Genome Sequence of *Sphingomonas* sp. S17, Isolated from an Alkaline, Hyperarsenic, and Hypersaline Volcano-Associated Lake at High Altitude in the Argentinean Puna[▽]

Maria Eugenia Farias,[1] Santiago Revale,[2] Estefania Mancini,[2] Omar Ordoñez,[1]
Adrian Turjanski,[3] Néstor Cortez,[4] and Martin P. Vazquez[2]*

*Laboratorio de Investigaciones Microbiológicas de Lagunas Andinas (LIMLA), Planta Piloto de Procesos Industriales Microbiológicos
(PROIMI), CCT-CONICET, Tucumán, Argentina[1]; Instituto de Agrobiotecnología Rosario (INDEAR)-CONICET,
Ocampo 210 bis, CCT-CONICET, Rosario, Argentina[2]; Facultad de Ciencias Exactas y Naturales,
Universidad de Buenos Aires, Buenos Aires, Argentina[3]; and Instituto de Biología Molecular y
Celular de Rosario (IBR-UNR-CONICET), Rosario, Argentina[4]*

The high-altitude Andean lakes (HAAL) in the Argentinean Puna-high Andes region represent an almost unexplored ecosystem exposed to extreme conditions (high UV irradiation, hypersalinity, drastic temperature changes, desiccation, and high pH). Here we present the first genome sequence, a *Sphingomonas* sp., isolated from this extreme environment.

Sphingomonas sp. S17 was isolated from a modern stromatolite community in Socompa Lake (3,800-m altitude), placed near the active volcano Socompa in northwest Argentina. The environmental conditions of these lakes include a pH of 9, a salinity of 22%, and an arsenic content of 32 mg/liter (2, 3, 4, 5, 11).

The genome sequence was obtained using a whole-genome shotgun (WGS) strategy with a 454 GS Titanium pyrosequencer at INDEAR, Argentina. Assembly was done using 454 Newbler, version 2.5.3, using the –urt option with $63\times$ genome coverage. This assembly generated 62 large contigs. The draft genome was 4,268,406 bases in length, with a mean GC content of 65.74%.

Genome annotation was done using the standard operating procedures (SOPs) for prokaryotic annotation from ISGA (7) and from the RAST annotation server (1). A total of 3,892 coding sequences (CDS) and 52 structural RNAs (49 tRNAs) were predicted. Annotation covered 340 RAST subsystems (43%) with 1,650 CDS, while 1,251 CDS (32%) were classified as hypothetical proteins. The complete 16S rRNA gene presented a maximum identity of 95% with that of *Sphingomonas wittichii* (10).

Consistent with the extreme environment and high UV irradiation, *Sphingomonas* sp. S17 presented a complete DNA repair system, including a photolyase gene and a SOS regulatory system, with the *umuCD* operon harbored for translesion synthesis. Neither of these CDS was found in the reference genome of *Sphingomonas wittichii* RW1 (10). Accordingly, this bacterial genome also presented 24 genes devoted to sulfur metabolism, compared with only 4 such genes in *S. wittichii*. A set of 95 genes was present in the subsystem for resistance to antibiotics and toxic compounds, compared with only 41 such genes in *S. wittichii*. The majority of these genes are devoted to the resistance of arsenic, chromium, and fluoroquinolones, with 18 genes devoted to multidrug resistance efflux pumps. The genome also contained two copies of a NhaA-type CDS for the Na^+/H^+ antiporter (12) and six subunits of the multisubunit cation antiporter (Na^+/H^+), compatible with its alkaline and hypersaline environment.

As in other *Sphingomonas* genomes, a rich set of genes (47 CDS) for metabolism of aromatic compounds was present in the subsystem. It is interesting to note two other features absent from the *S. wittichii* genome: (i) the presence of 5 CDS related to the prokaryotic immune system CRISPR (one Cas1, one Cas2, two Cas3, and one Csn1 CDS), with seven associated direct repeats of 36 bases separated by spacers of 31 bases (8), and (ii) the presence of the operon containing the LodA and LodB CDS for the production of marinocine, a broad-spectrum antibacterial protein (9).

Finally, *Sphingomonas* sp. S17 contains at least one megaplasmid of the IncF type, with a complete set of mobilization proteins (variant 1 in the conjugative transfer subsystem; RAST annotation) (6). The megaplasmid contains several resolvase/integrase proteins (8 CDS) and transposases (5 CDS). Interestingly, it also harbors the *umuCD* operon.

This represents the first genome sequence reported from the unexplored extreme high-altitude Andean lake (HAAL) ecosystem.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession number AFGG00000000. The version described in this paper is the first version, accession number AFGG01000000.

* Corresponding author. Mailing address: INDEAR-CONICET, Ocampo 210 bis, CCT-CONICET, Rosario, Argentina. Phone: (54 341) 4448535. Fax: (54 341) 4448557. E-mail: martin.vazquez@indear .com.
▽ Published ahead of print on 20 May 2011.

REFERENCES

1. **Aziz, R. K., et al.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9:**75.
2. **Dib, J. R., et al.** 2009. Isolation of bacteria from remote high altitude Andean wetlands able to grow in the presence of antibiotics. Recent Pat. Antiinfect. Drug Discov. **4:**66–76.
3. **Dib, J. R., V. Fenandez Zenoff, J. Motok, O. F. Ordoñez, and M. E. Farías.** 2008. Occurrence of resistance to antibiotics, UV-B and arsenic in bacteria isolated from extreme environment in high altitude Andean wetlands (4,560 m and 4,400 m altitude). Curr. Microbiol. **56:**510–517.
4. **Fernandez Zenoff, V., F. Siñeriz, and M. E. Farías.** 2006. Bacteria isolated from Andean high altitude lakes display diverse responses to UV-B radiation and repair mechanisms. Appl. Environ. Microbiol. **72:**7857–7863.
5. **Flores, M. R., O. F. Ordoñez, and M. E. Farías.** 2009. Isolation of UV-B resistant bacteria from two Andean wetlands (4,400 m) with saline and hypersaline conditions. J. Gen. Appl. Microbiol. **55:**447–458.
6. **Frost, L. S., and G. Koraimann.** 2010. Regulation of bacterial conjugation: balancing opportunity with adversity. Future Microbiol. **5:**1057–1071.
7. **Hemmerich, C., A. Buechlein, R. Podicheti, K. V. Revanna, and Q. Dong.** 2010. An Ergatis-based prokaryotic genome annotation web server. Bioinformatics **26:**1122–1124.
8. **Horvath, P., and R. Barrangou.** 2010. CRISPR/Cas, the immune system of bacteria and archaea. Science **327:**167–170.
9. **Lucas-Elio, P., P. Hernandez, A. Sanchez-Amat, and F. Solano.** 2005. Purification and partial characterization of marinocine, a new broad-spectrum antibacterial protein produced by *Marinomonas mediterranea*. Biochim. Biophys. Acta **1721:**193–203.
10. **Miller, T. R., et al.** 2010. Genome sequence of the dioxin-mineralizing bacterium *Sphingomonas wittichii* RW1. J. Bacteriol. **192:**6101–6102.
11. **Ordoñez, O. F., M. R. Flores, J. R. Dib, A. Paz, and M. E. Farías.** 2009. Extremophile culture collection from Andean lakes: extreme pristine environments that host a wide diversity of microorganisms with tolerance to UV radiation. Microb. Ecol. **58:**461–473.
12. **Padan, E., et al.** 2004. NhaA of Escherichia coli, as a model of a pH-regulated Na^+/H^+ antiporter. Biochim. Biophys. Acta **1658:**2–13.