The ISME Journal (2010) 4, 377–387

www.nature.com/ismej

ORIGINAL ARTICLE

Diversification of the gut symbiont *Lactobacillus reuteri* as a result of host-driven evolution

Phaik Lyn Oh[1], Andrew K Benson[1], Daniel A Peterson[1], Prabhu B Patil[1], Etsuko N Moriyama[2], Stefan Roos[3] and Jens Walter[1]

[1]*Department of Food Science and Technology, University of Nebraska, Lincoln, NE, USA;* [2]*School of Biological Sciences and Center for Plant Science Innovation, University of Nebraska, Lincoln, NE, USA and* [3]*Department of Microbiology, Swedish University of Agricultural Sciences, Uppsala, Sweden*

The vertebrate digestive tract, including that of humans, is the habitat to trillions of bacteria that are of significant importance to host biology and health. Although these communities are often postulated to have coevolved with their hosts, evidence is lacking, yet critical for our understanding of microbial symbiosis in vertebrates. To gain insight into the evolution of a gut symbiont, we have characterized the population genetic structure and phylogeny of *Lactobacillus reuteri* strains isolated from six different host species (human, mouse, rat, pig, chicken and turkey) using Amplified-Fragment Length Polymorphism (AFLP) and Multi-Locus Sequence Analysis (MLSA). The results revealed considerable genetic heterogeneity within the *L. reuteri* population and distinct monophyletic clades reflecting host origin but not provenance. The evolutionary patterns detected indicate a long-term association of *L. reuteri* lineages with particular vertebrate species and host-driven diversification. Results from a competition experiment in a gnotobiotic mouse model revealed that rodent isolates showed elevated ecological performance, indicating that evolution of *L. reuteri* lineages was adaptive. These findings provide evidence that some vertebrate gut microbes are not promiscuous, but have diversified into host-adapted lineages by a long-term evolutionary process, allowing the development of a highly specialized symbiosis.

The ISME Journal (2010) **4,** 377–387; doi:10.1038/ismej.2009.123; published online 19 November 2009
Subject Category: evolutionary genetics
Keywords: microbiome; evolution; population structure; *lactobacillus reuteri*

Introduction

The vertebrate gastrointestinal tract (GIT) is the habitat to a collection of microbes that is enormous in density and diversity, most of them bacteria. This community exerts an important influence on the development and performance of the vertebrate host through extensive immunological and metabolic interactions (Hooper and Gordon, 2001; Rakoff-Nahoum *et al.*, 2004; Peterson *et al.*, 2007; Li *et al.*, 2008; Round and Mazmanian, 2009). Despite the importance of gut microbes for vertebrate biology, little is known about how this symbiosis evolved and the selective pressures that shaped it. In general, animals evolved in a 'microbial soup' with pathogenic infection as a constant selection pressure, leading to the emergence of multiple immune strategies to confine microbes to the intestinal lumen (Cerutti and Rescigno, 2008). Beyond merely containing microbes, however, vertebrates developed mechanisms to preserve community composition as they gained benefits from the microbes (nutrients, defense, regulation of immune responses), and interestingly, some host responses to the gut microbiota have evolved early as they are conserved in mice and zebrafish (Rawls *et al.*, 2004). From the microbial perspective, inference from comparative genomic analysis of representative genera (*Bacteroides*, *Bifidobacterium*) revealed that gut bacteria possess features highly adapted for life in the digestive tract, having evolved sophisticated systems of energy utilization and coexistence with the host's immune system (Schell *et al.*, 2002; Xu *et al.*, 2003, 2007; Peterson *et al.*, 2007). An open question is whether vertebrate gut microbes share long-term evolutionary histories with particular host species or if they have evolved to become generalists for gut ecosystems.

Despite lack of definitive evidence, coevolution, the reciprocal adaptation occurring between interacting species (Moran, 2006), is now generally postulated to account for the development of the microbial symbiosis between humans and their gut microbes (Dethlefsen *et al.*, 2007; Claus *et al.*, 2008; Round and Mazmanian, 2009). Coevolution has largely been inferred by 16S rRNA surveys that revealed that different host species sustain

Correspondence: J Walter, Department of Food Science and Technology, University of Nebraska, 333 Food Industry Complex, Lincoln, NE 68583-0919, USA.
E-mail: jwalter2@unl.edu
Received 27 August 2009; revised 15 October 2009; accepted 16 October 2009; published online 19 November 2009

microbial communities containing lineages that are host specific (Dethlefsen *et al.*, 2007). However, this does not preclude the possibility that the specific environmental conditions in individual vertebrates simply select for phylotypes that are generalists for gut environments. In addition, many phylotypes are associated with several host species, and these lineages could be environmentally acquired. Nevertheless, Ley *et al.* (2008a), who analyzed microbial populations of 59 mammalian species by a 16S rRNA survey, provided evidence that the patterns of community similarity matched the mammal phylogeny and that some lineages co-diversified with their mammalian hosts. However, as many phylotypes were shared across host species, the authors concluded that gut microbes may be fairly promiscuous between hosts. This view contrasts with phenotypic characterizations of gut microbes that clearly showed that bacterial colonization and epithelial adherence can be host specific (Lin and Savage, 1984; Tannock *et al.*, 1984). An important question is whether host specialization, when observed, is the outcome of coevolution or just the result of short-term adaptation by a promiscuous gut microbe after invading a new vertebrate species.

The majority of studies that infer the evolution of gut microbes used 16S rRNA surveys of whole bacterial communities (Dethlefsen *et al.*, 2007; Ley *et al.*, 2008a, b). Their disadvantage is that they underestimate the real diversity of the vertebrate microbiome as 16S rRNA sequences possess insufficient genetic resolution to differentiate all bacterial groups that have ecologically distinct roles (Koeppel *et al.*, 2008). Therefore, many evolutionary events will remain undetected, especially if lineage diversification and specialization occurred more recently. A more sensitive and systematic strategy is, therefore, to examine the phylogeny of a single bacterial species that inhabits multiple vertebrate hosts by population genetic approaches. In this instance, the signatures of evolution can be compared directly because of the shared ancestry of the genomic content. If gut microbes have coevolved with their host species, one would hypothesize to find distinct subpopulations, each represented by a monophyletic clade associated with a specific vertebrate host species.

The Gram-positive bacterium *Lactobacillus reuteri* is an excellent model organism to study evolutionary mechanisms of a vertebrate gut symbiont and test the hypothesis of coevolution as this species stably inhabits the GIT of mammals as diverse as humans, pigs, mice and rats as well as different species of birds. In rodents, pigs and chickens, *L. reuteri* is one of the dominant species in the GIT (Leser *et al.*, 2002; Salzman *et al.*, 2002; Brooks *et al.*, 2003; Abbas Hilmi *et al.*, 2007). In these animals, *L. reuteri* colonizes the surfaces of the stratified squamous epithelial lining of the proximal regions of the digestive tract (Tannock, 1992; Walter *et al.*, 2007). The ecology of *L. reuteri* in the human gut is less clear and the organism is only detected in some subjects (Walter, 2008). Nevertheless, the type-strain of the species, DSM 20016^{T}, was detected in a human subject over several months, and *L. reuteri* has been considered autochthonous to the human gut (Reuter, 2001). In this study, we have investigated the population structure of a large collection of *L. reuteri* strains from multiple vertebrate hosts to gain an understanding of the evolution of this vertebrate symbiont. Our data provide remarkable evidence for the diversification of the *L. reuteri* population into ecotypes that are highly host-specific.

