Microbiology (2004), **150**, 1735–1740 DOI 10.1099/mic.0.26733-0

Diversity analysis of commensal porcine *Escherichia coli* – associations between genotypes and habitat in the porcine gastrointestinal tract

Sameer M. Dixit,[1,2] David M. Gordon,[3] Xi-Yang Wu,[1] Toni Chapman,[1] Kaila Kailasapathy[2] and James J.-C. Chin[1]†

[1]Department of Immunology, Elizabeth Macarthur Agriculture Institute, Woodbridge Rd, Menangle, NSW 2568, Sydney, Australia

[2]Center for Advanced Food Research, University of Western Sydney, Hawkesbury Campus, Penrith South DC, NSW 1797, Australia

[3]School of Botany and Zoology, Australia National University, Canberra, ACT 0200, Australia

Correspondence
James J.-C. Chin
james.chin@agric.nsw.gov.au

Received 27 August 2003
Revised 11 February 2004
Accepted 5 March 2004

Diversity studies of enteric *Escherichia coli* have relied almost entirely on faecal isolations on the assumption that they are representative of flora found throughout the gastrointestinal tract. The authors have addressed this belief by analysing isolates obtained from the duodenum, ileum, colon and faeces of pigs. *E. coli* isolates were obtained from eight pigs and characterized using multi-locus enzyme electrophoresis and PCR-based screening for a range of factors thought to be associated with intestinal and extra-intestinal disease. There are four main genetic groups of commensal *E. coli* (A, B1, B2, D). Group A strains represented 76 % of the isolates from the duodenum, ileum and colon compared to 58 % of the strains isolated from faeces. A nested molecular analysis of variance based on the allozyme and virulence factor screening results showed that differences among individual pigs accounted for 6 % of the observed genetic diversity, whilst 27 % of the genetic variation could be explained by clonal composition differences among gut regions. Finally, the absence of virulence genes in these commensals indicates that they may be suitable as a probiotic consortium, particularly if they also display increased adherence to enterocytes and antagonistic activity against pathogenic strains of *E. coli*.

INTRODUCTION

Escherichia coli is the dominant aerobe in the gastrointestinal tract (GIT) of humans and other mammals (Hartl & Dykhuizen, 1984; Selander *et al.*, 1987; Gordon & Cowling, 2003). *E. coli* is thought to inhabit the lower GIT; however, few studies have examined the distribution of *E. coli* among the different regions of the intestine. Still fewer studies have been undertaken with the aim of determining if *E. coli* inhabiting different regions of the GIT are genetically distinct. Bettelheim *et al.* (1992) serotyped *E. coli* from different regions of the guts of infants that had died of sudden infant death syndrome and concluded that the strains found in the faeces represented those found in other regions of the gut. The majority of diversity studies involving commensal *E. coli* have relied on faecal isolations (Gordon *et al.*, 2002; Pupo *et al.*, 2000; Duriez *et al.*, 2001). However, for some bacterial species such as *Streptococcus mitis*, faecal isolations may not accurately represent the population colonizing the gut mucosa (Hohwy *et al.*, 2001).

Most strains of *E. coli* are harmless commensal of mammals (Hartl & Dykhuizen, 1984; Selander *et al.*, 1987); however, some strains are capable of causing either intestinal or extra-intestinal disease (Ørskov & Ørskov, 1992). Strains responsible for intestinal diseases are thought to cause much of their pathology in the small intestine (Grauke *et al.*, 2002). This pathology is due, in part, to the ability of the strains to adhere to gut epithelial cells (Law, 2000). Some commensal *E. coli* isolates can also adhere to gut epithelial cells (Chin & Mullbacher, 2003; Chin, 2003).

Probiotic treatment is a disease-management strategy based on the use of bacterial strains incapable of causing disease in their host, but which can prevent the establishment

†Present address: NSW Agriculture, PMB 8, Camden, NSW 2570, Sydney, Australia.

