

**Insights into the distribution and abundance of the ubiquitous Candidatus
Saccharibacteria phylum following tag pyrosequencing**

Belinda Ferrari¹ #, Tristrom Winsley¹, Mukan Ji¹ and Brett Neilan¹

1. School of Biotechnology and Biomolecular Sciences, University of New South
Wales, Randwick, NSW, 2052, Australia.

Corresponding author:

Dr Belinda C Ferrari, School of Biotechnology and Biomolecular Sciences, University
of New South Wales, Randwick, NSW, 2052, Australia.

Tel: +612 9385 2032, Fax: +612 9385 1483

Email: b.ferrari@unsw.edu.au

1

2

3 **Figure S1. Maximum-likelihood circular cladogram of the Candidatus**

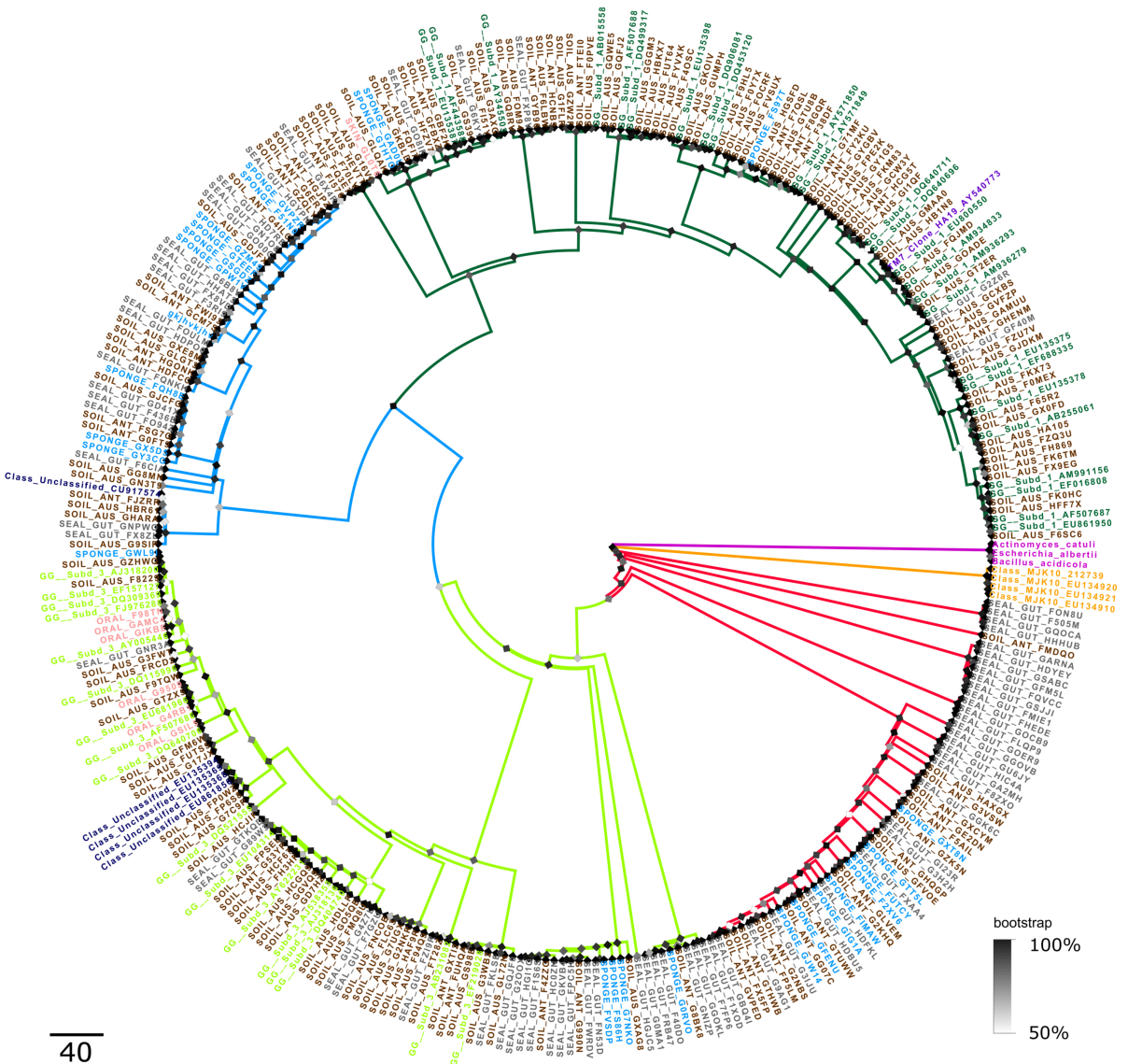
4 **Saccharibacteria phylum.**

5 This similarity tree was constructed with 229 new sequences (OTUs defined at 0.08
6 dissimilarity), as well as 46 sequences from GreenGenes
7 representing previously defined class level clades 1, 3, MJK10 and the outgroup
8 species, *E. albertii* str. 19982 (AY696662), *B. acidicola* str. TSAS-
9 1(GQ389780) and *A. catuli* str. CCUG 41709 (AJ276805). The reference
10 sequences from GreenGenes were prefixed with the subdivision affiliation assigned
11 by the GreenGenes taxonomy; Subdivision 1 (Dark Green), Subdivision 3 (Light
12 Green), Class MJK10 (Orange), or unclassified class (Dark Green). The sequences
13 obtained in this study (229) are colored according to their source, soil (brown),
14 sponge (blue), seal gut (grey), human (pink). The branches of the trees are colored
15 according to the 4 proposed clades; 1 (dark green), 2 (blue), 3 (light green), 4 (red).
16 Bootstrap confidence was calculated from 1000 replicates and displayed by
17 diamonds at each node. The color of the diamonds, ranging from white to black
18 corresponds to the bootstrap confidence values (between 50 and 100%). The image
19 was produced with Figtree software v1.4.0 (<http://tree.bio.ed.ac.uk/software/figtree/>)
