

# PRIMER: POPULATION AND FUNCTION

GREG CARTER  
ASSISTANT PROFESSOR  
THE JACKSON LABORATORY



# Visualizing Population and Function

LEVERAGING WALL-SIZED HIGH-RESOLUTION DISPLAYS FOR COMPARATIVE GENOMICS ANALYSES OF COPY NUMBER VARIATION

Roy Ruddle, Waleed Fateen, Darren Treanor, Peter Sondergeld, Phil Quirke

HUMMOD BROWSER: AN EXPLORATORY VISUALIZATION TOOL FOR THE ANALYSIS OF WHOLE-BODY PHYSIOLOGY SIMULATION DATA

Keqin Wu, Jian Chen, William Pruett, Robert Hester

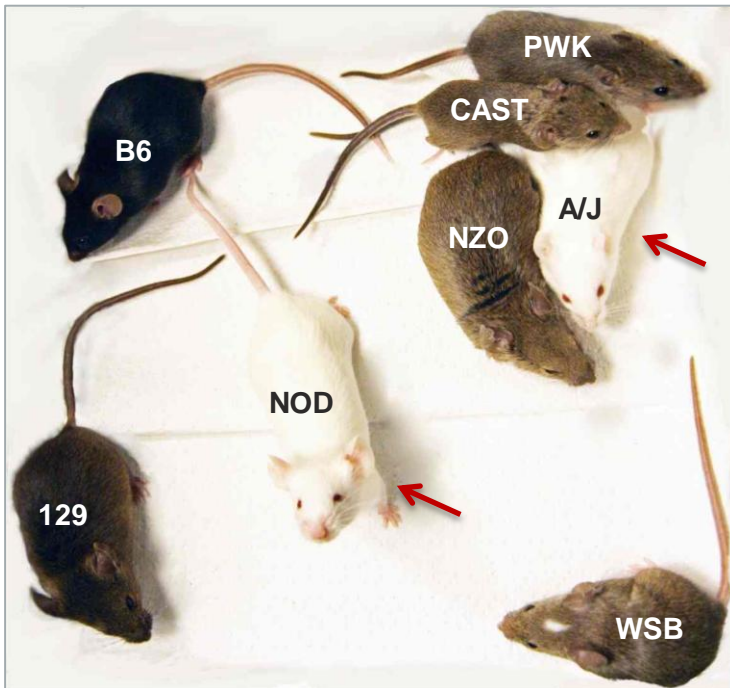
VISUAL CLEANING OF GENOTYPE DATA

Jessie Kennedy, Martin Graham, Trevor Paterson, Andy Law



# Statistical Genetics

## Associating genes with traits



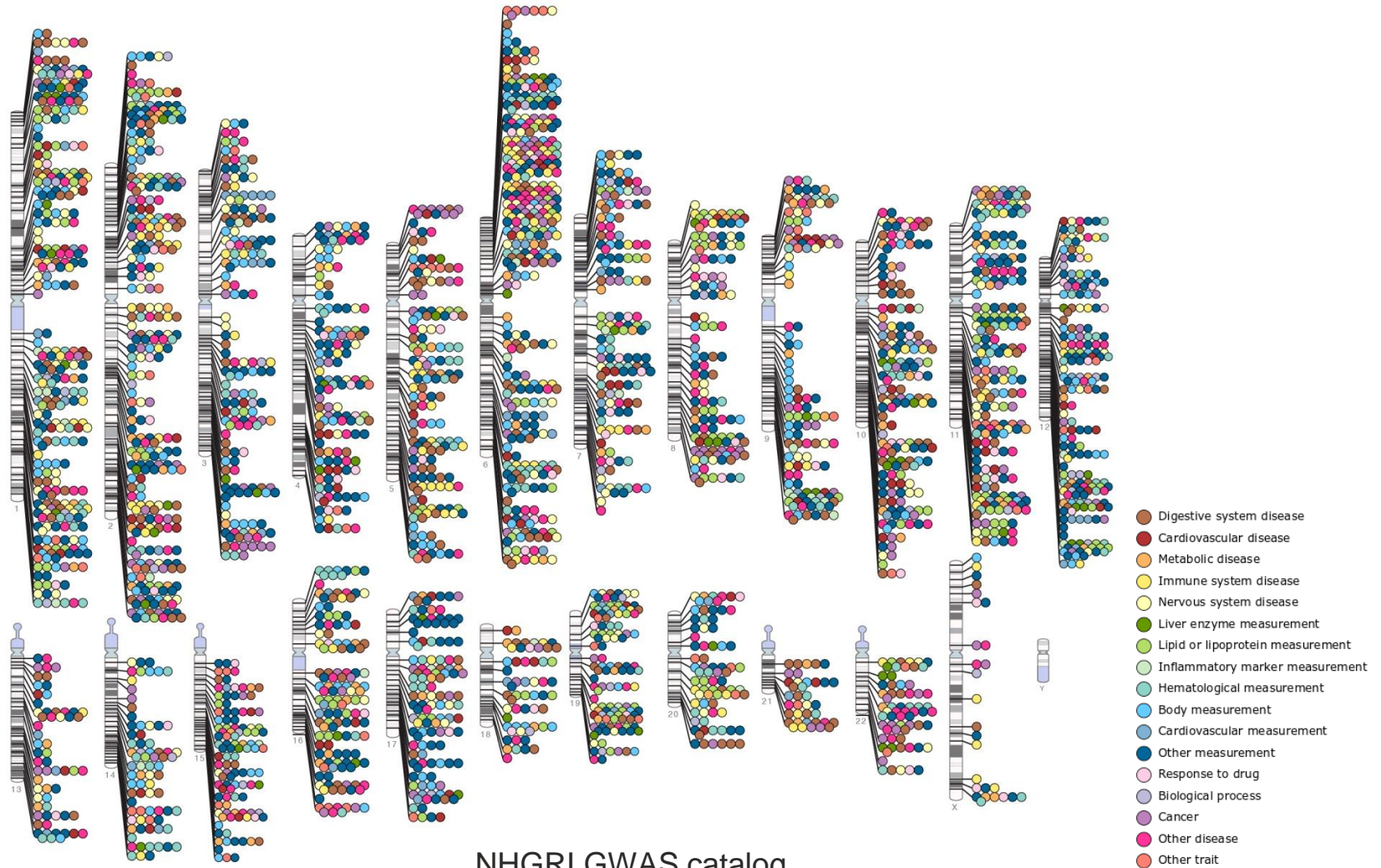
Gene	Chromosome	Position	dbSNP	Reference	129S1/svImJ	A/J	C57BL/6NJ	CAST/EiJ	NOD/ShiLJ	NZO/HILJ	PWK/PhJ	WSB/EiJ
Tyr	7	87,429,149	<a href="#">rs37137376</a>	A	-	-	-	-	-	-	T	-
Tyr	7	87,429,175	<a href="#">rs36559163</a>	G	-	-	C	-	-	-	C	-
Tyr	7	87,429,265	<a href="#">rs37390229</a>	A	-	-	G	-	-	-	G	-
Tyr	7	87,438,039	<a href="#">rs37996025</a>	T	-	-	-	-	-	-	G	-
Tyr	7	87,483,873	<a href="#">rs36285353</a>	T	-	-	C	-	-	-	-	-
