Genome Announcement

Complete genome sequence of *Metallosphaera cuprina*, a metal sulfide oxidizing archaeon from hotspring

Li-Jun Liu¹, Xiao-Yan You¹2, Hua-Jun Zheng³, Sheng-Yue Wang³, Cheng-Ying Jiang¹²*, Shuang-Jiang Liu¹2*

State Key Laboratory of Microbial Resource¹, Environmental Microbiology Research Center², at Institute of Microbiology, Beijing 100101; Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai³, People’s Republic of China

Running title: Complete genome of *Metallosphaera cuprina*

*For correspondence:
Dr. Shuang-Jiang Liu, Institute of Microbiology, Chinese Academy of Sciences, Bei-Chen-Xi-Lu No. 1, Chao-Yang Distrct, Beijing 100101, P. R. China.
Tel: +86-10-64807423, Fax: +86-10-64807421.
Email: liusj@sun.im.ac.cn

Or

Dr. Cheng-Ying Jiang
Institute of Microbiology, Chinese Academy of Sciences, Bei-Chen-Xi-Lu No. 1, Chao-Yang Distrct, Beijing 100101, P. R. China.
Tel: +86-10-64807581, Fax: +86-10-64807421.
Email: jiangey@im.ac.cn
Abstract:

The genome of metal sulfide oxidizing, thermoacidiphilic *Metallosphaera cuprina* Ar-4 has been completely sequenced and annotated. Originally isolated from a sulfuric hotspring, strain Ar-4 grows optimally at 65 °C and pH of 3.5. The *M. cuprina* genome has a 1,840,348 bp circular chromosome (2029 ORFs) and is 16% smaller than the previously sequenced *Metallosphaera sedula* genome. Compared to the *M. sedula* genome, there are no counterpart genes in the *M. cuprina* genome for about 480 ORFs in the *M. sedula* genome, of which 243 ORFs are annotated as hypothetical proteins. Still, there are 235 ORFs uniquely occurring in *M. cuprina*. Genome annotation supports that *M. cuprina* lives a facultative life on CO₂ and organics, and obtains energy from oxidation of sulfidic ores and reduced inorganic sulfuric compounds.

Extremely thermoacidophilic archaea play important roles in mobilizing metal sulfide deposits in natural bioleaching environments (5, 9). Due to the ability to oxidize RISCs under high temperature, *Metallosphaera* has attracted increasing interest from biomining industry (5, 10-13). The bioleaching *Metallosphaera sedula* were explored at genomic level (2). Here we present the complete genome of a newly isolated, bioleaching and thermoacidophilic *Metallosphaera cuprina* (7).

Genomic DNA of *M. cuprina* Ar-4 was purified from cells grown in modified Allen medium (3). Whole-genome was sequenced by the Roche 454 Genome
Sequencer FLX instrument. A total of 295,139 shotgun reads were produced, and assembled into 55 contigs, providing a 67-fold coverage. Gaps were closed by multiplex PCR and primer walking methods. The gap-spanning PCR products were sequenced with an ABI 3730 DNA Analyzer and the resulted sequences were assembled using phred/phrap/consed software. The final consensus quality of each base was above 64. Protein-coding genes were identified with Glimmer 3.02 program (4). Protein function was predicted by either homology searches in the GenBank and UniProt protein database or function assignment searches in CDD (COG) database or domain/motif searches in the Pfam databases. The KEGG tool was used to reconstruct metabolic pathways. Membrane proteins were predicted by LipoP, SignalP, and ConPred II programs. The tRNA genes were identified by using tRNAscan-SE tool and the rRNA genes were identified by using RNAmmer-1.2 and BLASTN program.

The *M. cuprina* Ar-4 grew chemolithotrophically on CO₂ with metal sulfide and RISCs as energy sources, or chemoheterotrophically on various organics (8). Its genome consisted of a 1,840,348 bp circular chromosome. The genome carried totally 2029 ORFs. Genome annotation and metabolic reconstruction supported that *M. cuprina* lived a facultative life. The *M. cuprina* fixed CO₂ via 3-hydroxypropionate/4-hydroxybutrate cycle, and it assimilated carbohydrates via the non-phosphorylated ED pathway. It had a complete TCA pathway and incomplete phosphate pentose pathway. Oxidation of RISCs by heterodisulfide reductase complex, sulfide:quinone oxidoreductase, thiosulfate:quinone oxidoreductase, tetrathionate hydrolase, and sulfite:acceptor oxidoreductase in *M. cuprina* was proposed. The
terminal oxidase complexes of *M. cuprina* that channels electrons from RISCs oxidation to oxygen were similar to those of the “*Metallosphaera yellowstonensis*” (7) and *M. sedula* (1).

The *M. cuprina* genome was 16% smaller than *M. sedula* genome. Analysis indicated that the counterpart genes of about 480 ORFs in the *M. sedula* genome were not found in the *M. cuprina* genome. Still, there were 235 ORFs uniquely occurring in *M. cuprina*. Most of those ORFs were annotated as hypothetical proteins. Gene redundancy in *M. cuprina* apparently kept low. For example, there was only one copy of the 4-hydroxybutyryl-CoA dehydratase gene in *M. cuprina*, but duplication of this function was observed in *M. sedula* genome (2). The information provided in the *M. cuprina* genome sequence will facilitate additional researches on this organism, as well as defining the core genome and key physiological features of the genus *Metallosphaera*.

**Nucleotide sequence accession number.** The *M. cuprina* genome sequence is available at GenBank under Accession No. CP002656.

**Acknowledgements**

The work was supported by Ministry of Science and Technology (973 program, 2010CB630903) and National Science Foundation of China (31070042, 30921065).

**REFERENCES**


