Supporting an Integrated Data Analysis across SEURAT-1 through the ToxBank Data Warehouse

OpenTox USA 2013 Meeting

Hamner Conference Center,

Research Triangle Park,

North Carolina, USA

29th October 2013

This project is jointly funded by Cosmetics Europe and the European Commission. Any opinions expressed in these slides are those of the authors. Cosmetics Europe is not liable for any use that may be made of the information contained therein.

ToxBank

Cosmetics Europe
the personal care association

SEVENTH FRAMEWORK PROGRAMME

Topics

- Background to SEURAT-1 and ToxBank projects
- Protocol and data warehousing
- Integrated data analysis
- Worked example using public data
- Summary

ToxBank

Cosmetics Europe
the personal care association

SEVENTH FRAMEWORK PROGRAMME

Background

- Legislation: The EU "Cosmetics Directive" 2013 deadline for ….

  …. animal testing of cosmetic products in the fields of repeated dose toxicity, reproductive toxicity and toxicokinetics.

- To overcome the lack of scientific knowledge for implementation of alternative testing solutions ….the Health Programme of DG Research and Innovation defined a long-term target: **Safety Evaluation Ultimately Replacing Animal Testing (SEURAT)** …. which will have an impact on many different areas including drug development, industrial chemicals, biocides etc….

SEURAT-1 objectives

Development of an **innovative concept for repeated dose systemic toxicity testing**.

Proof of concept for a future full implementation of a **mode-of-action** strategy.

Development of **innovative testing methods** more predictive than existing testing procedures.

The Building Blocks of SEURAT-1

ScrTox — Stem cell differentiation for providing human-based organ specific target cells

HeMiBio — Development of a hepatic microfluidic bioreactor

DETECTIVE (Detection of Endpoints and Biomarkers for Repeated Dose Toxicity Using In Vitro Systems) — Identification and investigation of human biomarkers

COSMOS — Delivery of computational tools to predict the effects of chemicals based on *in silico* calculations and estimation techniques

NOTOX — Development of systems biological tools for organotypic human cell cultures

ToxBank — Supporting integrated data analysis and servicing of alternative testing methods in toxicology

COACH — Cluster level Coordinating and Support Action

ToxBank

Cosmetics Europe — the personal care association

SEVENTH FRAMEWORK PROGRAMME

ToxBank

Establishment of a …

… cell and tissue banking information resource

… repository for the selected test compounds

… database of reference test compounds

… dedicated web-based data warehouse

MOA anchored 'Gold' compounds

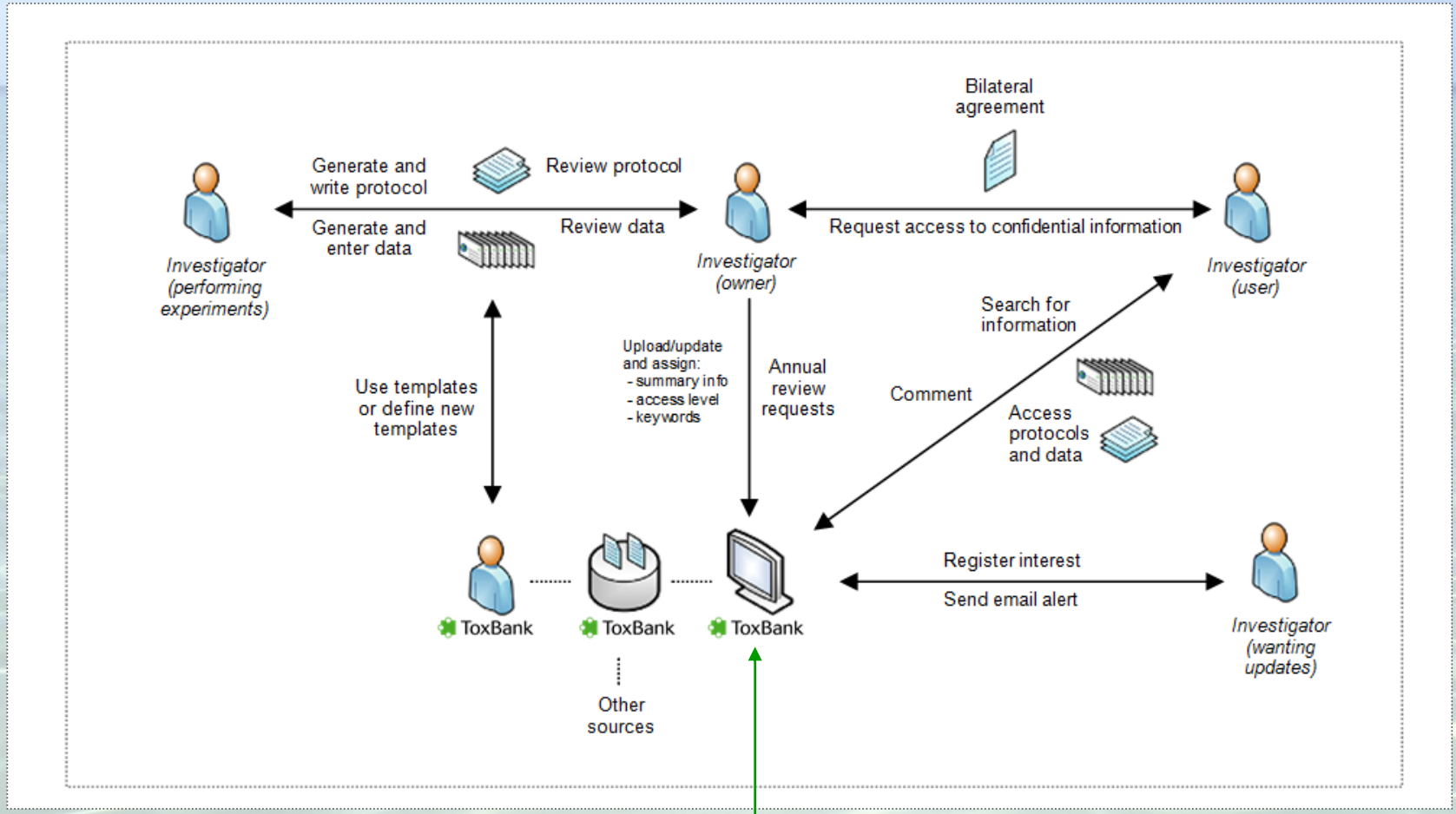
- Compounds are selected based on MOAs that are demonstrably relevant to human toxicity
- All SEURAT-1 partners will use this common set of compounds in their experiments
- Data on compounds is made available through a wiki (wiki.toxbank.net)

Compound	Target organ	MOA	Adverse event
Acetaminophen CAS # 103-90-2	Liver	Thiol reagent, oxidizing agent	Necrosis
Doxorubicin CAS # 23214-92-8	Heart	Redox cycling, DNA oxidation	Cellular lesions leading to heart failure
Allyl alcohol CAS # 107-18-6	Liver	Thiol reagent	Fibrosis
Carbon tetrachloride CAS # 56-23-5	Liver	Free radical	Fibrosis, steatosis
Aflatoxin B1 CAS # 1162-65-8	Liver	Lysine reagent	Apoptosis
Chlorpromazine CAS # 50-53-3	Liver	Thiol reagent, oxidizing agent, free radical, lipid binding, ATP synthase inhibition	Cholestasis, hepatitis
Iodoacetamide CAS # 144-48-9	All	Thiol reagent	(MOA standard)
DMNQ CAS # 6956-96-3	All	Redox cycling	(MOA standard)
Sodium valproate CAS # 99-66-1	Liver	Inhibition of multiple pathways, including β-oxidation	Steatosis, necrosis
Amiodarone CAS # 1951-25-3	Liver	Phospholipid binding	Steatosis, necrosis, phospholipidosis
E 4031 CAS # 113558-89-7	Heart	hERG channel blocker	Arrhythmias
Rotenone CAS # 83-79-4	All	Complex I (electron transport)	(MOA standard)
Oligomycin CAS # 1404-19-9	All	ATP synthase inhibitor	(MOA standard)

Compound	Target organ	MOA	Adverse event
FCCP CAS # 370-86-5	All	Proton gradient uncoupler	(MOA standard)
Bosentan CAS # 147536-97-8	Liver	BESP inhibition	Cholestasis
Dirlotapide CAS # 481658-94-0	Liver	MTTP inhibition	Steatosis
Fluoxetine CAS # 54910-89-3	Liver	Phospholipid binding	Phospholipidosis
Methotrexate CAS # 59-05-2	All	Antimetabolite	Hepatic fibrosis
Carbachol CAS # 51-83-2	Heart	Cholinergic agonist	(used for cell line characterization)
(-)Isoproterenol CAS # 7683-59-2	Heart	Adrenergic agonist	(used for cell line characterization)
Nifedipine CAS # 21829-25-4	Heart	L-type Ca channel blocker	(used for cell line characterization)
Hygromycin B CAS # 31282-04-9	All	Protein synthesis inhibitor	(standard for electron microscopy)
Tamoxifen CAS # 10540-29-1	Liver	Promiscuous ligand	Steatosis, cholestasis, epigenetics
TO901317 CAS # 293754-55-9	Liver	LXR and PXR agonist	Steatosis
Potassium Bromate CAS # 7758-01-2	Renal	Oxidative damage	Nephrotoxicity and Ototoxicity
Ochratoxin A CAS # 303-47-9	Renal	Non-genotoxic carcinogen	Renal carcinogenicity and nephrotoxicity

