G T C A

Seeing the Bigger Picture T G G A

Through Billions of Bases





Nancy J. Cox, Ph.D.

The University of Chicago

http://genemed.uchospitals.edu

G

Seeing the Bigger T G G



Through Billions of Bases





Nancy J. Cox, Ph.D.

The University of Chicago

http://genemed.uchospitals.edu

G

Seeing the Bigger T G G



Through Billions of Bases





Nancy J. Cox, Ph.D.

The University of Chicago

http://genemed.uchospitals.edu

Pictures and Art

- Association studies via highthroughput genotyping – Cave painting?
- Insights from exome and wholegenome sequencing – Minimalism?
- Omics beyond DNA sequence Abstraction?
- Prediction, causality Impressionism?
- A more complete synthesis postmodernism?

Relating Genotype to Phenotype





Genotyping

Sequencing

Relating Genotype to Phenotype



Genotyping





Sequencing

Genome Interrogation



GWAS Has ...

Given us 1000-fold more discovery of loci associated with common human diseases and related quantitative traits

GWAS Has Not ...

Given us much novel biological understanding; highly significant and reproducible signals tend to be non-coding and we are often unsure even what gene is implicated

Picking the Cherries ...



Orchards of Cherry Trees to Pick





Ellsworth Kelly

Whole Genome and Exome Sequencing Has...

Given us an unprecedented understanding of human variation and drawn attention to how recent demography has shaped our genome

Enabled a successful assault on the remainder of Mendelian disorders not yet ascribed to genes

Taught us lessons we have not been keen to learn about effect sizes for rare variants in common diseases and quantitative traits



Exome Sequencing Has Not ...

Frank Stella

- Enabled gene discovery for all "Mendelian" phenotypes or pedigrees attempted
- Been adequately powered to enable discovery of novel contributing genes in applications to complex disorders at 10K + cases and controls





Methylated DNA Locus



Unmethylated DNA Locus



Methylome

Protein-omics

GTEx: Genotype Tissue Exchange



Program Snapshot

The Common Fund's Genotype-Tissue Expression (GTEx) program aims to study human gene expression and regulation in multiple tissues, providing valuable insights into the mechanisms of gene regulation and, in the future, its disease-related perturbations. Genetic variation between individuals will be examined for correlation with differences in gene expression level to identify regions of the genome that influence whether and how much a gene is expressed. The GTEx project includes the following initiatives:

- Novel Statistical Methods for Human Gene Expression Quantitative Trait Loci (eQTL) Analysis
- Laboratory, Data Analysis, and Coordinating Center (LDACC)
- caHUB Acquisition of Normal Tissues in Support of the GTEx Project



Read more ...





Autoimmune Disorders

Hypertension and Adipose eQTLs



HT (eQTLs)





Polygenic Modeling and Prediction

REPORT

GCTA: A Tool for Genome-wide Complex Trait Analysis

Jian Yang,^{1,*} S. Hong Lee,¹ Michael E. Goddard,^{2,3} and Peter M. Visscher¹

ARTICLE

Estimating Missing Heritability for Disease from Genome-wide Association Studies

Sang Hong Lee,1 Naomi R. Wray,1 Michael E. Goddard,2,3 and Peter M. Visscher1,*



By Minor Allele Frequency

	Tourette Syndrome		Obsessive-Compulsive Disorder	
MAF	Number of SNPs	h2 (s.e.)	Number of SNPs	h2 (s.e.)
> 0.001 - 0.05	20,316	0.13 (0.04)	19,605	0 (0.03)
> 0.05 - 0.1	43,440	(0.02	47,970	0.04 (0.05)
> 0.1 -0.2	96,398	0.11 (0.07)	91,661	0.08 (0.08)
> 0.2-0.3	81,924	0.12 (0.07)	77,641	0.01 (0.01)
> 0.3-0.4	74,393	0.16 (0.07)	70,193	0.11 (0.05)
> 0.4 -0.5	70,911	0.07 (0.06)	66,770	0.11 (0.05)

Heritability by MAF (Imputed)

	Tourette Syndrome		Obsessive-Compulsive Disorder	
MAF	Number of SNPs	h2 (s.e.)	Number of SNPs	h2 (s.e.)
> 0.001 - 0.05	2,243,744	0.15 (0.09)	2,357,568	0.000001 (0.06)
> 0.05 – 0.5	5,538,943	0.34 (0.10)	5,492,973	0.32 (0.12)

Genetic Correlation



Bipolar Disorder

- Heritability estimated from previous twin and family studies ~ 0.7
- Heritability estimated from (GAIN + Bipolar Genome Study) ~ 0.35 (600K SNPs)
- Heritability estimated for same samples using 27K cis-eQTLs from brain ~ 0.2
- Heritability estimated for same samples using 27K cis-eQTLs from LCLs < 0.01



Causality ...

Remains hugely challenging for rare variant discoveries in Mendelian diseases and for common and rare variant discoveries in complex disorders

Cannot be assessed through knock-outs and knock-ins in every situation

Is particularly hard to prove when you have a single family (or individual) with a rare phenotype





Post-Modernism ?

