

Handling integrated biological data using Python, Jupyter, and InterMine

Rachel Lyne

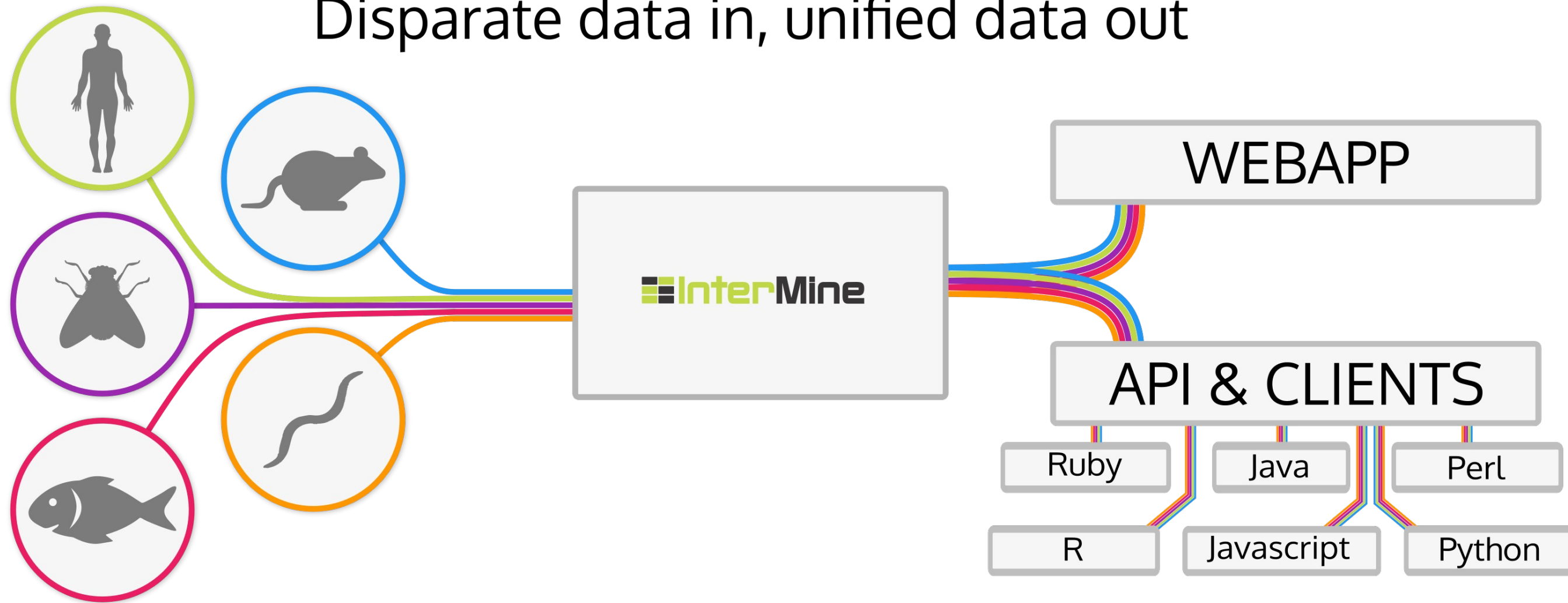
Yo Yehudi

Daniela Butano

Sergio Contrino

What is InterMine

Disparate data in, unified data out




















Model organism images Designed by Freepik and distributed by Flaticon

Who Uses InterMine?



<http://registry.intermine.org/>

InterMine Registry All InterMine instances up-to-date information in one place.

	Name	Description	Organisms
	BMAP	Brassicales Map Alignment Project	
	BeanMine	A mine with common bean data from the Legume Info tripal.chado database	A. ipaensis, A. duranensis, A. thalia...
	BovineMine	An integrated data warehouse for the Bovine Genome Database	B. taurus, C. hircus, O. aries
	CHOmine	An integrated database for Cricetulus griseus and CHO cells	C. griseus
	ChickpeaMine	A mine with chickpea data (both desi and kabuli varieties) from the Legume...	A. ipaensis, A. duranensis, A. thalia...
	CowpeaMine	A mine containing both cowpea genetic and genomic data, courtesy UC-Riv...	A. duranensis, A. ipaensis, C. arietin...
	FlyMine	An integrated database for Drosophila genomics	D. melanogaster
	GrapeMine	An integrated database for grapevine data	
	HumanMine	HumanMine integrates many types of data for Homo sapiens and Mus mus...	H. sapiens
	HymenopteraMine	An integrated data warehouse for the Hymenoptera Genome Database	A. dorsata, A. echinatio, A. florea, A...
	IndigoMine	INDIGO enables the integration of annotations for the exploration and analy...	Archae
	LegumeMine	Multi-organism mine integrates data from legume species: string bean, soy...	A. duranensis, A. ipaensis, C arietin...
	MaizeMine	An integrated data warehouse for MaizeGDB	
	MedicMine	MedicMine integrates many types of data for Medicago truncatula. You can...	A. thaliana, M. truncatula, M. trunca...
	MitoMiner	MitoMiner is an integrated web resource of mitochondrial localisation evide...	D. rerio, H. sapiens, M. musculus, R...
	ModMine	A data warehouse for the modENCODE project	
	MouseMine	MouseMine is a powerful new system for online access to mouse data fro...	M. musculus





ORTHOLOGUES

GENES UTRs

GENE ONTOLOGY

INTERACTIONS

EXONS

PATHWAYS

PROTEINS

Protein Domains

GWAS

REGULATORY

VARIANTS

SNPs MICROARRAY

ALLELES PHENOTYPES

RNA-seq

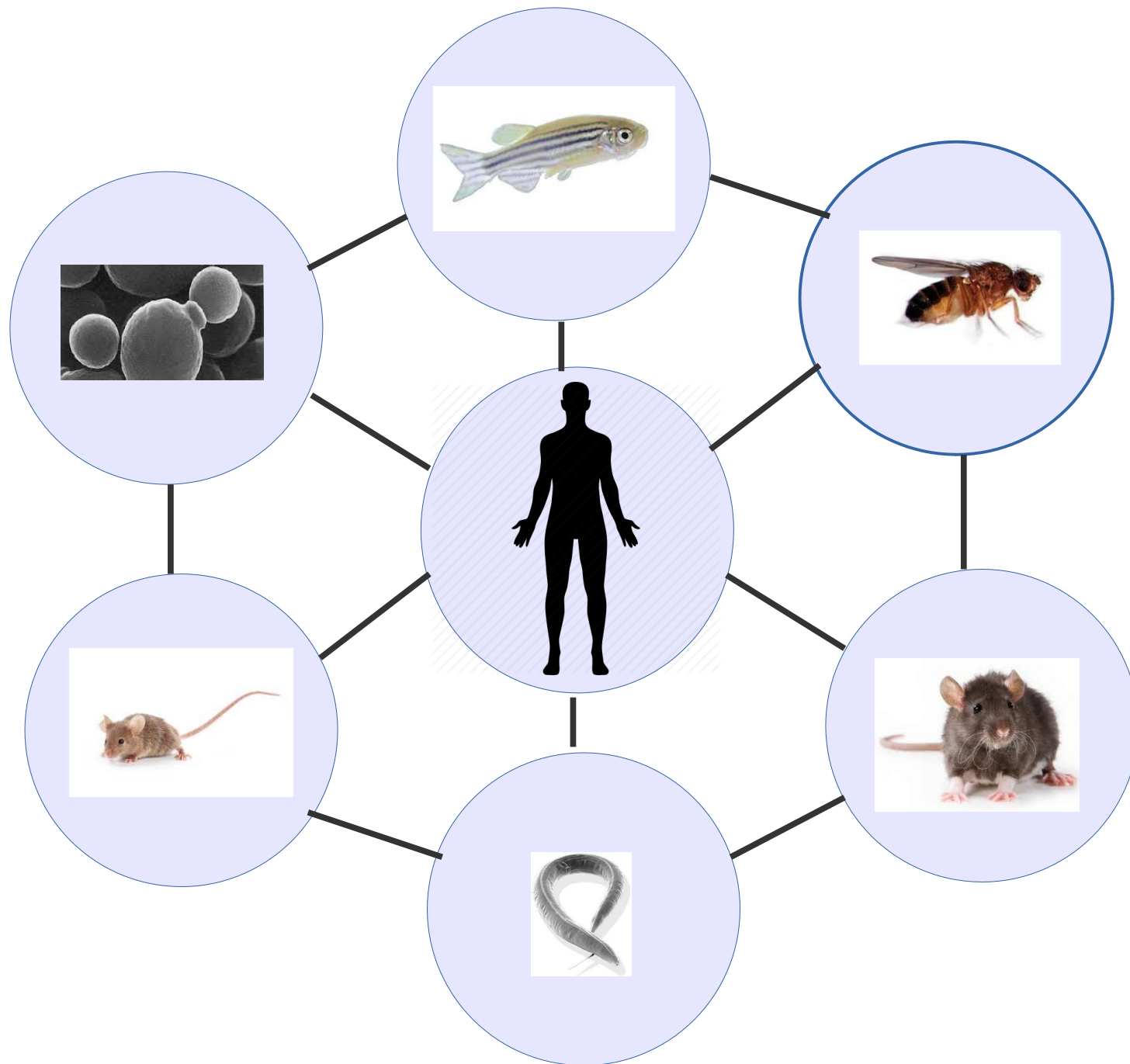
DISEASE

EXPRESSION

Why use InterMine?

