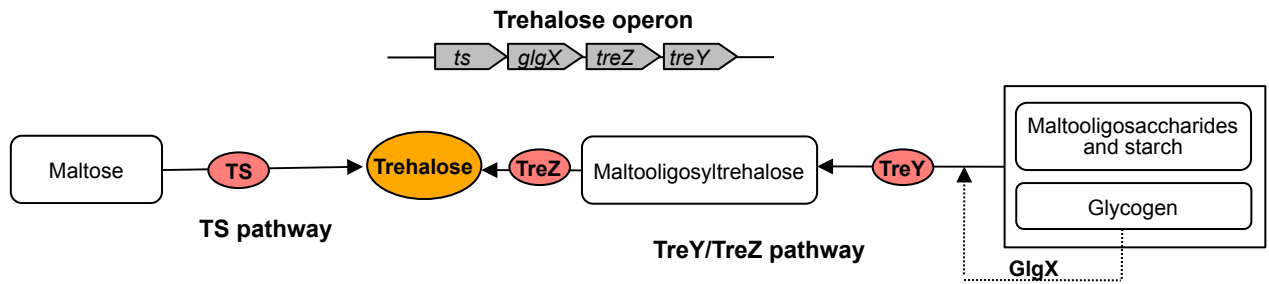


A.



B.

