

InterMine and InterMOD

Perspective of an Early Adapter

Ben Hitz

GMOD meeting

April 2012

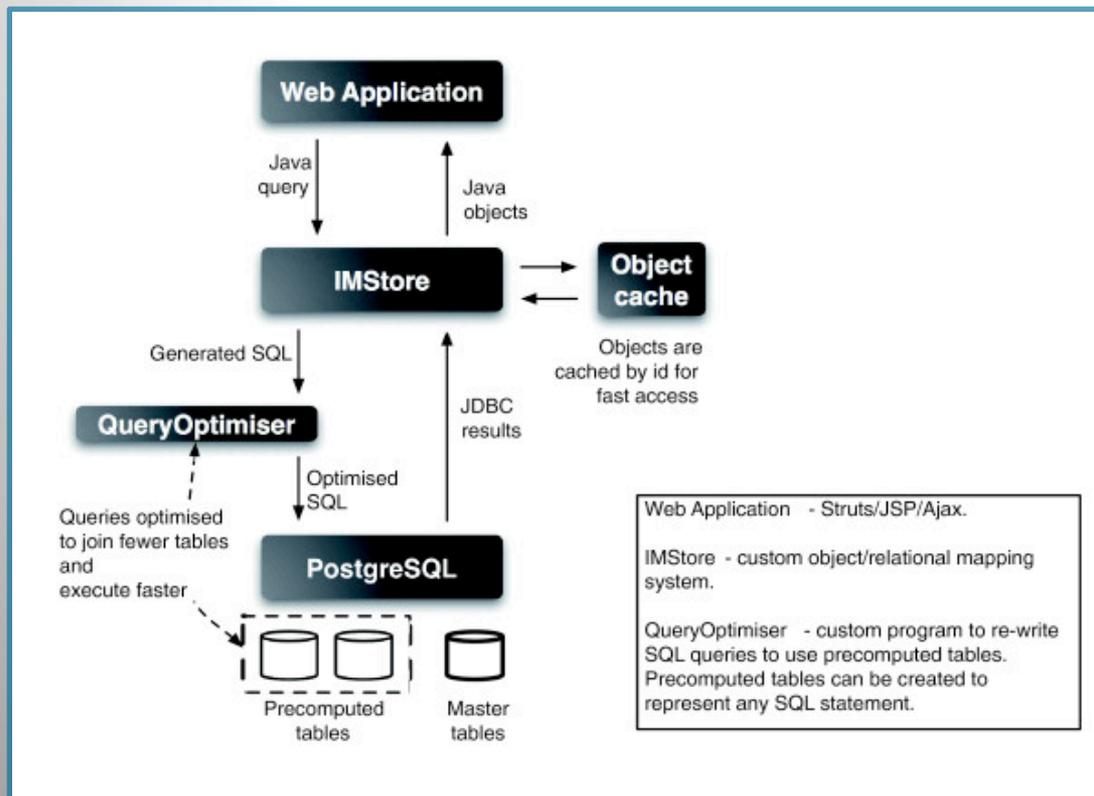
Overview

- ◆ What is InterMine?
- ◆ What is InterMOD
- ◆ Why did SGD get involved?
- ◆ The Good, The Bad, and The Ugly

