Biobank-scale Multi-organ Imaging Genetics and Beyond

University of North Carolina at Chapel Hill

Hongtu Zhu

Joint works with all members of the UNC BIG-S2 Lab,

Bingxin Zhao, Tengfei Li, Yun Li, Stephen Smith, and Jason Stein

https://www.med.unc.edu/big-s2

CONTENTS

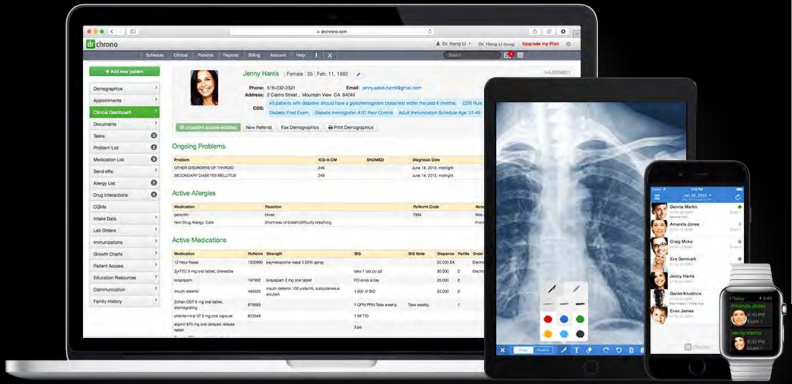
Part I

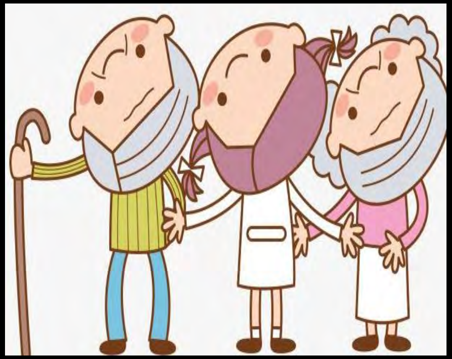
Methodological Challenges

EHR and PM

EHR is an information resource that takes residents' personal health as the core, runs through the entire life process, covers various health-related factors, realizes multi-channel information dynamic collection, and meets the needs of residents' self-care, health management and health decision-making.

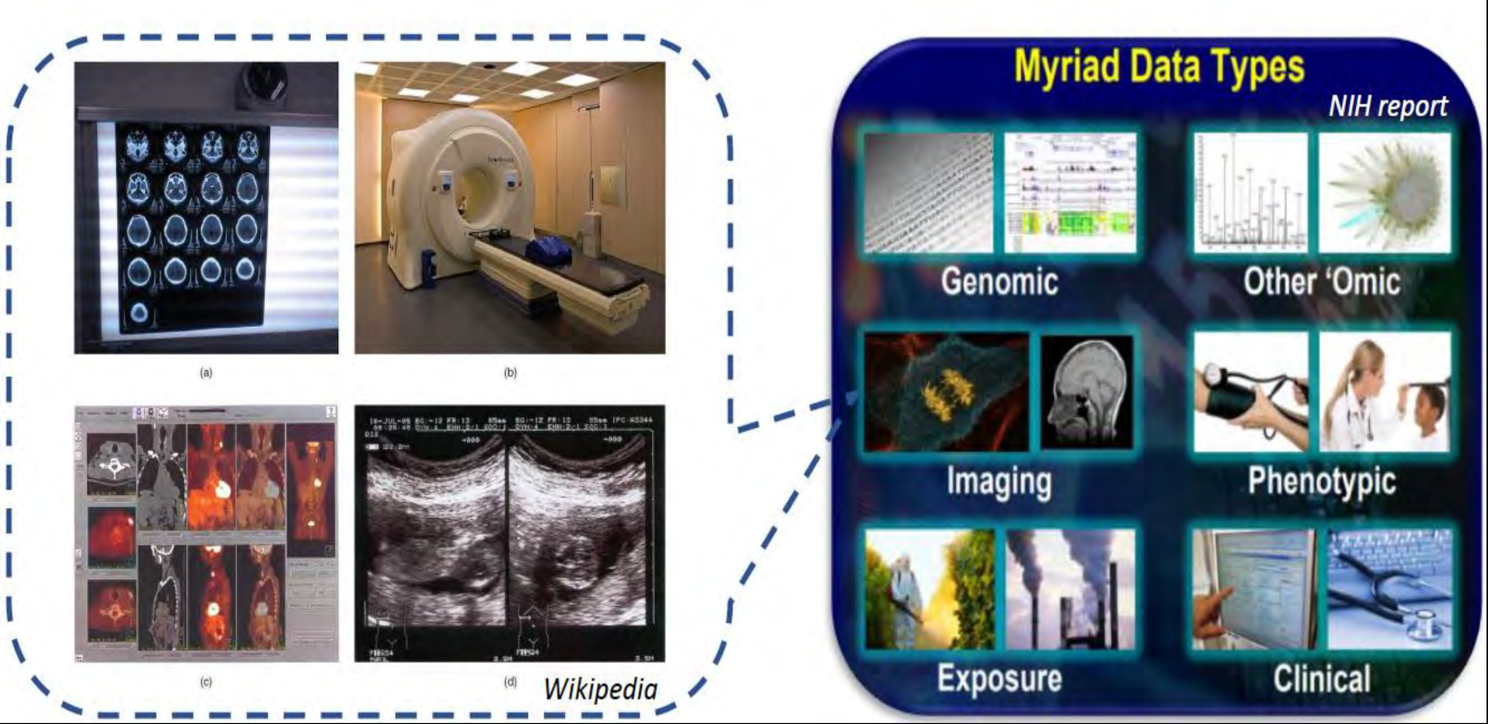
PM： Personalization, precision (time and plan), and health management. High-level medical technology is formed on the basis of in-depth understanding of people, diseases, and medicines. Analyze the health status of the entire population and improve the health of the general public.





Record the changes of all vital signs of an individual from birth to death, including personal living habits, past medical history, diagnosis and treatment, family medical history, current medical history, previous diagnosis and treatment history, previous physical examination results and other information, and accurately record digitally, so as to construct an integrated health service of prevention, diagnosis, treatment, rehabilitation, and health management.

Multi-modal Data

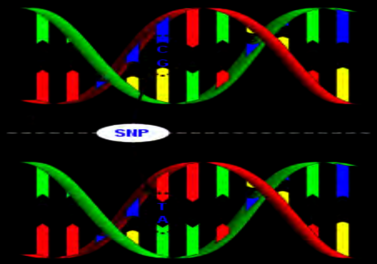


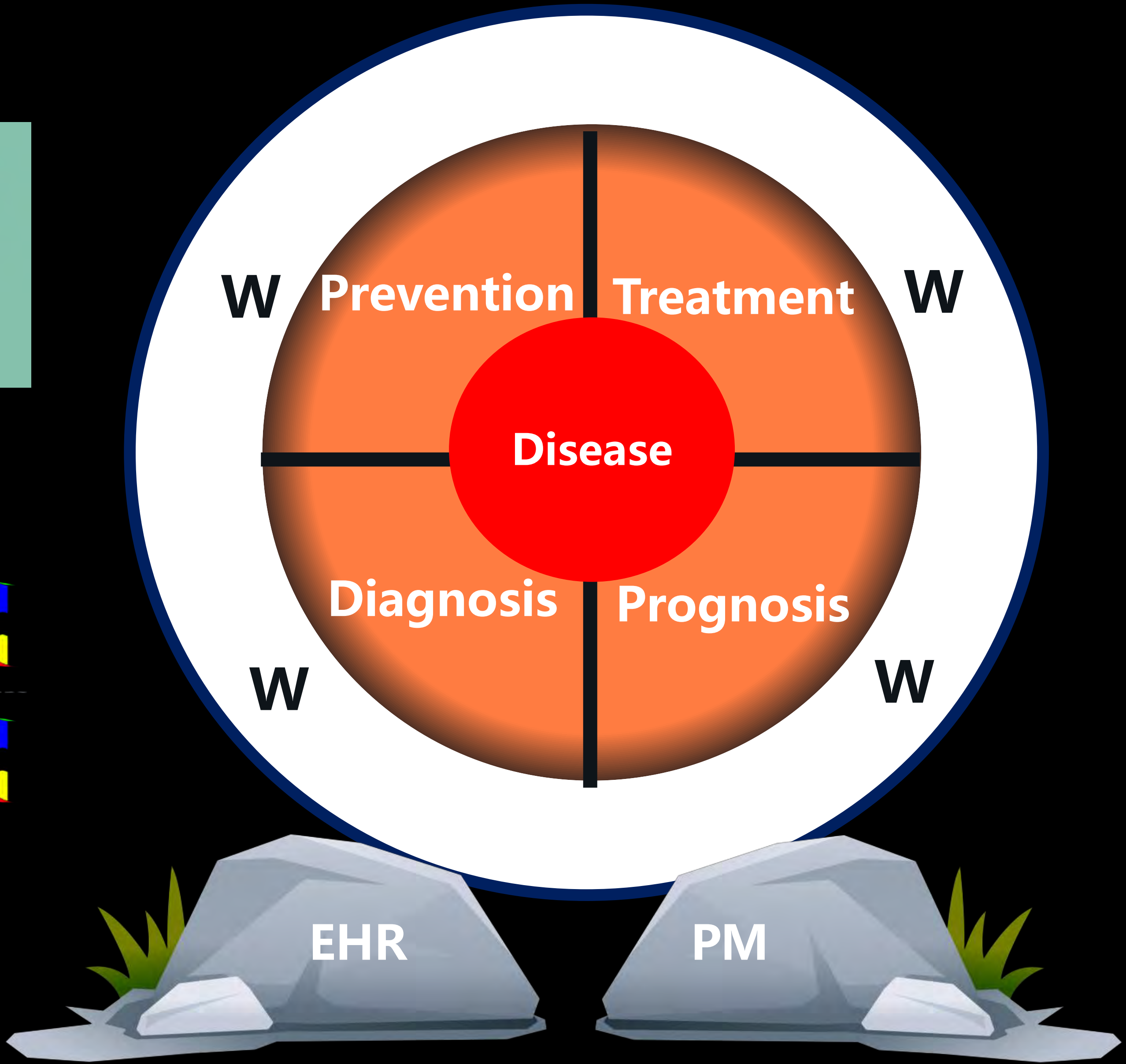
Clinical/Behavioral

Imaging

Genetics

Data Challenges



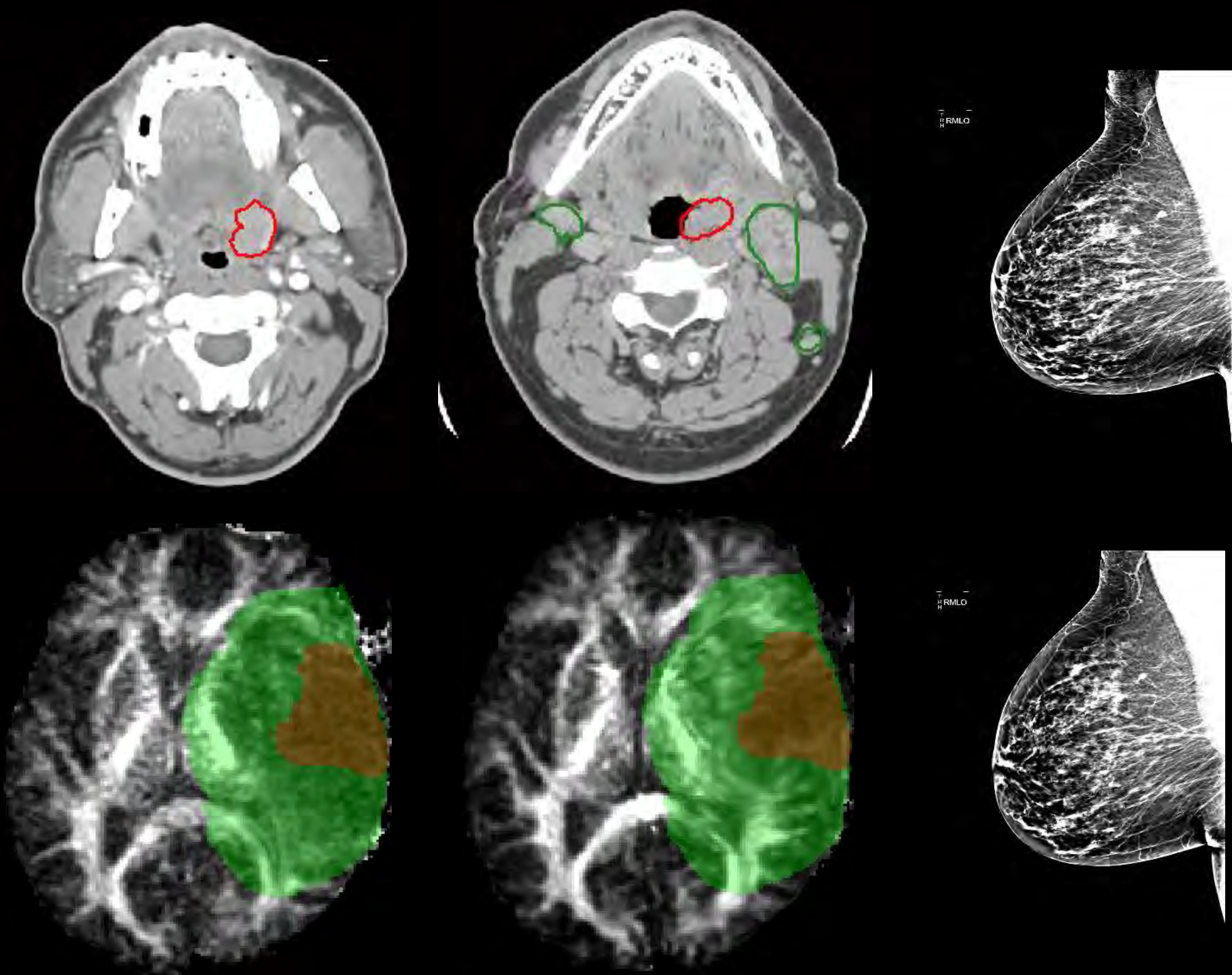


Data Challenges

IMAGENET

- Over 15M labeled high resolution images
- Roughly 80K categories
- Collected from web and labeled by Amazon Mechanical Turk

Lack of a large number of annotated data with high-quality



Method Challenges

D

Healthcare

Genetics

Imaging

Clinical

Shallow Information

Speech Recognition

NLP

Computer Vision

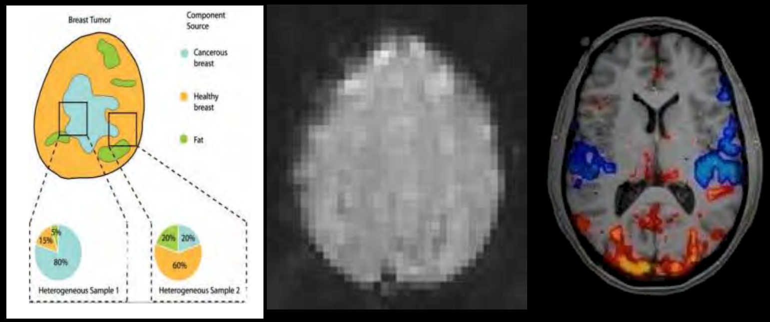
Prediction and Decision

IOT

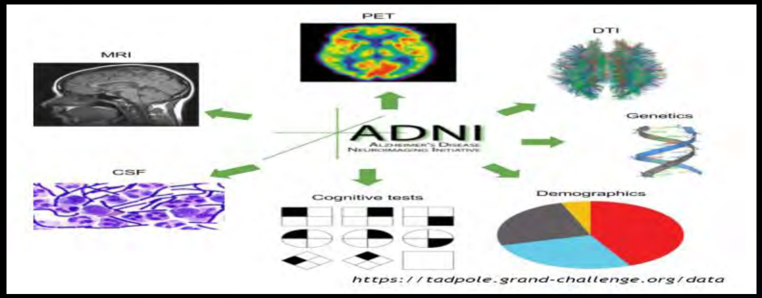
https://qz.com/989137/when-a-robot-ai-doctor-misdiagno

Source: 滴滴AI Labs, 滴滴战略部

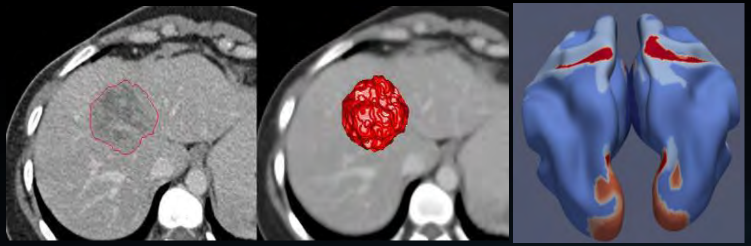
Ecological Layout



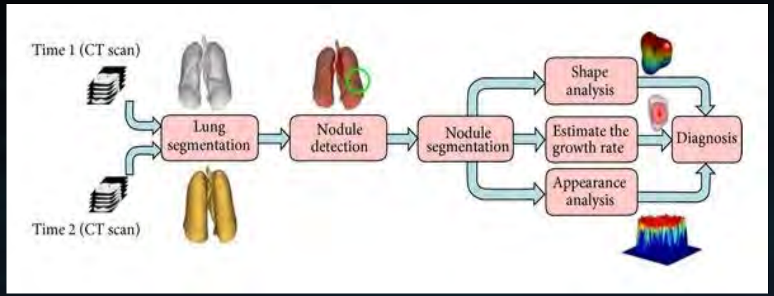
Deconvolution



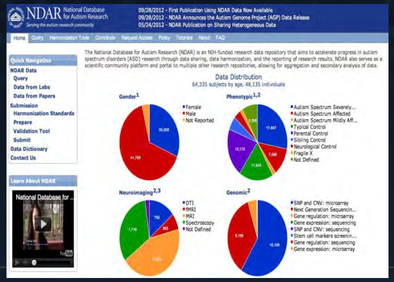
Integration



Learning



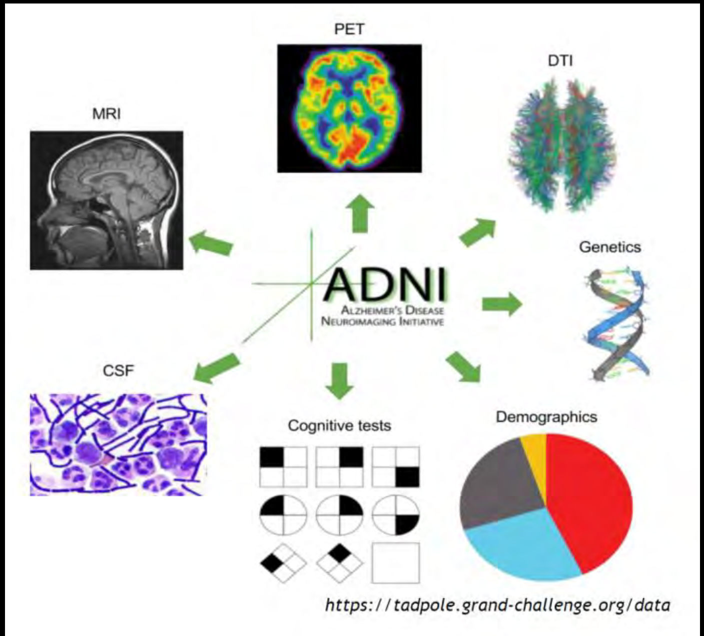
Prediction

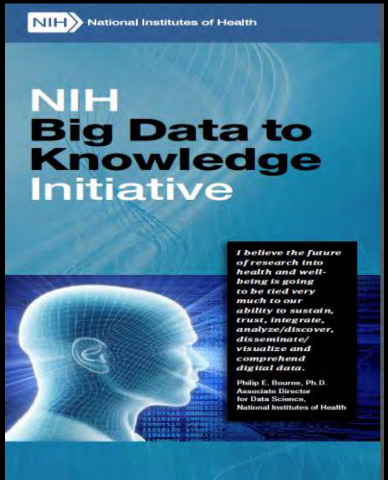


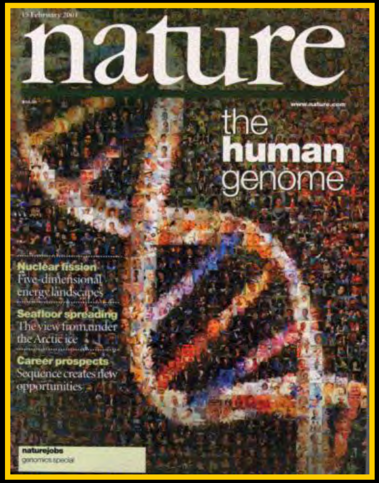
Ecological Layout

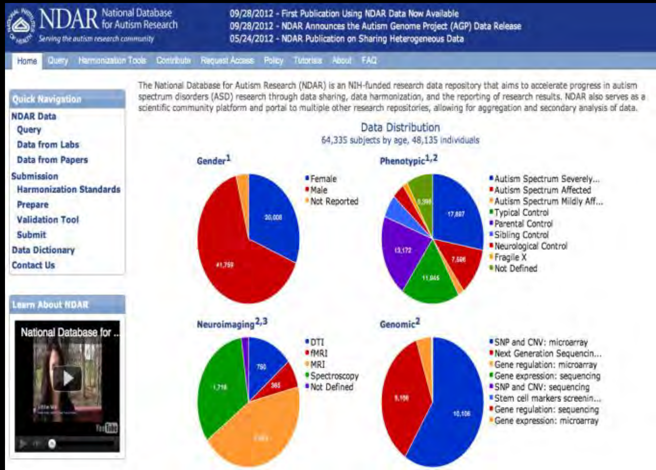
Large-scale Medical Studies

PING - 900 Pediatric Imaging, Neurocognition, and Genetics
BCP - 300 Baby Connectome Project
ADNI - 2000 Alzheimer’s Disease Neuroimaging Initiative
PNC - 1400 Philadelphia Neurodevelopmental Cohort
HCP - 1200 Human Connectome Project
ABCD - 10000 Adolescent Brain Cognitive Development
UKB - 500,000 UK Biobank Project
TCIA – 37,600 The Cancer Imaging Archive
NLST - 19,000 National Lung Screening Trial
OAI – 4800 Osteoarthritis Initiative
AllOfUs-1000,000+ All of us project









“Big Data” Brain Imaging Genetics Cohorts

“Big data” Brain imaging genetics datasets become available in recent few years

Systematically collect publicly available individual-level data for **> 50k** individuals

Build the largest database in this field

Aging Brain

BCP (*Age [0,5]*)

PING (*Age [3,21]*)

ABCD ***(n ~ 10k, Age [9,11])***

PNC (*Age [14,29]*)

HCP (*Age [22,35]*)

UK Biobank ***(n ~ 100k [Ongoing], Age [40,69])***

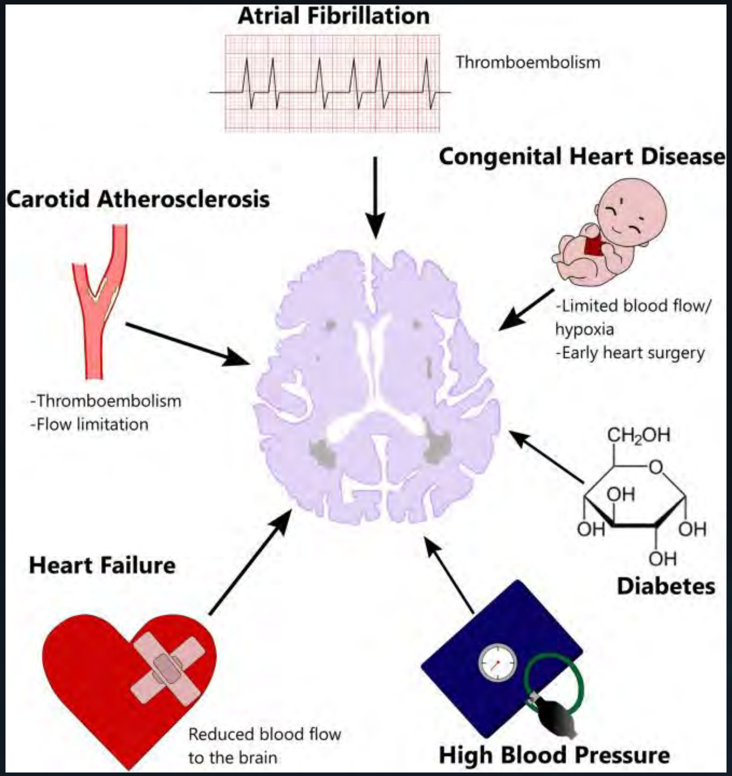
RADC (*Age > 65*)

ADNI (*Age [55,92]*)

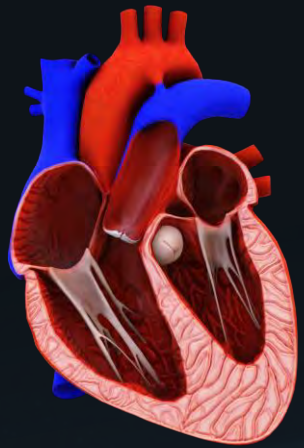
Brain Development

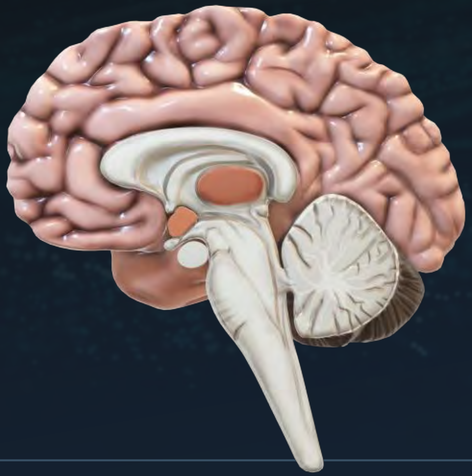
Cardiovascular Disease & Brain Health

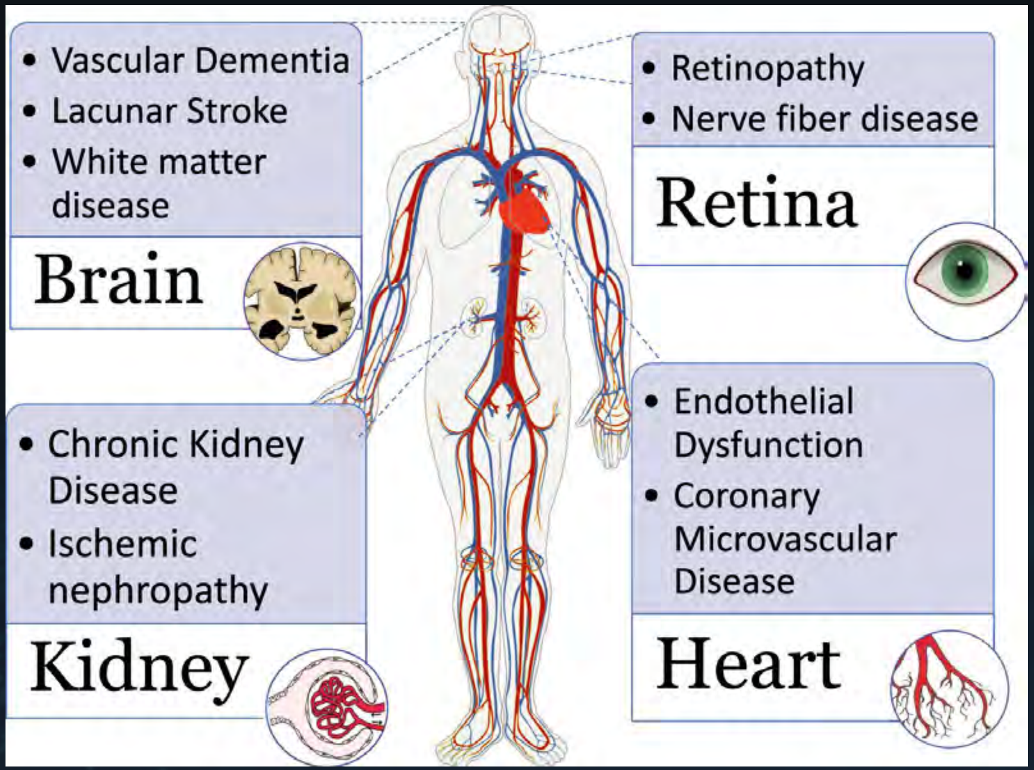
(Neuro)imaging: help understand the complex interplay between brain and other human organs and their underlying genetic overlaps



