BRIEF REPORT

Open-Source Genomic Analysis of Shiga-Toxin–Producing *E. coli* O104:H4

Holger Rohde, M.D., Junjie Qin, Ph.D., Yujun Cui, Ph.D., Dongfang Li, M.E., Nicholas J. Loman, M.B., B.S., Moritz Hentschke, M.D., Wentong Chen, B.S., Fei Pu, B.S., Yangqing Peng, B.S., Junhua Li, B.E., Feng Xi, B.E., Shenghui Li, B.S., Yin Li, B.S., Zhaoxi Zhang, B.S., Xianwei Yang, B.S., Meiru Zhao, M.S., Peng Wang, B.M., Yuanlin Guan, B.E., Zhong Cen, M.E., Xiangna Zhao, B.S., Martin Christner, M.D., Robin Kobbe, M.D., Sebastian Loos, M.D., Jun Oh, M.D., Liang Yang, Ph.D., Antoine Danchin, Ph.D., George F. Gao, Ph.D., Yajun Song, Ph.D., Yingrui Li, B.S., Huanming Yang, Ph.D., Jian Wang, Ph.D., Jianguo Xu, M.D., Ph.D., Mark J. Pallen, M.D., Ph.D., Jun Wang, Ph.D., Martin Aepfelbacher, M.D., Ruifu Yang, M.D., Ph.D., and the *E. coli* O104:H4 Genome Analysis Crowd-Sourcing Consortium*

SUMMARY

An outbreak caused by Shiga-toxin–producing *Escherichia coli* O104:H4 occurred in Germany in May and June of 2011, with more than 3000 persons infected. Here, we report a cluster of cases associated with a single family and describe an open-source genomic analysis of an isolate from one member of the family. This analysis involved the use of rapid, bench-top DNA sequencing technology, open-source data release, and prompt crowd-sourced analyses. In less than a week, these studies revealed that the outbreak strain belonged to an enteroaggregative *E. coli* lineage that had acquired genes for Shiga toxin 2 and for antibiotic resistance.

The authors' affiliations are listed in the Appendix. Address reprint requests to Dr. Pallen at the Centre for Systems Biology, University of Birmingham, B15 2TT, United Kingdom, or at m.pallen@bham.ac.uk; or to Dr. Yang at BGI-Shenzhen, Shenzhen 518083, China, or at the Beijing Institute of Microbiology and Epidemiology, 20 Dongda St., Beijing 100071, China, or at yangruifu@genomics.org.cn.

The following two groups of authors contributed equally to this article: Drs. Rohde, Qin, Cui, D. Li, and Loman; and Drs. Pallen, J. Wang, Aepfelbacher, and R. Yang.

*Members of the *E. coli* O104:H4 Genome Analysis Crowd-Sourcing Consortium are listed in the Supplementary Appendix, available at NEJM.org.

This article (10.1056/NEJMoa1107643) was published on July 27, 2011, at NEJM.org.

N Engl J Med 2011;365:718-24.

Escherichia coli is a widespread commensal of the mammalian gut and a versatile pathogen.[1,2] Enterovirulent strains of *E. coli* are classified into a number of overlapping pathotypes, which include Shiga-toxin–producing, enterohemorrhagic, and enteroaggregative varieties.[2] Enteroaggregative *E. coli* strains have been associated with sporadic and epidemic diarrhea and, in the laboratory, show a distinctive pattern of adherence to Hep-2 cells (termed aggregative, or "stacked brick").[3] In Shiga-toxin–producing *E. coli*, the toxin is encoded on a prophage and inhibits protein synthesis within susceptible eukaryotic cells. Strains of enterohemorrhagic *E. coli* produce Shiga toxin and a specific protein secretion system (called a type III secretion system) that is encoded by the locus of enterocyte effacement (LEE) and that is responsible for attachment to the intestine.[2] Shiga-toxin–producing and enterohemorrhagic *E. coli* strains are commonly associated with the hemolytic–uremic syndrome, a combination of renal impairment, thrombocytopenia, and hemolytic anemia that is often accompanied by neurologic and myocardial damage.