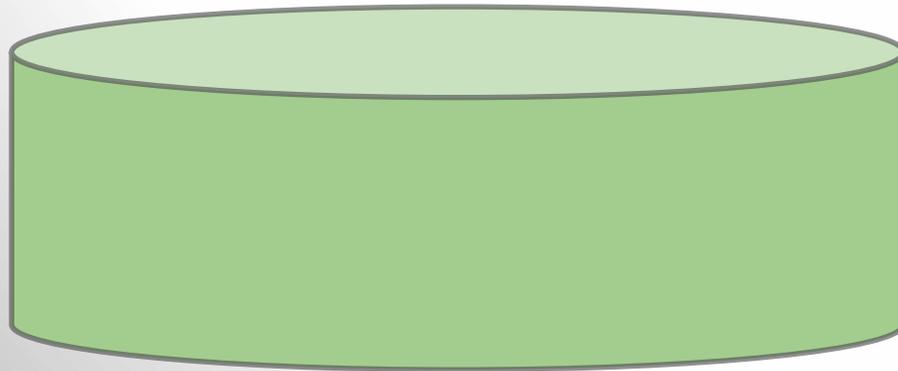
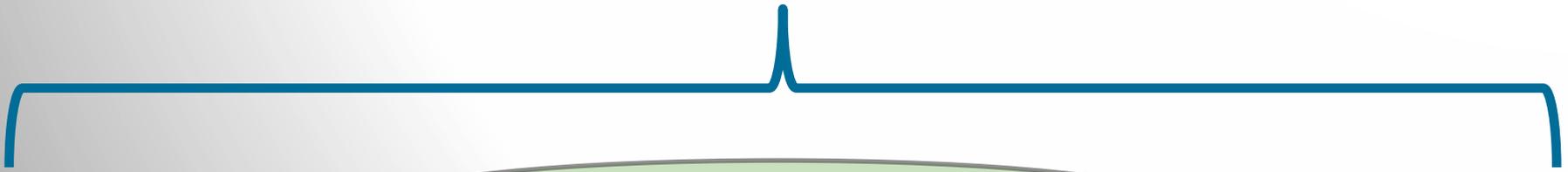
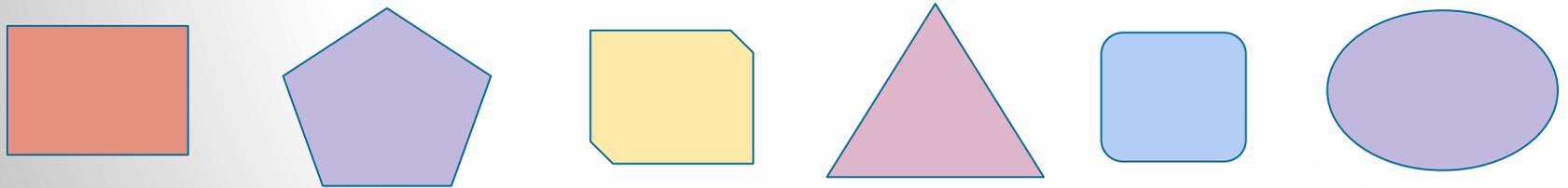


What is InterMine?

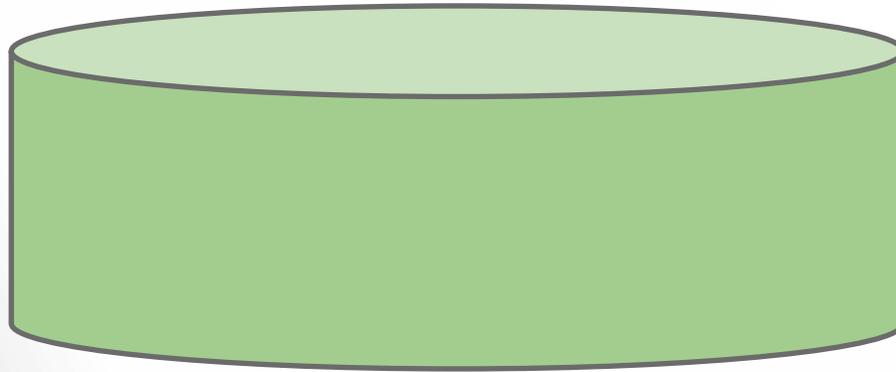
“Powerful Open Source Java Data Warehouse Project”



Data Sources



All For One, and One For All



**JSP Web
Application**

Web Services

Current Mines

Released and in development:

- ◆ MODs for **fruit fly**, mouse, **rat**, worm, zebrafish, slime mold and **budding yeast**
- ◆ modENCODE project (**modMine**)
- ◆ **metabolicMine**, HumanMine, GOMine
- ◆ Other MODs have expressed interest