Materials and methods

Bacterial isolates and growth conditions

A total of 165 *L. reuteri* strains isolated from different animal hosts from global geographic locations were used. Information on all strains is provided in the Supplementary Table S1. Isolates were obtained from humans ($n=35$), pigs ($n=41$), chickens ($n=28$), turkeys ($n=5$), mice ($n=35$) and rats ($n=26$). Human isolates were obtained mainly from fecal samples but also from breast milk, vagina, stomach and the oral cavity. Animal isolates were obtained from GIT or fecal samples. Bacteria were grown anaerobically using modified MRS (mMRS) medium at 37 °C for 48 h. mMRS is MRS medium (Difco, BD, Franklin Lakes, NJ, USA) supplemented with maltose (10 g per liter) and fructose (5 g per liter). DNA isolation and bacterial identification by 16S rRNA sequence analysis was performed as described in the Supplementary materials.

Amplified-Fragment Length Polymorphism (AFLP)

AFLP was performed using the AFLP template preparation kit available from LI-COR Biosciences (Lincoln, Nebraska, USA), with slight modifications to the manufacturer's protocol (details are given in Supplementary materials). For each strain, two individual AFLP electropherograms were generated with primers that had different specific nucleotide overhangs (Supplementary Table S4). AFLP images analysis and PCA was performed using Bionumerics software Version 5.0 (Applied Maths, Kortrijk, Belgium). AFLP dendrogram was constructed using the neighbor-joining (NJ) algorithm implemented in PAUP* 4.0b10.

Multi-Locus Sequence Analysis (MLSA)

MLSA was performed with seven housekeeping genes: D-alanine-D-alanine ligase (*ddl*), phosphoketolase (*pkt*), leucyl-tRNA synthetase (*leuS*), DNA gyrase B subunit (*gyrB*), D-alanine-D-alanyl carrier protein ligase (*dltA*), RNA polymerase alpha subunit (*rpoA*) and recombinase (*recA*) using standard methods (details are given in Supplementary materials). The nucleotide sequences of all MLSA

loci used during this study have been submitted to Genbank under the following database accession numbers: *ddl* (GQ378230—GQ378345); *pkt* (GQ378694—GQ378809); l*euS* (GQ378578—GQ378693); *gyrB* (GQ378462—GQ378577); *dltA* (GQ378346—GQ378461); *rpoA* (GQ378926—GQ379041); and *recA* (GQ378810—GQ378925). Descriptive and phylogenetic analysis of MLSA data and the test for recombination using SplitsTree4 (version 4.10) was performed as described in the Supplementary materials.

Reconstruction of genealogies using the ClonalFrame software

Sequences from all seven MLSA loci were analyzed using ClonalFrame (Didelot and Falush, 2007). ClonalFrame uses a Bayesian framework assuming a neutral coalescent model. Five independent runs were performed, each consisting of 300 000 iterations (100 000 burn-in iterations and 200 000 post-burn-in iterations). Phylogenies were sampled every 100 iterations after the burn-in, resulting in a total of 10 000 sampled trees. Analyses from five independent runs were concatenated to construct 50% majority-rule consensus tree. Genealogies were drawn using the radial tree option in MEGA4 software package. ClonalFrame analyses were performed with all strains and clade-by-clade to obtain evolutionary parameters such as the relative impact of recombination as compared with point mutation ($r\,m^{-1}$).

Competition experiment in ex-germ-free (GF) mice

Equal volumes of cell solutions of *L. reuteri* strains were administered to germ-free Swiss Webster mice housed under GF conditions at the University of Nebraska gnotobiotic facility (details are given in Supplementary materials). Fecal samples of individual mice were taken at days 1, 4, 8, 11, and mice were killed on day 11. The strain composition in fecal samples as well as on the forestomach epithelium and the cecum was determined by pyrosequencing of PCR amplicons of the leucyl-tRNA synthetase (*leuS*) gene using a 454 GS-FLX sequencer (Roche, Basel, Switzerland). The sequencing resulted in an average of 12 739 sequence tags per sample. A local nucleotide database was established in Bioedit v7.0.9 for each sample, and the BLASTn algorithm was used to determine the number of sequences that account for each individual strain.

Results

AFLP and MLSA genotyping revealed highly diverged, host-confined subpopulations of L. reuteri

To gain insight into the population structure of *L. reuteri*, we investigated 165 strains of *L. reuteri* originating from humans, pigs, mice, rats, chickens and turkeys from global geographic locations (Supplementary Table S1). Comparisons of partial 16S rRNA gene sequences (700 bp) of all strains revealed an identity of at least 99.4% to the sequence of the type strain (DSM20016^T), indicating that these strains fall within what is currently accepted as the species *L. reuteri*. As a first step to inferring the population structure within this species, AFLP analysis was performed using two independent primer pairs (the electropherograms are shown in Supplementary Figure S1). Subsequent phylogenetic analysis by the neighbor-joining (NJ) method revealed significant genetic diversity within the *L. reuteri* population (Figure 1a). Remarkably, the phylogeny revealed several distinct clusters that were almost completely host specific. Three rodent (B, D, E), two pig (A, Cii), and one human cluster (Ciii) were detected, and the poultry strains formed a large cluster together with a second group of human isolates (Ci). Isolates from rodents (mice and rats) and poultry (chicken and turkey) grouped in the same clades (Supplementary Table S1). Principal component analysis (PCA) of AFLP data revealed that the genetically distinct subpopulations of *L. reuteri* group according to host origin but not by geographic origin (Figure 1b).

To validate the phylogenetic relationships inferred from the AFLP data, 116 strains were examined by MLSA. The strains were selected to be equally distributed in the AFLP dendrogram while maintaining maximal geographic diversity (Supplementary Table S1). Seven conserved gene loci (*ddl*, *pkt*, *leuS*, *gyrB*, *dltA*, *rpoA* and *recA*) were amplified by PCR, sequenced and analyzed. The descriptive analysis of the MLSA data is summarized in Supplementary Table S2. Significant sequence diversity within the *L. reuteri* population was detected. The number of alleles per locus ranged from 16 to 39 and the number of polymorphic nucleotide sites from 16 to 88. The low dN/dS ratios indicated that all genes are under purifying selection as expected from highly conserved housekeeping genes.