Possible causal factors of brain structure changes, resulting in brain disorders like stroke, dementia and cognitive impairment

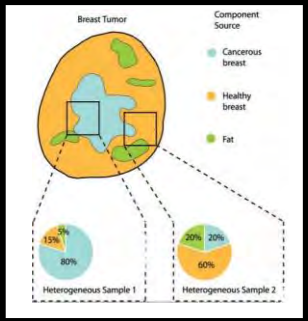


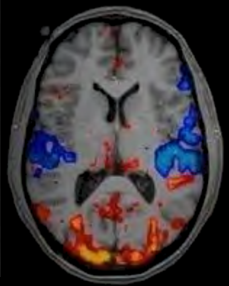




Many diseases (e.g., microvascular disease, high blood pressure) are multisystem disorders

Ecological Layout

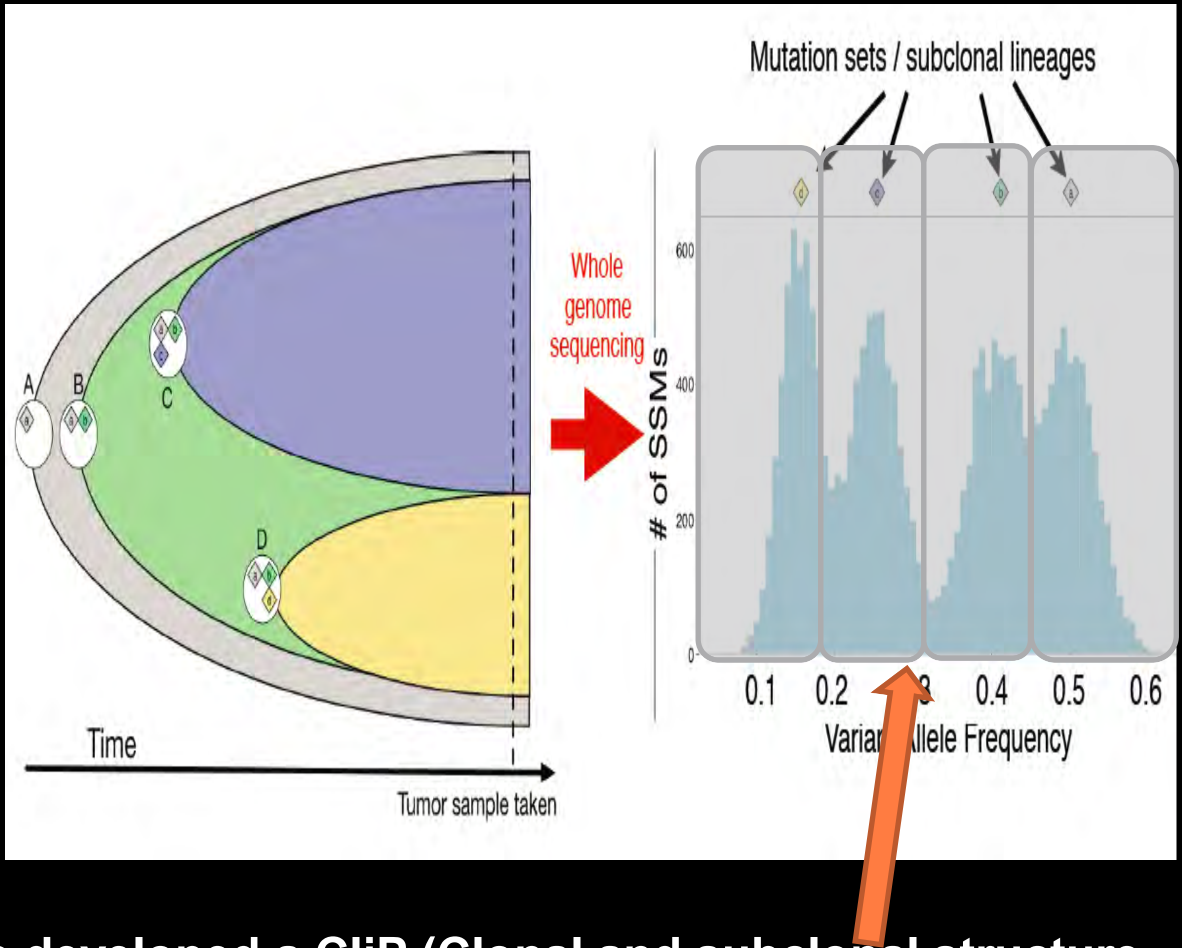


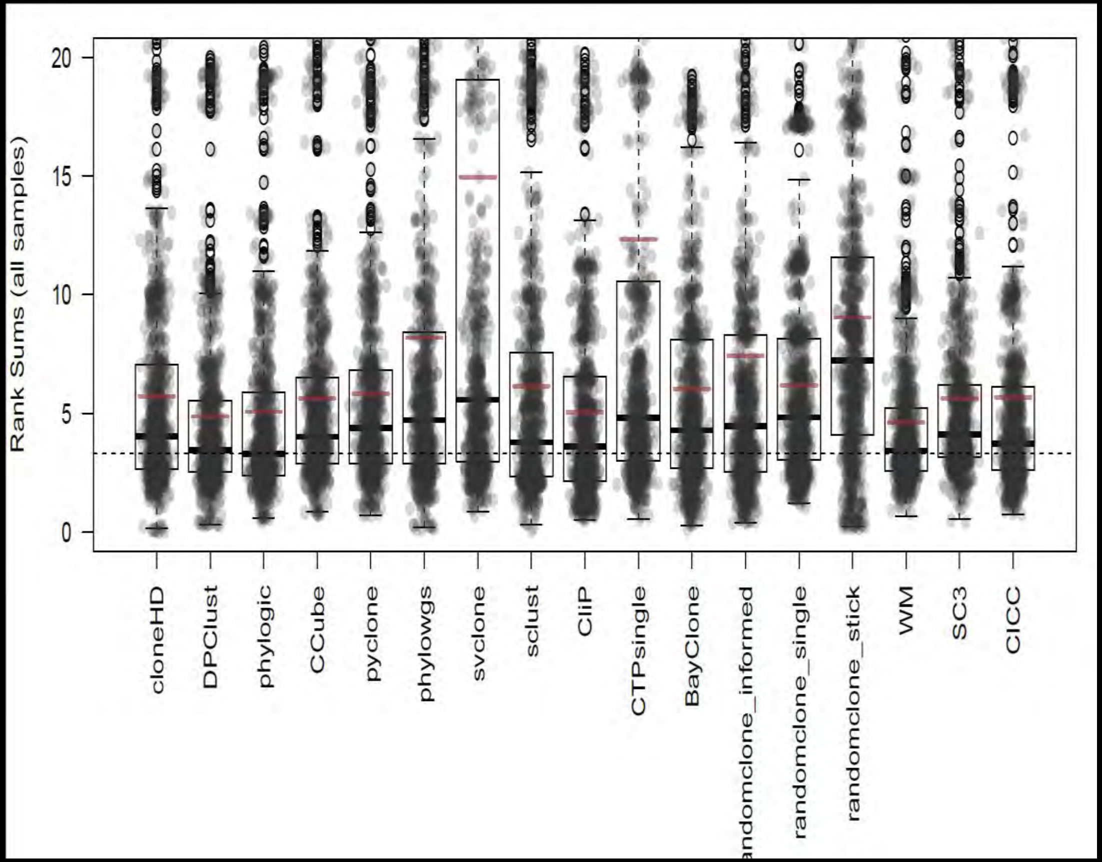


Deconvolution

1. **Gerstung, M. …, H. Zhu,…, P. V. Loo. and PCAWG network. The evolutionary history of 2,658 cancers. bioRxiv. *Nature*,** 578(7793):122-128, 2020.
2. **The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. *Nature*, 578, 82-93, 2020.**
3. **Y.Jiang, K.Yu, …., H. Zhu, W. Wang. CliP: subclonal architecture reconstruction of cancer cells in DNA sequencing data using a penalized likelihood model:** 36059962 (biorxiv.org)

Tumor Heterogeneity: identification





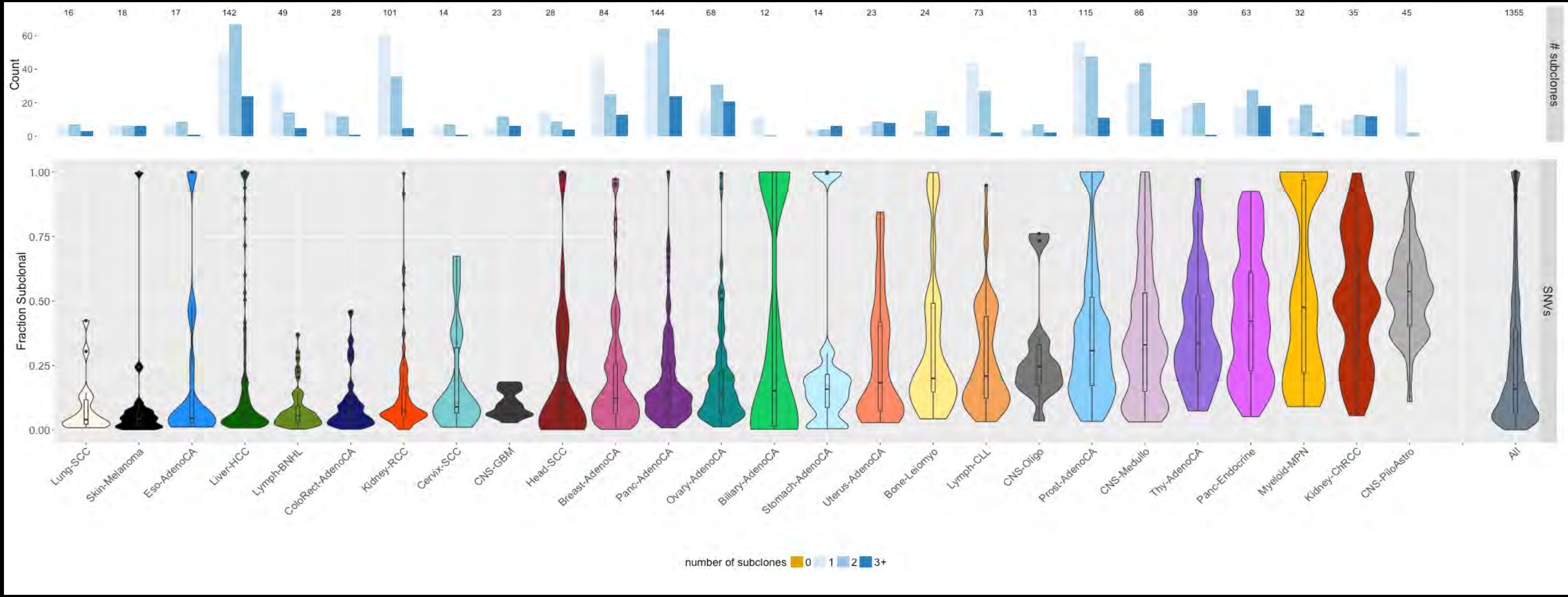
We developed a CliP (Clonal and subclonal structure identification through Pairwise difference penalization), to distinguish the sub-clones.

We tested CliP on 965 simulated samples generated by the Broad Institution, all samples are generated using copy number profiles from actual patients samples.

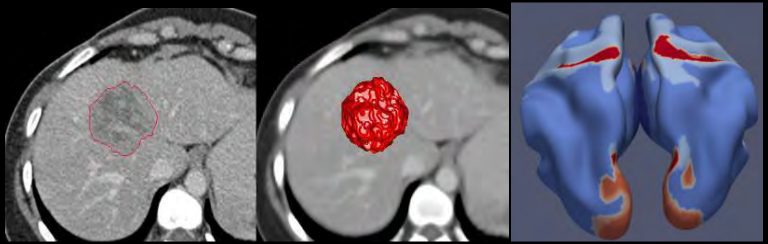
Results on ICGC samples

The International Cancer Genome Consortium has collected whole genome sequencing for over 2,700 samples. The clonality study shows that the clone/subclonality compositions are quite different across cancer types.

Figure: clonality composition of selected types of cancer. Both the number of subclones and subclonal fractions are different across tumor types



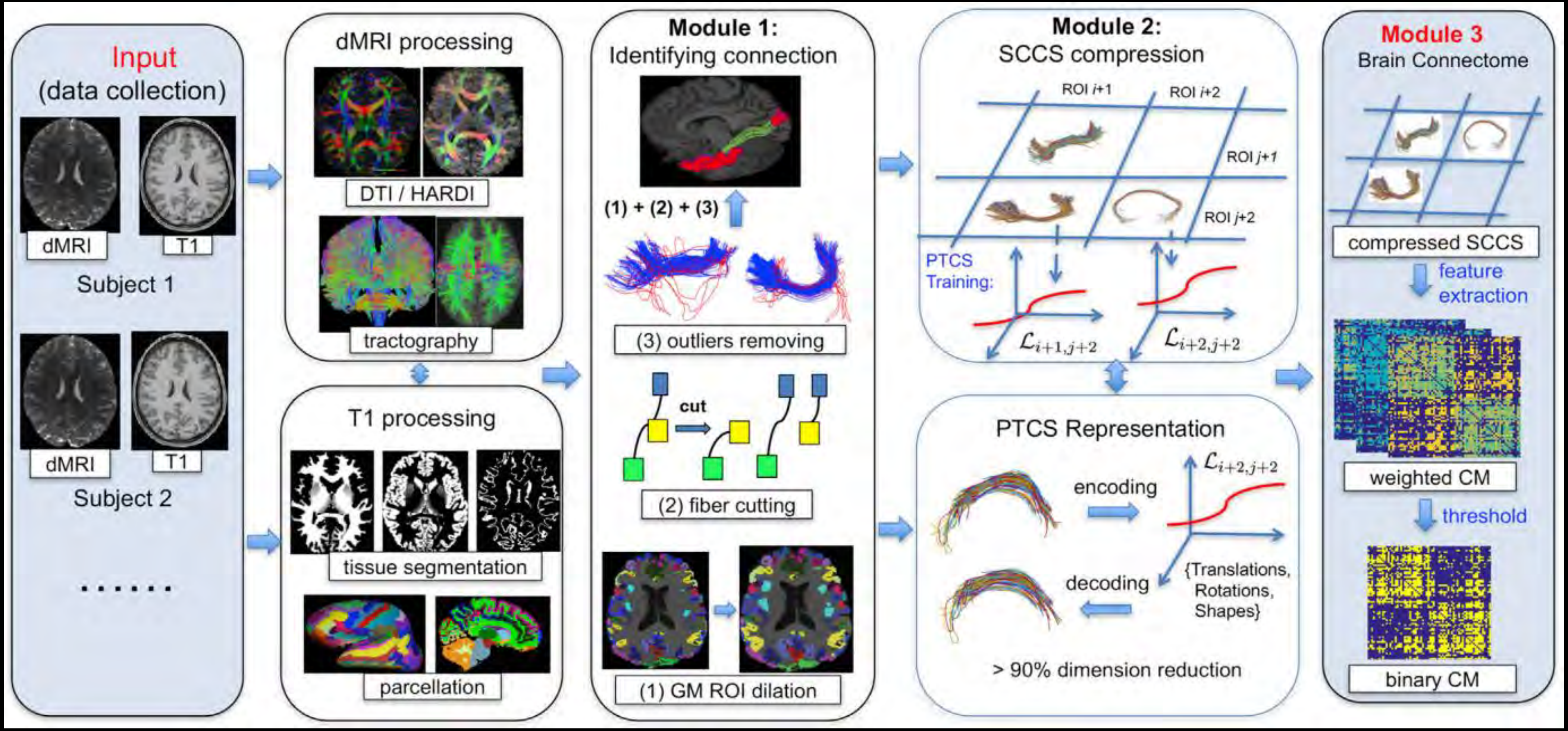
Ecological Layout



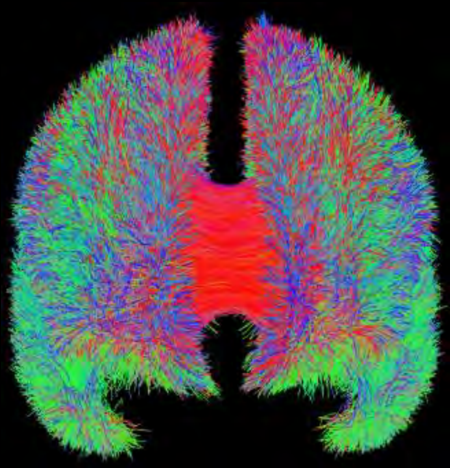
Learning

- W.Yin， T. Li， SC. Hung， H. Zhang， L. Wang， D. Shen， H. Zhu， P. J. Mucha， J. R. Cohen， Weili Lin. The Emergence of a Functionally Flexible Brain During Early Infancy. *PNAS*, 117 (38) 23904-23913.
- Smith, I. T., Townsend, L.B., Huh, R. H. Zhu, and Smith, S. L. (2017) Stream-dependent development of higher visual cortical areas. *Nature Neuroscience*, 20, 200–208.
- Z. Zhang, M. Descoteaux, J. Zhang, G. Girard, M. Chamberland, D. Dunson, A. Srivastava, and *H. Zhu*. (2018). Mapping Population based Structural Connectomes. *NeuroImage*, 172, 130-145.

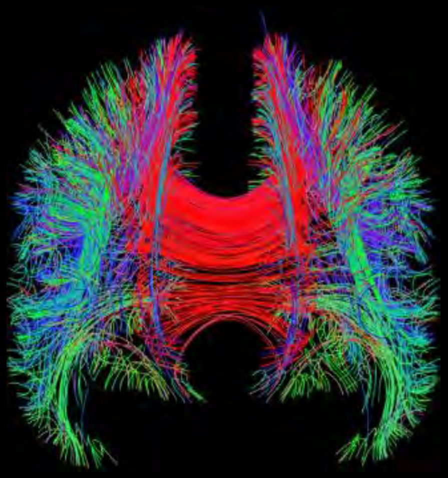
Population based Structural Connectomes



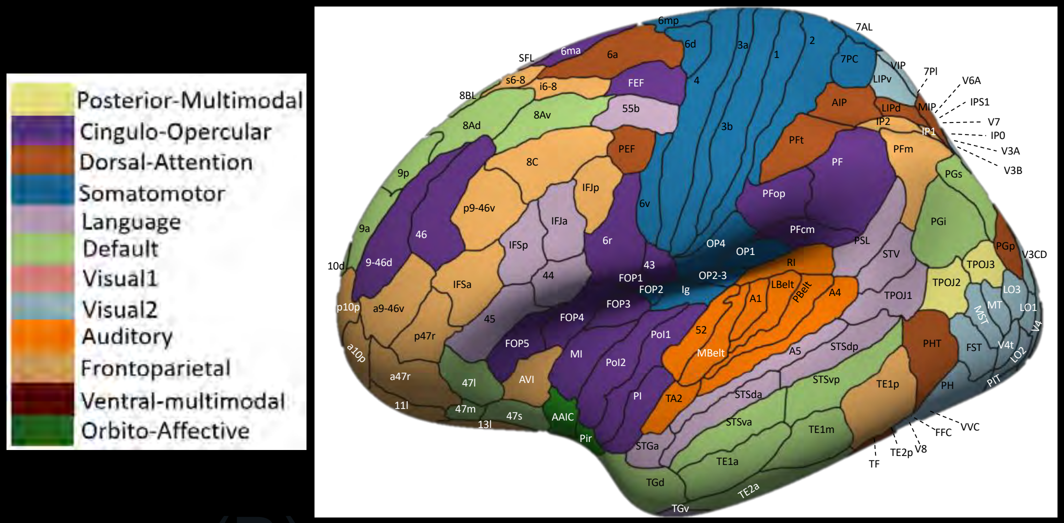
Brain Function-based Structural Connectome Atlas



Fiber atlas

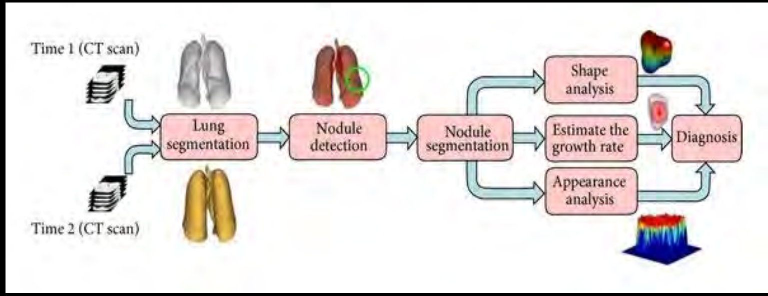


Fiber skeleton atlas



(B)

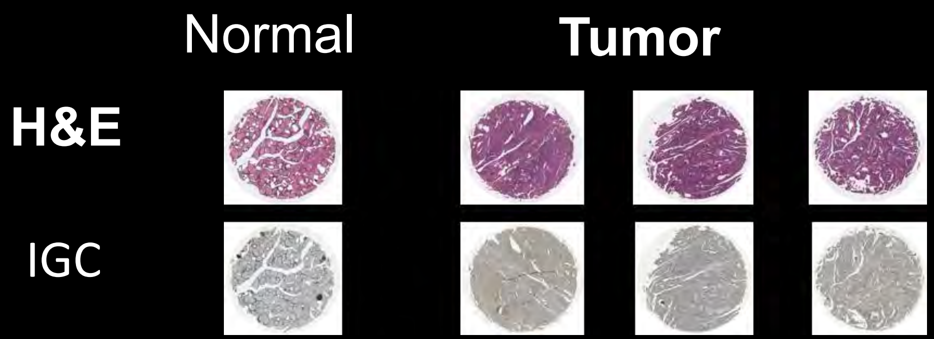
Ecological Layout



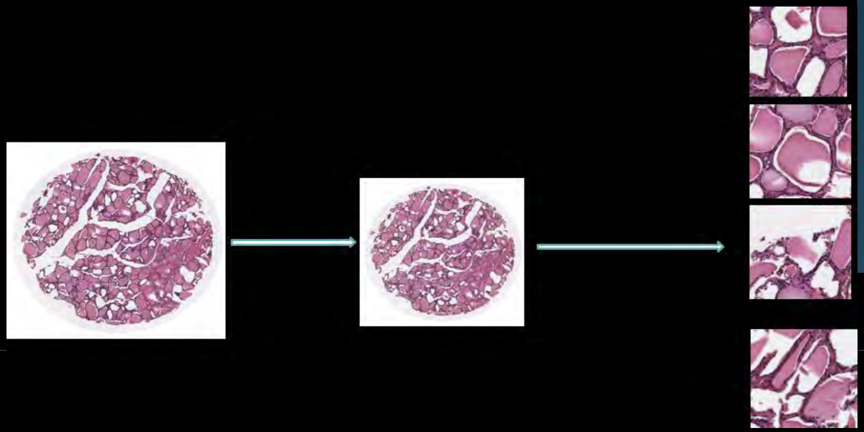
Prediction

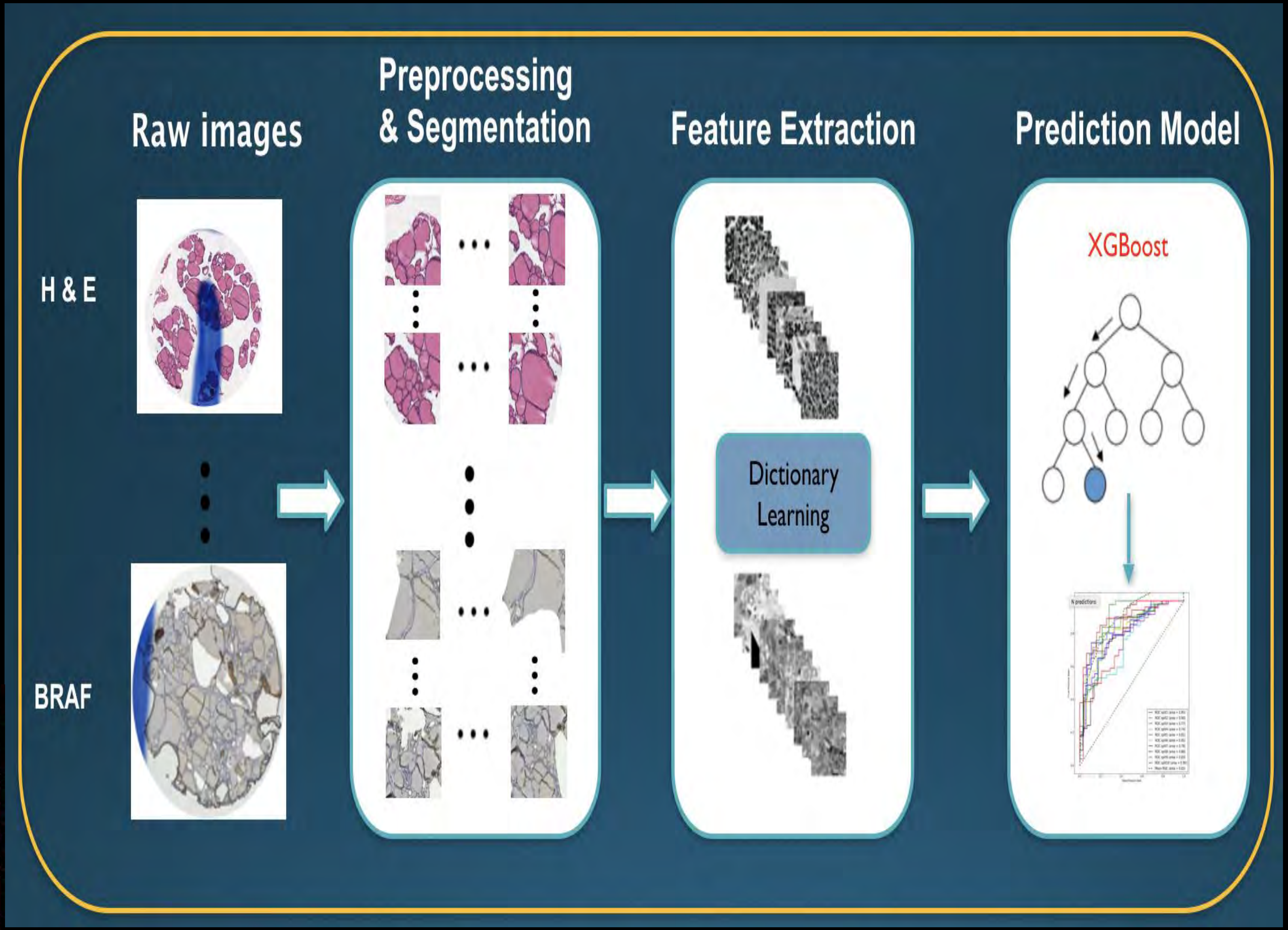
C. W. Wang, Y.C. Lee, E.Calista, **F. Zhou, <u>H. Zhu</u>**, R.Suzuki, D. Komura, S.Ishikawa, S.P. Cheng (2018).
A benchmark for comparing precision medicine methods in thyroid cancer diagnosis using tissue microarrays.
Bioinformatics, 34, 1767-1773.