Tyr	7	87,483,990	<a href="#">rs37591213</a>	G	-	-	A	-	-	-	-	-
Tyr	7	87,492,549	<a href="#">rs237830526</a>	A	-	-	-	-	-	-	G	-
Tyr	7	87,492,574	<a href="#">rs250107030</a>	A	-	-	-	-	-	-	C	-
Tyr	7	87,492,834	<a href="#">rs242698493</a>	A	-	-	-	-	-	-	G	-
Tyr	7	87,493,030	<a href="#">rs38343375</a>	A	-	-	-	-	-	-	G	-
Tyr	7	87,493,043	<a href="#">rs31191169</a>	C	-	G	-	-	G	-	-	-
Tyr	7	87,493,206	<a href="#">rs38358457</a>	G	-	-	-	A	-	-	-	-
Tyr	7	87,493,213	<a href="#">rs225453317</a>	G	-	-	-	A	-	-	-	-
Tyr	7	87,493,273	<a href="#">rs36558707</a>	G	-	-	-	A	-	-	-	-

Sanger mouse genomes database



# Human GWAS

Thousands of associations across the human genome

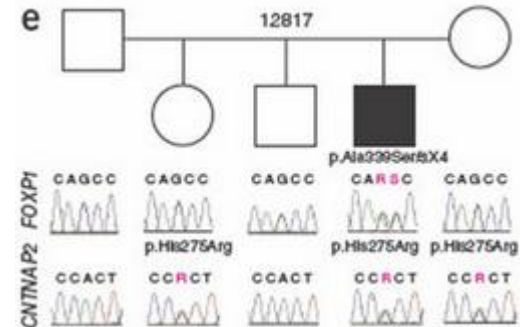


# Pedigree Studies

Familial disease

Rare variant associations

Accessible with exome sequencing



NATURE GENETICS | LETTER

日本語要約

Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations

Brian J O’Roak, Pelagia Deriziotis, Choli Lee, Laura Vives, Jerrod J Schwartz, Santhosh Girirajan, Emre Karakoc, Alexandra P MacKenzie, Sarah B Ng, Carl Baker, Mark J Rieder, Deborah A Nickerson, Raphael Bernier, Simon E Fisher, Jay Shendure & Evan E Eichler



