12-1-2017

Complete Sequencing and Comparative Analysis of the Genomes of the First Magnetotactic Gammaproteobacteria Isolated in Pure Culture: Strains BW-2 and SS-5

Corey Geurink

University of Nevada, Las Vegas, coreygeurink@gmail.com

Follow this and additional works at: https://digitalscholarship.unlv.edu/thesesdissertations

Part of the Microbiology Commons, and the Molecular Biology Commons

COMPLETE SEQUENCING AND COMPARATIVE ANALYSIS OF THE GENOMES OF THE FIRST

MAGNETOTACTIC *GAMMAPROTEOBACTERIA* ISOLATED IN PURE CULTURE: STRAINS BW-2 AND

SS-5

By

Corey Geurink

Bachelor of Science Biology

Grand Valley State University

2000

A dissertation submitted in partial fulfillment of the requirements for the

Doctor of Philosophy - Biological Sciences

School of Life Sciences

College of Sciences

The Graduate College

University of Nevada, Las Vegas

December 2017

Dissertation Approval

The Graduate College
The University of Nevada, Las Vegas

November 17, 2017

This dissertation prepared by

Corey Geurink

entitled

Complete Sequencing and Comparative Analysis of the Genomes of the First Magnetotactic *Gammaproteobacteria* Isolated in Pure Culture: Strains BW-2 and SS-5

is approved in partial fulfillment of the requirements for the degree of

Doctor of Philosophy - Biological Sciences
School of Life Sciences

Dennis Bazylinski, Ph.D.
Examination Committee Chair

Andrew Andres, Ph.D.
Examination Committee Member

Frank Van Breukelen, Ph.D.
Examination Committee Member

Eduardo Robleto, Ph.D.
Examination Committee Member

Ernesto Abel-Santos, Ph.D.
Graduate College Faculty Representative

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Interim Dean

Abstract

The genomes of the first two discovered magnetotactic bacteria (MTB) belonging to the *Gammaproteobacteria*, strains BW-2 and SS-5[1], were sequenced, sealed, annotated and compared to MTB of other phylogenetic groups. Cells of both strains are rod-shaped and biomineralize cuboctahedral and elongated octahedral crystals of magnetite, respectively, that are enveloped in a protein-embedded, lipid-bilayer membrane referred to as the magnetosome membrane or vesicle. The crystals and their associated membranes are known as magnetosomes. Magnetosome crystals consist of either magnetite (Fe_3O_4) or greigite (Fe_3S_4) and, because of their specific mineral compositions, crystal morphologies and sizes, the biomineralization processes involved in magnetosome formation are thought to be under strict genetic and (bio)chemical control through the action of the proteins associated with the magnetosome membrane. These biomineralization proteins are referred to as the Mam and Mms proteins, and the genes encoding them, the *mam* or *mms* genes. Known species of MTB phylogenetically belong to the *Alpha-, Gamma-, and Deltaproteobacteria* subgroups of the *Proteobacteria* phylum, the *Nitrospira* phylum and the candidate phyla *Latescibacteria* (WS3) and *Omnitrophica* (OP3). Complete or partial genomes sequences are available for MTB of most of these phylogenetic groups. A notable exception is genomic sequences from MTB from the *Gammaproteobacteria.*

The focus of this doctoral dissertation is the presentation of the first complete genome sequences of gammaproteobacterial MTB as well as a comparison of the genomes and, in particular, a comparison of the magnetosome genes of MTB of other phylogenetic groups. Specific goals were to determine the origin and evolution of the magnetosome genes as well as to determine the metabolic potential of these MTB strains.

The genomes of strains BW-2 and SS-5 are 4,103,727 bp and 3,729,439 bp in size, respectively, containing 3,838 and 3,724 coding sequences, respectively. Magnetosome genes form a single cluster about 32,781 kb in length in the genome of strain BW-2, while magnetosome genes make up two clusters in the strain SS-5 genome, of about 29,446 and 5,034 kb in length. Interspersed among the magnetosome genes of both MTB strains are unique coding sequences that share high homology to uncharacterized, hypothetical magnetosome-associated genes of other MTB. As to general metabolism, the genomes of both MTB strains contain the genes necessary for the utilization of the Calvin-Benson-Bassham cycle for autotrophy as well as the genes necessary for the oxidation of reduced sulfur compounds as a source of electrons. This is consistent with their ability to grow as microaerophilic chemolithoautotrophs using thiosulfate and sulfide as electron donors. Finally, numerous sets of genes are present in both genomes that explain other important and minor metabolic features of the strains.

In this work genes common to octahedral magnetite producers across phylogenetic groups are identified as well as potential uncharacterized magnetosome genes and the mode of acquisition and molecular evolution of the phenotype across phylogenies will be assessed.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Discovery and general features of magnetotactic bacteria

Magnetotactic bacteria (MTB) are a phylogenetically diverse group of Gram-negative, aquatic, motile prokaryotes that, although are metabolically and morphologically heterogeneous, share in common the ability to biomineralize intracellular, membrane-bounded, magnetic nanocrystals of magnetite (Fe_3O_4) and greigite (Fe_3S_4).[2–6] These magnetic nanocrystals crystals, together with the membrane they are formed within, are referred to as magnetosomes[7] and impart a permanent magnetic dipole moment to the cell. In turn, this magnetic dipole moment causes cells of MTB to align passively along the Earth's geomagnetic field lines, similarly to a compass needle, while they swim.[2,4,6,8–13] The combination of this passive magnetic orientation and active motility is referred to as magnetotaxis, the defining feature of MTB.[10,13]

Magnetotaxis was first observed by the Italian microbiologist Salvatore Bellini which he then described in 1963 in the University of Pavia publication Instituto di Microbiologia.[14] While observing a sample of mud and water, he noticed a large subset of cells swimming consistently in a northward direction. Referring to them as "batteri magnetosensibili," or magnetosensitive bacteria, he hypothesized that the magnetically influenced behavior was due to the presence of an internal "magnetic compass." This was confirmed in 1974 when Richard P. Blakemore independently rediscovered MTB and observed magnetosomes for the first time as described in his seminal paper on the topic.[6,12,15]

Ecology of MTB

MTB are ubiquitous in freshwater, brackish marine and hypersaline sedimentary habitats and water columns that are that are chemically stratified,[16] and have even been found in relatively moist soils.[17] Despite their cosmopolitan distribution, as generally microaerophilic gradient organisms[4,12] they are exquisitely sensitive to concentrations of O_2 and other compounds[12,18] that occur in vertical concentration gradients in natural habitats. In chemically-stratified habitats MTB occur in their highest numbers at the oxic-anoxic interface (OAI) or what has also been referred to as the oxic-anoxic transition zone (OATZ).[6,9,10,12,13,19–21] This is the microaerobic zone that lies between the oxic and the anoxic zones that is typified by a very low concentration of O_2 but overlaps with many reduced compounds some of which are used as electron donors for bacterial growth.[6,9,10,12,13,19–21] These reduced compounds include reduced sulfur species such as sulfide (S^{2-}).[3,4,12,13,21–28] A number of MTB have now been discovered below the OAI in the anoxic zone which is devoid of O_2.[12,13,29,30] In general, most magnetite-producing MTB are found at the OAI proper[4,12,26,29–34] while greigite-producing and some magnetite-producing species are present below the OAI in the anoxic zone[4,12,13,16,26,31,35] which is also often sulfidic.[4,6,35–37] This information suggests that MTB would thus likely be microaerophiles or anaerobes or both and the subsequent isolation and cultivation of many MTB confirms this supposition.[9,10,38–42] Although MTB, as a group, are very sensitive to atmospheric concentrations of O_2 (it is toxic to them) and tend to occupy niches at mesophilic temperatures and at near neutral pH[12,19,43,44], some strains are able to exploit extreme habitats

including those that are highly alkaline,[45,46] acidic,[47] hypersaline,[48–50] or have high temperatures (up to 63°C)[6,51] (e.g., hot springs) and low temperatures (e.g., the Antarctic)[52,53].

Function of magnetosomes and magnetotaxis

As described in the previous section, MTB are generally found in their highest numbers at the OAI.[10,12,54] For many species, this is important for their survival and growth as it is the only location where their electron donors (e.g., reduced sulfur species such as S^{2-}) and acceptors (O_2) overlap in many environments.[12,13] In addition, atmospheric concentrations of O_2 are toxic to these microaerophilic/anaerobic microorganisms,[2,12,16,37,41,55] and thus MTB must efficiently locate and maintain this optimal position. [11,13,38,56] The magnetosome mediated alignment along the Earth's geomagnetic field lines, coupled with chemotaxis, function to help the cell to more efficiently find their optimal O_2 concentration in a vertical concentration gradient systems.[4,6,9,11,36] The magnetic polar directionality of the geomagnetic field is positioned at an angle relative to the surface of the earth. The angle of the magnetic field lines vary across the globe from pole to pole. This angle of dip is the vector that results from the combined horizontal and vertical aspects of the magnetic field as it angles vertically away from the earth and horizontally arcing from the direction of one pole to the other as illustrated in Figure 1.[25,57] The field is horizontal at the magnetic equator (0°) and inclines more steeply closer to the magnetic poles such that it reaches a 90° inclination at magnetic north and south.[25,57,58] In both hemispheres, it is thought that MTB swim downward along the geomagnetic field lines by rotating their flagella counterclockwise and reverse their direction by

3

rotating their flagella clockwise.[13,25,59,60] MTB in opposite hemispheres have opposing

polarities, or a predominant or preferred swimming direction under oxic conditions [61] such that

in the northern hemisphere most are north-seeking and those in the southern hemisphere are

mostly south-seeking.[16] North-seeking (swimming towards a south pole) in the northern

hemisphere is also down-seeking and vice versa in the southern hemisphere. Thus, once

aligned along the inclined Earth's geomagnetic field lines, north-seeking MTB, under oxic

conditions, swim downwards along the chemogradient in response to it to the OAI. They then

need only reverse their swimming direction under anoxic conditions by reversing their flagellar

rotation enabling them to again find the OAI more readily.[16] At the magnetic equator there are

both north- and south-seeking MTB in equal amounts. [10,13,25,55]

Figure 1 Illustrations of the Earth's Geomagnetic Field Lines
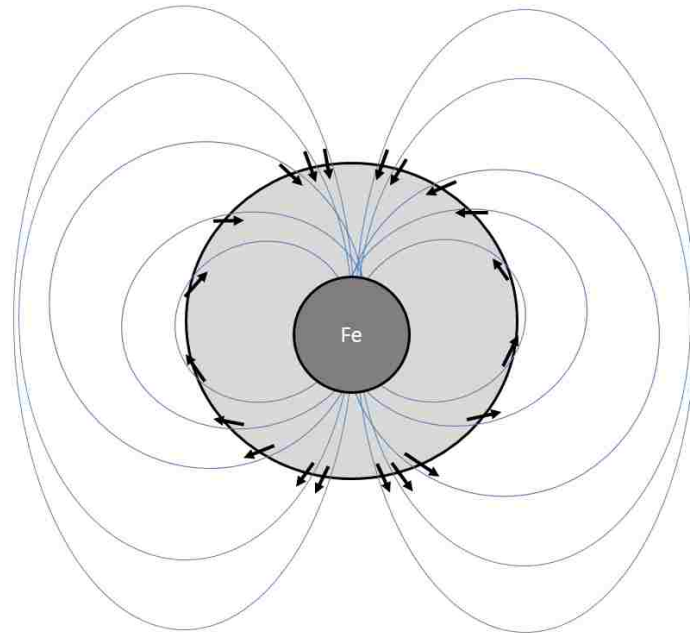
Diagram illustrating how magnetic field lines angle through the Earth's surface at different latitudes. Arrows indicate northward directionality of the geomagnetic field, blue lines represend magnetic field.

Once an MTB has passively aligned along the Earth's inclined geomagnetic field lines, it simply swims forward or in reverse, translating to up and down. This facilitates their movement in the water column or sediment along the opposing concentration gradients of O_2 decreasing from the surface with depth as reduced sulfur compounds increase in concentration. In doing so, the combination of magnetic field alignment and chemotaxis reduces a three dimensional search through the sediments, where the OAI is typically located, to one of a single dimension where cells swim forward and reverse (upward or downward) along the chemogradient, thereby conferring a selective advantage to the magnetotactic phenotype. [9,16]

Enrichment and separation of MTB for study

MTB generally are relatively easily enriched for in collected samples of mud and water by simply storing samples of collected sediment and water at room temperature in dim light.[27,62] The MTB in a sample can be concentrated and collected by first allowing the sample to settle with a magnet placed on the outside of the sample container at the approximate location of subsequent OAI formation as depicted in Figure 2, a process called magnetic enrichment.

Figure 2 Magnetic Enrichment of MTB from a Sample of Water and Sediment

Magnetic enrichment of MTB from an environmental sample of water and sediment using a magnetic stir bar placed near the OAI. Over time MTB aggregate near the magnet as indicated by the arrow. (Figure adapted from Bazylinski 2013)

After several hours, a drop is extracted from this location from the inside of the container and examined by employing the hanging drop technique.[12] When using this technique a drop is placed on a cover slip which is inverted and placed on an O-ring affixed to a microscope slide

such that the sample drop hangs from the cover slip as depicted in Figure 3. When a magnetic field is imposed by the placement of the south pole end of a magnet at one side of the drop and the north pole end of another magnet at the other end of the drop, any MTB present can be microscopically observed aggregating at the edge of the drop nearest the magnet pole corresponding to the geomagnetic field pole affinity of the strains present.

Figure 3 Illustration of the Hanging Drop Technique
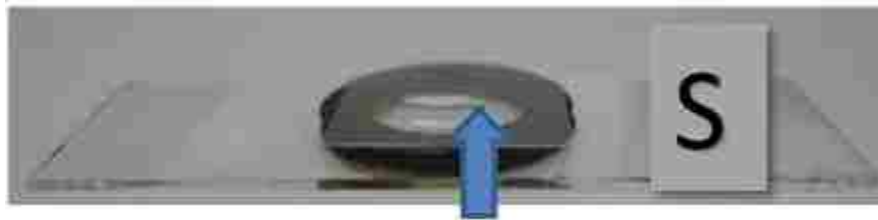
A drop of water and/or sediment from a magnetically-enriched environmental sample is placed on a cover slip and inverted on a rubber O ring affixed to a microscope slide. Arrow indicates the position geographically north seeking MTB will migrate to when the south pole of a magnet is placed at the position indicated by the S. (Figure adapted from Bazylinski 2013)

Collection and separation of MTB from non-magnetotactic bacteria in samples for various studies and experiments are achieved by exploitation of their magnetotactic behavior. The most commonly used method for this is through the use of the so-called capillary "magnetic racetrack," as depicted in Figure 4.[40,63] The racetrack is constructed by flame sealing the tapered end of a Pasteur pipette and inserting a cotton plug into the wide end, wedging it in at the site of taper. After sterilization of the racetrack, the sealed end is aseptically filled with filter-sterilized water past the cotton plug with a long syringe needle before addition of

the water and mud containing MTB to the wide end. To collect north-seeking MTB (i.e.,

generally those in the northern hemisphere), a magnetic field is then imposed on the racetrack

by positioning magnets at either end with the south pole of the magnet facing the sealed end of

the racetrack, and the north pole facing the wide opening of the race track. North seeking MTB

are directed to the south poles of magnets because the Earth's geographical north pole is in

fact, a south pole magnet.

Figure 4 Magnetic Racetrack
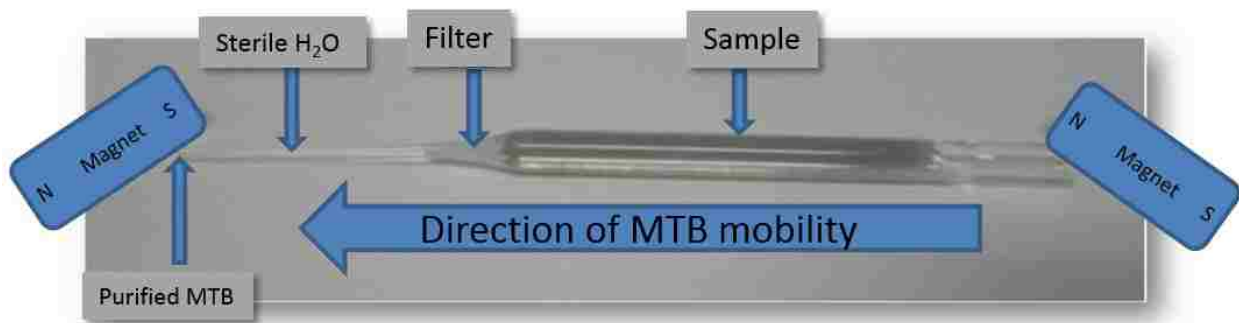
Pasteur pipette fashioned into a magnetic race track by heat sealing the tapered end, inserting a cotton plug at the taper, then

autoclaving. Magnets are arranged at either end to set up a magnetic field after aseptically injecting sterile water or culture medium

into the sealed end and environmental sample in the wide end. Small arrows point to location of contents, the long arrow indicates

the direction of MTB motility. (Figure adapted from Bazylinski 2013)

In the case of south-seeking MTB the magnetic field would be reversed, as south seeking MTB

are directed to the north poles of magnets because the Earth's south pole, is actually a north

pole magnet. When high numbers of MTB are present in the original sample, they can be

visualized accumulating at the sealed end using a dissecting microscope. As soon as sufficient

cells are collected, the end of the capillary is broken off and the cells removed with a sterile syringe. This technique works very well with fast swimming species (e.g., magnetococci), but for MTB that swim relatively slowly, the longer time allowed for migration increases the likelihood of contamination by non-magnetotactic bacteria in the racetrack.[40] MTB purified using this technique can be used in electron microscope studies, for DNA isolation or as inocula in growth experiments

Isolation in pure culture and cultivation

As MTB are generally very fastidious, they require specialized culture conditions for growth. It is often necessary to formulate custom O_2 concentration gradient semisolid medium based on the chemical conditions present in the sample where the particular MTB was found. [40,64] Once the medium has been prepared it is then dispensed into screw capped test tubes to simulate an aquatic gradient column. The O_2 concentration gradient is achieved by the addition of O_2 scavenging reducing agents such as cysteine. Thiosulfate or sulfide are often added as well, which also serve dual purposes as reducing agents as well as electron donors depending on the suspected metabolism of the cells targeted for enrichment. Once the medium is dispensed and the tube capped, the trapped O_2 in the headspace diffuses into the medium from the surface. The result is the production of a stable O_2 concentration and redox gradient in the tube with a concomitant establishment of an OAI. The OAI can be visualized by the

9

addition of resazurin to the growth medium which is pink in the presence of O_2 but in its

absence is colorless or light brown (the latter in the presence of some reduced sulfur species).

(Figure 5)

Figure 5 Visualization Representation of the Oxic/Anoxic Interface (OAI) in Semisolid O_2 Concentration Growth Medium When Resazurin is Present
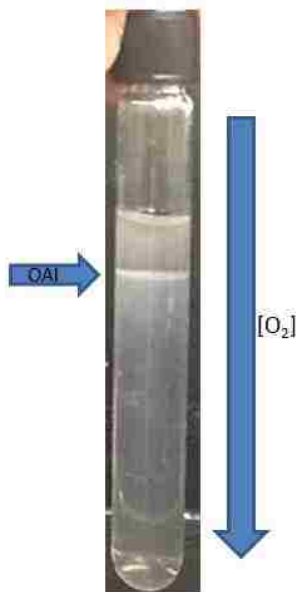
Tubes of semisolid O_2 concentration growth medium for the culture of strain BW-2. The OAI and band of cells are indicated by the arrow. The medium above the OAI is oxic due to the presence of O_2 which can be discerned by the pink color, while the medium below the OAI is anoxic due to the absence of O_2 which is discerned by the absence of the pink color. The downward arrow indicates the direction of O_2 diffusion.

Once the OAI has been established and is visible, cells collected from the magnetic racetrack

can be inoculated just below the OAI. Given the extremely fastidious nature of MTB, the major

challenge is deducing a medium formulation that actually results in the growth of the isolated

MTB.[12] In most cases MTB isolated from environmental samples do not survive or grow. In

cases where the strain grows in culture, occasionally it grows only to a low cell density that

10

presents a challenge when conducting characterization experiments or acquiring significant

amounts of DNA for cloning and sequencing.

Magnetosomes

Magnetosomes produced by MTB are defined as magnetic nanocrystals enveloped by a

phospholipid-bilayer membrane.[10,16,35,58] This membrane has been referred to as the

magnetosome membrane and the magnetosome membrane vesicle[10,13,25,65] and is the location

of the nucleation and growth of the mineral crystal.[4,30,35] It contains or is associated with a

number of proteins not found elsewhere in the cell and thought to be responsible for the

biomineralization and organization of magnetosomes in the cell.[12,13,16,66,67] The composition,

size and morphology of the magnetosome mineral crystals appears to be under strict genetic

and (bio)chemical control which will be further discussed later.[12,16,35] In addition, there is an

important correlation between magnetosome crystal composition and morphology and

phylogenetic affiliation in MTB that is discussed later.

Composition of magnetosome crystals

As previously stated, magnetosome crystals are composed of either the iron oxide magnetite (Fe_3O_4) or the iron sulfide greigite (Fe_3S_4). Generally a specific MTB will produce only one type of crystal although it has now been shown that some MTB produce both. [50]

Greigite-producing MTB have been shown to also contain some non-magnetic precursors to magnetite including [13,68] mackinawite (tetragonal FeS) and a sphalerite-type cubic FeS which have been observed to convert to magnetic greigite.[69] Interestingly, these precursor crystals are formed in a chain within the cell prior to them becoming magnetic.[12,13,68] The S structural arrangement remains unchanged during the transformation while only the Fe atoms rearrange in a step wise process from cubic FeS to mackinawite to finally form magnetic greigite.[68,69]

Figure 6 Conversion of the Putative Iron Sulfide Precursors of Greigite to Greigite in MTB
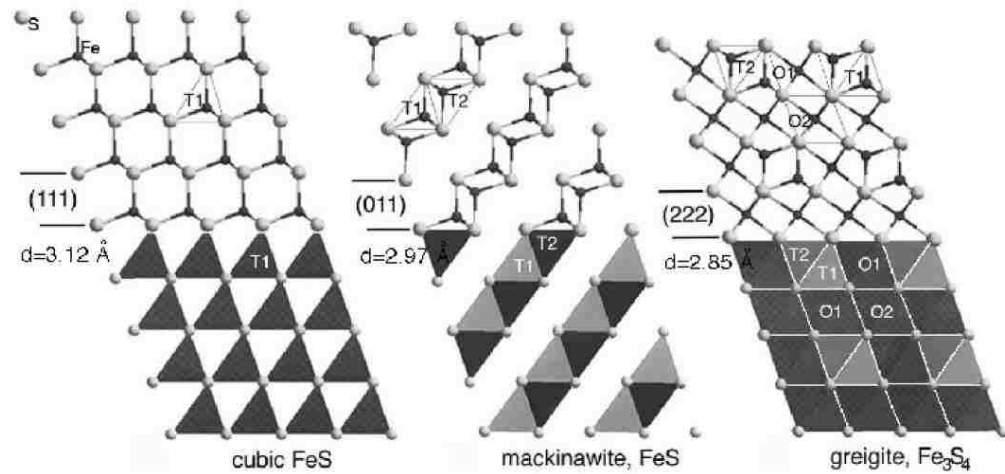
The putative stepwise chemical conversion of cubic FeS to mackinawite FeS to greigite in MTB. Light circles represent S atoms, dark circles represent Fe atoms, which rearrange around the S atom matrix in steps as FeS transforms incrementally to greigite. (Figure from Posfai 2006)

Magnetosome crystal size

Magnetosome crystals, regardless of whether they are composed of magnetite or greigite, fall within a narrow size range of between about 35 and 120 nm.[2] In this size range, crystals of both minerals have all of their atoms arranged such that their magnetic directionality is aligned in the same direction.[70,71] Magnetic material arranged in this way is referred to as a single magnetic domain. In excess of this size range, the crystal becomes unstable and its magnetic material may spontaneously repartition, dividing into a multi-domain magnet with smaller domains that will be within the stable size range, but may be oriented in different directions thereby reducing the net magnetic directionality of the whole. Crystals smaller than

the optimal size range of 35 to 120 nm are superparamagnetic, meaning that at ambient temperature they can spontaneously change magnetic directionality. [2,10,15,28,72] This is because there aren't enough atoms to influence one another to remain aligned and overcome the rearranging movement induced by thermal energy.[70,71] The amount of force a magnetic material can produce is defined as its dipole moment. Single magnetic domain crystals of magnetite and greigite have the maximum amount of iron atoms that can all remain aligned, they have the maximum magnetic dipole moment per mass. [2,35]

Magnetosome Crystal Morphology

Although some minor morphological variations have been observed within the same species under different growth conditions,[11] different strains of magnetite producing MTB produce different strain-specific magnetosome crystal morphologies that will be further discussed later. These crystal morphologies include cuboctahedral, bullet-shaped and elongated prismatic (Figure 7).[3,10,13,28,30,55]

Greigite magnetosome crystal formation appears to be less tightly controlled generally resulting in crystals that have irregularly shaped surfaces devoid of clear facets, that is, subhedral. [3,4,11,30,35] Even confirmed greigite crystals that generally lack clear facets, may have morphologies that are still species specific with respect to crystal size and dimension and that they are subhedral. For example, one strain might produce only equidimensional crystals of a certain size whereas another, only elongated crystals or yet another only bullet-shaped.[30] However there isn't yet a clear consensus as to whether all greigite magnetosome crystals are

14

species specific or whether they are just predictably subhedral.[30,69] Cuboctahedral and elongated prismatic greigite crystals have been observed in uncultured MTB from environmental samples taken from salt marshes in the Woods Hole, MA area.[69,73]

Figure 7 Magnetosome Crystal Morphologies in Cells of Different MTB
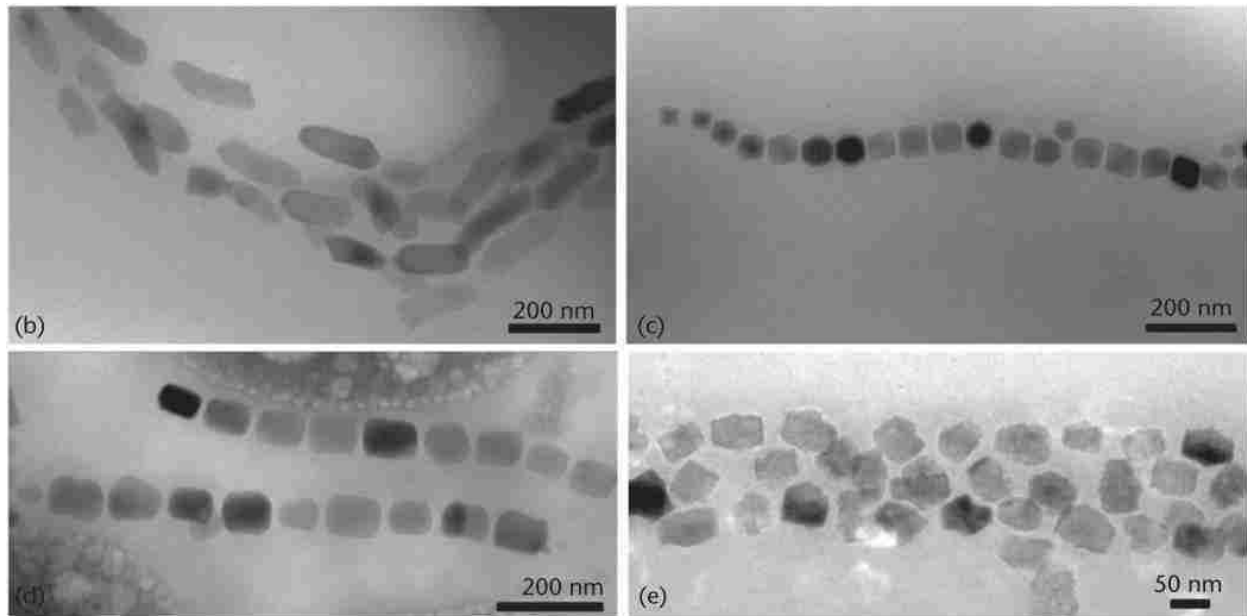
Transmission electron micrograph images of different magnetosome crystal morphologies. (b) Bullet- or tooth-shaped magnetite crystals (c), cuboctahedral magnetite crystals (d), elongated prismatic magnetite crystals (e) and irregularly-shaped greigite crystals. (Image modified from Lefevre 2011)

Magnetosome arrangement

Most MTB organize their magnetosomes via cytoskeletal elements into one or more relatively linear chains that align along the long axis of the cell, orienting their magnetic crystals end to end.[11,25,54,74,75] The resulting effect is that the magnetic fields of each overlap and join additively to combine the individual magnetic moments of each crystal into one larger magnetic

moment. What is gained is in effect a larger stable magnet than the cell could biomineralize otherwise, as can be seen by imaging the way the resulting magnetic field bends the electron beam during transmission electron microscopy[16,76,77] (Figure 8). Because the chains are fixed parallel to one another and to the cell, the torque applied by the geomagnetic field to the crystals is transferred to the cell, aligning it along geomagnetic north and south.

In light of the observations that the crystals at the peripheral ends of the magnetosome chains tend to be smaller than crystals more centrally located, it appears likely that they form at the ends of the chains.[16,78] Usually greigite crystal magnetosomes are less strictly controlled in their orientation than magnetite magnetosomes and are often found to be arranged in somewhat disordered chains, although still eliciting the effect of a single magnet.[2,4,30] Some magnetic cocci have also been found to form their magnetosomes in dispersed clusters, usually at the end of the cell corresponding to flagellar insertion.[10,13,79,80] These clusters are also still arranged in such a way as to result in the same behavior as a chain to produce a net magnetic dipole moment. [10,13,79,80]

Figure 8 Magnetic Field of a Magnetosome Chain in a Cell of Magnetospirillum

magnetotacticum
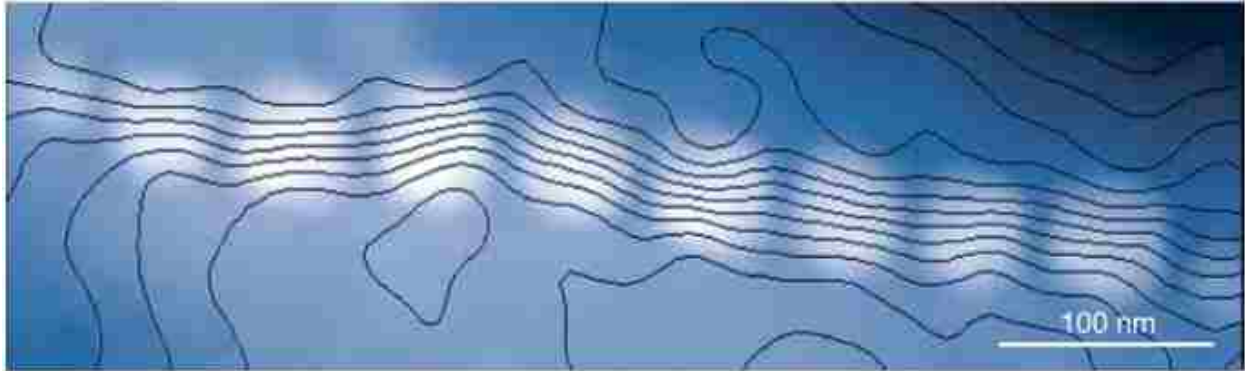
Magnetic field lines (black) superimposed on magnetosomes (bright regions) derived from electron holography, or mapping the

interference pattern observed during electron microscopy. The field lines passing through the magnetosomes indicates the chain

acts as a single magnetic dipole. (Figure from Bazylinski 2004)

Magnetosome Genes

While the characteristic that defines MTB as a group is the phenotype that results from

magnetosomes, the crystal mineral composition and arrangement vary with each species

producing crystals with a morphology unique to it.[9,12,16] The magnetosome crystal size, mineral

composition, purity, morphology and composition are tightly controlled by the proteins

embedded in the magnetosome membrane.[4,12,16,35] These embedded proteins are coded for by

the bacterium's compliment of magnetosome-associated membrane genes referred to as the *mam*, *mms* and *mtx* genes.[42]

The first studies to investigate the genetic determinants necessary for magnetite biomineralization in these prokaryotes were conducted with axenic cultures of the exclusively elongated cuboctahedral magnetite crystal producing[81] *Magnetospirillum* (*M.*) species of the *Alphaproteobacteria*. As a result the greatest degree of established knowledge is based on this group. These studies, primarily utilizing genetic systems established in *M. magneticum* and *M. grysphiswaldense*, [9,44,82] showed that there are, at least for the species of the genus *Magnetospirillum,* numerous genes involved in magnetosome formation that cluster together, some of which are transcribed as operons. [82,83] This includes the *mam* genes: *A,P,Q,R,B,S,T,U,H,E,J,K,L,M,N* and *O*. (Appendix Table 30)

The genes of the *mam*AB operon are involved in various necessary basic functions such as the formation of and attachment to the actin like cytoskeletal scaffolding that hold the magnetosomes in place (MamK),[84,85] invagination of the inner membrane to form the vesicle (MamIMBLQ),[2,86] vesicle activation and crystal maturation (MamA),[2,16] magnetite biomineralization (MamEPMBO).[83,87,88] Other functions include iron uptake and transport, (FeoA-like and FeoB-like,[89] MamN[83]) crystal nucleation (MamN[83]), electron transfer (MamT[33]), small solute transport (MamH[90]) and proteases (MamE[65]). It was later determined through knock out experiments that the genes of the mamAB operon made up what is seemingly the core set of *mam* genes required for the magnetic phenotype through complete magnetosome formation.[82,83,91] Magnetite biomineralization can be achieved with these genes alone, although notably the crystals are both smaller and misshapen.[82,83] Not all of the genes of the

18

mamAB operon are still generally considered essential however, as some are missing in MTB in groups other than the genus *Magnetospirillum*.   For example, although essential to the anchoring of magnetosomes to MamK, *mamJ,* which codes for a protein mediating this attachment, is not present in other MTB.[74,84,85]

 The fact that each strain produces crystals of a specific morphology unique to it, may be due to its complement of periphery magnetosome associated genes and its unique genetic background containing potentially pleiotropic genes also playing a role in biomineralization.[92]

In addition to the *mamAB* genes, there are also less characterized accessory clusters and genes associated with magnetosome formation in various MTB species.   The fact that crystals formed with only the *mamAB* operon available develop with aberrant morphology indicates that the additional genes play a role in the tight size and shape control of crystal morphology observed in MTB.   For example, the 3.6 kb five gene *mms* cluster is indeed involved in the nucleation and size control of magnetite crystals.[93,94] The approximately 2.1 kb *mamGFDC* cluster in magnetospirillum produces proteins that are also involved in size regulation of the crystals.[67] In the marine magnetococcus *Magnetococcus marinus* strain MC-1, *mamC* is, rather than part of a *mamGFDC* cluster, found in a *mamCEIH* cluster,[56] while *mamC* is found in the *mamCOP* clusters of strain BW-2 and SS-5.   Yet another cluster appearing to modulate magnetosome crystal size is the *mamXY* cluster in *Magnetospirillum*[90] and the orthologous *mamDXZ* cluster in *Magnetococcus marinus*.[56] Found apart from any other magnetosome genes, *mamW* is also likely to play a non-essential role in crystal morphology.[95] The *mtx* cluster that is found in various MTB species contains genes with TPR protein homology as well as one homologous to an adenylate cyclase and therefore may be involved in signal transduction in

some way.[56] TPR (tetratricopeptide repeat) proteins are involved in the formation of scaffolds to mediate protein–protein interactions, very often leading to the assembly of multiprotein complexes.[96]

Arrangement of magnetosome genes within genomes of MTB

In not all, but the majority of MTB, most of the genes involved in the magnetic phenotype are found clustered in a stretch of genome with all of the characteristics of a genomic island, a portion of a genome containing evidence of the transfer of genetic material from another species.   Such evidence includes: a different GC content from the rest of the genome, the genome section being flanked by mobile elements such as insertion sequences, transposases and integrases and tRNA genes that serve as insertion sites for integrases. [97,98] The magnetosome gene regions in *M. magneticum, M. gryphiswaldense* and *Desulfovibrio magneticus* have these characteristics.[38] In one of the first studied MTB, *M. gryphiswaldense*, for example, magnetosome associated genes as well as interspersed pseudogenes are contained within a genomic segment of about 45 kb and are flanked by mobile elements (Figure 9).