CAMELYON17



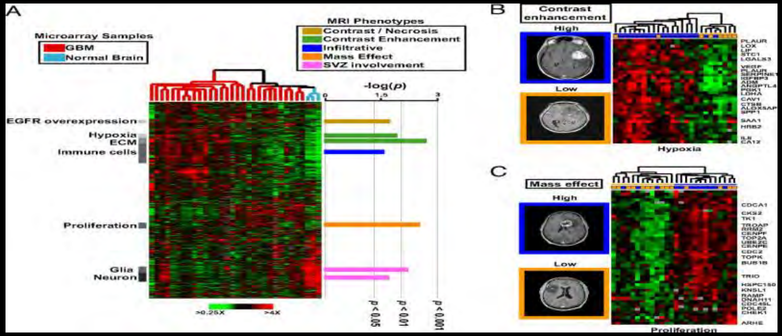
Normal
Tumor
H&E
IGC





Raw images
Preprocessing & Segmentation
Feature Extraction
Prediction Model
H & E
BRAF
Dictionary Learning
XGBoost

Ecological Layout

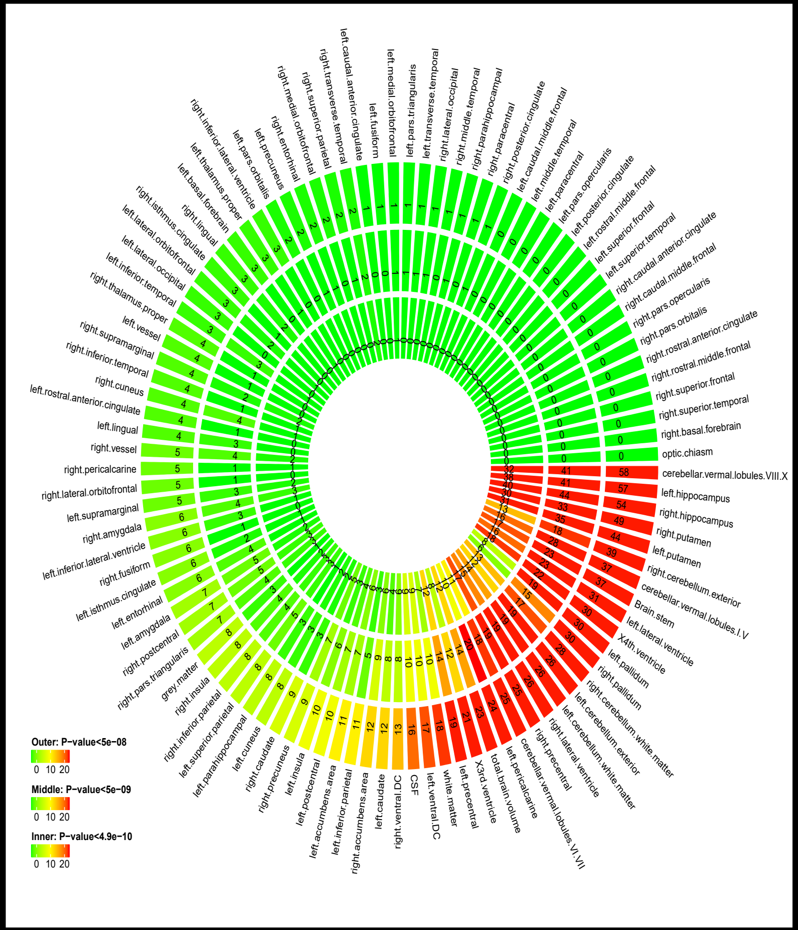


http://www.pnas.org/content/105/13/5213/F1.expansion.html

Integration

Image Genetics

Genome-wide association study (GWAS) of hundreds of imaging phenotypes with more than 50,000 subjects from five publicly available datasets (largest brain imaging GWAS so far)

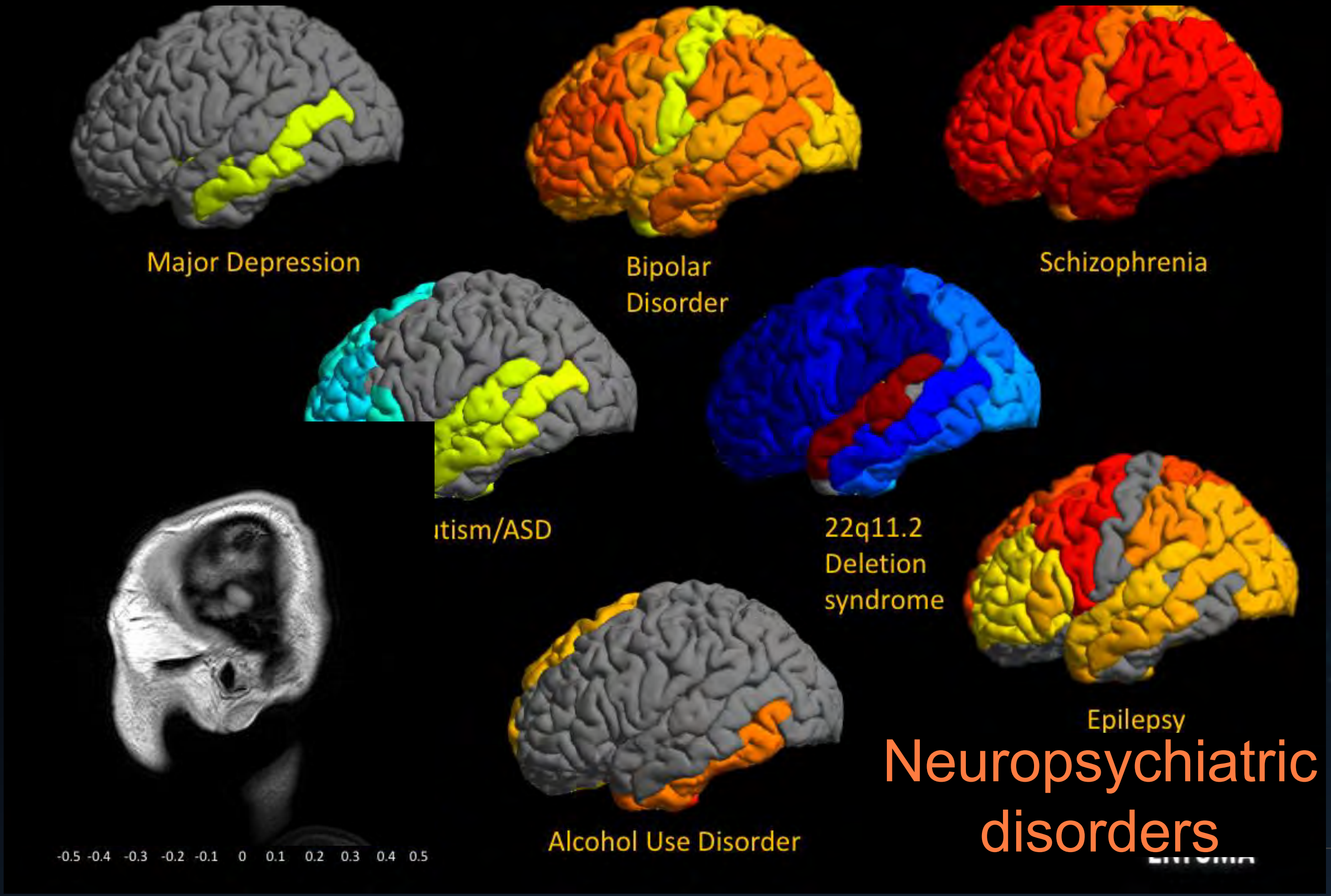


Part II

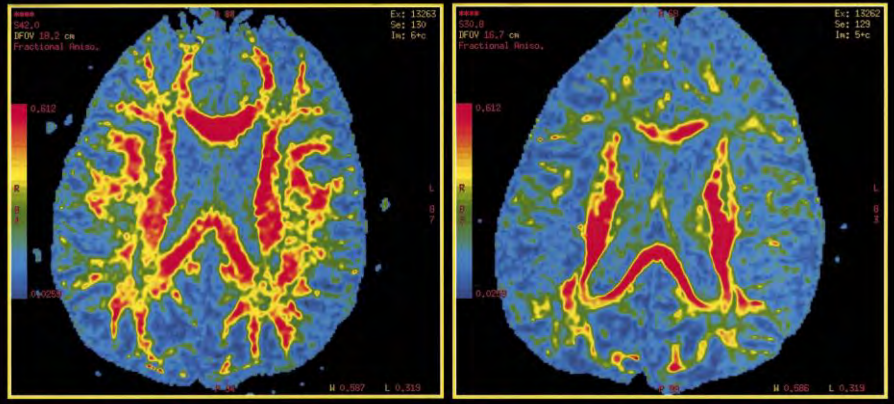
Big Data in Imaging Genetics

Brain Imaging for Brain Disorders

Capture the brain structure and function changes associated with major brain-related disorders and normal development

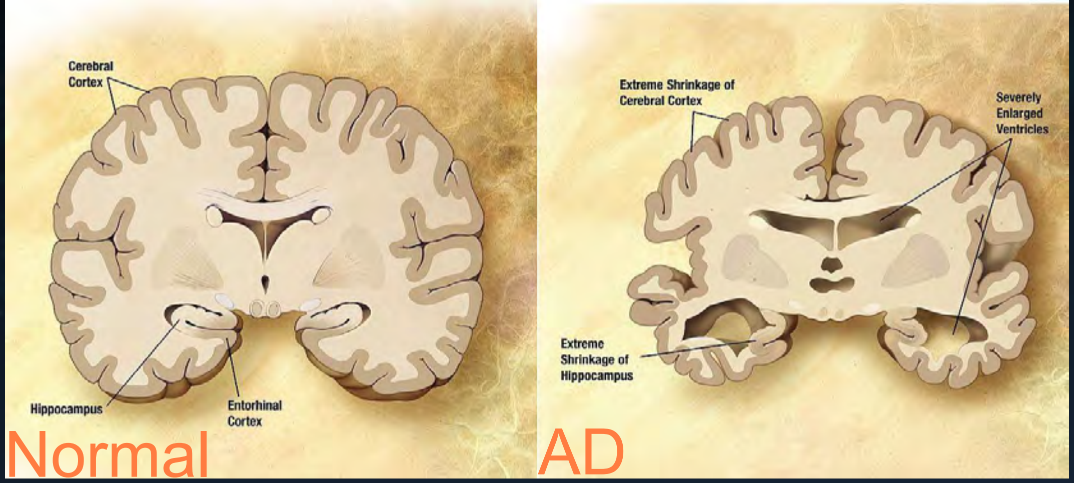


Neuropsychiatric disorders



Normal AD

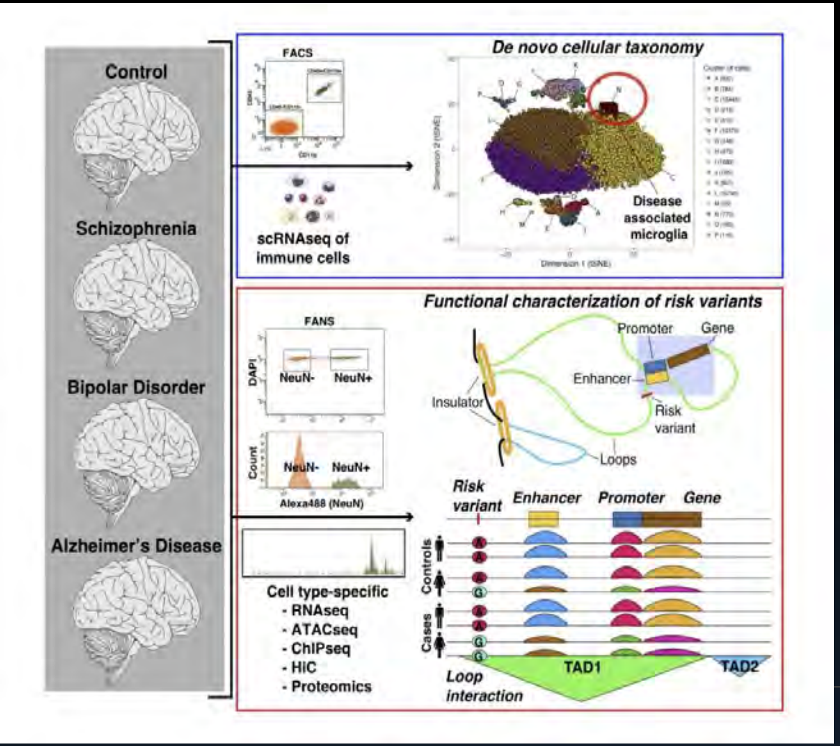
Alzheimer's disease (AD) is associated with brain shrinkage



Genetics of Brain Disorders

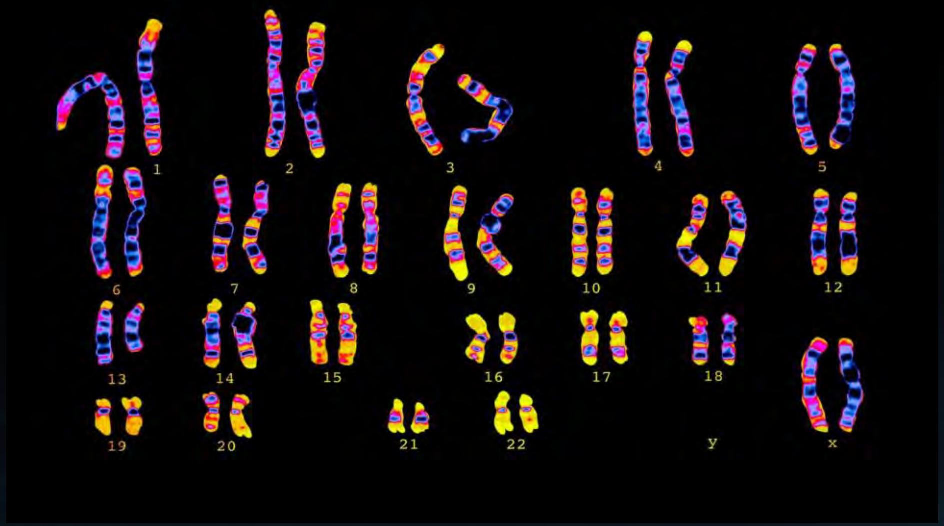
Most major brain disorders (like AD) are **heritable complex traits/diseases**

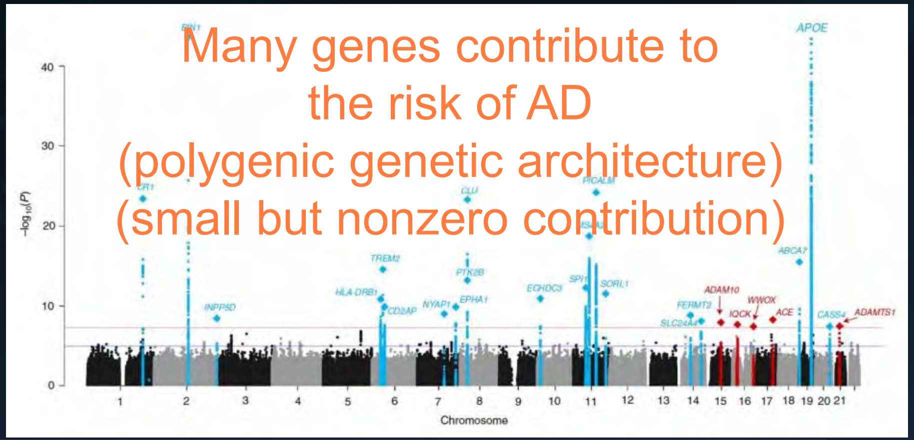
Together 50%-70% of AD risk
75%-90% of ADHD risk
60%-85% of Schizophrenia risk
~80% of Autism Spectrum Disorder (ASD) risk



Complex traits/diseases (many genes, environmental factors, complex functional mechanism)

Genetic signals are non-spare and weak:
Need large sample size to detect weak signals

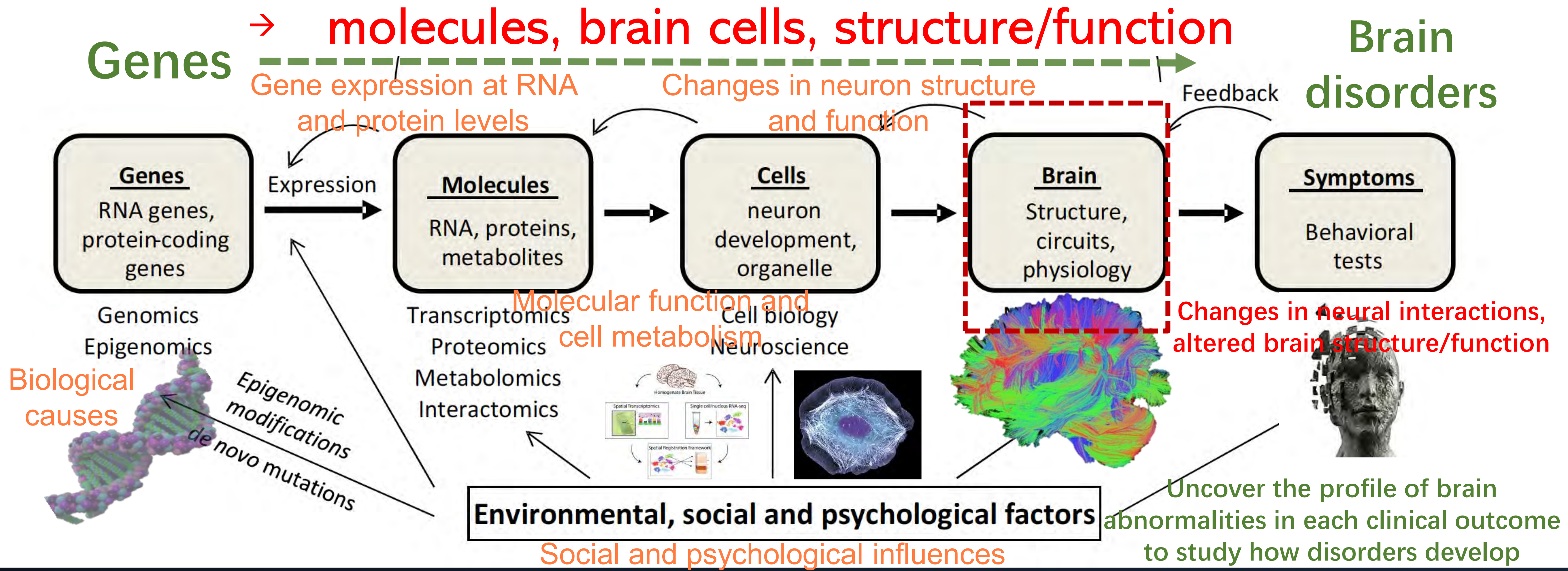




Many genes contribute to the risk of AD
(polygenic genetic architecture)
(small but nonzero contribution)

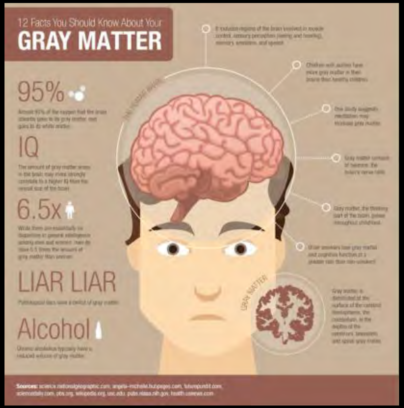
Brain Imaging Genetics Paradigm

Neuroimaging: an important component to help understand the complex biological pathways of brain disorders

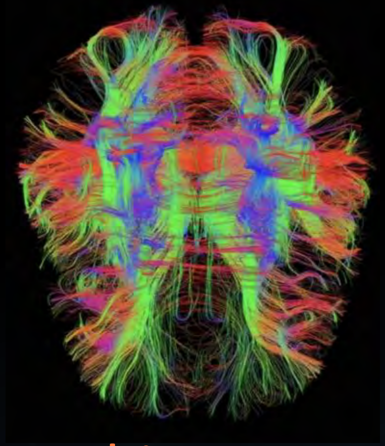


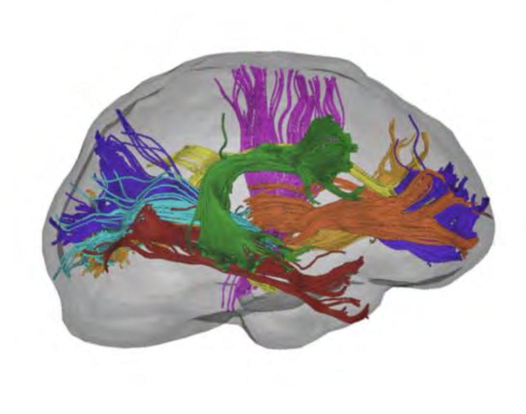
Brain Imaging Modality Examples

Harmonize tools/pipelines to consistently generate the full spectrum of neuroimaging features

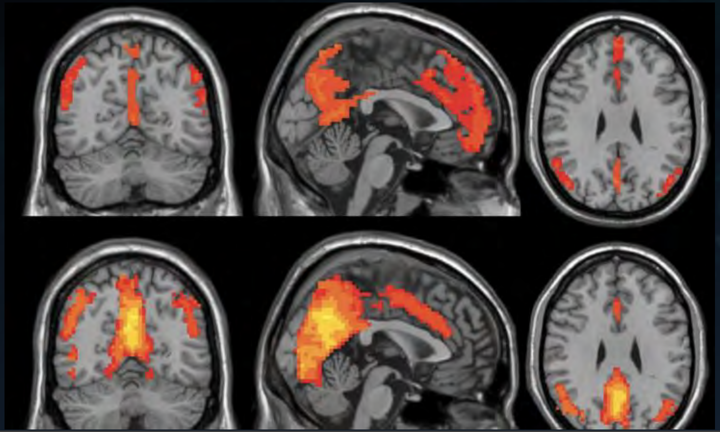


Cortical and subcortical structures





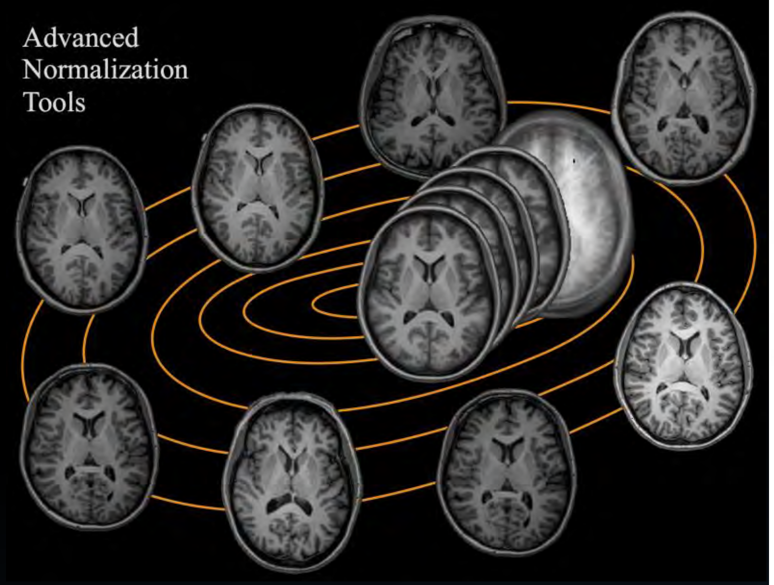
White matter microstructure (Structural connectivity, diffusion MRI)

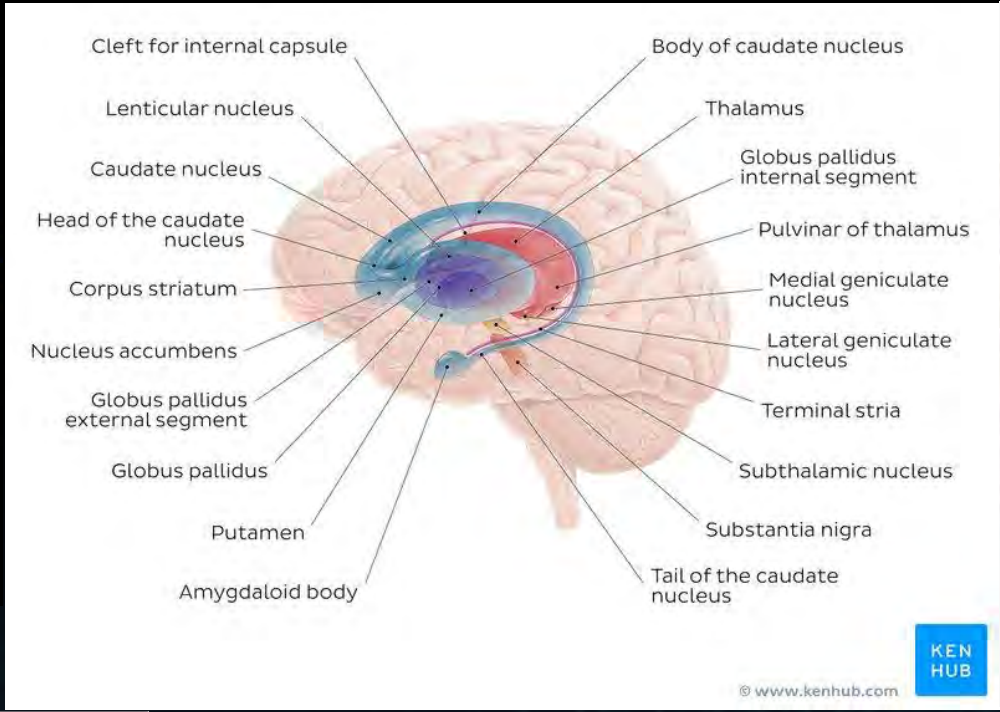


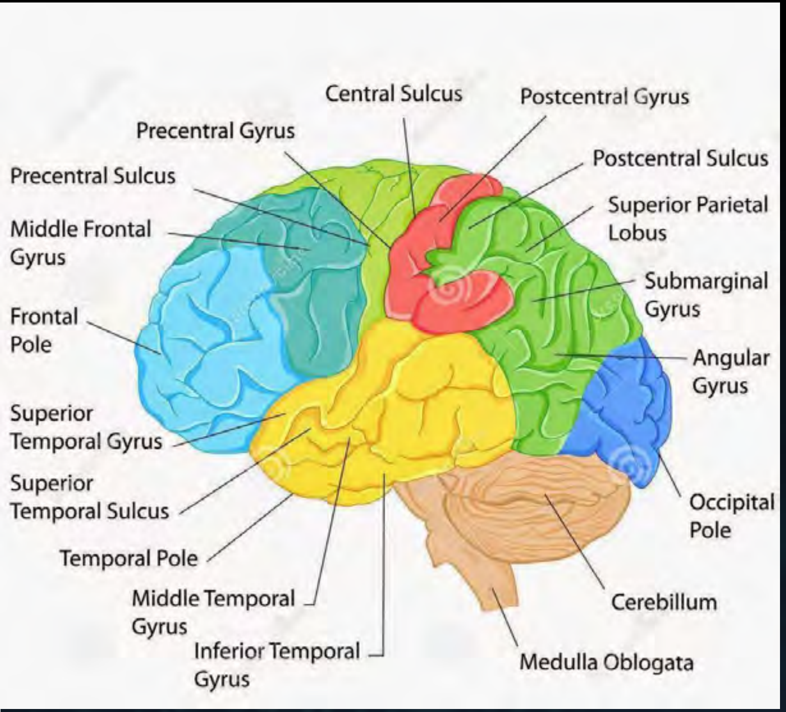
Functional networks (Functional connectivity, functional MRI)