More than 3000 cases of infection with an unusual strain of Shiga-toxin–producing *E. coli* O104:H4 were reported to the Robert Koch Institute in Berlin during a nationwide outbreak in Germany in May and June of 2011.[4] This outbreak resulted in more than 40 deaths, and associated cases were reported in more than a dozen countries in Europe and North America (mostly in travelers returning from Germany). Household transmission was described in the Netherlands, and life-threatening colonic ischemia was reported as a complication in addition to the hemolytic–uremic syndrome and bloody diarrhea.[5,6] Epidemiologic and microbiologic evidence indicated that the O104:H4 strain was distributed throughout Germany on bean sprouts.[7]

The outbreak was characterized by several unusual features: a high incidence in adults (especially women), a greatly increased incidence of the hemolytic–uremic syndrome (in approximately 25% of patients, as compared with 1 to 15% in previous outbreaks of Shiga-toxin–producing *E. coli*), a predominance of female patients among cases of the hemolytic–uremic syndrome, and a rare serotype of Shiga-toxin–producing *E. coli* that had been linked to only two sporadic cases of the hemolytic–uremic syndrome (one in Germany and the other in South Korea).[4,8,9] Recognition of infection during the outbreak was hampered by a laboratory approach that targeted phenotypes associated with the most common lineage of enterohemorrhagic *E. coli* (the non–sorbitol-fermenting O157:H7 serotype) rather than one aimed at finding all strains of Shiga-toxin–producing *E. coli*.[10] Here, we report a local cluster of cases associated with a family from northern Germany and describe an open-source genomic analysis of an isolate from the family cluster.

CASE REPORTS

On May 17, 2011, a 16-year-old girl was admitted to the pediatric emergency ward at the University Medical Center Hamburg–Eppendorf with bloody diarrhea and abdominal pain. Her laboratory values were normal. Later on the same day, her 12-year-old brother was admitted with a 2-day history of malaise and headache and a 1-day history of vomiting and nonbloody diarrhea. The boy presented with acute renal failure (serum creatinine level, 4.1 mg per deciliter [362 μmol per liter]; and potassium level, 6 mmol per liter), thrombocytopenia (22,000 platelets per cubic millimeter), and hemolytic anemia (hemoglobin, 11.6 g per deciliter; bilirubin, 2.8 mg per deciliter [49 μmol per liter]; and lactate dehydrogenase, 2297 U per liter). His hemoglobin level fell to 8.4 g per deciliter within 48 hours after admission, thereby fulfilling the case definition of the hemolytic–uremic syndrome.

The children, their parents, and a teenage friend had eaten a meal together a week earlier. The meal included a freshly prepared salad containing bean sprouts. The children's mother had no symptoms, and no Shiga-toxin–producing *E. coli* was isolated from her stool. However, the hemolytic–uremic syndrome developed in the father, and his stool sample was culture-positive for Shiga-toxin–producing *E. coli*. The teenage friend had diarrhea but was not admitted to the medical center.

Stool samples from the siblings were plated on Sorbitol–MacConkey agar and incubated in a liquid enrichment culture. The next day, supernatants from the liquid cultures tested positive for Shiga toxin on enzyme-linked immunosorbent assay. Uniformly sorbitol-positive colonies were identified as *E. coli* on MALDI-TOF (matrix-assisted laser desorption ionization–time of flight) mass spectrometry. Several single colonies were positive for the *stx2* gene and negative for the *stx1* and *eae* genes on polymerase-chain-reaction (PCR) assay. None of the isolates agglutinated with polyvalent serum samples directed against the serotypes that are most frequently associated with Shiga-toxin–producing *E. coli*. Subsequent analyses showed that the strain belonged to the rare serotype O104:H4 harboring an extended-spectrum beta-lactamase (ESBL) gene of the CTX-M-15 class.

