

RGD: A Tutorial for Data, Searches and Tools

<http://rgd.mcw.edu>

RGD: A Tutorial for Data, Searches and Tools

Contents:

RGD: A Tutorial for Data, Searches and Tools	3
Data and Searches.....	3
Search by keyword, external identifier or RGD ID	3
Search by Position.....	5
Ontology Search and RGD's Gene Pages.....	6
QTL Search and QTL Reports	10
Strain Search and Strain Reports	11
Reference/Author Search and Reference Reports	13
Pathway Annotations and Visualization	14
RGD Tools.....	16
Disease Portals.....	16
BioMart	17
GBrowse	18
FTP Downloads	21

RGD: A Tutorial for Data, Searches and Tools

Data and Searches

Search by keyword, external identifier or RGD ID

1. On any RGD web page, e.g. the RGD home page (<http://rgd.mcw.edu>), keyword searches either of the exact term or using a * as a wildcard can be made using the box at the top right corner of the page.

The screenshot shows the RGD home page with several search features highlighted:

- 1**: Keyword search box at the top right containing "blood".
- 2**: "Search RGD" menu item under the "HOME" tab.
- 3**: "Listing of Search Features" link on the "Search RGD" page.

The "Search RGD" page contains two main search boxes:

- Search RGD By Keyword**: Examples: Lepr, "blood pressure", SS BN Mcw, obesity and blood pressure, "high", pmid:15232614. The input field has "blood" and the search button is yellow.
- Search RGD By Position**: Examples: Chr 2, 12000000 - 130000000. The input fields show "Chr 1", "From 1 (bp)", "To 1 (bp)", and the search button is yellow.

On the right side of the page, there are sections for "Capitalization", "Automatic 'and' queries", "Searching by RGD ID", and "Searching with external identifiers".

2. RGD's Keyword search can also be accessed by clicking on the "Search RGD" menu item under the "HOME" tab. This will take you to the "Search RGD" page.
3. For detailed information about the RGD Keyword search click the "Listing of Search Features" on the "Search RGD" page.

The screenshot shows the "Search RGD" page with four search boxes labeled 4:

- Keyword 2004
- Keyword :232345
- Keyword P16599
- Keyword Tnf

A fifth search box labeled 5 is shown separately, containing "Keyword blood". Each search box has a magnifying glass icon to its right.

4. RGD's Keyword search includes searching by RGD ID. Any numeric value entered into these boxes will be searched for as an RGD ID. To search for a numeric external identifier such as an Entrez Gene or PubMed ID, include a colon (:) in front of the ID. If the search term is not numeric, such as a UniProtKB ID, an external identifier search is always attempted.
5. For the purpose of this example, enter the word "blood" in one of the keyword search boxes and click the magnifying glass icon.

	Rat	Mouse	Human
Genes	89	24	198
QTLs	377	24	77
Strains	50	0	0
SSLPs	0	0	0
References	586	0	0

6. The Search Result page gives a summary of all of the RGD objects which matched the search term. In addition the search returns all of the ontology terms which matched. To see a list of the genes, QTL and/or strains that are annotated to a particular ontology term, click on the term to go to the Ontology Report page. For more information on the Ontology Report page, see below.
7. To see a list of all of the rat genes that matched the search term "blood", click the "View Rat Genes Report" link or the number "89" in the Result Matrix.

RGD ID	Symbol	Name	Description	Chr	Start	Stop	Species	Annotations	Match	Type
1359650	Xk	X-linked Kx blood group (McLeod syndrome) homolog	human homolog is a red cell membrane protein whose deficiency causes the McLeod syndrome, a disorder characterized by blood group, neuromuscular and hematopoietic abnormalities	X	25591152	25624913	Rat	10	description, old_gene_name gene, name	
620451	Rhd	Rh blood group, D antigen	blood group gene; produces antigens on the surface of erythrocytes	5	153638642	153672532	Rat	5	description, old_gene_name gene, name	
628609	Abol1	ABO blood group-like 1	histo-blood group protein with A- and B-transferase activities	3	5253078	5270252	Rat	12	description, old_gene_name gene, name	

8. Information included in the gene result list includes the gene RGD ID, symbol, name, description and chromosomal position as well as the number of ontology annotations associated with that gene.
9. The box in the upper right corner of the page shows the current search term and allows the user to refine the search by changing the search term and/or selecting a particular chromosome. The genomic or genetic position information can be changed by using the drop-down box labeled "Assembly" and the results can be sorted by any of the columns simplifying the task of finding the particular result needed.
10. Tabs at the top of the results box allow users to easily move between results for rat, mouse, human or all three.
11. Results can be exported as CSV or tab-delimited text files, sent to a printer, or viewed in their genomic context using the RGD Genome Viewer. See below for more information on RGD's Genome Viewer.

Search by Position

1. On the RGD home page (<http://rgd.mcw.edu/>), under the tab which reads "HOME" click the menu item "Search RGD". This will take you to the "Search RGD" page.

The screenshot shows the RGD home page with various navigation links like HOME, DATA, GENOME TOOLS, DISEASES, PHENOTYPES & MODELS, and COMMUNITY. A red box labeled '1' highlights the 'Search RGD' link in the top left corner. Below it, there's a search bar for 'Keyword' and another for 'Search RGD By Keyword' with an example 'Lepr Hypertr MCS'. A second red box labeled '2' highlights the 'Search RGD By Position' section, which includes a chromosome map, input fields for 'Chr' (set to 12), 'From' (0.5e6 bp), 'To' (25e6 bp), and a 'Search by Position' button.

2. In the "Search RGD By Position" box, select chromosome 12 and enter 500000 into the "From:" box and 25000000 into the "To:" box. Note that for simplicity mega base pair positions can be entered using exponential notation making these positions 0.5e6 and 25e6, as shown. Click the "Search by Position" button.

The screenshot shows the RGD home page again. A red box labeled '3' highlights the search criteria for 'Search for genes, SSLPs and QTLs by position': 'Chr: 12', 'From: 500000 (bp)', 'To: 25000000 (bp)', and the 'Search' button. Below this, a table titled 'Objects returned by your search' shows the count of objects found for each type: GENES (425), SSLPs (108), QTLs (38), and ALL TYPES (57). A red box labeled '4a' highlights this table.

3. On the resulting page, notice that the search criteria are displayed at the top of the page. The search can be refined or expanded using these boxes.
4. Also notice that a summary table of the objects returned by the search is shown. Click on "QTLs" to navigate down to the table displaying all QTLs located within the search region. You can return to the top of the page using either your browser's "back" button, or the "Back to top of page" link located below each table.

The screenshot shows the results of a QTL search. At the top, a red box labeled '4b' highlights the search criteria: 'Chr: 12', 'From: 500000 (bp)', 'To: 25000000 (bp)', and the 'Search' button. Below this, a red box labeled '7' highlights the 'QTLs found in your search region (export table to spreadsheet)' link. A red box labeled '6' highlights the '61334' entry in the first table. A red box labeled '5' highlights the '5' entry in the second table, which is a detailed view of the QTL 'Gluco17'. Both tables have headers: RGD_ID, TYPE, SYMBOL, NAME, CHR., START, and STOP.

RGD_ID	TYPE	SYMBOL	NAME	CHR.	START	STOP
61331	QTL	Eau2	Experimental allergic uveoretinitis QTL 2	12	1	29130348
1598855	QTL	Bp294	Blood pressure QTL 294	12	1	35270263
1288081	QTL	Cia25	Collagen induced arthritis QTL 25	12	1	43432656
1357337	QTL	Gluc03	Glucose level QTL 3	12	1	12240883
61334	QTL	Gluc017	Glucose level QTL 17	12	1	12240883
2303568	QTL					
61421	QTL					
2303575	QTL					
737979	QTL					
631521	QTL					
2312418	QTL					
634350	QTL					
1300174	QTL					

RGD_ID	TYPE	SYMBOL	NAME	CHR.	START	STOP
61331	QTL	Eau2	Experimental allergic uveoretinitis QTL 2	12	1	29130348
61334	QTL	Gluco17	Glucose level QTL 17	12	1	12240883
61421	QTL	Cia12	Collagen induced arthritis QTL 12	12	1	44521457
2303575	QTL	Insu14	Insulin level QTL 14	12	1	43562201
1357337	QTL	Gluc03	Glucose level QTL 3	12	1	12240883
1598855	QTL	Bp294	Blood pressure QTL 294	12	1	35270263
2303568	QTL	Bw88	Body weight QTL 88	12	1	25457135
1288081	QTL	Cia25	Collagen induced arthritis QTL 25	12	1	43432656
737979	QTL	Pia22	Pristane induced arthritis QTL 22	12	530865	45530865
1626034	GENE	Vom2r-ps90	vomeronasal 2 receptor, pseudogene 90	12	534145	541373
35384	SSLP	D12Rat1	D12Rat1	12	555108	555654
631521	QTL	Pia12	Pristane induced arthritis QTL 12	12	555108	24748499
1583708	GENE	Vom2r568	vomeronasal 2 receptor, 58	12	600121	732796
2312418	QTL	Kidm41	Kidney mass QTL 41	12	665384	20932760
1588806	GENE	Vom2r-ps91	vomeronasal 2 receptor, pseudogene 91	12	736334	820564
1588795	GENE	Vom2r-ps93	vomeronasal 2 receptor, pseudogene 93	12	933091	935363

