points, from pig 81. BLASTclust was used to identify 100% identical sequences within the collection. Using this criterion, the 16S rRNA sequences reduced to 23 different sequences, and the cpn60 data set contained 43 different sequences. Three species (*E. hirae*, *E. faecalis*, and *E. faecium*) were identified based on cpn60 sequence identity to type strains. Resolution of the *E. hirae* and *E. faecium* strains by 16S rRNA sequence was unclear since the type strain sequences for these species are 99% identical to each other over the 510-bp region examined. The cpn60 sequences for these

Table 1 Total *Enterococcus* colony counts in pig feces

Pig ID	Litter	Total culturable *Enterococcus* (cfu/g)		
		3 weeks	9 weeks	15 weeks
57	5	2.39×10^3	1.33×10^4	3.00×10^3
58		$>1\times10^8$	No sample	Dead
61	6	0.00	No sample	3.33×10^3
64		$>1\times10^8$	No sample	0.00
67		4.00×10^5	3.00×10^4	3.33×10^2
68		4.91×10^3	1.70×10^5	6.66×10^2
70	7	8.00×10^5	Dead	Dead
71		1.27×10^5	1.33×10^4	3.66×10^3
77		5.63×10^5	0.00	3.20×10^3
78		3.53×10^5	1.67×10^4	3.70×10^3
81	8	2.92×10^7	4.33×10^4	4.60×10^3
82		$>1\times10^8$	No sample	1.90×10^3
87		2.43×10^7	9.43×10^4	No sample
88		1.33×10^5	1.52×10^6	3.47×10^4

two species are only 85% identical to each other. Interspecific and intraspecific distances were larger for the cpn60 sequences, resulting in clear resolution of species and the identification of two distinct subgroups with *E. hirae* (Fig. 1). Based on these initial results from pig 81, all further characterization of strains was done only with the cpn60 target.

A total of 694 isolates from pigs 57 (13 isolates), 67 (45 isolates), 68 (185 isolates), 71 (9 isolates), 78 (156 isolates), 81 (168 isolates), and 88 (118 isolates) were subjected to cpn60 PCR and sequencing. These pigs were selected since isolates were available at each time point. In most cases, all isolates cultured from a given pig at each time were collected for analysis. In cases were more than 96 isolates were available, only the first 96 randomly chosen isolates were examined. Although five species were identified (*E. faecium*, *E. hirae*, *E. faecalis*, *E. mundtii*, and *E. avium*), all but three isolates were *E. faecalis* or *E. hirae*. Figure 2 is a phylogenetic tree of the unique cpn60 sequences from each pig at each time (198 sequences representing 694 isolates). Representative sequences have been deposited in Genbank and assigned accession numbers GQ405204–

