

# PhenoGen Informatics

The site for quantitative genetics of the transcriptome  
<http://phenogen.ucdenver.edu>

## **The PhenoGen Group**

**Initiated and directed by Dr. Boris Tabakoff**

**Presented by Dr. Laura Saba**

**Anschutz Medical Campus**

**University of Colorado Denver**

**\*\*Please use Firefox for optimal results and to follow along  
in the workshop**

# Acknowledgements

## **The PhenoGen Group:**

Boris Tabakoff, PhD; Paula L. Hoffman, PhD; Spencer Mahaffey, MA; Laura Saba, PhD; Stephen Flink, PhD; Lauren Vanderlinden, MS; Yinni Yu, MS

## **Collaborators:**

Morton Printz, PhD; Michal Pravenec, PhD

## **Technical Support:**

Adam Chapman; James Huntley; Laura Breen; Donna Moye

## **Grant Support:**

National Institute on Alcohol Abuse and Alcoholism; the Banbury Fund (Robertson Family)

# Outline

1. Introduction to PhenoGen Informatics
2. Interactive demo of candidate gene approach
3. Summary of phenotype approach
4. Downloadable data
5. Future directions

# Goal for the PhenoGen Project

## Transcriptional Connectome

**To generate a new image of organs as networks of interacting elements (transcripts)**

- Collect genome sequence and full transcriptome information for organs (brain, liver, heart).
  - We have completed exon array analysis for these three organs
  - RNA-Seq for brain in 30 strains of the HXB/BXH panel completed
  - RNA-Seq for heart and liver of the RI progenitor strains completed



# Why Networks?

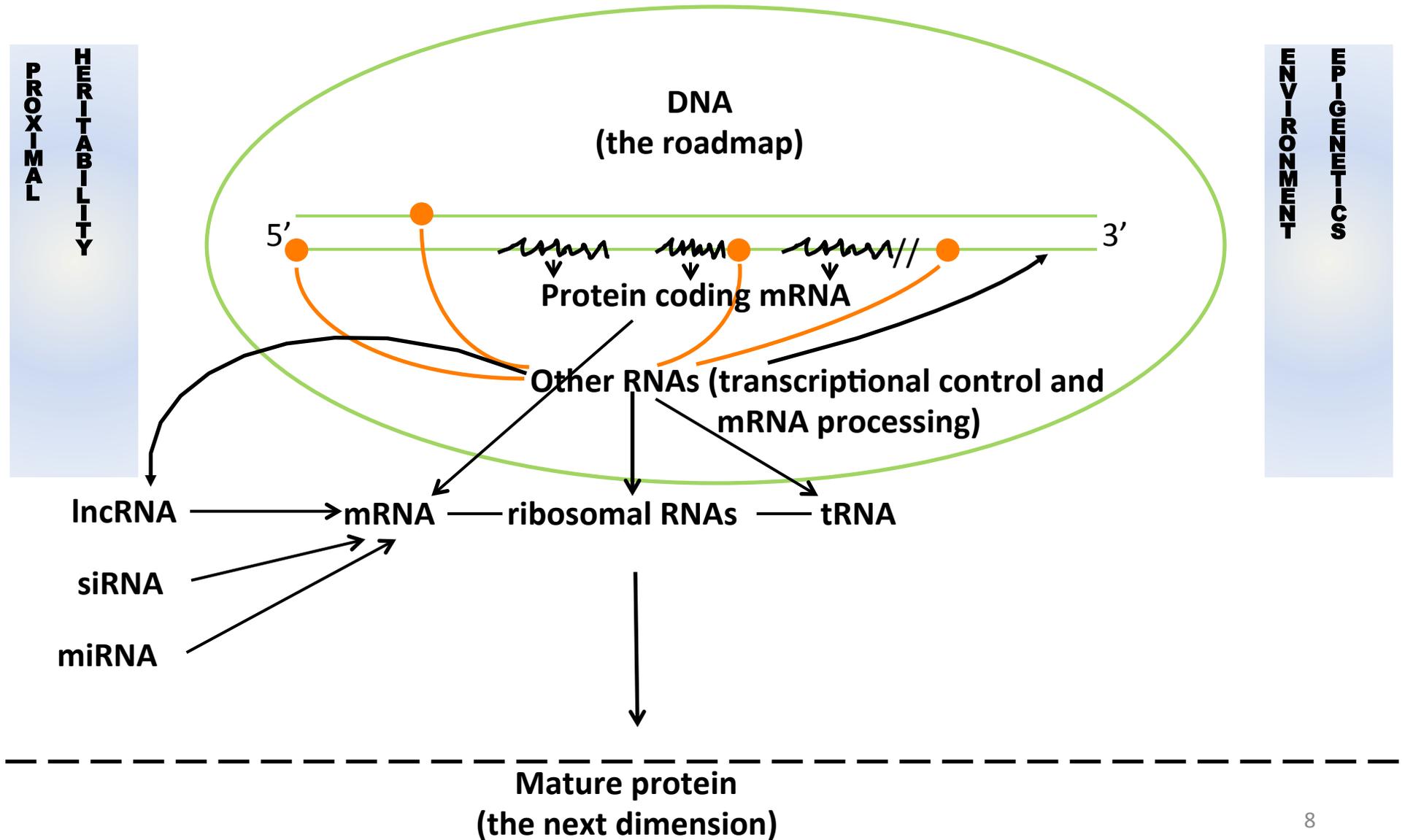
1. The brain is a complex hierarchical network spatio-temporally linked through structure and function.
2. Complex pathologic traits can be conceptualized as systems disorders of failed network regulation.
3. The brains structural and functional systems have features of complex networks that can be described through application of “graph theory” (small world topology).
4. The generation of a “Transcriptional Connectome” representing the Resting State transcriptional networks provides power for understanding predisposition to disease, etiology of organ or behavioral pathology and response to medications or toxins.

# Why Study the RNA Dimension

## Transcriptome links DNA and complex traits/diseases

- A. RNA is one of the first quantitative links between DNA sequence and phenotype (an endophenotype).**
- B. Transcriptome information addresses part of the GWAS Gap: how does an identified DNA polymorphic locus contribute to disease?**
- C. First step where DNA sequence and environment interact.**
- D. Implementation of graph theory at the transcript level provides insight into genetic/environmental interactions that are the basis for susceptibility to complex diseases.**

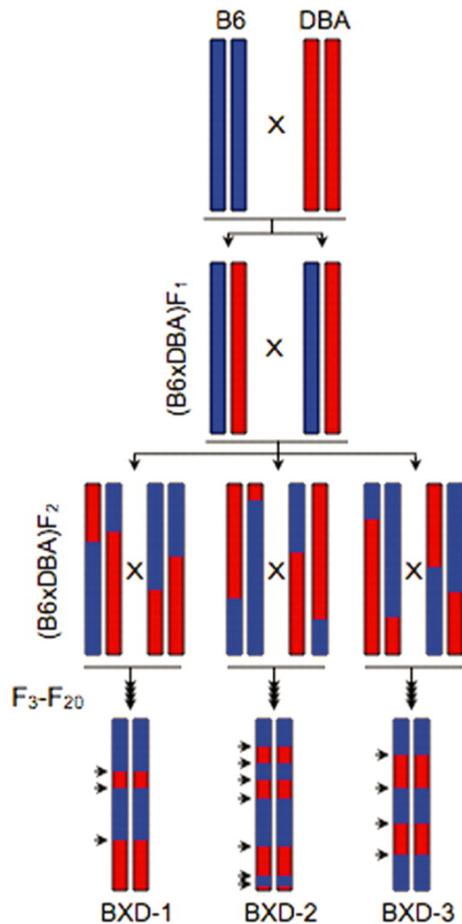
# The RNA Dimension (the true intermediate phenotype)



# Co-expression as a measure of the “connectome”

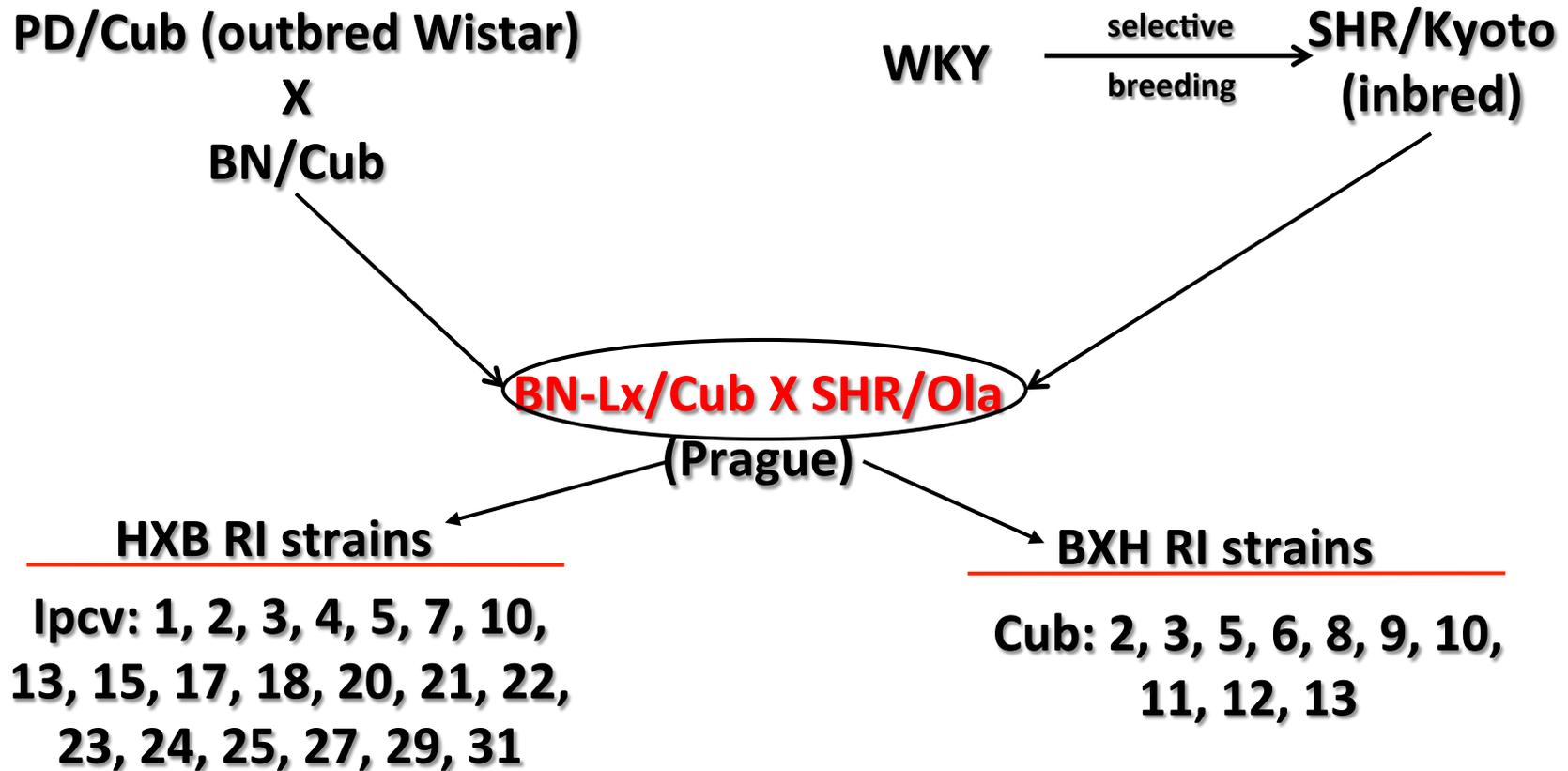
- **Theory** – if the magnitude of RNA expression of two transcripts correlates over multiple “environments” (genomes), then the two transcripts are involved in similar biological processes
- Caveats when multiple environments are multiple genetic backgrounds (false positive correlations)
  - Linkage Disequilibrium
  - Cell-type mixing proportions

# Recombinant Inbred Rodent Panel



- Genetic identity is retained over generations
- Cumulative genetic and phenotype data across labs
- Ideal genetic controls for studying interventions/ environmental effects

# Origin of HXB/BXH RI Panel



- > 4,500,000 SNPs/indels between progenitors
  - Haplotype map generated

## Current Data – RNA-Seq

Strains	Tissue	Sex	Number of Biological Replicates Per Strain	Number of Paired-End Reads (rRNA-depleted total RNA)	Number of Paired-End Reads (polyA+-selected RNA)	Number of Single- End Reads (small RNA)
SHR and BNLx	brain	male	3	645 million	192 million	96 million
SHR and BNLx	brain	female	4	982 million	---	297 million
SHR and BNLx	liver	male	3	583 million	---	342 million
SHR and BNLx	heart	male	4	790 million	---	300 million
30 RI Strains	brain	male	1 to 2	7.3 billion	----	1.9 billion

## Current Data - Microarrays

Panel	Platform	Tissue	Number of Biological Replicates
21 HXB/BXH RI Strains	Exon Array	Brain	2 to 4
21 HXB/BXH RI Strains	Exon Array	Liver	2 to 4
21 HXB/BXH RI Strains	Exon Array	Heart	2 to 4
21 HXB/BXH RI Strains	Exon Array	Brown Adipose	2 to 4
60 LXS Mouse RI Strains	Exon Array	Brain	6
30 BXD Mouse RI Strains	3' Array	Brain	4 to 7
26 Inbred Mouse Strains	3' Array	Brain	4 to 6

# **CANDIDATE GENE EXAMPLE**

# Transcriptome Information on a Gene of Interest

- Genome View
  - Annotated isoforms
  - DNA variants in parental strains
  - Published QTL
- Transcriptome View
  - Exon array information
  - RNA-Seq results
- Selected Feature
  - Gene details and external links
  - Gene eQTL
  - Weighted gene co-expression network analysis
  - Probe set level data

# Genome View

1. Click Genome/  
Transcriptome Data  
Browser
2. Type in Rat/Mouse  
Official Gene Symbol  
into “Gene Identifier or  
Region” box
3. Select **Rattus Norvegicus  
(rn5)** from “Species” box
4. Select **Genome  
(Predefined)** from  
“Initial View” box
5. Click “Go”

The image shows two screenshots of the PhenoGen Informatics website. The top screenshot shows the main navigation menu with the 'Genome / Transcriptome Data Browser' link circled in blue. The bottom screenshot shows the 'Genome/Transcriptome Data Browser' page with a search form. The search form has a 'Gene Identifier or Region' field containing 'Oprm1', a 'Species' dropdown menu set to 'Rattus norvegicus (rn5)', and an 'Initial View' dropdown menu set to 'Genome (Predefined)'. A blue oval highlights the search form area. Below the search form are two buttons: 'Quick Navigation Demonstration' and 'Custom View/Custom Track Demonstration'.

PhenoGen Informatics  
The site for quantitative genetics of the transcriptome.

Overview Genome / Transcriptome Data Browser Available Data Downloads Microarray Analysis Tools Gene List Analysis Tools QTL Tools About Help Login/ Register

Welcome to PhenoGen Informatics  
The site for quantitative genetics of the transcriptome.

Hover over or click on nodes in the graph below to see the tools/data available on the site.  
Green no login required.  
Blue sections require a login.

PhenoGen - Genome/Transcriptome Data Browser

PhenoGen Informatics  
The site for quantitative genetics of the transcriptome.

Overview Genome / Transcriptome Data Browser Available Data Downloads Microarray Analysis Tools Gene List Analysis Tools QTL Tools About Help Login/ Register

1. Enter a gene identifier(e.g. gene symbol, probe set ID, Ensembl ID, etc.) in the gene field.  
or  
Enter a region such as  
"chr1:1-50000" which would be Chromosome 1 @ bp 1-50,000.  
"chr1:5000+2000" which would be Chromosome 1 @ bp 3,000-7,000.  
"chr1:5000+2000" which would be Chromosome 1 @ bp 5,000-7,000.  
or  
Click on the Translate Region to Mouse/Rat to find regions on the Mouse/Rat genome that correspond to a region of interest in the Human/Mouse/Rat genome.  
2. Choose a species.  
3. Click Get Transcription Details.

