Investigating Schizophrenia GWAS Risk Variants in Cognitive and Neuroimaging Defects: A GENUS Consortium Update

Tracey L. Petryshen[1,2,*] and Gabriëlla A.M. Blokland[1,2], on behalf of the GENUS Consortium[3]

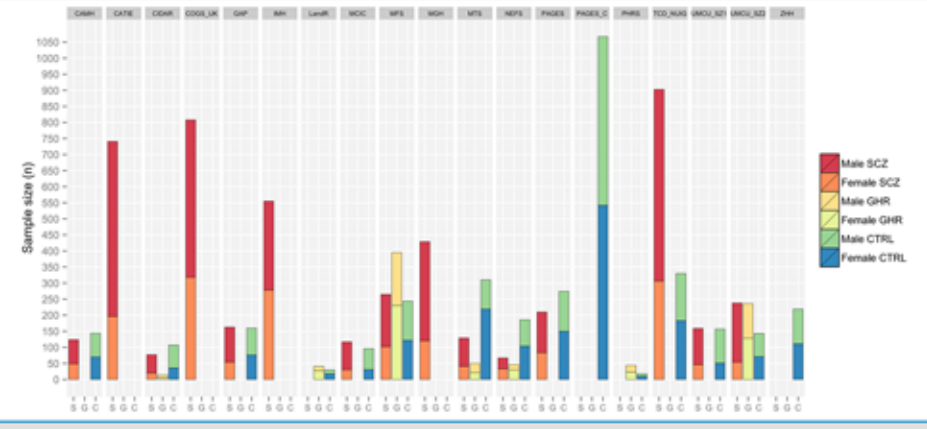
[1] Psychiatric and Neurodevelopmental Genetics Unit, Department of Psychiatry and Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States;
[2] Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, United States; [3] genus.mgh.harvard.edu/investigators. * tpetryshen@mgh.harvard.edu

Background

- The Genetics of Endophenotypes of Neurofunction to Understand Schizophrenia (GENUS) consortium is clarifying the neurobiological role of schizophrenia (SCZ) risk genes by examining associations with cognitive and neuroanatomical measures in a unique sample collection.
- Fifteen research groups contributed cognitive and genetic data from a total of 4,575 SCZ cases, 4640 controls, and 875 genetic risk individuals, of which a subset (~3,000) also has structural MRI data.
- A crucial first step has been the development of phenotype and genotype processing and quality assessment protocols to harmonize data across the samples, since the sites used different neuropsychological test batteries, SNP arrays, and MRI acquisition parameters.
- A parallel step has been extensive literature review and meta-analysis to select cognitive and brain structure traits for genetic analyses that have high heritability, reliability, and robust case-control differences.
- Initial genetic studies will focus on association analyses of each selected phenotype with SNPs having prior SCZ GWAS evidence ($p<5x10^{-8}$) and polygenic SNP sets that are functionally related, as well as multivariate regression analysis to identify phenotype-genotype profiles or clusters.

Samples - 4575 SCZ, 875 genetic high risk, 4640 control

Sample	Site	PI(s)
CAMH	Toronto, Canada	Voineskos
CATIE	Multi-site, USA	NIMH
CIDAR-VA	Boston, USA	McCarley, Kikinis
COGS-UK	Cardiff, UK	Walters
GAP, MTS, MFS	London, UK	Toulopoulou
IMH	Singapore	Keong, Sim, Yang
L&R	Boston, USA	DeLisi
MCIC	Multi-site, USA	Turner, Ehrlich, Gollub
MGH	Boston, USA	Roffman, Holt
NEFS	Boston, USA	Goldstein, Seidman, Buka
PAGES	Munich, Germany	Rujescu, Giegling
PHRS	Pittsburgh, USA	Keshavan, Montrose
TCD-NUIG	Dublin/Galway, Ireland	Donohoe
UMCU-SZ1,2	Utrecht, Netherlands	Kahn, van Haren
ZHH	New York, USA	Malhotra, Trampush



