

Hypothesis

Open Access

## Lateral Gene Transfer (LGT) between Archaea and *Escherichia coli* is a contributor to the emergence of novel infectious disease

David M Faguy\*

Address: Dept. of Biology, University of New Mexico, Albuquerque, NM 87131, USA

Email: David M Faguy\* - dfaguy@unm.edu

\* Corresponding author

Published: 19 June 2003

Received: 09 January 2003

*BMC Infectious Diseases* 2003, **3**:13

Accepted: 19 June 2003

This article is available from: <http://www.biomedcentral.com/1471-2334/3/13>

© 2003 Faguy; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

### Abstract

**Background:** Lateral gene transfer is the major mechanism for acquisition of new virulence genes in pathogens. Recent whole genome analyses have suggested massive gene transfer between widely divergent organisms.

**Presentation of the hypothesis:** Archeal-like genes acting as virulence genes are present in several pathogens and genomes contain a number of archaeal-like genes of unknown function. Archaea, by virtue of their very different evolutionary history and different environment, provide a pool of potential virulence genes to bacterial pathogens.

**Testing the hypothesis:** We can test this hypothesis by 1) identifying genes likely to have been transferred (directly or indirectly) to *E. coli* O157:H7 from archaea; 2) investigating the distribution of similar genes in pathogens and non-pathogens and 3) performing rigorous phylogenetic analyses on putative transfers.

**Implications of the hypothesis:** Although this hypothesis focuses on archaea and *E. coli*, it will serve as a model having broad applicability to a number of pathogenic systems. Since no archaea are known vertebrate pathogens, archaeal-like transferred genes that are associated with virulence in bacteria represent a clear model for the emergence of virulence genes.

### Background

The fact and process of lateral gene transfer (LGT) has been integral to the study of infectious disease since Griffith [1]. Most investigations have, however, centered on transfer of known virulence genes and genes involved in establishing infections (antibiotic resistance, toxins, capsule, etc.) between close relatives who are both pathogens. The explosion of gene sequence information and complete genomes in the last 5 years has reinforced and extended our view of virulence evolution. To quote a review [2]: "lateral transfers have effectively changed the ecological and pathogenic character of bacterial species." Analysis of gene sequence and complete genome informa-

tion [3–11] have led to the realization that LGT is not a rare exception to classical Darwinian evolution, but may be the predominant mode of evolutionary change in prokaryotes [2, 12–15].

Given the common perception of archaea as extremophiles, one might expect the opportunity for transfer between archaeal organisms and bacterial pathogens to be a rare event. However, many archaea are present in more "normal" environments. Methanogenic archaea are very common in vertebrate and invertebrate digestive systems. In fact, in dairy cattle, methanogenic archaea make up a substantial fraction of the micro biota [16] and *E. coli*

O157 is also very common, being cultured from 75% of herds [17]. Furthermore, there are similarities among bacterial and archaeal phages, plasmids, and other vectors for mediating transfer (IS elements; transposable elements)[18]. Transfer of DNA in natural environments has been extensively described for a number of different organisms: Bushman (Table 5.5,[19]) lists more than 30 studies. Many different functions – proteases, metabolic genes, oxygen protection genes, secretion systems, transporter genes, iron acquisition systems – can, under the right circumstances, contribute to virulence in an emerging pathogen.

*E. coli* O157:H7 causes hemorrhagic colitis and hemolytic uremic syndrome. It is widely recognized as a worldwide public health danger. However, it has only been associated with human disease since 1982. There is a complete genome sequence available for *E. coli* K-12[20], as well as much genetic information on many other well-studied non-pathogenic *E. coli*. There are two published complete genomes for *E. coli* O157:H7 enteropathogenic strains [8,21], as well as seven *E. coli* genomes in progress [22] (GOLD database [23]). Virulence also depends on many different genes in enterohemorrhagic *E. coli* [8]. For the purposes of testing the contribution of archaea to virulence genes, *E. coli* O157:H7 serves as an ideal model.

#### Testing the hypothesis

We can test this hypothesis by:

- 1) Identifying genes likely to have been transferred (directly or indirectly) to *E. coli* O157:H7 from archaea.
- 2) Investigating the distribution of similar genes in pathogenic and non-pathogens and performing rigorous phylogenetic analyses on putative transfers.

