RESEARCH　　　　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

Expanding magnetic organelle biogenesis in the domain *Bacteria*

Wei Lin[1,2,3]* , Wensi Zhang[1,2,3,4], Greig A. Paterson[5], Qiyun Zhu[6], Xiang Zhao[7], Rob Knight[6], Dennis A. Bazylinski[8], Andrew P. Roberts[7] and Yongxin Pan[1,2,3,4]*

Abstract

Background: The discovery of membrane-enclosed, metabolically functional organelles in *Bacteria* has transformed our understanding of the subcellular complexity of prokaryotic cells. Biomineralization of magnetic nanoparticles within magnetosomes by magnetotactic bacteria (MTB) is a fascinating example of prokaryotic organelles. Magnetosomes, as nano-sized magnetic sensors in MTB, facilitate cell navigation along the local geomagnetic field, a behaviour referred to as magnetotaxis or microbial magnetoreception. Recent discovery of novel MTB outside the traditionally recognized taxonomic lineages suggests that MTB diversity across the domain *Bacteria* are considerably underestimated, which limits understanding of the taxonomic distribution and evolutionary origin of magnetosome organelle biogenesis.

Results: Here, we perform the most comprehensive metagenomic analysis available of MTB communities and reconstruct metagenome-assembled MTB genomes from diverse ecosystems. Discovery of MTB in acidic peatland soils suggests widespread MTB occurrence in waterlogged soils in addition to subaqueous sediments and water bodies. A total of 168 MTB draft genomes have been reconstructed, which represent nearly a 3-fold increase over the number currently available and more than double the known MTB species at the genome level. Phylogenomic analysis reveals that these genomes belong to 13 Bacterial phyla, six of which were previously not known to include MTB. These findings indicate a much wider taxonomic distribution of magnetosome organelle biogenesis across the domain *Bacteria* than previously thought. Comparative genome analysis reveals a vast diversity of magnetosome gene clusters involved in magnetosomal biogenesis in terms of gene content and synteny residing in distinct taxonomic lineages. Phylogenetic analyses of core magnetosome proteins in this largest available and taxonomically diverse dataset support an unexpectedly early evolutionary origin of magnetosome biomineralization, likely ancestral to the origin of the domain *Bacteria*.

Conclusions: These findings expand the taxonomic and phylogenetic diversity of MTB across the domain *Bacteria* and shed new light on the origin and evolution of microbial magnetoreception. Potential biogenesis of the magnetosome organelle in the close descendants of the last bacterial common ancestor has important implications for our understanding of the evolutionary history of bacterial cellular complexity and emphasizes the biological significance of the magnetosome organelle.

Keywords: Magnetotactic bacteria, Magnetosome, Magnetotaxis, Prokaryotic organelle, Last bacterial common ancestor

---

* Correspondence: weilin0408@gmail.com; yxpan@mail.iggcas.ac.cn
[1]Key Laboratory of Earth and Planetary Physics, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China
Full list of author information is available at the end of the article

Background

It was accepted widely that intracellular, membrane-bounded, metabolically functional organelles are present exclusively in eukaryotic cells and that they are absent from *Bacteria* and *Archaea*. This long-held view was revised after numerous recent discoveries of a diverse group of highly organized, membrane-enclosed organelles in the domains *Bacteria* and *Archaea* associated with specific cellular functions [1–3]. However, the taxonomic distribution, origin and evolution of prokaryotic organelles remain largely elusive. It is still unclear whether organelle biogenesis emerged early or late during the evolution of *Bacteria* and *Archaea*, posing problems for elucidating the evolutionary history of cellular complexity.

Magnetosomes within magnetotactic bacteria (MTB) are a striking example of prokaryotic organelles [4]. Magnetosomes consist of a lipid bilayer-bounded membrane in which nanosized, ferrimagnetic magnetite (Fe_3O_4) and/or greigite (Fe_3S_4) crystals are biomineralized and are usually arranged in chain-like structure(s) that maximize the magnetic dipole moment [5, 6] (Fig. 1). The most accepted major function of magnetic magnetosomal crystals is to produce tiny compass needles that facilitate MTB navigation to their preferred low-O_2 or anaerobic microenvironments in chemically stratified aquatic systems, a behaviour referred to as magnetotaxis or microbial magnetoreception [7]. Additional suggested functions of magnetosomal crystals include detoxification/elimination of toxic reactive oxygen species (ROS), iron sequestration and storage in which they act as an electrochemical battery, or as a gravity sensor [8–10] (Fig. 1a). Understanding the phylogenetic and genomic diversity of MTB could advance our understanding of the evolutionary origin of bacterial organelle biogenesis in general. Moreover, considering that magnetoreception occurs widely in both micro- and macro-
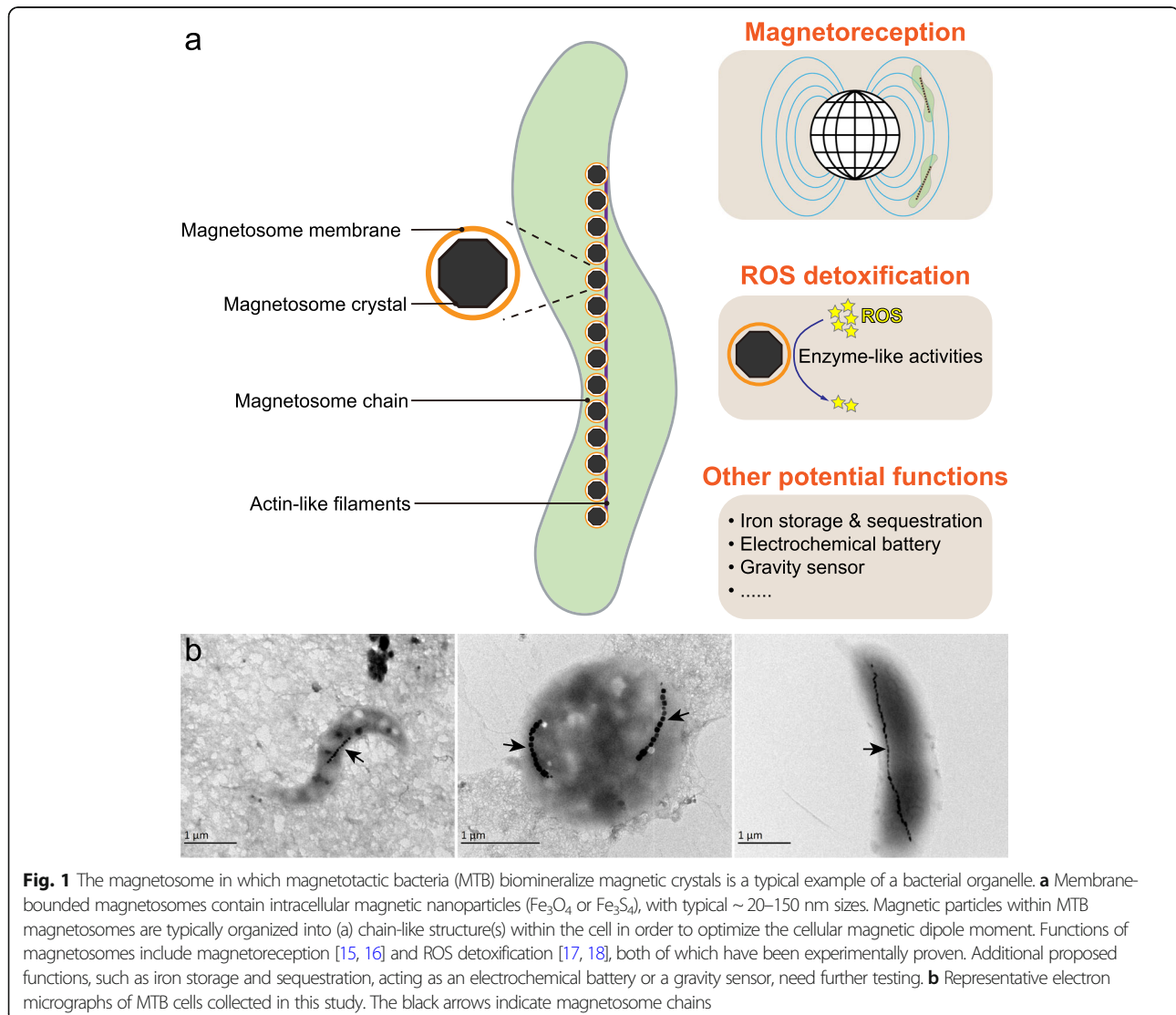

Fig. 1 The magnetosome in which magnetotactic bacteria (MTB) biomineralize magnetic crystals is a typical example of a bacterial organelle. **a** Membrane-bounded magnetosomes contain intracellular magnetic nanoparticles (Fe_3O_4 or Fe_3S_4), with typical ~ 20–150 nm sizes. Magnetic particles within MTB magnetosomes are typically organized into (a) chain-like structure(s) within the cell in order to optimize the cellular magnetic dipole moment. Functions of magnetosomes include magnetoreception [15, 16] and ROS detoxification [17, 18], both of which have been experimentally proven. Additional proposed functions, such as iron storage and sequestration, acting as an electrochemical battery or a gravity sensor, need further testing. **b** Representative electron micrographs of MTB cells collected in this study. The black arrows indicate magnetosome chains

organisms and that magnetosomal crystals are the only magnetoreceptors definitively characterized thus far, MTB also represent a valuable system for exploring the origin and early evolution of magnetoreception [11–14].