Figure 9 Genomic island of MSR-1
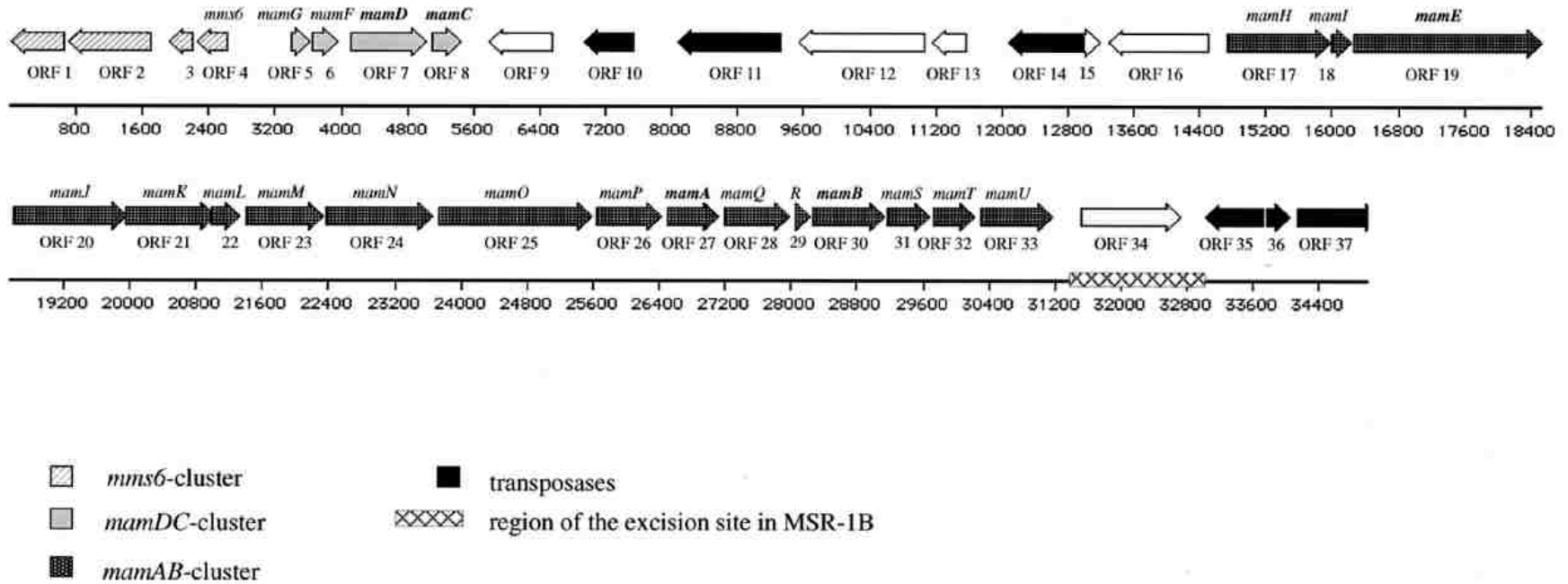
Genomic island of M. gryphiswaldense (MSR-1) flanked by transposases. Arrows indicate open reading frames (ORF), numbered hash marks indicate genomic position. Figure modified from Schubbe 2003

Spontaneous loss of magnetosome genomic islands or parts of them with a corresponding loss of magnetotactic phenotype have also been observed,[93] lending further support to the idea that some large magnetosome gene clusters are truly genomic islands.[91,95,97] Magnetosome associated genes are typically arranged collinearly as is seen with the genes of typical operons which are transcribed as single long polycistronic transcriptional units, suggesting along with some experimental evidence that magnetosome genes may be transcribed similarly.[91] This indicates that magnetotaxis may have been acquired by some MTB through horizontal gene transfer as well as by descent. [99]

Phylogeny and Physiology

Phylogenetically, representatives of MTB are found in the *Alphaproteobacteria*, *Gammaproteobacteria*, *Deltaproteobacteria classes of the Proteobacteria phylum, the Nitrospira* phylum and the candidate phyla *Latescibacteria* (formerly candidate phylum WS3) and *Omnitrophica* (formerly candidate phylum OP3), although the majority that have been so far described and cultured are of the *Alphaproteopacteria*.[30,100]

Alphaproteobacteria Class

MTB of the *Alphaproteobacteria* are facultative or obligate microaerophiles, some of which grow anaerobically. *Magnetococcus marinus* strain Mc-1, the coccoid strain MO-1, and

M. magnetotactcum strain MS-1 appear to be obligate microaerophiles.[101,102] All

alphaproteobacterial MTB   tested for nitrogenase activity (nitrogen fixation) possess this trait

and all necessary *nif* genes for the nitrogenase nitrogen fixation system were found to be in the

genomes of *M. magnetotacticum* and *M. magneticum*. [12] They appear to make up the majority

of MTB in many freshwater and marine environments,[103] and the most commonly present or

ubiquitous of these are the so-called magneto cocci of the *Alphaproteobacteria*.[104]   The two

strains of magneto cocci in culture, *Magnetococcus marinus*[56] and strain MO-1[12],   grow

autotrophically using reduced sulfur compounds as electron donors and the reverse

tricarboxylic acid cycle for autotrophy.[104,105] *Magnetococcus marinus* also grows

chemoorganoheterotrophically using acetate as a carbon and electron source.[56]

Members of the freshwater genus *Magnetospirillum*[106] appear to be the most

amenable MTB to culture (based on the overall number of isolated strains in culture) and the

least fastidious with regards to growth. This latter feature has allowed for the development of

genetic systems in several species of *Magnetospirillum* though which the function of

magnetosome related genes can be studied.[41]   Three of the most studied *Magnetospirillum*

species are *M. gryphiswaldense, M. magneticum* and *M. magnetotacticum* which have in

common the ability to respire as microaerophiles using O_2 or anaerobically with nitrate via

chemoorganoheterotrophic metabolism with organic acids as a carbon and electron

source.[106,107]   However, *M. gryphiswaldense* is also capable of autotrophic and mixotrophic

growth utilizing reduced sulfur compounds as electron donors. Ribulose-1,5-bisphosphate

carboxylase/oxygenase (RuBisCO) activity has been observed and the genes coding for RuBisCO

type II were identified in *M. gryphiswaldense* and *M. magnetotacticum*, suggesting that these

organisms use the Calvin-Benson-Bassham cycle for autotrophy.[22,108,109] RuBisCO, considered the most abundant protein on the planet, catalyzes the eventual conversion of CO_2 to the carbohydrate precursor 3-phosphoglycerate.[110]

Another well characterized MTB of the *Alphaproteobacteria* is the marine vibrio *Magnetovibrio blakemorei* strain MV-1.[111] Its metabolism is similar to freshwater magnetospirilla in that it also respires oxygen, nitrate or additionally nitrous oxide as terminal electron acceptors in respiration.[111] It grows chemoorganoheterotrophically with organic acids as well as some amino acids as carbon and electron sources,[111] and chemolithoautotrophically with reduced sulfur compounds as electron donors and the Calvin-Benson-Bassham autotrophy cycle,[36,111] with nitrogenase activity indicating nitrogen fixation in both modes of growth.[111]

Magnetospira thiophilia (strains MMS-1 and QH-2) are obligate microaerophiles that heterotrophically utilize organic acids as carbon and electron sources. They also exhibit chemolithoautotrophic growth using thiosulfate and nitrogenase activity. [12]

All known MTB of the *Alphaproteobacteria* biomineralize magnetite crystals that have cuboctahedral or elongated prismatic morphologies,[1,86,112] usually only producing the crystals under microaerobic or anaerobic conditions [16,101,102]

Deltaproteobacteria Class

The MTB of the *Deltaproteobacteria* include rods,[50,98] vibrios,[98] and even multicellular magnetotactic prokaryotes (MMPs) that consist of 10-60 genetically similar, communicating interdependent cells.[49,80,113] The Deltaproteobacterial MTB are able to grow heterotrophically by reducing sulfate and almost all cultured species exhibit nitrogenase activity and therefore likely fix atmospheric nitrogen. They are able to grow in alkaline, anaerobic, microaerophilic and saline environmets.[4,12]

Of most interest to this work, the MTB of the *Deltaproteobacteria* biomineralize magnetite, greigite or both, although notably the crystal morphology is bullet or tooth shaped.[49]

Gammaproteobacteria Class

The only finished MTB genomes of the *Gammaproteobacteria* are those of BW-2 and SS-5. They are both rods that grow chemolithoautotrophically, using thiosulfate and sulfide microaerobically as electron donors. Both use the Calvin-Benson-Bassham cycle for autotrophy, as evidenced by RuBisCO activity[1] and the presence of the necessary RubisCO genes.

Both notably biomineralize octahedral magnetite, similarly to the MTB of the *Alphaproteobacteria*.

Figure 10 Flagellar patterns on cells of strains BW-2 and SS-5
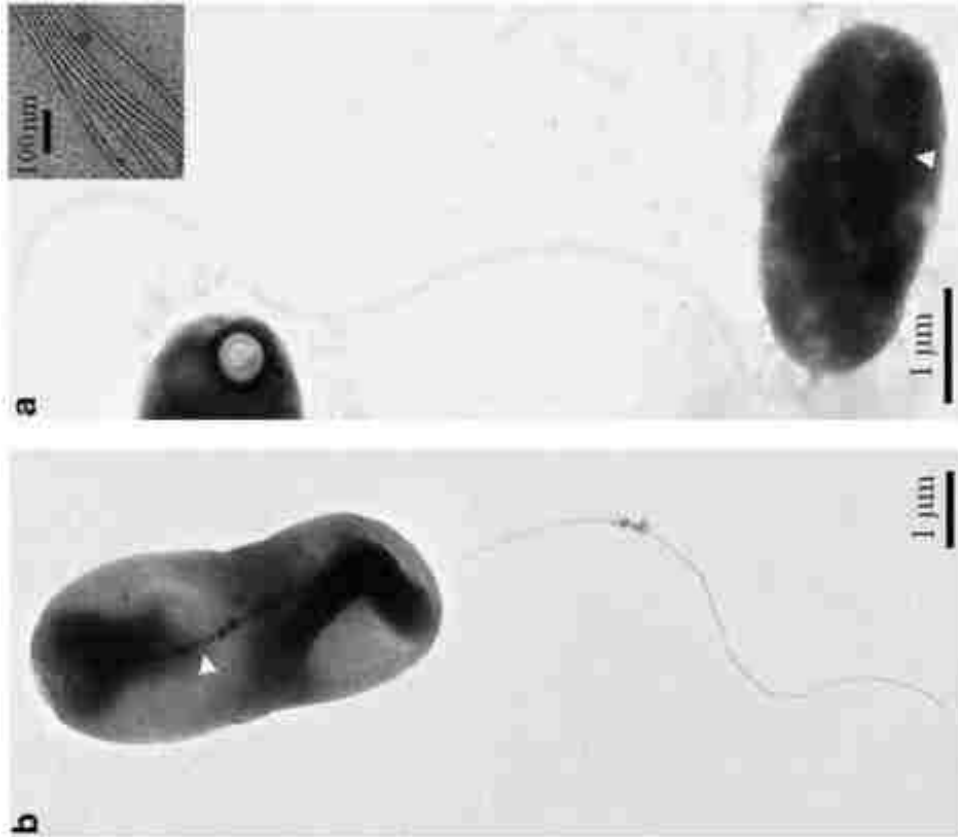
Transmission electron microscopy image of BW-2 (a) and SS-5 (b) cells. White arrows indicate magnetosome chains. The inset image depicts the bundle of seven flagella of BW-2.

Nitrospirae Phylum

As no MTB of the *Nitrospirae* have been successfully cultivated, less is known about their metabolism although they are likely as diverse as other groups as they inhabit the same niches and are observably as diverse in other regards. Morphologically, all so far described include rods,[39,114] ovoids[51] and vibrio.[98] They inhabit anoxic and microaerobic sediment

environments in fresh and brackish water of moderate to moderately high temperatures.

[12,51](up to 63°C)

Of importance to this work, all known MTB of the *Nitrospirae* biomineralize bullet

shaped magnetite.[27,51,53,115–118]

Phylogenetic significance of magnetosome morphology

It appears clear that in at least some cases the magnetic phenotype was acquired

through horizontal gene transfer events.[56] Evidence for this includes gene complement,

synteny, the fact that the magnetosome related genes of some MTB are clustered with all of

the characteristics of genomic islands including transposable elements and the entire gene set

has been observed to spontaneously be lost.[56,119]

Although the defining characteristic of MTB as a group is the formation of

magnetosomes, because this process is under such strict genetic control, there is considerable

variation of this phenotype between phylogenetically unrelated species. If it were generally the

case that the magnetotactic trait had been transferred in this way between different

phylogenetic groups, all of the various crystal morphologies would likely be present in the

various phylogenetic groups of MTB, which is not what has been observed.[30]

However, magnetosome crystal morphology distribution among MTB comports with

acquisition of the trait through descent. Production of the octahedral, prismatic and

rectangular crystals that require stricter control over the biomineralization process are localized

to the more recently diverging *Alpha* and *Gammaproteobacteria*. The bullet shaped magnetite

and irregular greigite crystals are produced within a monophyletic clade of the

Deltaproteobacteria, the *Nitrospirae*, and the candidate division OP3 which are the most deeply

branching groups.[100,120]

In addition, the core set of genes required for the magnetic phenotype is conserved in

all MTB with enough sequence variability to track the descent of the trait and evolutionary

differentiation of the different morphologies.[30] When a magnetosome gene phylogenetic tree

is constructed and compared with the 16s rRNA gene sequence phylogenetic tree the evolution

of the core magnetosome genes appears to be congruent with the evolution of the MTB species

themselves, lending strong support to the view that the magnetic phenotype evolved in current

MTB species from a common magnetic ancestor via descent.[100]

In light of these evidences it becomes clear that in the majority of cases the phenotype

is very likely the result of a monophyletic origin in a last common ancestor with many genes

controlling crystal morphology evolving after the branching of the groups.[86,100,116]

The MTB of the *Delaproteobacteria* and *Nitrospirae* are more deeply branching

phylogenetically and have fewer known magnetosome associated genes than the *Alpha* and

Gammaproteobacteria,[21] therefore it follows that they would have less precise control over the

biomineralization of their magnetosome crystals as is actually observed. Acquisition of the

phenotype through descent also comports with the strong correlation between crystal

morphology and phylogeny[21,86] with the *Nitrospirae* and *Deltaproteobacteria* producing bullet

or tooth shaped crystals and the MTB of the *Alphaproteobacteria* and *Gammaproteobacteria* producing cuboctahedral and elongated prismatic crystals.

Utility of Magnetotactic Bacteria Research

MTB cells and their magnetosomes have a unique combination of magnetic, physical and optical properties that can be used for many applications. Representing the oldest form of biomineralization on the planet, [9] these bacteria are an excellent system in which to study controlled biomineralization and the genetic components involved. The production of magnetosomes may have great potential in many biotechnological and nanotechnological applications due to the precise morphological and chemical nature of the magnetosome crystals they produce and the fact that they are membrane bound.[10,28] The membrane allows for aqueous dispersion in biological systems while enabling the binding of proteins via embedded antibodies for magnetic purification or chemiluminescence enzyme immunoassays. A potential clinically relevant application includes targeted magnetosome mediated drug delivery whereby a drug can be targeted within the body with a focused magnetic field. Additionally, in an oscillating focused magnetic field, magnetosome crystals produce heat, potentially facilitating the selective thermal destruction of malignant tissues. In addition, this form of prokaryotic biomineralization is a simpler and more tractable model system than bone in which to elucidate the basics of biomineralization.

MTB could also offer insight into some paleoecological and exobiological paradigms as these crystals are uniquely biomineralized such that they do not occur abiotically and remain intact over long periods of time under certain conditions in specific natural environments. MTB also present a unique opportunity to study molecular evolution as the phenotype has been identified in across taxonomical groups including the *Alpha* and *Deltaproteobacteria* [33] as well as the *Nitrospirae* [115] and the candidate OP3 division.[120] Only two so far have been identified in the *Gammaproteobacteria*, strains SS-5 and BW-2,[1] though neither has yet been thoroughly characterized. Genetic data regarding a number of other distantly related MTB are available, mainly from the cubooctahedral and elongated prismatic magnetite producing *Alphaproteobacteria* species and the bullet-shaped magnetite and/or greigite producing *Deltaproteobacteria*. There are only seven completed MTB genomes, limiting the genetic background information available for the magnetic phenotype.

In this study we finished, annotated and analyzed the complete genome sequences of strains BW-2 and SS-5, currently the only MTB of the *Gammaproteobacteria* isolated in pure culture. This enabled us to elucidate the gene content, sequence and synteny of the operons and MAI's of these new strains as they relate to each other and to other phylogenetically-unrelated characterized strains, as well as their genetic context, lending valuable insight into the molecular evolution and function of these unique genes as well as the mechanisms used to acquire these genes. . In chapters three and four BW-2 and SS-5 respectively will be assessed for metabolic potential based on genomic features and the gene compliment, synteny and presence or absence of genomic features indicative of the presence of a genomic island will be elucidated. In chapter five the metabolic potential and magnetosome gene compliment and

synteny of BW-2 and SS-5 will be compared to one another as well as to other MTB. The

phylogeny of magnetosome-related gene sequences will be compared with strain phylogeny

across phylogenetic groups in an effort to determine whether the magnetic phenotypes of

these strains were acquired via horizontal gene transfer or by vertical descent. Genes already

considered to be involved in magnetosome formation will be compared to identify those

common to magnetite biomineralizers across phylogenetic groups as well as any differences in

gene compliment between the octahedral crystal-producing BW-2 and the elongated prism-

producing SS-5. With these additional completed genomes available from the

Gammaproteobacteria, a further comparative genomics approach will be more powerful in

predicting novel candidate magnetotaxis genes which will be proposed in chapter five for

strains BW-2, SS-5 and in other MTB species. With a more complete metabolic background,

metabolic genes elsewhere in the genome could reveal pathways, reactions and enzymes

responsible for supportive, synergistic or redundant functions in the biomineralization of

magnetosomes, such as nitrate reductase (*nap*) gene. [121,122]

Chapter 2 Materials and Methods

Gammaproteobacterial MTB Strains Used in this Study

Strain BW-2 was isolated from sediment collected under permit from Badwater Basin in Death Valley National Park.[1] It is a rod-shaped bacterium that is motile via a polar bundle of seven flagella at one end of the cell.[1] Cells biomineralize cuboctahedral magnetite crystals in their magnetosomes, and sulfur- and phosphorus-rich inclusions when grown in semi-solid O_2-concentration gradient ([O_2]-gradient) medium with thiosulfate as electron donor[1] (Figure 11).

Figure 11 Inclusions within cells of strain BW-2
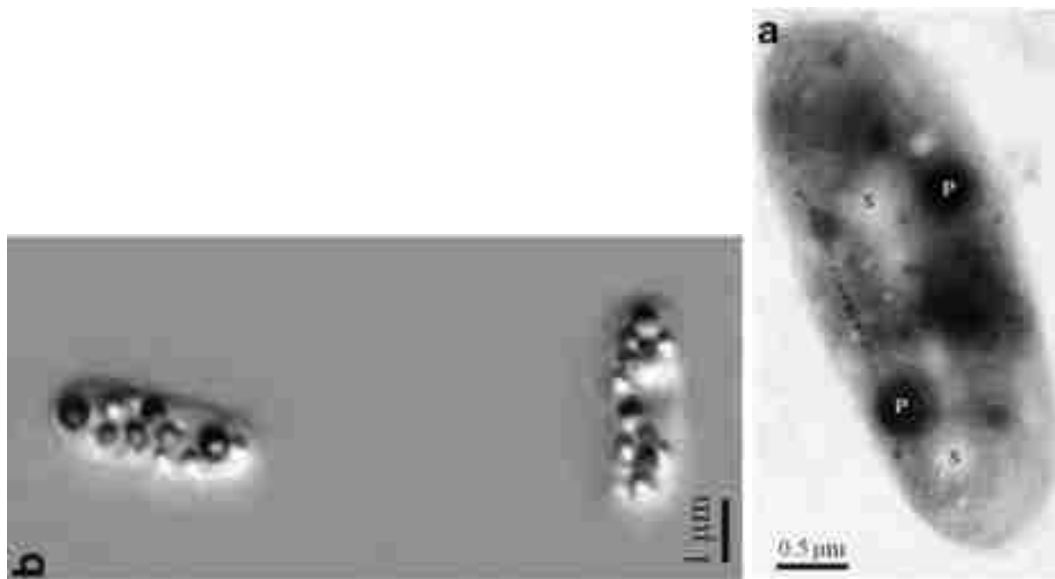
a) Transmission electron microscope image of a cell of strain BW-2 showing cell inclusions. B) Differential interference contrast image of cells of strain BW-2 showing cell inclusions. S = sulfur inclusions, P = phosphate inclusions, (From Lefevre 2012)

Strain SS-5 was isolated from sediment taken under permit from the shore of the Salton Sea in California.[1] It is also a rod-shaped bacterium and is motile but by means of a single polar flagellum. Cells of this strain also biomineralize magnetite magnetosome crystals but with an elongated, prismatic octahedral morphology and produce phosphorus-rich inclusions[1] (Figure 12).

Figure 12 Transmission electron microscope image of cells of strain SS-5 showing cell inclusions
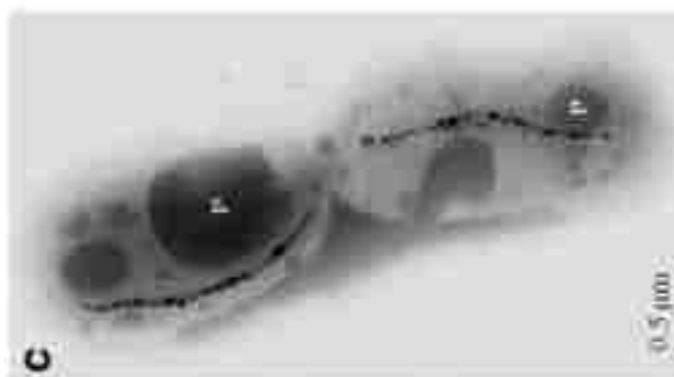
Differential interference contrast image of cells of strain SS-5. P = phosphate inclusions, (From Lefevre 2012)

Growth Conditions

BW-2 and SS-5 were grown in semi-solid [O_2]-gradient medium optimized for their individual growth. The most effective formula for the growth of strain BW-2 contained a base of salts consisting of 348 mM NaCl, 21.3 mM $MgCl_2 \bullet 6H_2O$, 22.8 mM Na_2SO_4, 3.25 mM $CaCl_2 \bullet 2H_2O$ and 7.73 mM KCl per liter of distilled, deionized H_2O (ddH_2O). To this was added per liter of medium, 0.2 mL aqueous resazurin, 5 mL Wolfe's mineral solution[123], 15 mM $NaHCO_3$, 5.6 mM NH_4Cl and 1.6 g Bacto-Agar (Difco, Burlington, NC). The pH was then adjusted to 7.0 and the medium was autoclaved. After autoclaving, the pH of the medium was again adjusted to pH 7.0 aseptically, if necessary, and the following sterile solutions were added

aseptically: 0.5 mL of vitamin solution[124], 1.8 mL of 0.5 M phosphate buffer (pH 7.0), 3 mL 0.01

M $FeSO_4$, 3 mL 40% sodium thiosulfate, 0.4 g of freshly made, filter-sterilized cysteine and 5 mL

100 mM Na_2S. The medium was again adjusted to pH 7.0 if necessary and dispensed into

screw-capped test tubes to approximately 80% of their volume, leaving 20% of the volume as

an air headspace. Cells were inoculated into this medium by syringe just below the OAI as

indicated by the colorless-pink interface (resazurin is pink in color when oxidized and colorless

when reduced ref). Cultures were then incubated at 25°C in the dark.

The growth medium used strain SS-5 was similar to that of BW-2 except the basal salts

consisted of: NaCl 646.8 mM, $MgCl_2 \bullet 6H_2O$ 26.56 mM, Na_2SO_4 38 mM, KCl 12 mM, and

$CaCl_2 \bullet 2H_2O$ 3.4 mM per liter of dd H_2O. The other reagents were the same as for strain BW-2

and the medium was inoculated and incubated as described above.

Isolation of DNA from Cells

When DNA was to be isolated, 825 mL of medium was prepared in a 1 L flask and 1.2 g

of low melting point agarose (ThermoFisher, Waltham, MA) was substituted for Bacto-Agar. In

the presence of sulfide this method facilitated dense cell growth as floating aggregates that

formed at the surface in addition to a layer of cells at the OAI.

DNA was extracted after pelleting cells by centrifugation at 13.4 RCF at 4°C then

removing the pellets by pipette from beneath the concentrated agarose that settled on top of

the pellet. Cell pellets were then combined and re-centrifuged and any additional agarose was

melted with warm sterile water and removed by pipette to keep it from clogging spin columns

used in genomic DNA (gDNA) extraction kits or from carrying over in the aqueous phase during

the phenol chloroform extraction.[125]   Kits used for isolation and cleanup were Qiagen Blood

and Tissue, (Germantown, MD) Qiaquick pcr purification kit (Qiagen, Germantown, MD) and

Zymo clean and concentrator. (Irvine, Ca.) DNA quality was determined by agarose gel

electrophoresis and Agilent Bioanalyzer and quantification was by Nanodrop ND-1000

spectrophotometer and Qubit 2.0 Fluorometer.

Genome sequencing, assembly, annotation and scaffolding.

Four-five-four genome sequences from strain SS-5 were obtained in one run on the GS

FLX sequencing system. For strain BW-2, 800x coverage of paired end Illumina sequence with

300 bp inserts were obtained in three runs on the Illumina MiSeq platform after library prep

with Illumina Nexterra XT (San Diego, Ca) and KAPA Biosciences (Wilmington, Ma) Library

Preparation Kits. Pacific Biosciences (PacBio)[126] long reads were obtained on the PacBio RS

Single Molecule Real-Time (SMRT) sequencing platform, using three SMRT cells for BW-2 to

achieve 333x coverage. CLC Genomics Workbench (CLC Bio-Qiagen, Aarhus, Denmark) was

then used to correct errors in the reads using a 75% subset of the input reads before de novo

assembly, which has been shown to yield results as accurate as correcting with the higher

accuracy Illumina reads.[127,128] The assembly was done using the automatic word size and contig

polishing settings, minimum word coverage of two, and minimum contig length of 2,000.

Scaffolding was then performed with PacBio long reads using CLC Bio Genomics Workbench

8.5.1 with the Genome Module 1.5.2 contig alignment, joining and splitting functions. This was then repeated 5 more times starting with different datasets with a consistent result before Illumina reads were mapped to it for further error correction. Illumina assemblies were also done with the SPAdes 3.5.0, [129] Velvet [130] and CLC Bio Genomics Workbench 8.5.1 (CLC Bio-Qiagen, Aarhus, Denmark) assemblers as well as the SABIA[131] and MaGe [132] annotation platforms in order to achieve the highest N50 and lowest number of contigs in order to compare assemblies for scaffolding purposes. Contig genomic features were annotated using the RAST[133,134] and MaGe genome annotation environments and potential promoters were identified using BPROM.[135] The genomes were searched for prophages using the PHAST server tool.[136] Circular chromosome maps were generated by the CG View web server application[137] and modified in Power Point (Microsoft, Redmond, WA).   Phylogenetic tree of MTB species was constructed in CLC Genomics workbench and modified in Power Point. At the scaffolding stage, this also enabled targeted PCR across contig gaps based on conserved genomic features that appeared to be split by contig breaks such as conserved operons and the magnetosome associated cluster. Singlet PCR was performed using GoTaq Green PCR master mix. (Promega, Madison, WI) Primers were designed with tools including Primer3, [138] OligoAnalyzer (PrimerQuest® program, IDT, Coralville, USA. Retrieved 12 December, 2012. http://www.idtdna.com/Scitools) and Primer-BLAST [139,140] utilizing nearest-neighbor, mono and divalent cation, oligo and dNTP concentration parameters. Primers were constructed by Integrated DNA Technologies. (Coralville, IA) Resulting amplicons were Sanger sequenced on the ABI 3730xl DNA Sequencer then aligned and joined to contigs using the Geneious 7.1.9 platform (http://www.geneious.com, Kearse et al., 2012). Contigs joined in this way were

designated as trusted contigs and were subsequently used as additional SPAdes input in re-

assembly and as quality checks for assembly and scaffolding based on PacBio reads. When

conserved regions at contig ends were no longer identified in BW-2, the Mauve whole genome

alignment algorithm [142] was employed to guide further targeted PCR attempts, using the

nearest 16s rRNA neighbor with a sealed genome, *Thioalkalivibrio versutus,*[143] which is 89.5%

similar in its 16s rRNA gene sequence to strain BW-2. When there were no longer data to guide

PCR primer placement, exploratory PCR methods were employed including a modification of a

universal multiplex primer PCR method that has been previously described[144–146] using the

Platinum Multiplex PCR master mix by Thermo Fisher Scientific. (Waltham, MA) In this method,

a sequence not occurring in the target genome is identified and its complimentary sequence is

incorporated at the 5' end of primers specific to sites of interest in the target genome to form a

compound universal primer/specific primer. The compound primers can then be pooled in a

multiplex PCR reaction at $1/20^{th}$ of standard concentration (0.5 µM) for a 20-plex reaction,

along with the usual standard concentration of the universal primer. (10 µM) As the reaction

proceeds, the specific primer portion of the compound primers generate amplicons during the

early phase to serve as additional template for regions of interest. These amplicons have

incorporated the complimentary sequence to the universal primer sequence. As the reaction

continues into exponential amplification the compound primers are exhausted enabling the

universal primer alone to take over in the amplification reaction (Figure 13). This eliminates

amplification bias among specific primers of differing efficiencies and allows amplicons to

equally amplify. In eliminating amplification bias during the exponential phase of the reaction

in this way, no primer is outcompeted for reagent and all possible amplicons are produced.

Figure 13 Universal Primer Multiplex PCR
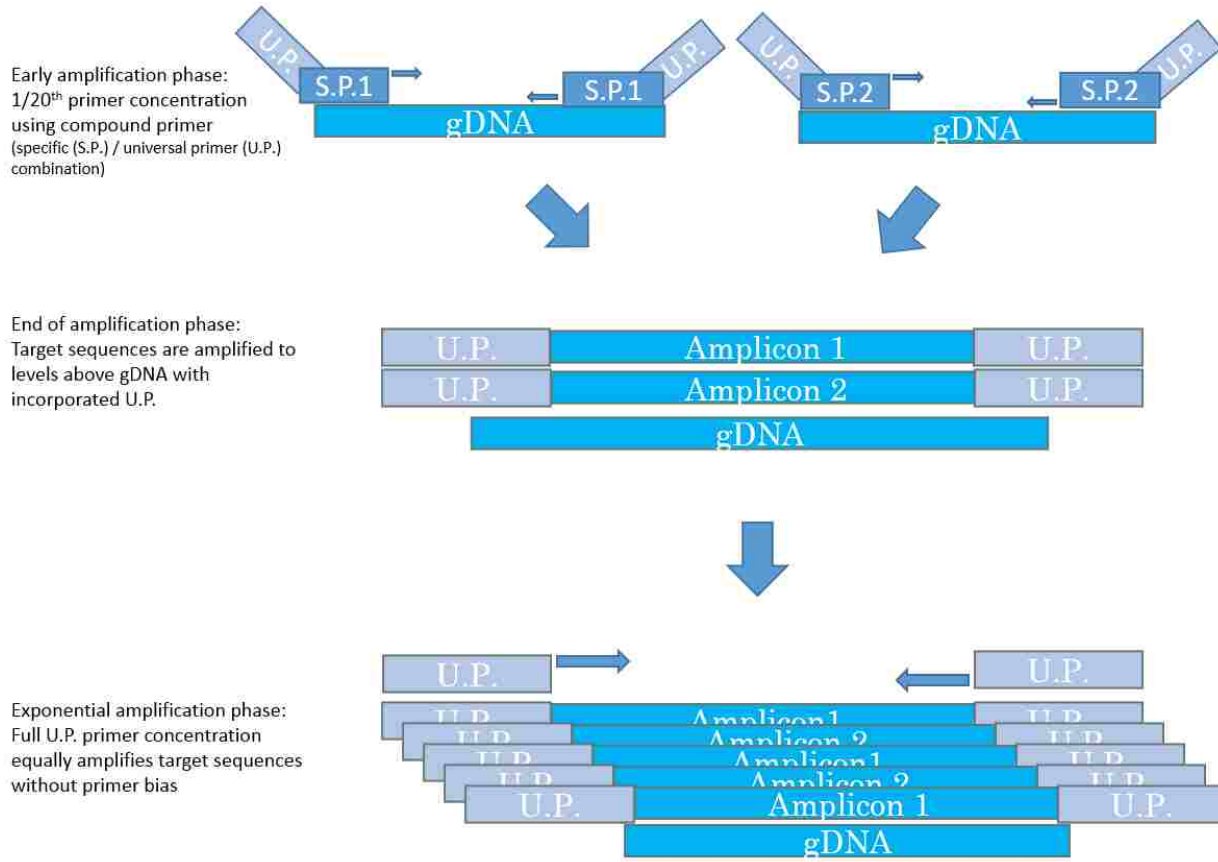
A suitably unique sequence for use as the universal primer was identified by using restriction sites indicated as non-cutting sites in a virtual digest performed in Geneious, allowing the universal primer to also serve as a cut site should the amplicon need to be incorporated into a plasmid for subsequent sequencing. Resulting amplicons were mapped to contig ends in Geneious after processing with Mixed Sequence Reader, [147] an algorithm designed to separate overlapping allelic amplicons using genomic sequence as a reference to call second peaks in sequencing chromatograms. PCR was then repeated with primers at locations indicated in this way in order to obtain more accurate amplicon sequences for use in

spanning the inter-contig gap. In order to increase the likelihood an amplicon would occur, a second universal primer was then incorporated as an adapter ligated to the ends of 1-2 kb gDNA fragments after digesting gDNA with restriction endonucleases that were determined to cut every 1-2kb in the genome, similar in principle to vectorette PCR. [148,149] Appropriate restriction sites were identified by importing the results of a virtual digest performed in Geneious into Excel (Microsoft, Redomond, WA) and graphing those with an amount of cut sites sufficient to digest the 3.2Mb genome into 1-2kb fragments. Plotting the fragments in this way identified restriction sites for elimination from consideration that clustered in the genome to yield an uneven fragment size distribution outside of the desired size range.

The three restriction enzymes indicated in this way as ideal candidates were AvaI, BglI (Promega, Madison, WI) and EcoRII. (New England Biolabs, Ipswich, MA) Single stranded PCR[150,151] was also employed with two primers situated close together on the same strand and in the same direction. The use of a second primer biased the reaction to amplify the intended target sequence rather than amplification due to nonspecific binding of the primer.   Finally, to resolve remaining scaffolding errors, Illumina paired end reads were mapped back to contigs in CLC Genomics Workbench to identify pronounced read disagreement sites at which contigs were split. The mapped, paired Illumina reads were then used in Genomics Workbench to align read overhangs and identify broken pairs so as to identify the final remaining contig end alignments.

Chapter 3 Results from the Genome of Strain BW-2

General genomic features

The genome of strain BW-2 consists of a circular chromosome of 4,103,727 bp with no evidence of extrachromosomal elements (Figure 14).   The average G+C content is 52.6% and there are 3,791 coding sequences including those encoding 41 tRNAs, 2 sets of 5S/16S/23S rRNAs and 1,572 hypothetical proteins. The average coding sequence length is 990.49 bp, representing 90.99% of the genome as coding sequence. (Table 1, 2, 3) The 269 nt region containing the probable OriC is positioned at noon on the circular chromosome map (Figure 14).   It is immediately downstream of the *dnaA* gene and followed by *dnaN* with a GC skew to 0 in the same area. The region contains 6 *Escherichia coli dnaA* box sequences with 0-1 mismatches and a 21 bp sequence of 90% identity to that found in the SS-5 probable OriC that produces no hits when blasted against NCBI databases (Figure 15).

