



Summer Institute  
in Statistical Genetics 2019

# Genetics and Genomics

## 1. Genes and Inheritance



[ggibson.gt@gmail.com](mailto:ggibson.gt@gmail.com)  
<http://www.cig.gatech.edu>



### Course Outline

1. Mon AM Genes and Inheritance (GG)
2. Mon PM Molecular Biology of the Genome (CQ)
3. Tue AM Association Studies and Gene Expression (GG)
4. Tue PM Epigenetics and Genome Editing (CQ)
5. Wed AM Evolutionary Genetics (GG) / Functional Genomics (CQ)

## Genotype and Phenotype

The **genotype** of an organism is the sequence of its genes.

The **phenotype** of an organism the way it appears.

In general, genes are not deterministic. Genotypic variation among organisms specifies the information that, in combination with the environment, influences the phenotype.

**Pleiotropy** refers to the ability of single genes to influence multiple phenotypes.

**Penetrance** is the proportion of individuals with a genotype who have the phenotype / disease.

**Expressivity** is the degree / severity of the phenotype in affected individuals.

## Mendelian Genetics

$$\frac{A}{A} \times \frac{a}{a}$$

F0: Pure breeding parents



$$\frac{A}{a} \times \frac{A}{a}$$

F1: Heterozygous offspring



$$\frac{A}{A} ; \frac{A}{a} ; \frac{a}{A} ; \frac{a}{a}$$

F2: Mendelian proportions of Homozygotes + Heterozygotes

### 3 Models of Complex Disease

#### CDCV: Common Disease / Common Variant

The proposition that most disease susceptibility can be attributed to 10 to 20 loci, each of which explain around 5% of disease risk.

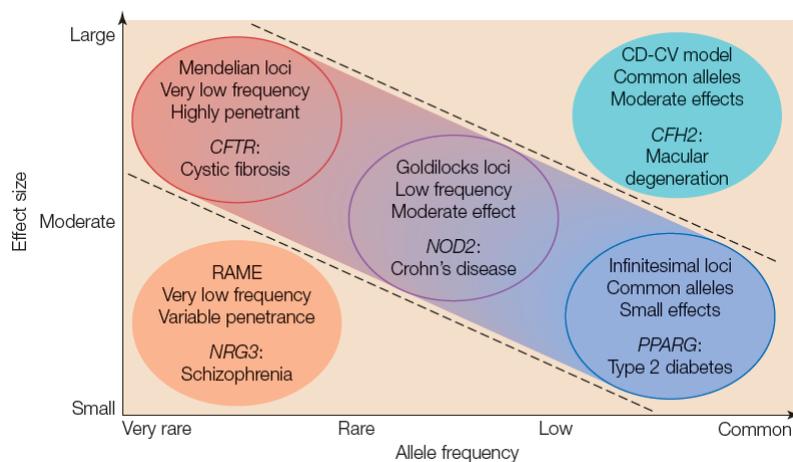
#### RAME: Rare alleles of Major Effect

The proposition that diseases are highly heterogeneous, with hundreds or thousands of rare mutations causing individual cases of disease.

#### Infinitesimal:

The proposition that we all carry thousands of very weak susceptibility alleles, and those unlucky enough to have too many are at highest risk, where rare variants or environmental triggers push us over the edge.

### Models of the Genetics of Complex Traits



Manolio et al (2009) *Nature* 461: 747-753

## Heritability

Heritability is the proportion of variance in a population that can be attributed to genotypic differences

$$h^2 = V_G/V_P \text{ where } V_P = V_G + V_E$$

The phenotypes may be **discrete**, such as disease status; **categorical**, such as number of digits; or **continuous**, such as height or a biochemical measure.

1. Heritability is not a statement about individuals.

A heritability of 50% for diabetes does not imply that half the reason why someone is diabetic is genetic, the other half environmental. Rather, it suggests that there would be half as much diabetes in the population if everyone was genetically identical.

2. Heritability is only a statement about a single population.

A heritability of 80% for height does not imply that most of the average difference in height between populations is due to genetic differences. Heritability estimates alone should not be used to draw inferences about genetic divergence between groups.

3. Heritability is not the same as inheritance.

Inheritance is the correspondence between children and their biological parents. It can be due to environmental, including cultural, factors that are shared by family members, or to effects. The only way to confidently interpret heritability is to actually measure the genotypic contribution.

