Effect of Rare Variation on Human Brain Structure

Carrie E. Bearden, Ph.D.
Psychiatry and Biobehavioral Sciences and Psychology
Sæmel Institute for Neuroscience and Human Behavior
University of California, Los Angeles

cbearden@mednet.ucla.edu

OHBM 2018, Singapore
I have no relevant financial disclosures or conflicts of interest
OUTLINE

- Definition of Terms – what do we mean by ‘rare variation’ and how do we assess it?

- Analysis of Copy Number Variants (CNV)

- Neuropsychiatric CNV examples – gene dosage effects

- Single Nucleotide Variants (SNVs)- exome /whole genome sequencing

- SNV examples – classes of pathogenic mutations

- Functional genomics /methods – Weighted Gene Co-Expression Network Analysis (WGCNA)
Common vs. Rare Genetic Variation

Minor Allele Frequency: frequency at which the least common allele occurs in population

Common variants: $\text{MAF} \geq 5\%$
Low frequency variants: $0.5\% < \text{MAF} \leq 5\%$
Rare variants: $\text{MAF} \leq 0.5\%$

GWAS - identification of common variants

But actually..

Most of human genetic variants are rare!
Landscape of Genetic Variation and Methods

(McCarthy, Nat Review Genet, 2008)
RGDs can result from alteration of single nucleotides in one gene, entire chromosomes with 100s of genes, or anywhere in between.

- Neuropsychiatric symptoms documented in 74% of RGDs
- Largest contribution to neuropsychiatric dx w/ high-heritability, early age-of-onset, impaired cognition, and/or reduced fecundity (Power et al JAMA Psych 2013)

OMIM: Online Mendelian Inheritance in Man; G2P: Genotype to Phenotype; HPO: Human Phenotype Ontology
Copy Number Variation

- Structural variation in the genome involving unbalanced rearrangements of DNA segments (≥50 bp) - microdeletions/duplications; number of repeats varies between individuals in the human population (Iafrate et al 2004; Zarrei et al Nat Rev Gen 2015)

- Benign vs. pathogenic CNVs? (Riggs et al, 2014)

- Specific genomic regions more likely to harbor CNVs (“hotspots”)

- Special class of CNVs - rearrangement breakpoints mediated by stretches of repetitive DNA - nonallelic homologous recombination; ‘genomic disorders’ (Lupski TIG 1998, Hastings et al. Nat Rev Genetics 2009)

Proportion of chromosome that is copy number variable
Continuous spectrum of phenotypic effects of CNVs:
– no phenotypic effect
– account for adaptive traits
– can underlie disease; many have neuropsychiatric phenotype

De Novo vs Inherited – de novo likely more deleterious (not subject to evolutionary pressure)

1, Moreno-De-Luca & Cubells (2011); 2, Merikangas et al (2009); 3, Malhotra & Sebat (2012)
Neuropsychiatric Phenotypes Associated With CNVs

Disease-Based Studies of Genetic Mutations
- Adult/late-onset diseases
  - Schizophrenia
  - Bipolar disorder
- Pediatric/developmental diseases
  - Epilepsy
  - Autism
  - Mental retardation
  - Birth defects

CNV-Based Studies of Expression in Families and General Population Samples
- Severe phenotypes (core)
- Observable if assessed adequately
- Mutation carrier (no observable phenotype)

Discovery of rare CNV associated with disease

Core phenotypes differ for specific CNVs
- Schizophrenia
- Epilepsy
- Autism
- Mental retardation
- Birth defects

22q11.2 deletion
15q13.3 deletion
1q21.1 deletion

Bassett et al  AJP 2010
Majority of cases of developmental neuropsychiatric dx of unknown etiology; But rare genomic copy number variants (CNVs) may account for a larger proportion of cases than previously believed

- ~2% of schizophrenia cases (Levinson et al. 2011; Bassett et al. 2010; Rees et al. 2014);
- 20% of cases of simplex ASD (Geschwind 2015, Devlin & Scherer 2012)

Also provide strong evidence for genetic pleiotropy, challenging widely held views of diagnostic ‘categories’
Lines of Evidence for CNV Contribution to Schizophrenia

