Relative Risk Example

What is the likelihood of seeing a trait given a genotype compared to the overall likelihood of seeing that trait?

Dry ear wax?
rs17822931

	Wet	Dry
TT	0	19
CT/CC	56	4

For TT, 19/19 have dry wax = 1

For entire class, 23/79 have dry ear wax = .29

Relative Risk = 1/.29 = 3.4

http://stanford.edu/class/gene210/files/writeups/2012/Odds_Ratio_and_Likelihood_Ratio.pdf

Class GWAS (n=79)

1. Allele counts

```
earwax (4988235):
      Wet       Dry
A      44         5
G      68         41
earwax (7495174):
      Wet       Dry
A      95         25
G      17         21
earwax (713598):
      Wet       Dry
C      56         20
G      56         26
earwax (17822931):
      Wet       Dry
C      89         8
T      23         38
earwax (4481887):
      Wet       Dry
A      36         10
G      76         36
```

Class GWAS (n=79)
2. Allele p-values

phenotype	rs4988235	rs7495174	rs713598	rs17822931	rs4481887
earwax	0.0004513	4.67E-05	0.4561	3.32E-13	0.191
eyes	0.0003015	7.26E-07	0.5007	5.43E-06	0.07006
asparagus	0.7053	0.1705	0.678	0.9334	0.08973
bitter	0.2804	0.8256	3.72E-05	0.5863	0.0229
lactose	0.0008853	0.3083	0.7809	0.1813	0.02298

Class GWAS
3. genotype counts

```
earwax (4988235):
      Wet        Dry
AA      12         2
AG/GG   44         21

GG      24         20
AA/AG   32         3

earwax (7495174):
      Wet        Dry
AA      40         9
AG/GG   16         14

GG      1          7
AA/AG   55         16

earwax (713598):
      Wet        Dry
CC      13         2
CG/GG   43         21

GG      13         5
CC/CG   43         18
```

```
earwax (17822931):
      Wet        Dry
CC      33         4
CT/TT   23         19

TT      0          19
CC/CT   56         4

earwax (4481887):
      Wet        Dry
AA      7          1
AG/GG   49         22

GG      27         14
AA/AG   29         9
```

Class GWAS
3. genotype counts

```
earwax (17822931):
       Wet        Dry
CC      33          4
CT/TT   23          19

TT      0          19
CC/CT   56          4
```

T is a null allele in ABC11
T/T has dry wax. T/C and C/C have wet earwax usually.

Class GWAS
3. genotype p-values

	17822931 (CC rec.)	(TT rec.)
earwax	0.00105	**1.002e-14**
eyes	0.001628	0.0001687
asparagus	0.3713	0.4271
bitter	0.4654	1
lactose	0.6059	0.229

Allelic p-value for rs17822931 is 3e-13
model where TT is recessive fits the data best.

Class GWAS
3. genotype p-values

	4988235 (AA rec.)	(GG rec.)	7495174 (AA rec.)	(GG rec.)	713598 (CC rec.)	(GG rec.)	17822931 (CC rec.)	(TT rec.)	4481887 (AA rec.)	(GG rec.)
earwax	0.2145	0.0003765	0.01061	0.000545	0.208	1	0.00105	**1.002e-14**	0.4258	0.3331
eyes	0.05695	0.003398	**1.675e-07**	0.04753	1	0.3937	0.001628	0.0001687	1	0.01573
asparagus	1	0.6544	0.3575	0.2798	0.779	0.7896	0.3713	0.4271	0.1284	**0.3643**
bitter	0.533	0.4685	1	1	**7.078e-07**	0.04713	0.4654	1	0.09783	0.09343
lactose	0.1023	**0.003507**	0.5947	0.4103	0.7485	0.538	0.6059	0.229	0.1916	0.07385

Class GWAS results

Known Associations:

Lactose intolerance: rs4988235, GG associated with lactose intolerance
Eye color: rs7495174, AA associated with blue/green eyes
Bitter taste: rs713598, CC associated with inability to taste bitterness
Earwax: rs17822931, TT associated with dry earwax

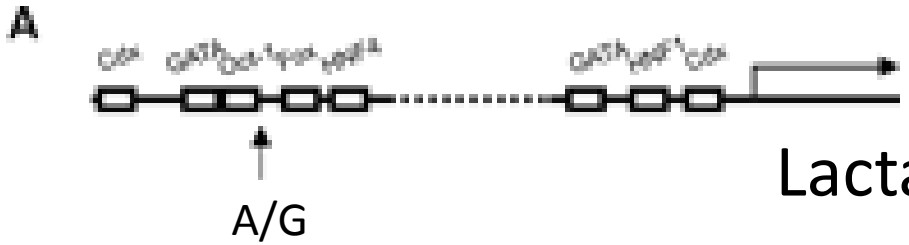
Class GWAS

Odds Ratio, Likelihood Ratio, Relative Risk

		P-value	**OR**	**LR**	**RR**
Lactose Intolerance	rs4988235	8e-4	5.5	1.45	1.5
Eye Color	rs7495174	1e-7	Large	1.55	1.6
Bitter Taste	rs713598	7e-7	4.5	1.95	2.8
Earwax	rs17822931	1e-14	18	4.02	3.4

