RGD: Data and Tools for Translational Research in Otitis Media

Jennifer R. Smith

Scientific Project Manager Rat Genome Database Medical College of Wisconsin, Milwaukee, WI, USA

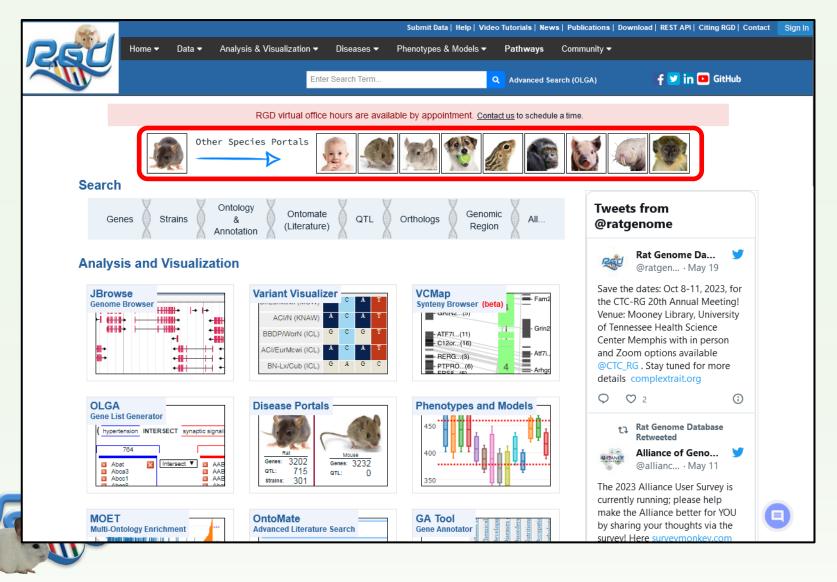
ISOM 22ND International Symposium on Recent Advances in Otitis Media



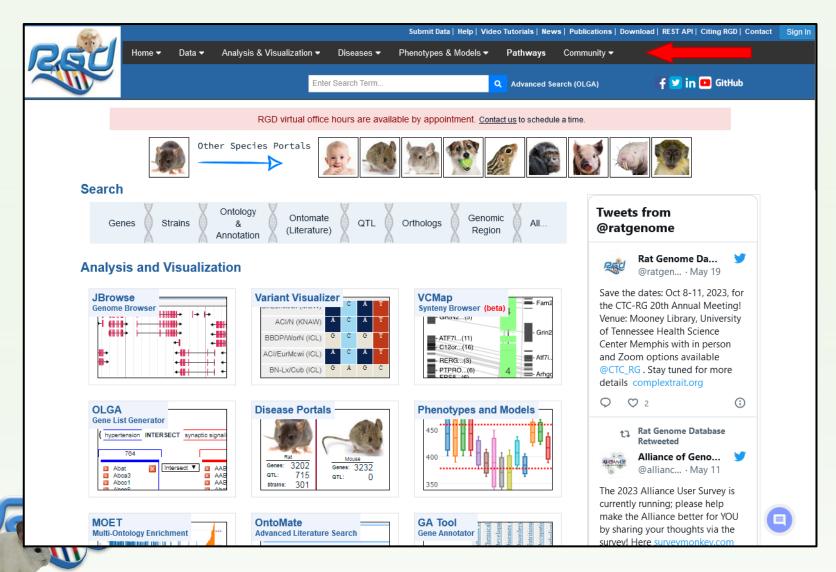






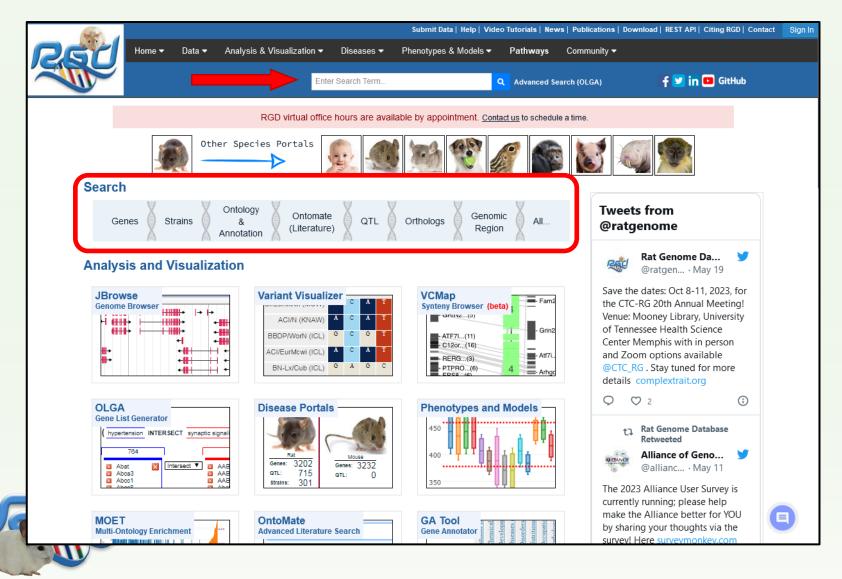


- Houses data for ten mammalian species including human and chinchilla.
- Integrates data derived from expert literature curation with data imported from other sources including other databases and direct submissions from researchers.
- Provides an extensive and well-rounded corpus of genetic, genomic, disease and physiological data.
- Provides innovative tools for data analysis.



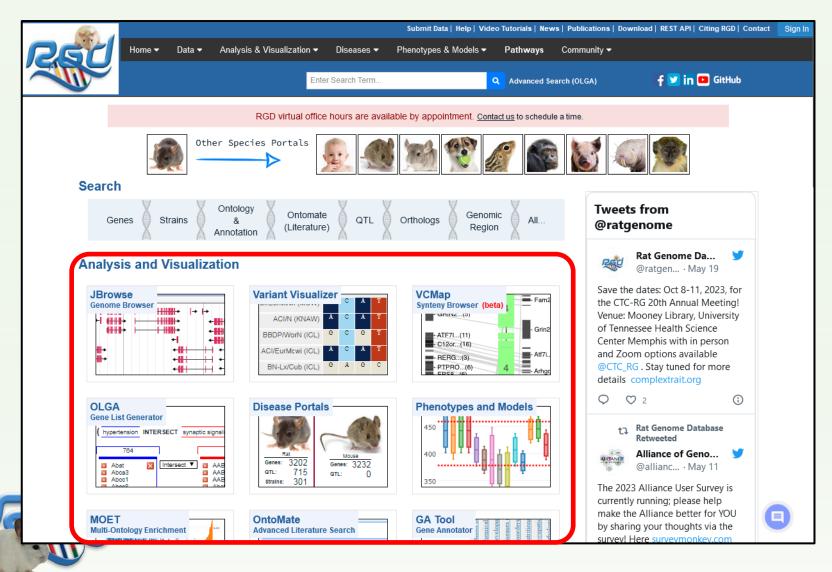
 Data and tools are accessible using the menus at the top of the page.





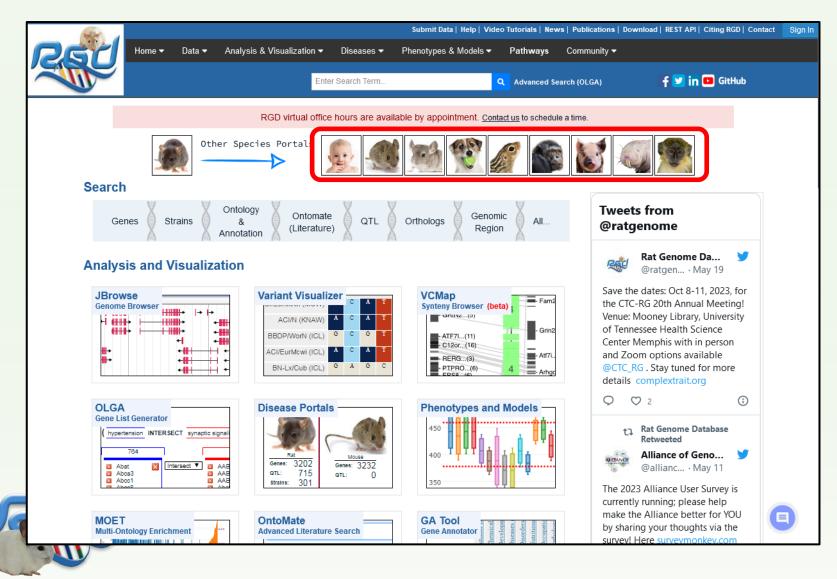
- Data and tools are accessible using the menus at the top of the page
- Data can also be accessed using the general search box or targeted searches.





- Data and tools are accessible using the menus at the top of the page
- Data can also be accessed using the general search box or targeted searches.
- Tools can also be accessed using the Analysis and Visualization links on the front page.





- Data and tools are accessible using the menus at the top of the page
- Data can also be accessed using the general search box or targeted searches.
- Tools can also be accessed using the Analysis and Visualization links on the front page.
- Species icons link to informational landing pages for each species.





Human

Homo sapiens

Welcome to the Human Portal within RGD. This portal contains links to data related to Homo sapiens, and to tools that clinical researchers can use to find and analyze that data. Please feel free to contact us with suggestions for additional data or tools that would help advance your research.



"Human diseases and conditions are routinely studied directly in human subjects, via controlled case studies and regulated clinical trials. However, pre-clinical studies are also vitally necessary. Animal models are employed that represent "a living organism in which normative biology or behavior can be studied, or in which a spontaneous or induced pathological process can be investigated, and in which the phenomenon in one or more respects resembles the same phenomenon in humans." Researchers need to be able to determine and utilize the most precise models for their disease of interest; that is, models that closely mimic the human phenotypic profile for, or aspects of a disease or condition. Sometimes despite promising results with preclinical treatments in one animal model, the same treatments do not always translate to successful human clinical trials. Thus, increasingly researchers want informal nabout more than one model organism, and research benefits greatly from being able to explore integrated data. Having both human and model organism data accessible in one place will help researchers find the data they need for both clinical and translational research. Whether a researcher is looking for models of a disease, or human variations, or is working with a particular species and is looking for corroboration in other species, the power of combined data can be applied to research within the Rat Genome Database using the tools found on this page."