MLSA revealed a significant association of multilocus genotypes with particular host species. The 116 strains represented 60 genotypes (or sequence types, STs), and virtually all STs were specific for a particular host (Supplementary Table S1). This indicates that specific clones of *L. reuteri* are dominant in particular host species. To infer the evolutionary relationships within *L. reuteri*, we reconstructed a phylogenetic tree with the concatenated sequence of the seven housekeeping genes (in total 3822 bp) using the maximum-likelihood (ML) method. The ML tree is shown in Figure 2a (the same ML tree with strain names is provided in Supplementary Figure S2a). The analysis revealed six subpopulations that were almost completely host specific and were referred to as I (rodents 2), II (human), III (rodents 1), IV (pig 1), V (pig 2) and VI (poultry/human). Despite the excellent reflection of host origin, the clusters identified by ML showed

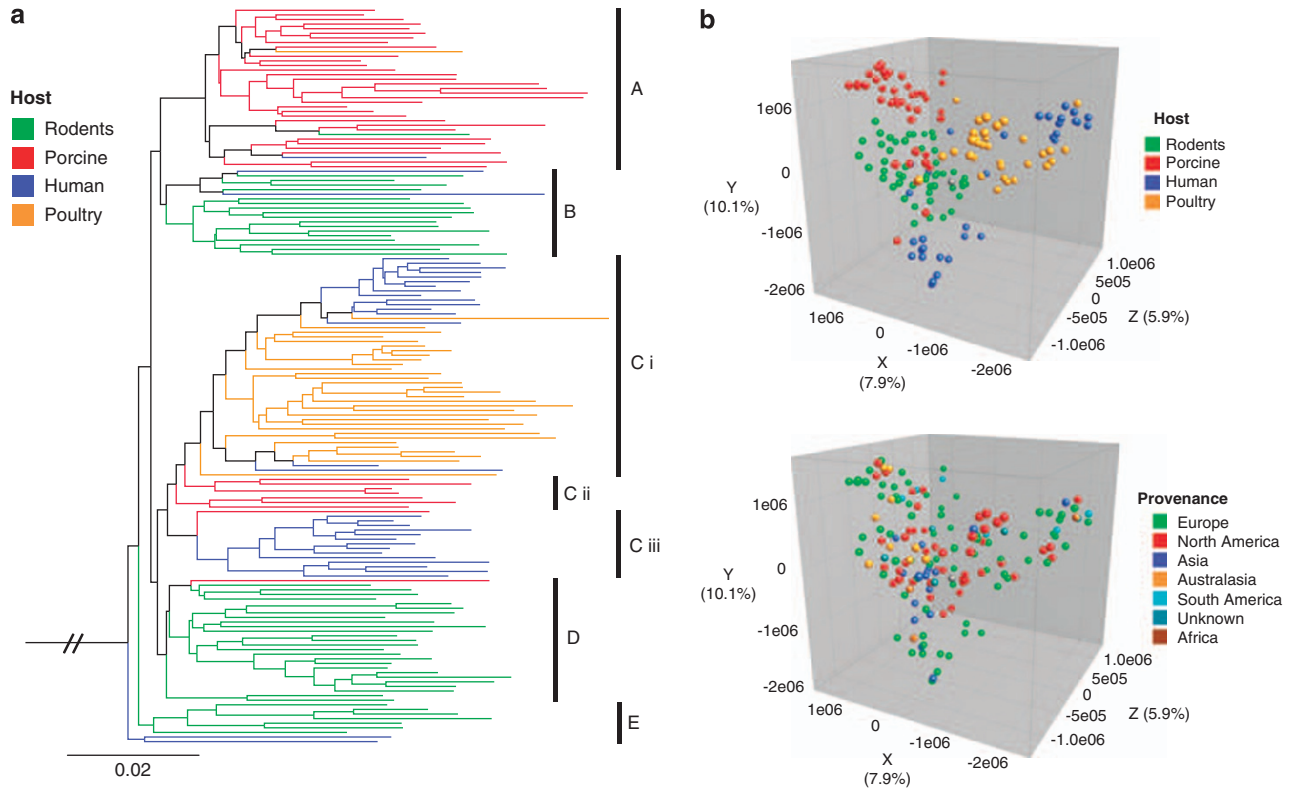

Figure 1 AFLP analysis of a global collection of *Lactobacillus reuteri* strains from humans, rodents (mice and rats), pigs and poultry (chicken and turkey). (**a**) The neighbor-joining phylogeny of 165 strains based on 439 markers obtained from two AFLP primer pairs. Branches are colored according to host origin. *Lactobacillus coryniformis* (Li146:1, not shown in the figure) was used as the outgroup to root the *L. reuteri* phylogeny. (**b**) Principal component analysis of AFLP data. Plots are based on the values of the first three principal components with each circle corresponding to an individual strain colored according to host origin or geographic origin (provenance).

some degree of 'fuzziness', meaning that not all strains of the subpopulations formed discrete phylogenetic groups. Nevertheless, there was excellent correlation between the MLSA and AFLP genotypes, as both techniques identified subpopulations that contained virtually the same strains (Supplementary Table S1). For instance, the composition of strains in MLSA clusters IV, V and VI was identical with the composition of the corresponding AFLP clusters A, Cii and Ci. In accordance with the AFLP analysis, isolates from rodents (mice and rats) and poultry (chicken and turkey) fall within the same MLSA clusters.

Host-specific clusters differ markedly in genetic heterogeneity

A remarkable feature of the *L. reuteri* population structure inferred by MLSA was the substantial difference in the genetic heterogeneity observed among individual host-specific clades. For example, clusters II (human), IV (Pig 1), V (Pig 2), and VI (poultry/human) had very low sequence diversity, whereas the two rodent-specific clusters I and III were highly heterogeneous (Supplementary Table S3). These differences were confirmed using the eBURST algorithm, which infers patterns of evolutionary descent by grouping strains into clonal complexes (groups of strains that share a minimum of five out of the seven alleles). This analysis classified the 60 STs into 9 clonal complexes (CCs) and 29 singletons. As shown in Figure 2a, all strains within MLSA cluster IV (pig 1) and most strains of cluster II (human) fell within the CC-3 and CC-47, respectively. Several clonal complexes and one ST (containing 13 isolates) could also be identified in clade VI (poultry/human). As expected, many of the identified clonal complexes correspond to the MLSA clusters supported by high bootstrap values (>94%) in the phylogeny. In contrast, most strains of the rodent-specific clusters I and III did not fall within CCs but represented 'singletons'.

The diversification of L. reuteri *took millions of years*

To estimate the time that was required for *L. reuteri* diversification, we determined the age of the last common ancestor using a molecular clock with the MLSA data set. The mean pairwise difference at synonymous sites (Ds) derived from the concatenated sequences of all seven loci from 116 *L. reuteri* strains was 0.122. Using the previously reported synonymous substitution rate of 4.7×10^{-9} per site per year for *Escherichia coli*/*Salmonella typhimurium* and 8.2×10^{-9} per site per year for obligate

