

# Unravelling the Complete Genome of *Archangium gephyra* DSM 2261<sup>T</sup> and Evolutionary Insights into Myxobacterial Chitinases

Gaurav Sharma<sup>1,2</sup> and Srikrishna Subramanian<sup>1,\*</sup>

<sup>1</sup>Protein Science and Engineering, CSIR-Institute of Microbial Technology, Sector-39A, Chandigarh, India

<sup>2</sup>Present address: Department of Microbiology and Molecular Genetics, University of California, Davis, CA

\*Corresponding author: E-mail: krishna@imtech.res.in.

Accepted: April 3, 2017

**Data deposition:** This genome has been deposited at DDBJ/ENA/GenBank under the accession number CP011509 with bioproject number PRJNA283574.

## Abstract

Family Cystobacteraceae is a group of eubacteria within order Myxococcales and class Deltaproteobacteria that includes more than 20 species belonging to 6 genera, that is, *Angiococcus*, *Archangium*, *Cystobacter*, *Hyalangium*, *Melittangium*, and *Stigmatella*. Earlier these members have been classified based on chitin degrading efficiency such as *Cystobacter fuscus* and *Stigmatella aurantiaca*, which are efficient chitin degraders, *C. violaceus* a partial chitin degrader and *Archangium gephyra* a chitin nondegrader. Here we report the 12.5 Mbp complete genome of *A. gephyra* DSM 2261<sup>T</sup> and compare it with four available genomes within the family Cystobacteraceae. Phylogeny and DNA–DNA hybridization studies reveal that *A. gephyra* is closest to *Angiococcus disciformis*, *C. violaceus* and *C. ferrugineus*, which are partial chitin degraders of the family Cystobacteraceae. Homology studies reveal the conservation of approximately half of the proteins in these genomes, with about 15% unique proteins in each genome. The total carbohydrate-active enzymes (CAZome) analysis reveals the presence of one GH18 chitinase in the *A. gephyra* genome whereas eight copies are present in *C. fuscus* and *S. aurantiaca*. Evolutionary studies of myxobacterial GH18 chitinases reveal that most of them are likely related to *Terrabacteria* and *Proteobacteria* whereas the *Archangium* GH18 homolog shares maximum similarity with those of chitin nondegrading *Acidobacteria*.

**Key words:** phylogeny, methylome, *Cystobacter*, *Stigmatella*, *Hyalangium*.

## Introduction

Order Myxococcales (Myxobacteria) is divided into 3 suborders that have been further classified into 10 families, more than 25 genera, and around 50 species (<https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=29>; last accessed April 10, 2017). These organisms are aerobic (with the exception of *Anaeromyxobacter* [Sanford et al. 2002; Thomas et al. 2008]) Gram-negative bacteria and exhibit social behavior, gliding motility, and fruiting body formation (Reichenbach 2005; Goldman et al. 2006; Huntley et al. 2011). Myxobacteria possess some of the largest known genomes among prokaryotes (typically 9–16 Mbp), with a large number of encoded genes (>7,000 genes) and regulatory repertoire (Schneiker et al. 2007; Han et al. 2013; Sharma et al. 2016a, 2016b). *Archangium gephyra* DSM 2261<sup>T</sup> is a Gram-negative, motile,

aerobic, and chitin nondegrading myxobacteria classified in the family Cystobacteraceae in the order Myxococcales within the class Deltaproteobacteria. The genus name “*Archangium*” was coined by Jahn in 1924 on the basis of the distinct fruiting body when compared with other known myxobacteria (Jahn 1924). The classified strains under genus “*Archangium*” were reported to have convoluted braided strings of rod-shaped myxobacteria dipped in hardened slime (Reichenbach 1965; dos Santos et al. 2014). This genus consists of a single species, that is, *A. gephyra*. In the past, several *Archangium* species have been described among which some were found identical to *A. gephyra*, whereas others were reclassified into other myxobacterial species (Reichenbach 2005). *Archangium gephyra* DSM 2261<sup>T</sup> is the type strain of this genus. *Archangium gephyra* is a common soil microorganism which shares a wide range of

habitats similar to the rest of the myxobacteria (Reichenbach 2005). They have been isolated from cold environments such as the Swedish arctic tundra, Alaska, Iceland, and soils of northern Canada as well as from hot biotopes like desert soils and also from the coastal samples from North and South America (dos Santos et al. 2014). *Archangium gephyra* has both proteolytic and bacteriolytic modes of nutrition, that is, they can use both proteins as well as other bacteria for their nutrition requirements (Reichenbach 2005). It has been reported earlier that they are not able to degrade chitin when compared with efficient chitin degraders of family Cystobacteraceae (dos Santos et al. 2014; Awal et al. 2017).

