

Abstract

Background

In October 2010, the first cholera outbreak in over a century began in Haiti, a country on the western peninsula of the Caribbean island of Hispaniola. Sixteen specimens were sent to the CDC for isolation and confirmation of *Vibrio cholerae*. Thirteen isolates were recovered and characterized as toxigenic *V. cholerae* serogroup O1, serotype Ogawa, biotype El Tor. The hypothesis was that Optical Mapping could accurately differentiate *V. cholerae* and subsequently be used to order, orient, and validate whole genome sequencing (WGS) contigs of the isolates.

Methods

Three *V. cholerae* isolates from Haiti were selected for WGS analysis. The Biotechnology Core Facility at the CDC constructed libraries for the 454 (Roche) and Illumina sequencing (Illumina) platforms, which allowed sequencing to be carried out simultaneously on both devices. Additionally, non-Haitian outbreak isolates from a wide geographic distribution were also selected and sent for sequencing for comparison. To assist in ordering and orienting contigs, 9 isolates were also submitted to OpGen for Optical Mapping.

Results

Optical Maps of 9 isolates were generated in <4 days from colony to Optical Map and the relationship between isolates was consistent with PFGE data, thereby demonstrating the compatibility of Optical Maps with epidemiologically relevant data. The CDC Core Facility completed WGS on 13 isolates in 4 weeks. Optical Maps were used as a scaffold to guide the assembly of contigs for each strain into pseudomolecules and were instrumental in identifying the regions of the genome that needed to be sequenced to close gaps between contigs.

Conclusions

The advent of high-throughput, rapid sequencing methods has allowed the utility of WGS data to be explored in the midst of an ongoing outbreak. The data support the hypothesis that Optical Mapping can differentiate isolates and provide a sequencing independent method to improve the quality of WGS assemblies.

Strain list

CDC ID or Key	WGS	Year	Strain description	Sfil Pattern	NotI Pattern
2010EL-1786	yes	2010	O1 Ogawa, El tor, typical Haiti strain, Haiti	KZGS12.0088	KZGN11.0092
2009V-1096	yes	2009	O1 Ogawa, El tor, VSP-1 variant, travel to India	KZGS12.0088	KZGN11.0092
3500-05	yes	2005	O1 Inaba, El tor, ctxAB variant, travel to India	KZGS12.0088	KZGN11.0092
2009V-1116	yes	2009	O1 Ogawa, El tor, ctxAB variant, travel to Pakistan	KZGS12.0088	KZGN11.0092
2010EL-1749	yes	2010	O1 Ogawa, PFGE match, Cameroon	KZGS12.0088	KZGN11.0092
C6706	yes	1991	O1 Inaba, typical Latin American strain, Peru	KZGS12.0114	KZGN11.0033
2010V-1031	yes	2010	O75, pansusceptible strain	KZGS12.0073	KZGN11.0084
3569-08	yes	2008	O1 Inaba, El Tor, typical Gulf Coast strain	KZGS12.0055	KZGN11.0029
3566-08	yes	2008	O141, pansusceptible strain	KZGS12.0098	KZGN11.0086

