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**

D7.3 Second Microme Annotation Jamboree

Due date of deliverable: M27
Actual submission date: April 2012 (M29)

Start date of project: 1.12.2009 Duration: 48 months

Organisation name of lead contractor for this deliverable: CNIO

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

Contributors

Responsible Beneficiary: CEA
Lead: CNIO
Organization: CSIC, EMBL-EBI
Coordination and preparation: SIB, KO & CEA
Staff initially involved in DOW: EMBL-EBI, CEA, SIB, DSMZ, WUR
Additional staff involved in the jamboree annotation: CERTH, WTSI, MN, KO and TAU.

INTRODUCTION

Deliverable reference number: ***D7.3 Second Microme Annotation Jamboree***

This document summarizes the work done during the second Microme Annotation Jamboree that took place at the Mediterranean Institute for Advanced Studies (IMEDEA) in Mallorca on the 17th of April 2012.
The objective of this jamboree was to perform vertical annotation on a selected pathway in order to get curated data and to refine the Microme curation pipeline.
We focused on the curation of the biosynthesis pathway of L-lysine amino acid for a set of 88 microbial genomes.
The result of several reaction projection strategies have been integrated. Based on this pre-compiled data set, the participants have curated gene-reaction associations (approving/rejecting predicted associations or adding missing associations).
The jamboree provided an opportunity to share expertise among partners.

ORGANISATION

Location : IMEDEA - Mediterranean Institute for Advanced Studies, C/ Miquel Marquès, 21, 07190 Esporles, Mallorca, Spain. *(http://www.imedea.uib.es/),*

Number of participants: 25 (3 from EBI, 1 ULB, 2 WTSI, 1 MN, 2 CERTH, 1 CNIO, 3 CSIC, 1 WUR, 2 DSMZ, 2 SIB, 2 AMB, 3 CEA, 1 NYU School of Medicine from the Microme SAB).

The session started by short presentations to provide all information relevant to the jamboree. *i.e* a short survey of the L-lysine biosynthesis pathway and the two platforms that will be used during the Jamboree.

Participants have worked in small groups during two curation sessions.

At the end of the day, each group presented its concluding remarks.

Agenda:

08h30-09h00 : Configuration of participant laptops
09h00-10h00 : Introduction (Lysine case study, data exploration in UniPathway)
10h00-13h00 : Free curation on selected genomes/subpathways - part 1 -
13h00-14h00 : Lunch
14h00 -15h00 : Sharing of expertise on the different curation strategies used by the participants
15h00-15h30 : Results of the curation process part 1
15h30-17h30 : Free curation on selected genomes/subpathways - part 2 -
17h30-18h30: Global results of the curation, discussions, and conclusions

DATA and METHODS

All the information related to the L-lysine biosynthesis case-study has been collected in the Microme confluence web site:

http://www.ebi.ac.uk/seqdb/confluence/display/Microme/Data+Jamboree+2012

1) Lysine biosynthesis – state of the art

The biosynthesis of the basic amino-acid L-lysine occurs via two distinct anabolic routes that evolved separately, the diaminopimelate (DAP) and aminoadipate (AAA) pathways :

- The **DAP pathway** synthesizes L-lysine from aspartate and uses diaminopimelate as an intermediate.
- The **AAA pathway** synthesizes L-lysine from 2-oxoglutarate and uses L-alpha-aminoadipate as an intermediate.

The table below gives the mappings between the descriptions of lysine biosynthesis that are provided by three metabolic pathway resources (UniPathway, KEGG, MetaCyc). The UniPathway description has been used as reference data set during this jamboree.

UniPathway	MetaCyc	KEGG
L-lysine biosynthesis (UPA00404)	Pathways Class : Lysine biosynthesis	Lysine biosynthesis map00300
L-lysine biosynthesis from DAP pathway (UPA00034)		
variant 1	lysine biosynthesis III (PWY-2942)	Lysine biosynthesis map00300
variant 2	lysine biosynthesis VI (PWY-5097)	
variant 3	lysine biosynthesis I (DAPLYSINESYN-PWY)	
variant 4	(lysine biosynthesis II PWY-2941)	
L-lysine biosynthesis from AAA pathway (UPA00033)		Lysine biosynthesis map00300
variant 1	lysine biosynthesis IV (LYSINE-AMINOAD-PWY)	
variant 2	lysine biosynthesis V (PWY-3081)	

