Arthrobacter sp. EpRS66 and Arthrobacter sp. EpRS71: Draft Genome Sequences from Two Bacteria Isolated from *Echinacea purpurea* Rhizospheric Soil

Luana Presta[1], Marco Fondi[1], Elena Perrin[1], Isabel Maida[1], Elisangela Miceli[1], Carolina Chiellini[1], Valentina Maggini[1], Patrizia Bogani[1], Vincenzo Di Pilato[2], Gian M. Rossolini[3, 4, 5, 6], Alessio Mengoni[1] and Renato Fani[1]*

[1] Department of Biology, University of Florence, Florence, Italy, [2] Department of Surgery and Translational Medicine, University of Florence, Florence, Italy, [3] Department of Medical Biotechnologies, University of Siena, Siena, Italy, [4] Department of Experimental and Clinical Medicine, University of Florence, Florence, Italy, [5] Clinical Microbiology and Virology Unit, Careggi University Hospital, Florence, Italy, [6] Don Carlo Gnocchi Foundation, Florence, Italy

INTRODUCTION

One of the most promising, but still overlooked fields of microbiological research is represented by endophytic microorganisms, i.e., those organisms living in the tissues of host plants and/or in their rhizosphere (Rosenblueth and Martínez-Romero, 2006; Reinhold-Hurek and Hurek, 2011). These microbes are emerging as a new potential source of secondary metabolites and products, for exploitation in medicine, agriculture, and industry.

From a biotechnological perspective, a controlled (engineered) colonization of plant's tissues by some bacteria may be desirable because of their ability to produce a variety of plant growth promoting (PGP) molecules, spanning from siderophores, nitrogenases, hormones, and so on. In addition, host-microbe interaction confers indirect advantage to the plant, resulting from the inhibition activity exerted by the associated microbial community toward potential pathogens.

Moreover, in the field of medicine, there are examples of very well-known molecules derived by endophytes like antibiotics, antimycotics, and anticancer drugs. Also, it is still unknown if plant-associated bacteria may enhance (or be responsible for) some of the effects exerted by the extracts of medicinal plants (essential oils) (Kloepper and Ryu, 2006; Hardoim et al., 2008).

In this regard, in October 2012, in Casola Valsenio (Italy), a collection of microorganisms was isolated from both internal tissues and the rhizospheric soil of the medicinal plant *Echinacea purpurea*, as reported in Chiellini et al. (2014). Among others, two strains sampled from the rhizosphere and belonging to *Arthrobacter* species were identified, i.e., *Arthrobacter* sp. EpRS66 and *Arthrobacter* sp. EpRS71. Based on their antibiotic resistance profile, reported in Mengoni et al. (2014), and on further tests performed on these two strains, they were selected as good candidates for genome sequencing analysis. The last, will constitute a resource to deeply investigate their genomic features and to perform comparative genomics analysis. Moreover, in the aim of new drugs discovery, the genome sequence will facilitate the identification of putative genes responsible for the production of bioactive compounds.

MATERIALS AND METHODS

DNA Extraction and Sequencing

Arthrobacter sp. EpRS66 and *Arthrobacter* sp. EpRS71 strains were inoculated overnight on TSB medium at 30°C. Their genomic DNA was then extracted using the CTAB method (Perrin et al., 2015). Furthermore, the authenticity of the genomic DNA was confirmed by 16S RNA gene sequencing.

MiSeq sequencing system (Illumina Inc., San Diego, CA) was used to perform the whole genome shot-gun of the two organisms. The method used a 2 × 300 bp paired-end approach, which produced a genome coverage of 246.0 x for *Arthrobacter* sp. EpRS66 and 91x for *Arthrobacter* sp. EpRS71.

Genome Assembly and Annotation

The quality of the obtained read pairs was evaluated by inspecting them with FastQC software package v. 0.52 (Kunde-Ramamoorthy et al., 2014). Poor quality bases were removed with StreamingTrim (Bacci et al., 2014). *De novo* assembly was performed by using SPAdes 3.5 software (Bankevich et al., 2012) with a k-mer length of 21, 33, and 55. After, those contigs with length inferior to 2000 bp were trimmed and the remaining (6 and 24 for *Arthrobacter* sp. EPRS66 and *Arthrobacter* sp. EPRS71, respectively) were launched in a multi-draft based analysis through MeDuSa scaffolder (Bosi et al., 2015), by using as references 5 *Arthrobacter* genomes retrieved at NCBI database (*Arthrobacter arilaitensis* Re117, *Arthrobacter* FB24, *Arthrobacter* Rue61a, *Arthrobacter aurescens* TC1, *Arthrobacter chlorophenolicus* A6).

