AEM Accepts, published online ahead of print on 10 July 2009
Appl. Environ. Microbiol. doi:10.1128/AEM.00443-09
Copyright © 2009, American Society for Microbiology and/or the Listed Authors/Institutions. All Rights Reserved.

AEM Accepts published online ahead of print

1 **Absence of *Escherichia coli* Phylogenetic Group B2 Strains**
2 **in Humans and Domesticated Animals from Jeonam Province, Korea**
3
4 Tatsuya Unno[1], Dukki Han[1], Jeonghwan Jang[1], Sun-Nim Lee[1], Gwangpyo Go[2], Ha
5 Young Choi[3], Joon Ha Kim[1], Michael J. Sadowsky[4], Hor-Gil Hur[1,*]
6
7 [1]Department of Environmental Science and Engineering, Gwangju Institute of Science
8 and Technology, Gwangju, Korea, [2]Department of Environmental Public Health, Seoul
9 National University, Yeoungeon-Dong, Jongro-Gu, Seoul, Korea, [3]Department of
10 Diagnostic Laboratory Medicine, Cheomdan General Hospital, Gwangsan-gu, Gwangju,
11 Korea, and [4]Department of Soil, Water and Climate and BioTechnology Institute,
12 University of Minnesota, St. Paul, Minnesota, USA.
13
14 Running title: Phylogenetic group B2 *E. coli* absence in South Korea
15
16 *Corresponding author: Department of Environmental Science and Engineering,
17 Gwangju Institute of Science and Technology, Oryong-dong, Buk-gu, Gwangju 500-712,
18 Republic of Korea.
19 Tel.: +82-62-970-2437; fax: +82-62-970-2434,
20 E-mail address: hghur@gist.ac.kr

AEM Accepts published online ahead of print

21 **Abstract**

22 Multiplex PCR analyses of DNAs from genotypically unique *Escherichia coli*
23 strains isolated from the feces of 138 humans and 376 domesticated animals from
24 Jeonam Province, Korea, done using primers specific for the *chuA* and *yjaA* genes, and
25 an unknown DNA fragment, TSPE4.C2, indicated that none of the strains belonged to *E.*
26 *coli* phylogenetic group B2. In contrast, phylogenetic group B2 strains were detected in
27 about 17% (8 of 48 isolates) from 24 wild geese feces and in 3% (3 of 96) isolates
28 obtained from the Yeongsan River in Jeonam Province, Korea. The distribution of *E. coli*
29 strain in phylogenetic groups A, B1, and D varied depending on the host examined and
30 there was no apparent seasonal variation in the distribution of strains in phylogenetic
31 group among the Yeongsan River isolates. The distribution of four virulence genes
32 (*eaeA, hlyA,* stx_1, and stx_2) in isolates was also examined by using multiplex PCR.
33 Virulence genes were detected in about 5% (38 of 707) of the total unique strains
34 examined, with 24, 13, 13, and 9 strains containing *hlyA*, *eaeA,* stx_2, and stx_1, respectively.
35 The virulence genes were most frequently present in the phylogenetic group B1 strains
36 isolated from beef cattle. Taken together, results of these studies indicate that *E. coli*
37 strains in phylogenic group B2 were rarely found in humans and domesticated animals
38 in Jeonam Province, Korea and that the majority of strains containing virulence genes
39 belonged to phylogenic group B1 and were isolated from beef cattle. Results of this
40 study also suggest that the relationship between the presence and types of virulence genes
41 and phylogenetic groupings may differ among geographically-distinct *E. coli* populations.
42

43 *Key words: Escherichia coli*, phylogenetic group, virulence genes, multiplex PCR,
44 Korea

AEM Accepts published online ahead of print

45 **Introduction**

46 *Escherichia coli* is a normal inhabitant of the lower intestinal tract of
47 warm-blooded animals and humans. While the majority of *E. coli* strains are
48 commensals, some are known to be pathogenic, causing intestinal and extra-intestinal
49 diseases, such as diarrhea and urinary tract infections (42). Phylogenetic studies done
50 using multilocus enzyme electrophoresis (MLEE) and 72 *E. coli* strains in the reference
51 ECOR collection showed that *E. coli* strains can be divided into four phylogenetic
52 groups (A, B1, B2, and D) (20, 41, 48). Recently, a potential fifth group (E) has also
53 been proposed (11). Since multiplex PCR was developed for analysis of phylogenetic
54 groups (6), a number of studies have analyzed a variety of *E. coli* strains for their
55 phylogenetic group association (10, 12, 17, 18, 23, 54). Duriez et al. (10) reported the
56 possible influence of geographic conditions, dietary factors, use of antibiotics, and/or
57 host genetic factors on the distribution of phylogenetic groups among 168 commensal *E.*
58 *coli* strains isolated from human stools from three geographically distinct populations in
59 France, Croatia, and Mali. Random amplified polymorphic DNA (RAPD) analysis of
60 the intraspecies distribution of *E. coli* in pregnant women and neonates indicated that
61 there was a correlation between the distribution of phylogenetic groups, RAPD groups,
62 and virulence factors (54). Moreover, based on comparisons of the distribution of *E. coli*
63 phylogenetic groups among humans of different sex and ages, it has been suggested that
64 *E. coli* genotypes are likely influenced by morphological, physiological, and dietary
65 differences (18). In addition, climate has also been proposed to influence the distribution
66 of strains within *E. coli* phylogenetic groups (12). There are now several reports
67 indicating that there is a potential relationship between *E. coli* phylogenetic groups, age,
68 and disease. For example, *E. coli* isolates belonging to phylogenetic group B2 have been

AEM Accepts published online ahead of print

69 shown to predominate in infants with neonatal bacterial meningitis (27), and among
70 urinary tract and rectal isolates (55). Also, Nowrouzian et al. (39) and Moreno et al. (37)
71 reported that strains belonging to phylogenetic group B2 persisted among the intestinal
72 microflora of infants and were more likely to cause clinical symptoms.

73 Boyd and Hartl (2) reported that among the *E. coli* strains in the ECOR and the
74 diarrheagenic *E. coli* (DEC) collections, strains in phylogenetic group B2 carry the
75 greatest number of virulence factors, followed by those in group D. Virulence factors
76 carried by group B2 strains are thought to contribute to their strong colonizing capacity,
77 a greater number of virulence genes have been detected in resident strains than in
78 transient ones (38). Moreover, a mouse model of extraintestinal virulence showed that
79 phylogenetic group B2 strains killed mice at greater frequency and possessed more
80 virulence determinants than strains in other phylogenetic groups, suggesting a link
81 between phylogeny and virulence genes in *E. coli* extraintestinal infection (45). In
82 contrast, Johnson and Kuskowski (25) suggested that a group B2 ancestral strain might
83 have simply acquired virulence genes by chance, and that these genes were vertically
84 inherited by group members during clonal expansion. However, numerous studies
85 published to date suggest that there is a relationship between the genomic background
86 of phylogenetic group B2 and its association with virulence factors (12, 28, 34, 39, 45).