Sources: PMID:10845564; PMID:31552413; PMID:3155241; NLM Unique ID:101117739; Hau J. (2008) Sourcebook of Models for Biomedical Research; Simmons, D. (2008) Nature Education 1(1):70

Ontomate: Pre-filtered to Human Human QTL Search Human Gene Search Human Genetic Variants Visualize RGD's Disease Annotations RGD's Disease Portals RGD's Pathway Annotations RGD's Pathway Diagrams GO: Molecular Function Annotations GO: Biological Process Annotations GO: Cellular Component Annotations RGD's Human Download Site RGD's Full Download Site Human Analysis Tools Human JBrowse Genome Browse OLGA: Object List Generator & Analyze MOET: Multi Ontology Enrichment Tool GA Tool: Annotation Search and Export GOLF: Gene and Ortholog Location Finder InterViewer - Protein Interactions

As an example, this is the top of the species page for human. The layout is essentially the same for all of the species portals, although depending on the data and tools available for each, there are differences. The next few slides show the sections of the chinchilla portal.

Human Data

Ontomate: Concept-Based Literature Search

Ontomate: pre-filtered for Human literature



Human QTLs at RGD

Human Genes at RGD

Human Genetic Variants Visualizer



RGD Disease Annotations



Human Genome Resources

RGD Video Tutorials





Ontomate: Mair

Ontomate: pre-filtered for Chinchilla

Chinchilla Otitis Media Annotations
RGD's Pathway Annotations

GO: Molecular Function Annotations

GO: Biological Process Annotations

GO: Cellular Component Annotations

Chinchilla JBrowse Genome Browser

OLGA: Object List Generator & Analyzer

MOET: Multi Ontology Enrichment Tool

GA Tool: Annotation Search and Export

GOLF: Gene and Ortholog Location Finder

RGD's Chinchilla Download Site

RGD's Full Download Site

hinchilla Analysis Tools

Chinchilla Genes at RGD

RGD Disease Annotations

RGD's Pathway Diagrams

Long-tailed Chinchilla

Chinchilla lanigera

Welcome to the Chinchilla Portal within RGD. This portal contains links to data related to the long-tailed chinchilla, Chinchilla lanigera, and to tools that chinchilla researchers can use to find and analyze that data. Please feel free to contact us with suggestions for additional data or tools that would help advance your research.



"The long-tailed chinchilla (Chinchilla lanigera) is used as a model for understanding the physiology, development and function of the auditory system. The chinchilla is used in several research areas, including ottlis media, upper respiratory tract infections, hearing, psychoacoustics and ofotoxicity. The chinchilla has been used for auditory or acoustics research because of the anatomical and physiological similarities between its inner ear and that of human (1-4). Its middle ear and Eustachian tube structures are also similar to that of humans and unlike other rodent models it is not susceptible to innate middle ear infections, making it an ideal model organism for otitis media (OM) studies (5-7). The presence of a large cephalid bulla, from which sufficient quantities of middle ear fluids can be withdrawn for microbiological and immunological assessment, is also an advantage. Most interestingly, the chinchilla has a permanently semipatulous Eustachian tube, which makes this tubal organ highly similar to the floppy' Eustachian tube of childhood (5). The chinchilla is recognized as an ideal model of otitis media and its use has produced critical information regarding the molecular mechanisms of pathogenesis of pneumococal OM, the immune response to S. pneumoniae, and the immunogenicity and protective efficacy of pneumococcal capsular polysaccharide vaccine anticens (8)."

Source: PMID:27173523

Chinchilla Data

Ontomate: Concept-Based Literature Search

Ontomate: pre-filtered for Chinchilla literature



Chinchilla Genes at RGD



RGD Disease Annotations

Chinchilla Otitis Media Annotations

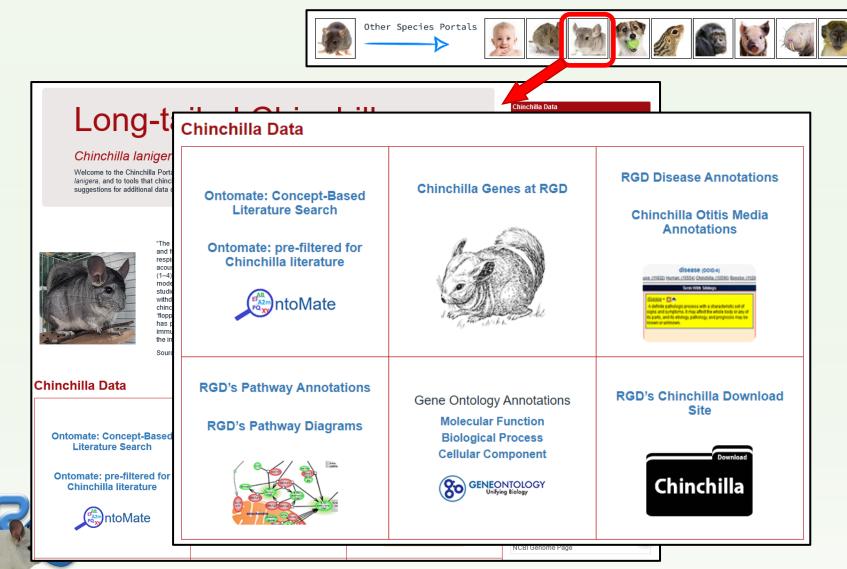
	Te .	m Witch Sil	Angs	
diseasi A petri signs ar ts parts snown o	e pathologic p d symptoms. It and its etiologi r unknown.	ocess with may affect pathology	a charact he whole and prop	Histic set of body or any of nosis may be

chilla Otitis Media

	Chinchilla PhenoMiner
2	
1	Chinchilla Genome Resources
	RGD's Genome Information pa
,	NCBI Genome Page

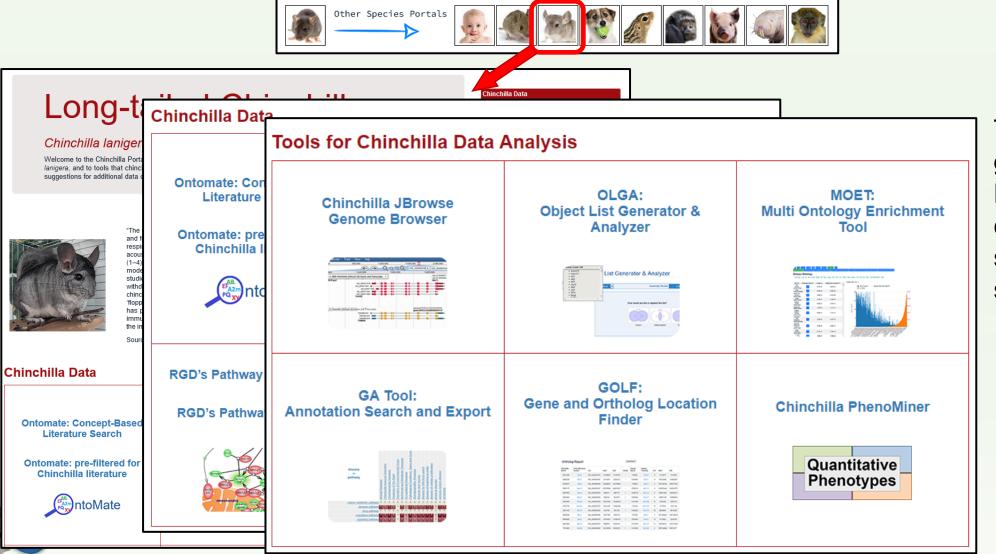
The top of the chinchilla portal gives information about chinchilla as a model of human disease





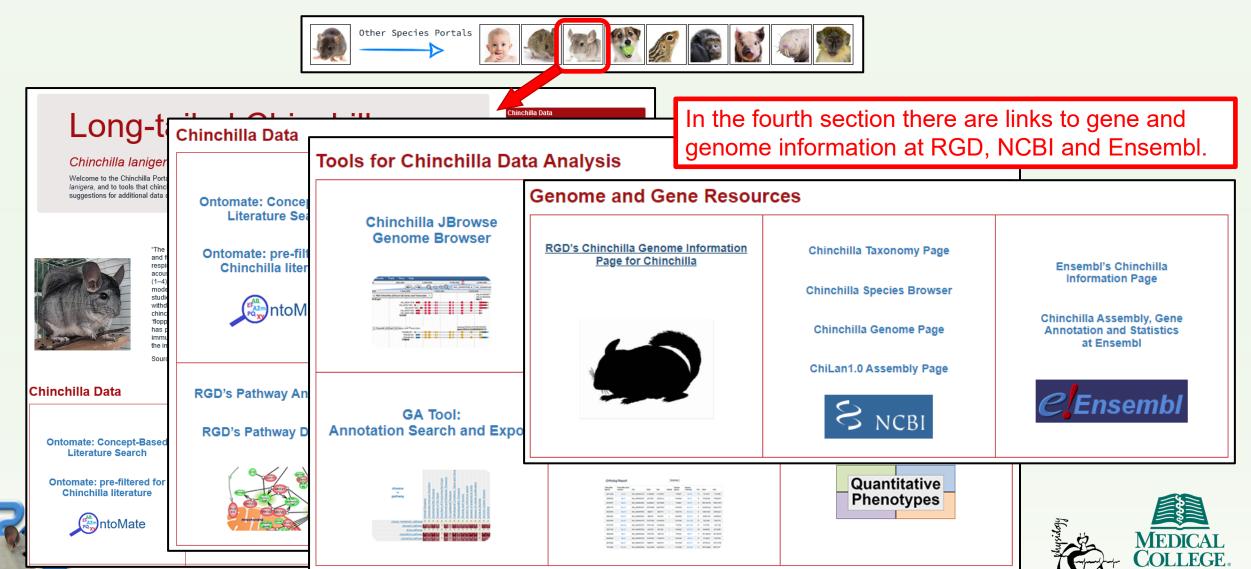
The second section provides links to chinchilla data, including a link to RGD's download site for chinchilla data.



