

## SUPPLEMENTARY TABLE

### Analysis of Biological Diversity between the cyanobacteria *Cylindrospermopsis* and *Sphaerospermopsis*

**Reno Nooblath**

Federal University of Pará (UFPA), [reno.bioinfo@gmail.com](mailto:reno.bioinfo@gmail.com), +55 (91) 98208-1316, Augusto Corrêa Street, 01 – Bairro Guamá – CEP 66075-110 Belém–Pará – Amazônia, Brasil.

**Daniel Gomes**

Federal University of Pará (UFPA), [daniel.h.gomes18@gmail.com](mailto:daniel.h.gomes18@gmail.com), +55 (91) 98126-7143, Augusto Corrêa Street, 01 – Bairro Guamá – CEP 66075-110 Belém – Pará – Amazônia, Brasil.

**Vinícius Abreu**

Federal University of Pará (UFPA), [vabreuufpa@gmail.com](mailto:vabreuufpa@gmail.com), +55 (91) 99810-7845, Augusto Corrêa Street, 01 – Bairro Guamá – CEP 66075-110 Belém –Pará – Amazônia, Brasil.

**Sintia Silva de Almeida**

Federal University of Pará (UFPA), [sintiaalmeida@gmail.com](mailto:sintiaalmeida@gmail.com), +55 (91) 98218-2015, Augusto Corrêa Street, 01 – Bairro Guamá – CEP 66075-110 Belém – Pará – Amazônia, Brasil.

**Supplementary Table S1 . Origin of the *Cylindrospermopsis* and *Sphaerospermopsis* strains used in this study and their taxonomic identification on NCBI, organized in ascending order based on the number of scaffolds.**

| Organism                    | Taxonomy ID | Scaffolds | Origin of Sample | Size    | GC%    |
|-----------------------------|-------------|-----------|------------------|---------|--------|
| <b>C. raciborskii KLL07</b> | 683357      | 1         | Israel           | 3.80818 | 40.2   |
| C. raciborskii N8           | 1524929     | 1         | China            | 3.85717 | 40.1   |
| C. raciborskii Cr2010       | 2108055     | 1         | Netherlands      | 3.75785 | 40.2   |
| C. curvispora GIHE G1       | 2666332     | 2         | South Korea      | 4.05721 | 40.117 |
| C. raciborskii DSH          | 3117726     | 2         | China            | 4.0     | 40     |
| C. raciborskii GIHE 2018    | 2588993     | 3         | South Korea      | 3.62982 | 40.2   |
| C. raciborskii CS-505       | 533240      | 6         | Australia        | 3.9     | 40     |
| R. brookii D9               | 533247      | 47        | Brazil           | 3.2     | 40     |
| C. raciborskii CENA 302     | 1170768     | 58        | Brazil           | 3.5     | 40     |
| C. raciborskii CENA 303     | 1170769     | 77        | Brazil           | 3.4     | 40.5   |
| C. raciborskii MVCC14       | 940191      | 99        | Uruguay          | 3.6     | 40.1   |
| C. raciborskii KL1          | 2787621     | 100       | USA              | 3.7     | 40     |
| C. raciborskii PAMP2012     | 2969974     | 134       | Brazil           | 3.2     | 40     |
| C. sp. CR12                 | 1747196     | 136       | Malaysia         | 3.7     | 40     |
| C. raciborskii MVCC 19      | 940192      | 155       | Uruguay          | 3.5     | 40     |
| C. raciborskii CS508        | 533243      | 162       | Australia        | 3.6     | 40     |
| C. raciborskii CYRF         | 2021698     | 166       | Brazil           | 4.2     | 40     |
| C. raciborskii 1523720      | 77022       | 173       | India            | 3.2     | 40     |
| C. raciborskii CYLP         | 2021699     | 188       | Brazil           | 4.2     | 40     |
| C. raciborskii CHAB3438     | 1480071     | 189       | China            | 3.5     | 40.5   |
| C. raciborskii ITEPA-A1     | 1810942     | 195       | Brazil           | 3.6     | 40     |
| S. torques-reginae ITEP-024 | 984208      | 1         | Brazil           | 5.3     | 37.5   |

|                         |         |     |             |     |      |
|-------------------------|---------|-----|-------------|-----|------|
| S. kisseleviana NIES-73 | 1973480 | 2   | Japan       | 5.4 | 37.5 |
| S. sp. SIO1G1           | 2607814 | 38  | Puerto Rico | 5.2 | 36.5 |
| S. sp. FACHB-1194       | 2692862 | 108 | China       | 5.4 | 37.5 |
| S. sp. LEGE 08334       | 1828651 | 119 | Mexico      | 5.5 | 37   |
| S. sp. LEGE 00249       | 1380707 | 177 | Portugal    | 5.3 | 37.5 |

**Supplementary Table S2. Proposed function of the proteins encoded by the saxitoxin biosynthetic gene cluster from the *Sphaerospermopsis* group.**

| Protein     | Amino acids | Product                                  | Strain                               | Organism                             | Identity (%) | Function                                  | Acession number |
|-------------|-------------|--|--------------------------------------|--------------------------------------|--------------|---|-----------------|
| <b>sxt</b>  | 266         | Phytanoyl-CoA dioxygenase                | <i>S. torques reginae</i><br>ITEP024 | <i>S. sp. LEGE 08334</i>             | 100          | phytanoyl-CoA dioxygenase                 | WP_194055575.1  |
| <b>sxtJ</b> | 747         | carbamoyltransferase                     | <i>S. torques reginae</i><br>ITEP024 | <i>S. sp. LEGE 08334</i>             | 100          | carbamoyltransferase                      | MBE9056840.1    |
| <b>sxtK</b> | 54          | DUF5989                                  | <i>S. torques reginae</i><br>ITEP024 | <i>Nostoc sp. ChiQUE01a</i>          | 89           | DUF5989                                   | MDZ8239473.1    |
| <b>sxtJ</b> | 747         | carbamoyltransferase                     | <i>S. torques reginae</i><br>ITEP024 | <i>S. torques reginae</i><br>ITEP024 | 100          | carbamoyltransferase                      | QYX30560.1      |
| <b>sxtK</b> | 54          | DUF5989                                  | <i>S. torques reginae</i><br>ITEP024 | <i>Nostoc sp. ChiQUE01a</i>          | 89           | DUF5989                                   | MDZ8239473.1    |
| <b>sxt</b>  | 747         | Nodulation protein noLO                  | <i>S. torques reginae</i><br>ITEP024 | <i>S. sp. LEGE 08334</i>             | 100          | SxtJ family membrane protein              | WP_194055576    |
| <b>sxt</b>  | 108         | hypothetical protein                     | <i>S. torques reginae</i><br>ITEP024 | <i>S. sp. LEGE 08334</i>             | 100          | macrolide family glycosyltransferase      | WP_194055579.1  |
| <b>sxt</b>  | 385         | Carbamoyl-phosphate synthase small chain | <i>S. torques reginae</i><br>ITEP024 | <i>S. sp. LEGE 08334</i>             | 99           | glutamine-hydrolyzing carbamoyl-phosphate | WP_194055594.1  |

**Supplementary Table S3. Proposed function of the proteins encoded by saxitoxin biosynthetic gene cluster in the genomes of the *C. raciborskii* strains *R. brookii* D9 , MVCC14, MVCC19,CENA302, ITEP-A1, and the CYRF.**

| Protein     | Amino acids | Product            | Strains                       | Organism                              | Identity(%) | Function                            | Acession number |
|-------------|-------------|--------------------|-------------------------------|---------------------------------------|-------------|-------------------------------------|-----------------|
| <b>SxtA</b> | 1245        | PKS                | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3              | 99          | Polyketide synthase-related protein | ABI75094.1      |
|             | 1245        |                    | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3              | 100         | Polyketide synthase-related protein | ABI75094.1      |
|             | 1245        |                    | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3              | 100         | Polyketide synthase-related protein | ABI75094.1      |
|             | 1245        |                    | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3              | 100         | Polyketide synthase-related protein | ABI75094.1      |
|             | 1245        |                    | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3              | 99          | Polyketide synthase-related protein | ABI75094.1      |
|             | 860         |                    | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3              | 99          | polyketide synthase-related protein | ABI75094.1      |
| <b>SxtB</b> | 318         | Cytidine deaminase | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3              | 100         | Cytidine deaminase                  | ABI75093.1      |
|             | 318         |                    | <i>C. raciborskii</i> CENA302 | <i>R. brookii</i> D9                  | 100         | SxtB                                | EFA72782.1      |
|             | 318         |                    | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3              | 100         | Cytidine deaminase                  | ABI75093.1      |
|             | 318         |                    | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3              | 100         | Cytidine deaminase                  | ABI75093.1      |
|             | 318         |                    | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3              | 89          | Cytidine deaminase                  | WP_254011061.1  |
|             | 318         |                    | <i>C. raciborskii</i> CYRF    | <i>Heteroscytonema crispum</i> UCFS10 | 92          | SxtB                                | AYN62266.1      |

|             |     |                                 |                               |                                       |     |  |            |
|-------------|-----|---------------------------------|-------------------------------|---------------------------------------|-----|--|------------|
| <b>SxtC</b> | 117 | Regulatory                      | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3              | 100 | SxtC   | ABI75092.1 |
|             | 94  |                                 | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3              | 100 | SxtC   | ABI75092.1 |
|             | 94  |                                 | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3              | 100 | sxtC   | ABI75092.1 |
|             | 117 |                                 | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3              | 100 | sxtC   | ABI75092.1 |
|             | 94  |                                 | <i>C. raciborskii</i> MVCC19  | <i>Heteroscytonema crispum</i> UCFS10 | 91  | sxtC   | AYN62265.1 |
|             | 117 |                                 | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3              | 100 | sxtC   | AYN62265.1 |
| <b>SxtD</b> | 252 | Sterole desaturase-like protein | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3              | 99  | Sterole desaturase   | ABI75089.1 |
|             | 252 |                                 | <i>C. raciborskii</i> CENA302 | <i>R. brookii</i> D9                  | 100 | SxtD   | EFA72785.1 |
|             | 252 |                                 | <i>C. raciborskii</i> ITEP-A1 | <i>R. brookii</i> D9                  | 100 | SxtD   | EFA72785.1 |
|             | 252 |                                 | <i>C. raciborskii</i> MVCC14  | <i>R. brookii</i> D9                  | 100 | SxtD   | EFA72785.1 |
| <b>SxtE</b> | 128 | Unknown protein                 | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3              | 99  | SxtE   | ABI75095.1 |
|             | 45  |                                 | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3              | 100 | SxtE   | ABI75095.1 |
|             | 45  |                                 | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3              | 100 | SxtE   | ABI75095.1 |
|             | 128 |                                 | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3              | 100 | SxtE   | ABI75095.1 |
|             | 127 |                                 | <i>C. raciborskii</i> MVCC19  | <i>Microseira wollei</i>              | 88  | SxtE   | ACZ26228.1 |
|             | 120 |                                 | <i>C. raciborskii</i> CYRF    | <i>Microseira wollei</i>              | 92  | SxtE   | ACZ26228.1 |
| <b>SxtF</b> | 471 | MATE                            | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3              | 99  | Sodium-driven multidrug and toxic compound extrusion protein | ABI75096.1 |
|             | 471 |                                 | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3              | 100 | Sodium-driven multidrug and toxic                            | ABI75096.1 |

|             |     |                              |                               |                                       |     |  |            |
|-------------|-----|------------------------------|-------------------------------|---------------------------------------|-----|--|------------|
|             |     |                              |                               |                                       |     | compound extrusion protein                                   |            |
|             | 471 |                              | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3              | 100 | Sodium-driven multidrug and toxic compound extrusion protein | ABI75096.1 |
|             | 471 |                              | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3              | 100 | Sodium-driven multidrug and toxic compound extrusion protein | ABI75096.1 |
|             | 471 |                              | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3              | 100 | Sodium-driven multidrug and toxic compound extrusion protein | ABI75096.1 |
|             | 220 |                              | <i>C. raciborskii</i> CYRF    | <i>Heteroscytonema crispum</i> UCFS10 | 83  | SxtM1  | AYN62272.1 |
| <b>SxtG</b> | 377 | Amidinotransferase           | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3              | 100 | Amidinotransferase   | ABI75097.1 |
|             | 377 |                              | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3              | 100 | Amidinotransferase   | ABI75097.1 |
|             | 377 |                              | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3              | 100 | Amidinotransferase   | ABI75097.1 |
|             | 377 |                              | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3              | 100 | Amidinotransferase   | ABI75097.1 |
|             | 377 |                              | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3              | 100 | Amidinotransferase   | ABI75097.1 |
|             | 377 |                              | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3              | 100 | Amidinotransferase   | ABI75097.1 |
| <b>SxtH</b> | 334 | Phenylpropionate dioxygenase | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3              | 100 | Phenylpropionate dioxygenase                                 | ABI75098.1 |
|             | 334 |                              | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3              | 100 | Phenylpropionate dioxygenase                                 | ABI75098.1 |
|             | 334 |                              | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3              | 100 | Phenylpropionate dioxygenase                                 | ABI75098.1 |
|             | 334 |                              | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3              | 100 | Phenylpropionate dioxygenase                                 | ABI75098.1 |
|             | 334 |                              | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3              | 100 | Phenylpropionate dioxygenase                                 | ABI75098.1 |
|             | 334 |                              | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3              | 100 | Phenylpropionate dioxygenase                                 | ABI75098.1 |

|             |     |                      |                               |                              |     |  |            |
|-------------|-----|----------------------|-------------------------------|------------------------------|-----|--|------------|
| <b>SxtI</b> | 612 | Carbamoyltransferase | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3     | 100 | NodU/CmcH-related carbamoyltransferase | ABI75099.1 |
|             | 596 |                      | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3     | 99  | NodU/CmcH-related carbamoyltransferase | ABI75099.1 |
|             | 612 |                      | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3     | 100 | NodU/CmcH-related carbamoyltransferase | ABI75099.1 |
|             | 612 |                      | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3     | 99  | NodU/CmcH-related carbamoyltransferase | ABI75099.1 |
|             | 612 |                      | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3     | 99  | NodU/CmcH-related carbamoyltransferase | ABI75099.1 |
| <b>SxtJ</b> | 147 | Regulatory           | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3     | 100 | SxtJ                                   | ABI75100.1 |
|             | 134 |                      | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3     | 99  | SxtJ                                   | ABI75100.1 |
|             | 147 |                      | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> MVCC14 | 100 | SxtJ                                   | OHY34958.1 |
|             | 147 |                      | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> MVCC14 | 100 | SxtJ                                   | OHY34958.1 |
|             | 147 |                      | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> MVCC14 | 100 | SxtJ                                   | OHY34958.1 |
| <b>SxtK</b> | 54  | Unknown protein      | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3     | 100 | SxtK                                   | ABI75101.1 |
|             | 54  |                      | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3     | 100 | SxtK                                   | ABI75101.1 |
|             | 54  |                      | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3     | 100 | SxtK                                   | ABI75101.1 |
|             | 54  |                      | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3     | 100 | SxtK                                   | ABI75101.1 |
|             | 54  |                      | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3     | 100 | SxtK                                   | ABI75101.1 |

|             |     |             |                               |                              |     |  |            |
|-------------|-----|-------------|-------------------------------|------------------------------|-----|--|------------|
|             |     |             |                               |                              |     |  |            |
| <b>SxtL</b> | 435 | GDSL-lipase | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3     | 100 | GDSL-lipase  | ABI75102.1 |
|             | 407 |             | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3     | 99  | GDSL-lipase  | ABI75102.1 |
|             | 420 |             | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3     | 99  | GDSL-lipase  | ABI75102.1 |
|             | 419 |             | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3     | 99  | GDSL-lipase  | ABI75102.1 |
|             | 419 |             | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3     | 99  | GDSL-lipase  | ABI75102.1 |
|             | 432 |             | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3     | 99  | GDSL-lipase  | ABI75102.1 |
| <b>SxtM</b> | 475 | MATE        | <i>R. brookii</i> D9          | <i>C. raciborskii</i> MVCC14 | 99  | SxtM, partial  | AFQ99040.1 |
|             | 429 |             | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> MVCC14 | 99  | SxtM, partial  | AFQ99040.1 |
|             | 482 |             | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> T3     | 99  | Sodium-driven multidrug and toxic compound extrusion protein | ABI75103.1 |
|             | 475 |             | <i>C. raciborskii</i> MVCC14  | <i>R. brookii</i> D9         | 99  | SxtSUL   | EFA72771.1 |
|             | 475 |             | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> MVCC14 | 99  | SxtM   | AFQ99040.1 |



|             |     |                        |                               |                               |     |  |            |
|-------------|-----|------------------------|-------------------------------|-------------------------------|-----|--|------------|
|             | 182 |                        | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3      | 99  | Sodium-driven multidrug and toxic compound extrusion protein | ABI75103.1 |
| <b>SxtN</b> | 302 | Sulfotransferase       | <i>R. brookii</i> D9          | <i>Microseira wollei</i>      | 89  | SxtSUL   | ACG63834.1 |
|             | 302 |                        | <i>C. raciborskii</i> CENA302 | <i>R. brookii</i> D9          | 100 | SxtSUL   | EFA72773.1 |
|             | 302 |                        | <i>C. raciborskii</i> MVCC14  | <i>R. brookii</i> D9          | 100 | SxtSUL   | EFA72773.1 |
|             | 270 |                        | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3      | 100 | SxtSUL   | EFA72773.1 |
| <b>SxtO</b> | 200 | Adenylylsulfate kinase | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3      | 100 | Adenylylsulfate kinase                                       | ABI75115.1 |
|             | 159 |                        | <i>C. raciborskii</i> CENA302 | <i>R. brookii</i> D9          | 100 | Adenylylsulfate kinase                                       | EFA72764.1 |
|             | 200 |                        | <i>C. raciborskii</i> MVCC14  | <i>R. brookii</i> D9          | 100 | Adenylylsulfate kinase                                       | EFA72764.1 |
|             | 200 |                        | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> MVCC14  | 100 | Adenylylsulfate kinase                                       | OHY34903.1 |
|             | 200 |                        | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> CENA302 | 100 | Adenylylsulfate kinase                                       | OPH09277   |
|             | 200 |                        | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3      | 100 | Adenylylsulfate kinase                                       | ABI75115.1 |
|             | 200 |                        | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3      | 100 | Adenylylsulfate kinase                                       | ABI75115.1 |
| <b>SxtP</b> | 408 | RTX toxin              | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3      | 99  | Putative saxitoxin-binding protein                           | ABI75114.1 |
|             | 408 |                        | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3      | 100 | Putative saxitoxin-binding protein                           | ABI75114.1 |
|             | 408 |                        | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3      | 100 | Putative saxitoxin-binding protein                           | ABI75114.1 |
|             | 408 |                        | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3      | 100 | Putative saxitoxin-binding protein                           | ABI75114.1 |
|             | 408 |                        | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3      | 100 | Putative saxitoxin-binding protein                           | ABI75114.1 |

|             |     |                              |                               |                          |     |                              |            |
|-------------|-----|------------------------------|-------------------------------|--------------------------|-----|------------------------------|------------|
|             |     |                              |                               |                          |     |                              |            |
| <b>SxtQ</b> | 258 | Unknown protein              | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3 | 99  | SxtQ                         | ABI75113.1 |
|             | 239 |                              | <i>C. raciborskii</i> CENA302 | <i>R. brookii</i> D9     | 100 | SxtQ                         | EFA72766.1 |
|             | 239 |                              | <i>C. raciborskii</i> MVCC14  | <i>R. brookii</i> D9     | 100 | SxtQ                         | EFA72766.1 |
|             | 239 |                              | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3 | 100 | SxtQ                         | EFA72766.1 |
|             | 239 |                              | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3 | 100 | SxtQ                         | EFA72766.1 |
| <b>SxtR</b> | 258 | Acyl transferase             | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3 | 100 | Acyl-CoA N-acyltransferase   | ABI75112.1 |
|             | 258 |                              | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> T3 | 100 | Acyl-CoA N-acyltransferase   | ABI75112.1 |
|             | 258 |                              | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> T3 | 100 | Acyl-CoA N-acyltransferase   | ABI75112.1 |
|             | 258 |                              | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3 | 100 | Acyl-CoA N-acyltransferase   | ABI75112.1 |
|             | 258 |                              | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3 | 100 | Acyl-CoA N-acyltransferase   | ABI75112.1 |
| <b>SxtS</b> | 241 | Phytanoyl-CoA dioxygenase    | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3 | 99  | Phytanoyl-CoA dioxygenase    | ABI75110.1 |
|             | 241 |                              | <i>C. raciborskii</i> CENA302 | <i>R. brookii</i> D9     | 100 | SxtS                         | EFA72768.1 |
|             | 241 |                              | <i>C. raciborskii</i> MVCC14  | <i>R. brookii</i> D9     | 100 | SxtS                         | EFA72768.1 |
|             | 241 |                              | <i>C. raciborskii</i> MVCC19  | <i>C. raciborskii</i> T3 | 99  | Phytanoyl-CoA dioxygenase    | ABI75110   |
|             | 241 |                              | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3 | 100 | Phytanoyl-CoA dioxygenase    | ABI75110   |
| <b>SxtT</b> | 334 | Phenylpropionate dioxygenase | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3 | 99  | Phenylpropionate dioxygenase | ABI75109.1 |
|             | 334 |                              | <i>C. raciborskii</i> CENA302 | <i>R. brookii</i> D9     | 100 | SxtT                         | EFA72769.1 |

|                |     |                          |                               |                          |     |                                   |            |
|----------------|-----|--------------------------|-------------------------------|--------------------------|-----|-----------------------------------|------------|
|                | 334 |                          | <i>C. raciborskii</i> MVCC14  | <i>R. brookii</i> D9     | 100 | SxtT                              | EFA72769.1 |
|                | 334 |                          | <i>C. raciborskii</i> MVCC19  | <i>R. brookii</i> D9     | 100 | SxtT                              | EFA72769.1 |
|                | 334 |                          | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3 | 100 | Phenylpropionate dioxygenase      | EFA72769.1 |
| <b>SxtU</b>    | 248 | Alcohol dehydrogenase    | <i>R. brookii</i> D9          | <i>C. raciborskii</i> T3 | 100 | Short-chain alcohol dehydrogenase | ABI75108.1 |
|                | 248 |                          | <i>C. raciborskii</i> CENA302 | <i>R. brookii</i> D9     | 100 | SxtU                              | EFA72770.1 |
|                | 248 |                          | <i>C. raciborskii</i> MVCC14  | <i>R. brookii</i> D9     | 100 | SxtU                              | EFA72770.1 |
|                | 248 |                          | <i>C. raciborskii</i> MVCC19  | <i>R. brookii</i> D9     | 93  | SxtU                              | EFA72770.1 |
|                | 249 |                          | <i>C. raciborskii</i> CYRF    | <i>C. raciborskii</i> T3 | 100 | Short-chain alcohol dehydrogenase | ABI75108.1 |
| <b>SxtDIOX</b> | 334 | (2Fe-2S)-binding protein | <i>R. brookii</i> D9          | <i>Microseira wollei</i> | 86  | SxtDIOX                           | ACG63835.1 |
|                | 334 |                          | <i>C. raciborskii</i> CENA302 | <i>R. brookii</i> D9     | 99  | SxtDIOX                           | EFA72772.1 |
|                | 334 |                          | <i>C. raciborskii</i> MVCC14  | <i>R. brookii</i> D9     | 99  | SxtDIOX                           | EFA72772.1 |

