Project No. *222886-2*

MICROME

The MICROME Project:
A Knowledge-Based Bioinformatics Framework for Microbial Pathway Genomics

Instrument: **Collaborative project**

Thematic Priority: **KBBE-2007-3-2-08: BIO-INFORMATICS - Microbial genomics and bio-informatics**

D7.4 Second MICROME Training Course

Due date of deliverable: November 30, 2012 (M36)
Actual submission date: September 4, 2012 (M34)

Start date of project: 1.12.2009 Duration: 48 months

Organisation name of lead contractor for this deliverable: SIB

Project co-funded by the European Commission within the Seventh Framework Programme (2009-2013)		
Dissemination Level		
PU	Public	PU

Contributors

Eugeni Belda (CEA/Genoscope)
Francois Le Fevre (CEA/Genoscope)
Damien Mornico (CEA/Genoscope)
David Vallenet (CEA/Genoscope)
Claudine Médigue (CEA/Genoscope)
Christos Ouzounis (CERTH)
Fotis Psomopoulos (CERTH)
Alessandro Vullo (EBI)
Paul Kersey (EBI)
Pierre-Yves Bourguignon (Isthmus)
Anne Morgat (SIB)
Raphy Zarecki (TAU)
Didier Croes (ULB)
Jacques van Helden (ULB)
Karoline Faust (Ex-ULB, currently VIB-VUB, Belgium)

INTRODUCTION

Deliverable reference number:

D7.4 Second MICROME Training Course

Expand deliverable by 2 or 3 lines

1. General introductory course
2. Training in metabolism-related bioinformatics with specific sessions for curators and bioinformaticians
3. Refinement of the teaching material that remains from course to course

Methods

Briefly outline in 5 – 10 lines the methods used for this deliverable

The goal of the training course was to build an organized resource for training users of the MICROME resources (databases, curation tools, analysis tools), and to use this resource for a 2-days training primarily (but not exclusively) focused on members of the MICROME project. The training course was organized by Jacques van Helden (ULB) and Claudine Médigue (CEA). It took place at Aix-Marseille Université (France) from March 28 to March 30, 2012 and gathered 20 participants. Teaching material and tutorial sessions (slides, tutorials, sample files) were contributed by 7 partners (CEA, CERTH, EBI, Isthmus, SIB, TAU, ULB). All the teaching material (slides, tutorials, sample files) was organized in a Web site (http://rsat.ulb.ac.be/MICROME/training/training_Marseille_2012-03/), which will remain available after the training, and will be progressively upgraded with future progress of the project.

Results (if applicable, interactions with other workpackages)

The report must have self contained evidence that this deliverable has been achieved; the report should be 1 to 2 pages long, but should ideally include one or two figures. If large data files are required, either make an appendix at the end of the report or, if the data files are very large, upload to the MICROME web site and show the first page in an appendix

Attendance

The course gathered 20 participants, including the trainers (PIs and MICROME postdocs) plus some trainee postdocs and PhD students from MICROME partners, plus 2 Master students (Jonathan Verneau and Quentin Da Costa) and one Professor (Laurent Tichit) from Aix-Marseille Université. These students participated to the MICROME project by developing client scripts for MICROME Web services, under the joined supervision of Jacques van Helden and Laurent Tichit. These Web service client scripts were shown during the last session.

Teaching material

Before the training course, the teaching material was organized in a cooperative way, by creating a shared folder (on Dropbox), with a sub-folder for each contributed partner. The final versions were placed on a Web site that was used by participants during the training sessions.

Web site of the Second MICROME Training Course (http://rsat.ulb.ac.be/MICROME/training/training_Marseille_2012-03/).

