

1 **Supplemental Files for:**

2
3 **Why orange Guaymas Basin *Beggiatoa* (*Maribeggiatoa*) are orange:**

4 **Single-filament genome-enabled identification of an abundant octaheme cytochrome**
5 **with hydroxylamine oxidase, hydrazine oxidase, and nitrite reductase activities**

6
7 Barbara J. MacGregor, Jennifer F. Biddle, Jason R. Siebert, Eric Staunton, Eric L. Hegg,
8 Ann G. Matthyse, and Andreas Teske

9
10
11 ¹Department of Marine Sciences, University of North Carolina - Chapel Hill, Chapel Hill
12 NC 27599

13 ²College of Earth, Ocean, and the Environment, University of Delaware, Lewes DE
14 19958

15 ³Department of Biochemistry & Molecular Biology, Michigan State University, East
16 Lansing MI 48824-1319

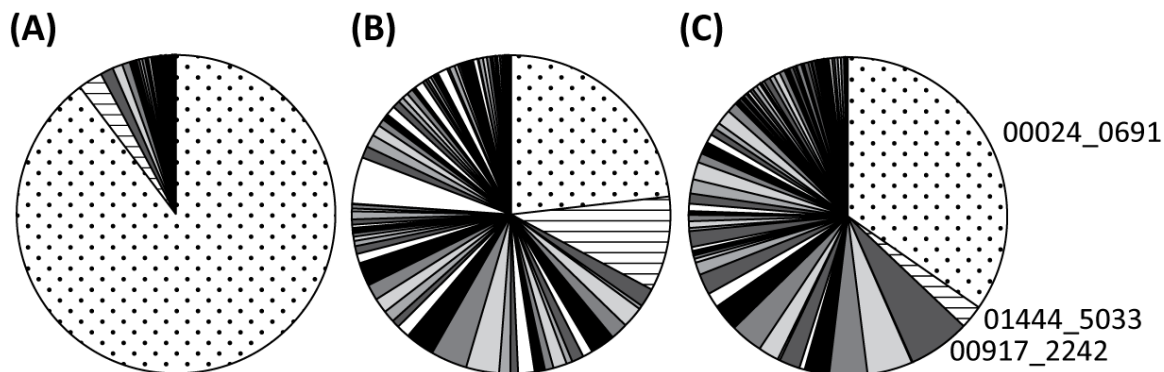
17 ⁴Department of Environmental and Health Sciences, University of North Carolina -
18 Chapel Hill, Chapel Hill NC 27599

19 ⁵Department of Biology, University of North Carolina - Chapel Hill, Chapel Hill NC
20 27599

21
22 #To whom correspondence should be addressed: bmacgreg@unc.edu

23
24
25
26

27



28
29

30 **Figure S1. Identification of ORF encoding the orange protein by μ LC/MS/MS. (A)**

31 **Sum of total ion currents per identified ORF (0.2% not assigned). (B) Average of total**

32 **ion currents for fragments assigned to each identified ORF (0.8% not assigned). (C)**

33 **Number of MS/MS spectra attributed to each identified ORF (total 1272, with 29 not**

34 **assigned)**

35



39 **Figure S2. Phylogenetic tree of inferred amino acid sequences of BOGUAY**
 40 **00024_0691 and BOGUAY 00935_1708.** The first 100 BLASTP/nr hits for each
 41 sequence (all but two of which were identical) were used to construct an initial tree,
 42 which was then trimmed to show only closest relatives. Evolutionary analyses were
 43 conducted in MEGA5 (4) with sequences aligned by MUSCLE (1). The evolutionary
 44 history was inferred using the Neighbor-Joining method (3). The percentage of replicate
 45 trees in which the associated taxa clustered together in the bootstrap test (500 replicates)
 46 are shown next to the branches (2). The evolutionary distances were computed using the
 47 Poisson correction method (5) and are in units of amino acid substitutions per site. All
 48 ambiguous positions were removed for each sequence pair. The analysis involved 35
 49 amino acid sequences, and 1132 positions were included in the final analysis.

50

51

52

53 **Table S1. NADPH-quinone oxidoreductase genes.**

54

Gene product	Complex	Contig_ORF
NuoA	Membrane complex	00285_1243
NuoB	FeS protein	00285_1242, 00322_3116
NuoC	FeS protein	00285_1236, 00322_3117
NuoD	FeS protein	00285_1234, 00322_3119
NuoE	FMN	00285_1225
NuoF	FMN	00552_2594, 01153_1100
NuoG	FMN	00432_3220
NuoH	Membrane complex	00432_3219, 00362_1731
NuoI	FeS protein	00432_3218
NuoJ	Membrane complex	00432_3216
NuoK	Membrane complex	00432_3215
NuoL	Membrane complex	00543_4484
NuoM	Membrane complex	01191_1496
NuoN	Membrane complex	00024_0688

55

56

56

57 **SUPPLEMENTAL REFERENCES**

- 58 1. **Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy
59 and high throughput. *Nucleic Acids Research* **32**:1792-1797.
- 60 2. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using
61 the bootstrap. *Evolution* **39**:783-791.
- 62 3. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method
63 for reconstructing phylogenetic trees. *Molecular Biology and Evolution*
64 **4**:406-425.
- 65 4. **Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar.**
66 2011. MEGA5: Molecular Evolutionary Genetics Analysis using maximum
67 likelihood, evolutionary distance, and maximum parsimony methods.
68 *Molecular Biology and Evolution* **28**:2731-2739.
- 69 5. **Zuckerandl, E., and L. Pauling.** 1965. Evolutionary convergence and
70 divergence in proteins, p. 97-166. *In* V. Bryson and H. J. Vogel (ed.), *Evolving*
71 *Genes and Proteins*. Academic Press, New York.
- 72
- 73