

*MICROSCOPE : A PLATFORM FOR MICROBIAL GENOME ANNOTATION AND
COMPARATIVE GENOMICS.*

D. Vallenet, S. Engelen, D. Mornico, S. Cruveiller, A. Lajus, Z. Rouy, D. Roche, G.
Savignol & C. Médigue

CEA/DSV/IG/Genoscope-CNRS UMR8030, Laboratoire de Génomique Comparative (LGC), 2
rue Gaston Crémieux, 91057 Evry Cedex, France

Keywords : *Genomics, Comparative genomics, Automatic and Expert Annotation, Biocuration*

Since 6 years, the LGC group is involved in the development of the MicroScope platform which allows to support microbial genomes (re)annotation projects and comparative analysis (<http://www.genoscope.cns.fr/agc/microscope>). It is made of three major components described in Vallenet *et al.* (NAR 34:53-65, 2006): (i) a wide range of syntactic, functional and relational annotation tools, (ii) a relational database, Prokaryotic Genome DataBase, (PkGDB) which is linked to metabolic pathway databases (<http://www.genoscope.cns.fr/agc/microcyc>) and (iii) a graphically oriented web interface, Magnifying Genome (MaGe, <http://www.genoscope.cns.fr/agc/mage>). A thematic database includes newly sequenced genomes together with related chromosomes/plasmids (finished or not) for which a complete automatic annotation, including synteny groups computation with available microbial proteomes, is performed. The data is made available through advanced graphical Web interfaces for manual refinement of the automatic assignment of gene products, and for the exploration of the annotation data and the comparative analysis results. Beside the 4 days training program proposed to the users, we also provide continuous assistance to experts in their annotation task.

We will present a general overview of the MicroScope platform and several new functionalities of the MaGe's interfaces: advanced searches interface allowing to query automatic/expert annotations and any results generated by the analysis pipeline, several dynamic procedures allowing to compute KEGG pathways for which genes are also involved in synteny groups, to compute genomic islands in the studied bacterial genome, and to compare metabolic networks of several selected genomes (using BioCyc or KEGG data). Moreover, a 'Gene card' interface allows to manipulate several set of genes obtained at different levels of the MaGe functionalities (Keywords, Phyloprofile synteny, Genomic islands). One of the future extensions of the MicroScope platform will be related to the curation of metabolic networks in microbial genomes: corresponding tools and interfaces will be developed in the context of a European PF7 project called MICROME.

The platform is used as a viewer in order to browse updated annotation information of available bacterial genomes (to date more than 300 organisms), and in the context of new annotation projects (92 bacterial genomes, of which 60 are sequenced at Genoscope, 12 at DOE, 6 at Sanger Institute). Since 2004, 21 complete bacterial genomes have been published in 12 papers, and 12 additional articles are related to specific analysis (comparative genomics, transcriptome and proteome analysis). The expert annotations gathered in PkGDB (to date, 240 000 created by more than 570 curators having a personal account on the thematic databases) contribute to improve the quality of the bacterial annotations, at least for genomes initially annotated in an automatic way only.

Acknowledgment : This work is supported by the French Ministry ANR PFTV 2007