#### Have any transfers been described between archaea and *E. coli* O157?

I have demonstrated that a gene coding for a bifunctional catalase-peroxidase is likely a transfer from archaea to a variety of pathogenic bacteria, including *E. coli* O157:H7[9]. Although not yet directly implicated in O157:H7 as a virulence factor, this enzyme has been implicated as a virulence factor in *Mycobacterium tuberculosis* [24,26], and in *Legionella pneumophila* [25]. Furthermore, this *E. coli* O157:H7 catalase-peroxidase has been associated with enterohaemorrhagic hemolysin in a variety of shiga-like toxin-producing (verotoxin-producing) *E. coli* [30,31]. This correlation of the presence of the catalase-peroxidase in many virulent but not in avirulent strains suggests a direct role in the virulence of enterohemorrhagic *E. coli*.

#### How can we identify other genes likely to have been transferred (directly or indirectly) to *E. coli* O157:H7 from archaea?

We can use the three complete *E. coli* genomes available[8,20,21] to identify a subset of genes present in one (or both) of the O157 strains but not in K-12. This subset of genes – although most have not been directly identified as virulence genes – are more likely to be virulence-associated. This subset will be the focus of our search for genes likely to have laterally transferred from archaea. Perna *et al* [8] identified 1,387 genes in the EDL933 O157 strain they sequenced that are O157-specific; This number does not include the O157-specific plasmids that were previously sequenced[32] from which some of the preliminary work[9] described below was derived.

As a preliminary step to testing this hypothesis I have searched the *E. coli* O157:H7 strain EDL933[8] genome for open reading frames (ORFs) meeting the following criteria:

- 1) present in O157 strain EDL933 but not in *E. coli* K-12
- 2) highly similar (having a BLASTP similarity bit score > 95) to ORFs found in at least two archaeal genomes
- 3) having few or no highly similar proteins (BLASTP score < 85) in other bacterial genomes

This search was facilitated by the Clusters of Orthologous Groups (COG) database at NCBI [33,34]. This preliminary, non-systematic search produced 6 ORFs worth considering as LGTs from archaea to pathogenic bacteria. Table 1 shows these ORFs with their location and, if any information was available, a possible function.

Although the putative functions in Table 1 are not generally associated with virulence, many genes, under the right circumstances, can facilitate infection or contribute to pathogenicity. There are examples in the literature where both ABC transporters (in *Streptococcus gordonii*:[35]) and helicase genes (in *Legionella pneumophila* :[36]) have been directly implicated in virulence.

Further testing of this hypothesis will require rigorous phylogenetic analyses of each suspected transfer. The procedure of comparing similarity scores to identify potential lateral transfers (used above) although commonly employed[6,7,37] it is fraught with potential errors [15,38–42,34,35] and must serve only as an initial screen. Ragan[15] recently wrote:

*This study demonstrates the need for a systematic, comprehensive approach to the study of LGT based on first principles, i.e. rigorous inference and statistically based comparison of*

**Table 1: Potential Archaeal to Bacterial Laterally Transferred ORFs in *E. coli* O157:H7 EDL933 Genome. Although ORFs Z5331 and Z0509 do not have archaeal genes as the most similar BLAST hit, they are included because in general they are absent from almost all bacterial genomes.**

<i>E. coli</i> O157:H7 designation	Location (O island)	Best BLAST hit to ORF from:	Best guess at function
Z4267	119	<i>Methanobacterium thermoautotrophicum</i>	ABC transporter : cations?
Z4271	119	<i>Methanobacterium thermoautotrophicum</i>	ABC transporter : cations?
Z5331	155	<i>Helicobacter pylori</i>	?
Z0509	25	<i>Yersinia pestis</i>	?
Z5989	172	<i>Halobacterium</i> sp. NRC-1	Helicase
Z5900	172	<i>Halobacterium</i> sp. NRC-1	Helicase

molecular phylogenetic trees. As more genomic sequences appear, a tree-based approach will become both more challenging and more rewarding.

### Implications of the hypothesis

Although this hypothesis focuses on archaea and *E. coli*, the model of distant gene transfer as a major contributor of "new" virulence genes to pathogens or potential pathogens has broad applicability to a large number of pathogenic systems. Archaea represent both the most distant source and, in many ways, the most unlikely source for virulence genes. If *E. coli* can acquire virulence genes from archaea, then potentially any organism is a reservoir of virulence genes for pathogens. [43]

The implications, should this hypothesis be proven, are myriad. Our understanding of potential sources of virulence genes will be expanded to include virtually all life on earth – at first glance a frightening prospect. At the same time, however, it would allow us to move from a descriptive, reactionary view of infectious disease towards a predictive science of infectious disease. It would be dramatic evidence of what some microbiologists suspect: that lateral gene transfer is the predominant engine of variation in prokaryotes and the catalyst for the emergence of new bacterial pathogens.

### Competing interests

None declared.