- CNVs contribute to risk for SCZ: genome-wide enrichment of rare deletions and duplications in SCZ cases vs controls (Walsh et al Science 2008; ISC Nature 2008)
- Higher rate of de novo CNVs in cases relative to controls (Malhotra et al Neuron 2011; Xu et al Nat Genetics 2008; Kirov et al Molec Psych 2012)
- Association evidence implicating a small number of specific loci
- All CNVs implicated in SCZ are rare in the population but confer significant risk (ORs 2–60).
Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects

- Overall CNV burden sig. greater among SCZ cases: total distance (kb) covered (OR = 1.12, P = 5.7x 10\(^{-15}\)), genes affected (OR = 1.21, P = 6.6x10\(^{-21}\)) or CNV number (OR = 1.03, P = 1 \times 10^{-3})

- Genome-wide significant evidence for 1q21.1, 2p16.3 (NRXN1), 3q29, 7q11.2, 15q13.3, distal/proximal 16p11.2, and 22q11.2

- CNV burden enriched for genes associated with synaptic function (OR = 1.68, P = 2.8x 10\(^{-11}\)), neurobehavioral phenotypes in mouse (OR = 1.18, P = 7.3 \times 10^{-5})

Marshall et al Nat Genetics 2017
Reciprocal recurrent CNVs at the 16p11.2 locus

4 rearrangements at the 16p11.2 locus

29 genes at 16p11.2 BP4-5

Zufferey et al (2012)
Disorders associated with the 16p11.2 CNV

**DELETION**
1/2000 individuals

- Macrocephaly
- Obesity¹

-2 SD lower Intelligent Quotient than controls⁴,⁵

- Autism Spectrum Disorder²
  10-fold increased risk

**DUPLICATION**
1/2000 individuals

- Microcephaly
- Underweight¹

-2 SD lower Intelligent Quotient than controls⁴,⁵

- Autism Spectrum Disorder²
  10-fold increased risk

- Schizophrenia³
  10-fold increased risk

A novel highly-penetrant form of obesity due to microdeletions on chromosome 16p11.2


Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus

Sébastien Jacquemont, Alexandre Reymond, Flore Zufferey, Louise Harewood, Robin G. Walters, Zoltán Kutalik, Danielle Martinet, Yiping Shen, Armand Valsesia, Noam D. Beckmann, Gudmar Thorleifsson, Marco Belfiore, Sonia

Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities

Published in final edited form as:

Opposing Brain Differences in 16p11.2 Deletion and Duplication Carriers


Neurobiology of Disease

Dose-Dependent Brain and Cognition Alterations in 15q11.2 CNV Carriers

Key Questions

➢ Reciprocal CNVs’: Investigation of Gene Dosage Effects

➢ Can we use these ‘genetics first’ models to probe circuit-level dysfunction translationally?

➢ Is there overlap between the genetic and neurobiological mechanisms underlying development of psychosis in a highly penetrant CNV and idiopathic illness?

➢ Mechanisms underlying divergent outcomes of neuropsychiatric CNV’s?
22q11.2 Microdeletion Syndrome

- Velocardiofacial/DiGeorge Syndrome
- Estimated incidence -~ 1/4000 live births (Grati et al 2015)
- Results from hemizygous deletion of chromosome 22q11 (~3 Mb)
- Cardiac defects, immune deficiency, craniofacial anomalies
- Morbid risk of schizophrenia ~ 25x the general population risk (Murphy et al. 1999; Bassett et al. 2010); account for ~1% of sporadic SZ cases (Stefansson et al., 2008; Karayiorgou et al. 2010)
- Elevated rates of Autistic Spectrum Disorder (12-50%), ADHD (33-40%) and anxiety disorder (40-50%) in childhood (Schneider et al. 2014; Richards et al. 2014)

CNV and SCZ Working Group of PGC, Nat Gen 2016

Kobrynski Lancet 2007; Kapadia & Bassett, CMAJ 2008
Machine Learning Classification of 22q11DS Case-Control Status

- GLMNET algorithm (20 10-fold cross-validations)
- Feature selection based on Random Forest algorithm

Accuracy peaks at 94.44% (125 features)
- Cuneus & lingual SA, and caudal anterior cingulate volume among most important features for classification

Accuracy of 86.36% achieved using bilateral cuneus and lingual SA for classification
Effects of Deletion Size on Cognition (n=1353)

Full Scale, Verbal and Performance IQ decreased by 0.3-0.5 SD in AD vs. AB deletions

Zhao et al AJMG in press
Are there underlying brain structural differences?

