

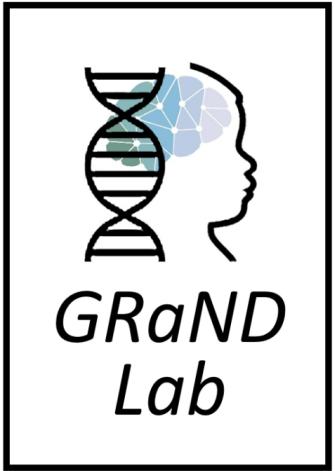


INTRODUCTION TO IMAGING GENETICS

Neurohackademy
August 1, 2024

Jennifer Forsyth, Ph.D.
Assistant Professor
University of Washington

ACKNOWLEDGMENTS



GRaND Lab

- Ariana Chavannes
- Mahnoor Hyat
- Emma Mirhashemi
- Kate Scheuer
- Sam Sievertsen
- Zac Trevorrorow
- Jinhan Zhu
- Mihin Wajayasundara
- Sarah D'Souza
- Nathan Chen
- Minkyung Cha



Collaborators

- Carrie Bearden
- Eva Mennigen
- Amy Lin
- Daqiang Sun
- Ariana Vajdi
- Leila Kushan-Wells
- Christopher Ching
- Julio Villalon Reina
- Paul Thompson
- **22q11.2 ENIGMA Working Group**



✉ jenforsy@uw.edu

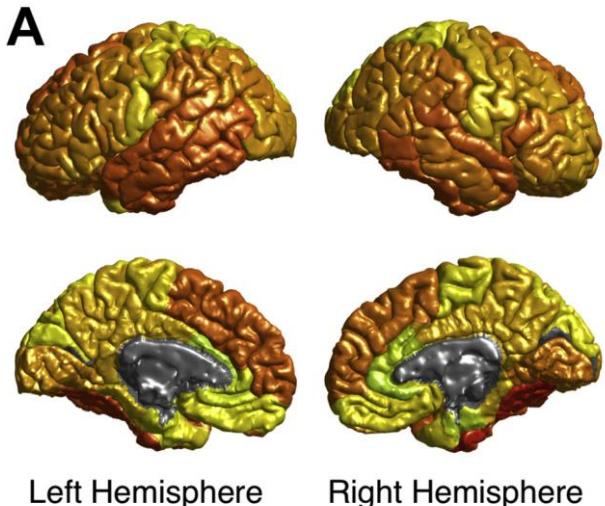
🐦 jenforsythphd

OVERVIEW

- Intro to central dogma of biology + human genome
- Research designs in imaging genetics
 - Genetic variants vs. imaging phenotypes
 - Genetic scores vs. imaging phenotypes
 - Imaging transcriptomics
- Hands-on imaging transcriptomics

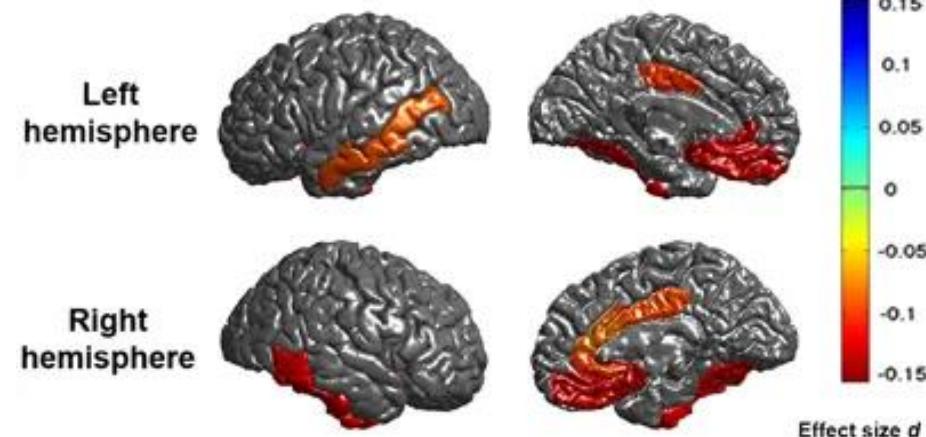
WHY GENETICS/GENOMICS/-OMICS?

SZ patients versus healthy controls:
regional cortical thickness



Van Erp et al., 2018, *Biological Psychiatry*

Adult MDD patients versus healthy controls:
regional cortical thickness



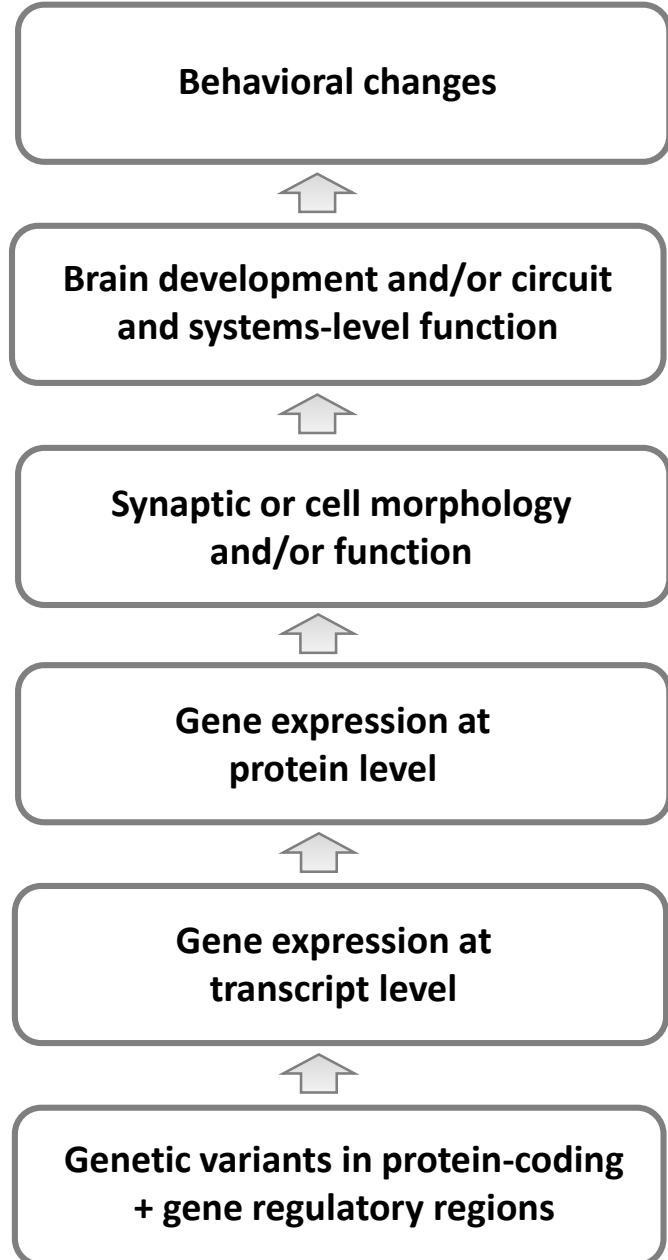
Schmaal et al., 2017, *Molecular Psychiatry*

What do we know now about these disorders?

What don't we know?

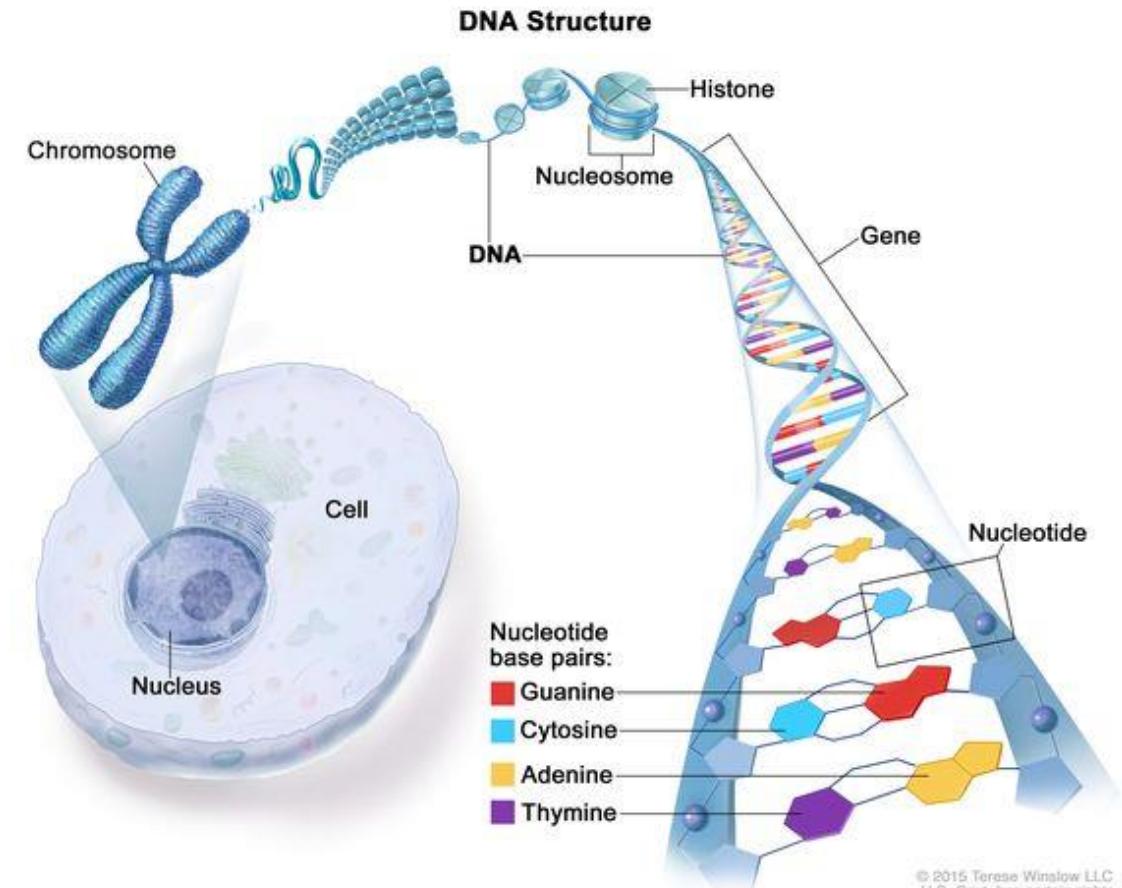
WHY GENETICS/GENOMICS/-OMICS?

- Genetics is **causal** + often contributes to neurobehavioral differences
- Genetics + transcriptomics offer **window into biology**
- Molecular genetic methods **increasingly affordable and reliable**, can leverage a wealth of resources from genetics + bioinformatics communities
 - Genome-wide genotyping array = ~\$50
 - Exome sequencing = ~\$300
 - Whole genome sequencing = ~\$1000**



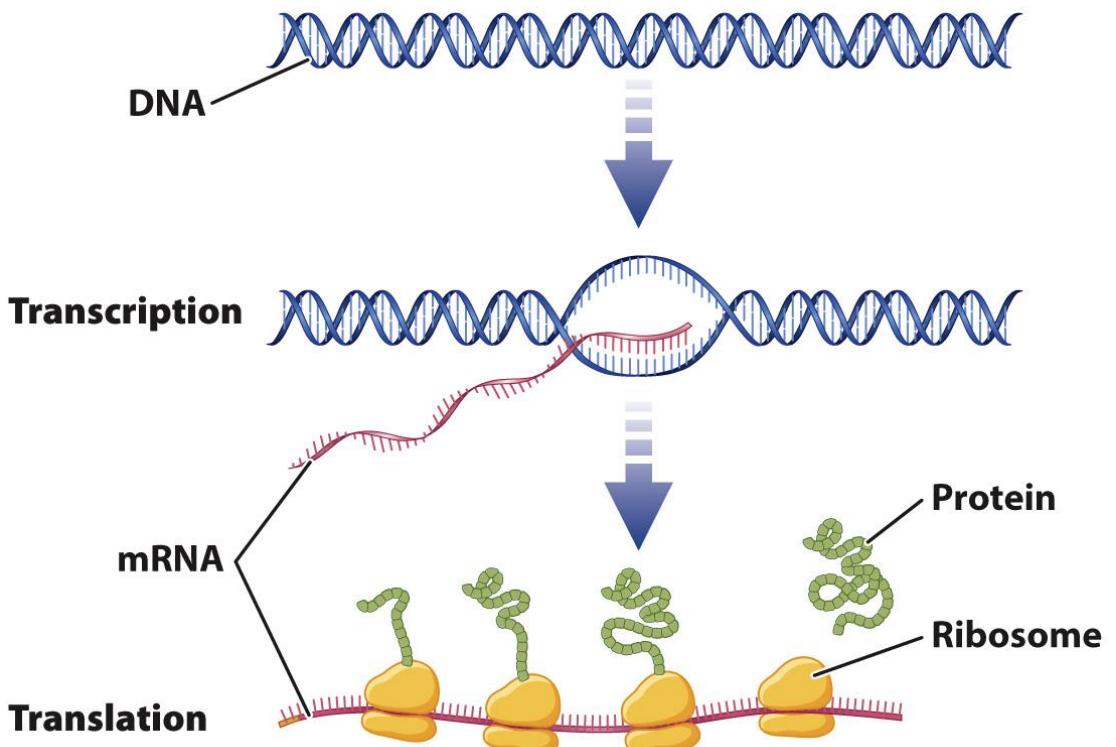
GENETICS AND THE CENTRAL DOGMA OF BIOLOGY

- DNA = 4 letter nucleotide genetic language
 - Adenine (A), cytosine (C), guanine (G), thymine (T)
 - Double-stranded helix
- Instructions to create a human from a single cell
- Carried in every cell nucleus



CENTRAL DOGMA OF MOLECULAR BIOLOGY

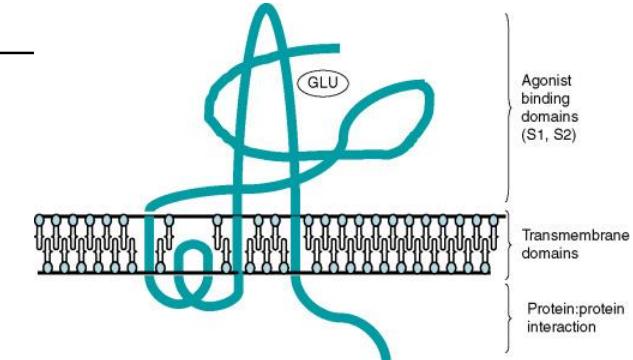
- Functional unit of DNA is genes, which encode proteins = building blocks of all cells in the body
- Central dogma describes flow of genetic information from DNA, to messenger RNA via transcription, to protein, via translation of mRNA into amino acids



DNA/RNA → AMINO ACIDS → PROTEIN

chr5: 153,490,738 - 153,490,944

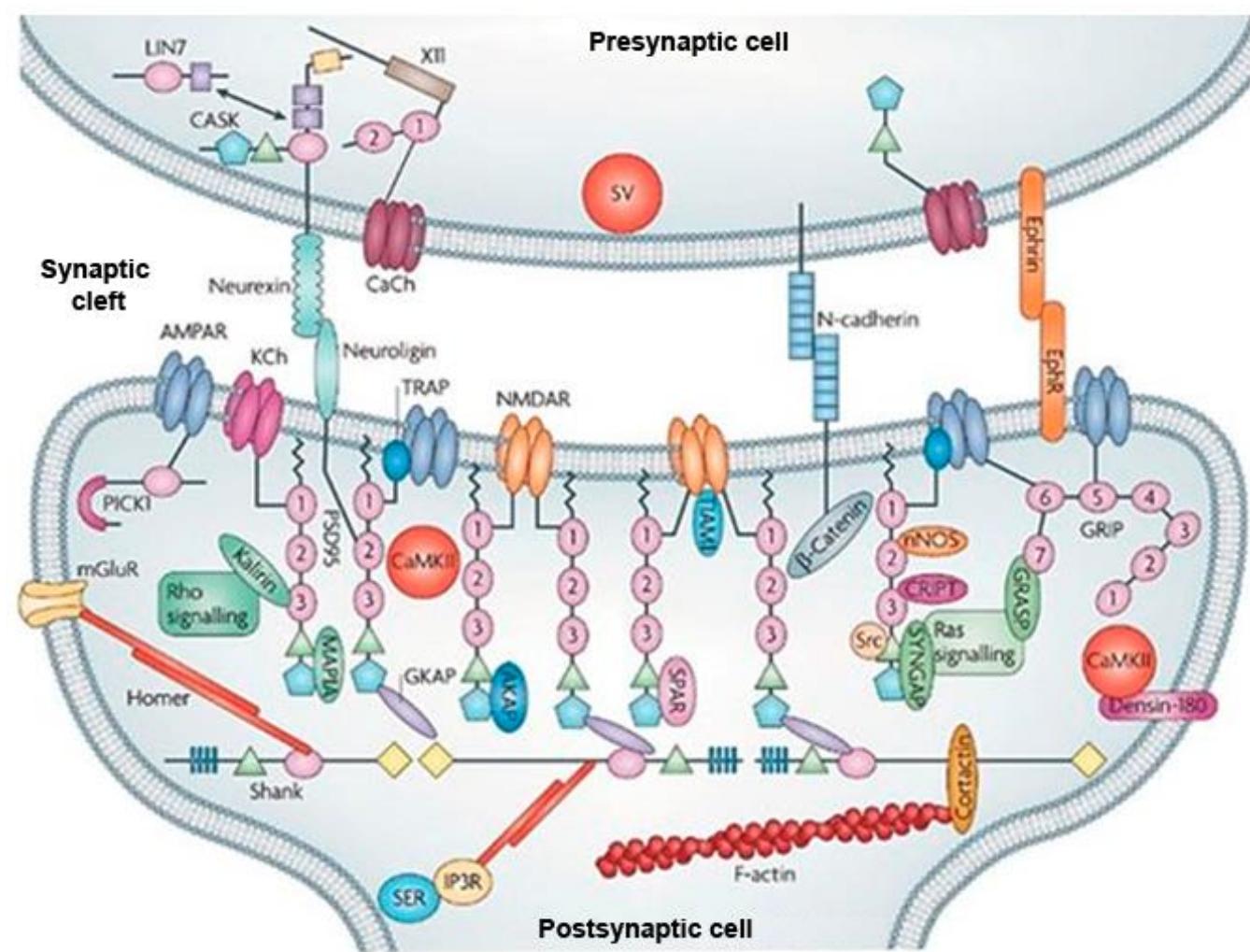
Start of GRIA1 gene = glutamate AMPA receptor subunit



...AACCTCACGAAAGGAAGGAAGCAAGCAAGCAAGGAAGGA
ACTGCAGGAGGAAAAAGAACAGGCAGAACAGCGAGAAGAATA
AAGGGAAAGGGGGGAAACACCAAATCTATGATTGGACCTGG
GCTTCTTTGCCAATGCAAAAAGGAATATGCAGCACATT
Met (M) Gln (Q) His (H) Ile (I) Phe (F)
GCCTTCTTCTGCACCGGTTCCCTAGGCGCGGTAGTAGGTGCC...
Ala (A) Phe (F) Phe (F) Cys (C) Thr (T) Gly (G) Phe (F) Leu (L) Gly (G) Ala (A) Val (V) ? ? ?