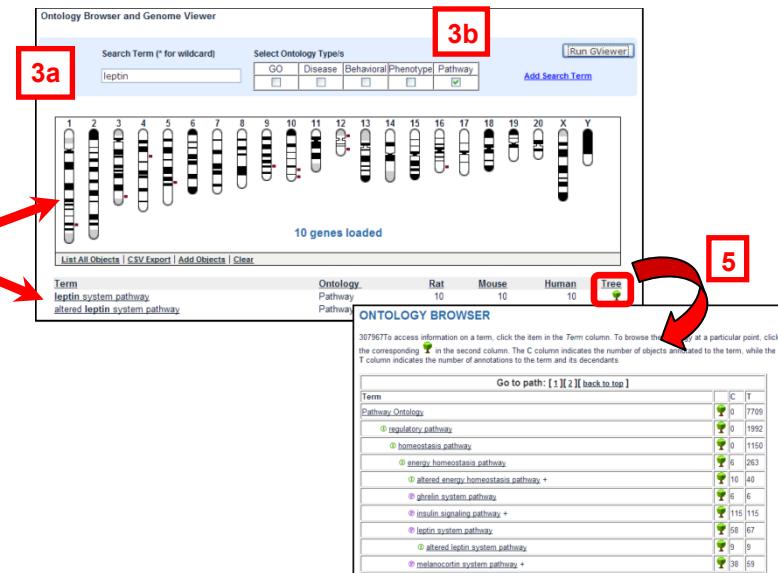
5. From the QTLS table, scroll down to the table which displays all of the objects located within the search region. Notice that the results are ordered by base pair position.
6. Click one of the RGD ID links to navigate to the report page for that gene, QTL or SSLP.
7. Returning to the results page, click the “export table to spreadsheet” link to download the results as a comma-separated text file which can then be opened in Excel for ease of analysis.

Ontology Search and RGD's Gene Pages

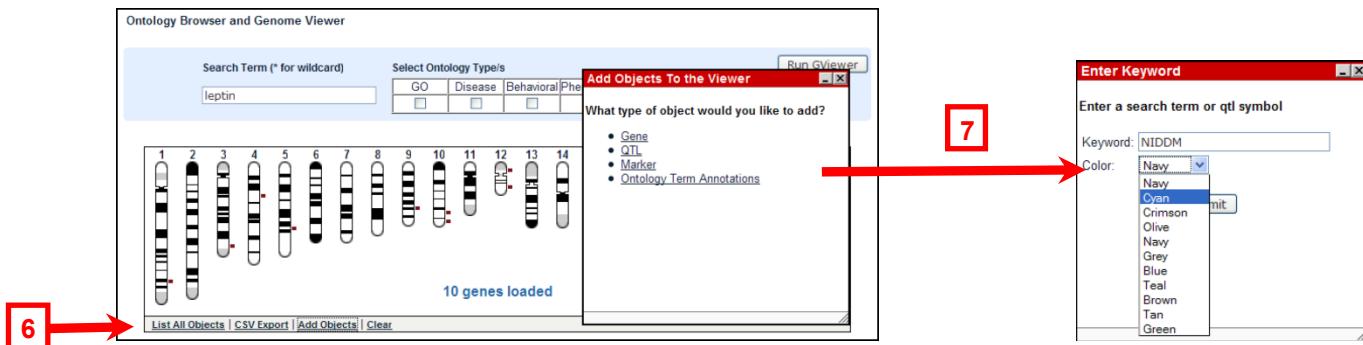
RGD is using several controlled vocabularies for the annotation of genes, QTL and strains. These include the three vocabularies of the Gene Ontology (GO function, process and component) along with Disease (DO), Phenotype (MP) and Pathway (PW) Ontologies. For rat genes all these vocabularies are used, for human genes DO, MP and PW and for mouse genes DO and PW are being used, respectively. DO and MP are also used for the annotations of rat and human QTL and of rat strains.

The image consists of two side-by-side screenshots of the RGD website. The left screenshot shows the main RGD Data page under the 'DATA' tab. A red box highlights the 'ONTROLOGIES' link in the menu. A red number '1' is placed over the 'ONTROLOGIES' link. Below it, another red box highlights the 'GVIEWER' link in the 'GENOME TOOLS' section. A red number '2' is placed over the 'GVIEWER' link. The right screenshot shows the 'Genome Tools' page under the 'GENOME TOOLS' tab. It lists various tools: Rat Genome Browser (highlighted with a red box and red number '2'), VCMas (highlighted with a red box and red number '3'), BioMart (highlighted with a red box and red number '4'), ACP Haplotype (highlighted with a red box and red number '5'), and Genome Scanner (highlighted with a red box and red number '6').

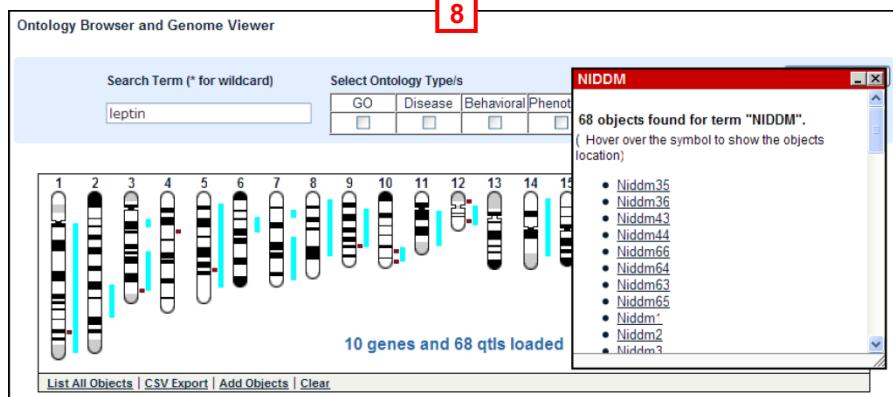
1. The **Ontology Browser** and **GViewer** can be used to search one or several vocabularies for terms of interest and to find what objects have been annotated to them. Click the “DATA” tab to access a page listing all of the data types in RGD. Click “Ontologies” in the list of data types or in the menu at the top to access a simple Ontology Browser.
2. Click the “GENOME TOOLS” tab to access a page listing all RGD’s tools. Click “GViewer” in the menu at the top or “Genome Viewer” in the list of tools to access the Ontology Search/GViewer.



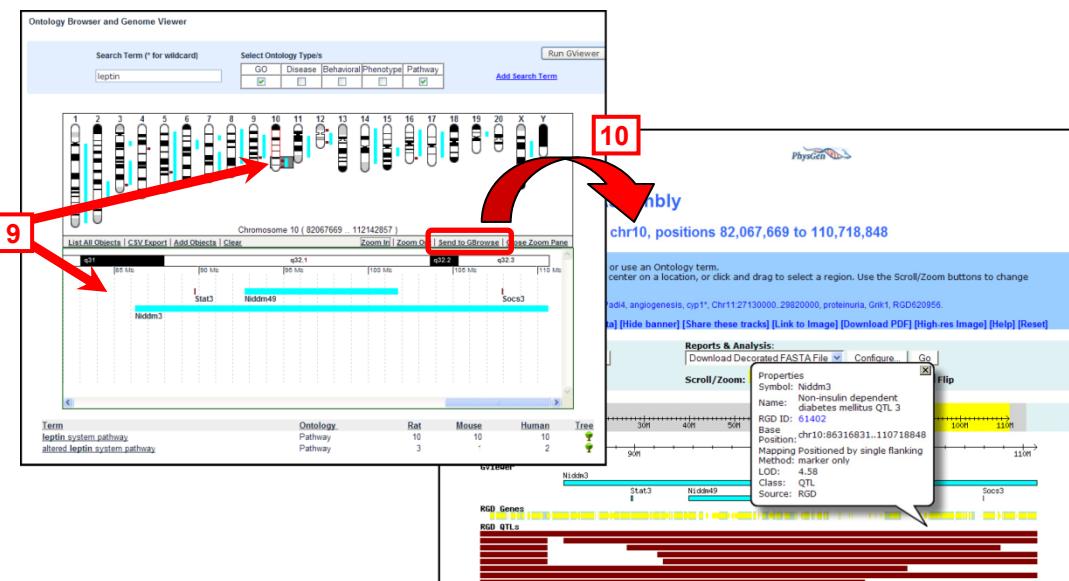
3. As an example, enter “leptin” into the search box. The default is for all of the ontologies to be selected. Click all of the check boxes except “pathway” to search only RGD’s Pathway Ontology. Click “Run GViewer”.
4. The search returns a screen with a GViewer image of all of the genes annotated to PW terms that include the word “leptin”. In addition a box contains a list of all of the ontology terms which match the search criteria with counts of the number of annotations to each term for rat, mouse and human.
5. Note that clicking on the tree icon in the entry for the ontology term will bring up a diagram of the placement of that term in the ontology.



