Bloody *coli*: a Gene Cocktail in *Escherichia coli* O104:H4

Fernando Baquero,[a] Raquel Tobes[b]

Department of Microbiology, Ramón y Cajal University Hospital, IRYCIS, CIBERESP, Madrid, Spain[a]; Research Group, Era7 Bioinformatics, Granada, Spain[b]

ABSTRACT A recent study published in *mBio* [Y. H. Grad et al., *mBio* 4(1):e00452-12, 2013] indicates that a rapid introgressive evolution has occurred in *Escherichia coli* O104:H4 by sequential acquisition of foreign genetic material involving pathogenicity traits. O104 genetic promiscuity cannot be readily explained by high population sizes. However, extensive interactions leading to cumulative assemblies of pathogenicity genes might be assured by small *K*-strategist populations exploiting particular intestinal niches. Next-generation sequencing technologies will be critical to detect particular "gene cocktails" as potentially pathogenic ensembles and to predict the risk of future outbreaks.

The cell is a unit of interaction. Bacterial cells from the species *Escherichia coli* specifically interact with the lower gastrointestinal tracts (including the microbiotas) of animals and humans, a relationship resulting from a long-term coevolutionary process that has shaped the well-defined *E. coli* core genome. We fully agree with the proposal that *E. coli* should be treated as a single microorganism, in spite of the fact that strains are often classified according to their intestinal pathotypes, including as enteropathogenic, enterotoxigenic, enterohemorrhagic, enteroaggregative, enteroinvasive, adherent-invasive (1), and, in the case of O104:H4, enteroaggregative hemorrhagic *E. coli* (EAHEC). All of these *E. coli* pathotypes have essentially the same core genome (comprising about 2,000 genes) maintained by vertical descent. During the course of *E. coli*'s evolution, a basic type of the *E. coli* genome seems to have been complemented by the acquisition by horizontal transfer of different adaptive (including "pathogenic") genes, genes that are propagated by introgressive descent (2). The outcome is the emergence of a variety of strains with different colonization or pathogenic abilities. From this perspective, it may be inappropriate to link the term "pathogenic" to particular serotypes or *E. coli* multilocus sequence types (MLSTs). Classic "pathogenic types" might correspond to those types where the acquisition of pathogenicity traits (PTs) has been documented with particular frequency. However, not all *E. coli* serotypes or sequence types (STs) are equally distributed in all habitats, indicating that a number of noncore genes have evolved to provide different degrees of ecological specificity, eventually leading to some kind of ecological barrier for intraspecies genetic exchanges among *E. coli* genomes from different ecological environments (3). Certainly not all pathogenicity traits are equally distributed among *E. coli* serotypes or STs. That probably means that a highly pathogenic clone emerges when an *E. coli* type able to sustain particular environmental interactions accumulates pathogenic genes. This is, in a sense, when virulence meets metabolism (4).

The strains belonging to the O104:H4 serotype (ST678, phylogroup B1), responsible for the severe German outbreak of bloody diarrhea and hemolytic-uremic syndrome in 2011, illustrate this concept. In their paper in *mBio*, "Comparative genomics of recent Shiga toxin-producing *Escherichia coli* O104:H4: short-term evolution of an emerging pathogen," Grad et al. (5) indicate that a rapid introgressive evolution has occurred in these strains by sequential acquisition of foreign genetic material, including such pathogenicity traits as Shiga toxin and aggregative-adherence fimbriae. The result is a highly pathogenic behavior for *E. coli* in humans, but what are the evolutionary benefits for the bacterial organism?