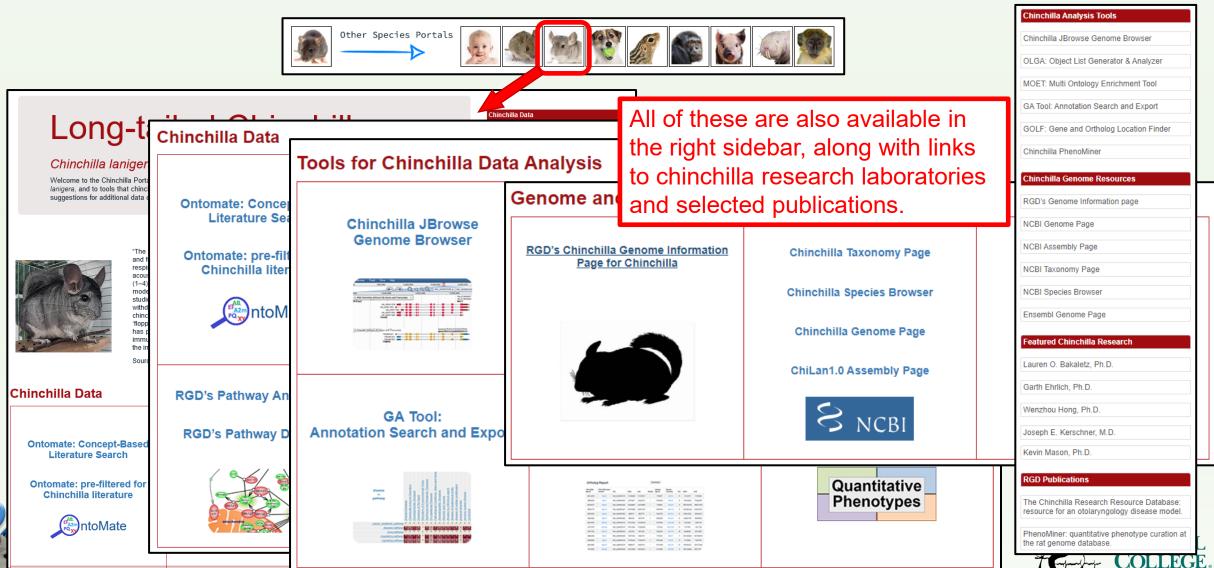


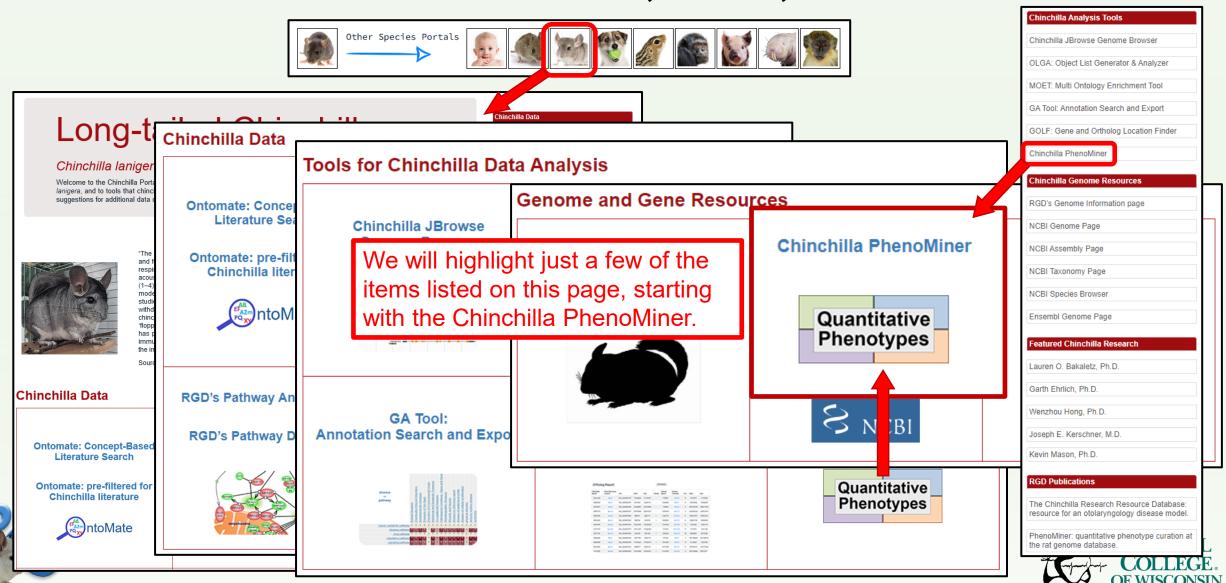
The third section gives links to RGD tools with chinchilla already selected as the species.

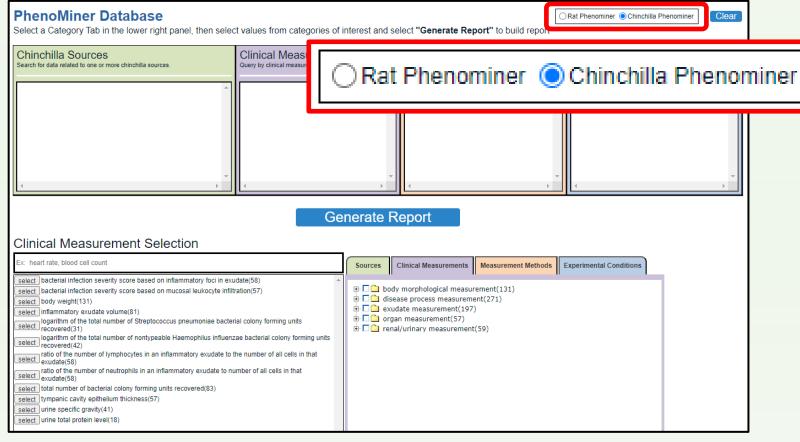








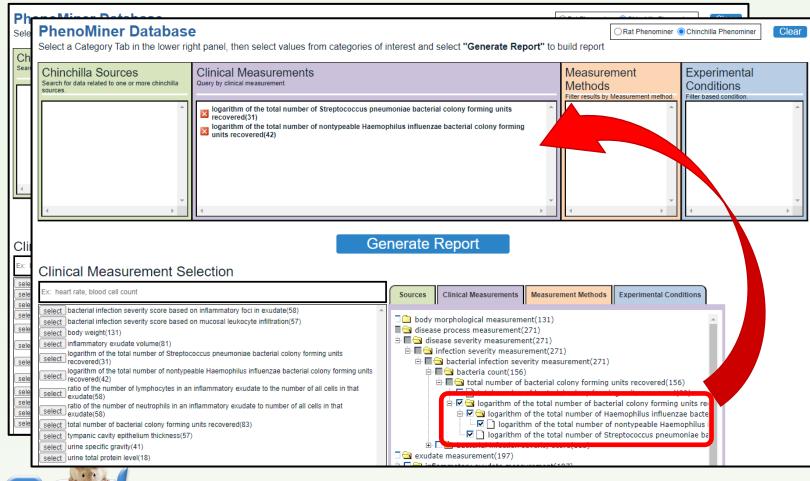




- The data in PhenoMiner is a combination of curated data from the literature and direct submissions from researchers.
- The link from the Chinchilla portal directly accesses the Chinchilla PhenoMiner.



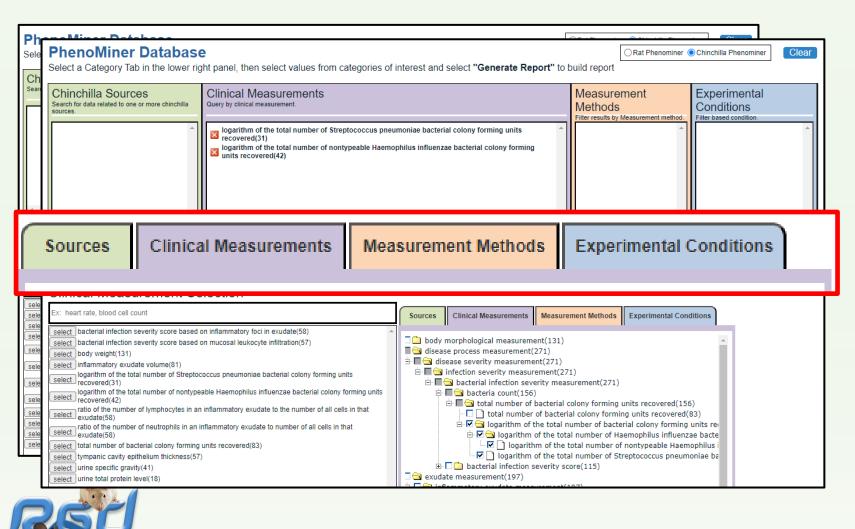




 Selecting terms in the bottom section of the PhenoMiner home page adds them to the corresponding bins in the top section.

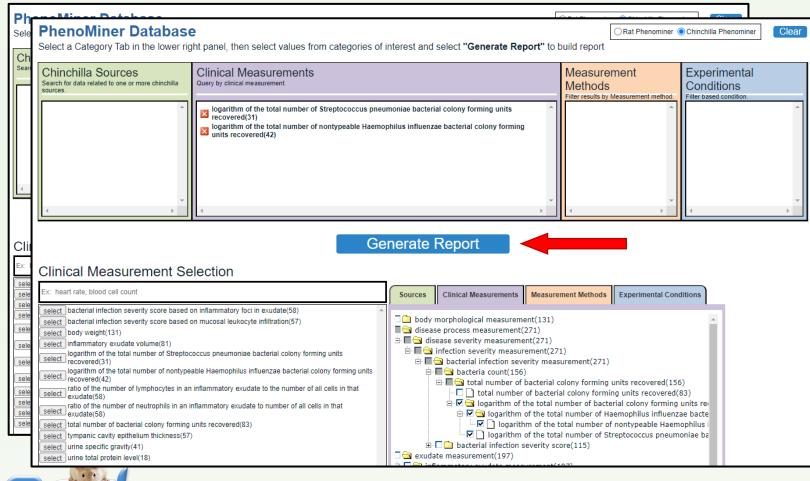






- Data can be chosen on the basis of any or all of four criteria:
 - > Chinchilla Sources
 - Clinical Measurement (what was measured)
 - Measurement Method (how it was measured)
 - Experimental Conditions (under what condition(s) was it measured)

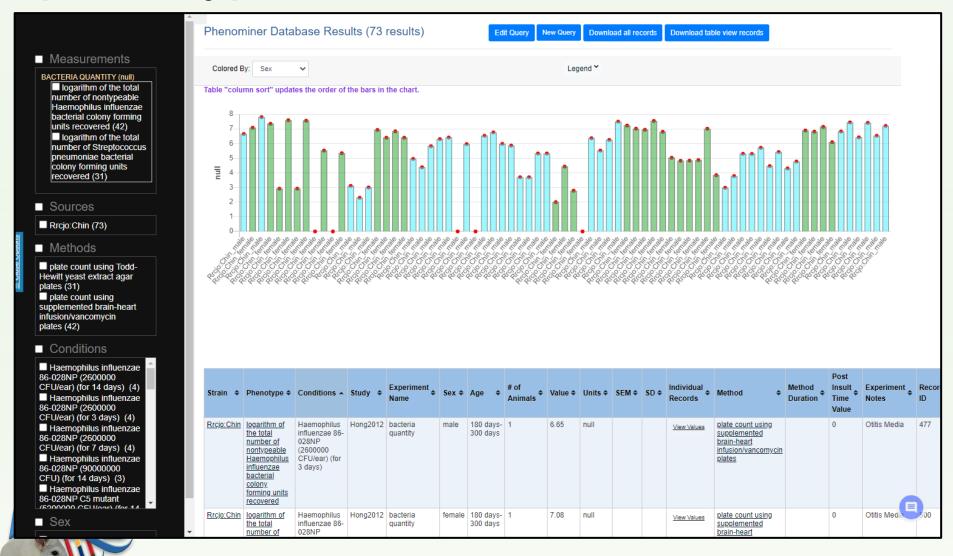




 Once selections have been made, click "Generate Report" to view the associated data.