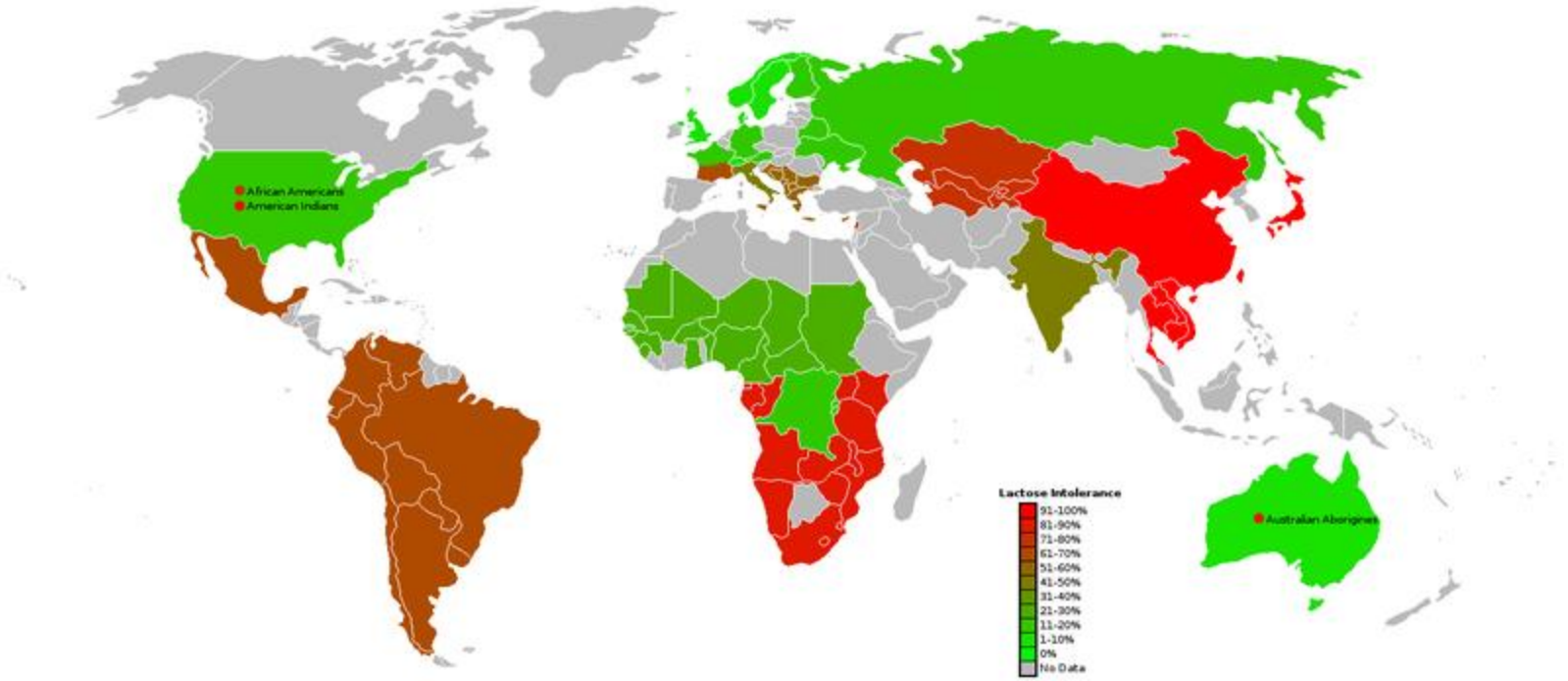
Lactose Intolerance

Rs4988235



A – lactase expressed in adulthood
G – lactase expression turns off in adulthood

Lactose Intolerance



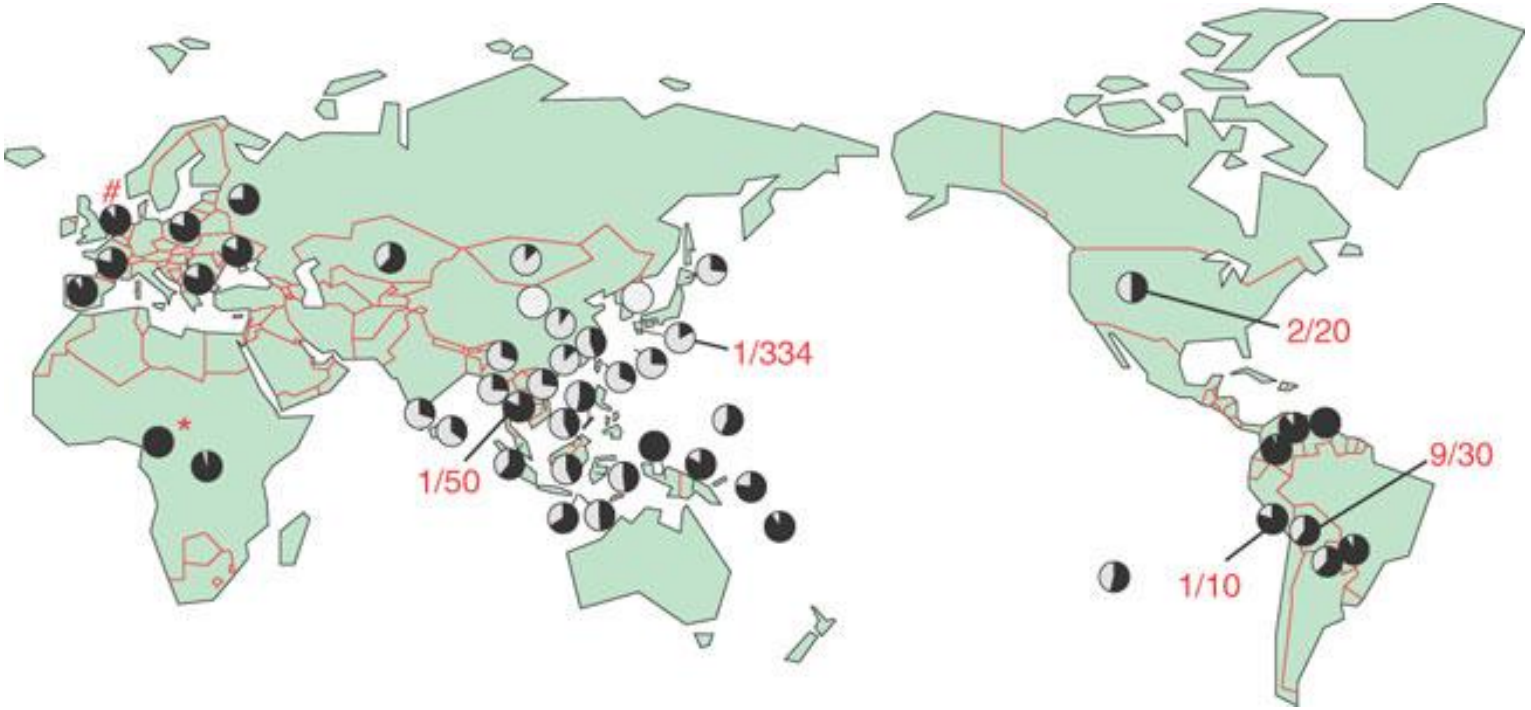
Eye Color

Rs7495174
In OCA2, the oculocutaneous albinism gene (also known as the human P protein gene).
Involved in making pigment for eyes, skin, hair.
accounts for 74% of variation in human eye color.
Rs7495174 leads to reduced expression in eye specifically.
Null alleles cause albinism

Ear Wax

Rs17822931
In *ABCC11* gene that transports various molecules across extra- and intra-cellular membranes.
The A allele is loss of function of the protein.
Phenotypic implications of wet earwax: Insect trapping, self-cleaning and prevention of dryness of the external auditory canal.
Wet earwax: linked to axillary odor and apocrine colostrum.

