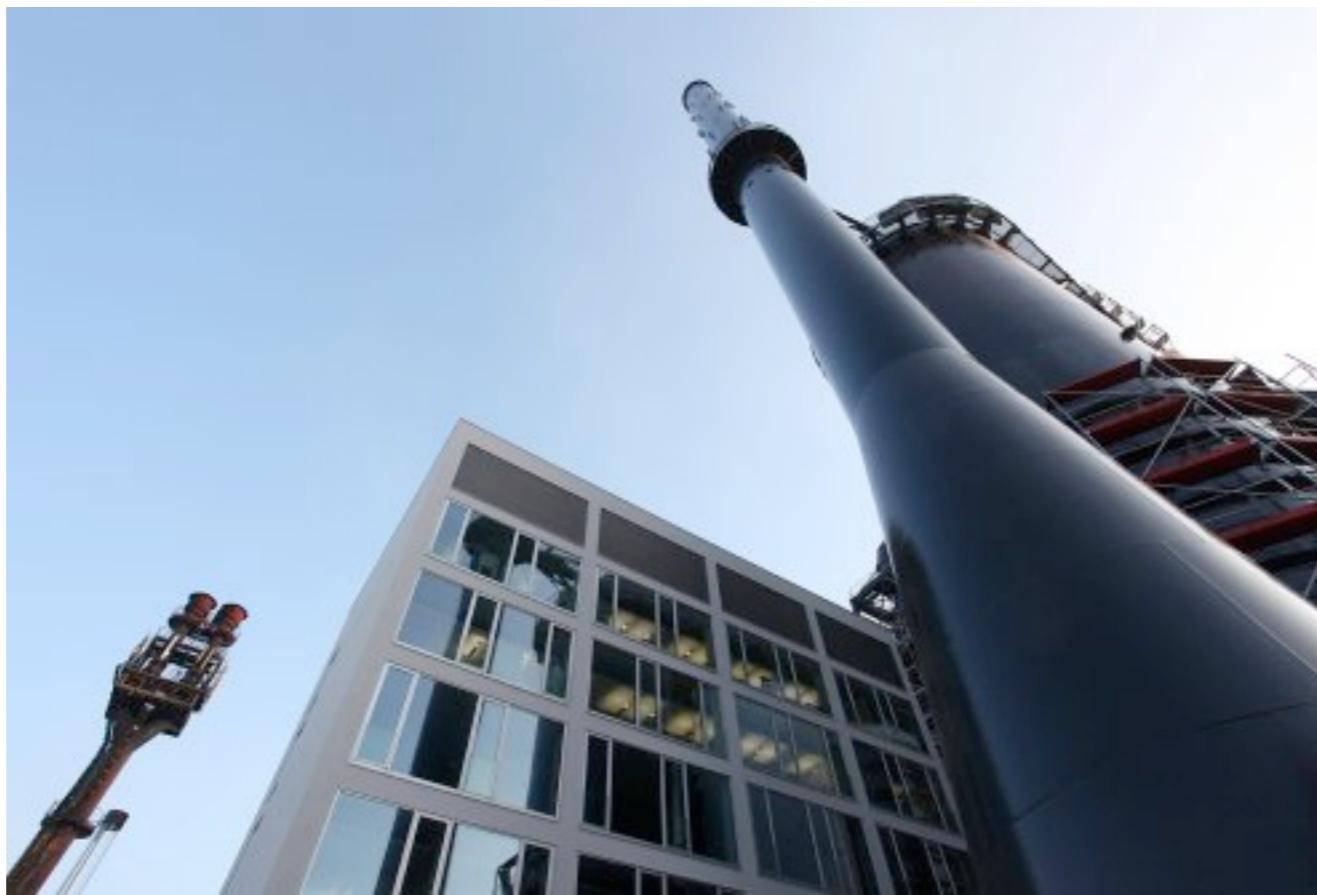


# Update about the BioInfoCore

Dr. Reinhard Schneider

**Luxembourg Centre for Systems  
Biomedicine (LCSB)**



# LCSB in a nutshell

part of  
the Biohealth  
Initiative of  
Luxembourg

Interdisciplinary  
research centre  
of the University  
of Luxembourg

Biomedical  
and  
Systems Biology  
Research

Aim:  
*Personalized  
medicine*

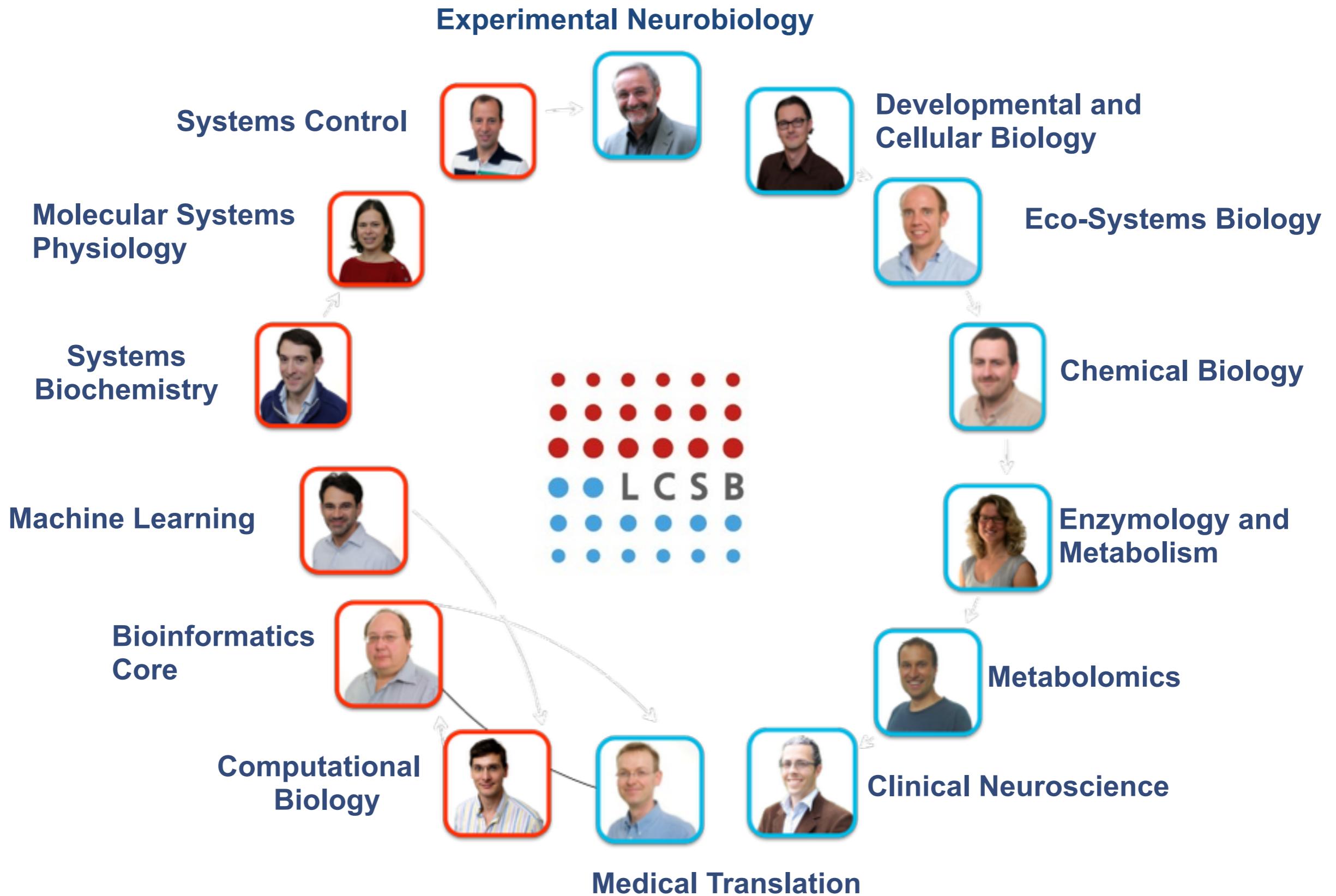
Founded  
in  
Sept. 2009



# LCSB status quo

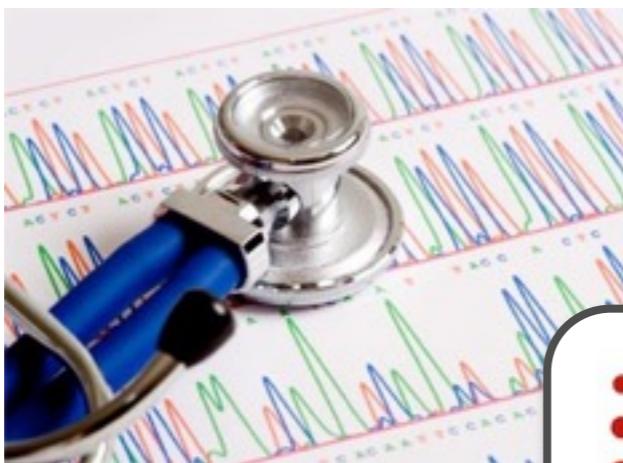
- Personnel
  - 14 research groups
  - 151 staff
  - 38 PhD students
  - 33 nationalities
- KPIs
  - 168 publications
  - 11 patents
  - 2 spin offs
  - 25,65 Mio€ external funds  
of which 315k€ industry funds

# Research groups



# Scientific strategy of LCSB

Clinical translational research



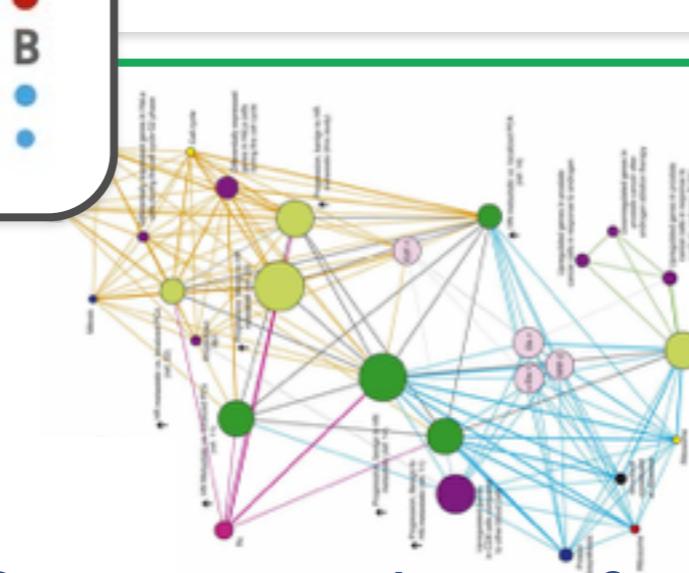
Experimental Biology



Technology platforms



Computer science & Bioinformatics



Discoveries

# Bioinformatics Core facilities (25+ FTE's)

## **Data integration and management**

Organize, store and categorize large amount of data (PetaByte scale). Providing access and management to large compute farms.

## **Automatic pipelines for large scale data-analysis**

Setup of automatic procedures to filter and extract the most relevant information out of large heterogeneous datasets

## **Network (re-)construction**

Extract known and predicted networks (protein, protein-protein, protein-chemical,...) from databases and by applying text-mining technologies

# **Large scale visualization tools for heterogenous data**

Development of 2D and 3D visualization tool for data exploration and hypothesis generation

## **Text-mining**

Crunching large scale full-text corpora of hundereds of thousand of articles to extract knowledge map and relationships between diseases, genes, proteins etc.

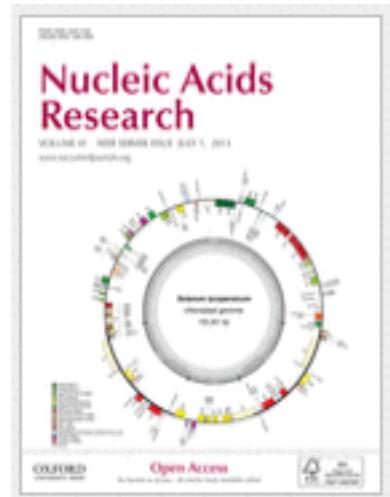
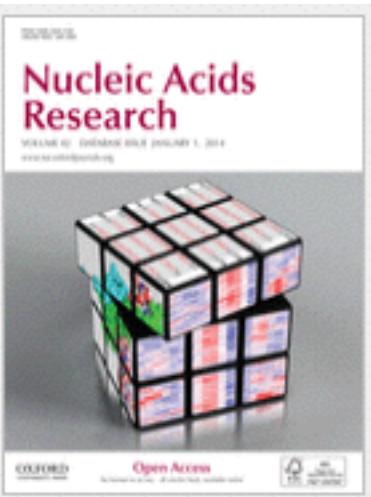
## **Data Analysis partner in several projects**

FP7: eTRIKS, EpiPGX; coGIE; betaJUDO

## **Development of dedicated problem oriented research tools**



# Data, Databases



**bioinformatics.ca**  
links directory

Pathguide

- 2014 NAR database issue reports on 1552 databases and 2013 web server issue reports on 95 web servers
- Currently mentions 173 resources, 623 databases, 1472 web server tools
- 547 biological pathway and interaction related resources

<http://www.oxfordjournals.org/nar/database/c/>

<http://www.oxfordjournals.org/nar/database/cap/>

[http://bioinformatics.ca/links\\_directory](http://bioinformatics.ca/links_directory)

<http://www.pathguide.org>

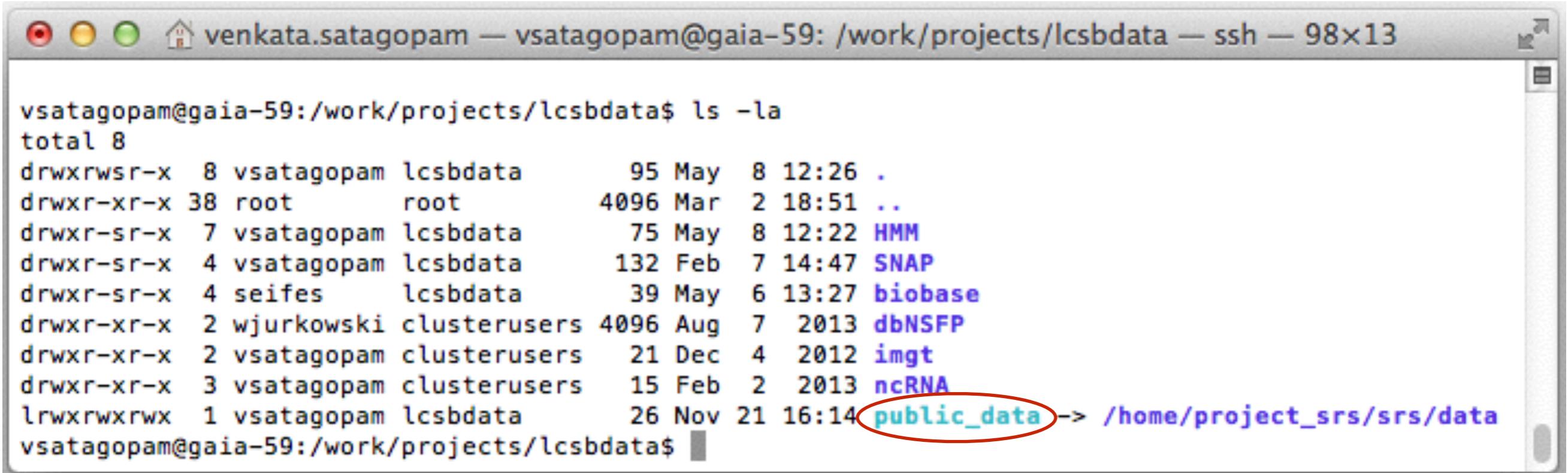
# Data, Databases

- ~80 important public databases are provided on gaia-cluster
- updated regularly using SRS Prisma

```
# Connect to Gaia (Linux/OS X):  
(yourmachine)$> ssh -p 8022 <user_name>@access-gaia.uni.lu
```

```
(gaia-frontend)$> oarsub -I -l nodes=1/core=1,walltime=00:30:00
```

```
(node)$> cd /work/projects/lcsbdata
```



```
venkata.satagopam — vsatagopam@gai-59: /work/projects/lcsbdata — ssh — 98x13  
vsatagopam@gai-59:/work/projects/lcsbdata$ ls -la  
total 8  
drwxrwsr-x 8 vsatagopam lcsbdata 95 May 8 12:26 .  
drwxr-xr-x 38 root root 4096 Mar 2 18:51 ..  
drwxr-sr-x 7 vsatagopam lcsbdata 75 May 8 12:22 HMM  
drwxr-sr-x 4 vsatagopam lcsbdata 132 Feb 7 14:47 SNAP  
drwxr-sr-x 4 seifes lcsbdata 39 May 6 13:27 biobase  
drwxr-xr-x 2 wjurkowski clusterusers 4096 Aug 7 2013 dbNSFP  
drwxr-xr-x 2 vsatagopam clusterusers 21 Dec 4 2012 imgt  
drwxr-xr-x 3 vsatagopam clusterusers 15 Feb 2 2013 ncRNA  
lrwxrwxrwx 1 vsatagopam lcsbdata 26 Nov 21 16:14 public_data -> /home/project_srs/srs/data  
vsatagopam@gai-59:/work/projects/lcsbdata$
```

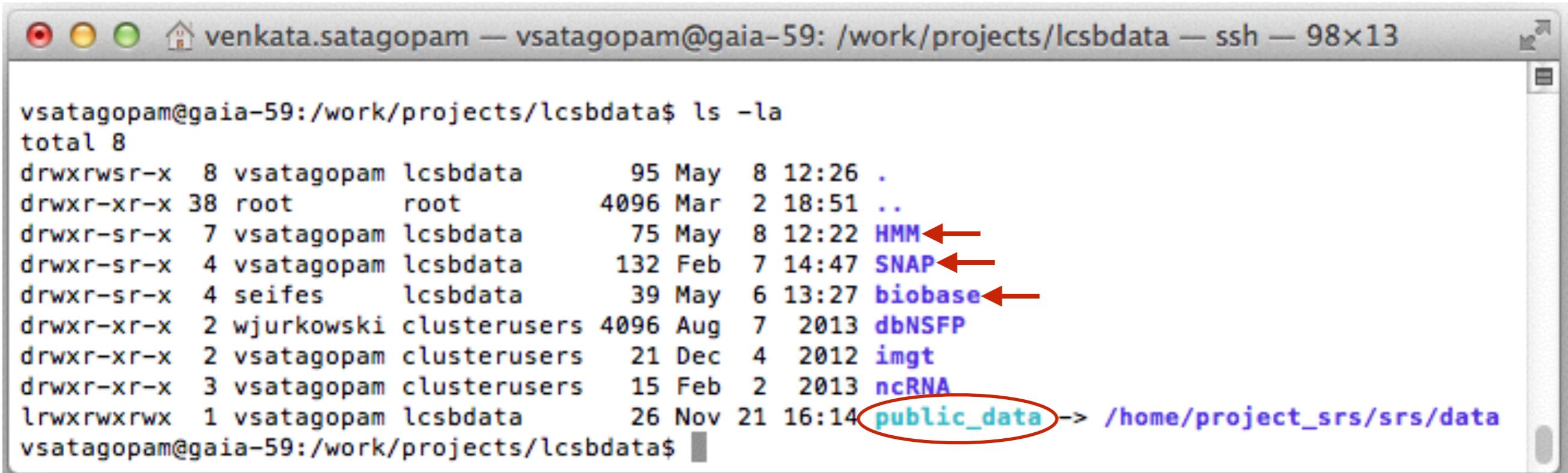
# Data, Databases

- ~80 important public databases are provided on gaia-cluster
- updated regularly using SRS Prisma

```
# Connect to Gaia (Linux/OS X):  
(yourmachine)$> ssh -p 8022 <user_name>@access-gaia.uni.lu
```

```
(gaia-frontend)$> oarsub -I -l nodes=1/core=1,walltime=00:30:00
```

```
(node)$> cd /work/projects/lcsbdata
```