*Archangium gephyra* DSM 2261<sup>T</sup> has been isolated from Essex county (Ontario, Canada) soil samples in 1969 using soil baiting method with sterile, antibiotic-free rabbit dung pellets (McCurdy 1969). In this study, we report the 12.5 Mbp complete genome sequence of *A. gephyra* DSM 2261<sup>T</sup> along with its genomic features and annotation. Phylogenetic analysis based on 16S rRNA and *in silico* DNA–DNA hybridization (DDH) values position *A. gephyra* close to *Cystobacter violaceus* and *C. disciformis*, partial chitin degraders belonging to the family Cystobacteraceae. We performed homology and orthology studies within the five available genomes of family Cystobacteraceae. Finally, we have compared the myxobacterial GH18 chitinases to get insights into their origin.

## Materials and Methods

### Growth Conditions and DNA Isolation

Actively growing plate culture of *A. gephyra* was procured from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) as strain number DSM 2261<sup>T</sup> (also known as strain number M18, ATCC 25201, NBRC 100087; as mentioned in DSMZ records). Genomic DNA for next generation sequencing (NGS) was obtained using both the Zymogen Research Bacterial/fungal DNA isolation kit and Phenol–Chloroform–Isoamyl alcohol-based manual method. The quantity and quality of the extraction were checked by gel electrophoresis along with Nanodrop method and followed by Qubit quantification.

### Genome Sequencing, Assembly, and Annotation

Sequencing was performed on a Pacific Biosciences RSII instrument at the McGill University and Genome Quebec Innovation Center, Montréal (Québec), Canada (Full protocol described in Sharma et al. 2016a). SMRTbell long library was constructed with 10 µg whole genomic DNA using a 20-kb template preparation method using Procedure and Checklist–20 kb Template Preparation using BluePippin™ Size Selection (<http://www.pacb.com/wp-content/uploads/2015/09/Shared-Protocol-20-kb-Template-Preparation-Using-BluePippin-Size-Selection-System-15-kb-Size-Cutoff.pdf>; last accessed April 10, 2017). The library was then loaded onto four single-

molecule real-time (SMRT) cells and sequenced using P6 polymerase and C4 chemistry (P6C4) with 180-min movie time. Sequencing yielded 104,260 raw reads of the average length of 11,124 bp; a total of 89,020 reads with minimum seed read length of 6 kb was used to perform read corrections during assembly. De novo assembly was carried out using the hierarchical genome assembly process protocol from SMRT Analysis v2.0, including consensus polishing with Quiver (Chin et al. 2013). Gene prediction and functional annotation were performed by Rapid Annotation using Subsystem Technology (RAST) (Aziz et al. 2008). RNAmmer 1.2 (Lagesen et al. 2007) and tRNAscan-SE-1.23 (Lowe and Eddy 1997) were used to predict rRNA and tRNA genes. The predicted coding DNA sequences (CDS) were translated and searched against the nonredundant database, TIGRFam, Pfam (Finn et al. 2014), COG, and CAZY databases (Yin et al. 2012; Lombard et al. 2014). Phobius (Kall et al. 2007) and TMHMM (Krogh et al. 2001) were used to predict signal peptides and transmembrane helices, respectively. CRISPR and IS elements were identified using CRISPRfinder (Grissa et al. 2007) and ISfinder servers (Siguier et al. 2006).

### Comparative Genomic Studies

The genomes of *A. gephyra* DSM 2261<sup>T</sup> (*Ag*: CP011509.1), *C. fuscus* DSM 2262<sup>T</sup> (*Cyb*: ANAH000000000), *C. violaceus* Cb vi76 (*Cyvi*: JPMI000000000) (Stevens et al. 2014), *Hyalangium minutum* DSM 14724<sup>T</sup> (*Hm*: JMCB000000000), and *Stigmatella aurantiaca* DW4/3-1 (*Sa*: NC\_014623.1) (Huntley et al. 2011) were downloaded from NCBI. For all these genomes, gene prediction and functional annotation were performed using RAST. Proteinortho (Lechner et al. 2011), which uses reciprocal best hits Basic Local Alignment Search Tool (BLAST) approach, was used to predict the orthologous proteins in various clusters among the five family Cystobacteraceae members. NCBI BLAST+(v 2.2.26+) was used throughout the study (Altschul et al. 1990) using *E*-value cutoff of  $1e^{-5}$ , query coverage of 50% and 35% identity. For homology studies, all proteomes were mapped against each other using BLASTp. A Venn diagram was generated to depict the number of homologous and orthologous proteins between different genome combinations. The total carbohydrate-active enzymes (CAZomes) for each genome was identified using profile searches against CAZY database (Lombard et al. 2014). GH18 characteristic motif "[LIVMFY][DN]G[LIVMF][DN][LIVMF][DN].E" (Prosite No. PS01095) was also searched among all the proteomes to retrieve putative GH18 chitinases.