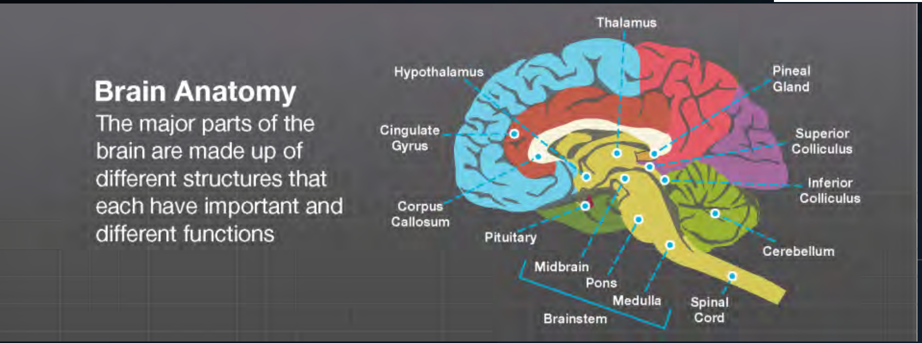
Regional Brain Volumes and Shape

Generate regional brain volumes and shape representations for 98 pre-specified brain regions and total grey matter, white matter, and brain volumes









Subcortical structures
(deep within the brain)

Cortical structures
(outer layer of the cerebrum)

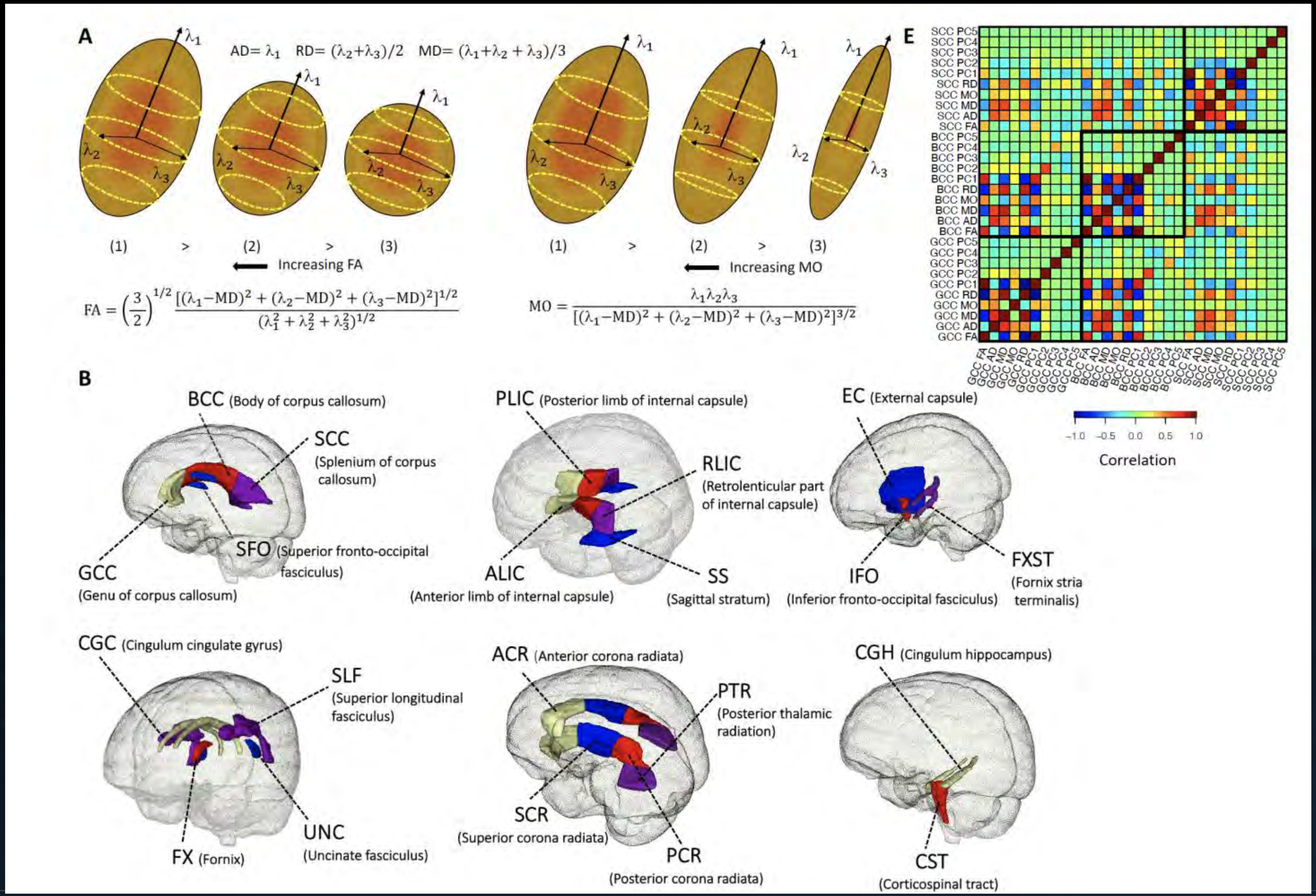
White Matter Microstructure

5 white matter microstructure measures (DTI parameters) for 21 white matter tracts

21 white matter tracts from ENIGAMA-DTI pipeline

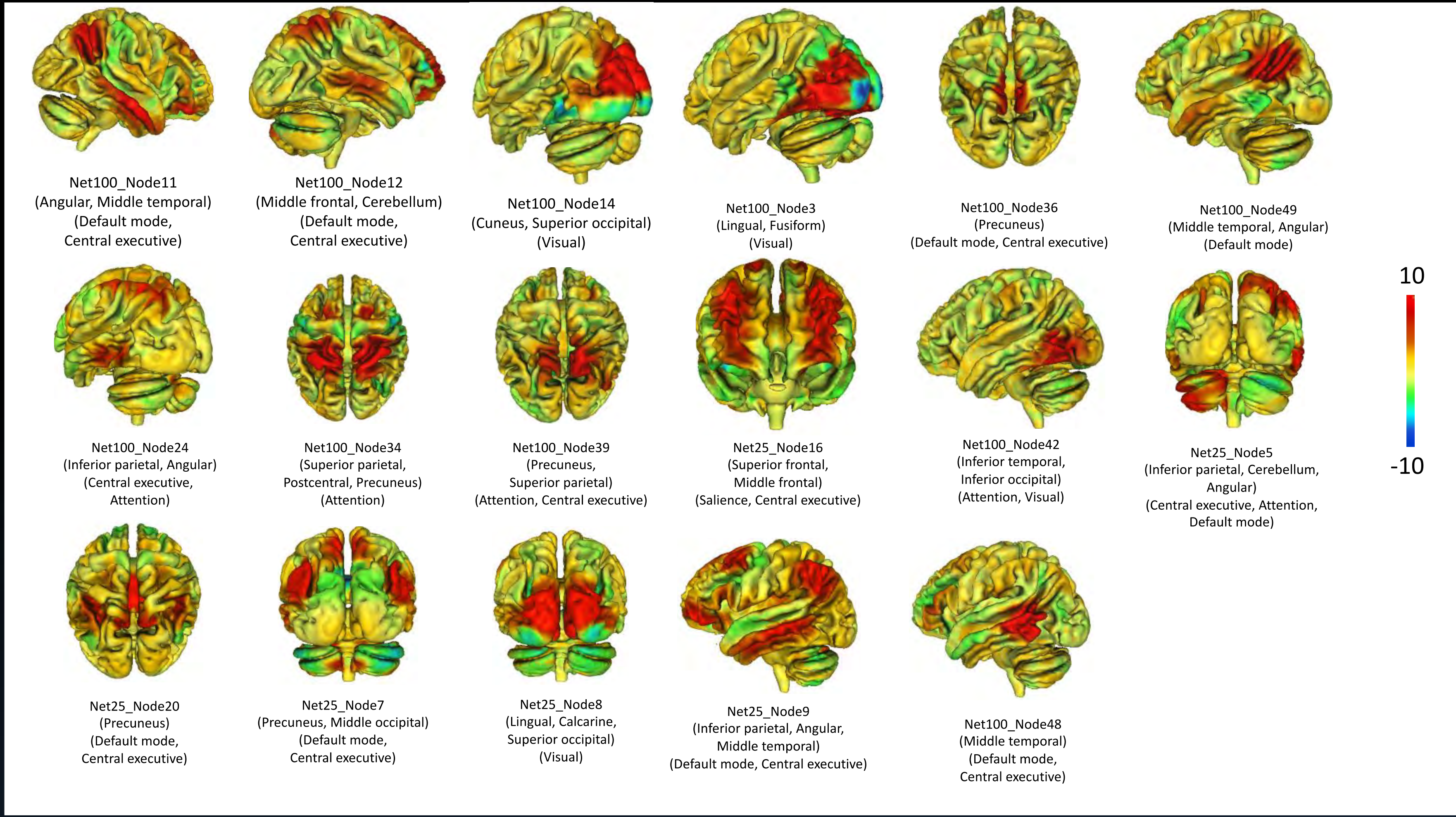
fractional anisotropy (FA)
mean diffusivity (MD),
axial diffusivity (AD),
radial diffusivity (RD), and
mode of anisotropy (MO)

sensitive to specific types of microstructural changes and have also been widely used in clinical research



Resting/task functional MRI (fMRI)

Independent component analysis (ICA)-based methods to form 76 functional regions and generate 1,701 functional connectivity traits



Characterize major functional brain regions and their connectivity

Brain Imaging Genetics: Learning Problems

(a) Learning Problems in Brain Imaging Genomics

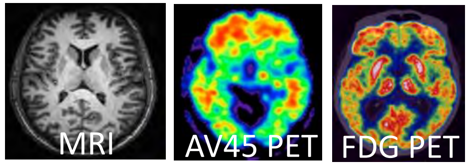
Heritability estimation of brain imaging phenotypes

Imaging genomics associations
- Fundamentals
- Meta-analysis
- Multivariate regression
- Bi-multivariate correlation

Integrating imaging and genomics for outcome prediction

(b) Biomedical Application Considerations

Brain imaging
- Voxels, ROIs, spatial correlation
- Multimodal, longitudinal studies
- Prior knowledge, brain network



Genomics
- SNPs, LD blocks
- Genes, pathways, networks
- Polygenic risk scores (PRSs)

Clinical outcome
- Diagnosis
- Progression
- Impairment score

Cognitive Condition

D A 2 End B 1 4 C Begin 3

Diagnosis

Normal	SMC
MCI	AD

(c) Statistical & Machine Learning Considerations

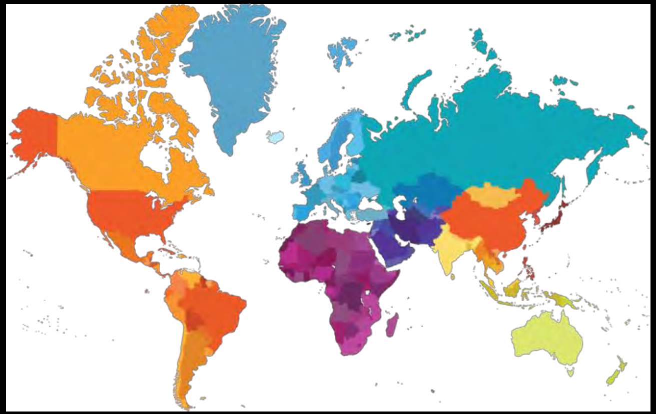
Increasing power
- Quantitative traits
- Multiple comparison
- Meta/mega analysis
- Multivariate models

Overfitting control
- Dimensionality reduction
- Regularization
- Knowledge-guided learning
- Outcome-guided learning

Other topics
- Biological interpretation
- Scalability
- Biased sampling
- Interaction

(Shen & Thompson Proc. of the IEEE 2020)

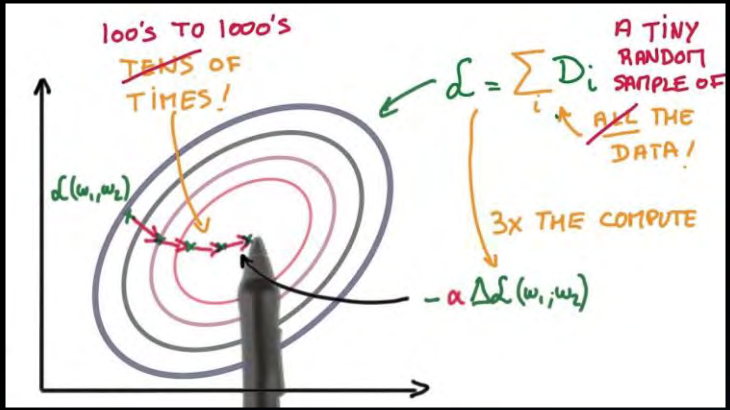
Methodological Challenges



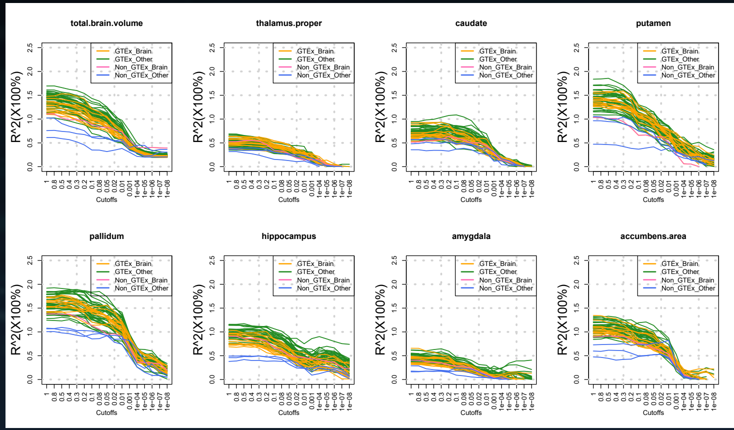
Multiple Biobanks Integration
(e.g., Heterogeneity in global populations)

Omics Data Integration
(e.g., new tech, biological pathway)

Database/ Tool/Theory for Brain Imaging Genetics



New Computational Tools
(e.g., challenge of dense signal in biobank-scale database)



Advanced Methods for Dense Signals
(e.g., deep learning)

Brain Imaging Genetics: Learning Problems

B. Zhao and ***H. Zhu*****.** On genetic correlation estimation with summary statistics from genome-wide association studies. *Journal of American Statistical Association*, in press, 2021.
Zhao, Y., Li, T.F., and ***Zhu, H.*** Bayesian sparse heritability analysis with high-dimensional neuroimaging phenotypes。 *Biostatistics*， in press， 2020.
Zhou, F., Zhou, H.B., **Li, T.**, and ***Zhu, H.*** Analysis of Secondary Phenotypes in Multi-group Association Studies. *Biometrics*, in press, 2020.
Kong，D. H., An, B. G., Zhang, J. W., and Zhu, H. L2RM: Low-rank Linear Regression Models for High-dimensional Matrix Responses. *Journal of American Statistical Association*, in press, 2020.
Benjamin, R. and Zhu, H.T. ACE of Space: Estimating Genetic Components of High-Dimensional Imaging Data. *Biostatistics*, in press， 2020.
Zhang, J., Xia, K., **Ahn, M**., Jha, S.C., Blanchett, R., Crowley, J.J., Szatkiewicz, J.P., Zou, F., Zhu, H., Styner, M., Gilmore, J.H., Knickmeyer, R.C. Genome-Wide Association Analysis of Neonatal White Matter Microstructure. *Cerebral Cortex*， in press， 2020.
Yize Zhao, Hongtu Zhu, **Zhaohua Lu**, Rebecca Knickmeyer, and Fei Zou. Bayesian Hierarchical Variable Selection for Structured Genome-wide Association Studies。 *Genetics,* 212, 397-415, 2019.
S.J. Lee, **J.W.Zhang**, M. C. Neale, M. Styner, H. Zhu, J.H. Gilmore. Quantitative tract-based white matter heritability in 1- and 2-year-old twins. *Human Brain Mapping*, 40, 1164-1173, 2019.
Zhang, J.W., Ibrahim, J. G., **Li, T.F.**, and **Zhu, H.T**. A Powerful Global Test Statistic for Functional Statistical Inference. *AAAI 2019.*
Farouk Nathoo, Linglong Kong, and Zhu, Hongtu. A Review of Statistical Methods in Imaging Genetics. *The Canadian Journal of Statistics*, 47, 108-131, 2019.
Huang, C., Thompson, P., Wang, Y., **Yu, Y., Zhang, J., Kong, D.**, Colen, R., Knickmeyer, R., ***Zhu, H. T.*** "FGWAS: Functional genome wide association analysis." *NeuroImage*, 159, 107-121, 2017.
Zhaohua Lu, Zakaria Khondker, Joseph G Ibrahim, **Yue Wang**, and ***H. Zhu.*** Bayesian longitudinal low-rank regression models for imaging genetic data from longitudinal studies. *NeuroImage,* 149:305-322, 2017.
Zhang, J.W., Ibrahim, J.G., R.C. Knickmeyer, M. Styner, Gilmore, J. H., and ***H. Zhu.*** HFPRM: Hierarchical Functional Principal Regression Model for Diffusion Tensor Image Bundle Statistics. *IPMI 2017*.
Zhu, W. S., **Y.Yuan, J. Zhang, F. Zhou,** R. C. Knickmeyer, and ***H.Zhu.*** Genome-wide Association Analysis of Secondary Imaging Phenotypes from the Alzheimer's Disease Neuroimaging Initiative Study. *NeuroImage,* 146:983-1002, 2016.
Lu, Z. H., ***Zhu, H.T.,*** R.C. Knickmeyer, P.F. Sullivan, W.N. Stephanie, and Fei Zou, Multiple SNP-sets Analysis for Genome-wide Association Studies through Bayesian Latent Variable Selection. *Genetic Epidemiology*, 39, 664-677, 2015.
D. Kong, K. S. Giovanello, Y.L. Wang, W. Lin, E. Lee, Y. Fan, P. M. Doraiswamy, and ***Hongtu Zhu***, ADNI (2015). Predicting Alzheimer's disease using combined imaging-whole genome SNP data. *Journal of Alzheimer's Disease,* 46: 695-702.
M.Huang, T.Nichols, C.Huang, Y.Yang, Z. Lu, Q. Feng, R.C. Knickmeyer, ***H.Zhu,*** and for ADNI. (2015). FVGWAS: Fast Voxelwise Genome Wide Association Analysis of Large-scale Imaging Genetic Data. *NeuroImage*, 118, 613-627. **Winner of Best Paper award in ASA SI Session, 2015.**
Lin, J, ***Zhu, H.T.,*** **Ahn, M**., Sun, W, and Ibrahim, J. G. Functional mixed effects models for imaging genetic data. *Genetic Epidemiology,* 38, 680-691, 2014.
Zhu, H.T., Khondker, Z. S., Lu, Z.H., and Ibrahim, J. G. Bayesian generalized low rank regression models for neuroimaging phenotypes and genetic markers. *Journal of American Statistical Association,* 109, 1084-1098, 2014.
Lin, J., ***Zhu, H.T.,*** Knickmeyer, R., Styner, M., Gilmore, J. and Ibrahim, J.G. Projection Regression Models for Multivariate Imaging Phenotype. *Genetic Epidemiology,* 36, 631-641, 2012.

Part III

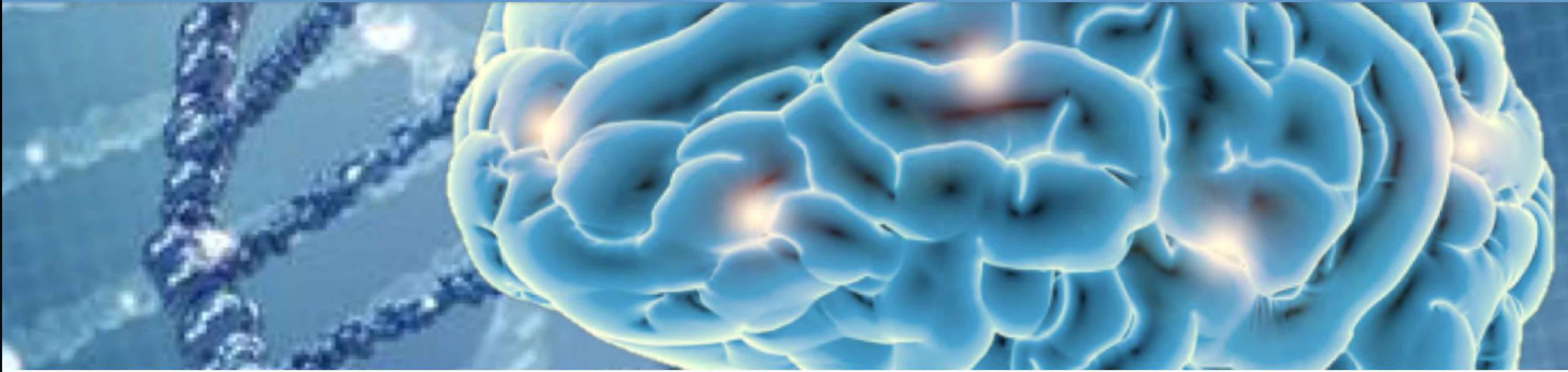
Novel Clinical findings

Brain Imaging Genetics Knowledge Portal (BIG-KP)

Genetics Discoveries in Human Brain by Big Data Integration

bigkp.org

Imaging Genetics Online Server | GWAS Summary Statistics Data Download | UNC BIG-S2 Lab | BIG-S2 Github | Other Resources



Aim to build the best knowledge database of neuroimaging genetics

GWAS Locus Browser

Brain Imaging Genetics Summary Statistics | Search for a variant, gene, or phenotype | Phenotypes | Top Hits | Random | About

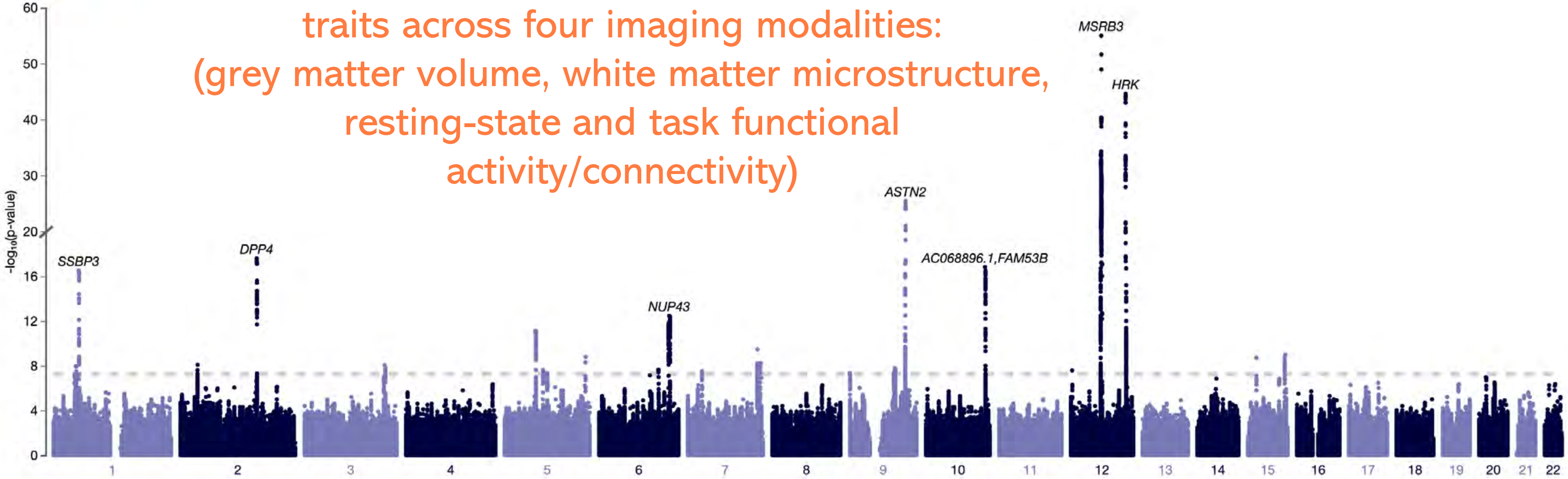
left.hippocampus

Category: **sMRI**

Download summary statistics

Manhattan QQ

Searchable database for 1593+ neuroimaging traits across four imaging modalities: (grey matter volume, white matter microstructure, resting-state and task functional activity/connectivity)



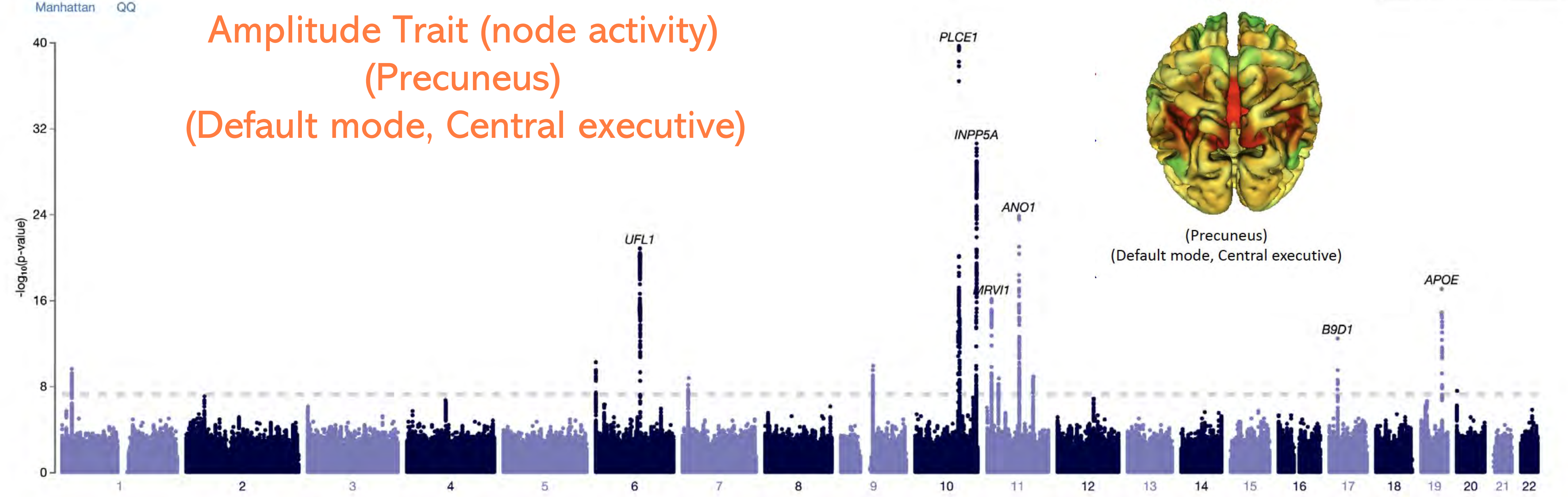
GWAS Locus Browser

Net25_Node20

Category: **rs-fMRI**

Download summary statistics

Manhattan QQ



GWAS Summary Statistics

The full set of GWAS summary statistics have been made freely available to the research community

GWAS Summary Statistics for Brain Imaging Phenotypes

Resources with the largest sample size

(>4,350 page views since Sep 2019)

Involved datasets: UK Biobank (UKB), Adolescent Brain Cognitive Development (ABCD) Study, Human Connectome Project (HCP), Philadelphia Neurodevelopmental Cohort (PNC), Alzheimer's Disease Neuroimaging Initiative (ADNI), Pediatric Imaging, Neurocognition, and Genetics (PING)

Terms of Use:

- By downloading these data, you acknowledge that they will be used for research purposes and that you are in compliance with applicable rules, policies and regulations.
- When reporting results of research that utilizes these data we request that you cite the original publication.