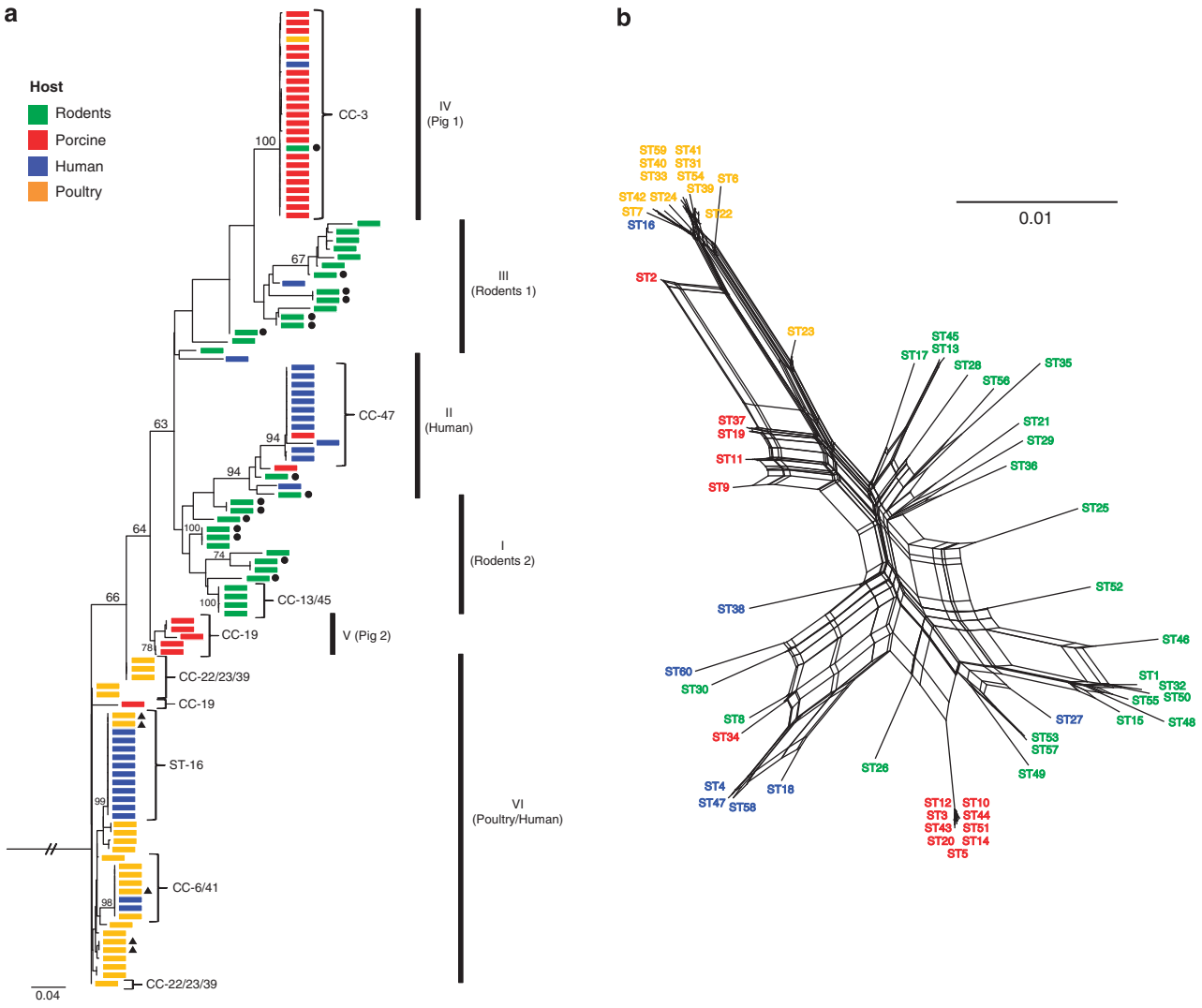

Figure 2 Phylogenetic analysis of 116 strains of *Lactobacillus reuteri* based on the concatenated sequences of seven loci. (**a**) The maximum-likelihood tree was reconstructed with PhyML using the GTR + I + G model of nucleotide substitution and 100 bootstrap replicates. *Lactobacillus vaginalis* ATCC49540 (not shown in the figure) was used as the outgroup to root the *L. reuteri* phylogeny. Bootstrap values above 50% are shown at the nodes. Clonal complexes (CC) are labeled with a number corresponding to its predicted founder genotype (as identified by eBURST). Only CCs that contain more than four strains are shown. Isolates from rats are labeled with a circle, and isolates from turkey are labeled with a triangle. (**b**) Phylogenetic network of the 60 sequence types (STs) are present in the *L. reuteri* population. Conflicting phylogenetic signals are visualized by the network structure in the center.

symbiotic bacteria, *Buchnera aphidicola* (Lawrence and Ochman, 1998; Ochman *et al.*, 1999), the average age of the cenancestor of the *L. reuteri* strains studied here was estimated to be 7.6–12.95 million years. Although time scales obtained with molecular clocks should be taken with caution as accelerated rates of evolution have been detected for some bacteria (Wilson *et al.*, 2009), our data nevertheless indicates a long-term association of *L. reuteri* lineages with particular vertebrate species.

Recombination is an important factor in the evolution of L. reuteri

It is now well established that bacteria do not necessarily conform to the clonal model of evolution (Spratt *et al.*, 2001). Recombination leads to the replacement of short pieces of DNA with the homologous segments from other strains, and in some species, it contributes more to diversification than point mutations. It can often confound the reconstruction of phylogenetic relationships, and the importance of recombination in *L. reuteri* was made manifest in the highly incongruent topologies of phylogenetic trees reconstructed from individual housekeeping genes (although host specificity was still observed for all individual loci) (Supplementary Figure S3). It is therefore likely that the 'fuzzy' topology of the ML tree (Figure 2a) was caused by recombination, as described previously for other recombinogenic bacteria (Hanage *et al.*, 2005). We used the split-decomposition method, implemented in the software SplitsTree4, to detect the signatures of recombination in the MLSA data set (Huson,

1998). Split-decomposition does not force sequence information into a tree-like phylogeny, and the presence of conflicting phylogenetic signals results in a net-like structure. This analysis revealed extensive networked evolution for *L. reuteri*, which is indicative of recombination, for the concatenated sequence (Figure 2b) and most of the individual genes (Supplementary Figure S4). The Pairwise Homoplasy Index test revealed statistically significant evidence for intragenic recombination for the following genes: *ddl*, *leuS*, *gyrB* and *dltA* (Supplementary Table S2). Remarkably, despite extensive recombination, host-specific groupings were still detected in the tree reconstructed with SplitsTree4 (Figure 2b).

Reconstruction of a more accurate genealogy

In an attempt to deduce a more accurate evolutionary history of *L. reuteri*, we reconstructed an ML phylogenetic tree with the concatenated sequence of the three loci that, according to the PHI statistical test, were not significantly affected by recombination (*pkt*, *rpoA* and *recA*). This tree showed a similar topology to the tree reconstructed with all seven loci (Supplementary Figure S2). We further applied the ClonalFrame software, which uses a coalescent-based Bayesian method to infer strained relationships. ClonalFrame excludes putative recombinant regions in its analysis, thus estimating real genealogies more accurately (Didelot and Falush, 2007). As shown in Figure 3, the genealogy inferred by ClonalFrame featured host-specific branches with clear demarcations with little reflection of geographic origin. A determination of relative recombination rates ($r\,m^{-1}$) with ClonalFrame revealed an overall $r\,m^{-1}$ value of 3.0, and variable levels of recombination in individual subpopulations (Table 1). Despite clear evidence for recombination in *L. reuteri*, the clusters identified by ML phylogeny and ClonalFrame were in almost complete agreement (Figures 2a and 3). Therefore, we conclude that although recombination is prominent in *L. reuteri*, it does not confound the ancestry of individual lineages.