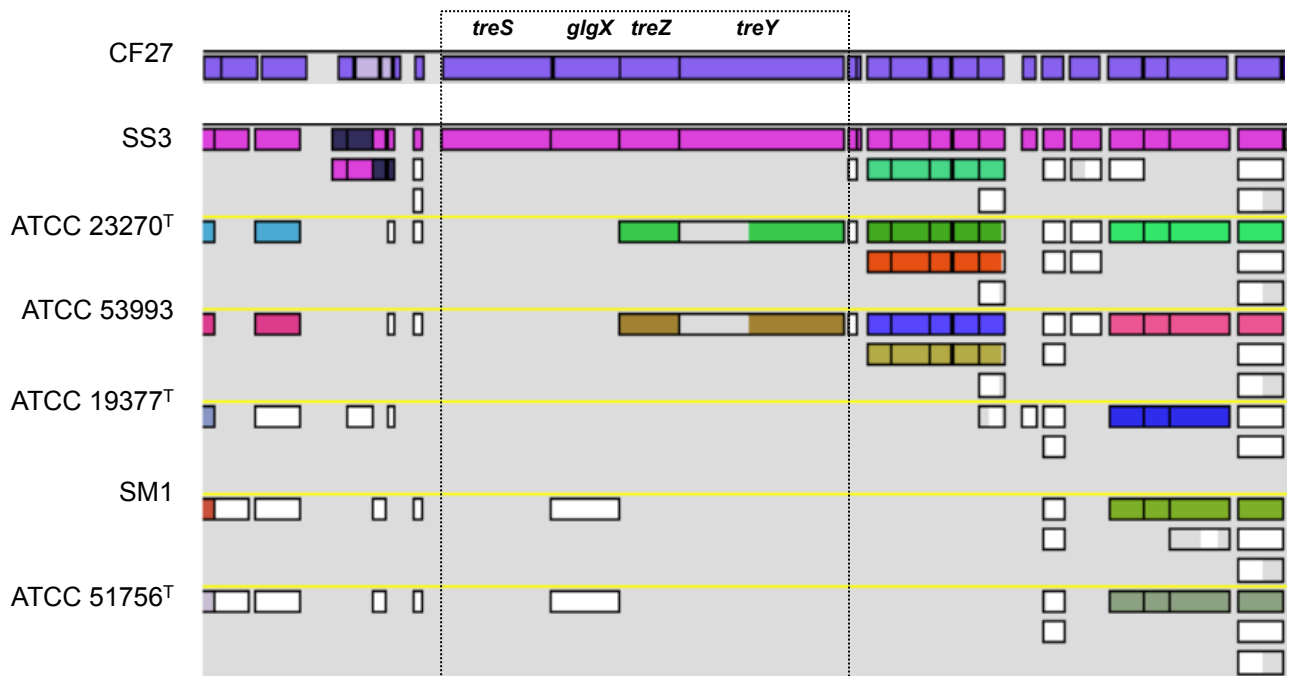


Figure S1. Trehalose biosynthesis in *Acidithiobacillus ferrivorans* CF27. (A) Various pathways for trehalose biosynthesis in *At. ferrivorans* CF27. Functional description of proteins are: TreS, trehalose synthase; GlgX, glycogen debranching protein; TreZ, malto-oligosyltrehalose trehalohydrolase; TreY, malto-oligosyltrehalose synthase. The trehalose operon is shown. (B) Synteny map of trehalose biosynthesis genes between *Acidithiobacillus* strains. The upper panel shows *At. ferrivorans* CF27 while the lower panel shows the other *Acidithiobacillus* strains. The genes predicted to be involved in trehalose biosynthesis are boxed. They include *ts* for trehalose synthase, *glgX* for glycogen debranching protein, *treZ* for malto-oligosyltrehalose trehalohydrolase and *treY* for malto-oligosyltrehalose synthase (see **Supplementary Table S5** for AFERRI_v2_ numbers). The same color tone indicates the same localization.

