Genomics Data 11 (2017) 73–74

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Data in Brief

Genome sequencing of three bacteria associated to black band disease from a Colombian reef-building coral

Juan Henao [a], Hermes Pérez [b], Deisy Abril [b], Katterine Ospina [b], Adriana Piza [b], Kelly Botero [a], Cristhian Rincón [b], Jhon Donato [b], Andrea Hurtado [b], Erika García [b], Vanessa Otero [b], Alexander Del Risco [b], Brenda Guerra [b], Yina Cifuentes [a], Alvaro Ordoñez [b], Daniel Rojas [b], Karen Suarez [b], Daniel Osorio [a], Andrés Pinzón [a,*]

[a] Bioinformatics and Systems Biology Group, National University of Colombia, Bogotá, Colombia
[b] Bioinformatics for Omics, Postgraduate Class, National University of Colombia, Bogotá, Colombia

ARTICLE INFO

Article history:
Received 21 November 2016
Received in revised form 2 December 2016
Accepted 7 December 2016
Available online 15 December 2016

Keywords:
Black band disease
Genome announcement
Coral reefs
Stappia indica
Nitratireductor aquibiodomus

ABSTRACT

We announce the draft genome sequence of three Gram-negative bacteria isolated from coral tissues affected with the black band disease (BBD), identified with the NCBI's Assembly Database accession numbers: MBQF, MAYB and MBQE. These genome drafts constitute an useful tool for the characterisation of these bacteria and for the understanding of the relationship between the microbial consortia associated with the disease and the onset and progression of the pathology.

Specifications	
Organism/cell line/tissue	*Stappia indica; Nitratireductor aquibiodomus*
Strain	*EBBD 17.2, EBB 35.1, SBBC 49*
Sequencer or array type	Illumina Hiseq 2000
Data format	Assembled
Experimental factors	Coral reef affected
Experimental features	Whole genome shotgun sequencing followed by genome assembly and a brief gene description
Consent	N/A
Sample source location	Santa Marta, Colombia

1. Direct link to deposited data

EBBD 17.2 (*Stappia indica*): https://www.ncbi.nlm.nih.gov/bioproject/328867

EBB 35.1 (*Nitratireductor aquibiodomus*): https://www.ncbi.nlm.nih.gov/bioproject/PRJNA327597

SBBC 49 (*Stappia indica*): https://www.ncbi.nlm.nih.gov/bioproject/PRJNA328866

2. Experimental design, materials and methods

Coral reefs are one of the most diverse, yet vulnerable, ecosystem in the world. Indeed, local and global environmental changes have increased the vulnerability of these ecosystems to develop coral diseases of probable infectious nature [1]. The black band disease (BBD) was one of the first documented coral diseases and affects coral reefs worldwide [2]. Although the etiology of the BBD is not completely understood it appears that is caused by a consortium of cyanobacteria that might infringe direct damage and/or facilitate the colonization by opportunistic bacteria [3]. Hereby we present three high quality draft genome assemblies from three bacteria isolated from corals affected with BBD. Coral samples were collected in Santa Marta Bay, Colombia, and were subsequently macerated and plated on LB or Marine agar.

The genomic DNA was extracted by *Salting Out* and resuspended in DEPC water. Whole-Genome sequencing was performed using the Illumina Hiseq 2000 system with paired-end reads (Table 1). In order to ensure the best computational results we tested different pre-processing and assembly methods. Sequence reads were preprocessed using the Trimmomatic [4], Sickle [5], and Trim Galore [6] tools. A *de novo* genome assembly was conducted using Spades [7], Velvet [8], and Abyss [9], iterating over different k-mers and other software-specific parameters. Overall three pre-processing and three assembly steps

* Corresponding author at: Institute for Genetics, National University of Colombia, UN. 11001, Colombia.
E-mail address: ampinzonv@unal.edu.co (A. Pinzón).

http://dx.doi.org/10.1016/j.gdata.2016.12.008

Table 1
Overall information of the assembly and annotation process to the three bacteria.

