



OPEN ACCESS

EDITED BY

Feng Gao,
Tianjin University, China

REVIEWED BY

Maurycy Daroch,
Peking University, China
Shawn E. McGlynn,
Tokyo Institute of Technology, Japan
Gen Enomoto,
University of Freiburg, Germany

*CORRESPONDENCE

Chih-Horng Kuo
chk@gate.sinica.edu.tw
Hsiu-An Chu
chuha@gate.sinica.edu.tw

†These authors have contributed
equally to this work

SPECIALTY SECTION

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

RECEIVED 30 April 2022

ACCEPTED 20 July 2022

PUBLISHED 11 August 2022

CITATION

Cheng Y-I, Lin Y-C, Leu J-Y, Kuo C-H
and Chu H-A (2022) Comparative
analysis reveals distinctive genomic
features of Taiwan hot-spring
cyanobacterium
Thermosynechococcus sp. TA-1.
Front. Microbiol. 13:932840.
doi: 10.3389/fmicb.2022.932840

COPYRIGHT

© 2022 Cheng, Lin, Leu, Kuo and Chu.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Comparative analysis reveals distinctive genomic features of Taiwan hot-spring cyanobacterium *Thermosynechococcus* sp. TA-1

Yen-I Cheng^{1†}, Yu-Chen Lin^{1†}, Jyh-Yih Leu²,
Chih-Horng Kuo^{1*} and Hsiu-An Chu^{1*}

¹Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan, ²Department of Life Science, Fu Jen Catholic University, New Taipei City, Taiwan

Thermosynechococcus is a genus of thermophilic unicellular cyanobacteria that dominates microbial mats in Asian non-acidic hot springs. These cyanobacteria are the major primary producers in their ecological niches and are promising sources of thermostable enzymes for biotechnology applications. To improve our understanding of these organisms, we conducted whole-genome sequencing of a novel strain for comparative analysis with other representatives in the same genus. This newly characterized strain, *Thermosynechococcus* sp. TA-1, was isolated from the Taian hot springs in Taiwan. Analyses based on average nucleotide identity (ANI) and genome-scale phylogeny suggested that TA-1 and another Taiwanese strain CL-1 belong to a novel species-level taxon. Two metagenome-assembled genomes (MAGs) originated from India represent the sister group, and *Thermosynechococcus elongatus* PKUAC-SCTE542 from China is the next closest lineage. All cultivated strains and MAGs from Japan form a separate monophyletic clade and could be classified into two species-level taxa. Intriguingly, although TA-1 and CL-1 share 97.0% ANI, the genome alignment identified at least 16 synteny breakpoints that are mostly associated with transposase genes, which illustrates the dynamic nature of their chromosomal evolution. Gene content comparisons identified multiple features distinct at species- or strain-level among these *Thermosynechococcus* representatives. Examples include genes involved in bicarbonate transportation, nitric oxide protection, urea utilization, kanamycin resistance, restriction-modification system, and chemotaxis. Moreover, we observed the insertion of type II inteins in multiple genes of the two Taiwanese strains and inferred putative horizontal transfer of an asparagine synthase gene (*asnB*) associated with exopolysaccharides gene cluster. Taken together, while previous work suggested that strains in this genus share a highly conserved genomic core and no clear genetic differentiation could be linked to environmental factors, we found that the overall pattern of gene content divergence is largely congruent with core genome phylogeny. However, it

is difficult to distinguish between the roles of phylogenetic relatedness and geographic proximity in shaping the genetic differentiation. In conclusion, knowledge of the genomic differentiation among these strains provides valuable resources for future functional characterization.

KEYWORDS

thermophilic cyanobacterium, *Thermosynechococcus*, genome, comparative genomics, hot-spring cyanobacteria

Introduction

Thermophilic cyanobacteria grow photosynthetically and are often the primary producers in non-acidic hot springs (Ward et al., 2012). The underlying mechanisms of how thermophilic cyanobacteria adapt to diverse ecological niches and various stressful conditions of hot springs are still not fully understood (Chen et al., 2021; Prondzinsky et al., 2021). The information would be valuable for developing feasible solutions to mitigate the impacts of global warming (Matos et al., 2021). Additionally, the thermostable enzymes and bioproducts of these bacteria have been utilized in numerous scientific research and biotechnology applications (Leu et al., 2013; Su et al., 2013; Liang et al., 2019; Patel et al., 2019; Zuliani et al., 2021), which further illustrates the importance of these organisms.

Among thermophilic cyanobacteria, the genus *Thermosynechococcus* dominates the microbial mats in Asian non-acidic hot springs in the temperature range of about 50–65°C (Liao et al., 2006; Everroad et al., 2012; Tang et al., 2018; Patel et al., 2019; Prondzinsky et al., 2021). The genome sequences of several representative *Thermosynechococcus* strains have been determined and analyzed. These include *Thermosynechococcus elongatus* BP-1, *Thermosynechococcus vulcanus* NIES-2134, and *Thermosynechococcus* sp. NK55 from Japan (Nakamura et al., 2002; Stolyar et al., 2014); *Thermosynechococcus elongatus* PKUAC-SCTE542 from China (Liang et al., 2019); and *Thermosynechococcus* sp. CL-1 from Taiwan (Hsueh et al., 2007; Cheng et al., 2020; Table 1). The closest known relative of this genus is the *Synechococcus* phylogenetic lineage C1 (Papke et al., 2003), represented by *Synechococcus lividus* PCC 6715, which was isolated from a hot spring in Yellowstone National Park in the United States (Dyer and Gafford, 1961).

Thermosynechococcus genomes typically have smaller sizes and higher GC content than most mesophilic cyanobacteria (Larsson et al., 2011; Alcorta et al., 2020; Prondzinsky et al., 2021). Recently, a study found that *Thermosynechococcus* strains have a highly conserved genome core, and no clear genetic differences or adaptive mechanisms could be detected when comparing strains isolated from hot

springs with different environmental factors (Prondzinsky et al., 2021). This observation contrasts with a study of *Synechococcus* phylogenetic lineage A/B strains from a hot spring in Yellowstone National Park, which identified genetic differentiation between high- and low-light adapted strains (Becraft et al., 2015; Olsen et al., 2015). This difference between these two distinct genera remains to be investigated.

For genomic characterization of *Thermosynechococcus*, we previously studied a strain CL-1 isolated from the Chin-Lun hot spring in Eastern Taiwan (Cheng et al., 2020). CL-1 was genetically divergent from other characterized strains in the same genus and likely represents a novel species-level taxon based on genome-based phylogeny and average nucleotide identity (ANI) analysis. Furthermore, as compared with other strains, CL-1 has distinct genes related to photosynthesis, transporters, signal transduction, the chaperone/usher system, nitric oxide protection, antibiotic resistance, prokaryotic immunity systems, and other physiological processes (Cheng et al., 2020). However, with CL-1 as the only representative of this novel species-level taxon, whether those genetic features are specific at the level of species or strains is unknown. To address this question, we studied another Taiwanese strain, TA-1, which was isolated from the Taian hot springs in Central Taiwan (Leu et al., 2013). TA-1 is able to grow photosynthetically under high levels of CO₂ (up to 80%) and has been applied to produce C-phycocyanin with thermo- and pH-stability for bio-industrial applications (Leu et al., 2013).

In this work, we determined the complete genome sequence of TA-1 and conducted a comparative analysis with other *Thermosynechococcus* strains to gain a better understanding of these cyanobacteria. In addition to providing a first look into the within-species level diversity of this species-level taxon in Taiwan, we improved upon our previous work (Cheng et al., 2020) by incorporating multiple metagenome-assembled genomes (MAGs) with diverse geographic origins to achieve a better sampling of *Thermosynechococcus* diversity (Table 1). Finally, the updated comparative analysis used newer versions of annotation of these *Thermosynechococcus* genomes. This update is important because the annotation of PKUAC-SCTE542 had a major update that changed the number of intact coding sequences from 1,625 to 2,508, which would

greatly impact the inference of the core genome composition of *Thermosynechococcus* (Prondzinsky et al., 2021).

Materials and methods

Biological material and growth conditions

The strain TA-1 was isolated from the Tai'an hot springs (40–63°C, pH 7.5) in Miaoli County, Taiwan (Leu et al., 2013). The water sample had a conductivity of 1.97 mS/cm and the concentration (unit: mEq/L) of major ions were: F⁻, 2.69; K⁺, 1.21; Ca²⁺, 0.99; Cl⁻, 26.96; Mg²⁺, 0.96; Na⁺, 478.12; Si⁴⁺, 11.83; SO₄²⁻, 5.47; HCO₃²⁻, 1044.43 (Central Geological Survey, Ministry of Economic Affairs, Taiwan). The strain underwent axenic culture at 50°C in BG-11 medium supplemented with 20 mM TES (pH 8.0) under continuous white LED light (20 μmol photons m⁻² s⁻¹). For antibiotic resistance experiments, the *Thermosynechococcus* cultures were grown on BG-11 solid agar plates (5 mM TES, pH 8.0) containing 5 μg/ml of kanamycin. For phototaxis experiments, the TA-1 and CL-1 cultures were grown on BG-11 solid agar plates (5 mM TES, pH 8.0) under directional white light.

Microscopy

The procedures for microscopy were modified from those described in our previous work (Cho et al., 2019). Briefly, for scanning electron microscopy, the samples were fixed in 2.5% glutaraldehyde and 4% paraformaldehyde in 0.1 M sodium phosphate buffer, pH 7.0 at room temperature for 1 h. After rinsing with buffer three times for 20 min, samples were dehydrated in an ethanol series. Critical point drying was done with a Leica EM CPD300 critical point dryer. After coating with a Hitachi E-1010 ion sputter, an FEI Quanta 200 scanning electron microscope at 20 kV was used for viewing and imaging.