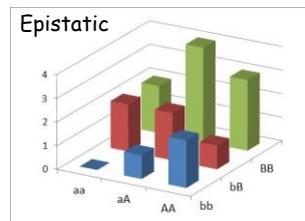
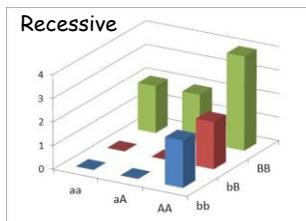
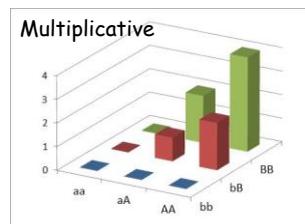
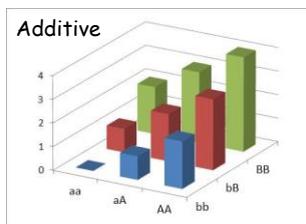
4. Very low heritability does not imply very little genetic contribution.

It may either be due to relatively high environmental variance (hence, a large denominator  $V_P$ ), or to an absence of variance in the genes that contribute. Many important genes, including drug targets, are not polymorphic and will only be discovered through other types of approach including model organism research.

## Broad Sense Heritability

$$V_G = V_A + V_D + V_I + V_{G \times E}$$

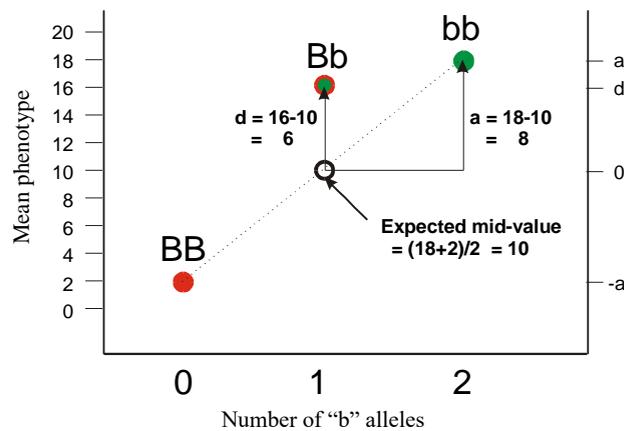
**Narrow sense** heritability is only the additive component whereas **Broad sense** heritability includes dominance, interaction and genotype-by-environment effects.



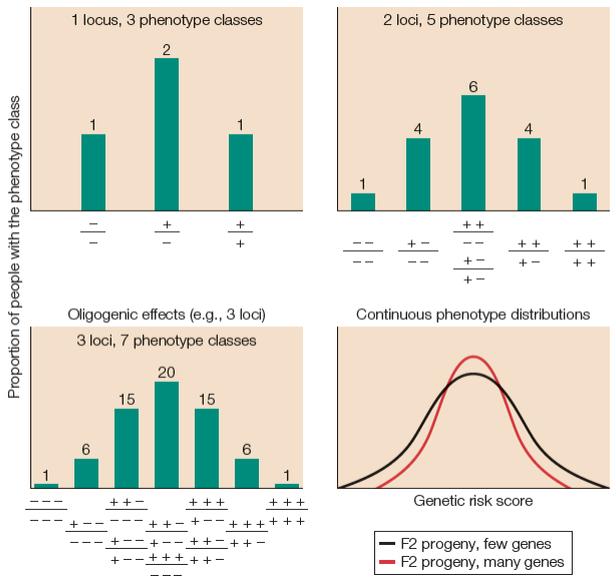
$$V_P = V_A + V_D + V_I + V_{G \times E} + V_E$$

- Loci are said to have *Additive* effects if the contributions of each individual allele can simply be added algebraically to arrive at a prediction of a phenotype given a genotype.
- *Dominance* refers to the observation that heterozygotes resemble one class of homozygotes more than the other.
- *Epistasis* refers to a locus-by-locus interaction, such as when alleles at two loci antagonize or synergize with one another.
- $V_E$  is the *environmental* variance