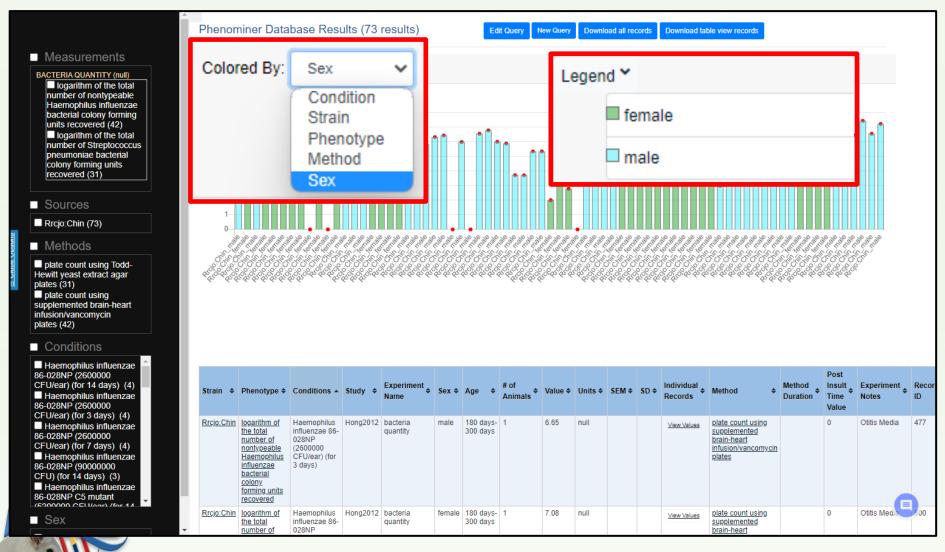






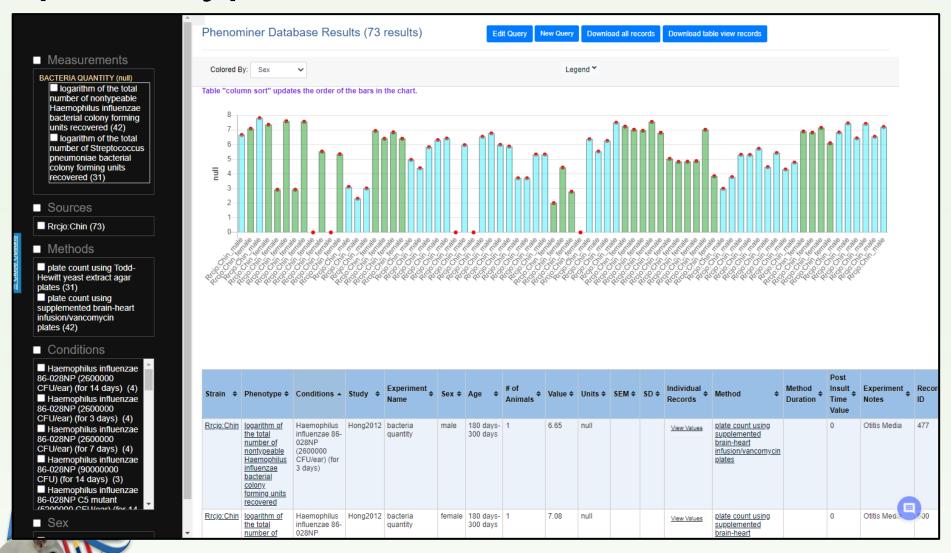
- Results consist of three parts:
 - > A graph
 - > A detailed table
 - Filters in the left sidebar.
- The three sections work together so filters selected on the left remove bars from the graph and rows from the table, and sorting the table reorders bars in the graph.





- The "Colored by"functionality gives options for coloring the bars of the graph.
- The default is to color by condition but any of the options can be selected.
- A dropdown legend is available to indicate what color corresponds to each possible value for the selected experimental variable.





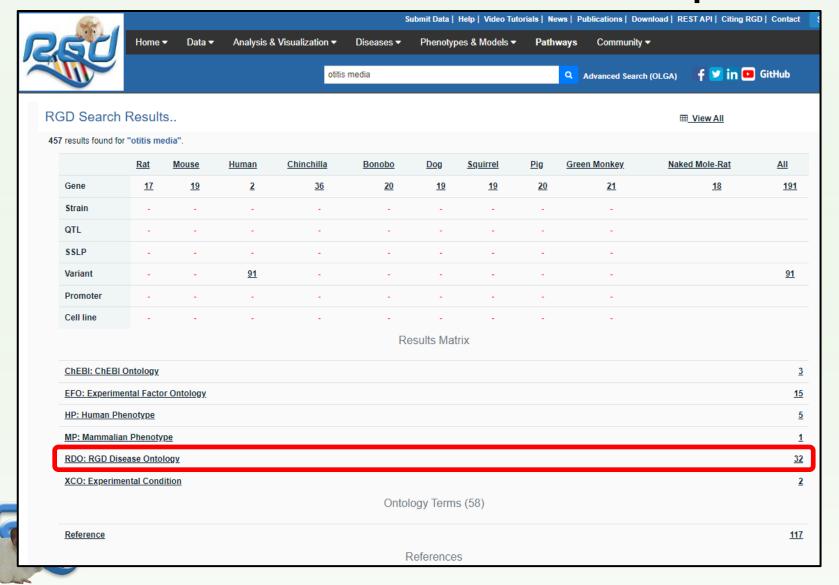
 Here the table is sorted by condition and bars are colored by the sex of the animals.





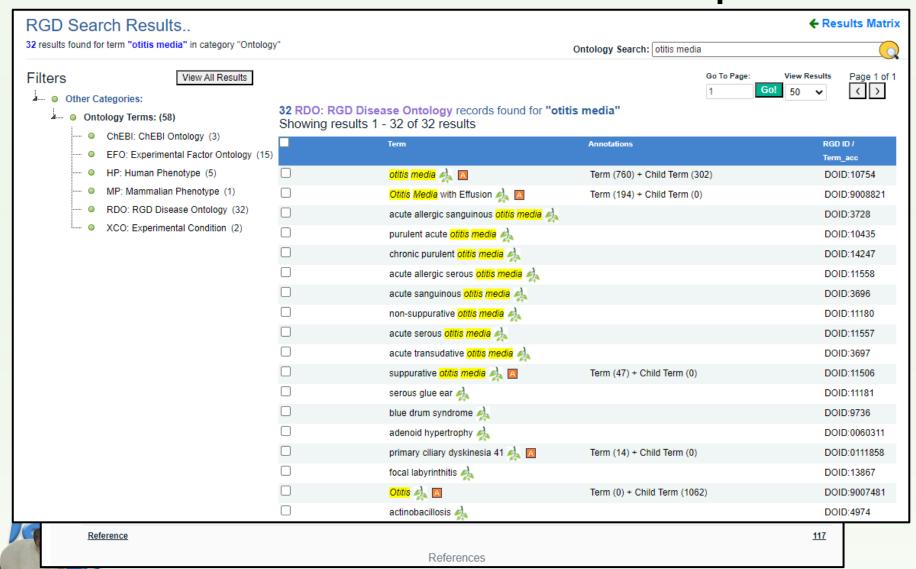
 Options are provided to download all records matching the original query or, if filters have been applied, just the filtered set of records.





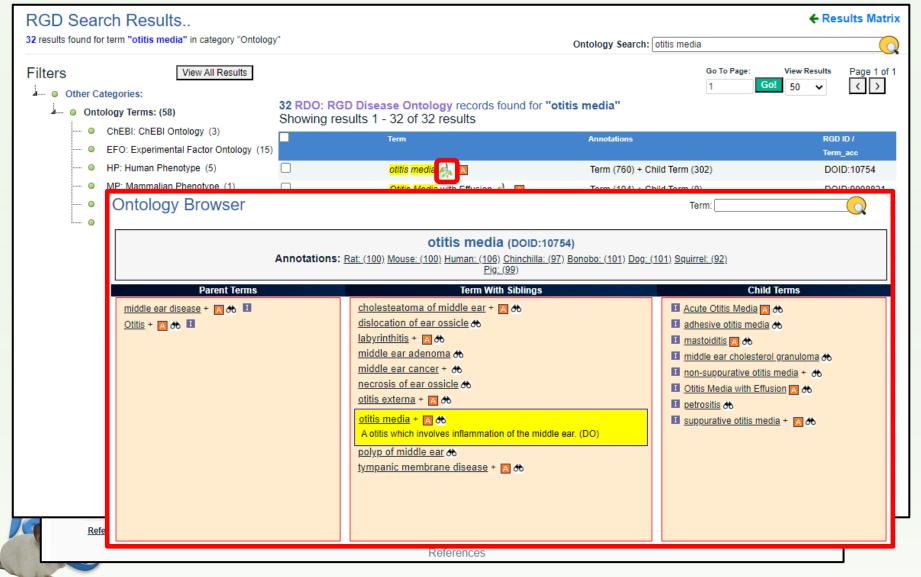
- A search for "otitis media" returns
 - A number of genes for each species
 - Human clinical variants from NCBI's ClinVar database
 - Links to vocabularies that RGD uses for functional annotations, including the Disease Ontology.
- Clicking the line for the Disease Ontology takes you to the list of terms in that vocabulary that match the search term "otitis media".



