

Phylogenomic and pangenomic assessment of a Mediterranean strain of *Raphidiopsis raciborskii* extends knowledge of the global distribution and characteristics of a potentially toxigenic cyanobacterium

Nico Salmaso, Leonardo Cerasino, Margherita Di Brizio, Massimo Pindo, Adriano Boscaini

Supplementary Tables

Table S1

Species belonging to the *Raphidiopsis* species analysed in this work. NCBI assembly codes: GenBank assembly accession codes. NCBI organism name: the taxon name and strain code in the NCBI GenBank taxonomy. GTDB organism name: the classification adopted in the GTDB (Genome Taxonomy Database) taxonomy. Country code: alpha-3 country codes indicating where the organism has been collected.

NCBI assembly code	NCBI Organism Name	GTDB Organism name	Country code
GCA_002893155.1	<i>Cylindrospermopsis raciborskii</i> C03	Raphidiopsis raciborskii	AUS
GCA_002893145.1	<i>Cylindrospermopsis raciborskii</i> C04	Raphidiopsis raciborskii	AUS
GCA_002893125.1	<i>Cylindrospermopsis raciborskii</i> C07	Raphidiopsis raciborskii	AUS
GCA_001676585.1	<i>Cylindrospermopsis raciborskii</i> CS-505	Raphidiopsis raciborskii	AUS
GCA_001858115.1	<i>Cylindrospermopsis raciborskii</i> CS-508	Raphidiopsis raciborskii	AUS
GCA_002893285.1	<i>Cylindrospermopsis raciborskii</i> S01	Raphidiopsis raciborskii	AUS
GCA_002893265.1	<i>Cylindrospermopsis raciborskii</i> S05	Raphidiopsis raciborskii	AUS
GCA_002893245.1	<i>Cylindrospermopsis raciborskii</i> S06	Raphidiopsis raciborskii	AUS
GCA_002893205.1	<i>Cylindrospermopsis raciborskii</i> S07	Raphidiopsis raciborskii	AUS
GCA_002893215.1	<i>Cylindrospermopsis raciborskii</i> S10	Raphidiopsis raciborskii	AUS
GCA_002893185.1	<i>Cylindrospermopsis raciborskii</i> S14	Raphidiopsis raciborskii	AUS
GCA_002027345.1	<i>Cylindrospermopsis raciborskii</i> CENA302	Raphidiopsis brookii	BRA
GCA_002114155.1	<i>Cylindrospermopsis raciborskii</i> CENA303	Raphidiopsis brookii	BRA
GCA_001586755.1	<i>Cylindrospermopsis raciborskii</i> ITEP-A1	Raphidiopsis brookii	BRA
GCA_026929835.1	<i>Cylindrospermopsis raciborskii</i> PAMP2011	Raphidiopsis brookii	BRA
GCA_026929755.1	<i>Cylindrospermopsis raciborskii</i> PAMP2012	Raphidiopsis brookii	BRA
GCA_000175855.1	<i>Raphidiopsis brookii</i> D9	Raphidiopsis brookii	BRA
GCA_018969285.1	Cyanobacteria bacterium REEB494	Raphidiopsis raciborskii	CHN
GCA_022496815.1	<i>Cylindrospermopsis raciborskii</i> CHAB3438	Raphidiopsis raciborskii	CHN
GCA_036409495.1	<i>Cylindrospermopsis raciborskii</i> DSH	NA	CHN
GCA_018139025.1	<i>Cylindrospermopsis raciborskii</i> N8	Raphidiopsis raciborskii	CHN
GCA_035282025.1	<i>Cylindrospermopsis raciborskii</i> 1523720	NA	IND
GCA_021650815.1	<i>Cylindrospermopsis raciborskii</i> KLL07	Raphidiopsis raciborskii	ISR
GCA_002368135.1	<i>Raphidiopsis curvata</i> NIES-932	Raphidiopsis curvispora_A	JPN
GCA_014489415.1	<i>Cylindrospermopsis curvispora</i> GIHE-G1	Raphidiopsis curvispora_A	KOR
GCA_006523545.1	<i>Cylindrospermopsis raciborskii</i> GIHE 2018	Raphidiopsis raciborskii	KOR
GCA_003367075.2	<i>Cylindrospermopsis raciborskii</i> Cr2010	Raphidiopsis raciborskii	NLD
GCA_045377515.1	<i>Cylindrospermopsis raciborskii</i> UAM/DH-BiRr	NA	POL
GCA_045377475.1	<i>Cylindrospermopsis raciborskii</i> UAM/DH-KmRr	NA	POL
GCA_045377415.1	<i>Cylindrospermopsis raciborskii</i> UAM/DH-MRr	NA	POL
GCA_045377345.1	<i>Cylindrospermopsis raciborskii</i> UAM/DH-ZRr	NA	POL
GCA_001432185.1	<i>Cylindrospermopsis</i> sp. CR12	Raphidiopsis raciborskii	SGP
GCA_001858125.1	<i>Cylindrospermopsis raciborskii</i> MVCC14	Raphidiopsis brookii	URY
GCA_012583295.1	<i>Cylindrospermopsis raciborskii</i> MVCC19	Raphidiopsis brookii	URY
GCA_015708265.1	<i>Cylindrospermopsis raciborskii</i> KL1	Raphidiopsis brookii	USA
GCA_044352285.1	<i>Raphidiopsis brookii</i>	NA	USA

Table S2

(A) Blast analysis of rRNA genes and selected housekeeping genes of taxonomic relevance extracted from (B) the Lake Trasimeno genome annotated by PGAP and bakta, and (C) from the assembled unbinned contigs (Megablast; Max target sequences=500). Q. Len., query length; QC, query cover; pident, percentage identity; Acc. Len., accession length (“genome” if the subject sequence was included in a genome). The rRNA and a few of the selected housekeeping genes (core_nt, in bold) have been used in several phylogenetic analyses based on target genes (see the section Introduction).

(A)

Database: Core nucleotide database (core_nt)

Gene	Q. Len. (bp)	QC (%)	pident (%)	Acc. Len. (bp)	Blast top pident accession	NCBI Taxonomy	Water body
5S rRNA	118	100	100	genomes	several	<i>C. raciborskii/curvispora</i>	global
5S rRNA	114	100	100	genomes	several	<i>C. raciborskii/curvispora</i>	global
16S rRNA	120	100	100	genomes	several	<i>C. raciborskii/curvispora</i>	global
23S rRNA	2821	99	99.4	2818	LC455654.1	<i>C. raciborskii</i> NIES-1260	Gonoike Pond, JPN
<i>rpoB</i>	3342	100	96.92	genomes	CP073250.1	<i>C. raciborskii</i> N8	Zhenhai Reservoir CHN
<i>rbcX</i>	402	81	97.26	582	several	<i>C. raciborskii/R. mediterranea/R. curvata</i>	several, Asia
<i>nifH</i>	888	33-38	100	297-345	several	<i>C. raciborskii</i> /Uncultured bacterium	ESP, GRC, USA, MEX
<i>rpoC1</i> ⁽¹⁾	1878	20	99.85	380	HG942518.1/19.1	<i>C. raciborskii</i> MM67 and MM69	Lake Whangape, NZL
<i>cpcA</i>	489	62	100	710	GU434457.1	<i>Cylindrospermopsis</i> W104094B	Lake Wingra, USA
<i>cpcB</i>	522	43	100	611	JN903642.1	<i>C. raciborskii</i> UAM520	Vega del Jabalon ESP
16S rRNA (unbinned)	1485	98	100	1461	FM177492.1	<i>C. raciborskii</i> 1LT32S01	Lake Trasimeno ITA

⁽¹⁾ QC 35% (accession length=660), pident=99.85%, *C. raciborskii* AZ60, Fish Pond, Arizona, USA.

