Genome Sequence of “Pedosphaera parvula” Ellin514, an Aerobic Verrucomicrobial Isolate from Pasture Soil

Ravi Kant,1 Mark W. J. van Passel,2 Parveen Sangwan,3 Airi Palva,1 Susan Lucas,4,5 Alex Copeland,4,6 Alla Lapidus,4,6† Tijana Glavina del Rio,4,6 Eileen Dalin,4,6‡ Hope Tice,4,6 David Bruce,7 Lynne Goodwin,7 Sam Pitluck,4,6 Olga Chertkov,1 Frank W. Larimer,8 Miriam L. Land,8 Loren Hauser,8 Thomas S. Brettn,8 John C. Detter,7 Shunsheng Han,7 Willem M. de Vos,1,2 Peter H. Janssen,9 and Hauke Smidt2*  

Department of Veterinary Biosciences, Faculty of Veterinary Medicine, FIN-00014 University of Helsinki, Finland; Laboratory of Microbiology, Wageningen University, Wageningen, Netherlands2; CSIRO Materials Science and Engineering, Private Bag 33, Clayton South, Victoria 3169, Australia; DOE Joint Genome Institute, Walnut Creek, California 94598; Lawrence Livermore National Laboratory, Livermore, California 94550; Lawrence Berkeley National Laboratory, Berkeley, California 94720; Los Alamos National Laboratory, Los Alamos, New Mexico 87545; Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831; and AgResearch Ltd., Grasslands Research Centre, Palmerston North, New Zealand9

Received 3 March 2011/Accepted 25 March 2011

“Pedosphaera parvula” Ellin514 is an aerobically grown verrucomicrobial isolate from pasture soil. It is one of the few cultured representatives of subdivision 3 of the phylum Verrucomicrobia. Members of this group are widespread in terrestrial environments.

Based on 16S rRNA gene-based surveys, as well as analyses of metagenomic libraries, members of the phylum Verrucomicrobia have been identified as prevalent colonizers of numerous different environments, ranging from aquatic to terrestrial ecosystems (10) to the mammalian gastrointestinal tract (1, 9).

Verrucomicrobiunm spinosum, Prosthecobacter deijongei, Chthonobacter flavus, and “Pedosphaera parvula” Ellin514 represent two classes (Verrucomicrobiae and Spartobacteria) and four genera within the Verrucomicrobia. Examination of these species after high-pressure freezing and cryosubstitution has shown cellular features that are consistent with compartmentalization of the cell cytoplasm by internal coated membranes (4).

Apart from eukaryotes, membrane coat (MC)-like proteins are found solely in the Planctomycetes-Verrucomicrobia-Chlamydiae bacterial superphylum. MC proteins that play key roles in shaping eukaryotic membranes were most likely present in the ancestral eukaryote but cannot be detected in prokaryotes using sequence-driven searches (8). “P. parvula” Ellin514, a representative of Verrucomicrobia subdivision 3 (7), possesses a major pirellulosome compartment separated by an intracytoplasmic membrane from an outer paraphoplasm (4), analogous to the planctomycete cell plan (5), and contains nine putative MC proteins (8). The MC proteins may be associated with the internal membrane structures (8).

The genome of strain Ellin514 was elucidated in the frame of a larger project encompassing five genomes within the Verrucomicrobia and Lentisphaerae, providing the necessary basis for propelling our knowledge in the fields of taxonomy and evolution for organisms from these phyla. Genome sequences will also inform studies of their functionality in aquatic, terrestrial, and intestinal environments through comparative and functional genomics.

“P. parvula” Ellin514, an obligate aerobic bacterium, was isolated from pasture soil from the Dairy Research Institute, Ellinbank, Victoria, Australia (7). Its genome was sequenced using a combination of Sanger and 454 sequencing platforms (GS FLX). All general aspects of library construction and sequencing can be found at the JGI website (http://www.jgi.doe.gov). Possible misassemblies were corrected, and gaps between contigs were closed by editing in Consed, by custom primer walks from subclones or PCR products. The combination of the Sanger and 454 sequencing platforms provided 24.73-fold coverage of the genome. The gene modeling program Prodigal (3) was run on the permanent draft genome, using default settings that permit overlapping genes and using ATG, GTG, and TTG as potential starts. The resulting protein translations were compared to GenBank’s nonredundant database and the Swiss-Prot/TrEMBL, PRIAM, Pfam, TIGRFam, Interpro, KEGG, and COGs databases using BLASTP or HMMER. From these results, product assignments were made, followed by manual corrections (6).

The unclosed draft genome of “P. parvula” Ellin514 contains 7,414,222 nucleotides. The overall G+C content of the chromosome is 52.6%. The chromosome contains 6,510 protein-coding sequences (CDS), 59 tRNA genes, and 4 rRNAs. Putative functions could be predicted for 3,483 CDS (53%), and 3,507 (53.35%) protein-coding genes were assigned to clusters of orthologous groups of proteins (COGs). The high percentage of CDS with predicted signal peptides (2,682, 40.8%) in-
icates an extensive array of interactions of “P. parvula” Ellin514 with its environment.

**Nucleotide sequence accession number.** The draft genome sequence of “P. parvula” Ellin514 has been assigned GenBank accession number ABOX00000000.

R.K. was supported by Center of Excellence in Microbial Food Safety Research (MiFoSa), Academy of Finland. M.W.J.V.P. is funded by the Netherlands Organization for Scientific Research (NWO) via a VENI grant.

**REFERENCES**

2. Reference deleted.