Table 1 General Genome Features of strain BW-2

General Genome Features	
Genome Feature	BW-2
Size (bp)	4,074,140
G+C content (%)	52.6
No. Genomic features (CDS, RNAs)	3,838
No. CDS	3,791
Avr. CDS length	990.49
Protein coding density (%)	90.99
No. rRNAs	6
No. tRNAs	41
No. Hypothetical/unknown function	1572
Avr. Intergenic length (bp)	116.37

Table 2 Coding Sequence Classification of Open Reading Frames (ORFs) of Strain BW-2

BW-2 CDS Classifications	No.	
o : ORF of unknown function	1572	41.51%
Unclassified	1563	41.27%
e : enzyme	330	8.71%
pe : putative enzyme	83	2.19%
s : structure	57	1.51%
f : factor	44	1.16%
t : transporter	38	1.00%
r : regulator	27	0.71%
cp : cell process	18	0.48%
pt : putative transporter	13	0.34%
c : carrier	10	0.26%
pf : putative factor	7	0.18%
pc : putative carrier	6	0.16%
ph : phenotype	4	0.11%
pm : putative membrane component	4	0.11%
m : membrane component	4	0.11%
pr : putative regulator	3	0.08%
lp : lipoprotein	2	0.05%
ps : putative structure	2	0.05%

Table 3 Distribution of Gene COG Functional Classes of Strain BW-2

Process	Class ID	Description	CDS	%
POORLY CHARACTERIZED	R	General function prediction only	347	9.16%
CELLULAR PROCESSES AND SIGNALING	T	Signal transduction mechanisms	333	8.79%
INFORMATION STORAGE AND PROCESSING	L	Replication, recombination and repair	327	8.63%
POORLY CHARACTERIZED	S	Function unknown	234	6.18%
METABOLISM	E	Amino acid transport and metabolism	209	5.52%
METABOLISM	C	Energy production and conversion	197	5.20%
CELLULAR PROCESSES AND SIGNALING	M	Cell wall/membrane/envelope biogenesis	196	5.18%
CELLULAR PROCESSES AND SIGNALING	N	Cell motility	172	4.54%
INFORMATION STORAGE AND PROCESSING	J	Translation, ribosomal structure and biogenesis	169	4.46%
CELLULAR PROCESSES AND SIGNALING	O	Posttranslational modification, protein turnover, chaperones	153	4.04%
METABOLISM	P	Inorganic ion transport and metabolism	151	3.99%
INFORMATION STORAGE AND PROCESSING	K	Transcription	149	3.93%
CELLULAR PROCESSES AND SIGNALING	D	Cell cycle control, cell division, chromosome partitioning	130	3.43%
CELLULAR PROCESSES AND SIGNALING	U	Intracellular trafficking, secretion, and vesicular transport	113	2.98%
METABOLISM	G	Carbohydrate transport and metabolism	113	2.98%
METABOLISM	H	Coenzyme transport and metabolism	111	2.93%
METABOLISM	Q	Secondary metabolites biosynthesis, transport and catabolism	70	1.85%
METABOLISM	F	Nucleotide transport and metabolism	53	1.40%
METABOLISM	I	Lipid transport and metabolism	48	1.27%
CELLULAR PROCESSES AND SIGNALING	V	Defense mechanisms	42	1.11%
CELLULAR PROCESSES AND SIGNALING	W	Extracellular structures	8	0.21%
CELLULAR PROCESSES AND SIGNALING	Z	Cytoskeleton	1	0.03%
INFORMATION STORAGE AND PROCESSING	A	RNA processing and modification	1	0.03%

Figure 14 Circular Genome Map of Strain BW-2
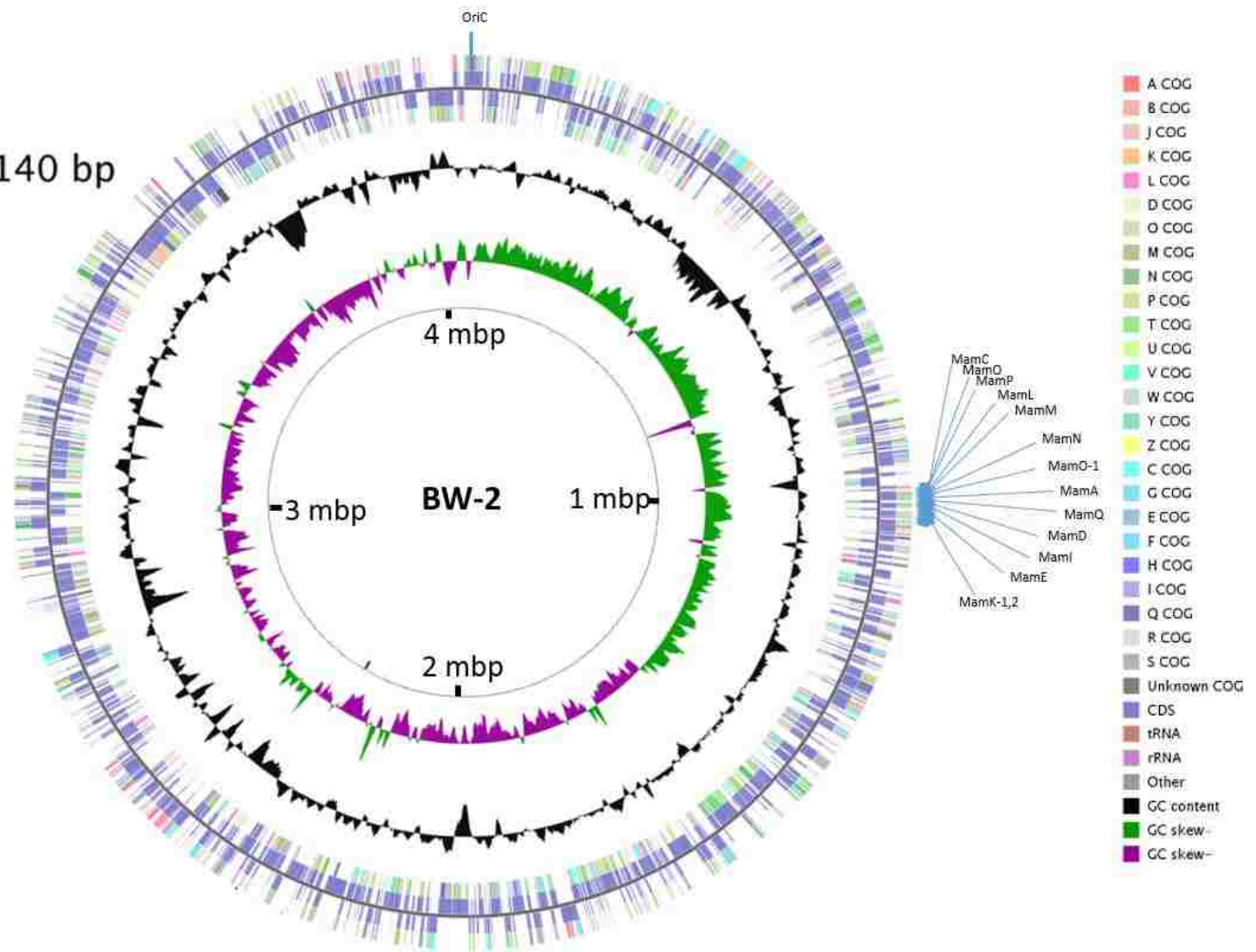
Figure 15 Probable Origin of Replication in the Genome of Strain BW-2

BW-2
269nt

DnaA   OriC   DnaN

cTGTGGATAAgatgtgaagcgatggggatagtgaacttatcccgatcttatcccaccCTATC
CACActcataagcgagtaataatacacagcgatggttaaacTAACTAAATGAATTTAAA
CAAaaaaattgacTTATCAACAtatcaacagccctaataataacaataggtattagatatata
taaacaatattattaataatagccTGTGGATAGTGTGGATAAgggTGTAGATAActtt
tgaataacttgttaataggatagaatc
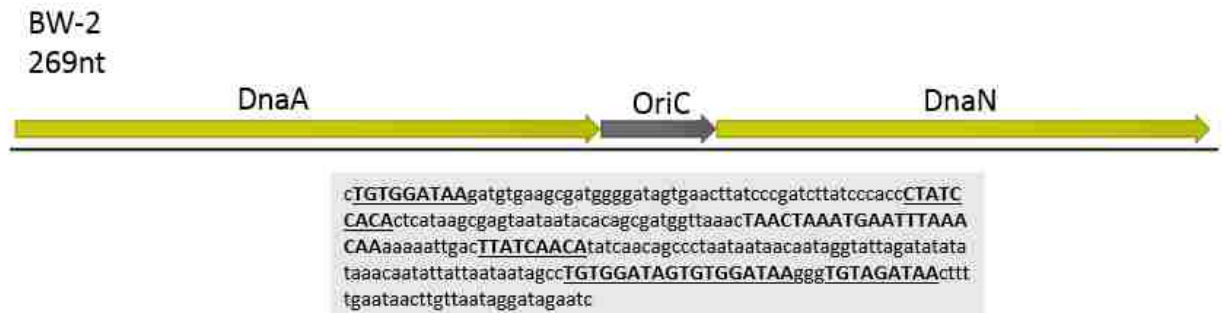
The *Escherichia coli* perfect DnaA box (ttatccaca) was searched for with no more than one mismatch. DnaA boxes are underlined and in bold.
21 bp alignment of BW-2 and SS-5 probable OriC sequences is in bold. (2 mismatches at 90.4% identity, no hits when blasted against NCBI)

Phylogenetically, strain BW-2 belongs to the order *Thiotricales*.[1] Its most closely related

MTB strain is SS-5 with a 16s rRNA gene sequence identity of 87.58%. The next closest MTB

strain is the newly isolated and as yet unfinished genome of another magnetic bacterium of the

Gammaproteobacteria, strain CG-1 which belongs to the order *Chromatiales*, at 85.12% 16s

rRNA gene sequence identity (Figure 16).

Figure 16 Phylogenetic Tree of 16 Species of MTB
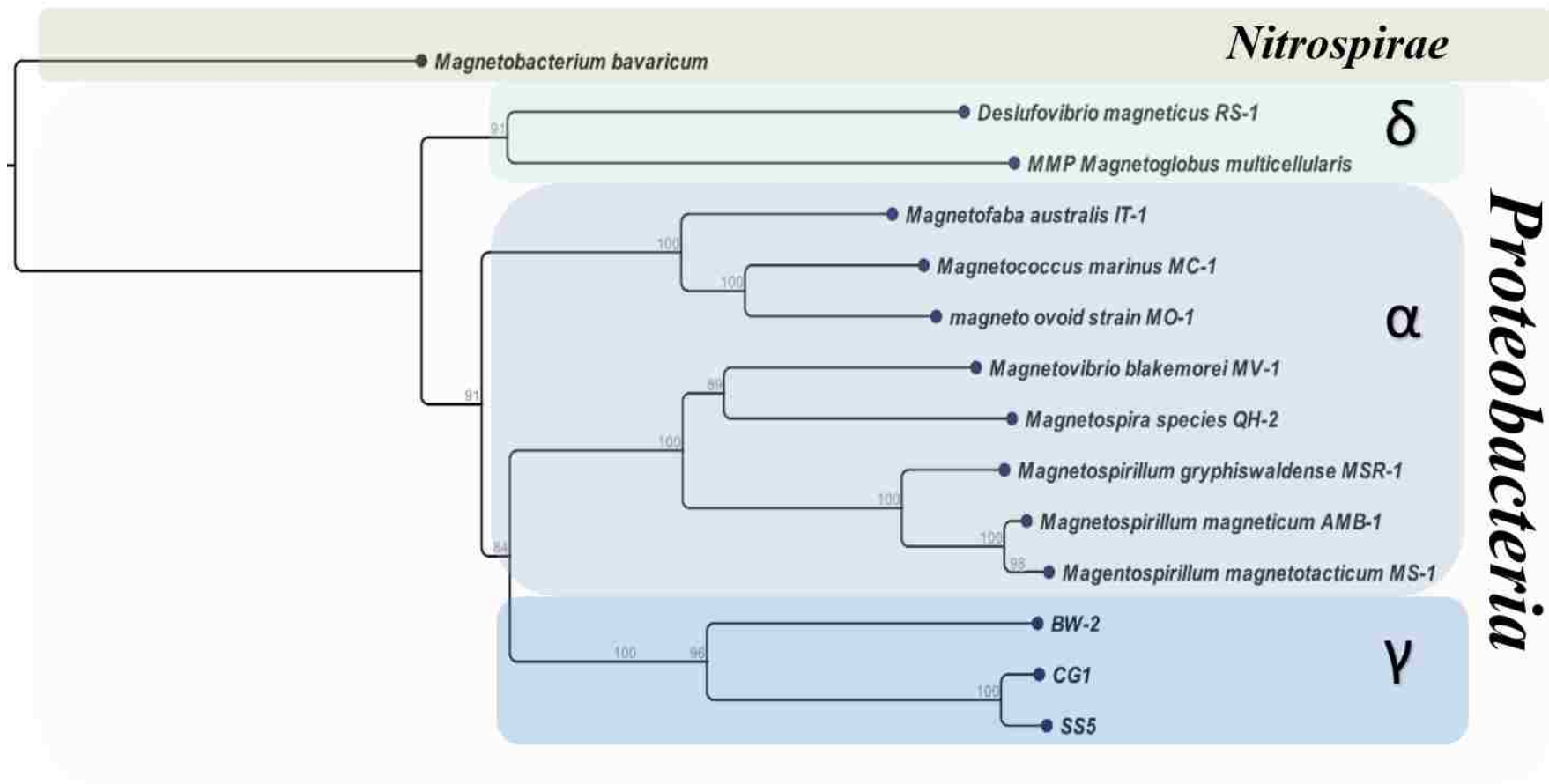
Tree based on 16s rRNA gene sequences using *Magnetobacterium bavaricum* to root. Bar represents 10% sequence divergence, numbers at branch points indicate bootstrap values out of 100 replicates.

Some Important Metabolic Features and the Genome of strain BW-2

Autotrophy: CO_2 Fixation

Strain BW-2 has been shown to grow chemolithoautotrophically using sulfide and thiosulfate as electron donors and sodium bicarbonate/CO_2 as the major source of carbon.[1] Coding sequences for 31 genes involved in carbon fixation were identified in the BW-2 genome, including two complete sets of form I "green-like" ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) genes for carbon fixation indicating it has the potential to employ the Calvin-Benson-Bassham cycle for autotrophy[1] (Figure 17). Both sets are in the same region 6,382 bp apart. The upstream set is incorporated into a full α-carboxysome operon[23,32,152–154] with shell protein gene *csoS2,* carbonic anhydrase encoding *csoS3,* shell protein genes *csoS4A, csoS4B, csoS1A and csoS1B,* a bacterioferritin, a possible dehydratase, a hypothetical gene and a RuBisCO operon transcriptional regulator gene in the given order (Figure 17). Carboxysomes are polyhedral intracellular compartments some of which are known to be involved in carbon fixation by enclosing and concentrating RuBisCO, (Figure 18) making up over 60% of the carboxysome protein.[32,154] The carboxysome specific carbonic anhydrase, *csoS3*, is thought to catalyze cytoplasmic bicarbonate to CO_2 which is concentrated within the carboxysome and is the main substrate for RuBisCO.[155]   This results in the favoring of the carboxylase reaction over the oxygenase reaction, leading to the production of glycerate-3-phosphate for use in central carbon metabolism.[152,153] One copy of a bicarbonate transporter gene *bicA* with high homology to the cyanobacterial Na^+-dependent HCO_3^- transporter *bicA* is also present 3 kb upstream of this carbon fixation gene cluster. Downstream is the second set

of *cbb* genes in the order *cbbL* large chain, *cbbL* small chain, a hypothetical CDS, *cbbQ, cbbO*, an

adenosylhomocysteinase and *csoS1D*. Although the full complement of genes necessary for

their expression is present, carboxysomes have not yet been visualized in BW-2.

Figure 17 Rubisco and Carboxysome gene Regions in the Genome of Strain BW-2
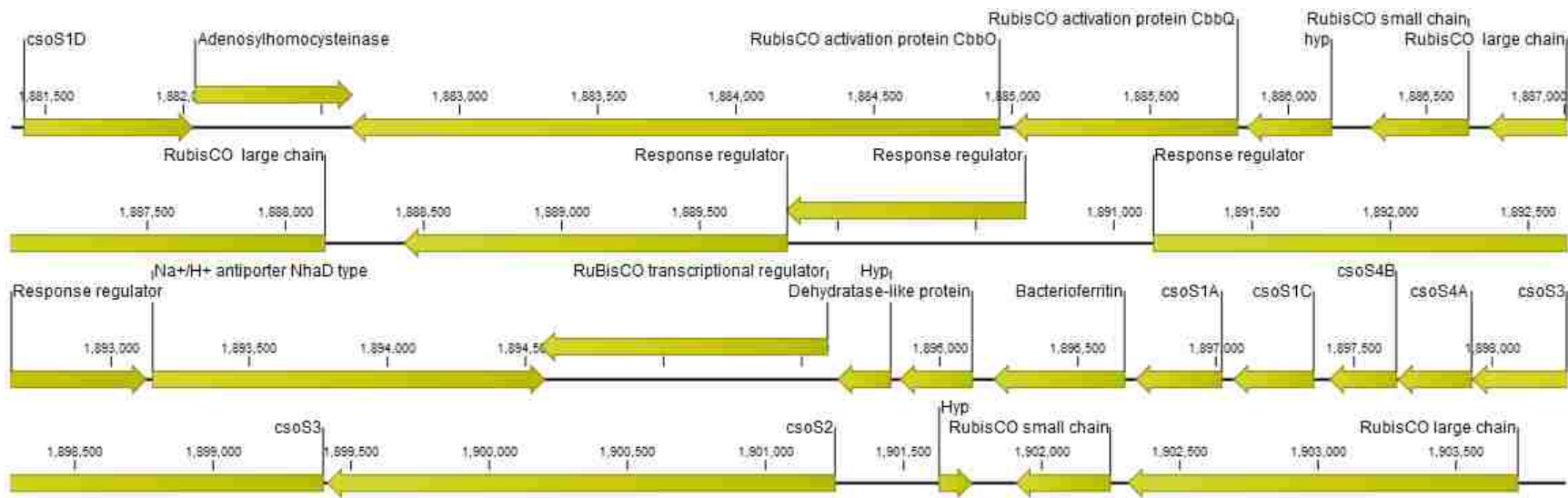
Arrows represent genes, numbed hash marks represent genomic location in bases relative to origin of replication. The genes csoS2, *csoS4A, csoS4B, csoS1A, csoS1B and csoS1D code for c*arboxysome shell proteins *csoS3 encodes carbonic anhydrase, hyp = hypothetical gene*

Figure 18 Metabolic Overview Diagram of Strain BW-2
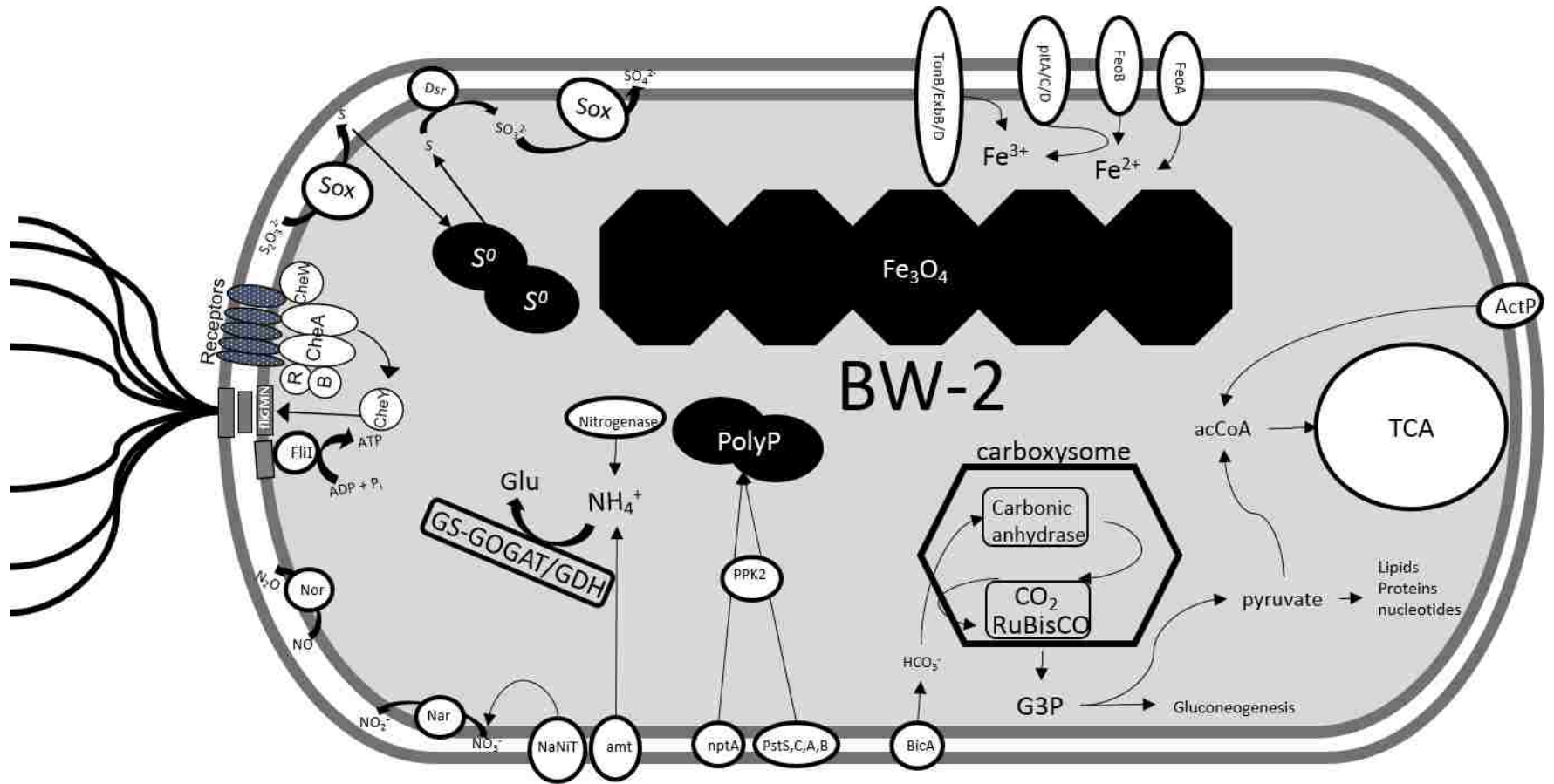
The gene compliment for the tricarboxylic acid cycle (TCA) cycle is complete in the genome of strain BW-2. Not surprisingly, in light of the presence of the complete RuBisCO system, genes enabling the TCA cycle to operate in the reductive direction were not found. Genes necessary for carbon flow through glycolysis are present, however gluconeogenesis is either nonfunctional due to a missing fructose-1,6-bisphosphatase or more likely catalyzed by an unknown nonorthologous enzyme, in that the largest group of confirmed and predicted non-homologous isofunctional enzymes is within carbohydrate metabolism.[156]

The gene for the acetate transporter acetate permease (*actP*) is present along with those encoding all necessary enzymes to convert acetate to acetyl-CoA making it available to the TCA cycle. Previous attempts to grow strain BW-2 heterotrophically on acetate were however unsuccessful[1] indicating this pathway may not be functional or is not so under currently used culture conditions. Some enzymes for sugar metabolism are present as would be expected for the resupply of intermediates in anabolic pathways. However, consistent with the observed obligate chemolithoautotrophic growth, no identifiable sugar transporter genes are present in the genome of strain BW-2.

Iron Metabolism

As biomineralizers of magnetosome iron crystals, MTB contain several orders of magnitude more iron than non-magnetotactic bacteria with their dry mass composed of up to 3% iron.[16] In fact, some MTB strains have been known to synthesize up to 16.7 mg magnetite liter^{-1} day^{-1}.[157] Therefore, Fe metabolism is obviously importance to MTB including its uptake, transport, oxidation and reduction.

The genome of strain BW-2 contains genes encoding an iron uptake ABC transporter, two copies each of *pitA* and *pitC* and three copies of *pitD*. In addition, there are ferrous iron transport proteins FeoA and FeoB.   BW-2 also has the *tonB/exbB/exbD* complex for ferric-siderophore transport and two non-ribosomal peptide synthetases coding sequences for siderophore synthesis, one contiguous with an amino acid ABC transport gene cluster and the other contiguous with another non-ribosomal peptide synthetase. The existence of other iron transport mechanisms can't be discounted. For example, there is an *yfeX* gene present, the product of which has been shown to promote Fe extraction from heme and participate in an encapsulated Fe acquisition system. [158] A heavy metal transporter was also identified as well as many uncharacterized proteins in BW-2 that are highly homologous to cation transporters in other bacteria. A coding sequence for the iron uptake, oxidation and storage protein bacterioferritin[159,160] was also identified adjacent to RuBisCO genes as well as a ferric iron uptake regulator protein (*fur*) gene. The BW-2 genome also contains the mam genes involved in iron redox control during magetosome synthesis, *mamBMEPTH*.

Nitrogen Metabolism

Cells of strain BW-2 have been shown to possess nitrogenase activity as evidenced by their ability to reduce acetylene under nitrogen-limited conditions suggesting they fix atmospheric dinitrogen (N_2).[1] The complete nitrogenase regulon is present in three genomic areas (Figure 19). The regulon consists of five separate operons including the *nifKDHTNY* operon which encodes the nitrogenase enzyme and is immediately upstream of *nifENX* which encodes nitrogenase cofactors. The *nifBQ* operon along with nitrogenase associated ferredoxins, is a little over 0.57 Mb upstream, and also encodes nitrogenase cofactors. The nitrogenase expression regulation operon *nifLA* is located immediately upstream of the *nifUSVM* operon and is responsible for the processing of a nitrogenase subunit and the *nifWZ* operon involved in electron transfer to nitrogenase when functioning. This cluster of *nifLA*, *nifUSVM* and *nifWZ* is about 1.76 Mb upstream of the *nifENX* and *nifHDKTNY* gene clusters.   A *nifJ* gene encoding the pyruvate-flavodoxin-oxidoreductase protein involved in the electron transfer to nitrogenase, does not appear to be present although there are three nitrogenase associated pyruvate-ferredoxin-oxidoreductase enzyme coding genes in strain BW-2 to fulfill this role.

Figure 19 Nif Regulon of Strain BW-2
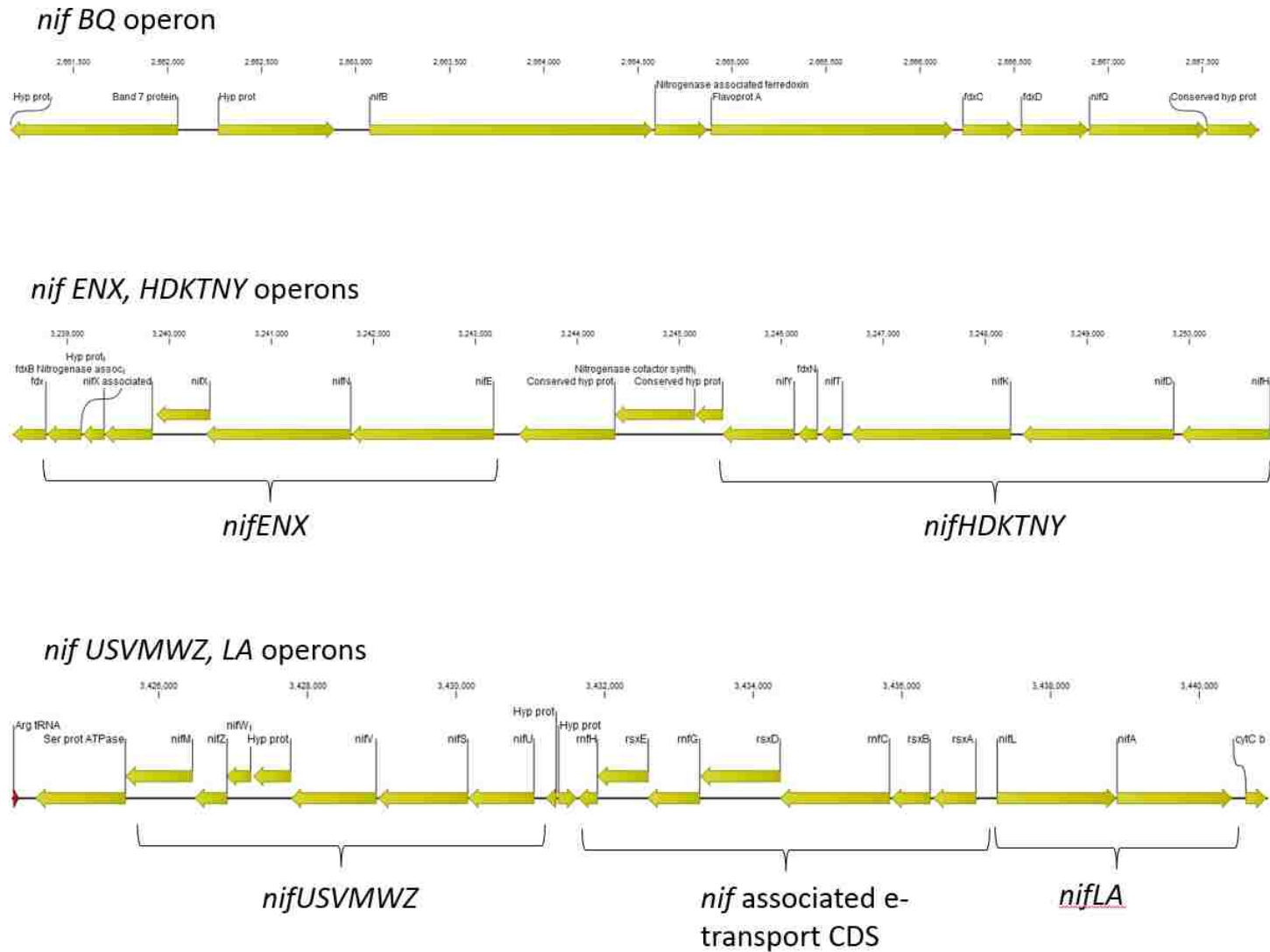
Arrows indicate genes, brackets indicate operons, and numbers at hash marks indicate genomic position by base pair numbered from origin of replication

Strain BW-2 utilizes ammonia as a source of nitrogen for growth. Coding sequences were found for an ammonium transporter gene, *amt*, as well as for glutamine synthetase/glutamate synthase (bacteria-like GS-GOGAT cycle) and glutamate dehydrogenase for the assimilation and fixation of ammonia when available in the environment (Figure 18). Strain BW-2 does not appear to be capable of using nitrate as a sole source of nitrogen for growth as genes for an assimilatory nitrate reductase were not identified.

Denitrification is the sequential dissimilatory reduction of nitrate or nitrite to a gaseous end product (nitrous oxide or dinitrogen) and is a globally important environmental biogeochemical process since it results in the loss of fixed nitrogen from an ecosystem.[101,121,161] Strain BW-2 appears to have only part of the complete denitrification pathway. According to its genome, strain BW-2 has the potential for anaerobic growth using certain nitrogen oxides as terminal electron acceptors. There are two nitrate/nitrite transporter genes present immediately upstream of the respiratory nitrate reductase *narGHJI* genes for the reduction of nitrate to nitrite (Figure 21). Despite the presence of the genes for the dissimilatory reduction of nitrate to nitrite, it is noteworthy that attempts to grow strain BW-2 anaerobically with nitrate in the laboratory were unsuccessful. Strain BW-2 appears to be missing any type of dissimilatory nitrite reductase preventing the further reduction of nitrite to nitric oxide. However, the genes for nitric oxide (NO) reductase *norCBDQD* are present and located immediately upstream of the nitric oxide reductase regulatory and activation genes potentially allowing strain BW-2 to reduce nitric oxide to nitrous oxide (N_2O) (Figure 20). No evidence for a nitrous oxide reductase (*nos* genes) was found in the genome of strain BW-2.

Figure 20. The Sequential Reductive Pathway of Denitrification in Prokaryotes Present in Strain BW-2.

Blue boxes indicate nitrogen oxides being reduced (except N_2), teal arrows indicate enzymes mediating the reduction, blue arrows indicate genes for those enzymes present in the genome of strain BW-2.
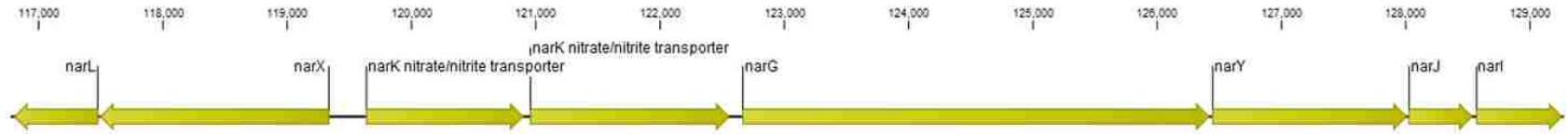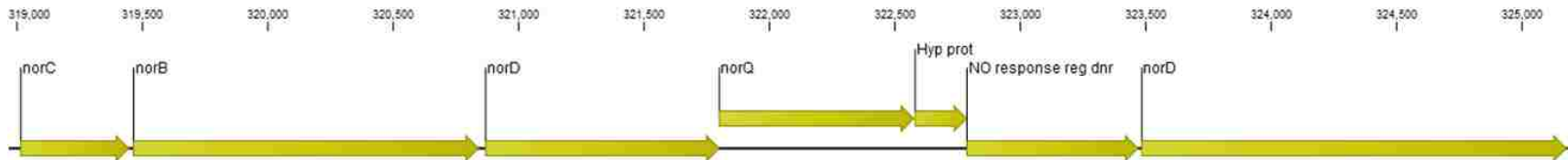
Arrows indicate genes, numbers at hash marks indicate genomic position by base pair numbered from origin of replication

As a sulfide- and thiosulfate-oxidizing bacterium, strain BW-2 has numerous genes related to sulfur metabolism scattered throughout its genome. Coding sequences were found for the sulfur-oxidation system SOX encoded by *soxAX, soxYZ, soxB*, the thioredoxin encoding gene *soxW* and the sulfide dehydrogenase encoding gene *soxF*. Five copies of *soxF*, three copies of *soxW* and a copy of homologous protein *soxH* are also present. The genes *soxCD* are absent*,* typical of other *Gammaproteobacteria* that oxidize sulfur using the SOX system.[5] It has been suggested that sulfur or polysulfide is the product of the SOX reactions without *soxCD*[5] when thiosulfate is the substrate. Although it has also been observed in some organisms that when thiosulfate is not present, *soxCD* is not expressed.[162] This might explain why cultures containing thiosulfate as the sole electron donor result in cells of BW-2 containing sulfur inclusions, but do not produce them initially in cultures when the energetically preferable sulfide is included in the medium. Once a culture with both sulfide and thiosulfate ages and becomes depleted of the more favorable sulfide, (as evidenced by the reduced intensity of the distinct sulfide odor), cells begin producing the inclusions presumably upon switching to the lower energy thiosulfate (Figure 18). As a culture with cells containing inclusions ages, the inclusions are then observed to deplete until they disappear, indicating a switch to stored S^0 once the higher energy exogenous sulfide and thiosulfate are depleted. It has been demonstrated in some strains of chemolithoautotrophic sulfur-oxidizing bacteria that these sulfur globules are formed as intermediates, and are then further oxidized by proteins encoded by the Dsr reverse-acting sulfite reductase system, that include proteins closely related to dissimilatory sulfite reductases that operate in the oxidative direction. [5,163]   A complete set of

the 16 Dsr system encoding genes including those for the membrane complex DsrMKJOP and the dissimilatory sulfite reductase DsrAB, as well as DsrCC4EFHLNRS. Once sulfur is converted to sulfite in this way it can then be further oxidized to sulfate by the SOX system.[5] It is also possible that the YedYZ membrane associated reductase (often annotated as sulfite oxidase) is active in the reduction of sulfur species as, although its functions are largely unknown it has been shown to act upon sulfur containing substrates.[164]

No genes for sulfate reduction, either assimilatory or dissimilatory, were identified. Reduced sulfur species required for assimilatory metabolism such as sulfide or cysteine are likely acquired exogenously.

Central Phosphate Metabolism

As phosphate is a major growth limiting resource in most aquatic ecosystems, microbes in these habitats generally have well-developed phosphate acquisition systems.   Cells of strain BW-2 produce numerous phosphate inclusions[1] and have numerous phosphorous metabolism genes. In the interest of acquiring phosphate from the environment, BW-2 has 13 phosphate transport system genes. Among them are coding sequences for the high inorganic phosphate affinity PstSCAB transporter complex as well as the PhoR and PhoB two component regulatory system that make up the PHO regulon. Limiting phosphate conditions lead to the autophosphorylation of the sensory kinase PhoR which activates PhoB by phosphorylation (Figure 22).   PhoB is a positive transcription regulator of the phosphate regulon that is present in 5 copies in BW-2.   In non-phosphate limiting conditions, PhoB may in turn inhibit the kinase

activity of PhoR, leading to interaction of PhoR and PhoU to inhibit PhoR and dephosphorylate and inactivate PhoB.[165,166] BW-2 also has multiple copies encoding an ATPase related to phosphate starvation-inducible protein, PhoH, that may be involved in lipid and phospholipid metabolism and RNA modification.[167]

Additionally, the BW-2 genome contains genes that encode an NptA Na^+-dependent phosphate transporter. Phosphate inclusions result from the polymerization of phosphate mediated by a polyphosphate kinase,[23,32,166,168] which is present in BW-2 as *ppk2*. (Figure 18)

Polyphosphate kinase and exopolyphosphatase are the enzymes typically involved in the production of phosphate granules. Coding sequences for both enzymes were found each as two copies in the genome of strain BW-2.

Figure 22 PHO System Regulation
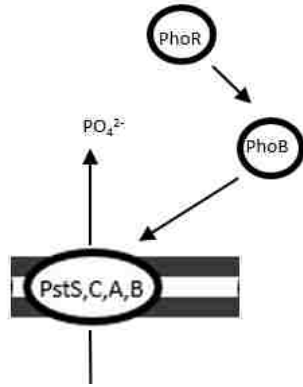
In phosphate limiting conditions PhoR activates by autophosphorylation, then in turn activating by phosphorylation PhoB which induces transcription of the Pst phosphate transporter regulon	When phosphate availability is not limiting to growth, PhoB may inhibit the autophosphorylation of PhoR, leading to its dimerization with PhoRU, further inhibiting the activity of PhoR
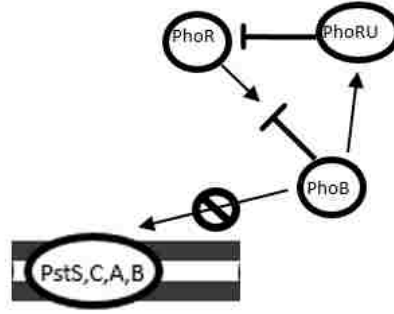
Transporters

There are 128 membrane transport related coding sequences in the genome of BW-2, typical of the number expected in an autotrophic prokaryote.[23,169] Of these, 11 are ABC transport genes coding for one full and two partial ABC transporters. Among those identified were complete set of genes representing a branched chain amino acid ABC transport system, *livFGMHJ* and the associated high-affinity leucine-specific periplasmic binding protein LivK. There is also a double copy of the PhnD phosphate binding periplasmic subunit of the alkylphosphonate ABC transporter present in BW-2, but without the other components. Another two coding sequences in BW-2 homologous to an ABC transporter involved in

oligopeptide transport was also found to be partially present in SS-5. In BW-2, there are permease subunits *oppB* and *oppC*, and no subunit *oppA*, indicating loss of the other components due to genomic rearrangement since the last common ancestor of the two strains.

Three Na^+ transporters were identified, including a complete set of a seven subunit Mrp (*A-G*) Na^+/H^+ transporter, the Na^+/H^+ antiporter NhaD in 2 copies, as well as the NptA Na^+ dependent phosphate transporter.

74 protein secretion transport related sequences are present in BW-2, including 31 type IV protein encoding genes for pilin and fimbriae synthesis.

31 type I proteins for autoaggregation and biofilm formation are present in BW-2 in multiple copies. There are 20 copies of T1SS secreted agglutinin alone in addition to multiple copies of *lapBCE*. These systems are clearly active when growing BW-2 in sulfide containing medium with a headspace of air, as cells form floating aggregations and adhere to the side of the jar at and near the surface of the liquid phase of the medium. There are also 4 type II protein secretion genes, 2 more that are involved in translocation to the membrane and 4 twin-arginine translocation system genes for export of folded proteins.

There are 9 additional protein encoding genes involved in the transport of cations, including the Mg/Co/Ni transporter gene *mgtE*, the Mg/Co efflux transporter gene *corC* in 3 copies, and the Cu efflux protein encoding sequence *copA* in 5 copies.

18 more coding sequences are members of the previously mentioned Ton/Tol transport system. Three of these are involved in Fe siderophore transport. This includes coding sequences for the periplasmic binding protein TonB in 2 copies, the transport protein TolQ in 3

copies, TonR Ferric receptor in 4 copies and TolR siderophore transport protein in 3 copies. Also identified are sequences coding for the TolQ proton channel, the TolB colicin uptake protein, and the heme uptake receptor TonR in 4 copies.

Chemotaxis and signal transduction

There are multiple protein coding sequences for the versatile cyclic second messenger cAMP signal in strain BW-2. Genes were identified that code for the cAMP receptor protein FNR, multiple copies of the cAMP binding protein CGA and 14 copies of an adenylate cyclase.

There were also genes identified as encoding proteins with redox sensor PAS/PAC domains along with GGDEF and EAL domains which are involved in the synthesis and phosphorylation of cyclic second messengers.[60] These proteins react to their redox environment as assessed by their PAS/PAC domains. In this case by synthesis (GGDEF) and phosphorylation (EAL) of the second messenger cyclic di-GMP.[170] 61 of these redox sensor signaling proteins were identified in BW-2 in addition to 2 with the PAS/PAC redox sensing domain only.

As an aerotactic microbe, BW-2 has the gene coding for the aerotaxis receptor protein Aer. The BW-2 genome also contains many copies of genes coding for the methyl accepting chemotaxis proteins. 35 are identified in BW-2, including *tsr*, the methyl transferase CheR, and the CheC inhibitor of methyl accepting chemotaxis proteins. There are many copies of genes encoding the CheW/CheA/CheY system for signal transduction to the flagellum including CheR

and CheB, as well as the chemotaxis response phosphatase CheZ. The Che system transduces

signal to flagellar motor switch proteins FliG, M, N, which were are present. [60] (Figure 18)

Histidine kinases are proteins that are typically transmembrane enzymes that mediate

signal transduction across cellular membranes.[60] In addition to the previously discussed CheA,

there are 45 more genes encoding histidine kinases in BW-2. 3 of these are part of two

component response regulator systems.[60,166,171] BW-2 also has one histidine kinase response

regulator. There are two histidine kinase transcriptional regulators specific to the sigma factor

54, which is unlike most prokaryote sigma factors in that it is not related to the more commonly

employed sigma 70, indicating these transcriptional regulators may be involved in the initiation

of specialized systems.[172] In addition, there are 30 single component, non-histidine kinase

response regulator genes in the genome of strain BW-2.