Although our 16-year-old patient had a mild course of disease without the hemolytic–uremic syndrome and was discharged from the hospital on the same day, the clinical picture for her brother was much less benign. The boy's renal function, hemoglobin level, and thrombocytopenia improved after 9 days of peritoneal dialysis, but severe neurologic symptoms, including somnolence, visual impairment, speech disturbances, hemiplegia, and incontinence, developed. He underwent four cycles of plasmapheresis and therapy with the anti–C5-antibody eculizumab. After this treatment, his clinical condition improved, and he was discharged after 24 days with serum creatinine levels just above the normal range. However, he was left with neurologic sequelae and required rehabilitation.

METHODS AND RESULTS

OPEN-SOURCE GENOMICS

To investigate the evolutionary origins and pathogenic potential of the outbreak strain, we set in motion an open-source genomics program of research that incorporated new high-throughput sequencing approaches, public data release, and rapid outsourcing of analyses to bioinformaticians worldwide (crowd-sourcing) (Fig. 1). Initially, we sequenced the genome of the isolate from the 16-year-old girl (TY2482), using the Ion Torrent Personal Genome Machine (PGM), and obtained an initial draft of the genome 3 days after receipt of the DNA sample. Three DNA libraries were prepared and seven sequencing runs performed, following the protocols of the manufacturer (Life Technologies), to generate 79 Mb of sequence data, with an average read length of 101 bp. (For details regarding the sequencing procedures, see the Supplementary Appendix, available with the full text of this article at NEJM.org.)

We released these data into the public domain under a Creative Commons 0 license, which elicited a burst of crowd-sourced, curiosity-driven analyses carried out by bioinformaticians on four continents.[11] Twenty-four hours after the release of the genome, it had been assembled; 2 days after its dissemination, it had been assigned to an existing sequence type. Five days after the release of the sequence data, we had designed and released strain-specific diagnostic primer sequences, and within a week, two dozen reports had been filed on an open-source wiki (a Web site that facilitates collaborative effort) dedicated to analysis of the strain. These analyses provided timely information on the strain's virulence and resistance genes, in addition to its phylogenetic lineage.

We also performed sequencing on the Illumina HiSeq platform in accordance with the manufacturer's instructions. An initial single-end run was used to correct errors in the Ion Torrent sequence, principally in homopolymeric tracts. We later performed paired-end and mate-pair sequencing on this platform, exploiting libraries with insert sizes of 470 bp, 2 kb, and 6 kb, and generated enough data (1 Gb, 576 Mb, and 576 Mb from each library, respectively) to create a high-quality draft genome sequence within 2 weeks after receipt of the DNA samples. (Additional details are provided in the Supplementary Appendix.) The reads were deposited in GenBank's Short Read Archive with accession numbers SRA037315 for Ion Torrent reads and SRA039136 for Illumina platform reads.

PHYLOGENETIC ANALYSIS

The assembled Ion Torrent data provided gene sequences that could be analyzed with an existing multilocus-sequence-typing scheme for *E. coli* that relied on sequence comparisons for seven conserved housekeeping genes (*adk*, *fumC*, *gyrB*, *mdh*, *purA*, *recA*, and *icd*).[12] This analysis revealed a close relationship to a strain, 01-09591, which was isolated in Germany in 2001 and which fell into sequence type ST678. The TY2482 sequences differed from the profile of the 2001 strain by a single base pair in the *adk* gene and a single-base difference in a homopolymeric sequence in the *recA* gene. (We subsequently discovered that the latter difference was a sequencing error generated by the PGM.) The 2001 strain, which produced Shiga toxin and was associated with the hemolytic–uremic syndrome, fell into the O104:H4 serotype but did not have the genes associated with type III secretion in typical enterohemorrhagic *E. coli*.[13,14] Additional scrutiny of the multilocus-sequence-typing database revealed that strains with the broad O104 serotype were scattered across several sequence types, whereas strains with the narrower O104:H4 serotype appeared to be limited to ST678.[10]

