

# Mapping addiction and behavioral traits: Getting close to causal variants using

*GeneNetwork.org*

Webinar 2

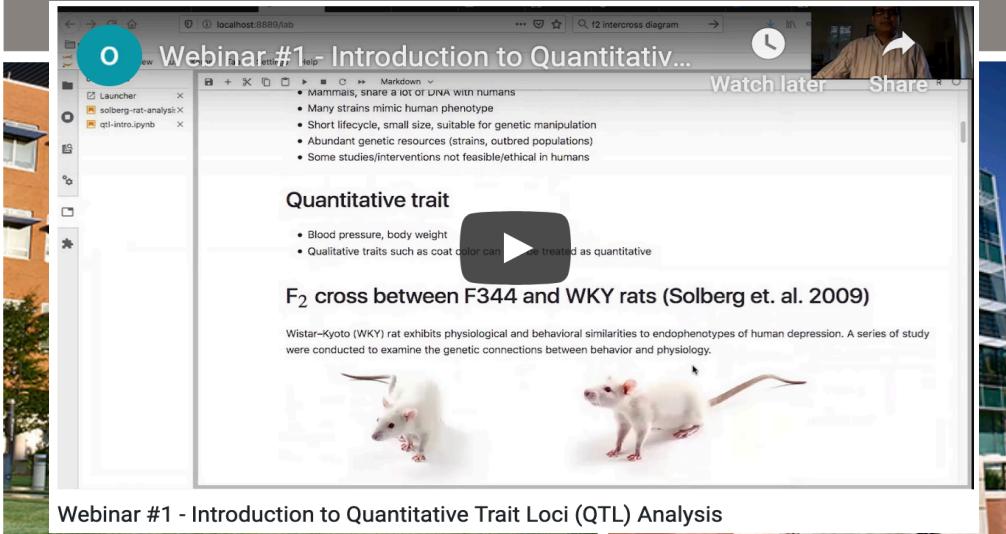
Williams RW, Centeno A, Chen H, Chitre A, Palmer AA, Prins J, Saba LM, Sen S, Sloan Z, Yousefi S, and P30 + P50 NIDA teams

Sponsored by the NIDA Center of Excellence in Omics, Systems Genetics, and the Addictome (NIDA P30 DA044223), with support from NIGMS Systems Genetics and Precision Medicine Project (R01 GM123489)

Presented by: Rob Williams  
[rwilliams@uthsc.edu](mailto:rwilliams@uthsc.edu)



# Summer of Rat Genetics and Genomics (SuRGe) Workshop



- 1 Quick review of rat, mouse and human molecular, SUD, and behavioral data in [GeneNetwork.org](#)
- 2 How to use GeneNetwork to **review and edit trait data**
- 3 How to **map QTLs**
- 4 How to **evaluate and rank candidate genes**

Please join us for two days in lovely Colorado to learn about rat genomics web resources including:

- ▶ Rat Genome Database
- ▶ GeneNetwork
- ▶ GeneWeaver
- ▶ PhenoGen
- ▶ Omics Portal for Addiction Research, etc.

Sponsored by the NIDA Center of Excellence in Omics, Systems Genetics, and the Addictome (<http://opar.io>) and in collaboration with researchers and data sponsored by NIAAA (PhenoGen), NIDA (GeneWeaver), NIGMS (GeneNetwork), and NHLBI (RGD).

Sessions will be separated into three phases:

1. identifying genomic regions of interest in rat
2. exploring a rat genomic region in the context of addiction
3. examining genes of interest for further biological/molecular/physiological/behavioral information.

Session will be led by content area experts and majority of time will be spent doing hands-on activities. Presenters and coaches include: Rob Williams (UTHSC), Saunak Sen (UTHSC), Laura Saba (CU-AMC), Katerina Kechris (CU-AMC), and Jennifer Smith (RGD/MCW).

All participants must come with a genomic area of interest in the rat and/or a set of relevant genes.

When: August 2nd and 3rd, 2019

Where: Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado Anschutz Medical Campus

Registration: [http://bit.ly/OSGA\\_2019](http://bit.ly/OSGA_2019)

GitHub Repository (with agenda): [https://github.com/OSGA-OPAR/SuRGe\\_2019](https://github.com/OSGA-OPAR/SuRGe_2019)

Contact: [Laura.Saba@cuanschutz.edu](mailto:Laura.Saba@cuanschutz.edu)

Phenotypes

Search All

nicotine cocaine morphine alcohol THC amphetamine impulsivity

GN searched for the term(s) "nicotine" in 51 datasets and 13763 traits across 10 species and found 105 results that match your query.  
You can filter these results by adding key words in the fields below and you can also sort results on most columns.

Genes / Molecules

Search All

TTC27 RASGRP3 DRD1A DRD2 TAAR1 MPDZ

GN searched for the term(s) "TTC27" in 754 datasets and 39,765,944 traits across 10 species and found 1487 results that match your query.  
You can filter these results by adding key words in the fields below and you can also sort results on most columns.

## Finding data

GeneNetwork Intro Help Tools Collections 4 Source Code S

Genes / Molecules Search All

### Select and search

Species: Mouse (mm10)

Human (hg19)  
Mouse (mm10)

Group: BXD Family

✓ Rat (rn6)  
Macaque monkey

Type: Hippocampus mRNA

Drosophila  
Barley

Dataset: Hippocampus Consortium M430v2 (Jun06) F

Arabidopsis thaliana  
Poplar

Get Any:

Soybean  
Tomato

Enter terms, genes, ID numbers in the **Search** field.

Use \* or ? wildcards (Cyp\*a?, synap\*).

Use quotes for terms such as "tyrosine kinase".

## Slide 3. Finding data

## Select and search

Species:

Mouse (mm10)

Group:

BXD Family

Type:

Nucleus Accumbens mRNA

Info

Dataset:

✓ VCU BXD NAc Sal M430 2.0 (Oct07) RMA

VCU BXD NAc CIE Air Masked ComBat M430 2.0 (Apr16) RMA

VCU BXD NAc CIE Air Masked ComBat M430 2.0 (Apr16) RMA Z-score

VCU BXD NAc CIE EtOH Masked ComBat M430 2.0 (Apr16) RMA

VCU BXD NAc CIE EtOH Masked ComBat M430 2.0 (Apr16) RMA Z-score

VCU BXD NAc EtOH M430 2.0 (Oct07) RMA

VCU BXD NAc EtOH vs Sal M430 2.0 (Oct07) Sscore\*

VCU BXD NAc CIE Air M430 2.0 (Jan13) RMA

VCU BXD NAc CIE EtOH M430 2.0 (Jan13) RMA

VCU BXD NAc EtOH vs CIE Air M430 2.0 (Jan13) Sscore\*

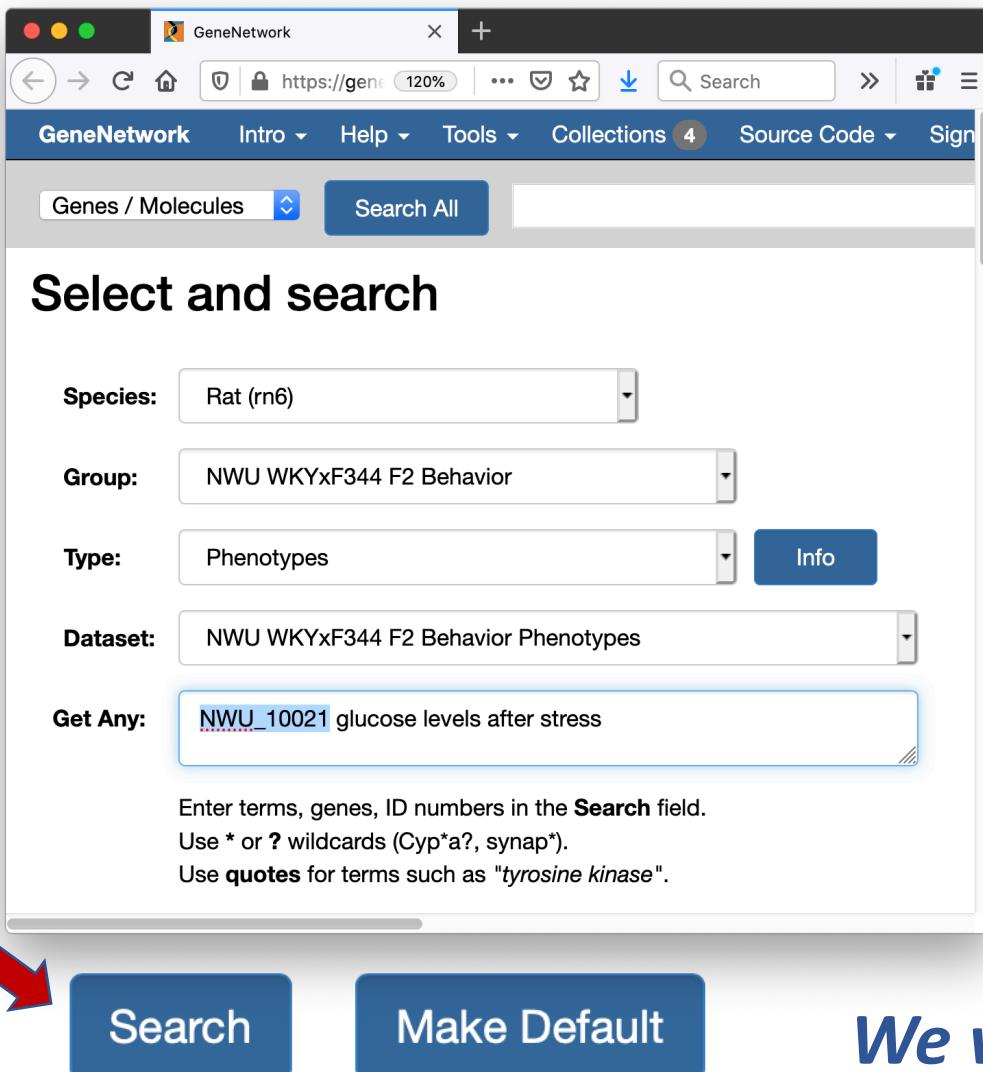
Combined:

Search

Make Default

# Data from our first webinar presentation by Saunak Sen

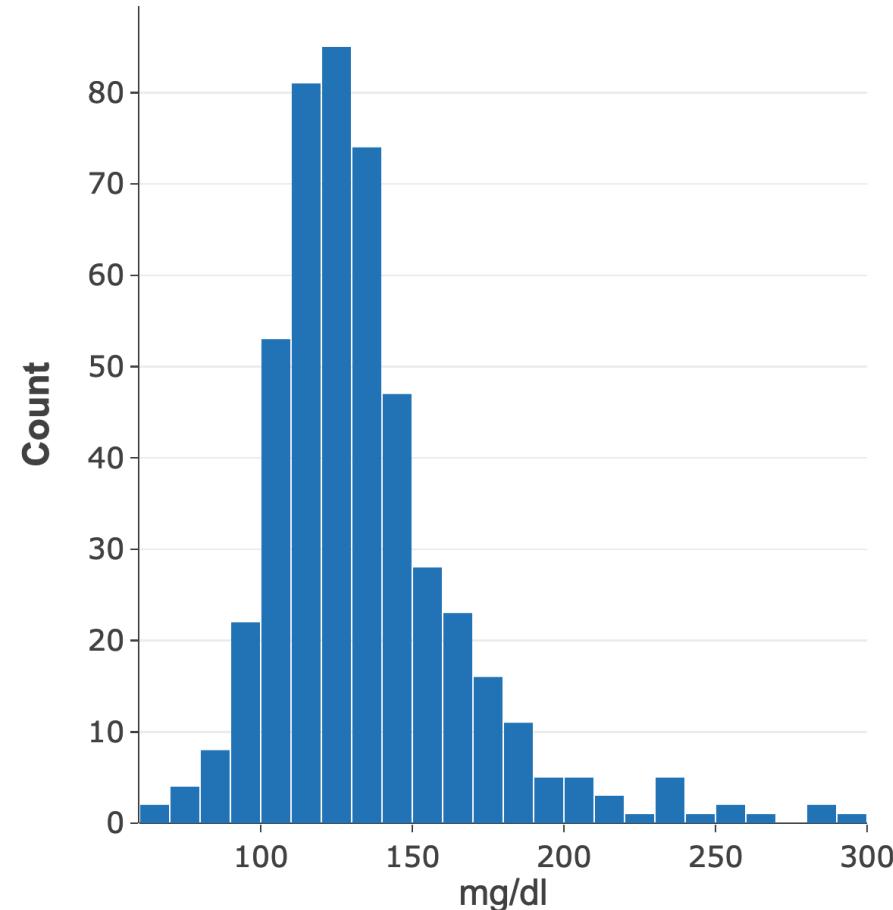
Link to [www.GeneNetwork.org](https://www.GeneNetwork.org)



The screenshot shows the 'Select and search' interface of the GeneNetwork website. The 'Species' dropdown is set to 'Rat (rn6)'. The 'Group' dropdown is set to 'NWU WKYxF344 F2 Behavior'. The 'Type' dropdown is set to 'Phenotypes', with an 'Info' button next to it. The 'Dataset' dropdown is set to 'NWU WKYxF344 F2 Behavior Phenotypes'. The 'Get Any:' input field contains the text 'NWU\_10021 glucose levels after stress'. Below the input field, instructions say: 'Enter terms, genes, ID numbers in the **Search** field. Use \* or ? wildcards (Cyp\*a?, synap\*). Use quotes for terms such as "tyrosine kinase".' At the bottom are two buttons: 'Search' and 'Make Default'.

Statistic	Value
N of Samples	480
Mean	134.228
Median	128.548
Standard Error (SE)	1.475
Standard Deviation (SD)	32.320
Minimum	62.310
Maximum	290.137
Range	227.827
Interquartile Range	31.361
Skewness	1.624
Excess Kurtosis	4.429