Outline of the ToxBank Data Warehouse

Phase 1: Unified data access



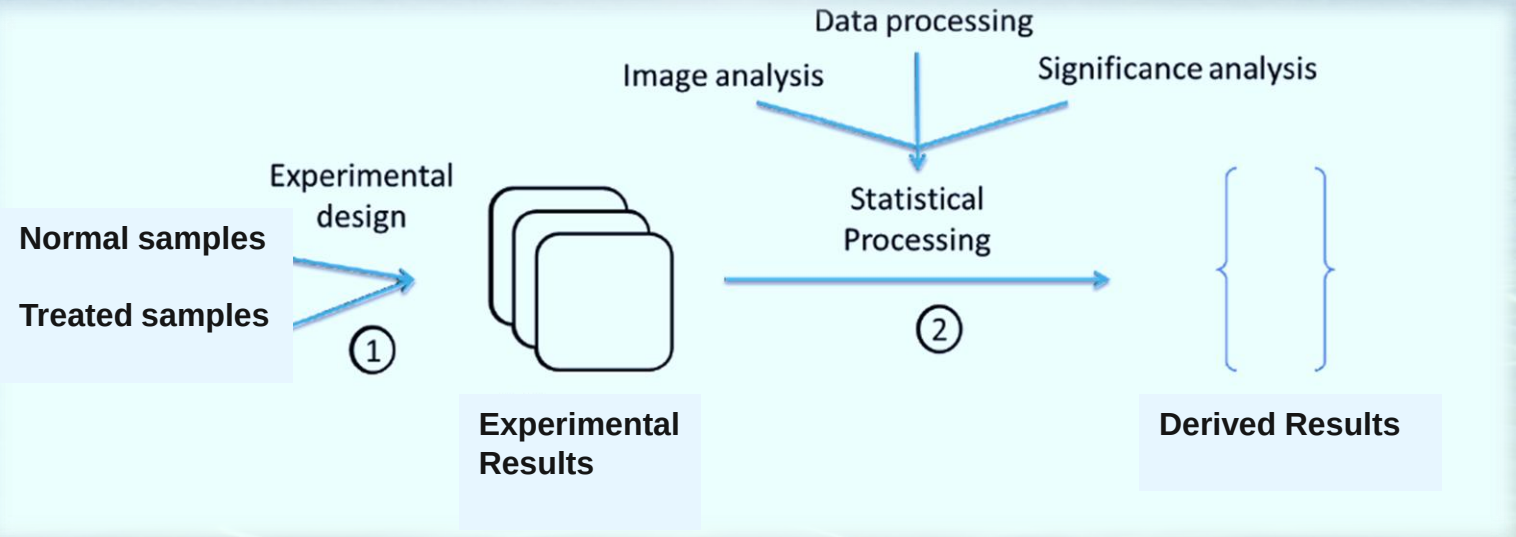
Phase 2: Integrated data analysis

ToxBank

Cosmetics Europe

SEVENTH FRAMEWORK PROGRAMME

The use of ISA-TAB Universal data exchange format



- ✓ the **investigation**: hypothesis, people & affiliations, timeline, publication
- ✓ the **experiment**: materials, methods and results
- ✓ the **materials**: subjects, samples, probes, equipment and software
- ✓ the **methods**: sample procurement and processing, measurement of gene expression, data processing and statistical testing
- ✓ the **results**: experimental data, normalized values, differential expression, significance, the list of differentially expressed genes

ToxBank

Cosmetics Europe

SEVENTH FRAMEWORK PROGRAMME

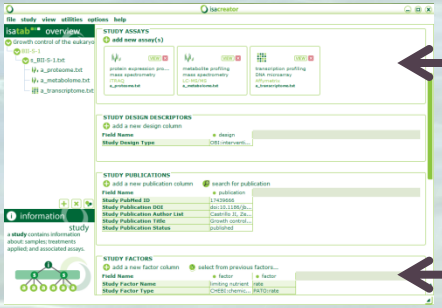
Use of SEURAT-configured ISAcreator to prepare datasets

Investigation information

SEURAT-1 information

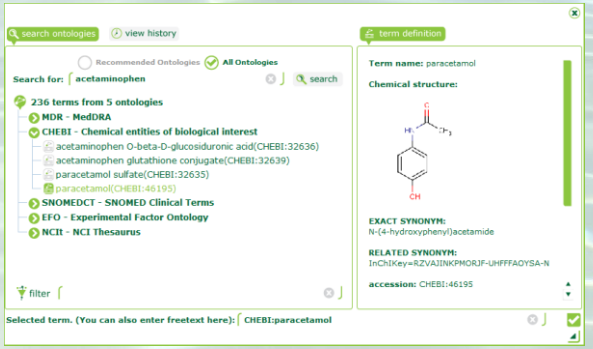
Publications

Each step linked to a SEURAT-1 protocol

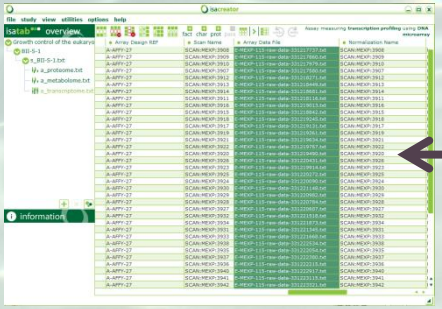


Templates for different assays

Specify experimental factors



Terms mapped to ontologies



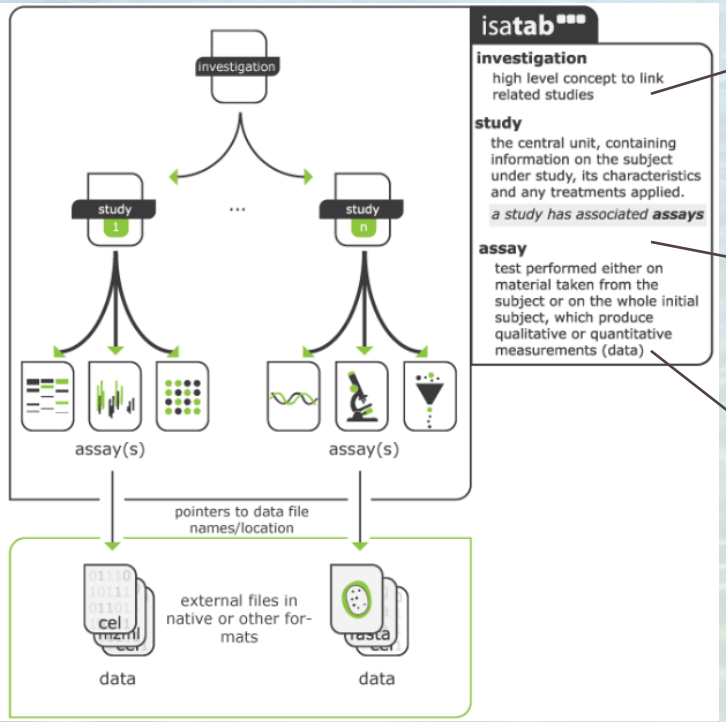
Materials and results, with links to files containing the raw or processed data

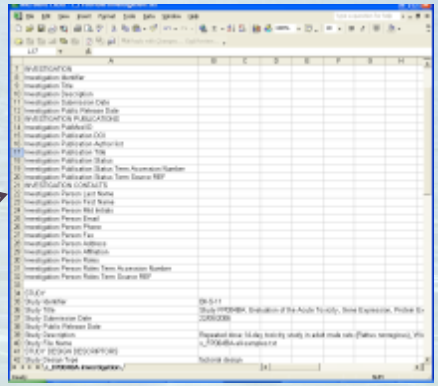
ToxBank

Cosmetics Europe

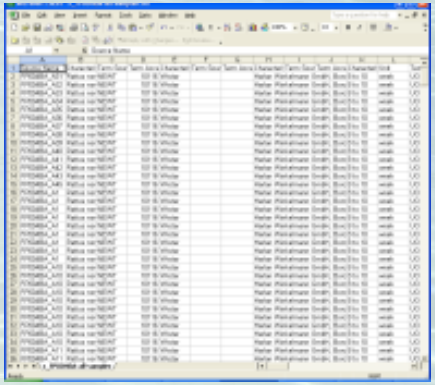
SEVENTH FRAMEWORK PROGRAMME

Create an ISA-tab zip archive for each investigation

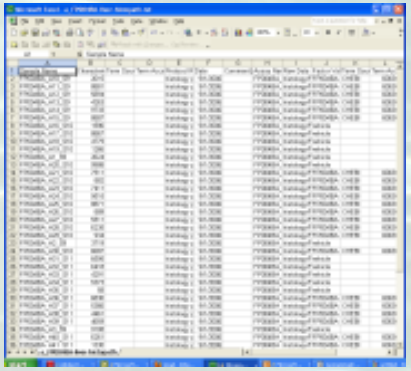




Overall investigation design and information (i… files)