Database: Whole Genome shotgun contigs (wgs)

Gene	Q. Len. (bp)	QC (%)	pident (%)	Acc. Len. (bp)	Blast top pident accession	NCBI Taxonomy	Water body
5S rRNA	118	100	100	several	several	<i>C. raciborskii/curvispora</i>	global
5S rRNA	114	100	100	several	several	<i>C. raciborskii/curvispora</i>	global
16S rRNA	120	100	100	several	several	<i>C. raciborskii/curvispora</i>	global
23S rRNA	2821	100	99.75	4749	VOIM01000076.1	<i>C. raciborskii</i> LB2897	Lemon Lake, USA
<i>rpoB</i>	3342	100	99.97	24858	JBILEC010000040.1	<i>R. brookii</i> ME2016-07-27	Lake Mendota, USA
<i>rbcX</i> ⁽¹⁾	402	100	100	several	several	<i>C. raciborskii / R. brookii</i>	Four water bodies, USA
<i>nifH</i> ⁽²⁾	888	100	100	82151	JBILEC010000040.1	<i>R. brookii</i> ME2016-07-27	Lake Mendota, USA
				259219	JADQCS010000002.1	<i>C. raciborskii</i> KL1	Kissena Lake, USA
<i>rpoC1</i>	1878	100	99.79	24858	JBILEC010000040.1	<i>R. brookii</i> ME2016-07-27	Lake Mendota, USA
				268160	JADQCS010000001.1	<i>C. raciborskii</i> KL1	Kissena Lake, USA
<i>cpcA</i>	489	100	100	61472	JBILEC010000007.1	<i>R. brookii</i> ME2016-07-27	Lake Mendota, USA
				2959	JBEAWA010000478.1	Cyan.NASQAN2011_262_B_bin.3	River water, USA
<i>cpcB</i>	522	100	100	64487	VOIM01000172.1	<i>C. raciborskii</i> LB2897	Lemon Lake, USA
16S rRNA (unbinned)	1485	100	99.87	4749	VOIM01000076.1	<i>C. raciborskii</i> LB2897	Lemon Lake, USA

(B)

Lake Trasimeno genome

5S ribosomal RNA

```
>5S ribosomal RNA TARS_fil_Rr_000000000265:2987-3105
CCTGGTGTCTATGGTGC GGTTGGAACCACTGATCCCTTCCCGAAGCTCAGAGGTGAAACGCTGTTGCGGCGACGATAGTA
TGGGGGTTGCCCTATGTCAAATAGTCTCGATGCCAGGT
```

```
>5S TARS_fil_Rr_0000000002920:28692-28806 (partial=true)
ACCTGGCATCGAGCTATTTTGACATAGGGCAACCCCATACTATCGTCGCCGCAACAGCGTTTCACCTCTGAGTTCGGGA
AGGGATCAGTGTGGTTCCACCGCACCATAGACAC
```

16S ribosomal RNA

```
>16S ribosomal RNA (partial) TARS_fil_Rr_0000000001183:3039-3159 (partial=true)
AAAACGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGTATGCTTAACACATGCAAGTGAACGGGATGCTTAGGC
ATCTAGTGGCGGACGGGTGAGTAACGCGTGAGAATCTGGC
```

23S ribosomal RNA

```
>23S ribosomal RNA TARS_fil_Rr_000000000265:88-2909
GGTCAAGCTAATAAGGGCTCATGGTGATACCTAGGCACACAGAGGCGAAGAAGGACGTGGTTACCTGCGAAAAGTTCGG
```

GGGAGTTGGAAGCAACAGAGAGCCGGAATATCCGAATGGGGCAACCCCTGAACACTGCCTGTTGAATATATAGACAGGA
AAGAGCGAACCTGGCGAATTGAAACATCTTAGTAGCCAGAGGAAAAAGAAATCAAAAGAGATTCCCCGAGTAGTGGTGAGG
GAAAGGGGAAAAAGCCTAAACCAGGGGGTATACCTTCTGGGGTAGTGGGACAGCAAAATCGAATCTAGAGACTAGATGAAG
CAGCT'AAATACTGCACCAGAGGGGGTGAAAGTCCCCGTAGTGC'AAAGTTGAAGGATAGAAGCTGAATCCCGAGTAGTAGCGG
GGCACGAGGAATCCCGTATGAATCAGCGAGGACCATCTCGTAAGGCTAAATACTACTGTGTGACCGATAGAGAACAAGTA
CCCGGAGGGAAGGTGAAAGGAAACCCCGGAGGGGAGTGAAAAAGAACATGAAACCGTGAGCTTACAAGCAGTGGGAGGTC
CGATTAAACGGATGACCCGCTGCCTGTTGAAGAATGAGCCGGCGACTTATAGGTACTGGTAGGTTAAAGCGGGGAATGCTG
GAGCCAAAGGGAAACCGAGTCTGAAAGGGGCGATAATCAGTATTTATAGACCCGAACCCCTGGTGATCTAACCATGGCCAG
GATGAAGCTTGGGTAAACCAAGTGGAGGTCCGAACCGACCGATGTTGAAAAATCGGCGGATGAGCGGTGGTTAGGGGTG
AAATGCCAATCGAACCGAGGAGCTAGCTGGTTCTCCCGAAATGTGTTGAGGCGCAGCGGTAATGAAAAAATTGGGGGGTA
AAGCACTGTTTCGGTGC GGCTGGGAGACCGGTACCAAAATCGAGGCAAACTAAGAATACCCAAGGAGCACATTGCCAGTG
AGACGTTGGGGGATAAGCTTCATCGTCAAGAGGGAAACAGCCCGAGACCACAGCTAAGGTCCCCAAATCATCACTAAGTG
AAAAAGGAGGTGGGGTTGCAAGACAACCTAGGAGGTTTGCCTAGAAGCAGCCACCCCTTGAAGAGTGCGTAATAGCTCAC
TAGTCAAGCATCTGCGCCGAAATCTGCTAGTGGAGTGTACCGCAAGCTGTGGGATTAGAAATAATCGGTAGGGG
AGCGTTCCGTAGTAGTGAGAAGCAGTAGCGGCAAGCAGCTGTGGACGAGACGGAAGTGAGAATGTGCGCTTGAGTAGCGC
AAACATTTGGTGAGAATCCAATGCCCCGAAACCCCTAAGGGTTCAGAGGCGAGGTTTCGTCCACTCTGGGTTAGTCGGGACCT
AAGGCGAGGTGCAAGGGCGTAGTCGATGGAGACCGGTCACAAATCCCGGACTACAATATGGGAGCAGAACTAGGGACGC
ATGAAGATAGCCACGCCCTGAATGGATTGGGAAGACCGTTACGACGGTCGAGTGGGGAAGGAAAGTGCCAAGAAAAAGCT
AGGGTTGTGATGAACATATAGTACCCGTACCCGAAACCGACACAGGTAGGGAGGTTGAGAATACCAAGGGGCGCGAGATA
ACTCTCTCTAAGGAACTCGGCAAAATGGCCCCGTAACTTAGGAAGAAGGGGTGCCAGCCGAAAGGCTGGTCGCAGTGAAG
AGATCCAGGCGACTGTTTACCAAAAACACAGGTCTCCGCTAACTCGAAAGAGGACGTATGGGGGCTGACGCCTGCCAGT
CGCGGAAGGTTAAGGCTTCTGCTAGTGGAAACATAAAGCTGACGACCGCAAGCCCGGTGAACGGCGGCCGTAACTATAA
CGGTCCTAAGTAGAGCGAAATTCCTTGTGCGGTAAGTTCCGACCCGACGAAAGGCGTAACGATCTGGATGGTGTCTCAGA
GAGAGACTCGGCGAAATAGGAATGTCTGTGAAGATACGGACTACCTGCACCTGGACAGAAAGACCCTATGAAGCTTTACT
GTAGCTGGAATTGTGTTGCGGCTTGGCTTGC GCAGATAGGTGGGAAGCGAAGAACTTCTCCTTGAGGGGGGAAGGGGAG
CTAACGGTGAATACCACTCTGGCGAAGCTAGAATTCTAATCTCGTCACCGTGAAGCGGTGAGAGGAAAGTTTCAGTGGG
CAGTTTGACTGGGGCGGTGCGCTCCTTAAAGGTAACGGAGGCGCAAGGTTCTCTCAGCAGCTTGGAAACCGTCGCA
CGAGTGTAAGGCAAGAAGAGAGCTTGACTGCAAGACCAACAAGTCGAGCAGGGACGAAAGTCGGCCTTAGTGATCCGAC
GGCGCAGCATGGAATGGCCGTGCGCTCAACGATATAAAGTTACTCTAGGGATAACAGGCTGATCTCCCCAAGAGTCCACA
TCGACGGGGAGGTTTGGCAGCTGATGTGCGCTCATCGCAACCTGGGGCGGAAGTACGTCCCAAGGGTTGGGCTGTTCCG
CCATTAAAGCGGTACGTAGCTGGGTTAGAACGTCGTGAGCAGTTCGGTCCATATCCGGTGCAGGCGAAAGAACATTG
AGAGGAGTCTCTCTTAGTACGAGAGGACCGGGAGGAACGAACCGCTGGTGTACCAGTTATTTCCGCCAGGAGTAGACGCTG
GGTAGCCAAGTTCCGAGAGGATAACCGCTGAAAGCATCTAAGTGGGAAGCCACCTTAAAGTAGAGTGTCTCACTACGAG
AGTAGGTAAGGTCACGGGGAGAAGACCCGTTGATAGGCTTTATGTGGAAGTACAGTAATGTATGTAGCAGAGGAGTCTTA
ACAGACCGAGGGCTTGACCTC