- Query across several data sources at once
- Data formatting issues resolved
- Identifier resolution system
- Collate information about items and sets
- Common platform to many organisms and type of data
- Information without visiting several sites

Cross-organism analysis



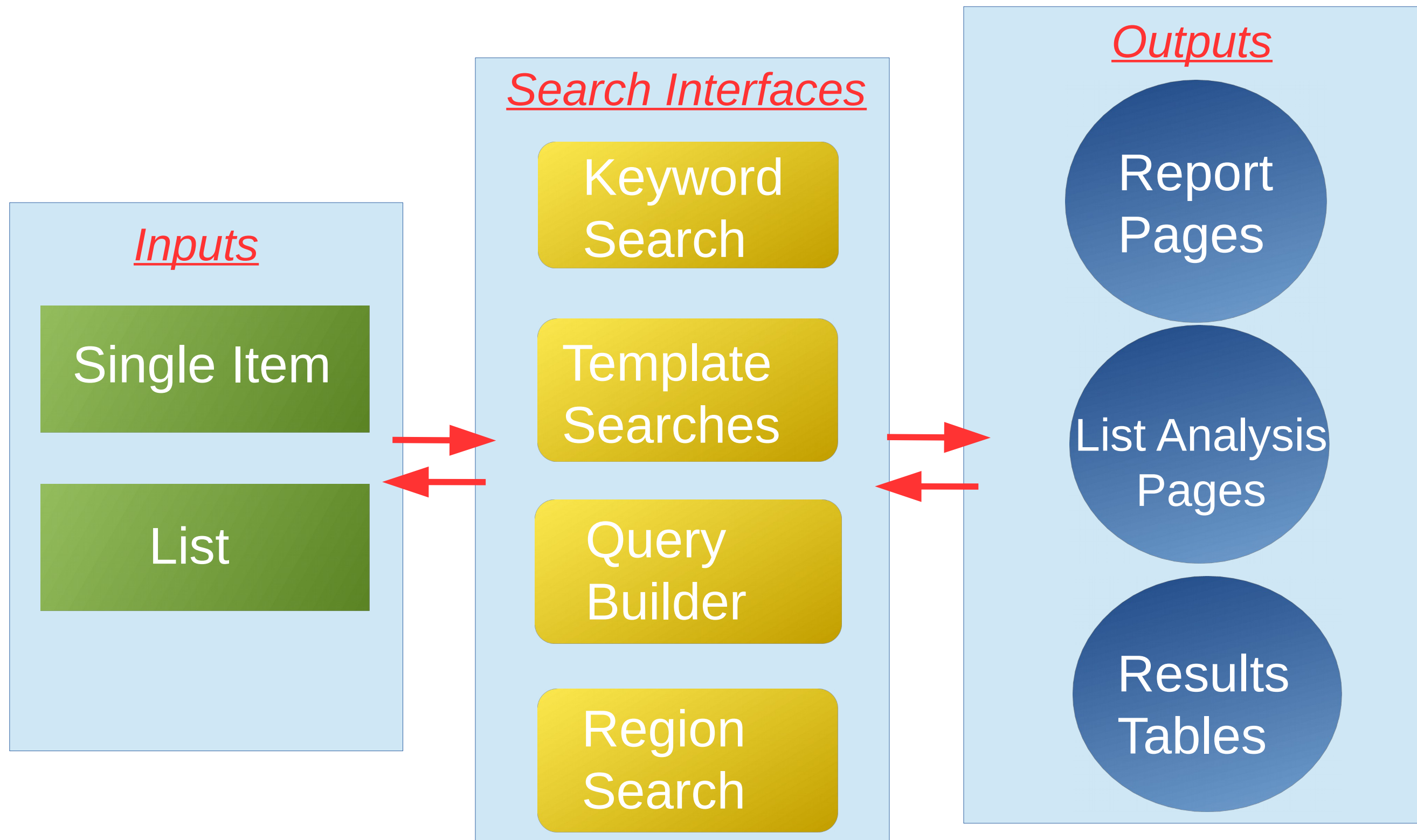
InterMine Accounts

InterMine is free to use without creating an account.

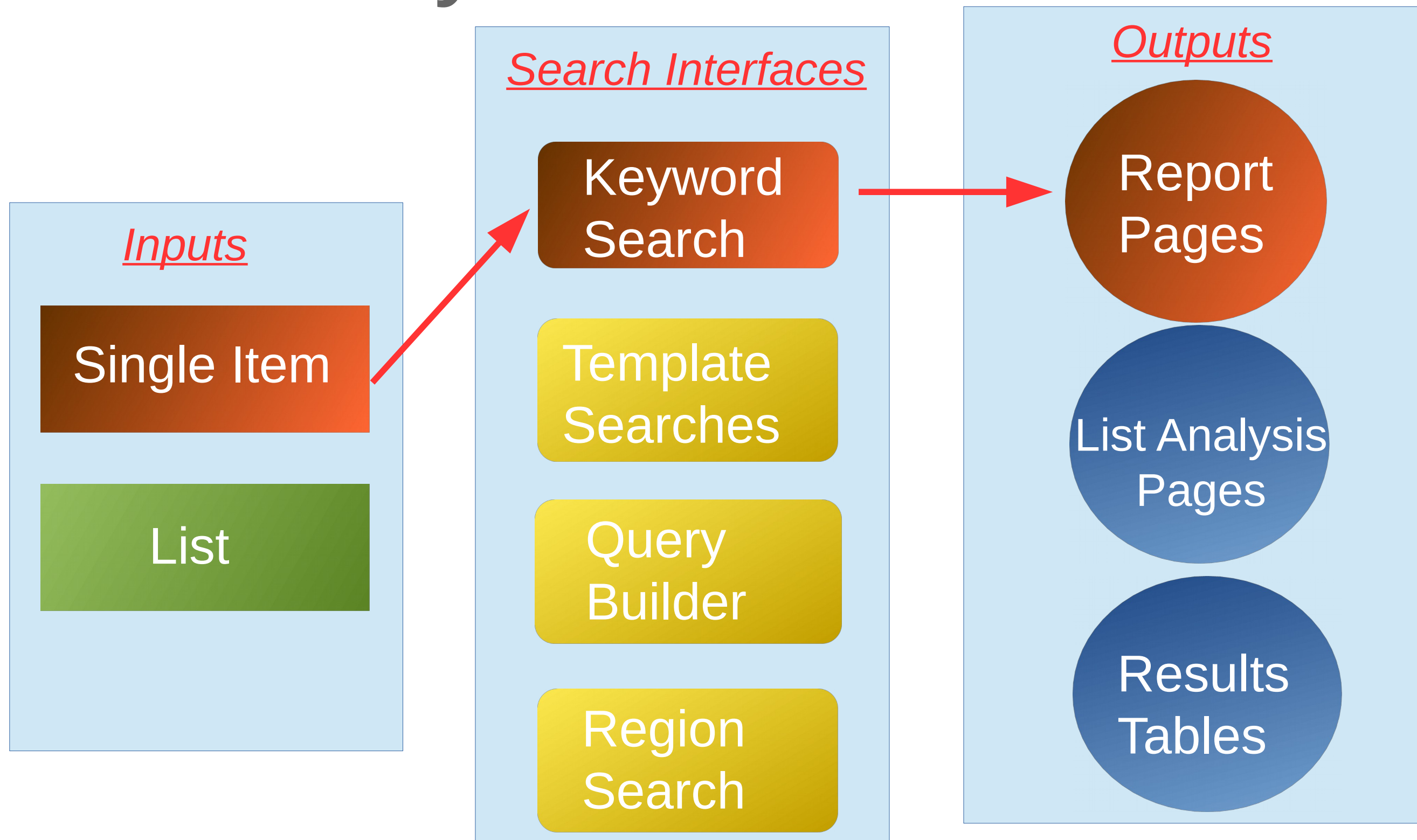
Creating an account allows you to save lists and searches permanently and share lists with your colleagues.

At the moment you have to make a separate account for each InterMine database

The Web Interface



Keyword Search

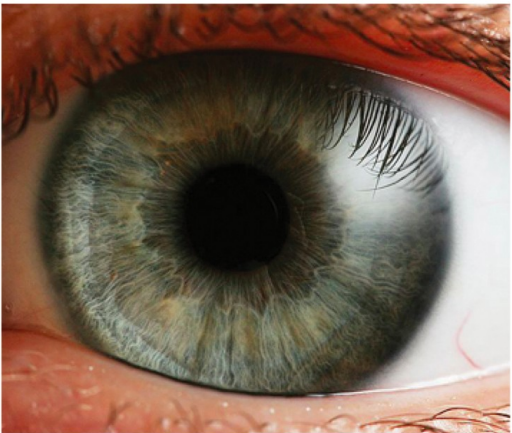
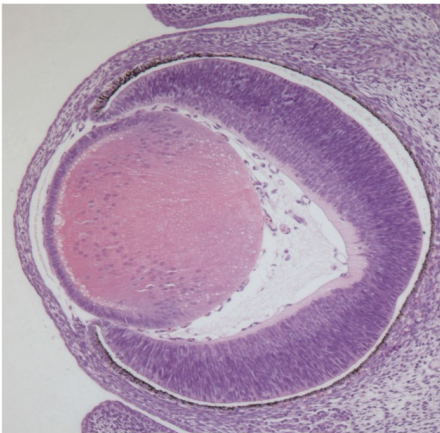