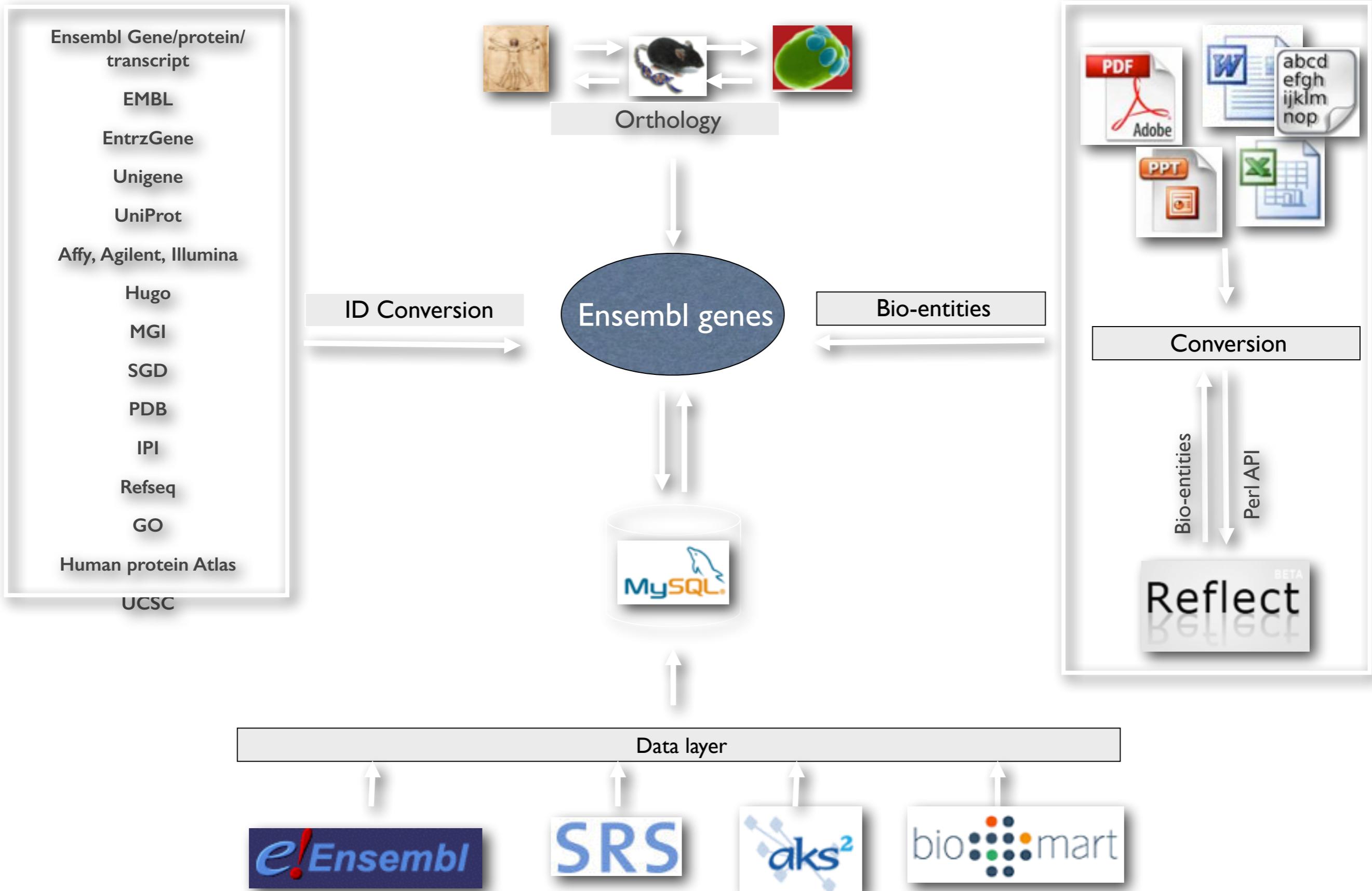


The screenshot shows a terminal window with the following details:

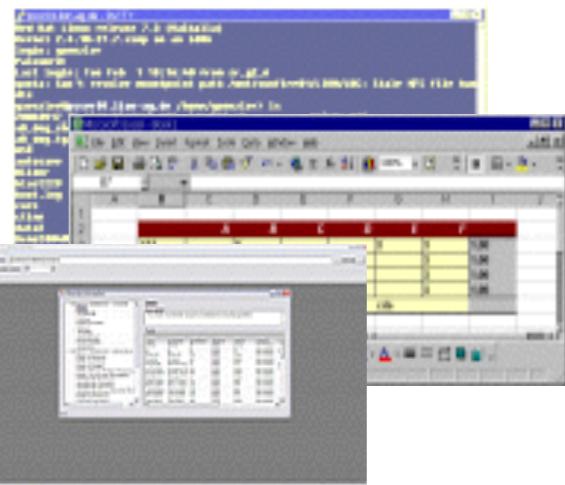
- Terminal title: venkata.satagopam — vsatagopam@gai... — ssh — 98x13
- User: vsatagopam
- Host: gaia-59
- Working Directory: /work/projects/lcsbdata
- Command: ls -la
- Output:

File/Folder	Owner	Group	Type	Date	Time	Size	Description
.	vsatagopam	lcsbdata	drwxrwsr-x	95	May	8 12:26	.
..	root	root	drwxr-xr-x	4096	Mar	2 18:51	..
HMM	vsatagopam	lcsbdata	drwxr-sr-x	75	May	8 12:22	HMM
SNAP	vsatagopam	lcsbdata	drwxr-sr-x	132	Feb	7 14:47	SNAP
biobase	seifes	lcsbdata	drwxr-sr-x	39	May	6 13:27	biobase
dbNSFP	wjurkowski	clusterusers	drwxr-xr-x	4096	Aug	7 2013	dbNSFP
imgt	vsatagopam	clusterusers	drwxr-xr-x	21	Dec	4 2012	imgt
ncRNA	vsatagopam	clusterusers	drwxr-xr-x	15	Feb	2 2013	ncRNA
public_data	vsatagopam	lcsbdata	lrwxrwxrwx	26	Nov	21 16:14	public_data → /home/project_srs/srs/data

# bioCompendium overview



Perl



Java

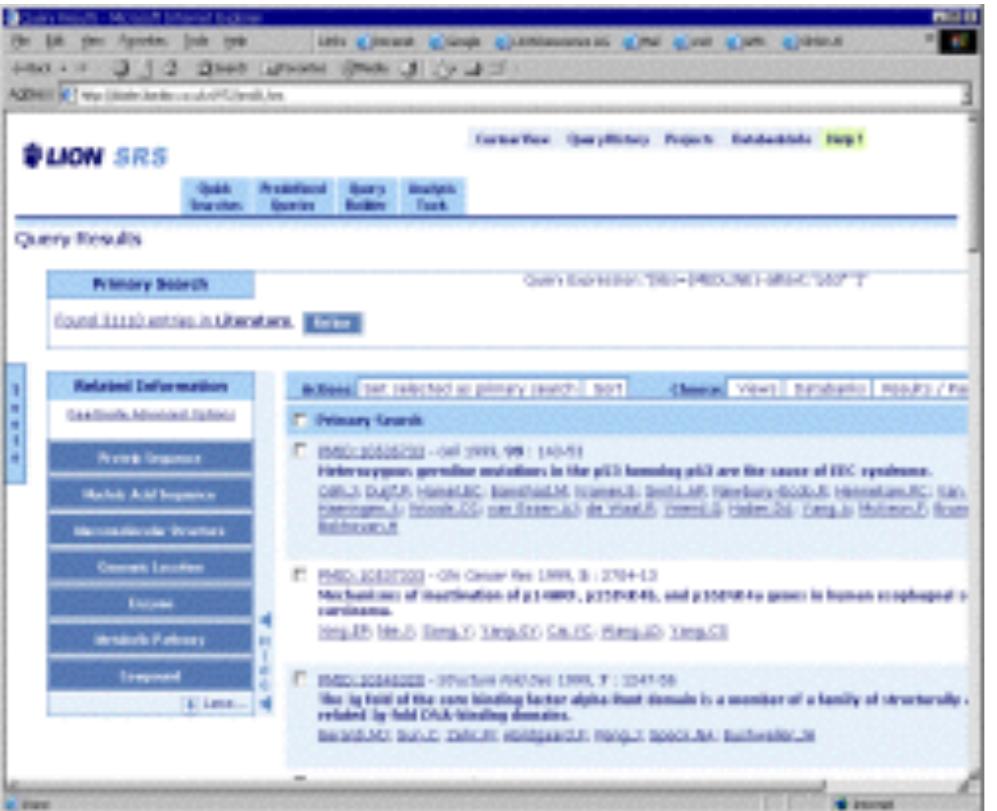
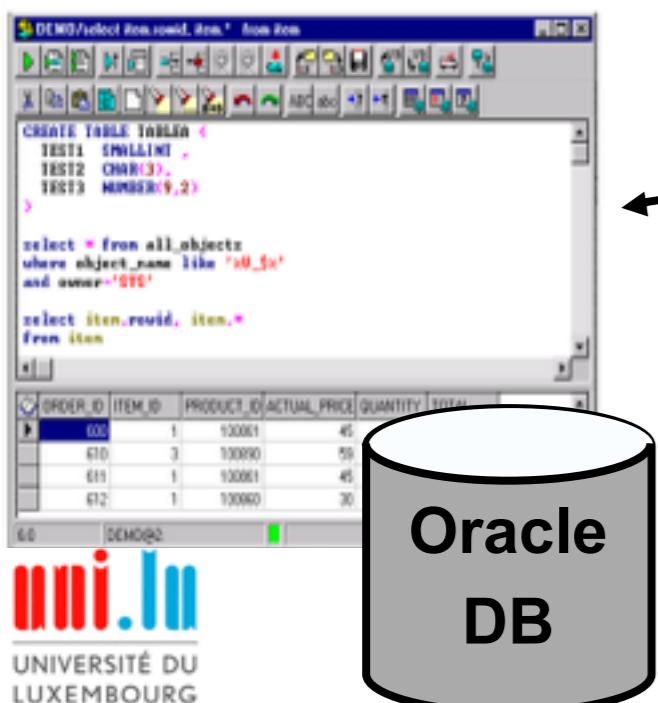
## SRS WS Objects

biowisdom® SRS					
Databank List					
Databank	No. of Entries	Indexing Date	Group	Availability	
BIND Complex	3699	16-Mar-2007	XML Databases	OK	
BIND Domains	213096	16-Mar-2007	XML Databases	OK	
BIND Pathways	7	16-Jul-2007	XML Databases	OK	
BLAST			Tool Results	No Entries	
BLASTZ			Tool Results	No Entries	
CUSTOMANOT			Sequence related	No Entries	
DRUGBANK	4772	23-Feb-2009	Drug databases	OK	
DRUGBANK DRUG TARGET	10575	24-Feb-2009	Drug databases	OK	
DRUGBANK METABOLIZING ENZYME	1043	24-Feb-2009	Drug databases	OK	
EMBL (Release)			Sequence databases - complete	Virtual Database	
EMBL (Updates)			Sequence databases - subsections	No Entries	
EMBL (Whole Genome Shotgun Sequences)			Sequence databases - subsections	No Entries	
EMBL (Whole Genome Shotgun Sequences - full)			Sequence databases - subsections	No Entries	
EMBL (Whole Genome Shotgun Sequences - updates)			Sequence databases - subsections	No Entries	
EMBL (Whole Genome Shotgun Sequences)			Sequence databases - subsections	Virtual Database	
ENTREZGENE	6607869	12-Jan-2010	Genome	OK	
ENTREZSITE	4859	16-Mar-2007	Sequence related	OK	
EPO (Release: 99)	17852	23-Apr-2009	Sequence related	OK	
EPO Proteins	1525988	17-Sep-2009	Patent databases	OK	
ENSEMBL			Sequence databases - complete	Virtual Database	
GENBANK (Release: 17.0)	7025290	01-Mar-2009	Sequence databases - subsections	OK	
GENBANK (Updates) (Release: 20090301)	694995	01-Mar-2009	Sequence databases - subsections	OK	
GENBANK (Whole Genome Shotgun)	49057	17-Mar-2009	Sequence databases - subsections	OK	
GO (Annotations)	28015	14-Feb-2009	Gene Ontology	OK	
HMDB	6888	03-Mar-2009	Metabolome Databases	OK	
HOMOLOGUE	158908	23-Feb-2009	Protein Databases	OK	
HOOMD	44562	23-Feb-2009	Protein 3D Structures	OK	
IPRDOMAIN	59582	21-Jan-2010	Protein 3D Structures	OK	
IPRSPROTEIN	225411	21-Jan-2010	Protein 3D Structures	OK	
IPRSPROTEIN	99549	20-Jan-2010	Protein 3D Structures	OK	
INTERACT	9401	14-Feb-2009	XHL Databases	OK	
INTERACT (Release: 1.0)	12323	17-Mar-2009	Protein Interaction Databases	OK	
INTACT Interactor	41603	14-Feb-2009	XHL Databases	OK	
INTERPRO (Release: 19.0)	17412	14-Feb-2009	XHL Databases	OK	
IPD	68089	14-Feb-2009	Protein Databases	OK	
KEGG Gene (Amino Acid)	4087163	14-Feb-2009	Sequence databases - complete	OK	
KEGG Gene (Nucleic Acid)	4087165	14-Feb-2009	Sequence databases - complete	OK	
KEGG LIGAND	599	14-Feb-2009	Chemical Databases	OK	
KEGG Orthology	11557	17-Feb-2009	Metabolic Pathways	OK	
KEGG Pathway	11399	17-Feb-2009	Metabolic Pathways	OK	
KEGG Reaction	10395	17-Feb-2009	Metabolic Pathways	OK	
LIGAND	8042	27-Feb-2009	Metabolic Pathways	OK	
LIGAND	8043	27-Feb-2009	Metabolic Pathways	OK	
LIGAND	10566	14-Jan-2010	Metabolic Pathways	OK	
REACTION	7824	14-Feb-2009	Metabolic Pathways	OK	
My Protein Sequences			User Owned Databases	No Entries	
My Protein Sequences			User Owned Databases	No Entries	
OMIM	20809	11-Jan-2010	Mutations	OK	
PATHWAY	70742	07-Mar-2009	Metabolic Pathways	OK	
PATHWAY	102140	19-Feb-2009	Metabolic Pathways	OK	
PATOMINE	1327384	19-Feb-2009	Patent databases	OK	
PCPDB	650	14-Feb-2009	Protein Databases	OK	
PDRUNIPROT	54955	14-Feb-2009	Protein 3D Structures	OK	
POBQUIV	55105	26-Feb-2009	Others	OK	
PROTEIN	11912	17-Jan-2010	Protein 3D Structures	No Entries	
PRAMA	142301	17-Jan-2010	Sequence related	OK	
PRAMB	142301	17-Jan-2010	Sequence related	No Entries	
PRAME	20680	23-Dec-2009	Sequence related	OK	
PRAMHHM	11912	17-Jan-2010	Sequence related	OK	
PRAMSEED	11912	17-Jan-2010	Sequence related	OK	
PRAMSEED	11912	10-Nov-2009	Tool Results	OK	
PRAMSEED	11912	10-Nov-2009	Tool Results	No Entries	
PRAMSEED	11912	10-Nov-2009	Tool Results	No Entries	
PRAMTOOLTESTA	2138	14-Feb-2009	Sequence related	OK	
PRAMTOOLTESTA	2138	14-Feb-2009	Sequence related	No Entries	
PRAMTOOLTESTA	2138	14-Feb-2009	Tool Results	OK	
PRAMTOOLTESTA	2138	14-Feb-2009	Tool Results	No Entries	
PROTEIN (Release: 26.42)	2138	14-Feb-2009	Sequence related	OK	
PROTEIN (Release: 19.19)	1542	14-Feb-2009	Sequence related	OK	
PSDB	4785957	21-Jan-2010	Protein 3D Structures	OK	
PSDB	195224	14-Feb-2009	Protein Databases	OK	
PSDB	48170	14-Feb-2009	Protein 3D Structures	OK	
PSDB	8808430	04-Jan-2009	Chemistry databases	OK	
PSDB	4724125	14-Feb-2009	Patent databases	OK	
PSDB	292251	14-Feb-2009	Patent databases	Virtual Database	
REFSEQ			Sequence databases - complete	Virtual Database	
RefSeq (Release)	4111233	16-Jan-2010	Sequence databases - complete	Virtual Database	
RefSeq (Updates)	200228	15-Jan-2010	Sequence databases - subsections	Virtual Database	
RefSeq (Updates)	200228	15-Jan-2010	Tool Results	Virtual Database	
RefSeq Protein (Release)	6413124	27-Feb-2009	Sequence databases - subsections	Virtual Database	
RefSeq Protein (Updates)	913159	21-Jan-2010	Sequence databases - subsections	Virtual Database	
SWISS-PROT	3148750	21-Jan-2010	Sequence databases - subsections	Virtual Database	
TAXONOMY	633941	11-Jan-2010	Sequence related	OK	
UNIPROT	1984496	14-Feb-2009	Tool Results	No Entries	
UNIPROT	1984496	14-Feb-2009	Sequence databases - complete	Virtual Database	
UNIPROT (Release: 14.7)	607876	22-Jan-2009	Sequence databases - complete	Virtual Database	
UNIPROT (Release: 14.7)	2218589	22-Jan-2009	Sequence databases - complete	Virtual Database	
UNIPROT90 (Release: 14.7)	4664726	22-Jan-2009	Sequence databases - complete	Virtual Database	
UNIPROT90 (Release: 14.7)	1310224	21-Jan-2010	Tool Results	Virtual Database	
UniprotKB/Swiss-Prote	514212	21-Jan-2010	Sequence databases - subsections	Virtual Database	
UniprotKB/Swiss-Prote (Release: 57.13)	10158056	21-Jan-2010	Sequence databases - subsections	Virtual Database	
UniprotKB/UniProt (Release: 40.13)	27169	14-Feb-2009	Sequence databases - complete	Virtual Database	
Uniprot Solice Variants			Sequence databases - complete	Virtual Database	
Uniprot Solice Variants			Tool Results	No Entries	
Uniprot Solice Variants			Sequence databases - complete	Virtual Database	
Uniprot Solice Variants			Tool Results	No Entries	

SRS

## SRS Gateway for Oracle

SQL



**Databank List**

Currently sorted by : Databank in ascending order.

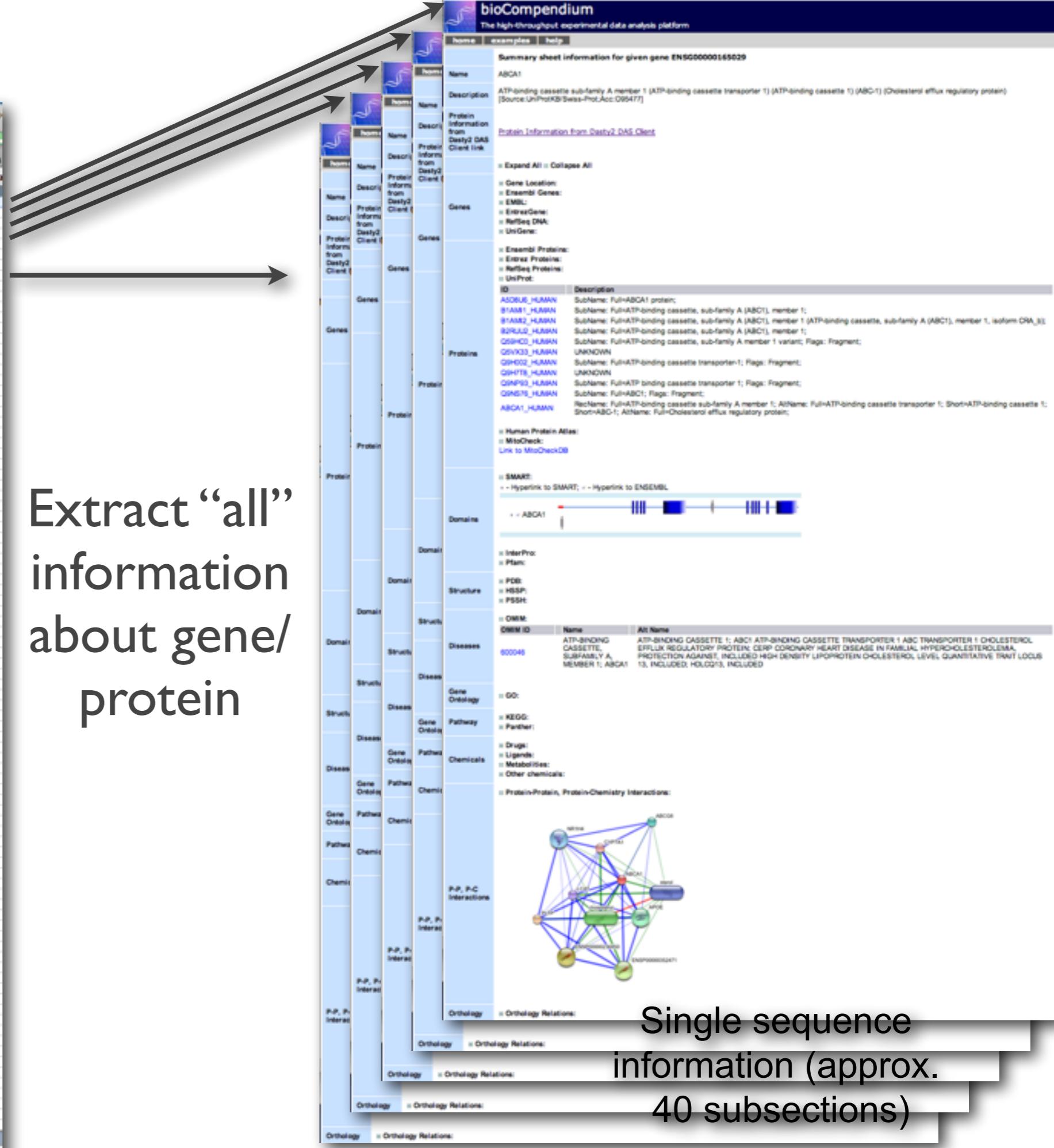
Include Tools : 

