University of Colorado Boulder

IBS INSTITUTE OF BEHAVIORAL SCIENCE ■ UNIVERSITY OF COLORADO AT BOULDER ■ BOULDER, COLORADO ■ CUPC

Genome-wide association (GWAS) methods for demographers

Jason D. Boardman

University of Colorado at Boulder

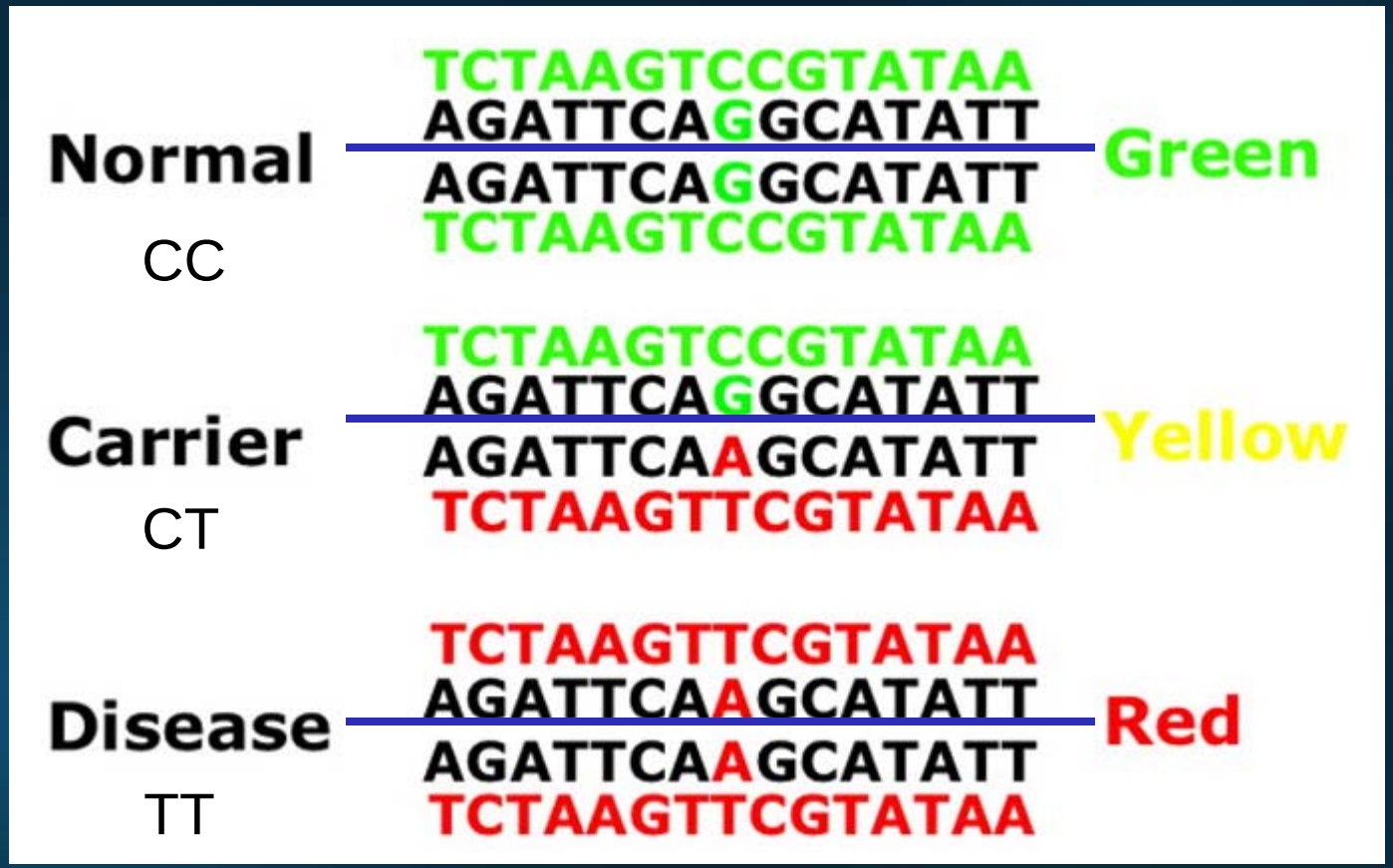
Department of Sociology and Institute of Behavioral Science

•*Presented at the* "Network on Biological Risk" at the annual meetings of the Population Association of America. San Francisco, CA. May 1st, 2012. **With a HUGE thank you to Matt McQueen for his generosity with some of the slides presented here**.

Overview of talk

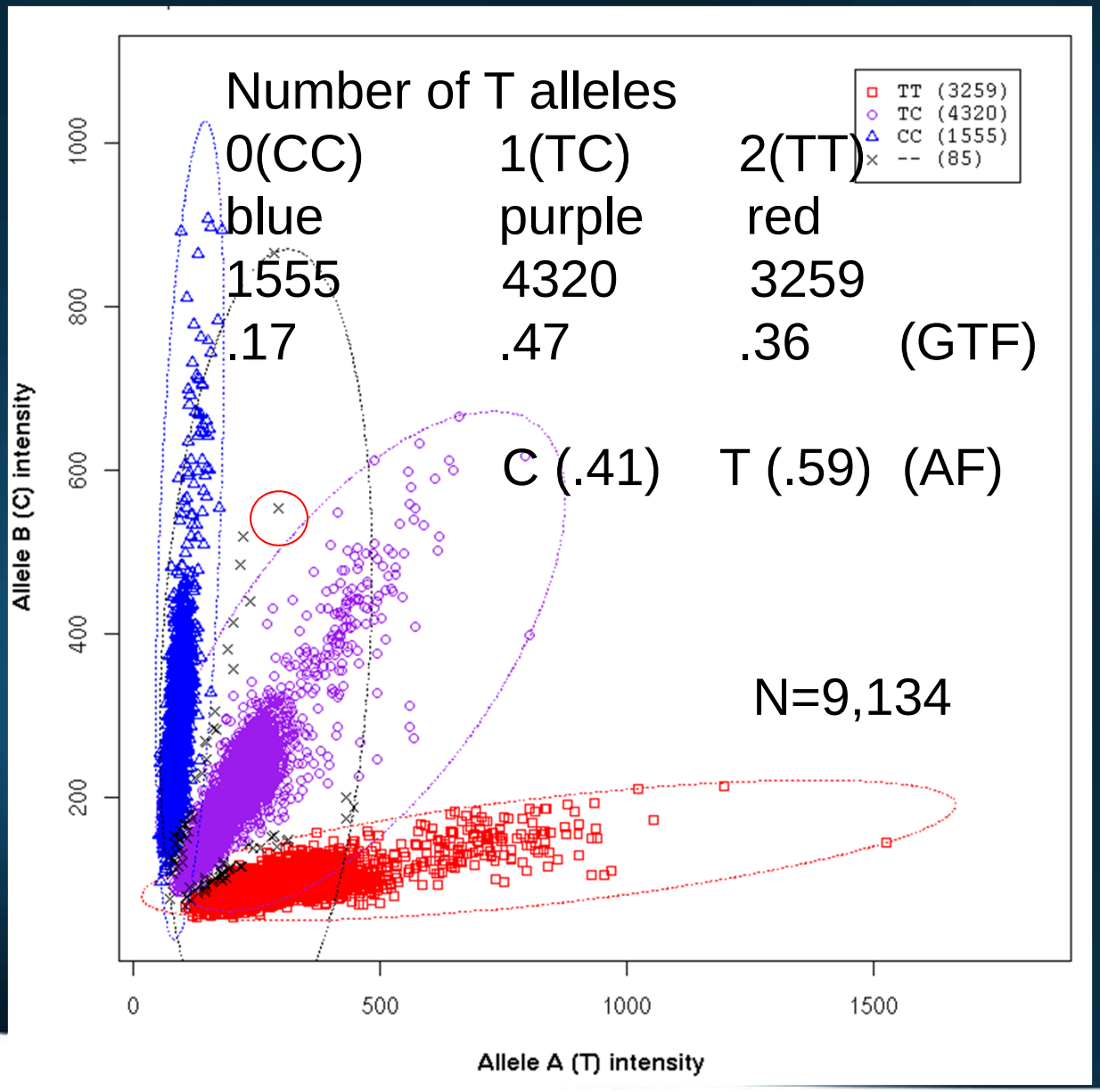
- What does genome wide data look like?
- What can demographers do with this data?
  - GWAS
  - Population stratification and importance of family data
  - GWGEI
  - Genome wide heritability models (GCTA)
  - Pathways analysis (Webgestalt2)