# Cancer Genetics and Genomics

## The Cancer Genome Atlas

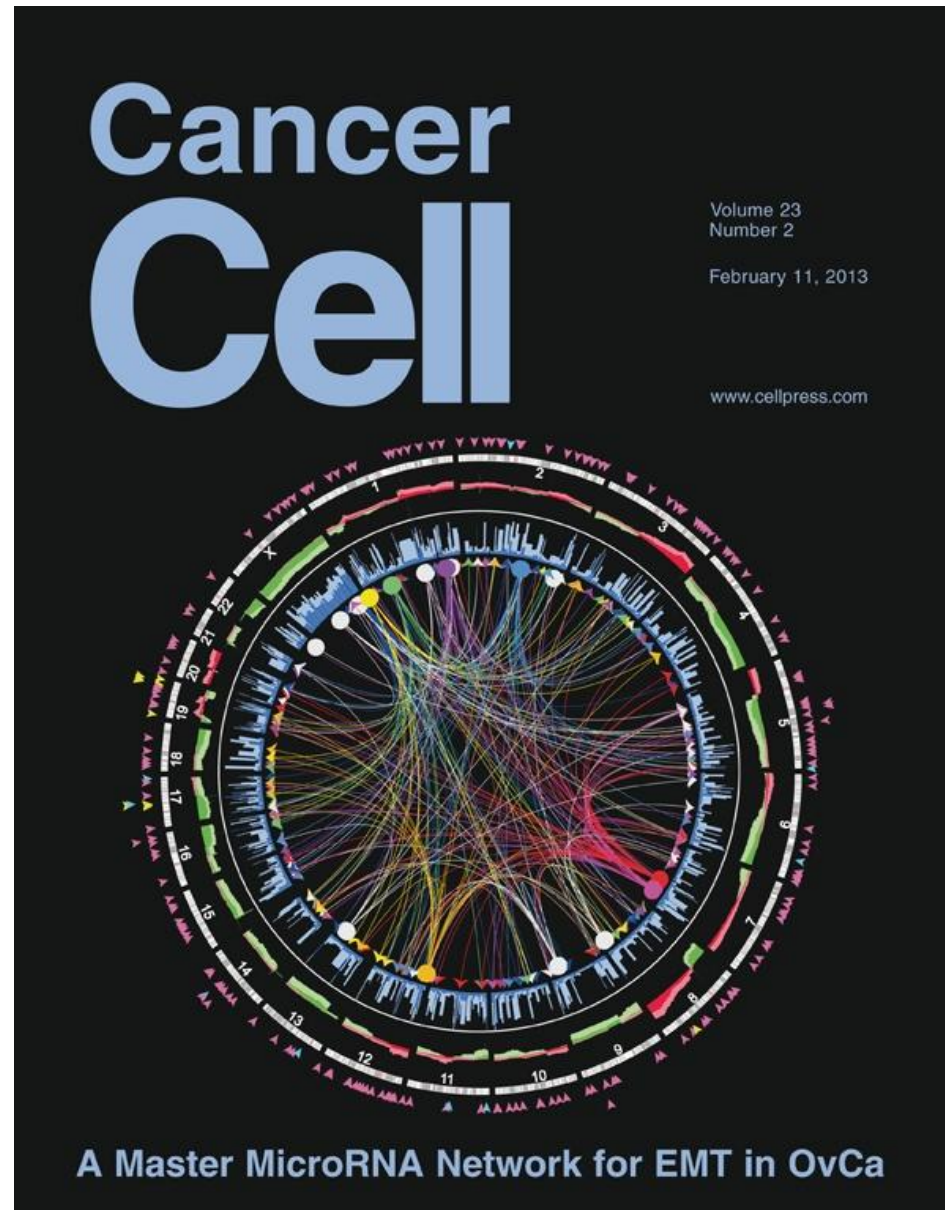
Primary tumor samples

dozens of cancers

thousands of samples

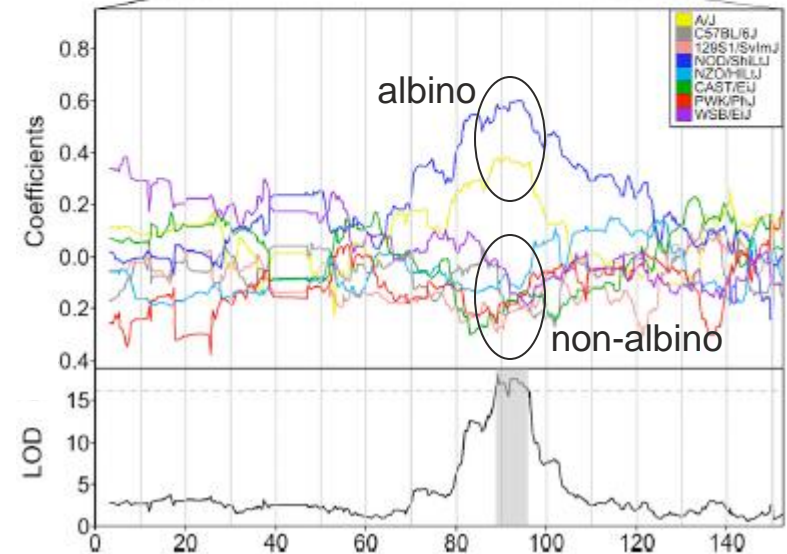