Databank	No. of Entries	Indexing Date	Group	Availability
BIND Complex	3698	16-Mar-2007	XML Databases	OK
BIND Interaction	213096	21-Jun-2007	XML Databases	OK
BIND Pathway	7	16-Jul-2007	XML Databases	OK
CUSTOMANNOT			Sequence related	No Entries
DRUGBANK	4774	07-Apr-2010	Others	OK
DrugBank Target	4554	07-Apr-2010	Sequence databanks - complete	OK
EMBL			Sequence databanks - complete	Virtual Databank
EMBL (Release) (Release:107)	138744224	07-Apr-2011	Sequence databanks - subsections	OK
EMBL (Updates) (Release:20110422)	18531116	23-Apr-2011	Sequence databanks - subsections	OK
EMBL (Whole Genome Shotgun Sequences - full release)	65390435	07-Apr-2011	Sequence databanks - subsections	OK
EMBL (Whole Genome Shotgun Sequences - updates)	884570	23-Apr-2011	Sequence databanks - subsections	OK
EMBL (Whole Genome Shotgun Sequences)			Sequence databanks - subsections	Virtual Databank
ENTREZGENE	8115435	18-Mar-2011	Genome	OK
ENZYME	5268	22-Mar-2011	Sequence related	OK
EPD (Release:105)	17852	10-Apr-2011	Sequence related	OK
EPO Proteins	2548827	22-Mar-2011	Patent databanks	OK
GENBANK			Sequence databanks - complete	Virtual Databank
GENBANK (Release) (Release:183.0)	141944676	04-May-2011	Sequence databanks - subsections	OK
GENBANK (Updates) (Release:20110430)	2007542	04-May-2011	Sequence databanks - subsections	OK
GENBANK (Whole Genome Shotgun)	62816097	04-May-2011	Sequence databanks - subsections	OK
GO (Release:20110305)	33699	09-Mar-2011	Gene Ontology	OK
HMDB Enzymes	7265	26-Jul-2010	Others	OK
HMDB Metabolites	8152	26-Jul-2010	Others	OK
HOMOLOGENE	43726	22-Mar-2011	XML Databases	OK
HSSP	67885	03-Mar-2011	Protein 3D Structures	OK
HSSPALIGN	225419173	05-May-2010	Protein 3D Structures	OK
HSSPCHAIRN	99549	05-May-2010	Protein 3D Structures	OK
INTACT Experiment	12025	22-Mar-2011	XML Databases	OK
INTACT Interaction	154553	22-Mar-2011	XML Databases	OK
INTACT Interactor	52297	22-Mar-2011	XML Databases	OK
INTERPRO (Release:21185)	21185	22-Mar-2011	XML Databases	OK
IPI	307585	12-Mar-2011	Sequence databanks - complete	OK
JPO Proteins	1092504	22-Mar-2011	Patent databanks	OK
KEGG Genes (Amino Acid)	6326661	18-Mar-2011	Sequence databanks - complete	OK
KEGG Genes (Nucleic Acid)	6326661	18-Mar-2011	Sequence databanks - complete	OK
KEGG Genome	1546	18-Mar-2011	Metabolic Pathways	OK
KEGG Orthology	14327	18-Mar-2011	Metabolic Pathways	OK
KIPO Proteins	113599	22-Mar-2011	Patent databanks	OK
LCOMPOUND	16948	27-Apr-2011	Metabolic Pathways	OK
LDRUG	9433	11-Mar-2011	Metabolic Pathways	OK
LENZYME	5342	11-Mar-2011	Metabolic Pathways	OK
LGLYCAN	10978	11-Mar-2011	Metabolic Pathways	OK
LREACTION	8451	11-Mar-2011	Metabolic Pathways	OK

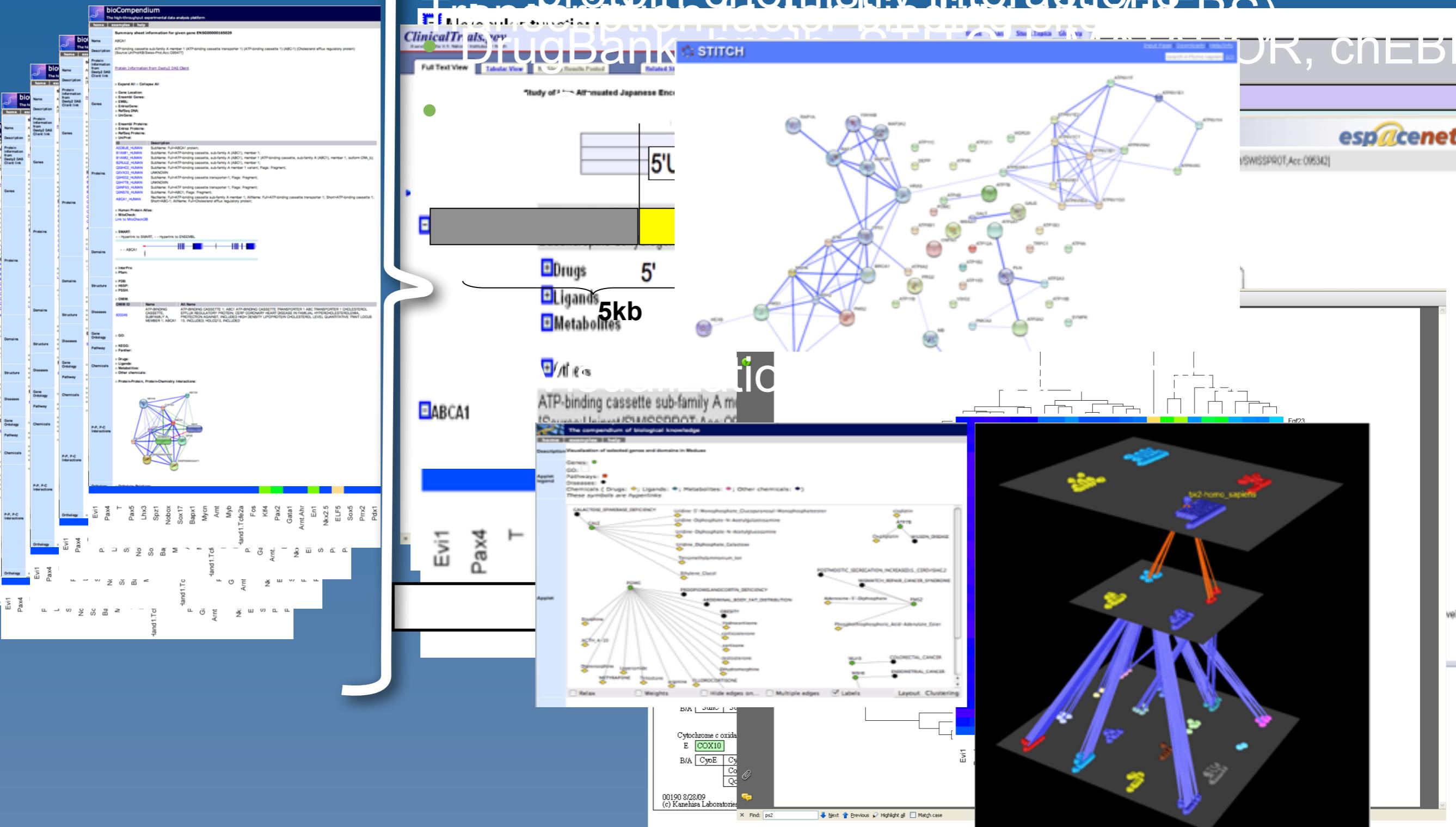
OMIM	21395
PANTHER	70742
PATHWAY	137569
PATOME	1327384
PDB	67981
PDB2UNIPROT	63956
PDBEQUIV	55101
PDBMODEL	
PFAMA	11912
PFAMB	142303
PFAMC	423
PFAMHMM	31912
PFAMSEED	11912
PRINTS (Release:41.1)	2050
PROSITE (Release:20.71)	2228
PROSITEDOC (Release:20.71)	1608
PSSH	4785957
PUBCHEM_COMPOUND	27392083
PUBCHEM_COMPOUND_MESH	61634
PUBCHEM_COMPOUND_SYNONYM	8808430
PUBCHEM_SUBSTANCE	69088009
Patent Proteins	
REACTOME	292251
RefSeq	
RefSeq (Release)	4814433
RefSeq (Updates) (Release:20110304)	103306
RefSeq_Protein	
RefSeq_Protein (Release)	11934213
RefSeq_Protein (Updates) (Release:20110304)	1072133
SWISSPFAM	3148759
TAXONOMY	738056
UNIGENE	2272090
UNIPROT	
UNIREF100 (Release:2011_03)	12162651
UNIREF50 (Release:2011_03)	3785756
UNIREF90 (Release:2011_03)	7920686
USPTO Proteins	1780942
UniProtKB/Swiss-Prot (Release:2011_03)	525997
UniProtKB/TrEMBL (Release:2011_03)	13897064
Uniprot Splice Variants	29994
VIEWS	

# List of genes/proteins

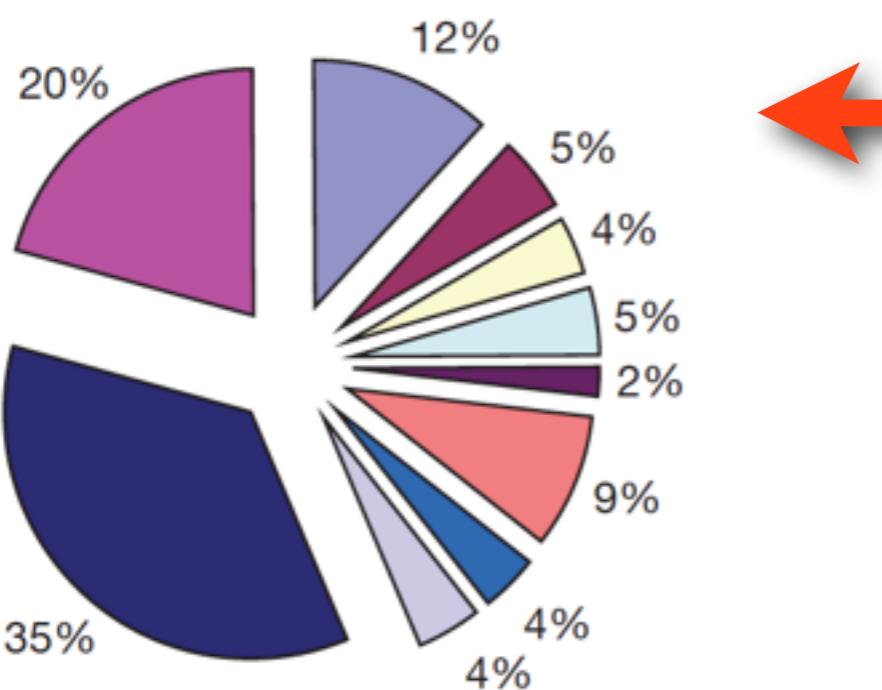
	A	B
1	EPO_MOUSE	Erythropoietin
2	EPOR_MOUSE	Erythropoietin receptor
3	JAK2_MOUSE	Tyrosine-protein kinase JAK2
4	PTN6_MOUSE	Tyrosine-protein phosphatase non-receptor type 6
5	PTN11_MOUSE	Tyrosine-protein phosphatase non-receptor type 11
6	GAB1_MOUSE	GAB2-associated-binding protein 1
7	GAB2_MOUSE	GAB2-associated-binding protein 2
8	SHPT1_MOUSE	Phosphatidylinositol-3,4,5-trisphosphate 5-phosphatase 1
9	SHIP2_MOUSE	Phosphatidylinositol-3,4,5-trisphosphate 5-phosphatase 2
10	PBSA_MOUSE	Phosphatidylinositol 3-kinase regulatory subunit alpha
11	P58B_MOUSE	Phosphatidylinositol 3-kinase regulatory subunit beta
12	P58G_MOUSE	Phosphatidylinositol 3-kinase regulatory subunit gamma
13	PK3CA_MOUSE	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha isoform
14	PK3CB_MOUSE	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit beta isoform
15	PK3CG_MOUSE	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit gamma isoform
16	PK3CD_MOUSE	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit delta isoform
17	PDPK1_MOUSE	3-phosphoinositide-dependent protein kinase 1
18	AKT1_MOUSE	RAC-alpha serine/threonine-protein kinase
19	AKT2_MOUSE	RAC-beta serine/threonine-protein kinase
20	AKT3_MOUSE	RAC-gamma serine/threonine-protein kinase
21	GSK3B_MOUSE	Glycogen synthase kinase-3 alpha
22	GSK3B_MOUSE	Glycogen synthase kinase-3 beta
23	FRAP_MOUSE	Serine/threonine-protein kinase mTOR
24	LST8_MOUSE	Target of rapamycin complex subunit LST8
25	RPTOR_MOUSE	Regulatory-associated protein of mTOR
26	AKTS1_MOUSE	Proline-rich AKT1 substrate 1
27	RICTR_MOUSE	Rapamycin-insensitive companion of mTOR
28	RHEB_MOUSE	GTP-binding protein Rheb
29	4EBP1_MOUSE	Eukaryotic translation initiation factor 4E-binding protein 1
30	TSC1_MOUSE	Hamartin
31	TSC2_MOUSE	Tuberin
32	K58B1_MOUSE	Ribosomal protein S6 kinase beta-1
33	K58B2_MOUSE	Ribosomal protein S6 kinase beta-2
34	IF4E_MOUSE	Eukaryotic translation initiation factor 4E
35	KPC2_MOUSE	Protein kinase C zeta type
36	PR15A_MOUSE	Protein phosphatase 1 regulatory subunit 15A
37	PPR3C_MOUSE	Protein phosphatase 1 regulatory subunit 3C
38	PHP1_MOUSE	PH domain leucine-rich repeat-containing protein phosphatase 1
39	PML_MOUSE	Probable transcription factor PML
40	2ASA_MOUSE	Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit alpha isoform
41	2ASG_MOUSE	Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit gamma isoform
42	2ASE_MOUSE	Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit epsilon isoform
43	PP2AA_MOUSE	Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform
44	PP2AB_MOUSE	Serine/threonine-protein phosphatase 2A catalytic subunit beta isoform
45	Q9WVH4_MOUSE	Forkhead protein FKH2
46	FOXO1_MOUSE	Forkhead box protein O1
47	CCND2_MOUSE	G1/S-specific cyclin-D2
48	CCND3_MOUSE	Cyclin-G2
49	CCNE1_MOUSE	G1/S-specific cyclin-E1
50	CDN18_MOUSE	Cyclin-dependent kinase inhibitor 1B
51	CDN1A_MOUSE	Cyclin-dependent kinase inhibitor 1
52	CCNC_MOUSE	Cyclin-C
53	EGR3_MOUSE	Early growth response protein 3
54	CC45IL_MOUSE	CDC45-related protein
55	BAD_MOUSE	Bcl2 antagonist of cell death
56	MDM2_MOUSE	E3 ubiquitin-protein ligase Mdm2
57	P53_MOUSE	Cellular tumor antigen p53
58	4EBP1_MOUSE	Eukaryotic translation initiation factor 4E-binding protein 1
59	MYC_MOUSE	Myc proto-oncogene protein
60	PIM1_MOUSE	Proto-oncogene serine/threonine-protein kinase pim-1
61	JUNB_MOUSE	Transcription factor jun-B
62	ACV1B_MOUSE	activin A receptor, type 1B
63	ACV1C_MOUSE	activin A receptor, type 1C
64	ACVR1_MOUSE	activin A receptor, type 1
65	AKT1_MOUSE	thymoma viral proto-oncogene 1
66	ARBK1_MOUSE	adrenergic receptor, kinase, beta 1
67	ATF2_MOUSE	activating transcription factor 2
68	ATF2_MOUSE	activating transcription factor 2
69	AVR2B_MOUSE	activin receptor IIB
70	BAMBI_MOUSE	BMP and activin membrane-bound inhibitor, homolog (Xenopus laevis)
71	BMP2_MOUSE	bone morphogenic protein 2
72	BMP4_MOUSE	bone morphogenic protein 4
73	BMP6_MOUSE	bone morphogenic protein 6
74	BMP7_MOUSE	bone morphogenic protein 7
75	BMPR2_MOUSE	bone morphogenic protein receptor, type II (serine/threonine kinase)
76	BMR1A_MOUSE	bone morphogenic protein receptor, type IA
77	BMR1B_MOUSE	bone morphogenic protein receptor, type IB
78	CER1_MOUSE	cerberus 1 homolog (Xenopus laevis)
79	CHRD_MOUSE	chordin
80	CITE1_MOUSE	Cbp/p300-interacting transactivator with Glu/Asp-rich carboxy-terminal domain 1
81	CITE2_MOUSE	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2
82	CRDL1_MOUSE	chordin-like 1
83	CRDL2_MOUSE	chordin-like 2
84	CREB1_MOUSE	cAMP responsive element binding protein 1
85	CTSP1_MOUSE	C-terminal binding protein 1
86	CTDS1_MOUSE	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase 1
87	CTDSL_MOUSE	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase-like
88	CUL1_MOUSE	cullin 1
89	DAB2_MOUSE	disabled homolog 2 (Drosophila)



# Protein–protein, Pathway and GO terms enrichment of the fusion gene–target gene interactions



# Use of BioCompendium



Phenotypic profiling of the human genome by time-lapse microscopy reveals **genes with functions in cell division, survival or migration**  
Ellenberg J. EMBL  
(Neumann et.al., Nature, 2010)

From experimental setup to bioinformatics: an RNAi screening platform to identify **host factors involved in HIV-1 replication**  
Kräusslich H.G., Dep. of Infect. Diseases, Univ. Heidelberg  
(Börner et.al., Biotech J., 2010)

Human-gpDB: A database of **human GPCRs, G-proteins, Effectors and their interactions**  
Hamodrakas S., Univ. Athens, Greece  
(Satagopam et.al., submitted)

Defective lamin A-Rb signaling in **Hutchinson-Gilford Progeria Syndrome** and reversal by farnesyltransferase inhibition  
Djabali K., Columbia Univ. NYC  
(Marji et.al., submitted)

**Signaling in insulin-producing cells** is altered when cells are grown in islets (3-D) compared to monolayer (2-D).  
Bergsten P., Medical Cell Biology, Univ. Uppsala

**Pain related targets**  
Kuner R., Institute of Pharmacology, Univ. Heidelberg

**Proliferation vs. Differentiation** in homeopathic cells  
Eils R., Klingmüller U., BioQuant, Univ. Heidelberg

Gene Ontology (GO), Biological process annotations of the 572 validated mitotic hits

Search gene-expression analysis results: search...

**Legend**

1. Click on analysis hyper link to see the analysis results
2. Select the checkboxes from different experiments to compare/analyze/enrich with **bioCompendium**

Expand All  Collapse All

Huntington's Disease

Reference	Title	P-Value cut off	Fold change significant cut off	No. of hits
<input type="checkbox"/> RF3	TAMAHUD screen	0.05	0.0	8521
<input checked="" type="checkbox"/> RF3		0.05	1.0	443
<input type="checkbox"/> RF3		0.05	1.5	117
<input type="checkbox"/> RF3		0.05	2.0	34
<input type="checkbox"/> RF3		0.01	0.0	4758
<input type="checkbox"/> RF3		0.01	1.0	440
<input type="checkbox"/> RF3		0.01	1.5	117
<input type="checkbox"/> RF3		0.01	2.0	34

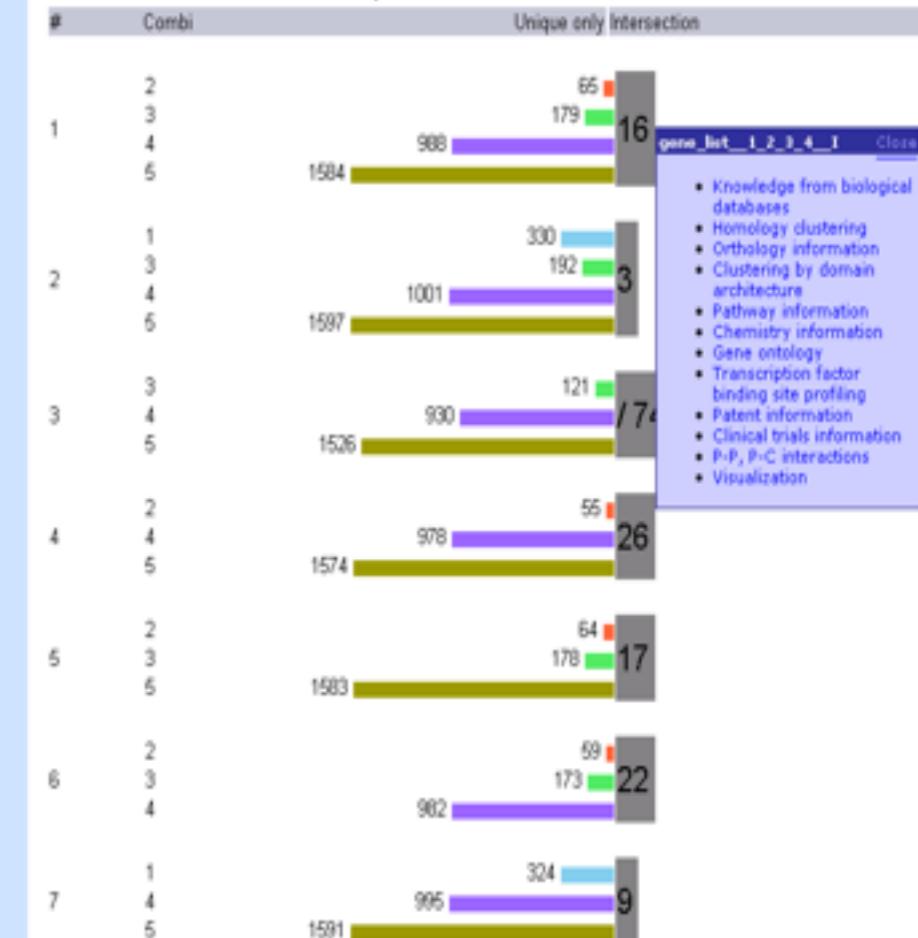
Analysis	Organism	GEO acc	PubMed ID
<input type="checkbox"/> BE3	human		
<input checked="" type="checkbox"/> RF3	mouse	GDS2912	17696994
<input type="checkbox"/> RF3	mouse	GDS2911	17708681
<input type="checkbox"/> Lee, Plos Genet 2007	mouse	GDS2911	17708681
<input type="checkbox"/> Apostol, Hum Mol Genet 2006	rat	GDS1236	16330479

Select the data sets to be analyzed

[Go to bioCompendium](#)

Links to GEO & PubMed

Combinations between maximum top 10 sets:



- Knowledge from biological databases
- Homology clustering
- Orthology information
- Clustering by domain architecture
- Pathway information
- Chemistry information
- Gene ontology
- Transcription factor binding site profiling
- Patent information
- Clinical trials information
- P-P, P-C interactions
- Visualization

**bioCompendium**  
The compendium of biological knowledge

home examples help

Description: Information from various biological sources in a short tabular view.

