# Deliverable 5.2

| | |
|---|---|
| Project Title: | Developing an efficient e-infrastructure, standards and data-flow for metabolomics and its interface to biomedical and life science e-infrastructures in Europe and world-wide |
| Project Acronym: | COSMOS |
| Grant agreement no.: | 312941 |
| | Research Infrastructures, FP7 Capacities Specific Programme; [INFRA-2011-2.3.2.] “Implementation of common solutions for a cluster of ESFRI infrastructures in the field of "Life sciences" |
| Deliverable title: | Implemented data-broadcast mechanism |
| WP No. | 5 |
| Lead Beneficiary: | 2. LU |
| WP Title | Dissemination Pipelines |
| Contractual delivery date: | 1 10 2014 |
| Actual delivery date: | 1 10 2014 |
| WP leader: | Thomas Hankemeier 2. LU |
| Contributing partner(s): | 1.EMBL-EBI, |

*Authors: Thomas Hankemeier Reza Salek, Michael van Vliet*

# Contents

# 1 Executive summary

For this deliverable D5.2 we have implemented a broadcast mechanism for MetabolomeXchange to inform the metabolomics community about new or updated data sets. The broadcast mechanism is based on the RSS 2.0 specifications (http://www.rssboard.org/rss-specification).

# 2 Project objectives

With this deliverable, the project has contributed the following objective:

| No. | Objective | Yes | No |
|---|---|---|---|
| 1 | Enable the metabolomics community to be kept up-to-date by implementing a data-broadcast mechanism. | X | |

# 3 Detailed report on the deliverable

## 3.1 Background

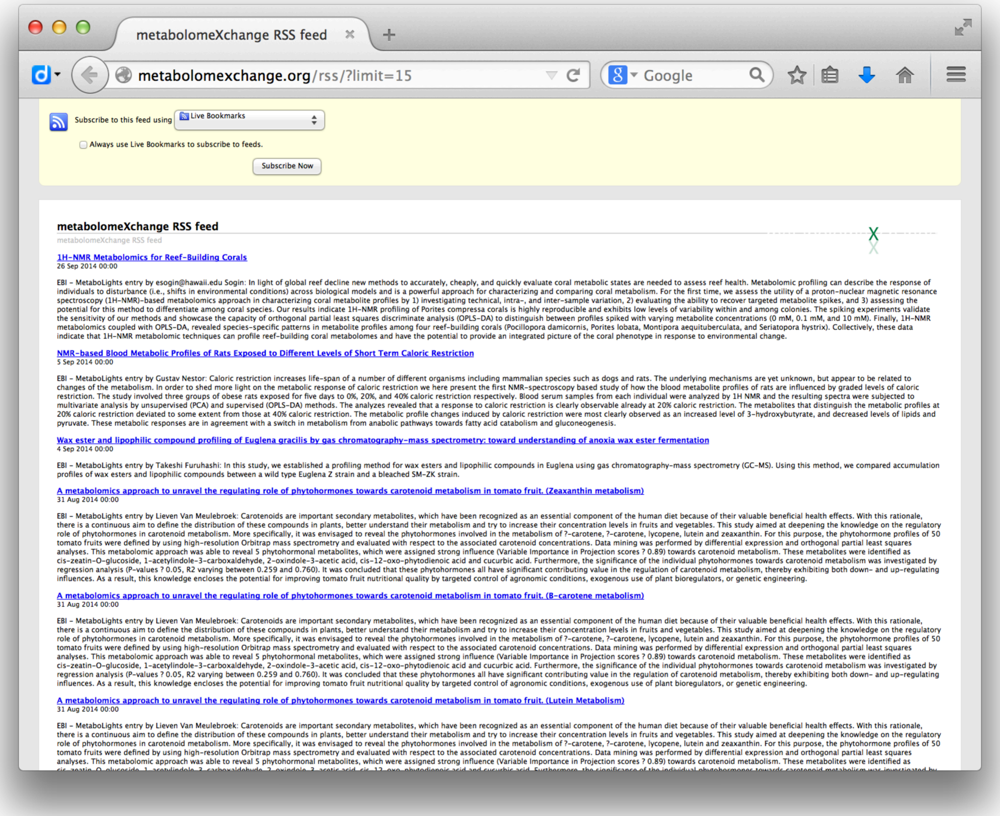
Within WP5.1 we developed a mechanism to aggregate and store meta-data about publicly available metabolomics data sets into a central register. This register can be accessed using a web browser and used to find data sets of interest. This requires actively searching and looking for updates by individual users. The idea behind WP5.2 is to make it even easier for regular visitors of MetabolomeXchange to be notified of changes. The most common and widely used technology on the Internet today to facilitate this is RSS. RSS is a web content syndication format that allows content providers to announce new or updated content.