Ecological fitness of L. reuteri *strains reflects host origin*

To gain insight into the evolutionary factors that led to the development of host-confined subpopulations of *L. reuteri*, we performed a competition experiment in mice. The objective of this experiment was to determine if evolution of *L. reuteri* was adaptive, which would result in specialization and fitness benefits, or merely neutral. We introduced an equal mixture of nine *L. reuteri* strains originating from a

Table 1 Relative recombination rate of *L. reuteri* population and individual lineages as inferred by ClonalFrame[a]

Population	$r\,m^{-1}$[b]
All strains	3.0 (2.1–4.2)
I (rodents 2)	4.1 (1.8–8.5)
II (human)	1.8 (0.6–4.1)
III (rodents 1)	3.1 (1.6–5.6)
IV (pig 1)	1.0 (0.003–7.9)
V (pig 2)	N/A
VI (poultry/human)	0.5 (0.05–1.6)

[a]Mean value of the parameters with the 95% credibility interval given in brackets.
[b]Relative impact of recombination as compared with point mutation in the genetic diversification of the lineage.

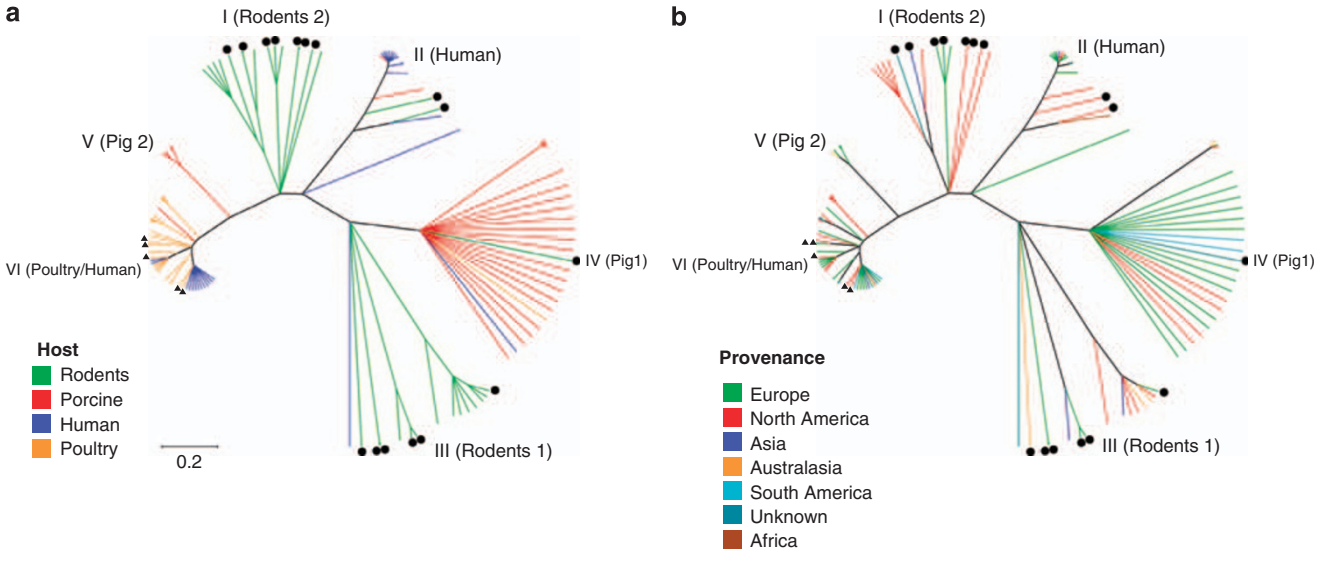

Figure 3 *L. reuteri* genealogy inferred by ClonalFrame using a 50% majority-rule consensus tree from five independent runs. Scale bar in the ClonalFrame genealogy represents time in coalescent units. Isolates from rats are labeled with a circle, and isolates from turkey are labeled with a triangle. Trees are color coded by host (**a**) and geographic origin (**b**).

mouse (1), rats (2), humans (2), chickens (2) and pigs (2) together with a mouse isolate *Lactobacillus johnsonii* DB-C4 into germ-free mice. *L. johnsonii* was added as this species co-exists with *L. reuteri* in natural populations of the rodent forestomach (Salzman *et al.*, 2002; Brooks *et al.*, 2003). Each of the *L. reuteri* strains used in the experiment carried a distinct allele of the *leuS* gene, allowing their differentiation. The composition of the *L. reuteri* population in the inoculum, fecal samples (days 1, 4, 8 and 11) and the forestomach and cecum (day 11) was characterized by 454 pyrosequencing of *leuS* amplicons obtained by PCR. This analysis showed that two of the rodent strains possessed a very obvious competitive advantage (Figure 4). The rodent isolates 6799jm-1 (mouse) and R2LC (rat) outgrew all other strains and comprised around 97% of the total population on the forestomach epithelium, the natural niche of *L. reuteri*. This competition experiment confirmed earlier studies on the host specificity of *L. reuteri* strains in gut colonization (Molin *et al.*, 1992; Carbajal *et al.*, 1999; Schreiber *et al.*, 2009), and together, the results clearly show that the phylogenetic relationships discovered for *L. reuteri* do reflect not only host origin but also ecological fitness.

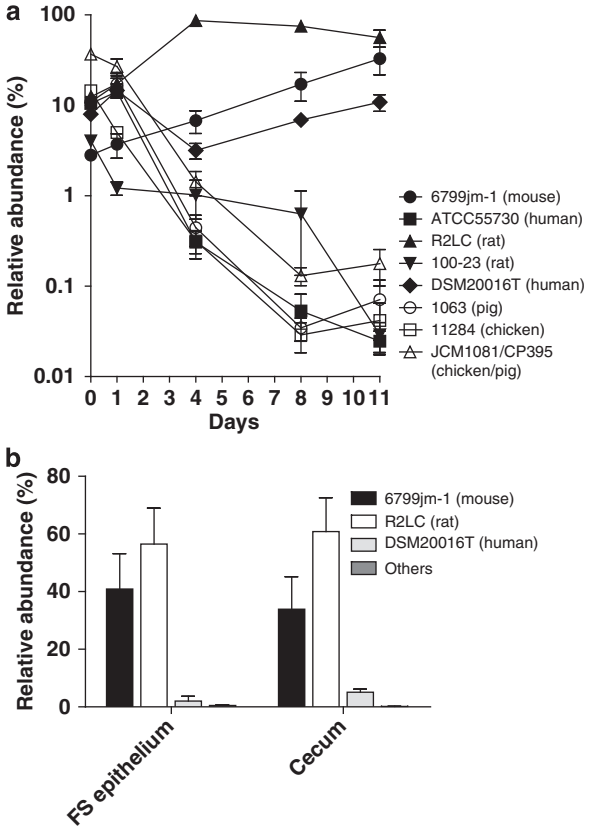

Figure 4 Test of ecological fitness of *L. reuteri* strains in the digestive tract of ex-germ-free mice. Germ-free mice were colonized with an equal mixture of *L. reuteri* strains 6799jm-1 (mouse), R2LC (rat), 100-23 (rat), DSM20016T (human), ATCC55730 (human), 1063 (pig), CP395 (pig), 11284 (chicken), and JCM1081 (chicken) and a mouse isolate *Lactobacillus johnsonii* DB-C4. Relative abundance of individual *L. reuteri* strains in populations was determined by pyrosequencing of the leucyl-tRNA synthetase (*leuS*) gene. (**a**) Abundance (%) of strains in fecal samples taken during the experiment (11 days). Means with s.e.m. are shown ($n = 5$). (**b**) Abundance after 11 days on the forestomach epithelium and cecum of mice. Means with s.e.m. are shown ($n = 5$).