Comparisons of the TY2482 genome with all previously sequenced complete genomes of *E. coli* isolates revealed a very close relationship to *E. coli* strain 55989, with an average nucleotide identity of 99.8% (see the Supplementary Appendix). This strain was isolated in the Central African Republic from a stool sample obtained from an adult with human immunodeficiency virus infection who had persistent watery diarrhea.[15] It has been classified as an enteroaggregative *E. coli*, but unlike TY2482, it does not have Shiga toxin genes.[15] However, it is worth noting that Mossoro et al.,[15] who first described *E. coli* strain 55989, also described strains of enteroaggregative *E. coli* with Shiga toxin genes in the same human population.[15]

COMPARISON OF THE CHROMOSOMES OF TY2482 AND 55989

Isolates from the German outbreak were initially described as enterohemorrhagic *E. coli*. However, the close relationship between TY2482 and 55989 led us to consider the likelihood that TY2482 is an enteroaggregative *E. coli*. Our analysis of the gene content of TY2482 showed that it, like 01-09591,

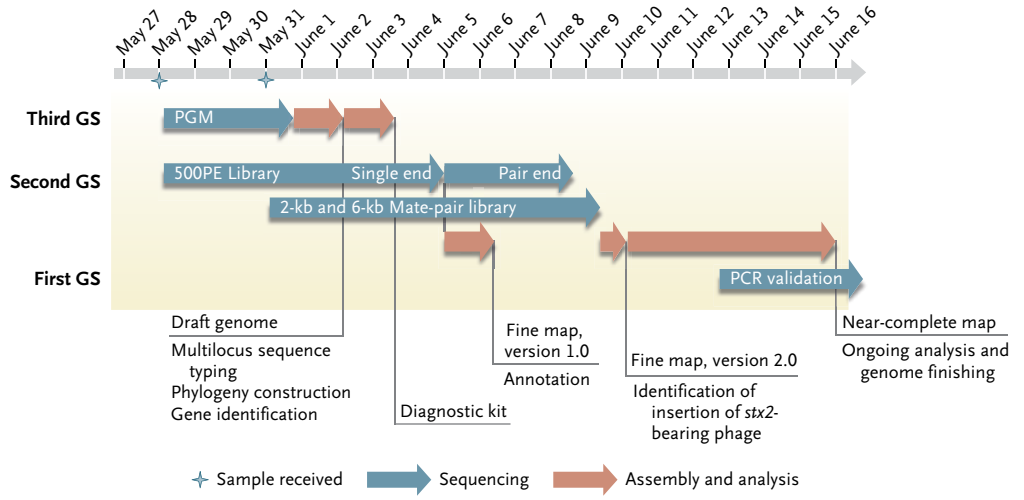

Figure 1. Timeline of the Open-Source Genomics Program.

After receiving the first batch of DNA samples on May 28, 2011, sequencing runs with the use of the Ion Torrent Personal Genome Machine (PGM) and Illumina (small-insert library) were initiated simultaneously. On May 31, the second batch of DNA was received and used for Illumina large-insert sequencing. An assembly of the Ion Torrent reads was released on June 2, which enabled subsequent analyses (multilocus sequence typing, phylogenetic analysis, and genome comparisons). Errors in the Ion Torrent data were corrected with the use of later Illumina data, and a high-quality draft genome sequence was created. GS denotes generation of sequencing technology. The symbols at May 28 and May 31 in the timeline indicate the arrival of DNA samples.

lacked the LEE and genes encoding effectors associated with type III secretion.[16] Instead, we found that the TY2482 genome encodes virulence factors that are typical of enteroaggregative *E. coli*. Other investigators working on the outbreak strain have also observed genes typically found in enteroaggregative strains on PCR assay and have noted a behavioral phenotype that is characteristic of this pathotype on cell-adherence assay.[17]