rpoB

>rpoB TARS_fil_Rr_000000002247:223036-226378

ATGAACAACGAAAAATTACACGGAACCATCTTTCCTATTACCAGATCTAATTGAAATTCAGCGCTCTAGCTTTTCGTGGTT
TTTAGAAGAGGGTTTTAATAGAAGAGCTGAACCTATTAGCTACTATTACGGACTATACGGGAAAAATTAGAACTCCATTTTT
TAGGACATAACTATAAGCTAAAAGAGCCGAAATATAGCGTAGAAGAGTCAAAAAGAAGAGATAGCACATACCGTGTACAG
ATGTATGTACCCACGAGATTGCTTAAATAAAGAAACAGGTGATATTAAAGAACAGGAAGTATTATAGGTGATCTACCGTT
AATGACGGATAGAGGAACCTTTATTATTAACGGAGCCGAGGCTAGTCAATAGTCAATCAAATCGTGCGATCGCCCCGAGTAT
ATTACAAATGAGGATTGCAAGAGATTGATAAAGAGGAGAAGAACAATCTGTCGCGAGTCTGATTTCCCAACCGGGGGCATGGCTAAAA
TTTGAGACAGACCGTAATGATTTAGTGTGGGTAAGAATAGATAAGACCCGCAAAATTGTCAATTACGGTACTACTTTAAAGC
ATTAGACTGTGAGATAACGAAATTTTAGATGCTTTAAGGATCCAGAATACTTTCAAAAGACCATTGAAAAGAGGGGAC
AATTTCTCCGAAGAGGAGGCCCTTCTAGAGCTGTATAGAAAGCTGAGGCCGGGGGAACACCCACCGTGATGGGGGGACAA
CAGCTATTAGAATCGCGCTTTTCGATCCCAACGTTACGACCTAGGTAGGGTAGGTTAGATATAAACTTAACAAGAACT
ACGCCTTTCCGTTCTGTACACCACTAGGGTTCTCACCCAGGGGATATCCTATCCGCTGTGGACTATCTCATCAACCTAG
AATACGACATTTGGTAGTATTGATGACATTGATCATTGGGAATCGCCGAGTTAGAAGCGTTGGTGAAGTGTTCAGAAC
CAAGTAAGAGTTGGTTTAAATCGTTTAGAGAGGATCATTAGAGAAAGAATGACCGTCTCCGATGCCGAAGTGTTAACACC
AGCATCACTAGTGAACCCCAAAACCCCTAGTAGCAGCCATAAAAGAGTTTTTTGGTTCTAGTCAGCTAAGTCAATTCATGG
ATCAGACCAACCCCTTAGCGGAACTAACCCATAAACGACGCTTGAGTGCTTTAGGTCTGGGGGACTAACCCAGAGAGAGA
GCTGGTTTTGCGGTGCGAGACATTCATCCAGCCACTACGGAGCGATTTGCCTATTGAAACGCCGGAAGGTCCAAATGCG
CGGATTAATTGGTTCTCTAGCTACCCATGCTCGTGTTAACCAAGTATGGATTTTGTAGAAACACCCCTTTAGACCCGTAGAAA
ACGGAAGAGTTTGCTACGAAACCCCTGCTGTTACATGACGACGAGGAAGATGACCTGCGAGTTGCACCTGGTGAC
ATTCCCGTAGATGACAACGGGCAAACTACTAGGAATCCAAGTACCCGTTTCGTTATCGCCAGGAATTTTCCACCACCACTCC
CGAACAGGTAGACTACGTAGCTGTGTACCAGTACAATTTGTCTCAGTAGCCACCAGCATGATTCCCTTTTTGGAGCATG
ATGATGCCAACCGAGCCTTAATGGGTTCTAACATGCAAAAGGCAAGCAGTACCCCTATTGTAAACCCGAGCGTCCCCTAGTG
GGAATCGGATGGAGGCCCAAGGTGCCAGAGATTCCAGGTAGGTAATGTTTCCCGTACTGATGGAGATGTGGTATATGT
AGATGCGACCGAGATTCTGTGTTGAGTTAAAGAAAAGGCCGCCCTTACCAACCGTGAAAGCGAAACTATTCCCCATAAAC
CCCAGGAAGTTAAATACGTTCTCTCCAAATATCAACGTTCTAACCCAGGACACCTGTCTCAATCAAAAACCCCTAGTGGG
ATTGGGGAAAAGGTCATAGCAGGTACAGGTATTAGCAGATGGTTATCCACAGAAGGGGGAGAATTAGCATTAGGGCAAAA
TGTTATTGTTGCCATATAGCCCTGGGAAGGGTATAACTACGAAGATGCAATTTTAAATTTCCGAGAGACTGGTACAAGAGG
ATGTTTACACATCAATACACATTTGAAAAATTTGAAATAGCAAGACAAACCAAACTAGGACCTGAAGAGATTACGAGA
GAAATACCAACGTTGGGGAAGACGCCCTCAGACAATTAGACGAACAGGGGATAATCAGAGTAGGGGCCTGGGTAGAATC
AGGAGACATTTCTAGTGGGGAAGTGACCCCTAAAGGAGAATCAGATCAACCACCGGAAGAAAACTACTCAGAGCGATTT
TCGGGGAGAAAGCTCGCAGCTGAGGGATAACTCCCTGCGGGTTCCCAATGGTGAAAAAGGAAGAGTAGTGGATGTGCGA
CTATTTACTAGGTCAGGGGATAACTTCCACCGGAGCCAATATGGTGGTGC GGGTGTATGTAGCCCAAGAACGTTAA
AATCCAGGTTGGTGACAAAATGGCTGGGAGACATGGAATAAAGGCATCATTTCCCGAATTTTGCCTTAGAAGACATGC
CATACCTTCCCGATGGCACCCAGTAGATATAGTCTTAACCCCTAGGTGTACCTAGTCGAATGAACGTAGGACAGGTG
TTTGAATGCTTATTGGGTTGGGCGGTCACAACCTGGGGGTACGCTTTAAATCACCCCTTTGATGAAATGTACGGGGA
AGAATCTCCCGAGCTATAGCTACCGCAAAATTAAGGCAAAATTAGGGAAGCAAGTACGGGATGGGTATATAATCCCG
AAAACCCCTGGCAAAATCATGGTTTATGATGGGCGCACGGGTGAACCCCTTGATCGTGCAGTACCGTGGGGATAGCCTAT
ATGCTTAAACTTGTCCATTTAGTTGATGATAAAATTCATGCTCGCTCCACAGGTCTTATTCTTAGTCACCCAGCAACC
CTTGGGTGGTAAAGCACACAGGGAGGTCAAAGATTTGGAGAGATGGAAGTGTGGGCACTAGAGGCCTTTGGTGCTGCTT
ATACATTACGGAACGTGCTAACTGTGAATCAGATGATATGACGGGACGTAATGAAGCAATTGAACGCCATTGTTAAAGGA
AAAGCTATTTCTCGTCTGCGACTCCTGAGTCTTCAAAGTGTTAATGAGAGAATTACAATCTTAGGATTGGATATTGC

