

Type of the Paper: Article

# Metagenome-assembled genome of *Cyanocohniella* sp. LLY from the cyanosphere of Llayta, an edible Andean cyanobacterial macrocolony

Claudia Vilo<sup>1</sup>, Qunfeng Dong<sup>2</sup>, Alexandra Galetovic<sup>1</sup>, Benito Gómez-Silva<sup>1\*</sup>

<sup>1</sup> Laboratory of Biochemistry, Biomedical Department, Health Sciences Faculty and Centre for Biotechnology and Bioengineering (CeBiB), Universidad de Antofagasta, Antofagasta, Chile; [claudiavilo@gmail.com](mailto:claudiavilo@gmail.com); [alexandra.galetovic@uantof.cl](mailto:alexandra.galetovic@uantof.cl); [benito.gomez@uantof.cl](mailto:benito.gomez@uantof.cl)

<sup>2</sup> Center for Biomedical Informatics, Department of Medicine, Stritch School of Medicine, Loyola University of Chicago, Maywood, Illinois, USA Affiliation 2; [qdong@luc.edu](mailto:qdong@luc.edu)

\* Correspondence: [benito.gomez@uantof.cl](mailto:benito.gomez@uantof.cl)

**Abstract:** Cyanobacterial macrocolonies known as Llayta are found at Andean wetlands and consumed since pre-Columbian times in South America. Macrocolonies of filamentous cyanobacteria are niches for colonization by other microorganisms; however, the microbiome of edible Llayta has not been explored. Based on a culture-independent approach, we report the presence, identification and metagenomic genome reconstruction of *Cyanocohniella* sp. LLY associated to Llayta trichomes. The assembled genome of strain LLY is now available for further inquiries, and may be instrumental for taxonomic advances on this genus. All known members of the *Cyanocohniella* genus have been isolated from salty European habitats. A biogeographic gap for the *Cyanocohniella* genus is partially filled by the existence of strain LLY at Andes Mountains wetlands in South America as a new habitat. This is the first genome available for members of this genus. Genes involved in primary and secondary metabolism are described providing new insights on the putative metabolic capabilities of *Cyanocohniella* sp. LLY. The reconstructed genome of strain LLY is now available and instrumental for further inquiries and taxonomic advances on the genus *Cyanocohniella*.

**Keywords:** cyanosphere, cyanobacteria, *Cyanocohniella*, Llayta, macrocolonies, metagenomics, metagenomic-assembled genome, microbiome

## 1. Introduction

Cyanobacterial and microalgal biomasses have been consumed for centuries and *Dunaliella*, *Chlorella* and *Arthrospira* (in Africa and North America), and *Nostoc* (in Asia and South America) are documented sources of essential amino acids, vitamins, polyunsaturated fatty acids, carotenoids, phycobiliproteins, and secondary metabolites with diverse biotechnological applications [1-16].

The Atacama Desert is considered the driest and oldest desert on Earth, with climatic, oceanic, latitudinal, and geomorphological factors that account for its existence [17-19]. Precipitations at the Andes Range are at least 40-50 fold higher than those observed at the dry core of the Atacama Desert [19]. Indigenous cultures have historically acknowledged the Andean biome as source of forage, food, and ethnomedicine [16,20,21]. The consumption of cyanobacterial macrocolonies made of filamentous diazotrophic species, known by the vernacular name of Llayta, is a centenary Andean alimentary practice that can be traced back to pre-Columbian times and it is an example of a neglected natural food resource [8-10,14,16]. Llayta macrocolonies are collected at wetlands at the Andes highlands

over 3,000 m above sea level; they are sun-dried and sold for human consumption as a dry, dark green, leaf-like biomass in food markets in Arica and Iquique, in northern Chile, and Tacna, in southern Peru [16,22]. Ethnographic, chemical composition, and cyanotoxic studies in addition to the absence of negative epidemiological evidence, support the notion that Llayta would be a safe, natural food ingredient [16,22,23].

The natural life cycle of primary producers may involve mutualistic, antagonistic, or commensal interactions with cyanophages, fungi, and heterotrophic bacteria, and positive or negative outcomes have been previously discussed in *Nostoc* macrocolonies [24-27]. However, identification of members of the bacterial community associated with edible Llayta macrocolonies have been a pending issue. Here we report the metagenome-assembled genome (MAG) of a previously undetected cyanobacterium strain associated to the cyanosphere of isolated Llayta trichomes. We have identified it as a member of the genus *Cyanocohniella*, referred here as strain *Cyanocohniella* sp. LLY, [expanding the presence of the genus \*Cyanocohniella\* to a new habitat, the Andes Mountains wetlands in South America](#). Bioinformatic analyses of MAG allowed the identification of this new strain, the annotation of functional genes, and insights on its putative metabolic capabilities. The reconstructed genome of strain LLY is now available and instrumental for further inquiries and taxonomic advances on the *Cyanocohniella* genus.

## 2. Materials and Methods

*Isolation Llayta filaments.* Dried Llayta colonies were obtained from the major farmers' market in Arica, Chile, during 2015 and maintained in the original plastic bag until used. Dry colonies were suspended in sterile deionized water for rehydration for 24 h. To isolate filaments, aliquots (5 mL) were transferred to liquid, nitrogen-free Arnon mineral medium [28] and cultured at 30°C under white fluorescent light ( $\mu\text{E } 180 \text{ m}^{-2} \text{ s}^{-1}$ ), continuous agitation (200 rpm), and aeration enriched in 1% v/v  $\text{CO}_2$  [14,22]. Culture aliquots were spread onto agar plates, observed under a stereo microscope, and isolated filaments were picked and transferred to fresh medium for growth as above. Aliquots (10-15 mL) of cultured Llayta filaments were collected at the exponential growth phase (day 21st), washed with fresh medium, the filaments were recovered as a pellet by centrifugation at 4,000 x g, for 5 min, at room temperature, and used to extract total genomic DNA.

*DNA extraction.* Total genomic DNA was extracted from the filament pellets with Ultra Clean Microbial DNA isolation kit (MoBio), following the manufacturer's instructions. DNA quality was evaluated by electrophoresis in 0.8 % agarose gel and quantified photometrically at 260 nm.