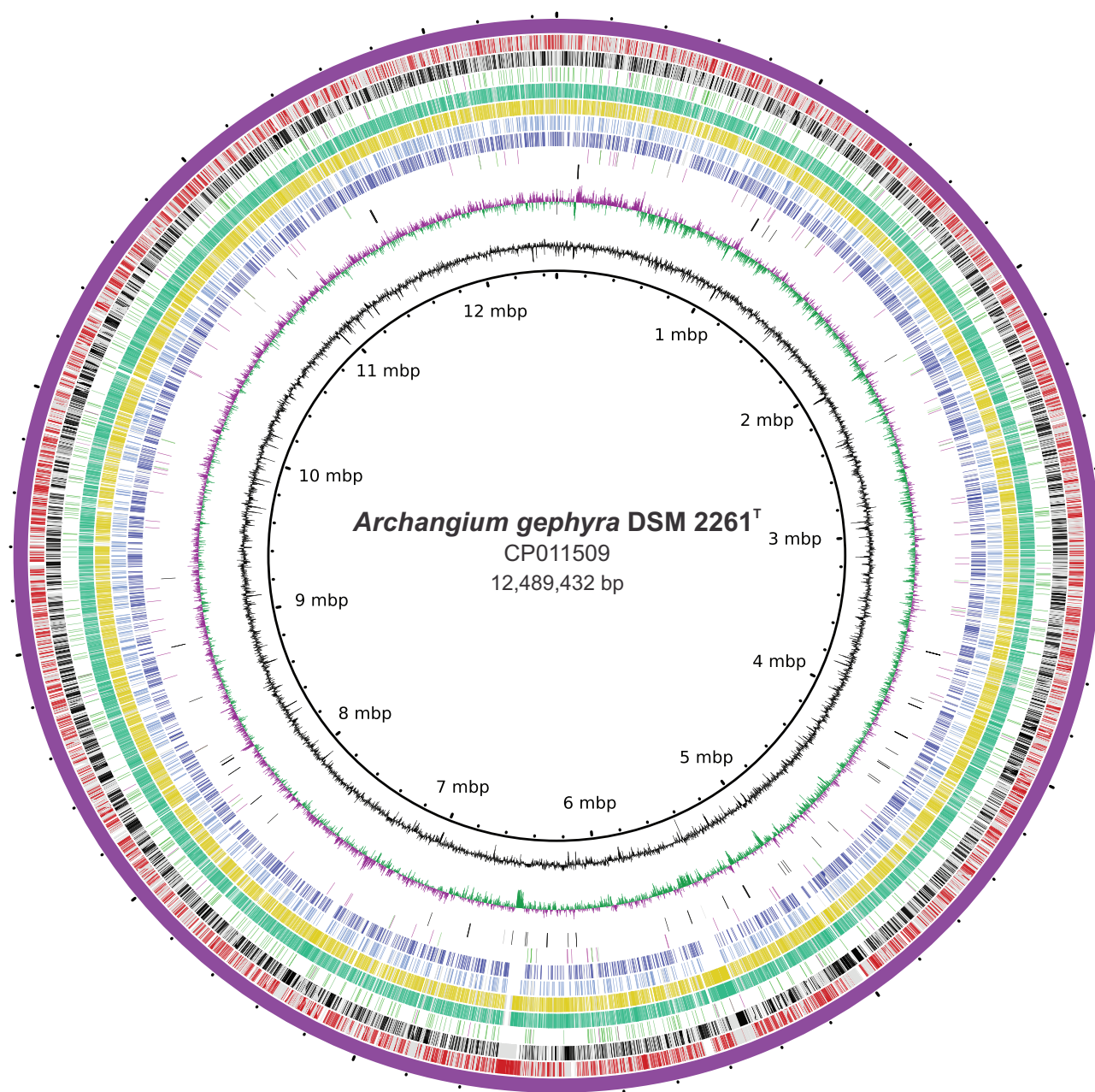
## Results and Discussion

### Genome Features

The genome of *A. gephyra* DSM 2261<sup>T</sup> was assembled into a complete chromosome of 12,489,432 bp having 69.5%

G + C content (fig. 1, supplementary table S1, Supplementary Material online). No plasmid sequence was detected in the assembly process. In the PacBio library preparation step, short DNA fragments are filtered out due to BluePippin™ Size Selection protocol (>20 K), it is possible that small plasmids may not be represented during sequencing. The genome

encodes 10,211 predicted genes that include 10,121 protein-coding genes, 90 RNA genes (9 rRNA genes and 81 tRNA genes), 3 CRISPR repeats, and 109 IS elements. This genome has tRNAs for the 20 amino acids along with 2 pseudo-tRNA genes, 1 tRNA for Selenocysteine (SeC), and 1 tRNA with undetermined function. Among the encoded 10,121



**FIG. 1.**—Circular representation of the complete genome of *Archangium gephyra* DSM 2261<sup>T</sup>. Rings from outside to the center: (1) *A. gephyra* DSM 2261<sup>T</sup> chromosome, (2) genes encoded on the positive strand, (3) genes encoded on the negative strand, (4) Unique genes when compared with nonredundant (nr) database, (5) genes mapped against Pfam database, (6) genes mapped against COG database, (7) genes having transmembrane domains ( $\geq 1$  domain), (8) genes having signal peptide sequences, (9) IS elements, (10) RNA sequences, (11) GC skew, (12) GC content. The image was generated using BRIG (Alikhan et al. 2011).



proteins, functions were assigned to 5,886 proteins (58.16%), whereas the rest 4,235 proteins (41.84%) have been annotated as hypothetical proteins (supplementary table S1, Supplementary Material online). Only 66.5% proteins of *A. gephyra* genome were mapped against Clusters of Orthologous Groups (COGs) functional categories as depicted in supplementary table S2, Supplementary Material online.