87 Both enteropathogenic (EPEC) and enterohemorrhagic (EHEC) *E. coli* are
88 among the most important food-borne pathogens worldwide, often causing severe
89 gastrointestinal disease and fatal infections (13). While EPEC strains cause diarrhea,
90 and generally do not produce enterotoxin, they possess adherence factor which is
91 controlled by the chromosomal gene, *eaeA*, encoding for intimin (8). Unlike the EPEC,
92 however, the EHEC typically contain the *hlyA*, stx_1, and stx_2 virulence genes, encoding

AEM Accepts published online ahead of print

93 for hemolysins, and Shiga-like type 1 and 2 toxins, respectively, and *eaeA*. The ability to
94 detect EHEC has been great facilitated by the use of multiplex PCR (13, 44, 53).
95 Several studies have shown that strains producing Shiga-like toxin 2 are more frequently
96 found in cases of hemolytic-uremic syndrome (HUS) than are those containing
97 Shiga-like toxin 1 (30, 43, 46, 49).
98 In the study reported here, we examined the distribution of phylogenetic groups
99 and the prevalence of virulence genes in 659 genotypically-unique *E. coli* strains
100 isolated from humans and domestic animals in Korea. In addition, we also tested 48 and
101 96 non-unique *E. coli* isolates from wild geese and the Yeongsan River, respectively for
102 phylogenetic distribution and virulence gene profiles. Here we report that contrary to
103 what has been previously reported in other parts of the world, no *E. coli* strains
104 belonging to phylogenetic group B2 were found in domesticated animals and in humans
105 from Jeonam Province, Korea. We also report that among the strains we examined,
106 virulence genes were mainly found in phylogenetic group B1 strains isolated from beef
107 cattle. Results of these studies may prove to be useful for the development of risk
108 management strategies to maintain public health.
109
110 **Materials and Methods**
111
112 **Isolation of *E. coli* from humans and domesticated animals, Jeonam Province,**
113 **Korea**
114 The sources of *E. coli* isolates, the number of isolates obtained from each
115 source, and the number of individual hosts sampled are listed in Table 1. The human
116 isolates were obtained from randomly-selected stool samples collected from healthy

AEM Accepts published online ahead of print

117 humans, and patient isolates were obtained in August 2008 from diarrheic patients at a
118 hospital located in Jeonam Province, Korea. The data obtained from studies done with *E.*
119 *coli* isolates from healthy humans and patients with diarrhea were analyzed separately.
120 The *E. coli* strains from domesticated animals were obtained in May 2006 by using
121 fecal swabings from chickens, ducks, swine, and beef and dairy cattle collected at farms
122 in Jeonam Province, Korea. According to the Korea Food and Drug Administration
123 (KFDA), antibiotics, such as tetracycline and penicillin, are regularly fed to
124 domesticated animals as feed additives (31). Wild geese isolates were obtained from
125 fecal swabs collected in December 2007 in Jeonam Province, Korea, where migrating
126 birds from Siberia rest every winter. Fecal swabs were stored in tubes on ice and
127 streaked within 6 hours of collection onto mFC agar (Difco, Detroit, MI) plates and
128 incubated at 44.5°C for 18 hours. Subsequently, three to five blue colonies appearing on
129 mFC agar plates, per fecal sample, were further streaked for purification onto mFC agar
130 plates and incubated overnight at 44.5°C. All isolates were verified to be *E. coli* as
131 previously described (9), and preserved at -70°C in LB freezing buffer (47).
132
133 **Isolation of *E. coli* from Yeongsan River, Jeonam Province, Korea**
134 One site on the Yeongsan River, in Jeonam Province, Korea, was selected for
135 these studies. The site is a part of a tributary upstream from the Yeongsan River and
136 surrounded by an urbanized area. Environmental *E. coli* strains were obtained using the
137 membrane filtration technique according to U. S. Environmental Protection Agency
138 (USEPA) method 1603 (52). Briefly, 500 ml of surface water was sampled every 3
139 months from November 2007 to August 2008, and 100 ml, 10 ml, and 1 ml aliquots of
140 surface water were individually filtered onto the surface of 0.45 µm pore-size

AEM Accepts published online ahead of print

141 membranes (Advantec, Tokyo, Japan). Filtrates were incubated on modified mTEC agar
142 (Difco, Detroit, MI) plates at 35°C for the first 2 hours, and then at 44.5°C for 16 hours.
143 Red- or magenta-colored colonies were considered to be *E. coli*, and 24 randomly
144 selected *E. coli* isolates were further streaked and incubated under the same condition
145 and used for subsequent species verification as described above.
146
147 **Horizontal fluorophore-enhanced rep-PCR DNA fingerprinting**
148 Horizontal fluorophore-enhanced rep-PCR (HFERP) DNA fingerprinting of the
149 *E. coli* strains was done as described previously (29). Briefly, a loopful of bacteria from
150 each strain was suspended in 0.05 N NaOH for 15 min at 95°C, and 1 μl was used as
151 template for PCR. The HFERP DNA fingerprinting was done using BOXA1R primers
152 labeled with 6-FAM (6-carboxyfluorescein; Genotech Co. Ltd., Korea) as previously
153 described (Johnson et al. 2004). All gel lanes contained Genescan-2500 ROX
154 (6-carboxy-X-rhodamine) (Applied Biosystems, Foster City, Calif.) as an internal size
155 standard. Gel images were captured using a Typhoon 9400 variable mode imager
156 (Molecular Dynamics/Amersham Biosciences, Sunnyvale, CA) using the fluorescence
157 acquisition mode, with the following settings: green excitation laser, 610 BP 30 and 526
158 SP emission filters in the autolink mode, normal sensitivity, 200-μm/pixel scan
159 resolution, +3-mm focal plane, and 800-V power. Scanned images of HFERP DNA
160 fingerprints were processed using Image Quant (Molecular Dynamics/Amersham
161 Biosciences, Sunnyvale, CA) and converted to 256 gray-scale tagged image file format
162 images. Gel images were normalized and analyzed using Bionumerics v.5.0 software
163 (Applied Maths, Sint-Martens-Latem, Belgium). Isolates which showed ≥ 92%
164 similarity from the same host were considered to be clones and removed from further

AEM Accepts published online ahead of print

165 analyses (22, 29). The percentages of known-phylogenetic group strains assigned to
166 their correct phylogenetic group were calculated by using Jackknife analysis with
167 maximum similarities.
168

169 **Phylogenetic grouping and virulence gene identification**

170 Only unique strains defined by the HFERP DNA fingerprint analysis were
171 subjected to analyses of phylogenetic groups and virulence gene identification.
172 Phylogenetic grouping was done as previously described by Clermont et al. (6). The
173 presence of the *ibeA* gene (invasion of brain epithelium) among Clermont phylogenetic
174 group D strains having a *chuA* $^{+}$, *yjaA* $^{-}$, and TSPE4.C2 $^{+}$ genotype, was examined as
175 previously described (14).

176 The presence of virulence genes in *E. coli* strains was determined by using
177 multiplex PCR as previously described (44). Genomic DNA from strains was extracted
178 from cells as described above, diluted 10-fold in TE buffer, and 1 µl was used as
179 template for multiplex PCR using a Labcycler (SensoQuest, City, Germany) instrument.
180

181 **Results and Discussion**

182

183 **Phylogenetic grouping patterns**

184 A total of 1,585 *E. coli* isolates obtained from humans and domesticated
185 animals were examined for their genetic relatedness by using HFERP DNA
186 fingerprinting as described by Johnson et al. (29). Strains sharing the same individual
187 host and having a genetic similarity ≥ 92% in HFERP banding patterns were considered
188 to be clones (29) and were removed from further analyses. Based on this definition, 659

AEM Accepts published online ahead of print

189 strains were considered to be unique and were subjected to further phylogenetic
190 grouping and virulence gene analyses. The *E. coli* isolates from migrating wild geese and
191 the Yeongsan River, however, were not subjected to HFERP analysis to remove clones.

192 The distribution of phylogenetic groups among genomically-unique *E. coli*
193 isolates obtained from humans and animals are summarized in Figure 1. The *E. coli*
194 strains from each host source showed a different distribution pattern of phylogenetic
195 groups. The *E. coli* strains from healthy humans were nearly equally represented in each
196 phylogenetic group, with 29, 34, and 36% of the strains in phylogenetic groups A, B1, and
197 D, respectively. There was a slightly greater number of isolates in phylogenetic group D
198 (42.9%) from human patients compared to the other phylogenetic groups, A (23.8%) and
199 B1 (33.3%). The majority of *E. coli* isolates from chickens were localized to phylogenetic
200 group A (55%), followed by strains in groups B1 (31.7%) and D (13.3%). A similar pattern
201 of distribution was also found among isolates from domesticated ducks, where about 63,
202 24, and 13% of strains were in phylogenetic groups A, B1, and D, respectively. In contrast,
203 *E. coli* isolates from beef cattle had the greatest percentage of group B1 strains (79.2%)
204 among all sources, and fewer isolates belonging to groups A (15.1%) and D (5.7%). A
205 similar trend was observed among *E. coli* isolates from dairy cattle, where 62% of the
206 isolates belonged to group B1, and a fewer percentage to groups A (32.0%) and D (5.7%).
207 Swine isolates showed unique phylogenetic group distribution, with an extremely low
208 percentage of group D (0.7%) strains, a relatively high percentage of group A (64.7%)
209 strains, and a moderate percentage of group B1 (34.5%) strains. The phylogenetic group
210 distribution of isolates from migrating wild geese isolates was the most distinctive, the
211 majority of isolates (60.4%) were in phylogenetic group B1, and 16.7, 14.6 and 8.3% of
212 the remaining isolates were in phylogenetic groups B2, A and D, respectively. It should

