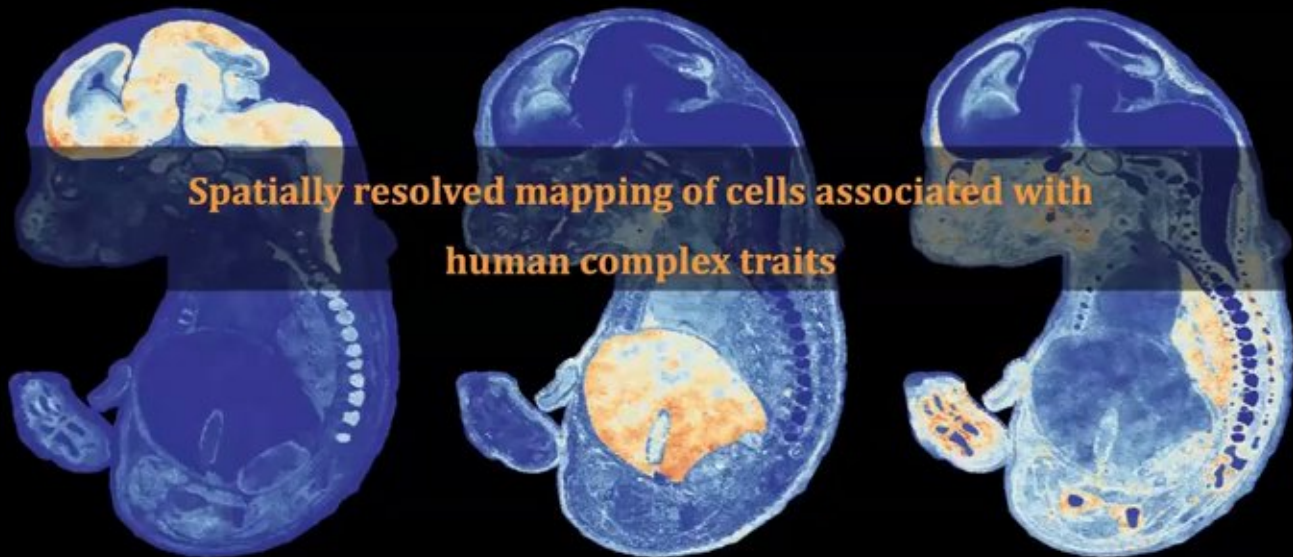




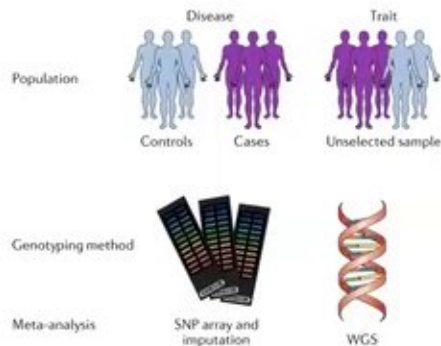
**Spatially resolved mapping of cells associated with
human complex traits**



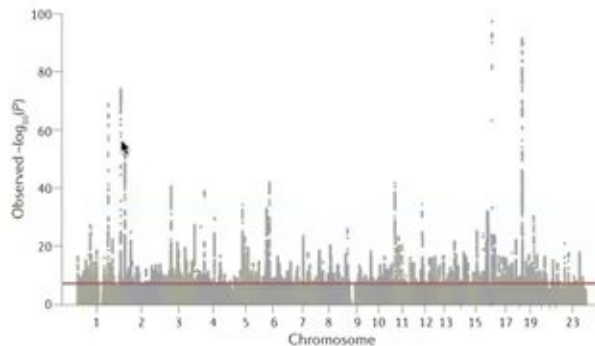
CGM, 16 May 2025

Liyang Song

Genome-wide association studies (GWAS)



Nat Rev Genet, 20(8):467-484 (2019)

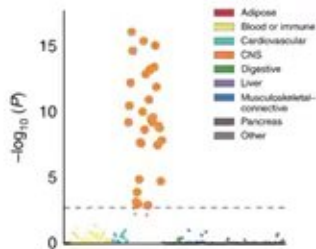


Nat Rev Methods Primers 1, 59 (2021)



Finding the relevant biological contexts for complex traits

Tissue/cell type enrichment
(LDSC-SEG, scLinker)



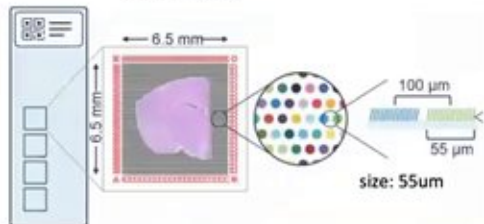
Nature Genetics, 50: 621-9 (2018)

Nature Genetics, 54: 1479-92 (2022)

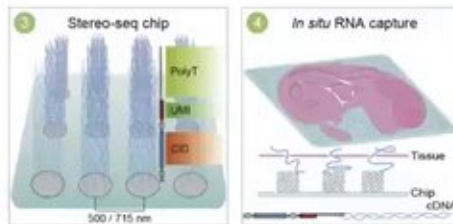


Spatial transcriptomics (ST)

10X Visium

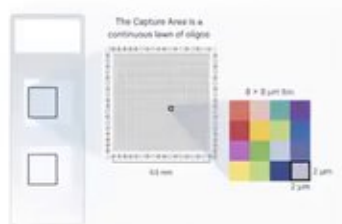


BGI Stereo-seq



size: 2um

10X Visium HD

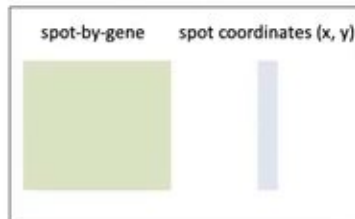


size: 2um

spot:

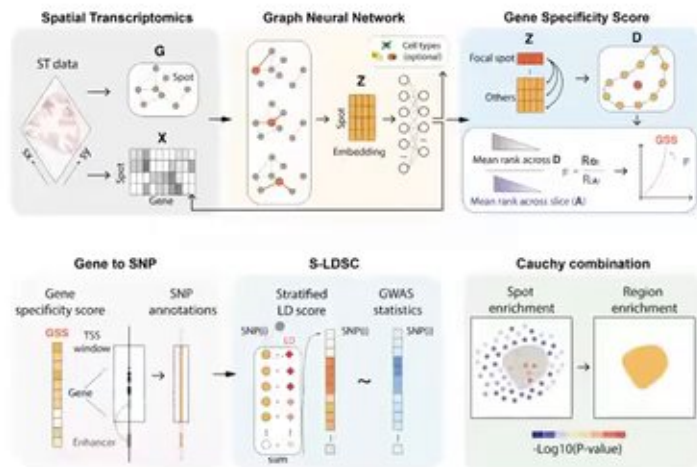
- cell-bin: single-cell level
- bin50: 50 X 50 bins
- bin100: 100 X 100 bins
- 10X Visium spot

h5ad file





gsMap: human genetics + spatial transcriptomics



Article

Spatially resolved mapping of cells associated with human complex traits

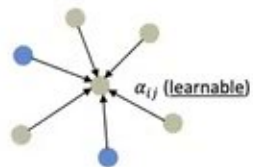
<https://doi.org/10.1038/s41586-025-08757-x> | Liang Song^{1,2}, Wenhan Chen^{1,2}, Xuan-Hua^{1,2}, Miao-Duo^{1,2} & Jin Yang^{1,2}
Received 17 May 2024

- Identifying marker genes for a spot in spatial transcriptomics (ST) data
- Assessing the enrichment of trait heritability in genetic variants near these genes

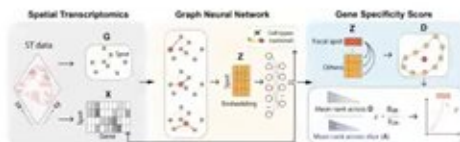
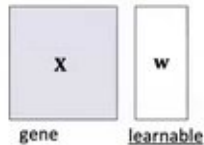
Spot-level gene specificity score

Graph attention autoencoder (GAT)

Spatial graphs



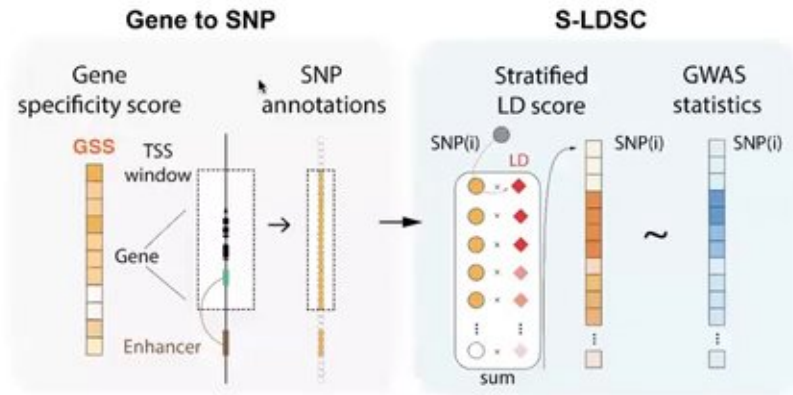
Expression matrix



leon song



Spot-level genomics annotation and associations



Associating mouse embryonic cells with human traits by gsMap

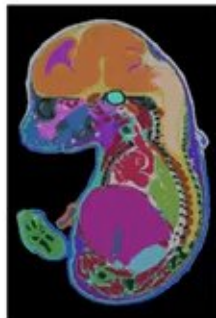
Mouse E16.5 embryo
(bin50 resolution)