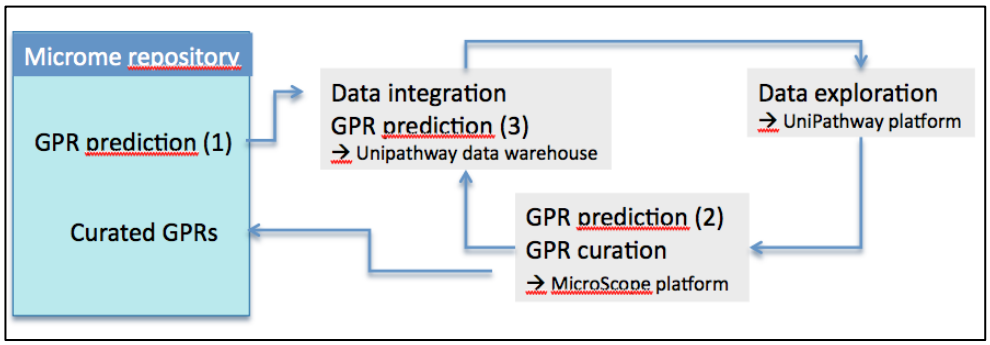
2) Pathway projection strategies

The table below describes the resources used during the jamboree.

Source	Method	Data
Microme genome-reaction matrix (D2.5)	GPR prediction (1) Projection strategy chaining InterPro - GO - EC - Rhea	ENA gene - UniProt - Rhea reaction
Microscope	GPR prediction (2) PathoLogic annotation process (excepted *B.*	MicroScope/Refseq gene - UniProt - MetaCyc reaction

	subtilis). Annotation strategy chaining GO terms, names, EC number mappings to MetaCyc reactions	
UniProtKB/Swiss-Prot	Manual curation	UniProt protein - UniPathway enz-reaction (UER) UniProt protein - EC numbers
UniProtKB/TrEMBL	UniProt automatic annotation process	UniProt protein - UniPathway enz-reaction (UER) UniProt protein - EC numbers
Microme/UniPathway	GPR prediction (3) Projection strategy chaining a curated set of predictors (InterPro, HAMAP) to annotate UniPathway/Rhea reactions These Reaction annotation rules differentiate isozymes and complexes	UniProt protein - UniPathway (UER) / Rhea

All these data sets have been integrated into UniPathway in order to provide mapping to the Microme reference data, namely UniProtKB for proteins and Rhea for reactions) and to facilitate navigation through different resources. In particular, to facilitate access to the Gene-Reaction curation page of the MicroScope Platform. The curation cycle is described in the figure below.



3) Genome data set

The idea was to have a limited but representative set of organisms that would be manageable in the time course of a jamboree.

We used the microbial UniProt reference proteome set (http://www.uniprot.org/taxonomy/complete-proteomes) and completed this list with two reference organisms for Microme (*Acinetobacter baylii*, *Pseudomonas putid*a), and obtained a list of 88 species for which we mapped proteome, genome, taxonomy identifiers in different resources (http://www.grenoble.prabi.fr/zeo4/obiwarehouse/microme/proteome_list).

4) Pathway completion

The projection of the reactions of the L-lysine biosynthesis pathway can give 3 results :

Species with complete pathway
- these species may have one apparently complete pathway variant
- or, more rarely, they may have multiple apparently complete pathway variants

Species with incomplete pathway
Depending on which of the reactions are missing, it may or may not be possible to predict which of the variants is likely present, but incomplete. Hence an incomplete pathway may be a:
- probable known variant, yet incomplete
- an unknown variant, which is also incomplete

Species missing the pathway
If no reactions (and hence no putative pathway variant) can be identified, then:
- either the species do not produce L-lysine at all,
- or the species do not produce L-lysine through the DAP pathway, but rather through the AAA pathway or another as yet unknown set of chemical transformations)

http://www.grenoble.prabi.fr/zeo4/obiwarehouse/unipathway/upa?upid=UPA00033

Obviously, the set of GPRs of apparently complete pathways also need to be evaluated too as some reactions may have been erroneously predicted.

We can explain 'missing' reactions (*i.e* reactions without any detected catalysts) in the following ways:

a) Prediction issue

• The protein exists but the sequence has diverged and has not been detected with the default threshold used by the detection methods.

• Functional convergence. The protein exists but it doesn't share sequence similarity with known proteins that perform this function.

