

Heritability of Human Structural Connectomes

Jaewon Chung

(he/him) - [NeuroData lab](#)

Johns Hopkins University - Biomedical Engineering

✉ j1c@jhu.edu

🐙 [@j1c \(Github\)](#)

🐦 [@j1c \(Twitter\)](#)



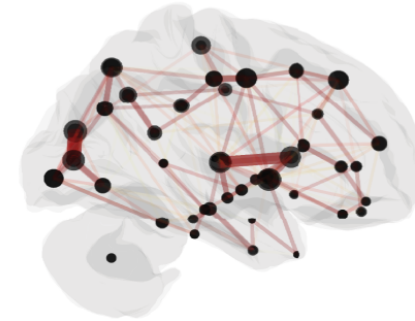
What is heritability?

- Variations in phenotype caused by variations in genotype.
- Potentially discover relationships between diseases and genetics.

Are the brain connectivity patterns heritable?

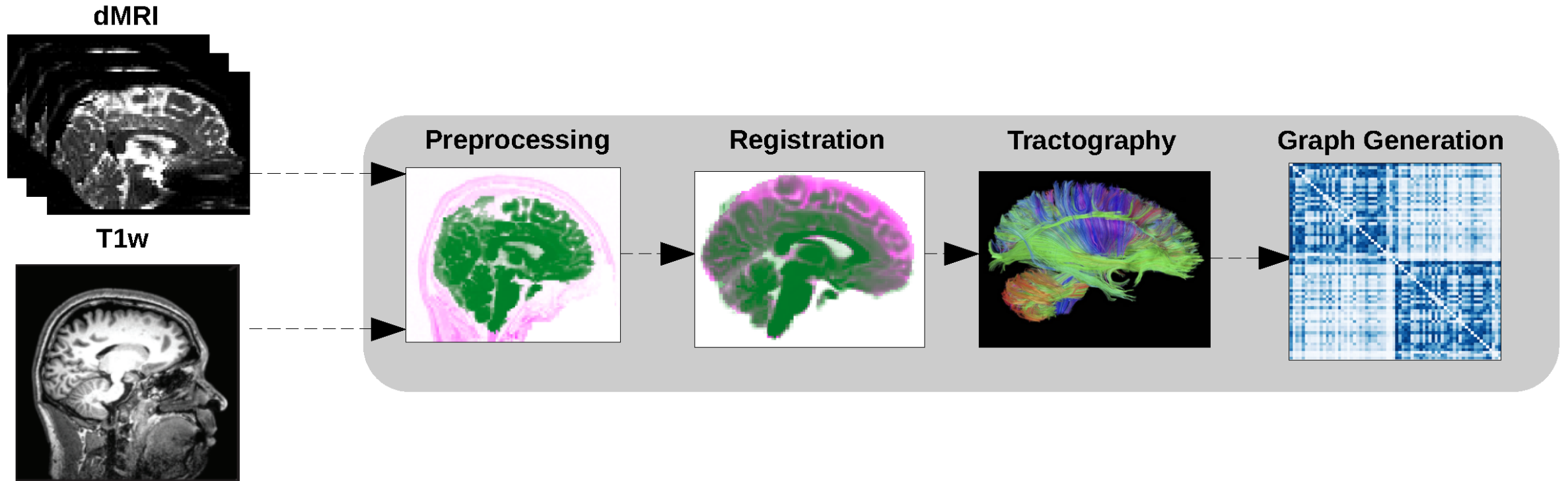
Brain connectivity as connectomes

- Vertex: region of interest
- Edges: connectivity measure between a pair of vertices
- Diffusion MRI: # of estimated neuronal fibers
- Undirected: neurons have no direction