Study description (s… files)



Test results (a… files) with links to data table or native file (e.g. CEL files)

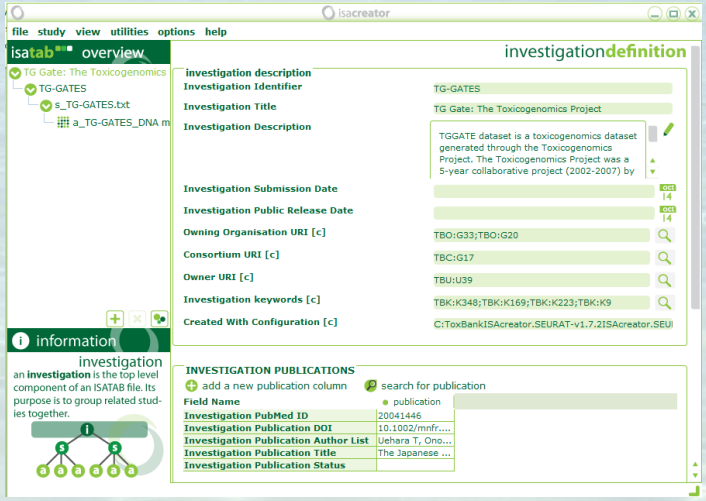
ToxBank

Cosmetics Europe

SEVENTH FRAMEWORK PROGRAMME

Generating the ISA-tab (TG-GATES* example)

Meta information on the study



Sample description and study factors

Field Name	row	row
Source Name	Hepatocyte_medium	Hepatocyte_medium
Characteristics[organism]	NEWT:Homo sapiens (Human)	NEWT:Homo sapiens (Human)
Characteristics[cell]	OBI:hepatocyte	OBI:hepatocyte
Characteristics[Technical Replicate]	2	1
Factor Value[compound]	CHEBI:DOXORUBICIN	CHEBI:DOXORUBICIN
StdInChIKey [c]		
Characteristics[control]	Negative	Negative
Factor Value[dose]	0	0
Unit	UO:micromolar	UO:micromolar
Factor Value[sample TimePoint]	8	24
Characteristics[sample TimePointU...	UO:hour	UO:hour
Protocol REF		
Sample Name	TGiv_DOX_Control_8hr_2	TGiv_DOX_Control_24hr_1

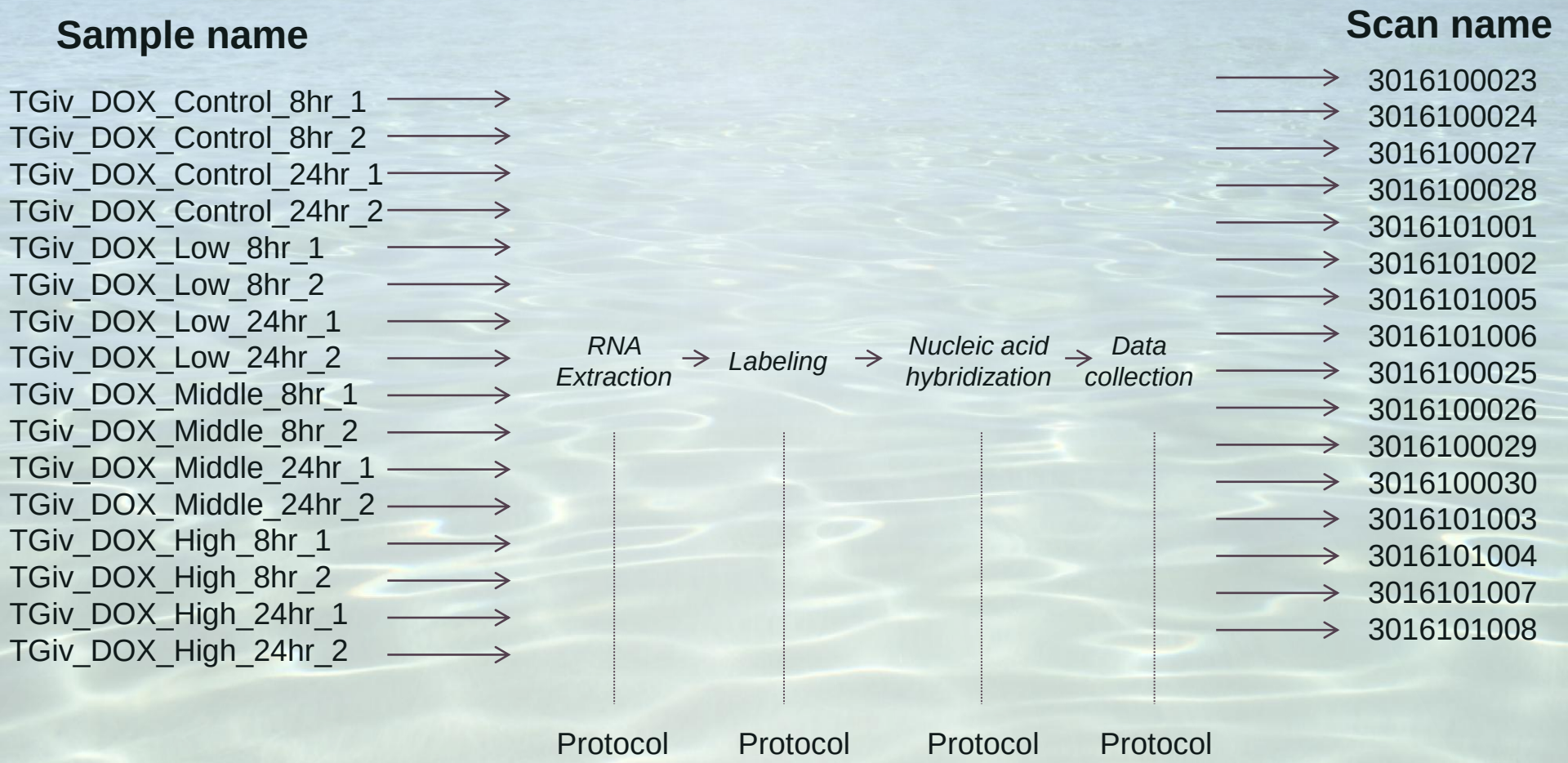
ToxBank

* 'Toxicogenomics Project and Toxicogenomics Informatics'

Cosmetics Europe

SEVENTH FRAMEWORK PROGRAMME

Generating the ISA-tab (TG-GATES example)



ToxBank

Cosmetics Europe

SEVENTH FRAMEWORK PROGRAMME

Generating the ISA-tab (TG-GATES example)

Scan name

3016100023
3016100024
3016100027
3016100028
3016101001
3016101002
3016101005
3016101006
3016100025
3016100026
3016100029
3016100030
3016101003
3016101004
3016101007
3016101008

normalization data transformation

data transformation

Data transformation name

LC8hr
MC8hr
HC8hr
ML8hr
HL8hr
HM8hr
LC24hr
MC24hr
HC24hr
ML24hr
HL24hr
HM24hr
LC8hr24hr
MC8hr24hr
HC8hr24hr
HL8hr24hr
ML8hr24hr
HM8hr24hr
CC8hr24hr

Processed data file

Ensembl	Entrez	Symbol	Log-average expression	FC'HC8hr'	FC'MC8hr'	FC'LC8hr'
ENSG00000000003	7105	TSPAN6	10.52	0.021	-0.112	0.005
ENSG00000000005	64102	TNMD	4.04	0.21	0.066	0.214
ENSG00000000419	8813	DPM1	12.31	0.168	0.316	0.184
ENSG00000000457	57147	SCYL3	7.19	-1.049	-0.206	0.101
ENSG00000000460	55732	C1orf112	5.26	-0.402	-0.497	-0.143
ENSG00000000938	2268	FGR	5.77	0.157	0.299	-0.026
ENSG00000000971	3075	CFH	10.1	0.571	0.232	0.035
ENSG00000001036	2519	FUCA2	10.46	0.036	-0.05	-0.041
ENSG00000001084	2729	GCLC	9.22	-0.377	-0.153	0.105
ENSG00000001167	4800	NFYA	6.88	-1.052	-0.966	-0.214
ENSG00000001460	90529	STPG1	6.42	0.046	0.025	0.005
ENSG00000001461	57185	NIPAL3	6.88	-0.048	0.223	0.056
ENSG00000001497	81887	LAS1L	8.9	0.303	0.129	-0.012
ENSG00000001561	22875	ENPP4	7.24	-0.059	-0.391	0.008

