

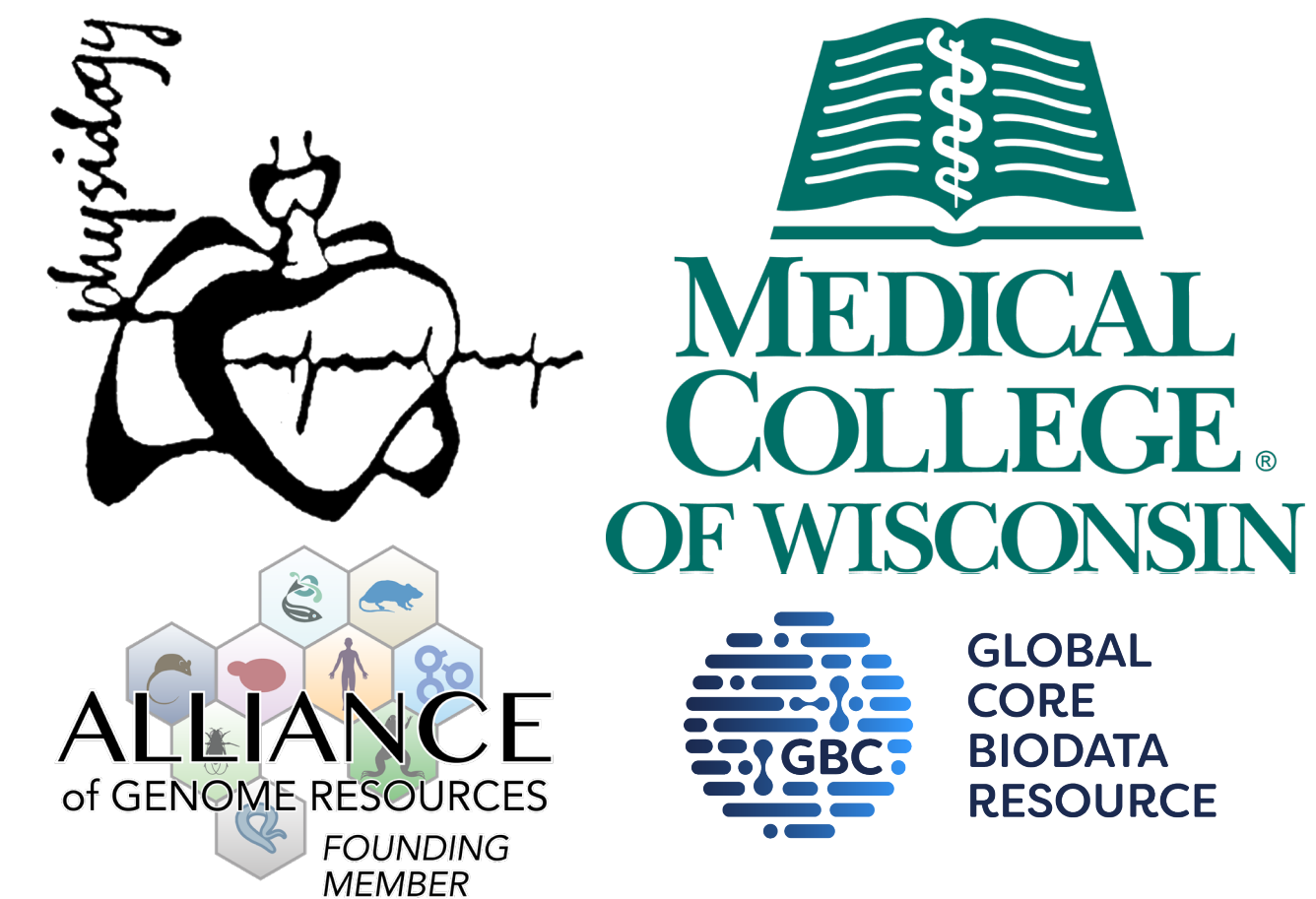


Upgraded genome browsers at the Rat Genome Database support comparative and translational studies

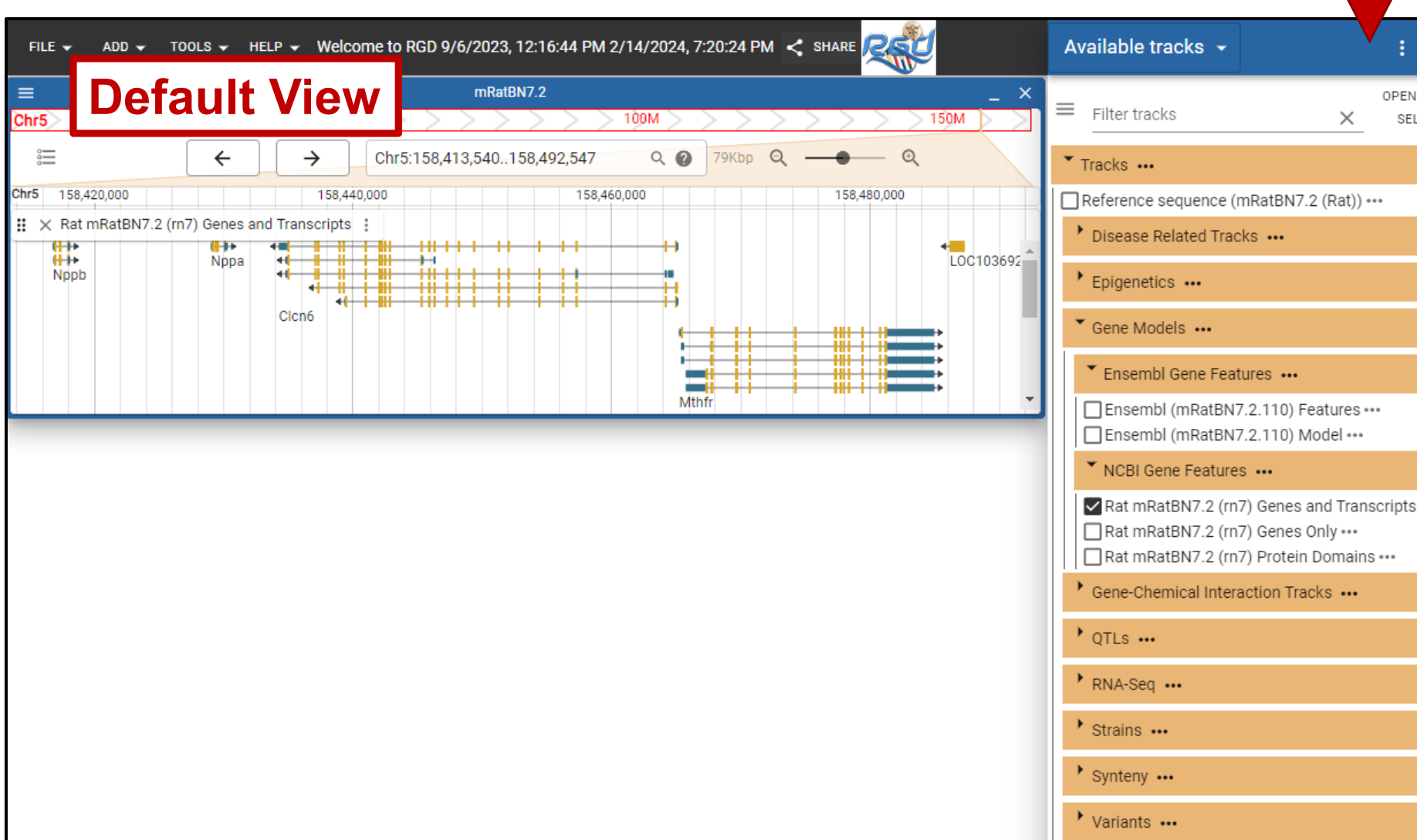
Jennifer R Smith¹, Stanley JF Laudekind¹, G Thomas Hayman¹, Shur-Jen Wang¹, Monika Tutaj¹, Mary L Kaldunski¹, Mahima VEDI¹, Wendy M Demos¹, Marek A Tutaj¹, Jyothi Thota¹, Logan Lamers¹, Adam C Gibson¹, Akhilanand Kundurthi¹, Varun Reddy Gollapally¹, Kent C Brodie², Stacy Zacher³, Jeffrey L De Pons¹, Melinda R Dwinell¹, Anne E Kwitck¹