Whole genome interrogation – genes as well as other functional units

Using more – ideally all – available information simultaneously

Apply Console Status

BIONIMBUS PROTECTED DATA CLOUD

Secure cloud services for the scientific community

What is the Bionimbus PDC?

P D (

The Bionimbus Protected Data Cloud (PDC) is a collaboration between the Open Science Data Cloud (OSDC) and the IGSB (IGSB,) the Center for Research Informatics (CRI), the Institute for Translational Medicine (ITM), and the University of Chicago Comprehensive Cancer Center (UCCCC). The PDC allows users authorized by NIH to compute over human genomic data from dbGaP in a secure compliant fashion. Currently, selected datasets from the The Cancer Genome Atlas (TCGA) are available in the PDC.

How can I get involved?

- · Apply for an Bionimbus PDC account and use the Bionimbus PDC to manage, analyze and share your data.
- · Partner with us and add your own racks to the Bionimbus PDC (we will manage them for you).
- Help us develop the open source Bionimbus PDC software stack.

You can contact us at info@opencloudconsortium.org.

How do I get started?

First, apply for an account. Once your account is approved, you can login to the console and get started. Support questions can be directed to support@opencloudconsortium.org.

Apply for the PDC Now

Login to the PDC Console









CHICAGO and our sponsors.

bionimbus-pdc.opensciencedatacloud.org

Apply For An Account

PDC Console Apply

oly Status

Apply for a PDC Account

Please fill out this form to apply for access to the Bionimbus Protected Data Cloud. To access the PDC you will need an eRA Commons account and dbGaP access to TCGA. A member from the PDC team will be in touch with you within 48 business hours. If you have any questions about the application process, please contact us at accounts@opencloudconsortium.org. Thank you.

eRA Commons Username*	
Name*	The Bionimbus PDC is
E-mail Address*	Currently in Deta and
Organization/University*	IOOKING FOR PILOT
Web Page	access to TCGA data.
Phone Number	
Address	

bionimbus-pdc.opensciencedatacloud.org



Anna Pluzhnikov

Cox Lab





Steven Zhang



Anuar Konkashbaev



Pat Evans

Jason Torres



Lea Davis (Bridget)







Anna Tikhomirov



Eric Gamazon

The "nitty gritty" analysis group!