AEM Accepts published online ahead of print

213 be noted that the phylogenetic group distribution pattern seen among *E. coli* isolates from
214 migrating wild geese was significantly different from that seen among isolates from
215 domesticated poultry chicken and duck, although the chicken and duck isolates showed
216 similar phylogenetic distribution patterns. Taken together, results of these studies indicate
217 that *E. coli* isolates belonging to phylogenetic group A were more frequently found in
218 chickens, ducks and swine, whereas those in phylogenetic group B1 were
219 predominantly found in isolates obtained from beef and dairy cattle. Results in Figure 1
220 also show that there was a different distribution pattern of *E. coli* phylogenetic group D
221 strains from humans and animals. While the majority of strains from healthy humans
222 (36%) and patients (42.9%) belonged to phylogenetic group D, strains in this
223 phylogenetic group generally comprised a small number of isolates obtained from all the
224 animals, including wild geese.

225 Among the 659 genomically-unique strains examined, 15 isolates (11, 2, 1, and 1
226 from healthy humans, ducks, chickens, and human patients, respectively) were found to
227 be members of Clermont phylogenetic group D and had a *chu*A^{+}, *yja*A^{-}, and TSPE4.C2^{+}
228 genotype. Recently, Gordon et al. (15) reported that strains having this genotype and
229 contain the *ibeA* gene are likely members of phylogenetic group B2. PCR analyses done
230 here indicated that 5 of the 15 isolates (3, 1, and 1 from healthy humans, a chicken, and a
231 human patient, respectively) contained the *ibeA gene.* While this result suggested that
232 these strains may possibly belong to phylogenetic group B2 as redefined , by Gordon et al.
233 (16), we propose to assign these 5 strains to phylogenetic group D until the method
234 proposed by Gordon et al. is evaluated by others using a larger number of
235 geographically-diverse isolates and becomes a more established method for the
236 assignment of *E. coli* strains to phylogenetic groups.

AEM Accepts published online ahead of print

237 In the study reported here, no *E. coli* Clermont phylogenetic group B2 strains,
238 using the classical, excepted, definition, were found among the isolates we obtained
239 from humans or domesticated animals in Korea (Figure 1). This is not likely a
240 methodological issue as the multiplex (triplex) PCR method used in this study was
241 previously shown to correctly assign 95% of strains to phylogenetic groups B1 and B2,
242 when compared to MLST (16). Moreover, since phylogenetic group B2 strains were
243 identified among the river water isolates and migrating geese isolates examined, this
244 indicates that the methods used was sufficiently robust to detect strains in the group.
245 Previously, Escobar-Páramo et al. (12) reported that the prevalence of phylogenetic
246 group B2 isolates among individuals in temperate regions of mainland France,
247 Michigan, and Tokyo was greater than that found among people in tropical populations
248 from Bògotá, Cotonou, and French Guyana. In contrast, our data shows that the
249 phylogenetic group distribution for human isolates from Jeonam Province, Korea was
250 nearly equally divided among phylogenetic groups A (30%), B1 (34%), and D (36.2%).
251 Interestingly, it was also previously reported that *E. coli* strains from humans in Tokyo
252 were predominantly in phylogenetic group B2, and no B1 strains were present (40). The
253 different phylogenetic group distribution among *E. coli* strains from Japan and Korea may
254 be due to differences in dietary habits. Moreover, distributional differences among
255 phylogenetic groups of human *E. coli* isolates are not static and were shown to change
256 in response to geographic shifts in populations, which typically result in subsequent
257 alterations to diet (49). For example, shifts in *E. coli* phylogenetic group were found
258 among 25 humans who expatriated from metropolitan France to French Guyana (50).
259 This data suggests that there is a strong environmental influence on the phylogenetic
260 group distribution of intestinal *E. coli* isolates in humans.

AEM Accepts published online ahead of print

261 The phylogenetic distribution of human *E. coli* isolates may also be impacted by
262 the use of antibiotics. It was previously reported that *E. coli* strains belonging to
263 phylogenetic group B2 were less likely to be resistant to antibiotics than non-B2 group
264 strains (24, 51). Skurnik et al. (51) reported that only 3.7% of group B2 strains carried
265 integrons, whereas greater than 16% of strains from other phylogenetic groups did. As
266 compared to other industrialized countries, the use of antibiotics in Korea is quite
267 extensive, with a defined daily dose (DDD) rate of 33.2 /1000 inhabitants/day. In
268 contrast, the DDD rate in OECD countries averaged 21.3 /1000 inhabitants/day (35).
269 Moreover, *E. coli* strains resistant to multiple antimicrobial substances are frequently
270 observed in Korea (4, 5, 36). Taken together, these factors may contribute to the absence
271 of phylogenetic group B2 strains among the Korean human populations we examined.
272 In addition to animals and humans, the phylogenetic distribution of
273 environmental *E. coli* isolates from the Yeongsan River, Jeonam province, Korea, was
274 also examined during the four seasons of an entire year. The Yeongsan River water
275 contained, on average, greater than 200 colony forming units (CFU) of *E. coli* per ml in
276 all seasons (data not shown). Results in Figure 2 show the seasonal variation in the
277 phylogenetic group distribution of *E. coli* strains in the Yeongsan River. Similar
278 distribution patterns were seen in November 2007 and May 2008 samples. A high
279 percentage of group B1 strains was found in both the November 2007 and May 2008
280 samples (45.8% and 54.2%, respectively), while a smaller percentage of strains were
281 shown to comprise phylogenetic groups A (25.0% and 33.3%, respectively) and D
282 (25.0% and 12.5%, respectively). In contrast, the February 2008 and August 2008
283 samples contained a high percentage of group A strains (87.5% and 83.3%, respectively).
284 In contrast to what was found with *E. coli* isolates from humans and domesticated

AEM Accepts published online ahead of print

285 animals, *E. coli strains* in phylogenetic group B2 were detected in the November 2007
286 (4.2%) and August 2008 (8.3%) water samples. However, these strains were only
287 infrequently isolated. Results in Figure 1 also show that a greater percentage of *E. coli*
288 strains obtained from chickens, ducks, and swine were in phylogenetic group A,
289 whereas a high percentage of strains in group B1 were observed among *E. coli* obtained
290 from beef and dairy cattle and wild geese.
291
292 **Virulence gene distribution**

293 The occurrence and distributional pattern of virulence genes among the
294 phylogenetic groups of unique *E. coli* isolates obtained from the various human and
295 animal hosts is shown in Table 2. Of the 659 unique strains and the 48 wild geese and
296 96 fresh water isolates examined, only 38 strains (4.7%) from healthy humans, human
297 patients, chickens, beef cattle, dairy cattle, and swine were found to contain virulence
298 genes. Approximately 20% of the beef cattle isolates in phylogenetic group B1 (17 of 84
299 strains) were found to carry virulence genes, and 16.7 and 12.5% of the strains were in
300 phylogenetic groups D and A, respectively. The distribution of virulence genes in dairy
301 cattle had a different pattern than those from beef cattle. While 23.5% of dairy cattle
302 strain containing virulence genes were in phylogenetic group A (4 out of 17 strains),
303 15.2% of the strains were in group B1 (5 out of 33 strains). None of the dairy cattle
304 strains in phylogenetic group D contained virulence genes. Taken together, our results
305 indicated that the percentage of *E. coli* strains carrying virulence genes was unequally
306 distributed among sources, and depended both on host source and the prevalence of
307 strains in each phylogenetic group. For example, phylogenetic group B1 strains from

AEM Accepts published online ahead of print

308 each host source generally had a great percentage of strains carrying virulence genes, and
309 those in group D had a lesser number.