### Acknowledgements

The author would like to thank E. S. Loker and W. F. Doolittle for helpful discussions and comments on an early version of the manuscript and Hannu Myllykallio and Jacques Ravel for helpful and insightful comments during review. This work was supported by funding from the Research Allocations Committee of the University of New Mexico.

### References

- Griffith F: **The significance of pneumococcal types** *Journal of Hygiene* 1928, **27**:113-159.
- Ochman H, Lawrence JG and Groisman EA: **Lateral gene transfer and the nature of bacterial innovation** *Nature* 2000, **405**:299-304.
- Syvanen M: **Cross-species gene transfer; implications for a new theory of evolution** *J Theor Biol* 1985, **112**:333-343.
- Hilario E and Gogarten JP: **Horizontal transfer of ATPase genes—the tree of life becomes a net of life** *Biosystems* 1993, **31**:111-119.
- Lawrence JG and Ochman H: **Molecular archaeology of the *Escherichia coli* genome** *Proc Natl Acad Sci U S A* 1998, **95**:9413-9417.
- Aravind L, Tatusov RL, Wolf YI, Walker DR and Koonin EV: **Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles** *Trends Genet* 1998, **14**:442-444.
- Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Nelson WC, Ketchum KA, McDonald L, Utterback TR, Malek JA, Linher KD, Garrett MM, Stewart AM, Cotton MD, Pratt MS, Phillips CA, Richardson D, Heidelberg J, Sutton GG, Fleischmann RD, Eisen JA, Fraser CM and et al.: **Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*** *Nature* 1999, **399**:323-329.
- Perna NT, Plunkett G., 3rd, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Posfai G, Hackett J, Klink S, Boutin A, Shao Y, Miller L, Grotbeck EJ, Davis NW, Lim A, Dimalanta ET, Potamousis KD, Apodaca J, Anantharaman TS, Lin J, Yen G, Schwartz DC, Welch RA and Blattner FR: **Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7** *Nature* 2001, **409**:529-533.
- Faguy DM and Doolittle WF: **Horizontal transfer of catalase-peroxidase genes between archaea and pathogenic bacteria** *Trends Genet* 2000, **16**:196-7. \_00002007\_00002007.
- Faguy DM: **The controlled chaos of shifty pathogens** *Curr Biol* 2000, **10**:R498-501.
- Koonin EV, Makarova KS and Aravind L: **Horizontal gene transfer in prokaryotes: quantification and classification** *Annu Rev Microbiol* 2001, **55**:709-742.
- Doolittle WF: **Phylogenetic classification and the universal tree** *Science* 1999, **284**:2124-2129.
- Doolittle WF: **Uprooting the tree of life** *Sci Am* 2000, **282**:90-95.
- Levin BR and Bergstrom CT: **Bacteria are different: observations, interpretations, speculations, and opinions about the mechanisms of adaptive evolution in prokaryotes** *Proc Natl Acad Sci U S A* 2000, **97**:6981-6985.
- Ragan MA: **Detection of lateral gene transfer among microbial genomes** *Curr Opin Genet Dev* 2001, **11**:620-626.
- Lin CZ, Raskin L and Stahl DA: **Microbial community structure in gastrointestinal tracts of domestic animals: Comparative analyses using rRNA-targeted oligonucleotide probes** *Fems Microbiology Ecology* 1997, **22**:281-294.
- Hancock DD, Rice DH, Herriott DE, Besser TE, Ebel ED and Carpenter LV: **Effects of farm manure-handling practices on *Escherichia coli* O157 prevalence in cattle** *Journal of Food Protection* 1997, **60**:363-366.
- Zillig W, Prangishvili D, Schleper C, Elferink M, Holz I, Albers S, Janevskovic D and Gotz D: **Viruses, plasmids and other genetic elements of thermophilic and hyperthermophilic Archaea** *FEMS Microbiol Rev* 1996, **18**:225-236.