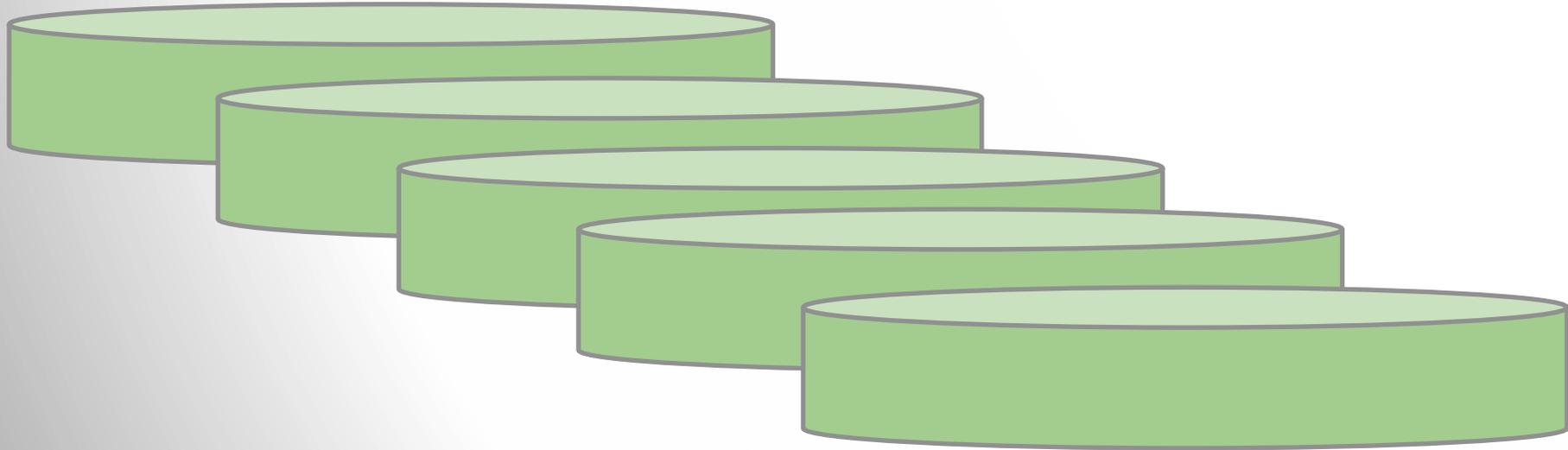
What is InterMOD?

2002: InterMine = new Project();

2009: InterMOD = new Collaboration(
InterMine, SGD, RGD, ZFIN);

2011: InterMOD.addCollaborators (Wormbase, MGI);

All For One, and One For All, For All



**JSP Web
Application**

Web Services

“There are two kinds of people in this world. Those with a rope around the neck, and the people who have the job of doing the cutting” – Tuco (“The Ugly”)

Why did SGD adopt InterMine?

“Solve All Our Problems”

- ◆ “End Around” for Legacy Code Base
 - ◆ Batch Download
 - ◆ Advanced Search
- ◆ MOD Interoperability
 - ◆ Web Services



Search

Search YeastMine. Enter name, identifier or keyword for genes, proteins, ontology terms, authors, abstract etc. (e.g. *rad54*, Act1p, DNA binding, Betel D).

[SEARCH](#)

Analyse

Enter a list of identifiers.

```
rad51; rad52; rad53; ddc1;
rad55; rad57; spo11; dmcl;
rad17; rad9; rad24; msh1;
msh5; mre11; xrs2; rad80;
```

[advanced](#)[ANALYSE](#)

Welcome Back!

See how YeastMine works from our video tour..

[TAKE A TOUR](#)[GENES](#)[PROTEINS](#)[GENE ONTOLOGY](#)[PHENOTYPES](#)[LITERATURE](#)[INTERACTIONS](#)[PATHWAYS](#)[GENE EXPRESSION](#)[SEQUENCE](#)[Read more](#)

Query for genes:

- Chromosome ➔ All genes.
- Gene ➔ Features within a specific distance
- Feature Type ➔ Features of a selected feature Type
- Gene ➔ Homologs
- Chromosomal Region ➔ All genes
- All genes in organism ➔ All Chromosomal Features
- Gene ➔ Chromosomal location
- All genes of a selected Feature Type ➔ Genes with introns

[» More queries](#)

popular templates



You can use '*' as a wildcard when entering text constraint values. Hide



Phenotype → Genes

List genes that are annotated to selected phenotypes. Genes include Dubious, Uncharacterized and Verified ORFs, pseudogenes, transposable element genes, RNAs, and genes Not in Systematic Sequence of S228C.

Phenotype > Observable

= Resistance to chemicals

Show Results

Edit Query



Powered by InterMine. YeastMine is a collaboration between SGD and the InterMine project at the Cambridge Systems Biology Centre. YeastMine is funded by NHGRI grant 2R01HG004834. Copyright © 1997-2012 The Board of Trustees of Leland Stanford Junior University. Permission to use the information contained in this database was given by the researchers/institutes who contributed or published the information. Users of the database are solely responsible for compliance with any copyright restrictions, including those applying to the author abstracts. Documents from this server are provided "AS-IS" without any warranty, expressed or implied. The SGD project at Stanford University is supported by a Genome Research Resource Grant from the US National Human Genome Research Institute, part of the US National Institutes of Health.