*DNA sequencing and analysis.* The metagenome of Llayta filaments was sequenced via MiSeq sequencing technology using shotgun paired-end libraries, with an average insert size of 250 bp. Reads had an average length of 300 bp, with good quality scores, as evaluated by the FastQC program (version 0.10.0). The sequencing produced a total of 17,137,246 reads. Metagenomic assembly was done using MEGAHIT assembler v.1.2.9 [29], and binning was conducted using the PATRIC web server [30]. Secondary metabolites were searched with PRISM [31] and AntiSMASH [32] software. Sequencing reads are available at the Sequence Read Archive (SRA) with accession number SRR17916224. The Whole Genome Shotgun project have been deposited at DDBJ/ENA/GenBank under the accession JAKOMP000000000.

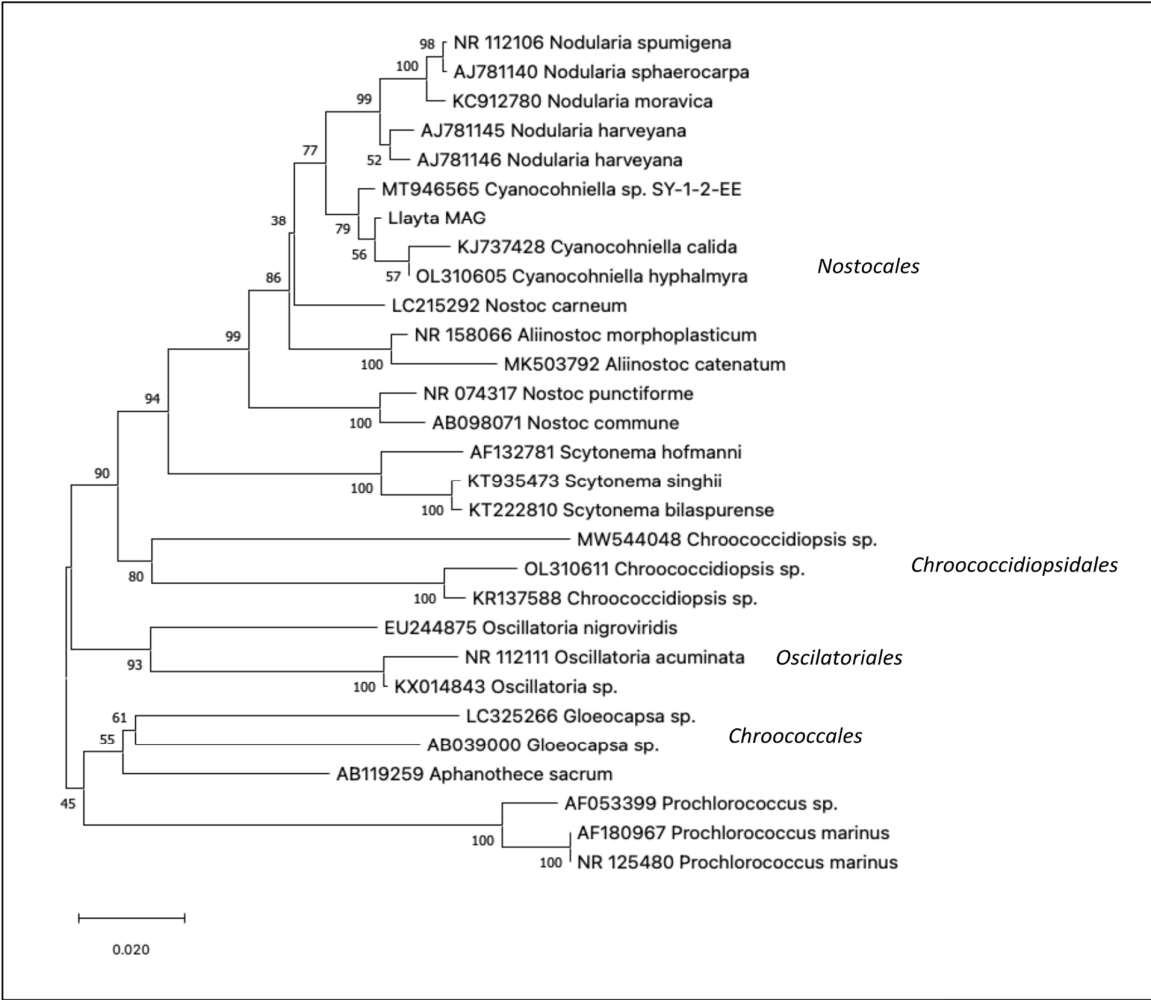
## 3. Results

### 3.1. Genome assembly and phylogeny of *Cyanocohniella* sp. LLY

The metagenome analysis of Llayta trichomes reads revealed predominant bacteria belonging to the phylum *Proteobacteria* (82%). Assembly and binning strategy resulted in four reconstructed bacterial genomes resembling the four most abundant bacterial genera, whose features will be published in the near future.

Most importantly, the metagenome analysis revealed a 16% of reads corresponding to Cyanobacteria phylum, from which one cyanobacterium genome was reconstructed. The cyanobacterium was assigned to the *Cyanocohniella* genera, which is consistent with the dominance of Nostocales reads (40%) among cyanobacteria phylum. The *Cyanocohniella* sp. LLY genome was reconstructed using 244 contigs with an N50 of 50,369, which rendered a total genome size of 6,781,030 bp with a GC content of 41.2%. Binning strategy accounted for a complete reconstructed genome with minimal contamination (0.7%)

Phylogenetic analyses of 16S rRNA gene confirmed that the MAG from Llayta filaments belonged to the Nostocales order. Moreover, a clear placement to the *Cyanocohniella* clade was observed (Fig. 1). Alignments of the Llayta MAG's 16S rRNA to the NCBI nr Database rendered a 99% identity to *Cyanocohniella* sp. SY-1-2-Y (NCBI accession number: MT946558.1)



**Figure 1.** Phylogenetic analysis by Maximum Likelihood method based on gene sequences of 16S rRNA. The horizontal bar at the figure's base represents 0.02 substitutions per nucleotide site. The percentage of trees in which the associated taxa clustered together is shown next to the branches, using a bootstrap of 1000. Evolutionary analysis was conducted in MEGA X.