Prophages

While there are several phage and related genes scattered throughout the genome of strain

BW-2 including 5 integrases, there are only two regions with clusters that represent nearly complete

prophage regions (Figure 23). The first is about 7 Kb in length with a GC percentage of 51.81%. It

contains 7 phage associated genes including 3 conserved hypothetical proteins, two phage tail proteins

and a conserved phage related protein. The second is about 35.8 Kb with an average GC content of

40.13%, containing 8 phage related proteins.

Figure 23 Prophage Genomic Regions of Strain BW-2
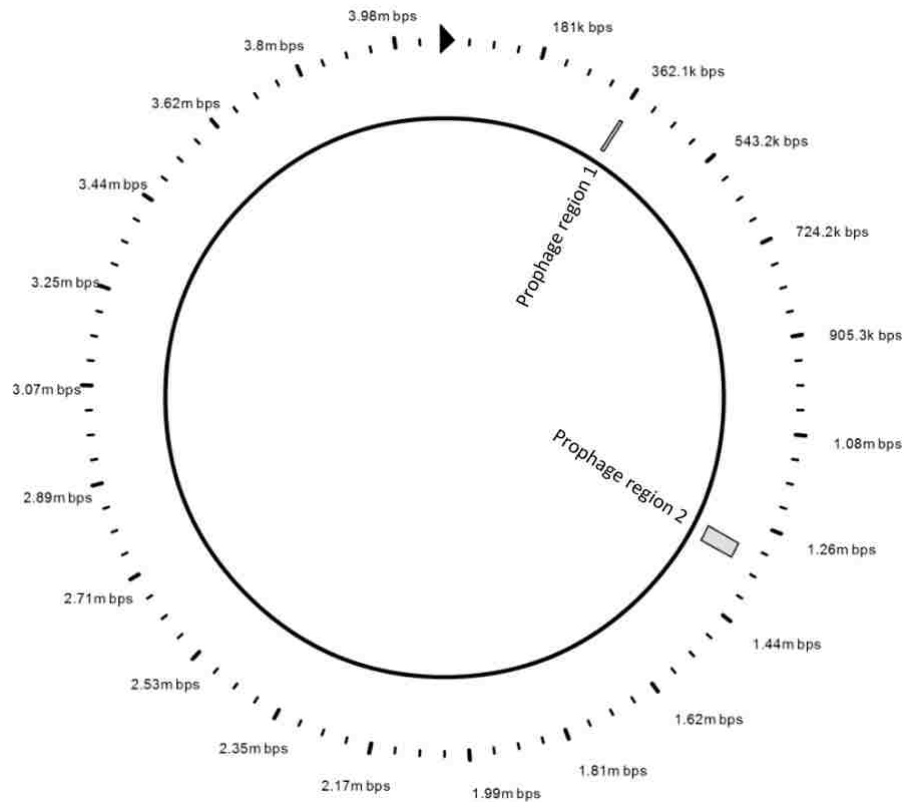
Numbered hash marks indicate genomic position relative to origin of replication as represented by the arrow at noon position. Prophage regions represented by shaded boxes.

Magnetosome Related Genes

An entire set of known magnetosome related genes is arranged in the genome of strain BW-2 as a single contiguous cluster (Figure 24, Table 4). Immediately downstream of this magnetosome gene cluster, are two coding sequences of high homology to a resolvase and a transposase. However, neither the resolvase nor transposase share significant homology to that found in any other strain of MTB. There are no other mobile element proteins or tRNAs in or near the other clusters. Additionally, there is no difference in surrounding region GC

content, and the phylogeny and synteny of the Mam proteins in BW-2 are highly homologous with those in the G*ammaproteobacterium* SS-5. Taken together, this seems to indicate that these genes were not acquired through the result of a recent horizontal gene transfer event, but rather through descent via a last common ancestor shared with SS-5. Further evidence of this is the fact that the majority of magnetosome proteins in BW-2 and SS-5 are most closely related to another unpublished magnetic *Gammaproteobacterium*, strain *Chromatiales* CG-1. The next most related magnetosome associated protein sequences are the corresponding proteins found in the genome of *M. magneticum*, of the *Alphaproteobacteria* (Table 4).

Figure 24 Magnetosome Gene Cluster of Strain BW-2
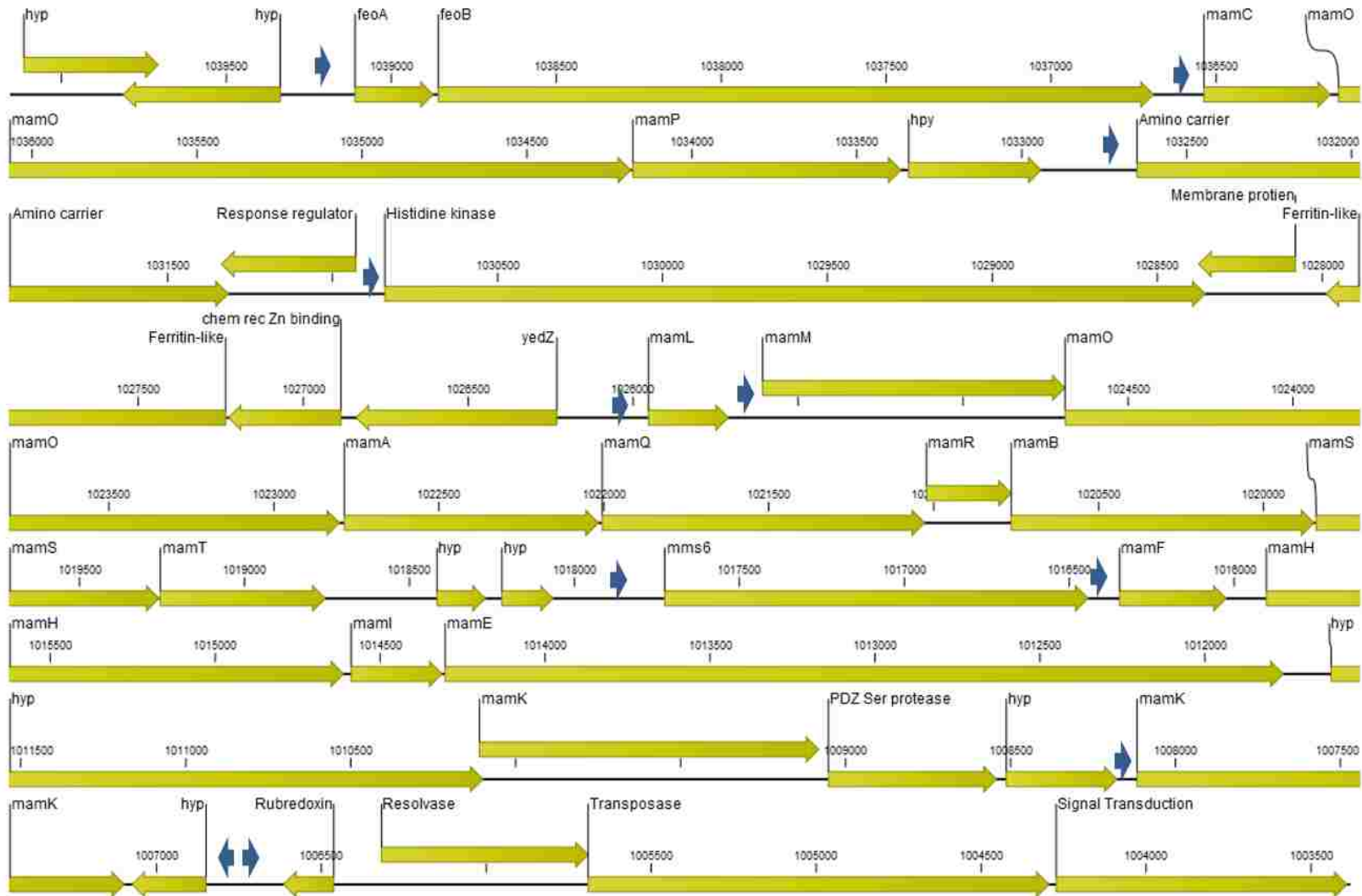
Green arrows represent genes, blue arrows represent characterized promoter sequences, and numbers at hash marks indicate genomic position relative to origin of replication.

Table 4 Magnetosome Cluster Gene of Interest Homologies in Strain BW-2

Mam protein	Magnetotactic bacterium with protein with highest sequence identity*	e-value	Identity (%)	Magnetotactic bacterium with protein with second highest sequence identity*	e-value	Identity (%)	Magnetotactic bacterium with protein with third highest sequence identity*	e-value	Identity (%)
feoA	Magnetite-containing magnetic vibrio MV-1	4.00E-16	44.44	Magnetospirillum sp. XM-1	2.00E-10	35.44	Magnetospirillum moscoviense BB-1	2.00E-10	37.5
feoB	Magnetite-containing magnetic vibrio MV-1	0	54.06	Magnetospirillum gryphiswaldense MSR-1 v2	0	46.72	Magnetospirillum sp. XM-1	0	46.86
MamC	Thiotrichales bacterium SS-5	2.00E-46	63.48	Unamed Chromatiales CG1	1.00E-45	63.48	Magnetospirillum sp. SO-1	4.00E-30	50
MamO	Thiotrichales bacterium SS-5	0	52.4	Unamed Chromatiales CG1	0	48.58	Thermobifida fusca YX	0.004	35.24
MamP	Thiotrichales bacterium SS-5	2.00E-111	73.13	Unamed Chromatiales CG1	3.00E-109	72.64	Magnetococcus sp. MC-1	9.00E-51	44.85
A.A. transporter	Chromatiales CG1	2E-161	53	Magnetococcus marinus Mc-1	9.00E-143	47	Magnetofaba australis IT-1	7.00E-137	46
His Kinase	Chromatiales CG2	2E-175	57.96	Thiotrichales bacterium SS-5	4.00E-167	56.85	None		
Ferritin	Chromatiales CG1	5.00E-60	50.25	Magnetococcus marinus MC-1	6.00E-53	53	None		
Serine protease	SS-5	3.00E-09	38	Magnetospira sp. QH-2	6.00E-08	40	Magnetospirillum magneticum AMB-1	5.00E-07	37
MamL	Thiotrichales bacterium SS-5	1.00E-27	61.25	Unamed Chromatiales CG1	3.00E-27	61.25	Magnetospirillum magneticum AMB-1 v2	2.00E-03	28.33
MamM	Unamed Chromatiales CG1	3.00E-164	76.71	Thiotrichales bacterium SS-5	6.00E-163	75.17	Magnetospirillum magneticum AMB-1 v2	4.00E-110	52.86
MamO	Thiotrichales bacterium SS-5	0.00E+00	52.08	Unamed Chromatiales CG1	0.00E+00	51.24	Magnetospirillum sp. XM-1	2.00E-146	39.47
MamA	Unamed Chromatiales CG1	6.00E-43	41.15	Thiotrichales bacterium SS-5	1.00E-42	40.78	Magnetospirillum magneticum AMB-1 v2	8.00E-23	27.1
MamQ	Unamed Chromatiales CG1	2.00E-117	64.62	Thiotrichales bacterium SS-5	3.00E-115	57.41	Magnetospirillum magneticum AMB-1 v2	3.00E-51	42
MamR	Unamed Chromatiales CG1	1.00E-23	42.35	Thiotrichales bacterium SS-5	3.00E-21	41.18	Magnetospirillum magneticum AMB-1 v2	1.00E-13	30.23
MamB	Thiotrichales bacterium SS-5	2.00E-179	83.22	Unamed Chromatiales CG1	2.00E-176	81.82	Magnetospirillum magneticum AMB-1 v2	9.00E-112	51.35
MamS	Thiotrichales bacterium SS-5	2.00E-46	45.79	Unamed Chromatiales CG1	1.00E-46	42.71	Magnetospirillum magneticum AMB-1 v2	4.00E-19	39.34
MamT	Unamed Chromatiales CG1	1.00E-70	61.64	Thiotrichales bacterium SS-5	3.00E-62	54.72	Magnetospirillum magneticum AMB-1 v2	1.00E-49	50.31
Mms6	Thiotrichales bacterium SS-5	1.00E-103	50.35	Unamed Chromatiales CG1	1.00E-98	46.97	Magnetospirillum magneticum AMB-1 v2	3.00E-17	37.42
MamF	Unamed Chromatiales CG1	2.00E-52	68.52	Thiotrichales bacterium SS-5	2.00E-51	68.52	Magnetospirillum magneticum AMB-1 v2	6.00E-32	47.96
MamH	Unamed Chromatiales CG1	0.00E+00	71.29	Thiotrichales bacterium SS-5	0.00E+00	70.12	Magnetospirillum magneticum AMB-1 v2	4.00E-169	56.65
MamI	Unamed Chromatiales CG1	5.00E-24	64.71	Thiotrichales bacterium SS-5	1.00E-21	57.35	Magnetospirillum magneticum AMB-1 v2	1.00E-18	60.38
MamE	Unamed Chromatiales CG1	0.00E+00	47.52	Thiotrichales bacterium SS-5	0.00E+00	45.29	Magnetospirillum magneticum AMB-1 v2	5.00E-113	33.56
MamK-1	Unamed Chromatiales CG1 MamK-1	4.00E-132	53.87	Thiotrichales bacterium SS-5 MamK-1	4.00E-131	54.76	Magnetospirillum magneticum AMB-1 v2	6.00E-105	44.78
MamK-2	Unamed Chromatiales CG1 MamK-2	2.00E-153	60.65	Thiotrichales bacterium SS-5 MamK-2	5.00E-149	60.06	Magnetospirillum magneticum AMB-1 v3	2.00E-119	48.8

In BW-2 the upstream end of the magnetosome associated gene cluster consists of the aforementioned *feoA* and *feoB* genes. The *feoA* and *feoB* genes encode a membrane protein FeoA/FeoB, a small soluble protein that is involved in ferrous iron (Fe_2^+) transport in microorganisms.[89] In most MTB there are at least two clusters *feoAB* type genes, one specific to the MTB (*FeoAB*-like) and another *feoAB* cluster that has the same origin as in other bacteria. However BW-2 has only one set of *feoAB*.[42]

The *mamCOPD* cluster, which is found apart from the main cluster of magnetotaxis genes in some MTB is upstream and contiguous in this order with the other mam genes in BW-2. The gene *mamP* encodes an archetypal MamP protein that comprises one PDZ domain, a conserved domain enabling protein-protein interaction, and two magnetochrome domains.[66] The gene *mamO* encodes a protein that contains one N-terminus trypsin-like protease domain and one C-terminus TauE domain shown to be involved in the transport of anions across the membrane during taurine metabolism.[173] The *mamC* gene encodes MamC, an MTB specific protein that does not contain a conserved domain. *mamD* encodes for MamD, a MTB specific protein that also does not contain known conserved domain.

Immediately downstream from the *mamCOPD* cluster is the *mamAB* cluster. This cluster contains two copies of the *mamK* gene, as well as *mam* genes, *L, M, O, A, Q, R, B, S, T, D, F, H, I* and E and an *mms6* (Figure 24). This cluster is analogous to the *mamAB* operon described in *Magnetospirillum* species, therefore is referred to as the *mamAB* cluster as there is no evidence whether the genes in this putative operon are also transcribed under the control of a single promoter,[82,91] and there are in fact 11 identifiable potential promoters within the BW-2 *mam*AB cluster (Figure 24). The *mamAB* cluster is contiguous with the main

magnetosome gene clusters in the genome of strain BW-2 which contains 15 different *mam* genes with *mamI, mamE* and *mamK* present as 2 copies. The gene order of *mam* genes in strain BW-2 is generally conserved when compared to the *mamAB* operon of strain SS-5, and *Magnetospirillum* species of the *Alphaproteobacteria*, *Magnetovibrio blakemorei*[119] or *Magnetospira* sp. QH-2.[174] Notably, the putative cytoskeletal attachment protein *mamJ,*[84] once thought to be essential for the construction of the magnetosome chain[74,82,84] appears to be absent in strain BW-2.

Summary

Taken together, evidence from previous experiments and possible metabolic systems identified in the genome of strain BW-2 described here indicate that this organism has the potential as to fix carbon dioxide autotrophically as well to fix dinitrogen from the atmosphere. In addition, this strain utilizes ammonia as a nitrogen source and respires using reduced sulfur compounds such as sulfide and thiosulfate as electron donors and nitrate and nitric oxide as terminal electron acceptors. 128 established transporters of various types were identified. Few prophage genes were identified and only two clusters that represent near complete prophages. A compliment of 21 previously characterized magnetosome-associated genes found in other MTB were identified in the genome of strain BW-2 in a single contiguous cluster without the hallmark features of a genomic island which include the presence of transposase genes, flanking tRNA genes and GC content differing from that of the rest of the genome. The metabolic and magnetosome-related genomic features of BW-2 will be compared with those of

SS-5 and of other MTB in chapter five in the interest of revealing common metabolic features,

as of yet uncharacterized genes that may be involved in the magnetotactic phenotype, as well

as any evidence of how magnetosome-associated genes are acquired by strain BW-2

Chapter 4 Results from the Genome of Strain SS-5

General genomic features

The SS-5 genome consists of a single circular chromosome of 3,729,439 bp. There is no evidence to indicating the presence of extrachromosomal elements (Figure 25). With an average G+C content of 61.65%, there are 3,724 predicted genes including those encoding 51 tRNAs, 2 sets of 5S/16S/23S rRNAs and 6 miscellaneous RNAs and 1,480 coding sequences of unknown function. The average coding sequence length is 908.25bp with 88.24% of the genome comprised of protein coding genes (Table 5). The *dna*A gene is immediately followed by an 181 nt intergenic sequence containing two *E. coli* DnaA box sequences with 0 and 1 mismatch, which is immediately upstream of *dnaN*. In addition to a nearby GC skew shift to 0%, this region also contains a 21 bp sequence of 90.3% identity to the probable OriC sequence of strain BW-2, while producing no hits when blasted against NCBI databases (Figure 26). This region was therefore designated the probable OriC[175] and positioned at noon on the circular map.

Table 5 General Genomic Features of strain SS-5

General Genome Features	
Genome Feature	SS-5
Size (bp)	3,729,439
G+C content (%)	61.65
No. Genomic features (CDS, RNAs)	3,724
No. CDS	3,669
Avr. CDS length	908.25
Protein coding density (%)	88.24
No. rRNAs	6
No. tRNAs	51
No. Hypothetical/unknown function	1,480
Avr. Intergenic length (bp)	147.31

Table 6 SS-5 Coding sequence Classifications of putative genes in the genome of strain SS-5

SS-5 CDS Classifications	No.	
o : ORF of unknown function	1,480	40.43%
Unclassified	1,299	35.48%
e : enzyme	444	12.13%
pe : putative enzyme	104	2.84%
s : structure	69	1.88%
f : factor	58	1.58%
r : regulator	53	1.45%
t : transporter	40	1.09%
cp : cell process	23	0.63%
prc : putative receptor	16	0.44%
pt : putative transporter	14	0.38%
c : carrier	13	0.36%
m : membrane component	10	0.27%
pr : putative regulator	9	0.25%
pf : putative factor	8	0.22%
pm : putative membrane component	8	0.22%
pc : putative carrier	5	0.14%
ph : phenotype	4	0.11%
h : extrachromosomal origin	2	0.05%
rc : receptor	1	0.03%
lp : lipoprotein	1	0.03%

Table 7 Distribution of COG Functional Classes of ORFs in the genome of strain SS-5

Process	Class ID	Description	CDS	%
POORLY CHARACTERIZED	R	General function prediction only	373	10.19%
CELLULAR PROCESSES AND SIGNALING	T	Signal transduction mechanisms	291	7.95%
POORLY CHARACTERIZED	S	Function unknown	267	7.29%
INFORMATION STORAGE AND PROCESSING	L	Replication, recombination and repair	229	6.26%
METABOLISM	C	Energy production and conversion	228	6.23%
CELLULAR PROCESSES AND SIGNALING	M	Cell wall/membrane/envelope biogenesis	200	5.46%
METABOLISM	P	Inorganic ion transport and metabolism	187	5.11%
INFORMATION STORAGE AND PROCESSING	J	Translation, ribosomal structure and biogenesis	173	4.73%
METABOLISM	E	Amino acid transport and metabolism	170	4.64%
INFORMATION STORAGE AND PROCESSING	K	Transcription	167	4.56%
CELLULAR PROCESSES AND SIGNALING	N	Cell motility	133	3.63%
CELLULAR PROCESSES AND SIGNALING	O	Posttranslational modification, protein turnover, chaperones	130	3.55%
METABOLISM	H	Coenzyme transport and metabolism	119	3.25%
METABOLISM	G	Carbohydrate transport and metabolism	101	2.76%
CELLULAR PROCESSES AND SIGNALING	U	Intracellular trafficking, secretion, and vesicular transport	86	2.35%
CELLULAR PROCESSES AND SIGNALING	V	Defense mechanisms	78	2.13%
METABOLISM	Q	Secondary metabolites biosynthesis, transport and catabolism	77	2.10%
CELLULAR PROCESSES AND SIGNALING	D	Cell cycle control, cell division, chromosome partitioning	64	1.75%
METABOLISM	F	Nucleotide transport and metabolism	53	1.45%
METABOLISM	I	Lipid transport and metabolism	52	1.42%
CELLULAR PROCESSES AND SIGNALING	W	Extracellular structures	3	0.08%
CELLULAR PROCESSES AND SIGNALING	Z	Cytoskeleton	1	0.03%
INFORMATION STORAGE AND PROCESSING	A	RNA processing and modification	1	0.03%

Figure 25 Circular Genome Map of Strain SS-5
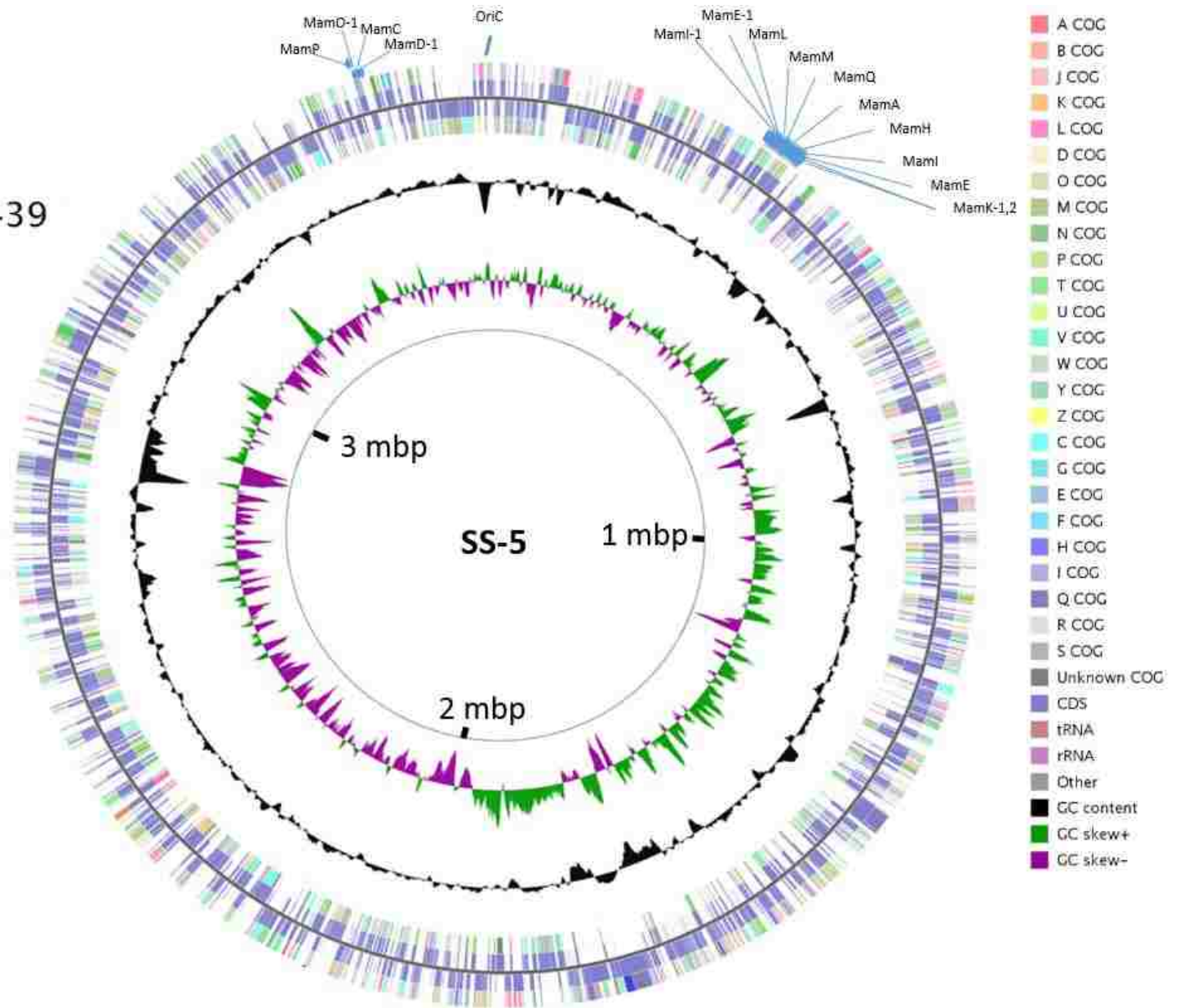
Figure 26 Probable Origin of Replication in the Genome of Strain SS-5
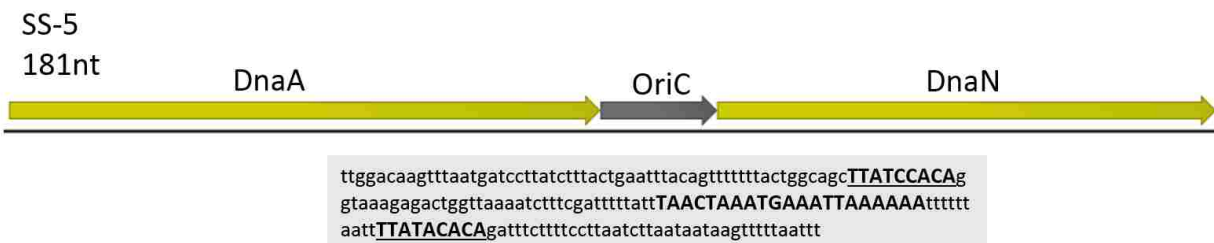
ttggacaagtttaatgatccttatctttactgaatttacagtttttttactggcagc**TTATCCACA**g
gtaaagagactggttaaaatctttcgattttttatt**TAACTAAATGAAATTAAAAAA**tttttt
aatt**TTATACACA**gatttcttttccttaatcttaataataagttttttaattt

Escherichia coli perfect DnaA box (ttatccaca) was searched for with no more than one mismatch. DnaA boxes underlined and in bold. 21 bp alignment of BW-2 and SS-5 probable OriC sequences is in bold. (2 mismatches at 90.4% identity, no hits when blasted against NCBI)

Strain SS-5 is phylogenetically most closely related to the uncultured, gammaproteobacterial MTB strain CG-1 of the *Chromatiales* order with a 94.93% 16s gene sequence identity. The next most related MTB is the cultured, gammaproteobacterial MTB strain BW-2 described in this work with an 87.58% 16s rRNA gene sequence identity (Figure 16).

Some Important Metabolic Features Based on Genomics of strain SS-5

Autotrophy: CO$_2$ Fixation

Strain SS-5 has been shown to grow chemolithoautotrophically using sulfide and thiosulfate as electron donors and sodium bicarbonate/CO$_2$ as the major source of carbon.[1] Coding sequences in SS-5 were identified as form I "green-like" ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) genes and indicating it has the genetic potential to employ the Calvin-Benson-Bassham cycle for autotrophy.[1] Additionally, the genome of strain SS-5 contains a full set of genes coding for carboxysomes.[23,32,152–154] As previously discussed, carboxysomes are polyhedral intracellular compartments (Figure 27) known to be involved in

carbon dioxide fixation by enclosing and concentrating RuBisCO, which makes up over 60% of the carboxysome protein.[32,154] The carboxysome-specific carbonic anhydrase, *csoS3*, likely catalyzes cytoplasmic bicarbonate to CO_2 which is concentrated within the carboxysome and is the main substrate for RuBisCO.[155] This results in the favoring of the carboxylase reaction over the oxygenase reaction leading to the production of glycerate-3-phosphate for use in central carbon metabolism.[152,153] In all, coding sequences for 38 genes involved in carbon dioxide fixation were identified in the SS-5 genome, including two sets of form I 'green-like' *cbbL* ribulose bisphosphate carboxylase (RuBisCO) genes (Figure 28). One set of RuBisCO genes are intact and integrated as part of the upstream end of the α-carboxysomal operon encoded by *csoS2, csoS3, csoS4A, csoS4B, csoS1A, csoS1B* a bacterioferritin gene, a possible dehydratase and a RuBisCO operon transcriptional regulator gene in the given order. A transposase and a sequence with homology to mobile element associated proteins is 7.9 kb upstream to the long chain *cbbL* gene (RuBisCO large subunit) as well as a tRNA gene, indicating possible rearrangement events mediated by transposase proteins that use tRNA genes as attachment sites.[176] There are also 4 sequences immediately upstream of the RubisCO large subunit gene that are closely related to restriction system genes involved in DNA modification, usually as a defense against exogenously introduced DNA.[177] Among these DNA modifying genes is a gene coding for an abi protein. Little is known about abi proteins except that they are associated with genes which mediate immunity against bacteria derived antibiotics known as bacteriocins.[178] Just over a Mb downstream there is a fragment of the RuBisCO operon adjacent to mobile element, phage and tRNA coding sequences, containing the *cbbL* large chain sequence, the *cbbQ, cbbO* and the *cbbR* transcriptional regulator. Another cluster of RuBisCO

genes is in a region roughly 450 Kb downstream that shows signs of previous genomic

rearrangement. This incomplete cluster contains the RuBisCO large subunit gene but not the

small subunit. Immediately downstream of the large subunit coding sequence there are two

genes coding for RuBisCO activation proteins. Another 5.2 Kb downstream is a tRNA and an

integrase in another 1.2 Kb. Another Mb downstream there is a third, complete cluster of

RuBisCO genes.   This region contains the RuBisCO large and small chain *cbbL* sequences

upstream of the RuBisCO activation protein genes *cbbQ* and *cbbO*, however there is an

intervening 23S rRNA CDS between the *cbbQ* and *cbbO* activation sequences. Immediately

upstream of the RuBisCO large subunit there is a tRNA.

Figure 27 Cell Metabolic Overview diagram based on the genomics of strain SS-5
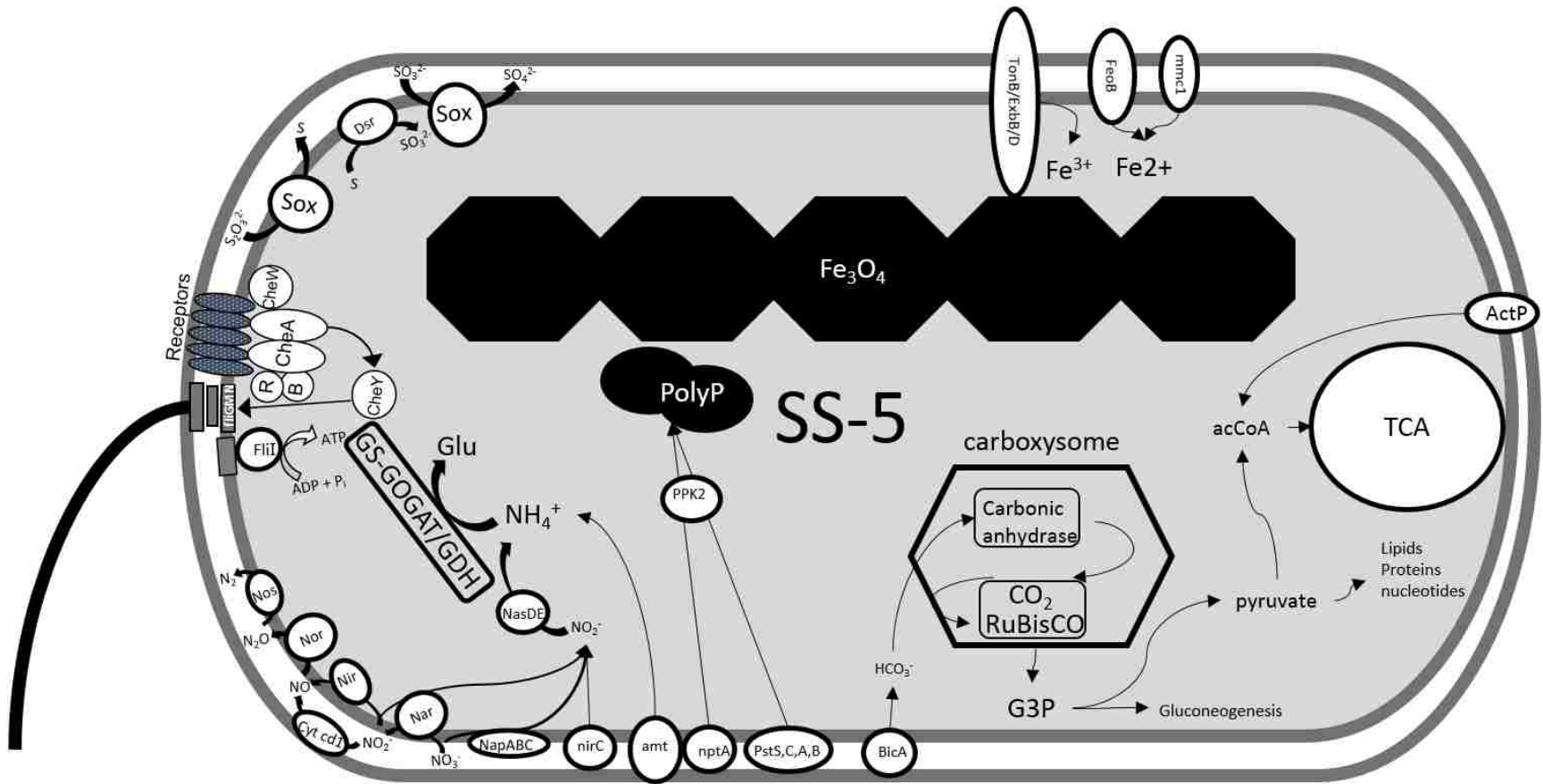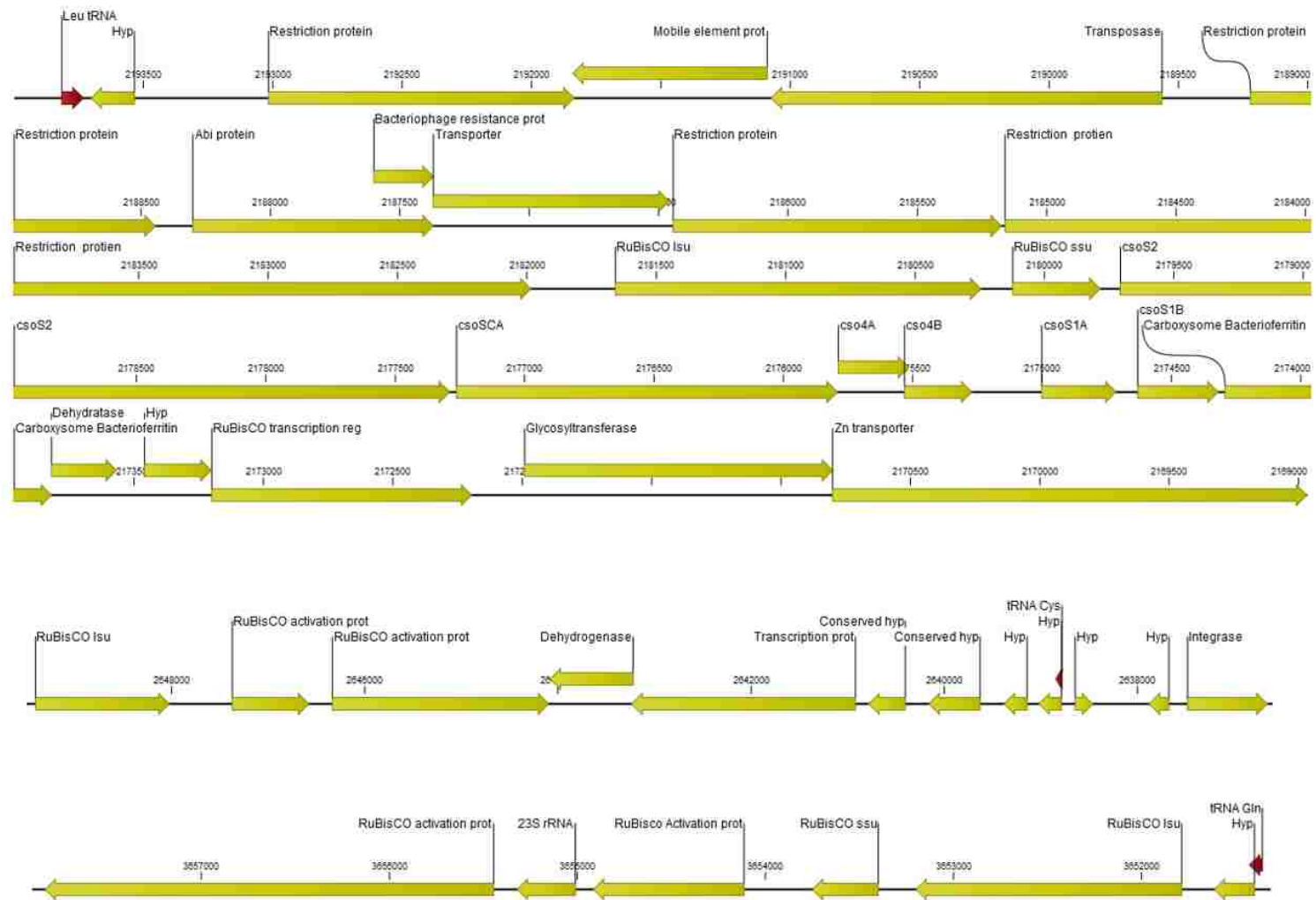
Arrows represent genes, numbered hash marks indicate genomic position in base pairs relative to origin of replication. The genes csoS2, csoS4A, csoS4B, csoS1A, csoS1B and csoS1D code for carboxysome shell proteins csoS3 encodes carbonic anhydrase, hyp = hypothetical gene

Central carbon metabolism

The gene compliment for the TCA cycle is complete in strain SS-5. Genes enabling the TCA cycle to operate in the reductive direction for autotrophy were not found and would be unnecessary in the presence of the complete Calvin-Benson-Bassham cycle. Genes required for carbon flow through glycolysis are present, however the gluconeogenesis doesn't appear to be complete due to a missing fructose-1, 6-bisphosphatase enzyme encoding gene. It may be that this reaction is catalyzed by an unknown nonorthologous enzyme, in that the largest group of confirmed and predicted non-homologous isofunctional enzymes is within carbohydrate metabolism.[156]

The acetate transporter acetate permease (*actP*) is present along with all necessary enzymes to convert acetate to acetyl-CoA making it available to the TCA cycle. Previous attempts to grow strain SS-5 heterotrophically on acetate were, however, unsuccessful,[1] indicating this pathway may not be functional or is not so under currently used culture conditions for this strain. Some enzymes for sugar metabolism are present, probably for the resupply of intermediates in anabolic pathways. Consistent with the observed obligate chemolithoautotrophic growth, no identifiable sugar transporters were identified in the genome of strain SS-5.