Discussion

Symbiotic microorganisms form intimate associations with most members of the animal kingdom (Moran, 2006). These partnerships vary in how they evolved, their mechanisms for achieving and maintaining associations, and the roles played by the microbe in host biology. To date, the evolution of microbial symbiosis is best understood in invertebrate endosymbionts, where evolutionary mechanisms have been characterized and coevolution has been convincingly proven in many cases (Dale and Moran, 2006; Moran *et al.*, 2008). The phylogenetic analysis of a global *L. reuteri* population presented here now allows, for the first time, insight into the evolutionary history of a vertebrate symbiont. The most striking finding with both AFLP and MLSA was the clear reflection of host origin within the *L. reuteri* population structure and phylogeny. As provenance was not reflected, it appears that the distinct ecological conditions within the digestive tract of vertebrate species were the major forces that drove diversification of the *L. reuteri* ancestral population into host-specific lineages. The results also predict a stable relationship of *L. reuteri* with vertebrate hosts, as the deep divisions between host-specific lineages indicate a long period of close association. One would assume that the maternal care of animals such as rodents and pigs would facilitate parental–offspring transmission, and for pigs and humans, maternal transmission of *L. reuteri* has been empirically shown (Tannock *et al.*, 1990; Casas and Dobrogosz, 2000). Although the data does not allow us to pinpoint the exact duration of stable symbiotic relationships, the high average sequence divergence between subpopulations (Supplementary Table S3) and the time span determined by a molecular clock suggested that *L. reuteri* lineages have evolved with particular hosts for millions of years.

How host specific is the evolution of L. reuteri*?*

The phylogeny of *L. reuteri* suggests that its long-term evolutionary relationships with vertebrates resulted in host-specific subpopulations. However, they do not appear to be absolutely constrained to individual host species. Isolates from rodents (mice and rats) and poultry (chicken and turkey) group in the same phylogenetic clusters. This suggests that the time for the divergence of truly host-specific lineages was not yet sufficient, or that *L. reuteri*

evolved with groups of closely related host species. Such a 'joint' evolution would result in the development of generalists, which are promiscuous in terms of their ability to colonize the GITs of all members of these groups. In fact, this can be observed in *L. reuteri* isolates from mice and rats, which show comparable ecological performance in the mouse digestive tract (Figure 4 and references: Molin *et al.*, 1992; Carbajal *et al.*, 1999; Schreiber *et al.*, 2009). In addition, murine *Lactobacillus* isolates show no host specificity in their adherence to epithelial tissues of mice and rats, whereas they generally do not adhere to epithelial cells of chickens and pigs (Savage, 1972; Lin and Savage, 1984; Tannock, 1992). Based on these findings, we hypothesize that some vertebrate gut symbionts evolve with groups of closely related host species that possess similar niches (that is, shared anatomy, host glycan composition, diet composition and innate immune receptors) in the GITs and whose social behavior allows horizontal transfer of bacteria.

Although the population analysis revealed clusters that reflect host origin, the clade VI contained not only poultry strains but also human strains. In addition, other essentially host-specific clusters contained some 'outliers' originating from unrelated hosts. These findings could indicate that some strains might be promiscuous and have recently switched hosts. However, we argue that a promiscuous lifestyle that entails common host jumps would not have resulted in the mostly host-specific subpopulations detected during this study. It is more likely that these 'outliers' are allochthonous to the particular host animal. Allochthonous bacteria are often present in the intestinal tract as they get introduced through the diet (Dal Bello *et al.*, 2003; Walter, 2008), but they are only transient and hence do not share an evolutionary history with the host (Savage, 1972). For example, *L. reuteri* strains originating from poultry might become transiently associated with humans just as *Salmonella* that often originate from chicken (Callaway *et al.*, 2008). Nevertheless, the detection of allochthonous strains originating from other species indicates that there is an opportunity for a horizontal transfer between hosts, and host switches are likely to have had an important role when *L. reuteri* became associated with a variety of vertebrate species. However, the population structure detected here indicates that such jumps between unrelated hosts have been rare events during the evolution of *L. reuteri*.

The evolution of L. reuteri *is adaptive*

As the genetic clusters of *L. reuteri* showed high cohesion and could be robustly linked to specific ecological settings (the gut of particular vertebrate species), we postulate that they represent ecotypes. Ecotypes are populations adapted to a particular set of environmental conditions whose members are both genetically and ecologically similar (Cohan and Perry, 2007). The population structure predicts the host environment as the main driver of diversification, although further unknown factors led to the diversification of rodent and pig lineages into two ecotypes. But what evolutionary process led to diversification of *L. reuteri*? Recent theoretical and empirical studies showed that diversification and genetic coherence within ecotypes can be facilitated by both natural selection and genetic drift (Fraser *et al.*, 2005, 2007; Cohan and Perry, 2007). It has been argued that genetic drift could be especially important for commensals, as host transmission and life in attached populations might reduce the effective population size (N_e) and generate population bottlenecks that favor neutral diversification (Cohan and Perry, 2007; Fraser *et al.*, 2007). Here, we provide evidence that there is more than neutral evolution and genetic drift in *L. reuteri* diversification. First, ecological fitness of *L. reuteri* reflects host origin in competition experiments. Second, the average ratio of homologous recombination to mutation rates ($r\,m^{-1}$) in *L. reuteri* is 3.0 (Table 1), which exceeds the clonal–sexual threshold allowing formation of clusters in the absence of selection (Fraser *et al.*, 2007). These findings are in agreement with conclusions made by Ley *et al.* (2008b) that selection and competition might have a significant role in gastrointestinal ecosystems.

Selective forces vary in different vertebrate hosts

The physiological, immunological and behavioral differences of host species are likely to constitute distinct challenges and selective pressures for their gut inhabitants. This is reflected in the *L. reuteri* population structure. The allelic diversity of housekeeping genes within the two rodent clades (1.72% and 1.53%) was much higher than the diversity within clusters of human, pig, and poultry origin (<0.87%), which were also much more clonal (Figure 2a). Diversity might have been purged in the human, porcine and poultry clusters through recent events of periodic selection in accordance with the stable ecotype model (Cohan and Perry, 2007). Periodic selection refers to a mutation event that improves the fitness of one strain such that its descendants out-compete all other strains within the ecotype. However, the rodent clusters were clearly too diverse to support a recent selective sweep, and processes such as clonal interference or the adaptation of individual strains to micro-niches might have contributed to the diversity within genetic clusters (Doolittle and Zhaxybayeva, 2009). Why diversity is much higher in the two rodent ecotypes remains to be evaluated. In *Helicobacter pylori*, genetic variability generated by both high mutation rates and recombination has long been thought to contribute to the adaptation to individual human hosts and bacterial transmission (Suerbaum and Josenhans, 2007). It is possible that the ability to introduce genetic variations, especially through a

high rate of intrastrain recombination (Table 1), is an important mechanism by which rodent *L. reuteri* lineages adapt to physiological differences of individual animals and, probably, mice and rats.