3.2. Functional capabilities of *Cyanocohniella* sp. LLY

The reconstructed genome of *Cyanocohniella* sp. LLY was analyzed to identify sequences of functional genes and gene clusters coding for enzymes involved in primary

and secondary metabolism. As expected, the cyanobacterium genomic capabilities included sequences related to the production of vitamins, prosthetic groups, and pigments (5.5%); DNA metabolism (5.3%); fatty acids, lipids, and isoprenoids (2.9%); dormancy and sporulation (0.14%); multiple gene copies associated to Photosystems I and II proteins; genes and pathways associated to DNA repair metabolism, and genes for cell division and cell cycle processes. In addition, *nif* gene cluster for nitrogen fixation were also identified in LLY MAG. Gene sequences involved in chemotaxis and motility were absent.

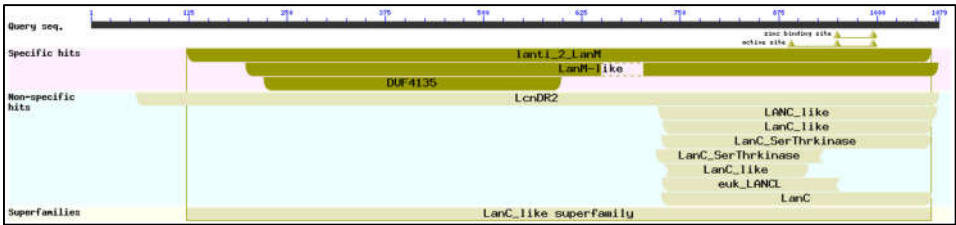
Genomic analysis in silico demonstrated the presence of conserved lycopene cyclase *CruA/CruP* genes, the complete *mysABCD* and *scyABCDEF* (and accompanied regulator and kinase sensor) gene clusters involved in carotenoids biosynthesis, mycosporine-like amino acids (MAAs), and scytonemin biosynthesis (Table 1). These putative capabilities are intimately related to protection against photooxidative damages triggered by high-UV solar insolation prevailing at the Andes plateau wetlands [33].

**Table 1.** *Cyanocohniella* sp. LLY open reading frames annotations and best Blast hit for *MysABC* and *ScyABCDEF* cluster genes

ORF	Description based on Subsystem annotations	Accession	Identity (%)
Orf 3147	Demethyl 4-deoxygadusol synthase MysA	WP_069074324.1	93
Orf 3148	O-methyltransferase MysB	BBC27542.1	85
Orf 3149	ATP-grasp ligase forming mycosporine-glycine, MysC	BBC27543.1	82
Orf 4592	Scytonemin biosynthesis protein ScyA	WP_206262883.1	83
Orf 4593	Tryptophan dehydrogenase ScyB	WP_086764771.1	84

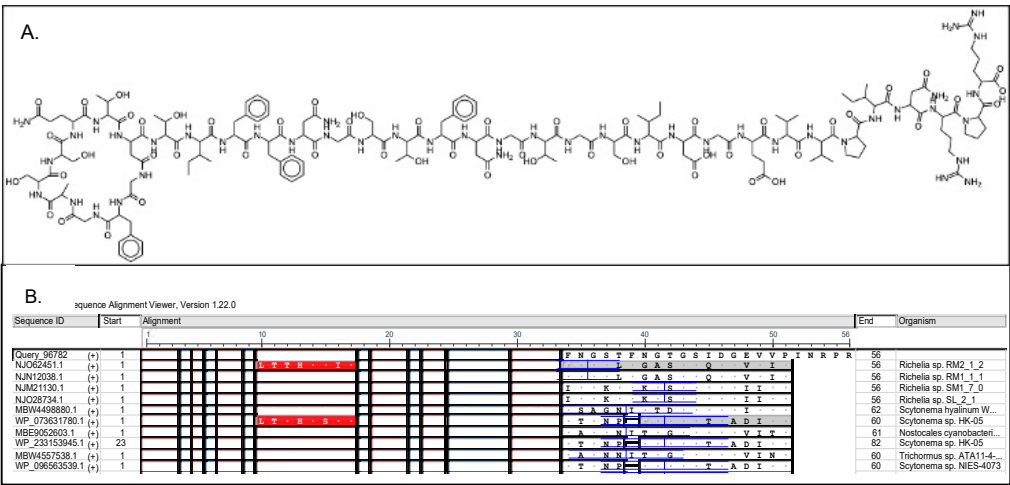
Bioinformatic analyses in the reconstructed genome of *Cyanocohniella* sp. LLY rendered two copies of gene *cphA1* coding for a putative non-ribosomal cyanophycin synthetase. Sequence alignment of *Cyanocohniella* sp. LLY cyanophycin synthetase to Nostocales CphA1 (NCBI accession number: MBE9052550.1 and MBE9053887.1) showed an 85% identity. The reconstructed genome of *Cyanocohniella* sp. LLY also contained three copies of the gene *cphB* coding for a putative cyanophycin-degrading cyanophycinase, which showed an 83% identity with Cyanobacteria cyanophycinase proteins (NCBI accession number: AHJ26983.1 and WP\_096726859.1).

*Cyanocohniella* sp. LLY MAG contained sequences of the gene *LanM*, involved in the biosynthesis of a family of ribosomally-synthesized and post-translationally modified peptides. This sequence and the CCG motif for Zn binding sites at the C-terminal showed a 70% identity after alignments with the *LanM* sequence of cyanobacteria bacterium UBA11372 (NCBI accession number HAX78725.1). Fig. 2 shows the domains for dehydratase and cyclization activities and three cysteine residues associated with zinc-binding sites, one cysteine residue alone, and the CCG conserved motif from the Llayta cyanobacterium.



**Figure 2.** Domain characterization of lantibiotics biosynthesis enzyme using the NCBI Conserved Domain Database.

The reconstructed genome of *Cyanocohniella* sp. LLY contains sequences for all genes coding for lasso peptide biosynthesis: a lasso peptide precursor lasso P (49% identity to *Nostocales*, cyanobacterium accession number MBE9052603.1), a lasso C isopeptide bond-forming cyclase (74% identity to *Nostocales*, cyanobacterium MBE9052604.1), a lasso peptide biosynthesis B2 protein (73% identity to *Nostocales*, cyanobacterium accession number MBE9052605.1), and lasso peptide biosynthesis B1 PqqD family protein (86% identity to *Nostocales*, cyanobacterium accession number MBE9052606.1). Fig 3 depicts the organization for the lasso peptide gene cluster in the genome of the Llayta cyanobacterium, resembling that from Actinobacteria, and includes the putative primary structure of the cleaved lasso peptide from *Cyanocohniella* sp. LLY, the location of the lasso structure with a glycine residue at the peptide N-terminal and bulky residues at its C-terminal, and a comparison of conserved sequences among several microbial genera.