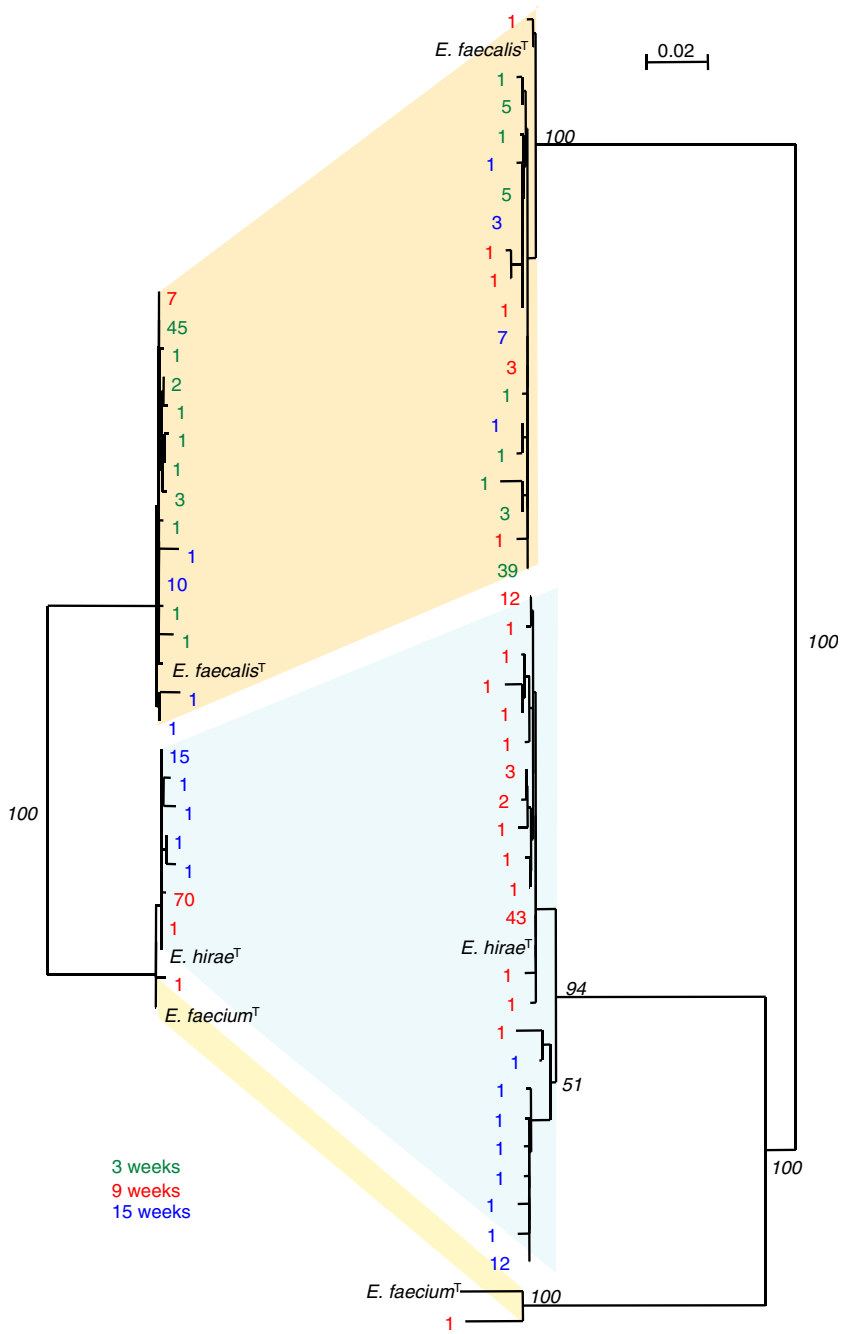

Figure 1 Phylogenetic trees based on partial 16S rRNA (*left*) and cpn60 (*right*) sequences for isolates from pig81 at 3, 9, and 15 weeks of age. Sequences from type strains of *Enterococcus hirae*, *Enterococcus faecalis*, and *Enterococcus faecium* are included for reference. *Labels* indicate the number of isolates with identical sequences at each point in the tree. *Label colors* indicate the age of pig from which the sequenced isolate was obtained. The tree is a consensus of 100 neighbor-joined trees, and bootstrap values are indicated at the major branch points

