Postprint of: Critical Reviews in Microbiology 42(1): 31-45 (2016)

Title:

Inferring pathways leading to organic-sulfur mineralization in the Bacillales

Margarida M. Santana+, Juan M. Gonzalez* and M. I. Clara+

+ Instituto de Ciências Agrárias e Ambientais Mediterrânicas (ICAAM), Universidade de Évora, Núcleo da Mitra, 7002-774, Évora, Portugal

*Instituto de Recursos Naturales y Agrobiología, IRNAS-CSIC, Av. Reina Mercedes, 10, E-41012, Sevilla, Spain

Corresponding author:

Margarida M. Santana, PhD

Instituto de Ciências Agrárias e Ambientais Mediterrânicas (ICAAM), Universidade de Évora, Núcleo da Mitra 7002-774, Évora, Portugal

Tel. + 351 266 760851; Fax + 351 266 760913 E-mail: msantana@uevora.pt

Key terms: enzymes of cysteine and methionine pathways; sulfite oxidase; aspartate transaminase; sulfite reductase; dissimilatory organic-sulfur oxidation

Abstract

Microbial organic sulfur mineralization to sulfate in terrestrial systems is poorly understood. The process is often missing in published sulfur cycle models. Studies on microbial sulfur cycling have been mostly centered on transformations of inorganic sulfur, mainly on sulfate-reducing and inorganic sulfur-oxidizing bacteria. Nevertheless, organic sulfur constitutes most sulfur in soils. Recent reports demonstrate that the mobilization of organic-bound-sulfur as sulfate in terrestrial environments occurs preferentially under high temperatures and thermophilic Firmicutes bacteria play a major role in the process, carrying out dissimilative organic-sulfur oxidation. So far, the determinant metabolic reactions of such activity have not been evaluated. Here, *in silico* analysis was performed on the genomes of sulfate-producing thermophilic genera and mesophilic low-sulfate producers, revealing that highest sulfate production is related to the simultaneous presence of metabolic pathways leading to sulfite synthesis, similar to the ones found in mammalian cells. Those pathways include reverse transsulfuration reactions (tightly associated with methionine cycling), and the presence of aspartate aminotransferases with the potential of 3-sulfinoalanine aminotransferase and cysteine aminotransferase activity, which ultimately leads to sulfite production. Sulfite is oxidized to sulfate by sulfite oxidase, this enzyme is determinant in sulfate synthesis, and it is absent in many mesophiles.

Introduction

Sulfur concentration varies between 100 and 1500 ppm in soils of temperate regions (Ehrlich, 1996). Organic sulfur represents more than 90% of total sulfur (Schlesinger 1997; Anandham et al. 2008) and it is present, for instance, in the form of the amino acids cysteine and methionine and their oxidation products, besides other cell components and their derivatives in natural environments.

In plants, sulfur is absorbed in the form of sulfate, therefore organic sulfur mineralization to sulfate is essential for plant growth. Earlier experiments carried out at 20-25ºC reported organic-sulfur mineralization to be highly limited (Ghani et al. 1993; Eriksen, 1996) but recently, it was shown that the mobilization of organic-bound-sulfur as sulfate in terrestrial environments occurs preferentially under high temperatures, and thermophilic bacteria, which have been suggested to be present in all soils (Marchant et al., 2002) play a major role in the process (Portillo et al, 2012). Several sulfate-producing thermophilic Bacillales isolates were retrieved from Iberian soils (Portillo et al, 2012; Santana et al, 2013). Temperatures of and above 40ºC occur frequently in soils, reaching maximum temperatures around 75°C in South Iberian soils; these temperature values allow the proliferation of those thermophiles. Sulfate released by these thermophilic bacteria induced enhanced seedling growth of *Nicotiana benthamiana,* recorded following treatments with culture supernatants of selected thermophilic isolates recovered from soil (Santana et al, 2013).

Most bacteria performing inorganic sulfur oxidation are able to oxidize thiosulfate, an intermediate of the oxidation of inorganic sulfur (Madigan et al. 2003) to sulfate. In a study carried out by Portillo et al. (2012), monitorization during growth of thermophilic isolates gave similar sulfate concentrations in thiosulfate-supplemented and unsupplemented cultures, showing that the isolates were not inorganic sulfur oxidizing bacteria. The production of

sulfide (H_2S) by the thermophilic isolates, which could be chemically oxidized to sulfate, was ruled out because it was undetectable by H_2S microelectrodes. On the other hand, the release of sulfate was related to organic nutrient availability, the lowest nutrient concentration showed the lowest level of sulfate (Portillo et al, 2012). Thus, this discovery established a novel step of sulfur cycling in terrestrial environments, a dissimilative organic-sulfur oxidation. The metabolic pathway involved in the production of sulfate by these thermophiles is not known, but it could be inferred from the genome annotation of sequenced Bacillales. In addition, a clue for the sulfate production must be found from the catabolism of amino acids cysteine and methionine, major sources of organic-sulfur. These amino acids are released from proteins after hydrolysis by proteases and their availability can be associated with the regulation of organic-sulfur transporters (see for example Erwin et al, 2005).

Accumulation of cysteine and homocysteine can pose a problem to the cell because they are highly reducing compounds. In mammalian cells, different pathways of cysteine degradation are important for the cell redox balance; in the cysteine-sulfinate (sulfinoalanine)-dependent pathway the oxidation of cysteine is performed by the cysteine dioxygenase (CDO) producing cysteine-sulfinate (sulfinoalanine). The latter can be catabolized in two different ways: either it is transaminated by an aspartate aminotransferase (AAT) with production of pyruvate and sulfate through sulfite oxidase (SUOX) (Figure 1) or it is decarboxylated by cysteine sulfinate decarboxylase (sulfinoalanine decarboxylase) into hypotaurine. Hypotaurine can be non-enzymatically oxidized into taurine (Bella, 1996). Interestingly, examination of the genomes of several Bacillales and their encoded proteins in NCBI databases indicates the presence of genes encoding for CDO, AAT and SUOX.

1. The *cdo* and *mtnD* genes

Cysteine dioxygenase has been reported as a mammalian non-heme iron enzyme that catalyzes the formation of cysteine sulfinate from L-cysteine by incorporation of dioxygen. Nevertheless, *cdo* gene is present, for instance, in *Bacillus subtilis* str. 168 and *Brevibacillus (Br.) brevis* NBRC 100599 genomes (Table 1). Two *Paenibacillus* hypothetical proteins, (e.g. the hypothetical protein PelgB in *Paenibacillus elgii* B69) show similarity to *B. subtilis* str. 168 CDO (57% similarity between CDO and PelgB, $E=4e^{-27}$). The Actinobacteria *Streptomyces sviceus* ATCC 29083, albeit possessing a cysteine dioxygenase gene in its genome, the CDO similarity with that of the Bacillales, in particular with the one of *B. subtilis* str. 168 is low (48%, $E = 4e^{-12}$ for only 57% coverage).

Every genome from the Bacillales analyzed in this study (Table 1) revealed the presence of a gene encoding another dioxygenase with different nomenclature, i.e. 1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase, methionine salvage pathway protein E-2/E-2', etc, these terms could be synonymous of acireductone dioxygenase (ARD) Fe^{2+} or Ni^{2+}, a product of the *mtnD* gene. According to Pfam database, the enzyme belongs to the Cupin 2 conserved barrel domain proteins and catalyzes the reaction between oxygen and the acireductone 1, 2-dihydroxy-3-keto-5-methylthiopentene (DHK-MTPene). However, the reaction yields different products, depending on the bound metal ion. When bound to iron, the reaction produces formate and 2-keto-4- methylthiobutyrate (KMTB), the alpha-ketoacid precursor of methionine in the methionine salvage pathway, present in mammals and some bacteria, for example, *B. subtilis* (Sekowska and Danchin, 2002). Besides the Bacillalles, a protein related to MtnD is present in *Klebsiella oxytoca* and *S. sviceus* ATCC 29083 (identity 29%, $E= 1e^{-16}$, 90% coverage).

2. The *mccA (yrhA)* and *mccB (yrhB)* genes

117 Methionine is usually degraded through its activated form, *S*-adenosylmethionine
118 (SAM) towards cysteine. SAM is subjected to the action of SAM-dependent methylases that
119 use SAM as a substrate and produce S-adenosyl homocysteine (SAH). SAH is further
120 hydrolyzed to homocysteine and adenosine by S-adenosylhomocysteine hydrolase (mostly in
121 eukaryotes) or it is first hydrolyzed to *S*-ribosyl-L-homocysteine (SRH) by SAH nucleosidase,
122 followed by conversion to L-homocysteine by *S*-ribosylhomocysteine lyase (mostly in
123 prokaryotes) (Figure 2). These reactions are part of the activated methyl cycle pathway,
124 whose function is the generation of the methyl group donor SAM and the regeneration of
125 methionine from SAH, through the activity of *MTA*/*SAH'ase*- (Pfs) enzymes. Pfs also converts
126 5'-methylthioadenosine (MTA) to methylthioribose (MTR) and adenine. Bacteria such as
127 *Klebsiella pneumoniae* and *B. subtilis,* convert MTR back to methionine (Greene, 1996;
128 Kredich 1996; Sekowska and Danchin, 2002) through the Methionine salvage pathway
129 (Figure 2).

130 The reverse transsulfuration pathway, leading from methionine and homocysteine to
131 cysteine, is present in several organisms (Figure 2). Mammalian cells possess an L-
132 homocysteine to cysteine pathway, where L-homocysteine reacts with serine and the
133 condensation product, cystathionine, formed by cystathione *β*-synthase, is further converted to
134 cysteine by cysthatione *γ*-lyase. However, bidirectional transsulfuration pathways that allow
135 the interconversion of homocysteine and cysteine via the intermediary formation of
136 cystathionine (Figure 2) are present in other organisms. For instance, in many enteric bacteria,
137 the synthesis of homocysteine from cysteine is the only means of transsulfuration (Greene,
138 1996), whereas *Lactococcus lactis* exhibits both pathways (Sperandio et al., 2005).

139 Prokaryotic *mccA* and *mccB* genes, encode the key enzymes of the reverse
140 transsulfuration pathway, cystathionine-βsynthase and cystathionine γ-lyase, respectively.

Cystathionine-β-synthase catalyzes the first step of the pathway, from homocysteine to cystathionine:

O-AcetylL-serine + L-homocysteine ⇌ L-cystathionine + H_2O (1)

The enzyme uses the cofactor pyridoxal-5'-phosphate (PLP) and can be allosterically regulated by effectors such as SAM. This enzyme belongs to the hydro-lyases family, cleaving carbon-oxygen bonds.

A homologous protein to *B. subtilis* 168 MccA in NCBI nr protein database is cysteine synthase, also attributed with the following terms: O-acetylserine lyase, O-acetylserine thiol lyase, CysK. In addition, more than one of these homologous were found; for instance in *Br. brevis* two cysteine synthases, besides cystathionine β-synthase, are reported from Blast comparison with the *B. subtilis* 168 cystathione β-synthase (E scores: $4e^{-77}$ and $2e^{-72}$, respectively). In *P. mucilaginosus* 3016, also two enzymes, Cysk 1 and Cysk 2, were identified (E scores of $6e^{-75}$ and $9e^{-77}$, respectively). Cysteine synthase is also named O-acetylserine sulfhydrylase and catalyzes the reaction shown in equation (2).

O-acetyl-L-serine + hydrogen sulfide ⇌ L-cysteine + acetate (2)

This suggests that MccA could be involved in the formation of sulfide from cysteine, as well as in the reverse reaction as indicated in Figure 2; indeed Hullo et al. (2007) have shown a low OAS thiol-lyase *in vitro* activity for *B. subtilis* MccA during the synthesis of cysteine.

Several homologous proteins to cystathione γ-lyase MccB of *B.subtilis* 168 are reported by PBlast, in the bacteria listed in Table 1. For instance, *Geobacillus thermoleovorans* has four proteins homologous to MccB: cystathionine γ-lyase (E=0), cystathionine β-lyase (E= $2e^{-145}$), cystathionine γ-synthase (E= $6e^{-124}$) and O-

165 acetylhomoserine sulfhydrylase (E= $2e^{-84}$). This is not surprising, as cystathionine γ-synthase,
166 cystathionine β-lyase, cystathionine γ-lyase, *O*-acetylhomoserine thiol-lyase, constitute a
167 family of evolutionary related proteins, whereas OAS thiol-lyase and cystathionine β-synthase
168 constitute a different family (Mehta and Christen, 2000). The MccB and MccA proteins of *B.*
169 *subtilis* belong to the cystathionine γ-synthase and to the OAS thiol-lyase family of proteins,
170 respectively. On the other hand, the homology to *O*-acetylhomoserine thiol-lyase suggests a
171 homocysteine γ-lyase activity for MccB, indicated by the dashed arrow in Figure 2, actually
172 reported by Hullo et al. (2007).

173 Considering the above, neither the presence of MccB or MccA as cystathionine γ-
174 lyase and β-synthase enzymes, respectively, can be undoubtfully proposed based simply on
175 genome annotation inherent to sequence analysis. The use of specific databases, for instance
176 of RegPrecise (a database centered on regulon analysis constructed by comparative genomics
177 in a large number of prokaryotic genomes), as well as the inspection of reported functional
178 studies on the same genera that are listed in Table 1, and gene-cluster analysis, indicate the
179 closest related sequences to MccA and MccB for this study (see Table 1 footnotes). For
180 example, the reverse transsulfuration pathway is present in *K. pneumoniae* (Seiflein and
181 Lawrence, 2006), whereas it has never been reported for *Escherichia.* Also, *Streptomyces*
182 *venezuelae* contains a cystathionine β-synthase implicated in reverse transsulfuration (Chang
183 and Vining, 2002). Nevertheless, (considering the high similarity found in Blast with *B.*
184 *subtilis* 168 Mcc proteins), *E. coli* and other Mcc-like sequences have been included in the
185 alignments and used for the phylogenetic approach (described in the section 6 below), which
186 corroborated the distinct function for several Mcc-like enzymes and the proposal of evolution
187 pathways.

