The open source ISA software suite and its international user community:

Knowledge management of experimental data

Alejandra González-Beltrán

Senior Software Engineer, ISATeam

Oxford e-Research Centre, University of Oxford

Oxford, UK

NETTAB 2012 – Integrated Bio-Search, Como, Italy, November 14-16

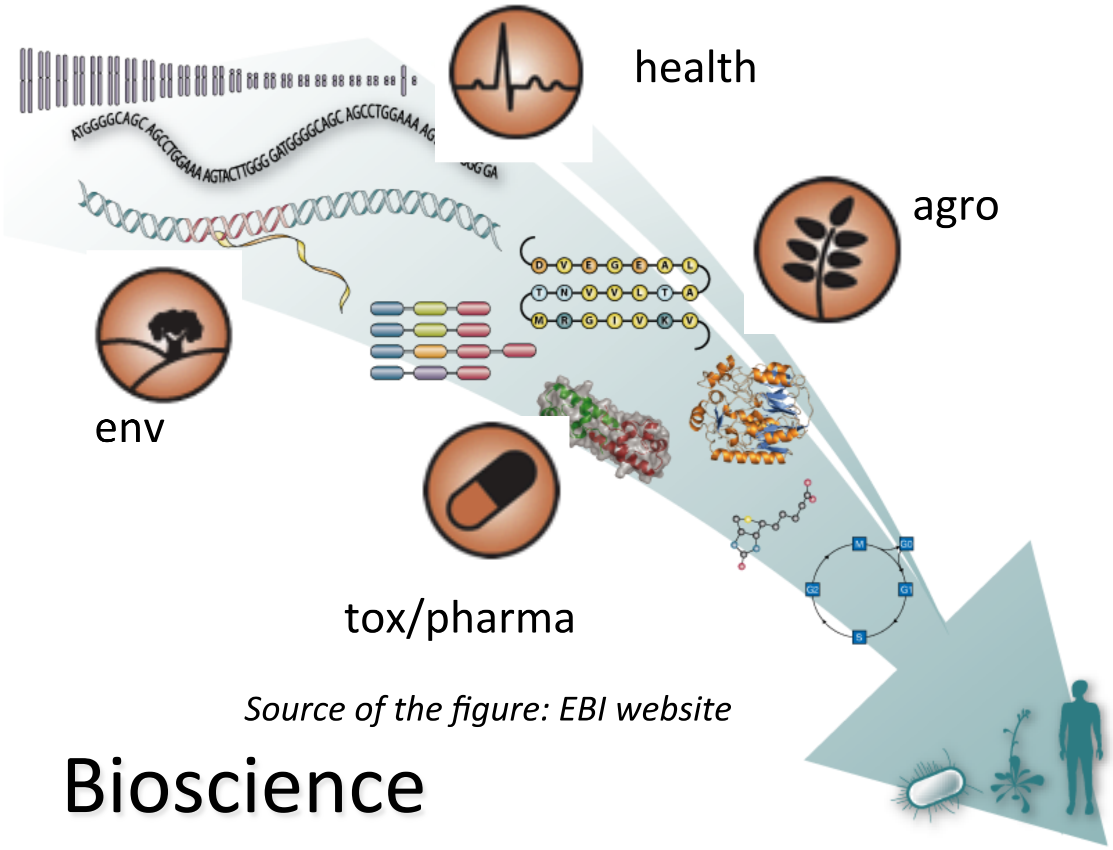
Outline

- Knowledge management of experimental data
  - Setting the scene
  - The isa ecosystem: ISA-tab, tools, community
  - Use case
- Latest additions

OntoMaton Risa isatoowl

- Related projects & main points

Setting the scene



health

agro

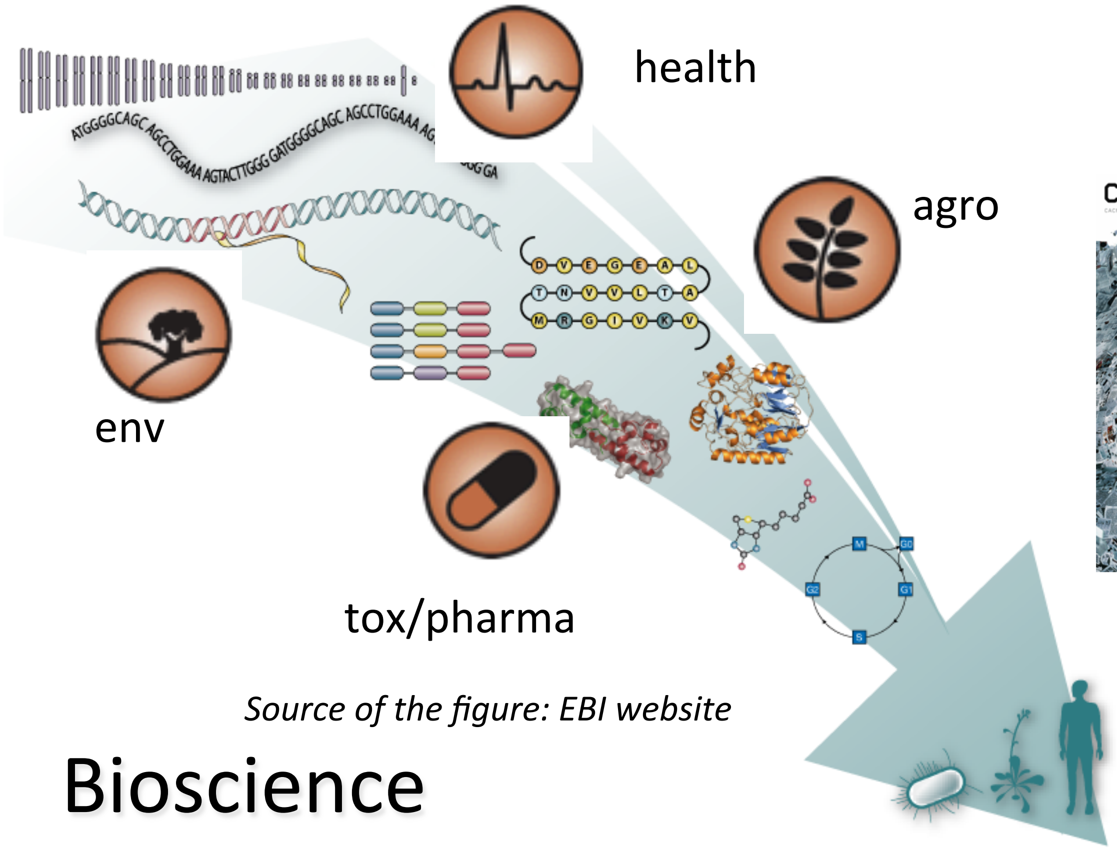
env

tox/pharma

Source of the figure: EBI website

Bioscience
is multi-domain…

Setting the scene

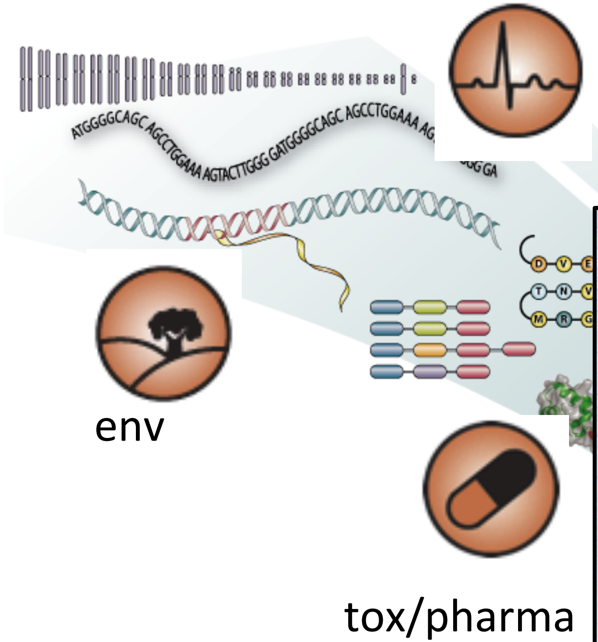


Source of the figure: EBI website

Bioscience
is multi-domain…

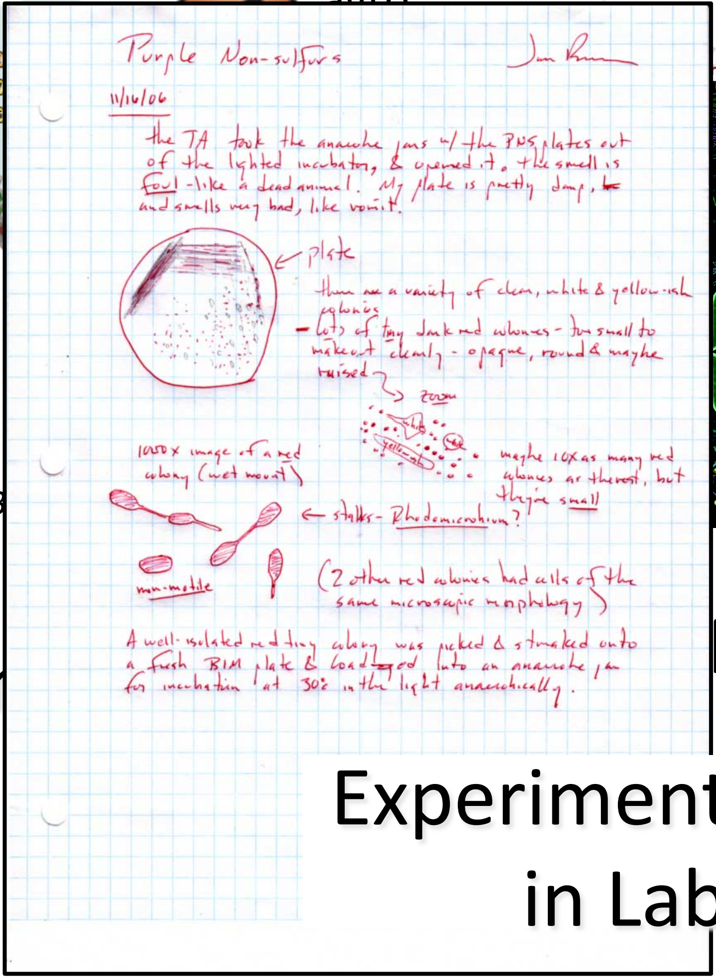
Petabytes of data

Setting the scene



Source of the figure: EB

Bioscience
is multi-domain



Purple Non-sulfurs

11/16/06

the TA took the anaerobe jars w/ the PNS plates out of the lighted incubator, & opened it. the smell is foul - like a dead animal. My plate is pretty damp, and smells very bad, like vomit.

← plate

there are a variety of clear, white & yellow-ish colonies

- lots of tiny dark red colonies - too small to make out clearly - opaque, round & maybe raised

zoom

maybe 10x as many red colonies as the rest, but they're small

1050x image of a red colony (wet mount)

← stalks - Rhodomicrobium?

non-motile

(2 other red colonies had cells of the same microscopic morphology)

A well-isolated red tiny colony was picked & streaked onto a fresh BIM plate & loaded into an anaerobe jar for incubation at 30°c in the light anaerobically.

Petabytes of data

Experimental metadata
in Lab books

isa

investigation study assay

- Assist in the annotation and management of experimental data at source
- Deal with data from high-throughput studies using one or a combination of *omics* and other technologies
- Empower users to uptake community-defined checklists and ontologies
- Facilitate data sharing, reuse, comparison and reproducibility of experiments, submission to international public repositories

The isa ecosystem

isatab

isatools
isa-tools.org

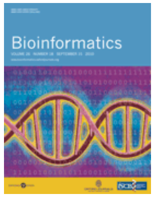
isacommons
isacommons.org

The isa ecosystem

isatab

isatools
isa-tools.org

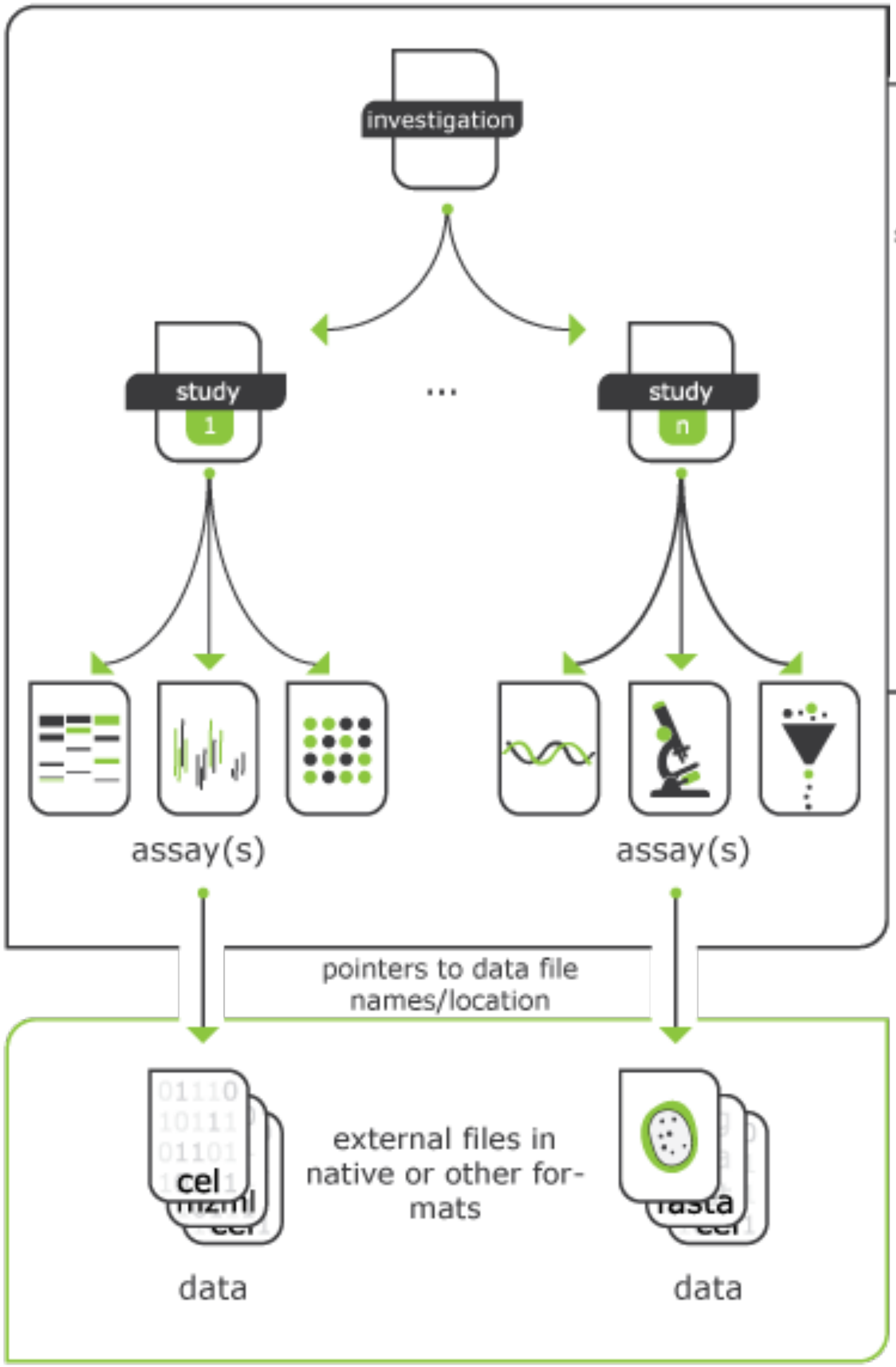
isacommons
isacommons.org



ISA software suite: supporting standards-compliant experimental annotation and enabling curation at the community level
Rocca-Serra et al, 2010
Bioinformatics