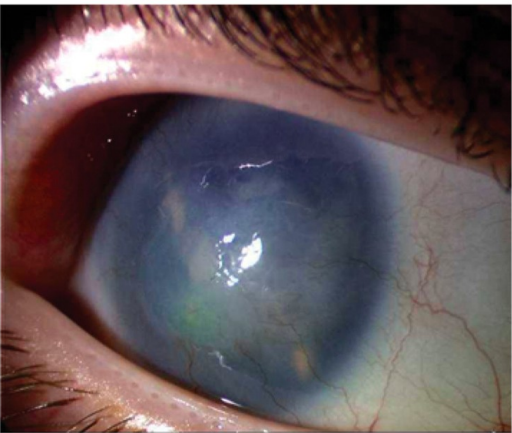
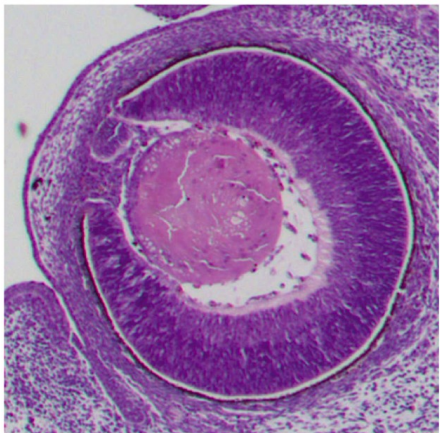
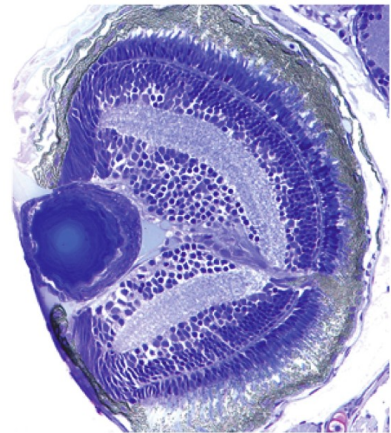
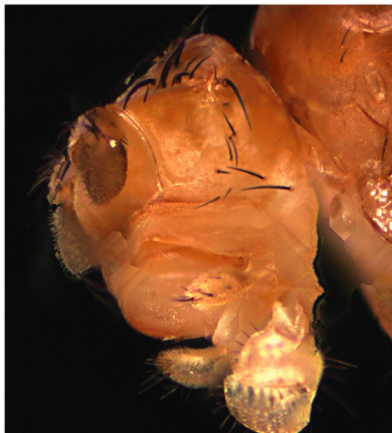


Data Exploration: report pages

Every object (item) in InterMine has a report page

- Collate all the data available for that object
- Contain a mixture of interactive tables, search results and graphical displays.
- Links to related data

Explore: Pax6


	Human	Mouse	Zebrafish	Drosophila
WT				
mut				
	<i>PAX6^{+/-}</i>	<i>Pax6^{-/-}</i>	<i>pax6b^{-/-}</i>	<i>ey^{-/-}</i>
EQs	cornea opaque iris absent retina degenerate lens opaque aqueous humor of eyeball increased pressure	eye decreased size lens fused_to cornea iris morphology anterior chamber absent	eye decreased size lens decreased size retina malformed	eye absent

Washington NL, Haendel MA, Mungall CJ, Ashburner M, Westerfield M, Lewis SE. - Figure 1 of Washington et al.:

"Linking Human Diseases to Animal Models Using Ontology-Based Phenotype Annotation." *PLoS Biol* 7(11): e1000247. doi:10.1371/journal.pbio.1000247



Pax6 Report Page

 **HumanMine** v3.0 2016 July An integrated database of *Homo sapiens* genomic data

help | about | cite | softw

Home Templates Lists QueryBuilder Regions Data Sources API MyMine

Contact Us | Log in

Search: GO

Gene : PAX6 *Homo sapiens*

Name	paired box 6	Brief Description	paired box 6
synonyms:	FVH1, D11S812E, aniridia, keratitis, AN, AN2, AN, PAX6, WAGR, AN2, uc021qfm.1, AY707088, Show more		
identifiers:	5080, ENSG00000007372, PAX6		
Region:	gene	Length: 33170	FASTA...
Location:	11:31784792-31817961 reverse strand		

9 **Pathways**
Reactome, KEGG

10 **Diseases**
OMIM

42 **Mouse Alleles (MGI)**
mouse alleles

75 **Gene Ontology**

Tissue
↑14↓20

Genes
Expression

Disease
↑15↓8

[SHARE](#)

Quick Links: [Summary](#) [Function](#) [Genomics](#) [Proteins](#) [SNPs](#) [Disease](#) [Homology](#) [Interactions](#) [Gene Ontology](#) [Other](#)

Curated comments from UniProt

Type	Comment
developmental stage	Expressed in the developing eye and brain. Expression in the retina peaks at fetal days 51-60. At 6-week old, in the retina, is predominantly detected in the neural layer (at protein level). At 8- and 10-week old, in the retina, the expression is strongest in the inner and middle layer of the neural part (at protein level).
disease	MIM:106210; Aniridia; AN; A congenital, bilateral, panocular disorder characterized by complete absence of the iris or extreme iris hypoplasia. Aniridia is not just an isolated defect in iris development but it is associated with macular and optic nerve hypoplasia, cataract, corneal changes, nystagmus. Visual acuity is generally low but is unrelated to the degree of iris hypoplasia. Glaucoma is a secondary problem causing additional visual loss over time. The disease is caused by mutations affecting the gene represented in this entry.
disease	MIM:120200; Coloboma, ocular, autosomal dominant; COAD; A set of malformations resulting from abnormal morphogenesis of the optic cup and stalk, and the fusion of the fetal fissure (optic fissure). The clinical presentation is variable. Some individuals may present with minimal defects in the anterior iris leaf without other ocular defects. More complex malformations create a combination of iris, uveoretinal and/or optic nerve defects without or with microphthalmia or even anophthalmia. The disease is caused by mutations affecting the gene represented in this entry.

[Show proteins](#)

Lists

This Gene isn't in any lists.
[Upload a list.](#)

Links to other Mines

RatMine

R. norvegicus
[Pax6](#)

YeastMine No results


FlyMine

D. melanogaster
[toy](#), [ey](#)

MouseMine



Pax6 Report Page

 **HumanMine** v3.0 2016 July An integrated database of *Homo sapiens* genomic data

help | about | cite | softw

Home Templates Lists QueryBuilder Regions Data Sources API MyMine

Contact Us | Log in

Search: GO

Gene : PAX6 *Homo sapiens*

Name	paired box 6	Brief Description	paired box 6
synonyms:	FVH1, D11S812E, aniridia, keratitis, AN, AN2, AN, PAX6, WAGR, AN2, uc021qfm.1, AY707088, Show more		
identifiers:	5080, ENSG00000007372, PAX6		
Region:	gene	Length: 33170 FASTA...	
Location:	11:31784792-31817961 reverse strand		

9 **Pathways**
Reactome, KEGG

10 **Diseases**
OMIM

42 **Mouse Alleles (MGI)**
mouse alleles

75 **Gene Ontology**

Tissue
↑14↓20

Genes
Expression

Disease
↑15↓8

[SHARE](#)

Quick Links: [Summary](#) [Function](#) [Genomics](#) [Proteins](#) [SNPs](#) [Disease](#) [Homology](#) [Interactions](#) [Gene Ontology](#) [Other](#)

Curated comments from UniProt

[Show proteins](#)

Type	Comment
developmental stage	Expressed in the developing eye and brain. Expression in the retina peaks at fetal days 51-60. At 6-week old, in the retina, is predominantly detected in the neural layer (at protein level). At 8- and 10-week old, in the retina, the expression is strongest in the inner and middle layer of the neural part (at protein level).
disease	MIM:106210; Aniridia; AN; A congenital, bilateral, panocular disorder characterized by complete absence of the iris or extreme iris hypoplasia. Aniridia is not just an isolated defect in iris development but it is associated with malformations of the optic nerve hypoplasia, cataract, corneal changes, nystagmus. Visual acuity is generally low but is unrelated to the degree of iris hypoplasia. Glaucoma is a secondary problem causing additional visual loss over time. The disease is caused by mutations affecting the gene represented in this entry.
disease	MIM:120200; Coloboma, ocular, autosomal dominant; COAD; A set of malformations resulting from abnormal morphogenesis of the optic cup and stalk, and the fusion of the fetal fissure (optic fissure). The clinical presentation varies. Some individuals may present with minimal defects in the anterior iris leaf without other ocular defects. More complex malformations create a combination of iris, uveoretinal and/or optic nerve defects without or with microphthalmia or even anophthalmia. The disease is caused by mutations affecting the gene represented in this entry.

Lists

This Gene isn't in any lists.
[Upload a list.](#)

Links to other Mines

[RatMine](#)
R. norvegicus
[Pax6](#)

