The Order of Trait Emergence in the Evolution of Cyanobacterial Multicellularity

Katrin Hammerschmidt [1],*, Giddy Landan [1],†, Fernando Domingues Kümmel Tria[1],†, Jaime Alcorta[2], and Tal Dagan[1]

[1]Genomic Microbiology Group, Institute of Microbiology, Kiel University, Germany

[2]Department of Molecular Genetics and Microbiology, Biological Sciences Faculty, Pontifical Catholic University of Chile, Santiago, Chile

[3]Present address: Institute for Molecular Evolution, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

†These authors contributed equally to this work.

*Corresponding author: E-mail: khammerschmidt@ifam.uni-kiel.de.

Abstract

The transition from unicellular to multicellular organisms is one of the most significant events in the history of life. Key to this process is the emergence of Darwinian individuality at the higher level: Groups must become single entities capable of reproduction for selection to shape their evolution. Evolutionary transitions in individuality are characterized by cooperation between the lower level entities and by division of labor. Theory suggests that division of labor may drive the transition to multicellularity by eliminating the trade off between two incompatible processes that cannot be performed simultaneously in one cell. Here, we examine the evolution of the most ancient multicellular transition known today, that of cyanobacteria, where we reconstruct the sequence of ecological and phenotypic trait evolution. Our results show that the prime driver of multicellularity in cyanobacteria was the expansion in metabolic capacity offered by nitrogen fixation, which was accompanied by the emergence of the filamentous morphology and succeeded by a reproductive life cycle. This was followed by the progression of multicellularity into higher complexity in the form of differentiated cells and patterned multicellularity.

Key words: N_2 fixation, division of labor, filament, complexity, transition in individuality.

Significance

The emergence of multicellularity is a major evolutionary transition. The oldest transition, that of cyanobacteria, happened more than 3–3.5 Ga. We find N_2 fixation to be the prime driver of multicellularity in cyanobacteria. This innovation faced the challenge of incompatible metabolic processes because the N_2 fixing enzyme (nitrogenase) is sensitive to oxygen, which is abundantly found in cyanobacteria cells performing photosynthesis. At the same time, N_2 fixation conferred an adaptive benefit to the filamentous morphology as cells could divide their labor into performing either N_2 fixation or photosynthesis. This was followed by the culmination of complex multicellularity in the form of differentiated cells and patterned multicellularity.

Introduction

Multicellularity is considered a characteristic trait of eukaryotes, but has evolved independently several times in diverse prokaryote taxa, including actinobacteria, myxobacteria, and cyanobacteria (Bonner 1998). Bacterial multicellularity ranges from transient associations, such as colonies, biofilms, and cellular aggregations, to permanent multicellular forms (Shapiro 1988).

Instances of multicellular bacterial species present the major traits of eukaryotic multicellularity, including cell-to-cell

adhesion, peri or cytoplasmic continuity, intercellular communication, patterning, programmed cell death (PCD), and division of labor (Claessen et al. 2014). Aggregative forms of multicellularity are common among bacterial species, for example, those that form a biofilm under specific external conditions (Tarnita et al. 2013). *Bacillus subtilis*, for instance, forms biofilms upon nutrient deprivation in which cells differentiate into motile, matrix producing, or spore cells depending on the environmental cues (Claessen et al. 2014). Notably, cell differentiation in aggregates is adaptive at the level of the individual cell as it directly confers a fitness benefit to that particular cell. In contrast, under true division of labor, cells are interdependent upon each other and specialize in performing complementary tasks. These tasks, for example, somatic functions or PCD, are not beneficial at the level of the individual cell, but are advantageous for the colony; thus, they are emergent properties at a higher level of organization (van Gestel et al. 2015).

True division of labor in bacteria is best described in actinobacteria and cyanobacteria (van Gestel et al. 2015). In cyanobacteria, the most complex of the filamentous species can differentiate up to five different cell types: vegetative (photosynthetic) cells, akinetes (spore-like cells), hormogonia (reproductive, motile filaments), necridia (dead cells resulting from PCD/apoptosis for hormogonia release), and heterocysts (Claessen et al. 2014; Herrero et al. 2016). Heterocysts differentiate under nitrogen deprivation and are specialized in nitrogen (N_2) fixation by the enzyme nitrogenase (Frías et al. 1994). As this enzyme is sensitive to oxygen (O_2), these cells are characterized by the absence of oxygenic photosynthesis and by a thick cell wall, which maintains an anaerobic environment. Heterocysts and vegetative cells in the filament are metabolically interdependent with the heterocysts providing combined nitrogen to the other cells within the filament and receiving fixed carbon compounds in return. Heterocysts cannot reproduce hence they represent a prime example for emergent traits on the level of a multicellular organism.

Cyanobacteria possess the hallmark traits reminiscent of complex eukaryotic multicellularity, making the order of trait emergence essential for understanding the origin of higher-level complexity in organismal evolution. Here, we infer the evolutionary trajectory of the emergence of traits in the evolution of multicellularity in cyanobacteria.

Materials and Methods

Data

The primary data underlying this study consist of the genomic sequences and phenotypic traits of 199 representative cyanobacterial species. These were selected from the available genomes so that the number of represented taxa will be as large as possible and genus-level redundancy will be reduced (see supplementary table S1, Supplementary Material online, for the complete list of species).

Phenotypic Traits

Phenotypic traits were chosen for their potential relevance to the evolution of multicellularity in cyanobacteria, such as environmental factors that might facilitate multicellularity and markers that are indicative for the transition to multicellularity (table 1).

Information on presence and absence of traits was obtained from the published literature and from the Pasteur Culture Collection of cyanobacteria, extending the work by Uyeda et al. 2016, and coded as binary trait states. Traits included morphology (unicellular, filamentous), nitrogen fixation (no N_2 fixation, N_2 fixation), habitat (marine/non marine), baeocytes, hormogonia, thermophilic, akinetes, heterocysts, true branching, epiphytic, microbial mats, attached/unattached, sheath, mucilage, gas vesicles, motility, and multiplane fission (table 1, supplementary table S1, Supplementary Material online).