No of Genes: 41

The following table is sortable.

Name	Description	Gene
AP1S1P1	AP1 subunit gamma-binding protein 1 (Damma-kyanopine) [Source:UniProtKB/Swiss-Prot;Acc:Q8UMZ2]	ENSG00000006114
ARF5	ADP-ribosylation factor 5 [Source:UniProtKB/Swiss-Prot;Acc:P84085]	ENSG00000004059
BANP2L1	Brain-specific angiogenesis inhibitor 1-associated protein 2-like protein 1 (BAI1-associated protein 2-like protein 1) [Source:UniProtKB/Swiss-Prot;Acc:Q8UH41]	ENSG00000006453
DCAP2B	D-cell receptor-associated protein 29 (DCR-associated protein Dap29) [Source:UniProtKB/Swiss-Prot;Acc:Q9UHQ4]	ENSG00000005790
C1orf601	UPF0490 protein C1orf601 [Source:UniProtKB/Swiss-Prot;Acc:Q5TH74]	ENSG00000001460
C2orf66	UPF0511 protein C2orf66, mitochondrial Precursor [Source:UniProtKB/Swiss-Prot;Acc:Q7L582]	ENSG00000001609
CACN02	Voltage-dependent calcium channel gamma-3 subunit (Neuronal voltage-gated calcium channel gamma-3 subunit) [Source:UniProtKB/Swiss-Prot;Acc:Q80359]	ENSG00000006116
CC1L8	C-C motif chemokine 18 (Small-inducible cytokine 18) (Macrophage inflammatory protein 4) (MIP-4) (Pulmonary and activation-regulated chemokine) (CC chemokine PARC) [Source:UniProtKB/Swiss-Prot;Acc:P55774]	ENSG00000006074
CC1L6	C-C motif chemokine 26 (Small-inducible cytokine 26) (Elastin-2) (Macrophage inflammatory protein 4-alpha) (MIP-4-alpha) (Thymic stroma chemokine-1) (TSC-1) [Source:UniProtKB/Swiss-Prot;Acc:P597268]	ENSG00000006605
CFTR	Cystic fibrosis transmembrane conductance regulator (CFTR) (cAMP-dependent chloride channel) (ATP-binding cassette transporter sub-family C member 7) [Source:UniProtKB/Swiss-Prot;Acc:P13669]	ENSG00000001626
CD3CL1	Fractalkine Precursor (C10-C1 motif chemokine 1) (Fractalkine) (CX3C membrane-anchored chemokine) (Small-inducible cytokine 11) (Processed fractalkine) [Source:UniProtKB/Swiss-Prot;Acc:P70423]	ENSG00000006210
CPD561	Cytochrome P450 26B1 (EC 1.14.-.-) (P450 26A2) (P450 retinoic acid-inactivating 2) (P450RA2) (Retinoic acid-metabolizing cytochrome) [Source:UniProtKB/Swiss-Prot;Acc:Q9W63]	ENSG00000003137
CPBP1A1	Cytochrome P450 14A1 (EC 1.14.13.70) (CYPL1) (P450L) (Sterol 14-alpha demethylase) (Lanosterol 14-alpha demethylase) (LDM) (P450 14DM) (P45014DM) [Source:UniProtKB/Swiss-Prot;Acc:Q16860]	ENSG00000001630
DRN001	Dynamin-domain-containing protein 1 [Source:UniProtKB/Swiss-Prot;Acc:Q9H691]	ENSG00000003240
DVL2	Segment polarity protein dishevelled homolog DVL-2 (Dishevelled-2) (DVL homolog 2) [Source:UniProtKB/Swiss-Prot;Acc:Q16841]	ENSG00000004695
ESR1A1	Steroid hormone receptor ESR1 (Estrogen-related receptor, alpha) (ERα-αpha) (Estrogen receptor-like 1) (Nuclear receptor subfamily 3 group B member 1) [Source:UniProtKB/Swiss-Prot;Acc:P11474]	ENSG000000173153
FHBP4	FHBP6-binding protein 4 (EC 5.2.1.8) (Peptidyl-prolyl cis-trans isomerase) (PPase) (Fistamase) (HSP-binding immunophilin) (HBP) (FHBP52 protein) [Source:UniProtKB/Swiss-Prot;Acc:Q02790]	ENSG00000004478
FUCA2	Plasma alpha-L-fucosidase Precursor (EC 3.2.1.51) (Alpha-L-fucosidase N-acetylhexosaminidase 2) (Alpha-L-fucosidase 2) [Source:UniProtKB/Swiss-Prot;Acc:Q9BYT2]	ENSG00000001036
HOMA11	Homeobox protein Hox-A11 (Hox-11) [Source:UniProtKB/Swiss-Prot;Acc:P21270]	ENSG00000005073
HS3ST1	Heparan sulfate glucosamine 3-O-sulfotransferase 1 Precursor (EC 2.6.2.22) (Heparan sulfate Glucosaminyl 3-O-sulfotransferase 1) (Heparan sulfate 3-O-sulfotransferase 1) (HS-OST-1) [Source:UniProtKB/Swiss-Prot;Acc:Q14792]	ENSG00000002567
HSPB6	Heat shock protein beta-6 (HspB6) (Heat shock 20 kDa-like protein p20) [Source:UniProtKB/Swiss-Prot;Acc:Q14598]	ENSG00000004776
UFR0378	UFR0378 protein HAAA100 Precursor (Breast cancer overexpressed gene 1 protein) (Antigen MAA-22) [Source:UniProtKB/Swiss-Prot;Acc:Q9YXW1]	ENSG00000002567

**bioCompendium**  
The high-throughput experimental data analysis platform

home examples help

**Legend**

Click on color bars to explore more knowledge

Expand All  Collapse All

**Selected background**

Other gene list(s)

#	Name	Size	Organism
1	RF3_RF3	333	human
2	GDS2912_18wkR6_1versuswildtype	81	mouse->human
3	GSE10202_StriatumCHL2versusstriatumwild-type	195	mouse->human
4	GSE9803_StriatumR6_2versusstriatumwild-type	1004	mouse->human
5	Strand_neuroscience_R6_2mouseversuscontrol	1600	mouse->human

**Input after conversion**

**Combinations**

Combinations between maximum top 10 sets:

# Environment Modules

- A software package that allows us to provide a multitude of **applications** and **libraries** in **multiple versions** on the UL HPC platform
- The tool itself is used to manage environment variables such as PATH, LD\_LIBRARY\_PATH and MANPATH, enabling the **easy loading** and **unloading** of application/library profiles and their **dependencies**
- New applications can be added using **EasyBuild** framework

<https://hpc.uni.lu/users/software>

<https://hpc.uni.lu/users/docs/modules.html>

<https://github.com/hpcugent/easybuild/wiki>

<https://hpc.uni.lu/users/docs/programming.html>

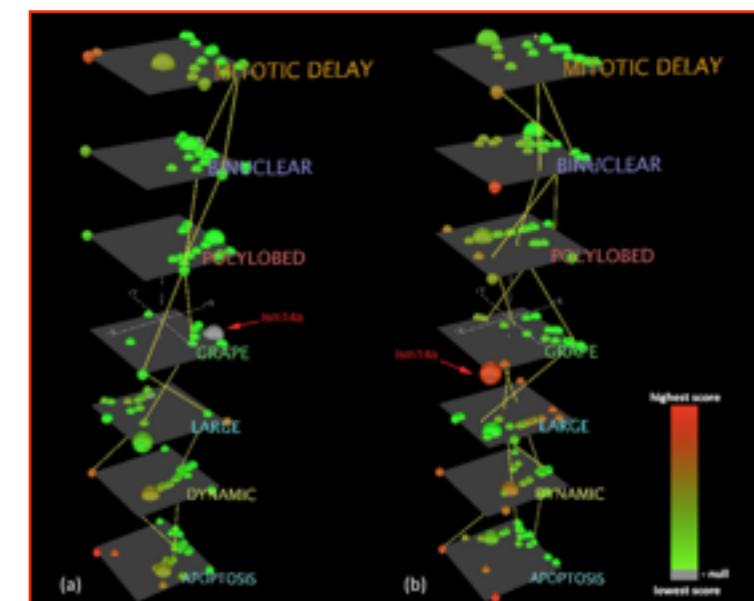
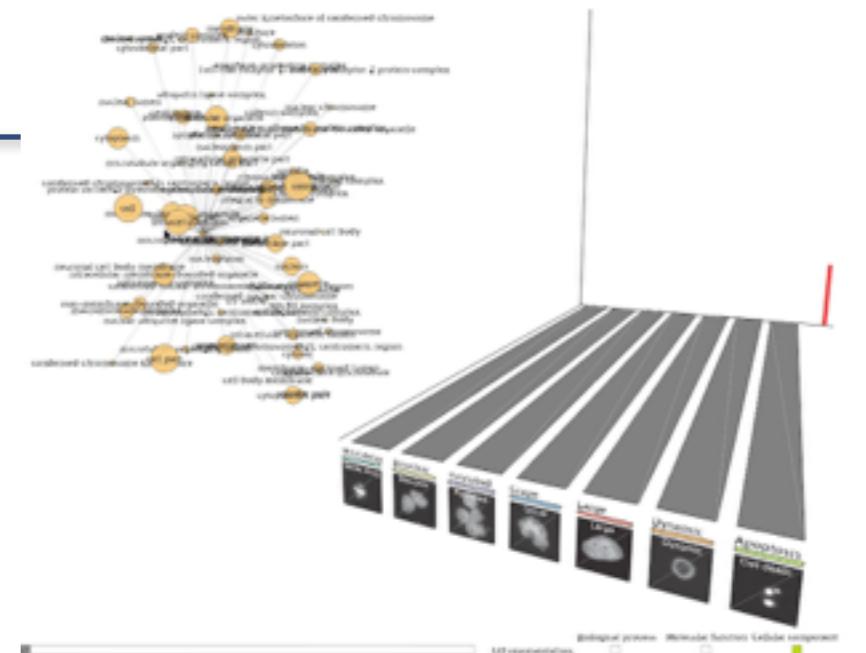
# Environment Modules - Biology category

## 1. Biology category <https://hpc.uni.lu/users/software/>

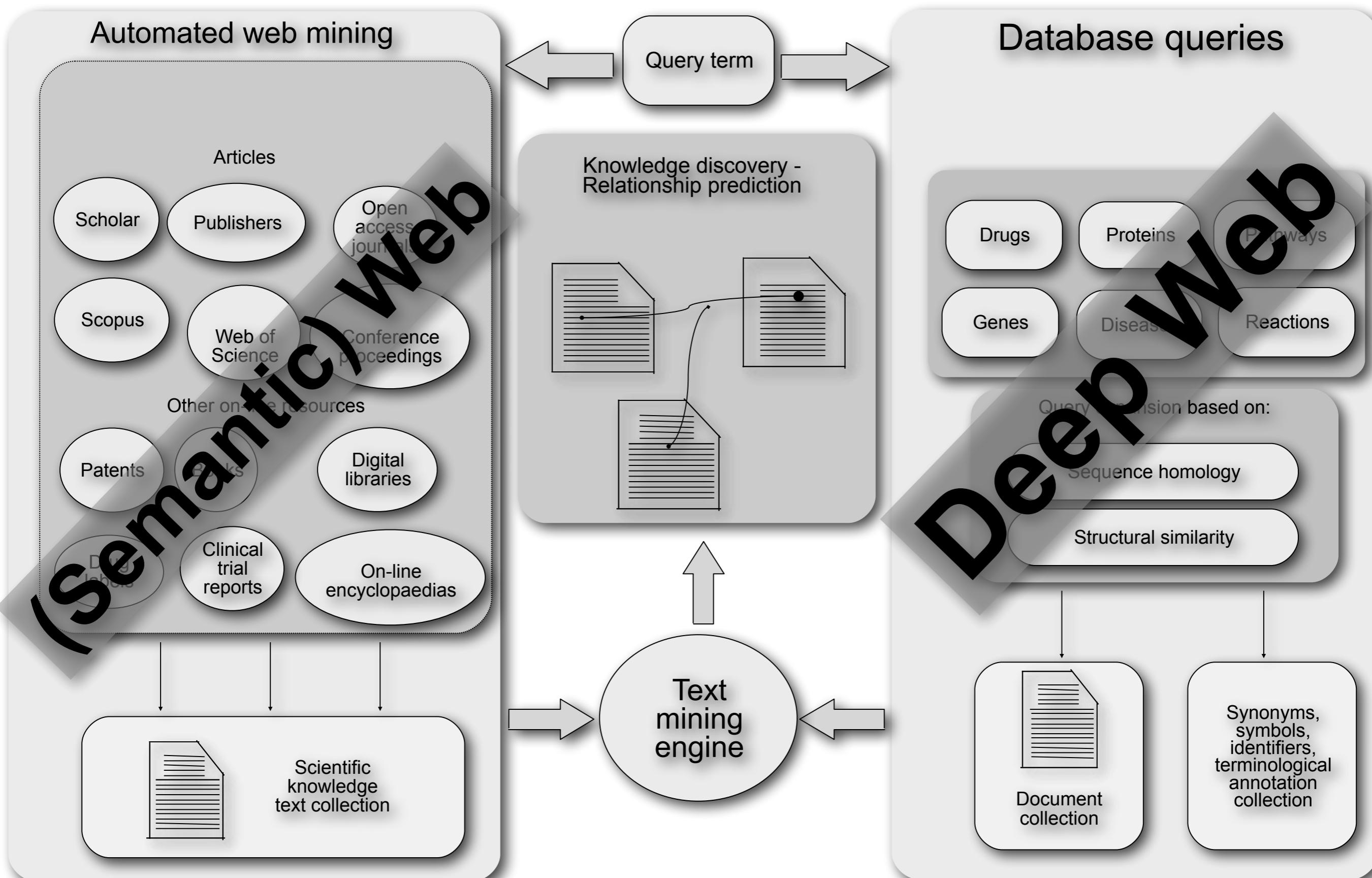
Software available in category: <a href="#">Biology</a>		
Name	Description	Version(s)
ABySS	Description: Assembly By Short Sequences - a de novo, parallel, paired-end sequence assembler	1.3.4-ictce-5.3.0-Python-2.7.3 1.3.4-goolf-1.4.10-Python-2.7.3
ALLPATHS-LG	Description: ALLPATHS-LG, the new short read genome assembler.	46968-goolf-1.4.10
AMOS	Description: The AMOS consortium is committed to the development of open-source whole genome assembly software	3.1.0-goolf-1.4.10 3.1.0-ictce-5.3.0
bam2fastq	Description: The BAM format is an efficient method for storing and sharing data from modern, highly parallel sequencers. While primarily used for storing alignment information, BAMs can (and frequently do) store unaligned reads as well.	1.1.0-goolf-1.4.10 1.1.0-ictce-5.3.0
BamTools	Description: BamTools provides both a programmer's API and an end-user's toolkit for handling BAM files.	2.2.3-goolf-1.4.10 2.2.3-ictce-5.3.0
BEDTools	Description: The BEDTools utilities allow one to address common genomics tasks such as finding feature overlaps and computing coverage. The utilities are largely based on four widely-used file formats: BED, GFF/GTF, VCF, and SAM/BAM.	2.17.0-goolf-1.4.10 2.18.1-goolf-1.4.10
BFAST	Description: BFAST facilitates the fast and accurate mapping of short reads to reference sequences. Some advantages of BFAST include: 1) Speed: enables billions of short reads to be mapped quickly. 2) Accuracy: A priori probabilities for mapping reads with defined set of variants. 3) An easy way to measurably tune accuracy at the expense of speed.	0.7.0a-goolf-1.4.10 0.7.0a-ictce-5.3.0
biodeps	Description: The purpose of this collection is to provide common dependencies in a bundle, so that software/modules can be mixed and matched easily for composite pipelines in Life Sciences	1.6-ictce-5.3.0-extended 1.6-goolf-1.4.10-extended 1.6-goolf-1.4.10 1.6-ictce-5.3.0
BioPerl	Description: BioPerl is the product of a community effort to produce Perl code which is useful in biology. Examples include Sequence objects, Alignment objects and database searching objects.	1.6.1-goolf-1.4.10-Perl-5.16.3 1.6.1-ictce-5.3.0-Perl-5.16.3
Biopython	Description: Biopython is a set of freely available tools for biological computation written in Python by an international team of developers. It is a distributed collaborative effort to develop Python libraries and applications which address the needs of current and future work in bioinformatics.	1.61-goolf-1.4.10-Python-2.7.3 1.61-ictce-5.3.0-Python-2.7.3
BLAST	Description: Basic Local Alignment Search Tool, or BLAST, is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of different proteins or the nucleotides of DNA sequences.	2.2.28-goolf-1.4.10-Python-2.7.3 2.2.28-ictce-5.3.0-Python-2.7.3
BLAT	Description: BLAT on DNA is designed to quickly find sequences of 95% and greater similarity of length 25 bases or more.	3.5-goolf-1.4.10
Bowtie2	Description: Bowtie 2 is an ultrafast and memory-efficient tool for aligning sequencing reads to long reference sequences.	2.0.2-goolf-1.4.10 2.0.2-ictce-5.3.0

# Web applications

- PathVar - <http://pathvar.embl.de/>
- EnrichNet - <http://www.enrichnet.org>
- ArrayMining - <http://www.arraymining.net>
- PathExpand - <http://www.pathexpand.net/>
- TopoGSA - <http://www.topogsa.net>
- GAT|CAT - <http://metnet.uni.lux/software/od/>
- METNET - <http://metnet.uni.lux:9088>
- bioCompendium - <http://biocompendium.embl.de>
- MapAnnotator - [http://rschneider.embl.de/map\\_annotator](http://rschneider.embl.de/map_annotator)
- Arena3D - <http://arena3d.org>
- Martini - <http://martini.embl.de>
- Caipirini - <http://caipirini.org>
- HivMut - <http://hivmut.org>
- Human-gpDB - [http://schneider.embl.de/human\\_gpdb](http://schneider.embl.de/human_gpdb)
- Reflect - <http://reflect.ws>
- OnTheFly - <http://onthefly.embl.de>



# Bridging the silos



# Primary Design Criteria

- Installing and using should be fast and easy
- Interactive (speed)

<http://reflect.embl.de>



Pafilis, O'Donoghue, Jensen, Horn, Kuhn, Brown, Schneider  
Reflect: Augmented Browsing for the Life Scientist  
Nature Biotechnology, 27, 508-510, 2009



Matched terms

Sequence identifier(s)

Synonyms

Domains in SMART

Draggable Sequence window

Protein description

Representative 3D structure In PDBsum

Genomic location In ENSEMBL

Major interaction Partners in STITCH

Related Medline Abstracts in iHOP

Green indicates Known subcellular location

Organism(s)