Ear Wax



Rs17822931
“the allele A arose in northeast Asia and thereafter spread through the world.”

Asparagus

Certain compounds in asparagus
are metabolized to yield ammonia and various sulfur-containing degradation products, including various thiols and thioesters, which give urine a characteristic smell.
Methanethiol (pungent)
dimethyl sulfide (pungent)
dimethyl disulfide
bis(methylthio)methane
dimethyl sulfoxide (sweet aroma)
dimethyl sulfone (sweet aroma)

rs4481887 is in a region containing 39 olfactory receptors

Bitter taste

Rs713598
Taste receptor 2 member 38
accounts for up to 85% of the variation in PTC tasting ability

Genetic principles are universal

Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms

DAVID BOTSTEIN,[1] RAYMOND L. WHITE,[2] MARK SKOLNICK,[3] AND RONALD W. DAVIS[4]

Am J Hum Genet. 1980 May;32(3):314-31.

Events
News
News Archive
Featured Video

News Archive

Back to home page.

David Botstein among 11 recipients of inaugural $3 million prize

Posted Feb 22, 2013 By by Staff

David Botstein, Princeton University's Anthony B. Evnin '62 Professor of Genomics and molecular biology and director of the Lewis-Sigler Institute for Integrative Genomics, was among 11 recipients of the inaugural Breakthrough Prize in Life Sciences. The $3 million award recognizes Botstein for his "linkage mapping of Mendelian disease [one caused by a single gene mutation such as sickle-cell anemia or cystic fibrosis] in humans using DNA polymorphisms," according to the Breakthrough Prize in Life Sciences Foundation. The foundation supports outstanding research intended to cure "intractable diseases and extend human life."

Different genetics for different traits

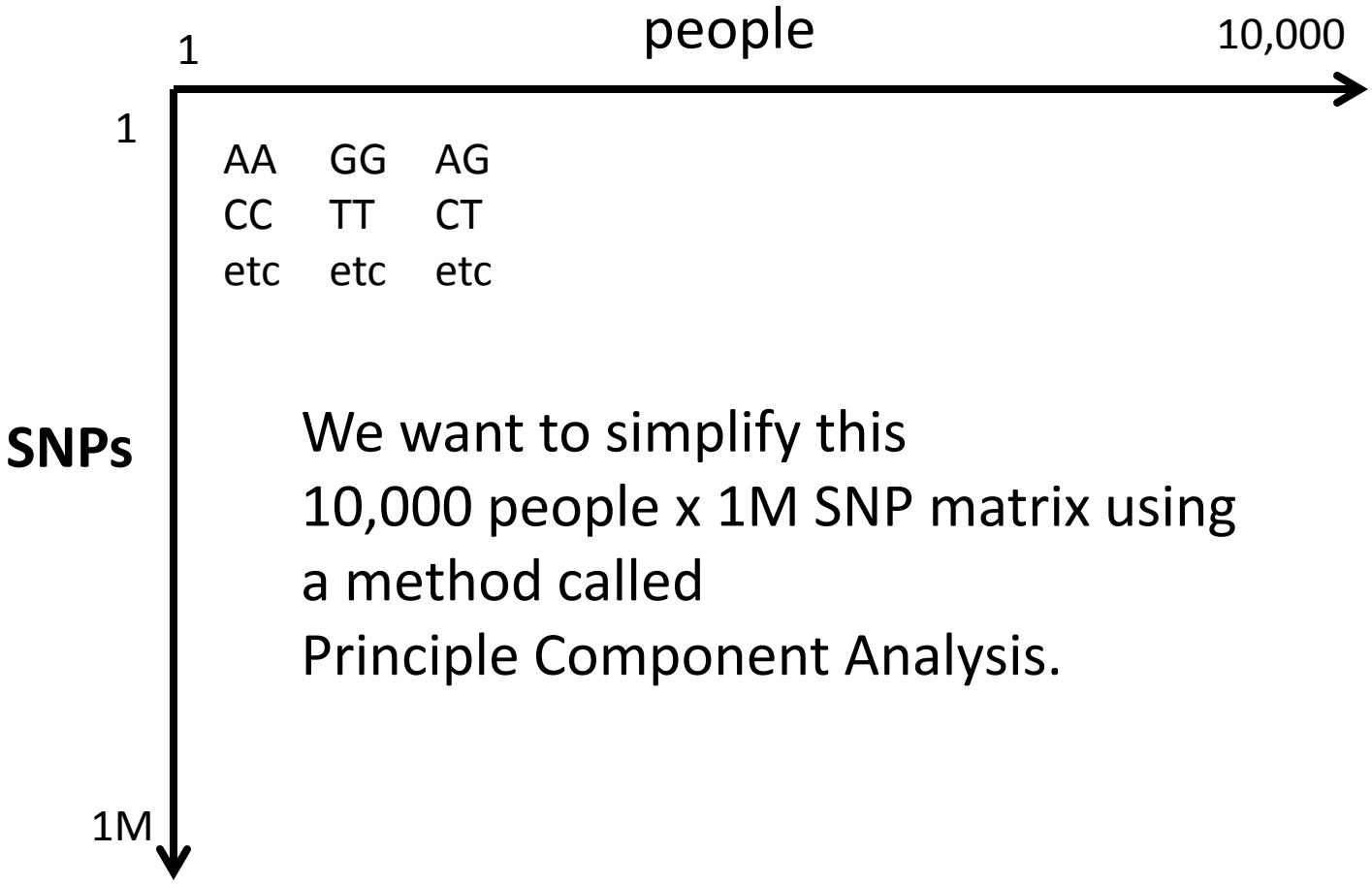
Simple: Lactose tolerance, asparagus smell, photic sneeze
Complex: T2D, CVD
Same allele: CFTR, sickle cell
Different alleles: BRCA1, hypertrophic cardiomyopathy