Automated annotation of the two draft genome sequences has then been performed with NCBI Prokaryotic Genome Annotation Pipeline.

RESULTS

The last version of *Arthrobacter* sp. EpRS66 genome has a total length of 3,707,708 bp and embeds only 2 scaffolds (L50 equal to 1), with a mean G+C content of 59.27%. The annotation analysis identified a total of 3485 genes, of which 3383 have been annotated as coding DNA sequences (CDS), 29 as pseudogenes, 4 as rRNAs, 68 as tRNAs, and 1 as ncRNA.

The draft genome sequence of *Arthrobacter* sp. EpRS71 24 is 4,849,450 bp long and its contigs are set-up in 10 scaffolds (L50 equal to 1). The G+C content is 61.60%, a value slightly higher than the previous but still perfectly comparable with that of other *Arthrobacter* genomes sequenced so far. The annotation of *Arthrobacter* sp. EpRS71 genome revealed the presence of 4515 genes. This total amount includes 4379 proteins coding sequences, 71 pseudogenes, and 62 RNA (6 rRNAs, 55 tRNAs, 1 ncRNA) coding sequences.

Both genome sequences have been deposited at NCBI database and are available in both fasta and GenBank format; the GenBank accession number of *Arthrobacter* sp. EPRS66 is LNUU00000000 and the version reported in this work was named LNUU01000000; the GenBank accession number of *Arthrobacter* sp. EPRS71 is LNUV00000000 and the version reported in this work is LNUV01000000.

AUTHOR CONTRIBUTIONS

This project was planned by RF and AM. The DNA extraction was performed by IM and EP. The DNA sequencing has been performed by GR and VD. The data processing has been performed by LP and MF. CC, VM, PB, and EM assisted substantially on the technical part of this work. All author contributed to writing and editing the present manuscript.

FUNDING

REFERENCES

Bacci, G., Bazzicalupo, M., Benedetti, A., and Mengoni, A. (2014). StreamingTrim 1.0: a Java software for dynamic trimming of 16S rRNA sequence data from metagenetic studies. *Mol. Ecol. Resour.* 14, 426–434. doi: 10.1111/1755-0998.12187

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Bosi, E., Donati, B., Galardini, M., Brunetti, S., Sagot, M.-F., Lió, P., et al. (2015). MeDuSa: a multi-draft based scaffolder. *Bioinformatics* 31, 2443–2451. doi: 10.1093/bioinformatics/btv171

Chiellini, C., Maida, I., Emiliani, G., Mengoni, A., Stefano, M., Fabiani, A., et al. (2014), Endophytic and rhizospheric bacterial communities isolated from the medicinal plants *Echinacea purpurea* and *Echinacea angustifolia*. *Int. Microbiol.* 17, 165–174. doi: 10.2436/20.1501.01.219

Hardoim, P. R., van Overbeek, L. S., and Elsas van, J. D. (2008). Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol.* 16, 463–471. doi: 10.1016/j.tim.2008.07.008

Kloepper, J. W., and Ryu, C.-M. (2006). "Bacterial endophytes as elicitors of induced systemic resistance," in *Microbial Root Endophytes*, eds B. Schulz, C. Boyle, and T. N. Sieber (Berlin; Heidelberg: Springer), 33–52. doi: 10.1007/3-540-33526-9_3

Kunde-Ramamoorthy, G., Coarfa, C., Laritsky, E., Kessler, N. J., Harris, R. A., Xu, M., et al. (2014). Comparison and quantitative verification of mapping algorithms for whole-genome bisulfite sequencing. *Nucleic Acids Res.* 42:e43. doi: 10.1093/nar/gkt1325

Mengoni, A., Maida, I., Chiellini, C., Emiliani, G., Mocali, S., Fabiani, A., et al. (2014). Antibiotic resistance differentiates *Echinacea purpurea* endophytic bacterial communities with respect to plant organs. *Res. Microbiol.* 165, 686–694. doi: 10.1016/j.resmic.2014.09.008

Perrin, E., Fondi, M., Maida, I., Mengoni, A., Chiellini, C., Mocali, S., et al. (2015). Genomes analysis and bacteria identification: the use of overlapping genes as molecular markers. *J. Microbiol. Methods* 117, 108–112. doi: 10.1016/j.mimet.2015.07.025

Reinhold-Hurek, B., and Hurek, T. (2011). Living inside plants: bacterial endophytes. *Curr. Opin. Plant Biol.* 14, 435–443. doi: 10.1016/j.pbi.2011.04.004

Rosenblueth, M., and Martínez-Romero, E. (2006). Bacterial endophytes and their interactions with hosts. *Mol. Plant. Microbe. Interact.* 19, 827–837. doi: 10.1094/MPMI-19-0827

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.