## 3.2 Description of Work

Implement a broadcast mechanism that allows visitors of MetabolomeXchange to be notified after subscribing to data set updates.

### 3.2.1 RSS feed (server side)

MetabolomeXchange provides access to all or only the latest data sets formatted as RSS 2.0. For each data set the RSS feed contains the link to the page of that data set, title, description and publication date. De description contains the name of the original repository, the name of the submitter and the data set abstract. Additional provider specific meta-data is not included in the feed but only accessible on the MetabolomeXchange site or API.



Subscribe to this feed using Live Bookmarks

☐ Always use Live Bookmarks to subscribe to feeds.

**metabolomeXchange RSS feed**

metabolomeXchange RSS feed

1H-NMR Metabolomics for Reef-Building Corals
26 Sep 2014 00:00

EBI – MetaboLights entry by esogin@hawaii.edu Sogin: In light of global reef decline new methods to accurately, cheaply, and quickly evaluate coral metabolic states are needed to assess reef health. Metabolomic profiling can describe the response of individuals to disturbance (i.e., shifts in environmental conditions) across biological models and is a powerful approach for characterizing and comparing coral metabolism. For the first time, we assess the utility of a proton-nuclear magnetic resonance spectroscopy (1H-NMR)-based metabolomics approach in characterizing coral metabolite profiles by 1) investigating technical, intra-, and inter-sample variation, 2) evaluating the ability to recover targeted metabolite spikes, and 3) assessing the potential for this method to differentiate among coral species. Our results indicate 1H-NMR profiling of Porites compressa corals is highly reproducible and exhibits low levels of variability within and among colonies. The spiking experiments validate the sensitivity of our methods and showcase the capacity of orthogonal partial least squares discriminate analysis (OPLS-DA) to distinguish between profiles spiked with varying metabolite concentrations (0 mM, 0.1 mM, and 10 mM). Finally, 1H-NMR metabolomics coupled with OPLS-DA, revealed species-specific patterns in metabolite profiles among four reef-building corals (Pocillopora damicornis, Porites lobata, Montipora aequituberculata, and Seriatopora hystrix). Collectively, these data indicate that 1H-NMR metabolomic techniques can profile reef-building coral metabolomes and have the potential to provide an integrated picture of the coral phenotype in response to environmental change.

NMR-based Blood Metabolic Profiles of Rats Exposed to Different Levels of Short Term Caloric Restriction
5 Sep 2014 00:00

EBI – MetaboLights entry by Gustav Nestor: Caloric restriction increases life-span of a number of different organisms including mammalian species such as dogs and rats. The underlying mechanisms are yet unknown, but appear to be related to changes of the metabolism. In order to shed more light on the metabolic response of caloric restriction we here present the first NMR-spectroscopy based study of how the blood metabolite profiles of rats are influenced by graded levels of caloric restriction. The study involved three groups of obese rats exposed for five days to 0%, 20%, and 40% caloric restriction respectively. Blood serum samples from each individual were analyzed by 1H NMR and the resulting spectra were subjected to multivariate analysis by unsupervised (PCA) and supervised (OPLS-DA) methods. The analyzes revealed that a response to caloric restriction is clearly observable already at 20% caloric restriction. The metabolites that distinguish the metabolic profiles at 20% caloric restriction deviated to some extent from those at 40% caloric restriction. The metabolic profile changes induced by caloric restriction were most clearly observed as an increased level of 3-hydroxybutyrate, and decreased levels of lipids and pyruvate. These metabolic responses are in agreement with a switch in metabolism from anabolic pathways towards fatty acid catabolism and gluconeogenesis.