Ancestry

Go to Genotation, Ancestry, PCA (principle components analysis)
Load in genome.
Start with HGDP world
Resolution 10,000
PC1 and PC2

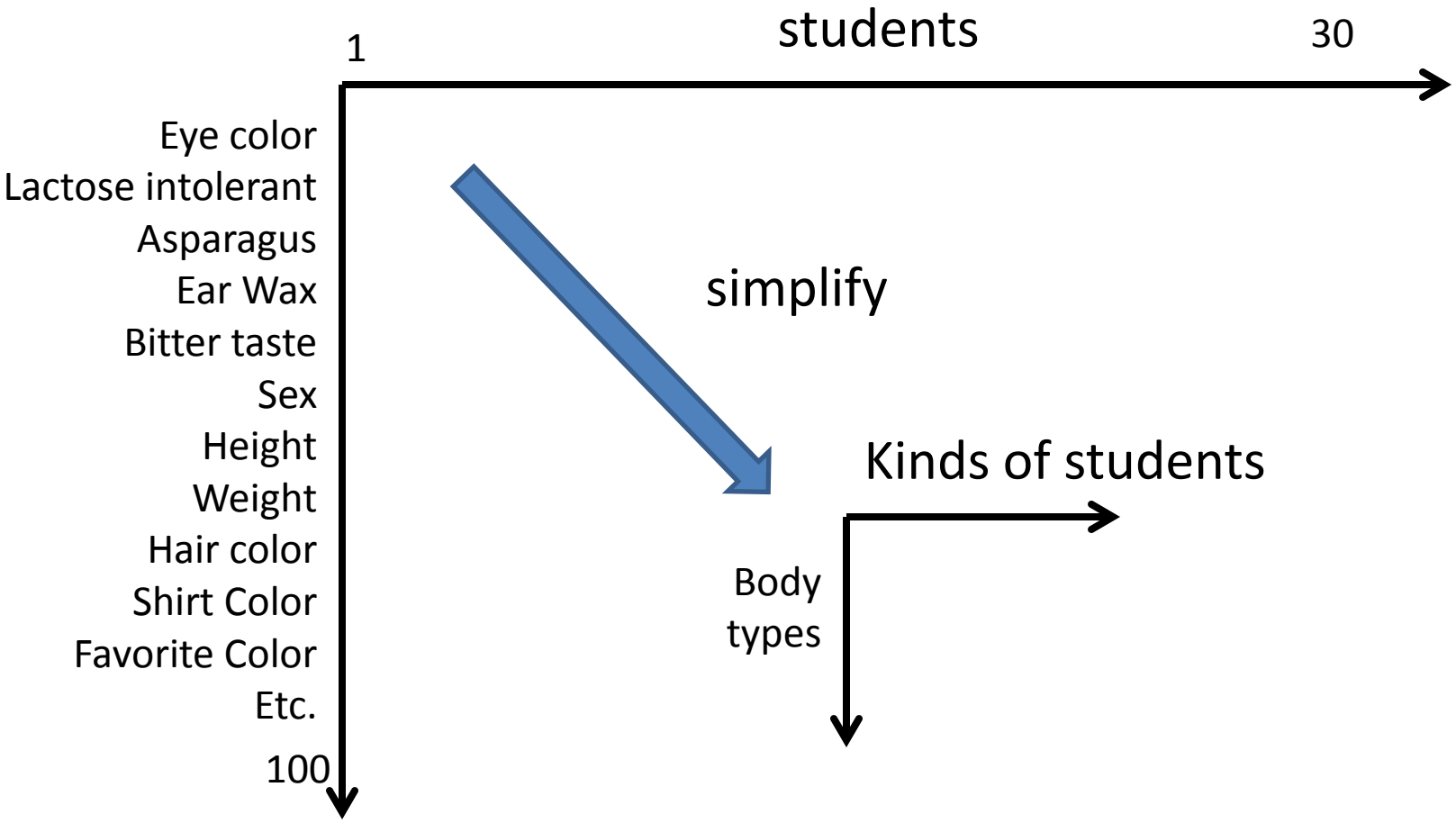
Then go to Ancestry, painting

Ancestry Analysis



We want to simplify this
10,000 people x 1M SNP matrix using
a method called
Principle Component Analysis.

PCA example


students
1
30
Eye color
Lactose intolerant
Asparagus
Ear Wax
Bitter taste
Sex
Height
Weight
Hair color
Shirt Color
Favorite Color
Etc.
100
simplify
Kinds of students
Body types

Informative traits

Skin color
eye color
height
weight
sex
hair length
etc.