What is a SNP (and base pair)



- Image from: http://www.arrayit.com/Services/SNP_Genotyping/a_disease_carriers.jpg

How to “call” a genotype: Intensity files



What the data look like

	DRD2_1	DRD2_2	DRD2_g~o
1.	1	2	1
2.	1	2	1
3.	1	2	1
4.	2	2	2
5.	2	2	2
6.	1	2	1
7.	2	2	2
8.	1	2	1
9.	2	2	2
10.	2	2	2
11.	1	1	0

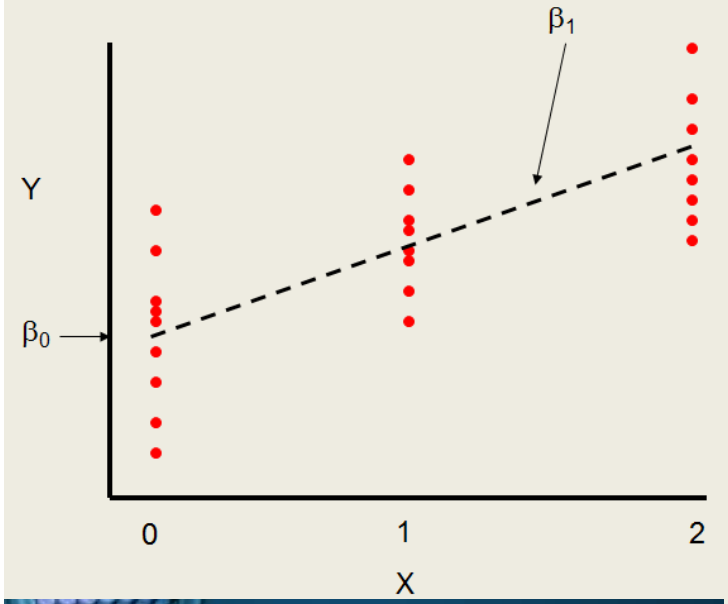
- TaqIa polymorphism in the DRD2 gene.
- DRD2_1 is strand 1
- DRD2_2 is strand 2
- DRD2_geno is the genotype.

Genotype and phenotype: now what

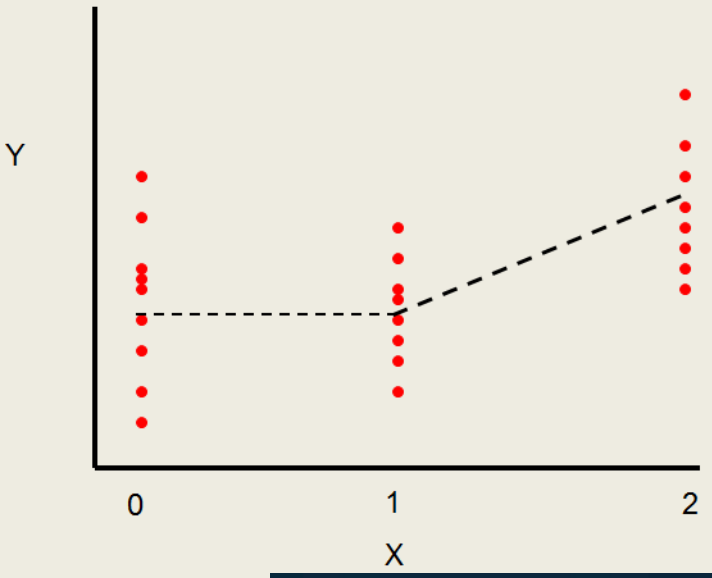
- It depends
  - Do you have genetic information of the parents
  - Do you have any siblings?
  - Is your dependent variable binary, count, time to onset, continuous.
  - Do you have multilevel data sources
  - Do you have complex sampling designs
- The take home is…..
  - That generalized linear models can be used to assess most GWAS approaches but it is important to understand the limitations and potential problems.

Just regress phenotype (y) on genotype (x)

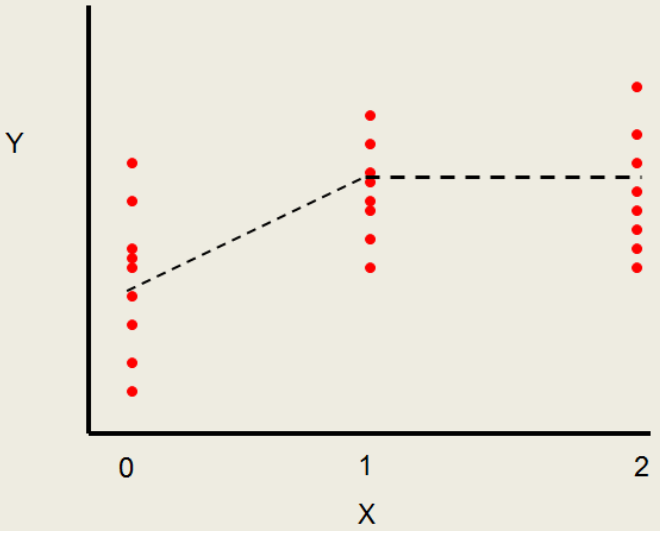
Additive Mode



Recessive Mode



Dominant Mode



NOW…it's just a regression 1m times.

- Additive (g = 0, 1, 2)

$$y = a + b_1 g + e_i$$

- Dominant (0=0, 1 = 1, 2=1)

$$y = a + b_1 g_{1,2} + e_i$$

- Recessive (0=0, 1 = 0, 2=1)

$$y = a + b_1 g_2 + e_i$$

- Heterozygote advantage (0=0, 1 = 1, 2=0)

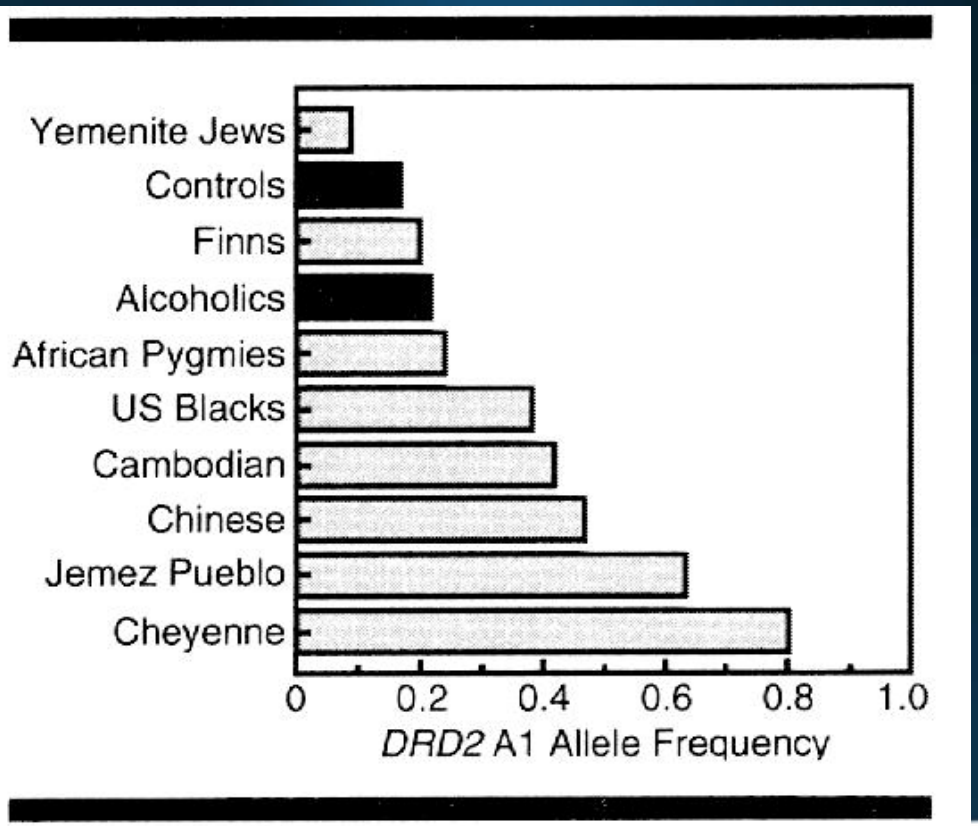
$$y = a + b_1 g_1 + e_i$$

- Any statistical program can run these models but there may be serious PROBLEMS with this approach.