19. Bushman Frederic: **Lateral DNA transfer : mechanisms and consequences** Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press; 2002:xiv, 448.
20. Blattner FR, Plunkett G., 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B and Shao Y: **The complete genome sequence of Escherichia coli K-12** Science 1997, **277**:1453-1474.
21. Hayashi T, Makino K, Ohnishi M, Kurokawa K, Ishii K, Yokoyama K, Han CG, Ohtsubo E, Nakayama K, Murata T, Tanaka M, Tobe T, Iida T, Takami H, Honda T, Sasakawa C, Ogasawara N, Yasunaga T, Kuhara S, Shiba T, Hattori M and Shinagawa H: **Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12 DNA Res** 2001, **8**:11-22.
22. Welch RA, Burland V, Plunkett G., 3rd, Redford P, Roesch P, Rasko D, Buckles EL, Liou SR, Boutin A, Hackett J, Stroud D, Mayhew GF, Rose DJ, Zhou S, Schwartz DC, Perna NT, Mobley HL, Donnenberg MS and Blattner FR: **Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli** Proc Natl Acad Sci U S A 2002, **99**:17020-17024.
23. Bernal A, Ear U and Kyrpides N: **Genomes OnLine Database (GOLD): a monitor of genome projects world-wide** Nucleic Acids Res 2001, **29**:126-127.
24. Zhang XL, Morris C and Hackett J: **Molecular cloning, nucleotide sequence, and function of a site-specific recombinase encoded in the major 'pathogenicity island' of Salmonella typhi** Gene 1997, **202**:139-146.
25. Manca C, Paul S, Barry C. E., 3rd, Freedman VH and Kaplan G: **Mycobacterium tuberculosis catalase and peroxidase activities and resistance to oxidative killing in human monocytes in vitro** Infect Immun 1999, **67**:74-79.
26. Yu S, Chouchane S and Magliozzo RS: **Characterization of the W321F mutant of Mycobacterium tuberculosis catalase-peroxidase KatG** Protein Sci 2002, **11**:58-64.
27. Bandyopadhyay P and Steinman HM: **Legionella pneumophila catalase-peroxidases: cloning of the katB gene and studies of KatB function** J Bacteriol 1998, **180**:5369-5374.
28. Cianciotto NP: **Pathogenicity of Legionella pneumophila** Int J Med Microbiol 2001, **291**:331-343.
29. Bandyopadhyay P and Steinman HM: **Catalase-peroxidases of Legionella pneumophila: cloning of the katA gene and studies of KatA function** J Bacteriol 2000, **182**:6679-6686.
30. Brunder WW, Schmidt H and Karch H: **KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic Escherichia coli O157:H7** Microbiology 1996, **142**:3305-3315.
31. Stephan R, Ragettli S and Untermann F: **Prevalence and characteristics of verotoxin-producing Escherichia coli (VTEC) in stool samples from asymptomatic human carriers working in the meat processing industry in Switzerland** J Appl Microbiol 2000, **88**:335-341.
32. Burland V, Shao Y, Perna NT, Plunkett G, Sofia HJ and Blattner FR: **The complete DNA sequence and analysis of the large virulence plasmid of Escherichia coli O157:H7** Nucleic Acids Res 1998, **26**:4196-4204.
33. Tatusov RL, Koonin EV and Lipman DJ: **A genomic perspective on protein families** Science 1997, **278**:631-637.
34. Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND and Koonin EV: **The COG database: new developments in phylogenetic classification of proteins from complete genomes** Nucleic Acids Res 2001, **29**:22-28.
35. Jakubovics NS, Smith AW and Jenkinson HF: **Oxidative stress tolerance is manganese (Mn(2+)) regulated in Streptococcus gordonii** Microbiology 2002, **148**:3255-3263.
36. Harb OS and Abu Kwaik Y: **Essential role for the Legionella pneumophila rep helicase homologue in intracellular infection of mammalian cells** Infect Immun 2000, **68**:6970-6978.
37. Faguy DM and Doolittle WF: **Lessons from the Aeropyrum pernix genome** Curr Biol 1999, **9**:R883-6.
38. Logsdon JM and Faguy DM: **Thermotoga heats up lateral gene transfer** Curr Biol 1999, **9**:R747-51.
39. Kyrpides NC and Ouzounis CA: **Whole-genome sequence annotation: 'Going wrong with confidence'** Mol Microbiol 1999, **32**:886-887.
40. Eisen JA: **Horizontal gene transfer among microbial genomes: new insights from complete genome analysis** Curr Opin Genet Dev 2000, **10**:606-611.
41. Wang B: **Limitations of compositional approach to identifying horizontally transferred genes** J Mol Evol 2001, **53**:244-250.
42. Lawrence JG and Ochman H: **Reconciling the many faces of lateral gene transfer** Trends Microbiol 2002, **10**:1-4.
43. Parkhill J, Wren BW, Thomson NR, Titball RW, Holden MT, Prentice MB, Sebaihia M, James KD, Churcher C, Mungall KL, Baker S, Basham D, Bentley SD, Brooks K, Cerdeno-Tarraga AM, Chillingworth T, Cronin A, Davies RM, Davis P, Dougan G, Feltwell T, Hamlin N, Holroyd S, Jagels K, Karlyshev AV, Leather S, Moule S, Oyston PC, Quail M, Rutherford K, Simmonds M, Skelton J, Stevens K, Whitehead S and Barrell BG: **Genome sequence of Yersinia pestis, the causative agent of plague** Nature 2001, **413**:523-527.

### Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2334/3/13/prepub>

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