Gene Identifier or Region:  Species:  Initial View:

Or

Quick Navigation Demonstration Custom View/Custom Track Demonstration

PhenoGen - Genome/Transcri... x

phenogen.ucdenver.edu/PhenoGen/gene.jsp

PhenoGen Informatics  
Genetics of the transcriptome.

Genome/Transcriptome Data Browser

Navigation Buttons

Control Displayed Tracks

Gene Identifier or Region: ENSRNOG00000018191 Species: Rattus norvegicus (rn5) Initial View: Genome (Predefined)

Or Translate Region to Mouse/Rat

Not sure where to start? watch a quick navigation demonstration or view the help images again

Region Image View: Genome Organism: Rat(rn5) Strain:BN PhenoGen v2.5.2(4/8/2015)

Navigation Hints: Hold mouse over areas of the image for available actions.

Draw Genes as Transcript. Select/Edit Views

chr1

44,810,000 44,820,000 44,830,000 44,840,000 44,850,000 44,860,000

Ensembl Protein Coding Transcripts

Ensembl Long Non-Coding Transcripts

Ensembl Brain Small RNA Transcripts

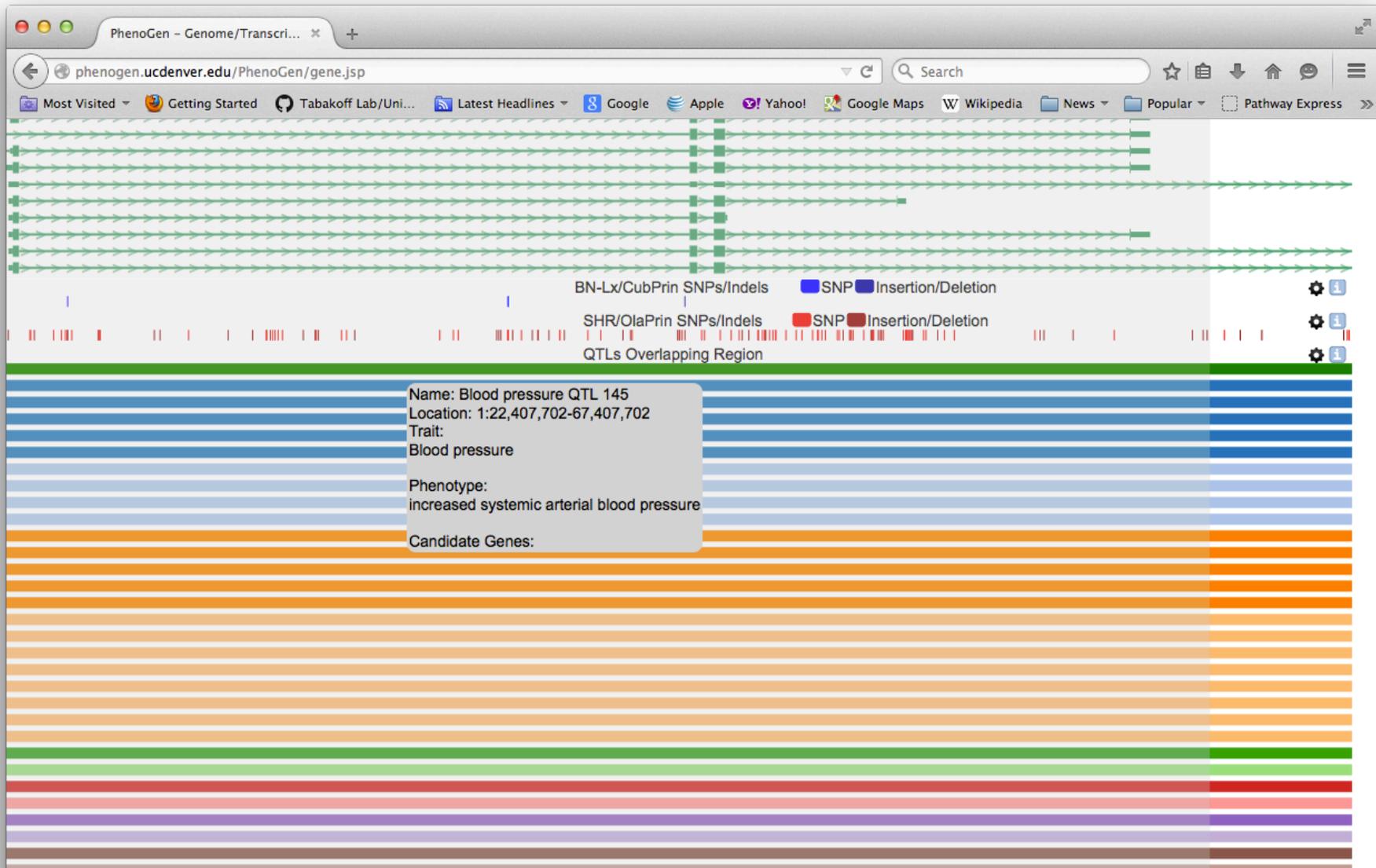
Ref Seq Transcripts

Reviewed Validated Provisional Inferred Predicted

# Genetic Variants Between Parental Strains of RI Panel



# Published Behavioral/Physiologic QTL



# Get Region Summary

The screenshot displays the PhenoGen website interface for a specific genomic region. The browser address bar shows the URL `phenogen.ucdenver.edu/PhenoGen/gene.jsp`. The main content area is titled "Selected Feature Image" and shows a genomic track for the gene ENSRNOG00000018191 (Oprm1). The track includes various annotations such as Ensembl Protein Coding Transcripts, Ensembl Long Non-Coding Transcripts, Ensembl Brain Small RNA Transcripts, Ref Seq Transcripts, and SNPs/Indels. A red box highlights the "Close Selected Feature Image" button, which is a small 'X' icon in the top right corner of the feature image area. A red circle is drawn around this button, and a red arrow points to it from the text "Close Selected Feature Image".

# Region Summary

PhenoGen - Genome/Transcri... x +

phenogen.ucdenver.edu/PhenoGen/gene.jsp

Search

Most Visited Getting Started Tabakoff Lab/Uni... Latest Headlines Google Apple Yahoo! Google Maps Wikipedia News Popular Pathway Express

Track Details Genes with an eQTL in this region WGCNA

**Region Summary**

**Track List:**

- Ensembl Protein Coding Transcripts: 3
- Ensembl Long Non-Coding Transcripts: 0
- Ensembl Brain Small RNA Transcripts: 0
- BN-Lx/CubPrin SNPs/Indels: 4
- Ref Seq Transcripts: 13
- SHR/OlaPrin SNPs/Indels: 132
- QTLs Overlapping Region: 41

**Break down of track count\***

\*Note: Depending on the track settings some features may not be displayed and will not be reflected in the image above.

**Features in Selected Track**

View Columns

Search:

RGD ID	QTL Name	Trait	Phenotype	Associated Diseases	References RGD Ref PubMed	Candidate Genes	bQTL Region	LOD Score
70225	Blood pressure QTL 58	Blood pressure - direct systolic	increased systemic arterial blood pressure	Hypertension	70067 69692		chr1:36,395,668-180,434,202	3.3
631494	Blood pressure QTL 95	Blood pressure - systolic	increased systemic arterial blood pressure	Hypertension	11082136 68885 11160999	Slc9a3	chr1:23,257,188-50,297,622	40.0

# Alternate Region Summary Under Different View

The screenshot shows a web browser window with the URL `phenogen.ucdenver.edu/PhenoGen/gene.jsp`. The page displays two tracks for read counts: "Brain Illumina PolyA+ RNA Read Counts" and "Brain Illumina Total RNA(rRNA depleted) Read Counts", both with a scale from <1 to 5,000+. Below the tracks are navigation tabs for "Track Details", "Genes with an eQTL in this region", and "WGCNA". The "Region Summary" tab is active, showing a "Track List" with the following items:

- Affy Exon 1.0 ST Probe Sets: 38
- Brain RNA-Seq Reconstruction Protein Coding / PolyA+ Transcripts: 0
- Brain RNA-Seq Reconstruction Long Non-Coding / Non-PolyA+ Transcripts: 1

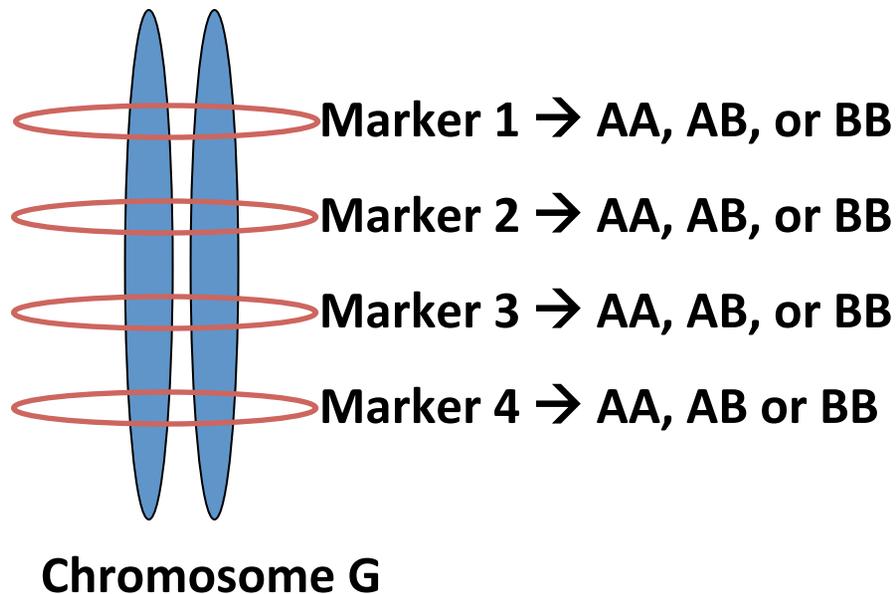
To the right of the track list is a pie chart titled "Break down of track count\*". The pie chart is divided into four segments: a black segment (14), a green segment (12), a red segment (8), and a blue segment (4). A tooltip for the red segment shows "Name: Core" and "Count: 8".

\*Note: Depending on the track settings some features may not be displayed and will not be reflected in the image above.

At the bottom of the page, there is a footer with the text "©2011-2015 Regents of the University of Colorado. All Rights Reserved." and social media links for GitHub, Facebook, Google+, and Twitter.

# Quantitative Trait Loci

- Definition – area of the genome where polymorphisms are associated with a quantitative trait

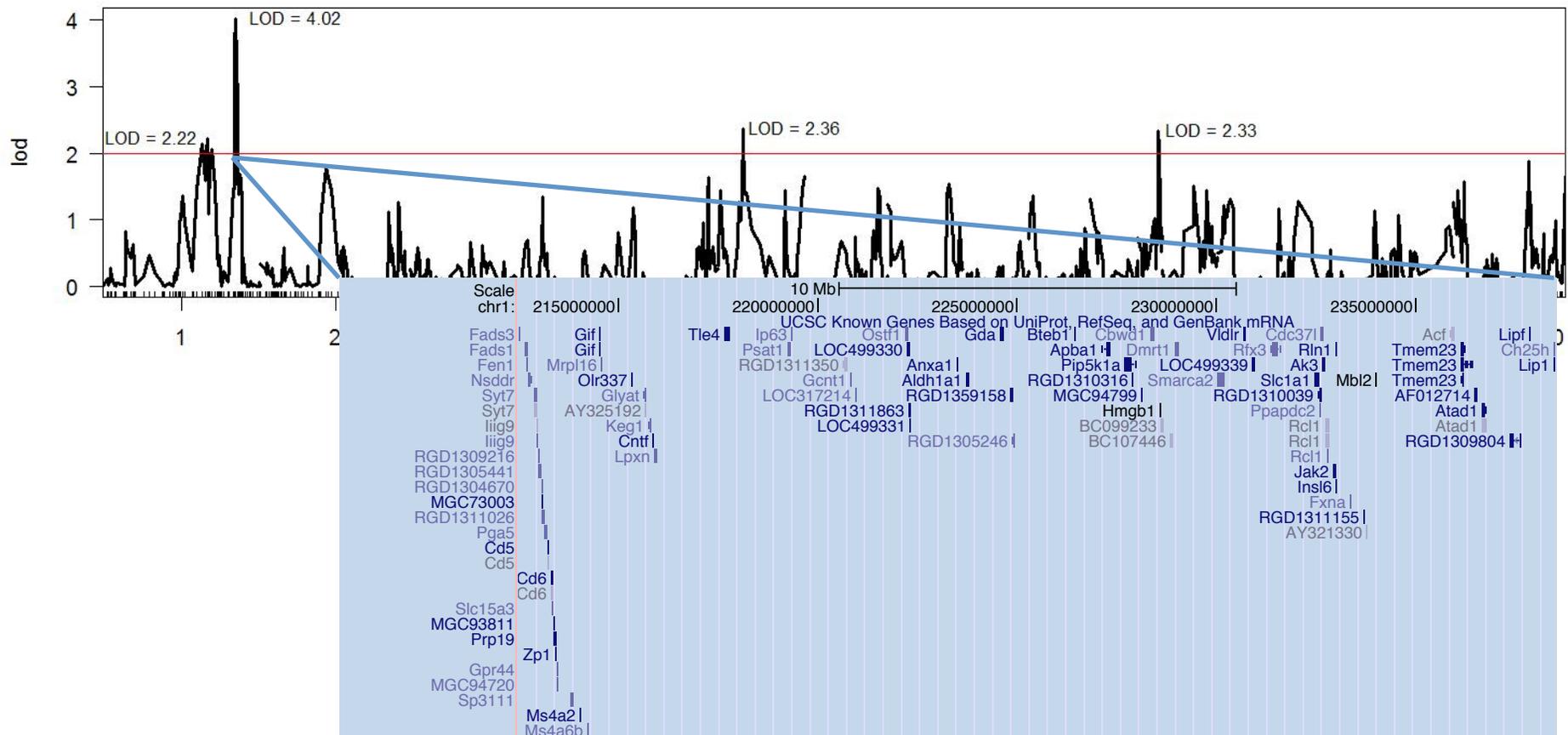


At each marker:

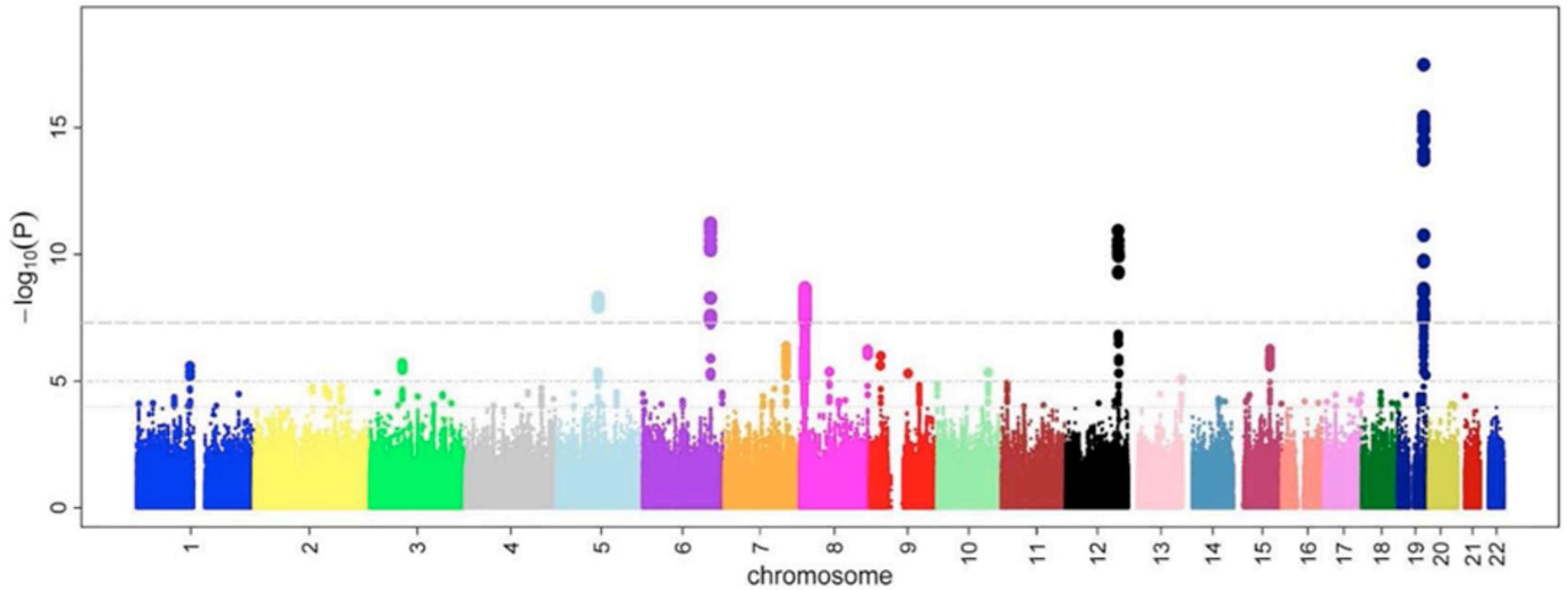
1. Split population into two/three groups based on genotype
2. Compare mean values for the quantitative trait (phenotype) between the groups
3. LARGE Statistically Significant Differences → QTL