¹Rat Genome Database, Department of Physiology, ²Clinical and Translational Science Institute, ³Finance and Administration, Medical College of Wisconsin, Milwaukee, WI, 53226, USA

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).

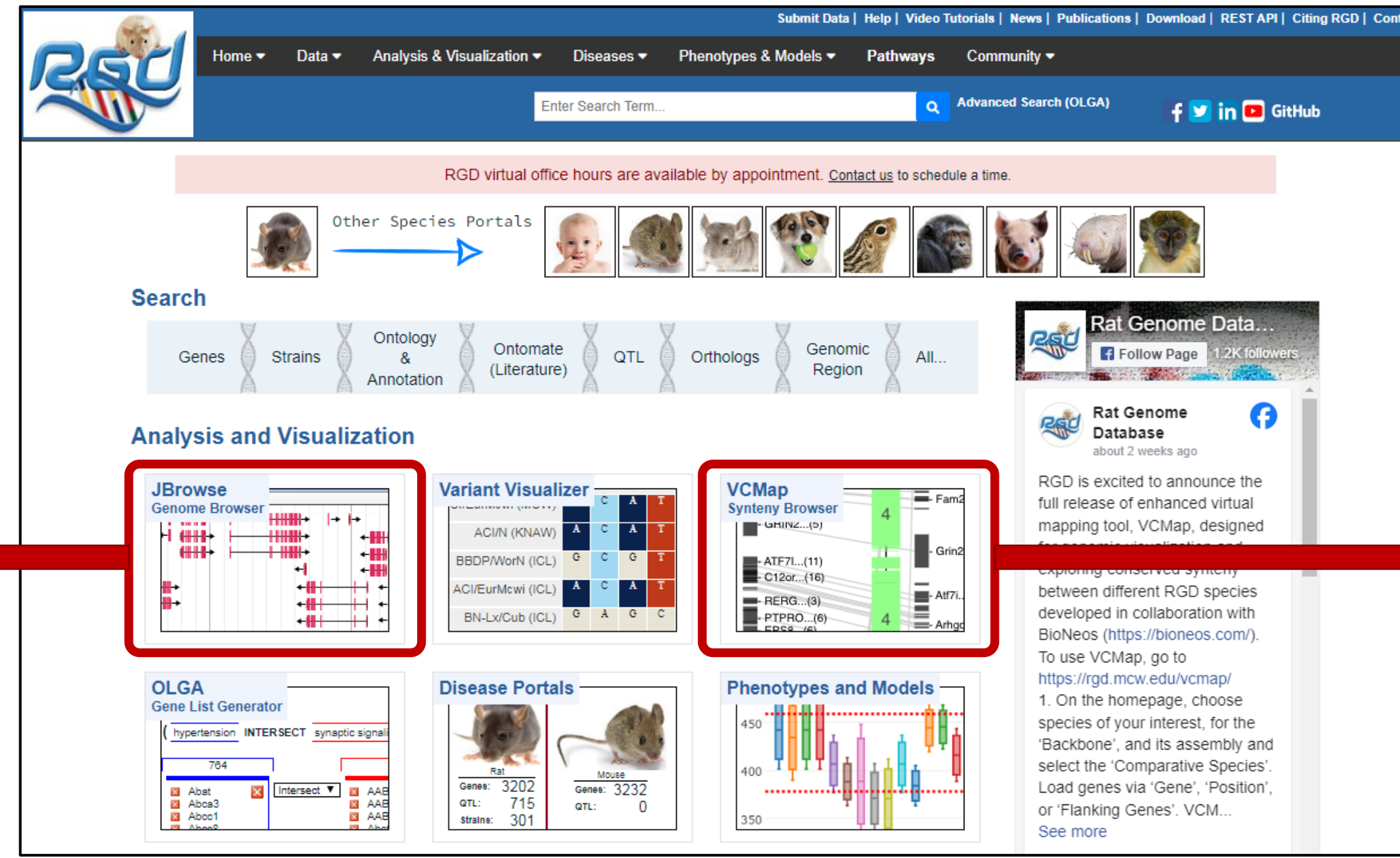
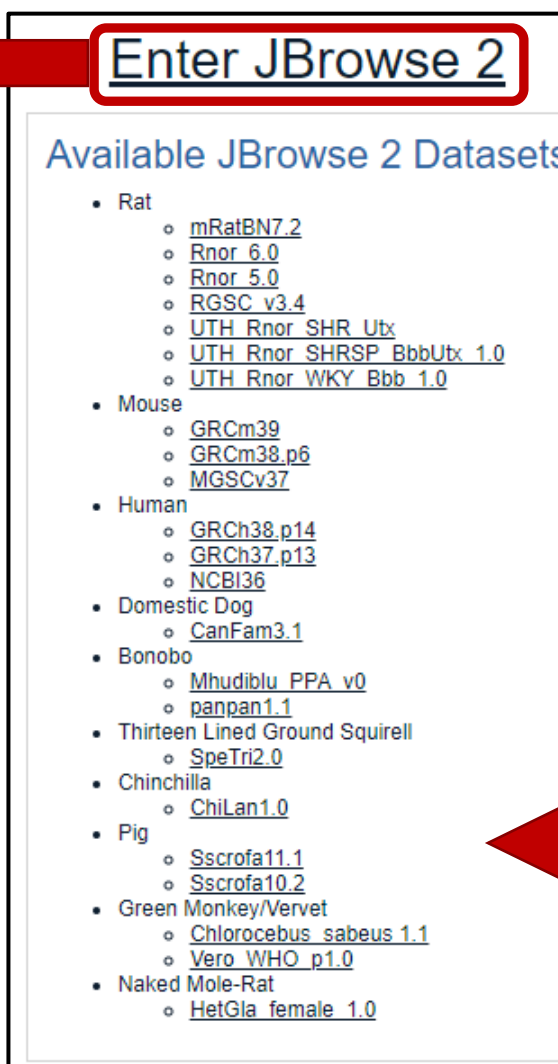


JBrowse 2



The "Enter JBrowse 2" link in the list of datasets accesses the default view. A subwindow shows the NCBI genes in a region of mRatBN7.2 and the track selector is open. Expand or collapse categories in the track selector using the arrows. Use the "Filter tracks" function to find specific tracks by keyword.