GGTACATAAAGTAGAAACCCAAACAGATGGTAGTTCTCTAGATGTAGAAGTTGACTTGATGGCCGATCAAGTATCTCGTC
GCACCTCTCCCGTCCAACATACGAATCCTTTTCTCGTGATTCCCCTGGATGAGGATGAGTAA

rpoC1

>rpoC1 TARS_fil_Rr_000000002247:226492-228370
ATGCGCTCCGTTCAATCTAACTCAATTTGATTACGTCAAAATTTGGCATTGCTTCACCAGAACGTATCCGCGCTTGGGGGGA
ACGTACCTTACCTAATGGTCAGGTGGTTGGTGAGGTGACAAAACCGGAAACCATTAACTACCGCACCTTCAAACAGAAA
TGGATGGTTTTGTTTTGCGAGCGGATCTTTGGCCCTGCTAAAGACTGGGAATGTCACGTGGTAAATATAAACGGGTTCTGT
CATAGAGGTATTGTCTGTGAGCGCTGTGGGGTAGAGGTAACGGAATCACGAGTTTCGCCGTCATAGAATGGGTTTCATTAA
GTTAGCAGCACCTGTAGCTCATGTTTTGGTACCTCAAAGGCATTTCCAGTTATATTGCTATACTATTAGACATGCCCTTAC
GAGATGTTGAGCAAACTCGTCTACTTTAACTCCTATGTAGTTTTAGCTCCTGGTAATGCCGATACACTCGTTTTACAAGCAA
TTATTAAACAGAAGACCAATGGTTAGAAATTGAGGATAGGATTTATAGTGAAGATTCTCAATTAGTAGGCGTAGAAGTGGG
AATTTGGTGCCGAAGCTTTACTACGTTTGTATCAGATATTAATTTAGAAGAAGAAGCAGAAAAACTGCGGGGGGAAATTG
AAGCAGCAAAGGGACAAAAACGGGCCAAATTGATTAACGCTCTGCGGGTAATTGACAATTTTCATTGCCACTGGTTCCCAA
CCGGAATGGATGGTGATGTGAGCCATTCCAGTCATTCCACCCGATCTCCGTCCTGGTACAACCTAGACGGTGGTAGATT
TGCTACTAGCGATTTTAAACGACCTCTACCGTCGGGTAATCAATCGCAATAATCGTTTTAGCTAGACTACAGGAAATCCTGG
CGCCAGAAATCATTGTTTCGAATGAAAAACGCATGTTACAGGAAGCGGTAGATGCTTTAATTGACAATGGTCGGAGAGGA
CGCACGTAGTAGGAGCCAATAACCGGCCATTAAAGTCCCTATCAGATATCATAGAAGGTAACAGGGGCGATTTCCGGCA
AAACCTCCTGGGTAAGGGTAGACTACTCAGGACGTTTCGGTTATTGTAGTTGGTCCCAAATTGCAAATTCACCAGTGTG
GACTCCCCAGGGAGATGGCCATAGAATTGTTTCAGCCATTTGTGATTAATCGTTTAAATTAGATCGGGAATAGTTAAACAT
ATTAAAGCAGCCAAAAAACTCATTTCTCGTAATGACCCAGTGTTTGGGACGTTCTGGAAGAAGTCATAGAAGGACACCC
GGTTATGTTTAAACCGGCTCCTACGTCATAGGTTGGGAATACAGGCTTTTGAACCCATATTAGTAGAAGGAAGAGCAA
TACAGTTACATCCTCTGTGTTGTCCGGCATTCAATGCTGACTTTGATGGAGATCAGATGGCTGTTACGTTCCCTTATCG
TTAGAAAGCCAAAGCGAAGCAAGACTGCTAATGTTAGCATCCAATAATATTCTTTCCGCCGTACGGGTAAACCCATCGT
CACACCCAGTCAAGACATGGTTTGGGAGCCTATTACCTCACTGCTGAAAACCCCAATGCTAGCAAGGGAGCAGGCAAT
ACTTTGCTCCTCAGATGATGTAATTTAGGCATATCAACAAGATCAAGTAGAAGTACATGCCTATATTATGTTTCGATT
GATGGCGATGTGCAACAGGACAACCGGATAACGAACCCCTTGGAAAGTTACAGAAAAACGACGATGGTAGCCGTACTGTCTT
GTATAAATTTTCGTCGAGTACGAGAAGATGCTCAAGGGAATTTAATTTCCCAATATATATATACAACCTCTGCGAGAGTGA
TTTATAACAAGCAATTCAAGATGCTTTAGCCAGCTAA

rbcX

>rbcX TARS_fil_Rr_000000001118:7103-7505
TTATGGAAAATTTCCCTTCTCACCAGCTGGGTTATCTAAACTGTTATCTGGGTAAGGTTCAAATACGCTCAGTTGGGTTA
TCCGCTCAAGATGTTGACGACGCTGCTCCATGTTGGCTTGCTGTATGTTAGTCCGTACCATCTCTGGCAAAAATTCGTGTG
ATTTCCGCCGCAATGTACTCCGCACTGTCATAATCCGCAAGGCTAAATCCGACTTCTCTTGAAACAGCTCTTCGATGTA
TCTCTCTCCATCTTGAACTTTCCCCGCAGAAAAGTTATGTAGCCAAACCGCCAAGGGTGGATTAGTTTCGCTTAACTGCA
CTAACACTGTCTCAACGCCCTGGTATGTTAGGTAGTCTCGGAGAGTTTGGCAGTCTCTTTGGCAATTTCCCGCAAAATTC
AT

