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**

D2.4
Microme SOPs

Due date of deliverable: 31/08/2011
Actual submission date: 26/08/2011

Start date of project: 1.12.2009

Duration: 48 months

Organisation name of lead contractor for this deliverable: CEA

Project co-funded by the European Commission within the Seventh Framework Programme (2009-2013)		
Dissemination Level		
PU	Public	x
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contributors
The document has been prepared by CEA, with contributions from SIB, AMB, ULB and MN

INTRODUCTION
Deliverable reference number: D2.4 Microme SOPs This document aims at describing specific Standard Operating Procedures (SOPs) to be used by Microme curators.

Methods

The SOPs will follow the recommendations by Angiuoli *et al.* (2008, http://www.liebertonline.com/doi/pdf/10.1089/omi.2008.0017). A SOP template was defined and the used model was inspired from the JCVI annotation procedures for genome annotation (http://cmr.jcvi.org/CMR/AnnotationSops.shtml)
 	
The primary target of the SOPs is the MICROME curators:
   - trained in biology
   - with no specific programming skills
The procedures should enable new curators to use all the facilities of the MICROME project in order to curate all the components of the project (compounds, reactions, enzymes with gene-protein-reaction associations, pathways, and genome-wide metabolic networks).
 	
After the first Microme annotation jamboree (see deliverable D7.2), the experience gained was used to define a list of SOPs. Two main categories were defines:
   - SOPs to curate Reference Data
   - SOPs to curate Projected Data
 	
The SOPs will be used during the training sessions, and improved by taking into account the feedback of the trainees.

Results

1 SOP template

A SOP template was defined as follow and will serve as a basis for the future SOPs.

MICROME SOPs		
Title:		Pages:
SOPNumber	Revision level:	Effective date:
Author(s):		

1-Overview

Short introduction with the scientif ic context.

1.1-Scope

Goal of the SOP with a description of the main biological entities that will be involved.

1.2-Related documents

Other related SOPs (SOPNumber – SOP title):

Publications:

Other documents:

1.3-Revision history

Author	Date	Change

2-Requirements

Materials and data that are required to correctly accomplish the SOP.

3-Procedure

Description of the dif f erent steps to achieve the procedure.

4-Data management

4.1-Quality control

Relevant indicators which are usef ul to conf irm the correct execution of the procedure.

2 SOPs listing

This SOP list was established after the first Microme jamboree (see deliverable D7.2).
*: A first version of the SOP is available on the Microme web site (http://www.microme.eu/pages/microme-sops).

2.1 SOPs to curate Reference Data

SOP01: Compound curation in ChEBI environment*

Submission of chemical compounds missing in the ChEBI database but necessary to create novel reactions in the Rhea reaction database.

SOP02: Reaction curation in Rhea environment

Creation of chemical reactions in Rhea database. The reactions chemically balanced at pH 7.3 are necessary for the subsequent steps of pathway curation and creation by Microme curators.

SOP03: *De-novo* pathway creation in Microme environment

Creation of pathway events in Microme central repository from the whole set of Rhea reactions and ChEBI compounds that define it. Association of reaction events to its corresponding catalyst activities based on UniProtKB entities.

SOP04: *De-novo* protein complex creation in Microme

Creation of complex entities in Microme central repository from the UniProtKB identifiers of its corresponding subunits with well defined subunit stoichiometry.

2.2 SOPs to curate Projected Data

2.2.1 SOPs to curate Projected Data SOPs related with Genome centric curation

SOP05: Pathway prediction curation in MicroScope environment

Curation of automatically generated pathway projections using the different tools available in the MicroScope platform, which allows curators to validate the existence of a pathway in a given organism and identify candidate gene-protein-reaction (GPR) associations for pathway holes in the initial projection.

SOP06: Gene-Protein-Reaction (GPR) association in MicroScope environment*

Use of GPR association functionality of the MicroScope platform to link specific MetaCyc reactions to the corresponding UniProtKB entries.

SOP07: Procedures for transport validation based on Biolog data

Reconciliation of growth phenotypes experimentally observed in Biolog experiments with Specie-Specific Pathway Assemblies through the curation of GPR assignment for transport reactions associated with metabolites for which a positive growth phenotype is observed in Biolog plates.

2.2.2 SOPs related with vertical propagation of GPR assignments to multiple genomes

SOP08: Gene families reconstruction protocol

Reconstruction of gene families across query genomes in Microme with the use of EnsemblCompara pipeline.

SOP09: Procedures for automatic reaction projection based on gene families

Projection of GPR assignments of reference genomes in Microme to other bacterial genomes based on the gene family profiles identified with EnsemblCompara pipeline.

SOP10: Procedures for automatic pathway projection

Projection of GPR on the Microme reference pathway dataset and evaluation of pathway existence in a given organism.

Perspectives

The description of the SOPs is an on-going process and they will be produced during the entire Microme project. Additional SOPs will be prepared for bioinformaticians / computer scientists involved in the project, in order to document the automatic procedures for global analyses (genome annotation, metabolic path finding, prediction of synteny groups, regulons, operons, comparisons of pathways ...).

The SOPs that may be useful for a wider audience will be reformatted in order to publish them in specialized journals (Methods in Molecular Biology, Nature Protocols) or on the specialized Web site Protocol Exchange (http://www.nature.com/protocolexchange/).

Annexes

The SOP template and a first version of SOP01 and SOP06 are available at http://www.microme.eu/pages/microme-sops.