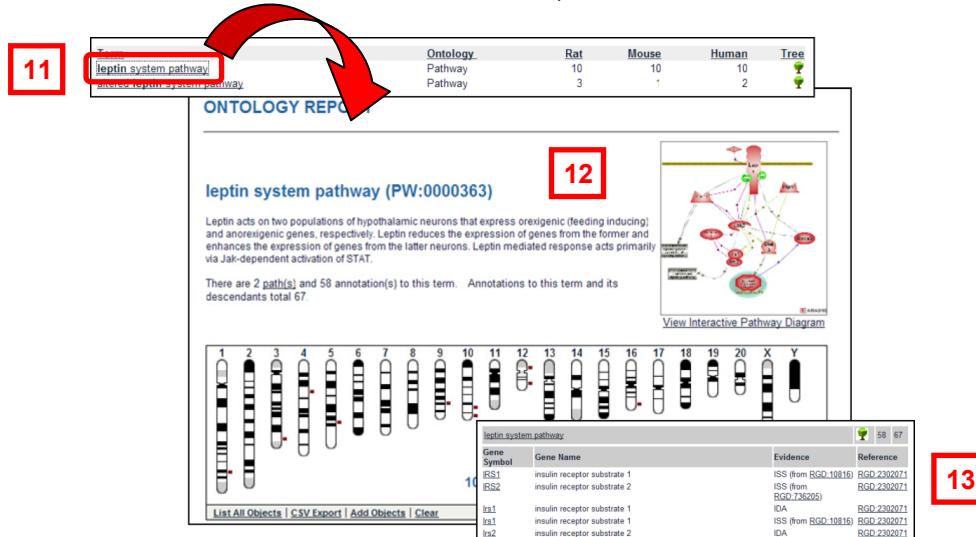
6. Looking at the GViewer tool, you can see a line of options along the bottom of the image which allow you to view or download a list of all of the objects displayed, add more objects to the display or clear the display and start over.
7. If you click the “Add Objects” link a box will open giving you the option of searching for objects such as genes, QTL or SSLP Markers, or for another ontology term to view where objects annotated to that term overlap with the results you are already viewing. For this example, click the link for "QTL" and enter "NIDDM" in the search box. Select "Cyan" from the drop down list of colors and click "Submit".



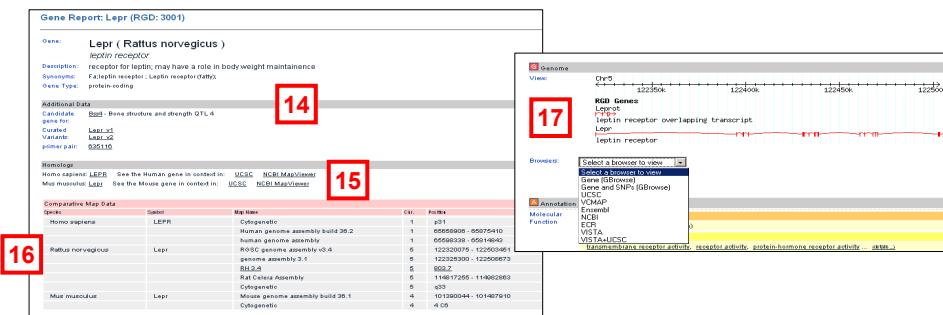
8. All of RGD's rat Niddm QTL are now displayed as cyan bars and a list of the QTL is displayed.



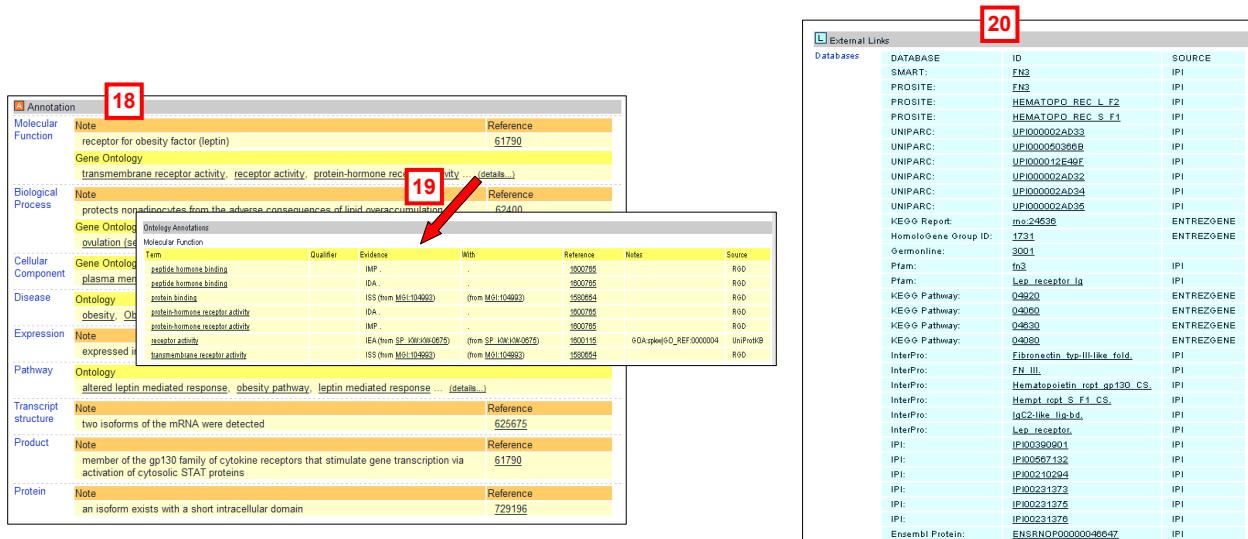
9. To zoom into an area of a chromosome, click on the chromosome at the position you would like to see more closely. A zoom pane showing the selected area with its genes and QTL appears below the main display pane. The area can be moved by sliding the selected area along the chromosome in the main display. You can double click the slider to lock it into place.
10. To see the area you have zoomed into in its broader genomic context, click the “Send to GBrowse” button. RGD’s rat Genome Browser will open up with the GViewer results displayed as a separate track. For more information on RGD’s GBrowse tool, see below.



- Going back to the Genome Viewer results page, click the “leptin system pathway” term in the list below the GViewer to access the corresponding Ontology Report page.
- Note that the report page includes a GViewer image of the rat genes annotated to this term as well as the definition of the term. In addition, when RGD has an interactive pathway diagram associated with the term, an icon linking to that diagram is included on the report page.
- The Ontology Report page also includes a list of genes annotated to this GO term. Clicking on any of the annotated genes brings up the report page for that gene. As an example, scroll down and click on Lepr – the receptor for leptin. This will bring up the gene report for the rat leptin receptor.



- The top of the gene report page for the rat Lepr gene includes information such as current and previous nomenclature, information on splice variants and a notation that this gene is a candidate gene for the Bss4 QTL.
- The “Homologs” section gives links to the RGD gene pages for the mouse and human orthologous genes. Clicking on “UCSC” or “NCBI MapViewer” will take you to the corresponding genome browser image for that gene.
- Mapping information for the rat gene and its orthologs, including map positions for genomic, genetic and RH maps, is included in the “Comparative Map Data section”.
- Scroll down to the “Genome” section of the report to see an intron-exon diagram of the gene. Also included is a selection of browsers for genomic viewing.



- Scroll down to the “Annotations” section of the page to view the manually curated annotations for this gene. These include both free text notes and ontology annotations.
- Only a partial list of terms is included on the gene page. Click “(details...)” to access the full list of ontology annotations for the Lepr gene.
- Returning to the gene page, scroll down again to the “External Links” section. This section provides links to external gene-specific records such as Entrez Gene and Ensembl gene pages, and to protein-specific data. Links to databases such as UniProtKB, IPI and Ensembl protein provide information on

the whole protein, while links such as those for InterPro, Pfam and SMART give information on domains, families and protein structure.

QTL Search and QTL Reports

- From the RGD home page (<http://rgd.mcw.edu>), or any internal page, click on the Data tab at the top of the page. The submenu in the blue bar lists the various types of biological data stored in RGD (A). Click on “QTLs” to open the QTL query page.

Left Screenshot (Search Form):

- 1: DATA tab selected in the blue bar.
- 2: Keyword search box containing "blood pressure".
- 3: Chromosome selection dropdown set to "5".
- 4: Species selection dropdown set to "Rat".
- 5: "Search QTLs" button.

Right Screenshot (Results Table):

- 6: Refine Term: blood pressure, Assembly: RSGC genome assembly v3.4, Sort By: Relevance, Descending.
- 7: Export options: CSV, TAB, Printer, Genome Viewer.
- 8: Row for Bp7 highlighted.
- 9: Column headers: RGID, Symbol, Name, LOD Value, Trait, Sub Trait, Chr, Start, Stop, Species, Annotations, Match, Type.

