De-novo Assembly of Limnospira fusiformis Using Ultra-Long Reads

McKenna Hicks¹, Thuy-Khanh Tran-Dao², Logan Mulroney² and David L. Bernick²*

¹ Department of Microbiology and Environmental Toxicology, University of California, Santa Cruz, Santa Cruz, CA United States, ² Department of Biomolecular Engineering, University of California, Santa Cruz, Santa Cruz, CA United States

The Limnospira genus is a recently established clade that is economically important due to its worldwide use in biotechnology and agriculture. This genus includes organisms that were reclassified from Arthrospira, which are commercially marketed as “Spirulina.” Limnospira are photoautotrophic organisms that are widely used for research in nutrition, medicine, bioremediation, and biomanufacturing. Despite its widespread use, there is no closed genome for the Limnospira genus, and no reference genome for the type strain, Limnospira fusiformis. In this work, the L. fusiformis genome was sequenced using Oxford Nanopore Technologies MinION and assembled using only ultra-long reads (>35 kb). This assembly was polished with Illumina MiSeq reads sourced from an axenic L. fusiformis culture; axenicity was verified via microscopy and rDNA analysis. Ultra-long read sequencing resulted in a 6.42 Mb closed genome assembled as a single contig with no plasmid. Phylogenetic analysis placed L. fusiformis in the Limnospira clade; some Arthrospira were also placed in this clade, suggesting a misclassification of these strains. This work provides a fully closed and accurate reference genome for the economically important type strain, L. fusiformis. We also present a rapid axenicity method to isolate L. fusiformis. These contributions enable future biotechnological development of L. fusiformis by way of genetic engineering.

Keywords: Limnospira, Arthrospira, Spirulina, genome, nanopore, axenicity, cyanobacteria, long-read

INTRODUCTION

Limnospira is a recently established genus that contains three species reclassified from Arthrospira based on morphological, molecular, and ecological differences between the two genera (Nowicka-Krawczyk et al., 2019). Spirulina (Arthrospira) platensis SAG 85.79 (SAG 85.79 = CCALA 026 = UTEX 2340), Arthrospira maxima UTEX 2720, and Arthrospira platensis PCC 8005 were reclassified to Limnospira fusiformis, L. maxima, and L. indica, respectively (Voronichin, 1934; Nowicka-Krawczyk et al., 2019). Limnospira fusiformis is the type species of this new genus (Nowicka-Krawczyk et al., 2019). Although more strains possibly fall under the new Limnospira genus, there is insufficient evidence to reclassify them (Nowicka-Krawczyk et al., 2019). Both Arthrospira and Limnospira comprise photosynthetic filamentous cyanobacteria that form multicellular cylindrical trichomes (Belay and Gershwin, 2007; Nowicka-Krawczyk et al., 2019). Their cultivation is simple, inexpensive, and their alkaline growth preference makes commercial production less prone to living contaminants (Kebede and Ahlgren, 1996; Nowicka-Krawczyk et al., 2019). These aspects make Limnospira/Arthrospira attractive for use in many biotechnological and agricultural applications, including wastewater treatment, biofuels, biomanufacturing, medicine,
and nutritional supplementation (Belay et al., 1993; Tokuşoğlu and Ünal, 2003; Belay and Gershwin, 2007; Abed et al., 2016; Jiang et al., 2017; Aikawa et al., 2018; Czerwionka et al., 2018; Mehani et al., 2018; Zinicovscaia et al., 2018). Both L. fusicomaris and L. maxima are mass-produced, economically important, and are approved by the Food and Drug Administration (FDA) for human consumption (Belay and Gershwin, 2007; Sili et al., 2012; Nowicka-Krawczyk et al., 2019). Both are, confusingly, known as “Spirulina” in commercial settings (Sili et al., 2012; Nowicka-Krawczyk et al., 2019).

Documented consumption of “Spirulina” dates to ancient Aztec civilization (Walker et al., 2014; Wan et al., 2016). Presently, “Spirulina” is globally consumed as a nutritional supplement because they are extremely nutrient dense (Campanella et al., 1999; Tokuşoğlu and Ünal, 2003; Babadzhanov et al., 2004). A recent market research report predicted the global “Spirulina” market to reach $629.6 million and 68,025.2 tons by 2025 (Spirulina Market - Global Opportunity Analysis and Industry Forecast, 2019–2025).

In addition to being a massively utilized commercial “Spirulina” strain, L. fusicomaris is one of two that have a publicly available genetic system (A. platensis C1 being the second) (Jeamton et al., 2017; Dehghani et al., 2018). Despite its economic importance and status as the type species for the Limnospira genus, a published reference genome for L. fusicomaris does not yet exist.

There are many sequenced genomes in the “Spirulina” grouping; many of these were sequenced using short-read technologies resulting in multi-contig assemblies due to their repeat ridden character (Fujisawa et al., 2010; Carriero et al., 2011; Cheevadhanarak et al., 2012; Sili et al., 2012; Lefort et al., 2014; Dong et al., 2015; Xu et al., 2016; Suzuki et al., 2019). A recent L. fusicomaris KN assembly produced a non-final 5.78 Mb genome composed of 401 scaffolds with an N50 of 32,852 (GCA_014489865.1), demonstrating the limitations of final 5.78 Mb genome composed of 401 scaffolds with an N50 25.3 Mb (Spirulina Market - Global Opportunity Analysis and Industry Forecast, 2019–2025).