188 In mammalian tissues, both cystathionine β-synthase and cystathionine γ-lyase are
189 primarily responsible for H_2S biogenesis. For instance, the first enzyme catalyzes also the

condensation of cysteine with homocysteine to form cystathionine and H_2S, the second produces H_2S also through reactions (3) and (4). H_2S can be converted to sulfite in the mitochondria by thiosulfate reductase, and the sulfite is further oxidized to thiosulfate and sulfate by sulfite oxidase. In bacteria, H_2S can be non-enzymatically converted to sulfite or eventually through the activity of sulfite reductase (see section 5). Therefore, one can speculate, that those Bacillales that do not possess a *cdo* gene, can use the methionine salvage/reverse transsulfuration pathway with synthesis of cysteine and H_2S, the latter can further be converted to sulfite used to generate sulfate through sulfite oxidase, whose gene has been detected in the genomes of most Bacillales (see Table 1 and Figure 1).

The activity of cystathionine γ-lyase also leads to ammonia production either from cysteine or homocysteine, according to the following equations:

$$\text{Cysteine} \longrightarrow \text{Pyruvate} + \text{ammonium} + \text{hydrogen sulfide} \quad (3)$$

$$\text{Homocysteine} \longrightarrow \alpha-\text{Ketobutyrate} + \text{ammonium} + \text{hydrogen sulfide} \quad (4)$$

Recently, Yoshida et al. (2010) reported the occurrence of reaction (3) in the gram-negative bacterium *Fusobacterium nucleatum*. Reactions (3) and (4) could represent a pathway (see Figure 2) in thermophilic Bacillales that would account for the hyperammonium production, equally observed in the isolated thermophiles (Portillo et al, 2012; Santana et al, 2013).

3. The *patA* gene

We have identified genes similar to the *patA* gene encoding an aminotransferase (AT) protein of *B. subtilis* 168 (Table 1). The nomenclature, also in this case, is not consensual, the

following terms were found: aminotransferase A, class I and II aminotransferase, aspartate transaminase. Despite the non-uniform nomenclature, all the enzymes belong to the aspartate aminotransferase family, which itself belongs to PLP-dependent aspartate aminotransferase superfamily. PLP combines with an alpha-amino acid to form an aldimine intermediate, the substrate in a transamination reaction. The enzyme aspartate aminotransferase (AAT), also called Glutamate oxaloacetate transaminase (GOT) is able to transfer an amino group from L-aspartate or L-alanine to α-ketoglutarate, the products of the reversible transamination reaction being oxaloacetate and glutamate, and pyruvate and glutamate, respectively. Enzymes GOT1 and GOT2, found in mammals, are isoenzymes from cytoplasm and mitochondria, respectively, which catalyze also the reaction sulfinoalanine + 2-oxoglutarate <=> 3-sulfinyl-pyruvate + L-glutamate (see Figure 1). 3-Sulfinoalanine is the product of cysteine dioxygenase activity and, interestingly, GOT2 (NP_037309.1 [*Rattus norvegicus*]) has 43% similarity (E= $1e^{-07}$) with an aminotransferase from *Bacillus cereus* ATCC 14579, whereas PatA is the most similar *B. subtilis* 168 enzyme to GOT2 (E= $4e^{-04}$, 42% similarity).

Bacterial ATs are involved in amino acid interconversion, however, being all ATs PLP-dependent enzymes, with formation of the PLP-aldimine intermediate, the latter actually enables a wide variety of further reactions such as C-S lyase activity by α, β- elimination (cysteine desulfurases) or decarboxylation (John,1995). As an example, the study of Marienhagen et al (2005) on the functional analysis of all twenty ATs proteins inferred from the genome sequence of *Corynebacterium glutamicum,* identified two cysteine desulfurase enzymes and a cystathionine β-lyase besides transaminases involved in amino acid synthesis. In addition, one AT (AlaT) with broad amino donor specificity and using pyruvate as acceptor was also identified. Thus, it was expectable to find several high similar enzymes to PatA in the same microorganism, as it happened before with Mcc enzymes. Also in this case, the choice of the closest related sequence to PatA was made based on Blast analysis (score,

structure domain, etc.) as well as on reported functional studies (e.g. Lee and Lee, 1993; Berger et al., 2003) and on gene-cluster analysis (see section 7). Still, due to the similarity in the PLP-catalyzed reactions and to the related structure of the different ATs one cannot confirm the presence of a Got2 enzyme in the listed bacteria (Table 1).

Cysteine desulfurases cleave L-cysteine and transfer the sulfur first to a cysteine residue in the active site and then to various other acceptors, and are involved in the biosynthesis of iron-sulfur clusters. Unlike most desulfurases, L-cysteine C-S-lyase from *Synechocystis sp.* PCC6714 does not have a conserved cysteine residue at the active site and catalyzes the breakdown of L-cysteine to yield sulfide, pyruvate and ammonia (Mihara and Esaki, 2002) instead of alanine and sulfur.

In mammals, AAT has a cysteine aminotransferase (CAT) activity, which, in combination with 3-mercaptopyruvate sulfurtransferase (MST) also leads to sulfide formation (see equations below).

L-Cysteine (or L-Aspartate) + PLP-Enz $\longleftrightarrow$ Mercaptopyruvate (or Oxaloacetate) + PMP-Enz (5)

MST transfers sulfane sulfur from mercaptopyruvate to an active cysteine site and a persulfide is formed (Equation 6). Sulfide is liberated from the MST-persulfide by a thiol reductant (RSH) or a thioredoxin (Westrop et al, 2009) as shown in Equation (7).

Mercaptopyruvate + MST-SH $\rightarrow$ Pyruvate + MST-SSH (6)

MST -SSH + RSH $\rightarrow$ MST -S- S-R + H_2S (7)

MSTs have been identified in bacteria (see Paul and Snyder 2012). *B. subtilis* 168 has an YbfQ protein, a putative rhodanese-related sulfurtransferase, which could be a 3-mercaptopyruvate sulfurtransferase with a rhodanese domain. Similar YbfQ rhodanese domain-containing proteins were identified by BlastP in the herein listed *Bacillus* and in *Br. brevis* NBRC 100599. A rhodanese domain protein was also found in *K. oxytoca* 10-5245, *E. coli* M863 and *C. carboxidivorans* P7. *Geobacillus thermodenitrificans* NG80-2, G. *kaustophilus* HTA426, G. *thermoleovorans* CCB_US3_UF5, G. sp. Y4.1MC1, T. *composti* KWC4 and *P. mucilaginosus* 3016 all present rhodanese-related sulfurtransferase-like proteins, although with lower BlastP scores (between 40 and 50), likely indicating significant differences between YbfQ and their protein structures.

Considering the above, it is possible that among the several ATs found by sequence analysis in Bacillales, there are cysteine desulfurases with similar function to L-cysteine C-S-lyase from *Synechocystis sp.*, thus accounting for ammonium production, and AAT enzymes with CAT activity. The CAT activity associated with MST activity, bypasses the requirement for CDO in the production of sulfite to be further oxidized by sulfite oxidase to sulfate.

4. The *yuiH* and *ylaL* genes

Sulfite oxidase named SUOX, SOR or YuiH, this last term derived from the *yuiH* coding gene in *Bacillus*, catalyze the two-electron oxidation of sulfite to sulfate (Equation 8), the electron acceptor is oxygen and/or heme-coordinated iron ions (Kelly et al, 1997) according to the enzyme structure (see section 6 below).

$$SO_3^{2-} + H_2O \rightarrow SO_4^{2-} + 2\ H^+ + 2\ e^- \quad (7)$$

289 A Blast analysis shows that a gene encoding a sulfite oxidase YuiH protein is present
290 in the Bacillalles, including all the thermophilic strains under study. In the case of *K.*
291 *oxytoca* 10-5245, a sulfoxide reductase catalytic subunit YedY is identified, with 25%
292 identity with YuiH from *B. subtilis* strain 168, for 71% query coverage (E= $1e^{-08}$). Also for
293 *E. coli* M683, a similar oxidoreductase molybdopterin binding domain-containing protein is
294 registered with 24% identity, for 69% query coverage (E= $1e^{-08}$). These differences are
295 related to differences in the enzymes primary structure and to the distinct role for YedY
296 (see sections 6 and 7).

297 A second putative sulfite oxidase encoded by *ylaL* is also present in all the listed
298 members (Table 1) of the genera *Bacillus*, *Geobacillus* and *Brevibacillus*, curiously the
299 ones that include the thermophilic sulfate producing strains.

300

301 **5. The *yvgQ* gene**

302 *B. subtilis yvgQ* gene codes for the hemeprotein subunit of hydrogen-sulfide: acceptor
303 oxidoreductase, commonly named sulfite reductase, an enzyme that catalyzes the six-
304 electron reduction of sulfite into hydrogen sulfide and water (Equation 9).

305 $$SO_3^{2-} + \text{electron donor} \rightleftharpoons H_2S + \text{oxidized donor} + 3\ H_2O \quad (9)$$

306 The electron donor may be an NADPH molecule, a bound flavin or a ferredoxin.

307 In *E. coli*, the enzyme is a complex composed of a flavoprotein component (FP) and a
308 hemeprotein component (HP) (see section 6), the latter is the catalytic subunit, with a 4Fe-
309 4S metal center. Table 1 indicates, among the selected microorganisms for this study, those
310 with a sulfite reductase containing both FP and HP subunits. Even though very similar

sequences to HP were found in the remaining organisms at the table, gene annotation and cluster-analysis (section 7) indicates they correspond to different enzymes, e.g. some genes are annotated as encoding ferredoxin-nitrite reductase. Indeed, both nitrite and sulfite reductase are part of the *NIR_SIR* (PF01077) 4Fe-4S domain family.

The ability of sulfite reductase for the synthesis of either hydrogen sulfide (which can be non-enzymatically converted to sulfite, as already pointed out) or sulfite in the reverse reaction that has been registered in some cases (see MetaCyc; http://www.metacyc.org), suggests that its activity can lead to sulfate synthesis through sulfite oxidase. Moreover, sulfite reductase activity comes with the potential for ATP generation (see section 8).

6. Structure and Phylogenetic analysis

Gene alignment phylogenetic analysis allowed the identification of motifs, and conserved versus non-conserved residues, which are the basis of potential differences in the structure stability and activity of the encoded enzymes. Particular attention was paid to the differences between thermophilic and mesophilic microorganisms.

The alignment of CDO sequences can be found in Supplementary Figures (Figure S1). CDO dioxygenase produces 3-sulfinoalanine using the iron center and the substrate (cysteine) to activate dioxygen and then carrying out the catalysis. The enzyme has a β-barrel structure that is characteristic of the cupin superfamily, whose members feature two conserved sequence motifs that provide the ligands for the metal binding site: $G(X)_5HXH(X)_{3,4}$-$E(X)_6G$ and $G(X)_5PXG(X)_2$-$H(X)_3N$. The active site is typically formed using a combination of the two His and Glu residues from the first motif and the His residue from the second more variable motif. In bacterial CDOs, Gly is found in the Glu position from the first motif.

As expected, the motif is found in the alignment of ARD sequences, also members of the cupin family (Supplementary material Figure S2). A high degree of similarity is in the conserved β-helix, particularly in the vicinity of the active site, where His 100, His 102, Glu 106 and His 145 (*B. subtilis* 168 numeration) provide the protein-based ligands for the metal of these metalloenzymes.

Neighbor-Joining (NJ) phylogenetic trees were obtained from CDO and ARD sequences (Figure 3). Three main branches can be established, one comprising CDO sequences, one including Firmicutes ARD sequences and one including the ARD sequences of Actinobacteria *S. sviceus* ATCC 29083 and the Proteobacteria *K. oxytoca* 10-5245. The pictured tree suggest that CDO and ARD could be out-paralogs, and distinct ARD sequences as well as some CDO sequences may have been the result of lateral transfer, as branch nodes with high bootstrap values are grouping distinct genera.

Supplementary Figures S3 and S4 show the alignment of MccA and MccB proteins, respectively. These are both PLP-dependent enzymes that share the common structural feature Ser-X-X-Lys. PLP combines with the Lys residue to form an aldimine bond (Schiff base linkage). However, some PLP-dependent enzymes do not have this Lys residue. Indeed, several sequences in Figure S3 have an Arg (R) residue at position 102 instead of Lys (K), both being basic polar amino acids. The presence of such differences in key residues can suggest a distinct functional role of MccA in different microorganisms.

The *B. subtilis* str. 168 MccB sequence is annotated as a cystathionine β-lyase, however it is reported as a cystathionine γ-lyase and homocysteine γ-lyase following the results of enzyme activity measurements (Hullo et al, 2007); the alignment in Figure S4 contains the sequences with highest Blast scores to *B. subtilis* str. 168 MccB sequence.

The analysis of NJ phylogenetic tree represented in Figure 4 indicates that MccA and MccB sequences are related and could be originated from a common ancestor. In addition,

three main groups can be defined, one including MccB sequences, one including MccA sequences from mesophilic Paenibacillaceae, *Thermobacillus composti* KWC4 and Enterobacteriaceae, and one including the MccA sequences from Bacillales including thermophilic genera. The group containing *E. coli* M863 MccA sequence, is consistent with the different PLP-binding motif S-X-X-R (found also in the Paenibacillaceae MccA subgroup) and with a distinct activity of MccA related *E. coli* enzyme- as cysteine synthase B (Zhao et al, 2006). Also, *E.coli* M863 "MccB" enzyme is an O-succinylhomoserine (thiol)-lyase responsible for cystathionine synthesis from cysteine and O-Succinylhomoserine (Tran et al, 1983; Duchange et al, 1983) and forms a separated branch from the remaining MccB sequences.

