

Complete genome sequence of the plant-associated *Serratia plymuthica* strain AS13

Saraswoti Neupane¹, Roger D. Finlay¹, Nikos C. Kyrpides², Lynne Goodwin^{2,3}, Sadhna Alström¹, Susan Lucas², Miriam Land^{2,4}, James Han², Alla Lapidus², Jan-Fang Cheng², David Bruce^{2,3}, Sam Pitluck², Lin Peters², Galina Ovchinnikova², Brittany Held^{2,3}, Cliff Han^{2,3}, John C. Detter^{2,3}, Roxanne Tapia^{2,3}, Loren Hauser^{2,4}, Natalia Ivanova², Ioanna Pagani², Tanja Woyke², Hans-Peter Klenk⁵ and Nils Högberg¹

¹ 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: Gram-negative, non-sporulating, motile, plant-associated, chemoorganotrophic, *Enterobacteriaceae*

Serratia plymuthica AS13 is a plant-associated Gammaproteobacteria, isolated from rapeseed roots. It is of special interest because of its ability to inhibit fungal pathogens of rapeseed and to promote plant growth. The complete genome of *S. plymuthica* AS13 consists of a 5,442,549 bp circular chromosome. The chromosome contains 4,951 protein-coding genes, 87 tRNA genes and 7 rRNA operons. This genome was sequenced as part of the project entitled “Genomics of four rapeseed plant growth promoting bacteria with antagonistic effect on plant pathogens” within the 2010 DOE-JGI Community Sequencing Program (CSP2010).

Introduction

The members of the genus *Serratia* are widely distributed in nature. They are commonly found in soil, water, plants, insects, and other animals including humans [1]. The genus includes biologically and ecologically diverse species – from those beneficial to economically important plants, to pathogenic species that are harmful to humans. The plant-associated species comprise both endophytes and free living taxa, such as *S. proteamaculans*, *S. plymuthica*, *S. liquefaciens* and *S. grimesii*. Most of them are of interest because of their ability to promote plant growth and inhibit plant pathogenic fungi [2-6].

There are currently 16 validly named *Serratia* species. However, there are several unidentified plant-associated *Serratia* strains that have an impact on agriculture by stimulating plant growth and/or inhibiting soil borne plant pathogens [3]. *S. plymuthica* AS13 was isolated from rapeseed roots from Uppsala, Sweden. Our interest in *S. plymuthica* AS13 is due

to its ability to stimulate rapeseed plant growth and to inhibit soil borne fungal pathogens such as *Verticillium dahlia* and *Rhizoctonia solani* [6]. Here we present a description of the complete genome of *S. plymuthica* AS13 and its annotation.

Classification and features

A representative sequence of the 16S rRNA gene of *S. plymuthica* AS13 was compared with the most recently released GenBank databases using NCBI BLAST [7] under default settings. It showed that the strain AS13 shares 99-100% similarity with the genus *Serratia*. When considering high-scoring segment pairs (HSPs) from the best 250 hits, the most frequent matches were several unspecified *Serratia* strains (17.2%) with maximum identity of 97-100%, while *S. plymuthica* (5.2%) had maximum identity of 97-100%, *S. proteamaculans* (4.8%) maximum identity of 97-99%, *S. marcescens* (4.8%) maximum identity of 96-97% and also different *Rahnella* strains (7%) maximum identity of 97-98%.

The phylogenetic relationship of *S. plymuthica* AS13 is shown in Figure 1 in a 16S rRNA based tree. All *Serratia* lineages clustered together and were distinct from other enterobacteria (except *Obesumbacterium proteus*). The tree also shows its very close relation with *S. plymuthica* strains AS9 and AS12, which was confirmed by digital DNA-DNA hybridization values [12] above 70% when compared with the (unpublished) draft genome sequence of the *S. plymuthica* type strain Breed K-7^T from a culture of DSM 4540, and when compared with the complete genome sequences of *S. plymuthica* AS9 [13] and *S. plymuthica* AS12 [14] using the GGDC web server [15].