**Supplementary Table S4. Proposed function of the proteins encoded by cylindrospermopsin biosynthetic gene cluster in the genomes of the *C. raciborskii* strains CS-505, CR12, DSH and CHAB-3438.**

| Protein     | Amino acids | Proposed function   | Strain                         | Organism                     | Identity(%) | Function            | Acession number |
|-------------|-------------|---------------------|--------------------------------|------------------------------|-------------|---------------------|-----------------|
| <b>CyrA</b> | 391         | Amidino transferase | <i>C. raciborskii</i> CS-505   | <i>C. raciborskii</i> AWT205 | 100         | Amidino transferase | ABX60160.1      |
|             | 391         |                     | <i>C. raciborskii</i> CR12     | <i>C. raciborskii</i> AWT205 | 99          | Amidino transferase | ABX60160.1      |
|             | 391         |                     | <i>C. raciborskii</i> CHAB3438 | <i>C. raciborskii</i> AWT205 | 100         | Amidino transferase | ABX60160.1      |
| <b>CyrB</b> | 2917        | NRPS/PKS            | <i>C. raciborskii</i> CS-505   | <i>C. raciborskii</i> AWT205 | 100         | Mixed NRPS/PKS      | ABX60161.1      |
|             | 2917        |                     | <i>C. raciborskii</i> CR12     | <i>C. raciborskii</i>        | 100         | Mixed NRPS/PKS      | ABX60161.1      |

|      |      |                       |                                   |   |     |                                |            |
|------|------|-----------------------|-----------------------------------|---|-----|--------------------------------|------------|
|      |      |                       |                                   | AWT205  |     |                                |            |
|      | 2917 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>Raphidiopsis</i><br><i>curvata</i> HB1         | 100 | Mixed NRPS/PKS                 | AHN91606.1 |
| CyrC | 1667 | PKS                   | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>AWT205                   | 99  | Polyketide synthase            | ABX60163.1 |
|      | 1667 |                       | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i><br>AWT205                   | 99  | Polyketide synthase            | ABX60163.1 |
|      | 1667 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>Raphidiopsis</i><br><i>curvata</i> HB1         | 100 | Polyketide synthase            | AHN91608.1 |
|      | 1667 |                       |                                   |   |     |                                |            |
| CyrD | 1851 | PKS                   | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>AWT205                   | 100 | Polyketide synthase            | ABX60152.1 |
|      | 1876 |                       | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i><br>AWT205                   | 99  | Polyketide synthase            | ABX60152.1 |
|      | 1876 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>Raphidiopsis</i><br><i>curvata</i> HB1         | 100 | Polyketide synthase            | AHN91609.1 |
| CyrE | 1888 | PKS                   | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>AWT205                   | 99  | Polyketide synthase            | ABX60162.1 |
|      | 1888 |                       | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i><br>AWT205                   | 99  | Polyketide synthase            | ABX60162.1 |
|      | 1888 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>Raphidiopsis</i><br><i>curvata</i> HB1         | 100 | Polyketide synthase            | AHN91607.1 |
| CyrF | 1357 | PKS                   | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>AWT205                   | 100 | Polyketide synthase            | ABX60153.1 |
|      | 1357 |                       | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i><br>AWT205                   | 99  | Polyketide synthase            | ABX60153.1 |
|      | 1357 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>Raphidiopsis</i><br><i>curvata</i><br>CHAB1150 | 100 | Polyketide synthase            | AFC35246.1 |
| CyrG | 478  | Uracil ring formation | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>AWT205                   | 100 | Putative uracil ring formation | ABX60154.1 |
|      | 478  |                       | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i><br>AWT205                   | 99  | Putative uracil ring formation | ABX60154.1 |

|             |     |                       |                                   |   |     |  |                |
|-------------|-----|-----------------------|-----------------------------------|---|-----|--|----------------|
|             | 478 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>Raphidiopsis</i><br><i>curvata</i> HB1         | 99  | Putative uracil ring formation                   | AHN91611.1     |
| <b>CyrI</b> | 276 | Hydroxylation of C-7  | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>CHAB3438                 | 100 | Putative 2-oxoglutarate-dependent iron oxygenase | AHN91588.1     |
|             | 276 |                       | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>CR12                     | 99  | Putative 2-oxoglutarate-dependent iron oxygenase | WP_057178790   |
|             | 278 |                       | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i><br>CHAB3438                 | 99  | Putative 2-oxoglutarate-dependent iron oxygenase | AHN91588.1     |
|             | 276 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>C. raciborskii</i><br>CR12                     | 99  | Putative 2-oxoglutarate-dependent iron oxygenase | WP_057178790.1 |
| <b>CyrK</b> | 465 | Exporter              | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>AWT205                   | 99  | Multidrug exporter MatE                          | ABX60156.1     |
|             | 465 |                       | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> CS-505                      | 99  | MATE family efflux transporter                   | OBU75962.1     |
|             | 451 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>Raphidiopsis</i><br><i>curvata</i><br>CHAB1150 | 100 | Multidrug exporter MatE                          | AFC35248.1     |
| <b>CyrL</b> | 249 | Transposase           | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>AWT205                   | 100 | Transposase                                      | ABX60157.1     |
|             | 153 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>C. raciborskii</i><br>CR12                     | 98  | Transposase                                      | WP_161808566.1 |
| <b>CyrH</b> | 476 | Uracil ring formation | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>AWT205                   | 100 | Amidohydrolase                                   | ABX60158.1     |
|             | 476 |                       | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i><br>AWT205                   | 99  | Amidohydrolase                                   | ABX60158.1     |
|             | 476 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>Raphidiopsis</i><br><i>curvata</i><br>CHAB1150 | 100 | Amidohydrolase                                   | AHN91614.1     |
| <b>CyrJ</b> | 259 | Sulfotransferase      | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i><br>AWT205                   | 100 | Putative sulfotransferase                        | ABX60159.1     |
|             | 259 |                       | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i><br>AWT205                   | 99  | Putative sulfotransferase                        | ABX60159.1     |
|             | 261 |                       | <i>C. raciborskii</i><br>CHAB3438 | <i>Raphidiopsis</i><br><i>curvata</i> HB1         | 100 | Putative sulfotransferase                        | AHN91615       |

|             |     |                         |                              |                              |     |                        |            |
|-------------|-----|-------------------------|------------------------------|------------------------------|-----|------------------------|------------|
| <b>CyrM</b> | 105 | Transposase             | <i>C. raciborskii</i> CS-505 | <i>C. raciborskii</i> AWT205 | 97  | Transposase protein    | ABX60157.1 |
| <b>CyrN</b> | 219 | Adenylyl-sulfate kinase | <i>C. raciborskii</i> CS-505 | <i>C. raciborskii</i> AWT205 | 100 | Adenylylsulfate kinase | ABX60164.1 |
| <b>CyrO</b> | 515 | Regulator               | <i>C. raciborskii</i> CS-505 | <i>C. raciborskii</i> AWT205 | 99  | Hypothetical protein   | AHN91599.1 |

**Supplementary Table S5. Proposed function of the proteins encoded by anabaenopeptin biosynthetic gene cluster of the Sphaerospermopsis group.**

| Protein      | Amino acids | Product   | Strain                         | Organism                             | Identity (%) | Function                         | Acession number |
|--------------|-------------|---|--------------------------------|--------------------------------------|--------------|----------------------------------|-----------------|
| <b>aptA1</b> | 4161        | Siderophore biosynthesis non-ribosomal peptide synthetase modules | <i>S. torques</i> ITEP024      | <i>Aphanizomenon flos-aquae</i>      | 90           | non-ribosomal peptide synthetase | WP_190382984.1  |
|              | 4109        |   | <i>S. Kisseleviana</i> NIES 73 | <i>Dolichospermum</i> sp. LEGE 00246 | 83           |                                  | WP_1939621      |
| <b>aptA2</b> | 2208        | Polyketide synthase   | <i>S. torques</i> ITEP024      | <i>S. LEGE</i> 08334                 | 97           | Polyketide synthase              | WP_194056036.1  |
|              | 2226        |   | <i>S. Kisseleviana</i> NIES 73 | <i>S. FACHB</i> 1194                 | 95           |                                  | WP_190346674.1  |
|              | 2186        |   | <i>S. sp</i> SIO1G1            | <i>S. LEGE</i> 08334                 | 74           |                                  | WP_194056036.1  |
|              | 2224        |   | <i>S. FACHB</i> 1194           | <i>S. Kisseleviana</i> NIES 73       | 96           |                                  | WP_096571779.1  |
|              | 2212        |   | <i>S. LEGE</i> 08334           | <i>S. torques reginae</i> ITEP024    | 97           |                                  | WP_220609901.1  |

|      |      |   |                                |  |    |                                     |                |
|------|------|---|--------------------------------|--|----|-------------------------------------|----------------|
|      |      |   |                                |  |    |                                     |                |
| aptB | 1069 | Siderophore biosynthesis<br>non-ribosomal peptide                       | <i>S. torques</i> ITEP024      | <i>S. LEGE</i> 08334   | 95 | non-ribosomal peptide<br>synthetase | WP_194056038.1 |
|      | 1087 |   | <i>S. Kisseleviana</i> NIES 73 | <i>S. FACHB</i> 1194   | 95 | amino acid<br>adenylation           | WP_242052557.1 |
|      | 1074 |   | <i>S. sp</i> SIO1G1            | <i>S. torques</i> ITEP024  | 72 | non-ribosomal peptide<br>synthetase | WP_220609902.1 |
|      | 1088 |   | <i>S. FACHB</i> 1194           | <i>S. Kisseleviana</i> NIES<br>73                                    | 95 | non-ribosomal peptide<br>synthetase | WP_096571778.1 |
|      | 1073 |   | <i>S. LEGE</i> 08334           | <i>S. torques reginae</i><br>ITEP024                                 | 95 | non-ribosomal peptide<br>synthetase | WP_220609902.1 |
| aptC | 2576 | Siderophore biosynthesis<br>non-ribosomal peptide<br>synthetase modules | <i>S. torques</i> ITEP024      | <i>Nodularia spumigena</i>   | 90 | non-ribosomal peptide<br>synthetase | WP_063874470.1 |
|      | 2587 |   | <i>S. Kisseleviana</i> NIES 73 | <i>Dolichospermum sp.</i><br><i>LEGE</i> 00246                       | 97 | non-ribosomal peptide<br>synthetase | WP_193962698.1 |
|      | 2562 |   | <i>S. sp</i> SIO1G1            | <i>Aphanizomenonaceae</i><br><i>cyanobacterium</i><br><i>TIOX110</i> | 74 | amino acid adenylation              | WZB86541.1     |
|      | 2570 |   | <i>S. FACHB</i> 1194           | <i>S. Kisseleviana</i> NIES<br>73                                    | 81 | non-ribosomal peptide<br>synthetase | WP_096571777.1 |
|      | 2200 |   | <i>S. LEGE</i> 08334           | <i>S. FACHB</i> 1194   | 78 |                                     | WP_190346675.1 |
|      |      |   |                                |  |    |                                     |                |

|      |      |   |                                |  |     |                                     |                |
|------|------|---|--------------------------------|--|-----|-------------------------------------|----------------|
|      |      |   |                                |  |     |                                     |                |
| aptD | 1415 | Siderophore biosynthesis<br>non-ribosomal peptide | <i>S. torques</i> ITEP024      | <i>S. LEGE</i> 08334   | 88  |                                     | WP_194056042.1 |
|      | 1363 |   | <i>S. Kisseleviana</i> NIES 73 | <i>Dolichospermum</i>  | 93  |                                     | WP_193962697.1 |
|      | 1397 |   | <i>S. sp</i> SIO1G1            | <i>Aphanizomenonaceae</i><br><i>cyanobacterium</i><br><i>TIOX110</i> | 74  | amino acid adenylation              | WZB86542.1     |
|      | 1419 |   | <i>S. FACHB</i> 1194           | <i>S. Kisseleviana</i> NIES<br>73                                    | 84  |                                     | WP_272110476.1 |
|      | 1420 |   | <i>S. LEGE</i> 08334           | <i>S. torques reginae</i><br><i>ITEP024</i>                          | 88  | non-ribosomal peptide<br>synthetase | WP_220609905.1 |
| aptE | 392  | 2-isopropylmalate<br>synthase                     | <i>S. torques</i> ITEP024      | <i>S. FACHB</i> 1194   | 98  | 2-isopropylmalate<br>synthase       | WP_190346677.1 |
|      | 392  |   | <i>S. LEGE</i> 08334           | <i>S. Kisseleviana</i> NIES<br>73                                    | 93  |                                     | WP_272110474.1 |
|      | 253  |   | <i>S. Kisseleviana</i> NIES 73 | <i>S. FACHB</i> 1194   | 100 |                                     | MBD2144629.1   |
|      | 392  |   | <i>S. LEGE</i> 08334           | <i>S. Kisseleviana</i> NIES<br>73                                    | 93  |                                     | WP_272110474.1 |
| aptF | 813  | ATP-binding cassette                              | <i>S. FACHB</i> 1194           | <i>Sphaerospermopsis</i><br><i>reniformis</i>                        | 95  | ATP-binding cassette                | GCL35044.1     |
|      | 761  |   | <i>S. sp</i> SIO1G1            | <i>Okeanomitos</i><br><i>corallinicola</i>                           | 73  |                                     | WZB86544.1     |
|      | 803  |   | <i>S. torques</i> ITEP024      | <i>S. FACHB</i> 1194   | 86  |                                     | WP_190346678   |
|      | 761  |   | <i>S. Kisseleviana</i> NIES 73 | <i>S. LEGE</i> 08334   | 94  |                                     | MBE9057071.1   |
|      | 769  |   | <i>S. LEGE</i> 08334           | <i>S. Kisseleviana</i> NIES<br>73                                    | 94  |                                     | BAZ82708.1     |



**Supplementary Table S6. Proposed function of the proteins encoded by Sphaerociclamida biosynthetic gene cluster**

| Protein     | Amino acids | Product   | Strain                   | Organism                                  | Identity (%) | Function  | Acession number |
|-------------|-------------|---|--------------------------|---|--------------|---|-----------------|
| <b>sphG</b> | 710         | cyanobactin maturation protease                   | <i>S. sp. LEGE 00249</i> | <i>Aphanizomenon gracile</i> PMC638.10    | 95           | cyanobactin maturation protease                   | MDM3844803.1    |
| <b>sphF</b> | 290         | LynF/TruF/PatF family peptide O-prenyltransferase | <i>S. sp. LEGE 00249</i> | <i>Chrysosporum ovalisporum</i> Ak1311    | 99           | LynF/TruF/PatF family peptide O-prenyltransferase | MDH6088854.1    |
| <b>sphE</b> | 47          | anacyclamide/piricyclamide                        | <i>S. sp. LEGE 00249</i> | <i>Dolichospermum sp. JUN01</i>           | 97           | anacyclamide/piricyclamide                        | MBO1057685.1    |
| <b>sphA</b> | 667         | cyanobactin maturation protease                   | <i>S. sp. LEGE 00249</i> | <i>Sphaerospermopsis aphanizomenoides</i> | 95           | cyanobactin maturation protease                   | MBE9236410.1    |

**Supplementary Table S7. Proposed function of the proteins encoded by Nocuolin biosynthetic gene cluster**

| Protein     | Aminoacids | Product                              | Strain                   | Organism                       | Identity (%) | Function                             | Acession number |
|-------------|------------|--------------------------------------|--------------------------|--------------------------------|--------------|--------------------------------------|-----------------|
| <b>nocT</b> | 597        | AarF/ABC1/UbiB kinase family protein | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>      | 99           | AarF/ABC1/UbiB kinase family protein | MBE9236862.1    |
| <b>nocS</b> | 2101       | tectonin domain-containing protein   | <i>S. sp. LEGE 00249</i> | <i>S. FACHB 1194</i>           | 96           | tectonin domain-containing protein   | WP_242052555.1  |
| <b>nocR</b> | 426        | NocR                                 | <i>S. sp. LEGE 00249</i> | <i>Nostoc sp. CCAP 1453/38</i> | 82           | NocR                                 | AKL71651.1      |
| <b>nocQ</b> | 2325       | Oxidoreductase                       | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>      | 99           | oxidoreductase                       | MBE9235326.1    |
| <b>nocP</b> | 1286       | Polyketide synthase                  | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>      | 99           | polyketide synthase                  | MBE9235325.1    |
| <b>nocO</b> | 452        | hypothetical protein                 | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>      | 87           | Hypothetical protein                 | MBE9235324.1    |
| <b>nocM</b> | 92         | acyl carrier protein                 | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>      | 92           | acyl carrier protein                 | MBE9235322.1    |
| <b>nocN</b> | 471        | hypothetical protein                 | <i>S. sp. LEGE 00249</i> | <i>Anabaena sp. PCC 7108</i>   | 87           | Hypothetical protein                 | WP_016949101.1  |
| <b>nocK</b> | 397        | DUF3419 family protein               | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>      | 99           | DUF3419 family protein               | MBE9235319      |
| <b>nocJ</b> | 349        | NocJ                                 | <i>S. sp. LEGE 00249</i> | <i>Nostoc sp. CCAP 1453/38</i> | 89           | NocJ                                 | AKL71643.1      |

|             |     |  |                          |                                     |     |   |              |
|-------------|-----|--|--------------------------|-------------------------------------|-----|---|--------------|
| <b>nocI</b> | 372 | acyl-CoA dehydrogenase                     | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>           | 100 | acyl-CoA dehydrogenase                        | MBE9235317.1 |
| <b>nocH</b> | 698 | AMP-binding protein                        | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>           | 100 | AMP-binding protein                           | MBE9235316.1 |
| <b>nocG</b> | 347 | 3-oxoacyl-[acyl-carrier-protein] synthase  | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>           | 99  | 3-oxoacyl-[acyl-carrier-protein] synthase III | MBE9235315.1 |
| <b>nocF</b> | 872 | aminotransferase                           | <i>S. sp. LEGE 00249</i> | <i>S. aphanizomenoides BCCUSP55</i> | 89  | aminotransferase                              | MBK1987674   |
| <b>nocE</b> | 486 | NAD(P)-binding                             | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>           | 100 | NAD(P)-binding                                | MBE9235313.1 |
| <b>nocD</b> | 231 | Isoprenylcysteine carboxymethyltransferase | <i>S. sp. LEGE 00249</i> | <i>S. Aphanizomenoids</i>           | 100 | Isoprenylcysteine carboxymethyltransferase    | MBE9235312.1 |
| <b>nocB</b> | 184 | NocB                                       | <i>S. sp. LEGE 00249</i> | <i>Nostoc sp. CCAP 1453/38</i>      | 89  | NocB  | AKL71634.1   |
| <b>nocA</b> | 214 | NocA                                       | <i>S. sp. LEGE 00249</i> | <i>Nostoc sp. CCAP 1453/38</i>      | 90  | NocA  | AKL71633.1   |