In both cases, the prediction could be improved by genomic context analysis (gene neighbouring, synteny, phylogenetic footprints) or phylogenetic profiles.

b) Pathway definition issue

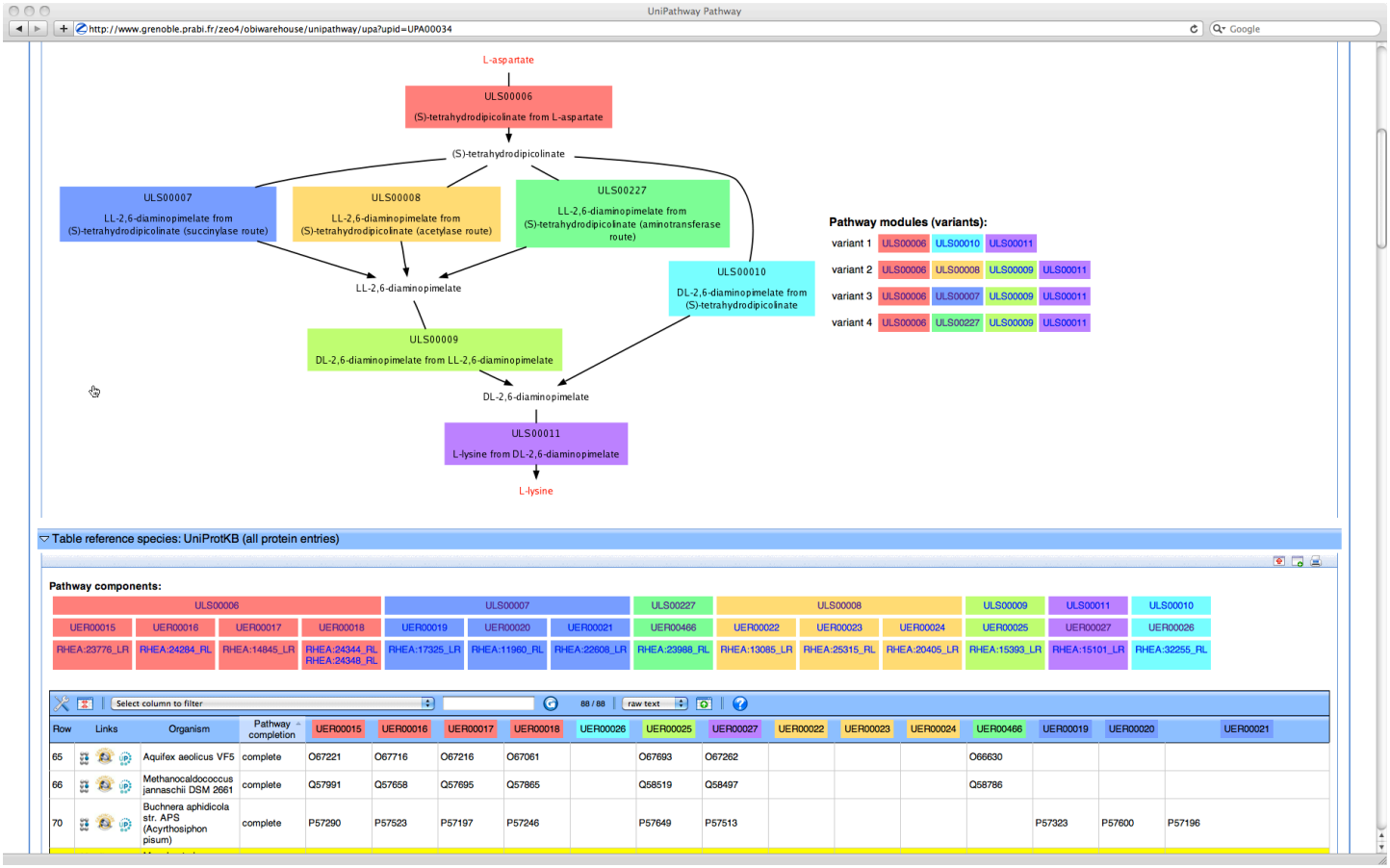
The organism uses an alternative route that has not yet been discovered. In such cases, we planned to use the tools developed by WP5 to identify putative new reactions. If a new pathway was discovered, it should be curated in the Microme/Reactome environment.

5) Data exploration

The UniPathway platform (http://www.grenoble.prabi.fr/zeo4/obiwarehouse/unipathway) was used to integrate the results of the reaction projection strategies and to explore the metabolic data in several contexts.