Towards interoperable bioscience data
Sansone et al, 2012
Nature Genetics

isatab



investigation
high level concept to link related studies

study
the central unit, containing information on the subject under study, its characteristics and any treatments applied.

a study has associated ***assays***

assay
test performed either on material taken from the subject or on the whole initial subject, which produce qualitative or quantitative measurements (data)

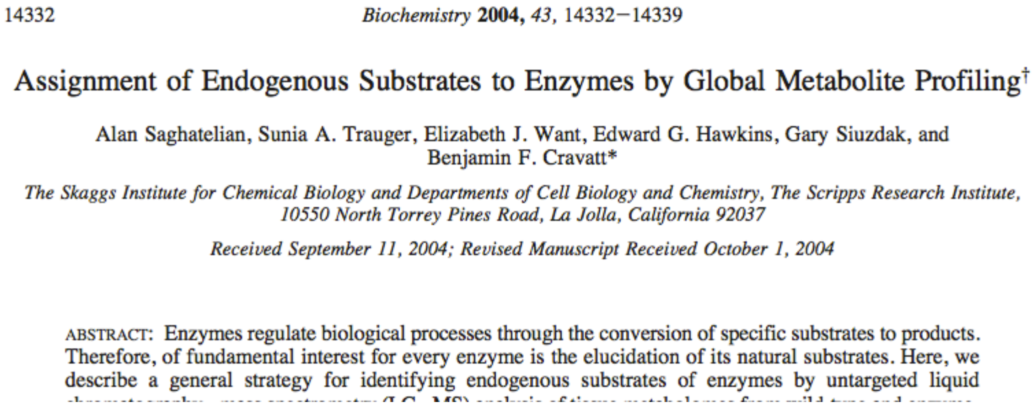
General purpose & flexible format

Domain agnostic

Captures metadata in omics experiments and traditional experiments (e.g. clinical chemistry and histology)

faahKO dataset

- Available in BioConductor
- Subset of the original data on global metabolite profiling



14332 *Biochemistry* **2004,** *43,* 14332−14339

Assignment of Endogenous Substrates to Enzymes by Global Metabolite Profiling[†]

Alan Saghatelian, Sunia A. Trauger, Elizabeth J. Want, Edward G. Hawkins, Gary Siuzdak, and Benjamin F. Cravatt*

The Skaggs Institute for Chemical Biology and Departments of Cell Biology and Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received September 11, 2004; Revised Manuscript Received October 1, 2004

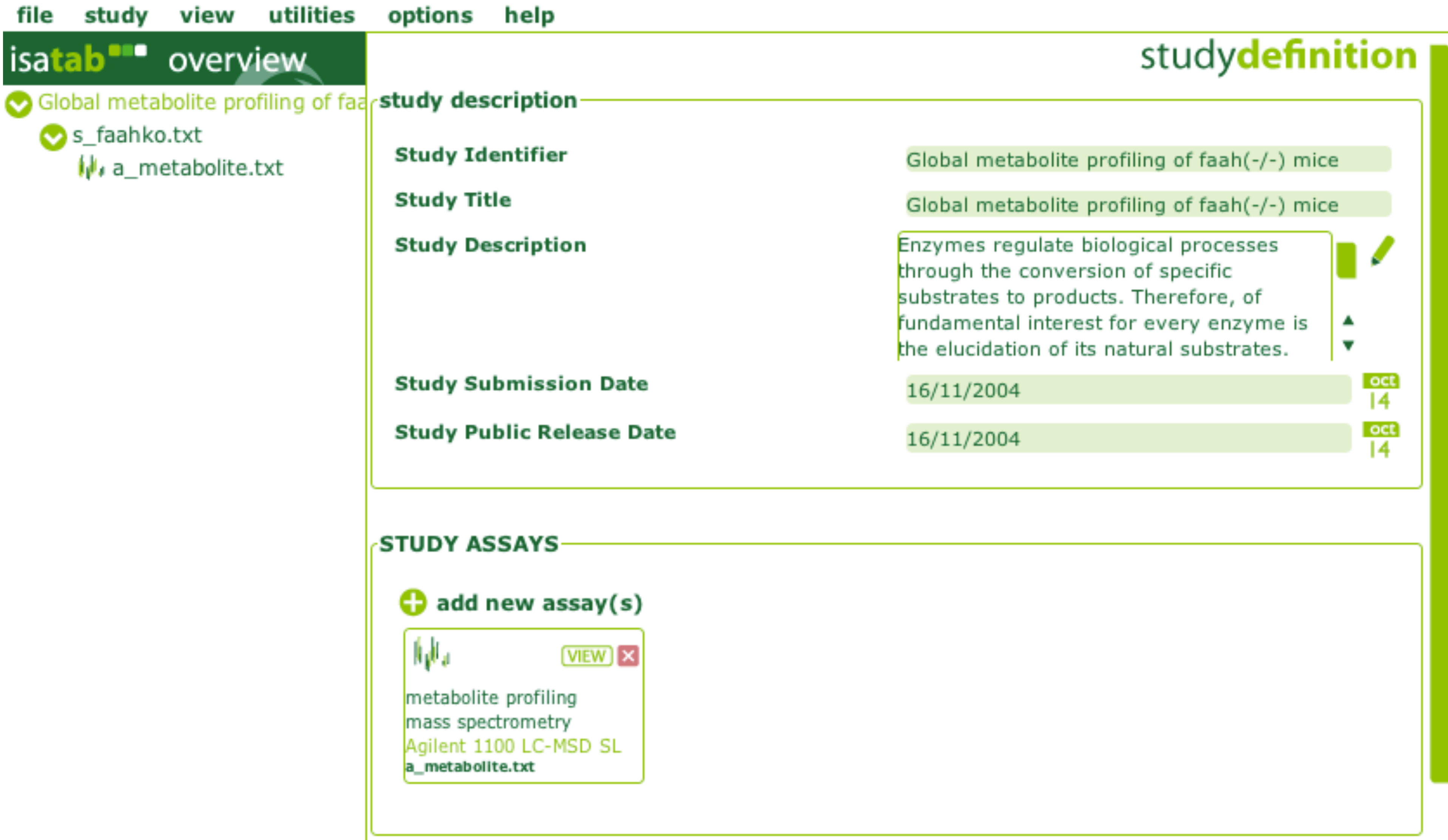
ABSTRACT: Enzymes regulate biological processes through the conversion of specific substrates to products. Therefore, of fundamental interest for every enzyme is the elucidation of its natural substrates. Here, we describe a general strategy for identifying endogenous substrates of enzymes by untargeted liquid

Saghatlian et al.
Biochemstry. 2004

- LC/MS peaks from the spinal cords of 6 wild-type and 6 FAAH (fatty acid amyde hydrolase) knockout mice

- Define key entities (e.g. factors, protocols, parameters)
- Grouping of studies
- Relate studies and assays

faahKO investigation



Study Identifier: Global metabolite profiling of faah(-/-) mice

Study Title: Global metabolite profiling of faah(-/-) mice

Study Description: Enzymes regulate biological processes through the conversion of specific substrates to products. Therefore, of fundamental interest for every enzyme is the elucidation of its natural substrates.

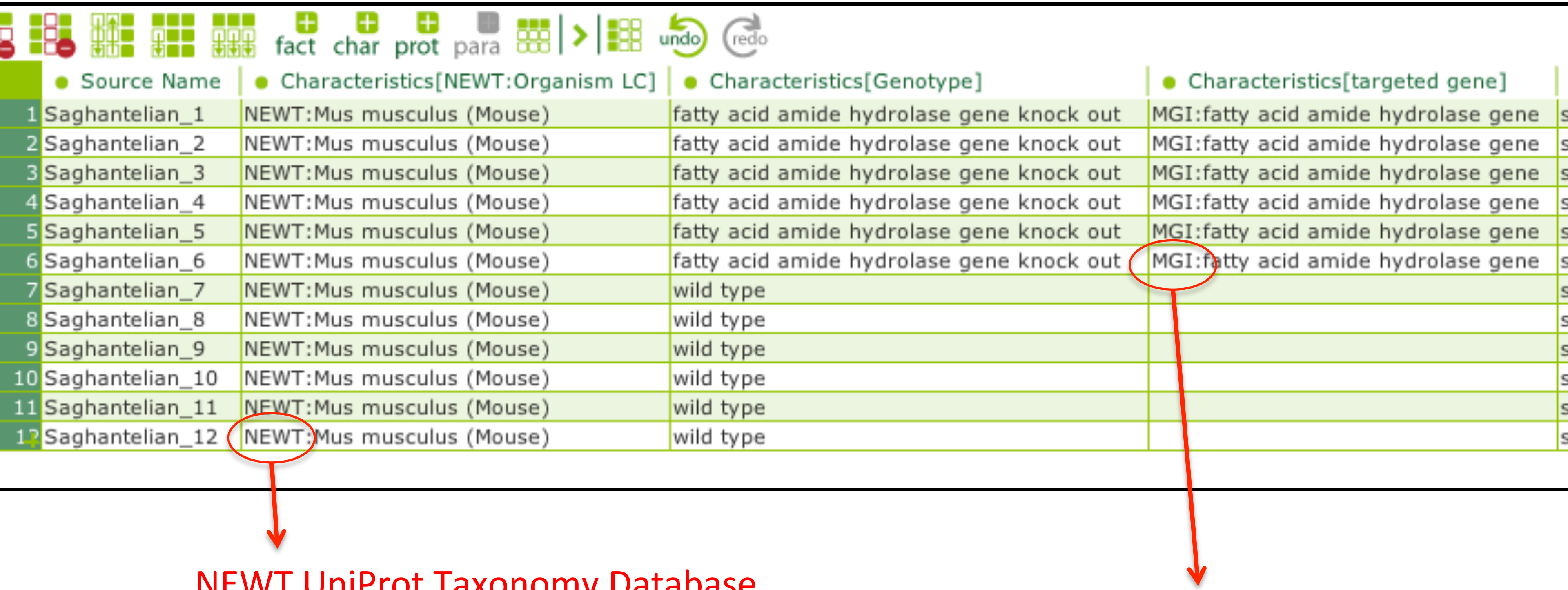
Study Submission Date: 16/11/2004

Study Public Release Date: 16/11/2004

metabolite profiling
mass spectrometry
Agilent 1100 LC-MSD SL
a_metabolite.txt

faahKO study

- Subjects studied: source(s), sampling methodology, characteristics
- treatments/manipulations performed to prepare the specimens