Cell, 185: 1777-92 (2022)

Associating mouse embryonic cells with human traits by gsMap

Mouse E16.5 embryo
(bin50 resolution)

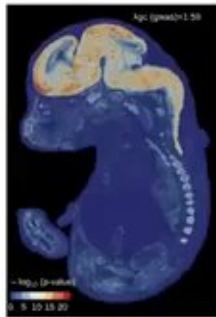


Cell, 185: 1777-92 (2022)

Trait:

IQ

Nature Genetics, 50:
912-9 (2018)

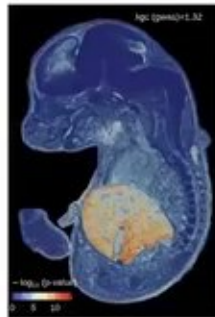


Tissue:

Brain & spinal cord

Haemoglobin
concentration

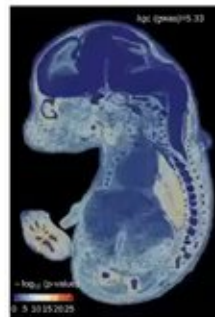
Cell, 182: 1214-31 (2020)



Liver

Height

Nature, 610: 704-12
(2022)

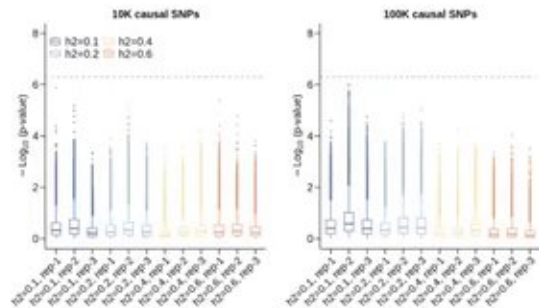


Cartilage & its
primordium

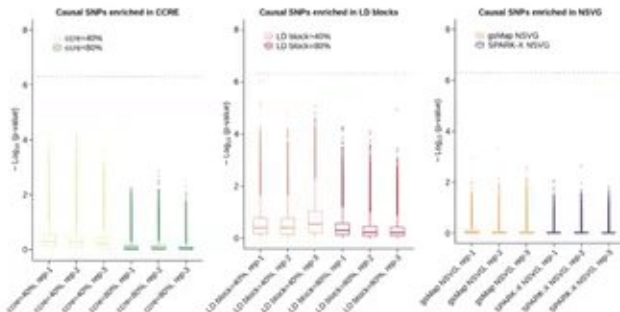
Negative controls: simulations under the null

Simulated traits based on UK Biobank genotype data ($n = 100,000$) + mouse embryonic spatial transcriptomics data

causal variants were randomly distributed



causal variants were concentrated in specific genomic regions



Male pattern baldness

Mouse E16.5 embryo
(single-cell resolution)



Cell, 185: 1777-92 (2022)

Epithelial cells



Male pattern baldness



Nature Communications,
9: 5407 (2018)

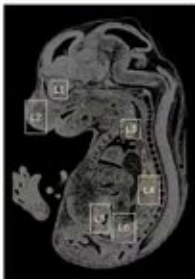
Male pattern baldness

Mouse E16.5 embryo
(single-cell resolution)



Cell, 185: 1777-92 (2022)

Epithelial cells

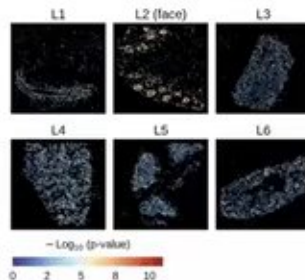


Male pattern baldness



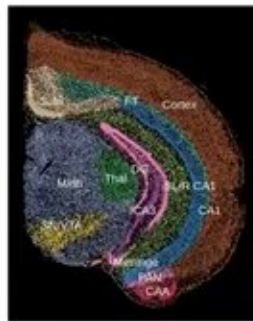
Nature Communications,
9: 5407 (2018)

Hair follicles like cell assemblies in L2 (face)

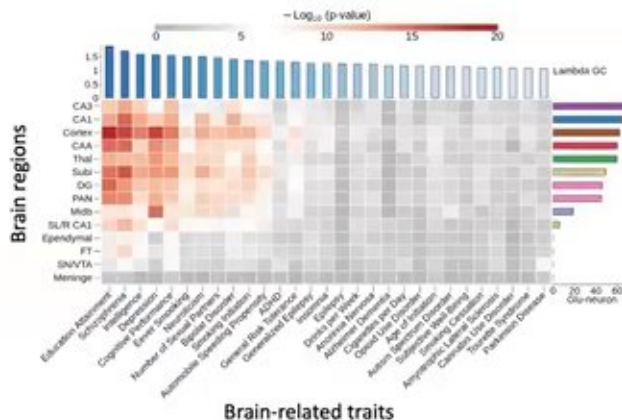


Associating cells in adult mouse brain with 29 human complex traits

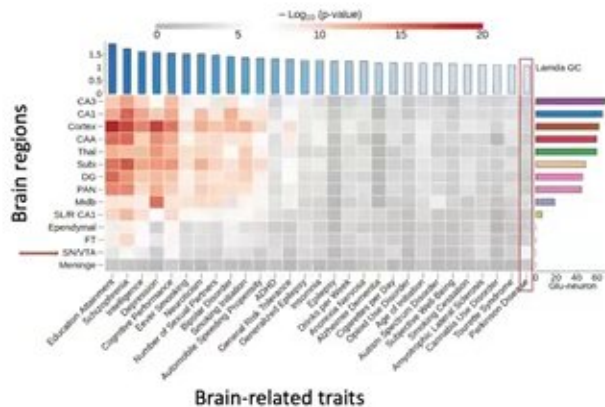
Adult mouse brain
(single-cell resolution)



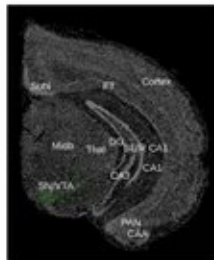
Cell, 185: 1777-92 (2022)



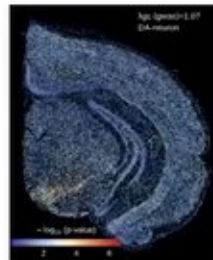
Most relevant brain region for Parkinson disease



Dopaminergic neurons



Parkinson disease



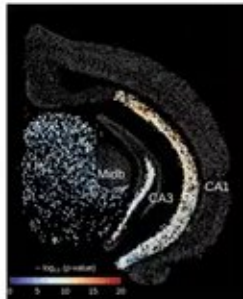
SN/VTA: substantia nigra/ventral tegmental area

Comparison among schizophrenia, depression, and IQ

Odds ratio (OR): strength of association with a brain region relative to the whole brain

Schizophrenia

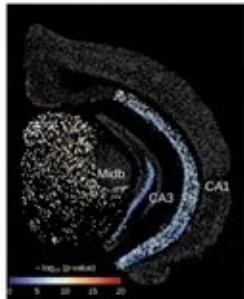
Nature, 604: 502-8 (2022)



Hippocampus CA1 + CA3: OR = 4.47
Midbrain: OR = 0.24

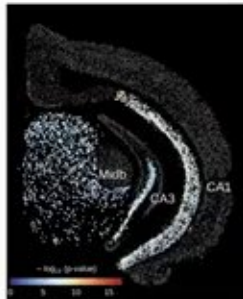
Depression

Nature Medicine, 29: 1832-44 (2023)



OR = 0.33
OR = 6.72

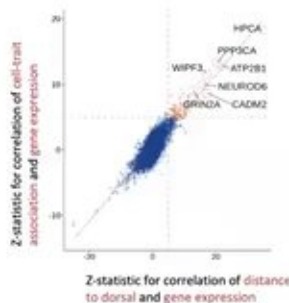
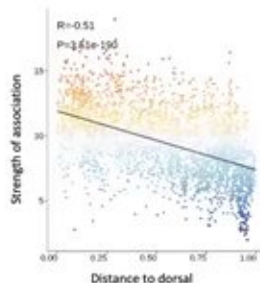
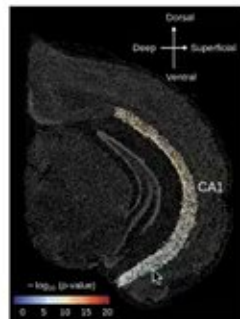
IQ