Some specific views have been developed in The UniPathway platform to facilitate the exploration of metabolic data.

http://www.grenoble.prabi.fr/zeo4/obiwarehouse/unipathway/upa?upid=UPA00034



6) Data curation

The MicroScope platform (https://www.genoscope.cns.fr/agc/microscope) was used to perform gene-reaction association curation. MicroScope accounts have been created for each participant.

Given the limited period of this jamboree, we focused on species with a limited number of missing reactions.

To define curation priorities the results of the reaction predictions have been compiled in an Excel file where organisms were sorted according to the number of missing reactions. Each group of curators were assigned 2 to 3 species to annotate.

		ULS00006				ULS0007			ULS00227	ULS0008			ULS00009	ULS00011	ULS00010	VARIANT			CURATION
		EC 2.7.2.4	EC 1.2.1.11	EC 4.2.1.52	EC 1.3.1.26	EC 2.3.1.117	EC 2.6.1.17	EC 3.5.1.18	EC 2.6.1.83	EC 2.3.1.89	EC 2.6.1.-	EC 3.5.1.47	EC 5.1.1.7	EC 4.1.1.20	EC 1.4.1.16	MetaCyc	KEGG	UniPathway	
Organisme	Lineage	UER00015	UER00016	UER00017	UER00018	UER00019	UER00020	UER00021	UER00466	UER00022	UER00023	UER00024	UER00025	UER00027	UER00026				
Bradyrhizobium japonicum USDA 110	Bacteria; **Proteobacteria**; **Alphaproteobacteria**; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium	bll0216	bll0501,blr4687	blr3302,blr3307,blr3884,bll5072,blr6784,bll7272,blr7382,bll7969	blr0685	blr8104	0	blr8106	blr4361,blr4296	0	0	blr5602,blr3021	bll0477	blr1383	0	Lysine biosynthesis I or Lysine biosynthesis VI	map00300	DAP3? Or DAP2	1 gap DAP3 - check candidats pour DAP2
Campylobacter jejuni subsp. jejuni NCTC 11168	Bacteria; **Proteobacteria**; **Epsilonproteobacteria**; Campylobacterales; Campylobacteraceae; Campylobacter	Cj0582	Cj1023c	Cj0806,Cj0481	Cj0197c	Cj1605c	0	Cj1048c	0	0	0	0	Cj1531	Cj0314	0	Lysine biosynthesis I	map00300	DAP3?	1 gap
Coxiella burnetii	Bacteria; **Proteobacteria**; **Gammaproteobacteria**; Legionellales; Coxiellaceae; Coxiella	CBU_1051	CBU_0875	DAPA_COXBU	DAPB_COXBU	DAPD_COXBU	CBU_0517	DAPE_COXBU	0	0	0	0	DAPF_COXBU	0	0	Lysine biosynthesis I	map00300	DAP3?	1 gap
Haemophilus influenzae Rd KW20	Bacteria; **Proteobacteria**; **Gammaproteobacteria**; Pasteurellales; Pasteurellaceae; Haemophilus	HI0089	HI0646	HI0255	HI1308	HI1634	0	HI0102	0	0	0	0	HI0750	HI0727	0	Lysine biosynthesis I	map00300	DAP3?	1 gap
Rhodobacter sphaeroides 2.4.1	Bacteria; **Proteobacteria**; **Alphaproteobacteria**; Rhodobacterales; Rhodobacteraceae; Rhodobacter	RSP_1849	RSP_1376	RSP_4002,RSP_3406,RSP_3456,RSP_0882	RSP_1105	RSP_1131	0	RSP_1128	0	0	0	0	RSP_0936	RSP_0729	0	Lysine biosynthesis I	map00300	DAP3?	1 gap
Rhodospirillum rubrum ATCC 11170	Bacteria; **Proteobacteria**; **Alphaproteobacteria**; Rhodospirillales; Rhodospirillaceae; Rhodospirillum	Rru_A0743	Rru_A1196	Rru_A1865,Rru_A2066,Rru_A3342	Rru_A0154	Rru_A3479	0	Rru_A3480	0	0	0	0	Rru_A1183	Rru_A0396,Rru_A3135	0	Lysine biosynthesis I	map00300	DAP3?	1 gap
Shewanella oneidensis MR-1	Bacteria; **Proteobacteria**; **Gammaproteobacteria**; Alteromonadales; Shewanellaceae; Shewanella	SO3415,SO4027,SO3986,SO4055	0	SO1879	SO1140	SO1625	SO0617	SO2471	0	0	0	0	SO4308	SO4309	0	Lysine biosynthesis I	map00300	DAP3?	1 gap