To identify strain-specific genes, we performed a detailed comparison of the chromosomes of TY2482 and enteroaggregative *E. coli* strain 55989. First, we aligned the TY2482 assembly against the 55989 chromosome (for details, see the Supplementary Appendix). We then adopted the gene predictions and annotation from the 55989 genome for these conserved sequences. Next, we identified several isolate-specific regions of difference (i.e., regions present in the TY2482 chromosome and absent from the 55989 genome or vice versa) that were more than 5 kb (Table 1 and Fig. 2, and the Supplementary Appendix). TY2482-specific regions of difference included prophage remnants or apparently intact prophages, such as the *stx2* prophage, which, like its close relatives in the genomes of O157:H7 strains EDL933 and Sakai, is inserted into the *wrbA* locus. The *stx2* genes differ by only one single-nucleotide polymorphism from the *stx2* allele seen in O157 enterohemorrhagic *E. coli* strain EDL933.

TY2482 PLASMIDS

From our de novo assembly (i.e., assembly without the use of a reference genome), we concluded that the TY2482 genome contains two large conjugative plasmids, pESBL TY2482 and pAA TY2482, and a small plasmid, pG2011 TY2482 (Fig. 2). From scrutiny of copy numbers of sequence reads, it was clear that the two large plasmids were replicating at an approximate ratio of 1:1 with the chromosome, whereas the small plasmid was maintained at a copy number at least nine times that of the other replicons. No phenotype could be ascribed to the small plasmid.

The largest plasmid, pESBL TY2482, was an IncI plasmid similar to pEC_Bactec, which was found in an *E. coli* strain isolated from the joint of a horse with arthritis.[18] The pESBL TY2482 plasmid encodes a CTX-M-15 ESBL, as well as a beta-lactamase from the TEM class. The second large

Table 1. Genetic Elements in Strain TY2482 of Shiga-Toxin–Producing *Escherichia coli* O104:H4.

Genetic Element	Notable Features or Functions	Size or 55989 Coordinates*
Plasmid		
pESBL TY2482	IncI1 plasmid, homologous to pEC_Bactec carrying *bla* CTX-M-15	88 kb
pAA TY2482	Plasmid encoding aggregative adherence fimbriae I	76 kb
pG2011 TY2482	Plasmid with no obvious phenotype	1.5 kb
Region of difference		
I-ROD1	Degenerate prophage	296227 (tRNA-*Thr*)
I-ROD2	*Stx2*-encoding prophage	1176265 (*wrbA*)
I-ROD3	Microcin gene cluster; tellurite resistance gene cluster	1207704 (tRNA-*Ser*)
I-ROD4	Prophage	1811905 (*ynfG*)
I-ROD5	Prophage	2102453 (*yecE*)
I-ROD6	Molybdate metabolism regulator; *yehL*	2426442 (IS1)
I-ROD7	Multidrug-resistant gene cluster (*dfA7, sulI, sulII, strA, strB, tetA*); mercury resistance	4211244 (tRNA-*Sec*)
D-ROD1	Prophage	1094587–1140306
D-ROD2	Prophage	1413924–1446834
D-ROD3	Prophage	1754689–1800354
D-ROD4	Prophage	2688656–2701228
D-ROD5	Type VI secretion genes	3401720–3427357
D-ROD6	Prophage	4944269–5004333

* Coordinates from the genome of *E. coli* strain 55989 are given for predicted boundaries of regions of difference, with the gene carrying the insertion site shown in parentheses for a region of difference involving an insertion into 55989 (I-ROD). D-ROD denotes a region of difference involving a deletion.

plasmid, pAA TY2482, resembled a plasmid from strain 55989 but carried a gene cluster encoding a rare type of aggregative adherence fimbria (AAF/I) instead of the more common type (AAF/III) encoded by genes in the 55989 plasmid. We exploited this AAF/I cluster as a target for strain-specific PCR primers as part of a suite of primers to identify the outbreak isolate.

DISCUSSION

Our genomic analyses suggest that the German outbreak strain evolved from a progenitor that belonged to the enteroaggregative pathotype and resembled strain 55989. The emergence of the outbreak strain depended on the acquisition of a *stx2* prophage and of a plasmid encoding a CTX-M-15 ESBL. Sometime during this process, the strain also appears to have lost one gene cluster, encoding AAF/III fimbriae, and gained another, encoding the rarer AAF/I fimbriae.