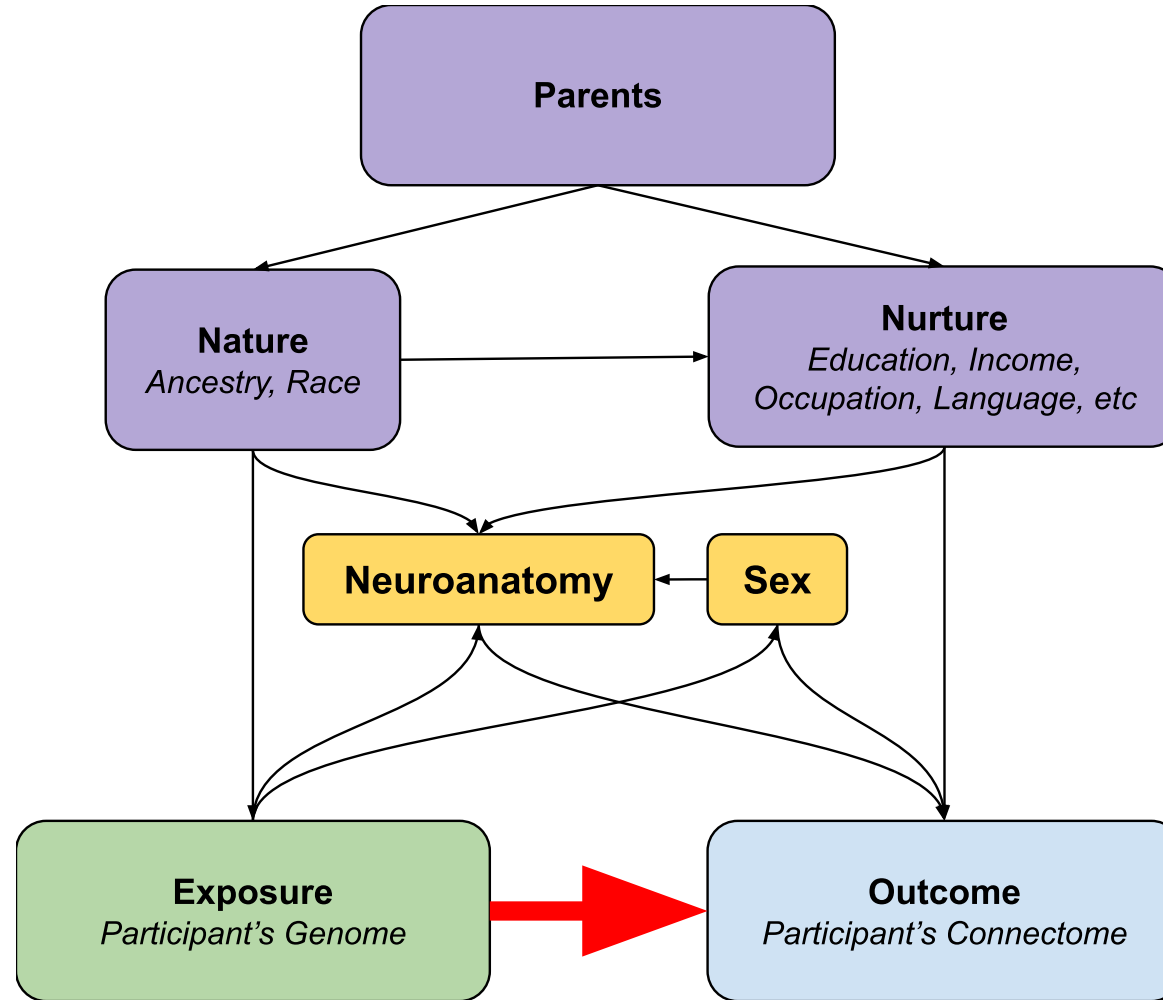


$$\begin{bmatrix} 0 & 3 & 0 & 0 & 1 & 0 \\ 3 & 0 & 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 6 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 & 0 \end{bmatrix}$$

How do we get structural connectomes?



Heritability as causal problem



Do genomes affect connectomes?