RGID	Symbol	Name	LOD Value	Trait	Sub Trait	Chr	Start	Stop	Species	Annotations	Match	Type
61396	Bp7	Blood pressure QTL	7.45	.0001	Blood pressure	salt-loaded systolic	5	62555131	168174140	Rat	2	trat, name, qtl
61380	Bp49	Blood pressure QTL	16.6		Blood pressure		5	62555131	103059258	Rat	4	trat, name, qtl
631498	Bp100	Blood pressure QTL	4.9	.001	Blood pressure	systolic	5	165063550	165074526	Rat	2	trat, name, qtl
631505	Bp103	Blood pressure QTL	3.2		Blood pressure	systolic	5	139690001	171801887	Rat	2	trat, name, qtl
631553	Bp131	Blood pressure QTL	4.2	.0005	Blood pressure	systolic	5	165074526	173096209	Rat	2	trat, name, qtl
631606	Bp119	Blood pressure QTL	3.8		Blood pressure	systolic	5	1	40017184	Rat	2	trat, name, qtl
634549	Bp139	Blood pressure QTL		.001	Blood pressure	systolic	5	135877124	173096209	Rat	2	trat, name, qtl
724525	Bp147	Blood pressure QTL	4.3	.0001	Blood pressure	systolic	5	132950669	173096209	Rat	2	trat, name, qtl
1298077	Bp157	Blood pressure QTL	157		Blood pressure	salt-loaded systolic	5	121964552	126386030	Rat	3	trat, name, qtl
1298086	Bp156	Blood pressure QTL	156		Blood pressure	salt-loaded systolic	5	90916908	135916908	Rat	2	trat, name, qtl

- In the keyword search box, enter “blood pressure”.
- The QTL search gives you several additional options that you can use to narrow your search. These include chromosomal position and species. For the genomic position you can select a chromosome number or both a chromosome and base pair position. You also have the option to change the genomic assembly that the position information refers to and/or choose to include or exclude objects that do not map to the selected genome assembly. For this example, select chromosome 5.
- You can also select a species. Notice that RGD also stores information on mouse and human quantitative trait loci and you can search for QTLs from any or all of these species. For this example, leave the default selection, Rat.
- Click on the “Search QTLs” button to run the search.
- Examine the QTL results page. The QTL search result page contains a list of all rat QTLs in RGD which match the search criteria. For each QTL, in addition to the RGD ID, symbol and name, the list gives the LOD score, the p value if available, trait and subtrait, and the base pair position on the chromosome.
- Options in the box at the top of the form allow you to refine your search term, change the genome assembly which positions are given for and sort the results by any of the available columns.
- Search results can be exported as a CSV or tab-delimited text file, sent to a printer, or viewed in the Genome Viewer by clicking on the appropriate link.
- Click on the symbol for Bp7 to go to the RGD report page for that QTL.

10. RGD QTL Report pages supply manually curated, detailed information about quantitative trait loci. In the Description section of the page, notice the trait and subtrait which are linked to this chromosomal locus. The “Method” line gives detailed information on the experimental methods used to measure the trait. Ontology annotations for associated disease and phenotype terms are shown. Click on “hypertension” or on “increased blood pressure” to go to the Disease or Mammalian Phenotype ontology report pages for the respective terms. RGD curates statistical measurements from QTL papers and supplies them in a searchable format. Notice that for this QTL both LOD score and P value are given.
11. RGD links QTL reports to the reports for the strains that were used to generate each QTL. Click on “SS/Jr” to navigate to the Strain Report page for that rat strain.
12. QTLs are defined by corresponding marker positions used to establish linkage. Notice that on the Map Summary line both peak and flanking markers are given. Scroll down to the Map Data section to view the map positions of the three markers. RGD uses genomic, genetic and RH maps to localize markers for QTLs.
13. Just above the Map Data section is a view of the Bp7 QTL on chromosome 5. Use the dropdown list labeled “Select a browser to view” to choose one of the genome browsers within RGD or on other websites such as UCSC or NCBI to view the Bp7 QTL in its genomic context.
14. Scroll down to the “Annotations” section. This section of the page gives free text notes which supply extra information about the QTL. Each note is linked to the reference from which that information was curated. Click on “61059” to read the abstract for the paper from which the note “SS allele associated with increased blood pressure” was derived.
15. The Related Objects section of the page lists any related QTLs and/or candidate genes. Click on “Edn2” to view the gene report page for the candidate gene “endothelin 2”. Note that the gene report page also supplies the information that Edn2 is a candidate gene for Bp7 and contains a link to the QTL.
16. “Other database links” provides links to information on Bp7 in external databases. Click “RatMap” to view the RatMap report or “Entrez Gene” to view the information from NCBI’s Entrez Gene database. The data at Entrez Gene is derived from RGD’s QTL data.
17. The References section links to a page which lists all of the references associated with Bp7. Click “Article(s) about Bp7 (3 curated references)” to view the citations for the three papers associated with Bp7. Links are provided to the abstracts at RGD and at PubMed.

Strain Search and Strain Reports

1. From the RGD home page (<http://rgd.mcw.edu>), or any internal page, click on the Data tab at the top of the page. The submenu in the blue bar lists the various types of biological data stored in RGD (A). Click on “Strains” to open the Strain query page.

The screenshot shows two versions of the RGD website side-by-side. The left version is the 'Strain Search' page, and the right version is the 'Strains search result' page for 'Rattus norvegicus'.

- Strain Search (Left):**
 - Header: RGD logo, PhysGen logo, Keyword search bar.
 - Navigation: HOME, DATA (highlighted), GENOME TOOLS, DISEASES, PHENOTYPES & MODELS, COMMUNITY.
 - Sub-navigation: Genes, QTLs, Strains (highlighted), Markers, ESTs, Maps, Ontologies, Sequences, References.
 - Section: Strain Search. Subtext: 'Strain reports include a comprehensive description of strain origin, disease, phenotype, gene sub-strains, and strain sources.' [Search Help]. Example searches: 'fh', 'blood pressure', 'SS BN Mow', 'g16rat5'.
 - Form: Keyword input 'cop', Search Strains button.
 - Text: 'Switch to classic strain search' and 'View all search features'.
- Strains search result (Right):**
 - Header: RGD logo, PhysGen logo, Keyword search bar.
 - Navigation: HOME, DATA, GENOME TOOLS, DISEASES, PHENOTYPES & MODELS, COMMUNITY.
 - Sub-navigation: Genes, QTLs, Strains, Markers, ESTs, Maps, Ontologies, Sequences, References, FTP Download, Submit Data.
 - Section: Strains search result for 'Rattus norvegicus'. Subtext: 'View Results for all Objects and Ontologies'.
 - Form: Refine Term: 'cop', Update button.
 - Text: 'Sort By: RGD ID, Ascending'.
 - Table: 'Strains' table with 29 records found for search term 'cop'.

Rat	Export This Report To	CSV	TAB	Printer			
RGD ID	Symbol	Name	Origin	Source	Type	Annotations	Match
10011	COP/OlaHsd		Hartley		inbred	5	symbol
60097	DA	DA			inbred	11	origin
60019	COP	COP			inbred	4	symbol
60019	CDRH	Cohen Rosenblat Diabetic Hypertension	Hebrew University Hospital and Hadassah-Hebrew Medical College, Jerusalem, Israel		inbred	11	origin

- The RGD Strain Search works in the same way as the general Keyword search at the top of the page except that it only searches for strains. For this example, enter "cop" in the "Keyword" search box under "Strain Search".
- Click the "Search Strains" button to navigate to the Strains Search results page.
- Notice that the search returns the list of all strains where "COP" occurs anywhere in the indexed fields of the strain record. In addition to the obvious options of symbol and name, these include previous nomenclature and origin. Note in the results list you will see substrains and congenic strains in addition to the parental COP strain.
- Click the link for "COP" to go to the Strain Report page for the parental strain.

The screenshot shows the 'Strain Report' page for the COP/Crl strain.

- Strain Report (Left):**
 - Header: RGD logo, PhysGen logo, Keyword search bar.
 - Navigation: HOME, DATA, GENOME TOOLS, DISEASES, PHENOTYPES & MODELS, COMMUNITY.
 - Sub-navigation: Genes, QTLs, Strains (highlighted), Markers, ESTs, Maps, Ontologies, Sequences, References.
 - Section: Strain Report. Subtext: 'See Also: COP/Crl, COP/OlaHsd'.
 - Form: Basic details section with information: Symbol: COP, Strain: COP, Type: inbred, Origin: COP/Crl, See Also: COP/Crl, COP/OlaHsd.
 - Section: Annotations: LifeSpan and Spontaneous Disease, Drugs and Chemicals.
- Strain Report (Right):**
 - Header: RGD logo, PhysGen logo, Keyword search bar.
 - Navigation: HOME, DATA, GENOME TOOLS, DISEASES, PHENOTYPES & MODELS, COMMUNITY.
 - Sub-navigation: Genes, QTLs, Strains, Markers, ESTs, Maps, Ontologies, Sequences, References.
 - Section: Strain Report for COP/Crl.
 - Form: Basic details section with information: Symbol: COP/Crl, Strain: COP, Substrain: Crl, Name: Copenhagen, Source: Charles River Laboratories, Type: inbred, Origin: Inbred strain is from Curtis in 1921 at Columbia University Institute for Cancer Research. To Charles River from the National Cancer Institute in 1998.
 - Section: Annotations: Lifespan and Spontaneous Disease, Drugs and Chemicals.
 - Table: Ontology table with rows for Disease (Hyperplasia) and Phenotype (Inherited mammary hyperplasia).
 - Table: Strain SSLP data table.
 - Table: Strain QTL data table with rows for COP/Crl, Crl, and Crl-1.
 - Section: Associated References (12 references listed).

