Journal of Virology

Draft Genome Sequence of the Coccolithovirus Emiliania huxleyi Virus 203

Jozef I. Nissimov, Charlotte A. Worthy, Paul Rooks, Johnathan A. Napier, Susan A. Kimmance, Matthew R. Henn, Hiroyuki Ogata and Michael J. Allen

J. Virol. **2011, 85(24):13468. DOI: 10.1128/JVI.06440-11.**

JOURNAL OF VIROLOGY, Dec. 2011, p. 13468–13469 Vol. 85, No. 24
0022-538X/11/$12.00 doi:10.1128/JVI.06440-11

GENOME ANNOUNCEMENTS

Draft Genome Sequence of the Coccolithovirus *Emiliania huxleyi* Virus 203

Jozef I. Nissimov,[1] Charlotte A. Worthy,[1,2] Paul Rooks,[1] Johnathan A. Napier,[2] Susan A. Kimmance,[1] Matthew R. Henn,[3] Hiroyuki Ogata,[4] and Michael J. Allen[1]*

Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth PL1 3DH, United Kingdom[1]; Department of Biological Chemistry, Rothamsted Research, Harpenden, Herts AL5 2JQ, United Kingdom[2]; The Broad Institute of MIT and Harvard, Cambridge, Massachusetts[3]; and Structural and Genomic Information Laboratory, CNRS-UPR2589, Mediterranean Institute of Microbiology (IFR-88), Aix-Marseille University, 163 avenue de Luminy Case 934, FR-13288 Marseille, France[4]

Received 30 September 2011/Accepted 3 October 2011

The *Coccolithoviridae* are a recently discovered group of viruses that infect the marine coccolithophorid *Emiliania huxleyi*. *Emiliania huxleyi* virus 203 (EhV-203) has a 160- to 180-nm-diameter icosahedral structure and a genome of approximately 400 kbp, consisting of 464 coding sequences (CDSs). Here we describe the genomic features of EhV-203 together with a draft genome sequence and its annotation, highlighting the homology and heterogeneity of this genome in comparison with the EhV-86 reference genome.

Coccolithoviruses infect the coccolithophore *Emiliania huxleyi*, a cosmopolitan marine microalga which forms blooms that can cover up to 100,000 km^2 (10). Coccolithoviruses are a major cause of bloom termination, and their role in global biogeochemical cycling is gaining increasing attention (2). Coccolithovirus abundances can reach 10^7 ml^{-1} in natural seawater under bloom conditions and 10^8 to 10^9 ml^{-1} under laboratory culture (8). *Emiliania huxleyi* virus 86 (EhV-86), the model virus, was isolated in 1999 from the English Channel (50°13.79′N, 04°9.59′W), and was sequenced in its entirety in 2005 (9, 10). *Emiliania huxleyi* virus 203 (EhV-203) was isolated from the English Channel (50°00.36′N, 04°18.87′W) from a depth of 15 m on 27 July 2001 (1, 10). The icosahedral EhV-203 virion structure and morphology are similar to those of other coccolithoviruses and phycodnaviruses in general (11). Phylogenetic analysis of available major capsid protein (MCP) gene sequences indicates that the closest relatives to EhV-203 are EhV-201, EhV-202, and EhV-207 (1, 10).

EhV-203 genome sequencing, finishing, and annotation were performed by the Broad Institute. The genome was sequenced using the 454 FLX pyrosequencing technology platform (Roche/454, Branford, CT). Library construction and sequencing were performed as previously described (4). General protocols for library construction can be found at www.broadinstitute.org/annotation/viral/Phage/Protocols.html. *De novo* genome assembly of the resulting reads was performed using the Newbler v2.3 assembly software package (4). A total of 71,325 reads were produced and assembled into 7 contigs comprising 400,520 bp with a maximum contig length of 142,770 bp, average contig length of 57,131 bp, and an average coverage of approximately 40. Genes were identified using the Broad Institute's Automated Phage Annotation Protocol (4). Additional gene prediction analysis and functional annotation were performed within the Integrated Microbial Genomes—Expert Review (IMG-ER) platform (5).

General features of the EhV-203 genome sequence include a nucleotide composition of 40.12% G+C, a total of 464 predicted protein coding sequences (CDSs), and six tRNA genes (Arg, Asn, Gln, Glu, Leu, and Lys). Of the 464 CDSs annotated by IMG-ER, 91 have been annotated with functional product predictions. A total of 412 CDSs have homologues (>20% identity) within the EhV-86 genome (31 are 100% identical). Of the 52 CDSs unique to EhV-203, the majority are of unknown function, although two CDSs have homology to glycosyltranferases and one to a zinc finger protein.

