Complete genome sequence of the rapeseed plant-growth promoting Serratia plymuthica strain AS9

Saraswoti Neupane1, Nils Högberg1, Sadhna Alström1, Susan Lucas2, James Han2, Alla Lapidus2, Jan-Fang Cheng2, David Bruce2,3, Lynne Goodwin2,3, Sam Pitluck2, Lin Peters2, Galina Ovchinnikova2, Megan Lu2,3, Cliff Han2,3, John C. Detter2,3, Roxanne Tapia2,3, Anne Fiebig5, Miriam Land2,4, Loren Hauser2,4, Nikos C. Kyrpides2, Natalia Ivanova2, Ioanna Pagani2, Hans-Peter Klenk2, Tanja Woyke2 and Roger D. Finlay1

1 Uppsala BioCenter, Department of Forest Mycology and Pathology, Swedish University of Agricultural Sciences, Uppsala, Sweden
2 DOE Joint Genome Institute, Walnut Creek, California, USA,
3 Los Alamos National Laboratory, Bioscience Division, Los Alamos, New Mexico, USA,
4 Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
5 Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany

Corresponding author: Saraswoti Neupane, saraswoti.neupane@slu.se

Keywords: motile, non-sporulating, mesophile, Gram-negative, free living, plant-associated, chemoorganotrophic, Enterobacteriaceae, CSP 2010

Introduction

The genus Serratia belongs to a group of Gammaproteobacteria, commonly found in soil, water, plants, insects and humans [1]. The genus includes antagonists of soil borne pathogens of different plant species, plant growth promoters and insect pathogens, as well as opportunistic human pathogens. The most common human pathogen in this genus is Serratia marcescens which causes nosocomial infections in humans, while other species are harmless. In agriculture, S. plymuthica is successfully used for control of many soil borne fungal pathogens of different crops (e.g. strawberry, rapeseed) [2,3], while S. proteamaculans promotes the growth of poplar trees [4].

S. plymuthica AS9 (= CCUG 61396) was isolated from field samples of rapeseed roots in Uppsala, Sweden. Our interest in S. plymuthica AS9 is attributed to its ability to stimulate rapeseed plant growth, to inhibit soil borne fungal pathogens and to increase oilseed production. Here we present a description of the complete genome sequencing of S. plymuthica AS9 and its annotation.

Classification and features

The bacterial strain AS9 was previously considered a member of the family Enterobacteriaceae [5]. Recently, comparison of 16S rRNA gene sequences with the most recent databases from GenBank using NCBI BLAST [6] under default settings showed that S. plymuthica AS9 shares 99% similarity with many Serratia species including S. plymuthica (AJ233433) and Serratia proteamaculans (CP000826.1). When considering high-scoring segment pairs (HSPs) from the best
250 hits, the most frequent matches were with various Serratia species (17.2% with maximum identity of 97-100%) with S. plymuthica (5.2% with maximum identity of 97-99%), S. proteamaculans (4.8% with maximum identity of 97-99%), S. marcescens (4.8% with maximum identity of 96-97%) and various Rahnella species. (7% with maximum identity of 97-98%).

Figure 1 shows the phylogenetic relationship of S. plymuthica AS9 with other species within the genus Serratia in a 16S rRNA based tree. The tree shows its close relationship with the type strain of S. plymuthica, which was confirmed by digital DNA-DNA hybridization values [11] above 70% with the (unpublished) draft genome sequence of the S. plymuthica type strain Breed K-7T from a DSM4540 culture using the GGDC web server [12].

S. plymuthica AS9 is a Gram-negative, rod shaped, motile bacterium, 1-2 µm long and 0.5-0.7 µm wide (Figure 2 and Table 1). It forms red to pink colored colonies 1-2 mm in diameter on tryptic soy agar and potato dextrose agar. The color of the bacterium is the result of its production of the red pigment, prodigiosin, but the colony color or production of pigment depends on the ingredients, pH of the medium and the incubation temperature [26-28]. S. plymuthica is a facultative anaerobe, grows between 4 °C and 40 °C and within the pH range 4 - 10. It can utilize a wide range of carbon sources and also has chitinolytic, proteolytic, cellulolytic, and phospholytic activity [5].