OR = 1.54
OR = 0.62

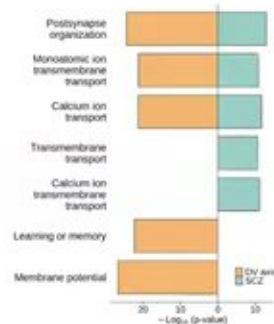


Spatially dependent associations of neurons with schizophrenia



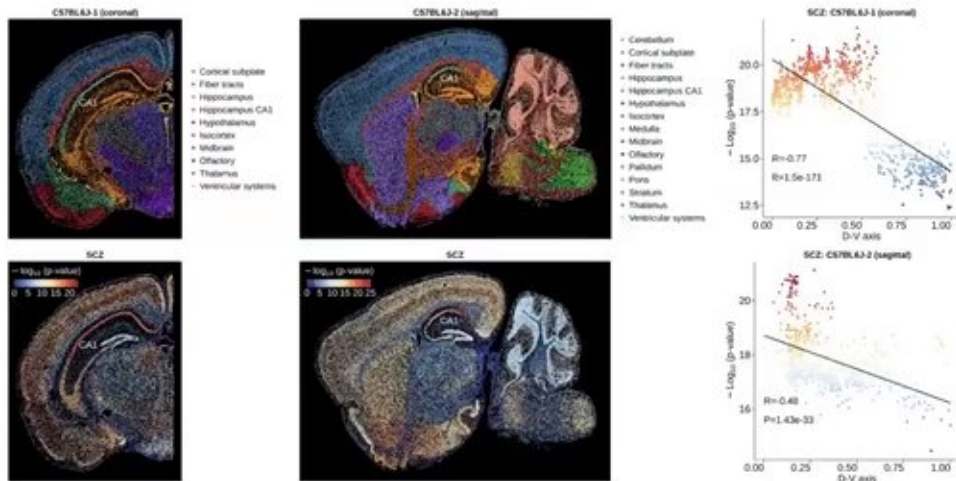
Top three genes: *HPCA*, *PPP3CA*, and *ATP2B1*, all related to calcium signaling and regulation

GO enrichment

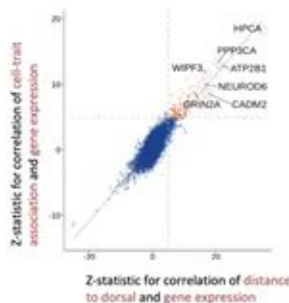
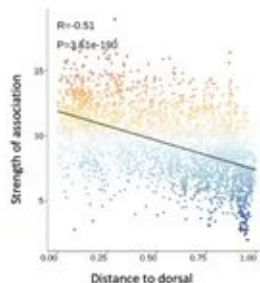
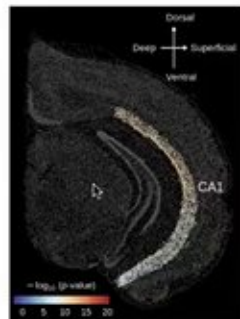


Spatially dependent associations of neurons with schizophrenia

Multiplexed error-robust fluorescence in situ hybridization (MERFISH) + scRNA-seq *Nature* 624: 343–354 (2023)

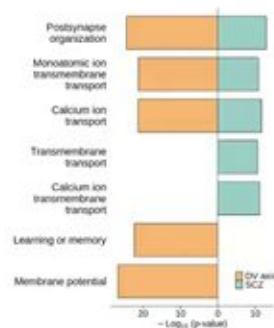


Spatially dependent associations of neurons with schizophrenia



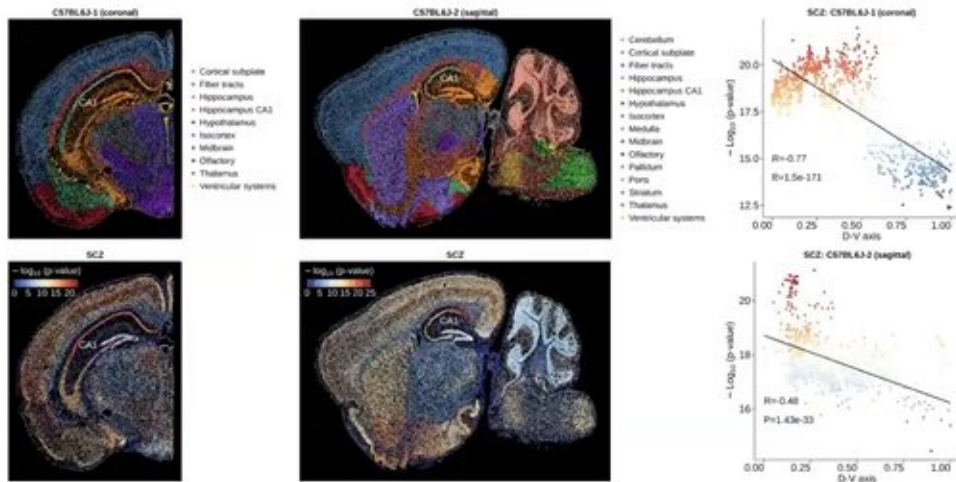
Top three genes: *HPCA*, *PPP3CA*, and *ATP2B1*, all related to calcium signaling and regulation

GO enrichment

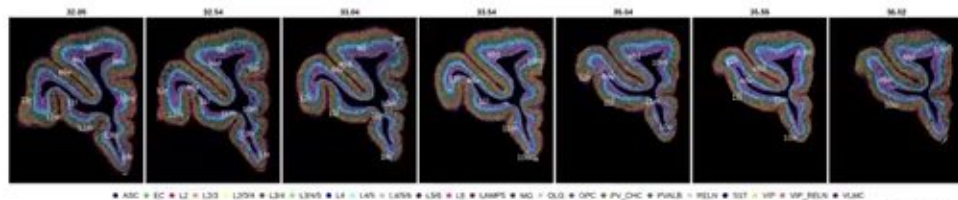


Spatially dependent associations of neurons with schizophrenia

Multiplexed error-robust fluorescence in situ hybridization (MERFISH) + scRNA-seq *Nature* 624: 343–354 (2023)



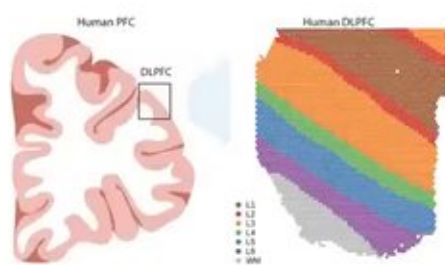
Applying gsMap to macaque cortex spatial transcriptomics data



Cell, 186: 3726-3743 (2023)



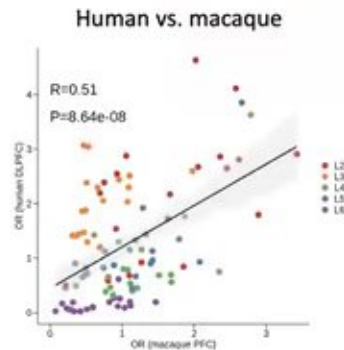
Adult human brain DLPFC spatial transcriptomics data



Nature Neuroscience, 24: 425-36 (2021)

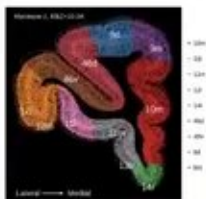


Cell, 186: 3726-43 (2023)



Associating cells in macaque prefrontal cortex with human traits

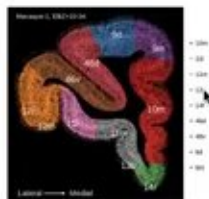
Macaque prefrontal cortex (PFC)
(single-cell resolution)



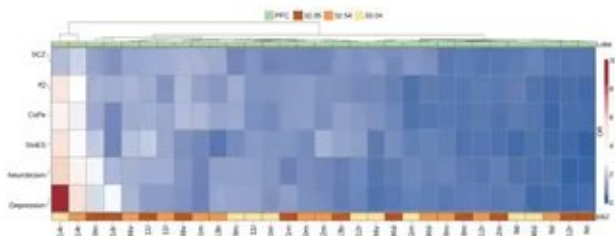


Associating cells in macaque prefrontal cortex with human traits

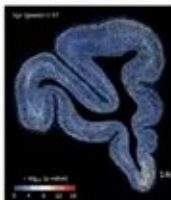
Macaque prefrontal cortex (PFC)
(single-cell resolution)



Associations of macaque PFC regions with human traits



Depression & 14r



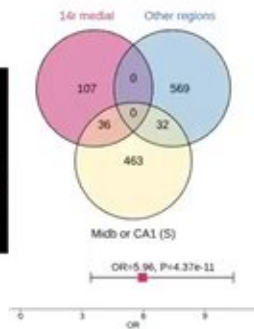
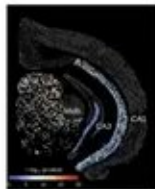
Gyrus rectus



Macaque 14r medial & depression

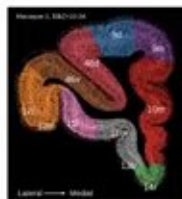
Highly expressed genes

Mouse ST data
& depression

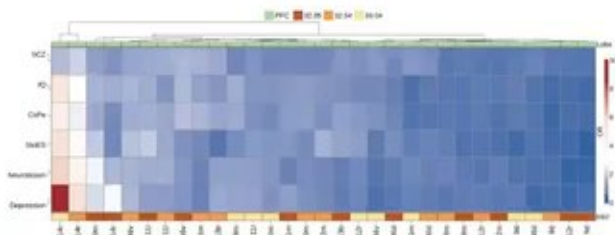


Associating cells in macaque prefrontal cortex with human traits

Macaque prefrontal cortex (PFC)
(single-cell resolution)



Associations of macaque PFC regions with human traits

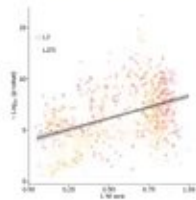


Depression & 14r

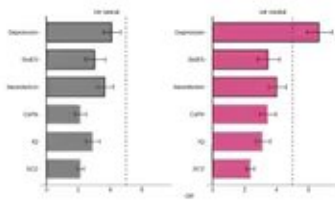


Gyrus rectus

Lateral to medial axis

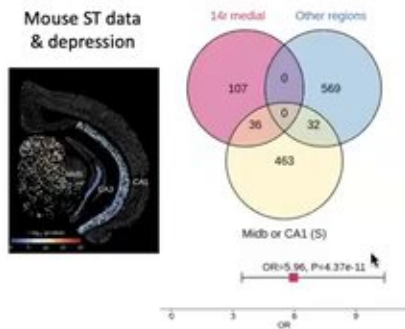


14r medial



Macaque 14r medial & depression

Highly expressed genes





Macaque 14r medial & psychiatric drugs

Drug targets enrichment of genes
highly expressed in 14r medial



Cell-relevance correlation (CRC)

SCZ & IQ: Cortex, Hippocampus CA

Depression: Midbrain, PFC 14r

The neurons associated with SCZ more resembles that for IQ than that for depression.