The big problem: Population stratification

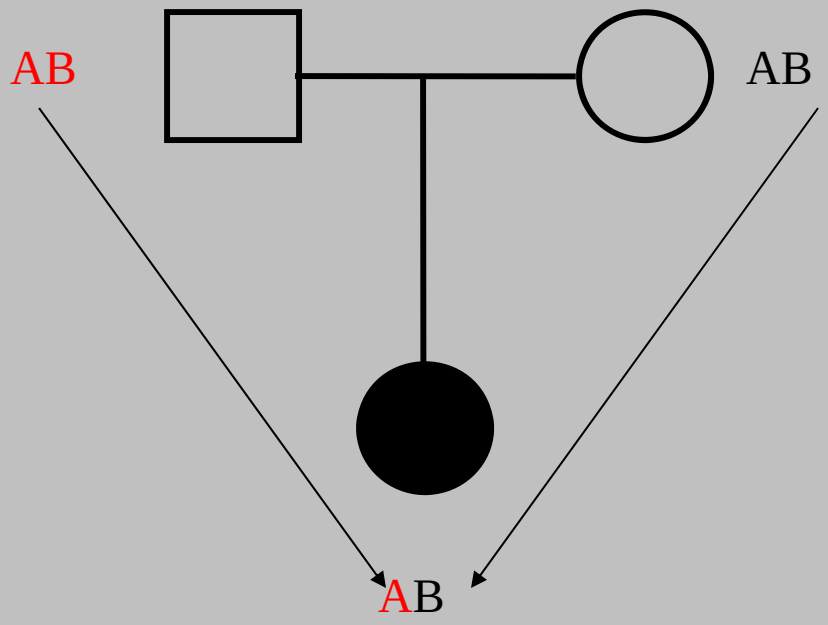
- What is it?
  - Allele frequency differences across socially defined racial and ethnic groups
- What to do about it
  - FAMILY studies
    - Trios
    - Siblings
  - Genome wide principal components

- Does it matter?
  - Yes.
  - Consider DRD2



STRATIFICATION SOLUTION 1

Transmission Disequilibrium Test



Transmitted	Not Transmitted A	Not Transmitted B
A	n_{AA}	n_{AB}
B	n_{BA}	n_{BB}

$$TDT = \frac{(n_{BA} - n_{AB})^2}{n_{BA} + n_{AB}} \sim \chi_1^2$$

Transmitted	Not Transmitted A	Not Transmitted B
A	0	1
B	1	0

Only affected offspring; only dichotomous phenotypes; biallelic markers; only an ADDITIVE model; must have full information on all members of the trio.

Stratification solution 2: more flexible family solutions

- Simply adjust for the mating type as a factor (e.g., fixed effect) in any standard generalized linear model.
- As such, the parameter estimate (the genetic effect) is conditional upon the likelihood of observing a particular genotype (I have an example shortly)

TABLE I. Potential informative strata of sufficient statistics for parental mating type under an additive genetic model

Trios (parents genotyped)	AA-AB, AB-AB, AB-BB
Sibpairs (parents missing)	(AA,AB), (AB,BB), (AA,BB)
Sibtrios (parents missing)	(AA,AA,AB), (AA,AA,BB), (AA,AB,AB), (AA,AB,BB), (AA,BB,BB), (AB,AB,BB), (AB,BB,BB)

For example, AA-AB indicates a parental mating type, and (AA,AB) indicates a configuration of sibpairs for each strata of sufficient statistic.

Mating type (6):

- AA-AA
- AA-AB
- AA-BB
- AB-AB
- AB-BB
- BB-BB

Stratification solution 3: sibling models

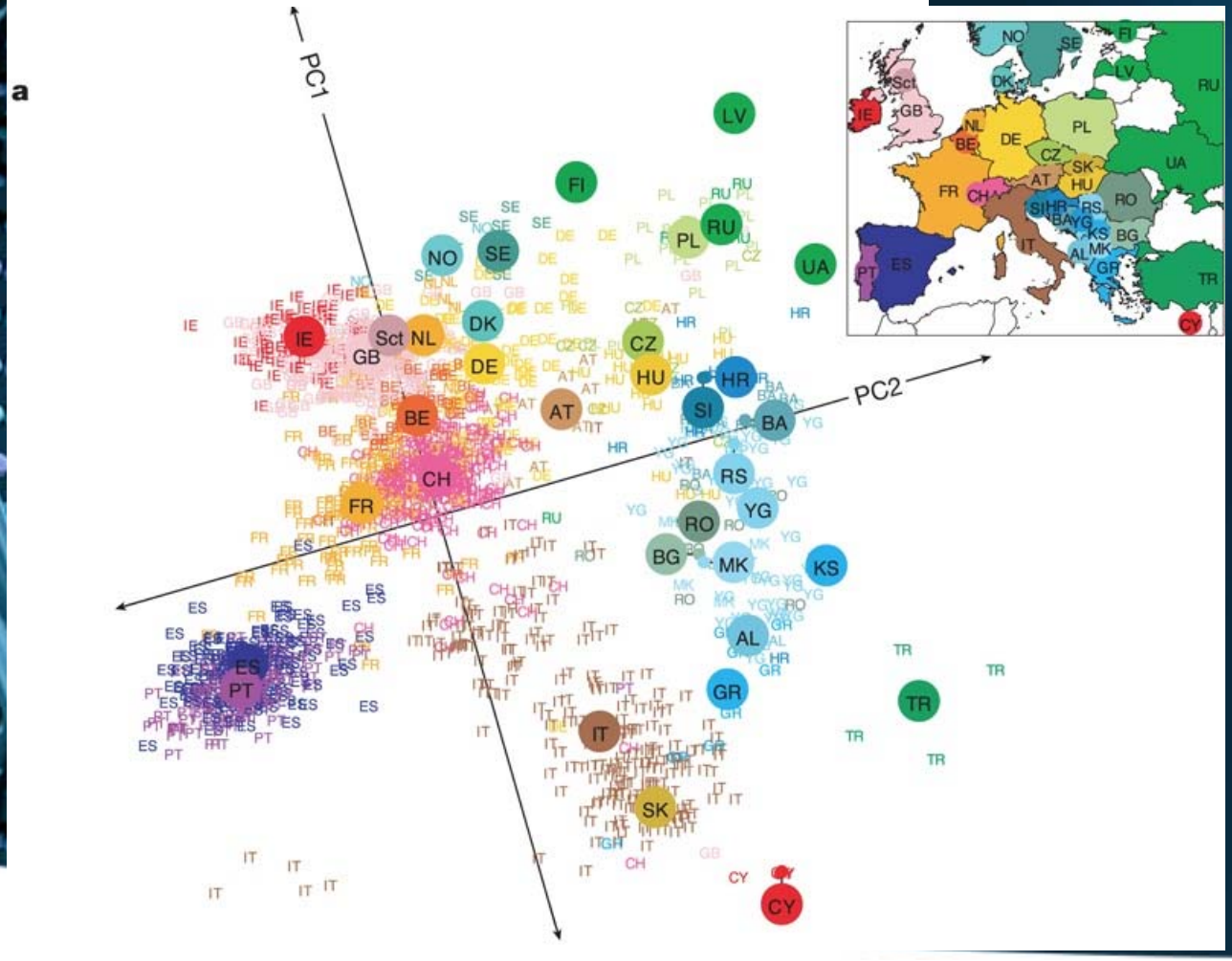
- SDT (sibling disequilibrium test) is comparable to TDT but only works with discordant pairs and is computational HUGE for GWAS data (See Horvath and Laird 1998)
- GLM with pair data is a VERY powerful model because the allocation of the ‘risk allele’ is random across siblings, they tend to share a fairly similar environment, and it can be used for many different traits, designs, etc.
- $y_{ij} = a + b_1\bar{g}_j + b_2\left(g_{ij} - \bar{g}_j\right) + u_j + e_{ij}$
- The between is subject to pop strat but not the within; comparable to a fixed effects model but allows for concordant genotypes and concordant phenotypes.