Edit synonyms or description

Domains – mouse over to show name

Sequence Scroll bars

Organism

TP53

Protein

ENSP00000269305

P53\_HUMAN

H. sapiens

[edit]

Antigen NY CO 13; Phosphoprotein p53; Tumor suppressor p53; A2522

Domains, Sequence, Structure, Locus, Literature

MEEPQSDPSVEPPLSQETFSDLWKLLENVNLSPQLPSQAMDDLM

Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and

Close (Esc)

Help

Domains, Sequence, Structure, Locus, Literature

Subcellular localization: Nucleus

Organism: Human

Sequence: MEEPQSDPSVEPPLSQETFSDLWKLLENVNLSPQLPSQAMDDLM

3D structure: PDBsum

Major interaction Partners: STITCH

Related Medline Abstracts: iHOP

Organism(s): H. sapiens

Sequence identifier(s): ENSP00000269305

Synonyms: Antigen NY CO 13; Phosphoprotein p53; Tumor suppressor p53; A2522

Domains in SMART: Domains, Sequence, Structure, Locus, Literature

Draggable Sequence window: MEEPQSDPSVEPPLSQETFSDLWKLLENVNLSPQLPSQAMDDLM

Protein description: Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and

Representative 3D structure In PDBsum: PDBsum

Genomic location In ENSEMBL: P53\_HUMAN

Major interaction Partners in STITCH: STITCH

Related Medline Abstracts in iHOP: iHOP

Organism(s): H. sapiens

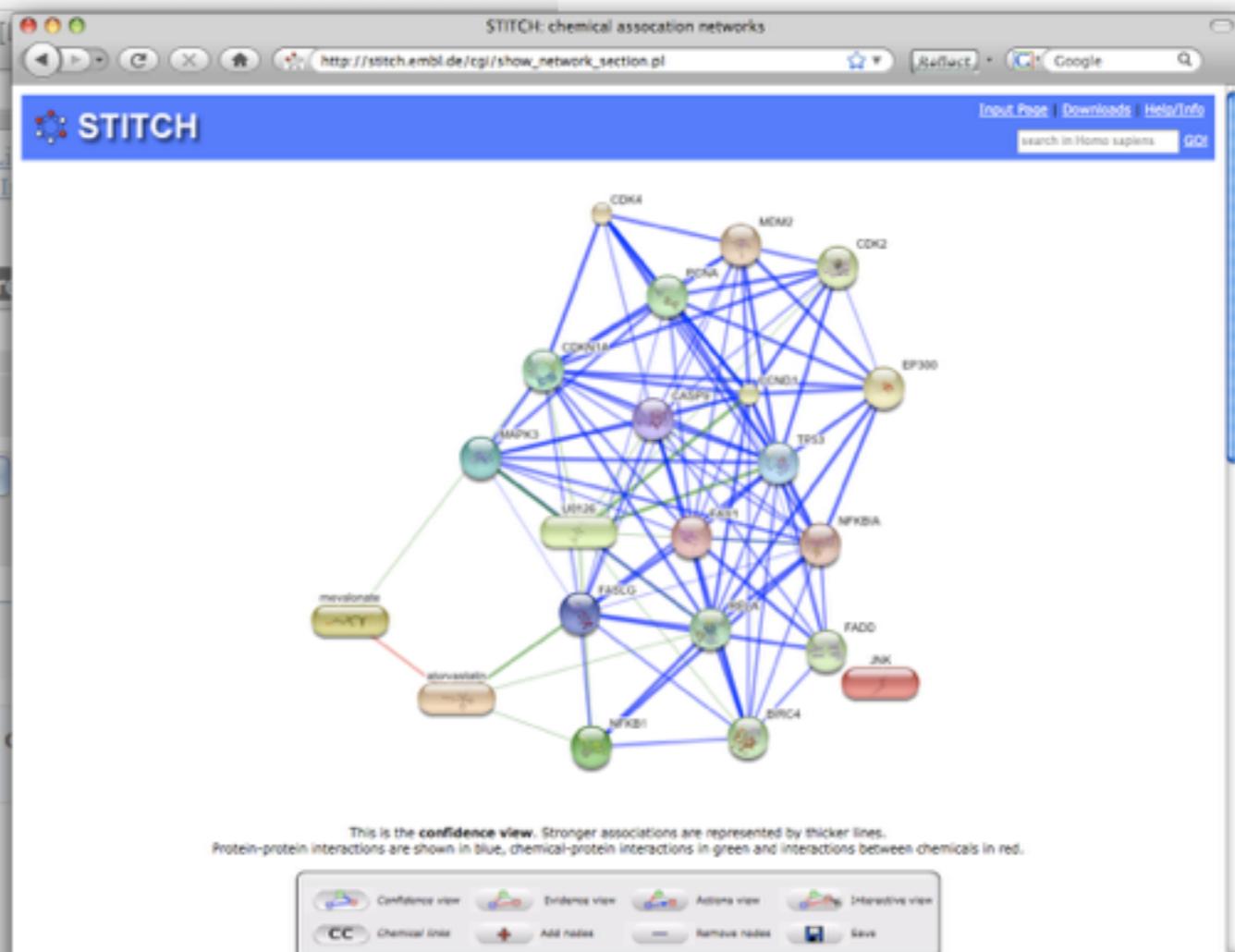
# Choose interactions

Atorvastatin induces apoptosis by a caspase-9-dependent pathway: an in vitro study on activated rat hepatic stellate cells.

Aprigliano I, Dudas J, Ramadori G, Saile B.

Department of Internal Medicine, Section of Gastroenterology and Endocrinology, University of Göttingen, Göttingen, Germany.

BACKGROUND: Statins are shown to have cholesterol-independent properties such as anti-inflammation and immunomodulation. Activated hepatic stellate cells (HSCs) acquire the capacity to synthesize matrix proteins in damaged liver. We tested the hypothesis that atorvastatin may be capable of inducing apoptosis in HSCs. METHODS: Primary cultures of rat HSCs were exposed to atorvastatin, mevalonic acid and U0126. Quantification of living, apoptotic and necrotic HSCs was performed by flow cytometry and laser-scan microscopy. Cell-cycle analysis was performed by flow cytometry. Pro- and anti-apoptotic factors were investigated by Western blot and electrophoresis mobility shift assay. Protease activity of caspases was calculated using a colorimetric kit. RESULTS: Atorvastatin leads to a G2 - arrest and induces apoptosis in activated HSCs. Atorvastatin-mediated apoptosis could be blocked by co-administration of mevalonic acid and U0126. No effects of atorvastatin on gene expression of CD95, CD95L, NF-kappaB, p53 and p21WAF1 could be observed. Atorvastatin - induced apoptosis in activated HSCs is related to an increased protease activity of caspase-9 and -3. Gene expression of the major proteins of the bcl-system shows that truncated Bid is involved in apoptosis mediated by atorvastatin. By blocking the extracellular signal-regulated protein kinase (ERK1 / 2) activation by adding U0126, we could prevent the apoptosis induced by atorvastatin. By Western blot we could not detect any change in the activation of c-jun N-terminal kinase (JNK). CONCLUSIONS: Atorvastatin induces apoptosis in activated HSCs acting through an ERK-dependent cleavage of Bid and a highly increased protease activity of caspase-9 and -3. JNK is not involved in atorvastatin-mediated apoptosis in HSCs.



- [View Page Source](#)
  - [View Page Info](#)
  - [Look up on Answers.com](#)
  - [biobank](#)
  - [Inspect Element](#)
  - [Reflect](#)
  - [Add Word to Reflect](#)
  - [Remove Word from Page](#)
  - [Interactions](#)
  - [translator](#)
  - [Web Developer](#)
- Recent Activity

# OnTheFly (<http://onthefly.embl.de>)

**OnTheFly 2.0 Beta**  
*Automated Annotation of Scientific Documents*

[Home](#) | [About](#) | [Help](#)

## Welcome to OnTheFly!

Automatically annotate files and extract biological information

Select Organism:

**Drag and Drop Files Here**



no files selected



[Contact Us](#) | [EMBL](#) | [MARBiGEN](#) | [HCMR](#) | [UOC](#)

**NCBI** **PubMed**

All Databases PubMed Nucleotide

Search PubMed for [PubMed]

Limits Preview/Index History Clipboard

Display AbstractPlus Show 20

All: 1 Review: 0

**1: Cell. 2008 Jun 13;133(6):1006-18.**

Comment in: Cell. 2008 Jun 13;133(6):958-61.

Chemokine signaling via the CXCR2 receptor reinforces senescence.

Acosta JC, O'Loghlen A, Banito A, Guijarro MV, Augert A, Raguz S, Fumagalli M, Da Costa M, Brown C, Popov N, Takatsu Y, Melamed J, d'Adda di Fagagna F, Bernard D, Hernando E, Gil J.

Cell Proliferation Group, MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College, Hammersmith Campus, W12 0NN London, UK.

Cells enter senescence, a state of stable proliferation arrest, in response to both cellular stresses, including telomere erosion, DNA damage, and oncogenic signaling, which acts as a barrier against malignant transformation *in vivo*. To identify genes controlling senescence, we conducted an unbiased screen for small hairpin RNAs that extend the life span of primary human fibroblasts. Here, we report that knocking down the chemokine receptor CXCR2 (IL8RB) alleviates both *exogenously* induced senescence (OIS) and diminishes the DNA-damage response. Conversely, ectopic expression of CXCR2 results in premature senescence via a per se mechanism. Cells undergoing OIS secrete multiple CXCR2-binding chemokines, a program that is regulated by the NF-kappaB and C/EBPbeta transcription factors and coordinately induces CXCR2 expression. CXCR2 upregulation is also observed in preneoplastic lesions *In vivo*. These results suggest that senescent cells activate a self-reinforcing network in which CXCR2-binding chemokines reinforce growth arrest.

PMID: 18555777 [PubMed - indexed for MEDLINE]

**Reflect**  
Add 'Word' to Reflect  
**Remove 'JUN' from Page**  
Interactions  
translator  
Web Developer

**Back** **Forward** **Reload** **Stop** **Details**

**Bookmark This Page** **Save Page As...** **Send Link...**

**View Background Image** **Select All**

**View Page Source** **View Page Info**

**Properties** **Look up on Answers.com**

**bioRxiv** **bioRxiv**

**Inspect Element**

**Reflect**  
Add 'Word' to Reflect  
**Remove 'JUN' from Page**  
Interactions  
translator  
Web Developer

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journal

Search PubMed for [PubMed]

Limits Preview/Index History Clipboard Details

Display AbstractPlus Show 20

All: 1 Review: 0

**1: Cell. 2008 Jun 13;133(6):1006-18.**

Comment in: Cell. 2008 Jun 13;133(6):958-61.

Chemokine signaling via the CXCR2 receptor reinforces senescence.

Acosta JC, O'Loghlen A, Banito A, Guijarro MV, Augert A, Raguz S, Fumagalli M, Da Costa M, Brown C, Popov N, Takatsu Y, Melamed J, d'Adda di Fagagna F, Bernard D, Hernando E, Gil J.

Cell Proliferation Group, MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College, Hammersmith Campus, W12 0NN London, UK.

Cells enter senescence, a state of stable proliferation arrest, in response to both cellular stresses, including telomere erosion, DNA damage, and oncogenic signaling, which acts as a barrier against malignant transformation *in vivo*. To identify genes controlling senescence, we conducted an unbiased screen for small hairpin RNAs that extend the life span of primary human fibroblasts. Here, we report that knocking down the chemokine receptor CXCR2 (IL8RB) alleviates both *exogenously* induced senescence (OIS) and diminishes the DNA-damage response. Conversely, ectopic expression of CXCR2 results in premature senescence via a per se mechanism. Cells undergoing OIS secrete multiple CXCR2-binding chemokines, a program that is regulated by the NF-kappaB and C/EBPbeta transcription factors and coordinately induces CXCR2 expression. CXCR2 upregulation is also observed in preneoplastic lesions *In vivo*. These results suggest that senescent cells activate a self-reinforcing network in which CXCR2-binding chemokines reinforce growth arrest.

PMID: 18555777 [PubMed - indexed for MEDLINE]

**Reflect**  
**Add 'NF-kappaB' to Reflect**  
Remove 'NF-kappaB' from Page

**Interactions**

**translator**

**Web Developer**

**NF-kappaB**

Protein Wikipedia Add Synonym Help

Add "NF-kappaB" as a:

Type: ▼ Protein Synonym

Identifier or Name:

Organism:

Matches:  IKBKG (ENSG00000073009)

Add to Reflect

# Reflect Dictionaries (43 GByte RAM)

	Entities (million)	Synonyms (million)	Orthographics expansion
Proteins (640 organisms)	2.6	21	82
Small molecules	7.4	24	

## Tagging Speed

(Sun Fire X4150 Server, 16 core, 64 GByte RAM)

Tagging time in seconds

Full scientific paper (~ 10,000 words)

0.3 seconds

Typical web page (<1,000 words)

0.075



Tagging usually faster than data transfer



**FuGENE® HD Transfection Reagent**  
Now available from Promega [Learn More](#)

Copyright © 2009 Elsevier Inc. All rights reserved.  
*Cell*, Volume 139, Issue 6, 1084-1095, 11 December 2009

doi:10.1016/j.cell.2009.11.015

<b>Article Outline</b>	X
Summary	
Introduction	
Results and Discussion	
Experimental Procedures	
Acknowledgments	
Accession Numbers	
Supplemental Data	
References	
Publication Information	

[Previous Article](#) [Table of Contents](#) [Next Article](#)

## Article

### The Structural Basis for mRNA Recognition and Cleavage by the Ribosome-Dependent Endonuclease RelE

Cajetan Neubauer<sup>1,4</sup>, Yong-Gui Gao<sup>1,4</sup>, Kasper R. Andersen<sup>2,4</sup>, Christine M. Dunham<sup>1,5</sup>, Ann C. Kelley<sup>1</sup>, Jendrik Hentschel<sup>1</sup>, Kenn Gerdes<sup>3</sup>, V. Ramakrishnan<sup>1</sup>, and Ditlev E. Brodersen<sup>2</sup>,

<sup>1</sup> MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK

<sup>2</sup> Department of Molecular Biology, Aarhus University, DK-8000 Aarhus C, Denmark

<sup>3</sup> Institute for Cell and Molecular Biosciences, The Medical School, University of Newcastle, Newcastle NE2 4HH, UK

Corresponding author

Corresponding author

<sup>4</sup> These authors contributed equally to the work

<sup>5</sup> Present address: Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA

#### Summary

Translational control is widely used to adjust gene expression levels. During the stringent response in bacteria, mRNA is degraded on the ribosome by the ribosome-dependent endonuclease, RelE. The molecular basis for recognition of the ribosome and mRNA by RelE and the mechanism of cleavage are unknown. Here, we present crystal structures of *E. coli* RelE in isolation (2.5 Å) and bound to programmed *Thermus thermophilus* 70S ribosomes before (3.3 Å) and after (3.6 Å) cleavage. RelE occupies the A site and causes cleavage of mRNA after the second nucleotide of the codon by reorienting and activating the mRNA for 2'-OH-induced hydrolysis. Stacking of A site codon bases with conserved residues in RelE and 16S rRNA explains the requirement for the ribosome in catalysis and the subtle sequence specificity of the reaction. These structures provide detailed insight into the translational regulation on the bacterial ribosome by mRNA cleavage.

#### INTRODUCTION

Rapid adaptation to environmental stress is vital for free-living bacteria. During deprivation of nutrients, uncharged transfer RNAs (tRNAs) bind to the stalled ribosomes and catalyzes synthesis of the signal nucleotide, (p)ppGpp. This alarmone regulates the stringent response, a far-reaching adaptation (Potrykus and Cashel, 2008). The stringent response also leads to activation of RelE, an effective inhibitor of protein synthesis. Under normal physiological conditions, RelE is inactive (Christensen and Gerdes, 2003; Christensen et al., 2001; Galvani et al., 2001; Gotfredsen and Gerdes, 1998; Li et al., 2009; Overgaard et al., 2009). Degradation of RelE by Lon protease, RelE is able to bind the ribosome and specifically cleave messenger RNA (mRNA) in the A site (Christensen and Gerdes, 2003; Pedersen et al., 2003).

Such toxin-antitoxin pairs are very common in bacteria, and the RelE superfamily also encompasses HigB, YoeB, YafQ, and YhaV, associated with a variety of cellular processes (Christensen and Gerdes, 2003; Grady and Hayes, 2003; Prysak et al., 2009; Schmidt et al., 2007). The crystal structure of YoeB showed that its fold and catalytic mechanism are similar to RelE (Li et al., 2009; Takagi et al., 2005) but lacks the conserved catalytic histidine and glutamic acid. RelE has an intrinsic nuclease activity of the ribosome (Garza-Sánchez et al., 2008; Hayes and Sauer, 2003; Kamada and Hanaoka, 2005; Li et al., 2009; Suncharoen et al., 2009). Pausing ribosomes are recovered for rescue by tmRNA in the absence of stringent response factors (Hayes and Sauer, 2003; Kamada and Hanaoka, 2005). RelE-induced cleavage is most likely a result of the combined action of RelE-like endonucleases and exonucleases like RNase II (Garza-Sánchez et al., 2009). RelE-induced cleavage occurs at the second nucleotide of the codon, although it is occasionally also seen after the third nucleotide and, upon peptide release, even in the E site (Pedersen et al., 2003). The mRNA cleavage efficiency, with the UAG and UGA stop codons and sense codons like UCG and CAG, among the most efficiently cleaved (Pedersen et al., 2003). Together,

**RelA**

Protein Chemical Help  
relA (b2784) *E. coli* Edit  
(p)ppGpp synthetase; ppGpp synthetase I; ATP:GTP 3'-pyrophosphotransferase  
Literature Sequence Structure Locus Domains  
MVAVRSAHINKAGEFDPEKNIASL GITSQKSCECLAETHAYCLOOTQ(4)  
  
No information available  
(P)ppGpp synthetase /GTP pyrophosphokinase; In eubacteria ppGpp (guanosine 3'-diphosphate 5'-diphosphate) is a mediator of the

#### Reflect

Reflect highlights proteins and small molecules and provides links to information-rich summaries.

Cell is using Reflect on a trial basis.  
[Tell us what you think.](#)

[Click here to turn off Reflect.](#)

Add/View Comments (0)

#### Article Information

[PDF \(1796 kb\)](#)

[Full Text with Large Figures](#)

[Supplemental Data](#)

[Export Citation](#)

[Request permission](#)

#### PubMed

[Articles by Cajetan Neubauer](#)

[Articles by Ditlev E. Brodersen](#)

#### Related Articles

[RAG Proteins Shepherd Double-Strand Breaks to a Specific Plat...](#)

[Joining-Deficient RAG1](#)

[Mutants Block V\(D\)J Recombination In...](#)

[Separation-of-Function Mutants Reveal Critical Roles for RAG...](#)

[...more](#)

#### Get Bookmark

[Add Bookmark](#)

#### Related Resources from Featured Advertisers

View datasheets, protocols and more for chromatin & nuclear signaling products

[Abcam](#)

QuikChange® site directed and random mutagenesis kits

[Agilent \(Stratagene Products\)](#)

[More»](#)

## Jackson Lab Finds Text-Mining Software Can Speed Some Database Curation Tasks

January 08, 2010

Newsletter: [BioInform](#)  
[BioInform - January 8, 2010](#)

By Vivien Marx

**Researchers at the Jackson Laboratory** have found that integrating text-mining tools into the Mouse Genome Informatics biocuration workflow has improved the productivity of staffers by around 50 percent for certain steps in the curation workflow.

The team found that, when combined, OntheFly and Reflect "offer functionality we are able to use in-house," Dowell said, adding that the MGI team has begun using them along with ProMiner.....