**Figure 3.** Characterization of the lasso peptides from the reconstructed genome of the *Cyanocohniella* sp. LLY cyanobacterium. (A) Primary chemical structure of the cleaved Lasso P, made of 36 residues. (B) Protein sequence alignment of Lasso P peptide against NCBI non-redundant database.

The genome of *Cyanocohniella* sp. LLY also has putative biosynthetic pathways for capreomycin peptides from non-ribosomal peptide synthases (NRPS) and microviridin peptides from ribosomally synthesized and post-translationally modified peptides (RiPPs). Capreomycin synthase gene from *Cyanocohniella* sp. LLY showed an 82% identity with capreomycin synthase gene (sequence ID: WP\_104904844.1) of a *Nostoc* sp. cyanobiont of *Lobaria pulmonaria* and 42.23% identity with capreomycin hydroxylase (sequence ID: WP\_051702360) from *Streptomyces vinaceus*. Also, microviridin precursor peptide sequence in the *Cyanocohniella* sp. LLY genome showed an 83% identity with the microviridin family tricyclic proteinase inhibitor gene of *Nostoc* sp. (sequence ID: WP\_224090535.1). The capreomycin and microviridin biosynthetic pathway included arginine hydrolase and condensation, adenylation, thiolation, epimerization, and reductase enzymes.

CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR-associated) operons are defensive systems found in archaea and bacteria [34-40] and were found in the *Cyanocohniella* sp. Llayta genome. One of them included Cmr2-Cmr3-Cmr4-Cmr5-Cmr6 and Cas1-Cas2 operons flanked by CRISPR arrays; the second system showed

the Csc3-Csc2-Csc1-Cas6-Cas4-Cas1-Cas2 genes and CRISPR array, and the third comprised Cmr2-Cmr3-Cmr4-Cmr5-Cmr6 genes flanked by CRISPR array (Table 2). Therefore, and according to the observed organization, the CRISPR-Cas system operons in *Cyanocohniella* sp. LLY, would indeed confer immunity defense against mobile elements.

**Table 2.** *Cyanocohniella* sp. LLY open reading frame annotations and best Blast hit for CRISPR-CAS system operons.

ORF	Description based on Subsystem annotations	Accession	Identity (%)
Orf 2275	CRISPR-associated RAMP Cmr2	WP_179048547.1	92
Orf 2276	CRISPR-associated RAMP Cmr3	WP_102220820.1	95
Orf 2277	CRISPR-associated RAMP Cmr4	WP_179048549.1	94
Orf 2278	CRISPR-associated RAMP Cmr5	WP_102220822.1	95
Orf 2279	CRISPR-associated RAMP Cmr6	WP_179048551.1	78
Orf 2288	CRISPR-associated protein Cas1	WP_218653184.1	94
Orf 2289	CRISPR-associated protein Cas2	WP_179048556.1	94
Orf 3613	CRISPR-associated negative autoregulator Cas7/Cst2	WP_194144297.1	94
Orf 3614	CRISPR-associated protein Cas5	MBW4428053.1	96
Orf 5125	CRISPR-associated protein Cas2	WP_190898341.1	95
Orf 5126	CRISPR-associated protein Cas1	MBN3899475.1	93
Orf 5227	CRISPR-associated RecB family exonuclease Cas4	MBW4675378.1	91
Orf 5128	CRISPR-associated endoribonuclease Cas6	WP_096682797.1	94
Orf 5137	CRISPR-associated helicase Cas3	WP_179075640.1	97
Orf 5452	CRISPR-associated RAMP Cmr2	MBD2365142.1	92
Orf 5453	CRISPR-associated RAMP Cmr3	WP_190709980.1	95
Orf 5454	CRISPR-associated RAMP Cmr4	WP_190709978.1	93
Orf 5455	CRISPR-associated RAMP Cmr5	WP_190709976.1	95
Orf 5456	CRISPR-associated RAMP Cmr6	WP_190709975.1	91

4. Discussion

Cyanobacterial biomasses are well-documented sources of metabolites for biotechnological applications and hopefully for humankind. Some cyanobacteria species are an inherited foodstuff from ancient cultures and are still consumed nowadays [4,5,7,11]. In South America, edible cyanobacterial macrocolonies known by the vernacular name of Llayta have been consumed since pre-Columbian times, and today they are available as dry biomass at food markets in northern Chile and southern Peru [8-10,14,16,22].

Andean wetlands over 4,000 m of altitude are habitats under daily and seasonal environmental stressors, where resilient life forms have evolved morphological and biochemical adaptive mechanisms to withstand high solar UV insolation, desiccation periods, and pH, osmotic and temperature changes and other environmental stressors [41]. The outer envelope and the mucilaginous inner matrix of cyanobacterial macrocolonies are physical protecting barriers against solar radiation and loss of intracellular water [24].



Then, genomic information on functional capabilities of the Llayta microbiome may be understood as part of the adaptive responses of this microbiome to survive and proliferate under the stressful environmental conditions found at the Andean wetlands where these macrocolonies can be found.

*Nostocales* are morphologically diverse and ubiquitous to almost any ecosystem on Earth, including those under extreme environmental conditions. Trichomes containing vegetative and N<sub>2</sub>-fixing heterocyst cells develop during the life cycle of diazotrophic species. Cyanobacterial macrocolonies can be found in nature as sheet-like or spherical forms, with green to brown colorations, depending upon the species, dehydration state, local environment, with a massive accumulation of filaments in a mucilaginous matrix, rich in a complex polysaccharide, enclosed by an outer envelope [7,24,42]. In natural environments, cyanobacterial macrocolonies are colonized by diverse microorganisms involved in biogeochemical cycles and ecosystem services, although differences in heterotrophic bacteria diversity among the inner matrix, the envelope, or the surrounding environment of *Nostoc* macrocolonies have been reported [25-27]. Secker et al. [25] reported absence of bacteria at the inner matrix of *Nostoc* macrocolonies from ephemeral wetlands in New Zealand, but high bacterial diversity at their external surface enriched with members of the genus *Sphingomonas*. Conversely, a different bacterial composition was found at the inner matrix, outer layer, and the littoral zone on *Nostoc* macrocolonies from Chungará Lake in northern Chile [26]. Comparatively, and using metagenomics, multiple taxonomic markers, and microscopic approaches, Satjarak et al. [27] reported high taxonomic diversity (cyanobacteria, microalgae, and anoxygenic bacterial genera) on the accompanying epimicrobiota of macroscopic dark-brown sheets of *Nostoc* from standing water pools at Lauca National Park in northern Chile.