For transmission electron microscopy, cultured cells were frozen using high-pressure freezer (Leica EMPACT2) at 2,000–2,050 bar. Freeze substitution was performed in anhydrous acetone (containing 2% OsO₄) using a Leica EM AFS2. The samples were kept at -90°C for 2 days, at -60°C for 12 h, at -20°C for 12 h, 0°C for 12 h, and raised to room temperature. After rinsing with acetone two times (2 h each time), the samples were infiltrated and embedding with the Spurr's resin. Ultra-thin sections (70–90 nm) were cut using a Reichert UltraCut S or Leica EM UC7 and collected using 100 mesh copper grids. The sections were stained with 5% uranyl acetate in 50% methanol for 10 min and 0.4% lead citrate for 4–6 min. Sections were examined using a Philips CM 100 transmission electron microscope at 80 kV and the images were captured with a Gatan Orius CCD camera.

Genome sequencing and analysis

The procedures for genome sequencing and analysis were based on those described in our previous work (Cheng et al., 2020; Weng et al., 2021). All bioinformatics tools were used with the default settings unless stated otherwise.

Briefly, total genomic DNA was prepared from 150 ml of liquid culture grown for 5 days using the DNeasy Plant Maxi Kit (QIAGEN, Germany). The quantity and quality of the DNA sample were assessed with a NanoDrop 2000 spectrophotometer (Thermo Fisher, United States) and 1% agarose gel electrophoresis. Two sequencing platforms were used for whole-genome shotgun sequencing, namely, Illumina short reads and Oxford Nanopore Technologies (ONT) long reads. For Illumina sequencing, a paired-end library was prepared by using KAPA LTP Library Preparation Kit (KK8232; Roche, Switzerland) without amplification, then sequenced by using a HiSeq X Ten sequencer. For ONT sequencing, the library was prepared by using the ONT Ligation Kit (SQK-LSK109) and sequenced by using MinION (FLO-MIN106; R9.4 chemistry and MinKNOW Core v3.6.0); Guppy v3.4.5 was used for base calling. The raw reads from both platforms were combined for *de novo* assembly by using Unicycler v0.4.8-beta (Wick et al., 2017). For validation, the Illumina and ONT raw reads were mapped to the assembly by using BWA v0.7.12 (Li and Durbin, 2009) and Minimap2 v2.15 (Li, 2018), respectively. The results were programmatically checked by using SAMtools v1.2 (Li et al., 2009) and manually inspected by using IGV v2.3.57 (Robinson et al., 2011). The finalized assembly was submitted to the National Center for Biotechnology Information (NCBI) and annotated by using their Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova et al., 2016).

For comparative analysis, a list of *Thermosynechococcus* strains with complete genome sequences available from GenBank (Benson et al., 2018) was compiled (Table 1). Furthermore, multiple MAGs that may be assigned to *Thermosynechococcus* (Alcorta et al., 2020; Prondzinsky et al., 2021) were included to improve the sampling of genetic, phylogenetic, and geographic diversity. One MAG, OHK43 from Japan, was omitted because this assembly (accession JACOMP000000000) has > 99.6% ANI with NK55, is highly fragmented with 273 contigs, and lacks annotation. *S. lividus* PCC 6715, which has been demonstrated as a sister lineage of *Thermosynechococcus* (Liang et al., 2019; Komarek et al., 2020; Tang et al., 2022), was included as an outgroup. The annotation was based on those provided in the GenBank files. The homologous gene clusters (HGCs) among these genomes were identified by using OrthoMCL v1.3 (Li et al., 2003).

For phylogenetic analysis, MUSCLE v3.8.31 (Edgar, 2004) was used to generate multiple sequence alignments and PhyML v3.3 (Guindon and Gascuel, 2003) for maximum likelihood inference. The proportion of invariable sites and the gamma distribution parameter were estimated from the dataset, and

the number of substitute rate categories was set to four. To quantify genome similarity, FastANI v1.1 (Jain et al., 2018) was used to calculate the percentage of genomic segments mapped and the ANI based on these segments for each genome pair. For pairwise genome alignment, the NUCleotide MUMmer (NUCmer) program of the MUMmer package v3.23 (Kurtz et al., 2004) was used with the setting “–maxmatch –mincluster 2000.” Transposase genes based on annotation keyword search were used as queries to search the entire genome sequences using TBLASTN v2.11.0 + (Camacho et al., 2009). Only hits with high-scoring pairs accounting for at least 70% of the query length and overall amino acid sequence identity of at least 90% were retained.

The packages ggplot2 v3.3.3 (Wickham, 2009) and genoplots v0.8.9 (Guy et al., 2010) were used to visualize the organization and synteny conservation of selected genes.

Results and discussion

Morphology of *Thermosynechococcus* sp. TA-1 cells

The cell morphology of *Thermosynechococcus* sp. TA-1 is shown in [Supplementary Figure 1](#). Consistent with the previous description (Leu et al., 2013), TA-1 is unicellular and has rod-shaped cells that are 1.2–2.5 μm in width and 6.0–9.0 μm in length.

Complete genome sequence of *Thermosynechococcus* sp. TA-1

The whole-genome shotgun sequencing of *Thermosynechococcus* sp. TA-1 generated 2,432,241*2 pair-end Illumina reads and 73,708 ONT long reads (average length = 15,070 bp, N50 length = 24,058 bp, and total length = 1,110,800,338 bp). These reads provided ~272-fold (Illumina) and ~401-fold (ONT) coverage for the single 2,658,717-bp circular chromosome produced by *de novo* assembly. No plasmid was found. The annotation includes 42 tRNA genes, one complete set of 16S-23S-5S rRNA genes, and 2,523 protein-coding genes.

Comparison with other strains

A total of 1,359 HGCs were shared by all of the genome sequences compared ([Table 1](#)). When the outgroup and the MAGs were excluded, the six *Thermosynechococcus* strains with complete genome sequences available shared 2,058 HGCs, which accounted for ~84% of the gene content in each individual strain. This finding differs substantially from our

previous result that only 1,264 HGCs could be identified among *Thermosynechococcus* representatives (Cheng et al., 2020), mainly because of the early version of PKUAC-SCTE542 genome annotation contains 944 pseudogenes and only 1,625 intact protein-coding genes. With the updated annotation of this genome, our new result is consistent with a recent inference that *Thermosynechococcus* has a highly conserved core genome (Prondzinsky et al., 2021).

According to the 1,311 HGCs present as single-copy genes in all strains analyzed, we obtained a concatenated alignment containing 383,962 aligned amino acid sites. A maximum-likelihood phylogeny inferred from this alignment received 100% bootstrap support for all except one branch ([Figure 1A](#)). Based on this phylogeny and all possible pairwise ANI analyses, TA-1 was most closely related to another Taiwanese strain, CL-1, and these two strains shared 97.0% ANI, which is above the 95% threshold suggested for delineating bacterial species (Jain et al., 2018). Two MAGs collected in India, M98_K2018_005 and M55_K2018_012, were inferred as the sister species of the Taiwanese lineage.

Among the cultivated strains, PKUAC-SCTE542 from China represented the most closely related species, with ~90.1% ANI when comparing the two Taiwanese strains. The three cultivated strains from Japan all had smaller genome sizes ([Table 1](#)) and belonged to a monophyletic clade that included two species-level taxa based on the 95% ANI threshold ([Figure 1A](#)). One MAG collected in India, M46_R2017_013, represented the basal lineage of *Thermosynechococcus*. For the comparison of overall genome sequence similarity, all *Thermosynechococcus* representatives shared > 85% ANI, while *S. lividus* PCC 6715 fell below this threshold. The ANI pattern, as well as the comparison of overall gene content ([Figure 1B](#)), both suggested that PCC 6715 is a suitable outgroup. Because M46_R2017_013 and PCC 6715 were both highly divergent from others in the overall gene content, and we conducted a second comparison that excluded these two genomes to improve resolution ([Figure 1C](#)). Although those two Indian MAGs represented the sister group to the Taiwanese strains based on the phylogeny of conserved genes, their gene content is quite different. This observation may be due to the nature of incompleteness (and possible contamination) of MAGs and requires future improvement in the availability of complete genome sequences for further investigation.

The taxonomy of these cyanobacteria has been under active revision in recent years (Liang et al., 2019; Komarek et al., 2020; Prondzinsky et al., 2021; Tang et al., 2022). To avoid confusion, we utilized strain names as the major identifiers in this work and described the scientific names based on the associated GenBank records ([Table 1](#)) when necessary for trackability. Nevertheless, based on the results of genome-scale phylogeny and overall genome similarity as quantified by ANI, it appeared that further taxonomic revision may be necessary. For example, given the high level of genome similarity (i.e., ANI = 99.0%) between