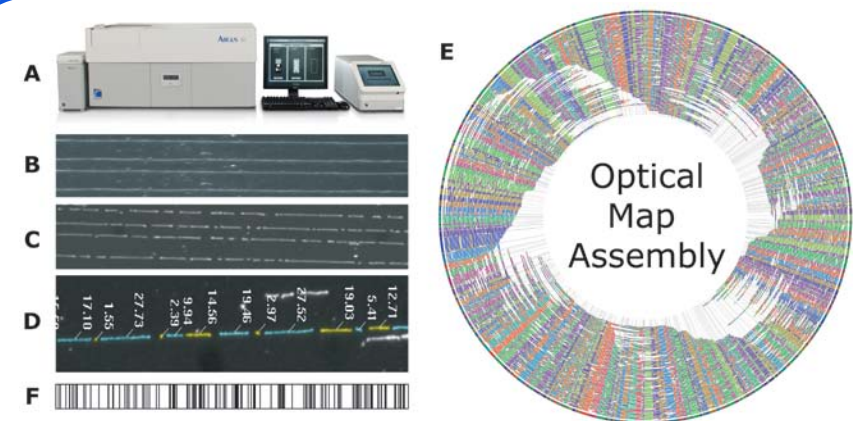


Figure 1. Argus Optical Mapping System (A). Single DNA molecules captured onto surface (B). Digested DNA molecules (C). Single molecule restriction maps (D). Whole-genome Optical Map Assembly (E). The consensus Optical Map (F).

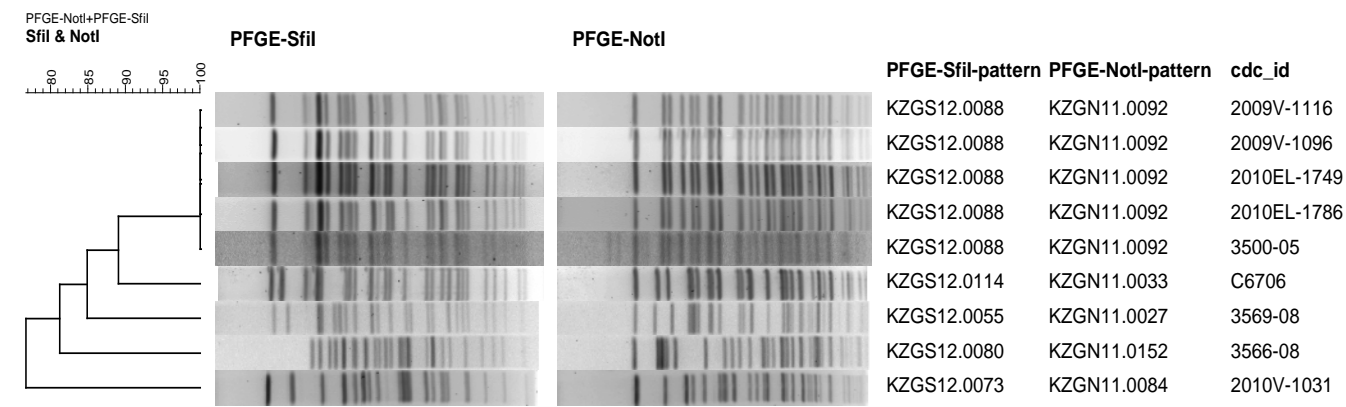


Figure 2. Dendrogram showing relationship between *V. cholerae* strains chosen for optical mapping and WGS based on *Sfil* and *NotI* PFGE patterns. A typical Haitian strain (2010EL-1786) is indistinguishable from a strain circulating in Cameroon (2010EL-1749) and three strains collected from U.S. travelers to South Asia (2009V-1096, 3500-05 and 2009V-1116). In contrast, non-O1 strains (2010V-1031 and 3566-08) and typical Latin American and Gulf Coast strains (C6706 and 3569-08) differ greatly.



Figure 3. Optical Maps of chromosomes 1 & 2 of 9 *Vibrio cholerae* strains generated with *NheI*. Dendrograms were generated using MapSolver and a developmental Map Distance algorithm that weighs genetic events like insertions, deletions, insertions and translocations. The default parameters were used except that Cut Mismatch Weight was changed to 0 and Minimum Indel Size was changed to 3.5 kb. Optical Mapping differentiates the Cameroon strain 2010EL-1749 from the clonal group by a ~6 kb relative insertion in chromosome 1.