310 The greatest number of strains carrying virulence genes was found in
311 phylogenetic group B1 strains obtained from beef cattle (15.7%), followed by group B1
312 strains from dairy cattle (9.4%). Ishii et al. (23) and Girardeau et al. (15) reported that
313 Shiga-like toxin producing *E. coli* (STEC) strains segregated mainly into phylogenetic
314 group B1. This is similar to the results we report here for isolates obtained from Jeonam
315 Province, Korea. No strains from ducks or wild geese were found to contain virulence
316 genes.

317 The distributional pattern of virulence genes tested in this study is shown in
318 Figure 3. The *eaeA* (an attaching and effacing (A/E) protein, intimin, responsible for
319 pathogenicity) was detected less in phylogenetic group A strains than those in the other
320 groups (A: 0.34%, B1:3.68%, and D 2.08%), which is in agreement with results from a
321 previous report (15). The intimin protein has been shown to be important for
322 enterohemorrhagic infection of *E. coli* (1, 3, 19, 32, 33). The *eaeA* has also been used to
323 detect a chromosomally-localized pathogenicity island, referred to as the locus of
324 enterocyte effacement (LEE), and strains containing *eaeA* and lacking stx_1 and stx_2 are
325 referred to as enteropathogenic *E. coli* (EPEC) (21). In our studies, potential EPEC
326 strains were detected in 2.2, 4.7, and 2.8% of isolates from healthy humans, human
327 patients, and beef cattle, respectively, while potential EHEC strains ($eaeA^+$ stx^+) were
328 detected in 4.7 and 1.9% of strains from beef and dairy cattle, respectively (Table 3). By
329 far, the greatest percentage of strains containing *eaeA*, *hlyA*, stx_1 and stx_2 belonged to
330 phylogenetic group B1 (Figure 3). Genes encoding for stx_1 and stx_2 were found in 2.5%

AEM Accepts published online ahead of print

331 (18 out of 707) of the strains examined, and the greatest number of *E. coli* strains
332 carrying virulence genes were seen in the beef (18.9%; 20 out of 104 strains) and dairy
333 cattle (17.0%; 9 out of 53) isolates (Table 3). A similar percentages of STEC strains
334 were reported to be present among *E. coli* isolates obtained from cattle fecal material in
335 Germany (56) and Australia (7).

336 **Population structure of *E. coli* strains obtained from human and domesticated**
337 **animal hosts**

338 The genetic relatedness of the unique *E. coli* strains containing virulence genes are
339 shown in Figure 4. Generally speaking, the strains could be divided into two major
340 groups (A and B) at the 45% similarity level. The group A strains could be further
341 subdivided into two subclusters, I and II. Subcluster II strains were further subdivided
342 into three subgroups (II.1, II.2, and II.3). The subgroup II.2 and II.3 strains were
343 separated at the 63% similarity level and comprised 47% (18 out of 38) of the analyzed
344 strains. Regardless of host and virulence profiles, 47% of the strains (18 of 38) were
345 related to each other at the ≥ 80% similarity level. Moreover, 22% (6 out of 27) of the
346 phylogenetic group B1 strains were clustered at a ≥ 88% similarity level. The majority
347 (71.1%) of strains carrying virulence genes belonged to phylogenetic group B1, and
348 71% (15 out of 21) of the cluster A, subgroup II strains were from beef cattle.

349 The patterns of virulence gene profiles were not uniformly distributed among
350 the strains examined by HFERP analysis. For example, while strains aa18, ak70, and
351 aa84 shared the same virulence gene profile (*hlyA*, stx_1, and stx_2 positive) they were
352 only distantly genetically related, at less than the 70% similarity level. It also should be
353 noted that one phylogenetic group D strain carried all four virulence genes tested and

AEM Accepts published online ahead of print

354 was not genetically related to any of the strains carrying similar virulence genes.

355 Multivariate analysis of variance (MANOVA) was used to determine if the
356 HFERP DNA fingerprint patterns of strains could be used to differentiate phylogenetic
357 groups. The percentage of strains correctly classified into each group was determined by
358 using Jackknife analysis (Table 4). Results in Figure 5 and Table 4 show that cluster
359 analysis separated the strains into three groups which did not correlated well with
360 phylogenetic groupings. Approximately 70 to 75% of group A and B1 strains were
361 correctly assigned to their respective phylogenetic groups, whereas about 20% of these
362 strains were misclassified. The phylogenetic group D strains showed the lowest
363 percentage of correct assignment (57.3%).
364

365 **Conclusions**

366 Six hundred and fifty-nine genomically-unique *E. coli* isolates obtained from
367 domesticated animals and humans were subjected to phylogenetic grouping analysis
368 using multiplex PCR. Of the strains examined, 291, 272, and 96 isolates were assigned
369 to phylogenetic groups A, B1, and D, respectively. No group B2 strains were found
370 among *E. coli* isolated from feces of any of domesticated animals and humans from
371 Jeonam Province, Korea. However, strains in phylogenetic group B2 were found in the
372 isolates obtained from wild geese and Yeongsan River water. The clustering of strains by
373 HFERP DNA fingerprint analysis did not correlate well with phylogenetic group
374 designations made based on PCR analyses, and the method misclassified about 20% of
375 group A and B1 strains, and about 40% of group D strains. While it was also previously
376 reported that BOX-PCR DNA fingerprinting may not be useful for differentiating

AEM Accepts published online ahead of print

377 strains within *E. coli* phylogenetic groups (26), the method has proven to be useful for
378 strain-level discrimination, to cluster genetically similar *E. coli* ecotypes, and to
379 differentiate sources and virotypes of *E. coli* (9, 23, 29).
380 The distribution of *E. coli* strains in the three phylogenetic groups varied depending on
381 the animal host from where the strains were obtained; beef and dairy cattle isolates
382 showed a relatively similar distributional pattern of phylogenetic groups, as did the duck
383 and chicken isolates. Our data also support previous suggestions that diet and antibiotic
384 usage may strongly influence the phylogenetic group distribution of *E. coli* strains (17,
385 18, 24, 51). Moreover, results from these studies indicate that the distribution of *E. coli*
386 strains in phylogenetic groups may be strongly influenced by geographical boundaries.
387 Therefore, further physiological and epidemiological studies are needed to clarify the
388 reason why phylogenetic group B2 strains are rare in South Korea.
389 More virulence genes were found in the Korean phylogenetic group B1 strains
390 we examined than in strains from the other phylogenetic groups. This suggests that these
391 strains may either share a common ancestor, or are subjected to intensive horizontal gene
392 transfer and recombination events. The relatively frequent occurrence of *eaeA* positive
393 strains among beef cattle isolates suggests that further surveillance studies are required
394 in order to properly assess risk associated with *E. coli* from different animal sources in
395 Korea.

AEM Accepts published online ahead of print

396 **Acknowledgements**

397 This study was supported, in part, from a grant from the Korean Ministry of
398 Environment as "The Eco-technopia 21 project" and by a grant from the Korean
399 Research Foundation funded by the Korean Government (MOEHRD) (Project No.
400 420035) to Tatsuya Unno. We thank Dr. Timothy Johnson for providing DNA of *E. coli*
401 strain carrying *ibeA*.

AEM Accepts published online ahead of print

402 **References**

403 1. **Boerlin, P., S. Chen, J. K. Colbourne, R. Johnson, S. De Grandis, and C.**
404 **Gyles.** 1998. Evolution of enterohemorrhagic *Escherichia coli* hemolysin
405 plasmids and the locus for enterocyte effacement in shiga toxin-producing *E.*
406 *coli*. Infect. Immun. **66:**2553-61.

407 2. **Boyd, E. F., and D. L. Hartl.** 1998. Chromosomal regions specific to
408 pathogenic isolates of *Escherichia coli* have a phylogenetically clustered
409 distribution. J. Bacteriol. **180:**1159-65.

410 3. **China, B., E. Jacquemin, A. C. Devrin, V. Pirson, and J. Mainil.** 1999.
411 Heterogeneity of the *eae* genes in attaching/effacing *Escherichia coli* from
412 cattle: comparison with human strains. Res. Microbiol. **150:**323-32.

413 4. **Choi, C., H. J. Ham, D. Kwon, J. Kim, D. S. Cheon, K. Min, W. S. Cho, H. K.**
414 **Chung, T. Jung, K. Jung, and C. Chae.** 2002. Antimicrobial susceptibility of
415 pathogenic *Escherichia coli* isolated from pigs in Korea. J. Vet. Med. Sci.
416 **64:**71-3.

