Complete genome sequence of *Serratia plymuthica* strain AS12

Saraswoti Neupane¹, Roger D. Finlay¹, Sadhna Alström¹, Lynne Goodwin²,³, Nikos C. Kyrpides², Susan Lucas², Alla Lapidus², David Bruce²,³, Sam Pitluck², Lin Peters², Galina Ovcinnikova², Olga Chertkov²,³, James Han², Cliff Han²,³, Roxanne Tapia²,³, John C. Detter²,³, Miriam Land², Loren Hauser²,³, Jan-Fang Cheng², Natalia Ivanova², Ioanna Pagani², Hans-Peter Klenk², Tanja Woyke² and Nils Höögberg¹

¹ Uppsala BioCenter, Department of Forest Mycology and Plant 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

Keywords: Facultative anaerobe, gram-negative, motile, non-sporulating, mesophilic, chemoorganotrophic, agriculture, *Enterobacteriaceae*, CSP 2010

A plant-associated member of the family *Enterobacteriaceae*, *Serratia plymuthica* strain AS12 was isolated from rapeseed roots. It is of scientific interest because it promotes plant growth and inhibits plant pathogens. The genome of *S. plymuthica* AS12 comprises a 5,443,009 bp long circular chromosome, which consists of 4,952 protein-coding genes, 87 rRNA genes and 7 rRNA operons. This genome was sequenced within the 2010 DOE-JGI Community Sequencing Program (CSP2010) as part of the project entitled “Genomics of four rapeseed plant growth promoting bacteria with antagonistic effect on plant pathogens”.

Introduction

Plant associated *Serratia* species are commonly found as free-living bacteria in rhizosphere soil and as endophytes within plant roots. They include strains with the ability to stimulate plant growth and to inhibit the growth of soil borne pathogens of economically important agricultural plants [1-3]. One *Serratia* strain, *S. plymuthica* HRO-C48, is successfully used as an alternative to chemical agents for control of soil-borne fungal diseases in different crops such as strawberry and rapeseed [3,4]. Its ability to degrade chitin, a fungal cell wall component, may be responsible for antifungal activity, whereas the production of the plant hormone indole-3-acetic acid (IAA) could be involved in plant growth promotion [3]. *S. plymuthica* AS12 has chitinolytic activity and was isolated from rapeseed roots from Uppsala, Sweden in 1998 [5]. The reason for our interest in *S. plymuthica* AS12 is its ability to inhibit *Verticillium longisporum* (earlier *V. dahliae*), a soil borne fungal pathogen of rapeseed, thus promoting the rapeseed growth both directly and indirectly [5]. Here we present a description of the complete genome of *S. plymuthica* AS12 and its annotation.

Classification and features

A representative 16S rRNA gene sequence of the strain AS12 genome was used for comparison using NCBI BLAST [6] under default settings with the most recent databases. The relative frequencies of taxa and BLAST scores were determined. The most frequently occurring genus is *Serratia* where some of the ‘hits’ share a 100% identity. When considering high-scoring segment pairs (HSPs) from the best 250 hits, the most frequent matches were *Serratia* sp. (17.2%) with a maximum identity of 97-100%, while *S. plymuthica* (5.2%) had a maximum identity of 97-100%, *S. proteomaculans* (4.8%) with a maximum identity of 97-99%, *S. marcescens* (4.8%) with a maximum identity of 96-97% and different strains of *Rahnella* (7%) with a maximum identity of 97-98%.
A phylogenetic tree (Figure 1) was constructed using 16S rRNA sequences of *S. plymuthica* AS12 with other genera within the family *Enterobacteriaceae* including two species within the genus *Serratia*. The tree shows the position of *S. plymuthica* AS12 within the genus *Serratia* and its distinct clustering with *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 DSM 4540 culture as well as with the complete genome sequence of *S. plymuthica* AS9 [12] using the GGDC web server [13].