GENES → PROTEIN BUILDING BLOCKS → SYNAPSE

- Presynaptic:
 - **Scaffold proteins** (e.g., **BSN**) tether vesicles carrying neurotransmitters near presynaptic membrane
 - When action potential arrives, vesicles docked by **SNARE proteins** (e.g., **SNAP25, STX1A**) fuse with membrane to release neurotransmitter
- Transsynaptic:
 - Pre + post synaptic sides held together by **neurexin** (e.g. **NRXN1**) + **neuroligin** (e.g., **NRLGN1**) cell adhesion molecules
- Postsynaptic:
 - Active zone in dendritic spines containing post-synaptic density = **neurotransmitter receptors** (e.g., **GRIA1, GRIA2, GRIA3, GRIA4** for AMPAR), signaling molecules, cytoskeleton, and scaffold proteins that bind components

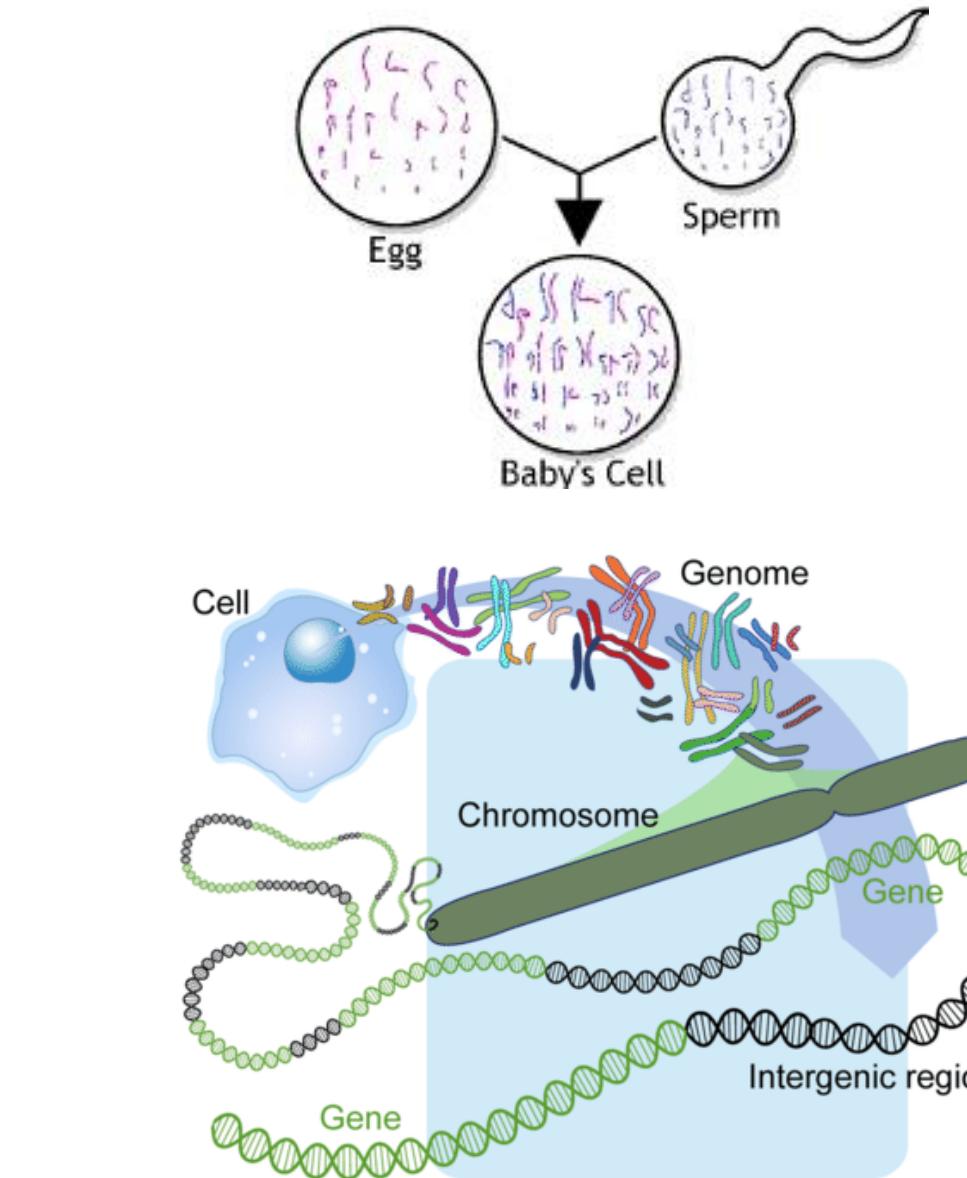


HUMAN GENOME

How long is the human genome?

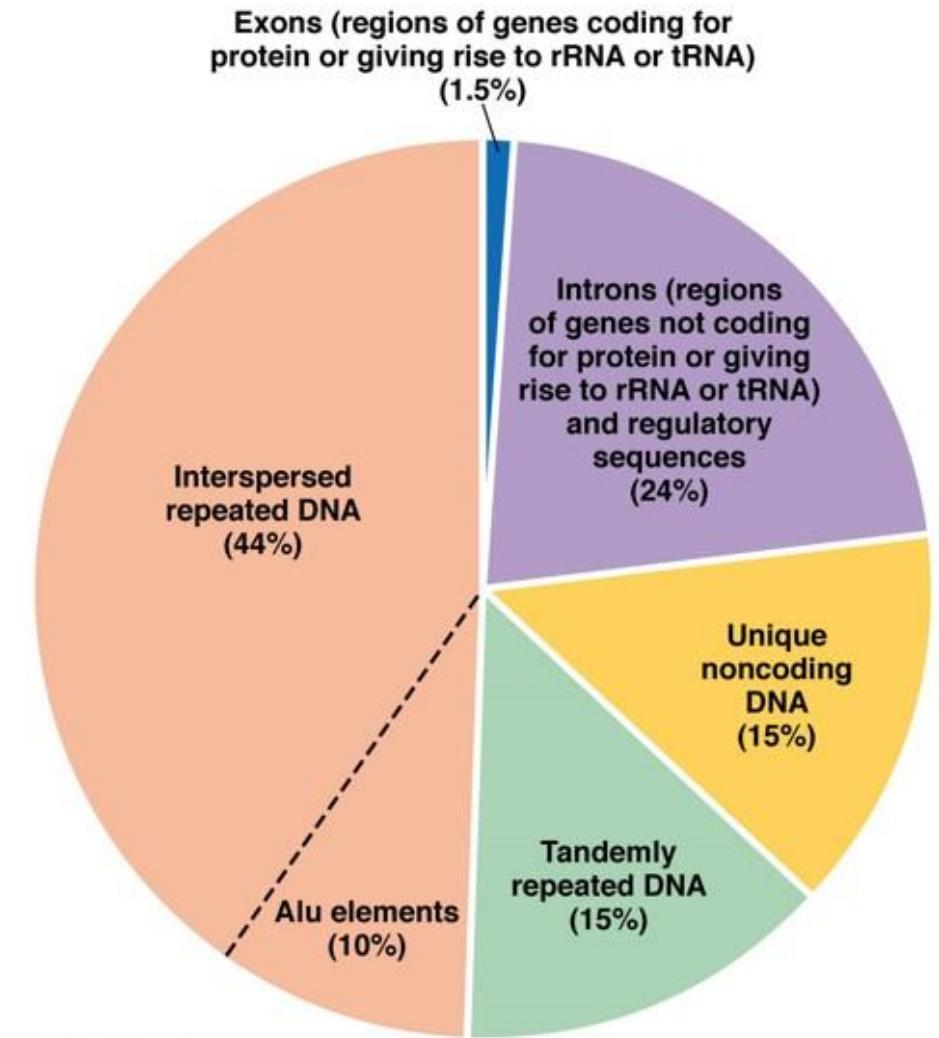
HUMAN GENOME STATS

- Haploid human genome is ~3 billion nucleotide pairs long
- Each cell carries 2 copies of the genome
 - 1 from mother's egg + 1 from father's sperm, originally combined in fertilized egg
- Total per cell = ~6 billion nucleotide pairs
 - Diploid human genome
 - ~6 feet stretched out



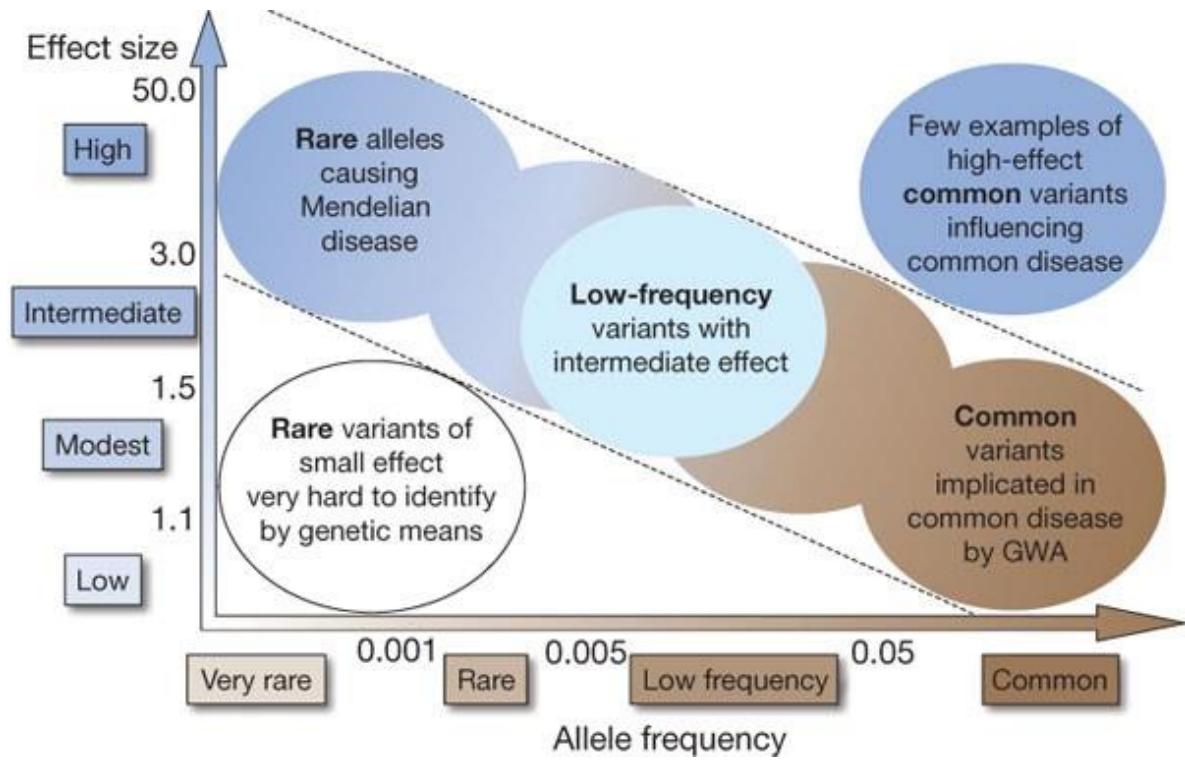
HUMAN GENOME STATS

- Approximately 20,000 protein-coding genes, 8,000 non-coding genes
- 1-1.5% of human genome directly encodes proteins
- Most genes are 10-100 kb in length
- Rest of DNA sequence influences how often genes are expressed or no known function



HUMAN GENOME STATS

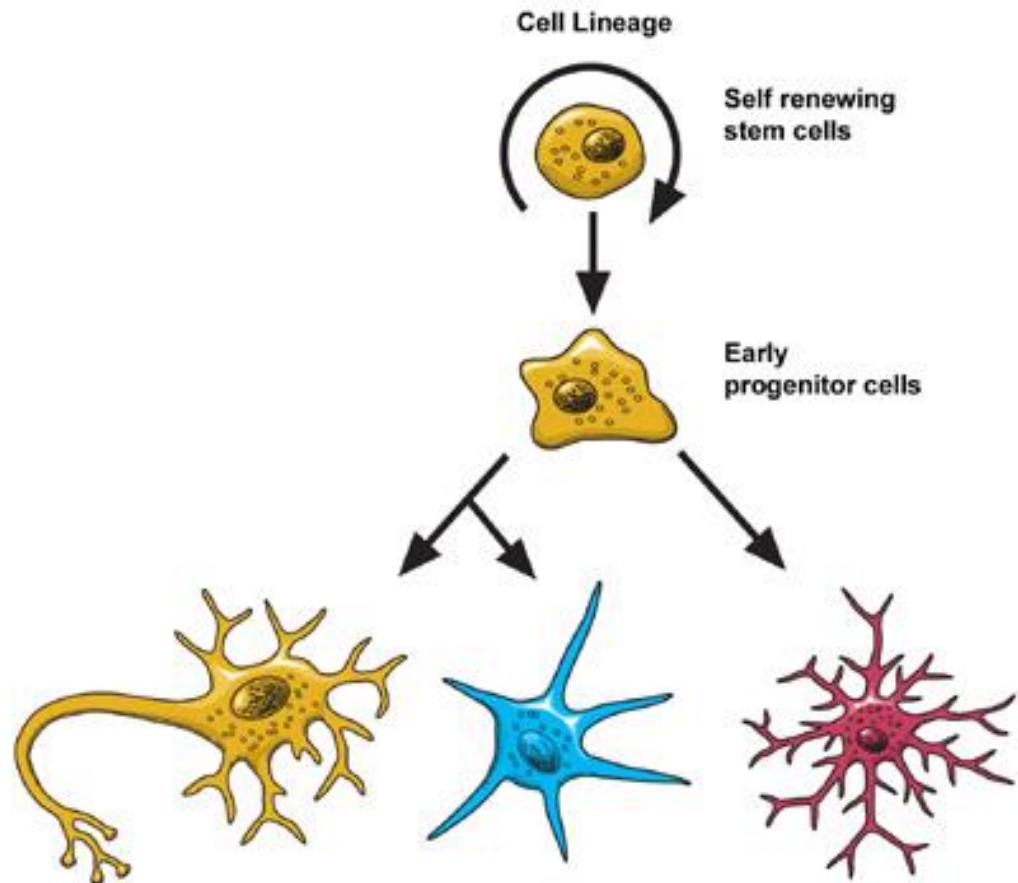
- Humans share ~99.7% of DNA sequence with one another
- Majority of variants carried are common
 - ~1-4% with frequency < 0.5%
 - 40,000 to 200,000 rare variants
- Rare variants have more impact because under negative selection or newly arisen
- Variant impact on phenotype depends on whether located in protein-coding or gene regulatory regions that impact gene expression



Manolio et al., 2009, *Nature*

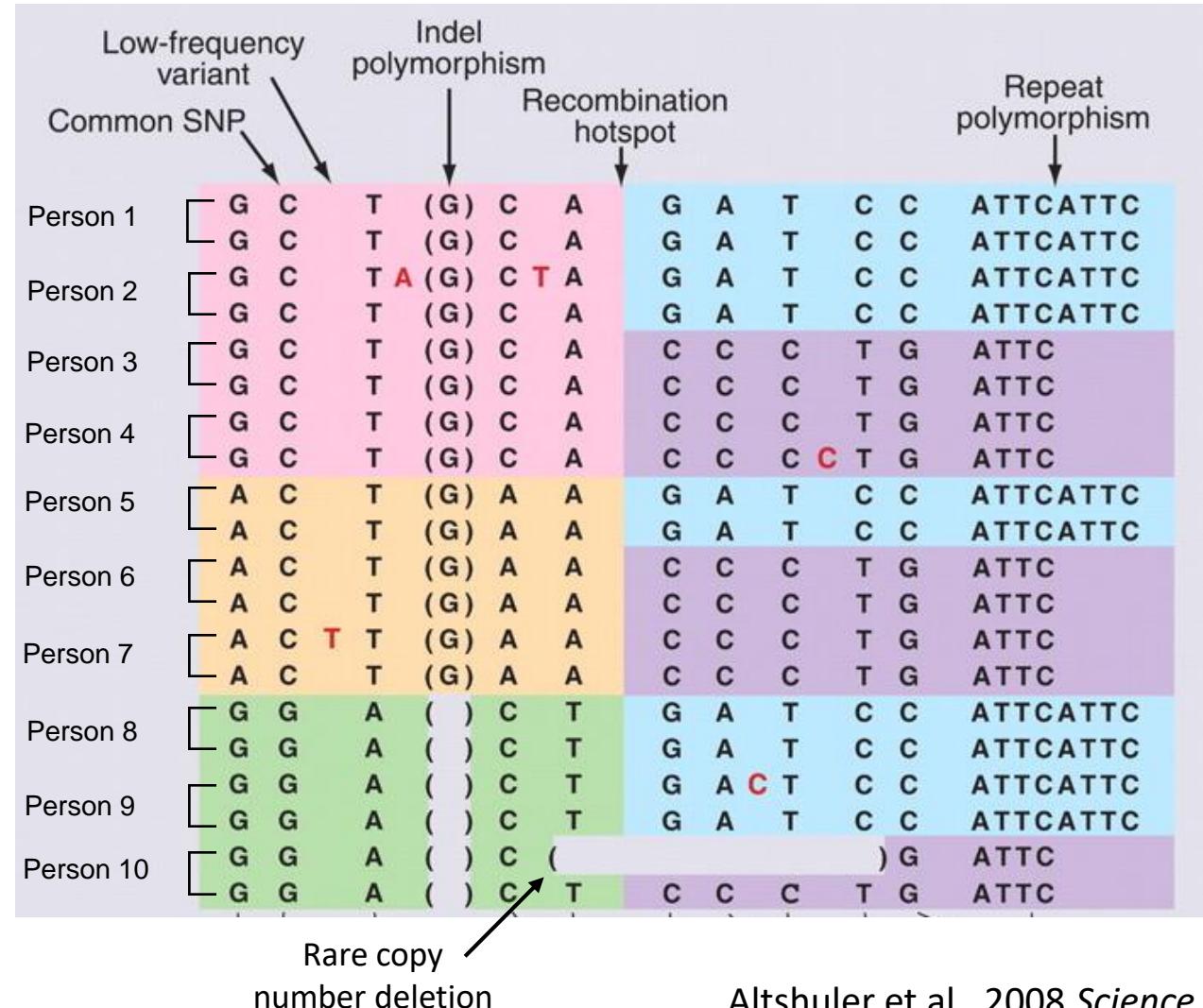
GENE EXPRESSION DRIVES CELL IDENTITY + FUNCTION

- Selective expression of different combinations of genes determines the function and characteristics of different cells
- Drives pluripotent stem cells to mature into different cell types
- Gives cells stable identity



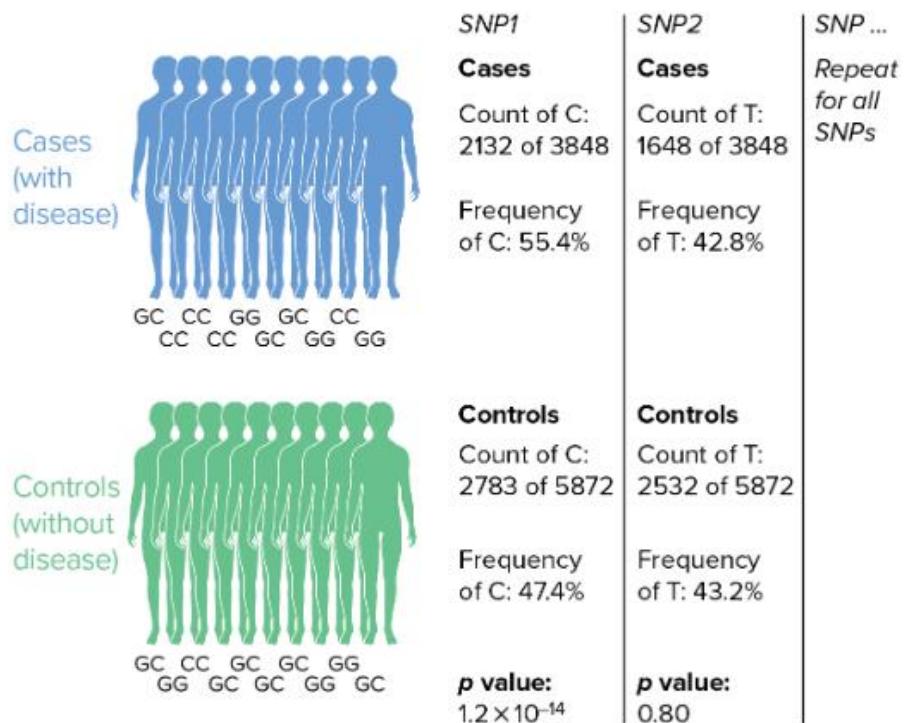
HUMAN GENETIC VARIATION

- DNA variants in 10 individuals
- Representative for 5 kb
- Common single nucleotide polymorphisms (SNPs) most researched
- Many classes of variants



GENOME WIDE ASSOCIATION STUDIES (GWAS)

- Test millions of sites across genome that commonly vary (i.e., SNPs) for association with phenotype
- Tests **additive model** for relationship between number of minor alleles and likelihood of trait
- Power of GWAS depends on sample size + fraction of trait variation explained by SNPs



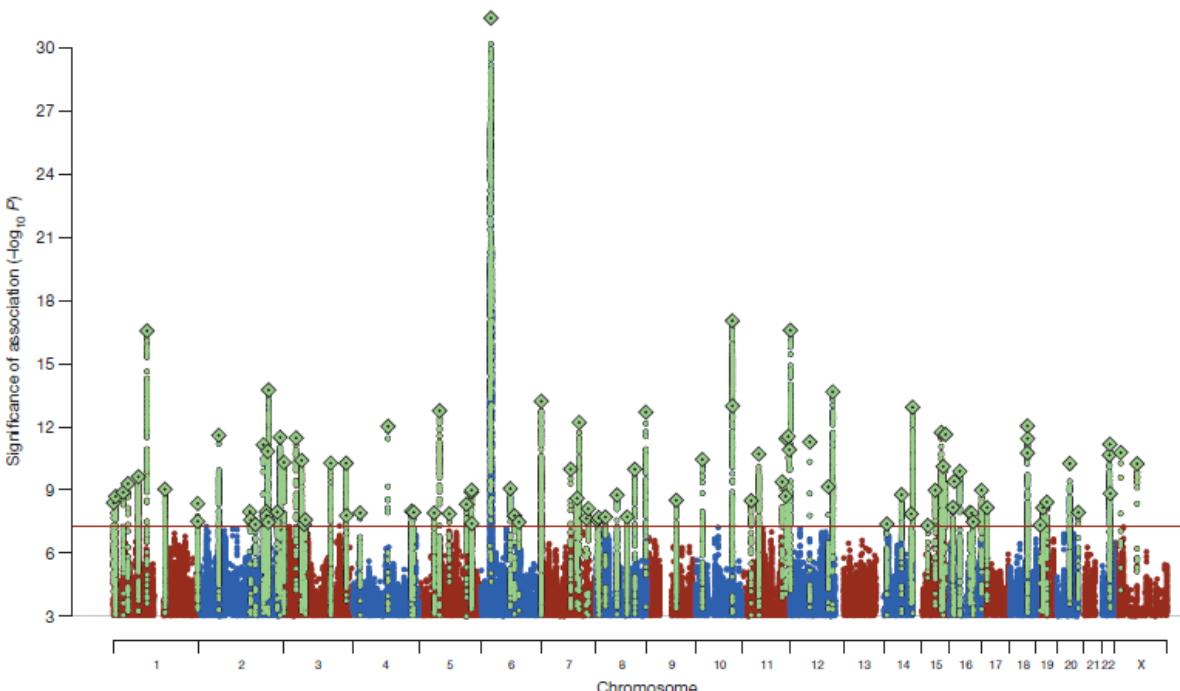
GENOME WIDE ASSOCIATION STUDIES (GWAS)

Genome-wide, what variants are associated with variation in a trait?