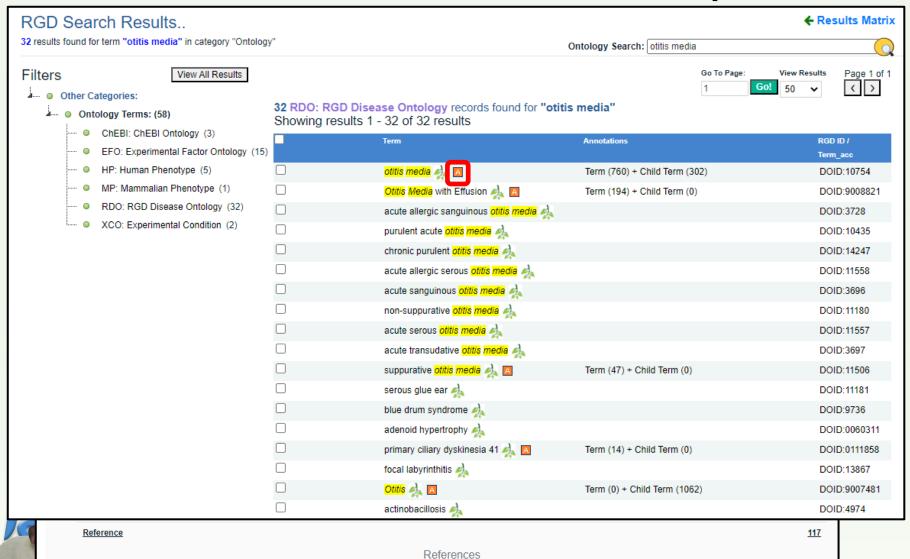


- RGD uses ontologies for data standardization.
- An ontology is a structured vocabulary which is often shown as an upside-down tree structure with fewer more general terms toward the top and larger numbers of more specific terms as you move down the tree.





- RGD uses ontologies for data standardization.
- An ontology is a structured vocabulary which is often shown as an upside-down tree structure with fewer more general terms toward the top and larger numbers of more specific terms as you move down the tree.
- The leaf icon in the search results list takes you to that term in RGD's ontology browser which makes it very easy to navigate up the tree to the more general term and down the tree to the more specific ones.



 The "A" icon links to the ontology report page as seen in the next slide.

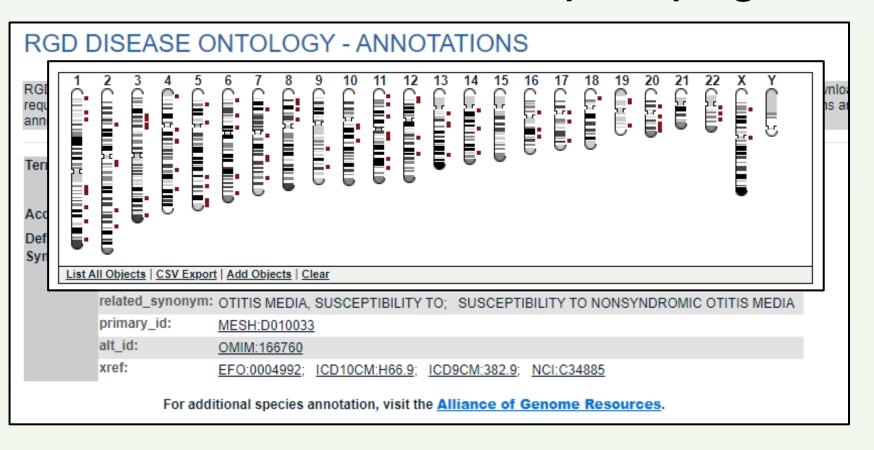


RGD DISEASE ONTOLOGY - ANNOTATIONS RGD uses the Human Disease Ontology (DO, https://disease-ontology.org/) for disease curation across species. RGD automatically download required for RGD's curation purposes but are not currently covered in the official version of DO have been added. As corresponding terms at annotations and subsequently used for curation. Term: go back to main search page otitis media Accession: DOID:10754 browse the term Definition: A otitis which involves inflammation of the middle ear. (DO) Synonyms: exact_synonym: COME/ROM: Middle Ear Inflammation; OMS narrow_synonym: FAMILIAL OTITIS MEDIA; Nonsyndromic otitis media; OTITIS MEDIA, CHRONIC/RECURRENT related_synonym: OTITIS MEDIA, SUSCEPTIBILITY TO; SUSCEPTIBILITY TO NONSYNDROMIC OTITIS MEDIA primary_id: MESH:D010033 alt id: OMIM:166760 xref: EFO:0004992; ICD10CM:H66.9; ICD9CM:382.9; NCI:C34885 For additional species annotation, visit the Alliance of Genome Resources.

 The top of the disease report page gives information about the term, including the definition, synonyms for the term and links to related terms in other vocabularies and information in other databases such as the Online Mendelian Inheritance in Man (OMIM) database.



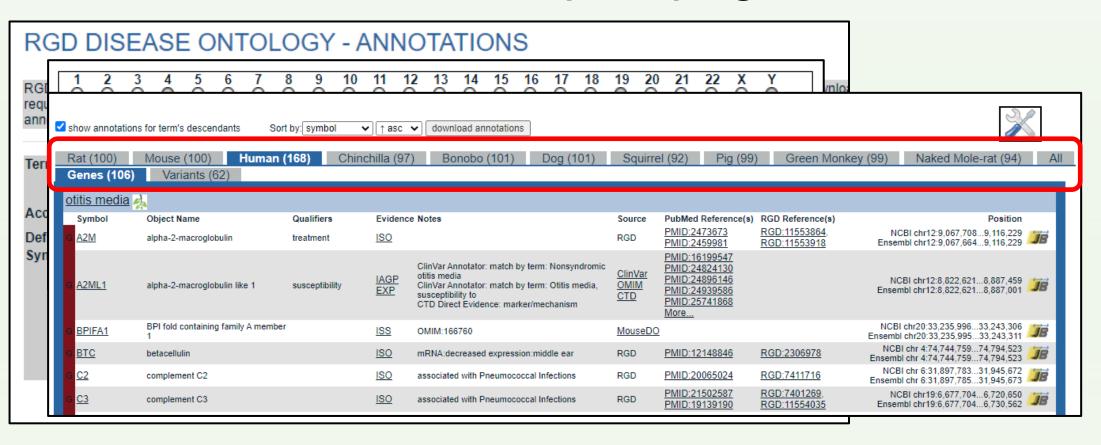




 Where possible a genome-level view of the genes annotated to the term is shown. For rat, the genome level view also shows positions for QTLs and mapped strains annotated to the selected term.



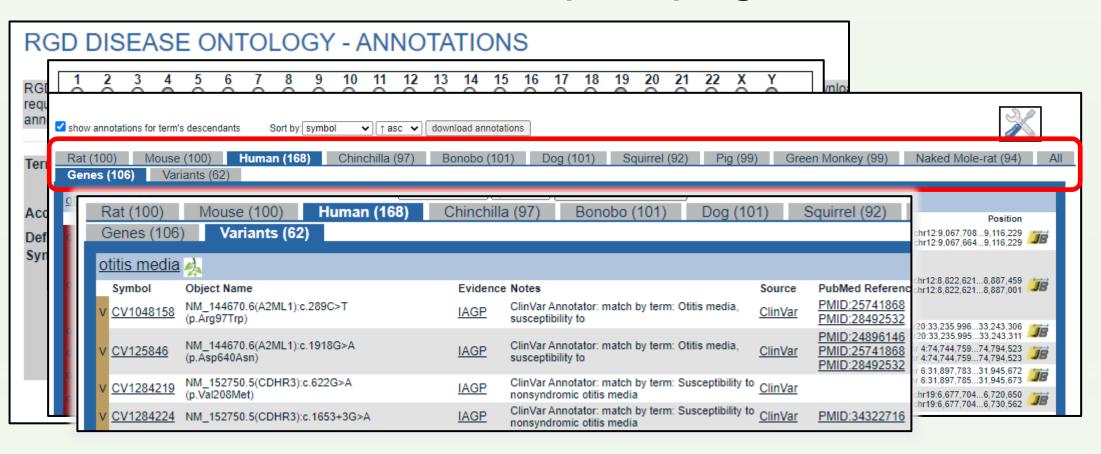




 For human, RGD has annotations to otitis media for both genes and ClinVar variants.



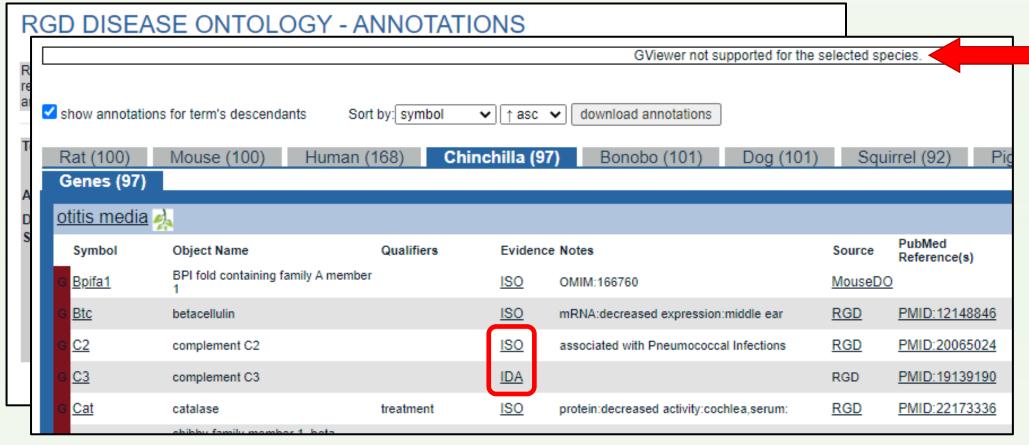




 For human, RGD has annotations to otitis media for both genes and ClinVar variants.