Stratification solution 3: PC models

Nature Genetics **40**, 646 - 649 (2008)
Published online: 20 April 2008 | doi:10.1038/ng.139

Interpreting principal component analyses of spatial population genetic variation

John Novembre[1,3] & Matthew Stephens[1,2]



Each individual in your study will have a value for the 1st and 2nd (3rd, 4th, etc.) principal component. These values can be used as controls in gwas models.

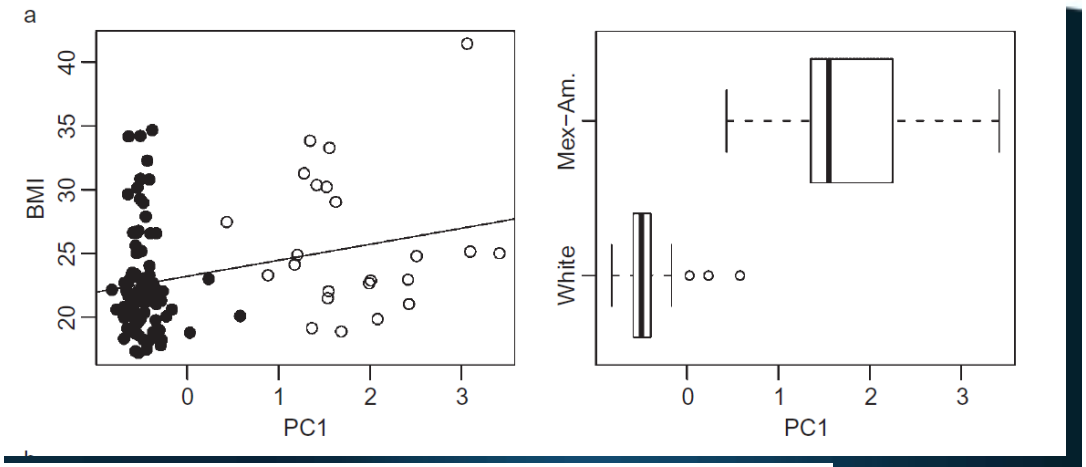
Biodemography and Social Biology, 56:123–136, 2010

ISSN: 1948-5565 print / 1948-5573 online
DOI: 10.1080/19485565.2010.524589

Routledge
Taylor & Francis Group

Ethnicity, Body Mass, and Genome-Wide Data

JASON D. BOARDMAN,[1] CASEY L. BLALOCK,[1]
ROBIN P. CORLEY,[2] MICHAEL C. STALLINGS,[2]
BENJAMIN W. DOMINGUE,[3] MATTHEW B. McQUEEN,[2]
THOMAS J. CROWLEY,[4] JOHN K. HEWITT,[1] YING LU,[5]
AND SAMUEL H. FIELD[6]



Principal components from genome-wide data: associations with BMI and self-reported ethnicity

	Variance explained		PC-BMI association		PC-ethnicity association		
	Unique	Cumulative	r (PC,BMI)	Pr ($r = 0$)	$\overline{PC}_W$	$\overline{PC}_M$	Pr (w-b = 0)
PC1	0.030	0.030	0.216	0.030	1.804	−0.472	0.000
PC2	0.014	0.044	−0.183	0.067	−0.043	0.011	0.768
PC3	0.013	0.057	0.037	0.710	0.130	−0.034	0.364
PC4	0.013	0.070	0.010	0.923	0.027	−0.007	0.843
PC5	0.012	0.083	−0.021	0.838	−0.044	0.011	0.892
PC6	0.012	0.095	0.116	0.249	0.108	−0.028	0.720
PC7	0.012	0.106	−0.063	0.531	−0.145	0.038	0.607
PC8	0.012	0.118	0.119	0.237	0.096	−0.025	0.718
PC9	0.011	0.129	0.145	0.148	−0.201	0.053	0.482
PC10	0.011	0.141	−0.015	0.881	−0.032	0.008	0.906

Note: Principal components were estimated from the 111 people across 5,490 single nucleotide polymorphisms. See the Methods section for a more detailed description of these techniques.

Genome wide heritability

- x- copies of the ref allele for the ith SNP in the jth/kth; p is the MAF; A jk is the relationship matrix for pair j,k

$$A_{jk} = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_{ij}-2p_i)(x_{ik}-2p_i)}{2p_i(1-p_i)}.$$

$$\mathbf{V} = \mathbf{A}\sigma_g^2 + \mathbf{I}\sigma_\varepsilon^2$$

gump.qimr.edu.au/gcta/

GCTA

a tool for Genome-wide Complex Trait Analysis

Overview

Download

Tutorial

FAQ

Table 1 Estimates of the variance explained by all autosomal SNPs for height, BMI, vWF and QTi

		No PC[a]		10 PCs[b]			
Trait	*n*	h^2_G (s.e.)[c]	*P*	h^2_G (s.e.)	*P*	Heritability[d]	GWAS[e]
Height	11,576	0.448 (0.029)	4.5×10^{-69}	0.419 (0.030)	7.9×10^{-48}	80–90%[32]	~10%[23]
BMI	11,558	0.165 (0.029)	3.0×10^{-10}	0.159 (0.029)	5.3×10^{-9}	42–80%[25,26]	~1.5%[14]
vWF	6,641	0.252 (0.051)	1.6×10^{-7}	0.254 (0.051)	2.0×10^{-7}	66–75%[33,34]	~13%[15]
QTi	6,567	0.209 (0.050)	3.1×10^{-6}	0.168 (0.052)	5.0×10^{-4}	37–60%[35,36]	~7%[16]

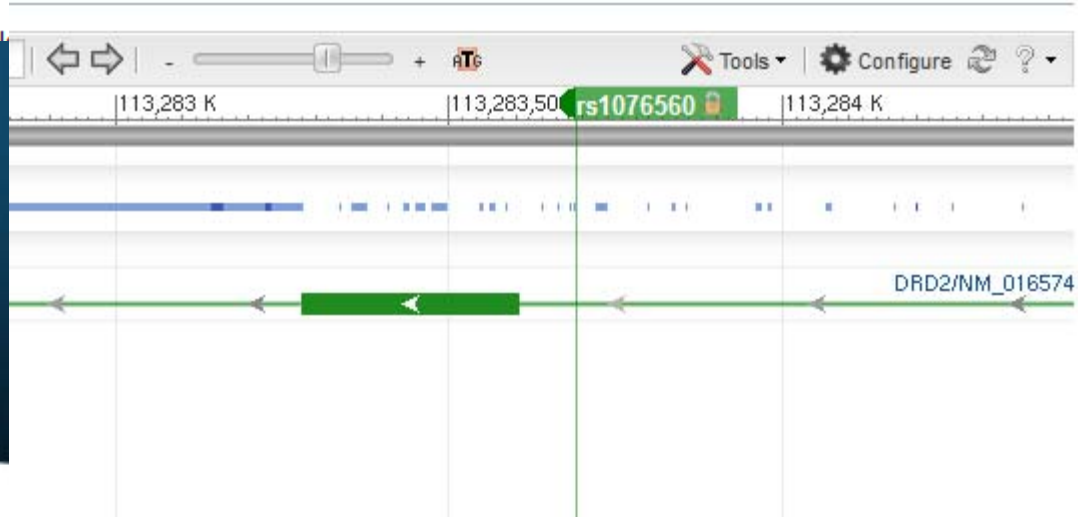
The traits vWF and QTi were available in the ARIC cohort only.
[a]Without principal component adjustment. [b]Adjustment with the first 10 principal components from principal component analysis. [c]Estimate of variance explained by all autosomal SNPs. [d]Narrow sense heritability estimate from family or twin studies from the literature. [e]Variance explained by GWAS associated loci from the literature. PC, principal component; s.e., standard error.

GWAS results, now what?