What biology is implicated by associated variants?

BASIC GWAS OUTPUT

- Association effect size and p-value stats for SNPs across genome
- Follow-up analyses to implicate genes + convergent biology



CHR	POS	SNP	Tested_Allele	Other_Allele	Freq_Testing_Allele_in_HRS	BETA	SE	P	N
7	92383888	rs10	A	C	0.06431	-0.0066	0.0035	6.00E-02	605309
12	126890980	rs1000000	A	G	0.2219	-0.0001	0.0017	9.50E-01	706961
4	21618674	rs10000010	T	C	0.5086	-0.0022	0.0014	1.20E-01	697417
4	1357325	rs10000012	C	G	0.8634	0.0143	0.0021	4.50E-12	709514
4	37225069	rs10000013	A	C	0.7708	0.0015	0.0017	3.90E-01	704912
4	84778125	rs10000017	T	C	0.2284	0.0019	0.0017	2.80E-01	703109
3	183635768	rs1000002	T	C	0.4884	0.004	0.0014	5.10E-03	709522
4	95733906	rs10000023	T	G	0.5817	0.0029	0.0015	4.60E-02	691905
4	156176217	rs10000027	C	G	0.771	-0.0013	0.0019	4.80E-01	526087
3	98342907	rs1000003	A	G	0.8404	0.0078	0.002	8.90E-05	707561
4	103374154	rs10000030	A	G	0.1351	0.0029	0.0021	1.70E-01	707602
4	139599898	rs10000033	T	C	0.5355	0.002	0.0014	1.70E-01	694621
4	38924330	rs10000037	A	G	0.2516	0.0099	0.0016	1.80E-09	708831
4	189176035	rs10000038	T	G	0.6681	-0.0037	0.0015	1.40E-02	688310
4	165621955	rs10000041	T	G	0.8555	-0.0005	0.0021	8.00E-01	706813
4	5237152	rs10000042	T	C	0.04321	0.0029	0.0037	4.30E-01	695530
21	34433051	rs1000005	C	G	0.5885	-0.0054	0.0015	3.50E-04	708718
4	189321617	rs10000056	A	T	0.1752	0.0007	0.0019	7.30E-01	694725

Schizophrenia Working Group of the PGC, 2014, *Nature*

GWAS OF BRAIN TRAITS

RESEARCH

RESEARCH ARTICLE SUMMARY

CORTICAL GENETICS

The genetic architecture of the human cerebral cortex

Katrina L. Grasby*† and Neda Jahanshad*† et al.

INTRODUCTION: The cerebral cortex underlies our complex cognitive capabilities. Variations in human cortical surface area and thickness are associated with neurological, psychological, and behavioral traits and can be measured *in vivo* by magnetic resonance imaging (MRI). Studies in model organisms have identified genes that influence cortical structure, but little is known about common genetic variants that affect human cortical structure.

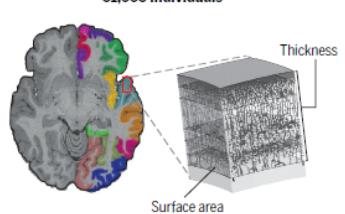
RATIONALE: To identify genetic variants associated with human cortical structure at both global and regional levels, we conducted a genome-wide association meta-analysis of brain MRI data from 51,665 individuals across 60 cohorts. We analyzed the surface area and

average thickness of the whole cortex and 34 cortical regions with known functional specializations.

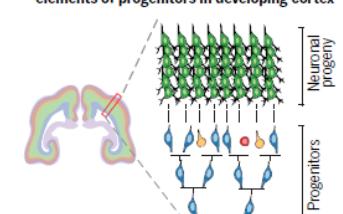
RESULTS: We identified 369 nominally genome-wide significant loci ($P < 5 \times 10^{-5}$) associated with cortical structure in a discovery sample of 33,992 participants of European ancestry. Of the 360 loci for which replication data were available, 241 loci influencing surface area and 66 influencing thickness remained significant after replication, with 237 loci passing multiple testing correction ($P < 8.3 \times 10^{-10}$; 187 influencing surface area and 50 influencing thickness).

Common genetic variants explained 34% ($SE = 3\%$) of the variation in total surface area

A Cortical structure from brain MRI in 51,665 individuals



C Surface area heritability enrichment in regulatory elements of progenitors in developing cortex



B Genomic locations of associated loci

D Genetic correlations with cortical surface area $r^2 = 0.4$

and 26% ($SE = 2\%$) in average thickness; surface area and thickness showed a negative genetic correlation ($r_G = -0.32$, $SE = 0.05$, $P = 6.5 \times 10^{-12}$), which suggests that genetic influences have opposing effects on surface area and thickness. Bioinformatic analyses showed that total surface area is influenced by genetic variants that alter gene regulatory activity in neural progenitor cells during fetal development.

By contrast, average thickness is influenced by active regulatory elements in adult brain samples, which may reflect processes that occur after mid-fetal development, such as myelination, branching, or pruning.

When considered together, these results support the radial unit hypothesis that different developmental mechanisms promote surface area expansion and increases in thickness.

To identify specific genetic influences on individual cortical regions, we controlled for global measures (total surface area or average thickness) in the regional analyses. After multiple testing correction, we identified 175 loci that influence regional surface area and 46 that influence regional thickness. Loci that affect regional surface area cluster near genes involved in the Wnt signaling pathway, which is known to influence areal identity.

We observed significant positive genetic correlations and evidence of bidirectional causation of total surface area with both general cognitive functioning and educational attainment. We found additional positive genetic correlations between total surface area and Parkinson's disease but did not find evidence of causation. Negative genetic correlations were evident between total surface area and insomnia, attention deficit hyperactivity disorder, depressive symptoms, major depressive

RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

Common genetic variation influencing human white matter microstructure

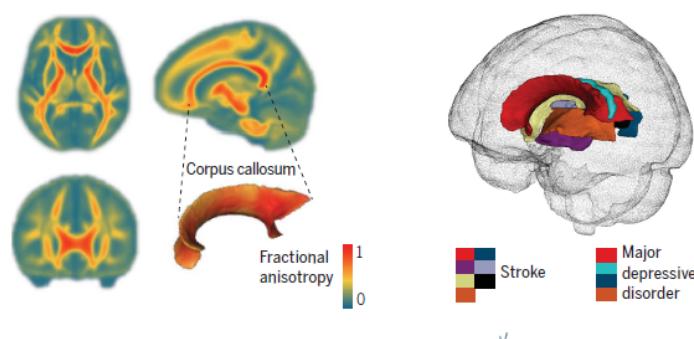
Bingxin Zhao†, Tengfei Li†, Yue Yang, Xifeng Wang, Tianyou Luo, Yue Shan, Ziliang Zhu, Di Xiong, Mads E. Hauberg, Jaroslav Bendl, John F. Fullard, Panagiotis Roussos, Yun Li, Jason L. Stein, Hongtu Zhu*

INTRODUCTION: White matter in the human brain serves a critical role in organizing distributed neural networks. Diffusion magnetic resonance imaging (dMRI) has enabled the study of white matter *in vivo*, showing that interindividual variations in white matter microstructure are associated with a wide variety of clinical outcomes. Although white matter differences in general population cohorts are known to be heritable, few common genetic variants influencing white matter microstructure have been identified.

RATIONALE: To identify genetic variants influencing white matter microstructure, we conducted a genome-wide association study (GWAS) of dMRI data from 43,802 individuals across

five data resources. We analyzed five major diffusion tensor imaging (DTI) model-derived parameters along 21 cerebral white matter tracts.

RESULTS: In the discovery GWAS with 34,024 individuals of British ancestry, we replicated 42 of the 44 genomic regions discovered in the largest previous GWAS and identified 109 additional regions associated with white matter microstructure ($P < 2.3 \times 10^{-10}$, adjusted for the number of phenotypes studied). These results indicate strong polygenic influences on white matter microstructure. Of the 151 regions, 52 passed the Bonferroni significance level ($P < 5 \times 10^{-5}$) in our analysis of nine



independent validation datasets, including four with subjects of non-European ancestry.

On average, common genetic variants explained 41% (standard error = 2%) of the variation in white matter microstructure. The 151 identified genomic regions can explain 32.3% of heritability for white matter microstructure, whereas the 44 previously identified genomic regions can only explain 11.7% of heritability. As a biological validation of our GWAS findings, we observed heritability enrichment within regulatory elements active in oligodendrocytes and other glia, whereas no enrichment was observed in neurons. These results are expected and suggest that genetic variation leads to changes in white matter microstructure by affecting gene regulation in glia.

We observed genetic correlations and localizations of white matter microstructure with a wide range of brain-related complex traits and diseases, such as cognitive functions, cardiovascular risk factors, as well as various neurological and psychiatric diseases. For example, of the 25 reported genetic risk regions of glioma, 11 were also associated with white matter microstructure, which illustrates the close genetic relationship between glioma and white matter integrity. Additionally, we found that 14 white matter microstructure-associated genes ($P < 1.2 \times 10^{-8}$) were targets for 79 commonly used nervous system drugs, such as antipsychotics, antidepressants, anti-convulsants, and drugs for Parkinson's disease and dementia.

CONCLUSION: This large-scale study of dMRI scans from 43,802 subjects improved our understanding of the highly polygenic genetic architecture of human brain white matter tracts. We identified 151 genomic regions associated with white matter microstructure. The GWAS findings were supported by enrichments within cell types that make up white matter microstructure. Moreover, we uncovered genetic relation-

GWAS OF BRAIN-RELATED TRAITS

Article

Mapping genomic loci implicates genes and synaptic biology in schizophrenia

<https://doi.org/10.1038/s41586-022-04434-5>

Received: 12 August 2020

Accepted: 10 January 2022

Published online: 8 April 2022

 Check for updates

Schizophrenia has a heritability of 60–80%, much of which is attributable to common risk alleles. Here, in a two-stage genome-wide association study of up to 76,755 individuals with schizophrenia and 243,649 control individuals, we report common variant associations at 287 distinct genomic loci. Associations were concentrated in genes that are expressed in excitatory and inhibitory neurons of the central nervous system, but not in other tissues or cell types. Using fine-mapping and functional genomic data, we identify 120 genes (106 protein-coding) that are likely to underpin associations at some of these loci, including 16 genes with credible causal non-synonymous or untranslated region variation. We also implicate fundamental processes related to neuronal function, including synaptic organization, differentiation and transmission. Fine-mapped candidates were enriched for genes associated with rare disruptive coding variants in people with schizophrenia, including the glutamate receptor subunit *GRIN2A* and transcription factor *SP4*, and were also enriched for genes implicated by such variants in neurodevelopmental disorders. We identify biological processes relevant to schizophrenia

LETTERS

<https://doi.org/10.1038/s41588-018-0151-7>

nature
genetics

Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways

Mats Nagel^{1,2,11}, Philip R. Jansen^{1,3,11}, Sven Stringer¹, Kyoko Watanabe¹, Christiaan A. de Leeuw¹, Julien Bryois⁴, Jeanne E. Savage¹, Anke R. Hammerschlag¹, Nathan G. Skene⁵, Ana B. Muñoz-Manchado⁵, 23andMe Research Team⁶, Tonya White³, Henning Tiemeier^{1,3,7}, Sten Linnarsson¹, Jens Hjerling-Leffler^{1,5,8}, Tinca J. C. Polderman¹, Patrick F. Sullivan^{1,4,9,10}, Sophie van der Sluis^{1,2,12} and Danielle Posthuma^{1,2,12*}

ARTICLES

<https://doi.org/10.1038/s41588-022-01024-z>

nature
genetics

 Check for updates

OPEN

New insights into the genetic etiology of Alzheimer's disease and related dementias

Characterization of the genetic landscape of Alzheimer's disease (AD) and related dementias (ADD) provides a unique opportunity for a better understanding of the associated pathophysiological processes. We performed a two-stage genome-wide association study totaling 111,326 clinically diagnosed/'proxy' AD cases and 677,663 controls. We found 75 risk loci, of which 42 were new at the time of analysis. Pathway enrichment analyses confirmed the involvement of amyloid/tau pathways and highlighted microglia implication. Gene prioritization in the new loci identified 31 genes that were suggestive of new genetically associated processes, including the tumor necrosis factor alpha pathway through the linear ubiquitin chain assembly complex. We also built a new genetic risk score associated with the risk of future AD/dementia or progression from mild cognitive impairment to AD/dementia. The improvement in prediction led to a 1.6- to 1.9-fold increase in AD risk from the lowest to the highest decile, in addition to effects of age and the *APOE ε4* allele.

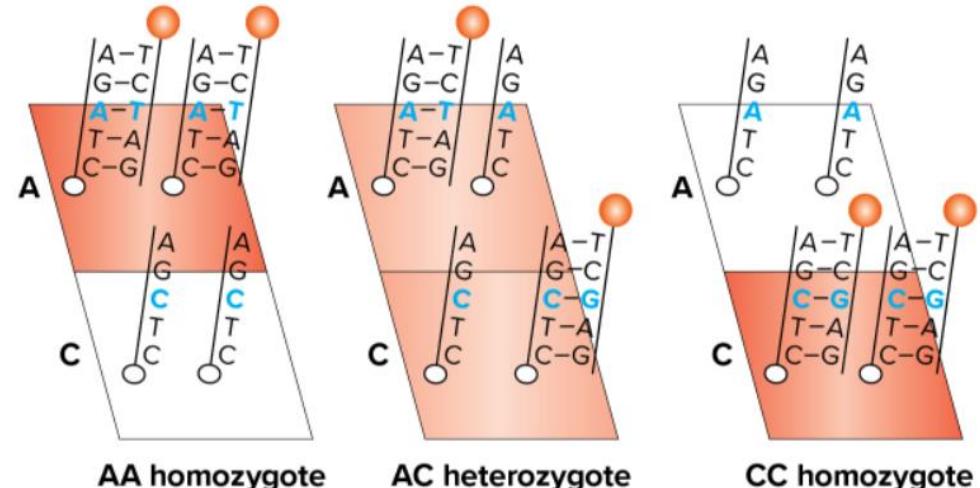
GWAS OF COMPLEX BRAIN-RELATED TRAIT THEMES

- Systematic examination of genome necessary in large samples
- Highly polygenic with 100s to 1,000s of independent associations
- Most SNPs have small effects
- GWAS largely inconsistent with candidate gene findings (Duncan et al., 2019)

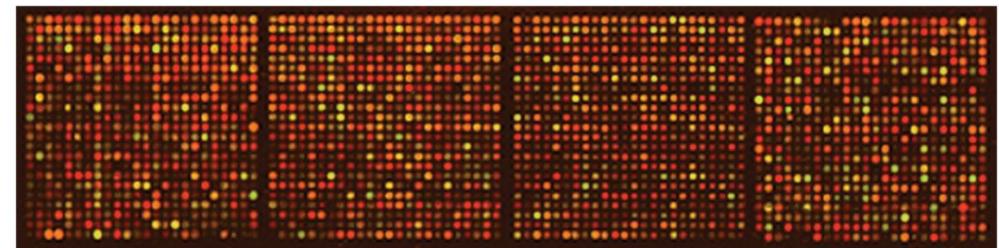
GWAS OF COMPLEX TRAITS – WHAT DO YOU NEED?