# Example LOD Plot

Alcohol Consumption in HXB/BXH Rats - Week 2

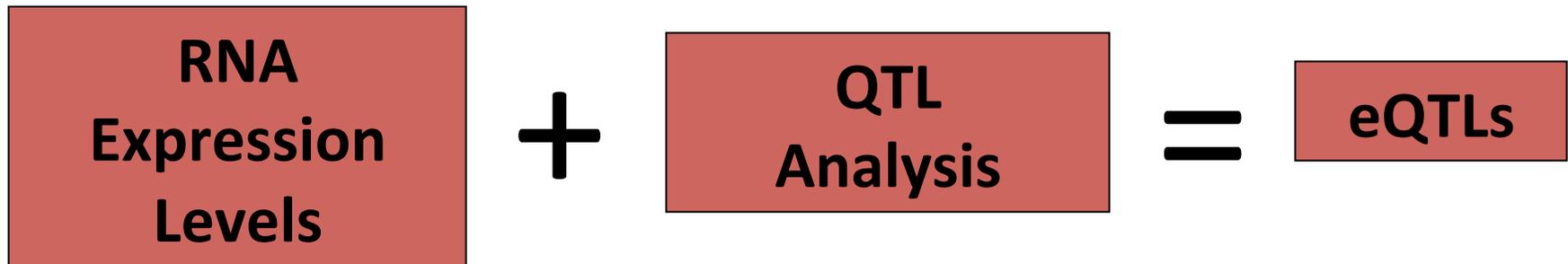


# Example Manhattan Plot

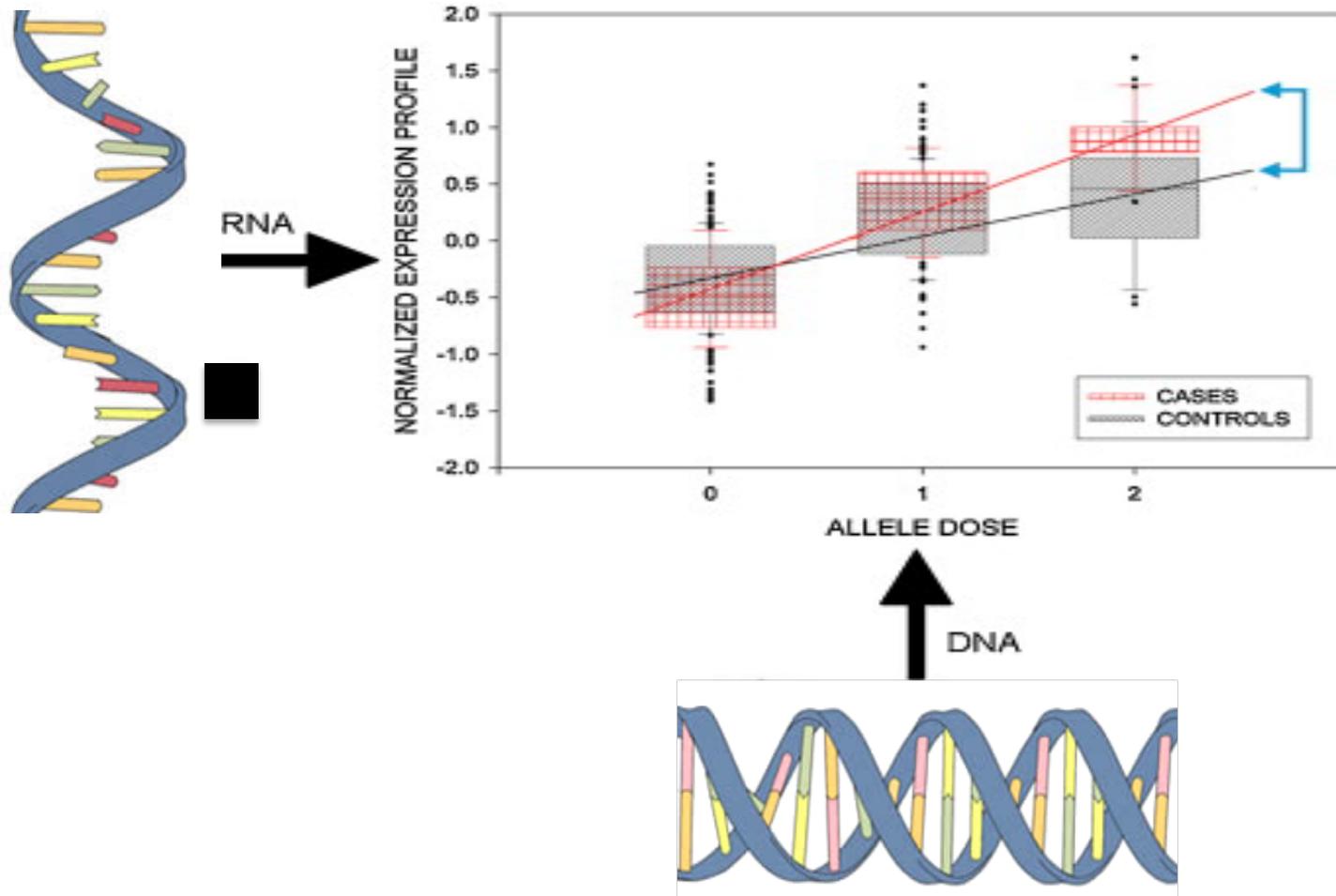


# Definition of Genetical Genomics

The study of the genomic location where transcription is controlled by determining expression Quantitative Trait Loci (eQTL).



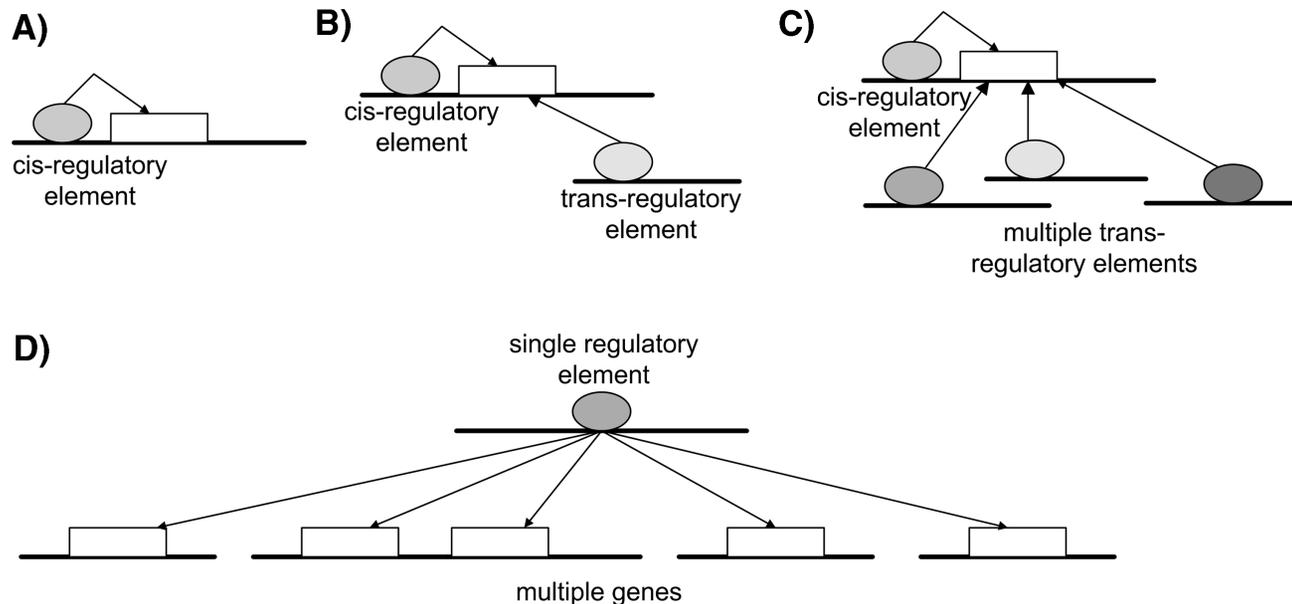
# eQTL Definition



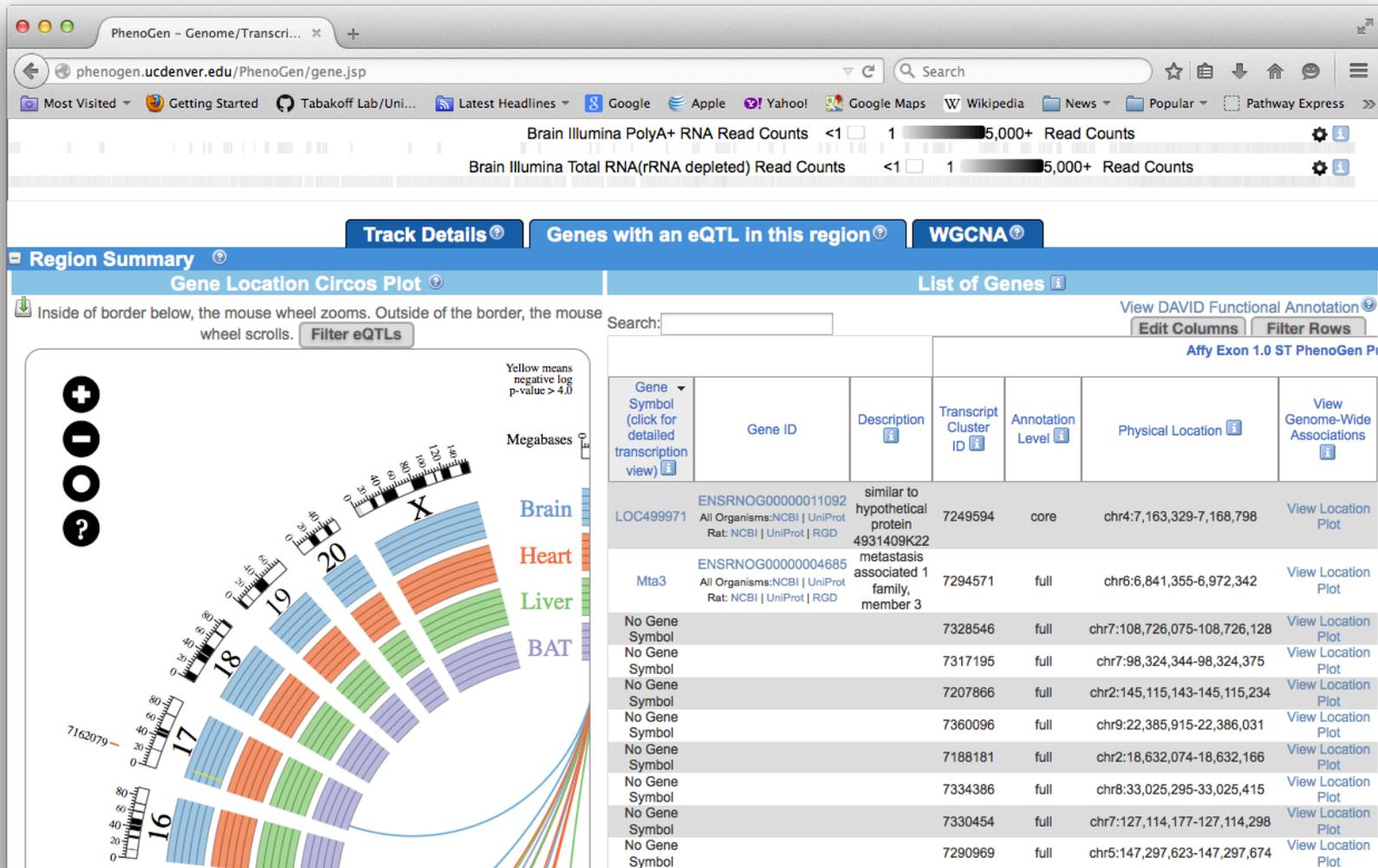
Myers, AJ. The age of the "ome": Genome, transcriptome and proteome data set collection and analysis. Brain Research Bulletin Volume 88, Issue 4 2012 294 - 301

# Types of eQTL

- *cis*-eQTL (or local eQTL) – the locus controlling transcription is “near” the physical location of the gene in the genome
- *trans*-eQTL (or distal eQTL) – the locus controlling the transcription is NOT “near” the physical location of the gene in the genome