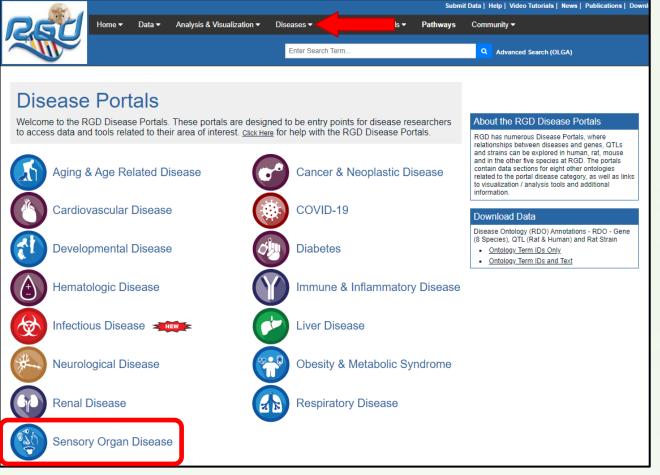


Note that Chinchilla does not have a genome-level view since the genome is only assembled to the point of scaffolds.



 For chinchilla, most gene annotations are derived from orthologous genes in other species ("ISO" evidence code), but RGD curators have reviewed the literature and made experimental annotations directly to chinchilla wherever possible (in this case indicated by the "inferred from direct assay" (IDA) evidence code.

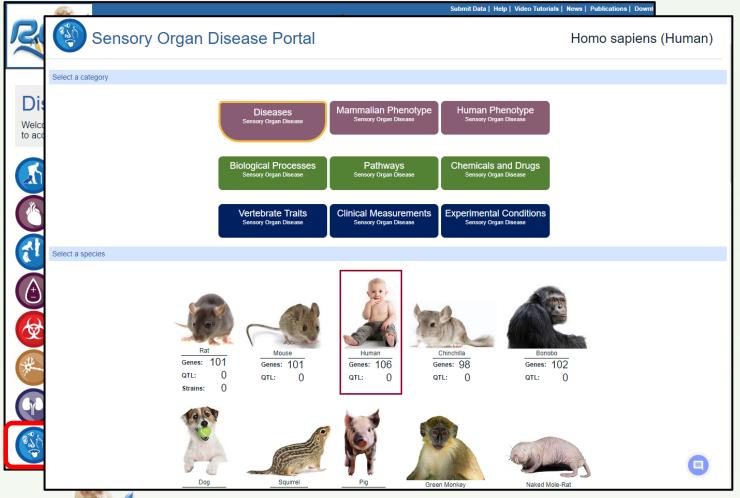




- Disease portals can be accessed via a dropdown list in the top menu bar (arrow), or by clicking the "Diseases" tab in the menu bar to go to the disease portal landing page shown here which provides links to all of the disease portals.
- We will use the Sensory Organ Disease Portal as our example.

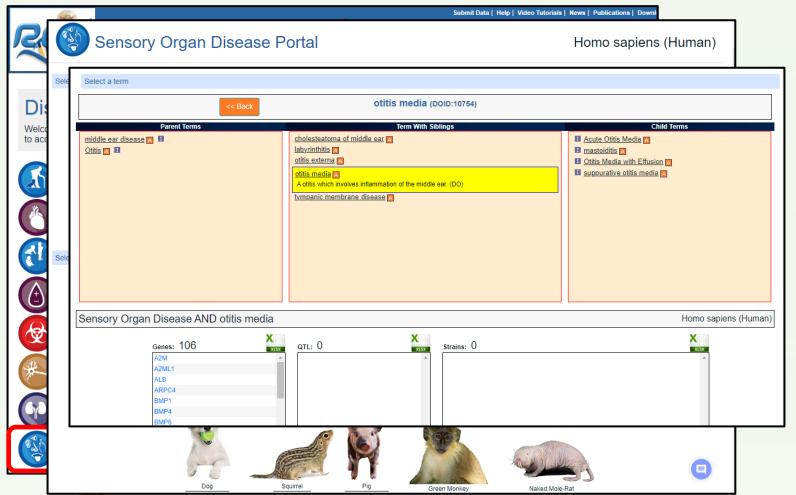






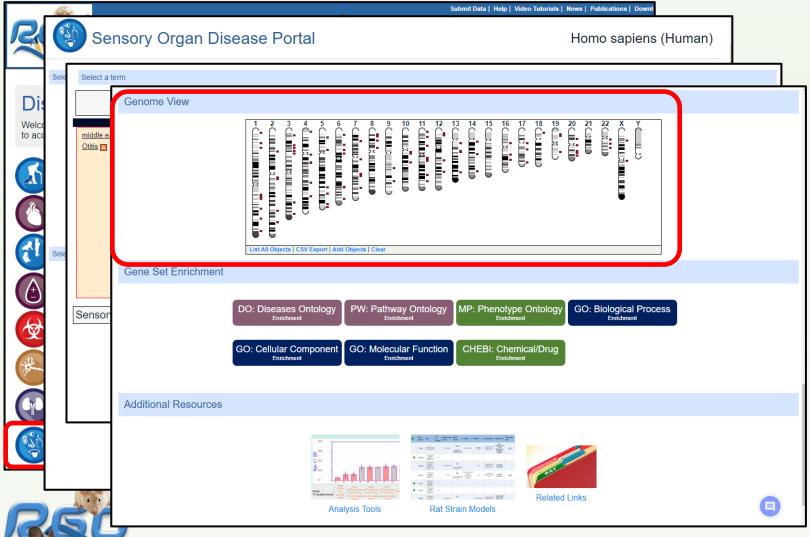
- When accessing a disease portal, first select a data type, in this case, "Diseases", which is the default.
- Select a species, here we have chosen human.





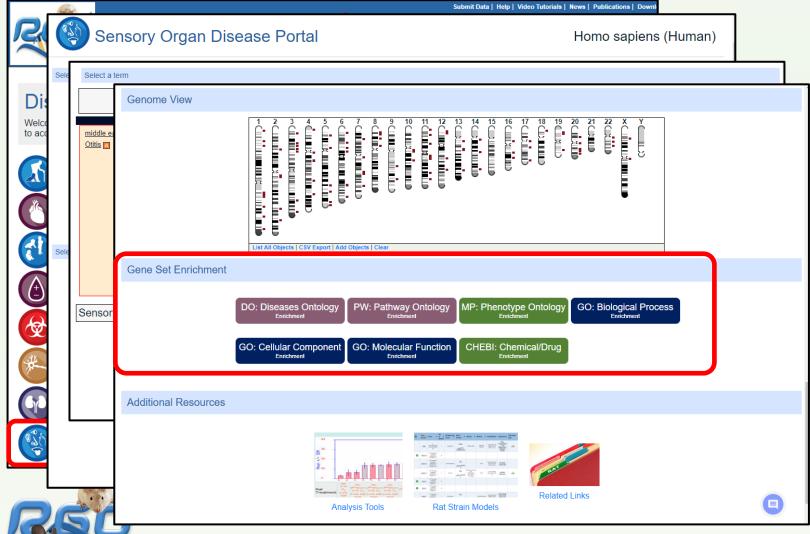
- Disease portals start with the most general term in the category, in this case, "sensory system disease".
- Use the embedded ontology browser to select a more specific term. The portals leverage the structure of the ontology to show all data, in this case, all data annotated to the selected term and it's more specific children.
- Here there are 106 human genes annotated to otitis media or a more specific child term in the list below the browser.
- As selections are made in the browser, the list of genes below and the object counts for each species above are updated on the fly.





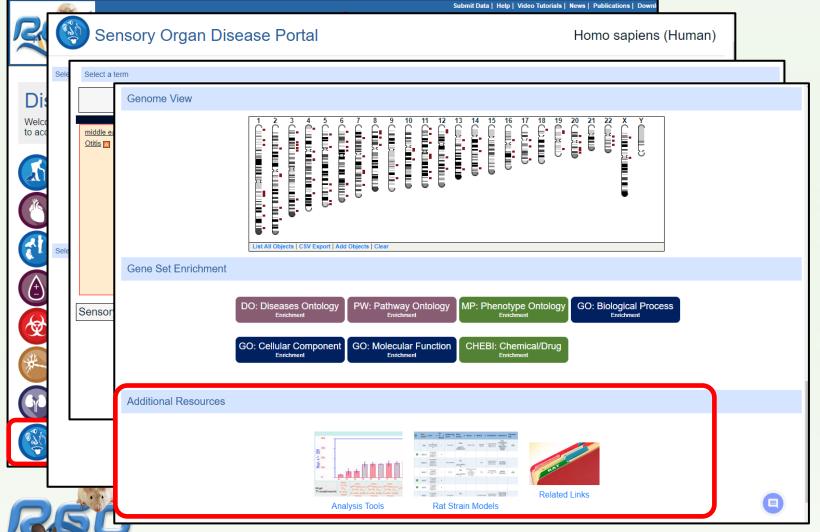
 For human and other species with chromosome-level assemblies, the genome-wide view of gene positions is shown.





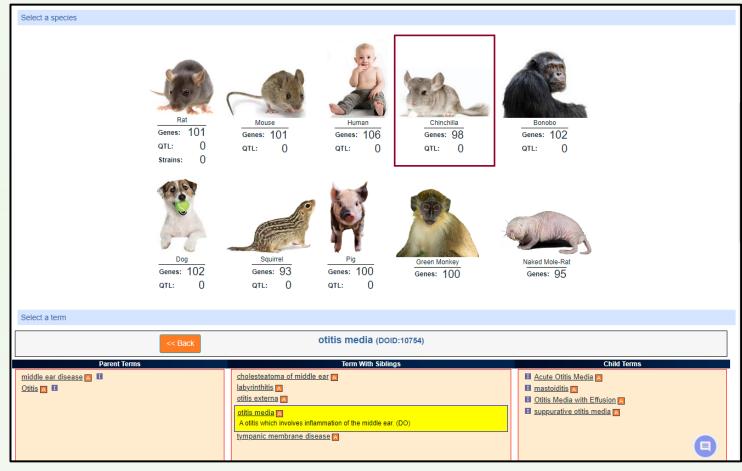
- For human and other species with chromosome-level assemblies, the genome-wide view of gene positions is shown.
- RGD's Multi-Ontology
 Enrichment Tool is embedded into the disease portals. (More about this later.)





- For human and other species with chromosome-level assemblies, the genome-wide view of gene positions is shown.
- RGD's Multi-Ontology
 Enrichment Tool is embedded into the disease portals. (More about this later.)
- Links to additional information are provided at the bottom of the page.
- The same functionality is available for all species.