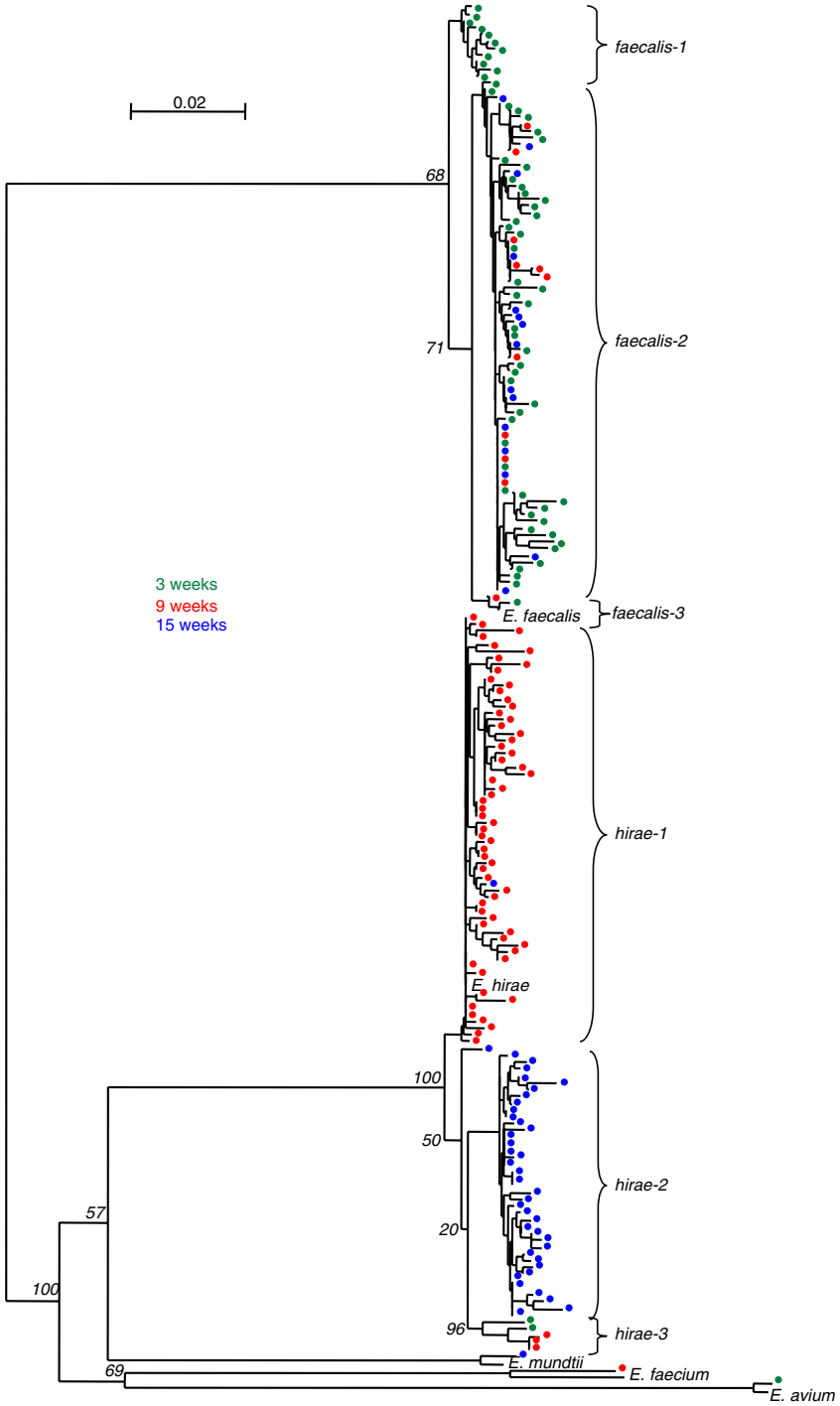

Figure 2 Phylogenetic trees based on alignment of 198 unique cpn60 gene from 694 *Enterococcus* isolates. Representative sequences from each pig at each time are included. Type strains of *Enterococcus faecalis*, *Enterococcus hirae*, *Enterococcus faecium*, *Enterococcus mundtii*, and *Enterococcus avium* are included for reference. The tree is a consensus of 100 neighbor-joined trees, and bootstrap values are indicated at the major branch points. Sub-species groups are also indicated. *Colored dots* indicate the age of pig from which the sequences were obtained

GQ405360. Sub-species divisions with solid bootstrap support are evident within *E. faecalis* and *E. hirae*, and the sequence identities between these intraspecific groups are 96–98%. All sub-species groups contain sequences from multiple pigs. A closer examination of the *E. hirae* species cluster (107 sequences) shows that of the 66 variable positions in an alignment of these sequences (552 bp), the ratio of codon positions (first:second:third) is 19:13:34, indicating a skew toward the third position. Within hirae-1 and hirae-2 subgroups, codon position of differences was 18:9:15 and 10:6:7, respectively.

Temporal Shifts in Sub-species Groups Within the Genus *Enterococcus*

Figures 2 and 3 describe the observed temporal changes in fecal *Enterococcus* populations. At 3 weeks, the isolate collection was dominated by *E. faecalis* which accounted

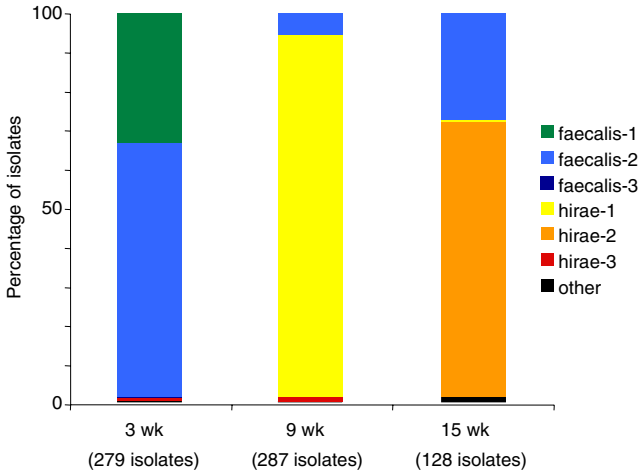

Figure 3 Proportional representation of the various species and strains in isolates from fecal samples at 3, 9, and 15 weeks of age. The total number of isolates analyzed at each age is indicated below the *x*-axis. Sub-species group names correspond to Fig. 2

for over 98% of characterized isolates. Two isolates were classified as hirae-3 and one as *E. faecium*. At 9 weeks, the vast majority of isolates were identified as *E. hirae* (93% hirae-1 and 1% hirae-3). By 15 weeks, *E. hirae* was still the dominant cultured species, but there was a pronounced shift to subgroup hirae-2, with only one hirae-1 isolate identified from pig71. No subgroup hirae-3 isolates were detected at 15 weeks. *E. faecalis* subgroup faecalis-2 was observed at all times in all pigs, while faecalis-1 was only detected in suckling pigs 68 and 88 (3 weeks).

Carbon Source Utilization Profiles of *Enterococcus* Sub-species Groups

To determine if the sub-species groups identified based on cpn60 sequence and phylogeny were also distinguishable based on phenotype, we determined carbon source utilization profiles from the 198 isolates representing the sequences shown in Fig. 2 as well as type strains of *E. faecalis* (ATCC 19433), *E. hirae* (ATCC 8043), *E. faecium* (ATCC 19434), *E. avium* (ATCC 14025), and *E. mundtii* (ATCC 43186), using the BIOLOG GP2 system. Ninety-five sole carbon sources are evaluated in each plate. None of the isolates had profiles identical to their corresponding type strains.

Among all *E. hirae* isolates, 34 of the carbon sources were utilized by at least one isolate (Table 2, Supplementary Fig. 1a). For 14 of these carbon sources, a significant difference in the proportion of isolates in each subgroup utilizing the carbon source was observed, based on binary logistic regression analysis with $p<0.05$ considered significant. Hirae-2 was collectively able to utilize 30 of the 34 carbon sources with 18 of these being utilized by at least 50% of the isolates. Although subgroup hirae-1 was collectively able to utilize 31 of the 34 carbon sources, only six were utilized by more than 50% of the isolates, indicative of the greater degree of variability within hirae-1 compared to hirae-2. Individual hirae-2 isolates utilized an average of 16.5 carbon sources (range from 3–27) compared to an average of 8.6 for hirae-1 isolates (range 1–24). The most differentiating carbon sources for hirae-1 and hirae-2 were α-D-lactose, D-ribose, and uridine, which were positively associated with hirae-2. Subgroup hirae-3 isolates were collectively only able to utilize 16 carbon sources, although only five isolates were examined.