- Tens of thousands to hundreds of thousands of individuals with DNA samples + phenotyping
- Genotyping array indexing hundreds of thousands of genomic spots

(a) Microarray schematic

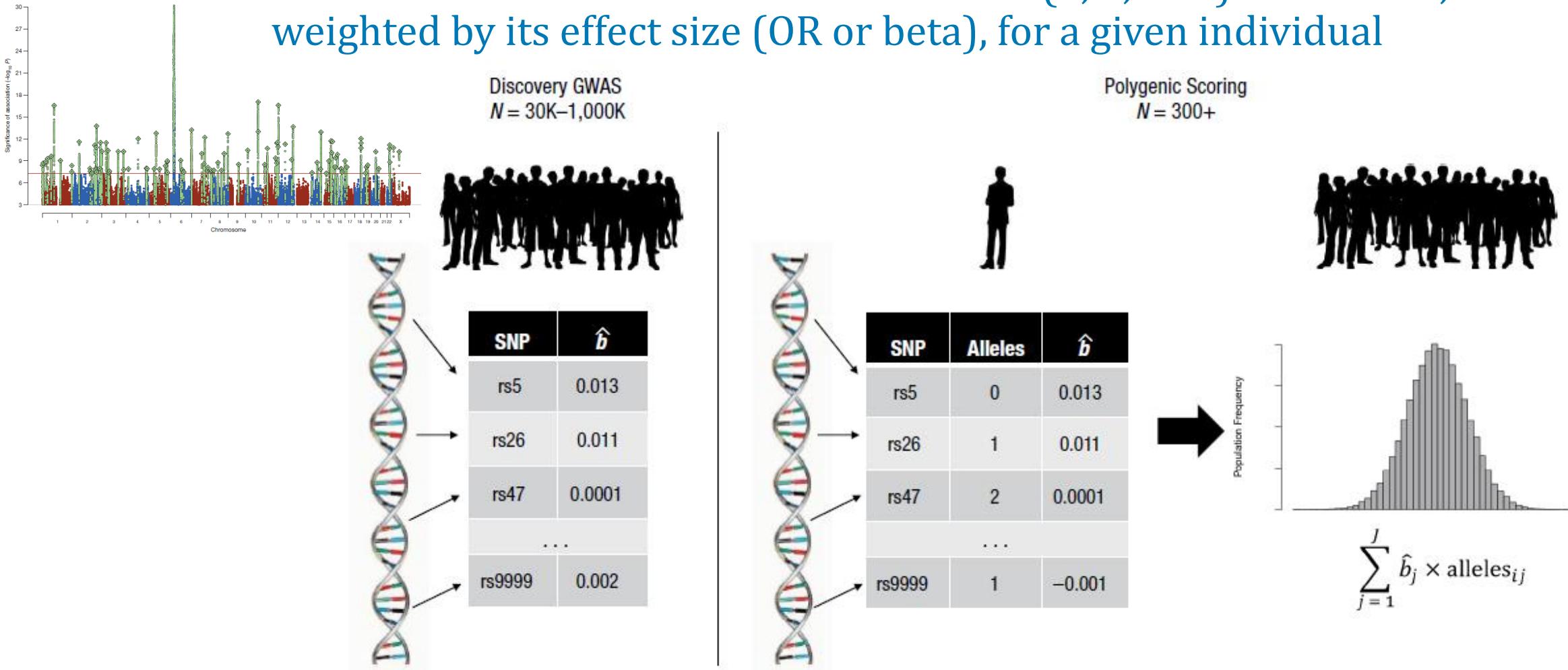


(c) Signal from part of a large microarray



PRS/GPS CONSTRUCTION FROM GWAS SUMMARY STATS

Number of risk alleles carried at each variant (0, 1, or 2) is summed, weighted by its effect size (OR or beta), for a given individual



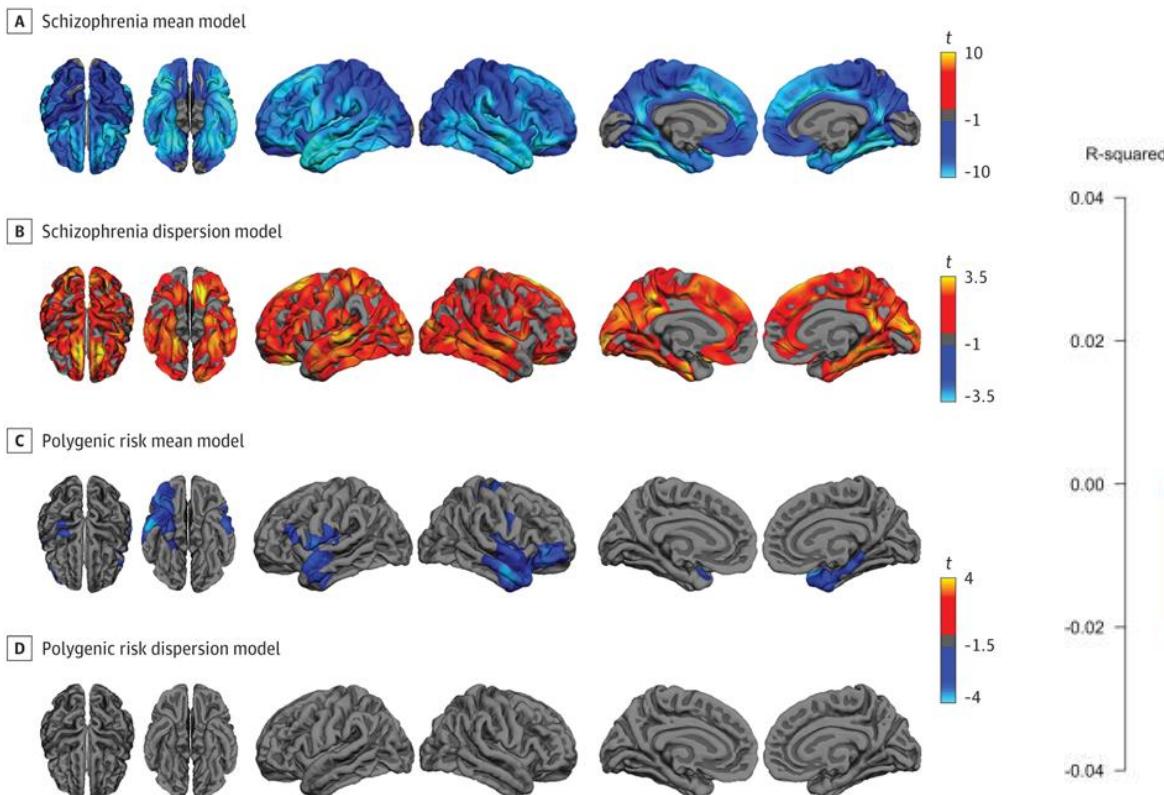
STUDIES USING PGS/GPS

Is there a link between an observed neuroimaging (or other) phenotype and genetic variants associated with that or other phenotypes?

What other phenotypes are associated with a given genetic score?

PRS VS BRAIN PHENOTYPES IN HEALTHY INDIVIDUALS

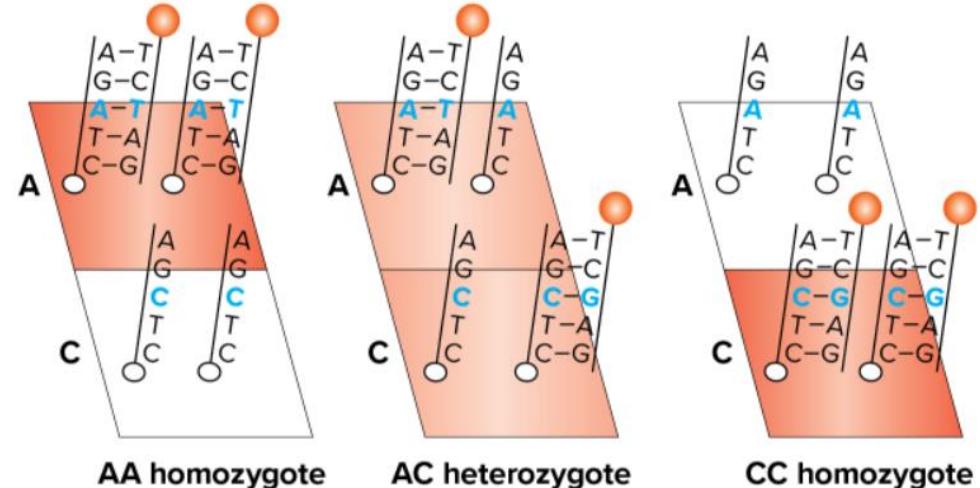
SCZ PRS associated with cortical thinning



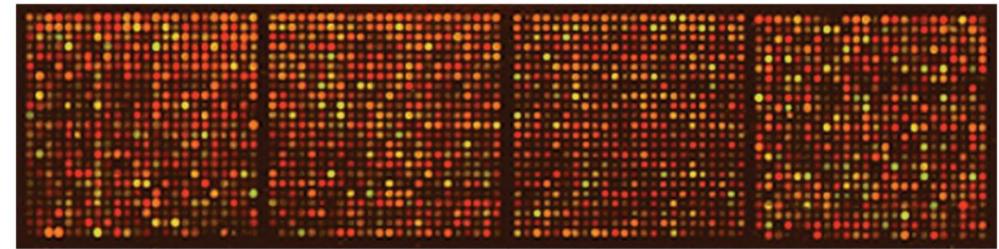
PRS/GPS VS. BRAIN TRAITS – WHAT DO YOU NEED?

- Hundreds to thousands of individuals with DNA samples + phenotyping
- Genotyping array indexing hundreds of thousands of genomic spots
- Summary statistics from GWAS for trait of interest (e.g., neuropsychiatric disorder, brain features)

(a) Microarray schematic



(c) Signal from part of a large microarray



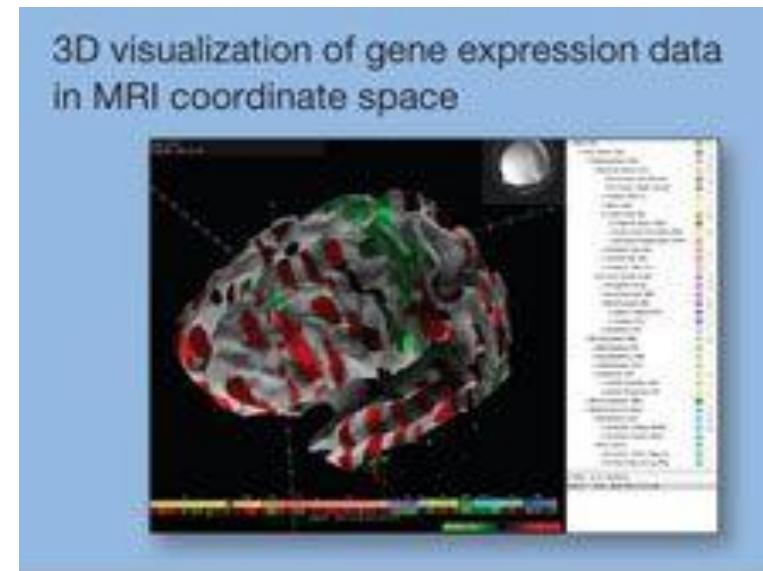
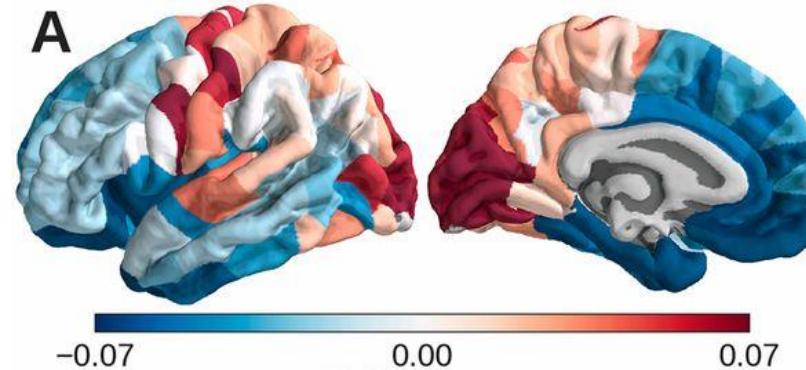
WHAT IF YOU DON'T HAVE GENOTYPED SUBJECTS?

But want to link brain features to
cellular or molecular features?

Imaging transcriptomics!

IMAGING TRANSCRIPTOMICS – WHAT DO YOU NEED?

- Neuroimaging data for trait of interest with effect size weights across regions
- Allen Human Brain Atlas (AHBA) gene expression data parcellated to same regions as neuroimaging data



USING AHBA TO GAIN MOLECULAR INSIGHTS

CrossMark
click for updates

Adolescence is associated with genetically patterned consolidation of the hubs of the human brain connectome

Kirstie J. Whitaker^{a,1,2}, Petra E. Vértes^{a,2}, Rafael Romero-García^a, František Váša^a, Michael Moutoussis^b, Gita Prabhu^b, Nikolaus Weiskopf^{b,c}, Martina F. Callaghan^b, Konrad Wagstyl^d, Timothy Rittman^d, Roger Tait^d, Cindy Ooi^d, John Suckling^{a,e,f}, Becky Inkster^a, Peter Fonagy^g, Raymond J. Dolan^{b,h}, Peter B. Jones^{a,e}, Ian M. Goodyer^{a,e}, the NSPN Consortiumⁱ, and Edward T. Bullmore^{a,e,f,j}

^aDepartment of Psychiatry, University of Cambridge, Cambridge CB2 0SZ, United Kingdom; ^bWellcome Trust Centre for Neuroimaging, University College London, London WC1N 3BG, United Kingdom; ^cDepartment of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany; ^dDepartment of Clinical Neurosciences, University of Cambridge, Cambridge CB2 3EB, United Kingdom; ^eCambridgeshire and Peterborough National Health Service Foundation Trust, Cambridge, CB21 5EF, United Kingdom; ^fMedical Research Council/Wellcome Trust Behavioural and Clinical Neuroscience Institute, University of Cambridge, Cambridge CB2 3EB, United Kingdom; ^gResearch Department of Clinical, Educational and Health Psychology, University College London, London WC1E 6BT, United Kingdom; ^hMax Planck University College London Centre for Computational Psychiatry and Ageing Research, University College London, London WC1B 5EH, United Kingdom; and ⁱImmunoPsychiatry, GlaxoSmithKline Research and Development, Stevenage SG1 2NY, United Kingdom

Edited by Michael S. Gazzaniga, University of California, Santa Barbara, CA, and approved May 26, 2016 (received for review February 16, 2016)

How does human brain structure mature during adolescence? We used MRI to measure cortical thickness and intracortical myelination in 297 population volunteers aged 14–24 y old. We found and replicated that association cortical areas were thicker and less myelinated than primary cortical areas at 14 y. However, association cortex had faster rates of shrinkage and myelination over the course of adolescence. Age-related increases in cortical myelination were maximized approximately at the internal layer of projection neurons. Adolescent cortical myelination and shrinkage were coupled and specifically associated with a dorsoventrally patterned gene expression profile enriched for synaptic, oligodendroglial- and schizophrenia-related genes. Topologically efficient and biologically expensive hubs of the brain anatomical network had greater rates of shrinkage/myelination and were associated with overexpression of the same transcriptional profile as cortical consolidation. We conclude that normative human brain maturation involves a genetically patterned process of consolid-

that shorter longitudinal (T_1) relaxation times reflect either a reduction in the fraction of “watery” cytoplasmic material, like cell bodies, synapses, or extracellular fluid, or an increase in the fraction of “fatty” myelinated material, like axons. Pruning models propose that cortical shrinkage in adolescence represents loss or remodeling of synapses, dendrites, or cell bodies (13). Myelination models propose that the cortex appears to shrink because of an increasing proportion of myelinated axons, without necessarily implying any loss or change of neuronal material (5).

In the macaque monkey, although the main phase of synaptic pruning and neuronal loss occurs earlier in development (14, 15), there is evidence for further synaptic remodeling during adolescence (16, 17). In rodents, there is histological evidence for increasing intracortical myelination during adolescence, especially at the deeper cytoarchitectonic layers of cortex (V and VI) (18, 19). At a cellular

USING AHBA TO GAIN MOLECULAR INSIGHTS

CrossMark
click for updates

Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome

Kirstie J. Whitaker^{a,1,2}, Petra E. Vértes^{a,2}, Rafael Romero-García^a, František Váša^a, Michael Moutoussis^b, Gita Prabhu^b, Nikolaus Weiskopf^{b,c}, Martina F. Callaghan^b, Konrad Wagstyl^d, Timothy Rittman^d, Roger Tait^d, Cindy Ooi^d, John Suckling^{a,e,f}, Becky Inkster^a, Peter Fonagy^g, Raymond J. Dolan^{b,h}, Peter B. Jones^{a,e}, Ian M. Goodyer^{a,e}, the NSPN Consortiumⁱ, and Edward T. Bullmore^{a,e,f,j}

^aDepartment of Psychiatry, University College London, London WC1N 3BG, United Kingdom; ^bDepartment of Clinical Neuroscience, University College London, London WC1N 3BG, United Kingdom; ^cNational Health Service Research and Development, University College London, London WC1N 3BG, United Kingdom; ^dCentre for Ageing and Neuroscience, University College London, London WC1N 3BG, United Kingdom; ^eDevelopmental Psychiatry, University College London, London WC1N 3BG, United Kingdom; ^fCentre for Ageing and Neuroscience, University College London, London WC1N 3BG, United Kingdom; ^gUniversity College London, London WC1N 3BG, United Kingdom; ^hDepartment of Psychology, University College London, London WC1N 3BG, United Kingdom; ⁱNSPN Consortium, United Kingdom; ^jEdmond and Musée M�dane Institute, United Kingdom

Molecular Psychiatry (2019) 24:1053–1064
<https://doi.org/10.1038/s41380-018-0023-7>

ARTICLE

How does human brain structure used MRI to measure cortex in 297 population volunteers? Estimated that association cortices than primary cortical areas have faster rates of shrinkage and growth. Age-related increases approximately at the internal cortical myelination and shrinkage associated with a dorsomedial enriched for synaptic, oligodendroglial genes. Topologically efficient brain anatomical network had and were associated with overall profile as cortical consolidation brain maturation increases.

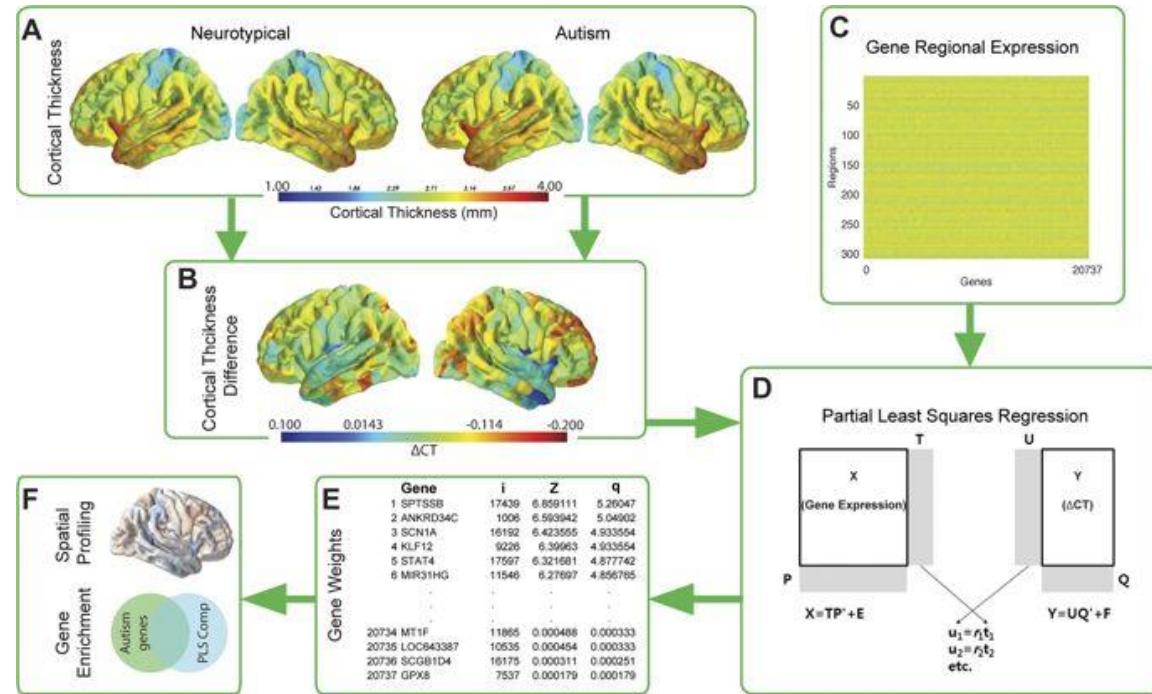
Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism

Rafael Romero-García¹ · Varun Warrier² · Edward T. Bullmore^{1,3,4} · Simon Baron-Cohen^{2,5} · Richard A. I. Bethlehem^{1,2}

Received: 24 October 2017 / Revised: 27 November 2017 / Accepted: 15 December 2017 / Published online: 26 February 2018
© The Author(s) 2018. This article is published with open access