Vibrio cholerae N16961	Bacteria; **Proteobacteria**; **Gammaproteobacteria**; Vibrionales; Vibrionaceae; Vibrio	VCA0822,VC0391,VC0547,VC2364,VC2684	VC2107,VC2036	VC2157	VC2391	VC2329	0	VC2152	0	0	0	0	VC0126	VC0125	0	Lysine biosynthesis I	map00300	DAP3?	1 gap
Salinispora arenicola CNS-205	Bacteria; **Actinobacteria**; **Actinobacteria** (class); Actinobacteridae; Actinomycetales; Micromonosporineae; Micromonosporaceae; Salinispora	Sare_0265	Sare_0266,Sare_3946	Sare_1347	Sare_1341	0	Sare_4140	Sare_4134	Sare_2784,Sare_4140	0	0	0	Sare_1404	Sare_0486,Sare_3058,Sare_4034	0	Lysine biosynthesis I? or Lysine biosynthesis VI	map00300	DAP3? Or DAP2	1 gap
Bacillus cereus ATCC 14579	Bacteria; **Firmicutes**; **Bacillales**; Bacillaceae; Bacillus; Bacillus cereus group	BC1748,BC3798	BC2363,BC3799	BC2833,BC3797	BC1532	0	BC4127	BC2978	0	BC3981	0	BC3980	BC4936	BC1419	0	Lysine biosynthesis II or Lysine biosynthesis I	map00300	DAP4? OU DAP3?	1 gap
Enterococcus faecalis	Bacteria; **Firmicutes**; **Lactobacillales**; Enterococcaceae; Enterococcus	EF0368	EF1183	DAPA_ENTFA	DAPB_ENTFA	EF1133	0	EF3178,EF2578,EF1157	0	DAPH_ENTFA	0	DAPEL_ENTFA	DAPF_ENTFA	Q834X3_ENTFA	0	biosynthesis II or Lysine	map00300	DAP4 or DAP3	1 gap
Lactobacillus plantarum WCFS	Bacteria; **Firmicutes**; **Lactobacillales**; Lactobacillaceae; Lactobacillus	lp_2306,lp_0979	lp_1346,lp_2570	lp_2123,lp_2685	lp_1874	0	0	lp_1923,lp_2855	0	lp_2264	0	lp_2263	lp_2185	lp_1713	0	Lysine biosynthesis II	map00300	DAP4?	1 gap
Synechocystis sp. PCC 6803	Bacteria; **Cyanobacteria**; Chroococcales; Synechocystis	slr0657	0	slr0550	sll1058	0	sll0936	0	sll0480	0	0	0	slr1665	sll0504	0	Lysine biosynthesis VI	map00300	DAP2?	1 gap
Methanosarcina acetivorans C2A	**Archaea**; **Euryarchaeota**; Methanomicrobia; Methanosarcinales; Methanosarcinaceae; Methanosarcina	MA0131	MA0430	MA4473	MA4474	0	0	0	MA1712	0	0	0	0	MA0726	0	Lysine biosynthesis VI	map00300	DAP2?	1 gap
Acidobacteria bacterium Ellin345	Bacteria; **Fibrobacteres/Acidobacteria group**; **Acidobacteria**; Acidobacteria (class); Acidobacteriales; Acidobacteriaceae; unclassified Acidobacteriaceae	Acid345_1482,Acid345_2491	Acid345_2358,Acid345_2490	Acid345_2493,Acid345_2666	Acid345_2492	Acid345_2087	0	Acid345_1040	0	0	0	0	Acid345_2622	Acid345_3440	0	Lysine biosynthesis I	map00300	DAP3?	1 gap
Chlamydia trachomatis D/UW-3/CX	Bacteria; **Chlamydiae**; **Chlamydiales**; Chlamydiaceae; Chlamydia/Chlamydophila group; Chlamydia	CT362	CT363	CT361	CT364	0	0	0	CT390	0	0	0	CT430	0	0	Lysine biosynthesis VI	map00300	DAP2?	1 gap
Thermotoga maritima MSB8	Bacteria; **Thermotogae**; **Thermotogales**; Thermotogaceae; Thermotoga	TM1518	TM1523	TM1521	TM1520	0	TM1785	TM1666	0	TM1519	0	TM1516	TM1522	TM1517	0	Lysine biosynthesis II or Lysine biosynthesis I	map00300	DAP4? OU DAP3?	1 gap
Salinibacter ruber DSM 13855	Bacteria; **Bacteroidetes**; **Order II. Incertae sedis**; Rhodothermaceae; Salinibacter	SRU_1745,SRU_1746	SRU_1838	SRU_1435,SRU_1436	SRU_1081	SRU_0969	0	SRU_2220	0	0	0	0	SRU_1415	SRU_0301	0	Lysine biosynthesis I	map00300	DAP3?	1 gap