Chemotaxonomy
The whole cell lipid pattern of S. plymuthica AS9 contains a mixture of saturated and unsaturated fatty acids. The main fatty acids in AS9 strain comprise C16:0 (24.13%), C16:1ω7c (19.41%), C18:1ω7c (18.76%), C14:0 (5.24%) along with other minor fatty acid components. Previously it has been shown that Serratia spp. contain a mixture of C14:0, C16:0, C16:1 and C18:1+2 fatty acids of which 50-80% of the total was C14:0 and other were less than 3% each [29]. This is consistent with the fact that the C14:0 3OH is characteristic of the family Enterobacteriaceae.

Figure 1. Phylogenetic tree highlighting the position of S. plymuthica AS9 in relation to other species within the genus Serratia, which is based on 1,479 characters of the 16S rRNA gene sequence aligned in ClustalW2 [7]. The tree was inferred under the maximum likelihood criterion [MEGA5, 8] and rooted with Yersinia pseudotuberculosis (a member of the family Enterobacteriaceae). The branches are scaled in terms of the expected number of substitutions per site. The numbers above branches are support values from 1,000 bootstrap replicates if larger than 60% [9]. Lineages with type strain genome sequences registered in GOLD [10] are shown in blue.
**Figure 2.** Scanning electron micrograph of *S. plymuthica* AS9

**Table 1.** Classification and general features of *S. plymuthica* AS9 according to the MIGS recommendations [13]

| MIGS ID | Property                | Term                                                                 | Evidence code |
|---------|-------------------------|----------------------------------------------------------------------|---------------|
|         | Current classification  | Domain *Bacteria*                                                      | TAS [14]      |
|         |                         | Phylum *Proteobacteria*                                               | TAS [15]      |
|         |                         | Class *Gammaproteobacteria*                                           | TAS [15,16]   |
|         |                         | Order “*Enterobacterales*”                                            | TAS [17]      |
|         |                         | Family *Enterobacteriaceae*                                           | TAS [18-20]   |
|         |                         | Genus *Serratia*                                                      | TAS [18,21,22]|
|         |                         | Species *Serratia plymuthica*                                         | TAS [18,23]   |
|         | Strain AS9              |                                                                       | IDA           |
|         | Gram stain              | Negative                                                              | IDA           |
|         | Cell shape              | Rod-shaped                                                            | IDA           |
|         | Motility                | Motile                                                                | IDA           |
|         | Sporulation             | Non-sporulating                                                       | IDA           |
|         | Temperature range        | Mesophilic                                                            | IDA           |
|         | Optimum temperature     | 28°C                                                                  | IDA           |
|         | Carbon source           | Glucose, mannitol, sucrose, arabinose, cellobiose                     | IDA           |
|         | Energy metabolism       | Chemoorganotrophic                                                   | NAS           |
|         | Terminal electron receptor | --                        | NAS           |
| MIGS-6  | Habitat                 | Rapeseed roots                                                        | NAS           |
| MIGS-6.3| Salinity                | Medium                                                                | IDA           |
| MIGS-22 | Oxygen                  | Facultative                                                           | IDA           |
| MIGS-15 | Biotic relationship     | Free living                                                           | NAS           |
| MIGS-14 | Pathogenicity           | Non-pathogenic                                                        | IDA           |
|         | Biosafety level         | 1+                                                                    | TAS [24]      |
| MIGS-4  | Geographic location     | Uppsala, Sweden                                                       | NAS           |
| MIGS-5  | Sample collection time  | Summer 1998                                                           | NAS           |
| MIGS-4.1| Latitude                | 59.8                                                                  | NAS           |
| MIGS-4.2| Longitude               | 17.65                                                                 | NAS           |
| MIGS-4.3| Depth                  | 0.1 m                                                                  | NAS           |
| MIGS-4.4| Altitude               | 24-25 m                                                                | NAS           |

a) Evidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [25]. If the evidence code is IDA, then the property was observed by one of the authors, or an expert mentioned in the acknowledgements.
Genome sequencing information