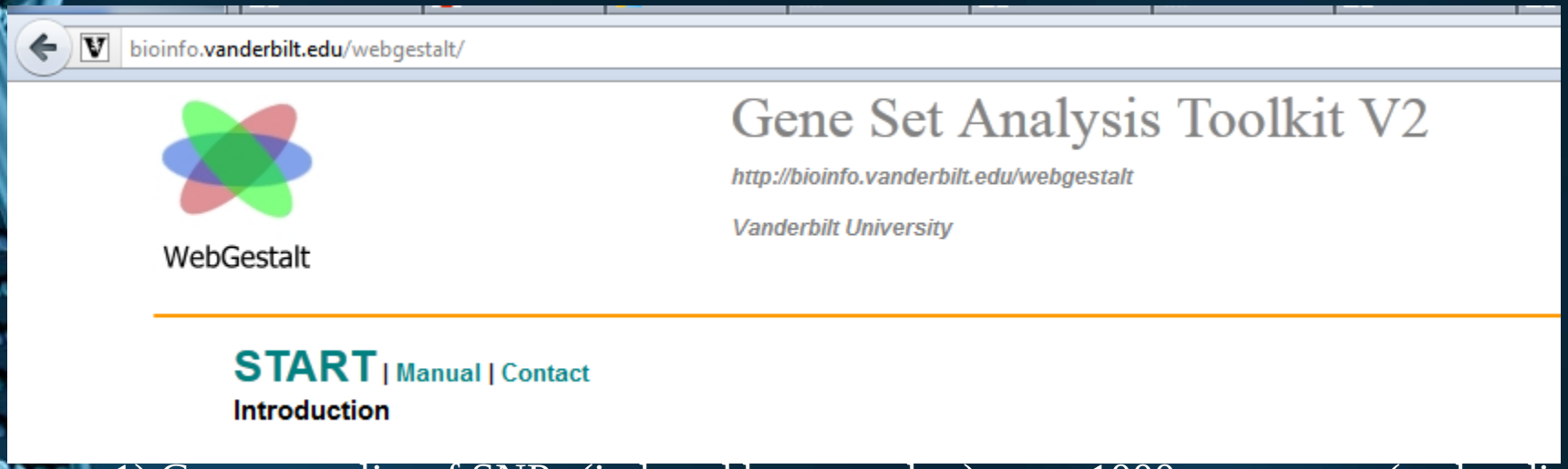
- I found a SNP that is associated with some outcome. Where is this SNP? Is it in a gene?
- Several great resources online: dbSNP

Reference SNP(refSNP) Cluster Report: rs1076560

RefSNP		Allele	
Organism:	human (Homo sapiens)	Variation Class:	SNV: single nucleotide variation
Molecule Type:	Genomic	RefSNP Alleles:	A/C
Created/Updated in build:	86/135	Allele Origin:	
Map to Genome Build:	37.3	Ancestral Allele:	C
Validation Status:		Clinical Source:	unknown
Citation:	PubMed	Clinical Significance:	NA
		MAF/MinorAlleleCount:	A=0.210/459



TONS of GWAS results: now what?



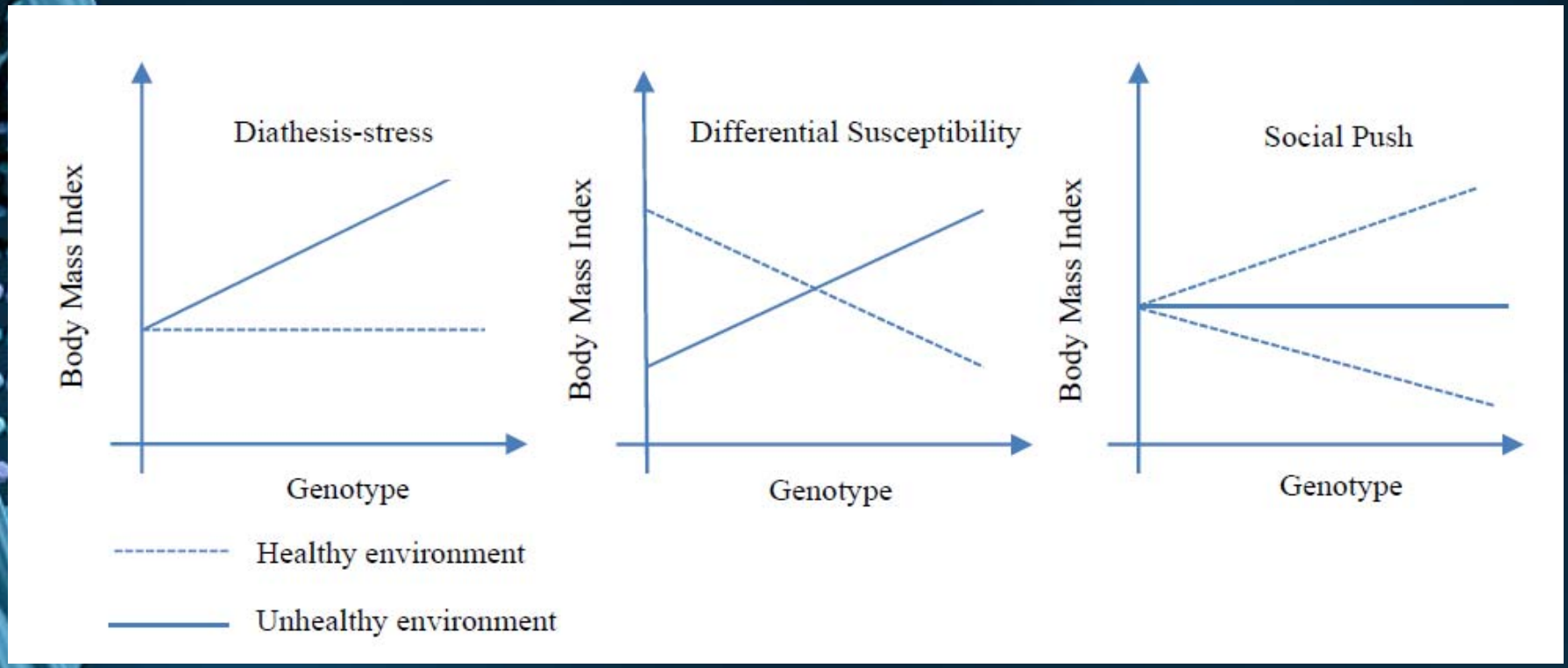
1) Generate a list of SNPs (indexed by rsnumber) – top 1000; save a .txt (no heading)
2) Go to http://bioinfo.vanderbilt.edu/webgestalt/
3) Click Start
4) Choose *hsapiens* as the organism and enter in the security code provided
5) Choose the text file saved in #1
6) For the gene ID type, select your sample and chip: I've used *hsapiens__affy_GenomeWideSNP_5_rsid*
7) Click on *Upload file*
8) Choose an Enrichment Analysis (popular ones include GO or KEGG)
9) For the Reference Set, again select *hsapiens__genome*
10) Leave everything else as default and hit *Run Enrichment Analysis*

Putting this all together: an example

- Framingham 3rd generation (G3)
- So I have their parents (G2) genotypes.
- Affy 5.0 ~ 500k snps
- Measured height and weight

- College education is the environmental variable
- I can do a GWAS and a GWGEI (genome wide gene-environment interaction model)

GxE models: stress-diathesis, differential susceptibility, and social push



Heritability estimates from GWAS data

Table 1. Descriptive statistics for all variables used in the analysis.

	College degree		
	No	Yes	Prob.
Body mass index	27.41	25.96	<.001
	(5.59)	(5.20)	
Sex [Female = 1]	0.50	0.54	<.103
Age (years)	40.32	38.83	<.001
	(7.69)	(7.47)	
Variance components			
V(G)	9.953	11.146	<.473
V(e)	15.325	11.887	<.009
Vp	25.278	23.033	<.059
V(G)/Vp	0.394	0.484	<.120
Sample Size	864	1103	

Note: all data obtained from the G3 respondents of the Framingham Heart Study (n = 1,967). P-values represent two-tailed t-tests for the difference in means between those with and without college education. The variance components describe the relative contribution of genetic (G) and environmental (E) influences to overall phenotypic variance (P) and are obtained using genetic similarity among related and unrelated individuals using genome wide data.