Protocol

Protocol

.cel files

Normalized data file

ToxBank

Cosmetics Europe

SEVENTH FRAMEWORK PROGRAMME

Uploading protocols and data

Main screen

Upload and set email alerts

Upload protocols

Upload data

Prepare datasets with ISAcreator

ToxBank

Cosmetics Europe

SEVENTH FRAMEWORK PROGRAMME

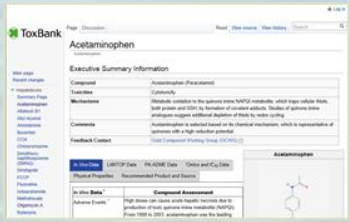
Searching and browsing

Main screen

Browse search results

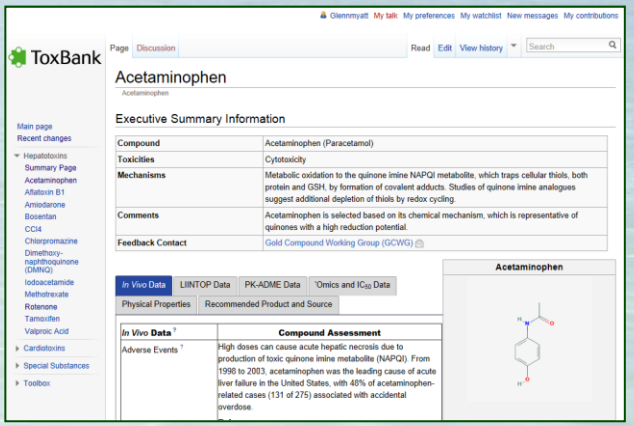
Download protocols and data

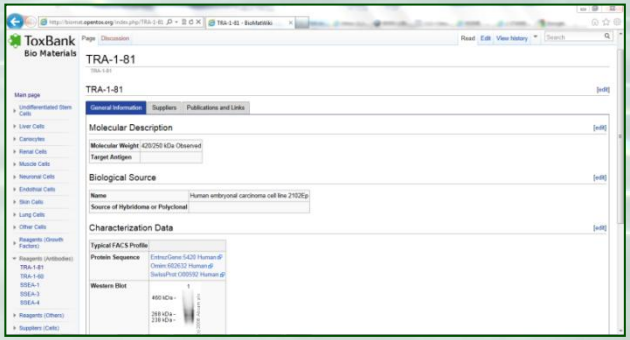
Access related information



Information resources

- **Gold compound wiki**
  - Information on selection criteria
  - In vivo, PB-PK data, ‘omics/IC50, physical data and sources
- **Biomaterials wiki**
  - Information on cells (stem cells, hES/iPS-derived cells, primary cells), reagents (e.g. antibodies, growth factors) and suppliers





wiki.toxbank.net

ToxBank

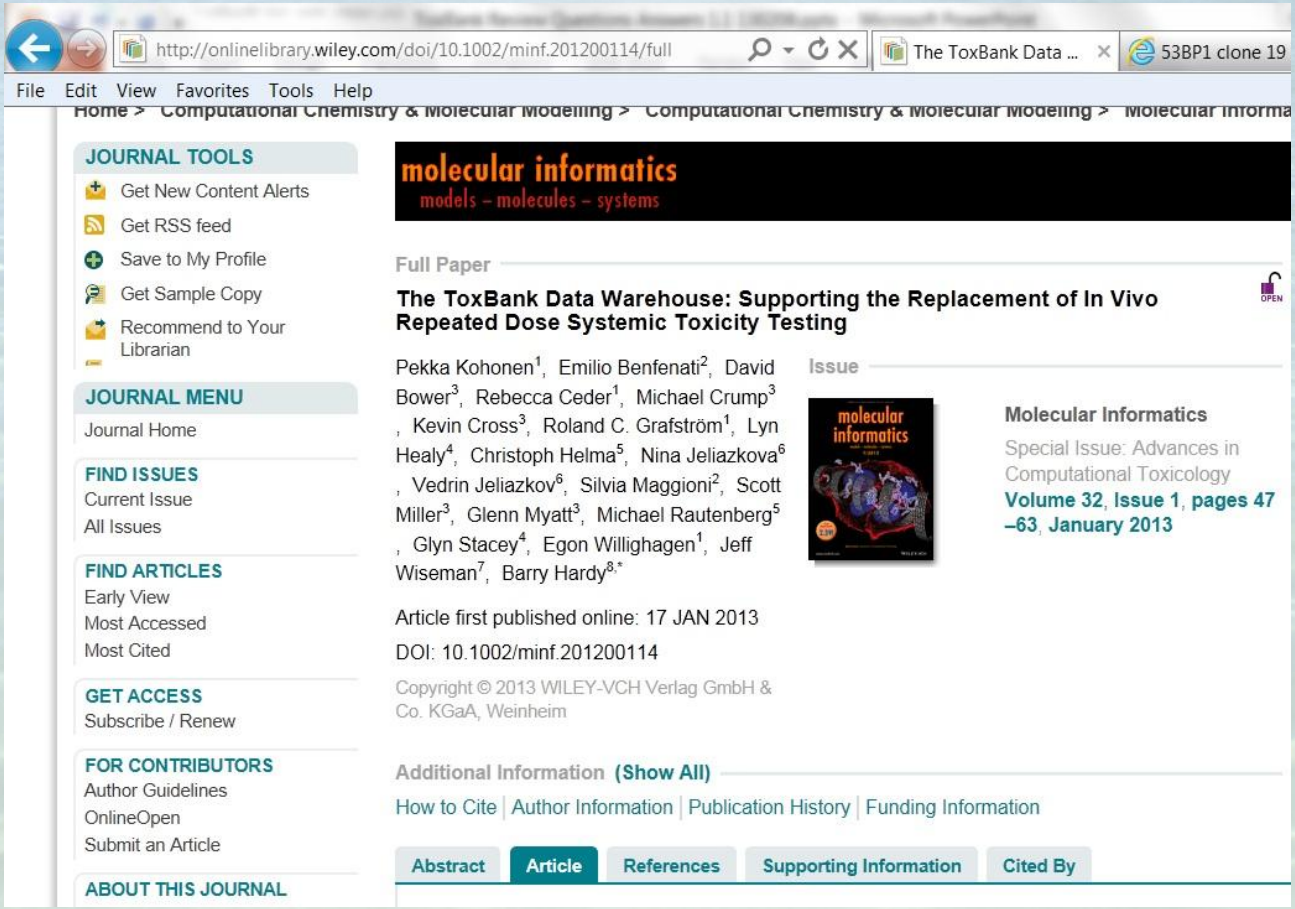
Cosmetics Europe
the personal care association

SEVENTH FRAMEWORK PROGRAMME

ToxBank technologies

- **ToxBank adopts the OpenTox framework design:**
  - Representational State Transfer (REST) software architecture style allowing platform and programming language independence and facilitating the implementation of new data and processing components
  - Formally defined common information model, based on the W3C Resource Description Framework (RDF)and communication through well-defined interfaces ensuring interoperability of the web components
  - 4store triple store as a backend for the investigation service
  - Authentication and authorization, allowing defining access policies of REST resources, based on OpenAM

ToxBank Phase I – Unified data access



onlinelibrary.wiley.com/doi/10.1002/minf.201200114/full

ToxBank

Cosmetics Europe

SEVENTH FRAMEWORK PROGRAMME

ToxBank phase II: Integrated data analysis

- **Use cases**
  - Supporting research questions, understanding biological context, assessing safety through read across (including using omics data), development of test battery, …
- **Queries to support hypotheses and integrated analysis**
  - Significant up or down regulated genes, proteins, …
  - Cells, metabolites and pathways
  - Chemical structure searching (exact, substructure and similarity)
- **Dashboard to explore multiple investigations**
  - Understand both the experimental factors, parameters and technologies used in producing the data across experiments
  - Export raw or standardized processed data to data analysis and bioinformatics/chemoinformatics tools

ToxBank Phase II – Integrated Data Analysis

Search

Integrate experimental data across consortia

SciToxHeMiBio
COSMOS NOTOX
ToxBank DETECTIVE

Protocols

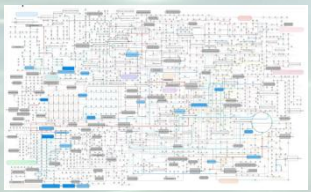
Investigation data

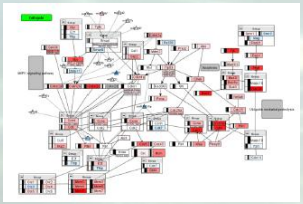
Integrate with public data (clinical, Tox21, …)