Protein Families and Alignments

The cyanobacteria protein families were constructed from completely sequenced genomes available in RefSeq database (O'Leary et al. 2016; ver. May 2016). For the construction of protein families, at the first stage, all protein sequences annotated in the genomes were blasted all-against-all using stand-alone BLAST (Altschul et al. 1990) ver. 2.2.26. Protein sequence pairs that were found as reciprocal best BLAST hits (rBBHs) (Tatusov et al. 1997) with a threshold of E-value $\leq 1 \times 10^{-5}$ were further compared by global alignment using needle (Rice et al. 2000). Sequence pairs having $\geq 30\%$ identical amino acids were clustered into protein families using the Markov clustering algorithm (MCL) (Enright et al. 2002) ver. 12-135 with the default parameters. Multiple-copy gene families were discarded, resulting in an initial data set of 18,873 single-copy gene families.

Gene families were then extended to include homologous sequences from noncyanobacteria species, serving as outgroups for rooting purposes. We identified outgroup homologues by an rBBH analysis of the *Scytonema hofmanni* PCC 7110 genome (the most widely present species in the initial gene family data set) against 26 high quality noncyanobacteria genomes: Vampirovibrionia (12 genomes) and Sericytochromatia (2) (Soo et al. 2017; Carnevali 2019), the closest phyla Margulisbacteria (6), Saganbacteria (2), Fusobacteria (1), and Firmicutes (1) (according to Carnevali 2019; Zhu et al. 2019); one reference anoxygenic photosynthetic genome from *Chloroflexus aurantiacus* J-10-fl and the *Escherichia coli* str. K-12 substr. MG1655 genome (see supplementary table S1, Supplementary Material online). The number of gene families with homologs ranged between 204 and 451 for the 26 outgroup genomes. We selected six

Table 1
Description of Cyanobacterial Cell Types, Morphological and Physiological Traits, Their Habitat and Life Style

Cell types	
Vegetative cells	Photosynthetic cells
HORMOGONIA[a]	Motile reproductive filaments that result from repeated rounds of fission without intermittent growth phases. They break off the mother filament, ensuring the reproduction and dispersal of benthic species
HETEROCYSTS[a]	Thick-walled cells that are specialized in fixing N_2
AKINETES[a]	Thick-walled, spore-like cells that provide reproduction, dormancy, and resilience
BAEOCYTES	Reproductive cells that result from repeated rounds of fission without intermittent growth phases
Morphological and physiological traits	
UNICELLULAR	Single-celled morphology. After cell division cells separate
FILAMENTOUS[a]	Multi-celled morphology. Cells remain attached after cell division
NO N FIXATION	No Fixation of N_2 into ammonia, in contrast to:
N FIXATION	Fixation of N_2 into ammonia
SHEATH	Part of the cell envelope, located outside the cell wall
MUCILAGE	Part of the envelope, located outside the cell wall, comprised extracellular polymeric substances (EPS), without a defined structure
GAS VESICLES[a]	Intracellular gas-filled chambers for regulating buoyancy in the water column
MOTILITY	Movement across surfaces or through a liquid medium
MULTIPLANE FISSION	Cell division in two or three perpendicular planes
TRUE BRANCHING[a]	Fission in multiple planes leads to branching filaments that remain attached to the main filament
Habitat and life style	
NON MARINE	Aquatic environments with salinity between 0 and 0.5 ppt, and terrestrial habitats
MARINE	Environments with salinity between 30 and 50 ppt
THERMOPHILIC	Optimal growth temperature above 45 °C
MICROBIAL MATS	Growth inside thick, laminated, microbial structures
UNATTACHED	Organism that lives unattached to a substrate (planktonic), in contrast to:
ATTACHED	Organism that lives attached to a substrate (sessile/benthic)
EPIPHYTIC	Growth on plants

NOTE.—Small caps indicate the traits that have been used in the analysis.
[a]Multicellularity markers: traits that are adaptations on the level of the filament.

of these outgroups for further analyses: *Vampirovibrio chlorellavorus*, *Chloroflexales*, *Obscuri-PALSA-1081*, *Sericytochromatia-UBA7694*, *B. subtilis*, and *Margulis-GWF2-35-9*. Protein sequences of these families were aligned using MAFFT version 7.027 b employing the L-INS-i strategy (Katoh and Standley 2013). The alignments are available in supplementary material online.

Species Tree Reconstruction

The sequence data for the reconstruction of the cyanobacterial species tree consisted of 14 single-copy gene families that are present in all 199 cyanobacteria genomes and any of the six outgroup genomes. The species tree was inferred using IQ-TREE (Nguyen et al. 2015) in a partitioned analysis over the concatenated alignment of the 14 gene alignments (iqtree version 1.6.6.b with parameters -t BIONJ -keep-ident -mset LG -madd LG4X –spp). The unrooted species tree was rooted on the branch leading to the outgroup. The species tree is available in supplementary material online.

Gene Trees Reconstruction

To evaluate the robustness of inferences drawn from the species tree, we also reconstructed gene trees to provide a large sample of comparisons to the species tree. The gene trees data set consisted of 553 single-copy gene families present in at least one genome from both sides of the root of the species tree, and at least one of the six outgroup species. Gene trees were inferred using IQ-TREE (Nguyen et al. 2015) version 1.6.6.b with parameters -t BIONJ -keep-ident -mset LG -madd LG4X. Trees were rooted on the branch separating the outgroup from the ingroup. A total of 138 trees where the outgroup sequences did not form a single partition were discarded. The gene trees are available in supplementary material online.