- Hypothesis:

$$H_0 : F(\text{Connectome}|\text{Genome}) = F(\text{Connectome})$$

$$H_A : F(\text{Connectome}|\text{Genome}) \neq F(\text{Connectome})$$

- Alternatively:

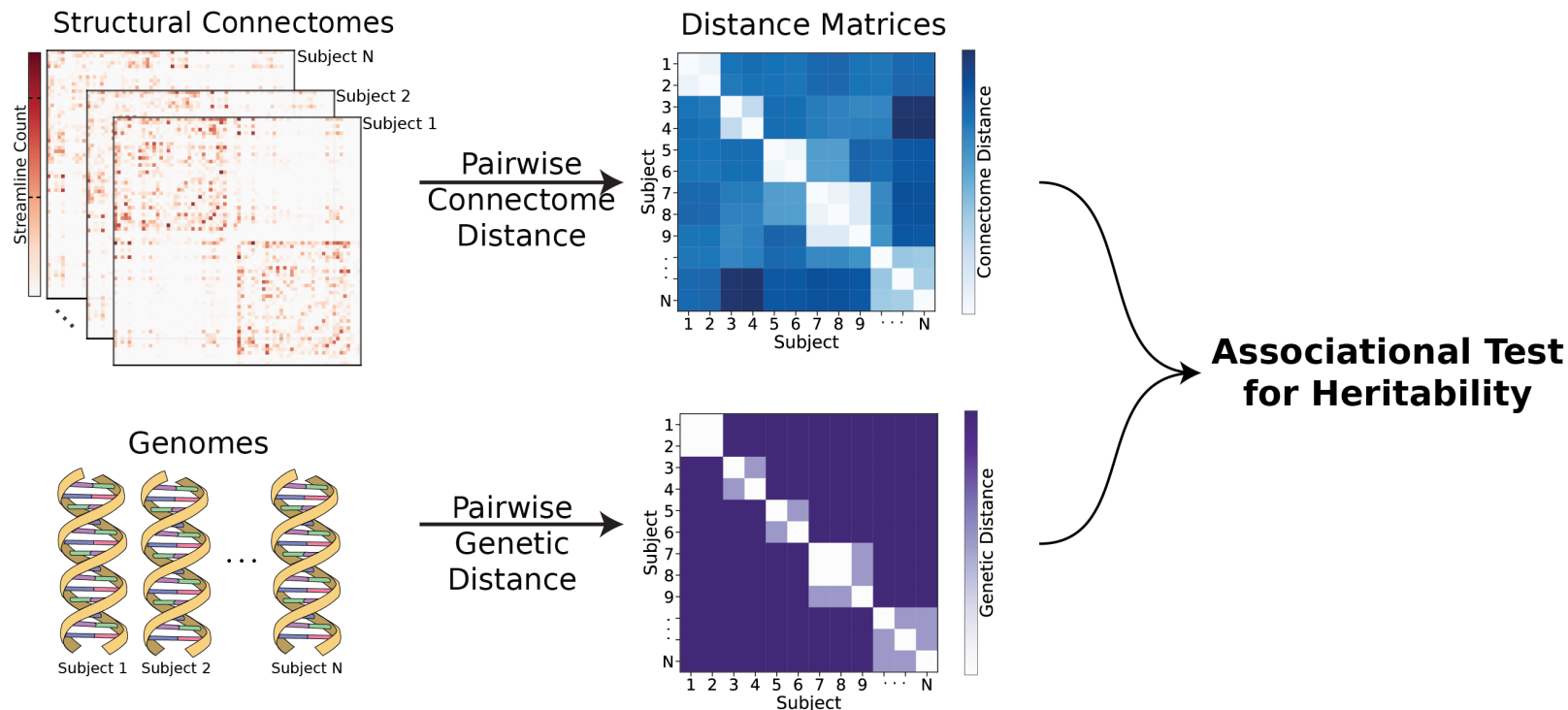
$$H_0 : F(\text{Connectome}, \text{Genome}) = F(\text{Connectome})F(\text{Genome})$$

$$H_A : F(\text{Connectome}, \text{Genome}) \neq F(\text{Connectome})F(\text{Genome})$$

- Known as independence testing
- Test statistic: *distance correlation* (*dcorr*)
- Implication if true: there exists an associational heritability.

What is distance correlation?

- Measures dependence between two multivariate quantities.
 - For example: connectomes, genomes.
- Can detect nonlinear associations.
- Measures correlation between pairwise distances.