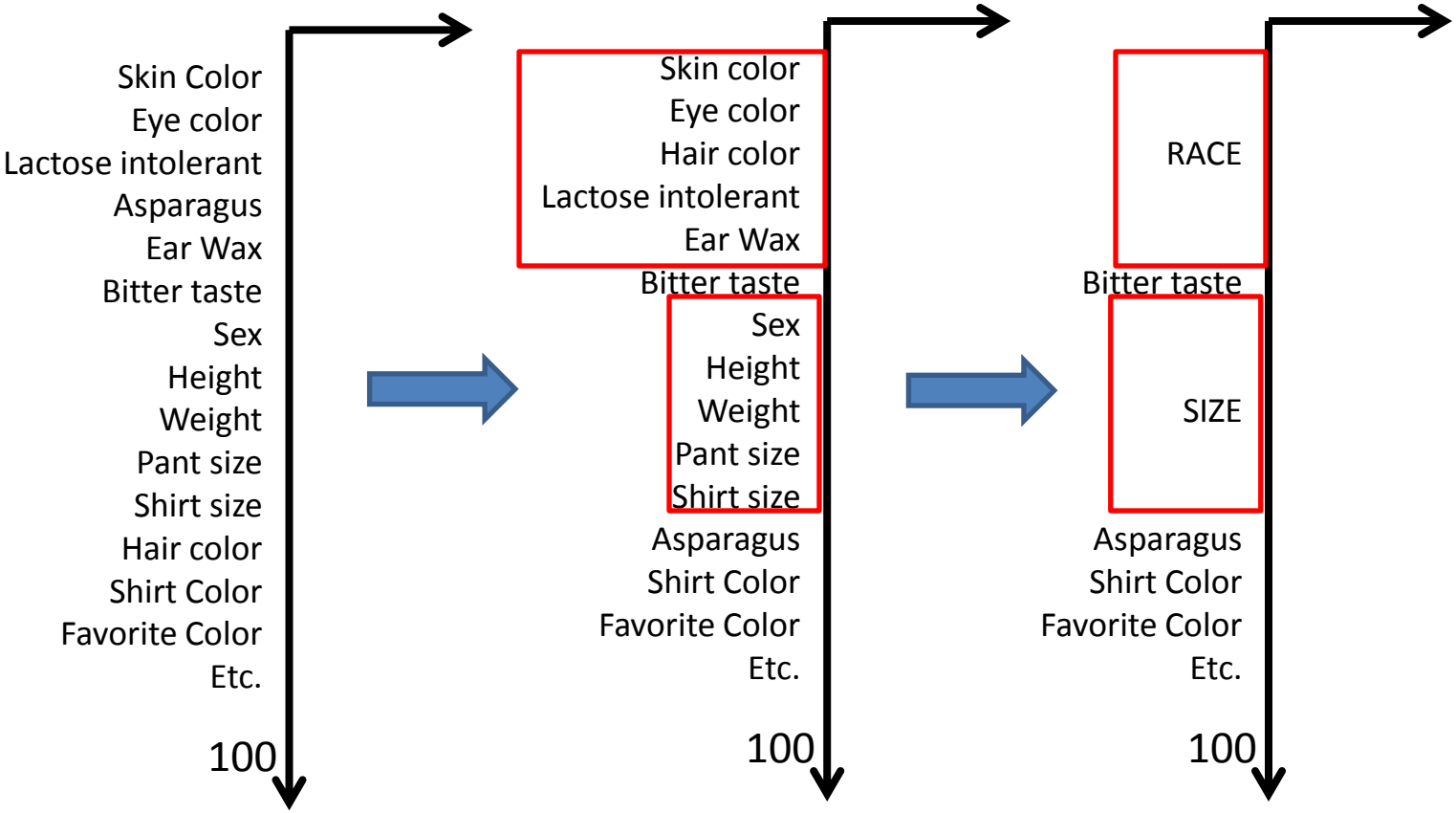
~SNPs informative for ancestry

Uninformative traits

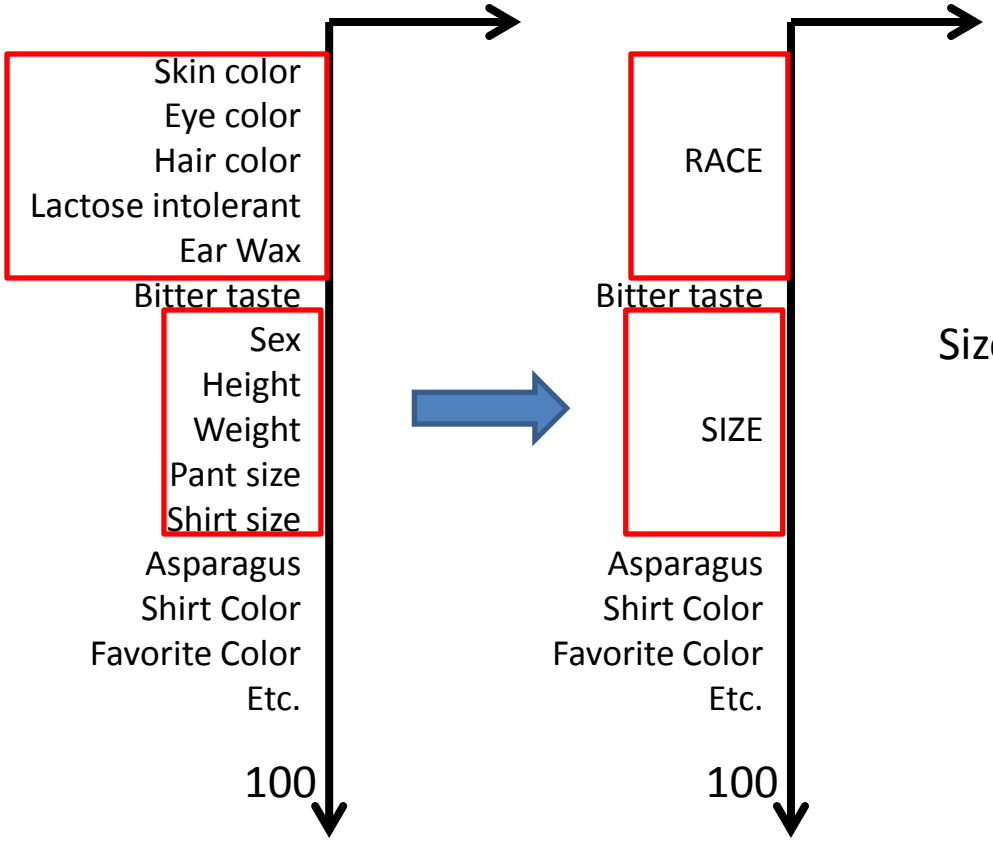
shirt color
Pants color
favorite toothpaste
favorite color
etc.

~SNPs not informative for ancestry

PCA example


Skin Color
Eye color
Lactose intolerant
Asparagus
Ear Wax
Bitter taste
Sex
Height
Weight
Pant size
Shirt size
Hair color
Shirt Color
Favorite Color
Etc.
100
Skin color
Eye color
Hair color
Lactose intolerant
Ear Wax
Bitter taste
Sex
Height
Weight
Pant size
Shirt size
Asparagus
Shirt Color
Favorite Color
Etc.
100
RACE
Bitter taste
SIZE
Asparagus
Shirt Color
Favorite Color
Etc.
100

