

Whole-Genome Sequences of Three Symbiotic *Endozoicomonas* Strains

Matthew J. Neave,^{a,b} Craig T. Michell,^a Amy Apprill,^b Christian R. Voolstra^a

Red Sea Research Center, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia^a; Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA^b

Members of the genus *Endozoicomonas* associate with a wide range of marine organisms. Here, we report on the whole-genome sequencing, assembly, and annotation of three *Endozoicomonas* type strains. These data will assist in exploring interactions between *Endozoicomonas* organisms and their hosts, and it will aid in the assembly of genomes from uncultivated *Endozoicomonas* spp.

Received 24 July 2014 Accepted 28 July 2014 Published 14 August 2014

Citation Neave MJ, Michell CT, Apprill A, Voolstra CR. 2014. Whole-genome sequences of three symbiotic *Endozoicomonas* strains. *Genome Announc.* 2(4):e00802-14. doi:10.1128/genomeA.00802-14.

Copyright © 2014 Neave et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Amy Apprill, apprill@whoi.edu, or Christian R. Voolstra, christian.voolstra@kaust.edu.sa.

Endozoicomonas spp. (*Gammaproteobacteria*; *Oceanospirillales*) are dominant members of the bacterial community associated with diverse marine invertebrates, including corals (1–6), sponges (7), gorgonians (8, 9), molluscs (10–13), and tubeworms (14), as well as a basal chordate (15). In some hosts, these bacteria have been observed intracellularly (2, 11, 13). However, despite the apparent importance of *Endozoicomonas* spp., it is not clear how they interact with their host. For example, they are the dominant bacteria in seemingly healthy animals (3, 8, 15, 16), although they have been implicated as the causal agent of disease in fish (17). Clarifying the functional capabilities of *Endozoicomonas* has been challenging because they reside in or on a host organism and can be difficult to culture (6). Only a handful of isolates are available in culture collections (7, 10, 12, 18, 19). Thus, metagenomic or single-cell analyses may be useful techniques for assessing the genomic capabilities of these bacteria; however, a lack of genetic resources hampers these approaches. To address this issue, we sequenced the genomes of three publically available *Endozoicomonas* type strains.

Endozoicomonas elysicola DSM 22380 (12) and *Endozoicomonas numazuensis* DSM 25634 (7) were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ) (Braunschweig, Germany), and *Endozoicomonas montiporae* LMG 24815 (19) was obtained from the Belgian Coordinated Collections of Microorganisms (BCCM) (Ghent, Belgium). We prepared small-insert libraries by shearing isolated DNA into 180-bp fragments and processing with the NEBNext library preparation kit (New

England BioLabs). Long libraries averaging approximately 2 kb were prepared according to the Nextera mate-pair sample preparation kit (Illumina). The small-insert libraries were sequenced using the Illumina HiSeq platform (100-bp paired-end reads), and the long mate-pair libraries were sequenced using the Illumina MiSeq platform (150-bp paired-end reads). Approximately 10 million paired-end reads were obtained for each library and each *Endozoicomonas* strain.

The small-insert reads were trimmed for quality, and the Illumina adapters were removed using Trimmomatic (20). Fragments with both surviving read pairs were then digitally normalized using the recommended protocol in khmer (21–23). The long mate-pair reads were trimmed using NextClip (24), and fragments with the junction adapter in at least one of the paired reads were used in the assembly. The small- and long-insert libraries were error corrected and assembled using the AllPaths-LG assembler (25), and the gaps in the resulting scaffolds were closed using GapFiller (26). A small number of scaffolds were further joined after a manual examination of mate-pair mappings using Circos (27). The genomes assembled into ≤ 31 scaffolds, with a scaffold N_{50} of >0.92 Mbp (Table 1). The whole-genome sequences were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/).

The *Endozoicomonas* genomes were large (>5 Mbp; Table 1) and contained versatile metabolic strategies, including the complete Embden-Meyerhof-Parnas glycolytic pathway, the complete tricarboxylic acid cycle, and genes for the conversion and assimilation

TABLE 1 Sequencing and assembly results for the *Endozoicomonas* type strains

| Strain | Species | Host organism | GenBank accession no. | Genome size (Mbp) | No. of scaffolds | Scaffold N_{50} (Mbp) | No. of contigs | No. of ORFs ^a | No. of 5S rRNAs | No. of 16S rRNAs | No. of 23S rRNAs | G+C content (%) |
|-----------|-----------------------|--|-----------------------|-------------------|------------------|-------------------------|----------------|--------------------------|-----------------|------------------|------------------|-----------------|
| DSM 22380 | <i>E. elysicola</i> | Sea slug, <i>Elysia ornata</i> | JOJP00000000 | 5.61 | 2 | 5.57 | 21 | 4,270 | 8 | 6 | 6 | 46.8 |
| LMG 24815 | <i>E. montiporae</i> | Coral, <i>Montipora aequituberculata</i> | JOKG00000000 | 5.60 | 20 | 1.02 | 83 | 4,362 | 8 | 4 | 4 | 48.5 |
| DSM 25634 | <i>E. numazuensis</i> | Sponge, cf. <i>Haliclona</i> spp. | JOKH00000000 | 6.34 | 31 | 0.92 | 131 | 4,650 | 3 | 5 | 2 | 47.1 |

^a ORF, open reading frame.

lation of nitrate. Although a genome sequence of *E. elysicola* was already available (28), we provide ordered contigs in an almost-closed scaffold for gene synteny studies, which was not available previously.