The cells of strain AS12 stain Gram-negative and are rod shaped, 1-2 μm long, 0.5-0.7 μm wide (Figure 2 and Table 1) and motile. The culture forms red to pink colored colonies of 1-2 mm diameter on tryptic soy agar and potato dextrose agar, but the colony color depends on different factors such as the growth substrate, pH of the medium and growth temperature. The organism is a facultative anaerobe and grows at temperatures ranging from 4 °C - 40 °C and within a pH range of 4 - 10. It has the ability to utilize a wide range of carbon sources such as glucose, sucrose, succinate, mannitol and arabinose. It also has cellulolytic, phospholysic, chitinolytic and proteolytic activity [5]. The strain is deposited in the Culture Collection, University of Göteborg, Sweden (CCUG) as *Serratia sp.* AS12 (= CCUG 61397).

**Chemotaxonomy**

The cells of *S. plymuthica* AS12 contain a mixture of saturated and unsaturated fatty acids. The dominant fatty acids in strain AS12 are C₁₆:0 (22.94%), C₁₆:1ω₇c (17.08%), C₁₈:1ω₇c (19.65%), C₁₄:0 (5.11%), along with other minor fatty acid components. No information is available for other compounds. Previously it has been shown that *Serratia* spp. contain a mixture of C₁₄:0, C₁₆:0, C₁₆:1 and C₁₈:1ω₂ fatty acids in which 50-80% of the total fatty acid in the cell is C₁₄:0 and others each less than 3% [28]. This is consistent with the observation that C₁₄:0 is a characteristic fatty acid of the family *Enterobacteriaceae*.

---

**Figure 1.** Phylogenetic tree highlighting the position of *S. plymuthica* AS12 in relation to selected *Serratia* strains and other genera within the family *Enterobacteriaceae*. The tree was based on 1,535 characters of the 16S rRNA gene sequence aligned in ClustalW2 [7]. The tree was inferred under the maximum likelihood criterion using MEGA5 software [8] and rooted with *Pseudomonas trivialis* (a member of the *Pseudomonadaceae* family). The branches are mapped by the expected number of substitutions per site. The numbers above the branches are support values from 1,000 bootstrap replicates if larger than 60% [9]. Lineages with genome sequences registered in GOLD [10] are shown in blue.
Figure 2. Scanning electron micrograph of *S. plymuthica* AS12

Table 1. Classification and general features of *S. plymuthica* AS12 according to MIGS recommendations [14]

| MIGS ID | Property                  | Term                                                                 | Evidence code |
|---------|---------------------------|----------------------------------------------------------------------|---------------|
|         | Current classification    | Domain *Bacteria*                                                      | TAS [15]      |
|         |                           | Phylum *Proteobacteria*                                               | TAS [16]      |
|         |                           | Class *Gammaproteobacteria*                                           | TAS [17,18]   |
|         |                           | Order “Enterobacteriales”                                             | TAS [19]      |
|         |                           | Family *Enterobacteriaceae*                                           | TAS [20-22]   |
|         |                           | Genus *Serratia*                                                      | TAS [20,23,24]|
|         |                           | Species *Serratia plymuthica*                                         | TAS [20,25]   |
|         | Strain AS12               |                                                                       | IDA           |
|         | Gram stain                | Negative                                                             | IDA           |
|         | Cell shape                | Rod-shaped                                                           | IDA           |
|         | Motility                  | Motile                                                               | IDA           |
|         | Sporulation               | Non-sporulating                                                      | IDA           |
|         | Optimum temperature       | Mesophilic, 4 – 40°C                                                 | IDA           |
|         | Carbon source             | Glucose, sucrose, fructose, succinate, trehalose, mannitol, inositol, arabinose | IDA           |
|         | Energy metabolism         | Chemoorganotrophic                                                  | IDA           |
|         | MIGS-6                    | Habitat                                                              | IDA           |
|         |                             | Rapeseed roots                                                       | IDA           |
|         | MIGS-6.3                  | Salinity                                                             | IDA           |
|         |                             | Medium                                                               | IDA           |
|         | MIGS-22                   | Oxygen                                                               | IDA           |
|         |                             | Facultative                                                          | IDA           |
|         | MIGS-15                   | Biotic relationship                                                  | TAS [5]       |
|         |                             | Endophyte                                                            |               |
|         | MIGS-14                   | Pathogenicity                                                        | NAS           |
|         |                             | None                                                                 |               |
|         | Biosafety level           | 1+                                                                   | TAS [26]      |
|         | MIGS-4                    | Geographic location                                                 | NAS           |
|         |                             | Uppsala, Sweden                                                      |               |
|         | MIGS-5                    | Sample collection time                                              | NAS           |
|         |                             | Summer 1998                                                          |               |
|         | MIGS-4.1                  | Latitude                                                             | NAS           |
|         |                             | 59.8                                                                 |               |
|         | MIGS-4.2                  | Longitude                                                            | NAS           |
|         |                             | 17.65                                                                |               |
|         | MIGS-4.3                  | Depth                                                                | NAS           |
|         |                             | 0.1 m                                                                |               |
|         | MIGS-4.4                  | Altitude                                                             | NAS           |
|         |                             | 24-25 m                                                              |               |