"Right now we are seeing about a 50 percent increase in the productivity of our staff indexers,"

One of the major difficulties in biomedical literature-based curation is the unambiguous identification of biological and chemical entities within text. Reflect and OnTheFly both offer a significant degree of assistance to curators in this area, providing candidate database identifiers and useful supplementary information on such entities which speeds their identification by curators and the creation of a curated entry. Reflect is also useful as an authoring tool. Curators can examine tagged versions of their own UniProt entries and quickly identify inconsistencies in nomenclature which when corrected would both significantly aid comprehension by human users and facilitate automatic tagging of the text entities by Reflect

# Mutation Finder

# OUTLINE

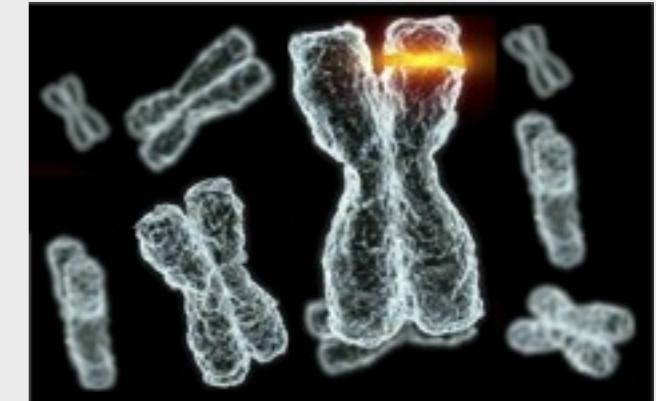
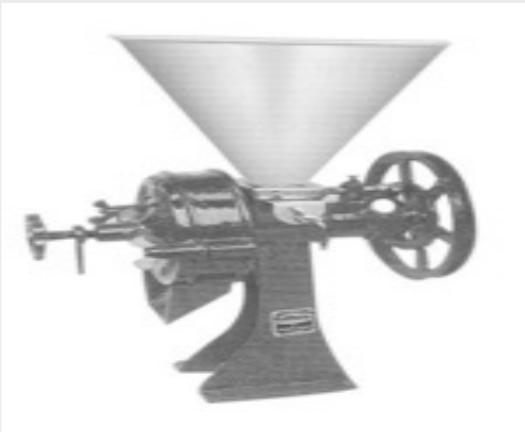
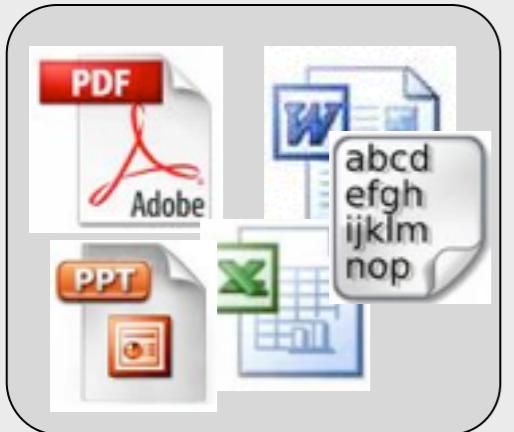
- Introduction and Motivation.
- MutationLocator Overview.
- Extracting mutations and Effects.
- Linking mutation to real Sequences.
- Validation of the Information.
- Representation of the information: 2D and 3d models.

# INTRODUCTION

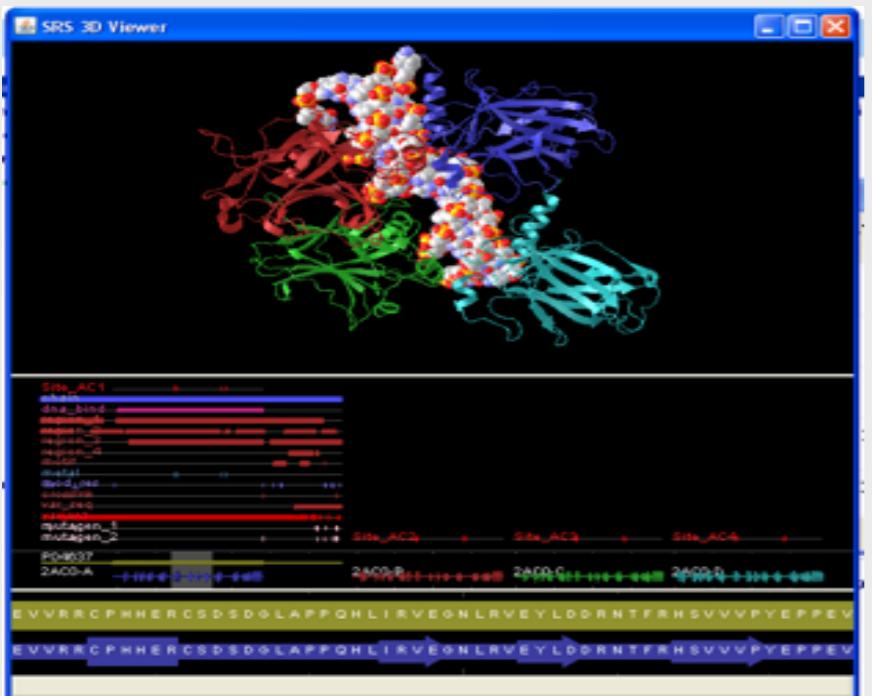
- Mutation information available in database like UniProt, OMIM etc
- But lots of mutations buried in the vast amount of literature (~ 19.9 million PubMed papers)
- Our aim: to extract the mutations from the literature and bring them to the 3D structure context
- Developed a web based system exclusively for HIV proteins, plan to extend to full spectrum of species

# MutationLocator Overview

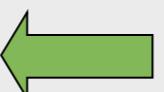
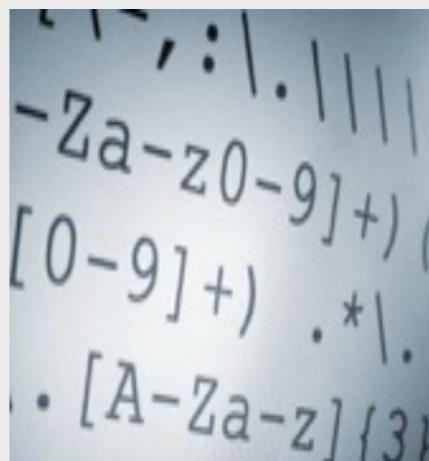
## Text mining



## SRS3D



## Matching mutations to sequence



A screenshot of a web-based application showing a list of mutations. Each entry includes a score, a mutation line, and a snippet of text from a document. The snippets provide context for each mutation, such as "A mutation in Alpha Helix 3 of CA Renders Human Immunodeficiency Virus Type 1 Cyclosporin A Resistant and Dependent: Residue by a Second-Gate Substitution in a Distal Region of CA".

## Document context

# Overview for MutationLocator

Total number of 100.000 papers extracted from pubmed central and other freely available sources which mention “HIV” or “AIDS” and other synonyms somewhere in the full text:

- Science,
- The Journal of general virology,
- The Journal of biological chemistry,
- The Journal of antimicrobial chemotherapy,
- The journal of medical investigation : JMI
- .....

# Extracting mutations and Effects

## Mutation normalization

Gold standard:

Format describing a mutation using the single-letter amino acid code of the wild type,  
followed by the residue position and its mutated amino acid form

e.g. E75D.

## Mutation recognition

Detect each description of a mutational change described in a document, including  
abbreviated and full-text forms, as well as nominal and pronominal references.

Distinguish between single-point mutations and mutation series, as well as insertions or  
deletions.

“The definitions of what precisely constitutes an individual mutation entity is even more  
difficult than for a protein or organism entity, due to the high number of variations  
possible for describing mutational changes performed on a protein.”

Rebholz –Schuhmann et al.

Automatic extraction of mutations from Medline and Cross-validation with OMIN.

# Extracting mutations and Effects

**Table 1.** Several examples of phrases describing mutations found in Medline

---

Arg506 to Gln
valine 804→leucine
Ile15 to Thr15
Pro12Ala
arginine(3500)–glutamine
C282Y
A1166→C
677C→T
1166A/C, 359 (Ile/Leu)
Nucleotide 383T→C
codon 113 and His→Arg
Cys/Val343
Val→Ala at codon 113
IVS1–2A→G
codon 241 and codon 247, where the single base changes from C to T
Methionine to threonine substitution at residue 235
Methionine for valine at position 30
Ser→Leu change at amino acid 217
Heterozygosity for the IVS-I-5 (G→C) mutation
A fourth mutation, 433–2(A→G) transition, was identified at the splice-acceptor site in intron 2

---

**Rebholz –Schuhmann et al.**

**Automatic extraction of mutations from Medline and Cross-validation with OMIM. 2003**

# Extracting mutations and Effects

## Three main conditions for a meaningful Effect:

### a) A function related to the virus cycle:

dimerization, packaging, uncoating, activity, maturation, stability, infectivity, replication, integrity, regulation, binding, assembly, formation, interaction, budding, etc

### b) An Action (As represented in uniprot):

Abolish, affect, abrogate, block, decrease, increase, defect, diminish, disrupt, enhance, impair, inhibit, loss (loose), prevent, produce, process, delete, restrict

### c) A Key word related to the protein of interest:

mutation, mutant, n-terminus, n-terminus, ca, cactd, cantd, capsid, nc, nucleo, ma, matrix, matrix, zinc, finger, linker, function, effect, terminal, p6, p6-gag

$$\text{Score} = f(a, b, c)$$

The mutation E24T in the CA stops virus replication.

# Extracting mutations and Effects

## Packaging

Most mutations inhibited gRNA packaging as much, or nearly as much, as deleting the whole finger. F16A and 2GAF1 appeared particularly inhibitory. One striking exception was N17K: it stimulated gRNA packaging 2 to 3-fold relative to WT ([Table 2](#)).

Virus stability was reduced  $\geq$  5-fold by mutation 3EF1. Most of the other mutations, including  $\Delta$ F1, had approximately a 2-fold inhibitory effect. RT packaging was little affected by mutations H23C and C28H, and reduced 1.5 to 3-fold, depending on the mutation, by the other mutations ([Table 2](#)).

## Zinc finger 2

Twelve mutations were studied ([Table 1](#)). Each mutation blocked viral replication, except 2KAF2 (substitution of alanine for Lys41 and Lys47) and K38N, which were about 100 times and 10,000 times less infectious than WT, respectively ([Fig. 2D](#)). Thus point mutations W37A, H44A, H44C, C49H or even W37F (W standing for tryptophan) had devastating effects on viral replication, even though W37F simply replaced an aromatic residue by another that was wild-type in zinc finger 1.

“Mapping of nucleocapsid residues important for HIV-1 genomic RNA dimerization and packaging”. Jafar Kafaie et al. March 2008

# Extracting mutations and Effects

Table 2

Effect of nucleocapsid protein mutations on virus infectivity, genomic RNA dimerization, genomic RNA packaging, packaging of reverse transcriptase activity, virus stability and Pr55gag proteolytic maturation

Region		Construct name <sup>a</sup>	Viral replication <sup>b</sup>	gRNA dimerization <sup>c</sup>	gRNA packaging <sup>c</sup>	RT packaging <sup>c</sup>	Virus stability <sup>c</sup>	Pr55gag processing <sup>d</sup>
	1	HXB2	+	100	100	100	100	97
	2	PR-	-	60±3	nd	nd	nd	0
N-Terminus	3	R7	-	77±2	48±2	nd	46±5	73
	4	R7E	-	79±5	8±2	19±2	32±5	60
	5	N+	+	102±3	23±4	93±5	239±33	100
Zinc finger 1	6	3AF1	-	79±1	42±2	28±4	nd	98
	7	3EF1	-	68±4	22±3	24±4	15±2	80
	8	N17K	+/-	97.5±2	250±73	59±4	69±7	97
	9	F16A	-	76±2	15±2	nd	nd	77
	10	2GAF1	-	73±3	11±2	nd	56±6	91
	11	H23A	-	69±4	41±3	39±6	36±4	89
	12	H23C	-	91±1.5	30±2	71±2	25±4	98
	13	C28H	-	81±1.5	36±11	82±3	nd	99
	14	C28S	-	81±3.5	50±5	60±5	39±12	96
	15	NC2-2	-	86±2	34±3	44±3.5	nd	93
	16	ΔF1	-	72.5±3	25±10	nd	43±7	54
Linker	17	S3	+/-	94±3.5	59±13	nd	97±14	98
	18	S3E	-	45±2	28±5	10±3	24±4	64
	19	FVI	-	74±2	40±6	nd	nd	97
	20	P31A	+	96.5±2.5	33±4	nd	84±10	98
	21	ΔAP	-	75±2	34±4	95±11	12±3	90
	22	LL	-	72±3	65±4	45±5	42±14	96
	23	ΔLinker	-	75±2	16±6	nd	21±4	89
	24	L+	+/-	96±3	20±7	nd	100±15	98

- Javascript must be enabled!
- Select the mutation and the give us the additional data (under the paper) to go on...

Mutation	Score	Mutation line
All		
None		
<input type="checkbox"/> A92E	1	<a href="#">In additional studies, we confirmed that the infectivity of the A92E mutant virus was enhanced by CsA in HeLa-P4 but not in HOS cells, as previously reported (C)</a>
<input type="checkbox"/> A92E	1	<a href="#">A105T rescued the impaired single-cycle infectivity and replication defects of both T54A and A92E mutants</a>
<input type="checkbox"/> A92E	1	<a href="#">The A92E and G94D mutations do not inhibit incorporation of CypA into HIV-1 particles and thus do not appear to block CypA binding to CA</a>
<input type="checkbox"/> A92E	1	<a href="#">We also identified the second-site mutation A105T that complements the T54A infectivity impairment, as well as that of the previously characterized CsA-resistant/dependent mutant A92E</a>
<input type="checkbox"/> A92E	1	<a href="#">Our results demonstrated that CsA enhances infection of viruses containing mutations in two distinct domains of CA: T54 in helix 3 and A92E in the CypA-binding loop</a>
<input type="checkbox"/> A92E	1	<a href="#">Thus, the A105T mutation enhanced the replication of both T54A and A92E HIV-1 mutants</a>
<input type="checkbox"/> A92E	1	<a href="#">Inhibition of the CA-CypA interaction by CsA is detrimental to the infectivity of most wild-type HIV-1 isolates, yet CsA promotes infection by the CA mutants A92E, G94D, and T54A in some cells</a>
<input type="checkbox"/> A92E	1	<a href="#">These results indicate that the infectivity of the T54A mutant, like that of A92E, is enhanced by CsA in a target cell-specific manner</a>
<input type="checkbox"/> A92E	1	<a href="#">Our findings that T54A replication is enhanced by CsA, and that the A105T mutation restores replication to T54A and A92E mutant viruses, suggest that these</a>

page 1 of 8    100%    Index Pages Text

► Page 1  
► Page 2  
► Page 3  
► Page 4  
► Page 5  
► Page 6  
► Page 7  
► Page 8

JOURNAL OF VIROLOGY, Apr. 2007, p. 3749–3756  
0022-538X/07/\$08.00 © doi:10.1128/JVI.02634-06  
Copyright © 2007, American Society for Microbiology. All Rights Reserved.  
Vol. 81, No. 8

## A Mutation in Alpha Helix 3 of CA Renders Human Immunodeficiency Virus Type 1 Cyclosporin A Resistant and Dependent: Rescue by a Second-Site Substitution in a Distal Region of CA

Ruifeng Yang and Christopher Aiken\*

Department of Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, Tennessee

Received 28 November 2006/Accepted 22 January 2007

The replication of many isolates of human immunodeficiency virus type 1 (HIV-1) is enhanced by binding of the host cell protein cyclophilin A (CypA) to the viral capsid protein (CA). The immunosuppressive drug cyclosporine A (CsA) and its nonimmunosuppressive analogs bind with high affinity to CypA and inhibit HIV-1 replication. Previous studies have identified two mutations, A92E and G94D, in the CypA-binding loop of CA that confer the ability of HIV-1 to replicate in the presence of CsA. Interestingly, CsA stimulates the replication of HIV-1 mutants containing either the A92E or G94D substitution in some human cell lines. Here, we show that substitution of alanine for threonine at position 54 of CA (T54A) also confers HIV-1 resistance to and dependence on CsA. Like the previously identified CsA-resistant/dependent mutants, infection by the T54A mutant was stimulated by CsA in a target cell-specific manner. RNA interference-mediated reduction of CypA expression enhanced the permissiveness of HeLa cells to infection by the T54A mutant. A suppressor mutation, encoding a substitution of threonine for alanine at position 105 of CA (A105T), was identified through adaptation of the T54A mutant virus for growth in CEM cells. A105T rescued the impaired single-cycle infectivity and replication defects of both T54A and A92E mutants. These results indicate that CA determinants outside the CypA-binding loop can modulate the dependence of HIV-1 infection on CypA.

The capsid (CA) protein of human immunodeficiency virus type 1 (HIV-1) consists of two domains, with the N- and C-terminal domains (NTD and CTD) connected by a flexible linker. The NTD contains an N-terminal -hairpin, seven -helices, and an extended loop that binds to the cellular protein cyclophilin A (CypA) (20, 29). Cryoelectron microscopy studies have revealed that the NTD forms a hexamer lattice, with the CA CTD making dimeric contacts that connect each ring to its six nearest neighbors (25). Mutagenesis studies have shown that both the NTD and CTD are essential for capsid formation and particle assembly (14, 28, 31). During particle maturation, the CA protein condenses to form a conical core around the ribonucleoprotein complex. Mutations that alter HIV-1 core morphology also reduce infectivity (14, 28, 31, 33, 38). These observations suggest that proper formation of the conical HIV-1 core is essential for the early postentry events in HIV infection.

CypA is a cellular peptidylprolyl isomerase that binds to the HIV-1 CA NTD and is incorporated into virions through interaction with an exposed loop between helices 4 and 5 in

the CypA-CA interaction is manifested following entry of the core into target cells, and incorporation of CypA into virions appears to be biologically irrelevant (21, 35). Despite many years of study, the precise role of CypA in promoting HIV-1 infection remains obscure.

G94D A92E Close

- PubMed ID : 17267487
- Keyword :
- Action :
- Function :

shown that substitutions in the CypA-G94D, confer HIV-1 resistance to infection by these mutants is enhanced in cell lines, such as HeLa and H9, but not Jurkat (1, 21). Positions 92 and 94 in the CypA-binding loop in CA, but these mutations do not affect CypA-CA binding (6), which suggests that CsA resistance is independent of this interaction. The CsA dependence implies that CypA binding to CA has a detrimental effect on infection by these mutants in some cell types.

In the present study, we analyzed in detail the phenotype of a poorly infectious HIV-1 mutant, T54A, encoding a Thr-to-Ala substitution in helix 3 of the CA NTD. We found that the effects of the T54A substitution resemble those of previously characterized CypA-binding loop mutations A92E and G94D.

# Linking mutation to real Sequences

## **Grounding tasks:**

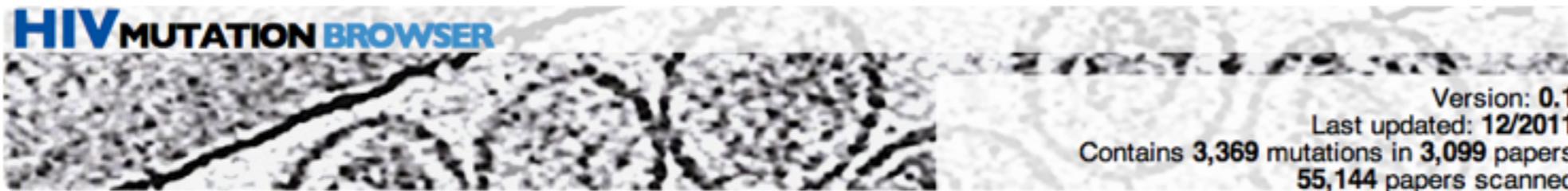
Grounding cross-links entities detected in documents with their real-world counterparts.

**Mutation Grounding: Verify and if necessary positionally correct each mutation location to match its corresponding protein's sequence as obtained from UniProtKB.**

# Linking mutation to real Sequences

- Mutations – Original Amino-acid, the position, and the change for the mutation.
- At least 1 sequence in 518415 (Swissprot Last Update)
- Identifying the organism ---->search space for HIV-1: 381 sequences.
- Score =  $f(\text{pb}, \text{N\_match}/\text{N\_Total}, \text{Distance})$
- Ranking:
  - The more mutations matching the better.
  - The closer the better.