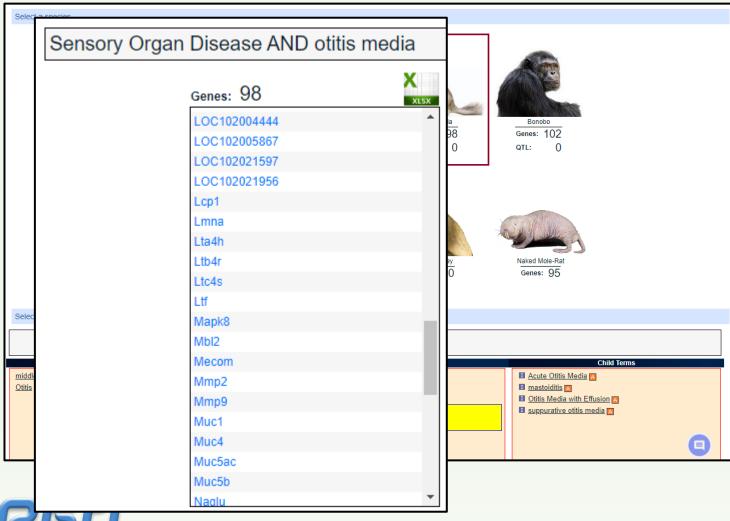




 Going back up to the top of the page, we will select chinchilla as the species.

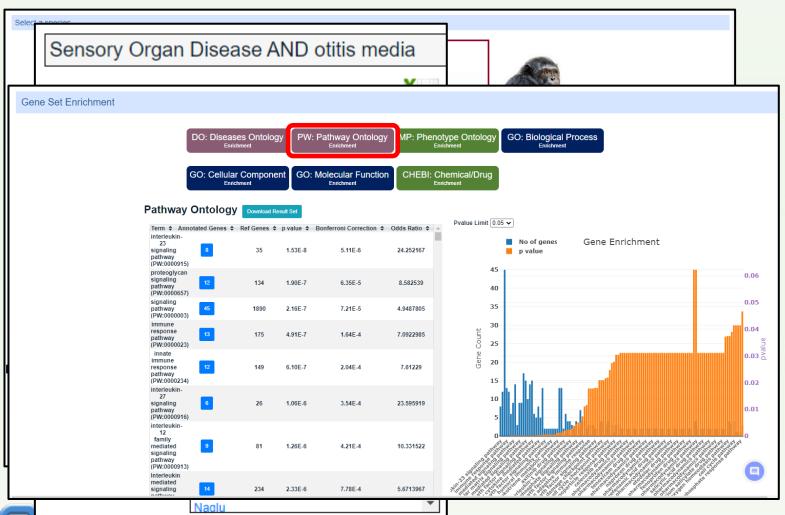






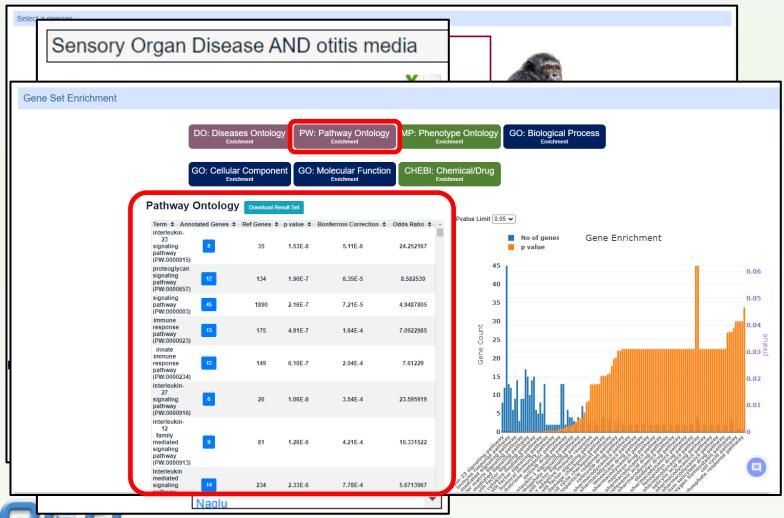
- Going back up to the top of the page, we will select chinchilla as the species.
- When Chinchilla is chosen the list shows the 98 genes in chinchilla annotated to otitis media or any of its more specific children.





- Going back up to the top of the page, we will select chinchilla as the species.
- When Chinchilla is chosen the list shows the 98 genes in chinchilla annotated to otitis media or any of its more specific children.
- When an ontology is selected in the "Gene Set Enrichment" section of the page, the list of genes is automatically submitted to the embedded MOET tool for enrichment analysis.
- Here RGD's Pathway Ontology has been selected.

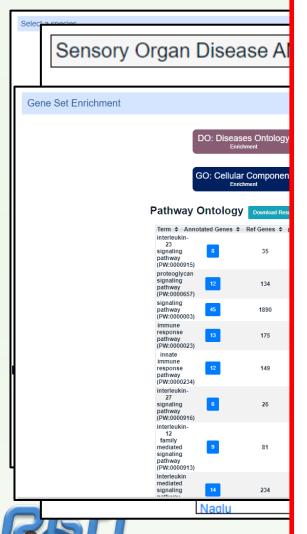




 The table in the enrichment results shows the list of terms over-represented in the annotations for the otitis mediarelated genes, the number of genes in the list annotated to that specific term or its children, uncorrected and Bonferronicorrected p-values and odds ratios for each term.



RGD's views of



23 signaling pathway (PW:0000915)

proteoglycan signaling pathway (PW:0000657)

signaling pathway (PW:0000003)

immune response pathway (PW:0000023)

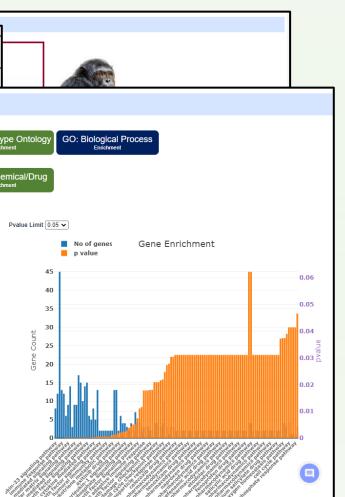
innate immune response pathway (PW:0000234)

interleukin-27 signaling pathway (PW:0000916)

interleukin-12 family

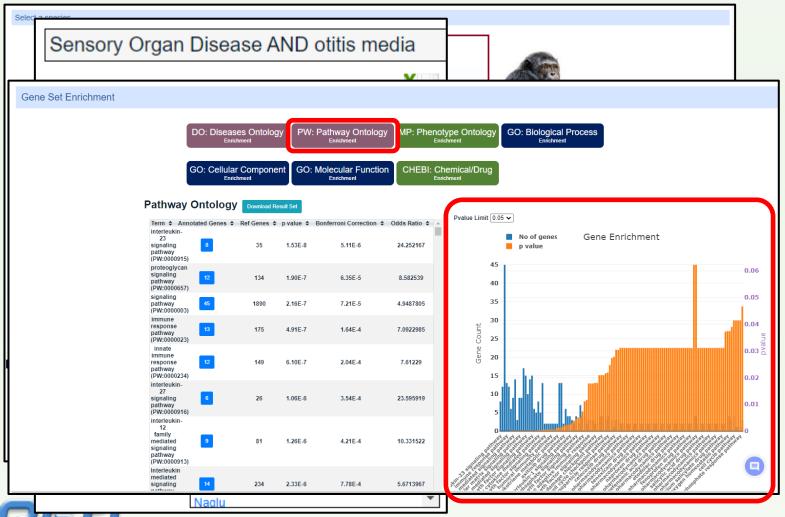
mediated signaling pathway (PW:0000913)

Portals provide consolidated lated to 15 disease categories



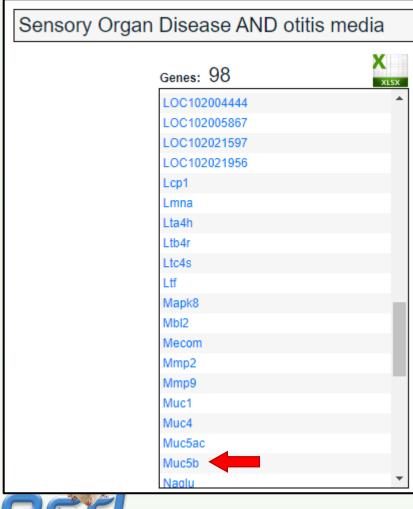
- The table in the enrichment results shows the list of terms over-represented in the annotations for the otitis mediarelated genes, the number of genes in the list annotated to that specific term or its children, uncorrected and Bonferronicorrected p-values and odds ratios for each term.
- Not surprisingly, the most enriched terms in the pathway ontology for the otitis media-related genes are interleukin signaling and immunerelated pathways.





- The table in the enrichment results shows the list of terms over-represented in the annotations for the otitis mediarelated genes, the number of genes in the list annotated to that specific term or its children, uncorrected and Bonferronicorrected p-values and odds ratios for each term.
- Not surprisingly, the most enriched terms in the pathway ontology for the otitis media-related genes are interleukin signaling and immunerelated pathways.
- The graph shows the number of genes in blue and p-values in orange for each term in the list.

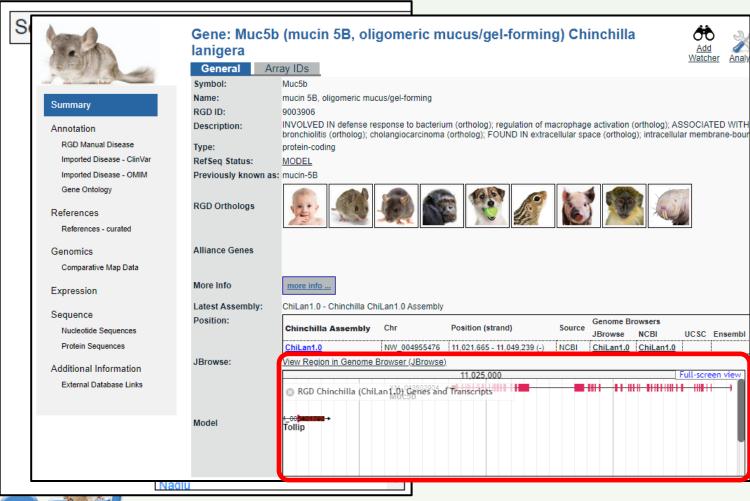




 Each gene in the disease portal list links to the respective gene report page, so clicking "Muc5b" takes you to the chinchilla mucin 5b gene page.