Iron Metabolism

In that MTB biomineralize iron crystals, cells of MTB contain several orders of magnitude more iron than non-magnetic bacteria. Some strains are known to synthesize up to 16.7 mg magnetite liter^{-1} day^{-1}.[157] The dry mass of MTB is typically composed of up to 3% iron.[16] Therefore, of obvious importance to MTB is iron metabolism, including its uptake, transport, oxidation and reduction.

The genome of strain SS-5 contains a ferrous iron transport protein coding FeoB sequence in addition to multiple copies of the *tonB/exbB/exbD* ferric-siderophore transport complex and a transporter with high homology to iron-chelate binding ATP-binding cassette (ABC) transporters. Two ferric iron uptake regulator protein (*fur*) genes were identified as was a coding sequence for the previously discussed bacterioferritin. Sequences encoding genes involved in Fe^{2+} and Fe^{3+}oxidation, detoxification, storage and release regulation [159] were also identified adjacent to the RuBisCO genes. Lastly, as a MTB, strain SS-5 possesses *mamBMEPTH* genes involved in redox control in the synthesis of magnetosome crystals[28].

Nitrogen Metabolism

Strain SS-5 is one of the few magnetotactic bacteria that does not demonstrate nitrogenase activity and thus it is not surprising that *nif* genes are not present in the genome of this organism. The strain does utilize ammonia as a source of cellular nitrogen and has an ammonium transporter gene *amt*, glutamine synthetase/glutamate synthase (bacteria-like GS-GOGAT cycle) and glutamate dehydrogenase for the assimilation of ammonia (Figure 27). *NasDE* genes were identified immediately upstream of the nitrite transporter gene *nirC* (Figure 30). The *nasDE* cluster codes for a nitrite reductase that is involved in both assimilatory and dissimilatory pathways and is involved in the reduction of nitrite to ammonia.[179,180] Assimilatory nitrate reductase genes *nasBC* were not identified. Together, these latter results suggest that strain SS-5 is capable of using nitrite but not nitrate as a source of cellular nitrogen although this remains to be tested in the laboratory.

Denitrification is the sequential dissimilatory reduction of nitrate or nitrite to a gaseous end product (nitrous oxide or dinitrogen) and is a globally important environmental biogeochemical process since it results in the loss of fixed nitrogen from an ecosystem.[101,121,161] The full pathway of denitrification is shown in Figure 29. Strain SS-5 possesses all the genes for the entire pathway of denitrification[181] including those encoding the nitrate reductases NarGHJC, the nitrite reductases NirSK, NasDE and cytochrome cd1, the nitric oxide reductase NorBC, and the nitrous oxide reductase NosZ. (Figure 30). Additionally, the genes encoding the enzymes for the reduction of nitrate to nitrite through a periplasmic respiratory nitrate reductase system,[121] *nap ABCDFGH,* were identified in two gene clusters ~55 Kb apart (Figure

30). However, nitrate transporters have not been identified in SS-5 and despite having these

nitrate reducing genes, attempts to grow strain SS-5 anaerobically in the laboratory with nitrate

as a terminal electron acceptor were unsuccessful.

Figure 29 The Sequential Reductive Pathway of Denitrification in Prokaryotes Present in Strain SS-5.

Blue boxes indicate nitrogen oxides being reduced (except N_2), teal arrows indicate enzymes mediating the reduction, blue arrows indicate genes for those enzymes present in the genome of strain SS-5.

Figure 30 N Metabolism Gene Clusters in the Genome of Strain SS-5
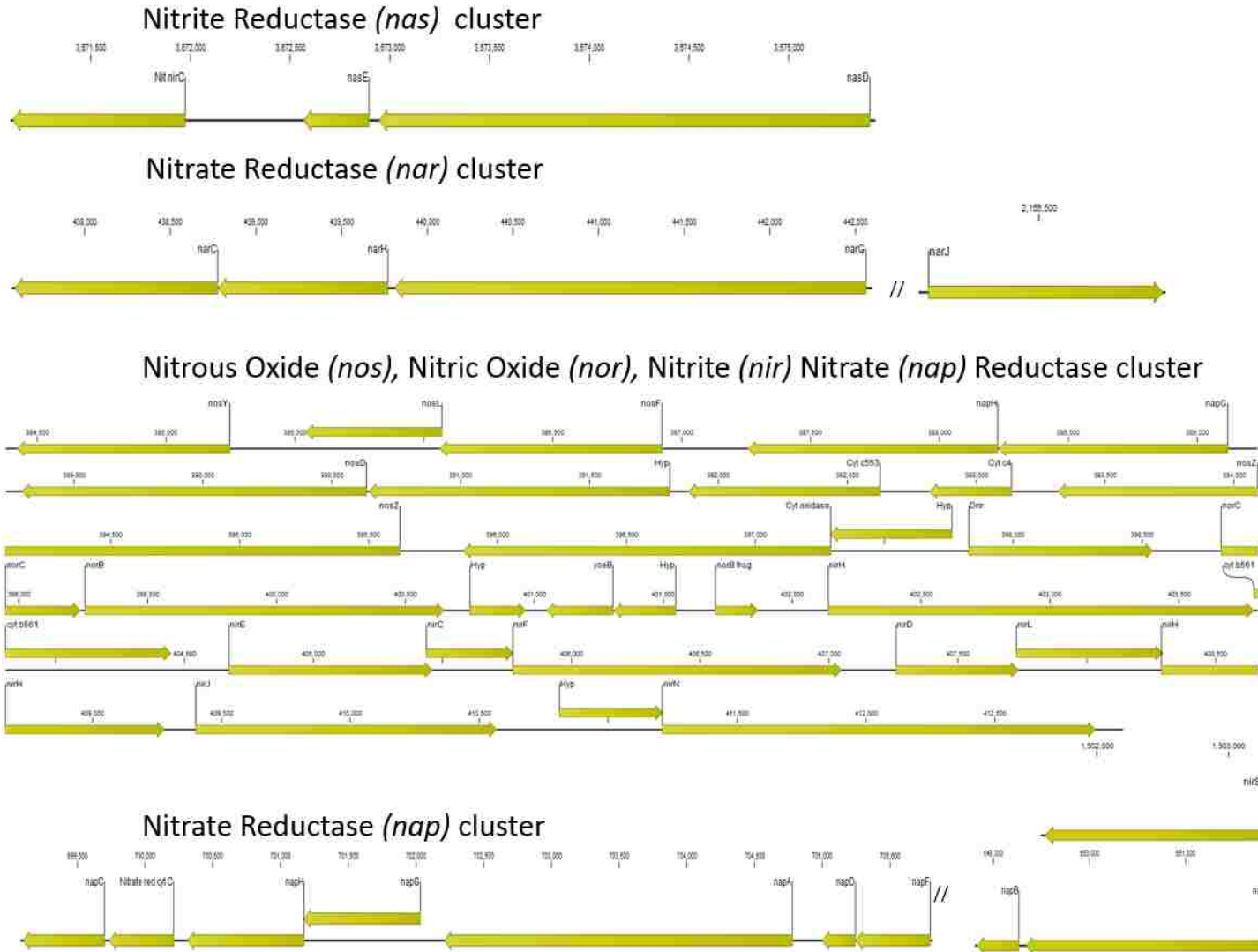
Arrows indicate genes, numbered hash marks indicate genomic position relative to Ori, Nit = nitrate transporter, Hyp = hypothetical protein, Dnr = Nitric oxide responding transcriptional regulator, yoeB = toxin

Sulfur Metabolism

SS-5 has several genes related to oxidation of sulfide and thiosulfate as sources of high energy electrons to drive respiration. This includes the sulfur oxidation enzymes encoded by *soxAX, soxYZ, soxB*, the thioredoxin encoding gene *soxW* and the sulfide dehydrogenase encoding gene *soxF*, all in single copies. SS-5 seems to be missing *soxCD*. Sulfur globules have not been observed in SS-5, indicating the sulfur intermediate often seen in thiosulfate-oxidizing bacteria missing *soxCD*[5,162] is not allowed to accumulate, but is instead immediately further processed. Subsequent oxidation can be achieved by proteins encoded by the reverse-acting (oxidizing) Dsr system that includes proteins closely related to dissimilatory sulfite reductases operating to reduce sulfite. [5,163] Strain SS-5 has the complete set of 16 Dsr encoding genes including the membrane complex *dsrMKJOP* and the reverse acting dissimilatory sulfite reductase *dsrAB*, as well as *dsrCC4EFHLNRS*.

No genes for sulfate reduction, either assimilatory or dissimilatory, were identified, indicating a reliance on exogenous reduced sulfur sources as a source of cellular S directed to anabolic pathways.

Central phosphate metabolism

Phosphate is often a major growth limiting resource for many microbes in natural habitats and thus many have evolved efficient phosphate acquisition systems. Cells of strain SS-5 produce numerous phosphate inclusions[1] and its genome contains numerous phosphorous metabolism genes. 10 of these are phosphate acquisition system genes, including the *pstSCAB* transporter complex, the PhoR and PhoB two component regulatory system that make up the PHO regulon, in addition to a low affinity inorganic phosphate transporter *lat* and a high affinity *pstABC* ABC transporter. Limiting phosphate conditions lead to the autophosphorylation of the sensory kinase PhoR which activates PhoB by phosphorylation (Figure 22). PhoB is a positive transcription regulator of the pstSCAB phosphate regulon that is present in 4 copies in SS-5. In non-phosphate limiting conditions, PhoB may in turn inhibit the kinase activity of PhoR, leading to interaction of PhoR and PhoU to dephosphorylate and inactivate PhoB.[165,166] SS-5 also has multiple copies encoding an ATPase related to phosphate starvation-inducible protein, PhoH, that may be involved in lipid and phospholipid metabolism and RNA modification.[167] Lastly, there is a *nptA* Na$^+$-dependent phosphate transporter gene present in the SS-5 genome. Polymerization of phosphate into the inclusions observed in SS-5 is mediated by a the polyphosphate kinase *ppk2*.[23,32,166,168]

There are 96 membrane transport related coding sequences in the genome of SS-5, in the low range of what is typical of the number expected in an autotrophic prokaryote.[23,169]

Of these, 6 are ABC transport genes. One gene, *phnD*, is present as a double copy of a phosphate binding periplasmic subunit but without the other components, indicating that either the transporter is incomplete or this protein is part of a non-ABC transporter system. Another two coding sequences are homologous to ABC transporter periplasmic oligopeptide binding subunit OppA and permease OppB that are involved in oligopeptide transport, although the other subunits are missing possibly lost to previous genomic rearrangement events. There is also an incomplete ABC dipeptide transporter, with subunits present for periplasmic dipeptide binding and a subunit for ATP binding but no permease subunit.

There are 19 uni, sym or antiporter genes present in the SS-5 genome. Among these is a Na^+/H^+ antiporter, *nhaD* and the aforementioned *nptA* Na^+ dependent phosphate transporter in two copies. Also present is a complete set of a seven subunit (A-G) Na^+/H^+ transporter coding sequences in multiple copies.

38 protein secretion transport related sequences are present in the genome of SS-5. 27 genes are type IV protein encoding genes for pili and fimbriae synthesis, while there are 4 type II protein secretion related proteins, 2 more involved in translocation to the membrane and 3 twin-arginine translocation system genes for export of folded proteins.

There are 8 additional protein encoding genes involved in the transport of cations and 18 more are members of the previously mentioned Ton/Tol transport system.

Additionally, there are 7 Tripartite ATP-independent periplasmic transporter (TRAP) genes, 2 encoding proteins related to C4-dicarboxylate transport. 5 are involved in the transport of as yet unknown substrates, although it is known that they are likely organic acids.[182]

Chemotaxis and signal transduction

There are multiple protein coding sequences for the cyclic second messenger cAMP, genes coding for the cAMP receptor protein FNR, 7 copies of cAMP binding proteins CGA and 7 copies of an adenylate cyclase present in the genome of strain SS-5.

As an aerotactic microbe, strain SS-5 contains genes coding for the O_2-sensitive aerotaxis receptor protein Aer. Genes encoding 21 methyl accepting chemotaxis proteins are also present in multiple copies, including those for Tsr and the methyl transferase CheR. The genes encoding the CheW/CheA/CheY system for signal transduction to the flagellum including CheR and CheB are present in several copies each, ranging from 6 to 9. As previously discussed, the Che system transduces signal to flagellar motor switch proteins FliG, M, N, which are also identified. [60] 53 genes were identified as encoding proteins with redox sensor PAS/PAC domains along with GGDEF and EAL domains which are involved in the synthesis and phosphorylation of cyclic second messengers like cAMP and cyclic di-GMP.[60] These proteins react to their redox environment as assessed by their PAS/PAC domains. In this case by synthesis (GGDEF domain) and phosphorylation (EAL domain) of the second messenger cyclic di-GMP.[170]

As previously discussed, histidine kinases are proteins that are typically transmembrane enzymes that mediate signal transduction across cellular membranes.[60] In addition to the previously discussed CheA, there are 26 genes encoding histidine kinases in the genome of SS-5. 2 of these are among the 14 two component response regulators and sensory proteins[60,166,171] in SS-5.   There are also 22 single component, non-histidine kinase response regulators in SS-5.

Prophages

While there are several phage related genes scattered throughout the genome including 45 transposase genes and 24 phage related genes, there does not appear to be a cluster representing a complete prophage.

Magnetosome-Related genes

There are four magnetosome-associated gene clusters in the genome of strain SS-5. The *mamAB* cluster is 29,731 bp, the *mamCOPD* cluster is 5,072 bp, a putative magnetosome associated hypothetical magnetosome protein *(hmp)* gene region is 12,255 bp and a single *feoB*-like gene is in a region of the chromosome not close to any other magnetosome genes (Figure 32, Table 8). Although the majority of the genes putatively involved in magnetosome formation are clustered with the *mamAB* cluster of the SS-5 genome, the smaller *mamCOPD* cluster and three *hmp* genes are located at a different region of its chromosome (Figure 32,

Table 8). These *hmp* genes contain an adenylate cyclase domain and are found only in several

MTB, although their role is yet known.

There is no evidence that the magnetosome gene clusters in strain SS-5 represent a

genomic island as there are no tRNA or mobile element genes such as resolvases or

transposases in or near the magnetosome gene clusters, thus it seems unlikely that the

magnetosome genes here originated from lateral gene transfer from another organism.

Additionally, there is no difference in surrounding region GC content, and the phylogeny and

synteny of the *mam* genes in both strains SS-5 and BW-2 are highly homologous. Taken

together, this seems to again indicate that, these genes were not acquired through the result of

a recent horizontal gene transfer event, but rather through descent via a last common ancestor

of strains SS-5 and BW-2. Further evidence of this is the fact that the majority of magnetosome

proteins in BW-2 and SS-5 are most closely related to another unpublished magnetic

gammaproteobacterium, strain CG-1 of the order *Chromatiales*. The next most related

magnetosome associated gene sequences are the corresponding proteins found in the genome

of *M. magneticum*, of the *Alphaproteobacteria* (Table 8).

Figure 31Magnetosome Gene Clusters in the genome of strain SS-5
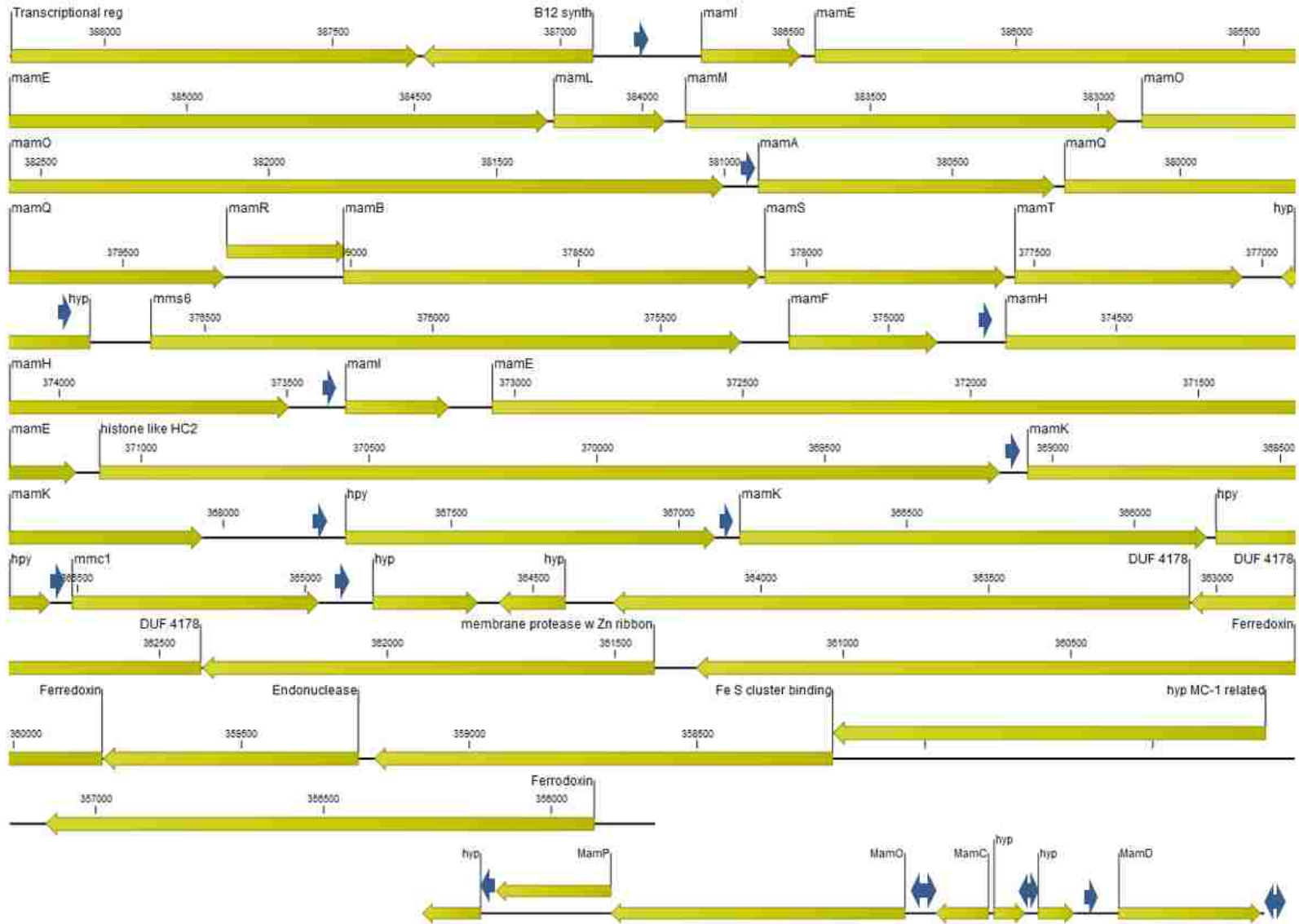
Green arrows indicate coding sequences, blue arrows indicate promoter sequences, numbers at hash marks indicate base location in genome relative to putative origin of replication.

Table 8 Genes with Homology to the Magnetosome Cluster Genes of Interest of Strain SS-5

Mam protein	Magnetotactic bacterium with protein with highest sequence identity*	Identity (%)	e-value	Magnetotactic bacterium with protein with second highest sequence identity*	Identity (%)	e-value	Magnetotactic bacterium with protein with third highest sequence identity*	Identity (%)	e-value
MamI	Chromatiales CG1	79.17	3.00E-37	Thiotrichales BW2	57.35	1.00E-21	Mv. blakemorei	51.52	2.79E-17
MamE	Chromatiales CG1	60.69	0	Thiotrichales BW2	46.84	0	Mv. blakemorei	41.25	5.05E-90
MamL	Chromatiales CG1	91.36	1.00E-48	Thiotrichales BW2	61.25	1.00E-27	Mv. blakemorei	41.43	6.75207E-15
MamM	Chromatiales CG1	92.7	0	Thiotrichales BW2	75.17	7.00E-163	M. magneticum	55.14	8.00E-116
MamO	Chromatiales CG1	80.57	0	Thiotrichales BW2	61.15	2.00E-125	M. magneticum	39.54	2.00E-147
MamA	Chromatiales CG1	69.9	5.00E-96	Thiotrichales BW2	40.78	1.00E-42	M. magneticum (mms24)	38.35	2.00E-46
MamQ	Chromatiales CG1	78.7	5.00E-180	Thiotrichales BW2	57.41	3.00E-115	Mv. blakemorei	44.87	1.15E-60
MamR	Chromatiales CG1	84.09	2.00E-49	Thiotrichales BW2	41.18	3.00E-21	Mv. blakemorei	42.86	2.28E-14
MamB	Chromatiales CG1	93.42	0	Thiotrichales BW2	83.22	2.00E-179	M. magneticum	53.82	5.00E-113
MamS	Chromatiales CG1	65.43	5.00E-85	Thiotrichales BW2	45.79	2.00E-46	M. magneticum	36.94	2.00E-24
MamT	Chromatiales CG1	72.29	7.00E-95	Thiotrichales BW2	54.72	3.00E-62	M. magneticum	51.76	1.00E-52
Mms6	Chromatiales CG1	66.6	0	Thiotrichales BW2	50.35	1.00E-103	M. magneticum	43.11	5.00E-26
MamF	Chromatiales CG1	95.37	2.00E-71	Thiotrichales BW2	68.52	2.00E-51	M. magneticum	54.29	4.00E-37
MamH	Chromatiales CG1	86.65	0	Thiotrichales BW2	70.12	0	M. magneticum	57.11	6.00E-169
MamI	Chromatiales CG1	72.46	3.00E-30	Thiotrichales BW2	57.35	8.00E-21	Mv. blakemorei	52.31	5.55E-18
MamE	Chromatiales CG1	55.1	0	Thiotrichales BW2	35.69	1.00E-101	M. marisnigri SP-1	35.61	8.00E-106
hyp prot	Chromatiales CG1	36.88	3.00E-91	Thiotrichales BW2	22.1	4.00E-14	Rhodococcus erythropolis CCM2595	28.4	4.00E-16
MamK	Chromatiales CG1	92.6	0.00E+00	Thiotrichales BW2	60.06	5.00E-149	M. magneticum	50.3	3.00E-117
hyp prot	Chromatiales CG1	35.5	1.00E-32	Streptomyces pratensis ATCC 33331	28.28	7.00E-05	Streptomyces sp. AA0539	31.06	2.00E-07
MamK	Chromatiales CG1	82.4	0	Thiotrichales BW2	60.29	2.00E-146	M. magneticum	52.68	1.00E-119
hyp prot	Chromatiales CG1	65.06	2.00E-33	Frankia sp. CN3	35.71	1.50E-02	None		
*Mmc1-like	Chromatiales CG1	76.27	3E-95	Magnetospira QH-2	3e-53	48	Mc. marinus	33	9E-17
MamP	Chromatiales CG1	81.82	2.00E-148	Thiotrichales BW2	73.13	2.00E-111	M. magnetotacticum	53.97	2.00E-64
MamO	Chromatiales CG1	75.19	0	Thiotrichales BW2	52.4	0	M. magnetotacticum	38.41	3.00E-130
MamC	Chromatiales CG1	80	9.00E-62	Thiotrichales BW2	63.48	2.00E-46	M. marisnigri SP-1	54.37	2.00E-33
MamD	Chromatiales CG1	73.54	3.00E-159	Magnetospira sp. QH-2 hyp. Prot.	49	2.00E-86	M. magneticum hyp. prot.	41.14	7.00E-57
	Chromatiales CG1								
membrane protease	Chromatiales CG1	99.56	3.00E-161	Bacillus subtilis subsp. natto	99.56	3.00E-161			
Hyp	Chromatiales CG1	53.07	3.00E-49	Chromatiales CG1	92.16	2.00E-25			
MamD	Chromatiales CG1	88.32	0	M. magnetotacticum	38.9	3.00E-58	Magnetospirillum sp. SO-1	39.48	6.20E-50
Hyp	Chromatiales CG1	74.12	3.00E-159						
Hmp4	Chromatiales CG1	64.1	5.00E-55	Magnetospirillum sp. SO-1	35.14	4.49501E-09	M. magneticum	37.84	1.00E-10
Hmp3	Chromatiales CG1	86.51	0	M. marisnigri	30.12	2.00E-74	M. magneticum	29.7	1.00E-69
Hmp2	Chromatiales CG1	94.44	0	M. marisnigri	36.56	1.00E-119	Geobacter metallireducens GS-15	46.46	1.00E-21

As previously discussed, the *feoAB*-like genes encode membrane proteins that are involved in iron transport and appear to be specific to MTB.[89] In most MTB there are at least two clusters *feoAB*, one specific to the MTB (*feoAB-like* in the magnetotactic *Alphaproteobacteria* and *feoBA-like* in the magnetotactic *Deltaproteobacteria*) and another *feoAB* cluster that has the same origin as in other bacteria.[33,89] However, in the genome SS-5 there is just one *feoB*. Interestingly, in the genome of SS-5 there is no *feoA* homolog.

An ~5kb *mamCOPD* cluster is present in the genome of strain SS-5 that contains magnetosome genes *mamPOCD* in that order, with an coding sequence between *mamC* and *mamD* that encodes a protein of unknown function that does not share homology with any protein from other bacteria. The gene *mamP* encodes an archetypal MamP protein that comprises one PDZ , a conserved domain enabling protein-protein interaction, and two magnetochrome domains.[66] The gene *mamO* encodes a protein that contains one N-terminus trypsin-like protease domain and one C-terminus TauE domain. The *mamC* gene encodes MamC, a MTB specific protein that does not contain domains conserved with other bacteria. *mamD* encodes for MamD, an MTB specific protein that does not contain known conserved domains.

There is another copy of *mamD* far from the magnetosome gene clusters, named *mamD-2*, which likely originated from a duplication of *mamD-1* or vice versa. In fact, phylogenetically, the proteins *mamD-1* and *mamD-2* are more closely related to each other than they are to *mamD* from other MTB. Just upstream of this *mamD-2*, there are three uncharacterized genes designated *hmp.* These are highly homologous to genes found in

93

Magnetospirillum. Interestingly, two of these sequences that have homologs in

Magnetospirillum are collinear and found in the vicinity of the magnetosome island.

The ~21.8kb SS-5 cluster contains the *mamAB* cluster which includes

mamIELMOAQRBSTFHIEKK and an *mms*6 gene. There are two copies each of *mamI, mamE* and

mamK (Figure 32).   The putative cytoskeletal attachment protein *mamJ*[84] appears to be absent.

As previously discussed, this cluster is analogous to the *mamAB* operon described in

Magnetospirillum species, and so is referred to it as *mamAB,* although, again, there is no

evidence whether the genes in this putative operon are also transcribed under the control of a

single promoter. In fact, there are 10 identifiable potential promoters within the *mam*AB

cluster of SS-5. (Figure 32)

Summary

Taken together, evidence from previous experiments and the genetic systems identified

in the genome of SS-5 here indicate that it has the potential as an autotroph to fix carbon and

ammonia from the environment, however, unlike most MTB appears to be unable to fix

nitrogen. The ability to respire by oxidizing reduced sulfur compounds such as sulfide and

thiosulfate and reducing nitrate and nitric oxide was also indicated. 96 established transporters

of various types were identified.   Few prophage genes were identified and none clustered as an

identifiable prophage. A compliment of 24 previously characterized magnetosome associated

genes found in other MTB were identified in strain SS-5 in two clusters separated in the

genome but without the hallmark features of a genomic island such as the presence of

transposases, flanking tRNAs and a GC content differing from that of the rest of the genome.

The metabolic and magnetosome related genomic features of BW-2 will be compared with

those of BW-2 and of other MTB in chapter five in the interest of uncovering common

metabolic features, as of yet uncharacterized genes that may be involved in the magnetic

phenotype, as well as any further evidence of the nature of the acquisition by strain SS-5 of

these magnetosome related genes.

Chapter 5 Comparative Analysis

Phylogeny

Based on 16S rRNA gene sequence homologies of strains BW-2 and SS-5, these prokaryotes belong clearly to the *Gammaproteobacteria*. Originally strain BW-2 was determined to be in the order *Thiotricales* while SS-5 was placed among members of the *Chromatiales*.[1] Upon subsequent phylogenetic analysis in this work, however, the phylogenetic position of strain BW-2 has been resolved and revised which shows it belongs to the *Chromatiales* rather than the *Thiotricales* (Figure 33).

Figure 32 Phylogenetic tree showing phylogenetic positions of strains BW-2 and SS-5BW-2 and SS-5
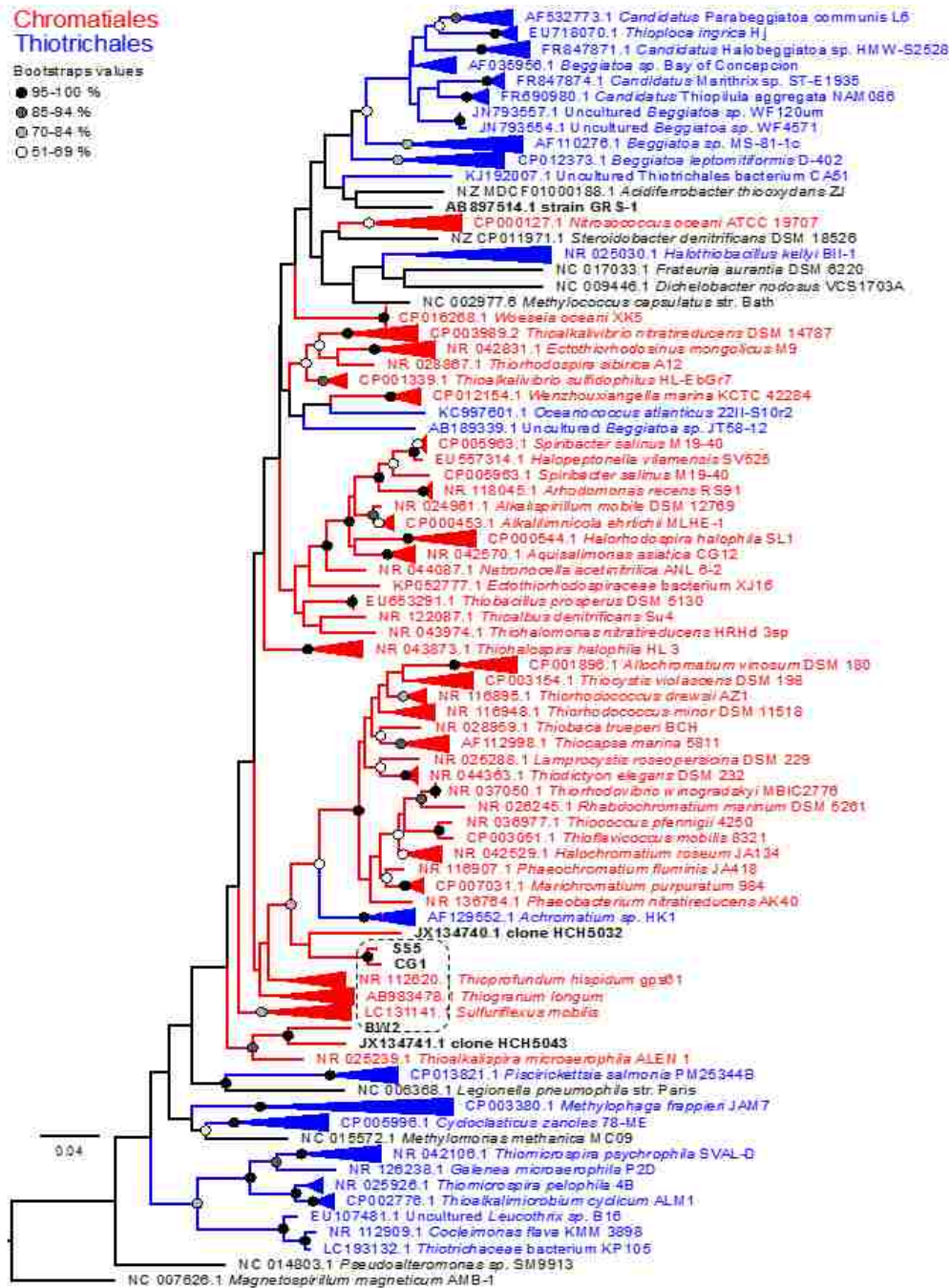
Unpublished tree from collaborator on this work, Caroline Monteil. Phylogeny of the *Gammaproteobacteria* class based on 16S

rRNA gene sequences. Analysis of the nucleotide sequences of 281 strains largely belonging to the *Thiotrichales* and *Chromatiales*,

aligned with MUSCLE [183], tree constructed with MEGA 7 [184] MTB names are in bold, 100 bootstraps were conducted to get

proportional support values. Branch lengths represent the number of nucleotide substitutions per site. *Chromatiales* and

Thiotrichales are colored in red and blue respectively according to the NCBI taxonomy database.

When compared to other MTB, strain BW-2 is most closely phylogenetically related to strain SS-5 with a 16s rRNA gene sequence identity of 87.58% (Figure 16). The next most related magnetotactic strain to BW-2 is the newly isolated and as yet unfinished genome of the gammaproteobacterial, *Chromatiales* strain CG-1. Strains BW-2 and CG-1 share an 85.12% 16s rRNA gene sequence identity. Strain SS-5 is most related to strain CG-1 with 94.93% 16s rRNA gene sequence identity then secondly with BW-2 at 87.58% 16s rRNA gene sequence identity. The next most related MTB to these magnetic *Gammaproteobacteria* is *Magnetococcus marinus* strain MC-1 of the *Alphaproteobacteria.*

Metabolic comparisons

BW-2 and SS-5 share a similar metabolic lifestyle as chemolithoautotrophs, using sulfide and thiosulfate as high energy electron sources, O_2 or possibly certain nitrogen oxides as electron acceptors, and sodium bicarbonate or CO_2 as primary carbon sources.[1]

Autotrophy

Both strains have intact genetic systems coding for RuBisCO-mediated CO_2 fixation in two sets. Both are found in the same region in strain BW-2 within 6,382 bases, (Figure 17) while in strain SS-5 one set is in two fragments over a megabase apart and interspersed and flanked with 23S rRNA, tRNA, mobile element and phage protein coding sequences (Figure 28). In each strain, one RuBisCO cluster is incorporated into a carboxysome gene operon while one

is not. Complete carboxysome gene sets are also present in both strains, although

carboxysomes or caboxysome-like structures yet to be observed in either strain under currently

used culture conditions. While these operons are present in both strains, the synteny of the

genes is different possibly due to a rearrangement of genes during their evolutionary

divergence.

Many MTB of the *Alphaproteobacteria* also grow chemolithoautotrophically. Genes

coding for RubisCO type II have also been identified in *M. gryphiswaldense, M.*

magnetotacticum and Magnetovibrio blakemorei, whereas *Magnetococcus marinus* and strain

MO-1 grow chemolithoautotrophically using the reverse tricarboxylic acid cycle. To date,

however, carboxysome operons have not been identified in other MTB, and MTB species

discovered so far belonging to the *Deltaproteobacteria* are chemoorganoheterotrophs.[6]

Central Carbon Metabolism

Intact TCA cycle and glycolysis gene sets are identifiable in both strains BW-2 and SS-5.

However, similarly to *Magnetococcus marinus*, both are missing any known fructose-1,6-

bisphosphatase isozyme,[156] an enzyme necessary to the gluconeogenesis pathway of

carbohydrate synthesis. *Magnetococcus marinus* has two inositol monophosphatase genes,

one of which was determined to be similar to an archaeal fructose-1,6-bisphosphatase gene,

possibly performing this function.[56] BW-2 and SS-5 also have two inositol monophosphatase

genes, but neither of them are more similar to archaeal fructose-1, 6-bisphosphatase than to

inositol monophosphatase of other bacteria.   One explanation for this absence of fructose-1, 6-

bisphosphatase could be that this reaction is catalyzed by unknown nonorthologous enzyme acquired in a common ancestor. In support of this possibility the largest group of confirmed and predicted non-homologous isofunctional enzymes is within carbohydrate metabolism.[156] There are also no identifiable sugar transporters in either strain, consistent with the lack of success in growing these strains heterotrophically. As previously discussed however, the acetate transporter acetate permease (*actP*) is present in both BW-2 and SS-5 along with all necessary enzymes, including acetyl-CoA synthetase, to convert acetate to acetyl-CoA making it available to the TCA cycle.

Nitrogen Metabolism

Both gammaproteobacterial strains can also acquire nitrogen in the form of ammonia from the environment for assimilation into cellular components via an ammonium transporter and the GS-GOGAT cycle to incorporate it into amino acids. In addition *NasDE* genes were identified in strain SS-5 immediately upstream of the nitrite transporter gene *nirC* (Figure 30). The *nasDE* cluster codes for a nitrite reductase that is involved in both assimilatory and dissimilatory pathways and is involved in the reduction of nitrite to ammonia.[179,180] Assimilatory nitrate reductase genes *nasBC* were not identified. Together this suggests that strain SS-5 is capable of using nitrite but not nitrate as a source of cellular nitrogen although this remains to be tested in the laboratory. BW-2 does not have an assimilatory nitrate reductase but does have a complete nitrogenase system. By contrast, while strain BW-2 does not have an assimilatory nitrate reductase and therefore does not appear to be capable of

using nitrate as a sole source of nitrogen for growth, it does have a complete nitrogenase

regulon enabling it to also fix nitrogen from the environment (Figure 19). Both BW-2 and SS-5

appear capable of reducing some nitrogen oxides in a dissimilatory fashion as terminal electron

acceptors in lieu of O_2 (Figure 34), however only strain BW-2 appears capable of taking up

nitrate via two copies of nitrate/nitrite transporters. No nitrate transporters were identified in

strain SS-5. Strain BW-2 has the genes for the enzymes dissimilatory nitrate reductase and

nitric oxide reductase, suggesting that this strain has the potential to use nitrate and nitric

oxide as terminal electron acceptors in respiration. Strain SS-possesses the complete suite of

genes encoding the enzymes for the complete pathway of denitrification and thus has the

potential to reduce nitrate to dinitrogen gas.