nifH

>nifH TARS_fil_Rr_000000003117:340865-341753
CTACTTGCTTGCAGAGCAGGCTTGCCAATTACTTCAGCGTTTTTGCTGTCTATCATCAAGAATACCATACTCAATCAACA
GAGCTTCTAACTCATCCATTTCCATGGGAGTAGGAATAGTTAGCTTGGTGTTATTAATGATCTTTTTACCCAATGCACGG
TATTTCATGAGCCTGGTTACTATCGGGTGCATACTCGTTAACTGTCTATACGACGCAGTTCTGCGTGTTGAACAATATTGTC
ACGAGGTACGAAGTGAATCATTTGGGTGTTTAAACGTTTCAGCTAAGTTTTCAATCAATTACGCTTCACGGTCAACCTTAC
GGCTATTACAGATTAACACCTAAACGCACACCACCGGAGTGAGCATATTTCAAATACCACGAGCGATGTTGTTTGCA
CGGTACATCGCCATCATTTTACCAGGAGTAACGATGTAGATTTCTTGCTTTACCTTCACGAATGGGCATAGCGAAACC
ACCACATACAACGTCAACCAATACGTCGTAGGAAACGAAGTCTAAATCTTGGTAAGCACCATTTTCTTCCAAAAAGTTAA
TGGCAGTGATAATACACGACCCAGCACAACCTACACCGGGTCTGGAACACAGATTCACGCATTTAACACCACGGGAAT
CCGGTCAGCATTACTTCTTCCAATTCAGGTCTTCTACAGCACCTTCTCAGCAGCCAAGTGAATACGGTAGTTTGGAGC
TTTAGAGTGACGATCAAAACGGGTAGAGTCAGCTTTGGGTGCGCAACCTACAAATTAGAATGCGTTGGCCCATTTACGCCA
TAGCTGCTAGAGTGTTTTGAGAGGTGGTAGATTACCAATACCGCCTTTGCCGTAGAATGCTATCTGTCTAATTCTTTTCG
TCGCTCAT

cpcA

>cpcA TARS_fil_Rr_000000002618:45755-46244
CTAGCTTAGAGCGTTGATAGCGTAGTCAATGTAGGTATTAGCTTCATTAGCAGCTTGACCGCCCAATCCGTGGTTAGCTT
TGATATATTTTAAAGCTTCCACATACCAACTGGGAGATAGGTTAAAGCACCGTTAATTTACGCCAAACAGCAATCAGG
AATTCATCCAAGGGACCTGTGCCACCAAGCACTAAGCTAAAGTAACAATGCGGAGGTAGTGACCAACGTCACGAGCACA
CTTGGATTATACCACGAGCATCAGCAGCATACTGAGCGCCGGGGGTGGAAGTGGTGAGGGAACTTTTGGTATACAGCGT
TAGTAGCACCATCAATTAATTTTGGAGATTAGCTGTTAAACCACGGGCGGCTTCCATGCTGGCAGCAGCAGTACAAAA
CGACCGTTAACTGCTTGTAATTCAGTATTGCTCAAAAAGCGTCCTTGGGTATCAGCAGCTGCAATAGCTTCGGTAATGGG
GTTTTTCAT

cpcB

>cpcB TARS_fil_Rr_000000002618:46357-46879
CGATGCAACAGCTGCTGCTGCACGATCAAAGTAGCCAGCCAATTACAGAACTATTTGACTGCAATCGCCCTTGGTAATAC
CTTTGGGATCGTTAACTATCTTGATTGTCAGCTTCTTTCAATTTGCCAAACACCAACAGCTACGGATGCACAGGAGTACCC
AAAGCTATATAGGTTTTCGCGTAAGCCATTTAAGCAACGGTCATCGAGAACACTAGCGTCACCAGCTAATGCAGCGTAGGT
AACATAGCGTAAGATGATTTCCATATCGCGCAGACAAGCAGCCATGCGACGGTTGGTGTAAGCATTACCACAGGAGCAA
TTAGTTGGGGCTGTTCTTCAAACAGTGCACGAGCAGCATCGGTAACGATAGCAGAAGCATTGCTTGTAATACGGTTAACA
GTGCTAGACGTTTTGCTGCCGGAAGCTACAACCTGCTGTTAAAGCGTCCAACCTGTTTCAGTGCTCAAAAATTCGCCTCTGGC
GTCAGCTTGGGAAACAACCTTGGAAAAATACATCTAATGTCTAT

(C)

Lake Trasimeno, unbinned *R. raciborskii* 16S rRNA

```
>16S_ribosomal_RNA_16S_rRNA::TARS_fil_Rr_000000002915:24-1509(+)
ACGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGTATGCTTAACACATGCAAGTCGAACGGGATGCTTAGGCATCTAGTGGCGGACGGGTGAGTA
ACGCGTGAGAATCTGGCTCCAGGTCGGGGATAACAGTTGGAAACGACTGCTAATACCGGATGTGCCGAGAGGTGAAAAGATTTATCGCCTGGAGATGAGCT
CGCGTCTGATTAGCTAGTTGGCGGTGTAAGGGACCACCAAGGCGTCGATCAGTAGCTGGTCTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCC
CAGACTCCTACGGGAGGCAGCAGTGGGGAATTTCCGCAATGGGCGAAAGCCTGACGGAGCAATACCGCGTGAGGGAGGAAGGCTCTTGGGTCGTAAACC
TCTTTTCTCAAGGAAGAAGAAAGTGACGGTACTTGAGGAATAAGCATCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGATGCAAGCGTTATCC
GGAATGATTGGGCGTAAAGGCTCTGCAGGTGGAAGTAAAGTCTGCTGTTAAAGAGTTTGGCTTAACCAAATAAAAGCGGTGGAAACTACAGAACTAGAG
TGCGGTAGGGGCAAAAGGAATTCCTGGTGTAGCGGTGAAATGCGTAGATATCAGGAAGAACACCGGTGGCGAAAGCGTTTGTCTAGACCGTAACTGACAC
TGAGGGACGAAAGCTAGGGGAGCGAATGGGATTAGATACCCAGTAGTCCTAGCCGTAAACGATGGATACTAGGCGTGGCTTGTATCGACCCGAGCCGTG
CCGGAGCTAACGCGTTAAGTATCCCGCTGGGGAGTACGCACGCAAGTGTGAAACTCAAAGGAATTGACGGGGGCCCGCACAGCGGTGGAGTATGTGGT
TTAATTCGATGCAACGCGAAGAACCCTTACCAAGGCTTGACATCCTGCGAATCCTGGTGAAAGCTGGGAGTGCCTTAGGGAGCGCAGAGACAGGTGGTGCA
TGGCTGTGTCAGCTCGTGTCTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTCGTTTTTAGTTGCCAGCATTAAAGTTGGGCACTCTAGAGAG
ACTGCCGGTGACAAACCGGAGGAAGGTGAGGATGACGTCAAGTCAGCATGCCCTTACGTCTTGGGCTACACACGTACTACAATGCTACGGACAGAGGGC
AGCGAGCCAGGGATGGCAAGCGAATCCAGAAACCGTAGCTCAGTTTCAAGTCAAGGCTGCAACTCGCCTTCGTGAAGGAGGAATCGCTAGTAATTGCAGG
TCAGCATACTGCAGTGAATTCGTTCCCGGGCCTTGTAACACCGCCCGTCACACCATGGAAGTTGGTCACGCCCAGATCATTACCCCAACCGAAAGGAG
GGGGATGCCTAAGGTAGGACTGGTGAAGTGAAGTTCGTAACAAGGTAGCCGTACCGGAAGGTGTGGCTGGATCACCTCCTTT
```

Table S3

Module pairwise completeness matrix produced by anvio. For every genome analysed in this work, completeness is estimated by computing the fraction of components KOs identified in the genome. In anvio, a threshold of 0.75 is adopted to assess the completeness of a module. A single M number or a combination of M numbers can be used for characterizing phenotypic features encoded in the genome (Kanehisa and Sato 2020). The genomes are reported in the order given in the phylogenomic tree (Fig. 1). Modules not represented in the strains are highlighted.

