APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Oct. 2010, p. 6698–6700 Vol. 76, No. 19
0099-2240/10/$12.00 doi:10.1128/AEM.00707-10


# Phylogenetic Groups and Pathogenicity Island Markers in Fecal *Escherichia coli* Isolates from Asymptomatic Humans in China[▽]

Bin Li,[1] Jing-yong Sun,[1] Li-zhong Han,[1] Xin-hong Huang,[2] Qiang Fu,[2] and Yu-xing Ni[1]*

*Department of Clinical Microbiology, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, 197 Ruijin 2nd Road, Shanghai 200025, China,[1] and Departments of Clinical Laboratory, the Affiliated Union Hospital, Fujian Medical University, Fuzhou, Fujian 350001, China[2]*

Received 21 March 2010/Accepted 7 August 2010

**The study of phylogenetic groups and pathogenicity island (PAI) markers in commensal *Escherichia coli* strains from asymptomatic Chinese people showed that group A strains are the most common and that nearly half of all fecal strains which were randomly selected harbor PAIs.**

*Escherichia coli* is a well-diversified commensal species in the intestine of healthy humans but also includes intestinal or extraintestinal pathogens. It has been reported that pathogenic *E. coli* may be derived from fecal strains by acquisition of virulence determinants (11). The relationship between the *E. coli* genetic background and the acquisition of virulence factors is now better understood (1, 5). Extraintestinal *E. coli* strains may harbor several virulence factors, such as adhesins, fimbriae, and hemolysin, which can contribute to bacterial pathogenesis. These traits are usually encoded on pathogenicity islands (PAIs), which have been studied in pathogenic *E. coli* previously (15). The *E. coli* population includes 4 major phylogroups (A, $B_1$, $B_2$, and D) (2). Pathogenic strains belong mainly to groups $B_2$ and D, while most fecal isolates belong to groups A and $B_1$. Strains of groups $B_2$ and D often carry virulence factors that are lacking in group A and $B_1$ strains (3, 9, 13).

In this study, we examined the distribution of phylogroups and the prevalence of PAIs in commensal *E. coli* strains isolated from asymptomatic persons in one region of China.

**Bacterial strains.** The asymptomatic individuals (174 males and 151 females, with an age range from 18 to 75 years) were recruited from those who underwent annual personal physical examination in one hospital in Fuzhou, China, from February to May 2009. All had no confirmed diagnosis of digestive tract diseases. No information on antibiotic history was available. All of the participants gave their informed consent. Samples were cultured on MacConkey agar plates directly. One *E. coli* isolate was collected per person. All isolates were identified by biochemical methods (indole–methyl red–Voges-Proskauer–citrate [IMViC] tests and urease production, $H_2S$ production, and various sugar fermentation tests).

**Phylogroup analysis.** Phylogroups were examined by a PCR-based method (2). The results showed a slightly greater number of strains in group A than in the other three groups (Table 1). Statistical analysis demonstrated that sex and age factors had no effect on distribution of phylogroups (data not shown). Only two studies showed concern about the influence of these two factors on distribution of phylogroups, but they did not reach a consensus (4, 6). Further studies are required to better describe the disparity.

It was reported that geographic and climate factors might affect the distribution of phylogroups of commensal *E. coli* (4). The prevalence of group A isolates among persons in temperate regions was half that among persons in tropical regions in one study (4). In contrast, such strains predominated in this study, though China lies in the temperate belt. However, it would be prudent to analyze the difference between these studies, because no background data of hosts, such as genetic factors, diet, etc., were included in any of this research. As