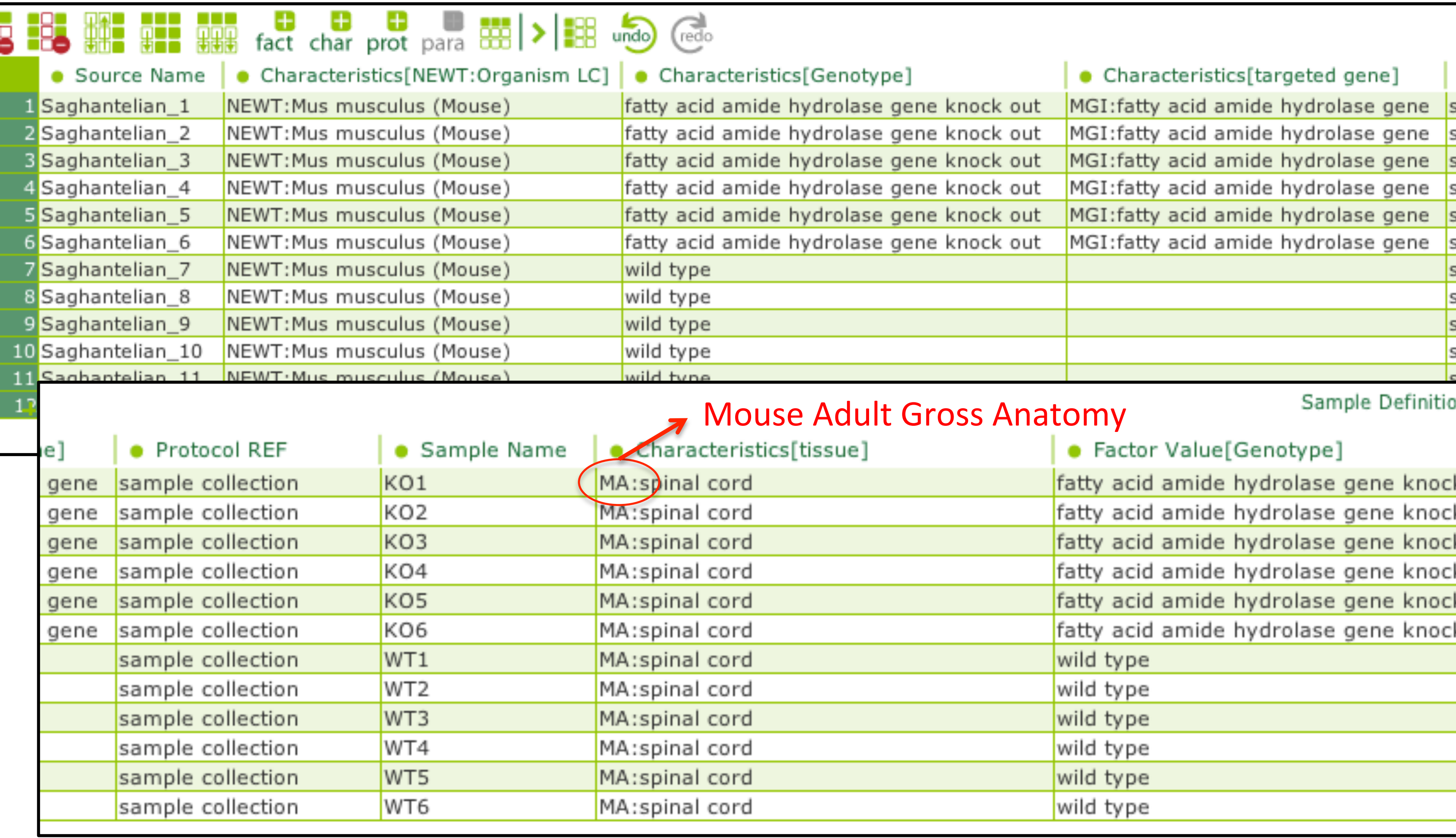
	Source Name	Characteristics[NEWT:Organism LC]	Characteristics[Genotype]	Characteristics[targeted gene]
1	Saghantelian_1	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
2	Saghantelian_2	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
3	Saghantelian_3	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
4	Saghantelian_4	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
5	Saghantelian_5	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
6	Saghantelian_6	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
7	Saghantelian_7	NEWT:Mus musculus (Mouse)	wild type	
8	Saghantelian_8	NEWT:Mus musculus (Mouse)	wild type	
9	Saghantelian_9	NEWT:Mus musculus (Mouse)	wild type	
10	Saghantelian_10	NEWT:Mus musculus (Mouse)	wild type	
11	Saghantelian_11	NEWT:Mus musculus (Mouse)	wild type	
12	Saghantelian_12	NEWT:Mus musculus (Mouse)	wild type	

NEWT UniProt Taxonomy Database

Mouse Genome Informatics

faahKO study

- Subjects studied: source(s), sampling methodology, characteristics
- treatments/manipulations performed to prepare the specimens



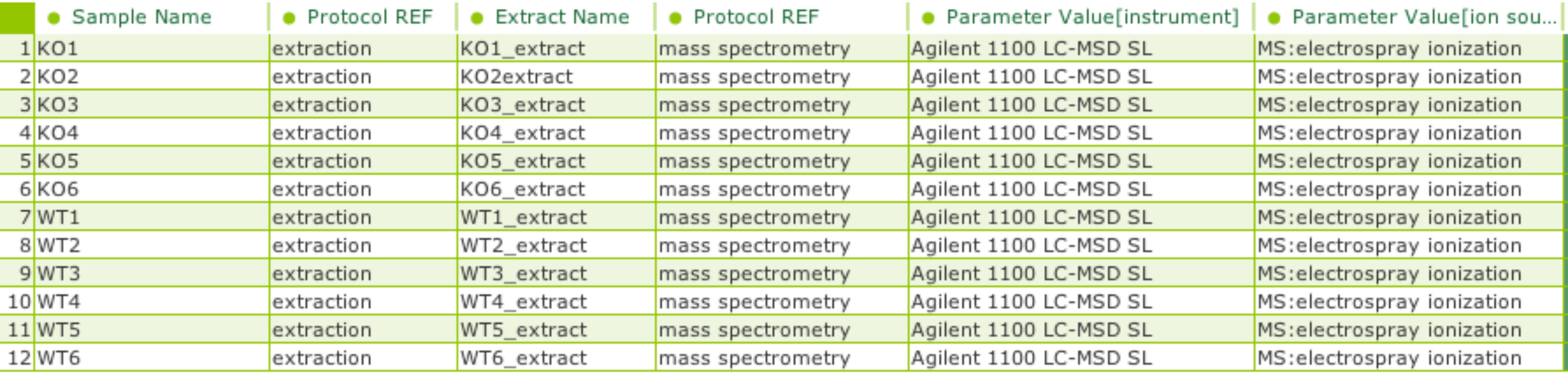
	Source Name	Characteristics[NEWT:Organism LC]	Characteristics[Genotype]	Characteristics[targeted gene]
1	Saghantelian_1	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
2	Saghantelian_2	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
3	Saghantelian_3	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
4	Saghantelian_4	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
5	Saghantelian_5	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
6	Saghantelian_6	NEWT:Mus musculus (Mouse)	fatty acid amide hydrolase gene knock out	MGI:fatty acid amide hydrolase gene
7	Saghantelian_7	NEWT:Mus musculus (Mouse)	wild type	
8	Saghantelian_8	NEWT:Mus musculus (Mouse)	wild type	
9	Saghantelian_9	NEWT:Mus musculus (Mouse)	wild type	
10	Saghantelian_10	NEWT:Mus musculus (Mouse)	wild type	
11	Saghantelian_11	NEWT:Mus musculus (Mouse)	wild type	

Mouse Adult Gross Anatomy

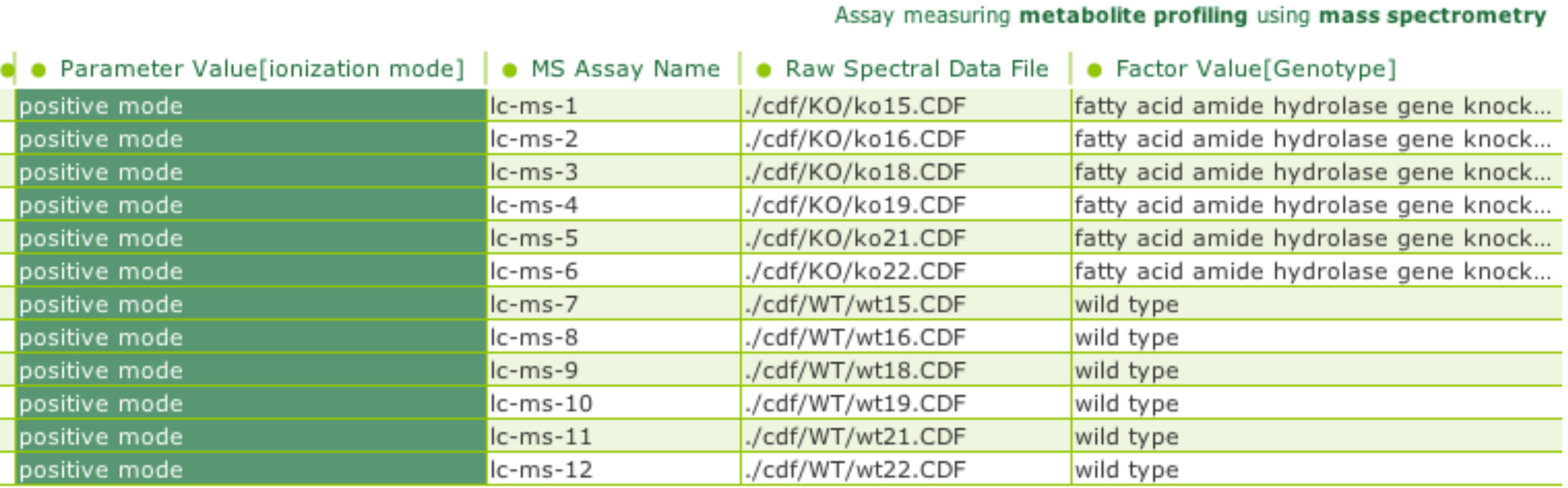
e]	Protocol REF	Sample Name	Characteristics[tissue]	Factor Value[Genotype]
gene	sample collection	KO1	MA:spinal cord	fatty acid amide hydrolase gene knoc
gene	sample collection	KO2	MA:spinal cord	fatty acid amide hydrolase gene knoc
gene	sample collection	KO3	MA:spinal cord	fatty acid amide hydrolase gene knoc
gene	sample collection	KO4	MA:spinal cord	fatty acid amide hydrolase gene knoc
gene	sample collection	KO5	MA:spinal cord	fatty acid amide hydrolase gene knoc
gene	sample collection	KO6	MA:spinal cord	fatty acid amide hydrolase gene knoc
	sample collection	WT1	MA:spinal cord	wild type
	sample collection	WT2	MA:spinal cord	wild type
	sample collection	WT3	MA:spinal cord	wild type
	sample collection	WT4	MA:spinal cord	wild type
	sample collection	WT5	MA:spinal cord	wild type
	sample collection	WT6	MA:spinal cord	wild type

- measurement type, e.g. metabolite profiling
- technology, e.g. mass spectrometry

faahKO assay



	Sample Name	Protocol REF	Extract Name	Protocol REF	Parameter Value[instrument]	Parameter Value[ion sou…
1	KO1	extraction	KO1_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
2	KO2	extraction	KO2extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
3	KO3	extraction	KO3_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
4	KO4	extraction	KO4_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
5	KO5	extraction	KO5_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
6	KO6	extraction	KO6_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
7	WT1	extraction	WT1_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
8	WT2	extraction	WT2_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
9	WT3	extraction	WT3_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
10	WT4	extraction	WT4_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
11	WT5	extraction	WT5_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization
12	WT6	extraction	WT6_extract	mass spectrometry	Agilent 1100 LC-MSD SL	MS:electrospray ionization

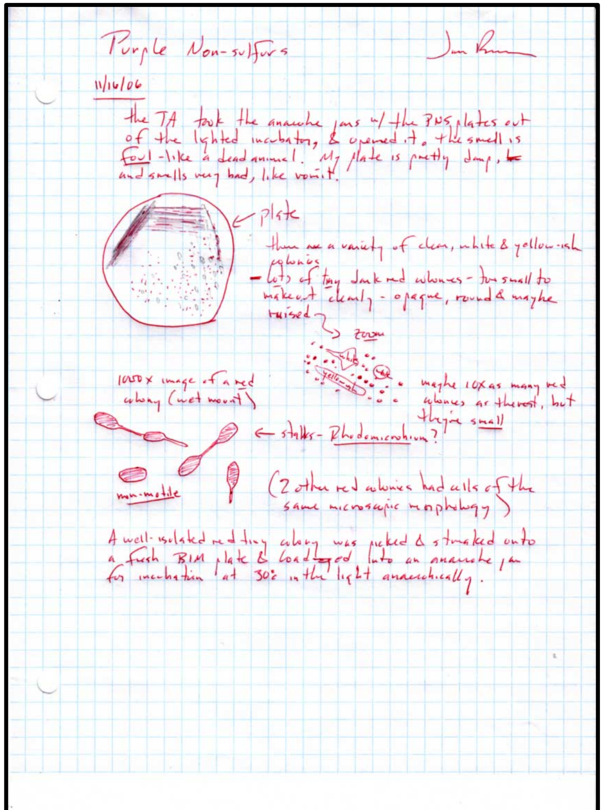


Assay measuring **metabolite profiling** using **mass spectrometry**

Parameter Value[ionization mode]	MS Assay Name	Raw Spectral Data File	Factor Value[Genotype]
positive mode	lc-ms-1	./cdf/KO/ko15.CDF	fatty acid amide hydrolase gene knock…
positive mode	lc-ms-2	./cdf/KO/ko16.CDF	fatty acid amide hydrolase gene knock…
positive mode	lc-ms-3	./cdf/KO/ko18.CDF	fatty acid amide hydrolase gene knock…
positive mode	lc-ms-4	./cdf/KO/ko19.CDF	fatty acid amide hydrolase gene knock…
positive mode	lc-ms-5	./cdf/KO/ko21.CDF	fatty acid amide hydrolase gene knock…
positive mode	lc-ms-6	./cdf/KO/ko22.CDF	fatty acid amide hydrolase gene knock…
positive mode	lc-ms-7	./cdf/WT/wt15.CDF	wild type
positive mode	lc-ms-8	./cdf/WT/wt16.CDF	wild type
positive mode	lc-ms-9	./cdf/WT/wt18.CDF	wild type
positive mode	lc-ms-10	./cdf/WT/wt19.CDF	wild type
positive mode	lc-ms-11	./cdf/WT/wt21.CDF	wild type
positive mode	lc-ms-12	./cdf/WT/wt22.CDF	wild type

isatab

isatab



Purple Non-sulfurs

11/16/06

the TA took the anaerobe jars w/ the PNS plates out of the lighted incubator, & opened it. the smell is foul - like a dead animal. My plate is pretty damp, and smells very bad, like vomit.