Strain AS13 is a rod shaped bacterium, 1-2 µm long, 0.5-0.7 µm wide (Figure 2 and Table 1), is Gram-

negative, motile, and a member of the family *Enterobacteriaceae*. The bacterium is a facultative anaerobe and grows within the temperature range 4 °C - 40 °C and within a pH range of 4 - 10. It has chitinolytic, cellulolytic, proteolytic, and phospholytic activity [6] and can easily grow on different carbon sources such as glucose, cellobiose, succinate, mannitol, arabinose and inositol. It forms red to pink colored colonies that are 1-2 mm in diameter on potato dextrose agar at low temperature. The color of the bacterium depends on the growth substrate, temperature and pH of the culture medium [30]. The bacterium is deposited in the Culture Collection, University of Göteborg, Sweden (CCUG) as *S. plymuthica* AS13 (= CCUG 61398).

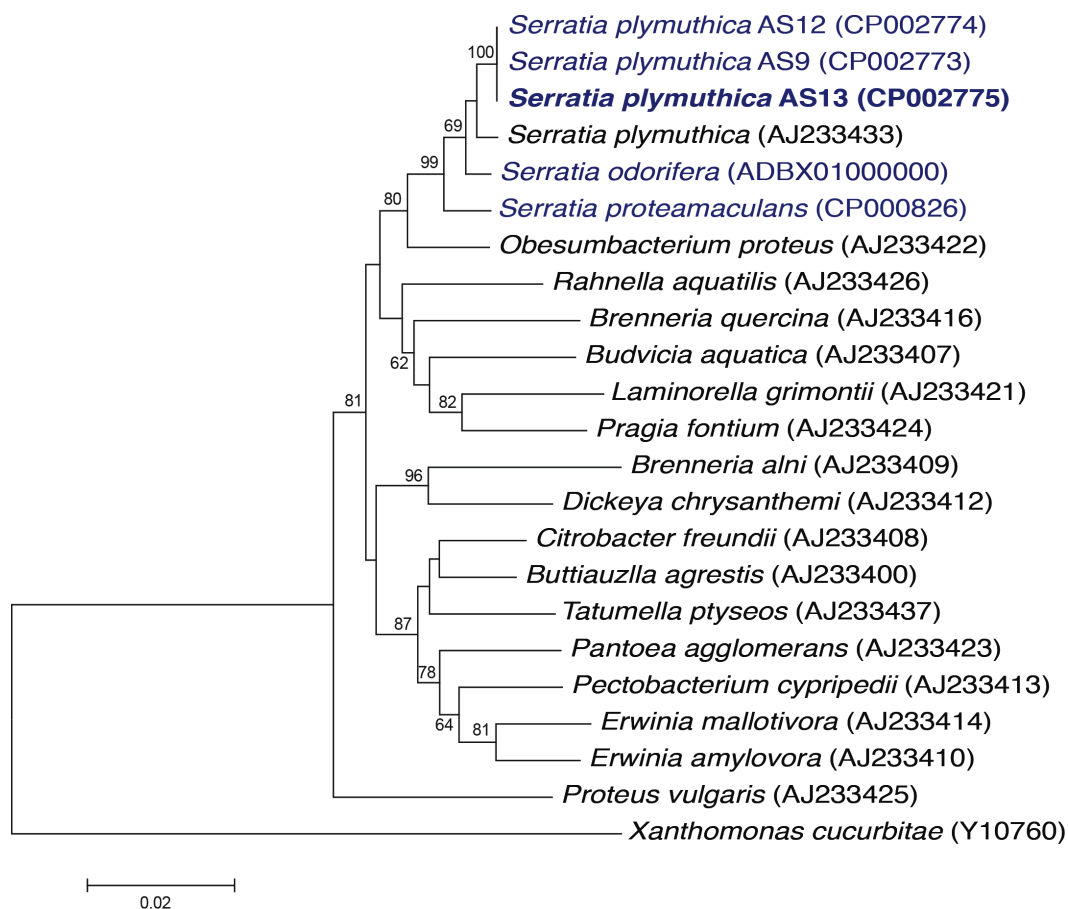


Figure 1. Phylogenetic tree highlighting the position of *S. plymuthica* AS13 in relation to other genera within the family *Enterobacteriaceae*, based on 1,472 characters of the 16S rRNA gene sequence aligned in ClustalW2 [8]. The tree was constructed under the maximum likelihood criterion using MEGA5 software [9] and rooted with *Xanthomonas cucurbitae* (a member of the *Xanthomonadaceae* family). The branches are scaled based on the expected number of substitutions per site. The numbers above branches are support values from 1,000 bootstrap replicates if larger than 60% [10]. The lineages shown in blue color are the genome sequences of bacterial strains that are registered in GOLD [11].