Although this outbreak strain has surprised the general public and public health officials, related potential progenitor strains have been reported from three continents. The appearance of an O104:H4 strain associated with the hemolytic–uremic syndrome in Korea in 2005 is unexplained, and its link to the German outbreak is unclear.[9] Also, the O104:H4 strain 01-09591 that was isolated in Germany in 2001 urgently requires further investigation. Both strains should undergo genome sequencing and comparison with TY2482. The link to strain 55989, which was isolated in the Central African Republic in the late 1990s, is also intriguing. Genome sequencing of additional Central African isolates from the study that yielded 55989 is likely to illuminate the evolution of this lineage and of enterovirulent *E. coli* in general (see the article by Rasko et al. elsewhere in this issue of the *Journal*[19]).

Although the genome sequence alone cannot provide a full explanation for the high degree of virulence of this strain, it prompts a reassessment of our assumptions and provides a framework for future hypothesis-driven research. Both commensal and pathogenic varieties of *E. coli* have to survive

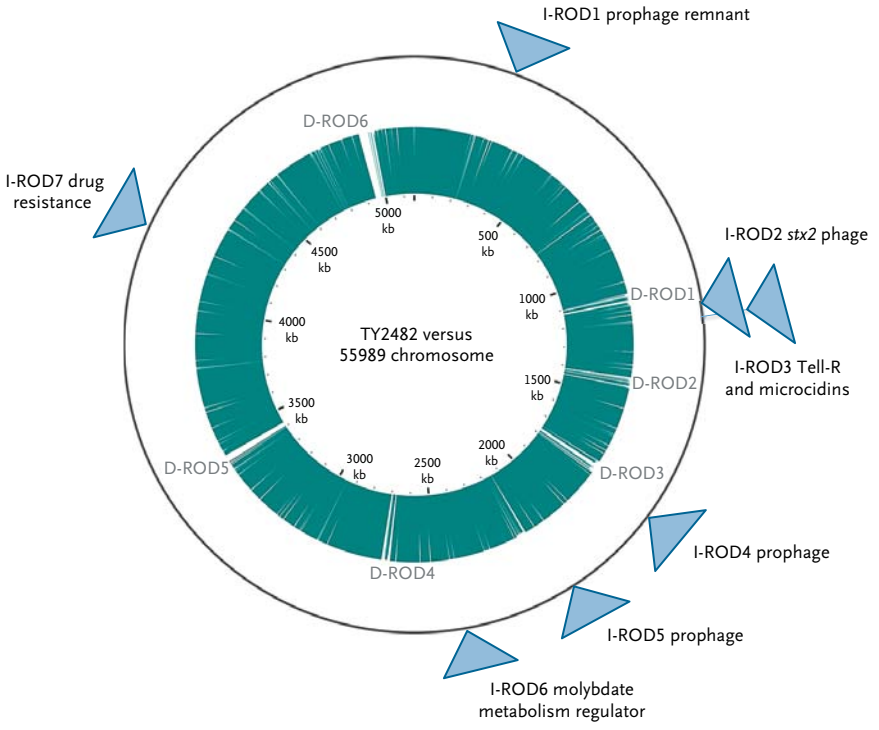

Figure 2. Comparison of the TY2482 and 55989 Genomes.

The outer circle depicts the *Escherichia coli* 55989 chromosome. The inner circle represents the TY2482 assembly mapped against the 55989 chromosome. Regions of difference (ROD) that are present in 55989 but not TY2482 (D-RODs) are shown as gaps in the inner circle. The positions of RODs that were found in TY2482 but not 55989 (I-RODs) are shown as wedges on the outer circle at positions corresponding to the predicted insertion sites. Tell-R denotes tellurite resistance.

in the gut. However, mere survival, even if twinned with the production of Shiga toxin, is probably not enough to cause the hemolytic–uremic syndrome or bloody diarrhea. For that, the bacteria would probably need to adhere to the gut mucosa. In the past, much research has been concentrated on the adhesion systems of typical enterohemorrhagic *E. coli*, particularly the LEE-encoded type III secretion system.[16,20] This German outbreak strain shows us that Shiga-toxin–producing *E. coli* can exploit alternative adhesion mechanisms, very likely including aggregative adherence fimbriae, to the same end. This strain also shows that pathotypes of *E. coli* can overlap and that they evolve rather than stand as fixed archetypes.