E. faecalis isolates were collectively able to utilize 50 of the carbon sources in the GP2 plate (Table 3, Supplementary Fig. 1b). Subgroups faecalis-1 and faecalis-2 could be distinguished on the basis of eight carbon sources whose utilization was positively associated with faecalis-2 (based on chi-square or Fisher's exact test, $p<0.05$). Faecalis-3 had only two members, providing inadequate information for meaningful comparison with the other groups. In subgroup faecalis-2, 29 of the 50 carbon sources were utilized by at least 50% of the isolates and of those, 24 were utilized by 90–100% of isolates. This is consistent with the more homogenous nature of carbon source utilization profiles observed for *E. faecalis* subgroups compared to *E. hirae*. Within faecalis-2 were isolates from all three ages of pig, but isolates from the age categories or individual pigs were not distinguishable by carbon source utilization profile (Supplementary Fig. 1b).

Discussion

There is ample evidence that diversity in microbial communities (at various taxonomic levels) is a significant parameter related to community function. Comparative studies where a community is examined over time or conditions are still relatively rare, but there are examples of changes in diversity that are either the cause or effect of biological changes. Metal contamination resulted in at least 99.9% reduction in diversity in a pristine soil, although the total number of bacteria remained approximately the same [13]; diversity in colonic microbial populations in inflammatory bowel disease patients is 30–50% of the diversity observed in healthy patients [29]; and diet composition and antimicrobial use affect microbial community sequence diversity in the intestines of pigs and chickens [10, 16]. Recently, based on genome sequencing of cultured prokaryotes, it has become evident that biologically significant genomic diversity exists within traditionally defined species, and this diversity may also be an important factor in microbial community dynamics. Questions regarding genomic diversity in natural microbial communities have been more difficult to address since rather than comparisons

Table 2 Carbon source utilization profile of *Enterococcus hirae* strains

Carbon source	Number (percent) positive isolates		
	hirae-1 (63 isolates)	hirae-2 (39 isolates)	hirae-3 (5 isolates)
N-Acetyl-D-glucosamine	22 (34.9) a	35 (89.7) b	2 (40.0) a
N-Acetyl-β-D-mannosamine	24 (38.1) a	33 (84.6) b	2 (40.0) a
Amygdalin	40 (63.5) a	37 (94.9) b	5 (100.0) ab
α-D-Glucose	32 (50. 8) a	34 (87.2) b	2 (40.0) a
D-Cellobiose	13 (20.6) a	31 (79.5) b	1 (20.0) a
D-Fructose	27 (42.9) a	35 (89.7) b	2 (40.0) a
Gentiobiose	15 (23.8) a	27 (69.2) b	2 (40.0) ab
α-D-Lactose	10 (15.9) a	36 (92.3) b	0 (0.0) ab
D-Mannose	30 (47.6) a	34 (87.2) b	1 (20.0) a
D-Melibiose	4 (6.3) a	17 (43.6) b	0 (0.0) ab
Sucrose	24 (38.1) a	34 (87.2) b	1 (20.0) a
D-Ribose	1 (1.6) a	37 (94.9) b	0 (0.0) ab
Salicin	46 (73.0) ab	34 (87.2) a	2 (40.0) b
D-Trehalose	6 (9.5) a	23 (59.0) b	1 (20.0) ab
Uridine	1 (1.6)	26 (66.7)	1 (20.0)
D-Gluconic acid	1 (1.6)	4 (10.3)	0 (0.0)
Arbutin	56 (88.9)	37 (94.9)	4 (80.0)
D-Galactose	57 (90.4)	39 (100.0)	5 (100.0)
Lactulose	1 (1.6)	1 (2.6)	0 (0.0)
Maltose	1 (1.6)	2 (5.1)	0 (0.0)
Maltotriose	1 (1.6)	2 (5.1)	0 (0.0)
D-Melezitose	1 (1.6)	3 (7.7)	0 (0.0)
β-Methyl-D-glucoside	56 (88.9)	38 (97.4)	4 (80.0)
Stachyose	2 (3.2)	0 (0.0)	0 (0.0)
α-Methyl-D-mannoside	1 (1.6)	0 (0.0)	0 (0.0)
D-Raffinose	1 (1.6)	0 (0.0)	0 (0.0)
α-Ketovaleric acid	1 (1.6)	0 (0.0)	0 (0.0)
Pyruvic acid methyl ester	1 (1.6)	1 (2.6)	0 (0.0)
Pyruvic acid	1 (1.6)	2 (5.1)	0 (0.0)
Glycerol	62 (98.4)	39 (100.0)	5 (100.0)
Adenosine	1 (1.6)	1 (2.6)	0 (0.0)
2′-Deoxy adenosine	0 (0.0)	1 (2.6)	0 (0.0)
Inosine	0 (0.0)	1 (2.6)	0 (0.0)
Thymidine	0 (0.0)	1 (2.6)	0 (0.0)

Numbers with different lowercase letters within a row indicate significant ($p<0.05$) differences determined by binary logistic regression

of two or a few fully sequenced genomes, a complex community of hundreds to thousands of different taxa must be considered. Although micro-diversity is discernable even in some carefully constructed 16S rRNA sequence libraries that include nearly full-length gene sequences (approximately 1.5 kb) [2], we sought to demonstrate that this fine-scale microbial community architecture is more clearly resolved using cpn60 universal target sequences.