List Analysis for interacting_gene_list_copy2 (38 Genes)

Export Edit Find

Gene > Primary SGDID	Gene > Systematic Name	Gene > Standard Name	Gene > Name	Gene > Alias	Gene > Length
<input type="checkbox"/> S000000009	[ORF] YAL011W	[ORF] SWC3	[ORF] SWr Complex	SWC1	1878
<input type="checkbox"/> S000000199	[ORF] YBL103C	[ORF] RTG3	[ORF] ReTroGrade regulation	[no value]	1461
<input type="checkbox"/> S000000292	[ORF] YBR088C	[ORF] POL30	[ORF] POLymerase	PCNA	777
<input type="checkbox"/> S000000435	[ORF] YBR231C	[ORF] SWC5	[ORF] SWr Complex	AOR1	912
<input type="checkbox"/> S000000566	[ORF] YCL061C	[ORF] MRC1	[ORF] Mediator of the Replication Checkpoint	YCL060C	3291
<input type="checkbox"/> S000000662	[ORF] YCR066W	[ORF] RAD18	[ORF] RADiation sensitive	[no value]	1464
<input type="checkbox"/> S000000975	[ORF] YER173W	[ORF] RAD24	[ORF] RADiation sensitive	RS1	1980
<input type="checkbox"/> S000001412	[ORF] YIL150C	[ORF] MCM10	[ORF] MiniChromosome Maintenance	DNA43	1716
<input type="checkbox"/> S000001602	[ORF] YKL119C	[ORF] VPH2	[ORF] Vacuolar pH	VMA12 CLS10	648
<input type="checkbox"/> S000001738	[ORF] YKR030W	[ORF] GMH1	[ORF] Gea1-6 Membrane-associated High-copy suppressor	MSG1	822

Convert

Orthologues

A. gambiae

Convert

No conversions available

External Links

RatMine
 D.erio orthologues at ZFIN protomine
 D.mel orthologues in FlyMine

Selected: << First < Previous | Next > Last >> | View all 38 records as re:

Widgets displaying properties of 'interacting_gene_list_copy2'

Click to select widgets you would like to display: Gene Ontology Enrichment Publication Enrichment Pathway Enrichment Interactions

Gene Ontology Enrichment

close

GO terms enriched for items in this list. For more information about the math used in these calculations, see [here](#).
 Number of Genes in this list not analysed in this widget: 0

Options
 Multiple Hypothesis Test Correction: Holm-Bonferroni
 Ontology: biological_process
 Maximum value to display: 0.05

View Download

GO Term	p-Value
<input type="checkbox"/> DNA repair [GO:0006281]	4.9717E-12 19
<input type="checkbox"/> response to DNA damage stimulus [GO:0006974]	5.739E-12 20
<input type="checkbox"/> postreplication repair [GO:0006301]	4.1462E-11 9
<input type="checkbox"/> DNA metabolic process [GO:0006259]	5.3172E-11 23
<input type="checkbox"/> DNA replication [GO:0006260]	3.7018E-9 14
<input type="checkbox"/> cellular response to stimulus [GO:0051716]	1.4907E-7 24

Publication Enrichment

close

Publications enriched for genes in this list. For more information about the math used in these calculations, see [here](#).
 Number of Genes in this list not analysed in this widget: 0

Options
 Multiple Hypothesis Test Correction: Holm-Bonferroni
 Maximum value to display: 0.05

View Download

Publication	p-Value
<input type="checkbox"/> Mgs1 and Rad18/Rad5/Mms2 are required for survival of Saccharomyces cerevisiae mutants with novel temperature/cold sensitive alleles of the DNA polymerase delta subunit, Pol31. [PubMed:16949354]	3.7924E-22 12
<input type="checkbox"/> Post-replication repair suppresses duplication-mediated genome instability. [PubMed:20463880]	5.6811E-14 11
<input type="checkbox"/> The F-Box Protein Dia2 Overcomes Replication Impedance to Promote Genome Stability in Saccharomyces cerevisiae.	2.589E-13 12

Good

- ◆ It Works!
- ◆ Extremely Fast, Extremely Flexible
- ◆ List Manipulation
- ◆ Excellent Support and Collaboration
- ◆ Web Services and APIs
 - ◆ iPhone app
 - ◆ Cytoscape web toy