Trait 10021: Gluc\_PostStress

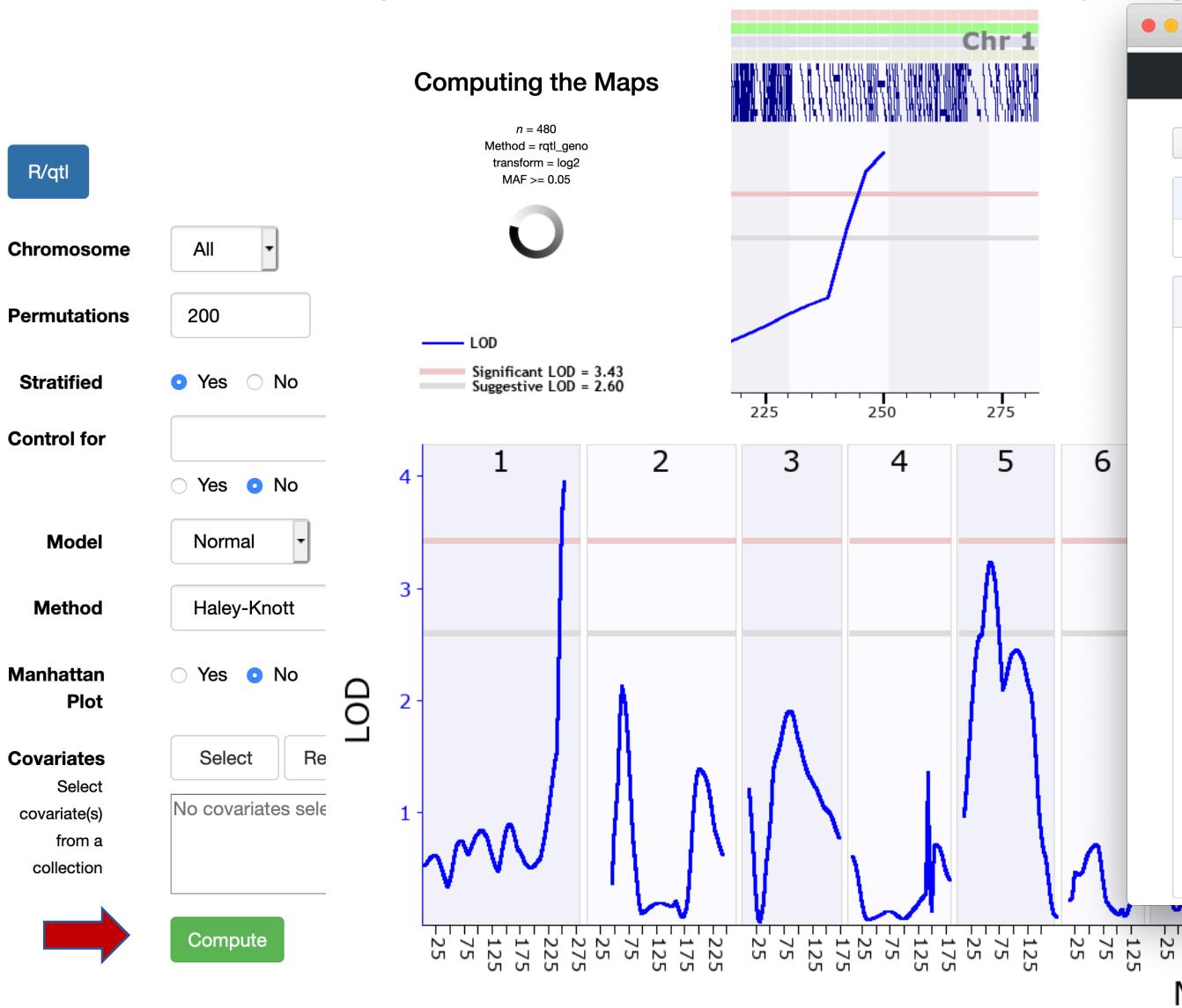


We were here on May 8, 2020, but using R/qtl

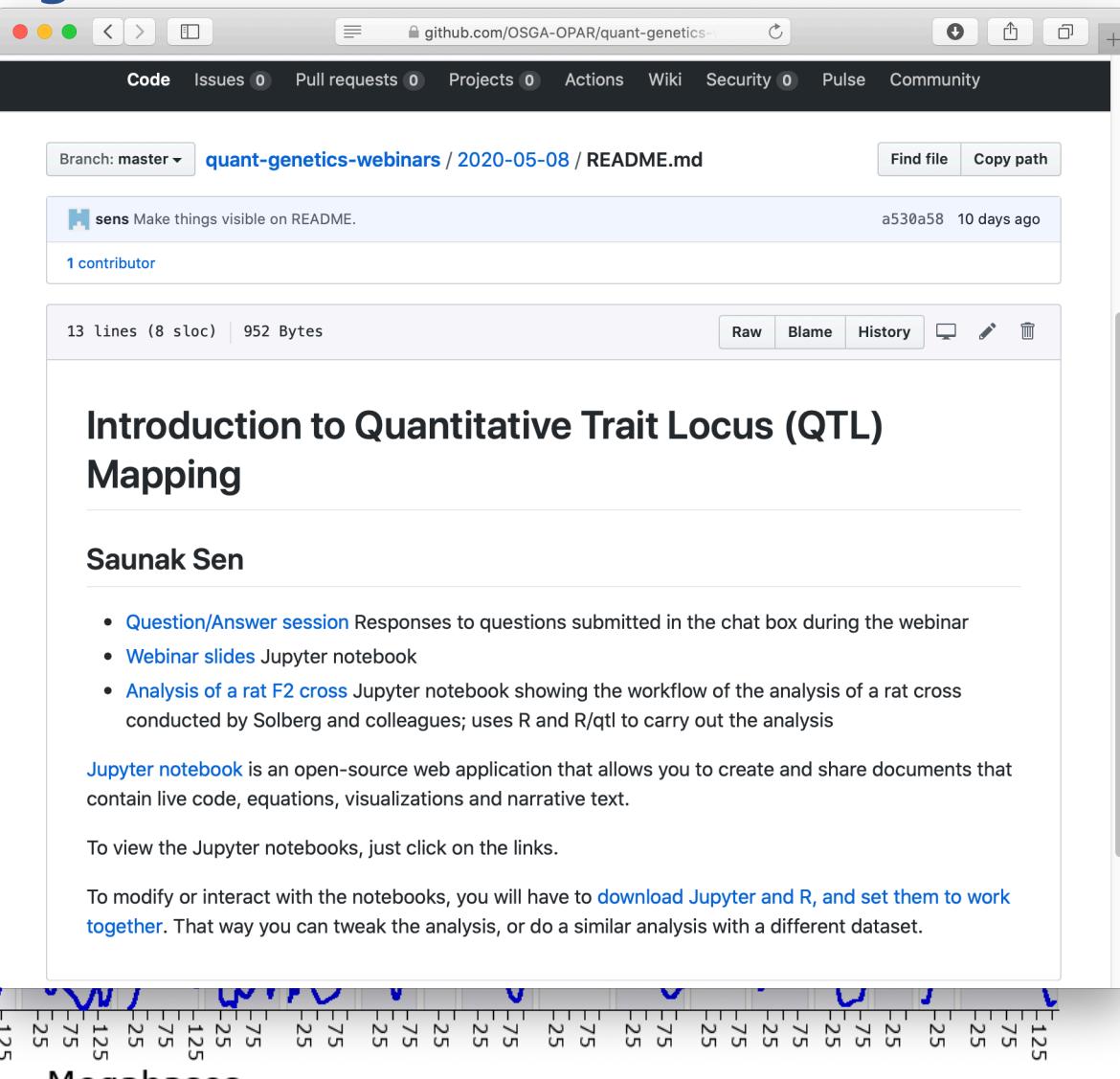
Slide 4. Configure the *Select and search* page as shown. Click on **Make Default**. Enter a wildcard (\*) in the **Get Any** field (red arrow). Click on **Search**.

# *Quick review of 1<sup>st</sup> webinar on QTL mapping*

[http://opar.io/webinar\\_series\\_1.html](http://opar.io/webinar_series_1.html)



**Slide 5.** Click on *Compute*. Just that easy.

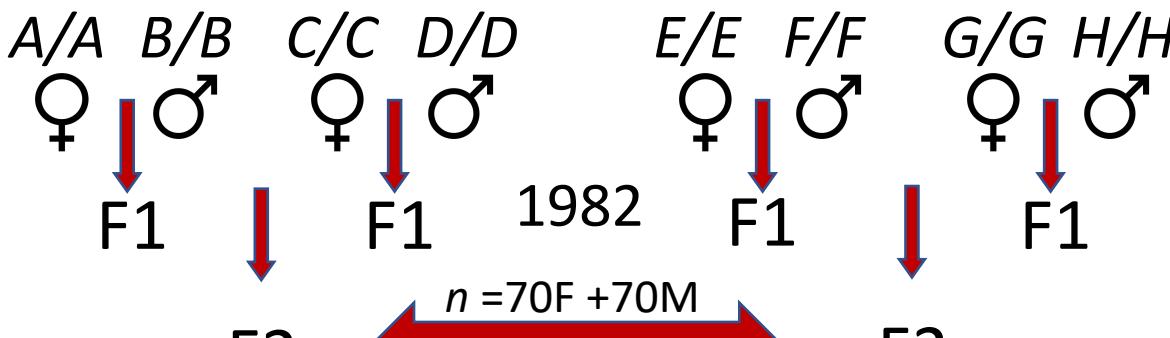


Link to course material: <https://github.com/OSGA-OPAR/quant-genetics-webinars/blob/master/2020-05-08/README.md>

# Today's focus: NIH HS rat cross

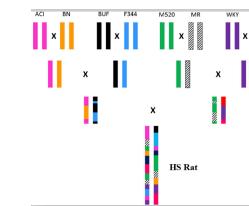


Genes and Addiction **RatGenes.org**  
NIDA Center for GWAS in Outbred Rats



♀ ♂ Reciprocal G3 ♀ ♂

A=BN/SsN      60 pairs  
B=MR/N      12 units x 5 cages  
C=BUF/N      effective population  
D-M520/N      of 120



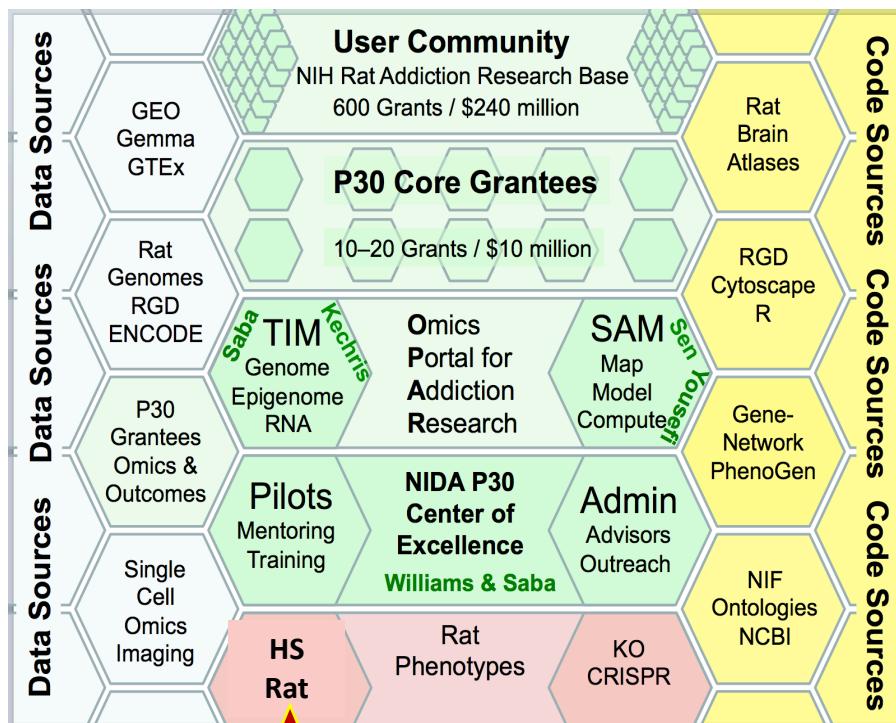
- Rat genome-wide mapping studies: great precision
- Extensive behavioral data and RNA-seq for brain regions maintained in relational database
- Very high density SNP markers (genotyping-by-sequencing)
- Powerful mapping methods (GEMMA, FastLMM)
- Data copied into GeneNetwork for companion analysis and for enhanced longterm FAIR data sharing.

# Getting to HS P50 data in GeneNetwork

Link to [www.GeneNetwork.org](http://www.GeneNetwork.org)

The screenshot shows the GeneNetwork homepage with a navigation bar at the top. The 'Collections' section has a red notification badge with the number '0'. Below the navigation bar is a search bar with dropdown menus for 'Genes / Molecules' and 'Search All'. The main area is titled 'Select and search' and contains four dropdown menus: 'Species: Rat (rn6)', 'Group: NIH Heterogeneous Stock (Palmer)', 'Type: Phenotypes', and 'Dataset: HSNIH Phenotypes'. At the bottom is a 'Get Any:' field containing a wildcard character (\*), a red arrow pointing to it, a 'Search' button, and a 'Make Default' button.

## Select and search



HS  
Rat

↑

You are here

Slide 7. Configure the *Select and search* page as shown. Click on **Make Default**. Enter a wildcard (\*) in the **Get Any** field (red arrow). Click on **Search**. Just that easy.

# Getting HS data into Your Collection

Select     Add    Download CSV    cofactor 

Select

Add

 Search This Table For ...

Select Top ...

Deselect

Show/Hide Columns:

Index

Record

Description

Mean

Authors

Year

Max LRS?

Showing 1 to 4 of 4 entries (filtered from 446 total entries)

	Index	Record	Description
<input checked="" type="checkbox"/>	443	HSR_10443	Cofactor, metadata: Sex (0=female, 1=male) [binary]
<input checked="" type="checkbox"/>	444	HSR_10444	Cofactor, metadata: Phenotyping batch within each phenotyping center
<input checked="" type="checkbox"/>	445	HSR_10445	Cofactor, metadata: Coat color (1=brown, 2=brownhood, 3=blackhood, 4=black, 5=albino)
<input checked="" type="checkbox"/>	446	HSR_10446	Cofactor, metadata: Phenotyping batch (1=University of Michigan, 2= University at Buffalo, 3= University of Tennessee)

## Define or Add to Collection

You have two choices: Name a new collection or add to an existing collection.

### 1. Create a new collection

HS\_RAT\_GWAS

Create collection

**Slide 8. Search Results. Left.** GeneNetwork returns all 508 HS traits and cofactors. These traits are not yet linked to PMID citations. The *Palmer AA et al.* citation is often a placeholder. The **Select** and **Add** functions (top red arrow) are used in sequence to add traits and cofactors to **Your Collection**. **Top right.** Type the search term **cofactor** into the text field (arrow). Click on **Select** and **Add** all of the cofactors to define a **Collection—bottom right**—here named **HS\_RAT\_GWAS**. Collections are needed in order to use cofactors when mapping (*sex, batch, site, year...*).

# Collection of traits and cofactors

Tools ▾

Collections 2

Source Code ▾

HS\_RAT\_GWAS

Change Collection Name

This collection has 27 records

Correlation Matrix Network Graph WGCNA Analysis CTL Analysis Heatmap Comparison Bar Chart WebGestalt GeneWeaver BNW

✓ Select All Copy Download CSV Search Table For ... Select Top ... ✘ Deselect Delete Rows Delete Collection

Index	Dataset	Record	Symbol	Description	Location	Mean	Max LRS ?	Max LRS Location	Additive Effect ?
	Index	Record		Description		Mean			
✓	1	HSR_10444		Cofactor, metadata: Phenotyping batch within each phenotyping center		8.912			
✓	2	HSR_10443		Cofactor, metadata: Sex (0=female, 1=male) [binary]		0.507			
✓	3	HSR_10446		Cofactor, metadata: Phenotyping batch (1= University of Michigan, 2= University at Buffalo, 3= University of Tennessee)		2.192			
□	4	HSR_10445		Cofactor, metadata: Coat color (1=brown, 2=brownhood, 3=blackhood, 4=black, 5=albino)		2.797			
□	5	HSR_10009		Psychiatry, motivation: Total distance of locomotion during session 7 for males and females [meters]		16699.275			
□	6	HSR_10004		Psychiatry, motivation: Average velocity of locomotion during session 2 for males and females [meters/sec]		80.254			
□	7	HSR_10020		Central nervous system, behavior, learning and memory: Pavlovian Conditioned Approach, conditioned reinforcement, number of active nose-poke hole entries during the test		96.532			
□	8	HSR_10002		Psychiatry, motivation: Average velocity of locomotion during session 1 for males and females [meters/sec]		81.260			
□	9	HSR_10017		Psychiatry, motivation: Contextual conditioning distance difference score, difference of total distance of locomotion during session 8 and session 2		3145.250			
□	10	HSR_10011		Psychiatry, motivation: Stereotypy head waving bouts on seventh day [n]		504.274			
□	11	HSR_10021		Central nervous system, behavior, learning and memory: Pavlovian Conditioned Approach, conditioned reinforcement, number of inactive nose-poke hole entries during the test		4 2007			
□	12	HSR_10001		Central nervous system, behavior: Distance traveled in open field (locomotion) during session 1 for males and females [meters/session]		8147.933			
□	13	HSR_10018		Central nervous system, behavior: Distance traveled in open field (locomotion) during session 1 for males and females [meters/session] minus its memory: Pavlovian Conditioned Approach, conditioned reinforcement, active test		55.133			
□	14	HSR_10008		Psychiatry, motivation: Duration on third day [sec]		132.752			
□	15	HSR_10010		Psychiatry, motivation: Average velocity of locomotion during session 7 for males and females [meters/sec]		84.115			
□	16	HSR_10015		Psychiatry, motivation: Degree of sensitization distance, difference of total distance of locomotion during sessions 7 and 3		-7972.985			
□	17	HSR_10016		Psychiatry, motivation: Decree of sensitization stereotypy. difference of head waving bouts during sessions 7 and 3		321.106			

**Slide 9. Left.** The yellow highlighted rows are cofactors. Each trait has a permanent **Record identifier** with format **HSR\_10505** that are citable in papers and figures. To view data, click on blue ID numbers. **Top. Collections, Tools, and Code** are available from the **menu**. **Right.** Any open HS data can be accessed directly with the URL. Please be patient when loading data into your browser—the rat GWAS is one of the largest cohorts in GeneNetwork.

Your now have 2 collections

This link will take you directly to the HS body weight data in GeneNetwork:

[http://genenetwork.org/show\\_trait?trait\\_id=10505&dataset=HSNIH-PalmerPublish](http://genenetwork.org/show_trait?trait_id=10505&dataset=HSNIH-PalmerPublish)

**Known issue:** Because HS rat GWAS data set are so large, opening some windows may take a web browser a few minutes. Don't be surprised by this window:



# Motivation to obtain nicotine or subjective value of nicotine—in brief—seeking

## Trait Data and Analysis for HSR\_10505

▼ Details and Links

Group	Rat: HSNIH-Palmer group
Phenotype	Central nervous system, behavior: Ratio of licks on the active over inactive spout during a progressive ratio (PR) session tested after 10 fixed ratio sessions in male and female adolescents (41 to 51 days old) trained using a socially acquired nicotine intravenous self-administration procedure where licking was used as the operant response (Chen lab cohort, UTHSC) [ratio]
Authors	Wang T, Chitre A, Garcia Martines A, Han W, Polesskaya O, Solberg Woods L, Palmer AA, Chen H
Title	Genetics of nicotine use and the influence of social cues in adolescent rats
Journal	None (2020)
Database	HSNIH Phenotypes

Add View in GN1

Index	Record	Description	Mean	Authors	Year
505	HSR_10505	Central nervous system, behavior: Ratio of licks on the active over inactive spout during a progressive ratio (PR) session tested after 10 fixed ratio sessions in male and female adolescents (41 to 51 days old) trained using a socially acquired nicotine intravenous self-administration procedure where licking was used as the operant response (Chen lab cohort, UTHSC) [ratio]	7.204	Wang T, Chitre A, Garcia Martines A, Han W, Polesskaya O, Solberg Woods L, et al.	2020

▼ Statistics

▼ Transform and Filter Data

▼ Calculate Correlations

▼ Mapping Tools

▲ Review and Edit Data

Search This Table For ... Export Excel Reset

Samples

ID	Sample	Value	Sex	Color	Batch	DOB	Site
1	0007A0008B	7.250	M	brown	B11	10/9/17	TN

▲ Review and Edit Data

Search This Table For ... Export Excel Reset

**Page bottom**

Samples

ID	Sample	Value	Sex	Color	Batch	DOB	Site	DOW	DOD
1	0007A0008B	7.250	M	brown	B11	10/9/17	TN	10/30/17	12/12/17
2	0007A00024	0.596	F	brown	B12	11/9/17	TN	11/30/17	1/18/18
3	0007A000DB	0.000	M	brownhood	B13	2/13/18	TN	3/6/18	4/23/18
4	0007A001C5	x	F	blackhood	B09	4/6/17	TN	4/28/17	6/30/17
5	0007A0059F	0.000	F	black	B09	3/10/17	TN	4/1/17	5/11/17
6	0007A00263	1.500	F	brown	B12	12/6/17	TN	12/27/17	2/6/18
7	0007A00670	0.400	M	brownhood	B10	7/5/17	TN	7/26/17	9/14/17
8	0007A00716	0.246	F	brown	B10	4/22/17	TN	5/13/17	6/27/17
9	0007A01A7C	x	F	black	B09	2/27/17	TN	3/20/17	5/26/17
10	0007A01C94	x	F	brown	B09	4/23/17	TN	5/14/17	7/20/17
11	0007A06B58	0.000	M	albino	B11	8/28/17	TN	9/18/17	10/31/17
12	0007A06BBD	x	F	brownhood	B11	10/9/17	TN	10/30/17	1/4/18
13	0007A06D20	0.010	M	brown	B12	11/9/17	TN	11/30/17	1/19/18
14	0007A011E3	0.273	F	black	B11	10/9/17	TN	10/30/17	12/20/17
15	0007A017AB	0.455	F	brown	B10	4/23/17	TN	5/14/17	6/27/17
16	0007A022BB	0.000	M	albino	B09	2/27/17	TN	3/20/17	5/5/17
17	0007A027D6	x	F	albino	B11	10/25/17	TN	11/15/17	1/19/18
18	0007A037D1	x	M	brown	B11	10/9/17	TN	10/30/17	1/4/18
19	0007A0153B	19.700	F	blackhood	B09	3/13/17	TN	4/3/17	5/17/17