Gene–environment interaction tests for family studies with quantitative phenotypes: A review and extension to longitudinal measures

Hortensia Moreno-Macias,[1,2]*

$$E(Y_i|X,Z) = \beta_{0M} + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i \quad (6)$$

Where $M = 1, 2, \ldots, 6$ are the six possible

```
lm(formula = bmi ~ as.factor(snp.parent)+ snp*college + age + female, data=x)
```

But what about rGE: gene environment correlation?

$$E(Y_i|X,Z) = \beta_{0M} + \beta_1 X_i + \beta_{2M} Z_i + \beta_3 Z_i X_i \quad (9)$$

```
lm(formula = bmi ~ as.factor(snp.parent)*college+ snp*college + age + female, data=x)
```

What does this look like?

$$E(Y_i|X,Z) = \beta_{0M} + \beta_1 X_i + \beta_{2M} Z_i + \beta_3 Z_i X_i \quad (9)$$

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	25.01828	0.77096	32.451	< 2e-16
college	-1.58628	0.55161	-2.876	0.00408
snp.parent0.1	0.13055	0.53046	0.246	0.80563
snp.parent0.2	1.15436	0.82668	1.396	0.16277
snp.parent1.1	0.43227	0.67087	0.644	0.51944
snp.parent1.2	-1.11777	0.97262	-1.149	0.25061
snp.parent2.2	-1.58830	1.94385	-0.817	0.41398
snp	0.03924	0.38298	0.102	0.91841
age	0.08360	0.01627	5.140	3.05e-07
female	-2.23860	0.24498	-9.138	< 2e-16
college:snp.parent0.1	0.35279	0.71364	0.494	0.62112
college:snp.parent0.2	-1.21633	1.12536	-1.081	0.27992
college:snp.parent1.1	-0.03286	0.91738	-0.036	0.97143
college:snp.parent1.2	0.86114	1.28515	0.670	0.50290
college:snp.parent2.2	2.99966	2.54135	1.180	0.23802
college:snp	0.27914	0.50708	0.550	0.58206

E

Pop strat

G

rGE

GxE

Start with Main Genetic effects then look for interactions. Nothing.

	Model 1				Model 2							
Rs.	g1	g1pv	e1	e1pv	g2	g2pv	e2	e2pv	ge	gepv	maf	chisq
rs2180961	-1.61	1.7E-06	-1.22	8.3E-07	-1.99	1.2E-04	-1.48	1.5E-05	0.66	3.3E-01	0.15	0.25
rs949534	-1.23	4.1E-06	-1.35	6.8E-08	-1.23	1.9E-03	-1.11	2.8E-02	-0.04	9.4E-01	0.31	0.82
rs1744661	-1.50	8.2E-06	-1.22	6.0E-07	-1.85	3.7E-04	-1.44	1.8E-05	0.60	3.8E-01	0.15	0.33
rs4800019	-1.21	1.9E-05	-1.24	4.8E-07	-1.38	8.6E-04	-1.22	5.0E-03	0.23	6.9E-01	0.25	3.10
rs949092	-1.22	2.0E-05	-1.22	8.6E-07	-1.35	1.2E-03	-1.15	8.3E-03	0.17	7.6E-01	0.25	2.52
rs1107614	-1.10	2.3E-05	-1.17	6.0E-06	-1.57	8.6E-05	-1.56	6.9E-02	0.78	1.4E-01	0.44	2.31
rs1104888	1.49	2.5E-05	-1.17	1.8E-06	2.09	5.5E-05	-0.88	7.3E-03	-1.08	1.3E-01	0.13	1.67
rs949093	-1.19	2.5E-05	-1.23	6.2E-07	-1.34	1.1E-03	-1.18	6.6E-03	0.21	7.1E-01	0.25	3.19
rs1683950	-1.57	3.1E-05	-1.21	8.4E-07	-1.29	2.6E-02	-1.32	2.1E-05	-0.48	5.3E-01	0.11	2.46
rs473325	1.61	4.1E-05	-1.25	3.3E-07	2.14	3.0E-04	-1.36	7.9E-06	-0.96	2.2E-01	0.10	1.07
rs7903156	1.32	4.5E-05	-1.16	2.8E-06	1.39	3.9E-03	-1.30	1.7E-04	-0.12	8.5E-01	0.16	0.29
rs4816887	1.07	4.5E-05	-1.11	6.8E-06	0.69	7.6E-02	-1.91	2.6E-03	0.64	2.2E-01	0.37	3.73

Run GWGEI and sort by main genetic effects (Stress-diathesis social trigger).