The development of pig intestinal microbiota from birth to the "climax community" [35] of the adult offers an interesting model system. Dramatic changes in the structure and function of the microbiota occur, with the most dramatic changes associated with weaning when piglets are abruptly converted from a milk diet to solid food. Characterization of microbial succession in pigs to date has involved methods that do not detect changes occurring below the level of traditionally defined species [20, 38].

Our observation of a decrease in the number of culturable enterococci in feces from weaning to 15 weeks is consistent with previous observations of the progression from a neonatal intestinal microbiota specialized for the digestion of a milk diet to the more complex and diverse microbiota required for the digestion of complex diet ingredients. In pigs, as in other animals including humans, there is a succession from coliforms and streptococci,

Table 3 Carbon source utilization profile differences between *Enterococcus faecalis* strains

Carbon source	Number (percent) positive isolates		
	faecalis-1 (11 isolates)	faecalis-2 (68 isolates)	faecalis-3 (2 isolates)
Uridine-5′-monophosphate	11 (100) a	38 (55.9) b	1 (50)
Thymidine-5′-monophosphate	0 (0)	17 (25)	0 (0)
Adenosine-5′-monophosphate	0 (0) a	29 (42.6) b	0 (0)
Uridine	10 (91.0)	67 (98.5)	2 (100)
Thymidine	11 (100)	65 (95.6)	2 (100)
Inosine	6 (54.5) a	61 (89.7) b	2 (100)
2′-Deoxy adenosine	11 (100)	64 (94.1)	1 (50)
D-L-α-Glycerol phosphate	0 (0)	1 (1.5)	0 (0)
D-Glucose-6-phosphate	0 (0)	2 (2.9)	0 (0)
Adenosine	11 (100)	67 (98.5)	2 (100)
Glycerol	11 (100)	68 (100)	2 (100)
Pyruvic acid	1 (9.1) a	52 (76.5) b	1 (50)
Pyruvic acid methyl ester	11 (100)	65 (95.6)	2 (100)
L-Malic acid	1 (9.1) a	38 (55.9) b	0 (0)
Turanose	0 (0) a	20 (29.4) b	0 (0)
D-Trehalose	10 (91.0)	64 (94.1)	2 (100)
α-Ketovaleric acid	0 (0) a	22 (32.4) b	0 (0)
Salicin	11 (100)	66 (97.1)	2 (100)
D-Ribose	10 (91.0)	66 (97.1)	2 (100)
L-Rhamnose	0 (0)	2 (2.9)	0 (0)
D-Raffinose	1 (9.1)	1 (1.5)	0 (0)
Palatinose	1 (9.1)	1 (1.5)	0 (0)
Sucrose	11 (100)	67 (98.5)	2 (100)
Stachyose	1 (9.1)	1 (1.5)	0 (0)
D-Sorbitol	0 (0)	2 (2.9)	0 (0)
β-Methyl-D-glucoside	11 (100)	59 (86.8)	2 (100)
α-Methyl-D-galactoside	1 (9.1)	0 (0)	0 (0)
D-Melibiose	1 (9.1)	1 (1.5)	0 (0)
D-Melezitose	11 (100)	67 (98.5)	2 (100)
D-Mannose	11 (100)	68 (100)	2 (100)
D-Mannitol	0 (0)	2 (2.9)	0 (0)
Maltotriose	11 (100)	66 (97.1)	2 (100)
Maltose	11 (100)	67 (98.5)	2 (100)
Lactulose	0 (0)	4 (5.9)	0 (0)
α-Methyl-D-glucoside	0 (0)	1 (1.5)	0 (0)
α-D-Lactose	1 (9.1) a	54 (79.4) b	0 (0)
Gentiobiose	10 (91.0)	68 (100)	1 (50)
D-Galacturonic acid	1 (9.1)	0 (0)	0 (0)
D-Galactose	11 (100)	68 (100)	2 (100)
D-Fructose	11 (100)	68 (100)	2 (100)
D-Cellobiose	11 (100)	67 (98.5)	2 (100)
Arbutin	11 (100)	68 (100)	2 (100)
m-Inositol	1 (9.1)	3 (4.4)	0 (0)
α-D-Glucose	10 (91.0)	68 (100)	2 (100)
D-Gluconic acid	10 (91.0)	65 (95.6)	2 (100)
b-Cyclodextrin	0 (0)	5 (7.4)	0 (0)
α-Cyclodextrin	0 (0)	6 (8.8)	0 (0)

Table 3 (continued)

Carbon source	Number (percent) positive isolates		
	faecalis-1 (11 isolates)	faecalis-2 (68 isolates)	faecalis-3 (2 isolates)
Amygdalin	1 (9.1)	3 (4.4)	0 (0)
N-Acetyl-β-D-mannosamine	11 (100)	68 (100)	2 (100)
N-Acetyl-D-glucosamine	11 (100)	66 (97.0)	2 (100)