COSMOS
Tox21
ToxBank

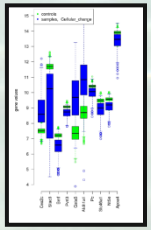
DASHBOARD*

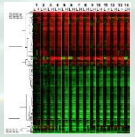
Understand/export multiple experiments

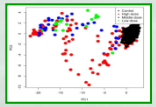




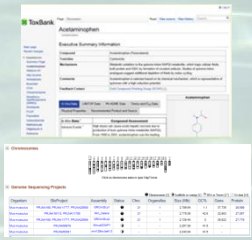
Pathway integration across projects, omics platforms, and experiments







Analysis and visualization



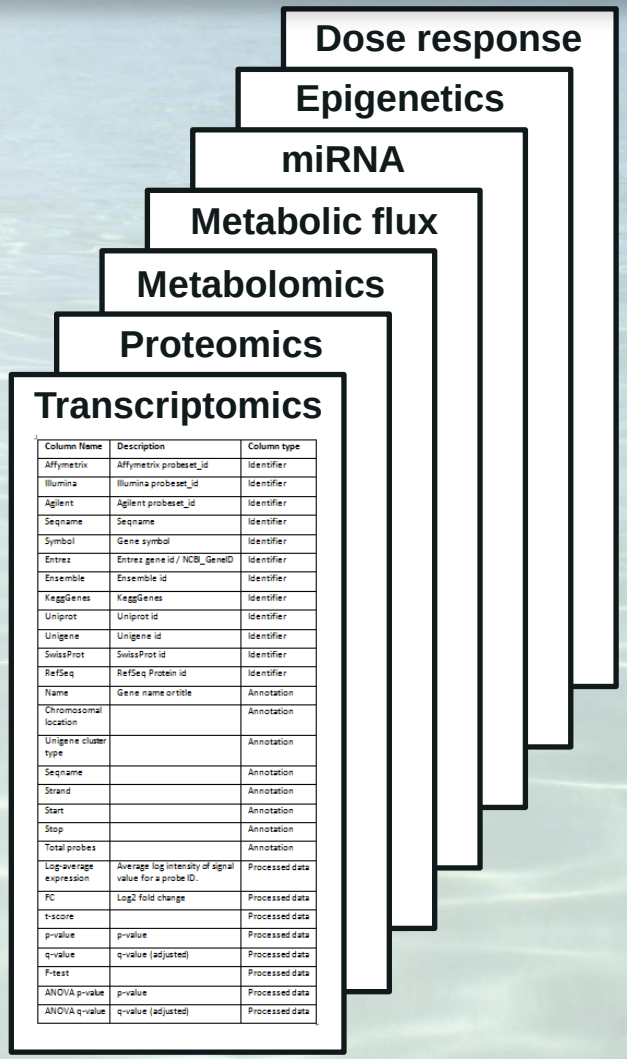
Links to other resources

ToxBank * Under development

Cosmetics Europe the personal care association
SEVENTH FRAMEWORK PROGRAMME

Standardization of processed data

- To support ToxBank integrated data analysis objectives (precise searching, meta analysis, …)
- The columns will
  - (1) uniquely identify the *material* (e.g. the Affymetrix probeset_id),
  - (2) annotate the *material* (e.g. the name of the gene),
  - (3) describe the processed results (e.g. fold change comparing genes expressed in the treated sample to the control).



Column Name	Description	Column type
Affymetrix	Affymetrix probeset_id	Identifier
Illumina	Illumina probeset_id	Identifier
Agilent	Agilent probeset_id	Identifier
Seqname	Seqname	Identifier
Symbol	Gene symbol	Identifier
Entrez	Entrez gene id / NCBI_GeneID	Identifier
Ensemble	Ensemble id	Identifier
KeggGenes	KeggGenes	Identifier
Uniprot	Uniprot id	Identifier
Unigene	Unigene id	Identifier
SwissProt	SwissProt id	Identifier
RefSeq	RefSeq Protein id	Identifier
Name	Gene name ortitle	Annotation
Chromosomal location		Annotation
Unigene cluster type		Annotation
Seqname		Annotation
Strand		Annotation
Start		Annotation
Stop		Annotation
Total probes		Annotation
Log-average expression	Average log intensity of signal value for a probe ID.	Processed data
FC	Log2 fold change	Processed data
t-score		Processed data
p-value	p-value	Processed data
q-value	q-value (adjusted)	Processed data
F-test		Processed data
ANOVA p-value	p-value	Processed data
ANOVA q-value	q-value (adjusted)	Processed data

Doxorubicin

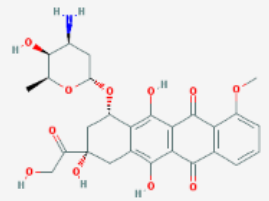
Executive Summary Information

Compound	Doxorubicin
Toxicities	Cytotoxicity
Mechanisms	Toxicity is initiated by oxidative damage associated both with the hydroquinone moiety and with iron-complexes of the parent compound. The major metabolic product is more toxic than the parent, but metabolism is not a requirement for toxicity. This compound intercalates with DNA and thus causes direct damage to DNA as well as to proteins. Toxicity is both acute and chronic and is life-threatening.
Comments	This compound was selected as an archetypical repeated dose cardiotoxin.
Feedback Contact	Gold Compound Working Group (GCWG)

In Vivo Data | PK-ADME Data | 'Omics and IC_{50} Data | Physical Properties | Recommended Product and Source

***In Vivo* Data** [?]	**Compound Assessment**
Adverse Events [?]	**Acute cardiotoxicity** Arrythmias during or within 24 hours of doxorubicin administration. Histopathological features of acute cardiotoxicity include increased hyaline material, contraction band necrosis and an infiltrate of neutrophils, lymphocytes and histiocytes. **Subacute cardiotoxicity** Myopericarditis days to weeks after administration. **Chronic cardiotoxicity**

Doxorubicin



Identifiers

http://wiki.toxbank.net/wiki/Doxorubicin

ISA-tab TG-GATES example

Scan name

3016100023
3016100024
3016100027
3016100028
3016101001
3016101002
3016101005
3016101006
3016100025
3016100026
3016100029
3016100030
3016101003
3016101004
3016101007
3016101008

normalization data transformation → *data transformation* →

Protocol

Protocol

Data transformation name

LC8hr
MC8hr
HC8hr
ML8hr
HL8hr
HM8hr
LC24hr
MC24hr
HC24hr
ML24hr
HL24hr
HM24hr
LC8hr24hr
MC8hr24hr
HC8hr24hr
HL8hr24hr
ML8hr24hr
HM8hr24hr
CC8hr24hr

Processed data file

Ensembl	Entrez	Symbol	Log-average expression	FC'HC8hr'	FC'MC8hr'	FC'LC8hr'
ENSG00000000003	7105	TSPAN6	10.52	0.021	-0.112	0.005
ENSG00000000005	64102	TNMD	4.04	0.21	0.066	0.214
ENSG00000000419	8813	DPM1	12.31	0.168	0.316	0.184
ENSG00000000457	57147	SCYL3	7.19	-1.049	-0.206	0.101
ENSG00000000460	55732	C1orf112	5.26	-0.402	-0.497	-0.143
ENSG00000000938	2268	FGR	5.77	0.157	0.299	-0.026
ENSG00000000971	3075	CFH	10.1	0.571	0.232	0.035
ENSG00000001036	2519	FUCA2	10.46	0.036	-0.05	-0.041
ENSG00000001084	2729	GCLC	9.22	-0.377	-0.153	0.105
ENSG00000001167	4800	NFYA	6.88	-1.052	-0.966	-0.214
ENSG00000001460	90529	STPG1	6.42	0.046	0.025	0.005
ENSG00000001461	57185	NIPAL3	6.88	-0.048	0.223	0.056
ENSG00000001497	81887	LAS1L	8.9	0.303	0.129	-0.012
ENSG00000001561	22875	ENPP4	7.24	-0.059	-0.391	0.008