JBrowse 2 browsers are available for all ten species for which RGD houses data, including multiple assemblies for many.

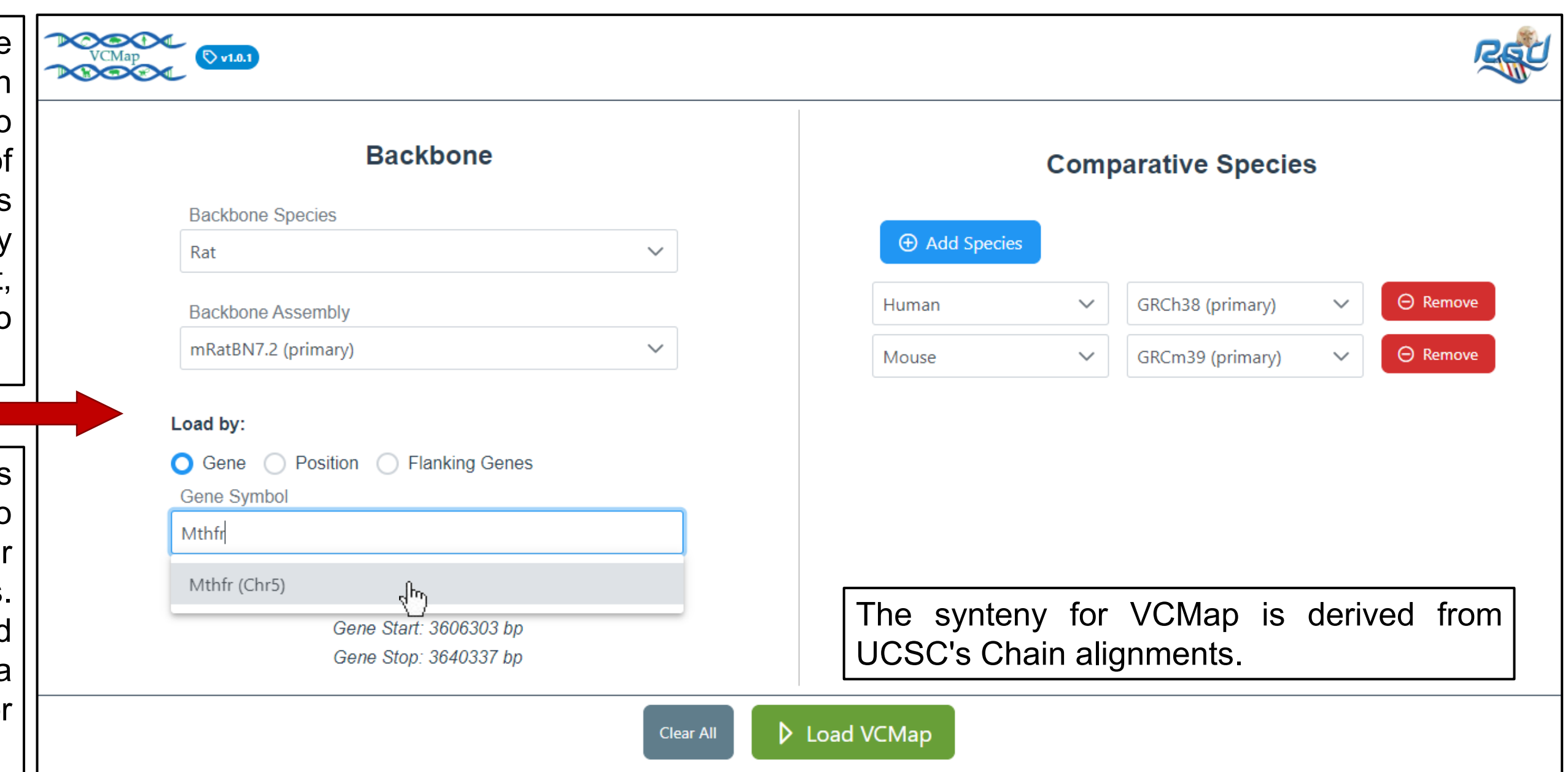


The Virtual Comparative Map (VCM) is an innovative tool designed to allow comparisons of syntenic regions across multiple species. Currently the tool has data for rat, human, mouse, pig, bonobo and green monkey.