PCA example



Size = Sex + Height + Weight +
Pant size + Shirt size …

Ancestry Analysis

	1	2	3	4	5	6	7
Snp1	A	A	A	A	A	A	T
Snp2	G	G	G	G	G	G	G
Snp3	A	A	A	A	A	A	T
Snp4	C	C	C	T	T	T	T
Snp5	A	A	A	A	A	A	G
Snp6	G	G	G	A	A	A	A
Snp7	C	C	C	C	C	C	A
Snp8	T	T	T	G	G	G	G
Snp9	G	G	G	G	G	G	T
Snp10	A	G	C	T	A	G	C
Snp11	T	T	T	T	T	T	C
Snp12	G	C	T	A	A	G	C

Reorder the SNPs

	1	**2**	**3**	**4**	**5**	**6**	**7**
Snp1	**A**	**A**	**A**	**A**	**A**	**A**	**T**
Snp3	**A**	**A**	**A**	**A**	**A**	**A**	**T**
Snp5	**A**	**A**	**A**	**A**	**A**	**A**	**G**
Snp7	**C**	**C**	**C**	**C**	**C**	**C**	**A**
Snp9	**G**	**G**	**G**	**G**	**G**	**G**	**T**
Snp11	**T**	**T**	**T**	**T**	**T**	**T**	**C**
Snp2	G	G	G	G	G	G	G
Snp4	C	C	C	T	T	T	T
Snp6	G	G	G	A	A	A	A
Snp8	T	T	T	G	G	G	G
Snp10	A	G	C	T	A	G	C
Snp12	G	C	T	A	A	G	C

Ancestry Analysis

	1	2	3	4	5	6	7
Snp1	**A**	**A**	**A**	**A**	**A**	**A**	**T**
Snp3	**A**	**A**	**A**	**A**	**A**	**A**	**T**
Snp5	**A**	**A**	**A**	**A**	**A**	**A**	**G**
Snp7	**C**	**C**	**C**	**C**	**C**	**C**	**A**
Snp9	**G**	**G**	**G**	**G**	**G**	**G**	**T**
Snp11	**T**	**T**	**T**	**T**	**T**	**T**	**C**
Snp4	C	C	C	T	T	T	T
Snp6	G	G	G	A	A	A	A
Snp8	T	T	T	G	G	G	G
Snp2	G	G	G	G	G	G	G
Snp10	A	G	C	T	A	G	C
Snp12	G	C	T	A	A	G	C

Ancestry Analysis

	1	2	3	4	5	6	7
Snp1	A	A	A	A	A	A	T
Snp3	A	A	A	A	A	A	T
Snp5	A	A	A	A	A	A	G
Snp7	C	C	C	C	C	C	A
Snp9	G	G	G	G	G	G	T
Snp11	T	T	T	T	T	T	C

	1-6	7
Snp1	A	T
Snp3	A	T
Snp5	A	G
Snp7	C	A
Snp9	G	T
Snp11	T	C

	1
Snp1	A
Snp3	A
Snp5	A
Snp7	C
Snp9	G
Snp11	T

=X

	7
Snp1	T
Snp3	T
Snp5	G
Snp7	A
Snp9	T
Snp11	C

=X

Ancestry Analysis

	1	2	3	4	5	6	7
Snp1	A	A	A	A	A	A	T
Snp3	A	A	A	A	A	A	T
Snp5	A	A	A	A	A	A	G
Snp7	C	C	C	C	C	C	A
Snp9	G	G	G	G	G	G	T
Snp11	T	T	T	T	T	T	C

	M	N
PC1	X	x

Ancestry Analysis

	1	2	3	4	5	6	7
Snp4	C	C	C	T	T	T	T
Snp6	G	G	G	A	A	A	A
Snp8	T	T	T	G	G	G	G

	1-3	4-7
Snp4	C	T
Snp6	G	A
Snp8	T	G

	1-3
Snp4	C
Snp6	G
Snp8	T

=Y

	4-7
Snp4	T
Snp6	A
Snp8	G

=y

	1-3	4-7
PC2	Y	y

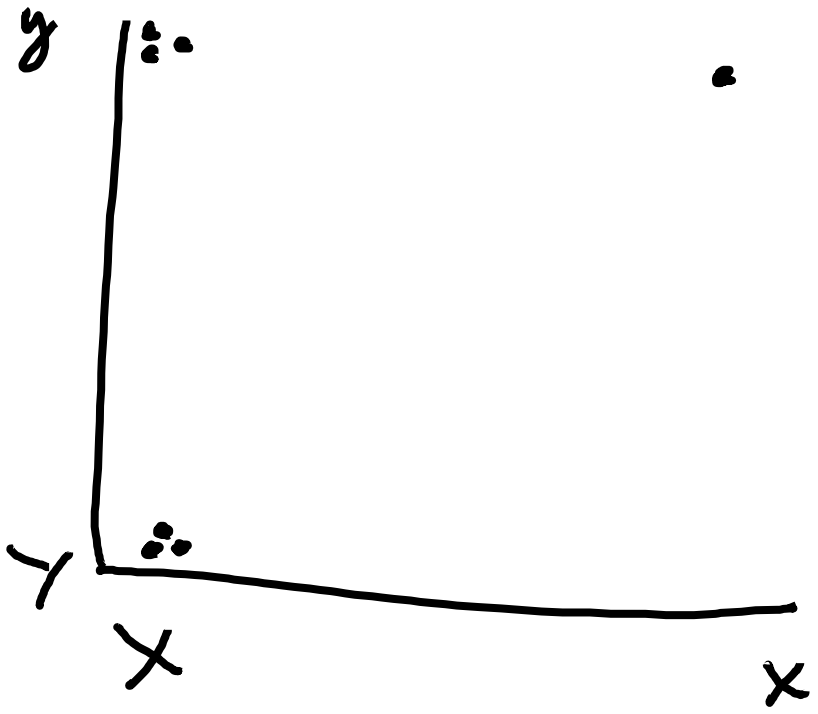
Ancestry Analysis

	1	2	3	4	5	6	7
PC1	**X**	**X**	**X**	**X**	**X**	**X**	**x**
PC2	Y	Y	Y	y	y	y	y
Snp2	G	G	G	G	G	G	G
Snp10	A	G	C	T	A	G	C
Snp12	G	C	T	A	A	G	C

	1-3	4-6	7
PC1	**X**	**X**	**x**
PC2	Y	y	y
Snp2			
Snp10			
Snp12			

PC1 and PC2 inform about ancestry

	1-3	4-6	7
PC1	**X**	**X**	**x**
PC2	Y	y	y
Snp2	G	G	G
Snp10	A	T	C
Snp12	G	A	C



Complex traits: height heritability is 80%

nature.com/naturegenetics

Sizing up human height variation

Peter M Visscher

Genome-wide association studies have identified many variants affecting susceptibility to disease. Now, three studies use this approach to study adult height variation in a combined sample size of ~63,000 individuals and report a total of 54 validated variants influencing this trait.