- Small deletion associated with larger cortical surface area (p=.003)

Sun et al, Molecular Psychiatry in press
IMMEDIATE COMMUNICATION

Evidence that duplications of 22q11.2 protect against schizophrenia

E Rees¹, G Kirov¹, A Sanders²,³, JTR Walters¹, KD Chambert⁴, J Shi⁵, J Szatkiewicz⁶, C O’Dushlaine⁶, AL Richards¹, EK Green¹,², I Jones¹, G Davies¹, SE Legge¹, JL Moran⁷, C Pato⁸, M Pato⁹, G Genovese⁸, D Levinson¹, J Duan¹,², W Moy⁷, HHH Göring¹⁰, D Morris¹¹, P Cormican¹¹, KS Kendler¹², FA O’Neill¹³, B Riley¹⁵, M Gill¹¹, A Corvin¹⁴, Wellcome Trust Case Control Consortium¹⁵, N Craddock¹, P Sklar¹, C Hultman¹⁵, PF Sullivan¹⁶,¹⁷,¹⁸, PV Gejman²,³, SA McCarroll²,³, MC O’Donovan¹ and MJ Owen¹

A number of large, rare copy number variants (CNVs) are deleterious for neurodevelopmental disorders, but large, rare, protective CNVs have not been reported for such phenotypes. Here we show in a CNV analysis of 47 005 individuals, the largest CNV analysis of schizophrenia to date, that large duplications (1.5–3.0 Mb) at 22q11.2—the reciprocal of the well-known, risk-inducing deletion of this locus—are substantially less common in schizophrenia cases than in the general population (0.014% vs 0.085%, OR = 0.17, P = 0.00086). 22q11.2 duplications represent the first putative protective mutation for schizophrenia.

Molecular Psychiatry (2014) 19, 37–40; doi:10.1038/mp.2013.156; published online 12 November 2013

Keywords: 22q11.2; CNV; duplication; protective; schizophrenia

---

<table>
<thead>
<tr>
<th>Study</th>
<th>Case 22q11.2dup frequency (N CNVs/N samples)</th>
<th>Control 22q11.2dup frequency (N CNVs/N samples)</th>
<th>P value (Fisher’s exact test)</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery</td>
<td>0% (0/6 882)</td>
<td>0.089% (10/11 255)</td>
<td>0.017 (2-Tail)</td>
<td></td>
</tr>
<tr>
<td>Replication</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGS EA</td>
<td>0.090% (2/2 215)</td>
<td>0.16% (4/2 556)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGS AA</td>
<td>0% (0/977)</td>
<td>0.23% (2/881)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISC</td>
<td>0% (0/3 395)</td>
<td>0.031% (1/3 185)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irish/WTCCC2</td>
<td>0% (0/1 377)</td>
<td>0.10% (1/992)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>0.061% (1/1 637)</td>
<td>0% (0/960)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swedish</td>
<td>0% (0/4 655)</td>
<td>0.066% (4/6 038)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total replication</td>
<td>0.0021% (3/14 256)</td>
<td>0.082% (12/14 612)</td>
<td>0.020 (1-Tail)</td>
<td></td>
</tr>
<tr>
<td>Total discovery + replication</td>
<td>0.014% (3/21 138)</td>
<td>0.085% (22/25 867)</td>
<td>0.00086 (2-Tail)</td>
<td>0.17 (0.050–0.56)</td>
</tr>
</tbody>
</table>