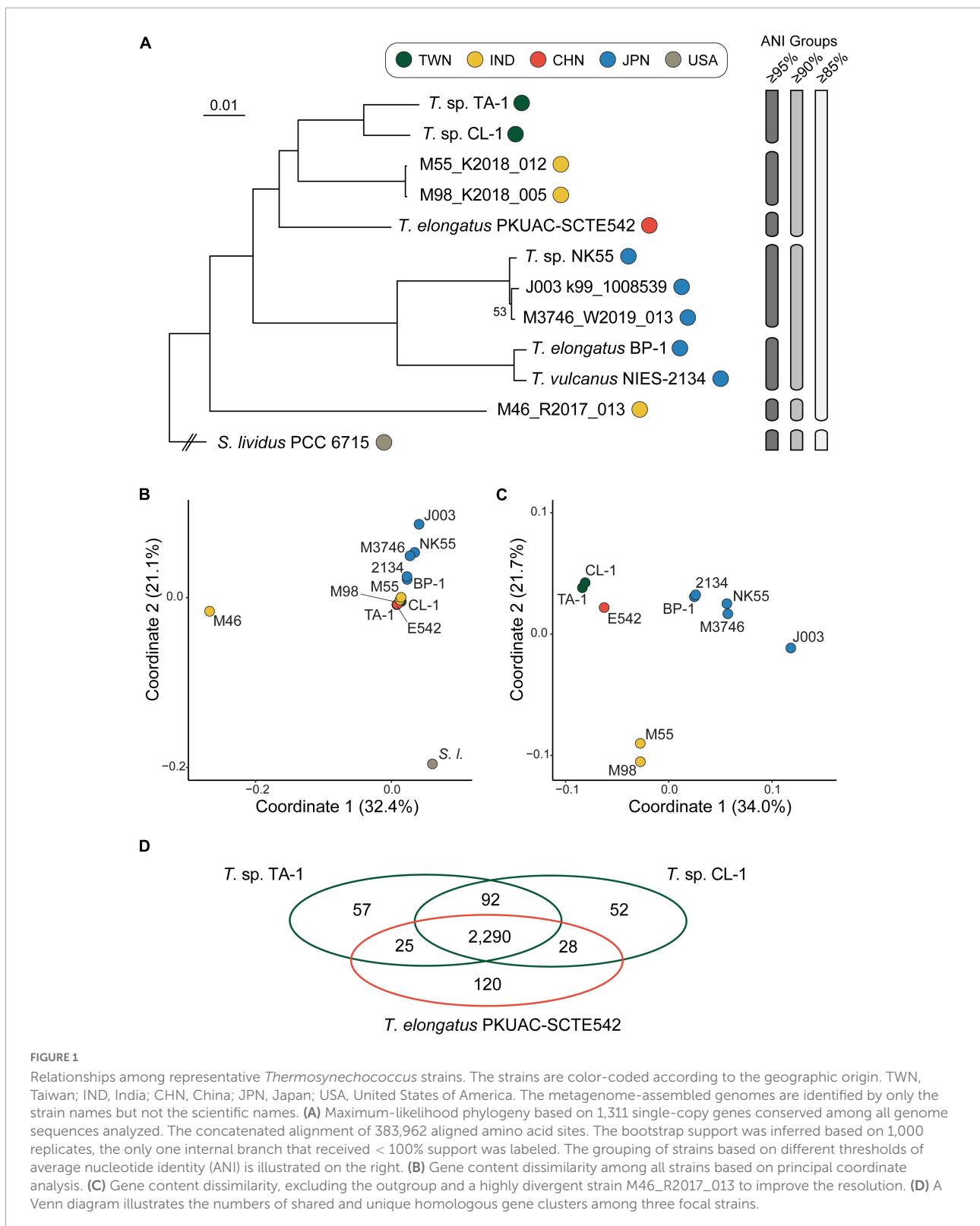


FIGURE 1

Relationships among representative *Thermosynechococcus* strains. The strains are color-coded according to the geographic origin. TWN, Taiwan; IND, India; CHN, China; JPN, Japan; USA, United States of America. The metagenome-assembled genomes are identified by only the strain names but not the scientific names. **(A)** Maximum-likelihood phylogeny based on 1,311 single-copy genes conserved among all genome sequences analyzed. The concatenated alignment of 383,962 aligned amino acid sites. The bootstrap support was inferred based on 1,000 replicates, the only one internal branch that received < 100% support was labeled. The grouping of strains based on different thresholds of average nucleotide identity (ANI) is illustrated on the right. **(B)** Gene content dissimilarity among all strains based on principal coordinate analysis. **(C)** Gene content dissimilarity, excluding the outgroup and a highly divergent strain M46_R2017_013 to improve the resolution. **(D)** A Venn diagram illustrates the numbers of shared and unique homologous gene clusters among three focal strains.

the representative strains of *T. elongatus* and *T. vulcanus*, the relationship between these two species needs to be clarified. Furthermore, with BP-1 as the type strain of *T. elongatus* and

its low level of genome similarity (i.e., ANI = 87.7%) with the Chinese strain PKUAC-SCTE542, the species assignment of PKUAC-SCTE542 needs to be re-examined. Finally, based

on the comparisons of ANI and gene content (Figure 1), it is reasonable to classify all in-group strains to the same genus (i.e., *Thermosynechococcus*), while excluding *S. lividus* PCC 6715.

Another interesting observation is that there appeared to be some links between the phylogenetic and genomic diversity of these cyanobacteria with geography (Papke et al., 2003). Specifically, all Japanese lineages form a strongly supported monophyletic clade and are relatively similar in overall genome sequence and gene content (Figure 1). However, it is difficult to distinguish between the roles of phylogenetic relatedness and geographic proximity in shaping the genetic differentiation. For future work, more extensive sampling from China and other regions are necessary to better understand the genetic diversity and phylogeography of these cyanobacteria.

For comparison among TA-1 and its two close relatives (i.e., CL-1 and PKUAC-SCTE542), a total of 2,290 HGCs were identified (Figure 1D). Consistent with the patterns observed from molecular phylogeny and ANI analysis, the two Taiwanese strains share more genes with each other than with the Chinese strain. Genes exhibiting strain- or species-specific patterns of presence/absence that may have functional significance are discussed in the following sections. The two Indian MAGs were not used in this comparison due to the concern of genome assembly quality, as well as the limitation of lacking cultured biological materials for follow-up experiments.

Intriguingly, despite the close genetic relatedness between the two Taiwanese strains, pairwise genome alignment identified at least 16 synteny breakpoints corresponding to the translocation or inversion of large chromosomal segments (Figure 2). Examination of gene content identified large numbers of putative transposase genes scattered across the entire chromosomes of these strains (i.e., 36 in TA-1 and 24 in CL-1), and nearly all of the synteny breakpoints were associated with transposase genes. The dynamic nature of the chromosomal organization in this genus was reported previously from between-species comparisons (Cheng et al., 2020). The new result from this study further establishes that chromosomal rearrangements are rampant at the within-species level and may be facilitated by mobile genetic elements.

Overview of notable genetic differentiation

Notable genetic differentiation among these strains are summarized in Figure 3. Consistent with the phylogeny, the two Taiwanese strains TA-1 and CL-1 share several derived traits. For example, these strains appear to have acquired multiple genes involved in nitric oxide protection and kanamycin resistance. In contrast, these strains seem to have lost genes involved in urea transport and chemotaxis. More detailed information regarding the inferred function, distribution patterns, and molecular evolution of these genes are provided in the following sections.

Bicarbonate transporters

Bicarbonate transporters play an essential role in the CO₂-concentrating mechanism (CCM) of cyanobacteria to actively import dissolved inorganic carbon (Ci) from aquatic environments for their photosynthetic growth (Price et al., 2008; Gupta et al., 2020). Three types of bicarbonate transporters (BicA, BCT1, and SbtA) have been identified in cyanobacteria (Liu et al., 2021). BicA is a constitutive sodium-dependent, low-affinity, high-flux transporter (Wang et al., 2019), and the corresponding gene *bicA* was conserved in all *Thermosynechococcus* strains analyzed except for one MAG (Figure 3). In comparison, BCT1 (encoded by the *cmpABCD* operon) is an inducible, medium-affinity, low-flux ABC transporter (Omata et al., 1999). The *cmpABCD* operon is present in CL-1 (locus tags FFX45_RS03210-25) and all three Japanese strains but not TA-1, PKUAC-SCTE542, or the outgroup *S. lividus*. According to the syntenic conservation (Figure 4A), this gene cluster may be present in the most recent common ancestor (MRCA) of *Thermosynechococcus*, and there were two independent losses in TA-1 and PKUAC-SCTE542. Finally, SbtA is an inducible sodium-dependent, high-affinity, low-flux bicarbonate transporter (Liu et al., 2021). The *sbtA* gene forms an operon with the neighboring *sbtB* gene. SbtB is a PII-like signaling protein that functions as a Ci sensor. The *sbtA/B* cluster is present in CL-1 (FFX45_RS07275-80), TA-1 (JW907_RS11360-65), PKUAC-SCTE542 (D3A95_RS10120-25), and two Indian MAGs, but absent in all Japanese strains (Figure 3). However, a pseudogenized *sbtB* was found in NK55 (Figure 4B), suggesting that this gene cluster might be present in the ancestor of those Japanese strains.

Of note, the *sbtA* homolog in TA-1 is pseudogenized owing to a 14-bp insertion that resulted in a premature stop codon. TA-1 was originally selected under a high CO₂ (20% CO₂) growth condition for high CO₂ mitigation (Leu et al., 2013). This growth condition may have reduced the need for a high-affinity bicarbonate transporter and relaxed the selection pressure to prevent mutation accumulation in this gene. In comparison, CL-1 was isolated directly from a high-pH and high-temperature environment in the Chin-Lun hot spring (pH 9.3, 62°C) without further selection (Cheng et al., 2020). CL-1 is the only cultivated *Thermosynechococcus* strain known to have all three types of bicarbonate transporters, which suggests that this strain may have a high capacity or flexibility for bicarbonate uptake. Consistent with this expectation, we found that CL-1 had a faster growth rate than TA-1, particularly when supplemented with 100 mM NaHCO₃ (Supplementary Figure 2).

Nitric oxide protection

Nitric oxide (NO) is an important signaling biomolecule as well as an intermediate produced in the process of

TABLE 1 List of genome sequences included in the comparative analysis.