ToxBank

.cel files **Normalized data file**

Cosmetics Europe

SEVENTH FRAMEWORK PROGRAMME

TG-GATES analysis example

Ensembl	Entrez	Symbol	Log-average expression	FC'HC8hr'	FC'MC8hr'	FC'LC8hr'
ENSG00000000003	7105	TSPAN6	10.52	0.021	-0.112	0.005
ENSG00000000005	64102	TNMD	4.04	0.21	0.066	0.214
ENSG00000000419	8813	DPM1	12.31	0.168	0.316	0.184
ENSG00000000457	57147	SCYL3	7.19	-1.049	-0.206	0.101
ENSG00000000460	55732	C1orf112	5.26	-0.402	-0.497	-0.143
ENSG00000000938	2268	FGR	5.77	0.157	0.299	-0.026
ENSG00000000971	3075	CFH	10.1	0.571	0.232	0.035
ENSG00000001036	2519	FUCA2	10.46	0.036	-0.05	-0.041
ENSG00000001084	2729	GCLC	9.22	-0.377	-0.153	0.105
ENSG00000001167	4800	NFYA	6.88	-1.052	-0.966	-0.214
ENSG00000001460	90529	STPG1	6.42	0.046	0.025	0.005
ENSG00000001461	57185	NIPAL3	6.88	-0.048	0.223	0.056
ENSG00000001497	81887	LAS1L	8.9	0.303	0.129	-0.012
ENSG00000001561	22875	ENPP4	7.24	-0.059	-0.391	0.008
ENSG00000001617	6405	SEMA3F	6.55	0.139	0.207	0.111

Pathway enrichment*

ID	Name	List ratio	BG ratio	P-value	Q-value	Genes/Compounds
path:hsa04668	TNF signaling pathway	16/799	86/14867	9.073E-6	1.778E-3	TNFRSF1A, TRAF1, FADD, NFKBIA, CREB1, CX3CL1, JUNB, MAPK14, BAG4, CCL2, CASP3, JUN, MAP3K5, CEBPB, FOS, CASP8
path:hsa05161	Hepatitis B	17/799	110/14867	5.492E-5	5.382E-3	IL8, FADD, NFKBIA, CREB1, SMAD4, TLR4, CCNA2, MYC, DDX58, CASP3, TGFB2, JUN, TBK1, TICAM1, EGR2, FOS, CASP8
path:hsa05164	Influenza A	16/799	112/14867	2.167E-4	0.0142	TNFRSF1A, IL8, NFKBIA, PLG, IVNS1ABP, EIF2AK3, TLR4, RSAD2, MAPK14, IL18, JAK2, CCL2, DDX58, JUN, TBK1, TICAM1
path:hsa04110	Cell cycle	15/799	105/14867	3.31E-4	0.0162	CDC20, CHEK1, CDKN2B, TTK, SMAD4, CDC7, CCNA2, MYC, ORC2, TGFB2, CDK1, MAD2L1, CDC6, ATR, CUL1
path:hsa05142	Chagas disease (Ameri...	12/799	77/14867	5.688E-4	0.0223	TNFRSF1A, CCL2, IL8, FADD, NFKBIA, TGFB2, JUN, TICAM1, TLR4, MAPK14, FOS, CASP8
path:hsa05168	Herpes simplex infection	16/799	126/14867	7.573E-4	0.0247	TNFRSF1A, TRAF1, FADD, NFKBIA, EIF2AK3, JAK2, CCL2, DDX58, CASP3, CDK1, JUN, TBK1, TICAM1, CUL1, FOS, CASP8
path:hsa05323	Rheumatoid arthritis	10/799	62/14867	1.194E-3	0.0334	CCL2, IL8, TGFB2, CXCL6, JUN, TNFSF11, TLR4, IL18, MMP1, FOS
path:hsa04620	Toll-like receptor signal...	11/799	76/14867	1.664E-3	0.0408	IL8, FADD, NFKBIA, TBK1, JUN, TICAM1, CXCL11, TLR4, MAPK14, FOS, CASP8

*InCroMAP software (http://www.ra.cs.uni-tuebingen.de/software/InCroMAP/)

TG-GATES analysis example

Ensembl	Entrez	Symbol	Log-average expression	FC'HC8hr'	FC'MC8hr'	FC'LC8hr'
ENSG00000000003	7105	TSPAN6	10.52	0.021	-0.112	0.005
ENSG00000000005	64102	TNMD	4.04	0.21	0.066	0.214
ENSG00000000419	8813	DPM1	12.31	0.168	0.316	0.184
ENSG00000000457	57147	SCYL3	7.19	-1.049	-0.206	0.101
ENSG00000000460	55732	C1orf112	5.26	-0.402	-0.497	-0.143
ENSG00000000938	2268	FGR	5.77	0.157	0.299	-0.026
ENSG00000000971	3075	CFH	10.1	0.571	0.232	0.035
ENSG00000001036	2519	FUCA2	10.46	0.036	-0.05	-0.041
ENSG00000001084	2729	GCLC	9.22	-0.377	-0.153	0.105
ENSG00000001167	4800	NFYA	6.88	-1.052	-0.966	-0.214
ENSG00000001460	90529	STPG1	6.42	0.046	0.025	0.005
ENSG00000001461	57185	NIPAL3	6.88	-0.048	0.223	0.056
ENSG00000001497	81887	LAS1L	8.9	0.303	0.129	-0.012
ENSG00000001561	22875	ENPP4	7.24	-0.059	-0.391	0.008

Pathway enrichment summarization

Pathway class	Pathways	FC'LC8hr'	FC'MC8hr'	FC'HC8hr'	FC'ML8hr'	FC'HL8hr'	FC'HM8hr'	FC'LC24hr'	FC'MC24hr'	FC'HC24hr'	FC'ML24hr'	FC'HL24hr'	FC'HM24hr'	FC'LC8hr24hr'	FC'MC8hr24hr'	FC'HC8hr24hr'	FC'HL8hr24hr'	FC'ML8hr24hr'	FC'HM8hr24hr'	FC'CC8hr24hr'
Cellular Processes; Cell growth and death	Cell cycle	*						*	*	*				*	*					
Cellular Processes; Cell growth and death	p53 signaling pathway								*											
Cellular Processes; Cell growth and death	Oocyte meiosis							*						*						
Environmental Information Processing; Signal transduction	TNF signaling pathway									*										
Genetic Information Processing; Replication and repair	DNA replication							*	*					*	*					
Genetic Information Processing; Replication and repair	Mismatch repair								*											
Genetic Information Processing; Replication and repair	Fanconi anemia pathway							*	*											
Human Diseases; Cancers	Viral carcinogenesis							*												
Human Diseases; Immune diseases	Rheumatoid arthritis									*										*
Human Diseases; Infectious diseases	Influenza A									*		*								
Human Diseases; Infectious diseases	Chagas disease (American trypanosomiasis)									*		*								
Human Diseases; Infectious diseases	Hepatitis B									*		*	*							
Human Diseases; Infectious diseases	Herpes simplex infection									*										
Metabolism; Nucleotide metabolism	Pyrimidine metabolism							*	*						*					
Organismal Systems; Endocrine system	Progesterone-mediated oocyte maturation													*						
Organismal Systems; Immune system	Toll-like receptor signaling pathway									*		*								