Tips for using gsMap

Mouse Embryo (Quick Mode)

The `quick_mode` option provides a simplified and efficient way to execute the entire `gsMap` pipeline. It minimizes running time and configuration complexity by utilizing pre-calculated weights based on the 1000G EUR reference panel and protein-coding genes from gtf file of Gencode v46. This mode is ideal for users who prefer a streamlined approach. For a more customizable experience, such as using custom GTF files, reference panels, and adjustable parameters, please refer to the [Step by Step](#) guide.

```
gsmap quick_mode \  
--workDir './example_quick_mode/Mouse_Embryo' \  
--homolog_file 'gsMap_resource/homologs/mouse_human_homologs.txt' \  
--sample_name 'E16.5_E151.MOSTA' \  
--gsMap_resource_dir 'gsMap_resource' \  
--hdfs_path 'gsMap_example_data/ST/E16.5_E151.MOSTA.h5ad' \  
--annotation 'annotation' \  
--data_layer 'count' \  
--sumstats_file 'gsMap_example_data/GMS/IO_N6_2018.sumstats.gz' \  
--trait_name 'IO'
```

```
gsMap_resource  
├── genome_annotation  
│   ├── enhancer  
│   └── gtf  
├── homologs  
│   ├── macaque_human_homologs.txt  
│   └── mouse_human_homologs.txt  
├── LD_Reference_Panel  
│   └── 1000G_EUR_Phase3_gLink  
├── LDSC_resource  
│   ├── hapmap3_snps  
│   └── weights_hm3_no_hla  
├── quick_mode  
│   ├── baseline  
│   ├── SNP_gene_pair  
│   └── snp_gene_weight_matrix.h5ad
```

Mouse Embryo (Step by Step)

This tutorial guides you through running `gsMap` step by step, with user-defined parameters and resources, granting greater flexibility and control over the analysis. This mode is suited for users who require detailed customization of their pipeline. For a faster, one-command execution, please see the [Quick Mode](#) tutorial.

Documents:

<https://yanglab.westlake.edu.cn/gsmmap/document/software>

Results report:

[web report](#)

leon song



E16.5_E1S1.MOSTA Genetic Spatial Mapping Report

Genetic Spatial Mapping Plot

This plot shows the spatial genetic mapping results across different tissues.

IQ (gsMap)



annotation

- Cavity
- Epidermis
- Connective tissue
- Muscle
- Adipose tissue
- Cartilage primordium
- Submandibular gland
- Jaw and tooth
- Bone
- Cartilage
- Lung
- Kidney
- Meninges
- GI tract
- Liver
- Inner ear
- Adrenal gland
- Dorsal root ganglion
- Mucosal epithelium
- Smooth muscle
- Heart
- Sympathetic nerve
- Spinal cord
- Brain
- Choroid plexus

$-\log_{10}(p)$

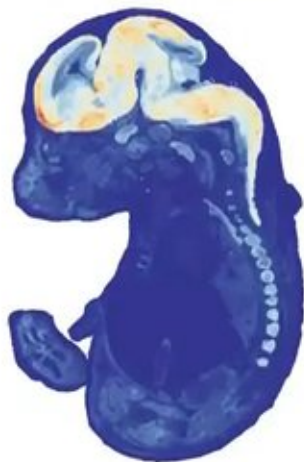
leon song

E16.5_E1S1.MOSTA Genetic Spatial Mapping Report

Genetic Spatial Mapping Plot

This plot shows the spatial genetic mapping results across different tissues.

IQ (gsMap)



annotation

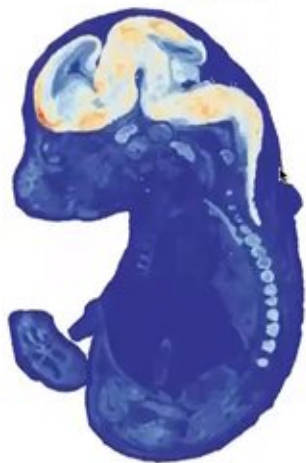
- Cavity
- Epidermis
- Connective tissue
- Muscle
- Adipose tissue
- Cartilage primordium
- Submandibular gland
- Jaw and tooth
- Bone
- Cartilage
- Lung
- Kidney
- Meninges
- GI tract
- Liver
- Inner ear
- Adrenal gland
- Dorsal root ganglion
- Mucosal epithelium
- Smooth muscle
- Heart
- Sympathetic nerve
- Spinal cord
- Brain
- Choroid plexus

$-\log_{10}(p)$

Genetic Spatial Mapping Plot

This plot shows the spatial genetic mapping results across different tissues.

IQ (gsMap) 🔍 + □ ○ ▢ ▣ ▤ ▥ ▦ ▧ ▨



annotation

- Cavity
- Epidermis
- Connective tissue
- Muscle
- Adipose tissue
- Cartilage primordium
- Submandibular gland
- Jaw and tooth
- Bone
- Cartilage
- Lung
- Kidney
- Meninges
- GI tract
- Liver
- Inner ear
- Adrenal gland
- Dorsal root ganglion
- Mucosal epithelium
- Smooth muscle
- Heart
- Sympathetic nerve
- Spinal cord
- Brain
- Choroid plexus