Inference of Trait Order

The traits presence/absence pattern was mapped onto the rooted species tree. Most of the traits in our study are rather complex (i.e., they involved multiple genes); hence, their emergence in the evolution of cyanobacteria is expected to

be a rare event. Accordingly, we used a parsimonious reconstruction approach and assigned the origin of a trait to the most recent species tree node where the trait is present in any of the node's descendants. Consistency index (CI) and retention index (RI) for each trait were calculated using the PHYLIP program pars (Felsenstein 2005) as described in Farris (1989). The species tree was traversed from root to tips to determine the order of trait emergence. For each pair of traits we tested whether the order observed in the species tree was reproduced in gene trees with at least two species displaying each of the two traits. For that purpose, we repeated the trait-order analysis with the set of single-copy rooted gene trees, including gene families that do not span the full taxa set. The agreement of the gene trees with the conclusion based on the species trees is calculated as the proportion of gene trees where the trait order is the same as in the species tree.

Results and Discussion

To reconstruct the order of trait emergence in the evolution of cyanobacterial multicellularity, we evaluated 21 phenotypic traits variably present in 199 cyanobacterial species (table 1, supplementary table S1, Supplementary Material online). The ability to perform photosynthesis is not included in our study as it is universal to all cyanobacteria hence it is considered an ancestral trait (Garcia Pichel et al. 2020). We inferred a species tree from a partitioned analysis of 14 single-copy core genes that are present in all 199 species. The species tree was rooted by inclusion of outgroup sequences from six bacterial species (see Materials and Methods), with the root separating the genus *Gloeobacter* from all other cyanobacteria.

The rooting of cyanobacteria is a thorny issue, with two competing clades put forward as basal lineages. Phylogenetic studies based on single gene phylogenies indicated *Gloeobacter* to be a basal lineage within cyanobacteria (e.g., Bhattacharya and Medlin 1995; Honda et al. 1999). The *Gloeobacter* basal position is in agreement with the cellular characteristics of species in this genus, which are lacking well-defined thylakoids (Rippka et al. 1974; Mareš et al. 2019) that are considered a derived structure in the phylum. Indeed, several studies in the literature thus used *Gloeobacter* for rooting the cyanobacterial species tree (e.g., Shi and Falkowski 2008; Shih et al. 2012; Dagan et al. 2013; Sánchez-Baracaldo et al. 2014). Nonetheless, other studies, which used midpoint rooting or minimal ancestor deviation (MAD) to root the cyanobacteria species tree, position the root on the branch separating the pico-cyanobacteria (*Synechococcus* & *Prochlorococcus*, and hereafter SynPro clade) and the remaining species (Szöllősi et al. 2012; Tria et al. 2017). This branch, however, is typically long in gene trees, as well as the species tree, hence it may reflect a Long Branch Attraction (LBA) artifact (defined in Felsenstein 1978).

To evaluate whether the rooting placement is robust, we conducted gene-tree support analyses by reconstructing trees
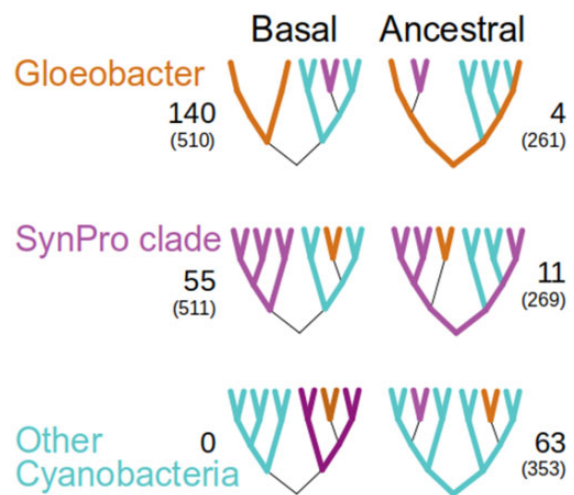

FIG. 1.—Support for three possible basal/ancestral cyanobacterial groups in 273 rooted gene trees. The number of gene families supporting each type of rooted topology is given (median alignment length is shown in parenthesis).

for 553 cyanobacterial single-copy gene families along with homologs from six different outgroup species (fig. 1, and Materials and Methods). We then extracted the rooted cyanobacterial subtree, while discarding 138 trees where the outgroup species formed more than one group. We next characterized the trees by the pattern of the three cyanobacterial subgroups: *Gloeobacter* (2 spp.), SynPro clade (32 spp.), and Other Cyanobacteria (165 spp.), discarding 142 genes that are present in less than two genomes for each group. First, we considered gene trees in which no group is present on both sides of the root. The group appearing on its own as a lineage originating at the root is considered a basal group, and the tree is supporting a root located on the branch separating it from the others (left column in fig. 1). Next, if only one group appears on both sides of the root, this group is labeled as ancestral, and supports a root position within it (right column in fig. 1). We found no gene trees where more than one group appears on both sides of the root. We note that the gene trees may be discordant with the species tree, and that any of the groups may seem to be paraphyletic, either due to methodological artifacts (e.g., LBA involving the SynPro clade) or due to biological processes, such as lateral gene transfer (LGT).

Our results reveal that the majority of gene trees identify the *Gloeobacter* lineage as a distinct basal lineage stemming from the root, thus supporting the outgroup rooting of the species tree on that branch.

The Order of Trait Emergence in Cyanobacterial Evolution

We infer the order of trait emergence in cyanobacterial evolution by mapping the traits onto the rooted species tree. The *origin node* corresponding to each trait was assigned as the most recent node where the trait is present in any of the

node's descendants. This is a conservative approach in that for each trait it allows for a single origin with possible subsequent losses (Dagan and Martin 2007). Our approach furthermore assumes that the traits in our analysis are vertically inherited rather than acquired by lateral transfer. Indeed, previous phylogenomic studies indicated that gene acquisition by lateral gene transfer is frequent in the evolution cyanobacteria (Zhaxybayeva et al. 2009; Dagan et al. 2013). The traits we included in our analysis, however, are complex phenotypic traits that are the product of multiple genes whose expression is well coordinated with other physiological processes in the cell. Such traits are unlikely to be acquired by lateral gene transfer, because their retention requires that multiple coding and regulation elements will be functional on arrival (Cohen et al. 2011). Nonetheless, since complex systems may have been sporadically and rarely transferred (e.g., Brinkmann et al. 2018), we quantified the level of homoplasy and synapomorphy in the evolution of the studied traits by the CI and RI (Farris 1989). Our results for most of the traits revealed high RI values (median 0.61) and low CI values (median 0.048)

(supplementary table S1, Supplementary Material online). This pattern is as expected for complex traits whose evolution is best characterized by a single origin and differential loses.