The accompanying microflora of edible Llayta macrocolonies has been poorly explored and its presence becomes evident during attempts to isolate axenic trichomes. Vilo et al. [42] reported the identification and the draft genome of a *Bacillus* bacterium from the microbiota associated with Llayta colonies, as a first approach to address physiological relationships within the Llayta microbiome and to inquire into the survival and adaptive strategies to arsenic, metals, UV radiation, and other prevalent extreme environmental stressors at the Andes wetlands. In this context, the present report provides new metagenomics-based information, as a culture-independent approach to improve our understanding of the Llayta cyanosphere and to gain insights into the metabolic capabilities of Llayta cyanosphere and natural products prospection.

The metagenomics analyses on the cyanosphere of Llayta filaments provided proper sequencing data for identification and genome reconstruction of strain *Cyanocohniella* sp. LLY. The *Cyanocohniella* genus, Order Nostocales, is a recently described and revised genus with a filamentous diazotrophic holotype species *Cyanocohniella calida*, isolated from a thermal spring, Czech Republic [43]. The complex life cycle of *C. calida* shows similarity with *Pseudanabaena*-like stage, *Nostoc*-like stage, and *Chlorogloeopsis*-like stage, which makes the taxon especially difficult to identify from environmental samples. More recently, *Cyanocohniella crotaloides* sp. nov. was obtained after sampling sandy beach algal mats in Netherlands and described as showing a polymorphic life cycle resembling *Nostoc*-like thallus structures with curved, constricted filaments embedded in a nonlamellated, fine colorless sheath [44]. These authors also provide a detailed morphological comparison between these two *Cyanocohniella* species. In addition, strain TAU-MAC 3117 has been proposed as a third new species, *Cyanocohniella hyphalmyra*, isolated from epilithic mats on the surface of rocks in the shore of a brackish lake in Greece, and found positive for HCN production [45]. Finally, two other isolates, SY-1-2-EE and SY-1-2-Y, from biological soil crusts at the surroundings of saline potash tailing piles, were assigned to the *Cyanocohniella* genus based on SSU rRNA and ITS sequences [46]. More recently, Jung et

al. [47] have point out that all known members of the *Cyanocohniella* genus have been isolated from salty habitats, this genus might have had a marine origin and expanded to colonize salty terrestrial habitats. Based on our work, strain *Cyanocohniella* sp. LLY was identified as a new member of the *Cyanocohniella* genus and found at the cyanosphere of filaments of Llayta macrocolonies developing naturally in brackish water at Andean wetlands.

Gene mining conducted on the reconstructed genome of *Cyanocohniella* sp. LLY demonstrated the presence of conserved lycopene cyclase *cruA/cruP* genes and complete *mysABCD* and *scyABCDEF* gene clusters involved in mycosporine-like amino acids and scytonemin biosynthesis, two UV-absorbing molecules located at the outer cell membrane and the intracellular compartment, respectively [48]. This genomic analysis shows putative synthetic capabilities of *Cyanocohniella* sp. LLY needed for protection against photooxidative damages triggered by high-UV solar insolation prevailing at the Andes wetlands [33].

Lasso peptides are ribosomally-synthesized and post-translationally modified small peptides with antimicrobial, antiviral, and enzymatic inhibitors [49]. The gene cluster for Lasso peptide biosynthesis includes the precursor peptide, an ATP-dependent cysteine protease, and one ATP-dependent macrolactam synthetase. The general cluster organization has been studied in Actinobacteria, Firmicutes, and Proteobacteria phyla showing differences in the presence of an ABC-transporter gene and a splitting of the B proteins in Actinobacteria in contrast to the Proteobacteria cluster; however, both gene organizations can produce the Lasso peptide [49-51]. Gene mining on the genome of *Cyanocohniella* sp. LLY showed the presence of all genes needed for lasso peptide biosynthesis with a similar organization found in Actinobacteria [51]. Although the gene cluster in *Cyanocohniella* sp. LLY has a reverse direction for transcription (Fig. 3A) compared to the Actinobacteria cluster, it might produce the necessary peptides for the biosynthesis of Lasso peptide. According to phylogeny, the precursor peptide sequence is conserved in Nostocales (Fig. 3C) and would confirm the genomic potential of strain LLY for Lasso peptides biosynthesis.

Gene mining of the assembled genome of *Cyanocohniella* sp. LLY showed the potential capability for capreomycin biosynthesis and the production of members of the tuberactinomycin family of antibiotic peptides such as viomycin and capreomycin [52], but also for the biosynthesis of a microviridin precursor that would render a serine-protease inhibitor peptide and [53].

The bioinformatic analysis of the metagenome-assembled genome from *Cyanocohniella* sp. LLY showed the presence of CRISPR-Cas operons conferring an adaptive immunity defense against mobile genetic elements such as phages to the Andean cyanobacterium. Such defensive systems have been reported in archaea and bacteria [54,55], organized as CRISPR arrays (short, direct repeats of 20–40 bp separated by variable spacers) and Cas operons [56-57]. According to Shao et al. [58], the Cmr complex binding and cleaving of foreign RNA may involve a crRNA endoribonuclease activity coded at one of the Cmr1, Cmr3, Cmr4, or Cmr6 genes; in contrast, the Cmr2 gene is involved in binding specific targets. The crRNA transcribed by the CRISPR array would bind to CAS proteins to recognize the foreign nucleic acid, which would be cleaved by Cas1 and Cas2 nucleases [58-60]. Sequences for CRISPR-CAS system operons found at the metagenome-assembled genome from *Cyanocohniella* sp. LLY were substantially similar sequences from CRISPR-CAS operons available at public databases.

## 5. Conclusions

Metagenomics analyses on the cyanosphere of filaments isolated from edible Llayta macrocolonies, allowed us to identify the presence of the cyanobacterium strain *Cyano-*



*cohniiella* sp. LLY. Its reconstructed genome provided information on its putative capabilities for the biosynthesis of UV-absorbing molecules (mycosporine-like amino acids and scytonemin), Lasso peptides, microviridins, and the presence of CRISPR-CAS system operons. The assembled genome of strain LLY is now available for further inquiries, and may be instrumental for taxonomic advances on this genus as well as on its origin and spatio-temporal distribution [47, 61]. All known members of the *Cyanocohniella* genus have only been isolated from salty European habitats; the existence of strain *Cyanocohniella* sp. LLY comes to fill a biogeographic gap including the Andes Mountains wetlands in South America as a new habitat for the genus.