**Supplementary Table S8. Proposed function of proteins encoded by hassallidin gene cluster identified in the genomes of the *C. raciborskii* strains CENA303, PAMP 2012, KL1, KLL07, CR 2010, CS-505, CS-508, GIHE-2018, GIHE-G1, DSH, N8, CR12, 1523720, CYRF and *S. torques* reginae ITP-024 como a única do gênero *Sphaerospermopsis*.**

| Protein     | Aminoacids | Proposed function      | Strain                        | Organism                     | Identity (%) | Function                            | Acession number |
|-------------|------------|------------------------|-------------------------------|------------------------------|--------------|-------------------------------------|-----------------|
| <b>HasA</b> | 555        | <i>ABC-transporter</i> | <i>C. raciborskii</i> CS-505  | <i>C. raciborskii</i> CR12   | 99           | ABC transporter ATP-binding protein | KRH96585.1      |
|             | 555        |                        | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505 | 99           | ABC transporter-like protein        | EFA70507.1      |
|             | 555        |                        | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-505 | 99           | ABC transporter-like protein        | EFA70507.1      |
|             | 555        |                        | <i>C. raciborskii</i> CENA303 | <i>C. raciborskii</i> CR12   | 88           | ABC transporter ATP-binding protein | KRH96585.1      |
|             | 555        |                        | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CR2010 | 99           | ABC transporter ATP-binding protein | UJL33519.1      |

|      |     |               |                                 |                                |     |                                     |                |
|------|-----|---------------|---------------------------------|--------------------------------|-----|-------------------------------------|----------------|
|      |     |               |                                 |                                |     |                                     |                |
|      | 555 |               | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> KLL07    | 99  | ABC transporter ATP-binding protein | UJS03222.1     |
| HasB | 670 | Alpha amylase | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12     | 97  | Alpha amylase                       | KRH96572.1     |
|      | 670 |               | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505   | 99  | Alpha-amylase                       | EFA70488.1     |
|      | 670 |               | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505   | 97  | Alpha amylase                       | EFA70488.1     |
|      | 571 |               | <i>R. brookii</i> D9            | <i>C. raciborskii</i> CR12     | 79  | Alpha amylase                       | KRH96572.1     |
|      | 670 |               | <i>C. raciborskii</i> CENA303   | <i>R. brookii</i> D9           | 99  | Alpha amylase                       | EFA73364.1     |
|      | 329 |               | <i>C. raciborskii</i> CENA302   | <i>C. raciborskii</i> CR12     | 79  | Alpha-amylase                       | KRH96572.1     |
|      | 670 |               | <i>C. raciborskii</i> CR12      | <i>C. curvispora</i> GIHE G1   | 98  | Alpha-amylase                       | WP_187706034.1 |
|      | 210 |               | <i>C. raciborskii</i> PAMP2012  | <i>C. raciborskii</i> CR12     | 73  | Alpha-amylase                       | KRH96572.1     |
|      | 571 |               | <i>R. brookii</i> D9            | <i>C. raciborskii</i> PAMP2012 | 98  | Alpha-amylase                       | MCZ2202904.1   |
|      | 670 |               | <i>C. raciborskii</i> GIHE 2018 | <i>C. curvispora</i> GIHE G1   | 99  | Alpha-amylase                       | WP_187706034.1 |
|      | 670 |               | <i>C. raciborskii</i> GIHE 2018 | <i>C. raciborskii</i> CR12     | 98  | Alpha-amylase                       | WP_057178215.1 |
|      | 670 |               | <i>C. curvispora</i> GIHEG1     | <i>C. raciborskii</i> CR12     | 99  | Alpha-amylase                       | WP_057178215.1 |
|      | 670 |               | <i>C. raciborskii</i> KLL07     | <i>C. raciborskii</i> CR12     | 98  | Alpha-amylase                       | WP_057178215.1 |
|      | 670 |               | <i>C. raciborskii</i> KLL07     | <i>C. curvispora</i> GIHEG1    | 98  | Alpha-amylase                       | WP_187706034.1 |
|      | 670 |               | <i>C. raciborskii</i> N8        | <i>C. raciborskii</i> CR12     | 100 | Alpha-amylase                       | WP_057178215.1 |
|      | 670 |               | <i>C. raciborskii</i> N8        | <i>C. curvispora</i> GIHEG1    | 99  | Alpha-amylase                       | WP_187706034.1 |
|      | 502 |               | <i>C. raciborskii</i> PAMP2012  | <i>C. curvispora</i> GIHEG1    | 81  | Alpha-amylase                       | WP_187706034.1 |
|      | 502 |               | <i>C. raciborskii</i> PAMP2012  | <i>C. raciborskii</i> CR12     | 81  | Alpha-amylase                       | WP_057178215.1 |
|      | 523 |               | <i>C. raciborskii</i> 1523720   | <i>C. raciborskii</i> CR12     | 98  | Alpha-amylase                       | WP_057178215.1 |
|      | 523 |               | <i>C. raciborskii</i> 1523720   | <i>C. curvispora</i> GIHEG1    | 98  | Alpha-amylase                       | WP_187706034.1 |

|      |     |                             |                                 |                                 |     |   |                |
|------|-----|-----------------------------|---------------------------------|---------------------------------|-----|---|----------------|
| HasC | 342 | Methyltransferase           | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12      | 95  | Methyltransferase domain-containing protein | KRH96618.1     |
|      | 342 |                             | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505    | 98  | Hypothetical protein                        | EFA70501.1     |
|      | 361 |                             | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505    | 95  | Hypothetical protein                        | EFA70501.1     |
|      | 342 |                             | <i>R. brookii</i> D9            | <i>C. raciborskii</i> CS-505    | 83  | Methyltransferase domain-containing protein | OBU78171.1     |
|      | 357 |                             | <i>C. raciborskii</i> CENA303   | <i>C. raciborskii</i> CR12      | 95  | Methyltransferase domain-containing protein | KRH96618.1     |
|      | 344 |                             | <i>C. raciborskii</i> CENA302   | <i>R. brookii</i> D9            | 93  | Hypothetical protein                        | EFA73367.1     |
|      | 342 |                             | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CHAB 3438 | 99  | Methyltransferase domain-containing protein | MCH4903615.1   |
|      | 342 |                             | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CHAB 3438 | 97  | Methyltransferase domain-containing protein | MCH4903615.2   |
|      | 71  |                             | <i>C. raciborskii</i> PAMP2012  | <i>C. raciborskii</i> CHAB 3438 | 100 | Methyltransferase domain-containing protein | MCH4903615.1   |
|      | 372 | acetylglucosamine           | <i>C. raciborskii</i> CR2010    | <i>C. curvispora</i> GIHEG1     | 97  | acetylglucosamine                           | WP_187706020.1 |
|      | 372 |                             | <i>C. raciborskii</i> GIHE 2018 | <i>C. raciborskii</i> CR12      | 100 | acetylglucosamine                           | WP_057178233.1 |
|      | 372 |                             | <i>C. raciborskii</i> GIHE 2018 | <i>C. curvispora</i> GIHE G1    | 97  | acetylglucosamine                           | WP_187706020.1 |
|      | 372 |                             | <i>C. curvispora</i> GIHEG1     | <i>C. raciborskii</i> CR12      | 97  | acetylglucosamine                           | WP_057178233.1 |
|      | 372 |                             | <i>C. raciborskii</i> N8        | <i>C. raciborskii</i> CR12      | 99  | undecaprenyldiphospho-muramoylpentapeptide  | WP_057178233   |
|      | 372 |                             | <i>C. raciborskii</i> N8        | <i>C. curvispora</i> GIHEG1     | 97  | undecaprenyldiphospho-muramoylpentapeptide  | WP_187706020.1 |
| HasD | 265 | Glycosyl transferase family | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12      | 99  | Glycosyl transferase family 2               | KRH96581.1     |
|      | 265 |                             | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505    | 99  | Glycosyl transferase family 2               | EFA70500.1     |
|      | 275 |                             | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505    | 99  | Glycosyl transferase family 2               | EFA70500.1     |
|      | 267 |                             | <i>C. raciborskii</i> ITEP-A1   | <i>Anabaena</i> sp. 90          | 83  | Glycosyl transferase family 2               | AFW95739.1     |
|      | 266 |                             | <i>C. raciborskii</i> CR12      | <i>C. curvispora</i> GIHE G1    | 100 | Glycosyl transferase family 2               | QNP29407.1     |
|      | 266 |                             | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> GIHE 2018 | 100 | Glycosyl transferase family 2               | TPX27091.1     |

|      |     |   |                                    |  |     |  |                |
|------|-----|---|------------------------------------|--|-----|--|----------------|
|      | 266 | Dolichol-phosphate<br>mannosyltransferase | <i>C. raciborskii</i> GIHE<br>2018 | <i>C. raciborskii</i> CENA303            | 76  | polyprenol<br>monophosphomannose<br>synthase | OSO97128.1     |
|      | 266 |   | <i>C. raciborskii</i> GIHE<br>2018 | <i>C. raciborskii</i> KL1                | 76  | polyprenol<br>monophosphomannose<br>synthase | MBG0742668.1   |
|      | 266 |   | <i>C. raciborskii</i> GIHE<br>2018 | <i>C. raciborskii</i> KL1                | 76  | polyprenol<br>monophosphomannose<br>synthase | MBG0742668.1   |
|      | 266 |   | <i>C. raciborskii</i> GIHE<br>2018 | <i>C. raciborskii</i> CENA303            | 76  | polyprenol<br>monophosphomannose<br>synthase | OSO97128.1     |
|      | 266 |   | <i>C. curvispora</i> GIHEG1        | <i>C. raciborskii</i> CENA303            | 76  | polyprenol<br>monophosphomannose<br>synthase | WP_141303791.1 |
|      | 266 | Dolichol-phosphate<br>mannosyltransferase | <i>C. curvispora</i> GIHEG1        | <i>C. raciborskii</i> KL1                | 76  | polyprenol<br>monophosphomannose<br>synthase | MBG0742668.1   |
|      | 265 | Dolichol-phosphate<br>mannosyltransferase | <i>C. raciborskii</i> KLL07        | <i>C. raciborskii</i> KL1                | 77  | polyprenol<br>monophosphomannose<br>synthase | MBG0742668.1   |
|      | 265 | Dolichol-phosphate<br>mannosyltransferase | <i>C. raciborskii</i> KLL07        | <i>C. raciborskii</i> CENA303            | 77  | polyprenol<br>monophosphomannose<br>synthase | OSO97128.1     |
|      | 266 | Dolichol-phosphate<br>mannosyltransferase | <i>C. raciborskii</i> N8           | <i>C. raciborskii</i> KL1                | 76  | polyprenol<br>monophosphomannose<br>synthase | MBG0742668.1   |
|      | 266 | Dolichol-phosphate<br>mannosyltransferase | <i>C. raciborskii</i> N8           | <i>C. raciborskii</i> CENA303            | 76  | polyprenol<br>monophosphomannose<br>synthase | OSO97128.1     |
| HasE | 154 | GtrA family protein                       | <i>C. raciborskii</i> CS-505       | <i>C. raciborskii</i> CR12               | 100 | Hypothetical protein                         | KRH96580.1     |
|      | 151 |   | <i>C. raciborskii</i> CS-508       | <i>C. raciborskii</i> CR12               | 99  | Hypothetical protein                         | KRH96580.1     |
|      | 151 |   | <i>C. raciborskii</i> CR12         | <i>C. raciborskii</i> CS-505             | 100 | Hypothetical protein                         | OBU77708.1     |
|      | 184 |   | <i>C. raciborskii</i> CENA303      | <i>Scytonema hofmanni</i> UTEX<br>B 1581 | 70  | Hypothetical protein                         | WP_051502819.1 |
|      | 184 |   | <i>C. raciborskii</i> ITEP-A1      | <i>Anabaena</i> sp. 90                   | 68  | GtrA-like protein                            | AFW95740.1     |

|      |     |                                 |                               |                                 |    |  |                |
|------|-----|---------------------------------|-------------------------------|---------------------------------|----|--|----------------|
|      | 197 | Polyketide synthase             | <i>C. raciborskii</i> 1523720 | <i>C. raciborskii</i> CR12      | 96 | non-ribosomal peptide synthetase                       | WP_057178252.1 |
|      | 197 | Polyketide synthase             | <i>C. raciborskii</i> 1523720 | <i>C. curvispora</i> GIHEG1     | 96 | non-ribosomal peptide synthetase                       | WP_187706030.1 |
| HasF | 547 | Putative membrane protein       | <i>C. raciborskii</i> CS-505  | <i>C. raciborskii</i> CR12      | 99 | Dolichyl-phosphate-mannose-protein mannosyltransferase | KRH96579.1     |
|      | 547 |                                 | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505    | 99 | Dolichyl-phosphate-mannose-protein mannosyltransferase | OBU77709.1     |
|      | 547 |                                 | <i>C. raciborskii</i> CR12    | <i>C. curvispora</i> GIHEG1     | 99 | Dolichyl-phosphate-mannose-protein mannosyltransferase | WP_187706028.1 |
|      | 547 |                                 | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-505    | 99 | Conserved hypothetical protein                         | EFA70498.1     |
|      | 545 |                                 | <i>C. raciborskii</i> CENA303 | <i>Anabaena</i> sp. 90          | 68 | Putative membrane protein                              | AFW95741.1     |
|      | 62  |                                 | <i>C. raciborskii</i> CENA302 | <i>Anabaena</i> sp. Syke748     | 75 | Putative membrane protein                              | AHZ20765.1     |
|      | 547 |                                 | <i>C. curvispora</i> GIHEG1   | <i>C. raciborskii</i> CR12      | 99 | Dolichyl-phosphate-mannose-protein mannosyltransferase | KRH96579.1     |
| HasG | 312 | AMP-dependent synthetase/ligase | <i>C. raciborskii</i> CENA303 | <i>Anabaena</i> sp. 90          | 87 | AMP-dependent synthetase/ligase                        | AFW95742.1     |
| HasK | 452 | Major facilitator transporter   | <i>C. raciborskii</i> CS-505  | <i>C. raciborskii</i> CR12      | 98 | MFS transporter  | KRH96587.1     |
|      | 456 |                                 | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CR2010    | 98 | MFS transporter  | UJL33521.1     |
|      | 456 |                                 | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS- 508   | 97 | MFS transporter  | OHY34101.1     |
|      | 456 |                                 | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CHAB 3438 | 97 | MFS transporter  | MCH4903619.1   |
|      | 456 |                                 | <i>C. raciborskii</i> CR12    | <i>C. curvispora</i> GIHEG1     | 96 | MFS transporter  | TPX27098.1     |
|      | 456 |                                 | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505    | 99 | Hypothetical protein                                   | EFA70509.1     |
|      | 452 |                                 | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-505    | 98 | MFS transporter  | OBU77700.1     |
|      | 449 |                                 | <i>R. brookii</i> D9          | <i>C. raciborskii</i> CS-505    | 66 | MFS transporter  | OBU77700.1     |

|      |     |  |                                 |                                   |     |   |                |
|------|-----|--|---------------------------------|-----------------------------------|-----|---|----------------|
|      | 449 |  | <i>C. raciborskii</i> CENA303   | <i>C. raciborskii</i> CS-505      | 66  | MFS transporter                               | OBU77700.1     |
|      | 449 |  | <i>C. raciborskii</i> ITEP-A1   | <i>R. brookii</i> D9              | 96  | Major facilitator superfamily MFS_1           | EFA73366.1     |
|      | 445 | Long-chain-fatty-acid--CoA ligase        | <i>C. raciborskii</i> KLL07     | <i>Nostocaceae cyanobacterium</i> | 92  | MAG TPA: fatty acid-CoA ligase family protein | HLO88498       |
|      | 444 | hypothetical protein                     | <i>C. raciborskii</i> KLL07     | <i>Nostocaceae cyanobacterium</i> | 80  | alpha/beta hydrolase-fold protein             | HLO88501.1     |
| HasL | 254 | 3-oxoacyl-acyl-carrier-protein reductase | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12        | 99  | 3-oxoacyl-ACP reductase                       | KRH96575.1     |
|      | 254 |  | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505      | 100 | 3-oxoacyl-ACP reductase                       | EFA70493.1     |
|      | 254 |  | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505      | 99  | 3-oxoacyl-ACP reductase                       | EFA70493.1     |
|      | 254 |  | <i>C. raciborskii</i> CENA303   | <i>C. raciborskii</i> CS-505      | 87  | 3-oxoacyl-ACP reductase                       | EFA70493.1     |
|      | 254 |  | <i>C. raciborskii</i> CR12      | <i>S. torques reginae</i> ITEP024 | 87  | 3-oxoacyl-ACP reductase                       | WP_220609295.1 |
|      | 254 |  | <i>C. raciborskii</i> CS-505    | <i>S. torques reginae</i> ITEP024 | 87  | 3-oxoacyl-ACP reductase                       | WP_220609295.1 |
|      | 254 |  | <i>C. raciborskii</i> CR2010    | <i>S. torques</i> ITEP024         | 87  | 3-oxoacyl-ACP reductase                       | WP_220609295.1 |
|      | 254 |  | <i>C. raciborskii</i> GIHE 2018 | <i>S. torques</i> ITEP024         | 87  | 3-oxoacyl-ACP reductase                       | WP_220609295.1 |
|      | 254 |  | <i>C. raciborskii</i> KLL07     | <i>S. torques</i> ITEP024         | 87  | 3-oxoacyl-ACP reductase                       | WP_220609295.1 |
|      | 254 |  | <i>C. raciborskii</i> N8        | <i>S. torques</i> ITEP024         | 87  | 3-oxoacyl-ACP reductase                       | WP_220609295.1 |
|      | 254 |  | <i>C. raciborskii</i> PAMP2012  | <i>S. torques</i> ITEP024         | 89  | 3-oxoacyl-ACP reductase                       | WP_220609295.1 |
|      | 254 |  | <i>C. raciborskii</i> 1523720   | <i>S. torques</i> ITEP024         | 87  | oxidoreductase                                | WP_220609295.1 |
|      | 254 |  | <i>S. torques</i> ITEP024       | <i>Cronbergia</i> sp. UHCC 0137]  | 95  | oxidoreductase                                | WP_323283647   |
|      |     |  |                                 |                                   |     |   |                |