417 5. **Chong, Y., and K. Lee.** 2000. Present situation of antimicrobial resistance in
418 Korea. J. Infect. Chemother. **6:**189-95.

419 6. **Clermont, O., S. Bonacorsi, and E. Bingen.** 2000. Rapid and simple
420 determination of the *Escherichia coli* phylogenetic group. Appl. Environ.
421 Microbiol. **66:**4555-8.

422 7. **Cobbold, R., and P. Desmarchelier.** 2000. A longitudinal study of
423 Shiga-toxigenic *Escherichia coli* (STEC) prevalence in three Australian diary
424 herds. Vet. Microbiol. **71:**125-37.

425 8. **Cocolin, L., M. Manzano, C. Cantoni, and G. Comi.** 2000. A multiplex-PCR

AEM Accepts published online ahead of print

426 method to detect enterohemorrhagic (EHEC) and enteropathogenic (EPEC)
427 Escherichia coli in artificially contaminated foods. Int J Hyg Environ Health
428 **203:**159-64.

429 9. **Dombek, P. E., L. K. Johnson, S. T. Zimmerley, and M. J. Sadowsky.** 2000.
430 Use of repetitive DNA sequences and the PCR To differentiate *Escherichia coli*
431 isolates from human and animal sources. Appl. Environ. Microbiol. **66:**2572-7.

432 10. **Duriez, P., O. Clermont, S. Bonacorsi, E. Bingen, A. Chaventre, J. Elion, B.**
433 **Picard, and E. Denamur.** 2001. Commensal *Escherichia coli* isolates are
434 phylogenetically distributed among geographically distinct human populations.
435 Microbiology **147:**1671-6.

436 11. **Escobar-Paramo, P., O. Clermont, A. B. Blanc-Potard, H. Bui, C. Le**
437 **Bouguenec, and E. Denamur.** 2004. A specific genetic background is required
438 for acquisition and expression of virulence factors in *Escherichia coli*. Mol. Biol.
439 Evol. **21:**1085-94.

440 12. **Escobar-Paramo, P., K. Grenet, A. Le Menac'h, L. Rode, E. Salgado, C.**
441 **Amorin, S. Gouriou, B. Picard, M. C. Rahimy, A. Andremont, E. Denamur,**
442 **and R. Ruimy.** 2004. Large-scale population structure of human commensal
443 *Escherichia coli* isolates. Appl. Environ. Microbiol. **70:**5698-700.

444 13. **Fagan, P. K., M. A. Hornitzky, K. A. Bettelheim, and S. P. Djordjevic.** 1999.
445 Detection of shiga-like toxin (*stx1* and *stx2*), intimin (*eaeA*), and
446 enterohemorrhagic Escherichia coli (EHEC) hemolysin (EHEC *hlyA*) genes in
447 animal feces by multiplex PCR. Appl. Environ. Microbiol. **65:**868-72.

448 14. **Germon, P., Y. H. Chen, L. He, J. E. Blanco, A. Bree, C. Schouler, S. H.**
449 **Huang, and M. Moulin-Schouleur.** 2005. ibeA, a virulence factor of avian

AEM Accepts published online ahead of print

450 pathogenic Escherichia coli. Microbiology **151:**1179-86.

451 15. **Girardeau, J. P., A. Dalmasso, Y. Bertin, C. Ducrot, S. Bord, V. Livrelli, C.**
452 **Vernozy-Rozand, and C. Martin.** 2005. Association of virulence genotype with
453 phylogenetic background in comparison to different seropathotypes of Shiga
454 toxin-producing *Escherichia coli* isolates. J. Clin. Microbiol. **43:**6098-107.

455 16. **Gordon, D. M., O. Clermont, H. Tolley, and E. Denamur.** 2008. Assigning
456 *Escherichia coli* strains to phylogenetic groups: multi-locus sequence typing
457 versus the PCR triplex method. Environ. Microbiol. **10:**2484-96.

458 17. **Gordon, D. M., and A. Cowling.** 2003. The distribution and genetic structure of
459 *Escherichia coli* in Australian vertebrates: host and geographic effects.
460 Microbiology **149:**3575-86.

461 18. **Gordon, D. M., S. E. Stern, and P. J. Collignon.** 2005. Influence of the age and
462 sex of human hosts on the distribution of *Escherichia coli* ECOR groups and
463 virulence traits. Microbiology **151:**15-23.

464 19. **Gyles, C., R. Johnson, A. Gao, K. Ziebell, D. Pierard, S. Aleksic, and P.**
465 **Boerlin.** 1998. Association of enterohemorrhagic *Escherichia coli* hemolysin
466 with serotypes of shiga-like-toxin-producing *Escherichia coli* of human and
467 bovine origins. Appl. Environ. Microbiol. **64:**4134-41.

468 20. **Herzer, P. J., S. Inouye, M. Inouye, and T. S. Whittam.** 1990. Phylogenetic
469 distribution of branched RNA-linked multicopy single-stranded DNA among
470 natural isolates of *Escherichia coli*. J. Bacteriol. **172:**6175-81.

471 21. **Hornitzky, M. A., K. Mercieca, K. A. Bettelheim, and S. P. Djordjevic.** 2005.
472 Bovine feces from animals with gastrointestinal infections are a source of
473 serologically diverse atypical enteropathogenic *Escherichia coli* and Shiga

AEM Accepts published online ahead of print

474 toxin-producing *E. coli* strains that commonly possess intimin. Appl. Environ.
475 Microbiol. **71:**3405-12.
476 22. **Ishii, S., W. B. Ksoll, R. E. Hicks, and M. J. Sadowsky.** 2006. Presence and
477 growth of naturalized *Escherichia coli* in temperate soils from Lake Superior
478 watersheds. Appl. Environ. Microbiol. **72:**612-21.
479 23. **Ishii, S., K. P. Meyer, and M. J. Sadowsky.** 2007. Relationship between
480 phylogenetic groups, genotypic clusters, and virulence gene profiles of
481 *Escherichia coli* strains from diverse human and animal sources. Appl. Environ.
482 Microbiol. **73:**5703-10.
483 24. **Johnson, J. R., P. Goullet, B. Picard, S. L. Moseley, P. L. Roberts, and W. E.**
484 **Stamm.** 1991. Association of carboxylesterase B electrophoretic pattern with
485 presence and expression of urovirulence factor determinants and antimicrobial
486 resistance among strains of *Escherichia coli* that cause urosepsis. Infect. Immun.
487 **59:**2311-5.
488 25. **Johnson, J. R., and M. Kuskowski.** 2000. Clonal origin, virulence factors, and
489 virulence. Infect. Immun. **68:**424-5.
490 26. **Johnson, J. R., and T. T. O'Bryan.** 2000. Improved repetitive-element PCR
491 fingerprinting for resolving pathogenic and nonpathogenic phylogenetic groups
492 within Escherichia coli. Clin Diagn Lab Immunol **7:**265-73.
493 27. **Johnson, J. R., E. Oswald, T. T. O'Bryan, M. A. Kuskowski, and L.**
494 **Spanjaard.** 2002. Phylogenetic distribution of virulence-associated genes
495 among *Escherichia coli* isolates associated with neonatal bacterial meningitis in
496 the Netherlands. J. Infect. Dis. **185:**774-84.
497 28. **Johnson, J. R., and A. L. Stell.** 2000. Extended virulence genotypes of