GWAS summary statistics for 200 resting-state functional MRI (rs-fMRI) traits

- **Sample size:** n=34,691
- **Version**: July 15, 2020
- Download Summary Statistics:

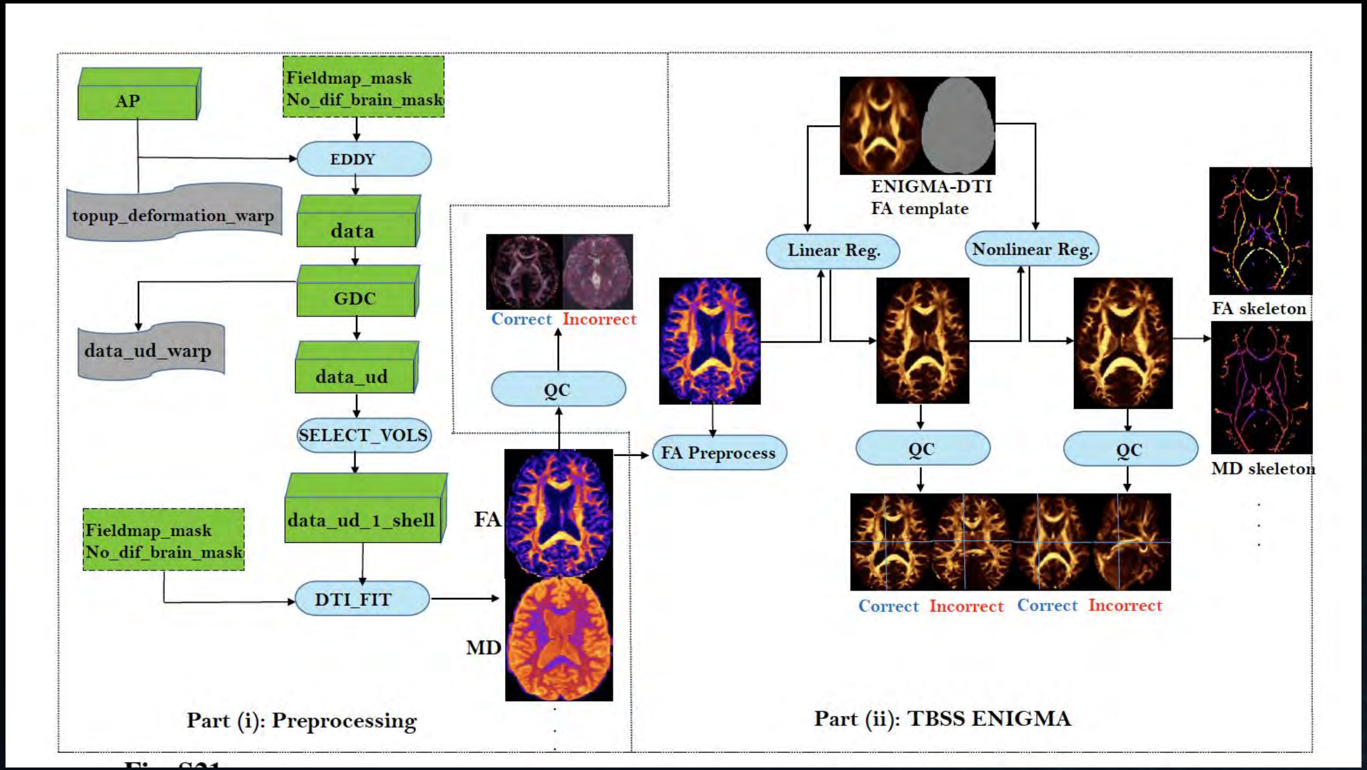
```
wget --no-check-certificate --content-disposition https://raw.
githubusercontent.com/stat-yyang/sumstats/master/fMRI.list
wget -i fMRI.list
```

- **Description**: readme
- **Citation**: Zhao et al (2020) Common variants contribute to intrinsic

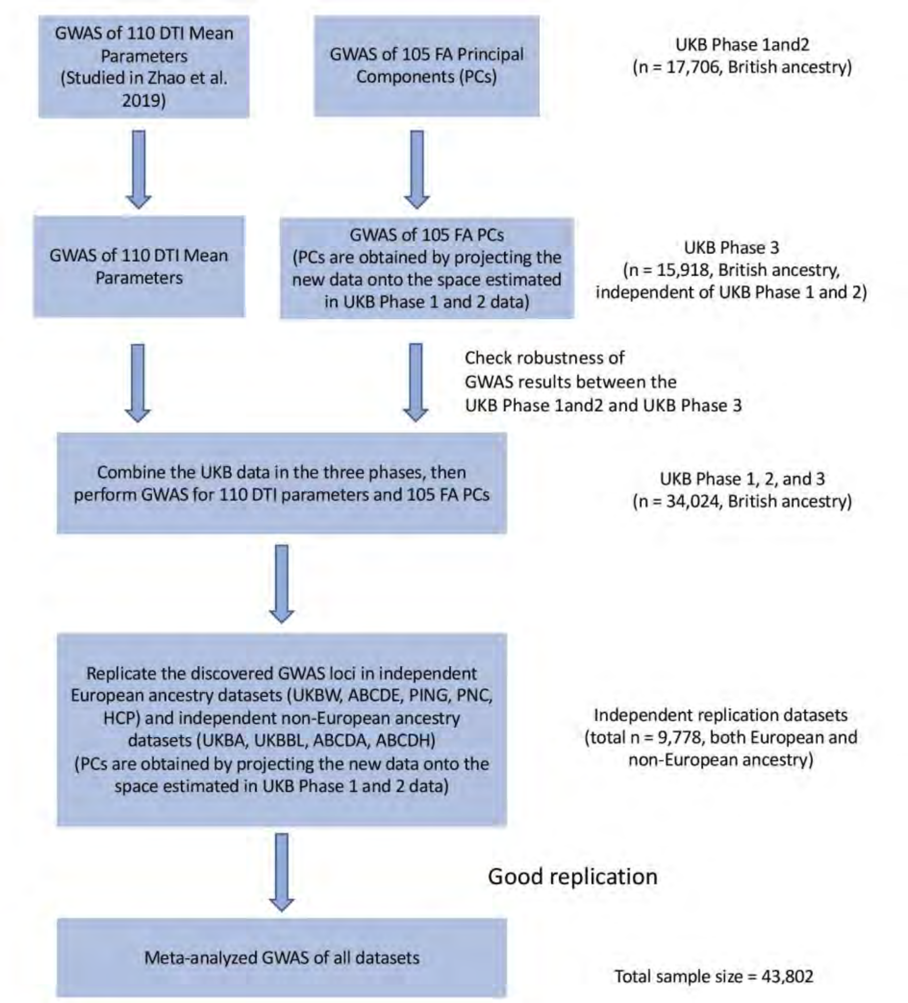
Contents [hide]

GWAS of White Matter Tracts

Overview of the ENIGMA-DTI pipeline and the multiple-stage design in GWAS



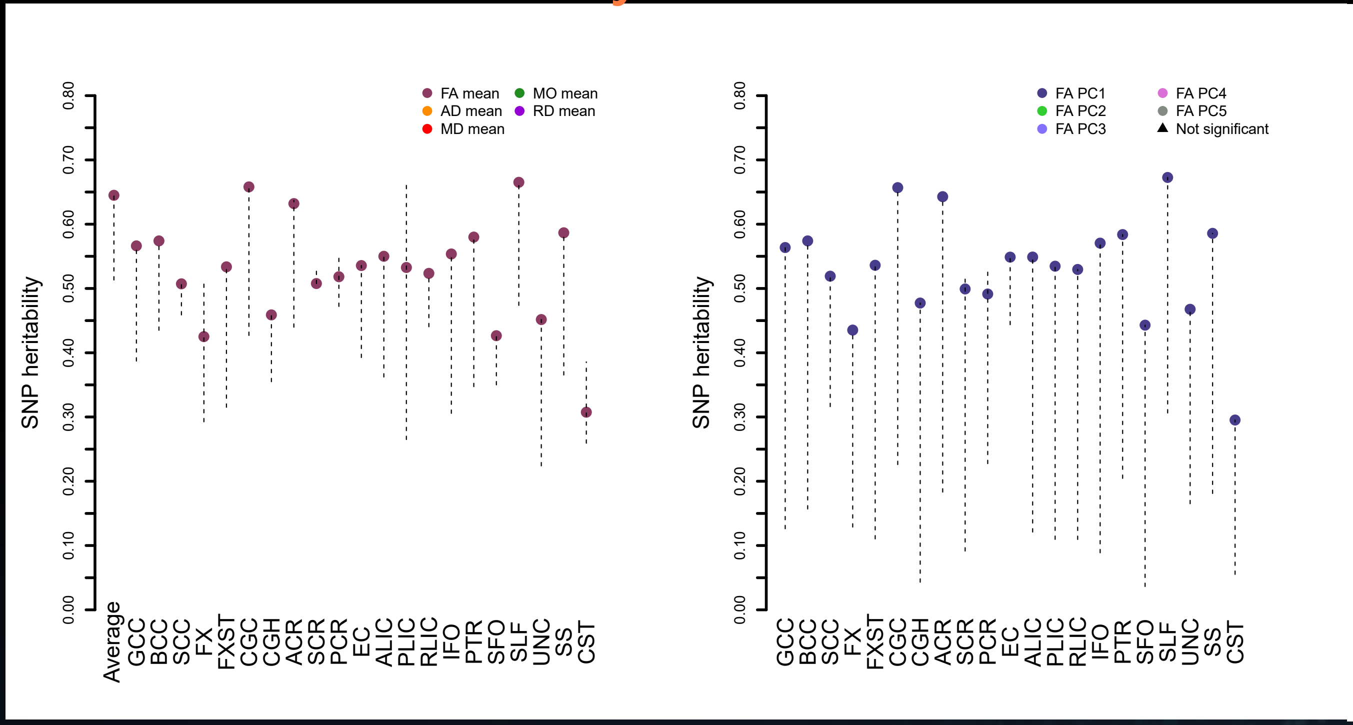
Apply the same pipeline in different datasets (UKB, ABCD, PING, PNC, HCP)



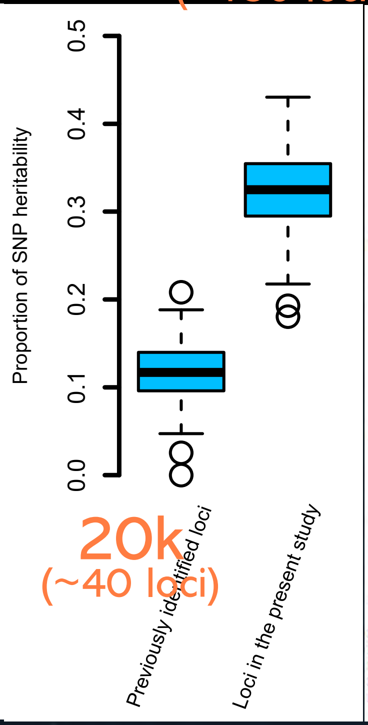
Genetic Architecture of White Matter

We observed 109 novel genomic regions (151 in total, $P < 2.3\text{e-}10$, 5e-8/215) associated with white matter microstructure

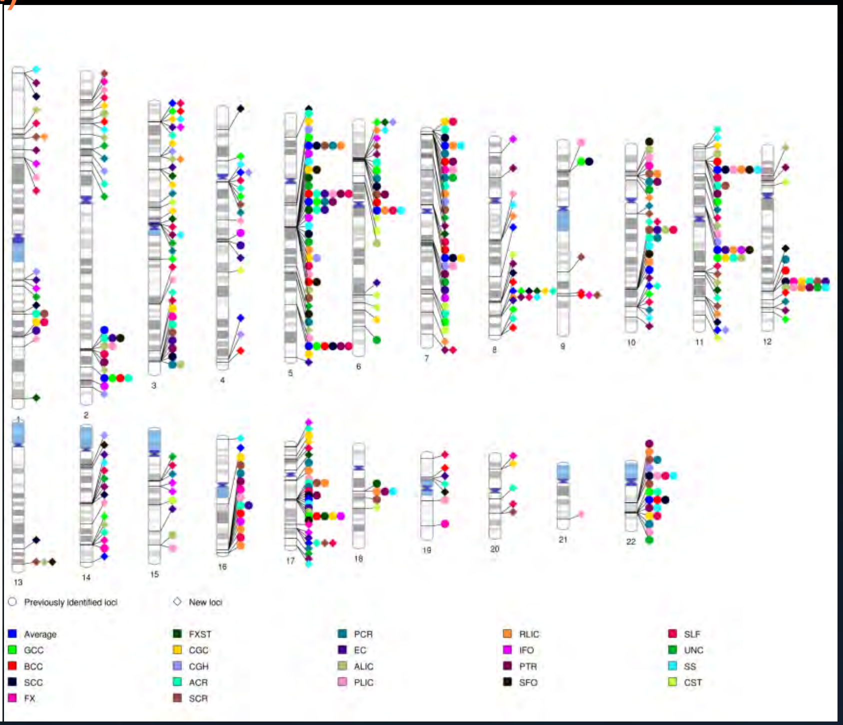
Heritability $h^2 \sim 45\%$



40k
(~150 loci)



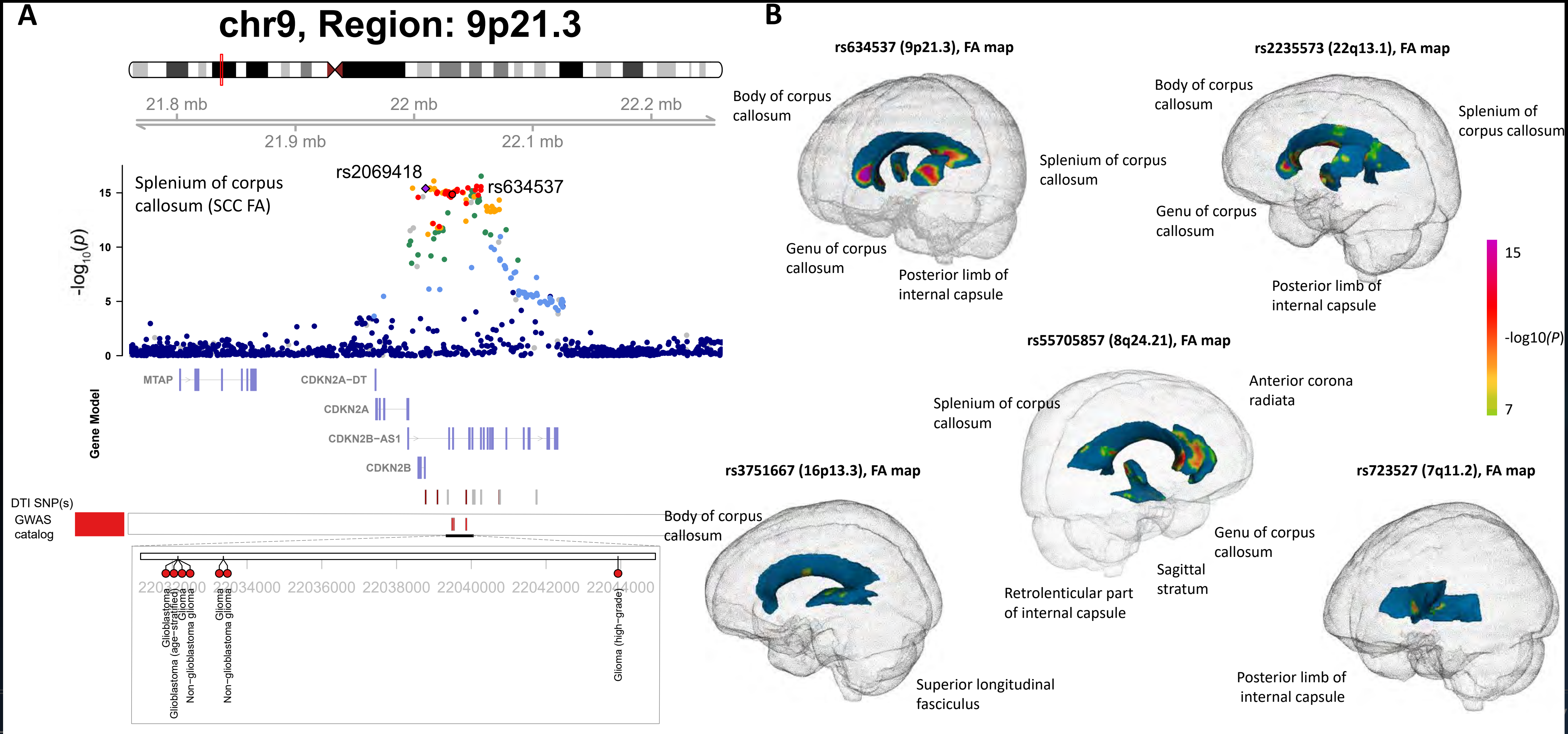
Ideogram of genomic regions



Sample size is essential for gene discovery of traits with highly polygenic genetic architecture

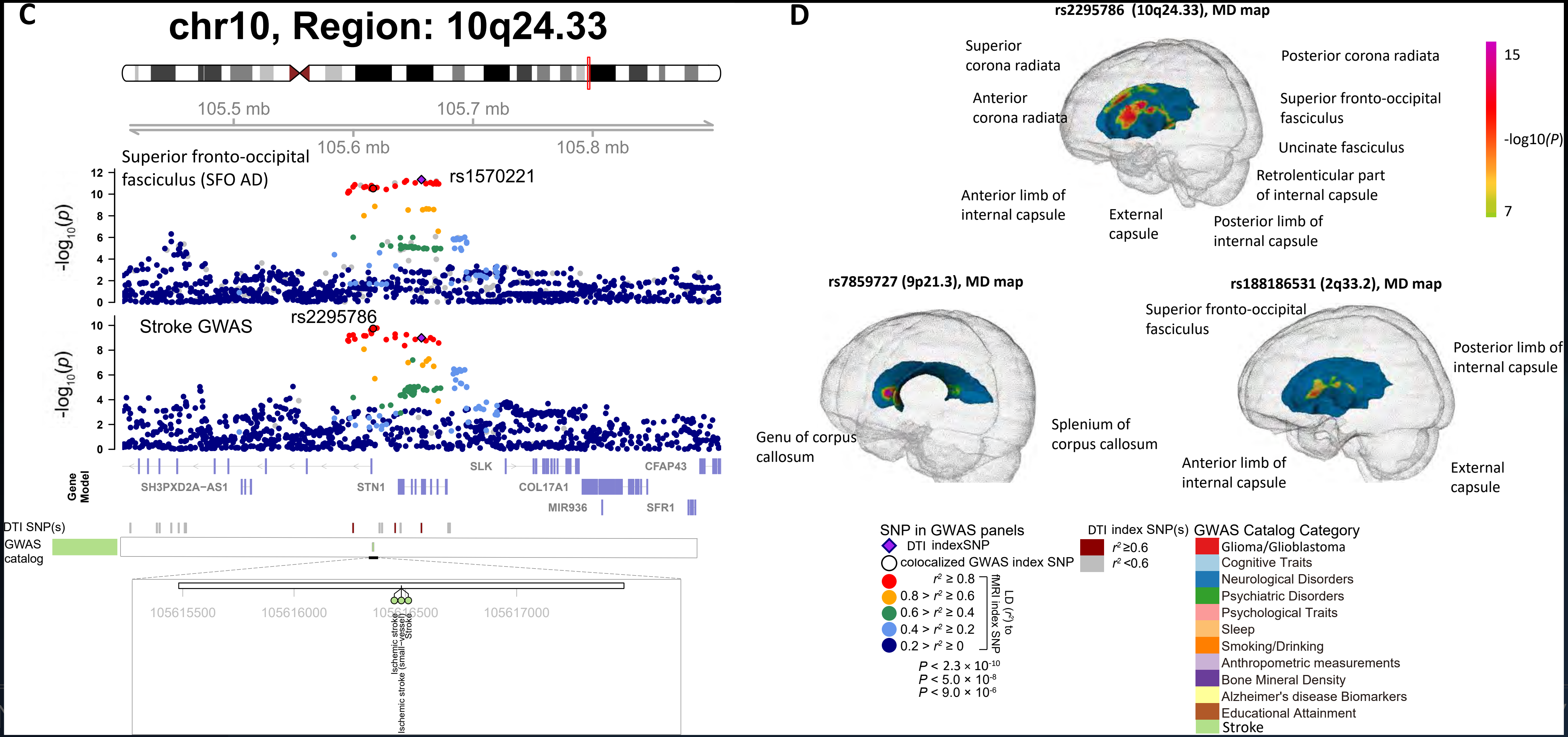
Colocalization with Glioma/GBM

For the 25 known genomic risk regions of Glioma/GBM, 11 are associated with white matter microstructure



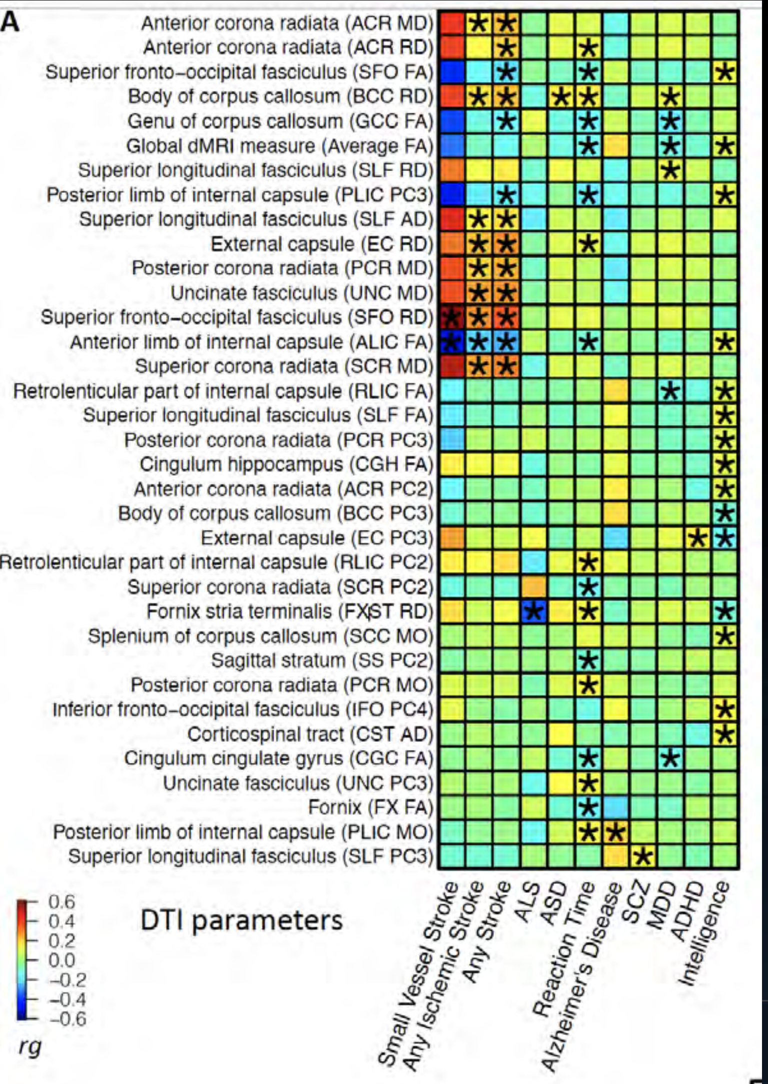
Colocalization with Stroke

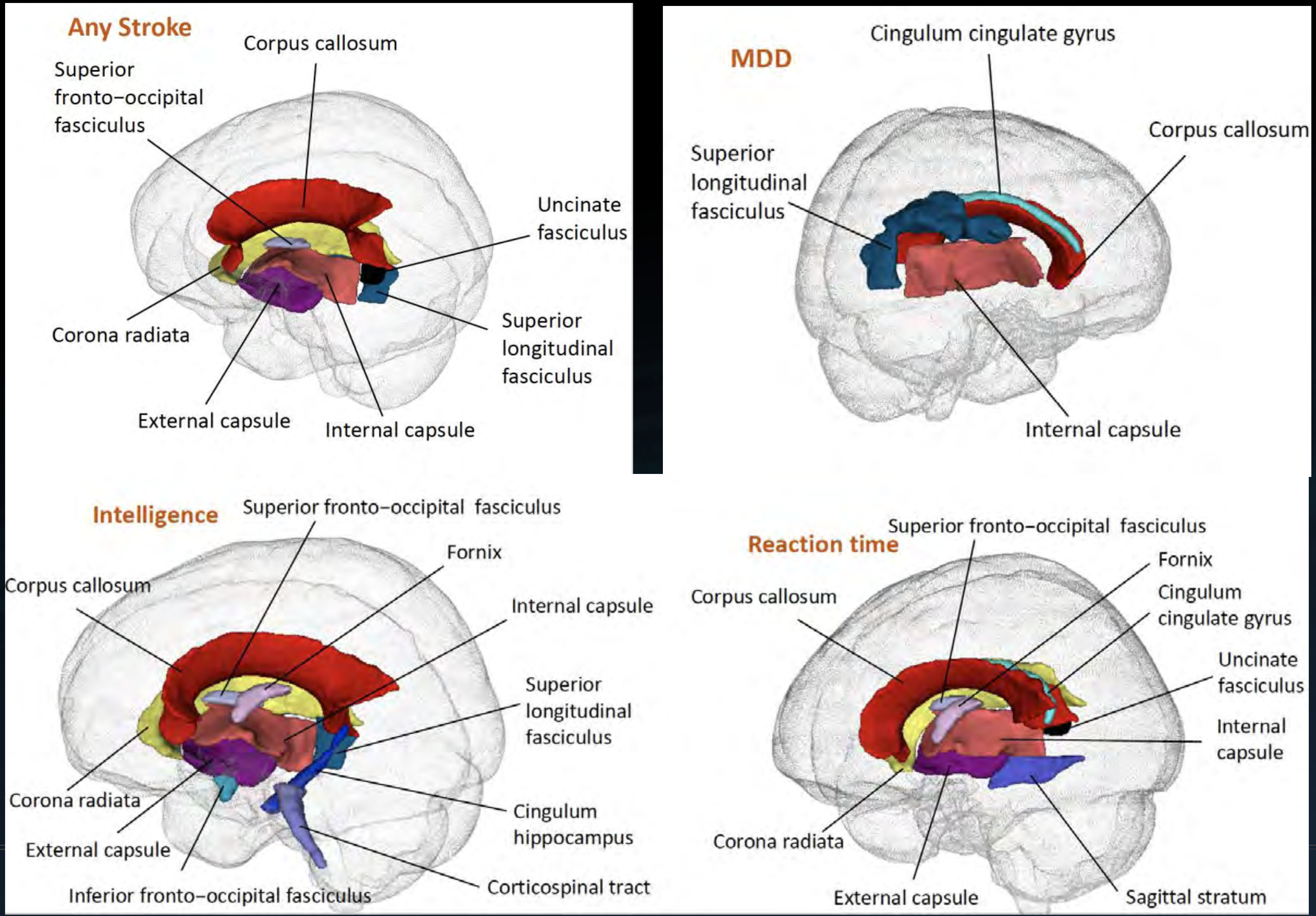
Genetic colocalizations among vascular risk factors (e.g., obesity, diabetes, high blood pressure), white matter microstructure, and stroke



Genetic Correlations with Brain Disorders

Strong genetic correlation between white matter microstructure and small vessel stroke subtype