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 [27]. 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.
**Genome sequencing and assembly**

The genome of *S. plymuthica* AS12 was sequenced using a combination of Illumina [30] and 454 sequencing platforms [31]. The detailed information on library construction and sequencing can be found at the JGI website [29]. The sequence data from Illumina GAii (1,800 Mb) were assembled with Velvet [32] and the consensus sequence was computationally shredded into 1.5 kb overlapping fake reads. The sequencing data from 454 pyrosequencing (81.6 Mb) were assembled with Newbler. The initial draft assembly contained 61 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 parallel Phrap [33]. Possible mis-assemblies were corrected with gapResolution [29], Dupfinisher [34], or by sequencing cloned bridging PCR fragments with subcloning or transposon bomb (Epicentre Biotechnologies, Madison, WI). The gaps between contigs were closed by editing in the software Consed [35-37], by PCR and by Bubble PCR (J.-F. Chang, unpublished) primer walks. A total of 160 additional reactions was 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 [38]. The final assembly is based on 47.4 Mb of 454 draft data which provides an average 8.8 × coverage of the genome and 315 Mb of Illumina draft data which provides an average 59 × coverage of the genome.

**Genome annotation**

The *S. plymuthica* AS12 genes were identified using Prodigal [39] as part of the genome annotation pipeline at Oak Ridge National Laboratory.

---

**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 library, one Illumina library |
| MIGS-29 | Sequencing platforms      | Illumina GAii, 454 GS FLX Titanium        |
| MIGS-31.2 | Fold coverage          | 59.0 × Illumina; 8.8 × pyrosequencing     |
| MIGS-30 | Assemblers                | Velvet v. 1.0.13, Newbler v. 2.3, Phrap version SPS – 4.24 |
| MIGS-32 | Gene calling method       | Prodigal 1.4, GenePRIMP                   |
|         | NCBI project ID           | 60453                                     |
|         | INSDC ID                  | CP002774                                  |
|         | Genbank Date of Release   | October 12, 2011                          |
|         | GOLD ID                   | Gc01771                                   |
| MIGS-13 | Source material identifier| CCUG 61397                                |
|         | Project relevance         | Biocontrol, Agricultural                  |