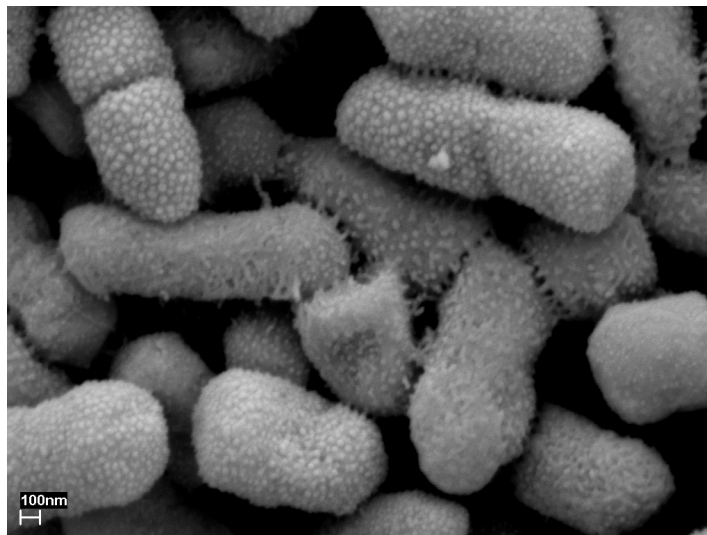


Figure 2. Scanning electron micrograph of *S. plymuthica* AS13

Chemotaxonomy

Little is known about the chemotaxonomy of *S. plymuthica* AS13. Fatty acid methyl ester (FAME) analysis showed the main fatty acid in strain AS13 comprises C_{16:0} (25.27%), C_{16:1 ω 7c} (15.41%), C_{18:1 ω 7c} (18.17%), C_{14:0} (5.21%), C_{17:0} cyclo (18.53%), along with other minor fatty acid components. Previously it has been shown that *Serratia* spp. contain a mixture of C_{14:0}, C_{16:0}, C_{16:1} and C_{18:1+2} fatty acids in which 50-80% of the total fatty acid in the cell is C_{14:0} and other fatty acids are less than 3% each [31]. This is consistent with the fact that C_{14:0} fatty acid is characteristic of the family *Enterobacteriaceae*.

Genome sequencing information

S. plymuthica AS13, a bacterial strain isolated from rapeseed roots was selected for sequencing on the basis of its biocontrol activity against fungal pathogens of rapeseed and its plant growth promoting ability. The genome project is deposited in the Genomes On Line Database [11] (GOLD ID = Gc01776) and the complete genome sequence is deposited in GenBank (INSDC ID = CP002775). 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 MGS identifiers.

Growth conditions and DNA isolation

S. plymuthica AS13 was grown in Luria Broth (LB) medium at 28 °C until early stationary phase. The DNA was extracted from the cells by using a standard CTAB protocol for bacterial genomic DNA isolation that is available at JGI [32].