TABLE 1. Prevalence of the four major *E. coli* phylogenetic groups in fecal samples from different human populations[a]

| Population | No. of samples | Prevalence (%) of phylogenetic group: | | | |
|---|---|---|---|---|---|
| | | A | $B_1$ | $B_2$ | D |
| French people | | | | | |
| Paris area residents I[b] | 56 | 61.0 | 12.5 | 10.5 | 16.0 |
| Paris area residents II[c] | 27 | 29.6 | 11.1 | 37.1 | 22.2 |
| Brittany (BIW)[e] residents | 25 | 24.0 | 24.0 | 32.0 | 20.0 |
| Brittany (PF)[f] residents | 25 | 32.0 | 28.0 | 16.0 | 24.0 |
| Brest University students | 21 | 14.3 | 23.8 | 33.3 | 28.6 |
| Tours residents | 24 | 25.0 | 21.0 | 29.0 | 25.0 |
| Michigan residents | 88 | 20.5 | 12.5 | 47.7 | 19.3 |
| Tokyo, Japan, residents | 61 | 28.0 | 0.0 | 44.0 | 28.0 |
| Bogota, Colombia, residents | 28 | 57.1 | 3.6 | 14.3 | 25.0 |
| Cotonou residents | 46 | 50.0 | 32.6 | 17.4 | 0.0 |
| Guyana Amerindians | 93 | 63.4 | 20.4 | 3.2 | 12.9 |
| Malian people | 55 | 23.6 | 58.2 | 1.8 | 16.4 |
| Croatian people | 57 | 35.1 | 31.6 | 19.3 | 14.0 |
| Australian people | 266 | 19.5 | 12.4 | 45.1 | 22.9 |
| Korean people | 141 | 29.8 | 34.0 | 0.0 | 36.2 |
| Chinese people[d] | 325 | 43.7 | 23.4 | 16.0 | 16.9 |

[a] Data from studies by Duriez et al. (3), Escobar-Paramo et al. (4), Gordon et al. (6), and Unno et al. (18) and from this study.
[b] Data from study by Duriez et al. (3).
[c] Data from study by Escobar-Paramo et al. (4).
[d] Data from this study.
[e] BIW, bank and insurance workers.
[f] PF, pig farmers.

* Corresponding author. Mailing address: Department of Clinical Microbiology, Ruijin Hospital, School of Medicine, Shanghai Jiaotong University, 197 Ruijin 2nd Rd., Shanghai 200025, People's Republic of China. Phone: 86-21-64370045-665541. Fax: 86-21-64333548. E-mail: yuxing_ni@yahoo.com.cn.

▽ Published ahead of print on 13 August 2010.

TABLE 2. Number of pathogenicity islands (PAIs) detected in 220 commensal *Escherichia coli* isolates, classified according to phylogenetic group

| Phylogenetic group | Source[a] | No. (%) of isolates with the following number of PAIs detected: | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| A | E | 73 (73) | 21 (21) | 5 (5) | 1 (1) | | | | |
| | O | 7 (70) | 3 (30) | | | | | | |
| $B_1$ | E | 35 (53) | 20 (31) | 4 (6) | 5 (8) | 1 (2) | | | |
| | O | 14 (82) | 3 (18) | | | | | | |
| $B_2$ | E | 5 (17) | 4 (14) | 11 (38) | 7 (24) | 1 (3) | 1 (3) | | |
| | O | | | 2 (22) | 3 (34) | | 2 (22) | 2 (22) | |
| D | E | 4 (15) | 13 (50) | 9 (35) | | | | | |
| | O | 9 (64) | 3 (22) | 2 (14) | | | | | |

[a] E, our work; O, another work (14).

discussed by Duriez et al. (3), this difference may result from factors related to the external environment or may be due to cultural differences in diet or food processing and preparation practices.

Notably, certain phylogroups were absent in certain populations, such as $B_2$ types in South Korea (18), $B_1$ types in Japan, and D types in Cotonou (Benin, Africa) (Table 1). These data suggested a strong genetic influence on the phylogroup distribution in commensal *E. coli* strains. Curiously, the levels of prevalence of group A strains were quite different in Paris, France, in different studies (Table 1). As discussed by Escobar-Paramo et al. (4), this was probably the result of different dietary habits and hygiene factors.

Factors such as those discussed above may influence the distribution of phylogroups in fecal isolates in humans. It would be of interest to identify strains with genotypes that may possess particular characteristics which could help them to survive in and adapt to various environments. Moreover, since groups A and $B_1$ are the most prevalent among *E. coli* strains in the environment (19), it is possible that human populations that acquire more bacteria from the environment will also have an overabundance of these types of *E. coli*. Further study is needed to confirm these hypotheses.