20 from a newick file generated by FastTree software.

21

22

23



24
25
26
27
28
29
30
31
32
33
34

40

35 **Table S1. DNA Sample Description**

Sample	Sample Type	Sample Source	DNA extraction method
1	Marine Sponge	Botany Bay, NSW, Australia	FastDNA spin Kit
2	Marine Sponge	Botany Bay, NSW, Australia	FastDNA spin Kit
3	Marine Sponge	Botany Bay, NSW, Australia	FastDNA spin Kit
4	Marine Sponge	Botany Bay, NSW, Australia	FastDNA spin Kit
5	Marine Sponge	Botany Bay, NSW, Australia	FastDNA spin Kit
7	Marine Sponge	Great Barrier Reef, QLD, Australia	FastDNA spin Kit
8	Marine Sponge	Great Barrier Reef, QLD, Australia	FastDNA spin Kit
9	Marine Sponge	Great Barrier Reef, QLD, Australia	FastDNA spin Kit
10	Marine Sponge	Great Barrier Reef, QLD, Australia	FastDNA spin Kit
11	Marine Sponge	Great Barrier Reef, QLD, Australia	FastDNA spin Kit
13	Marine Sponge	Great Barrier Reef, QLD, Australia	FastDNA spin Kit
14	Marine Sponge	Great Barrier Reef, QLD, Australia	FastDNA spin Kit
17	Soil	Pagewood, NSW, Australia	FastDNA spin Kit
18	Soil	Pulya Pulya Dam, NT, Australia	FastDNA spin Kit
19	Soil	Eastlakes, NSW, Australia	FastDNA spin Kit
20	Soil	UNSW Australia, NSW, Australia	FastDNA spin Kit
21	Soil	UNSW Australia, NSW, Australia	FastDNA spin Kit
22	Soil	UNSW Australia, NSW, Australia	FastDNA spin Kit
23	Soil	Macquarie University, NSW, Australia	FastDNA spin Kit
24	Polar soil	Herring Island, Antarctica	FastDNA spin Kit
25	Polar soil	Herring Island, Antarctica	FastDNA spin Kit
26	Polar soil	Casey Station, Antarctica	FastDNA spin Kit
27	Polar soil	Casey Station, Antarctica	FastDNA spin Kit
28	Seal faeces	Leopard seal, Danco Coast, WA, Australia	QIAamp stool kit
29	Seal faeces	Leopard seal, Danco Coast, WA, Australia	QIAamp stool kit
30	Seal faeces	Southern elephant seal, South Shetland Islands, WA, Australia	QIAamp stool kit
31	Seal faeces	Captive leopard seal, Taronga Zoo, NSW, Australia	QIAamp stool kit