MTB are distributed globally across a broad range of O_2-limited or anaerobic aquatic habitats, ranging from freshwater lakes to oceans and even to some extreme environments [19–21]. However, few studies have reported MTB in waterlogged soils [22, 23]. Magnetosomal biogenesis by MTB is recognized as a key component of the global iron cycle [24, 25], and MTB cells are potentially important in the biogeochemical cycling of carbon, phosphorus, nitrogen, and sulphur [9, 26–28]. Historically, magnetosomal biomineralization has been viewed as a specialized type of metabolism restricted to two Bacterial phyla: the *Proteobacteria* (*Alphaproteobacteria*, *Deltaproteobacteria*, and *Gammaproteobacteria* classes) and the *Nitrospirae* [29]. Initial attempts to explain this restricted yet scattered distribution of magnetosomal biogenesis based on a few taxa gave rise to two alternative hypotheses: polyphyletic origin in different taxonomic lineages [30], or extensive horizontal gene transfers (HGTs) [31, 32].

In recent years, the known extent of MTB diversity has undergone a significant expansion due to methodological advances, such as the successful cultivation of novel MTB strains [33–36], 16S rRNA gene-based characterization [37–39], genome mining of public repositories such as GenBank and IMG/ER [21, 40], and cultivation-independent surveys of magnetosome gene clusters (MGCs, which are physically clustered groups of genes that are together responsible for magnetosomal biogenesis) containing genomes [41, 42]. MTB have now been further identified in other Bacterial taxa, including the *Betaproteobacteria*, *Zetaproteobacteria*, "*Candidatus* Etaproteobacteria", and "*Candidatus* Lambdaproteobacteria" classes of the *Proteobacteria* phylum, the candidate phylum *Omnitrophica* (previously known as the candidate division OP3), the candidate phylum *Latescibacteria* (previously known as the candidate division WS3), and the phylum *Planctomycetes*, according to the NCBI taxonomy. Analyses including data from these latter groups suggest that magnetosomal biogenesis among different Bacterial lineages has a monophyletic origin from a common ancestor, which occurred prior to divergence of the *Nitrospirae* and *Proteobacteria* phyla, or perhaps even earlier, in the last common ancestor of five MTB-containing Bacterial phyla: *Proteobacteria*, *Nitrospirae*, *Omnitrophica*, *Latescibacteria*, and *Planctomycetes* [19, 41, 43–45].

Discovery of novel MTB outside the traditionally recognized taxonomic lineages suggests that the diversity and distribution of MTB across the domain *Bacteria* are considerably underestimated. This raises important questions regarding the taxonomic distribution and evolutionary origin of magnetosome organelle biosynthesis. Here we present the most comprehensive metagenomic analysis available of MTB communities from geographically, physically and chemically diverse sites to evaluate two questions: (1) Are MTB taxonomically widespread among the phyla of the domain *Bacteria*? (2) Did the magnetosome organelle originate earlier than previously realized? Findings of this study provide novel insights into the evolutionary origins of prokaryotic organelles.

Results and discussion
Survey of MTB from diverse environments
We performed a light microscopy survey of MTB from wide-ranging environments across China and Australia (Fig. 2a, b), including sediments from freshwater lakes, ponds, rivers, creeks, paddy fields, and intertidal zones, and soils from acidic peatlands. MTB were observed microscopically in these habitats with a salinity range of < 0.1–37.0 ppt and a pH range of 4.3–8.6 (Supplementary Table 1). Unexpectedly, we find living MTB cells in acidic peatland soils (pH 4.3–5.7) with high water contents (> 60%) and organic matter contents (typically > 20%) [46–48]. MTB have been found broadly in diverse aquatic ecosystems, including some extreme environments such as hot springs [49, 50], saline-alkaline lakes [51], acidic lagoons and mine drainage systems [37, 52], and deep-sea sediments [53], but reports of MTB in soils are limited to a few studies that were published almost 30 years ago [22, 23]. Whether MTB can survive in waterlogged soil environments remains unresolved, and the taxonomic diversity of MTB in these environments has not been elucidated. Our finding represents, to the best of our knowledge, the first discovery of MTB in soils with relatively acidic pH. MTB in acidic peatland soils are represented mainly by deep-branching lineages such as the *Nitrospirae* and *Omnitrophica* phyla, which is markedly different from other environments (Fig. 3, discussed below). Our finding, together with the previous studies [22, 23], suggests potential widespread MTB occurrence in different types of waterlogged soils, in addition to subaqueous sediments and water bodies.

Expanded genomic diversity of MTB
Metagenomic DNA from magnetically enriched MTB was sequenced, and metagenome-assembled genomes were reconstructed with a single-sample assembly and binning strategy. Reconstructed genomes were checked manually for the presence of MGCs. From this analysis, we recover a total of 168 MGC-containing genomes with a quality score [54] above 50 (Supplementary Table 2). Of these putative MTB genomes, 69 (41%) are high-quality (with > 90% completeness and < 5% contamination), 64 (38%) are medium-quality (70–90 % complete
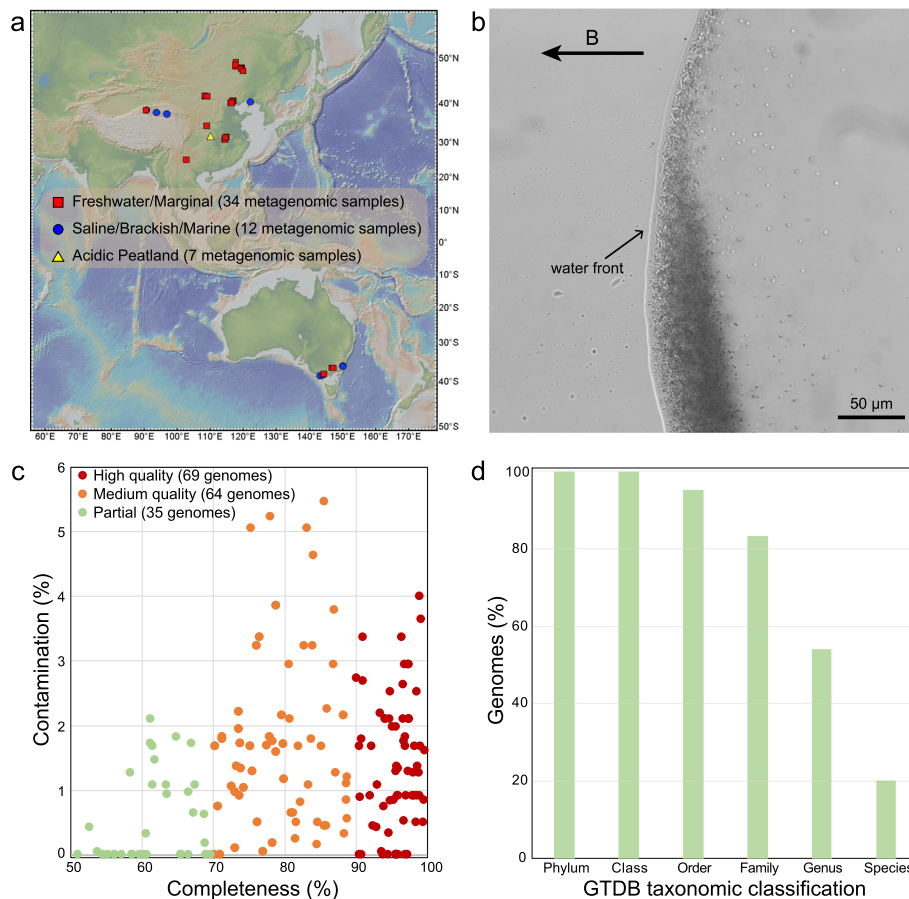
Fig. 2 Recovery of 168 MTB genomes from various environments. **a** Map of sampling locations (generated using the GeoMapApp 3.6.0, http://www.geomapapp.org/). Further site details are given in Supplementary Table 1. **b** A micrograph of MTB cells (cocci and rods) from Lake Dianchi, China, as observed under a light microscope (Olympus BX51, Olympus, Tokyo, Japan). The applied field (B) direction is from right to left. **c** Estimated completeness and contamination of MTB genomes reconstructed in this study. CheckM was used to estimate completeness and contamination. Of these genomes, 69 are high-quality (> 90% completeness and < 5% contamination), 64 are medium-quality (70–90% completeness and < 6% contamination), and 35 are partial (50–70% completeness and < 5% contamination) genomes. **d** Relative abundance of recovered MTB genomes that can be classified according to the GTDB taxonomy (database Release 04-RS89). Of the 168 recovered genomes, 34 were classified at the species level, 91 were classified at the genus level, 140 were classified at the family level, 160 were classified at the order level, and 168 could be classified at the class and phylum levels. Details are given in Supplementary Table 2

with < 6% contamination), and 35 (21%) are partial (with 50–70% completeness and < 5% contamination) genomes (Fig. 2c and Supplementary Table 2). These genomes increase substantially the known genomic diversity of MTB, expanded from 59 (as of January 2020, Supplementary Table 3) to 227 MTB genomes, which can now be clustered into 164 unique species-level genomes based on a 95% average nucleotide identity (ANI) [55]. Of these species, 110 (67%) are reported here for the first time.