The alignment of AAT sequences is shown in Figure S5 (supplementary material). We have included in the PatA alignment two sequences for *C. carboxidivorans* P7 (a and b) and for *P.elgii* B69 (PatA and YugH); sequences a) and PatA are closest to *B. subtilis* PatA, however sequences b) and YugH, in the constructed phylogeny (Figure 5), group together with the highest similar sequence to PatA from *S. sviceus* ATCC 29083 and from *T. composti* KWC4, respectively, suggesting a distinct role or activity for these enzymes. Indeed, the phylogeny presented in Figure 5, reveals four main groups of similar sequences, one comprising the *Geobacillu*s, *Bacillus* and the *P. elgii* B69 PatA sequences, clustered close to *C. carboxidovorans* P7 a) sequence, a second group comprising the *Brevibacillus*, *T. composti* KWC4, *P. mucilaginosus* 3016 and *P. elgii* B69 YugH sequences, and two distant remaining groups, one comprising the Enterobacteriaceae, another *C. carboxidovorans* P7 b) and *S. sviceus* sequences. The active site motif in the alignment differs accordingly to this distribution, i.e. in the first group the SKSHSM stretch, which comprises the Lys (K) 231 residue (*B. subtilis 168* numeration) where PLP links, is replaced by SK[S/A]FAM in the

second group, whereas the Enterobacteriaceae have the motif SKSYNM and the remaining group the sequence [S/A] KTY[S/A]M.

The sulfite oxidase metalloenzyme is found in the mitochondria of all eukaryotes and in prokaryotic organisms and catalyzes the last step of the metabolism of sulfur-compounds before the sulfate is generated.

Mitochondrial SOR contains two identical subunits with an N-terminal domain, which has a heme cofactor, and a C-terminal domain hosting a molybdopterin cofactor with a Mo(VI) center, where the catalytic oxidation of sulfite takes place. Upon reacting with sulfite, one oxygen atom of the center is transferred to sulfite to produce sulfate, generating a two-electron-reduced Mo(IV). Water then displaces sulfate and the removal of two protons and two electrons to the exogenous cytochrome *c*, returns the active site to its original state. The electrons transferred via cytochrome *c* to the electron transport chain allow generation of ATP by oxidative phosphorylation (Cohen et al, 1972). In bacteria, two SOR groups have been defined. Group 1 is a complex of two subunits, one with the molybdenum cofactor and the other with the heme; SORs group 2 contains only the molybdenum (Kappler, 2011). The presence of a heme relay, which is an integral part of mitochondrial SOR or an additional *c*-type cytochrome subunit in bacteria, allows the electron transfer from the molybdenum cofactor to a subsequent acceptor (e.g. cytochrome *c*), while its absence in SORs group 2 makes oxygen the direct and final electron acceptor. The purified moderate thermophilic sulfite oxidase enzyme of *Deinococcus radiodurans* (D'Errico et al, 2006) is an example of type 2 SORs, unable to use horse heart cytochrome *c*, routinely used as an electron acceptor for sulfite oxidoreductases.

The alignment of YuiH -sulfite oxidase- sequences is shown in Figure S6. For *Br. brevis* NBRC100599, *S sviceus* ATCC29083 and *P. mucilaginosus* 3016, more than one YuiH sequence was found and included in the alignment. Both *Br. brevis* sequences are highly

similar to the one from *B. subtilis* (E scores $1e^{-79}$ and $4e^{-67}$), while *P. mucilaginosus* sequence a) is more similar to *B. subtilis* YuiH (E= $1e^{-88}$) than sequence b) (E= $2e^{-63}$). *S. sviceus* a) and b) sequences are the less similar with E scores of $8e^{-44}$ and $2e^{-41}$, respectively. These differences are translated into different groupings in the phylogenetic tree derived from S6 alignment (Figure 6). The *Geobacillus* sequences group together, as well as both *S. sviceus* and Enterobacteriaceae sulfite oxidase sequences; a fourth group comprises the sulfite oxidase sequences from genus *Bacillus* and Paenibacillaceae. Two main branches can be defined in the fourth group: one includes *T. composti* KWC4 and *Br. brevis* b) sequences, the other includes *Br. brevis* a) sequence, among others. The different nodes in the tree, separate SUOX-like enzymes having a distinct mode of catalysis or even another function. The Enterobacteriaceae proteins have a quite different primary sequence compared to the rest of the chosen sequences presented herein, and therefore they have the most dissimilar structure.

Actually, most of the sequences herein correspond to SOR type 2 enzymes, while the Enterobacteriaceae sequences are from YedY molybdoenzyme; YedY is part of the YedYZ complex, a heme-molybdoenzyme complex which was isolated from *E. coli* (Brokx et al, 2005), where YedZ is the heme-binding subunit. YedYZ has reductase activity with trimethylamine N-oxide and dimethyl sulfoxide and no oxidase activity with sulfite. The crystal structure of YedY was solved by Loschi et al. 2004 who observed that YedY has aromatic amino acid residues (Tyr91, Tyr275 and Trp290, Figure S6) at the active site, not present in typical sulfite oxidase sequences. Important residues in sulfite oxidase catalysis (Kisker et al, 1997) and molybdopterin binding (e.g. Cys75, Arg 149, Tyr183, His144, this last two are not present in the Enterobacteriaceae YedY enzymes) (D'Errico et al, 2006) are also outlined in Figure S6.

Supplementary Figure S7 shows the alignment for the second putative sulfite oxidase YlaL, already mentioned in section 4. Although a C-terminal molybdopterin binding

domain is found, as shown in the alignment, the Cys75 residue, important for catalysis in sulfite oxidase is not present, suggesting another function or a different sulfite binding for this protein. A neighbor-joining consensus tree performed for the alignment of both YuiH/ YedY and YlaL sequences indicates a common ancestor, from where three branch nodes defining these three sequence groups derive (results not shown).

E. coli has a sulfite reductase enzyme, YvgQR, whose structure is known (Crane et al., 1995). FP alpha component contains FAD and FMN, HP (YvgQ) beta-component contains a metal center with a [4Fe-4S] cluster coupled, through a cysteine, to siroheme. Electrons are transferred from NADPH to FAD and on to FMN in FP, and from there to the HP metal center, where they reduce siroheme-bound sulfite.

The NJ neighbor phylogenetic tree represented in Figure 7 was obtained from the alignment of YvgQ sequences (Figure S8). Clearly, the Bacillaceae and the Paenibacillaceae form two distinct groups, a third group is formed by both Enterobacteriaceae sequences. All the sequences possess a conserved region at the C-terminus, which correspond to the 4Fe-4S and siroheme binding domain; the prosite motif PS00365, characteristic of the beta subunit of sulfite reductase, [STV]-G-C-x(3)-C-x(6)-[DE]-[LIVMF]-[GAT]-LIVMF] is found, where the two C's are presumptive the iron-S ligands (Valdés et al, 2003).

7. Gene-cluster organization

Loci for the reported gene sequences were analyzed with respect to their organization in the genome. The *cdo* gene was found in the complementary strand of that of adjacent genes, except in the *S sviceus* ATCC29083 genome, where it is upstream of a gene important

456 in the biosynthesis of F420, cofactor for enzymes involved in redox reactions. On the
457 contrary, most of the remaining gene sequences showed an operon-like organization.

458 *mtnD* gene encoding acireductone dioxygenase is clustered together with three to four
459 other genes in all the listed members of genera *Bacillus, Geobacillus and Brevibacillus*
460 (Figure 8A). In *B. subtilis*, *B.megaterium*, *B.cereus*, and *Geobacillus,* the cluster *mtnEWXBD*
461 comprises the genes encoding the following enzymes: aminotransferase, 2,3-diketo-5-
462 methylthiopentyl-1-phosphate-enolase, 2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate
463 phosphatase, methylthioribulose-1-phosphate dehydratase and acireductone dioxygenase,
464 respectively. It should be noted that in *G. thermoleovorans* CCBUS3-UF5 *MtnB* is mistakenly
465 annotated as 2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase. In G.
466 *thermoglucosidasius, mtnW* gene is replaced for an *mtnW-like* pseudogene. *S. sviceus* and *K.*
467 *oxytoca* show also a cluster organization with fewer ORF (see Figure 8A), while in
468 *Paenibacillus* members, *mtnD* gene is not part of an operon. MtnD (ARD) is an enzyme that
469 participates in methionine metabolism producing the alpha-ketoacid precursor of methionine
470 (KMTB) in the methionine cycling pathway. Therefore, it is expected a location of *mtnD* gene
471 in the vicinity of genes *mtnE-B*, which code for other enzymes of this pathway; this operon
472 organization provides a fast cellular mechanism in methionine synthesis. Consequently,
473 *Paenibacillus* and *E.coli,* the latter does not possess an *mtnD* gene, are inefficient in the
474 supply of a methionine pool for subsequent sulfate formation, an observation that could
475 account for the recorded weak production of sulfate both in *E.coli* and *Paenibacillus*
476 mesophilic members (Portillo et al., 2012; Santana et al, 2013).

477 *mccA* and *mccB* genes are clustered together in all the *Bacillus*, *Geobacillus* and
478 *Brevibacillus*, as well as in *C. carboxidivorans* and *K. oxytoca* genomes. In *Bacillus* and
479 *Geobacillus*, *mcc* genes are within a cluster of at least four genes (Figure 8B) with the
480 following order *yrrT-mtn-mccA-mccB*, encoding respectively AdoMet-dependent

methyltransferase, 5'-methylthioadenosine-S-adenosylhomocysteine nucleosidase (Pfs), cystathionine-βsynthase and cystathionine γ-lyase. Bacteria that possess *mcc* genes are able to produce cysteine and sulfide by the reverse transsulfuration pathway. The MccA substrate-homocysteine- can be synthesized by homocysteine methyltransferase enzyme, from methionine, in the AdoMet Recycling pathway. Thus, it is not surprising to find *mcc* genes clustered together with *yrrT* and *mtn*; AdoMet-dependent methyltransferase catalyzes methyl group transfer from the ubiquitous cofactor SAM, fundamental both in AdoMet Recycling pathway and in Methionine salvage via MTR pathway; Pfs is equally important in Methionine salvage via MTR pathway (see section 2. and Figure 2). Therefore, the reported gene-cluster organization indicates a tight association and a common regulation for the mentioned methionine synthesis pathways and reverse transsulfuration. In those listed bacteria with no adjacent *mccB* genes, e.g. *Paenibacillus, T. composti* and *E. coli*, the MccA and the MccB–like sequences would indeed correspond to enzymes with a different function (see Figure 8B), respectively, cysteine synthase and cystathionine γ-synthase or cystathionine β-lyase, in agreement with the NCBI annotation and the reported function of Mcc-like enzymes in *E. coli*, where reverse transsulfuration is absent. Interestingly, the *yrrT-mtn-mccA-mccB* cluster is found in those genera comprising the moderate and thermophilic Firmicutes members, e.g. *Geobacillus,* which have been shown to be active sulfate producers (Portillo et al., 2012) suggesting that the presence of the above mentioned pathways association could account for such production.

patA gene encoding the *B. subtilis* aspartate aminotransferase protein is clustered together with *yhzT* gene, which encodes a short hypothetical protein. *patA* homologues, present in the analyzed genomes, are also in the immediate vicinity of one gene encoding a hypothetical protein in all the *Bacillus* and *Geobacillus* here referred. Once more, a similar organization is found for *Bacillus* and *Geobacillus* members (Figure 8C). In *P. elgii* and *P.*

mucilaginosus, we have found a *patA* homologue and a pseudogene *patA*, respectively. They are both preceded by gene *ylyB* encoding pseudouridine synthase, an enzyme responsible for converting uridine in RNA to its *C*-glycoside isomer, pseudouridine. Probably, the adjacent *patA*-like gene encodes a transferase with a role in the posttranscriptional modification of RNAs. The closest sequence to PatA in *T. composti* derives from the same branch node as *P. elgii* YugH and *P. mucilaginosus* aminotransferase sequences (Figure 5). The three corresponding genes are included in gene-clusters with similar organization, with *patA* framed between two transcriptional regulators coding sequences, one encoding a protein of Lrp/AsnC family, the other encoding a SpoE-like protein. The protein members of the Lrp/AsnC family affect cell metabolism in response to exogenous amino acids; thus, the location of these sequences suggests that these *pat* genes have a role in amino acid metabolism. In addition, both *Paenibacillus pat* sequences are part of a large cluster including a cysteine synthase coding gene, therefore a cluster possibly implicated in cysteine metabolism (see Figure 8C). Also in the *Brevibacillus* members, the *patA* genes are found downstream of a transcriptional regulator of Lrp/AsnC family. The remaining AT sequences in Figure 5 are codified by genes either isolated in one DNA strand (i.e. the Enterobacteriaceae sequences) or in the vicinity of other genes, as it is the case in S. *sviceus*, where the *patA*-like gene is nearby adenosine deaminase encoding gene, which is necessary for nucleic acid turnover. Considering the diversity of *patA*-like *loci* organization for the chosen sequences, as well as the versatility of ATs performed reactions, as mentioned in section 3, it is not possible to infer which of the *patA* sequences are encoding a 3-sulfinoalanine aminotransferase.

Gene-cluster organization in neighbor *loci* of sulfite oxidase encoding gene *yuiH* is quite distinct for the different listed microorganisms (Figure 8D). *yuiH* gene is found in the vicinity of same strand *loci* for genes encoding cell amino acid transporters in members of the genera *Bacillus, T. composti, Br. brevis and P. mucilaginosus* (sequences b) (see also Figure

6). *Brevibacillus brevis* (sequence a) and *Br. laterosporus yuiH* sequences are in gene-clusters with upstream *loci* for genes encoding a cytosol aminopeptidase and a CoA-binding protein, indicating an associated role of YuiH to peptide degradation and amino acid recycling. In *P.elgii and P. mucilaginosus* (sequence a), *yuiH* is part of a gene-cluster with two upstream genes encoding a two-component sensor histidine kinase and a putative oxidoreductase; this system may be implicated in the regulation of the uptake and metabolism of amino acids. The loci organization is different in the *Geobacillus* as well in *S. sviceus* (sequence a), where *yuiH* gene is isolated. In summary, *yuiH* gene is either clustered with genes encoding proteins as peptidases and/or amino acid transporters, suggesting that its role is coupled with protein entrance and further degradation, or in a single *locus* in *Geobacillus*, where it could be implicated in more than one pathway. In *K. oxytoca* 10-5245 genome, *yuiH* like-gene (*yedY*) is located next to *yedZ*, the latter encoding the heme binding subunit of the YedYZ reductase (see section 6), also present in *E. coli* M863; this surely accounts for different chemical properties of the enzyme and for the distinct function of this sulfite oxidase–like enzyme in the Enterobacteriaceae, as mentioned in the above section.