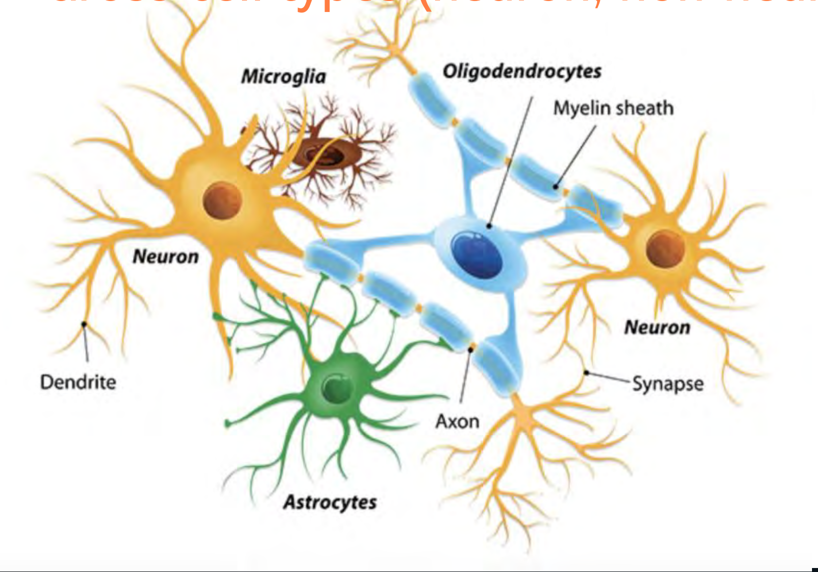


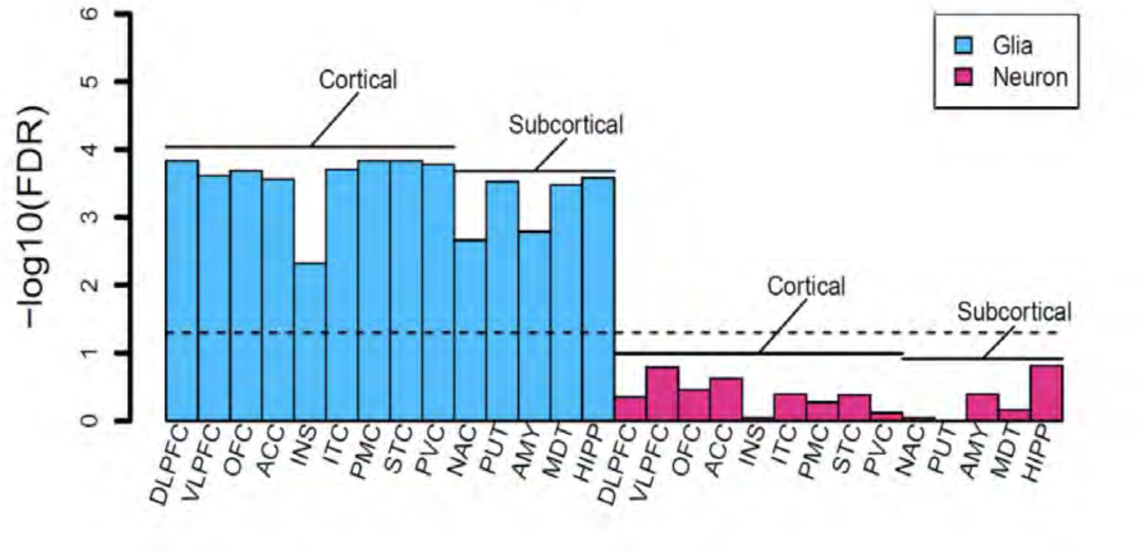


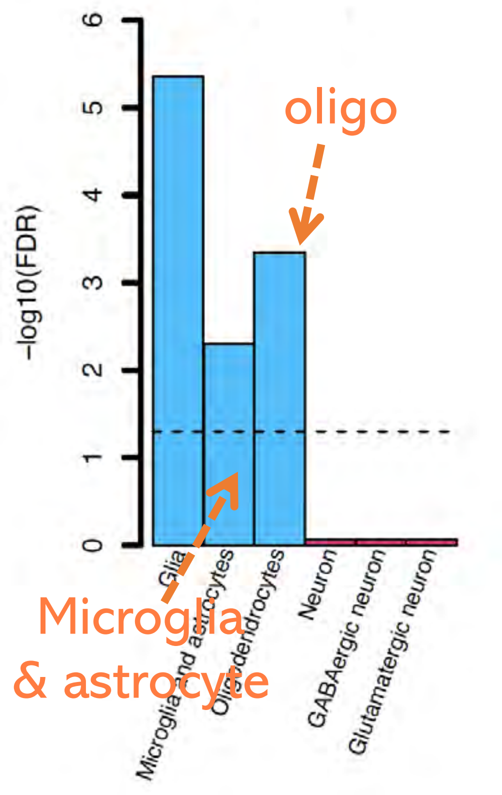
Heritability Enrichment in Brain Cells

Identify brain cell types where genetic variation leads to changes in white matter connectivity

Gross cell types (neuron, non-neuron [glia, including oligo, microglia, astrocyte])





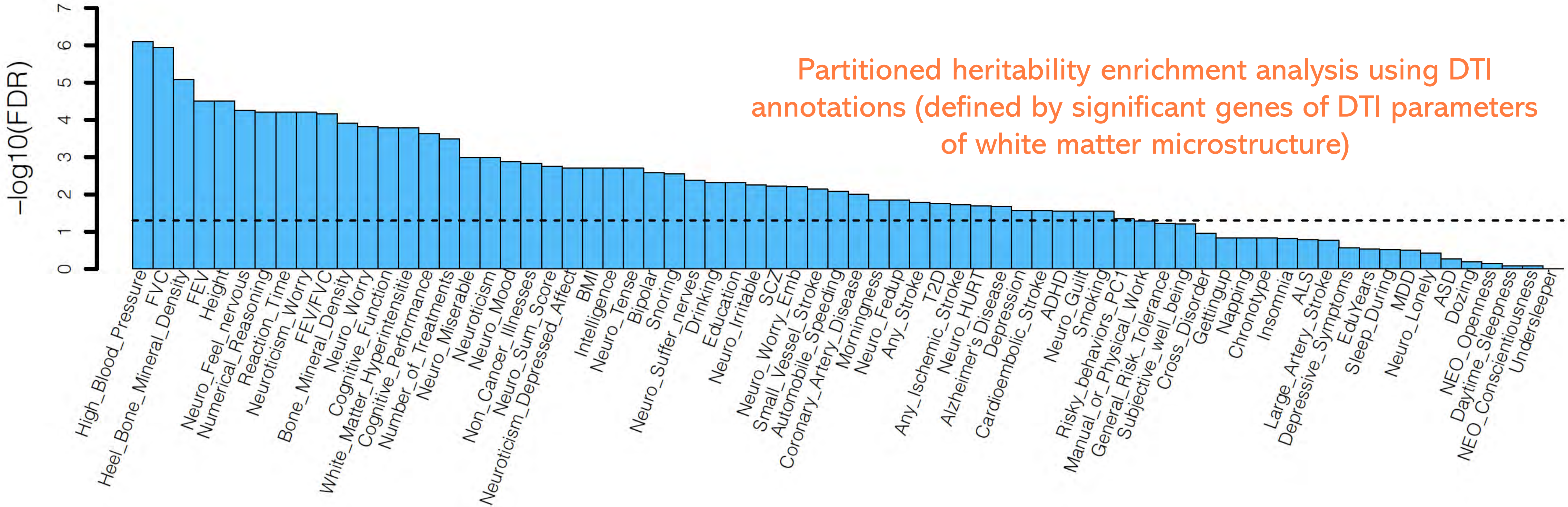


White matter is largely composed of glial cell types (oligo, microglia, astrocyte)

Oligo annotation accounted for 10.4% heritability while only composed 0.3% of the genetic variants

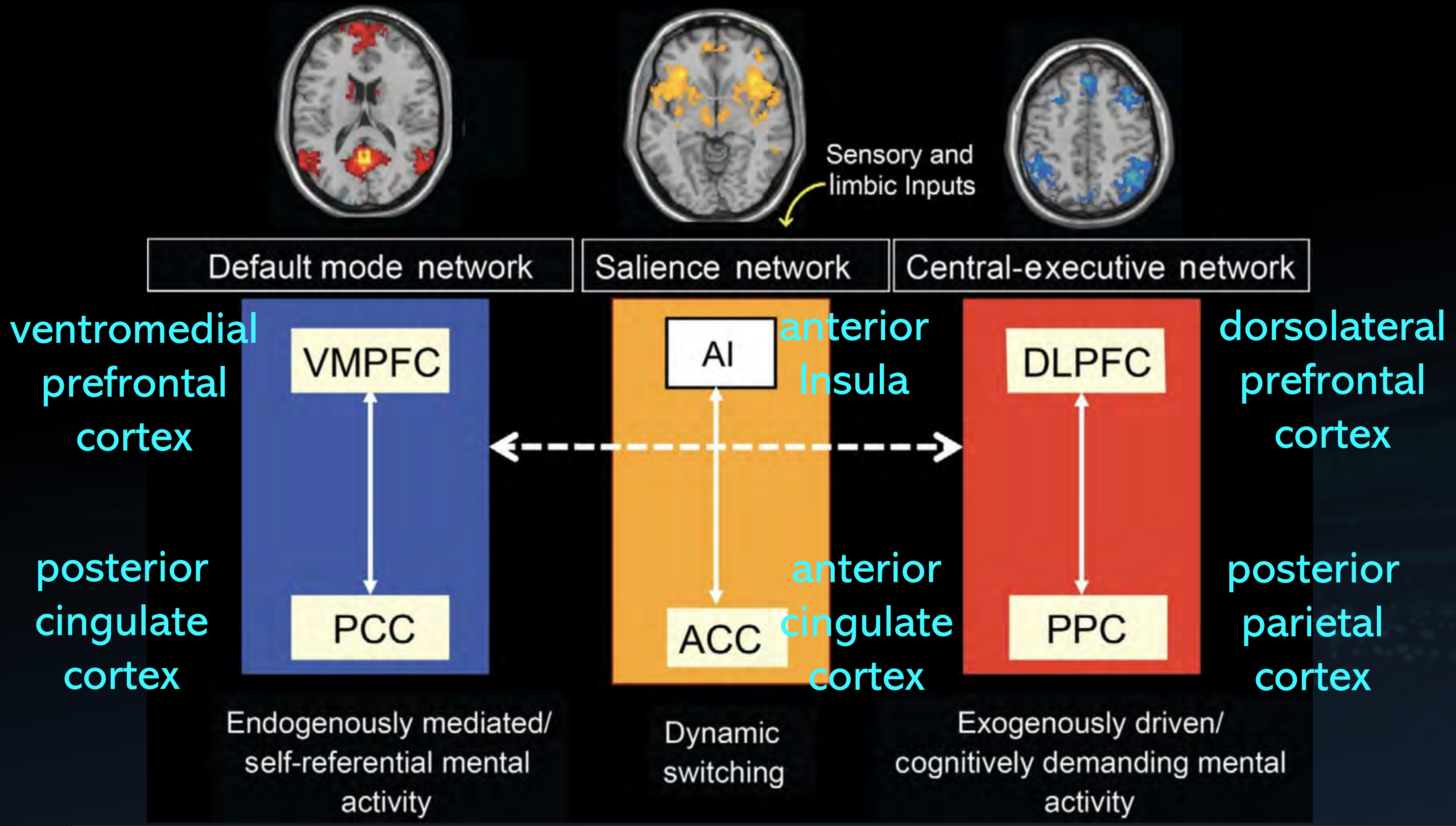
DTI annotation enrichment

Heritability of 49 complex traits was significantly enriched in genetic regions influencing white matter microstructure, such as stroke, schizophrenia, ADHD, bipolar Alzheimer's Disease, T2D, high blood pressure, and coronary artery disease



Triple Network Model of Psychopathology

The salience network (SN) plays a crucial role in dynamic switching between the central executive (CE) and default mode (DM) networks

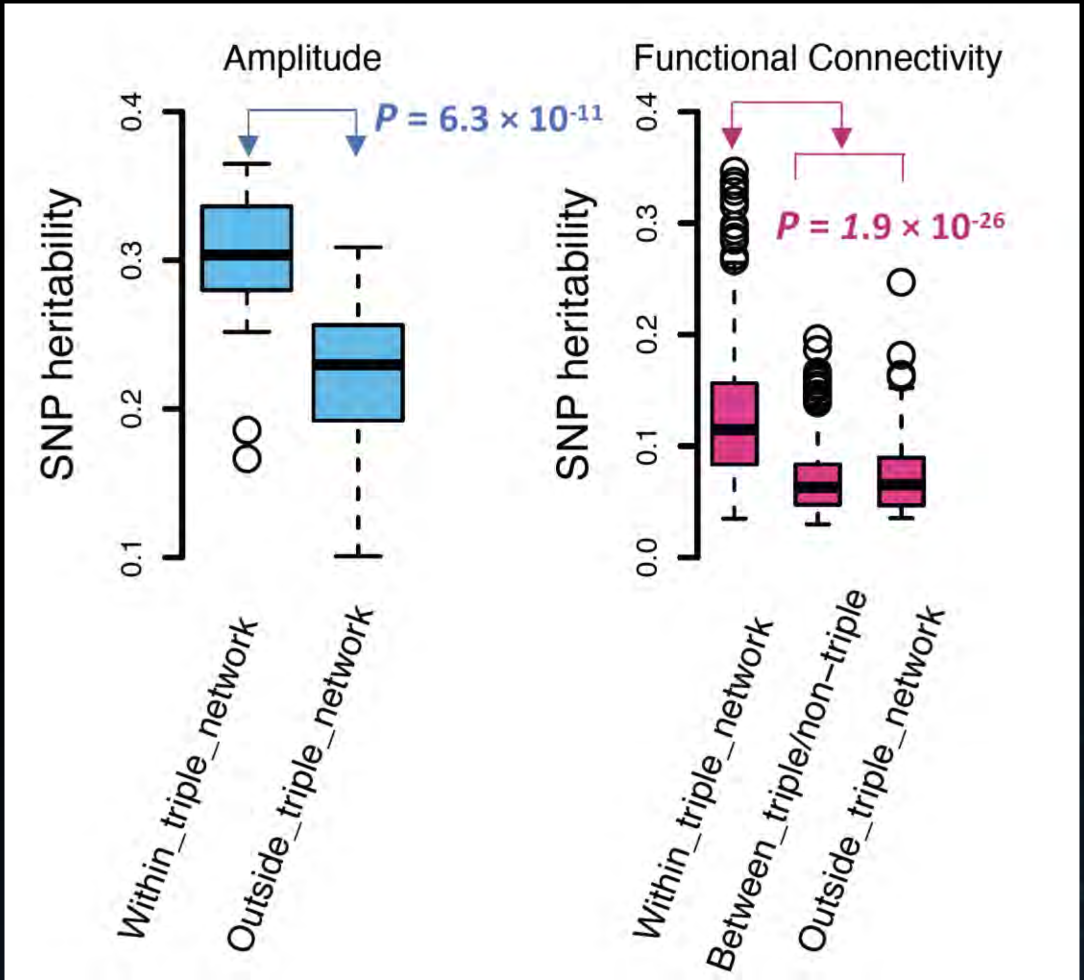


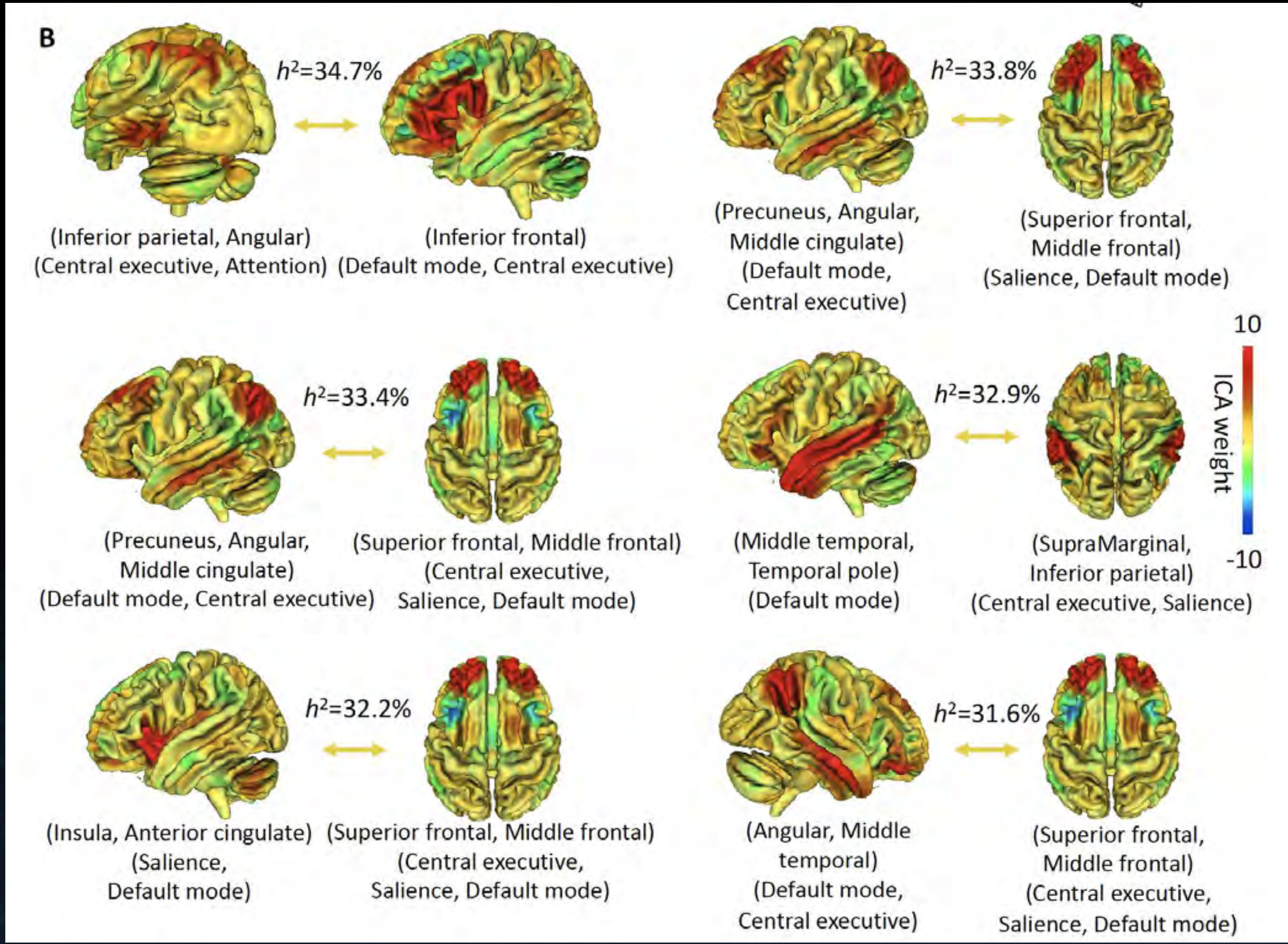
Three core functional networks that support efficient cognition

Related to major brain disorders, such as Alzheimer's disease (AD), Parkinson's disease (PD), and major depressive disorder (MDD)

Genetics of the Triple Networks

Higher heritability than other functional networks (e.g., motor, vision)

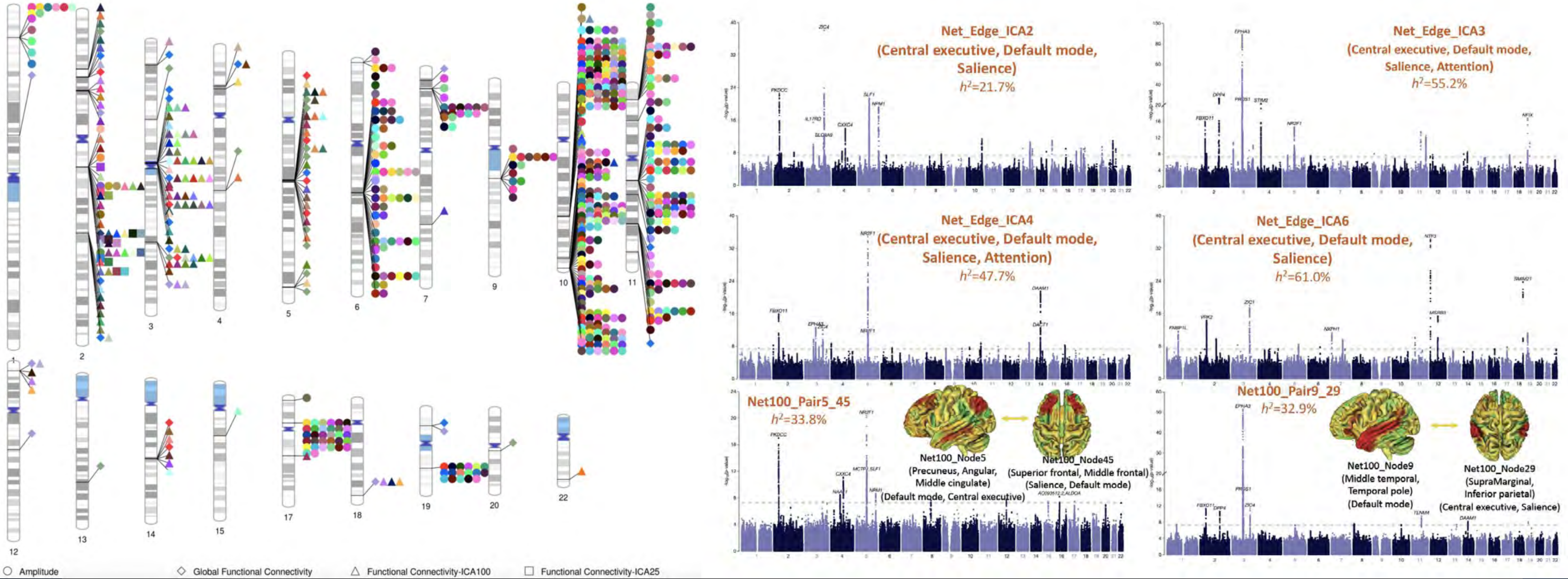




The level of genetic control is higher in the triple networks, which closely control multiple cognitive functions and affect major brain disorders

Genetics of Functional Brain

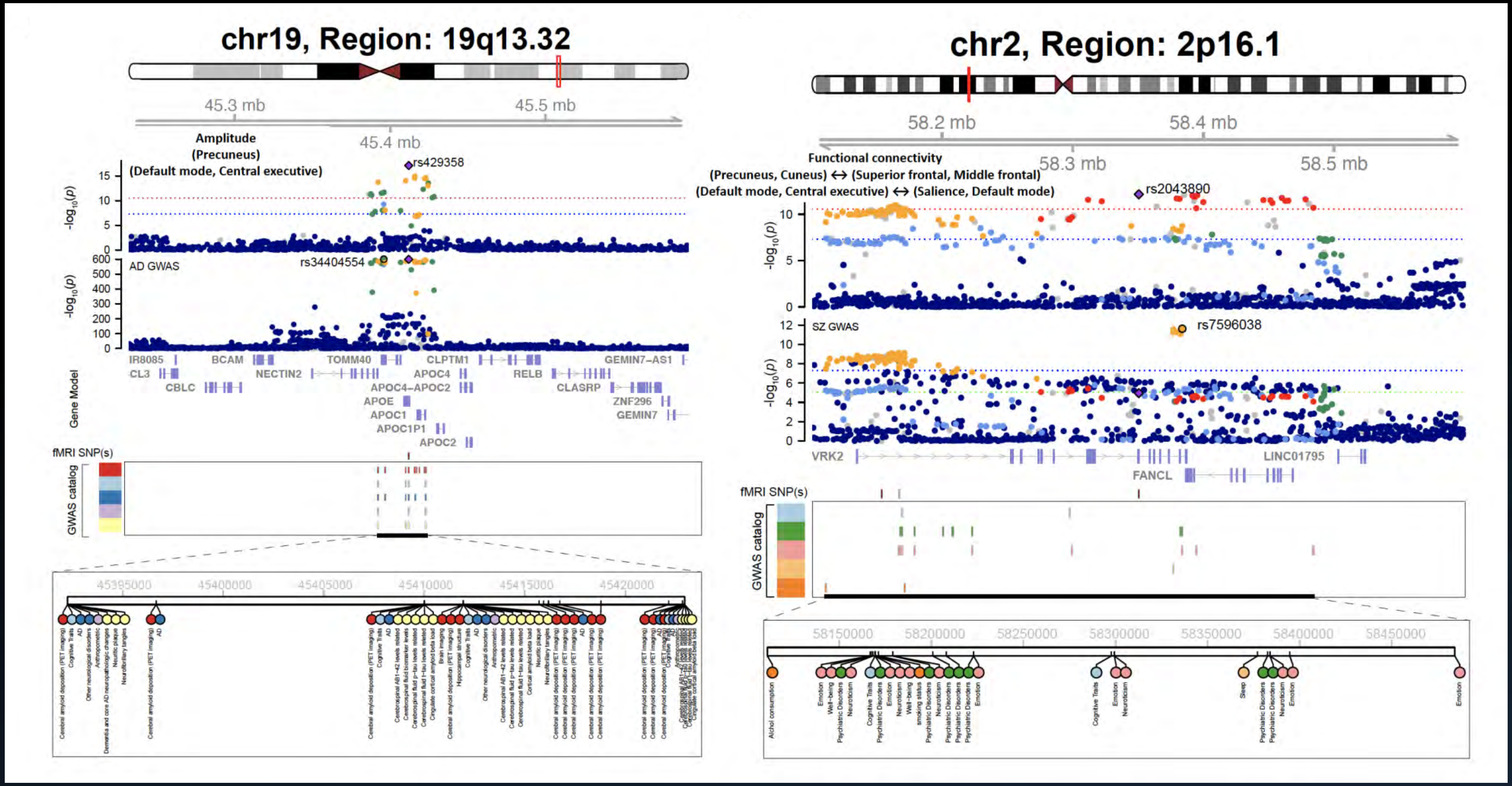
Ideogram of the loci influencing rsfMRI traits of intrinsic brain activity at the significance level 2.8e-11 (5e-8/1777)



Colocalization with AD and SCZ

Colocalization between brain function in the default mode (DM) and central executive (CE) networks with Alzheimer's disease (AD) and Schizophrenia (SCZ)