Figure 33 Nitrogen oxide reduction in strains BW-2 and SS-5
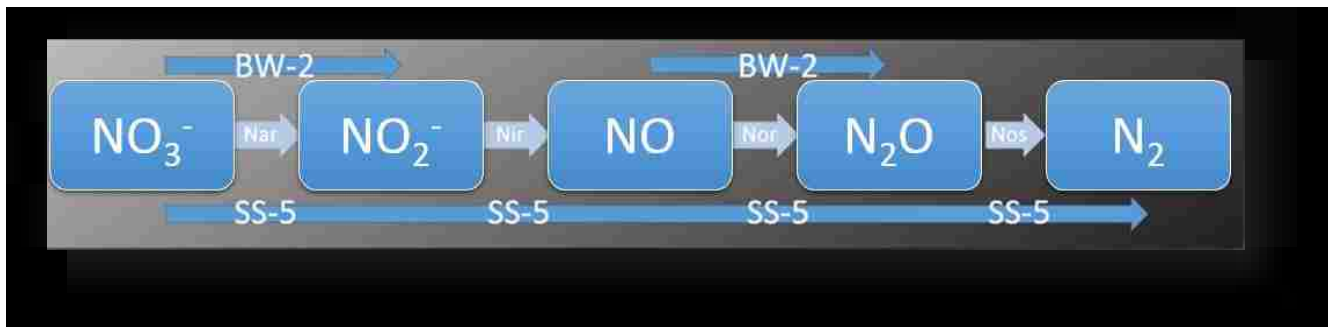
Blue boxes indicate nitrogen species being reduced, teal arrows indicate enzymes mediating the reduction, blue arrows indicate

genes present in respective strains.

Nitrogenase activity is typically confirmed in MTB including most MTB of the

Alphaproteobacteria including *M. magneticum, M. gryphiswaldense, Magnetococcus marinus,*

Magnetovibrio Blakemorei, Magnetospira thiophila and most affiliated with the

Deltaproteobacteria.[12] In fact, dissimilatory nitrate reduction has been functionally linked to

nitrogenase activity[185] and to magnetite biomineralization in *M. gryphiswaldense,* and related

enzymes participate in redox reactions required for magnetite biomineralization.[121] However,

as of yet unidentified enzymes likely perform these functions in *Magnetospira* sp. QH-2 as

genes for nitrogenase are absent from its genome.[174] Similarly to BW-2 and SS-5,

Magnetococcus *marinus* additionally has the GS-GOGAT cycle and glutamate dehydrogenase

for the assimilation of ammonia.

Sulfur Metabolism

Both strains BW-2 and SS-5 are sulfide and thiosulfate-oxidizing bacteria. As such, they

have several genes related to sulfur metabolism scattered throughout its genome including the

sulfur oxidation (SOX) system and the reverse acting (oxidative [5,163]) dissimilatory sulfite

reductase (Dsr) systems, as previously discussed. While both strains are missing *soxCD*, as is

typical of other *Gammaproteobacteria* that oxidize sulfur using the SOX system,[5] only BW-2 is

observed to accumulate sulfur inclusions when cultured with thiosulfate as is observed in many

other bacteria that perform SOX reactions without *soxCD*.[5] Sulfur globules have not been

observed in SS-5, indicating the sulfur intermediate often resulting in SOX reactions without

soxCD is not allowed to accumulate, but is instead immediately further processed by the Dsr

system into sulfite, then further oxidized to sulfate by the SOX system. [5,163]

Genes for dissimilatory sulfate reduction were not identified in either strain. Assimilatory sulfate reduction genes were also not identified indicating a reliance on exogenous reduced sulfur sources as a source of cellular S directed to anabolic pathways.   YedYZ , a putative membrane associated reductase (often annotated as sulfite oxidase) is present in BW-2 and could be active in the reduction of sulfur species as, although its functions are largely unknown it has been shown to act upon sulfur containing substrates.[164] The genome of *Magnetococcus marinus* also encodes YedYZ, however its sulfate assimilation pathway is incomplete with only the first two enzymes present, likely also acquiring reduced sulfur exogenously.[56] By contrast, the genomes of M. magneticum and M. magnetotacticum contain all genes necessary for assimilatory sulfate reduction[56] Although the MTB of the Deltaproteobacteria do reduce sulfate in a dissimilatory fashion, they are chemoorganoheterotrophs.

Sulfur compound oxidation is also typical of other MTB. The two strains of magnetococci in culture, *Magnetococcus marinus*[56] and strain MO-1,[12] grow chemolithoautotrophically using reduced sulfur compounds as electron donors. *Magnetococcus marinus*, similarly to BW-2 and SS-5, is also missing *soxCD*.[56] As observed in BW-2, it also produces sulfur globules that accumulate intracellularly that may be further converted to sulfite via its Dsr system. *M. gryphiswaldense* is also capable of autotrophic and mixotrophic growth utilizing reduced sulfur compounds and *Magnetovibrio blakemorei*[111] grows chemolithoautotrophically with reduced sulfur compounds. *Magnetospira thiophila* (strains MMS-1 and QH-2) grow chemolithoautotrophically using thiosulfate.

Phosphate Metabolism

Both SS-5 and BW-2 produce numerous phosphate inclusions and have numerous

phosphorous metabolism genes. In the interest of acquiring phosphate from the environment,

BW-2 has 13 phosphate transport system genes while SS-5 has 10. Although each has the same

genes present, the SS-5 genome additionally encodes a low affinity inorganic phosphate

transporter. Both BW-2 and SS-5 have the PstSCAB transporter complex as well as the PhoR and

PhoB two component regulatory system. Limiting phosphate conditions lead to the

autophosphorylation of the sensory kinase PhoR which phosphorylates the PhoB positive

transcription regulator of the phosphate regulon that is present in multiple copies.   In non-

phosphate limiting conditions, PhoB can in turn inhibit the kinase activity of PhoR, leading to

interaction of PhoR and PhoU to dephosphorylate and inactivate PhoB.[165,166] Both SS-5 and

BW-2 also have multiple copies encoding an ATPase related to phosphate starvation-inducible

protein, PhoH, that may be involved in lipid and phospholipid metabolism and RNA

modification.[167] Lastly, both genomes encode an nptA Na^+-dependent phosphate transporter.

Unlike SS-5, BW-2 produces phosphate inclusions. Predictably, coding sequences for

polyphosphate kinase and exopolyphosphatase, the enzymes involved in the production of

phosphate granules, were found in two copies each in BW-2.

Polyphosphate inclusions are commonly found in environmental samples of MBT.[12]

These inclusions are also common in cultured MTB, including all three *Magnetospirillum* strains

and *Magnetococcus marinus* which also contain genes encoding polyphosphate kinase and

exopolyphosphatase. [186]

Transport Genes

There are 128 membrane transport related coding sequences in the genome of BW-2, typical of the number expected in an autotrophic prokaryote,[23,169] and 96 membrane transport related coding sequences in the genome of SS-5, a somewhat low number but still in the low range typical of the number expected in an autotrophic prokaryote.[23,169] In BW-2, 11 of these are ABC transport genes while in SS-5, 6 are ABC transport genes. In both BW-2 and SS-5 the double copy of the *phnD* phosphate binding periplasmic subunit is present without the other components, indicating that either the transporter is incomplete or this protein is part of a non-ABC transporter system. Another two coding sequences homologous to the ABC transporter involved in oligopeptide transport that was found to be partially present in SS-5 are present in BW-2. In both cases, there are permease subunits *oppB* and *oppC*, and no subunit *oppA*, indicating loss of the other components due to genomic rearrangement in the last common ancestor of the two strains. There is also an incomplete ABC dipeptide transporter in SS-5 that was not identified in BW-2, with subunits present for periplasmic dipeptide binding and a subunit for ATP binding but no permease subunit. In BW-2, a complete set of genes representing a branched chain amino acid ABC transport system, *livFGMHJ* and the high-affinity leucine-specific periplasmic binding protein LivK, were found that were not found in SS-5. There are also 16 uni, sym or antiporters in BW-2 and 19 uni, sym or antiporter genes present in the SS-5 genome. As in SS-5 there is a Na$^+$/H$^+$ antiporter, *nhaD*, but in three copies instead of the one in SS-5, as well as the *nptA* Na$^+$ dependent phosphate transporter in one copy but with three found in SS-5. Additionally, in both BW-2 and SS-5 there is a complete set of a seven subunit (*A-G*) Na$^+$/H$^+$ transporter, with five of these in duplicate. There are 74 protein secretion

transport related sequences are present in BW-2 and 38 protein secretion transport related

sequences are present in the genome of SS-5. In BW-2, this includes 31 type IV protein

encoding genes for pilin and fimbriae synthesis and 27 in SS-5. However, unlike SS-5 that

encodes no type I proteins, 31 type I proteins for autoaggregation and biofilm formation are

present in BW-2 in multiple copies. In BW-2 there are 20 copies of T1SS secreted agglutinin

alone in addition to multiple copies of *lapBCE*. These systems are clearly active when growing

BW-2 in sulfide containing medium with a headspace of air, as cells form floating aggregations

and adhere to the side of the jar at and near the surface of the liquid phase of the medium

(Figure 35). In both strains there are also 4 type II protein secretion genes, 2 more that are

involved in translocation to the membrane and twin-arginine translocation system genes for

export of folded proteins, 4 in BW-2 and 3 in SS-5. BW-2 has 9 additional protein encoding

genes involved in the transport of cations while SS-5 has 8. Both strains have 18 genes that

encode members of the previously discussed Ton/Tol transport system. Present in SS-5 but not

found in BW-2, are 7 Tripartite ATP-independent periplasmic transporter (TRAP) genes, 2

encoding proteins related to C4-dicarboxylate transport. 5 are involved in the transport of as

yet unknown substrates, although it is known that they are likely organic acids.[182]

The 12and 96 membrane transport related coding sequences found in BW-2 and SS-5

respectively represent a somewhat low number, but still in the typical range expected in an

autotrophic prokaryote.[23,169] Heterotrophic bacteria often have a higher number of

transporters as they take up multiple sources of carbon such as sugars, organic acids and amino

acids.[169] By comparison, there are 259 in the genomes of *M. magneticum*, 694 in *M.

magnetotacticum* and *Magnetococcus marinus* contains 115.[56] With respect to iron uptake,

MTB have a high demand, engaging in the synthesis of up to 16.7 mg magnetite liter^{-1} day^{-1}.[157]

As previously discussed, MTB have in common the *feoA-like*, *feoB-like*[89] and *mamN*[83] genes involved in iron transport, in addition to the *feoAB* genes that are more widespread among prokaryotes.[89] The MTB specific genes *mamB* and *mamM* have homology to cation diffusion facilitator family genes[56] but as they are magnetosome-associated proteins, they are involved in transport of cytoplasmic iron into magnetosome vesicles rather than into the cell.

Figure 34 Growth of strain BW-2 in Sulfide Containing Medium

Strain BW-2 growth in 825 mL sulfide containing medium with a headspace of 175 mL air. White accumulations on and just below the surface are BW-2 cell aggregates.

Chemotaxis and signal transduction

There are multiple protein coding sequences for the versatile cyclic second messenger cAMP signal in BW-2 and SS-5. Both have genes coding for the cAMP receptor protein FNR, multiple copies each of cAMP binding proteins CGA and 14 and 7 copies for BW-2 and SS-5 respectively of an adenylate cyclase. As aerotactic microbes, both SS-5 and BW-2 have the gene coding for the aerotaxis receptor protein Aer. Both genomes also contain many copies of genes coding for methyl accepting chemotaxis proteins. 21 are identified in SS-5 and 35 in BW-2, including genes encoding Tsr and the methyl transferase CheR. The genome of strain BW-2 also contains the gene coding for the CheC inhibitor of methyl accepting chemotaxis **proteins**

as well. Both strains have many copies of genes encoding the CheW/CheA/CheY system for signal transduction to the flagellum including CheR and CheB, with BW-2 additionally encoding the chemotaxis response phosphatase CheZ. The Che system transduces signal to flagellar motor switch proteins FliG, M, N, which are also identified in both strains. [60] There were also genes identified as encoding proteins with redox sensor PAS/PAC domains along with GGDEF and EAL domains which are involved in the synthesis and phosphorylation of cyclic second messengers. (eg. cAMP) 53 of these were identified in SS-5 and 61 were identified in BW-2 in addition to 2 with the PAS/PAC domain only. Histidine kinases, proteins that are typically transmembrane enzymes that mediate signal transduction across cellular membranes. In addition to the previously discussed CheA, there are also 45 genes encoding histidine kinases in BW-2 and 26 in SS-5. Some of these are part of two component response regulator systems. 2 of these are among the 14 two component response regulators and sensory proteins in SS-5. BW-2 only has 3 two component systems, one response regulator and two sigma 54 specific transcriptional regulators. There are 22 single component, non-histidine kinase response regulators in SS-5 and 30 in BW-2.

As previously discussed, MTB rely upon magnetic orientation coupled with aerotaxis to find the optimal redox conditions or O_2 concentration in a water or sediment column. High numbers of chemotaxis genes would therefore be expected. In total BW-2 has 57 and SS-5 has 53. By comparison *M. magneticum* contains 162 chemotaxis coding sequences, *M. magnetotacticum contains* 326 and *Magnetococcus marinus contains* 65.[56] Among all motile prokaryotes, the average number of chemotaxis genes per Mb has been roughly estimated to

be 12.2.[56] As expected, BW-2 and SS-5 contain slightly higher compliment of genes with 14.25

and 14.24 per Mb respectively.

Prophages

Despite the presence of 45 transposase genes and 24 phage-related genes, there are no

predicted prophages in the genome of strain SS-5. The genome of starin BW-2 contains two

regions containing clusters of phage-related genes that appear to represent incomplete

prophages (Figure 23). This is a low number when compared to the genomes of *M.*

magneticum which contains 12 putative prophages, *M. gryphiswaldense*, which contains 6 and

a plasmid containing 1, and *Magnetococcus marinus* which contains 14.[56] Therefore it appears

unlikely that when undergoing stress, phage genes in these strains could mediate horizontal

gene transfer events by becoming activated and encapsulating genes to pass them on to new

microbes. Nor does it appear that such an event has recently introduced genes into the

genomes of strains BW-2 or SS-5.

Magnetosome genetics comparisons

Gene compliment and synteny

Genes that are involved in the same process or pathway tend to be found clustered in

the chromosomes of prokaryotes, which is useful in consolidating support for the inference of

the function of putative genes found among them.[187] In the genome of strain BW-2, the

magnetosome genes are arranged as a single contiguous cluster (Figure 36). By contrast,

although the majority of the genes putatively involved in magnetosome formation are clustered

with the *mamAB* cluster of the SS-5 genome, there is a smaller *mamCOPD* cluster and three

putative *hmp* magnetosome-associated genes located at a different region of its chromosome

(Figure 36).

Figure 35 Comparison of synteny of Mam gene clusters in strains BW-2 and SS-5
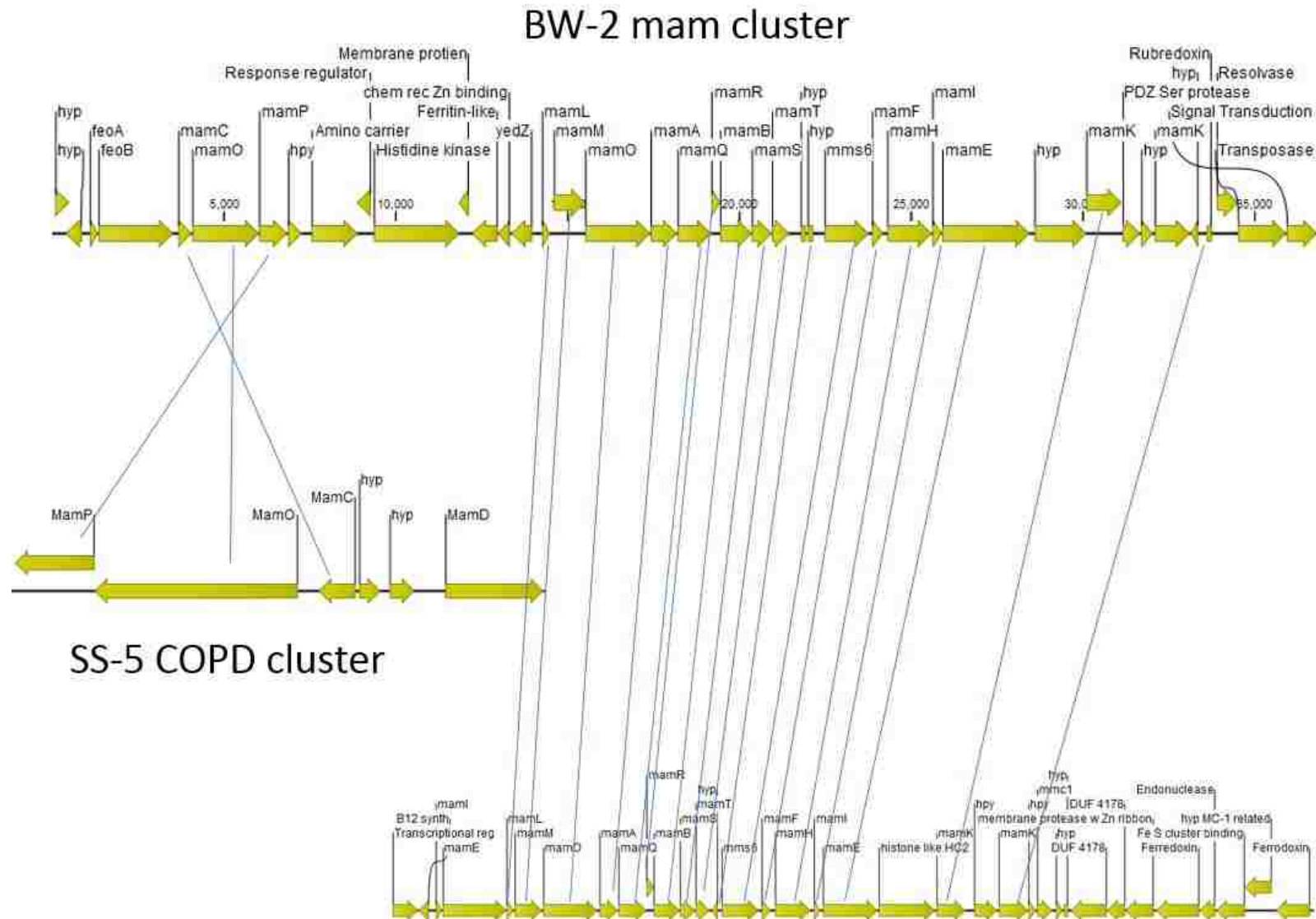
Figure 36 Magnetosome gene Cluster Comparison Across MTB
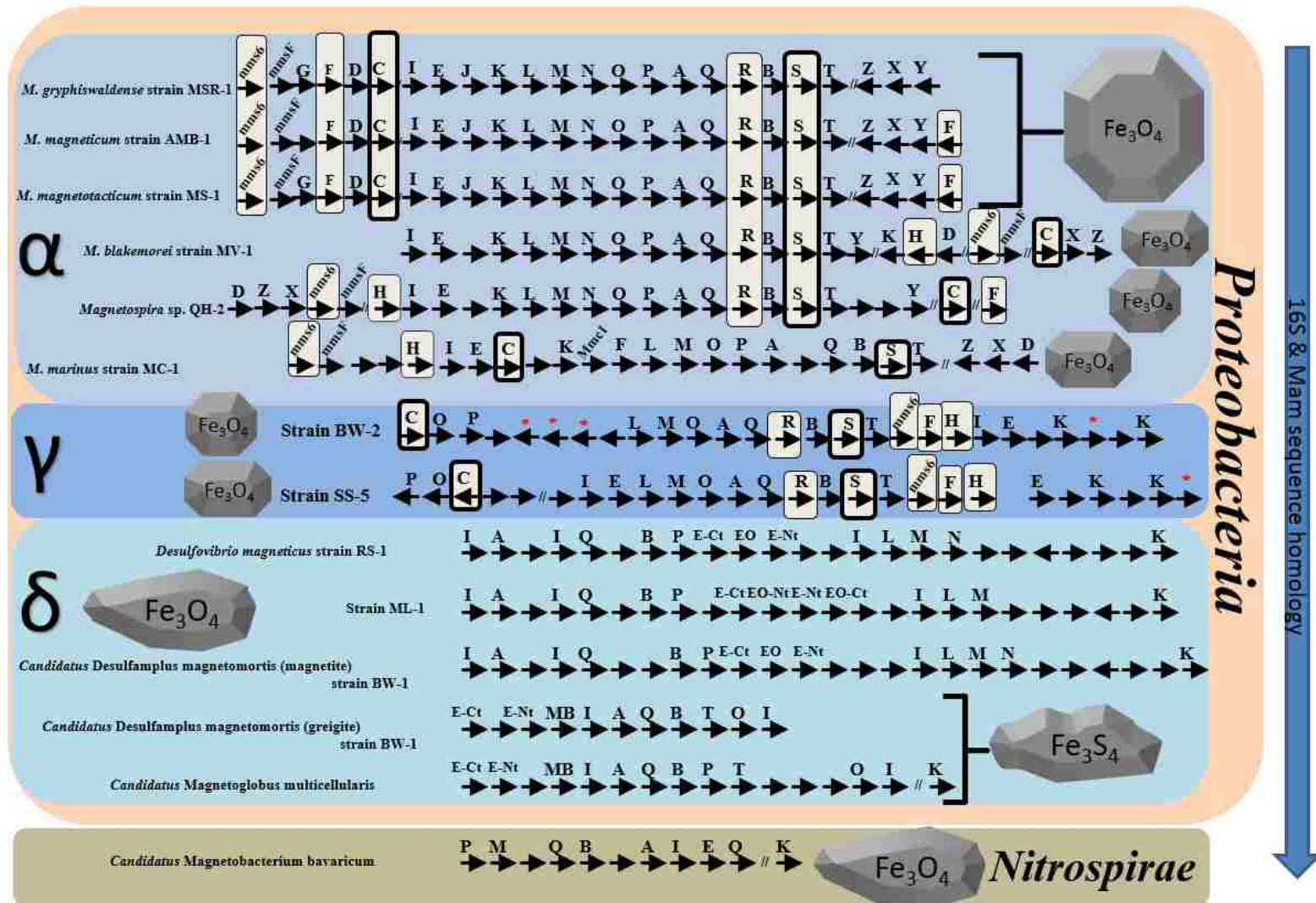
Black arrows indicate magnetosome genes, gray shapes illustrate crystal morphology and composition, colored fields indicate phylogenetic groups, vertical blue arrow indicates phylogenetic divergence, white boxes indicate genes shared only between MTB of *Gammaproteobacteria* and *Alphaproteobacteria*, bolded boxes indicate gene shared by all MTB of the *Alphaproteobacteria* and *Gammaproteobacteria* exclusively and possible uncharacterized magnetosome genes are indicated by asterisk

The synteny of mam genes in the genomes of strains SS-5 and BW-2 is generally

conserved when compared to the *mamAB* operons of MTB of the *Alphoproteobacteria* such as

Magnetospirillum species, *Magnetovibrio blakemorei*[119] or *Magnetospira* sp. QH-2[174] (Figure

37). Some of the differences between SS-5/BW-2 and *Magnetospirillum* species are that *mamN*

is absent, *mamK* is positioned at the downstream end of the cluster, and *mamP* is in another

cluster containing *mamPO* and *C* in SS-5 while they are present in the order *mamC, mamO,*

mamP at the upstream end in strain BW-2 along with the FeoAB MTB specific Fe transport

cluster. Notably, the previously discussed putative cytoskeletal attachment protein *mamJ*[84]

appears to be absent in both strains as is the case with *Magnetovibrio blakemorei*,

Magnetospira sp. QH-2, and *Magnetococcus marinus*. Strains SS-5 and BW-2 have eighteen

magnetosome-associated genes in common only with the different species of the magnetic

Alphaproteobacteria, which produce clearly faceted, cuboctahedral or elongated prismatic

magnetite crystals, but not the MTB of the *Deltaproteobacteria* or *Nitrospirae* (Figure 37).

Notably, of these genes, only *mamC* and *mamS* are common to BW-2, SS-5 and each of the MTB

from the well characterized magnetotactic *Alphaproteobacteria* (*M. magneticum, M.*

gryphiswaldense, M. magnetotacticum, Magnetovibrio blakemorei, Magnetococcus marinus

and Magnetospira sp. QH-2).

Magnetosome gene and morphology phylogeny

When compared to the magnetite-producing magnetotactic *Delta-* and

Alphaproteobacteria, strains BW-2 and SS-5 share with them 11 homologous genes involved in

magnetosome formation (Figure 37). A phylogeny based on the concatenated protein

sequences that are encoded by these genes indicates that their evolution is congruent with the

evolution of the 16S rRNA gene through these strains.[100] Therefore, in light of phylogenetic

analyses, gene synteny and gene conservation, the molecular mechanisms involved in

magnetosome formation in strains BW-2 and SS-5 of the *Gammaproteobacteria* appear to be

more related to the magnetosome formation processes of the magnetotactic

Alphaproteobacteria. This comports with the more similar crystal morphologies between the

groups in that they all produce octahedral or elongated octahedral crystals rather than the

tooth- or bullet-shaped crystals of other phylogenetic groups of MTB (Figure 37). The most

likely mechanism behind the tight crystal morphology control exhibited by the *Alpha-* and

Gammaproteobacteria is mediated by the genes they have in common that are not found in the

bullet-or tooth-shaped magnetite or greigite producers of the *Deltaproteobacteria*.   There is

evidence from gene deletion studies in the *Magnetospirillum* species of the

Alphaproteobacteria that some of these genes are in fact involved in control of magnetite

crystal morphology.[66,83]   Of the *Alpha-* and *Gammaproteobacteria*-specific *mam* genes, *mamC,*

mamH and *mamR* deletion results in size defects only. Deletion of *mamS, mamX, mms6* and

mmsF however, results in defects in shape as well as size, indicating these genes are likely

involved in controlling the observed octahedral magnetite crystal morphology. In particular,

the nature of the aforementioned *mamS* commonality to MTB that produce clearly faceted

magnetite crystals as well as the fact that deletions of *mamS* have been shown to result in the

severe loss of crystal morphology,[83] indicate that *mamS* likely plays a prominent role in

controlling magnetite crystal morphology during magnetosome biomineralization.

There is no genomic evidence in SS-5 for recent horizontal gene transfer events, although immediately downstream of the BW-2 magnetosome gene cluster there are two coding sequences of high homology to a resolvase and a transposase. However, neither the resolvase nor transposase share significant homology to that found in any other strain of MTB. There are no other mobile element proteins or tRNAs in or near the other clusters. Additionally, there is no difference in surrounding region GC content, and the phylogeny and synteny of the Mam proteins in strains SS-5 and BW-2 are highly homologous (Figure 36). Taken together, this seems to indicate that these genes were not acquired through the result of a recent horizontal gene transfer event, but rather through descent via a last common ancestor. Further evidence of this is the fact that the majority of magnetosome proteins in BW-2 and SS-5 are most closely related to another unpublished magnetic *Gammaproteobacterium*, strain *Chromatiales* CG-1. The next most related magnetosome-associated proteins sequences are the corresponding proteins found in the genome of *M. magneticum* of the *Alphaproteobacteria*.   (Table 4, 8 Figure 37).

Potential uncharacterized magnetosome proteins

Genes in BW-2 with homology exclusive to other MTB

Both strains BW-2 and SS-5 contain genes that appear to encode uncharacterized proteins involved in magnetosome formation based on observations related to their homology with genes found in other MTB and their close proximity to other characterized magnetosome

genes. Within the magnetosome gene cluster of strain BW-2 there is a stretch of eight genes

between the *mamCOP* cluster and the *mamAB* cluster that contain three genes which share

identity and synteny with other MTB (Figure 38). After a hypothetical amino acid carrier

protein immediately downstream from *mamP* there is a protein with high homology to the

Na$^+$/Gly symporter GlyP and proteins of the Na$^+$/Ala symporter superfamily. Next, there is a

chemotaxis response regulator with high homology to CheY, a protein that transmits

chemoreceptor signals to flagellar motor components. The next putative protein contains a

histidine kinase domain with multiple sensor domains. Then, a membrane transport protein, a

ferritin protein, a chemoreceptor Zn binding protein, and a Yedz heme binding type protein

encoding gene.

Figure 37 Uncharacterized Genes of Interest in BW-2 Mam gene
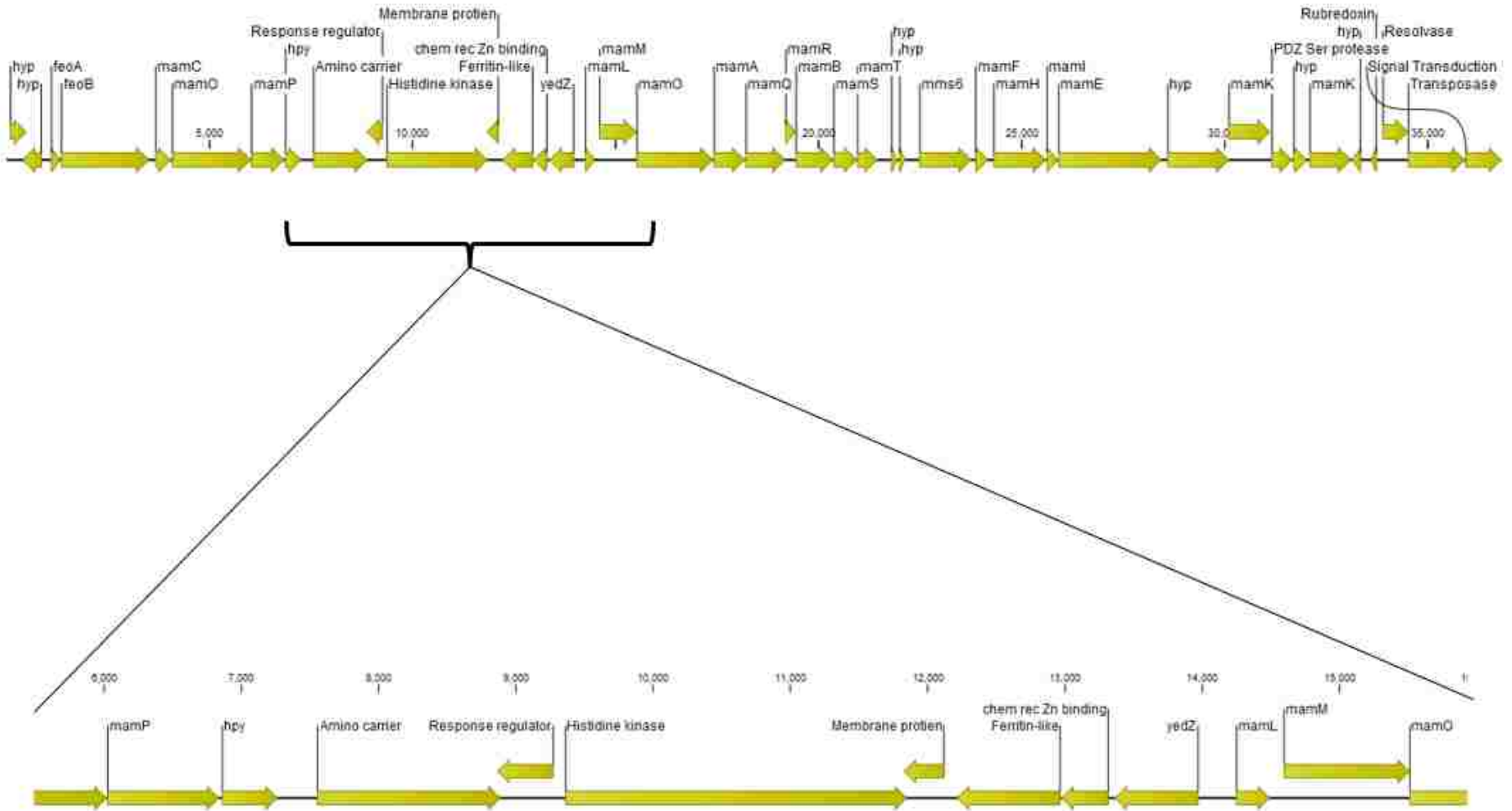
BW-2 magnetosome gene cluster (top) and expanded view of uncharacterized genes of interest within the cluster (bottom)

The amino acid transporter type protein is most highly related to genes found in six other magnetic bacteria and most related to one found in the *Chromatiales* strain CG-1 with an e-value of 2.00E-161, then to *Magnetococcus marinus* with an e-value of 9.00E-143 (Table 11). It is less closely matched to an amino acid symport gene found in strain SS-5 with an e-value of 2.00E-106 with no nearby characterized magnetosome genes. In *Magnetococcus marinus*, the amino acid symport type protein is located within the magnetosome gene cluster immediately downstream of *mamF* and just upstream of *mamH* with two hypothetical coding sequences intervening. It is located with some synteny in *Magnetovibrio blakemorei* immediately downstream from the Fe transporter like cluster *feoAB*-like, which is found in all other MTB to date.[33] It is also upstream of a hypothetical protein with homology to the hypothetical protein immediately upstream of the *feoAB*-like cluster of strain BW-2 (Figure 39).

Table 9 Sequences with Homology to the Uncharacterized Amino Acid Carrier within the Magnetosome Cluster of Strain BW-2

Gene homology	E-val	Identity
Chromatiales CG1	2.00E-161	53%
Magnetococcus marinus Mc-1	9.00E-143	47%
Magnetofaba australis IT-1	7.00E-137	46%
Magnetovibrio blakemorei MV-1	6.00E-135	47%
Magnetospira sp. QH-2	6.00E-131	45%
Magnetovibrio blakemorei	1.00E-130	47%
Terasakiella sp. PR1	2.00E-129	45%
Streptococcus suis	2.00E-81	38%

Figure 38 Amino Acid Symporter Synteny Between Strain BW-2 and Magnetovibrio blakemorei
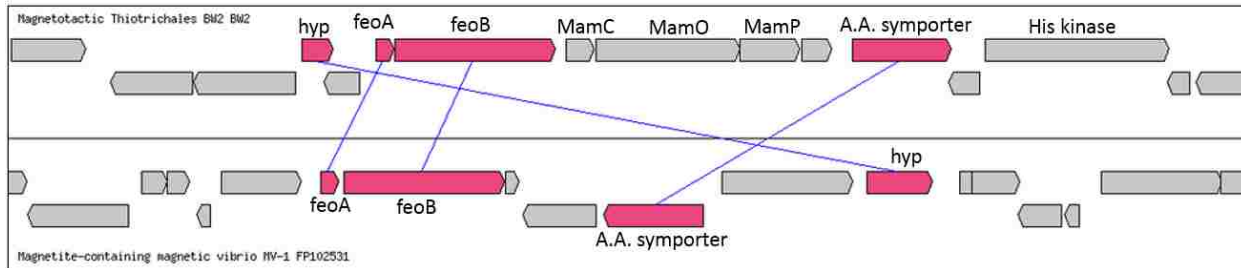
Arrows indicate genes and colored arrows indicate genes displaying synteny with the compared strains. Blue lines connect the homologous genes discussed to illustrate synteny.

The histidine kinase just upstream to the amino acid symporter is a complex, multifaceted protein also found in strains CG1 and SS-5 (Table 12). In SS-5 it is in the same orientation and proximity with the adjacent CheY- like protein, although unlike as in BW-2 it is not clustered with known magnetosome genes (Figure 40). The array of homologous domains within this protein include transcriptional regulators, multiple signaling mechanism domains, an ATP binding domain and photo sensors (Table 13). Interestingly, this histidine kinase in SS-5 is followed by a gene containing a BLUF domain protein involved in sensing blue light, although it is not found associated with magnetosome genes.

Table 10 Sequences with Homology to the mam Cluster Associated His Kinase of Strain BW-2

Gene homology	E-val	Identity
Chromatiales CG-1	2.00E-175	57.96
Thiotrichales SS-5	4.00E-167	56.85

Figure 39 Synteny of the mam Cluster Associated His Kinase between Strain BW-2 and Strain SS-5
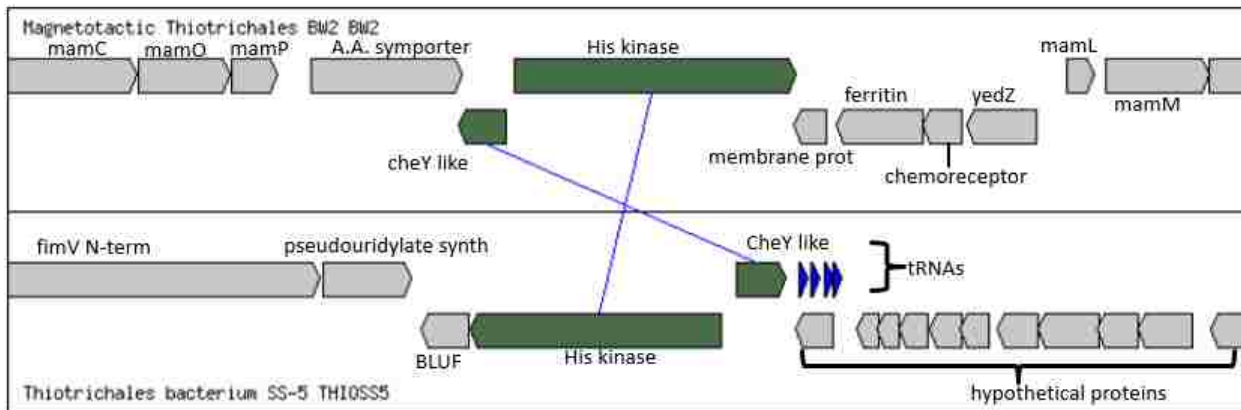
Arrows indicate genes and colored arrows indicate genes displaying synteny with the compared strains. Blue lines connect the homologous genes discussed to illustrate synteny.

Table 11 Characterized Protein Domains of the mam cluster associated His Kinase in strain BW-2

Domain superfamily	Identified functions
PAS	Bind ligands, and to act as sensors for light and oxygen in signal transduction
GAF_2	Upregulate ATP and GTP conversion to cAMP and cGMP Present in Phytochromes - regulatory photoreceptors FhlA and NifA transcriptional regulators (often for NifA)
HisKA	conserved signaling mechanisms in prokaryotes and eukaryotes
HATPase_c	Histidine kinase-like ATPases, ATP binding
REC	Signal receiver containing a phosphoacceptor site phosphorylated by histidine kinase
Hpt	Signaling through a two part component systems

After the histidine kinase encoding sequence, there is another hypothetical gene

unrelated to any currently characterized protein, and then a coding sequence coding for a

protein containing a ferritin domain. This coding sequence is unrelated to sequences in non-

MTB, and is most related to one found in strain CG-1, and is also found in *Magnetococcus*

marinus found between *mamC* and *mamK* (Table 14).