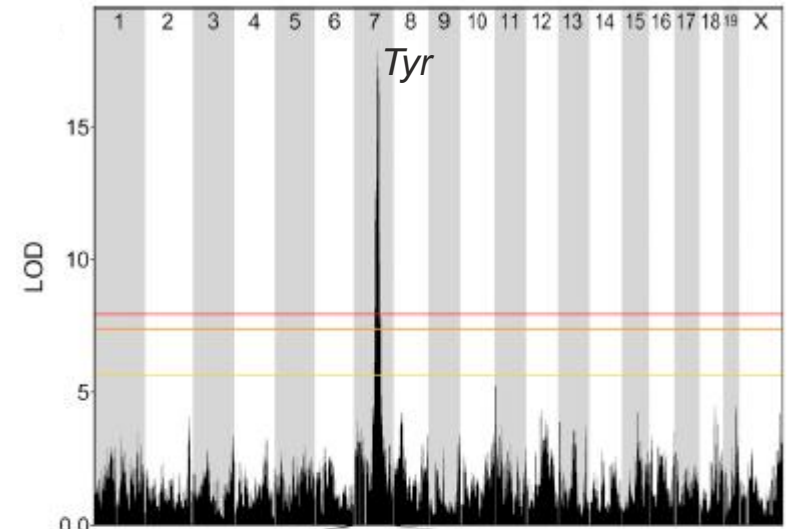
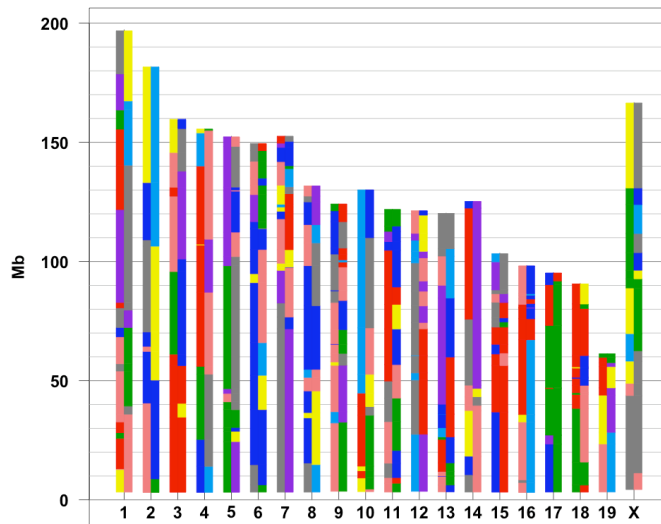
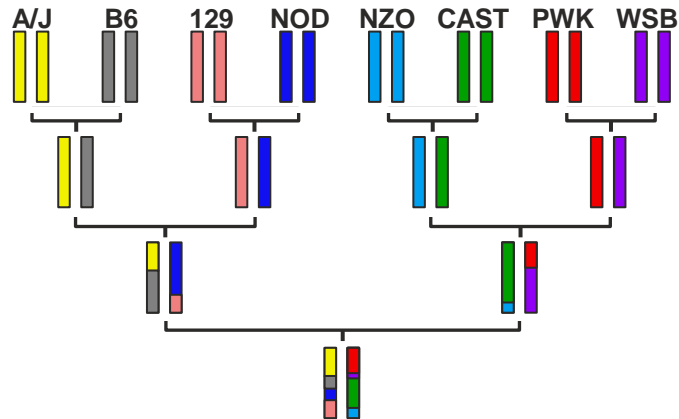
Multiple data types

CNV, SNP, RNAseq, histone  
modifications, etc.