**Slide 10.** Progressive ratio test, ratio of active to inactive lever presses: trait [HSR\\_10505](#). **Left.** The top of the *Trait Data and Analysis* window. **Right.** Scroll down to review the data and cofactors. Most columns are self-explanatory. **DOB**, **DOW**, and **DOD** are dates of birth, experimental work, and death. In this screenshot, all individuals are shown, and for this trait, all phenotypes generated by Hao Chen and team at **Site = TN**

# Resorting: high to low

## Trait Data and Analysis

▼ Details and Links

▼ Statistics

▼ Transform and Filter Data

▼ Calculate Correlations

▼ Mapping Tools

▲ Review and Edit Data

▲ Review and Edit Data

albino Export Excel Reset

Samples

	ID	Sample	Value	Sex	Color	Batch	DOB	Site	DOW	DOD
■	1	000792A2AE	707.000	F	albino	B09	5/1/17	TN	5/25/17	7/12/17
■	2	00078A0153	449.500	M	albino	B03	8/21/15	TN	9/11/15	10/23/15
■	3	00078A17AC	211.167	F	albino	B03	8/5/15	TN	8/26/15	10/7/15
■	4	00078A19F2	143.143	F	albino	B03	8/5/15	TN	8/26/15	10/7/15
■	5	00078A260C	112.000	F	albino	B02	5/5/15	TN	5/26/15	7/7/15
■	6	00077E78A5	64.000	F	albino	B01	1/28/15	TN	2/18/15	4/16/15
■	7	000789993E	60.333	M	albino	B10	5/26/17	TN	6/16/17	8/8/17
■	8	0007929B26	24.000	F	albino	B07	9/21/16	TN	10/12/16	12/13/16
■	9	00078A0070	19.182	M	albino	B07	8/10/16	TN	8/31/16	10/17/16
■	10	00078A23B1	12.333	M	albino	B02	5/8/15	TN	5/29/15	7/9/15
■	11	00078A072B	12.294	M	albino	B02	5/5/15	TN	5/26/15	7/9/15
■	12	00077E8308	12.250	M	albino	B01	2/9/15	TN	2/28/15	4/16/15
■	13	00078A23F4	11.700	M	albino	B02	5/5/15	TN	5/26/15	7/10/15
■	14	00078A0331	10.500	F	albino	B03	8/21/15	TN	9/11/15	10/27/15
■	15	00078A242C	8.000	M	albino	B03	9/28/15	TN	10/22/15	11/30/15
■	16	00077E811A	7.806	F	albino	B01	2/9/15	TN	2/28/15	4/16/15
■	17	00078A0A00	6.667	F	albino	B02	5/5/15	TN	5/26/15	7/7/15
■	18	00078A22E6	5.714	F	albino	B03	9/28/15	TN	10/22/15	12/1/15
■	19	0007D60B77	5.575	F	albino	B14	4/27/18	TN	NA	6/26/18

Slide 11. **Left.** The *Trait Data and Analysis* window has six major sections. The last is *Review and Edit Data* (large red arrow) and is open by default. **Middle.** The *Value* column (small red arrow) has been sorted from high to low. Sorting will take a minute. Once sorted, rows of outlier cases—those with very high or low values—will be flagged as potential outliers in orange. The Value fields can be edited. You can winsorize outliers.

# Normalizing and blocking data

## Transform and Filter Data

Edit or delete values in the Trait Data boxes, and use the **Reset** option as needed.

Block samples by index: Example: 3, 5-10, 12    HSNIH-Palmer    Block

Block samples by group: Sex    F    Block

Hide No Value    Block Outliers    Reset    Export    Excel

Normalize    Log2

- ✓ Sex
- Color
- Batch
- DOB
- Site
- DOW
- DOD

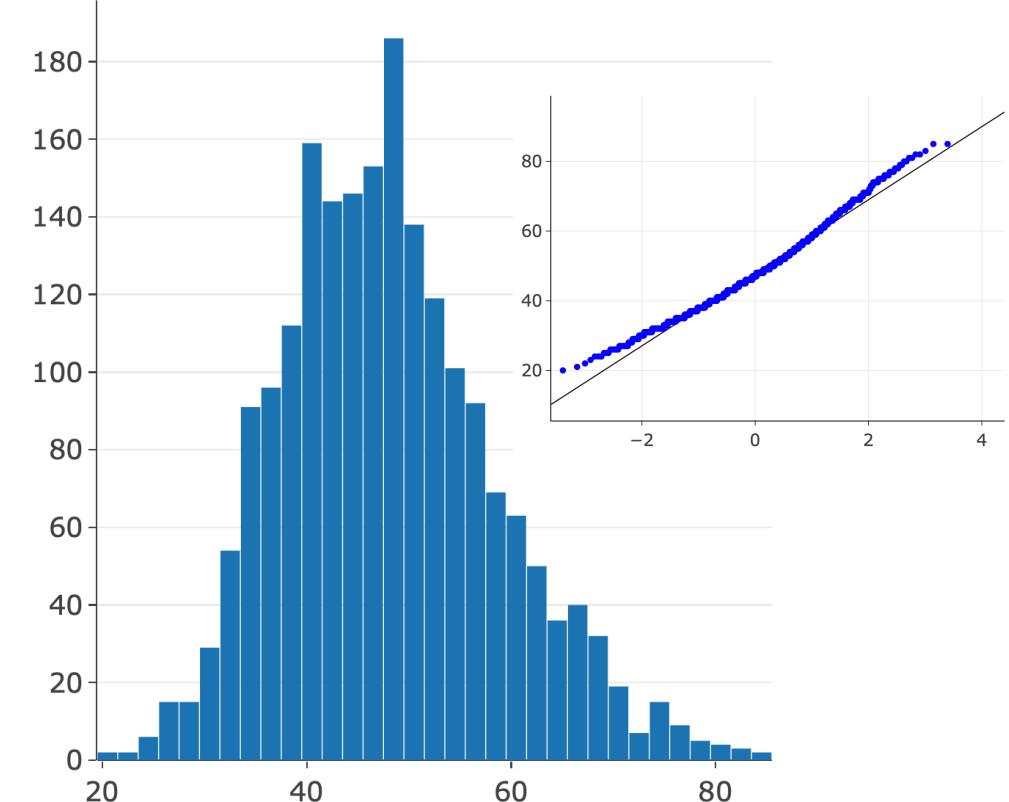
Normalize by changing scales first. This affects all values. Blocking samples. Blocking is only additive.

Block samples by group: Sex    M    Block

Reset and Export    Reset    Export    Excel

## HSNIH-Palmer Only

ID	Sample	Value	Sex	Color
1	00077E82D2	x	F	blackhood
2	00077E8431	x	F	brown
3	00077E8E49	85.000	F	black
4	00077E9272	85.000	F	blackhood



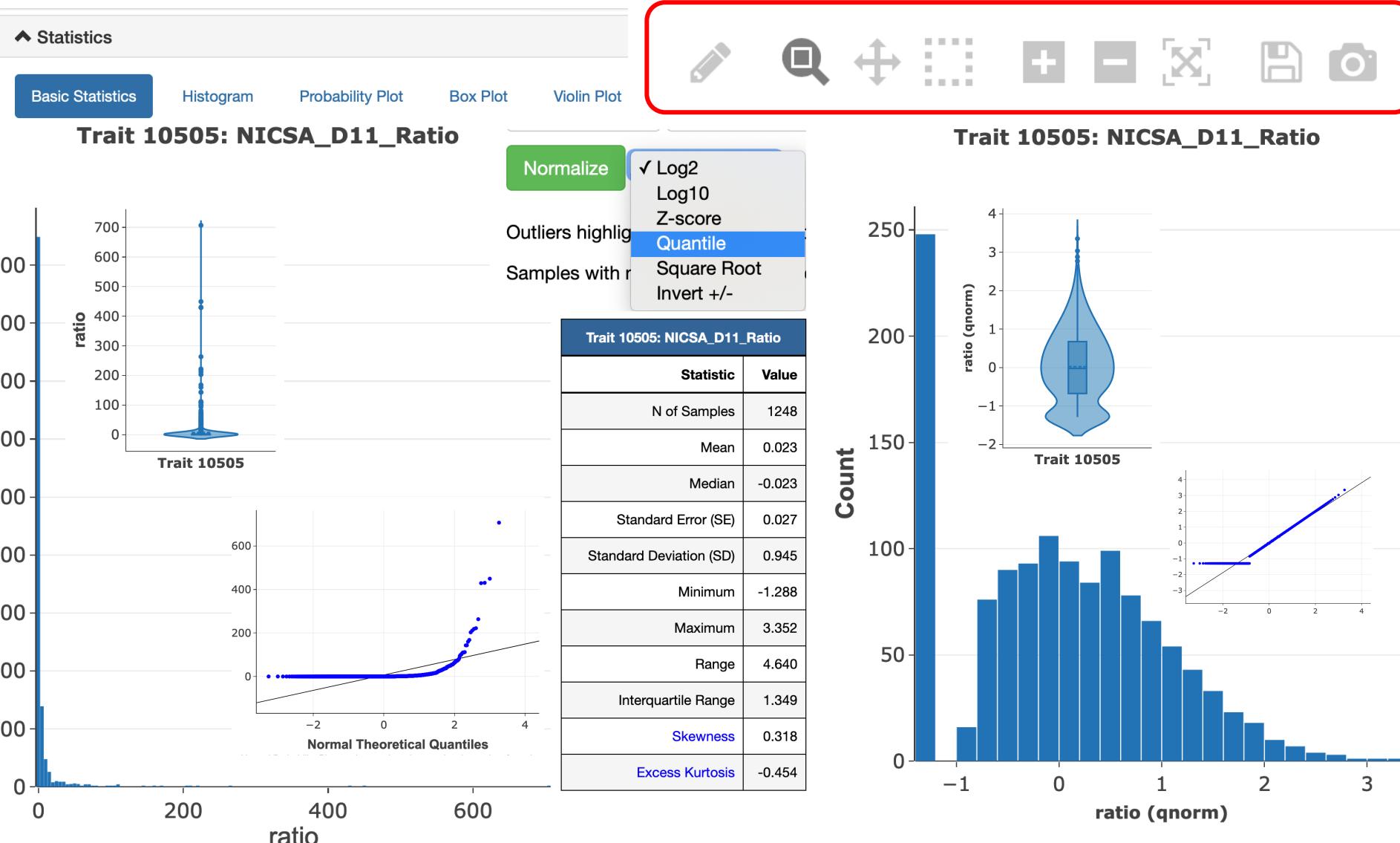
**Slide 12. Normalizing and blocking.** In this example the goal is to map body weight of male and female cases separately and then jointly. **Left.** The first step is to use the **Transform and Filter Data** tab functions, specifically **Block samples by group: Sex = M**. This action replaces male **Values** with an **x**. You can block by **Site, Batch, Color**. **Right.** Review **Statistics** again, especially **Basic Statistics** and **Histogram**. Censor the two highest body weights (92 and 94 g).

# Statistics windows

## Statistics

Basic Statistics Histogram Prc

Trait 10505: NICSA_D11_Ratio	
Statistic	Value
N of Samples	1248
Mean	7.204
Median	0.333
Standard Error (SE)	1.004
Standard Deviation (SD)	35.474
Minimum	0.000
Maximum	707.000
Range	707.000
Interquartile Range	2.132
Skewness	11.785
Excess Kurtosis	177.995



**Slide 13. Left.** Summary statistics. **Middle.** Motivation to seek nicotine measured by a progressive ratio test is skewed (non-normal) and cannot be mapped correctly using the original scale. **Insets.** Quantile probability plots in the two histograms, a *violin plot*, and statistics. **Top.** Graphing tools in the **red box**. **Left.** Application of a quantile transform gets data into a nearly normal form that is more suitable for mapping.



# Maps for nicotine seeking motivation using GEMMA with out cofactors

▲ Mapping Tools

GEMMA Haley-Knott Regression

Chromosome All

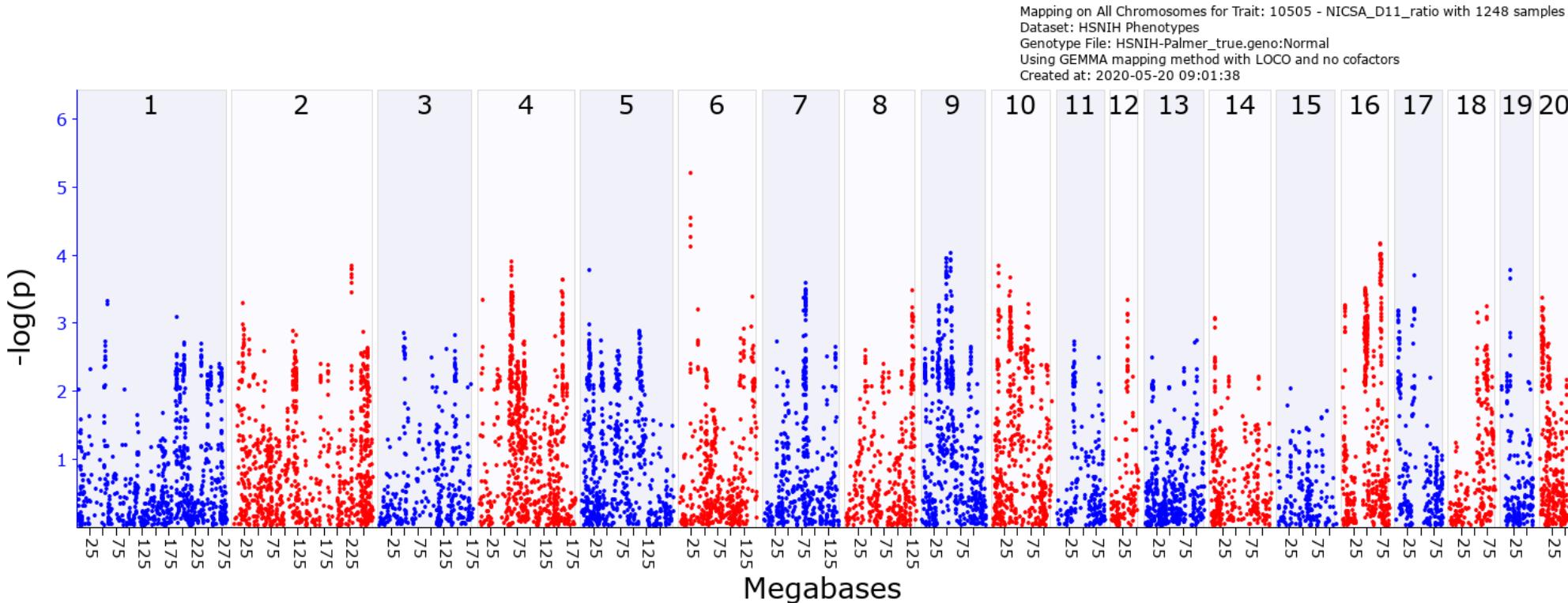
Genotypes Normal

MAF >= 0.05

Use LOCO  Yes  No

Covariates Select Remove  
Select covariate(s) from a collection  
No covariates selected

Compute



**Slide 14. Left.** Maps for nicotine seeking without cofactors ([HSR\\_10505](#)). The x-axis of the map lists chromosomes on a megabase (Mb) scale. The y-axis is *minus the logarithm of the linkage probability* ( $-\log p$ ). The Chr 6 locus has a peak linkage of about 5.19. These maps can be zoomed to single chromosome level or even small parts of chromosomes.

# Selecting covariates for mapping (all cases)

[GEMMA](#)[Haley-Knott Regression](#)[R/qtI](#)

## Mapping Tools

[GEMMA](#)[Haley-Knott Regression](#)[R/qtI](#)

Chromosome

Genotypes

MAF &gt;=

Use LOCO

 Yes    No 

Select covariate(s) from a collection  
10443 - sex  
10444 - batchnumber

## Computing the Maps

$n = 1248$   
Method = GEMMA  
transform = qnorm  
MAF  $\geq 0.05$



Chromosome

Permutations

Stratified

 Yes    No

Control for

 Yes    No

Model

Method

Manhattan Plot

 Yes    No

Covariates

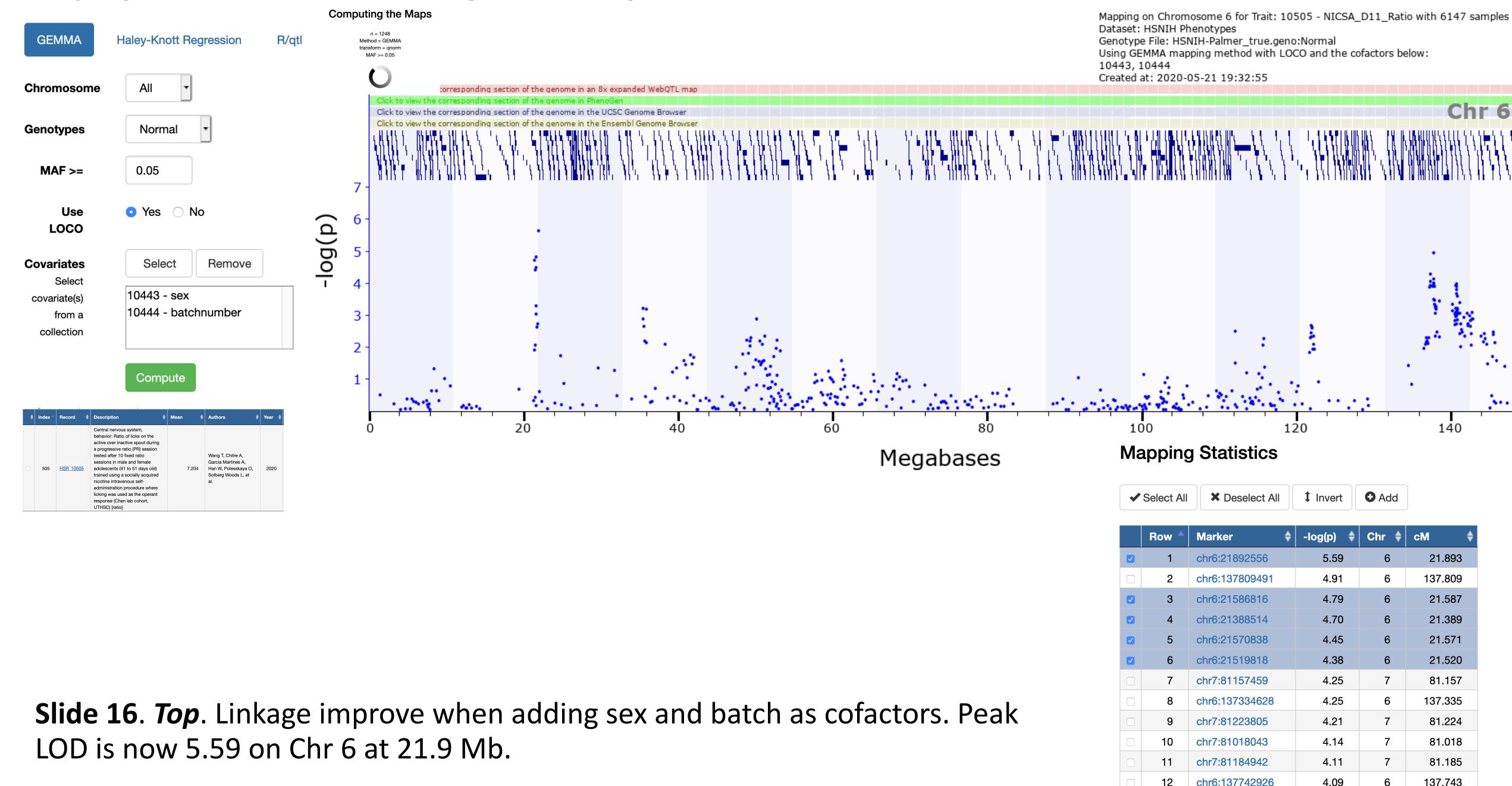
Select covariate(s) from a collection  
10443 - sex  
10444 - batchnumber