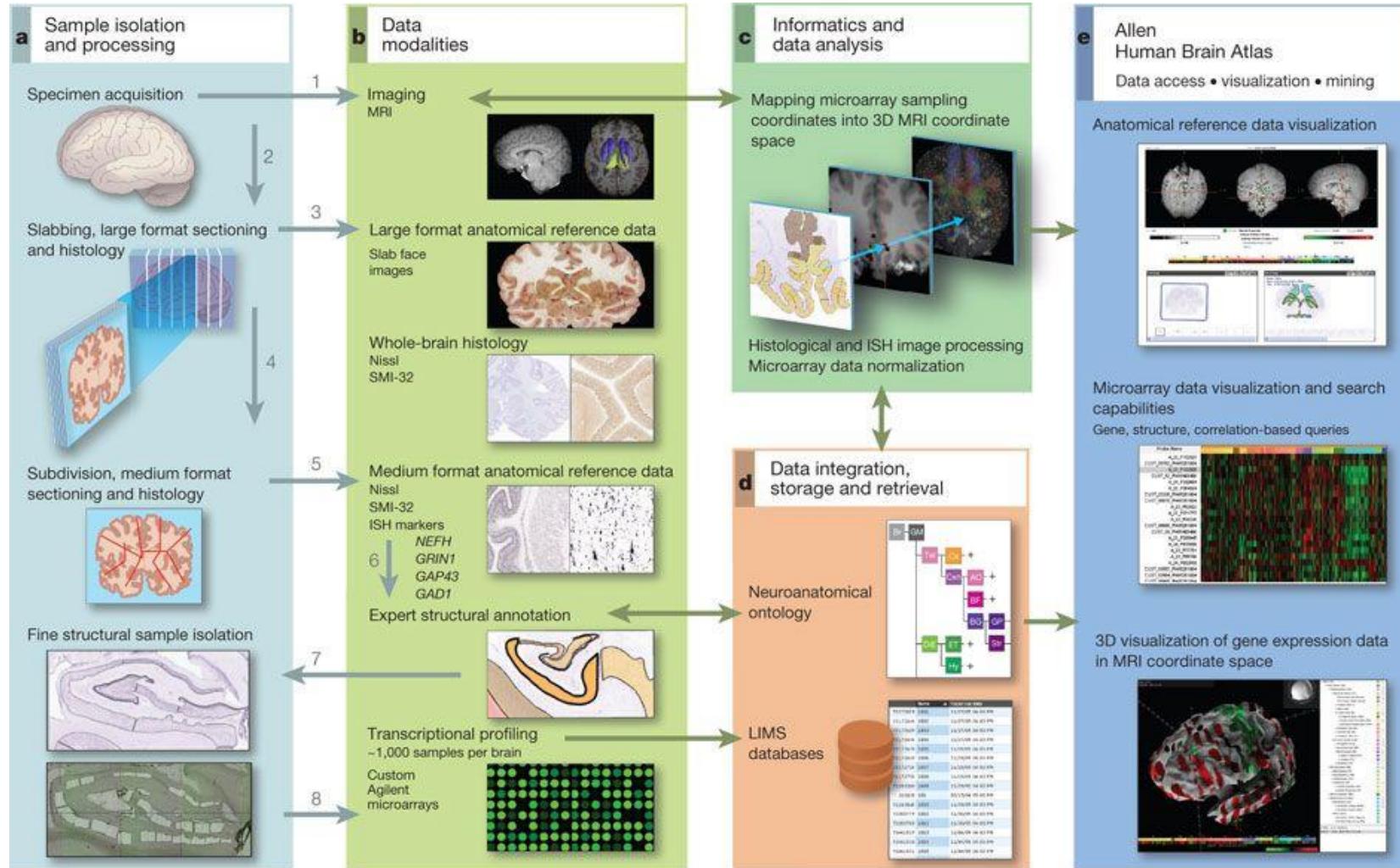
Abstract

Differences in cortical morphology—in particular, cortical volume, thickness and surface area—have been reported in individuals with autism. However, it is unclear what aspects of genetic and transcriptomic variation are associated with these differences. Here we investigate the genetic correlates of global cortical thickness differences (ΔCT) in children with autism. We used Partial Least Squares Regression (PLSR) on structural MRI data from 548 children (166 with autism, 295 neurotypical children and 87 children with ADHD) and cortical gene expression data from the Allen Institute for Brain Science to identify genetic correlates of ΔCT in autism. We identify that these genes are enriched for synaptic transmission pathways and explain significant variation in ΔCT . These genes are also significantly enriched for genes dysregulated in the autism post-mortem cortex (Odd Ratio (OR) = 1.11, $P_{\text{corrected}} 10^{-14}$), driven entirely by downregulated genes (OR = 1.87, $P_{\text{corrected}} 10^{-15}$). We validated the enrichment for downregulated genes in two independent data sets: Validation 1 (OR = 1.44, $P_{\text{corrected}} = 0.004$) and Validation 2 (OR = 1.30; $P_{\text{corrected}} = 0.001$). We conclude that transcriptionally downregulated genes implicated in autism are robustly associated with global changes in cortical thickness variability in children with autism.



Cortical thickness differences in ASD associated with genes enriched for synaptic transmission pathways and downregulated in ASD post-mortem cortex

ALLEN HUMAN BRAIN ATLAS <http://human.brain-map.org/>



Use correlations in spatial patterning of gene expression and brain phenotypes to identify associated genes and mechanisms

HI-RES ALLEN HUMAN BRAIN ATLAS DATA

ARTICLE

doi:10.1038/nature11405

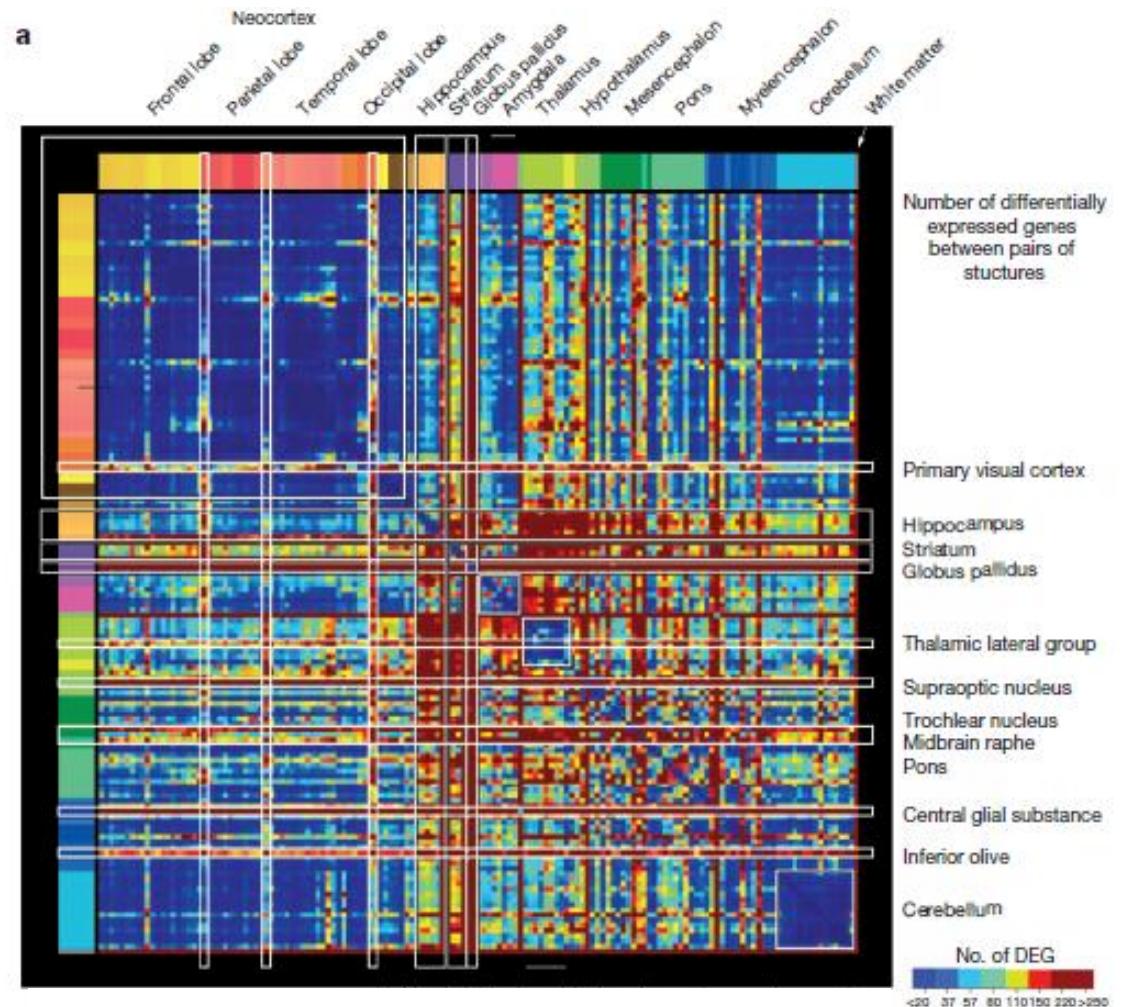
An anatomically comprehensive atlas of the adult human brain transcriptome

Michael J. Hawrylycz^{1*}, Ed S. Lein^{1*}, Angela L. Guillozet-Bongaarts¹, Elaine H. Shen¹, Lydia Ng¹, Jeremy A. Miller¹, Louie N. van de Lagemaat², Kimberly A. Smith¹, Amanda Ebbert¹, Zackery L. Riley¹, Chris Abajian¹, Christian F. Beckmann³, Amy Bernard¹, Darren Bertagnolli¹, Andrew F. Boe¹, Preston M. Cartagena⁴, M. Mallar Chakravarty^{1,5}, Mike Chapin¹, Jimmy Chong¹, Rachel A. Dalley¹, Barry David Daly⁶, Chinh Dang¹, Suvro Datta¹, Nick Dee¹, Tim A. Dolbeare¹, Vance Faber¹, David Feng¹, David R. Fowler⁷, Jeff Goldy¹, Benjamin W. Gregor¹, Zeb Haradon¹, David R. Haynor⁸, John G. Hohmann¹, Steve Horvath⁹, Robert E. Howard¹, Andreas Jeromin¹⁰, Jayson M. Jochim¹, Marty Kinnunen¹, Christopher Lau¹, Evan T. Lazarz¹, Changkyu Lee¹, Tracy A. Lemon¹, Ling Li¹¹, Yang Li¹, John A. Morris¹, Caroline C. Overly¹, Patrick D. Parker¹, Sheana E. Parry¹, Melissa Reding¹, Joshua J. Royall¹, Jay Schulkin¹², Pedro Adolfo Sequeira¹³, Clifford R. Slaughterbeck¹, Simon C. Smith¹⁴, Andy J. Sodt¹, Susan M. Sunkin¹, Beryl E. Swanson¹, Marquis P. Vawter¹³, Derrick Williams¹, Paul Wohnoutka¹, H. Ronald Zielke¹⁵, Daniel H. Geschwind¹⁶, Patrick R. Hof¹⁷, Stephen M. Smith¹⁸, Christof Koch^{1,19}, Seth G. N. Grant² & Allan R. Jones¹

Neuroanatomically precise, genome-wide maps of transcript distributions are critical resources to complement genomic sequence data and to correlate functional and genetic brain architecture. Here we describe the generation and analysis of a transcriptional atlas of the adult human brain, comprising extensive histological analysis and comprehensive microarray profiling of ~900 neuroanatomically precise subdivisions in two individuals. Transcriptional regulation varies enormously by anatomical location, with different regions and their constituent cell types displaying robust molecular signatures that are highly conserved between individuals. Analysis of differential gene expression and gene co-expression relationships demonstrates that brain-wide variation strongly reflects the distributions of major cell classes such as neurons, oligodendrocytes, astrocytes and microglia. Local neighbourhood relationships between fine anatomical subdivisions are associated with discrete neuronal subtypes and genes involved with synaptic transmission. The neocortex displays a relatively homogeneous transcriptional pattern, but with distinct features associated selectively with primary sensorimotor cortices and with enriched frontal lobe expression. Notably, the spatial

HI-RES ADULT HUMAN BRAIN TRANSCRIPTOME

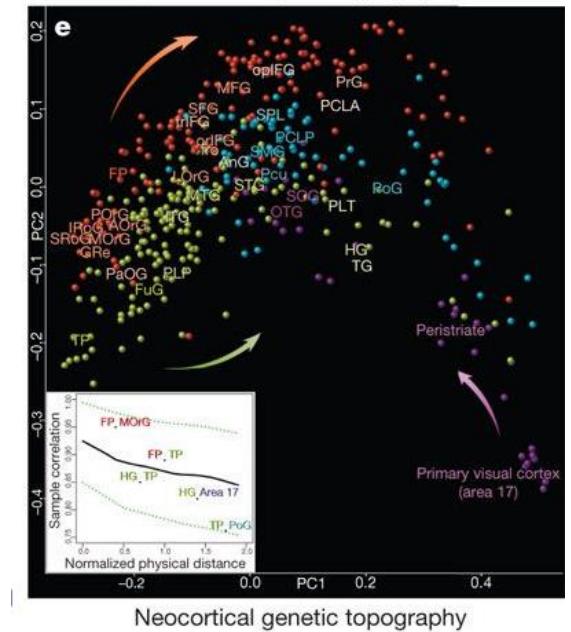
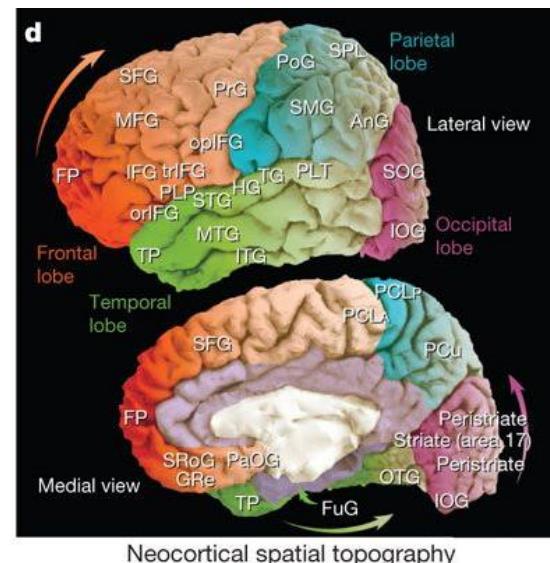
- Differentially expressed genes between structures:
 - Low variation within cortex, cerebellum
 - Primary sensory and visual cortex more distinct
 - Striatum, globus pallidum, hippocampus have more distinct profiles



Hawrylycz et al., 2012, *Nature*

HI-RES ADULT HUMAN BRAIN TRANSCRIPTOME

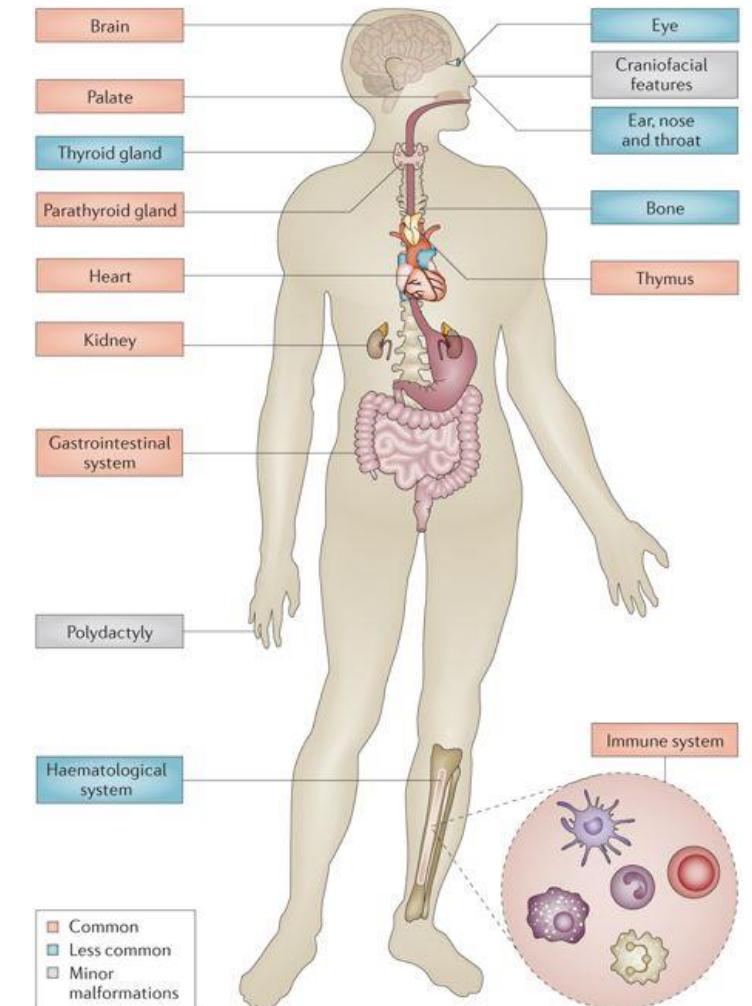
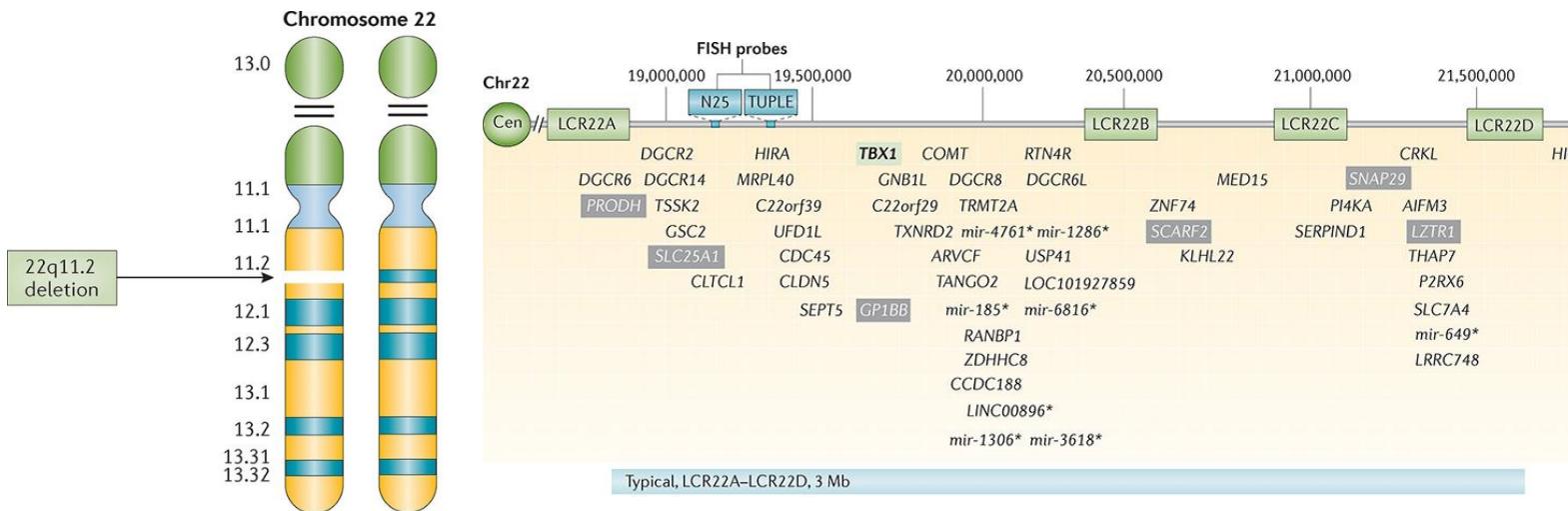
- Differentially expressed genes between structures:
 - Low variation within cortex, cerebellum
 - Primary sensory and visual cortex more distinct
 - Striatum, globus pallidum, hippocampus have more distinct profiles
- Neocortical transcriptome reflects *in vivo* spatial topography



EXAMPLE STUDY

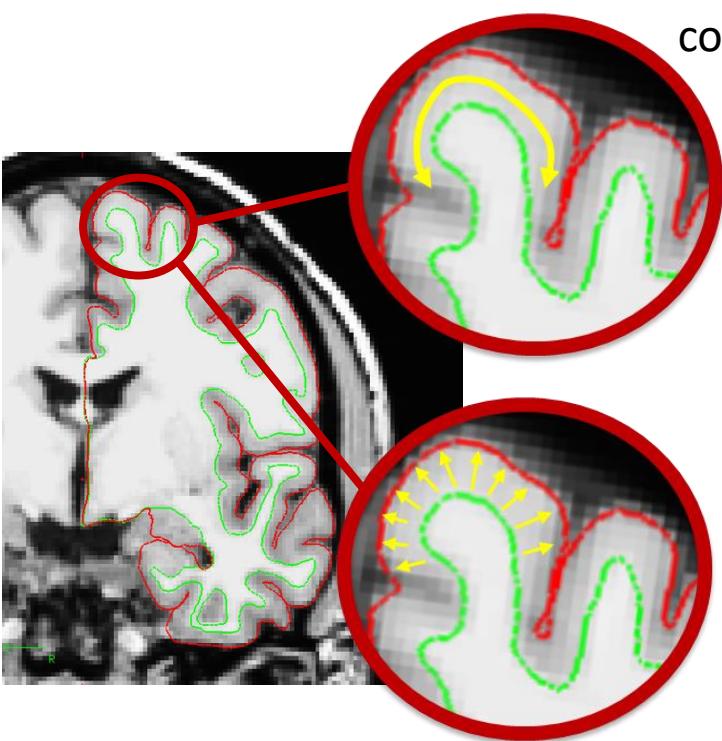
22q11.2 DELETION SYNDROME

- ~2.6 Mb deletion spans 46 protein-coding genes
- Broad Phenotype
- Psychiatric and neurodevelopmental disorders (e.g., SCZ, ASD, ID)