Numbers with different lowercase letters within a row indicate significant ($p<0.05$) differences determined by chi-square or Fisher's exact test, where appropriate. Comparisons were made only between faecalis-1 and faecalis-2

including *Enterococcus*, to the obligate anaerobes that dominate the climax community [19, 24]. The high degree of variability we observed in the suckling pig colony counts (from undetectable to $>10^8$ cfu/g) may reflect that our sampling represented only a "snapshot" at a single time of a complex and inherently unstable environment. However, by 15 weeks, the range of total culturable *Enterococcus* colony counts had decreased (3.33×10^2 to 3.47×10^4 cfu/g) suggesting that the microbiota of suckling piglets is more variable between pigs than that of adult animals.

One approach to describing genomic diversity is comparative genomics based on whole genome sequencing. Studies of *Enterococcus* spp., particularly *E. faecalis*, have demonstrated the genomic plasticity of these organisms and that they are replete with strain-specific genes [1, 27, 30]. A subset of 1,722 genes has been proposed for the *E. faecalis* core genome, with at least another 1,000 ORFs identified as being variably associated with individual strains characterized to date [36]. Strain-specific genes include putative virulence factors, antimicrobial resistance factors, as well as those involved in homeostasis, nutrient uptake, and metabolism. Selection acting on genes related to niche specialization is thought to be a major driver of strain diversification [22]. Our understanding of the range of genomic diversity in this genus is expanding rapidly with approximately 50 *Enterococcus* genome projects in progress (at the time of writing) and expansion of focus from pathogens to include commensal *Enterococcus* strains. However, current logistical constraints limit comparative genomic studies to selected isolates, rather than in-depth studies of natural microbial populations, which are more readily studied using a combination of targeted sequencing coupled with phenotypic characterization.

Sequence-based detection of sub-species level diversity in a population requires the use of a gene target with sufficient resolving power. Cpn60 sequences have previously been demonstrated to be useful for identification of phenotypically similar *Enterococcus* spp. of clinical importance [14]. A direct comparison of partial 16S rRNA and cpn60 sequences from isolates from pig 81 demonstrates the relative resolution of these two targets (Fig. 1). Pairwise interspecies sequence identities are lower for cpn60, allowing clear discrimination of the species included (85% pairwise identities between *E. faecalis*, *E. faecium*, and *E. hirae*). The amplified region of 16S rRNA (including the first three variable regions) was insufficient to discriminate between *E. faecium* and *E. hirae.* In fact, a comparison of published full-length 16S rRNA sequences for the type strains of these species (Genbank Accessions AJ276355 and AJ276356) shows that they are 99% identical over the entire length of the gene. Using cpn60 sequences, we were also able to identify two sub-species groups within *E. hirae*. Average pairwise sequence identities between these two groups were 96–98%, but they were nonetheless distinguishable with a high bootstrap value. These divisions, initially recognized in pig 81, were supported by additional isolates from all other pigs (Fig. 2). We also detected two sub-species groups within *E. faecalis* with similar intergroup sequence similarities.

The species represented in the culture collection were likely influenced by the isolation methods we employed. For example, while we did not isolate any *E. cecorum*, a species previously reported in pigs [7], there have been reports that this species does not grow well on common selective media for enterococci [8]. However, regardless of the culture method biases, isolation conditions were consistent at each time, making it possible to compare isolates selected. Dramatic shifts in the *Enterococcus* population profile were observed over the three time points, and these changes occurred simultaneously among all pigs examined (Figs. 2 and 3).

The fact that *Enterococcus* population changes occurred simultaneously indicates that the cause of these changes is likely something in common between all individual animals, probably diet. Fecal collection times corresponded to three different feed regimens: sow milk at 3 weeks and two different grower diets at 9 and 15 weeks. Between 9 and 15 weeks, dietary changes included a 50% decrease in wheat content and the introduction of corn and canola meal. The shift from hirae-1 to hirae-2 between 9 and 15 weeks may be the result of diet-based selection. The hirae-2 isolates were generally capable of utilizing more of the sole

carbon sources in the BIOLOG GP2 assay, perhaps indicating that they are more metabolically capable and were positively selected for with a more complex diet. This would be analogous to the changes in *Lactobacillus* metabolic profiles observed in the small intestine of weaning pigs where carbon source utilization pattern changes mirrored changes in diet composition [33]. The coordinated shifts in *Enterococcus* populations within the animals in this study is in interesting contrast to descriptions of microbial succession in human intestinal microbiota where much more individual-to-individual variation occurs, likely due to the individual environments and diets of the subjects [23, 39].