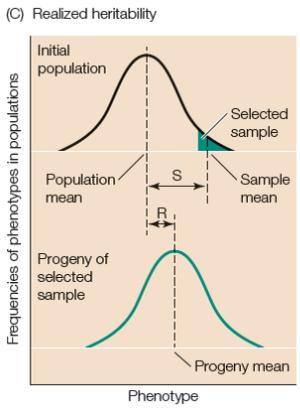
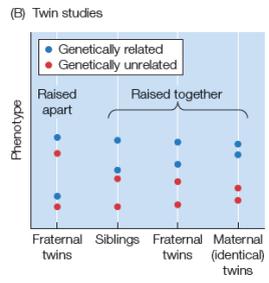
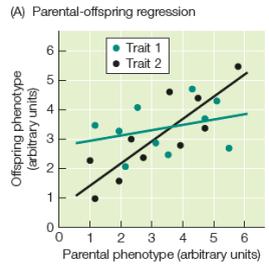
### Dominance ratio



### From Mendelian to Quantitative genetics



### Estimating Heritability



### Twin Studies

Identical / Maternal



Tiki and Ronde Barber

$$r_{mz} = A + C$$

Dizygotic / Fraternal

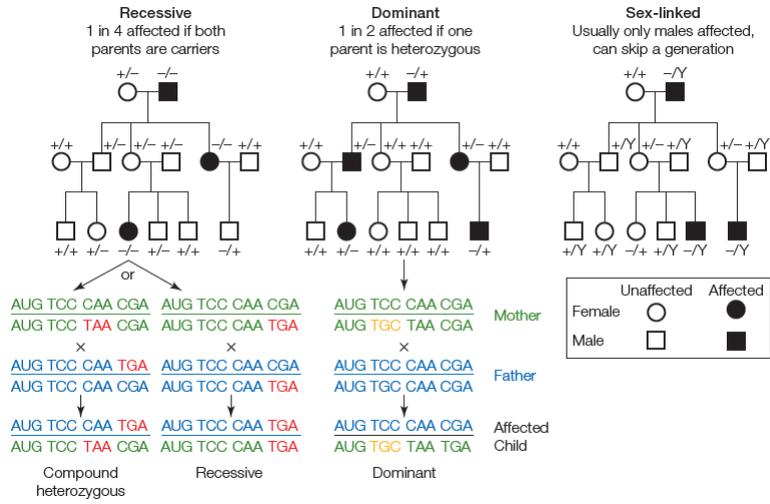


Jenna and Barbara Bush

$$r_{dz} = \frac{1}{2}A + C$$

A = Additive Genetic component; C = Common Environment (smaller if reared apart)  
 E = unique environment =  $1 - r_{mz}$   
 $r_{dz}$  should be greater than  $r_{sib}$  since C is larger where the womb/upbringing is shared

### Mendelian Pedigree Studies



## Monogenic Disorders

Approximately 1 in 3,700 Americans have Cystic Fibrosis

Assuming  $p^2 = 0.00027$ , then  $p = 0.016$ , the mutant allele frequency

That is, 1 in 30 people are carriers (which is 120 times as many people as have CF),

that is, **3% of Caucasians are carriers**, and less than 0.03% sufferers.

It is very likely that someone in this class is a carrier of a CF mutation

A CF carrier has a 1 in 30 chance of marrying another carrier by chance, and

1 in 4 of their children will be expected to have CF, and only half of all 2-child families will have an affected child

*There are hundreds of similar conditions (rare recessives with  $p \sim 0.01$ ), so we are all carriers for multiple Mendelian disease genes. Collectively, as many as 1 in 25 couples should expect to be dual carriers for a recessive Mendelian disorder, corresponding to an approximate 1% affected rate in all children*

Around 1 in 400 children have an inherited Inborn Error of Metabolism, namely an enzyme deficiency affecting amino acid, lipid or other organic molecule biosynthesis. For example:

- Phenylketonuria	1/15,000	mental retardation syndrome
- Galactosemia	1/40,000	liver dysfunction and cataracts
- Gaucher's Disease	1/60,000	facial dysmorphism, liver disease
- Zellweger Syndrome	1/50,000	seizures, low muscle tone
- Lesch-Nyhan Syndrome	1/380,000	self-inflicted injury, gout / kidney disease

## Online Mendelian Inheritance in Man (OMIM)



**OMIM®** Online Mendelian Inheritance in Man®  
An Online Catalog of Human Genes and Genetic Disorders  
Updated 24 June 2016

Search OMIM...