### Taxonomic Classification

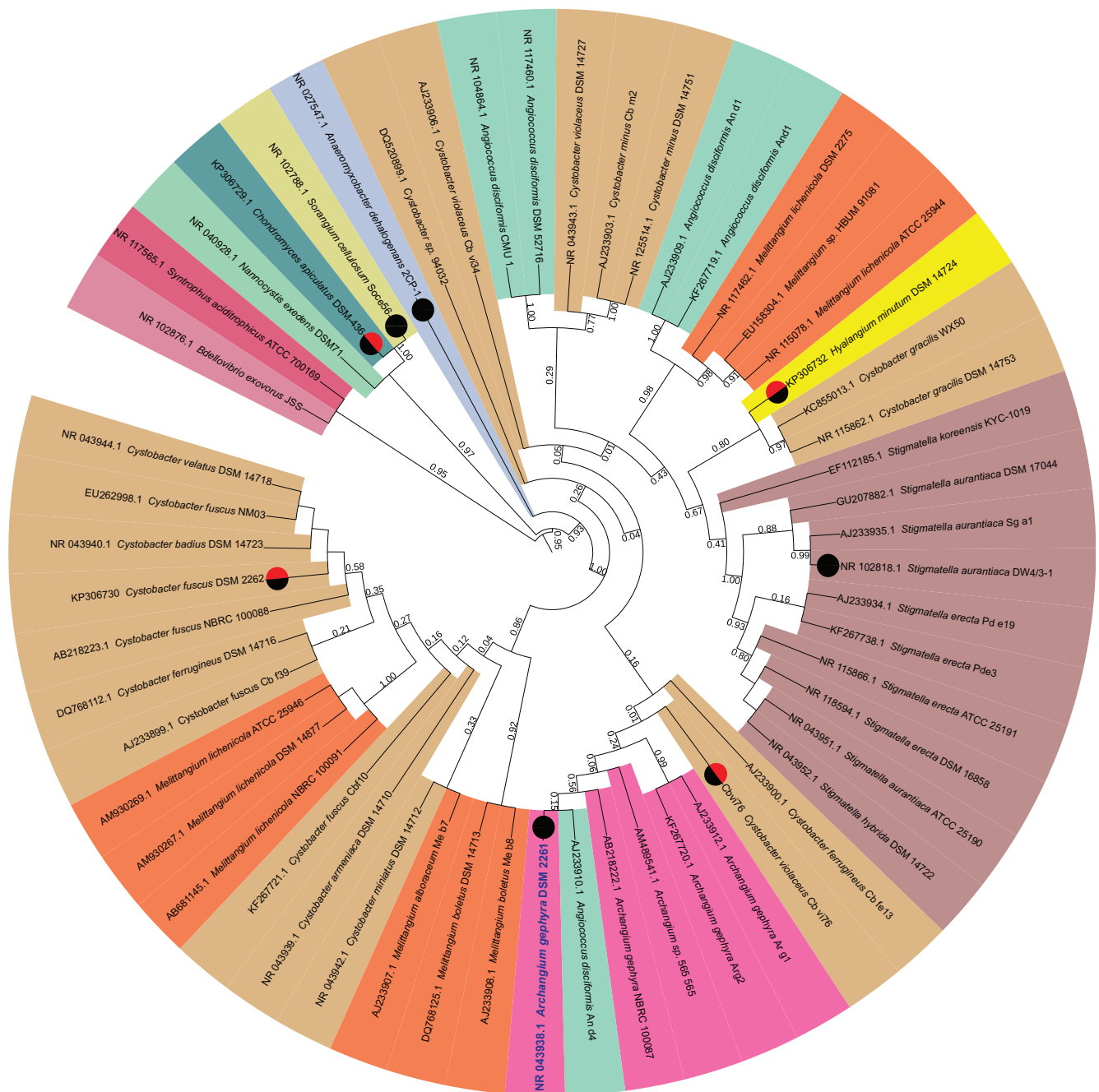
The genome of *A. gephyra* DSM 2261<sup>T</sup> has three operons of identical 16S rRNA sequences (location on genome: 110,548–112,085, 1,059,494–1,061,031, and 5,230,715–5,232,252 bp). The 16S rRNA sequences of *A. gephyra* show maximum identity with 16S rRNA sequence of *C. violaceus* Cbvi76 and *Angiococcus disciformis* An-d4 strain (98.64%) followed by *C. ferrugineus* Cb fe13 (98.44%). Phylogenetic analysis also reveals that it forms a separate clade along with other *Archangium* strains, *C. violaceus* Cbvi76, *Angiococcus disciformis* An-d4, and *C. ferrugineus* Cb fe13 (fig. 2). Family Cystobacteraceae members are well known for chitinase activity and the present species determination is based on this along with morphological features (Reichenbach 2005; dos Santos et al. 2014). Based on this, family Cystobacteraceae has been distributed in three groups; the representative species for each group are *C. fuscus* (efficient degrader), *C. disciformis* (partial degrader), and *C. gracilis* (nondegrader) (dos Santos et al. 2014). These three groups form separate clades in the phylogenetic tree (fig. 2) as well as earlier studies (Sproer et al. 1999). *Archangium gephyra* DSM 2261<sup>T</sup> along with other *Archangium* strains falls in the close vicinity of *C. violaceus*, *C. ferrugineus*, and *Angiococcus disciformis*, which represent the second group based on chitinase activity. These second group members, as well as *A. gephyra* can only partially degrade chitin when compared with the efficient chitin degraders of the first group. Recently, it has been proposed to reclassify the above mentioned close relatives of *A. gephyra*, that is, *C. violaceus* and *Angiococcus disciformis* as species within genus *Archangium* (*A. violaceus* and *A. disciformis*, respectively) owing to its phylogenetic and morphological distinctness (Lang et al. 2015). Phylogeny based on 28 conserved housekeeping gene sequences (Wu and Eisen 2008; Sharma et al. 2016b) also suggests the same (supplementary fig. S1, Supplementary Material online). Besides phylogeny analysis, we also calculated *in silico* DDH values of *A. gephyra* DSM 2261<sup>T</sup> compared with all available myxobacterial genomes using Genome-To-Genome Distance Calculator server (Auch et al. 2010). We found that among the family Cystobacteraceae genomes, it shows maximum DDH value of 46 with *C. violaceus* Cbvi76 followed by 25.1 DDH value for *C. fuscus* DSM 2262<sup>T</sup>. These results are in accordance with the phylogenetic analysis.