We next compare the origin nodes for pairs of traits. When the origin of two traits is assigned to the same node, we label the two traits as "simultaneous" at the resolution of the current taxa sample. When an origin node of a trait is a descendent of a second trait's origin node, we conclude that the first trait emerged earlier. A third possibility is that the origin nodes of two traits are not nested, but this relationship was not observed in any of the 210 trait pairs. The order of trait emergence is depicted in figure 2.

The order of trait emergence is inferred from a species tree, yet, phylogenetic inference of species phylogenies based on the concatenation of only few core genes may suffer from a low resolution of the phylogenetic signal (e.g., Dagan and Martin 2006; Thiergart et al. 2014). To evaluate the robustness of the trait order derived from the species tree, we repeated the analysis by testing the trait order inference in individual gene trees. For that purpose, we considered the
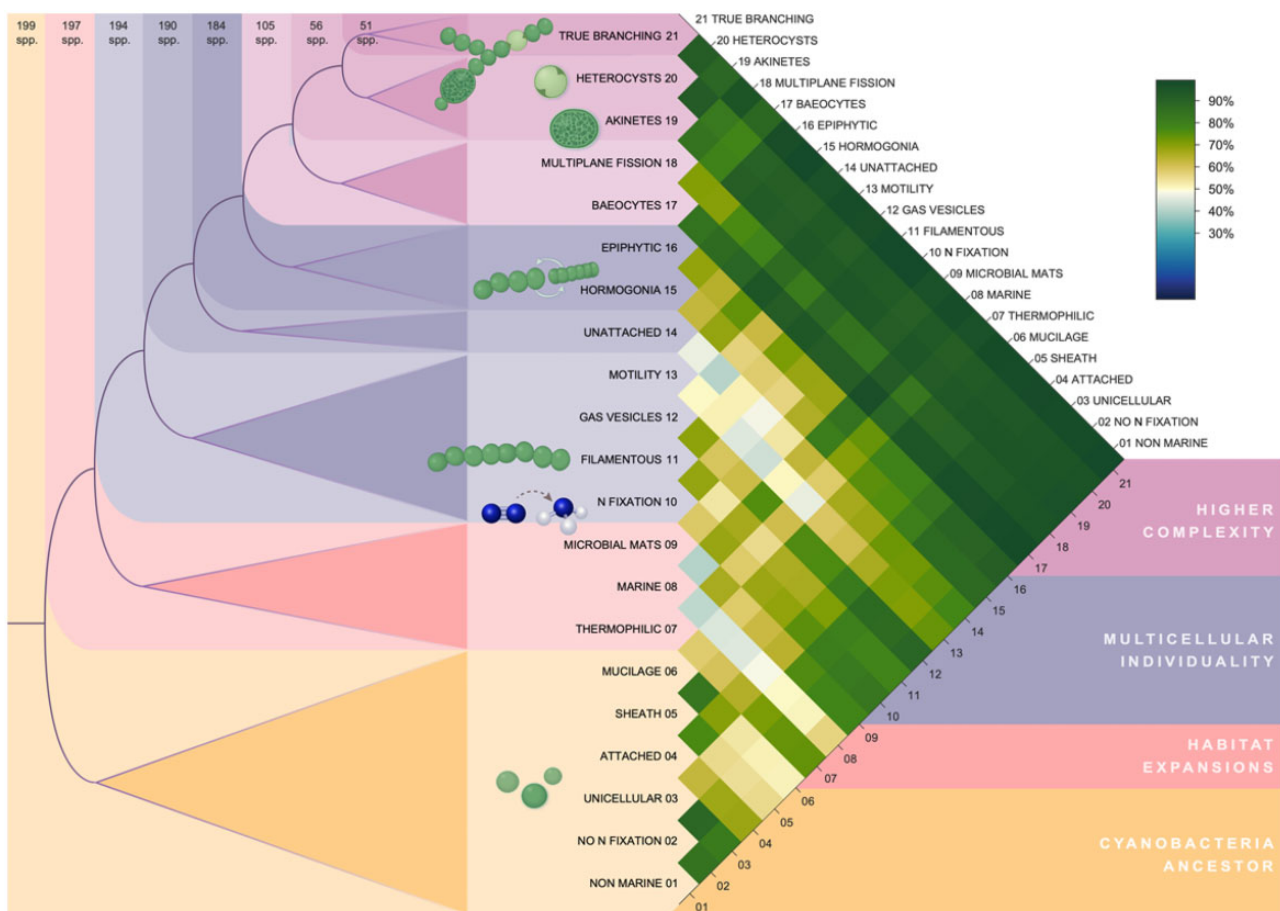

Fig. 2.—Order of trait emergence. Left: traits and their inferred origin node from the rooted species tree. Colors mark traits with a common origin node (note that the order of traits within the colored blocks is arbitrary). Colored boxes are nested, that is, earlier traits are present also in the nested colors. Right: Frequency of gene trees in agreement with the relative order of pairs of traits. Cells in the matrix are shaded according to the color bar on the right.

set of single-copy gene families where the gene is present in at least one species from both sides of the species tree root and at least one outgroup species. The gene trees were rooted by the outgroup, and the order of pairs of traits determined. The analysis of a large sample of gene trees provides a statistical view not possible in a single species tree, but potentially introduces contradicting inferences due to both biological reasons (LGT) and methodological uncertainties (alignment quality and phylogenetic artifacts, such as LBA). In figure 2, we report the percentage of gene trees that reproduces the species tree ordering. The vast majority of trait pair orderings are observed also in >50% of the individual gene trees. Excluding the sequences from the SynPro clade from the gene trees led to slightly better agreement of the gene trees with the order observed in the species tree (supplementary fig. S1, Supplementary Material online), suggesting that LBA artifacts associated with the SynPro clade may lead to disagreement between the gene and species trees. We note that gene trees that are discordant with the species tree generally had shorter alignment length and lower bootstrap support (supplementary tables S2 and S3, Supplementary Material online). Our results thus show a high level of agreement between the species tree and the gene trees, as is expected under an overall homogeneous yet low frequency of LGT during bacterial evolution (Dagan and Martin 2007).