32	Seal faeces	Southern elephant seal, Taronga Zoo, Sydney, NSW, Australia	QIAamp stool kit
34	Seal faeces	Southern elephant seal, Taronga Zoo, Sydney, NSW, Australia	QIAamp stool kit
35	Human oral	Extracted teeth; Sydney, NSW, Australia	prepGEM tissue kit
37	Human oral	Extracted teeth; Sydney, NSW, Australia	prepGEM tissue kit
38	Human oral	Extracted teeth; Sydney, NSW, Australia	prepGEM tissue kit
39	Human oral	Extracted teeth; Sydney, NSW, Australia	prepGEM tissue kit
40	Human skin	Sydney, NSW, Australia	prepGEM tissue kit

36

37

38

39 **Table S2. PCR primer sequences targeting Saccharibacteria and *in silico* matches is RDP.**
 40 Number of matches to accessions in each phylum in the RDP Release 10 Update 28. Included are
 41 published primers that target the Candidate Division (Hugenholtz *et al.*, 2001; Brinig, *et al.*, 2003) as
 42 well as the 4 taxon-specific primers designed in this study.

43
 44

Phylum	Number of Accessions	Hugenholtz <i>et al.</i> , 2001		Brinig <i>et al.</i> , 2003			This study			
		TM7-314F	TM7-580F	TM7-910F	TM7-1093F	TM7-1177R	TM7-211F	TM7-686R	TM7-590F	TM7-965R
<i>Actinobacteria</i>	187010	0	2	1	0	0	1	2	2	1
<i>Aquificae</i>	907	0	0	0	0	0	0	0	0	0
<i>Bacteroidetes</i>	88998	5	0	0	9	76	3	1	1	0
<i>Caldiserica</i>	53	0	0	0	0	0	0	0	0	0
<i>Chlamydiae</i>	361	0	0	0	0	0	0	0	0	0
<i>Chlorobi</i>	809	1	0	0	0	0	0	0	0	0
<i>Chloroflexi</i>	5505	123	0	0	0	14	3	0	0	0
<i>Chrysiogenetes</i>	9	0	0	0	0	0	0	0	0	0
<i>Deferribacteres</i>	287	0	0	0	0	0	0	0	0	0
<i>Deinococcus-Thermus</i>	1329	0	0	0	0	0	0	0	0	0
<i>Dictyoglomi</i>	22	0	0	0	0	0	0	0	0	0
<i>Elusimicrobia</i>	139	0	0	0	0	0	0	0	0	0
<i>Fibrobacteres</i>	256	0	0	0	0	0	0	0	0	0
<i>Fusobacteria</i>	4659	1	0	0	0	0	1	0	0	0
<i>Gemmatimonadetes</i>	1025	0	0	0	0	0	0	0	0	0
<i>Lentisphaerae</i>	224	0	0	0	0	0	0	0	0	0
<i>Nitrospira</i>	1064	0	0	0	0	0	0	0	0	0
<i>Planctomycetes</i>	4142	0	0	0	0	0	1	0	0	0
<i>Proteobacteria</i>	275869	36	2	2	4	4	15	2	2	8
<i>Spirochaetes</i>	3563	0	0	0	0	0	0	0	0	0
<i>Synergistetes</i>	944	0	0	0	0	0	0	0	0	0
<i>Tenericutes</i>	2534	0	0	0	0	0	0	0	0	0
<i>Thermodesulfobacteria</i>	116	0	0	0	0	0	0	0	0	0
<i>Thermotogae</i>	528	0	0	0	0	0	0	0	0	0
BRC1	70	0	0	0	0	0	0	0	0	0
OD1	127	0	0	0	0	0	0	0	0	0
OP11	84	0	0	0	0	0	0	0	0	0
SR1	247	0	0	0	0	0	0	0	0	0
Saccharibacteria	1650	692	1442	1541	1311	1052	1308	1560	1564	1527
WS3	135	0	0	0	0	0	0	0	0	0
<i>Armatimonadetes</i>	528	11	0	0	1	0	0	0	0	0
<i>Verrucomicrobia</i>	5263	1	1	0	0	1	1	1	1	0
<i>Acidobacteria</i>	10124	0	1	0	0	0	0	1	1	0
<i>Firmicutes</i>	404557	4	0	3	6	4	4	3	1	3
<i>Cyanobacteria/Chloroplast</i>	11341	0	0	0	0	0	0	0	0	0
unclassified Bacteria	17744	63	6	7	3	9	58	13	7	71
Total	1032223	937	1454	1554	1334	1160	1395	1583	1579	1610
% Specificity		73.9	99.2	99.2	99.8	90.7	93.8	98.5	99.1	94.8
% Saccharibacteria Accessions matched		41.9	87.4	93.4	79.5	63.8	79.2	94.5	94.7	92.5