Genome sequencing and assembly

The genome of *S. plymuthica* AS13 was sequenced using a combination of Illumina and 454 sequencing platforms. The details of library construction and sequencing can be found at the JGI [32]. The sequence data from Illumina GAii (1,457.3 Mb) were assembled with Velvet [33] and the consensus sequence was computationally shredded into 1.5 kb overlapping fake reads. The sequencing data from 454 pyrosequencing (79.5 Mb) were assembled with Newbler and consensus sequences were computationally shredded into 2 kb overlapping fake reads. The initial draft assembly contained 86 contigs in 1 scaffold. The 454 Newbler consensus reads, the Illumina Velvet consensus reads and the read pairs in the 454 paired end library were assembled and quality assessment performed in the subsequent finishing process by using software phrap package [34-37]. Possible mis-assemblies were corrected with gapResolution [32], Dupfinisher [38], or by sequencing cloned bridging PCR fragments with subcloning. The gaps between contigs were closed by editing in the software Consed [37], by PCR and by Bubble PCR primer walks (J.-F. Chang, unpublished). Fifty one 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 46.8 Mb of 454 draft data which provides an average 8.7 × coverage of the genome and 1,415.6 Mb of Illumina draft data which provides an average 262.2 × coverage of the genome.

Genome annotation

The *S. plymuthica* AS13 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 GenePRIMP pipeline [41]. The predicted CDS were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant database, Uniport, TIGR-Fam, Pfam, PRIAM, KEGG, COG and InterPro databases. Non-coding genes and miscellaneous features were predicted using tRNAscan-SE [42], RNAmmer [43], Rfam [44], TMHMM [45], and signalP [46]. Additional gene prediction analysis and functional annotation was performed within the Integrated Microbial Ge-

nomes – Expert Review (IMG-ER) platform developed by the Joint Genome Institute, Walnut Creek, CA, USA [47].

Genome properties

The genome of *S. plymuthica* AS13 has a single circular chromosome of 5,442,549 bp with 55.96% GC content (Table 3 and Figure 3). It has 5,139 predicted genes, of which 4,951 were assigned as protein-coding genes. Among them, most of the protein coding genes (84.41%) were functionally assigned while the remaining ones were annotated as hypothetical proteins. 112 genes were assigned as RNA genes and 76 as pseudogenes. The distribution of genes into COG functional categories is presented in Table 4.

Table 1. Classification and general features of *S. plymuthica* AS13 according to the MIGS recommendations [16]

| MIGS ID | Property | Term | Evidence code |
|----------|------------------------|--|----------------|
| | | Domain <i>Bacteria</i> | TAS [17] |
| | | Phylum <i>Proteobacteria</i> | TAS [18] |
| | | Class Gammaproteobacteria | TAS [19,20] |
| | Current classification | Order “ <i>Enterobacteriales</i> ” | TAS [21] |
| | | Family <i>Enterobacteriaceae</i> | TAS [22-24] |
| | | Genus <i>Serratia</i> | TAS [22,25,26] |
| | | Species <i>Serratia plymuthica</i> | TAS [22,27] |
| | | Strain AS13 | 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, inositol, arabinose, succinate, sucrose, fructose | IDA |
| | Energy metabolism | Chemoorganotrophic | IDA |
| MIGS-6 | Habitat | Rapeseed roots | IDA |
| MIGS-6.3 | Salinity | Medium | IDA |
| MIGS-22 | Oxygen | Facultative | IDA |
| MIGS-15 | Biotic relationship | Plant associated | TAS [6] |
| MIGS-14 | Pathogenicity | None | IDA |
| | Biosafety level | 1 | TAS [28] |
| MIGS-4 | Geographic location | Uppsala, Sweden | NAS |
| MIGS-5 | Sample collection time | Summer 1998 | IDA |
| 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 [29]. If the evidence code is IDA, then the property should have been directly observed, for the purpose of this specific publication, for a live isolate by one of the authors, or an expert or reputable institution mentioned in the acknowledgements.