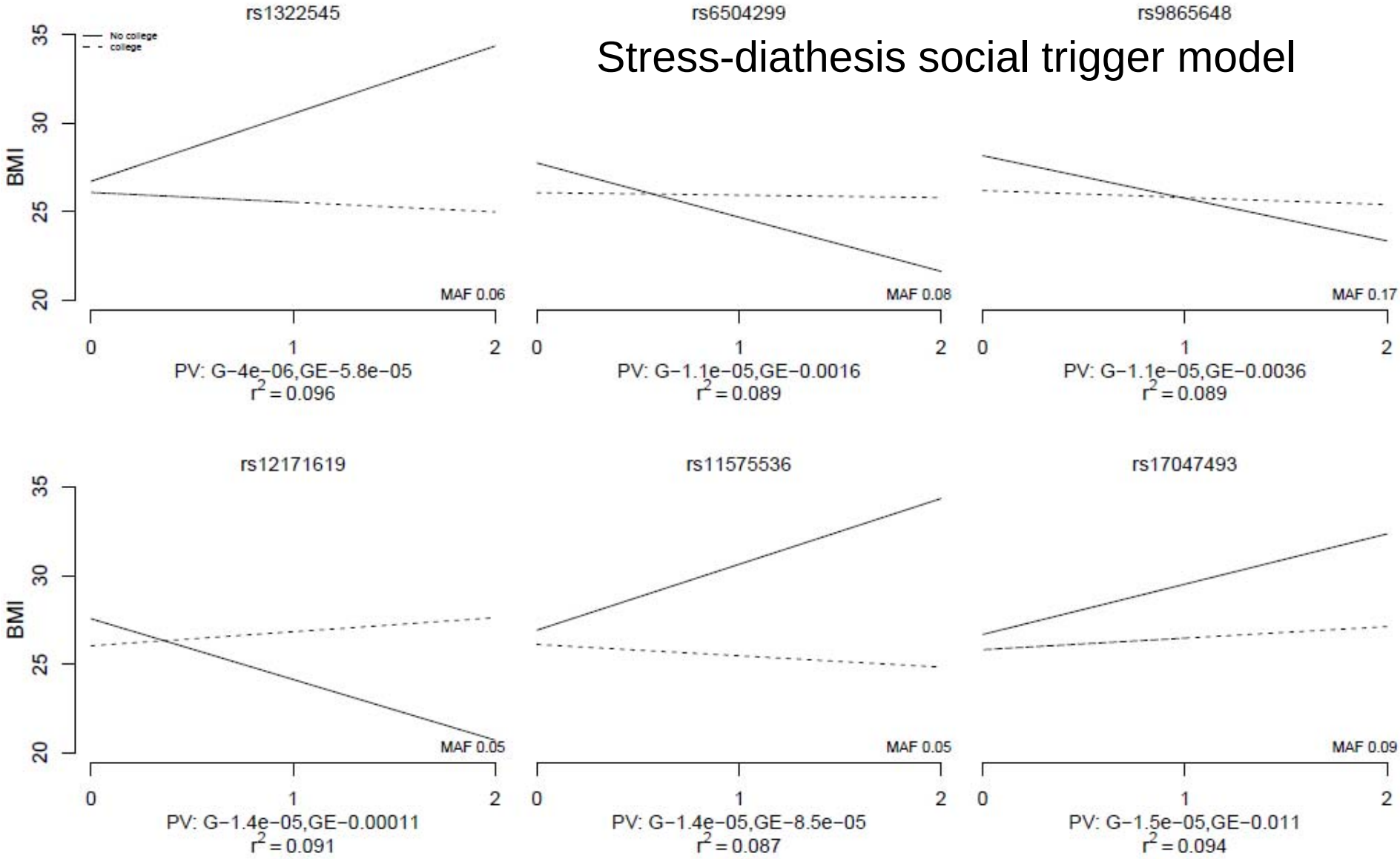
Model 1 | Model 2

Rs.	g1	g1pv	e1	e1pv	g2	g2pv	e2	e2pv	ge	gepv	maf	chisq
rs1322545	1.25	1.9E-02	-1.14	1.5E-05	3.82	4.0E-06	-1.24	2.3E-05	-4.37	5.8E-05	0.06	0.87
rs6504299	-1.47	1.4E-03	-1.21	1.3E-06	-3.06	1.1E-05	-1.09	2.1E-04	2.92	1.6E-03	0.08	0.85
rs9865648	-1.13	6.9E-04	-1.26	4.7E-07	-2.40	1.1E-05	-1.29	2.4E-04	2.01	3.6E-03	0.17	4.11
rs1217161	-1.19	2.9E-02	-1.13	7.1E-06	-3.43	1.4E-05	-0.79	4.5E-03	4.22	1.1E-04	0.05	0.85
rs1157553	1.07	4.8E-02	-1.20	3.6E-06	3.71	1.4E-05	-1.06	2.4E-04	-4.35	8.5E-05	0.05	1.88
rs1704749	1.54	2.9E-04	-1.22	1.7E-06	2.84	1.5E-05	-1.07	4.5E-04	-2.19	1.1E-02	0.09	0.94
rs363323	-0.70	1.4E-01	-1.22	7.6E-07	-3.39	1.6E-05	-1.30	4.1E-06	4.00	4.7E-05	0.07	5.19
rs1093608	-0.88	1.4E-03	-1.24	4.9E-07	-1.84	1.7E-05	-1.79	1.4E-04	1.65	3.2E-03	0.27	0.21
rs1522392	-0.86	1.7E-03	-1.24	4.6E-07	-1.84	1.8E-05	-1.70	2.5E-04	1.68	2.6E-03	0.27	0.13
rs7315837	0.67	2.5E-02	-1.19	1.5E-06	1.87	2.1E-05	-1.24	1.7E-03	-2.27	1.5E-04	0.20	1.85
rs1081896	-0.96	2.3E-04	-1.23	6.5E-07	-1.71	2.5E-05	-1.46	7.7E-03	1.28	1.6E-02	0.37	6.52
rs1682732	-0.83	1.4E-03	-1.22	7.1E-07	-1.69	2.6E-05	-1.69	1.8E-03	1.50	4.1E-03	0.32	0.44

Compare these: genetic effect is offset for college graduates

Sort by “main genetic” p-value in GxE
No genetic inf. among Coll. educated

Stress-diathesis social trigger model

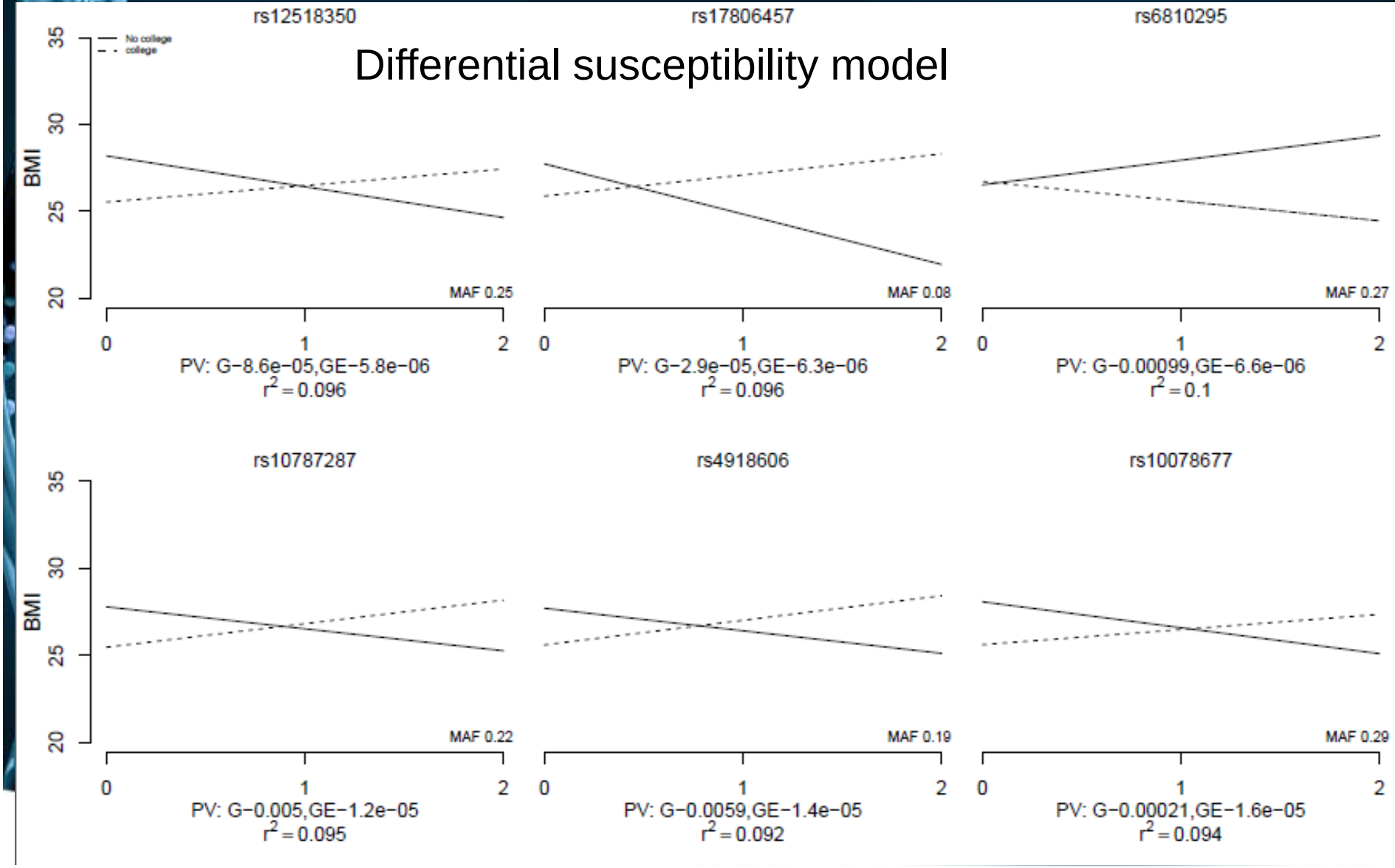


Run GWGEI and sort by GxE pvalues (differential susceptibility).

Model 1 | Model 2

Rs.	g1	g1pv	e1	e1pv	g2	g2pv	e2	e2pv	ge	gepv	maf	chisq
rs1251835	-0.23	4.4E-01	-1.30	5.6E-07	-1.77	8.6E-05	-1.66	2.3E-04	2.72	5.8E-06	0.25	1.64
rs1780645	-0.50	2.7E-01	-1.22	7.6E-07	-2.89	2.9E-05	-1.13	9.0E-05	4.10	6.3E-06	0.08	1.74
rs6810295	-0.05	8.6E-01	-1.17	1.0E-05	1.41	9.9E-04	-1.09	4.1E-02	-2.54	6.6E-06	0.27	0.27
rs1078728	0.23	4.4E-01	-1.19	1.9E-06	-1.26	5.0E-03	-1.25	1.6E-03	2.62	1.2E-05	0.22	0.26
rs4918606	0.29	3.4E-01	-1.10	1.0E-05	-1.30	5.9E-03	-1.09	3.5E-03	2.71	1.4E-05	0.19	3.28
rs1007867	-0.21	4.4E-01	-1.11	9.4E-06	-1.49	2.1E-04	-1.37	5.1E-03	2.36	1.6E-05	0.29	3.81
rs1761856	-0.09	7.4E-01	-1.17	2.7E-06	-1.42	6.3E-04	-1.04	1.8E-02	2.41	1.6E-05	0.26	0.74
rs1268430	-0.15	5.9E-01	-1.28	5.0E-07	1.17	5.7E-03	-0.91	3.0E-02	-2.45	1.9E-05	0.22	2.84
rs4358837	0.49	3.5E-01	-1.28	3.9E-07	-2.17	7.8E-03	-1.35	1.5E-06	4.56	2.0E-05	0.06	0.17
rs9572541	-0.34	1.6E-01	-1.21	9.0E-07	0.83	2.3E-02	-1.21	7.7E-02	-2.09	2.1E-05	0.38	0.69
rs1741088	0.29	6.3E-01	-1.17	5.5E-05	-2.88	2.6E-03	-1.22	1.8E-04	5.22	2.1E-05	0.06	0.74
rs1049495	-0.10	7.4E-01	-1.36	1.3E-07	-1.60	6.1E-04	-1.21	3.6E-03	2.61	2.2E-05	0.20	0.90