Materials and Methods

Bacterial Strains and Culture Conditions

A xenic culture of Limnospira fusicomaris was purchased from the University of Texas Culture Collection of Algae (UTEX 2340). The cells were grown in modified SAG medium (162 mM NaHCO3, 38 mM Na2CO3, 2.9 mM K2HPO4, 29.4 mM NaNO3, 5.74 mM K2SO4, 17.1 mM NaCl, 0.81 mM MgSO4 · 7H2O, 0.27 mM CaCl2 · 2H2O, 12.13 µM Na2EDTA · 2H2O, 2.16 µM FeCl3 · 6H2O, 1.32 µM MnCl2 · 4H2O, 0.22 µM ZnCl2, 0.134 µM CoCl2 · 6H2O, 0.154 µM Na2MoO4 · 2H2O, 0.08 µM CuSO4 · 5H2O, 0.15 µM ZnSO4 · 7H2O, 10 µM H3BO3, 0.1 µM cyanocobalamin) shaking at 150 rpm in 30°C with 12 h 80 µmol photons per m² light cycles using a 90 CRI light source with a color temperature of 3,000 K.

Physical L. fusicomaris Purification Treatments

Xenic L. fusicomaris cultures were inoculated (OD750:0.2) in modified SAG medium. Culture growth was measured in 24 h intervals via OD750. Once the cultures reached the log phase of growth, they were incubated without shaking overnight to allow the formation of a layer of floating cells. Floating filaments were collected from the culture's liquid surface and vortexed at medium speed for 15 s followed by 15 s on ice for 1 min. The vortexed culture was filtered using a 40 µm nylon cell strainer and washed with sterile SAG media to remove contaminants smaller than 40 µm. Filaments that remained on the strainer were suspended in 1 mL of sterile SAG, briefly vortexed, then centrifuged for 10 min at 8,000 g to separate the cellular mass into two phases: filaments pelleted with contaminants and buoyant filaments at the surface. The surface layer of filaments were microscopically examined for changes in gas vesicles via phase-contrast microscopy at 600× magnification using a Nikon Eclipse E400.
cells were treated with further chemical purification to establish an axenic culture.

**Chemical L. fusiformis Purification Treatments**

Sterile SAG media was supplemented with 65 μg/mL ampicillin, 77 μg/mL cefoxitin, and 100 μg/mL meropenem (Choi et al., 2008; Sena et al., 2011). The pH of the media was adjusted to 12.15 using NaOH, then re-sterilized by 0.22 μm filtration. Cells that were physically treated were inoculated into this medium and incubated in the dark for 4 days at 150 rpm and 30°C.

After 4 days, the entire culture was filtered through a 1 μm polycarbonate filter. Cells retained on the membrane were gathered and suspended in 1 mL sterile SAG medium. The cell suspension was centrifuged and the surface layer was collected. These surface layer cells were examined for contaminating species using dark-field and phase-contrast microscopy using a Nikon Eclipse E400 at 200× and 600× magnification. DNA was extracted from these cells and sequenced with Illumina sequencing to polish the genome assembly and to verify axenicity.

**DNA Extraction**

We adapted a CTAB-based *A. platensis* DNA extraction method to purify and preserve high molecular weight DNA (Morin et al., 2010). Throughout the method, to preserve high molecular weight DNA we used wide-bore tips, pipetted slowly, and only mixed with gentle inversion and finger-flicking to prevent mechanical DNA shearing. *L. fusiformis* cells were resuspended in 0.5 mL sterile extraction buffer (0.15 M NaCl, 0.1 M EDTA, pH 8.0) and subjected to three freeze-thaw cycles using dry ice and a 37°C bath to damage the cell walls and increase the efficiency of cell lysis. The cells were centrifuged for 10 min at 8,000 g, collected, and resuspended in CTAB buffer (75 mM Tris-HCl, 2% CTAB, 1.4 M NaCl, 1 mM EDTA, H2O, pH 8). These cells were enzymatically lysed with 50 mg of lysozyme at 37°C for 30 min. The lysed cells were incubated at 37°C for 1 h with 2% SDS, 5 mg/mL proteinase K, and 100 μg/mL RNase A. Following this incubation, the lysed cells were gently mixed by slow inversion and incubated at 65°C for 10 min to optimize the formation of CTAB-protein- and -polysaccharide complexes.

The sample was incubated with 1 volume of 24:1 chloroform:isoamyl alcohol on ice for 30 min. The sample was centrifuged for 10 min at 3,500 g and the aqueous phase was transferred to a fresh tube. One volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added, mixed with gentle inversion, and centrifuged for 3 min at 3,500 g. The aqueous phase was transferred to a fresh tube and the phenol:chloroform:isoamyl wash was repeated until the interphase was cleared of flocculent material. The aqueous phase was gently washed with one volume of 24:1 chloroform:isoamyl alcohol, centrifuged for 3 min at 3,500 g, and transferred to a fresh tube. The chloroform:isoamyl alcohol wash was repeated 4 more times.

The final aqueous phase was gently mixed with 1/10 volume 3 M NaOAc (pH 5.4) and 2.5 volumes of 100% ethanol to precipitate DNA. The sample was incubated at −20°C overnight, then centrifuged at 3,500 g for 1 h at 4°C to pellet the DNA. The supernatant was removed, and the DNA was washed twice with 70% ethanol and mixed by inversion. Following the last wash, the DNA pellets were allowed to air dry. TE (1X) buffer was added to the DNA pellet and it was then incubated at 37°C until fully dissolved. The DNA purity was assessed with a Nanodrop UV/VIS spectrophotometer and the size was observed on a 0.5% agarose gel alongside a NEB 1 kb extended ladder (#N3239S). We consistently extracted pure high molecular weight DNA using this method (Supplementary Figures 1, 2). High molecular weight DNA was selected for Nanopore sequencing via gel extraction.