Abbreviations: AMOVA, molecular analysis of variance; ECOR, *E. coli* Reference collection; GIT, gastrointestinal tract; MLEE, multi-locus enzyme electrophoresis.

MLEE and virulence factor screening results for all *E. coli* isolates examined are available as supplementary data with the online version of this paper at http://mic.sgmjournals.org.

of pathogenic bacteria (Ouwehand *et al.*, 1999; Kailasapathy & Chin, 2000). Pathogens may be prevented from establishing a population via a number of mechanisms. The probiotic strain may pre-empt attachment sites in the gut, produce antimicrobial compounds that kill the invading pathogen, or outcompete the pathogen for nutrients (Todoriki *et al.*, 2001; Gismondo *et al.*, 1999). Strains of the same species are more likely to have a similar ecological niche than are strains of different species. Most intestinal disease in pigs is caused by pathogenic strains of *E. coli* (Osek, 1999). Thus *E. coli* is a likely candidate for use as a probiotic in pigs. However, selecting a suitable strain or suite of strains for use in a probiotic formulation requires finding strains that are incapable of causing disease but which are able to establish populations in the small intestine of pigs.

The results of mathematical studies modelling the intestine as a plug-flow reactor have suggested that the ability of *E. coli* to establish populations in the large and small intestine should be different (Ballyk & Smith, 1999). This difference results from the higher velocities that occur in the small intestine as compared to the colon (Chivers & Langer, 1994). The models suggest that *E. coli* is unlikely to establish a population in the small intestine unless it is capable of wall growth. If incapable of wall growth, *E. coli* will be more likely to establish a population in the colon. Thus, the theoretical studies suggest that *E. coli* inhabiting different gut regions may well be genetically distinct. This is particularly important in the selection of commensal strains as candidate probiotic bacteria.

The goal of this study was to determine if strains of *E. coli* inhabiting different regions of the gut represent genetically distinct populations. *E. coli* isolates were obtained from the major gut regions, duodenum, ileum, colon and faeces, of pigs. The isolates were characterized using the technique of multi-locus enzyme electrophoresis (MLEE), screened for the presence/absence of a number of genes implicated as virulence factors and assigned to one of the four main groups of *E. coli* found in the *E. coli* Reference collection (ECOR) (Ochman & Selander, 1984). The virulence gene assays were also conducted with a view to investigating the safety factor of commensal isolates if they were to be used as probiotic bacteria.

METHODS

Animals, bacterial isolation and identification. Eight 13-week-old male pigs (hybrids of Large White and Landrace) from eight different litters, weighing 35–45 kg at the start of the experiment, were used in this study. The pigs were housed in separate pens and provided with water and commercial pelleted diet (Vella Stock Feeds, NSW) *ad libitum*. After 3 weeks the animals were killed and the GITs removed. A section of gut was removed from each of three regions: duodenum, the proximal 4 cm region immediately after the stomach; ileum, the distal 4 cm region immediately before the ileo-caecal junction; colon, a 4 cm region immediately past the caecum. A sample of faecal material was removed from the distal region of the colon. The contents of each region, as well as the faecal sample, were weighed and diluted (using peptone-buffered water) to obtain a final concentration of 50 mg sample ml^{-1}. The samples were then serially diluted and aliquots from each dilution were plated onto a MacConkey agar plate. Following incubation, five colonies with the appropriate colour and morphology for *E. coli* were selected at random from the plate with an appropriate number of colonies. Thus, on average, four or five isolates were taken per gut region and faecal sample for each individual pig. For two of the animals no *E. coli* were detected in the duodenum samples. All isolates were tested for indole production and growth on Simmons citrate agar plates and minimal lactose agar plates. Isolates with a Lac^+ Cit^- Ind^+ phenotype were presumed to be *E. coli*. This procedure resulted in the recovery of 158 isolates of *E. coli*, of which 151 were subsequently used for further analysis.