ACGI MERT MERT MERT ACGT ACGT ACGT ACGT A	CGT ACGT ACGT ACGT ACGT ACGT ACGT
ACGT ACGT NOFT NOT NOT NOT NOT	T NOT ACCT ACCT CUT/CUT/CUT/CUT/CUT/
ACGIACGIACGIACGI ACCT ACCT ACGIACG	TACGIACGIACGIACGIACGIACGIACGI
ACGUACGU ADT ADT ADTACGUACG	TACIERO ALEO ALEO ALEO ALEO ALEO ALEO ALEO ALE
ACCT ACCT ACCT ACCT & COT ATTACT AC	TATETATETACCE ACCE ACCE MOST MOST ACCE.
NCTION LOOP LOOP LOOP LOOP LOOP L	CONTRACT ACCT FOR ADD ADD
ACONTLATATOT ACGT ACGT ACGT A	CCTACT TALL ANT ANT ANT ANTICAL
ACGININ ADTACT ACGINEGIA	Lincon acon acon acon acon acon
ACOT ACOT ACOT ACOT ACOT ACOT ACOT ACOT	AGE AGET AGET AUGIAOST AOST AUGIAUGI
ACGIACGIACGTACGTACGTACGTACGTACCTA	CGT ACCT ACCT ACCT ACCT ACCT
ACCE ACCE ACCE ACCE ACCE ACCE ACCE ACCE	EDD THAT THAT THAT THAT THAT THAT THAT
ACCRETATING AND AND AND	NOT NOT NOT NOT NOT NOT AN LONG T
Multicon Loop Loop Loop Mile	ALC: ALC: ALC: MODELOW LOCE LOCE
AUGUNEUST NEUST NEUST NEUST NEUST NEUST NEUST	A DEFAURT ACGT ACGT ACGT ACGT ACGT
ACGT ACGT ACGT ACGT ACGT ACGT ACGT	WARE WARE WARE BOAT ACCT. HUMOWAR
AGET AGET ACGT ACGT ACGT ACGT ACGT ACGT ACGT ACG	ATT MANE ALLE ACOT ACOT ACOT ACOT
ACGIACGIACGIACGIACGIACGI ACGI ACGI	GIACGIACGTACGT
ACET ACET ACET ACETACITY ACCT ACET	TALL AND A LOCAL ACCULATED
NOT NOT THE NEEDED AND AND ADDRESS	TACKED AND ANT ADDITION TO ADDITION
ACCRECATION AND AND ADDRACTIONS	TATUTA A CASE AND ADDRESS AND A CASE AND A CASE
ACCT ACCT IN THE THE THE THE THE	
ALTI ALTI ALTIALTIALTIALTIALTIALTIALTIALTIALTIALTI	COLDER MODEL MODEL MODEL MODEL AND ADDRESS
ACGT ACGT ACGT ACGT ACGT ACGT ACGT ACGT	TROAT BOAT BOAT BOAT BOAT BOAT BOAT
ACGTACGTACGTACGTACGTACGT #05/2 #0	BY AGET MENT ACCURACET ACCT
ACCTACGAOGLICUTACITACITACI	ACIENCE AND A REAL PROPERTY AND A REAL PROPERT
NOT NOT NOT ACCURACED ACED ACED A	CCTACCT ACCT ACCT ACCT
ACCEACCE ACCE ACCEACUE	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1
ACCE ACCE ACCE ACCE ACCE ACCE	ALLE ALLE ALLE ALLE ALLE ALLE ALLE
ACGT ACGT AGAT ACGT ACGT ACGT ACGT ACGT	DI SUI AUGURI AUGURIS I AUGURIS I
AUGTAUGTAUGTAUGTAUGTAUGT	ACCHARGE ACGT ACGT ACGT
ACGTACGTACGTACGTACGTACGTACGTACGTACGTACGT	ACGUARTNEETACGTACGTACGT ACGT ACGT
ACGU ANT ACCU ACCU ACCU ACCU ACCU	I MEET ADDIT ACCESSION ACCESSION
NOT NOT NOT NET NET ALL AND ADD TO DE	THE ALL PROPERTY AND A DATA DATA
A LET A LET A LET A LET A LET A	A CELACE ACCURACE
NET NET NET NET NET NET NET NET	
ACDIACDIACDIACDIACDIACDIACDIA	3.0 SUBJELET RUTAUS AUSTAUS AUST AUST
AGE AGE AGE AGE AGE ACCT /6.5	ACTERNAL PROVENUES AND ALL ALL ALL ALL ALL ALL ALL ALL ALL AL
ACGTACG ACGTACGTACGT ACCT ACCT	T ACOT ADOT ACGIACGIACGIACGT ACCT
ACGTACH INCITACISTACIST ACCT NOW	AUSTACGIACGY ACCY ACGIACGIACGIACGE
ACGA CELACELACELACELACELACELACELACELACELACELA	TY ACT ADD ADD ADD ADD ADD ADD ADD ADD ADD AD
NOT THE REAL PROPERTY OF THE R	LOCAL MORT LOCAL MORT MORT MORT
Martin Martin Martin and Martin	Charles March March March States March 19
HEAT AND	A STATE AND A STATE AND A STATE OF A STATE O
ALL CONTROL OF COTACLEACED ACSE	ALST ALST ACGE ACGE ACGE ACGE ACGE ACGE
197 ACGT ACGT ACGT ACGT ACGT	ADM/AUGIA/USIAUGI ACGT ACGT ACGT
A TET ATET ATET ATET	ADSTACHT ACCELAUGT ACCIT ACCIT ACCIT
ATT ATT ALL.	NOS ACCENCE ACCE ACCE ACCE
100 ACCT	ADDY ADDITADOL SCIEPADOL AND ADD
A STATE OF A	ALL ALL PROPERTY AND A REAL PROPERTY AND A REAL PROPERTY.
ALC: ALC: ALC: ALC: ALC: ALC: ALC: ALC:	MENTALLINES WAS WAS WAS
CLI AC.	LACIELACIELACIELACIELACIELACIELACIEL
UNC ENAUGRACI	GIADGAREGMODI/ACGT ACGT
ACCENCE ACCE ACCE ACCE ACCE ACCE	ACCEP & COST MOXES NOT A COST ACCET
TATACG	ACCIDENT ACCESSION ACCESSION ACCESSION
ACCT A	CCLACGI ACCU ACCU ACCUACE
ALC: NO ALC: NO ALC: NO ALC: NO ALC: NO	
A COLOR AC	CI HAV MAN MONTH AND ALL AND
A REAL PROPERTY AND A REAL PROPERTY AND A REAL PROPERTY AND A	MATTER TENT
ACCEL ACCEL ACCEL ACCEL ACCEL ACCEL	ALGMACHAUGIACGT ALUT ALGT ACGT
ALL AND ALL AND ALL AND GALCAN	OF AGY AGY ACGE ACGE ACGE ACGE
THE REPORT OF THE PARTY OF THE STATE OF	ACGTACGISCERACEIACCIACCIACGIACEI
As 100 YOM THE COMPANY OF THE STATE	ACCUT ACCUT ACCUT ACCUT ACCUT ACCUT ACCUT
	COT SCOT SCOT SCOT SCOT
	the set of the point of the set of the set of the set
OT ALL MALE ADDRESS OF A DAMAGE AND A DAMAGE	CHARGE ACETACET ACCT
ALCONTRACT AND ADDRESS AND ADDRESS	ALL ALL ALL ALL ALL ALL ALL ALL
	T ATT NUSACGAOGIACTIACUACG
	GIACGIACGI A CALL
	AND A STATE AND A STATE AND A STATE
	A DE ANTRACES AND A LET
	TOT ACCT AGET
	AND MARK AND ADDRESS OF THE AND





