The YvgQ (also named CysI) sequences in Figure 7 are those whose genes are located immediately downstream the gene *cysJ* encoding the sulfite reductase flavoprotein alpha-component (Figure 8E). Therefore, it is expected that, for the microbes with this operon structure, a functional sulfite reductase complex is synthetized likewise that in *E. coli,* where it catalyzes the reduction of sulfite to sulfide and is required for the synthesis of L-cysteine from sulfate (see Wu et al., 1991) (see Equation 2). The gene-cluster organization of the remaining sequences (see Figure 8E) also suggests a function in amino acid metabolism and assimilation of inorganic sulfate, due to the proximity of the sulfite reductase–like gene to Adenosine 5'-phosphosulfate kinase (APSK) or to PAPS reductase coding genes. APSK catalyzes the phosphorylation of adenosine 5'-phosphosulfate to form 3'-phosphoadenosine 5'-

556 phosphosulfate (PAPS), the latter being a sulfate form important for the assimilation of
557 inorganic sulfate. PAPS reductase uses thioredoxin as an electron donor for the reduction of
558 PAPS to phospho-adenosine-phosphate (PAP).

559

560 **8. Proposed pathways for the high-sulfate production in thermophilic Firmicutes**

561

562 H_2S was undetectable in the isolated thermophilic sulfate-producers studied by Portillo
563 et al, (2012). This is likely to be related to the existence of enzymes in these bacteria capable
564 of rapidly convert sulfide in other sulfur forms. Also, although gas transport through the
565 membrane occurs by simple diffusion (Mathai et al, 2009), rapid ionization of the gas at
566 physiological pH prevents permeation through the membrane. This might have been
567 determinant in evolution, dictating the presence of such enzymes.

568 Several main conclusions can be draw from the overall genome analysis on the
569 presence of particular genes in the microorganisms under analysis in this study, together with
570 their location and the phylogeny of their encoded enzymes:

571 - the high sulfate production that has been found for members of genus *Bacillus*, and
572 mainly of genera *Geobacillus* and *Brevibacillus* (Portillo et al, 2012; Santana et al, 2013) is
573 related to the number and the type of metabolic pathways involved in that production, along
574 with the structure of the enzymes involved in those pathways. While the presence of certain
575 enzymes seems to be not determinant (e.g. the cysteine dioxygenase Cdo and the sulfide
576 oxidoreductase CysIJ), others- as the Mcc enzymes of the reverse transsulfuration pathway-
577 are important for the synthesis of sulfide together with ammonium (see Equations 3 and 4),
578 the latter also produced in large amounts by the thermophilic members of the aforementioned
579 genera. The produced sulfide is further oxidized to sulfite and this to sulfate, by sulfite
580 oxidase YuiH, the key enzyme for sulfate production.

581 - the reactions associated with ARD activity and Mcc enzymes bypass the need for the
582 presence of CDO, since ARD catalyzes the synthesis of KMTB, precursor of methionine
583 synthesis, the amino acid can be further degraded towards cysteine through the reverse
584 transsulfuration pathway with production of sulfide (Figure 2). The phylogeny of ARD
585 sequences establishes two main branches, one derived from a common ancestor for the
586 Firmicutes sequences, the second for *K. oxytoca* and *S. sviceus* (Figure 3). Moreover, the
587 gene-cluster analysis shows the same operon organization for the Firmicutes high sulfate-
588 producing genera *Bacillus*, *Geobacillus* and *Brevibacillus,* where the transcription regulation
589 of *mtnD* is common in genes of the methionine recycling pathway, which are in the same
590 operon.
591 - due to the large similarity between Mcc sequences and other Mcc-like enzymes,
592 which are related in terms of evolution (section 2), the presence of cystathionine-β-synthase
593 (MccA) and cystathionine-γ –lyase (MccB) is inferred through a joint analysis, including
594 knowledge on physiology and correspondent gene-cluster organization. Also in this case, the
595 operon organization found for the *Bacillus*, *Geobacillus* and *Brevibacillus* sequences is
596 related to the phylogeny, with grouping of the corresponding Mcc sequences (Figure 4). In the
597 genera *Bacillus* and *Geobacillus, mccA* and *mccB* locate next to genes codifying the enzymes
598 of the Ado Met recycling and methionine salvage pathways, this location indicates an
599 association between methionine recycling and the reverse transsulfuration pathway.
600 - as it happens with Mcc sequences, the presence of a GOT2 homologue (PatA) cannot
601 be inferred through genome analysis because of the related structure of different
602 aminotransferases and of the catalytic versatility of the PLP-catalyzed reactions (section 3).
603 The existence of one or more AT, able to form sulfunyl-pyruvate from sulfinoalanine pool
604 (Figure 1) and/or to use cysteine and aspartate as substrates will largely contribute to extra
605 sulfide production (Equations 5-7) to be further oxidized to sulfate. In the case of cysteine

606 aminotransferase activity, the ulterior activity of MST, a rhodanese sulfurtransferase, is
607 required. Putative rhodanese domain MSTs encoding genes have indeed been found in the
608 genomes of most of the listed Bacillales (see section 3). On the other hand, the juncture
609 analysis of PatA homologous sequences phylogeny and spatial arrangement of their genes,
610 suggests that they may correspond to distinct enzymes with different substrate specificity. A
611 common ancestor can define a main branch node for all the Firmicutes sequences (Figure 5),
612 with a subgroup including all the Bacillaceae, other the Paenibacillaceae sequences. The first
613 group presents a common *patA* gene spatial organization, while the location of *patA* gene
614 sequences in the second group (in the vicinity of genes implicated in amino acid metabolism,
615 particularly cysteine metabolism) suggest a role for the latter Pat proteins as CAT enzymes. It
616 is noteworthy, that other AT sequences in Bacillaceae members, with lower Blast scores to *B.*
617 *subtilis* PatA, show a spatial arrangement of the corresponding coding genes similar to the
618 second group. Probably, the existence of many AT translates in a global large potential to
619 produce sulfide by different *vias.*

620 - sulfite oxidase, YuiH, is undoubtefully determinant for sulfate production, as it is
621 precisely the enzyme that catalyzes its synthesis in the last step of sulfur-organic
622 dissimilation. The presence in all of the listed Firmicutes, except in *B. cereus*, namely in those
623 genera with high sulfate-producers (in some cases, as in *Br. brevis,* two YuiH-like sequences
624 are found) corroborates its importance. The enzyme uses oxygen as electron acceptor, readily
625 available to the aerobic Firmicutes who could therefore rapidly oxidize sulfite. The *yuiH* gene
626 shows quite different *locus* positioning in the different Firmicutes. In spite of that, in all cases
627 except for the *Geobacillus* members, the location points to the associated role of YuiH to
628 amino acid metabolism. The simultaneous presence of a second putative sulfite oxidase,
629 YlaL, for the genera comprising thermophilic high sulfate-producers, is intriguing and a
630 subject of further study.

- sulfite reductase produces sulfide, contributing to sulfate production through the ulterior sulfite production and the reaction by YuiH. Therefore, the Firmicutes containing sulfite reductase have an "extra source" for sulfate production. It may be that under excess hydrogen sulfide the enzyme will produce sulfite and an electron carrier (such as reduced ferredoxin) that might introduce electrons in the respiratory chain with further ATP generation. Interestingly, the sulfite reductase complex, which is encoded by two neighbor genes *cysI* and *cysJ* (section 7), is not common to all members of the same genus, suggesting that the common ancestor for *cysI* Firmicutes sequences has been lost in some members during evolution. Because properties of each of these enzymes remain to be known (for instance, the ability to catalyze the reverse reaction) we cannot infer on the reason for that lost, which could be also related to the sulfite transport systems in those microbes.

Based on the *in silico* analysis herein, a general metabolic pathway is proposed (Figure 9) for the high sulfate production in thermophilic Firmicutes, which have been characterized among *Geobacillus*, *Brevibacillus*, and also *Bacillus* genera (Portillo et al, 2012; Santana et al, 2013). However, the authors realize that not only several intersected pathways can exist in the same cell, as their operational status might be modulated according to different conditions, for instance under normoxic versus oxidative stress. At this regard, it is noteworthy to mention that few of the listed members (i.e. *B. cereus, G. Kaustophilus, P. elgii, Br. brevis and Br. laterosporus*) have a putative sulfide quinone oxidoreductase, a membrane-bound flavoprotein that oxidizes sulfide to protein-bound persulfide. The reaction results in soluble polysulfide formation and reduced ubiquinone (Griesbeck et al, 2002), which transfers the electrons to a bc_1 complex or quinol oxidase in the respiratory chain. The general pathway, pictured in Figure 9, is similar to sulfide oxidation in mitochondria (Kabil and Banerjee, 2010), which is connected to ATP generation. This fact not only reinforces the origin of the ancestral sulfite oxidase in bacteria (see Paul and Snyder 2012), considering the

656 endosymbiotic theory of mitochondria origin, as it may explain why in many thermophiles
657 sulfate production is higher under stress conditions, i.e. at temperatures near the maximum or
658 minimum growth temperatures, since ATP production could be a way to cope with such
659 stress.
660

661 **Acknowledgments**

662 The corresponding author, Margarida Santana, was supported by the program "Ciência
663 2008" from "Fundação para a Ciência e a Tecnologia". JMG acknowledges funding from the
664 Spanish Ministry of Economy and Productivity (Consolider CSD2009-00006) which included
665 cofinancing through FEDER funds.
666

667 **Declaration of interest**

668 The authors report no declarations of interest.
669
670
671
672
673
674
675
676

677 **References**

678 Anandham R, Indiragandhi P, Madhaiyan M, Ryu KY, Jee HJ, Sa TM. (2008).
679 Chemolithoautotrophic oxidation of thiosulfate and phylogenetic distribution of S
680 oxidation gene (soxB) in rhizobacteria isolated from crop plants. Res Microbiol, 159, 579-
681 89.

682 Bella DL, Stipanuk MH. (1996). High levels of dietary protein or methionine have different
683 effects on cysteine metabolism in rat hepatocytes. Adv Exp Med Biol, 403, 73-4.

684 Berger BJ, English S, Chan G, Knodel MH. (2003). Methionine regeneration and
685 aminotransferases in *Bacillus subtilis*, *Bacillus cereus*, and *Bacillus anthracis*. J Bacteriol,
686 185, 2418-31.

687 Brokx SJ, Rothery RA, Zhang G, Ng DP, Weiner JH. (2005). Characterization of an
688 *Escherichia coli* sulfite oxidase homologue reveals the role of a conserved active site
689 cysteine in assembly and function. Biochem, 44, 10339-48.

690 Chang Z, Vining LC. (2002). Biosynthesis of S-containing amino acids in *Streptomyces*
691 *venezuelae* ISP5230: roles for cystathionine β-synthase and transsulfuration. Microbiol,
692 148, 2135-47.

693 Cohen HJ, Betcher-Lange S, Kessler DL, Rajagopalan KV. (1972). Hepatic sulfite oxidase.
694 Congruency in mitochondria of prosthetic groups and activity. J Biol Chem, 247, 7759-66.

695 Crane BR, Siegel LM, Getzoff ED. (1995). Sulfite reductase structure at 1.6 Å: evolution and
696 catalysis for reduction of inorganic anions. Science, 270, 59-7.

697 D'Errico G, Di Salle A, La Cara F, Rossi M, Cannio R. (2006). Identification and
698 characterization of a novel bacterial sulfite oxidase with no heme binding domain from
699 *Deinococcus radiodurans*". J Bacteriol, 188, 694-01.

700 Duchange N, Zakin MM, Ferrara P, Saint-Girons I, Park I, Tran SV, Py MC, Cohen GN.
701 (1983). Structure of the *metJBLF* cluster in *Escherichia coli* K12. Sequence of the *metB*
702 structural gene and of the 5'- and 3'-flanking regions of the *metBL* operon. J Biol Chem,
703 258, 14868-71.

704 Ehrlich HL. (1996). Geomicrobiology. New York, USA: Marcel Dekker Publisher and
705 Distributor.

706 Eriksen J. (1996). Incorporation of S into soil organic matter in the field as determined by the
707 natural abundance of stable S isotopes. BiolFertil Soils, 22, 149–55.

708 Erwin KN, Nakano S, Zuber P. (2005). Sulfate-dependent repression of genes that function in
709 organosulfur metabolism in *Bacillus subtilis* requires Spx. J Bacteriol, 187, 4042-49.

710 Ghani A, McLaren RG, Swift RS. (1993). Mobilization of recently formed soil organic
711 sulphur. Soil Biol Biochem, 25, 1739–44.

712 Greene RC. (1996). Biosynthesis of methionine. In: Neidhardt FC, Curtiss III R, Ingraham JL,
713 Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger JE,
714 eds. *Escherichiacoli* and *Salmonella*: cellular and molecular biology, 2nd ed. Washington
715 DC: ASMPress, 542-60.

716 Griesbeck C, Schutz, M, Schodl T, Bathe S, Nausch L, Mederer N, Vielreicher M, Hauska G.
717 (2002). Mechanism of sulfide-quinone reductase investigated using site-directed
718 mutagenesis and sulfur analysis. Biochem, 4, 11552-65.

719 Hullo M-F, Auger S, Soutourina O, Barzu O, Yvon M, Danchin A, Martin-Vertraete I.
720 (2007). Conversion of methionine to cysteine in *Bacillus subtilis* and its regulation. J
721 Bacteriol, 189, 187-97.

722 John RA. (1995). Pyridoxal phosphate-dependent enzymes. Biochim Biophys Acta, 1248, 81-
723 96.

724 Kabil O, Banerjee R. (2010). Redox biochemistry of hydrogen sulfide. J Biol Chem, 285,
725 21903-07.

726 Kappler U. (2011). Bacterial sulfite-oxidizing enzymes. Biochim Biophys Acta, 1807, 1–10.

727 Kelly DP, Shergill JK, Lu WP, Wood AP. (1997). Oxidative metabolism of inorganic sulfur
728 compounds by bacteria. AVan Leeuw, 71, 95-107.

