

**Supplementary Table 1.** Sequenced *E. coli* strains used for phylogenetic analysis.

<i>E. coli</i> strain label	GenBank Accession	Strain Characteristics	Reference
<b>Extraintestinal pathogenic strains including uropathogens</b>			
536	<a href="#">NC_008253</a>	Uropathogen	(8)
CFT073	<a href="#">NC_004431</a>	Uropathogen	(20)
UMN026	<a href="#">NC_011751</a>	Uropathogen	(19)
IAI39	<a href="#">NC_011750</a>	Uropathogen	
S88	<a href="#">NC_011742</a>	Neonatal meningitis strain	
UTI89	<a href="#">NC_007946</a>	Uropathogen	(1)
APEC	<a href="#">NC_008563</a>	Avian pathogenic strain	(11)
<b>Human commensal &amp; environmental strains</b>			
HS	<a href="#">NC_009800</a>	Human commensal	(17)
ED1a	<a href="#">NC_011745</a>	Human commensal	(19)
IAI1	<a href="#">NC_011741</a>	Human commensal	
SE11	<a href="#">NC_011415</a>	Human commensal	(15)
SMS-3-5	<a href="#">NC_010498</a>	Multi-drug resistant environmental strain	(6)
<b>Intestinal pathogenic strains</b>			
O111:H- Str 11128	<a href="#">NC_013364</a>	Enterohaemorrhagic <i>E. coli</i> (EHEC)	(14)
O26:H11 Str 11368	<a href="#">NC_013361</a>	EHEC	
O103:H2 Str 12009	<a href="#">NC_013353</a>	EHEC	
O157:H7 Str TW14359	<a href="#">NC_013008</a>	EHEC	(12)
O157:H7 Str EDL933	<a href="#">NC_002655</a>	EHEC	(16)
O55:H7 Str CB9615	<a href="#">NC_013941</a>	Precursor of O157:H7 EHEC	(21)
O157: H7 Str Sakai	<a href="#">NC_002695</a>	EHEC	(7)
O157:H7 Str EC4115	<a href="#">NC_011353</a>	EHEC	(4)
E24377A	<a href="#">NC_009801</a>	ETEC	(17)
S5989	<a href="#">NC_011748</a>	Enteroaggregative <i>E. coli</i> (EAEC)	(19)
O127:H6 Str E2348/69	<a href="#">NC_011601</a>	Enteropathogenic <i>E. coli</i> (EPEC)	(9)
<b>Laboratory &amp; other strains</b>			
K-12 MG1655	<a href="#">NC_000913</a>	Laboratory strain	(18)
BW2952	<a href="#">NC_012759</a>	Laboratory strain	(5)
REL606	<a href="#">NC_012967</a>	Laboratory strain, similar to K-12 MG1655	(10)
K-12 substr DH10B	<a href="#">NC_010473</a>	Laboratory strain, derived from K-12 MG1655	(3)
<i>E. coli</i> 'BL21-Gold(DE3)pLysS AG'	<a href="#">NC_012947</a>	No information available	(13)
ATCC 8739 strain	<a href="#">NC_010468</a>	No information available	(2)

## References

1. **Chen, S. L., C. S. Hung, J. Xu, C. S. Reigstad, V. Magrini, A. Sabo, D. Blasiar, T. Bieri, R. R. Meyer, P. Ozersky, J. R. Armstrong, R. S. Fulton, J. P. Latreille, J. Spieth, T. M. Hooton, E. R. Mardis, S. J. Hultgren, and J. I. Gordon.** 2006. Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. *Proceedings of the National Academy of Sciences of the United States of America* **103**:5977-5982.