The Atacama Desert is a biodiversity hotspot deserving extensive research and environmental protection; edible Llayta macrocolonies from high-altitude Andean wetlands are another example of its microbial richness [62]. Further studies on the cyanosphere associated to Llayta will render insights into microbial functional capabilities, biosynthetic pathways, and adaptive strategies to the environmental conditions at high-altitude Andean wetlands.

**Author Contributions:** Conceptualization, C.V., B.G-S.; methodology, C.V., A.G., Q.D.; validation, A.G. C.V., Q.D.; formal analysis, C.V.; investigation, C.V., A.G., and B.G-S.; data curation, C.V., A.G.; writing-original draft preparation, B.G-S., C.V.; writing-review and editing, B.G-S, C.V., A.G., Q.D.; project administration, B.G-S.; funding acquisition, B.G.-S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by CONICYT, Chile, grant CeBiB FB-0001.

**Acknowledgments:** Data was generated in the Genome Sequencing Facility, which is supported by UT Health San Antonio, NIH-NCI P30 CA054174 (Cancer Center at UT Health San Antonio), NIH Shared Instrument grant 1S10OD021805-01.

**Conflicts of Interest:** The authors declare no conflict of interest. The funding organization had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Aaronson, S. (1986). A role for algae as human food in antiquity. *Food Foodways*. 1, 311-315. <https://doi.org/10.1080/07409710.1986.9961891>.
2. Niccolai, A., Bigagli, E., Biondi, N., Rodolfi, L., Cinci, L., Luceri, C., Tredici, M.R. (2017). In vitro toxicity of microalgal and cyanobacterial strains of interest as food source. *J. Appl. Phycol.* 29, 199–209. <https://doi.org/10.1007/s10811-016-0924-2>.
3. Wells, M.L., Potin, P., Craigie, J.S., Raven, J.A., Merchant, S.S., Helliwell, K.E., Smith, A.G., Camire, M.E., Brawley, S.H. (2017). Algae as nutritional and functional food sources: revisiting our understanding. *J. Appl. Phycol.* 29, 949-982 DOI 10.1007/s10811-016-0974-5.
4. Wells ML, Philippe Potin 2, James S Craigie 3, John A Raven 4 5, Sabeeha S Merchant 6, Katherine E Helliwell 7 8, Alison G Smith 7, Mary Ellen Camire 9, Susan H Brawley (2017). Algae as nutritional and functional food sources: revisiting our understanding. *J Appl Phycol.* 29(2):949-982. doi: 10.1007/s10811-016-0974-5.
5. Li, Z., Guo, M. (2018). Oncotarget. Healthy efficacy of Nostoc commune Vaucher. <https://doi.org/10.18632/oncotarget.23620>.
6. Nandagopal, P., Steven, A.N., Chan, L-W., Rahmat, Z., Jamaluddin, H., Noh, N.I.M. (2021). Bioactive metabolites produced by cyanobacteria for growth adaptation and their pharmacological properties. *Biology* 10, 1061. <https://doi.org/10.3390/biology10101061>.
7. Gao, K. (1998). Chinese studies on the edible blue-green alga, *Nostoc flagelliforme*: a review. *J. Appl. Phycol.* 10, 37-49. <https://doi.org/10.1023/A:1008014424247>.
8. Bertonio, L. Vocabulario de la Lengua Aymara. Colección Biblioteca Nacional de Chile: Impreso en la Compañía de Jesús, Perú. Electronic Document. 1612. Available online: <http://www.memoriachilena.cl/602/w3-article-8656.html> (accessed 11 October 2021).
9. Lagerheim, M.G. (1892). La Yuyucha. *La Nuova Notarisia* 3:1376-1377. <https://ia601307.us.archive.org/23/items/lanotarisiacomme7293levi/lanotarisiacomme7293levi.pdf>. (accessed 11 October 2021).
10. Mannheim, B. (1986). Poetic form in Guaman Poma's Wariqa Arawi. *Amerindian*, 11, 41–67. [https://sedyl.cnrs.fr/amerindia/articles/pdf/A\\_11\\_03.pdf](https://sedyl.cnrs.fr/amerindia/articles/pdf/A_11_03.pdf). (accessed 11 October 2021).
11. Becker, E.W. (2013). "Microalgae for Human and Animal Nutrition", in *Handbook of Microalgal Culture*, ed. A. Richmond (Oxford: John Wiley & Sons, Ltd.), 461-503. <https://doi.org/10.1002/9780470995280.ch18>.

