PRIMER: POPULATION AND FUNCTION

GREG CARTER ASSISTANT PROFESSOR THE JACKSON LABORATORY



Leading the search for tomorrow's cures

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	C/V	C57BL/6NJ	CAST/EIJ	NOD/ShiLtJ	NZO/HILtJ	CH4/XW4	WSB/Eij
Tyr	7	87,429,149	<u>rs37137376</u>	A	-	-	-	-	-	-	Т	-
Tyr	7	87,429,175	<u>rs36559163</u>	G	-	-	-	С	-	-	С	-
Tyr	7	87,429,265	<u>rs37390229</u>	A	-	-	-	G	-	-	G	-
Tyr	7	87,438,039	<u>rs37996025</u>	Т	-	-	-	-	-	-	G	-
Tyr	7	87,483,873	<u>rs36285353</u>	Т	-	-	-	С	-	-	-	-
Tyr	7	87,483,990	<u>rs37591213</u>	G	-	-	-	Α	-	-	-	-
Tyr	7	87,492,549	<u>rs237830526</u>	A	-	-	-	-	-	-	G	-
Tyr	7	87,492,574	<u>rs250107030</u>	Α	-	-	-	-	-	-	С	-
Tyr	7	87,492,834	<u>rs242698493</u>	A	-	-	-	-	-	-	G	-
Tyr	7	87,493,030	<u>rs38343375</u>	A	-	-	-	-	-	-	G	-
Tyr	7	87,493,043	<u>rs31191169</u>	С	-	G	-	-	G	-	-	-
Tyr	7	87,493,206	<u>rs38358457</u>	G	-	-	-	A	-	-	-	-
Tyr	7	87,493,213	<u>rs225453317</u>	G	-	-	-	A	-	-	-	-
Tyr	7	87,493,273	<u>rs36558707</u>	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



Svenson, et al., *Genetics* 2012



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 visualiazation

Current Tools: Browsers

TCGA data browser

Home Cancer Browser Help	Tumor Images Genome Browser	
Datasets Genesets Signatures	H Chromosomes Genes Search: chr1-chrY 3,080,419,480 bp (View in Genome Browser)	T Bookmarks PDF
Star SU2C Breast Public		
TCGA acute myeloid leukemia		
TCGA bladder urothelial carcinoma	chr1 chr2 chr3 chr4 chr5 chr6 chr7 chr8 chr9 chr10 chr11 chr12 13 14 15 16 17 18 19 20 chrX Y	
TCGA brain lower grade glioma	N=1032 Notes that the second (BRCA) gene expression (IlluminaHISeq)	×
 TCGA breast invasive carcinoma 		·
Copy Number		
Copy Number (delete germline cnv)		
Copy Number (gistic2)		
Copy Number (gistic2_focal)		
Copy Number (gistic2_thresholded)		
DNA Methylation (Methylation27)		
DNA Methylation (Methylation450)		
Exon Expression (IlluminaHiSeq)		
Gene Expression (AgilentG4502A_07_3)		
Gene Expression (IlluminaHiSeq)		
Mutation		
Paradigm.RNAseq		
Paradigm.RNAseq+CNV		
Paradigm.mRNA		
Paradigm.mRNA+CNV		
Protein		
TCGA cervical and endocervical SCC		
TCGA colon & rectum adenocarcinoma	chr1 chr2 chr2 chr5 chr4 chr5 chr6 chr7 chr6 chr7 chr8 chr9 chr10 thr11 12 13 14 15 16 17 18 19 chrX Y	ed e Harris Kara
TCGA colon adenocarcinoma		
TCGA diffuse large B-cell lymphoma		W W20
TCGA esophageal carcinoma		by by
TCGA glioblastoma multiforme	TCGA breast invasive carcinoma (BRCA) copy number gistic2 estimate •	
TCGA head & neck squamous cell carcinoma	N=913	×
TCGA kidney chromophobe		
TCGA kidney clear cell carcinoma		
 TCGA kidney papillary cell carcinoma 		
 TCGA liver hepatocellular carcinoma 		
TCGA lung adenocarcinoma		
 TCGA lung cancer 		
TCGA lung squamous cell carcinoma		
TCGA ovarian serous cystadenocarcinoma		
 TCGA pancreatic adenocarcinoma 		
 TCGA prostate adenocarcinoma 		
TCGA rectum adenocarcinoma		
TCGA sarcoma	chr1 chr2 chr3 chr4 chr5 chr6 chr7 chr8 chr9 chr10chr11 12 13 14 15 16 17 1819 chrX Y	ed se m n n n
 TCGA skin cutaneous melanoma 		
TCGA stomach adenocarcinoma		Wa Wa
TCGA thyroid carcinoma	*	s d
TOOL 1 1 1 1 1 1		

UCSC Cancer Genome Browser



Current Tools: Browsers

TCGA summary statistics



cBioPortal for Cancer Genomics, MSKCC



Current Tools: Paper Figures

TCGA publication





TCGA Network, Nature 2012

Moving Towards Interactivity

eQTL: genetic studies with thousands of traits



Figure 4. Expression-QTL map for transcripts with significant (P < 0.05, 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