Module class

Module category

Central carbohydrate metabolism

Central carbohydrate metabolism

Other carbohydrate metabolism

Energy metabolism

Carbon fixation

Methane metabolism

Nitrogen metabolism

Sulfur metabolism

Phosphorus

ATP synthesis

Lipid metabolism

Lipid metabolism

Nucleotide metabolism

Pyrimidine metabolism

Amino acid metabolism

Cysteine and threonine metabolism

Cysteine and threonine metabolism

Branch-chain amino acid metabolism

Lysoine metabolism

Module	Module name	000001	000002	000003	000004	000005	000006	000007	000008	000009	000010	000011	000012	000013	000014	000015	000016	000017	000018	000019	000020	000021	000022	000023	000024	000025	000026	000027	000028	000029	000030	000031	000032	000033	000034	000035	000036	000037	000038	000039	000040	000041	000042	000043	000044	000045	000046	000047	000048	000049	000050	000051	000052	000053	000054	000055	000056	000057	000058	000059	000060	000061	000062	000063	000064	000065	000066	000067	000068	000069	000070	000071	000072	000073	000074	000075	000076	000077	000078	000079	000080	000081	000082	000083	000084	000085	000086	000087	000088	000089	000090	000091	000092	000093	000094	000095	000096	000097	000098	000099	000100	000101	000102	000103	000104	000105	000106	000107	000108	000109	000110	000111	000112	000113	000114	000115	000116	000117	000118	000119	000120	000121	000122	000123	000124	000125	000126	000127	000128	000129	000130	000131	000132	000133	000134	000135	000136	000137	000138	000139	000140	000141	000142	000143	000144	000145	000146	000147	000148	000149	000150	000151	000152	000153	000154	000155	000156	000157	000158	000159	000160	000161	000162	000163	000164	000165	000166	000167	000168	000169	000170	000171	000172	000173	000174	000175	000176	000177	000178	000179	000180	000181	000182	000183	000184	000185	000186	000187	000188	000189	000190	000191	000192	000193	000194	000195	000196	000197	000198	000199	000200	000201	000202	000203	000204	000205	000206	000207	000208	000209	000210	000211	000212	000213	000214	000215	000216	000217	000218	000219	000220	000221	000222	000223	000224	000225	000226	000227	000228	000229	000230	000231	000232	000233	000234	000235	000236	000237	000238	000239	000240	000241	000242	000243	000244	000245	000246	000247	000248	000249	000250	000251	000252	000253	000254	000255	000256	000257	000258	000259	000260	000261	000262	000263	000264	000265	000266	000267	000268	000269	000270	000271	000272	000273	000274	000275	000276	000277	000278	000279	000280	000281	000282	000283	000284	000285	000286	000287	000288	000289	000290	000291	000292	000293	000294	000295	000296	000297	000298	000299	000300	000301	000302	000303	000304	000305	000306	000307	000308	000309	000310	000311	000312	000313	000314	000315	000316	000317	000318	000319	000320	000321	000322	000323	000324	000325	000326	000327	000328	000329	000330	000331	000332	000333	000334	000335	000336	000337	000338	000339	000340	000341	000342	000343	000344	000345	000346	000347	000348	000349	000350	000351	000352	000353	000354	000355	000356	000357	000358	000359	000360	000361	000362	000363	000364	000365	000366	000367	000368	000369	000370	000371	000372	000373	000374	000375	000376	000377	000378	000379	000380	000381	000382	000383	000384	000385	000386	000387	000388	000389	000390	000391	000392	000393	000394	000395	000396	000397	000398	000399	000400	000401	000402	000403	000404	000405	000406	000407	000408	000409	000410	000411	000412	000413	000414	000415	000416	000417	000418	000419	000420	000421	000422	000423	000424	000425	000426	000427	000428	000429	000430	000431	000432	000433	000434	000435	000436	000437	000438	000439	000440	000441	000442	000443	000444	000445	000446	000447	000448	000449	000450	000451	000452	000453	000454	000455	000456	000457	000458	000459	000460	000461	000462	000463	000464	000465	000466	000467	000468	000469	000470	000471	000472	000473	000474	000475	000476	000477	000478	000479	000480	000481	000482	000483	000484	000485	000486	000487	000488	000489	000490	000491	000492	000493	000494	000495	000496	000497	000498	000499	000500	000501	000502	000503	000504	000505	000506	000507	000508	000509	000510	000511	000512	000513	000514	000515	000516	000517	000518	000519	000520	000521	000522	000523	000524	000525	000526	000527	000528	000529	000530	000531	000532	000533	000534	000535	000536	000537	000538	000539	000540	000541	000542	000543	000544	000545	000546	000547	000548	000549	000550	000551	000552	000553	000554	000555	000556	000557	000558	000559	000560	000561	000562	000563	000564	000565	000566	000567	000568	000569	000570	000571	000572	000573	000574	000575	000576	000577	000578	000579	000580	000581	000582	000583	000584	000585	000586	000587	000588	000589	000590	000591	000592	000593	000594	000595	000596	000597	000598	000599	000600	000601	000602	000603	000604	000605	000606	000607	000608	000609	000610	000611	000612	000613	000614	000615	000616	000617	000618	000619	000620	000621	000622	000623	000624	000625	000626	000627	000628	000629	000630	000631	000632	000633	000634	000635	000636	000637	000638	000639	000640	000641	000642	000643	000644	000645	000646	000647	000648	000649	000650	000651	000652	000653	000654	000655	000656	000657	000658	000659	000660	000661	000662	000663	000664	000665	000666	000667	000668	000669	000670	000671	000672	000673	000674	000675	000676	000677	000678	000679	000680	000681	000682	000683	000684	000685	000686	000687	000688	000689	000690	000691	000692	000693	000694	000695	000696	000697	000698	000699	000700	000701	000702	000703	000704	000705	000706	000707	000708	000709	000710	000711	000712	000713	000714	000715	000716	000717	000718	000719	000720	000721	000722	000723	000724	000725	000726	000727	000728	000729	000730	000731	000732	000733	000734	000735	000736	000737	000738	000739	000740	000741	000742	000743	000744	000745	000746	000747	000748	000749	000750	000751	000752	000753	000754	000755	000756	000757	000758	000759	000760	000761	000762	000763	000764	000765	000766	000767	000768	000769	000770	000771	000772	000773	000774	000775	000776	000777	000778	000779	000780	000781	000782	000783	000784	000785	000786	000787	000788	000789	000790	000791	000792	000793	000794	000795	000796	000797	000798	000799	000800	000801	000802	000803	000804	000805	000806	000807	000808	000809	000810	000811	000812	000813	000814	000815	000816	000817	000818	000819	000820	000821	000822	000823	000824	000825	000826	000827	000828	000829	000830	000831	000832	000833	000834	000835	000836	000837	000838	000839	000840	000841	000842	000843	000844	000845	000846	000847	000848	000849	000850	000851	000852	000853	000854	000855	000856	000857	000858	000859	000860	000861	000862	000863	000864	000865	000866	000867	000868	000869	000870	000871	000872	000873	000874	000875	000876	000877	000878	000879	000880	000881	000882	000883	000884	000885	000886	000887	000888	000889	000890	000891	000892	000893	000894	000895	000896	000897	000898	000899	000900	000901	000902	000903	000904	000905	000906	000907	000908	000909	000910	000911	000912	000913	000914	000915	000916	000917	000918	000919	000920	000921	000922	000923	000924	000925	000926	000927	000928	000929	000930	000931	000932	000933	000934	000935	000936	000937	000938	000939	000940	000941	000942	000943	000944	000945	000946	000947	000948	000949	000950	000951	000952	000953	000954	000955	000956	000957	000958	000959	000960	000961	000962	000963	000964	000965	000966	000967	000968	000969	000970	000971	000972	000973	000974	000975	000976	000977	000978	000979	000980	000981	000982	000983	000984	000985	000986	000987	000988	000989	000990	000991	000992	000993	000994	000995	000996	000997	000998	000999	001000	001001	001002	001003	001004	001005	001006	001007	001008	001009	001010	001011	001012	001013	001014	001015	001016	001017	001018	001019	001020	001021	001022	001023	001024	001025	001026	001027	001028	001029	001030	001031	001032	001033	001034	001035	001036	001037	001038	001039	001040	001041	001042	001043	001044	001045	001046	001047	001048	001049	001050	001051	001052	001053	001054	001055	001056	001057	001058	001059	001060	001061	001062	001063	001064	001065	001066	001067	001068	001069	001070	001071	001072	001073	001074	001075	001076	001077	001078	001079	001080	001081	001082	001083	001084	001085	001086	001087	001088	001089	001090	001091	001092	001093	001094	001095	001096	001097	001098	001099	001100	001101	001102	001103	001104	001105	001106	001107	0011
--------	-------------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	------