Table 12 Sequences with Homology to the mam Cluster Associated Ferritin of Strain BW-2

Organism with gene homology	e-value	Identity
Chromatiales CG-1	5e-60	50.25
Magnetococcus marinus	6e-53	53

There is also an uncharacterized gene that could play a role in magnetosome

biomineralization located in the magnetosome gene cluster between the duplicated *mamK*

genes of BW-2, a duplication which is also found in several MTB of the *Alphaproteobacteria*,

and in the genome of strain SS-5 but present some distance away from other magnetosome

genes. This coding sequence encodes a PDZ serine protease domain similar to the previously

discussed MamP. There is a 38% identity match with an e value = 3e-09 between the gene in

BW-2 and that of SS-5, a 40% match with e value = 6e-08 with a coding sequence in

Magnetospira sp. QH-2 and a 37% match with e value = 5e-07 with a coding sequence in *M.*

magneticum (Table 15).

Table 13 BW-2 mam Cluster Associated Serine Protease Homology

Organism with gene homology	E-val	% Identity
SS-5	3e-09	38
Magnetospira sp. QH-2	6e-08	40
Magnetospirillum magneticum	5e-07	37

Genes in SS-5 with homology exclusive to other MTB

There is also the second gene immediately following the second *mamK* of SS-5 with homology to one in the *Chromatiales* strain CG-1, one in *Magnetospira* sp. QH-2 and two found in *Magnetococcus marinus* (Table 16). In *Magnetococcus marinus*, QH-2 and CG-1 these are not associated with magnetosome genes and this coding sequence does not contain sequence homologous to a currently characterized protein domain.

Table 14 SS-5 Genes with Homology to the Uncharacterized Gene Following mamK in the mam Cluster of Strain SS-5

Organism with gene homology	E-val	% Identity
Chromatiales CG-1	3.00E-95	76.27
Magnetospira QH-2	3.00E-53	48
Magnetococcus marinus	9.00E-17	33
Magnetococcus marinus	1.00E-12	34

As previously stated, SS-5 has three hypothetical magnetosome protein encoding genes based on conservation in other magnetotactic species designated *hmp2*, *hmp3 and hmp4*. These genes have similar localization and organization with homologous genes in other MTB with sequences in National Center for Biotechnology Information (NCBI) database (Tables 17-19). They are clustered in SS-5 with a *mamD* gene away from the *mamAB* or *mamCOPD* clusters. This arrangement is conserved with the arrangement found in the unfinished *gammaproteobacterium* of the *Chromatiales* designated as CG-1 (Figure 41). *hmp2* contains a region that codes for an adenylate cyclase domain, whereas *hmp3* and *hmp4* have no known protein domain coding homology. *hmp2* and *hmp3* are additionally highly homologous with uncharacterized genes in other MTB of the *Alphaproteobacteria* with sequence stored in the European Molecular Biology Laboratory database (EMBL) TrEMBL that do not share similar synteny (Tables 20,21).

Table 15 SS-5 Sequences with Homology to the hmp4 of Strain SS-5

Organisms with homology	% Identity	e-value
Chromatiales CG-1	64.1	5.00E-55
Magnetospirillum sp. SO-1	35.14	4.49501E-09
Magnetospirillum magneticum	37.84	1.00E-10
Magnetospirillum magnetotacticum	36.49	1E-10
Magnetospirillum moscoviense	35.14	2E-10
Magnetospirillum marisnigri	37.88	1.00E-09

Table 16 SS-5 Sequences with Homology to the hmp3 of Strain SS-5

Organism with homology	% Identity	e-value
Chromatiales CG-1	86.51	0
Magnetospirillum marisnigri	30.12	2.00E-74
Magnetospirillum magneticum	29.7	1.00E-69
Magnetospirillum magnetotacticum	29.64	2.00E-68
Magnetospirillum moscoviense	29.29	2.00E-66
Magnetospirillum sp. SO-1	29.54	3.99E-59

Table 17 SS-5 Sequences with Homology to the hmp2 of Strain SS-5

Organisms with homology	% Identity	e-value
Chromatiales CG-1	94.44	0
Magnetospirillum marisnigri	36.56	1.00E-119
Geobacter metallireducens	46.46	1.00E-21

Figure 40 Synteny of mamD/hmp Clusters of Strains SS-5 and CG-1
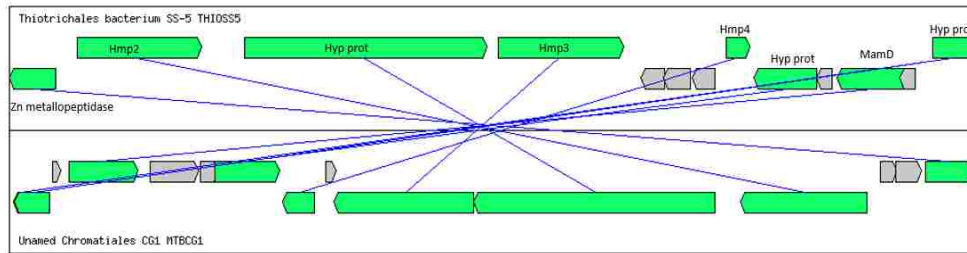
Arrows indicate genes and colored arrows indicate genes displaying synteny with the compared strains. Blue lines connect the

homologous genes discussed to illustrate synteny.

Table 19 SS-5 hmp2 Genes of Homology TrEMBL Hits

Organism	PB id	Ident %	Eval	Gene	Description	PubMedId
Spirochaetes bacterium DG_61	A0A0S7ZRD3	38.23	2.00E-122	Hypothetical	Uncharacterized protein	25922666
Magnetospirillum moscoviense	A0A178MSU3	37.05	1.00E-117	Hypothetical	Uncharacterized protein	NA
Magnetospirillum gryphiswaldense	A4U2F6	37.09	2.00E-116	Hypothetical	Adenylate cyclase, family 3	17449609
Magnetospirillum marisnigri	A0A178MMH7	36.56	8.00E-116	Hypothetical	Uncharacterized protein	NA
Magnetospirillum sp. XM-1	A0A0U5MKG1	36.67	2.00E-115	Hypothetical	Putative adenylate cyclase (Partial)	NA
Magnetospirillum sp. SO-1	M2Z763	36.52	2.00E-114	Hypothetical	Calphotin	24723706
Magnetospirillum magnetotacticum	A0A0C2Z0X5	36.52	4.00E-112	Hypothetical	Uncharacterized protein	NA
Magnetospirillum magneticum	Q2W1X4	39.78	5.00E-112	Hypothetical	Calphotin Microtubule-associated protein	16303747
Aphanocapsa montana	A0A0C1WJJ6	36.53	2.00E-107	Hypothetical	Uncharacterized protein	NA
Lyngbya confervoides	A0A0C1YKM2	36.53	2.00E-107	Hypothetical	Uncharacterized protein	NA
Gammaproteobacteria bacterium SG8_31	A0A0S8FKL1	35.65	2.00E-97	Hypothetical	Uncharacterized protein	25922666
Sphingomonas paucimobilis	A2PZP1	29.49	5.00E-26	Hypothetical	Uncharacterized protein	17928721
Chlamydomonas reinhardtii	Q5YLC2	28.08	0.008	CYG12	Guanylate cyclase	15472039
Anabaena sp	P94180	25.31	0.12	cyaA	Adenylate cyclase	9171404

Table 18 SS-5 hmp3 Genes of Homology TrEMBL Hits

Organism	PB id	Ident %	Eval	Gene	Description	PubMedId
Magnetospirillum marisnigri	A0A178MKG7	30.12	2.00E-70	Hypothetical	Uncharacterized protein	NA
Magnetospirillum sp. XM-1	A0A0U5MFK6	29.77	6.00E-67	Hypothetical	Uncharacterized protein	NA
Magnetospirillum magneticum	Q2W8W3	29.7	9.00E-66	Hypothetical	Uncharacterized protein	16303747
Magnetospirillum magnetotacticum	A0A0C2YAJ1	29.64	2.00E-64	Hypothetical	Uncharacterized protein	NA
Magnetospirillum moscoviense	A0A178MYP7	29.29	2.00E-62	Hypothetical	Uncharacterized protein	NA
Magnetospirillum sp. SO-1	M3A5I1	29.54	1.00E-61	Hypothetical	Uncharacterized protein	24723706
Magnetospira sp. QH-2	W6KB61	30.89	4.00E-58	Hypothetical	Uncharacterized protein	23841906
uncultured bacterium	C5JAI1	28.49	1.00E-44	Hypothetical	Conserved uncharacterized protein	NA
Magnetovibrio blakemorei	C4RAE8	30.57	1.00E-35	Hypothetical	Uncharacterized protein	19220405
marine sediment metagenome Expedition CK06-06	X0XMQ8	37.13	7.00E-29	Hypothetical	Uncharacterized protein	24624126
Trypanosoma brucei	Q99356	73.44	4.00E-24	Hypothetical	Procyclin (pPS-1)	2011148
Rattus norvegicus	Q9JLE9	71.43	1.00E-22	Gabre	GABA-A receptor epsilon-like subunit	10804200

There are hypothetical proteins interspersed with these SS-5 and CG-1 *hmp* clusters that

are not significantly similar in sequence to those of non MTB species. The arrangement of this

cluster is also conserved between SS-5, CG-1 and the unfinished *Magnetospirillum* strain SO-1,

but with SO-1 limited to *hmp3, hmp4* and *mamD*. Another notable difference is that the SO-1

cluster is immediately flanked on either side by four phage proteins, indicating this strain may

have acquired this cluster through a more recent horizontal gene transfer event (Figure 42).

When compared with *M. magneticum*, the synteny is again conserved with that of SS-5, CG-1

and SO-1, and again this cluster is flanked by an integrase and a transposase, but limited to

hmp3 and *hmp4* (Figure 43). Synteny of the *hmp/mamD* cluster is conserved not only between

SS-5, CG-1, SO-1, *M. magneticum* but with *M. magnetotacticum* as well, again with three phage

proteins similar in type to those found flanking the same cluster of SO-1 (Figure 44). Yet again,

this cluster arrangement is also observed in the unfinished *M. moscoviense* strain BB-1. In this

instance, rather than flanked by phage proteins, the cluster is adjacent upstream to a *mamM*

fragment, *mamO* and *mamB* (Figure 45). Finally, this synteny is yet again conserved between

SS-5 and the unfinished *M. marisnigri* strain SP-1. *M. marisnigri* has the *hmp3, hmp4* and

mamD genes, but they are accompanied by *mamC* immediately upstream of *mamD* and flanked

by a mobile element protein and a virulence protein on the other side of the cluster (Figure 46).

The fact that this cluster is so highly conserved in gene compliment and synteny across only

species of MTB and flanked with magnetosome genes in three of them seems to indicate a

potential role in the magnetic phenotype. That phage genes flank this cluster for four of these

species indicates potential genomic rearrangement of this cluster and conceivably horizontal

gene transfer. Adding to this possibility is the fact that this cluster is found in MTB across phylogenetic groups, but not in all MTB members of those groups.

Notably, although they are not found in the *Gammaproteobacteria*, there are eight additional hypothetical proteins of interest between *mamC* and the *hmp* genes of this cluster as found in *M. marisnigri*.   These genes could be involved in magnetosome formation in the *Alphaproteobacteria* as they are highly conserved in sequence similarity among *Magnetospirillum* species, but not found in other species (Figure 46, Tables 22-29).

Figure 41 Synteny of mamD/hmp Clusters in Strains SS-5 and SO-1
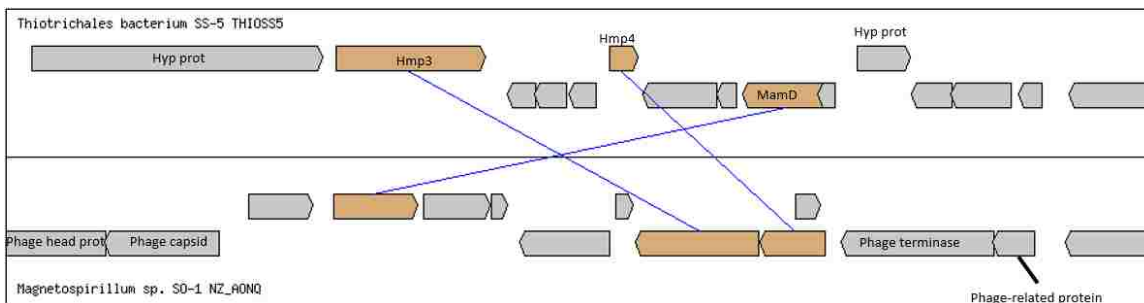
Arrows indicate genes and colored arrows indicate genes displaying synteny with the compared strains. Blue lines connect the homologous genes discussed to illustrate synteny.

Figure 42 Synteny of mamD/hmp Clusters of Strains SS-5 and M. magneticum
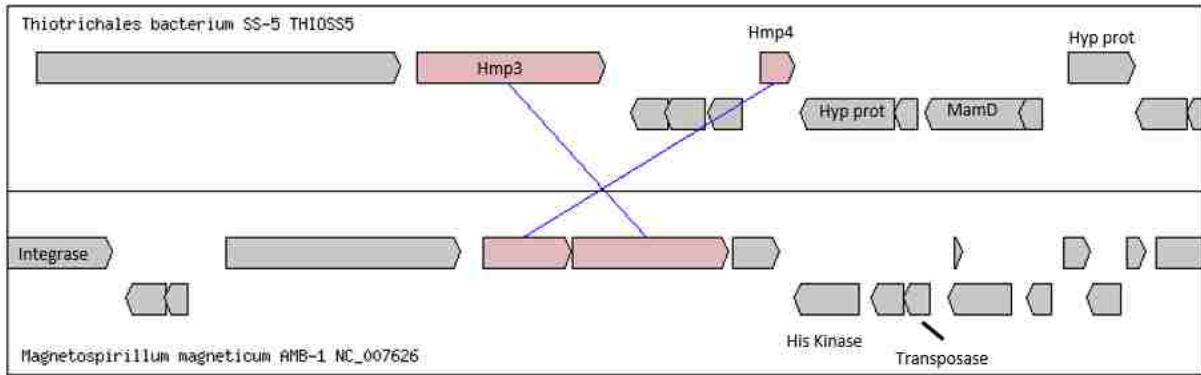
Arrows indicate genes and colored arrows indicate genes displaying synteny with the compared strains. Blue lines connect the homologous genes discussed to illustrate synteny.

Figure 43 Synteny of mamD/hmp Clusters of Strians SS-5 and M. magnetotacticum
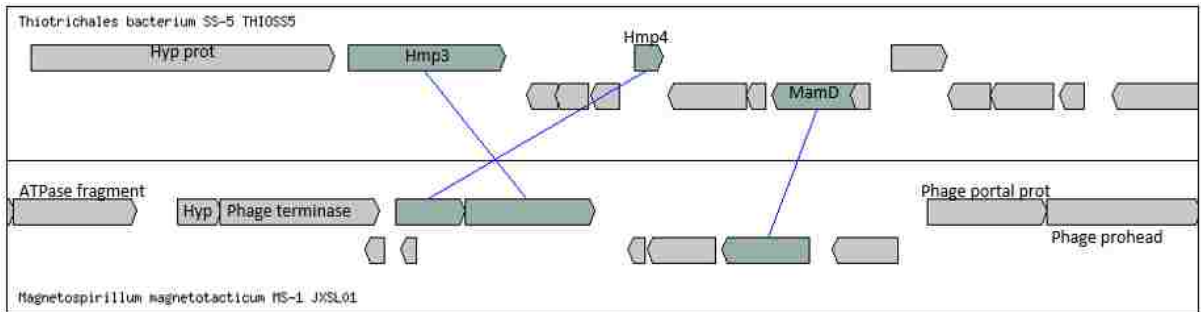
Arrows indicate genes and colored arrows indicate genes displaying synteny with the compared strains. Blue lines connect the homologous genes discussed to illustrate synteny.

Figure 44 Synteny of mamD/hmp Clusters of Strains SS-5 and M. moscoviense
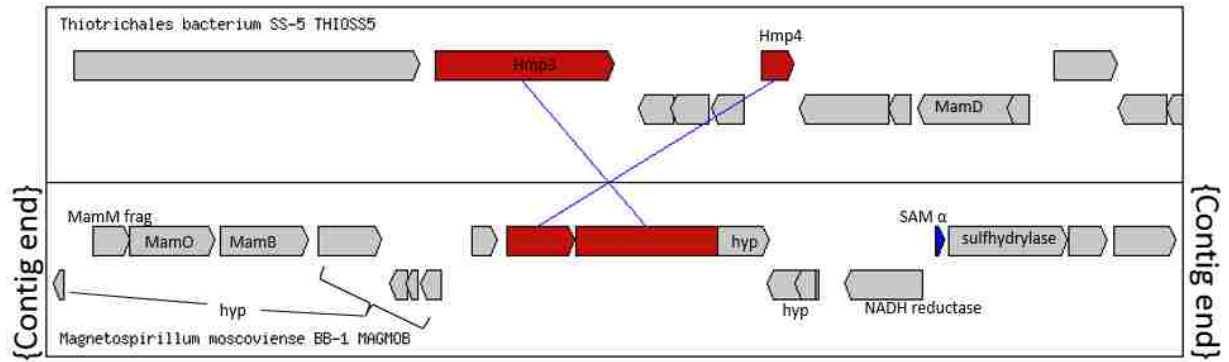
Arrows indicate genes and colored arrows indicate genes displaying synteny with the compared strains. Blue lines connect the homologous genes discussed to illustrate synteny.

Figure 45 Synteny of mamD/hmp Clusters of SS-5 and M. marisnigri
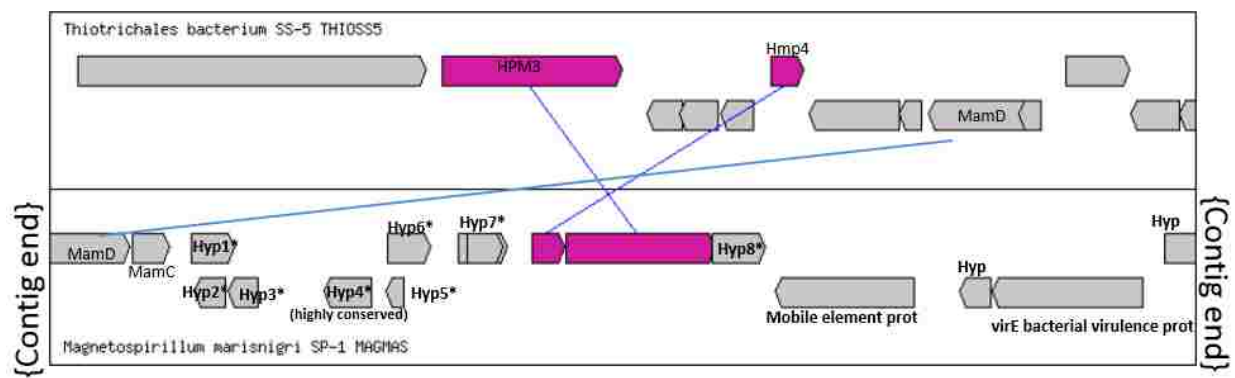
Arrows indicate genes and colored arrows indicate genes displaying synteny with the compared strains. Blue lines connect the homologous genes discussed to illustrate synteny.

Table 20 SP-1 Hypothetical Protein 1

*SP-1 hyp prot 1 homologies		
	% Identity	e-value
Magnetospirillum sp. SO-1	100	3.00E-105
Magnetospirillum sp. XM-1	100	3.00E-105
Magnetospirillum magnetotacticum	96	3.00E-47

Table 21 SP-1 Hypothetical Protein 2

*SP-1 hyp prot 2 homologies		
	% Identity	e-value
Magnetospirillum sp. SO-1	97.17	9.00E-74
Magnetospirillum sp. XM-1	97.17	1.00E-72
Magnetospirillum magnetotacticum	95.28	1.00E-71
Magnetospirillum magneticum	96.77	2.00E-63
Magnetospirillum sp. XM-1	46.6	1.00E-28
Magnetospirillum magnetotacticum	44.44	1.00E-17

Table 22 SP-1 Hypothetical Protein 3

*SP-1 hyp prot 3 homologies		
	% Identity	e-value
Magnetospirillum magnetotacticum	90.91	3.00E-37
Magnetospirillum sp. XM-1	88.06	7.00E-37
Magnetospirillum sp. XM-1	63.08	4.00E-24
Magnetospirillum magnetotacticum	63.08	2.00E-23
Magnetospirillum sp. SO-1	39.68	5.00E-07
Magnetospirillum sp. XM-1	39.68	6.00E-07
Magnetospirillum magnetotacticum	39.68	6.00E-07

Table 23 SP-1 Hypothetical Protein 4

*SP-1 hyp prot 4 homologies		
	% Identity	e-value
Magnetospirillum sp. XM-1	98.11	1.00E-112
Magnetospirillum magneticum	98.11	1.00E-112
Magnetospirillum sp. SO-1	98.11	1.00E-112
Magnetospirillum magnetotacticum	93.67	2.00E-107

Table 24 SP-1 Hypothetical Protein 5

*SP-1 hyp prot 5 homologies		
	% Identity	e-value
Magnetospirillum magnetotacticum	87.93	9.00E-29
Magnetospirillum sp. SO-1	95	6.00E-21

Table 25 SP-1 Hypothetical Protein 6

*SP-1 hyp prot 6 homologies		
	% Identity	e-value
Magnetospirillum sp. SO-1	96.53	5.00E-102
Magnetospirillum sp. XM-1	94.93	5.00E-94
Magnetospirillum magneticum	93.48	1.00E-92
Magnetospirillum magnetotacticum	88.41	3.00E-88

Table 26 SP-1 Hypothetical Protein 7

*SP-1 hyp prot 7 homologies		
	% Identity	e-value
Magnetospirillum sp. XM-1	90.85	3.00E-108
Magnetospirillum magnetotacticum	87.58	2.00E-98
Magnetospirillum sp. SO-1	88.5	7.00E-72
Magnetospirillum magneticum	83.64	1.00E-27

Table 27 SP-1 Hypothetical Protein 8

*SP-1 hyp prot 8 homologies		
	% Identity	e-value
Magnetospirillum magneticum	90.85	9.00E-103
Magnetospirillum magneticum	94.63	3.00E-97
Magnetospirillum moscoviense	79.5	6.00E-88
Magnetospirillum sp. XM-1	84.56	4.00E-87

Upon finishing the genome of *M. marisnigri* it will be possible to investigate further whether there are additional genes shared by other MTB in this vicinity.

Although such preliminary evidence of potentially uncharacterized magnetosome genes is of interest, functional studies would of course need to be performed to confirm these observations. While these studies would be difficult to conduct in the *Gammaproteobacteria* as there is yet no established genetic system, this could be explored further in the case of the eight genes noted above that were identified in the *Magnetospirillum* as there are tractable genetic systems in place for the *Magnetospirillum*.

As previously stated, the evidence gathered here supports the likelihood that the magnetic phenotype was acquired by descent in the currently known *Gammaproteobacteria* in that only traces of evidence for the magnetosome genes of either BW-2 or SS-5 having ever been associated with a genomic island capable of mediating horizontal gene transfer are detectable. The resolvase and transposase coding sequences at one end of the magnetosome gene cluster of BW-2 also don't share homology to resolvase and transposase genes of known magnetic bacteria. Additionally, the homology of most of the characterized magnetosome genes generally correlate to the phylogeny of the bacterium in which they are found. The yet

uncharacterized genes found among the magnetosome genes of BW-2 and SS-5 appear to be

exceptions, though, in that they share sequence identity with magnetic bacteria they are not

most closely related to in terms of genome phylogeny. One explanation for this could be that

those not shared between BW-2 and SS-5 were lost from one or the other during

rearrangement events that led to the lack of synteny in these areas.

Chapter 6 Conclusion

In conclusion, the genomes of strains BW-2 and SS-5, the first two discovered magnetotactic bacteria (MTB) belonging to the *Gammaproteobacteria*, were sequenced, sealed, annotated and compared to one another and to MTB of other phylogenetic groups. Specific goals were to determine the metabolic potential of these two strains as well as the compliment, synteny, origin and evolution of the magnetosome genes, whether there may be sequences that can be identifiable as potentially novel magnetosome associated genes and whether there are genes conspicuously associated with the octahedral magnetosome crystal morphology.

As to general metabolism, the genomes of both MTB strains contain the genes necessary for the utilization of the Calvin-Benson-Bassham cycle for autotrophy and carboxysome formation as well as the genes necessary for the oxidation of reduced sulfur compounds as a source of electrons and nitrogen fixation and dissimilatory nitrogen oxide reduction. This is consistent with their ability to grow as microaerophilic chemolithoautotrophs using thiosulfate and sulfide as electron donors.

Characterized magnetosome genes form a single cluster about 2.3 kb in length in the genome of strain BW-2 while magnetosome genes make up two clusters of about 21.8 and 5 kb in length in the genome of strain SS-5. The gene compliment and synteny of these strains is more closely conserved with one another, then with the octahedral magnetite producers of the *Alphaproteobacteria*, mirroring the phylogeny of the rest of their genomes and indicating the phenotype was likely acquired by descent.   Interspersed among the magnetosome genes of

both MTB strains are unique coding sequences that share high homology to uncharacterized, hypothetical magnetosome-associated genes of other MTB. Some of these sequences are associated with uncharacterized coding sequences in MTB of *Alphaproteobacteria* that are highly homologous with one another, only found in these MTB and are similarly organized.

The highly conserved in gene compliment and synteny of this cluster across only species of MTB, flanked with magnetosome genes in three of them seems to indicate a potential role in the magnetic phenotype. In four of these species phage genes flank this cluster and it is found in MTB across phylogenetic groups, but not in all MTB members of those groups. This indicates the potential genomic rearrangement of this cluster independently of the characterized magnetosome gene clusters and conceivably horizontal gene transfer.

Further characterization of these newly identified hypothetical magnetosome genes would require the establishment of a genetic system for experimentally derived generation of data. While significant progress was made during this study in the refinement of culturing technique resulting in the robust growth in strain BW-2, a genetic system is not yet in place for the MTB of the *Gammaproteobacteria*. There are, however, established genetic systems in place for *M. magneticum* and *M. magnetotacticum* of the *Alphaproteobacteria* in which novel hypothetical magnetosome genes were identified, presenting the immediate opportunity for further investigation.

Appendix

Table 28 Magnetosome Associated Genes and Functions

Protein	Localization	Process	Function	Deletion Effects
MamA	Cytosol. Dynamic, surrounding vesicles	Invagination of cell membrane	It has multiple domains with TPR motifs (protein-protein interactions); may act as multi-protein assembly site; stabilizes magnetosome chain.	Invagination is not affected. Reduction in the number of magnetosomes and changes in iron accumulation.
MamB	Transmembrane in MM	Iron transport and magnetite nucleation	May be involved in iron transport since has homology to CDF (cation diffusion facilitator). Contains TPR domain (protein-protein interactions) and interacts with MamE; requires MamM for stabilization.	Loss of magnetosome vesicles and of crystal formation.
MamC	Transmembrane in MM	Crystal shape and size	Its loop interacts with magnetosome crystal. It is not essential to biomineralization but may control chemical conditions inside vesicles.	Changes in size and organization of chains and size of vesicles. No effects observed in crystal size or shape.
MamD	Transmembrane in MM, N-terminal in ML	Crystal shape and size	Associated with control over size of magnetosome crystal.	Changes in crystal size.
MamE	Transmembrane in MM, C-terminal towards ML	Iron transport and nucleation	Acts as a serine protease and has PDZ domain (protein-protein interaction) which interacts with MamB and I. Magnetochrome might control the magnetosome redox state and balance between Fe^{2+}/Fe^{3+}.	Formation of empty magnetosome vesicles, loss of magnetite synthesis, mislocation of MamI and other Mam proteins.
MamF	Transmembrane MM	Crystal shape and size	Associated to control of magnetosome size; interacts with crystal.	Changes in crystal size.
MamG	Transmembrane in MM	Crystal shape and size	Associated to control of magnetosome size.	Changes in crystal size.
MamH	Transmembrane in MM	Iron transport and nucleation	Contains conserved domains homologous to MFS proteins (membrane transporters) and might function as phosphate transporter during magnetite biomineralization.	Reduced number and size of magnetosomes.
MamI	Transmembrane in MM	Invagination	Involved in the formation and bending of the MM.	Absence of MM.
MamJ	Cytosol	Arrangement of chains	Acts as an anchor between MamK filaments and vesicle membrane to arrange magnetosomes in a chain.	Magnetosomes arranged in clusters and no longer in chains. Reduced magnetotactic response.
MamK	Cytosol	Arrangement of chains	Controls chain assembly and position along the cell axis; positions chain for cellular division; homologous to MreB (actin-like).	Lack of filaments near the magnetosomes. Shorter chains and wrong position of MamJ.
MamL	Transmembrane in MM	Invagination	Involved in the formation of MM; similar to MamI.	Absence of MM.
MamM	Transmembrane in MM	Iron transport and magnetite nucleation	Involved in iron transport and may use H^+/cation antiporter mechanism. Involved in the begining of crystalization and localization of other Mam proteins; stabilizes MamB; homologous to CDF (cation diffusion facilitator).	Loss of magnetite crystals, formation of empty vesicles.

MamN	Transmembrane in MM	Iron transport and magnetite nucleation	Homologous to Na$^+$/H$^+$ antiporter and might be involved in the extrusion of H$^+$ from the vesicle.	Formation of empty magnetosome vesicles. Does not affect localization of other proteins.
MamO	Transmembrane in MM; *C*-terminal in ML	Iron transport and magnetite nucleation	Composed of two domains: (1) transmembrane, homologous to proteins involved in transport of anions across cell membrane and (2) similar to a trypsin-like peptidase, but possibly with no protease function.	Formation of empty magnetosome vesicles.
MamP	Transmembrane in MM with active sites towards ML	Iron transport and magnetite nucleation	Involved in control of crystal number and size and in electron transfer necessary to magnetosome assembly and magnetite formation; similar to MamE and MamT; may contain an iron-binding site.	Defects in crystal size, fewer magnetosomes per cell.
MamQ	Transmembrane in MM; *C*-terminal in ML	Invagination	Unknown function; homologous to LemA.	Complete loss of magnetosome formation in AMB-1.
MamR	Cytosol	Crystal shape and size	Controls the number and size of crystals; predicted to have a DNA-binding domain.	Smaller magnetosome and weaker magnetotactic response.
MamS	Transmembrane in MM; *C*-terminal in ML	Crystal shape and size	Controls the number and size of crystals.	Defects in crystal size and morphology, weaker magnetotactic response.
MamT	Transmembrane in MM; *C*-terminal in ML	Iron transport and magnetite nucleation	Involved in regulation of crystal size and morphology; has a magnetochrome domain.	Defects in crystal maturation and loss of magnetotactic response.
MamU	Cytosol	Invagination	Unknown function. Homologous to DGK Family, that includes kinase involved in regulation of cell response.	None observed.
MamV	Transmembrane in MM	Iron transport and magnetite nucleation	Putative CDF transporter.	None observed.
MamW	MM (structure unknown)	Iron transport and magnetite nucleation	Implicated it magnetite synthesis or not associated to magnetosomes.	None observed.
MamX	Transmembrane in MM; *C*-terminal in ML	Iron transport and magnetite nucleation	Involved in electron transport, with Cytochrome c-like domain; weak similarity to MamS and E.	Smaller crystals and with irregular shapes. Weaker magnetotactic cell response.
MamY	Transmembrane in MM; *C*-terminal in cytosol	Invagination	Constricts the MM and consequently affects crystal growth; homologous to BAR proteins (involved in membrane dynamics).	Enlarged magnetosome vesicles with smaller crystals.
MamZ	Transmembrane in MM; *C*-terminal in ML	Iron transport and magnetite nucleation	Involved in redox control for magnetosome formation; creates an iron oxidoreductase and transport complex with MamX and MamH.	Smaller size of crystals and higher proportion of twinned crystals.
Mms6	Transmembrane in MM	Crystal shape and size	Involved in the initiation of magnetite synthesis and control of crystal shape; presents *in vitro* activity.	Smaller magnetosomes with heterogeneous shapes. Irregular alignment of chains.
MmsF	Transmembrane in MM	Crystal shape and size	Involved in the control of size and shape of magnetite crystal during maturation.	Formation of elongated crystals and of non-magnetotactic cells.

Magnetosome genes and functions as determined by knock-out experiments in *M. gryphiswaldense*. BAR = Bin/Amphiphysin/Rvs domain related to membrane dynamics; CDF = cation diffusion facilitator; DGK = diacylglycerol kinase; MM = magnetosome membrane associated; ML = magnetosome lumen associated; PDZ = conserved domain enabling protein-protein interaction; MFS = major facilitator superfamily of transporters; TPR = Tetratricopeptide repeat domain enabling protein-protein interactions. (Table modified from Araujo 2015)

References

1.   Lefèvre, C. T. *et al.* Novel magnetite-producing magnetotactic bacteria belonging to the

     Gammaproteobacteria. *ISME J.* **6,** 440–450 (2012).

2.   Bazylinski, D. a & Schu, D. Biomineralization and Assembly of the Bacterial Magnetosome

     Chain. *Microbe* **4,** 124–130 (2009).

3.   Frankel, R. B. & Bazylinski, D. a. Biologically induced mineralization by bacteria. *Rev.*

     Mineral. Geochemistry **54,** 95–114 (2003).

4.   Bazylinski, D. a. & Frankel, R. B. Biologically Controlled Mineralization in Prokaryotes.

     Rev. Mineral. Geochemistry **54,** 217–247 (2003).

5.   Friedrich, C. G., Bardischewsky, F., Rother, D., Quentmeier, A. & Fischer, J. Prokaryotic

     sulfur oxidation. *Curr. Opin. Microbiol.* **8,** 253–259 (2005).

6.   Lefevre, C. T. & Bazylinski, D. a. Ecology, Diversity, and Evolution of Magnetotactic

     Bacteria. *Microbiol. Mol. Biol. Rev.* **77,** 497–526 (2013).

7.   Balkwill, D. L., Maratea, D. & Blakemore, R. P. Ultrastructure of a magnetotactic spirillum.

     J. Bacteriol. **141,** 1399–1408 (1980).

8.   Frankel, R. B. & Bazylinski, D. a. How magnetotactic bacteria make magnetosomes queue

     up. *Trends Microbiol.* **14,** 329–331 (2006).

9.   Lower, B. H. & Bazylinski, D. a. The Bacterial Magnetosome: A Unique Prokaryotic

     Organelle. *J. Mol. Microbiol. Biotechnol.* **23,** 63–80 (2013).

10. Yan, L. *et al.* Magnetotactic bacteria, magnetosomes and their application. *Microbiol. Res.* **167,** 507–519 (2012).

11. Bazylinski, D. a. & Schübbe, S. Controlled Biomineralization by and Applications of Magnetotactic Bacteria. **62,** 21–62 (2007).

12. Bazylinski, D. A. & Lefe, C. T. *The Prokaryotes*. (2013). doi:10.1007/978-3-642-30141-4

13. Lefèvre, C. T., Frankel, R. B. & Bazylinski, D. a. Magnetotaxis in Prokaryotes. *eLS* (2011). doi:10.1002/9780470015902.a0000397.pub2

14. Bellini, S. On a unique behavior of freshwater bacteria. *Chinese J. Oceanol. Limnol.* **27,** 3–5 (2009).

15. Cox, B. L. *et al.* Organization and Elemental Analysis of P-, S-, and Fe-rich Inclusions in a Population of Freshwater Magnetococci. *Geomicrobiol. J.* **19,** 387–406 (2002).

16. Bazylinski, D. a & Frankel, R. B. Magnetosome formation in prokaryotes. *Nat. Rev. Microbiol.* **2,** 217–230 (2004).

17. Fassbinder, J. W. E., Stanjekt, H. & Vali, H. Occurrence of magnetic bacteria in soil. *Nature* **343,** 161–163 (1990).

18. Jimenez-Lopez, C., Romanek, C. S. & Bazylinski, D. a. Magnetite as a prokaryotic biomarker: A review. *J. Geophys. Res.* **115,** G00G03 (2010).

19. Bazylinski, D. & Lefèvre, C. Magnetotactic Bacteria from Extreme Environments. *Life* **3,** 295–307 (2013).

20. Kodama, K. P., Moeller, R. E., Bazylinski, D. a., Kopp, R. E. & Chen, a. P. The mineral magnetic record of magnetofossils in recent lake sediments of Lake Ely, PA. *Glob. Planet. Change* **110,** 350–363 (2013).

21. Lefevre, C. T. *et al.* Insight into the Evolution of Magnetotaxis in Magnetospirillum spp., Based on mam Gene Phylogeny. *Appl. Environ. Microbiol.* **78,** 7238–7248 (2012).

22. Geelhoed, J. S., Kleerebezem, R., Sorokin, D. Y., Stams, A. J. M. & van Loosdrecht, M. C. M. Reduced inorganic sulfur oxidation supports autotrophic and mixotrophic growth of Magnetospirillum strain J10 and Magnetospirillum gryphiswaldense. *Environ. Microbiol.* **12,** 1031–1040 (2010).

23. Scott, K. M. *et al.* The genome of deep-sea vent chemolithoautotroph Thiomicrospira crunogena XCL-2. *PLoS Biol.* **4,** 2196–2212 (2006).

24. Thomas-Keprta, K. L. *et al.* Truncated hexa-octahedral magnetite crystals in ALH84001: presumptive biosignatures. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 2164–2169 (2001).

25. Frankel, R. B., Polytechnic, C. & Obispo, S. L. Magnetotaxis : Microbial. **4,** (2002).

26. Lefèvre, C. T. *et al.* Diversity of Magneto-Aerotactic Behaviors and Oxygen Sensing Mechanisms in Cultured Magnetotactic Bacteria. *Biophys. J.* **107,** 527–538 (2014).

27. Flies, C. B., Peplies, J. & Schüler, D. Combined Approach for Characterization of Uncultivated Magnetotactic Bacteria from Various Aquatic Environments Combined Approach for Characterization of Uncultivated Magnetotactic Bacteria from Various Aquatic Environments. *Appl Env. Microbiol* **71,** 2723–2731 (2005).

28.   Araujo, A., Abreu, F., Silva, K., Bazylinski, D. & Lins, U. *Magnetotactic Bacteria as Potential Sources of Bioproducts*. *Marine Drugs* **13,** (2015).