In what follows we divide the inferred order of trait emergence into four temporal phases: (phase i) the cyanobacteria ancestor (traits 1–6); (phase ii) habitat expansions (traits 7–9); (phase iii) the transition to multicellular individuality (traits 10–16); and (phase iv) the evolution of higher complexity (traits 17–21).

The Cyanobacterial Ancestor and Subsequent Habitat Expansion

The rooted tree topology supports the view that the cyanobacterial ancestor was characterized by traits that include UNICELLULAR and NO N FIXATION (fig. 2). The ancestral state of both traits as preceding the emergence of filamentous forms is debated in the literature—whereas one study suggested the ancestor to be unicellular and the filamentous morphology to arise in independent lineages of the cyanobacterial tree (Sánchez-Baracaldo et al. 2005), another view posed that the filamentous morphology evolved early during cyanobacterial evolution and was subsequently lost and regained several times (Schirrmeister et al. 2011). There are also claims that the last cyanobacterial common ancestor already fixed N2 (Tomitani et al. 2006), there are however others that concluded that it could not fix N2 and that cyanobacteria must have acquired this trait several times independently (Sánchez-Baracaldo et al. 2005).

We can further deduce from our data that the cyanobacterial ancestor lived ATTACHED and possessed a SHEATH and

MUCILAGE. Whether the cyanobacteria ancestor lived ATTACHED is a matter of debate and opposing views on the topic have been published (Schopf 1993; Garcia-Pichel 1998; Sánchez-Baracaldo et al. 2005; Sánchez-Baracaldo 2015; Uyeda et al. 2016).

SHEATH and MUCILAGE are forms of extracellular polymeric substances (EPS), located outside the cell wall, which in today's cyanobacteria are mainly involved in protecting the cell from various stresses, such as UV and desiccation (Potts 1994; Ehling-Schulz and Scherer 1999). Furthermore, we find the cyanobacteria ancestor to have most likely inhabited a NON MARINE environment, agreeing with studies that suggest that early cyanobacteria lived in freshwater or terrestrial habitats and subsequently diverged into marine environments (Dagan et al. 2013; Uyeda et al. 2016).

The second phase in cyanobacterial evolution is expansion of the cyanobacterial habitat, indicated by the traits MARINE and THERMOPHILIC, which are both inferred to simultaneously occur with the ability to form MICROBIAL MATS. MICROBIAL MATS are dense communities (Stal 1995) that typically present a laminated segregation of functional types. They are often formed by cyanobacteria and are frequently found in extreme habitats, such as deserts or hot springs, characterized by temperatures between 30 °C to 73 °C (e.g., Cox et al. 2011).

The Emergence of N_2 Fixation is at the Origin of Cyanobacterial Multicellularity

Phase iii in our reconstruction comprises three sets of cyanobacterial traits (fig. 2). First, the simultaneous emergence of the FILAMENTOUS morphology, N FIXATION, GAS VESICLES, and MOTILITY, followed by the trait UNATTACHED, and lastly by the co-occurrence of HORMOGONIA and EPIPHYTIC. During cyanobacterial N_2 fixation, molecular dinitrogen (N_2) is reduced to ammonia (NH_3), a process that is catalyzed by the enzyme nitrogenase. Whereas present day cyanobacteria, other microorganisms, and most plants are able to take up nitrogen in various combined forms, such as nitrate, ammonium, organic nitrogen, or urea, these combined forms of nitrogen are scarce in most environments (e.g., open oceans or terrestrial habitats [Zehr 2011]). Combined nitrogen, which is critical for the biosynthesis of amino and nucleic acids, was likely a limiting resource in the early Earth environment (Kasting and Siefert 2001).

The realization of the full metabolic potential of N_2 fixation, however, faced the challenge of the incompatibility of nitrogenase with intracellular oxygen (Gallon 1981). When the cyanobacterial ancestor first acquired the capacity of N_2 fixation, it must have imposed a strong selection pressure on the individual cells. The trade off between photosynthesis and nitrogen fixation led to the evolution of multiple solutions, which are still present in today's cyanobacteria: the circadian rhythm of N_2 fixation in unicellular cyanobacteria (Mitsui et al. 1986), specific cells devoted to N_2 fixation in undifferentiated filaments (Berman-Frank et al. 2003; Bergman et al. 2013),

and the differentiation of the highly specialized heterocyst in filamentous cyanobacteria (Flores et al. 2019).

The need to compartmentalize the two incompatible functions, photosynthesis and N₂ fixation, has been proposed to drive the emergence of multicellular forms in cyanobacteria (Ispolatov et al. 2012). More specifically, the result from a quantitative theoretical model predicts that within a population of genetically identical unicellular nitrogen fixing cyanobacteria, cell differentiation and phenotypic heterogeneity would have been adaptive if this increased the fitness of the organisms in multicellular groups. In the case of unicellular cyanobacteria this means that cells evolved adhesion and exchanged fixed nitrogen and carbon products within early cell groups, such as filaments.

In filamentous cyanobacteria, dividing cells remain linked in a chain, resulting in a localization of cells in close spatial proximity, facilitating metabolite exchange between the individual cells. When compared with the more transient associations in spatially structured communities, such as in EPS imbedded biofilms, the development of filaments opens possibilities for a more direct and permanent exchange of molecules between neighboring cells with high specificity. Metabolic exchange could have evolved as described for the evolution of metabolic cross-feeding (D'Souza et al. 2018), as the exchange of carbon and nitrogen against other products is generally common in photosynthetic or nitrogen-fixing organisms (Kaiser et al. 2015).