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 paired end 454 library (9.0 kb insert size) and one Illumina library) |
| MIGS-29 | Sequencing platforms | Illumina GAii, 454 GS FLX Titanium |
| MIGS-31.2 | Fold coverage | 262.2 × Illumina, 8.7 × pyrosequencing |
| MIGS-30 | Assemblers | Newbler version 2.3, Velvet 1.0.13, phrap version SPS - 4.24 |
| MIGS-32 | Gene calling method | Prodigal 1.4, GenePRIMP |
| | NCBI project ID | 60455 |
| | INSDC ID | CP002775 |
| | Genbank Date of Release | October 12, 2011 |
| | GOLD ID | Gc01776 |
| | Project relevance | Biocontrol, Agriculture |

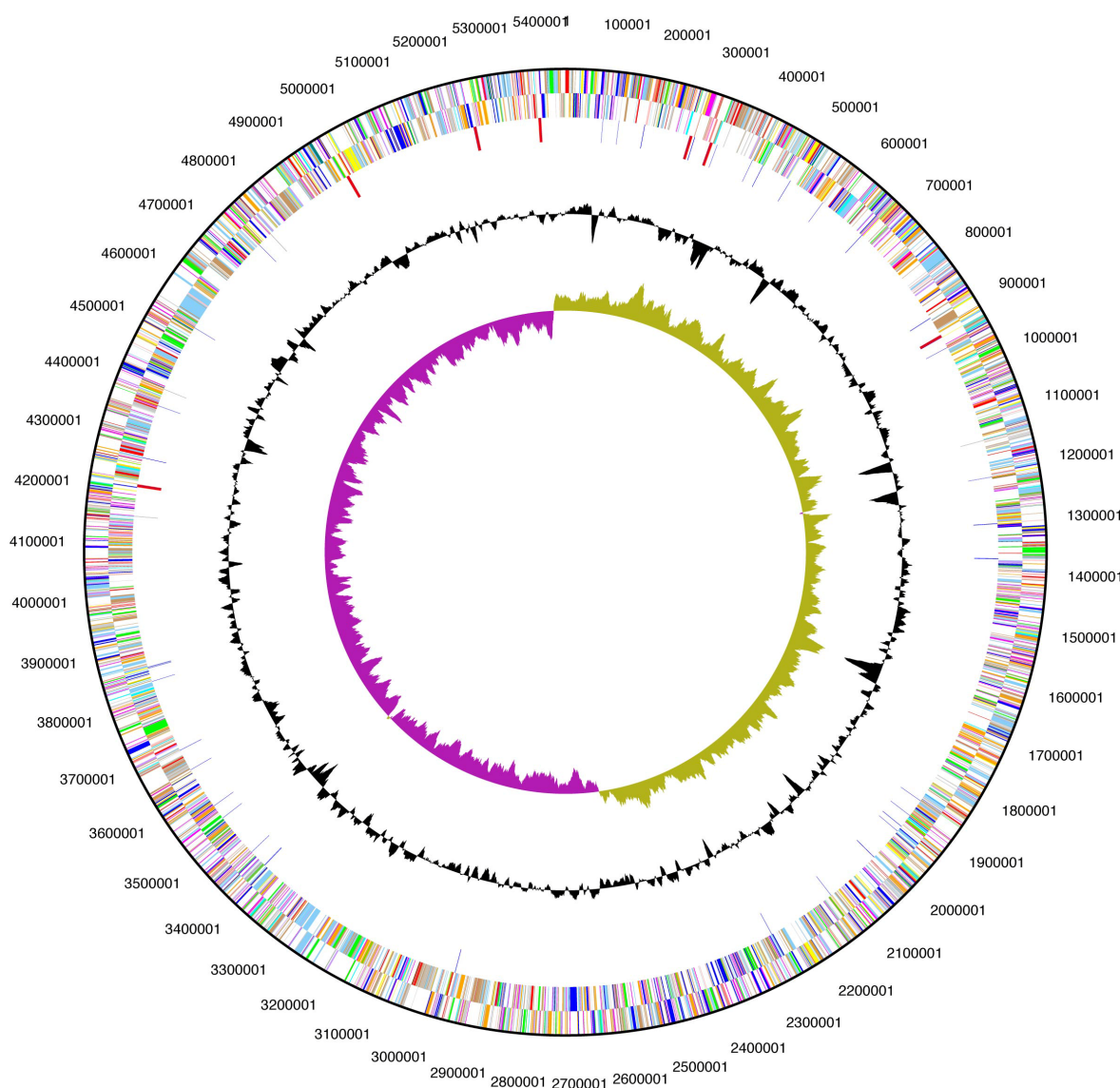


Figure 3. Graphical circular map of the chromosome. From outside to the center: Genes on forward strand (color by COG categories), Genes on reverse strand (color by COG categories), RNA genes (tRNAs blue, rRNAs red, other RNAs black), GC content, GC skew.