| HasM | 192  | Aspartyl/asparaginyl beta-hydroxylase | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12        | 99 | Hydroxylase                           | KRH96574.1     |
|------|------|---------------------------------------|---------------------------------|-----------------------------------|----|---------------------------------------|----------------|
|      | 192  |                                       | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505      | 99 | Aspartyl/asparaginyl beta-hydroxylase | EFA70492.1     |
|      | 192  |                                       | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505      | 99 | Hydroxylase                           | OBU77714.1     |
|      | 204  |                                       | <i>C. raciborskii</i> CENA303   | <i>Anabaena</i> sp. 90            | 73 | Aspartyl/asparaginyl beta-hydroxylase | AFW95748.1     |
|      | 192  |                                       | <i>C. raciborskii</i> CR12      | <i>S. torques reginae</i> ITEP024 | 70 | Aspartyl/asparaginyl beta-hydroxylase | WP_220609294.1 |
|      | 197  | Polyketide synthase                   | <i>C. raciborskii</i> 1523720   | <i>C. raciborskii</i> CR12        | 96 | Non-ribosomal peptide synthetase      | WP_057178252.1 |
|      | 197  | Polyketide synthase                   | <i>C. raciborskii</i> 1523720   | <i>C. curvispora</i> GIHEG1       | 96 | Non-ribosomal peptide synthetase      | WP_187706030.1 |
|      | 3051 | NRPS                                  | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12        | 98 | Non-ribosomal peptide synthetase      | KRH96616.1     |
|      | 2824 |                                       | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CR12        | 96 | Non-ribosomal peptide synthetase      | KRH96616.1     |
|      | 3049 |                                       | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505      | 98 | Non-ribosomal peptide synthetase      | OBU77715.1     |
|      | 3050 |                                       | <i>C. raciborskii</i> CENA303   | <i>C. raciborskii</i> CS-505      | 71 | Non-ribosomal peptide synthetase      | OBU77715.1     |
|      | 3049 |                                       | <i>C. raciborskii</i> CR2010    | <i>C. raciborskii</i> CR12        | 99 | Non-ribosomal peptide synthetase      | WP_072149121.1 |
|      | 3049 |                                       | <i>C. raciborskii</i> CR2010    | <i>C. raciborskii</i> CS505       | 97 | Non-ribosomal peptide synthetase      | EFA70490.1     |
|      | 3049 |                                       | <i>C. raciborskii</i> CR2010    | <i>S. torques</i> ITEP024         | 70 | Non-ribosomal peptide synthetase      | WP_220609293.1 |
|      | 3049 |                                       | <i>C. raciborskii</i> GIHE 2018 | <i>C. curvispora</i> GIHE G1      | 99 | Non-ribosomal peptide synthetase      | WP_187706032.1 |
|      | 3049 |                                       | <i>C. raciborskii</i> GIHE 2018 | <i>C. raciborskii</i> CR12        | 99 | Non-ribosomal peptide synthetase      | WP_072149121.1 |
|      | 3049 |                                       | <i>C. curvispora</i> GIHEG1     | <i>C. raciborskii</i> CR12        | 99 | Non-ribosomal peptide synthetase      | WP_072149121.1 |
|      | 3049 |                                       | <i>C. curvispora</i> GIHEG1     | <i>C. raciborskii</i> CS505       | 97 | Non-ribosomal peptide synthetase      | EFA70490.1     |



|      |      |      |                                   |                                   |     |                                  |                |
|------|------|------|-----------------------------------|-----------------------------------|-----|----------------------------------|----------------|
| HasN | 3050 | NRPS | <i>C. raciborskii</i> KLL07       | <i>C. curvispora</i> GIHEG1       | 99  | Non-ribosomal peptide synthetase | WP_187706032.1 |
|      | 3050 |      | <i>C. raciborskii</i> KLL07       | <i>C. raciborskii</i> CR12        | 98  | Non-ribosomal peptide synthetase | WP_072149121.1 |
|      | 3050 |      | <i>C. raciborskii</i> KLL07       | <i>C. raciborskii</i> CS505       | 96  | Non-ribosomal peptide synthetase | EFA70490       |
|      | 3050 |      | <i>C. raciborskii</i> KLL07       | <i>S. torques reginae</i> ITEP024 | 71  | Non-ribosomal peptide synthetase | WP_220609293.1 |
|      | 3049 |      | <i>C. raciborskii</i> N8          | <i>C. curvispora</i> GIHEG1       | 99  | Non-ribosomal peptide synthetase | WP_187706032.1 |
|      | 3049 |      | <i>C. raciborskii</i> N8          | <i>C. raciborskii</i> CR12        | 100 | Non-ribosomal peptide synthetase | WP_072149121.1 |
|      | 3049 |      | <i>C. raciborskii</i> N8          | <i>C. raciborskii</i> CS505       | 98  | Non-ribosomal peptide synthetase | EFA70490.1     |
|      | 3050 |      | <i>C. raciborskii</i> PAMP2012    | <i>C. curvispora</i> GIHEG1       | 71  | Non-ribosomal peptide synthetase | WP_187706032.1 |
|      | 3050 |      | <i>C. raciborskii</i> PAMP2012    | <i>C. raciborskii</i> CR12        | 71  | Non-ribosomal peptide synthetase | WP_072149121.1 |
|      | 3055 |      | <i>C. raciborskii</i> 1523720     | <i>C. raciborskii</i> CR12        | 94  | Non-ribosomal peptide synthetase | WP_072149121.1 |
|      | 3055 |      | <i>C. raciborskii</i> 1523720     | <i>C. raciborskii</i> CS505       | 95  | Non-ribosomal peptide synthetase | EFA70490.1     |
|      | 3056 |      | <i>S. torques reginae</i> ITEP024 | <i>Cronbergia</i> sp. UHCC 0137]  | 85  | Non-ribosomal peptide synthetase | WP_323283649.1 |
| HasO | 3661 | NRPS | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> CS-505      | 100 | Non-ribosomal peptide synthase   | EFA70489.1     |
|      | 3670 |      | <i>C. raciborskii</i> CENA303     | <i>C. raciborskii</i> CR12        | 72  | Non-ribosomal peptide synthase   | KRH96573.1     |
|      | 3049 |      | <i>C. raciborskii</i> CR12        | <i>C. curvispora</i> GIHEG1       | 97  | Non-ribosomal peptide synthase   | WP_187706032.1 |
|      | 3049 |      | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> CS-505      | 100 | Non-ribosomal peptide synthase   | EFA70490.1     |
|      | 3049 |      | <i>C. raciborskii</i> CR12        | <i>S. torques reginae</i> ITEP024 | 71  | Non-ribosomal peptide synthase   | WP_220609293.1 |
|      | 3662 |      | <i>C. raciborskii</i> CR2010      | <i>C. raciborskii</i> CR12        | 97  | Non-ribosomal peptide synthase   | WP_057178216.1 |

|      |      |  |                                 |                              |     |                                     |                |
|------|------|--|---------------------------------|------------------------------|-----|-------------------------------------|----------------|
|      | 3659 |  | <i>C. raciborskii</i> GIHE 2018 | <i>C. raciborskii</i> CR12   | 97  | Non-ribosomal peptide synthase      | WP_057178216.1 |
|      | 3659 |  | <i>C. curvispora</i> GIHEG1     | <i>C. raciborskii</i> CR12   | 97  | Non-ribosomal peptide synthase      | WP_057178216.1 |
|      | 3662 |  | <i>C. raciborskii</i> KLL07     | <i>C. curvispora</i> GIHEG1  | 99  | Non-ribosomal peptide synthase      | WP_187706033.1 |
|      | 3662 |  | <i>C. raciborskii</i> KLL07     | <i>C. raciborskii</i> CR12   | 97  | Non-ribosomal peptide synthase      | WP_057178216.1 |
|      | 3661 |  | <i>C. raciborskii</i> N8        | <i>C. raciborskii</i> CR12   | 100 | Non-ribosomal peptide synthase      | WP_057178216.1 |
|      | 3661 |  | <i>C. raciborskii</i> N8        | <i>C. curvispora</i> GIHEG1  | 97  | Non-ribosomal peptide synthase      | WP_057178216.1 |
|      | 3670 |  | <i>C. raciborskii</i> PAMP2012  | <i>C. raciborskii</i> CR12   | 72  | Non-ribosomal peptide synthase      | WP_057178216.1 |
|      | 3670 |  | <i>C. raciborskii</i> PAMP2012  | <i>C. curvispora</i> GIHEG1  | 72  | Non-ribosomal peptide synthase      | WP_187706033.1 |
|      | 3660 |  | <i>C. raciborskii</i> 1523720   | <i>C. raciborskii</i> CR12   | 95  | non-ribosomal peptide synthetase    | WP_057178216.1 |
|      | 3660 |  | <i>C. raciborskii</i> 1523720   | <i>C. curvispora</i> GIHEG1  | 94  | non-ribosomal peptide synthetase    | WP_187706033   |
| HasP | 354  | NAD-dependent epimerase dTDP-glucose 4,6-dehydratase | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12   | 99  | NAD-dependent epimerase             | KRH96583.1     |
|      | 354  |  | <i>C. raciborskii</i> CR12      | <i>C. curvispora</i> GIHEG1  | 100 | NAD-dependent epimerase/dehydratase | WP_187706027.1 |
|      | 354  |  | <i>C. raciborskii</i> CS-505    | <i>C. curvispora</i> GIHEG1  | 99  | NAD-dependent epimerase/dehydratase | WP_187706027.1 |
|      | 354  |  | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505 | 99  | NAD-dependent epimerase/dehydratase | EFA70505.1     |
|      | 352  |  | <i>C. raciborskii</i> CENA303   | <i>Anabaena</i> sp. 90       | 84  | NAD dependent epimerase/dehydratase | AFW95751.1     |
|      | 370  |  | <i>C. raciborskii</i> CR2010    | <i>C. raciborskii</i> CR12   | 99  | NAD-dependent epimerase/dehydratase | WP_057178224.1 |
|      | 370  |  | <i>C. raciborskii</i> GIHE 2018 | <i>C. raciborskii</i> CR12   | 99  | NAD-dependent epimerase/dehydratase | WP_057178224.1 |
|      | 370  |  | <i>C. raciborskii</i> GIHE 2018 | <i>C. curvispora</i> GIHE G1 | 99  | NAD-dependent epimerase/dehydratase | WP_187706027.1 |
|      | 370  |  | <i>C. curvispora</i> GIHEG1     | <i>C. raciborskii</i> CR12   | 94  | NAD-dependent epimerase/dehydratase | WP_057178224.1 |

|      |     |                          |                                   |                                     |     |                                     |                |
|------|-----|--------------------------|-----------------------------------|-------------------------------------|-----|-------------------------------------|----------------|
|      | 370 |                          | <i>C. raciborskii</i> KLL07       | <i>C. curvispora</i> GIHEG1         | 100 | NAD-dependent epimerase/dehydratase | WP_187706027   |
|      | 370 |                          | <i>C. raciborskii</i> KLL07       | <i>C. raciborskii</i> CR12          | 100 | NAD-dependent epimerase/dehydratase | WP_057178224.1 |
|      | 354 |                          | <i>C. raciborskii</i> N8          | <i>C. raciborskii</i> CR12          | 100 | NAD dependent epimerase/dehydratase | WP_057178224.1 |
|      | 354 |                          | <i>C. raciborskii</i> N8          | <i>C. curvispora</i> GIHEG1         | 99  | NAD dependent epimerase/dehydratase | WP_187706027.1 |
|      | 352 |                          | <i>C. raciborskii</i> PAMP2012    | <i>Anabaena cylindrica</i>          | 82  | NAD-dependent epimerase/dehydratase | WP_323309738.1 |
| HasQ | 402 | Glycosyltransferase      | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i> CR12          | 99  | MGT family glycosyltransferase      | KRH96584.1     |
|      | 402 |                          | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> CS-505        | 99  | Hypothetical protein                | EFA70506.1     |
|      | 73  |                          | <i>C. raciborskii</i> CENA303     | <i>C. raciborskii</i> CS-505        | 90  | MGT family glycosyltransferase      | OBU77703.1     |
|      | 402 |                          | <i>C. raciborskii</i> CS-505      | <i>S. torques reginae</i> ITEP024   | 83  | MGT family glycosyltransferase      | WP_220609292.1 |
|      | 402 |                          | <i>C. raciborskii</i> CR2010      | <i>S. torques reginae</i> ITEP024   | 82  | Glycosyl transferase                | WP_220609292.1 |
|      | 402 |                          | <i>C. raciborskii</i> GIHE 2018   | <i>S. torques reginae</i> ITEP024   | 72  | Glycosyl transferase                | WP_220609292.1 |
|      | 402 |                          | <i>C. curvispora</i> GIHEG1       | <i>S. torques reginae</i> ITEP024   | 83  | Glycosyl transferase                | WP_220609292.1 |
|      | 402 |                          | <i>C. raciborskii</i> KLL07       | <i>S. torques reginae</i> ITEP024   | 83  | Glycosyl transferase                | WP_220609292.1 |
|      | 402 |                          | <i>C. raciborskii</i> CR12        | <i>S. torques reginae</i> ITEP024   | 83  | Glycosyl transferase                | WP_220609292.1 |
|      | 402 |                          | <i>C. raciborskii</i> N8          | <i>S. torques reginae</i> ITEP024   | 82  | Glycosyl transferase                | WP_220609292.1 |
|      | 408 |                          | <i>S. torques reginae</i> ITEP024 | <i>Dolichospermum</i> sp. UHCC 0352 | 91  | Glycosyl transferase                | WP_168652019.1 |
| HasR | 206 | Putative acyltransferase | <i>C. raciborskii</i> CENA303     | <i>Anabaena</i> sp. 90              | 68  | Hypothetical protein                | WP_015080900.1 |

|      |     |                            |                                 |                                   |     |                                    |                |
|------|-----|----------------------------|---------------------------------|-----------------------------------|-----|------------------------------------|----------------|
|      |     |                            |                                 |                                   |     |                                    |                |
| HasT | 245 | Glycosyl transferase       | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12        | 99  | Glycosyl transferase               | KRH96578.1     |
|      | 245 |                            | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505      | 100 | Glycosyl transferase               | EFA70497.1     |
|      | 245 |                            | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505      | 99  | Putative glycosyltransferase       | EFA70497.1     |
|      | 245 |                            | <i>C. raciborskii</i> CR12      | <i>C. curvispora</i> GIHEG1       | 99  | Glycosyl transferase               | WP_187706029.1 |
|      | 258 |                            | <i>C. raciborskii</i> CENA303   | <i>C. raciborskii</i> CR12        | 79  | Glycosyl transferase               | KRH96578.1     |
|      | 258 |                            | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> PAMP2012    | 79  | Glycosyl transferase               | MCZ2202044.1   |
|      | 315 |                            | <i>S. torques</i> ITEP024       | <i>C. raciborskii</i>             | 88  | Glycosyl transferase               | WP_324140027.1 |
|      |     |                            |                                 |                                   |     |                                    |                |
| HasU | 72  | Putative MbtH-like protein | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12        | 100 | Antibiotic synthesis protein MbtH  | KRH96576.1     |
|      | 72  |                            | <i>Cylindrospermopsis</i> CR12  | <i>S. torques reginae</i> ITEP024 | 84  | Antibiotic synthesis protein MbtH  | WP_220609290.1 |
|      | 72  |                            | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505      | 98  | Putative MbtH-like protein         | EFA70494.1     |
|      | 72  |                            | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> DSH         | 100 | Putative MbtH-like protein         | MEE6162471     |
|      | 72  |                            | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12        | 100 | Putative MbtH-like protein         | KRH96576.1     |
|      | 72  |                            | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CS-508      | 100 | Antibiotic synthesis protein MbtH  | OHY35853.1     |
|      | 72  |                            | <i>C. raciborskii</i> CS-505    | <i>C. curvispora</i> GIHEG1       | 100 | Putative MbtH-like protein         | TPX27085.1     |
|      | 72  |                            | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR2010      | 100 | MbtH family NRPS accessory protein | UJL33506.1     |
|      | 72  |                            | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> KLL07       | 100 | MbtH family NRPS accessory protein | UJS03236.1     |
|      | 72  |                            | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505      | 100 | Putative MbtH-like protein         | EFA70494.1     |
|      | 73  |                            | <i>C. raciborskii</i> CENA303   | <i>Anabaena</i> sp. 90            | 87  | MbtH-like protein                  | AFW95756.1     |
|      | 72  |                            | <i>C. raciborskii</i> CR2010    | <i>S. torques reginae</i> ITEP024 | 84  | MbtH family protein                | WP_220609290.1 |
|      | 72  |                            | <i>C. raciborskii</i> GIHE 2018 | <i>S. torques reginae</i> ITEP024 | 72  | MbtH family protein                | WP_220609290.1 |
|      | 72  |                            | <i>C. curvispora</i> GIHEG1     | <i>S. torques reginae</i> ITEP024 | 84  | MbtH family protein                | WP_220609290.1 |
|      | 72  |                            | <i>C. raciborskii</i> KLL07     | <i>S. torques reginae</i> ITEP024 | 84  | MbtH family protein                | WP_220609290.1 |
|      | 72  |                            | <i>C. raciborskii</i> N8        | <i>S. torques reginae</i> ITEP024 | 84  | MbtH family protein                | WP_220609290.1 |
|      |     |                            |                                 |                                   |     |                                    |                |

|      |      |      |                                    |                                   |     |   |                |
|------|------|------|------------------------------------|-----------------------------------|-----|---|----------------|
|      | 73   |      | <i>C. raciborskii</i><br>PAMP2012  | <i>S. torques reginae</i> ITEP024 | 94  | MbtH family protein                                 | WP_220609290   |
|      | 72   |      | <i>C. raciborskii</i> 1523720      | <i>S. torques reginae</i> ITEP024 | 84  | MbtH family protein                                 | WP_220609290.1 |
|      | 71   |      | <i>S. torques</i> ITEP024          | <i>Cronbergia sp.</i> UHCC 0137]  | 96  | MbtH-like NRPS<br>chaperone                         | WP_323283663.1 |
| HasV | 2153 | NRPS | <i>C. raciborskii</i> CS-505       | <i>C. raciborskii</i> CR12        | 97  | Non-ribosomal peptide<br>synthase                   | KRH96577.1     |
|      | 2127 |      | <i>C. raciborskii</i> CS-508       | <i>C. raciborskii</i> CR12        | 99  | Non-ribosomal peptide<br>synthase                   | KRH96577.1     |
|      | 2153 |      | <i>C. raciborskii</i> CR12         | <i>C. curvispora</i> GIHEG1       | 99  | Non-ribosomal peptide<br>synthase                   | WP_187706031.1 |
|      | 2153 |      | <i>C. raciborskii</i> CR12         | <i>C. raciborskii</i> CR2010      | 99  | amino acid adenylation<br>domain-containing protein | UJL33507.1     |
|      | 2153 |      | <i>C. raciborskii</i> CR12         | <i>C. raciborskii</i> KLL07       | 99  | amino acid adenylation<br>domain-containing protein | UJS06337.1     |
|      | 2153 |      | <i>C. raciborskii</i> CR12         | <i>S. torques reginae</i> ITEP024 | 74  | Non-ribosomal peptide<br>synthase                   | WP_220609291.1 |
|      | 2153 |      | <i>C. raciborskii</i> CR12         | <i>C. raciborskii</i> KL1         | 98  | amino acid adenylation<br>domain-containing protein | MBG0742686.1   |
|      | 2153 |      | <i>C. raciborskii</i> CR12         | <i>C. raciborskii</i> CS-505      | 97  | Non-ribosomal peptide<br>synthase                   | EFA70495.1     |
|      | 2147 |      | <i>C. raciborskii</i> CENA303      | <i>Anabaena sp.</i> 90            | 73  | Non-ribosomal peptide<br>synthase                   | AFW95757.1     |
|      | 2153 |      | <i>C. raciborskii</i> GIHE<br>2018 | <i>C. raciborskii</i> CR2010      | 100 | amino acid adenylation                              | UJL33507.1     |
|      | 2153 |      | <i>C. raciborskii</i> GIHE<br>2018 | <i>C. raciborskii</i> CR12        | 99  | Non-ribosomal peptide<br>synthase                   | WP_057178219   |
|      | 2153 |      | <i>C. raciborskii</i> GIHE<br>2018 | <i>C. curvispora</i> GIHE G1      | 99  | Non-ribosomal peptide<br>synthase                   | WP_187706031.1 |
|      | 2153 |      | <i>C. raciborskii</i> GIHE<br>2018 | <i>C. raciborskii</i> KLL07       | 100 | Non-ribosomal peptide<br>synthase                   | UJS06337.1     |
|      | 2153 |      | <i>C. raciborskii</i> GIHE<br>2018 | <i>S. torques</i> ITEP024         | 74  | Non-ribosomal peptide<br>synthase                   | WP_220609291.1 |
|      | 2153 |      | <i>C. curvispora</i> GIHEG1        | <i>C. raciborskii</i> CR12        | 100 | Non-ribosomal peptide<br>synthase                   | WP_057178219.1 |
|      | 2153 |      | <i>C. curvispora</i> GIHEG1        | <i>C. raciborskii</i> CR2010      | 100 | Non-ribosomal peptide<br>synthase                   | UJL33507.1     |
|      | 2153 |      | <i>C. curvispora</i> GIHEG1        | <i>C. raciborskii</i> KLL07       | 99  | Non-ribosomal peptide                               | UJS06337.1     |