Programme

Day	From	To	Trainer name(s)	Partner	Title
Wed 28 Mar	13:30	14:00	Jacques van Helden	ULB	Welcome + configuration of the accounts and laptops
Wed 28 Mar	14:00	16:00	Anne Morgat	SIB	Browsing metabolic pathways in UniPathway
Wed 28 Mar	16:00	16:15			Coffe break
Wed 28 Mar	16:15	18:15	Damien Mornico , Eugenio Belda, David Vallenet, François Lefèvre	CEA/LAB GeM	Searching for gene reaction gaps and GPRs curation using MetaCyc and Rhea resources in MicroScope
Wed 28 Mar	18:15				End of day 1 training sessions
Wed 28 Mar	19:30	22:00			Workshop dinner
Thu 29 Mar	09:00	11:00	Alessandro Vullo	EBI	Pathway curation and analysis in MICROME (using the Reactome toolset)
Thu 29 Mar	11:00	11:15			Coffe break
Thu 29 Mar	11:15	13:15	Karoline Faust, Didier Croes, Jacques van Helden	ULB + AMU	Predicting operons, regulons and metabolic pathways from bacterial genomes
Thu 29 Mar	13:15	14:15			Lunch

Thu 29 Mar	14:15	16:15	Fotis Psomopoulos	CERTH	Constructing ancestral gene content from genome trees and phylogenetic profiles - Attempts to profile ancestral pathway content for MICROME
Thu 29 Mar	16:15	16:30			Coffe break
Thu 29 Mar	16:30	17:00	Raphy Zarecki, Alik Peltinovich	TAU	Reconstructing metabolic models from the MICROME database (short lecture to gather feedback from participants)
Thu 29 Mar	17:00	18:00	All MICROME partners	MICROME	Methodology for annotating bacterial metabolism
Thu 29 Mar	18:00				End of day 2 training sessions
Fri 30 Mar	08:30	09:45	Pierre-Yves Bourguignon	Isthmus	Metabolic CAD
Fri 30 Mar	09:45	10:30	François Le Fevre, Jacques van Helden, Paul Kersey	CEA/LAB GeM,ULB, EBI	MICROME Web services
Fri 30 Mar	10:30	10:45			Coffe break
Fri 30 Mar	10:45	11:45	All MICROME developers	MICROME	Practical: developing clients for Web services.
Fri 30 Mar	11:45	12:45	All MICROME partners	MICROME	Discussion Integrating MICROME tools and databases
Fri 30 Mar	12:45	13:15	All MICROME partners	MICROME	Evaluation of this training workshop and perspectives for future events
Fri 30 Mar	13:15				End of training

Interactions between partners

The organization of the course fostered collaborations between MICROME partners. For example, the course was the occasion to test and show the first Web service connection between two partners (CEA and ULB). Sample scripts were developed in Python and in Perl, and made available in the teaching material. Beyond teaching, these web clients were used to extract metabolic networks and Gene-Protein-Reaction (GPR) associations from Genoscope (developed by partner CEA), and install them in the pathway extraction tool of the Network Analysis Tools (developed by partner ULB). The workshop was also the opportunity to extend this idea of concretizing interactions, and we started a collaborative document listing the interactions between resources produced by some partners and required by other partners.

You must also explicitly mention the milestones achieved, which in this case are
M n.n. Number of participants: 20.