729 Kisker C, Schindelin H, Pacheco A, Wehbi WA, Garrett RM, Rajagopalan KV, Enemark JH,
730 Rees DC. (1997). Molecular basis of sulfite oxidase deficiency from the structure of sulfite
731 oxidase. Cell, 91, 973-83.

732 Kredich NM. **(**1996). Biosynthesis of cysteine. In: Neidhardt FC, Curtiss III R, Ingraham JL,
733 Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger JE,
734 eds. *Escherichiacoli* and *Salmonella*: cellular and molecular biology, 2nd ed. Washington
735 DC: ASMPress, 514-27.

736 Lee SH, Lee KJ. (1993). Aspartate aminotransferase and tylosin biosynthesis in *Streptomyces*
737 *fradiae.* Appl Environ Microbiol, 59, 822-27.

738 Loschi L, Brokx SJ, Hills TL, Zhang G, Bertero MG, Lovering AL, Weiner JH, Strynadka
739 NC. (2004). Structural and biochemical identification of a novel bacterial oxidoreductase.
740 J Biol Chem, 279, 50391-400.

741 Madigan M, Martinko JM, Parker J. (2003). Brock biology of microorganisms. New Jersey,
742 USA: Prentice Hall Publisher and Distributor.

743 Marchant R, Banat IM, Rahman TJ, Berzano M. (2002). The frequency and characteristics of
744 highly thermophilic bacteria in cool soil environments. Environ Microbiol, 4, 595–602.

745 Marienhagen J, Kennerknecht N, Sahm H, Eggeling L. (2005). Functional analysis of all
746 aminotransferase proteins inferred from the genome sequence of *Corynebacterium*
747 *glutamicum*. J. Bacteriol, 187, 7639–46.
748 Mathai JC, Missner A, Kugler P, Saparov SM, Zeidel ML, Lee JK, Pohl P. (2009). No
749 facilitator required for membrane transport of hydrogen sulfide. Proc Natl Acad Sci, 106,
750 16633-38.
751 Mehta PK, Christen P. **(**2000). The molecular evolution of pyridoxal-5'phosphate-dependent
752 enzymes. Adv Enzymol, 74, 129-84.
753 Mihara H, Esaki N. (2002). Bacterial cysteine desulfurases: their function and mechanisms.
754 Appl Microbiol Biotechnol, 60, 12-23.
755 Paul BD, Snyder HS. (2012). H_2S signalling through protein sulfhydration and beyond. Nat
756 Rev Mol Cell Biol, 13**,** 499-507.
757 Portillo MC, Santana M, Gonzalez J. (2012). Presence and potential role of thermophilic
758 bacteria in temperate terrestrial environments. Naturwissenschaften 99, 43-53.
759 Santana MM, Portillo M, Gonzalez JM, Clara MI. (2013) Characterization of new soil
760 thermophilic bacteria potentially involved in soil fertilization. J Plant Nutr Soil Sci, 176,
761 47-56.
762 Schlesinger, WH. (1997). Biogeochemistry, 2nd edn. San Diego, CA, USA: Academic Press.
763 Seiflein TA, Lawrence JG. (2006). Two transsulfuration pathways in *Klebsiella pneumoniae*.
764 J Bacteriol, 188, 5762-74.
765 Sekowska A, Danchin A. (2002). The methionine salvage pathway in *Bacillus subtilis*. BMC
766 Microbiol, 2, 8-21.
767 Sperandio B, Polard P, Ehrlich DS, Renault P, Guédon E. (2005). Sulfur amino acid
768 metabolism and its Control in *Lactococcus lactis* IL1403. J Bacteriol, 187, 3762-78.

769 Tran SV, Schaeffer E, Bertrand O, Mariuzza R, Ferrara P. (1983). Appendix. Purification,
770 molecular weight, and NH2-terminal sequence of cystathionine gamma-synthase of
771 *Escherichia coli*. J Biol Chem, 258, 14872-73.

772 Valdés J, Veloso F, Jedlicki E, Holmes D. (2003). Metabolic reconstruction of sulfur
773 assimilation in the extremophile *Acidithiobacillus ferrooxidans* based on genome analysis.
774 BMC Genomics, 4, 51-66.

775 Westrop GD, Georg I, Coombs GH. (2009). The mercaptopyruvate sulfurtransferase of
776 *Trichomonas vaginalis* links cysteine catabolism to the production of thioredoxin
777 persulfide. J Biol Chem, 284, 33485-94.

778 Wu J-Y, Siegel LM, Kredich NM. (1991). High-level expression of *Escherichia coli*
779 NADPH-sulfite reductase: requirement for a cloned *cysG* plasmid to overcome limiting
780 siroheme cofactor. J Bacteriol, 173, 325-33.

781 Yoshida Y, Ito S, Kamo M, Kezuka Y, Tamura H, Kunimatsu K, Kato H. (2010). Production
782 of hydrogen sulfide by two enzymes associated with biosynthesis of homocysteine and
783 lanthionine in *Fusobacterium* nucleatum subsp. nucleatum ATCC 25586. Microbiol, 156,
784 2260-69.

785 Zhao C, Kumada Y, Imanaka H, Imamura K, Nakanishi K. (2006). Cloning, overexpression,
786 purification, and characterization of O-acetylserine sulfhydrylase-B from *Escherichia coli*.
787 Protein Expr Purif, 47, 607-613.

788

789

790

791

Table captions

Table 1. Presence of genes involved in sulfate production.

Figure captions

Figure 1. Schematic representation of a proposed pathway for sulfate synthesis from cysteine in the Bacillalles based on metabolic pathways occurring in mammals and on the annotation of sequenced genomes.

Figure 2. Cysteine and methionine pathways in prokaryotes. After uptake, sulfate is reduced to sulfide (through enzymes, e.g. ATP sulfurylase and adenosine phosphosulfate kinase, etc) and the sulfur incorporated into cysteine. The latter reaction is catalyzed by CysK (O-acetylserine (thiol)-lyase-A). In many bacteria, cysteine is used for methionine *de novo* synthesis. Cysteine reacts with *O*-succinylhomoserine to form cystathionine, which is then cleaved to yield homocysteine. These reactions are catalyzed by MetB (cystathionine γ-synthase) and MetC (cystathionine β-lyase), respectively. Some bacteria use *O*-acetylhomoserine instead of *O*-succinylhomoserine in the MetB reaction. The conversion of homocysteine to methionine is catalyzed by either MetE (cobalamin-independent methionine synthase) or MetH (cobalamin -vitamin B_{12}-dependent methionine synthase); not all bacteria contain both enzymes. The reverse transsulfuration pathway is indicated by brown arrows. The symbol* refers to the putative production of sulfide through reactions catalyzed by YrhA (MccA) cystathionine-βsynthase; and YrhB (MccB), cystathionine γ-lyase (and homocysteine γ-lyase). Other abbreviations in the Figure are: KMTB (keto-4-methylthiobutyrate); OAS (O-acetylserine); MHF (4-hydroxy-5-methyl-3(2H)-furanone); MetA (homoserine O-succinyltransferase); MetK (SAM synthetase); YhcE (homocysteine methyltransferase); CysE

816 (serine *O*-acetyltransferase); LuxS (*S*-ribosylhomocysteine lyase). Underlined products,
817 pyruvate, α-ketobutyrate and ammonium are the results from MccB catalysis (see text).

818

819 **Figure 3**. Phylogenetic tree of both CDO and ARD sequences. Tree branches with bootstrap
820 below 50% were collapsed (the probability of nodes is indicated). Accession numbers of the
821 phylogenetic tree CDO members are: *B. subtilis subsp subtilis* str 168 NP_390992.1,
822 *Brevibacillus* (*Br.*) *brevis* NBRC 100599 BAH45804.1, *P. mucilaginosus* 3016
823 YP_005316048.1, *P. elgii* B69 ZP_09077176.1 and *S. sviceus* ATCC 29083 ZP_06917524.1.
824 Accession numbers of the phylogenetic tree ARD members are: *B. subtilis subsp subtilis* str
825 168 NP_389245.1, *B. megateri*um QM B1551 YP_003561720.1, *B. cereus* ATCC14579
826 NP_833757.1, *B. amyloliquefacciens* DSM7 YP_003920039.1, *G. thermoleovorans*
827 CCBUS3-UF5 YP_004981546.1, *G. thermoglucosidasius* C6-YS93 YP_004588899.1, *G.*
828 *thermodenitrificans* NG80-2 YP_001124967.1, *G. kaustophilus* HTA426 YP_146809.1, *G. sp*
829 Y4.1MC1 YP_003990211.1, *T. composti* KWC4 YP_007212873, *Br. Brevis* NBRC 100599
830 BAH45870.1, *Br. laterosporus* LMG15441 ZP_08639777.1, *P. mucilaginosus* 3016
831 YP_005313165.1, *P. elgii* B69 ZP_09074737.1 and *S. sviceus* ATCC 29083 ZP_06915925.1,
832 *K. oxytoca* 10-5245 EHT02396.1.

833

834 **Figure 4**. Phylogenetic tree of Mcc sequences. The represented tree is a consensus tree where
835 branches with bootstrap below 50% were collapsed (the probability of nodes is indicated).
836 Accession numbers of tree members are: *B. subtilis subsp subtilis* str 168 MCCA
837 NP_390604.1; MCCB NP_390603.1, *B. megateri*um QM B1551 MCCA YP_003565033.1;
838 MCCB YP_003565032.1, *B. cereus* ATCC14579 MCCA NP_834079.1; MCCB
839 NP_834078.1, *B. amyloliquefacciens* DSM7 MCCA YP_003921130.1; MCCB

YP_003921129.1, *G. thermoleovorans* CCBUS3-UF5 MCCA YP_004983267.1; MCCB YP_004983266.1; *G. thermoglucosidasius* C6-YS93 MCCA YP_004587174.1; MCCB YP_004587175.1, *G. thermodenitrificans* NG80-2 MCCA YP_001126564.1; MCCB YP_001126563.1, *G. kaustophilus* HTA426 MCCA YP_148394.1; MCCB YP_148393.1, *Geobacillus sp* Y4.1MC1 MCCA YP_003988449.1; MCCB YP_003988450.1, *T. composti* KWC4 MCCA YP_007214693.1; MCCB YP_007213800.1, *Br. Brevis* NBRC 100599 MCCA YP_002774267.1; MCCB YP_002774266., *Br. laterosporus* LMG15441 MCCA ZP_08639879.1; MCCB ZP_08639880.1, *P. mucilaginosus* 3016 MCCA YP_005316842.1; MCCB YP_005316168.1, *P. elgii* B69 MCCA ZP_09079537.1; MCCB ZP_09077242.1, *C. carboxidivorans* P7 MCCA ZP_05392126.1; MCCB ZP_05392125.1, *S. sviceus* ATCC 29083 MCCA ZP_06917587.1; MCCB ZP_06919429.1, *K. oxytoca* 10-5245 MCCA EHS91301; MCCB EHS91302.1. and *E. coli* M863 MCCA EGB63267; MCCB EGB59775.1.

Figure 5. Phylogenetic condensed tree of AAT-like sequences. Branches with bootstrap below 50% were collapsed (the probability of nodes is indicated). Accession numbers of tree members are: *B. subtilis subsp subtilis* str 168 NP_389283.2, *B. megateri*um QM B1551 YP_003561779.1, *B. cereus* ATCC14579 NP_833728.1, *B. amyloliquefacciens* DSM7 YP_003920072.1, *G. thermoleovorans* CCBUS3-UF5 YP_004981624.1, *G. thermoglucosidasius* C6-YS93 YP_004588853.1, *G. thermodenitrificans* NG80-2 YP_001125020.1, *G. kaustophilus* HTA426 YP_146888.1, *G. sp* Y4.1MC1 YP_003990169.1, *T. composti* KWC4 YP_007212287, *Br. Brevis* NBRC 100599 YP_002770592.1, *Br. laterosporus* LMG15441 ZP_08639552.1, *P. mucilaginosus* 3016 YP_005316197.1, *P. elgii* B69 PatA ZP_09073383.1; YugH ZP_09076625.1, *C. carboxidivorans* P7 a) ZP_05394073.1; b) ZP_05391167.1, *S. sviceus* ATCC 29083 ZP_06919100.1, *K. oxytoca* 10-5245 ZP_17103151.1 and *E. coli* M863 ZP_16848169.

865 **Figure 6**. Phylogenetic tree of YuiH sequences. Branches with Bootstrap below 50% were
866 collapsed (the probability of nodes is indicated). Accession numbers of tree members are: *B.*
867 *subtilis subsp subtilis* str 168 NP_391082.1, *B. megateri*um QM B1551 YP_003564237.1, *B.*
868 *amyloliquefacciens* DSM7 YP_003921594.1, *G. thermoleovorans* CCBUS3-UF5
869 YP_004982456.1, *G. thermoglucosidasius* C6-YS93 YP_004588129.1, *G.*
870 *thermodenitrificans* NG80-2 YP_001125765.1, *G. kaustophilus* HTA426 YP_147612.1, *G. sp*
871 Y4.1MC1 YP_003989380.1, *T. composti* KWC4 YP_007214134.1, *Br. Brevis* NBRC 100599
872 a) YP_002774226.1; b) YP_002775048.1, *Br. laterosporus* LMG15441 ZP_08639908.1, *P.*
873 *mucilaginosus* 3016 a) YP_005310784.1; b) YP_005314616.1, *P. elgii* B69 ZP_09074339.1,
874 *S. sviceus* ATCC 29083 a) ZP_06916490.1; b) ZP_06915020.1, *K. oxytoca* 10-5245
875 EHS90788.1and *E. coli* M863 EGB63665.1.