10443 - sex  
10444 - batchnumber

10446 - center

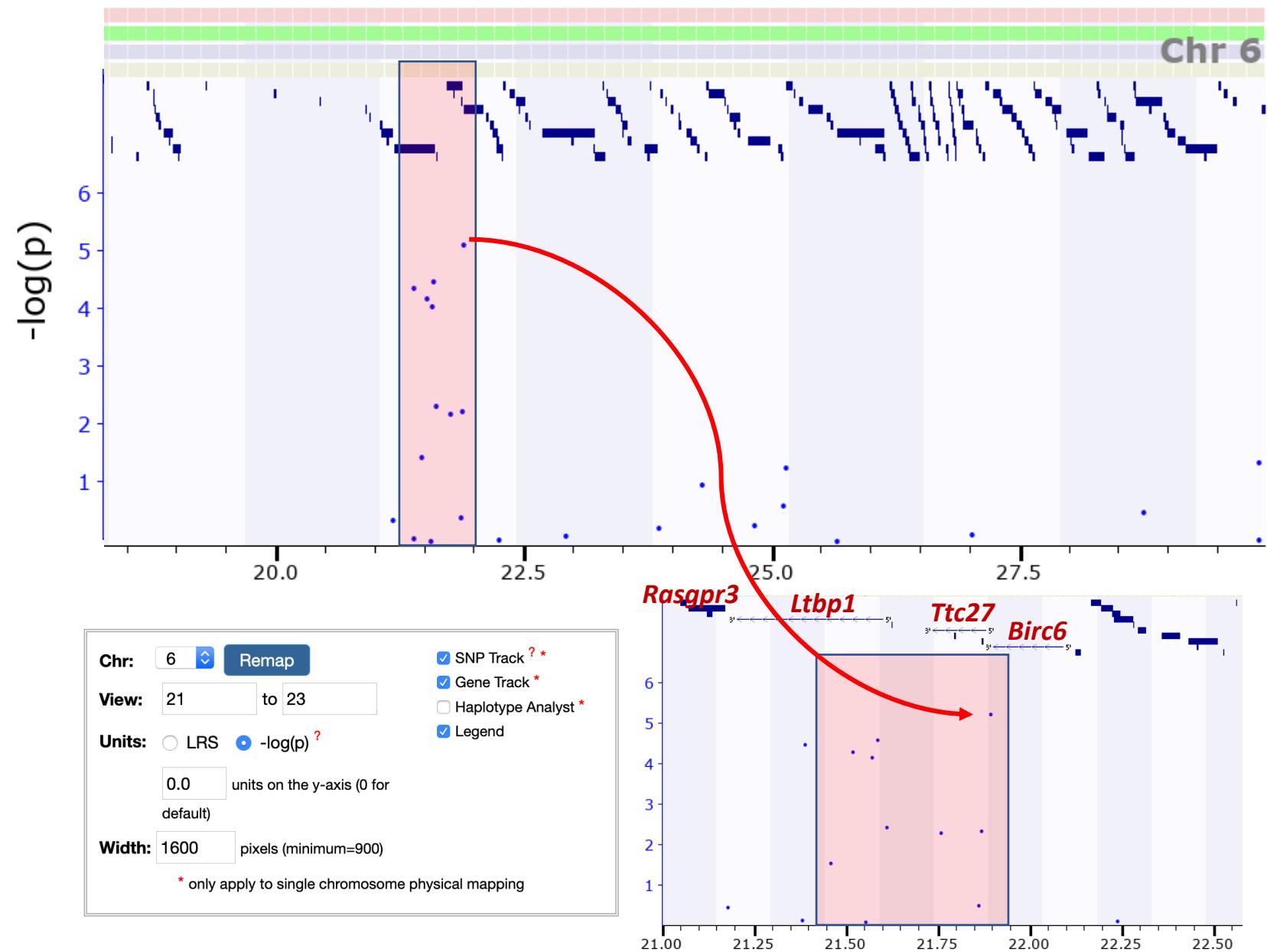
**Slide 15.** Mapping variation in nicotine seeking using sex and batch as addictive cofactors. We select appropriate **covariates** **sex** and **batchnumber** to include in the analysis from our custom **HS\_RAT\_GWAS** collection (red arrow). **GEMMA** or (Right) **R/qtI** can be used for mapping the HS. Click **GEMMA Compute**. **Computing the Maps** progress bar will appear. Time for a coffee break.

# Maps for nicotine seeking with cofactors: sex and batch



# Progressive zoom of the Chr 6 region of best linkage

Select All    Deselect All    Invert    Add



Row	Marker	-log(p)	Chr	Mb
1	chr6:21892556	5.19	6	21.892556
2	chr6:21586816	4.53	6	21.586816
3	chr6:21388514	4.42	6	21.388514
4	chr6:21519818	4.24	6	21.519818
5	chr16:73507202	4.15	16	73.507202
6	chr16:73524442	4.14	16	73.524442
7	chr6:21570838	4.10	6	21.570838
8	chr9:54925611	4.01	9	54.925611
9	chr16:73512789	3.99	16	73.512789
10	chr16:73820086	3.99	16	73.820086
11	chr16:73517286	3.97	16	73.517286
12	chr16:73529666	3.96	16	73.529666
13	chr9:45735969	3.94	9	45.735969
14	chr9:45681347	3.93	9	45.681347
15	chr16:74004047	3.92	16	74.004047
16	chr9:54855118	3.92	9	54.855118
17	chr16:73848841	3.92	16	73.848841
18	chr16:73528029	3.91	16	73.528029
19	chr9:54889453	3.90	9	54.889453
20	chr16:73778401	3.89	16	73.778401
21	chr4:62306712	3.89	4	62.306712
22	chr9:45798972	3.88	9	45.798972
23	chr16:74024582	3.86	16	74.024582
24	chr16:73609382	3.85	16	73.609382
25	chr2:225847493	3.83	2	225.847493
26	chr10:11450644	3.81	10	11.450644
27	chr9:45910519	3.80	9	45.910519
28	chr4:62275182	3.80	4	62.275182

Slide 17. Zoomed to a small region of Chr 6 near markers with the highest linkage scores (LOD 5.19 at 21.9 Mb).

# Interval analyst table of Chr 6 locus

	Index	Symbol	Mb Start	Length (Kb)	Avg Expr	Mouse Chr	Mb Start (mm9)	Human Chr	Mb Start (hg19)	Gene Description
<input type="checkbox"/>	1	Fam98a	21.051327	14.659		17	75.537085			family with sequence similarity 98, member A
<input checked="" type="checkbox"/>	2	Rasgrp3	21.072634	98.985		17	75.435904	2	33.573066	RAS guanyl releasing protein 3
<input type="checkbox"/>	3	LOC102549280	21.124268	11.651						60S ribosomal protein L32-like
<input checked="" type="checkbox"/>	4	Ltbp1	21.203502	396.939		17	75.005528	2	33.084042	latent transforming growth factor beta binding protein 1
<input type="checkbox"/>	5	LOC108351356	21.626809	0.443						40S ribosomal protein S18 pseudogene
<input checked="" type="checkbox"/>	6	Ttc27	21.735833	144.175		17	74.717749	2	32.628035	tetratricopeptide repeat domain 27
<input type="checkbox"/>	7	LOC108351178	21.797147	2.614						uncharacterized LOC108351178
<input type="checkbox"/>	8	LOC102546793	21.872392	2.264						uncharacterized LOC102546793
<input checked="" type="checkbox"/>	9	Birc6	21.900763	191.721		17	74.528294	2	32.493380	baculoviral IAP repeat-containing 6
<input type="checkbox"/>	10	Yipf4	22.126870	11.416		17	74.489492			Yip1 domain family, member 4
<input type="checkbox"/>	11	Nlrc4	22.167874	26.881		17	74.426294			NLR family, CARD domain containing 4
<input type="checkbox"/>	12	Slc30a6	22.197003	29.361		17	74.395607	2	32.302583	solute carrier family 30 member 6
<input type="checkbox"/>	13	LOC103692563	22.226452	20.005						uncharacterized LOC103692563
<input checked="" type="checkbox"/>	14	Spast	22.230067	52.099		17	74.338986	2	32.200330	spastin
<input type="checkbox"/>	15	LOC685651	22.284066	0.806						similar to ribosomal protein L18a
<input type="checkbox"/>	16	Dpy30	22.296128	20.766		17	74.299473			dpy-30 histone methyltransferase complex regulatory subunit
<input type="checkbox"/>	17	Memo1	22.362483	47.100		17	74.200699			mediator of cell motility 1

Slide 18. Top. A list of genes in a 1 Mb interval.

# The review of candidate genes begins

## Addition of useful prior data

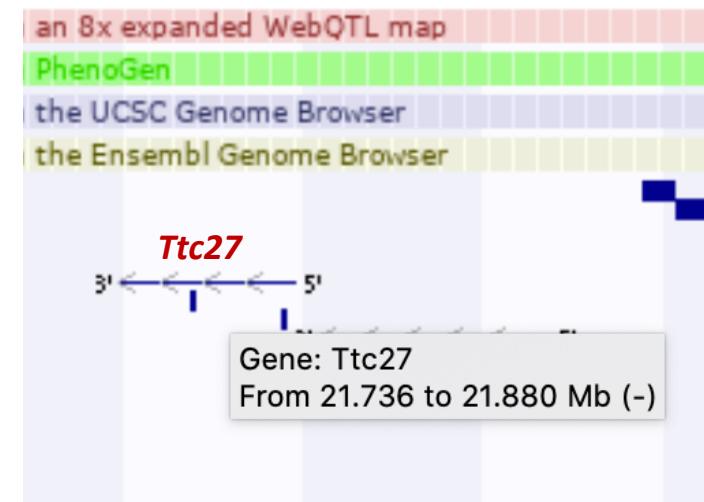
### Intrusion of bias

### The pressure of narrative potential

A screenshot of the GeneNetwork website interface. At the top, there's a navigation bar with links like 'GeneNetwork', 'Intro', 'Help', 'Tools', 'Collections (4)', 'Source Code', and 'Sign in'. Below the bar, there's a search bar with dropdown menus for 'Genes / Molecules' and 'Search All'. The main area is titled 'Select and search' and contains several dropdown menus:

- Species: Rat (rn6)
- Group: NIH Heterogeneous Stock (Palmer)
- Type: Nucleus Accumbens mRNA
- Dataset: HSNIH-Palmer Nucleus Accumbens Core RNA-Seq (Aug18) rlog
- Get Any: Fam98a Rasgrp3 Ltbp1 Ttc27 Birc6 Nrcc4 Slc30a6

At the bottom of this section, there's a note: 'Enter terms, genes, ID numbers in the Search field. Use \* or ? wildcards (Cyp\*a?, synap\*). Use quotes for terms such as "tyrosine kinase".'



A screenshot of the GeneNetwork 'Get Any' search results for the gene *Ttc27*. The results are organized into several sections:

- Species:** Rat (rn6)
- Group:** NIH Heterogeneous Stock (Palmer)
- Type:** ✓ Phenotypes  
Infralimbic Cortex mRNA  
Lateral Habenula mRNA  
Nucleus Accumbens mRNA  
Orbitofrontal Cortex mRNA  
Prelimbic Cortex mRNA
- Dataset:** (Listed below Type)
- Get Any:** (Listed below Dataset)

A large blue 'Info' button is located on the right side of the page. The 'Get Any' section lists the following genes:

Fam98a Rasgrp3 Ltbp1 Ttc27 Birc6 Nrcc4 Slc30a6

Slide 19. Top. Review expression of candidates in nucleus accumbens of HS rats

# The review of candidate genes begins

GeneNetwork Search Results

https://genenetwork.org/se

Search

Show/Hide Columns:

Index Record Symbol Description Location Mean Max LRS? Max LRS Location Additive Effect?

Showing 1 to 7 of 7 entries

	Index	Record	Symbol	Description	Location	Mean
<input checked="" type="checkbox"/>	1	ENSRNOG00000027191	Birc6	baculoviral IAP repeat-containing 6	Chr6: 21.900763	9.996
<input type="checkbox"/>	2	ENSRNOG00000030328	Fam98a	family with sequence similarity 98, member A	Chr6: 21.051327	8.939
<input checked="" type="checkbox"/>	7	ENSRNOG00000042932	Ttc27	tetratricopeptide repeat domain 27	Chr6: 21.735833	8.514
<input type="checkbox"/>	5	ENSRNOG00000032703	Rasgrp3	RAS guanyl releasing protein 3	Chr6: 21.072634	8.281
<input checked="" type="checkbox"/>	6	ENSRNOG00000005856	Slc30a6	solute carrier family 30 member 6	Chr6: 22.197003	7.658
<input type="checkbox"/>	3	ENSRNOG00000033090	Ltbp1	latent transforming growth factor beta binding protein 1	Chr6: 21.203502	5.329
<input checked="" type="checkbox"/>	4	ENSRNOG00000005810	Nlrc4	NLR family, CARD domain containing 4	Chr6: 22.167874	3.829

Showing 1 to 7 of 7 entries

**Slide 20. Top.** Review expression of candidates in brain regions—here in the core of the nucleus accumbens of HS rats

# Do any of these transcripts covary?

GeneNetwork X genenetwork.org/corr\_matrix +

genenetwork.org/corr\_matrix 110% Search

Lower-left cells list Pearson product-moment correlations, upper-right cells list Spearman rank order correlations. Each cell also contains the n of cases. Values ranging from 0.4 to 1.0 range from orange to white, while values ranging from -0.4 to -1.0 range from dark blue to white. Select any cell to generate a scatter plot. Select trait labels for more information.

		Spearman Rank Correlation (rho)						
		1	2	3	4	5	6	7
		✓ Short Labels	Long Labels					
Pearson	Trait 1: HSNIH-Rat-Acbc-RSeq-0818 ENSRNOG00000032703	n 77	0.03 77	-0.03 77	0.19 77	0.33 77	0.17 77	-0.01 77
	Gene Symbol: Rasgrp3							
	Trait 2: HSNIH-Rat-Acbc-RSeq-0818 ENSRNOG0000005856	0.08 77	n 77	0.00 77	-0.23 77	0.15 77	-0.06 77	-0.19 77
	Gene Symbol: Slc30a6							
	Trait 3: HSNIH-Rat-Acbc-RSeq-0818 ENSRNOG00000027191	-0.04 77	0.04 77	n 77	-0.14 77	0.21 77	0.19 77	-0.32 77
	Gene Symbol: Birc6							
	Trait 4: HSNIH-Rat-Acbc-RSeq-0818 ENSRNOG00000033090	0.18 77	-0.20 77	-0.03 77	n 77	0.02 77	0.07 77	0.13 77
	Gene Symbol: Ltbp1							
Spearman	Trait 5: HSNIH-Rat-Acbc-RSeq-0818 ENSRNOG0000005810	0.31 77	0.16 77	0.24 77	0.12 77	n 77	0.01 77	0.07 77
	Gene Symbol: Nlrc4							
	Trait 6: HSNIH-Rat-Acbc-RSeq-0818 ENSRNOG00000030328	0.19 77	-0.08 77	0.18 77	0.11 77	0.02 77	n 77	-0.02 77
	Gene Symbol: Fam98a							
	Trait 7: HSNIH-Rat-Acbc-RSeq-0818 ENSRNOG00000042932	-0.06 77	-0.13 77	-0.31 77	0.15 77	0.06 77	-0.03 77	n 77
	Gene Symbol: Ttc27							

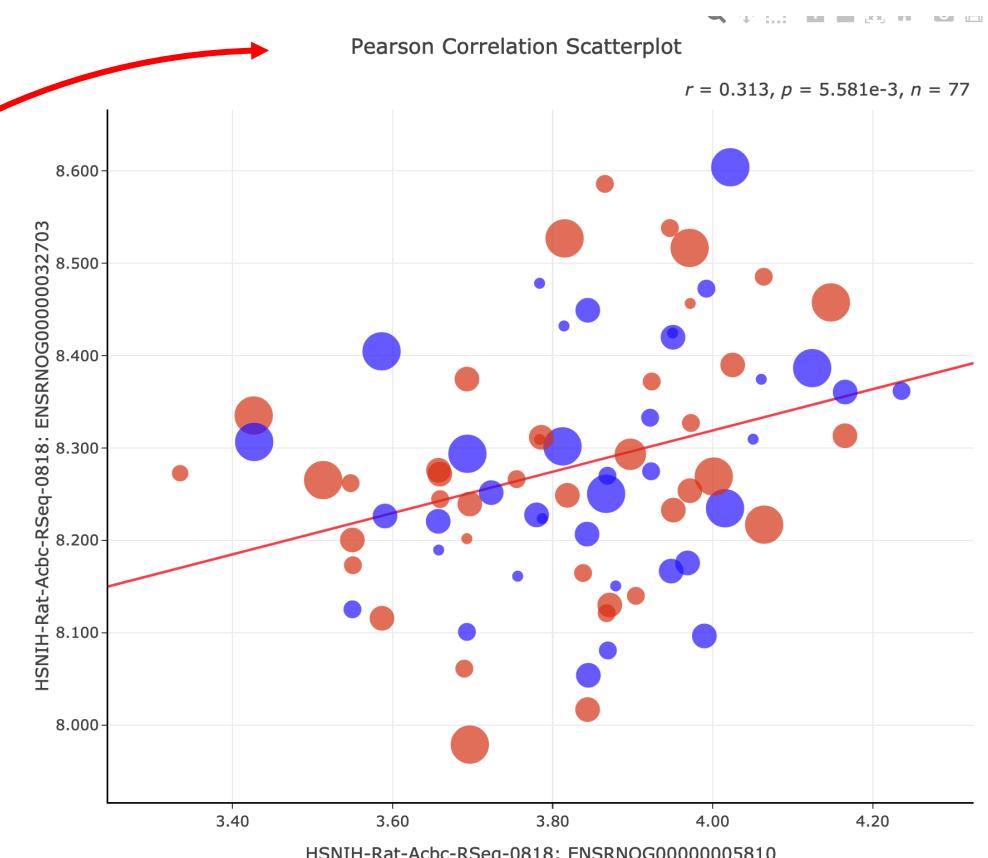
Chen\_HS\_Nicotine\_SA

Change Collection Name

This collection has eight records

Correlations Networks WebGestalt GeneWeaver BNW WGCNA CTL Maps

✓ Select All Copy Download CSV Search Table For ... Select Top ...