Among the EhV-203 CDSs displaying the lowest similarity to their EhV-86 homologues are a lectin protein (63.62% identity) and two endonucleases (43.24% and 87.38% identity). The CDSs with the highest similarity (100% identity) include predicted DNA-directed RNA polymerase, DNA-binding protein, and transcription factor S-II and ERV1/ALR family proteins (9). EhV-203 encodes the same sphingolipid LCB biosynthetic machinery as EhV-86, with homologues for a sterol desaturase, serine palmitoyltransferase, transmembrane fatty acid elongation protein, lipid phosphate phosphatase, and two desaturases (3, 6, 7, 9). Like EhV-86, EhV-203 also lacks the critical sphingolipid LCB biosynthetic activity 3-ketosphinganine reductase (6). Further sequencing of related strains in the future will, no doubt, reveal more about the genetic and functional diversity of these environmentally important viruses.

* Corresponding author. Mailing address: Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth PL1 3DH, United Kingdom. Phone: 44 123 456 7890. Fax: 44 1752 633 101. E-mail: mija@pml.ac.uk.

Nucleotide sequence accession number. The nucleotide sequence for the draft genome sequence has been deposited in GenBank under accession no. JF974291.

This research was funded in part by the Gordon and Betty Moore Foundation through a grant to the Broad Institute (M.R.H.) and through the NERC Oceans 2025 program (M.J.A.). Sample G4014 was sequenced, assembled, and annotated at the Broad Institute. J.I.N. is supported by an NERC studentship, and C.W. is supported by a BBSRC Industrial CASE studentship sponsored by PML Applications. H.O. is supported by IGS/CNRS and ANR (grant no. ANR-09-PCS-GENM-218 and ANR-08-BDVA-003).

We thank Konstantinos Mavromatis from JGI who assisted with information regarding the IMG/ER platform and the Broad Institute's Genome Sequencing Platform, Finishing Team, and Annotation Team for their efforts to generate the genomic data. Jean Devonshire and the Centre for Bioimaging at Rothamsted provided technical support for transmission electron microscopy.

REFERENCES

1. **Allen, M. J., J. Martinez-Martinez, D. C. Schroeder, P. J. Somerfield, and W. H. Wilson.** 2007. Use of microarrays to assess viral diversity: from genotype to phenotype. Environ. Microbiol. **9:**971–982.
2. **Coyle, K. O., and A. I. Pinchuk.** 2002. Climate-related differences in zooplankton density and growth on the inner shelf of the southeastern Bering Sea. Prog. Oceanogr. **55:**177–194.
3. **Han, G. S., et al.** 2006. Expression of a novel marine viral single-chain serine palmitoyltransferase and construction of yeast and mammalian single-chain chimera. J. Biol. Chem. **281:**39935–39942.
4. **Henn, M. R., et al.** 2010. Analysis of high-throughput sequencing and annotation strategies for phage genomes. Plos One **5:**e9083.
5. **Markowitz, V. M., et al.** 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25:**2271–2278.
6. **Michaelson, L. V., T. M. Dunn, and J. A. Napier.** 2010. Viral trans-dominant manipulation of algal sphingolipids. Trends Plant Sci. **15:**651–655.
7. **Monier, A., et al.** 2009. Horizontal gene transfer of an entire metabolic pathway between a eukaryotic alga and its DNA virus. Genome Res. **19:** 1441–1449.
8. **Schroeder, D. C., J. Oke, M. Hall, G. Malin, and W. H. Wilson.** 2003. Virus succession observed during an Emiliania huxleyi bloom. Appl. Environ. Microbiol. **69:**2484–2490.
9. **Wilson, W. H., et al.** 2005. Complete genome sequence and lytic phase transcription profile of a Coccolithovirus. Science **309:**1090–1092.
10. **Wilson, W. H., et al.** 2002. Isolation of viruses responsible for the demise of an Emiliania huxleyi bloom in the English Channel. J. Mar. Biol. Assoc. U. K. **82:**369–377.
11. **Wilson, W. H., J. L. Van Etten, and M. J. Allen.** 2009. The Phycodnaviridae: the story of how tiny giants rule the world. Curr. Top. Microbiol. Immunol. **328:**1–42.