876

877 **Figure 7**. Phylogenetic tree of YvgQ sequences. The represented tree is a consensus tree
878 where branches with bootstrap below 50% were collapsed (the probability of nodes is
879 indicated). Accession numbers of tree members are: *B. subtilis subsp subtilis* str 168
880 NP_391223.1, *B. megateri*um QM B1551 YP_003563552.1, *B. amyloliquefacciens* DSM7
881 YP_003921786.1, *G. thermoleovorans* CCBUS3-UF5 YP_004982050.1, *G.*
882 *thermodenitrificans* NG80-2 YP_001125385.1, *G. kaustophilus* HTA426 YP_147263.1, *T.*
883 *composti* KWC4 YP_007211273.1, *Br. Brevis* NBRC 100599 YP_002775330.1, *P.*
884 *mucilaginosus* 3016 YP_005314415.1, *P. elgii* B69 MCCA ZP_09074926.1, *K. oxytoca* 10-
885 5245 ZP_17103638.1 and *E. coli* M863 ZP_16849177.1.

886

887 **Figure 8**. Schematic representation of *mtnD* (**A**), *mcc*(**B**), *patA* (**C**), *yuiH* (**D**) and *yvgQ* (**E**)
888 gene clusters for the analyzed genomes. **A:** *S. sviceus mtnC* codes for 2,3 diketo-5-

889 methylthio-phosphopentane phosphatase(enolase phosphatase E1), which regenerates
890 methionine from methylthioadenosine (MTA) **B:** *Br. brevis yutD* codes for a hypothetical
891 protein. *P. mucilaginosus mccA*-like *is* cysK2, which codes for cysteine synthase,
892 *ppIase_ppiC* codes for a putative PpiC-type peptidyl-prolyl cis-trans isomerase and *trpE*
893 codes for anthranilate synthase. **C:** *Br. brevis patA*-like gene *is alaT, alaR* codes for an
894 transcriptional regulator of AsnC family. *P. mucilaginosus patA*-like is *yugH, yugG* codes for
895 a transcriptional regulator of AsnC family, *cysK* codes for cysteine synthase, PM3016_6417
896 codes for a SpoOE-like sporulation regulatory protein and PM3016_6420 codes for a
897 hypothetical protein. **D***: B. subtilis yuiF* codes for an amino acid transporter, *yuiE* codes for
898 leucylaminopeptidase. *Br. brevis yneT* codes for a putative CoA- binding protein, *pepA* codes
899 for a cytosol aminopeptidase and BBR47_47480 codes for a hypothetical protein. **E:** *G.*
900 *thermoglucosidasius YvgQ*-like is Geoth_3459 and codes for ferredoxin-nitrite reductase,
901 Geoth_3457 codes for sulfate adenyltransferase, Geoth_3458 codes for adenosine 5'-
902 phosphosulfate kinase (APSK), Geoth_3460 codes for a hypothethical protein, Geoth_3461
903 codes for uroporphyrin-III C-methyltransferase, Geoth_3462 codes for sirohydrochlorin
904 ferrochelatase and Geoth_3463 codes for acylphosphatase.
905

906 **Figure 9**. H_2S biogenesis and sulfate production. A. The enzymes cystathionine-βsynthase
907 (MccA), cystathionine γ-lyase (MccB) of the transsulfuration pathway, and aspartate
908 aminotransferase(s) of cysteine catabolic pathway catalyze H_2S and sulfite-producing
909 reactions. B. At the membrane level, the sulfide oxidation pathways couple H_2S oxidation to
910 the electron transfer chain and chemiosmotic ATP production.
911

912 **Table 1.** Presence of genes involved in sulfate production*

913

Bacteria (Family)	*mccA* (Cystathione β- synthase)	*mccB* (Cystathionine γ- lyase)	*mtnD* (Acireductone dioxygenase)	*cdo* (Cysteine dioxygenase)	*patA* (Aspartate aminotransferase)	*yuiH/ylaL* (Sulfite Oxidase)	*yvgQ* (Sulfite Reductase)
Bacillus subtilis subsp. *subtilis* str 168 (Bacillaceae)	+	+	+	+	+	+/+	+
Bacillus megaterium QM B1551 (Bacillaceae)	+	+	+		+	+/+	+
Bacillus cereus ATCC14579 (Bacillaceae)	+	+	+		+	/+	
Bacillus amyloliquefacciens DSM7 (Bacillaceae)	+	+	+		+	+/+	+
Geobacillus thermoleovorans CCBUS3-UF5 (Bacillaceae)	+	+	+		+	+/+	+
Geobacillus thermoglucosidasius C6-YS93 (Bacillaceae)	+	+	+		+	+/+	
Geobacillus thermodenitrificans NG80-2 (Bacillaceae)	+	+	+		+	+/+	+
Geobacillus kaustophilus HTA426 (Bacillaceae)	+	+	+		+	+/+	+
Geobacillus sp Y4.1MC1 (Bacillaceae)	+	+	+		+	+/+	
Thermobacillus composti KWC4 (Paenibacillaceae)	(+)[1]	(+)[1]	+		+	+/	+
Brevibacillus brevis NBRC 100599 (Paenibacillaceae)	+	+	+	+	+	+/+	+
Brevibacillus	+	+	+		+	+/+	

laterosporus LMG1544.1 (Paenibacillaceae)							
Paenibacillus mucilaginosus 3016 (Paenibacillaceae)	(+)[1]	(+)[1]	+	+	+	+/	+
Paenibacillus elgii B69 (Paenibacillaceae)	(+)[1]	+	+	+	+	+/	+
Clostridium carboxidivorans strain P7 (Clostridiaceae)	+	+			(+)[1]		
Streptomyces sviceus ATCC 29083 Streptomycetaceae)	+	+	+	+	(+)[1]	+/	
Klebsiella oxytoca 10-5245 (Enterobacteriaceae)	+	+	+		(+)[1]	(+)[1]/	+
Escherichia coli M863 (Enterobacteriaceae)	(+)[1,2]	(+)[1,2]			(+)[1]	(+)[1]/	+

914

915 *Gene presence was based on Blastp similarity scores to *B. subtilis* 168 sequences. Gene-cluster analysis and functional reports were also
916 considered. Symbols: + indicates presence (black: alignment score over 80, green: alignment score between 50 and 80). Empty boxes represent
917 absence (alignment score less than 50).

918 [1]The positive sign between brackets indicates distinct or possible distinct function. A putatively distinct funtion.

919 [2]Lack the enzymes of the reverse transsulfuration pathway (see text).

920

921

Figure 1

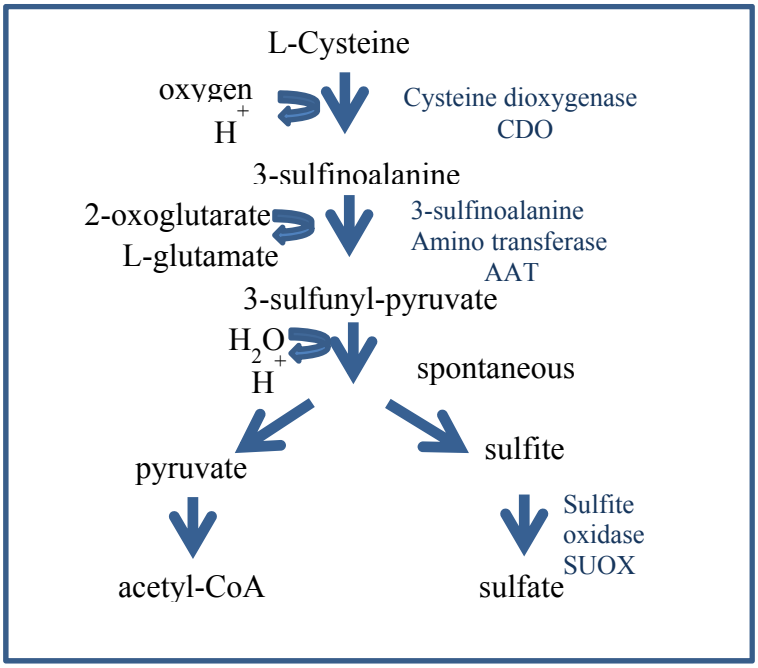
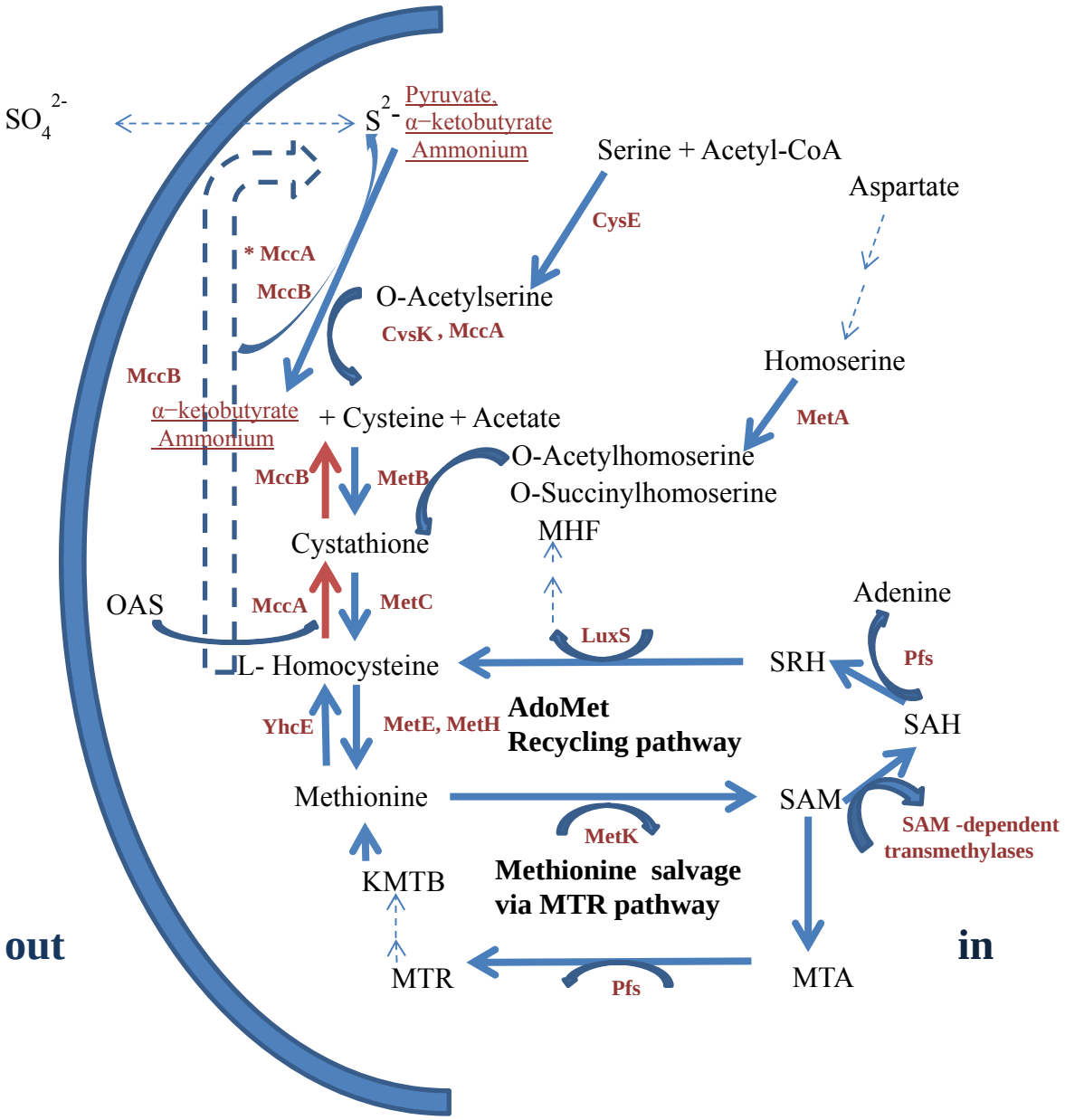

L-Cysteine
oxygen
H+
Cysteine dioxygenase
CDO
3-sulfinoalanine
2-oxoglutarate
L-glutamate
3-sulfinoalanine
Amino transferase
AAT
3-sulfunyl-pyruvate
H2O
H+
spontaneous
pyruvate
sulfite
Sulfite
oxidase
SUOX
acetyl-CoA
sulfate

