Rat Genome Database

an introduction

Simon N. Twigger, Ph.D.
Bioinformatics Research Center
Medical College of Wisconsin, Milwaukee
Background

- The Rat is a very good model of Human complex genetic diseases (eg. Hypertension, cancer, diabetes, etc)
- Rat is a highly developed physiological model system
- Ultimate goal of Rat research is to translate results from Rat to Human

RGD Background:
- One of 8 NIH funded Model Organism Databases (MOD)
- Established in 1999 (grown from 4 people to 20+)
- Goal - to facilitate rat research, access to data, in particular mapped phenotypes (QTLs).
Various rat strains are used as experimental models of Human disease.

Rat studies identify regions associated with disease in Rat.

Comparative Genomics allows translation to Mouse and Human
The Rat Genome Database (RGD) curates and integrates rat genetic and genomic data and provides access to this data to support research using the rat as a genetic model for the study of human disease.

Related links: About RGD, Article describing RGD

RGD Highlight
The Rat Genome Database Requests Your Input

The Rat Genome Database requests your participation in a web survey to help improve our website. Survey topics include: current tool and data use, desired tools and data, advanced searching needs and general user information. It will take approximately 5-10 minutes to complete. Click to start.

Strain registration form now available

You can register your strains in RGD by using the online strain registration form. RGD will check the name of the strain for its compliance with the rules for strain nomenclature. If you need help please contact strains@rgd.mcw.edu

What's New
- RGD releases The Pied Piper Vol 2.
- RGD strain registration form now are available.
- RGD finished the loading of rat updated gene information ... more.
- Click to view the list of new and updated gene reports with links to associated RefSeq mRNA sequences in RGD.
- New rat genome sequence annotation released.
- RGD completed nomenclature updates for many genes, click here to view.
- Accessing the sequence data are now online.

Events around the RGD ...
Anatomy of the site

<table>
<thead>
<tr>
<th>SEARCH</th>
<th>LIST OF MAJOR OBJECTS</th>
<th>HELP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PAGE CONTENT: TOOL, WEB PAGE, etc.</td>
<td></td>
</tr>
</tbody>
</table>
Anatomy of the site

There are two options to view the RGD strain records:

Option 1. Search by keyword or Option 2. By direct selection from a list of strains in RGD:

**Option 1. Search strain records by Keyword**

Enter Keyword

Search Fields

**Option 2. View specific strain record**

I3M
A28607
A35322
A7322
A990

Note: Selecting a 'primary' strain such as BN (no substring designation) will show the data for BN and all associated substrains of BN in RGD (BN/Cub-x, BN/SsNHsd, etc). Selecting a strain with a specific substring designation will show you the data for that strain alone. This is to facilitate comparison of strains and substrain information.
Lots of Help

RGD Help Pages

The help pages are divided up as shown in the diagram opposite. The "How Do I...?" section provides a guide to using the data and tools in RGD to perform common tasks. The Help Directory below lists more general help areas for the database contents, tools, and includes a FAQ list and glossary of terms.

Can't find the help you need?

- Fill in the RGD User Request form to get email help, or
- Call us on 414-456-8871 (10-Noon, 2-4pm CST), ask for RGD User Support.

Help Directory - information organized by subject

- **How Do I...?**
  - case studies using RGD

- Searching RGD
  - Quick searches, targeted searches...

- Glossary of Terms

- **What's in RGD**
  - Genes, Strains, ESTs, QTLs, Maps...

- Bioinformatic Tools
  - VMap, Genome Scanner, RH Mapserver...

- Frequently Asked Questions
  - Scientific, Database, Nomenclature...
Data is from multiple sources

Regular Journal Screening And Curation

Databases
Websites
Informatic data mining

Data Sources

Internal Data
External Data

RGD Database

Data Pipeline

Ongoing Data Curation and QA/QC

Data is from multiple sources:
- Databases
- Websites
- Informatic data mining

 RGD Database

Ongoing Data Curation and QA/QC
Data and Tools

Genes
QTLs
SSLPs
Maps
Strains
Sequences

Current Tools:
- Genome Scanner - Genetic Mapping/Genome Scans
- RHMap Server - Radiation Hybrid Mapping
- VCMap - Comparative Mapping
- MetaGene - Gene Prediction

Coming Soon:
- Ontology Browser and searches
- ACP Haplotyper
- Blast/Blat

PGA Workshop August 2003
RGD: References

Reference Report

Objects linked to reference

Search by Author.
RGD: Genes

Gene report provides key information about a gene: Name and aliases, curated homologs, functional annotations, mapping data, external database links and curated references.

External Database Links

- GenBank Nucleotide: M15890, M15793, NM_012614
- GenBank Protein: 205761, 205757, 205738
- Homologene: Npy
- LocusLink: Npy
- MGD: Npy
- RatMap: Npy
- UniGene: Rn.9714

Search Genbank for sequences associated with Npy

References

RGD Tools

Helping researchers access and use the data within RGD
Major Tools

- **Genome Scanner**
  Polymorphic marker selection

- **VCMap**
  Comparative mapping between Rat, Mouse and Human

- **MetaGene**
  Gene prediction and visualization tool

- **RH Mapserver**
  Map RH vectors to the genome *in silico*
Genome Scanner

Polymorphic marker identification between two strains
Easily select marker set for genome-wide scans for disease association