Sort by “GxE” p-value; all cross overs



Gene-environment correlation as a potential source of bias

American Journal of Epidemiology

Vol. 175, No. 3
DOI: 10.1093/aje/kwr368
Advance Access publication:
December 22, 2011

Practice of Epidemiology

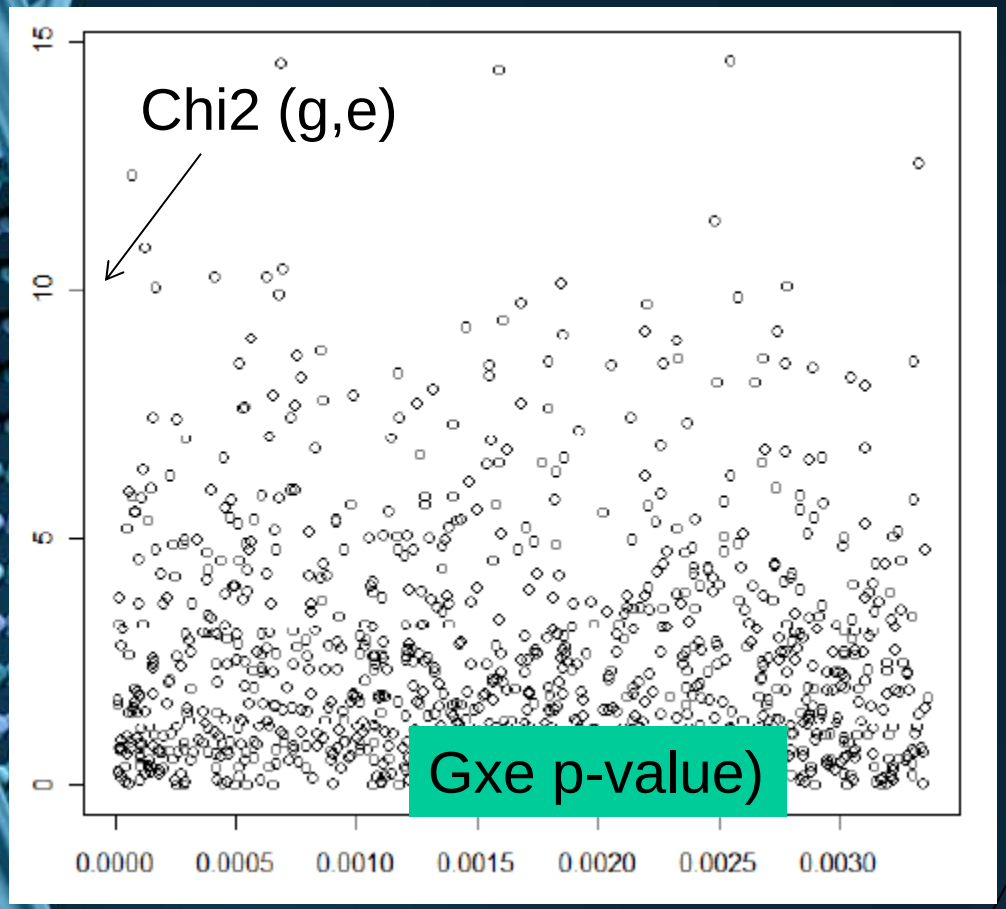
Gene-Environment Interactions in Genome-Wide Association Studies: A Comparative Study of Tests Applied to Empirical Studies of Type 2 Diabetes

Marilyn C. Cornelis*, Eric J. Tchetgen Tchetgen, Liming Liang, Lu Qi, Nilanjan Chatterjee, Frank B. Hu, and Peter Kraft

SNP	Region	No. of SNPs in LD[a]	MAF ("risk") for Cases	MAF ("risk") for Controls	*G-E* Association		Marginal *G* Effect		Standard	
					β Coefficient	*P* Value	β Coefficient	*P* Value	β Coefficient	*P* Value
Standard										
rs2110169	12p13	0	0.37	0.39	0.34	2.2×10^{-5}	−0.103	0.06	−0.55	2.4×10^{-6}
rs10512802	3q21	0	0.03	0.03	0.88	6.6×10^{-5}	0.037	0.82	−1.52	4.1×10^{-6}
rs9992452	4q35	0	0.16	0.15	0.44	2.5×10^{-5}	0.030	0.67	−0.68	6.5×10^{-6}
rs11669653	19q13.1	0	0.30	0.30	0.23	0.0058	0.002	0.97	−0.54	8.7×10^{-6}

What about rGE?

```
lm(formula = bmi ~ as.factor(snp.parent)+ snp*college + age + female, data=x)
```



Rs.	maf	chisq	with rGE	without rGE
rs7609257	0.08	5.33	1.3E-04	2.7E-06
rs4234656	0.06	3.56	7.2E-02	3.1E-06
rs17805193	0.06	0.95	8.3E-03	5.3E-06
rs4861426	0.30	7.40	1.8E-02	9.3E-06
rs11138365	0.21	1.95	8.6E-03	1.1E-05
rs2042861	0.28	4.96	2.8E-04	1.3E-05
rs7304582	0.34	0.39	1.4E-01	1.3E-05
rs6766503	0.06	4.08	1.2E-01	1.3E-05
rs6976501	0.19	2.28	2.9E-03	1.3E-05
rs17074039	0.07	0.56	9.7E-05	1.4E-05
rs11993516	0.21	3.24	4.6E-04	1.7E-05
rs17005475	0.08	2.39	1.3E-03	2.4E-05
rs1882111	0.31	8.59	2.7E-03	2.6E-05
rs2179766	0.13	0.18	1.8E-01	3.0E-05
rs12312031	0.10	0.86	3.2E-02	3.1E-05
rs17834425	0.17	0.42	9.1E-03	3.4E-05
rs6585026	0.12	0.38	1.3E-04	3.6E-05
rs17001408	0.20	4.30	2.8E-03	3.6E-05
rs6489945	0.35	6.58	1.1E-01	4.0E-05
rs10849918	0.46	1.10	1.2E-03	4.0E-05

```
lm(formula = bmi ~ as.factor(snp.parent)*college+ snp*college + age + female, data=x)
```

What about physiological function?

Table 4. Functional gene networks by GWAS and GWGEI significance of top 500 SNPs.

SNPs in cluster	Gene	Name, description of gene
Main Genetic Associations, No GxE Specified: Metabolic Pathways (p<.0119)		
rs12585267, rs682573, rs1170103, rs1170102, rs7981733	DGKH	diacylglycerol kinase, eta
rs542721	DMGDH	dimethylglycine dehydrogenase
rs367766	PRODH	proline dehydrogenase (oxidase) 1
rs1466427	AMPD3	adenosine monophosphate deaminase (isoform E)
rs1374670, rs4764407, rs10743273, rs7305694	PIK3C2G	phosphoinositide-3-kinase, class 2, gamma polypeptide
rs315076	ST6GALNAC3	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3
rs9638627	WBSCR17	Williams-Beuren syndrome chromosome region 17
rs486408	AGL	amylo-1, 6-glucosidase, 4-alpha-glucanotransferase
rs10831631	GALNTL4	UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-like 4
GxE Model, sort by Main Genetic Associations: Pathways in Cancer (p<.0010)		
rs990395, rs9585802, rs9513980	FGF14	fibroblast growth factor 14
rs1011814, rs6451758	FGF10	fibroblast growth factor 10
rs17446614, rs2180961	FOXO1	forkhead box O1
rs11576886	JAK1	Janus kinase 1
rs2281879, rs11191347, rs10883728, rs7094366	SUFU	suppressor of fused homolog (Drosophila)
rs655988	ZBTB16	zinc finger and BTB domain containing 16
rs7611561	FGF12	fibroblast growth factor 12
rs10029339, rs7668374, rs6827698	MAPK10	mitogen-activated protein kinase 10
rs2916528	CTNNA2	catenin (cadherin-associated protein), alpha 2
GxE Model Sort by GxE Associations: Axon Guidance (p<.0082)		
rs296005	SLIT3	slit homolog 3 (Drosophila)
rs9682089	ROBO2	roundabout, axon guidance receptor, homolog 2 (Drosophila)
rs3775823	SLIT2	slit homolog 2 (Drosophila)
rs7295635, rs6581526	SRGAP1	SLIT-ROBO Rho GTPase activating protein 1