### DNA Methylome Analysis

Using the complete genome as a reference, methylome of the *A. gephyra* DSM 2261<sup>T</sup> genome was determined using base modifications and enriched motifs identification protocol in SMRT portal. The Open Reading Frame (ORF) coding for deoxyadenosine methyltransferase (dam) (AG\_08591; having TIGR00571) was identified in *A. gephyra* genome, which functions as m6A methylase. Type I methylases (specific to Adenine) were also identified through Type I R&M system analysis (AG\_05605, AG\_05606, AG\_07205, AG\_08251, and AG\_08454; having N6\_Mtase [PF02384]) whereas Type II methylase corresponding to Type II R&M systems were not found in the genome. The m6A-based methylation was observed in >96% of the motifs AACNNNNNNCTGG, CCAGNNNNNGTT, CCAANNNNNNNCTC, GAGNNNNNNNTTGG and CCCGCA in DSM 2261<sup>T</sup> genome at second, third, fourth, second and sixth positions, respectively. We also identified m4C-based methylation in motif GAGCTC at C4 position. We could not detect corresponding N4-methylcytosine (m4C) methylase homolog while other methylases having Pfam domain N6\_N4\_Mtase (PF01555), which could function as both N-4 cytosine-specific and N-6 Adenine-specific DNA methylase, were identified in DSM 2261<sup>T</sup> genome (AG\_00686, AG\_00776, AG\_03646 and AG\_07508).

### Protein Homology and Orthology Studies

Five myxobacterial genomes are available for the family Cystobacteraceae (supplementary table S3, Supplementary Material online). Out of these five, *Stigmatella* has been widely studied for its secondary metabolites, motility, and sporulation (Neumann et al. 1993; Silakowski et al. 1999, 2000; Tan et al. 2013; Kunze et al. 2005). The aim of this study was to identify the set of commonly shared and unique genes among the family Cystobacteraceae genomes. These members have >9 Mbp genome size and 8,500–10,000 proteins. Homology studies suggest that 51% of the total proteins are well conserved across these organisms (supplementary fig. S2, Supplementary Material online). We also identified several proteins in different genome combinations such as 10.61% proteins are present in two genome combinations, 10.26% in three genome combinations, and 12% in four genome combinations. Out of total proteins, 16% proteins are present as unique proteins among these five organisms. The presence of a large number of unique proteins in each genome suggests wide differences in terms of genomic diversity at the genus level. Orthology studies identified the presence of 3,368 clusters conserved across all 5 organisms (supplementary fig. S2, Supplementary Material online). These clusters account for 33–40% of each proteome and these proteins assist in homeostasis, housekeeping functions and in maintaining morphological, developmental, and physiological features of the organism. Based on one-to-one orthology studies, we identified 70.55% (*Ag*), 59.57% (*Cyb*), 69.16% (*Cyvi*), 63.90% (*Hm*), and 63.95% (*Sa*) proteins within respective



**Fig. 2.**—16S rRNA-based phylogenetic tree highlighting the position of *A. gephyra* DSM 2261<sup>T</sup>. The 16S rRNA sequences from family Cystobacteraceae as well as from other outgroup strains have respective NCBI nucleotide ids and organism names. The black circle represents complete genomes and red-black circle shows the draft genomes. *Archangium gephyra* is depicted in violet bold fonts. Genus-wise colors have been used for visual demarcation. Bootstrap values are provided corresponding to the tree nodes.

genomes. The relative patterns of homology and orthology studies are consistent with the 16S rRNA tree.