2. **Copeland, A., Lucas, S., Lapidus, A., Glavina del Rio, T., Dalin, E., Tice, H., Bruce, D., Goodwin, L., Pitluck, S., Kiss, H., Brettin, T., Detter, J.C., Han, C., Kuske, C.R., Schmutz, J., Larimer, F., Land, M., Hauser, L., Kyrpides, N., Mikhailova, N., Ingram, L. and Richardson, P.** 2008. Complete sequence of *Escherichia coli* C str. ATCC 8739 (direct submission). NCBI Genome Project, <http://www.jgi.doe.gov>.
3. **Durfee, T., R. Nelson, S. Baldwin, G. Plunkett, 3rd, V. Burland, B. Mau, J. F. Petrosino, X. Qin, D. M. Muzny, M. Ayele, R. A. Gibbs, B. Csorgo, G. Posfai, G. M. Weinstock, and F. R. Blattner.** 2008. The complete genome sequence of *Escherichia coli* DH10B: insights into the biology of a laboratory workhorse. *Journal of bacteriology* **190**:2597-2606.
4. **Eppinger, M., M. K. Mammel, J. E. Leclerc, J. Ravel, and T. A. Cebula.** 2011. Genomic anatomy of *Escherichia coli* O157:H7 outbreaks. *Proceedings of the National Academy of Sciences of the United States of America* **108**:20142-20147.
5. **Ferenci, T., Z. Zhou, T. Betteridge, Y. Ren, Y. Liu, L. Feng, P. R. Reeves, and L. Wang.** 2009. Genomic sequencing reveals regulatory mutations and recombinational events in the widely used MC4100 lineage of *Escherichia coli* K-12. *Journal of bacteriology* **191**:4025-4029.
6. **Fricke, W. F., M. S. Wright, A. H. Lindell, D. M. Harkins, C. Baker-Austin, J. Ravel, and R. Stepanauskas.** 2008. Insights into the environmental resistance gene pool from the genome sequence of the multidrug-resistant environmental isolate *Escherichia coli* SMS-3-5. *Journal of bacteriology* **190**:6779-6794.
7. **Hayashi, T., K. Makino, M. Ohnishi, K. Kurokawa, K. Ishii, K. Yokoyama, C. G. Han, E. Ohtsubo, K. Nakayama, T. Murata, M. Tanaka, T. Tobe, T. Iida, H. Takami, T. Honda, C. Sasakawa, N. Ogasawara, T. Yasunaga, S. Kuhara, T. Shiba, M. Hattori, and H. Shinagawa.** 2001. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA research : an international journal for rapid publication of reports on genes and genomes* **8**:11-22.
8. **Hochhut, B., C. Wilde, G. Balling, B. Middendorf, U. Dobrindt, E. Brzuszkiewicz, G. Gottschalk, E. Carniel, and J. Hacker.** 2006. Role of pathogenicity island-associated integrases in the genome plasticity of uropathogenic *Escherichia coli* strain 536. *Molecular microbiology* **61**:584-595.
9. **Iguchi, A., N. R. Thomson, Y. Ogura, D. Saunders, T. Ooka, I. R. Henderson, D. Harris, M. Asadulghani, K. Kurokawa, P. Dean, B. Kenny, M. A. Quail, S. Thurston, G. Dougan, T. Hayashi, J. Parkhill, and G. Frankel.** 2009. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. *Journal of bacteriology* **191**:347-354.
10. **Jeong, H., V. Barbe, C. H. Lee, D. Vallenet, D. S. Yu, S. H. Choi, A. Couloux, S. W. Lee, S. H. Yoon, L. Cattolico, C. G. Hur, H. S. Park, B. Segurens, S. C. Kim, T. K. Oh, R. E. Lenski, F. W. Studier, P. Daegelen, and J. F. Kim.** 2009. Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *Journal of molecular biology* **394**:644-652.
11. **Johnson, T. J., S. Kariyawasam, Y. Wannemuehler, P. Mangiamele, S. J. Johnson, C. Doetkott, J. A. Skyberg, A. M. Lynne, J. R. Johnson, and L. K. Nolan.** 2007. The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. *Journal of bacteriology* **189**:3228-3236.