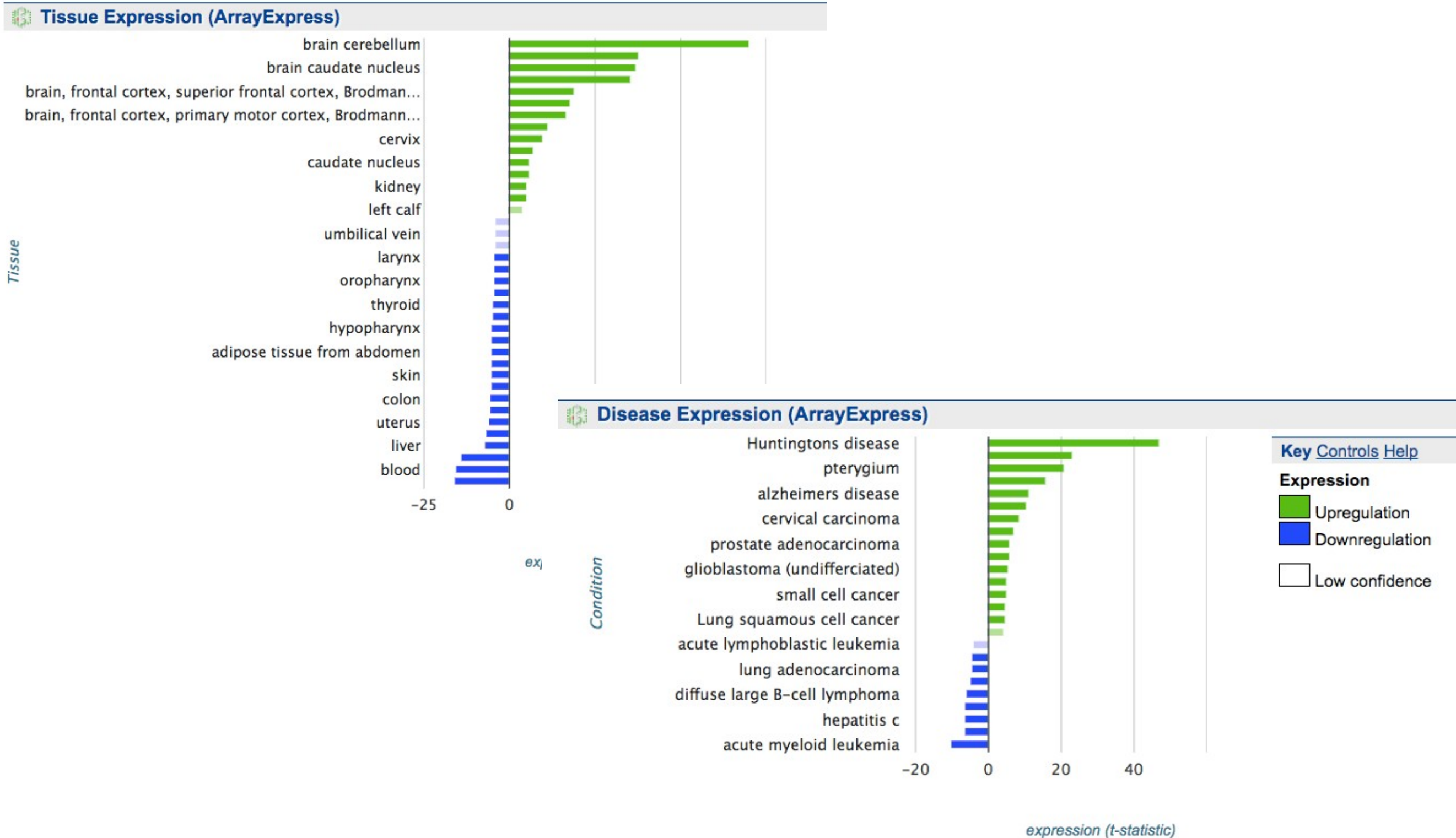
[YeastMine](#) No results

[FlyMine](#)
D. melanogaster
[toy](#), [ey](#)

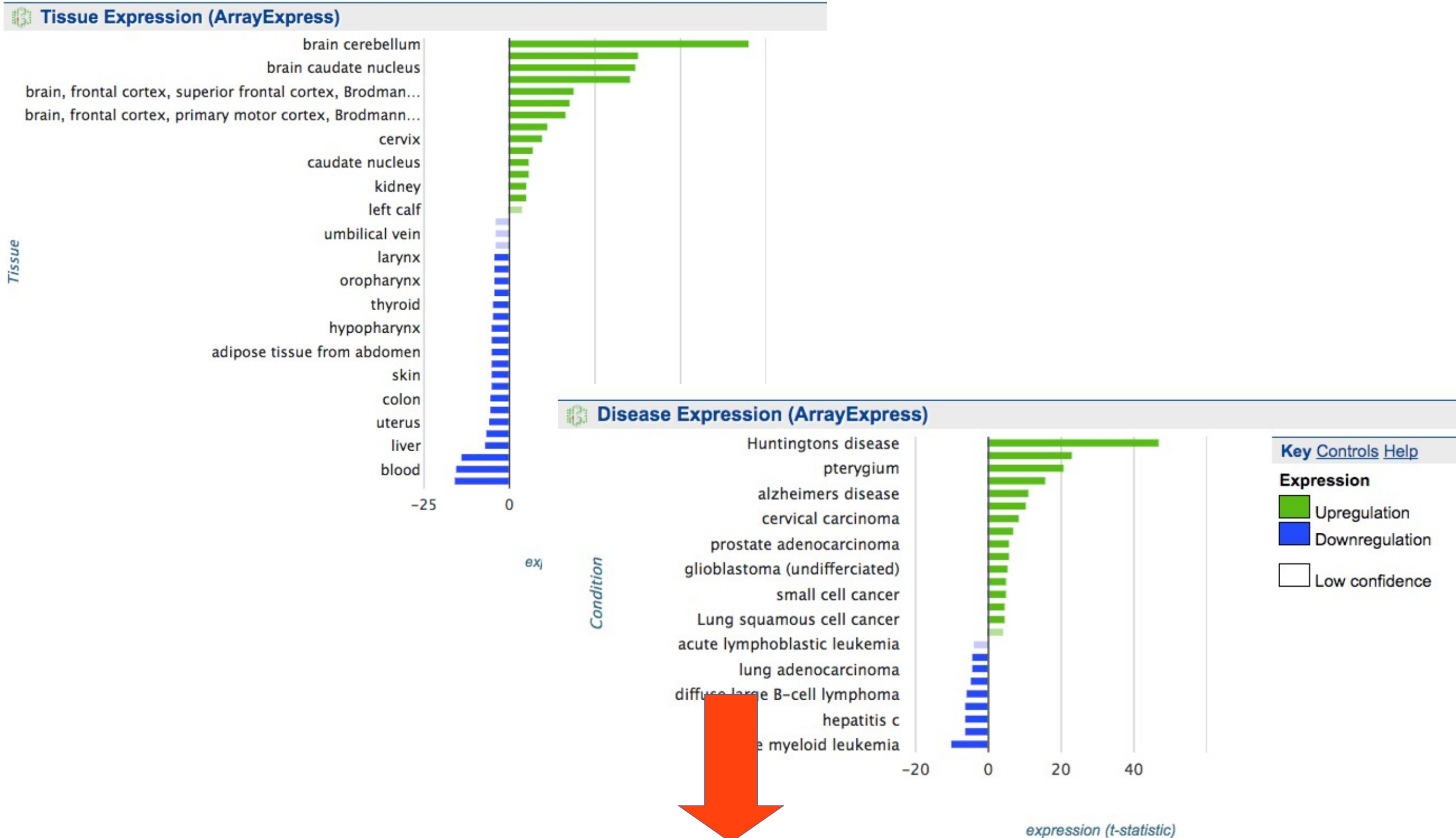
[MouseMine](#)



Pax6 – HumanMine Expression Data



Pax6 – HumanMine Expression Data



Pax6 – humanMine protein expression data

Overall	Organ	Cell types	Antibody staining
<input type="checkbox"/>	Blood and immune system (Hematopoietic)	8	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	Breast and female reproductive system (Female tissues)	9	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	Cardiovascular system (Heart and blood vessels)	1	<input type="checkbox"/>
<input type="checkbox"/>	Central nervous system (Brain)	11	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	Digestive tract (GI-tract)	11	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	Endocrine glands	3	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<input checked="" type="checkbox"/>	Liver and pancreas	5	<input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	Male reproductive system (Male tissues)	5	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	Placenta	2	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	Respiratory system (Lung)	4	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	Skin and soft tissues	2	<input type="checkbox"/> <input type="checkbox"/>
<input type="checkbox"/>	Urinary tract (Kidney and bladder)	3	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

High level expression
in the pancreas

Key Help

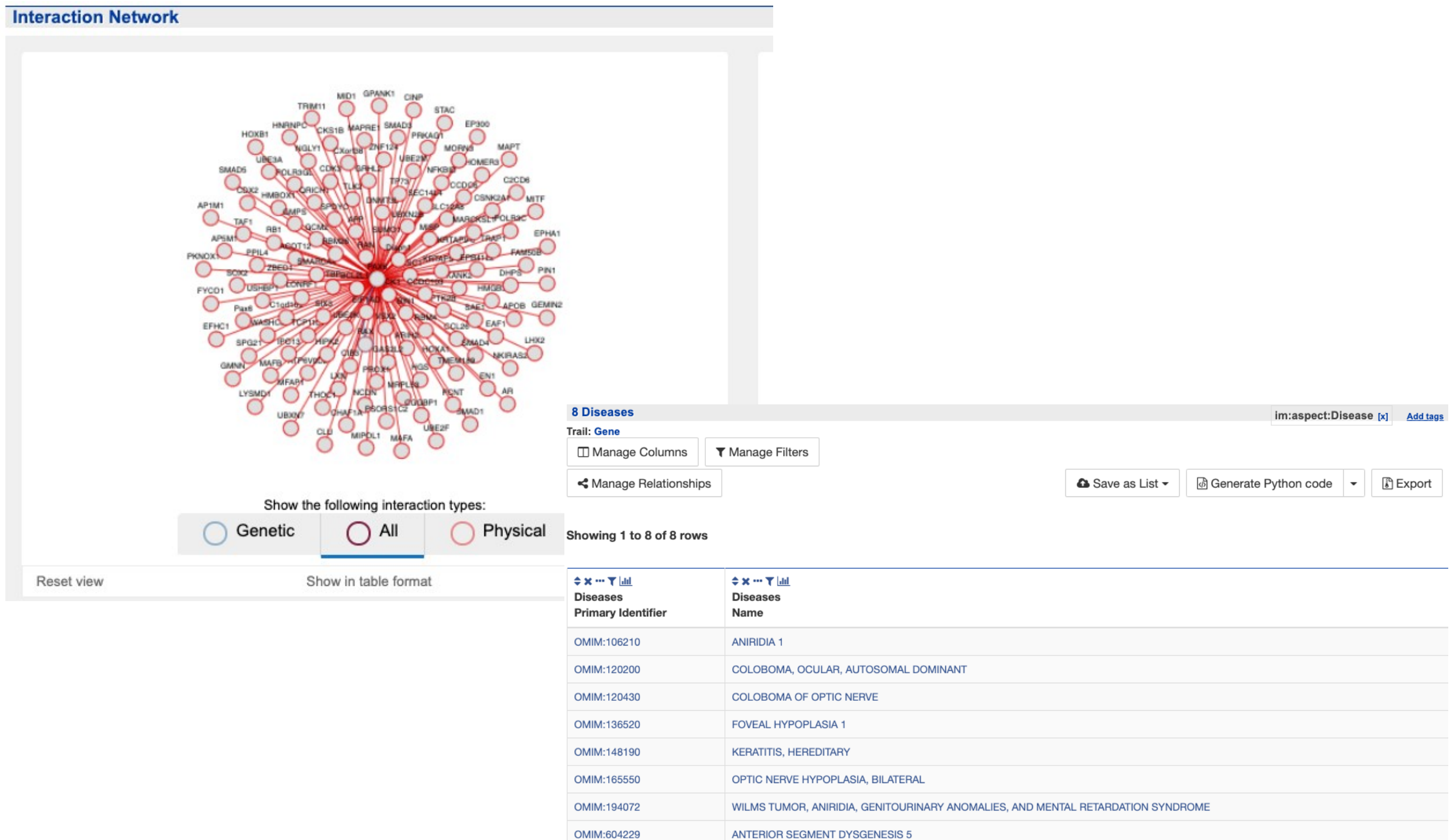
Reliability: **Supportive** (APE - two or more antibodies)