### CAZome Analysis

We did not find any correlation between genome size and CAZome. In *Ag* genome, 391 proteins (3.9%) were identified

with a different combination of CAZY domains, putatively functioning as carbohydrate active enzymes. When compared with other family Cystobacteraceae members, *Cyb* has the largest CAZome with 6% proteins whereas *Cyvi* has the lowest (3.2%) (supplementary table S4, Supplementary Material online). In spite of being the smallest genome (10.26

Mbp) among the family, *Sa* has 5.2% carbohydrate active enzymes. Although these genomes are about the same size and contain a comparable number of proteins, *Cyb* CAZome is two times as large as that of *Ag* and *Cyvi*; which is reflected in all CAZY categories.

We also identified specific CAZY proteins (supplementary table S5, Supplementary Material online) known to degrade chitin, cellulose, starch, and agar. Among the chitinases, GH18 is a well-known family (Funkhouser and Aronson 2007) and we find that *Cyb* and *Sa* have multiple GH18 domains. Other genomes such as *Ag* and *Hm* have just a single copy whereas we could not find any GH18 in *Cyvi*. The distribution of GH18 homologs suggests that *Cyb* and *Sa* are likely to degrade chitin efficiently when compared with *Ag*, *Hm*, and *Cyvi*. Family Cystobacteraceae members do not have any GH96 agarase, consistent with their agar nondegrading behavior (dos Santos et al. 2014). Cellulose degradation has not been reported in family Cystobacteraceae members (dos Santos et al. 2014); however, *Cyb* has all three types of cellulases ( $\beta$ -glucosidase, cellobiohydrolase/exoglucanase, and endoglucanase) in multiple copies and may thus be able to degrade cellulose. Exoglucanase is only present in *Cyb* (as GH48 homolog) and absent in all others. Among the amylases,  $\beta$ -amylase does not have any representation in any of the family Cystobacteraceae genomes. GH13 and GH15 domains representing  $\alpha$ - and  $\gamma$ -amylase, respectively, have been identified in multiple copies within the studied genomes ( $\sim 15$  and  $\sim 5$ , respectively), which suggest that they may be able to partially degrade starch.

### Evolution of Chitinases among Myxobacteria

*Archangium gephyra* has been characterized as a chitin nondegrader (Reichenbach 2005; dos Santos et al. 2014; Awal et al. 2017) whereas 16S rRNA and housekeeping phylogeny along with genome–genome distance depict its closeness with partial chitin degraders such as *C. violaceus* and *C. disciformis* (fig. 2; supplementary fig. S1, Supplementary Material online). To find the presence of chitinases in *Ag* proteome, we looked for GH18 CAZY family that is widely present in archaea, prokaryotes, and eukaryotes. We identified one protein, AG\_02459, which contains the GH18 module as well as the "[LIVMFY][DN]G[LIVMF][DN][LIVMF][DN].E" chitinase signature motif. We identified multiple chitinase homologs in *Cyb* draft assembly and *Sa* complete genome (eight homologs each), which likely enable them to degrade chitin efficiently when compared with *Ag*, which has just one GH18 chitinase. The presence of multiple chitinases has been reported earlier to be correlated with the efficiency of chitin degradation (Svitil et al. 1997; Beier and Bertilsson 2013; Bai et al. 2016). Moreover, we extended this study to all genomes of the order Myxococcales and identified all myxobacterial GH18 homologs. We found that all members of family Myxococcaceae such as *M. xanthus* species (*MxDK1622*,

*MxDZ2*, and *MxDZF1*) (Muller et al. 2013a, 2013b; Goldman et al. 2006), *M. hansupus* (Sharma et al. 2016b), *M. fulvus* (Sukharnikov et al. 2011), *M. stipitatus* (Huntley et al. 2013), and *Coralloccoccus corraloides* have five to six GH18 homologs, although these organisms have not been reported to have chitinase activity (Garcia and Müller 2014a). Other *Sorangium* and *Chondromyces* genomes also have two to three GH18 homologs. It has been reported that *Sorangium* species degrade chitin whereas *Chondromyces* species do not (Garcia and Müller 2014b). We generated a maximum likelihood phylogenetic tree using all putative GH18 homologs and their top non-Myxococcales BLAST hits as identified via BLASTp search against nonredundant NCBI database (supplementary fig. S3, Supplementary Material online). The tree depicts the evolutionary relationships of myxobacterial chitinases. All clades are clearly demarcated family and suborder wise. The closest identified homologs of most of the clades have representation from diverse taxa such as *Bacilli*, *Actinobacteria*, *Chloroflexi*, and *Deinococci*, which belong to the *Terrabacteria* group. Among all clades, the one containing the *Ag* GH18 homolog has the maximum closeness to GH18 proteins from chitin nondegrading *Acidobacteria* (Kielak et al. 2016). This suggests that although most of the myxobacterial chitinases share affinities with the *Terrabacteria* and *Proteobacteria* groups, the *Ag* GH18 chitinase was possibly acquired via horizontal gene transfer from *Acidobacteria*. It has been earlier reported that *Acidobacteria* have GH18 homologs but they do not show any chitinase activity (Rawat et al. 2012) suggesting the similar nature of *Archangium*.

