QUANTITATIVE TRAIT LOCI: HERITABILITY, LINKAGE,AND ASSOCIATION

J. ERIC SCHMITT MD, PhD

DEPARTMENT OF RADIOLOGY

UNIVERSITY OF PENNSYLVANIA

eric.schmitt@stanfordalumni.org

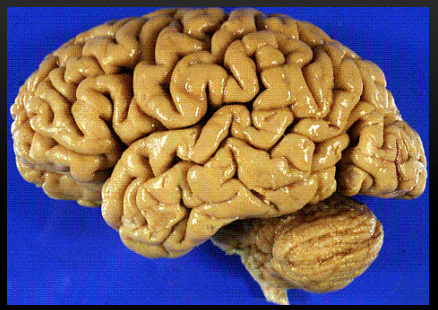
GOALS AND OBJECTIVES

- DEFINE CORE QUANTITATIVE GENETIC CONCEPTS
  - Heritability, linkage, association
  - Acronym parade: QTL, IBD, EEA, SEM, LOD
- Give an overview of Structural Equation Modeling (SEM) as a tool for genetic analysis.
- Provide examples of classic twin and family modeling and several extensions of basic models.
- Demonstrate flexibility of likelihood-based statistical modeling
- Distinguish between linkage and association studies and discuss ways to combine them
- Provide a transition from basic genetics to later lectures

BASIC STATISTICS

The Human Genome

- 22 +2 chromosomes
- 3,200,000,000 (3 x 10^9) base pairs
- 1/100 – 1/1500 bases exhibit polymorphisms (SNPs)
- 5-23 genes/1,000,000 bp
- 20,000-150,000 genes in the genome
- Relatively simple function (retain molecular information)



The Human Brain

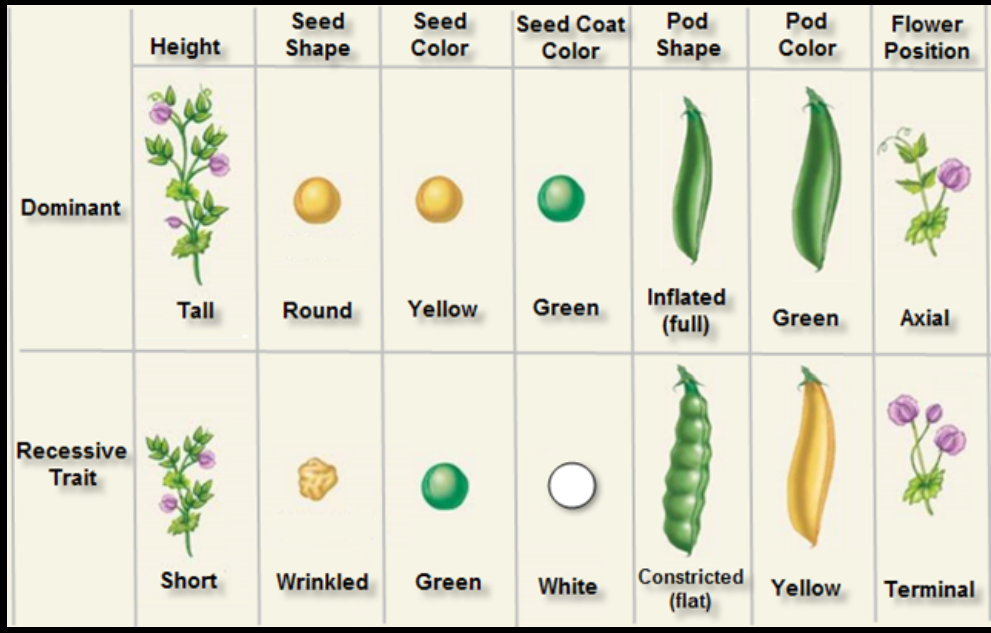
Function →

- 1,000,000,000,000 (1 x 10^{12}) cells
  - 100,000,000,000 (1 x 10^{11}) neurons
- 1,000-10,000: average number of connections per neuron in brain, ~30,000 for cortex
- Each neuron with a specific 3D spatial location in and role in the neural network
- The most complex structure known & extraordinarily complex function
- ~50% of the genome expressed in the brain

HOW DOES THIS…..MAKE THIS (among other things)???

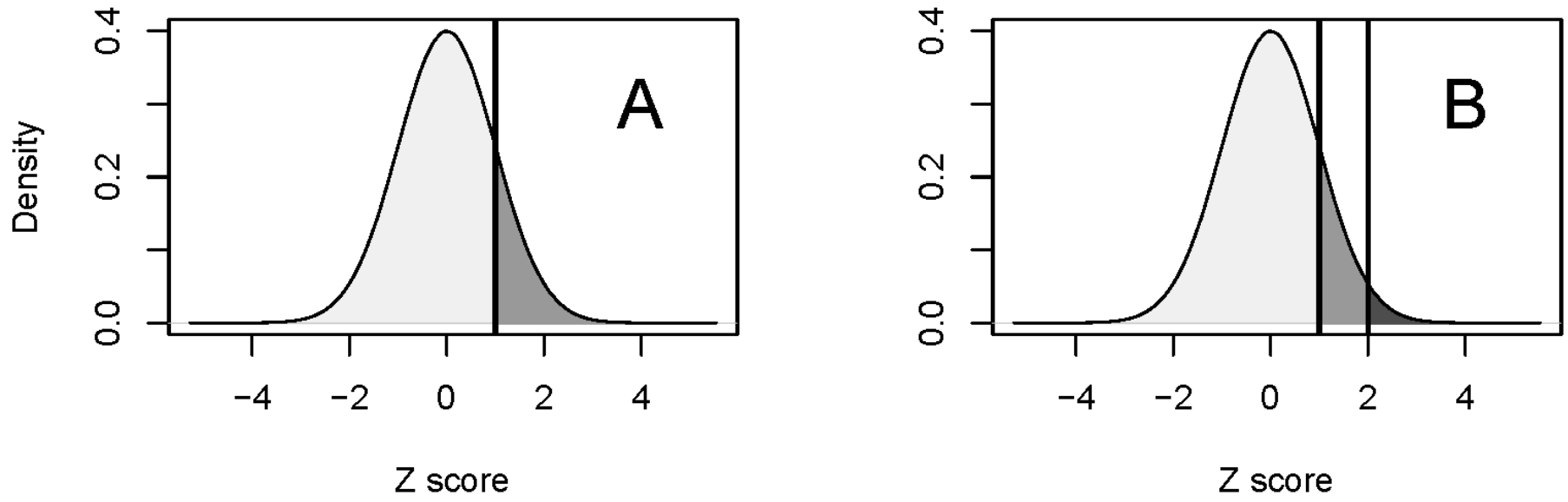
CLASSICAL GENETICS

- Classical Mendelian genetics deals with qualitative, usually dichotomous traits
- Many interesting phenotypes are not continuous (e.g brain volumes, BOLD signal, anisotropy, EEG patterns).
- Many complex poorly understood traits (i.e. diseases such as schizophrenia, Alzheimer's disease) are treated as dichotomous but are likely multifactorial.

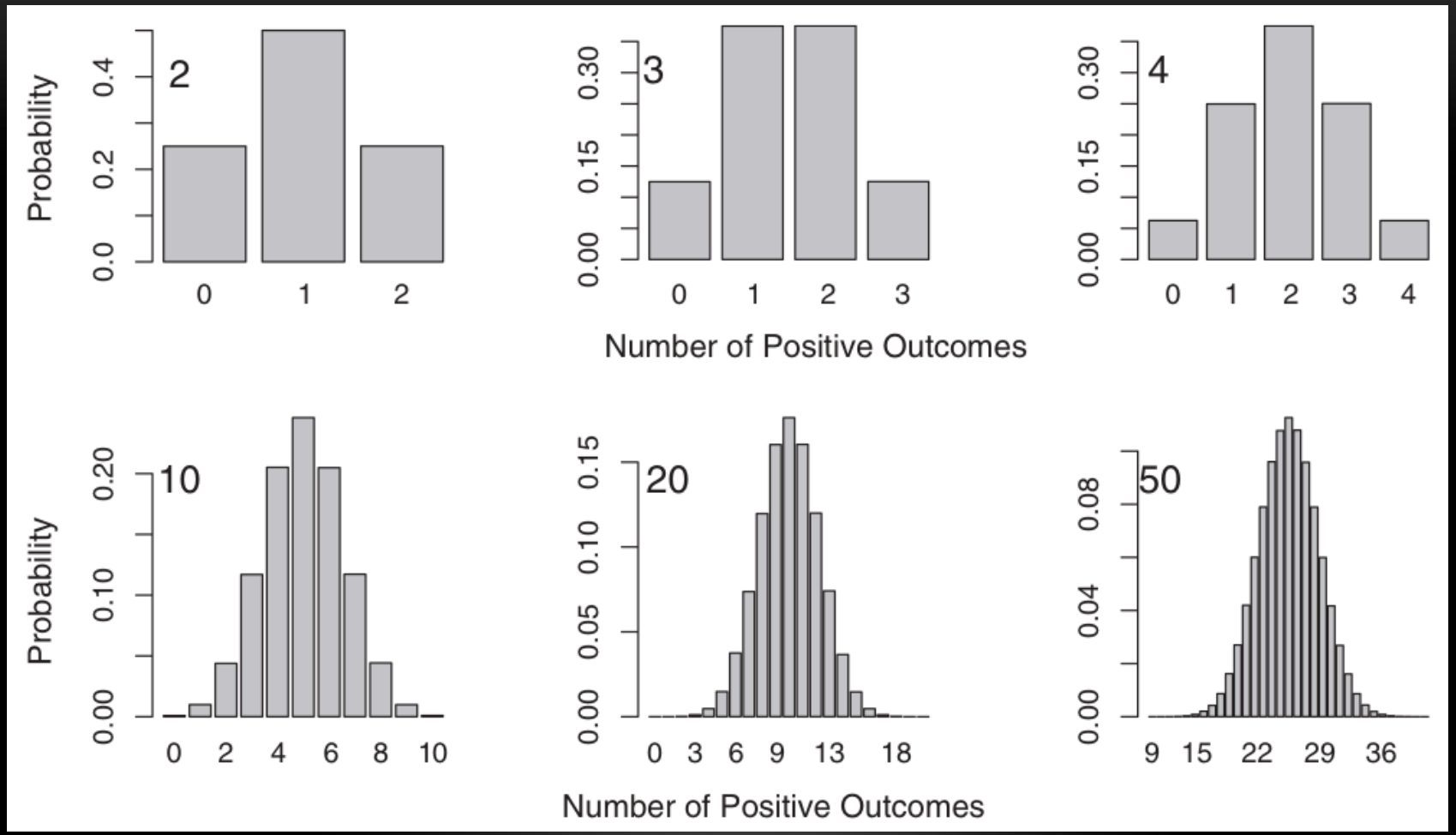


http://www.takdangaralin.com/science/genetics-the-beginning-of-genetics/

LIABILITY THRESHOLD MODEL FOR COMPLEX TRAITS:
COMPLEX DISEASES LIKELY DUE TO MULTIPLE ADDITIVE EFFECTS



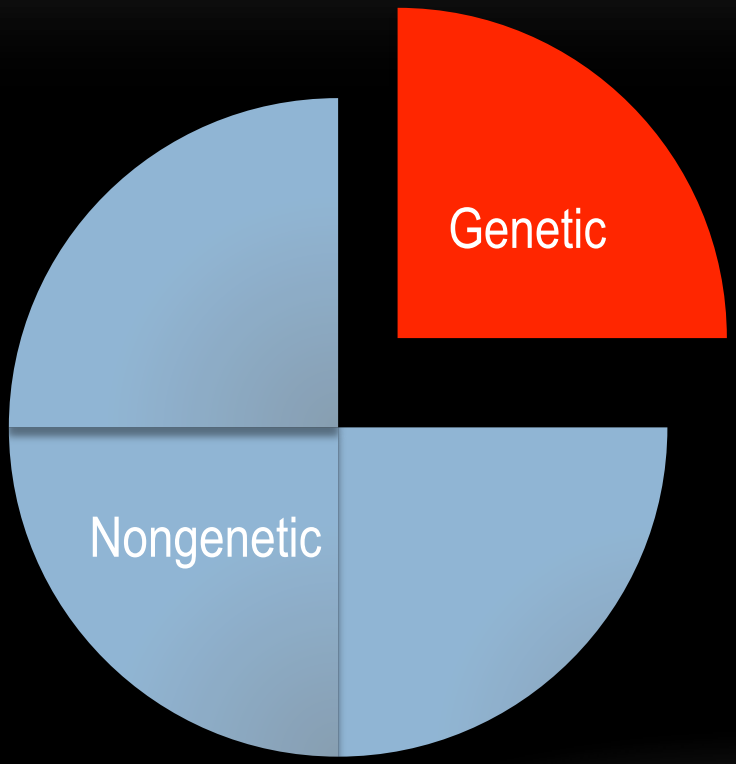
MULTIPLE BINARY FACTORS APPROXIMATE A NORMAL DISTRIBUTION



QUANTITATIVE GENETICS

- Branch of genetics “concerned with inheritance of those differences between individuals that are of degree rather than kind.”
  - D.S. Falconer, *Quantitative Genetics*
- Origins of the field a natural byproduct of discoveries in evolutionary and Mendelian genetics
- Based on genetic effects that are small at individual genetic loci, but large when added together
- Historically intertwined with the development of statistics, e.g. Fisher, Pearson
- Modern ties with molecular genetics and bioinformatics, as well as psychology and psychiatry.
- Non-quantitative analyses also can be performed, but usually less powered if quantitative data are available.
- Acknowledges that understanding the phenotype is as important as understanding genotype

VARIANCE DECOMPOSITION



$$Var(x) = Var_{Genetic} + Var_{Nongenetic}$$

Heritability (a^2)

- DEFINED AS THE PROPORTION OF TOTAL PHENOTYPIC VARIATION EXPLAINED GENETIC FACTORS
  - Ranges from 0-1.
  - BROAD SENSE: Includes all genetic variance
  - NARROW SENSE: Additive genetic variance only—distinction important for evolutionary genetics.
  - Population specific
  - Not only influenced by genetic variation: don't forget the denominator.
    - Environmental effects
    - Measurement Error
- Note that heritability does not account for genes that affect the phenotypic mean for the population if the genetic variance is zero; these genes are undetectable via traditional quantitative genetic analyses.
- Identifies phenotypes with significant genetic contributions
  - Targets for linkage/association
  - Shared genetic factors between traits, data reduction.
- Other measures of genetic variance exist
  - Raw genetic variance when continuous phenotypes are used.
  - Based on coefficient of variation (heritability standardized by the mean)

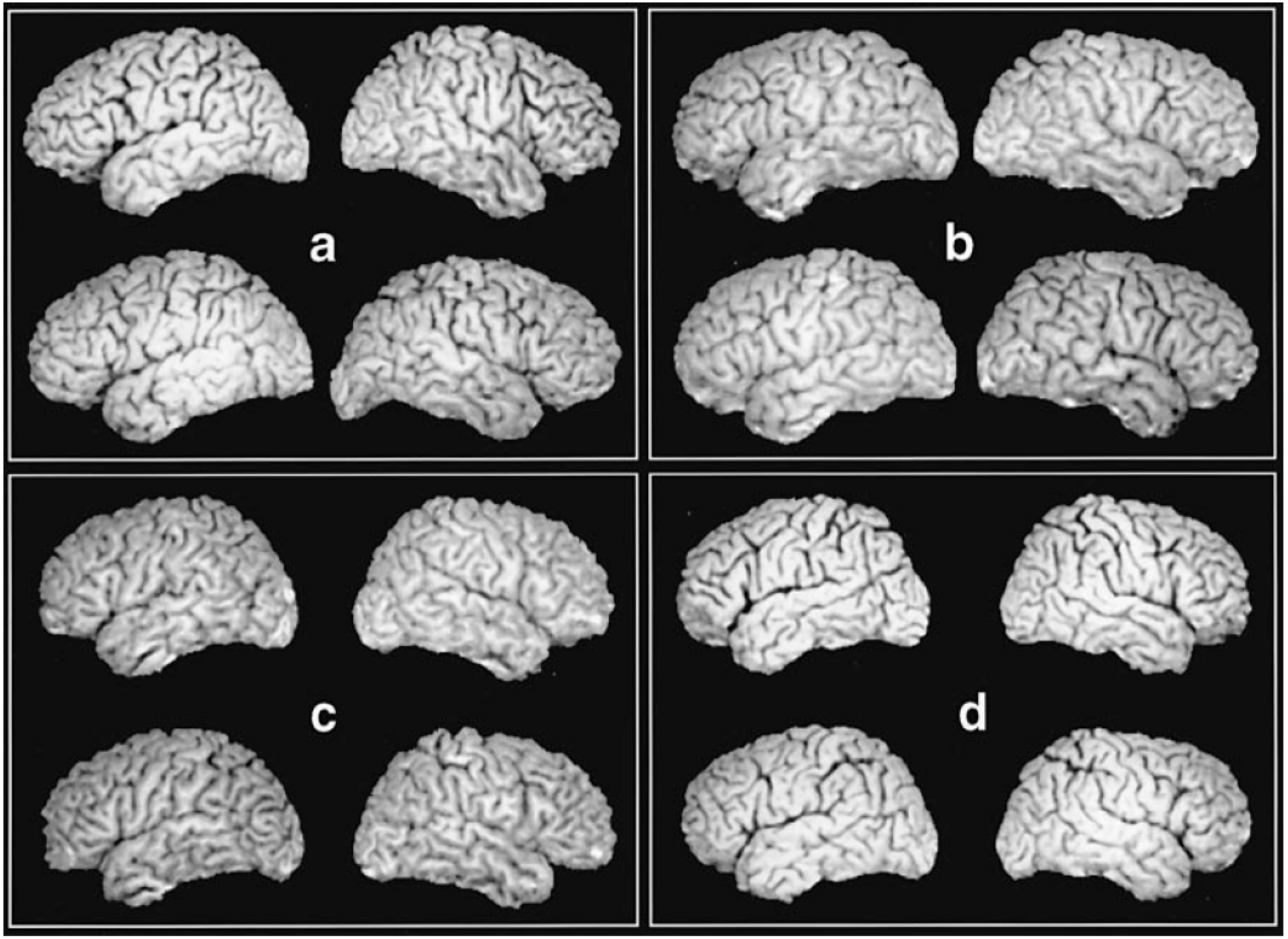
STRATEGIES FOR VARIANCE DECOMPOSITION IN HUMAN POPULATIONS

- Classical Twin Studies
- Adoption Studies
- Twins Reared Apart
- Family Studies

- All approaches use some statistical triangulation based on simple assumptions.

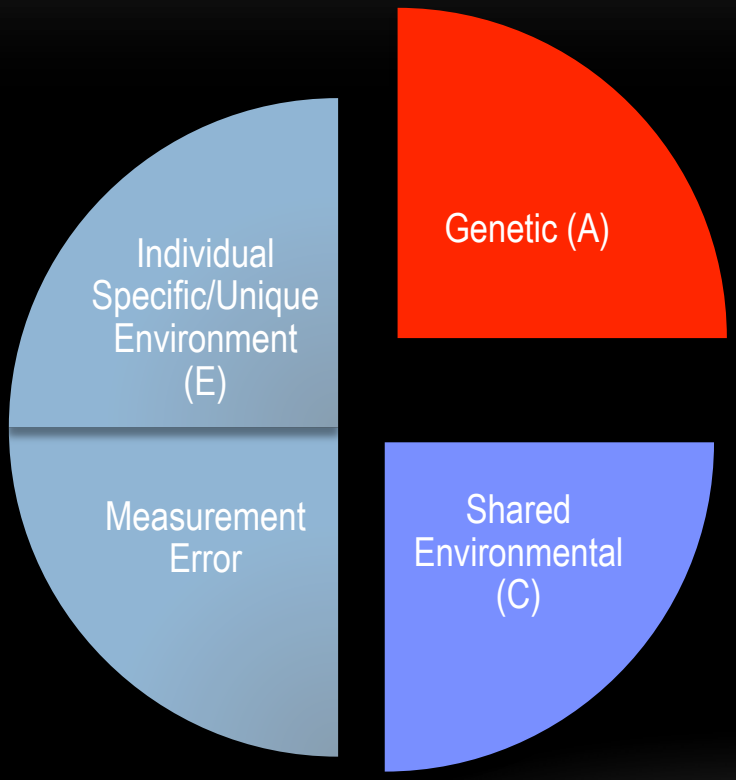
CLASSIC TWIN STUDIES

- Estimate heritability based on differences in zygosity between monozygotic (MZ) and dizygotic (DZ) twins.
- While MZ twins are genetically identical, DZ twins, on average, share ½ of their genes identical by descent. Based on this fact, mathematical formulae can be derived to calculate heritability based on differences in covariance/correlation between family members.
- Basic twin models assume that the environment is on average the same in MZ and DZ families *with respect to the phenotype of interest*. This is called the Equal Environment Assumption (EEA) and is a potential pitfall of twin research.
  - Example: similarities in choice of clothing very different between MZ and DZ twin pairs.
  - The EEA has been shown to hold for most other exophenotypes that have been tested.



