



Data Annotation & Exploration with Ratmine

Andrew K Vallejos¹, Jennifer R Smith², Richard Smith³, Julie Sullivan³, Gos Micklem³, and Simon N Twigger^{1,2}

¹ Biotechnology and Bioengineering Center and ²: Human and Molecular Genetics Center, Medical College of Wisconsin, Milwaukee, WI, USA, ³ Department of Genetics, University of Cambridge, UK.

<http://ratmine.mcw.edu>

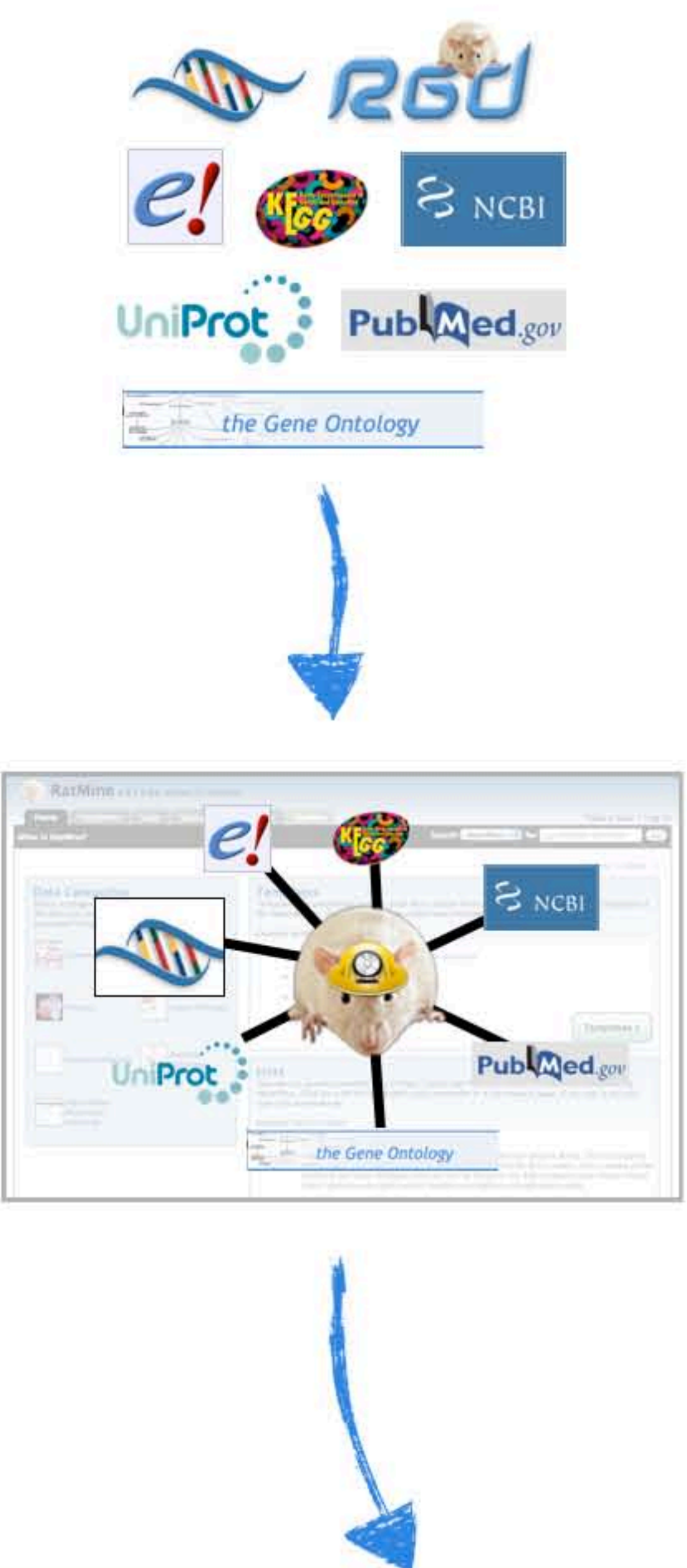


RGD data is combined with annotation from a wide variety of other sources and integrated within the Ratmine data warehouse.

Ratmine builds on RGD's curated data and provides flexible data analysis and reporting tools.