12. Qiu, B., Liu, J., Liu, Z., Sliu. (2002). Distribution and ecology of the edible cyanobacterium Ge-Xian-Mi (*Nostoc*) in rice fields of Hefeng County in China. *J. Appl. Phycol.* 14, 423-429. <https://doi.org/10.1023/A:1022198605743>.
13. Gantar, M., Svirčev, Z. (2008). Microalgae and cyanobacteria: food for thought. *J. Phycol.* 44, 260-268. doi: 10.1111/j.1529-8817.2008.00469.x.
14. Gómez-Silva, B., Mendizabal, I., Tapia, I., Olivares, H. (1994). Microalgas del norte de Chile. IV. Composición química de *Nostoc commune* Llaíta. *Rev. Invest. Cient. Tecnol. Cs. Mar* 3, 19-25. <http://bioquimica.uantof.cl/files/Publicaciones/1994%20Rev%20Invest%20%20Nostoc%20Llaita25.pdf>. (accessed 11 October 2021).
15. Vargas, M.A., Rodríguez, H., Moreno Fernández, J., Olivares, H., Del Campo, J.A., Rivas, J., Guerrero, M.G. (1998). Biochemical composition and fatty acid content of filamentous nitrogen-fixing cyanobacteria. *J. Phycol.* 34, 812-817. <https://doi.org/10.1046/j.1529-8817.1998.340812.x>.
16. Rivera, M., Galetovic, A., Licuime, R., Gómez-Silva, B. (2018). A microethnographic and ethnobotanical approach to Llayta consumption among the Andes feeding practices. *Foods* 7, 202. doi:10.3390/foods7120202.
17. Bull, A.T., J.A. Asenjo, M. Goodfellow, Gómez-Silva, B. (2016). The Atacama Desert: Technical Resources and the Growing Importance of Novel Microbial Diversity. *Ann. Rev. Microbiol.* 70, 215-234. doi: 10.1146/annurev-micro-102215-095236.
18. Gómez-Silva, B. (2018). Lithobiontic life: Atacama rocks are well and alive. *Antonie van Leeuwenhoek* 111, 1333-1345. <https://doi.org/10.1007/s10482-018-1033-9>.
19. Houston, J. (2006). Variability of precipitation in the Atacama Desert: its causes and hydrological impact. *Int. J. Climatol.* 26, 2181-2198. DOI: 10.1002/joc.1359.
20. Villagrán C, Castro V, Sánchez G, Hinojosa F, Latorre C. The Altiplano tradition: ethnobotanic study at the Iquique Andes, I Region, Chile. *Chungara* 1999; 31: 81-186.
21. Pardo O, Pizarro JL. Chile: Pre-Hispanic Edible Plants. Ediciones Parina. Arica, Chile; 2003.
22. Galetovic, A., Araya, J.E., Gómez-Silva, B. (2017). Composición bioquímica y toxicidad de colonias comestibles de la cianobacteria andina *Nostoc* sp. Llayta. *Rev. Chil. Nutr.* 44, 360-370. <http://dx.doi.org/10.4067/s0717-75182017000400360>.
23. Galetović A., J. Azevedo, R. Castelo-Branco, F. Oliveira, B. Gómez-Silva, V. Vasconcelos. (2020). Absence of Cyanotoxins in Llayta, Edible Nostocaceae Colonies from the Andes Highlands. *Toxins* 12(6), 382; <https://doi.org/10.3390/toxins12060382>
24. Dodds, W., Gudder, D.A., Mollenhauer, D. (1995). The ecology of *Nostoc*. *J. Phycol.* 31, 2-18. <https://doi.org/10.1111/j.0022-3646.1995.00002.x>.
25. Secker, N.H., Chua, J.P.S., Laurie, R.E. (2017). Characterization of the Cyanobacteria and Associated Bacterial Community From an Ephemeral Wetland in New Zealand. *J. Phycol.* 52, 761-773. doi: 10.1111/jpy.12434.
26. Aguilar, P., Dorador, C., Vila, I., Sommaruga, R. (2019). Bacterial communities associated with spherical *Nostoc* macrocolonies. *Front. Microbiol.* 10,483. doi: 10.3389/fmicb.2019.00483.
27. Satjarak, A., Graham, L.E., Piotrowski, M.J., Trest, M.T., Wilcox, L.W., Cook, M.E., Knack, J.J., Arancibia-Avila, P. (2021). Shotgun metagenomics and microscopy indicate diverse cyanophytes, other bacteria, and microeukaryotes in the epimicrobiota of a northern Chilean wetland *Nostoc* (cyanobacteria). *J. Phycol.* 57, 39–50. DOI: 10.1111/jpy.13084-20-040.
28. Arnon, D.I., McSwain, B.D., Tsujimoto, H.Y., Wada, K. (1974). Photochemical activity and components of membrane preparations from blue-green algae. I. Coexistence of two photosystems in relation to chlorophyll a and removal of phycocyanin. *Biochim. Biophys. Acta*, 357, 231-245.
29. Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., Yamashita, H., Lam, T-W. (2016). MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*, 102, 3-11.
30. Wattam A.R., Abraham, D., Dalay, O., Disz, T.L., Driscoll, T., Gabbard, J.L., Gillespie, J.J., Gough, R., Hix, D., Kenyon, R., Machi, D., Mao, C., Nordberg, E.K., Olson, R., Overbeek, R., Pusch, G.D., Shukla, M., Schulman, J., Stevens, R.L., Sullivan, D.E., Vonstein, V., Warren, A., Will, R., Wilson, M.J., Yoo, H.S., Zhang, C., Zhang, Y., Sobral, B.W. (2014). PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res.* D581-D591. doi: 10.1093/nar/gkt1099.
31. Skinnider, M.A., Merwin, N.J., Johnston, C.W., Magarvey, N.A. (2017). PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res.* 45(W1), W49-W54. doi: 10.1093/nar/gkx320.
32. Blin, K., Shaw, S., Kloosterman, A.M., Charlop-Powers, Z., van Wezel, G.P., Medema, M.H., Weber, T. (2021). antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 49(W1): W29–W35. <https://doi.org/10.1093/nar/gkab335>.
33. Cordero, R.R., Damiani, A., Jorquera, J., Sepúlveda, E., Caballero, M., Fernandez, S., Feron, S., Llanillo, P.J., Carrasco, J., Laroze, D., Labbe, F. (2018). Ultraviolet radiation in the Atacama Desert. *Antonie Van Leeuwenhoek*. 111(8), 1301-1313. doi: 10.1007/s10482-018-1075-z.
34. Garneau, J. E., Dupuis, M.-V., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. doi: 10.1038/nature09523
35. Takeuchi, N., Wolf, Y. I., Makarova, K. S., and Koonin, E. V. (2012). Nature and intensity of selection pressure on CRISPR-associated genes. *J. Bacteriol.* 194, 1216–1225. doi: 10.1128/JB.06521-11
36. Marraffini, L. A. (2015). CRISPR-Cas immunity in prokaryotes. *Nature* 526, 55–61. doi: 10.1038/nature15386
37. Grissa, I., Vergnaud, G., Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8:172. doi: 10.1186/1471-2105-8-172
38. Shmakov, S. A., Sitnik, V., Makarova, K. S., Wolf, Y. I., Severinov, K. V., Koonin, E. V. (2017). The CRISPR spacer space is dominated by sequences from species-specific mobilomes. *mBio* 8:e01397-17. doi: 10.1128/mBio.01397-17