Mutualism as a result of coevolution?

A fascinating aspect of the evolution of vertebrate symbionts is that the host itself is under selective pressure (Ley *et al.*, 2006). Therefore, beneficial traits of gut bacteria that share long evolutionary history with their host species could have been shaped by natural selection as they promote host fitness, which in turn facilitates the microbe's dissemination (Dethlefsen *et al.*, 2007; Ley *et al.*, 2008b; Round and Mazmanian, 2009). *L. reuteri* is documented to benefit humans and animals and can therefore be considered a true symbiont of vertebrates (Shornikova *et al.*, 1997; Madsen *et al.*, 1999; Casas and Dobrogosz, 2000; Rosenfeldt *et al.*, 2003; Moller *et al.*, 2005; Pena *et al.*, 2005; Weizman *et al.*, 2005; Savino *et al.*, 2007). The phylogenetic relationships discovered in this study suggest that the beneficial attributes of *L. reuteri* strains could be the outcome of a long-term evolutionary process that resulted in a mutualistic relationship between microbe and host. An important practical value of the phylogeny established here is its use for the selection of strains for probiotics. It is a logical working hypothesis that strains that share a long evolutionary history with particular vertebrate species possess adaptive traits that benefit their hosts (Walter, 2008). Clearly, our findings establish a framework for future phylogenomic investigations to identify adaptive traits of *L. reuteri* important for both the microbe's ecological performance and its effects on the host.

In conclusion, this study provided crucial knowledge for our understanding of microbial symbiosis in the vertebrate gut, as it revealed that a gut microbe can share a long-term evolutionary history with particular vertebrate host species. By theory, such an evolutionary process would favor the development of a mutualistic relationship and might account for the important beneficial attributes of the gut microbiota (Ley *et al.*, 2006; Dethlefsen *et al.*, 2007). The findings obtained during this work strengthen the notion that a disruption of the ancient host–microbial partnership through a modern lifestyle (antibiotics, diet, formula feeding, hospital deliveries, hygiene, etc.) could have significant consequences for health and might have contributed to the recent increase in diseases that have been linked to the aberration of the gut microbiome (Ley *et al.*, 2005; Nicholson *et al.*, 2005; Cani *et al.*, 2007). Clearly, future research should be targeted to identify evolutionary strategies for other members of the vertebrate gut microbiota.

Acknowledgements

This study would not have been possible without the generous contributions of well characterized strains from many colleagues (the names are provided in the Supplementary information section). We further thank Dr Xavier Didelot (University of Warwick, Coventry, UK) for very helpful discussions and advice in the application of the ClonalFrame software. We are grateful to Drs Sheila Scheideler, Merlyn Nielsen and Thomas Burkey (University of Nebraska-Lincoln, USA) for the provision of gut samples from chicken, mice and pig. We thank Jaehyoung Kim, Inés Martínez, Christi Toa and Ryan Talley (University of Nebraska-Lincoln, USA) for excellent technical support. This work was funded by grants of the University of Nebraska and through support of BioGaia AB, Sweden.

References

Abbas Hilmi HT, Surakka A, Apajalahti J, Saris PE. (2007). Identification of the most abundant *Lactobacillus* species in the crop of 1- and 5-week-old broiler chickens. *Appl Environ Microbiol* **73**: 7867–7873.

Brooks SP, McAllister M, Sandoz M, Kalmokoff ML. (2003). Culture-independent phylogenetic analysis of the faecal flora of the rat. *Can J Microbiol* **49**: 589–601.

Callaway TR, Edrington TS, Anderson RC, Byrd JA, Nisbet DJ. (2008). Gastrointestinal microbial ecology and the safety of our food supply as related to Salmonella. *J Anim Sci* **86**: E163–E172.

Cani PD, Neyrinck AM, Fava F, Knauf C, Burcelin RG, Tuohy KM *et al.* (2007). Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a mechanism associated with endotoxaemia. *Diabetologia* **50**: 2374–2383.

Carbajal N, Casas IA, Dobrogosz WJ. (1999). Effect of host-specific *Lactobacillus reuteri* on ileal tissue development in gnotobiotic BALB/c mice. *Microbial Ecol Health Dis* **11**: (Abstract) 184.

Casas IA, Dobrogosz WJ. (2000). Validation of the probiotic concept: *Lactobacillus reuteri* confers broad-spectrum protection against disease in humans and animals. *Microb Ecol Health Dis* **12**: 247–285.

Cerutti A, Rescigno M. (2008). The biology of intestinal immunoglobulin A responses. *Immunity* **28**: 740–750.

Claus SP, Tsang TM, Wang Y, Cloarec O, Skordi E, Martin FP *et al.* (2008). Systemic multicompartmental effects of the gut microbiome on mouse metabolic phenotypes. *Mol Syst Biol* **4**: 219.

Cohan FM, Perry EB. (2007). A systematics for discovering the fundamental units of bacterial diversity. *Curr Biol* **17**: R373–R386.

Dal Bello F, Walter J, Hammes WP, Hertel C. (2003). Increased complexity of the species composition of lactic acid bacteria in human feces revealed by alternative incubation condition. *Microb Ecol* **45**: 455–463.

Dale C, Moran NA. (2006). Molecular interactions between bacterial symbionts and their hosts. *Cell* **126**: 453–465.

Dethlefsen L, McFall-Ngai M, Relman DA. (2007). An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* **449**: 811–818.

Didelot X, Falush D. (2007). Inference of bacterial microevolution using multilocus sequence data. *Genetics* **175**: 1251–1266.

Doolittle WF, Zhaxybayeva O. (2009). On the origin of prokaryotic species. *Genome Res* **19**: 744–756.

Fraser C, Hanage WP, Spratt BG. (2005). Neutral microepidemic evolution of bacterial pathogens. *Proc Natl Acad Sci USA* **102**: 1968–1973.

Fraser C, Hanage WP, Spratt BG. (2007). Recombination and the nature of bacterial speciation. *Science* **315**: 476–480.

Hanage WP, Fraser C, Spratt BG. (2005). Fuzzy species among recombinogenic bacteria. *BMC Biol* **3**: 6.

Hooper LV, Gordon JI. (2001). Commensal host-bacterial relationships in the gut. *Science* **292**: 1115–1118.

Huson DH. (1998). SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* **14**: 68–73.

Koeppel A, Perry EB, Sikorski J, Krizanc D, Warner A, Ward DM *et al.* (2008). Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc Natl Acad Sci USA* **105**: 2504–2509.

Lawrence JG, Ochman H. (1998). Molecular archaeology of the *Escherichia coli* genome. *Proc Natl Acad Sci USA* **95**: 9413–9417.