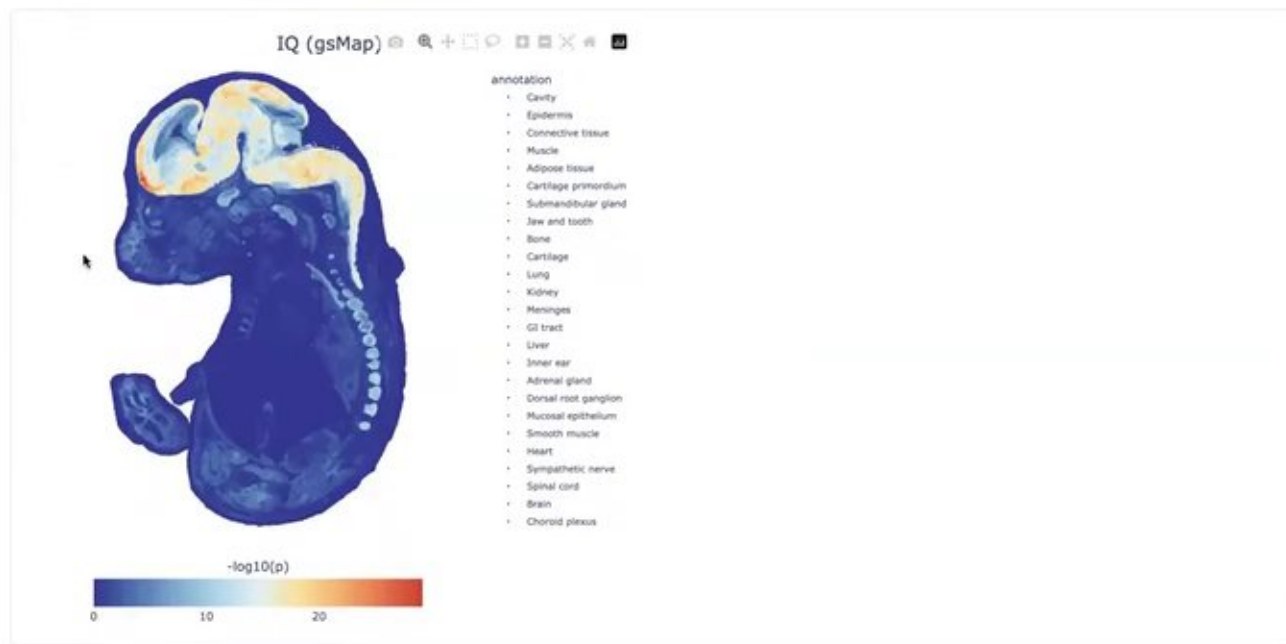
$-\log_{10}(p)$

0 10 20

leon song



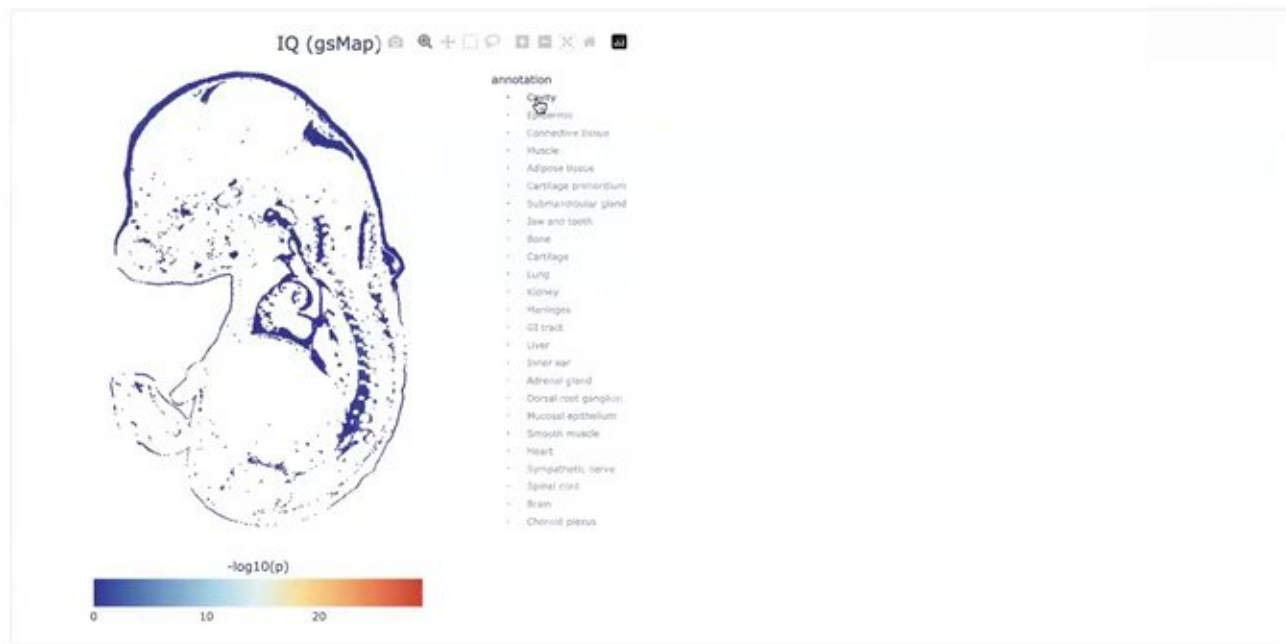
This plot shows the spatial gene expression results across different tissues.



leon song

Cauchy Combination Result

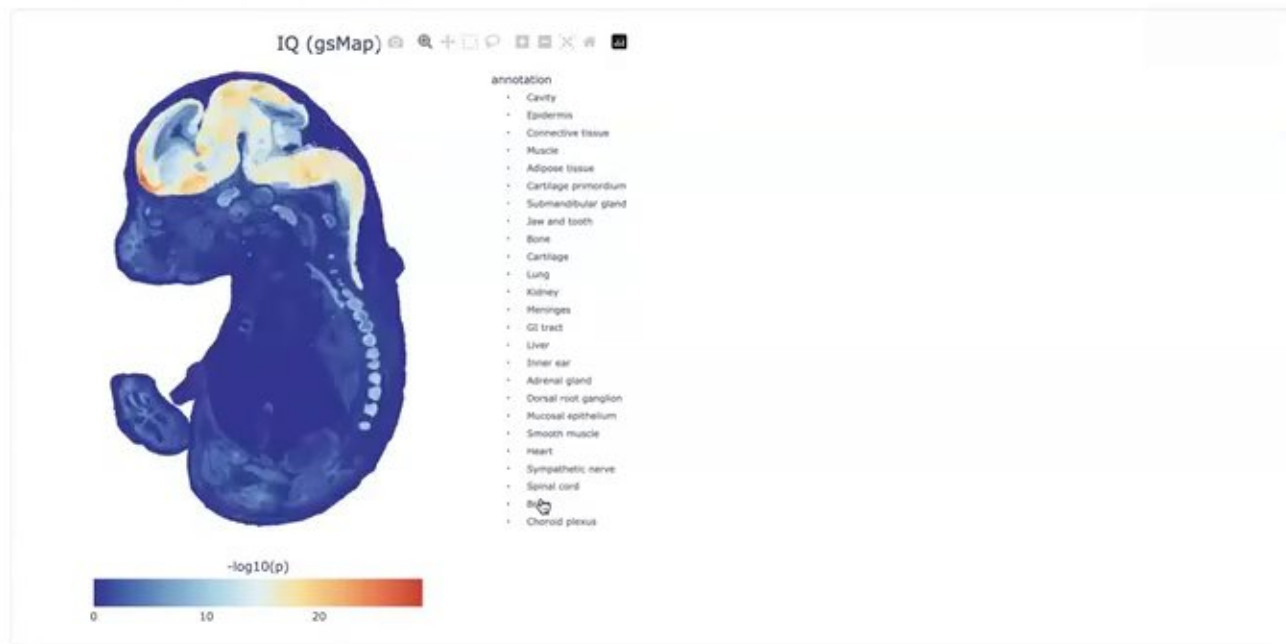
This plot shows the spatial gene expression results across different tissues.



leon song

Cauchy Combination Result

This plot shows the spatial gene expression results across different tissues.



leon song

Cauchy Combination Result

This plot shows the spatial gene clustering results across different tissues.



lean song

Cauchy Combination Result

Cauchy Combination Result

This table presents the results of the Cauchy combination test, summarizing the genetic associations.

Annotation	P Cauchy	P Median
Adipose tissue	9.6544e-01	6.9921e-01
Adrenal gland	3.6902e-01	4.4669e-01
Bone	9.9280e-01	6.9308e-01
Brain	5.1527e-26	1.2053e-15
Cartilage	9.9993e-01	2.4000e-01
Cartilage primordium	2.2454e-05	3.3773e-02
Cavity	9.9996e-01	8.0438e-01
Choroid plexus	6.9592e-10	5.1547e-06
Connective tissue	2.2180e-03	3.5259e-01

Diagnosis Manhattan Plot

The Manhattan plot shows the association of SNPs with the top associated gene across the genome.

gsMap Diagnosis Manhattan Plot

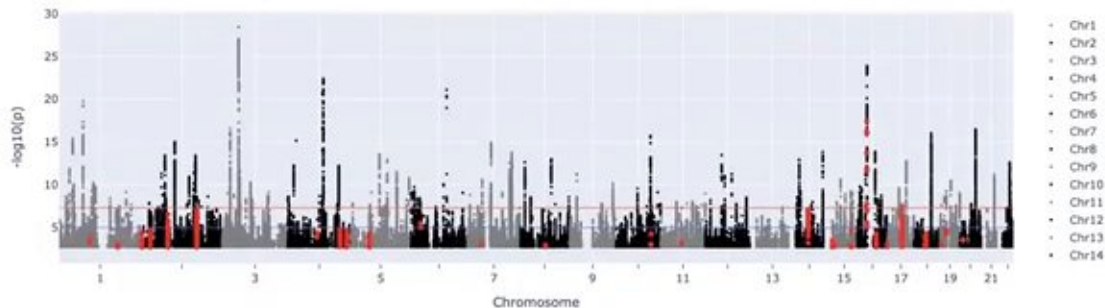
leon song



Diagnosis Manhattan Plot

The Manhattan plot shows the association of SNPs with the top associated gene across the genome.

gsMap Diagnosis Manhattan Plot



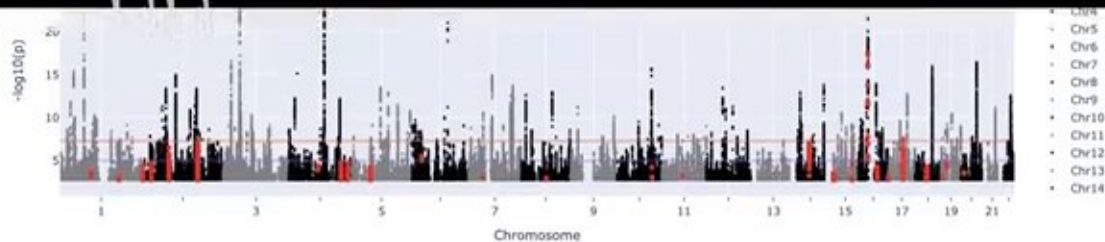
Gene Expression and GSS Distribution

Select a gene to view its expression distribution and gene specificity score (GSS).

Select a gene:

MAP2

leon song



Gene Expression and GSS Distribution

Select a gene to view its expression distribution and gene specificity score (GSS).

Select a gene:

MAP2

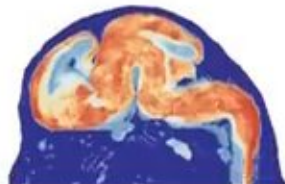
Expression Distribution

MAP2 (Expression)



Gene Specificity Score (GSS)

MAP2 (GSS)



leon song

Select a gene to view its expression distribution and gene specificity score (GSS).