**DNA Sequencing**

*L. fusiformis* cells that had undergone only physical purification were sequenced using one Oxford Nanopore Technologies’ (ONT) MinION flowcell. The DNA was prepared for sequencing using the SQK-LSK109 protocol following the manufacturer’s instructions. DNA extracted from the axenic *L. fusiformis* culture—which had undergone both physical and chemical treatments—was shipped to the University of California Davis DNA Technologies and Expression Analysis Cores for Illumina library preparation and sequencing. The DNA was sheared and size selected for 500 bp fragments and sequenced using a single Illumina MiSeq 2 × 300 run using index AAGGTACA, which accounted for 17% of the 21.8 M reads with an overall Q30 > 80%

**Long-Read Genome Assembly**

Default settings were used for all tools unless otherwise specified. The MinION reads were basecalled using ONT’s basecaller, Bonito v0.1.5. The basecalled MinION reads were filtered to include reads > 35 kb and these were assembled using Shasta v0.4.0 (Shafin et al., 2019). This assembly was polished using PEPPER v0.1.1, followed by three rounds of polishing with the Illumina data using Pilon v1.22 (Walker et al., 2014). NCBI’s Prokaryotic Genome Annotation Pipeline was used to annotate the genome. BUSCO v4.1.4 was used to assess genome quality by tabulating the 773 single-copy core cyanobacterial orthologs represented in the cyanobacteria_odb10 model (Seppey et al., 2019). NCBI tblastn and blastp (Altschul et al., 1990) were used to identify missing orthologs reported by BUSCO.

The genome was oriented using dnaA as a marker for the origin of replication; it was placed on the top strand and as the first gene using Geneious v11.0.5. Repeats were identified using the Repeat Finder v1.0.1 Geneious plugin. PlasmidSPAdes v3.12.0 and Bandage v0.8.1 were used to identify potential plasmids within the Illumina data (Wick et al., 2015; Antipov et al., 2016). PhyML v2.2.4 with 1,000 bootstrap steps and a Jukes-Cantor genetic distance model was used to build a maximum-likelihood phylogenetic tree using published cpcAB gene sequences (Guindon et al., 2010). The phylogenetic tree was visualized with SeaView v5.0.4 (Gouy et al., 2010).

---

1 https://github.com/nanoporetech/bonito
2 https://github.com/kishwarshafin/pepper
3 https://github.com/ncbi/pgap
Short-Read Genome Assembly

Illumina paired-end sequences were merged and adapters were trimmed using SeqPrep\(^4\), specifying a minimum overlap of 50 bases and enabling merging with the -s option. Of the 3.7 M original read-pairs, 2/3 were able to be merged, resulting in 2.36 M merged reads and 1.21 M unmerged read-pairs. The 1.2 M read-pairs and 2.36 M merged reads were assembled with the SPAdes-bwa mem v3.12.0 assembler using parameters -t 16 (16 threads) -v 1 (verbose level errors). This assembly was used only for visualization purposes (Figure 1B).

Axenicity Verification

Axenicity was verified using RNAmmer 1.2 to detect ribosomal RNA genes in axenic Illumina contigs (Lagesen et al., 2007). Organisms were identified by their ribosomal RNA genes using NCBI BLAST (Altschul et al., 1990).

RESULTS

Genome Sequencing Approach Produced a Closed Genome Assembly

The \textit{L. fusiformis} (Table 1) genome is highly repetitive, which may cause misassembly with short-read data, such as that produced by Illumina sequencing. We used long-read nanopore sequencing to assemble the genome to circumvent issues posed by repeated sequences. We performed nanopore sequencing on DNA that was extracted from a physically treated culture that resulted in predominantly \textit{L. fusiformis} DNA (Figure 2B and Supplementary Figure 13). Ultra-long read libraries maximize the overlaps between reads, minimize the opportunities for misassemblies caused by common repeats, and thereby reduce the need for a completely axenic culture. We acquired 1.72 M reads from a single MinION flowcell and filtered for reads > 35 kb (17,345 reads). The filtered reads resulted in a fully closed genome assembly with 114x coverage, at a predicted base accuracy of ~96.4% (Jain et al., 2018; Mulroney et al., 2020). To ensure the most reliable genome assembly, we isolated and

\(^4\)https://github.com/jstjohn/SeqPrep

| Property | Term | References |
|----------|------|------------|
| Current classification | Domain | TAS\(^a\) |
| | Bacteria | TAS\(^a\) |
| | Phylum | TAS\(^b\) |
| | Cyanobacteria | NAS |
| Class | Cyanophyceae | TAS\(^a\) |
| Order | Oscillatoriiales | TAS\(^b\) |
| Genus | Limnospira | TAS\(^b\) |
| Species | \textit{fusiformis} SAG 85.79 | TAS\(^b\) |
| Cell shape | Spiral | NAS |
| Sporulation | None | NAS |
| Encoded traits | Antibiotic resistance | IDA |
| | beta-lactam and fluoroquinolones | NAS |
| Temperature range | 20–40\(^\circ\)C | TAS\(^c\) |
| Optimum temperature | 30–45\(^\circ\)C | TAS\(^c\) |
| pH | 8.0–10.0 | TAS\(^c\) |
| Carbon source | Phototroph, mixotroph | TAS\(^c\) |
| Energy source | Phototroph | TAS\(^c\) |
| Relationship to oxygen | Aerobic | TAS\(^c\) |
| Pathogenicity | None | NAS |
| Origin | Natron Lake, Chad | NAS |
| Habitat | Freshwater | NAS |
| Latitude | 14.306969 N | NAS |
| Longitude | 18.581542 E | NAS |
| Obtained from | University of Texas, Strain UTEX 2340 | NAS |

Evidence codes: IDA, Inferred from Direct Assay (first time in publication); TAS, Traceable Author Statement. TAS\(^a\) : (Muhling, 2000), TAS\(^b\) : (Furmaniak et al., 2017), TAS\(^c\) : (Vonshak, 1997); NAS, Non-traceable Author Statement. These evidence codes are from the Gene ontology Project (Ashburner et al., 2000).
sequenced an axenic culture using selective gas vesicle collapse, filtration, antibiotic treatments and alkaline pH selection. DNA extracted from this culture was sequenced using short-read Illumina (157× coverage). The nanopore genome assembly was polished using this short-read library, with a predicted > 99% average base accuracy (Jain et al., 2018) (see section “Materials and Methods”). This produced an NCBI-classified complete assembly level genome (Kitts et al., 2016) with an N50 of 6.42 Mb. In contrast, the short-read only assembly had an N50 of 51.6 kb, thus it was discarded and not used outside of Figure 1B. The advantage of the ultra-long read library is clearly apparent when comparing assemblies using the two techniques independently (Figure 1). When combined, we produced a closed 6.42 Mb circular genome assembled as a single chromosome (Table 2).