29.   Moskowitz, B. M., Bazylinski, D. a., Egli, R., Frankel, R. B. & Edwards, K. J. Magnetic properties of marine magnetotactic bacteria in a seasonally stratified coastal pond (Salt Pond, MA, USA). *Geophys. J. Int.* **174,** 75–92 (2008).

30.   Pósfai, M., Lefèvre, C. T., Trubitsyn, D., Bazylinski, D. a. & Frankel, R. B. Phylogenetic significance of composition and crystal morphology of magnetosome minerals. *Front. Microbiol.* **4,** 1–15 (2013).

31.   Simmons, S. L., Bazylinski, D. a. & Edwards, K. J. Population dynamics of marine magnetotactic bacteria in a meromictic salt pond described with qPCR. *Environ. Microbiol.* **9,** 2162–2174 (2007).

32.   Shively, J. M. *et al.* Bacterial and Archaeal Inclusions. *eLS* (2011). doi:10.1002/9780470015902.a0000302.pub3

33.   Lefèvre, C. T. *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.* **15,** n/a-n/a (2013).

34.   Bazylinski, D. a *et al.* Controlled Biomineralization of Magnetite (Fe(inf3)O(inf4)) and Greigite (Fe(inf3)S(inf4)) in a Magnetotactic Bacterium. *Appl. Environ. Microbiol.* **61,** 3232–3239 (1995).

35.   Bazylinski, D. a. Synthesis of the bacterial magnetosome: The making of a magnetic

personality. *Int. Microbiol.* **2,** 71–80 (1999).

36.   Bazylinski, D. a. *et al.* Chemolithoautotrophy in the marine, magnetotactic bacterial

          strains MV-1 and MV-2. *Arch. Microbiol.* **182,** 373–387 (2004).

37.   Abreu, F. *et al.* Deciphering unusual uncultured magnetotactic multicellular prokaryotes

          through genomics. *ISME J.* **8,** 1055–1068 (2014).

38.   Morillo, V. *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.* **5,** 1–12 (2014).

39.   Jogler, C. *et al.* Cultivation-independent characterization of 'Candidatus

          Magnetobacterium bavaricum' via ultrastructural, geochemical, ecological and

          metagenomic methods. *Environ. Microbiol.* **12,** 2466–2478 (2010).

40.   Schüler, D., Spring, S. & Bazylinski, D. a. Improved technique for the isolation of

          magnetotactic spirilla from a freshwater sediment and their phylogenetic

          characterization. *Syst. Appl. Microbiol.* **22,** 466–471 (1999).

41.   Schultheiss, D. & Schuler, D. Development of a genetic system for Magnetospirillum

          gryphiswaldense. *Arch Microbiol* **179,** 89–94 (2003).

42.   Schuler, D. Genetics and cell biology of magnetosome formation in magnetotactic

          bacteria. *Int. Microbiol.* **32,** 654–672 (2008).

43.   Bacteria, E. M. Polyextremophiles. **27,** 581–595 (2013).

44.   Prozorov, T., Bazylinski, D. a., Mallapragada, S. K. & Prozorov, R. Novel magnetic

nanomaterials inspired by magnetotactic bacteria: Topical review. *Mater. Sci. Eng. R Reports* **74,** 133–172 (2013).

45. Lefèvre, C. T., Frankel, R. B., Pósfai, M., Prozorov, T. & Bazylinski, D. a. Isolation of obligately alkaliphilic magnetotactic bacteria from extremely alkaline environments. *Environ. Microbiol.* **13,** 2342–2350 (2011).

46. Pikuta, E. V. Desulfonatronum thiodismutans sp. nov., a novel alkaliphilic, sulfate-reducing bacterium capable of lithoautotrophic growth. *Int. J. Syst. Evol. Microbiol.* **53,** 1327–1332 (2003).

47. Goltsman, D. S. A., Comolli, L. R., Thomas, B. C. & Banfield, J. F. Community transcriptomics reveals unexpected high microbial diversity in acidophilic biofilm communities. **9,** 1014–1023 (2014).

48. Sorokin, D. Y., Tourova, T. P., Muyzer, G. & Kuenen, G. J. Thiohalospira halophila gen. nov., sp. nov. and Thiohalospira alkaliphila sp. nov., novel obligately chemolithoautotrophic, halophilic, sulfur-oxidizing gammaproteobacteria from hypersaline habitats. *Int. J. Syst. Evol. Microbiol.* **58,** 1685–1692 (2008).

49. Abreu, F. *et al.* 'Candidatus magnetoglobus multicellularis', a multicellular, magnetotactic prokaryote from a hypersaline environment. *Int. J. Syst. Evol. Microbiol.* **57,** 1318–1322 (2007).

50. Lefèvre, C. T. *et al.* A cultured greigite-producing magnetotactic bacterium in a novel group of sulfate-reducing bacteria. *Science* **334,** 1720–173 (2011).

51. Lefevre, C. T. *et al.* Moderately Thermophilic Magnetotactic Bacteria from Hot Springs in Nevada. *Appl. Environ. Microbiol.* **76,** 3740–3743 (2010).

52. Abreu, F. *et al.* Culture-independent characterization of novel psychrophilic magnetotactic cocci from Antarctic marine sediments. *Environ. Microbiol.* **18,** 4426–4441 (2016).

53. Vali, H., Förster, O., Amarantidis, G. & Petersen, N. Magnetotactic bacteria and their magnetofossils in sediments. *Earth Planet. Sci. Lett.* **86,** 389–400 (1987).

54. Kcaa, C. H andbook of Biom i neralization A l s p e c t s F o rma ti o n Biol ogic a a n d S tru ctu re Edited by Bauerlein Verlag.

55. Edwards, K. J. & Bazylinski, D. a. Intracellular minerals and metal deposits in prokaryotes. *Geobiology* **6,** 309–317 (2008).

56. Schubbe, S. *et al.* Complete Genome Sequence of the Chemolithoautotrophic Marine Magnetotactic Coccus Strain MC-1. *Appl. Environ. Microbiol.* **75,** 4835–4852 (2009).

57. Lohmann, K. J., Putman, N. F. & Lohmann, C. M. F. Geomagnetic imprinting: A unifying hypothesis of long-distance natal homing in salmon and sea turtles. *Proc. Natl. Acad. Sci.* **105,** 19096–19101 (2008).

58. Frankel, R. B. & Bazylinski, D. a. Magnetosomes and magneto-aerotaxis. *Bact. Sens. Signal.* **16,** 182–193 (2009).

59. Frankel, R. B., Bazylinski, D. a, Johnson, M. S. & Taylor, B. L. Magneto-aerotaxis in marine coccoid bacteria. *Biophys. J.* **73,** 994–1000 (1997).

60. Wadhams, G. H. & Armitage, J. P. Making Sense of it All: Bacterial Chemotaxis. *Nat. Rev. Mol. Cell Biol.* **5,** 1024–1037 (2004).

61. Uebe, R. & Schüler, D. Magnetosome biogenesis in magnetotactic bacteria. *Nat. Publ. Gr.* **14,** 621–637 (2016).

62. Blakemore, R. P. Magn'etotactic bacteria. 217–238 (1982).

63. Wolfe, R. A 'capillary racetrack' method for isolation of magnetotactic bacteria. *FEMS Microbiol. Lett.* **45,** 31–35 (1987).

64. Blakemore, R. P., Maratea, D. & Wolfe, R. S. Isolation and pure culture of a freshwater magnetic spirillum in chemically defined medium. *J. Bacteriol.* **140,** 720–729 (1979).

65. Schu, D. 4 Biochemical and Genetic Analysis of the Magnetosome Membrane in Magnetospirillum gryphiswaldense. **70,** 1040–1050 (2004).

66. Nudelman, H. & Zarivach, R. Structure prediction of magnetosome-associated proteins. *Front. Microbiol.* **5,** 1–17 (2014).

67. Scheffel, A., Gärdes, A., Grünberg, K., Wanner, G. & Schüler, D. The major magnetosome proteins MamGFDC are not essential for magnetite biomineralization in Magnetospirillum gryphiswaldense but regulate the size of magnetosome crystals. *J. Bacteriol.* **190,** 377–386 (2008).

68. Posfai, M. Sulfides in Biosystems. *Rev. Mineral. Geochemistry* **61,** 679–714 (2006).

69. Pósfai, M., Buseck, P. R., Bazylinski, D. A. & Frankel, R. B. Iron sulfides from magnetotactic bacteria: Structure, composition, and phase transitions. *Am. Mineral.* **83,** 1469–1481

(1998).

70.   Misiorny, M., Hell, M. & Wegewijs, M. R. Spintronic magnetic anisotropy. *Nat. Phys.* **9,**

       801–805 (2013).

71.   Papaefthymiou, G. C. Nanoparticle magnetism. *Nano Today* **4,** 438–447 (2009).

72.   Pósfai, M. *et al.* Properties of intracellular magnetite crystals produced by Desulfovibrio

       magneticus strain RS-1. *Earth Planet. Sci. Lett.* **249,** 444–455 (2006).

73.   Heywood, B. R., Mann, S. & Frankel, R. B. Structure , Morphology and Growth of

       Biogenic. *Mater. Res. Soc.* **218,** 93–108 (1991).

74.   Frankel, R. B. & Bazylinski, D. a. How magnetotactic bacteria make magnetosomes queue

       up. *Trends Microbiol.* **14,** 329–331 (2006).

75.   Frankel, R. B. & Bazylinski, D. a. Magnetosome mysteries. *ASM News* **70,** 176–183 (2004).

76.   McCartney, M. R., Lins, U., Farina, M., Buseck, P. R. & Frankel, R. B. Magnetic

       microstructure of bacterial magnetite by electron holography. *Eur. J. Mineral.* **13,** 685–

       689 (2001).

77.   Dunin-borkowski, A. R. E. *et al.* Magnetic Microstructure of Magnetotactic Bacteria by

       Electron Holography Published by : American Association for the Advancement of

       Science Stable URL : http://www.jstor.org/stable/2897012 Linked references are

       available on JSTOR for this article : Magnet. **282,** 1868–1870 (2017).

78.   Kalirai, S. S., Bazylinski, D. a. & Hitchcock, A. P. Anomalous Magnetic Orientations of

       Magnetosome Chains in a Magnetotactic Bacterium: Magnetovibrio blakemorei Strain

MV-1. *PLoS One* **8,** e53368 (2013).

79.   Rodgers, F. *et al.* Intercellular structure in a many-celled magnetotactic prokaryote. *Arch. Microbiol.* **154,** 18–22 (1990).

80.   Wenter, R., Wanner, G., Schüler, D. & Overmann, J. Ultrastructure, tactic behaviour and potential for sulfate reduction of a novel multicellular magnetotactic prokaryote from North Sea sediments. *Environ. Microbiol.* **11,** 1493–1505 (2009).

81.   Devouard, B. *et al.* Magnetite from magnetotactic bacteria: Size distributions and twinning. *Am. Mineral.* **83,** 1387–1398 (1998).

82.   Lohße, A. *et al.* Functional Analysis of the Magnetosome Island in Magnetospirillum gryphiswaldense: The mamAB Operon Is Sufficient for Magnetite Biomineralization. *PLoS One* **6,** e25561 (2011).

83.   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* **107,** 5593–5598 (2010).

84.   Katzmann, E., Scheffel, A., Gruska, M., Plitzko, J. M. & Sch??ler, D. Loss of the actin-like protein MamK has pleiotropic effects on magnetosome formation and chain assembly in Magnetospirillum gryphiswaldense. *Mol. Microbiol.* **77,** 208–224 (2010).

85.   Scheffel, A. *et al.* An acidic protein aligns magnetosomes along a filamentous structure in magnetotactic bacteria. *Nature* **440,** 110–114 (2006).

86.   Abreu, F. *et al.* Common ancestry of iron oxide- and iron-sulfide-based biomineralization

in magnetotactic bacteria. *ISME J.* **5,** 1634–1640 (2011).

87.   Komeili, A., Vali, H., Beveridge, T. J. & Newman, D. K. Magnetosome vesicles are present before magnetite formation, and MamA is required for their activation. *Proc Natl Acad Sci U S A* **101,** 3839–3844 (2004).

88.   Zeytuni, N. *et al.* Self-recognition mechanism of MamA, a magnetosome-associated TPR-containing protein, promotes complex assembly. *Pnas* **108,** 480–487 (2011).

89.   Cartron, M. L., Maddocks, S., Gillingham, P., Craven, C. J. & Andrews, S. C. Feo - Transport of ferrous iron into bacteria. *BioMetals* **19,** 143–157 (2006).

90.   Richter, M. *et al.* Comparative Genome Analysis of Four Magnetotactic Bacteria Reveals a Complex Set of Group-Specific Genes Implicated in Magnetosome Biomineralization and Function. *J. Bacteriol.* **189,** 4899–4910 (2007).

91.   Schübbe, S. *et al.* Transcriptional organization and regulation of magnetosome operons in Magnetospirillum gryphiswaldense. *Appl. Environ. Microbiol.* **72,** 5757–5765 (2006).

92.   Dubbels, B. L. Evidence for a copper-dependent iron transport system in the marine, magnetotactic bacterium strain MV-1. *Microbiology* **150,** 2931–2945 (2004).

93.   Schu, S. *et al.* Characterization of a Spontaneous Nonmagnetic Mutant of Magnetospirillum gryphiswaldense Reveals a Large Deletion Comprising a Putative Magnetosome Island. **185,** 5779–5790 (2003).

94.   Tanaka, M., Mazuyama, E., Arakaki, A. & Matsunaga, T. MMS6 protein regulates crystal morphology during nano-sized magnetite biomineralization in vivo. *J. Biol. Chem.* **286,**

6386–6392 (2011).

95.   Ullrich, S. *et al.* A Hypervariable 130-Kilobase Genomic Region of. *Microbiology* **187,**

7176–7184 (2005).

96.   Blatch, G. L. & Lässle, M. The tetratricopeptide repeat: A structural motif mediating

protein-protein interactions. *BioEssays* **21,** 932–939 (1999).

97.   Wawer, C., Tebo, B. M., Gru, K. & Schu, D. A Large Gene Cluster Encoding Several

Magnetosome Proteins Is Conserved in Different Species of Magnetotactic Bacteria. **67,**

4573–4582 (2001).

98.   Sakaguchi, T., Arakaki, A. & Matsunaga, T. Desulfovibrio magneticus sp. nov., a novel

sulfate-reducing bacterium that produces intracellular single-domain-sized magnetite

particles. *Int. J. Syst. Evol. Microbiol.* **52,** 215–221 (2002).

99.   Jogler, C. *et al.* Toward cloning of the magnetotactic metagenome: identification of

magnetosome island gene clusters in uncultivated magnetotactic bacteria from different

aquatic sediments. *Appl. Environ. Microbiol.* **75,** 3972–3979 (2009).

100.   Lefèvre, C. T. *et al.* Monophyletic origin of magnetotaxis and the first magnetosomes.

Environ. Microbiol. **15,** 2267–2274 (2013).

101.   Bazylinski, D. a & Blakemore, R. P. Denitrification and Assimilatory Nitrate Reduction in

Aquaspirillum magnetotacticum. *Appl. Environ. Microbiol.* **46,** 1118–24 (1983).

102.   Blakemore, R. P., Short, K. a., Bazylinski, D. a., Rosenblatt, C. & Frankel, R. B.

Microaerobic conditions are required for magnetite formation within *aquaspirillum*

magnetotacticum. Geomicrobiol. J. **4,** 53–71 (1985).

103. Spring, S. *et al.* Phylogenetic affiliation and ultrastructure of uncultured magnetic bacteria with unusually large magnetosomes. *Arch. Microbiol.* **169,** 136–147 (1998).

104. Bazylinski, D. a. *et al.* Magnetococcus marinus gen. nov., sp. nov., a marine, magnetotactic bacterium that represents a novel lineage (Magnetococcaceae fam. nov., Magnetococcales ord. nov.) at the base of the Alphaproteobacteria. *Int. J. Syst. Evol. Microbiol.* **63,** 801–808 (2013).

105. Williams, T. J., Zhang, C. L., Scott, J. H. & Bazylinski, D. a. Evidence for autotrophy via the reverse tricarboxylic acid cycle in the marine magnetotactic coccus strain MC-1. *Appl. Environ. Microbiol.* **72,** 1322–1329 (2006).

106. Burgess, J. G., Kawaguchi, R., Sakaguchi, T., Thornhill, R. H. & Matsunaga, T. Evolutionary relationships among Magnetospirillum strains inferred from phylogenetic analysis of 16S rDNA sequences. *J. Bacteriol.* **175,** 6689–6694 (1993).

107. Schleifer, K. H. *et al.* The Genus Magnetospirillum gen. nov. Description of Magnetospirillum gryphiswaldense sp. nov. and Transfer of Aquaspirillum magnetotacticum to Magnetospirillum magnetotacticum comb. nov. *Syst. Appl. Microbiol.* **14,** 379–385 (1991).

108. Bazylinski, D. a. & Blakemore, R. P. Nitrogen fixation (acetylene reduction) in Aquaspirillum magnetotacticum. *Curr. Microbiol.* **9,** 305–308 (1983).

109. Bazylinski, D. a, Dean, a J., Schüler, D., Phillips, E. J. & Lovley, D. R. N2-dependent growth

and nitrogenase activity in the metal-metabolizing bacteria, Geobacter and Magnetospirillum species. *Environ. Microbiol.* **2,** 266–73 (2000).

110. Feller, U., Anders, I. & Mae, T. Rubiscolytics: Fate of Rubisco after its enzymatic function in a cell is terminated. *J. Exp. Bot.* **59,** 1615–1624 (2008).

111. Bazylinski, D. a. *et al.* Magnetovibrio blakemorei gen. nov., sp. nov., a magnetotactic bacterium (Alphaproteobacteria: Rhodospirillaceae) isolated from a salt marsh. *Int. J. Syst. Evol. Microbiol.* **63,** 1824–1833 (2013).

112. Lefèvre, C. T. *et al.* Morphological features of elongated-anisotropic magnetosome crystals in magnetotactic bacteria of the Nitrospirae phylum and the Deltaproteobacteria class. *Earth Planet. Sci. Lett.* **312,** 194–200 (2011).

113. Zhou, K. *et al.* A novel genus of multicellular magnetotactic prokaryotes from the Yellow Sea. *Environ. Microbiol.* **14,** 405–413 (2012).

114. Hanzlik, M., Winklhofer, M. & Petersen, N. Spatial arrangement of chains of magnetosomes in magnetotactic bacteria. *Earth Planet. Sci. Lett.* **145,** 125–134 (1996).

115. Lefèvre, C. T., Frankel, R. B., Abreu, F., Lins, U. & Bazylinski, D. a. Culture-independent characterization of a novel, uncultivated magnetotactic member of the Nitrospirae phylum. *Environ. Microbiol.* **13,** 538–549 (2011).

116. Lin, W., Li, J. & Pan, Y. Newly isolated but uncultivated magnetotactic bacterium of the phylum Nitrospirae from Beijing, China. *Appl. Environ. Microbiol.* **78,** 668–675 (2012).

117. Isambert, A., Menguy, N., Larquet, E., Guyot, F. & Valet, J. P. Transmission electron

microscopy study of magnetites in a freshwater population of magnetotactic bacteria. *Am. Mineral.* **92,** 621–630 (2007).

118. Li, J. *et al.* Biomineralization, crystallography and magnetic properties of bullet-shaped magnetite magnetosomes in giant rod magnetotactic bacteria. *Earth Planet. Sci. Lett.* **293,** 368–376 (2010).

119. Jogler, C. *et al.* Comparative analysis of magnetosome gene clusters in magnetotactic bacteria provides further evidence for horizontal gene transfer. *Environ. Microbiol.* **11,** 1267–1277 (2009).

120. Kolinko, S. *et al.* Single-cell analysis reveals a novel uncultivated magnetotactic bacterium within the candidate division OP3. **14,** 1709–1721 (2012).

121. Li, Y., Katzmann, E., Borg, S. & Schüler, D. The periplasmic nitrate reductase Nap is required for anaerobic growth and involved in redox control of magnetite biomineralization in Magnetospirillum gryphiswaldense. *J. Bacteriol.* **194,** 4847–4856 (2012).

122. Rioux, J. B. *et al.* A second actin-like mamk protein in magnetospirillum magneticum AMB-1 encoded outside the genomic magnetosome island. *PLoS One* **5,** e9151 (2010).

123. Biological, O. F. Formation of methane. *J. Franklin Inst.* **176,** 737 (1913).

124. Bazylinski, D. a, Wirsen, C. O. & Jannasch, H. W. Microbial Utilization of Naturally Occurring Hydrocarbons at the Guaymas Basin Hydrothermal Vent Site. *Appl. Environ. Microbiol.* **55,** 2832–2836 (1989).

125. Zumbo, P. Phenol - chloroform Extraction. *Weill Cornell Med. Coll.* 1–7 (1979). doi:10.1038/nprot.2006.83.

126. Eid, J. *et al.* Real-time DNA sequencing from single polymerase molecules. *Science* **323,** 133–138 (2009).

127. Liao, Y., Lin, S. & Lin, H. assemblies : strategy and performance comparisons. (2015). doi:10.1038/srep08747

128. Scott, D. & Ely, B. Comparison of Genome Sequencing Technology and Assembly Methods for the Analysis of a GC-Rich Bacterial Genome. 338–344 (2015). doi:10.1007/s00284-014-0721-6

129. Bankevich, A. *et al.* and Its Applications to Single-Cell Sequencing. **19,** 455–477 (2012).

130. Zerbino, D. R. & Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18,** 821–9 (2008).

131. Almeida, L. G. P. *et al.* A System for Automated Bacterial (genome) Integrated Annotation--SABIA. *Bioinformatics* **20,** 2832–3 (2004).

132. Vallenet, D. *et al.* MicroScope--an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. *Nucleic Acids Res.* **41,** D636–D647 (2013).

133. Aziz, R. K. *et al.* The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9,** 75 (2008).

134. Overbeek, R. The Subsystems Approach to Genome Annotation and its Use in the Project

to Annotate 1000 Genomes. *Nucleic Acids Res.* **33,** 5691–5702 (2005).

135. Promoters, E., Convolutional, U., Learning, D. & Networks, N. V . Solovyev , A Salamov (

2011) Automatic Annotation of Microbial Genomes and Metagenomic Sequences . In

Metagenomics and its Applications in Agriculture , Biomedicine and Environmenta ...

(2016).

136. Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J. & Wishart, D. S. PHAST : A Fast Phage Search

Tool. **39,** 347–352 (2017).

137. Grant, J. R. & Stothard, P. The CGView Server: a comparative genomics tool for circular

genomes. *Nucleic Acids Res.* **36,** 181–184 (2008).

138. Untergasser, A. *et al.* Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40,** 1–12

(2012).

139. Ye, J. *et al.* Primer-BLAST: A tool to design target-specific primers for polymerase chain

reaction. *BMC Bioinformatics* **13,** 134 (2012).

140. Owczarzy, R., Moreira, B. G., You, Y., Behlke, M. A. & Walder, J. A. Predicting stability of

DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry*

47, 5336–53 (2008).

141. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software

platform for the organization and analysis of sequence data. *Bioinformatics* **28,** 1647–9

(2012).

142. Darling, A. C. E., Mau, B., Blattner, F. R. & Perna, N. T. Mauve: multiple alignment of

conserved genomic sequence with rearrangements. *Genome Res.* **14,** 1394–403 (2004).

143. Sorokin, D. Y. *et al.* nov . and Thioalkalimicrobium sibericum sp . nov ., and Thioalkalivibrio versutus gen . nov ., sp . nov ., Thioalkalivibrio nitratis sp . nov . and Thioalkalivibrio denitrificans sp . nov ., novel obligately alkaliphilic and obligately chemolithoautotroph. 565–580 (2017).

144. Xu, W. *et al.* A Novel Universal Primer-Multiplex-PCR Method with Sequencing Gel Electrophoresis Analysis. *PLoS One* **7,** e22900 (2012).

145. Lee, S. H. *et al.* A multiplex real-time PCR for differential detection and quantification of Salmonella spp., Salmonella enterica serovar Typhimurium and Enteritidis in meats. *J. Vet. Sci.* **10,** 43 (2009).

146. Zhang, Y. *et al.* A novel real-time quantitative PCR method using attached universal template probe. *Nucleic Acids Res* **31,** e123 (2003).

147. Chang, C.-T. *et al.* Mixed Sequence Reader: A Program for Analyzing DNA Sequences with Heterozygous Base Calling. *Sci. World J.* **2012,** 1–10 (2012).

148. Riley, J. *et al.* A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.* **18,** 2887–2890 (1990).

149. Devon, R. S., Porteous, D. J. & Brookes, a J. Splinkerettes--improved vectorettes for greater efficiency in PCR walking. *Nucleic Acids Res.* **23,** 1644–1645 (1995).

150. Priyanka, A. & Nazar, R. Benchmarks 10. *Biotechniques* **29,** 1176–1180 (2000).

151. linear ss pcr.pdf.

152. Rae, B. D., Long, B. M., Badger, M. R. & Price, G. D. Functions, Compositions, and Evolution of the Two Types of Carboxysomes: Polyhedral Microcompartments That Facilitate CO2 Fixation in Cyanobacteria and Some Proteobacteria. *Microbiol. Mol. Biol. Rev.* **77,** 357–379 (2013).

153. Cannon, G. C. *et al.* Organization of carboxysome genes in the thiobacilli. *Curr. Microbiol.* **46,** 115–119 (2003).

154. Shively, J. M. *et al.* Bacterial Inclusions. *eLS* (2006). doi:10.1038/npg.els.0004268

155. Abt, G. *et al.* A challenging interpretation of a hexagonally layered protein structure. *Plant Physiol.* **18,** 784–93 (2013).

156. Omelchenko, M., Galperin, M., Wolf, Y. & Koonin, E. Non-homologous isofunctional enzymes: A systematic analysis of alternative solutions in enzyme evolution. *Biol. Direct* **5,** 31 (2010).

157. Sun, J. B. *et al.* High-yield growth and magnetosome formation by Magnetospirillum gryphiswaldense MSR-1 in an oxygen-controlled fermentor supplied solely with air. *Appl. Microbiol. Biotechnol.* **79,** 389–397 (2008).

158. Letoffe, S., Heuck, G., Delepelaire, P., Lange, N. & Wandersman, C. Bacteria capture iron from heme by keeping tetrapyrrol skeleton intact. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 11719–11724 (2009).

159. Andrews, S. C. The Ferritin-like superfamily: Evolution of the biological iron storeman from a rubrerythrin-like ancestor. *Biochim. Biophys. Acta* **1800,** 691–705 (2010).

160. Rivera, M. Bacterioferritin: Structure, Dynamics, and Protein – Protein Interactions at Play in Iron Storage and Mobilization. (2017). doi:10.1021/acs.accounts.6b00514

161. Bazylinski, D. a., Palome, E., Blakemore, N. a. & Blakemore, R. P. Denitrification by Chromobacterium violaceum. *Appl. Environ. Microbiol.* **52,** 696–699 (1986).

162. Friedrich, C. G., Rother, D., Bardischewsky, F., Ouentmeier, A. & Fischer, J. Oxidation of Reduced Inorganic Sulfur Compounds by Bacteria: Emergence of a Common Mechanism? *Appl. Environ. Microbiol.* **67,** 2873–2882 (2001).

163. Grein, F., Pereira, I. A. C. & Dahl, C. Biochemical characterization of individual components of the Allochromatium vinosum DsrMKJOP transmembrane complex aids understanding of complex function in vivo. *J. Bacteriol.* **192,** 6369–6377 (2010).

164. Loschi, L. *et al.* Structural and biochemical identification of a novel bacterial oxidoreductase. *J. Biol. Chem.* **279,** 50391–400 (2004).

165. Wanner, B. L. Gene regulation by phosphate in enteric bacteria. *J. Cell. Biochem.* **51,** 47–54 (1993).

166. Kocan, M. *et al.* Two-component systems of *Corynebacterium glutamicum*: deletion analysis and involvement of the PhoS-PhoR system in the phosphate starvation response. *J. Bacteriol.* **188,** 724–32 (2006).

167. Kazakov, A. E., Vassieva, O., Gelfand, M. S., Osterman, A. & Overbeek, R. Bioinformatics classification and functional analysis of PhoH homologs. *In Silico Biol.* **3,** 3–15 (2003).

168. Zhang, H., Ishige, K. & Kornberg, A. A polyphosphate kinase (PPK2) widely conserved in

bacteria. (2002).

169. Ken, Q. & Pauisers, J. T. Comparative analyses of fundamental differences in membrane transport Capabilities in Prokaryotes and Eukaryotes. *PLoS Comput. Biol.* **1,** 0190–0201 (2005).

170. Römling, U., Gomelsky, M. & Galperin, M. Y. MicroReview C-di-GMP : the dawning of a novel bacterial signalling system. **57,** 629–639 (2005).

171. Saffarini, D. *Bacteria- Metal Interactions*.

172. Merrick, M. J. MicroReview In a class of its own — the RNA polymerase sigma factor ^. **10,** 903–909 (1993).

173. Weinitschke, S., Denger, K., Cook, A. M. & Smits, T. H. M. The DUF81 protein TauE in Cupriavidus necator H16 , a sulfite exporter in the metabolism of C 2 sulfonates. 3055–3060 (2017). doi:10.1099/mic.0.2007/009845-0

174. Ji, B. *et al.* Comparative genomic analysis provides insights into the evolution and niche adaptation of marine Magnetospira sp. QH-2 strain. *Environ. Microbiol.* **16,** 525–544 (2014).

175. Gao, F. & Zhang, C.-T. Ori-Finder: a web-based system for finding oriCs in unannotated bacterial genomes. *BMC Bioinformatics* **9,** 79 (2008).

176. Words, K. & Mcclintock, B. AND T RANSPOSABLE E LEMENTS. *Annu. Rev. Genet.* 389–410 (2002). doi:10.1146/annurev.genet.36.040202.092802

177. Murray, N. E. Type I Restriction Systems : Sophisticated Molecular Machines (a Legacy of

Bertani and Weigle). **64,** 412–434 (2000).

178. Self-immunity, B. *et al.* The Abi Proteins and Their Involvement in. **192,** 2068–2076

(2010).

179. Nakano, M. M., Hoffmann, T., Zhu, Y. I. & Jahn, D. Nitrogen and Oxygen Regulation of

Bacillus subtilis nasDEF Encoding NADH-Dependent Nitrite Reductase by TnrA and ResDE.

Am. Soc. Microbiol. **180,** 5344–5350 (1998).

180. Cabello, N., Marti, M., Moreno-vivia, C., Blasco, R. & Castillo, F. MINIREVIEW Prokaryotic

Nitrate Reduction : Molecular Properties and Functional Distinction among Bacterial

Nitrate Reductases. **181,** 6573–6584 (1999).

181. Al-Mamun, A. & Baawain, M. S. Accumulation of intermediate denitrifying compounds

inhibiting biological denitrification on cathode in Microbial Fuel Cell. *J. Environ. Heal. Sci.

Eng.* **13,** 81 (2015).

182. Mulligan, C., Fischer, M. & Thomas, G. H. Tripartite ATP-independent periplasmic (TRAP)

transporters in bacteria and archaea. *FEMS Microbiol. Rev.* **35,** 68–86 (2011).

183. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high

throughput. *Nucleic Acids Res.* **32,** 1792–1797 (2004).

184. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis

version 7.0 for bigger datasets. *Mol. Biol. Evol.* msw054 (2016).

doi:10.1093/molbev/msw054

185. Peter, N. D. Nitrate reduction and nitrogenase activity in Spirillum lipoferum '. (1977).

186. Schultheiss, D., Jendrossek, D., Hanzlik, M. & Schu, D. The Presumptive Magnetosome Protein Mms16 Is a Poly (3-Hydroxybutyrate) Granule-Bound Protein (Phasin) in Magnetospirillum gryphiswaldense. **187,** 2416–2425 (2005).

187. Osterman, A. & Overbeek, R. Missing genes in metabolic pathways: A comparative genomics approach. *Curr. Opin. Chem. Biol.* **7,** 238–251 (2003).

Curriculum Vitae

University of Nevada, Las Vegas	Email: CoreyGeurink@gmail.com
4505 S. Maryland Pkwy.	
Las Vegas, NV 89154	https://www.linkedin.com/in/coreygeurink

EDUCATION:

Ph.D. Cell and Molecular Biology, University of Nevada Las Vegas   2011-2017
   "Genomic Characterization of Magnetic *Gammaproteobacteria* Strain BW-2"

B.S. Biology,   Grand Valley State University   1998-2000

WORK EXPERIENCE:

Research Assistant, University of Nevada Las Vegas   2011-2017
   - Genomic characterization and molecular mechanisms of magnetosome formation in bacteria
   - Next Gen Sequencing and analysis
   - Genome annotation, scaffolding, finishing
   - Teaching and development of undergraduate courses (Biology, Anatomy and Physiology)
   - Lab manager (logistics, purchasing, 14 undergraduate students)
   - Research Experiences for Undergraduates (REU) Mentor
   - Preparation of manuscripts, presentations at meetings, guest lectures
   Supervisor: Prof. Dennis Bazylinski, Las Vegas, NV; Phone: 702-895-2053

Part Time Instructor, Nevada State College   2015-present
   - BIOL 189 Fundamentals of Life Science (lecture and lab)

Mental Health Technician, State of Nevada – Desert Willow Treatment Centre   2003-2011
   - Training and supervising new hires
   - Teaching, leading patient group therapy
   - HIPPA, patient rights
   - Documentation, chart audits
   - Physical management and de-escalation techniques

- Infection control procedures
- Behavior modification techniques

Supervisor: Linda Boyd, RN, Las Vegas, NV; Phone: 702-489-8900
   Reason for leaving: UNLV PhD program

Commercial fisherman, Fields and Sons, Kodiak, AK   2002-2002
- All duties related to operation including daily interaction with , handling of all varieties of ocean life

 Supervisor: Wallace Fields, Kodiak, AK; Phone: 907-486-8370
   Reason for leaving: Salmon season ended

QA In Process Chemistry Lab, Mead Johnson Pharmaceuticals   2001-2002
- Chemical analysis of raw materials, in process and finished product

   Supervisor: Marcia Kuipers, Zeeland, Mi; Phone: 616-748 - 7100
   Reason for leaving: Temporary employment

SCIENTIFIC SKILLS:

General chemical and microbiological techniques:
- Isolation from various environments, enrichment, identification, culturing of fastidious microaerophilic and anaerobic bacteria
- Transformation, conjugation, cloning
- Preparation of standard, microaerophilic and anaerobic custom medium

Molecular techniques:
- DNA isolation, purification, quantification, gel electrophoresis,
- Extensive PCR techniques (including multiplex, universal primer/dual universal primer-multiplex, vectorette, single stranded, asymmetrical, transposome mediated dual universal primer multiplex and long PCR)
- Cloning, transformation
- Restriction, ligation

Genome sequencing:
- Isolation of bacterial DNA from difficult samples, sequencing and data analysis (Sanger, Illumina, 454, PacBio)
- Genome assembly and scaffolding
- Automated and manual annotation of bacterial genomes
- Bioinformatic interpretation of metabolic pathways

Microscopy techniques:
- Light, confocal, fluorescence

Laboratory management:
- Procurement, implementing Health and Safety policies
- Management of 14 undergraduate research assistants
- Hazardous waste disposal
- Lab equipment setup, calibration, maintenance

Bioinformatics Skills:
- Sequence analysis software (CLC Genomics Workbench, Artemis, Geneious, Clustal, MAUVE, BLAST)
- Genome assembly, scaffolding and annotation software (CLC Genomics Workbench, Geneious, SPAdes, Velvet, BLAST, RAST, MaGe, Artemis, OrthoVenn, DNA plotter, CG view, etc.)

TEACHING EXPERIENCE:

Part Time Instructor, Nevada State College   Fall 2015-present
- BIOL 189 Fundamentals of Life Science (7 sections lecture and lab)

Lab Instructor, University of Nevada – Las Vegas   Sum 2011-summer 2016

- BIOL 189 (14 sections)

- 224 Anatomy and Physiology (6 sections)

Guest lecturer, University of Nevada – Las Vegas   Spring 2013 and 2014

- BIOL 418/618 Microbial Ecology

REU Mentor, University of Nevada – Las Vegas   Sum 2014

- Microbiology Research Experience for Undergraduates

Guest lecturer, University of Nevada – Las Vegas   Spring 2013

- BIOL 478/730 Cancer Cell Biology

Conference Speaker, American Society for Microbiology   Spring 2014

 "Genomic Characterization of Magnetic *Gammaproteobacteria* Strain BW-2"

Graduate Seminar, University of Nevada – Las Vegas   Fall 2011-Spr 2013

- Magnetotactic bacteria
- Molecular evolution
- Genomics
- Alternative splicing
- Autophagy
- Epigenetics
- Neurodegeneration
- Cell physiology of hibernation and caloric restriction

PUBLICATIONS:

Culture-independent characterization of novel psychrophilic magnetotactic cocci from Antarctic marine sediments (2016) Abreu F, Araujo AC, Leão P, Silva KT, Marques de Carvalho F, de Lima Cunha O, Almeida LG, Geurink C, Farina M, Rodelli D, Jovane L, Pellizari VH, de Vasconcelos AT, Bazylinski DA, Lins U Environ Microbiol. (2016) doi: 10.1111/1462-2920.13388

Draft Genome Sequence of the Obligately Alkaliphilic Sulfate-Reducing Bacterium Desulfonatronum thiodismutans Strain MLF1 (2014) Trubitsyn D, Geurink C, Pikuta E, Lefèvre CT, McShan WM, Gillaspy AF, Bazylinski DA. Genome announcements, doi: 10.1128/genomeA.00741-14

REFERENCES:

Dennis Bazylisnki, Ph.D.
School of Life Sciences
University of Nevada, Las Vegas

4505 S. Maryland Parkway
Las Vegas, NV 89154-4004
USA
Email: dennis.bazylinski@unlv.edu

Denis Trubitsyn, Ph.D.
Department of Biological Sciences
Southwestern Oklahoma State University
100 Campus Dr.
Weatherford, OK 73096
USA
Email: denis.trubitsyn@swosu.edu

Sean Niesenwenter, Ph.D.
School of Life Sciences

University of Nevada, Las Vegas
4505 S. Maryland Parkway
Las Vegas, NV 89154-4004
USA
Email: neiswent@unlv.nevada.edu

Viviana Morrillo-Lopez, Ph.D.
School of Life Sciences
University of Nevada, Las Vegas
4505 S. Maryland Parkway
Las Vegas, NV 89154-4004
USA
Email: viviana.morillolopez@unlv.edu

Nathan Silva
Nevada State College
1125 Nevada State Dr
Henderson, NV 89002
LAS-264
Email: Nathan.Silva@nsc.edu