How to compare genomes?

- Typical twin studies do not sequence genomes.
- Coefficient of kinship (ϕ_{ij})
 - Probabilities of finding particular genes as identical among subjects.

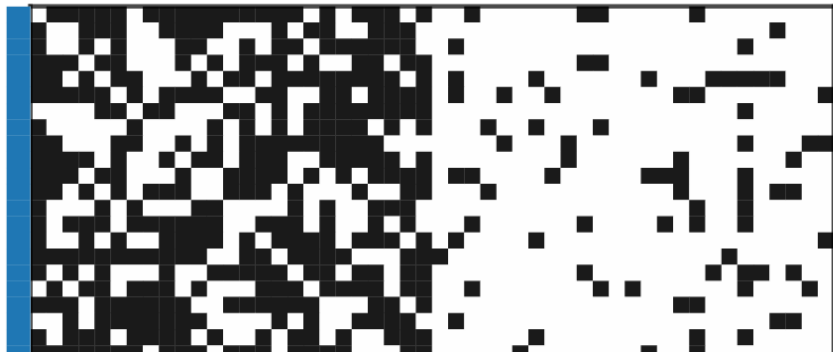
Table here

- $d(\text{Genome}_i, \text{Genome}_j) = 1 - 2\phi_{ij}$.

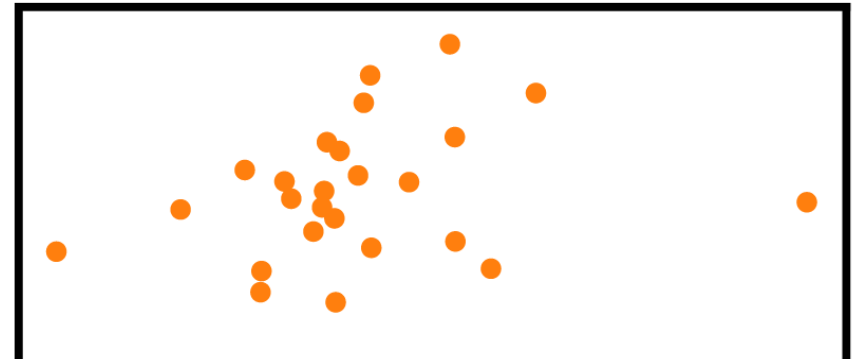
How to compare connectomes?

- Use statistical models for networks!
- Random dot product graph (RDPG) - each vertex
 - Probability from dot product of source node's latent vector, target node's latent vector.
 - x_i, x_j in d dimensions
- $d(\text{Connectome}_i, \text{Connectome}_j) = ||X^{(i)} - X^{(j)} R||_F$

Adjacency Matrix



Estimated Latent positions



Adjacency Spectral

Human Connectome Project

- Brain scans from identical (monozygotic), fraternal (dizygotic), non-twin siblings.
- Which parcellation?

Zygosit y	Monozygoti c	Dizygotic	Non-twin siblings
N	322	212	490
Sex	196 F, 126 M	125 F, 87 M	237 F, 253 M
Age (mean)	29.6 (3.3)	28.9 (3.4)	28.3 (3.9)
Age (range)	22-36	22-36	22-37

Monozygotic vs Dizygotic

- Assumptions:
 - Controls environment variable

Insert figure

All three groups

- Assumptions:
 - Add in environmental and genetic variance

Insert figure

Neuroanatomy (effect mediator)

- Literature show neuroanatomy (e.g. brain volume) is highly heritable.
- Want to test:

$$H_0 : F(\text{Neuroanatomy}, \text{Genome}) = F(\text{Neuroanatomy})F(\text{Genome})$$

$$H_A : F(\text{Neuroanatomy}, \text{Genome}) \neq F(\text{Neuroanatomy})F(\text{Genome})$$

New DAG

One with neuroanatomy

Statistical problem

- Want a conditional independence test!