Genome completeness is an important measure for assembly quality. One method to assess genome completeness is to identify core genes of the clade. We used BUSCO to assess genome completeness by tabulating the 773 single-copy core cyanobacterial orthologs (Seppey et al., 2019). The BUSCO score was 98.9% (765/773 complete). Of these complete genes, five were duplicated. Additionally, there were three genes that were fragmented and five that were reported missing. With further analysis of the five missing genes using NCBI tblastn and blastp (Altschul et al., 1990): two were present in alternate forms (inositol monophosphatase and n-acetyltransferase); recF and riboflavin synthase subunit alpha had 3′ truncations possibly derived from sequencing errors; and like all other *Limnospira* species, miniribonuclease 3 is also not present in *L. fusiformis* but ribonuclease 3 is present.
The genomes of *L. fusiformis* and its close relative *Arthrospira* sp. TJSD092 are within 10.7 kb in size and share the same 44.9% GC content (GGA_003060805.1). The DNA coding region (CDS) comprises 82.8% of the genome, with a total of 5,994 genes, 5,344 of which are protein-coding and 51 are RNA genes. The genome contains two rRNA operons and 41 tRNAs.

Selfish DNA elements tend to be repetitive, and we found that 29.6% of the genome is annotated as repetitive DNA (Genomes v11.0.5). Selfish DNA elements were abundant in the *L. fusiformis* genome and included 518 transposases, 12 recombinases, 89 toxin/antitoxin elements, and seven CRISPR arrays. We also found 37 reverse transcriptases and eight phage annotations, toxin/antitoxin elements, and seven CRISPR arrays. We also found 37 reverse transcriptases and eight phage annotations.

Selfish DNA elements were not detected. A report from 1993 (Song et al., 1993) indicated the presence of a plasmid in *Spirulina platensis*, however, this finding has not been described elsewhere (Fujisawa et al., 2010; Cheevadhanarak et al., 2012; Lefort et al., 2014). Therefore, we conducted a search for circular extrachromosomal DNA elements within the *L. fusiformis* genome. We used PlasmidSPAdes and Bandage to identify circular contigs within the Illumina library and screened the elements that were present in both outputs (Wick et al., 2015; Antipov et al., 2016). Our search for extrachromosomal elements was conducted with the short-read axenic library in order to exclude contaminating elements present in the long-read axenic library. We did not identify any extrachromosomal elements belonging to *L. fusiformis* (the PhiX Control v3 circular spike-in Illumina control library was present, as expected).

**Axenicity**

**Selective Gas Vesicle Collapse Is an Effective Physical Treatment**

Gas vesicles allow *A. platensis* cells to float during log-phase (Kim et al., 2005). If these vesicles proved to be more robust than the vesicles of contaminating species, this property could be used as a separation method. Under phase-contrast microscopy, gas vesicles appear as bright irregular shapes within these ∼80 µm buoyant trichomes (Cohen-Bazire et al., 1969; Sili et al., 2012). Gas vesicles collapse under pressure, causing cells to lose their buoyancy and appear darker and more hollow (Cohen-Bazire et al., 1969; Walsby and Bleything, 1988). We cultured xenic *L. fusiformis* under standard growth conditions and observed buoyancy from day three to at least day eight. These inhibit the stable transformation of *Arthrospira* (Cheevadhanarak et al., 2012; Jeamton et al., 2017).

**Phylogenetic Analysis Confirmed L. fusiformis Placement in the Limnospira Clade**

We built a phylogenetic tree based on the phycocyanin alpha and beta subunits, cpcAB (Figure 3; Manen and Falquet, 2002). The tree includes all *Arthrospira* and *Limnospira* genome assemblies, as well as additional strains that were previously placed in the *Limnospira* clade with publicly available cpcAB sequences (Nowicka-Krawczyk et al., 2019).

*Figure 3* shows *L. fusiformis* and most of the “Arthrospira” strains placed in the *Limnospira* clade (Nowicka-Krawczyk et al., 2019). *A. platensis* Paraca and *Arthrospira* sp. PLM2.Bin9 are the only strains not clearly in the *Limnospira* clade. Consistent with previous results (Cheevadhanarak et al., 2012; Nowicka-Krawczyk et al., 2019), the *Limnospira* clade is split into two major groups (Figure 3). Of the sequenced genomes (Table 3), *Arthrospira* sp. PCC 9108 and *A. platensis* NIES-39, NIES-46, YZ, FACHB-971, FACHB-439, and FACHB-835 are grouped together in one branch while a second group contains *L. fusiformis* SAG 85.79, *Arthrospira* sp. TSJD092, *Limnospira* sp. BM01, *L. fusiformis* KN, *Arthrospira* sp. TSJD091, *A. platensis* C1, *A. maxima* CS-328, *L. indica* PCC 8005, and *Arthrospira* sp. O9.13F (Figure 3).

**Extrachromosomal Elements Were Not Detected**

A report from 1993 (Song et al., 1993) indicated the presence of a plasmid in *Spirulina platensis*, however, this finding has not been described elsewhere (Fujisawa et al., 2010; Cheevadhanarak et al., 2012; Lefort et al., 2014). Therefore, we conducted a search for circular extrachromosomal DNA elements within the *L. fusiformis* genome. We used PlasmidSPAdes and Bandage to identify circular contigs within the Illumina library and screened the elements that were present in both outputs (Wick et al., 2015; Antipov et al., 2016). Our search for extrachromosomal elements was conducted with the short-read axenic library in order to exclude contaminating elements present in the long-read axenic library. We did not identify any extrachromosomal elements belonging to *L. fusiformis* (the PhiX Control v3 circular spike-in Illumina control library was present, as expected).