Analysis of the carbon source utilization profiles of representative isolates from all sub-species groups defined by cpn60-based phylogeny demonstrates that these groups also have distinguishable phenotypes and represent what have been described as ecotypes: groups of closely related strains occupying the same ecological niche [6]. This is best illustrated by the comparison of hirae-1 and hirae-2. The utilization patterns for 14 of the carbon sources evaluated showed significant differences between the *E. hirae* sub-species groups making it possible to classify isolates as hirae-1 or hirae-2 based on carbon source utilization parameters. For *E. faecalis*, faecalis-1 isolates were only found at 3 weeks in pigs 68 and 88. These isolates were distinguishable from faecalis-2 isolates based on cpn60 sequence and eight carbon sources whose utilization was positively association with faecalis-2 (Table 3, Supplemental Fig. 1a). Unlike the hirae-1 to hirae-2 shift, *E. faecalis* strains belonging to the faecalis-2 group were isolated throughout the study and consistent with their definition based on cpn60 sequence; very little variation in carbon source utilization pattern was observed within this group (Supplemental Fig. 1b). These observations suggest that while ecological conditions in the intestine selected for hirae-2 strains between 9 and 15 weeks, faecalis-2 strains occupy a niche unaffected by the diet change in this time frame.

Taken together, our results demonstrate that it is possible to use cpn60 sequences to identify phenotypically distinct strains or ecotypes within an intestinal microbial community. Lack of recognition of micro-diversity in many microbial ecology studies is attributable in part to the size of sequence collections produced. If a relatively small number of sequences from a single gene are generated from a population, it is unlikely that phylogenetically discrete sub-species groups will be identified. However, with the advent of next-generation sequencing technologies, much deeper data sets can be produced, increasing the likelihood of identifying these groups, assuming that the target sequence chosen is sufficiently informative. If micro-diversity is a significant parameter in microbial population dynamics, then exploitation of the cpn60 universal target, as demonstrated in this study, is a useful approach for its elucidation.

Acknowledgments The authors gratefully acknowledge John Harding and Crissie Baker for collection of pig fecal samples and Champika Fernando for assistance with statistical analysis. Funding for this work was provided by an NSERC Discovery Grant and a grant from the WCVM Vitamin Fund to JEH. Additional support was provided by NSERC Undergraduate Research Award (AHR) and the University of Saskatchewan (CJV & ARD).

References

1. Aakra A, Nyquist OL, Snipen L, Reiersen TS, Nes IF (2007) Survey of genomic diversity among *Enterococcus faecalis* strains by microarray-based comparative genomic hybridization. Appl Environ Microbiol 73:2207
2. Acinas SG, Klepac-Ceraj V, Hunt DE, Pharino C, Ceraj I, Distel DL, Polz MF (2004) Fine-scale phylogenetic architecture of a complex bacterial community. Nature 430:551
3. Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V, Polz MF (2005) PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. Appl Environ Microbiol 71:8966
4. Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997) The complete genome sequence of *Escherichia coli* K-12. Science 277:1453
5. Brousseau R, Hill JE, Prefontaine G, Goh SH, Harel J, Hemmingsen SM (2001) *Streptococcus suis* serotypes characterized by analysis of chaperonin 60 gene sequences. Appl Environ Microbiol 67:4828
6. Cohan FM (2002) What are bacterial species? Annu Rev Microbiol 56:457
7. Devriese LA, Hommez J, Pot B, Haesebrouck F (1994) Identification and composition of the streptococcal and enterococcal flora of tonsils, intestines and faeces of pigs. J Appl Bacteriol 77:31
8. Devriese LA, Pot B, Collins MD (1993) Phenotypic identification of the genus *Enterococcus* and differentiation of phylogenetically distinct enterococcal species and species groups. J Appl Bacteriol 75:399
9. Dorsch M, Stackebrandt E (1992) Some modifications in the procedure of direct sequencing of PCR amplified 16S rDNA. J Microbiol Methods 16:271
10. Dumonceaux TJ, Hill JE, Hemmingsen SM, Van Kessel AG (2006) Characterization of intestinal microbiota and response to dietary virginiamycin supplementation in the broiler chicken. Appl Environ Microbiol 72:2815
11. Dumonceaux TJ, Hill JE, Pelletier C, Paice MG, Van Kessel AG, Hemmingsen SM (2006) Molecular characterization of microbial communities in Canadian pulp and paper activated sludge and quantification of a novel *Thiothrix eikelboomii*-like bulking filament. Can J Microbiol 52:494
12. Felsenstein J (1993) PHYLIP (Phylogeny Infererence Package) version 3.5c, distributed by the author. Department of Genetics, University of Washington, Seattle
13. Gans J, Wolinsky M, Dunbar J (2005) Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science 309:1387