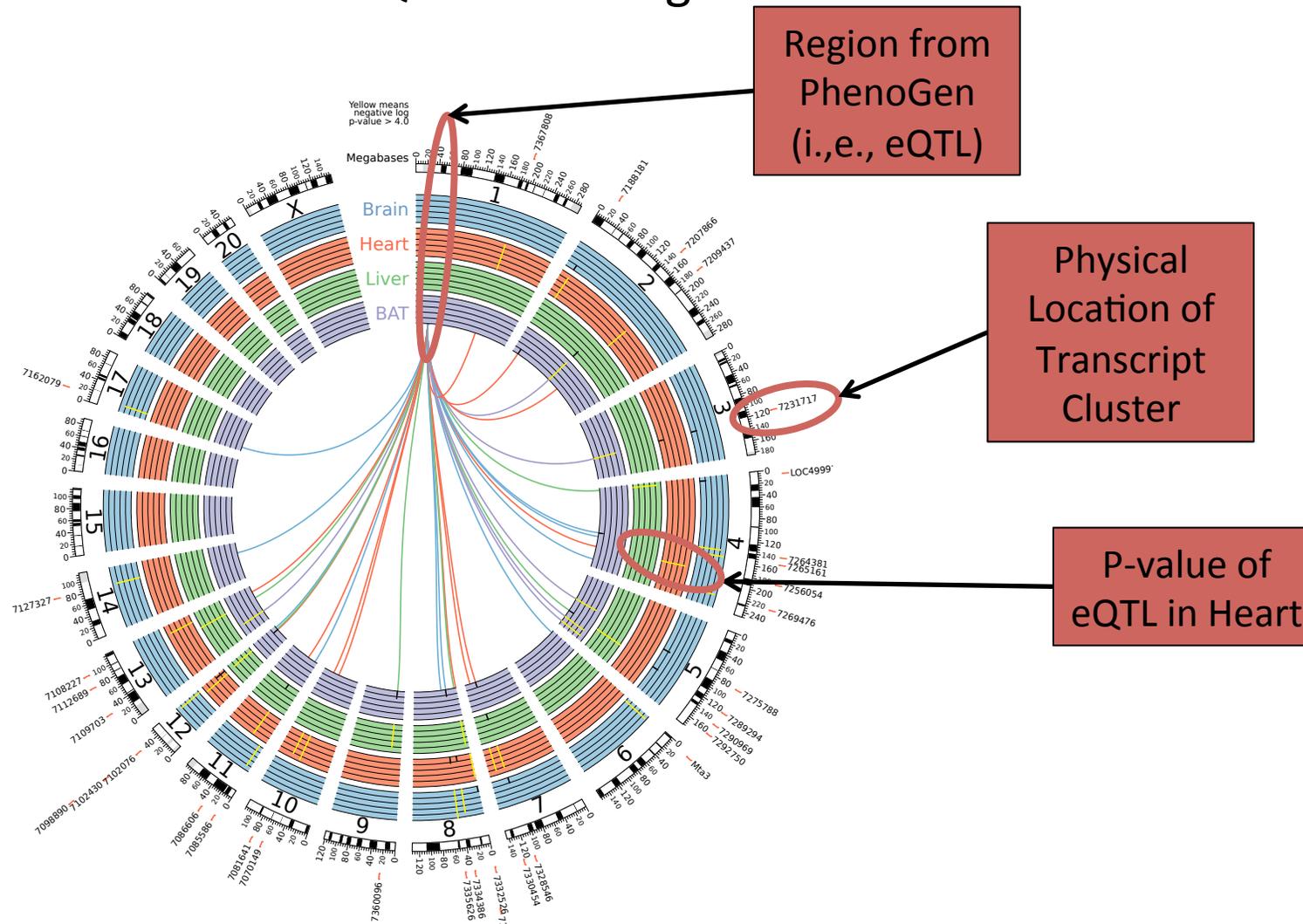


# Genes with an eQTL in this region



# Circos Plot

Genes within an eQTL in this region



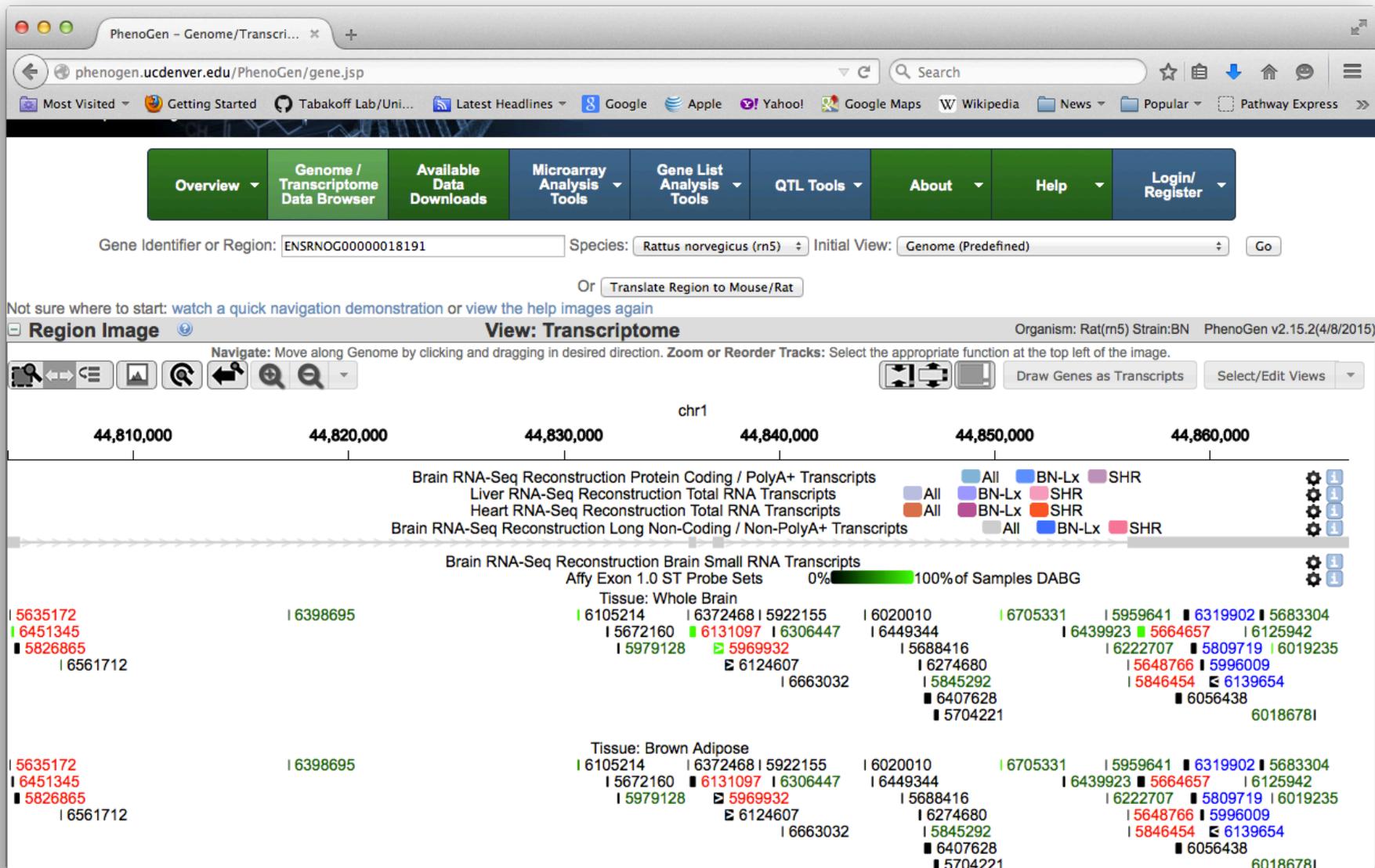
# Transcriptome View

PhenoGen - Genome/Transcri...  
phenogen.ucdenver.edu/PhenoGen/gene.jsp  
Gene Identifier or Region: ENSRNOG00000018191 Species: Rattus norvegicus (rn5) Initial View: Genome (Predefined) Go  
Or Translate Region to Mouse/Rat  
Region Image View: Genome Organism: Rat(m5) Strain:BN PhenoGen v2.15.2(4/8/2015)  
Navigation Hints: Hold mouse over areas of the image for available actions.  
Draw Genes as Transcripts Select/Edit Views

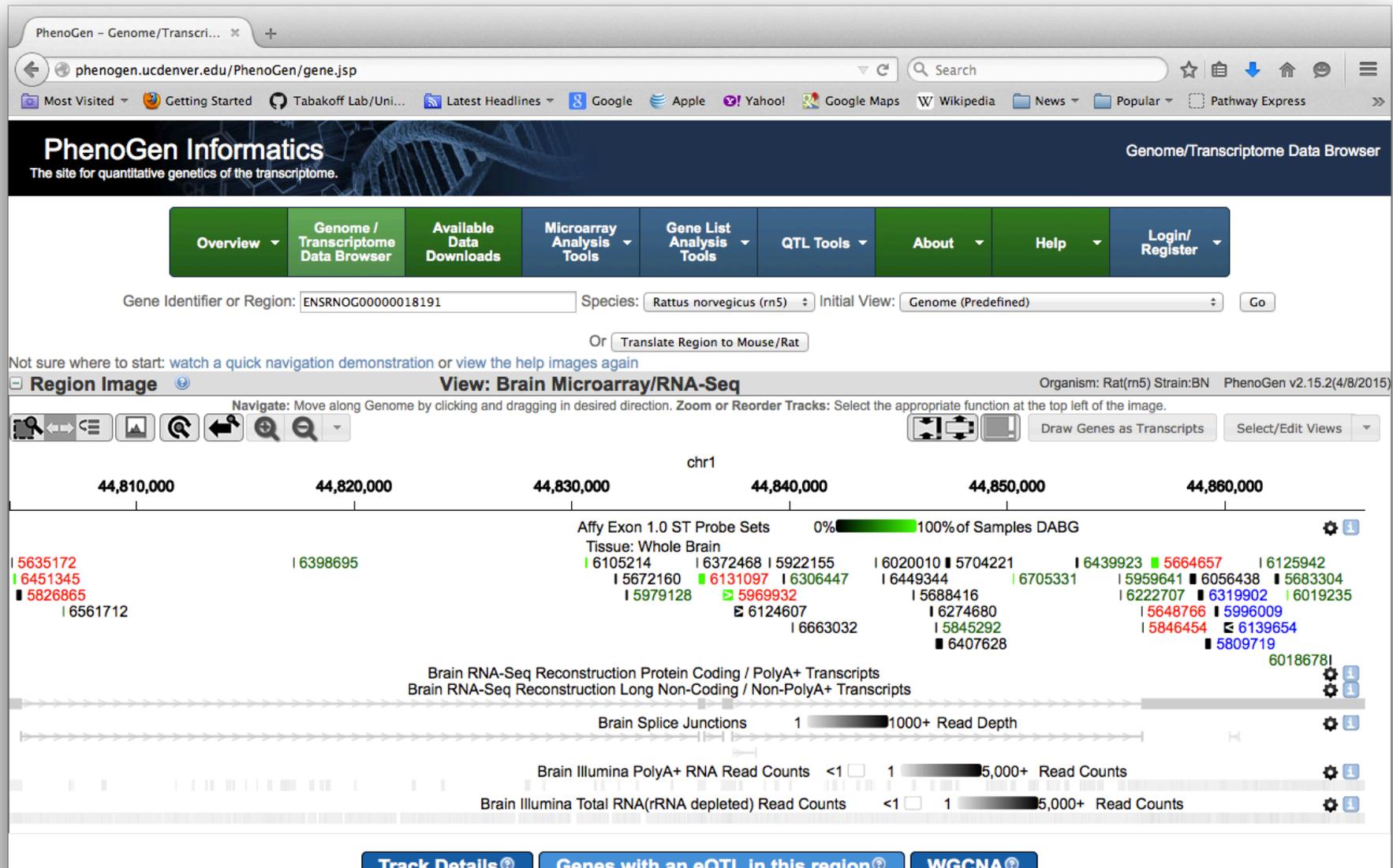
1. Click "Select/Edit Views" box.
2. Select "Transcriptome (Predefined) (6 tracks)" in Select a view below box
3. Click "Apply View" button

Select/Edit Views  
Sign in to see views/tracks not created on this computer.  
Click on a view to select it and view preview/details.  
View types: All Views  
Select a view below (click apply to display the view and return to the browser):  
Genome (Predefined) (8 tracks)  
**Transcriptome (Predefined) (6 tracks)**  
Genome/Transcriptome (Predefined) (14 tracks)  
Liver RNA-Seq (Predefined) (6 tracks) (Rat Only)  
Brain RNA-Seq (Predefined) (8 tracks)  
Liver Microarray/RNA-Seq (Predefined) (7 tracks) (Rat Only)  
Brain Microarray/RNA-Seq (Predefined) (7 tracks)  
Heart RNA-Seq (Predefined) (6 tracks) (Rat Only)  
Description/Preview View/Edit Track List  
Provides general Transcriptome Data including RNA-Seq transcriptomes, Affymetrix probesets, and RNA-Seq read count depth tracks.  
Preview  
chr1  
44,820,000 44,840,000 44,860,000  
Brain RNA-Seq Reconstruction Protein Coding / PolyA+ Transcripts

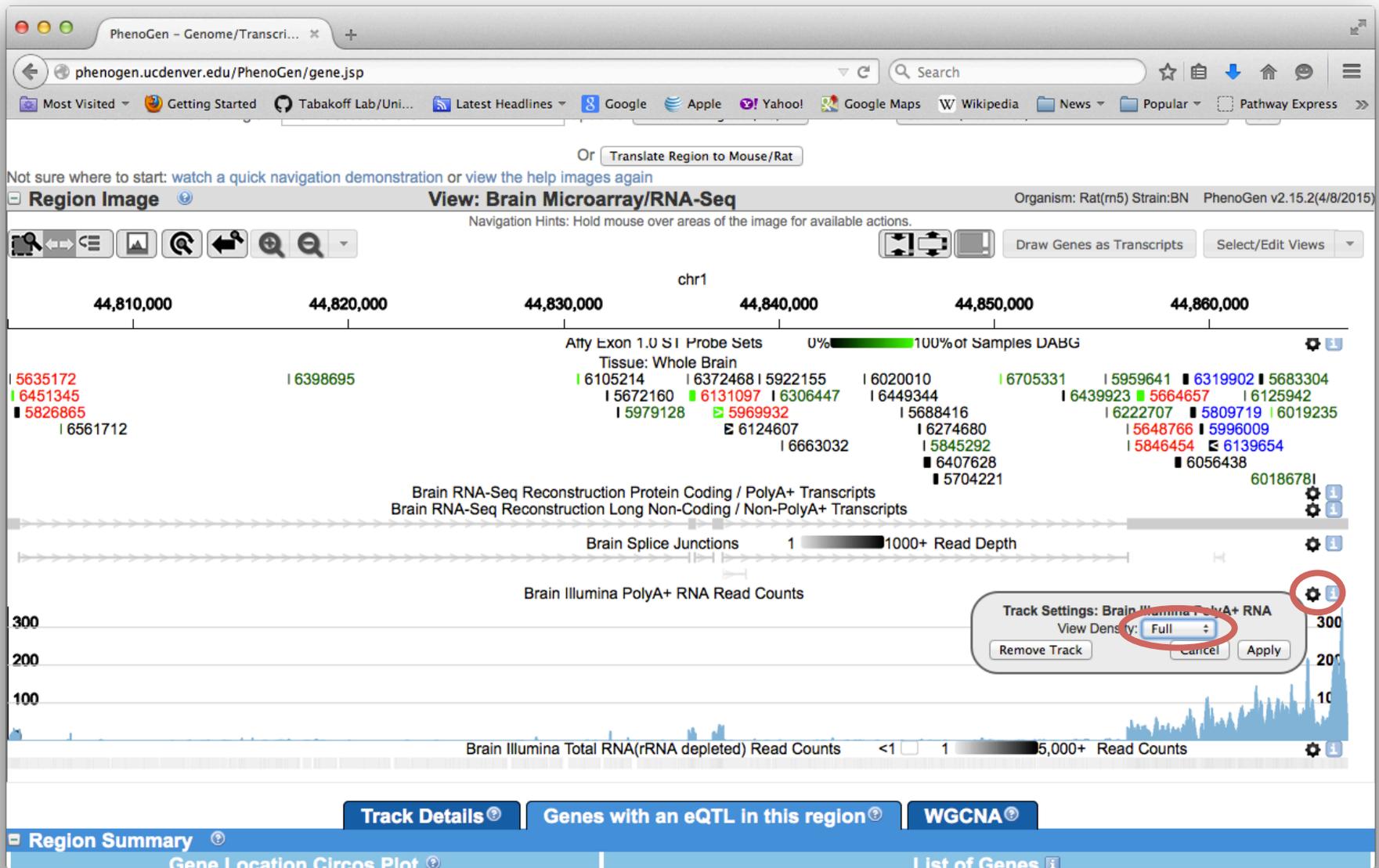
# Oprm1 – Only in Brain Reconstruction



# Brain Microarray/RNA-Seq View



# Expand RNA-Seq Read Counts



# Add Ensembl Track

The screenshot shows the PhenoGen website interface. At the top, the browser address bar displays `phenogen.ucdenver.edu/PhenoGen/gene.jsp`. Below the navigation bar, the search area contains the following information:

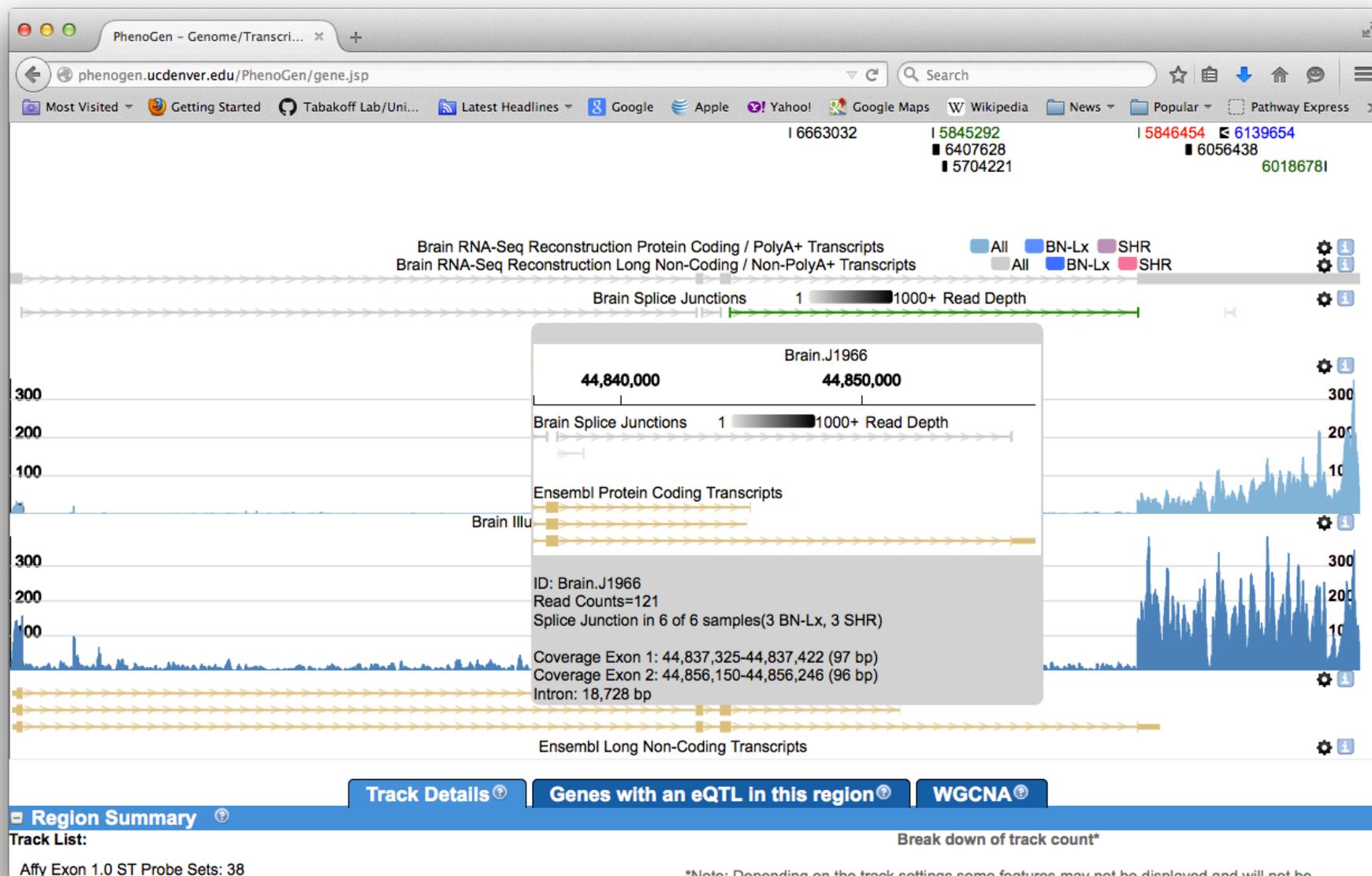
- Gene Identifier or Region: `ENSRNOG00000018191`
- Species: `Rattus norvegicus (rn5)`
- Initial View: `Genome (Predefined)`

The main content area shows the **Region Image** for `View: Brain Microarray/RNA-Seq (Modified)` in the `Organism: Rat(m5) Strain:BN` species. Two modal windows are open:

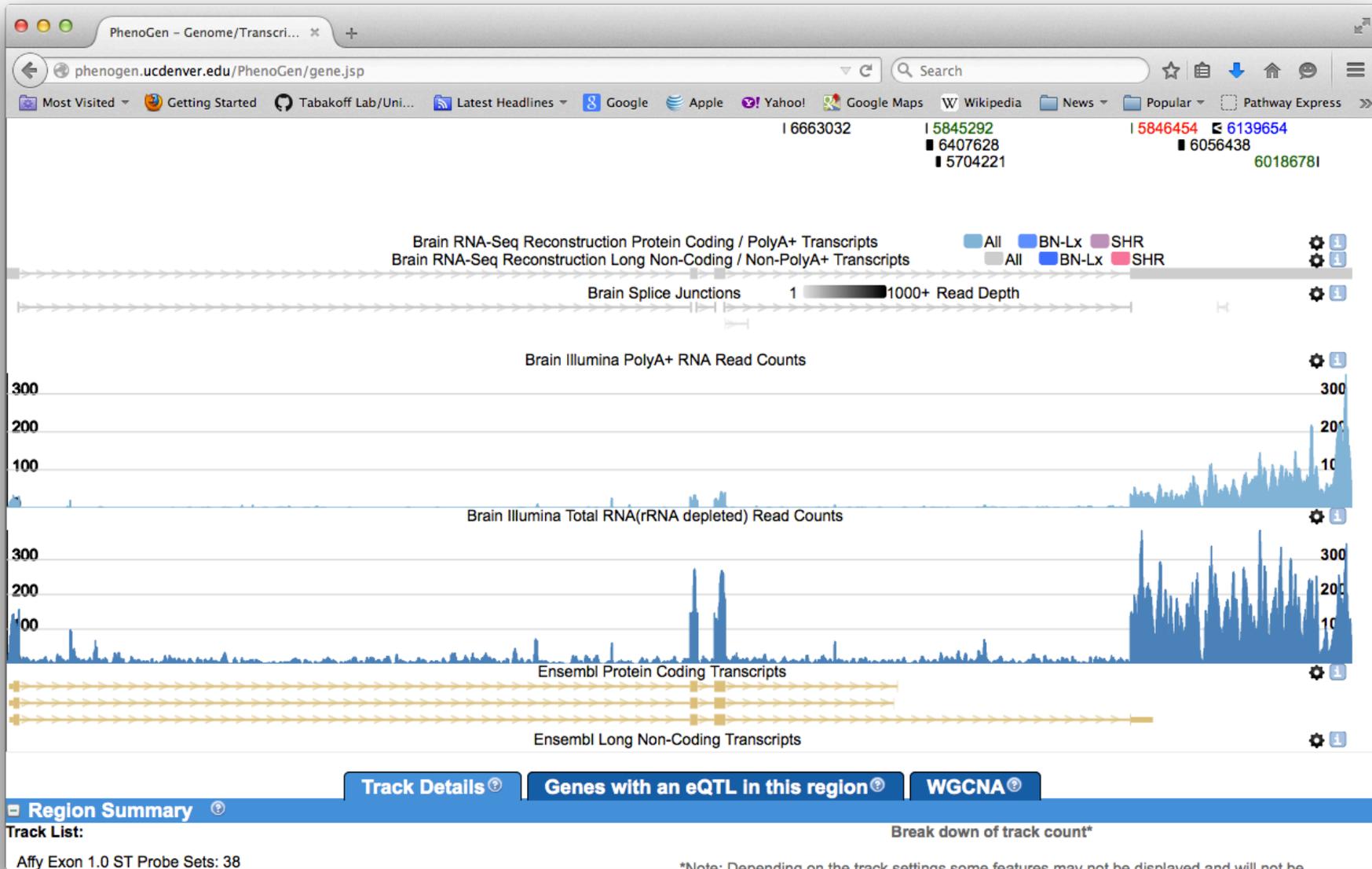
- Select a Track to add to Brain Microarray/RNA-Seq**: This window displays a list of available tracks. A search bar is present, and the tracks are organized by view type (All Tracks). The tracks listed include Ensembl Protein Coding Genes, Ensembl Long Non-Coding Genes, Ensembl Small RNA Genes, Ref Seq Genes, SHR SNPs small Insertion / Deletions, Brain Transcriptome Small RNA, BN-Lx SNPs small Insertion / Deletions, Liver Total-RNA Transcriptome (BN-Lx/SHR), and Liver Splice Junction Support (BN-Lx/SHR).
- Select/Edit Views**: This window shows a list of predefined views. A red circle highlights the **Description/Preview** tab, which contains a green plus icon in a square button. Below this is a table of tracks in the current view:

Order	Track Name	Organism	Edit
1	Reference Genomic Sequence		↑ ↓ ⚙ ×
2	Affymetrix Exon 1.0ST Probes	Rat only	↑ ↓ ⚙ ×
3	Brain Transcriptome Protein Coding Genes	Rat only	↑ ↓ ⚙ ×
4	Brain Transcriptome Long Non-Coding Genes	Rat only	↑ ↓ ⚙ ×
5	Brain Splice Junction Support	Rat only	↑ ↓ ⚙ ×

# Junction Reads



# Alternative 3' UTR



# Selected Feature Summary

PhenoGen - Genome/Transcri... x +

phenogen.ucdenver.edu/PhenoGen/gene.jsp

Most Visited Getting Started Tabakoff Lab/Uni... Latest Headlines Google Apple Yahoo! Google Maps Wikipedia News Popular Pathway Express

**Gene Details** **Gene eQTLs** **Probe Set Level Data** **WGCNA**

**Selected Feature Summary**

Gene Symbol: Oprm1  
 Location: chr1: 44,804,261-44,857,206  
 Strand: +  
 Description: opioid receptor, mu 1  
 Links: [ENSRNOG00000018191](#)  
 All Organisms: [NCBI](#) | [UniProt](#)  
 Rat: [NCBI](#) | [UniProt](#) | [RGD](#)

Exonic Variants: **Common:** 0 (SNPs) / 0(Insertions/Deletions)  
**BN-Lx/CubPrin:** 0 (SNPs) / 0(Insertions/Deletions)  
**SHR/OlaPrin:** 4 (SNPs) / 7 (Insertions/Deletions)  
**SHR/NCrPrin:** 4 (SNPs) / 7 (Insertions/Deletions)  
**F344:** 4 (SNPs) / 6 (Insertions/Deletions)

Transcripts:  
**ENSRNOT00000024682**  
**ENSRNOT00000045144**  
**ENSRNOT00000051837**  
**Brain\_T18812** - Transcript Match: ENSRNOT00000051837 2 Perfect Exon Matches,3' Extended, 5' Extended,

**Affy Probe Set Data: Overlapping Probe Set Count:23**

**Probe sets detected above background\*:**

Tissue	Number of probe sets detected above background* in more than 1% of samples (out of 23 probe sets for this gene)	Avg % of samples DABG*	Range
Brain	12	57 %	2.78 - 100 %
Heart	3	20 %	1.9 - 55.24 %
Liver	2	84 %	72.64 - 96.23 %
Brown Adipose	5	18 %	2.08 - 56.25 %

\*DABG is based on Affymetrix software that assigns a P-value to the probe sets detection above background. Using a comparison of RNA-Seq data probe sets that overlap a high confidence exon in the transcriptome are not detected above background roughly 5% of the time. Increasing the P-value cutoff of 0.0001 can reduce this but only at the expense of greatly elevated false positives.

**Probe Set Heritability:**

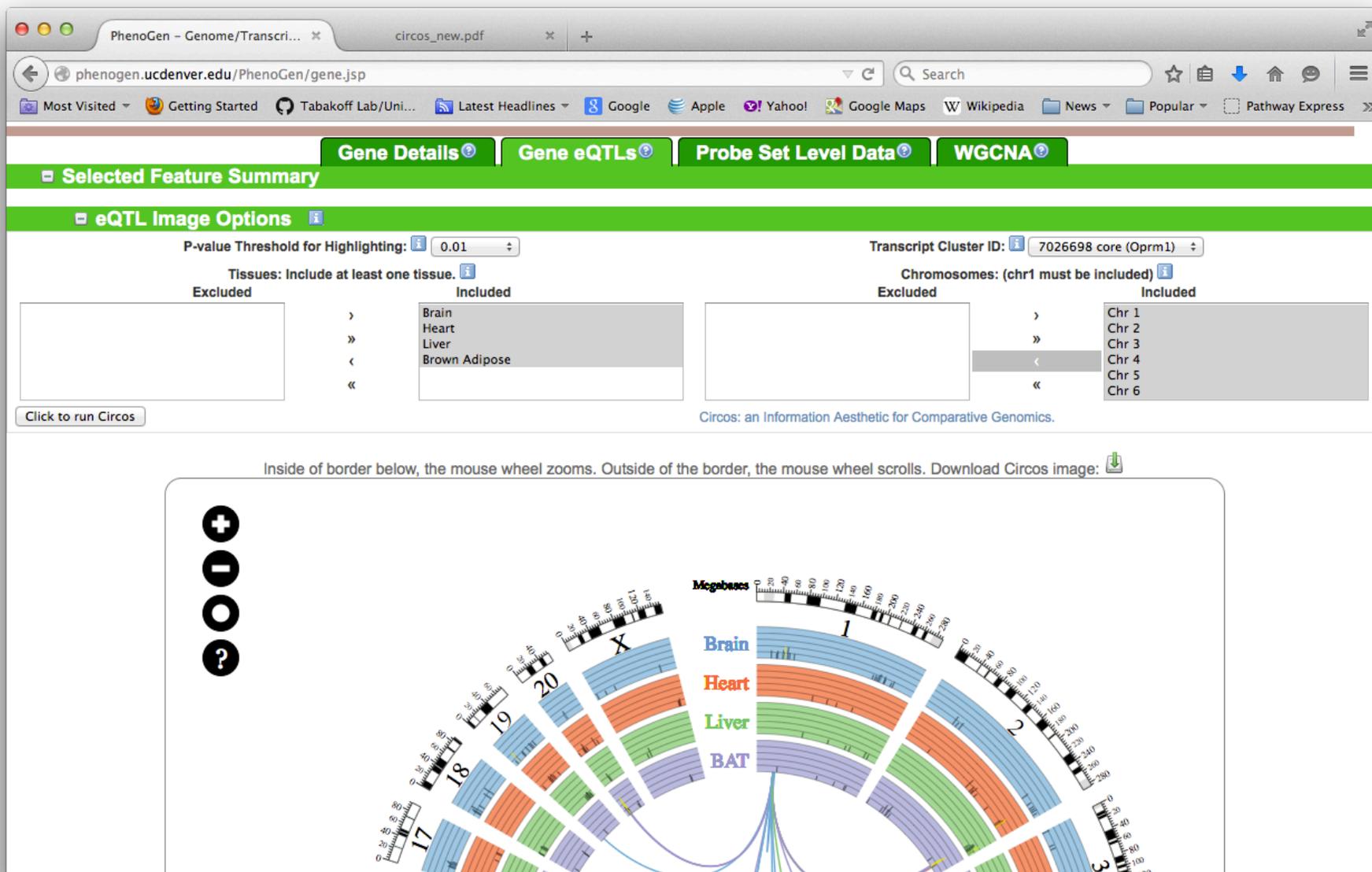
Tissue	Number of probe sets with a heritability greater than 0.33 (out of 23 probe sets for this gene)	Avg Herit	Range
Brain	5	0.36	0.33 - 0.4
Heart	9	0.39	0.33 - 0.47
Liver	11	0.38	0.34 - 0.44
Brown Adipose	6	0.41	0.34 - 0.5