Strain	Abbr.	Geographic origin	GenBank accession	Annotation date	Genome size (N50 of draft assembly) (bp)	G + C content (%)	Coding density (%)	No. of protein-coding genes	No. of pseudogenes	No. of rRNA Genes	No. of tRNA Genes	References
<i>T. sp.</i> TA-1	TA-1	Miaoli, Taiwan	NZ_CP070960.1	05-MAR-2021	2,658,717	53.6	90.4	2,523	33	3	42	This study
<i>T. sp.</i> CL-1	CL-1	Taitung, Taiwan	NZ_CP040671.1	25-SEP-2020	2,647,823	53.5	90.3	2,514	37	3	41	Cheng et al., 2020
<i>T. elongatus</i> PKUAC-SCTE542	E542	Sichuan, China	NZ_CP032152.1	03-AUG-2020	2,650,294	53.3	91.0	2,508	35	3	41	Liang et al., 2019
<i>T. elongatus</i> BP-1	BP-1	Beppu, Japan	NC_004113.1	13-DEC-2020	2,593,857	53.9	88.4	2,446	35	3	40	Nakamura et al., 2002
<i>T. vulcanus</i> NIES-2134	2134	Wakayama, Japan	NZ_AP018202.1	14-DEC-2020	2,571,271	53.9	88.9	2,422	36	3	40	NA
<i>T. sp.</i> NK55	NK55	Nagano, Japan	NC_023033.1	17-JAN-2021	2,520,064	53.8	87.1	2,333	78	3	41	Stolyar et al., 2014
J003 k99_1008539	J003	Tokyo, Japan	GCA_003696925.1	29-OCT-2018	2,311,962 (17,448)	53.8	85.0	2,182	148	3	35	Ward et al., 2019
M3746_W2019_013	M3746	Tokyo, Japan	GCA_015296045.1	09-NOV-2020	2,387,168 (53,495)	53.3	89.7	2,258	34	3	39	Alcorta et al., 2020
M46_R2017_013	M46	Ladakh, India	GCA_015296025.1	09-NOV-2020	2,396,365 (16,361)	52.3	72.3	2,155	324	3	41	Alcorta et al., 2020
M55_K2018_012	M55	Chhattisgarh, India	GCA_015295915.1	09-NOV-2020	2,399,209 (27,694)	53.6	90.3	2,343	36	2	39	Alcorta et al., 2020
M98_K2018_005	M98	Chhattisgarh, India	GCA_015295825.1	09-NOV-2020	2,370,950 (35,735)	53.7	90.0	2,290	43	3	39	Alcorta et al., 2020
<i>S. lividus</i> PCC 6715	<i>S. l.</i>	Yellowstone National Park, United States	NZ_CP018092.1	20-DEC-2020	2,659,739	53.5	80.1	2,383	196	3	41	Tang et al., 2022

The sampling included all cultivated strains in the genus *Thermosynechococcus* with complete genome sequences available. Additionally, metagenome-assembled genomes assigned to this genus were included; these assemblies are identified by only the strain id but not the scientific name. *Synechococcus lividus* PCC 6715 was included as an outgroup. All strains were collected from hot springs.

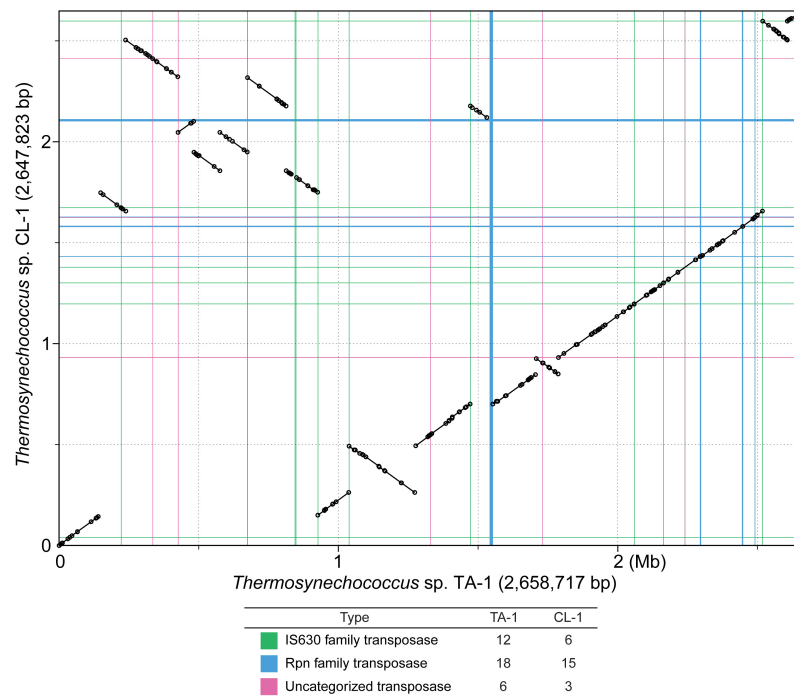


FIGURE 2

Genome alignment between two *Thermosynechococcus* strains collected in Taiwan. The circular chromosomes are plotted in linear form to illustrate syntenic regions. Positions of putative transposase genes are plotted using colored lines to illustrate the correspondence with synteny breakpoints. Counts of putative transposase genes are provided below the synteny plot. The locations and counts include intact transposase genes listed in the GenBank annotation (Table 1), as well as un-annotated fragments that were identified by TBLASTN searches (see section "materials and methods").

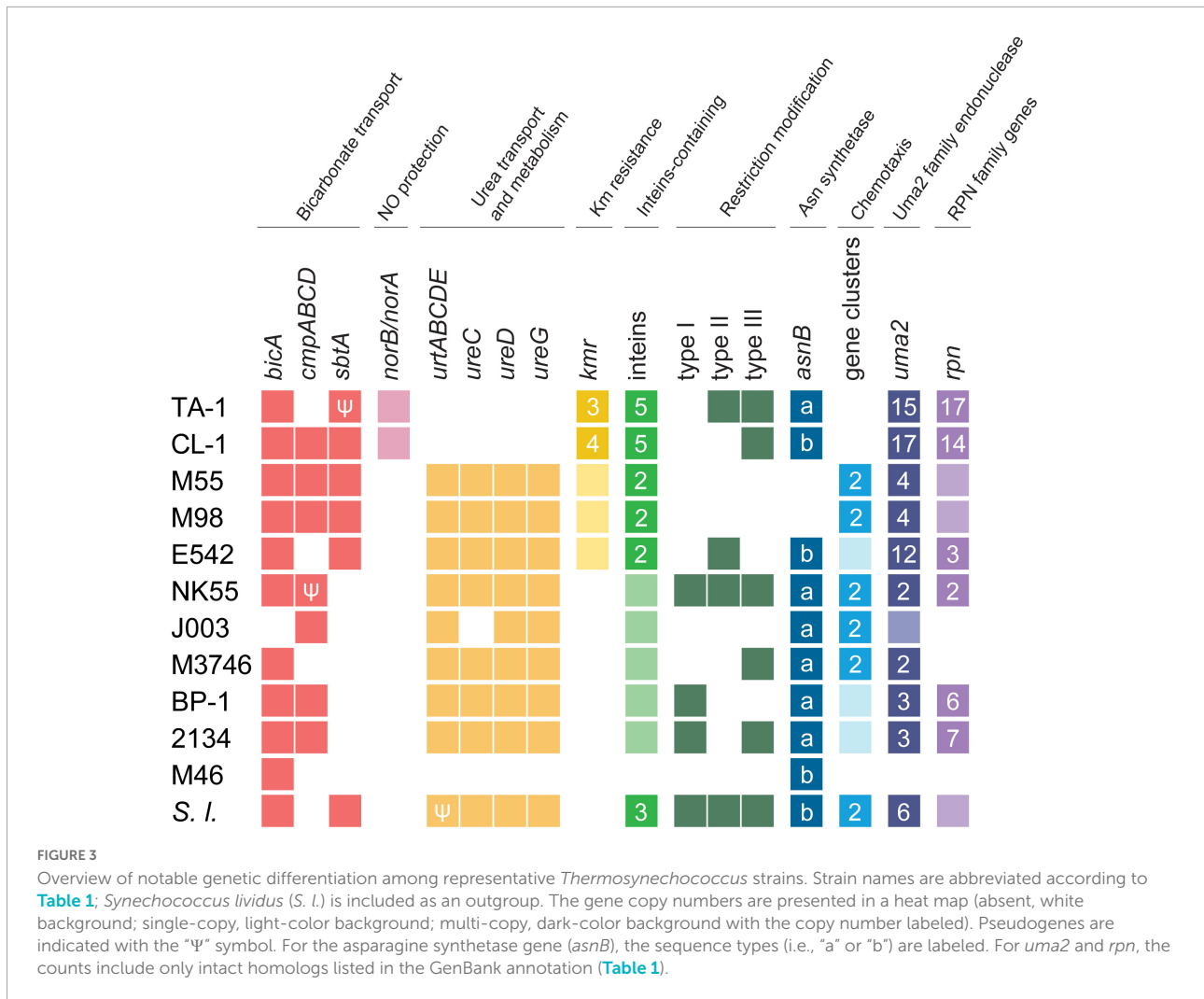
denitrification (Shiro, 2012). NO reductases catalyze the reduction of NO to N₂O, which function in nitrogen metabolism and a defense against NO toxicity (Büsch et al., 2002; Shiro, 2012). The product of NO reductase, N₂O, is a powerful green-house gas that contributes to ozone level depletion and global warming (Ravishankara et al., 2009). TA-1 and CL-1 share a distinct gene cluster for NO protection, which is absent in other *Thermosynechococcus* strains (Figure 3). This cluster contains two genes: *norA* (or *dnrN*; JW907_RS02870 in TA-1 and FFX45_RS09820 in CL-1), which encodes an iron-sulfur cluster repair di-iron protein, and *norB* (JW907_RS02875 and FFX45_RS09815), which encodes a quinol type NO reductase. According to the syntenic regions (Figure 4C), these two genes and three other adjacent genes may have been acquired in the MRCA of these two Taiwanese strains. These genes may provide defense against the toxicity of NO molecules released by other microbes in the environments of Taiwan hot springs (Büsch et al., 2002).

Urea and other nitrogen sources

TA-1 and CL-1 both lack the *urtABCDE* operon that encodes the urea ABC transport system and several urease

genes (i.e., *ureA*, *ureD*, and *ureG*) (Veaudor et al., 2019), which are conserved in other *Thermosynechococcus* strains. According to the gene distribution pattern (Figure 3) and the syntenic conservation (Figure 4D), the gene absence is best explained by gene losses in these two Taiwanese strains. Consequently, these two strains are expected to lose the ability to directly use urea as a nitrogen source. However, one previous study reported that TA-1 cells were able to grow slowly in the growth medium supplemented with 17.5 mM urea as the sole nitrogen source (Leu et al., 2013). Because urea is known to decompose slowly into cyanate and ammonia in solution at higher temperature (Hagel et al., 1971; Sun et al., 2014), the derived ammonia may be used by TA-1 cells.

