LEGE Culture Collection and its cyanobacterial diversity: valuing an increasingly important biological resource

Ramos V. [1,2], Morais J. [1], Vasconcelos V. [1,2]

[1] BBE, CIIMAR, Porto, Portugal [2] Faculty of Sciences of the University of Porto, Porto, Portugal

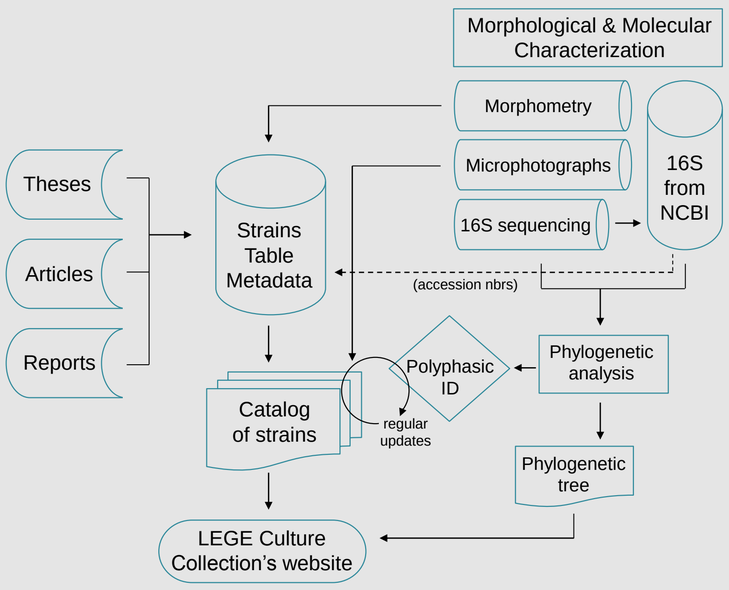
Introduction

The Blue Biotechnology and Ecotoxicology group (CIIMAR, University of Porto) has undertaken a process of organizing its cyanobacterial isolates by creating a culture collection. This led to the recently membership of LEGE CC in the World Federation for Culture Collections (WFCC). The collection is also part of the Research Infrastructure EMBRC.PT. At present it comprises more than 350 different cyanobacterial strains, several of them unique among the (phylogenetic) diversity of the group. These isolates have been obtained since 1991, from samples collected in different environments and locations mainly in Portugal (including Madeira and Azores Islands), but also elsewhere (e.g. Australia, Brazil, Colombia, Morocco, Mexico). As in other collections worldwide, LEGE CC seeks to provide starter cultures for a diverse range of aims (R&D, science education & dissemination). In fact, this has being done with local, national and international entities, under different types of partnerships. Several studies, most of them from BBE group's own research, have shown the potential or the effective capacity of different LEGE strains to produce a myriad of chemical compounds, including toxins or newly discovered bioactive molecules. Soon, a number of strains (at least three) will have their genomes sequenced and annotated. Others studies revealed that some strains are phylogenetically distinct, and thus taxonomic challenging.

In this work, while presenting the Culture Collection, the main findings from a survey of published and unpublished data available on the total LEGE strains are given. The methodological approach followed for this meta-data survey is summarized and presented in a systematic manner. Preliminary results are presented, linking the phylogenetic placement of LEGE strains in the "cyanobacterial Tree of Life" (based on 16S rRNA gene sequences), along with relevant information retrieved from the data compilation process (i.e. morphological features and identification, geographic and ecological origin, ecophysiological data, toxicity and bioactivity (bioassays, molecular, analytical) analyses performed; production/type of secondary metabolite, publications, year of publication, etc.). It is believed that connecting the phylogeny of each LEGE strain to the aforementioned data creates awareness and capture interest of the scientific community and of members of the general public, such as pharmaceutical and biotechnological enterprises. Full information about each LEGE strain will be made available through the online version (currently under construction) of the collection.

Material & Methods

➢ Workflow diagram for the data survey and/or data collection process, and analysis.