Other disorders
ID/DD/CM
ASD

---

Table 1. Frequencies of 22q11.2 duplications in cases and controls

---

Li et al. Biol Psych 2016
Microduplication 22q11.2: A new chromosomal syndrome

Marie-France Portnoï

<table>
<thead>
<tr>
<th>Type</th>
<th>Band</th>
<th>Location (NCBI 36/hg18)</th>
<th>Size (kb)</th>
<th>Recurrence (del/dup)</th>
<th>Frequency (n = 3,816)</th>
<th>p value (C = 232)</th>
<th>Studies</th>
<th>Genes (RefSeq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16q22.3</td>
<td>chr16:69,987,425–70,647,241</td>
<td>660</td>
<td>2 (1/1)</td>
<td>0.05%</td>
<td>1.00</td>
<td>1,2</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>20q13.33</td>
<td>chr20:61,056,624–61,076,763</td>
<td>20</td>
<td>3 (1/2)</td>
<td>0.08%</td>
<td>0.53</td>
<td>2,4</td>
<td>SLC17A9</td>
<td></td>
</tr>
<tr>
<td>22q11.21</td>
<td>chr22:17,265,500–19,786,200</td>
<td>2,521</td>
<td>5 (3/2)</td>
<td>0.13%</td>
<td>0.002</td>
<td>1,2,3,4</td>
<td>56</td>
<td></td>
</tr>
</tbody>
</table>

- Frequency of 0.05% in nonsyndromic simplex ASD

Opposing Effects on Cortical Surface Area and Thickness in Deletion vs. Duplication

\[\text{dup} < \text{CTL} < \text{del}\]

\[\text{del} < \text{CTL} < \text{dup}\]

Lin et al J Neurosci 2017
Whole exome sequencing (WES) Technology that assesses ~45 million individual nucleotides in regions of DNA that encode proteins (exons; ~1% of genome!), enabling the detection of SNVs and indels.

Whole genome sequencing (WGS) Assesses ~3.2 billion individual nucleotides across the entire genome, enabling the detection of SNVs, indels, large CNVs, and large SVs.

• WGS can improve detection of common variants in GWAS by statistically inferring SNPs not directly genotyped (imputation) and by identifying specific risk variants within region (fine mapping).

• May allow detection of common structural variants, including CNVs, missed by current SNP-based approaches
**Glossary- Key Terms Relevant to Rare Variation**

**Single Nucleotide Variant (SNV)** A genetic variant changing one nucleotide (e.g. C) to another (e.g. T).

**Insertion/deletion (indel)** The gain or loss of 1-50 nucleotides, usually detected by sequencing.

**Protein Truncating Variant (PTV)** An SNV or indel that disrupts one copy of a gene resulting in a premature stop codon that elicits nonsense mediated decay so that no protein is formed.

**Germline de novo mutation** A new genetic variant observed in a child, but not in either parent.

**Somatic de novo mutation** A new genetic variant observed in some cells, but not others.

**Chromosomal microarray (CMA)** Technology that enables the number of copies of DNA to be assessed at thousands of locations across the genome, enabling the detection of CNVs. Many SNP genotyping microarrays detect SNPs at these locations too.

**Mendelian disorder** A disorder caused by mutations of high effect at a single locus.

**Complex disorder** A disorder caused by multiple variants with a range of effect sizes at multiple loci, often in combination with environmental factors.

**Loss of Function (LoF) variant** Genetic variants predicted to seriously disrupt the function of protein-coding genes.

**pLI score** Probability that a given gene is haploinsufficient, and thus likely intolerant of LoF variation. Genes with high pLI scores ($\geq 0.9$) are extremely LoF intolerant, whereby genes with low pLI scores ($\leq 0.1$) are LoF tolerant.
What is the difference between a CNV and SNV?

a. CNV is a segment of DNA gain/loss; SNV is a single base pair change

b. CNV involves gene expression, not gene structure

c. SNVs are bigger

d. None of the above
Rare de novo ‘likely gene disrupting’ mutations most relevant to ASD