---

**TABLE 2 | Genome assembly details and summary statistics for *Limnospira fusiformis* SAG 85.79 (MIGS standard, Field et al., 2008).**
xenic cultures were pressurized with centrifugal forces at 2,000, 4,000, 8,000, and 16,000 g, and selective separation was found at 8,000 g (Supplementary Figures 3–11).

We enriched for *L. fusiformis* using selective gas vesicle collapse to deplete contaminating microbes. *L. fusiformis* cells were briefly vortexed to remove contaminating bacteria attached to the trichomes (Figure 2A), filtered at 40 µm, centrifuged to pellet the contaminants (Figure 2C), then transferred from the surface layer to sterile media (Figure 2B). Dark-field microscopy images at each stage confirmed the purification (Figures 2A–C and Supplementary Figures 12–14). We implemented these findings as a novel axenicity technique—selective gas vesicle collapse.

An Axenic *L. fusiformis* Culture Was Established Under Antibiotic and Alkaline Conditions

*L. fusiformis* tolerates elevated alkaline conditions and specific antibiotics, unlike microcystin-producing species (i.e., *Microcystis aeruginosa*, *Oscillatoria* spp., and *Nostoc* spp.), which often contaminate “Spirulina” supplements (Moreno et al., 2004; Sena et al., 2011; Roy-Lachapelle et al., 2017; Borowitzka, 2018; Khong et al., 2019). We found the highest tolerable pH (12.15) also caused contaminating species (i.e., *Microcystis* spp.) to lyse (Figures 2D,E and Supplementary Figures 15, 16). This high pH combined with an antibiotic cocktail selected against any residual contaminating species while minimizing damage to *L. fusiformis*, resulting in an axenic culture (Figure 2F and Supplementary Figure 17; Choi et al., 2008; Sena et al., 2011).

Microscopy and rDNA Sequence Analysis Support Axenicity of Prepared Culture

The treated culture was analyzed under dark-field microscopy and no contaminating organisms were detected. We then compared sequencing data from samples that had only undergone physical treatments—represented by the nanopore data—with samples that were subjected to physical and chemical treatments—represented by the Illumina data. We conducted a computational search for 16s and 23s ribosomal RNA gene sequences within these respective data sets. Four contaminating ribosomal gene sequences were identified within the nanopore data, with best matches to: *Coraliomargarita akajimensis*, *Lishizhenia caseinilytica*, *Halomonas desiderata*, and a member of **TABLE 3 | Summary of *Limnospira* and “Arthrospira” genome assemblies.**

| Strain                  | Assembly level | Sequencing technology | Coverage | N50 (bp) | References                         |
|------------------------|----------------|-----------------------|----------|----------|-----------------------------------|
| *L. fusiformis* SAG 85.79 | Complete       | Nanopore; Illumina    | 114 x    | 6,423,694| This study                        |
| “Arthrospira” sp. TJSD92 | Complete       | Illumina; PacBio      | 190 x    | 6,434,389| Genbank accession:                |
| “Arthrospira” sp. PCC 9108 | Complete       | PacBio                | 191 x    | 6,763,964| Genbank accession:                |
| *L. indica* PCC 8005   | Chromosome     | 454; Sanger           | n/a      | 1,412,831| Janssen et al., 2010:             |
| *Limnospira* sp. BM01  | Chromosome     | Illumina              | 6.43 x   | 61,454   | Genbank accession:                |
| “A. platensis” NIES-39 | Chromosome     | ABI 3730              | 11 x     | 619,347  | Fujisawa et al., 2010:            |
| “A. platensis” C1      | Chromosome     | 454; Sanger           | 28 x     | 206,210  | Cheevadhanarak et al., 2012:      |
| “A. platensis” YZ      | Chromosome     | ABI 3730; Illumina    | 86 x     | 1,054,592| Xu et al., 2016:                  |
| *Limnospira* fusiformis KN | Scaffold       | Illumina              | 350 x    | 32,852   | Genbank accession:                |
| “A. maxima” CS-328     | Contig         | n/a                   | n/a      | 92,573   | Carrieri et al., 2011:            |
| “A. platensis” NIES-46 | Contig         | MiSeq                 | 42 x     | 40,752   | Suzuki et al., 2019:              |
| *A. platensis* Paraca  | Contig         | Illumina              | 36 x     | 72,660   | Lefort et al., 2014:              |
| “Arthrospira” sp. TJSD091 | Contig       | Illumina              | 130 x    | 49,578   | Dong et al., 2015:                |
| “A. platensis” O9.13F  | Contig         | Illumina              | 20 x     | 6,514    | Genbank accession:                |
| “A. platensis” FACHB-971 | Contig        | Illumina              | 50 x     | 39,263   | Genbank accession:                |
| “A. platensis” FACHB-439 | Contig        | Illumina              | 50 x     | 44,370   | Genbank accession:                |
| “A. platensis” FACHB-835 | Contig        | Illumina              | 50 x     | 34,814   | Genbank accession:                |
| *Arthrospira* sp. PLM2.Bin9 | Contig       | Illumina              | 18 x     | 36,014   | Genbank accession:                |

Assembly level as described previously (Kitts et al., 2016).
**FIGURE 3** | Limnospira phylogenetic tree based on cpcAB loci. The sequence sources of the cpcAB loci are denoted: complete genome (red ****), chromosome assembly (blue **), scaffold assembly (green *), contig assembly (purple *), cpcAB sequence only (black). Bootstrap values of major branch points are shown. Quotation marks are used to indicate species proposed for *Limnospira* membership.
**DISCUSSION**

Long-read sequencing enabled closure of the *Limnospira fusiformis* genome in a single pass. The 6.42 Mb genome, assembled as a single circular chromosome, required the scaffolding power of long reads (>35 kb) and the accuracy afforded by the short-read library. Assemblies of this highly repetitive “Spirulina” family have been previously attempted without the guidance of long-reads; multi-contig drafts or incorrect genome structure can be the result (Alkan et al., 2011; Torresen et al., 2019). Indeed, the only closed genomes in the group have been possible with long-read produced scaffolds (Table 3). We believe our approach allows for a simple assembly that is less error prone when compared to using short-read data alone.