The emergence of GAS VESICLES and MOTILITY traits signifies the evolution from a stationary to a more active lifestyle, enabling cells to regulate their buoyancy in the water column. This result is further supported by the subsequent inference of UNATTACHED, which indicates the transition from a benthic to a planktonic lifestyle.

Thereafter, the traits HORMOGONIA and EPIPHYTIC are inferred to occur simultaneously. The differentiation of HORMOGONIA can be induced by environmental stimuli, such as nitrogen deprivation (Flores and Herrero 2010). HORMOGONIA are released from the mother filament through the formation of necridia, dead cells resulting from PCD (Nürnberg et al. 2014). After their release from the main trichome, HORMOGONIA disperse via gliding motility or float thanks to GAS VESICLES, ensuring the reproduction of benthic species (Rippka et al. 1979). HORMOGONIA with GAS VESICLES are thus important for distribution in aquatic environments as known from modern cyanobacteria (*Fischerella, Hapalosiphon, Tolypothrix*) (Komárek 2013). The close local association with plants, as indicated by EPIPHYTIC, however might have been the first step toward the initiation of one of the many symbioses between higher organisms and cyanobacteria, where HORMOGONIA serve as the infection units (Meeks and Elhai 2002).

Notably, the differentiation into HORMOGONIA is reversible, as they develop a sessile lifestyle, where they grow into a new vegetative filament (Flores and Herrero 2010). Here we observe the emergence of a life cycle with two distinct cell types,

which is important for the transition to multicellularity (Hammerschmidt et al. 2014; Rose 2020). Such a life cycle results in selection operating at the higher, the filament, level and thus ensures the reproduction of the newly formed collective entity.

The Evolution of Cell Differentiation Leads to Higher Cyanobacterial Complexity

A central innovation that is associated with this phase (iv) in the species tree is MULTIPLANE FISSION. This trait co-occurred with the ability to produce BAEOCYTES, differentiated cells, which are the reproductive stages in the order Pleurocapsales (Waterbury and Stanier 1978). Notably, baeocyte-forming cyanobacteria, that have been traditionally grouped together with unicellular cyanobacteria (Rippka et al. 1979), appear to immediately predate the evolution of spore-like AKINETES and nitrogen-fixing HETEROCYSTS and thus emerged much later than filamentous forms. Indeed, a recent study suggested that *Gloeocapsopsis* sp., a baeocytous cyanobacterium, harbors several characteristics that are in common with filamentous cyanobacteria, including mechanisms of cell–cell communication (Urrejola et al. 2020).

The late timing and the co-occurrence of the two traits AKINETES and HETEROCYSTS, indicative of higher complexity, are in line with the view that the evolution of the heterocyst was relatively late in the history of filamentous cyanobacteria (Tomitani et al. 2006), and where a common origin of akinetes and heterocysts has been proposed (Adams and Duggan 1999). HETEROCYSTS represent not only a morphological adaptation to the obstacle of N₂ fixation under oxic conditions but also an elaborate and highly specialized communication and metabolite exchange system. In *Anabaena* sp., for example, where several hundred cells communicate within a filament, a regular heterocyst formation pattern along the filament must be achieved to guarantee that every cell is adequately supplied with fixed nitrogen compounds (Herrero et al. 2016). For this, the inhibitory signaling peptide PatS needs to be distributed along the filament with heterocyst formation occurring only in cells with low PatS concentration (Yoon and Golden 1998). Whether the exchange of metabolites and regulators happens via the continuous periplasm (Flores et al. 2006) or through septal junctions (Mullineaux et al. 2008) is still not fully resolved.

The trait that evolved last, based on the analysis, is TRUE-BRANCHING, where cells in a filament perform MULTIPLANE FISSION. TRUE-BRANCHING is characteristic for the members of the Haphalosiphon/Stigonematales clade. Our results confirm previous morphological and phylogenetic studies that found TRUE-BRANCHING to be the latest evolutionary innovation in cyanobacteria (Rippka et al. 1979; Dagan et al. 2013; Koch et al. 2017).

Trade Offs between Incompatible Processes Lead to Division of Labor and Stable Multicellularity

Common features of evolutionary transitions in individuality comprise cooperation between the lower level units (Bonner 1998) and the division of labor (Michod 2007). The latter might be of particular advantage, and serve as the driver of the transition to multicellularity when there is a strong trade off between processes that cannot be performed in a single cell at one time (Michod 2007; Ispolatov et al. 2012). Our current findings support this theory and point to nitrogen fixation, and its incompatibility with photosynthesis, as the trigger for the evolution of multicellularity in cyanobacteria. One open question concerns how the underlying genetics of novel traits, such as the division of labor, arise within a newly emerging multicellular individual. In the case of cyanobacteria multicellularity, as also suggested for animal multicellularity (Brunet and King 2017), we conjecture that no new genes were required and that higher complexity was achieved by regulatory changes in gene expression patterns. Basic communication and metabolite exchange was pre-existing as single-celled bacteria frequently engage in cell–cell communication and cross-feeding of metabolites via the external environment (D'Souza et al. 2018). Division of labor between photosynthesis and nitrogen fixation was likely first established by the regulatory mechanism of temporal switching. Once simple forms of division of labor and metabolic exchange existed, the transition into spatial separation in differentiated cells could have evolved mainly by regulatory modifications.

Differentiated cells are one of the hallmarks of complex multicellularity. It is therefore significant that we observe six distinct cell types in cyanobacteria: Photosynthetic, hormogonia, necridia, akinetes, baeocytes, and heterocysts. Such a plurality indicates that the underlying regulatory mechanisms are well developed and that their plasticity and adaptability are a matter of course. It is also significant that three of the differentiated cell types, hormogonia, akinetes, and baeocytes, offer novel reproductive potential and the establishment of a multicellular life cycle. Moreover, signs of a nascent developmental plan can be observed in both the distribution of heterocysts along filaments and in the patterning of true branching cyanobacteria. These elements have no fitness value for the individual cell, but are selectable adaptations at the higher level, the filament. The chronology of the evolution of multicellularity in cyanobacteria shows that, once established, multicellular individuality opens new vistas of opportunities.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Author Contributions

K.H. and T.D. conceived the study. K.H. collected the traits data. G.L., F.D.K.T., and J.A. performed the analyses. K.H., G.L., and T.D. wrote the manuscript with contributions from F.D.K.T. and J.A.