Alzheimer's disease *(APOE)*



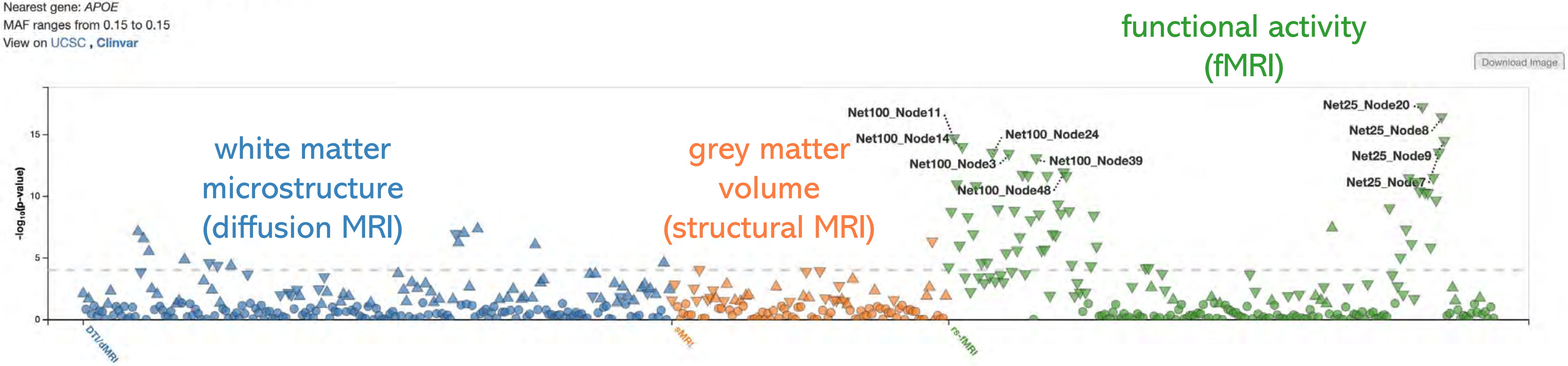
Schizophrenia

Colocalization at *APOE*

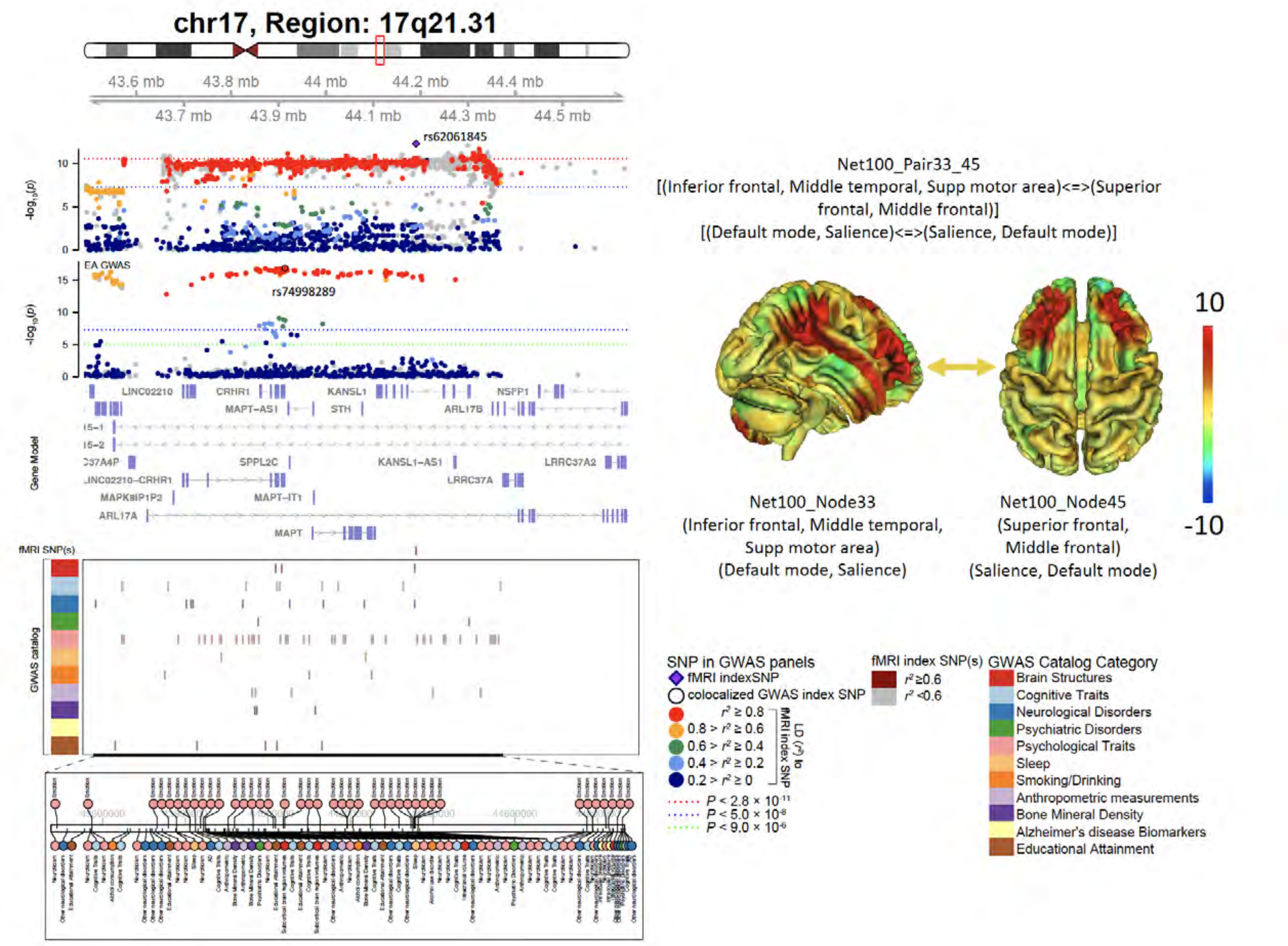
APOE gene has stronger genetic relationships with brain function than brain structures

19 : 45,411,941 C / T

Nearest gene: *APOE*
MAF ranges from 0.15 to 0.15
View on UCSC , Clinvar



Colocalization at 17q21.31 regions



Neurological disorders
(e.g., Parkinson's disease, Alzheimer's disease, corticobasal degeneration)

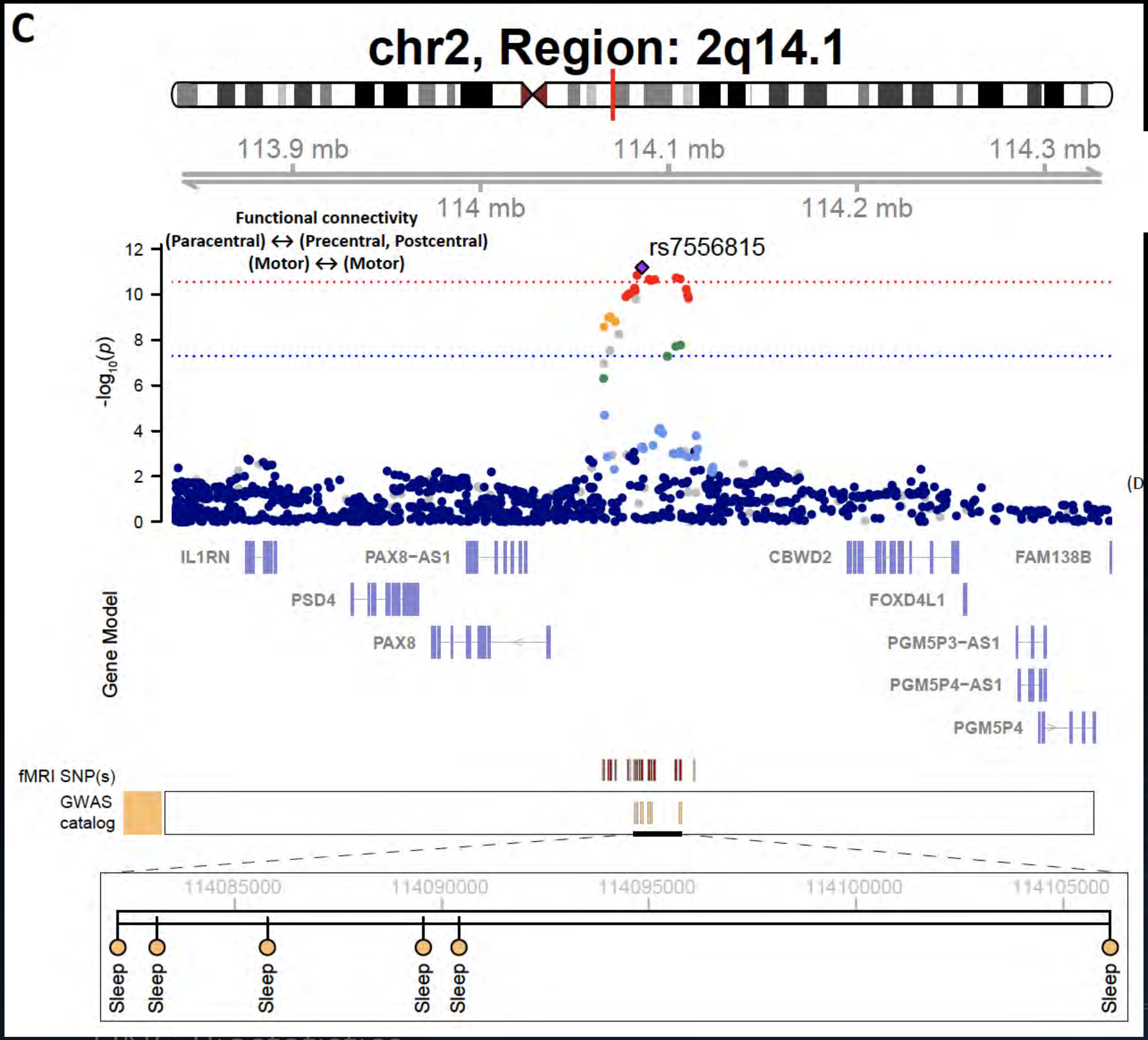
Psychiatric disorders
(e.g., autism spectrum disorder, depression)

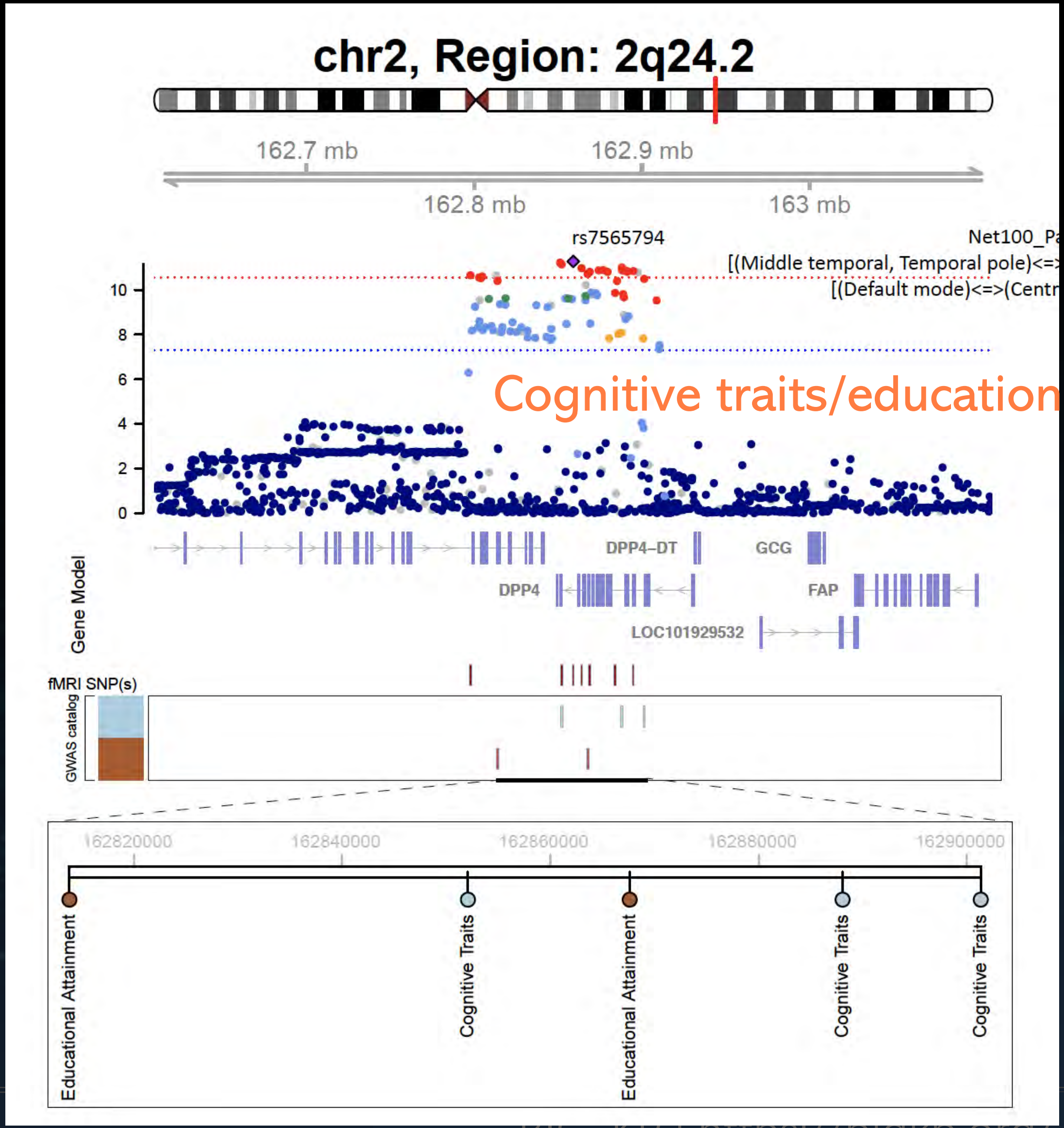
Education, cognitive ability

Psychological traits (e.g., neuroticism)

Alcohol use disorder

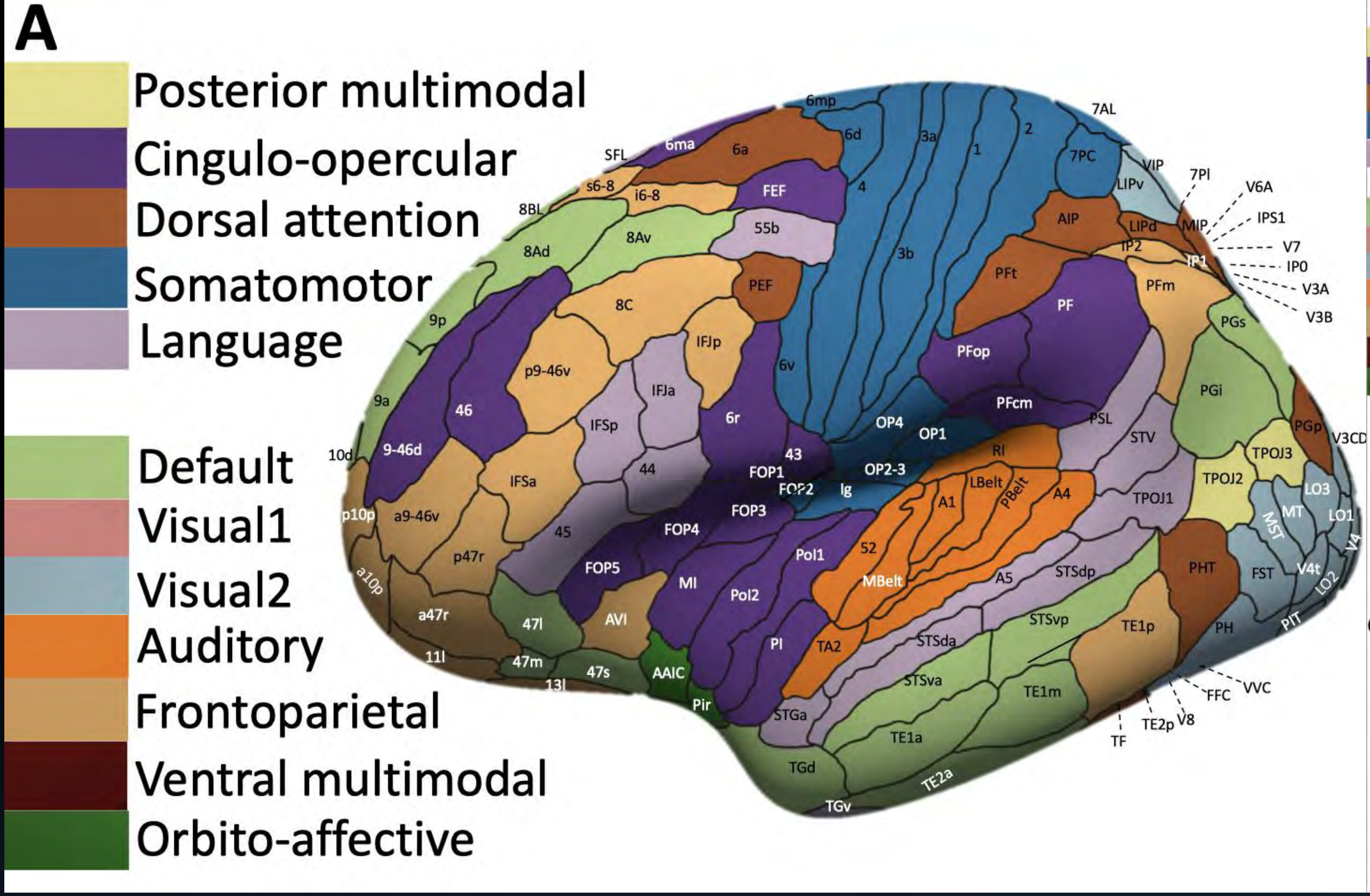
Colocalization with Sleep and Cognition

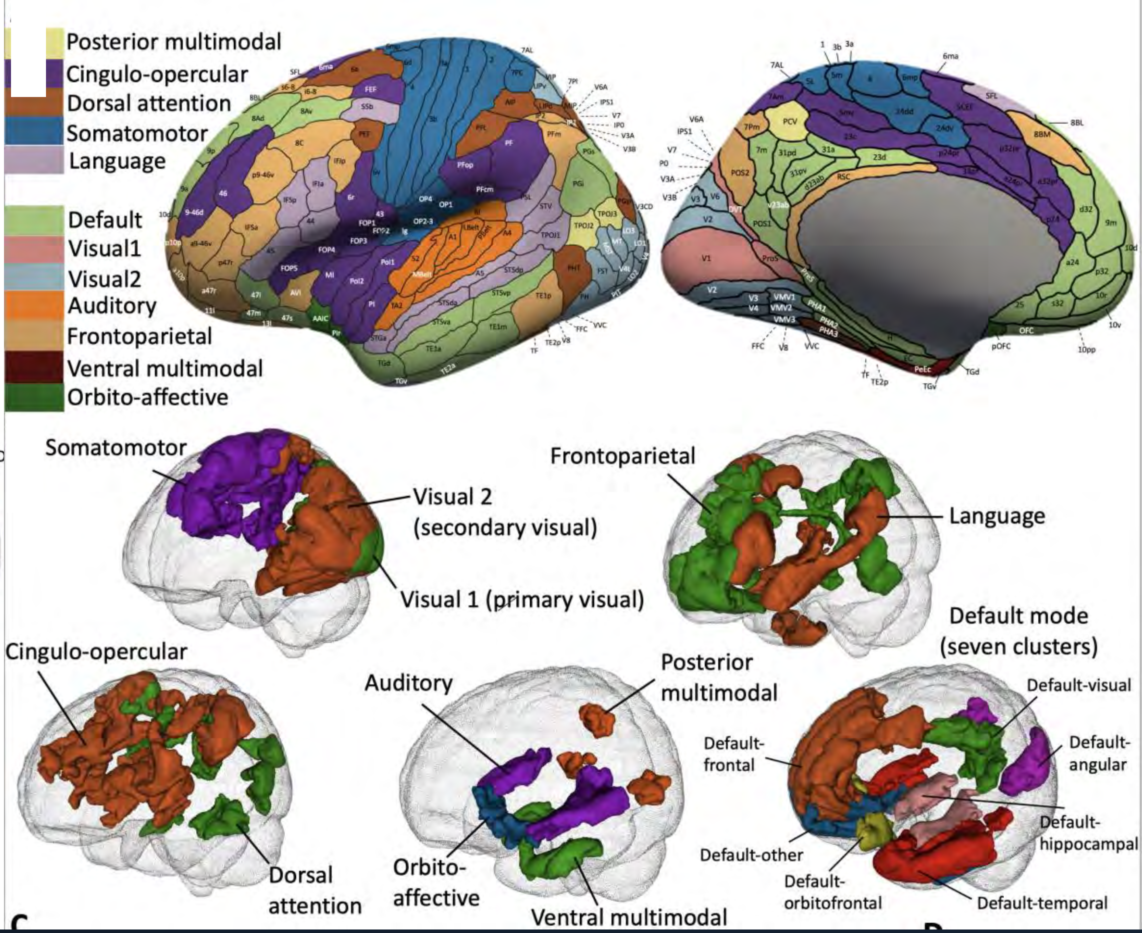




Resting/task functional MRI (fMRI)

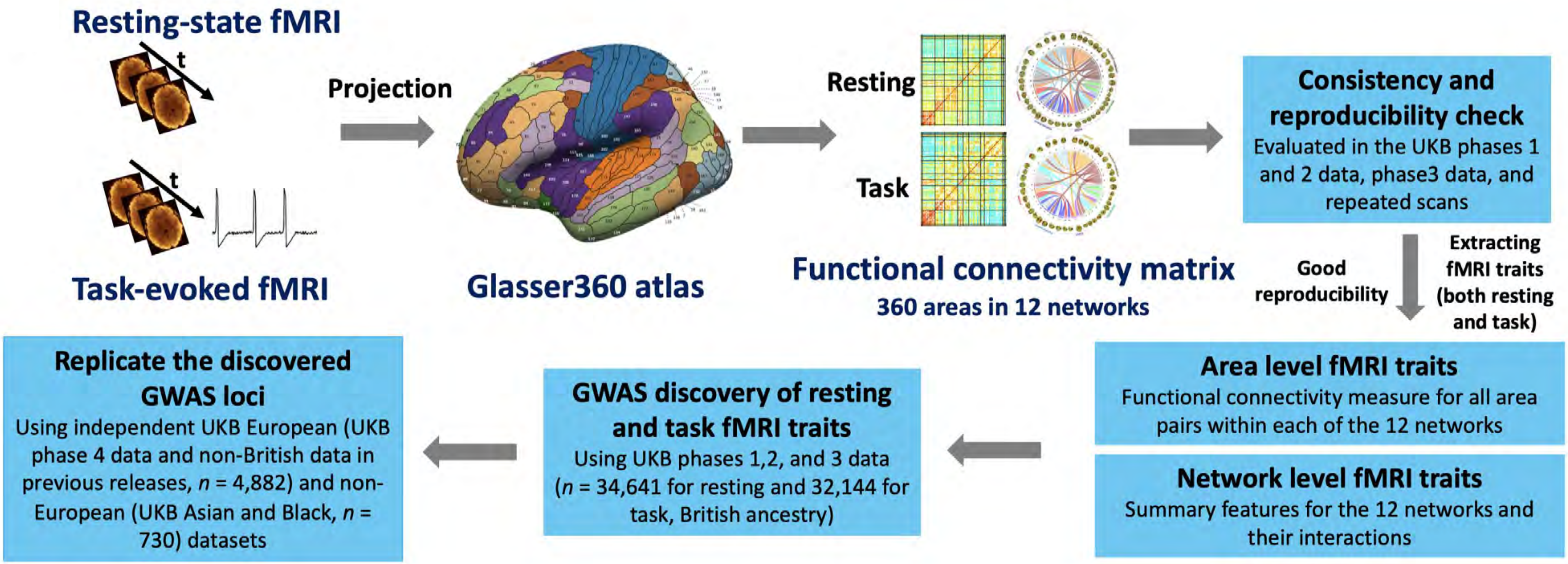
Parcellation-based approach to provide fine-grained details about the cerebral cortex functional organizations





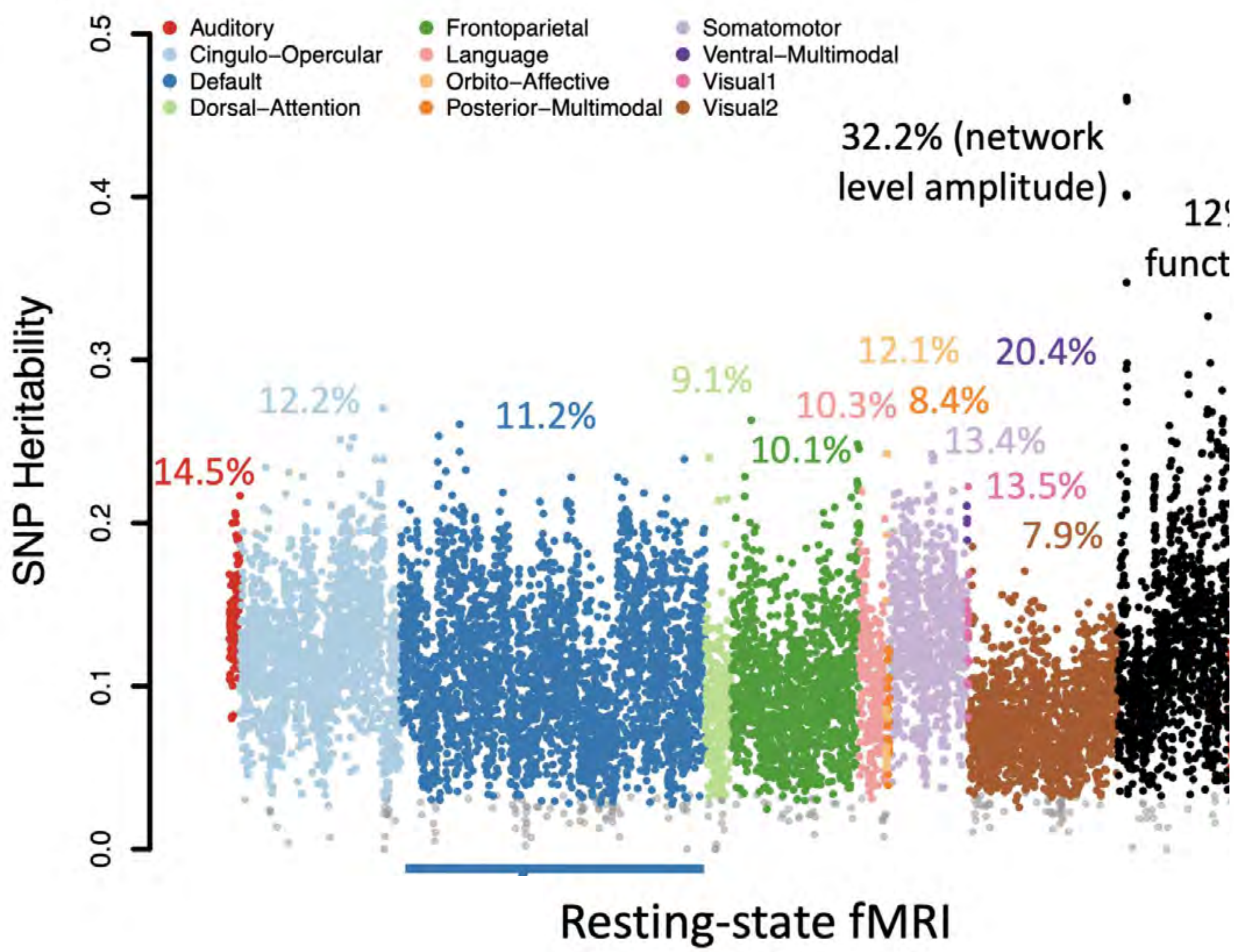
Partitioned the cerebral cortex into 360 well-defined functional areas with known biological functions