Nucleotide sequence accession numbers. These whole-genome shotgun projects have been deposited at DDBJ/EMBL/GenBank under the accession numbers given in Table 1. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

This work was supported by a KAUST-WHOI Special Academic Partnership Fellow Award to M.J.N. awarded by the King Abdullah University of Science and Technology (KAUST).

We thank Karie Holtermann for assistance with the cultures.

REFERENCES

- Apprill A, Huguen K, Mincer T. 2013. Major similarities in the bacterial communities associated with lesioned and healthy *Fungiidae* corals. *Environ. Microbiol.* 15:2063–2072. <http://dx.doi.org/10.1111/1462-2920.12107>.
- Bayer T, Neave MJ, Alsheikh-Hussain A, Aranda M, Yum LK, Mincer T, Huguen K, Apprill A, Voolstra CR. 2013. The microbiome of the Red Sea coral *Stylophora pistillata* is dominated by tissue-associated *Endozoicomonas* bacteria. *Appl. Environ. Microbiol.* 79:4759–4762. <http://dx.doi.org/10.1128/AEM.00695-13>.
- Jessen C, Villa Lizcano JF, Bayer T, Roder C, Aranda M, Wild C, Voolstra CR. 2013. *In-situ* effects of eutrophication and overfishing on physiology and bacterial diversity of the Red Sea coral *Acropora hemprichii*. *PLoS One* 8:e62091. <http://dx.doi.org/10.1371/journal.pone.0062091>.
- Lema KA, Willis BL, Bourne DG. 2013. Amplicon pyrosequencing reveals spatial and temporal consistency in diazotroph assemblages of the *Acropora millepora* microbiome. *Environ. Microbiol.*, in press. <http://dx.doi.org/10.1111/1462-2920.12366>.
- Lesser MP, Jarett JK. 2014. Culture-dependent and culture-independent analyses reveal no prokaryotic community shifts or recovery of *Serratia marcescens* in *Acropora palmata* with White pox disease. *FEMS Microbiol. Ecol.* 88:457–467. <http://dx.doi.org/10.1111/1574-6941.12311>.
- Speck MD, Donachie SP. 2012. Widespread *Oceanospirillaceae* bacteria in *Porites* spp. *J. Mar. Biol.* 2012:746720. <http://dx.doi.org/10.1155/2012/746720>.
- Nishijima M, Adachi K, Katsuta A, Shizuri Y, Yamasato K. 2013. *Endozoicomonas numazuensis* sp. nov., a gammaproteobacterium isolated from marine sponges, and emended description of the genus *Endozoicomonas* Kurahashi and Yokota 2007. *Int. J. Syst. Evol. Microbiol.* 63:709–714. <http://dx.doi.org/10.1099/ijs.0.042077-0>.
- Bayer T, Arif C, Ferrier-Pagès C, Zoccola D, Aranda M, Voolstra CR. 2013. Bacteria of the genus *Endozoicomonas* dominate the microbiome of the Mediterranean gorgonian coral *Eunicella cavolini*. *Mar. Ecol. Prog. Ser.* 479:75–84. <http://dx.doi.org/10.3354/meps10197>.
- Correa H, Haltli B, Duque C, Kerr R. 2013. Bacterial communities of the gorgonian octocoral *Pseudopterogorgia elisabethae*. *Microb. Ecol.* 66:972–985. <http://dx.doi.org/10.1007/s00248-013-0267-3>.
- Hyun DW, Shin NR, Kim MS, Oh SJ, Kim PS, Whon TW, Bae JW. 2014. *Endozoicomonas atrinae* sp. nov., a bacterium isolated from the intestine of a comb pen shell *Atrina pectinata*. *Int. J. Syst. Evol. Microbiol.* 64:2312–2318. <http://dx.doi.org/10.1099/ijs.0.060780-0>.
- Jensen S, Duperron S, Birkeland NK, Hovland M. 2010. Intracellular *Oceanospirillales* bacteria inhabit gills of *Acesta bivalves*. *FEMS Microbiol. Ecol.* 74:523–533. <http://dx.doi.org/10.1111/j.1574-6941.2010.00981.x>.
- Kurahashi M, Yokota A. 2007. *Endozoicomonas elysicola* gen. nov., sp. nov., a gamma-proteobacterium isolated from the sea slug *Elysia ornata*. *Syst. Appl. Microbiol.* 30:202–206. <http://dx.doi.org/10.1016/j.syapm.2006.07.003>.