It remains unclear why this strain has proved to be so virulent. As noted, a novel suite of adhesins might provide an explanation. Alternatively, perhaps this strain exploits more efficient mechanisms for toxin release. It is worth remembering that strains of enteroaggregative *E. coli* have caused large sprout-associated outbreaks before, including one outbreak[21] that affected more than 2000 persons in Japan in 1993. Thus, there is clearly an urgent need to understand how the German outbreak strain and other strains of enteroaggregative *E. coli* adhere to and colonize seeds and seedlings.

Our rapid open-source analysis of an outbreak-associated bacterial pathogen was characterized by a propitious confluence of high-throughput genomics, crowd-sourced analyses, and a liberal approach to data release. Although phenotypic or molecular analyses that exploit known virulence, resistance, or epidemiologic targets are useful in diagnostic and public health microbiology, genome sequencing offers the advantages of open-endedness (revealing the "unknown unknowns"), universal applicability, and the ultimate in resolution. Our study shows how benchtop sequencing platforms can generate data with sufficient speed to have an important effect on clinical and epidemiologic problems.

Supported by grants from the State Key Development Program for Basic Research of China (2009CB522600), the National Key Program for Infectious Diseases of China (2008ZX10004-009), Shenzhen Biological Industry Development Special Foundation–Basic Research Key Projects (JC201005250088A), Key Laboratory Project Supported by Shenzhen City (ZD200806180054A), the European Union Microme Program (FP7-KBBE-2007-3-2-08-222886), the Alexander von Humboldt Foundation (to Dr. L. Yang), the Medical Faculty of the University Medical Center Hamburg–Eppendorf, and the British Biotechnology and Biological Sciences Research Council (BB/E011179/1).

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank David Vallenet, Claudine Médigue, Xiaoning Wang, and Jennifer Gardy for their helpful discussions.

APPENDIX

The authors' affiliations are as follows: the Institute of Medical Microbiology, Virology and Hygiene (H.R., M.H., M.C., L.Y., M.A.) and the Department of Pediatrics (R.K., S. Loos, J.O.), University Medical Center Hamburg–Eppendorf, Hamburg, Germany; BGI-Shenzhen, Shenzhen, China (J.Q., Y.C., D.L., W.C., F.P., Y.P., J.L., F.X., S. Li, Yin Li, Z.Z., X.Y., M.Z., P.W., Y.G., Z.C., Yingrui Li, H.Y., Jian Wang, Jun Wang, R.Y.), and the School of Bioscience and Biotechnology, South China University of Technology, Guangzhou, China (J.L.); the State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology (Y.C., X.Z., Y.S., R.Y.), CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Science (G.F.G.), and the State Key Laboratory for Infectious Disease Prevention and Control and National Institute for Communicable Diseases Control and Prevention, Chinese Center for Disease Control and Prevention (G.F.G., J.X.) — all in Beijing; the Centre for Systems Biology, University of Birmingham, Birmingham, United Kingdom (N.J.L. M.J.P.); and AMAbiotics, Evry, France (A.D.).

REFERENCES

1. Hobman JL, Penn CW, Pallen MJ. Laboratory strains of Escherichia coli: model citizens or deceitful delinquents growing old disgracefully? Mol Microbiol 2007; 64:881-5.

2. Kaper JB, Nataro JP, Mobley HL. Pathogenic Escherichia coli. Nat Rev Microbiol 2004;2:123-40.

3. Weintraub A. Enteroaggregative Escherichia coli: epidemiology, virulence and detection. J Med Microbiol 2007;56:4-8.

4. Frank C, Werber D, Cramer JP, et al. Epidemic profile of Shiga-toxin–producing *Escherichia coli* O104:H4 outbreak in Germany — preliminary report. N Engl J Med 2011. DOI: 10.1056/NEJMoa1106483.