Leser TD, Amenuvor JZ, Jensen TK, Lindecrona RH, Boye M, Moller K. (2002). Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. *Appl Environ Microbiol* **68**: 673–690.

Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. (2005). Obesity alters gut microbial ecology. *Proc Natl Acad Sci USA* **102**: 11070–11075.

Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS *et al.* (2008a). Evolution of mammals and their gut microbes. *Science* **320**: 1647–1651.

Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. (2008b). Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol* **6**: 776–788.

Ley RE, Peterson DA, Gordon JI. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**: 837–848.

Li M, Wang B, Zhang M, Rantalainen M, Wang S, Zhou H *et al.* (2008). Symbiotic gut microbes modulate human metabolic phenotypes. *Proc Natl Acad Sci USA* **105**: 2117–2122.

Lin JH-C, Savage DC. (1984). Host specificity of the colonization of murine gastric epithelium by lactobacilli. *FEMS Microbiol Letters* **24**: 67–71.

Madsen KL, Doyle JS, Jewell LD, Tavernini MM, Fedorak RN. (1999). *Lactobacillus* species prevents colitis in interleukin 10 gene-deficient mice. *Gastroenterology* **116**: 1107–1114.

Molin G, Andersson R, Ahrne S, Lonner C, Marklinder I, Johansson ML *et al.* (1992). Effect of fermented oatmeal soup on the cholesterol level and the *Lactobacillus* colonization of rat intestinal mucosa. *Antonie Van Leeuwenhoek* **61**: 167–173.

Moller PL, Paerregaard A, Gad M, Kristensen NN, Claesson MH. (2005). Colitic scid mice fed *Lactobacillus* spp. show an ameliorated gut histopathology and an altered cytokine profile by local T cells. *Inflamm Bowel Dis* **11**: 814–819.

Moran NA. (2006). Symbiosis. *Curr Biol* **16**: R866–R871.

Moran NA, McCutcheon JP, Nakabachi A. (2008). Genomics and evolution of heritable bacterial symbionts. *Annu Rev Genet* **42**: 165–190.

Nicholson JK, Holmes E, Wilson ID. (2005). Gut microorganisms, mammalian metabolism and personalized health care. *Nat Rev Microbiol* **3**: 431–438.

Ochman H, Elwyn S, Moran NA. (1999). Calibrating bacterial evolution. *Proc Natl Acad Sci USA* **96**: 12638–12643.

Pena JA, Rogers AB, Ge Z, Ng V, Li SY, Fox JG *et al.* (2005). Probiotic *Lactobacillus* spp. diminish *Helicobacter hepaticus*-induced inflammatory bowel disease in interleukin-10-deficient mice. *Infect Immun* **73**: 912–920.

Peterson DA, McNulty NP, Guruge JL, Gordon JI. (2007). IgA response to symbiotic bacteria as a mediator of gut homeostasis. *Cell Host Microbe* **2**: 328–339.

Rakoff-Nahoum S, Paglino J, Eslami-Varzaneh F, Edberg S, Medzhitov R. (2004). Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* **118**: 229–241.

Rawls JF, Samuel BS, Gordon JI. (2004). Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. *Proc Natl Acad Sci USA* **101**: 4596–4601.

Reuter G. (2001). The *Lactobacillus* and *Bifidobacterium* microflora of the human intestine: composition and succession. *Curr Issues Intest Microbiol* **2**: 43–53.

Rosenfeldt V, Benfeldt E, Nielsen SD, Michaelsen KF, Jeppesen DL, Valerius NH *et al.* (2003). Effect of probiotic *Lactobacillus* strains in children with atopic dermatitis. *J Allergy Clin Immunol* **111**: 389–395.

Round JL, Mazmanian SK. (2009). The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* **9**: 313–323.

Salzman NH, de Jong H, Paterson Y, Harmsen HJ, Welling GW, Bos NA. (2002). Analysis of 16S libraries of mouse gastrointestinal microflora reveals a large new group of mouse intestinal bacteria. *Microbiology* **148**: 3651–3660.

Savage DC. (1972). Associations and physiological interactions of indigenous microorganisms and gastrointestinal epithelia. *Am J Clin Nutr* **25**: 1372–1379.

Savino F, Pelle E, Palumeri E, Oggero R, Miniero R. (2007). *Lactobacillus reuteri* (American Type Culture Collection Strain 55730) versus simethicone in the treatment of infantile colic: a prospective randomized study. *Pediatrics* **119**: e124–e130.

Schell MA, Karmirantzou M, Snel B, Vilanova D, Berger B, Pessi G *et al.* (2002). The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci USA* **99**: 14422–14427.

Schreiber O, Petersson J, Phillipson M, Perry M, Roos S, Holm L. (2009). *Lactobacillus reuteri* prevents colitis by reducing P-selectin-associated leukocyte- and platelet-endothelial cell interactions. *Am J Physiol Gastrointest Liver Physiol* **296**: G534–G542.

Shornikova AV, Casas IA, Isolauri E, Mykkanen H, Vesikari T. (1997). *Lactobacillus reuteri* as a therapeutic agent in acute diarrhea in young children. *J Pediatr Gastroenterol Nutr* **24**: 399–404.

Spratt BG, Hanage WP, Feil EJ. (2001). The relative contributions of recombination and point mutation to the diversification of bacterial clones. *Curr Opin Microbiol* **4**: 602–606.

Suerbaum S, Josenhans C. (2007). *Helicobacter pylori* evolution and phenotypic diversification in a changing host. *Nat Rev Microbiol* **5**: 441–452.

Tannock GW. (1992). Lactic microbiota of pigs, mice and rats. In: Wood BJB (ed). *The Lactic Acid Bacteria in Health and Disease*. Elsevier Applied Science: London, pp 21–48.

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

Tannock GW, Miller JR, Savage DC. (1984). Host specificity of filamentous, segmented microorganisms adherent to the small bowel epithelium in mice and rats. *Appl Environ Microbiol* **47**: 441–442.

Walter J. (2008). Ecological role of lactobacilli in the gastrointestinal tract: implications for fundamental and biomedical research. *Appl Environ Microbiol* **74**: 4985–4996.

Walter J, Loach DM, Alqumber M, Rockel C, Hermann C, Pfitzenmaier M *et al.* (2007). D-alanyl ester depletion of teichoic acids in *Lactobacillus reuteri* 100-23 results in impaired colonization of the mouse gastrointestinal tract. *Environ Microbiol* **9**: 1750–1760.

Weizman Z, Asli G, Alsheikh A. (2005). Effect of a probiotic infant formula on infections in child care centers: comparison of two probiotic agents. *Pediatrics* **115**: 5–9.

Wilson DJ, Gabriel E, Leatherbarrow AJ, Cheesbrough J, Gee S, Bolton E *et al.* (2009). Rapid evolution and the importance of recombination to the gastroenteric pathogen *Campylobacter jejuni*. *Mol Biol Evol* **26**: 385–397.

Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, Chiang HC *et al.* (2003). A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science* **299**: 2074–2076.

Xu J, Mahowald MA, Ley RE, Lozupone CA, Hamady M, Martens EC *et al.* (2007). Evolution of symbiotic bacteria in the distal human intestine. *PLoS Biol* **5**: e156.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)