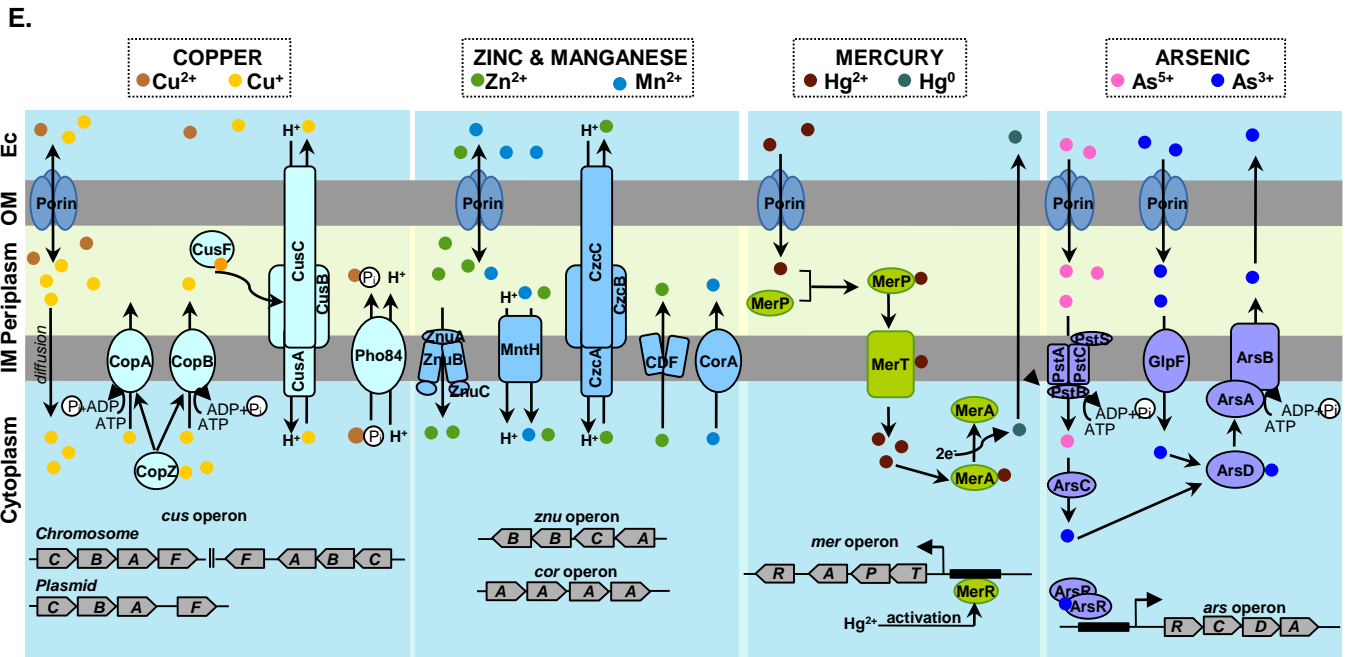
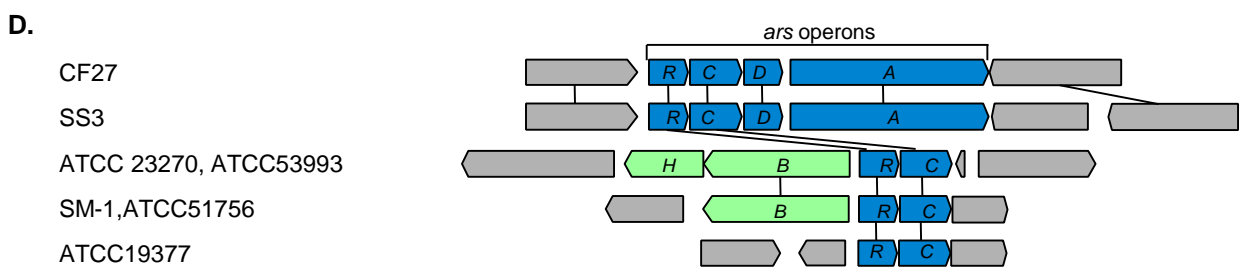
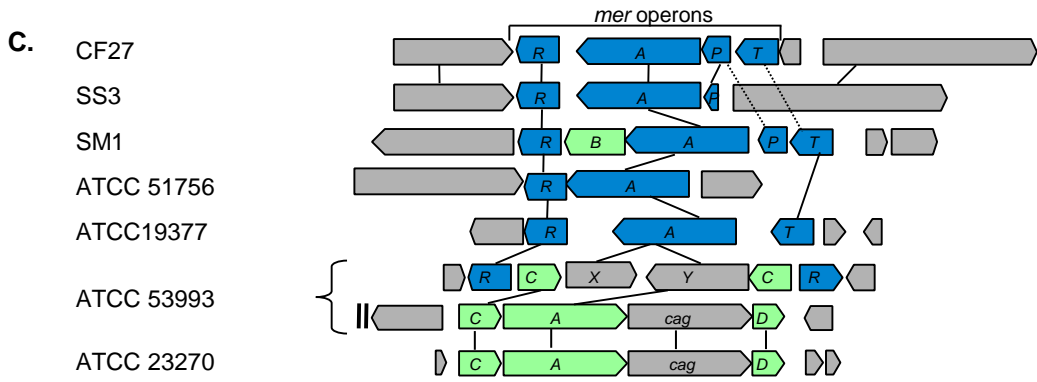
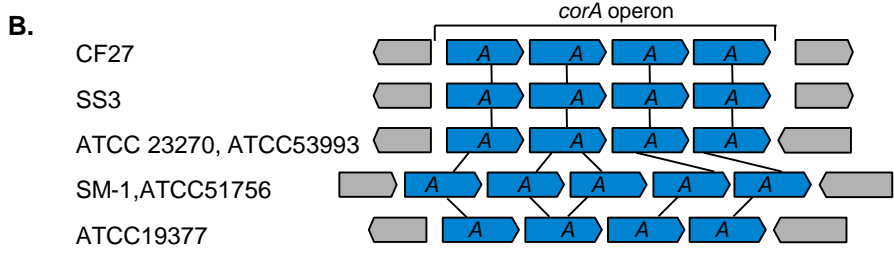
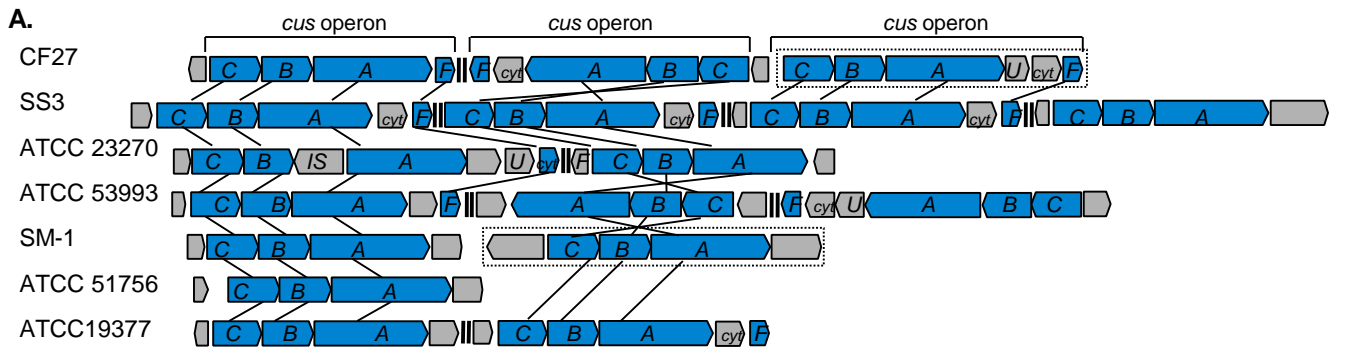