Genes, QTL, GO, disease, Ensembl SNPs, phenotypes, pathways, proteins, literature and more are available.

Human and Mouse orthologs, Protein-Protein interactions and dbSNP data sets are in development.



Template examples
Running a template creates a list of search results. These can be saved and analyzed further using the List Analysis feature

Disease Term [and children of this term] -> Genes
Show genes associated with Disease term (and any children of this Disease term)

Pathway Term [and children of this pathway] -> Genes
Show genes associated with this pathway (and any subpathways of this pathway)

Region -> SNP
Return a list of SNPs located in a specified region.

QTL -> Genes in QTL region
Return the genes located with a particular QTL region

QTLs for Trait -> Genes
List of genes located within the QTLs for a particular trait

Protein Mass -> Protein and Gene
Return a list of Proteins and Genes with a Protein Weight range.

Chromosomal Region -> All Genes
Show the genes located on a chromosome.

CD term name [and children of this term] -> Genes
Find all genes that are associated with a particular CD term.

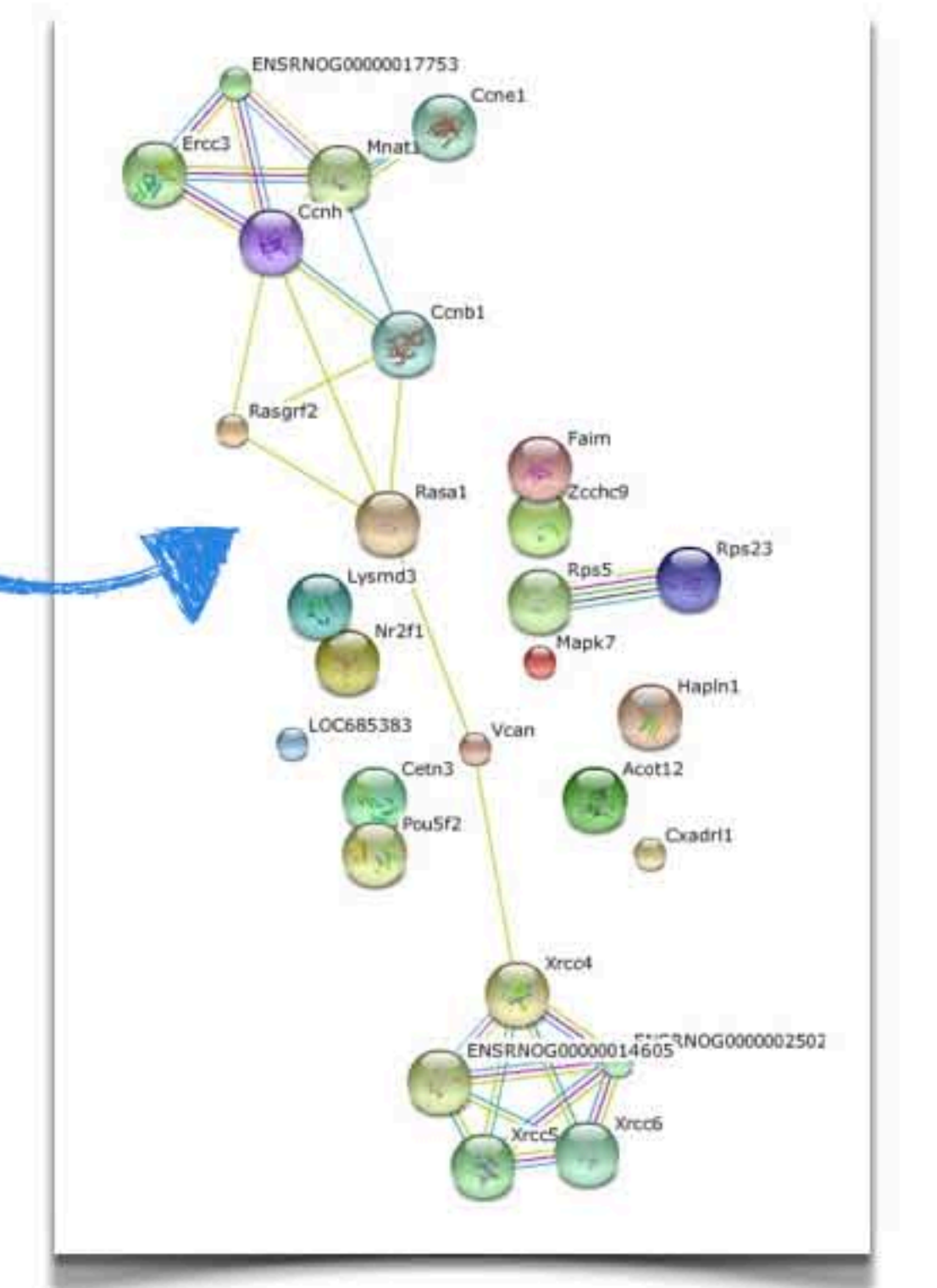
Protein -> Protein Domains
Return a list of protein domains for the given protein