Wax ester and lipophilic compound profiling of Euglena gracilis by gas chromatography-mass spectrometry: toward understanding of anoxia wax ester fermentation
4 Sep 2014 00:00

EBI – MetaboLights entry by Takeshi Furuhashi: In this study, we established a profiling method for wax esters and lipophilic compounds in Euglena using gas chromatography-mass spectrometry (GC-MS). Using this method, we compared accumulation profiles of wax esters and lipophilic compounds between a wild type Euglena Z strain and a bleached SM-ZK strain.

A metabolomics approach to unravel the regulating role of phytohormones towards carotenoid metabolism in tomato fruit. (Zeaxanthin metabolism)
31 Aug 2014 00:00

EBI – MetaboLights entry by Lieven Van Meulebroek: Carotenoids are important secondary metabolites, which have been recognized as an essential component of the human diet because of their valuable beneficial health effects. With this rationale, there is a continuous aim to define the distribution of these compounds in plants, better understand their metabolism and try to increase their concentration levels in fruits and vegetables. This study aimed at deepening the knowledge on the regulatory role of phytohormones in carotenoid metabolism. More specifically, it was envisaged to reveal the phytohormones involved in the metabolism of ?-carotene, ?-carotene, lycopene, lutein and zeaxanthin. For this purpose, the phytohormone profiles of 50 tomato fruits were defined by using high-resolution Orbitrap mass spectrometry and evaluated with respect to the associated carotenoid concentrations. Data mining was performed by differential expression and orthogonal partial least squares analyses. This metabolomic approach was able to reveal 5 phytohormonal metabolites, which were assigned strong influence (Variable Importance in Projection scores ? 0.89) towards carotenoid metabolism. These metabolites were identified as cis-zeatin-O-glucoside, 1-acetylindole-3-carboxaldehyde, 2-oxindole-3-acetic acid, cis-12-oxo-phytodienoic acid and cucurbic acid. Furthermore, the significance of the individual phytohormones towards carotenoid metabolism was investigated by regression analysis (P-values ? 0.05, R2 varying between 0.259 and 0.760). It was concluded that these phytohormones all have significant contributing value in the regulation of carotenoid metabolism, thereby exhibiting both down- and up-regulating influences. As a result, this knowledge encloses the potential for improving tomato fruit nutritional quality by targeted control of agronomic conditions, exogenous use of plant bioregulators, or genetic engineering.

A metabolomics approach to unravel the regulating role of phytohormones towards carotenoid metabolism in tomato fruit. (B-carotene metabolism)
31 Aug 2014 00:00

EBI – MetaboLights entry by Lieven Van Meulebroek: Carotenoids are important secondary metabolites, which have been recognized as an essential component of the human diet because of their valuable beneficial health effects. With this rationale, there is a continuous aim to define the distribution of these compounds in plants, better understand their metabolism and try to increase their concentration levels in fruits and vegetables. This study aimed at deepening the knowledge on the regulatory role of phytohormones in carotenoid metabolism. More specifically, it was envisaged to reveal the phytohormones involved in the metabolism of ?-carotene, ?-carotene, lycopene, lutein and zeaxanthin. For this purpose, the phytohormone profiles of 50 tomato fruits were defined by using high-resolution Orbitrap mass spectrometry and evaluated with respect to the associated carotenoid concentrations. Data mining was performed by differential expression and orthogonal partial least squares analyses. This metabolomic approach was able to reveal 5 phytohormonal metabolites, which were assigned strong influence (Variable Importance in Projection scores ? 0.89) towards carotenoid metabolism. These metabolites were identified as cis-zeatin-O-glucoside, 1-acetylindole-3-carboxaldehyde, 2-oxindole-3-acetic acid, cis-12-oxo-phytodienoic acid and cucurbic acid. Furthermore, the significance of the individual phytohormones towards carotenoid metabolism was investigated by regression analysis (P-values ? 0.05, R2 varying between 0.259 and 0.760). It was concluded that these phytohormones all have significant contributing value in the regulation of carotenoid metabolism, thereby exhibiting both down- and up-regulating influences. As a result, this knowledge encloses the potential for improving tomato fruit nutritional quality by targeted control of agronomic conditions, exogenous use of plant bioregulators, or genetic engineering.