*S. plymuthica* AS9, one of the strains isolated from rapeseed roots and rhizosphere soils was selected for sequencing on the basis of its ability to promote rapeseed growth and inhibit soil borne fungal pathogens. The genome project is deposited in the Genomes On Line Databases [10] and the complete genome sequence is deposited in GenBank. Sequencing, finishing and annotation were performed by the DOE Joint Genome Institute (JGI). A summary of the project information is shown in Table 2 and its association with MIGS identifiers.

Growth conditions and DNA isolation

*S. plymuthica* AS9 was grown in Luria Broth (LB) medium at 28°C for 12 hours (cells were in the early stationary phase) and the DNA was isolated using a standard CTAB protocol for bacterial genomic DNA isolation which is available at JGI [30].

Table 2. Genome sequencing project information

| MIGS ID  | Property          | Term                                                                 |
|---------|-------------------|----------------------------------------------------------------------|
| MIGS-31 | Finishing quality | Finished                                                             |
| MIGS-28 | Libraries used    | Three libraries: one 454 standard library, one 454 PE library (12.5 kb insert size), one Illumina library |
| MIGS-29 | Sequencing platforms | Illumina GAii, 454 GS FLX Titanium                                    |
| MIGS-31.2| Sequencing coverage | 323.5 × Illumina; 8.8 × pyrosequencing                               |
| MIGS-30 | Assemblers        | Velvet v. 0.7.63, Newbler v. 2.3 pre-release, phrap version SPS – 4.24 |
| MIGS-32 | Gene calling method | Prodigal 1.4, GenePRIMP                                               |
| NCBI project ID | 60457          |                                                                      |
| INSDC ID    | CP002773         |                                                                      |
| Genbank Date of Release | October 12, 2011 |                                                                      |
| GOLD ID     | Gc01772           |                                                                      |
| MIGS-13 | Source material identifier | CCUG 61396                                                  |
| Project relevance | Biocontrol, Agricultural |                                                  |

Genome sequencing and assembly

The genome of strain AS9 was sequenced using a combination of Illumina [31] and 454 sequencing platforms [32]. The details of library construction and sequencing are available at the JGI website [30]. The sequence data from Illumina GAii (1,790.7 Mb) were assembled with Velvet [33] and the consensus sequence computationally shredded into 1.5 kb overlapping fake reads. The sequencing data from 454 pyrosequencing (102.2 Mb) were assembled with Newbler (Roche). The initial draft assembly contained 41 contigs in one scaffold and consensus sequences were computationally shredded into 2 kb overlapping fake reads. The 454 Newbler consensus reads, the Illumina velvet consensus reads and the read pairs in the 454 paired end library were integrated using a software phrap (High Performance Software, LLC) [34]. Possible mis-assemblies were corrected with gapResolution [30], Dupfinisher [35], or by sequencing cloned bridging PCR fragments with subcloning or transposon bombing (Epicentre Biotechnologies, Madison, WI). The gaps between contigs were closed by editing in the software Consed [36-38], by PCR and by Bubble PCR (J.-F. Chang, unpublished) primer walks. Thirty seven additional reactions were necessary to close gaps and to raise the quality of the finished sequence. The sequence reads from Illumina were used to correct potential base errors and increase consensus quality using the software Polisher, developed at JGI [39]. The final assembly is based on 47.3 Mb of 454 draft data which provides an average 8.8× coverage of the genome and 1,746.8 Mb of Illumina draft data which provides an average 323.5× coverage of the genome.