List Analysis View
Edit or download your list of genes or proteins, view enrichment analyses to spot trends in your data

RatMine v 0.1 A Rat version of Intermine

List Analysis for Genes in Mcs1 Region (33 Genes)

Gene ID	Gene Symbol	Gene Name	Gene Length
1305526	1305526	Fam172a	family with sequence similarity 172, member A
1306703	1306703	Mblac2	metallo-beta-lactamase domain containing 2
1307450	1307450	Zfch9	zinc finger, CCHC domain containing 9
1308805	1308805	Lysm3	LysM, positive peptidoglycan-binding, domain containing 3
1309660	1309660	Tmem1610	transmembrane protein 1610
1315572	1315572	Xic2c4	X-ray repair complementing defective repair in Chinese hamster cells 4
1362101	1362101	Rco1562101	similar to very large G-protein coupled receptor 1
1362968	1362968	Stop2	single-stranded DNA binding protein 2
1363119	1363119	Mezf2	myocyte enhancer factor 2C
1364936	1364936	Atrap1	ATRAse, Hlx transporting, lysosomal accessory protein 1-like



Protein-protein interaction graph for your list (String.embl.de)

Widgets displaying properties of 'Genes in Mcs1 Region'

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.

Gene Ontology Enrichment
GO terms enriched for items in this list. Number of Genes in this list not analysed in this widget: 17

Pathway Ontology Enrichment
Pathway terms enriched for items in this list. Number of Genes in this list not analysed in this widget: 29

Disease Ontology Enrichment
DO terms enriched for items in this list. Number of Genes in this list not analysed in this widget: 28

Mammalian Phenotype Ontology Enrichment
Mammalian Phenotype terms enriched for items in this list. Number of Genes in this list not analysed in this widget: 31

Publication Enrichment
Publications enriched for genes in this list. Number of Genes in this list not analysed in this widget: 15

Widgets
These show various statistics and analyses of the genes or proteins in your list. You can quickly see popular pathways or functions.

- Things to try...**
- Use the MyMine tab to see your current lists and queries.
 - Create an account to save queries and lists between sessions.
 - Click on the [icon] icon in a list view table header...
 - Edit an old query in Query Builder, change the output fields to see how it works.
 - Click 'Embed this query' on a Template query page...

<http://ratmine.mcw.edu>

Data Categories
Select a category to see more information about the data sets included. Each category includes associated templates and lists.

- Genomics
- Quantitative Trait Loci
- Single Nucleotide Polymorphism
- Proteins
- Gene Ontology
- Disease
- Pathway
- Mammalian Phenotype

Templates
Templates are predefined queries, each with a simple form and a description. You can create new templates yourself.

Lists
You can run queries on whole lists of data. Create lists from the results of identifiers. Click on a list to view graphs and summaries in a list analysis view permanently.

Query Builder
You can use the flexible query interface to construct your own data mining queries. A good way to start is by editing an existing template.

Templates
Popular queries have been saved as templates for you to use. You can also create and save your own.

Lists
Upload your own data to form your own list or run a template query to find matching data or explore RGD's curated Cardiovascular, Cancer and other disease gene lists

Lists can be combined or split using powerful set operations

Query Builder
A powerful way to create your own custom queries across the data within the Ratmine database

Templates
Other queries that might be useful have already been run for you and are available at the bottom of the list.