Bartley et al., 1997, *Brain*

VARIANCE DECOMPOSITION: FALCONER ESTIMATION



$$Var(x) = Var_{Genetic} + Var_{Shared} + Var_{Unique}$$

$$V_p = a^2 + c^2 + e^2$$

$$Cor_{MZ} = a^2 + c^2$$

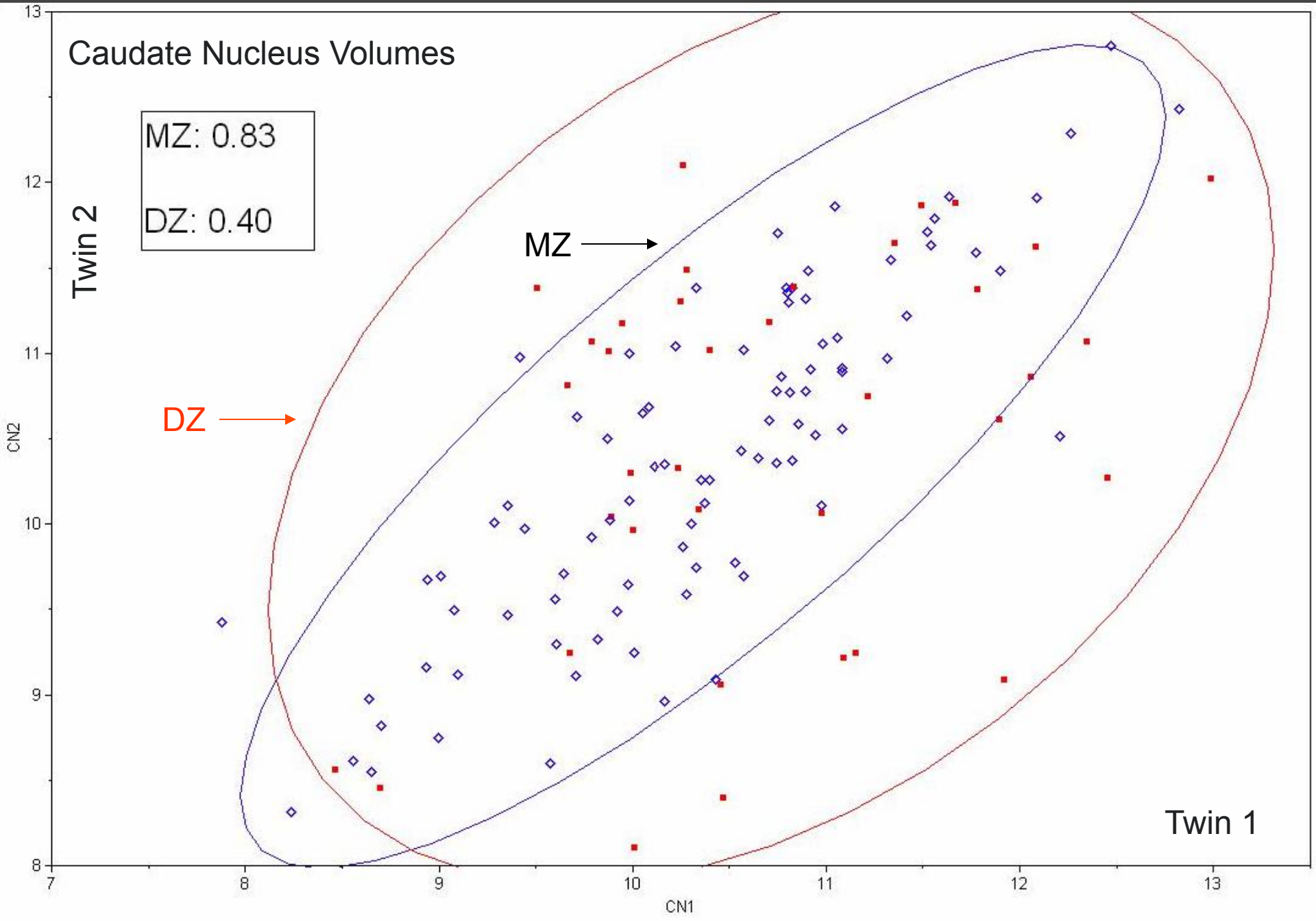
$$Cor_{DZ} = \frac{1}{2}a^2 + c^2$$

$$a^2 = 2(Cor_{MZ} - Cor_{DZ})$$

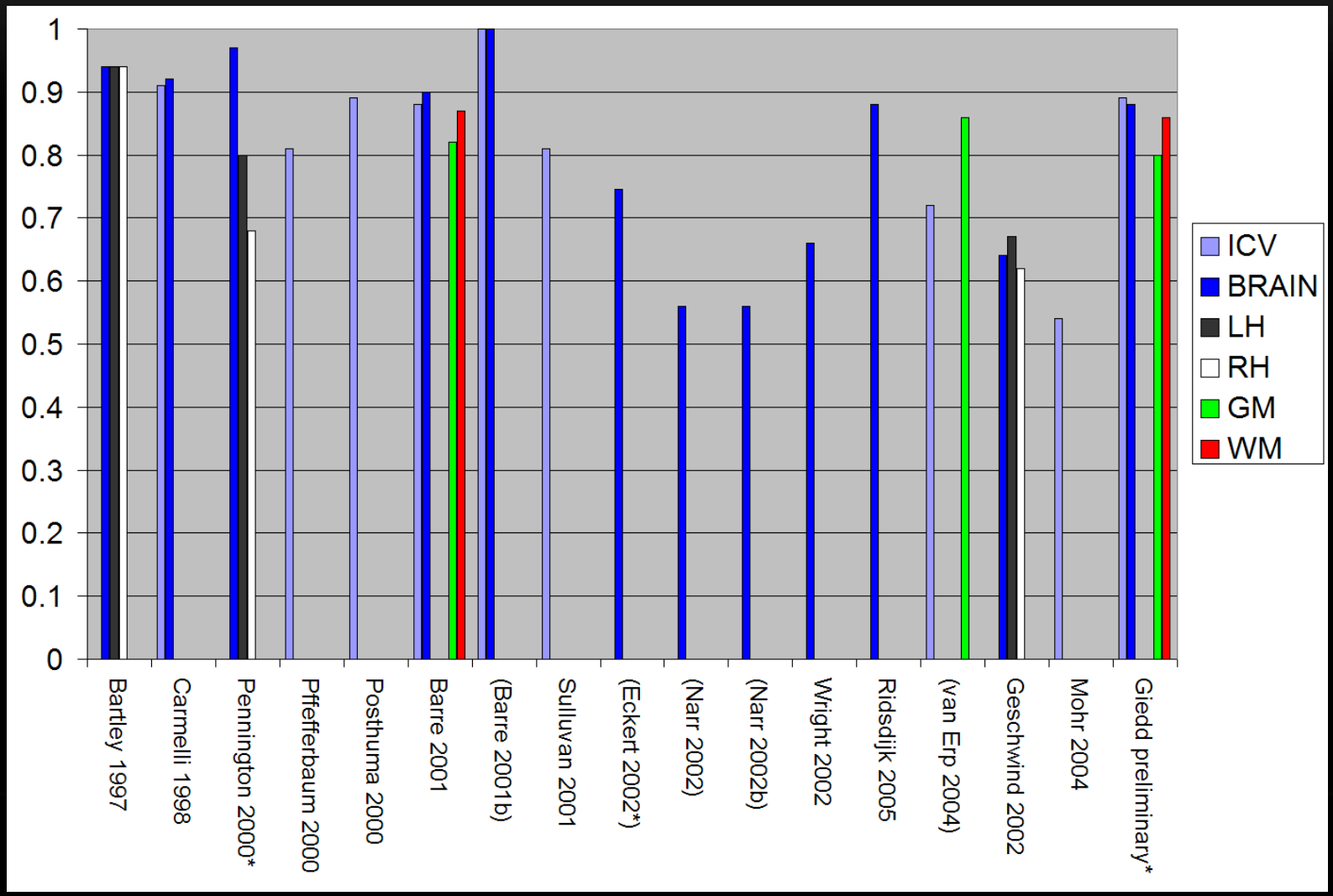
$$c^2 = (2 * Cor_{DZ}) - Cor_{MZ}$$

$$e^2 = 1 - a^2 - c^2$$

Covariances also can be used with continuous data


Caudate Nucleus Volumes
MZ: 0.83
DZ: 0.40
MZ
DZ
Twin 2
Twin 1
CN2
CN1
13
12
11
10
9
8
7
8
9
10
11
12
13

SURVEY OF HERITABILITY ESTIMATES OF LARGE NEUROANATOMIC STRUCTURES IN TYPICAL POPULATIONS (VOLUMETRIC)



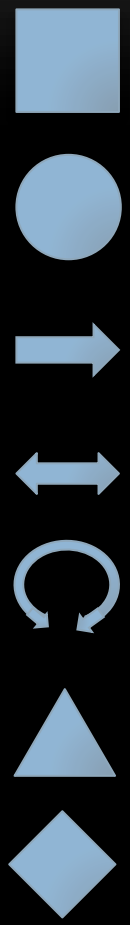
FALCONER ESTIMATION: LIMITATIONS

- Falconer estimation usually provides a good estimate of heritability, but has several limitations relative to other methods
  - Can obtain nonsensical estimates of heritability
  - Do not account for relative precision between MZ and DZ groups (based on both N and r)
  - Not easily generalizable
  - Inefficient with missing data
  - Not suitable for selected samples

STRUCTURAL EQUATION MODELS (SEM) AND PATH ANALYSIS

- Explicitly model influences on both observed and unobserved (i.e. latent) factors based on combinations of causal and correlational relationships.
  - General framework for customizable statistical analysis and hypothesis testing.
  - Expectations for observed variances, covariances, and means can be calculated based on parameter esstimates and compared to the real data.
- Most common inferential statistics (e.g. regression, ANOVA, correlation, factor analysis) can be considered specific cases of SEM).
- SEM is widely used in quantitative and behavioral genetics, but also has applications in numerous branches of science.
- Generally uses likelihood-based optimization strategies
  - Allows for straightforward hypothesis testing.
- Path analysis represents SEM's mathematically-equivalent visual analog.

PATH DIAGRAMS AND PATH ANALYSIS



Observed Variable

Latent Variable

Causal Path

Correlational Path

Self-correlation/Variance

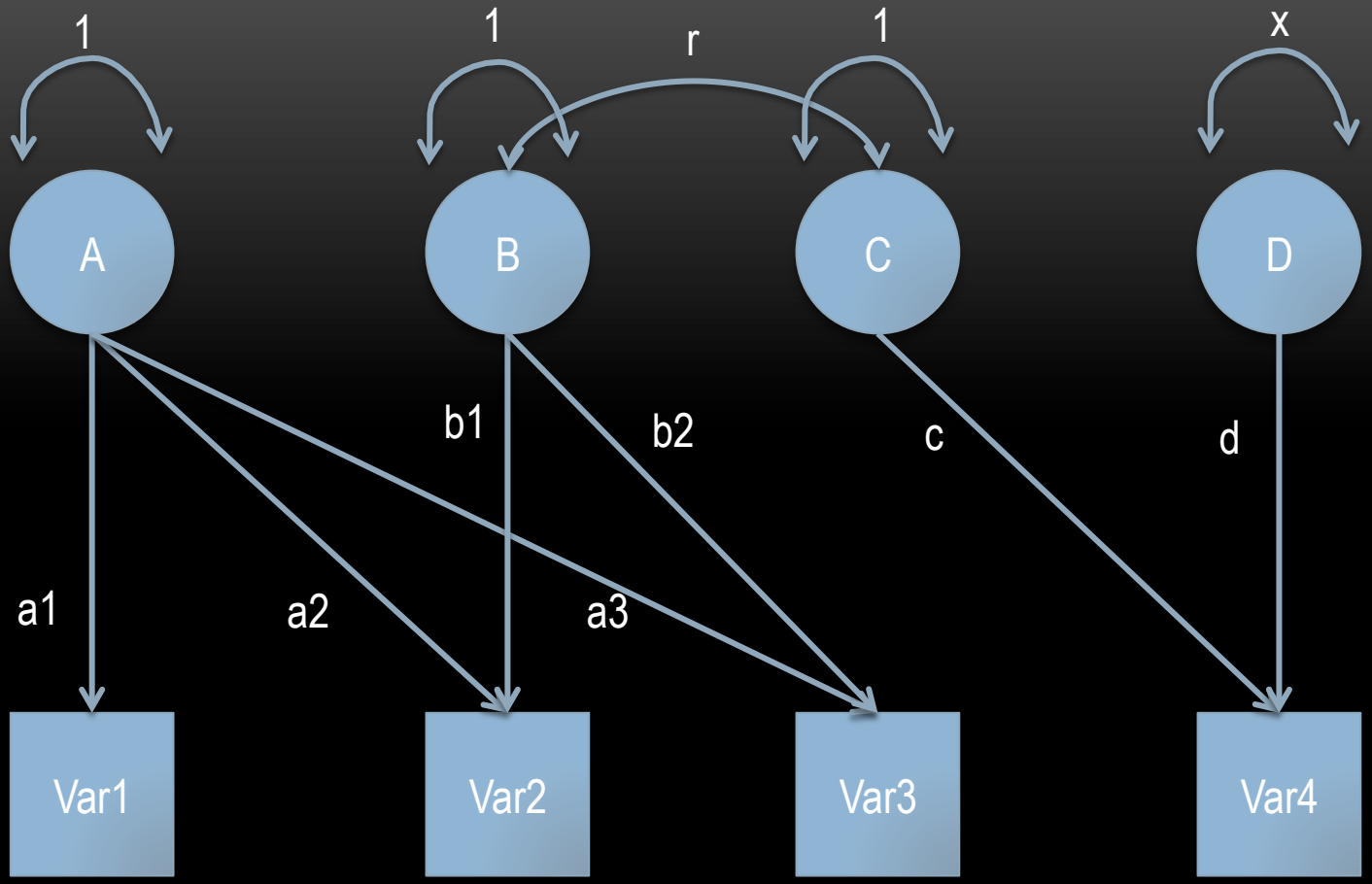
Mean

Definition Variable

- Wright's Rules:
  - Expected Covariance between two variables is the sum of all possible paths between them
  - Each path represents the product of all parameters along the path
  - First move back along causal paths (variables can be correlated by common causal factors, not shared effects)
  - Can only change directions once for each path
  - Can only traverse one correlational arrow along the path
  - Latent variables usually standardized to variance=1

Contrived Example:

Path Diagram



Expected Covariance

	Var1	Var2	Var3	Var4
Var1	a_1^2	$a_1 a_2$	$a_1 a_3$	0
Var2	$a_1 a_2$	$a_2^2 + b_1^2$	$a_2 a_3 + b_1 b_2$	b_1 rc
Var3	$a_1 a_3$	$a_2 a_3 + b_1 b_2$	$a_3^2 + b_2^2$	b_2 rc
Var4	0	b_1 rc	b_2 rc	$c^2 + xd^2$

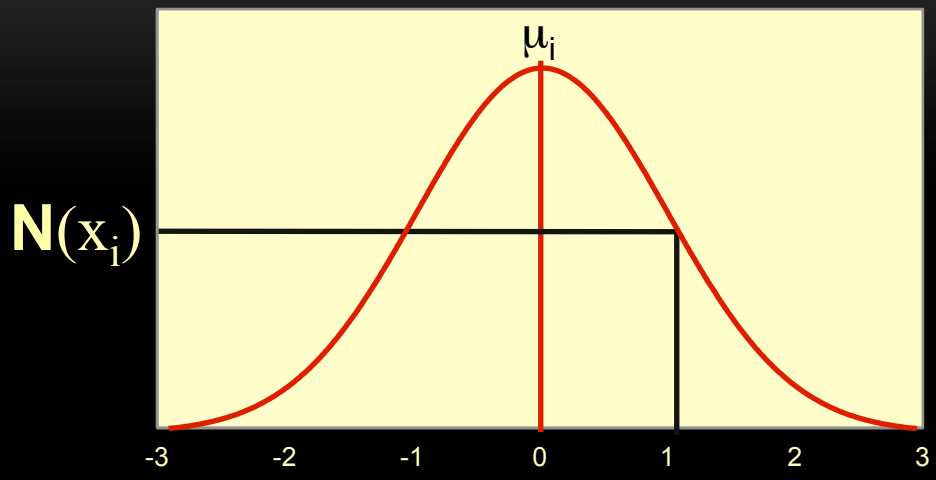
OPTIMIZATION

- Multiple SEM programs exist: OpenMx, AMOS, LISREL, EQS, Mplus, etc.
- Nearly all SEM methods employ maximum likelihood optimization
- User provides reasonable starting values for model parameters
- Program calculates likelihood of observing data given parameters
- Program then calculates likely changes in parameter estimates to improve model fit to data and changes them. Likelihood is then recalculated.
- Iterations continue until likelihood statistic cannot be improved.
- Generally computationally intense
- Other optimization statistics (e.g. least squares) also can be used.

MAXIMUM LIKELIHOOD

- Generalizable
- Precise and unbiased parameter estimates with large sample sizes
- Approximately normal error distributions of parameter estimates
- Robust and maximally informative to situations of missing data and ascertainment biases.
- ln(likelihood) often used to facilitate computation
- Often allow for straightforward hypothesis testing
  - Difference in -2ln(ML) between a model and submodel generally follows χ^2 distribution with df = difference in number of parameters
- Computationally expensive relative to other methods

Simple: Univariate Mean

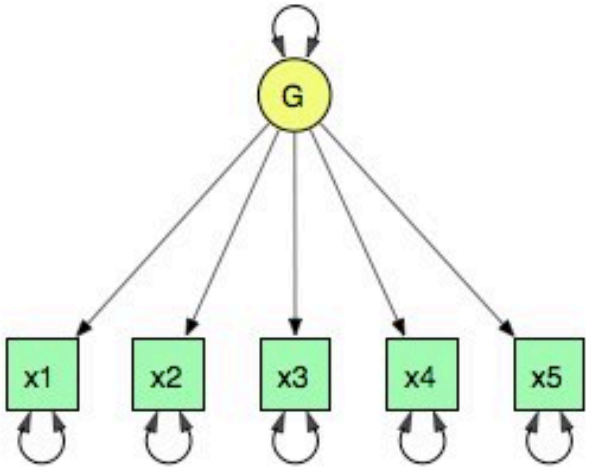


Complex: Continuous Multivariate

$$L = |2\pi\Sigma|^{-2m/2} \exp\left[-\frac{1}{2}(x_i - \mu_i)'\Sigma - 1(x_i - \mu_i)\right]$$

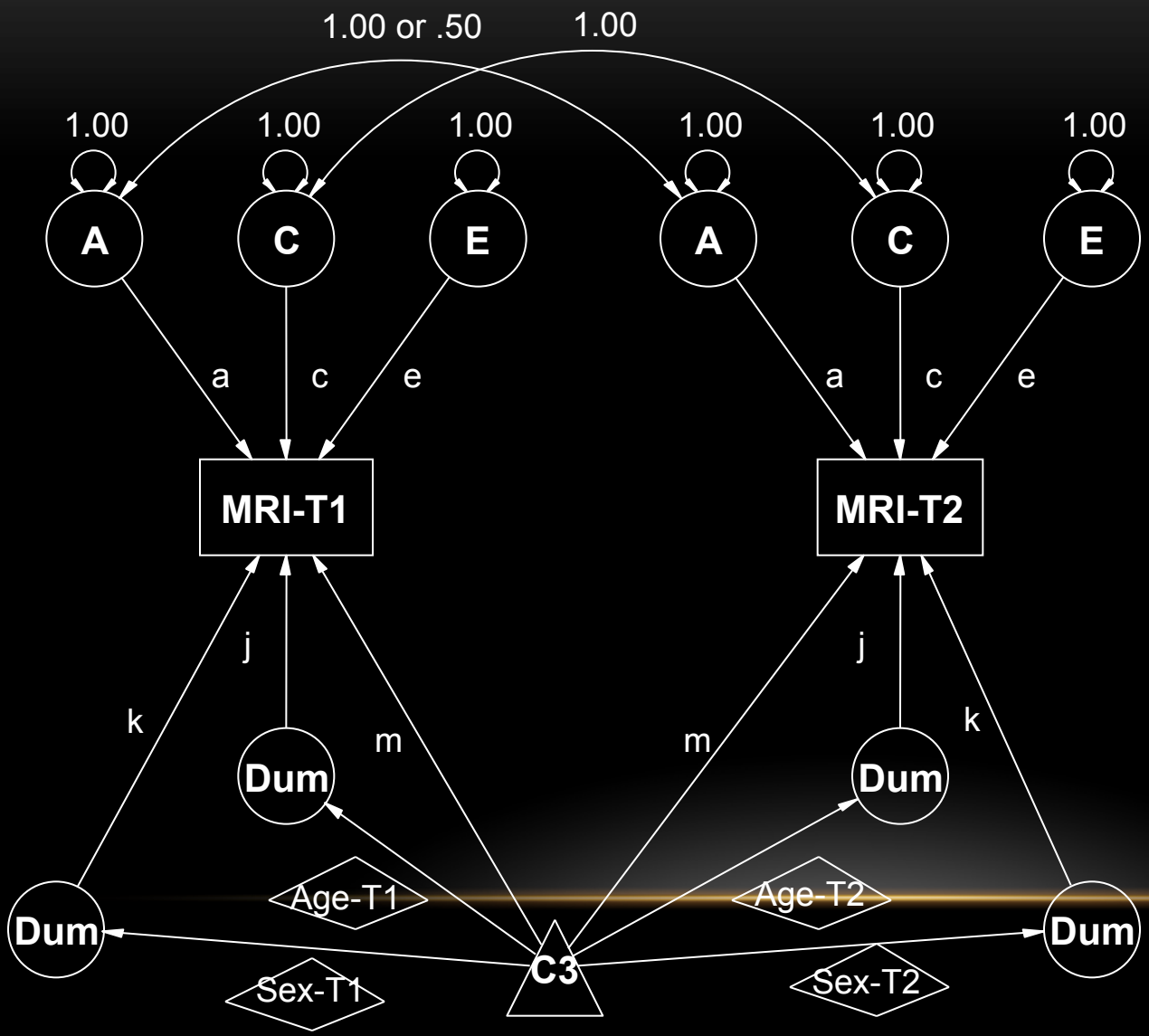
MX/OPEN MX

- Mx is a free structural equation modeling software program designed by Mike Neale with features that make multigroup modeling (e.g. twins, families, cases/controls) easy.
- Open Mx is represents a new iteration of Mx spearheaded by Steve Boker (UVA) and Mike Neale (VIPBG) that is open source and fully integrated into the R statistical programming environment.
- Can specify models based both on matrix formulations or path coefficients
- http://openmx.psyc.virginia.edu/




```
require(OpenMx)
data(demoOneFactor)
manifests <- names(demoOneFactor)
latents <- c("G")
factorModel <- mxModel("One Factor",
    type="RAM",
    manifestVars = manifests,
    latentVars = latents,
    mxPath(from=latents, to=manifests),
    mxPath(from=manifests, arrows=2),
    mxPath(from=latents, arrows=2,
        free=FALSE, values=1.0),
    mxData(cov(demoOneFactor), type="cov",
        numObs=500))
summary(mxRun(factorModel))
```

MODIFIED "CLASSIC" UNIVARIATE TWIN MODEL



Variance Components

a^2 = additive genetic

c^2 = shared env.

e^2 = unique env.

