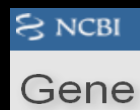# Data-Mining with InterMine

from queries, to analysis, to display

InterMine is:

- An integrated data-warehouse

- An optimising and caching query engine

- A set of web services

- **A suite of embeddable widgets**

You are here

You are here

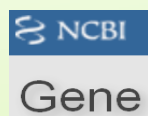| User Scripts and Programs | List Widgets | Tables |

| Java | Perl | Python | Ruby | JS |

Web-Application

Web-Services

Internal API

Object Store

Who runs InterMine?

InterMOD

- SGD (*S. cerevisiae*)
- RGD (*R. norvegicus*)
- ZFIN (*D. rerio*)
- Wormbase (*C. elegans*)
- MGI (*M. musculus*)

- FlyMine (*D. melanogaster*)
- modMine
- metabolicMine
- TargetMine
- mitoMine

Also: S. pombe, FlyBase, TAIR

Goals

- Flexible to do what I want

- Quick to get it done

- Usable for ordinary mortals

Features

- Queries - predefined and custom

- Saved result lists

- Reports and Analysis

- Flexible export

- API to all features

Queries

- What are all the genes in the same pathway as my favourite gene, and annotated with a certain ontology term?

- What are the features in this set of regions?

Templates and Tables

Gene ➡ Pathway. ⭐

Show the Pathway identifier(s) and name for the selected gene. (Data Source: KEGG, REACTOME or FlyReactome).

query history: 8 states▾

⬇ Download   ⚒ Get Code ▾   ▤ Create / Add to List ▾

Showing 1 to 25 of 9,815 rows   ⊞ Columns   ▼ Filters

⊩ ⟵ ⟵   p. 1   ⟶ ⟶ ⟶⊪   Rows per page: 25 ⇅

Gene > Pathways > **Name**	Gene > Pathways > Genes > **Symbol**	Gene > Pathways > Genes > GO Annotation > **Ontology Term . Name**
Activation of Downstream Effectors	Cka	JNK cascade
Activation of Downstream Effectors	Cka	compound eye ph
Activation of Downstream Effectors	Cka	cytoplasm
Activation of Downstream Effectors	Cka	dorsal closure
Activation of Downstream Effectors	Cka	mitotic spindle
Activation of Downstream Effectors	Cka	nucleus
Activation of Downstream Effectors	Cka	phagocytosis, eng
Activation of Downstream Effectors	Cka	positive regulation
Activation of Downstream Effectors	Cka	protein binding
Activation of Downstream Effectors	DAAM	Rho GTPase bind
Activation of Downstream Effectors	DAAM	actin binding
Activation of Downstream Effectors	DAAM	actin cytoskeleton
Activation of Downstream Effectors	DAAM	actin filament bun
Activation of Downstream Effectors	DAAM	axonal growth con

All 33 Ontology Term Names (filtered from 1,179)

↻ | cyto

Item

- ☐ cytoplasm
- ☐ phagocytosis, engulfment
- ☐ actin cytoskeleton organization
- ☐ cytokinesis
- ☐ immune response-regulating cell surface receptor si
- ☐ microtubule cytoskeleton organization involved in est
- ☐ cytoskeleton organization

Filter   Reset   Select All

Lists

- I have found a list of genes - can I

  - query over all of them at once?

  - See what characterises them?

  - See which are present in another list?

Reports and Analysis

- For individual items: its properties and connections

- For groups of items: What is most significantly associated with these items, as a group.