The taxonomy of the newly acquired MTB genomes is classified using a standardized phylogenomic curated taxonomy system Genome Taxonomy Database (GTDB) Toolkit [56, 57]. Around 80% of the recovered genomes cannot be assigned at the species level, and more than

45% cannot be assignable at the genus level (Fig. 2d), which confirms that many of these genomes represent previously unknown populations. The phylogenetic relationships of these genomes were further determined by phylogenomic analysis. The reconstructed genome tree is congruent with the GTDB taxonomy as shown in Fig. 3a and Supplementary Figure 1; therefore, the GTDB taxonomy is used for taxonomic classification of MTB throughout, unless otherwise noted. The corresponding NCBI taxonomy of each lineage is also given in Fig. 3a.

The 168 genomes belong to organisms from 13 distinct Bacterial phyla as defined in the GTDB taxonomy and 7 phyla according to the NCBI taxonomy (Fig. 3a and Supplementary Table 2). These genomes are recovered from several previously poorly characterized MTB
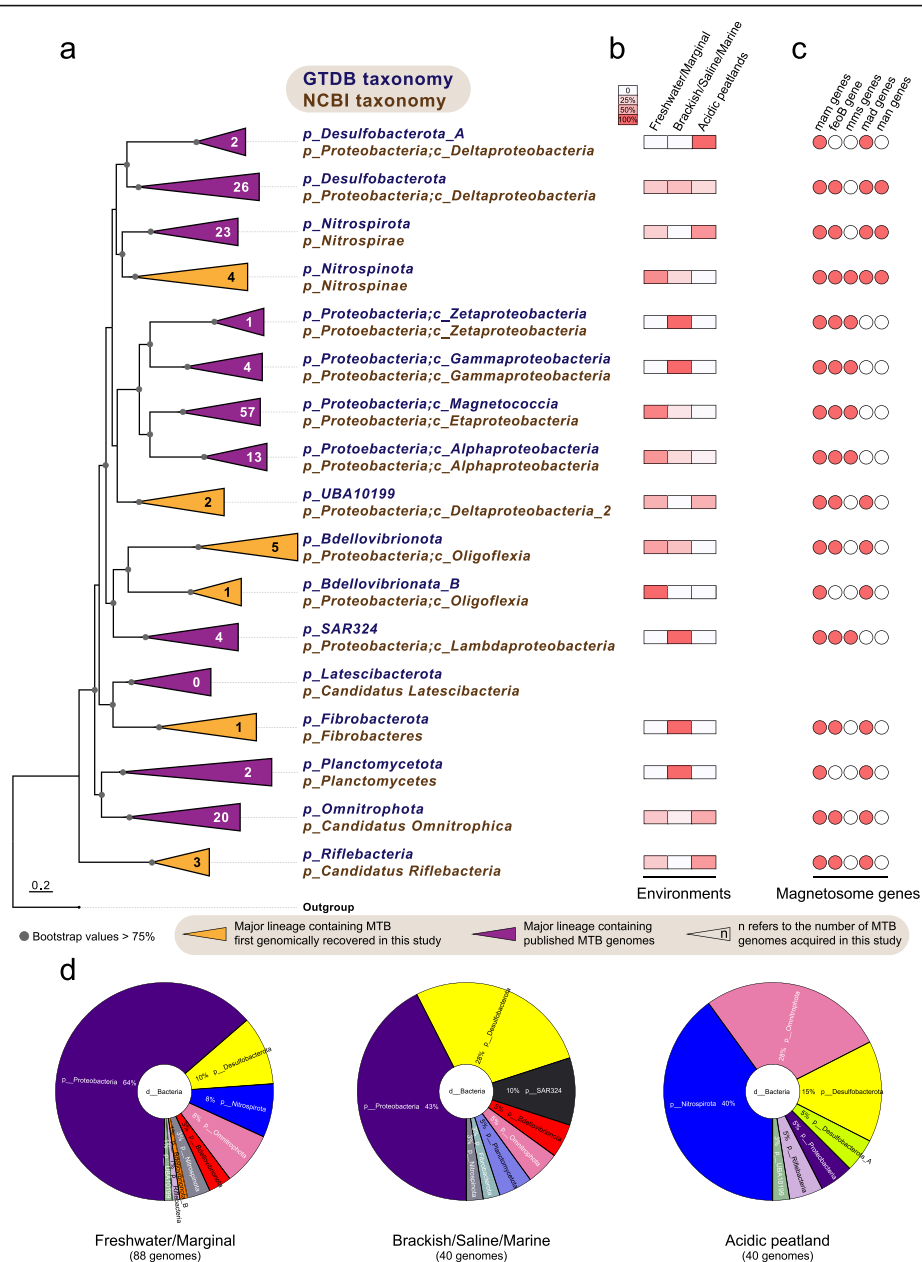
Fig. 3 Distribution of MTB genomes across Bacterial phyla and distinct environments. **a** The maximum-likelihood phylogenomic tree of MTB genomes and their close non-MTB relatives inferred from concatenated 120 bacterial single-copy marker proteins [54], which was constructed using IQ-TREE under the LG+I+G4 substitution model. The number in each clade refers to the number of MTB genomes reconstructed in this study. The complete tree is shown in Supplementary Figure 1. **b** Relative abundances of reconstructed MTB genomes in this study across different environments within each lineage. **c** Distribution of magnetosome genes (*mam*, *mms*, *mad*, and *man*) and *feoB* gene within MGCs across different lineages. **d** Distribution of acquired MTB genomes at the phylum level across different environments, including freshwater/marginal (< 1 ppt) and saline/brackish/marine (> 1 ppt) sediments, and soils from acidic peatland. Details are given in Supplementary Table 1

groups, including 26 genomes from the *Desulfobacterota* phylum, 23 genomes from the *Nitrospirota* phylum, and 20 genomes from the *Omnitrophota* phylum (Fig. 3a). These novel genomes also expand substantially the representation of common MTB lineages, such as the *Magnetococcia* (57 genomes) and *Alphaproteobacteria* (13 genomes) classes of the

Proteobacteria phylum. More importantly, we identify 16 genomes that are affiliated within 6 phyla that were not known until January 2020 to contain MTB, including the *Nitrospinota*, UBA10199, *Bdellovibrionota*, *Bdellovibrionata*_B, *Fibrobacterota*, and *Riflebacteria* phyla (Fig. 3a and Supplementary Figure 1). This expands greatly the number of Bacterial lineages

associated with magnetosomal biogenesis and magnetic navigation.

To deepen our understanding of the environmental distribution patterns of MTB, the taxonomic diversity of novel MTB genomes was compared across environments. Genomes belonging to the phyla *Proteobacteria* and *Desulfobacterota* are found predominantly in both freshwater/marginal (< 1 ppt) and saline/brackish/marine (> 1 ppt) sediments, whereas genomes affiliated with the *Nitrospirata* and *Omnitrophota* phyla represent the dominant MTB groups in acidic peatland soils (Fig. 3d). MTB from the *Proteobacteria*, *Desulfobacterota* and *Omnitrophota* exist in all three environmental sample types, while those of the *Desulfobacterota*_A phylum and of the SAR324, *Fibrobacterota* and *Planctomycetota* phyla are observed exclusively in acidic peatland soils and brackish/saline/marine environments, respectively (Fig. 3b). For *Proteobacteria* classes, MTB genomes from the *Magnetococcia* and *Alphaproteobacteria* are found in wide-ranging environments, while those with *Zetaproteobacteria* and *Gammaproteobacteria* genomes occur solely in brackish/saline/marine environments.

Diverse MGCs across distinct taxonomic lineages
Genes for the metabolic pathway responsible for magnetosomal biogenesis have been found in contiguous gene clusters in MTB genomes [58], which are referred to as MGCs [21]. These gene clusters are not only the key to deciphering the mechanisms and evolutionary origin of magnetosome formation and magnetotaxis [5], but they also provide a wealth of gene resources for biosynthesis of membrane-bounded, single-domain magnetic nanoparticles with diverse properties for various applications [59]. The genomes acquired here contain diverse MGC types in terms of gene content and synteny (Fig. 4), which expands our knowledge of MGC diversity considerably. Discovery of various MGCs suggests the potential for diverse magnetosomal biogenesis and magnetotaxis across the domain *Bacteria*. The *mam* [60–62] genes that play essential roles in magnetosomal biogenesis are present in all MGCs identified here. Remarkably, we note that the *feoB* gene, which is responsible for iron transport into the cell, is also shared by most MTB lineages and is usually included in the MGCs (Figs. 3c and 4). Deletion of *feoB* in *Magnetospirillum* strains results in reduced magnetite biomineralization [63, 64], which indicates its potentially significant role in magnetosomal biomineralization. The *mad* [65, 66] and *man* [67] genes, which have been proposed to play important accessory functions in magnetosomal biomineralization, represent a much wider distribution across MTB genomes than previously thought: *mad* genes are present in the genomes of 11 phyla and *man* genes occur in the genomes of the *Desulfobacterota* (nTS_bin18 and nDJH15_bin4),

Nitrospirota, and *Nitrospinota* (nPCR_bin9 and nNGH_bin12) phyla. The *mms6* operon, which contains magnetosome genes *mms6*, *mmsF*, *mms36*, and *mms48*, controls the size and/or number of magnetic magnetosomal crystals in *Magnetospirillum* strains [61, 62, 68, 69], and these genes have previously only been found in the *Proteobacteria* phylum. Here we find that MGCs from genomes of the phyla *Nitrospinota* (nNGH_bin12) and SAR324 (nKLK_bin6 and nPCR_bin7) also contain *mms* genes (*mms6* and/or *mmsF*). In general, the gene content and orientation of MGCs vary considerably across different taxonomic lineages but are generally conserved within the same lineage (Fig. 4), which is indicative of lineage specific MGC evolution without extensive interphylum HGTs and inter-class HGTs among the *Proteobacteria* phylum. Considering the potentially high metabolic cost of maintaining such complex gene clusters in MTB, the widely observed MGC distribution across different lineages suggests that magnetosomal biogenesis and magnetotaxis must confer selective advantages on these organisms.