Virulence factor screening. A total of 18 virulence factor genes were assessed using modifications of the multiplex PCR protocols described by Paton & Paton (1998) and Johnson & Stell (2000). Each reaction was carried out using a 25 μl mixture containing 2·5 μl 10× buffer, $MgCl_2$ at a final concentration of 1·5 mM, 20 pmol of each primer, each deoxynucleoside triphosphate (Astral Scientific) at a concentration of 400 μM, 2·5 U HotStarTaq polymerase (Qiagen) and 2 μl boiled bacterial cell lysate. The PCR was performed with a PC960 Air Cooled Thermal Cycler (Corbett research). The 18 virulence factor genes were: *PAI*, *iha*, *papA*, *fimH*, *sfa/focDE*, *kpsMT II*, *ibeA*, *hlyA*, *gafD*, *fyuA*, *ompT*, *cdtB*, *traT*, *kpsMT K5* (group 1); *stx1*, *stx2*, *eae*A and *ehx*A (group 2). The PCR procedure for group 1 genes involved the following conditions: 1 cycle of 15 min at 95 °C followed by 30 s at 94 °C; 30 s at 63 °C and 3 min at 68 °C for 25 cycles; and a final extension step of 10 min at 72 °C. The steps involved in group 2 PCR conditions were as follows: denaturing for 3 min at 95 °C followed by 10 cycles of 1 min at 95 °C, 2 min at 72 °C; 1 cycle of 1 min at 95 °C, 2 min at 72 °C; 1 cycle of 1 min at 95 °C, 2 min at 63 °C, 1·5 min at 72 °C; 1 cycle of 1 min at 95 °C, 2 min at 63 °C, 1·5 min at 72 °C; 1 cycle of 1 min at 95 °C, 2 min at 62 °C, 1·5 min at 72 °C; 10 cycles of 1 min at 95 °C, 2 min at 61 °C, 1·5 min at 72 °C; 10 cycles of 1 min at 95 °C, 2 min at 60 °C, 2·5 min at 72 °C.

ECOR assignment. The method of Clermont *et al.* (2000) was used to assign the *E. coli* isolates to one of the four main groups of *E. coli* identified in the ECOR collection (Ochman & Selander, 1984; Herzer *et al.*, 1990). The groups are designated A, B1, B2 and D. The Clermont method is based on a multiplex PCR protocol that determines the presence or absence of two genes (*chuA* and *yjaA*) and an anonymous DNA fragment (TSPE4.C2). The presence/absence of the three PCR products is used in the manner of a dichotomous key to assign an unknown isolate to one of the ECOR groups. A fraction of ECOR group A strains are negative for the three DNA products (Clermont *et al.*, 2000). Strains failing to produce any PCR products were repeated twice more using newly prepared template DNA. Any strain consistently negative for all three PCR products was identified using the BBL Crystal Enteric/Nonfermenter Identification System (Becton Dickinson) according to the manufacturer's protocol. Any strain identified as *E. coli* and negative for the three DNA products expected from the Clermont method was considered to represent an ECOR group A strain (Gordon & Cowling, 2003).

Allozyme analysis. All *E. coli* isolates were purity streaked onto Luria Broth agar plates from MacConkey plates. A single colony was inoculated into 10 ml Luria Broth (LB) culture and grown overnight with agitation at 37 °C. Cells were harvested by centrifugation, resuspended in 5 ml wash buffer (100 mM Tris, 1 mM EDTA pH 6·8) and harvested by centrifugation. Pellets were resuspended in 600 μl extraction buffer (0·005 mM NADP, 7·2 mM mercaptoethanol) and sonicated with a Branson 200 Sonifier for 10 s using two bursts of 5 s each. Lysed cell culture was centrifuged at 20 000 ***g*** for 5 min at