|  |      |  |                                   |                                   |     |  |                |
|--|------|--|-----------------------------------|-----------------------------------|-----|--|----------------|
|  |      |  |                                   |                                   |     | synthase   |                |
|  | 2153 |  | <i>C. curvispora</i> GIHEG1       | <i>S. torques</i> ITEP024         | 74  | Non-ribosomal peptide synthase                   | WP_220609291.1 |
|  | 2153 |  | <i>C. raciborskii</i> KLL07       | <i>C. raciborskii</i> CR2010      | 100 | Non-ribosomal peptide synthase                   | UJL33507.1     |
|  | 2153 |  | <i>C. raciborskii</i> KLL07       | <i>C. raciborskii</i> CR12        | 99  | Non-ribosomal peptide synthase                   | WP_057178219.1 |
|  | 2153 |  | <i>C. raciborskii</i> KLL07       | <i>C. curvispora</i> GIHEG1       | 99  | Non-ribosomal peptide synthase                   | WP_187706031.1 |
|  | 2153 |  | <i>C. raciborskii</i> KLL07       | <i>S. torques reginae</i> ITEP024 | 74  | Non-ribosomal peptide synthase                   | WP_220609291.1 |
|  | 2153 |  | <i>C. raciborskii</i> N8          | <i>C. raciborskii</i> CR12        | 100 | Non-ribosomal peptide synthase                   | WP_057178219.1 |
|  | 2153 |  | <i>C. raciborskii</i> N8          | <i>C. curvispora</i> GIHEG1       | 99  | Non-ribosomal peptide synthase                   | WP_187706031.1 |
|  | 2153 |  | <i>C. raciborskii</i> N8          | <i>C. raciborskii</i> CR2010      | 99  | amino acid adenylation domain-containing protein | UJL33507.1     |
|  | 2153 |  | <i>C. raciborskii</i> N8          | <i>C. raciborskii</i> KLL07       | 99  | amino acid adenylation domain-containing protein | UJS06337.1     |
|  | 2153 |  | <i>C. raciborskii</i> N8          | <i>S. torques</i> ITEP024         | 74  | Non-ribosomal peptide synthase                   | WP_220609291.1 |
|  | 2147 |  | <i>C. raciborskii</i> PAMP2012    | <i>C. raciborskii</i> KL1         | 98  | amino acid adenylation                           | MBG0742686.1   |
|  | 2147 |  | <i>C. raciborskii</i> PAMP2012    | <i>S. torques reginae</i> ITEP024 | 77  | Non-ribosomal peptide synthase                   | WP_220609291.1 |
|  | 2153 |  | <i>C. raciborskii</i> 1523720     | <i>C. curvispora</i> GIHEG1       | 96  | Non-ribosomal peptide synthase                   | WP_187706031.1 |
|  | 2153 |  | <i>C. raciborskii</i> 1523720     | <i>C. raciborskii</i> CR12        | 95  | Non-ribosomal peptide synthase                   | WP_057178219.1 |
|  | 2153 |  | <i>C. raciborskii</i> 1523720     | <i>C. raciborskii</i> CR2010      | 95  | amino acid adenylation                           | UJL33507.1     |
|  | 2153 |  | <i>C. raciborskii</i> 1523720     | <i>C. raciborskii</i> KLL07       | 95  | amino acid adenylation                           | UJS06337.1     |
|  | 2142 |  | <i>S. torques reginae</i> ITEP024 | <i>C. raciborskii</i> KL1         | 76  | Non-ribosomal peptide synthase                   | WP_220609288.1 |
|  | 1961 |  | <i>S. torques reginae</i> ITEP024 | <i>C. curvispora</i> GIHEG1       | 74  | Non-ribosomal peptide synthase                   | WP_187706033.1 |
|  | 1961 |  | <i>S. torques reginae</i>         | <i>C. raciborskii</i> CR12        | 74  | Non-ribosomal peptide                            | WP_057178216.1 |

|      |      |                      |                                 |                                       |     |                                |                |
|------|------|----------------------|---------------------------------|---------------------------------------|-----|--------------------------------|----------------|
|      |      |                      | ITEP024                         |                                       |     | synthase                       |                |
| HasX | 310  | Glycosyltransferase  | <i>C. raciborskii</i> CR12      | <i>Anabaena</i> sp. 90                | 73  | Glycosyl transferase family 2  | AFW95759.1     |
|      | 310  |                      | <i>C. raciborskii</i> N8        | <i>S. torques</i> ITEP024             | 88  | Glycosyltransferase            | WP_220609289.1 |
|      | 267  |                      | <i>C. raciborskii</i> CENA303   | <i>Anabaena</i> sp. 90                | 83  | Family 2 glycosyl transferase  | AFW95739.1     |
|      | 82   | Hypothetical protein | <i>C. raciborskii</i> CENA302   | <i>Scytonema hofmanni</i> UTEX B 1581 | 74  | Hypothetical protein           | WP_051502821.1 |
|      | 82   | Glycosyltransferase  | <i>C. raciborskii</i> PAMP2012  | <i>S. torques reginae</i> ITEP024     | 83  | Glycosyltransferase            | WP_220609289.1 |
| HasY | 2874 | NRPS                 | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12            | 90  | Non-ribosomal peptide synthase | KRH96617.1     |
|      | 2870 |                      | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CR12            | 96  | Non-ribosomal peptide synthase | KRH96617.1     |
|      | 2864 |                      | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505          | 90  | Non-ribosomal peptide synthase | EFA70496.1     |
|      | 2864 |                      | <i>C. raciborskii</i> CR12      | <i>S. torques reginae</i> ITEP024     | 76  | Non-ribosomal peptide synthase | WP_220609288.1 |
|      | 2864 |                      | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> KL1             | 72  | Non-ribosomal peptide synthase | MBG0742682.1   |
|      | 2864 |                      | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CENA303         | 72  | Non-ribosomal peptide synthase | OSO88008.1     |
|      | 2863 |                      | <i>C. raciborskii</i> CENA303   | <i>C. curvispora</i> GIHEG1           | 72  | Non-ribosomal peptide synthase | WP_187706030.1 |
|      | 2864 |                      | <i>C. raciborskii</i> GIHE 2018 | <i>C. curvispora</i> GIHEG1           | 100 | Non-ribosomal peptide synthase | WP_187706030.1 |
|      | 2864 |                      | <i>C. raciborskii</i> GIHE 2018 | <i>C. raciborskii</i> CR12            | 97  | Non-ribosomal peptide synthase | WP_057178252.1 |
|      | 2864 |                      | <i>C. raciborskii</i> GIHE 2018 | <i>C. raciborskii</i> KL1             | 72  | Non-ribosomal peptide synthase | MBG0742682.1   |
|      | 2864 |                      | <i>C. raciborskii</i> GIHE 2018 | <i>S. torques</i> ITEP024             | 76  | Non-ribosomal peptide synthase | WP_220609288.1 |
|      | 2864 |                      | <i>C. raciborskii</i> GIHE 2018 | <i>C. raciborskii</i> CENA303         | 72  | Non-ribosomal peptide synthase | OSO88008.1     |
|      | 2864 |                      | <i>C. curvispora</i> GIHEG1     | <i>C. raciborskii</i> CR12            | 97  | Non-ribosomal peptide synthase | WP_057178252.1 |
|      | 2864 |                      | <i>C. curvispora</i> GIHEG1     | <i>S. torques reginae</i> ITEP024     | 76  | Non-ribosomal peptide synthase | WP_220609288.1 |

|      |      |             |                                   |                                   |     |  |                |
|------|------|-------------|-----------------------------------|-----------------------------------|-----|--|----------------|
|      | 2864 |             | <i>C. raciborskii</i> KLL07       | <i>C. curvispora</i> GIHEG1       | 100 | Non-ribosomal peptide synthase                   | WP_220609288.1 |
|      | 2864 |             | <i>C. raciborskii</i> KLL07       | <i>C. raciborskii</i> CR12        | 98  | Non-ribosomal peptide synthase                   | WP_220609288.1 |
|      | 2864 |             | <i>C. raciborskii</i> KLL07       | <i>S. torques reginae</i> ITEP024 | 76  | Non-ribosomal peptide synthase                   | WP_220609288.1 |
|      | 2864 |             | <i>C. raciborskii</i> N8          | <i>C. curvispora</i> GIHEG1       | 99  | Non-ribosomal peptide synthase                   | WP_220609288.1 |
|      | 2864 |             | <i>C. raciborskii</i> N8          | <i>C. raciborskii</i> CR12        | 97  | Non-ribosomal peptide synthase                   | WP_220609288.1 |
|      | 2864 |             | <i>C. raciborskii</i> N8          | <i>S. torques reginae</i> ITEP024 | 76  | Non-ribosomal peptide synthase                   | WP_220609288.1 |
|      | 2864 |             | <i>C. raciborskii</i> N8          | <i>C. raciborskii</i> KL1         | 72  | amino acid adenylation domain-containing protein | MBG0742682.1   |
|      | 2864 |             | <i>C. raciborskii</i> N8          | <i>C. raciborskii</i> CENA303     | 72  | Non-ribosomal peptide synthase                   | OSO88008.1     |
|      | 2862 |             | <i>C. raciborskii</i> PAMP2012    | <i>S. torques reginae</i> ITEP024 | 84  | Non-ribosomal peptide synthase                   | WP_220609288.1 |
|      | 2862 |             | <i>C. raciborskii</i> PAMP2012    | <i>C. raciborskii</i> KL1         | 98  | Non-ribosomal peptide synthase                   | MBG0742682.1   |
|      | 2862 |             | <i>C. raciborskii</i> PAMP2012    | <i>C. raciborskii</i> CENA303     | 99  | Non-ribosomal peptide synthase                   | OSO88008.1     |
|      | 2855 |             | <i>S. torques reginae</i> ITEP024 | <i>C. raciborskii</i> KL1         | 84  | Non-ribosomal peptide synthase                   | MBG0742686.1   |
|      | 2855 |             | <i>S. torques reginae</i> ITEP024 | <i>C. raciborskii</i> CENA303     | 84  | Non-ribosomal peptide synthase                   | OSO88008.1     |
|      | 2855 |             | <i>S. torques reginae</i> ITEP024 | <i>C. curvispora</i> GIHEG1       | 76  | Non-ribosomal peptide synthase                   | WP_187706030.1 |
|      | 2855 |             | <i>S. torques reginae</i> ITEP024 | <i>C. raciborskii</i> CR12        | 76  | Non-ribosomal peptide synthase                   | WP_057178252.1 |
| HasZ | 431  | Transporter | <i>C. raciborskii</i> CS-505      | <i>Cylindrospermopsis</i> CR12    | 97  | MFS transporter                                  | KRH96586.1     |
|      | 431  |             | <i>C. raciborskii</i> CS-505      | <i>C. curvispora</i> GIHE G1      | 97  | MFS transporter                                  | QNP29400.1     |
|      | 431  |             | <i>C. raciborskii</i> CS-505      | <i>C. curvispora</i> GIHE G1      | 97  | MFS transporter                                  | QNP29400.1     |
|      | 431  |             | <i>C. raciborskii</i> CS-505      | <i>C. curvispora</i> GIHE G1      | 97  | Major facilitator superfamily MFS_0              | QNP29400.1     |
|      | 427  |             | <i>C. raciborskii</i> CS-508      | <i>C. raciborskii</i> CS-505      | 96  | Major facilitator superfamily MFS_1              | EFA70508.1     |



|  |     |  |                               |                              |    |                                     |                |
|--|-----|--|-------------------------------|------------------------------|----|-------------------------------------|----------------|
|  | 431 |  | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-505 | 97 | Major facilitator superfamily MFS_1 | EFA70508.1     |
|  | 431 |  | <i>C. raciborskii</i> CS-505  | <i>C. raciborskii</i> CR12   | 97 | MFS transporter                     | WP_057178227.1 |
|  | 429 |  | <i>C. raciborskii</i> CENA303 | <i>C. raciborskii</i> CS-505 | 68 | MFS transporter                     | OBU77701.1     |

**Supplementary Table S9. Proposed function of the proteins encoded by heterocyte glycolipid biosynthetic gene cluster and two proteins involved in the regulation of heterocyte differentiation in the genomes of the *C. raciborskii* .**

| Protein | Aminoacids | Product                    | Strain                        | Organism                           | Identity (%) | Function                        | Acession number |
|---------|------------|----------------------------|-------------------------------|------------------------------------|--------------|---------------------------------|-----------------|
| HglA    | 577        | 2-nitropropane dioxygenase | <i>C. raciborskii</i> CS-505  | <i>C. raciborskii</i> CS-505       | 99           | 2-nitropropane dioxygenase, NPD | EFA69444.1      |
|         | 577        |                            | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505       | 99           |                                 | EFA69444.1      |
|         | 577        |                            | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-505       | 99           |                                 | EFA69444.1      |
|         | 577        |                            | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> MVCC14       | 99           |                                 | OHY31851.1      |
|         | 577        |                            | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> MVCC14       | 99           |                                 | OHY31851.1      |
|         | 577        |                            | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> CR12         | 97           |                                 | EFA69444.1      |
|         | 577        |                            | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CHAB3438     | 100          |                                 | MCH4905695.1    |
|         | 577        |                            | <i>C. raciborskii</i> CR12    | <i>C. curvispora</i> GIHE-G8       | 98           |                                 | QNP28543.1      |
|         | 577        |                            | <i>C. raciborskii</i> CR12    | <i>S. sp.</i> SIO1G1               | 80           |                                 | NET02811.1      |
|         | 577        |                            | <i>C. raciborskii</i> CR12    | <i>S. kisseleviana</i> NIES-73     | 84           |                                 | BAZ83237.1      |
|         | 577        |                            | <i>C. raciborskii</i> CR12    | <i>S. torques reginae</i> ITEP-024 | 84           |                                 | QYX32837.1      |
|         | 577        |                            | <i>C. raciborskii</i> CR12    | <i>S. sp.</i> LEGE 08334           | 84           |                                 | MBE9056527.1    |
|         | 577        |                            | <i>C. raciborskii</i> CR12    | <i>S. sp.</i> LEGE 00249           | 84           |                                 | MBC5793826.1    |
| HglB    | 507        | Polyketide synthase        | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505       | 100          | Polyketide synthase HetM        | EFA69445.1      |
|         | 507        |                            | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505       | 99           | Polyketide synthase HetM        | EFA69445.1      |
|         | 507        |                            | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-505       | 100          | Polyketide synthase HetM        | EFA69445.1      |
|         | 507        |                            | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> MVCC14       | 100          | Polyketide synthase             | OHY31850.1      |
|         | 507        |                            | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> MVCC14       | 9            | Polyketide synthase             | OHY31850.1      |
|         | 507        |                            |                               |                                    | 9            |                                 |                 |

|             |      |                     |                               |                              |     |                                     |            |
|-------------|------|---------------------|-------------------------------|------------------------------|-----|-------------------------------------|------------|
|             | 507  |                     | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> CS-505 | 96  | Polyketide synthase HetM            | EFA69445.1 |
| <b>HgIC</b> | 1104 | Polyketide synthase | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505 | 100 | HgIC (Beta-ketoacyl synthase)       | EFA69443.1 |
|             | 1104 |                     | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CR12   | 99  | Polyketide synthase                 | KRH97007.1 |
|             | 1104 |                     | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-508 | 99  |                                     | OHY37860.1 |
|             | 1100 |                     | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> MVCC14 | 99  |                                     | OHY31852.1 |
|             | 1100 |                     | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> MVCC14 | 99  |                                     | OHY31852.1 |
|             | 1100 |                     | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> CS-508 | 97  |                                     | OHY37860.1 |
|             | 1100 |                     |                               |                              |     |                                     |            |
| <b>HgID</b> | 453  | Polyketide synthase | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505 | 100 | Heterocyst glycolipid synthase HgID | EFA69442.1 |
|             | 453  |                     | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505 | 99  | Heterocyst glycolipid synthase HgID | EFA69442.1 |
|             | 453  |                     | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-505 | 100 | Heterocyst glycolipid synthase HgID | EFA69442.1 |
|             | 453  |                     | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> CS-505 | 98  | Heterocyst glycolipid synthase HgID | EFA69442.1 |
|             | 453  |                     | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> MVCC14 | 99  | Polyketide synthase                 | OHY31853.1 |
|             | 453  |                     |                               |                              |     |                                     |            |
|             | 453  |                     | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> CS-505 | 98  | Heterocyst glycolipid synthase HgID | EFA69442.1 |
| <b>HgIE</b> | 1771 | Polyketide synthase | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505 | 100 | Beta-ketoacyl synthase              | EFA69440.1 |
|             | 1771 |                     | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505 | 99  | Beta-ketoacyl synthase              | EFA69440.1 |
|             | 1771 |                     | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-505 | 98  | Beta-ketoacyl synthase              | EFA68870.1 |
|             | 1751 |                     | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> MVCC14 | 99  | Beta-ketoacyl synthase              | OHY34297.1 |
|             | 1751 |                     |                               |                              |     |                                     |            |

|      |      |  |                               |                               |     |   |            |
|------|------|--|-------------------------------|-------------------------------|-----|---|------------|
|      | 1751 |  | <i>C. raciborskii</i> ITEP-A1 | <i>R. brookii</i> D9          | 95  | Beta-ketoacyl synthase                      | EFA74388.1 |
|      | 1751 |  | <i>C. raciborskii</i> MVCC14  | <i>R. brookii</i> D9          | 96  | Beta-ketoacyl synthase                      | EFA74388.1 |
| HglG | 585  | Polyketide synthase                        | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505  | 100 | Putative ketoreductase and dehydrase (HglG) | EFA69441.1 |
|      | 585  |  | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505  | 99  | Putative ketoreductase and dehydrase (HglG) | EFA69441.1 |
|      | 585  |  | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-508  | 99  | Polyketide synthase                         | OHY37862.1 |
|      | 585  |  | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> MVCC14  | 100 | Polyketide synthase                         | OHY31854.1 |
|      | 585  |  | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> MVCC14  | 100 | Polyketide synthase                         | OHY31854.1 |
|      | 585  |  | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> CS-508  | 97  | Polyketide synthase                         | OHY37862.1 |
|      |      |  |                               |                               |     |   |            |
| HetR | 299  | Heterocyte differentiation control protein | <i>C. raciborskii</i> CS-505  | <i>C. raciborskii</i> CS-505  | 100 | Heterocyst differentiation control protein  | EFA68362.1 |
|      | 299  |  | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505  | 100 | Heterocyst differentiation control protein  | EFA68362.1 |
|      | 299  |  | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-505  | 100 | Heterocyst differentiation control protein  | EFA68362.1 |
|      | 299  |  | <i>R. brookii</i> D9          | <i>C. raciborskii</i> CENA303 | 99  | Heterocyst differentiation control protein  | OSO91885.1 |
|      | 299  |  | <i>C. raciborskii</i> CENA303 | <i>R. brookii</i> D9          | 99  | Heterocyst differentiation control protein  | EFA73115.1 |
|      | 299  |  | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> CS-505  | 100 | Heterocyst differentiation control protein  | EFA68362.1 |
|      | 299  |  | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> CS-505  | 100 | Heterocyst differentiation control protein  | EFA68362.1 |
|      | 299  |  | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> CS-505  | 100 | Heterocyst differentiation                  | EFA68362.1 |

|      |     |  |                               |                                    |     |  |              |
|------|-----|--|-------------------------------|------------------------------------|-----|--|--------------|
|      |     |  |                               |                                    |     | control protein                            |              |
|      | 299 |  | <i>C. raciborskii</i> CS-505  | <i>C. curvispora</i> GIHE-G8       | 100 | Heterocyst differentiation control protein | QNP30904.1   |
|      | 299 |  | <i>C. raciborskii</i> CS-505  | <i>S. LEGE</i> 08334               | 90  | Heterocyst differentiation control protein | MBE9057199.1 |
|      | 299 |  | <i>C. raciborskii</i> CS-505  | <i>S. torques reginae</i> ITEP-024 | 90  | Heterocyst differentiation control protein | QYX31586.1   |
|      | 299 |  | <i>C. raciborskii</i> CS-505  | <i>S. sp. LEGE</i> 00249           | 89  | Heterocyst differentiation control protein | MBC5794768.1 |
|      | 299 |  | <i>C. raciborskii</i> CS-505  | <i>S.sp. SIO1G1</i>                | 88  | Heterocyst differentiation control protein | NET01886.1   |
| HetN | 241 | SDR family NAD(P)-dependent oxidoreductase | <i>C. raciborskii</i> CS-505  | <i>C. raciborskii</i> CS-505       | 100 | SDR family NAD(P)-dependent oxidoreductase | EFA70572.1   |
|      | 241 |  | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CS-505       | 100 | SDR family NAD(P)-dependent oxidoreductase | EFA70572.1   |
|      | 241 |  | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CS-505       | 99  | SDR family NAD(P)-dependent oxidoreductase | EFA70572.1   |
|      | 241 |  | <i>R. brookii</i> D9          | <i>C. raciborskii</i> CENA303      | 99  | Short-chain dehydrogenase                  | OSO94107.1   |
|      | 241 |  | <i>C. raciborskii</i> CENA303 | <i>R. brookii</i> D9               | 99  | Short-chain dehydrogenase/reductase SDR    | EFA73007.1   |
|      | 241 |  | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> MVCC14       | 100 | Short-chain dehydrogenase                  | OHY31780.1   |
|      | 241 |  | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> MVCC14       | 99  | Short-chain dehydrogenase                  | OHY31780.1   |
|      | 241 |  | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> CENA302      | 100 | Short-chain dehydrogenase                  | OPH09154.1   |