5. Kuijper E, Soonawala D, Vermont C, van Dissel J. Household transmission of haemolytic uraemic syndrome associated with Escherichia coli O104:H4 in the Netherlands. Euro Surveill 2011;16:pii: 19897.

6. Cordesmeyer S, Peitz U, Godde N, Kasper H, Hoffmann M, Allemeyer E. Colonic ischaemia as a severe Shiga toxin/verotoxin producing Escherichia coli O104:H4 complication in a patient without haemolytic uraemic syndrome, Germany, June 2011. Euro Surveill 2011;16: pii:19895.

7. Robert Koch Institute. Information update on EHEC/HUS outbreak, 2011. (http://www.rki.de/nn_217400/EN/Home/PM082011.html.)

8. Scheutz F, Moller Nielsen E, Frimodt-Moller J, et al. Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing Escherichia coli O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. Euro Surveill 2011;16:pii:19889.

9. Bae WK, Lee YK, Cho MS, et al. A case of hemolytic uremic syndrome caused by Escherichia coli O104:H4. Yonsei Med J 2006;47:437-9.

10. Chattaway MA, Dallman T, Okeke IN, Wain J. Enteroaggregative E. coli O104 from an outbreak of HUS in Germany 2011, could it happen again? J Infect Dev Ctries 2011;5:425-36.

11. GitHub. *E. coli* O104:H4 genome analysis crowd sourcing, 2011. (https://github.com/ehec-outbreak-crowdsourced/BGI-data-analysis/wiki.)

12. University College Cork. *Escherichia coli* MLST Database, 2011. (http://mlst.ucc.ie/mlst/dbs/Ecoli.)

13. Mellmann A, Bielaszewska M, Kock R, et al. Analysis of collection of hemolytic uremic syndrome-associated enterohemorrhagic Escherichia coli. Emerg Infect Dis 2008;14:1287-90.

14. Creuzburg K, Middendorf B, Mellmann A, et al. Evolutionary analysis and distribution of type III effector genes in pathogenic Escherichia coli from human, animal and food sources. Environ Microbiol 2011;13:439-52.

15. Mossoro C, Glaziou P, Yassibanda S, et al. Chronic diarrhea, hemorrhagic colitis, and hemolytic-uremic syndrome associated with HEp-2 adherent Escherichia coli in adults infected with human immunodeficiency virus in Bangui, Central African Republic. J Clin Microbiol 2002;40:3086-8.

16. Tobe T, Beatson SA, Taniguchi H, et al. An extensive repertoire of type III secretion effectors in Escherichia coli O157 and the role of lambdoid phages in their dissemination. Proc Natl Acad Sci U S A 2006;103:14941-6.

17. Bielaszewska M, Mellmann A, Zhang W, et al. Characterisation of the Escherichia coli strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. Lancet Infect Dis 2011 June 22 (Epub ahead of print).

18. Smet A, Van Nieuwerburgh FV, Vandekerckhove TTM, et al. Complete nucleotide sequence of CTX-M-15-plasmids from clinical Escherichia coli isolates: insertional events of transposons and insertion sequences. PLoS ONE 2010;5(6):e11202.

19. Rasko DA, Webster DR, Sahl JW, et al. Origins of the *E. coli* strain causing an outbreak of hemolytic–uremic syndrome in Germany. N Engl J Med 2011;365:709-17.

20. Schmidt MA. LEEways: tales of EPEC, ATEC and EHEC. Cell Microbiol 2010;12: 1544-52.

21. Itoh Y, Nagano I, Kunishima M, Ezaki T. Laboratory investigation of enteroaggregative Escherichia coli O untypeable:H10 associated with a massive outbreak of gastrointestinal illness. J Clin Microbiol 1997;35:2546-50.