Two types of magnetosomal mineral crystals have been identified to date: magnetite (Fe_3O_4) and greigite (Fe_3S_4). Some MTB biomineralize both minerals within the same cell [70–72]. Before this study, Fe_3O_4-type MGCs had been identified in all MTB lineages except for the *Latescibacterota* and *Planctomycetota* [21], while Fe_3S_4-type MGCs were only found in the *Desulfobacterota* (e.g., *Candidatus* Magnetoglobus multicellularis [73] and *Candidatus* Desulfamplus magnetomortis strain BW-1 [71]), *Latescibacterota* [40], and *Planctomycetota* [21] phyla. Here, most of the identified MGCs contain Fe_3O_4-type magnetosome genes (including the first identification of Fe_3O_4-type MGCs in the *Planctomycetota* phylum) (Fig. 4). Only a small fraction (nER2_bin1, nHLH_bin7, nN2-2_bin5, nS315_bin9, nS315_bin20, nS315_bin24, nTS_bin4, and nXX_bin1) harbor putative Fe_3S_4-type or both Fe_3O_4- and Fe_3S_4-type magnetosome genes, which suggests that Fe_3O_4-producing MTB are distributed more widely than Fe_3S_4-producing MTB in present-day habitats. The coexistence of Fe_3O_4- and Fe_3S_4-type MGCs in the same genomes of UBA10199 (nXX_bin1) and *Omnitrophota* (nS315_bin20 and nS315_bin24) phyla has never been reported previously, which not only indicates an unexpected phylogenetic diversity of Fe_3S_4-producing MTB, but, more importantly, challenges the traditional hypothesis regarding the origin of Fe_3S_4-producing MTB (discussed below). Thus, this study reveals that MTB in the domain *Bacteria* contain many more MGCs than anticipated, which captures a more complete picture of the genomic diversity of MTB. The diverse MGCs reconstructed here also represent a promising new gene resource for magnetic bio-nanoparticles that can be used to modify magnetosomal
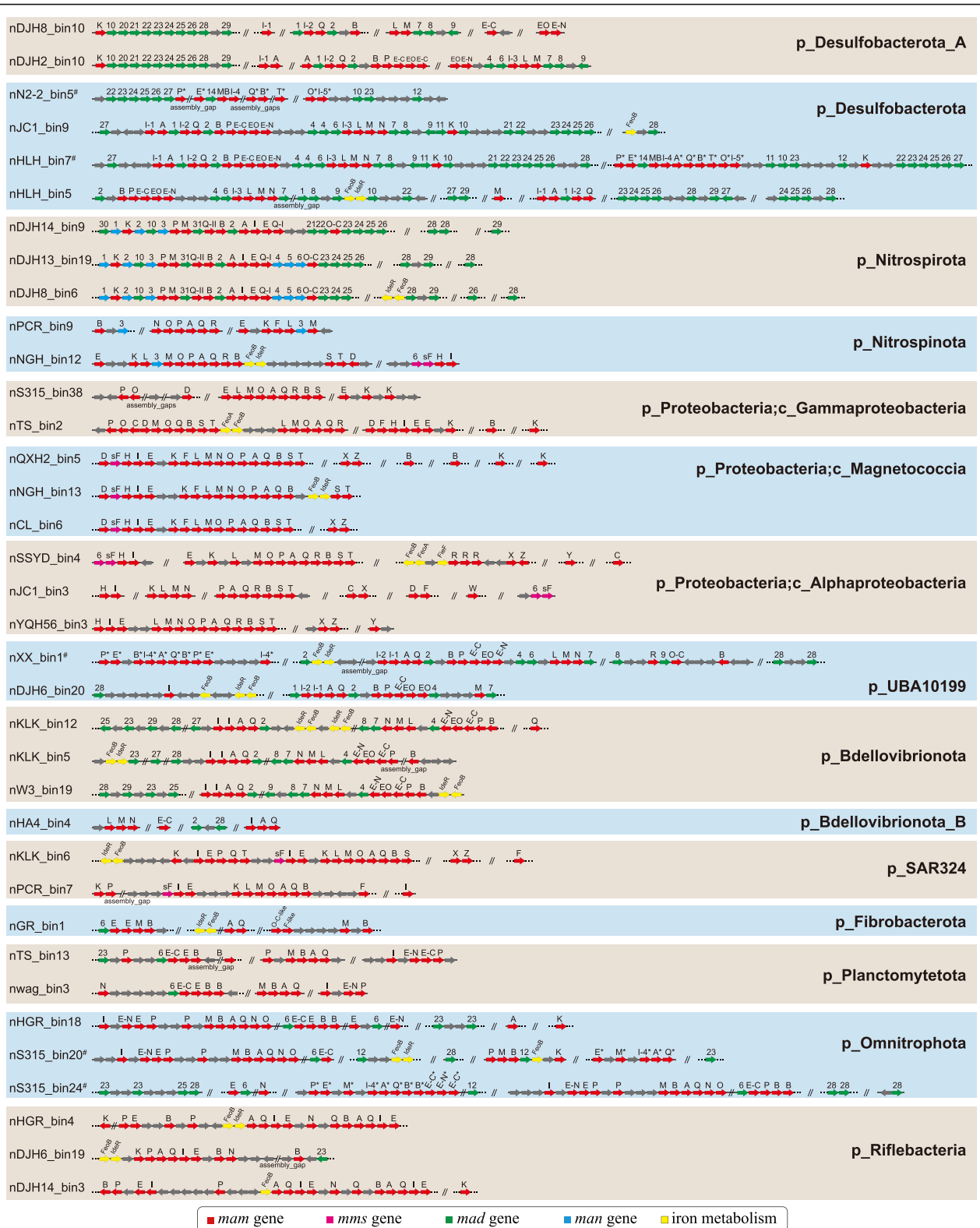
Fig. 4 Representative magnetosome gene clusters (MGCs) from distinct MTB lineages recovered in this study. Genomes containing Fe₃S₄-type MGCs are highlighted with # and putative Fe₃S₄-type magnetosome genes in MGCs are denoted by *
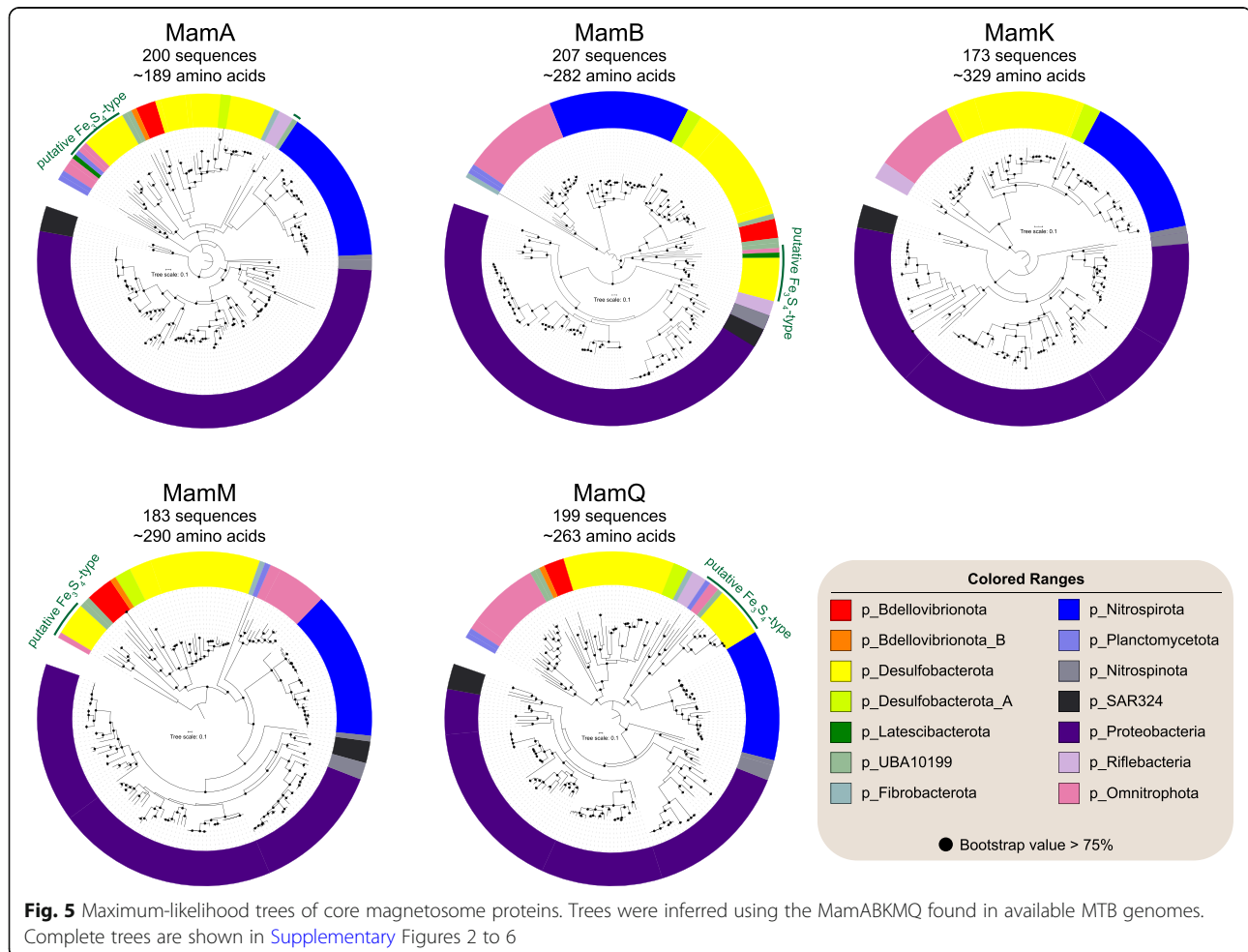
biogenesis pathways in MTB or even build new pathways in non-MTB [59].

