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”

http://intermine.org/wiki/TechnicalOverview
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

http://mods.intermine.org/wiki/Interoperability
“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).

e.g. act1

Search

Analyse

Enter a list of identifiers.

Gene

advanced

ANALYSE

Welcome Back!

See how YeastMine works from our video tour.

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
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 S288C.

Phenotype > Observable

- Resistance to chemicals

Show Results

< embed results />

web service URL

Perl | Python | Ruby | Java (help)

export XML

Powered by InterMine
## List Analysis for interacting_gene_list_copy2 (38 Genes)

<table>
<thead>
<tr>
<th>Gene &gt;</th>
<th>Systematic Name</th>
<th>Gene &gt;</th>
<th>Standard Name</th>
<th>Gene &gt;</th>
<th>Name</th>
<th>Gene &gt;</th>
<th>Alias</th>
<th>Gene &gt;</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>S000000009</td>
<td>YAL011W</td>
<td>S000000099</td>
<td>[ORF]</td>
<td>SWC3</td>
<td>[ORF]</td>
<td>ReTroeGade regulation</td>
<td>PCNA</td>
<td>777</td>
<td></td>
</tr>
<tr>
<td>S000000009</td>
<td>YBR088C</td>
<td>S000000023</td>
<td>[ORF]</td>
<td>POL30</td>
<td>[ORF]</td>
<td>Polymerase</td>
<td>[no value]</td>
<td>1461</td>
<td></td>
</tr>
<tr>
<td>S000000009</td>
<td>YBR231C</td>
<td>S000000045</td>
<td>[ORF]</td>
<td>SWC5</td>
<td>[ORF]</td>
<td>Swr Complex</td>
<td>AOR1</td>
<td>912</td>
<td></td>
</tr>
<tr>
<td>S000000009</td>
<td>YCL061C</td>
<td>S000000056</td>
<td>[ORF]</td>
<td>MRC1</td>
<td>[ORF]</td>
<td>Mediator of the Replication Checkpoint</td>
<td>YCL060C</td>
<td>3291</td>
<td></td>
</tr>
<tr>
<td>S000000009</td>
<td>YCR066W</td>
<td>S000000062</td>
<td>[ORF]</td>
<td>RAD18</td>
<td>[ORF]</td>
<td>Radiation sensitive</td>
<td>[no value]</td>
<td>1464</td>
<td></td>
</tr>
<tr>
<td>S000000009</td>
<td>YLR150C</td>
<td>S0000001412</td>
<td>[ORF]</td>
<td>MCM10</td>
<td>[ORF]</td>
<td>MiniChromosome Maintenance</td>
<td>DNA3</td>
<td>1716</td>
<td></td>
</tr>
<tr>
<td>S000000009</td>
<td>YKL119C</td>
<td>S0000001602</td>
<td>[ORF]</td>
<td>VPH1</td>
<td>[ORF]</td>
<td>Vascular pH</td>
<td>VMA12 CLS10</td>
<td>648</td>
<td></td>
</tr>
<tr>
<td>S000000009</td>
<td>YKR030W</td>
<td>S0000001738</td>
<td>[ORF]</td>
<td>GMH1</td>
<td>[ORF]</td>
<td>Gea1-6 Membrane-associated High-copy suppressor</td>
<td>MSG1</td>
<td>822</td>
<td></td>
</tr>
</tbody>
</table>

### Widgets displaying properties of 'interacting_gene_list_copy2'

- **Gene Ontology Enrichment**
- **Publication Enrichment**
- **Pathway Enrichment**
- **Interactions**

**Gene Ontology Enrichment**

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

**Publication Enrichment**

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
- Ontology: biological_process
- Maximum value to display: 0.05
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
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”)