### Supplementary Material

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

### Authors' Contributions

G.S. isolated genomic DNA, carried out strain identification, performed assembly and annotation and comparative analysis. G.S. and S.S. wrote the manuscript. Both authors have read and approved the manuscript.

### Acknowledgments

This work is supported by a project "Expansion and modernization of Microbial Type Culture Collection and Gene Bank (MTCC) jointly supported by Council of Scientific and Industrial Research (CSIR) Grant No. BSC0402 and Department of Biotechnology (DBT) Govt. of India Grant No. BT/PR7368/INF/22/177/2012". We thank MTCC for providing help in procuring the cultures from DSMZ. We also thank Dr Ramya TNC for providing us workplace to purify the genomic DNA. G.S. acknowledges CSIR for the research fellowship. We

thank the Génome Québec Innovation Centre, McGill University, Canada for help in obtaining PacBio NGS data.

## Literature Cited

- Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215:403–410.
- Auch AF, von Jan M, Klenk HP, Goker M. 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci.* 2:117–134.
- Awal RP, Garcia R, Gemperlein K, Wink J, Kunwar B, Parajuli N, Müller R. 2017. *Vitiosangium cumulatum* gen. nov., sp. nov. and *Vitiosangium subalbum* sp. nov., novel soil myxobacteria from Nepal, and emended descriptions of genus *Archangium* and *Angiococcus*, and of *Cystobacteraceae* family. *Int J Syst Evol Microbiol.* doi: 10.1099/ijsem.0.001829.
- Aziz RK, et al. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75.
- Bai Y, Eijsink VG, Kielak AM, van Veen JA, de Boer W. 2016. Genomic comparison of chitinolytic enzyme systems from terrestrial and aquatic bacteria. *Environ Microbiol.* 18:38–49.
- Beier S, Bertilsson S. 2013. Bacterial chitin degradation—mechanisms and ecophysiological strategies. *Front Microbiol.* 4:149.
- Chin C-S, et al. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Meth.* 10:563–569.
- dos Santos D, et al. 2014. The family *Cystobacteraceae*. In: Rosenberg E, DeLong E, Lory S, Stackebrandt E, Thompson F, editors. *The prokaryotes*. Berlin, Heidelberg: Springer. p. 19–40.
- Finn RD, et al. 2014. Pfam: the protein families database. *Nucleic Acids Res.* 42:D222–D230.
- Funkhouser JD, Aronson NN Jr. 2007. Chitinase family GH18: evolutionary insights from the genomic history of a diverse protein family. *BMC Evol Biol.* 7:96.
- García R, Müller R. 2014a. The family *Myxococcaceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. *The prokaryotes: Deltaproteobacteria and Epsilonproteobacteria*. Berlin, Heidelberg: Springer. p. 191–212.
- García R, Müller R. 2014b. The family *Polyangiaceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. *The prokaryotes: Deltaproteobacteria and Epsilonproteobacteria*. Berlin, Heidelberg: Springer. p. 247–279.
- Goldman BS, et al. 2006. Evolution of sensory complexity recorded in a myxobacterial genome. *Proc Natl Acad Sci U S A.* 103:15200–15205.
- Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35:W52–W57.
- Han K, et al. 2013. Extraordinary expansion of a *Sorangium cellulosum* genome from an alkaline milieu. *Sci Rep.* 3:2101.
- Huntley S, et al. 2011. Comparative genomic analysis of fruiting body formation in *Myxococcales*. *Mol Biol Evol.* 28:1083–1097.
- Huntley S, Kneip S, Treuner-Lange A, Sogaard-Andersen L. 2013. Complete genome sequence of *Myxococcus stipitatus* strain DSM 14675, a fruiting myxobacterium. *Genome Announc.* 1:e0010013.
- Jahn E. 1924. Beiträge zur botanischen Protistologie. I. Die Polyangiden. Leipzig: Verlag von Gebrüder Borntraeger.
- Kall L, Krogh A, Sonnhammer EL. 2007. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Res.* 35:W429–W432.
- Kielak AM, Barreto CC, Kowalchuk GA, van Veen JA, Kuramae EE. 2016. The ecology of Acidobacteria: moving beyond genes and genomes. *Front Microbiol.* 7:744.
- Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 305:567–580.
- Kunze B, Reichenbach H, Müller R, Hofle G. 2005. Aurafuron A and B, new bioactive polyketides from *Stigmatella aurantiaca* and *Archangium gephyra* (Myxobacteria). Fermentation, isolation, physico-chemical properties, structure and biological activity. *J Antibiot. (Tokyo)* 58:244–251.
- Lagesen K, et al. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.
- Lang E, Schumann P, Tindall BJ, Mohr KI, Sproer C. 2015. Reclassification of *Angiococcus disciformis*, *Cystobacter minus* and *Cystobacter violaceus* as *Archangium disciforme* comb. nov., *Archangium minus* comb. nov. and *Archangium violaceum* comb. nov., unification of the families *Archangiaceae* and *Cystobacteraceae*, and emended descriptions of the families *Myxococcaceae* and *Archangiaceae*. *Int J Syst Evol Microbiol.* 65(11):4032–4042.
- Lechner M, et al. 2011. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* 12:124.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42:D490–D495.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25:955–964.
- McCurdy HD. 1969. Studies on the taxonomy of the Myxobacteriales. I. Record of Canadian isolates and survey of methods. *Can J Microbiol.* 15:1453–1461.
- Müller S, Willett JW, Bahr SM, Darnell CL, Hummels KR, Dong CK, Vlamakis HC, Kirby JR. 2013a. Draft genome sequence of *Myxococcus xanthus* wild-type strain DZ2, a model organism for predation and development. *Genome Announc.* 1(3):e00217–13.
- Müller S, Willett JW, Bahr SM, Scott JC, Wilson JM, Darnell CL, Vlamakis HC, Kirby JR. 2013b. Draft genome of a Type 4 Pilus defective *Myxococcus xanthus* strain, DZF1. *Genome Announc.* 1(3):e00392–13.
- Neumann B, Pospiech A, Schairer HU. 1993. A physical and genetic map of the *Stigmatella aurantiaca* DW4/3.1 chromosome. *Mol Microbiol.* 10:1087–1099.
- Rawat SR, Männistö MK, Bromberg Y, Häggblom MM. 2012. Comparative genomic and physiological analysis provides insights into the role of Acidobacteria in organic carbon utilization in Arctic tundra soils. *FEMS Microbiol Ecol.* 82:341–355.
- Reichenbach H. 2005. Order VIII. *Myxococcales*. In: Brenner DJ, Krieg NR, Staley JT, Garrity GM, editors. *Bergey's manual of systematic bacteriology*. Vol. 2, 2nd ed part C. New York: Springer-Verlag. p. 1074–1079.
- Reichenbach H. 1965. [Studies on *Archangium violaceum*. Contribution to the study of myxobacteria]. *Arch Mikrobiol.* 52:376–403.
- Sanford RA, Cole JR, Tiedje JM. 2002. Characterization and description of *Anaeromyxobacter dehalogenans* gen. nov., sp. nov., an aryl-halorespiring facultative anaerobic *Myxobacterium*. *Appl Environ Microbiol.* 68:893–900.
- Schneider S, et al. 2007. Complete genome sequence of the myxobacterium *Sorangium cellulosum*. *Nat Biotechnol.* 25:1281–1289.
- Sharma G, Khatri I, Subramanian S. 2016a. Complete genome of the starch-degrading myxobacteria *Sandaracinus amylolyticus* DSM. *Genome Biol Evol.* 8(8):2520–2529.
- Sharma G, Narwani T, Subramanian S. 2016b. Complete Genome Sequence and Comparative Genomics of a Novel Myxobacterium *Myxococcus hansupus*. *PLoS One* 11(2):e0148593.
- Siguié P, Perochon J, Lestrade L, Mahillon J, Chandler M. 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 34:D32–D36.