Figure S2. Transition metal and arsenic resistance mechanisms in *Acidithiobacillus* spp.. Organization of *cus* (A), *corA* (B), *mer* (C) and *ars* (D) operons in *At. ferrivorans* CF27 (see **Supplementary Table S5** for AFERRI_v2_numbers) and other *Acidithiobacillus* genomes. The strain names are: CF27, *At. ferrivorans* CF27; SS3, *At. ferrivorans* SS3; ATCC 23270, *At. ferrooxidans* ATCC 23270^T; ATCC 53993, *At. ferrooxidans* 53993; ATCC 19377, *At. thiooxidans* ATCC 19377^T; SM-1, *At. caldus* SM-1; and ATCC 51756, *At. caldus* ATCC 51756^T. The genes and operons are: *cusABCF* involved in copper resistance; *corA* involved in zinc resistance; *merABCDTR* involved in mercury resistance; *arsABCDRH* involved in arsenic resistance; *cyt*, cytochrome *b*₅₆₁; U, protein of unknown function; X, heavy metal transport/detoxification protein; Y, FAD-dependent pyridine nucleotide-disulfide oxidoreductase; *cag*, glucan 1,4-a-glucosidase; *IS*, ISafe3 transposase; *MFS*, major facilitator superfamily. Gene synteny is displayed by connecting lines between the orthologous genes. Genes involved in metal(loid) resistance presenting similarities with *At. ferrivorans* CF27 genes are shown in blue (same transcriptional direction) or in green (opposite transcriptional direction). Other genes are illustrated in grey. Operons located in plasmids were boxed in dotted line. (E) Copper, zinc, mercury and arsenic resistance mechanisms present in *At. ferrivorans* CF27. Copper resistance components: CopZ, putative cytoplasmic copper chaperone, CopA1 and CopB, copper-exporting P-type ATPases; CusCFBA, copper efflux pump system; Pho84, copper-phosphate complex transporter. Zinc and manganese resistance components: CzcCBA, Zn²⁺ efflux pump system; CzcD, Co²⁺-Zn²⁺-Cd²⁺ efflux protein; CorA, Mg²⁺/Co²⁺/Zn²⁺ transporter protein. Mercury resistance components: MerP, periplasmic mercuric ion binding protein; MerT, mercuric ion transporter; MerA, mercuric reductase. Arsenic resistance components: GlpF, glycerol MIP channel; PstACS, high affinity phosphate transport system; ArsC, arsenate reductase; ArsD, arsenical resistance operon trans-acting repressor and arsenic chaperone; ArsAB, arsenic efflux pump. Gene clusters are represented in grey arrows. OM, Outer Membrane; IM, Inner Membrane; Ec, Extracellular.