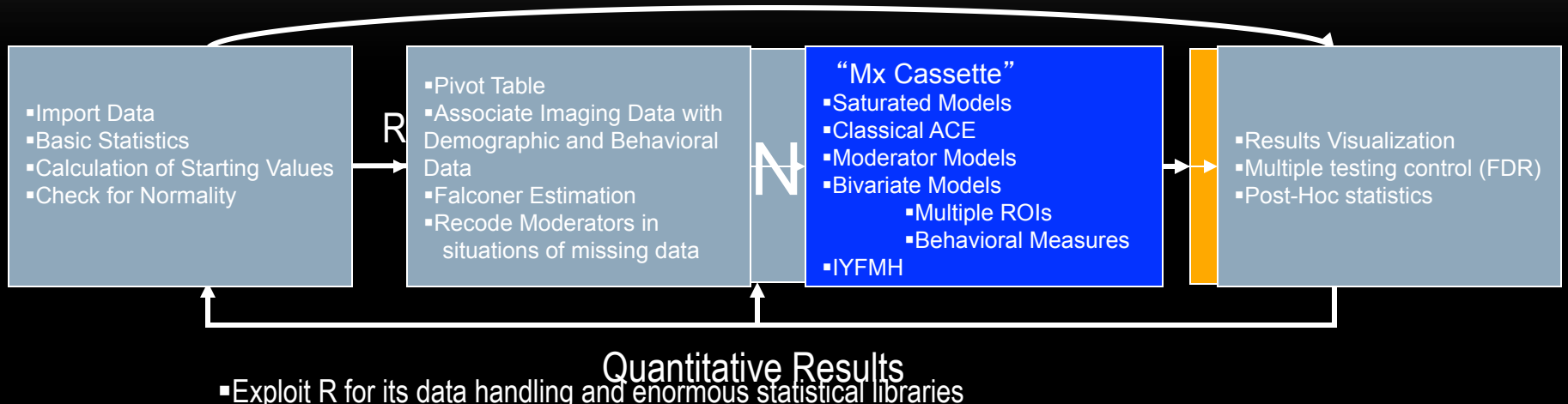
Means Regression

age

gender

SEM CAN EASILY BE INTEGRATED INTO IMAGING PIPELINES

Statistical Genetic Pipeline: A Glorified For Loop in R

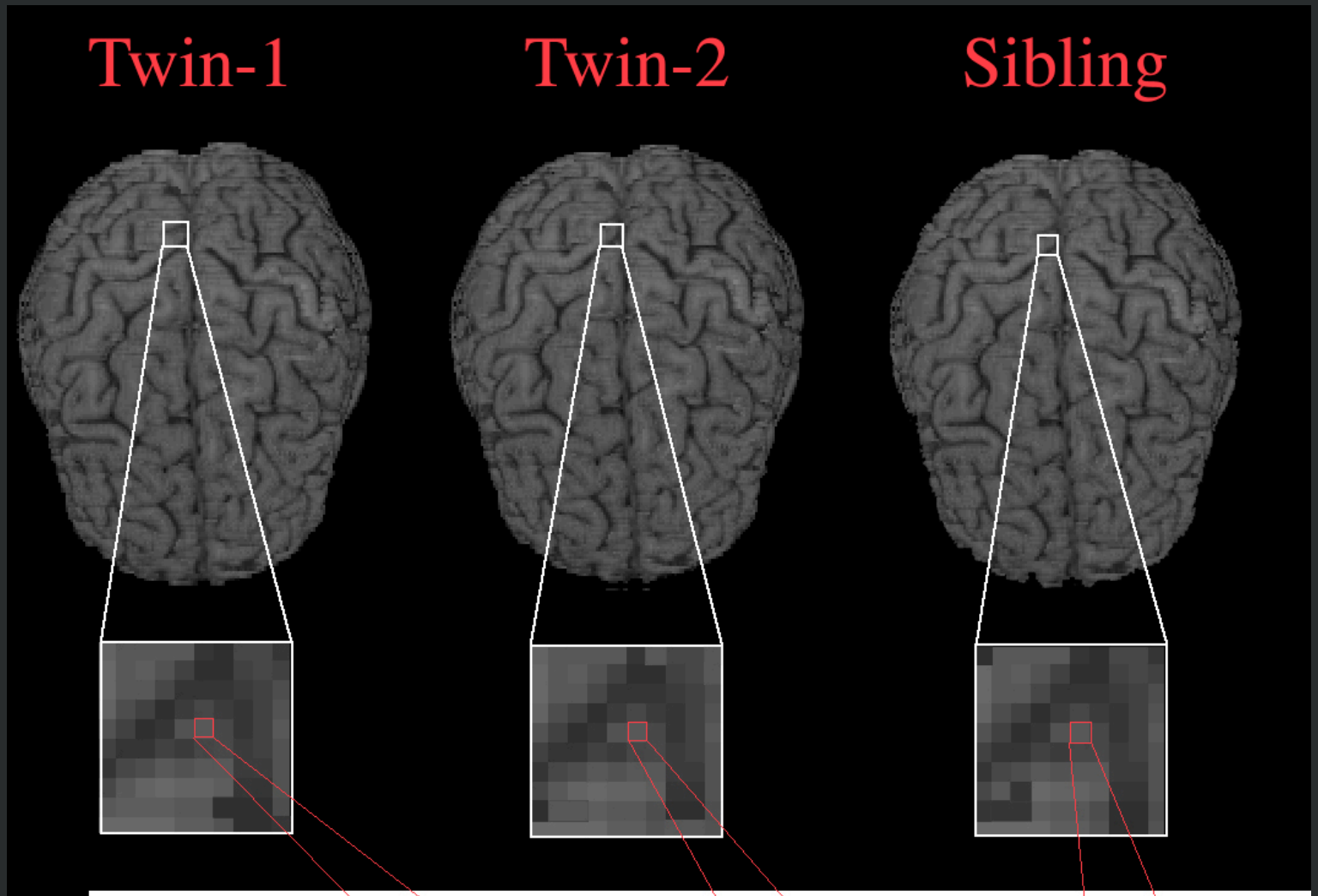


- Exploit R for its data handling and enormous statistical libraries
- Use Mx for optimizer and flexibility in SEM design

National Institutes for Mental Health
· Subject Recruitment
· Data Acquisition
· Behavioral and Phenotypic Analyses

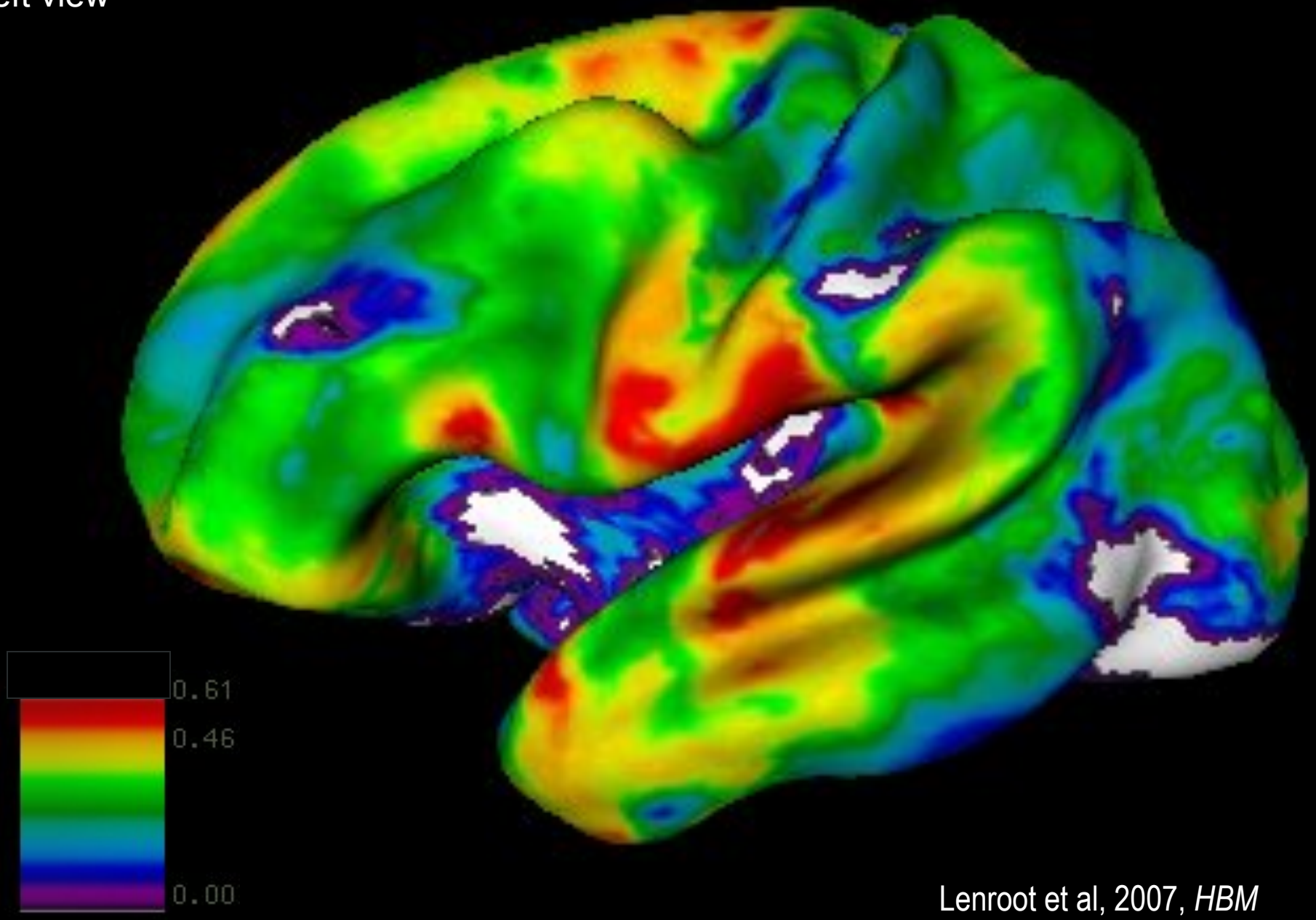
Montreal Neurological Institute
· Neuroimaging Methods
· Image Processing

Virginia Institute For Psychiatric and Behavioral Genetics
·Statistical Methods Development
· Quantitative Genetic Analyses

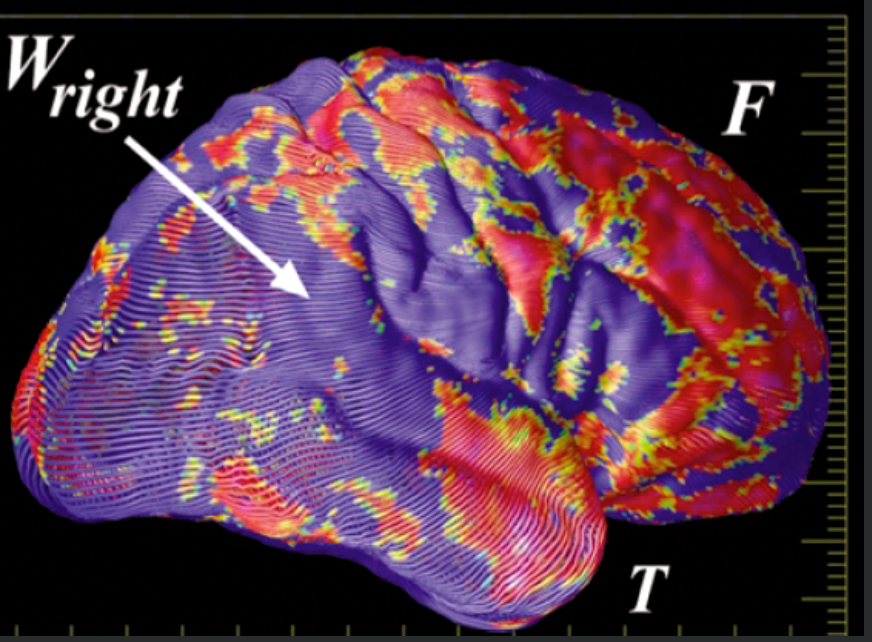


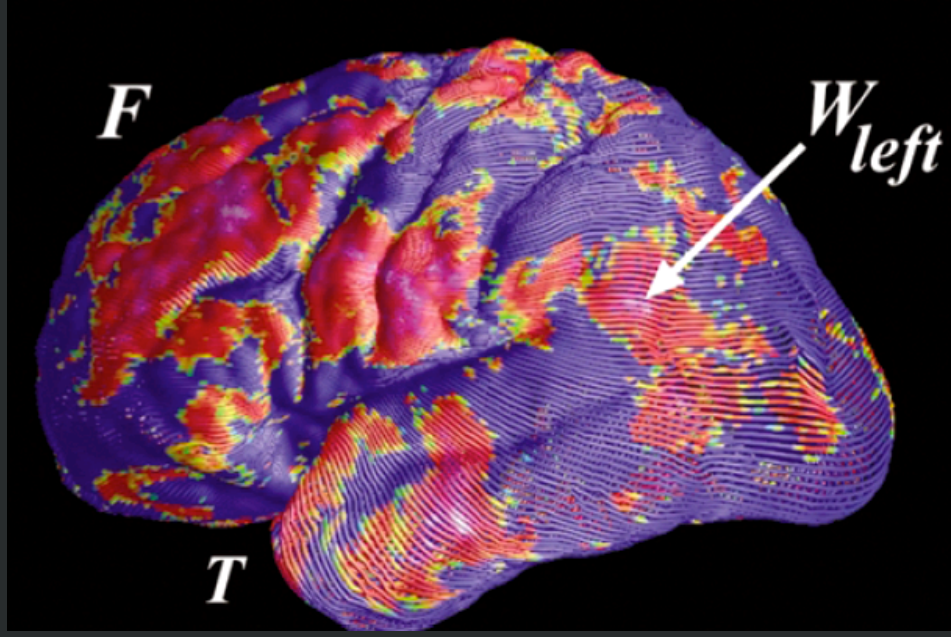
Famid	Twinzyg	Twin1-CT	Twin2-CT	Sib-CT
1010	0	4.1402	3.9789	4.0132
0915	1	3.9982	4.1264	3.4587
7827	0	4.1232	3.9948	4.3972
...				

a^2 MLE
left view

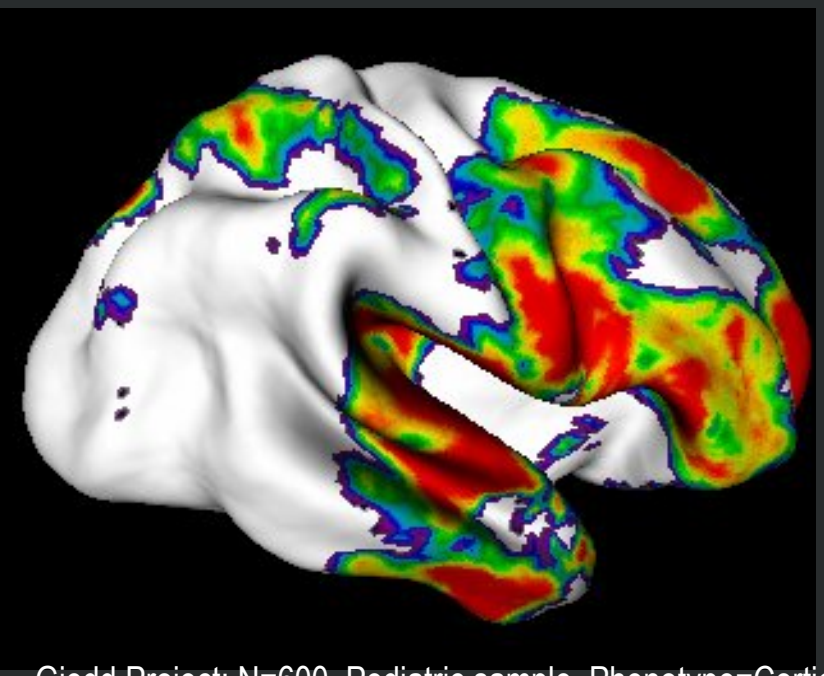


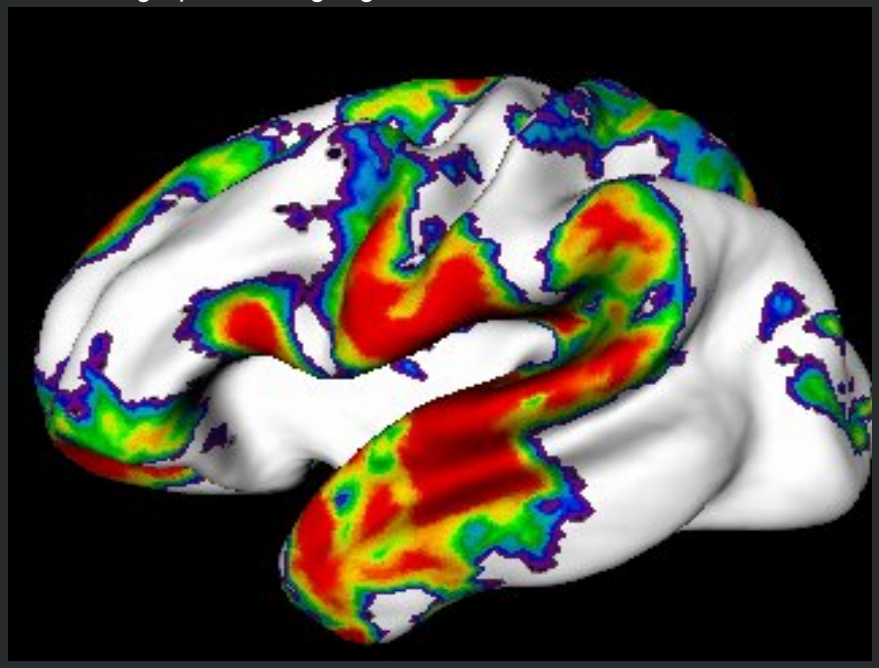
Lenroot et al, 2007, *HBM*





Thompson et al.:N=40, Adult sample, Phenotype=Gray Matter Density, UCLA image processing algorithms, Falconer Estimation



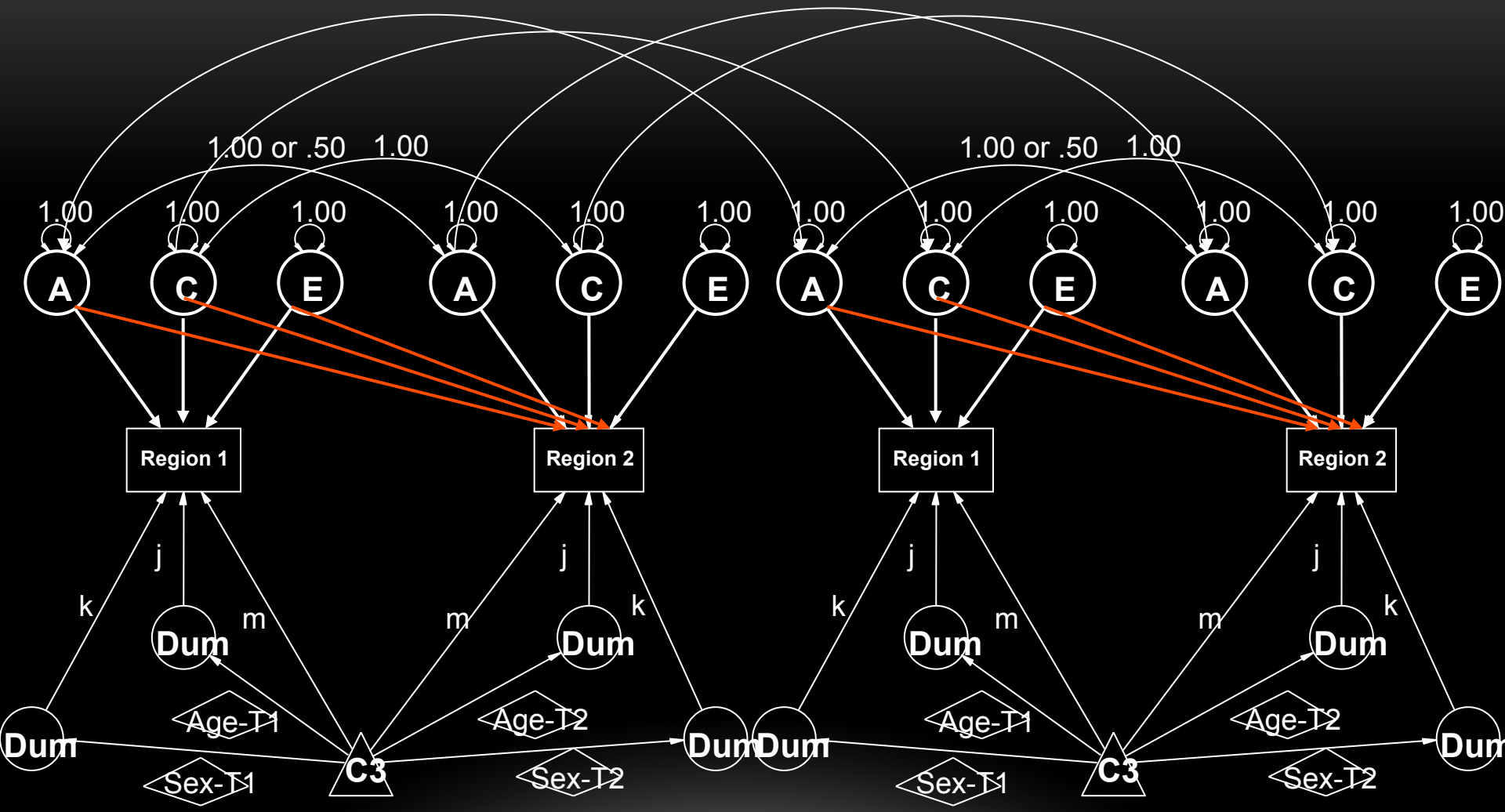


Giedd Project: N=600, Pediatric sample, Phenotype=Cortical Thickness, MNI image processing algorithms, SEM in Mx