Phylogenetic analyses of the core magnetosome proteins

A group of nine genes (*mamABEIKMOPQ*) was identified previously as the core magnetosome gene set shared by both Fe_3O_4- and Fe_3S_4-producing MTB [65]. The products of these genes are thought to have important functions in magnetosomal biomineralization and for magnetosome chain construction. To trace the evolutionary history of magnetosomal biogenesis, we first performed a comparative genome analysis of 83 representative high-quality MTB genomes (complete genomes and those with > 90% completeness and < 5% contamination) to identify core magnetosome genes that are defined such that > 90% of input genomes (i.e., ≥ 74 genomes) must contain these genes, which allows for missing or fragmented genes due to the incomplete nature of draft genomes. Six magnetosome genes (*mamABIKMQ*) meet these criteria. Among these genes, proteins encoded by *mamBIMQ* are identified to be essential for magnetosomal biogenesis in *Magnetospirillum* strains; deletion of these genes results in non-magnetotactic mutants [5, 60, 62].

Although *mamA* and *mamK* are not essential for magnetic mineral formation in *Magnetospirillum* species based on previous studies, the proteins they encoded are both involved in fine-tuning the magnetic dipole moment of the cell and thus magnetotaxis: MamA is responsible for magnetosome membrane assembly and MamK is involved in biomineral chain formation [74–77].

We then examined the phylogeny of each magnetosome protein of MamABKMQ across available MTB genomes (Fig. 5), with the exception of MamI due to its < 60 aa positions after alignment trimming. The resulting trees are congruent overall with the genome-based phylogeny in that they are consistent with monophyly of the major phyla. This provides good evidence that the current MGC distribution across major phyla is not due to extensive recent HGT events, but that it is due mainly to vertical inheritance. Coupled with the sharing of a common set of magnetosome genes, this strengthens the scenario of a single MGC emergence with later multiple independent losses during Bacterial diversification [12]. However, a few discrepancies must be noted, such as separation between the UBA10199 phylum of nDJH6_

Fig. 5 Maximum-likelihood trees of core magnetosome proteins. Trees were inferred using the MamABKMQ found in available MTB genomes. Complete trees are shown in Supplementary Figures 2 to 6

bin20 and nXX_bin1 (Fe_3O_4-type), separate branching of nPCR_bin9 from the other *Nitrospinota* lineages, clustering of the Fe_3O_4-type *Planctomycetota* and *Omnitrophota* phyla, and clustering of *Desulfobacterota*_A and *Desulfobacterota* phyla in most protein trees, and clustering of *Nitrospinota* (except nPCR_bin9) and SAR324 within the *Proteobacteria* in trees of MamB and MamM (Fig. 5 and Supplementary Figures 2 to 6). These discrepancies could have resulted from HGT events. Alternatively, considering that only limited MTB genomes are available for the UBA10199, SAR324, *Nitrospinota*, and *Planctomycetota* phyla, these groupings could also be an artefact of tree reconstruction [78]. Additional MTB genomes from these phyla will help to differentiate between these possibilities.

Initial discovery of both Fe_3O_4- and Fe_3S_4-type magnetosomes and MGCs within the same cell in the *Desulfobacterota* phylum (*Deltaproteobacteria* class in the NCBI taxonomy) led to the proposal that genes for Fe_3S_4 biomineralization in magnetosomes originated in this lineage [19]. Our finding of the coexistence of Fe_3O_4- and Fe_3S_4-type MGCs in the genomes of UBA10199 and *Omnitrophota* phyla weakens this hypothesis. Fe_3S_4-type magnetosome proteins form a monophyletic group in all MamABMQ trees with only the exception of Fe_3S_4-type MamA of nXX_bin1 that is clustered with *Nitrospirota* Fe_3O_4-type counterparts with low bootstrap support (< 75%, Supplementary Fig. 2). The structures of Fe_3S_4-type protein trees generally represent their phylum-level taxonomic phylogenies (Fig. 5 and Supplementary Figures 2 to 6), which suggests that all Fe_3S_4-type magnetosome proteins likely shared a common ancestor and then evolved separately in each phylum. These findings suggest that Fe_3S_4-type MGCs likely arose before the common ancestor of the Fe_3S_4-MTB containing phyla of *Desulfobacterota*, *Latescibacterota*, *Planctomycetota*, UBA10199, and *Omnitrophota*. The Fe_3S_4-type protein cluster is not associated robustly with any other Fe_3O_4-producing MTB lineages in protein trees, which precludes determination of the evolutionary origin of Fe_3S_4-producing MTB. Considering that Fe_3O_4-type MGCs are the most widespread, with identification in all major MTB lineages, it is likely that the Fe_3O_4-types are the ancestral form of magnetosome biosynthesis and that Fe_3S_4-type MGCs arose from Fe_3O_4-type MGCs through gene cluster duplication and divergence [19]. However, both Fe_3O_4- and Fe_3S_4-type MGCs originating from an ancient unknown MGC type (Fe_3S_4 or other iron-containing biominerals) is also plausible [41].

Biogenesis of a magnetic organelle in the close descendants of the last bacterial common ancestor (LBCA)?

Magnetosomal biogenesis is now considered to have an ancient origin [19, 41, 45, 79]. Discovery in this study of divergent MGCs across various Bacterial phyla expands significantly the database of MTB genomes and allows development of a more comprehensive scenario for the origin of magnetosomal biogenesis. The phylogenies of core magnetosome proteins (Fig. 5) strengthens the notion of an ancient MGC origin, which then spread through a combination of vertical inheritance over geological times followed by multiple independent losses, HGT events, and gene/cluster duplications. This should have occurred from a more basal ancestor than previously thought prior to the divergence of all 14 known MTB-containing phyla (Fig. 3). Interestingly, these phyla are scattered within the Bacterial tree of life and their common ancestor can be traced to near the base of the tree (Supplementary Figure 7), which indicates parsimoniously that close descendants of LBCA or even the LBCA itself may have already contained ancestors of magnetosome genes and were, thus, capable of biomineralizing primitive magnetosomes.

Potential biogenesis of the magnetosome organelle in the close descendants of LBCA has two major implications. First, the LBCA could have already possessed a relatively complex subcellular organization and, thus, was not as "primitive" as is usually imagined. Early organisms on Earth are described typically as simple organisms that lacked complex subcellular structures; however, a conceptual model of a complex LBCA or even a complex last universal common ancestor (LUCA) has been proposed based on the presence of eukaryote-like features present in some members of the *Planctomycetes* [80, 81]. Our results imply the formation of a magnetosome organelle in the close descendants of LBCA, which supports the idea of relatively complex early organisms. On early Earth, the lack of a protective ozone layer resulted in higher harmful ultraviolet radiation than the present-day Earth, which would have been a major challenge for life in surface and shallow-water conditions [82]. The intrinsic enzyme-like properties of magnetosomal iron nanoparticles [17, 18] and the stability of Fe_3O_4 nanozyme under a wide range of temperatures (4 to 90 °C) and pH (1 to 12) [83] might have helped life to cope with environmental stresses on early Earth (e.g., detoxification of ultraviolet radiation (UVR) and free-iron-generated ROS [12]). In addition, magnetotaxis behaviour may have also protected early MTB from lethal UVR by allowing efficient geomagnetic-field-directed swimming from near surface and shallow-water microenvironments to deeper water or sediment [84]. Second, a potential magnetosome-forming LBCA implies that this feature may have been inherited by a taxonomically wide group of Bacterial phyla, although some phyla could have lost this trait during evolution. Thus, we argue that many MTB-containing lineages await discovery. During review of this manuscript, a new study was published that generated 38 novel MTB draft genomes

from public databases [85], including five genomes belonging to the phyla of *Elusimicrobia*, *Candidatus* Hydrogenedentes, and *Nitrospinae* according to NCBI taxonomy. Future exploration of additional MTB affiliated within other Bacterial phyla, especially those near the base of the Bacterial tree [86, 87], will help to understand and trace the evolutionary origin of the magnetosome organelle.

Conclusions

We reconstruct 168 MTB genomes here, which expand substantially the genomic representation of MTB and indicate a much more widespread distribution of magnetosome organelle biogenesis across the domain *Bacteria*. Analysis of the core magnetosome proteins in the largest available taxonomic representation strengthens the notion of an ancient origin for magnetosome organelle biogenesis, which may date back to the base of the *Bacteria*. Genomes from this study will enable a better understanding of the biology and biomineralization of the magnetosome organelle and offer clues to assist in cultivation of uncultured MTB from different lineages.