Select a gene:

MAP2

Expression Distribution

MAP2 (Expression)

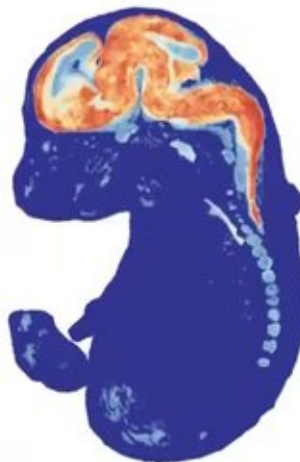


Expression



Gene Specificity Score (GSS)

MAP2 (GSS)



GSS



leon song



Gene	Annotation	Median GSS	PCC
NSG2	Brain	6.0117	0.9220
ELAVL3	Brain	5.3438	0.9217
CNIH2	Spinal cord	5.6523	0.9195
CRMP1	Spinal cord	6.4258	0.9174
CELF4	Brain	5.9809	0.9151
MYT1L	Brain	4.6562	0.9138
MLLT11	Spinal cord	6.0703	0.9120
MAPT	Spinal cord	7.7773	0.9118
RAB3C	Spinal cord	4.4414	0.9067

Running Info

Click to view detailed run information and parameters:

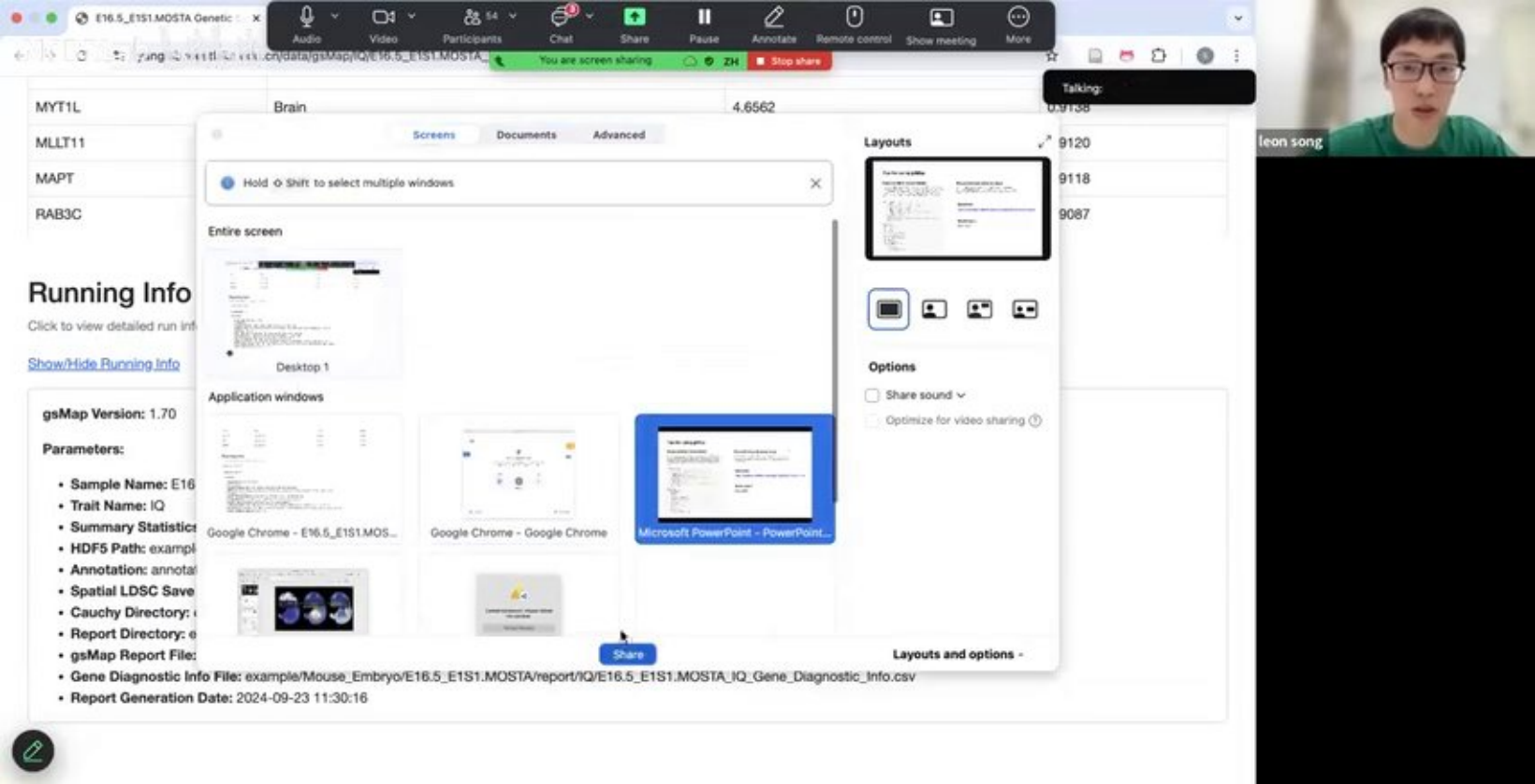
[Show/Hide Running Info](#)

gsMap Version: 1.70

Parameters:

- **Sample Name:** E16.5_E1S1.MOSTA
- **Trait Name:** IQ
- **Summary Statistics File:** gsMap_example_data/GWAS/IQ_NG_2018.sumstats.gz
- **HDF5 Path:** example/Mouse_Embryo/E16.5_E1S1.MOSTA/ind_latent_representations/E16.5_E1S1.MOSTA_ind_latent.n5ad
- **Annotations:** annotation
- **Spatial LDSC Save Directory:** example/Mouse_Embryo/E16.5_E1S1.MOSTA/spatial_ldsc

leon song



Tips for using gsMap

Mouse Embryo (Quick Mode)

The `quick_mode` option provides a simplified and efficient way to execute the entire `gsMap` pipeline. It minimizes running time and configuration complexity by utilizing pre-calculated weights based on the 1000G EUR reference panel and protein-coding genes from gtf file of Gencode v46. This mode is ideal for users who prefer a streamlined approach. For a more customizable experience, such as using custom GTF files, reference panels, and adjustable parameters, please refer to the [Step by Step](#) guide.

```
gsmap quick_mode \  
--workdir './example_quick_mode/Mouse_Embryo' \  
--homolog_file 'gsMap_resource/homologs/mouse_human_homologs.txt' \  
--sample_name 'E16.5_E151.MOSTA' \  
--gsMap_resource_dir 'gsMap_resource' \  
--hdfs_path 'gsMap_example_data/ST/E16.5_E151.MOSTA.hsad' \  
--annotation 'annotation' \  
--data_layer 'count' \  
--sumstats_file 'gsMap_example_data/GAS/10_M5_2018.sumstats.gz' \  
--trait_name 'IQ'
```

```
gsMap_resource  
├── genome_annotation  
│   ├── enhancer  
│   └── gtf  
├── homologs  
│   ├── macaque_human_homologs.txt  
│   └── mouse_human_homologs.txt  
├── LD_Reference_Panel  
│   └── 1000G_EUR_Phase3_gLink  
├── LDSC_resource  
│   ├── hapmap3_snps  
│   └── weights_hm3_no_hla  
├── quick_mode  
│   ├── baseline  
│   ├── SNP_gene_pair  
│   └── snp_gene_weight_matrix.hsad
```

Mouse Embryo (Step by Step)

This tutorial guides you through running `gsMap` step by step, with user-defined parameters and resources, granting greater flexibility and control over the analysis. This mode is suited for users who require detailed customization of their pipeline. For a faster, one-command execution, please see the [Quick Mode](#) tutorial.

Documents:

<https://yanglab.westlake.edu.cn/gsmmap/document/software>

Results report:

[web report](#)

leon song

Acknowledgments

- Dr. Jian Yang
- Wenhao Chen
- Junren Hou
- Minmin Guo
- All members in Yang Lab
- Dr. Bo Li
- Dr. Qiufu Ma

<https://yanglab.westlake.edu.cn/>



高性能计算中心
HIGH-PERFORMANCE COMPUTING CENTER



National Natural Science
Foundation of China
国家自然科学基金

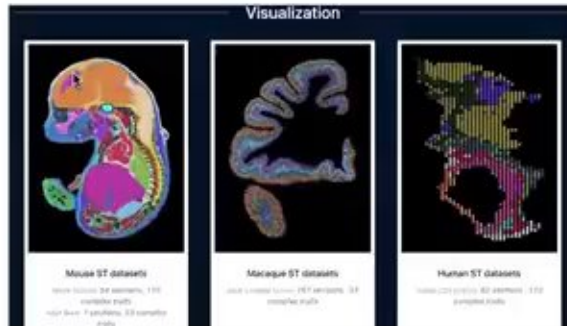
浙江省科学技术厅

gsMap software

gsMap is available at <https://github.com/JianYang-Lab/gsMap>



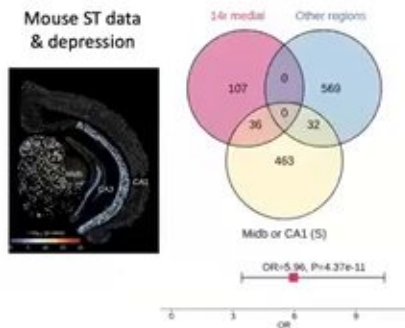
Visualization portal: <https://yanglab.westlake.edu.cn/gsmmap/visualize>



leon song

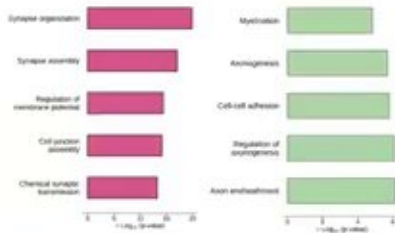
Macaque 14r medial & depression

Highly expressed genes



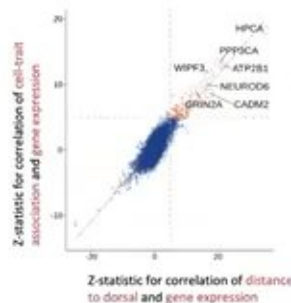
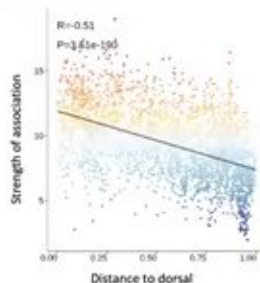
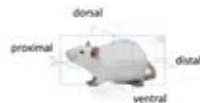
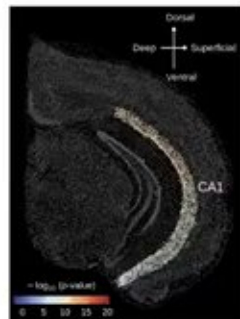
GO enrichment of highly expressed genes