4 °C. Supernatants were stored at −80 °C until further use. Titan-III cellulose acetate plates (Helena Laboratories) were used for enzyme electrophoresis, which was carried out at a constant voltage (200 V) for between 20 and 40 min. A total of 10 enzyme systems were tested, based on similar work involving *E. coli* carried out in other laboratories (Gordon & Lee, 1999). The following enzyme systems were used: SDH (shikimate dehydrogenase, EC 1.1.1.25), AK (adenylate kinase, EC 2.7.4.3), ADH (alcohol dehydrogenase, EC 1.1.1.1), MR (menadione reductase, EC 1.6.99.2), G6PDH (glucose-6-phosphate dehydrogenase, EC 1.1.1.49), MDH (malate dehydrogenase (EC 1.1.1.37), ME (malic enzyme, EC 1.1.1.40), IDH (isocitrate dehydrogenase, EC 1.1.1.42), PGD (6-phosphogluconate dehydrogenase, EC 1.1.1.44) and PGM (phosphoglucomutase, EC 5.4.2.2). ADH, MDH, PGD, PGI, PGM were run in 0·01 M Tris/glycine (TG) buffer, pH 8·5; all others were run in 0·1 M Tris/EDTA/maleate/$MgCl_2$ (TEM) buffer, pH 7·2.

All strains were run at least twice against strains giving similar band mobility in order to confirm their genotype. When no enzyme activity was detected at a particular locus, a null value was assigned to that locus. Nulls were included in the calculations of the genetic distances between strains and diversity estimates but were discarded when determining the total number of alleles per locus. Occasionally two protein bands were detected at a locus in some strains. These cases were treated as representing a different allele when calculating genetic diversity but not for the calculations of the number of alleles.

Statistical analysis. The genetic diversity of a locus was calculated as

$$h_j = [1 - \sum p_i^2]\ [\frac{n}{n-1}]$$

where p_i is the frequency of allele i at locus j and n is the number of isolates. Genome diversity (H) is estimated as the arithmetic mean of h_j over all loci.

Molecular analysis of variance (AMOVA) (Excoffier *et al.*, 1992) was used to determine how much of the variation among strains was explained by a population-level effect (pig and gut region). AMOVA is built on the principles of classic analysis of variance. However, the data analysed with AMOVA is the matrix of pairwise genetic distances (the number of loci at which two strains differ). The estimated variance components are expressed as the proportion of the total variance explained by the population-level effect (ϕst). ϕst statistics have an interpretation equivalent to the classic single-locus F statistics of Wright (1943), which are used for assessing the extent of population differentiation. Although AMOVA is based on analysis of variance, F statistics cannot be used to assess significance levels. Instead a randomization approach is used, whereby strains are randomly assigned to populations and ϕst recalculated. This procedure is repeated 99 times and the probability of obtaining a ϕst value as extreme as that observed is determined. The AMOVA analysis was done using the software WINAMOVA 1.5. Principal components analysis was used to reduce the dimensions of the data matrix and the first three axes were used as the response variables in a standard nested analysis of variance.

RESULTS

Genetic variation

The 151 isolates when characterized for 10 MLEE loci gave rise to 106 electrophoretic types. There was a mean of 2·6 alleles per locus and the overall level of allelic diversity was $H=0·41$ (Table 1). Ten of the 18 virulence factors screened for were detected (Table 2). The virulence factors *PAI*, *papA*, *sfa/focDE*, *ibeA*, *gafD*, *cdtB*, *stx2* and *ehx*A could not be detected. Inclusion of the virulence factor results with the MLEE data revealed 123 haplotypes among the 151 isolates. A supplementary data file for all strains examined for this study can be found at http://mic.sgmjournals.org. The supplementary file contains the MLEE and virulence factor profiles and ECOR group of all strains, as well as the animal and gut region from which the strain was recovered.

Table 1. Allelic diversity estimates of each locus for *E. coli* isolated from different gut regions of pigs

Locus*	Alleles per locus	Allelic diversity (h_j)
ADH	2	0·29
AK	2	0·30
IDH	3	0·59
MDH	3	0·35
ME	3	0·37
MR-1	2	0·12
M2-2	2	0·49
M2-3	2	0·31
PGD	4	0·77
SDH	3	0·50
Mean	2·6	0·405

*See text for definition of locus abbreviations.