Advanced Search: OMIM, Clinical Synopses, Gene Map  
Need help? Example Searches, OMIM Search Help, OMIM Tutorial  
Mirror site: mirror.omim.org

OMIM is supported by a grant from NHGRI, licensing fees, and generous contributions from people like you.

Home About Statistics Downloads Help External Links Terms of Use Contact Us MIMmatch Disease

Advanced Search Search

Table of Contents for #300623

**#300623** FRAGILE X TREMOR/ATAXIA SYNDROME; FXTAS

Phenotype-Genes Relationships

Phenotype	Inheritance	Phenotype	Gene/Locus	Gene/Locus	MIM number	OMIM number
Xq27.3	Fragile X trinucleotide expansion	300623	NCD	3	FMR1	309550

Clinical Synopsis

**TEXT**  
A number sign (#) is used with this entry because fragile X tremor/ataxia syndrome (FXTAS) is caused by an expanded trinucleotide repeat in the FMR1 gene (309550.0004).

In FXTAS, the expanded repeats range in size from 55 to 300 repeats and are referred to as "premutations"; full repeat expansions with greater than 300 repeats results in fragile X mental retardation syndrome (309243) (Jacquemont et al., 2003). [6]

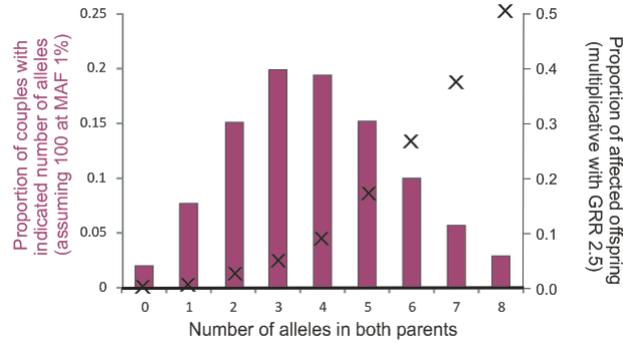
**Description**  
Jacquemont et al. (2007) provided a review of fragile X syndrome, which they characterized as a neurodevelopmental disorder, and FXTAS, which they characterized as a neurodegenerative disorder. Amici et al. (2008) provided a review of FXTAS and noted that the pathogenesis of the disorder is distinct from that in fragile X syndrome. FXTAS results from a toxic gain of function of FMR1 RNA, whereas fragile X syndrome results from a loss of FMR1 function. [6]

The penetrance of FXTAS in male carriers aged 50 years and over, ascertained through families with a fragile X syndrome proband, is at least 33% (Hagerman and Hagerman, 2004); its penetrance in female carriers is approximately 5-10% (Gooco et al., 2008). [6]

**Clinical Features**  
Hagerman et al. (2001) reported 5 men with a fragile X premutation, ranging from 78 to 98 repeats, who presented in the sixth decade with

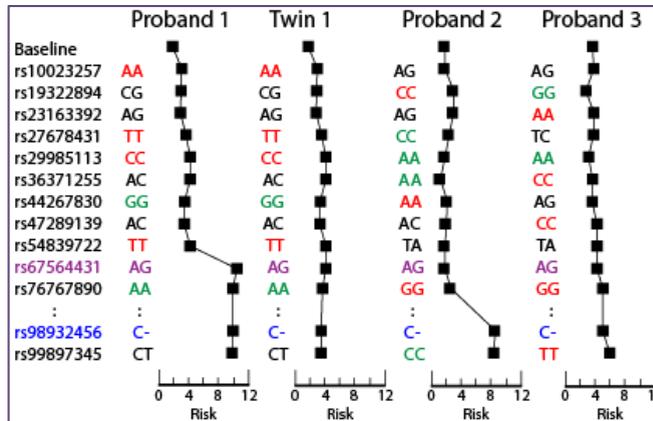
External Links  
Protein  
Clinical Resources  
Animal Models