Data Availability

The alignments and trees are available as supplementary data at *Genome Biology and Evolution* online.

Literature Cited

Adams DG, Duggan PS. 1999. Heterocyst and akinete differentiation in cyanobacteria. New Phytol. 144(1):3–33.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215(3):403–410.

Bergman B, Sandh G, Lin S, Larsson J, Carpenter EJ. 2013. *Trichodesmium*—a widespread marine cyanobacterium with unusual nitrogen fixation properties. FEMS Microbiol Rev. 37(3):286–302.

Berman-Frank I, Lundgren P, Falkowski P. 2003. Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. Res Microbiol. 154(3):157–164.

Bhattacharya D, Medlin L. 1995. The phylogeny of plastids: a review based on comparisons of small-subunit ribosomal RNA coding regions. J Phycol. 31(4):489–498.

Bonner JT. 1998. The origins of multicellularity. Integr Biol. 1(1):27–36.

Brinkmann H, Göker M, Koblížek M, Wagner-Döbler I, Petersen J. 2018. Horizontal operon transfer, plasmids, and the evolution of photosynthesis in *Rhodobacteraceae*. ISME J. 8:a018127.

Brunet T, King N. 2017. The origin of animal multicellularity and cell differentiation. Dev Cell. 43(2):124–140.

Carnevali PBM. 2019. Hydrogen-based metabolism as an ancestral trait in lineages sibling to the cyanobacteria. Nat Commun. 10:1–15.

Claessen D, Rozen DE, Kuipers OP, Søgaard-Andersen L, van Wezel GP. 2014. Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. Nat Rev Microbiol. 12(2):115–124.

Cohen O, Gophna U, Pupko T. 2011. The complexity hypothesis revisited: connectivity rather than function constitutes a barrier to horizontal gene transfer. Mol Biol Evol. 28(4):1481–1489.

Cox A, Shock EL, Havig JR. 2011. The transition to microbial photosynthesis in hot spring ecosystems. Chem Geol. 280(3-4):344–351.

Dagan T, et al. 2013. Genomes of Stigonematalean cyanobacteria (Subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. Genome Biol Evol. 5(1):31–44.

Dagan T, Martin W. 2006. The tree of one percent. Genome Biol. 7(10):118.

Dagan T, Martin W. 2007. Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. Proc Natl Acad Sci USA. 104(3):870–875.

D'Souza G, et al. 2018. Ecology and evolution of metabolic cross-feeding interactions in bacteria. Nat Prod Rep. 35(5):455–488.

Ehling-Schulz M, Scherer S. 1999. UV protection in cyanobacteria. Eur J Phycol. 34(4):329–338.

Enright AJ, Van Dongen S, Ouzounis CA. 2002. An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res. 30(7):1575–1584.

Farris JS. 1989. The retention index and the rescaled consistency index. Cladistics 5(4):417–419.

Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst Zool. 27(4):401–410.

Felsenstein J. 2005. PHYLIP (Phylogeny Inference Package) version 3.697. Distributed by the author. Seattle (WA): Department of Genome Sciences. University of Washington.

Flores E, Herrero A. 2010. Compartmentalized function through cell differentiation in filamentous cyanobacteria. Nat Rev Microbiol. 8(1):39–50.

Flores E, Herrero A, Wolk CP, Maldener I. 2006. Is the periplasm continuous in filamentous multicellular cyanobacteria? Trends Microbiol. 14(10):439–443.

Flores E, Picossi S, Valladares A, Herrero A. 2019. Transcriptional regulation of development in heterocyst-forming cyanobacteria. Biochim Biophys Acta Gene Regul Mech. 1862(7):673–684.

Frías JE, Flores E, Herrero A. 1994. Requirement of the regulatory protein NtcA for the expression of nitrogen assimilation and heterocyst development genes in the cyanobacterium Anabaena sp. PCC7120. Mol Microbiol. 14(4):823–832.

Gallon JR. 1981. The oxygen sensitivity of nitrogenase: a problem for biochemists and micro-organisms. Trends Biochem Sci. 6:19–23.

Garcia-Pichel F. 1998. Solar ultraviolet and the evolutionary history of cyanobacteria. Origins Life Evol Biosphere. 28(3):321–347.

Garcia Pichel F, Zehr JP, Bhattacharya D, Pakrasi HB. 2020. What's in a name? The case of cyanobacteria. J Phycol. 56(1):1–5.

Hammerschmidt K, Rose CJ, Kerr B, Rainey PB. 2014. Life cycles, fitness decoupling and the evolution of multicellularity. Nature 515(7525):75–79.

Herrero A, Stavans J, Flores E. 2016. The multicellular nature of filamentous heterocyst-forming cyanobacteria. FEMS Microbiol Rev. 40(6):831–854.

Honda D, Yokota A, Sugiyama J. 1999. Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine Synechococcus strains. J Mol Evol. 48(6):723–739.

Ispolatov I, Ackermann M, Doebeli M. 2012. Division of labour and the evolution of multicellularity. Proc R Soc B. 279(1734):1768–1776.

Kaiser D, Kowalski N, Böttcher ME, Yan B, Unger D. 2015. Benthic nutrient fluxes from mangrove sediments of an anthropogenically impacted estuary in Southern China. J Mar Sci Eng. 3(2):466–491.

Kasting JF, Siefert JL. 2001. Biogeochemistry. The nitrogen fix. Nature 412(6842):26–27.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

Koch R, et al. 2017. Plasticity first: molecular signatures of a complex morphological trait in filamentous cyanobacteria. BMC Evol Biol. 17(1):209.