The PCR-based method developed by Clermont *et al.* (2000) assigns *E. coli* isolates to one of the four main clusters of strains (A, B1, B2, D) found in the ECOR collection. Application of this method to the isolates from pigs showed that 29 % of the strains belonged to the B1 cluster and the balance to the A cluster. No B2 or D strains were detected.

Strains designated B1 had a significantly greater median number of virulence factors (2 per strain) than did A strains (1 per strain) (Kruskall–Wallis test: $\chi^2_{(1)}=21·2$, $P<0·001$). This was due to the significantly greater prevalence of *fimH*, *fyuA* and *ompT* in B1 strains (contingency table analysis: *fimH*, $\chi^2_{(1)}=17·5$, $P<0·001$; *fyuA*, $\chi^2_{(1)}=28·9$, $P<0·001$; *ompT*, $\chi^2_{(1)}=14·1$, $P<0·001$).

Variation within and among pigs

Nested AMOVA based on the allozyme and virulence factor screening results revealed that variation among pigs accounted for 6 % (AMOVA, $P<0·001$) of the observed genetic variation in this collection of strains. Variation among gut regions (within pigs) accounted for, on average, 27 % (AMOVA, $P<0·001$) of the variation observed. In two pigs (166 and 181) no statistically significant ($P>0·05$) differentiation among the *E. coli* populations inhabiting the different gut regions could be detected (AMOVA). In the pigs where significant differentiation was observed, the extent of population differentiation among gut regions was: pig 129, 42 %; pig 130, 30 %; pig 167, 46 %; pig 172, 31 %; pig 173, 25 % and pig 180, 24 %. Excluding the

Table 2. Virulence factor profiles and their frequency for the 151 commensal *E. coli* isolates from the pig GIT and their *E. coli* group membership

ECOR	*iha*	*fimH*	*kpsMT II*	*hlyA*	*fyuA*	*ompT*	*traT*	*K5*	*stx1*	*eaeA*	Frequency
A	−	−	−	−	−	−	−	−	−	−	10
A	−	−	−	−	−	−	+	−	−	−	11
A	−	−	−	−	−	−	+	−	−	+	2
A	−	+	−	−	−	−	−	−	−	−	35
A	−	+	−	−	−	−	−	−	−	+	1
A	−	+	−	−	−	−	+	−	−	−	37
A	−	+	−	−	−	−	+	−	−	+	1
A	−	+	−	−	−	+	−	−	−	−	1
A	−	+	−	−	−	+	+	−	+	−	2
A	−	+	−	−	+	−	−	−	−	−	1
A	−	+	−	−	+	−	+	−	−	−	3
A	−	+	+	−	−	−	+	+	−	−	1
A	+	+	−	−	−	−	+	−	−	−	1
A	+	+	−	+	+	−	−	−	−	−	1
B1	−	+	−	−	−	−	−	−	−	−	3
B1	−	+	−	−	−	−	+	−	−	−	13
B1	−	+	−	−	−	+	−	−	−	−	5
B1	−	+	−	−	−	+	+	−	−	−	4
B1	−	+	−	−	−	+	+	−	−	+	1
B1	−	+	−	−	+	−	−	−	−	−	18

virulence factor data from the AMOVA analysis had little effect on the outcome of the analysis (6 % among pigs and 26 % among gut regions nested within pigs). The first three principal components accounted for 12 %, 11 % and 10 %, respectively, of the variation in the MLEE and virulence factor data matrix. For the first and third principal components, significant ($P<0{\cdot}01$) among pigs and among gut region (nested within pigs) variation was detected using standard analysis of variance methods (results not presented). AMOVA was used to calculate the amount of genetic differentiation occurring among gut regions combining the data for all pigs. The resulting population differentiation statistics were used in a UPGMA analysis to illustrate the genetic similarity of strains from the different gut regions (Fig. 1). Overall, *E. coli* isolated from the duodenum and ileum were genetically similar and distinct from strains isolated from the colon and faeces. Strains from the colon and faeces were, on average, more distinct from each other compared to strains from the two regions of the small intestine.