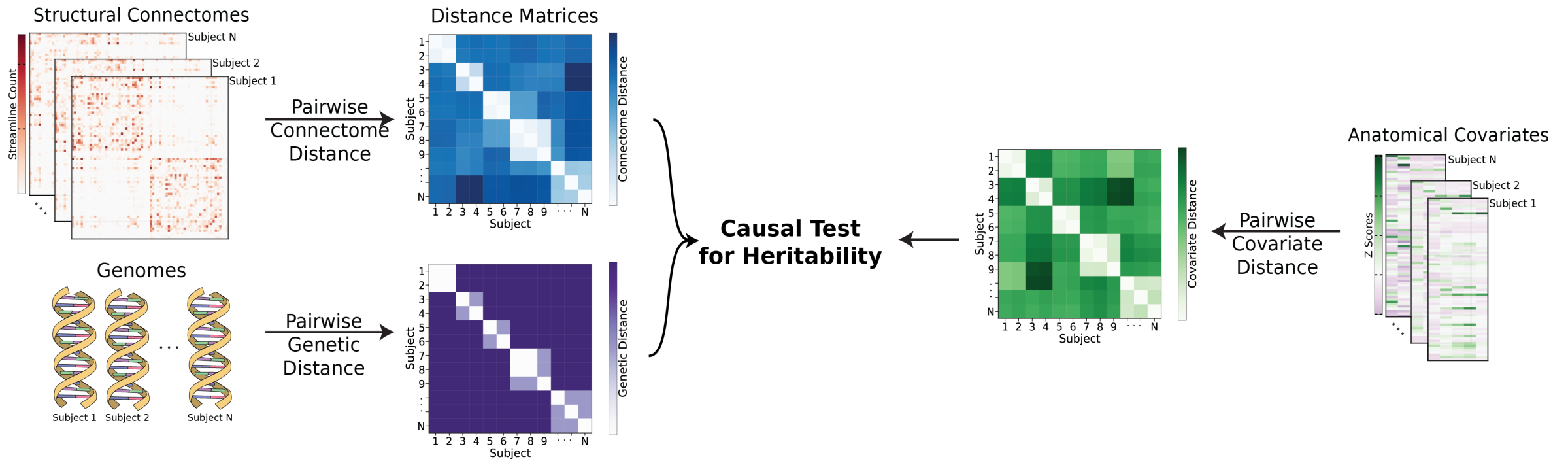
$$H_0 : F(\text{Conn.}, \text{Genome} | \text{Neuro.}) = F(\text{Conn.} | \text{Neuro.}) F(\text{Genome} | \text{Neuro.})$$

$$H_A : F(\text{Conn.}, \text{Genome} | \text{Neuro.}) \neq F(\text{Conn.} | \text{Neuro.}) F(\text{Genome} | \text{Neuro.})$$

- Test statistic: Conditional distance correlation (cdcorr)

What is conditional distance correlation?

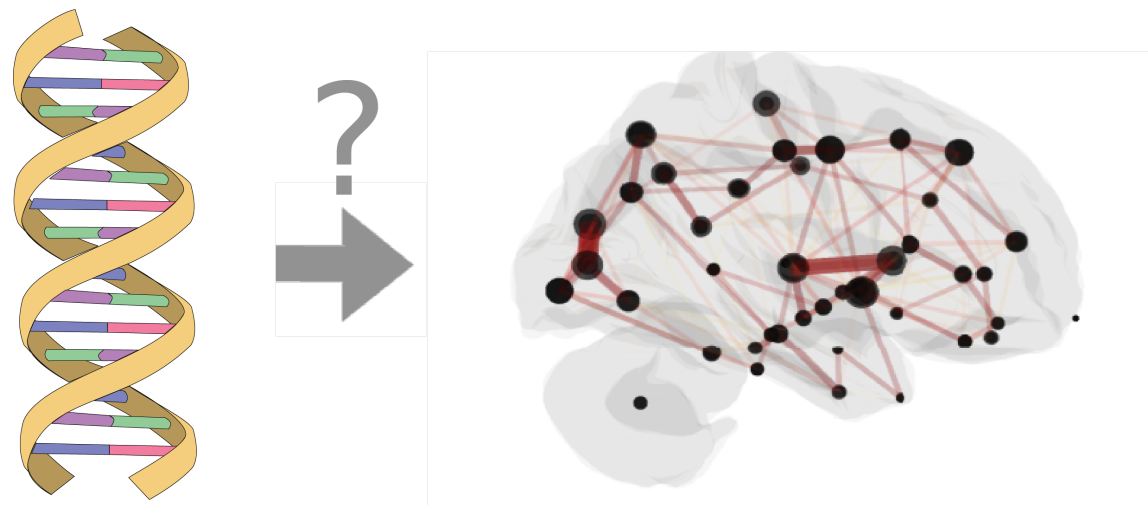
- Augment distance correlation procedure with third distance matrix.
- $d(\text{Neuroanatomy}_i, \text{Neuroanatomy}_j) = \|\text{Neuroanatomy}_i - \text{Neuroanatomy}_j\|_F$