EXTENSIONS

- Assortative mating
- Sex Limitation
- Gene x Environment Interactions
- Gene x Environmental Correlations
- Multivariate Models
  - Biometric
  - Psychometric
  - Longitudinal
  - Customizable

BIVARIATE CHOLESKY DECOMPOSITION



TWIN 1

TWIN 2

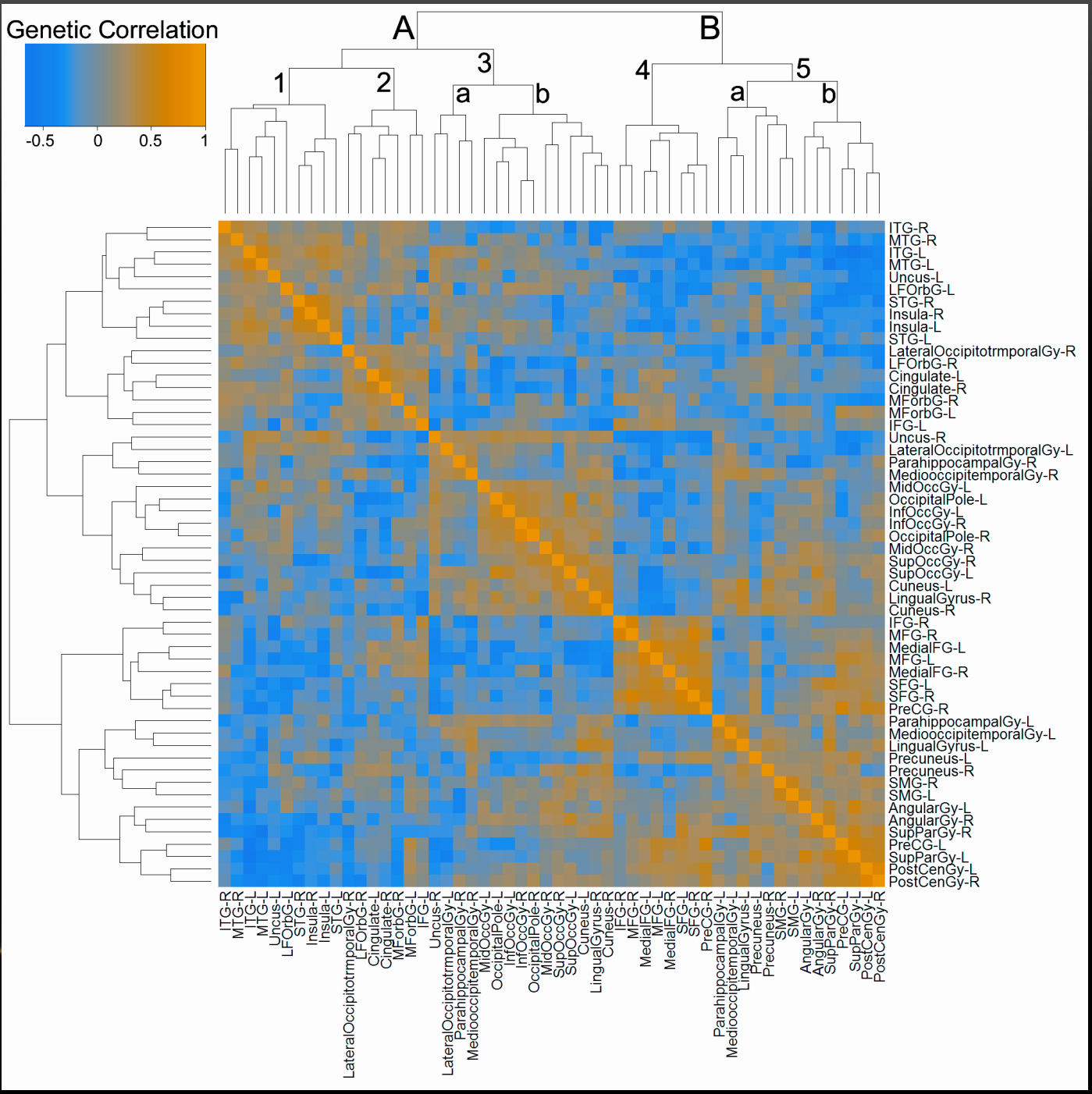
(plus singleton sib relationships modeled→)

Genetic correlations
For average CT at 54 gyral ROIs, adjusted for mean CT.

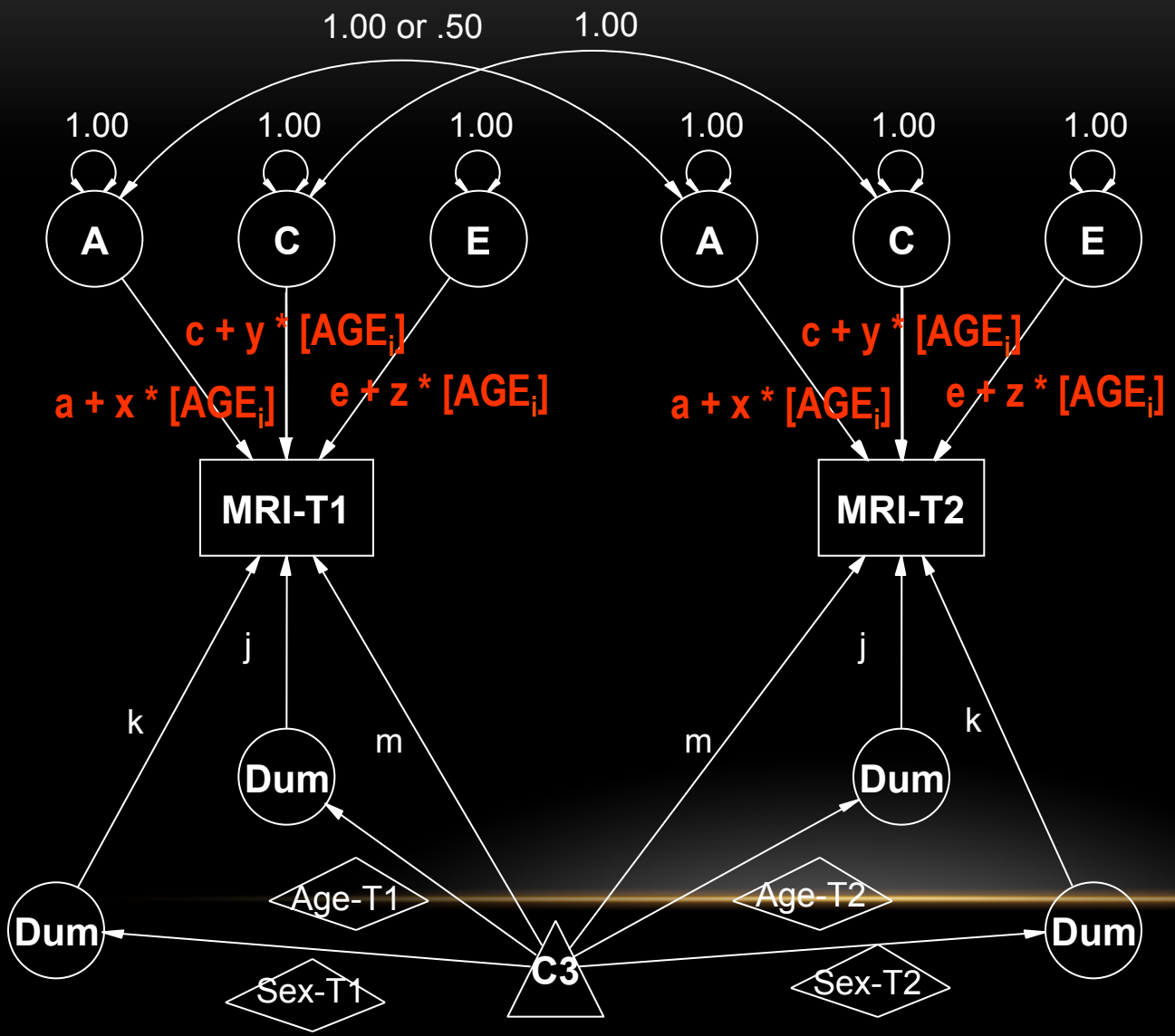
$$r_{x,y} = \frac{A_{xy}}{\sqrt{(A_x * A_y)}}$$

Methods similar to
I.C. Wright et al. (2002)