Each participant was free to use the web tools that he or she was most familiar with

The MicroScope implements several functionalities that guide the curation process, such as synteny maps, the phyloprofile tool (which performs searches searches for co-evolved genes), and the CanOE strategy (see MicroScope online tutorial, https://www.genoscope.cns.fr/agc/website/spip.php?rubrique189).

All the curation work carried out during the jamboree has been stored into the Microscope platform. This includes:

- the association of genes to Rhea and MetaCyc reactions
- the update of gene annotations (including the identity of the product, EC number, bibliographical references, etc)

To help each participant in the annotation task, a second Excel document was prepared that describes the data that has to be added in each field of the gene editor (reactions, EC numbers, etc).

Genomic Object Editor: SRU_2217
Salinibacter ruber DSM 13855 **DSM 13855; M31 - chromosome SRU_ NC_007677**

5'/3' | TrEMBL alignments | SwissProt alignments | PhyloProfile | PubMed | KEGG | BRENDA | MicroCyc

» CURRENT ANNOTATION — MaGe curated annotation — Status: finished — Annotator: deustp01

Type	Begin	End	Length	Frame	Mutation	Gene	Synonyms	Date	Status
CDS	2749223	2750410	1188 (395aa)	+2	no	dapC	argD, lysJ	2012-04-17 12:04:42	inProgress

Note	identified by match to protein family HMM PF00202
Product	succinyldiaminopimelate aminotransferase
Product Type	e : enzyme
EC number	2.6.1.17
MetaCyc Reaction	ACETYLORNTRANSAM-RXN: acetylornithine transaminase
Rhea Reaction	RHEA:11960: 2-oxoglutarate + N-succinyl-LL-2,6-diaminopimelate L-2-succinylamino-6-oxopimelate + L-glutamate
Localization	2 : Cytoplasmic
BioProcess	1.2 : Aspartate family
Roles	1.5.1.7 : Lysine, diaminopimelate
PubMedId	10074354, 20418392
Comments	The E. coli homologue of this protein appears to have catalytic activity for (at least) two substrates, rather than to contain two separate catalytic domains (PMID:10074354). Disruption of the homologous Methanococcus maripaludis open reading frame yields a strain incapable of lysine biosynthesis that lacks diaminopimelate aminotransferase activity in vitro (PMID:20418392).
Class	2a : Function of homologous gene experimentally demonstrated in an other organism

CANCEL — SAVE

Curation of the gene coding for the succinyldiaminopimelate aminotranferase in Salinibacter ruber

RESULTS (if applicable, interactions with other workpackages)

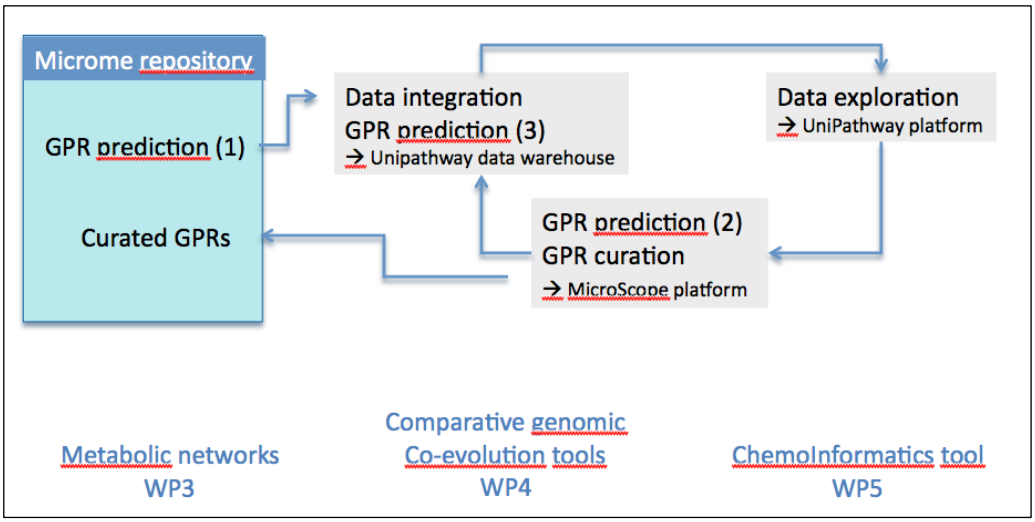
At the end of the jamboree, about 40 gene-reaction associations were curated (in black, curation of genes predicted by at least one method, in red curation of gene candidate found to fill in gaps).