Macaque 14r medial Mouse Midb and CA1 (S)



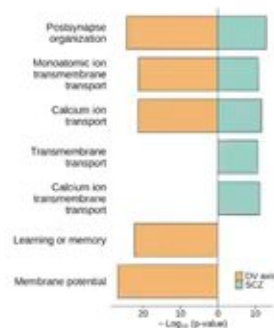
Enriched in biological processes related to neuron-plasticity

Spatially dependent associations of neurons with schizophrenia



Top three genes: *HPCA*, *PPP3CA*, and *ATP2B1*, all related to calcium signaling and regulation

GO enrichment



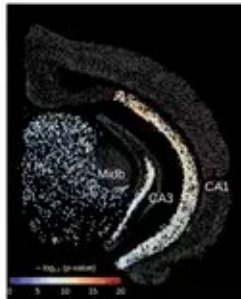


Comparison among schizophrenia, depression, and IQ

Odds ratio (OR): strength of association with a brain region relative to the whole brain

Schizophrenia

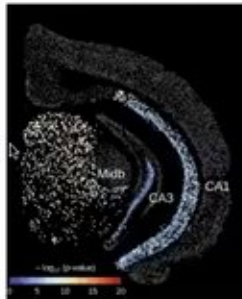
Nature, 604: 502-8 (2022)



Hippocampus CA1 + CA3: OR = 4.47
Midbrain: OR = 0.24

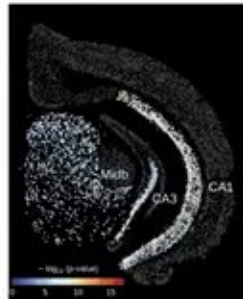
Depression

Nature Medicine, 29: 1832-44 (2023)



OR = 0.33
OR = 6.72

IQ



OR = 1.54
OR = 0.62

leon song

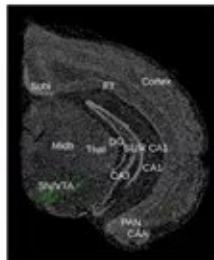
Heatmap showing the association of 15 brain regions with 25 neuropsychiatric and cognitive disorders. The color scale represents $-\log_{10}(p\text{-value})$ from 0 (grey) to 20 (dark red). The regions are listed on the y-axis, and the disorders are on the x-axis. A bar chart at the top shows the Λ_{GC} for each region. A bar chart on the right shows the number of disorders associated with each region.

Brain Region	Λ_{GC}	Number of Disorders
CAJ	~1.5	1
CAL	~1.2	1
Cortex	~1.0	1
CAA	~0.8	1
Thal	~0.7	1
Sub	~0.6	1
DG	~0.5	1
PAN	~0.4	1
Medi	~0.3	1
SUR CAL	~0.2	1
Ependymal	~0.1	1
FT	~0.1	1
SNVTA	~0.1	1
Meninge	~0.1	1

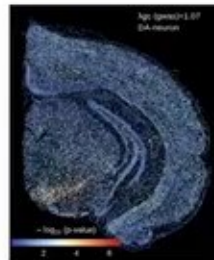
Disorders (x-axis):

- Educational Attainment
- Schizophrenia
- Intelligence
- Depression
- Cognitive Performance
- Executive Function
- Neuroticism
- Number of Siblings
- Big Five Personality
- Smoking Cessation
- Spending Propensity
- General Risk Tolerance
- ADHD
- Generational Epilepsy
- Epilepsy
- Drinks per Week
- Alcoholism
- Alcoholism
- Cognitive per Day
- Open Use Disorder
- Autism
- Age at Onset
- Subjective Well-Being
- Smoking Cessation
- Amphetamine Use Disorder
- Cannabis Use Disorder
- Trauma Syndrome
- Parkinson Disease

- **Cortex** showed the highest relevance to most traits
- The second-highest associated region varied across traits
 - **Hippocampus CA1** for IQ and schizophrenia
 - **Midbrain** for depression

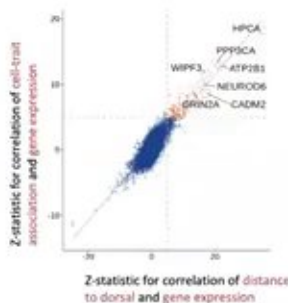
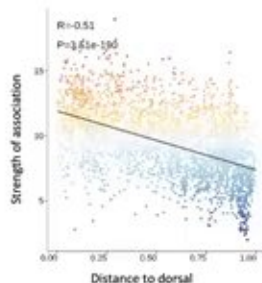
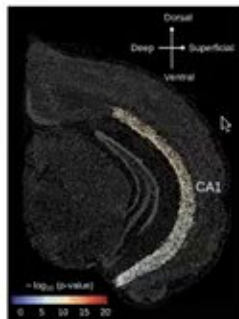


Parkinson disease



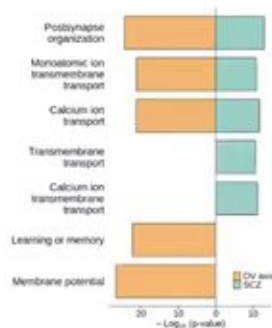
13

Spatially dependent associations of neurons with schizophrenia



Top three genes: *HPCA*, *PPP3CA*, and *ATP2B1*, all related to calcium signaling and regulation

GO enrichment

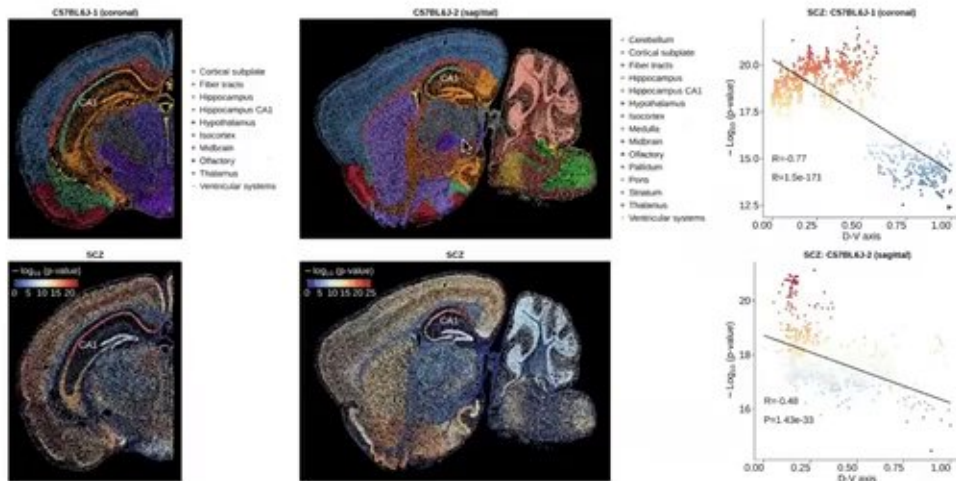




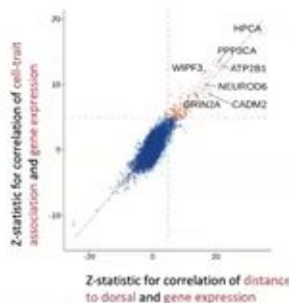
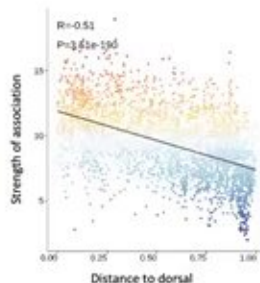
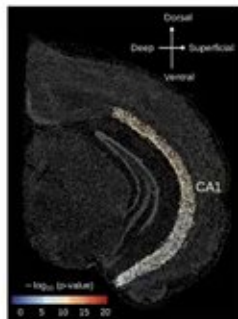
leon song

Spatially dependent associations of neurons with schizophrenia

Multiplexed error-robust fluorescence in situ hybridization (MERFISH) + scRNA-seq *Nature* 624: 343–354 (2023)