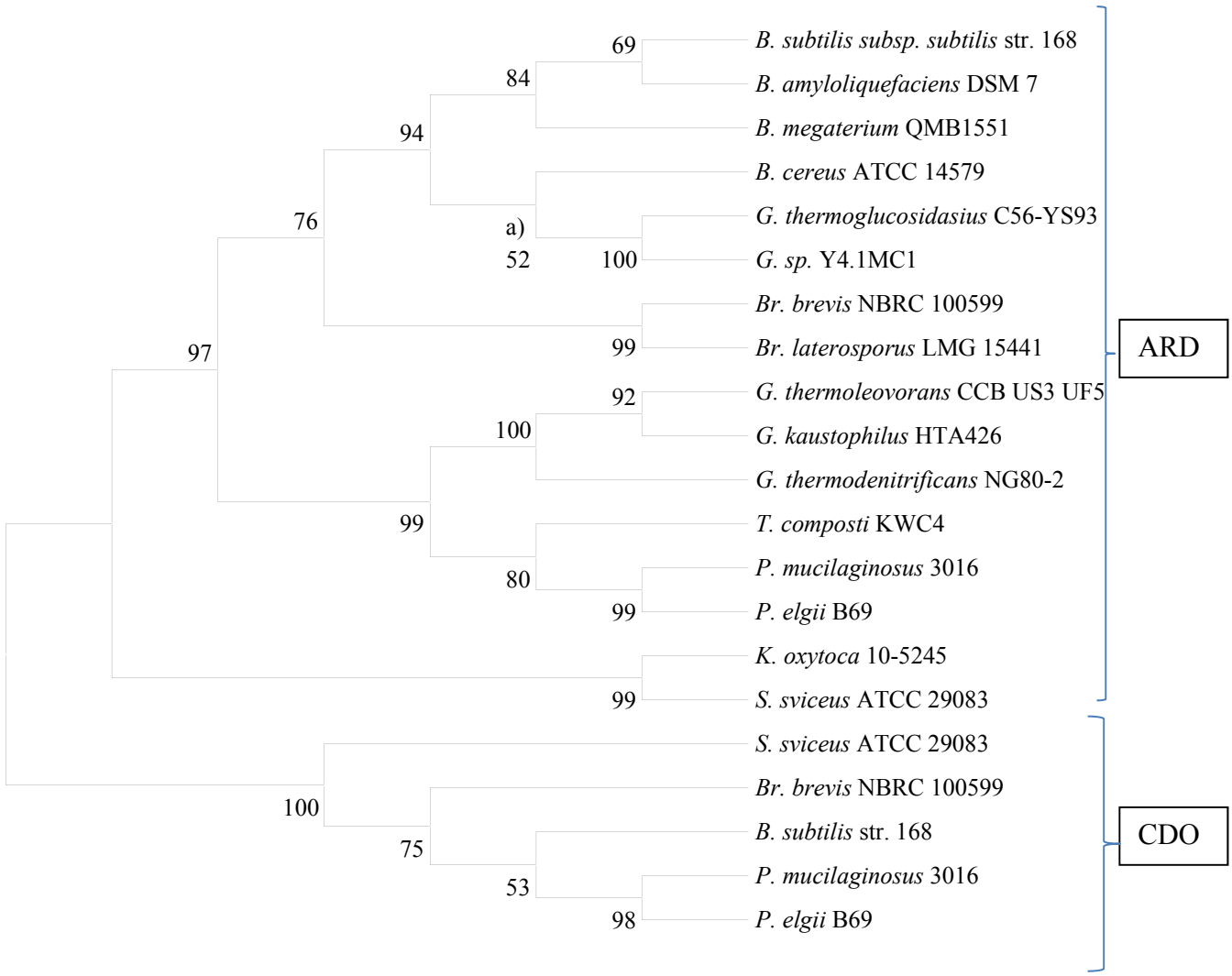
Figure 2


SO4 2-
S2-
Pyruvate,
α–ketobutyrate
Ammonium
Serine + Acetyl-CoA
CysE
Aspartate
* MccA
MccB
O-Acetylserine
CysK , MccA
Homoserine
MccB
α–ketobutyrate
Ammonium
+ Cysteine + Acetate
MetA
MccB
MetB
O-Acetylhomoserine
O-Succinylhomoserine
MHF
Cystathione
OAS
MccA
MetC
Adenine
LuxS
SRH
Pfs
L- Homocysteine
YhcE
MetE, MetH
AdoMet
Recycling pathway
SAH
Methionine
SAM
MetK
SAM -dependent
transmethylases
KMTB
Methionine salvage
via MTR pathway
out
in
MTR
Pfs
MTA

961 Figure 3

962

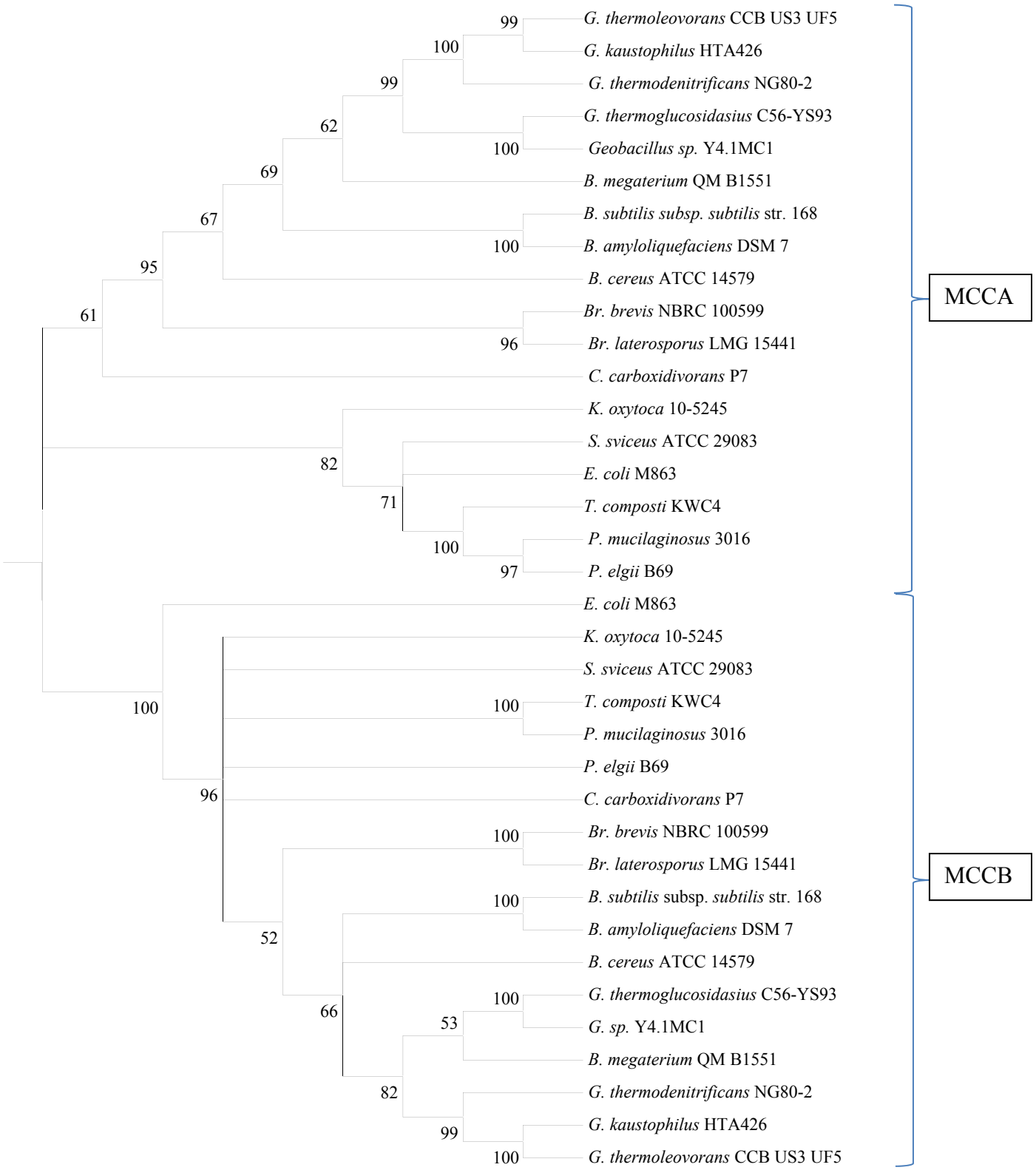
963



964

965

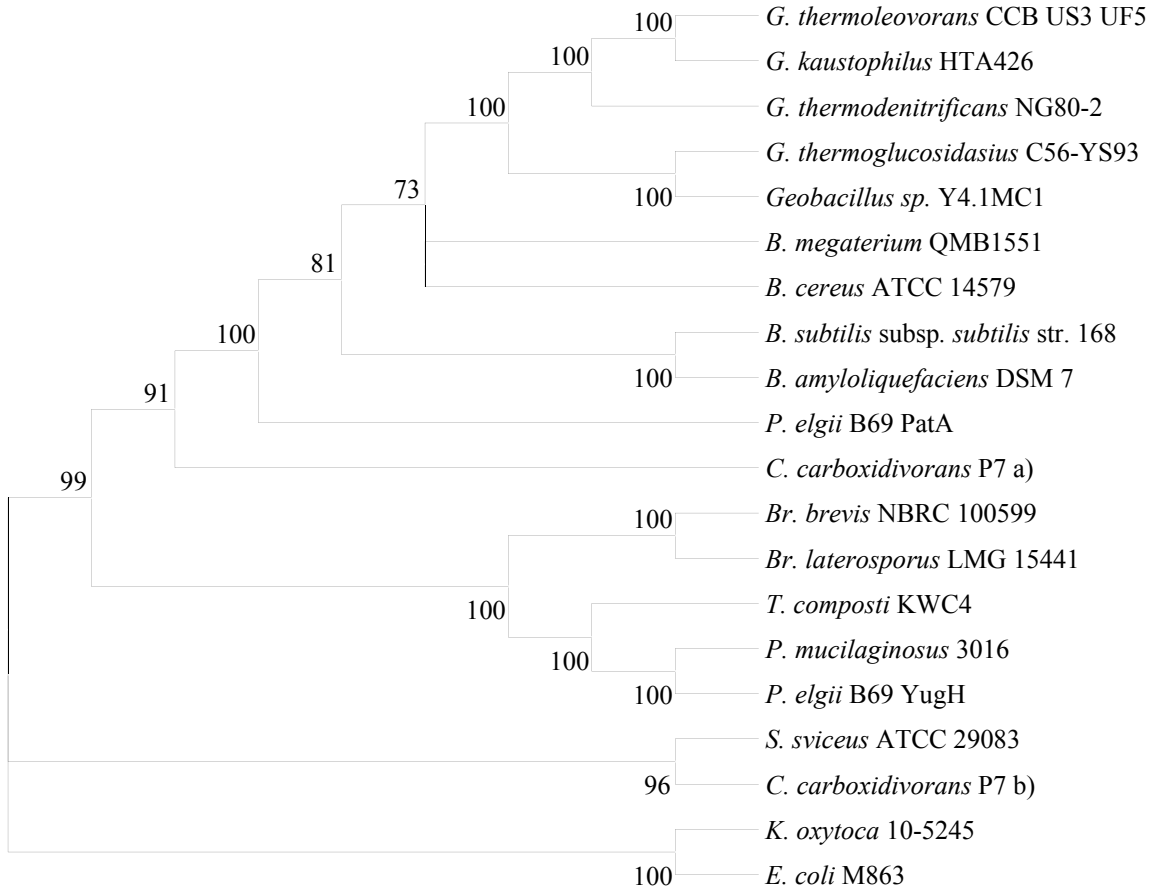
966 Figure 4



967

968

969 Figure 5

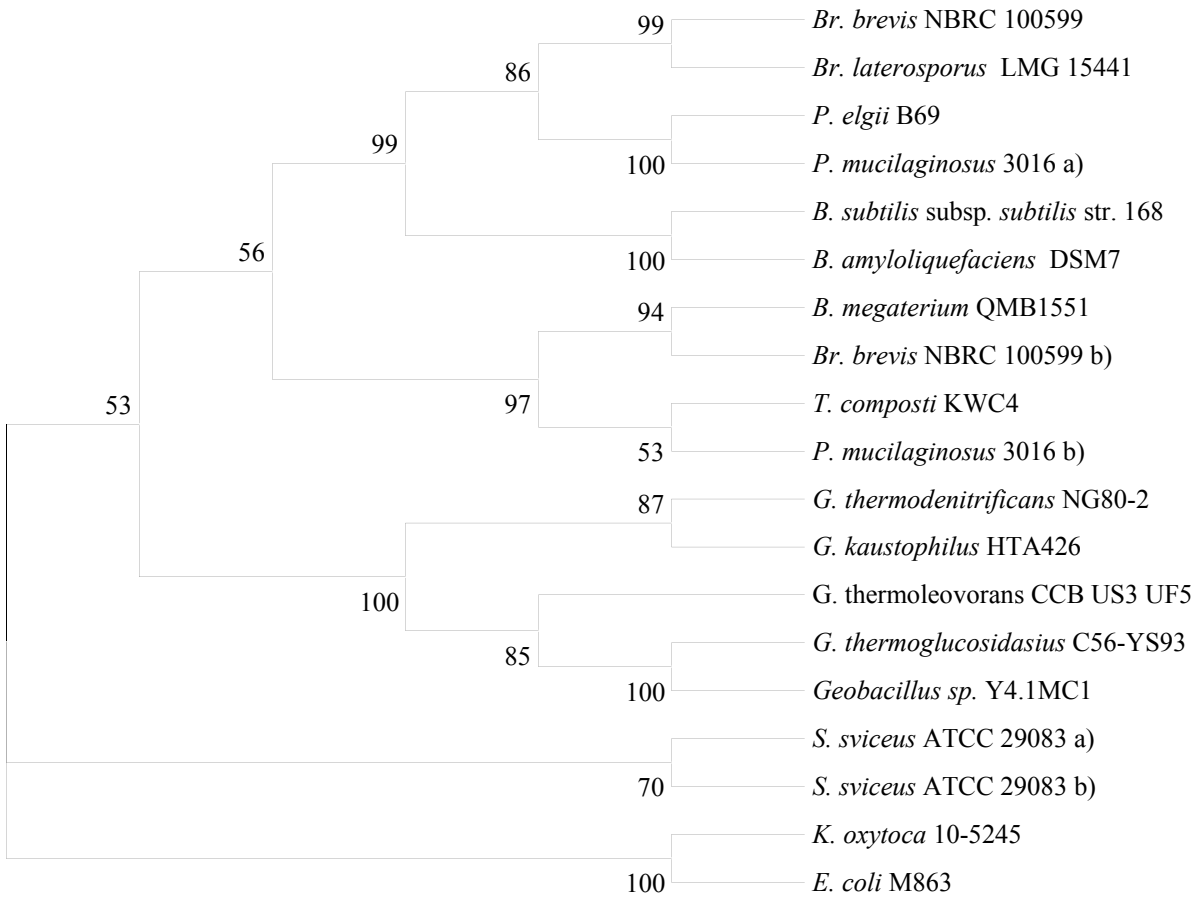


970

971

Fig 3

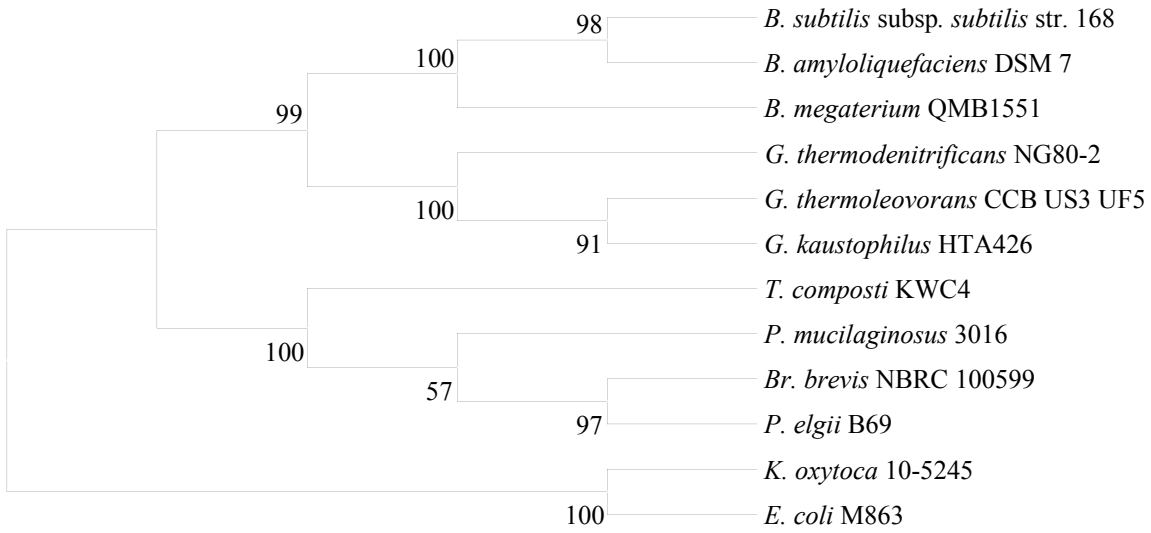
Figure 6


99
86
99
100
56
100
94
53
97
53
87
100
85
100
70
100
Br. brevis NBRC 100599
Br. laterosporus LMG 15441
P. elgii B69
P. mucilaginosus 3016 a)
B. subtilis subsp. subtilis str. 168
B. amyloliquefaciens DSM7
B. megaterium QMB1551
Br. brevis NBRC 100599 b)
T. composti KWC4
P. mucilaginosus 3016 b)
G. thermodenitrificans NG80-2
G. kaustophilus HTA426
G. thermoleovorans CCB US3 UF5
G. thermoglucosidasius C56-YS93
Geobacillus sp. Y4.1MC1
S. sviceus ATCC 29083 a)
S. sviceus ATCC 29083 b)
K. oxytoca 10-5245
E. coli M863