Slide 21. Top. Review expression of candidates in brain regions—here in the core of the nucleus accumbens of HS rats

# Relevant biology—links from GeneWiki

## Trait Data and Analysis for ENSRNOG00000042932

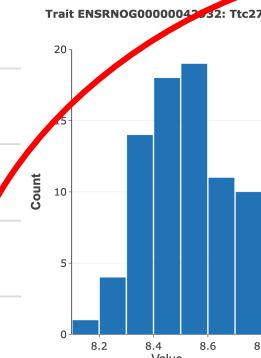
Tetratricopeptide repeat domain 27

### ▼ Details and Links

Group	Rat: HSNIH-Palmer group
Tissue	Nucleus Accumbens mRNA
Gene Symbol	Ttc27
Aliases	Wikidata: 2610511O17Rik; RGD1305264 GeneNetwork: Ttc27
Location	Chr 6 @ 21.735833 Mb on the minus strand
Database	HSNIH-Palmer Nucleus Accumbens Core RNA-Seq (Aug18) rlog
Resource Links	<a href="#">Gene</a> <a href="#">GeneMANIA</a> <a href="#">Protein Atlas</a> <a href="#">Rat Genome DB</a> <a href="#">GTEx Portal</a> <a href="#">BioGPS</a> <a href="#">STRING</a> <a href="#">PANTHER</a> <a href="#">Gemma</a> <a href="#">EBI GWAS</a>

Add Find GeneWiki SNPs Probes View in GN1

### ▼ Statistics



### GeneWiki for Ttc27: [New GeneWiki Entry](#)

#### GeneNetwork:

There is no GeneWiki entry for this gene.

#### GeneRIF from NCBI:

1. Clinical trial of gene-disease association and gene-environment interaction. (HuGE Navigator) ([Homo sapiens](#)) [PubMed](#)

A	B	C	D	E	F	G	H	I	J
<b>Genome-Wide Association for Smoking Cessation Success in a Trial of Precessation Nicotin</b>									
Abstract									
Abilities to successfully quit smoking and molecular genetic studies. Genetic single-nucleotide polymorphisms (SNPs) from individuals who were unable to quit smoking were compared with those from community samples. Many of the SNPs were located in regions of the genome that have been associated with smoking cessation success previously. Compared with chance, these results suggest that SNPs in these regions are more likely to be associated with smoking cessation success than those identified in six prior studies. These results may be useful for identifying individuals who are most likely to succeed in quitting smoking.									
3 George R Uhl, Tomas Drgn, Catherine Johnson, Marco F Ramoni, Frederique M Behm, and Jed E Rose									
4 Mol Med. 2010 Nov-Dec; 16(11-12): 513-526.									
5 Published online 2010 Aug 24. doi: 10.2119/molmed.2010.00052									
6 PMCID: PMC2972392									
7 PMID: 20811658									
8 <a href="https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2972392/">https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2972392/</a>									
9 Exported 21May2020 by RW Williams									
10									
11 <b>Table 1</b>									
12									
13 Genomic regions that contain clustered, nominally positive SNPs for success in smoking cessation.									
14 Columns list the chromosome and bp coordinates for the beginning and end of the genomic region identified by cluster analysis.									
15	Index	Chromosom	bp: Start	bp: End	No. SNPs	Gene(s)	P <sub>min</sub> SNP	P <sub>min</sub>	
16	1	1	4,514,682	4,527,839	5		rs241275	5.10E-04	
17	2	1	6,632,197	6,693,727	9	DNAJC11	rs7549198	9.61E-04	
18	3	1	10,231,218	10,268,359	5	KIF1B	rs17034615	4.40E-04	
19	4	1	23,620,353	23,638,820	6	DDEF1 and TCEA3	rs1077514	6.38E-03	
20	5	1	34,455,349	34,457,938	4	C1orf94	rs10158529	1.09E-03	
21	6	1	37,211,157	37,308,655	8	GRIK3	rs12077898	7.80E-05	
22	7	1	57,609,677	57,646,636	9	DAB1	rs2405994	8.50E-05	
23	8	1	67,964,539	67,970,259	5	GNG12	rs2803462	1.93E-03	
345	330	23	65,661,456	65,675,958	4		rs6624988	5.23E-03	
346	331	23	68,704,238	68,751,820	6	EDA	rs4844179	2.41E-03	
347	332	23	83,765,563	83,775,407	4		rs830240	3.81E-03	
348	333	23	83,802,507	83,866,535	12		rs707677	3.85E-03	
349	334	23	86,001,607	86,017,818	4		rs1936029	4.13E-03	
350	335	23	86,669,781	86,714,253	16	KLHL4	rs6617426	1.10E-03	
351	336	23	111,364,987	111,397,689	5	ZCCHC16	rs17307753	7.49E-03	
352	337	23	120,290,236	120,347,792	16		rs7054144	1.00E-03	
353	338	23	144,107,310	144,120,517	6		rs9792699	5.30E-03	
354									
355									

Slide 22. Left. Ttc27 mRNA trait page—expression data for the core of the nucleus accumbens of HS rats. Link from [GeneWiki](#) to a single [PubMed](#) article harvested daily from NCBI [Gene-Reference-Into-Function](#) (RIF).

# Relevant biology—links from GeneWiki

GeneWiki Entries

GeneWiki enables you to enrich the annotation of genes and transcripts. Please submit or edit a GeneWiki note (500 characters max) related to a gene, its transcripts, or proteins. When possible include PubMed identifiers or web resource links (URL addresses). Please ensure that the additions will have widespread use. For additional information, check the GeneWiki [help document](#).

**GeneWiki for Rasgrp3:** [New GeneWiki Entry](#)

**GeneNetwork:**  
There is no GeneWiki entry for this gene.

**GeneRIF from NCBI:**

- Expressed in developing blood vessels, Rasgrp3 contributes to the incidence of cardiovascular defects found in embryos from diabetic mothers. ([Mus musculus](#)) [PubMed](#)
- RasGRP1, but not RasGRP3, is required for thymocyte positive selection and invariant natural killer T cell selection. ([Mus musculus](#)) [PubMed](#)
- RasGRP3 limits inflammatory response by activating Rap1 on low-intensity pathogen infection, setting a threshold for preventing excessive inflammatory response. ([Mus musculus](#)) [PubMed](#)
- RasGRP1/3-deficient progenitors show impaired migration toward the CCR9 ligand, CCL25, suggesting that RasGRP1 and RasGRP3 may regulate progenitor entry into the thymus through a CCR9-dependent mechanism. ([Mus musculus](#)) [PubMed](#)
- These results indicate that RasGRP3 is implicated in phorbol ester-induced, PKC-independent exocytosis. ([Rattus norvegicus](#)) [PubMed](#)
- Together these data demonstrate a novel mechanism in which the balance between stability and plasticity in dendritic spines depends on binding of drebrin to actin filaments in a manner that is regulated by Ras. ([Rattus norvegicus](#)) [PubMed](#)
- Observational study and genome-wide association study of gene-disease association. (HuGE Navigator) ([Homo sapiens](#)) [PubMed](#)

> *Eur J Neurosci.* 2008 Jun;27(11):2847-59. doi: 10.1111/j.1460-9568.2008.06269.x.

## Interactions Between Drebrin and Ras Regulate Dendritic Spine Plasticity

Virginie Biou <sup>1</sup>, Heike Brinkhaus, Robert C Malenka, Andrew Matus

Affiliations + expand

PMID: 18588530 DOI: [10.1111/j.1460-9568.2008.06269.x](https://doi.org/10.1111/j.1460-9568.2008.06269.x)

### Abstract

Dendritic spines are major sites of morphological plasticity in the CNS, but the molecular mechanisms that regulate their dynamics remain poorly understood. Here we show that the association of drebrin with actin filaments plays a major role in regulating dendritic spine stability and plasticity. Overexpressing drebrin or the internal actin-binding site of drebrin in rat hippocampal neurons destabilized mature dendritic spines so that they lost synaptic contacts and came to resemble immature dendritic filopodia. Drebrin-induced spine destabilization was dependent on Ras activation: expression of constitutively active Ras destabilized spine morphology whereas drebrin-induced spine destabilization was rescued by co-expressing dominant negative Ras. Conversely, RNAi-mediated drebrin knockdown prevented Ras-induced destabilization and promoted spine maturation in developing neurons. Together these data demonstrate a novel mechanism in which the balance between stability and plasticity in dendritic spines depends on binding of drebrin to actin filaments in a manner that is regulated by Ras.

### Similar articles

[Drebrin A regulates dendritic spine plasticity and synaptic function in mature cultured hippocampal neurons.](#)

Slide 23. Left. *Rasgrp3 GeneWiki* page. RIF link to a paper on CNS plasticity.

## Trait Data and Analysis for ENSRNOG00000023688

### Dopamine receptor D1

▼ Details and Links	
<b>Group</b>	Rat: HSNIH-Palmer group
<b>Tissue</b>	Infralimbic Cortex mRNA
<b>Gene Symbol</b>	Drd1
<b>Aliases</b>	Wikidata: DADR; DRD1A; C030036C15Rik; Drd-1; Drd1a; Gpcr15; D1a GeneNetwork: Drd1
<b>Location</b>	Chr 17 @ 11.099736 Mb on the plus strand
<b>Summary</b>	induces dopamine sensitive adenylate cyclase activation; plays a role in regulating the expression of genes involved in the dopaminergic system.
<b>Database</b>	<a href="#">HSNIH-Palmer Infralimbic Cortex RNA-Seq (Aug18) rlog</a>
<b>Resource Links</b>	<a href="#">Gene</a> <a href="#">GeneMANIA</a> <a href="#">Protein Atlas</a> <a href="#">Rat Genome DB</a> <a href="#">GTEx Portal</a> <a href="#">PhenoGen</a> <a href="#">BioGPS</a> <a href="#">STRING</a> <a href="#">PANTHER</a> <a href="#">Gemma</a> <a href="#">EBI GWAS</a>

[Gene](#) [GeneMANIA](#) [Protein Atlas](#) [Rat Genome DB](#) [GTEx Portal](#) [PhenoGen](#)  
[BioGPS](#) [STRING](#) [PANTHER](#) [Gemma](#) [EBI GWAS](#)

# Relevant biology—links from GeneWiki

## Trait Data and Analysis for ENSRNOG00000042932

### Tetratricopeptide repeat domain 27

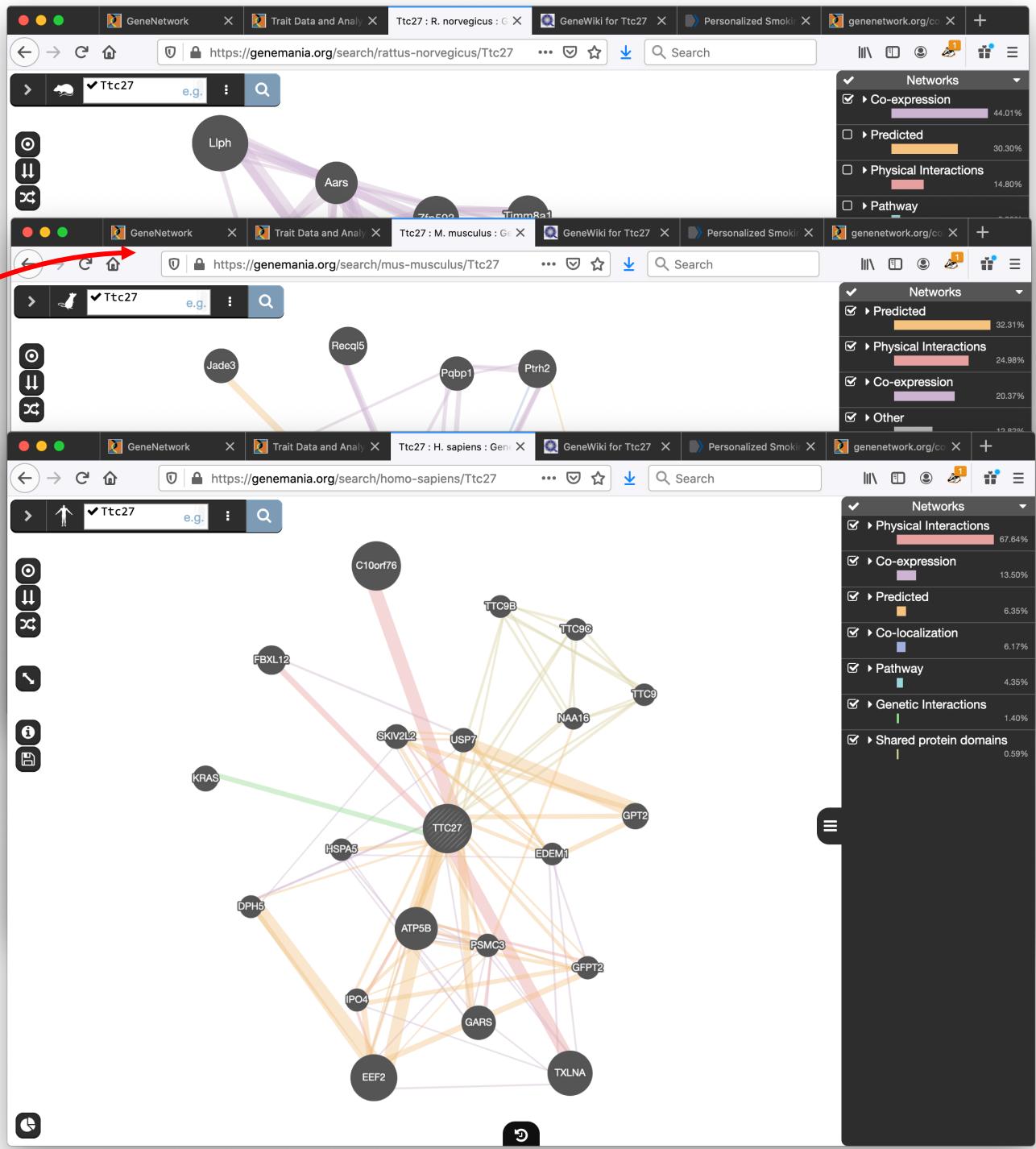
#### ▼ Details and Links

Group	Rat: HSNIH-Palmer group
Tissue	Nucleus Accumbens mRNA
Gene Symbol	Ttc27
Aliases	Wikidata: 2610511O17Rik; RGD1305264 GeneNetwork: Ttc27
Location	Chr 6 @ 21.735833 Mb on the minus strand
Database	HSNIH-Palmer Nucleus Accumbens Core RNA-Seq (Aug18) rlog
Resource Links	<a href="#">Gene</a> <a href="#">GeneMANIA</a> <a href="#">Protein Atlas</a> <a href="#">Rat Genome DB</a> <a href="#">GTEx Portal</a> <a href="#">BioGPS</a> <a href="#">STRING</a> <a href="#">PANTHER</a> <a href="#">Gemma</a> <a href="#">EBI GWAS</a>

[Add](#)[Find](#)[GeneWiki](#)[SNPs](#)[Probes](#)[View in GN1](#)

#### ▼ Statistics

Slide 25. *Left.* Ttc27 GeneMania pages.