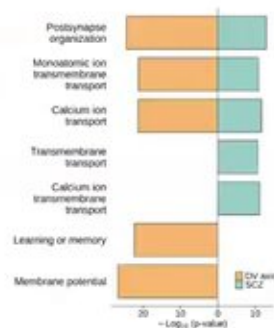


Spatially dependent associations of neurons with schizophrenia



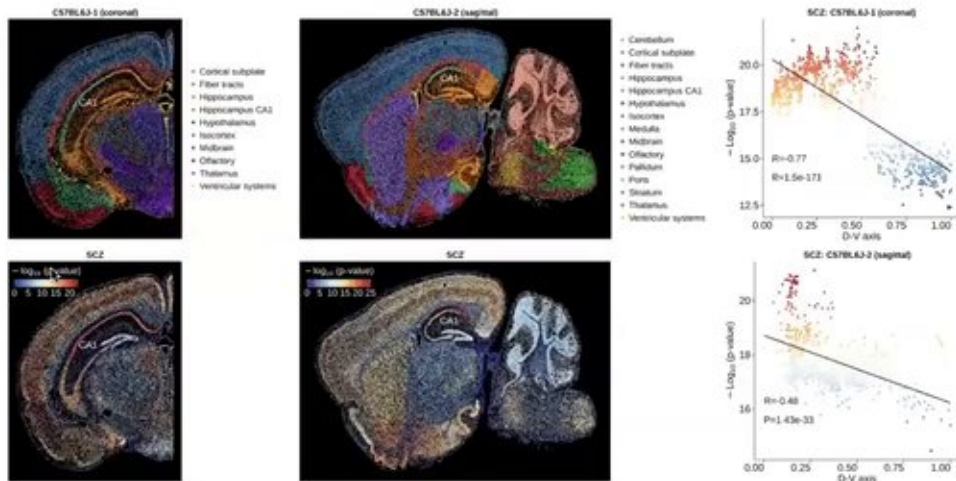
Top three genes: *HPCA*, *PPP3CA*, and *ATP2B1*, all related to calcium signaling and regulation

GO enrichment



Spatially dependent associations of neurons with schizophrenia

Multiplexed error-robust fluorescence in situ hybridization (MERFISH) + scRNA-seq *Nature* 624: 343–354 (2023)

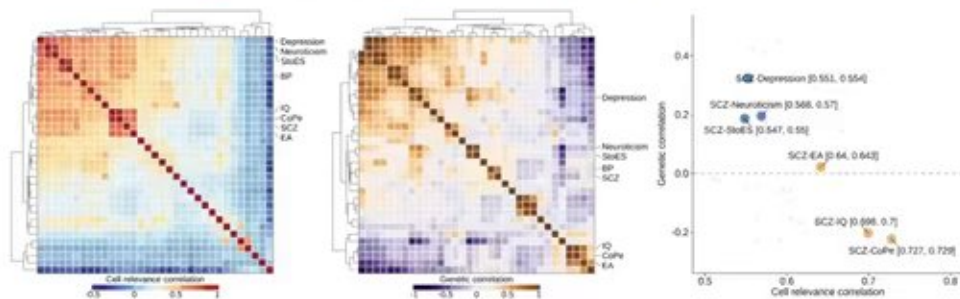


Cell-relevance correlation (CRC)

SCZ & IQ: Cortex, Hippocampus CA

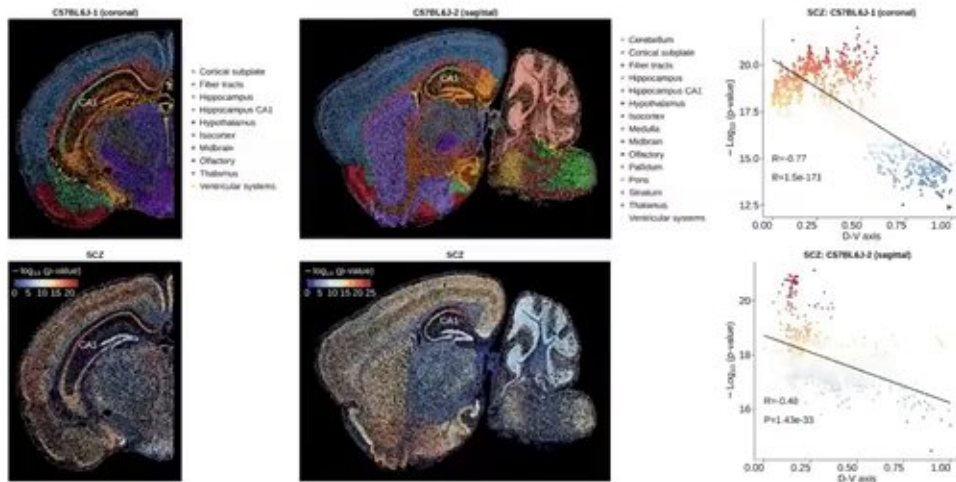
Depression: Midbrain, PFC 14r

The neurons associated with **SCZ** more resembles that for **IQ** than that for **depression**.



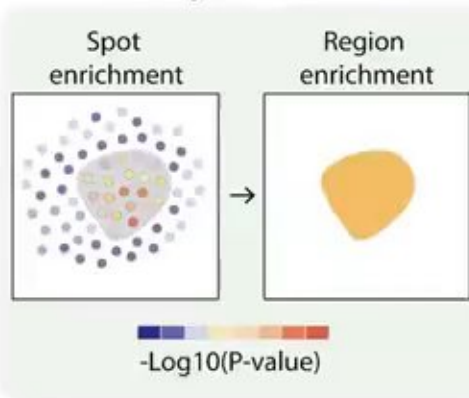
Spatially dependent associations of neurons with schizophrenia

Multiplexed error-robust fluorescence in situ hybridization (MERFISH) + scRNA-seq *Nature* 624: 343–354 (2023)



Region-level associations

Cauchy combination

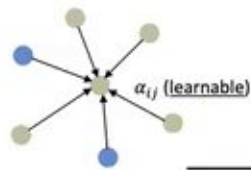


leon song

Spot-level gene specificity score

Graph attention autoencoder (GAT)

Spatial graphs



Embedding matrix



Homogenous spots D_i

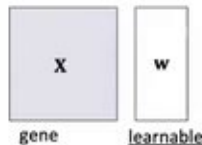


Exponential projection of relative rank

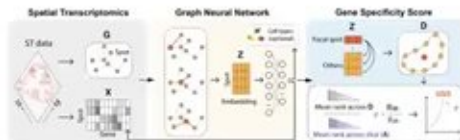
$$F_{ig} = \exp \left\{ \frac{n \sum_{k \in D_i} \log(R_{kg})}{d \sum_{j=1}^n \log(R_{jg})} \right\}$$

$$GSS_{ig} = \exp[F_{ig}^2] - 1$$

Expression matrix



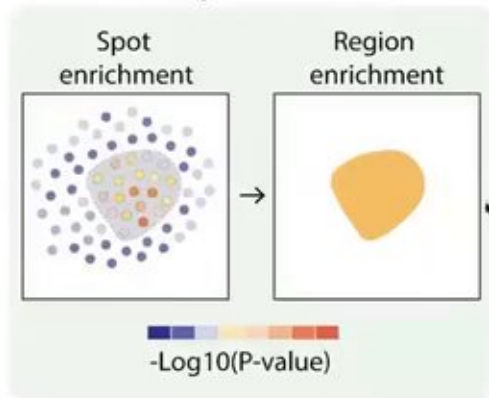
Spatial neighboring spots



leon song

Region-level associations

Cauchy combination

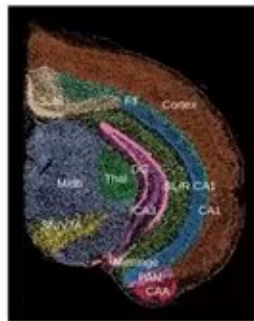


朱麗蓉

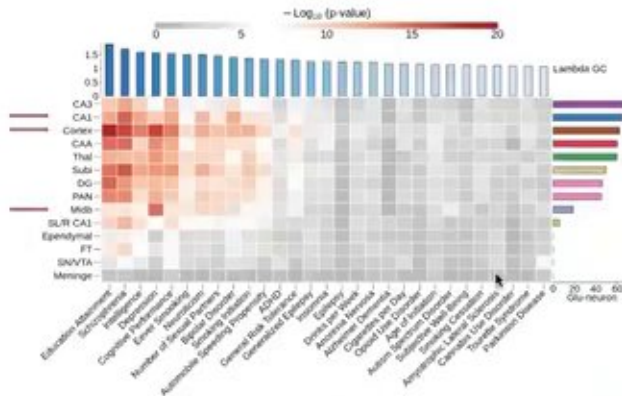


Associating cells in adult mouse brain with 29 human complex traits

Adult mouse brain
(single-cell resolution)



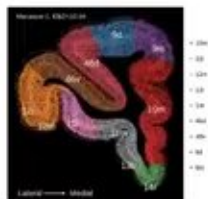
Cell, 185: 1777-92 (2022)



- **Cortex** showed the highest relevance to most traits
- The second-highest associated region varied across traits
 - **Hippocampus CA1** for IQ and schizophrenia
 - **Midbrain** for depression

Associating cells in macaque prefrontal cortex with human traits

Macaque prefrontal cortex (PFC)
(single-cell resolution)



Macaque 14r medial & psychiatric drugs

leon song



蔻享学术-直播