The differences in the relative frequencies of the strains belonging to the ECOR groups A and B1 reflected the results obtained using the allozyme and virulence factor data. There were significant differences in the relative frequencies of A and B1 strains among pigs (contingency table analysis: $\chi^2_{(7)}=21{\cdot}7$, $P<0{\cdot}003$). The relative frequency of B1 and A strains differed among gut regions (contingency table analysis: $\chi^2_{(3)}=7{\cdot}9$, $P<0{\cdot}05$). B1 strains represented 42 % of the isolates from faeces and 24 % of the isolates from the duodenum, ileum and colon were A strains.

DISCUSSION

Bacteria found in the faeces may not be fully representative of their counterparts in different regions of the GIT. This is because the stomach, duodenum, small intestine, caecum and large intestine represent very different environments evolved specifically for different tasks. For example, the stomach (transport regulation of food, storage, sorting, mixing and autoenzymic digestion of food bolus) has a different task than small intestine (absorption of nutrients) and the large intestine (absorption of nutrients, faecal-pellet formation) (Chivers & Langer, 1994). Within these regions, bacterial species can be epithelium adherent or mucus adherent. The intestinal epithelium and the overlying

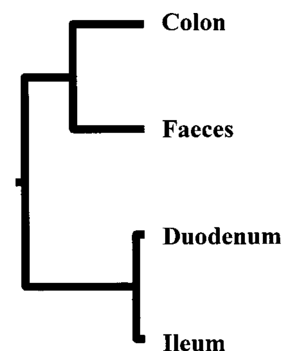

Fig. 1. Genetic similarity of *E. coli* populations from different parts of the pig GIT. Dendrogram created from the pairwise population differentiation estimates derived from an AMOVA.

mucus layer turn over continuously through life, and peristaltic activity leads to continuous exposure of the gut regions to a large number of bacterial species, dietary molecules and secretions from other organs (Falk *et al.*, 1998). It is known that different species of bacteria reside in different regions of the gut, and a species occupies a particular region in order to meet its nutritional and environmental requirements (Ewing & Cole, 1994).

The results of this study endorse the concept that different gut regions represent different ecological niches for *E. coli*. The mean amount of among-gut region differentiation was 27 %. This value is comparable to the degree of population differentiation observed (29 %) among the *E. coli* populations inhabiting different human hosts (Gordon, 2001; Gordon *et al.*, 2002). It far exceeds the values of 2–5 % differentiation observed among populations isolated from different geographical localities (Whittam *et al.*, 1983; Gordon, 1997; Gordon & Lee, 1999; Gordon *et al.*, 2002). It also exceeds the 5 % differentiation observed in *E. coli* populations isolated from different mammalian taxonomic families (Gordon & Lee, 1999).

The differentiation that is observed among the *E. coli* populations inhabiting different gut regions may simply be a consequence of the dynamics of strain colonization and extinction within an individual pig. That is, different regions of the gut are colonized at random by strains and once a region has been occupied other strains are prevented from establishing. However, the observation that ECOR group A strains are more likely to be isolated from the upper GIT and B1 strains from the faeces indicates that the colonization of different regions of the GIT is not a random process. In addition, the results of the AMOVA analysis (Fig. 1) show that the *E. coli* populations in the duodenum and ileum are very similar to another, but divergent from the populations inhabiting the lower GIT. These outcomes suggest that strains inhabiting different regions of the GIT may possess unique suites of traits that lead to niche specialization with respect to the different environments that exist along the length of the GIT. Further work is required to identify these traits and to determine how they lead to habitat specialization.