<table>
<thead>
<tr>
<th>Study design</th>
<th>Platform</th>
<th>Variant detected</th>
<th>Disorder</th>
<th>Patients</th>
<th>Controls</th>
<th>Genome-wide hits</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case-control</td>
<td>Genotyping microarray</td>
<td>SNP (GWAS)</td>
<td>ASD</td>
<td>16,539</td>
<td>157,234</td>
<td>1</td>
<td>Anney et al., 2017(^64)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SCZ</td>
<td>36,989</td>
<td>113,075</td>
<td>108</td>
<td>Ripke et al., 2014(^7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BPD</td>
<td>11,974</td>
<td>51,792</td>
<td>2</td>
<td>Sklar et al., 2011(^39)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MDD</td>
<td>121,380</td>
<td>338,101</td>
<td>15</td>
<td>Hyde et al., 2016(^65)</td>
</tr>
<tr>
<td></td>
<td>CNV</td>
<td></td>
<td>SCZ</td>
<td>21,094</td>
<td>20,227</td>
<td>8</td>
<td>Marshall et al., 2017(^66)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BPD</td>
<td>9,129</td>
<td>81,802</td>
<td>1</td>
<td>Green et al., 2015(^67)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MDD</td>
<td>2,591</td>
<td>8,842</td>
<td>0</td>
<td>Rucker et al., 2015(^68)</td>
</tr>
<tr>
<td>Exome sequencing</td>
<td>Rare PTV</td>
<td></td>
<td>ASD</td>
<td>5,563</td>
<td>1,881</td>
<td>0</td>
<td>Sanders et al., 2015(^5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SCZ</td>
<td>2,536</td>
<td>2,543</td>
<td>0</td>
<td>Purcell et al., 2014(^69)</td>
</tr>
<tr>
<td></td>
<td>Ultra-rare PTV</td>
<td></td>
<td>SCZ</td>
<td>4,877</td>
<td>6,203</td>
<td>0</td>
<td>Genovese et al., 2016(^70)</td>
</tr>
<tr>
<td>Family-based</td>
<td>Genotyping microarray</td>
<td>CNV</td>
<td>ASD</td>
<td>4,687</td>
<td>2,100</td>
<td>8(^*)</td>
<td>Sanders et al., 2015(^5)</td>
</tr>
<tr>
<td></td>
<td>Exome sequencing</td>
<td>De novo PTV</td>
<td>ASD</td>
<td>5,563</td>
<td>1,881</td>
<td>65(^*)</td>
<td>Sanders et al., 2015(^5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SCZ</td>
<td>617</td>
<td>731</td>
<td>0</td>
<td>Fromer et al., 2014(^71)</td>
</tr>
<tr>
<td>Meta-analysis</td>
<td>Exome sequencing</td>
<td>Rare and de novo</td>
<td>SCZ</td>
<td>7,776</td>
<td>13,028</td>
<td>1</td>
<td>Singh et al., 2016(^72)</td>
</tr>
</tbody>
</table>

SCZ, schizophrenia; BPD, bipolar disorder; MDD, major depressive disorder. \(^*\)False discovery rate (FDR) ≤ 0.1.

Sanders et al Nat neuro 2017
Non-recurrent CNVs (in aggregate) associated with cognitive impairment/reduced educational achievement (Mannik et al JAMA 2015)

-Multiple linear regression model to predict IQ based on genes within CNV: pLI score, score rate for intolerance for deletions and duplications, number of protein-protein interactions, and differential stability score of regional patterns of brain gene expression
- Best model- mean (SE) decrease of 2.74 (0.68) points per unit of the probability of being loss-of-function intolerant (P = 8x 10⁻⁵).
- Suggests CNV effects are polygenic - attributable to sum of small individual effects of each gene

Huguet et al. (2018)
Is there a common variant contribution to neuropsychiatric dx risk in high-penetrance CNVs?