39. Shao, Y., Cocozaki, A. I., Ramia, N. F., Terns, R. M., Terns, M. P., Li, H. (2013). Structure of the Cmr2-Cmr3 subcomplex of the Cmr RNA silencing complex. *Structure*, 21(3), 376-384.
40. Jungblut, A.D., Raymond, F., Dion, S. Moineau, V. Mohit, G. Quang Nguyen, M. Déraspe, É. Francovic-Fontaine, C. Lovejoy, A.I. Culley M.B., Corbeil, J., Vincent, W.F. (2021). Genomic diversity and CRISPR-Cas systems in the cyanobacterium *Nostoc* in the High Arctic. *Environ. Microbiol.* 23, 2955-2968. doi: 10.1111/1462-2920.15481.
41. Albarracín, V.H., Kurth, D., Ordoñez, O.F., Belfiore, C., Luccine, E., Salum, G.M., Piacentini, R.D., Farías, M.E. (2015). High-up: a remote reservoir of microbial extremophiles in Central Andean Wetlands. *Front. Microbiol.* 6, 1404. Doi: 103389/fmicb.2015.01404.
42. Vilo, C., Galetovic, A., Araya, J. E., Gómez-Silva, B., Dong, Q. (2015). Draft Genome Sequence of a *Bacillus* Bacterium from the Atacama Desert Wetlands Metagenome. *Genome announcements*, 3(4), e00955-15. <https://doi.org/10.1128/genomeA.00955-15>
43. Kaštovský, J., Gomez, E.B., Hladil, J., Johansen, J.R. (2014). *Cyanocohniella calida* gen. et sp. nov. (Cyanobacteria: Aphani-zomenonaceae) a new cyanobacterium from the thermal springs from Karlovy Vary, Czech Republic. *Phytotaxa* 181 (5): 279–292. <http://dx.doi.org/10.11646/phytotaxa.181.5.3>
44. Jung, P., Mikhailyuk, T., Emrich, D., Baumann, K., Dultz, S., B. Büdel. (2020). Shifting Boundaries: Ecological and Geographical Range extension Based on Three New Species in the Cyanobacterial Genera *Cyanocohniella*, *Oculatella*, and, *Aliterella*. *J. Phycol.* 56, 1216–1231. <https://doi.org/10.1111/jpy.13025>.
45. Panou, M., Gkelis, S. (2022). Unravelling unknown cyanobacteria diversity linked with HCN production. *Mol. Phylogenetics Evol.* 166, 1077322. <https://doi.org/10.1016/j.ympev.2021.107322>
46. Sommer, V., Mikhailyuk, T., Glaser, K., Karsten, U. (2020). Uncovering Unique Green Algae and Cyanobacteria Isolated from Biocrusts in Highly Saline Potash Tailing Pile Habitats, Using an Integrative Approach. *Microorganisms* 8, 1667. <http://dx.doi.org/10.3390/microorganisms8111667>
47. Jung, P.; Sommer, V.; Karsten, U.; Lakatos, M. Salty Twins: Salt-Tolerance of Terrestrial *Cyanocohniella* Strains (Cyanobacteria) and Description of *C. rudolphia* sp. nov. Point towards a Marine Origin of the Genus and Terrestrial Long Distance Dispersal Patterns. *Microorganisms* 2022, 10,968. <https://doi.org/10.3390/microorganisms10050968>.
48. Garcia-Pichel, F., Castenholz, R.W. (1993). Occurrence of UV-Absorbing, Mycosporine-Like Compounds among Cyanobacterial Isolates and an Estimate of Their Screening Capacity. *Appl. Environ. Microbiol.* 59, 163-169.
49. Tietz, J.I., Schwalen, C.J., Patel, P.S., Maxson, T., Blair, P.M., Tai, H.C., Zakai, U.I., Mitchell, D.A. (2017). A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat Chem Biol.* 13(5), 470-478. doi: 10.1038/nchembio.2319.
50. Martin-Gómez, H., Tulla-Puche, J. (2018). Lasso peptides: chemical approaches and structural elucidation. *Org. Biomol. Chem.* 16(28), 5065-5080. DOI <https://doi.org/10.1039/C8OB01304G>.
51. Mevaere, J., Goulard, C., Schneider, O., Sekurova, O.N., Ma, H., Zirah, S., Afonso, C., Rebuffat, S., Zotchev, S.B., Li, Y. (2018). An orthogonal system for heterologous expression of actinobacterial lasso peptides in *Streptomyces* hosts. *Sci Rep.* 29;8(1), 8232. doi: 10.1038/s41598-018-26620-0.
52. Fei, X., Yin, X., Zhang, L. Zabriskie, T.M. (2007). Roles of VioG and VioQ in the Incorporation and Modification of the Capreomycin Residue in the Peptide Antibiotic Viomycin. *J. Nat. Prod.* 70(4), 618–622. doi:10.1021/np060605u.
53. Ziemert, N., Ishida, K., Weiz, A., Hertweck, C., Dittmann, E. (2010). Exploiting the Natural Diversity of Microviridin Gene Clusters for Discovery of Novel Tricyclic Depsipeptides. *Appl. Environ. Microbiol.* 76, 3568-3574. doi:10.1128/AEM.02858-09.
54. Sorek, R., Lawrence, C. M., and Wiedenheft, B. (2013). CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.* 82, 237–266. doi: 10.1146/annurev-biochem-072911-172315.
55. Marraffini, L. A. (2015). CRISPR-Cas immunity in prokaryotes. *Nature* 526, 55–61. doi: 10.1038/nature15386
56. Grissa, I., Vergnaud, G., Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8:172. doi: 10.1186/1471-2105-8-172
57. Shmakov, S. A., Sitnik, V., Makarova, K. S., Wolf, Y. I., Severinov, K. V., Koonin, E. V. (2017). The CRISPR spacer space is dominated by sequences from species-specific mobilomes. *mBio* 8:e01397-17. doi: 10.1128/mBio.01397-17
58. Shao, Y., Cocozaki, A. I., Ramia, N. F., Terns, R. M., Terns, M. P., Li, H. (2013). Structure of the Cmr2-Cmr3 subcomplex of the Cmr RNA silencing complex. *Structure*, 21(3), 376-384.
59. Garneau, J. E., Dupuis, M.-V., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. doi: 10.1038/nature09523
60. Takeuchi, N., Wolf, Y. I., Makarova, K. S., and Koonin, E. V. (2012). Nature and intensity of selection pressure on CRISPR-associated genes. *J. Bacteriol.* 194, 1216–1225. doi: 10.1128/JB.06521-11.
61. Komárek, J., Kaštovský, J., Mareš, J. & Johansen, J.R. (2014). Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach. *Preslia* 86, 295–335.
62. Gómez-Silva, B., Batista-García, R.A. (2022). The Atacama Desert: A Biodiversity Hotspot and Not Just a Mineral-Rich Region. *Front. Microbiol.* 13:812842. <https://doi.org/10.3389/fmicb.2022.812842>.