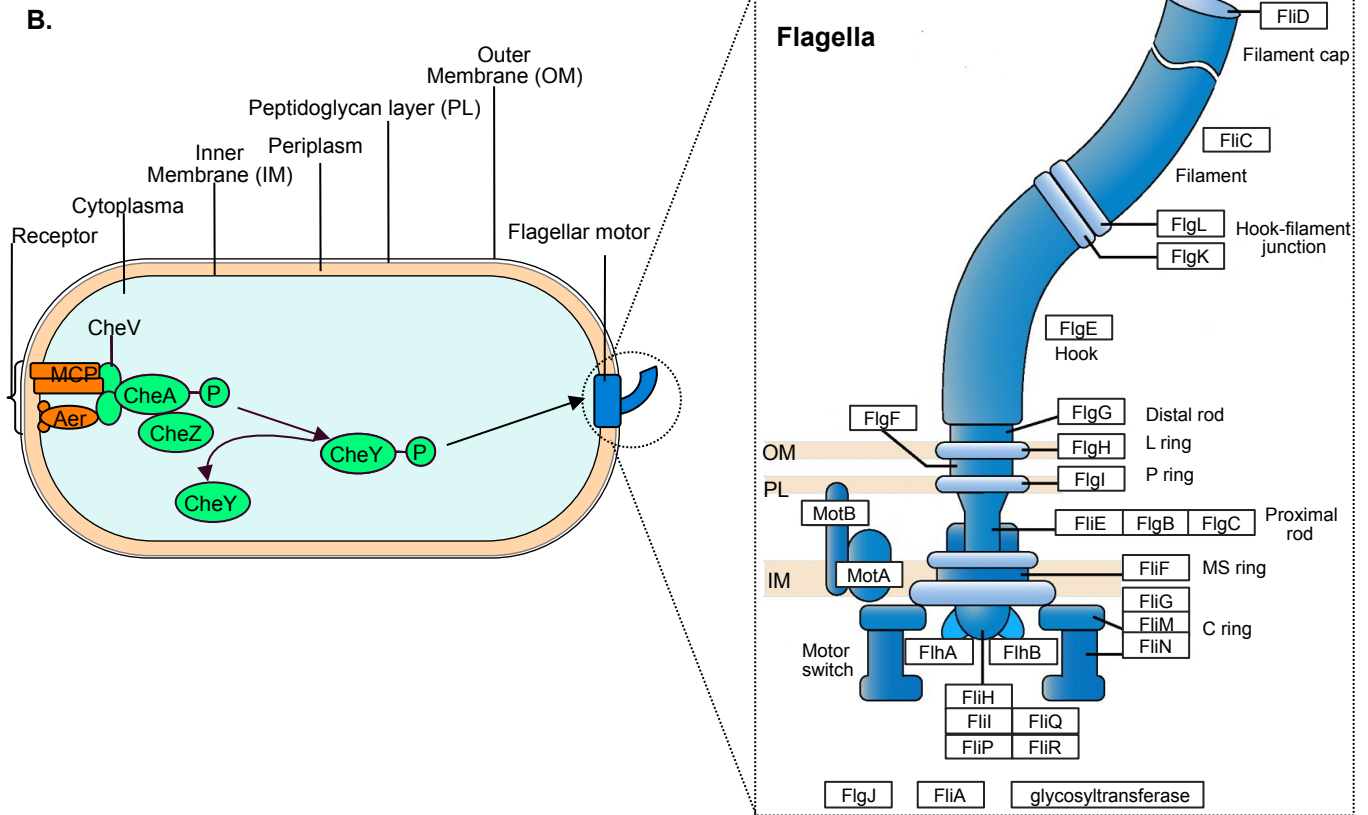
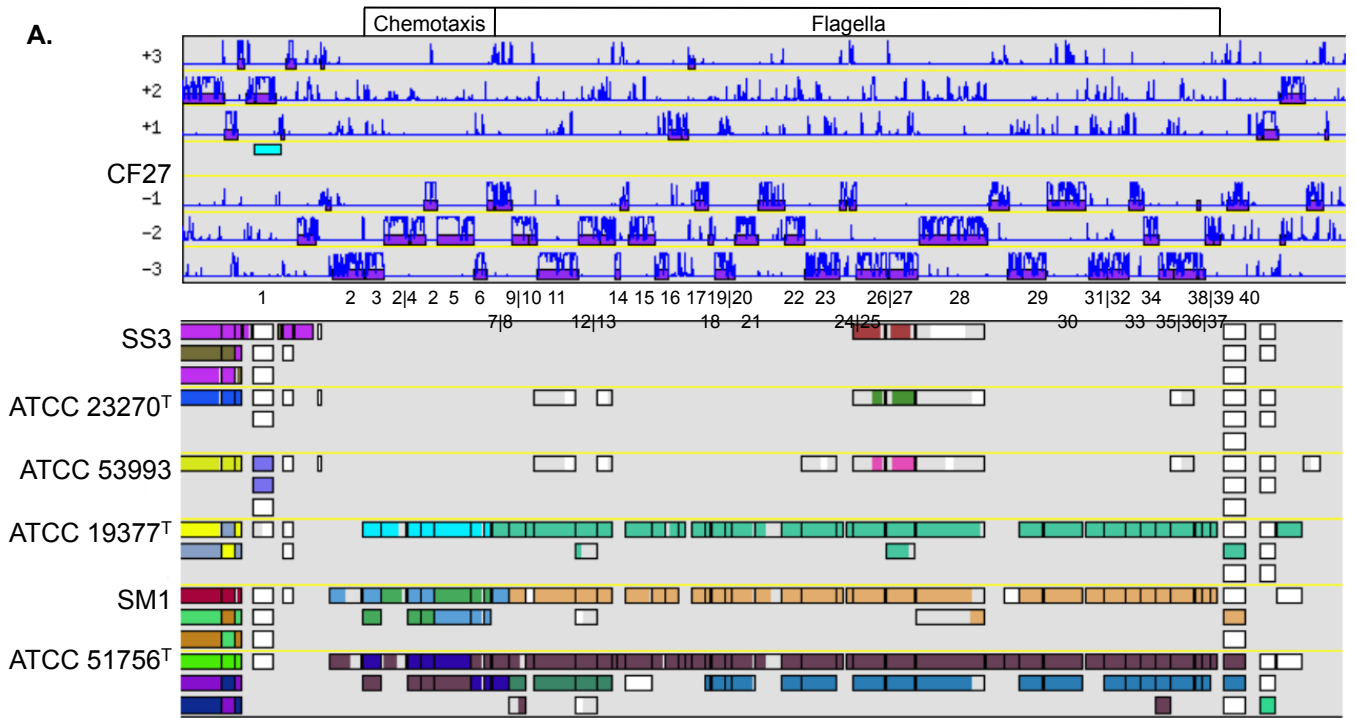