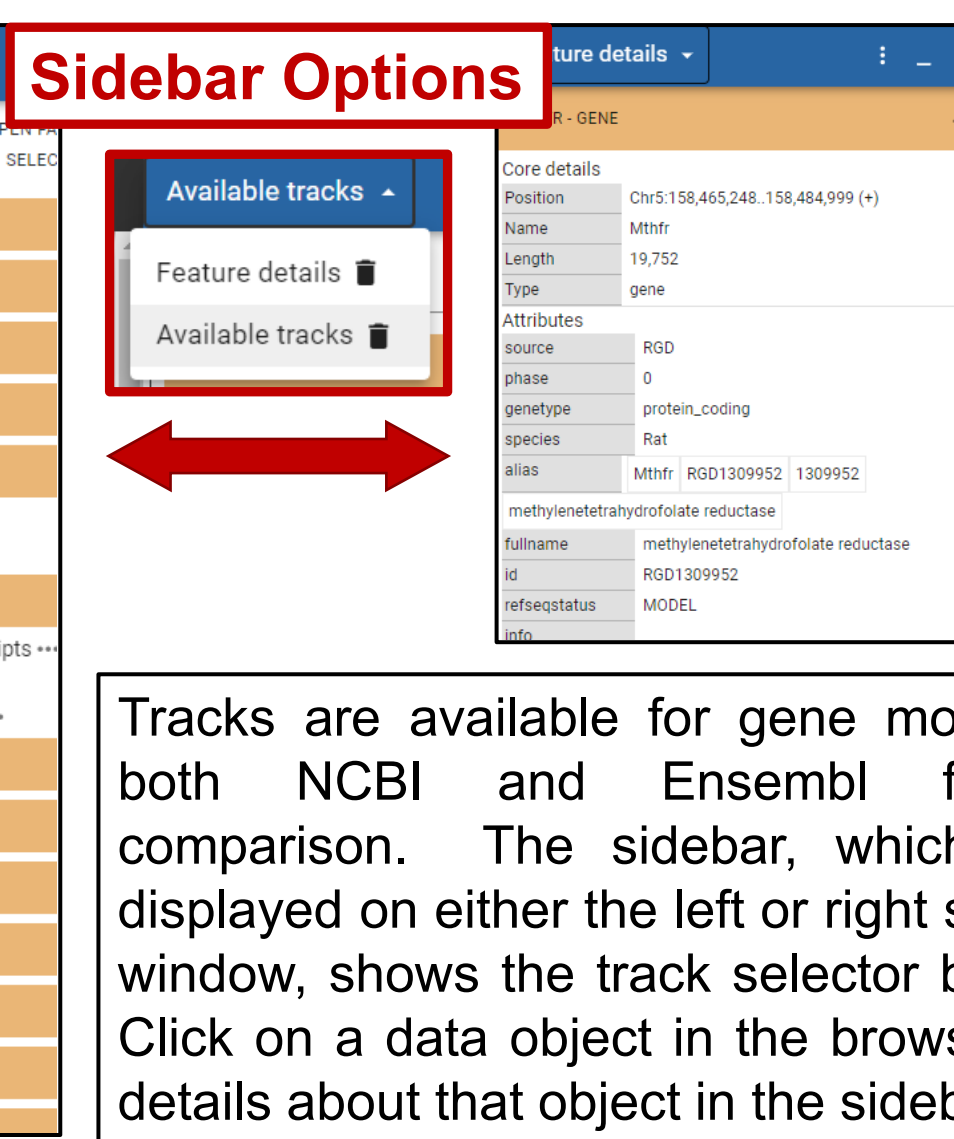
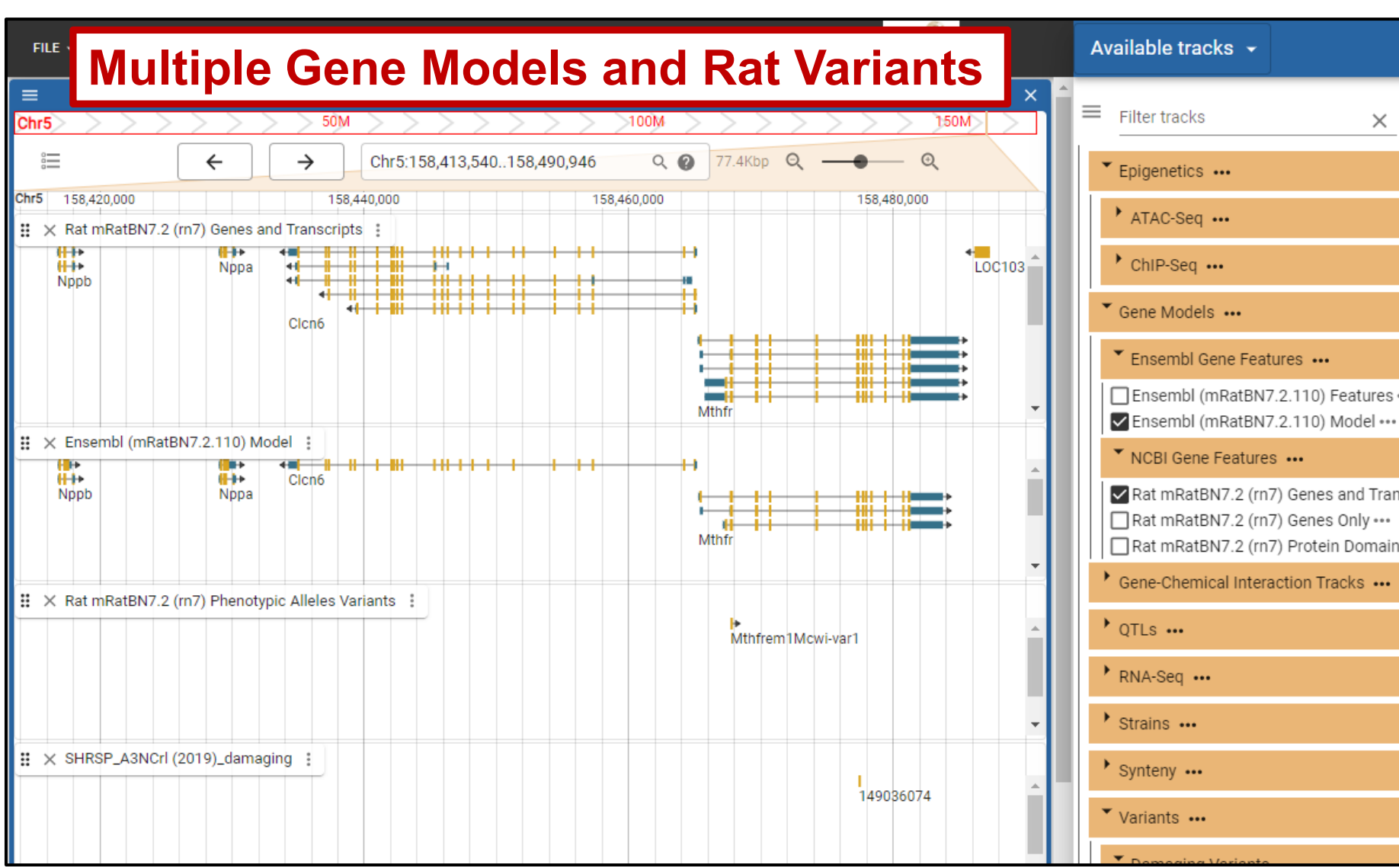
Select a backbone species and a genomic region to view, then choose one or more comparator species. The region can be specified by entering a gene, a chromosomal position or two flanking genes.