	EBBD 17.2	EBB 35.1	SBBC 49
# Of reads	11.126.882	11.305.818	11.154.437
Length of reads	101 pb	101 pb	101 pb
# Of high-quality reads	11.019.579	11.191.685	11.045.752
# Of contigs	75	71	28
Length largest contig	1.427.481 pb	744.415 pb	1.710.053 pb
N50	507.072	637.544	718.983
Length final assembly	4.827.765 pb	4.495.976 pb	5.008.263 pb
GC content	66,95%	61,46%	67,05%
Coverage	36×	50,26×	48,73×
Species found	*Stappia indica*	*Nitratireductor aquibiodomus*	*Stappia indica*
# Of predicted protein-coding	4429	4378	4658
# Of predicted protein-coding related with virulence	66	95	692

were tested. An assembly based on a reference genome was not feasible due to the lack of published genomes for the species under study.

The taxonomical classification of the organisms under study were analysed with the RNAmmer tool [10] to find the 16S sequence and the Ez Taxon tool [11] to find the most exact taxon position based on their similarity value. Based on this analysis the bacteria under study were identified as *Stappia indica* to strain EBBD 17.2 and SBBC 49, and *Nitratireductor aquibiodomus* to strain EBB 35.1.

An automated functional annotation using the RAST server [12] was also performed. This annotation process allowed us to identify 4429 protein-coding sequences to the strain EBBD 17.2, 66 of which were associated with the virulence/resistance process (Table 1). For strains EBB 35.1 and SBBC 49, we found 4378 and 4658 protein-coding sequences respectively, where 95 and 692 sequences were associated with virulence/resistance for strains EBB 35.1 and SBBC 49, respectively (Table 1).

Genomes accession numbers

These bacterial genomes were deposited at DDBJ/ENA/GenBank Databases under the accession MBQF to the bacterium 1 (*Stappia indica*), MAYB to the bacterium 2 (*Nitratireductor aquibiodomus*) and MBQE to the bacterium 3 (*Stappia indicia*), with their corresponding versions MBQF01, MAYB01 and MBQE01.

Funding

This work was funded by "Dirección de investigación y extensión" and "Dirección académica", at National University of Colombia (grant number 32687), and by a grant from the Colombian "Departamento Administrativo de Ciencia, Tecnología e Innovación – COLCIENCIAS" to Dr. Catalina Arévalo-Ferro (Contract 382-2011, grant number 110152128415).

Acknowledgments

Bacteria isolation and DNA extraction was performed by Laura Ripe Jaime, Erika Diaz, Catalina Arévalo-Ferro, y Luis F. Cadavid from the Evolutive Inmunology and Inmunogenetics Group at Institute for Genetics, National University of Colombia.

References

[1] L. Arotsker, E. Kramarsky-Winter, E. Ben-Dov, N. Siboni, A. Kushmaro, Changes in the bacterial community associated with black band disease in a Red Sea coral, *Favia* sp., in relation to disease phases. Dis. Aquat. Org. 116 (1) (2015) 47–58.
[2] J.L. Meyer, S.P. Gunasekera, R.M. Scott, V.J. Paul, M. Teplitski, Microbiome shifts and the inhibition of quorum sensing by black band disease cyanobacteria. ISME J. 10 (5) (2016) 1204–1216.
[3] Y. Sato, M. Civiello, S.C. Bell, B.L. Willis, D.G. Bourne, Integrated approach to understanding the onset and pathogenesis of black band disease in corals. Environ. Microbiol. 18 (3) (2016) 752–765.
[4] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina Sequence Data. Bioinformatics 30 (15) (2014) 2114–2120.
[5] N.A. Joshi, J.N. Fass, Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]Available at https://github.com/najoshi/sickle2011.
[6] F. Krueger, Trim Galore!: A Wrapper Tool around Cutadapt and FastQC to Consistently Apply Quality and Adapter Trimming to FastQ Files. 2015.
[7] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, ... A.V. Pyshkin, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19 (5) (2012) 455–477.
[8] D.R. Zerbino, E. Birney, Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18 (5) (2008) 821–829.
[9] J.T. Simpson, K. Wong, S.D. Jackman, J.E. Schein, S.J. Jones, I. Birol, ABySS: a parallel assembler for short read sequence data. Genome Res. 19 (6) (2009) 1117–1123.
[10] K. Lagesen, P. Hallin, E.A. Rødland, H.H. Stærfeldt, T. Rognes, D.W. Ussery, RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35 (9) (2007) 3100–3108.
[11] J. Chun, J.H. Lee, Y. Jung, M. Kim, S. Kim, B.K. Kim, Y.W. Lim, EzTaxon: a web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences. Int. J. Syst. Evol. Microbiol. 57 (10) (2007) 2259–2261.
[12] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, ... F. Meyer, The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9 (1) (2008) 1.