← plate

there are a variety of clear, white & yellow-ish colonies

- lots of tiny dark red colonies - too small to make out clearly - opaque, round & maybe raised

zoom

maybe 10X as many red colonies as the rest, but they're small

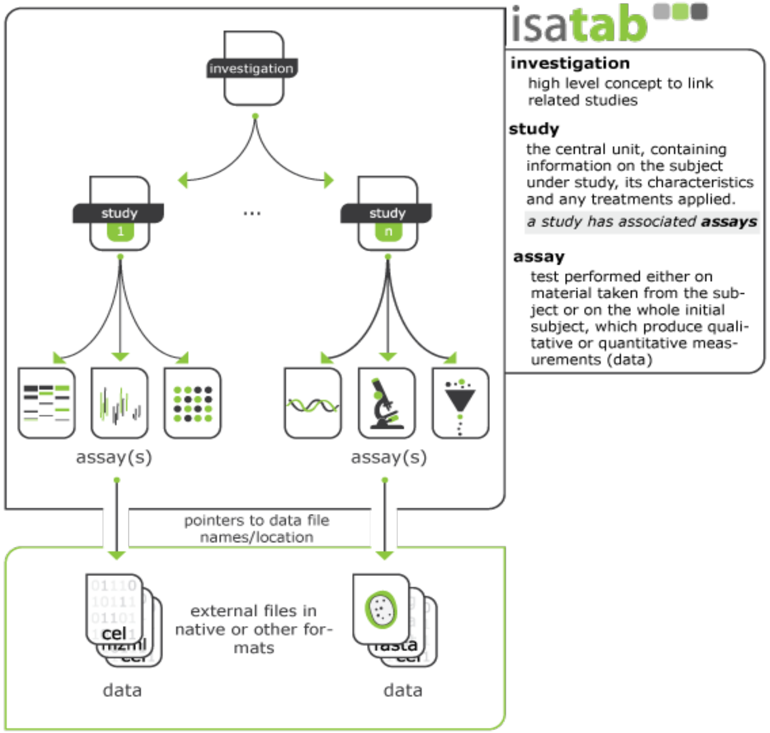
1000x image of a red colony (wet mount)

← stalks - Rhodomicrobium?

non-motile

(2 other red colonies had cells of the same microscopic morphology)

A well-isolated red tiny colony was picked & streaked onto a fresh BIM plate & loaded into an anaerobe jar for incubation at 30°C in the light anaerobically.



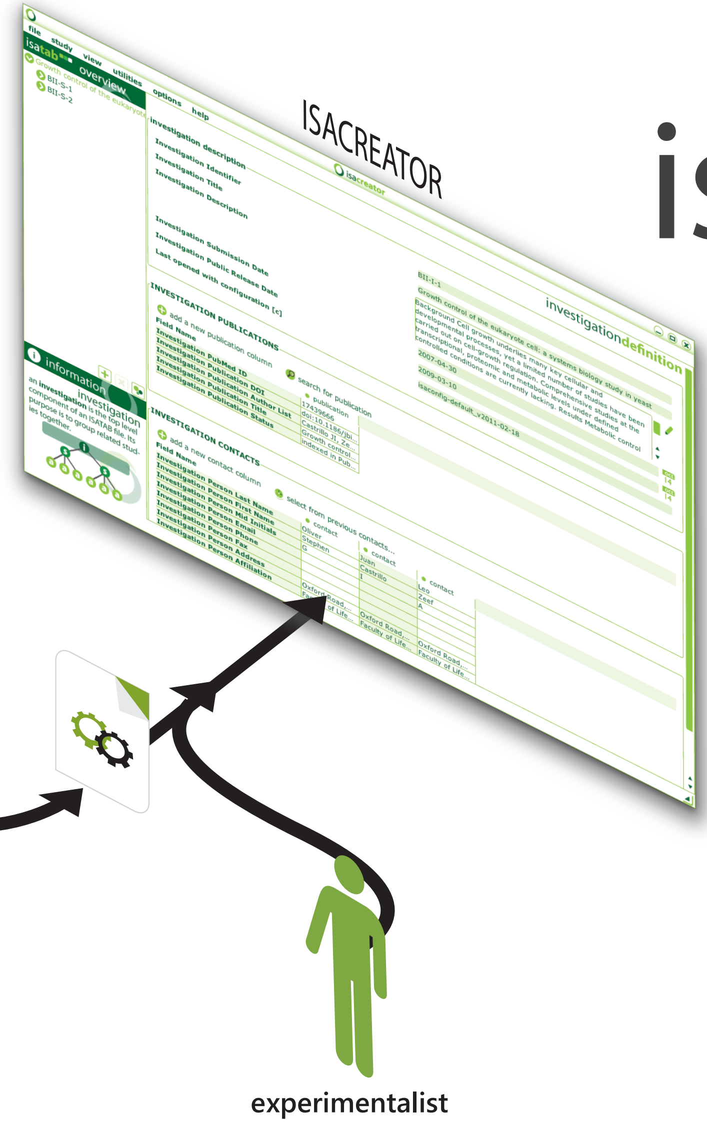
investigation
high level concept to link related studies

study
the central unit, containing information on the subject under study, its characteristics and any treatments applied.
*a study has associated **assays***

assay
test performed either on material taken from the subject or on the whole initial subject, which produce qualitative or quantitative measurements (data)

isacreator

isatab



Growth control of the eukaryote
BII-S-1
BII-S-2

an **investigation** is the top level component of an ISATAB file. Its purpose is to group related studies together.

investigation description
Investigation Identifier: BII-I-1
Investigation Title: Growth control of the eukaryote cell: a systems biology study in yeast
Investigation Description: Background Cell growth underlies many key cellular and developmental processes, yet a limited number of studies have been carried out on cell-growth regulation. Comprehensive studies at the transcriptional, proteomic and metabolic levels under defined controlled conditions are currently lacking. Results Metabolic control
Investigation Submission Date: 2007-04-30
Investigation Public Release Date: 2009-03-10
Last opened with configuration [c]: isaconfig-default_v2011-02-18

INVESTIGATION PUBLICATIONS
add a new publication column
search for publication

Field Name	publication
Investigation PubMed ID	17439666
Investigation Publication DOI	doi:10.1186/jbi...
Investigation Publication Author List	Castrillo JI, Ze...
Investigation Publication Title	Growth control...
Investigation Publication Status	indexed in Pub...

INVESTIGATION CONTACTS
add a new contact column
select from previous contacts...

Field Name	contact	contact	contact
Investigation Person Last Name	Oliver	Juan	Leo
Investigation Person First Name	Stephen	Castrillo	Zeef
Investigation Person Mid Initials	G	I	A
Investigation Person Email			
Investigation Person Phone			
Investigation Person Fax			
Investigation Person Address	Oxford Road,...	Oxford Road,...	Oxford Road,...
Investigation Person Affiliation	Faculty of Life...	Faculty of Life...	Faculty of Life...

STUDY ASSAYS
add new assay(s)

protein expression profiling
mass spectrometry
iTRAQ
a_proteome.txt

metabolite profiling
mass spectrometry
LC-MS/MS
a_metabolome.txt

transcription profiling
DNA microarray
Affymetrix
a_transcriptome.txt

metabolites in S. ce
nutritional or other
we carried out our

STUDY DESIGN DESCRIPTORS

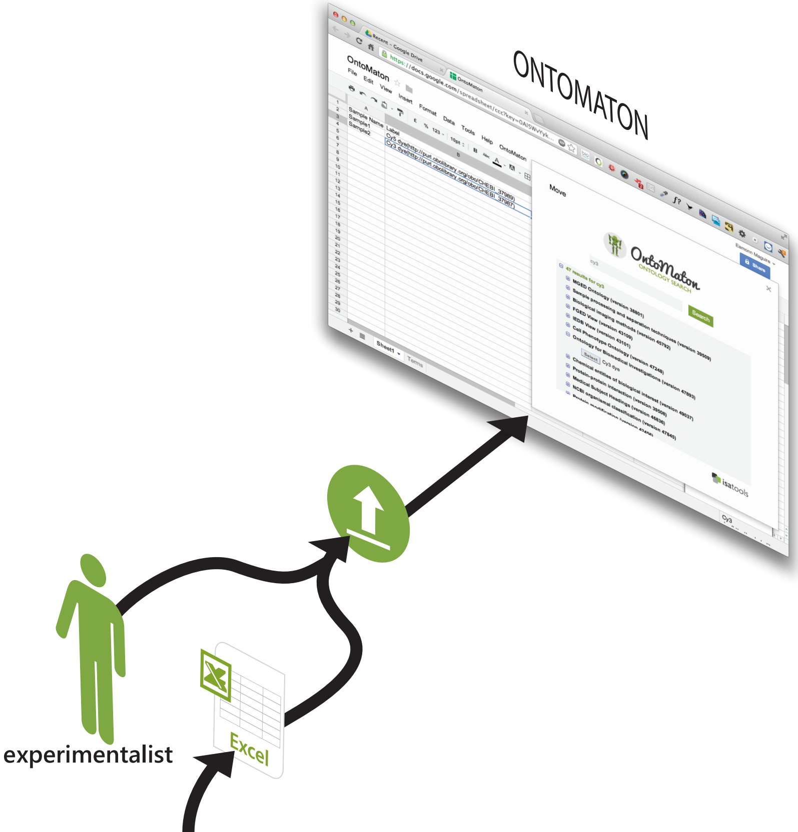
1 recommended ontologies
Chemical entities of biological interest
subatomic particle (CHEBI:36342)
role (CHEBI:50906)
chemical entity (CHEBI:24431)
atom (CHEBI:33250)
chemical substance (CHEBI:59999)
pure substance (CHEBI:60003)
mixture (CHEBI:60004)
phlobaphene (CHEBI:61197)
GSL-4A glycosphingolipid (CHEBI:61857)
lecithin (CHEBI:61995)
mineral (CHEBI:46662)

term not found? just enter freetext: CHEBI:acetylsalicylic acid

Term name: acetylsalicylic acid
Chemical structure:
EXACT SYNONYM: Acetylsalicylic acid

OntoMaton

isatab



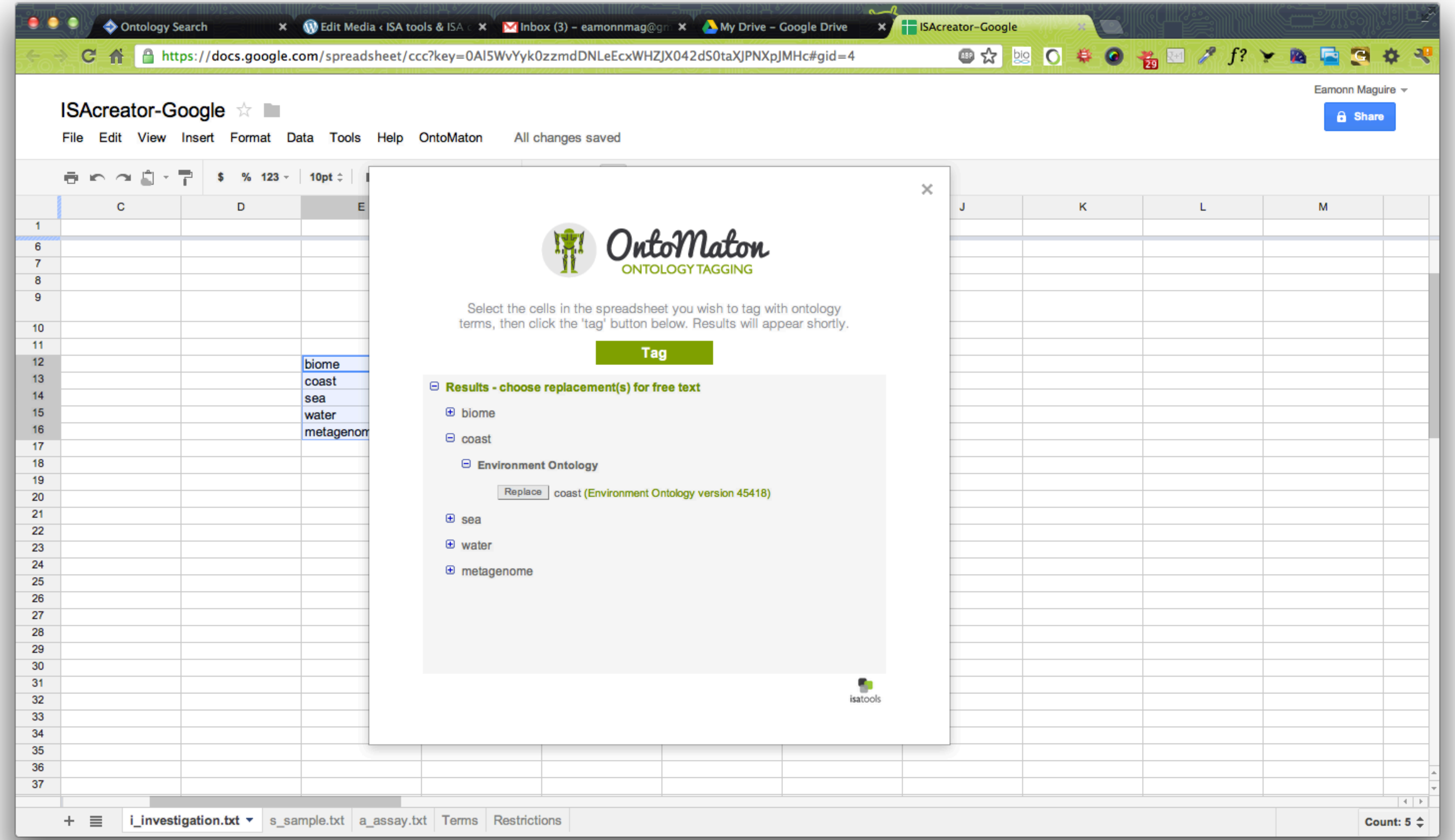
Report and edit the description of the investigation using Google Spreadsheets.

Use Google Spreadsheets in combination with ISA-Tab templates (created through importing the Excel file from the ISAconfigurator) and **OntoMaton** (for ontology search and tagging support) to report an investigation.

OntoMaton

Ontology Search and Tagging in Google Spreadsheets

- collaborative annotation
- distributed groups of users
- version control & history



Select the cells in the spreadsheet you wish to tag with ontology terms, then click the 'tag' button below. Results will appear shortly.