Analysis Widgets
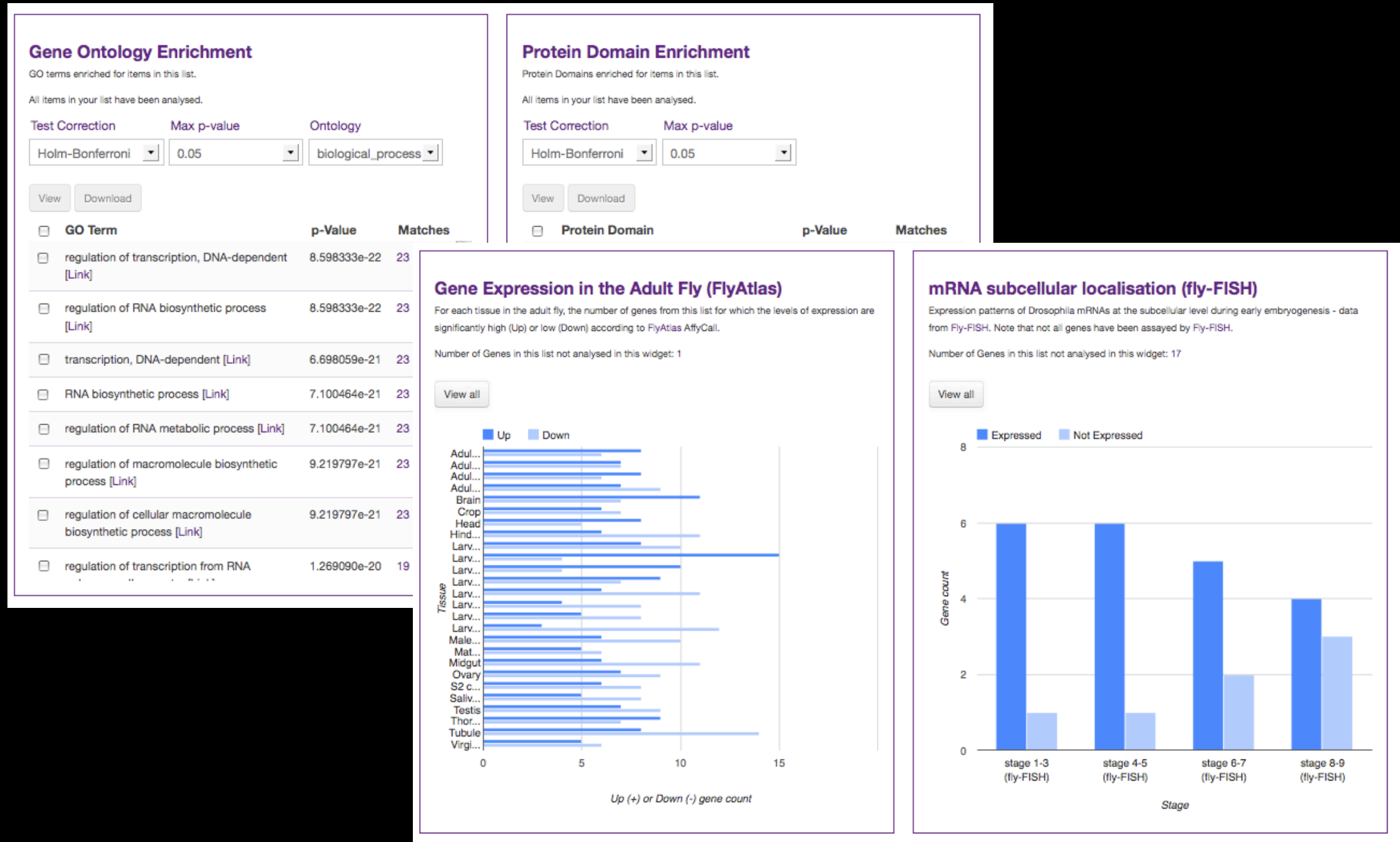
APIs and Libraries

- Support for:
  - ✓ Python
  - ✓ Perl
  - ✓ Ruby
  - ✓ Java
  - ✓ JavaScript (cross-browser & node.js)

```python
from intermine.webservice import Service
s = Service('www.flymine.org/query')
q = s.model.Gene.select('*', 'proteins.proteinDomains.*')\
                .where('pathways', 'LOOKUP', 'pentose*')

for gene in q:
    print gene.symbol
    for protein in gene.proteins:
        for domain in protein.proteinDomains:
            print domain.name
```

```
from intermine.webservice import Service
s = Service('www.flymine.org/query')
q = s.model.Gene.select('*', 'proteins.proteinDomains.*')\
            .where('pathways', 'LOOKUP', 'pentose*')

for row in q.rows():
    print row['name'], row['proteins.proteinDomains.name']
```

Over to You

- www.flymine.org

- yeastmine.yeastgenome.org

- ratmine.mcw.edu

- intermine.modencode.org

Over to You

```
<div id="my-table"></div>


<script>
(function($) {
  var q = { from: 'Gene', select: ['symbol', 'proteins.*'], where: [{pathways: {LOOKUP: 'pentose*'}}] };
  $('#my-table').imWidget({ type: 'table', url: 'www.flymine.org/query', token: 'MY-TOKEN', query: q });
)( jQuery );
</script>
```

Toolkit

One stop shop:

- All files served from cdn.intermine.org

  ➡ *Dependencies (jquery, backbone, bootstrap)*

  ➡ im.js

  ➡ im-tables.js

  ➡ widgets.js

InterMine

Alex Kalderimis, Daniela Butano, Adrian Carr , Sergio Contrino, Hu Fengyuan, Mike Lyne, Rachel Lyne, Julie Sullivan, Radek Štěpán, Gos Micklem

intermine.org

Systems Biology Centre, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK
+44 1223 760262
dev@intermine.org

UNIVERSITY OF CAMBRIDGE

National Human Genome Research Institute

Supported by **wellcome**trust