AEM Accepts published online ahead of print

498 *Escherichia* coli strains from patients with urosepsis in relation to phylogeny
499 and host compromise. J. Infect. Dis. **181:**261-72.
500 29. **Johnson, L. K., M. B. Brown, E. A. Carruthers, J. A. Ferguson, P. E.**
501 **Dombek, and M. J. Sadowsky.** 2004. Sample size, library composition, and
502 genotypic diversity among natural populations of *Escherichia coli* from different
503 animals influence accuracy of determining sources of fecal pollution. Appl.
504 Environ. Microbiol. **70:**4478-85.
505 30. **Kleanthous, H., H. R. Smith, S. M. Scotland, R. J. Gross, B. Rowe, C. M.**
506 **Taylor, and D. V. Milford.** 1990. Haemolytic uraemic syndromes in the British
507 Isles, 1985-8: association with verocytotoxin producing *Escherichia coli*. Part 2:
508 Microbiological aspects. Arch. Dis. Child. **65:**722-7.
509 31. **Korea-Food-and-Drug-Administration.** 2003. Report on National
510 Antimicrobial Resistance Control. *In* S. Korea Food and Drug Administration,
511 Korea. (ed.).
512 32. **Kudva, I. T., K. Blanch, and C. J. Hovde.** 1998. Analysis of *Escherichia coli*
513 O157:H7 survival in ovine or bovine manure and manure slurry. Appl. Environ.
514 Microbiol. **64:**3166-74.
515 33. **Kudva, I. T., P. G. Hatfield, and C. J. Hovde.** 1997. Characterization of
516 *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli* serotypes
517 isolated from sheep. J. Clin. Microbiol. **35:**892-9.
518 34. **Le Gall, T., O. Clermont, S. Gouriou, B. Picard, X. Nassif, E. Denamur, and**
519 **O. Tenaillon.** 2007. Extraintestinal virulence is a coincidental by-product of
520 commensalism in B2 phylogenetic group Escherichia coli strains. Mol Biol Evol
521 **24:**2373-84.

AEM Accepts published online ahead of print

522 35. **Lee, E., J. Bae, and K. Park.** 2000. Drug Use Evaluation. Seoul, Korea. Korea
523 Institute for Health and Social Affairs (in Korean).
524 36. **Lim, S. K., H. S. Lee, H. M. Nam, Y. S. Cho, J. M. Kim, S. W. Song, Y. H.**
525 **Park, and S. C. Jung.** 2007. Antimicrobial resistance observed in *Escherichia*
526 *coli* strains isolated from fecal samples of cattle and pigs in Korea during
527 2003-2004. Int. J. Food Microbiol. **116:**283-6.
528 37. **Moreno, E., I. Planells, G. Prats, A. M. Planes, G. Moreno, and A. Andreu.**
529 2005. Comparative study of *Escherichia coli* virulence determinants in strains
530 causing urinary tract bacteremia versus strains causing pyelonephritis and other
531 sources of bacteremia. Diagn. Microbiol. Infect. Dis. **53:**93-9.
532 38. **Nowrouzian, F. L., I. Adlerberth, and A. E. Wold.** 2006. Enhanced persistence
533 in the colonic microbiota of *Escherichia coli* strains belonging to phylogenetic
534 group B2: role of virulence factors and adherence to colonic cells. Microbes
535 Infect. **8:**834-40.
536 39. **Nowrouzian, F. L., A. E. Wold, and I. Adlerberth.** 2005. *Escherichia coli*
537 strains belonging to phylogenetic group B2 have superior capacity to persist in
538 the intestinal microflora of infants. J. Infect. Dis. **191:**1078-83.
539 40. **Obata-Yasuoka, M., W. Ba-Thein, T. Tsukamoto, H. Yoshikawa, and H.**
540 **Hayashi.** 2002. Vaginal *Escherichia coli* share common virulence factor profiles,
541 serotypes and phylogeny with other extraintestinal *E. coli*. Microbiology
542 **148:**2745-52.
543 41. **Ochman, H., and R. K. Selander.** 1984. Standard reference strains of
544 *Escherichia coli* from natural populations. J. Bacteriol. **157:**690-3.
545 42. **Orskov, F., and I. Orskov.** 1992. *Escherichia coli* serotyping and disease in

AEM Accepts published online ahead of print

546 man and animals. Can. J. Microbiol. **38:**699-704.

547 43. **Ostroff, S. M., P. I. Tarr, M. A. Neill, J. H. Lewis, N. Hargrett-Bean, and J.**
548 **M. Kobayashi.** 1989. Toxin genotypes and plasmid profiles as determinants of
549 systemic sequelae in *Escherichia coli* O157:H7 infections. J. Infect. Dis.
550 **160:**994-8.

551 44. **Paton, A. W., and J. C. Paton.** 1998. Detection and characterization of Shiga
552 toxigenic *Escherichia coli* by using multiplex PCR assays for *stx1*, *stx2*, *eaeA*,
553 enterohemorrhagic *E. coli hlyA*, *rfbO111*, and *rfbO157*. J. Clin. Microbiol.
554 **36:**598-602.

555 45. **Picard, B., J. S. Garcia, S. Gouriou, P. Duriez, N. Brahimi, E. Bingen, J.**
556 **Elion, and E. Denamur.** 1999. The link between phylogeny and virulence in
557 Escherichia coli extraintestinal infection. Infect Immun **67:**546-53.

558 46. **Proulx, F., E. G. Seidman, and D. Karpman.** 2001. Pathogenesis of Shiga
559 toxin-associated hemolytic uremic syndrome. Pediatr. Res. **50:**163-71.

560 47. **Sambrook, J., and W. D. Russell.** 2001. Molecular Cloning. A laboratory
561 manual. The third edition. Cold Spring Harbor Laboratory Press, Cold Spring
562 Harbor, New York.

563 48. **Selander, R. K., D. A. Caugant, H. Ochman, J. M. Musser, M. N. Gilmour,**
564 **and T. S. Whittam.** 1986. Methods of multilocus enzyme electrophoresis for
565 bacterial population genetics and systematics. Appl. Environ. Microbiol.
566 **51:**873-84.

567 49. **Siegler, R. L., T. G. Obrig, T. J. Pysher, V. L. Tesh, N. D. Denkers, and F. B.**
568 **Taylor.** 2003. Response to Shiga toxin 1 and 2 in a baboon model of hemolytic
569 uremic syndrome. Pediatr. Nephrol. **18:**92-6.

AEM Accepts published online ahead of print

570 50. **Skurnik, D., D. Bonnet, C. Bernede-Bauduin, R. Michel, C. Guette, J. M.**
571 **Becker, C. Balaire, F. Chau, J. Mohler, V. Jarlier, J. P. Boutin, B. Moreau, D.**
572 **Guillemot, E. Denamur, A. Andremont, and R. Ruimy.** 2008. Characteristics
573 of human intestinal *Escherichia coli* with changing environments. Environ.
574 Microbiol. **10:**2132-7.

575 51. **Skurnik, D., A. Le Menac'h, D. Zurakowski, D. Mazel, P. Courvalin, E.**
576 **Denamur, A. Andremont, and R. Ruimy.** 2005. Integron-associated antibiotic
577 resistance and phylogenetic grouping of *Escherichia coli* isolates from healthy
578 subjects free of recent antibiotic exposure. Antimicrob. Agents Chemother.
579 **49:**3062-5.

580 52. **U.S., E. P. A.** 2002. Method 1603: *Escherichia coli* (*E. coli*) in water by
581 membrane filtration using membrane-thermo-tolerant *Escherichia coli* agar
582 (modified mTEC). EPA-821-R-02-023. Office of Water, Washington DC.

583 53. **Wang, G., C. G. Clark, and F. G. Rodgers.** 2002. Detection in *Escherichia coli*
584 of the genes encoding the major virulence factors, the genes defining the
585 O157:H7 serotype, and components of the type 2 Shiga toxin family by
586 multiplex PCR. J. Clin. Microbiol. **40:**3613-9.

587 54. **Watt, S., P. Lanotte, L. Mereghetti, M. Moulin-Schouleur, B. Picard, and R.**
588 **Quentin.** 2003. *Escherichia coli* strains from pregnant women and neonates:
589 intraspecies genetic distribution and prevalence of virulence factors. J. Clin.
590 Microbiol. **41:**1929-35.

591 55. **Zhang, L., B. Foxman, and C. Marrs.** 2002. Both urinary and rectal
592 *Escherichia coli* isolates are dominated by strains of phylogenetic group B2. J.
593 Clin. Microbiol. **40:**3951-5.