14. Goh SH, Facklam RR, Chang M, Hill JE, Tyrrell GJ, Burns EC, Chan D, He C, Rahim T, Shaw C, Hemmingsen SM (2000) Identification of *Enterococcus* species and phenotypically similar *Lactococcus* and *Vagococcus* species by reverse checkerboard hybridization to chaperonin 60 gene sequences. J Clin Microbiol 38:3953
15. Hill JE, Goh SH, Money DM, Doyle M, Li A, Crosby WL, Links M, Leung A, Chan D, Hemmingsen SM (2005) Characterization of vaginal microflora of healthy, nonpregnant women by chaperonin-60 sequence-based methods. Am J Obstet Gynecol 193:682
16. Hill JE, Hemmingsen SM, Goldade BG, Dumonceaux TJ, Klassen J, Zijlstra RT, Goh SH, Van Kessel AG (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
17. Hill JE, Paccagnella A, Law K, Melito PL, Woodward DL, Price DL, Ng L-K, Hemmingsen SM, Goh SH (2006) Identification of *Campylobacter* spp. and discrimination from *Helicobacter* and *Arcobacter* spp. by direct sequencing of PCR-amplified cpn60 sequences and comparison to cpnDB, a chaperonin reference sequence database. J Med Microbiol 55:393
18. Hill JE, Penny SL, Crowell KG, Goh SH, Hemmingsen SM (2004) cpnDB: a chaperonin sequence database. Genome Res 14:1669
19. Hill JE, Seipp RP, Betts M, Hawkins L, Van Kessel AG, Crosby WL, Hemmingsen SM (2002) Extensive profiling of a complex microbial community by high-throughput sequencing. Appl Environ Microbiol 68:3055
20. Inoue R, Tsukahara T, Nakanishi N, Ushida K (2005) Development of the intestinal microbiota in the piglet. J Gen Appl Microbiol 51:257
21. Jian W, Zhu L, Dong X (2001) New approach to phylogenetic analysis of the genus *Bifidobacterium* based on partial HSP60 gene sequences. Int J Syst Evol Microbiol 51:1633
22. Jordan IK, Rogozin IB, Wolf YI, Koonin EV (2002) Microevolutionary genomics of bacteria. Theor Popul Biol 61:435
23. Kohara T, Hayashi S, Hamatsuka J, Nishitani Y, Masuda Y, Osawa R (2007) Intra-specific composition and succession of *Bifidobacterium longum* in human feces. Syst Appl Microbiol 30:165
24. Mackie RI, Sghir A, Gaskins HR (1999) Developmental microbial ecology of the neonatal gastrointestinal tract. Am J Clin Nutr 69:1035S
25. Maidak BL, Cole JR, Lilburn TG, Parker CT Jr, Saxman PR, Farris RJ, Garrity GM, Olsen GJ, Schmidt TM, Tiedje JM (2001) The RDP-II (Ribosomal Database Project). Nucleic Acids Res 29:173
26. Makarova KS, Koonin EV (2007) Evolutionary genomics of lactic acid bacteria. J Bacteriol 189:1199
27. McBride SM, Fischetti VA, Leblanc DJ, Moellering RC Jr, Gilmore MS (2007) Genetic diversity among *Enterococcus faecalis*. PLoS ONE 2:e582
28. Minana-Galbis D, Urbizu-Serrano A, Farfan M, Fuste MC, Loren JG (2009) Phylogenetic analysis and identification of *Aeromonas* species based on sequencing of the cpn60 universal target. Int J Syst Evol Microbiol. doi:ijs.0.005413-005410
29. Ott SJ, Musfeldt M, Wenderoth DF, Hampe J, Brant O, Folsch UR, Timmis KN, Schreiber S (2004) Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. Gut 53:685
30. Paulsen IT, Banerjei L, Myers GS, Nelson KE, Seshadri R, Read TD, Fouts DE, Eisen JA, Gill SR, Heidelberg JF, Tettelin H, Dodson RJ, Umayam L, Brinkac L, Beanan M, Daugherty S, DeBoy RT, Durkin S, Kolonay J, Madupu R, Nelson W, Vamathevan J, Tran B, Upton J, Hansen T, Shetty J, Khouri H, Utterback T, Radune D, Ketchum KA, Dougherty BA, Fraser CM (2003) Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. Science 299:2071
31. Perna NT, Plunkett G 3rd, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Posfai G, Hackett J, Klink S, Boutin A, Shao Y, Miller L, Grotbeck EJ, Davis NW, Lim A, Dimalanta ET, Potamousis KD, Apodaca J, Anantharaman TS, Lin J, Yen G, Schwartz DC, Welch RA, Blattner FR (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature 409:529
32. Perriere G, Gouy M (1996) WWW-query: an on-line retrieval system for biological sequence banks. Biochimie 78:364
33. Pieper R, Janczyk P, Zeyner A, Smidt H, Guiard V, Souffrant WB (2008) Ecophysiology of the developing total bacterial and *Lactobacillus* communities in the terminal small intestine of weaning piglets. Microb Ecol 56:474
34. Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 16:276
35. Savage DC (1977) Microbial ecology of the gastrointestinal tract. Annu Rev Microbiol 31:107
36. Solheim M, Aakra A, Snipen LG, Brede DA, Nes IF (2009) Comparative genomics of *Enterococcus faecalis* from healthy Norwegian infants. BMC Genomics 10:194
37. Staden R, Beal KF, Bonfield JK (2000) The Staden package, 1998. Methods Mol Biol 132:115
38. Swords WE, Wu CC, Champlin FR, Buddington RK (1993) Postnatal changes in selected bacterial groups of the pig colonic microflora. Biol Neonate 63:191
39. Wall R, Fitzgerald G, Hussey S, Ryan T, Murphy B, Ross P, Stanton C (2007) Genomic diversity of cultivable *Lactobacillus* populations residing in the neonatal and adult gastrointestinal tract. FEMS Microbiol Ecol 59:127
40. Xu J (2006) Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. Mol Ecol 15:1713