Schmitt et al, 2008,
Cerebral Cortex

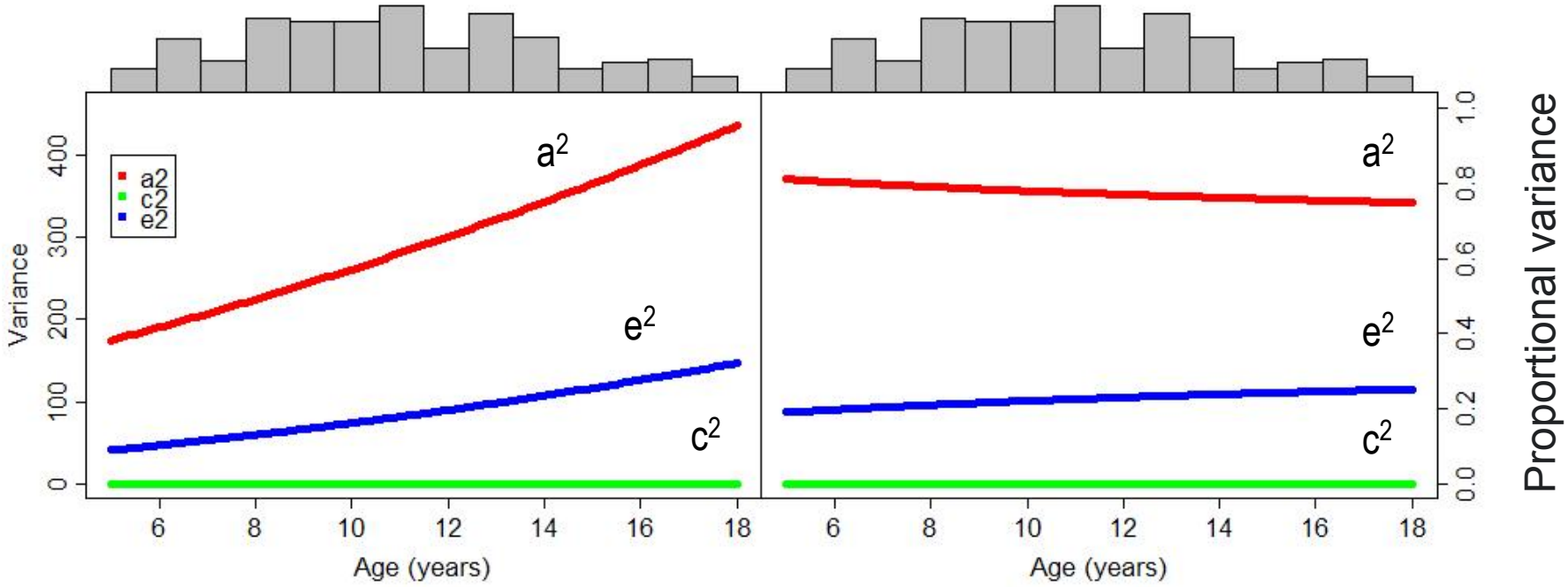


HETEROGENIETY / GENE X ENVIRONMENT INTERACTION

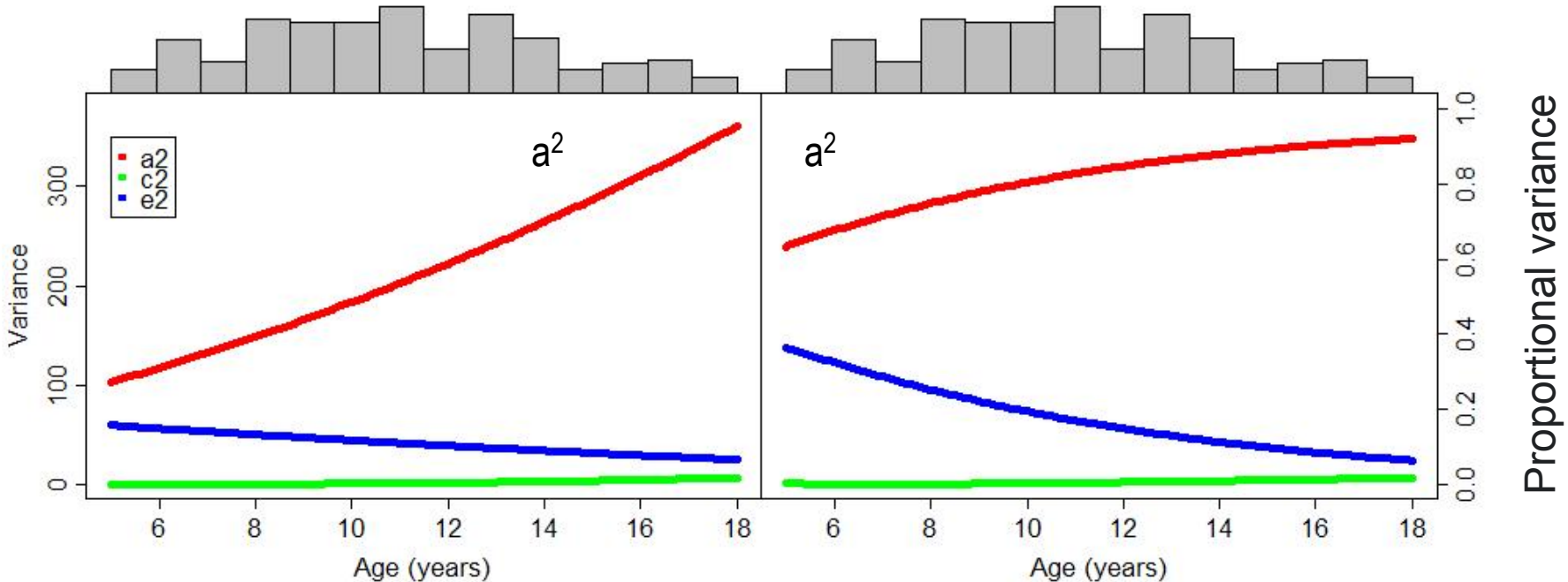


Purcell S, Sham P. 2002

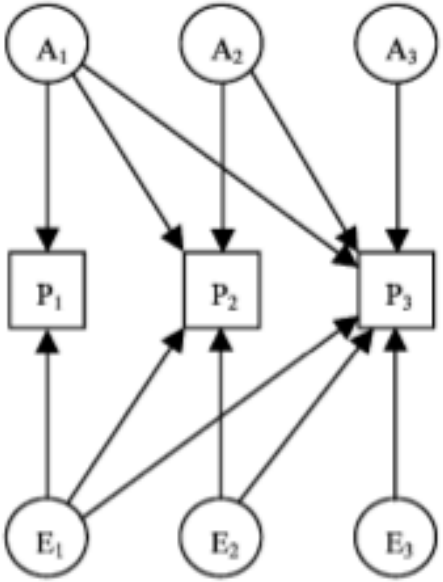
FRONTAL GRAY MATTER



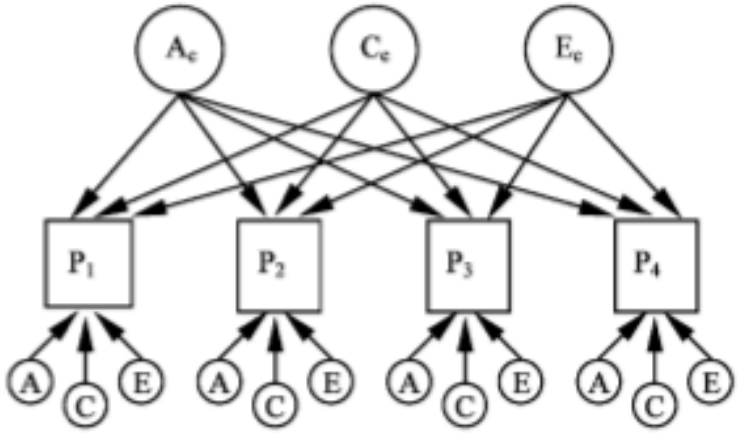
FRONTAL WHITE MATTER



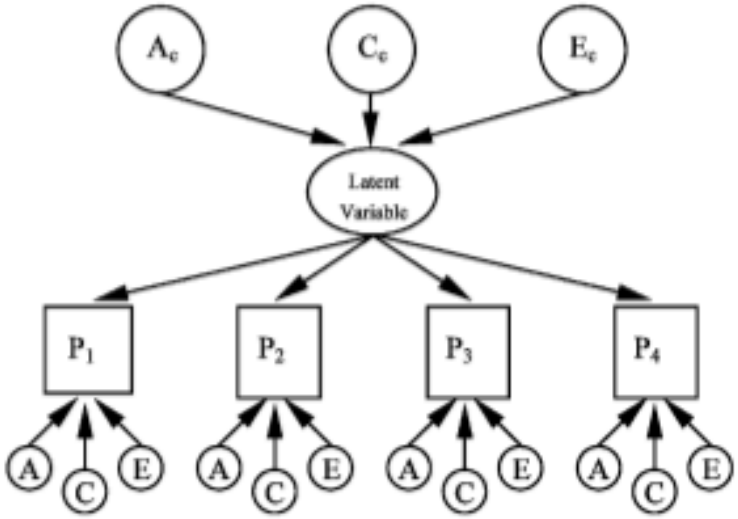
(A) Cholesky Model



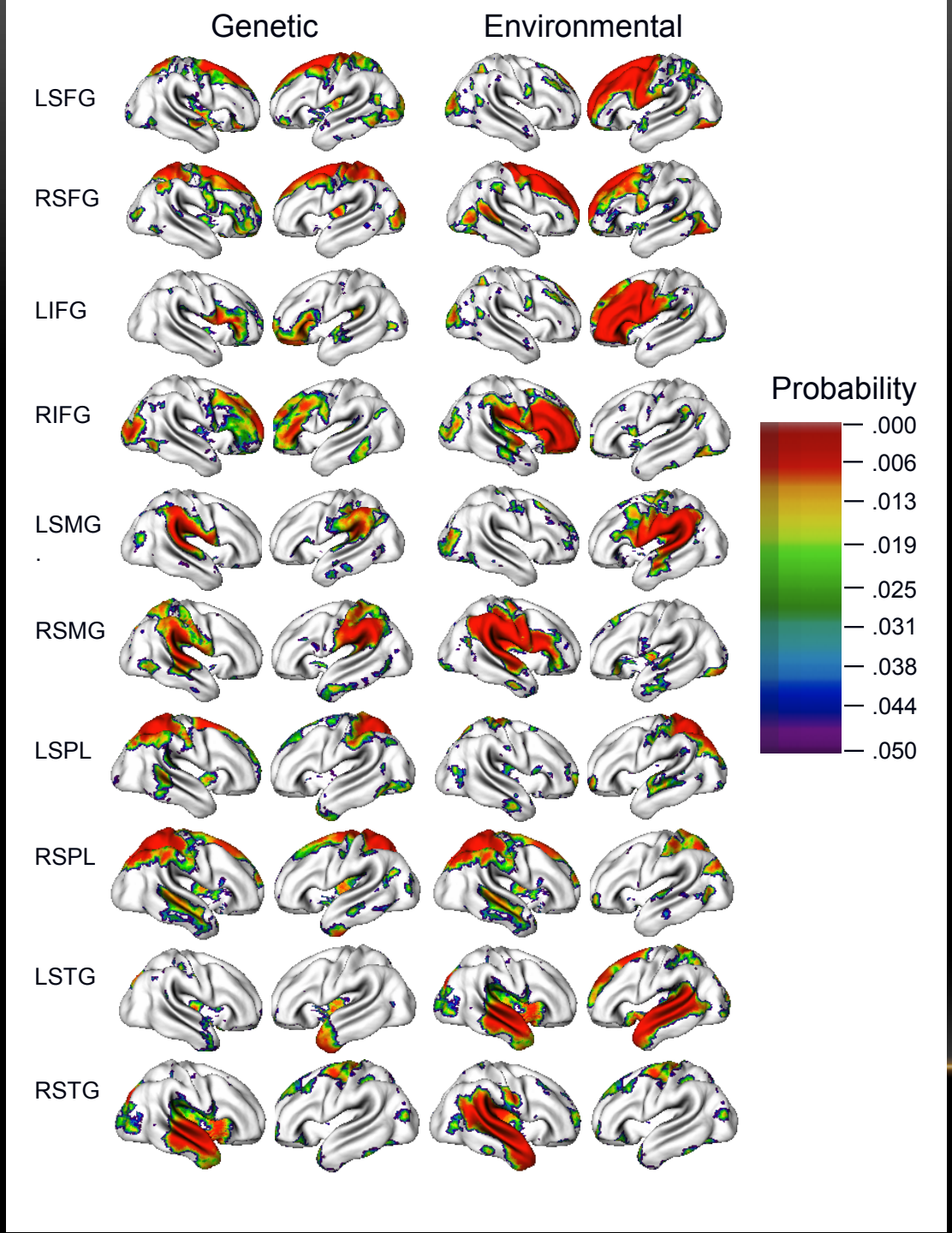
(B) Independent Pathways Model



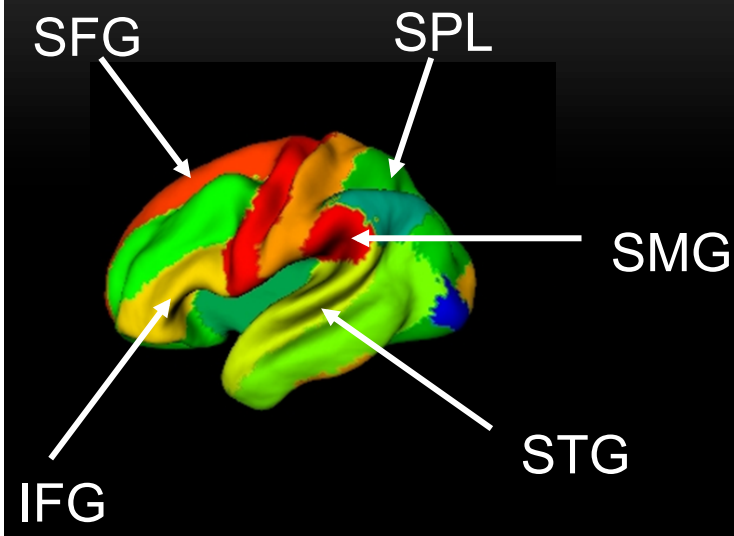
(C) Common Pathways Model



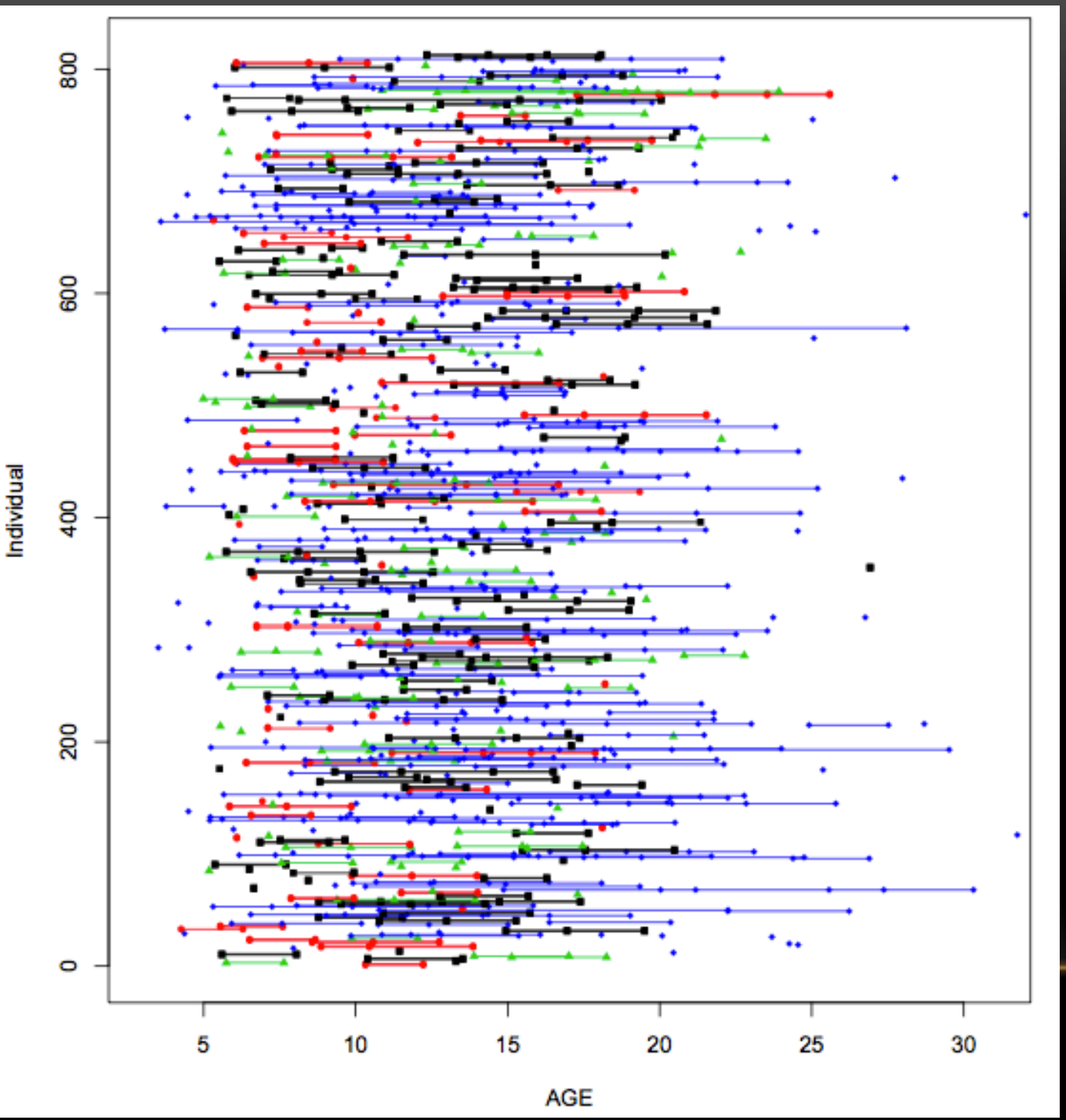
Evans DM et al. “Biometrical Genetics.” *Biological Psychology* 61:33-51 (2002).



Semi-Multivariate approaches to Sidestep the “curse of dimensionality”

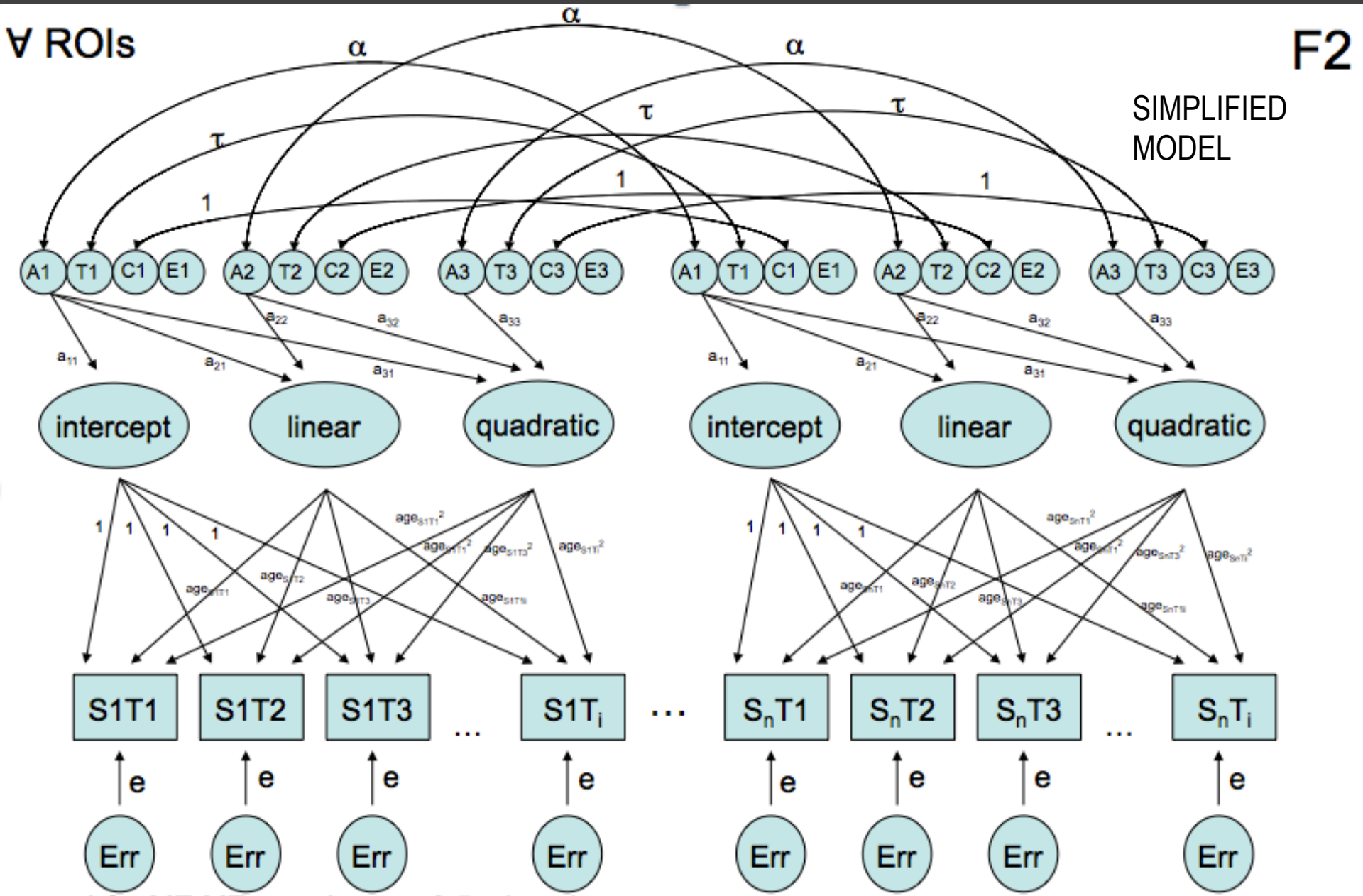


Schmitt et al (2008), *NeuroImage*

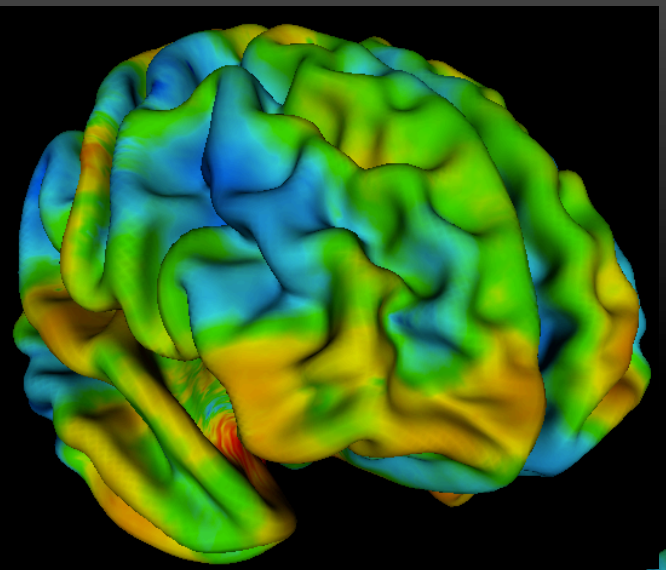


<u>1748 brain MRI:</u>
249 MZ twins (black)
131 DZ twins (red)
110 sibs of twins (green)
302 singleton (blue)

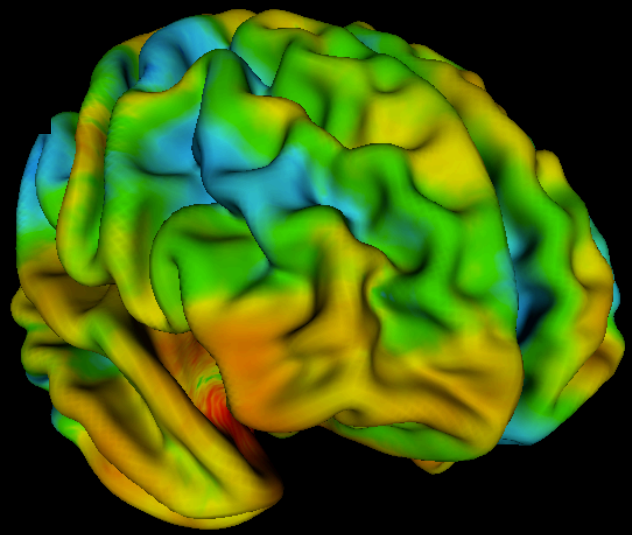
- up to 5 individuals/family
- up to 8 scans/individual


∀ ROIs
F2
SIMPLIFIED MODEL
α
τ
1
A1 T1 C1 E1
A2 T2 C2 E2
A3 T3 C3 E3
a11 a21 a31 a22 a32 a33
intercept
linear
quadratic
S1T1 S1T2 S1T3 … S1Ti
SnT1 SnT2 SnT3 … SnTi
e
Err
α = 1 for MZ-MZ covariances, 0.5 otherwise
τ = 1 for twin-twin covariances, 0 otherwise