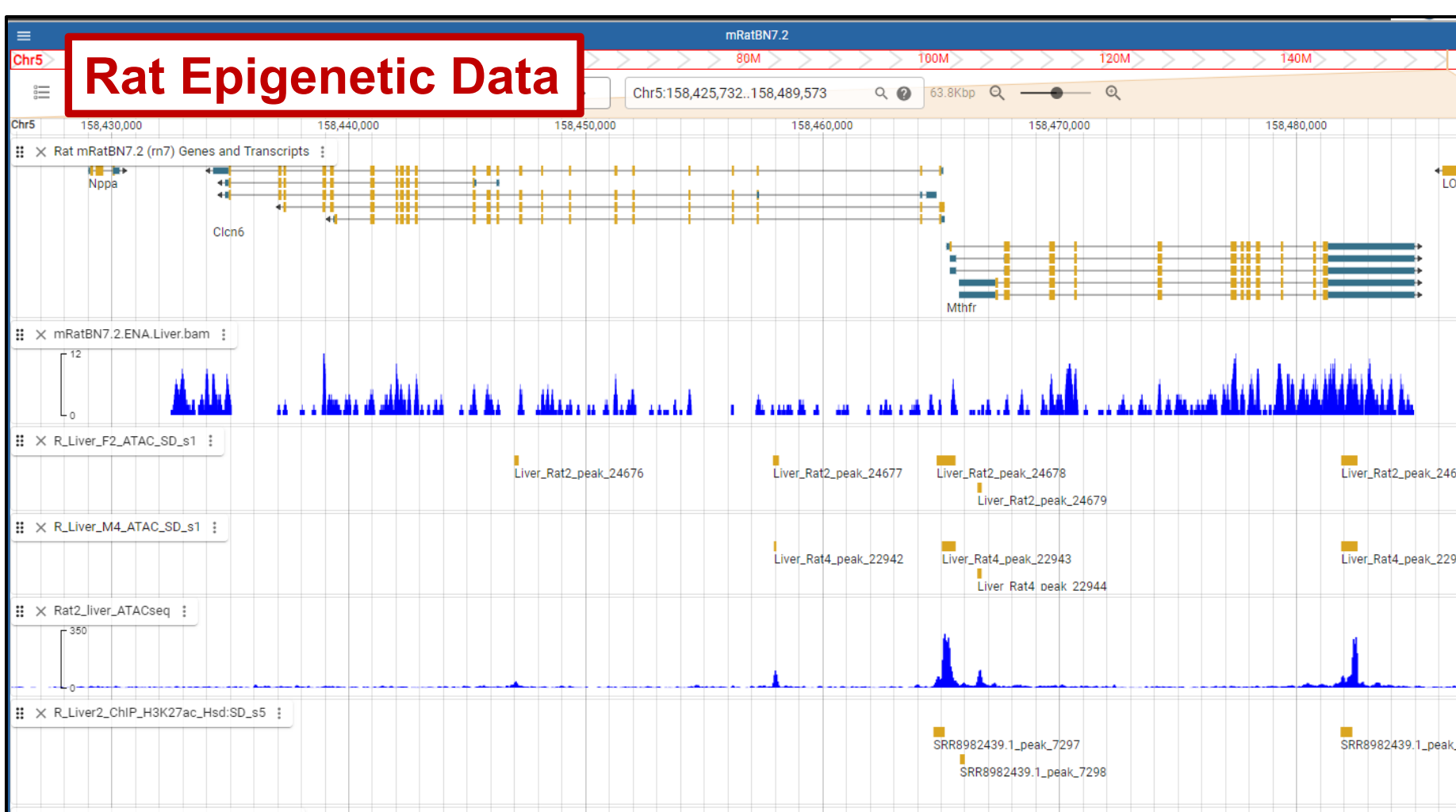
VCM Map



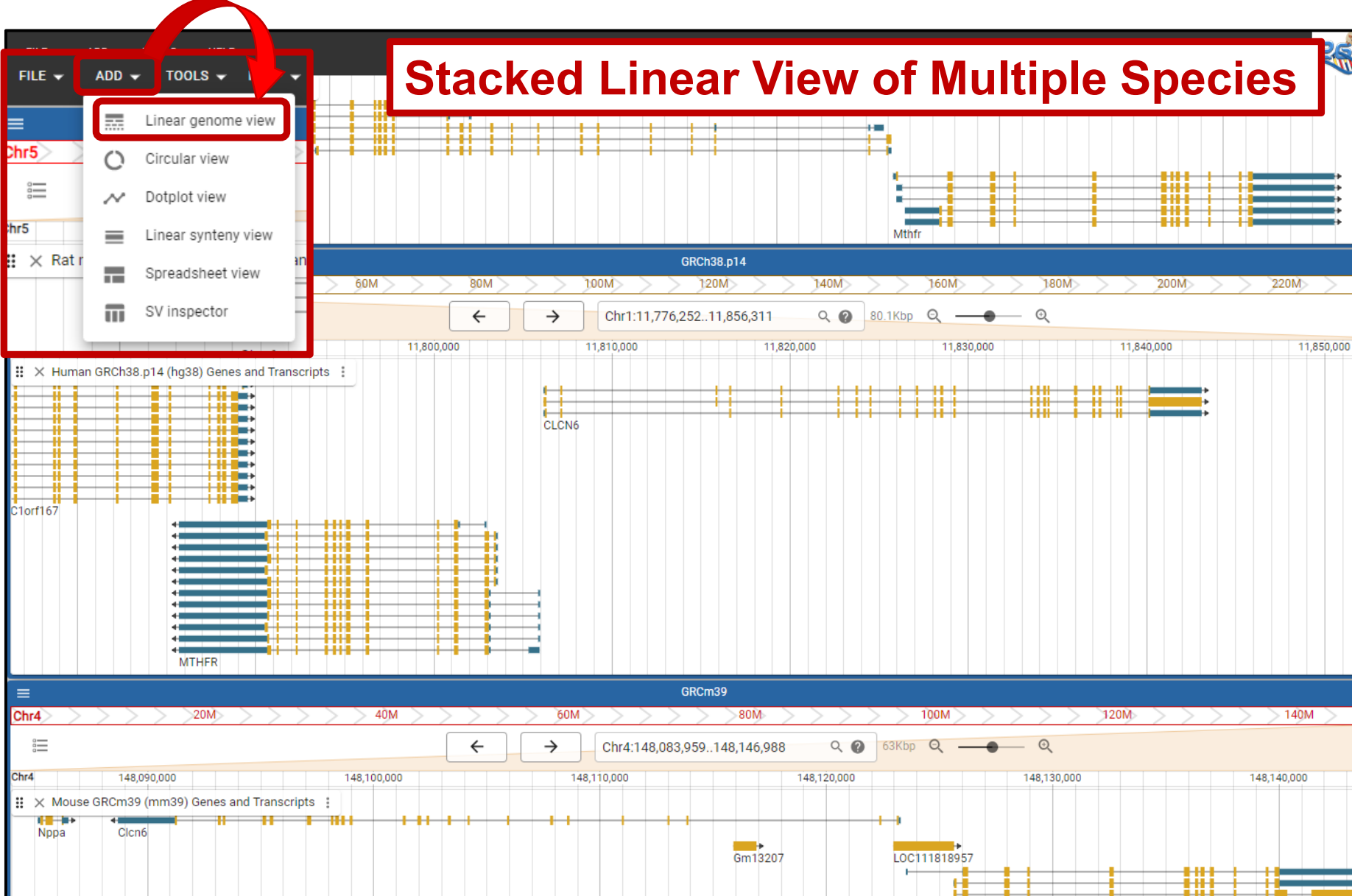
The syntenic for VCM is derived from UCSC's Chain alignments.



Tracks are available for gene models from both NCBI and Ensembl for easy comparison. The sidebar, which can be displayed on either the left or right side of the window, shows the track selector by default. Click on a data object in the browser to see details about that object in the sidebar.