## HIV MUTATION BROWSER

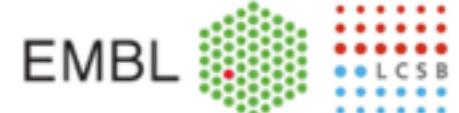


Version: 0.1

Last updated: 12/2011

Contains 3,369 mutations in 3,099 papers

55,144 papers scanned



### Browse

[gag](#)

[pol](#)

[env](#)

[tat](#)

[nef](#)

[rev](#)

[vif](#)

[vpr](#)

[vpu](#)

### News

**2/12 - Missing papers** There was a bug stopped certain papers from being shown but this has been fixed.

**1/12 - New UI** We've updated the user interface.

**28/11 - Structure offsets** Some of the offsets in the structures were wrong but this has been fixed.

### Featured Paper

[HIV-1 Vpu neutralizes the antiviral factor Tetherin/BST-2 by binding it and directing its beta-TrCP2-dependent degradation.](#)   
**Mangeat et al.** PLoS Pathogens

### Recently added

[p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease.](#)   
**Huang et al.** J Virol

### Database Statistics

Gene	Publications	Mutations
<a href="#">gag</a>	438	460
<a href="#">pol</a>	1934	1745
<a href="#">env</a>	340	672
<a href="#">tat</a>	128	106
<a href="#">nef</a>	107	134
<a href="#">rev</a>	23	40
<a href="#">vif</a>	51	106
<a href="#">vpr</a>	64	84
<a href="#">vpu</a>	14	22

Home Search About

gene: gag gag-pol env tat nef rev vif vpr vpu

view: Sequence Table

```

1 M G A R A S V L S G G E L D R W E K I R L R P G G K K K Y K L K H I T W A S R E L E R F A T N P G L L E T S E G C R Q I 60
61 L G Q L Q P S L Q T G S E E L R S L Y N T T A T L Y C V H Q R I E I K D T K E A L D K I E E E Q N K S K K K A Q Q A A A 120
121 D T G H S H Q T S Q N Y P I T Q N I Q G Q M T H Q A I S P R T L H A V V K Y T E E K A F S P E V I P M F S A L S E G A T 180
181 P Q D L N T H L N T T G G H Q A A M Q H L K E T I N E E A A E W D R T H P T H A G P I A P G Q M R E P R O S D I A G T T 240
241 S T L Q E Q I Q G W H T N N P P I P T G E I Y K R W I I L G L N K I V R M Y S P T S I L D I R Q Q P K E P F R D Y T D R F 300
301 Y K T L R A E Q A S Q E V E N V M T E T L L T Q N A N P D C K T I L K A L G P A A T L E E M M T A C Q G Y D G P G K E A 360
361 R Y T L A E A M S Q T T H S A T I M M Q R G N F R N Q R K I T K C F N C G K E G H T A R N C R A P R K K G C W K C G K E G 420
421 H Q M K D C T E R Q A N F L G K I W P S T K G R P G N F L Q S R P E P T A P P E E S F R S G T E T T T P P Q K Q E P I D 480
481 K E L Y P F L T S L R S I F G N D P S S Q

```

Restrict by mutation + reset

Show all sentences reset

Info: Structure Conservation

## K30A

[Show sentences](#)Title: Identification of residues of the Moloney murine leukemia virus nucleocapsid critical for viral DNA synthesis in vivo. [eP](#)

Authors: Gonsky, J Bacharach, E Goff, SP

Journal: J Virol

## K30E

[Show sentences](#)Title: HIV-1 Vpr oligomerization but not that of Gag directs the interaction between Vpr and Gag. [eP](#)

Authors: Fritz, JV Dujardin, D Godet, J Didier, P De Mey, J Darlix, JL Mely, Y de Boccard, H

Journal: J Virol

## K30R

[Show sentences](#)Title: Impaired processing and presentation of cytotoxic-T-lymphocyte (CTL) epitopes are major escape mechanisms from CTL immune pressure in human immunodeficiency virus type 1 infection. [eP](#)

Authors: Yokomaku, Y Miura, H Tomiyama, H Kawana-Tachikawa, A Takiguchi, M Kojima, A Nagai, Y Iwamoto, A Matsuda, Z Ariyoshi, K

Journal: J Virol

[Show sentences](#)Title: Antiretroviral drug therapy alters the profile of human immunodeficiency virus type 1-specific T-cell responses and shifts the immunodominant cytotoxic T-lymphocyte response from Gag to Pol. [eP](#)

Authors: Karlsson, AC Chapman, JM Heiken, BD Hoh, R Kallas, EG Martin, JN Hecht, RM Deeks, SG Nixon, DF

Journal: J Virol

[Show sentences](#)Title: Gag determinants of fitness and drug susceptibility in protease inhibitor-resistant human immunodeficiency virus type 1. [eP](#)

Authors: Parry, CM Kohli, A Boinett, CJ Towers, GJ McCormick, AI Pillay, D

Journal: J Virol



Restrict by mutation + reset

Show all sentences reset

Info: Structure Conservation

## K30A

[Hide sentences](#)Title: Identification of residues of the Moloney murine leukemia virus nucleocapsid critical for viral DNA synthesis in vivo. [eP](#)  
Authors: Gonsky, J Bacharach, E Goff, SP  
Journal: J Virol

- Perhaps the most likely mechanism whereby the L21A and K30A NC mutations may cause at least the initial reduction in viral DNA synthesis in vivo is by impairing the process of virion uncoating, or other steps occurring before the initiation of reverse transcription.
- K30A virus-infected cells showed a 6-day delay while L21A virus-infected cells showed an 11- to 15-day delay compared to cultures infected with wild-type or R17A control viruses (Fig. 2B and 3B).
- preceding the Cys-His box (L21A) or a lysine within the CysHis box (K30A) significantly delayed retroviral replication
- This notion is consistent with our observation that, at 20 h postinfection, cells infected by the L21A and K30A mutants show

## K30E

[Show sentences](#)Title: HIV-1 Vpr oligomerization but not that of Gag directs the interaction between Vpr and Gag. [eP](#)  
Authors: Fritz, JV Dujardin, D Godet, J Didier, P De Mey, J Darlix, JL Mely, Y de Boccard, H  
Journal: J Virol

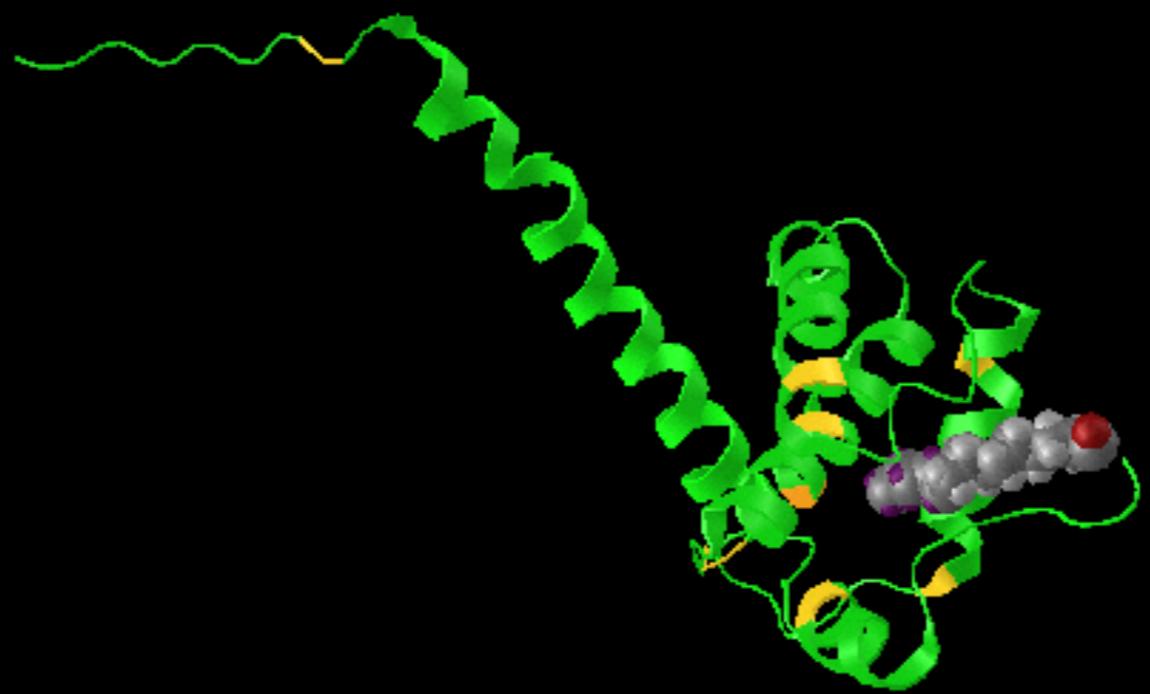
R:68.8% / other:31.2%

R L R P G G K K K Y	K L K H I V W A S R E	HXB2
R L R P G G K K K Y	M I K H L V W A R S E	ANT
R L R P G G K K K Y	K L K H V V W A S R E	U14788
R L K P G S K K K Y	M L K H I V W A S R E	RBF168
R L R P G S K K A Y	R L K H L V V W A S R E	MVP5180
R L K P G C K K K Y	R L K H L V V W A S R E	99SE_MP1300
R L K P G S K K K Y	R L K H L V V W A S R E	ANT70
R L K P G S K K K Y	R L K H L V V W A S R E	98CMU2901
R L R P G G K K K Y	M M K H L V V W A S R E	US_Manlyn
R L R P G G K K K Y	M M K H L V V W A S R E	SIVcpzMT145
Y L R P G G K K K Y	R M K H L V V W A S R E	DJO0131
V L R P G G K K K Y	R M K H L V V W A S R E	YBF30
V L R P G G K K K Y	R M K H L V V W A S R E	MPF106

# 3D Visualization

SRS 3D





Coloring: Sequence Similarity

- identical
- conserved
- non-conserved
- unaligned

lection  
empty>

site\_AC1

it\_met

chain\_1

chain\_2

peptide

n\_fing

motif

te

cod\_res

pid

05887

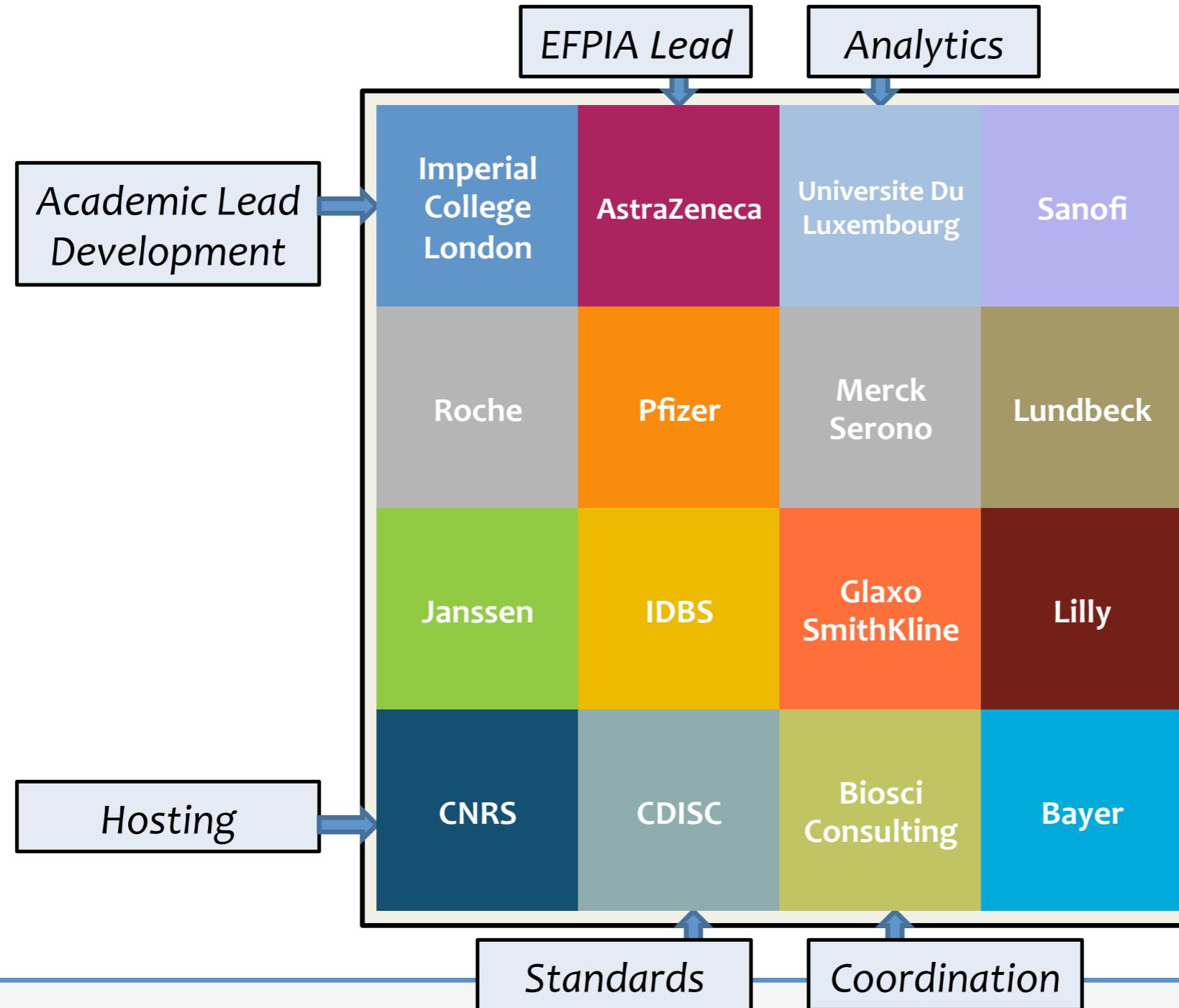
H3I-A

05887

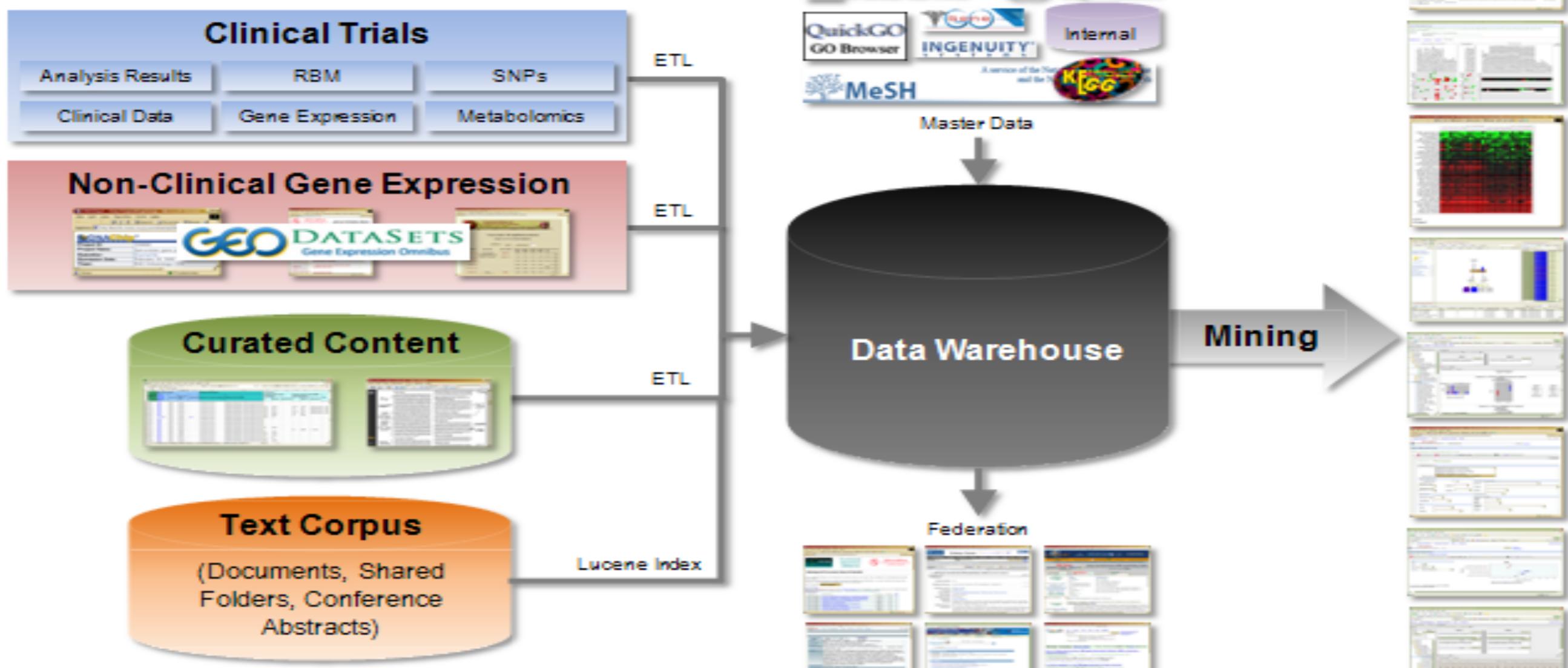
H3I-B

05887

H3I-C



## tranSMART



Measurment	Lab 1
Adiponectin	⊕
ALAT (alaninaminotransferase)	⊕
Albumin	⊕
Apolipoprotein A1	⊕
Apolipoprotein B	⊕
Apolipoprotein B/ApoA1 ratio	⊕
ASAT (aspartateaminotransferase)	⊕
Cholesterol total	⊕
C-Peptide	⊕
C-Reactive Protein hs-CRP	⊕
Creatinine	⊖
Cystatin C	⊕
Estradiole E2	⊕
Follicle Stimulating Hormone FSH	⊕
FT3 (free trijodtyronin)	⊖
FT4 (free tyroxin)	⊖
GFR (Glomerular filtration rate)	⊕
Glucose	⊕
Glutamyltransferase GGT	⊕
GOT glutamic oxaloacetate transaminase	
GPT glutamic pyrovate transaminase	
HbA1c	
HDL-cholesterol	⊕
Hemolysis	⊕
Ikterus (bilirubin)	⊕
Insulin	⊕
Interleukin 1-beta IL-1b	
Interleukin 6 IL-6	
LDL-cholesterol	⊖
Leptin LEPT	⊖
LH Lutropin	⊕
Lipoprotein A LPA	⊖
Proinsulin	⊕
Prolactin PROL	⊖
SHBG (steroid hormone binding globulin)	⊕
Testosterone	⊕
TNF-alpha (tumour necrosis factor)	
Triglyceride	⊕
TSH (thyroid stimulating hormone)	⊕
UREA (Harnstoff HST)	⊖
Uric acid (Urat Harnsaure HRS)	⊕
UU - Pt-GFR(CystC-beräkn) (see also GFR)	⊕
UU - GLP-1 (glucagon like peptide -1 in P800 tubes)	⊖
UU - FFA Free fatty acids	⊕
UU - steroid hormone pattern (12 different)	⊕

unit	Lab 2	unit	Lab 3	unit	STATUS	comment
ukat/L	⊖	µg/mL	⊕	mg/L	?	
g/L	⊕	%	⊕	g/L	OK	is it really %?
g/L	⊕	mg/dL	⊕	g/L	OK	
g/L	⊕	mg/dL	⊕	g/L	OK	
-	⊕	-	⊕	-	OK	was calculated for S and L
ukat/L	⊖		⊕	ukat/L	OK	
mmol/L	⊕	mg/dL	⊕	mmol/L	OK	
nmol/L	⊕	pmol/L?	⊕	nmol/L	OK	or pmol/mL?
mg/L	⊕	mg/dL	⊕	mg/L	OK	
mg/L	⊕	mg/L	⊕	mg/L	OK	
pmol/L	⊕	pg/mL	⊕	pmol/L	OK	
IE/L	⊕	mU/mL	⊕	U/L	OK	
	⊕	pmol/L	⊕	pmol/L	OK	
	⊕	ng/dL	⊕	ng/dL	OK	
mL/mi/1,73	⊖		⊖		?	
mmol/L	⊕	mg/dL	⊕	mmol/L	OK	
ukat/L	⊕	U/L	⊕	ukat/L	OK	
	⊕	U/L			?	
	⊕	U/L			?	
	⊖	%			?	
mmol/L	⊕	mg/dL	⊕	mmol/L	OK	exclude?
-						recalc
-						
mE/L	⊕	µU/mL	⊕	pmol/L	?	
	⊖					exclude?
	⊕	?				
	⊕	mg/dL	⊕	mmol/L	?	
	⊕	ng/mL	⊕	ng/mL	OK	
IE/L	⊕	mU/mL	⊕	U/L	OK	
	⊕	mg/dL	⊕	g/L	OK	
pmol/L	⊕	pmol/L	⊕	pmol/L	OK	Lab 1 high wrong units?
	⊕	µU/mL	⊕	mU/L	OK	
nmol/L	⊕	nmol/L	⊕	nmol/L	OK	
ng/mL	⊕	ng/mL	⊕	nmol/L	OK	recalc
	⊕	?				
mmol/L	⊕	mg/dL	⊕	mmol/L	OK	recalc
mIE/L	⊕	mU/L	⊕	mU/L	OK	
	⊕	mg/dL	⊕	mmol/L	OK	recalc
umol/L	⊕	mg/dL	⊕	umol/L	OK	recalc
ml/mi/	?					
	?					when?
	mg/dL					
	pg-ng/mL					