NATURE GENETICS | VOLUME 40 | NUMBER 5 | MAY 2008

p://www.nature.com/naturegenetics

Genome-wide association analysis identifies 20 loci that influence adult height

Michael N Weedon[1,2,23], Hana Lango[1,2,23], Cecilia M Lindgren[3,4], Chris Wallace[5], David M Evans[6], Massimo Mangino[7], Rachel M Freathy[1,2], John R B Perry[1,2], Suzanne Stevens[7], Alistair S Hall[8], Nilesh J Samani[7], Beverly Shields[2], Inga Prokopenko[3,4], Martin Farrall[9], Anna Dominiczak[10], Diabetes Genetics Initiative[21], The Wellcome Trust Case Control Consortium[21], Toby Johnson[11–13], Sven Bergmann[11,12], Jacques S Beckmann[11,14], Peter Vollenweider[15], Dawn M Waterworth[16], Vincent Mooser[16], Colin N A Palmer[17], Andrew D Morris[18], Willem H Ouwehand[19,20], Cambridge GEM Consortium[22], Mark Caulfield[5],

Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index

aturegenetics

Many sequence variants affecting diversity of adult human height

63K people
54 loci
~5% variance explained.

NATURE GENETICS VOLUME 40 [NUMBER 5 [MAY 2008
Nature Genetics VOLUME 42 | NUMBER 11 | NOVEMBER 2010

Calculating RISK for complex traits

- Start with your population prior for T2D: for CEU men, we use 0.237 (corresponding to LR of 0.237 / (1 – 0.237) = 0.311).

- Then, each variant has a likelihood ratio which we adjust the odds by.

dbSNP	Genotype	Imputed from	R squared	Likelihood Ratio	Running Total LR	Running Total Probability
Prior				0.311	0.311	23.700%
2283228	AA			1.031	0.320	24.252%
2237897	CC			1.023	0.328	24.675%
7903146	CC			0.753	0.247	19.789%
985694	CC			0.918	0.226	18.463%
726281	AA			0.909	0.206	17.071%
3020317	TT			0.928	0.191	16.032%

Slide by Rob Tirrell, 2010

LETTER

doi:10.1038/nature09410

Hundreds of variants clustered in genomic loci and biological pathways affect human height

A full list of authors and their affiliations appears at the end of the paper.

183K people
180 loci
~10% variance explained

832 | NATURE | VOL 467 | 14 OCTOBER 2010

Missing Heritability

Table 1 | Estimates of heritability and number of loci for several complex traits

Disease	Number of loci	Proportion of heritability explained	Heritability measure
Age-related macular degeneration[72]	5	50%	Sibling recurrence risk
Crohn's disease[21]	32	20%	Genetic risk (liability)
Systemic lupus erythematosus[73]	6	15%	Sibling recurrence risk
Type 2 diabetes[74]	18	6%	Sibling recurrence risk
HDL cholesterol[75]	7	5.2%	Residual* phenotypic variance
Height[15]	40	5%	Phenotypic variance
Early onset myocardial infarction[76]	9	2.8%	Phenotypic variance
Fasting glucose[77]	4	1.5%	Phenotypic variance

* Residual is after adjustment for age, gender, diabetes.

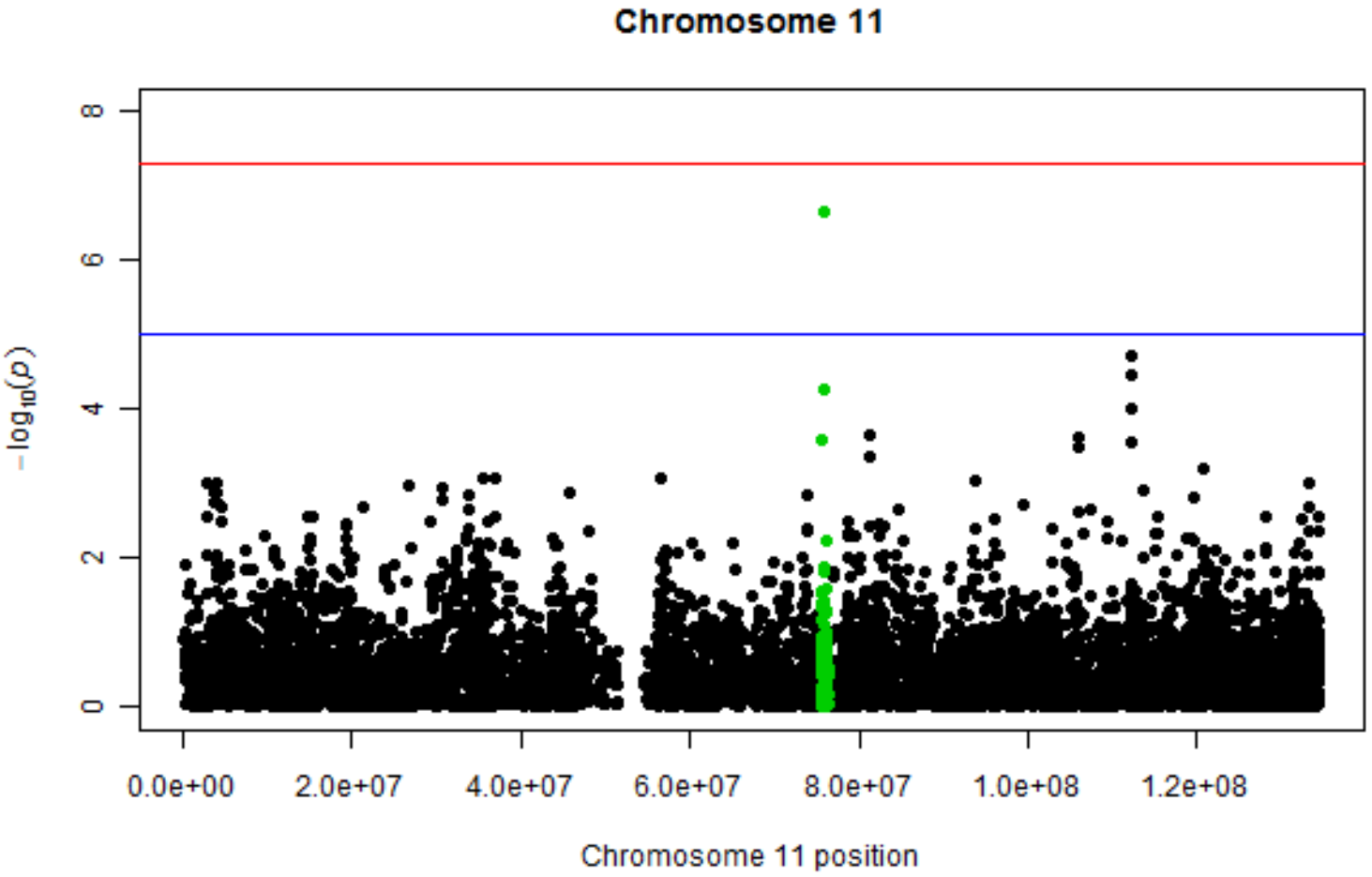
Where is the missing heritability?