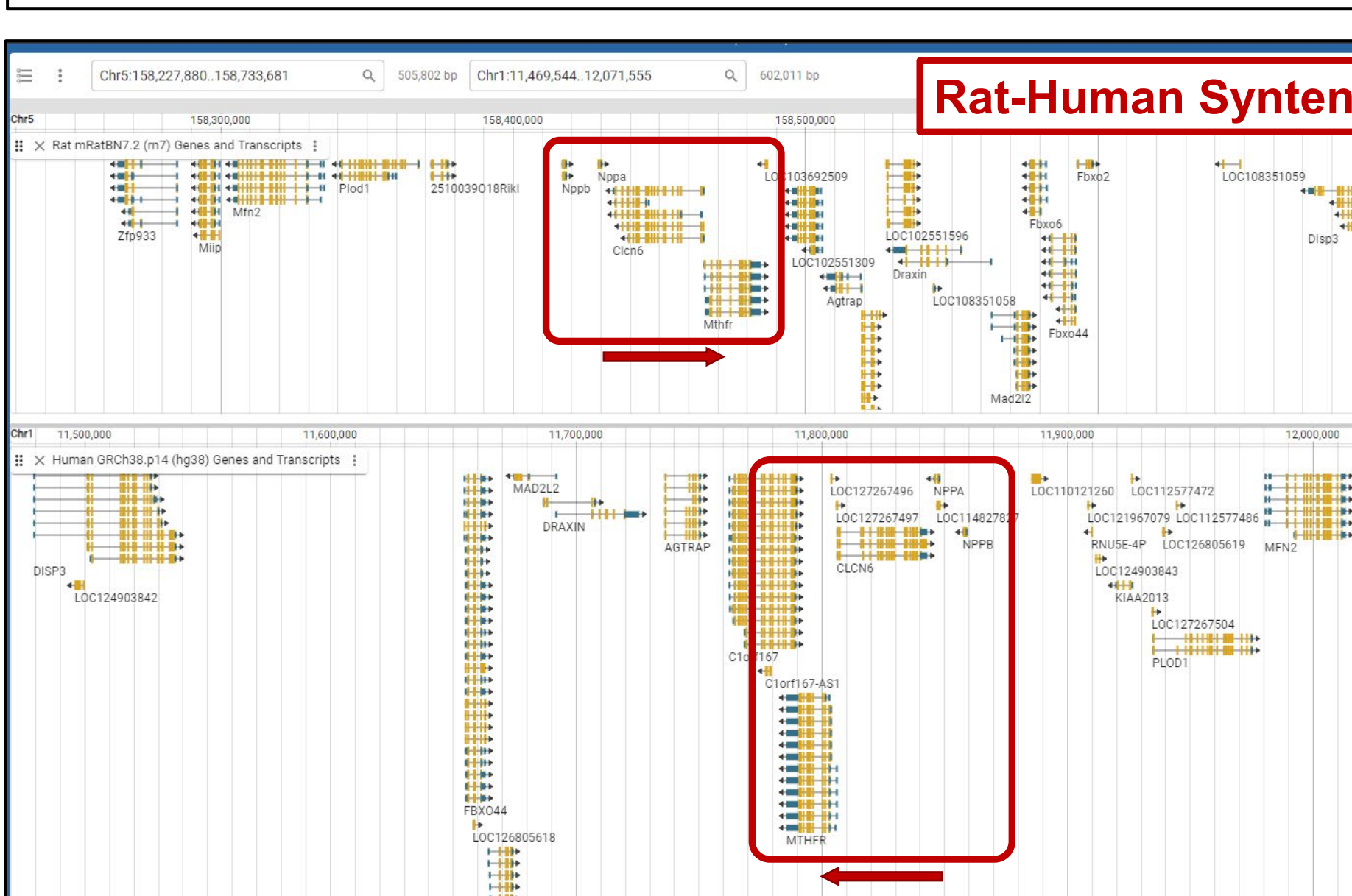


RNA-Seq BAM alignments and epigenetic data from a collaborative project here at MCW are incorporated into RGD's JBrowse 2. Tracks include ATAC-Seq peaks for 25 tissues and BAM alignments for liver and ChIP-Seq peaks for three tissues. Here the tracks have been filtered to only give options for "liver".

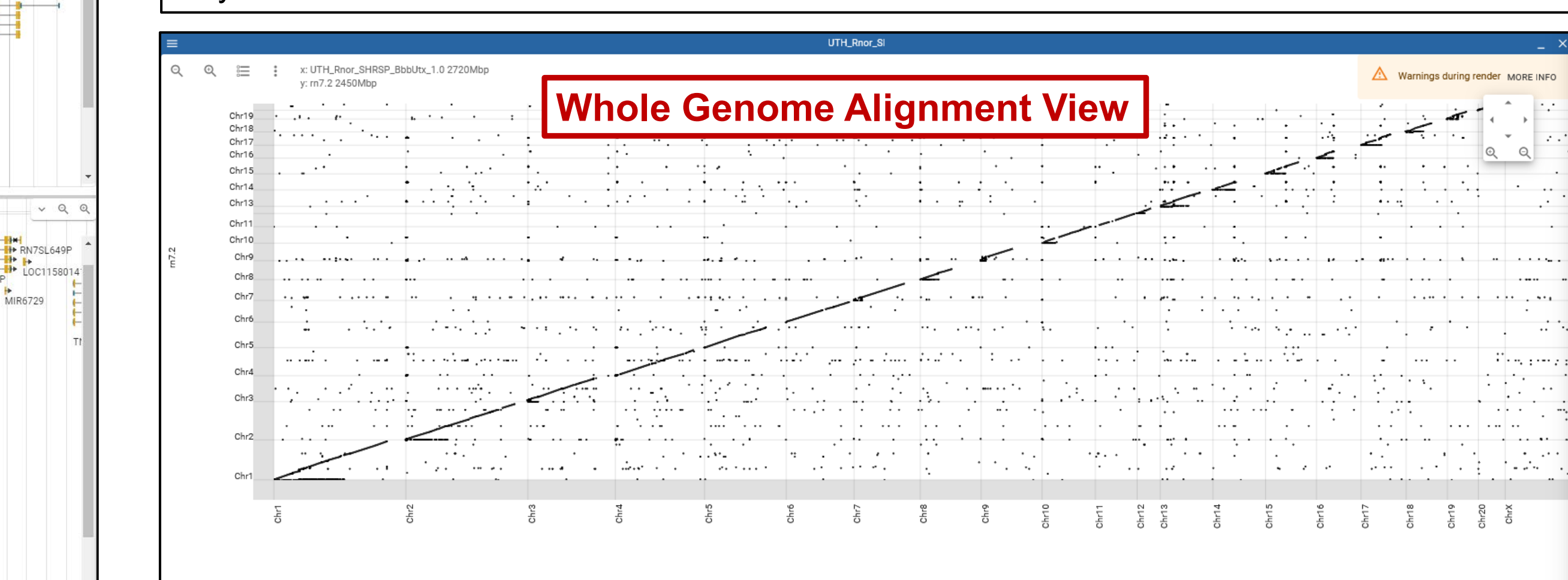


Several views are available to facilitate cross-species comparisons. This shows the stacked linear genome view. Because the browser functions as a subwindow within the larger tool, multiple browsers can be displayed simultaneously. Although each browser functions independently with its own search, navigation and track selector, displaying them within the same larger window makes it easy to compare across species or across assemblies within a species..

The linear syntenic view shows syntenic regions in two species or two assemblies. The view below clearly shows that the regions around the Mthfr gene in mRatBN7.2 and human GRCh38 are reversed.