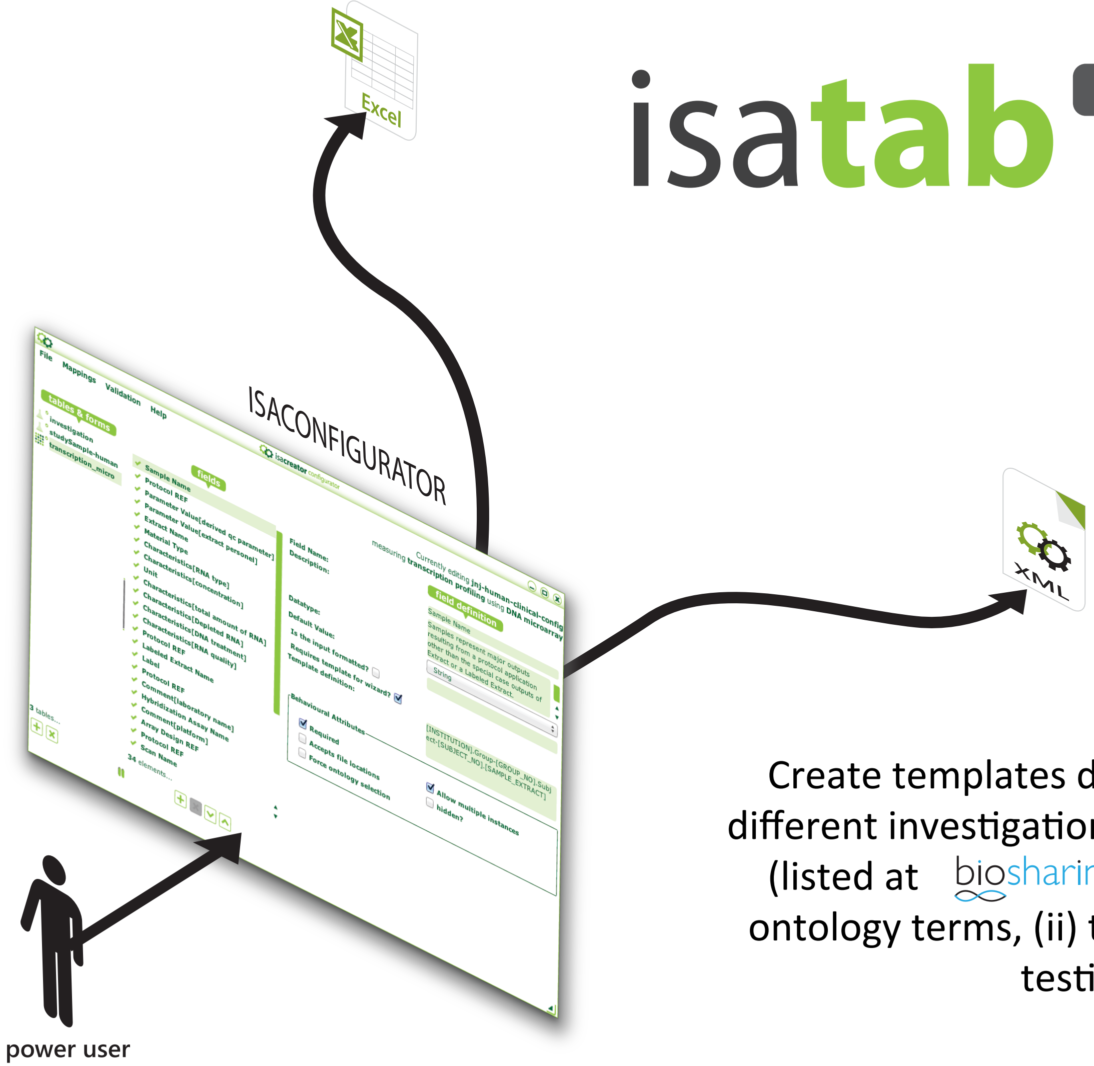
Tag

⊟ **Results - choose replacement(s) for free text**

- ⊞ biome
- ⊟ coast
  - ⊟ **Environment Ontology**
    - Replace coast (Environment Ontology version 45418)
- ⊞ sea
- ⊞ water
- ⊞ metagenome

isacreator
configurator

isatab



investigation
studySample-human
transcription_micro

Sample Name
Protocol REF
Parameter Value[derived qc parameter]
Parameter Value[extract personel]
Extract Name
Material Type
Characteristics[RNA type]
Characteristics[concentration]
Unit
Characteristics[total amount of RNA]
Characteristics[Depleted RNA]
Characteristics[DNA treatment]
Characteristics[RNA quality]
Protocol REF
Labeled Extract Name
Label
Protocol REF
Comment[laboratory name]
Hybridization Assay Name
Comment[platform]
Array Design REF
Protocol REF
Scan Name

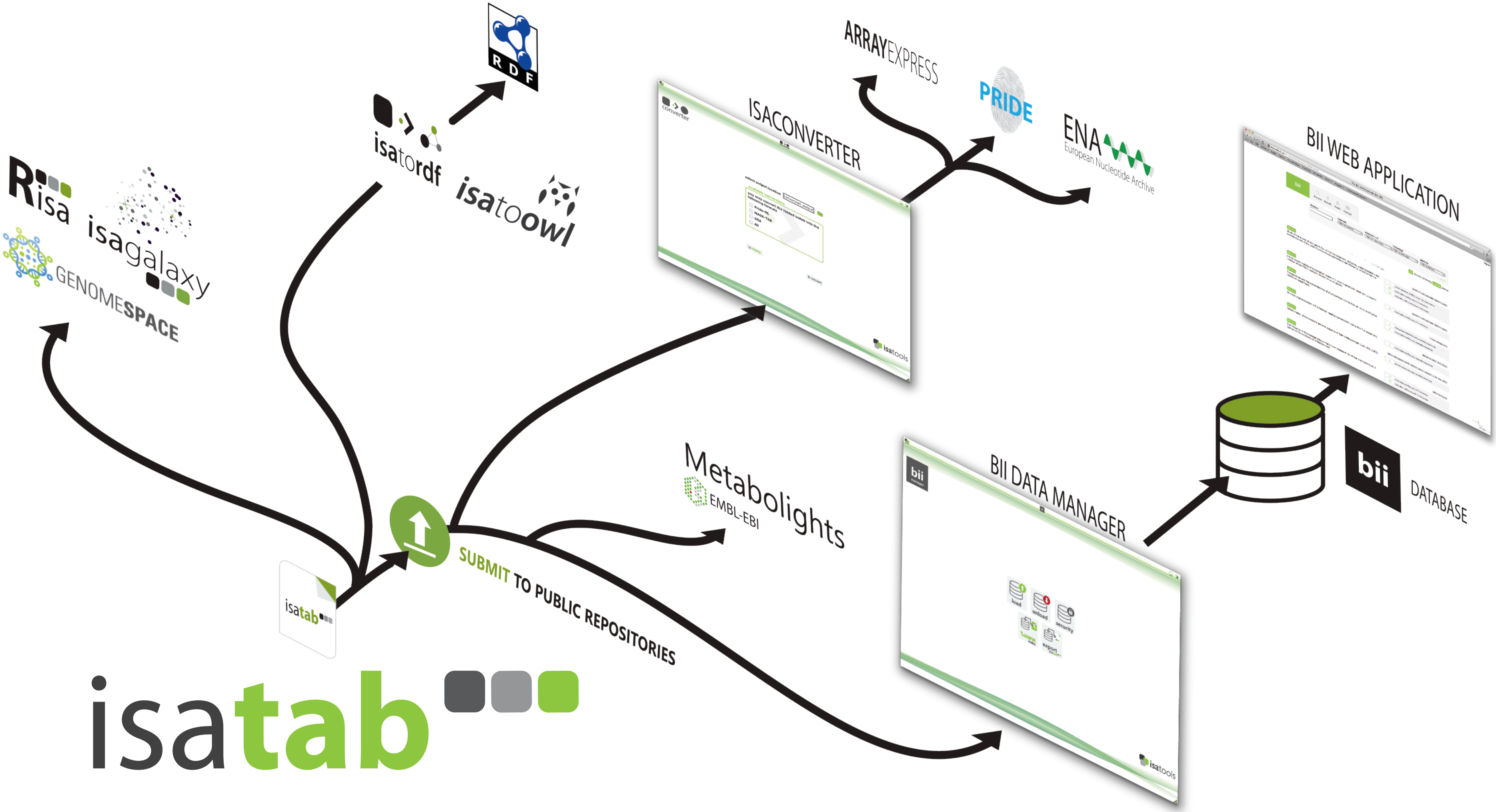
Currently editing **jnj-human-clinical-config** measuring **transcription profiling** using **DNA microarray**

Field Name: Sample Name
Description: Samples represent major outputs resulting from a protocol application other than the special case outputs of Extract or a Labeled Extract.
Datatype: String
Default Value:
Is the input formatted? ☐
Requires template for wizard? ☑
Template definition: [INSTITUTION].Group-[GROUP_NO].Subject-[SUBJECT_NO].[SAMPLE_EXTRACT]

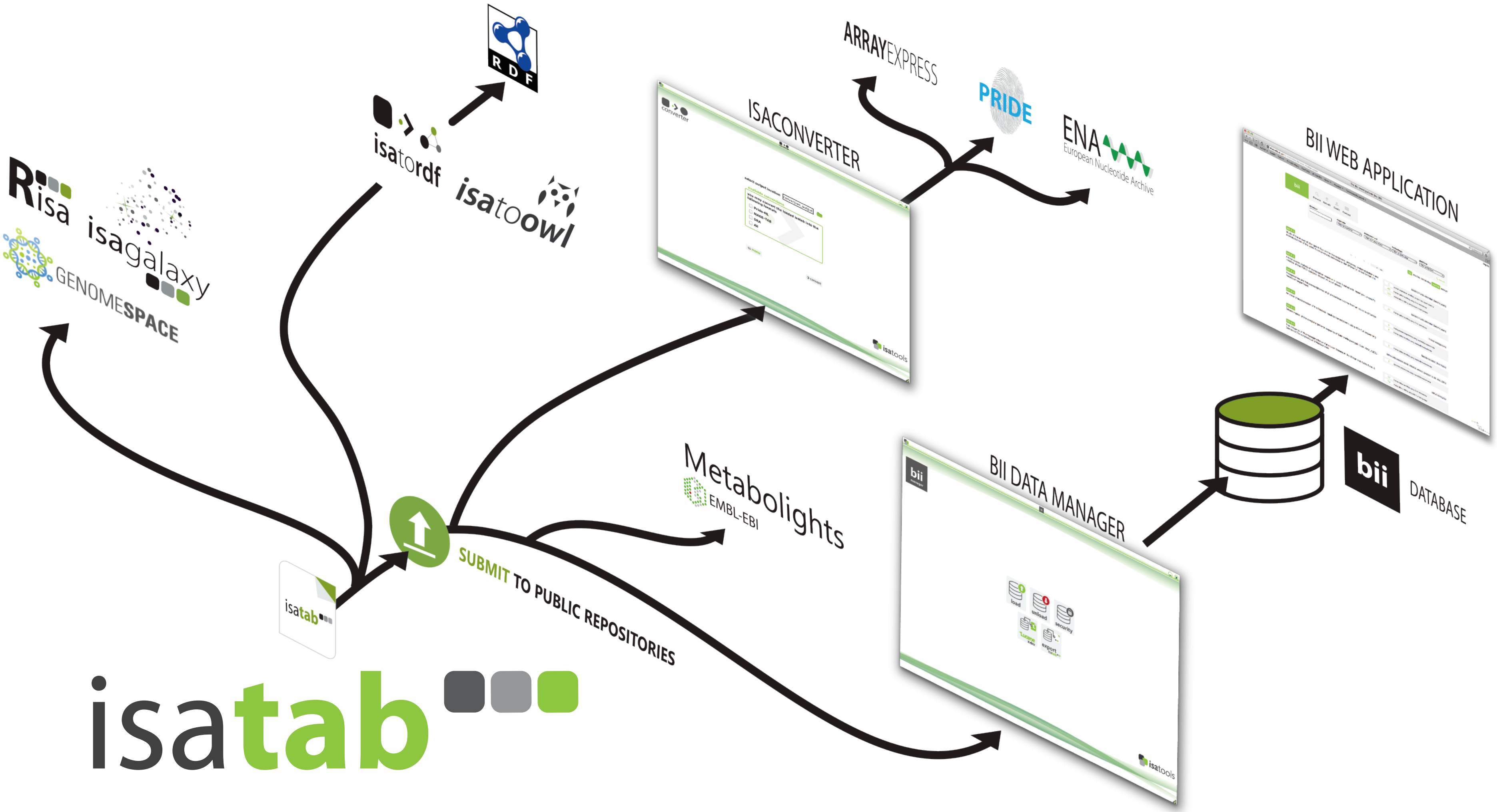
Behavioural Attributes
☑ Required
☐ Accepts file locations
☐ Force ontology selection
☑ Allow multiple instances
☐ hidden?

Create templates detailing the steps to be reported for different investigations, complying to community standards (listed at biosharing), e.g. configuring fields to be (i) ontology terms, (ii) text (with/without regular expression testing), (iii) numbers etc.