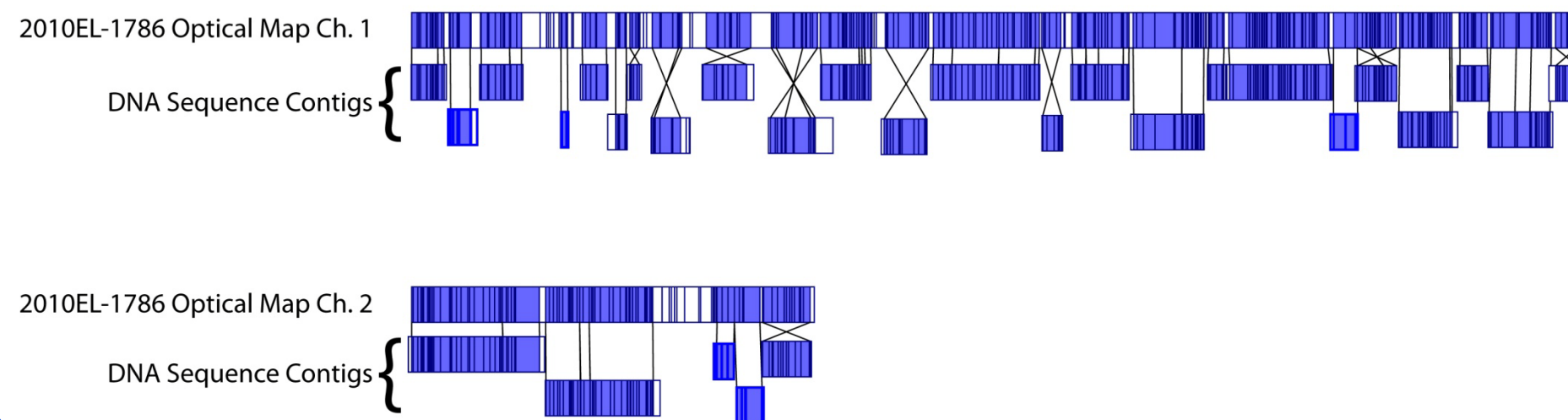


Figure 4. DNA sequence contigs were converted to *in silico* maps and aligned to the Optical Map scaffolds using MapSolver software to determine order, orientation, and validate the sequence assemblies.

Results

- Five strains collected over 5 years and 4 countries were indistinguishable by *Sfil* and *NotI* PFGE pattern combinations (KZGS12.0088/KZGN11.0092).
- Four strains were indistinguishable by *NheI* Optical Mapping.
- Strain 2010EL-1749 contained a ~6 kb insertion identified by Optical Mapping that was not present in 2009V-116, 2009V-1096, 2010EL-1786, or 3600-05.
- Outlier strains (non-O1, Latin American and Gulf Coast) were also significantly different based on optical maps.
- The DNA sequence contigs placed by MapSolver covered 88% of the 2010EL-1786 genome.
- Optical Mapping identified 4 contig gaps in project 2010EV-1786 where the contigs are estimated to overlap.
- Optical Mapping identified ~36% of the gaps between placed contigs in project 2010EV-1786 that were less than 2 kb.

Conclusions

- Optical Mapping can detect differences in strains that show identical PFGE patterns.
- Optical Mapping distinguished the Cameroon 2010EL-1749 strain from the other strains containing the same PFGE pattern by identifying a ~6 kb insertion in chromosome 1.
- Relationships between *V. cholerae* strains based on PFGE patterns and Optical Maps are in agreement, suggesting Optical Map data could be epidemiologically useful.
- Optical Mapping is a higher resolution technology that can be used to differentiate closely related strains that PFGE can not.
- The ease and speed of generating Optical Maps suggest the same Optical Map data used for differentiating strains could also be ideal for assisting with efficient whole-genome sequence assembly and validation.

Contact Information

Centers for Disease Control and Prevention
PFGE Reference Unit
1600 Clifton Rd. MSC03
Atlanta, GA 30329

Molly Freeman: MMFreeman@cdc.gov
Trevor Wagner: twagner@opgen.com
Efrain Ribot: EMRibot@cdc.gov
Mike Frace: MFrace@cdc.gov