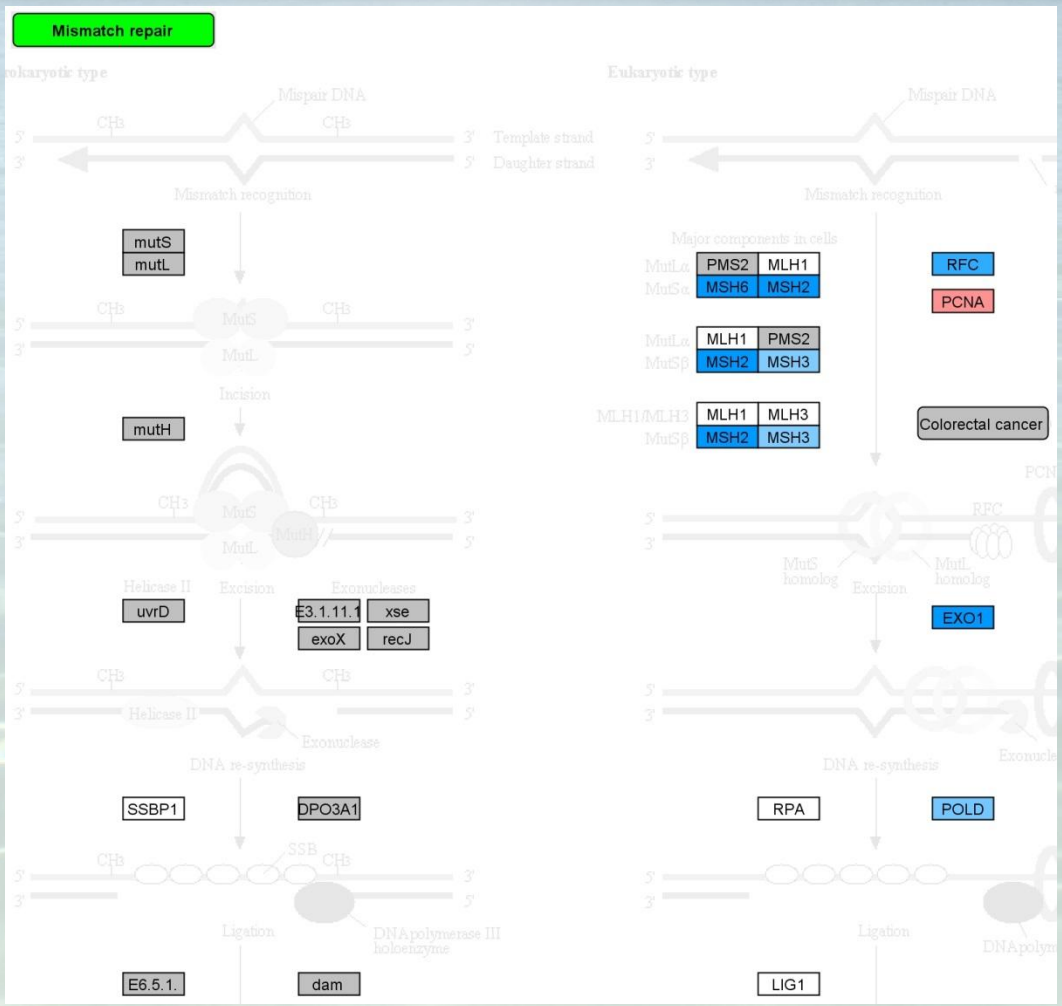
TG-GATES analysis example

Ensembl	Entrez	Symbol	Log-average expression	FC'HC8hr'	FC'MC8hr'	FC'LC8hr'
ENSG00000000003	7105	TSPAN6	10.52	0.021	-0.112	0.005
ENSG00000000005	64102	TNMD	4.04	0.21	0.066	0.214
ENSG00000000419	8813	DPM1	12.31	0.168	0.316	0.184
ENSG00000000457	57147	SCYL3	7.19	-1.049	-0.206	0.101
ENSG00000000460	55732	C1orf112	5.26	-0.402	-0.497	-0.143
ENSG00000000938	2268	FGR	5.77	0.157	0.299	-0.026
ENSG00000000971	3075	CFH	10.1	0.571	0.232	0.035
ENSG00000001036	2519	FUCA2	10.46	0.036	-0.05	-0.041
ENSG00000001084	2729	GCLC	9.22	-0.377	-0.153	0.105
ENSG00000001167	4800	NFYA	6.88	-1.052	-0.966	-0.214
ENSG00000001460	90529	STPG1	6.42	0.046	0.025	0.005
ENSG00000001461	57185	NIPAL3	6.88	-0.048	0.223	0.056
ENSG00000001497	81887	LAS1L	8.9	0.303	0.129	-0.012
ENSG00000001561	22875	ENPP4	7.24	-0.059	-0.391	0.008

Pathway enrichment summarization

Pathway class	Pathways	FC'LC8hr'	FC'MC8hr'	FC'HC8hr'	FC'ML8hr'	FC'HL8hr'	FC'HM8hr'	FC'LC24hr'	FC'MC24hr'	FC'HC24hr'	FC'ML24hr'	FC'HL24hr'	FC'HM24hr'	FC'LC8hr24hr'	FC'MC8hr24hr'	FC'HC8hr24hr'	FC'HL8hr24hr'	FC'ML8hr24hr'	FC'HM8hr24hr'	FC'CC8hr24hr'
Cellular Processes; Cell growth and death	Cell cycle	*						*	*	*				*	*					
Cellular Processes; Cell growth and death	p53 signaling pathway								*											
Cellular Processes; Cell growth and death	Oocyte meiosis							*						*						
Environmental Information Processing; Signal transduction	TNF signaling pathway									*										
Genetic Information Processing; Replication and repair	DNA replication							*	*					*	*					
Genetic Information Processing; Replication and repair	Mismatch repair								*											
Genetic Information Processing; Replication and repair	Fanconi anemia pathway							*	*											
Human Diseases; Cancers	Viral carcinogenesis							*												
Human Diseases; Immune diseases	Rheumatoid arthritis									*										*
Human Diseases; Infectious diseases	Influenza A									*		*								
Human Diseases; Infectious diseases	Chagas disease (American trypanosomiasis)									*		*								
Human Diseases; Infectious diseases	Hepatitis B									*		*	*							
Human Diseases; Infectious diseases	Herpes simplex infection									*										
Metabolism; Nucleotide metabolism	Pyrimidine metabolism							*	*						*					
Organismal Systems; Endocrine system	Progesterone-mediated oocyte maturation													*						
Organismal Systems; Immune system	Toll-like receptor signaling pathway									*		*								

TG-GATES analysis example



InCroMAP software (http://www.ra.cs.uni-tuebingen.de/software/InCroMAP/)

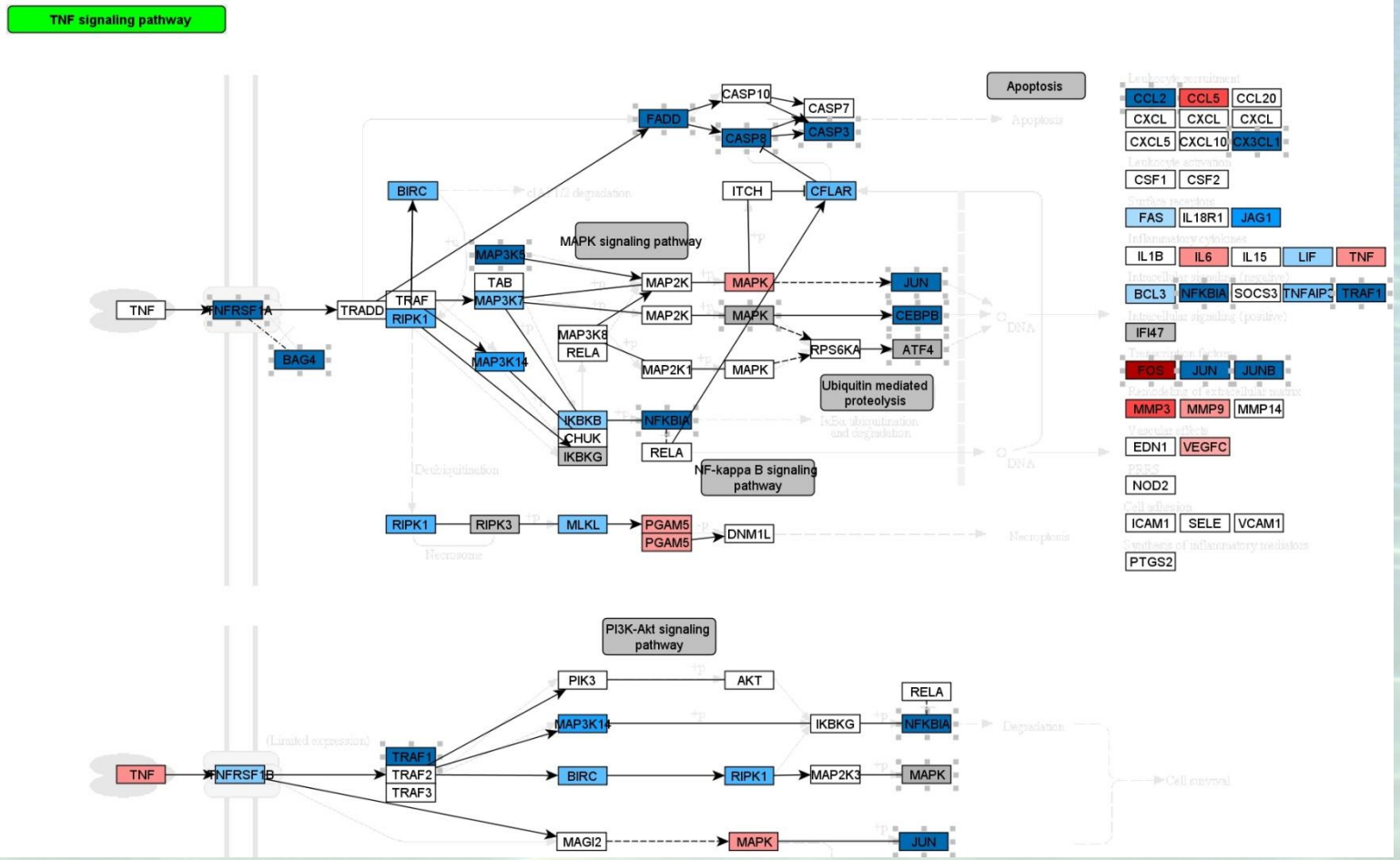
TG-GATES analysis example

Ensembl	Entrez	Symbol	Log-average expression	FC'HC8hr'	FC'MC8hr'	FC'LC8hr'
ENSG00000000003	7105	TSPAN6	10.52	0.021	-0.112	0.005
ENSG00000000005	64102	TNMD	4.04	0.21	0.066	0.214
ENSG00000000419	8813	DPM1	12.31	0.168	0.316	0.184
ENSG00000000457	57147	SCYL3	7.19	-1.049	-0.206	0.101
ENSG00000000460	55732	C1orf112	5.26	-0.402	-0.497	-0.143
ENSG00000000938	2268	FGR	5.77	0.157	0.299	-0.026
ENSG00000000971	3075	CFH	10.1	0.571	0.232	0.035
ENSG00000001036	2519	FUCA2	10.46	0.036	-0.05	-0.041
ENSG00000001084	2729	GCLC	9.22	-0.377	-0.153	0.105
ENSG00000001167	4800	NFYA	6.88	-1.052	-0.966	-0.214
ENSG00000001460	90529	STPG1	6.42	0.046	0.025	0.005
ENSG00000001461	57185	NIPAL3	6.88	-0.048	0.223	0.056
ENSG00000001497	81887	LAS1L	8.9	0.303	0.129	-0.012
ENSG00000001561	22875	ENPP4	7.24	-0.059	-0.391	0.008