Genome annotation

Genes were identified using Prodigal [40] as part of the genome annotation pipeline at Oak Ridge National Laboratory (ORNL), Oak Ridge, TN, USA, followed by a round of manual curation using the JGI GenPRIMP pipeline [41]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) non-redundant database, Uniport, TIGR-Fam, Pfam,
PRIAM, KEGG, COG and InterPro databases. The tRNAscanSE tool [42] was used to find tRNA genes. Additional gene prediction analysis and functional annotation were performed within the Integrated Microbial Genomes – Expert Review (IMG-ER) platform [43].

**Genome properties**

The *S. plymuthica* AS9 genome includes a single circular chromosome of 5,442,880 bp with 55.96% GC content. The genome had 5,139 predicted genes of which 4,952 were assigned as protein-coding genes, 113 RNA genes and 75 pseudogenes [Figure 3]. The majority of protein coding genes (87.42%) was assigned as a putative function while those remaining were annotated as hypothetical proteins [Table 3]. The distribution into COG functional categories is presented in Table 4.

![Graphical circular map of the chromosome. From outside to the center: Genes on forward strand (color by COG categories), Genes on reverse strand (color by COG categories), RNA genes (tRNAs green, rRNAs red, other RNAs black), GC content, GC skew.](image-url)
Table 3. Genome statistics

| Attribute                           | Value   | % of totala |
|-------------------------------------|---------|-------------|
| Genome size (bp)                    | 5,442,880 | 100.00%     |
| DNA Coding region (bp)              | 4,739,233 | 87.07%      |
| DNA G+C content (bp)                | 3,045,898 | 55.96%      |
| Total genesb                         | 5,139   | 100.00%     |
| RNA genes                           | 113     | 2.19%       |
| rRNA operons                        | 7       |             |
| Protein-coding genes                | 4,952   | 96.36%      |
| Pseudo genes                        | 75      | 1.46%       |
| Genes in paralog clusters           | 124     | 2.4%        |
| Genes assigned to COGs              | 3,807   | 74.08%      |
| Genes assigned in Pfam domains      | 4,185   | 81.43%      |
| Genes with signal peptides          | 677     | 13.17%      |
| Genes with transmembrane helices    | 1,227   | 23.87%      |
| CRISPR repeats                      | 1       |             |

a) The total is based on either the size of the genome in base pairs or the total number of protein coding genes in the annotated genome.

Table 4. Number of genes associated with the 25 general COG functional categories

| Code | Value | %agea | Description                                             |
|------|-------|-------|---------------------------------------------------------|
| J    | 201   | 4.27  | Translation, ribosomal structure and biogenesis         |
| A    | 1     | 0.02  | RNA processing and modification                         |
| K    | 481   | 10.22 | Transcription                                           |
| L    | 160   | 3.40  | DNA replication, recombination and repair               |
| B    | 1     | 0.02  | Chromatin structure and dynamics                        |
| D    | 37    | 0.79  | Cell division and chromosome partitioning              |
| Y    | 0     | 0.00  | Nuclear structure                                       |
| V    | 64    | 1.36  | Defense mechanisms                                      |
| T    | 187   | 3.97  | Signal transduction mechanisms                          |
| M    | 265   | 5.63  | Cell envelope biogenesis, Outer membrane                |
| N    | 94    | 2.00  | Cell motility and secretion                             |
| Z    | 0     | 0.00  | Cytoskeleton                                            |
| W    | 0     | 0.00  | Extracellular structure                                 |
| U    | 116   | 2.47  | Intracellular trafficking and secretion                 |
| O    | 153   | 3.25  | Posttranslational modification, protein turnover, chaperones |
| C    | 272   | 5.78  | Energy production and conversion                        |
| G    | 424   | 9.01  | Carbohydrate transport and metabolism                   |
| E    | 470   | 9.99  | Amino acid transport and metabolism                     |
| F    | 106   | 2.25  | Nucleotide transport and metabolism                     |
| H    | 185   | 3.93  | Coenzyme metabolism                                    |
| I    | 135   | 2.87  | Lipid metabolism                                       |
| P    | 285   | 6.06  | Inorganic ion transport and metabolism                  |
| Q    | 133   | 2.83  | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 537   | 11.41 | General function prediction only                        |
| S    | 398   | 8.46  | Function unknown                                        |
| -    | 917   | 17.85 | Not in COG                                             |