AEM Accepts published online ahead of print

594 56. **Zschock, M., H. P. Hamann, B. Kloppert, and W. Wolter.** 2000.
595 Shiga-toxin-producing *Escherichia coli* in faeces of healthy dairy cows, sheep
596 and goats: prevalence and virulence properties. Lett. Appl. Microbiol. **31:**203-8.
597

AEM Accepts published online ahead of print

598 **Figure Legends**

599

600 Figure 1. Distribution of phylogenetic groups among *E. coli* isolates obtained from
601 humans and domesticated animals: (□) group A; (▨) group B1; (■) group B2 and;
602 (⊠) group D.

603

604 Figure 2. Seasonal variations in phylogenetic group distribution among *E. coli* obtained
605 from the Yeongsan River, Jeonam province, Korea: (□) group A; (▨) group B1; (■)
606 group B2 and; (⊠) group D.

607

608 Figure 3. Distribution of virulence genes among phylogenetic groups of *E. coli* obtained
609 from humans and domesticated animals: (■) *eaeA*; (▨); *hlyA* (□) stx_1; and (⊠) stx_2.

610

611 Figure 4. Genetic relatedness of *E. coli* strains possessing virulence genes. The
612 dendrogram was generated from HFERP DNA fingerprints using Pearson's
613 product-moment correlation coefficient and the unweighted-pair group method with
614 arithmetic means clustering method.

615

616 Figure 5. Phylogenetic grouping analysis of HFERP DNA fingerprints using MANOVA:
617 group A (●);group B1 (●);group D (●). HFERP DNA fingerprints from *E. coli* strains
618 obtained from animal and human sources were numerically converted to binary
619 band-matching character tables and analyzed by MANOVA accounting for the
620 covariance structure.

AEM Accepts published online ahead of print

621 Table 1. Sources and numbers of *E. coli* isolates used in this study

Source	No. of individual hosts sampled	No. of isolates	No. of unique strains
Human	122	442	141
Patient	16	83	21
Chicken	57	154	60
Duck	93	220	139
Beef Cattle	71	266	106
Dairy Cattle	38	194	53
Swine	117	226	139
Wild Geese	24	48	ND[1]
Fresh water	4 times a year	96	ND

622 [1]ND, not determined

AEM Accepts published online ahead of print

623 Table 2. The occurrence of *E. coli* strains with virulence genes and phylogenetic groups

Source	Phylogenetic group	No. of unique strains in phylogenetic groups	No. of unique strains with virulence genes	% unique strains with virulence genes isolated from each source	% unique strains with virulence genes in phylogenetic group
Human	A	42	0	0	0
	B1	48	2	1.4	4.2
	D	51	1	0.7	2.0
Patient	A	5	0	0	0
	B1	7	1	4.8	14.3
	D	9	0	0	0
Chicken	A	33	1	1.7	3.0
	B1	19	0	0	0
	D	8	0	0	0
Duck	A	88	0	0	0
	B1	33	0	0	0
	D	18	0	0	0
Beef cattle	A	16	2	1.9	12.5
	B1	84	17	15.7	20.2
	D	6	1	0.9	16.7
Dairy cattle	A	17	4	7.6	23.5
	B1	33	5	9.4	15.2
	D	3	0	0	0
Swine	A	90	2	1.4	2.2
	B1	48	2	1.4	4.2
	D	1	0	0	0

624

625

AEM Accepts published online ahead of print

626 Table 3. Comparison of the virulence gene patterns in unique *E. coli* strains obtained from different human and domesticated animal
627 sources.

Virulence gene pattern				No. of unique strains with virulence gene pattern found in following sources:								
eaeA	*hlyA*	*stx*$_1$	*stx*$_2$	Humans	Patients	Chickens	Ducks	Beef cattle	Dairy cattle	Swine	Wild Geese	No. strains with virulence gene profiles
+	+	+	+	0	0	0	0	1	0	0	0	1
+	+	+	-	0	0	0	0	2	0	0	0	2
+	+	-	-	0	0	0	0	1	1	0	0	2
+	-	+	-	0	0	0	0	1	0	0	0	1
+	-	-	-	3	1	0	0	3	0	0	0	7
-	+	+	+	0	0	0	0	3	0	0	0	3
-	+	-	-	0	0	1	0	1	8	1	0	11
-	+	-	+	0	0	0	0	5	0	0	0	5
-	-	+	-	0	0	0	0	1	0	1	0	2
-	-	-	+	0	0	0	0	2	0	2	0	4
-	-	-	-	138	20	59	139	84	44	135	48	667
No. unique strains tested				141	21	60	139	104	53	139	48[a]	705
No. strains from each host with virulence genes				3	1	1	0	20	9	4	0	

628 [a]Clonal isolates not removed

629 Table 4. Assignment of unique strains to phylogenetic groups by using HFERP DNA

630 fingerprint and Jackknife analyses

631

Assigned phylogenetic group	Percent *E. coli* isolates in assigned group:		
	A	B1	D
A	73.2	18.4	22.9
B1	19.2	75.7	19.8
D	7.6	5.9	57.3