Komárek J. 2013. Süßwasserflora von Mitteleuropa, Bd. 19/3: cyanoprokaryota; 3rd part: heterocystous Genera. Berlin (Germany): Springer Spektrum.

Mareš J, Strunecký O, Bučinská L, Wiedermannová J. 2019. Evolutionary patterns of thylakoid architecture in cyanobacteria. Front Microbiol. 10:277.

Meeks JC, Elhai J. 2002. Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. Microbiol Mol Biol Rev. 66(1):94–121.

Michod RE. 2007. Evolution of individuality during the transition from unicellular to multicellular life. Proc Natl Acad Sci USA. 104(Suppl 1):8613–8618.

Mitsui A, et al. 1986. Strategy by which nitrogen-fixing unicellular cyanobacteria grow photoautotrophically. Nature 323(6090):720–722.

Mullineaux CW, et al. 2008. Mechanism of intercellular molecular exchange in heterocyst-forming cyanobacteria. EMBO J. 27(9):1299–1308.

Nguyen L-T, Schmidt HA, Haeseler von A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274.

Nürnberg DJ, et al. 2014. Branching and intercellular communication in the section V cyanobacterium Mastigocladus laminosus, a complex multicellular prokaryote. Mol Microbiol. 91(5):935–949.

O'Leary NA, et al. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 44:D733–745.

Potts M. 1994. Desiccation tolerance of prokaryotes. Microbiol Rev. 58(4):755–805.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. Trends Genet. 16(6):276–277.

Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY. 1979. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. Microbiology 111(1):1–61.

Rippka R, Waterbury J, Cohen-Bazire G. 1974. A cyanobacterium which lacks thylakoids. Arch Microbiol. 100(1):419–436.

Rose CJ. 2020. Germ lines and extended selection during the evolutionary transition to multicellularity. J Exp Zool (Mol Dev Evol). 1–7. doi:10.1002/jez.b.22985

Sánchez-Baracaldo P. 2015. Origin of marine planktonic cyanobacteria. Sci Rep. 5:17418.

Sánchez-Baracaldo P, Hayes PK, Blank CE. 2005. Morphological and habitat evolution in the Cyanobacteria using a compartmentalization approach. Geobiology 3(3):145–165.

Sánchez-Baracaldo P, Ridgwell A, Raven JA. 2014. A neoproterozoic transition in the marine nitrogen cycle. Curr Biol. 24(6):652–657.

Schirrmeister BE, Antonelli A, Bagheri HC. 2011. The origin of multicellularity in cyanobacteria. BMC Evol Biol. 11(1):45–44.

Schopf JW. 1993. Microfossils of the early Archean Apex chert: new evidence of the antiquity of life. Science 260(5108):640–646.

Shapiro J. 1988. Bacteria as multicellular organisms. Sci Am. 258(6):82–89.

Shi T, Falkowski PG. 2008. Genome evolution in cyanobacteria: the stable core and the variable shell. Proc Natl Acad Sci USA. 105(7):2510–2515.

Shih PM, et al. 2012. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. Proc Natl Acad Sci USA. 110:201217107–201211058.

Soo RM, Hemp J, Parks DH, Fischer WW, Hugenholtz P. 2017. On the origins of oxygenic photosynthesis and aerobic respiration in cyanobacteria. Science 355(6332):1436–1440.

Stal LJ. 1995. Physiological ecology of cyanobacteria in microbial mats and other communities. New Phytol. 131(1):1–32.

Szöllősi GJ, Boussau B, Abby SS, Tannier E, Daubin V. 2012. Phylogenetic modeling of lateral gene transfer reconstructs the pattern and relative timing of speciations. Proc Natl Acad Sci USA. 109(43):17513–17518.

Tarnita CE, Taubes CH, Nowak MA. 2013. Evolutionary construction by staying together and coming together. J Theor Biol. 320:10–22.

Tatusov RL, Koonin EV, Lipman DJ. 1997. A genomic perspective on protein families. Science 278(5338):631–637.

Thiergart T, Landan G, Martin WF. 2014. Concatenated alignments and the case of the disappearing tree. BMC Evol Biol. 14(1):266.

Tomitani A, Knoll AH, Cavanaugh CM, Ohno T. 2006. The evolutionary diversification of cyanobacteria: molecular–phylogenetic and paleontological perspectives. Proc Natl Acad Sci USA. 103(14):5442–5447.

Tria FDK, Landan G, Dagan T. 2017. Phylogenetic rooting using minimal ancestor deviation. Nat Ecol Evol. 1(1):0193–0197.

Urrejola C, et al. 2020. Loss of filamentous multicellularity in cyanobacteria: the extremophile *Gloeocapsopsis* sp. Strain UTEX B3054 retained multicellular features at the genomic and behavioral levels. J Bacteriol. 202(12):e00514–19.

Uyeda JC, Harmon LJ, Blank CE. 2016. A comprehensive study of cyanobacterial morphological and ecological evolutionary dynamics through deep geologic time. PLoS ONE 11(9):e016253932.

van Gestel J, Vlamakis H, Kolter R. 2015. Division of labor in biofilms: the ecology of cell differentiation. Microbiol Spec. 3:1–24.

Waterbury JB, Stanier RY. 1978. Patterns of growth and development in pleurocapsalean cyanobacteria. Microbiol Rev. 42(1):2–44.

Yoon H-S, Golden JW. 1998. Heterocyst pattern formation controlled by a diffusible peptide. Science 282(5390):935–938.

Zehr JP. 2011. Nitrogen fixation by marine cyanobacteria. Trends Microbiol. 19(4):162–173.

Zhaxybayeva O, Doolittle WF, Papke RT, Gogarten JP. 2009. Intertwined evolutionary histories of marine *Synechococcus* and *Prochlorococcus marinus*. Genome Biol Evol. 1:325–339.

Zhu Q, et al. 2019. Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains bacteria and archaea. Nat Commun. 10(1):1–14.

Associate editor: Sandra Baldauf