Are the commensal strains isolated from pigs in this study capable of causing disease? They seem to be unlikely to cause extra-intestinal disease as most strains that do so belong to the B2 and D ECOR groups and carry virulence genes such as *papA* and *hlyA* (Johnson & Stell, 2000). The strains in this study were negative for most virulence genes associated with extra-intestinal infection and were found to be either B1 or A strains. It also appears that they are unlikely to be capable of causing intestinal disease because they lack many virulence gene combinations found in *E. coli* isolated from clinical cases of both neonatal and postweaning diarrhoea (Wu *et al.*, 2003). Although *stx1* and *eaeA* are thought to be virulence factors for the VTEC group of diarrhoeic *E. coli*, these genes were not found to co-occur in a strain and their co-occurrence is thought to be necessary for their virulence to be expressed (Whittam, 1998; Law, 2000). The adherence characteristics of some of the commensal isolates have also been analysed (T. Chapman & J. J.-C. Chin, unpublished results) and a number of these have displayed a high capacity to bind to enterocytes in culture. These properties together with the lack of virulence genes increase the probability that commensal *E. coli* from the upper and lower GIT of healthy pigs may be suitable as a probiotic consortium for the rapid colonization of the intestines of newborn piglets.

ACKNOWLEDGEMENTS

The authors would like to thank Shoko Okada for her help in MLEE work at ANU and also Mr Paul Young for technical assistance.

REFERENCES

Ballyk, M. & Smith, H. (1999). A model of microbiological growth in a plug flow reactor with wall attachment. *Math Biosci* **158**, 95–126.

Bettelheim, K. A., Goldwater, P. N., Evangelidis, H., Pearce, J. L. & Smith, D. L. (1992). Distribution of toxigenic *Escherichia coli* serotypes in the intestines of infants. *Comp Immun Microbiol Infect Dis* **15**, 65–70.

Chin, J. C. (2003). Revisiting the "hygiene hypothesis" in gastrointestinal allergy. *Curr Opin Gastroenterol* **18**, 705–710.

Chin, J. C. & Mullbacher, A. (2003). Immune activation versus hyporesponsiveness and tolerance in the gut. In *Gut Flora, Nutrition, Immunity and Health*, pp. 178–195. Edited by R. Fuller & G. Perdigon. Oxford: Blackwell.

Chivers, D. J. & Langer, P. (1994). *The Digestive Systems in Mammals: Food, Form and Function.* Cambridge, UK: Cambridge University Press.

Clermont, O., Bonacorsi, S. & Bingen, E. (2000). Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* **66**, 4555–4558.

Duriez, P., Clermont, O., Bonacorsi, S., Bingen, E., Chaventre, A., Elion, J., Picard, B. & Denamur, E. (2001). Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology* **147**, 1671–1676.

Ewing, W. N. & Cole, D. J. A. (1994). The gastrointestinal tract. In *The Living Gut – an Introduction to Micro-Organisms in Nutrition*, pp. 10–28. Edited by W. Ewing & D. Cole. Nottingham, UK: Context Publication.

Excoffier, L., Smouse, P. E. & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491.

Falk, P. G., Hooper, L. V., Midtvedt, T. & Gordon, J. I. (1998). Creating and maintaining the gastrointestinal ecosystem: what we know and need to know from gnotobiology. *Microbiol Mol Biol Rev* **62**, 1157–1170.

Gismondo, M. R., Drago, L. & Lombardi, A. (1999). Review of probiotics available to modify gastrointestinal flora. *Int J Antimicrob Agents* **12**, 287–292.

Gordon, D. M. (1997). The genetic structure of *Escherichia coli* populations in feral house mice. *Microbiology* **143**, 2039–2046.

Gordon, D. M. (2001). Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. *Microbiology* **147**, 1079–1085.

Gordon, D. M. & Cowling, A. (2003). The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology* **149**, 3575–3586.

Gordon, D. M. & Lee, J. (1999). The genetic structure of enteric bacteria from Australian mammals. *Microbiology* **145**, 2673–2682.

Gordon, D. M., Bauer, S. & Johnson, J. R. (2002). The genetic structure of *Escherichia coli* populations in primary and secondary habitats. *Microbiology* **148**, 1513–1522.

Grauke, L. J., Kudva, I. T., Yoon, J. W., Hunt, C. W., Williams, C. J. & Hovde, C. J. (2002). Gastrointestinal tract location of *Escherichia coli* O157 : H7 in ruminants. *Appl Environ Microbiol* **68**, 2269–2277.