Level of antibody staining*

☒ High
☒ Medium
☐ Low
☐ None







Pax6 – interaction and disease data





Pax6 – jump to another InterMine


Links to other Mines


modMine
D. melanogaster
toy , ey 
C. elegans
vab-3 








RatMine
R. norvegicus
Pax6 

YeastMine No results

FlyMine
D. melanogaster
toy , ey 

MouseMine
M. musculus
Pax6 

ZebrafishMine
D. rerio
ZDB-GENE-990415-200 

External Links
HuGe 
ArrayExpress Atlas 
ensembl 
BioGPS 
HapMap 
Entrez Gene 
BioGRID 

Link to orthologous
gene in MouseMine

Pax6 – mouseMine Report Page

Gene : Pax6 paired box 6

Primary Identifier	MGI:97490	Organism	Mus musculus
Chromosome	2	NCBI Gene Number	18508
Mgi Type	protein coding gene		
description	<p>FUNCTION: This gene encodes a homeobox-containing protein that functions as a regulator of transcription. It plays a key role in the development of neural tissues, particularly the eye. Activity of this protein is also required for expression of glucagon in the pancreas. This gene is regulated by multiple enhancers located up to tens or hundreds of kilobases upstream and downstream of the transcription start sites. Mutations in this gene or deletion of these regulatory elements results in severe defects in eye development. Alternative splicing and the use of alternative promoters results in multiple transcript variants, some of which encode proteins that lack the N-terminal paired domain. [provided by RefSeq, Jul 2015]</p> <p>PHENOTYPE: Null and hypomorphic mutants show a range of phenotypes from viable with small eyes and lens/cornea fusion to microphthalmia and cataract to embryonic or perinatal lethality with anophthalmia and severe craniofacial and forebrain defects. [provided by MGI curators]</p>		
synonyms:	RIKEN cDNA 1500038E17 gene, Sey, MGD-MRK-13158, 1500038E17Rik, paired box 6, Show more		
earliestPublication:	Roberts RC (1967) Small-eyes, a new dominant mutant in the mouse. Genet Res 9:121-122		



Quick Links: [Summary](#) [Genome](#) [Proteins](#) [Function](#) [Homology](#) [Interactions](#) [Expression](#) [Phenotype](#) [Disease](#) [Literature](#)

[Features --> Cross References](#) (159 rows)

Genome

Genome feature			
Region:	protein_coding_gene	Length:	28465
Location:	2:105668900-105697364 forward strand		

Lists

This Gene isn't in any lists.
[Upload a list.](#)

Links to other Mines

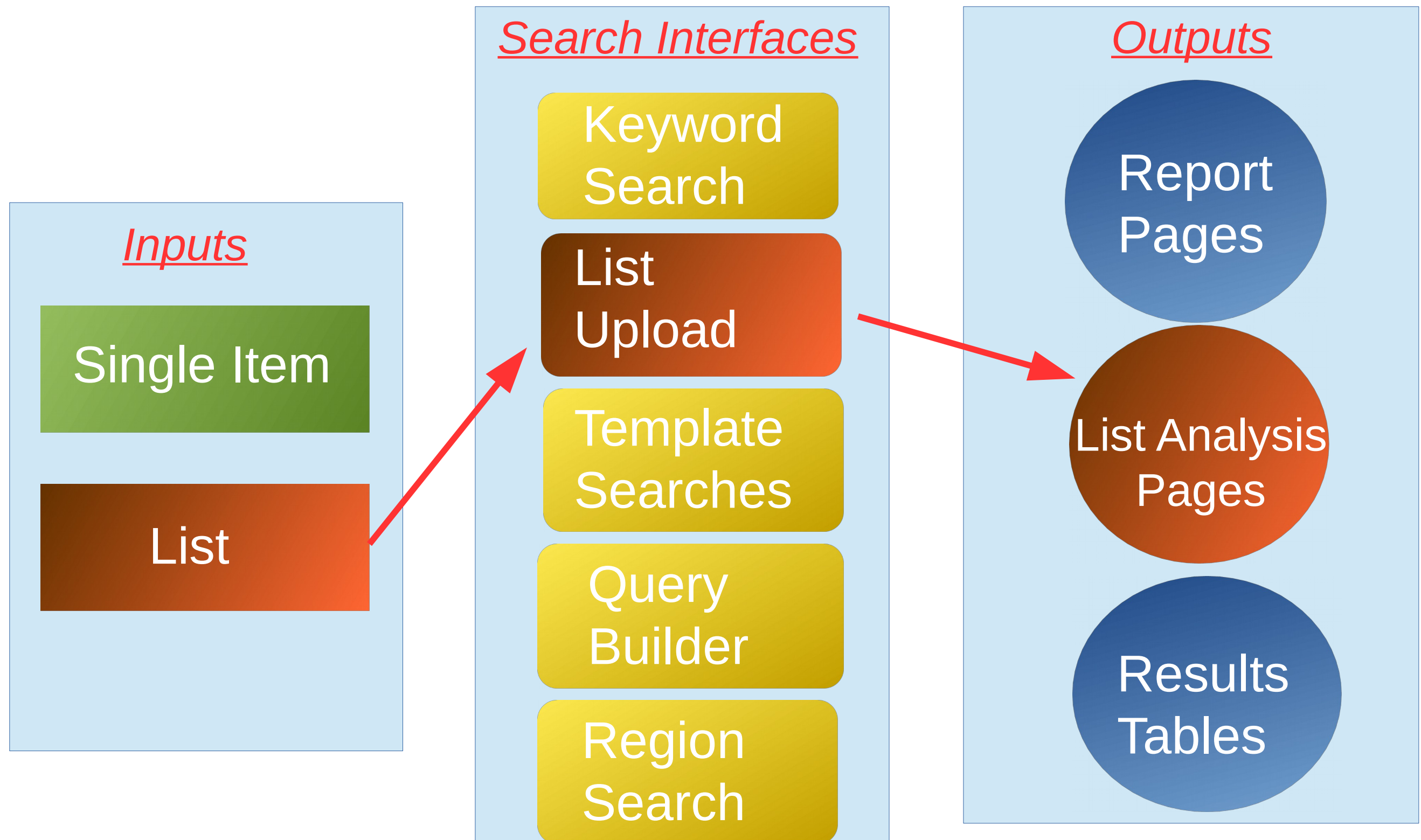
FlyMine

D. melanogaster
[ey](#), [toy](#)

HumanMine

H. sapiens
[PAX6](#)

Lists



List Analysis: Uploading a list

- Upload your own lists to InterMine
- Powerful identifier resolution system
- Convert old identifiers into an up-to-date set

List Analysis

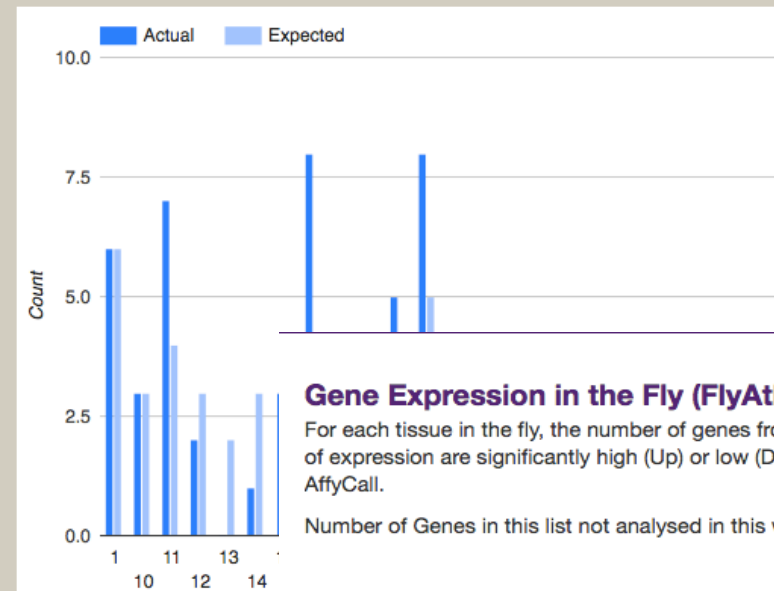
Chromosome Distribution

Actual: number of items in this list found on each chromosome. Expected: given the total number of items on the chromosome and the number of items in this list, the number of items expected to be found on each chromosome.

All items in your list have been analysed.

Organism

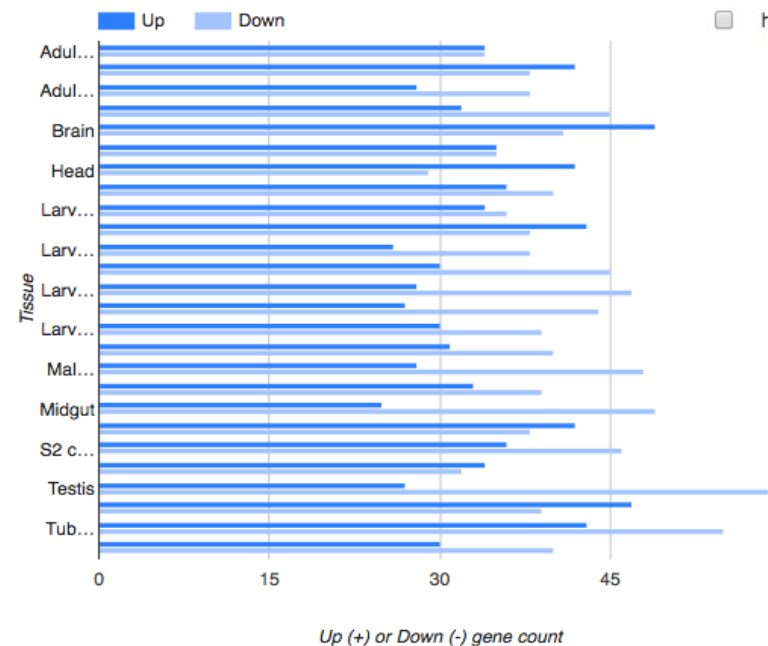
Homo sapiens



Gene Expression in the Fly (FlyAtlas)

For each tissue in the fly, the number of genes from this list for which of expression are significantly high (Up) or low (Down) according to FlyAtlas.