A metabolomics approach to unravel the regulating role of phytohormones towards carotenoid metabolism in tomato fruit. (Lutein Metabolism)
31 Aug 2014 00:00

EBI – MetaboLights entry by Lieven Van Meulebroek: Carotenoids are important secondary metabolites, which have been recognized as an essential component of the human diet because of their valuable beneficial health effects. With this rationale, there is a continuous aim to define the distribution of these compounds in plants, better understand their metabolism and try to increase their concentration levels in fruits and vegetables. This study aimed at deepening the knowledge on the regulatory role of phytohormones in carotenoid metabolism. More specifically, it was envisaged to reveal the phytohormones involved in the metabolism of ?-carotene, ?-carotene, lycopene, lutein and zeaxanthin. For this purpose, the phytohormone profiles of 50 tomato fruits were defined by using high-resolution Orbitrap mass spectrometry and evaluated with respect to the associated carotenoid concentrations. Data mining was performed by differential expression and orthogonal partial least squares analyses. This metabolomic approach was able to reveal 5 phytohormonal metabolites, which were assigned strong influence (Variable Importance in Projection scores ? 0.89) towards carotenoid metabolism. These metabolites were identified as

Figure A: RSS feed showing the latest 15 data sets

### 3.2.2 RSS feed (client side)

One of the biggest advantages about using the RSS web content syndication format is the huge collection of compatible clients. Most popular mail clients can handle RSS and present it as if it is an email. Others prefer an online-hosted service like Feedly to read RSS feeds. The RSS feed has been thoroughly tested with several RSS clients and based on feedback from users it works as expected.

### 3.2.3 Access and documentation

The MetabolomeXchange broadcast mechanism is available and accessible at http://metabolomexchange.org/rss. All source files are available on the project Github pages, together with accompanying readme files and license (Apache License, Version 2.0):

**GitHub** (application): https://github.com/leidenuniv-lacdr-abs/metabolomexchange

## 3.3 Next steps

As part of WP5.3, a "Tool that allows checking predefined information in the broadcast", we will provide to our users a more fine-grained filtering and search option to access the RSS feed. Users will be able to subscribe to predefined information using key words to filter out only the data sets of interest.

# 4 Publications

None.

# 5 Delivery and schedule

The delivery is delayed: ☐Yes ☑No

# 6 Adjustments made

None.

# 7 Efforts for this deliverable

| **Institute** | **Person-months (PM)** | |
|---|---|---|
| | **actual** | **estimated** |
| 2: UL | 6 | |
| 1: EMBL-EBI | 1 | |
| 8:MPG | 1 | |
| 11:IPB | 0.5 | |
| 13:UB2 | 0.5 | |
| 6:VTT | 0.19 | |
| **Total** | **9.19** | **12** |

# Appendices

1. N/A

# Background information

This deliverable relates to WP5; background information on this WP as originally indicated in the description of work (DoW) is included below.