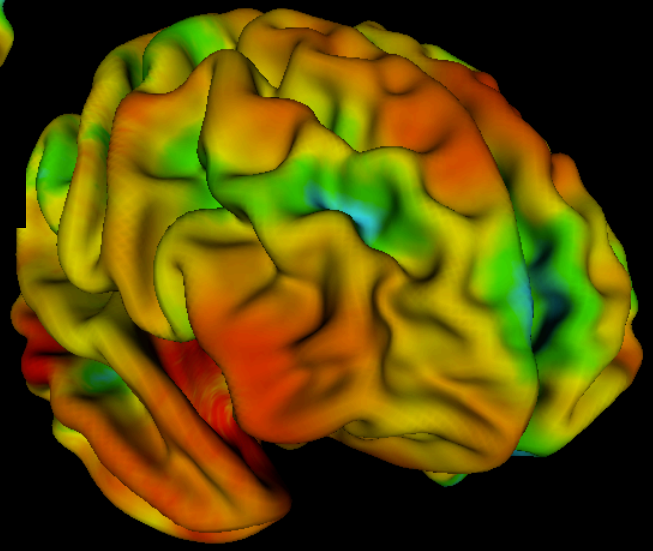
HERITABILITY



Age 5

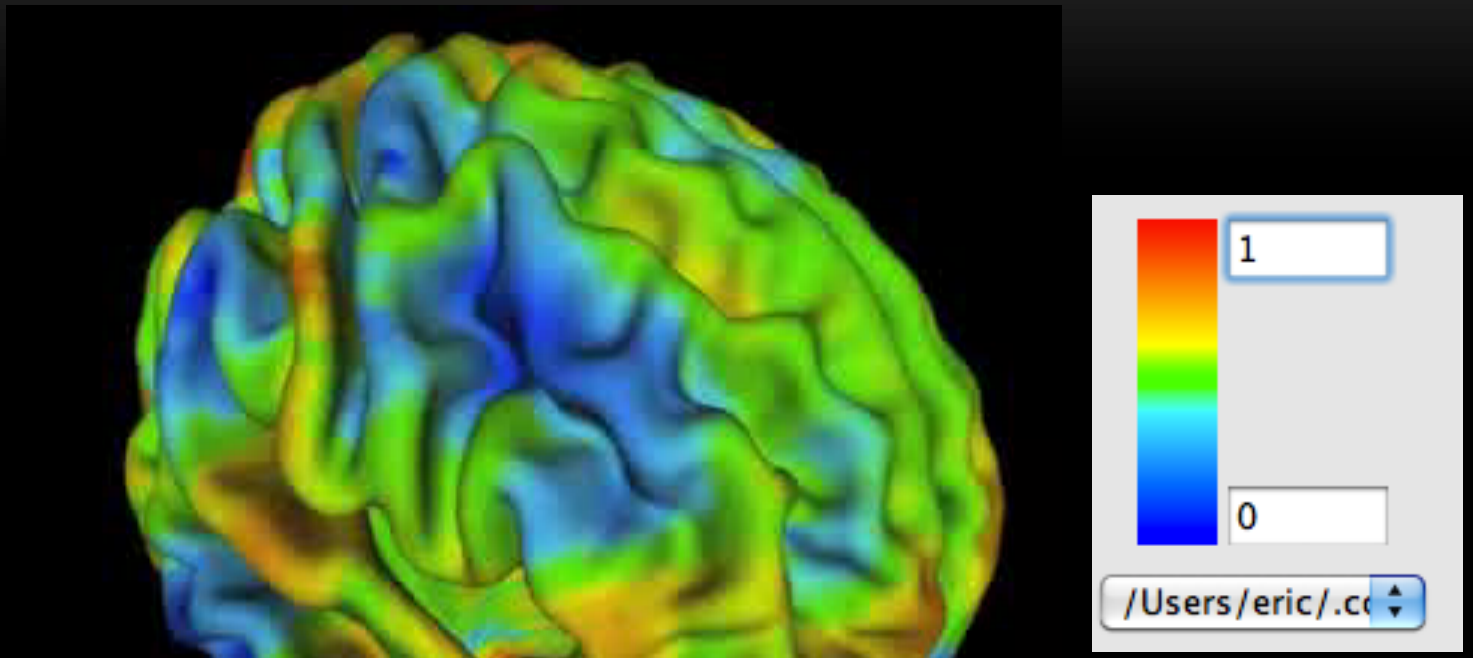


Age 10

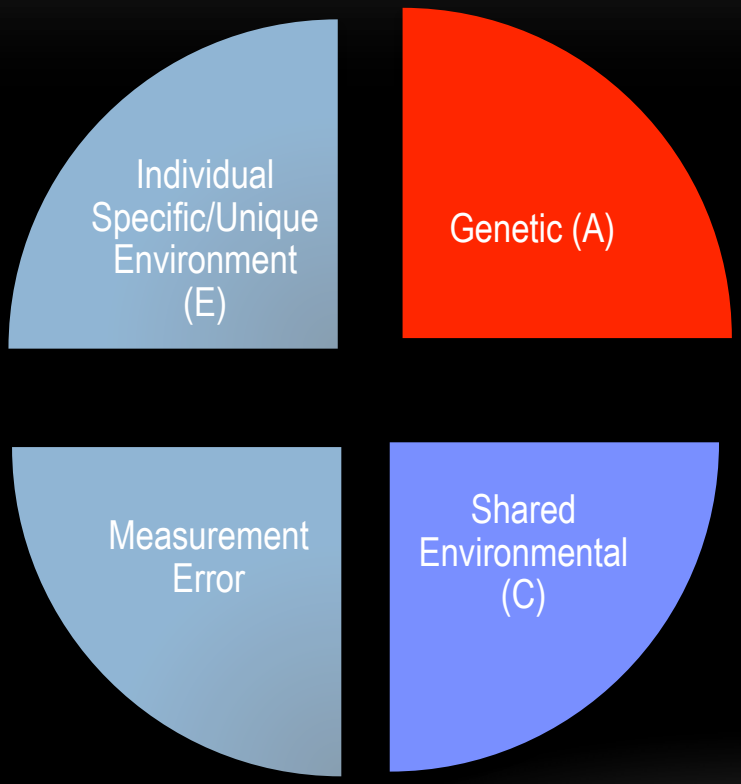


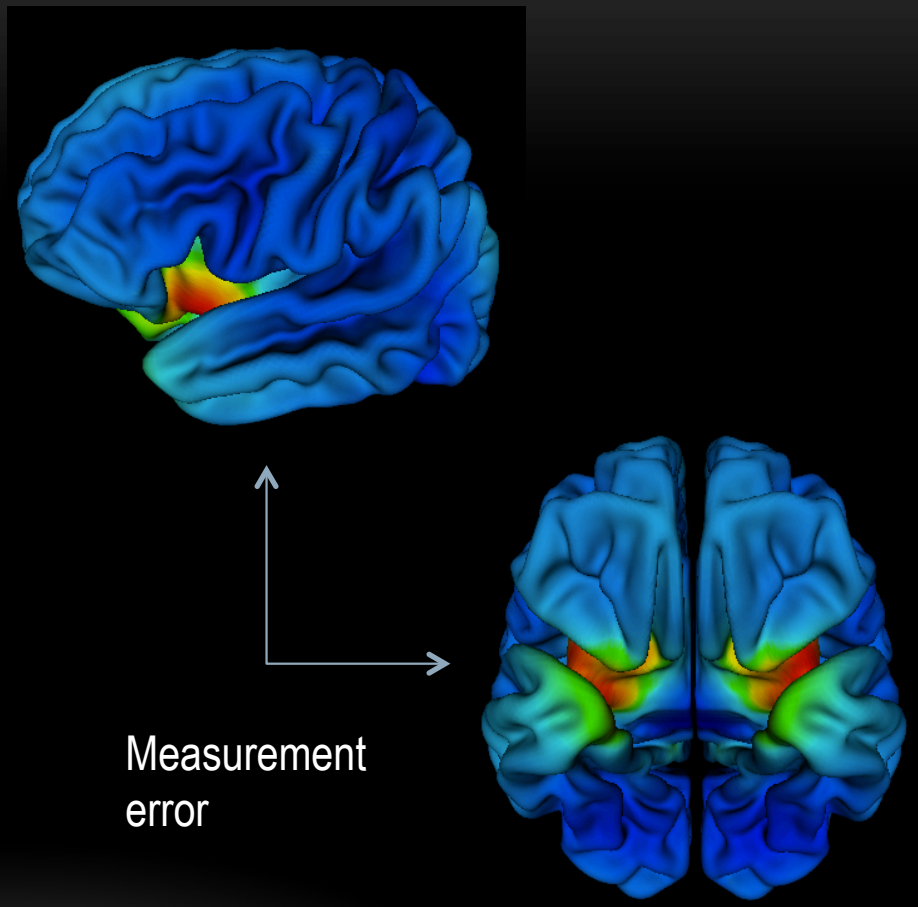
Age 15

DYNAMIC HERITABILITY MAP:AGES 3-18

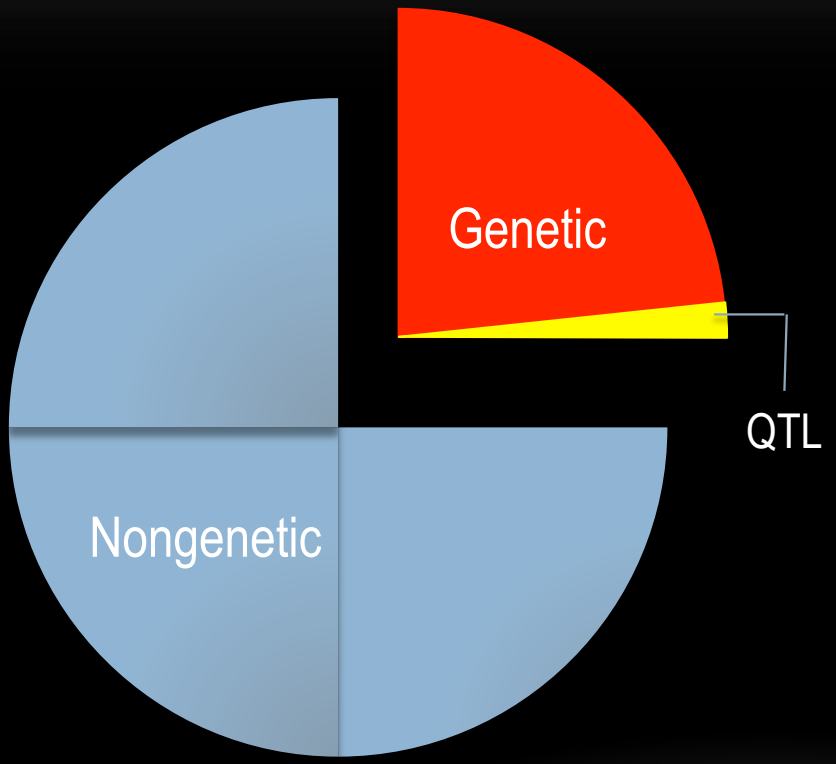


LONGITUDINAL MODELS: DISSOCIATE ERROR FROM E





VARIANCE DECOMPOSITION: QUANTITATIVE TRAIT LOCI

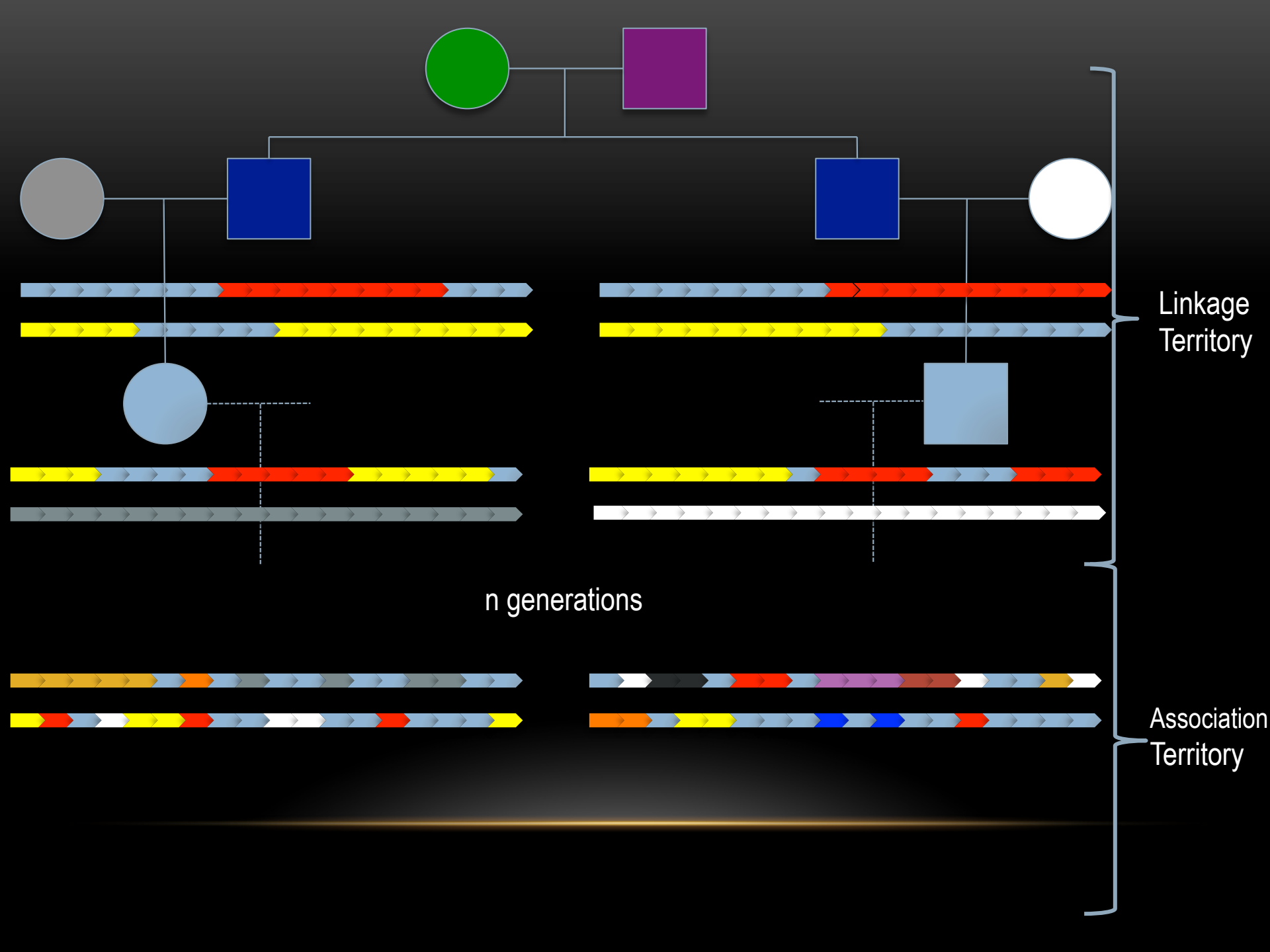


$$Var(x) = Var_{Genetic} + Var_{QTL} + Var_{Nongenetic}$$

$$Var(x) = Var_{A} + Var_{QTL} + Var_{C} + Var_{E}$$

QUANTITATIVE TRAIT LOCI: MAPPING

- QUANTITATIVE TRAIT LOCUS (QTL): A genetic locus that influences a quantitative trait. QTLs can be identified via both linkage and association.
- LINKAGE
  - Trace allelic markers through families and correlate with phenotype
  - Specific allele doesn't matter, just how alleles segregate through families
  - Requires pedigree information/family structure
- ASSOCIATION
  - Correlate presence of specific alleles with variation in quantitative trait
  - Case-control designs
    - Powerful if assumptions met, susceptible to bias from population stratification.
  - Family designs also exist
- COMBINED LINKAGE AND ASSOCIATION
- MULTIPLE STATISTICAL APPROACHES
  - BASIC (T-tests, Linear Regression, Nonlinear Regression, Multiple Linear Regression, Likelihood-based, Logistic)
  - SIB-PAIR METHODS

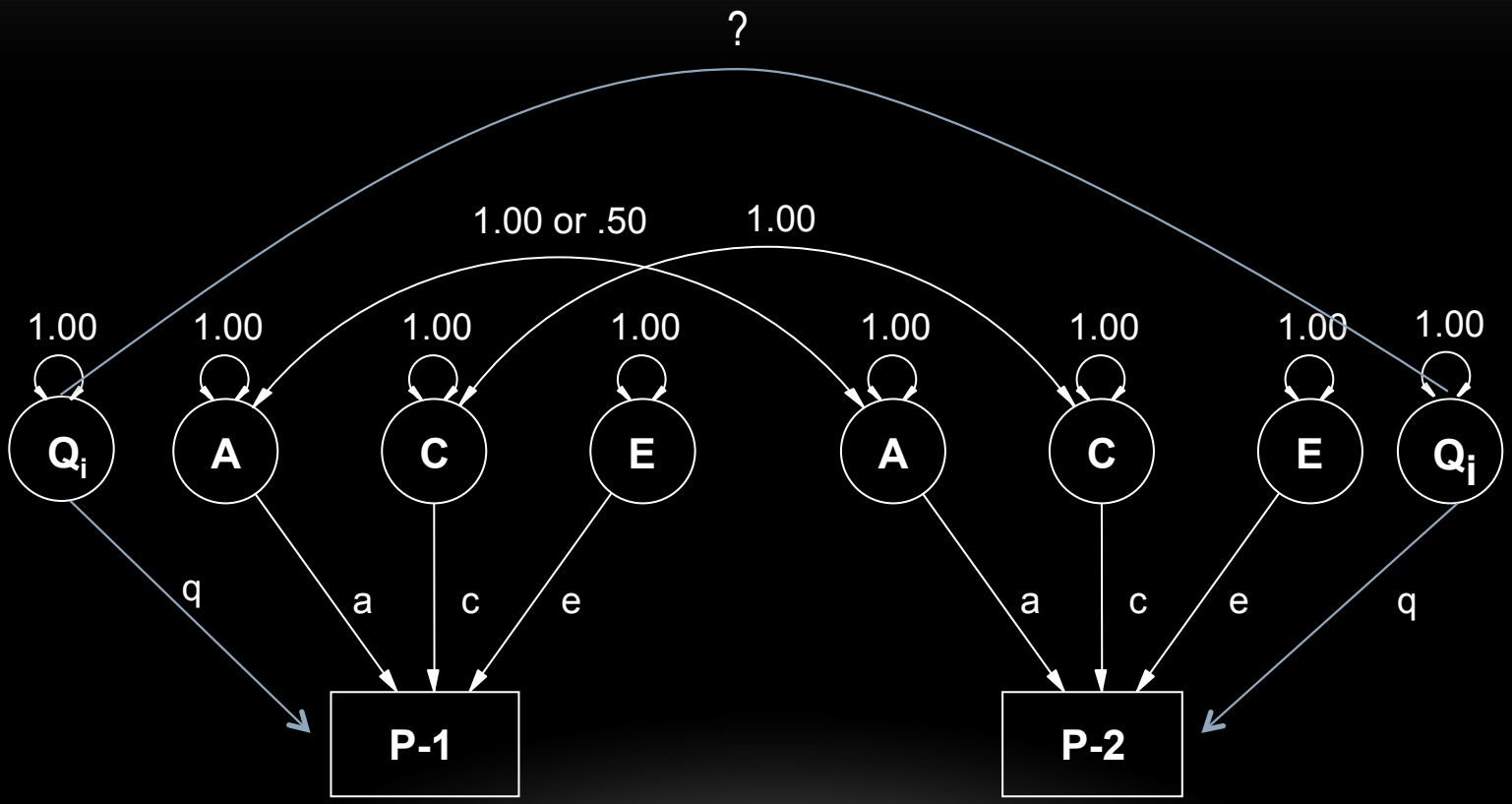

Linkage Territory
n generations
Association Territory

PRINCIPLES OF LINKAGE

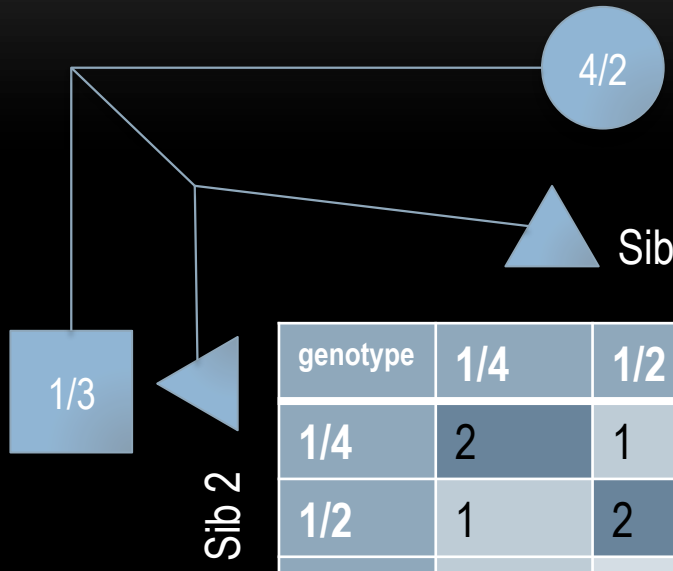
- Mendel's law of assortment states that separate loci segregate independently. We now know that for loci on the same chromosome, independent assortment requires recombination events between loci.
- Morgan discovered that assortment does not always hold for loci on the same chromosome.
- Loci that are in proximity on a chromosome are less likely to be separated by a recombination event.
- θ is defined as the recombination fraction, the probability of a recombination event between two loci. θ between two loci approaches 0.5 the closer loci are to independent assortment.
- Loci with $\theta < 0.5$ are in linkage.
- LOD scores represent log of the likelihood of linkage (H1:$\theta < 0.5$) relative to the null hypothesis (H0:$\theta = 0.5$).
  - LOD > 3.0 usually the standard for statistical significance
  - 3 > LOD > 2 suggestive of linkage---can add more data & reanalyze
- Parametric and Nonparametric models exist
  - Parametric models are very powerful if you know the model
  - Nonparametric models better for complex traits/the ignorant.

$$LOD = \log_{10} \frac{(1-\theta)^{NR}\theta^{R}}{1/2^{(NR+R)}}$$

LINKAGE EXAMPLE: EXTENSION TO TWIN MODEL



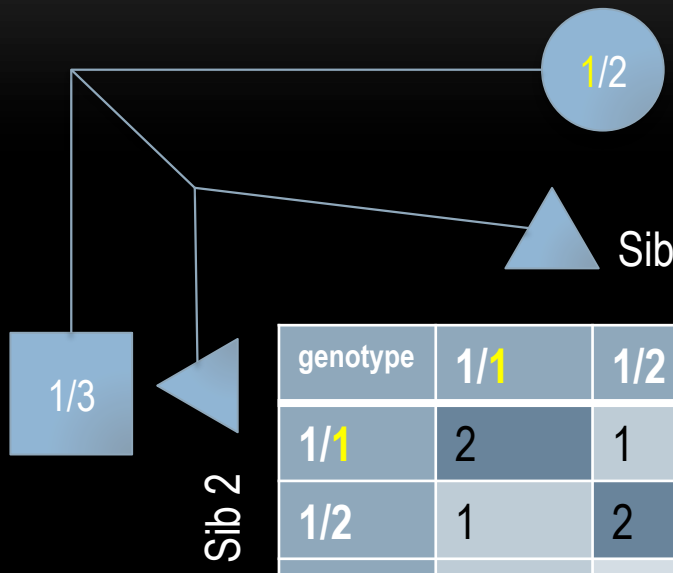
IDENTITY BY DESCENT (IBD)



genotype	1/4	1/2	3/4	3/2
1/4	2	1	1	0
1/2	1	2	0	1
3/4	1	0	2	1
3/2	0	1	1	2

- Alleles are identical by descent (IBD) if they are identical copies of the same ancestral allele.
- Alleles are identical by state (IBS) if they are identical but there is no evidence that they are copies of the same ancestral allele.
- Number of alleles identical by descent can take values 0, 1, or 2
- Conceptually, at *each locus* siblings are genetically unrelated, share half their genes IBD, or are genetically identical.
- At a given QTL, ¼ of sibs will be IBD2, ½ will be IBD1 and ¼ will be IBD0

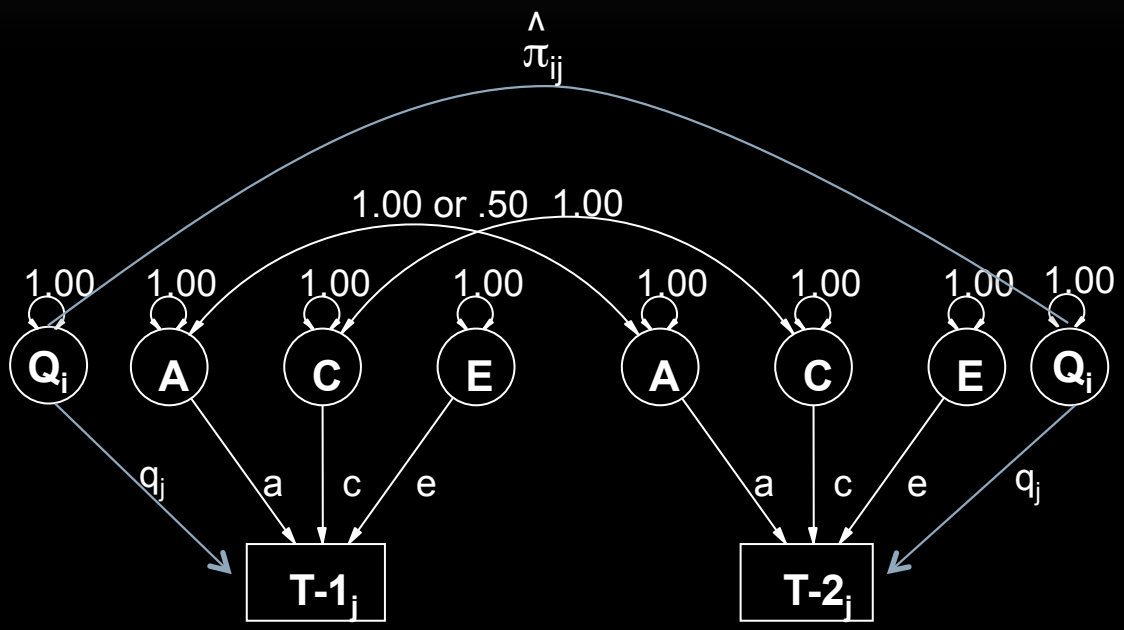
IDENTITY BY DESCENT (IBD)



genotype	1/1	1/2	3/1	3/2
1/1	2	1	1	0
1/2	1	2	0	1
3/1	1	0	2	1
3/2	0	1	1	2

- Alleles are identical by descent (IBD) if they are identical copies of the same ancestral allele.
- Alleles are identical by state (IBS) if they are identical but there is no evidence that they are copies of the same ancestral allele.
- Number of alleles identical by descent can take values 0, 1, or 2
- Conceptually, at *each locus* siblings are genetically unrelated, share half their genes IBD, or are genetically identical.
- At a given QTL, ¼ of sibs will be IBD2, ½ will be IBD1 and ¼ will be IBD0

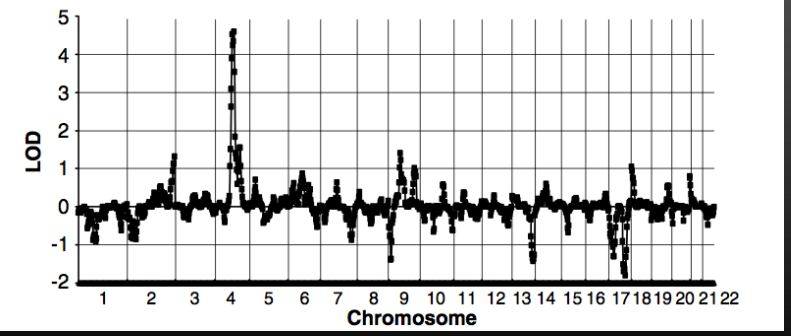
“PI-HAT”



- Pi-Hat represents the probability that QTL i is IBD for family pair j
- Calculated via programs such as Merlin or GeneHunter that analyze pedigree and genotype data.

$$\hat{\pi} = p(IBD2) + \frac{1}{2}p(IBD1)$$

- Weighted likelihood approaches also exist and are more elegant, powerful, and robust to selective sampling but also computationally expensive.



Semi-quantitative linkage analysis of alcohol dependence
484 families, primarily affected sib pairs

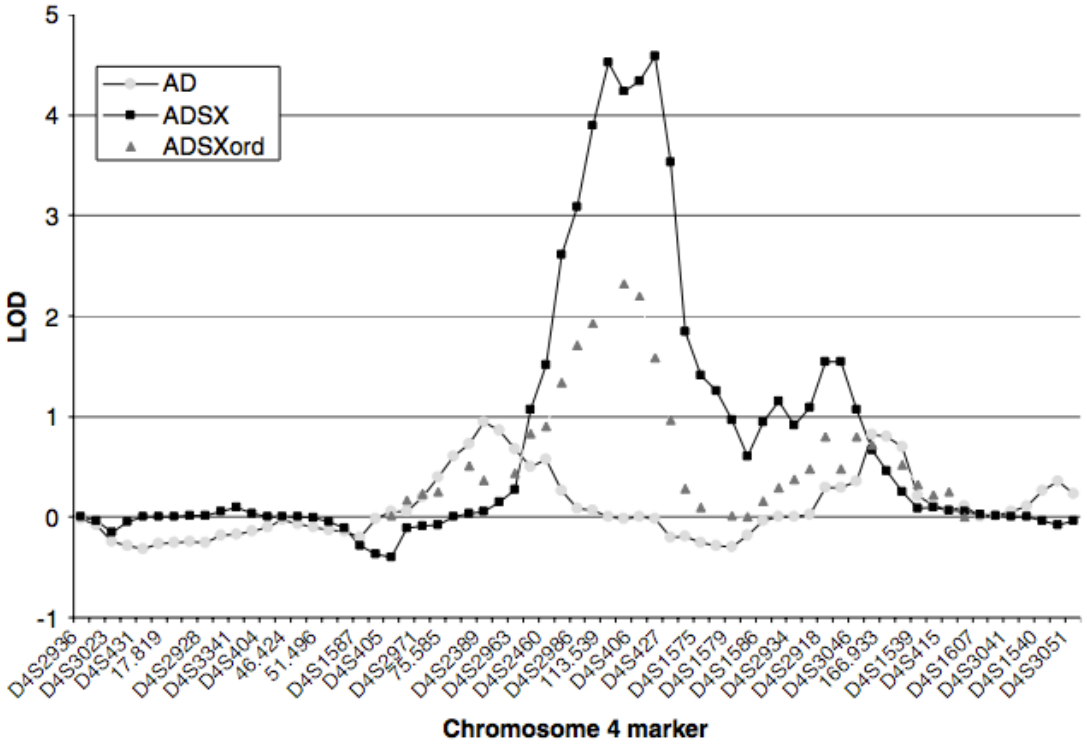


Figure 3 Chromosome 4 NPL LOD scores for alcohol dependence (AD) and AD symptom count analyzed as a continuous (ADSX) and ordinal (ADSXord) variable.