Number of Genes in this list not analysed in this widget: 3



Gene Ontology Enrichment

GO terms enriched for items in this list.

Number of Genes in this list not analysed in this widget: 17

Test Correction: Holm-Bonferroni, Max p-value: 0.05, Ontology: biological_process

Background population

Default Change

View Download

GO Term	p-Value	Matches
cell adhesion [GO:0007155]	8.077144e-5	13
biological adhesion [GO:0022610]		
Wnt signaling pathway [GO:0016055]		
cell-cell adhesion [GO:0098609]		
wing disc development [GO:0035220]		
cell-matrix adhesion [GO:0007160]		
tissue development [GO:0009888]		
heart development [GO:0007507]		

View homologues in other Mines:

RatMine

R. norvegicus

YeastMine

S. cerevisiae

MouseMine

M. musculus

HumanMine

H. sapiens

ZebrafishMine

D. rerio

Mammalian Phenotype Ontology Enrichment

MP terms enriched for items in this list.

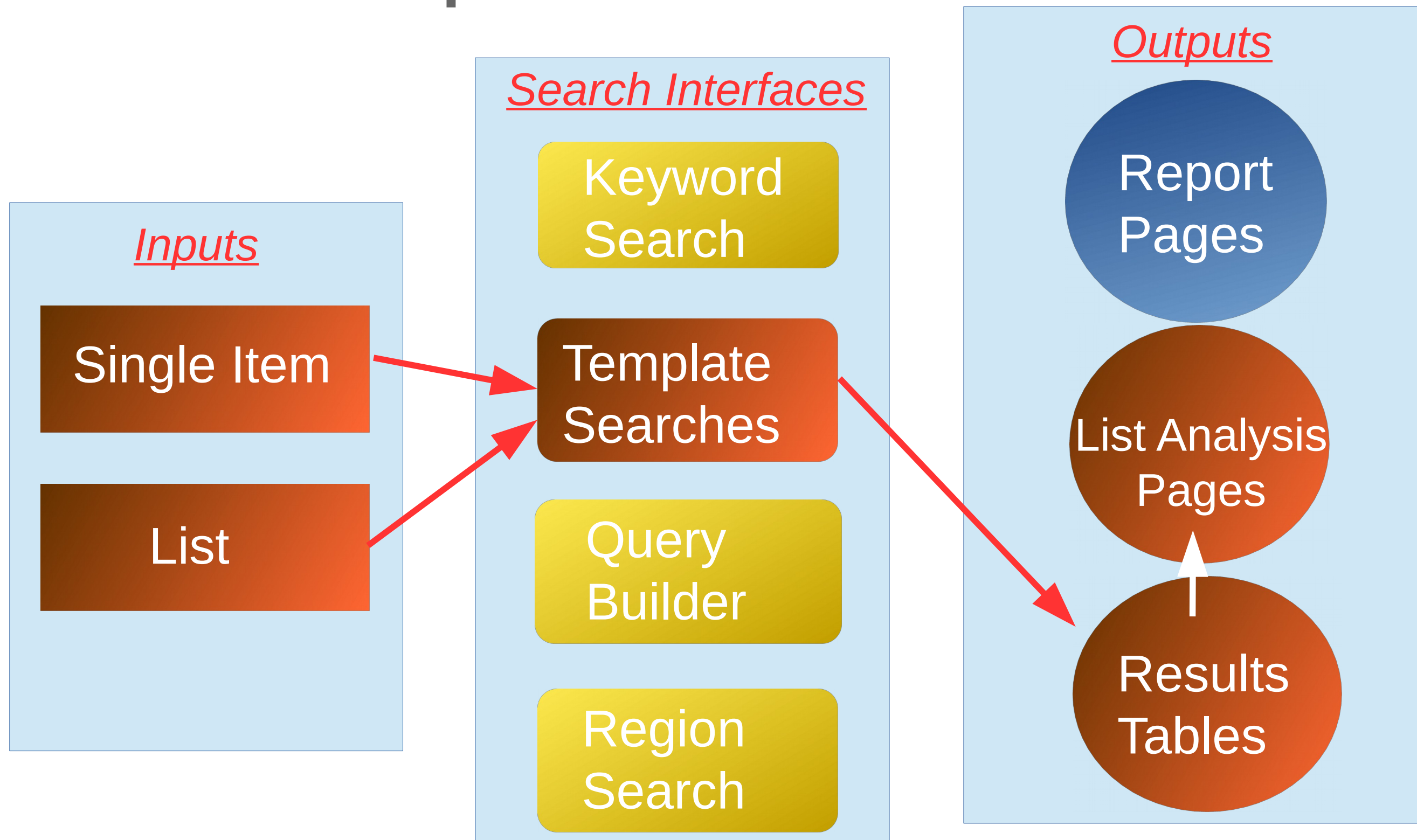
Number of Genes in this list not analysed in this widget: 72

Test Correction: Holm-Bonferroni, Max p-value: 0.05, Background population: Default Change

View Download

MP Term	p-Value	Matches
abnormal DNA repair [MP:0008058]	1.831867e-66	52
increased sensitivity to induced cell death [MP:0008943]	1.383230e-62	59
abnormal induced cell death [MP:0008942]	1.175928e-60	65
abnormal chromosome stability [MP:0010094]	4.405108e-58	48
chromosomal instability [MP:0008866]	1.743731e-56	47
abnormal cell physiology [MP:0005621]	2.197659e-53	163
cellular phenotype [MP:0005384]	1.305734e-51	168
chromosome breakage [MP:0004028]	4.007892e-50	37
abnormal cell death [MP:0000313]	1.184952e-44	106

Template Searches



Data Analysis: Template Searches

The template searches allow a more refined search than the keyword search and report pages but are still quick and easy to access.

- Pre-defined searches with simple filters
- Range from simple searches to more complex searches spanning several data types
- Run with single item or list
- Results are returned in sophisticated results tables
- Easy to add - just ask

Many Many Searches.....

- Which other genes have this GO annotation?
- Are there mutant phenotypes for this gene?
- Where is this gene expressed?
- What does this gene interact with?
- Do any of the interacting genes share the mutant phenotypes?
- Does this gene have a human orthologue with a disease association?
- Have any variants been associated with this gene/disease?
- Which organisms have models for this disease/gene?


Data Analysis: Template Searches

- What other genes are involved in pancreatic function?
- Are there potential targets of pax6 in pancreatic tissue?
- Have these genes been implicated in pancreatic disease?
- What published data is there about these genes?

Data Analysis: Template Searches

Are there potential targets of pax6 in pancreatic tissue?

Gene + Tissue Expression → Interactors that are expressed in that tissue

**Gene + Tissue Expression** → **Interactors that are expressed in that tissue**
For a specified gene or genes, show genes that interact with this gene that are also expressed in the specified tissue. (Interactors could be physical or genetic; Expression is based on tissue expression from the Protein Atlas Project and should be further interrogated for reliability level).

Gene – For the gene:
LOOKUP: for Organism:
☐ constrain to be saved Gene list

Tissue > Name – Show interactors that are expressed in:

[web service URL](#) [Perl](#) | [Python](#) | [Ruby](#) | [Java \[help\]](#) [export XML](#)

Data Analysis: Results Tables

Results tables allow further interactive analysis of the data through:

- Column summaries
- Column sorting
- Adding additional columns of data
- Filtering
- List creation
- Export

Data Analysis: Results Tables

Gene + Tissue Expression ➡ **Interactors that are expressed in that tissue** ☆

For a specified gene or genes, show genes that interact with this gene that are also expressed in the specified tissue. (Interactors could be physical or genetic; Expression is based on tissue expression from the Protein Atlas Project and should be further interrogated for reliability level).

 Manage Columns

▼ Manage Filters

➤ Manage Relationships

Save as List

 Generate Python code

 Export

Showing 1 to 25 of 191 rows

Rows per page: 25


 page 1

<div><div><div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div></div><div>Gene Primary Identifier</div></div></div><div><div><div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div></div><div>Gene Symbol</div></div></div><div><div><div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div></div><div>Interactions Participant 2 . Primary Identifier</div></div></div><div><div><div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div></div><div>Participant 2 Symbol</div></div></div><div><div><div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div></div><div>Protein Atlas Expression Tissue . Name</div></div></div><div><div><div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div><div></div><div></div><div></div><div></div><div></div></div></div></div><div>Protein Atlas Expression Level</div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div>

Data Analysis: Results Tables

Participant 2			as Expression		
Symbol					
CHAF1A					
CHAF1A					
SAE1					
SAE1					
TRAP1					
TRAP1					
CDK3					
CDK3					
CDX2					
CDX2					
POLR3C					
POLR3C	Pancreas				Not detected
EFHC1	Pancreas				Low
EFHC1	Pancreas				Medium
CKS1B	Pancreas				High

4 Protein Atlas Expression Levels



94 Items Selected

Filter values

Protein Atlas Expression Level	Count
<input checked="" type="checkbox"/> Medium	71
<input type="checkbox"/> Not detected	62
<input type="checkbox"/> Low	35
<input checked="" type="checkbox"/> High	23

Filter Filter values Download data

Data Analysis: Results Tables

➔ Interactors that are expressed in that tissue

Interactors that interact with this gene that are also expressed in the specified tissue. (Interactors could be phylogenetic and should be further interrogated for reliability level).