From the **ISA-Tab** we can perform analysis, convert to RDF/OWL and other formats for submission/ sharing to local/remote repositories,



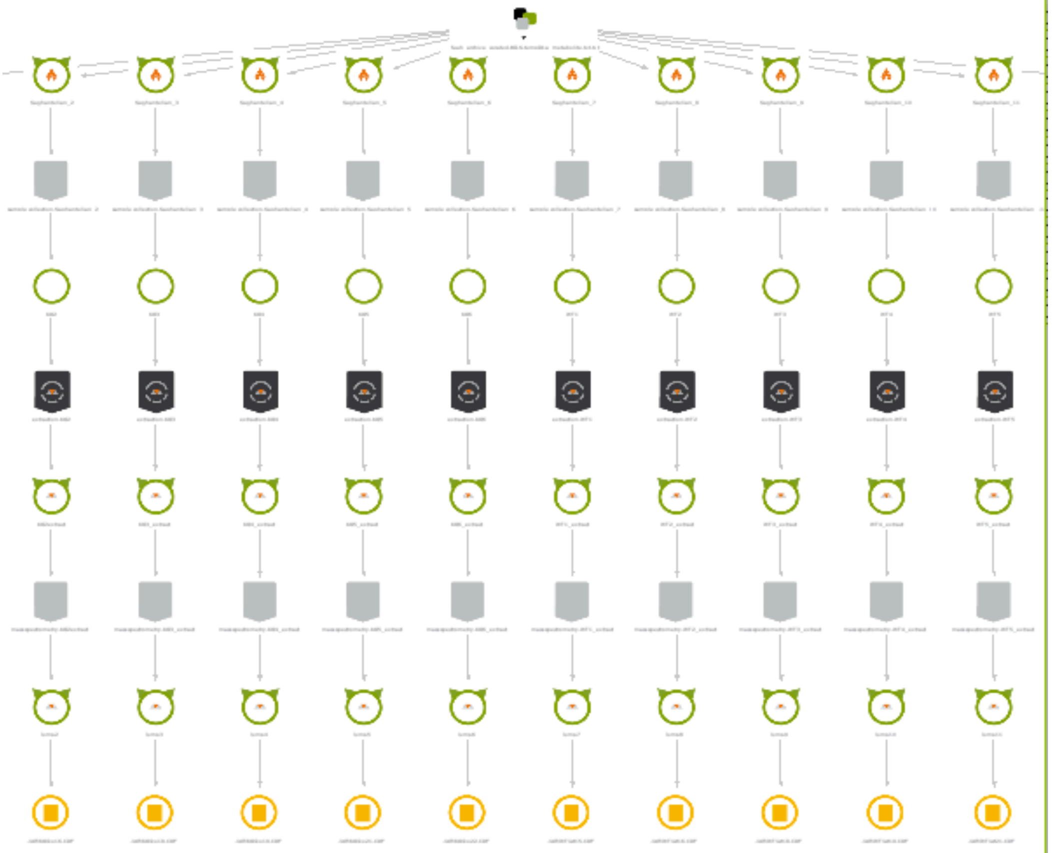
From the **ISA-Tab** we can perform analysis, convert to RDF/OWL and other formats for submission/ sharing to local/remote repositories,

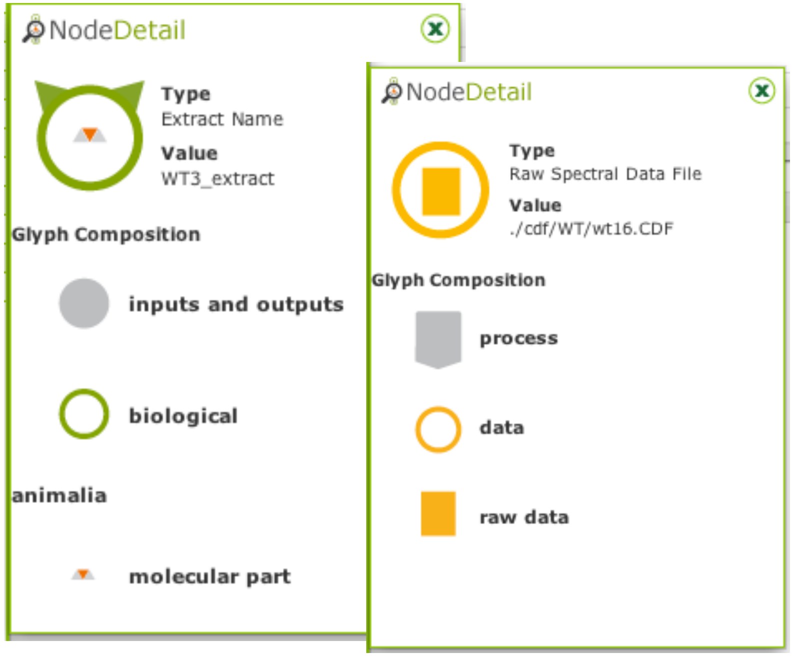


+ Visualisation Methods

faahKO Groups

faahKO Workflow





NodeDetail

Type
Extract Name
Value
WT3_extract

Glyph Composition

- inputs and outputs
- biological

animalia

- molecular part

NodeDetail

Type
Raw Spectral Data File
Value
./cdf/WT/wt16.CDF

Glyph Composition

- process
- data
- raw data

IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS
A publication of the IEEE Computer Society

Maguire E, Rocca-Serra P, Sansone SA, Davies J and Chen M.
Taxonomy-based Glyph Design -- with a Case Study on Visualizing Workflows of Biological Experiments,
IEEE Transactions on Visualization and Computer Graphics, volume 18, 2012 (in press)

Risa

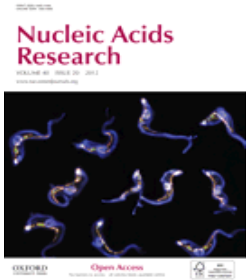
- R package available in BioConductor 2.11
  http://bioconductor.org/packages/release/bioc/html/Risa.html
- ISAtab class
- Read ISAtab files into ISAtab objects and save ISAtab files
- Build xcmsSet (xcms package) objects from mass spectrometry assays
- Augment the ISAtab dataset after analysis
- github SOCIAL CODING source & issues tracking
  GITHUB.COM/ISA-TOOLS
  https://github.com/ISA-tools/Risa

Risa

- faahKO package v. 2.12 contains ISAtab files describing the experiment

```
faahkoISA = readISAta(find.package("faahKO"))
assay.filename <- faahkoISA["assay.filenames"][[1]]
xset = processAssayXcmsSet(faahkoISA, assay.filename)
…
updateAssayMetadata(faahkoISA, assay.filename,"Derived Spectral
Data File","faahkoDSDF.txt" )
```

- MTBLS2 processing and analysis using Risa, xcms and CAMERA BioConductor packages

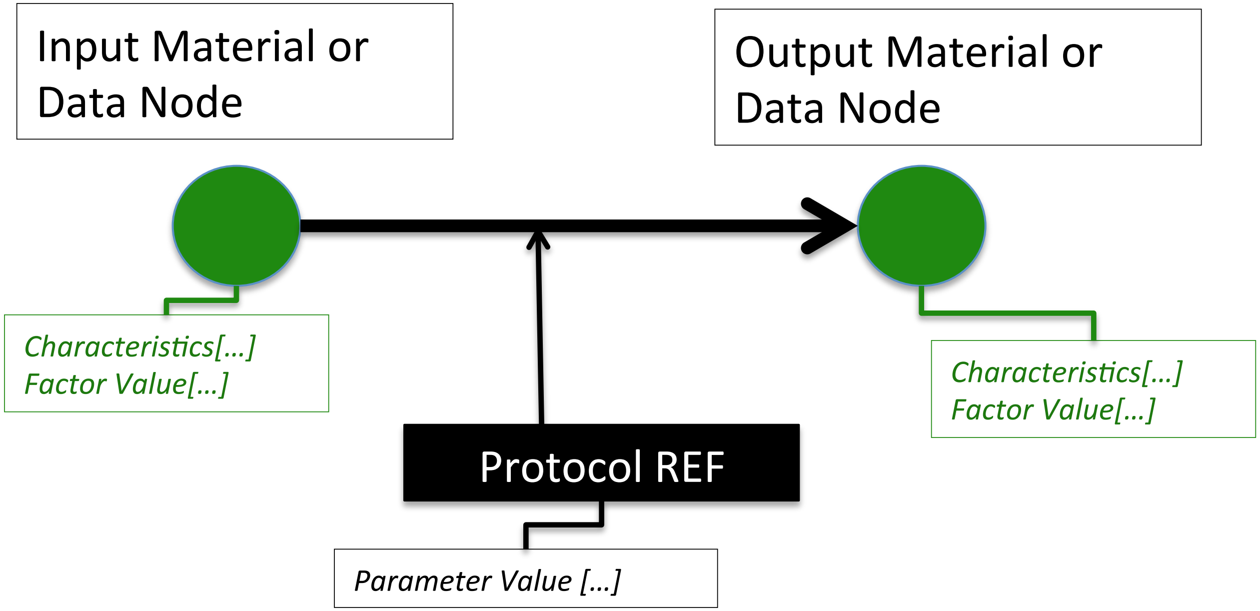


Nucleic Acids Research

Metabolights – an open access general-purpose repository for metabolomics studies and associated meta-data
Haug et al, 2012
Nucleic Acids Research

ISA syntax
& Underlying Material/Data workflows

metaboliteprofiling_ms



Input Material or Data Node

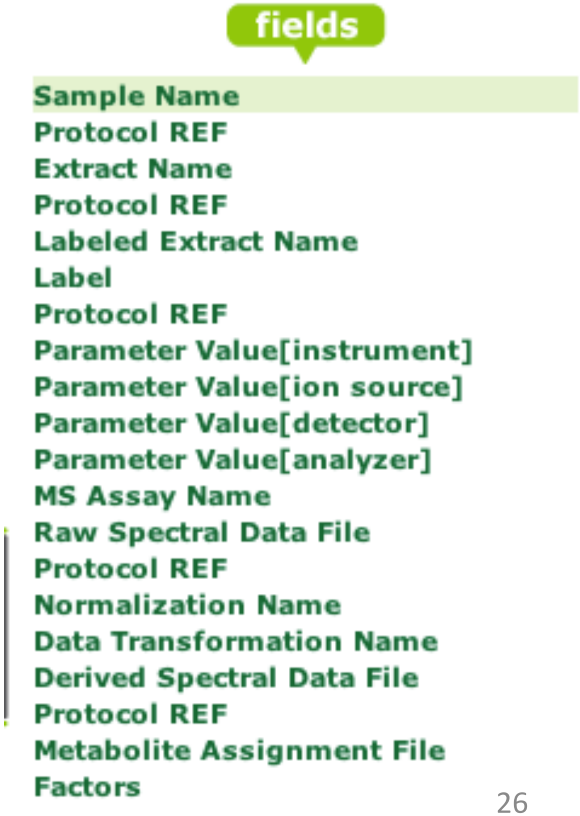
Output Material or Data Node

Characteristics[…]
Factor Value[…]

Characteristics[…]
Factor Value[…]

Protocol REF

Parameter Value […]



fields

Sample Name
Protocol REF
Extract Name
Protocol REF
Labeled Extract Name
Label
Protocol REF
Parameter Value[instrument]
Parameter Value[ion source]
Parameter Value[detector]
Parameter Value[analyzer]
MS Assay Name
Raw Spectral Data File
Protocol REF
Normalization Name
Data Transformation Name
Derived Spectral Data File
Protocol REF
Metabolite Assignment File
Factors

isatordf isatoowl

- Make the semantics of ISAtab explicit, including materials & data entities & processes
- Exploit the semantic annotations available in ISAtab datasets
- Augment ISA syntax with new elements (e.g. groups), facilitating the understanding & querying of experimental design
- Facilitate data integration & knowledge discovery/reasoning

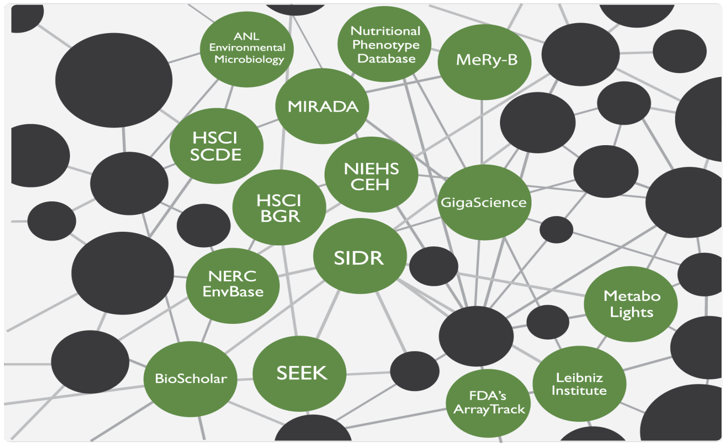
ISAtab datasets as linked data

- Connect to the growing Linked Data universe

  RDF = Resource Description Framework, OWL = Web Ontology Language

- Collaborations with Toxbank (isatordf)

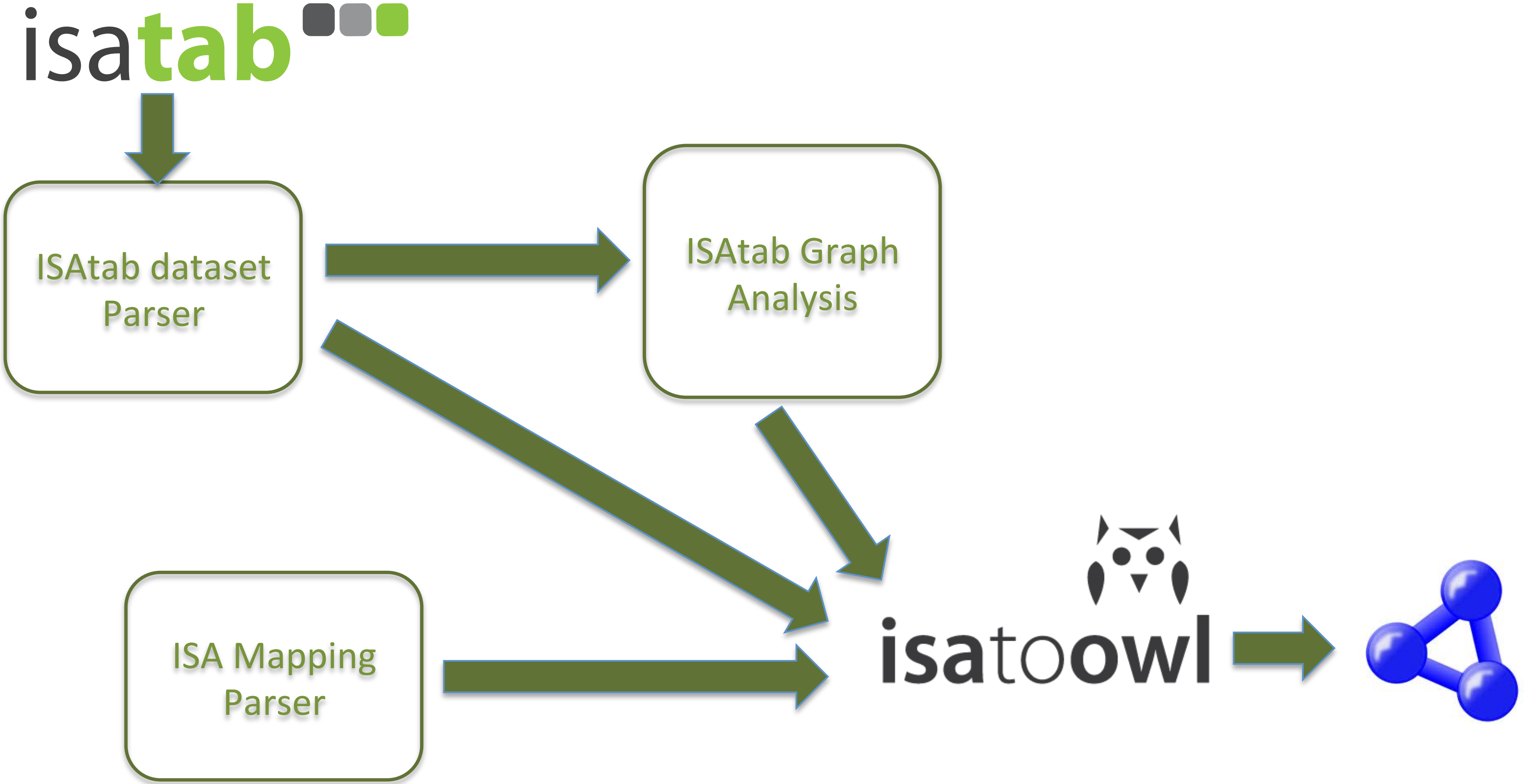
  & W3C Health Care & Life Sciences Interest Group (HCLSIG)