The whole genome alignment view shown below gives an overview of the correspondence between two assemblies. Here the reference assembly mRatBN7.2 is aligned to a de novo assembly of another rat strain, the SHRSP/BbbUx. The dot plot view demonstrates that there is good correspondence between the two assemblies with no major structural variations. The JBrowse 2 software also provides a breakpoint split viewer and a circular view for visualization of structural variations (not shown). These functions have not yet been implemented at RGD but may be added in the future.



Abstract:

Introduction:

The Rat Genome Database (RGD, <https://rgd.mcw.edu>) is a cross-species knowledgebase and the premier online resource for rat genetic and physiologic data. In addition to rat, RGD houses a diverse set of data for nine additional species and provides mining and analysis tools to assist researchers in effectively utilizing the information.

Methods and Results:

RGD has recently added two upgraded genome browsers—JBrowse 2 and VCM—to our suite of innovative analysis tools. JBrowse 2 (<https://rgd.mcw.edu/jbrowse2/>) is a cutting-edge genome browser with improved functionality for structural variant and comparative genomics visualization. While sharing many features with its predecessors GBrowse and JBrowse 1—the ability to select and view multiple data types in a single view, easy zooming and navigation across a chromosome, the ability to search for the name or ID of a genome feature such as a gene or QTL and go directly to that region of the genome—JBrowse 2 provides expanded functionality designed to facilitate comparative studies. For example, JBrowse 2's browsers are opened in sub-windows embedded in the larger browser window, allowing the user to simultaneously view multiple regions, either within a single assembly or from different assemblies/species, in a single view. JBrowse 2 also supports multiple advanced "view types". For instance, a stacked linear genome configuration is utilized by the linear syntenic viewer to show syntenic alignments and by the breakpoint split viewer to delineate connections between long split alignments or paired end reads across multiple chromosomes. A circular view gives an overview of chromosomal translocations, and a dot plot view provides a comparison of whole genome alignments.

JBrowse 2 browsers have been set up for all ten RGD species, including browsers for multiple assemblies for each species where those are available. For rat in particular, a substantial set of tracks is available. In addition to gene, QTL and strain tracks, RGD's JBrowse 2 provides tracks for RNA-Seq BAM alignments, strain-specific variants, consolidated variants from the European Variant Archive, and most recently, ATAC-Seq and ChIP-Seq epigenetics data aligned onto the mRatBN7.2 assembly.

In addition, for those interested in synteny across more than two species, RGD is developing the Virtual Comparative Map tool (VCM, <https://rgd.mcw.edu/vcm/>). VCM provides a bird's-eye view of synteny between two or more species/assemblies, whereas JBrowse2's synteny viewer is limited to only two. The most recent version of VCM has improved performance and navigation for comparisons of human, rat, mouse, pig, bonobo, and green monkey syntenic regions, and now includes heatmaps of genomic variant densities for human, rat, mouse, and pig.

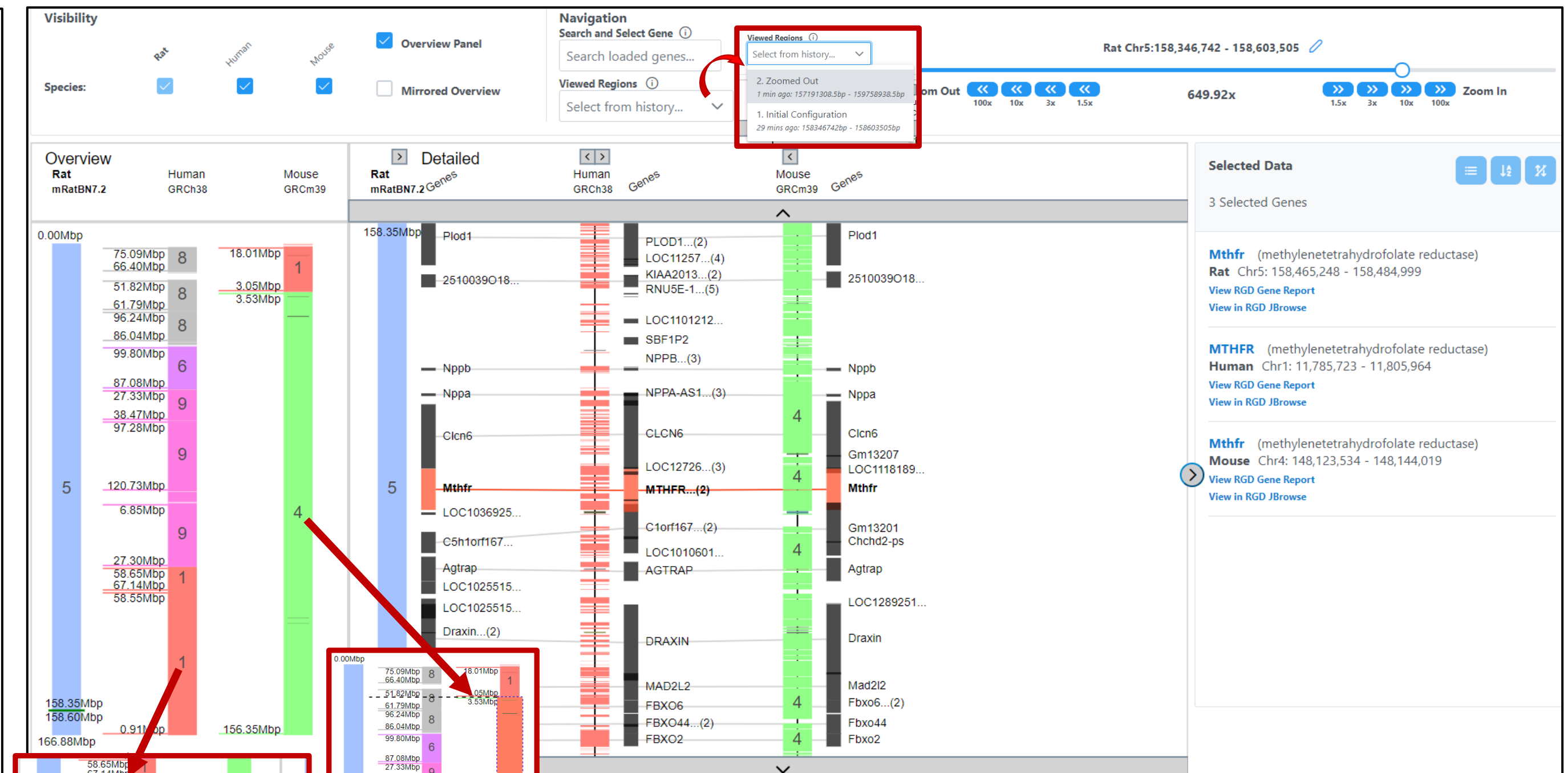
Conclusion:

RGD's JBrowse 2 and VCM provide valuable functionality for researchers engaging in comparative genomics and translational medicine.