- RGD's Strain Report pages display a wealth of valuable information on individual rat strains. In the "Basic Details" section, notice the information on Genetic markers, Coat Color and number of inbred Generations. In the Origin section, parent strain pages give information on the origins of not only the parental strain, but also all of the corresponding substrains.
- In the section above "Basic Details" click the link for "COP/Crl" to go to the page for this substrain. Notice that this Strain Report only gives information for the substrain, but includes a link to both the parent strain and the COP/OlaHsd substrain.
- Scroll down to the "Annotations" section.
- Detailed notes on topics such as "Lifespan and Spontaneous Disease" and "Drugs and Chemicals" have been written by RGD curators. Click the "1334449" link in the "Reference" column beside the first note under "Drugs and Chemicals" to read the abstract for the paper from which the note "Strain has decreased sensitivity to DES-induced pituitary growth" was curated. Notice that the Reference report page shows a link to the PubMed page which includes a link to the free full text of this article.

10. Returning to the Strain Report, scroll down to the Ontology section to view the Disease and Phenotype Ontology annotations that have been assigned to this strain. Click on the link for “mammary alveolar hyperplasia” to review the information for this Mammalian Phenotype term.
11. Returning once again to the strain page, scroll down to view the list of RGD QTLs which were generated using this strain. Click on the link for Ept10 to navigate to the QTL report page for Estrogen-induced pituitary tumorigenesis QTL 10. (See information on the QTL report pages elsewhere in this handout for details on using RGD QTL reports.)

The screenshot shows the RGD Strain Report page for the COP/OlaHsd strain. The 'Ontology' section lists Disease and Phenotype annotations. The 'Strain SSLP data' section shows markers typed in this strain, with one entry highlighted (12). The 'Strain QTL data' section lists QTLs for various traits across chromosomes 1-16. The 'Associated References' section contains a list of 13 references, with the last one (7) highlighted (13).

12. Returning to the COP/Crl strain page, either click the link at the top of the page or use your browser's "Back" button to navigate back to the COP parental strain page. Scroll down to the Strain SSLP data section. Notice that information on the markers typed in the COP/OlaHsd substrain is listed there. Directly beneath this section you will see that QTL data for the parental strain and both substrains is listed in the Strain QTL data segment. The QTLs listed for the COP/Crl substrain here are the same as those listed on that substrain's strain report.
13. Finally, scroll down to the "Associated References" section. This section lists all of the references associated with the COP strain and its substrains. Click on the last link in the list for COP/Crl, "7: Strecker TE et al..." to view the same Reference Report that you accessed from the COP/Crl report page.

Reference/Author Search and Reference Reports

1. From the RGD home page (<http://rgd.mcw.edu>), or any internal page, click on the Data tab at the top of the page. The submenu in the blue bar lists the various types of biological data stored in RGD (A). Click on “References” to open the Reference query page.

The screenshot shows the RGD Reference Search results for the strain Rattus norvegicus. A search term 'blood pressure rapp 1998' is entered in the Keyword search box (1). The results show 2 records found (2). The first record (4) is a reference to a paper by Rapp JP. The second record (10) is another reference to the same paper. Red arrows (3) point to the 'Refine Term' and 'Sort By' dropdown menus. Red boxes highlight the search term in the Keyword box (1), the record numbers (4, 10), and the abstract text (4).

2. Multiple search terms can be entered into the strain keyword search box. So for instance, enter “blood pressure rapp 1998” and the tool will automatically search for articles in RGD that meet all of the search criteria. Click “Search References” to locate the relevant references.

- On the results page, note once again that the search can be refined or the results sorted using the options in the box at the top of the form. The search results can be exported as a CSV or tab-delimited text file or sent to a printer. Results include the RGD ID, title and citation of the reference. The full abstract can be viewed directly on the search results page by clicking "...(more)". The result in this list also gives one-click access to the abstract in PubMed via the link in the "PubMed" column.
- Click the link for the article by Rapp et al ("Linkage analysis and construction...") to view the abstract for that article.

Reference Report

Linkage analysis and construction of a congenic strain for a blood pressure QTL on rat chromosome 9.

5 Rapp JP, Garrett MR, Dene H, Meng H, Hoebee B, Lathrop GM.

Citation: Rapp JP, et al., Genomics 1998 Jul 15;51(2):191-6.

Status: COMPLETED

Abstract (2)

A blood pressure quantitative trait locus was found (LOD = 5.0) on rat chromosome 9 derived from Dahl salt-sensitive (S) and Dahl salt-resistant (R) rats. The F2 strain intermediate between the R (low blood pressure) and S (high blood pressure) designated S/R(chr 9), had a lower blood pressure (19 mm Hg) than S rats (2% NaCl diet for 24 days), proving the existence of a blood pressure QTL on rat chromosome 9.

Show data curated from this reference **7**

External Database Links

6 PubMed Rapp JP, et al., Genomics 1998 Jul 15;51(2):191-6 (9722941)

Related Objects Report

There are 7 entries in RGD covered by this reference **8**

GENES	Symbol Name
Inha	inhibin alpha

STRAINS	Symbol Name
SR/Jr	Salt Resistant
SS/Jr	Salt Sensitive
SS-SR-Inha/Jr	-na-

QTLs	Symbol Name
Bp34	Blood pressure QTL 34

SSLPS	Symbol Name
D9Ua6	-na-
D9Ua9	-na-

- On the Reference report page, note that each author's name is a link. Click a name to view all of the articles in RGD by that author.
- The reference report page also includes a link to the abstract record in PubMed. Click the citation under "External Database Links" to view the abstract at PubMed.
- Finally, on the reference report page click the link labeled "Show data..." to view all of the data objects that have been curated from this reference.
- The Related Objects report shows that 7 objects have annotations derived from the Rapp et al reference. These include one gene, three strains, one QTL and two SSLPs. Click on the symbol for any of these objects to view the report page for that object.

Pathway Annotations and Visualization

- Similar to a search for GO, a search using the keyword 'leptin' within the Pathway Ontology brings up "leptin system pathway" and "altered leptin system pathway" – the two terms in the vocabulary that contain the word "leptin".

1 Term: leptin system pathway
altered leptin system pathway

Ontology	Rat	Mouse	Human	Tree
Pathway	10	10	10	
Pathway	3	1	2	

ONTOLOGY REPORT

2 There are 2 path(s) and 58 annotation(s) to this term. Annotations to this term and its descendants total 67.

3 **leptin system pathway (PW:0000363)**

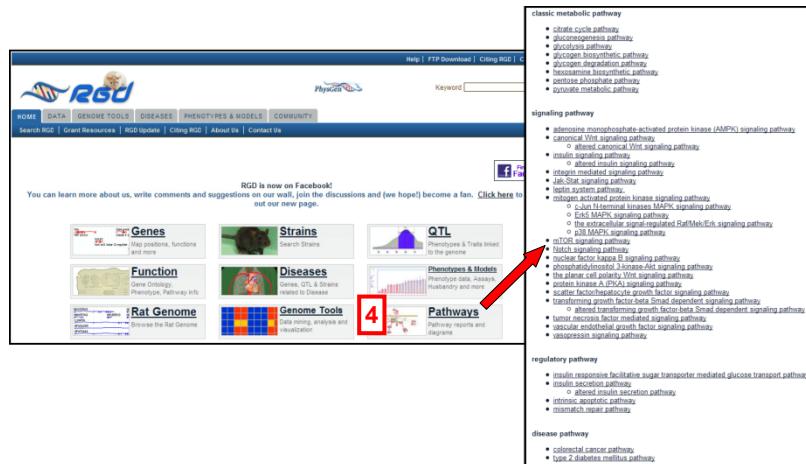
Leptin acts on two populations of hypothalamic neurons that express orexigenic (feeding inducing) and anorexigenic genes, respectively. Leptin reduces the expression of genes from the former and enhances the expression of genes from the latter neurons. Leptin mediated response acts primarily via Jak-dependent activation of STAT.

