# Systematic functional analysis of *Bacillus subtilis* with Micado, an integrative database for microbes

Véronique Biaudet, Shahinaz Gas, Franck Samson, Étienne Dervyn
S. Dusko Ehrlich & Philippe Bessières

*Génétique Microbienne – INRA – 78350 – Jouy-en-Josas*
*v: 33 1 34 65 25 32 – f: 25 21 – philb@biotec.jouy.inra.fr*
*http://locus.jouy.inra.fr/micado*

Micado is a relational database for microbial genomes, that is vertically integrating — from sequence to phenotype — annotated chromosome sequence, genetic map, genes metabolism classification, and functional analysis data of *Bacillus subtilis*. Physiological information concerning the disruption of 1200 unknown genes has been collected into the database, by an European consortium of 17 laboratories. The comparative analysis takes into account growth characteristics of mutant strains, reporter gene activity, mRNA expression, links on protein 2D gels data (Sub2D, at Greifswald U. [2]), and phenotype tests, by far the most diverse item to organize.

20 protocols, structured in 5 large classes of metabolism, have been defined in the database, for testing the mutant phenotypes at a first level: small molecules and inorganics, macromolecules, cell structure and mobility, stress and stationnary phase, and finally, cell processes. Surprisingly, half of the "knock-outed" unknown genes have been giving an observable phenotype, and an important part is implied in the response to osmotic and starvation stresses.

This invaluable new knowledge is opening to stimulating research perspectives, and should benefit first to bacterial genomics, by example to begin with the gram-positive lactic acid bacteria, actively studied for their applied interest. To help elucidating the function of unknown genes, and filling the gap from the protein activity to the phenotype, is calling for new extensions of the database. This is done by setting up diverse multidisciplinary and collaborative projects, to cross the phenotypes data with information such as the phylogeny of sequences, related gene expression and regulatory motifs on DNA, or structural homologies, found by protein threading.

1. Biaudet V, Samson F, Bessières P. (1997). Micado: a network oriented database for microbial genomes. *Comput. Applic. Biosci.* **13**, 431-438.
2. http://microbio2.biologie.uni-greifswald.de:8880