**Supplementary Table S10. Proposed function of the proteins encoded by the nitrogen fixation biosynthetic gene cluster in the genomes of the *Cylindrospermopsis* strains.**

| Protein     | Aminoacids | Proposed function   | Strain                          | Organism                     | Identity (%) | Function              | Acession number |
|-------------|------------|---------------------|---------------------------------|------------------------------|--------------|-----------------------|-----------------|
| <b>FeoA</b> | 86         | FeoA family protein | <i>C. raciborskii</i> CS-505    | <i>R. brookii</i> CR12       | 99           | Iron transporter FeoA | KRH97483.1      |
|             | 86         |                     | <i>C. raciborskii</i> CS-508    | <i>R. brookii</i> CR12       | 100          | Iron transporter FeoA | KRH97483.1      |
|             | 86         |                     | <i>C. raciborskii</i> GIHE 2018 | <i>R. brookii</i> CR12       | 99           | Iron transporter FeoA | TPX29624.1      |
|             | 86         |                     | <i>C. raciborskii</i> CR2010    | <i>R. brookii</i> CR12       | 99           | Iron transporter FeoA | UJL32789.1      |
|             | 86         |                     | <i>C. raciborskii</i> KLL07     | <i>R. brookii</i> CR12       | 99           | Iron transporter FeoA | UJS05248.1      |
|             | 86         |                     | <i>C. raciborskii</i> CHAB3438  | <i>R. brookii</i> CR12       | 99           | Iron transporter FeoA | MCH4904046.1    |
|             | 86         |                     | <i>C. raciborskii</i> CS-508    | <i>R. brookii</i> CR12       | 100          | Iron transporter FeoA | KRH97483.1      |
|             | 86         |                     | <i>C. raciborskii</i> DSH       | <i>R. brookii</i> CR12       | 99           | Iron transporter FeoA | WP_057177527.1  |
|             | 86         |                     | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505 | 99           | Iron transporter FeoA | EFA68726.1      |
|             | 86         |                     | <i>C. raciborskii</i> CENA302   | <i>C. raciborskii</i> CS-505 | 99           | Iron transporter FeoA | EFA68726.1      |
|             | 86         |                     | <i>C. raciborskii</i> CS-505    | <i>S. sp</i> SIO1G1          | 71           | Iron transporter FeoA | NET00482.1      |
|             | 86         |                     | <i>C. raciborskii</i> CS-505    | <i>S. sp. LEGE</i> 00249     | 72           | Iron transporter FeoA | MBC5796910.1    |
|             | 86         |                     | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> DSH    | 99           | Iron transporter FeoA | MEE6161293.1    |
| <b>FdxH</b> | 99         | Ferredoxin (2Fe-2S) | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12   | 100          | Ferredoxin            | KRH97484.1      |
|             | 99         |                     | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505 | 100          | Ferredoxin (2Fe-2S)   | EFA68727.1      |
|             | 99         |                     | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505 | 100          | Ferredoxin (2Fe-2S)   | EFA68727.1      |
|             | 99         |                     | <i>C. raciborskii</i> CENA302   | <i>C. raciborskii</i> CS-505 | 99           | Ferredoxin (2Fe-2S)   | EFA68727.1      |
|             | 99         |                     | <i>C. raciborskii</i> ITEP-A1   | <i>C. raciborskii</i> MVCC14 | 100          | Ferredoxin            | OHY36357.1      |

|             |     |                                       |                                 |                                |     |   |              |
|-------------|-----|---------------------------------------|---------------------------------|--------------------------------|-----|---|--------------|
|             | 99  |                                       | <i>C. raciborskii</i> MVCC14    | <i>C. raciborskii</i> CS-505   | 99  | Ferredoxin (2Fe-2S)                                 | EFA68727.1   |
|             | 99  |                                       | <i>C. raciborskii</i> CHAB 3438 | <i>C. raciborskii</i> CS-505   | 99  | Ferredoxin (2Fe-2S)                                 | MCH4904045.1 |
|             | 99  |                                       | <i>C. raciborskii</i> DSH       | <i>C. raciborskii</i> CS-505   | 99  | Ferredoxin (2Fe-2S)                                 | MEE6161294.1 |
|             | 99  |                                       | <i>C. raciborskii</i> GIHE 2018 | <i>C. raciborskii</i> CS-505   | 99  | Ferredoxin (2Fe-2S)                                 | TPX29625     |
|             | 99  |                                       | <i>C. raciborskii</i> KLL07     | <i>C. raciborskii</i> CS-505   | 99  | Ferredoxin (2Fe-2S)                                 | UJS05249.1   |
|             | 99  |                                       | <i>C. raciborskii</i> CR2010    | <i>C. raciborskii</i> CS-505   | 99  | Ferredoxin (2Fe-2S)                                 | UJL32790.1   |
| <b>fdxN</b> | 130 | Ferredoxin (2Fe-2S)                   | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12     | 100 | Ferredoxin  | KRH97498.1   |
|             | 130 |                                       | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505   | 100 | Ferredoxin-like protein in nif region protein, fdxN | EFA68741.1   |
|             | 130 |                                       | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505   | 100 | Ferredoxin-like protein in nif region protein, fdxN | EFA68741.1   |
|             | 112 |                                       | <i>C. raciborskii</i> CENA302   | <i>C. raciborskii</i> CS-505   | 94  | Ferredoxin-like protein in nif region protein, fdxN | EFA68741.1   |
|             | 112 |                                       | <i>C. raciborskii</i> ITEP-A1   | <i>C. raciborskii</i> CS-505   | 94  | Ferredoxin-like protein in nif region protein, fdxN | EFA68741.1   |
|             | 112 |                                       | <i>C. raciborskii</i> MVCC14    | <i>C. raciborskii</i> CS-505   | 94  | Ferredoxin-like protein in nif region protein, fdxN | EFA68741.1   |
|             | 112 |                                       | <i>C. raciborskii</i> DSH       | <i>C. curvispora</i> GIHE G1   | 96  | Ferredoxin-like protein in nif region protein, fdxN | MEE6161308.1 |
|             | 112 |                                       | <i>C. raciborskii</i> CR12      | <i>S. Kisseleviana</i> NIES 73 | 85  | Ferredoxin-like protein in nif region protein, fdxN | BAZ80946.1   |
|             | 112 |                                       | <i>C. raciborskii</i> CR12      | <i>S. sp</i> SIO1G1            | 80  | Ferredoxin-like protein in nif region protein, fdxN | NET02523.1   |
|             | 112 |                                       | <i>C. raciborskii</i> CR12      | <i>S. sp</i> SIO1G1            | 80  | Ferredoxin-like protein in nif region protein, fdxN | NET02523.1   |
| <b>HesA</b> | 259 | UBA/THIF-type NAD/FAD binding protein | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12     | 100 | Protein hesA  | KRH97486.1   |
|             | 259 |                                       | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505   | 100 | MoeZ/MoeB   | EFA68729.1   |

|             |     |  |                                 |                                   |     |                                    |              |
|-------------|-----|--|---------------------------------|-----------------------------------|-----|------------------------------------|--------------|
|             | 259 |  | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505      | 100 | MoeZ/MoeB                          | EFA68729.1   |
|             | 259 |  | <i>C. raciborskii</i> CENA302   | <i>C. raciborskii</i> CS-505      | 99  | MoeZ/MoeB                          | EFA68729.1   |
|             | 259 |  | <i>C. raciborskii</i> ITEP-A1   | <i>C. raciborskii</i> CS-505      | 98  | MoeZ/MoeB                          | EFA68729.1   |
|             | 259 |  | <i>C. raciborskii</i> MVCC14    | <i>C. raciborskii</i> CS-505      | 99  | MoeZ/MoeB                          | EFA68729.1   |
|             | 259 |  | <i>C. raciborskii</i> DSH       | <i>C. curvispora</i> GIHE G1      | 99  | HesA/MoeB                          | MEE6161296.1 |
|             | 259 |  | <i>C. raciborskii</i> CHAB 3438 | <i>C. curvispora</i> GIHE G1      | 99  | Protein hesA                       | MCH4904043.1 |
|             | 259 |  | <i>C. raciborskii</i> GIHE 2018 | <i>C. curvispora</i> GIHE G1      | 99  | HesA/MoeB                          | TPX29627.1   |
|             | 259 |  | <i>C. raciborskii</i> CR2010    | <i>C. curvispora</i> GIHE G1      | 99  | HesA/MoeB                          | UJL32792.1   |
|             | 259 |  | <i>C. raciborskii</i> KLL07     | <i>C. curvispora</i> GIHE G1      | 99  | HesA/MoeB                          | UJS05251.1   |
|             | 259 |  | <i>C. raciborskii</i> CR12      | <i>S. sp. LEGE</i> 08334          | 94  | HesA/MoeB                          | MBE9054908   |
|             | 259 |  | <i>C. raciborskii</i> CR12      | <i>S. torques reginae</i> ITEP024 | 93  | HesA/MoeB                          | CP080598.1   |
|             | 259 |  | <i>C. raciborskii</i> CR12      | <i>S. sp</i> SIO1G1               | 93  | HesA/MoeB                          | NET00479.1   |
| <b>HesB</b> | 121 | Iron-sulfur cluster assembly accessory protein | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12        | 99  | Fe-S cluster assembly protein HesB | KRH97485.1   |
|             | 121 |  | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CR12        | 100 | Fe-S cluster assembly protein HesB | KRH97485.1   |
|             | 121 |  | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505      | 99  | HesB/YadR/YfhF                     | EFA68728.1   |
|             | 121 |  | <i>C. raciborskii</i> CENA302   | <i>C. raciborskii</i> CR12        | 97  | Fe-S cluster assembly protein HesB | KRH97485.1   |
|             | 138 |  | <i>C. raciborskii</i> ITEP-A1   | <i>C. raciborskii</i> MVCC14      | 100 | Fe-S cluster assembly protein HesB | OHY36358.1   |
|             | 121 |  | <i>C. raciborskii</i> MVCC14    | <i>C. raciborskii</i> CR12        | 97  | Fe-S cluster assembly protein HesB | KRH97485.1   |
|             | 121 |  | <i>C. raciborskii</i> CHAB 3438 | <i>C. curvispora</i> GIHE G1      | 92  | Fe-S cluster assembly protein HesB | MCH4904044.1 |
|             | 121 |  | <i>C. raciborskii</i> DSH       | <i>C. curvispora</i> GIHE G1      | 92  | Fe-S cluster assembly protein HesB | MEE6161295.1 |

|             |     |  |                                 |                                   |     |   |              |
|-------------|-----|--|---------------------------------|-----------------------------------|-----|---|--------------|
|             | 121 |  | <i>C. raciborskii</i> CR2010    | <i>C. curvispora</i> GIHE G1      | 92  | Fe-S cluster assembly protein HesB              | UJL32791.1   |
|             | 121 |  | <i>C. raciborskii</i> KLL07     | <i>C. curvispora</i> GIHE G1      | 92  | Fe-S cluster assembly protein HesB              | UJS05250     |
|             | 121 |  | <i>C. raciborskii</i> GIHE 2018 | <i>C. curvispora</i> GIHE G1      | 92  | Fe-S cluster assembly protein HesB              | TPX29626.1   |
| <b>NifD</b> | 501 | Nitrogenase molybdenum-ion protein alpha chain | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12        | 99  | Nitrogenase molybdenum-iron protein alpha chain | KRH97494.1   |
|             | 501 | Nitrogenase molybdenum-ion protein alpha chain | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505      | 100 | Nitrogenase molybdenum-iron protein alpha chain | EFA68737.1   |
|             | 501 | Nitrogenase molybdenum-ion protein alpha chain | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505      | 99  | Nitrogenase molybdenum-iron protein alpha chain | EFA68737.1   |
|             | 501 | Nitrogenase molybdenum-ion protein alpha chain | <i>C. raciborskii</i> CENA302   | <i>C. raciborskii</i> CS-505      | 99  | Nitrogenase molybdenum-iron protein alpha chain | EFA68737.1   |
|             | 501 | Nitrogenase molybdenum-ion protein alpha chain | <i>C. raciborskii</i> ITEP-A1   | <i>C. raciborskii</i> MVCC14      | 100 | Nitrogenase molybdenum-iron protein alpha chain | OHY36367.1   |
|             | 501 | Nitrogenase molybdenum-ion protein alpha chain | <i>C. raciborskii</i> MVCC14    | <i>C. raciborskii</i> CS-505      | 99  | Nitrogenase molybdenum-iron protein alpha chain | EFA68737.1   |
|             | 501 | Nitrogenase molybdenum-ion protein alpha chain | <i>C. raciborskii</i> DSH       | <i>C. curvispora</i> GIHE G1      | 100 | Nitrogenase molybdenum-iron protein alpha chain | MEE6161304.1 |
|             | 501 | Nitrogenase molybdenum-ion protein alpha chain | <i>C. raciborskii</i> GIHE 2018 | <i>C. curvispora</i> GIHE G1      | 100 | Nitrogenase molybdenum-iron protein alpha chain | TPX29635.1   |
|             | 501 | Nitrogenase molybdenum-ion protein alpha chain | <i>C. raciborskii</i> CR12      | <i>S. torques reginae</i> ITEP024 | 96  | Nitrogenase molybdenum-iron protein             | QYX31286.1   |



|             |     |  |                                 |                                  |     |  |              |
|-------------|-----|--|---------------------------------|----------------------------------|-----|--|--------------|
|             |     |  |                                 |                                  |     | alpha chain<br>Nitrogenase<br>molybdenum-iron protein<br>alpha chain |              |
|             | 501 | Nitrogenase molybdenum-<br>ion protein alpha chain   | <i>C. raciborskii</i> CR12      | <i>S. sp. LEGE 00249</i>         | 96  |  | MBC5795059.1 |
| <b>NifE</b> | 460 | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12       | 99  | Nitrogenase<br>molybdenum-cofactor<br>biosynthesis protein NifE      | KRH97492.1   |
|             | 460 |  | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-<br>505 | 100 | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein NifE            | EFA68735.1   |
|             | 460 |  | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-<br>505 | 99  | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein NifE            | EFA68735.1   |
|             | 460 |  | <i>C. raciborskii</i> CENA302   | <i>C. raciborskii</i> CR12       | 99  | Nitrogenase<br>molybdenum-cofactor<br>biosynthesis protein NifE      | KRH97492.1   |
|             | 460 |  | <i>C. raciborskii</i> ITEP-A1   | <i>C. raciborskii</i> CS-<br>505 | 99  | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein NifE            | EFA68735.1   |
|             | 460 |  | <i>C. raciborskii</i> MVCC14    | <i>C. raciborskii</i> CS-<br>505 | 99  | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein NifE            | EFA68735.1   |
|             | 460 |  | <i>C. raciborskii</i> DSH       | <i>C. curvispora</i><br>GIHE-G1  | 98  | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein NifE            | MEE6161302.1 |
|             | 460 |  | <i>C. raciborskii</i> GIHE 2018 | <i>S. FACHB 1194</i>             | 95  | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein NifE            | TPX29633.1   |
|             | 460 |  | <i>C. raciborskii</i> CR2010    | <i>C. curvispora</i><br>GIHE- G1 | 98  | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein NifE            | UJL32798.1   |
|             | 460 |  | <i>C. raciborskii</i> KLL07     | <i>C. curvispora</i><br>GIHE- G1 | 98  | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein NifE            | UJS05257.1   |
|             | 295 | Nitrogenase iron protein                             | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12       | 99  | Nitrogenase reductase  | KRH97495.1   |
|             | 295 |  | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-        | 100 | Nitrogenase iron protein   | EFA68738.1   |

|             |     |   |                                 |                              |     |  |              |
|-------------|-----|---|---------------------------------|------------------------------|-----|--|--------------|
| <b>NifH</b> |     | <b>NifH</b>                               |                                 | 505                          |     |  |              |
|             | 295 |   | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505 | 99  | Nitrogenase iron protein                       | EFA68738.1   |
|             | 295 |   | <i>C. raciborskii</i> CENA302   | <i>C. raciborskii</i> CR12   | 99  | Nitrogenase reductase                          | KRH97495.1   |
|             | 295 |   | <i>C. raciborskii</i> ITEP-A1   | <i>C. raciborskii</i> CR12   | 99  | Nitrogenase reductase                          | KRH97495.1   |
|             | 295 |   | <i>C. raciborskii</i> MVCC14    | <i>C. raciborskii</i> CR12   | 99  | Nitrogenase reductase                          | KRH97495.1   |
|             | 295 |   | <i>C. raciborskii</i> DSH       | <i>C. curvispora</i> GIHE G1 | 99  | Nitrogenase reductase                          | MEE6161305.1 |
|             | 295 |   | <i>C. raciborskii</i> GIHE 2018 | <i>C. curvispora</i> GIHE G1 | 99  | Nitrogenase reductase                          | TPX29636.1   |
|             | 295 |   | <i>C. raciborskii</i> CR2010    | <i>C. curvispora</i> GIHE G1 | 99  | Nitrogenase reductase                          | UJL32801.1   |
|             | 295 |   |                                 |                              |     |  |              |
|             | 295 |   | <i>C. raciborskii</i> KLL07     | <i>C. curvispora</i> GIHE G1 | 99  | Nitrogenase reductase                          | UJS05260.1   |
|             |     |   |                                 |                              |     |  |              |
| <b>NifB</b> | 477 | Nitrogenase cofactor biosynthesis protein | <i>C. raciborskii</i> CS-505    | <i>C. raciborskii</i> CR12   | 100 | Nitrogen fixation protein NifB                 | KRH97499.1   |
|             | 477 |   | <i>C. raciborskii</i> CS-508    | <i>C. raciborskii</i> CS-505 | 100 | Nitrogenase cofactor biosynthesis protein NifB | EFA68742.1   |
|             | 477 |   | <i>C. raciborskii</i> CR12      | <i>C. raciborskii</i> CS-505 | 100 | Nitrogenase cofactor biosynthesis protein NifB | EFA68742.1   |
|             | 478 |   | <i>C. raciborskii</i> CENA302   | <i>C. raciborskii</i> MVCC14 | 100 | Nitrogenase cofactor biosynthesis protein NifB | OHY36372.1   |
|             | 478 |   | <i>C. raciborskii</i> ITEP-A1   | <i>C. raciborskii</i> CS-505 | 99  | Nitrogenase cofactor biosynthesis protein NifB | EFA68742.1   |

|             |     |                           |                                   |                                   |     |  |                |
|-------------|-----|---------------------------|-----------------------------------|-----------------------------------|-----|--|----------------|
|             | 478 |                           | <i>C. raciborskii</i> MVCC14      | <i>C. raciborskii</i> CS-505      | 99  | Nitrogenase cofactor biosynthesis protein NifB | EFA68742.1     |
|             | 477 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> DSH         | 99  | Nitrogenase cofactor biosynthesis protein NifB | MEE6161309.1   |
|             | 477 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> CS-508      | 98  | Nitrogenase cofactor biosynthesis protein NifB | OHY32534.1     |
|             | 477 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> CR2010      | 98  | Nitrogenase cofactor biosynthesis protein NifB | UJL32805.1     |
|             | 477 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> KLL07       | 98  | Nitrogenase cofactor biosynthesis protein NifB | UJS05264.1     |
|             | 477 |                           | <i>S. sp. LEGE</i> 08334          | <i>S. torques reginae</i> ITEP024 | 100 | Nitrogenase cofactor biosynthesis protein NifB | WP_220609327.1 |
|             | 477 |                           | <i>S. torques reginae</i> ITEP024 | <i>S. sp. LEGE</i> 00834          | 100 | Nitrogenase cofactor biosynthesis protein NifB | WP_194059990.1 |
|             | 477 |                           | <i>S. sp. LEGE</i> 08334          | <i>S. sp. LEGE</i> 00249          | 99  | Nitrogenase cofactor biosynthesis protein NifB | WP_187039462   |
|             | 477 |                           | <i>S. torques reginae</i> ITEP024 | <i>S. sp. SIO1G1</i>              | 94  | Nitrogenase cofactor biosynthesis protein NifB | NET02524.1     |
|             | 477 |                           |                                   |                                   |     |  |                |
| <b>NifW</b> | 105 | Nitrogen fixation protein | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i> CR12        | 100 | Nitrogenase stabilizing/protective protein     | KRH97487.1     |
|             | 105 |                           | <i>C. raciborskii</i> CS-508      | <i>C. raciborskii</i> CS-505      | 100 | Nitrogen fixation protein NifW                 | EFA68730.1     |
|             | 105 |                           | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> CS-505      | 100 | Nitrogen fixation protein NifW                 | EFA68730.1     |
|             | 105 |                           | <i>C. raciborskii</i> CENA302     | <i>C. raciborskii</i> CS-505      | 96  | Nitrogen fixation protein NifW                 | EFA68730.1     |
|             | 105 |                           | <i>C. raciborskii</i> ITEP-A1     | <i>C. raciborskii</i> CS-505      | 96  | Nitrogen fixation protein NifW                 | EFA68730.1     |
|             | 105 |                           | <i>C. raciborskii</i> MVCC14      | <i>C. raciborskii</i> CS-505      | 96  | Nitrogen fixation protein NifW                 | EFA68730.1     |
|             | 105 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> CHAB 3438   | 94  | Nitrogenase stabilizing/protective protein     | KRH97487.1     |
|             | 105 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> DSH         | 94  | Nitrogenase stabilizing/protective             | MEE6161297     |
|             | 105 |                           |                                   |                                   |     |  |                |