View Interactive Pathway Diagram

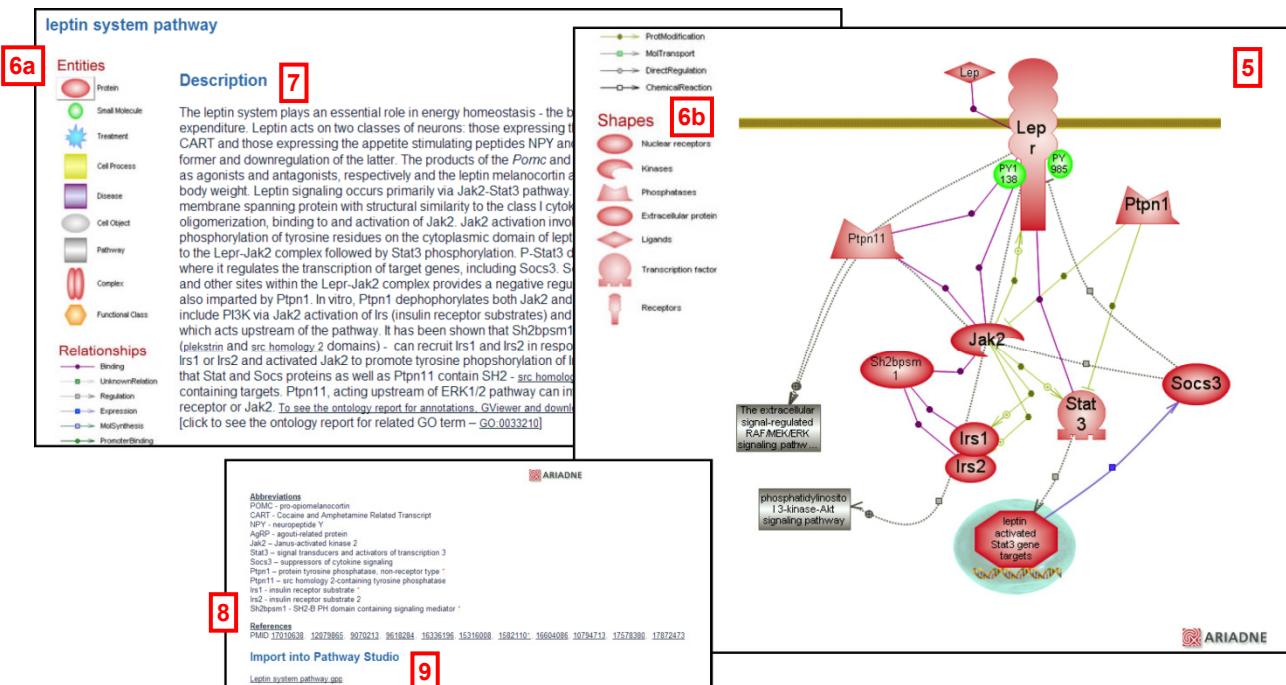
10
List All Objects | CSV Export | Add Objects | Clear

Gene Symbol	Gene Name	Evidence	Reference
IRS1	insulin receptor substrate 1	ISS (from RGD-10816)	RGD-2302071
IRS2	insulin receptor substrate 2	ISS (from RGD-10816)	RGD-2302071
Isr1	insulin receptor substrate 1	IDA	RGD-2302071
Isr2	insulin receptor substrate 2	IDA	RGD-2302071

- Click on the term “leptin system pathway” to view all the objects annotated to that term. The ontology report shows a GViewer image of the rat genes annotated to that term, as well as a list of rat, human and mouse genes annotated to the term and the paths to the term. Unlike GO annotations which are done on a gene-by-gene basis, annotations to pathway terms aim at capturing the network information, i.e., the identification of all genes which according to published reviews are components of a particular pathway and the subsequent annotation of these genes to the term.
- In addition to the annotation of genes to pathway terms, RGD builds interactive pathway diagrams using the Pathway Studio 5.0 from Ariadne Genomics. A growing list of classic metabolic, signaling, regulatory and disease pathway diagrams have already been incorporated into RGD. These diagrams are accessible from the Ontology report page by clicking on the pathway diagram icon at the top right side of the page.



- Alternatively, the entire list of pathway diagrams can be viewed by clicking on the “Pathways” button on the RGD home page. From the list, click on the name of a pathway to explore its diagram.



- The leptin system pathway diagram is shown here. Icons for individual genes link to the RGD records for those genes. Those for related pathways link to the diagrams for the respective pathways. Where a single icon represents multiple objects, such as target genes, a link is provided to a separate page with the complete list to simplify the diagram structure.

6. Along with the diagram itself, all diagram pages provide a legend for entities, relationships and shapes as designed by Ariadne Genomics.
7. Note the detailed description of the pathway
8. A list of abbreviations and a list of references that are hyperlinked to PubMed are provided.
9. In addition to viewing the diagram, users can import the diagram.gpp file into their own Pathway Studio tool. The Ariadne Genomics logo is displayed on all diagram pages.

RGD Tools

Disease Portals

RGD is providing disease and disease related information for rat, human and mouse genes, rat and human QTL and rat strains. In an effort to offer better ways for organizing and presenting these data, RGD has begun implementing the Disease Portals project. Portals so far offered are for the cardiovascular, neurological and obesity/metabolic syndrome disorders.

RGD Disease Portals

RGD disease portals are designed to be entry points for researchers to access data and tools related to their area of interest.

Cardiovascular Disease Portal

RGD has released its Cardiovascular Disease Portal to provide researchers with easy access to data on genes, QTLs, strain models, biological processes and pathways related to cardiovascular diseases. This resource funded by NIBIB also includes dynamic data analysis tools to make it a one stop resource for cardiovascular researchers. The user chooses a disease category to get a pull-down list of diseases. A single click on a disease will provide a list of related genes, QTLs, and strains as well as a genome wide view of these across the genome via GViewer and access to GBrowse results showing the genes and QTLs within the genomic context. Additional pages for Biological Processes, Pathways and Phenotypes provide one-click access to data of interest. A Tools section and a Links section provide additional resources.

Neurological Disease Portal

RGD has released its Neurological Disease Portal to provide researchers with easy access to data on genes, QTLs, strain models, biological processes and pathways related to neurological diseases. This resource, partially funded by NINDS also includes dynamic data analysis tools to make it a one stop resource for neuroscience researchers. The user chooses a disease category to get a pull-down list of diseases. A single click on a disease will provide a list of related genes, QTLs, and strains as well as a genome wide view of these across the genome via GViewer and access to GBrowse results showing the genes and QTLs within the genomic context. Additional pages for Biological Processes, Pathways and Phenotypes provide one-click access to data of interest. A Tools section and a Links section provide additional resources.

Obesity / Metabolic Syndrome Portal

RGD has released its Obesity / Metabolic Syndrome Portal to provide researchers with easy access to data on genes, QTLs, strain models, biological processes and pathways related to obesity diseases. The user chooses a disease category to get a pull-down list of diseases. A single click on a disease will provide a list of related genes, QTLs, and strains as well as a genome wide view of these across the genome via GViewer and access to GBrowse results showing the genes and QTLs within the genomic context. Additional pages for Biological Processes, Pathways and Phenotypes provide one-click access to data of interest. A Tools section and a Links section provide additional resources.

1. From the main RGD webpage click on the **DISEASES** tab to access a list of available portals
2. Select “Obesity / Metabolic Syndrome Portal” either in the menu bar at the top or in the portals list.

Obesity/Metabolic Syndrome Portal

Search RGD | disease | Go

1. Choose a disease category | 2. Choose a disease | OR | Show all

3. Summary

	Rat	Human	Mouse
Genes	713	687	672
QTLs	527	109	
Strains	128		

4. Rat Synteny

5. GViewer for rat chromosomes (highlighted with a red box)

6. Human Synteny

7. Gene Info, QTL Info, Strain Info tables

3. The default view for all portals is a display of all objects (genes, QTL and strains) annotated to disease terms within the portal. Note that a summary of the number of objects of each species is included at the top of the page.
4. A rat GViewer image shows a genome-wide view of annotated genes and QTLs. Click on “Human Synteny” to view an image with the human syntenic regions denoted by color.
5. Mouse over one of the objects in the diagram to see the symbol and position information for that object.

6. Click “View annotation Data” to download a list of the annotated objects with their genomic positions.
7. Lists of annotated objects are located beneath the GViewer image. You can scroll up and down the list. Also, each object is hyperlinked to its report page, i.e., clicking on it one gets to the gene, QTL or strain report page.

The screenshot shows the RGD home page with the 'Pathways' tab highlighted (red box 8). Below it, a dropdown menu is open with 'signaling pathway' selected (red box 9). The main content area displays a list of signaling pathways for Rat, including 'A1 protein signaling pathway' and 'MAPK/STAT signaling pathway'. At the bottom, a GViewer image shows chromosomes 1 through Y with various genomic features. A legend on the left identifies these features: Genes (black), RNAs (grey), QTLs (blue), and Strains (green).