However, the Clermont method is not perfect in classifying *E. coli* strains into phylogroups; in particular, group A strains (with genotype "−−−" [i.e., lacking *chuA*, *yjaA*, and Tspe4.C2]) are seldom classified correctly (7). In this study, the majority of strains belonged to group A, and 9.5% (31/325) of those were the "−−−" genotype. To validate our results, PCR with *E. coli*-specific *lacZ* primers was performed (8), and all of the strains generated the predicted amplicon. Although the strains in this study were relatively rare, as reported by other researchers (7), they should be characterized further using a multilocus sequence typing method.

**Distribution of PAIs.** PAIs have been investigated widely in pathogenic bacteria, but little attention has been paid to commensal strains. In this study, PAI markers were detected by PCR, as described previously, for 220 *E. coli* strains randomly selected from each group (14). Overall, 46.8% (103/220) of these strains carried PAIs, which was a little higher than that determined in Spain (14), and 161 PAIs were detected in total. PAI $IV_{536}$ was the most common one (38.2%), followed by PAI $I_{CFT073}$ (20.9%), PAI $II_{CFT073}$ (10.9%), PAI $II_{536}$ (1.8%), and PAI $I_{536}$ (1.3%), respectively. Meanwhile, three PAIs ($III_{536}$, $I_{J96}$, and $II_{J96}$) were not detected.

PAI $IV_{536}$ was reported to be the most ubiquitous PAI in *Enterobacteriaceae* (16, 17) and is supposed to be a fitness island rather than a pathogenic one (10, 12). Interestingly, it was previously always detected together with PAI $I_{CFT073}$ (14). However, PAI $IV_{536}$ was frequently detected alone in this study.

Among all of the strains, 20.5% (45/220) had multiple PAIs, which was higher than the results for strains from another nation (Table 2). The percentages of fecal *E. coli* strains with PAIs and with multiple PAIs were higher in group $B_2$ than in the other three groups (Table 3). Because data on the prevalence of PAIs in human intestinal flora are so far quite limited, the difference may be of limited significance from the pathogenicity standpoint. The majority of group $B_2$ strains had more than two PAIs, which revealed that these strains were highly virulent even in the human gut.

In summary, to our knowledge, this is the first report to

TABLE 3. Distribution of pathogenicity islands (PAIs) among 220 commensal isolates of *Escherichia coli*, classified according to phylogenetic group

| Phylogenetic group | % of isolates (no./total tested) with: | | No. of PAIs | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Single PAI | Multiple PAIs | $I_{536}$ | $II_{536}$ | $III_{536}$ | $IV_{536}$ | $I_{J96}$ | $II_{J96}$ | $I_{CFT073}$ | $II_{CFT073}$ |
| A | 27 (27/100) | 6 (6/100) | 0 | 1 | 0 | 23 | 0 | 0 | 5 | 6 |
| $B_1$ | 46 (30/65) | 15 (10/65) | 0 | 1 | 0 | 28 | 0 | 0 | 10 | 8 |
| $B_2$ | 83 (24/29) | 69 (20/29) | 3 | 2 | 0 | 22 | 0 | 0 | 21 | 8 |
| D | 54 (14/26) | 35 (9/26) | 0 | 0 | 0 | 11 | 0 | 0 | 10 | 2 |

analyze the phylogroups of fecal *E. coli* from asymptomatic humans in China. The results indicate that commensal *E. coli* strains from group A predominate in the gut flora. Moreover, some fecal *E. coli* strains appeared to be potentially virulent.

We thank the subjects and staff of the physical examination center of the Affiliated Union Hospital of Fujian Medical University for their kind support.