In comparison, the *nir-nrtABCD-narB* operon (Flores et al., 2005), encodes a nitrite reductase, nitrate ABC transport system, and nitrate reductase, is conserved among all of the *Thermosynechococcus* genomes examined, which suggests that nitrate/nitrite are conserved nitrogen sources among these cyanobacteria. Consistent with this gene content analysis, TA-1 grew equally well in the growth medium supplemented with 5 mM nitrate or nitrite as the nitrogen source (Supplementary Figure 3). Furthermore, all strains have genes that encode ammonium transporters. No nitrogenase gene was identified in



any of these genomes, which suggests that they lack the ability for nitrogen fixation.

Kanamycin resistance

Genes involved in antibiotic production and resistance play important roles in microbial competition and survival in natural environments (Kunhikannan et al., 2021). However, our understanding of the occurrence and evolution of antibiotic resistance genes in thermophilic cyanobacteria that reside in hot-spring environments is still limited (Najar et al., 2020; Sharma et al., 2020). In this work, we found putative bipartite aminoglycoside nucleotidyltransferase gene clusters that may confer kanamycin resistance (Lehmann et al., 2003) in TA-1 (three sets: JW907_RS08580-85, JW907_RS09145-50, and JW907_RS11865-70), CL-1 (four sets; FFX45_RS04515-20, FFX45_RS05120-25, FFX45_RS07775-80, and FFX45_RS11395-400), and PKUAC-SCTE542 (one set: D3A95_RS08325-30)

but were absent in all Japanese strains (Figure 3). Each of these clusters contains two genes: one encoding a substrate binding domain and the other nucleotide-binding domains of kanamycin nucleotidyltransferases. Consistent with the expectation derived from gene content, we found that both TA-1 and CL-1 can grow in medium containing 5 μg/ml of kanamycin (Supplementary Figure 4). The gene copy number increases observed in these two Taiwanese strains may suggest a higher selective pressure for antibiotic resistance in their natural environments.

Intein-containing genes

Intein is a polypeptide segment that can splice itself spontaneously and precisely from the precursor protein and ligate its flanking polypeptide segments (Baralle and Buratti, 2017). Intein-containing genes are distributed sporadically in all domains of life, but their biological function remains elusive

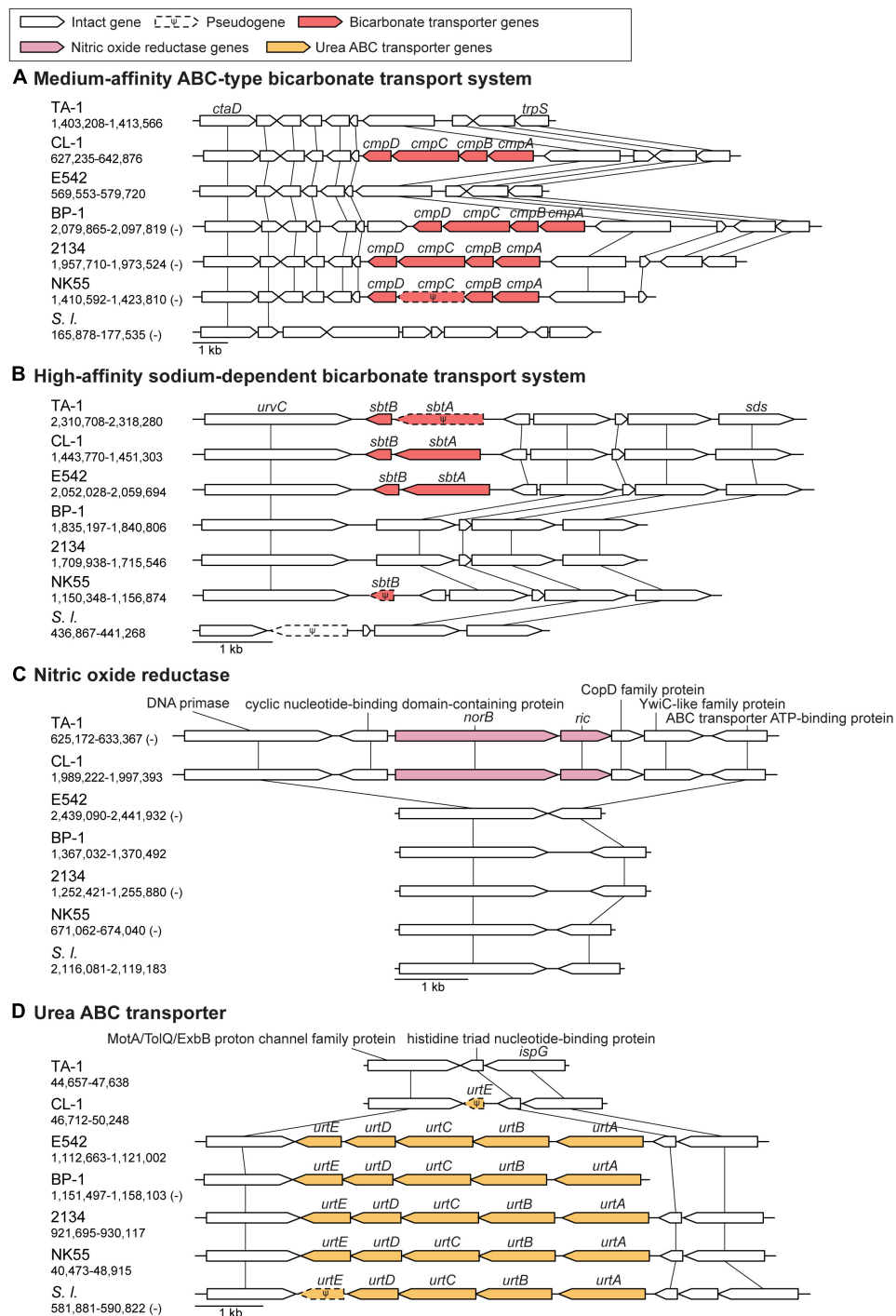
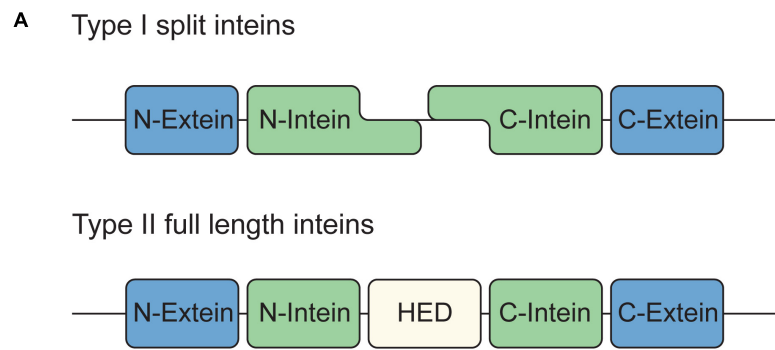


FIGURE 4

Synteny plots of notable genes. Chromosomal locations are labeled below strain name abbreviations on the left. Homologous genes are linked by vertical lines. (A) ABC-type bicarbonate transport system. (B) Sodium-dependent bicarbonate transport system. (C) Nitric oxide reductase gene cluster. (D) Urea ABC transporter gene cluster.

(Nanda et al., 2020). According to the structures, inteins can be classified into three major types (Nanda et al., 2020). Type I split inteins are expressed as two separate polypeptides and

undergo splicing in trans (Figure 5A). Type II (full length) inteins have a homing endonuclease domain (HED) embedded in between the splicing domains (Figure 5A); the HED may



B

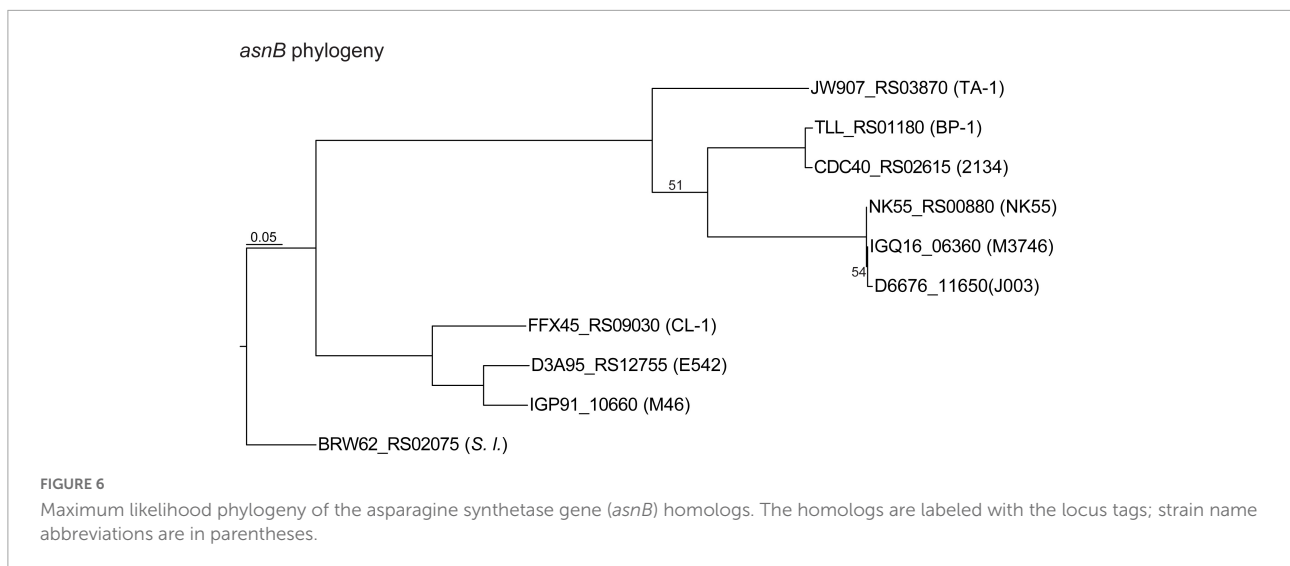
	<i>dnaE</i>	<i>nrdJ</i>	<i>gmd</i>	<i>gyrB</i>	DUF559
TA-1	I	II	II	II	II
CL-1	I	II	II	II	II
M55	I		II		
M98	I		II		
E542	I	II			
NK55	I				
J002	I				
M3746	I				
BP-1	I				
2134	I				
M46					
<i>S. I.</i>	I				

dnaE: DNA polymerase III subunit alpha DnaE-N end and DnaE-C end
nrdJ: ribonucleoside-triphosphate reductase, adenosylcobalamin-dependent with intein
gmd: GDP-D-mannose dehydratase with intein-containing DNA gyrase subunit A
gyrB: DNA gyrase subunit B with intein
DUF559: DUF559 domain-containing protein with intein

FIGURE 5
Intein-containing genes. **(A)** Structure of type I and II inteins. **(B)** Distribution of intein-containing genes in strains. Gene presence and absence are indicated as a colored and white background, respectively. For genes that contain inteins, the intein types are indicated by Roman numerals.

function in the mobilization and propagation of inteins (Nanda et al., 2020). Type III mini-inteins have a contiguous protein splicing domain but lack a HED.