Conditional Test as causal effect estimator

-

Summary



- Present a causal model for heritability of connectomes.
- Leveraged recent advances:
 - i. Statistical models for networks, allowing meaningful comparison of connectomes.
 - ii. Distance and conditional distance correlation as test statistic for causal analysis¹.

¹ Bridgeford, Eric W., et al. "Batch Effects are Causal Effects: Applications in Human Connectomics." (2021).

Acknowledgements

Team



Mike Powell



Eric Bridgeford



Carey Priebe

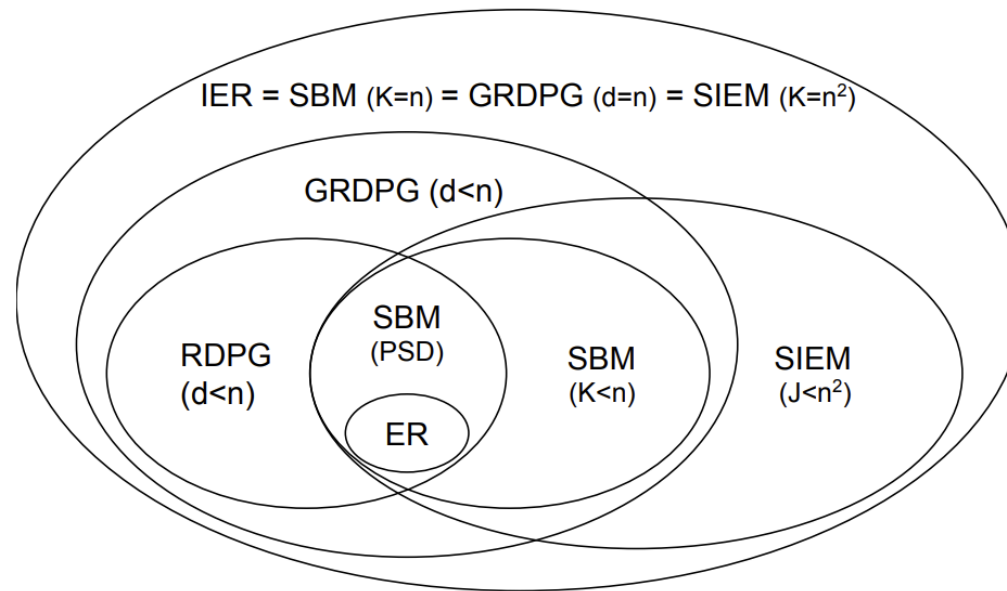


Joshua Vogelstein

Additional slides

Shortcomings - Network model

- Problems with connectome estimation.
 - Inability to determine the precise origin/termination of connections in the cortex.
 - -> false negatives
 - Crossing fibers
 - -> false positives
- RDPG can only represent subset of independent edge networks.



Shortcomings - Model assumptions

- No interaction between genome and environment
- No epistasis
 - Effect of one gene is dependent on another
 - Ex: black hair and baldness
- No dominance effects
- Strong assumptions in genetic distances

What are environmental effects?

- Shared
 - Common experiences of siblings living in the same household.
 - household income, the family's living situation, the dynamics between the parents, food consumed
- Non-shared
 - Everything else
 - Epigenetics
 - Luck
 - schools, peers

Random dot product graphs