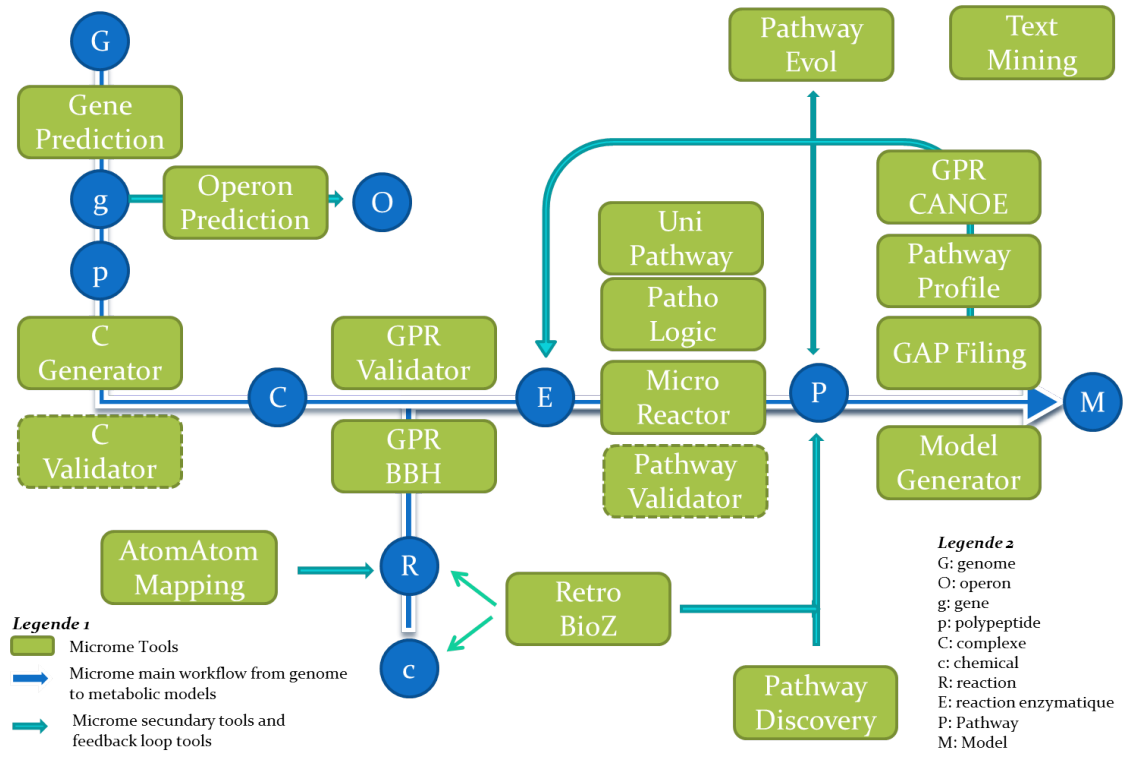

Figure 1: different steps dedicated to generate metabolic models directly from Genome. It is symbolized by a blue/white arrow going from Genome (G) to Model (M). Each processing step could be realized by different tools (green box). A green box with dashes represents a tool in development and not in production. Box inputs are mainly located at the left of each box whereas output at the right. Input/Output are symbolized by blue circle elements.

MICROME partners have developed a set of tools able to process each step. The first step is a gene prediction tool (Gene prediction), which parses the input Genome (G) to find out genes (g). CEA has developed its own tools, but MICROME pipeline relies also directly on genes defined by NCBI RefSeq. Gene protein products (p) are directly linked through their Uniprot identifier (Uniprot/Trembl). A complex generator (C Generator) is able to propose protein complexes (C) based on polypeptides subunits homology between a set of reference complexes and the study organism. These proteins complexes could be validated by an expert human annotation through a future dedicated web interface tool (C Validator). Chemicals (c) are defined in ChEBI reference database to build reactions (R) in Rhea reference database. Those reactions are then linked to an enzyme (E) by an automatic tool (GPR BBH) based mainly on sequence homology and could be also validated by an expert human annotation (GPR Validator), a dedicated web interface. These enzymatic reactions are then projected on reference Pathways (P) provided by the couple, Reactome (MetaCyc based) and Unipathway resources. MICROME partners have developed several tools for this projection module, they are currently under evaluation in order to find the best one between Unipathway, PathoLogic and MicroReactor modules. For instance MicroReactor, module made by EBI, links automatically proteins to Rhea reaction based on InterPro domain profiles. Nevertheless a unique tool is under development in order to assert the status of each projected pathway (Pathway Validator). Finally these pathways are aggregated by a Model Generator tool to define a draft metabolic model (M). This draft metabolic model is then processed by GPR Canoe, Pathway Profile and Gap filling modules in order to full fill the missing enzymatic reactions. These modules produce as output additional enzymatic reactions that could serve again as seed to find out new projected pathways. There is also a set of additional modules dedicated to Pathway analyses from discovery, to evolution (Pathway Evol /Pathway Discovery coupled with Operon Prediction). The RetroBio Synthesis module integrates information extracted from the AtomAtomMapping module, in charge to resolve reaction definition at the level of atom transfer, in order to design new pathways. At term the Text mining module should be able to propose input proposal for several modules.

Perspectives

3-5 lines on future plans in this area

This second MICROME Training course was primarily addressed to MICROME partners, in order to stimulate the cooperative elaboration of a corpus of teaching material, test the validity of the tutorials, and learn to use each other's resources. A third and last training course should be organized before the end of the project, and open to external trainees. The course should target biologists (with the user-friendly interfaces of the databases and tools) as well as bioinformaticians (learning to use the programmatic interfaces to the tools).

Publications
Please put in all publications which have had MICROME involvement This deliverable did not give rise to any publication, but resulted in the creation of a publicly available Web site with all the teaching material.