# Complex Genetic Model Organisms

## Mouse Intercross Studies



# Current Visualization Tools

Two main classes

Online tools for data access

Data browsers, gene and ontology databases

Project-based, publication display for dissemination

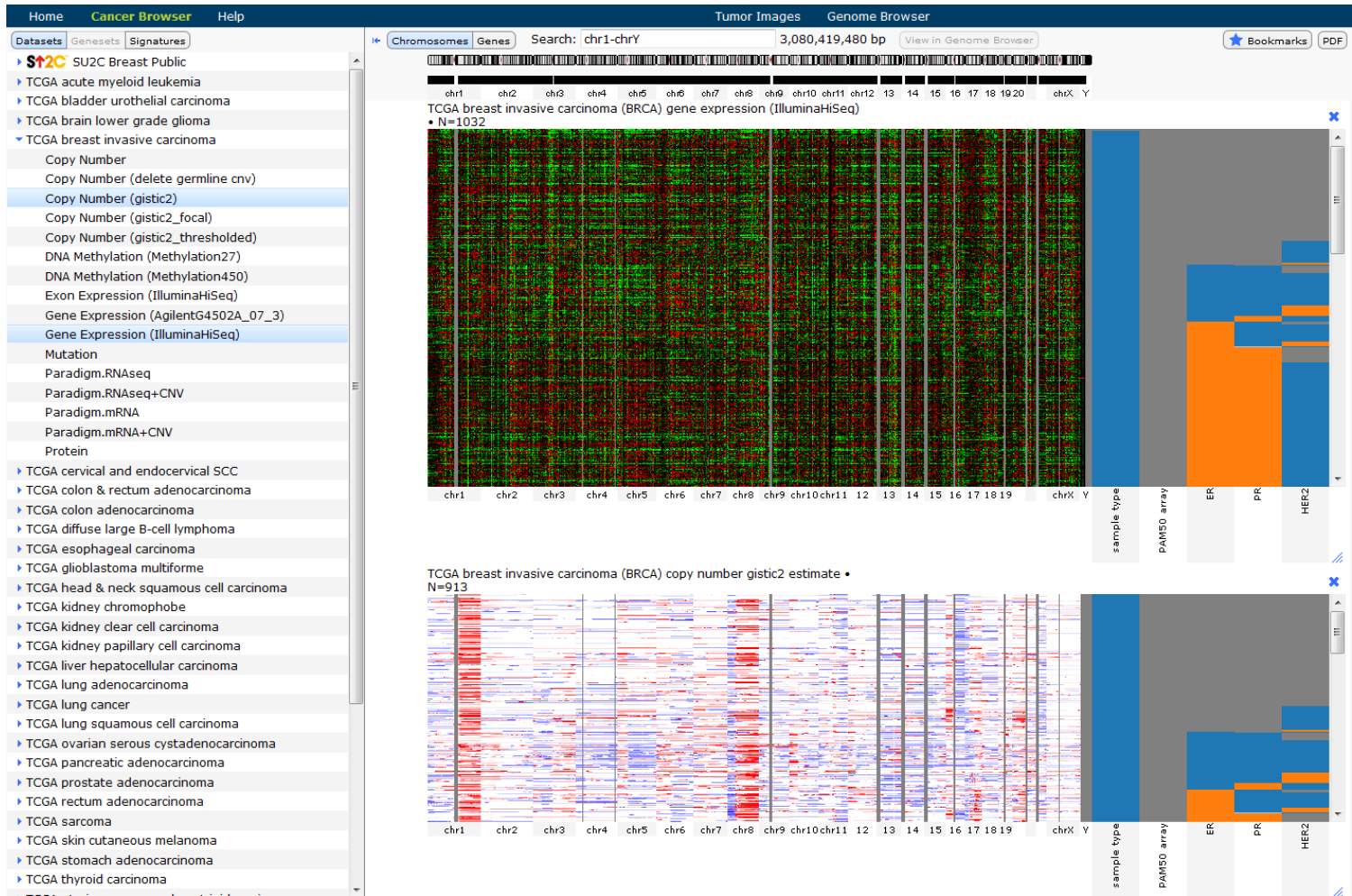
Professional science, amateur visualization