		ULS00006				ULS0007			ULS00227	ULS0008			ULS00009	ULS00011	ULS00010	VARIANT		
		EC 2.7.2.4	EC 1.2.1.11	EC 4.2.1.52	EC 1.3.1.26	EC 2.3.1.117	EC 2.6.1.17	EC 3.5.1.18	EC 2.6.1.83	EC 2.3.1.89	EC 2.6.1.-	EC 3.5.1.47	EC 5.1.1.7	EC 4.1.1.20	EC 1.4.1.16	MetaCyc	KEGG	UniPathway
Organisme	Lineage	UER00015	UER00016	UER00017	UER00018	UER00019	UER00020	UER00021	UER00466	UER00022	UER00023	UER00024	UER00025	UER00027	UER00026			
Bradyrhizobium japonicum USDA 110	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium	blr0216	bll0501,blr4687	blr3302,blr3307,blr3884,bll5072,blr6784,blr7272,blr7382,blr7969	blr0685	blr6104	blr4134	blr6106	blr4361,blr4296	0	0	blr5602,blr3021	bll0477	blr1383	0	Lysine biosynthesis I or Lysine biosynthesis VI	map00300	DAP3 Or DAP2?
Campylobacter jejuni subsp. jejuni NCTC 11168	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Campylobacteraceae; Campylobacter	Cj0582	Cj1023c	Cj0806,Cj0481	Cj0197c	Cj1605c	Cj0227	Cj1048c	0	0	0	0	Cj1531	Cj0314	0	Lysine biosynthesis I	map00300	DAP3
Coxiella burnetii	Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales; Coxiellaceae; Coxiella	CBU_1051	CBU_0875	DAPA_COXBU	DAPB_COXBU	DAPD_COXBU	CBU_0517	DAPE_COXBU	0	0	0	0	DAPF_COXBU	0	0	Lysine biosynthesis I	map00300	DAP3?
Haemophilus influenzae Rd KW20	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales; Pasteurellaceae; Haemophilus	HI0089	HI0646	HI0255	HI1308	HI1634	0	HI0102	0	0	0	0	HI0750	HI0727	0	Lysine biosynthesis I	map00300	DAP3?
Rhodobacter sphaeroides 2.4.1	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Rhodobacter	RSP_1849	RSP_1376	RSP_4002,RSP_3406,RSP_0882	RSP_1105	RSP_1131	RSP_2008	RSP_1128	0	0	0	0	RSP_0936	RSP_0729	0	Lysine biosynthesis I	map00300	DAP3
Rhodospirillum rubrum ATCC 11170	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales; Rhodospirillaceae; Rhodospirillum	Rru_A0743	Rru_A1196	Rru_A1865,Rru_A2066,Rru_A3342	Rru_A0154	Rru_A3479	0	Rru_A3480	Rru_A2411	0	0	0	Rru_A1183	Rru_A0396,Rru_A3135	0	Lysine biosynthesis I	map00300	DAP3? DAP2
Shewanella oneidensis MR-1	Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Shewanellaceae; Shewanella	SO3415,SO3427,SO3986,SO4055	SO3070	SO1879	SO1140	SO1625	SO0617	SO2471	0	0	0	0	SO4308	SO4309	0	Lysine biosynthesis I	map00300	DAP3
Vibrio cholerae N16961	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio	VCA0822,VC0391,VC0547,VC2364,VC2684	VC2107,VC2036	VC2157	VC2391	VC2329	VC2618	VC2152	0	0	0	0	VC0126	VC0125	0	Lysine biosynthesis I	map00300	DAP3
Salinispora arenicola CNS-205	Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Micromonosporineae; Micromonosporaceae; Salinispora	Sare_0265	Sare_0266,Sare_3946	Sare_1347	Sare_1341	0	Sare_4140	Sare_4134	Sare_2784,Sare_4140	0	0	0	Sare_1404	Sare_0486,Sare_3058,Sare_4034	0	Lysine biosynthesis I? or Lysine biosynthesis VI	map00300	DAP3? Or DAP2
Bacillus cereus ATCC 14579	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group	BC1746,BC3799	BC2363,BC3799	BC2833,BC3797	BC1532	0	BC4127	BC2978	0	BC3981	0	BC3980	BC4936	BC1419	0	Lysine biosynthesis II or Lysine biosynthesis I	map00300	DAP4? OU DAP3?