8. In addition to this default display, you can choose a disease category and a specific disease within the category or make such selections for phenotypes, biological processes or pathways within the portal to see the annotated objects. All annotations are based on controlled vocabularies: the Disease Ontology (DO), the Mammalian Phenotype Ontology (MP), GO Process Ontology (one of the three vocabularies of GO), and the Pathway Ontology (PW). Click on the “Pathway link” to access this section of the portal.
9. In the drop down lists at the top of the page, select “signaling pathway”. Notice that the items in the second list change according to what you choose for the first term. Also note that the lists of objects and the GViewer image change according to the ontology and the terms selected.

BioMart

1. RGD's BioMart tool allows the user to query RGD for specific data based on both input and output criteria. Once a dataset is selected, that data is filtered based on the input criteria and the data retrieved is determined by the output criteria. From the RGD home page (<http://rgd.mcw.edu>), or any internal page, click the **TOOLS** tab at the top of the page and select **BioMart** on the menu bar to access the MCWMart home page.

The screenshot shows the MCWMart home page. The 'TOOLS' tab is highlighted (red box 1) and the 'BioMart' link under the 'MCWMart' heading is also highlighted (red box 2). At the bottom right, a large red box highlights the 'Start MartView session' button (red box 3).

2. Notice the **RGD Cookbook** links in the left margin. These pages walk you through some common problems.
3. To access the RGD BioMart tool, click **Start MartView session**.

The screenshot shows the RGD BioMart interface. A red box labeled '4' highlights the 'Dataset' dropdown menu, which currently displays '[None selected]'. To its right is another dropdown menu labeled 'CHOOSE DATABASE' with several options listed: 'RGD Mart', 'IPI Rat Mart', 'SSLP Mart', and 'UniProt Prototype (EBI)'.

4. There are several datasets to choose from. Select **RGD Mart**.

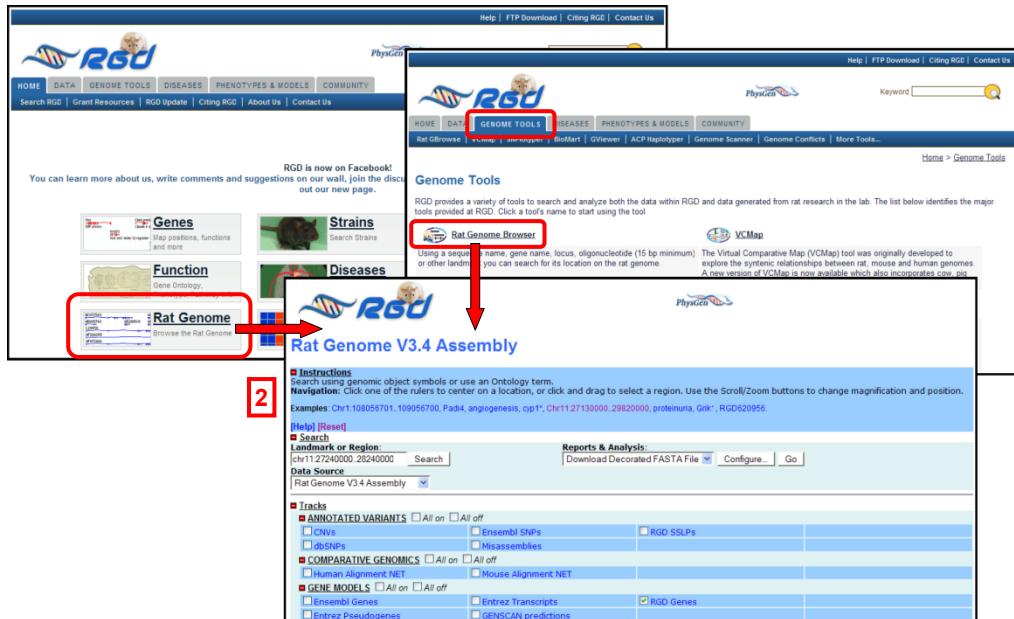
5. Click Filters to access a form which will allow you to limit the genes for which you get information. A subset of the RGD genes can be selected based on a single symbol and/or name, map position, a list or file containing gene, sequence or ontology identifiers, and/or a variety of ontology terms.
6. For this example, click the plus sign (+) next to “External Database Identifiers”, click the check box for “Limit to Genes with” and type “2004” into the text box under “RGD ID(s)”. If you wanted to upload a file of IDs you would click the “Browse...” button to locate and upload your file.

7. Select “Attributes” in the menu on the left. Choose the type of data that you wish to obtain for the genes you have selected, either “GENE AND FUNCTION” data or “DATABASE ACCESSIONS”.
8. For this example, keep the default selection of gene and function. This includes gene IDs and nomenclature, map data and annotations for any or all of the four ontologies that RGD uses—Gene, Disease, Phenotype and Pathway.
9. Click the check boxes for Symbol, RGD ID, Entrez Gene ID, Chromosome and Start and Stop positions.

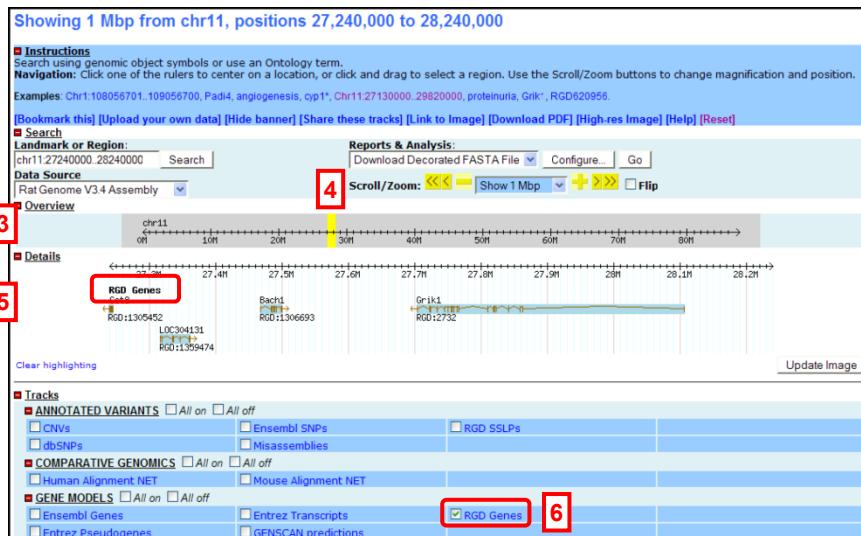
10. Click the Results button at the top of the page to view the selected data for RGD ID 2004.
11. Results can be viewed on the website or downloaded to a file as HTML, Comma or Tab separated Values or Excel spreadsheet. For both the download function and viewing on the website, you can choose “Unique results only” to avoid redundant annotations.

GBrowse

1. You can access the RGD Rat Genome Browser, or GBrowse, from the RGD home page (<http://rgd.mcw.edu>) by clicking on the button labeled “Rat Genome”, or from any RGD page, by clicking on the “GENOME TOOLS” tab at the top of the page. On the Genome Tools page, click the Rat Genome Browser link to access GBrowse.

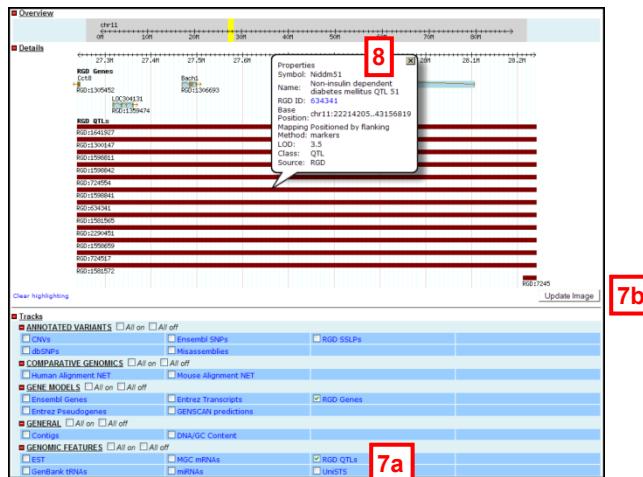


2. Notice at the top of the page there is a section with instructions for using GBrowse including examples of possible searches. Below this section, type “chr11:27240000..28240000” into the “Landmark or Region” box under Search. Note that you can also enter symbols or keywords into this box. Directly under this, the default for “Data Source” is Rat Genome V3.4 Assembly. Keep this as the source, but notice that other data sources are available including the Human genome. Click the Search button to display the region entered.