### Multiplicative Rare Alleles of Major Effect

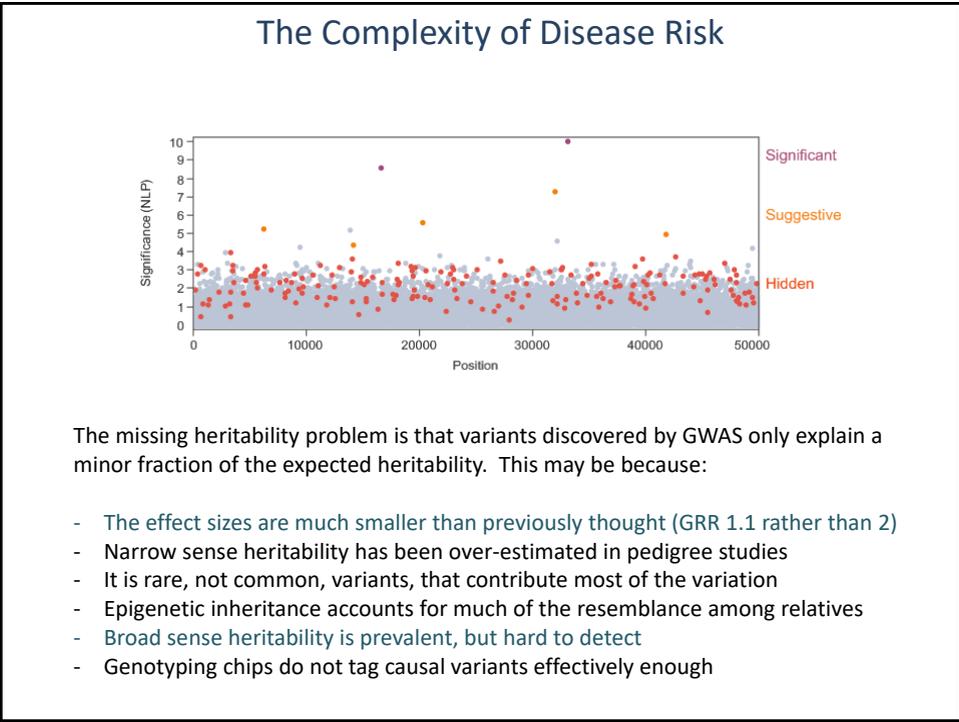
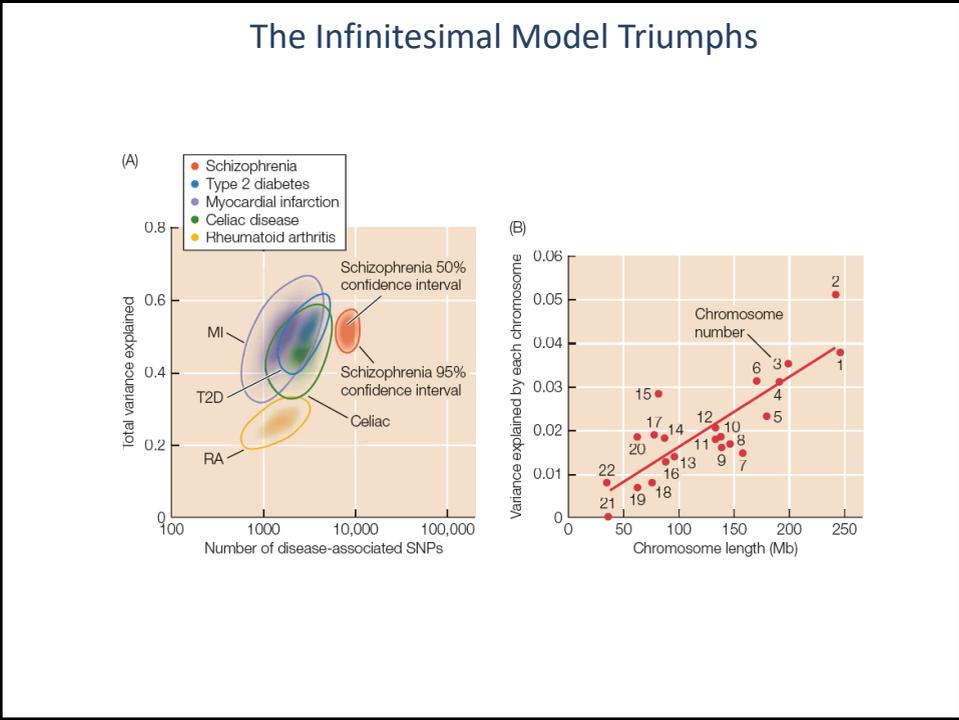


Assume there are 100 mutations at 1% frequency, each of which increases the risk of disease 2.5-fold over a baseline environmental risk of 1%. Whence 0 alleles have a risk of 1%, 1 of 2.5%, 2 of 6%, 3 of 15%, 4 of 39%, 5 or more is highly penetrant.