Enterococcus faecalis	Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus	EF0368	EF1183	DAPA_ENTFA; EF1183	DAPB_ENTFA	EF1133	0	EF3178,EF2578,EF1157	0	DAPH_ENTFA; EF1133	EF1706	DAPEL_ENTFA	DAPF_ENTFA; EF0464	Q834X3_ENTFA; EF1504	0	Lysine biosynthesis II or Lysine biosynthesis I	map00300	DAP4
Lactobacillus plantarum WCFS	Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae; Lactobacillus	lp_2308,lp_0979	lp_1346,lp_2570	lp_2123,lp_2685	lp_1874	0	0	lp_1923,lp_2855	0	lp_2264	Lp_1280,lp_2684	lp_2263	lp_2185	lp_1713	0	Lysine biosynthesis II	map00300	DAP4
Synechocystis sp. PCC 6803	Bacteria; Cyanobacteria; Chroococcales; Synechocystis	slr0657	0	slr0550	sll1058	0	sll0938	0	sll0480	0	0	0	slr1665	sll0504	0	Lysine biosynthesis VI	map00300	DAP2?
Methanosarcina acetivorans C2A	Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales; Methanosarcinaceae; Methanosarcina	MA0131	MA0430	MA4473	MA4474	0	0	0	MA1712	0	0	0	0	MA0726	0	Lysine biosynthesis VI	map00300	DAP2
Acidobacteria bacterium Ellin345	Bacteria; Fibrobacteres/Acidobacteria group; Acidobacteria; Acidobacteria (class); Acidobacteriales; Acidobacteriaceae; unclassified Acidobacteriaceae	Acid345_1482,Acid345_2491	Acid345_2358,Acid345_2490	Acid345_2493,Acid345_2666	Acid345_2492	Acid345_2087	0	Acid345_1040	0	0	0	0	Acid345_2622	Acid345_3440	0	Lysine biosynthesis I	map00300	DAP3?
Chlamydia trachomatis D/UW-3/CX	Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia/Chlamydophila group; Chlamydia	CT362	CT363	CT361	CT364	0	0	0	CT390	0	0	0	CT430	0	0	Lysine biosynthesis VI	map00300	DAP2?
Thermotoga maritima MSB8	Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga	TM1518	TM1523	TM1521	TM1520	0	TM1785	TM1666	0	TM1519	TM1255	TM1516	TM1522	TM1517	0	Lysine biosynthesis II or Lysine biosynthesis I	map00300	DAP4
Salinibacter ruber DSM 13855	Bacteria; Bacteroidetes; Bacteroidetes Order II. Incertae sedis; Rhodothermaceae; Salinibacter	SRU_1745,SRU_1838	SRU_1838	SRU_1435,SRU_1081	SRU_1081	SRU_0969	SRU_2217	SRU_2220	0	0	0	0	SRU_1415	SRU_0301	0	Lysine biosynthesis I	map00300	DAP3

Each group presented its results. Minutes of the discussion are stored on the Microme confluence web site.
http://www.ebi.ac.uk/seqdb/confluence/display/Microme/Discussions+Jamboree+2012

Conclusions & Perspectives

The main objective of this jamboree was to refine the Microme curation cycle, with the production of actual curated data a secondary outcome. The jamboree allowed different groups from Microme to become familiar with and to curation strategies, tools and methods developed by the participating Microme partners.



The WP2 will continue the data curation initiated during this jamboree, and the resulting data will be sent to WP4 partners to perform evolution studies.
To complete 'The evolutionary history of lysine biosynthesis pathways within eukaryotes' published by Torruella et al in 2009, we envisage a publication on the 'The evolutionary history of lysine biosynthesis pathways within prokaryotes'.