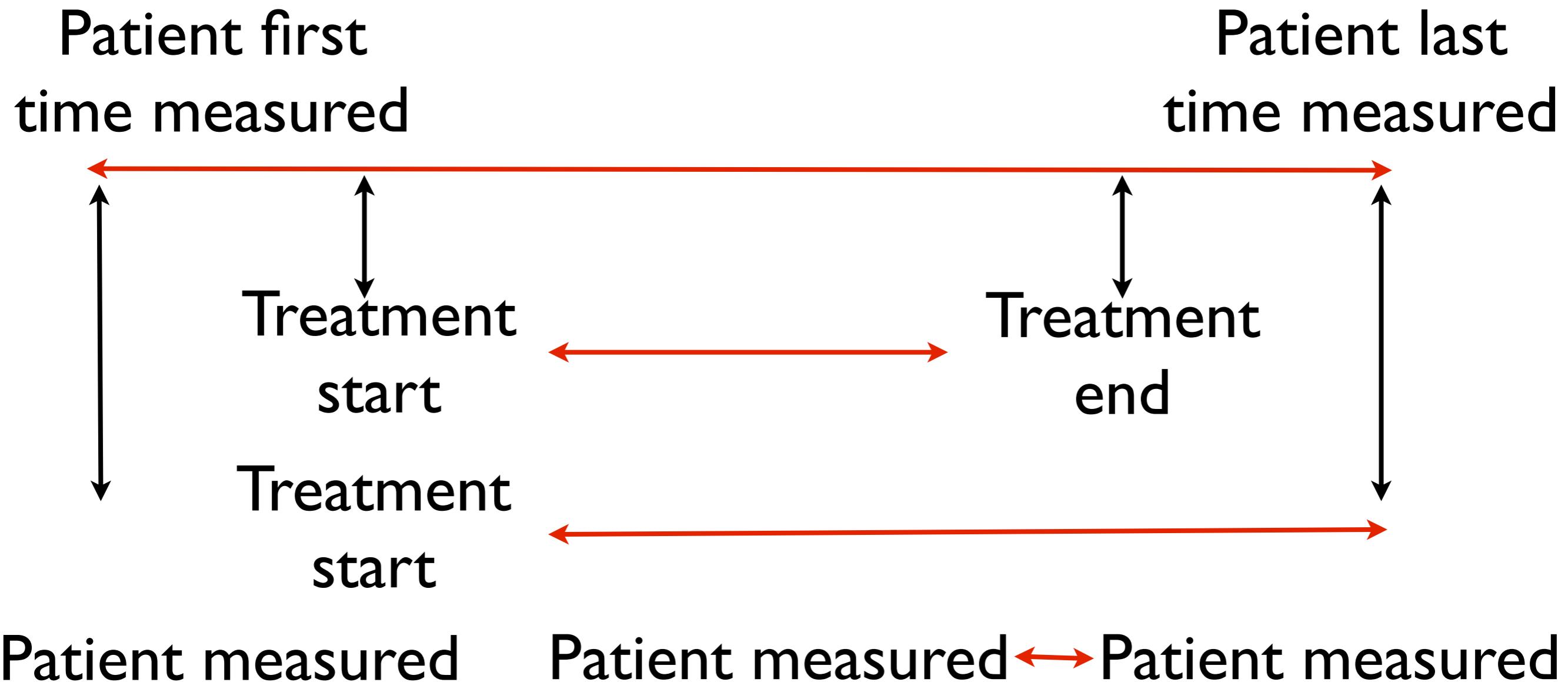
⊕ = values reported

⊖ = values NOT reported

? = values reported but without unit

STATUS = OK (good correlation between the reporting laboratories, at least 2/3)

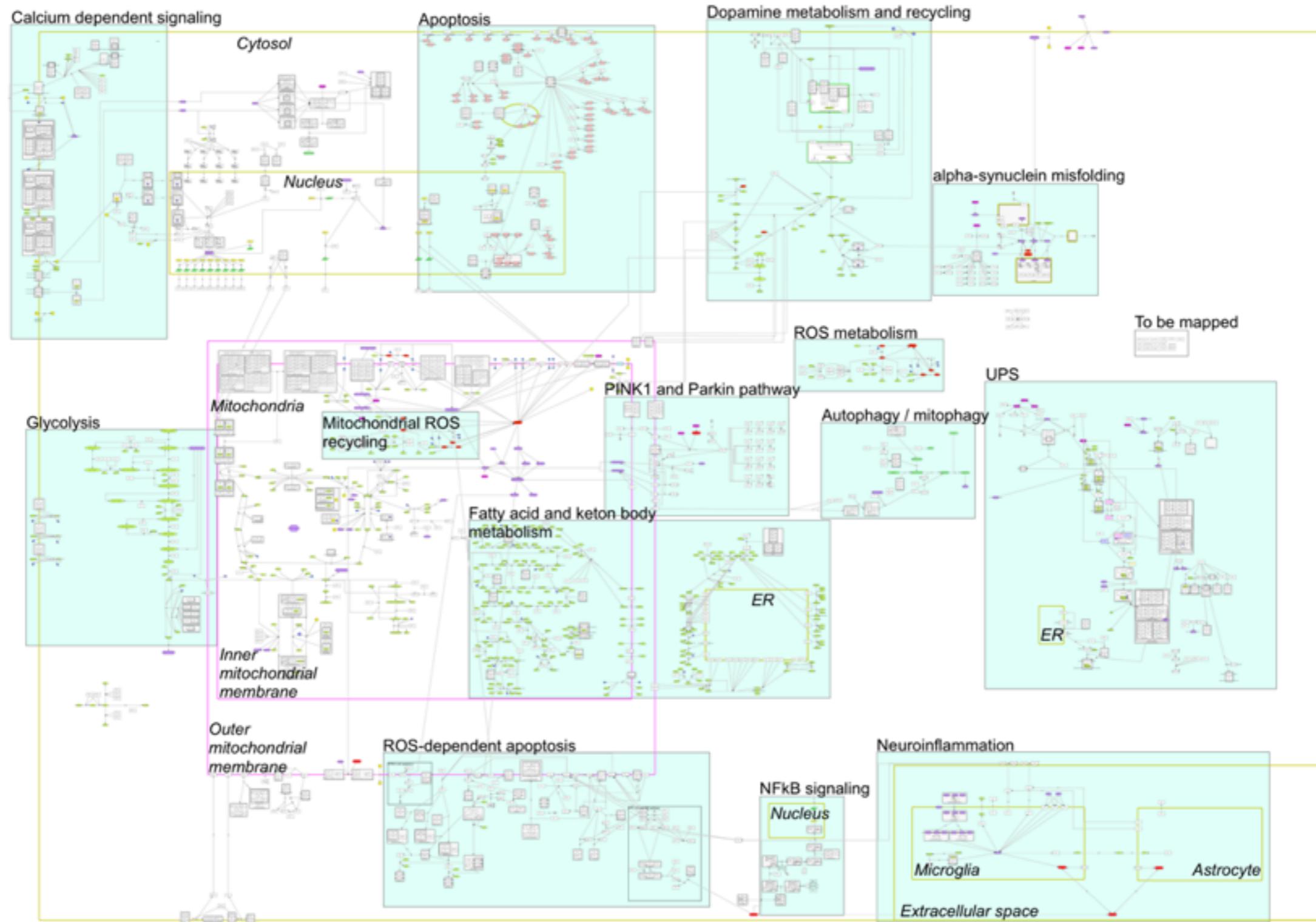
# One easy measurement: weight loss of patient !



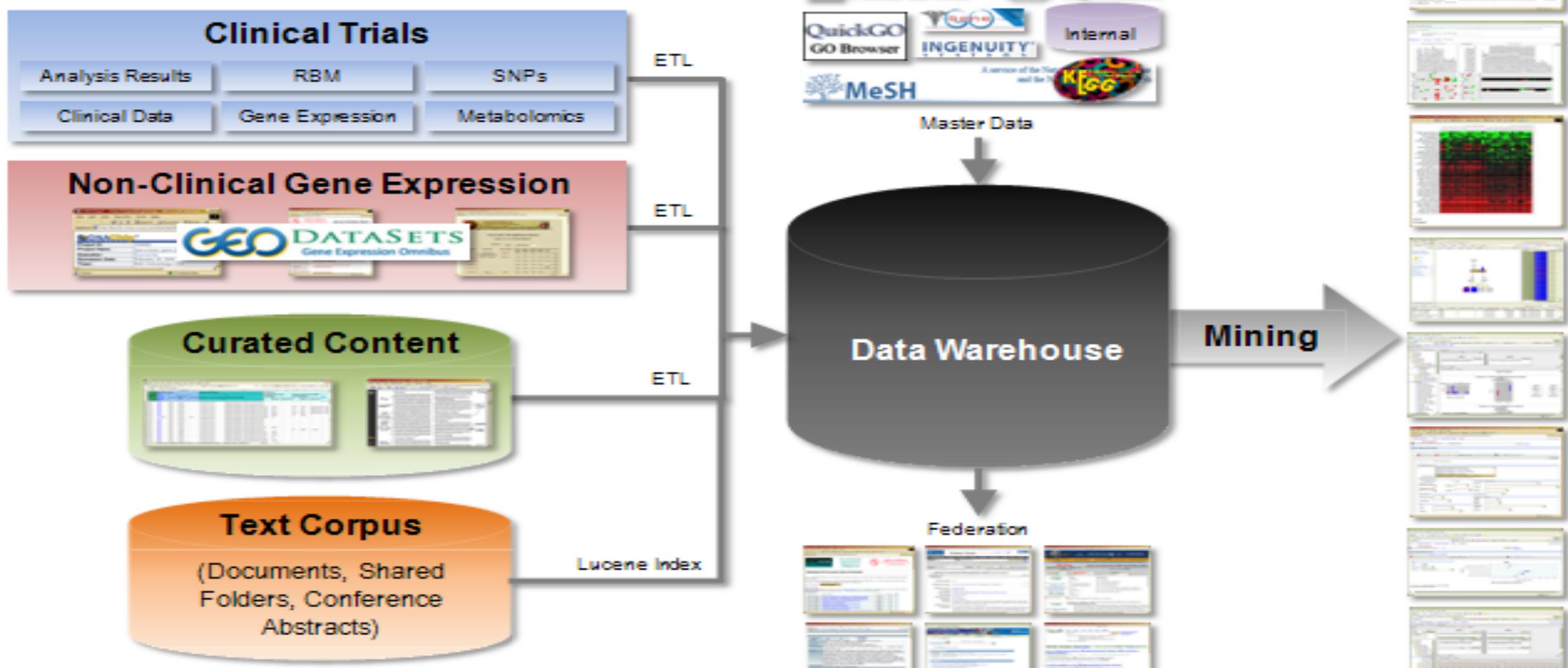
in Kg, Gramm, pounds, percentage ??

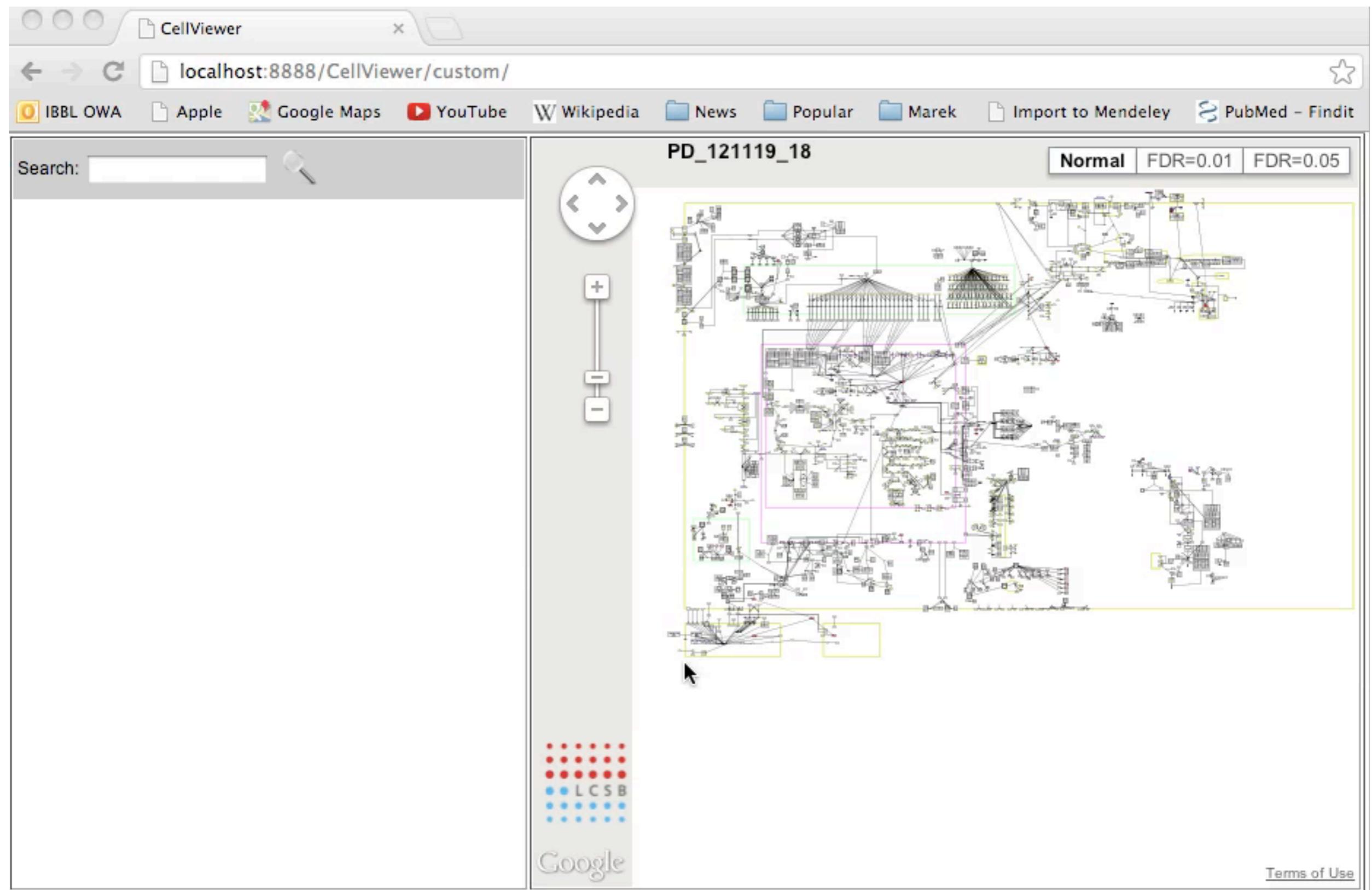
# Parkinson Disease Map

- <http://minerva.uni.lu/CellViewer/map?id=custom>



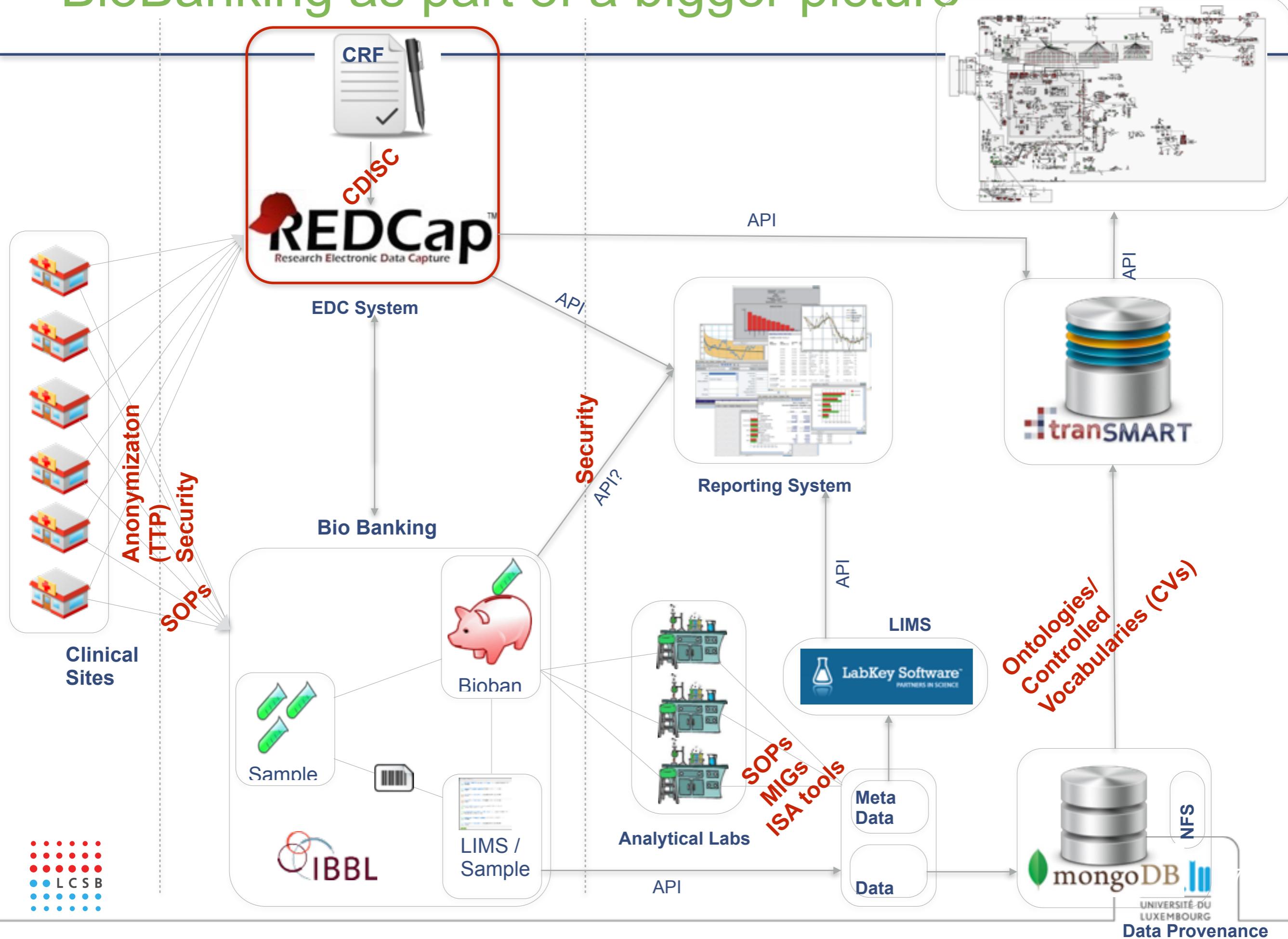
## tranSMART



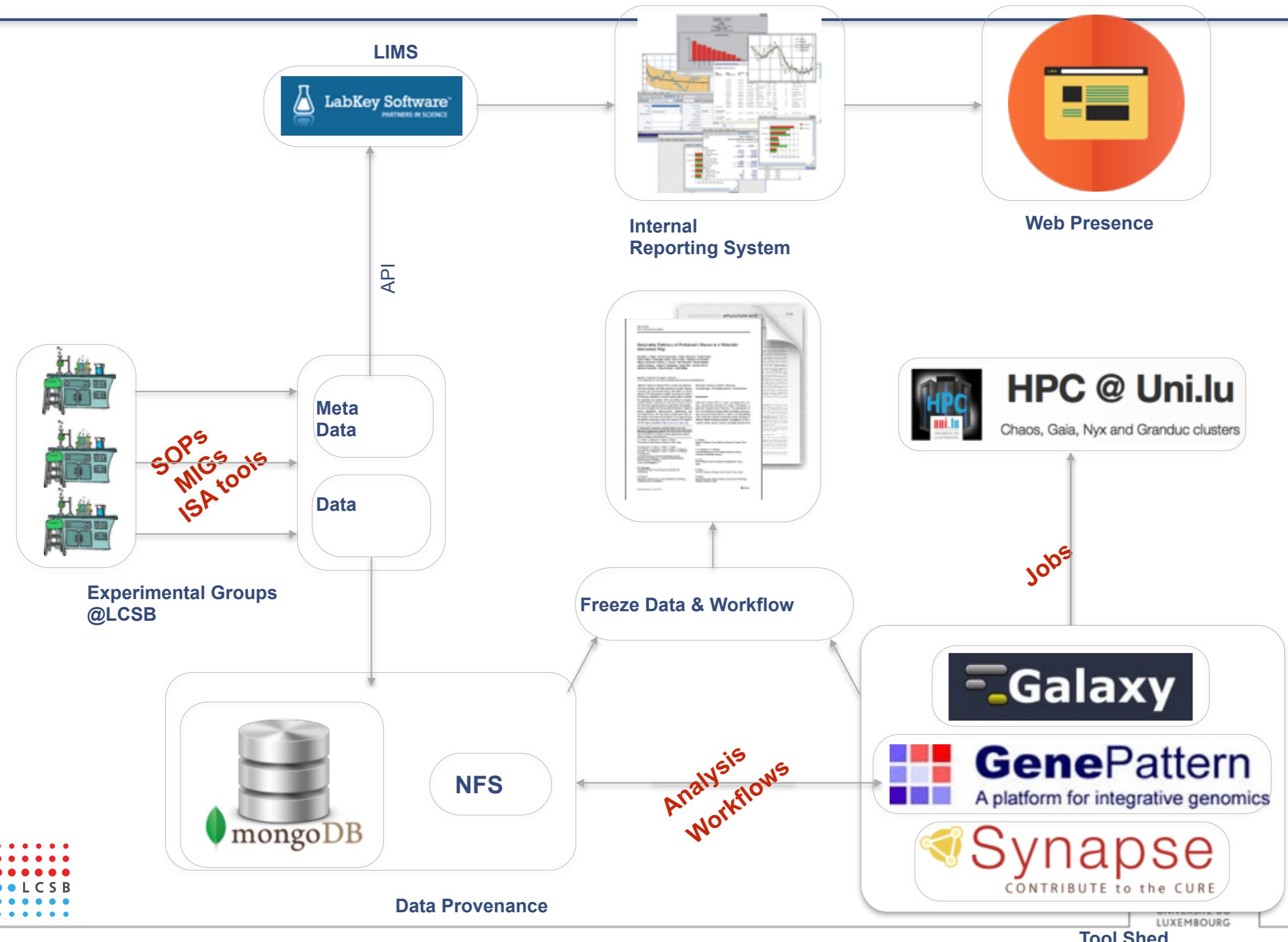


# BioBanking as part of a bigger picture

PD Map



# Towards ‘Reproducible Science’



**"healthy"**

**"diseased"**