Kanehisa, Minoru, and Yoko Sato. "KEGG Mapper for Inferring Cellular Functions from Protein Sequences." *Protein Science: A Publication of the Protein Society* 29, no. 1 (January 2020): 28–35. <https://doi.org/10.1002/pro.3711>.

Table S4

Functional enrichment analysis carried out on the R1 and R2 clades using the pangenomic workflow by anvio (Eren *et al.*, 2021). The analysis reports the main differences in the COG20 functions. Only entries with adjusted_q values < 0.01 are reported, distinguishing the COG20 classified categories (black) from the unclassified or "general function" categories (gray). p_R1, p_R2, and N_R1, N_R2: fraction of assemblies with the COG20 function, and number of assemblies in the R1 and R2 clades, respectively.

COG20_FUNCTION	COG20_CATEGORY	Adjusted_q	associated_groups	p_R1	p_R2	N_R1	N_R2
Nicotinic acid phosphoribosyltransferase	Coenzyme transport and metabolism	0.00966	R1	0.8846	0.2727	26	11
Antitoxin component HicB of the HicAB toxin-antitoxin system	Defense mechanisms	0.00003	R1	0.9615	0.0909	26	11
Virulence-associated protein VapD, endoribonuclease (VapD)	Defense mechanisms	0.00323	R1	0.7308	0	26	11
Cu/Ag efflux pump CusA (CusA)	Inorganic ion transport and metabolism	0.00002	R1	0.9231	0	26	11
Uncharacterized phage-associated protein, contains DUF4065 domain (GepA)	Mobilome: prophages, transposons	0.00008	R1	0.8846	0	26	11
Cytosine/adenosine deaminase or related metal-dependent hydrolase (SsnA)	Nucleotide transport and metabolism General function prediction only	0.00001	R1	0.9615	0	26	11
Phospholipase/lecithinase/hemolysin (PDB:6JL1) Uncharacterized conserved protein YhjY, contains autotransporter beta-barrel domain (YhjY)	Nucleotide transport and metabolism General function prediction only	0.00001	R1	0.9615	0	26	11
Predicted ATP-dependent endonuclease of the OLD family, contains P-loop ATPase and TOPRIM domains (YbjD)	Replication, recombination and repair	0.00711	R1	0.6923	0	26	11
NADH/NAD ratio-sensing transcriptional regulator Rex (Rex)	Transcription	0.00169	R1	0.7692	0	26	11
Ribonuclease PH (Rph)	Translation, ribosomal structure and biogenesis	0.00000	R1	1	0	26	11
RNase P protein component (RnpA)	Translation, ribosomal structure and biogenesis	0.00069	R1	0.9231	0.1818	26	11
Pectate lyase (PelB)	Carbohydrate transport and metabolism	0.00349	R2	0.0385	0.6364	26	11
6-phosphogluconolactonase, cycloisomerase 2 family (Pgl) (PDB:1JOF) Ca2+-binding protein, RTX toxin-related	Carbohydrate transport and metabolism Secondary metabolites biosynthesis, transport and catabolism	0.00271	R2	0	0.5455	26	11
Chromosome segregation ATPase Smc (Smc) (PDB:5XG3) 5-methylcytosine-specific restriction endonuclease McrBC, GTP-binding regulatory subunit McrB (McrB)	Cell cycle control, cell division, chromosome partitioning Defense mechanisms	0.00966	R2	0	0.4545	26	11
Surface polysaccharide O-acetyltransferase WecH (WecH)	Cell wall/membrane/envelope biogenesis	0.00966	R2	0	0.4545	26	11
Glycosyltransferase involved in cell wall biosynthesis (WcaA) (PDB:5MLZ) Predicted SAM-dependedent methyltransferase	Cell wall/membrane/envelope biogenesis General function prediction only	0.00069	R2	0	0.6364	26	11
Pyrimidine reductase, riboflavin biosynthesis (RibD)	Coenzyme transport and metabolism	0.00271	R2	0	0.5455	26	11
Toxin component of the Txe-Axe toxin-antitoxin module, Txe/YoeB family	Defense mechanisms	0.00003	R2	0	0.8182	26	11
Antitoxin component HigA of the HigAB toxin-antitoxin module, contains an N-terminal HTH domain (HigA)	Defense mechanisms	0.00069	R2	0	0.6364	26	11
Acetate kinase (AckA)	Energy production and conversion	0.00069	R2	0	0.6364	26	11
NAD(P)-dependent dehydrogenase, short-chain alcohol dehydrogenase family (FabG) Cyclopropane fatty-acyl-phospholipid synthase and related methyltransferases (Cfa) Acyl transferase domain in polyketide synthase (PKS) enzymes (PksD)	Lipid transport and metabolism Lipid transport and metabolism Secondary metabolites biosynthesis, transport and catabolism	0.00271	R2	0	0.5455	26	11
Acyl carrier protein (AcpP) EntF, seryl-AMP synthase component of non-ribosomal peptide synthetase (EntF) Uncharacterized conserved protein, contains a NRPS condensation (elongation) domain	Lipid transport and metabolism Secondary metabolites biosynthesis, transport and catabolism	0.00271	R2	0	0.5455	26	11