- Zielinski FU, Pernthaler A, Duperron S, Raggi L, Giere O, Borowski C, Dubilier N. 2009. Widespread occurrence of an intranuclear bacterial parasite in vent and seep bathymodiolid mussels. *Environ. Microbiol.* 11:1150–1167. <http://dx.doi.org/10.1111/j.1462-2920.2008.01847.x>.
- Forget NL, Kim Juniper S. 2013. Free-living bacterial communities associated with tubeworm (*Ridgeia piscesae*) aggregations in contrasting diffuse flow hydrothermal vent habitats at the Main Endeavour Field, Juan de Fuca Ridge. *MicrobiologyOpen* 2:259–275. <http://dx.doi.org/10.1002/mbo3.70>.
- Dishaw LJ, Flores-Torres J, Lax S, Gemayel K, Leigh B, Melillo D, Mueller MG, Natale L, Zucchetti I, De Santis R, Pinto MR, Litman GW, Gilbert JA. 2014. The gut of geographically disparate *Ciona intestinalis* harbors a core microbiota. *PLoS One* 9:e93386. <http://dx.doi.org/10.1371/journal.pone.0093386>.
- Morrow KM, Moss AG, Chadwick NE, Liles MR. 2012. Bacterial associates of two Caribbean coral species reveal species-specific distribution and geographic variability. *Appl. Environ. Microbiol.* 78:6438–6449. <http://dx.doi.org/10.1128/AEM.01162-12>.
- Mendoza M, Güiza L, Martínez X, Caraballo X, Rojas J, Aranguren LF, Salazar M. 2013. A novel agent (*Endozoicomonas elysicola*) responsible for epitheliocystis in cobia *Rachycentrum canadum* larvae. *Dis. Aquat. Organ* 106:31–37. <http://dx.doi.org/10.3354/dao02636>.
- Pike RE, Haltli B, Kerr RG. 2013. *Endozoicomonas uniceicola* sp. nov. and *Endozoicomonas gorgoniicola* sp. nov., bacteria isolated from the octocorals, *Eunicea fusca* and *Plexaura* sp. *Int. J. Syst. Evol. Microbiol.* 63:4294–4302. <http://dx.doi.org/10.1099/ijs.0.051490-0>.
- Yang CS, Chen MH, Arun AB, Chen CA, Wang JT, Chen WM. 2010. *Endozoicomonas montiporae* sp. nov., isolated from the encrusting pore coral *Montipora aequituberculata*. *Int. J. Syst. Evol. Microbiol.* 60:1158–1162. <http://dx.doi.org/10.1099/ijs.0.014357-0>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <http://dx.doi.org/10.1093/bioinformatics/btu170>.
- Brown CT, Crusoe MR. 2014. The Khmer software package: enabling efficient sequence analysis. *Figshare* <http://dx.doi.org/10.6084/m9.figshare.979190>.
- Brown CT, Howe AC, Zhang Q, Pyrkosz AB, Brom TH. 2012. A reference-free algorithm for computational normalization of shotgun sequencing data. *arXiv:1203.4802*. <http://arxiv.org/abs/1203.4802>.
- Zhang Q, Pell J, Canino-Koning R, Howe AC, Brown CT. 2013. These are not the *k*-mers you are looking for: efficient online *k*-mer counting using a probabilistic data structure. *arXiv:1309.2975*. <http://arxiv.org/abs/1309.2975>.
- Leggett RM, Clavijo BJ, Clissold L, Clark MD, Caccamo M. 2014. NextClip: an analysis and read preparation tool for Nextera long mate pair libraries. *Bioinformatics* 30:566–568. <http://dx.doi.org/10.1093/bioinformatics/btt702>.
- Ribeiro FJ, Przybylski D, Yin S, Sharpe T, Gnerre S, Abouelleil A, Berlin AM, Montmayeur A, Shea TP, Walker BJ, Young SK, Russ C, Nusbaum C, MacCallum I, Jaffe DB. 2012. Finished bacterial genomes from shotgun sequence data. *Genome Res.* 22:2270–2277. <http://dx.doi.org/10.1101/gr.141515.112>.
- Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. *Genome Biol.* 13:R56. <http://dx.doi.org/10.1186/gb-2012-13-6-r56>.
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Res.* 19:1639–1645. <http://dx.doi.org/10.1101/gr.092759.109>.
- Kyrpides NC, Woyke T, Eisen JA, Garrity G, Lilburn TG, Beck BJ, Whitman WB, Hugenholtz P, Klenk H-P. 2014. Genomic encyclopedia of type strains, phase I: the one thousand microbial genomes (KMG-I) project. *Stand. Genomic. Sci.* 9:628–634. <http://dx.doi.org/10.4056/sigs.5068949>.