Cognitive Phenotypes

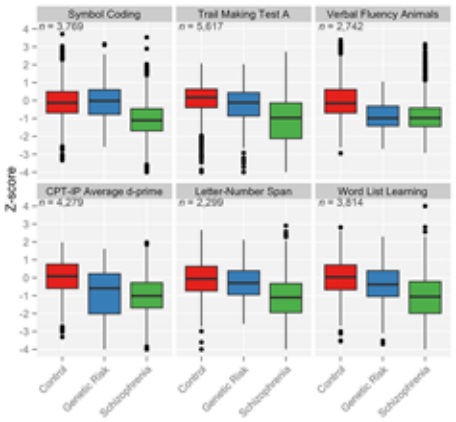
MATRICS domain	Neuropsychological tests	Effect size[2] standardized mean difference	Heritability[3]
Speed of Processing	BACS Symbol Coding*; Trail Making A*; Category Fluency*; Letter Fluency	-0.96	0.52 (0.45 - 0.59)
Attention/ Vigilance	Continuous Performance Test – Identical Pairs*; Other CPTs	-0.71	0.30 (0.23 - 0.37)
Working Memory	UMD Letter-Number Span/ WAIS Letter-Number Sequencing*; WMS-III Spatial Span/ WMS-R Visual Memory Span*; WAIS Digit Span; Arithmetic	-0.79	0.42 (0.37 - 0.47)
Verbal Learning	Hopkins Verbal Learning Test-R/ Other word list learning tests*; WMS Logical Memory/Stories; WMS Verbal Paired Associates	-1.20	0.34 (0.24 - 0.44)
Visual Learning	Brief Visuospatial Memory Test-R*; Benton Visual Retention Test; WMS Visual Paired Associates; Visual Reproduction; Figural Memory; Faces	-0.91	0.35 (0.30 - 0.41)
Reasoning/ Problem Solving	NAB Mazes; Wisconsin Card Sorting Test; Trail Making B	-0.83	0.26 (0.19 - 0.34)

BACS = Brief Assessment of Cognition for Schizophrenia; UMD = University of Maryland
HVLT = Hopkins Verbal Learning Test
CVLT = California Verbal Learning Test
CPT-IP = Continuous Performance Test - Independent Pairs

Schizophrenia-High Risk-Control Group Differences

All variables corrected for sex, age, and their interactions

All SCZ vs CTRL comparisons significant at p<0.0001 (Tukey's posthoc)



Symbol Coding = BACS Symbol Coding, WAIS Digit Symbol Coding, RBANS Symbol Coding

Letter-Number Span = UMD Letter-Number Span and WAIS Letter-Number Sequencing

Word List Learning = Total words recalled across trials from HVLT, CVLT, CVLT-SF, COGTEST Word List Memory, RBANS Word List Memory, Verbal Learning Test, Verbal Learning & Memory Test

Data Processing Pipelines

NEUROPSYCHOLOGICAL DATA

Literature review and Meta-analyses

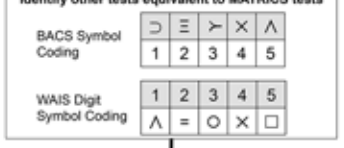
Identify cognitive domains having:
- Robust case-control differences (see Table to left)
- Significant heritability (see Table to left)

↓

MATRICS consensus cognitive battery
- BACS Symbol Coding
- Category Fluency
- Trail Making Test A
- Continuous Performance Test – Independent Pairs
- WMS-III Spatial Span
- UMD Letter-Number Span
- Hopkins Verbal Learning Test-R
- Brief Visuospatial Memory Test-R
- NAB Mazes

↓

Identify other tests equivalent to MATRICS tests



↓

Harmonize and standardize tests
- Pool controls for each test version
- Regress out age and sex effects
- Calculate standardized residuals for cases/controls
  Z = [(x - predicted x) / residuals SD controls]

↓

Phenotypes selected for genetic analyses
- MATRICS (equivalent) individual test scores
- Cognitive domain scores
- Spearman's "g"

↓

Quality Checking
- Verify case-control differences for each phenotype
- Identify significant covariate effects

GENETIC DATA

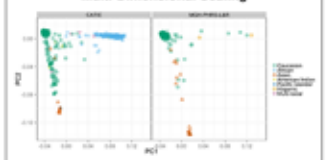
Genome-wide genotype data

↓

Quality Checking and Filtering (Plink)
- SNP missingness < 5% (before sample removal)
- Subject missingness < 2%
- Autosomal heterozygosity deviation (Fhet < 0.2)
- SNP missingness < 2% (after sample removal)
- Case/control difference SNP missingness < 2%
- SNP HWE: $p > 10^{-6}$ controls or $p > 10^{-10}$ cases

↓

Multi-Dimensional Scaling



↓

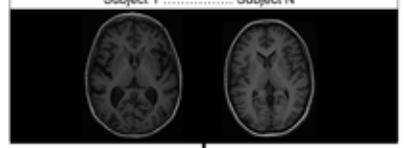
Imputation (IMPUTE2/SHAPEIT2)

↓

Quality Checking and Filtering (Plink)
- SNPs with high imputation quality (INFO > 0.8)
- SNPs with low missingness (<1%)
- Linkage disequilibrium pruning ($r^2 > 0.02$)
- Frequency filtering (MAF > 0.05)