**EQTLs Affymetrix Transcript Cluster(Confidence Level): 7026698 (core)**

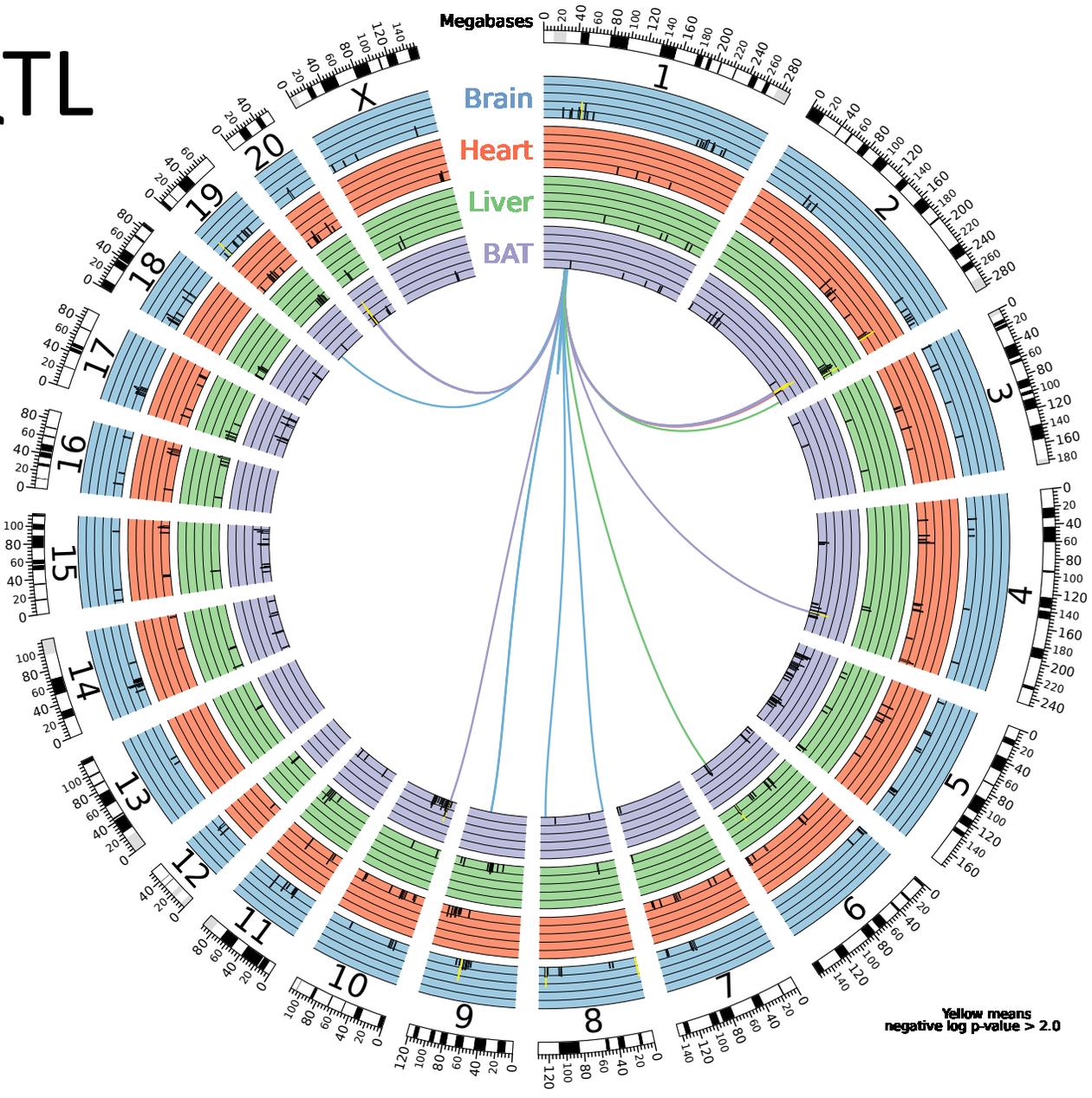
Tissue	Number of eQTLs	Minimum P-value EQTL	
		P-value	Location
Whole Brain	6	0.0024	<a href="#">chr9:76,343,169</a>
Heart	1	0.0096	<a href="#">chr2:259,028,982</a>
Liver	2	0.0052	<a href="#">chr6:119,885,523</a>
Brown Adipose	8	0.0006	<a href="#">chr20:12,052,918</a>

Click the Gene eQTLs Tab above to view a Circos Plot of the eQTLs listed above.

# Gene eQTL



# Gene eQTL



# **WEIGHTED GENE CO-EXPRESSION NETWORK ANALYSIS (WGCNA)**

What are we missing by considering each candidate gene individually?

- Define relationships among genes
- Infer biological function from other co-expressed genes
- Define the context in which the gene exerts its effect
- Find multiple therapeutic targets within the same pathway

What do we gain by building networks and identifying modules?

- No gene product acts independently in the cell
- Information about biological function in cell
- Information on causes/ consequences of differential expression

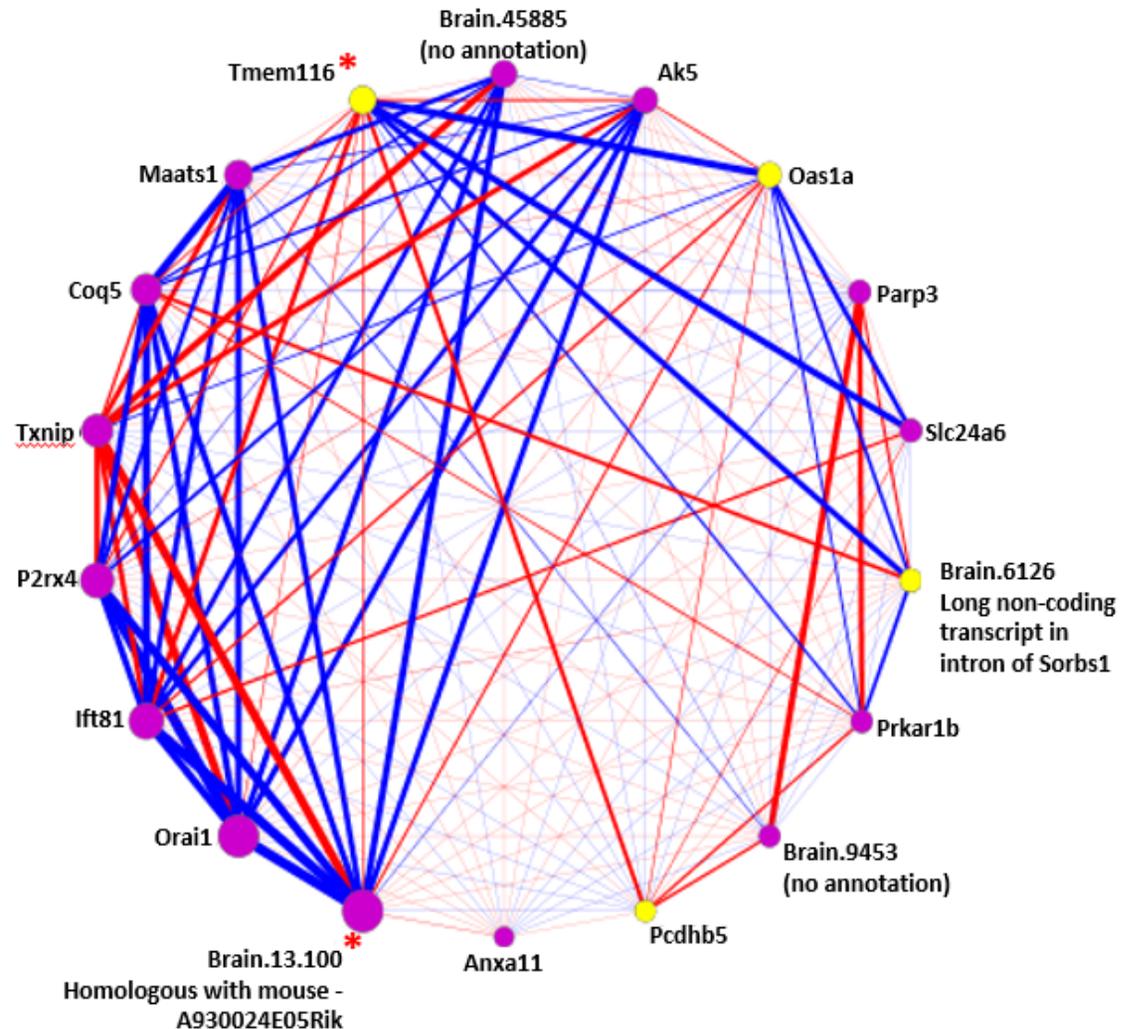
# Weighted Gene Co-Expression Network Analysis

## Why Not Just Use Correlation?

1. Simple correlation does not give connectivity.
2. How are we measuring co-expression?
  - Scale-Free Network
    - Network has few highly connected genes rather than each gene have similar connectivity
    - **Biologically motivated**, fewer highly connected genes means that a system is more robust to failure of any one gene
3. How do we get a **robust** measure of connectivity for identifying modules?
  - Topological Overlap Measure
    - Includes a measure of how many “friends” two genes have in common
    - Protects against spurious correlations among genes

# Example Co-Expression Module

- Edge thickness is weighted based on magnitude of correlation between nodes.
- Blue edges represent a positive correlation between nodes.
- Red edges represent a negative correlation between nodes.
- Node size is weighted based on connectivity within module.



# Summary of Expression Pattern Within Co-Expression Module

- Eigengene
  - First Principal Component
    - Maximize the amount of variance in expression captured by a single value per strain/sample
- Hub Gene
  - Most ‘connected’ gene within a module
  - May have biological implications

# WGCNA

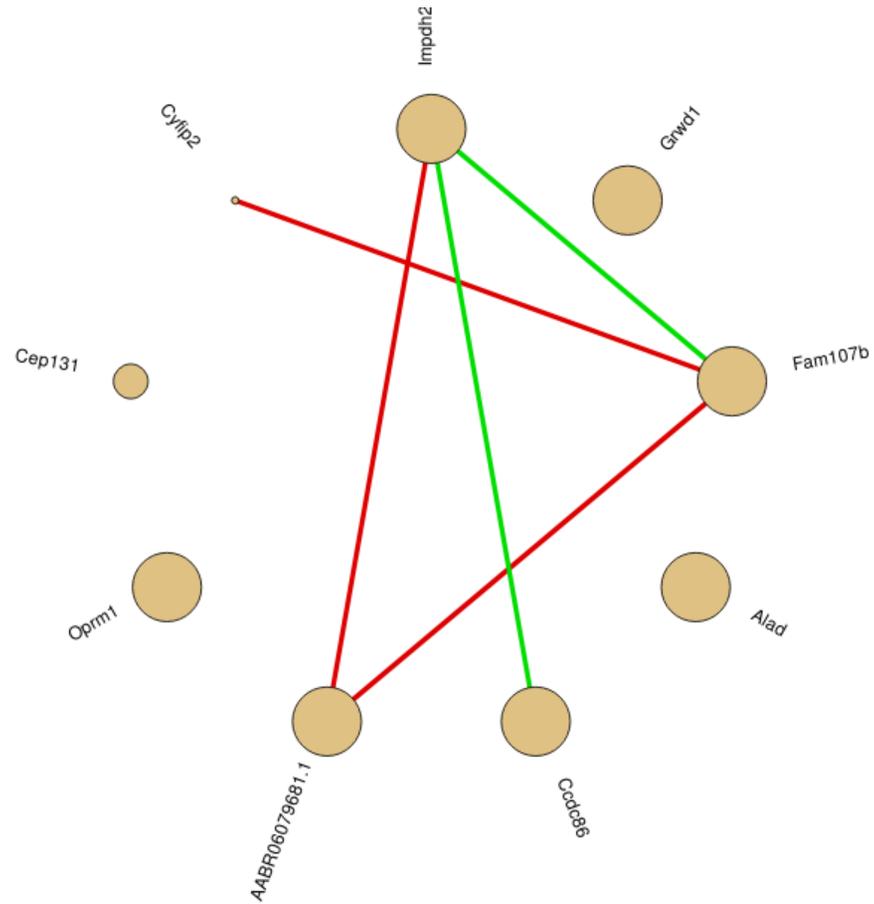
The screenshot shows a web browser window with the URL `phenogen.ucdenver.edu/PhenoGen/gene.jsp`. The page features a navigation bar with tabs for **Gene Details**, **Gene eQTLs**, **Probe Set Level Data**, and **WGCNA**. Below the navigation bar is a green header for the **Selected Feature Summary**. The main content area includes a **Tissue: Whole Brain** dropdown and a **Views:** section with radio buttons for **Module** (selected), **Eigengene eQTL**, and **Gene Ontology**. A tooltip is displayed over a small green circle labeled '1', showing the following information: **Module Name: azure**, **Transcript Size: 9**, and **Selected Gene Transcripts in Module: 1**. To the left of the tooltip is a large blue circle labeled '1' with the text 'blue' below it. Below the main content area, the footer contains the text **©2011-2015 Regents of the University of Colorado. All Rights Reserved.** and social media links for **Source Code (GitHub)**, **Legal Notices**, **Privacy Policy**, **Follow** (Facebook), **Follow** (Google+), and **Follow** (Twitter).

# Azure Module

azure

Legend

- Positive Correlation
- Negative Correlation
- Lower Correlation
- Higher Correlation



# Azure Module Table

PhenoGen - Genome/Transcri... x    circos\_new.pdf x +

phenogen.ucdenver.edu/PhenoGen/gene.jsp    Search

Most Visited   Getting Started   Tabakoff Lab/Uni...   Latest Headlines   Google   Apple   Yahoo!   Google Maps   Wikipedia   News   Popular   Pathway Express

**Transcripts in azure Module**    [Export CSV](#)

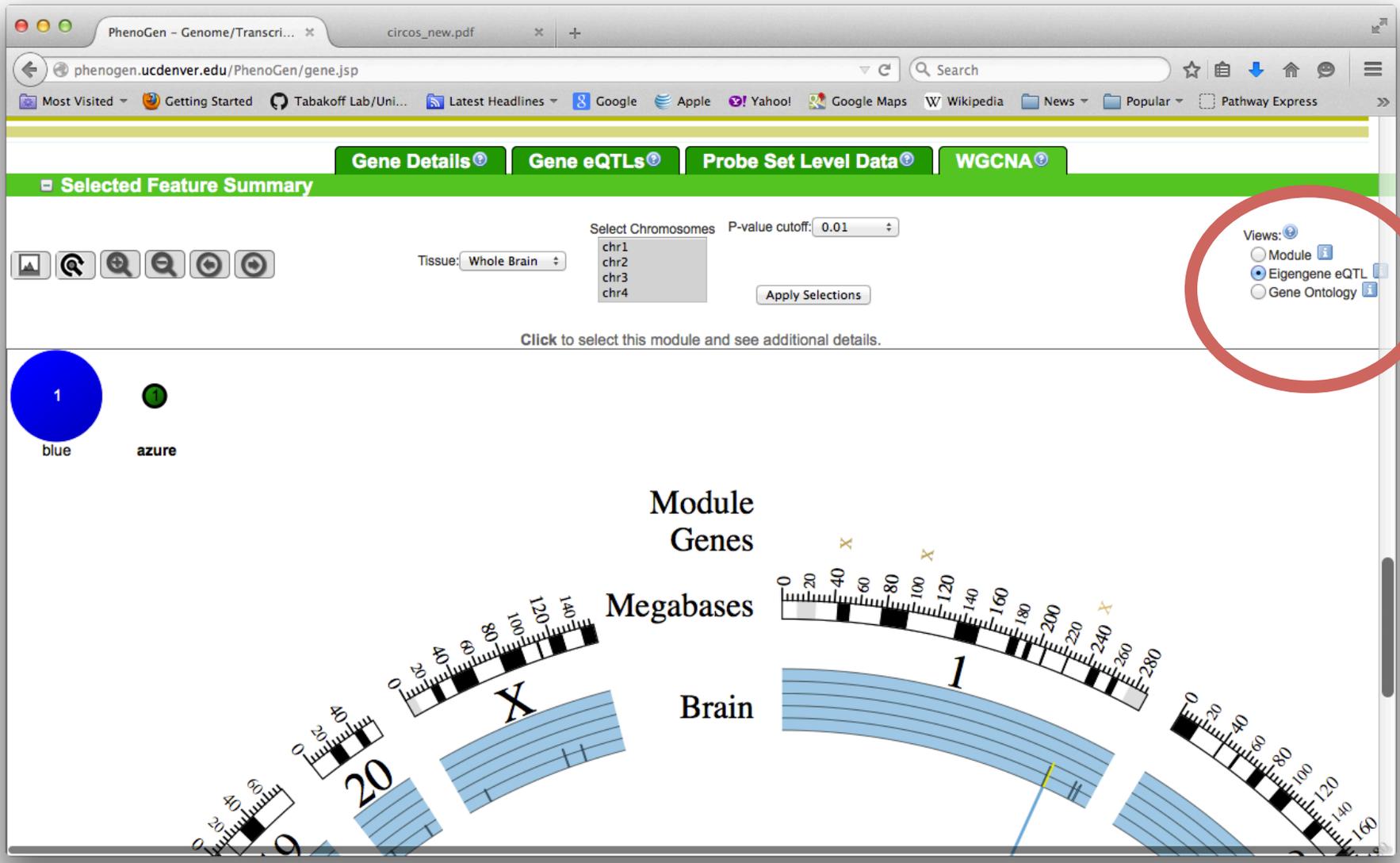
Search:

Showing 1 to 9 of 9 entries

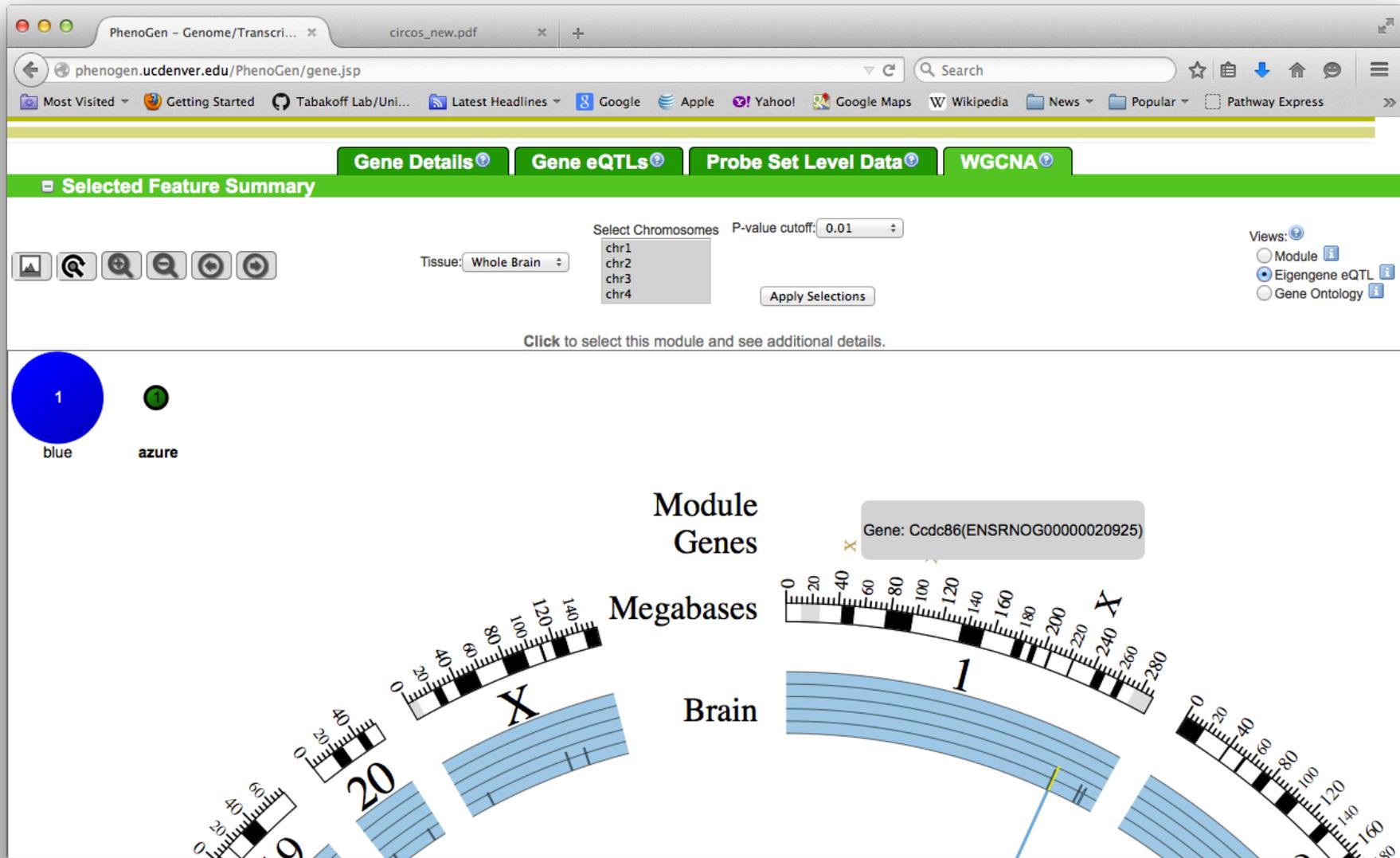
Gene Symbol	Gene ID	Transcript	Probe Sets	Link Total	Link Rank
Impdh2	ENSRNOG00000031965	Brain_C8804.3	1	4.90	1
Grwd1	ENSRNOG00000021058	Brain_C9091.1	1	3.48	2
Fam107b	ENSRNOG00000014886	Brain_C6674.1	5	3.21	3
Alad	ENSRNOG00000015206	Brain_C6729.2	1	3.14	4
Ccdc86	ENSRNOG00000020925	Brain_C4185.1	1	2.48	5
AABR06079681.1	ENSRNOG00000036375	Brain_C4097.4	1	1.41	6
Oprm1	ENSRNOG00000018191	Brain_C29606.1	1	1.29	7
Cep131	ENSRNOG00000004430	Brain_C1276.2	1	0.55	8
Cyfp2	ENSRNOG00000006557	Brain_C1260.2	2	0.00	9

©2011-2015 Regents of the University of Colorado. All Rights Reserved.    [Source Code \(GitHub\)](#)   [Legal Notices](#)   [Privacy Policy](#)   [Follow](#)   [Follow](#)   [Follow](#)

# Eigengene eQTL

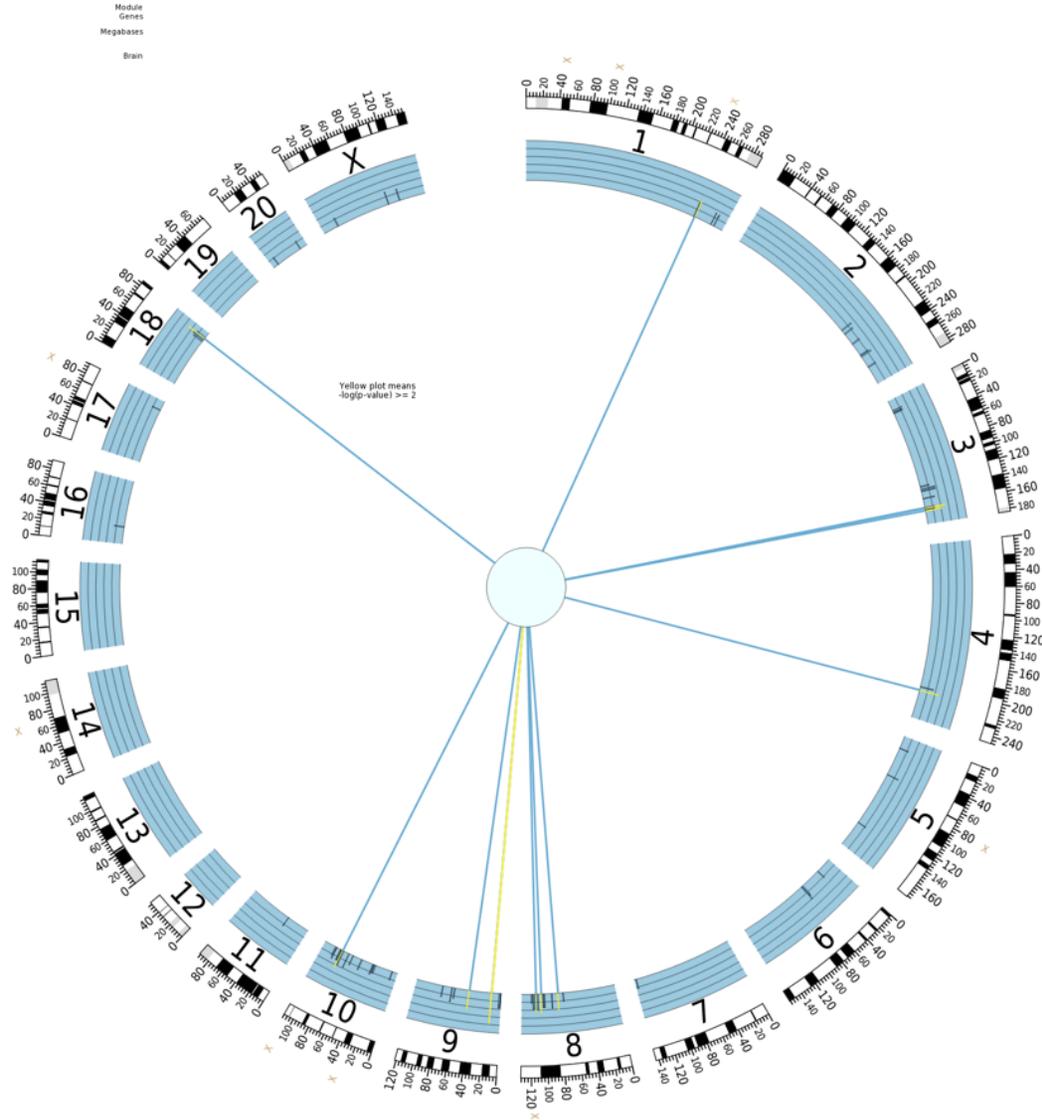


# Physical Location of Genes with Module



# Circos

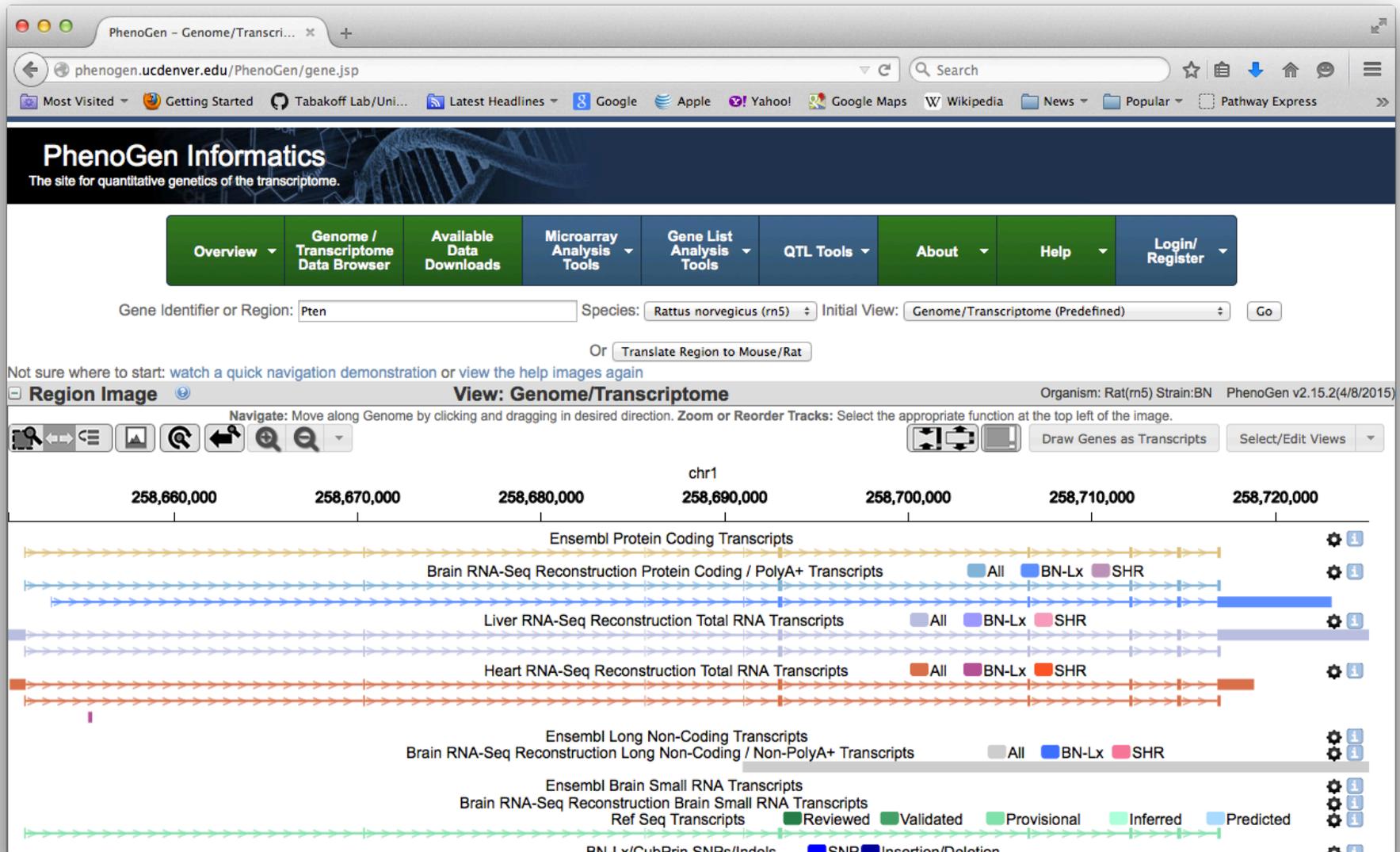
## Eigengene eQTL



# Gene Ontology

The screenshot shows a web browser window with the URL `phenogen.ucdenver.edu/PhenoGen/gene.jsp`. The page features a navigation bar with tabs for **Gene Details**, **Gene eQTLs**, **Probe Set Level Data**, and **WGCNA**. Below this is a **Selected Feature Summary** section with controls for **Tissue** (Whole Brain), **GO Domain** (Molecular Function), and **Number of Levels to Display** (3). A **Views** menu is circled in red, showing options for **Module**, **Eigengene eQTL**, and **Gene Ontology** (selected). Below the navigation hints, there are two circular icons labeled **blue** and **azure**. The text **7 genes in module with GO annotation** is displayed above a semi-circular sunburst chart. The chart segments represent various GO terms, including **signaling receptor activity**, **receptor activity**, **oxidoreductase activity**, **transmembrane transporter activity**, **transporter activity**, **substrate transporter activity**, and **catalytic activity**.

# RNA Expression Across Tissues



# Probe Set Level Data

Probe Set Level Data needs to collect and parse lots of data.

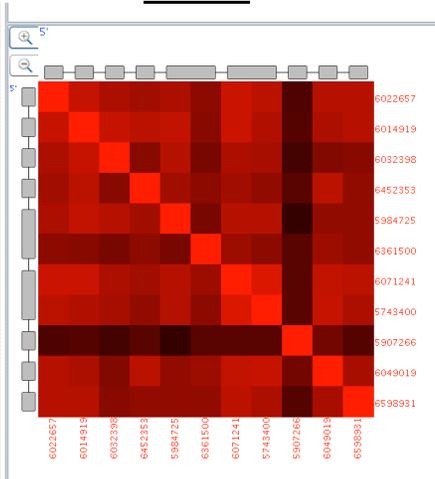
Click on the tab, and wait a minute before pushing the View Affy Probe Set Details button

The screenshot shows a web browser window with the URL `phenogen.ucdenver.edu/PhenoGen/gene.jsp`. The browser's address bar and search bar are visible. Below the browser window, a series of horizontal bars in various colors (pink, grey, yellow, cyan, blue, orange, green, red) represent data rows. At the bottom of the page, there is a navigation bar with four tabs: **Gene Details**, **Gene eQTLs**, **Probe Set Level Data** (which is selected), and **WGCNA**. Below the navigation bar, a green box contains the text: **Selected Feature Summary**. Below this, a message states: "This feature requires Java which will open in a separate window, when you click the button below. Java will be automatically detected and directions will be displayed on the next page if there are any issues to correct before proceeding." A button labeled **View Affy Probe Set Details** is positioned below the message. At the very bottom of the page, there is a footer with copyright information: "©2011-2015 Regents of the University of Colorado. All Rights Reserved." and social media links for GitHub, Facebook, Google+, and Twitter.