EXAMPLE STUDY

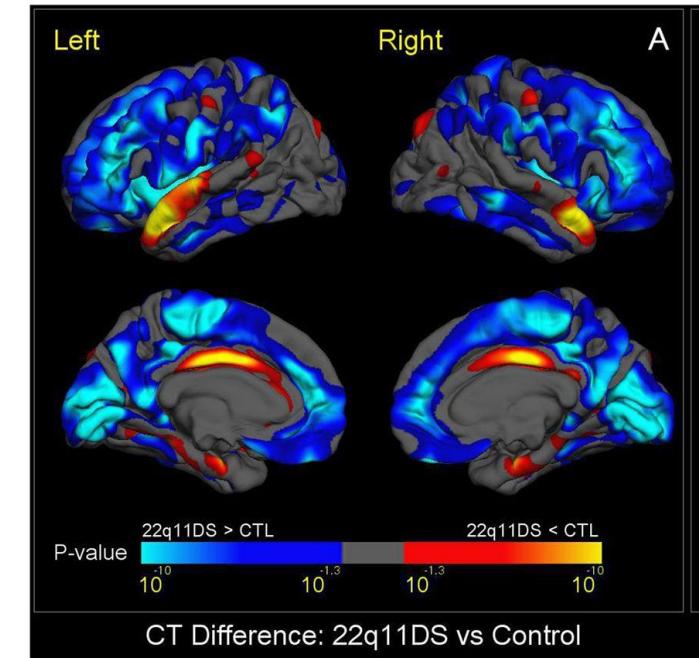
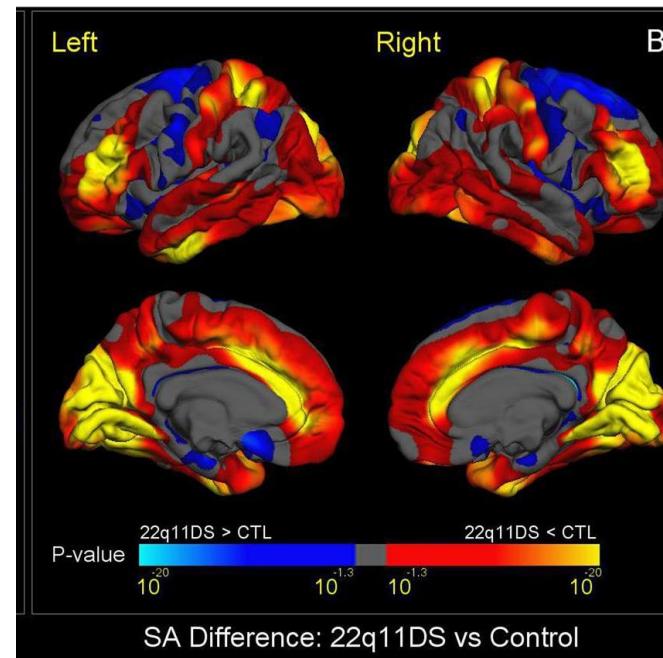
22q11.2 ENIGMA CORTICAL FINDINGS



Surface Area (SA):
total area covered by
cortex in a region

Cortical Thickness (CT):
average thickness of 6
cortical layers in region

Widespread SA reductions in
22q11DS and increased CT in
majority of regions



RESEARCH QUESTION

Can imaging transcriptomics help us prioritize genes within the 22q11.2 locus that may drive alterations in surface area and cortical thickness?

METHODS OVERVIEW

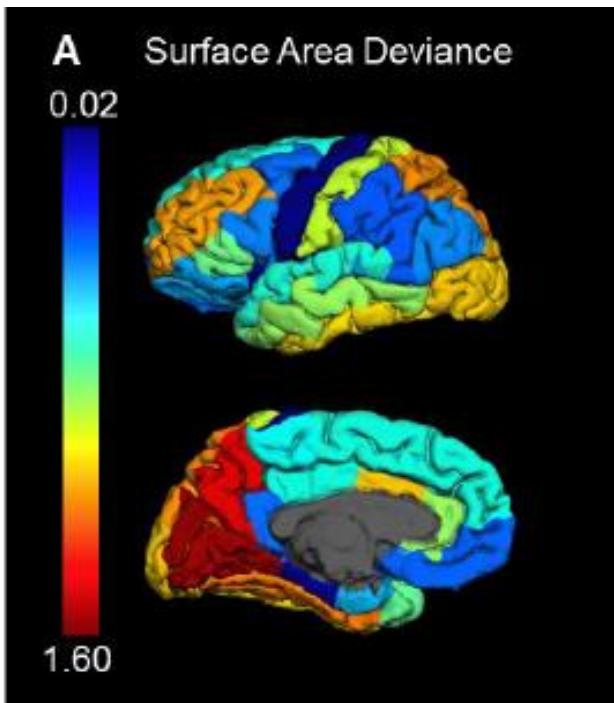
- T1 structural MRI images from 232 patients and 290 controls
 - Processed with Freesurfer - SA and CT for 34 Desikan-Killiany regions (Sun et al., 2020)
- Severity of 22q11DS Z-score deviance per region
 - Difference in mean SA or CT per group/control standard deviation
- Allen Human Brain Atlas (AHBA)
 - 10,318 protein-coding, brain-expressed genes summarized to DK regions (French & Paus, 2015)
- Correlate severity of SA or CT deviance across regions vs. regional expression of each protein-coding, brain expressed 22q11.2 gene
 - p -values from empirical distribution of all protein-coding, brain expressed AHBA genes
- Partial least square regression (PLSR) to identify top principal component maximizing regional spatial covariance between SA or CT deviance and gene expression
 - Followed by gene ontology (GO) term enrichment analysis of top loading genes on PC1

EXAMPLE STUDY

WIDESPREAD SURFACE AREA REDUCTIONS IN 22Q11DS

22q11DS SA reductions most prominent in midline posterior regions & detectable by childhood

Controls >
22q11DS



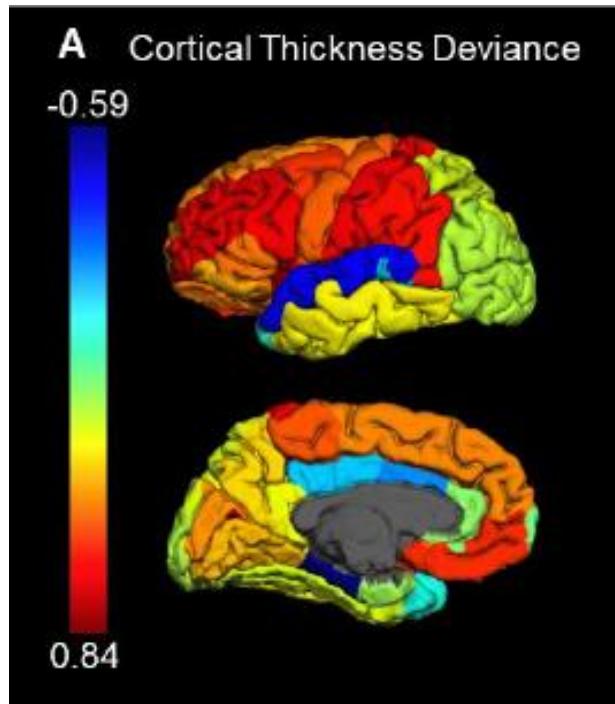
Region	F value	FDR p	HC mean (SD)	22q11DS mean (SD)	22q11DS ΔSA Z-Score
<i>lh_bankssts</i>	39.89	9.36E-10	981.28 (173.51)	887.42 (162.53)	0.54
<i>lh_caudalanteriorcingulate</i>	214.51	4.39E-40	627.8 (125.83)	487.36 (82.87)	1.12
<i>lh_caudalmiddlefrontal</i>	20.28	9.7E-06	2261.02 (396.98)	2098.3 (426.29)	0.41
<i>lh_cuneus</i>	360.80	5.76E-60	1413.03 (196.53)	1098.31 (176.99)	1.60
<i>lh_entorhinal</i>	39.49	1.08E-09	335.47 (76.88)	293.62 (74.01)	0.54
<i>lh_fusiform</i>	187.42	5.77E-36	3198.6 (391.2)	2714.94 (413.13)	1.24
<i>lh_inferiorparietal</i>	26.65	4.72E-07	4325.56 (604.88)	4048.42 (615.26)	0.46
<i>lh_inferiotemporal</i>	132.65	5.78E-27	2881.38 (441.43)	2407.97 (496.41)	1.07
<i>lh_isthmuscingulate</i>	21.42	5.88E-06	898.72 (162.01)	832.87 (160.98)	0.41
<i>lh_lateraloccipital</i>	158.51	2.45E-31	4592.46 (557.87)	3975.36 (554.68)	1.11
<i>lh_lateralorbitofrontal</i>	31.58	4.42E-08	2148.21 (299.9)	1998.06 (307.57)	0.50
<i>lh_lingual</i>	315.74	2.55E-54	3036.16 (379.58)	2442.35 (379.16)	1.56
<i>lh_medialorbitofrontal</i>	21.23	6.23E-06	1503.51 (244.82)	1406.25 (233)	0.40
<i>lh_middletemporal</i>	84.38	2.54E-18	2763.24 (372.12)	2441.76 (426.76)	0.86
<i>lh_parahippocampal</i>	0.97	0.346232	672.46 (100.71)	663.91 (96.25)	0.08
<i>lh_paracentral</i>	58.46	1.81E-13	1277.3 (197.62)	1149.3 (180.12)	0.65
<i>lh_parsopercularis</i>	26.58	4.72E-07	1708.55 (271.1)	1587.91 (258.79)	0.45
<i>lh_parsorbitalis</i>	32.07	3.64E-08	589.23 (81.07)	548.01 (84.55)	0.51
<i>lh_parstriangularis</i>	81.21	8.76E-18	1376.83 (192.32)	1222.77 (196.27)	0.80
<i>lh_pericalcarine</i>	234.18	5.52E-43	1387.16 (225.07)	1101.39 (194.42)	1.27
<i>lh_postcentral</i>	131.37	8.78E-27	4042.32 (485.77)	3586.84 (403.68)	0.94
<i>lh_posteriorcingulate</i>	62.54	3.35E-14	1141.39 (160.94)	1031.7 (153)	0.68
<i>lh_precentral</i>	0.06	0.799853	4486.01 (486.16)	4474.28 (570.24)	0.02
<i>lh_precuneus</i>	296.95	6.4E-52	3698.12 (431.73)	3064.5 (398.85)	1.47
<i>lh_rostralanteriorcingulate</i>	110.69	3.8E-23	675.52 (154.59)	541.23 (131.81)	0.87
<i>lh_rostralmiddlefrontal</i>	173.85	7.99E-34	5762.14 (675.66)	4977.04 (676.43)	1.16
<i>lh_superiorfrontal</i>	60.70	6.86E-14	6704.59 (773.24)	6177.75 (760.68)	0.68
<i>lh_superiorparietal</i>	205.15	1.04E-38	5566.71 (659.58)	4771.96 (590.76)	1.20
<i>lh_superiortemporal</i>	62.10	3.84E-14	3584.36 (411.59)	3296.29 (419.25)	0.70
<i>lh_supramarginal</i>	18.14	2.67E-05	3704.31 (543.53)	3497.84 (558.66)	0.38
<i>lh_frontalpole</i>	19.33	1.51E-05	205.1 (33.55)	191.99 (34.25)	0.39
<i>lh_temporalpole</i>	82.79	4.71E-18	425.6 (62.22)	376.49 (60.09)	0.79
<i>lh_transversetemporal</i>	52.17	3.1E-12	433.91 (70.69)	387.06 (77.18)	0.66
<i>lh_insula</i>	0.53	0.479074	1835.47 (240.73)	1818.75 (281.51)	0.07

EXAMPLE STUDY

THICKER CORTEX IN MOST REGIONS IN 22Q11DS

Subtler increases in CT in most regions, with focal thinning in caudal anterior cingulate, superior temporal cortex, & parahippocampus

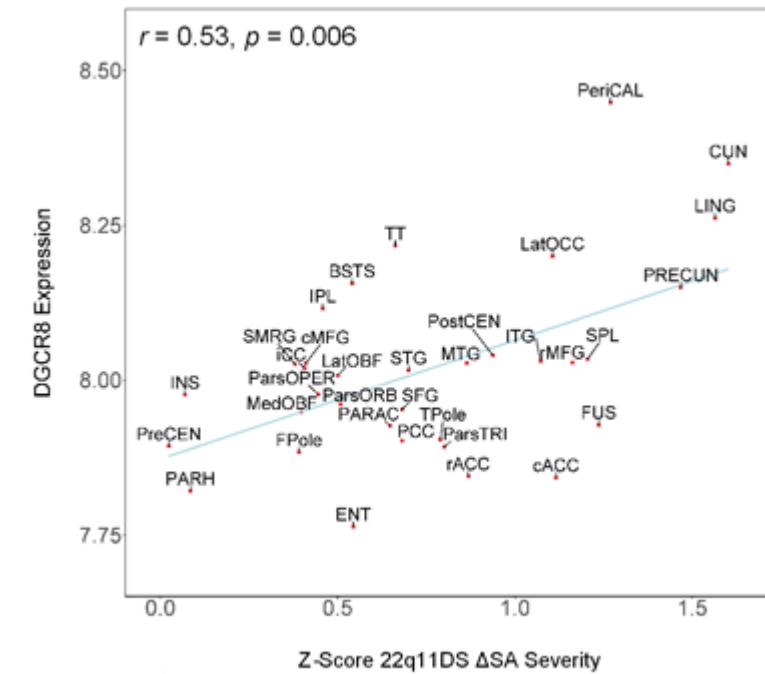
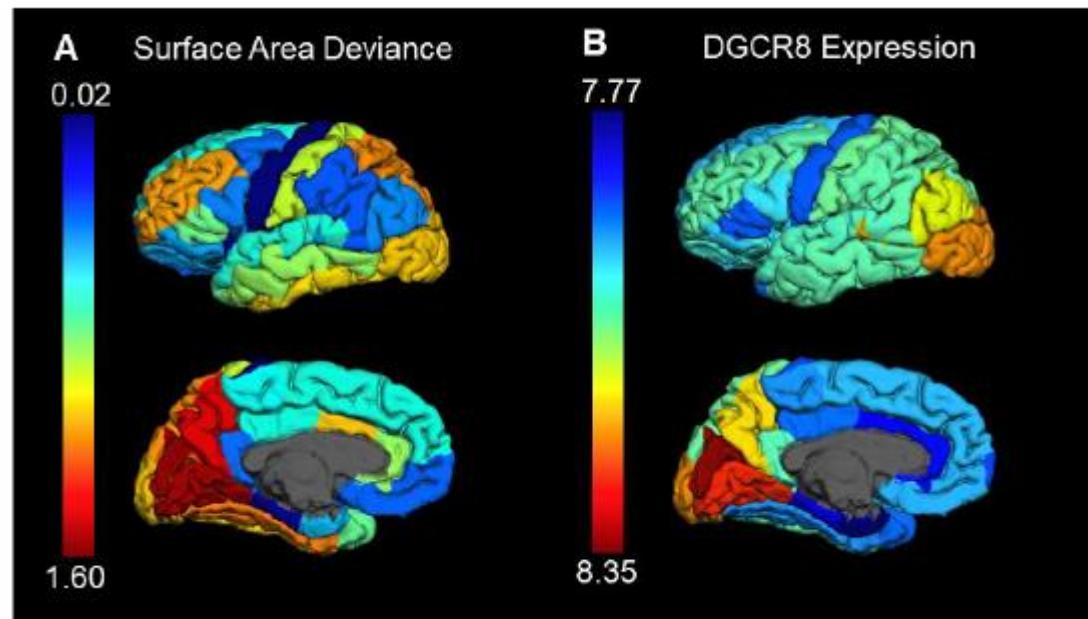
22q11DS >
Controls



Region	F value	FDR p	HC mean (SD)	22q11DS mean (SD)	22q11DS Z-Score
lh_bankssts	0.70	4.15E-01	3.23 (0.17)	3.22 (0.2)	-0.08
lh_caudalanteriorcingulate	4.35	4.71E-02	3.74 (0.24)	3.69 (0.26)	-0.19
lh_caudalmiddlefrontal	44.05	3.05E-10	2.98 (0.14)	3.06 (0.15)	0.61
lh_cuneus	30.76	1.13E-07	2.5 (0.15)	2.57 (0.15)	0.48
lh_entorhinal	4.12	5.21E-02	3.52 (0.35)	3.59 (0.4)	0.19
lh_fusiform	8.86	4.32E-03	3.31 (0.14)	3.34 (0.15)	0.27
lh_inferiorparietal	6.36	1.63E-02	3.13 (0.14)	3.17 (0.15)	0.23
lh_inferiotemporal	13.11	5.22E-04	3.4 (0.17)	3.46 (0.18)	0.32
lh_isthmuscingulate	12.21	7.98E-04	3.43 (0.19)	3.48 (0.19)	0.30
lh_lateraloccipital	5.96	1.95E-02	2.73 (0.13)	2.76 (0.13)	0.21
lh_lateralorbitofrontal	31.84	7.20E-08	3.41 (0.17)	3.49 (0.18)	0.51
lh_lingual	22.54	5.32E-06	2.67 (0.13)	2.73 (0.13)	0.42
lh_medialorbitofrontal	47.40	8.14E-11	3.28 (0.18)	3.39 (0.19)	0.62
lh_middletemporal	15.56	1.54E-04	3.55 (0.18)	3.61 (0.16)	0.34
lh_parahippocampal	46.81	9.37E-11	3.06 (0.3)	2.89 (0.28)	-0.59
lh_paracentral	39.58	2.28E-09	2.96 (0.16)	3.05 (0.16)	0.56
lh_parsopercularis	61.62	2.71E-13	3.12 (0.14)	3.22 (0.15)	0.70
lh_parsorbitalis	24.14	2.55E-06	3.43 (0.24)	3.54 (0.24)	0.44
lh_parstriangularis	34.39	2.28E-08	3.13 (0.17)	3.22 (0.18)	0.52
lh_pericalcarine	67.28	3.16E-14	1.94 (0.14)	2.05 (0.16)	0.77
lh_postcentral	56.50	1.69E-12	2.55 (0.13)	2.64 (0.15)	0.71
lh_posteriorcingulate	1.57	2.24E-01	3.48 (0.17)	3.46 (0.15)	-0.10
lh_precentral	36.18	1.05E-08	2.86 (0.13)	2.93 (0.15)	0.55
lh_precuneus	18.24	4.15E-05	3.11 (0.13)	3.16 (0.14)	0.38
lh_rostralanteriorcingulate	2.88	1.02E-01	3.75 (0.24)	3.78 (0.28)	0.16
lh_rostralmiddlefrontal	56.65	1.69E-12	2.98 (0.14)	3.08 (0.17)	0.73
lh_superiorfrontal	30.60	1.14E-07	3.49 (0.15)	3.57 (0.15)	0.50
lh_superiorparietal	9.63	2.98E-03	2.63 (0.14)	2.66 (0.13)	0.27
lh_supriortemporal	22.01	6.55E-06	3.34 (0.17)	3.27 (0.16)	-0.40
lh_supramarginal	55.53	2.19E-12	3.21 (0.15)	3.31 (0.16)	0.66
lh_frontalpole	2.31	1.41E-01	3.66 (0.35)	3.71 (0.37)	0.14
lh_temporalpole	0.00	9.72E-01	3.53 (0.36)	3.53 (0.41)	0.00
lh_transversetemporal	4.01	5.38E-02	3.16 (0.24)	3.12 (0.23)	-0.17
lh_insula	102.19	1.60E-20	3.76 (0.17)	3.9 (0.15)	0.85