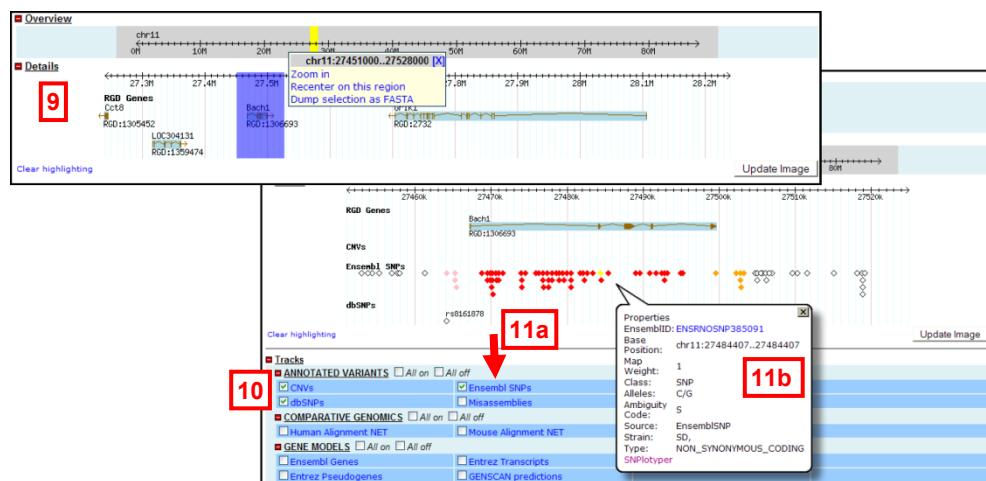


3. The Overview section shows a diagram of the entire chromosome 11 with a yellow box/line showing the section which is displayed in the Details section. Clicking on the overview diagram will move the Details display to the location where you clicked. The number of base pairs displayed remains the same. Use your browser’s “Back” button to return to the original position.
4. You can also use the arrows to scroll up or down the chromosome, and the dropdown box or plus/minus buttons to increase or decrease the number of base pairs displayed. Use your browser’s “Back” button to return to the original size and/or position.
5. In the Details section, if you have used GBrowse before the tool will use your previous settings. To clear these settings, use the “Reset” button at the top of the page. If you have not used the tool before or if you have cleared the settings, the default view is only the RGD Genes track.

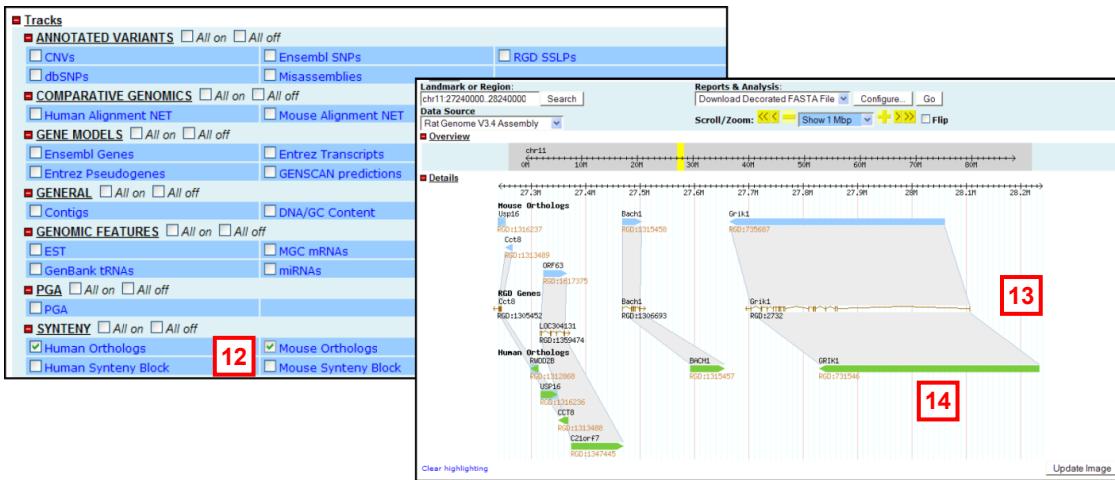
6. If you scroll down to the “Tracks” section of the page you can see the variety of different data types that GBrowse can display. Because only RGD genes are presently being displayed, only that box is checked.



7. To see all of the QTL that overlap your region, scroll down to the GENOMIC FEATURES segment of the Tracks section. Check the box for RGD QTLs and click the “Update Image” button.
 8. Note that when you mouse over any object in GBrowse, such as a QTL or gene, a balloon pops up with detailed information about that object. What information is displayed depends on the object. In this case, the QTL nomenclature, RGD ID, base pair position, mapping method and LOD score are shown. Notice that the RGD ID is a link that will take you to the corresponding QTL report.



9. RGD's GBrowse has tracks for a number of annotated variants including SSLPs, copy number variants and SNPs from both dbSNP and Ensembl. To see the SNPs for the gene Bach1, click and drag your mouse along the scale line in the Details section to select the portion that contains just that gene. A popup box gives you the option to Zoom in, Recenter the image or download the selection as a FASTA file. Select “Zoom in”. The image is automatically updated.
 10. Scrolling down to the Tracks section, uncheck the box for RGD QTLs, check the one for Ensembl SNPs and click “Update Image”.
 11. The diamond shapes that represent SNPs in GBrowse are color coded to show the type of SNP each is. For a list of the colors and their meaning, click the text (not the check box) for “Ensembl SNPs”. A new window will open up which gives more information about the various tracks. Mousing over a SNP will open a popup balloon that displays more information about that SNP, including what type of SNP it is. If you mouse over the yellow SNP in the second exon of the Bach1 gene you will see that it is a “Non-synonymous coding” SNP.



12. RGD's rat GBrowse now includes information on rat-mouse-human synteny. Using your browser's "Back" button, go back to the original image of the rat genes within chr11:27240000..28240000, or alternatively, click the "Reset" button and re-enter the genomic position of interest. Notice that the rat genes that are displayed have blue shading behind them. This indicates that those genes have mouse and or human orthologous genes that are currently not being displayed. Scrolling down to the bottom of the Tracks section, select the tracks for "Human Orthologs" and "Mouse Orthologs" and click "Update Image".
13. The display now shows a number of orthologous mouse and human genes with grey shaded areas linking the rat genes to their orthologs. Hover over a mouse gene to open a popup balloon that displays more information about that gene including its symbol, name, RGD, MGI and Entrez Gene IDs, and its genomic position in mouse. The RGD, MGI and Entrez Gene IDs are links to that gene's record on the corresponding website.
14. Note that the sizes of the genes in the different species have not been adjusted. If the rat gene is longer or shorter than the orthologous mouse or human gene, this is reflected in the sizes of the bars representing those genes.

FTP Downloads

1. Data can be downloaded in bulk from our ftp site. Click "Download Data via FTP" above the Keyword search box in the top right corner of each RGD page.

File	Date	Size	Description
NAR-00425-web-S-2005.R1	4/20/2005	12	
RGD genome annotations	10/7/2004	12	
data release	1/29/2008	3	
hpd	1/25/2005	12	
mirror	9/14/2004	12	
publications	1/10/2003	12	
rhmep	7/6/2005	12	
test	1 KB	10/27/2006	12
workshops	10/13/2006	12	

File	Date	Size	Description
4a	Up to higher level directory	12/9/2002	12:00:00 AM
EVS		3812 KB	7/19/2002 12:00:00 AM
ESTS		1024 KB	7/19/2002 12:00:00 AM
ESTS_NCBI		11980 KB	5/3/2007 12:00:00 AM
GENES_ARCHIVE		1681 KB	2/26/2008 3:09:00 AM
GENES_HUMAN		6800 KB	5/3/2007 12:00:00 AM
GENES_HOG		14361 KB	2/26/2008 3:09:00 AM
GENES_HUS		14705 KB	4/23/2007 12:00:00 AM
GENES_MOUSE		11071 KB	2/26/2008 3:09:00 AM
GENES_NAT		615 KB	6/21/2005 12:00:00 AM
GENE_SFT_ID_2_PUBLISHED_ID		887 KB	3/24/2007 12:00:00 AM
HISTORY		159 KB	2/28/2008 10:13:00 PM
HS_FW.xls		4775 KB	2/28/2008 10:13:00 PM
MAP_DATA		32 KB	hsome_genes_do.obo
MAP_FW.xls		167 KB	hsome_genes_sp.obo
PR20070321.obo		188 KB	hsome_genes_pv.obo
PR20071012.obo		188 KB	hsome_qtlis_do.obo
PR20071109.obo		28 KB	hsome_qtlis_sp.obo
PRM_Introduction.doc		1 KB	hsome_genes_do.obo
QTLs		2 KB	hsome_genes_pv.obo
README.txt		201 KB	tattus_genes_sp.obo
REF_ID_2_PUBLISHED_ID		218 KB	tattus_genes_pv.obo
Rn_FW.xls		3472 KB	tattus_qtlis_sp.obo
Strains		549 KB	tattus_qtlis_pv.obo
Strains		360 KB	tattus_qtlis_pv.obo
Strains		302 KB	tattus_qtlis_sp.obo
Strains		170 KB	tattus_strains_sp.obo
Strains		49 KB	tattus_strains_pv.obo
Strains		124 KB	tattus_strains_pv.obo

2. Another link to "FTP Download" appears in the menu bar at the very top of each RGD page.
3. On the FTP site, select the "data release" directory.
4. The data release directory contains a variety of files containing data compiled from the RGD database for rat, mouse and human genes, rat and human QTLs and rat strains. In addition, a subdirectory holds individual files with annotations of various objects to particular ontologies.