Among these *Thermosynechococcus* genomes, a type I split intein is conserved in the *dnaE* homologs of all strains except for the basal Indian MAG M46_R2017_013 (Figure 5B).



This intein-containing *dnaE* is also found in a few other cyanobacteria, namely, *Synechocystis* sp. PCC6803 (Caspi et al., 2003). Notably, multiple insertion events of type II inteins likely occurred in the MRCA of the Taiwanese strains (in *nrdJ* and *gyrB*) and the MRCA of the Taiwanese/Indian strains (in *gmd*, which encodes GDP-D-mannose dehydratase) (Figure 5B). Also, a separate event of type II intein insertion may have occurred to the Chinese strain in *nrdJ*. The DUF559-containing gene belongs to the EDxHD family, which represents a novel homing endonuclease family (Taylor et al., 2011), was identified in only those two Taiwanese strains. The origin and function of these inteins are unclear, and no type III mini-intein was identified in these *Thermosynechococcus* genomes.

Prokaryotic immunity systems

There are diverse distributions of CRISPR-Cas and restriction-modification systems among these *Thermosynechococcus* strains (Figure 3) that can confer resistance to phages (Makarova et al., 2011; Oliveira et al., 2014). Type I CRISPR-Cas systems were found in TA-1 (JW907_RS04965-04995), CL-1 (FFX45_RS02320-02350), PKUAC-SCTE542 (D3A95_RS09885-09915), and NK55 (NK55_RS04045-04075). Type III-B CRISPR-Cas systems were found in TA-1 (JW907_RS00465-00485), CL-1 (FFX45_RS00470-00490), and PKUAC-SCTE542 (D3A95_RS05650-05670). BP-1 and NIES-2134 do not have any CRISPR-Cas system. Furthermore, TA-1 has one additional gene for CRISPR-associated helicase Cas3' (JW907_RS09100), which has an HD-like endonuclease domain located in the N-terminal region of Cas3. This gene is absent in all other *Thermosynechococcus* genomes.

For restriction-modification systems (Oliveira et al., 2014), genes that encode types I, II, and III systems have a patchy

distribution among the genomes compared (Figure 3). These genes are mostly absent in the MAGs analyzed. However, such gene absence may be artifacts of incomplete assemblies. Among the strains with complete genome sequences available, only the Japanese strain NK55 and the outgroup *S. lividus* PCC 6715 have all three systems. These patterns suggest that multiple gene gain/loss events occurred in the evolutionary history of *Thermosynechococcus*.

Exopolysaccharide and associated asparagine synthetase genes

Exopolysaccharides (EPSs) on the cell surfaces of cyanobacteria may act as barriers to different types of stress and play significant roles in intra- as well as inter-species interactions (Kehr and Dittmann, 2015). Most *Thermosynechococcus* genomes (except for NK55) have one uncharacterized EPS gene cluster, which includes a gene that encodes a glutamine-hydrolyzing asparagine synthase (AsnB; also known as an exosortase). AsnB may be involved in N-linked protein glycosylation of EPS molecules (Bai et al., 2008). The size and composition in the EPS gene cluster vary among these *Thermosynechococcus* strains. Particularly, the components in the EPS gene cluster of TA-1 differ from those in CL-1 and other *Thermosynechococcus* strains. Among the 17 genes in this TA-1 gene cluster, five are uniquely present in this strain and lack identifiable homologs in all other *Thermosynechococcus* strains, suggesting that horizontal gene transfer may be involved in shaping the molecular evolution of this cluster. In addition, the *Thermosynechococcus* AsnB phylogeny (Figure 6) indicated that even for a conserved gene, the homologs among different strains may be classified into two sequence types (Figure 3). Notably, the TA-1 homolog is more similar to those of the Japan strains

(71–74% amino acid sequence identity) than the more closely related CL-1 and PKUAC-SCTE542 (about 60% amino acid sequence identity). Although recombination may provide one explanation for this conflict between the gene tree (Figure 6) and species tree (Figure 1A), the exact mechanism is unclear given the apparent geographic isolation and overall genome divergence. Alternatively, the conflict may also be explained by selection or incomplete lineage sorting. It is unclear which of these hypotheses is more plausible.

Phototaxis and motility

Most *Thermosynechococcus* genomes have three or four sets of gene clusters for motility-related two-component signaling systems (Nakamura et al., 2002; Bhaya, 2004; Ishizuka et al., 2006; Cheng et al., 2020). However, TA-1 and CL-1 contain only the *pixJ*- and *pilJ*-associated two-component signaling systems that are conserved in other *Thermosynechococcus* strains except for NIES-2134. *PixJ* homolog-associated (blue-light-responsive) two-component signaling systems are truncated in NIES-2134 (Ishizuka et al., 2006). The presence of a blue-light-responsive signaling system in cyanobacteria enables them to move toward optimal light conditions and away from UV and excess light conditions (Bhaya, 2004; Conradi et al., 2020). Of note, TA-1 has a truncated *cheA* (*pilL* homolog) in a *pilJ*-associated two-component signaling system (Yoshihara et al., 2002), which may explain its non-motile phenotype (Supplementary Figure 4). Moreover, CL-1 has a truncated *pilC* of the type IV pilus machinery (Bhaya et al., 2000) and is also non-motile (Supplementary Figure 4). Finally, TA-1 and CL-1 both lack the entire gene cluster for an uncharacterized chemotaxis two-component signaling system (Wadhams and Armitage, 2004), which is well conserved in the other *Thermosynechococcus* genomes.

Other notable genetic features

Another notable genetic feature shared by TA-1 and CL-1 is apparent expansions of Uma2 (DUF820) family endonuclease and recombination promoting nuclease/putative transposase genes. The Uma2 domain-containing genes encode a family of putative restrictive endonucleases (Uma2 family endonucleases) (Kinch et al., 2005). TA-1, CL-1, and PKUAC-SCTE542 all have high copy numbers (16, 17, and 15, respectively) of Uma2 family endonuclease genes, whereas the three Japanese strains all have only four or fewer copies (Figure 3). In addition, the RPN family genes that encode recombination-promoting nuclease/putative transposases (Kingston et al., 2017) are greatly expanded in the two Taiwanese strains (Figure 3). RPN family genes in the TA-1 genome are located in three tandem gene arrays (containing 11, 3, and 3 copies, respectively). In comparison, the

outgroup *S. lividus* has only one copy of the RPN family gene. Furthermore, RPN family genes are absent in genomes of JA-3-3-Ab and JA-2-3Ba (*Synechococcus* phylogenetic A/B lineage) from Yellowstone National Park (Bhaya et al., 2007). Taken together, these findings indicate the expansion of the *uma2* gene family in the MRCA of the Chinese/Taiwanese lineages and further expansion of the *rpn* gene family in the MRCA of the two Taiwanese strains.

Conclusion

In this work, we reported the complete genome sequence of a thermophilic cyanobacterium (i.e., strain TA-1) and conducted comparative analysis with other representatives within the genus *Thermosynechococcus*. Compared with a previous work that examined the genetic differentiation of *Thermosynechococcus* in the context of environmental factors and found no clear link (Prondzinsky et al., 2021), we found that the genetic differentiation of *Thermosynechococcus* are mostly congruent with phylogeny. Moreover, the two strains collected in Taiwan represent a novel species-level taxon in this genus, and this taxon has several distinct genetic features (e.g., antibiotic resistance, EPS biosynthesis, and motility). However, despite the high genetic relatedness between these two strains, we observed extensive genome rearrangements and multiple differentiation of gene content, which indicates the dynamic nature of the genome evolution of these bacteria. Taken together, these findings improve our understanding of these cyanobacteria, which are the primary producers in non-acidic hot springs, and provide genomic resources for future functional characterization that can benefit biotechnology applications.

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/genbank/>, CP070960 and <https://www.ncbi.nlm.nih.gov/>, PRJNA702872.

Author contributions

C-HK and H-AC: conceptualization, funding acquisition, project administration, and supervision. Y-IC and Y-CL: investigation, validation, and visualization. J-YL: biological materials. Y-IC, Y-CL, C-HK, and H-AC: writing – original draft. Y-IC, Y-CL, J-YL, C-HK, and H-AC: writing—review and editing. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Ministry of Science and Technology of Taiwan (MOST 110-2311-B-001-017) to H-AC and Academia Sinica to C-HK and H-AC. The funders had no role in study design, data collection, and interpretation, or the decision to submit the work for publication.