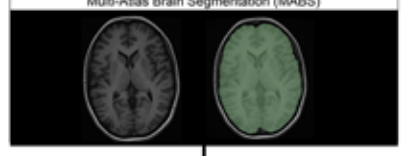
STRUCTURAL NEUROIMAGING DATA

T1-weighted structural images
Subject 1 ..... Subject N



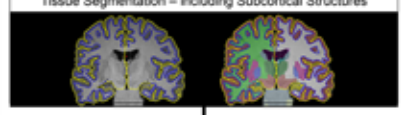
↓

MRI image processing
Multi-Atlas Brain Segmentation (MABS)



↓

FreeSurfer with MABS brain mask
Tissue Segmentation – including Subcortical Structures

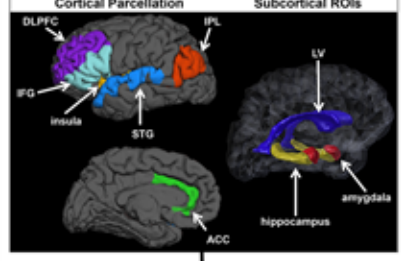


↓

Identify ROIs with:
- Robust case-control differences (see Table to right)
- Significant h^2 (see Table to right)

↓

Cortical Parcellation **Subcortical ROIs**



↓

Quality Checking

Genetic association to cognitive phenotypes

N=~10,000

SCZ ~4,575
CTRL ~4,640
GHR ~875

Genetic association to imaging phenotypes

N=~3,000

SCZ ~1,235
CTRL ~1,465
GHR ~370

Covariates
- Age
- Sex
- Age²
- Age-x-Sex
- Age²-x-Sex
- MDS
- Case-control status
- Medication/illness duration
- Education level (cog)
- ICV (MRI)
- Handedness (MRI)
- Scanner (MRI)

Meta-analysis across samples
- Fixed effects meta-analysis
- Random effects meta-analysis

NeuroImaging Phenotypes

Region of Interest	Effect size[4] Cohen's *d*	Heritability[5]
Superior temporal gyrus	-0.58	0.53 – 0.65
Transverse temporal gyrus	-0.29	0.44 – 0.58
Inferior frontal gyrus	-0.41	0.17 – 0.51
Middle frontal gyrus	-0.32	0.29 – 0.57
Anterior cingulate	-0.34	0.12 – 0.50
Inferior parietal lobule	--	0.41 – 0.65
Insula	-0.44	0.50 – 0.65
Hippocampus	-0.52	0.56 – 0.64
Amygdala	-0.31	0.65
Lateral Ventricles	+0.45	0.44 – 0.70
Total cortical GM volume	-0.43	0.67
Total brain volume	-0.30	0.83
Intracranial volume	-0.17	0.79

Discussion

- Due to the high cross-site variability in the neuropsychological test batteries, the GWAS arrays, and the MRI acquisition parameters, we have developed comprehensive data harmonization approaches for the phenotypic and genotypic data.
- We have demonstrated that our neuropsychological data harmonization approach is successful, with significant case-control differences for individual tests and Spearman's "g", and genetic risk individuals performing at a level between that of cases and controls, as expected.
- Genotype data processing and structural MRI processing are expected to be completed in summer 2015.
- Association analyses of known SCZ risk variants and polygene sets with cognitive and brain structure traits in one of the largest sample collections of its kind may contribute towards understanding the function of existing SCZ risk variants in neural processes underlying SCZ.

References

1. Schizophrenia Working Group of the Psychiatric Genomics Consortium, *Biological insights from 108 schizophrenia-associated genetic loci*, Nature. 2014;511(7510):421-7
2. Mesholam-Gately, R.I., et al. *Neurocognition in first-episode schizophrenia: a meta-analytic review.* Neuropsychology 2009;23(3):315
3. Blokland, G.A.M., Petryshen, T.L., et al., in preparation.
4. Haijma, S.V., et al., *Brain volumes in schizophrenia: A meta-analysis in over 18000 subjects.* Schizophrenia Bulletin, 2012; doi:10.1093/schbul/sbs118.
5. Blokland, G.A.M., et al., *Genetic and environmental influences on neuroimaging phenotypes: A meta-analytical perspective on twin imaging studies.* Twin Res Hum Genet, 2012; 15(3): 351-71.

Acknowledgements

This collaborative project "Genetic Determinants of Schizophrenia Intermediate Phenotypes" is funded by the National Institute of Mental Health (NIMH) of the National Institutes of Health (NIH) under Award Number R01MH092380 (PI: Tracey Petryshen). The content of this poster is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.