Such evolutionary benefits should exist. It seems likely that *E. coli* O104:H4 has undergone selection in some way or another in the recent past and has enlarged its population size and/or improved its adaptation to multiple habitats. The short-term evolution indicated by Grad et al. (5) and the cumulative acquisition of pathogenic traits require frequent and extensive ecological and genetic interactions with other bacterial (donor) cells, probably requiring a large number of cells and/or very effective dispersal, according to the genetic-capitalism principle (6). Studies based on whole-genome sequencing of several *E. coli* O104:H4 (ST678) strains isolated along the last few years revealed strong genetic differences in chromosomal and plasmid content (7). An unexpectedly high number of recombinant genes (125 genes) was found, and interestingly, the possible donors of these genes were not clustered in a single *E. coli* phylogenetic group. In fact, in half of cases, the recombinant genes contained sequences from donors in six phylogenetic groups. Even though the possibility of extensive recombination with a highly mosaic donor strain of another phylogroup cannot be totally excluded, differences among *E. coli* O104:H4 strains suggest separate sites and events in their recent evolutionary history. This implies a dense network of interactions with other bacteria. What could have been the necessary context for these interactions?

The simple answer is that organisms evolve from their natural reservoirs, where a sufficient population size can be reached. This natural reservoir, that is, the optimal environment for the reproduction, maintenance, and evolution of *E. coli* O104:H4, remains uncertain (8). *E. coli* is a normal component of the microbiotas of humans and animals, the natural place where *E. coli* strains might genetically interact with other *E. coli* strains and with many other *Proteobacteria*. Strains of serotype O104:H4 have been found mostly in association with sporadic cases of persistent or severe diarrhea since the 1990s, but little is known about their prevalence in healthy humans. Classic and modern studies from the 1960s indicate that O104 strains were scarcely found among normal hu-

Published 19 February 2013

Citation Baquero F, Tobes R. 2013. Bloody *coli*: a gene cocktail in *Escherichia coli* O104:H4. mBio 4(1):e00066-13. doi:10.1128/mBio.00066-13.

Address correspondence to Fernando Baquero, baquero@bitmailer.net.

is expected to occur. Increased connectivity facilitates further "community genetic exchanges" building, as in the case of *Shigella* (an *E. coli* derivative), a quasi-species group (Fig. 1).

Other possibilities to explain the promiscuous life of *E. coli* O104:H4 in spite of its seemingly low population density remain to be explored. This strain was consistently found in a common food (salad) vehicle involved in the Germany and France outbreaks, fenugreek sprouts (*Trigonella foenum-graecum*, family Fabaceae), presumably of African origin. Might this legume, used on this continent both for animal pasture and for human food, be something more than a vehicle? Could O104:H4 be ecologically associated (and genetically interact) with the root nodule bacteria of this plant? The possibility of an endemic status of *E. coli* O104:H4 in humans in Central Africa has recently been suggested (8). Was this plant exposed during agricultural farming to animal and human sewage? We should clarify the field (network) of interaction giving rise to the outbreak strain.

As in a crowdsourcing software project, the evolution of O104:H4 starts with an initial plastic bacterium project onto which new elements, including pathogenicity genes, are anonymously contributed by similar bacteria. At the end, we have different combinations of elements that develop and grow continuously within the crowdsourcing community (Fig. 1). How to follow that process is the aim of predicting risks for future outbreaks.

Surveilling a reduced number of conserved genes can produce deceiving results, because pathogenicity genes are frequently allocated to mobile modules with little sequence conservation. The availability of next-generation sequencing (NGS) technologies allows us to reveal the complete gene landscape of a bacterium and, moreover, of a community of bacteria. This new perspective on outbreaks circumvents the old bias of looking only at specific conserved genes belonging to the causative agents. NGS enables scientists to trace the importance of gene units that move freely between strains and to analyze the flux of these genes between bacterial communities. In the postgenomics era, NGS technologies are providing public health efforts with advanced and quick tools that allow researchers not only to retrospectively analyze the epidemiological evolution of an outbreak but also to predict its future evolution.

ACKNOWLEDGMENTS

F.B.'s research is sponsored by the European Union FP7 projects PAR-241476 and EvoTAR-282004 and the Regional Government of Madrid, Spain (grant PROMPT-S2010/BMD2414). R.T.'s research is sponsored by CDTI projects NEXTMICRO (grant IDI-20120242) and INNPACTO (brachVac; grant IPT-2011-0735-010000).