978 Figure 7

979

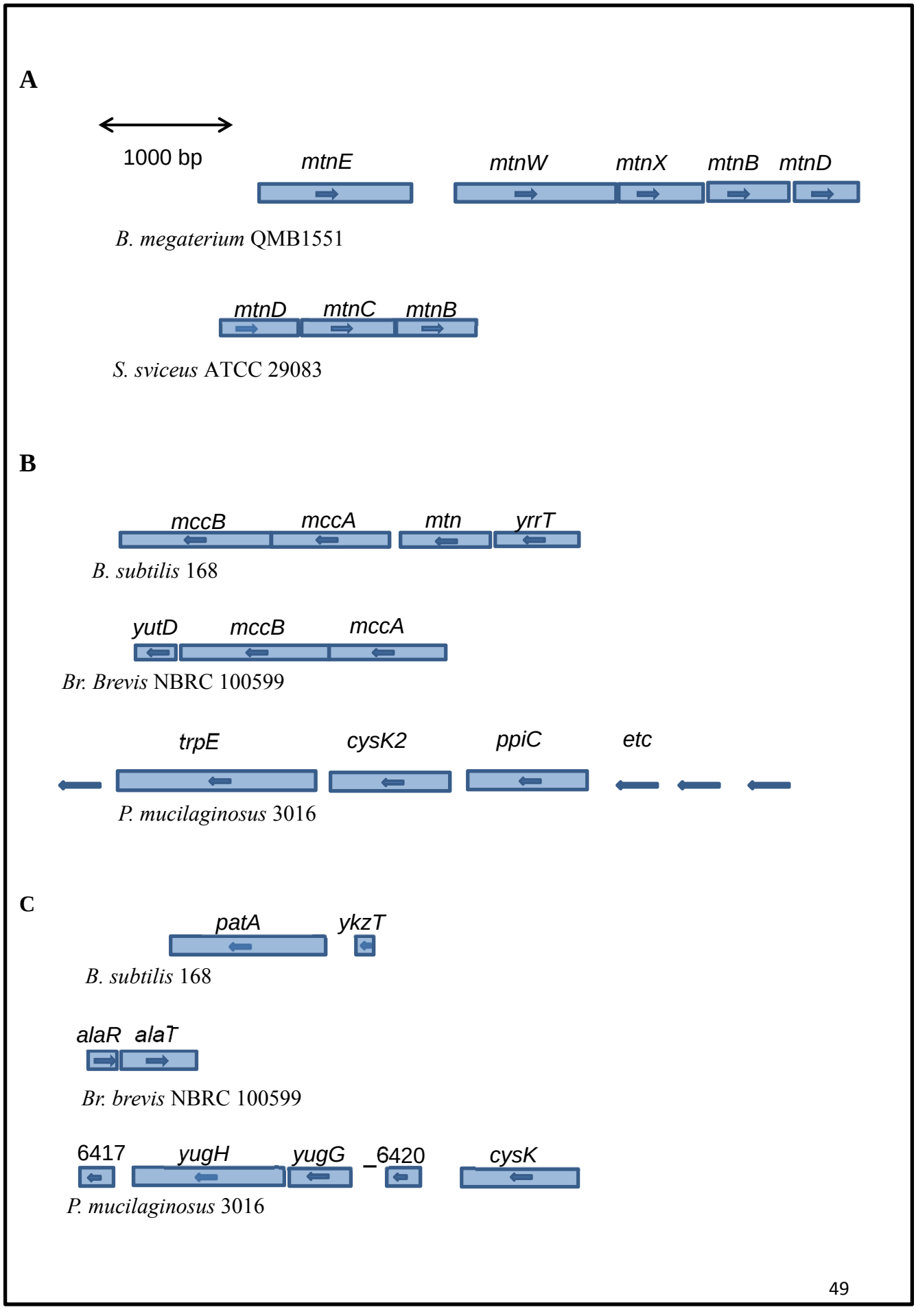
980

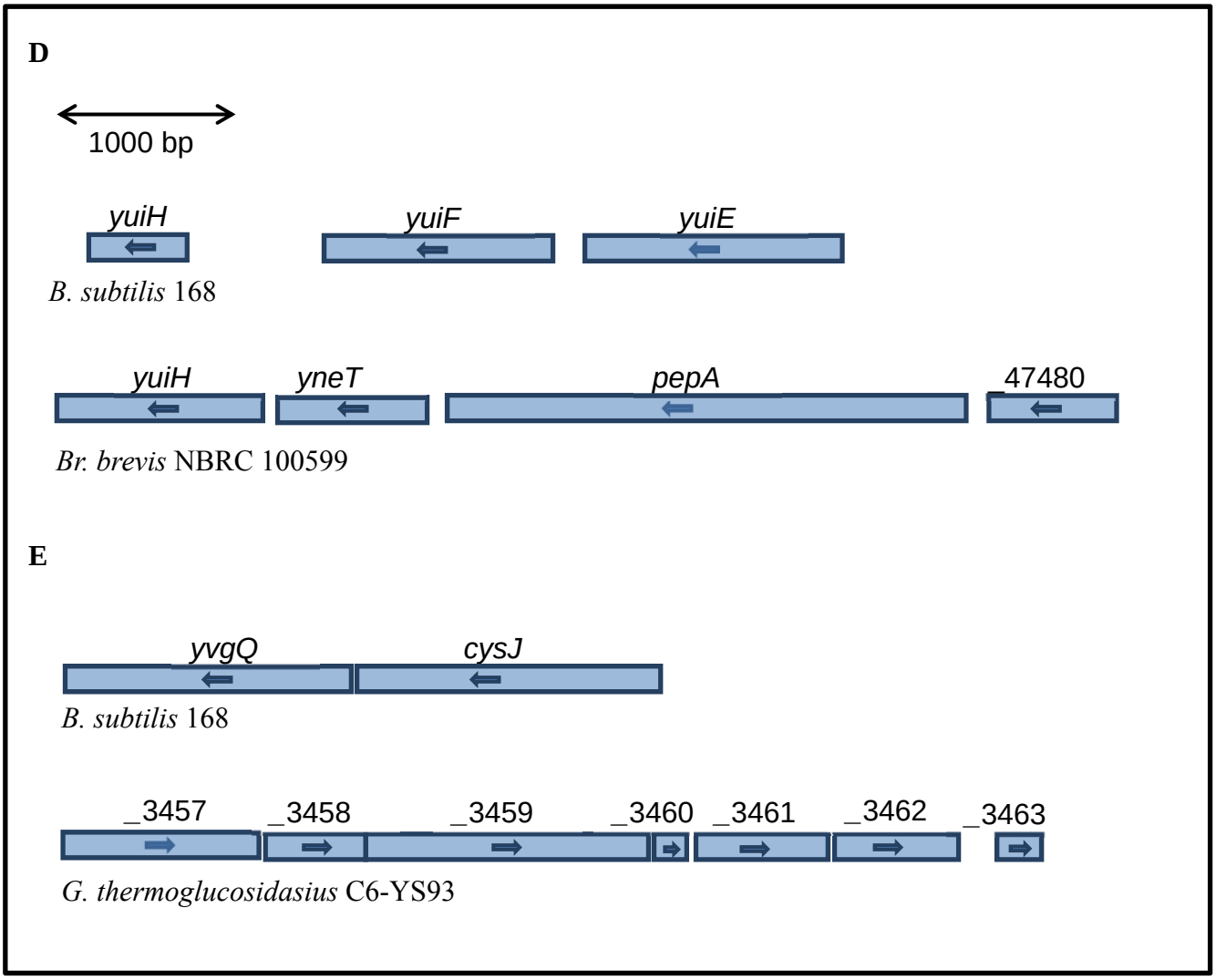


981

982

Figure 8

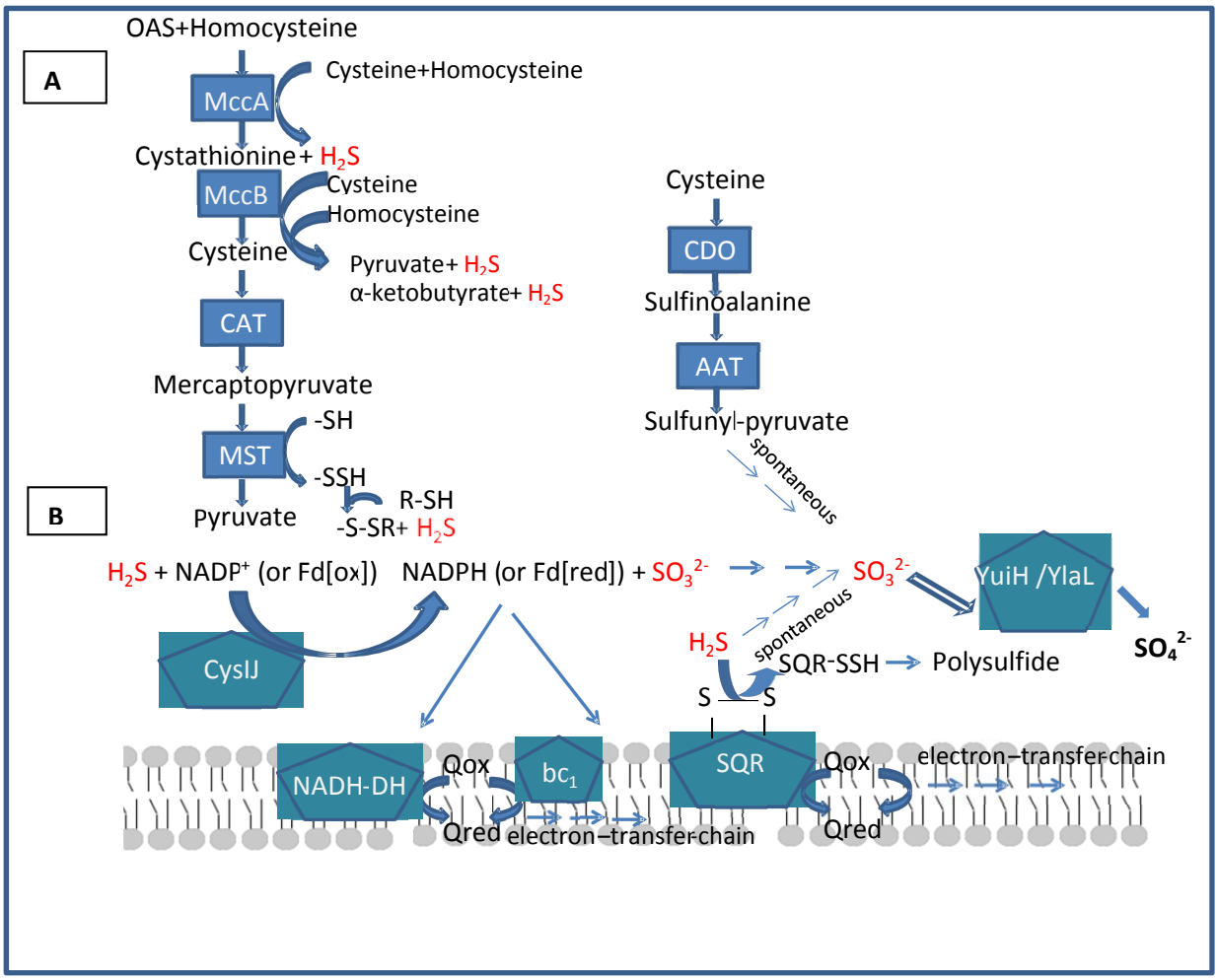

A
1000 bp
mtnE
mtnW
mtnX
mtnB
mtnD
B. megaterium QMB1551
mtnD
mtnC
mtnB
S. sviceus ATCC 29083
B
mccB
mccA
mtn
yrrT
B. subtilis 168
yutD
mccB
mccA
Br. Brevis NBRC 100599
trpE
cysK2
ppiC
etc
P. mucilaginosus 3016
C
patA
ykzT
B. subtilis 168
alaR
alaT
Br. brevis NBRC 100599
6417
yugH
yugG
_6420
cysK
P. mucilaginosus 3016


D
1000 bp
yuiH
yuiF
yuiE
B. subtilis 168
yuiH
yneT
pepA
_47480
Br. brevis NBRC 100599
E
yvgQ
cysJ
B. subtilis 168
_3457
_3458
_3459
_3460
_3461
_3462
_3463
G. thermoglucosidasius C6-YS93

1028 Figure 9

1029

1030



1031

1032