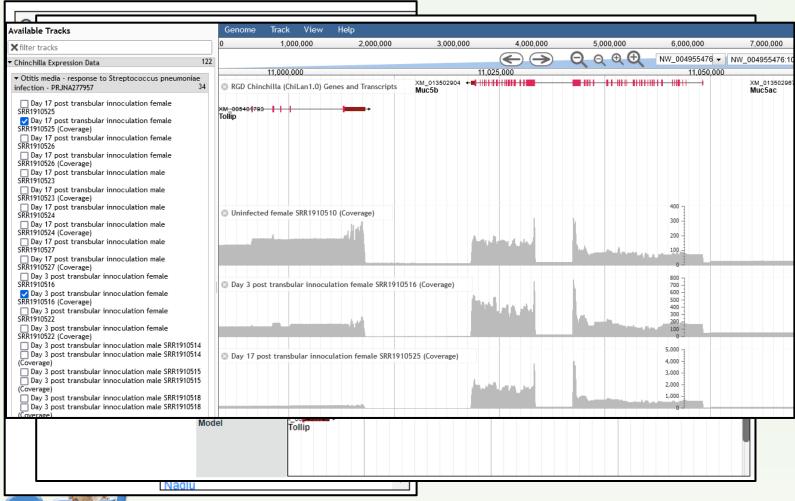




- Each gene in the disease portal list links to the respective gene report page, so clicking "Muc5b" takes you to the chinchilla mucin 5b gene page.
- The Muc5b gene page in turn links to that gene's position in RGD's chinchilla JBrowse genome browser.



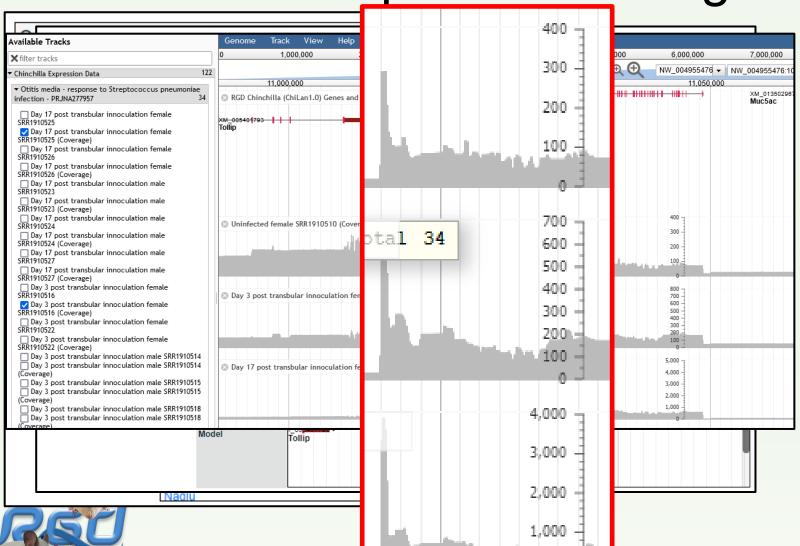




- RGD has loaded genes from NCBI and Ensembl and RNA-Seq BAM alignments and coverage tracks from two projects:
 - Uninfected and infected middle ear mucosa from the Medical College of Wisconsin
 - 14 tissues from samples provided by Dr. Lauren Bakaletz's group at the Research Institute at Nationwide Children's Hospital

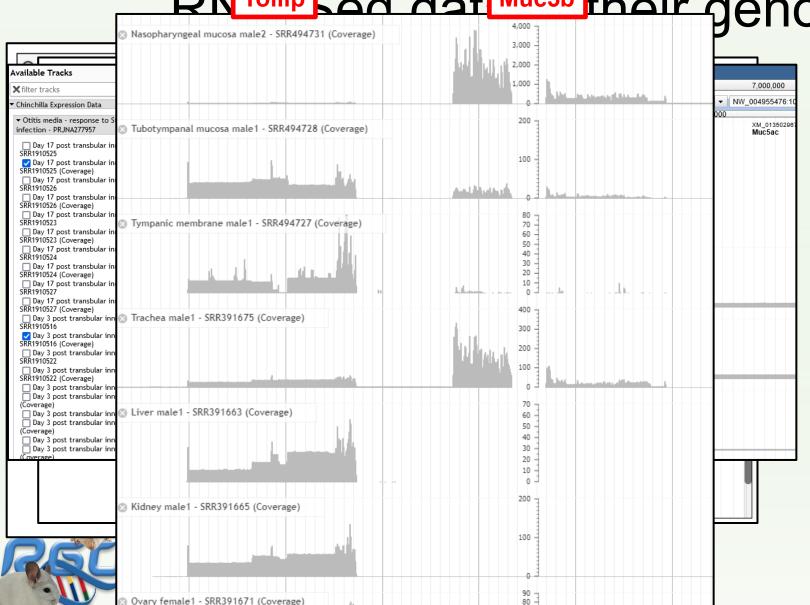






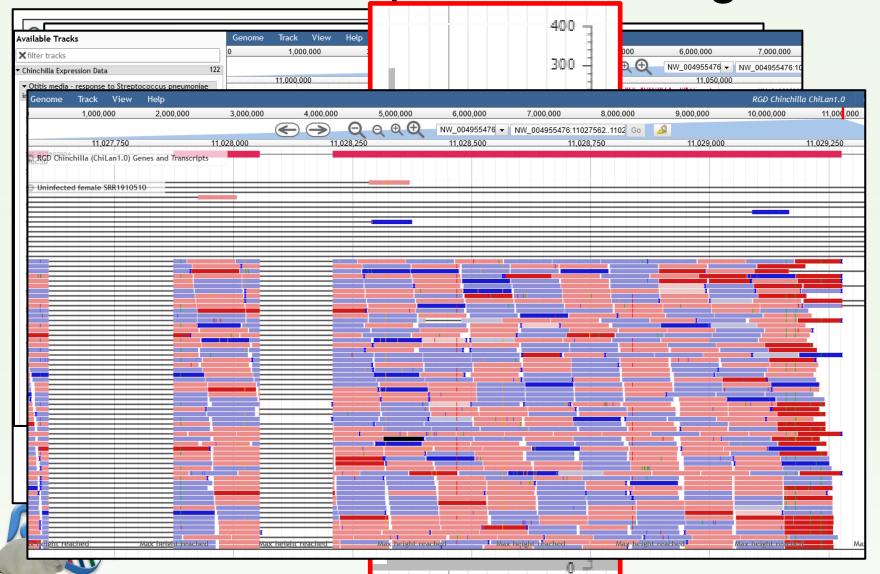
- RGD has loaded genes from NCBI and Ensembl and RNA-Seq BAM alignments and coverage tracks from two projects:
 - Uninfected and infected middle ear mucosa from the Medical College of Wisconsin
 - 14 tissues from samples provided by Dr. Lauren Bakaletz's group at the Research Institute at Nationwide Children's Hospital
- Muc5b increases in expression from uninfected (top graph) through 3 and 17 days of Strep infection.
- Note that although the patterns are similar, the scales are different.

RGD's JBrowse genome browser shows genes and RN Tollip Sea dat Muc5b their genomic context



- Muc5b shows tissue-specific expression. As shown here, there is strong expression in the nasopharyngeal mucosa, lower expression in trachea, and little or none in tubotympanal mucosa, tympanic membrane, liver, kidney, ovaries or testes.
- Again, although the coverage patterns are similar, the levels are an order of magnitude different between nasopharyngeal mucosa and trachea





- This display shows the RNA-Seq BAM alignments for two exons of Muc5b from one of the uninfected middle ear mucosa samples.
- The RNA-Seq data from these studies was used to improve the gene model predictions during the genome annotation process by NCBI and Ensembl.



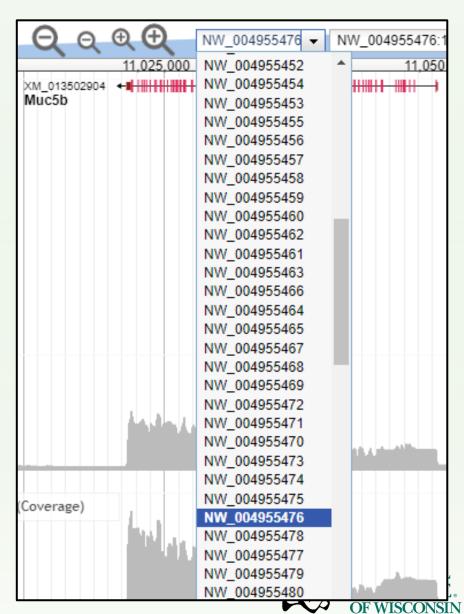
Conclusions and Questions for the research community

- This brief overview is only a subset of the rich set of data and tools that RGD provides for both clinical and translational researchers.
- If you have questions or are interested in more information, I am available to answer questions or give a more complete or targeted tour of what RGD has to offer.
- Contact me at irsmith@mcw.edu, or through the "Contact" link at the top or the feedback link in the bottom right of most RGD pages.
- The Chinchilla genome is only assembled to the point of scaffolds. While this is sufficient for some applications, for true comparative genomics, phylogeny, etc. chromosome-level assembly would be required.

Is there interest in improving the chinchilla genome assembly?

Would groups be willing to do additional sequencing to support chromosome-level assembly?







Thank you!

The RGD Team:

Principal Investigator:
Anne Kwitek, PhD

Co-Investigator:
Mindy Dwinell, PhD

Curation Team:

Jennifer Smith, MSc Stan Laulederkind, PhD Tom Hayman, PhD Shur-Jen Wang, PhD Mary Kaldunski, MSc Mahima Vedi, PhD Wendy Demos, MSc Monika Tutaj, PhD Development Team:

Jeff De Pons, BSc Marek Tutaj, MSc Jyothi Thota, MSc Logan Lamers, BSc Adam Gibson, BSc

We gratefully acknowledge our funders:

RGD is funded by the National Heart, Lung, and Blood Institute (NHLBI; R01HL064541), and the National Human Genome Research Institute (NHGRI) as part of the Alliance of Genome Resources (U24HG010859). and the researchers who contribute data and who faithfully use our website and data!



jrsmith@mcw.edu

https://rgd.mcw.edu