Pathway enrichment summarization

Pathway class	Pathways	FC'LC8hr'	FC'MC8hr'	FC'HC8hr'	FC'ML8hr'	FC'HL8hr'	FC'HM8hr'	FC'LC24hr'	FC'MC24hr'	FC'HC24hr'	FC'ML24hr'	FC'HL24hr'	FC'HM24hr'	FC'LC8hr24hr'	FC'MC8hr24hr'	FC'HC8hr24hr'	FC'HL8hr24hr'	FC'ML8hr24hr'	FC'HM8hr24hr'	FC'CC8hr24hr'
Cellular Processes; Cell growth and death	Cell cycle	*						*	*	*				*	*					
Cellular Processes; Cell growth and death	p53 signaling pathway								*											
Cellular Processes; Cell growth and death	Oocyte meiosis							*						*						
Environmental Information Processing; Signal transduction	TNF signaling pathway									*										
Genetic Information Processing; Replication and repair	DNA replication							*	*					*	*					
Genetic Information Processing; Replication and repair	Mismatch repair								*											
Genetic Information Processing; Replication and repair	Fanconi anemia pathway							*	*											
Human Diseases; Cancers	Viral carcinogenesis							*												
Human Diseases; Immune diseases	Rheumatoid arthritis									*										*
Human Diseases; Infectious diseases	Influenza A									*		*								
Human Diseases; Infectious diseases	Chagas disease (American trypanosomiasis)									*		*								
Human Diseases; Infectious diseases	Hepatitis B									*		*	*							
Human Diseases; Infectious diseases	Herpes simplex infection									*										
Metabolism; Nucleotide metabolism	Pyrimidine metabolism							*	*						*					
Organismal Systems; Endocrine system	Progesterone-mediated oocyte maturation													*						
Organismal Systems; Immune system	Toll-like receptor signaling pathway									*		*								

TG-GATES analysis example

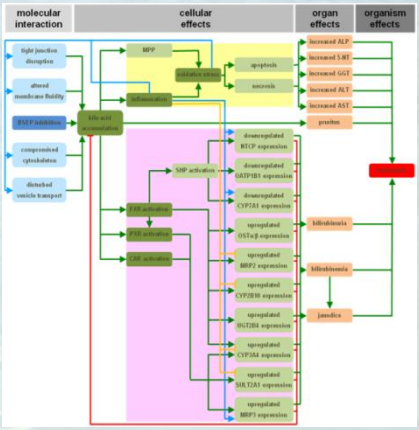


Analysis examples

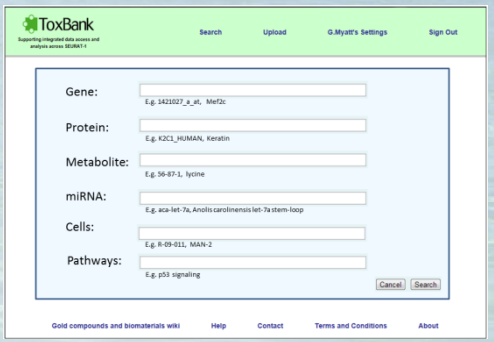
Multi-omics pathway enrichment

Pathway class	Pathways
Cellular Processes; Cell growth and death	Cell cycle
Cellular Processes; Cell growth and death	p53 signaling pathway
Cellular Processes; Cell growth and death	Oocyte meiosis
Environmental Information Processing; Signal transduction	TNF signaling pathway
Genetic Information Processing; Replication and repair	DNA replication
Genetic Information Processing; Replication and repair	Mismatch repair
Genetic Information Processing; Replication and repair	Fanconi anemia pathway
Human Diseases; Cancers	Viral carcinogenesis
Human Diseases; Immune diseases	Rheumatoid arthritis
Human Diseases; Infectious diseases	Influenza A
Human Diseases; Infectious diseases	Chagas disease (American trypanosomiasis)
Human Diseases; Infectious diseases	Hepatitis B
Human Diseases; Infectious diseases	Herpes simplex infection
Metabolism; Nucleotide metabolism	Pyrimidine metabolism
Organismal Systems; Endocrine system	Progesterone-mediated oocyte maturation
Organismal Systems; Immune system	Toll-like receptor signaling pathway

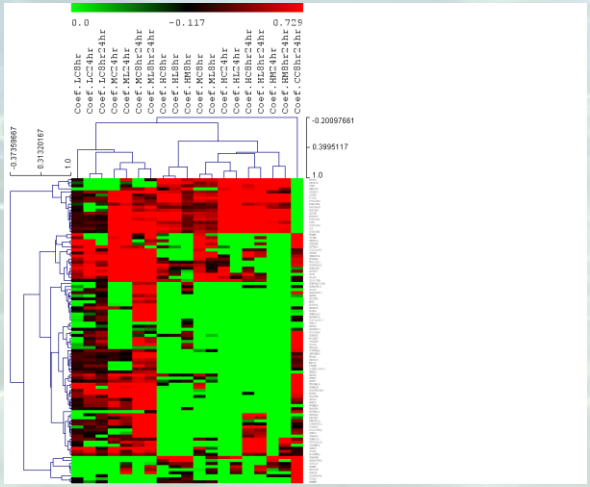
Development of AOPs



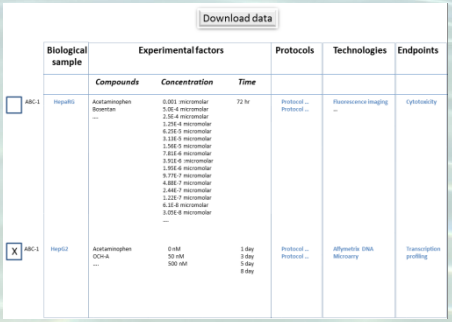
Search other investigations



Analysis and visualization



Understanding experiments



Understanding kinetics

PK-ADME [7]	Compound Assessment
PK parameters [7]	The most commonly used dose schedule when used as a single agent is 60 to 75 mg/m2 as a single intravenous injection administered at 21-day intervals.
	Protein binding 70%
	Half life 55 hours
	V_d 20-30 L/kg (700-100 L/m^2)
	C_{max} 3 uM for 30 mg/m^2 intravenous bolus dose. Cellular levels are about 30–100-fold higher than that of the plasma.
	Excretion predominantly in bile, 40-50% in feces within 7 days (50% as unchanged drug).
	Plasma clearance 324 to 809 mL/min/m^2, biphasic
	Metabolism ~50% metabolized by the liver
	References:
	- http://www.drugbank.ca/drugs/DB00997
	- http://reference.medscape.com/drug/doxorubicin-342120
	- A.-K. Souid et al. "Immediate effects of anticancer drugs on mitochondrial oxygen consumption", Biochemical Pharmacology 66 (2003) 977–987

ToxBank summary

- **Supporting the replacement of the repeated dose toxicity test**
  - Provides immediate access to existing and new protocols and data
    - Precisely documented protocols
    - The use of standardized templates and semantic annotation to ensure minimal information is collected in a consistent way
    - Store for legacy data
- **Technical/scientific integration with ToxCast and Tox21 data**
- **Enabling an integrated data analysis through**
  - Research hypothesis queries
  - Integration with pathways enrichment/mapping and data analysis/mining/visualization applications
  - Supporting safety assessment use cases

ToxBank Acknowledgements

DouglasConnect

in silico toxicology

Karolinska Institutet

Leadscope®

PHARMATROPE

UK StemCellBank

MARIO NEGRI ISTITUTO DI RICERCHE FARMACOLOGICHE

UK Stem Cell Bank, NIBSC-HPA

Ideaconsult Ltd

ToxBank

Cosmetics Europe

SEVENTH FRAMEWORK PROGRAMME