Figure S3. A. Synteny of chemotaxis and flagellar biosynthesis genes between *Acidithiobacillus* species. The upper panel shows the cluster of genes involved in chemotaxis and flagellar biosynthesis in *At. ferrivorans* CF27 (see **Supplementary Table S5** for AFERRI_v2_numbers) while the lower panel shows these clusters in the other *Acidithiobacillus* genomes. In between the two panels are given numbers corresponding to the genes with the following functional descriptions: 1, transposase; 2, methyl-accepting chemotaxis sensory transducers; 3, chemotaxis protein (CheV); 4, flavoprotein; 5, signal transduction histidine kinase (CheA); 6, putative chemotaxis phosphatase (CheZ); 7, chemotaxis regulator transmitting signal to flagellar motor component (CheY); 8, flagellar motor rotation protein (MotB); 9, flagellar motor component (MotA); 10, flagellar basal body-associated protein (FliL); 11, putative flagellar export pore protein (FlhA); 12, flagellar biosynthesis protein (FlhB); 13, putative flagellar biosynthetic protein (FliR); 14, putative flagellar biosynthetic protein (FliQ); 15, flagellum-specific ATP synthase (FliI); 16, putative flagellar assembly protein (FliH); 17, putative RNA polymerase sigma factor WhiG (FliA); 18, flagellar biosynthesis protein (FliQ); 19, flagellar biosynthesis protein (FliP); 20, flagellar motor switch protein (FliN); 21, flagellar motor switch protein (FliM); 22, flagellar motor switch protein (FliG); 23, flagellar M-ring protein (FliF); 24, flagellar basal-body component (FliE); 25, flagellin-specific chaperone (FliS); 26, flagellar hook-associated protein (FliD); 27, flagellin protein (FliC); 28, glycosyltransferase; 29, flagellar hook-associated protein (FlgL); 30, flagellar hook-associated protein (FlgK); 31, putative peptidoglycan hydrolase (FlgJ); 32, flagellar basal body P-ring protein (FlgI); 33, flagellar L-ring protein (FlgH); 34, flagellar component of cell-distal portion of basal-body rod (FlgG); 35, flagellar basal-body rod protein (FlgF); 36, flagellar hook protein (FlgE); 37, flagellar hook capping protein (FlgD); 38, flagellar component of cell-proximal portion of basal-body rod (FlgC); 39, flagellar basal body rod protein (FlgB); 40, transcriptional regulator. The genes encoding proteins of unknown function are not indicated. The same colour tone indicates the same localization. **B. Model of chemotaxis and flagella in *At. ferrivorans* CF27.** Flagellar picture was adapted from Liu and Ochaman (2007).

I- Common genes in *Acidithiobacillus* except CF27

ATCC 23270[20]/ ATCC 53993[25]/ ATCC 19377[20] / SM1[20]/ ATCC 51756[20]/ SS3[19]

- Putative phosphate transport proteins
- Cytochrome c-552 (modular protein)
- Proteins of unknown function
- ...