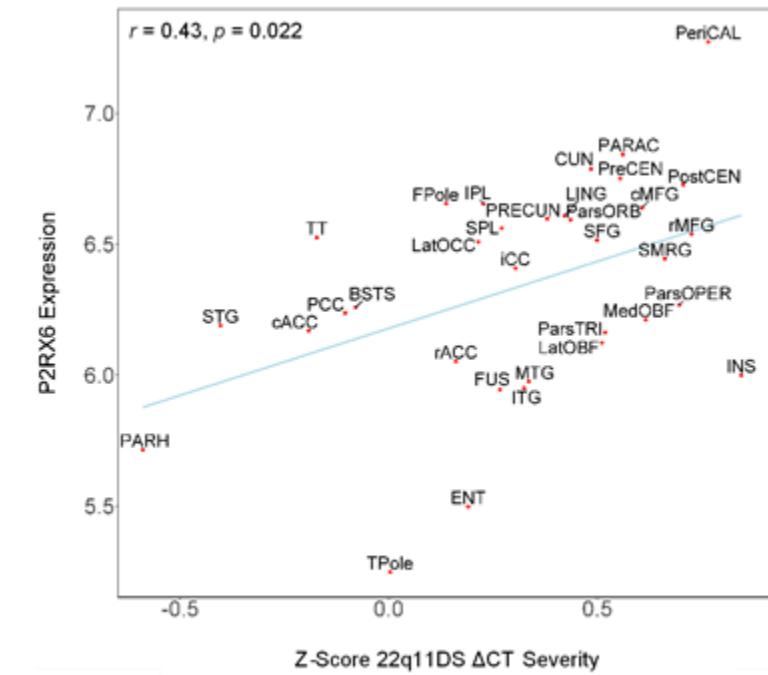
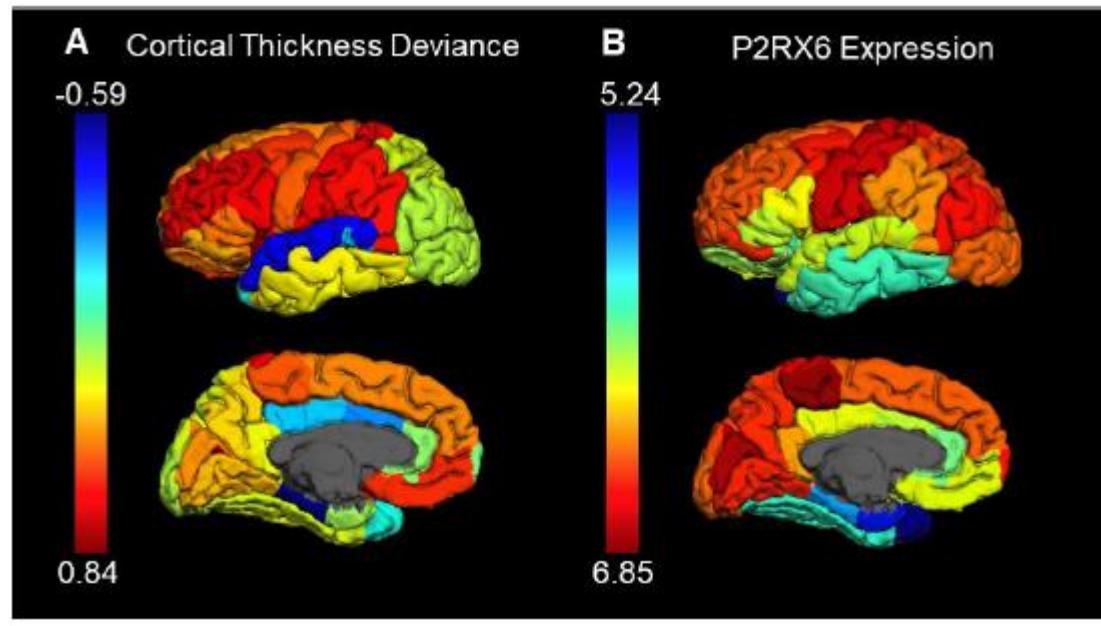
PRIORITIZED DRIVER OF SURFACE AREA REDUCTIONS

Spatial variation in severity of surface area reductions and DGCR8 expression are correlated



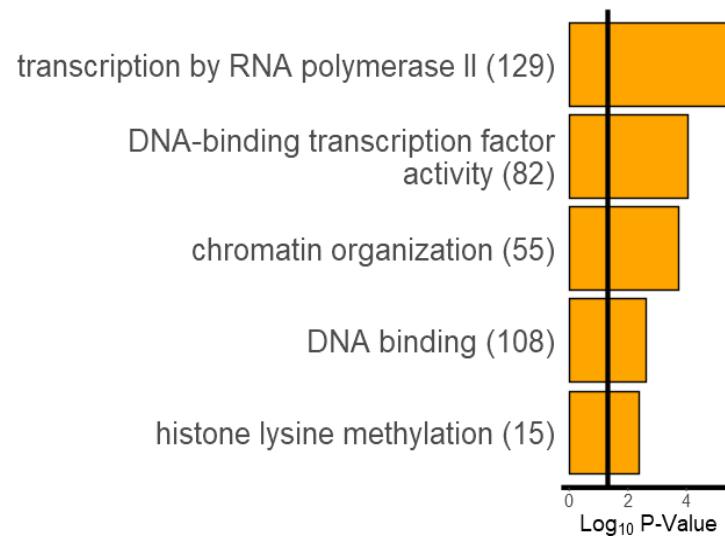
PRIORITIZED DRIVER OF ALTERED CORTICAL THICKNESS

Spatial variation in cortical thickness alterations
and P2RX6 expression are correlated



SIMILAR PRIORITIZED GENES WITH PLSR FOR SA

First principal component from partial least squares regression across AHBA genes and SA deviance prioritized similar 22q11.2 genes and implicates gene regulatory processes

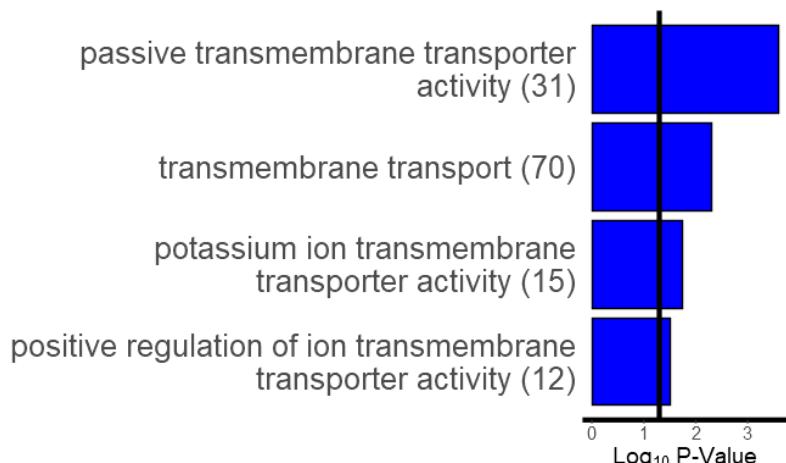


Gene	Bootstrap Mean Loading Weight (SD)	Bootstrap Loading Weight Z-Score	Bootstrap Loading Weight Z-Score AHBA Rank	Bootstrap Loading Weight Z-Score P _{AHBA}
DGCR8	0.0138 (0.0023)	5.8986	0.9983	0.0017
AIFM3	0.0167 (0.005)	3.3301	0.9565	0.0435
SCARF2	0.0062 (0.0034)	1.8487	0.8409	0.1591
CLDN5	0.0061 (0.0035)	1.7427	0.8290	0.1710
DGCR2	0.0046 (0.003)	1.5740	0.8064	0.1936
P2RX6	0.0142 (0.0103)	1.3831	0.7781	0.2219
TANGO2	0.0053 (0.0042)	1.2422	0.7573	0.2427
RANBP1	0.0023 (0.0026)	0.8670	0.6929	0.3071
UFD1	0.0008 (0.0035)	0.2446	0.5712	0.4288
ARVCF	0.0009 (0.0075)	0.1262	0.5467	0.4533
HIRA	0.0004 (0.004)	0.1076	0.5439	0.4561
COMT	0.0005 (0.0047)	0.0983	0.5417	0.4583
MED15	-0.0007 (0.0022)	-0.3165	0.4566	0.5434
PRODH	-0.0043 (0.0113)	-0.3781	0.4446	0.5554
GNB1L	-0.0014 (0.003)	-0.4686	0.4270	0.5730
SLC25A1	-0.0031 (0.0042)	-0.7447	0.3704	0.6296
MRPL40	-0.0039 (0.0026)	-1.4997	0.2340	0.7660
RIMBP3	-0.0099 (0.0057)	-1.7253	0.2008	0.7992
PI4KA	-0.0051 (0.0029)	-1.7364	0.1987	0.8013
KLHL22	-0.0053 (0.003)	-1.7545	0.1967	0.8033
GP1BB	-0.0082 (0.0042)	-1.9298	0.1732	0.8268
RTN4R	-0.008 (0.004)	-2.0103	0.1628	0.8372
DGCR6	-0.0068 (0.0031)	-2.1921	0.1418	0.8582
SEPT5	-0.0114 (0.0039)	-2.9506	0.0788	0.9212
C22orf39	-0.0094 (0.003)	-3.1750	0.0646	0.9354
DGCR6L	-0.0079 (0.0021)	-3.7103	0.0391	0.9609
SNAP29	-0.0095 (0.0018)	-5.2629	0.0074	0.9926
SLC7A4	-0.0199 (0.0035)	-5.6512	0.0044	0.9956

EXAMPLE STUDY

SIMILAR PRIORITIZED GENES WITH PLSR FOR CT

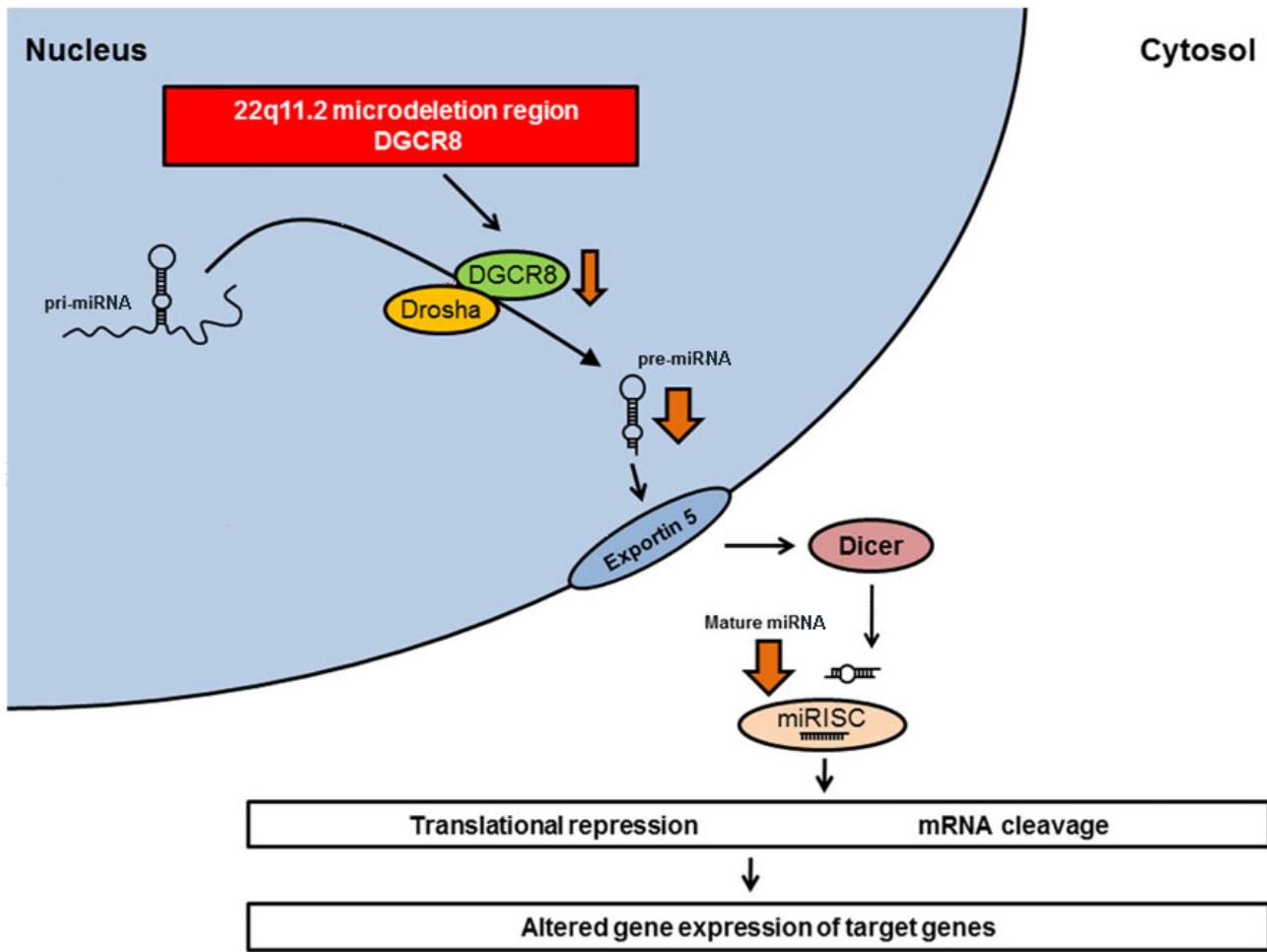
First principal component from partial least squares regression across all AHBA genes and CT deviance prioritized similar 22q11.2 genes and implicates transmembrane and ion transmembrane transport



Gene	Bootstrap Mean Loading Weight (SD)	Bootstrap Loading Weight Z-Score	Bootstrap Loading Weight Z-Score AHBA Rank	Bootstrap Loading Weight Z-Score P _{AHBA}
P2RX6	0.0332 (0.0002)	4.6932	0.9933	0.0067
GNB1L	0.0071 (0.0001)	2.1180	0.8792	0.1208
TANGO2	0.0068 (0.0001)	1.6862	0.8315	0.1685
AIFM3	0.0115 (0.0002)	1.6488	0.8272	0.1728
DGCR8	0.0061 (0.0001)	1.5251	0.8066	0.1934
RANBP1	0.0041 (0.0001)	1.4578	0.7963	0.2037
SCARF2	0.005 (0.0001)	1.3994	0.7874	0.2126
CLDN5	0.0048 (0.0001)	1.3078	0.7744	0.2256
MRPL40	0.0025 (0.0001)	1.1663	0.7498	0.2502
MED15	0.0018 (0.0001)	0.8498	0.6948	0.3052
DGCR2	0.0034 (0.0001)	0.8380	0.6920	0.3080
HIRA	0.0026 (0.0001)	0.7343	0.6692	0.3308
KLHL22	0.0019 (0.0001)	0.4686	0.6167	0.3833
COMT	0.0016 (0.0001)	0.3959	0.5994	0.4006
RTN4R	0.0022 (0.0002)	0.3706	0.5932	0.4068
ARVCF	0.0012 (0.0002)	0.1725	0.5440	0.4560
RIMBP3	0.0005 (0.0002)	0.0727	0.5182	0.4818
UFD1	0.0002 (0.0001)	0.0645	0.5154	0.4846
PI4KA	0 (0.0001)	-0.0052	0.4975	0.5025
DGCR6L	-0.0013 (0.0001)	-0.4193	0.3977	0.6023
SEPT5	-0.0041 (0.0002)	-0.7666	0.3247	0.6753
DGCR6	-0.0037 (0.0001)	-1.2169	0.2396	0.7604
GP1BB	-0.0079 (0.0002)	-1.3150	0.2233	0.7767
PRODH	-0.0165 (0.0003)	-1.8444	0.1439	0.8561
C22orf39	-0.0068 (0.0001)	-1.8653	0.1411	0.8589
SLC25A1	-0.009 (0.0001)	-2.0659	0.1170	0.8830
SLC7A4	-0.0128 (0.0002)	-2.1730	0.1035	0.8965
SNAP29	-0.0074 (0.0001)	-2.9221	0.0452	0.9548

EXAMPLE STUDY

DGCR8 GENE REGULATORY ROLE VIA MIRNA BIOGENESIS



miRNAs play a key gene regulatory role by repressing protein translation for ~50% of genes

Gene targets of 59 miRNAs down-regulated in DGCR8^{+/−}-mouse cortex modulate cell cycle and are enriched for early brain development

nature
genetics



Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model

Adapted from Forstner et al., 2013, *Front Mol Neurosci*

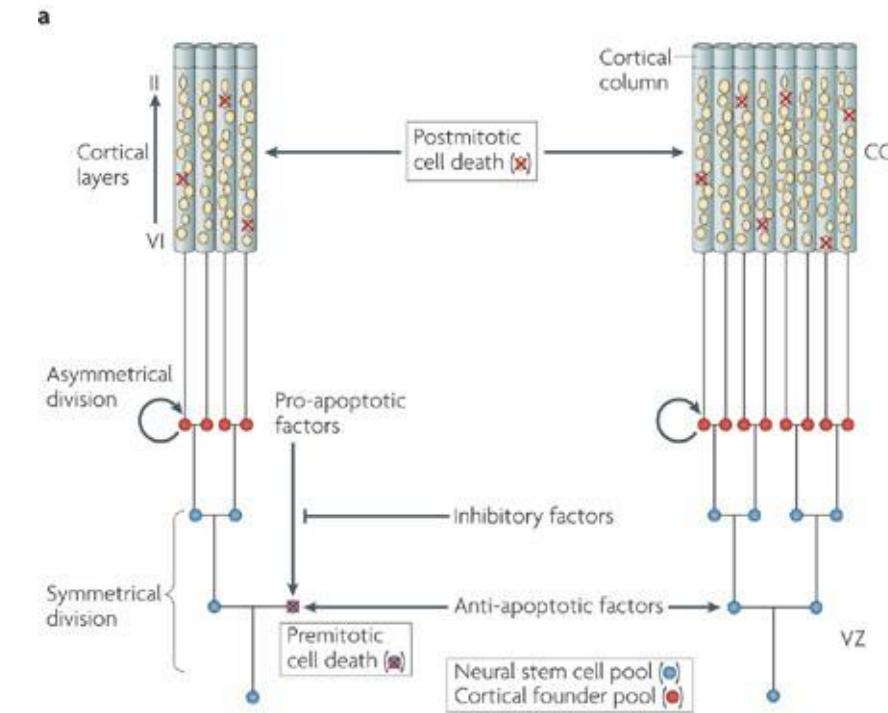
Kimberly L Stark^{1,6}, Bin Xu^{2,6}, Anindya Bagchi³, Wen-Sung Lai², Hui Liu¹, Ruby Hsu⁴, Xiang Wan⁵, Paul Pavlidis⁵, Alea A Mills³, Maria Karayiorgou¹ & Joseph A Gogos^{2,4}

EXAMPLE STUDY

SUMMARY

- Cell cycle parameters critical for cortex size (Rakic, 1988; Dehay & Kennedy, 2007; Grasby et al., 2020)
 - miRNAs downregulated by DGCR8+/- regulate cell proliferation
- DGCR8 involved in neurodevelopmental processes thought to fundamentally modulate brain size
- P2RX6 encodes ATP-gated ion channel (Khakh & North, 2012; Motahari et al., 2019)
- Approach highlights a well-characterized gene (DGCR8) and relatively understudied gene (P2RX6) as potential drivers of cortical abnormalities

Radial unit hypothesis of cortical neurogenesis: cell proliferation vs. death defines surface expansion