Manage Relationships

↶ Undo

Save as List

Gene (1 Gene)

Gene > Interactions > Participant 2 (56 Genes)

Gene > Interactions > Participant 2 > Protein Atlas Expression > Tissue (1 Tissue)

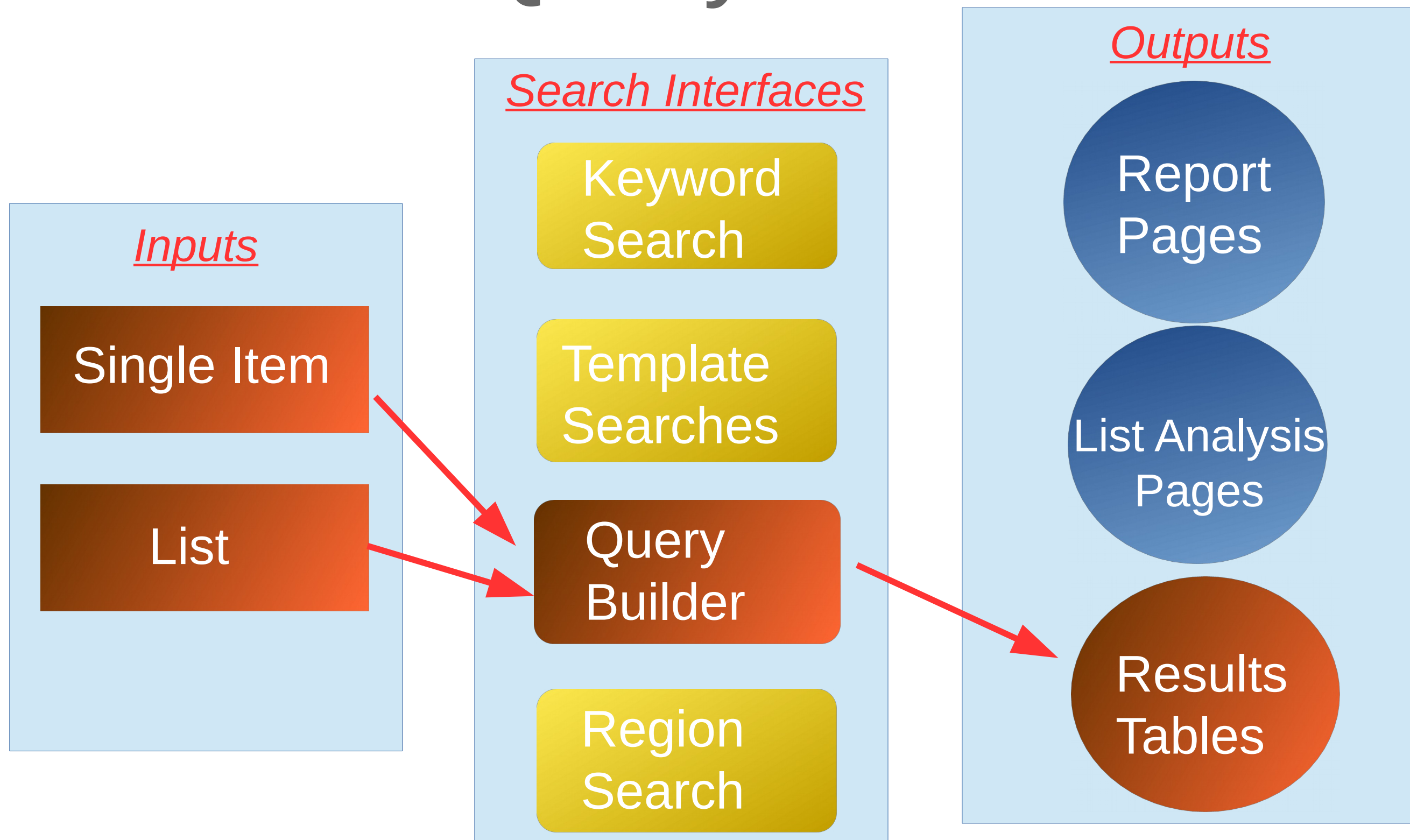
Gene > Interactions > Participant 2 > Protein Atlas Expression (94 Protein Atlas Expressions)

Pick items from the table

Create ListAdd to List

10055	SAE1	Pancreas
10131	TRAP1	Pancreas
10131	TRAP1	Pancreas
1018	CDK3	Pancreas
114327	EFHC1	Pancreas
1163	CKS1B	Pancreas
1163	CKS1B	Pancreas
116540	MRPL53	Pancreas

The Query Builder



Advanced Search: Query Builder

The Query Builder is InterMine's custom query builder, allowing you to create and save your own searches.

- Build your own Searches
- Modify template searches
- Combine any data:
 - And, Or,
 - Intersect; Union

Data Analysis: Query Builder

Three steps to construct a query:

1. Navigate the data model to find the class or attribute you need
2. Add the appropriate constraint (filter) to the class/attribute
3. Decide on the columns you want to view in your results

Data Analysis: Query Builder

Show results

Model browser

- Gene
- Brief Description
 - Cytological Location
 - Description
 - Id Integer
 - Length Integer
 - Name
 - Primary Identifier
 - Secondary Identifier
 - Symbol
- ☒ Alleles Allele
- ☒ Atlas Expression Atlas Expression
- ☒ Chromosome Chromosome
- ☒ Chromosome Location Location
- ☒ Cross References Cross Reference
- ☒ Data Sets Data Set
- ☒ Diseases Disease
- ☒ ☒ Genes Gene
- Id Integer
 - Name
 - Primary Identifier
- ☒ Alleles Allele
- ☒ Data Sets Data Set
- ☒ Hpo Annotations HPO Annotation
- ☐ Publications Publication

☐ Show empty fields

Fields selected for output

Query Overview

Gene

LOOKUP Pax6 IN H. sapiens (A)

Primary Identifier

Symbol

Diseases Disease collection

Name

Constraint logic:

one constraint





Data Analysis: Query Builder



Fields selected for output



Columns to Display

Use the [SHOW](#) or [SUMMARY](#) links to add fields to the results table. Click and drag the blue output boxes to choose the output column order. Click [A↓](#) to choose a column to sort results by, click again to select ascending [A↓](#) or descending [A↑](#). Use the [REMOVE ALL](#) link to remove all fields from the results table.

[REMOVE ALL](#)

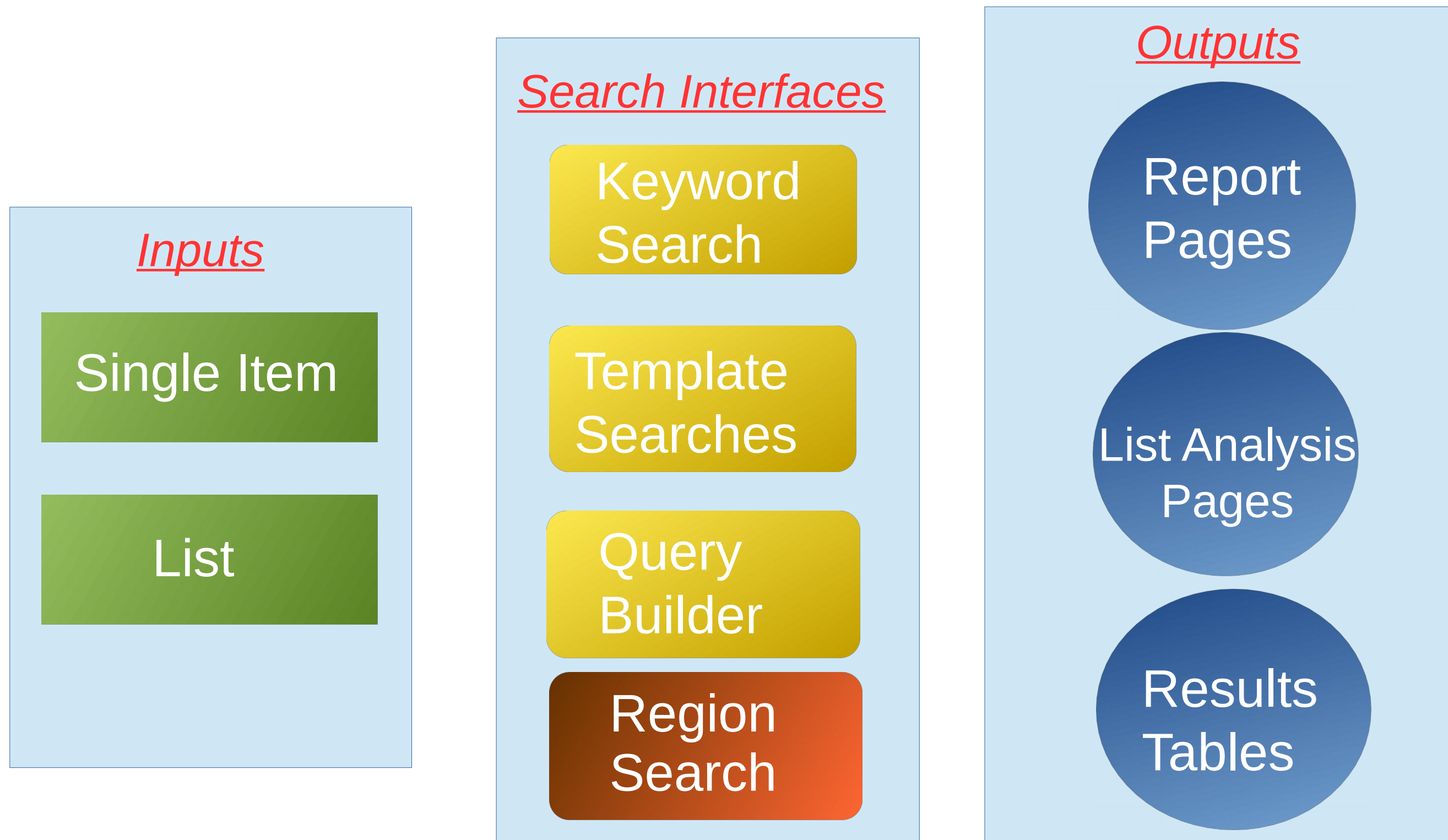
Gene > Primary Identifier 
(no description) 
[A↓](#)

Gene > Symbol 
(no description) 
[A↓](#)

Gene > Diseases > Name 
(no description) 
[A↓](#)

[Show results](#)[web service URL](#)[Perl](#) | [Python](#) | [Ruby](#) | [Java \[help\]](#)[Export XML](#)[Start building a template query](#) [Save query](#)

The Web Interface



Data Analysis: RegionSearch

The Region Search allows you to search for features that overlap a list of genome coordinates.