-SNPs detected in GWAS characterized by p-value and odds ratio

-Polygenic Risk Score (PRS) = sum of OR for each variant

\[ PRS = (\text{Allele}^{\text{SNP}_1} \times \text{Effect}^{\text{SNP}_1}) + (\text{Allele}^{\text{SNP}_2} \times \text{Effect}^{\text{SNP}_2}) + (\text{Allele}^{\text{SNP}_3} \times \text{Effect}^{\text{SNP}_3}) + \ldots \]

Example of PRS population distributions

Williams et al., iBBC
Possible Models of Disease Risk

Hypothesis 1:
Deletion $\rightarrow$ SCZ

Hypothesis 2:
High PRS $\rightarrow$ SCZ

Hypothesis 3:
Deletion + High PRS $\rightarrow$ SCZ

- Risk score calculated with PRSIce v1.25 (11 progressive thresholds ($P_T$))
- Exclude MHC region
- Logistic regression (PRS$\sim$ Phenotype)
- Covary for PCA components + sex

Slide courtesy of Nigel Williams
PRS is relevant to disease risk in Large Effect CNVs

Analysis exploring contribution of common variants to risk of severe developmental disorders (6,987 children with DD)

Niemi et al 2018
biorxiv.org/content/biorxiv/early/2018/05/06/309070
Nagelkerke R² = proportion of variance b/t groups explained by PRS

Total variance = \textbf{CNV (if present)} + PRS (R²) + unexplained variance

- 7.7\% (SE=2.1\%) of variance in risk for neurodevelopmental disorders attributable to common genetic variants; similar to common disorders such as autism (h²=11.8\%, SE=1.0\%) and MDD

- PRS explains ~18\% of variance in idiopathic SCZ \textit{(PGC 2014)}

- pTDT - investigate transmission disequilibrium of effect alleles for traits within trios. Test compares the means of PRS for probands vs. average parent-pair scores (one-sample t-test)

- Common variant signal positively correlated with genetic predisposition to fewer years of schooling, decreased intelligence, and schizophrenia risk

- Common variant scores for autism, height, birth weight, and intracranial volume were correlated with those traits in DDD cohort ~phenotypic expression in monogenic disorders affected by same variants as the general population

\textbf{DD risk ~ combination of CNV +PRS}
Central Dogma of Genetics
Genomic Mechanisms Underlying Phenotypic Variability

DNA

RNA

Protein

Genome

Transcriptome

Proteome

© 2003 Mosby, Inc. All rights reserved.
Weighted Gene Co-expression Network Analysis (WGCNA; Zhang and Horvath 2005)

- Unbiased way of defining genes w/ functional relationships; can reveal underlying organization of the transcriptome based on degree of gene neighborhood sharing

- Construct correlation matrix between all transcripts. Transform into weighted network using soft thresholding approach (Horvath et al. PNAS 2006)

- Network structure is robust and reproducible (Oldham et al. Nat Neurosci 2008; Winden et al. 2009 Mol Sys Biol; Miller et al. PNAS 2010).

- A gene’s network position is biologically meaningful
  - Identify groups of co-expressed genes (modules) that correspond to key elements of biological function (Oldham et al. 2006, 2008; Winden et al. 2009).
  - Within modules, identify the most central, “hub” genes (Horvath et al. 2006; Oldham et al. 2008).
  - Addresses multiple comparison problem in gene expression studies

GESCHWIND AND KONOPKA NATURE 2009
Overview of WGCNA Analysis

- **Construct a gene co-expression network**
  - Rationale: make use of interaction patterns among genes
  - Tools: correlation as a measure of co-expression

- **Identify modules**
  - Rationale: module (pathway) based analysis
  - Tools: hierarchical clustering, Dynamic Tree Cut

- **Relate modules to external information**
  - Array Information: clinical data, SNPs, proteomics
  - Gene Information: ontology, functional enrichment
  - Rationale: find biologically interesting modules

- **Study module relationships**
  - Rationale: biological data reduction, systems-level view
  - Tools: Eigengene Networks

- **Find the key drivers in interesting modules**
  - Rationale: experimental validation, biomarkers
  - Tools: intramodular connectivity, causality testing

- [https://labs.genetics.ucla.edu/horvath/CoexpressionNetwork/Rpackages/WGCNA/](https://labs.genetics.ucla.edu/horvath/CoexpressionNetwork/Rpackages/WGCNA/)
Examine Relationship Between Gene Modules and Phenotypic traits
Module–Trait Relationships