|             |     |                           |                                   |                                   |     |   |                |
|-------------|-----|---------------------------|-----------------------------------|-----------------------------------|-----|---|----------------|
|             |     |                           |                                   |                                   |     | protein   |                |
|             | 105 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> CS-508      | 94  | Nitrogenase stabilizing/protective protein          | OHY32546.1     |
|             | 105 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> GIHE 2018   | 94  | Nitrogenase stabilizing/protective protein          | TPX29628.1     |
|             | 105 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> CR2010      | 94  | Nitrogenase stabilizing/protective protein          | UJL32793.1     |
|             | 105 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> KLL07       | 94  | Nitrogenase stabilizing/protective protein          | UJS05252.1     |
|             | 105 |                           | <i>S. torques reginae</i> ITEP024 | <i>S. sp. LEGE 00249</i>          | 98  | Nitrogenase stabilizing/protective protein          | WP_187040960.1 |
|             | 105 |                           | <i>S. sp. LEGE 00249</i>          | <i>S. torques reginae</i> ITEP024 | 98  | Nitrogenase stabilizing/protective protein          | WP_220609369.1 |
| <b>NifX</b> | 139 | Nitrogen fixation protein | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i> CR12        | 100 | Nitrogen fixation protein NifX                      | KRH97490.1     |
|             | 139 |                           | <i>C. raciborskii</i> CS-508      | <i>C. raciborskii</i> CS-505      | 100 | Dinitrogenase iron-molybdenum cofactor biosynthesis | EFA68733.1     |
|             | 139 |                           | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> CS-505      | 100 | Dinitrogenase iron-molybdenum cofactor biosynthesis | EFA68733.1     |
|             | 139 |                           | <i>C. raciborskii</i> CENA302     | <i>C. raciborskii</i> CS-505      | 97  | Dinitrogenase iron-molybdenum cofactor biosynthesis | EFA68733.1     |
|             | 139 |                           | <i>C. raciborskii</i> ITEP-A1     | <i>C. raciborskii</i> CS-505      | 97  | Dinitrogenase iron-molybdenum cofactor biosynthesis | EFA68733.1     |
|             | 139 |                           | <i>C. raciborskii</i> MVCC14      | <i>C. raciborskii</i> CS-505      | 97  | Dinitrogenase iron-molybdenum cofactor biosynthesis | EFA68733.1     |
|             | 139 |                           | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> CHAB 3438   | 97  | Nitrogen fixation protein NifX                      | MCH4904039     |

|             |     |   |                               |                                 |     |  |                |
|-------------|-----|---|-------------------------------|---------------------------------|-----|--|----------------|
|             | 139 |   | <i>C. curvispora</i> GIHE G1  | <i>C. raciborskii</i> DSH       | 97  | Nitrogen fixation protein NifX                                 | MEE6161300.1   |
|             | 139 |   | <i>C. curvispora</i> GIHE G1  | <i>C. raciborskii</i> CR2010    | 97  | Nitrogen fixation protein NifX                                 | UJL32796.1     |
|             | 139 |   | <i>C. curvispora</i> GIHE G1  | <i>C. raciborskii</i> KLL07     | 97  | Nitrogen fixation protein NifX                                 | UJS05255.1     |
|             | 139 |   | <i>C. curvispora</i> GIHE G1  | <i>S. sp</i> SIO1G1             | 85  | Nitrogen fixation protein NifX                                 | NET02489.1     |
|             | 139 |   | <i>C. curvispora</i> GIHE G1  | <i>S. sp. LEGE</i> 00249        | 92  | Nitrogen fixation protein NifX                                 | WP_187040961.1 |
| <b>NifN</b> | 443 | Nitrogenase molybdenum-ion cofactor biosynthesis protein NifN | <i>C. raciborskii</i> CS-505  | <i>C. raciborskii</i> CR12      | 100 | Nitrogenase molybdenum-cofactor biosynthesis protein NifN      | KRH97491.1     |
|             | 443 |   | <i>C. raciborskii</i> CS-508  | <i>C. raciborskii</i> CR12      | 100 | Nitrogenase molybdenum-cofactor biosynthesis protein NifN      | KRH97491.1     |
|             | 443 |   | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CR12      | 100 | Nitrogenase molybdenum-cofactor biosynthesis protein NifN      | KRH97491.1     |
|             | 443 |   | <i>C. raciborskii</i> CENA302 | <i>C. raciborskii</i> CR12      | 100 | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | KRH97491.1     |
|             | 443 |   | <i>C. raciborskii</i> ITEP-A1 | <i>C. raciborskii</i> CR12      | 100 | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | KRH97491.1     |
|             | 443 |   | <i>C. raciborskii</i> MVCC14  | <i>C. raciborskii</i> CR12      | 99  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | KRH97491.1     |
|             | 443 |   | <i>C. raciborskii</i> CS-505  | <i>C. raciborskii</i> CR12      | 100 | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | KRH97493.1     |
|             | 443 |   | <i>C. raciborskii</i> CR12    | <i>C. raciborskii</i> CHAB 3438 | 100 | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | MCH4904038.1   |

|  |     |  |                                   |                                   |     |  |                |
|--|-----|--|-----------------------------------|-----------------------------------|-----|--|----------------|
|  | 443 |  | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> DSH         | 100 | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | MEE6161301.1   |
|  | 443 |  | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> GIHE 2018   | 100 | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | TPX29632.1     |
|  | 443 |  | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> CR2010      | 100 | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | UJL32797.1     |
|  | 443 |  | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> KLL07       | 100 | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | UJS05256.1     |
|  | 443 |  | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> CHAB 3438   | 99  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | MCH4904038.1   |
|  | 443 |  | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> DSH         | 99  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | MEE6161301.1   |
|  | 443 |  | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> GIHE 2018   | 99  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | TPX29632.1     |
|  | 443 |  | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> CR2010      | 99  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | UJL32797.1     |
|  | 443 |  | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> KLL07       | 99  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | UJS05256.1     |
|  | 442 |  | <i>S. sp.</i> LEGE 00249          | <i>S. torques reginae</i> ITEP024 | 99  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | WP_220609367.1 |
|  | 442 |  | <i>S. sp.</i> LEGE 00249          | <i>S. sp.</i> FACHB 1194          | 99  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | WP_190346642.1 |
|  | 442 |  | <i>S. torques reginae</i> ITEP024 | <i>S. sp.</i> LEGE 08334          | 99  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | WP_194051913.1 |
|  | 442 |  | <i>S. sp.</i> FACHB 1194          | <i>S. sp.</i> LEGE 00249          | 99  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | WP_187040962.1 |

|  |     |  |                                   |                                   |     |  |                |
|--|-----|--|-----------------------------------|-----------------------------------|-----|--|----------------|
|  | 511 |  | <i>C. raciborskii</i> CS-508      | <i>C. raciborskii</i> CS-505      | 100 | Nitrogenase molybdenum-iron protein beta chain | EFA68736.1     |
|  | 511 |  | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> CS-505      | 100 | Nitrogenase molybdenum-iron protein beta chain | EFA68736.1     |
|  | 511 |  | <i>C. raciborskii</i> CENA302     | <i>C. raciborskii</i> CS-505      | 99  | Nitrogenase molybdenum-iron protein beta chain | EFA68736.1     |
|  | 511 |  | <i>C. raciborskii</i> ITEP-A1     | <i>C. raciborskii</i> CS-505      | 99  | Nitrogenase molybdenum-iron protein beta chain | EFA68736.1     |
|  | 511 |  | <i>C. raciborskii</i> MVCC14      | <i>C. raciborskii</i> CS-505      | 99  | Nitrogenase molybdenum-iron protein beta chain | EFA68736.1     |
|  | 511 |  | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> DSH         | 99  | Nitrogenase molybdenum-iron protein beta chain | MEE6161303.1   |
|  | 511 |  | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> GIHE 2018   | 99  | Nitrogenase molybdenum-iron protein beta chain | TPX29634.1     |
|  | 511 |  | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> CR2010      | 99  | Nitrogenase molybdenum-iron protein beta chain | UJL32799.1     |
|  | 511 |  | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> KLL07       | 99  | Nitrogenase molybdenum-iron protein beta chain | UJS05258.1     |
|  | 511 |  | <i>S. sp.</i> LEGE 08334          | <i>S. torques reginae</i> ITEP024 | 100 | Nitrogenase molybdenum-iron protein beta chain | WP_220609365.1 |
|  | 511 |  | <i>S. torques reginae</i> ITEP024 | <i>S. sp.</i> LEGE 08334          | 100 | Nitrogenase molybdenum-iron protein beta chain | WP_194051907.1 |
|  | 308 |  | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i> CR12        | 99  | Fe-S cluster assembly protein NifU             | KRH97496.1     |
|  | 308 |  | <i>C. raciborskii</i> CS-508      | <i>C. raciborskii</i> CS-505      | 100 | Fe-S cluster assembly protein NifU             | EFA68739.1     |
|  | 308 |  | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> CS-505      | 99  | Fe-S cluster assembly protein NifU             | EFA68739.1     |

|      |     |                                    |                                   |                                   |     |                                    |                |
|------|-----|------------------------------------|-----------------------------------|-----------------------------------|-----|------------------------------------|----------------|
| NifU | 308 | Fe-S cluster assembly protein NifU | <i>C. raciborskii</i> CENA302     | <i>C. raciborskii</i> CS-505      | 97  | Fe-S cluster assemblyprotein NifU  | EFA68739.1     |
|      | 308 |                                    | <i>C. raciborskii</i> ITEP-A1     | <i>C. raciborskii</i> CS-505      | 97  | Fe-S cluster assembly protein NifU | EFA68739.1     |
|      | 308 |                                    | <i>C. raciborskii</i> MVCC14      | <i>C. raciborskii</i> CS-505      | 97  | Fe-S cluster assembly protein NifU | EFA68739.1     |
|      | 308 |                                    | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> DSH         | 97  | Fe-S cluster assembly protein NifU | MEE6161306.1   |
|      | 308 |                                    | <i>S. sp. LEGE 08334</i>          | <i>S. torques reginae</i> ITEP024 | 100 | Fe-S cluster assembly protein NifU | WP_220609329.1 |
|      | 308 |                                    | <i>S. torques reginae</i> ITEP024 | <i>S. LEGE 08334</i>              | 100 | Fe-S cluster assembly protein NifU | WP_194059988.1 |
|      | 308 |                                    | <i>S.sp. LEGE 08334</i>           | <i>S. sp SIO1G1</i>               | 92  | Fe-S cluster assembly protein NifU | NET02521.1     |
|      |     |                                    |                                   |                                   |     |                                    |                |
|      | 401 |                                    | <i>C. raciborskii</i> CS-505      | <i>C. raciborskii</i> CR12        | 100 | Cysteine desulfurase NifS          | KRH97497.1     |



|      |     |   |                                   |                                   |     |   |                |
|------|-----|---|-----------------------------------|-----------------------------------|-----|---|----------------|
| NifS | 401 | Nitrogenase<br>metalloclusters<br>biosynthesis protein NifS | <i>C. raciborskii</i> CS-508      | <i>C. raciborskii</i> CR12        | 100 | Cysteine desulfurase NifS                                     | KRH97497.1     |
|      | 401 |   | <i>C. raciborskii</i> CR12        | <i>C. raciborskii</i> CS-505      | 100 | Aromatic amino acid beta-eliminating lyase/threonine aldolase | EFA68740.1     |
|      | 401 |   | <i>C. raciborskii</i> CENA302     | <i>C. raciborskii</i> CS-505      | 99  | Aromatic amino acid beta-eliminating lyase/threonine aldolase | EFA68740.1     |
|      | 401 |   | <i>C. raciborskii</i> ITEP-A1     | <i>C. raciborskii</i> CS-505      | 99  | Aromatic amino acid beta-eliminating lyase/threonine aldolase | EFA68740.1     |
|      | 401 |   | <i>C. raciborskii</i> MVCC14      | <i>C. raciborskii</i> CS-505      | 99  | Aromatic amino acid beta-eliminating lyase/threonine aldolase | EFA68740.1     |
|      | 401 |   | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> DSH         | 99  | Cysteine desulfurase NifS                                     | MEE6161307.1   |
|      | 401 |   | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> GIHE 2018   | 99  | Cysteine desulfurase NifS                                     | TPX29638       |
|      | 401 |   | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> CR2010      | 99  | Cysteine desulfurase NifS                                     | UJL32803.1     |
|      | 401 |   | <i>C. curvispora</i> GIHE G1      | <i>C. raciborskii</i> KLL07       | 99  | Cysteine desulfurase NifS                                     | UJS05262.1     |
|      | 403 |   | <i>S. LEGE</i> 08334              | <i>S. torques reginae</i> ITEP024 | 99  | Cysteine desulfurase NifS                                     | WP_220609328.1 |
|      | 403 |   | <i>S. Kisseleviana</i> NIES 73    | <i>S. sp</i> SIO1G1               | 94  | Cysteine desulfurase NifS                                     | NET02522.1     |
|      | 403 |   | <i>S. torques reginae</i> ITEP024 | <i>S. sp. LEGE</i> 08334          | 99  | Cysteine desulfurase NifS                                     | WP_194059993.1 |

**Supplementary Table S11. Proposed function of the proteins encoded by heterocyte glycolipid biosynthetic gene cluster and two proteins involved in the regulation of heterocyte differentiation in the genomes of the group *Sphaerospermopsis*.**

| Protein | Aminoacids | Product                   | Strain                            | Organism                          | Identity (%) | Function                          | Acession number |
|---------|------------|---------------------------|-----------------------------------|-----------------------------------|--------------|-----------------------------------|-----------------|
| HglA    | 556        | 2-nitropropane dioxygenas | <i>S. torques reginae</i> ITEP024 | <i>S. sp</i> LEGE 08334           | 99           | 2-nitropropane dioxygenas         | WP_194054992.1  |
|         | 556        |                           | <i>S. Kisseleviana</i> NIES 73    | <i>S. sp.</i> LEGE 00249          | 99           | family polyunsaturated fatty acid | MBC5793826.1    |
|         | 556        |                           | <i>S. sp</i> SIO1G1               | <i>S. torques reginae</i> ITEP024 | 88           |                                   | WP_220610685.1  |
|         | 556        |                           | <i>S. sp.</i> FACHB 1194          | <i>S. sp.</i> LEGE 08334          | 99           |                                   | WP_194054992.1  |
|         | 556        |                           | <i>S. sp.</i> LEGE 08334          | <i>S. sp.</i> LEGE 00249          | 99           |                                   | MBC5793826.1    |
|         | 556        |                           | <i>S. sp.</i> FACHB 1194          | <i>S. sp.</i> LEGE 08334          | 99           |                                   | WP_194054992.1  |
|         | 556        |                           |                                   |                                   |              |                                   | MBC5793826.1    |
|         | 556        |                           | <i>S. sp.</i> LEGE 08334          | <i>S. sp.</i> LEGE 00249          | 99           |                                   |                 |
|         | 556        |                           | <i>S. sp.</i> LEGE 00249          | <i>S. Kisseleviana</i> NIES 73    | 99           |                                   | WP_096572229.1  |
|         | 556        |                           | <i>S. torques reginae</i> ITEP024 | <i>S. sp.</i> LEGE 08334          | 99           |                                   | WP_194054992.1  |
|         | 556        |                           |                                   |                                   |              |                                   |                 |
|         | 556        |                           | <i>S. Kisseleviana</i> NIES 73    | <i>S. sp.</i> LEGE 00249          | 99           |                                   | MBC5793826.1    |
| HglB    | 497        | Polyketide synthase       | <i>S. sp</i> SIO1G1               | <i>S. sp.</i> FACHB 1194          | 84           | thioester                         | WP_190647549.1  |
|         | 505        |                           | <i>S. sp.</i> FACHB 1194          | <i>S. Kisseleviana</i> NIES 73    | 100          |                                   | WP_096572228.1  |

|      |      |                                |                                   |                                   |                                   |                                     |                          |
|------|------|--------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-------------------------------------|--------------------------|
|      | 505  |                                | <i>S. sp. LEGE 08334</i>          | <i>S. torques reginae</i> ITEP024 | 98                                | reductase                           | WP_220610686.1           |
|      | 505  |                                | <i>S. sp. LEGE 00249</i>          | <i>S. Kisseleviana NIES 73</i>    | 99                                |                                     | WP_096572228.1           |
| HgIC | 1106 | omega-3                        | <i>S. Kisseleviana NIES 73</i>    | <i>S. FACHB 1194</i>              | 99                                | polyketide synthase                 | WP_242027320.1           |
|      | 1106 |                                | <i>S. sp. FACHB 1194</i>          | <i>S. Kisseleviana NIES 73</i>    | 100                               |                                     | WP_272109449.1           |
|      | 1106 |                                | <i>S. torques reginae ITEP024</i> | <i>S. sp. LEGE 08334</i>          | 97                                |                                     | WP_194054993.1           |
|      | 1106 |                                | <i>S. sp. LEGE 08334</i>          | <i>S. torques reginae ITEP024</i> | 98                                |                                     | WP_220610684.1           |
|      | 1106 |                                | <i>S. sp. LEGE 00249</i>          | <i>S. Kisseleviana NIES 73</i>    | 98                                |                                     | WP_272109449.1           |
|      | 1089 |                                | <i>S. sp SIO1G1</i>               | <i>S. torques reginae ITEP024</i> | 85                                |                                     | WP_220610684.1           |
|      | HgID |                                | 450                               | polyketide synthase               | <i>S. torques reginae ITEP024</i> |                                     | <i>S. sp. LEGE 08334</i> |
| 450  |      | <i>S. Kisseleviana NIES 73</i> | <i>S. sp. LEGE 00249</i>          |                                   | 99                                | MBC5793824.1                        |                          |
| 451  |      | <i>S. sp SIO1G1</i>            | <i>S. sp. LEGE 08334</i>          |                                   | 89                                | WP_194054994.1                      |                          |
| 450  |      | <i>S. sp. FACHB 1194</i>       | <i>S. sp. LEGE 00249</i>          |                                   | 99                                | MBC5793824.1                        |                          |
| 450  |      | <i>S. sp. LEGE 08334</i>       | <i>S. torques reginae ITEP024</i> |                                   | 98                                | WP_220610683                        |                          |
| 450  |      | <i>S. sp. LEGE 00249</i>       | <i>S. torques reginae ITEP024</i> |                                   | 98                                | WP_220610683                        |                          |
| 450  |      | <i>S. sp. LEGE 00249</i>       | <i>S. torques reginae ITEP024</i> |                                   | 98                                | WP_220610683                        |                          |
|      |      |                                |                                   |                                   |                                   |                                     |                          |
| HgIE | 1826 | polyketide synthase            | <i>S. torques reginae ITEP024</i> | <i>S. LEGE 08334</i>              | 99                                | polyketide synthase                 | WP_220611797.1           |
|      | 1807 |                                | <i>S.Kisseleviana NIES 73</i>     | <i>S.Aphanizomenoids</i>          | 99                                |                                     | WP_193941129.1           |
|      | 1788 |                                | <i>S. sp SIO1G1</i>               | <i>S. torques reginae ITEP024</i> | 80                                |                                     | WP_220611797.1           |
|      | 1801 |                                | <i>S. sp. FACHB 1194</i>          | <i>S.Kisseleviana NIES 73</i>     | 99                                |                                     | WP_096572232.1           |
|      | 1793 |                                | <i>S. sp. LEGE 08334</i>          | <i>S. torques reginae ITEP024</i> | 99                                |                                     | WP_220611797.1           |
|      | 1802 |                                | <i>S. sp. LEGE 00249</i>          | <i>S.kisseleviana NIES73</i>      | 99                                |                                     | WP_096572232.1           |
| HgIG | 334  | Polyketide synthase            | <i>S. torques reginae ITEP024</i> | <i>S. sp. LEGE 08334</i>          | 99                                | NAD-dependent epimerase/dehydratase | WP_194054998.1           |
|      | 334  |                                | <i>S. Kisseleviana NIES 73</i>    | <i>S. sp. LEGE 00249</i>          | 100                               |                                     | WP_18703495.1            |
|      | 334  |                                | <i>S. sp SIO1G1</i>               | <i>S. sp. LEGE 00249</i>          | 85                                |                                     | WP_187038495.1           |
|      | 334  |                                | <i>S. sp. FACHB 1194</i>          | <i>S. sp. LEGE 00249</i>          | 100                               |                                     | WP_187038495.1           |
|      | 334  |                                | <i>S. sp. LEGE 08334</i>          | <i>S. torques reginae ITEP024</i> | 99                                |                                     | WP_220610678.1           |