- Any or selected genome features can be searched.
- Accepts base or interbase coordinates
- Region to be searched can be extended upstream and downstream

Other tools: RegionSearch

The Region Search allows you to search for features that are near to / overlap a list of genome coordinates.

1. Select Organism: *genome build: not available*

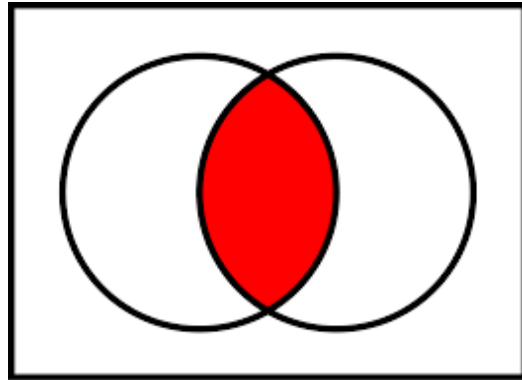
2. ☒ Select Feature Types:

<input checked="" type="checkbox"/> cDNA Clone i	<input checked="" type="checkbox"/> Exon i	<input checked="" type="checkbox"/> Probe Set i
<input checked="" type="checkbox"/> CDS i	<input checked="" type="checkbox"/> Five Prime UTR i	<input checked="" type="checkbox"/> rRNA i
<input checked="" type="checkbox"/> CRM i	<input checked="" type="checkbox"/> Gene i	<input checked="" type="checkbox"/> Regulatory Region i
<input checked="" type="checkbox"/> Chromosomal Deletion i	<input checked="" type="checkbox"/> Intergenic Region i	<input checked="" type="checkbox"/> snRNA i
<input checked="" type="checkbox"/> Chromosomal Duplication i	<input checked="" type="checkbox"/> Intron i	<input checked="" type="checkbox"/> snoRNA i
<input checked="" type="checkbox"/> Chromosomal Inversion i	<input checked="" type="checkbox"/> mRNA i	<input checked="" type="checkbox"/> TF Binding Site i
<input checked="" type="checkbox"/> Chromosomal Translocation i	<input checked="" type="checkbox"/> miRNA i	<input checked="" type="checkbox"/> tRNA i
<input checked="" type="checkbox"/> Chromosomal Transposition i	<input checked="" type="checkbox"/> miRNA Target i	<input checked="" type="checkbox"/> Three Prime UTR i
<input checked="" type="checkbox"/> Chromosome Band i	<input checked="" type="checkbox"/> Microarray Oligo i	<input checked="" type="checkbox"/> Transcript i
<input checked="" type="checkbox"/> Chromosome Structure Variation i	<input checked="" type="checkbox"/> Natural Transposable Element i	<input checked="" type="checkbox"/> Transposable Element i
<input checked="" type="checkbox"/> EST i	<input checked="" type="checkbox"/> ncRNA i	<input checked="" type="checkbox"/> Transposable Element Insertion Site i
<input checked="" type="checkbox"/> Enhancer i	<input checked="" type="checkbox"/> Point Mutation i	

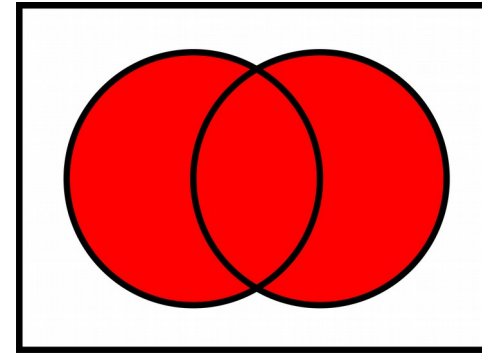
0 1k 10k 100k 1M 10M

5. ☐ Check this box to perform a strand-specific region search (search + strand if region start<end; search - strand if region end<start)

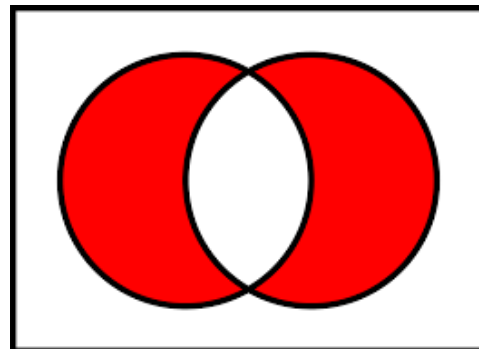
Lists: Set Analysis



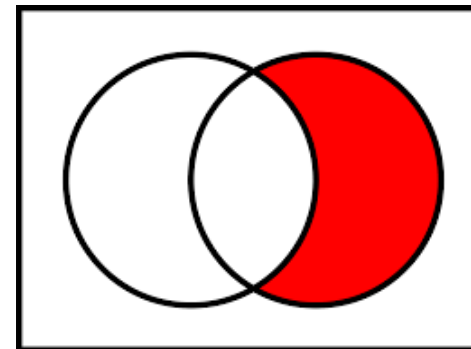
Intersect



Union

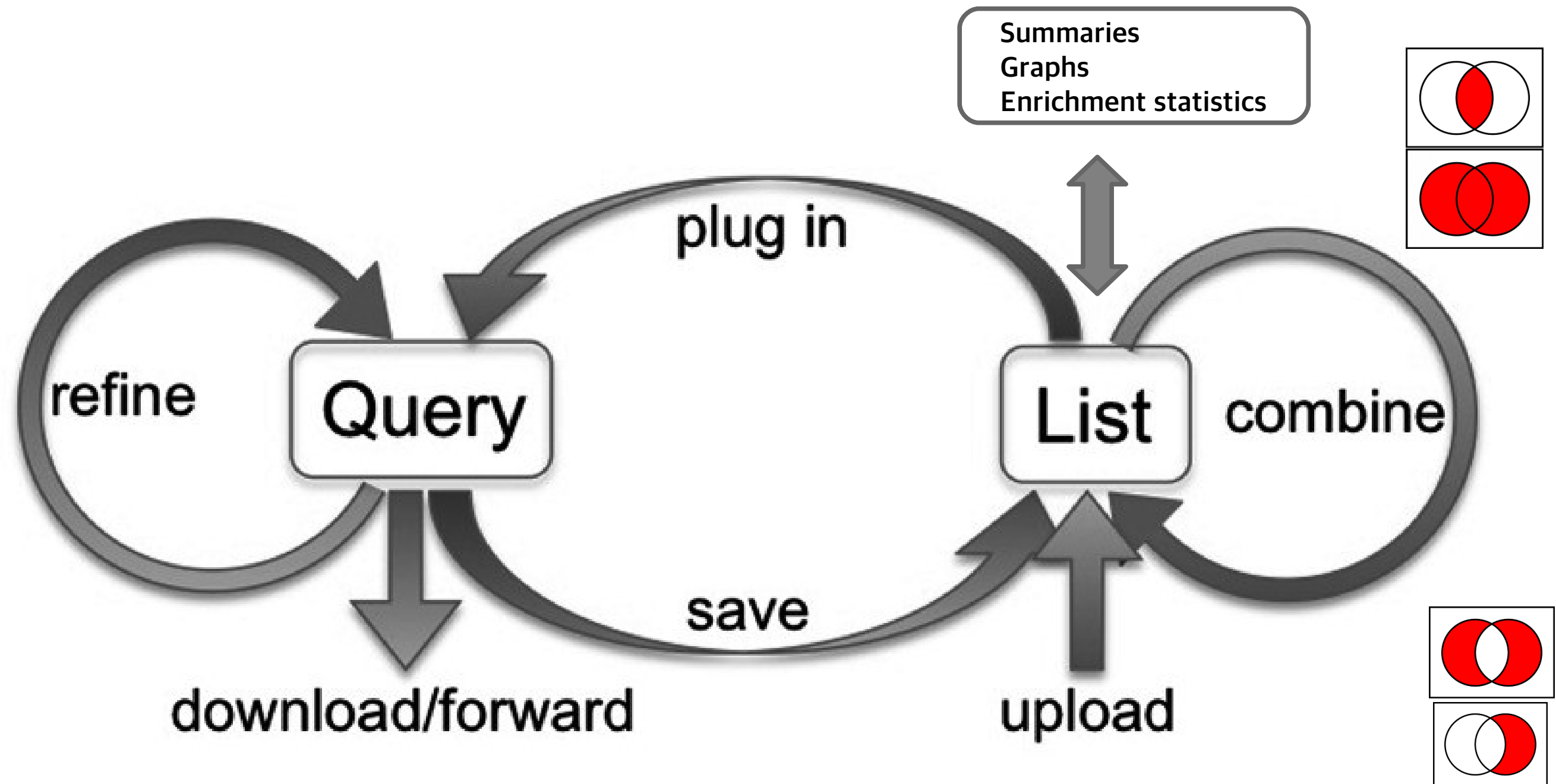


Subtraction



Asymmetric difference

Analysis Workflows



Motenko H, Neuhauser SB, O'Keefe M, Richardson JE. MouseMine: a new data warehouse for MGI. Mamm Genome. 2015 Aug;26(7-8):325-30. doi: 10.1007/s00335-015-9573-z. PubMed PMID: 26092688; PubMed Central PMCID: PMC4534495

The InterMine Team

PI

Gos Micklem

Software
Developers

Daniela Butano
Sergio Contrino
Kevin Herald Reierskog
Yo Yehudi
Adrian Rodriguez Bazaga

Biologist

Rachel Lyne