Methods

Sample collection

A total of 53 sediment and soil samples were collected from a wide range of natural environments across China and Australia (Fig. 2a), including 13 sediment samples that have been described previously [41, 88] (for details, see Supplementary Table 1). Each sample was examined for the presence of living MTB by light microscopy using the hanging-drop method [89]. MTB cells from sediments and peatland soils were enriched magnetically using a "MTB trap" and enriched MTB cells were then subjected to metagenomic analyses, the detailed procedures for which are described elsewhere [45, 90]. The morphologies of enriched cells were examined using a JEM-2100 HR transmission electron microscope (JEOL, Japan) at 200 kV.

Metagenome assembly, population genome binning, and comparative genomic analyses

Metagenomic DNA from each location was sequenced on Illumina HiSeq 2000, 2500, or 4000 platforms. Sequencing data were processed through a single-sample assembly and binning strategy using a MetaWRAP pipeline [91]. The individual metagenomic datasets were assembled separately de novo using metaSPAdes (version 3.13.0) [92] with default parameters. Assembled scaffolds ≥ 2000 bp were binned separately using MetaBAT2 (version 2.12.1) [93], MaxBin2 (version 2.2.4) [94], and CONCOCT [95]. Results of three binning methods for each sample were refined using MetaWRAP's Bin_refinement and Reassemble_bins [91]. Genome completeness and contamination

were estimated with CheckM [96] using the 'lineage_wf' workflow. Only genomes with an estimated completeness > 50% and contamination < 10% and an estimated quality of > 50 (defined as completeness − 5 × contamination) [54] were retained. Statistics for each genome were obtained using QUAST (version 4.2) [97], including the genome length, number of scaffolds, largest scaffold, GC content, and N50. Resultant genomes were annotated using Prokka (version 1.11) [98]. Reconstructed genomes were checked manually for the presence of MGCs, followed by extensive manual verification of candidate magnetosome genes using NCBI PSI-BLAST [99]. Notably, genome sequences with > 99% ANI of previously reconstructed MTB genomes [41] from the same samples were recovered here using a different approach, which emphasizes the reproducibility of different genome-resolved metagenomic approaches. Taxonomic annotation of all acquired MTB genomes was performed using the Genome Taxonomy Database Toolkit GTDB-Tk [57] (version 0.3.2, database Release 04-RS89) with the 'classify_wf' function and default parameters. The genomes reconstructed here were combined with published MTB genomes (Supplementary Table 3). These genomes were dereplicated at 95% ANI for species [55] delineation using dRep [100] with '-sa 0.95'.

To identify the core magnetosome gene set shared by MTB in their respective genomes, only complete genomes and those with > 90% completeness and < 5% contamination are considered. These genomes were dereplicated using dRep [100] with '-sa 0.99' for dereplication at 99% ANI, and finally 83 high-quality representative MTB genomes were selected. Most are draft genomes, so we define a core magnetosome protein as being present in > 90% of the input genomes (i.e., at least 74 of 83 genomes) to minimize exclusion of potential core proteins due to the incomplete nature of draft genomes. Core proteins were calculated with COGtriangles [101] using GET_HOMOLOGUES [102] and were checked manually for magnetosome proteins. MamA-BIKMQ protein sequences were then searched with hidden Markov models [103] (HMM) across available MTB genomes.

Phylogenetic analyses

The phylogenetic tree composed of MTB genomes and those of relatively closely related non-MTB was inferred from 120 concatenated Bacterial single-copy marker proteins [54]. Maximum-likelihood phylogeny was calculated using IQ-TREE (version 1.6.9) [104] under the LG+I+G4 substitution model with 1000 ultrafast bootstraps. The genome tree was rooted with the genome from the candidate phyla radiation (accession number LCFW00000000). For each MamABKMQ protein, sequences were aligned using MAFFT (version 7.407)

[105] in 'auto' mode and filtered using trimAL [106] with '-gappyout' option. Maximum-likelihood phylogenetic protein trees were then constructed using IQ-TREE under the TEST option for best model selection with 1000 ultrafast bootstraps. Protein trees were rooted at the midpoint. All trees were visualized using FigTree version 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/) and Interactive Tree Of Life (iTOL) v4 [107]. AnnoTree [108] was used for phylogenomic visualization of the distribution of MTB-containing phyla across the Bacterial tree of life at the taxonomic phylum level.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s40168-020-00931-9.

Additional file 1: Supplementary Table 1. Summary of sampled sites.

Additional file 2: Supplementary Table 2. General characteristics of the 168 MTB genomes reported in this study. Genome completeness and contamination were estimated using CheckM and genome statistics were obtained using QUAST (version 4.2). Genome quality was defined as (completeness - 5 × contamination).

Additional file 3: Supplementary Table 3. Previously published MTB genomes included in this study.

Additional file 4: Supplementary Figure 1. Maximum likelihood phylogenomic tree of MTB genomes and their close non-MTB relatives.

Additional file 5: Supplementary Figure 2. Maximum-likelihood tree of magnetosome protein MamA.

Additional file 6: Supplementary Figure 3. Maximum-likelihood tree of magnetosome protein MamB.

Additional file 7: Supplementary Figure 4. Maximum-likelihood tree of magnetosome protein MamK.

Additional file 8: Supplementary Figure 5. Maximum-likelihood tree of magnetosome protein MamM.

Additional file 9: Supplementary Figure 6. Maximum-likelihood tree of magnetosome protein MamQ.

Additional file 10: Supplementary Figure 7. Phylogenetic distribution of MTB-containing phyla across the Bacterial tree of life. The phylum level Bacterial tree of life with MTB-containing phyla highlighted in blue. The Bacterial tree was made using the AnnoTree server.

Abbreviations
ANI: Average nucleotide identity; GTDB: Genome taxonomy database; HGTs: Horizontal gene transfers; LBCA: Last bacterial common ancestor; LUCA: Last universal common ancestor; MGCs: Magnetosome gene clusters; MTB: Magnetotactic bacteria; ROS: Reactive oxygen species; UVR: Ultraviolet radiation

Authors' contributions
W.L. and Y.X.P. conceived and planned the study. W.L., W.S.Z., G.A.P., X.Z., A.P.R., and Y.X.P. performed field sampling. W.L. and W.S.Z. enriched environmental magnetotactic bacteria and analyzed the data. W.L. wrote the manuscript with input from G.A.P., Q.Y.Z., R.K., D.A.B., A.P.R., and Y.X.P. The authors approved the final manuscript.

Availability of data and materials
The reconstructed genomes have been deposited under NCBI BioProject PRJNA400260 (BioSample accession numbers SAMN15005448-SAMN15005615). These genomes have also been deposited in the National Microbiology Data Center (http://nmdc.cn, NMDC Accession NMDC10017683).

Ethics approval and consent to participate
Not applicable

Consent for publication
Not applicable

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Key Laboratory of Earth and Planetary Physics, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China. [2]Innovation Academy for Earth Science, Chinese Academy of Sciences, Beijing 100029, China. [3]France-China Joint Laboratory for Evolution and Development of Magnetotactic Multicellular Organisms, Chinese Academy of Sciences, Beijing 100029, China. [4]College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. [5]Department of Earth, Ocean and Ecological Sciences, University of Liverpool, L69 7ZE, Liverpool, UK. [6]Department of Pediatrics, University of California San Diego, La Jolla, CA 92037, USA. [7]Research School of Earth Sciences, Australian National University, ACT, Canberra 2601, Australia. [8]School of Life Sciences, University of Nevada at Las Vegas, Las Vegas, NV 89154-4004, USA.

References
1. Yeates TO, Crowley CS, Tanaka S. Bacterial microcompartment organelles: protein shell structure and evolution. Annu Rev Biophys. 2010;39:185–205.
2. Grant CR, Wan J, Komeili A. Organelle formation in Bacteria and Archaea. Annu Rev Cell Dev Biol. 2018;34:217–38.
3. Saier MH, Bogdanov MV. Membranous organelles in Bacteria. J Mol Microbiol Biotechnol. 2013;23:5–12.
4. Lower BH, Bazylinski DA. The bacterial magnetosome: a unique prokaryotic organelle. J Mol Microbiol Biotechnol. 2013;23:63–80.
5. Uebe R, Schüler D. Magnetosome biogenesis in magnetotactic bacteria. Nat Rev Microbiol. 2016;14:621–37.
6. McCausland HC, Komeili A. Magnetic genes: studying the genetics of biomineralization in magnetotactic bacteria. PLOS Genet. 2020;16:e1008499.
7. Frankel RB, Blakemore RP. Magnetite and magnetotaxis in microorganisms. Bioelectromagnetics. 1989;10:223–37.
8. Vali H, Kirschvink JL. Observations of magnetosome organization, surface structure, and iron biomineralization of undescribed magnetotactic bacteria: evolutionary speculations. In: Frankel RB, Blakemore RP (eds). Iron Biominerals. 1990. Plenum Press, New York, pp 97–115.
9. Simmons SL, Edwards KJ. Geobiology of magnetotactic bacteria. In: Schüler D (ed). Magnetoreception and Magnetosomes in Bacteria. 2007. Springer, Berlin, pp 77–102.
10. Kopp RE, Kirschvink JL. The identification and biogeochemical interpretation of fossil magnetotactic bacteria. Earth-Sci Rev. 2008;86:42–61.
11. Kirschvink JL, Walker MM, Diebel CE. Magnetite-based magnetoreception. Curr Opin Neurobiol. 2001;11:462–7.