Nowicka-Krawczyk proposed the *Limnospira* clade using rRNA nucleotide sequences of 64 “Arthrospira” strains and established *L. fusiformis*, *L. indica*, and *L. maxima* as the founding members of the genus (Nowicka-Krawczyk et al., 2019). Phylogenetic analysis using published cpcAB sequences (Manen and Falquet, 2002) from that clade combined with the currently available *Arthrospira* genomes (Table 3) shows that most of these are grouped with *Limnospira*, with the exception of *A. platensis* Paraca and *Arthrospira* sp. PLM2.Bin9 (Figure 3). Furthermore, the optimal pH conditions derived from publications, media recipes, or the isolate geography (exception of *A. platensis* Paraca and *Arthrospira* sp. PLM2.Bin9, 2015; Shiraishi, 2015; Dineshkumar et al., 2016; Xu et al., 2016; 1679); (BioSample: SAMN10237416; Vonshak, 1997; ¸ Seker (pH 8–10), as described by Nowicka-Krawczyk (ATCC medium 1926)). Phylogenetic analysis using published 16s and 23s rDNA identification). Targeted mutations to the restriction modification system in *L. fusiformis* would be an ideal target for improving its transformation efficiency, as has been done in other organisms (Hoshino et al., 1980; Kretz et al., 1991; Hobson et al., 2015; Chen and Wang, 2019). Targeted mutations to the restriction modification system in *L. fusiformis* would be an ideal target for improving its transformation efficiency, as has been done in other organisms (Hoshino et al., 1980; Kretz et al., 1991; Hobson et al., 2008; Ferri et al., 2010). With the rapid proliferation of genetic engineering tools, new methods for customizing *L. fusiformis* will become applicable as it finds continued success in the global market.

**DATA AVAILABILITY STATEMENT**

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA623410.

**AUTHOR CONTRIBUTIONS**

MH, LM, and DB devised the project and designed the experiments. DB and LM supervised the project. MH performed the genome experiments. MH and T-KT-D performed
ACKNOWLEDGMENTS

We would like to thank Chris Vollmers and his lab for performing the DNA size selection, providing material for an ONT library preparation, and for their valuable advice. We also thank Mark Akeson’s lab, especially Miten Jain, for providing materials and advice for Nanopore sequencing. Much gratitude toward Marilou Sison-Mangus’ lab for providing CBAT and advice on cyanobacteria culturing and DNA purification. Sincere thanks to Elliot Roth of Spira Inc. for strain and culturing advice. We would also like to thank our teammates from the UCSC iGEM team of 2017, who we began this project with. We give a special acknowledgment to team members Mark Lund, Marissa Taub, Tyler Myers, Hailey Loehde-Woolard, Sarah Medor, and Jethro Marasigan who helped with culturing and brainstorming during the early stages of this research. We thank corporate sponsors Zymo Research, Biomatters Inc., and Avantor Biosciences for providing research materials.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.657995/full#supplementary-material

REFERENCES

Abed, E., Ihab, N., Suliman, E., and Mahmoud, A. (2016). Impact of Spirulina on nutritional status, haematological profile and anaemia status in malnourished children in the gaza strip: randomized clinical trial. *Matern. Pediatri. Nutr.* 2:2. doi: 10.4172/2472-1182.1000011

Aikawa, S., Inokuma, K., Wakai, S., Sasaki, K., Ogino, C., Chang, J.-S., et al. (2018). Direct and highly productive conversion of cyanobacteria *Arthrospira platensis* to ethanol with CaCl2 addition. *Biotechnol. Biofuels* 11:50. doi: 10.1186/s13068-018-1050-y

Allan, C., Sajadian, S., and Eichler, E. E. (2011). Limitations of next-generation genome sequence assembly. *Nat. Methods* 8, 61–65. doi: 10.1038/nmeth.1527

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(95)80360-2

Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., and Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* 21:30. doi: 10.1186/s13059-020-1935-5

Antipov, D., Hartwick, N., Shen, M., Raiko, M., Lapidus, A., and Pevzner, P. A. (2016). plasmidSPAdes: assembling plasmids from whole genome sequencing data. *Bioinforma. Oxf. Engl.* 32, 3380–3387. doi: 10.1093/bioinformatics/btw493

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene Ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Babadzhyan, A. S., Abdusamatova, N., Yusupova, F. M., Faizullaeva, N., Mezhlumyan, L. G., and Malikova, M. K. (2004). Chemical Composition of Spirulina platensis Cultivated in Uzbekistan. *Chem. Nat. Compd.* 40, 276–279. doi: 10.1023/B:CONC.0000039141.98247.e8

Belay, A., and Gershwin, M. E. (2007). *Spirulina in Human Nutrition and Health*. Boca Raton, FL: CRC Press.

Belay, A., Ota, Y., Miyakawa, K., and Shimamatsu, H. (1993). Current knowledge on potential health benefits of *Spirulina*. *J. Appl. Phycol.* 5, 235–241. doi: 10.1007/BF00004024

Borowitzka, M. A. (2018). "Chapter 3 - Biology of microalgae,” in *Microalgae in Health and Disease Prevention*, eds I. A. Levine and J. Fleurence (Cambridge, MA: Academic Press), 23–72. doi: 10.1016/B978-0-12-814405-6.00003-7

Campanella, L., Crescentini, G., and Avino, P. (1999). Chemical composition and nutritional evaluation of some natural and commercial food products based on *Spirulina*. *Analyst* 27, 533–540. doi: 10.1039/a999910z

Carriero, D., Ananyev, G., Lenz, O., Bryant, D. A., and Dismukes, G. C. (2011). Contribution of a Sodium Ion Gradient to Energy Conservation during Fermentation in the Cyanobacterium *Arthrospira (Spirulina)* maxima CS-328. *Appl. Environ. Microbiol.* 77, 7185–7194. doi: 10.1128/AEM.00612-11