<table>
<thead>
<tr>
<th></th>
<th>Autism Spectrum Disorder</th>
<th>Gender</th>
<th>Psychotic Disorder</th>
<th>Antipsychotic Use</th>
<th>Positive Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>dark turquoise</td>
<td>0.4 (0.006)</td>
<td>-0.077</td>
<td>-0.062</td>
<td>0.010 (0.5)</td>
<td>-0.021 (0.9)</td>
</tr>
<tr>
<td>royal blue</td>
<td>0.046 (0.8)</td>
<td>0.012</td>
<td>0.17 (0.3)</td>
<td>0.90 (0.6)</td>
<td>0.17 (0.3)</td>
</tr>
<tr>
<td>red</td>
<td>0.069 (0.6)</td>
<td>-0.015</td>
<td>0.066</td>
<td>0.12 (0.4)</td>
<td>0.16 (0.2)</td>
</tr>
<tr>
<td>salmon</td>
<td>0.047 (0.8)</td>
<td>0.018</td>
<td>0.0089</td>
<td>0.06 (0.7)</td>
<td>0.18 (0.2)</td>
</tr>
<tr>
<td>blue</td>
<td>0.14 (0.3)</td>
<td>-0.077</td>
<td>0.22 (0.2)</td>
<td>0.16 (0.3)</td>
<td>0.37 (0.01)</td>
</tr>
<tr>
<td>dark red</td>
<td>-0.061 (0.7)</td>
<td>-0.22</td>
<td>0.14 (0.3)</td>
<td>0.03 (0.8)</td>
<td>0.36 (0.02)</td>
</tr>
<tr>
<td>yellow</td>
<td>0.015 (0.9)</td>
<td>-0.18</td>
<td>0.2 (0.2)</td>
<td>0.03 (0.8)</td>
<td>0.42 (0.004)</td>
</tr>
<tr>
<td>green-yellow</td>
<td>0.019 (0.9)</td>
<td>-0.16</td>
<td>0.052 (0.7)</td>
<td>-0.06 (0.6)</td>
<td>0.15 (0.3)</td>
</tr>
<tr>
<td>green</td>
<td>0.0034 (1)</td>
<td>-0.043</td>
<td>0.058 (0.7)</td>
<td>0.14 (0.3)</td>
<td>-0.1 (0.5)</td>
</tr>
<tr>
<td>pink</td>
<td>0.072 (0.6)</td>
<td>0.03 (0.6)</td>
<td>-0.16 (0.3)</td>
<td>0.15 (0.3)</td>
<td>-0.23 (0.6)</td>
</tr>
<tr>
<td>light green</td>
<td>-0.08 (0.6)</td>
<td>0.12 (0.4)</td>
<td>0.18 (0.2)</td>
<td>0.12 (0.4)</td>
<td>0.26 (0.08)</td>
</tr>
<tr>
<td>magenta</td>
<td>-0.17 (0.3)</td>
<td>0.19 (0.2)</td>
<td>-0.013 (0.9)</td>
<td>-0.02 (0.9)</td>
<td>-0.06 (0.7)</td>
</tr>
<tr>
<td>midnight blue</td>
<td>0.0036 (1)</td>
<td>-0.12 (0.4)</td>
<td>0.26 (0.08)</td>
<td>-0.07 (0.7)</td>
<td>-0.41 (0.004)</td>
</tr>
<tr>
<td>purple</td>
<td>0.029 (0.8)</td>
<td>-0.16 (0.3)</td>
<td>0.32 (0.03)</td>
<td>0.04 (0.8)</td>
<td>0.15 (0.3)</td>
</tr>
<tr>
<td>cyan</td>
<td>-0.017 (0.9)</td>
<td>-0.059 (0.7)</td>
<td>0.03 (0.8)</td>
<td>-0.2 (0.2)</td>
<td>0.15 (0.3)</td>
</tr>
<tr>
<td>light yellow</td>
<td>0.0047 (1)</td>
<td>-0.18 (0.2)</td>
<td>0.17 (0.3)</td>
<td>-0.09 (0.6)</td>
<td>0.41 (0.004)</td>
</tr>
<tr>
<td>grey</td>
<td>0.0036 (1)</td>
<td>-0.28 (0.06)</td>
<td>-0.013 (0.9)</td>
<td>-0.17 (0.5)</td>
<td>0.11 (0.4)</td>
</tr>
<tr>
<td>light cyan</td>
<td>-0.043 (0.8)</td>
<td>0.063 (0.7)</td>
<td>0.024 (0.9)</td>
<td>-0.22 (0.1)</td>
<td>-0.06 (0.7)</td>
</tr>
<tr>
<td>turquoise</td>
<td>0.02 (0.9)</td>
<td>0.12 (0.4)</td>
<td>-0.19 (0.2)</td>
<td>-0.22 (0.1)</td>
<td>-0.38 (0.009)</td>
</tr>
<tr>
<td>dark green</td>
<td>-0.049 (0.7)</td>
<td>-0.14 (0.3)</td>
<td>-0.048 (0.8)</td>
<td>-0.04 (0.8)</td>
<td>-0.088 (0.6)</td>
</tr>
<tr>
<td>dark grey</td>
<td>-0.073 (0.6)</td>
<td>-0.14 (0.4)</td>
<td>-0.28 (0.06)</td>
<td>-0.17 (0.3)</td>
<td>-0.33 (0.03)</td>
</tr>
<tr>
<td>brown</td>
<td>-0.14 (0.4)</td>
<td>0.061 (0.7)</td>
<td>-0.041 (0.8)</td>
<td>0.01 (0.9)</td>
<td>-0.21 (0.2)</td>
</tr>
<tr>
<td>black</td>
<td>0.02 (0.9)</td>
<td>0.13 (0.4)</td>
<td>-0.25 (0.09)</td>
<td>-0.1 (0.5)</td>
<td>-0.42 (0.004)</td>
</tr>
<tr>
<td>tan</td>
<td>0.067 (0.7)</td>
<td>0.072 (0.6)</td>
<td>-0.22 (0.1)</td>
<td>0.04 (0.8)</td>
<td>-0.29 (0.05)</td>
</tr>
</tbody>
</table>