# Links to EBI GWAS (for TTC27)

The screenshot shows the EBI GWAS Catalog page for the gene TTC27. The table displays the following data:

Variant and risk allele	P-value	P-value annotation	RAF	OR	Beta	CI	Mapped gene	Reported trait	Trait(s)	Study accession
rs1031261-C	$2 \times 10^{-6}$	(White matter hyperintensities)	0.11	-	-	-	TTC27	Hippocampal atrophy	hippocampal atrophy	GCST001530
rs2710619-T	$2 \times 10^{-9}$		NR	-	-	-	TTC27	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference	GCST009867
rs72787555-G	$2 \times 10^{-20}$		NR	-	0.145315 unit decrease	NR	TTC27	Male-pattern baldness	Alopecia	GCST006661
rs72787535-T	$2 \times 10^{-19}$		NR	-	0.144612 unit decrease	NR	TTC27	Male-pattern baldness	Alopecia	GCST006661
rs150489034-T	$5 \times 10^{-18}$		NR	-	0.14408 unit decrease	NR	TTC27	Male-pattern baldness	Alopecia	GCST006661
rs7599125-G	$9 \times 10^{-10}$		0.44	-	0.23 unit decrease	-	TTC27	Blood protein levels in cardiovascular risk	interleukin 18 measurement	GCST009731
rs72787520-G	$4 \times 10^{-20}$		NR	-	0.145234 unit decrease	NR	TTC27, BIRC6	Male-pattern baldness	Alopecia	GCST006661
rs2568464-T	$1 \times 10^{-8}$		0.5731	-	0.013693864 unit increase	[0.009-0.0184]	BIRC6, TTC27	Number of sexual partners	risky sexual behaviour measurement	GCST007326
rs6723226-A	$4 \times 10^{-8}$		NR	-	0.017145976	[0.011-0.023]	BIRC6, TTC27	Intelligence	intelligence	GCST005316

# Links to PheWAS (for TTC27)

pheweb.sph.umich.edu/gene/TTC27

phewas catalog

## TTC27

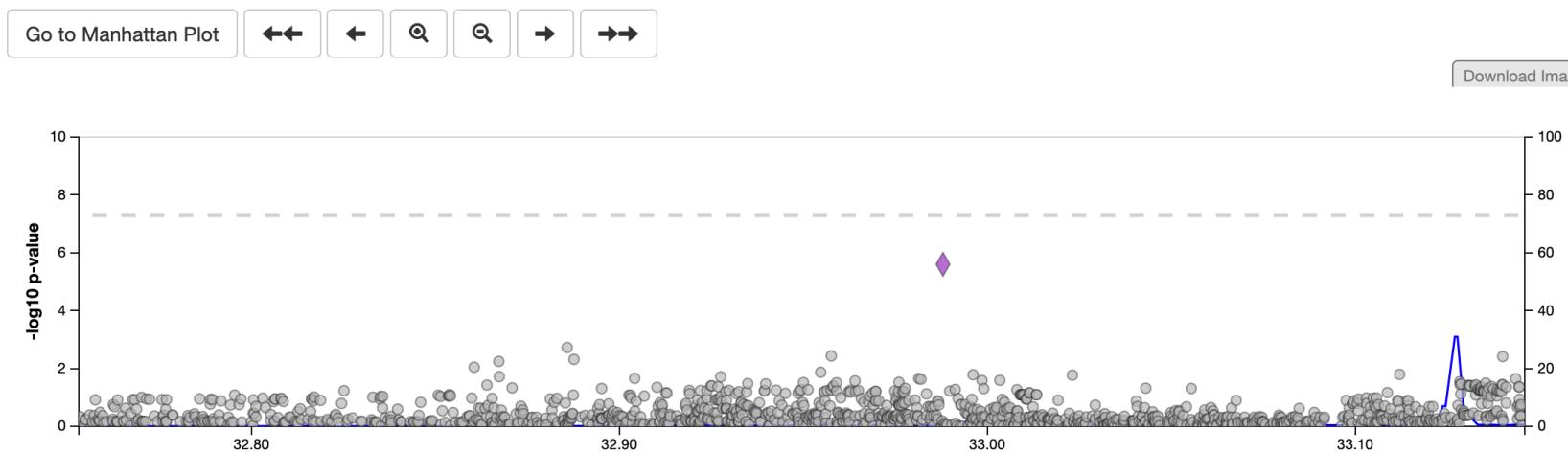
Phenotypes with the most-significant associations for this locus:

Top p-value in gene	Phenotype
2.5e-6	Personal history of allergy to medicinal agents
8.6e-6	Mycoses
8.6e-6	Optic atrophy
9.1e-6	Delirium dementia and amnestic and other cognitive disorders

Warning: This phenotype only has 182 cases.

182 cases, 14720 controls.

Category: injuries & poisonings

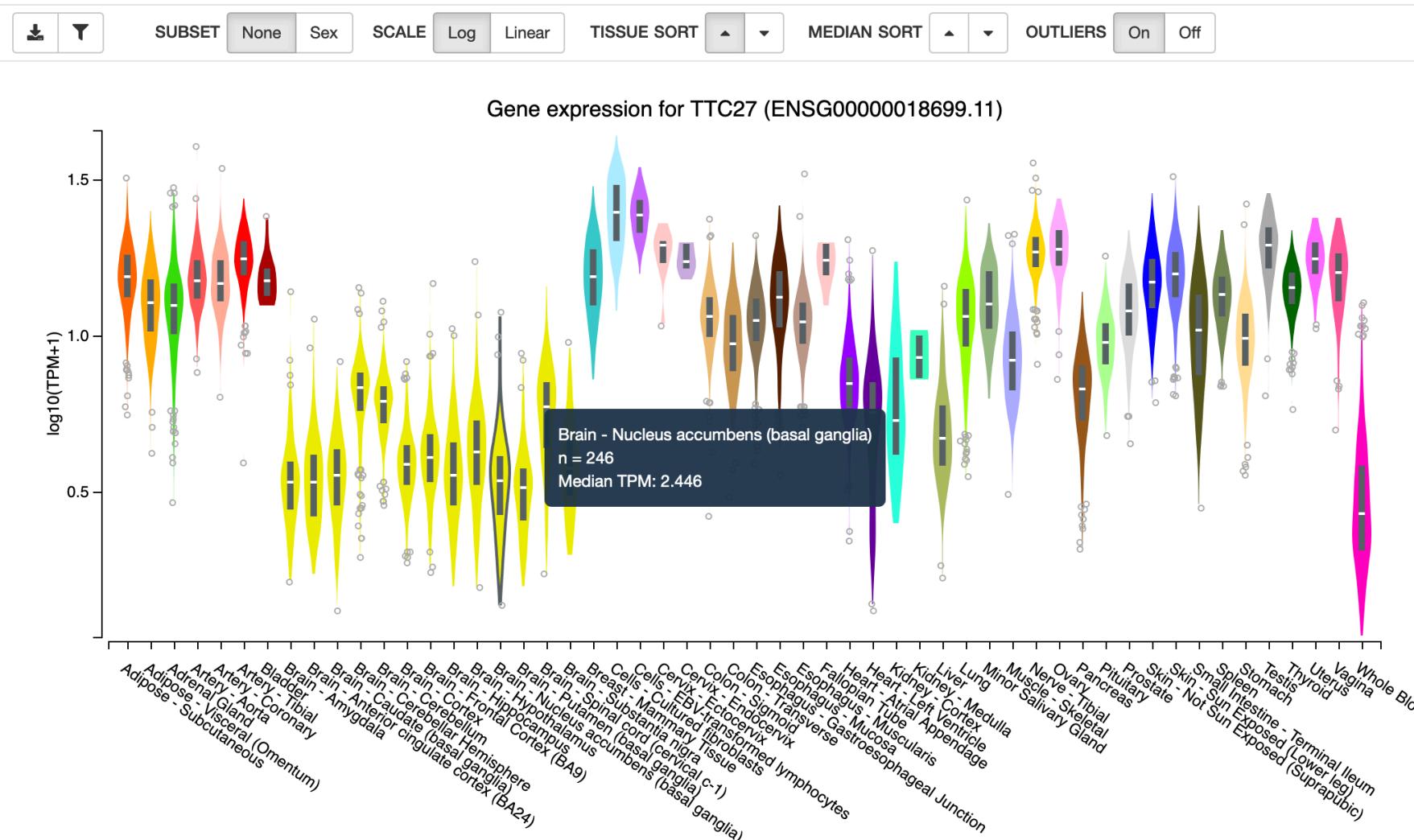


# Links to GTExPortal.org

## Gene expression for TTC27 (ENSG00000018699.11)

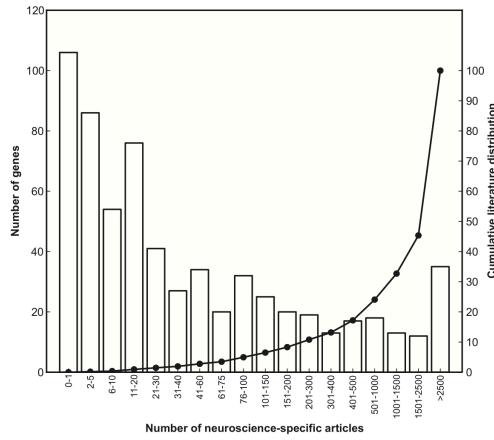
Data Source: GTEx Analysis Release V8 (dbGaP Accession phs000424.v8.p2)

Data processing and normalization [i](#)



## Summary comments

1. Please send me bug reports or suggestions for improvements or additions
2. Please contact me if you have data sets and projects that might be suitable for entry into GeneNetwork (human, rat, mouse, drosophila....)
3. Candidate gene evaluation really needs an AI system with tunable parameters
4. We are still comparatively ignorant about the functions of most genes and their many variants and isoforms

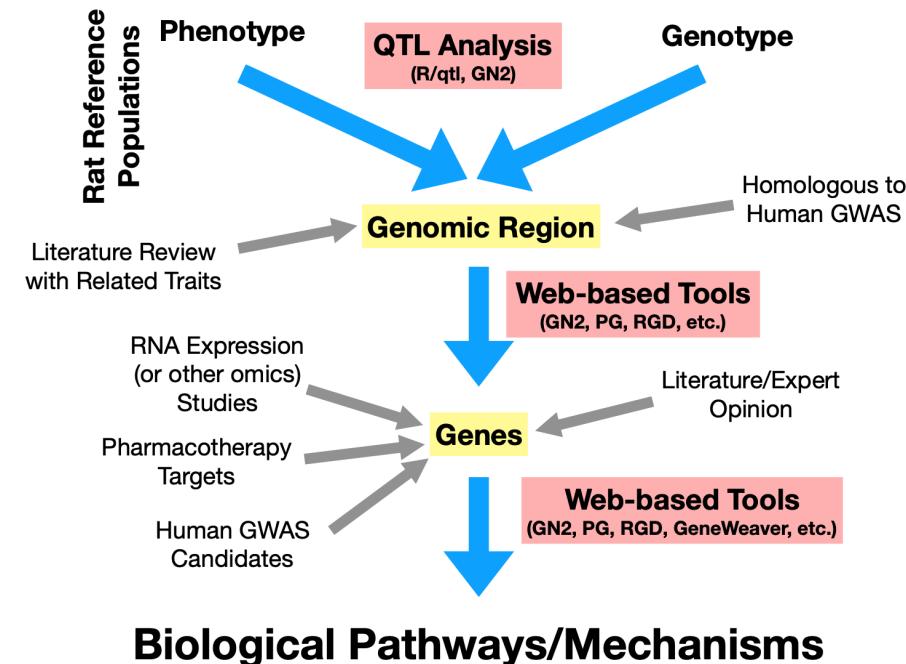


**Not covered yet—should be full talk**

Sequence analysis of gene variants and eQTLs within loci.

Functionally enigmatic genes: A case study of the brain ignorome

- 1 Quick review of rat, mouse and human molecular, SUD, and behavioral data in [GeneNetwork.org](#)
- 2 How to use GeneNetwork to **review and edit trait** data
- 3 How to **map QTLs**
- 4 How to **evaluate and rank** candidate genes (at least some qualitative methods)



# Some examples

*Comt*

## A Transposon in *Comt* Generates mRNA Variants and Causes Widespread Expression and Behavioral Differences among Mice

Zhengsheng Li<sup>1\*</sup>, Megan K. Mulligan<sup>1\*</sup>, Xusheng Wang<sup>1</sup>, Michael F. Miles<sup>2</sup>, Lu Lu<sup>1</sup>, Robert W. Williams<sup>1\*</sup>

<sup>1</sup> Department of Anatomy and Neurobiology, Center for Integrative and Translational Genomics, University of Tennessee Health Science Center, Memphis, Tennessee, United States of America, <sup>2</sup> Department of Pharmacology and Toxicology, Virginia Commonwealth University, Richmond, Virginia, United States of America

*Gabra2*

## Identification of a Functional Non-coding Variant in the GABA<sub>A</sub> Receptor α2 Subunit of the C57BL/6J Mouse Reference Genome: Major Implications for Neuroscience Research

Megan K. Mulligan<sup>1\*</sup>, Timothy Abreo<sup>1</sup>, Sarah M. Neuner<sup>2,3</sup>, Cory Parks<sup>1</sup>, Christine E. Watkins<sup>1</sup>, M. Trevor Houseal<sup>1</sup>, Thomas M. Shapaker<sup>1</sup>, Michael Hook<sup>1</sup>, Haiyan Tan<sup>4</sup>, Xusheng Wang<sup>4</sup>, Jesse Ingels<sup>1</sup>, Junmin Peng<sup>4</sup>, Lu Lu<sup>1</sup>, Catherine C. Kaczorowski<sup>3</sup>, Camron D. Bryant<sup>5</sup>, Gregg E. Homanics<sup>6</sup> and

[Taar1 gene variants have a causal role in methamphetamine intake and response and interact with Oprm1.](#)

Stafford AM, Reed C, Baba H, Walter NA, Mootz JR, Williams RW, Neve KA, Fedorov LM, Janowsky AJ, Phillips TJ.

Elife. 2019 Jul 9;8:e46472. doi: 10.7554/Elife.46472.

PMID: 31274109    [Free PMC article.](#)

Both methamphetamine intake and the thermal response mapped to **Taar1** and the independent effect of **Taar1** was dependent on genotype at Oprm1. Our findings encourage investigation of the contribution of **Taar1** and Oprm1 variants to human methamphetamine addiction ...

Slide 30. Successes—from QTLs to functional variants

The screenshot shows the GeneNetwork web interface. At the top, there are navigation links for GeneNetwork, Intro, Help, Tools, Collections (with a count of 4), Source Code, and Sign in. Below this is a search bar labeled "Genes / Molecules". To the right is a dropdown menu for "Species" set to "Rat (rn6)", with other options like "Human", "Mouse", and "Drosophila" available. The menu also includes links to "References", "Tutorials/Primers", "Glossary of Terms", "FAQ", "Policies", "Links", "Environments", and "GN1 News".

**Select and se**

Species: Rat (rn6)

*Fmn2*

## Dissection of a QTL Hotspot on Mouse Distal Chromosome 1 that Modulates Neurobehavioral Phenotypes and Gene Expression

Khyobeni Mozhui, Daniel C. Ciobanu, Thomas Schikorski, Xusheng Wang, Lu Lu, Robert W. Williams, Zhengsheng Li, Michael F. Miles, M. Trevor Houseal, Thomas M. Shapaker, Michael Hook, Haiyan Tan, Cory Parks, Sarah M. Neuner, Christine E. Watkins, Camron D. Bryant, Gregg E. Homanics, Catherine C. Kaczorowski, Timothy Abreo, and Megan K. Mulligan

*Pcdh15*

## Quantitative trait loci mapping and gene network analysis implicate protocadherin-15 as a determinant of brain serotonin transporter expression

R. Ye<sup>†</sup>, A. M. D. Carneiro<sup>†</sup>, Q. Han<sup>†</sup>, D. Airey<sup>†</sup>, E. Sanders-Bush<sup>†</sup>, B. Zhang<sup>‡</sup>, L. Lu<sup>§</sup>, R. Williams<sup>§</sup> and R. D. Blakely<sup>\*,†,||</sup>

Keywords: Antidepressant, BXD, microarray, network, protocadherin, QTL, serotonin, SLC5A7, SLC6A4, SLC6A3, TPH2, transcriptome, transporter, VMAT2

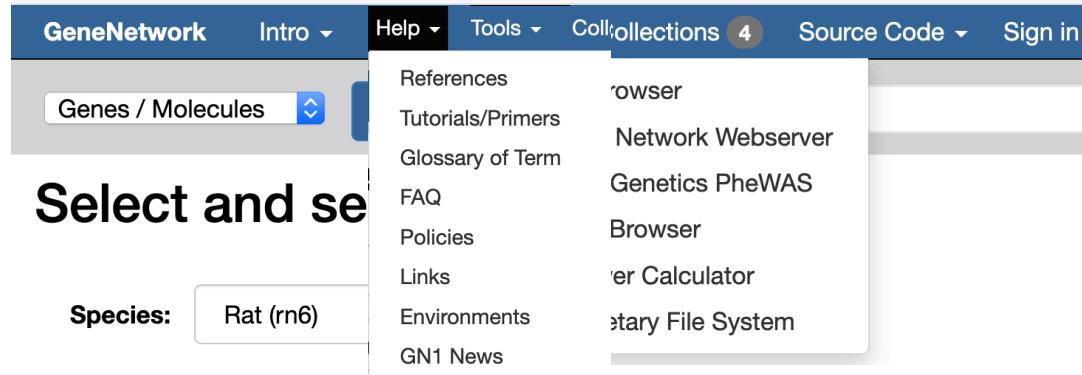
*Cacna2d1*

DOI: 10.1038/s41467-017-00837-5    OPEN

## Systems genetics identifies a role for *Cacna2d1* regulation in elevated intraocular pressure and glaucoma susceptibility

Sumana R. Chintalapudi<sup>1</sup>, Doaa Maria<sup>1,2,3</sup>, Xiang Di Wang<sup>1</sup>, Jessica N. Cooke Bailey<sup>4</sup>, NEIGHBORHOOD consortium, International Glaucoma Genetics consortium, Pirro G. Hysi<sup>5</sup>, Janey L. Wiggs<sup>6</sup>, Robert W. Williams<sup>1,7,8</sup> & Monica M. Jablonski<sup>1,2,7</sup>

# Test questions



The screenshot shows the GeneNetwork website's header with links for GeneNetwork, Intro, Help, Tools, Collections (with 4 items), Source Code, and Sign in. A dropdown menu is open under 'Help' for 'Select and search'. It includes options like Species (set to Rat (rn6)), Genes / Molecules, References, Tutorials/Primers, Glossary of Terms, FAQ, Policies, Links, Environments, and GN1 News.