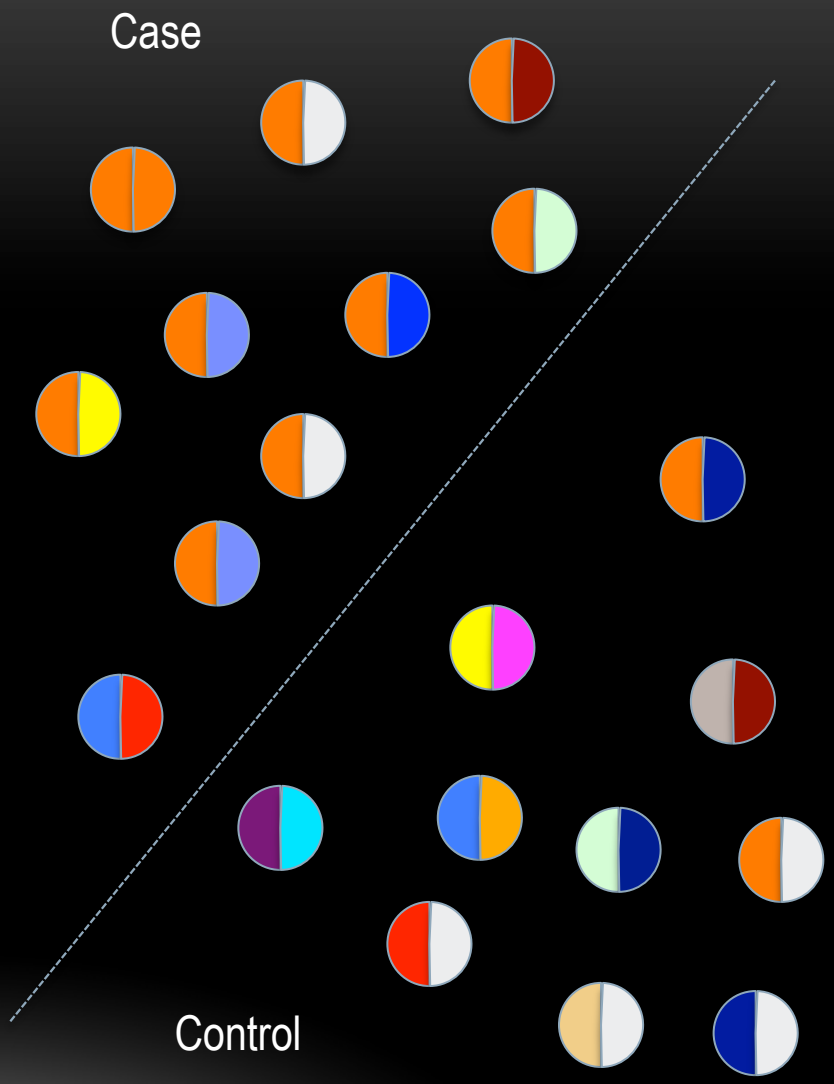
Prescott et al.

Molecular Psychiatry (2006) 11, 603–611

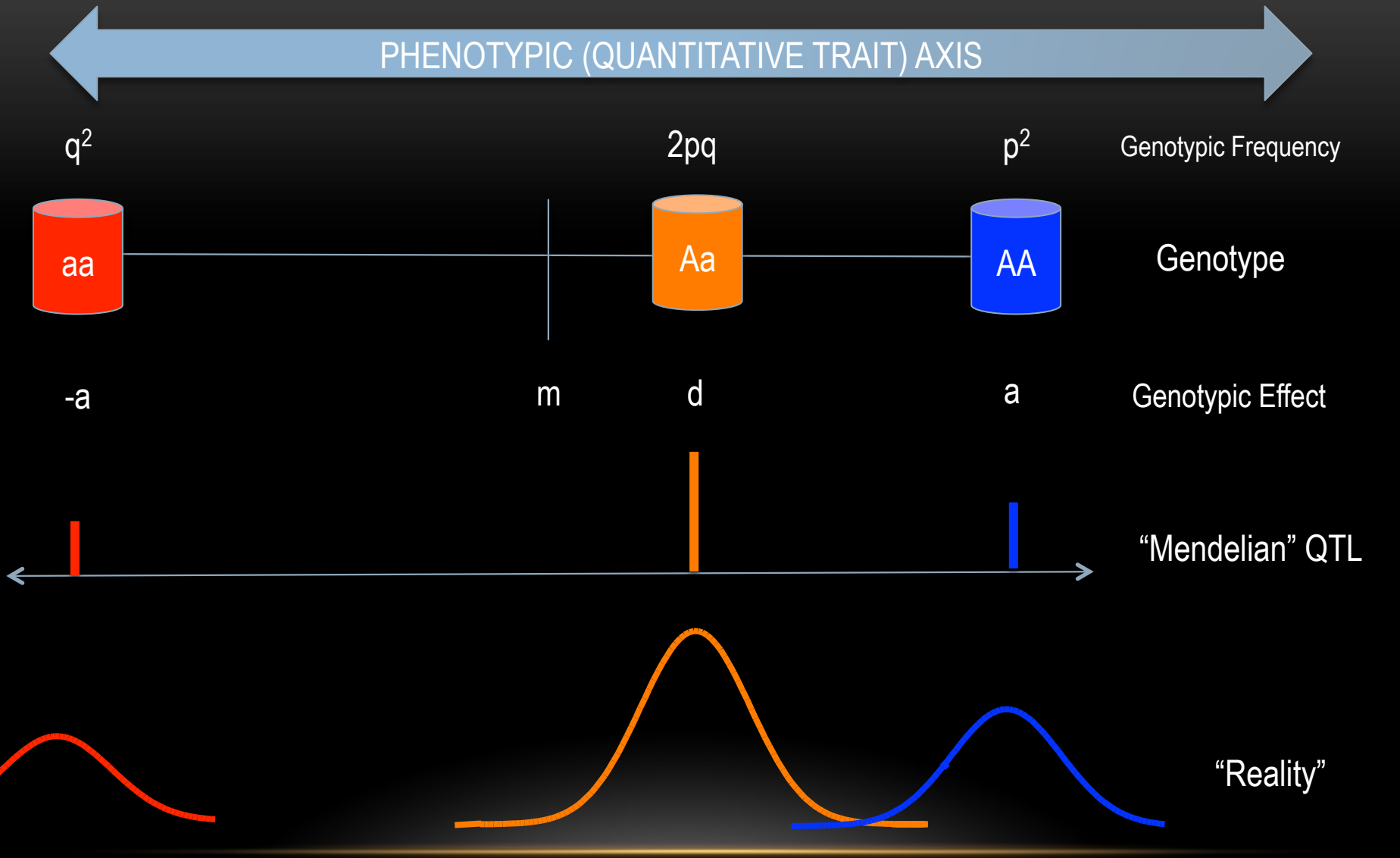
www.nature.com/mp

ASSOCIATION STUDIES

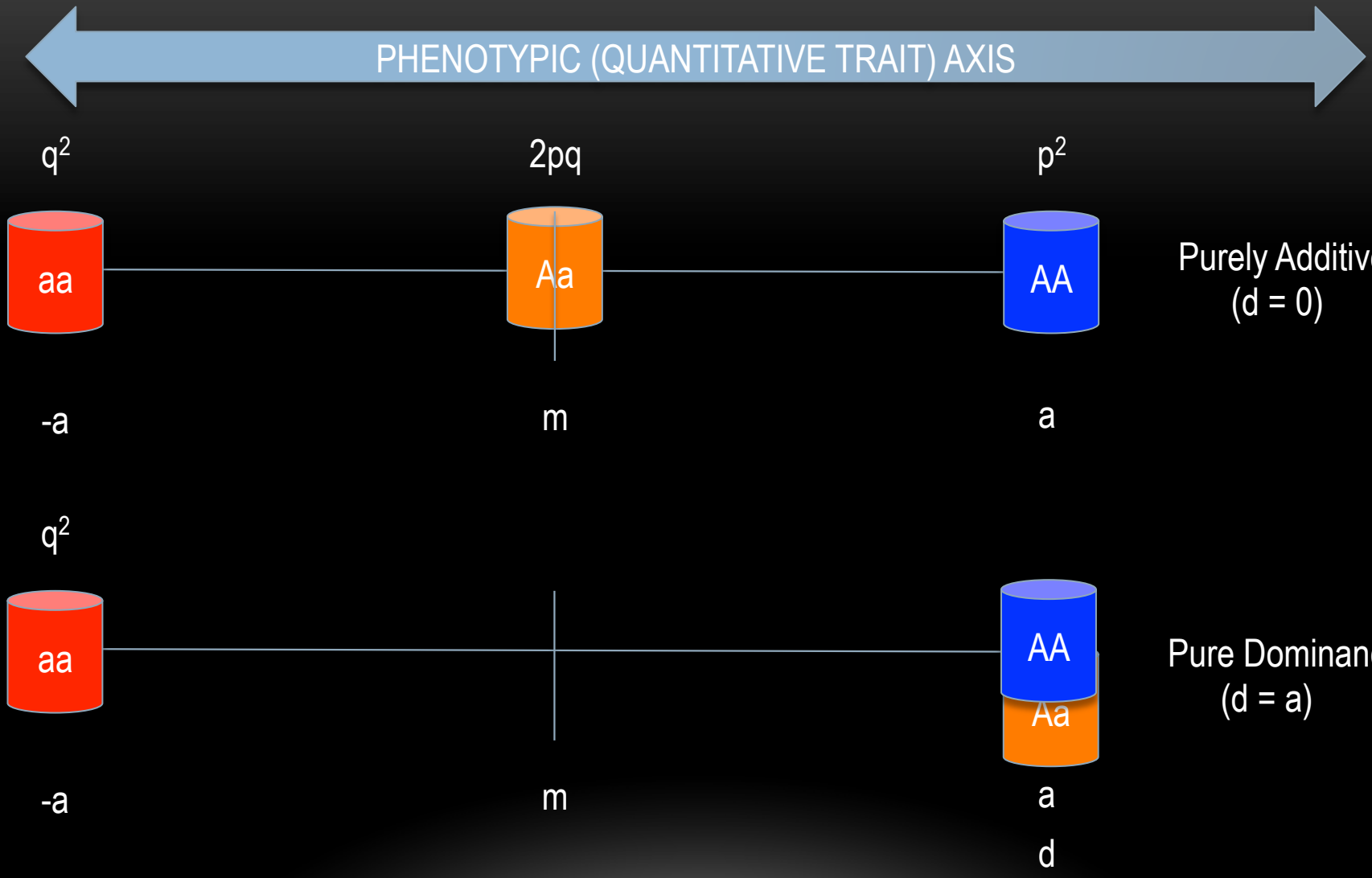
- In their most basic form, case control studies comparing genetic variability at a locus (SNL, RFLP, VNTR, etc. to phenotype.
- While linkage looks for relationships between markers and phenotypes within a family, association looks between families.
- Theoretically powered to detect additive high frequency polymorphisms (common genetic variants)
- Powerful if assumptions are met.
- Family based association studies exist, reduce potential for some spurious results
  - TDT test