Hartl, D. L. & Dykhuizen, D. E. (1984). The population genetics of *Escherichia coli*. *Annu Rev Genet* **18**, 31–68.

Herzer, P. J., Inouye, S., Inouye, M. & Whittam, T. S. (1990). Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *J Bacteriol* **172**, 6175–6181.

Hohwy, J., Reinholdt, J. & Mogens, K. (2001). Population dynamics of *Streptococcus mitis* in its natural habitat. *Infect Immun* **69**, 6055–6063.

Johnson, J. R. & Stell, A. L. (2000). Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *J Infect Dis* **181**, 261–272.

Kailasapathy, K. & Chin, J. C. (2000). Survival and therapeutic potential of probiotics organisms with reference to *Lactobacillus acidophilus* and *Bifidobacterium* spp. *Immunol Cell Biol* **78**, 80–88.

Law, D. (2000). Virulence factors of *Escherichia coli* O157 and other shiga toxin-producing *E. coli*. *J Appl Microbiol* **88**, 729–745.

Ochman, H. & Selander, R. K. (1984). Standard reference strains of *Escherichia coli* from natural populations. *J Bacteriol* **157**, 690–692.

Ørskov, F. & Ørskov, I. (1992). *Escherichia coli* serotyping and disease in man and animals. *Can J Microbiol* **38**, 699–704.

Osek, J. (1999). Prevalence of virulence factors of *Escherichia coli* strains isolated from diarrheic and healthy piglets after weaning. *Vet Microbiol* **68**, 209–217.

Ouwehand, A. C., Kirjavainen, P. V., Gronlund, M. M., Isolauri, E. & Salminen, S. J. (1999). Adhesion of probiotic micro-organisms to intestinal mucus. *Int Dairy J* **9**, 623–630.

Paton, A. W. & Paton, J. C. (1998). Detection and characterization of shiga toxigenic *Escherichia coli* by using multiplex PCR assays for stx_1, stx_2, *eaeA*, enterohemorrhagic *E. coli hlyA*, rfb_{O111} and rfb_{O157}. *J Clin Microbiol* **36**, 598–602.

Pupo, G. M., Lan, R., Reeves, R. & Baverstock, P. R. (2000). Population genetics of *Escherichia coli* in a natural population of native Australian rats. *Environ Microbiol* **2**, 594–610.

Selander, R. K., Caugant, D. A. & Whittam, T. S. (1987). Genetic structure and variation in natural populations *of Escherichia coli*. In *Escherichia coli and Salmonella Typhimurium*: *Cellular and Molecular Biology*, pp. 1625–1648. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.

Todoriki, K., Mukai, T., Sato, S. & Toba, T. (2001). Inhibition of adhesion of food-borne pathogens to Caco-2 cells by *Lactobacillus* strains. *J Appl Microbiol* **91**, 154–159.

Whittam, T. S. (1998). Evolution of *Escherichia coli* O157 : H7 and other Shiga toxin-producing *E. coli* strains. In *Escherichia coli O157 : H7 and Other Shiga-Toxin-Producing E. coli Strains*, pp. 195–209. Edited by J. B. Kaper & A. D. O'Brien. Washington, DC: American Society for Microbiology.

Whittam, T. S., Ochman, H. & Selander, R. K. (1983). Geographical components of linkage disequilibrium in natural populations of *Escherichia coli*. *Mol Biol Evol* **1**, 67–83.

Wright, S. (1943). Isolation by distance. *Genetics* **28**, 114–138.

Wu, X.-Y., Chapman, T., Gordon, D., Thuy, D. N., Driesen, S., Walker, M. & Chin, J. (2003). Molecular virulence gene typing of clinical *E. coli* isolates from pigs with post weaning diarrhoea. In *Manipulating Pig Production, Proceedings of the Ninth Biennual Conference of the Australasian Pig Science Association*, p. 59. Edited by J. E. Paterson. Werribee, Victoria: Australasian Pig Science Association (APSA).