Cytidine deaminase (Cdd)	Nucleotide transport and metabolism	0.00271	R2	0	0.5455	26	11
V8-like Glu-specific endopeptidase (eMpr)	Posttranslational modification, protein turnover, chaperones	0.00014	R2	0	0.7273	26	11
Periplasmic serine protease, S1-C subfamily, contain C-terminal PDZ domain (DegQ) Flp pilus assembly protein TadD, contains TPR repeats (TadD)	Posttranslational modification, protein turnover, chaperones Intracellular trafficking, secretion, and vesicular transport Extracellular structures	0.00966	R2	0	0.4545	26	11
Replication fork clamp-binding protein CrfC (dynamin-like GTPase family) (CrfC)	Replication, recombination and repair	0.00000	R2	0	1	26	11
DNA-binding transcriptional regulator, MerR family (SoxR)	Transcription	0.00966	R2	0	0.4545	26	11
tRNA isopentenyl-2-thiomethyl-A-37 hydroxylase MiaE (synthesis of 2-methylthio-cis-ribozeatin) (MiaE)	Translation, ribosomal structure and biogenesis	0.00002	R2	0.0769	1	26	11
Uncharacterized protein, contains SWIM-type Zn finger domain	Function unknown	0.00014	R1	1	0.2727	26	11
Predicted RNA binding protein YcfA, dsRBD-like fold, HicA-like mRNA interferase family (YcfA)	General function prediction only	0.00014	R1	1	0.2727	26	11
Predicted extracellular nuclease	General function prediction only	0.00069	R1	0.8077	0	26	11
Predicted nucleotidyltransferase	General function prediction only	0.00711	R1	0.6923	0	26	11
Uncharacterized conserved protein YcgG, contains conserved FPC and CPF motifs (YcgG)	Function unknown	0.00271	R2	0	0.5455	26	11
Uncharacterized CoxE-like protein, contains von Willebrand factor type A (vWA) domain (CoxE2)	Function unknown	0.00966	R2	0	0.4545	26	11
Uncharacterized membrane protein YhaH, DUF805 family (yhaH)	Function unknown	0.00966	R2	0	0.4545	26	11
Predicted carbamoyl transferase, NodU family	General function prediction only	0.00081	R2	0.0385	0.7273	26	11
Predicted sugar epimerase, cupin superfamily (CFF1)	General function prediction only	0.00271	R2	0	0.5455	26	11
SET domain-containing protein (function unknown) (SET)	General function prediction only	0.00966	R2	0	0.4545	26	11
Predicted antitoxin, contains HTH domain	General function prediction only	0.00966	R1	1	0.5455	26	11
Predicted extracellular nuclease Ca2+-binding protein, RTX toxin-related (PDB:1AF0)	General function prediction only Secondary metabolites biosynthesis, transport and catabolism	0.00966	R2	0	0.4545	26	11
GTPase SAR1 family domain (Gem1) (PDB:4C0J) Leucine-rich repeat (LRR) protein (LRR)	General function prediction only Transcription	0.00966	R2	0	0.4545	26	11

Adaptations to different geographic regions can be reflected by the presence of different defence mechanisms in the two groups, such as HicB and VapD (R1), and Txe and HigA (R2), which are widely present in bacterial genomes, often in association with mobile genetic elements, and involved in the control of stress response (Leplae *et al.*, 2011; Ren *et al.*, 2012; Thomet *et al.*, 2019; Fraikin *et al.*, 2020; Song *et al.*, 2021). Furthermore, the R1 group showed an exclusive enrichment part of the mobilome associated with prophages and transposons, indicative of specific interactions with phages in different geographical regions (Jaskulska *et al.*, 2022). Similarly, regarding metabolism, differences in functions in the carbohydrate transport and metabolism category can be linked to nutrient cycling and utilization of external organic compounds (Schwarz *et al.*, 2020) differentially developed in the R2 group. Regarding the inorganic ion transport, the exclusive presence of metal efflux pumps (CusA) in R1 is indicative of adaptations to specific resistance to excess of different toxic metal ions (Hahn *et al.*, 2013). Nevertheless, homeostasis maintenance for

essential and non-essential metals (Cavet *et al.*, 2003) in both groups is identified by the presence of other efflux pumps. Other differences between the two groups are linked to adaptations to specific environments like regulation of gene expression (category transcription and translation in R1 and R2), and maintenance of cell wall functionality (category cell wall/membrane/envelope biogenesis in R2). The remaining differences are indicative of a wide variety of metabolic and cellular processes that express different abilities to utilize different resources under different ecological environments or stress conditions, e.g. including, in the R2 group, sustained metabolite biosynthesis and survival under dark conditions favoured by Acetate kinase (AckA) (Ueda *et al.*, 2016).

- Cavet, J.S., Borrelly, G.P.M., and Robinson, N.J. (2003) Zn, Cu and Co in cyanobacteria: selective control of metal availability. *FEMS Microbiol Rev* **27**: 165–181.
- Eren, A.M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S.E., Schechter, M.S., et al. (2021) Community-led, integrated, reproducible multi-omics with anvi'o. *Nat Microbiol* **6**: 3–6.
- Fraikin, N., Goormaghtigh, F., and Van Melderren, L. (2020) Type II Toxin-Antitoxin Systems: Evolution and Revolutions. *J Bacteriol* **202**: 10.1128/jb.00763-19.
- Hahn, A., Stevanovic, M., Mirus, O., Lytvynenko, I., Pos, K.M., and Schleiff, E. (2013) The Outer Membrane TolC-like Channel HgdD Is Part of Tripartite Resistance-Nodulation-Cell Division (RND) Efflux Systems Conferring Multiple-drug Resistance in the Cyanobacterium *Anabaena* sp. PCC7120. *J Biol Chem* **288**: 31192–31205.
- Jaskulska, A., Šulčius, S., Kokociński, M., Koreivienė, J., Nájera, A.F., and Mankiewicz-Boczek, J. (2022) Cyanophage Distribution Across European Lakes of the Temperate-Humid Continental Climate Zone Assessed Using PCR-Based Genetic Markers. *Microb Ecol* **83**: 284–295.
- Leplae, R., Geeraerts, D., Hallez, R., Guglielmini, J., Drèze, P., and Van Melderren, L. (2011) Diversity of bacterial type II toxin–antitoxin systems: a comprehensive search and functional analysis of novel families. *Nucleic Acids Res* **39**: 5513–5525.
- Ren, D., Walker, A.N., and Daines, D.A. (2012) Toxin-antitoxin loci vapBC-1 and vapXD contribute to survival and virulence in nontypeable *Haemophilus influenzae*. *BMC Microbiol* **12**: 263.
- Schwarz, A., Walther, J., Geib, D., Witthohn, M., Strieth, D., Ulber, R., and Muffler, K. (2020) Influence of heterotrophic and mixotrophic cultivation on growth behaviour of terrestrial cyanobacteria. *Algal Res* **52**: 102125.
- Song, Y., Zhang, S., Luo, G., Shen, Y., Li, C., Zhu, Y., et al. (2021) Type II Antitoxin HigA Is a Key Virulence Regulator in *Pseudomonas aeruginosa*. *ACS Infect Dis* **7**: 2930–2940.
- Thomet, M., Trautwetter, A., Ermel, G., and Blanco, C. (2019) Characterization of HicAB toxin-antitoxin module of *Sinorhizobium meliloti*. *BMC Microbiol* **19**: 10.
- Ueda, S., Kawamura, Y., Iijima, H., Nakajima, M., Shirai, T., Okamoto, M., et al. (2016) Anionic metabolite biosynthesis enhanced by potassium under dark, anaerobic conditions in cyanobacteria. *Sci Rep* **6**: 32354.