## REFERENCES

1. **Bidet, P., S. Bonacorsi, O. Clermont, C. De Montille, N. Brahimi, and E. Bingen.** 2005. Multiple insertional events, restricted by the genetic background, have led to acquisition of pathogenicity island IIJ96-like domains among *Escherichia coli* strains of different clinical origins. Infect. Immun. **73:**4081–4087.
2. **Clermont, O., S. Bonacorsi, and E. Bingen.** 2000. Rapid and simple determination of the *Escherichia coli* phylogenetic group. Appl. Environ. Microbiol. **66:**4555–4558.
3. **Duriez, P., O. Clermont, S. Bonacorsi, E. Bingen, A. Chaventré, J. Elion, B. Picard, and E. Denamur.** 2001. Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. Microbiology **147:**1671–1676.
4. **Escobar-Páramo, P., K. Grenet, A. Le Menac', L. Rode, E. Salgado, C. Amorin, S. Gouriou, B. Picard, M. C. Rahimy, A. Andremout, E. Denamur, and R. Ruimy.** 2004. Large-scale population structure of human commensal *Escherichia coli* isolates. Appl. Environ. Microbiol. **70:**5698–5700.
5. **Escobar-Paramo, P., O. Clermont, A. B. Blanc-Potard, H. Bui, C. Le Bouguénec, and E. Denamur.** 2004. A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. Mol. Biol. Evol. **21:**1085–1094.
6. **Gordon, D. M., S. E. Stern, and P. J. Collignon.** 2005. Influence of the age and sex of human hosts on the distribution of Escherichia coli ECOR groups and virulence traits. Microbiology **151:**15–23.
7. **Gordon, D. M., O. Clermont, H. Tolley, and E. Denamur.** 2008. Assigning Escherichia coli strains to phylogenetic groups: multi-locus sequence typing versus the PCR triplex method. Environ. Microbiol. **10:**2484–2496.
8. **Higgins, J., H. Christian, S. Hornor, M. Frana, M. Denver, and R. Joerger.** 2007. Genotyping of *Escherichia coli* from environmental and animal samples. J. Microbiol. Methods **70:**227–235.
9. **Johnson, J. R., P. Delavari, M. Kuskowski, and A. L. Stell.** 2001. Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*. J. Infect. Dis. **183:**78–88.
10. **Karch, H., S. Schubert, D. Zhang, W. Zhang, H. Schmidt, T. Olschläger, and J. Hacker.** 1999. A genomic island, termed high-pathogenicity island, is present in certain non-O157 Shiga toxin-producing *Escherichia coli* clonal lineages. Infect. Immun. **67:**5994–6001.
11. **Ochman, H., J. G. Lawrence, and E. A. Groisman.** 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405:**299–304.
12. **Oelschlaeger, T. A., U. Dobrindt, and J. Hacker.** 2002. Pathogenicity islands of uropathogenic *E. coli* and the evolution of virulence. Int. J. Antimicrob. Agents **19:**517–521.
13. **Picard, B., J. S. Garcia, S. Gouriou, P. Duriez, N. Brahimi, E. Bingen, J. Elion, and E. Denamur.** 1999. The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. Infect. Immun. **67:**546–553.
14. **Sabaté, M., E. Moreno, T. Pérez, A. Andreu, and G. Prats.** 2006. Pathogenicity island markers in commensal and uropathogenic *Escherichia coli* isolates. Clin. Microbiol. Infect. **12:**880–886.
15. **Schmidt, H., and M. Hensel.** 2004. Pathogenicity islands in bacterial pathogenesis. Clin. Microbiol. Rev. **17:**14–56.
16. **Schubert, S., A. Rakin, and J. Heesemann.** 2004. The Yersinia high pathogenicity island (HPI): evolutionary and functional aspects. Int. J. Med. Microbiol. **294:**83–94.
17. **Schubert, S., B. Picard, S. Gouriou, J. Heesemann, and E. Denamur.** 2002. Yersinia high-pathogenicity island contributes to virulence in *Escherichia coli* causing extraintestinal infections. Infect. Immun. **70:**5335–5337.
18. **Unno, T., D. Han, J. Jang, S. N. Lee, G. Ko, H. Y. Choi, J. H. Kim, M. J. Sadowsky, and H. G. Hur.** 2009. Absence of *Escherichia coli* phylogenetic group $B_2$ strains in humans and domesticated animals from Jeonnam Province, Republic of Korea. Appl. Environ. Microbiol. **75:**5659–5666.
19. **Walk, S. T., E. W. Alm, L. M. Calhoun, J. M. Mladonicky, and T. S. Whittam.** 2007. Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. Environ. Microbiol. **9:**2274–2288.