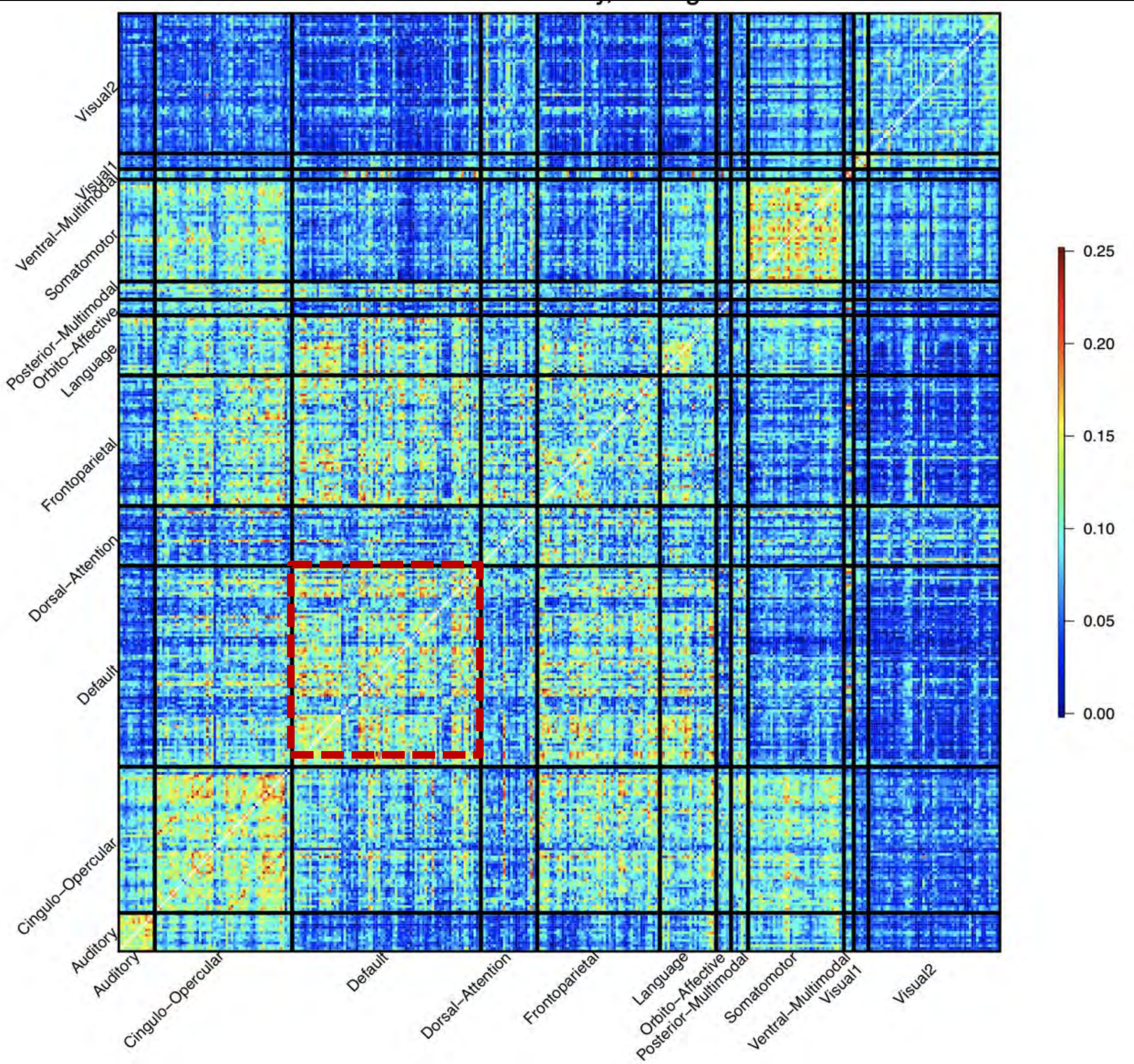
GWAS of Brain Functions



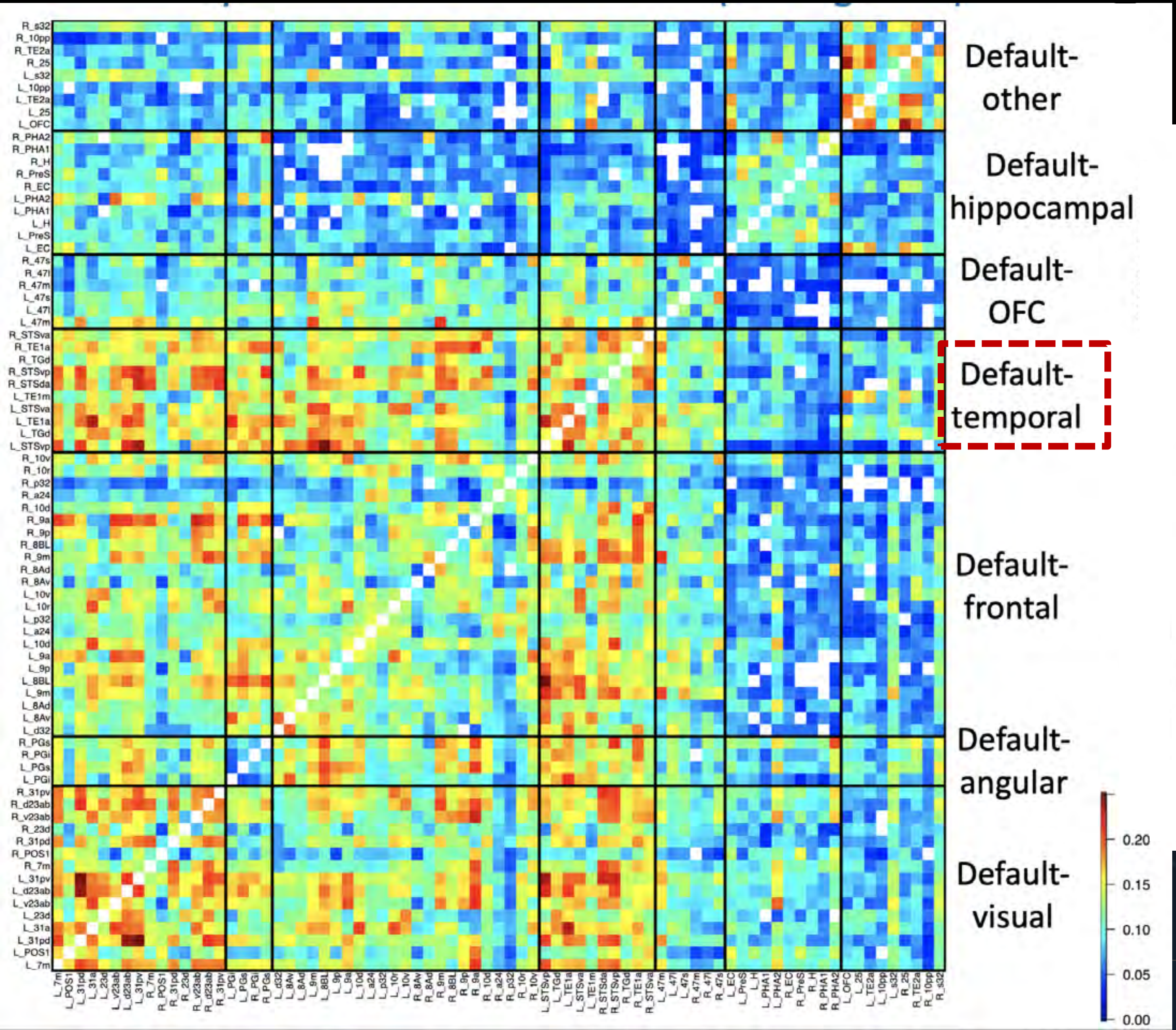
Area-level Heritability Pattern of Functional Brain

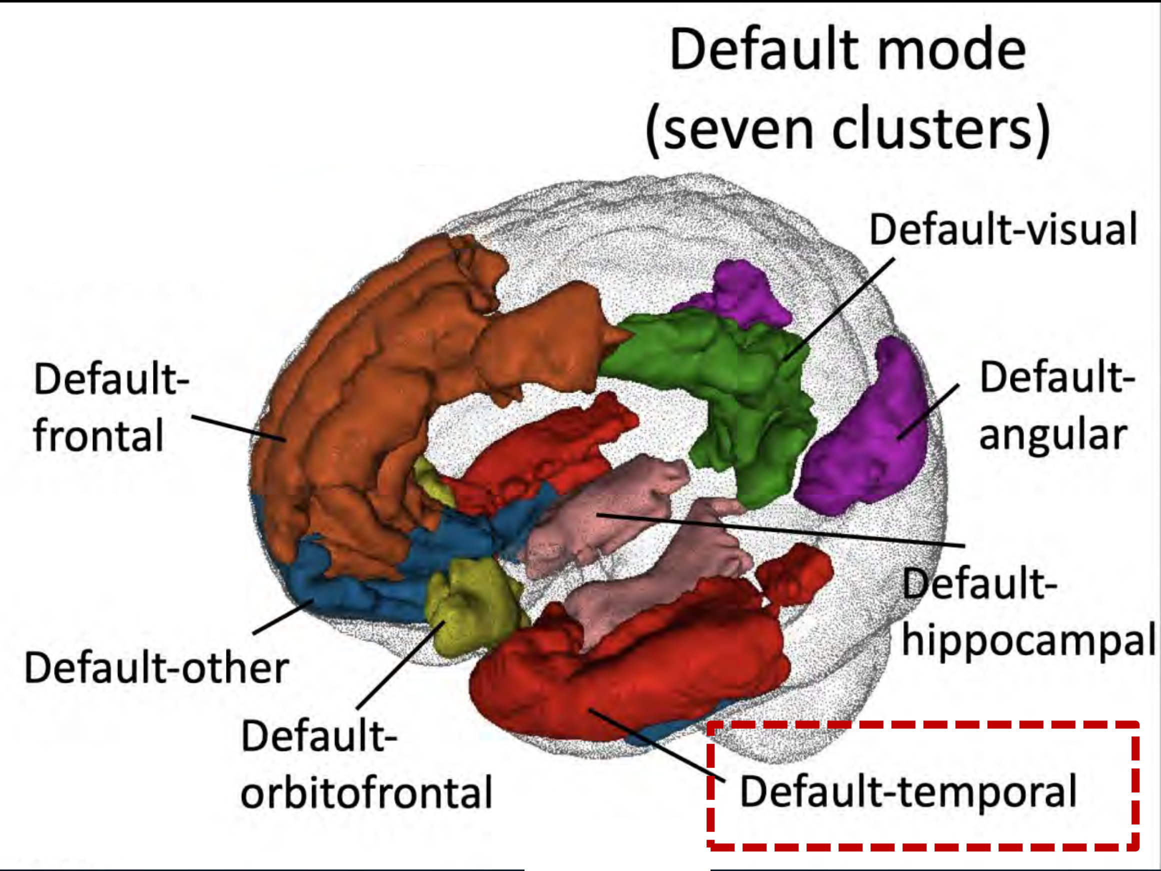
Fine details about the heritability pattern (> 64k fMRI connectivity traits among 360 regions)





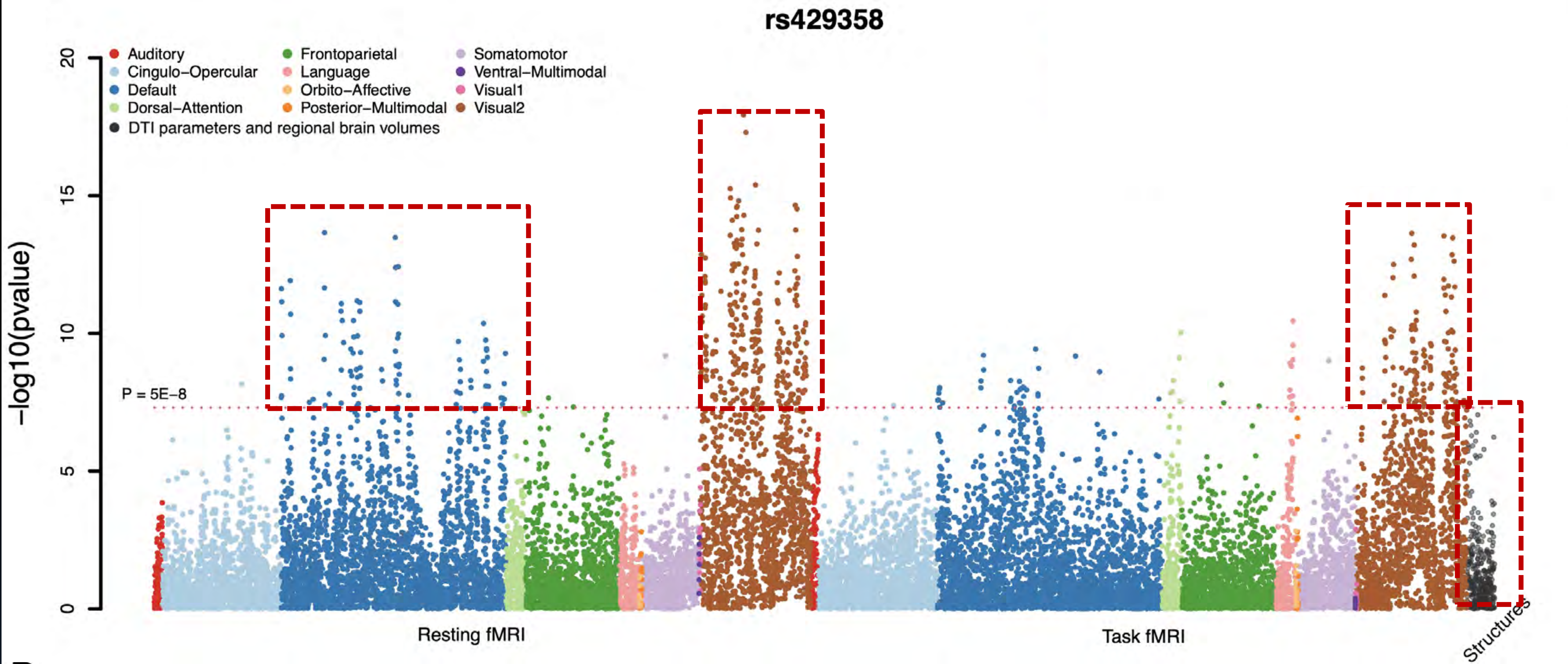
Heritability Pattern in the Default Mode Network





APOE-associations across functional networks

observations: 1) Enriched in the secondary visual and default mode networks;
2) Stronger connections in fMRI than in structural MRI.



It's just a beginning

Hundreds of associated genetic variants for 1593+ neuroimaging traits in three modalities: (grey matter volume, white matter microstructure, resting-state functional connectivity+rfMRI, task fMRI, shape, heart)

Publications (2018+)

Genetic influences on the intrinsic and extrinsic functional organizations of the cerebral cortex (2021). **medRxiv**, 21261187. LINK

Common genetic variation influencing human white matter microstructure (2021). *Science*, 372-6548. LINK

Science AAAS

Transcriptome-wide association analysis of brain structures yields insights into pleiotropy with complex neuropsychiatric traits (2021). *Nature Communications*, 842872. LINK

nature communications

Common variants contribute to intrinsic functional architecture of human brain (2020). *bioRxiv*, 229914. LINK

nature genetics

Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits (2019). *Nature Genetics*, 51(11), 1637-1644. LINK

nature genetics [Cover Feature]

Large-scale GWAS reveals genetic architecture of brain white matter microstructure and genetic overlap with cognitive and mental health traits (n= 17,706) (2019). *Molecular Psychiatry*, in press. LINK

Molecular Psychiatry

Heritability of regional brain volumes in large-scale neuroimaging and genetic studies (2018). *Cerebral Cortex*, 29(7), 2904-2914. LINK

Cerebral CORTEX

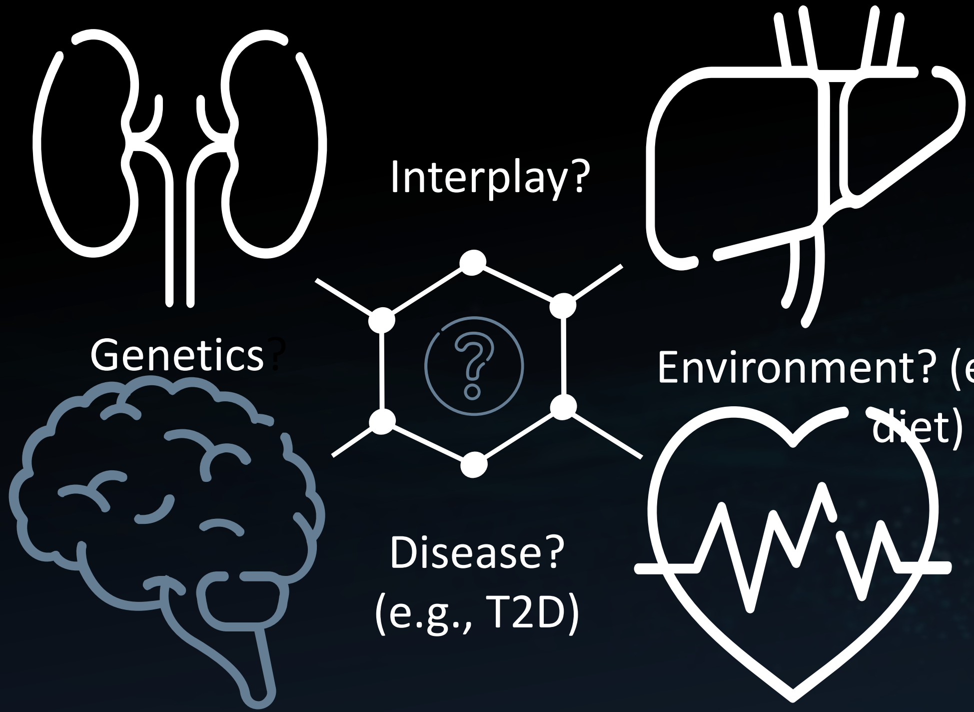
Genetics discovery in human brain by big data integration

Human Body Imaging Genetics Network (1-5 yrs)

Multi-organ images (abdominal, brain, and heart)

Kidney

Liver MRI



Other images (eye, face, knee, hip, etc.)

Retinal images

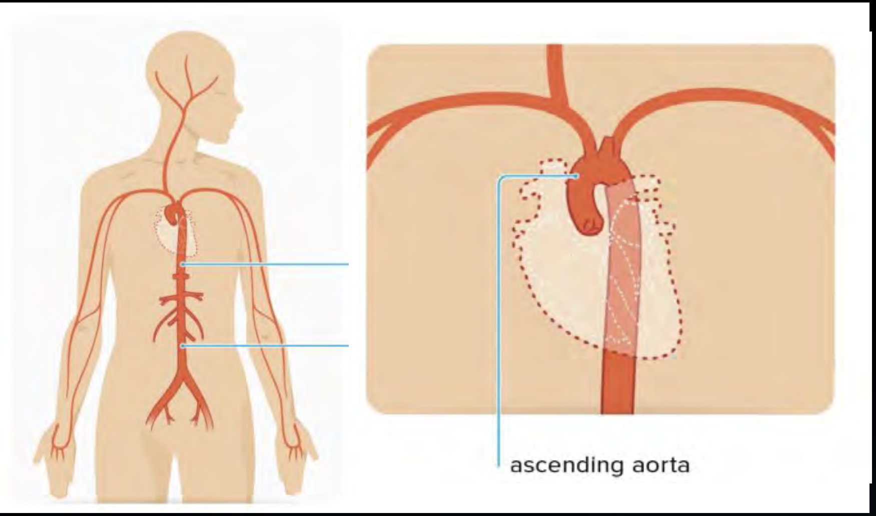
DXA bone images

Data from > 100k subjects (by 2024)

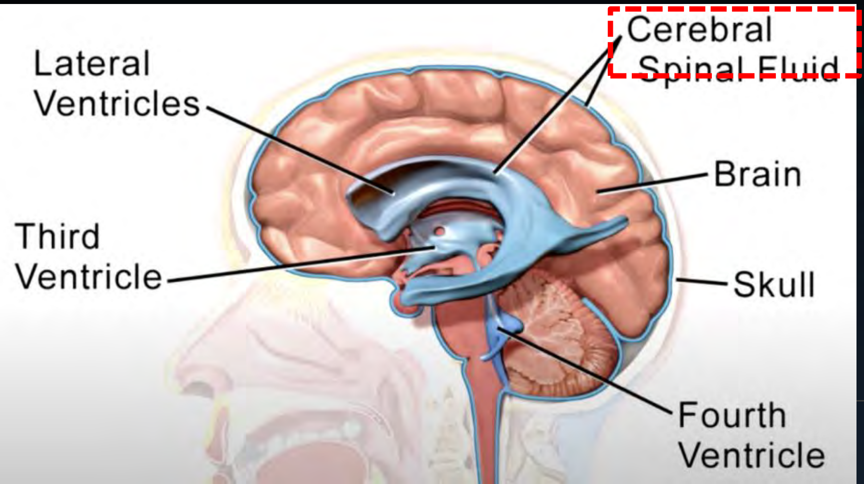
Heart-Brain Connections

Zhao et al. *medRxiv* , 2021, multi-organ images from 40k subjects

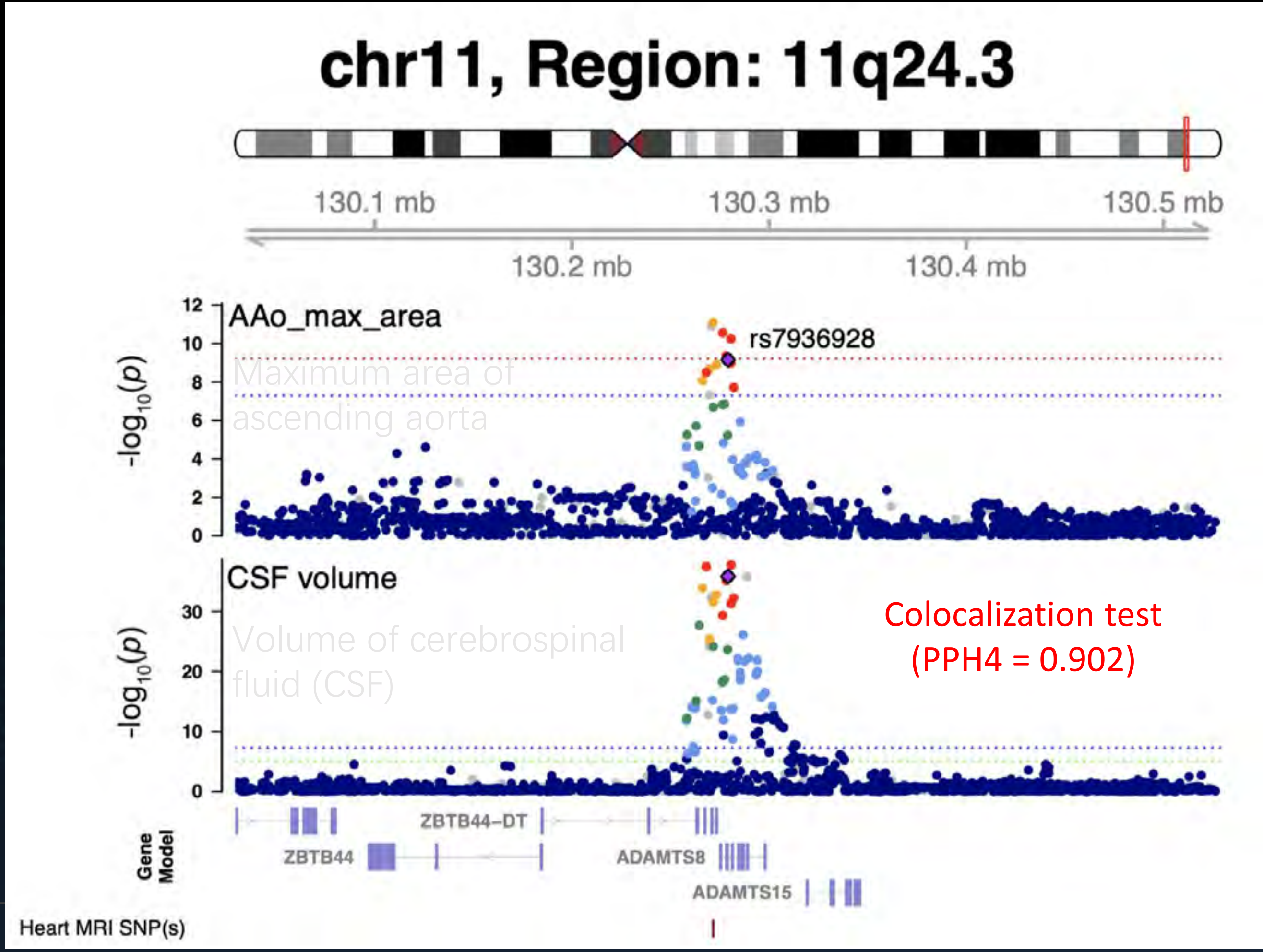
Shared genetic influence between heart and brain structures



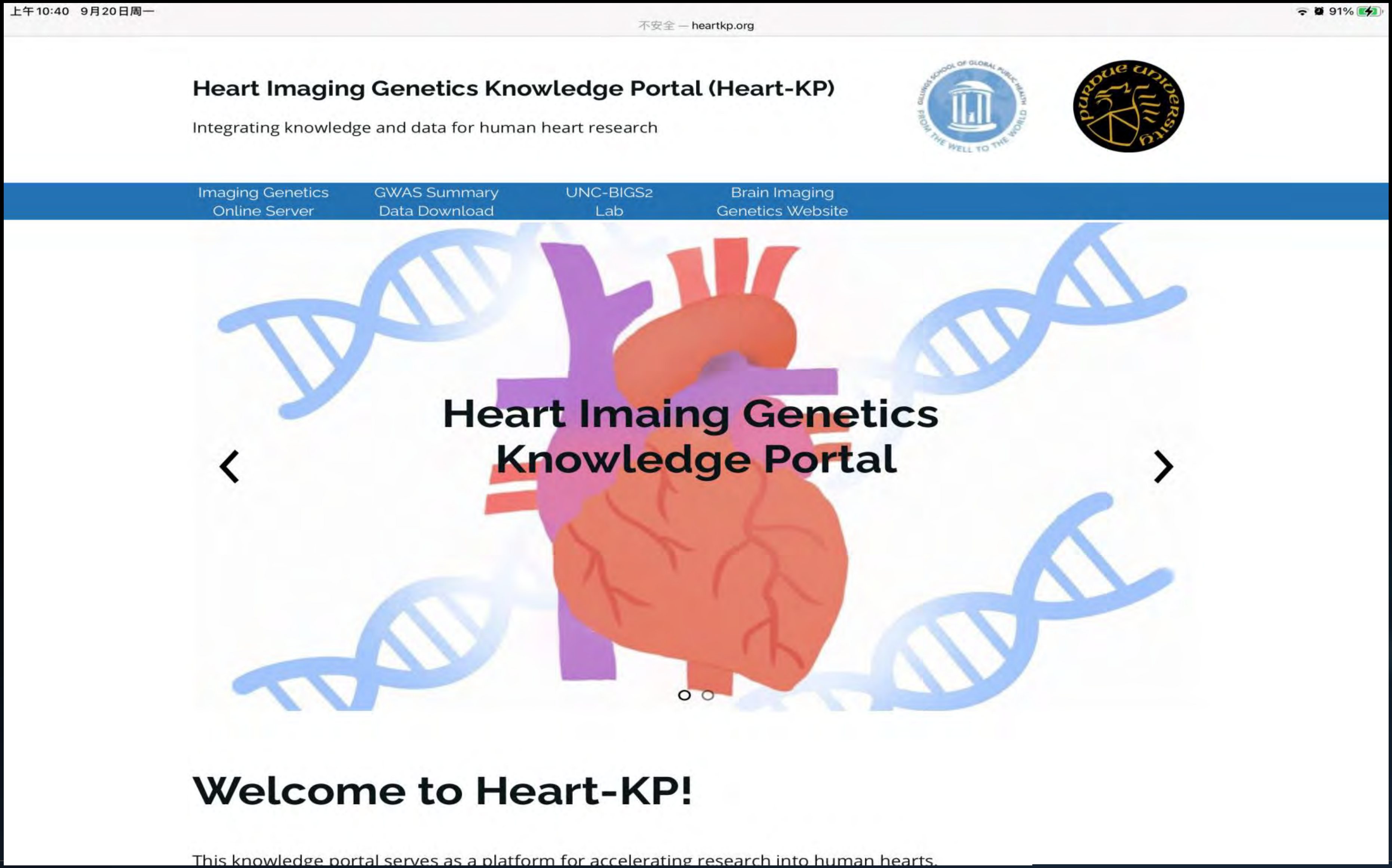
Heart MRI



Brain MRI



Heart Knowledge Portal



Ongoing/Future Directions

- Causal relationships among disease, brain structures, and brain functionalities (e.g., the genetic pathway among vascular risk factors, white matter, and stroke)
- Build optimal models for complex traits and diseases prediction using imaging and genetics data (e.g., deep learning)
- Compare and identify the best practical strategy and pipelines to process different neuroimaging modalities (e.g., ICA for fMRI)
- Model brain changes and genetics effects across the life span
- Align and integrate different neuroimaging modalities

Acknowledgement

UNC GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH

PURDUE UNIVERSITY | Department of Statistics

Brain Imaging Genetics Knowledge Portal (BIG-KP)

Genetics Discoveries in Human Brain by Big Data Integration

bigkp.org

Funding: U.S. NIH Grants MH086633 and MH116527

Pictures: Copyrights belong to their own authors and/or holders.

Data: We thank Bingxin Zhao, Tengfei Li and other members of the UNC BIG-S2 lab (https://med.unc.edu/bigs2/) for processing the neuroimaging data.

UK Biobank resource application number: 22783.