**WP5** **Title: Dissemination Pipelines**
Lead: Thomas Hankemeier, UL
Participants: EBI-EMBL, LU-NMC, MRC, VTT, UB, MPG, IPB, UB2 and UBHam,

This work package will focus on developing and coordinating the infrastructure to easily access, to process, store, and exchange metabolomics measurement and associated experimental metadata.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Work package number** | WP5 | **Start date or starting event:** | | | | | Month 1 | | |
| **Work package title** | Dissemination Piplines | | | | | | | | |
| **Activity Type** | COORD | | | | | | | | |
| **Participant number** | 1: EMBL-EBI | 2: LU/NC | 3:MRC | 6:VTT | 7:UB | 8:MPG | 11:IPB | 12:UB2 | 13:UBHAM |
| **Person-months per participant** | 7 | 15 | 2 | 2 | 3 | 2 | 1 | 1 | 2 |

**Objectives**

This work package will develop the mechanisms for disseminating the data submitted to all COSMOS partners to the other participating Metabolomics resources in the consortium, and the community at large. The desired setup will enable users to submit their data and metadata to any of the participating resources, whereupon it will be made available automatically to all other repositories or participants who wish to access the data, providing different, added value views of the data. Efficient user notification of new datasets and access to metadata will be provided through RSS notifications, and a central archive of such notifications. Reprocessed views of the data will also be announced and registered through this mechanism.

**Description of work and role of participants**

Task 1: Dissemination pipeline Once metabolomics data acquired by one of the COSMOS partners has been approved for public release (e.g. after assuring a certain quality level or after statistical analysis or publication), specific metadata will be automatically sent to all interested parties (all COSMOS partners and anyone interested in the metabolomics community) through RSS notifications. Checking the content of the metadata allows the receiver to decide if the dataset will be downloaded. The RSS feed does contain information (e.g. an URL) how to access the metabolomics data, possibly after checking authentication and authorization. The use-cases for this mechanism are manifold and of high interest to our user communities. One case would be experimentally derived standards. If a party is interested in a particular class of compounds, say eicosanoids, it will be alerted whenever a new structure was submitted so an update of their local database can be triggered. Secondly, based on a grouping of metabolites according to tissue type, researchers interested in, for example, adipose tissue will be alerted whenever a new metabolite in adipose tissue is found. Finally this will have obvious benefits for any large-scale model organism studies - e.g. yeast, C. elegans, flies etc.

Task 2: Development of MetaboStore, a metadata archive for Metabolomics, serving as an intermediate general-purpose component to feed into the stakeholder repositories. In a later stage an RSS receiving party will be able to specify up front what kind of data is of their interest. A tool will be developed that will alert the interested party only after finding certain predefined information after processing the metadata. The same tool can be used to query over all COSMOS studies ever released to the public by searching the MetaboStore, a metadata archive for metabolomics data. Such a federated query could, e.g., together with semantic queries, relieve individual Databases from managing SOAP/REST/custom query interfaces. Standardized metadata together with WP3 allows querying over studies on sample level, metabolite level (identities), on quantitative level (content of the dataset, reference data), on statistical data analysis result level or certain combinations of these levels. TNO will give input on the development of biological relevant queries and will develop essential ontologies, to facilitate data exchange. With the standards defined in WP 2 and 4 this will actually be a phenotype database on metabolism, and will be embedded in large e-infrastructures such as ELIXIR and BioMedBridges to allow the data integration and interoperability with important European initiatives. The data warehouse within the LU/NMC-DSP, developed together with NuGO, consists of the generic study capturing framework (GSCF), a simple assay module (for clinical chemistry data) and a metabolite centric module, and is a candidate repository to store the relevant study (meta) data.

The user acceptance will be monitored through usage and download statistics provided by the source code management site of our choice (SourceForge/Google Code).

In addition we will perform surveys as part of the last two annual stakeholder meetings.

**Deliverables**

| No. | Name | Due month |
|---|---|---|
| D 5.1 | Tool that enables uploading of specific metadata to the MetaboStore | 24 |
| D5.2 | Implemented data-broadcast mechanism | 24 |
| D5.3 | Tool that allows checking predefined information in broadcast | 30 |
| D5.4 | Tool that allows querying MetaboStore | 30 |
| D5.5 | Usage statistic and downloads report | 36 |