12. Lin W, Kirschvink JL, Paterson GA, Bazylinski DA, Pan Y. On the origin of microbial magnetoreception. Natl Sci Rev. 2020;7:472–9.

13. Monteil CL, Vallenet D, Menguy N, Benzerara K, Barbe V, Fouteau S, et al. Ectosymbiotic bacteria at the origin of magnetoreception in a marine protist. Nat Microbiol. 2019;4:1088–95.

14. Monteil CL, Lefevre CT. Magnetoreception in microorganisms. Trends Microbiol. 2020;28:266–75.

15. Blakemore RP. Magnetotactic bacteria. Science. 1975;190:377–9.

16. Frankel RB, Bazylinski DA, Johnson MS, Taylor BL. Magneto-aerotaxis in marine coccoid bacteria. Biophys J. 1997;73:994–1000.

17. Guo FF, Yang W, Jiang W, Geng S, Peng T, Li JL. Magnetosomes eliminate intracellular reactive oxygen species in *Magnetospirillum gryphiswaldense* MSR-1. Environ Microbiol. 2012;14:1722–9.

18. Li K, Wang P, Chen C, Chen C, Li L, Song T. Light irradiation helps magnetotactic bacteria eliminate intracellular reactive oxygen species. Environ Microbiol. 2017;19:3638–48.

19. Lefevre CT, Bazylinski DA. Ecology, diversity, and evolution of magnetotactic bacteria. Microbiol Mol Biol Rev. 2013;77:497–526.

20. Bazylinski DA, Lefevre CT. Magnetotactic bacteria from extreme environments. Life. 2013;3:295–307.

21. Lin W, Pan Y, Bazylinski DA. Diversity and ecology of and biomineralization by magnetotactic bacteria. Environ Microbiol Rep. 2017;9:345–56.

22. Fassbinder JWE, Stanjek H, Vali H. Occurrence of magnetic bacteria in soil. Nature. 1990;343:161–2.

23. Fassbinder JWE, Stanjek H. Magnetic properties of biogenic soil greigite (Fe_3S_4). Geophys Res Lett. 1994;21:2349–52.

24. Lin W, Bazylinski DA, Xiao T, Wu L-F, Pan Y. Life with compass: diversity and biogeography of magnetotactic bacteria. Environ Microbiol. 2014;16:2646–58.

25. Amor M, Tharaud M, Gélabert A, Komeili A. Single-cell determination of iron content in magnetotactic bacteria: implications for the iron biogeochemical cycle. Environ Microbiol. 2020;22:823–31.

26. Schulz-Vogt HN, Pollehne F, Jürgens K, Arz HW, Beier S, Bahlo R, et al. Effect of large magnetotactic bacteria with polyphosphate inclusions on the phosphate profile of the suboxic zone in the Black Sea. ISME J. 2019;13:1198–208.

27. Rivas-Lamelo S, Benzerara K, Lefevre CT, Monteil CL, Jézéquel D, Menguy N, et al. Magnetotactic bacteria as a new model for P sequestration in the ferruginous Lake Pavin. Geochemical Perspect Lett. 2017;35–41.

28. Cox BL, Popa R, Bazylinski DA, Lanoil B, Douglas S, Belz A, et al. Organization and elemental analysis of P-, S-, and Fe-rich inclusions in a population of freshwater magnetococci. Geomicrobiol J 2002; 19: 387–406.

29. Amann R, Peplies J, Schüler D. Diversity and taxonomy of magnetotactic bacteria. In: Schüler D (ed). Magnetoreception and Magnetosomes in Bacteria. 2007. Springer, Berlin, pp 25–36.

30. DeLong EF, Frankel RB, Bazylinski DA. Multiple evolutionary origins of magnetotaxis in bacteria. Science. 1993;259:803–6.

31. Jogler C, Kube M, Schübbe S, Ullrich S, Teeling H, Bazylinski DA, et al. Comparative analysis of magnetosome gene clusters in magnetotactic bacteria provides further evidence for horizontal gene transfer. Environ Microbiol. 2009;11:1267–77.

32. Jogler C, Schüler D. Genomics, genetics, and cell biology of magnetosome formation. Annu Rev Microbiol. 2009;63:501–21.

33. Morillo V, Abreu F, Araujo AC, de Almeida LGP, Enrich-Prast A, Farina M, et al. Isolation, cultivation and genomic analysis of magnetosome biomineralization genes of a new genus of South-seeking magnetotactic cocci within the *Alphaproteobacteria*. Front Microbiol. 2014;5:72.

34. Monteil CL, Perrière G, Menguy N, Ginet N, Alonso B, Waisbord N, et al. Genomic study of a novel magnetotactic *Alphaproteobacteria* uncovers the multiple ancestry of magnetotaxis. Environ Microbiol. 2018;20:4415–30.

35. Du H, Zhang W, Zhang W, Zhang W, Pan H, Pan Y, et al. Magnetosome gene duplication as an important driver in the evolution of magnetotaxis in the *Alphaproteobacteria*. mSystems. 2019;4:e00315–9.

36. Geurink C, Lefevre CT, Monteil CL, Morillo-Lopez V, Abreu F, Bazylinski DA, et al. Complete genome sequence of strain BW-2, a magnetotactic Gammaproteobacterium in the family *Ectothiorhodospiraceae*, isolated from a brackish spring in Death Valley, California. Microbiol Resour Announc. 2020;9:e01144–19.

37. Abreu F, Leão P, Vargas G, Cypriano J, Figueiredo V, Enrich-Prast A, et al. Culture-independent characterization of a novel magnetotactic member affiliated to the Beta class of the *Proteobacteria* phylum from an acidic lagoon. Environ Microbiol. 2018;20:2615–24.

38. Li J, Zhang H, Liu P, Menguy N, Roberts AP, Chen H, et al. Phylogenetic and structural identification of a novel magnetotactic Deltaproteobacteria strain, WYHR-1, from a freshwater lake. Appl Environ Microbiol. 2019;85:e00731–19.

39. Kolinko S, Jogler C, Katzmann E, Wanner G, Peplies J, Schüler D. Single-cell analysis reveals a novel uncultivated magnetotactic bacterium within the candidate division OP3. Environ Microbiol. 2012;14:1709–21.

40. Lin W, Pan Y. A putative greigite type magnetosome gene cluster from the candidate phylum *Latescibacteria*. Environ Microbiol Rep. 2015;7:237–42.

41. Lin W, Zhang W, Zhao X, Roberts AP, Paterson GA, Bazylinski DA, et al. Genomic expansion of magnetotactic bacteria reveals an early common origin of magnetotaxis with lineage-specific evolution. ISME J. 2018;12:1508–19.

42. Koziaeva V, Dziuba M, Leão P, Uzun M, Krutkina M, Grouzdev D. Genome-based metabolic reconstruction of a novel uncultivated freshwater magnetotactic coccus "Ca. Magnetaquicoccus inordinatus" UR-1, and proposal of a candidate family "Ca. Magnetaquicoccaceae". Front Microbiol 2019; 10: 2290.

43. Abreu F, Cantao ME, Nicolas MF, Barcellos FG, Morillo V, Almeida LG, et al. Common ancestry of iron oxide- and iron-sulfide-based biomineralization in magnetotactic bacteria. ISME J. 2011;5:1634–40.

44. Lefevre CT, Trubitsyn D, Abreu F, Kolinko S, de Almeida LGP, de Vasconcelos ATR, et al. Monophyletic origin of magnetotaxis and the first magnetosomes. Environ Microbiol. 2013;15:2267–74.

45. Lin W, Paterson GA, Zhu Q, Wang Y, Kopylova E, Li Y, et al. Origin of microbial biomineralization and magnetotaxis during the Archean. Proc Natl Acad Sci U S A. 2017;114:2171–6.

46. Huang X, Pancost RD, Xue J, Gu Y, Evershed RP, Xie S. Response of carbon cycle to drier conditions in the mid-Holocene in central China. Nat Commun. 2018;9:1369.

47. Tian W, Wang H, Xiang X, Wang R, Xu Y. Structural variations of bacterial community driven by sphagnum microhabitat differentiation in a subalpine peatland. Front Microbiol. 2019;10:1661.

48. Zhao M, Zhang Y, Zhang Z, Huang X. Comparison of microbial community in topsoil among different habitats in Dajiuhu, Hubei Province: evidence from phospholipid fatty acids (in Chinese with English abstract). Earth Sci. 2020;45:1877–86.

49. Nash C. Mechanisms and evolution of magnetotactic bacteria. California Institute of Technology . 2008. California Institute of Technology.

50. Lefevre CT, Abreu F, Schmidt ML, Lins U, Frankel RB, Hedlund BP, et al. Moderately thermophilic magnetotactic bacteria from hot springs in Nevada. Appl Environ Microbiol. 2010;76:3740–3.

51. Lefevre CT, Frankel RB, Posfai M, Prozorov T, Bazylinski DA. Isolation of obligately alkaliphilic magnetotactic bacteria from extremely alkaline environments. Environ Microbiol. 2011;13:2342–50.