Lots of minor loci
Rare alleles in a small number of loci
Gene-gene interactions
Gene-environment interactions

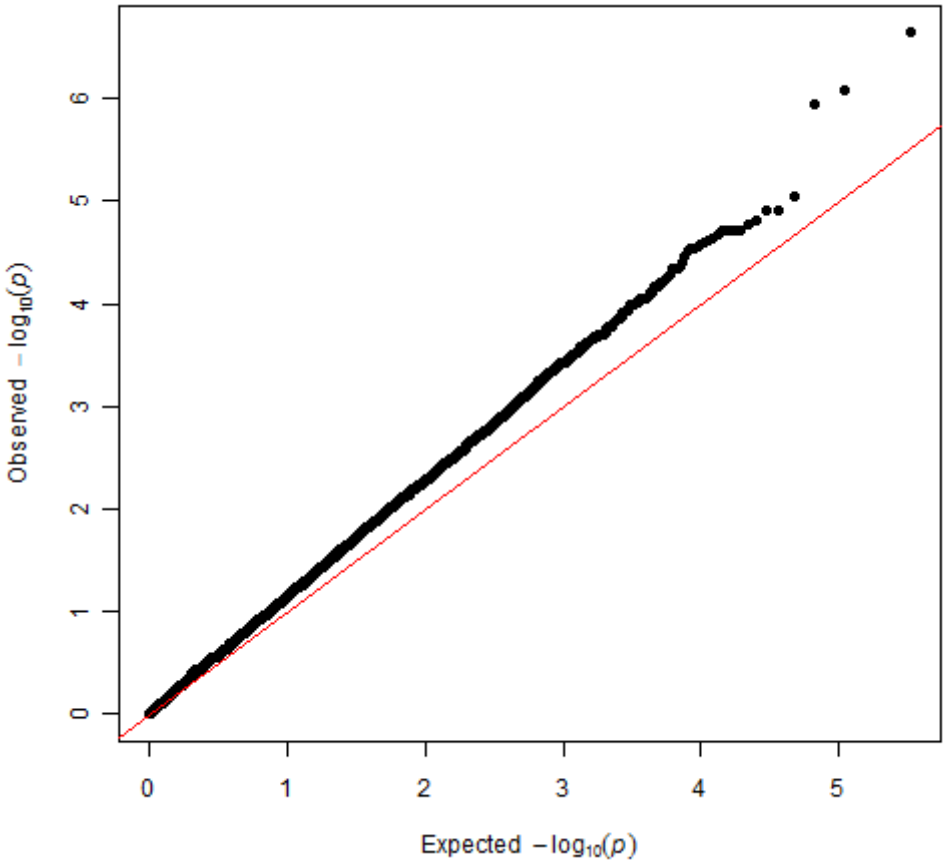
Common SNPs explain a large proportion of the heritability for human height

Jian Yang[1], Beben Benyamin[1], Brian P McEvoy[1], Scott Gordon[1], Anjali K Henders[1], Dale R Nyholt[1], Pamela A Madden[2], Andrew C Heath[2], Nicholas G Martin[1], Grant W Montgomery[1], Michael E Goddard[3] & Peter M Visscher[1]

Nature Genetics VOLUME 42 | NUMBER 7 | JULY 2010


Chromosome 11
$-\log_{10}(p)$
0
2
4
6
8
0.0e+00
2.0e+07
4.0e+07
6.0e+07
8.0e+07
1.0e+08
1.2e+08
Chromosome 11 position

Q-Q plot for human height



This approach explains 45% variance in height.

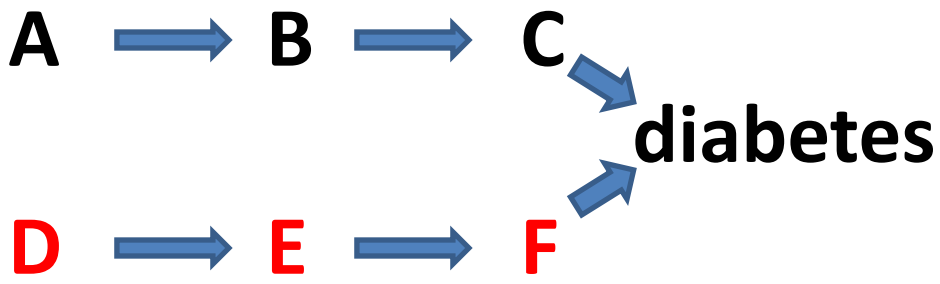
Rare alleles

Cases

Controls

1. You wont see the rare alleles unless you sequence
2. Each allele appears once, so need to aggregate alleles in the same gene in order to do statistics.

Gene-Gene



A^- not affected
D^- not affected

A^- D^- affected
A^- E^- affected
A^- F^- affected

A^- B^- not affected
D^- E^- not affected

Gene-environment

1. Height gene that requires eating meat
2. Lactase gene that requires drinking milk

These are SNPs that have effects only under certain environmental conditions