II- Common genes in *Acidithiobacillus* except SS3

ATCC 23270[20]/ ATCC 53993[25]/ ATCC 19377[20] / SM1[20]/ ATCC 51756[20]/ CF27[19]

- Na⁺/H⁺ antiporter
- Type IV pilin
- D12 class N6 adenine-specific DNA methyltransferase
- Proteins of unknown function
- ...

III- Common genes except *At. ferrivorans* (*At. ferrooxidans* + *At. thiooxidans* + *At. caldus*)

ATCC 23270[56]/ ATCC 53993[57]/ ATCC 19377[58] / SM1[62]/ ATCC 51756[67]

- TonB-dependent receptor
- Putative MotA/TolQ/ExbB proton channel
- Putative Biopolymer transport exbD protein
- Putative TonB protein
- C4-dicarboxylate transporter
- Multimodular transpeptidase-transglycosylase
- Heat shock protein Hsp20
- ABC transporter, ATP-binding protein
- Putative Quercetin 2,3-dioxygenase
- Cytochrome *d* terminal oxidase
- Proteins of unknown function
- ...



IV- Core *Acidithiobacillus* genome (All *Acidithiobacillus* strains)

CF27[1,466] (37.7%) ATCC 23270[1,442] (40.0%) ATCC 19377[1,441] (46.1%) SM1[1,486] (41.4%)
 SS3[1,461] (38.6%) ATCC 53993[1,446] (44.8%) ATCC 51756[1,479] (50.4%)

V- Common genes in *At. ferrivorans* except other *Acidithiobacillus*

CF27[258]/ SS3[257]

- Cellulose synthase operon C domain protein
- Tetratricopeptide TPR_1 repeat-containing protein
- Endo-1,4-D-glucanase
- Cellulose synthase catalytic subunit

- Response regulator receiver modulated diguanylate cyclase/phosphodiesterase
- Putative transcriptional regulator, Crp/Fnr family
- Regulatory protein ArsR
- Putative Crp/Fnr family transcriptional regulator
- Protein of unknown function
- Putative circadian clock protein, KaiC

- Class II aldolase/adducin family protein
- Transcriptional regulator, PadR-like family
- Carboxymuconolactone decarboxylase
- Rhodanese-like protein

- Putative PAS/PAC sensor protein
- Sulfate transporter/antisigma-factor antagonist STAS
- Ammonium transporter

- Diguanylate cyclase/phosphodiesterase
- 4Fe-4S ferredoxin iron-sulfur binding domain-containing protein
- PBS lyase HEAT domain protein repeat-containing protein
- Catalase
- Transposase
- Lipid A biosynthesis acyltransferase
- Cytochrome *c*₅₅₁ peroxidase
- Ribosylidihydroxycotinamide dehydrogenase [quinone]
- Acyltransferase 3
- G-D-S-L family lipolytic protein
- (Ni/Fe) hydrogenase, *b*-type cytochrome subunit
- Rieske (2Fe-2S) iron-sulfur domain-containing protein
- Cytochrome *c* prime
- Nickel-dependent hydrogenase, *b*-type cytochrome subunit
- ABC transporter permease
- UbiA prenyltransferase
- Glyoxalase/bleomycin resistance protein/dioxygenase
- Transcriptional modulator of MazE/toxin
- Nitrogen regulatory protein P-II
- Cold shock DNA-binding domain-containing protein
- Trehalose synthesis
- Proteins of unknown function
- ...

Figure S4. Venn diagram of orthologous groups (OG) from seven *Acidithiobacillus* genomes. The strain names are: CF27, *At. ferrivorans* CF27; SS3, *At. ferrivorans* SS3; ATCC 23270, *At. ferrooxidans* ATCC 23270^T; ATCC 53993, *At. ferrooxidans* ATCC 53993; ATCC 19377, *At. thiooxidans* ATCC 19377^T; SM1, *At. caldus* SM-1; and ATCC 51756, *At. caldus* ATCC 51756^T. OGs are displayed by the roman numerals (from I to IV). The numbers in brackets indicate the number of genes in each strain and their percentage (compared to the corresponding proteome) are shown in parentheses. Main functional proteins are shown in colour boxes and gene cluster is shown in an open bracket.