Cheevasadhanarak, S., Pathioonrangsanik, K., Prommenenate, P., Kaewngam, W., Musigkain, A., Tragoonrung, S., et al. (2012). Draft genome sequence of *Arthrospira platensis* C1 (PCC9348). *Stand. Genomic Sci.* 6:43. doi: 10.4056/sigs.252955

Chen, S. P., and Wang, H. H. (2019). An engineered cas-transposon system for programmable and precise DNA transpositions. *bioRxiv [Preprint]* doi: 10.1101/654996

Choi, G.-G., Bae, M.-S., Ahn, C.-Y., and Oh, H.-M. (2008). Induction of axenic culture of *Arthrospira (Spirulina)* platensis based on antibiotic sensitivity of contaminating bacteria. *Biotechnol. Lett.* 30, 87–92. doi: 10.1007/s10529-007-9323-2

Cohen-Bazire, G., Kunisawa, R., and Pfennig, N. (1969). Comparative study of the structure of gas vacuoles. *J. Bacteriol.* 100, 1049–1061. doi: 10.1128/JB.100.2.1049-1061.1969

Czerwonka, A., Kalawaj, K., Sławińska-Brych, A., Lemieszek, M., Bartnik, M., Kw, K., et al. (2018). Anticancer effect of the water extract of a commercial *Spirulina (Arthrospira)* platensis product on the human lung cancer A549 cell line. *Biomed. Pharmacother.* 106, 292–302. doi: 10.1016/j.biopha.2018.06.116

Dehghani, J., Adibkia, K., Movafeghi, A., Barzegari, A., Pourseifi, M., Maleki-Kak哈尔, H., et al. (2018). Stable transformation of *Spirulina (Arthrospira)* platensis: a promising microalgae for production of edible vaccines. *Appl. Microbiol. Biotechnol.* 102, 9267–9278. doi: 10.1007/s00253-018-9296-7

Dineshkumar, R., Narendran, R., and Sampathkumar, P. (2016). Cultivation of *Spirulina* platensis in different selective media. *Indian J. Mar. Sci.* 45, 1749–1754.

Dong, S., Chen, J., Wang, S., Wu, Y., Hou, H., Li, M., et al. (2015). Draft genome sequence of cyanobacteria *Arthrospira* sp. TSD091 isolated from seaside wetland. *Mar. Genomics* 24, 197–198. doi: 10.1016/j.margen.2015.05.008

Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y., et al. (2015). Efficient CRISPR/Cas9-mediated targeted mutagenesis in Populus in the first generation. *Sci. Rep.* 5:12217. doi: 10.1038/srep12217

Fermentation in the Cyanobacterium *Arthrospira (Spirulina)* maxima CS-328. *Appl. Environ. Microbiol.* 77, 7185–7194. doi: 10.1128/AEM.00612-11

Limnospira fusiformis Complete Genome Assembly
Kim, S.-G., Choi, A., Ahn, C.-Y., Park, C.-S., Park, Y.-H., and Oh, H.-M. (2012). Elucidation of the gas vesicle gene clusters in cyanobacteria of the genus Arthrospira (Oscillatoriales, Cyanophyta) and correlation with ITS phylogeny. Eur. J. Phycol. 47, 233–244. doi: 10.1080/09670262.2012.692317

Moreno, I. M., Maraver, J., Agaete, E. C., Leao, M., Gago-Martínez, A., and Cameán, A. M. (2004). Decomposition of microcystin-LR, microcystin-RR, and microcystin-YR in water samples submitted to in vitro dissolution tests. J. Agric. Food Chem. 52, 5933–5938. doi: 10.1021/jf0489668

Morin, V., Vallaesy, T., Hendrickx, L., Natalie, L., and Wilmotte, A. (2010). An efficient DNA isolation protocol for filamentous cyanobacteria of the genus Arthrospira. J. Microbiol. Methods 80, 148–154. doi: 10.1016/j.mimet.2009.11.012

Mühling, M. (2000). Characterization of Arthrospira (Spirulina) strains. Available online at: http://theses.dur.ac.uk/1188/ (accessed December 29, 2020)

Munroley, L., Wulf, M. G., Schildkraut, L., Teeritzinis, G., Buswell, J., Jain, M., et al. (2020). Identification of high confidence human poly(A) RNA isoform scaffolds using nanopore sequencing. bioRxiv [Preprint]. doi: 10.1101/2020.11.18.389049

Nowicka-Krawczyk, P., Mühlensteinerová, R., and Hauer, T. (2019). Detailed characterization of the Arthrospira type species separating commercially growing taxon into the new genus Limnospora (Cyanobacteria). Sci. Rep. 9:694. doi: 10.1038/s41598-018-36831-0

Roy-Lachapelle, A., Sollicc, M., Bouchard, M. F., and Sauré, S. (2017). Detection of Cytotoxins in algae dietary supplements. Toxins 9:76. doi: 10.3390/toxins9030076

Sant’Anna, C. L. A., Teresade, P., Werner, V. R., Dogo, C. R., Rios, F. R., and de, C. (2008). Review of toxic species of Cyanobacteria in Brazil. Algal Stud. 126, 251–265. doi: 10.1127/1864-1318/2008/0126-0251

Şeker, A., Shahwan, T., Eroğlu, A. E., Yılmaz, S., Demirel, Z., and Dalay, M. C. (2008). Equilibrium, thermodynamic and kinetic studies for the biosorption of aqueous lead(II), cadmium(II) and nickel(II) ions on Spirulina platensis lipids. J. Hazard. Mater. 154, 973–980. doi: 10.1016/j.jhazmat.2007.11.007

Sena, L., Rojas, D., Montiel, E., González, H., Morot, J., and Naranjo, L. (2011). A strategy to obtain axenic cultures of Arthrospira spp. cyanobacteria. World J. Microbiol. Biotechnol. 27, 1045–1053. doi: 10.1007/s11274-010-0549-6