---

**Growth conditions and DNA isolation**

The cells of *S. plymuthica* AS12 were grown in Luria Broth (LB) medium at 28°C with constant shaking at 200 rpm. The cells were harvested after 12 hours when the cells were in the early stationary phase. The cells were pelleted and resuspended in TE buffer (Sigma Aldrich). The DNA was extracted from the resuspended cells by following the standard CTAB protocol for bacterial genomic DNA isolation, which is available at JGI [29].
Neupane et al. (ORNL), Oak Ridge, TN, USA, followed by a round of manual curation using the JGI GenePRIMP pipeline [40]. 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. The miscellaneous functions were predicted using tRNAscan-SE [41], RNAmmer [42], TMHMM [43], and signalP [44]. Additional gene prediction analysis and functional annotation was performed within the Integrated Microbial Genomes – Expert Review (IMG-ER) platform developed by the Joint Genome Institute, Walnut Creek, CA, USA [45].

Genome properties
The genome of *S. plymuthica* AS12 comprises a single circular chromosome of 5,443,009 bp with 55.96% GC content (Figure 3 and Table 3) and 5,140 predicted genes. Among those predicted genes, 4,952 were assigned as protein-coding genes and 88.71% of protein coding genes were assigned for putative function and the remaining ones were annotated as hypothetical proteins. There were 76 pseudogenes and 113 RNA genes with seven rRNA operons. The distribution of genes into the COG functional categories is presented in Table 4.

**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 green, rRNAs red, other RNAs black), GC content, GC skew.
Table 3. Genome statistics

| Attribute                        | Value      | % of total |
|----------------------------------|------------|------------|
| Genome size (bp)                 | 5,443,009  | 100.00%    |
| DNA Coding region (bp)           | 4,772,809  | 87.69%     |
| DNA G+C content (bp)             | 3,045,986  | 55.96%     |
| Total genes\(^b\)                | 5,139      | 100.00%    |
| RNA genes                        | 112        | 2.18%      |
| rRNA operons                     | 7          | 0.14%      |
| Protein-coding genes             | 4,952      | 96.36%     |
| Pseudo genes                     | 75         | 1.46%      |
| Genes in paralog clusters        | 2721       | 52.95%     |
| Genes assigned to COGs           | 3,808      | 74.10%     |
| Genes assigned in Pfam domains   | 4,184      | 81.41%     |
| Genes with signal peptides       | 675        | 13.13%     |
| Genes with transmembrane helices | 1,228      | 23.89%     |
| 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 | % age | 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  | 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 COGs                                                                  |
Acknowledgements
We gratefully acknowledge the help of Elke Lang for providing a culture of reference bacterial strains, Evelyne-Marie Brambilla for extraction of DNA, and Anne Fiebig for assembly of the reference genomes required for digital DNA-DNA (all at DSMZ). The work was 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 PAD, Grimont F, Starr MP. Serratia species isolated from plants. Curr Microbiol 1981; 5:317-322. http://dx.doi.org/10.1007/BF01567926
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
3. 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
4. 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
5. 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
6. 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
7. 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
8. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGAS: 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
9. 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
10. 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 2010; 38:D346-D354. PubMed http://dx.doi.org/10.1093/nar/gkp848
11. 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
12. 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.
13. 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
14. 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
15. 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
16. Garrity GM, Bell JA, Lilburn T. Phyllum XIV. Proteobacteria phyl. nov. In: Garrity GM, Brenner DJ, Krieg NR, Staley JT (eds), Bergey’s Manual of
Serratia plymuthica strain AS12

Systematic Bacteriology, Second Edition, Volume 2, Part B, Springer, New York, 2005, p. 1.

17. 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

18. 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.

19. 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.

20. 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

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

22. 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. http://dx.doi.org/10.1099/0096266X-8-1-73

23. 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, 1994, p. 326-326.

24. 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.

25. 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.

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

27. 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

28. 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

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

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

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

32. 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

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

34. 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.

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

39. 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

40. 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

41. Schattner P, Brooks AN, Lowe TM. The tRNA scan-SE, 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

42. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007; 35:3100-3108. PubMed http://dx.doi.org/10.1093/nar/gkm160

43. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 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

44. 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

45. 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