Rakic, 2009, *Nat Rev Neurosci*

IMAGING TRANSCRIPTOMICS RESOURCES

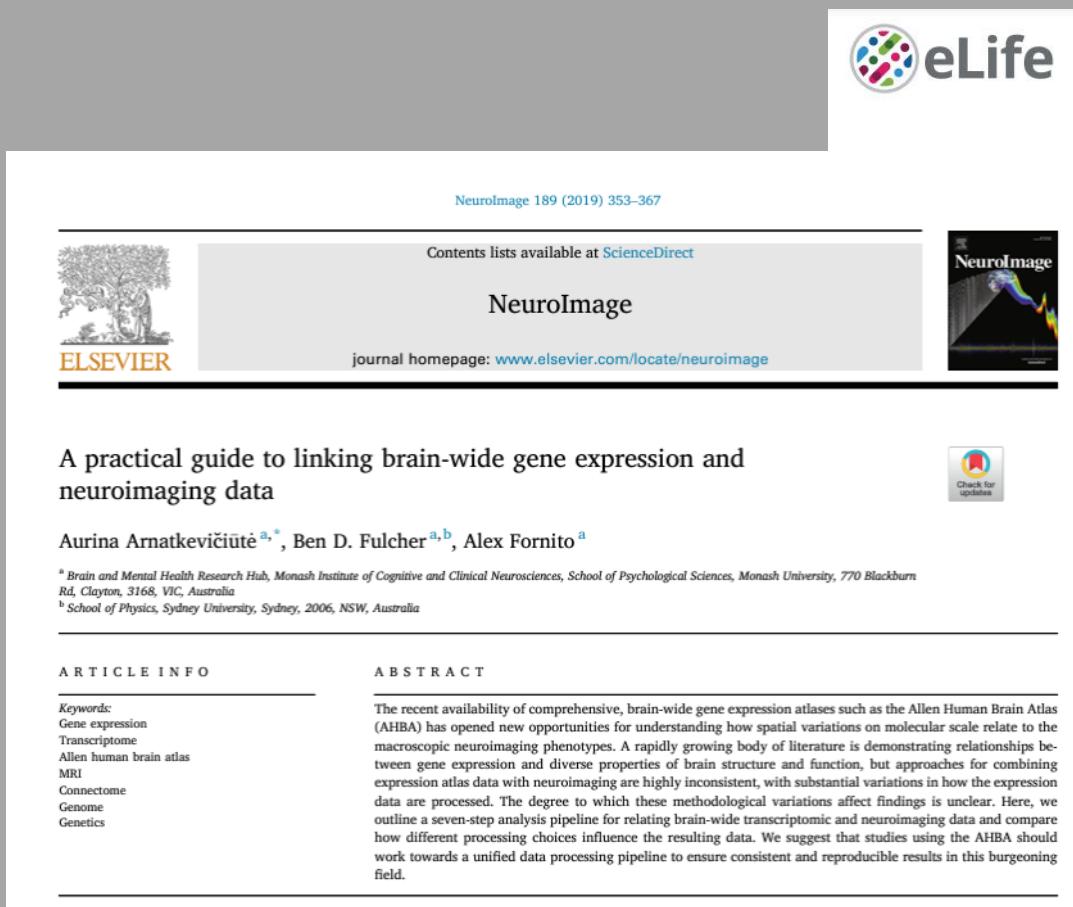
Allen Human Brain Atlas

<http://human.brain-map.org/>

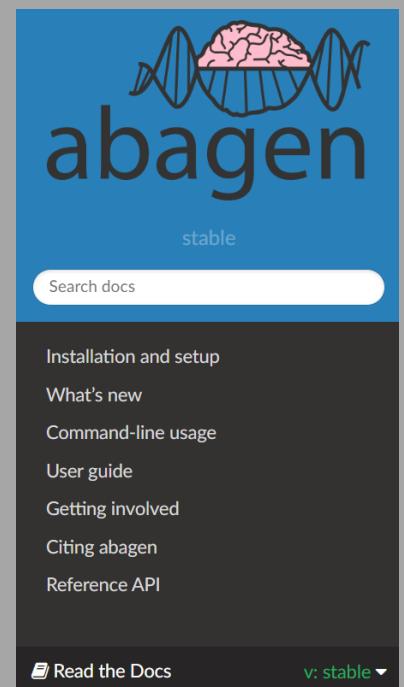
The screenshot shows the homepage of the Allen Human Brain Atlas Data Portal. At the top, there's a navigation bar with tabs for HOME, GET STARTED, HUMAN BRAIN (which is currently selected), and TOOLS. Below the navigation bar, there's a search bar with the placeholder "Enter Gene Name, Gene Symbol, NCBI Accession Number or Entrez Gene ID". To the left of the search bar are three radio buttons: "Gene Search" (selected), "Differential Search", and "Mouse Differential Search". To the right of the search bar is a "Search..." button. Below the search bar, there's a checkbox labeled "Show exact matches only". The main content area features a "Browse by Gene Category" section with a tag cloud of biological terms. Some prominent terms include "Alzheimer disease", "G-protein coupled receptor", "Hedgehog signaling pathway", "MAPKKK cascade", "Neurodegenerative", "Nervous system development", "Neurological system process", "NF-kappaB Cascade", "Notch signaling pathway", "Oxidative stress response", "Parkinson disease", "PDGF signaling pathway", "Potassium channel", "Schizophrenia associated", "Seizure", "Sodium channel", "Stress response", and "TGF-beta signaling". To the right of this section is a sidebar titled "Allen Human Brain Atlas" with links for "Search the data", "Find Correlates", "About the Microarray data", and a "Download" section. The bottom of the sidebar features a "Data Summary" table with two rows of information.

Category	Value
Number of genes	> 62,000 gene probes per profile
Number of samples	~ 500 samples per hemisphere across

IMAGING TRANSCRIPTOMICS RESOURCES



The screenshot shows a journal article from *NeuroImage* published in *eLife*. The article is titled "Standardizing workflows in imaging transcriptomics with the abagen toolbox". The authors are Ross D Markello, Aurina Arnatkevičiūtė, Jean-Baptiste Poline, Ben D Fulcher, Alex Fornito, and Bratislav Misic. The article is from the McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Canada; School of Psychological Sciences & Monash Biomedical Imaging, Monash University, Clayton, Australia; and School of Physics, University of Sydney, Sydney, Australia. The abstract discusses the development of the abagen toolbox, an open-access software package for working with transcriptomic data, and its application to the Allen Human Brain Atlas. The toolbox allows researchers to examine how methodological variability influences research outcomes using three prototypical analyses. The results show that choice of pipeline has a large impact on research findings, with parameters commonly varied in the literature influencing correlations between derived gene expression and other imaging phenotypes by as much as $\rho \geq 1.0$.



The screenshot shows the abagen project landing page. The header features the *abagen* logo with a brain and DNA helix icon, and links for "TOOLS AND RESOURCES", an open access symbol, and a Creative Commons license symbol. The main content area includes a search bar, navigation links for "Installation and setup", "What's new", "Command-line usage", "User guide", "Getting involved", "Citing abagen", and "Reference API". At the bottom, there are links for "Read the Docs" and "v: stable".

<https://abagen.readthedocs.io//>

IMAGING TRANSCRIPTOMICS RESOURCES

Introduction to the **pls** Package

Bjørn-Helge Mevik
University Center for Information Technology, University of Oslo
Norway

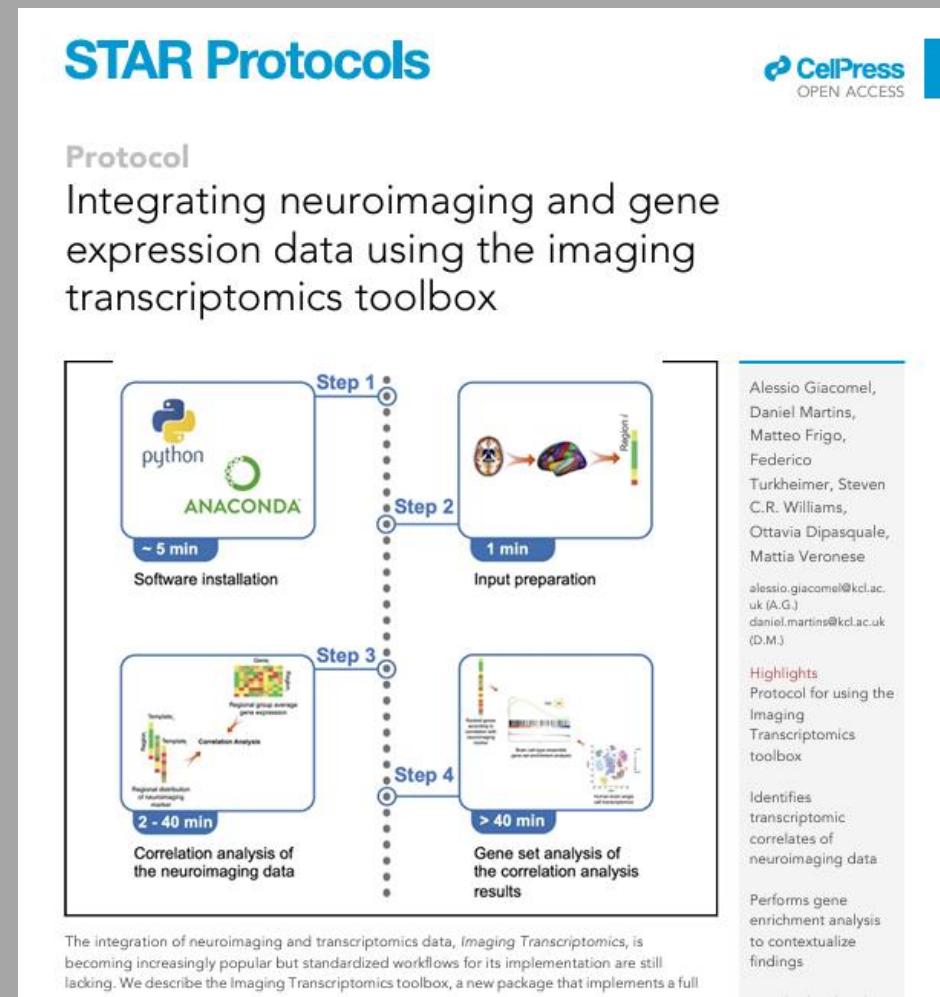
Ron Wehrens
Biometris, Wageningen University & Research
The Netherlands

November 17, 2023

Abstract

The **pls** package implements Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR) in R, and is freely available from the CRAN website, licensed under the Gnu General Public License (GPL).

The user interface is modelled after the traditional formula interface, as exemplified by `lm`. This was done so that people used to R would not have to learn yet another interface, and also because we believe the formula interface is a good way of working interactively with models. It thus has methods for generic functions like `predict`, `update` and `coef`. It also has more specialised functions like `scores`, `loadings` and `RMSEP`, and a flexible cross-validation system. Visual inspection and assessment is important in chemometrics, and the **pls** package has a number of plot functions for plotting scores, loadings, predictions, coefficients and RMSEP estimates.



GENETICS/TRANSCRIPTOMICS RESOURCES

Gene Ontology:

Enrichr

(<https://amp.pharm.mssm.edu/Enrichr/>)

G:Profiler

(<https://biit.cs.ut.ee/gprofiler/gost>)

The screenshot shows the Enrichr web interface. At the top right, it displays statistics: 21,683,692 lists analyzed, 318,171 terms, and 156 libraries. The main navigation bar includes links for Analyze, What's new?, Libraries, Gene search, About, and Help. Below the navigation, a section titled "Input data" contains two options: "Choose an input file to upload. Either in BED format or a list of genes." and "Paste a list of valid Entrez gene symbols on each row in the text-box below. Try a gene set example." A "Choose File" button shows "No file chosen". To the right, there is a large text input box. A sidebar on the right lists various analysis tools: g:GOST Gene Group Functional Profiling, g:Cocoa Compact Compare of Annotations, g:Convert Gene ID Converter, g:Orth Orthology search, and g:SNPense Convert rsID. Below this, a "Welcome!" message and links for Contact, FAQ, R / APIs, Beta, and Archive are shown. The footer contains copyright information and a list of supported databases and ontologies.

Enrichr

Login | Register
21,683,692 lists analyzed
318,171 terms
156 libraries

Analyze What's new? Libraries Gene search About Help

Input data

Choose an input file to upload. Either in BED format or a list of genes.
Try an example [BED file](#).

Choose File No file chosen

Paste a list of valid Entrez gene symbols on each row in the text-box below. [Try a gene set example](#).

g:Profiler

g:GOST Gene Group Functional Profiling
g:Cocoa Compact Compare of Annotations
g:Convert Gene ID Converter
g:Orth Orthology search
g:SNPense Convert rsID

Welcome! Contact FAQ R / APIs Beta Archive

J. Reimand, T. Arak, P. Adler, L. Kolberg, S. Reisberg, H. Peterson, J. Vilo: g:Profiler -- a web server for functional interpretation of gene lists (2016 update) Nucleic Acids Research 2016; doi: 10.1093/nar/gkw199 ([PDF](#), [more](#))

Organism: Homo sapiens

Query (genes, proteins, probes):

Options:

- [?] Significant only (checked)
- [?] Ordered query
- [?] No electronic GO annotations
- [?] Chromosomal regions
- [?] Hierarchical sorting (checked)
- [?] Hierarchical filtering
- Show all terms (no filtering)
- [?] Output type: Graphical (PNG)
- Show advanced options

[?] or Term ID:

g:Profiler! Clear

Example or random query

g:Profiler version r1760_e93_eg40. Version info

Gene Ontology (checked)
Biological process
Cellular component
Molecular function
Inferred from experiment [IDA, IPI, IMP, IGI, IEP]
Direct assay [IDA] / Mutant phenotype [IMP]
Genetic interaction [IGI] / Physical interaction [IPI]
Inferred from High Throughput Experiment [HDA, HMP, HGI, HEP]
High Throughput Direct Assay [HDA] / High Throughput Mutant Phenotype [HEP]
High Throughput Genetic Interaction [HGI] / High Throughput Expression pattern [HEP]
Traceable author [TAS] / Non-traceable author [NAS] / Inferred by curator [IC]
Expression pattern [IEP] / Sequence or structural similarity [ISS] / Genomic context [IGC]
Sequence Model [ISM] / Sequence Alignment [ISA] / Sequence Orthology [ISO]
Biological aspect of ancestor [IBA] / Rapid divergence [IRD]
Reviewed computational analysis [RCA] / Electronic annotation [IEA]
No biological data [ND] / Not annotated or not in background [NA]
Biological pathways
KEGG
Reactome
Regulatory motifs in DNA
TRANSFAC TFBS
miRTarBase
Protein databases
Human Protein Atlas
CORUM protein complexes
Human Phenotype Ontology (sequence homologs in other species)

GENETICS RESOURCES

GWAS Summary Statistics: PGC Downloads

(<https://www.med.unc.edu/pgc/data-index/>)

The screenshot shows the homepage of the Psychiatric Genomics Consortium (PGC) Data Index. The header includes the UNC School of Medicine logo, navigation links for UNC Chapel Hill, UNC Health Care, and International, and a search bar with options to search the site or the School of Medicine. The main title "Psychiatric Genomics Consortium" is displayed prominently. Below the title is a navigation menu with links to Home, PGC Groups, People, Data Access, Worldwide Lab, and FAQ. A breadcrumb trail indicates the current location is Home / Data Index. The main content area is titled "Data Index" and features a section for "PGC Data" with a note that all datasets are available for download. It lists several specific datasets under "Disorders": ADHD June 2017, ADHD – June 2017 (Full GWAS), ADHD – European GWAS, and ADHD – sex-specific.

UNC SCHOOL of MEDICINE

UNC Chapel Hill | UNC Health Care | Int

Search...

Search this site Search UNC School of Medicine

Psychiatric Genomics Consortium

Home / Data Index

Data Index

PGC Data

all the below datasets are available to download by visiting the PGC Downloads page

Disorders

ADHD June 2017
[ADHD – June 2017 \(Full GWAS\)](#)
[ADHD – European GWAS](#)
ADHD – sex-specific

GWAS + PRS/GPS RESOURCES

PROTOCOL

Basic statistical analysis in genetic case-control studies

Geraldine M Clarke¹, Carl A Anderson², Fredrik H Pettersson¹, Lon R Cardon³, Andrew P Morris¹ & Krina T Zondervan¹

¹Genetic and Genomic Epidemiology Unit, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. ²Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK. ³GlaxoSmithKline, King of Prussia, Pennsylvania, USA. Correspondence should be addressed to G.M.C. (gclarke@well.ox.ac.uk).

Published online 3 February 2011; doi:10.1038/nprot.2010.182

This protocol describes how to perform basic statistical analysis in a population-based genetic association case-control study. The steps described involve the (i) appropriate selection of measures of association and relevance of disease models; (ii) appropriate selection of tests of association; (iii) visualization and interpretation of results; (iv) consideration of appropriate methods to control for multiple testing; and (v) replication strategies. Assuming no previous experience with software such as PLINK, R or Haploview, we describe how to use these popular tools for handling single-nucleotide polymorphism data in order to carry out tests of association and visualize and interpret results. This protocol assumes that data quality assessment and control has been performed, as described in a previous protocol, so that samples and markers deemed to have the potential to introduce bias to the study have been identified and removed. Study design, marker selection and quality control of case-control studies have also been discussed in earlier protocols. The protocol should take ~1 h to complete.

© 2011 The Authors
Journal compilation © 2011
Association for
Laboratory and
Clinical
Pathology

nature
protocols

REVIEW ARTICLE

<https://doi.org/10.1038/s41596-020-0353-1>



Tutorial: a guide to performing polygenic risk score analyses

Shing Wan Choi^{1,2}, Timothy Shin-Heng Mak^{1,3} and Paul F. O'Reilly^{1,2}✉

A polygenic score (PGS) or polygenic risk score (PRS) is an estimate of an individual's genetic liability to a trait or disease, calculated according to their genotype profile and relevant genome-wide association study (GWAS) data. While present PRSs typically explain only a small fraction of trait variance, their correlation with the single largest contributor to phenotypic variation—genetic liability—has led to the routine application of PRSs across biomedical research. Among a range of applications, PRSs are exploited to assess shared etiology between phenotypes, to evaluate the clinical utility of genetic data for complex disease and as part of experimental studies in which, for example, experiments are performed that compare outcomes (e.g., gene expression and cellular response to treatment) between individuals with low and high PRS values. As GWAS sample sizes increase and PRSs become more powerful, PRSs are set to play a key role in research and stratified medicine. However, despite the importance and growing application of PRSs, there are limited guidelines for performing PRS analyses, which can lead to inconsistency between studies and misinterpretation of results. Here, we provide detailed guidelines for performing and interpreting PRS analyses. We outline standard quality control steps, discuss different methods for the calculation of PRSs, provide an introductory online tutorial, highlight common misconceptions relating to PRS results, offer recommendations for best practice and discuss future challenges.

HANDS-ON ACTIVITY: PLSR for 22q

<https://neurohackademy.2i2c.cloud/hub/spawn>

→ R machine → /curriculum/forsyth-genetics/PLSR_22q_CT_AHBA_Example.R

The screenshot shows the JupyterHub spawn interface. At the top, there is a navigation bar with the jupyterhub logo, 'Home', and 'Token' links. Below this is a section titled 'Server Options' containing four radio button choices:

- Regular CPU instance
- GPU machine

Start a container on a dedicated node with a GPU
- R machine

Start a container with R available
- Bring your own image

Specify your own docker image (must have python and jupyterhub installed in it)

Image: Other...

HANDS-ON ACTIVITY: PARCELLATING AHBA WITH ABAGEN

<https://neurohackademy.2i2c.cloud/hub/spawn> →
Regular CPU Instance → Python3 Notebook →
/curriculum/forsyth-genetics/Example_Abagen_Use.ipynb

Server Options

- Regular CPU instance

- GPU machine

Start a container on a dedicated node with a GPU

- R machine

Start a container with R available

- Bring your own image

Specify your own docker image (must have python and jupyterhub installed in it)

Image

Other...



Custom image



Thank you!

✉ jenforsy@uw.edu  jenforsythphd