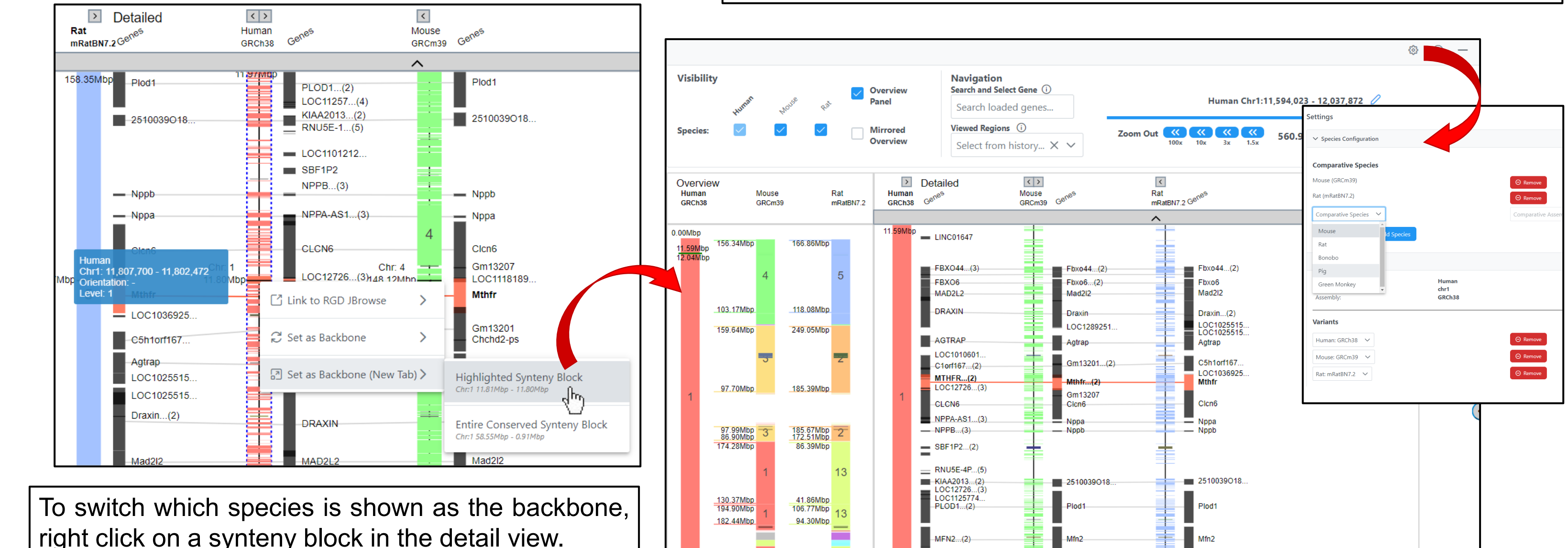


The browser window consists of four sections, navigation at the top, an overview panel at the left, the detail panel in the center and a sidebar on the right which shows information about selected objects such as genes or orthology groups. In the overview panel, hover over a syntenic block to see information about the block, including its orientation relative to the corresponding block on the backbone.

Show/hide species, change the order of the species, search for a gene within the current view, access previous views from the current session, and zoom the detail display using functions in the top section of the page.

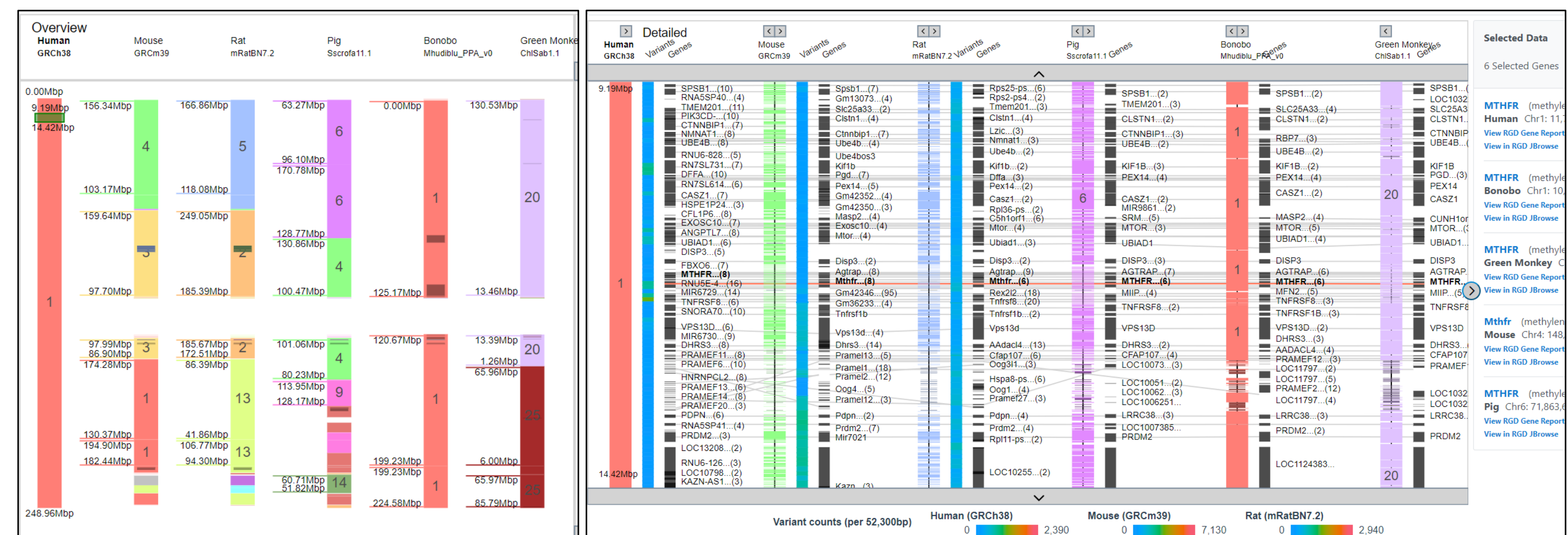


In VCM, syntenic blocks are aligned so that the orthologs appear in the same order for the comparison species as they do for the backbone. Since the syntenic block that contains the Mthfr gene is reversed in human compared to both rat and mouse, it is shown in the opposite direction as demonstrated by the higher chromosomal position at the top of the block and the lower number at the bottom in the overview and the crossed lines between the rat and human blocks (left callout).



To switch which species is shown as the backbone, right click on a syntenic block in the detail view.

Here the backbone is changed from rat to human. As shown in the callout, the "Settings" icon provides access to a popup page where species can be added or removed and additional data tracks, currently variant density tracks, can be added to the display.



With human as the backbone, all five additional species can be compared showing MTHFR orthologs in all of the species. For human, mouse and rat, heatmaps show variant density. Human variants are imported from ClinVar and the GWAS Catalog and mouse variants are imported from the European Variant Archive. Rat variants are derived from whole genome sequencing of multiple rat strains as part of Dr. Mindy Dwinell's Hybrid Rat Diversity Project here at the Medical College of Wisconsin. Mouse over a bin to see the number of variants in that segment.