Results

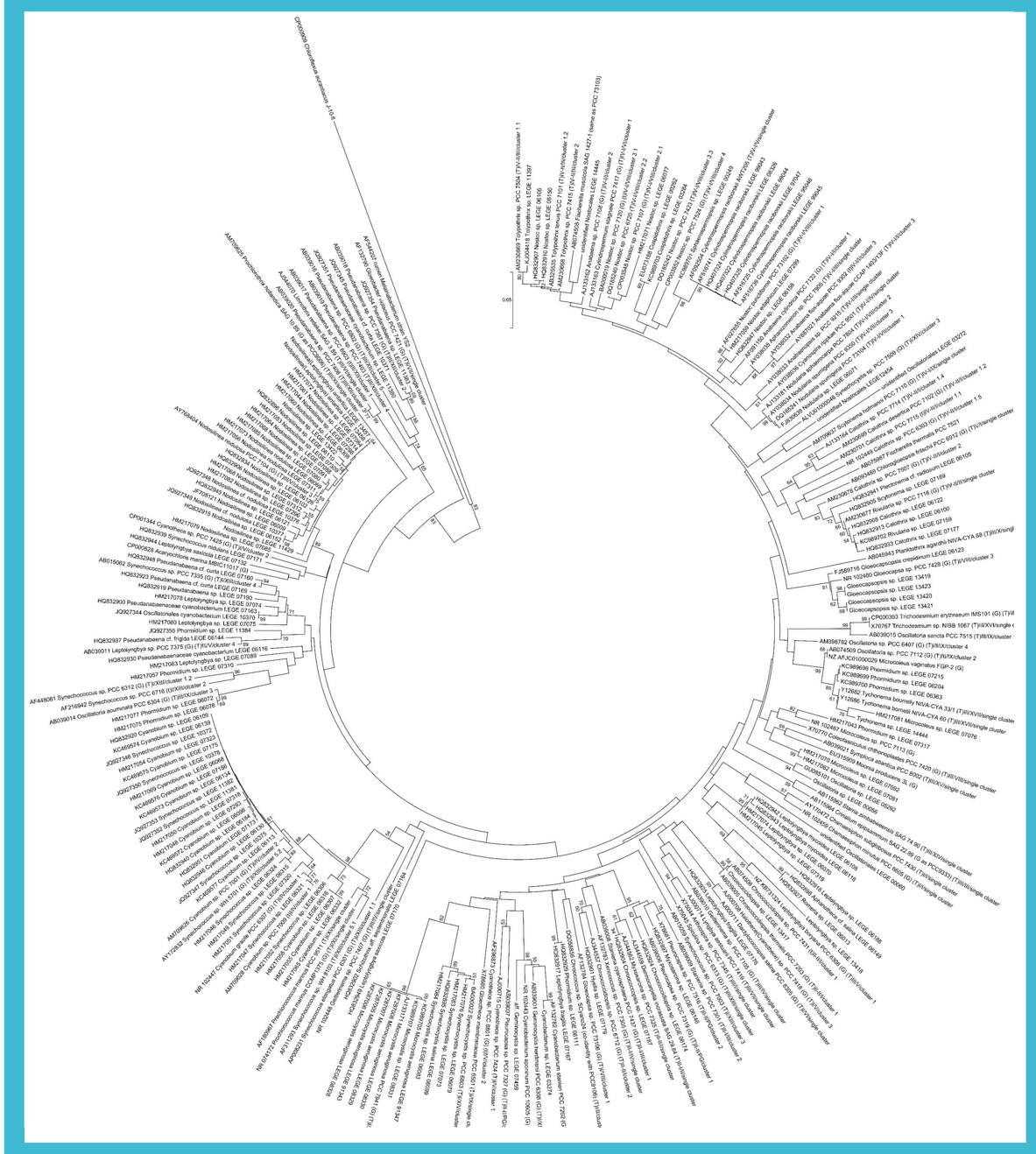
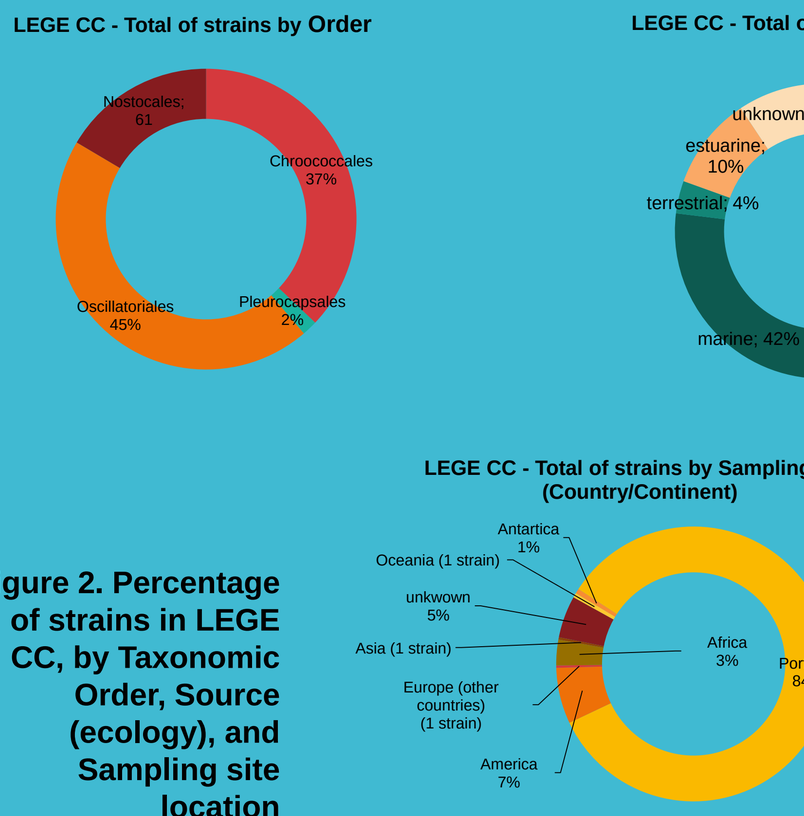


Figure 1. Preliminary phylogenetic tree of partial 16S rRNA gene sequences from LEGE CC strains and their placement in the cyanobacterial "Tree of Life".