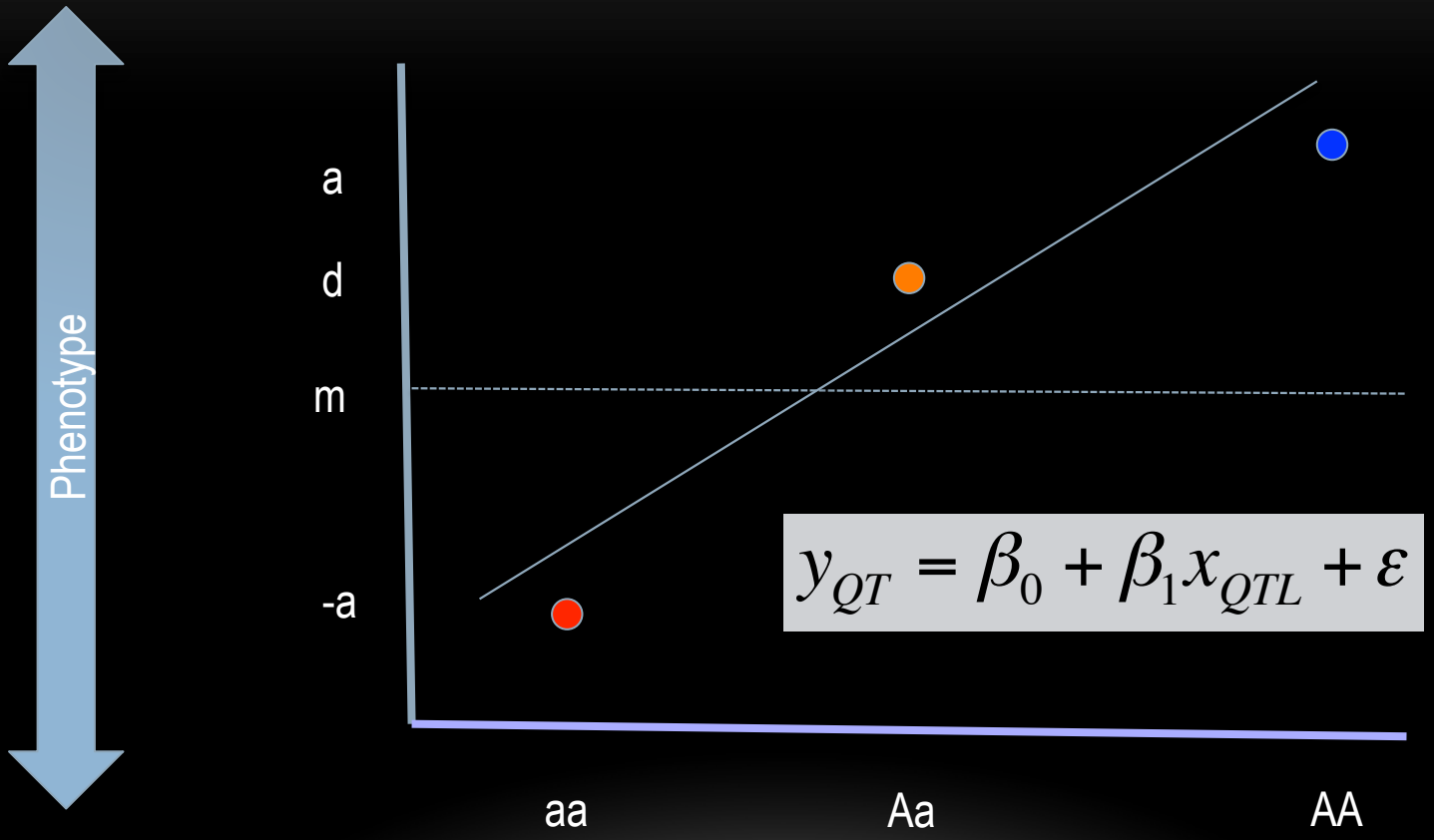
Single Two-Allele Locus



Single Two-Allele Locus

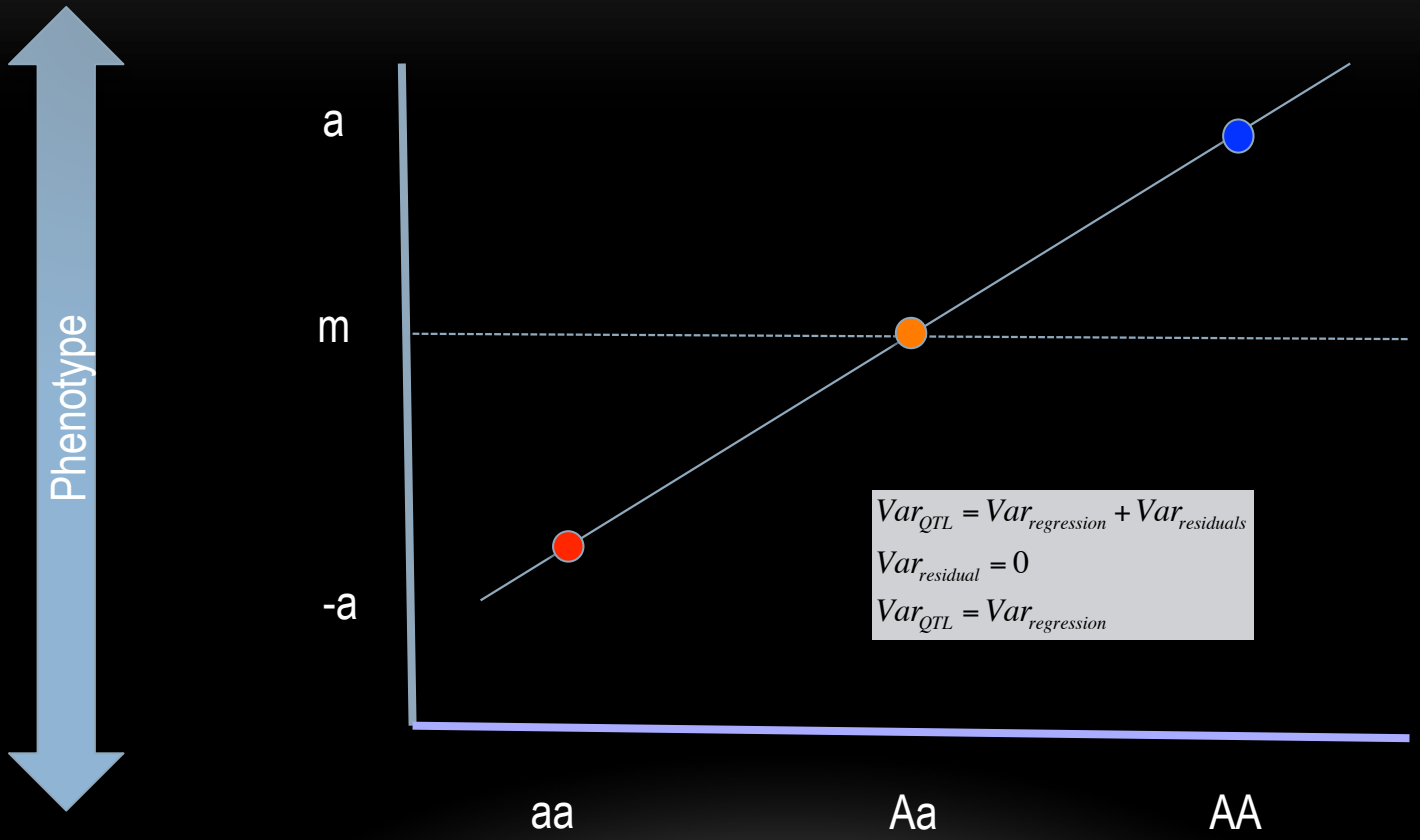


SIMPLE LINEAR REGRESSION MODEL FOR BIALLELIC QTL

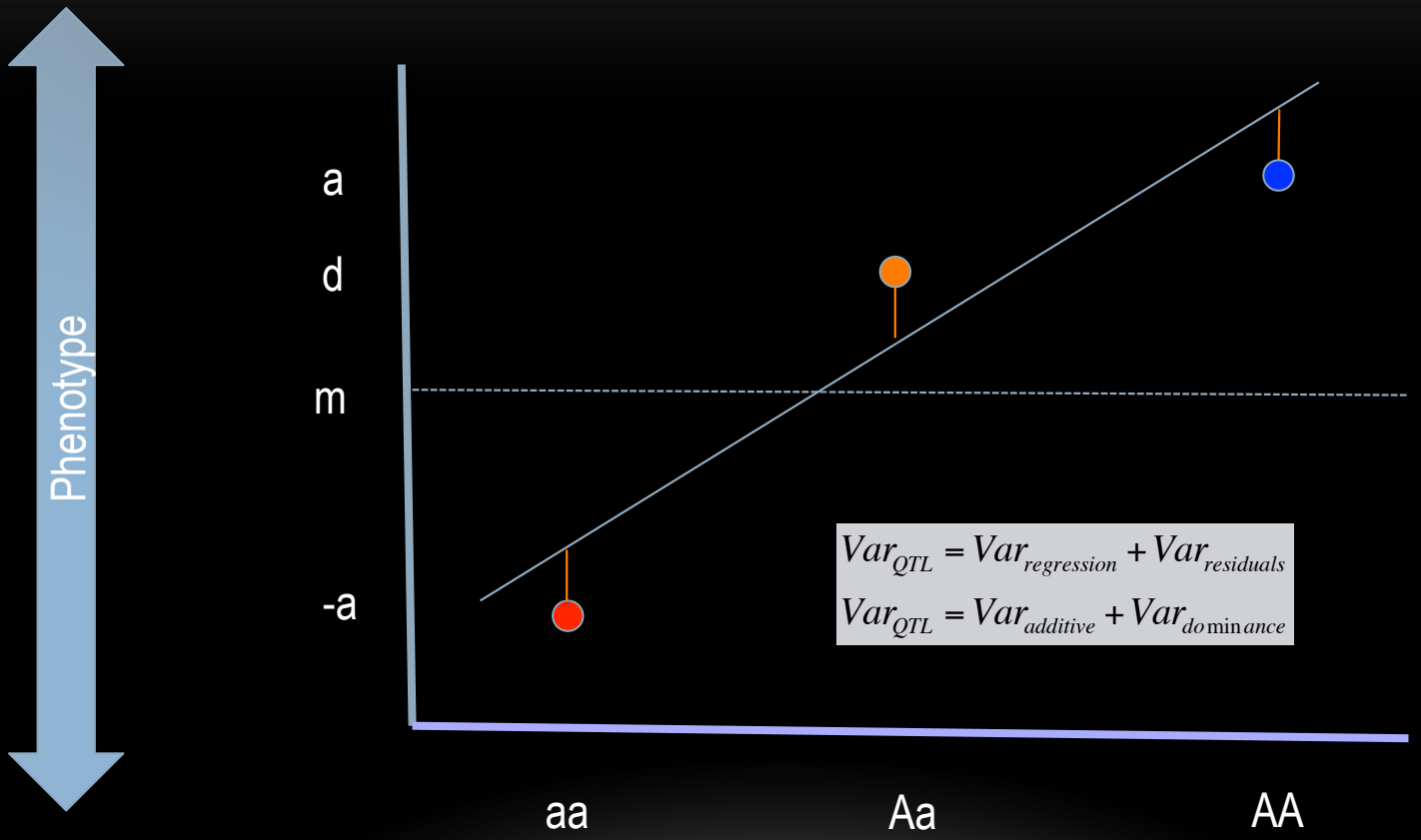


$$y_{QT} = \beta_0 + \beta_1 x_{QTL} + \varepsilon$$

PURELY ADDITIVE QTL

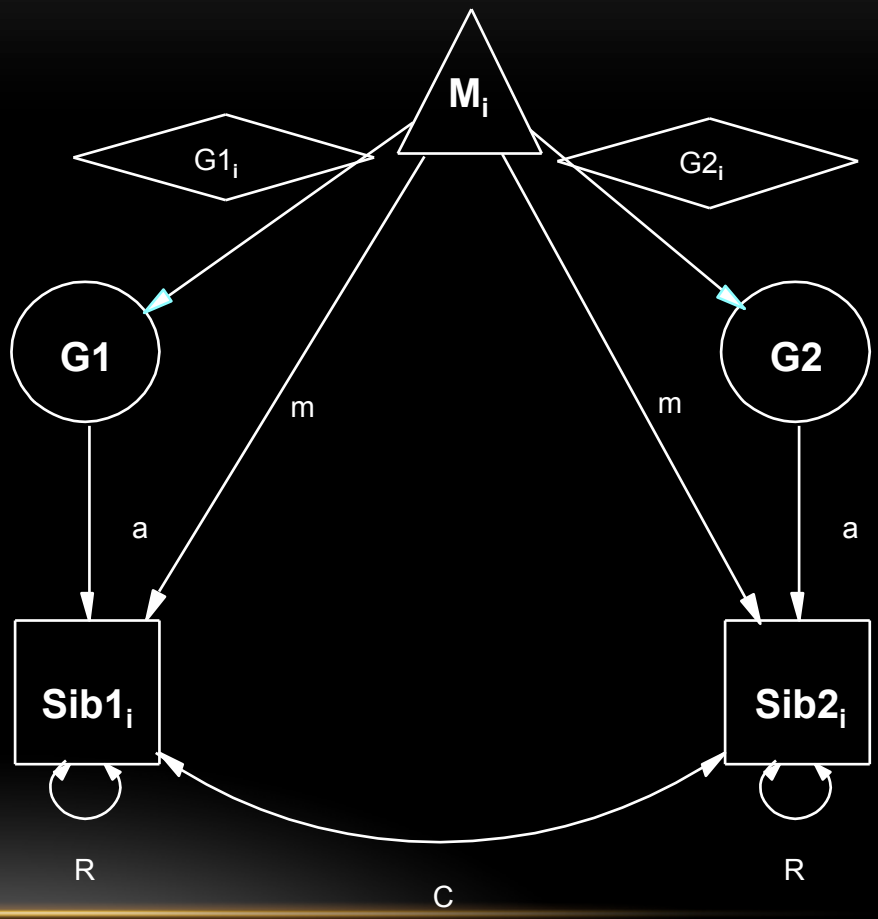


MIXED ADDITIVE AND DOMINANCE QTL

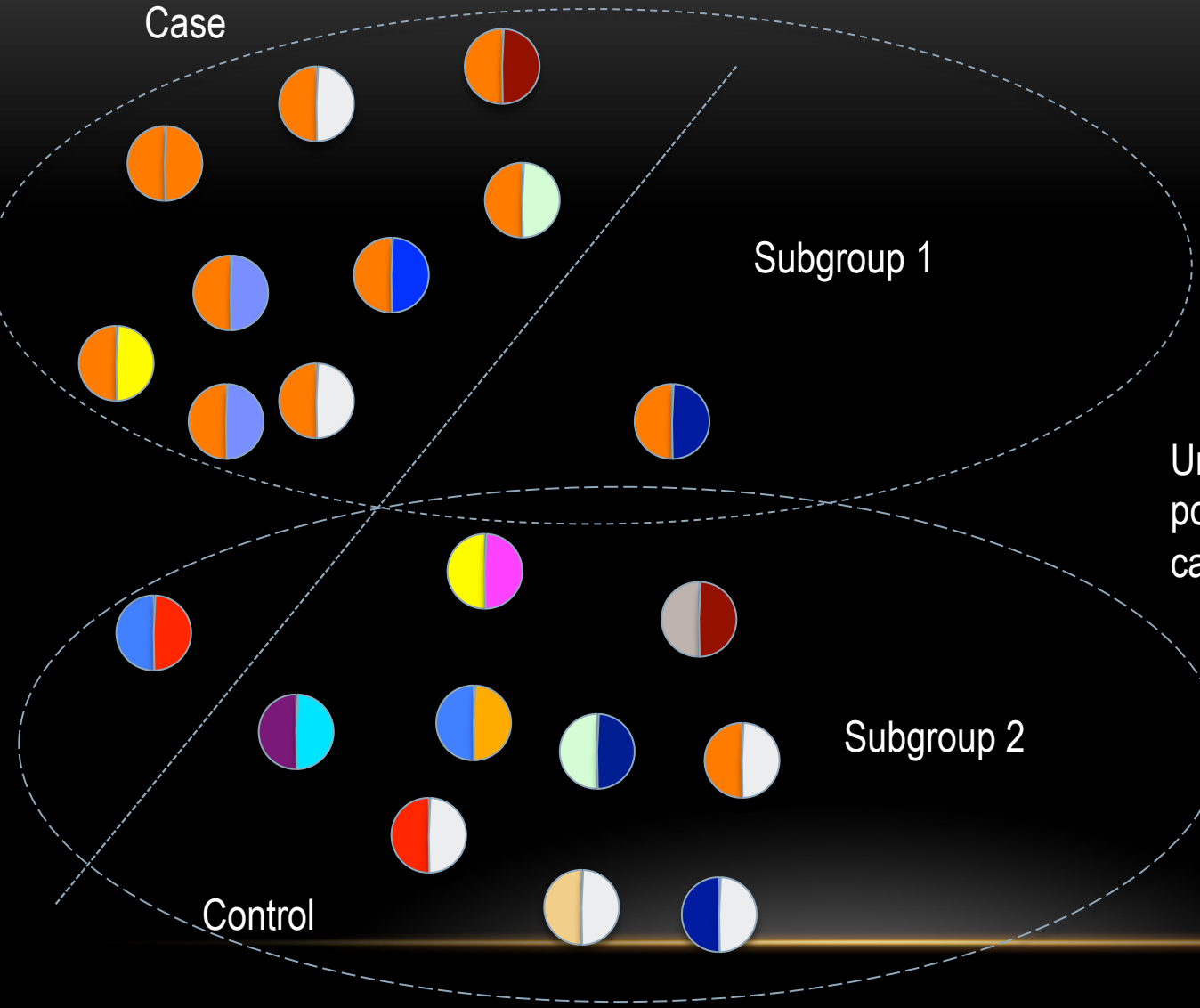


QUANTITATIVE ASSOCIATION MODEL

- SEM equivalent of familial additive QTL association study.
- Differences in genotype modeled as variation from expected mean
- In this case, variance/covariance is not explicitly modeled
- In practice, simple QTL association analyses are usually not performed using SEM, but rather using specially designed software packages
- PLINK: http://pngu.mgh.harvard.edu/~purcell/plink/



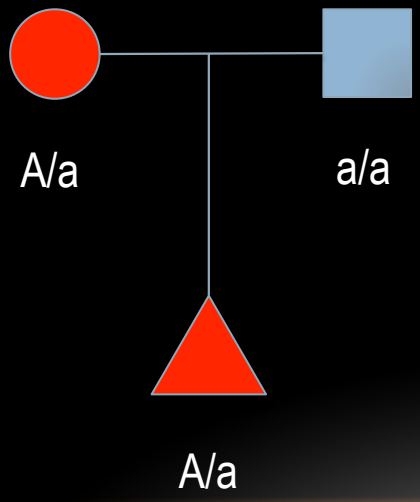
POPULATION STRATIFICATION



Unknown or unaccounted
population structure
can lead to spurious associations

TRANSMISSION DISEQUILIBRIUM (TDT) TEST

- Devised by Spielman based on McNemar's test
- Uses multiple parent sib triplets.
- Controls for population stratification by using parent as an internal control.
- ½ Probability that each allele is transmitted from a heterozygous parent (test is conditional on parental genotype). Generally assumes no segregation distortion.
- Effectively the simplest case of combined linkage and association
- Can be inefficient as many genotypes not informative.
- Original TDT was based on dichotomous traits



- Original TDT extended to quantitative phenotypes by Allison (1997)
- 5 models of varying complexity (most useful is TDTQ5)
  - Regresses quantitative trait on offspring genotypes after controlling for parental mating types.
  - 3 possible offspring genotypes
    - AA, Aa, aa
  - 3 parental mating types of interest
    - Aa/Aa, Aa/AA, Aa/aa

$$y_i = \beta_0 + (\beta_1 genDum1 + \beta_2 genDum2) + (\beta_3 matDum1 + \beta_4 matDum2) + \varepsilon$$

$$vs$$

$$y_i = \beta_0 + (\beta_3 matDum1 + \beta_4 matDum2) + \varepsilon$$

2 df difference between model and submodel

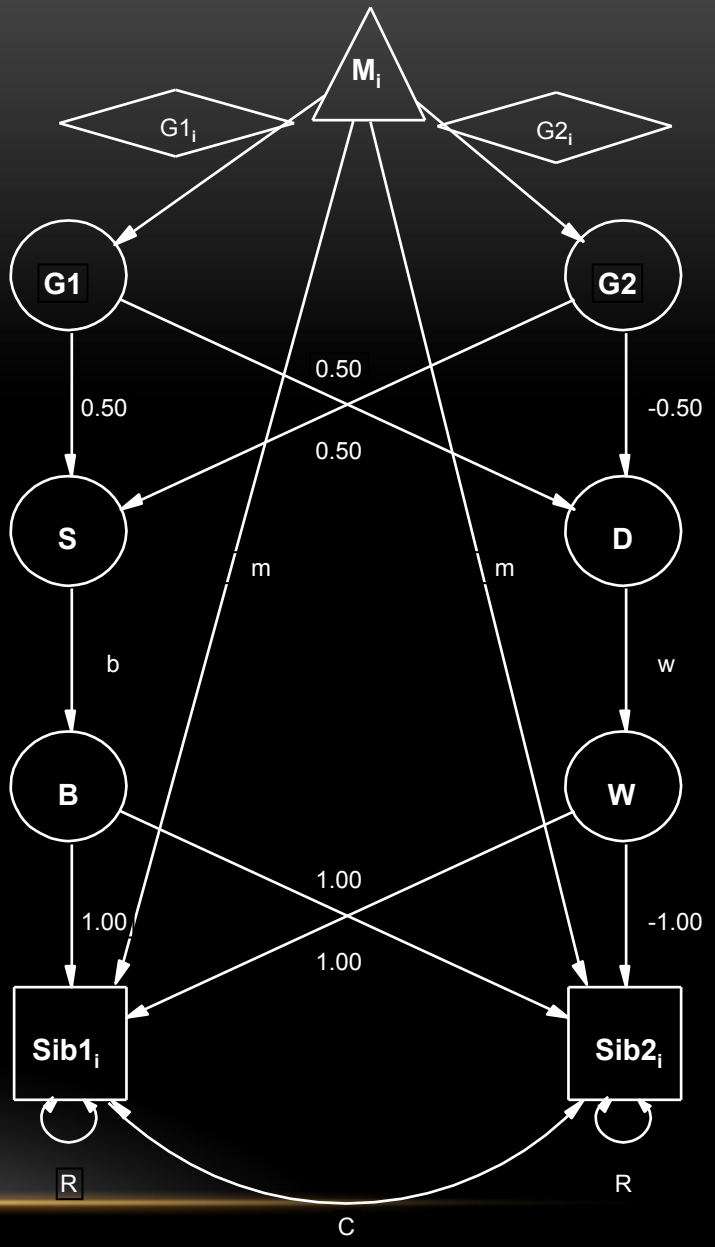
EXTENSIONS/VARIATIONS

TDTQ:

- Additive only Models
- Multipoint/Epistasis
- Multiple sibling Models
- Extended Pedigrees
- Sibling Only Designs
- Multivariate Models
- Imprinting Effects
- Logistic Regression/nonparametric models

FULKER MODEL

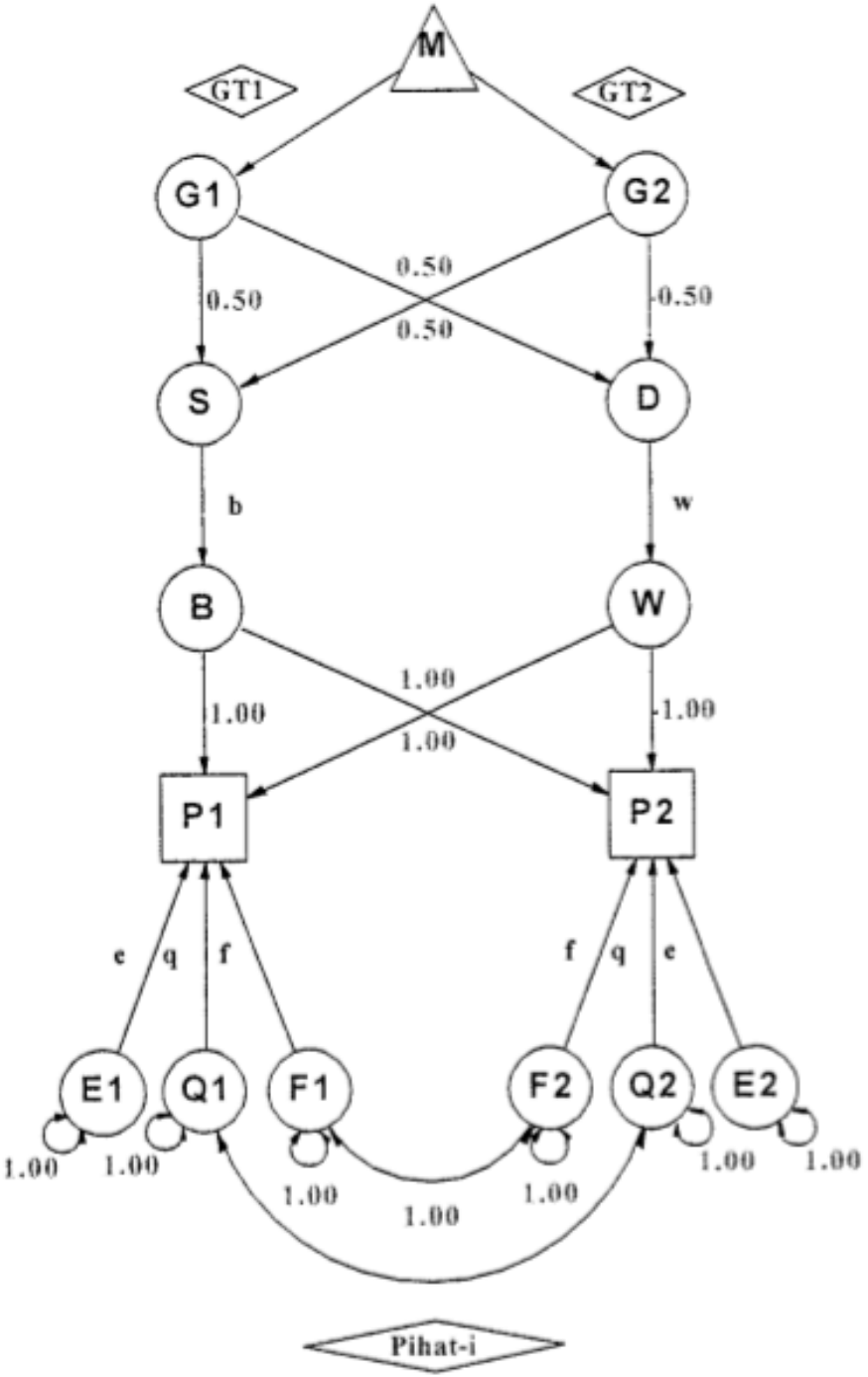
- Controls for population stratification by comparing mean differences between families relative to within families
- Stratification should cause differences between families but not within them
- True allelic effects on the mean should occur independent of between versus within family differences
- B = W when no stratification exists.



COMBINED LINKAGE/ ASSOCIATION

- Linkage and Association analyses are not mutually exclusive and can be complimentary.
- Combining approaches can increase power
- It should be apparent that decomposing familial variance (F) can be added to the model if genetically informative (i.e. twin, pedigree) data is available

- Allison and Neale, *Theoretical Population Biology* 60: 239-251 (2001)



PROS/CONS OF COMMON GENETIC ANALYSIS METHODS

Technique	Pros	Cons
Twin/Family Studies	•Population Based •Assessment of global genetic effects •Better characterize phenotype(s) •Relatively well replicated	•No localization of genetic effects •Equal Environment Assumption
Linkage	•Optimally powered to detect rare penetrant recessive genes with large effect on phenotype. •Relatively few markers needed.	•Relatively low genetic resolution. •Recruitment more challenging.
Association	•Optimally powered to detect additive effects from multiple common genetic variants with small effect size •High spatial resolution. •May detect actual variants responsible for disease.	•Prone to ascertainment bias and population stratification •Many markers needed. •Major multiple testing issues

THE VIRGINIA INSTITUTE FOR PSYCHIATRIC AND BEHAVIORAL GENETICS

Molecular Genetics

- Linkage and Association Scans
  - Depression
  - Schizophrenia
  - Drug use
- Candidate Gene Studies
  - COMT
  - Dysbindin
  - Neuroregulin
  - Etc…
- Methodological improvements
- Functional Classification

Software Development

- Structural Equation Modeling
  - Mx
- Neural Networks for Model finding
- Multistep LD study optimization
- Application of generalized statistical programs to BG problems
  - SAS
  - M-Plus
  - BUGS

Bioinformatics/Stat Gen

- Data Mining Tools
- Microarray Analysis
- Optimization of Sampling for Association Studies
- Strategies to Control Type I Error
- Joint Linkage/Association
- QTL analysis
- Systems Biology

Phenotypes

- Adult Psychopathology
  - Depression
  - Schizophrenia
  - Anxiety Disorders
  - Personality Disorders
  - Eating Disorders
  - Antisocial Behavior
  - Phobia
- Drug Use/Addiction
  - Illicit
  - Licit
- Neurological Dysfunction
  - Epilepsy
- Childhood Psychiatric Disorders
  - Conduct Disorder
  - Depression
  - ADHD
  - others
- Personality and Behavior
  - Neuroticism
  - Religiosity
- Stress
- Sexual Abuse
- Others
  - Cancer
  - Allergy
  - Obesity
  - Irritable Bowel Syndrome
- …

Psychometrics

- Measurement Invariance
- Item Response Theory
- Design and Testing of Mathematical Models for Complex Traits
- Latent Distribution Analyses

Quantitative Genetics

- Nature of Comorbidity
- Family Study Methodology
- Cultural Transmission
- Gene by Environment Interaction
- Epistasis
- rGE
- Gender Effects
- Evolutionary Biology/Psychology
- Extended Phenotypes

"Endophenotypes"

- Structural MRI
  - GAD | Hettema
  - Epilepsy | Corey
  - Child Neurodevelopment |Giedd/NIH
  - Adults and Aging | Kremen/UCSD
  - Infants| Gilmore/UNC
- A few others

CONCLUSIONS:

- Quantitative genetics allows for the measurement of genetic variance on continuous traits of interest from the phenotypic to the molecular level.
  - Quantitative trait as important to the analysis as the genetics.
- Heritability and variance components analyses estimate the contributions of genetic variability on phenotypic variability.
- SEM represents a flexible tool that can address numerous problems in quantitative genetics, neuroimaging, and psychometrics.
- Quantitative trait loci (QTL) represent molecular genetic variants that influence a (continuous or ordinal) trait of interest and can be identified via linkage-based, association-based, or combined approaches.
- Each genetic analysis methodology has its strengths and weaknesses, and, are in general complementary.

ACKNOWLEDGEMENTS

- Mike Neale: VIPBG
- Ken Kendler: VIPBG
- Jay Giedd: NIMH
- Mary Scanlon: Penn
- Children’s Hospital of Pennsylvania