Acknowledgments

The Illumina and Oxford Nanopore sequencing library preparation services were provided by the Genomic Technology Core (Institute of Plant and Microbial Biology, Academia Sinica). The Illumina HiSeq sequencing service was provided by Yourgene Health (New Taipei City, Taiwan). The English editing service was provided by Bio Med Editing. We thank Dr. Wen-Dar Lin in the Bioinformatics Core and Dr. Wann-Neng Jane in the Cell Biology Core (Institute of Plant and Microbial Biology, Academia Sinica) for technical assistance.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Alcorta, J., Alarcón-Schumacher, T., Salgado, O., and Díez, B. (2020). Taxonomic novelty and distinctive genomic features of hot spring cyanobacteria. *Front. Genet.* 11:568223. doi: 10.3389/fgene.2020.568223
- Bai, L., Jiang, R., Shan, J., Guo, L., Zhang, Y., Zhang, R., et al. (2008). Purification, characterization and functional analysis of asparagine synthetase encoding by *ste10* gene in Ebosin biosynthesis of *Streptomyces* sp. 139. *Enzyme Microb. Technol.* 42, 548–553.
- Baralle, M., and Buratti, E. (2017). *Protein splicing*. *Brenner's Encyclopedia of Genetics (Second Edition)*, (Amsterdam: Elsevier), 492–494. doi: 10.1016/B978-0-12-374984-0.01226-2
- Becraft, E. D., Wood, J. M., Rusch, D. B., Kühl, M., Jensen, S. I., Bryant, D. A., et al. (2015). The molecular dimension of microbial species: 1. Ecological distinctions among, and homogeneity within, putative ecotypes of *Synechococcus* inhabiting the cyanobacterial mat of Mushroom Spring, Yellowstone National Park. *Front. Microbiol.* 6:590. doi: 10.3389/fmicb.2015.00590
- Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Ostell, J., Pruitt, K. D., et al. (2018). GenBank. *Nucleic Acids Res.* 46:D41–D47. doi: 10.1093/nar/gkx1094
- Bhaya, D. (2004). Light matters: Phototaxis and signal transduction in unicellular cyanobacteria. *Mol. Microbiol.* 53, 745–754. doi: 10.1111/j.1365-2958.2004.04160.x
- Bhaya, D., Bianco, N. R., Bryant, D., and Grossman, A. (2000). Type IV pilus biogenesis and motility in the cyanobacterium *Synechocystis* sp. PCC6803. *Mol. Microbiol.* 37, 941–951. doi: 10.1046/j.1365-2958.2000.02068.x
- Bhaya, D., Grossman, A. R., Steunou, A. S., Khuri, N., Cohan, F. M., Hamamura, N., et al. (2007). Population level functional diversity in a microbial community revealed by comparative genomic and metagenomic analyses. *ISME J.* 1, 703–713. doi: 10.1038/ismej.2007.46
- Büsch, A., Friedrich, B., and Cramm, R. (2002). Characterization of the *norB* gene, encoding nitric oxide reductase, in the nondenitrifying cyanobacterium *Synechocystis* sp. strain PCC6803. *Appl. Environ. Microbiol.* 68, 668–672. doi: 10.1128/AEM.68.2.668-672.2002
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: Architecture and applications. *BMC Bioinform.* 10:421. doi: 10.1186/1471-2105-10-421
- Caspi, J., Amitai, G., Belenkiy, O., and Pietrokovski, S. (2003). Distribution of split DnaE inteins in cyanobacteria. *Mol. Microbiol.* 50, 1569–1577. doi: 10.1046/j.1365-2958.2003.03825.x
- Chen, M.-Y., Teng, W.-K., Zhao, L., Hu, C.-X., Zhou, Y.-K., Han, B.-P., et al. (2021). Comparative genomics reveals insights into cyanobacterial evolution and habitat adaptation. *ISME J.* 15, 211–227. doi: 10.1038/s41396-020-00775-z
- Cheng, Y.-I., Chou, L., Chiu, Y.-F., Hsueh, H.-T., Kuo, C.-H., and Chu, H.-A. (2020). Comparative genomic analysis of a novel strain of Taiwan hot-spring

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.932840/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Morphology of *Thermosynechococcus* sp. TA-1 cells. (A) Light microscopy. (B) Scanning electron microscopy. (C) Transmission electron microscopy, longitudinal section. (D) Transmission electron microscopy, partial enlargement of the longitudinal section. (E) Transmission electron microscopy, transverse section.

SUPPLEMENTARY FIGURE 2

Growth of *Thermosynechococcus* sp. CL-1 and TA-1 in BG-11 medium with or without bicarbonate. Error bars are plotted but not discernible at this scale.

SUPPLEMENTARY FIGURE 3

Growth of *Thermosynechococcus* sp. TA-1 in BG-11 medium with nitrate or nitrite as the nitrogen source. Error bars are plotted but not discernible at this scale.

SUPPLEMENTARY FIGURE 4

Growth of *Thermosynechococcus* sp. (A) TA-1 and (B) CL-1 on solid agar plates containing 5 µg/ml of kanamycin under directional white light.

- cyanobacterium Thermosynechococcus* sp. CL-1. *Front. Microbiol.* 11:82. doi: 10.3389/fmicb.2020.00082
- Cho, S.-T., Lin, C.-P., and Kuo, C.-H. (2019). Genomic characterization of the periwinkle leaf yellowing (PLY) phytoplasmas in Taiwan. *Front. Microbiol.* 10:2194. doi: 10.3389/fmicb.2019.02194
- Conradi, F. D., Mullineaux, C. W., and Wilde, A. (2020). The role of the cyanobacterial type IV pilus machinery in finding and maintaining a favourable environment. *Life* 10:252. doi: 10.3390/life10110252
- Dyer, D. L., and Gafford, R. D. (1961). Some characteristics of a thermophilic blue-green alga. *Science* 134, 616–617. doi: 10.1126/science.134.3479.616
- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Everroad, R. C., Otaki, H., Matsuura, K., and Haruta, S. (2012). Diversification of bacterial community composition along a temperature gradient at a thermal spring. *Microbes Environ.* 27, 374–381. doi: 10.1264/jsm.2.me11350
- Flores, E., Frias, J. E., Rubio, L. M., and Herrero, A. (2005). Photosynthetic nitrate assimilation in cyanobacteria. *Photosynth. Res.* 83, 117–133. doi: 10.1007/s11120-004-5830-9
- Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704. doi: 10.1080/10635150390235520
- Gupta, J. K., Rai, P., Jain, K. K., and Srivastava, S. (2020). Overexpression of bicarbonate transporters in the marine cyanobacterium *Synechococcus* sp. PCC 7002 increases growth rate and glycogen accumulation. *Biotechnol. Biofuels* 13:17. doi: 10.1186/s13068-020-1656-8
- Guy, L., Kultima, J. R., and Andersson, S. G. (2010). genoPlotR: Comparative gene and genome visualization in R. *Bioinformatics* 26, 2334–2335. doi: 10.1093/bioinformatics/btq413
- Hagel, P., Gerding, J. J. T., Fieggen, W., and Bloemendal, H. (1971). Cyanate formation in solutions of urea: I. Calculation of cyanate concentrations at different temperature and pH. *Biochim. Biophys. Acta Protein Structure* 243, 366–373. doi: 10.1016/0005-2795(71)90003-1
- Hsueh, H. T., Chu, H., and Chang, C. C. (2007). Identification and characteristics of a cyanobacterium isolated from a hot spring with dissolved inorganic carbon. *Environ. Sci. Technol.* 41, 1909–1914. doi: 10.1021/es0620639
- Ishizuka, T., Shimada, T., Okajima, K., Yoshihara, S., Ochiai, Y., Katayama, M., et al. (2006). Characterization of cyanobacteriochrome TePixJ from a thermophilic cyanobacterium *Thermosynechococcus elongatus* strain BP-1. *Plant Cell Physiol.* 47, 1251–1261. doi: 10.1093/pcp/pcj095
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9
- Kehr, J.-C., and Dittmann, E. (2015). Biosynthesis and function of extracellular glycans in cyanobacteria. *Life* 5, 164–180. doi: 10.3390/life5010164
- Kinch, L. N., Ginalski, K., Rychlewski, L., and Grishin, N. V. (2005). Identification of novel restriction endonuclease-like fold families among hypothetical proteins. *Nucleic Acids Res.* 33, 3598–3605. doi: 10.1093/nar/gki676
- Kingston, A. W., Ponkratz, C., and Raleigh, E. A. (2017). Rpn (YhgA-Like) Proteins of *Escherichia coli* K-12 and their contribution to RecA-independent horizontal transfer. *J. Bacteriol.* 199:e00787–16. doi: 10.1128/jb.00787-16
- Komarek, J., Johansen, J., Smarda, J., and Strunecký, O. (2020). Phylogeny and taxonomy of *Synechococcus*-like cyanobacteria. *Fottea* 20, 171–191. doi: 10.5507/fof.2020.006
- Kunhikannan, S., Thomas, C. J., Franks, A. E., Mahadevaiah, S., Kumar, S., and Petrovski, S. (2021). Environmental hotspots for antibiotic resistance genes. *Microbiologyopen* 10:e1197. doi: 10.1002/mbo3.1197
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12. doi: 10.1186/gb-2004-5-2-r12
- Larsson, J., Nylander, J. A., and Bergman, B. (2011). Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. *BMC Evol. Biol.* 11:187. doi: 10.1186/1471-2148-11-187
- Lehmann, C., Lim, K., Chalamesetty, V. R., Krajewski, W., Melamud, E., Galkin, A., et al. (2003). The HI0073/HI0074 protein pair from *Haemophilus influenzae* is a member of a new nucleotidyltransferase family: Structure, sequence analyses, and solution studies. *Proteins* 50, 249–260. doi: 10.1002/prot.10260
- Leu, J.-Y., Lin, T.-H., Selvamani, J., Chen, H.-C., Liang, J.-Z., and Pan, K.-M. (2013). Characterization of a novel thermophilic cyanobacterial strain from Taian hot springs in Taiwan for high CO₂ mitigation and C-phycoerythrin extraction. *Process Biochem.* 48, 41–48. doi: 10.1016/j.procbio.2012.09.019
- Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, L., Stoeckert, C. J. Jr., and Roos, D. S. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13, 2178–2189. doi: 10.1101/gr.1224503
- Liang, Y., Tang, J., Luo, Y., Kaczmarek, M. B., Li, X., and Daroch, M. (2019). *Thermosynechococcus* as a thermophilic photosynthetic microbial cell factory for CO(2) utilisation. *Bioresour. Technol.* 278, 255–265. doi: 10.1016/j.biortech.2019.01.089
- Liao, C.-C., Liu, S.-L., and Wang, W.-L. (2006). Effects of temperature and pH on growth and photosynthesis of the thermophilic cyanobacterium *Synechococcus lividus* as measured by pulse-amplitude modulated fluorometry. *Psychol. Res.* 54, 260–268. doi: 10.1111/j.1440-1835.2006.00432.x
- Liu, X. Y., Hou, W. T., Wang, L., Li, B., Chen, Y., Chen, Y., et al. (2021). Structures of cyanobacterial bicarbonate transporter SbtA and its complex with PII-like SbtB. *Cell Discov.* 7:63. doi: 10.1038/s41421-021-00287-w
- Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., et al. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477. doi: 10.1038/nrmicro2577
- Matos, M., Simões, P., Fragata, I., Quina, A. S., Kristensen, T. N., and Santos, M. (2021). Coping with climate change: A genomic perspective on thermal adaptation. *Front. Genet.* 11:619441. doi: 10.3389/fgene.2020.619441
- Najar, I. N., Sherpa, M. T., Das, S., Das, S., and Thakur, N. (2020). Diversity analysis and metagenomic insights into antibiotic and metal resistance among Himalayan hot spring bacteriobiome insinuating inherent environmental baseline levels of antibiotic and metal tolerance. *J. Glob. Antimicrob. Resist.* 21, 342–352. doi: 10.1016/j.jgar.2020.03.026
- Nakamura, Y., Kaneko, T., Sato, S., Ikeuchi, M., Katoh, H., Sasamoto, S., et al. (2002). Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1. *DNA Res.* 9, 123–130. doi: 10.1093/dnares/9.4.123
- Nanda, A., Nasker, S. S., Mehra, A., Panda, S., and Nayak, S. (2020). Inteins in science: Evolution to application. *Microorganisms* 8:2004. doi: 10.3390/microorganisms8122004
- Oliveira, P. H., Touchon, M., and Rocha, E. P. (2014). The interplay of restriction-modification systems with mobile genetic elements and their prokaryotic hosts. *Nucleic Acids Res.* 42, 10618–10631. doi: 10.1093/nar/gku734
- Olsen, M. T., Nowack, S., Wood, J. M., Becraft, E. D., LaButti, K., Lipzen, A., et al. (2015). The molecular dimension of microbial species: 3. Comparative genomics of *Synechococcus* strains with different light responses and in situ diel transcription patterns of associated putative ecotypes in the Mushroom Spring microbial mat. *Front. Microbiol.* 6:604. doi: 10.3389/fmicb.2015.00604
- Omata, T., Price, G. D., Badger, M. R., Okamura, M., Gohta, S., and Ogawa, T. (1999). Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *Proc. Natl. Acad. Sci. U.S.A.* 96, 13571–13576. doi: 10.1073/pnas.96.23.13571
- Papke, R. T., Ramsing, N. B., Bateson, M. M., and Ward, D. M. (2003). Geographical isolation in hot spring cyanobacteria. *Environ. Microbiol.* 5, 650–659. doi: 10.1046/j.1462-2920.2003.00460.x
- Patel, A., Matsakas, L., Rova, U., and Christakopoulos, P. (2019). A perspective on biotechnological applications of thermophilic microalgae and cyanobacteria. *Bioresour. Technol.* 278, 424–434. doi: 10.1016/j.biortech.2019.01.063
- Price, G. D., Badger, M. R., Woodger, F. J., and Long, B. M. (2008). Advances in understanding the cyanobacterial CO₂-concentrating-mechanism (CCM): Functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. *J. Exp. Bot.* 59, 1441–1461. doi: 10.1093/jxb/erm112
- Prondzinsky, P., Berkemer, S. J., Ward, L. M., and McGlynn, S. E. (2021). The *Thermosynechococcus* genus: Wide environmental distribution, but a highly conserved genomic core. *Microbes Environ.* 36:ME20138. doi: 10.1264/jsm.2021.38
- Ravishankara, A. R., Daniel, J., and Portmann, R. (2009). Nitrous oxide (N₂O): The dominant ozone-depleting substance emitted in the 21st century. *Science* 326, 123–125. doi: 10.1126/science.1176985

- Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat Biotechnol.* 29, 24–26. doi: 10.1038/nbt.1754
- Sharma, N., Kumar, J., Abedin, M. M., Sahoo, D., Pandey, A., Rai, A. K., et al. (2020). Metagenomics revealing molecular profiling of community structure and metabolic pathways in natural hot springs of the Sikkim Himalaya. *BMC Microbiol.* 20:246. doi: 10.1186/s12866-020-01923-3
- Shiro, Y. (2012). Structure and function of bacterial nitric oxide reductases: Nitric oxide reductase, anaerobic enzymes. *Biochim. Biophys. Acta Bioenerg.* 1817, 1907–1913. doi: 10.1016/j.bbabi.2012.03.001
- Stolyar, S., Liu, Z., Thiel, V., Tomsho, L. P., Pineda, N., Nelson, W. C., et al. (2014). Genome sequence of the thermophilic cyanobacterium *Thermosynechococcus* sp. strain NK55a. *Genome Announc.* 2:e01060–13. doi: 10.1128/genomeA.01060-13
- Su, C. M., Hsueh, H. T., Li, T. Y., Huang, L. C., Chu, Y. L., Tseng, C. M., et al. (2013). Effects of light availability on the biomass production, CO₂ fixation, and bioethanol production potential of *Thermosynechococcus* CL-1. *Bioresour. Technol.* 145, 162–165. doi: 10.1016/j.biortech.2013.02.092
- Sun, S., Zhou, J. Y., Yang, W., and Zhang, H. (2014). Inhibition of protein carbamylation in urea solution using ammonium-containing buffers. *Anal. Biochem.* 446, 76–81. doi: 10.1016/j.ab.2013.10.024
- Tang, J., Jiang, D., Luo, Y., Liang, Y., Li, L., Shah, M. M. R., et al. (2018). Potential new genera of cyanobacterial strains isolated from thermal springs of western Sichuan, China. *Algal Res.* 31, 14–20. doi: 10.1016/j.algal.2018.01.008
- Tang, J., Zhou, H., Yao, D., Riaz, S., You, D., Klepacz-Smółka, A., et al. (2022). Comparative genomic analysis revealed distinct molecular components and organization of CO₂-concentrating mechanism in thermophilic cyanobacteria. *Front. Microbiol.* 13:876272. doi: 10.3389/fmicb.2022.876272
- Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., et al. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 44, 6614–6624. doi: 10.1093/nar/gkw569
- Taylor, G. K., Heiter, D. F., Pietrovskiy, S., and Stoddard, B. L. (2011). Activity, specificity and structure of I-Bth0305I: A representative of a new homing endonuclease family. *Nucleic Acids Res.* 39, 9705–9719. doi: 10.1093/nar/gkr669
- Veaudor, T., Cassier-Chauvat, C., and Chauvat, F. (2019). Genomics of urea transport and catabolism in cyanobacteria: Biotechnological implications. *Front. Microbiol.* 10:2052. doi: 10.3389/fmicb.2019.02052
- Wadhams, G. H., and Armitage, J. P. (2004). Making sense of it all: Bacterial chemotaxis. *Nat. Rev. Mol. Cell Biol.* 5, 1024–1037. doi: 10.1038/nrm1524
- Wang, C., Sun, B., Zhang, X., Huang, X., Zhang, M., Guo, H., et al. (2019). Structural mechanism of the active bicarbonate transporter from cyanobacteria. *Nat. Plants* 5, 1184–1193. doi: 10.1038/s41477-019-0538-1
- Ward, D. M., Castenholz, R. W., and Miller, S. R. (2012). “Cyanobacteria in geothermal habitats,” in *Ecology of Cyanobacteria II: Their Diversity in Space and Time*, ed. B. A. Whitton (Dordrecht: Springer), 39–63. doi: 10.1007/978-94-007-3855-3_3
- Ward, L. M., Idei, A., Nakagawa, M., Ueno, Y., Fischer, W. W., and McGlynn, S. E. (2019). Geochemical and metagenomic characterization of Jinata Onsen, a Proterozoic-analog hot spring, reveals novel microbial diversity including iron-tolerant phototrophs and thermophilic lithotrophs. *Microbes Environ.* 34, 278–292. doi: 10.1264/jsme2.ME19017
- Weng, L.-W., Lin, Y.-C., Su, C.-C., Huang, C.-T., Cho, S.-T., Chen, A.-P., et al. (2021). Complete genome sequence of *Xylella taiwanensis* and comparative analysis of virulence gene content with *Xylella fastidiosa*. *Front. Microbiol.* 12:684092. doi: 10.3389/fmicb.2021.684092
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13:e1005595. doi: 10.1371/journal.pcbi.1005595
- Wickham, H. (2009). *Ggplot2*. New York, NY: Springer, doi: 10.1007/978-0-38798141-3
- Yoshihara, S., Geng, X., and Ikeuchi, M. (2002). pilG gene cluster and split pilL genes involved in pilus biogenesis, motility and genetic transformation in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol.* 43, 513–521. doi: 10.1093/pcp/pcf061
- Zuliani, L., Serpico, A., De Simone, M., Frison, N., and Fusco, S. (2021). Biorefinery gets hot: Thermophilic enzymes and microorganisms for second-generation bioethanol production. *Processes* 9:1583. doi: 10.3390/pr9091583