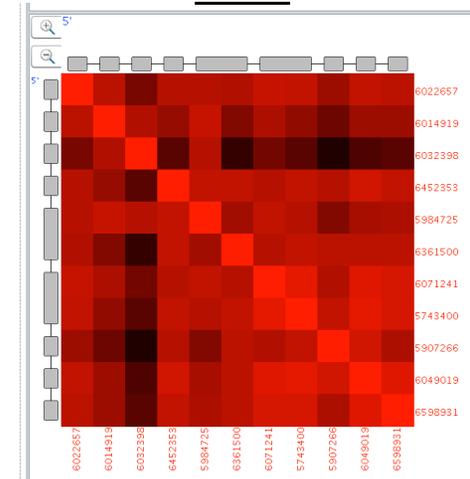


# Panel Exon Correlation In Different Tissues

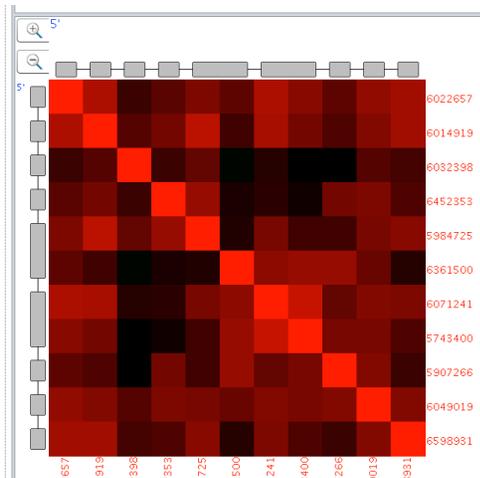
**Brain**



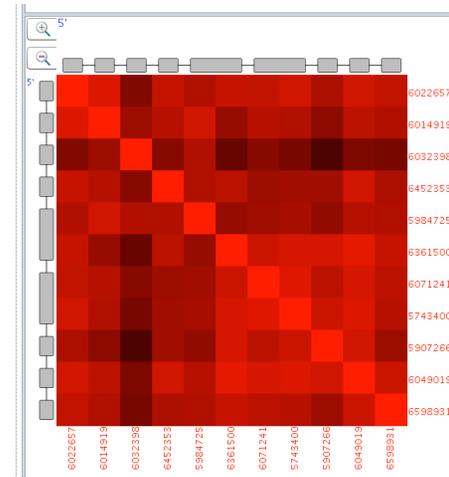
**Liver**



**Heart**



**Brown Adipose**



# **FROM PHENOTYPE TO CANDIDATE GENES**

# Genetical Genomics/Phenomics Approach

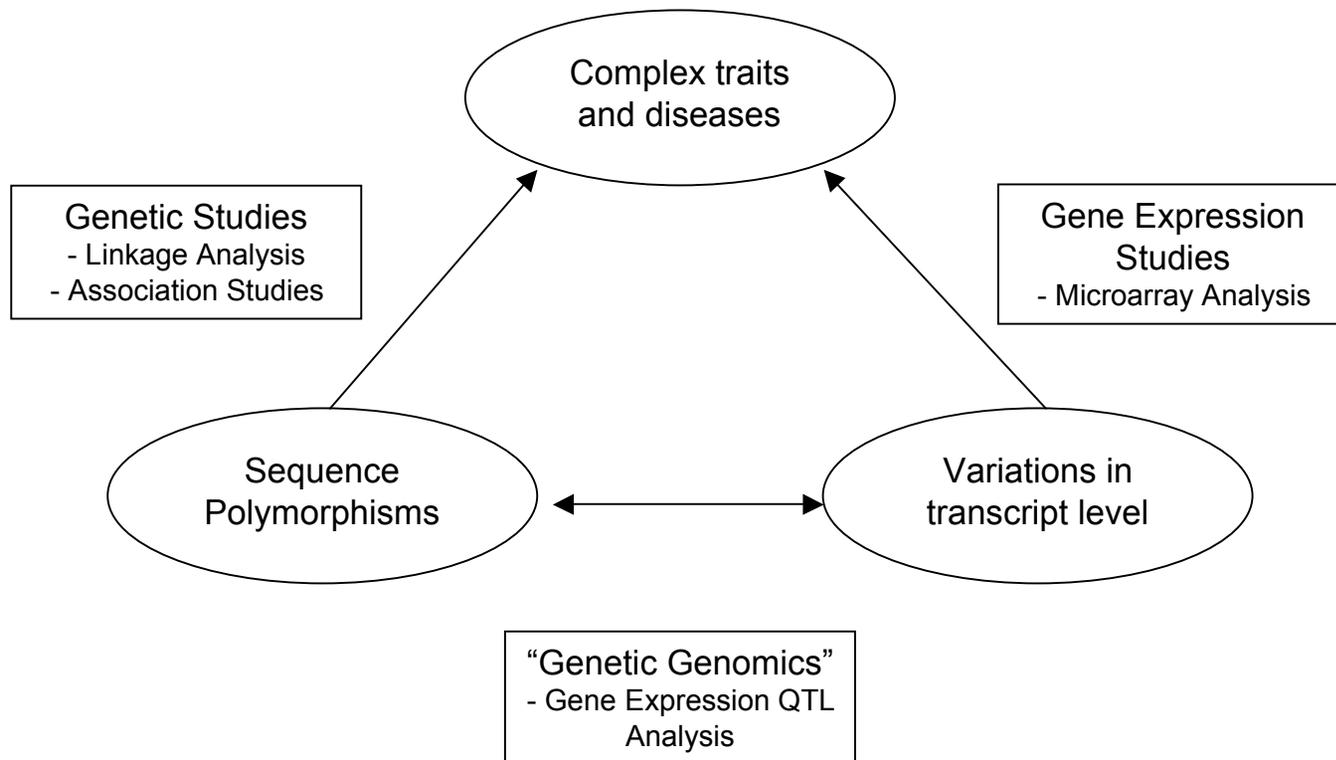


Image copied from "The Marriage of Phenomics and Genetical Genomics: A Systems Approach to Complex Trait Analysis" in *Systems Biology in Psychiatric Research: From High-Throughput Data to Mathematical Modeling*, edited by Tetter F, Winterer G, Gebicke-Haerter PG, and Mendoza E. Wiley-VCH 2010.

# Data

- What you provide
  - Quantitative phenotype data from one of the following panels:
    - BXD RI mice
    - LXS RI mice
    - HXB/BXH RI rats
- What we provide
  - RNA expression levels
  - SNP information

# Genetic Correlation

## RNA Expression Levels Correlated with Phenotype

The screenshot shows the PhenoGen Informatics website interface. The browser address bar displays the URL: `phenogen.ucdenver.edu/PhenoGen/web/datasets/correlation.jsp?datasetID=707&datasetVersion=4&analysisType=correlation`. The page header includes the site name "PhenoGen Informatics" and the tagline "The site for quantitative genetics of the transcriptome." A navigation menu contains links for Home, Genome / Transcriptome Data Browser, Available Data Downloads, Microarray Analysis Tools, Gene List Analysis Tools, QTL Tools, About, Help, and Account.

The main content area displays the current analysis: "You are Analyzing: Public ILSXISS RI Mice v4". Below this, a "Steps to run a correlation analysis:" workflow is shown with icons for: Choose Dataset, Choose Dataset Version, Choose Type of Analysis, Choose Phenotype Data (highlighted), Filter Probe(set)s, Run Statistical Test, Correct for Multiple Testing, and Save Gene List. To the right of the workflow are links for "Dataset Version Details" and "Create New Phenotype".

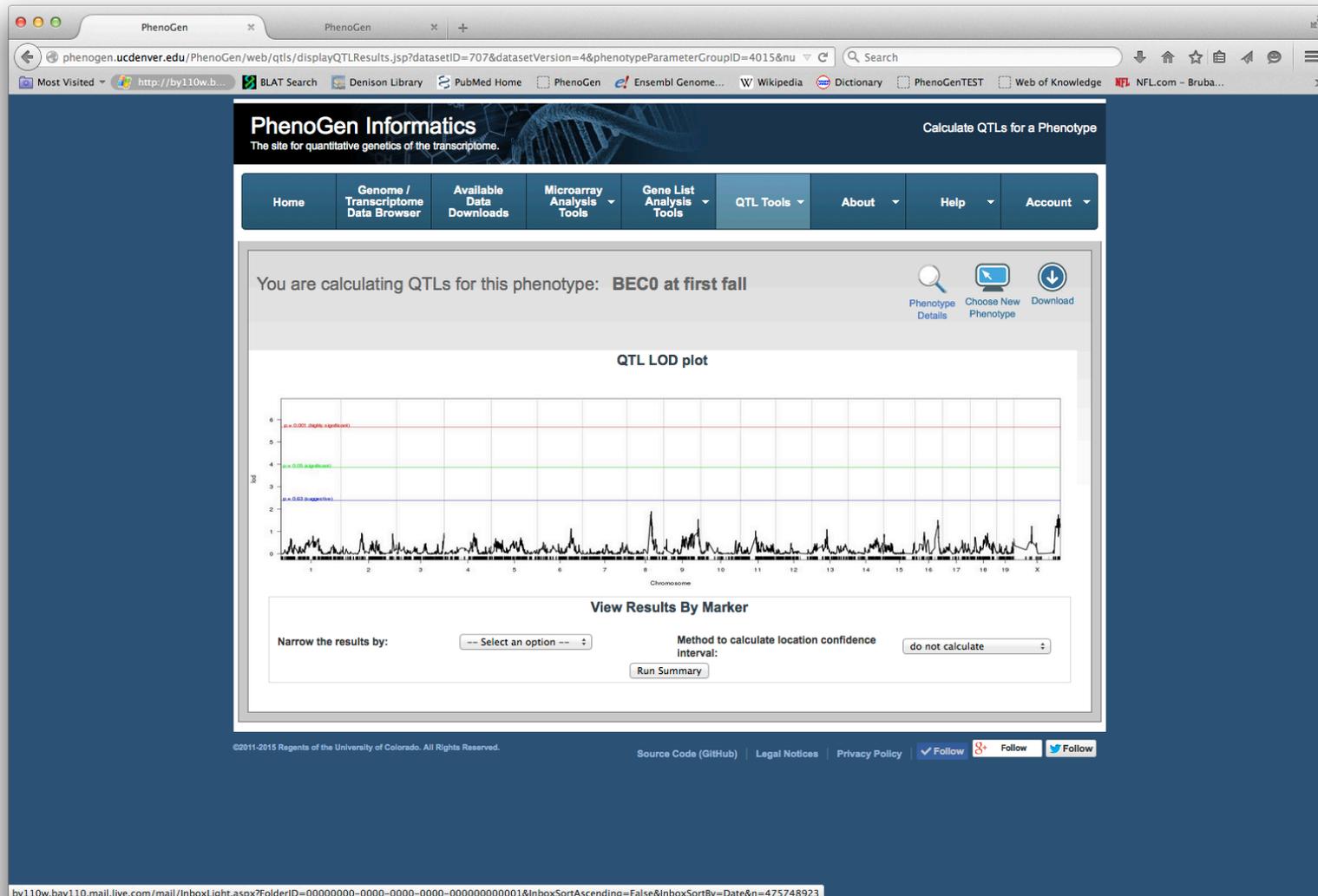
Below the workflow, a text prompt says: "Click on the phenotype data you would like to use, or enter new phenotype data." This is followed by a section titled "Phenotype Values (Matching 5 or more strains)" which contains a table with the following data:

Phenotype Name	Description	Details	Delete	Download
BECO at first fall	from GeneNetwork	<a href="#">View</a>	✗	
Longevity	longevity in days for male mice from MPD	<a href="#">View</a>	✗	
LORR from GeneNetwork	loss of righting reflex	<a href="#">View</a>	✗	
ME1130	eigengene for Mecp2 module	<a href="#">View</a>	✗	

At the bottom of the page, there is a footer with copyright information: "©2011-2015 Regents of the University of Colorado. All Rights Reserved." and social media links for GitHub, Legal Notices, Privacy Policy, and Follow buttons for various platforms.

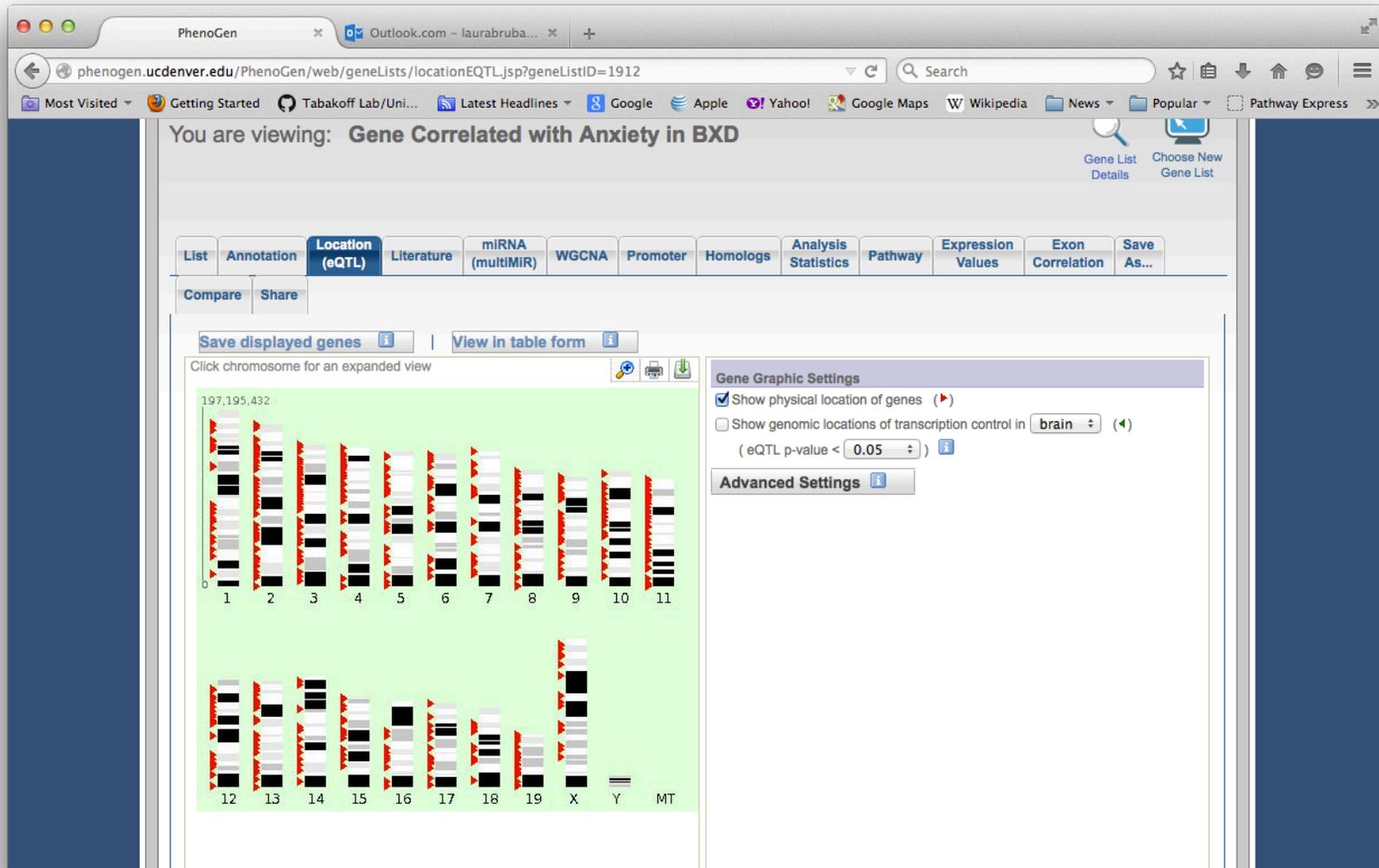
# Phenotypic QTL

## DNA Variants Associated with Phenotype



# Expression QTL

DNA Variants Associated with RNA Expression Levels



**DOWNLOADS**

# Downloadable Files

- Microarray Expression Data
  - Raw
  - Processed
  - eQTL/heritability
- Genotype Data
- RNA-Seq Data
  - Raw
  - Processed
- Code
  - Github

# **FUTURE DIRECTIONS**

# Future Data Plans

Strains	Tissue	Sex	Number of Biological Replicates Per Strain	Number of Paired-End Reads (rRNA-depleted total RNA)	Number of Single-End Reads (small RNA)
30 Classic Inbred Rat Strains	brain	male	4	18 trillion	5.5 billion
30 RI Rat Strains	brain	male	4	18 trillion	5.5 billion
30 Classic Inbred Rat Strains	liver	male	4	18 trillion	5.5 billion
30 RI Rat Strains	liver	male	4	18 trillion	5.5 billion

# Functions to Add in Future

- multiMiR (miRNA/mRNA pairs) for rat
  - Currently implemented for mouse
- WGCNA for uploaded data sets
- Causal inference in co-expression modules
- Google drive integration