“Bad”

- ◆ Not trivial to install or maintain
- ◆ Flexible only to a point
- ◆ Mod interoperability is not there yet

Ugly

Search: GO

Show results

Model browser

- With Text [SHOW](#) [CONSTRAIN](#)
- Data Sets Data Set [SUMMARY](#) [CONSTRAIN](#)
- Evidence GO Evidence [SUMMARY](#) [CONSTRAIN](#)
- Ontology Term Ontology Term [SUMMARY](#) [CONSTRAIN](#)
- Ontology Annotations Ontology Annotation [SUMMARY](#) [CONSTRAIN](#)
- Description [SHOW](#) [CONSTRAIN](#)
- Identifier [SHOW](#) [CONSTRAIN](#)
- Name [SHOW](#) [CONSTRAIN](#)
- Namespace [SHOW](#) [CONSTRAIN](#)
- Obsolete Boolean [SHOW](#) [CONSTRAIN](#)
- Data Sets Data Set [SUMMARY](#) [CONSTRAIN](#)
- Ontology Ontology [SUMMARY](#) [CONSTRAIN](#)
- Parents Ontology Term [SUMMARY](#) [CONSTRAIN](#)
- Description [SHOW](#) [CONSTRAIN](#)
- Identifier [SHOW](#) [CONSTRAIN](#)
- Name [SHOW](#) [CONSTRAIN](#)
- Namespace [SHOW](#) [CONSTRAIN](#)
- Obsolete Boolean [SHOW](#) [CONSTRAIN](#)
- Data Sets Data Set [SUMMARY](#) [CONSTRAIN](#)
- Ontology Ontology [SUMMARY](#) [CONSTRAIN](#)
- Ontology Annotations Ontology Annotation [SUMMARY](#) [CONSTRAIN](#)
- Ontology Term Ontology Term [SUMMARY](#) [CONSTRAIN](#)
- Qualifier [SHOW](#) [CONSTRAIN](#)

Query Overview

Gene

LOOKUP YAL018C IN *S. cerevisiae* (A)

Primary SGDID (X)

Qualifier (X)

!= Dubious (D) (X)

IS NULL (E) (X)

Systematic Name (X)

Status (X)

= Active (B) (X)

IS NULL (C) (X)

Standard Name (X)

GO Annotation GO Annotation collection (X)

Qualifier (X)

Evidence GO Evidence collection (X)

Code GO Evidence Code (X)

Code (X)

Ontology Term Ontology Term (X)

Identifier (X)

Name (X)

Namespace (X)

Constraint logic: A and (B or C) and (D or E)

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 [?](#) to choose a column to sort results by, click again to select ascending [?](#) or descending [?](#). Use the [REMOVEALL](#) link to remove all fields from the results table.

[REMOVEALL](#)

A)

traint

column order. Click [?](#) to choose a results table.

[REMOVEALL](#)

Phenotype > Genes > Primary SGDID (no description) ? ? ?	Phenotype > Genes > Systematic Name (no description) ? ? ?	Phenotype > Genes > Standard Name (no description) ? ? ?	Phenotype > Genes > Qualifier (no description) ? ? ?
Phenotype > Genes > Alias (no description) ? ? ?	Phenotype > Experiment Type (no description) ? ? ?	Phenotype > Mutant Type (no description) ? ? ?	Phenotype > Observable (no description) ? ? ?
Phenotype > Allele (no description) ? ? ?	Phenotype > Strain Background (no description) ? ? ?	Phenotype > Chemical (no description) ? ? ?	Phenotype > Condition (no description) ? ? ?
Phenotype > Details (no description) ? ? ?	Phenotype > Reporter (no description) ? ? ?		

Future Directions

- ◆ Interactive result tables
- ◆ Graphical widgets
- ◆ Custom display components
- ◆ Dynamic query-builder
- ◆ More federation

*“There are two kinds of people in this world.
Those with loaded guns, and those who dig”
– Blondie (“The Good”)*



Cambridge University

Julie Sullivan
Richard Smith
Alex Kalderimis
Radek Stepan

Gos Micklem

<http://yeastmine.yeastgenome.org/>

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

<http://intermine.org>

<http://www.yeastgenome.org>



Stanford University

Kalpana Karra
Gail Binkley
Stuart Miyasato
SGD Project

Mike Cherry