# Current Tools: Browsers

## TCGA data browser



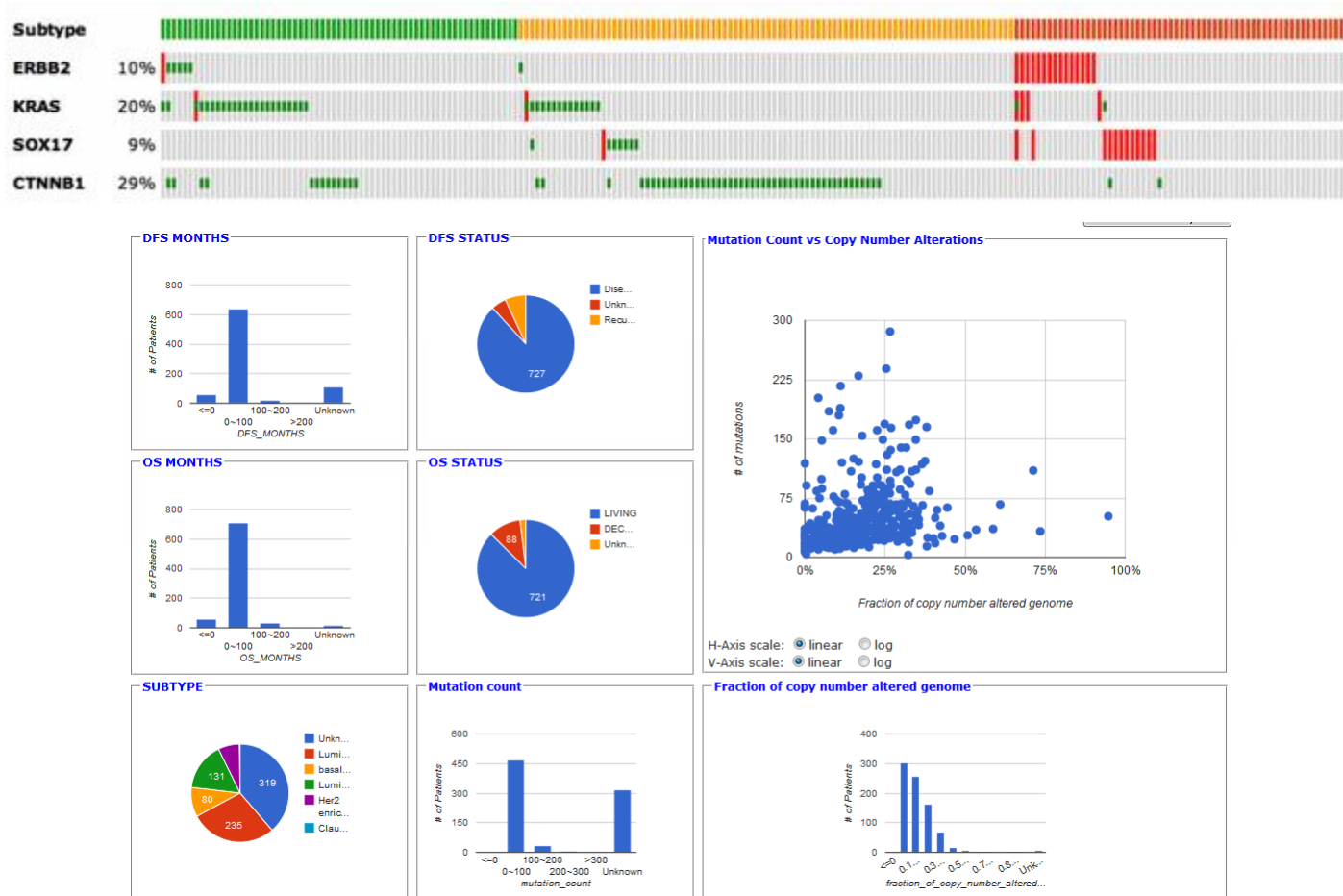
UCSC Cancer Genome Browser



THE JACKSON LABORATORY

# Current Tools: Browsers

## TCGA summary statistics

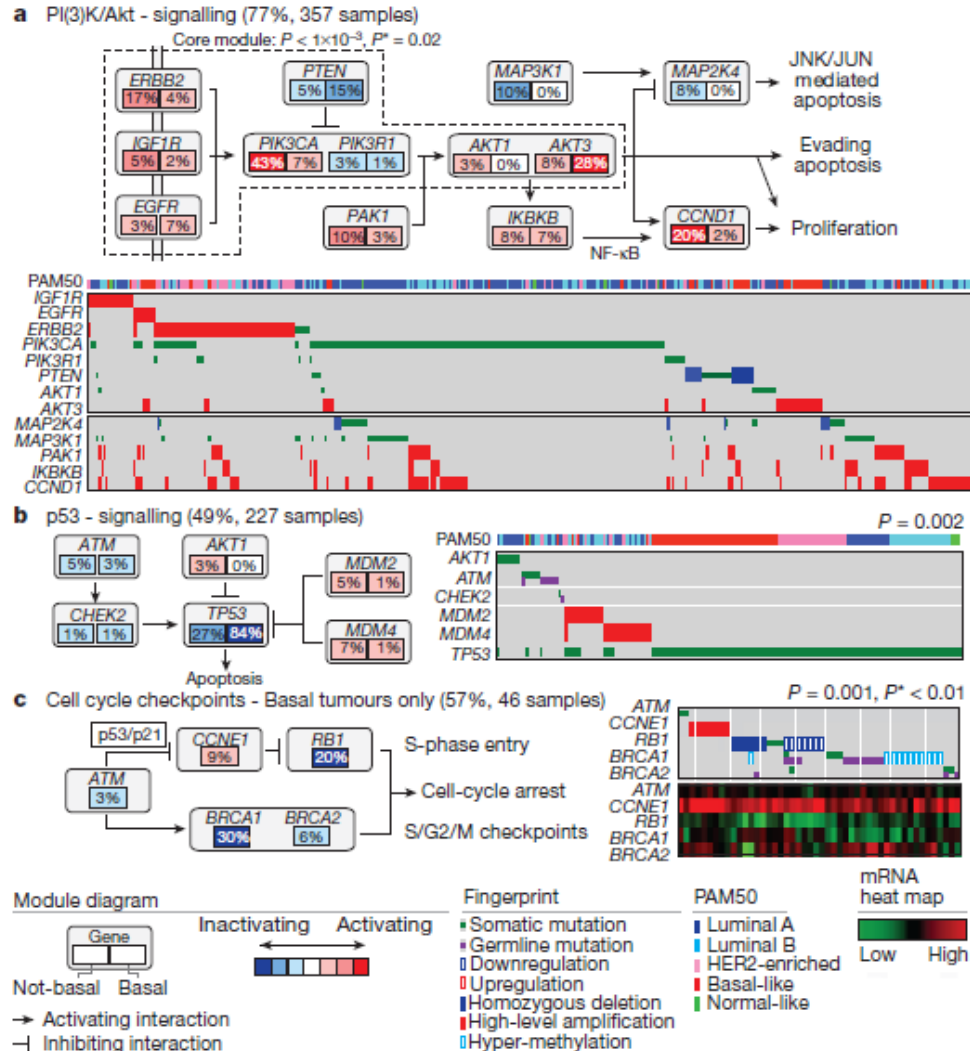


cBioPortal for Cancer Genomics, MSKCC



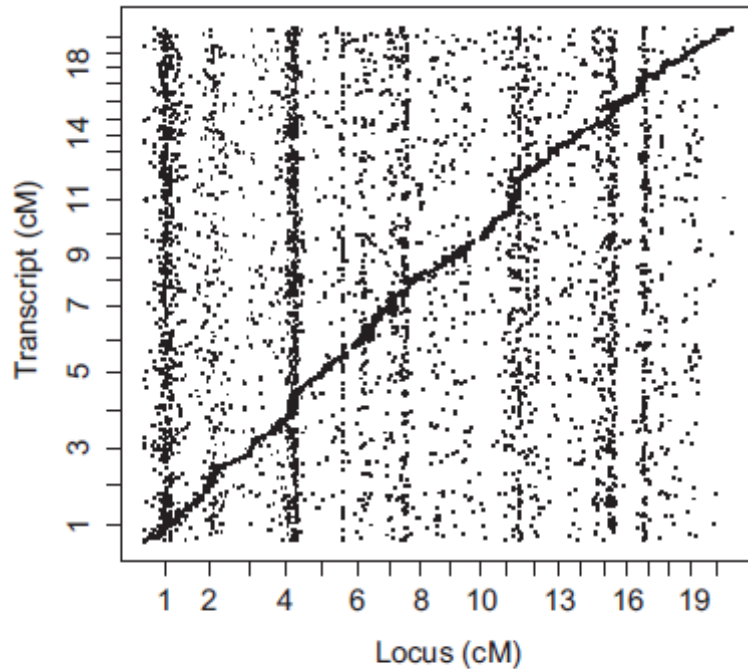
# Current Tools: Paper Figures

## TCGA publication



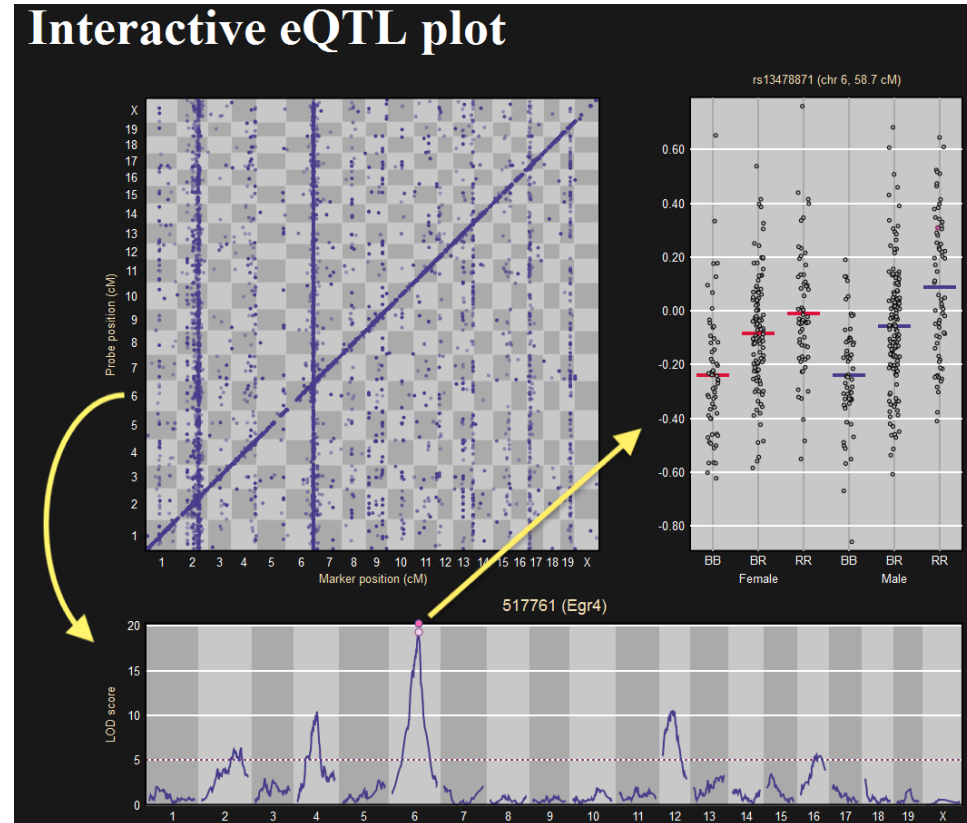
# Moving Towards Interactivity

eQTL: genetic studies with thousands of traits



**Figure 4.** Expression-QTL map for transcripts with significant ( $P < 0.05$ ,  $\text{LOD} > 3.85$ ) QTL. The x-axis indicates the position of the QTL locus, and the y-axis identifies the physical position of the transcript.