Network construction of co-expressed genes in relation to psychosis and autism spectrum disorders in 22q11 deletion.

Jalbrzikowski et al 2015
4) Infer functional relevance of interesting gene sets by conducting enrichment analysis via hypergeometric overlap testing with Gene Ontology (GO) terms, by asking if clusters include more genes than expected from groups of genes with established biological pathways and functions.
Developmental Gene Expression (BrainSpan): 22q11 genes overlap with gene modules implicated in idiopathic schizophrenia and ASD

* * **p<.05; **FDR <.05 (corrected for number of modules)
† Module enriched for neuronal markers (ZHANG et al., 2014/SEYFRIED et al., 2016)

Forsyth et al., in preparation
Questions!!

1. Pros and cons of Exome vs. Whole Genome Sequencing?

2. Say you have a gene with pLI of 90. What would you predict would be the effect on IQ?

3. Why are some segments of DNA particularly susceptible to recombination?

4. What is the central dogma of genetics?

5. What are some questions you might want to answer with a study of rare variation on the brain?
NAPLS Consortium
Alan Anticevic, Charlie Schleifer, Jean Addington, Ty Cannon, Diana Perkins, Larry Seidman, Kristin Cadenhead, Barbara Cornblatt, Daniel Mathalon, Thomas McGlashan, Diana Perkins, Ming Tsuang, Scott Woods, Elaine Walker

22q11.2 International Brain Behavior Consortium (Raquel Gur, Thomas Lehner)

22q11 ENIGMA Working Group

Neurobehavioral Genetics
Giovanni Coppola
Nelson Freimer
Dan Geschwind

USC Laboratory of NeuroImaging
Julio Villalon
Paul Thompson
Stanford University
Joachim Hallmayer
Sergiu Pasca
Acknowledgements

This work supported by: the National Institute of Mental Health (NIMH), IMHRO, Brain-Behavior Research Foundation (NARSAD), Simons Foundation, and all the patients and their families who participated.