a) The total is based on the total number of protein coding genes in the annotated genome.
Serratia plymuthica strain AS9

Acknowledgements
We would like to gratefully acknowledge the help of Elke Lang for providing cell pastes of reference material and Evelyne-Marie Brambilla for extraction of DNA for digital DNA-DNA hybridizations with the reference strains (both at DSMZ). The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

References
1. Grimont F, Grimont PAD. The genus *Serratia*. In: *The Prokaryotes*, Balows A, Trüper HG, Dworkin M, Harder W, Schleifer K-H (eds), New York: Springer. 1992;2822-2848.
2. Müller H, Berg G. Impact of formulation procedures on the effect of the biocontrol agent *Serratia plymuthica* HRO-C48 on *Verticillium* wilt in oilseed rape. *BioControl* 2008; 53:905-916. [PubMed](http://dx.doi.org/10.1007/s10526-007-9111-3)
3. Kalbe C, Marten P, Berg G. Strains of genus *Serratia* as beneficial rhizobacteria of oilseed rape with antifungal properties. *Microbiol Res* 1996; 151:433-439. [PubMed](http://dx.doi.org/10.1016/S0944-5013(96)80014-0)
4. Taghavi S, Garafola C, Monchy S, Newman L, Hoffmann A, Weyens N, Barac T, Vangronsveld J, van der Lelie D. Genome survey and characterization of endophytic bacteria exhibiting a beneficial effect on growth and development of poplar trees. *Appl Environ Microbiol* 2009; 75:748-757. [PubMed](http://dx.doi.org/10.1128/AEM.02239-08)
5. Alström S. Characteristics of bacteria from oilseed rape in relation to their biocontrol activity against *Verticillium dahliae*. *J Phytopathol* 2001; 149:57-64. [PubMed](http://dx.doi.org/10.1046/j.1439-0434.2001.00585.x)
6. Alschsul SF, Thomas LS, Alejandro AS, Jingui Z, Webb M, David JL. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 1997; 25:3389-3402. [PubMed](http://dx.doi.org/10.1093/nar/25.17.3389)
7. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valent F, Wallace IM, Wilm A, Lopez R, et al. Clustal W and Clustal X version 2.0. *Bioinformatics* 2007; 23:2947-2948. [PubMed](http://dx.doi.org/10.1093/bioinformatics/btm404)
8. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* 2011; 28:2731-2739.
17. Garrity GM, Holt JG. Taxonomic Outline of the Archaea and Bacteria. In: Garrity GM, Boone DR, Castenholz RW (eds), Bergey’s Manual of Systematic Bacteriology, Second Edition, Volume 1, Springer, New York, 2001, p. 155-166.

18. Skerman VBD, McGowan V, Sneath PHA. Approved Lists of Bacterial Names. Int J Syst Bacteriol 1980; 30:225-420. http://dx.doi.org/10.1099/00207713-30-1-225

19. Rahn O. New principles for the classification of bacteria. Zentralbl Bakteriol Parasitenkd Infektionskr Hyg 1937; 96:273-286.