Hageman, et al, 2011



Karl Broman (@kwbroman), U Wisconsin



# Population and Function

## Ongoing Challenges

### Data Management

Genotype and phenotype data quality assessment	Paper 3
Allelic complexity in human and model populations	
Ancestry, haplotype structure, pedigree maps	Paper 3
Big data, diverse data	Paper 1

### Representing Complex Genetic Models

Gene-gene and gene-environment interactions	
Networks of phenotypes	Paper 2
Levels of biological complexity: sequence variants, transcripts, proteins, metabolites, tissues, physiology, ecology, ...	Paper 2



1. LEVERAGING WALL-SIZED HIGH-RESOLUTION DISPLAYS FOR COMPARATIVE GENOMICS ANALYSES OF COPY NUMBER VARIATION
2. HUMMOD BROWSER: AN EXPLORATORY VISUALIZATION TOOL FOR THE ANALYSIS OF WHOLE-BODY PHYSIOLOGY SIMULATION DATA
3. VISUAL CLEANING OF GENOTYPE DATA



# Acknowledgements

## Carter Lab

Anna Tyler  
Vivek Philip  
Justin Hendrick  
Robert Valenzuela  
Guru Ananda

## The Jackson Laboratory

Gary Churchill  
Ron Korstanje  
Rachael Hageman  
Wayne Frankel  
Karen Svenson

## Institute for Systems Biology

Michelle Hays  
Amir Sherman  
Tim Galitski

## Funding

NIGMS K25  
NIGMS Center for Genome Dynamics  
NCI Jackson Laboratory Cancer Center