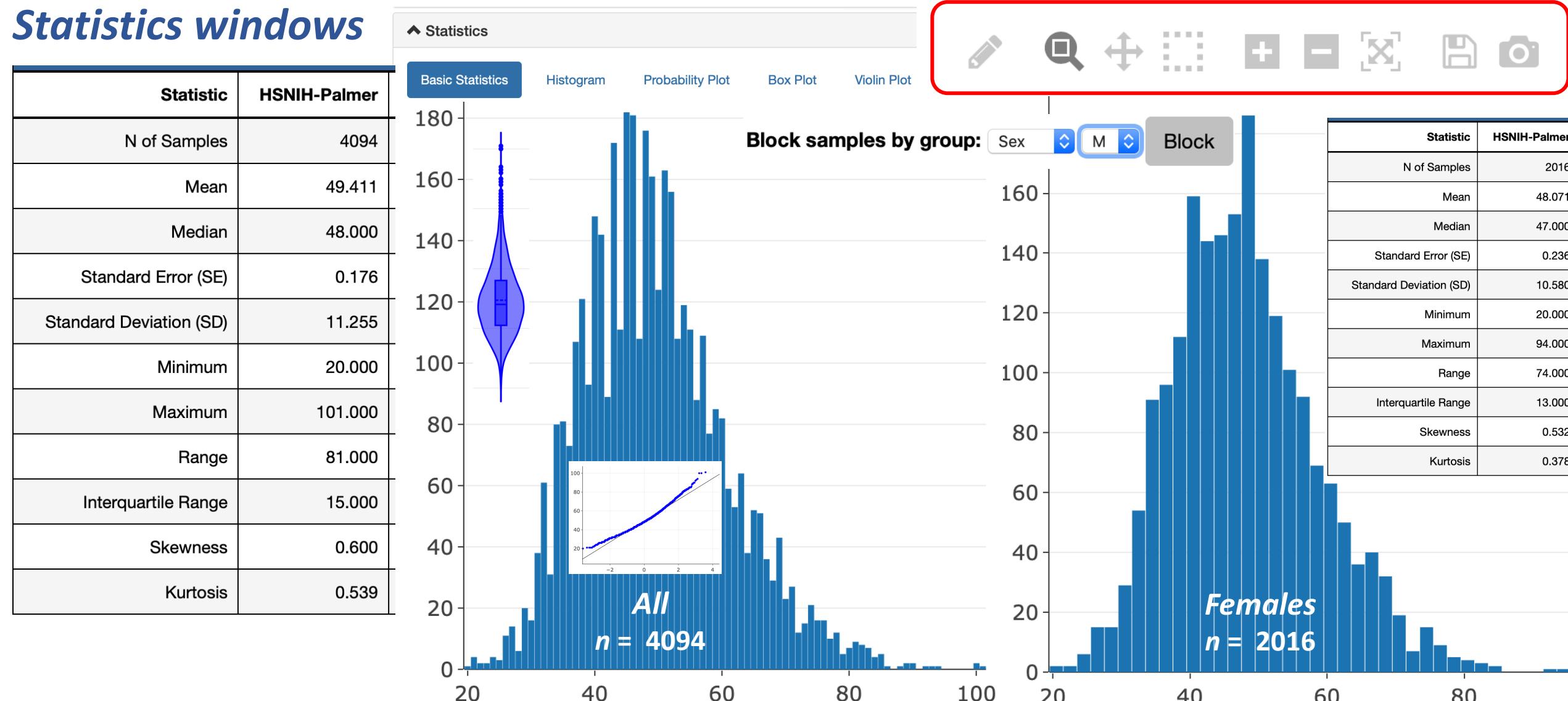
1. Find body weight data for weaning rats in the HS rats. Is the data normal enough to map?
2. The P50 Rat GWAS has several major sites. Map body weight of weanling using both sex and site as covariates. Map males and females separately and then together. How do results compare?
3. What is the correlation of expression of Drd1 and Drd2 in the infralimbic cortex of the HS rats?
4. Map stress response (corticosterone levels) 20 minutes after immobilization stress in young male rats using the HRDP HXB family AND using R/qtl. Where is your QTL? How sensitive is the result to different data transformations?
5. Organize and rank a list of genes within the QTL for the serum cort level QTL. (Or is it too big to bother with?)
6. Map morphine locomotor response 45–60 minutes after an injection of 50 mg/kg in male and female BXD mice. Define the best candidate gene.

[https://genenetwork.org/show\\_trait?trait\\_id=10441&dataset=HSNIH-PalmerPublish](https://genenetwork.org/show_trait?trait_id=10441&dataset=HSNIH-PalmerPublish)

[https://genenetwork.org/show\\_trait?trait\\_id=10061&dataset=HXBBXHPublish](https://genenetwork.org/show_trait?trait_id=10061&dataset=HXBBXHPublish) (LOD 4.37, Chr 11 at 69–78 Mb, ~120 genes)

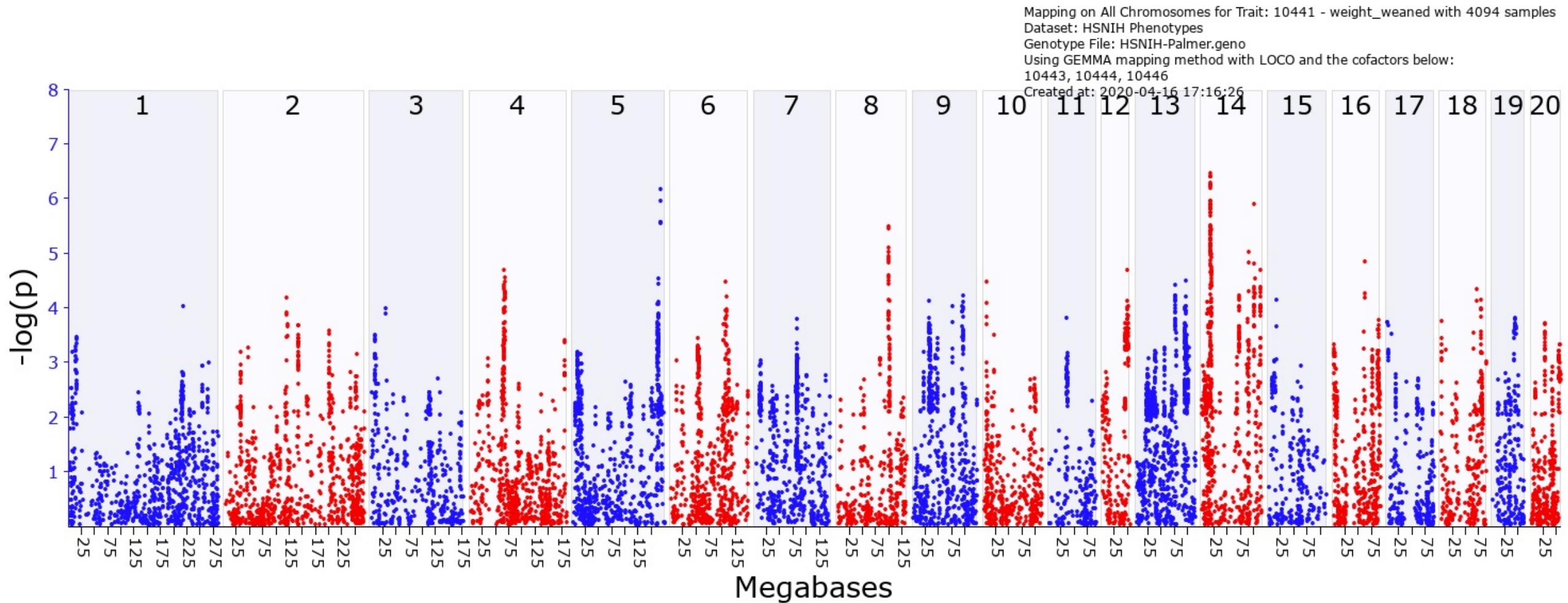
[https://genenetwork.org/show\\_trait?trait\\_id=11845&dataset=BXDPublish](https://genenetwork.org/show_trait?trait_id=11845&dataset=BXDPublish)

# Statistics windows



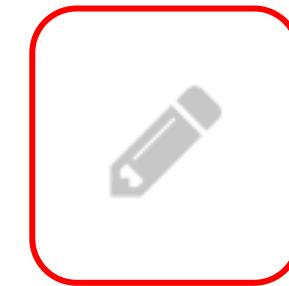
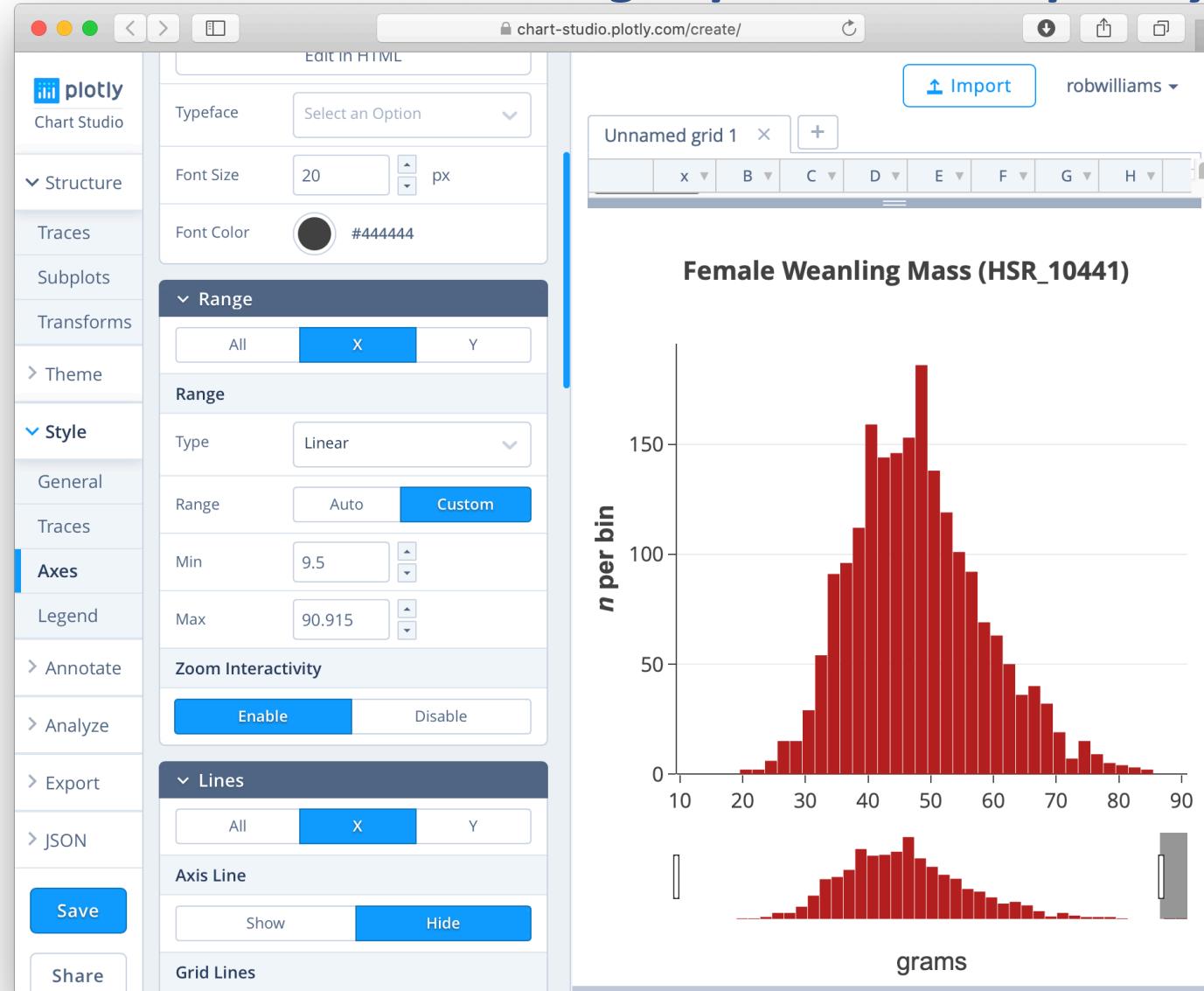
**Slide 32. Statistics.** **Left.** Summary table. **Middle.** Histogram of weanling weight generated using the **Statistics Histogram** tab. **Insets.** Quantile probability plots in the two histograms, a *violin plot*, and statistics. **Top.** Graphing tools are highlighted in the **red box**: **Edit**, **Pan**, **Zoom In**, **Zoom Out**, **Autoscale**, **Reset**, **Export SVG**, and **Export JPEG**. **Left.** Histogram of 2016 females. To analyze only females, **Block samples by group: Sex = M**

## *Maps for body weight with cofactors with GEMMA (about 45 minute job)*



**Slide 33. Top.** Maps for both males and females with three cofactors: *batch*, *sex*, and *center* (HSR\_10443, 10444, 10446). The x-axis lists chromosomes on a megabase (Mb) scale, except Chr X and Y. The y-axis is minus the logarithm of the linkage probability ( $-\log p$ ). The Chr 14 locus has a peak linkage of about 6.4. These maps can be zoomed to single chromosome level or even small parts of chromosomes.

# Edit in ChartStudio to get publication quality



**Slide 34.** Any of the statistic plots generated by **Plotly** can be modified in **Chart Studio**. Click on the **Edit pencil** icon. In this example, a default histogram has been modified in many small ways, including bar color, axes labels, tick marks, and the x-axis has been modified to show bins between 10 and 90 grams only. Free.

# HS genotypes are complex

## Trait Data and Analysis for chr6:21892556

Not available

### ▼ Details and Links

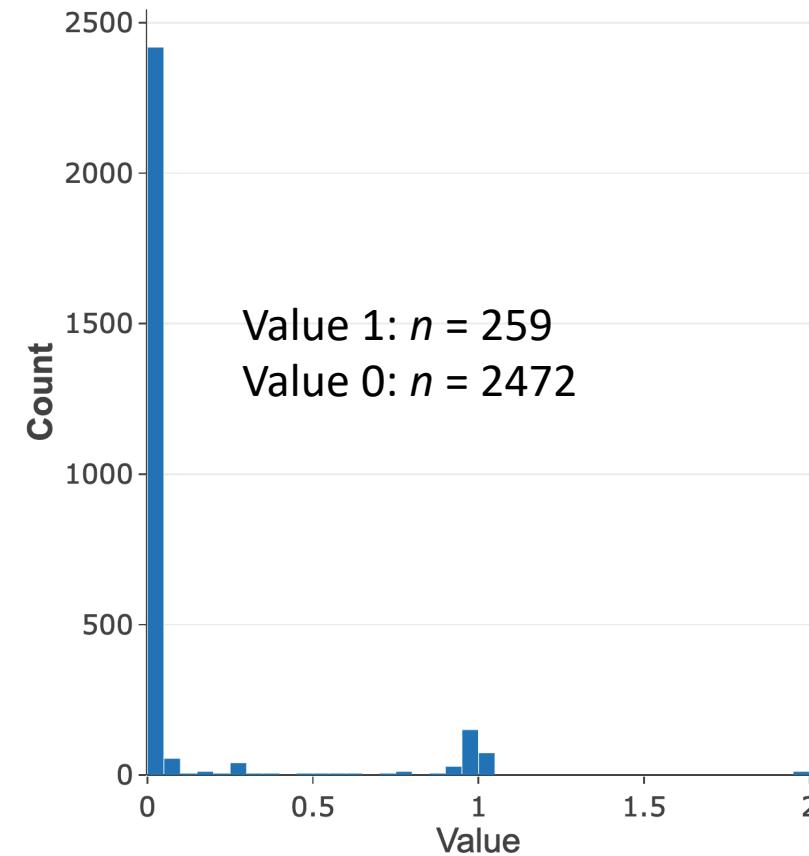
Group	Rat: HSNIH-Palmer group
Tissue	N/A
Location	Chr 6 @ 21.892556 Mb
Database	NIH Heterogeneous Stock (Palmer) Genotypes

Add

View in GN1

Marker chr6:21892556	
Statistic	Value
N of Samples	2807
Mean	0.116
Median	0.013
Standard Error (SE)	0.006
Standard Deviation (SD)	0.306
Minimum	0.003
Maximum	1.978
Range	1.975
Range (fold)	1.975
Interquartile Range	0.019
Skewness	2.973
Excess Kurtosis	8.426

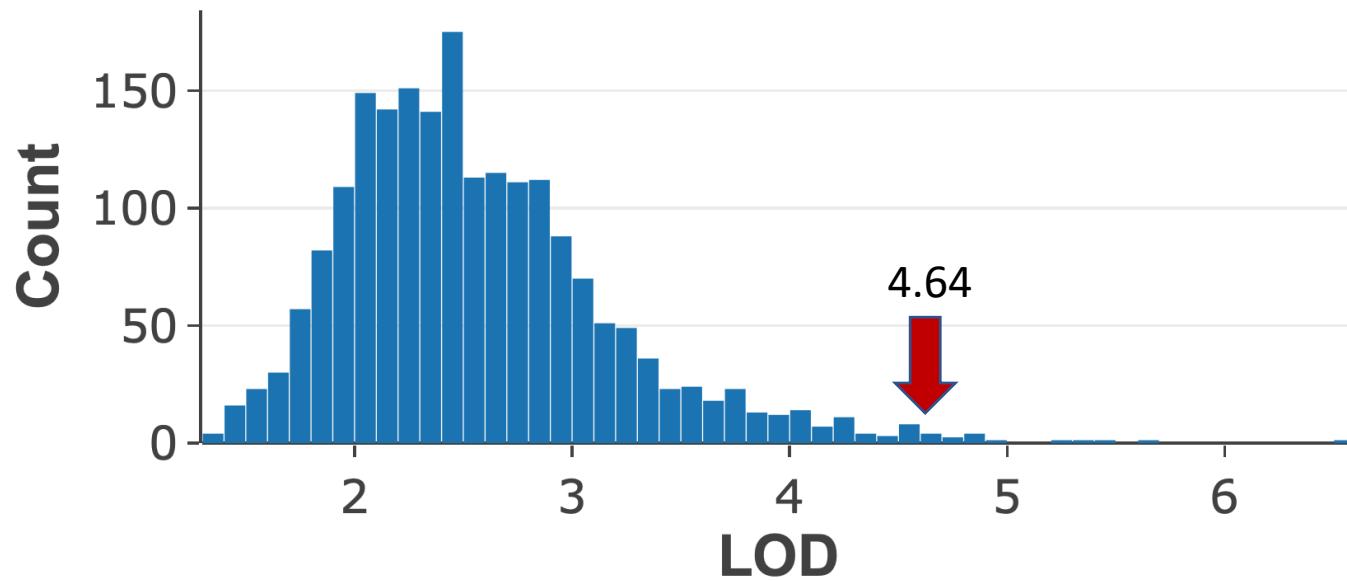
[genenetwork.org/show\\_trait?trait\\_id=chr6:21892556  
&dataset=HSNIH-PalmerGeno#redirect](http://genenetwork.org/show_trait?trait_id=chr6:21892556&dataset=HSNIH-PalmerGeno#redirect)



**Slide 35. Left.** Table of the  $-\log p$  linkage scores sorted from high to low by marker and position. The highest score for body weight is 6.4 on Chr 14 at [16.851818](#) at Mb. **Middle** and **Right**. A 1.5 confidence interval drop in both proximal (top arrows) and distal (bottom arrows) directions from the peak is used to define an approximate confidence interval (CI). Sets of red arrows mark approximate 1.5 CIs.

# Stratified permutation testing

Total of 2000 permutations [Download Permutation Results](#)



**Slide 36. Left.** Peak linkage computed for 2000 stratified permutations of phenotypes within **cofactors** when using R/qtl. In this example, only 15 of 2000 permutations have LOD peaks greater than 4.64. The genome-wide alpha of linkage to Chr 11 is therefore  $p < 0.01$ . **Right.** The permutation results can be downloaded as comma-separated-value text files. Here data have been entered into Excel to highlight the few permutations that exceed the value of the locus of 4.64. Note that each permutation run will provide somewhat different thresholds. Before publication compute thresholds using at 1000 – 2000 permutations—currently an overnight calculation of HS rats.