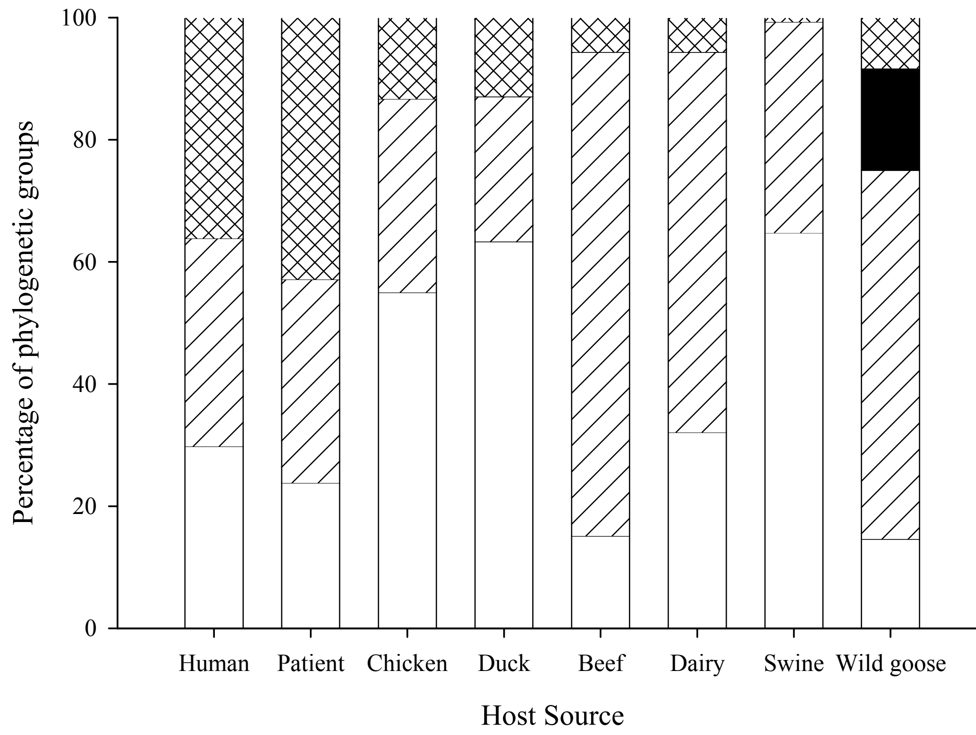
632

633

AEM Accepts published online ahead of print

AEM Accepts published online ahead of print

634 Figure 1



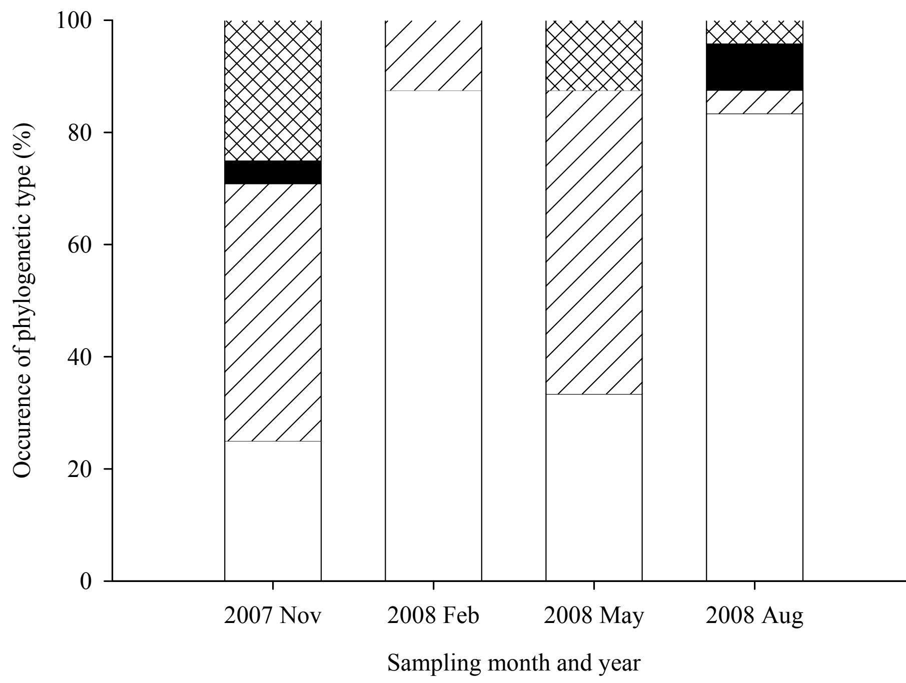
635

636

637

AEM Accepts published online ahead of print

638 Figure 2

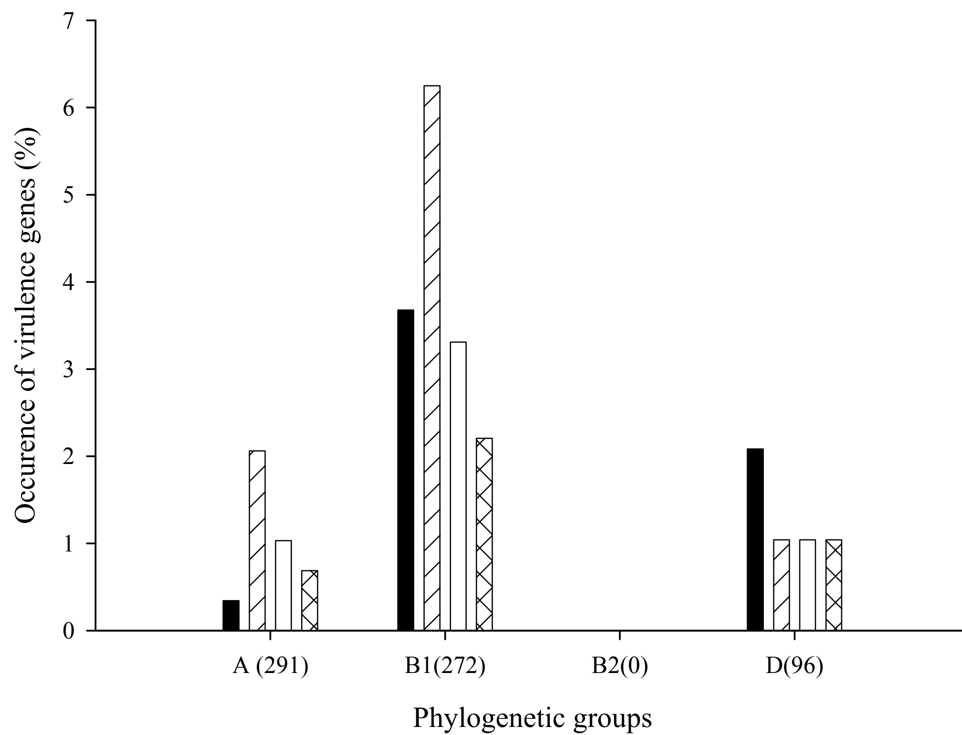


639
640
641
642
643
644
645
646
647
648

AEM Accepts published online ahead of print

649 Figure 3.

650



651

652

653

654

655

656

657

658

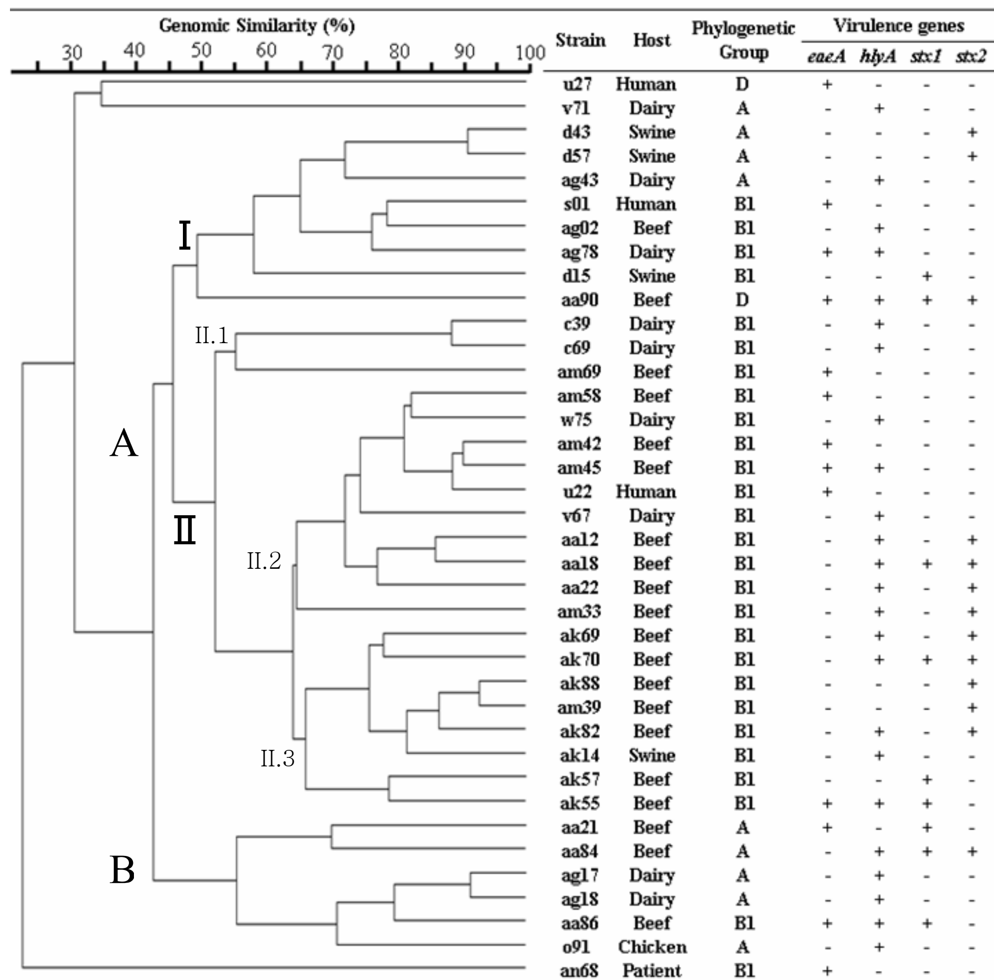
659

660

AEM Accepts published online ahead of print

661 Figure 4

662



663

664

665

666

667

668

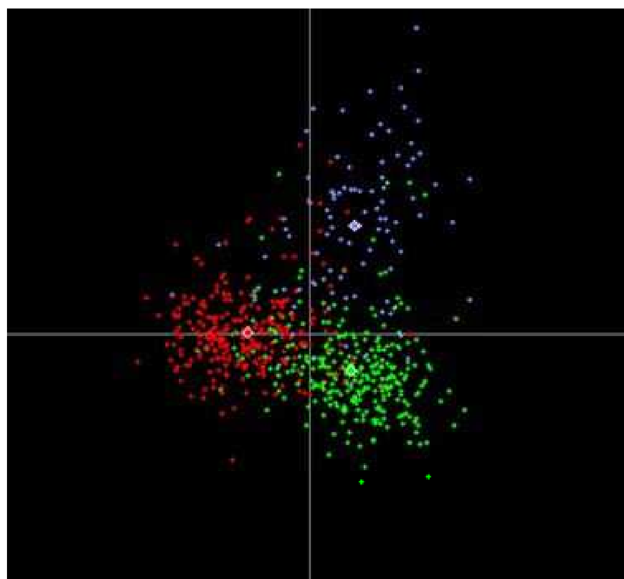
669

AEM Accepts published online ahead of print

670 Figure 5

671

672



673

674

675

676