Seppey, M., Manni, M., and Zdobnov, E. M. (2019). “BUSCO: assessing genome assembly and annotation completeness,” in Gene Prediction: Methods and Protocols Methods in Molecular Biology, ed. M. Kollmar (New York, NY: Springer), 227–245. doi: 10.1007/978-1-4939-9173-0_14

Shafin, K., Pesout, T., Lorig-Roach, R., Haukness, M., Olsen, H. E., Bosworth, C., et al. (2019). Efficient de novo assembly of eleven human genomes using Prometheus sequencing and a novel nanopore toolkit. bioRxiv [Preprint]. doi: 10.1101/201715722

Shiraishi, H. (2015). Association of heterotrophic bacteria with aggregated Arthrospira platensis exopolysaccharides: implications in the induction of axenic cultures. Biosci. Biotechnol. Biochem. 79, 331–341. doi: 10.1007/10808/1945.2014.972333

Sili, C., Torzillo, G., and Vonshak, A. (2012). “Arthrospira (Spirulina),” in Ecology of Cyanobacteria II: Their Diversity in Space and Time, ed. B. A. Whitten (Dordrecht: Springer Netherlands), 677–705. doi: 10.1007/978-94-007-3855-3_25

Song, Q., Shun, T., Peijun, Z., and Tseng, C. K. (1993). Isolation of plasmid from the blue-green alga Spirulina platensis. Chin. J. Oceanol. Limnol. 11, 285–288. doi: 10.1007/BF02850862

Spirulina Market - Global Opportunity Analysis and Industry Forecast. (2019–2025). Available online at: https://www.meticulousresearch.com/product/spirulina-market-5070/?utm_source=Globenewswire.com&utm_campaign=PressRelease&utm_medium=PressRelease&utm_campaign=Paid (accessed December 21, 2020)

Suzuki, S., Yamaguchi, H., and Kawachi, M. (2019). The draft genome of a hydrogen-producing Cyanobacterium, Arthrospira platensis NIES-46. J. Genomics 7, 56–59. doi: 10.7150/jagen.38149

Tokugou, Ö. and Ünal, M. K. (2003). Biomass nutrient profiles of three microalgae: Spirulina platensis, Chlorella vulgaris, and Isochrisis galbana. J. Food Sci. 68, 1144–1148. doi: 10.1111/j.1365-2621.2003.tb0965x.x

Torresen, O. K., Star, B., Mier, P., Andrade-Navarro, M. A., Bateman, A., Jarnot, P., et al. (2019). Tandem repeats lead to sequence assembly errors and impose
multi-level challenges for genome and protein databases. *Nucleic Acids Res.* 47, 10994–11006. doi: 10.1093/nar/gkz841

Tzfira, T., Frankman, L. R., Vaidya, M., and Citovsky, V. (2003). Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. *Plant Physiol.* 133, 1011–1023. doi: 10.1104/pp.103.032128

Vergunst, A. C., Jansen, L. E. T., and Hooykaas, P. J. J. (1998). Site-specific integration of Agrobacterium T-DNA in *Arabidopsis thaliana* mediated by Cre recombinase. *Nucleic Acids Res.* 26, 2729–2734. doi: 10.1093/nar/26.11.2729

Vonshak, A. (1997). *Spirulina Platensis Arthrospira: Physiology, Cell-Biology and Biotechnology*. Boca Raton, FL: CRC Press.

Voronichin, N. N. (1934). *Spirulina fusiformis Voronichin 1934: Algaebase*. Available online at: https://www.algaebase.org/search/species/detail/?tc=accept&species_id=47838 (accessed December 21, 2020)

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakhthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. doi: 10.1371/journal.pone.0112963

Walsby, A. E., and Bleything, A. (1988). The dimensions of Cyanobacterial gas vesicles in relation to their efficiency in providing buoyancy and withstanding pressure. *Microbiology* 134, 2635–2645. doi: 10.1099/00221287-134-10-2635

Wan, D., Wu, Q., and Kuča, K. (2016). "Spirulina," in *Nutraceuticals*, Chap. 42, ed. R. C. Gupta (Boston: Academic Press), 569–583. doi: 10.1016/B978-0-12-802147-7.00042-5

Wick, R. R., Schultz, M. B., Zobel, J., and Holt, K. E. (2015). Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* 31, 3350–3352. doi: 10.1093/bioinformatics/btv383

Xu, T., Qin, S., Hjä, Y., Song, Z., Ying, J., Li, P., et al. (2016). Whole genomic DNA sequencing and comparative genomic analysis of *Arthrospira platensis*: high genome plasticity and genetic diversity. *DNA Res.* 23, 325–338. doi: 10.1093/dnares/dsw023

Yadav, A., Monsieurs, P., Misztak, A., Waleron, K., Leys, N., Cuypers, A., et al. (2020). Helical and linear morphotypes of *Arthrospira* sp. PCC 8005 display genomic differences and respond differently to 60Co gamma irradiation. *Eur. J. Phycol.* 55, 129–146. doi: 10.1080/09670262.2019.1675763

Zhao, W., Wang, J., Xu, S., Lei, Y., Yang, R., Shi, L., et al. (2020). Bacterioplankton community variation in Bohai Bay (China) is explained by joint effects of environmental and spatial factors. *MicrobiologyOpen* 9:e997. doi: 10.1002/mbo3.997

Zinicovscaia, I., Cepoi, L., Povar, I., Chiriac, T., Rodlovskaya, E., and Culicov, O. A. (2018). Metal uptake from complex industrial effluent by cyanobacteria *Arthrospira platensis*. *Water, Air, Soil Pollut.* 229:220. doi: 10.1007/s11270-018-3873-3

Conflict of Interest: LM has received reimbursement for travel, accommodation and conference fees to speak at events organized by ONT. The authors declare that this study received materials from Zymo Research, Biomatters Inc., and Avantor Biosciences. These sponsors were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Copyright © 2021 Hicks, Tran-Dao, Mulroney and Bernick. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.