2000 Permutations		
Index	LOD_Max	<i>p</i>
1	1.3109	1.00000
2	1.3352	0.99950
3	1.3596	0.99900
4	1.3789	0.99850
5	1.4113	0.99800
1983	4.5770	0.00900
1984	4.5775	0.00850
<b>1985</b>	<b>4.6015</b>	<b>0.00800</b>
1986	4.6567	0.00750
1987	4.6662	0.00700
1988	4.6808	0.00650
1989	4.7241	0.00600
1990	4.7418	0.00550
1991	4.8167	0.00500
1992	4.8175	0.00450
1993	4.8497	0.00400
1994	4.8967	0.00350
1995	4.9994	0.00300
1996	5.2717	0.00250
1997	5.3127	0.00200
1998	5.4980	0.00150
1999	5.6551	0.00100
2000	6.5332	0.00050

# HS expression data in GeneNetwork: a ton of RNA-seq data

**Group:** NIH Heterogeneous Stock (Palmer)

**Type:** Nucleus Accumbens mRNA

**Dataset:** HSNIH-Palmer Nucleus Accumbens Core RNA-Seq (Aug18) rlog

**Get Any:** endorphin dopamine acetylcholine 5HT

**Search Results:** We searched HSNIH-Palmer Nucleus Accumbens Core RNA-Seq (Aug18) rlog to find all records that "endorphin" and "dopamine" and "acetylcholine" and "5HT"

8 records are shown below.

Correlations Networks WebGestalt GeneWeaver BNW WGCNA CTL Maps MultiMap Comparisor

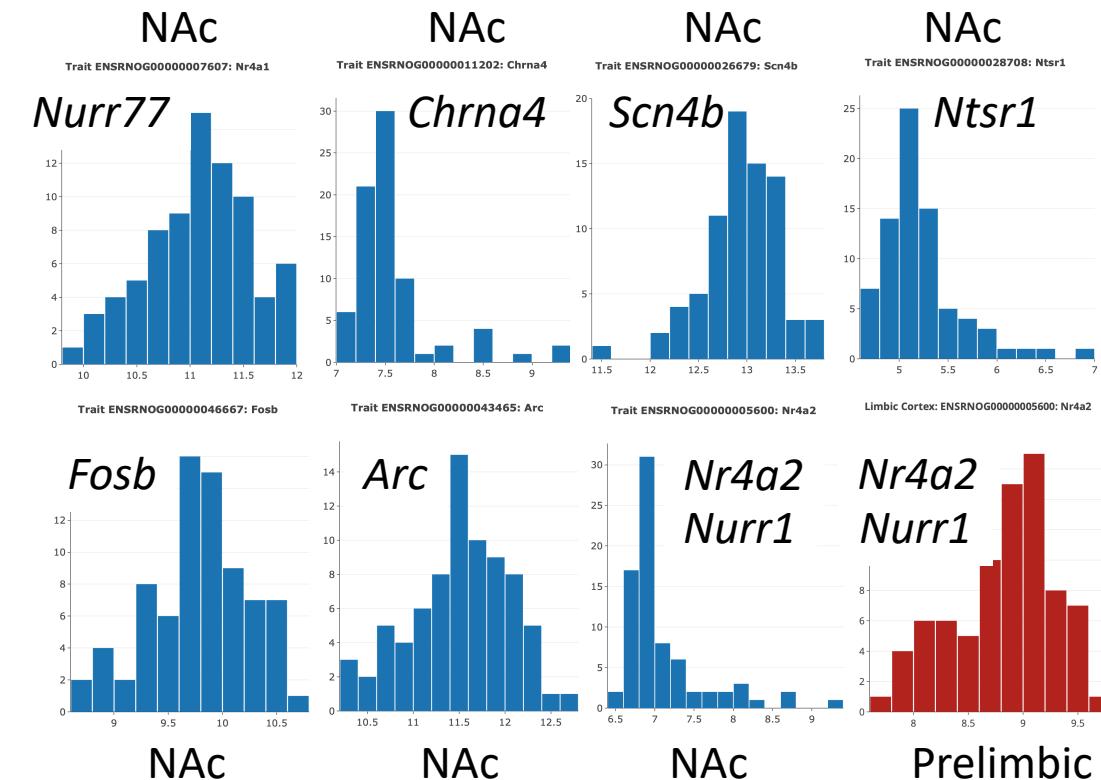
Select  Download CSV Search This Table For ... Select Top ...

Show/Hide Columns:

Index Record Symbol Description Location Mean Max LRS? Max LRS Location Additive Effect?

Showing 1 to 8 of 8 entries

	Index	Record	Symbol	Description	Location	Mean
<input type="checkbox"/>	1	ENSRNOG00000026493	Cdnf	cerebral dopamine neurotrophic factor	Chr17: 78.721230	3.924
<input type="checkbox"/>	2	ENSRNOG00000006641	Dbh	dopamine beta-hydroxylase	Chr3: 5.709236	1.552
<input type="checkbox"/>	3	ENSRNOG00000023688	Drd1	dopamine receptor D1	Chr17: 11.099736	11.234
<input type="checkbox"/>	4	ENSRNOG00000008428	Drd2	dopamine receptor D2	Chr8: 53.678777	11.279
<input type="checkbox"/>	5	ENSRNOG000000060806	Drd3	dopamine receptor D3	Chr11: 61.819102	3.652
<input type="checkbox"/>	6	ENSRNOG00000017927	Drd4	dopamine receptor D4	Chr1: 214.278296	0.503
<input type="checkbox"/>	7	ENSRNOG00000005338	Drd5	dopamine receptor D5	Chr14: 77.220579	4.060
<input type="checkbox"/>	8	ENSRNOG00000014691	Ric3	RIC3 acetylcholine receptor chaperone	Chr1: 173.629218	9.803



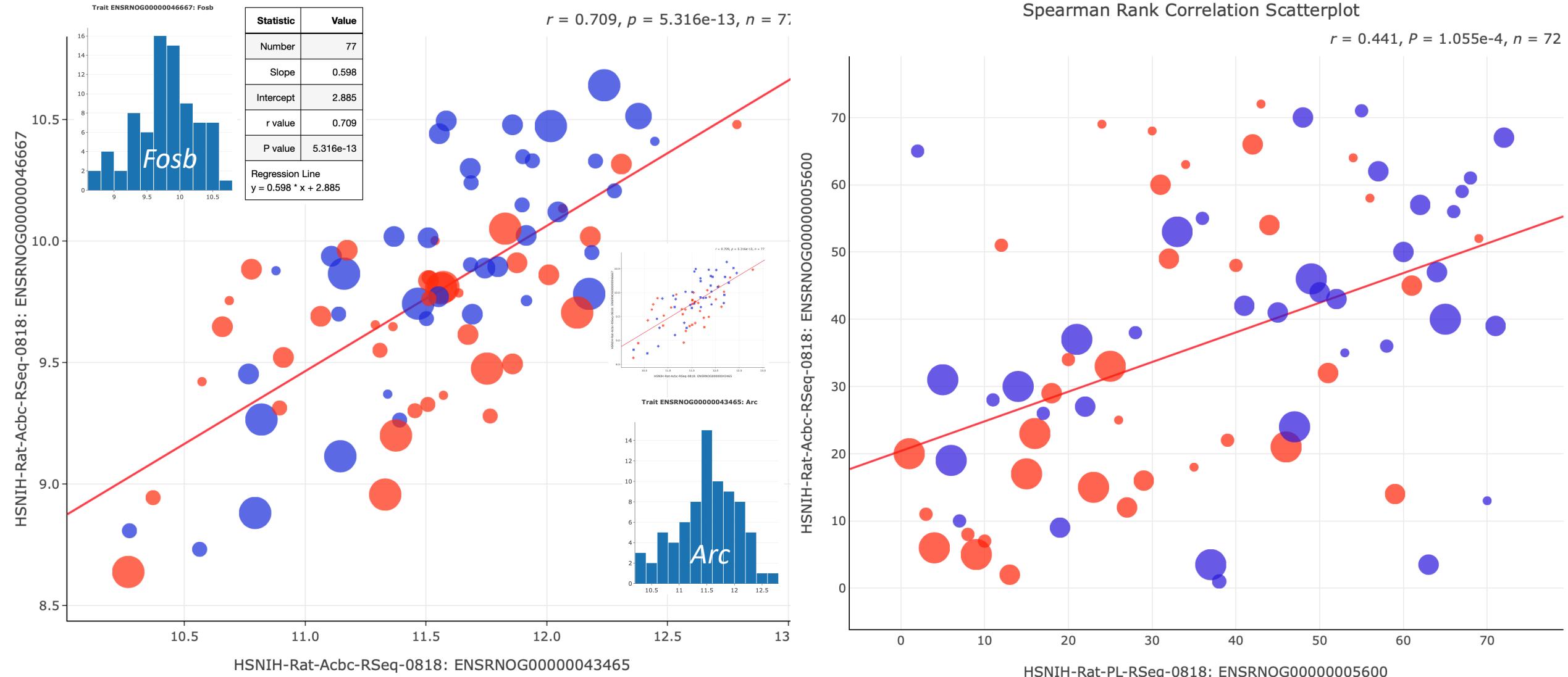
## Computing the Maps

$n = 77$   
Method = GEMMA  
MAF  $\geq 0.05$



**Slide 37. Left.** The same tools work with all HS omics data. Expression distribution require care, but are highly revealing. What are the correlations of **Arc** to **Fosb** in nucleus accumbens (Nac)? What is the correlation of **Nr4a2** (aka **Nurr1**) in NAc and prelimbic cortex (**red** bimodal histogram)?

# HS expression covariation: 4D scatterplots of Pearson r and Spearman rho



**Slide 38. Left.** *Arc* and *Fosb* are both immediate-early genes and have Pearson's correlation of 0.7. Males in blue, females red, and the batches as circle size 1, 2, 3, 4. **Right.** Spearman rank correlation of 0.44 between the key dopaminergic gene *Nurr1* (aka *Nr4a2*) in prelimbic cortex (x-axis) and NAc (y-axis).

# Nr4a2 expression in prelimbic cortex: what other genes covary?

## Calculate Correlations

Method: Sample r

Database: HSNIH-Palmer Prelimbic Cortex RNA-Seq

Return: Top 500

Samples: HSNIH-Palmer

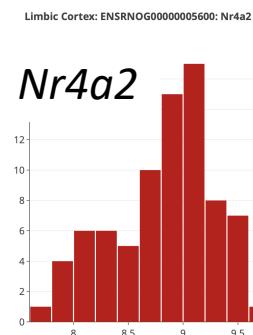
Type: Pearson

Min Expr: 5

Location: Chr: [ ] Mb: [ ] to [ ]

Range: [-1.00, 1.00]

Compute



Prelimbic

## Correlation Table

Trait: ENSRNOG00000005600

Values of record ENSRNOG00000005600 in the HSNIH-Palmer Prelimbic Cortex RNA-Seq (Aug18) rlog dataset were compared to all records in the HSNIH-Palmer (Aug18) rlog dataset. The top 500 correlations ranked by the Pearson's r are displayed. You can resort this list by clicking the headers. Select the first row to go to the data and analysis page.

Correlations Networks WebGestalt GeneWeaver BNW WGCNA CTL Maps MultiMap Comparison Bar Chart

Select All  Add Download CSV Search Table For ... Select Top ...  Deselect Reset Columns

More Options...

Show/Hide Columns:

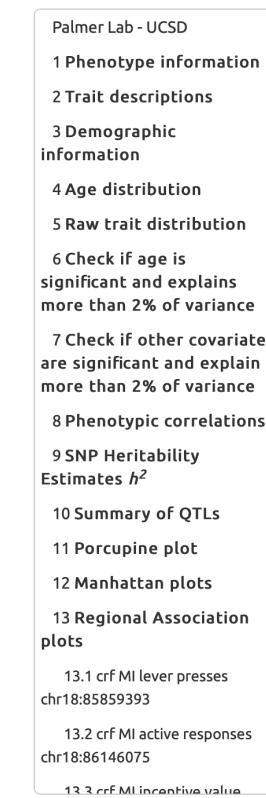
Index	Record	Symbol	Description	Location	Mean	Sample r	N	Sample p(r)	Lit r	Tissue r	Tissue p(r)	Max LRS	Max LRS L
1	ENSRNOG00000005600	Nr4a2	nuclear receptor subfamily 4, group A, member 2	Chr3: 43.111258	8.829	1.000	80						
2	ENSRNOG00000005964	Nr4a3	nuclear receptor subfamily 4, group A, member 3	Chr5: 63.781801	9.718	0.870	80						
3	ENSRNOG00000018714	Arl5b	ADP-ribosylation factor like GTPase 5B	Chr17: 82.065893	8.971	0.815	80						
4	ENSRNOG00000011238	Tiparp	TCDD-inducible poly(ADP-ribose) polymerase 1	Chr2: 157.316293	7.733	0.809	80						
5	ENSRNOG00000046667	Fosb	FosB proto-oncogene, AP-1 transcription factor subunit B	Chr1: 80.214691	9.863	0.800	80						
6	ENSRNOG00000014456	Coq10b	coenzyme Q10B	Chr9: 61.655963	9.032	0.798	80						
7	ENSRNOG0000000640	Egr2	early growth response 2	Chr20: 22.452170	9.155	0.772	80						

Slide 39. Correlation of one transcript with all others. It is possible to set a minimum threshold of 5 log<sub>2</sub> units.

# Technical conclusions (mainly for P50 team)

Crf\_MI  
01Apr2020

1. Much of the mapping workflow is in place. We need smoother data and metadata entry methods.
2. The HS Rat GWAS data and genetic analysis is FAIR-compliant in GeneNetwork.
3. All code is accessible on GitHub  
<https://github.com/genenetwork/genenetwork2>.
4. You can replicate and tweak models when mapping any of the P50 traits. You can do this on your own without much hassle.
5. It is a short step from the P50 reports you get from Apurva to the P30 reports you can generate in GN. Our goal is to make sure these reports converge perfectly (replicable work flows).
6. As the rat genome assembly and sequence data for the eight parents of the HS improve, all traits will be remappable on the new assemblies and with "overlay" tracks of sequence variants and haplotypes in every locus.

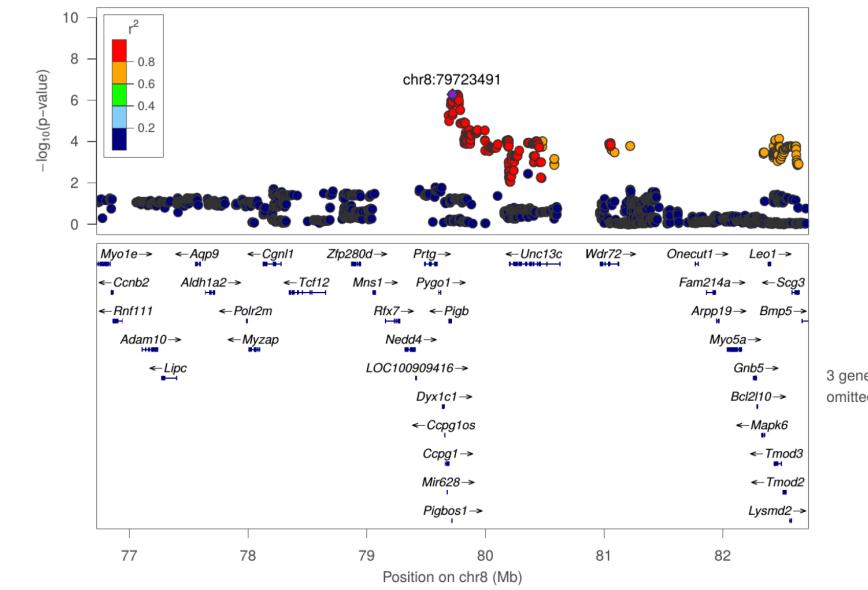


## P50 reports rock

13.7 crf MI inactive responses chr8:79723491

Size of interval : 7,984,956 bp

Number of genes in interval: 103



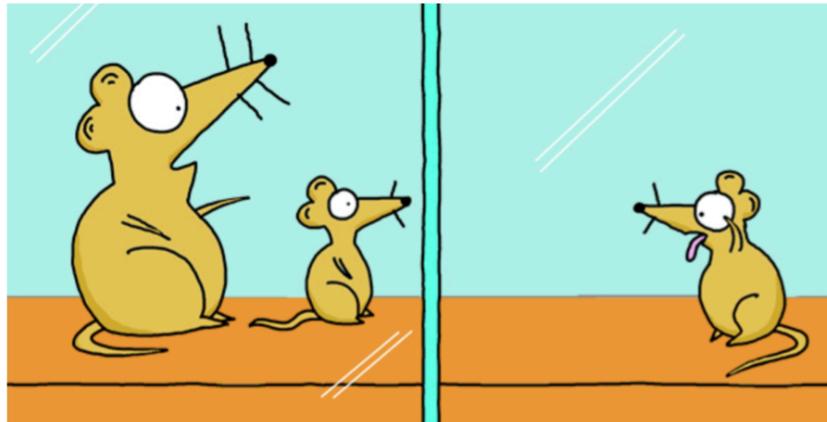
# Your suggestions

1. *Your ideas*
2. *Your ideas*
3. Faster mapping and permuting

Error messages in GN are cool ;-)

KOs

Us



"Don't play with him, he is Wild Type."

Slide 41. What to do next

genenetwork / genenetwork2

Code Issues 11 Pull requests 1 Actions Projects 5 Wiki Security Insights

GeneNetwork (2nd generation) <http://gn2.genenetwork.org/>

- 2,485 commits 5 branches 0 packages 1 release 11 contributors AGPL-3.0

Branch: testing New pull request Create new file Upload files Find file Clone or download

zsloan Fixed issue that caused submitted traits to not open properly when su... Latest commit cf8b4c2 2 hours ago

bin Fixed issue that caused logging in to stop working with new GUIX build 7 months ago

doc Updated Readme to reflect a couple changes 3 months ago

May 6, 2012 – Apr 16, 2020 Contributions: Commits

Contributions to testing, excluding merge commits