REFERENCES

1. **Moriel DG, Rosini R, Seib KL, Serino L, Pizza M, Rappuoli R.** 2012. Escherichia coli: great diversity around a common core. mBio **3**(3): e00118-12. http://dx.doi.org/10.1128/mBio.00118-12.
2. **Bapteste E, Lopez P, Bouchard F, Baquero F, McInerney JO, Burian RM.** 2012. Evolutionary analyses of non-genealogical bonds produced by introgressive descent. Proc. Natl. Acad. Sci. U. S. A. **109**:18266–18272.
3. **Luo C, Walk ST, Gordon DM, Feldgarden M, Tiedje JM, Konstantinidis KT.** 2011. Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. Proc. Natl. Acad. Sci. U. S. A. **108**:7200–7205.
4. **Njoroge JW, Nguyen Y, Curtis MM, Moreira CG, Sperandio V.** 2012. Virulence meets metabolism: cra and KdpE gene regulation in enterohemorrhagic *Escherichia coli*. mBio **3**(5):e00280-12. http://dx.doi.org/10.1128/mBio.00280-12.
5. **Grad YH, Godfrey P, Cerquiera GC, Mariani-Kurkdjian P, Gouali M, Bingen E, Shea TP, Haas BJ, Griggs A, Young S, Zeng Q, Lipsitch M, Waldor MK, Weill FX, Wortman JR, Hanage WP.** 2013. Comparative genomics of recent Shiga toxin-producing *E. coli* O104:H4: short-term evolution of an emerging pathogen. mBio **4**(1):e00452-12. http://dx.doi.org/10.1128/mBio.00452-12.
6. **Baquero F.** 2004. From pieces to patterns: evolutionary engineering of bacterial pathogens. Nat. Rev. Microbiol. **2**:510–518.
7. **Hao W, Allen VG, Jamieson FB, Low DE, Alexander DC.** 2012. Phylogenetic incongruence in *E. coli* O104: understanding the volutionary relationships of emerging pathogens in the face of homologous recombination. PLoS One **7**:e33971. http://dx.doi.org/10.1371/journal.pone.0033971.
8. **Beutin L, Martin A.** 2012. Outbreak of Shiga toxin-producing *Escherichia coli* (STEC) O104:H4 infection in Germany causes a paradigm shift with regard to human pathogenicity of STEC Strains. J. Food Prot. **75**: 408–418.
9. **Grad YH, Lipsitch M, Feldgarden M, Arachchi HM, Cerqueira GC, Gofrey P, Hass BJ, Murphy CI, Russ C, Sykes S, Walker BJ, Wortman JR, Young S, Zeng Q, Hanage WP, Hung DT, Berren BW, Nasbaum C, Lander ES.** 2011. Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe. Proc. Natl. Acad. Sci. U. S. A. **109**:3065–3070.
10. **Karch H.** 2012. The enemy within us: lessons from the 2011 European *Escherichia coli* O104:H4 outbreak. EMBO Mol. Med. **4**:841–848.
11. **Sin MA, Takla A, Flieger A, Prager R, Fruth A, Tietze E, Fink E, Korte J, Schink S, Höhle M, Eckmanns T.** 2013. Carrier prevalence, secondary household transmission, and long-term shedding in 2 districts during the *Escherichia coli* O104:H4 outbreak in Germany. J. Infect. Dis. **207**:-432–438.
12. **Urdahl AM, Solheim HT, Vold L, Hasseltvedt V, Wasteson Y.** 17 September 2012. Shiga toxin-encoding genes (*stx* genes) in human faecal samples. APMIS [Epub ahead of print.] http://dx.doi.org/10.1111/j.1600-0463.2012.02957.x.

The views expressed in this article do not necessarily reflect the views of the journal or of ASM.