|      |     |  |                                   |                                   |     |  |                |
|------|-----|--|-----------------------------------|-----------------------------------|-----|--|----------------|
|      | 334 |  | <i>S. sp. LEGE 00249</i>          | <i>S. sp. LEGE 08334</i>          | 99  |  | WP_194054998.1 |
| HetR | 299 | Heterocyst differentiation control protein | <i>S. torques reginae ITEP024</i> | <i>S. sp. LEGE 08334</i>          | 100 | Heterocyst differentiation control protein | QYX31586.1     |
|      | 299 |  | <i>S. torques reginae ITEP024</i> | <i>S. sp. LEGE 00249</i>          | 99  | Heterocyst differentiation control protein | MBC5794768.1   |
|      | 299 |  | <i>S. torques reginae ITEP024</i> | <i>S. sp. SIO1G1</i>              | 96  | Heterocyst differentiation control protein | NET01886.1     |
| HetN | 263 | SDR family NAD(P)-dependent oxidoreductase | <i>S. torques reginae ITEP024</i> | <i>S. sp. LEGE 08334</i>          | 99  | SDR family NAD(P)-dependent oxidoreductase | WP_194057199.1 |
|      | 263 |  | <i>S. torques reginae ITEP024</i> | <i>S. sp. LEGE 08334</i>          | 99  |  | WP_194057199.1 |
|      | 267 |  | <i>S. torques reginae ITEP024</i> | <i>S. Kisseleviana NIES 73</i>    | 95  |  | WP_096572227.1 |
|      | 267 |  | <i>S. Kisseleviana NIES 73</i>    | <i>S. torques reginae ITEP024</i> | 94  |  | WP_220610687.1 |
|      | 267 |  | <i>S. sp SIO1G1</i>               | <i>S. sp. LEGE 08334</i>          | 80  |  | WP_194054985.1 |
|      | 267 |  | <i>S. sp. FACHB 1194</i>          | <i>S. Kisseleviana NIES 73</i>    | 99  |  | WP_096572227.1 |
|      | 267 |  | <i>S. sp. LEGE 08334</i>          | <i>S. Kisseleviana NIES 73</i>    | 94  |  | WP_096572227.1 |
|      | 267 |  | <i>S. sp. LEGE 00249</i>          | <i>S. Kisseleviana NIES 73</i>    | 100 |  | WP_096572227.1 |
|      |     |  |                                   |                                   |     |  |                |

|  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|

**Supplementary Table S12. Proposed function of the proteins encoded by the nitrogen fixation biosynthetic gene cluster of the group *Sphaerospermopsis*.**

| Protein     | Aminoacids | Proposed function      | Strain                              | Organism                            | Identity (%) | Function              | Acession number |
|-------------|------------|------------------------|-------------------------------------|-------------------------------------|--------------|-----------------------|-----------------|
| <b>FeoA</b> | 90         | FeoA family protein    | <i>S. sp. LEGE 00249</i>            | <i>S. kisseleviana NIES-73</i>      | 93           | Iron transporter FeoA | BAZ80902.1      |
|             | 115        | FeoA family protein    | <i>S. kisseleviana NIES-73</i>      | <i>S. sp. LEGE 00249</i>            | 93           | Iron transporter FeoA | MBC5796910.1    |
|             | 83         | FeoA family protein    | <i>S. sp. SIO1G1</i>                | <i>S. sp. LEGE 00249</i>            | 75           | Iron transporter FeoA | MBC5796910.1    |
|             | 99         | Ferredoxin (2Fe-2S)    | <i>S. torquees reginae ITEP-024</i> | <i>S. sp. LEGE 08334</i>            | 99           | Iron transporter FeoA | MBE9054910.1    |
|             | 99         | Ferredoxin (2Fe-2S)    | <i>S. torquees reginae ITEP-024</i> | <i>S. sp. SIO1G1</i>                | 82           | Iron transporter FeoA | NET00481.1      |
| <b>FdxH</b> | 99         | Ferredoxin (2Fe-2S)    | <i>S. sp. LEGE 08334</i>            | <i>S. torquees reginae ITEP-024</i> | 99           | Ferredoxin (2Fe-2S)   | QYX31309.1      |
|             | 99         | Ferredoxin (2Fe-2S)    | <i>S. torquees reginae ITEP-024</i> | <i>S. sp. LEGE 08334</i>            | 99           | Ferredoxin (2Fe-2S)   | MBE9054910.1    |
| <b>FdxN</b> | 112        | 4Fe-4S binding protein | <i>S. Kisseleviana NIES 73</i>      | <i>S. sp SIO1G1</i>                 | 85           | Ferredoxin(4Fe-4S)    | NET02523.1      |

|             |     |  |                                       |                                |     |  |              |
|-------------|-----|--|---------------------------------------|--------------------------------|-----|--|--------------|
| <b>HesA</b> | 259 | UBA/THIF-type<br>NAD/FAD binding protein             | <i>S. torques reginae</i><br>ITEP-024 | <i>S. sp. LEGE 08334</i>       | 100 | HesA/MoeB  | MBE9054908.1 |
|             | 259 | UBA/THIF-type<br>NAD/FAD binding protein             | <i>S. torques reginae</i><br>ITEP-024 | <i>S. sp SIO1G1</i>            | 95  | HesA/MoeB  | NET00479.1   |
| <b>HesB</b> | 121 | Iron-sulfur cluster<br>assembly accessory<br>protein | <i>S. sp. SIO1G1</i>                  | <i>S. sp. LEGE 00249</i>       | 85  | Fe-S cluster<br>assembly protein<br>HesB                               | MBC5796912.1 |
| <b>NifD</b> | 460 | Nitrogenase<br>molybdenum-ion protein<br>alpha chain | <i>S. torques reginae</i> ITEP-024    | <i>S. LEGE 00249</i>           | 99  | Nitrogenase<br>molybdenum-iron<br>protein alpha chain                  | MBC5795059.1 |
| <b>NifE</b> | 455 | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein | <i>S. torques reginae</i> ITEP-024    | <i>S. kisseleviana</i> NIES-73 | 99  | Nitrogenase<br>molybdenum-<br>cofactor<br>biosynthesis<br>protein NifE | BAZ80912.1   |
|             | 455 | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein | <i>S. torques reginae</i> ITEP-024    | <i>S. sp. LEGE 08334</i>       | 99  | Nitrogenase MoFe<br>cofactor<br>biosynthesis<br>protein NifE           | MBE9054902.1 |
|             | 455 | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein | <i>S. torques reginae</i> ITEP-024    | <i>S. sp. FACHB 1194</i>       | 99  | Nitrogenase MoFe<br>cofactor<br>biosynthesis<br>protein NifE           | MBD2144569.1 |
|             | 455 | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein | <i>S. torques reginae</i> ITEP-024    | <i>S. sp. LEGE 00249</i>       | 99  | Nitrogenase MoFe<br>cofactor<br>biosynthesis<br>protein NifE           | MBC5796919.1 |
|             | 455 | Nitrogenase MoFe<br>cofactor biosynthesis<br>protein | <i>S. torques reginae</i> ITEP-024    | <i>S. sp SIO1G1</i>            | 99  | Nitrogenase MoFe<br>cofactor<br>biosynthesis<br>protein NifE           | NET02491.1   |
| <b>NifB</b> | 478 | Nitrogenase cofactor<br>biosynthesis protein         | <i>S. torques reginae</i><br>ITEP-024 | <i>S. sp. LEGE 08334</i>       | 100 | Nitrogenase<br>cofactor<br>biosynthesis                                | MBE9059234.1 |

|             |     |   |  |                          |     |  |              |
|-------------|-----|---|--|--------------------------|-----|--|--------------|
|             |     |   |  |                          |     | protein NifB   |              |
|             | 478 | Nitrogenase cofactor biosynthesis protein                     | <i>S. torquees reginae</i><br>ITEP-024 | <i>S. sp. LEGE 00249</i> | 99  | Nitrogenase cofactor biosynthesis protein NifB                 | MBC5795050.1 |
|             | 478 | Nitrogenase cofactor biosynthesis protein                     | <i>S. torquees reginae</i><br>ITEP-024 | <i>S. sp SIO1G1</i>      | 94  | Nitrogenase cofactor biosynthesis protein NifB                 | NET02524.1   |
| <b>NifW</b> | 105 | Nitrogen fixation protein                                     | <i>S. torquees reginae</i><br>ITEP-024 | <i>S. sp. LEGE 00249</i> | 98  | Nitrogenase stabilizing/protective protein                     | MBC5796914.1 |
|             | 105 | Nitrogen fixation protein                                     | <i>S. torquees reginae</i><br>ITEP-024 | <i>S. sp SIO1G1</i>      | 78  | Nitrogenase stabilizing/protective protein                     | NET00478.1   |
| <b>NifX</b> | 139 | Nitrogen fixation protein                                     | <i>S. sp SIO1G1</i>                    | <i>S. sp. LEGE 00249</i> | 88  | Dinitrogenase iron-molybdenum cofactor biosynthesis            | MBC5796917.1 |
| <b>NifN</b> | 442 | Nitrogenase molybdenum-ion cofactor biosynthesis protein NifN | <i>S. torquees reginae</i><br>ITEP-024 | <i>S. sp. LEGE 08334</i> | 100 | Nitrogenase molybdenum-cofactor biosynthesis protein NifN      | MBE9054903.1 |
|             | 442 | Nitrogenase molybdenum-ion cofactor biosynthesis protein NifN | <i>S. torquees reginae</i><br>ITEP-024 | <i>S. sp. LEGE 00249</i> | 98  | Nitrogenase molybdenum-cofactor biosynthesis protein NifN      | MBC5796918.1 |
|             | 442 | Nitrogenase molybdenum-ion cofactor biosynthesis protein NifN | <i>S. torquees reginae</i><br>ITEP-024 | <i>S. FACHB-1194</i>     | 98  | Nitrogenase iron-molybdenum cofactor biosynthesis protein NifN | MBD2144570.1 |
|             | 442 | Nitrogenase molybdenum-ion cofactor biosynthesis protein NifN | <i>S. torquees reginae</i><br>ITEP-024 | <i>S. sp SIO1G1</i>      | 91  | Nitrogenase iron-molybdenum cofactor biosynthesis              | NET02490.1   |

|             |     |   |                                     |                                |     |  |              |
|-------------|-----|---|-------------------------------------|--------------------------------|-----|--|--------------|
|             |     |   |                                     |                                |     | protein NifN                                     |              |
| <b>NifK</b> | 511 | Nitrogenase molybdenum-ion protein beta chain         | <i>S. torquees reginae</i> ITEP-024 | <i>S. sp</i> LEGE 08334        | 100 | Nitrogenase molybdenum-iron protein subunit beta | MBE9054900.1 |
| <b>NifU</b> | 300 | Fe-S cluster assembly protein NifU                    | <i>S. torquees reginae</i> ITEP-024 | <i>S. sp. LEGE 08334</i>       | 100 | Fe-S cluster assembly protein NifU               | MBE9059231.1 |
|             | 308 | Fe-S cluster assembly protein NifU                    | <i>S. torquees reginae</i> ITEP-024 | <i>S. sp SIO1G1</i>            | 92  | Fe-S cluster assembly protein NifU               | NET02521.1   |
| <b>NifS</b> | 403 | Nitrogenase metalloclusters biosynthesis protein NifS | <i>S. torquees reginae</i> ITEP-024 | <i>S. sp. LEGE 08334</i>       | 99  | Cysteine desulfurase NifS                        | MBE9059232.1 |
|             | 403 | Nitrogenase metalloclusters biosynthesis protein NifS | <i>S. torquees reginae</i> ITEP-024 | <i>S. kisseleviana NIES-73</i> | 99  | Cysteine desulfurase NifS                        | BAZ80945.1   |
|             | 403 | Nitrogenase metalloclusters biosynthesis protein NifS | <i>S. torquees reginae</i> ITEP-024 | <i>S. sp. LEGE 00249</i>       | 98  | Cysteine desulfurase NifS                        | MBC5795052.1 |
|             | 403 | Nitrogenase metalloclusters biosynthesis protein NifS | <i>S. torquees reginae</i> ITEP-024 | <i>S. sp SIO1G1</i>            | 94  | Cysteine desulfurase NifS                        | NET02522.1   |



**Supplementary Table S13. Comparison of genomic features and subsystem annotation of the strains *C. raciborskii* located in the Americas, Non Americas and Sphaerospermopsis.**

| <b>Genomic statistics</b> | N° Contigs | Total Size (bp) | Longest contig length (bp) | Shortest contig length (bp) | Mean contigs size (bp) | Median contigs size (bp) | GC content % | N50     | <b>Subsystem statistics - SEED</b> | N° of subsystems | N° of coding sequences | Coding sequences in subsystems | Coding sequences not in subsystems |
|---------------------------|------------|-----------------|----------------------------|-----------------------------|------------------------|--------------------------|--------------|---------|------------------------------------|------------------|------------------------|--------------------------------|------------------------------------|
| <i>R. brookii</i> D9      | 47         | 3,186,511       | 526,79                     | 3,5                         | 67,8                   | 29,59                    | 40.06        | 127,75  |                                    | 347              | 3,12                   | 1,307 or 41%                   | 1,813 or 58%                       |
| CENA 302                  | 58         | 3,476,418       | 324,204,0                  | 580                         | 59,938                 | 18,63                    | 40.08        | 162,402 |                                    | 350              | 3,392                  | 1,363 or 41%                   | 2,029 or 59%                       |
| CENA 303                  | 77         | 3,398,605       | 403,228                    | 511                         | 43,581                 | 7,321                    | 40.26        | 135,818 |                                    | 354              | 3,360                  | 1,342 or 40%                   | 2,018 or 60%                       |
| MVCC14                    | 99         | 3,594,524       | 299,478                    | 1,035                       | 36,308                 | 4,917                    | 40.08        | 150,437 |                                    | 351              | 3,53                   | 1,384 or 40%                   | 2,149 or 60%                       |
| KL1                       | 100        | 3,657,391       | 268160                     | 2084                        | 36573.9                | 8955                     | 40.1         | 164738  |                                    | 248              | 3512                   | 752 or 22%                     | 2760 or 78%                        |
| PAMP2012                  | 134        | 3,234,024       | 187192                     | 1501                        | 241.34.5               | 11409                    | 39.9         | 21339   |                                    | 250              | 3210                   | 747 or 24%                     | 2463 or 76%                        |
| MVCC19                    | 155        | 3,515,201       | 259706                     | 221                         | 22678.7                | 2284                     | 40.1         | 116886  |                                    | 249              | 3555                   | 766 or 22%                     | 2789 or 78%                        |
| CYRF                      | 166        | 4,150,048       | 37600889                   | 637                         | 25000.3                | 1704                     | 40.2         | 44019   |                                    | 253              | 5352                   | 1071 or 21%                    | 4281 or 79%                        |
| CYLP                      | 188        | 3,972,140       | 3476922                    | 496                         | 21128.4                | 1530                     | 40.1         | 53585   |                                    | 247              | 5508                   | 1149 or 21%                    | 4359 or                            |

|          |     |           |           |         |           |         |       |           |  |     |       |              |              |
|----------|-----|-----------|-----------|---------|-----------|---------|-------|-----------|--|-----|-------|--------------|--------------|
|          |     |           |           |         |           |         |       |           |  |     |       |              | 79%          |
| ITEP-A1  | 195 | 3,605,836 | 266,816   | 1,008   | 18,491    | 2,823   | 40.15 | 91,008    |  | 349 | 3,391 | 1,364 or 41% | 2,027 or 59% |
| KLL07    | 1   | 3,808,183 | 3808183   | 3808183 | 3808183   | 3808183 | 40.2  | -----     |  | 257 | 3675  | 757 or 21%   | 2918 or 79%  |
| N8       | 1   | 3,857,174 | 3857174   | 3857174 | 3857174   | 3857174 | 40.1  | -----     |  | 257 | 3760  | 778 or 21%   | 2982 or 79%  |
| CR2010   | 1   | 3,757,845 | 3757845   | 3757845 | 3757845   | 3757845 | 40.2  | -----     |  | 257 | 3744  | 761 or 21%   | 2983 or 79%  |
| GIHEG1   | 2   | 4,057,210 | 3910736   | 146474  | 2028605   | 3910736 | 40.1  | 3910736   |  | 257 | 3943  | 781 or 20%   | 3162 or 80%  |
| DSH      | 2   | 3,971,566 | 3350285   | 621281  | 1985783.0 | 3350285 | 40.1  | 621281    |  | 256 | 4163  | 776 or 19%   | 3387 or 81%  |
| GIHE2018 | 3   | 3,629,822 | 1305142   | 1046095 | 120994.07 | 1278585 | 40.2  | 1305142   |  | 250 | 3723  | 735 or 20%   | 2988 or 80%  |
| CS505    | 6   | 4,159,26  | 4,011,384 | 2,52    | 693,21    | 14,64   | 40.28 | 4,011,384 |  | 360 | 4,07  | 1,590 or 40% | 2,483 or 60% |
| CR12     | 136 | 3,723,955 | 279,631   | 1,058   | 27,383    | 6,201   | 40.03 | 79,912    |  | 356 | 3,475 | 1,366 or 40% | 2,109 or 60% |
| CS508    | 162 | 3,556,598 | 193,915   | 1,027   | 21,954    | 5,661   | 40.15 | 62,25     |  | 346 | 3,346 | 1,355 or 41% | 1,991 or 59% |

|                            |     |           |         |         |           |         |      |        |  |     |      |            |             |
|----------------------------|-----|-----------|---------|---------|-----------|---------|------|--------|--|-----|------|------------|-------------|
| 1523720                    | 173 | 3,199,554 | 123795  | 2501    | 18494.5   | 11735   | 40.0 | 31186  |  | 250 | 3123 | 748 or 24% | 2375 or 76% |
| CHAB3438                   | 189 | 3,544,092 | 293012  | 200     | 18751.8   | 1107    | 40.5 | 36962  |  | 255 | 3508 | 778 or 23% | 2730 or 77% |
| S. torques reginae ITP-024 | 1   | 5,254,542 | 5254542 | 5254542 | 5254542   | 5254542 | 37.4 | -----  |  | 269 | 5451 | 871 or 16% | 4580 or 84% |
| S. kisseleviana NIES 73    | 2   | 5,609,991 | 5354839 | 255152  | 2804995.5 | 5354839 | 37.6 | 68245  |  | 267 | 5761 | 894 or 16% | 4867 or 84% |
| S. sp. SIO1G1              | 38  | 5,187,854 | 1244078 | 1041    | 136522.5  | 28839   | 36.3 | 397515 |  | 259 | 4883 | 808 or 17% | 4075 or 83% |
| S. sp. FACHB1194           | 108 | 5,405,915 | 262430  | 1482    | 50054.8   | 31150   | 37.6 | 86753  |  | 265 | 5896 | 889 or 16% | 5007 or 84% |
| S. sp. LEGE 08334          | 119 | 5,548,733 | 229001  | 1002    | 46628     | 26281   | 37.2 | 97259  |  | 267 | 5894 | 893 or 16% | 5001 or 84% |
| S. sp. LEGE 00249          | 177 | 5,298,876 | 225832  | 1027    | 29937.2   | 17777   | 37.5 | 56781  |  | 261 | 5736 | 889 or 16% | 4847 or 84% |