<subject, predicate, object>

<lipoprotein> <participates_in> <inflammatory response>

<PRO:212342352> <BFO_0000056> <GO:0006954>



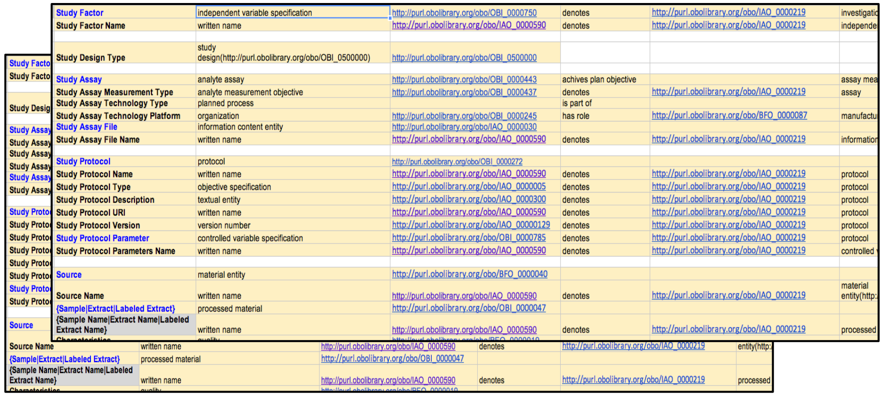
isatab

ISAtab dataset Parser

ISAtab Graph Analysis

ISA Mapping Parser

isatoowl



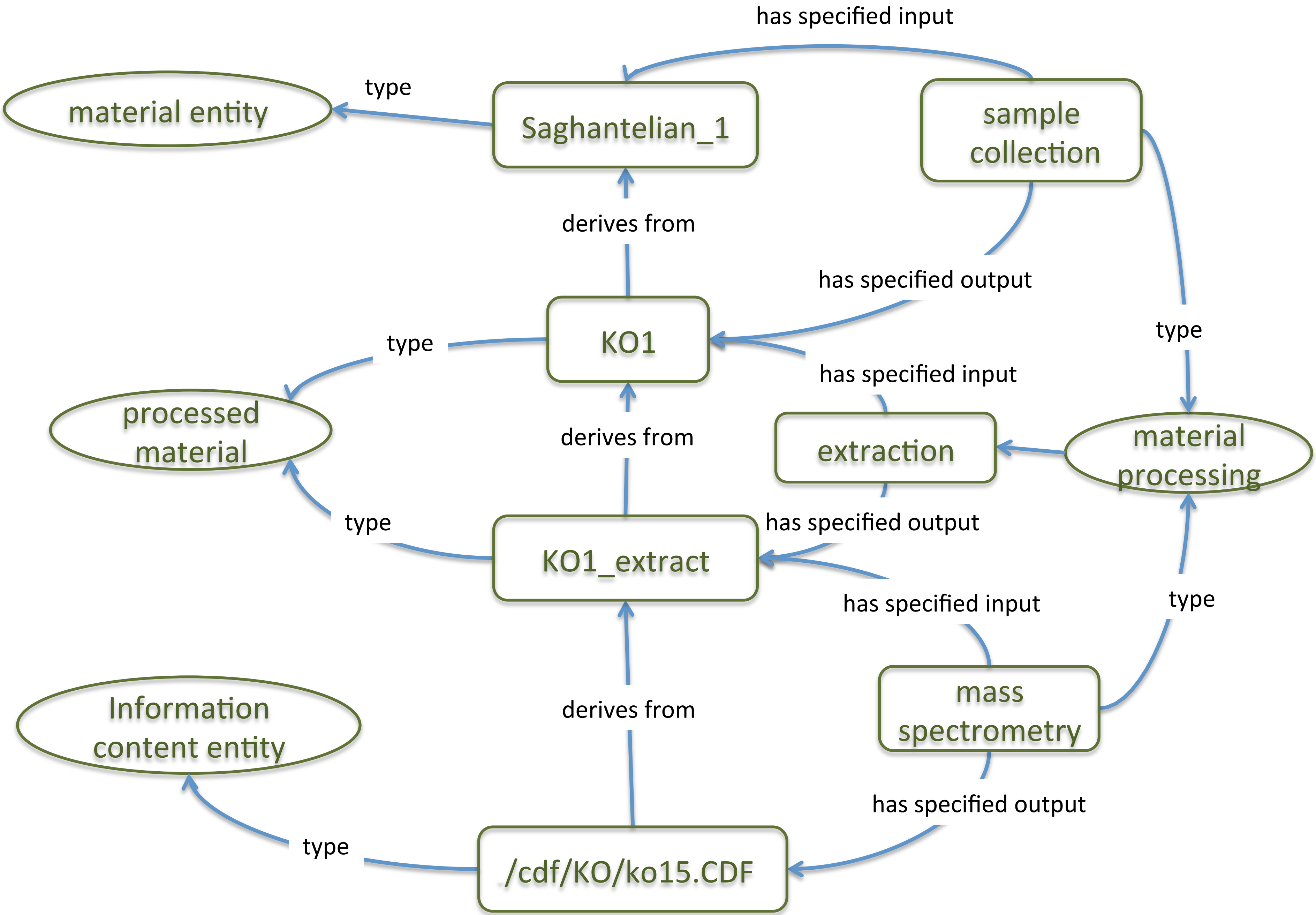
Study Factor	independent variable specification	http://purl.obolibrary.org/obo/OBI_0000750	denotes	http://purl.obolibrary.org/obo/IAO_0000219	investigatio
Study Factor Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes	http://purl.obolibrary.org/obo/IAO_0000219	independe
Study Design Type	study design(http://purl.obolibrary.org/obo/OBI_0500000)	http://purl.obolibrary.org/obo/OBI_0500000			
Study Assay	analyte assay	http://purl.obolibrary.org/obo/OBI_0000443	achives plan objective		assay mea
Study Assay Measurement Type	analyte measurement objective	http://purl.obolibrary.org/obo/OBI_0000437	denotes	http://purl.obolibrary.org/obo/IAO_0000219	assay
Study Assay Technology Type	planned process		is part of		
Study Assay Technology Platform	organization	http://purl.obolibrary.org/obo/OBI_0000245	has role	http://purl.obolibrary.org/obo/BFO_0000087	manufactu
Study Assay File	information content entity	http://purl.obolibrary.org/obo/IAO_0000030			
Study Assay File Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes	http://purl.obolibrary.org/obo/IAO_0000219	information
Study Protocol	protocol	http://purl.obolibrary.org/obo/OBI_0000272			
Study Protocol Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes	http://purl.obolibrary.org/obo/IAO_0000219	protocol
Study Protocol Type	objective specification	http://purl.obolibrary.org/obo/IAO_0000005	denotes	http://purl.obolibrary.org/obo/IAO_0000219	protocol
Study Protocol Description	textual entity	http://purl.obolibrary.org/obo/IAO_0000300	denotes	http://purl.obolibrary.org/obo/IAO_0000219	protocol
Study Protocol URI	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes	http://purl.obolibrary.org/obo/IAO_0000219	protocol
Study Protocol Version	version number	http://purl.obolibrary.org/obo/IAO_00000129	denotes	http://purl.obolibrary.org/obo/IAO_0000219	protocol
Study Protocol Parameter	controlled variable specification	http://purl.obolibrary.org/obo/OBI_0000785	denotes	http://purl.obolibrary.org/obo/IAO_0000219	protocol
Study Protocol Parameters Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes	http://purl.obolibrary.org/obo/IAO_0000219	controlled
Source	material entity	http://purl.obolibrary.org/obo/BFO_0000040			
Source Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes	http://purl.obolibrary.org/obo/IAO_0000219	material entity(http:
{Sample|Extract|Labeled Extract}	processed material	http://purl.obolibrary.org/obo/OBI_0000047			
{Sample Name|Extract Name|Labeled Extract Name}	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes	http://purl.obolibrary.org/obo/IAO_0000219	processed

Source Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes	http://purl.obolibrary.org/obo/IAO_0000219	entity(http:
{Sample|Extract|Labeled Extract}	processed material	http://purl.obolibrary.org/obo/OBI_0000047			
{Sample Name|Extract Name|Labeled Extract Name}	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes	http://purl.obolibrary.org/obo/IAO_0000219	processed

OntoMaton

ISA-OBO-mapping

Study Factor	independent variable specification	http://purl.obolibrary.org/obo/OBI_0000750	denotes
Study Factor Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes
Study Design Type	study design(http://purl.obolibrary.org/obo/OBI_0500000)	http://purl.obolibrary.org/obo/OBI_0500000	
Study Assay	analyte assay	http://purl.obolibrary.org/obo/OBI_0000443	achives plan objecti
Study Assay Measurement Type	analyte measurement objective	http://purl.obolibrary.org/obo/OBI_0000437	denotes
Study Assay Technology Type	planned process		is part of
Study Assay Technology Platform	organization	http://purl.obolibrary.org/obo/OBI_0000245	has role
Study Assay File	information content entity	http://purl.obolibrary.org/obo/IAO_0000030	
Study Assay File Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes
Study Protocol	protocol	http://purl.obolibrary.org/obo/OBI_0000272	
Study Protocol Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes
Study Protocol Type	objective specification	http://purl.obolibrary.org/obo/IAO_0000005	denotes
Study Protocol Description	textual entity	http://purl.obolibrary.org/obo/IAO_0000300	denotes
Study Protocol URI	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes
Study Protocol Version	version number	http://purl.obolibrary.org/obo/IAO_00000129	denotes
Study Protocol Parameter	controlled variable specification	http://purl.obolibrary.org/obo/OBI_0000785	denotes
Study Protocol Parameters Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes
Source	material entity	http://purl.obolibrary.org/obo/BFO_0000040	
Source Name	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes
{Sample|Extract|Labeled Extract}	processed material	http://purl.obolibrary.org/obo/OBI_0000047	
{Sample Name|Extract Name|Labeled Extract Name}	written name	http://purl.obolibrary.org/obo/IAO_0000590	denotes
Characteristics	quality	http://purl.obolibrary.org/obo/BFO_0000019	



has specified input

material entity

type

Saghantelian_1

sample collection

derives from

has specified output

type

type

KO1

has specified input

processed material

derives from

extraction

material processing

type

has specified output

KO1_extract

has specified input

type

derives from

mass spectrometry

Information content entity

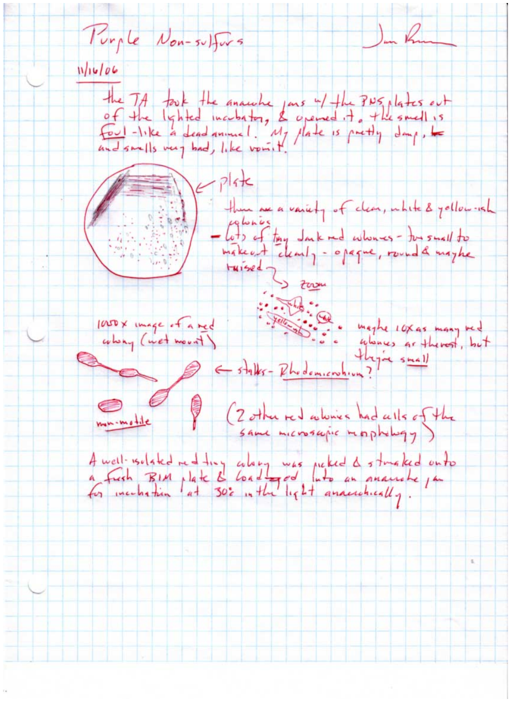
has specified output

type

/cdf/KO/ko15.CDF

Increasing level of structure…

…different target audiences



Purple Non-sulfurs

11/16/06

the TA took the anaerobe jars w/ the PNS plates out of the lighted incubator, & opened it. the smell is foul - like a dead animal. My plate is pretty damp, ~~a~~ and smells very bad, like vomit.