The analysis involved 248 nucleotide sequences, from which 142 belong to LEGE strains (c.a. 40% of the collection).
Selected strains include Reference strains (T) from the Bergey's [1] classification scheme (key path is given at the end of taxon label) and/or strains with their genome available (G). GenBank accession numbers (if available) are given for each strain.
The evolutionary history was inferred using the Neighbor-Joining method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method (the model that best fit the dataset) and are in the units of the number of base substitutions per site. The rate variation among sites was modeled with a gamma distribution (shape parameter = 0.63). Fewer than 10% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 359 positions in the final dataset. *Chloroflexus aurantiacus* J-10-fl and the rumen Melainabacterium strain YS2 were used as outgroup.
Evolutionary analyses were conducted in MEGA6 [2].



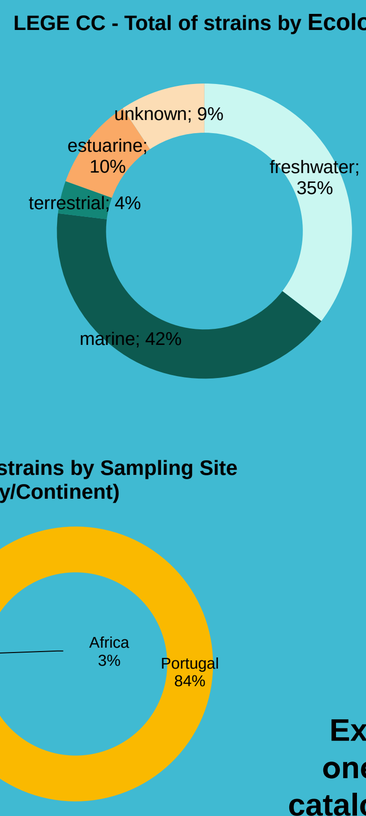


Figure 2. Percentage of strains in LEGE CC, by Taxonomic Order, Source (ecology), and Sampling site location

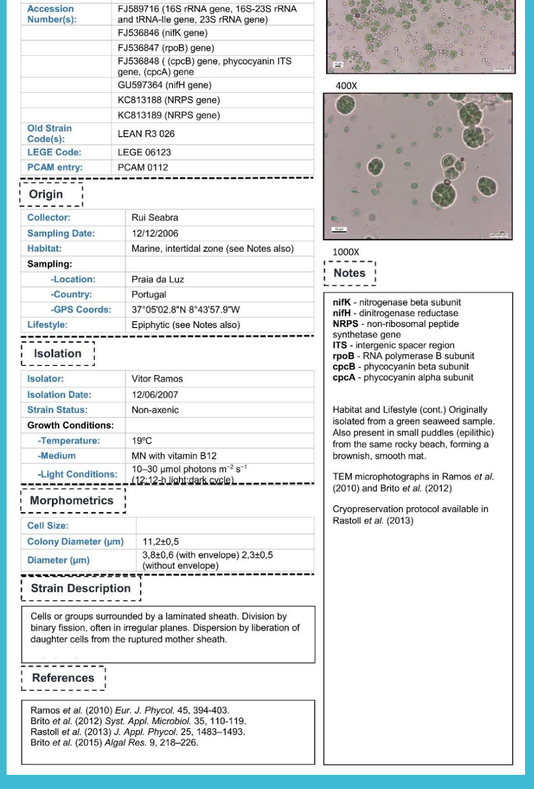


Figure 3. Example of one strain's catalog sheet.

Conclusion

LEGE strains (see Fig. 2 for some statistics) are phylogenetically diverse and widespread across the tree (Fig. 1). It is believed that linking their phylogenetic placement along with relevant data being retrieved from the survey (see M&M and Fig. 3) will capture interest from the scientific community and/or bioenterprises. Data gathered from this process (e.g. Fig. 3) will be made available in the website of LEGE CC (www.ciimar.up.pt/legeculturecollection), aiming to give awareness and visibility to the Collection.

Future work

- To finish the sequencing of 16S rRNA genes from all LEGE CC strains.
- To identify the cyanobacterial taxa according to the most recent (polyphasic) taxonomy (e.g.[3])
- To populate the site's database.

References

[1] Boone, D.R., Castenholz, R.W., Garrity, G.M. (2001). Bergey's Manual of Systematic Bacteriology – The Archaea and the Deeply Branching and Phototrophic Bacteria, 1, 2nd ed., Springer, New York, USA, pp. 473–599.

[2] Tamura K., Stecher G., Peterson D., Filipski A., and Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725-2729.

[3] Guiry, M.D. & Guiry, G.M. 2015. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. http://www.algaebase.org

Acknowledgments & Funding sources

This work was partially funded by the FCT project **PTDC/MAR-BIO/2818/2012**; by the Project MARBIOTECH **(reference NORTE-07-0124-FEDER-000047)** within the SR&TD Integrated Program MARVALOR - Building research and innovation capacity for improved management and valorization of marine resources, supported by the Programa Operacional Regional do Norte **(ON.2 – O Novo Norte)**; by the European Regional Development Fund; and also by NOVOMAR **(reference 0687-NOVOMAR-1P)**, supported by the European Regional Development Fund.