Cross: DA/Pit x F344/Pit
Available Maps: SHRSP x BN genetic map
Bin Size: 10 cM (for Genetic maps, cR for RH maps)
Chromosome: All
Highlight allele size differences between: 4 bp and 20 bp
Output Content: Chr. Stats and Table of Polymorphic Markers
Output Format: Web Page (HTML)

PGA Workshop August 2003
### Genome Scanner results

<table>
<thead>
<tr>
<th>Marker</th>
<th>Distance</th>
<th>LOD</th>
<th>DA/Pit Size</th>
<th>F344/Pit Size</th>
<th>Difference</th>
<th>Add to Runlist</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chromosome 4 (Bin 1, 0 - 10) (top)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D4Rat142</td>
<td>0</td>
<td>3.1</td>
<td>220</td>
<td>238</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>D4Rat3</td>
<td>0</td>
<td>3.1</td>
<td>158</td>
<td>148</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>D4Rat2</td>
<td>0</td>
<td>3.1</td>
<td>145</td>
<td>149</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>D4Rat143</td>
<td>0</td>
<td>3.1</td>
<td>124</td>
<td>126</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>D4Rat117</td>
<td>0</td>
<td>3.1</td>
<td>122</td>
<td>120</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>D4Arb14</td>
<td>0</td>
<td>3.1</td>
<td>213</td>
<td>207</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>D4Rat145</td>
<td>0</td>
<td>3.1</td>
<td>156</td>
<td>172</td>
<td>16</td>
<td>checked</td>
</tr>
<tr>
<td>D4Rat144</td>
<td>0</td>
<td>3.1</td>
<td>246</td>
<td>242</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>D4Arb13</td>
<td>.08</td>
<td>3.1</td>
<td>134</td>
<td>138</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>D4Rat147</td>
<td>.08</td>
<td>0</td>
<td>159</td>
<td>151</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>D4Rat149</td>
<td>6.7499</td>
<td>1.1</td>
<td>188</td>
<td>198</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>D4Rat221</td>
<td>7.05</td>
<td>1.2</td>
<td>152</td>
<td>146</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>D4Rat126</td>
<td>7.55</td>
<td>1.3</td>
<td>163</td>
<td>157</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>D4Rat150</td>
<td>7.92</td>
<td>2.8</td>
<td>193</td>
<td>195</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>D4Rat222</td>
<td>9.11</td>
<td>2.4</td>
<td>225</td>
<td>223</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Chromosome 4 (Bin 2, 10 - 20)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D4Rat126</td>
<td>11.3899</td>
<td>3.1</td>
<td>142</td>
<td>136</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>D4Rat11</td>
<td>18.1199</td>
<td>1.1</td>
<td>138</td>
<td>124</td>
<td>14</td>
<td>checked</td>
</tr>
<tr>
<td>D4Rat89</td>
<td>18.2</td>
<td>1.1</td>
<td>180</td>
<td>184</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>D4Rat115</td>
<td>19.87</td>
<td>1.4</td>
<td>234</td>
<td>245</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

Select one marker per 10cM bin.
Virtual Comparative Maps

Virtual Comparative Map (VC Map)

- EST-based comparative mapping algorithm
- Uses Rat, Mouse and Human RH maps
- Allows location predictions based on synteny (Virtual mapping)
- Web/Java maps and reports.
Virtual Comparative Maps
VCMap: Virtual Mapping

Within a defined syntenic region the location of homologous markers, previously unmapped in Rat, can be predicted.
Community Resources

- Rat Community Forum - online bulletin board
- Pied Piper - quarterly newsletter
- Rat Strain and Gene nomenclature guides
- User support - for RGD tools and general rat genomics/bioinformatics questions
- Visiting Scientist program - work with RGD to enhance the resources in a specialized area
# Acknowledgements

<table>
<thead>
<tr>
<th>RGD Admin</th>
<th>Data Mining and Advanced Tool Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>Howard Jacob</td>
<td>Norie de la Cruz</td>
</tr>
<tr>
<td>Simon Twigger</td>
<td>Jed Mathis</td>
</tr>
<tr>
<td>Peter Tonellato</td>
<td>Aubrey Hughes</td>
</tr>
<tr>
<td></td>
<td>Wenhua Wu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Database and Tool Management</th>
<th>Genome Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dean Pasko</td>
<td>Jeff Nie</td>
</tr>
<tr>
<td>Jiali Chen</td>
<td>Victor Ruotti</td>
</tr>
<tr>
<td>Chunyu Fan</td>
<td></td>
</tr>
<tr>
<td>Lan Zhao</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Integration and Comparative Analysis</th>
<th>System Admin &amp; WebMaster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Susan Bromberg</td>
<td>Yuan Ji</td>
</tr>
<tr>
<td>Chin-Fu Chen</td>
<td>Weihong Jin</td>
</tr>
<tr>
<td>Rajni Nigam</td>
<td></td>
</tr>
<tr>
<td>Gopal Gopinathrao</td>
<td></td>
</tr>
<tr>
<td>Angela Zuniga-Meyer</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Exploration and Discovery</th>
<th>Mouse Genome Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary Shimoyama</td>
<td>Janan Eppig</td>
</tr>
<tr>
<td>Nataliya Nenasheva</td>
<td>Lois Maltais</td>
</tr>
<tr>
<td>Weiye Wang</td>
<td></td>
</tr>
</tbody>
</table>

| | NCBI |
| | Greg Schuler |
| | Donna Maglott |
| | Svetlana Dracheva |

PGA Workshop August 2003