← plate

there are a variety of clear, white & yellow-ish colonies

- lots of tiny dark red colonies - too small to make out clearly - opaque, round & maybe raised

zoom

maybe 10X as many red colonies as the rest, but they're small

1000x image of a red colony (wet mount)

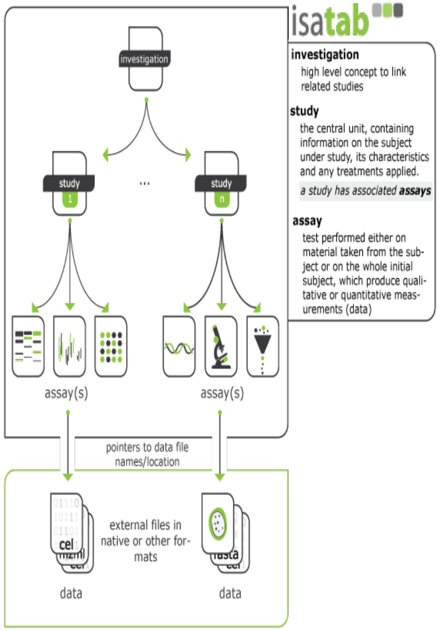
← stalks - Rhodomicrobium?

non-motile

(2 other red colonies had cells of the same microscopic morphology)

A well-isolated red tiny colony was picked & streaked onto a fresh BIM plate & loaded into an anaerobe jar for incubation at 30°C in the light anaerobically.

Notes in Lab books
(information for humans)

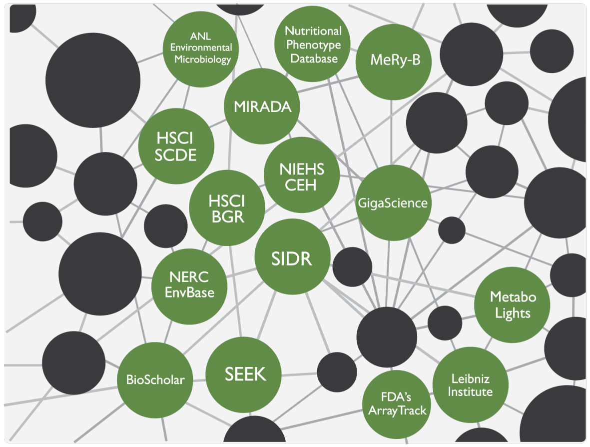


investigation
high level concept to link related studies

study
the central unit, containing information on the subject under study, its characteristics and any treatments applied.
*a study has associated **assays***

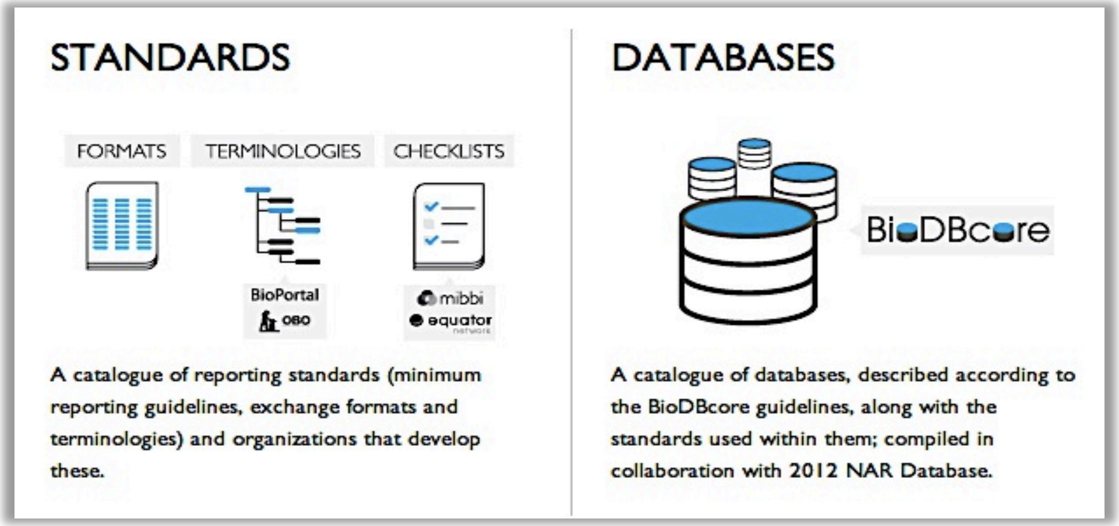
assay
test performed either on material taken from the subject or on the whole initial subject, which produce qualitative or quantitative measurements (data)

Spreadsheets & Tables
(ISAtab metadata)



Facts as RDF statements
(information for machines)

biosharing.org



STANDARDS

A catalogue of reporting standards (minimum reporting guidelines, exchange formats and terminologies) and organizations that develop these.

DATABASES

A catalogue of databases, described according to the BioDBcore guidelines, along with the standards used within them; compiled in collaboration with 2012 NAR Database.

SEVENTH FRAMEWORK PROGRAMME

BBSRC bioscience for the future

NATURAL ENVIRONMENT RESEARCH COUNCIL

UNIVERSITY OF OXFORD
OXFORD e-Research CENTRE
core organization in the
elixir Data for Life
UK Node

FDA Food and Drug Administration NCTR National Center for Toxicological Research

HSCI HARVARD STEM CELL INSTITUTE

HARVARD School of Public Health

EMBL-EBI

HARVARD MEDICAL SCHOOL

DUTCH TECHCENTRE FOR LIFE SCIENCES

Janssen Research & Development, LLC

Argonne NATIONAL LABORATORY

cgfb BIOINFORMATIQUE

CSIRO

Bioplatforms Australia

NATURAL ENVIRONMENT RESEARCH COUNCIL

nbic

isacommons
isacommons.org

The Novartis Institutes for BioMedical Research

cnrs

UNIVERSITY OF OXFORD

BiGCaT department of bioinformatics

MARINE BIOLOGICAL LABORATORY WOODS HOLE

SysMO Systems Biology of Microorganisms

INRA

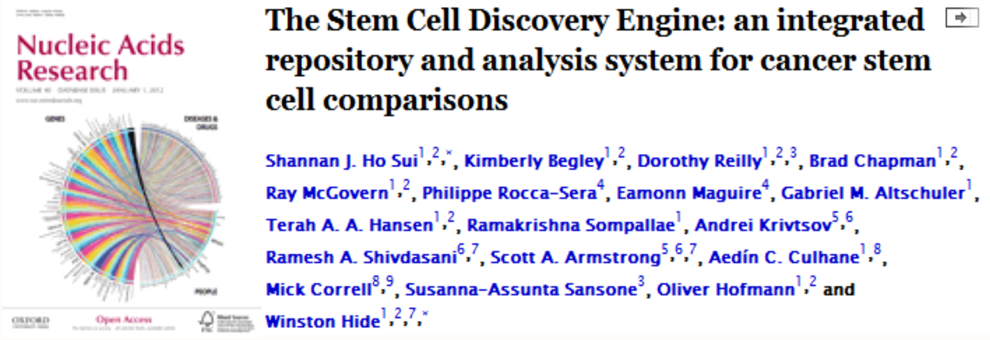
inist

$(GIGA)^n$ SCIENCE

Implementation at Harvard

HSCI
HARVARD STEM CELL INSTITUTE

SCDE STEM CELL DISCOVERY ENGINE


Nucleic Acids Research

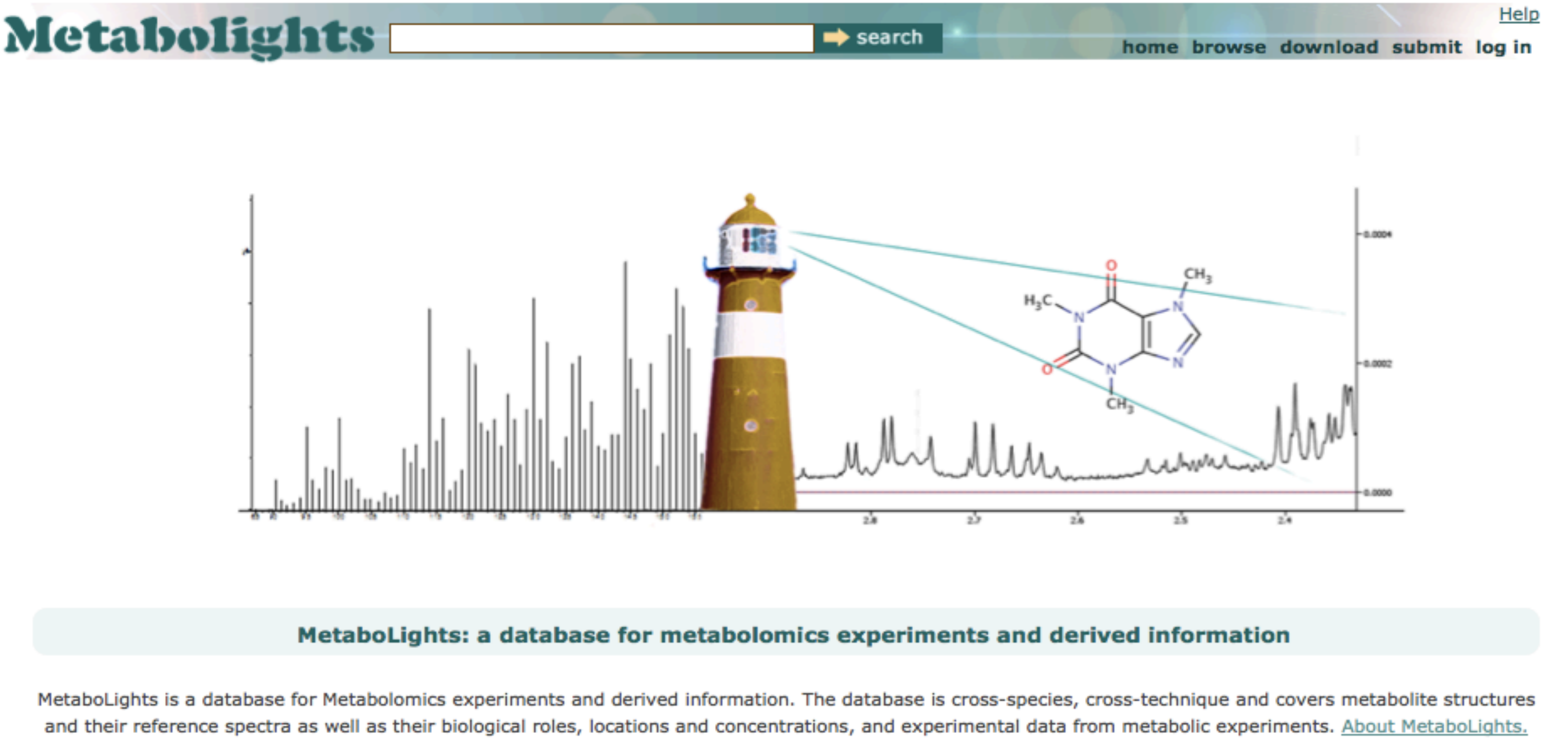
The Stem Cell Discovery Engine: an integrated repository and analysis system for cancer stem cell comparisons

Shannan J. Ho Sui[1,2,*], Kimberly Begley[1,2], Dorothy Reilly[1,2,3], Brad Chapman[1,2], Ray McGovern[1,2], Philippe Rocca-Sera[4], Eamonn Maguire[4], Gabriel M. Altschuler[1], Terah A. A. Hansen[1,2], Ramakrishna Sompallae[1], Andrei Krivtsov[5,6], Ramesh A. Shivdasani[6,7], Scott A. Armstrong[5,6,7], Aedín C. Culhane[1,8], Mick Correll[8,9], Susanna-Assunta Sansone[3], Oliver Hofmann[1,2] and Winston Hide[1,2,7,*]

http://discovery.hsci.harvard.edu/

Implementation at the EBI

http://www.ebi.ac.uk/metabolights



MetaboLights: a database for metabolomics experiments and derived information

MetaboLights is a database for Metabolomics experiments and derived information. The database is cross-species, cross-technique and covers metabolite structures and their reference spectra as well as their biological roles, locations and concentrations, and experimental data from metabolic experiments. About MetaboLights.

Metabolights – an open access general-purpose repository for metabolomics studies and associated meta-data
Haug et al, 2012
Nucleic Acids Research

The isa ecosystem

isatab

isatools
isa-tools.org

isacommons
isacommons.org

Risa

OntoMaton

isatoowl

github
SOCIAL CODING
GITHUB.COM/ISA-TOOLS

coreisateam


Susanna **Sansone**
Team Leader
University of Oxford


Philippe **Rocca-Serra**
Technical Co-ordinator
University of Oxford

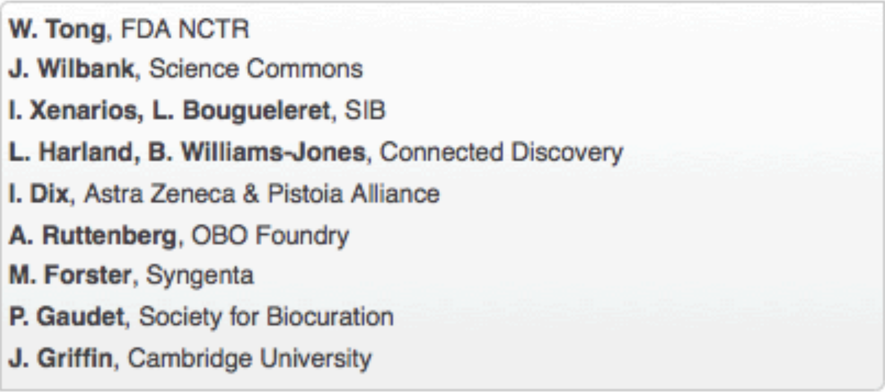

Eamonn **Maguire**
Lead Software Engineer
University of Oxford


Alejandra **Gonzalez-Beltran**
Software Engineer
University of Oxford

@isatools @biosharing

Isa-tools.org isacommons.org biosharing.org

advisory**board**


W. Tong, FDA NCTR
J. Wilbank, Science Commons
I. Xenarios, L. Bougueleret, SIB
L. Harland, B. Williams-Jones, Connected Discovery
I. Dix, Astra Zeneca & Pistoia Alliance
A. Ruttenberg, OBO Foundry
M. Forster, Syngenta
P. Gaudet, Society for Biocuration
J. Griffin, Cambridge University

funding

2011-2014, the ISA project has received another new funding stream from BBSRC and NERC.

2010-2013, the ISA project has received new funding stream from BBSRC and NERC.

2007-2010, the ISA project has been mainly supported by funds from the EU IP CarcinoGenomics, with contribution from EU NoE NuGO, NERC-NEBC, BBSRC and EMBL-EBI.