Table 3. Genome statistics

| Attribute | Value | % of total^a |
|----------------------------------|--------------|-------------------------------|
| Genome size (bp) | 5,442,549 | 100.00% |
| DNA Coding region (bp) | 4,770,475 | 87.65% |
| DNA G+C content (bp) | 3,045,680 | 55.96% |
| Total genes | 5,139 | 100.00% |
| RNA genes | 112 | 2.18% |
| rRNA operons | 7 | 0.14% |
| Protein-coding genes | 4,951 | 96.34% |
| Pseudogenes | 76 | 1.48% |
| Genes in paralog clusters | 112 | 2.18% |
| Genes assigned to COGs | 3,805 | 74.04% |
| Genes assigned in Pfam domains | 4,183 | 81.39% |
| Genes with signal peptides | 676 | 13.15% |
| Genes with transmembrane helices | 1,228 | 23.89% |
| CRISPR repeats | 1 | % of totala |

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 | % age | Description |
|-------------|--------------|--------------|--|
| J | 201 | 4.27 | Translation, ribosomal structure and biogenesis |
| A | 1 | 0.02 | RNA processing and modification |
| K | 480 | 10.20 | Transcription |
| L | 161 | 3.42 | 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 metabolite biosynthesis, transport and catabolism |
| R | 537 | 11.41 | General function prediction only |
| S | 398 | 8.46 | Function unknown |
| - | 918 | 17.86 | Not in COG |

Acknowledgements

We gratefully acknowledge the help of Elke Lang for providing cell cultures of the reference bacterial strain, Evelyne-Marie Brambilla for extraction of DNA and Anne Fiebig for assembly of the reference genome required for digital DNA-DNA hybridizations (all at

DSMZ). The work conducted by the US Department of Energy Joint Genome Institute is supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231.

References

1. Grimont F, Grimont PAD. (1992). The genus *Serratia*. In: Balows A, Trüper HG, Dworkin M, Harder W, Schleifer KH (eds) *The Prokaryotes*, pp 2822-2848. Springer, New York.
2. 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](http://dx.doi.org/10.1016/S0944-5013(96)80014-0)
3. 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. <http://dx.doi.org/10.1007/s10526-007-9111-3>
4. Kurze S, Bahl H, Dahl R, Berg G. Biological control of fungal strawberry diseases by *Serratia plymuthica* HRO-C48. *Plant Dis* 2001; **85**:529-534. <http://dx.doi.org/10.1094/PDIS.2001.85.5.529>
5. Taghavi S, Garafola C, Monchy S, Newman L, Hoffman 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>
6. Alström S. Characteristics of bacteria from oilseed rape in relation to their biocontrol activity against *Verticillium dahliae*. *J Phytopathol* 2001; **149**:57-64. <http://dx.doi.org/10.1046/j.1439-0434.2001.00585.x>
7. Altschul 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>
8. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin 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>
9. 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. [PubMed](#) <http://dx.doi.org/10.1093/molbev/msr121>
10. Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A. How many bootstrap replicates are necessary? *Lect Notes Comput Sci* 2009; **5541**:184-200. http://dx.doi.org/10.1007/978-3-642-02008-7_13
11. Liolios K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, Kyrpides NC. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res* 2009; **38**:D346-D354. [PubMed](#) <http://dx.doi.org/10.1093/nar/gkp848>
12. Auch AF, von Jan M, Klenk HP, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2010; **2**:117-134. [PubMed](#) <http://dx.doi.org/10.4056/sigs.531120>
13. Neupane S, Högberg N, Alström S, Lucas S, Han J, Lapidus A, Cheng JF, Bruce D, Goodwin L, Pitluck S, et al. Complete genome sequence of the rapeseed plant-growth promoting *Serratia plymuthica* strain AS9. *Stand Genomic Sci* 2012; **6**:54-62. [PubMed](#) <http://dx.doi.org/10.4056/sigs.2595762>
14. Neupane S, Finlay RD, Alström S, Goodwin L, Kyrpides NC, Lucas S, Lapidus A, Bruce D, Pitluck S, Peters L, et al. Complete genome sequence of *Serratia plymuthica* strain AS12. *Stand Genomic Sci* 2012; **6**:165-173. [PubMed](#) <http://dx.doi.org/10.4056/sigs.2705996>
15. Auch AF, Klenk HP, Göker M. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. *Stand Genomic Sci* 2010; **2**:142-148. [PubMed](#) <http://dx.doi.org/10.4056/sigs.541628>

16. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008; **26**:541-547. [PubMed](#) <http://dx.doi.org/10.1038/nbt1360>
17. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria* and *Eucarya*. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579. [PubMed](#) <http://dx.doi.org/10.1073/pnas.87.12.4576>
18. Garrity GM, Bell JA, Lilburn T. Phylum XIV. *Proteobacteria* phyl. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 2, Part B, Springer, New York, 2005, p. 1.
19. List Editor. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. List no. 106. *Int J Syst Evol Microbiol* 2005; **55**:2235-2238. <http://dx.doi.org/10.1099/ijs.0.64108-0>
20. Garrity GM, Bell JA, Lilburn T. Class III. Gammaproteobacteria class. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 2, Part B, Springer, New York, 2005, p. 1.
21. 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.
22. 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>
23. Rahn O. New principles for the classification of bacteria. *Zentralbl Bakteriol Parasitenkd Infektionskr Hyg* 1937; **96**:273-286.
24. Judicial Commission. Conservation of the family name *Enterobacteriaceae*, of the name of the type genus, and designation of the type species OPINION NO. 15. *Int Bull Bacteriol Nomencl Taxon* 1958; **8**:73-74.
25. 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.
26. 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.
27. 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.
28. BAuA. 2010, Classification of bacteria and archaea in risk groups. <http://www.baua.de> TRBA 466, p. 200.
29. 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>
30. 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>
31. 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>
32. DOE Joint Genome Institute. <http://www.jgi.doe.gov>
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. Ewing B, Green P. Base-Calling of automated sequencer traces using Phred. II. error probabilities. *Genome Res* 1998; **8**:186-194. [PubMed](#)
36. 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](#)
37. Gordon D, Abajian C, Green P. *Consed*: a graphical tool for sequence finishing. *Genome Res* 1998; **8**:195-202. [PubMed](#)
38. Han C, Chain P. 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.
39. Lapidus A, LaButti K, Foster B, Lowry S, Trong S, Goltsman E. POLISHER: An effective tool for us-

- ing 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) <http://dx.doi.org/10.1186/1471-2105-11-119>
 41. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides 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) <http://dx.doi.org/10.1038/nmeth.1457>
 42. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997; **25**:955-964. [PubMed](http://dx.doi.org/10.1093/nar/gkm160)
 43. Lagesen K, Hallin P, Rødland EA, Stærfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent annotation of rRNA genes in genomic sequences. *Nucleic Acids Res* 2007; **35**:3100-3108. [PubMed](http://dx.doi.org/10.1093/nar/gkm160) <http://dx.doi.org/10.1093/nar/gkm160>
 44. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. Rfam: an RNA family database. *Nucleic Acids Res* 2003; **31**:439-441. [PubMed](http://dx.doi.org/10.1093/nar/gkg006) <http://dx.doi.org/10.1093/nar/gkg006>
 45. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J Mol Biol* 2001; **305**:567-580. [PubMed](http://dx.doi.org/10.1006/jmbi.2000.4315) <http://dx.doi.org/10.1006/jmbi.2000.4315>
 46. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004; **340**:783-795. [PubMed](http://dx.doi.org/10.1016/j.jmb.2004.05.028) <http://dx.doi.org/10.1016/j.jmb.2004.05.028>
 47. Markowitz VM, Mavromatis K, 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) <http://dx.doi.org/10.1093/bioinformatics/btp393>