12. **Kulasekara, B. R., M. Jacobs, Y. Zhou, Z. Wu, E. Sims, C. Saenphimmachak, L. Rohmer, J. M. Ritchie, M. Radey, M. McKeivitt, T. L. Freeman, H. Hayden, E. Haugen, W. Gillett, C. Fong, J. Chang, V. Beskhlebnaya, M. K. Waldor, M. Samadpour, T. S. Whittam, R. Kaul, M. Brittnacher, and S. I. Miller.** 2009. Analysis of the genome of the *Escherichia coli* O157:H7 2006 spinach-associated outbreak isolate indicates candidate genes that may enhance virulence. *Infection and immunity* **77**:3713-3721.
13. **Lucas, S., Copeland, A., Lapidus, A., Glavina del Rio, T., Dalin, E., Tice, H., Bruce, D., Goodwin, L., Pitluck, S., LaButti, K. M., Clum, A., Larimer, F., Land, M., Hauser, L., Kyrpides, N., Anderson, I., Sorek, R. and Rubin, E.** 2009. Complete sequence of *Escherichia coli* BL21'-Gold(DE3)pLysS AG' (direct submission). NCBI Genome Project, <http://www.jgi.doe.gov>.
14. **Ogura, Y., T. Ooka, A. Iguchi, H. Toh, M. Asadulghani, K. Oshima, T. Kodama, H. Abe, K. Nakayama, K. Kurokawa, T. Tobe, M. Hattori, and T. Hayashi.** 2009. Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **106**:17939-17944.
15. **Oshima, K., H. Toh, Y. Ogura, H. Sasamoto, H. Morita, S. H. Park, T. Ooka, S. Iyoda, T. D. Taylor, T. Hayashi, K. Itoh, and M. Hattori.** 2008. Complete genome sequence and comparative analysis of the wild-type commensal *Escherichia coli* strain SE11 isolated from a healthy adult. *DNA research : an international journal for rapid publication of reports on genes and genomes* **15**:375-386.
16. **Perna, N. T., G. Plunkett, 3rd, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner.** 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**:529-533.
17. **Rasko, D. A., M. J. Rosovitz, G. S. Myers, E. F. Mongodin, W. F. Fricke, P. Gajer, J. Crabtree, M. Sebaihia, N. R. Thomson, R. Chaudhuri, I. R. Henderson, V. Sperandio, and J. Ravel.** 2008. The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E. coli* commensal and pathogenic isolates. *Journal of bacteriology* **190**:6881-6893.
18. **Riley, M., T. Abe, M. B. Arnaud, M. K. Berlyn, F. R. Blattner, R. R. Chaudhuri, J. D. Glasner, T. Horiuchi, I. M. Keseler, T. Kosuge, H. Mori, N. T. Perna, G. Plunkett, 3rd, K. E. Rudd, M. H. Serres, G. H. Thomas, N. R. Thomson, D. Wishart, and B. L. Wanner.** 2006. *Escherichia coli* K-12: a cooperatively developed annotation snapshot--2005. *Nucleic acids research* **34**:1-9.
19. **Touchon, M., C. Hoede, O. Tenaillon, V. Barbe, S. Baeriswyl, P. Bidet, E. Bingen, S. Bonacorsi, C. Bouchier, O. Bouvet, A. Calteau, H. Chiapello, O. Clermont, S. Cruveiller, A. Danchin, M. Diard, C. Dossat, M. E. Karoui, E. Frapy, L. Garry, J. M. Ghigo, A. M. Gilles, J. Johnson, C. Le Bouguenec, M. Lescat, S. Mangenot, V. Martinez-Jehanne, I. Matic, X. Nassif, S. Oztas, M. A. Petit, C. Pichon, Z. Rouy, C. S. Ruf, D. Schneider, J. Tourret, B. Vacherie, D. Vallenet, C. Medigue, E. P. Rocha, and E. Denamur.** 2009.

Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS genetics* **5**:e1000344.

20. **Welch, R. A., V. Burland, G. Plunkett, 3rd, P. Redford, P. Roesch, D. Rasko, E. L. Buckles, S. R. Liou, A. Boutin, J. Hackett, D. Stroud, G. F. Mayhew, D. J. Rose, S. Zhou, D. C. Schwartz, N. T. Perna, H. L. Mobley, M. S. Sonnenberg, and F. R. Blattner.** 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **99**:17020-17024.
21. **Zhou, Z., X. Li, B. Liu, L. Beutin, J. Xu, Y. Ren, L. Feng, R. Lan, P. R. Reeves, and L. Wang.** 2010. Derivation of *Escherichia coli* O157:H7 from its O55:H7 precursor. *PloS one* **5**:e8700.