- Silakowski B, et al. 2000. The myxochelin iron transport regulon of the myxobacterium *Stigmatella aurantiaca* Sg a15. *Eur J Biochem.* 267:6476–6485.
- Silakowski B, et al. 1999. New lessons for combinatorial biosynthesis from myxobacteria. The myxothiazol biosynthetic gene cluster of *Stigmatella aurantiaca* DW4/3-1. *J Biol Chem.* 274:37391–37399.
- Sproer C, Reichenbach H, Stackebrandt E. 1999. The correlation between morphological and phylogenetic classification of myxobacteria. *Int J Syst Bacteriol.* 49(Pt 3):1255–1262.
- Stevens DC, Young J, Carmichael R, Tan J, Taylor RE. 2014. Draft genome sequence of gephyronic acid producer *Cystobacter violaceus* strain Cb vi76. *Genome Announc.* 2(6):e01299–14.
- Sukharnikov LO, Cantwell BJ, Podar M, Zhulin IB. 2011. Cellulases: ambiguous nonhomologous enzymes in a genomic perspective. *Trends Biotechnol.* 29:473–479.
- Svitil AL, Chadhain S, Moore JA, Kirchman DL. 1997. Chitin degradation proteins produced by the marine bacterium *Vibrio harveyi* growing on different forms of chitin. *Appl Environ Microbiol.* 63:408–413.
- Tan Z, et al. 2013. Characterization of four type IV pilin homologues in *Stigmatella aurantiaca* DSM17044 by heterologous expression in *Myxococcus xanthus*. *PLoS One* 8:e75105.
- Thomas SH, et al. 2008. The mosaic genome of *Anaeromyxobacter dehalogenans* Strain 2CP-C suggests an aerobic common ancestor to the delta-proteobacteria. *PLoS One* 3:e2103.
- Wu M, Eisen JA. 2008. A simple, fast, and accurate method of phylogenomic inference. *Genome Biol.* 9:R151.
- Yin Y, et al. 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 40:W445–W451.

**Associate editor:** Howard Ochman