45
 46
 47
 48

49 **Table S3. *In silico* matches is RDP to the new Saccharibacteria taxon-specific primer set.** The
50 number of matches to accessions in each phylum in the RDP Release 10 Update 28 are reported,
51 including those after 1 and 2 mismatches.

52

	Total number of accessions	0 mismatches	1 mismatch	2 mismatches
Domain <i>Bacteria</i>	1235311	1720	2097	11850
<i>Actinobacteria</i>	180827	0	0	26
<i>Aquificae</i>	945	0	0	4
<i>Bacteroidetes</i>	141057	0	0	9
<i>Caldiserica</i>	220	0	0	0
<i>Chlamydiae</i>	504	0	0	0
<i>Chlorobi</i>	1058	0	0	0
<i>Chloroflexi</i>	20447	0	0	17
<i>Chrysiogenetes</i>	12	0	0	0
<i>Deferribacteres</i>	378	0	0	0
<i>Deinococcus-Thermus</i>	1844	0	0	0
<i>Dictyoglomi</i>	22	0	0	0
<i>Elusimicrobia</i>	172	0	0	0
<i>Fibrobacteres</i>	308	0	0	39
<i>Fusobacteria</i>	9378	0	0	0
<i>Gemmatimonadetes</i>	1256	0	0	0
<i>Lentisphaerae</i>	1707	0	0	0
<i>Nitrospira</i>	1433	0	0	0
<i>Planctomycetes</i>	11233	0	0	1
<i>Proteobacteria</i>	341501	0	0	9192
<i>Spirochaetes</i>	9606	0	0	0
<i>Synergistetes</i>	1120	0	0	0
<i>Tenericutes</i>	3214	0	0	0
<i>Thermodesulfobacteria</i>	107	0	0	0
<i>Thermotogae</i>	582	0	0	0
Candidate phylum BRC1	399	0	0	0
Candidate phylum OD1	153	0	0	6
Candidate phylum OP11	97	0	0	0
Candidate phylum SR1	234	0	0	0
<i>Saccharibacteria</i>	2167	1719	2072	2092
Candidate phylum WS3	529	0	0	0
<i>Armatimonadetes</i>	1016	0	0	2
<i>Verrucomicrobia</i>	9586	0	0	24
<i>Acidobacteria</i>	14151	0	0	99
<i>Firmicutes</i>	420503	0	0	6
<i>Cyanobacteria</i>	21330	0	0	70
Unclassified <i>Bacteria</i>	36214	1	25	263
<i>Archaea</i>	23532	0	0	0
% <i>Saccharibacteria</i> sequences targeted		79.3	95.6	96.5