52. Aliaga Goltsman DS, Comolli LR, Thomas BC, Banfield JF. Community transcriptomics reveals unexpected high microbial diversity in acidophilic biofilm communities. ISME J. 2015;9:1014–23.

53. Liu J, Zhang W, Li X, Li X, Chen X, Li JH, et al. Bacterial community structure and novel species of magnetotactic bacteria in sediments from a seamount in the Mariana volcanic arc. Sci Rep. 2017;7:17964.

54. Parks DH, Rinke C, Chuvochina M, Chaumeil P-A, Woodcroft BJ, Evans PN, et al. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nat Microbiol. 2017;2:1533–42.

55. Klappenbach JA, Goris J, Vandamme P, Coenye T, Konstantinidis KT, Tiedje JM. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol. 2007;57:81–91.

56. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol. 2018;36:996–1004.

57. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics. 2019:1–3.

58. Grünberg K, Wawer C, Tebo BM, Schüler D. A large gene cluster encoding several magnetosome proteins is conserved in different species of magnetotactic bacteria. Appl Environ Microbiol. 2001;67:4573–82.

59. Kolinko I, Lohße A, Borg S, Raschdorf O, Jogler C, Tu Q, et al. Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters. Nat Nanotechnol. 2014;9:193–7.

60. Murat D, Quinlan A, Vali H, Komeili A. Comprehensive genetic dissection of the magnetosome gene island reveals the step-wise assembly of a prokaryotic organelle. Proc Natl Acad Sci U S A. 2010;107:5593–8.

61. Lohße A, Ullrich S, Katzmann E, Borg S, Wanner G, Richter M, et al. Functional analysis of the magnetosome island in *Magnetospirillum*

gryphiswaldense: the mamAB operon is sufficient for magnetite biomineralization. PLoS One. 2011;6:e25561.

62. Lohße A, Borg S, Raschdorf O, Kolinko I, Tompa É, Pósfai M, et al. Genetic dissection of the mamAB and mms6 operons reveals a gene set essential for magnetosome biogenesis in *Magnetospirillum gryphiswaldense*. J Bacteriol. 2014;196:2658–69.

63. Rong C, Huang Y, Zhang W, Jiang W, Li Y, Li J. Ferrous iron transport protein B gene (feoB1) plays an accessory role in magnetosome formation in *Magnetospirillum gryphiswaldense* strain MSR-1. Res Microbiol. 2008;159:530–6.

64. Rong C, Zhang C, Zhang Y, Qi L, Yang J, Guan G, et al. FeoB2 functions in magnetosome formation and oxidative stress protection in *Magnetospirillum gryphiswaldense* strain MSR-1. J Bacteriol. 2012;194:3972–6.

65. Lefèvre CT, Trubitsyn D, Abreu F, Kolinko S, Jogler C, de Almeida LGP, et al. Comparative genomic analysis of magnetotactic bacteria from the *Deltaproteobacteria* provides new insights into magnetite and greigite magnetosome genes required for magnetotaxis. Environ Microbiol. 2013;15:2712–35.

66. Rahn-Lee L, Byrne ME, Zhang M, Le Sage D, Glenn DR, Milbourne T, et al. A genetic strategy for probing the functional diversity of magnetosome formation. Plos Genet. 2015;11:e1004811.

67. Lin W, Deng A, Wang Z, Li Y, Wen T, Wu L-F, et al. Genomic insights into the uncultured genus '*Candidatus* Magnetobacterium' in the phylum *Nitrospirae*. ISME J. 2014;8:2463–77.

68. Murat D, Falahati V, Bertinetti L, Csencsits R, Kornig A, Downing K, et al. The magnetosome membrane protein, MmsF, is a major regulator of magnetite biomineralization in *Magnetospirillum magneticum* AMB-1. Mol Microbiol. 2012;85:684–99.

69. Tanaka M, Mazuyama E, Arakaki A, Matsunaga T. Mms6 protein regulates crystal morphology during nano-sized magnetite biomineralization in vivo. J Biol Chem. 2011;286:6386–92.

70. Bazylinski DA, Heywood BR, Mann S, Frankel RB. Fe_3O_4 and Fe_3S_4 in a bacterium. Nature. 1993;366:218.

71. Lefèvre CT, Menguy N, Abreu F, Lins U, Pósfai M, Prozorov T, et al. A cultured greigite-producing magnetotactic bacterium in a novel group of sulfate-reducing bacteria. Science. 2011;334:1720–3.

72. Chen Y-R, Zhang R, Du H-J, Pan H-M, Zhang W-Y, Zhou K, et al. A novel species of ellipsoidal multicellular magnetotactic prokaryotes from Lake Yuehu in China. Environ Microbiol. 2015;17:637–47.

73. Abreu F, Morillo V, Nascimento FF, Werneck C, Cantao ME, Ciapina LP, et al. Deciphering unusual uncultured magnetotactic multicellular prokaryotes through genomics. ISME J. 2014;8:1055–68.

74. Komeili A, Li Z, Newman DK, Jensen GJ. Magnetosomes are cell membrane invaginations organized by the actin-like protein MamK. Science. 2006;311:242–5.

75. Zeytuni N, Ozyamak E, Ben-Harush K, Davidov G, Levin M, Gat Y, et al. Self-recognition mechanism of MamA, a magnetosome-associated TPR-containing protein, promotes complex assembly. Proc Natl Acad Sci. 2011;108:E480–7.

76. Katzmann E, Scheffel A, Gruska M, Plitzko JM, Schüler D. Loss of the actin-like protein MamK has pleiotropic effects on magnetosome formation and chain assembly in *Magnetospirillum gryphiswaldense*. Mol Microbiol. 2010;77:208–24.

77. Komeili A, Vali H, Beveridge TJ, Newman DK. Magnetosome vesicles are present before magnetite formation, and MamA is required for their activation. Proc Natl Acad Sci U S A. 2004;101:3839–44.

78. Yang Z, Rannala B. Molecular phylogenetics: principles and practice. Nat Rev Genet. 2012;13:303–14.

79. Zeytuni N, Cronin S, Lefevre CT, Arnoux P, Baran D, Shtein Z, et al. MamA as a model protein for structure-based insight into the evolutionary origins of magnetotactic bacteria. PLoS One. 2015;10:e0130394.

80. Forterre P. A new fusion hypothesis for the origin of Eukarya: better than previous ones, but probably also wrong. Res Microbiol. 2011;162:77–91.

81. Fuerst J. Keys to eukaryality: planctomycetes and ancestral evolution of cellular complexity. Front Microbiol. 2012;3:167.

82. Cnossen I, Sanz-Forcada J, Favata F, Witasse O, Zegers T, Arnold NF. Habitat of early life: Solar X-ray and UV radiation at Earth's surface 4–3.5 billion years ago. J Geophys Res. 2007;112:E02008.

83. Gao L, Fan K, Yan X. Iron oxide nanozyme: a multifunctional enzyme mimetic for biomedical applications. Theranostics. 2017;7:3207–27.

84. Shapiro OH, Hatzenpichler R, Buckley DH, Zinder SH, Orphan VJ. Multicellular photo-magnetotactic bacteria. Environ Microbiol Rep. 2011;3:233–8.

85. Uzun M, Alekseeva L, Krutkina M, Koziaeva V, Grouzdev D. Unravelling the diversity of magnetotactic bacteria through analysis of open genomic databases. Sci Data. 2020;7:252.

86. Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, et al. Unusual biology across a group comprising more than 15% of domain Bacteria. Nature. 2015;523:208–11.

87. Zhu Q, Mai U, Pfeiffer W, Janssen S, Asnicar F, Sanders JG, et al. Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains Bacteria and Archaea. Nat Commun. 2019;10:5477.

88. Zhang W, Ji R, Liu J, Pan Y, Wu L, Lin W. Two metagenome-assembled genome sequences of magnetotactic bacteria in the order *Magnetococcales*. Microbiol Resour Announc. 2020;9:7–9.

89. Greenberg M, Canter K, Mahler I, Tornheim A. Observation of magnetoreceptive behavior in a multicellular magnetotactic prokaryote in higher than geomagnetic fields. Biophys J. 2005;88:1496–9.

90. Jogler C, Lin W, Meyerdierks A, Kube M, Katzmann E, Flies C, et al. Towards cloning the magnetotactic metagenome: identification of magnetosome island gene clusters in uncultivated magnetotactic bacteria from different aquatic sediments. Appl Environ Microbiol. 2009;75:3972–9.

91. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome. 2018;6:158.

92. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 2017;27:824–34.

93. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. 2019;7:e7359.

94. Wu Y-W, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 2016; 32: 605–607.

95. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. Nat Methods. 2014;11:1144–6.

96. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25:1043–55.

97. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29:1072–5.

98. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.

99. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389–402.

100. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J. 2017;11:2864–8.

101. Kristensen DM, Kannan L, Coleman MK, Wolf YI, Sorokin A, Koonin EV, et al. A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. Bioinformatics. 2010;26:1481–7.

102. Contreras-Moreira B, Vinuesa P. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Appl Environ Microbiol. 2013;79:7696–701.

103. Johnson LS, Eddy SR, Portugaly E. Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinformatics. 2010;11:431.

104. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74.

105. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.

106. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3.

107. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47:W256–9.

108. Mendler K, Chen H, Parks DH, Lobb B, Hug LA, Doxey AC. AnnoTree: visualization and exploration of a functionally annotated microbial tree of life. Nucleic Acids Res. 2019;47:4442–8.

Publisher's Note