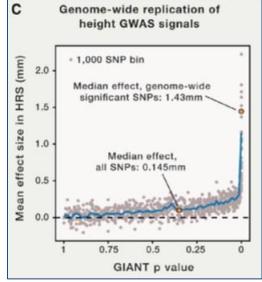
### Individualization of Polygenic Risk



Individual deviation may occur if a SNP has an outsized effect in a person, or they live in a traumatic environment.

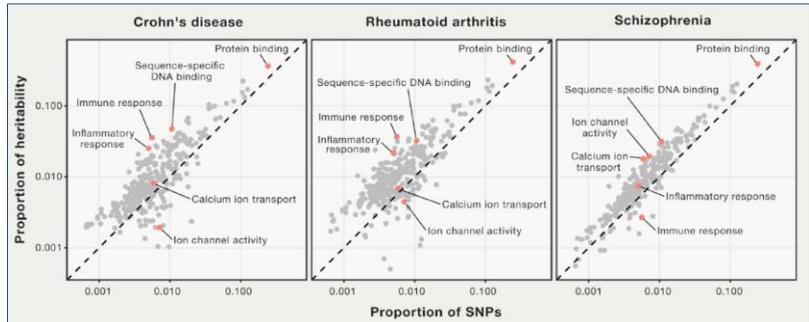


### Omnigenetics



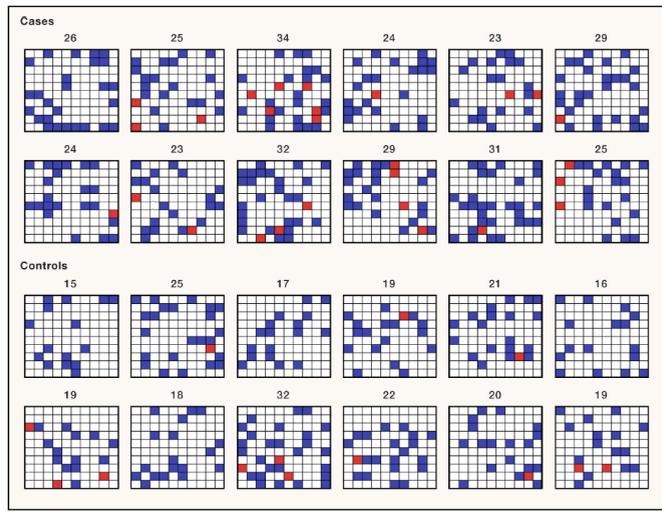
← Estimate that ~4% of all common SNPs influence height

Many core biochemical processes associate to different degrees with various traits



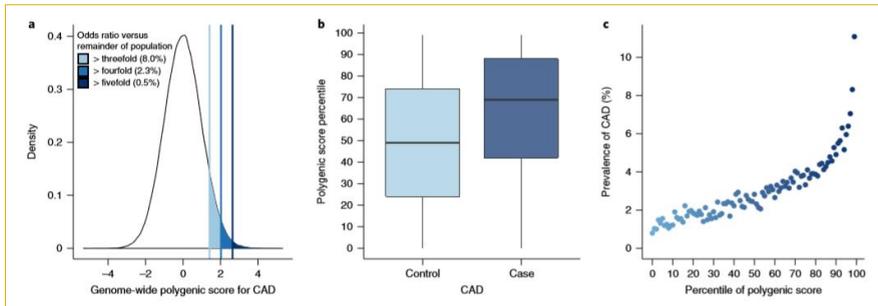
Boyle, Li and Pritchard (2017) Cell 169: 1177-1186

### Polygenic Heterogeneity and Personalized Medicine



Wray et al (2018) Cell 173: 1573-1580

### Polygenic Risk Scores



Khera et al (2018) *Nature Genetics* 50: 1219–1224

More people are at elevated risk due to common variants than rare mutations

Some people believe that PRS are the future of personalized medicine

I think they are also useful for negative prediction, namely finding those at low risk