20. Judicial Commission. Conservation of the family name Enterobacteriaceae, of the name of the type genus, and designation of the type species OPUS-ION NO. 15. Int Bull Bacteriol Nomencl Taxon 1958; 8:73-74. http://dx.doi.org/10.1099/0096266X-8-1-73

21. Bizio B. Lettera di Bartolomeo Bizio al chiarissimo canonico Angelo Bellani sopra il fenomeno della polenta porporina. Biblioteca Italiana o sia Giornale di Letteratura. [Anno VIII]. Scienze e Arti 1823; 30:275-295.

22. Sakazaki R. Genus IX. Serratia Bizio 1823, 288. In: Buchanan RE, Gibbons NE (eds), Bergey’s Manual of Determinative Bacteriology, Eighth Edition, The Williams and Wilkins Co., Baltimore, 1974, p. 326-326.

23. Breed RS, Murray EGD, Hitchens AP. In: Breed RS, Murray EGD, Hitchens AP (eds), Bergey’s Manual of Determinative Bacteriology, Sixth Edition, The Williams and Wilkins Co., Baltimore, 1948, p. 481-482.

24. BAuA. 2010, Classification of bacteria and archaea in risk groups. http://www.baua.de TRBA 466, p. 200.

25. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene Ontology: tool for the unification of biology. Nat Genet 2000; 25:25-29. PubMed http://dx.doi.org/10.1038/75556

26. Bennett JW, Bentley R. Seeing red: the story of prodigiosin. Adv Appl Microbiol 2000; 47:1-32. PubMed http://dx.doi.org/10.1016/S0065-2164(00)47000-0

27. Alström S, Gerhardson B. Characteristics of a Serratia plymuthica isolate from plant rhizospheres. Plant Soil 1987; 103:185-189. http://dx.doi.org/10.1007/BF02370387

28. Khanañarì A, Assadi MM, Fakhri FA. Review of prodigiosin pigmentation in Serratia marcescens. J Biol Sci 2006; 6:1-13.

29. Bergan T, Grimont AD, Grimont F. Fatty acids of Serratia determined by gas chromatography. Curr Microbiol 1983; 8:7-11. http://dx.doi.org/10.1007/BF01567306

30. DOE Joint Genome Institute. http://www.jgi.doe.gov

31. Bennett S. Solexa Ltd. Pharmacogenomics 2004; 5:433-438. PubMed http://dx.doi.org/10.1517/14622416.5.4.433

32. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 2005; 437:326-327. PubMed

33. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 2008; 18:821-829. PubMed http://dx.doi.org/10.1101/gr.074492.107

34. Phrap and Phred for Windows. MacOS, Linux, and Unix. http://www.phrap.com

35. Han C, Chain P. 2006. Finishing repeat regions automatically with Dupfinisher. In: Proceeding of the 2006 international conference on bioinformatics & computational biology. Arabina HR, Valafar H (eds), CSREA Press. June 26-29, 2006: 141-146.

36. Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res 1998; 8:195-202. PubMed

37. Ewing B, Green P. Base-calling of automated sequencer traces using Phred. II. error probabilities. Genome Res 1998; 8:186-194. PubMed

38. Ewing B, Hillier L, Wendl MC, Green P. Base-Calling of automated sequencer traces using Phred. I. accuracy assessment. Genome Res 1998; 8:175-185. PubMed

39. Lapidus A, LaButti K, Foster B, Lowry S, Trong S, Goltsman E. POLISHER: An effective tool for using ultra short reads in microbial genome assembly and finishing. AGBT, Marco Island, FL, 2008

40. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 2010; 11:119. PubMed http://dx.doi.org/10.1186/1471-2105-11-119

41. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kypirides NC.
GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods* 2010; 7:455-457. PubMed http://dx.doi.org/10.1038/nmeth.1457

42. Schattner P, Brooks AN, Lowe TM. The tRNAscanSE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* 2005; 33:W686-W689. PubMed http://dx.doi.org/10.1093/nar/gki366

43. Markowitz VM, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 2009; 25:2271-2278. PubMed http://dx.doi.org/10.1093/bioinformatics/btp393