

BGI 华大

Genome-wide association studies of brain imaging phenotypes in UK Biobank

陈逸希 曾镜琚 曹晨音 任宇晗

华大生命科学研究院
BGI • research

前提：大脑结构和功能的遗传基础在很大程度上是未知的。

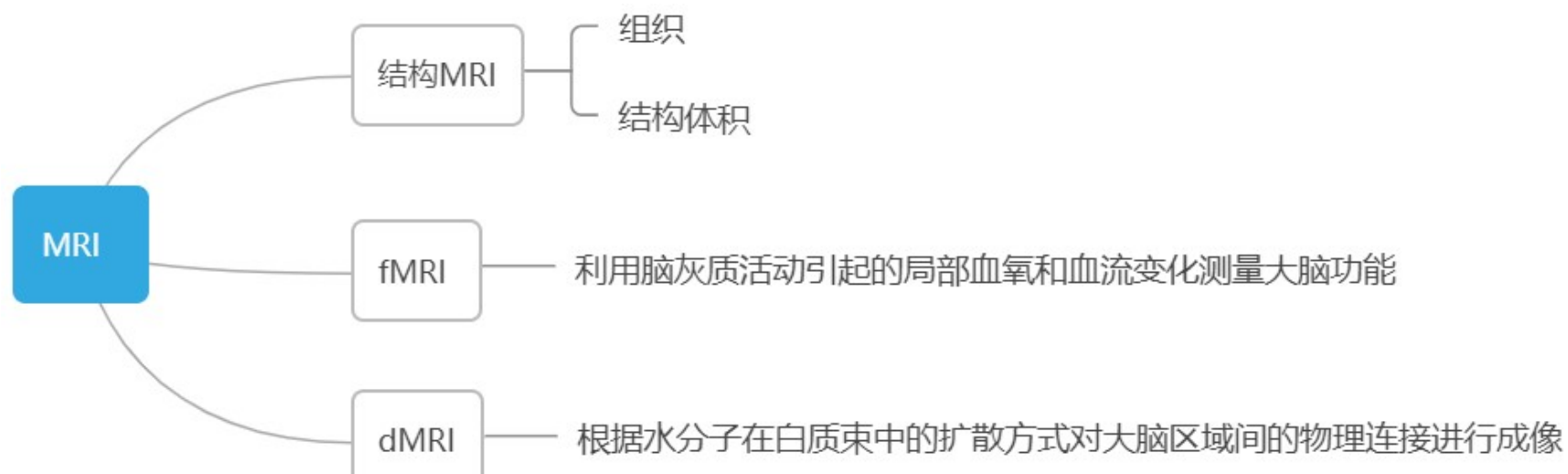
样本：UK BIOBANK 8428受试者的3144种结构和功能脑影像学表型（IDPs）

结果：许多表型是可以遗传的，鉴定了多个SNP与影像学表型之间的关联

可解释的关联包括：

- 1.铁转运和储存相关基因 与 皮层下脑组织磁化率
- 2.细胞外基质和表皮生长因子基因 与 白质微结构病变
- 3.调节中线轴突发育的基因 与 脑桥交叉纤维束

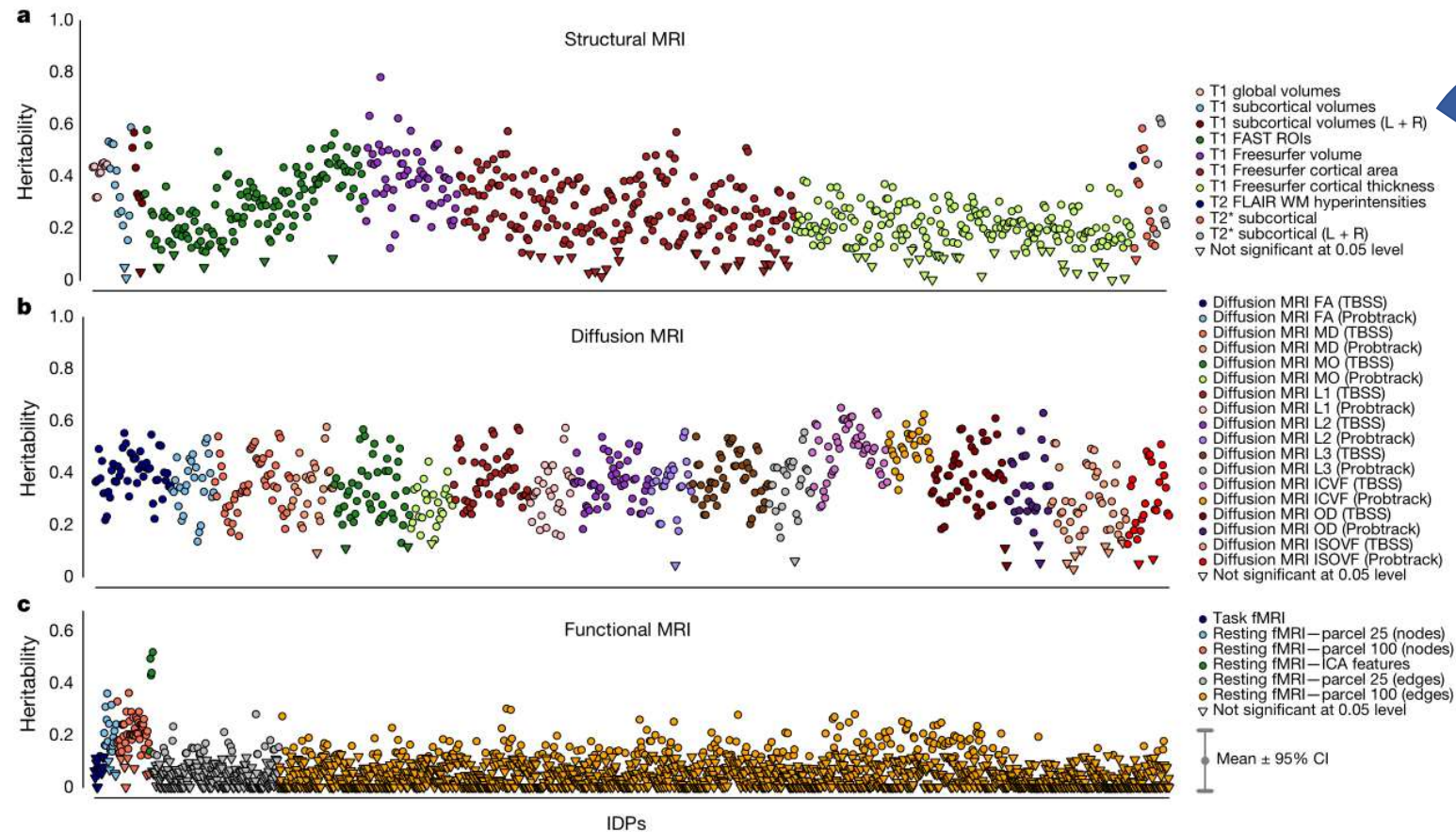
MRI：磁共振成像



UK BIOBANK：同时拥有人脑IDPs和全基因组SNP遗传数据

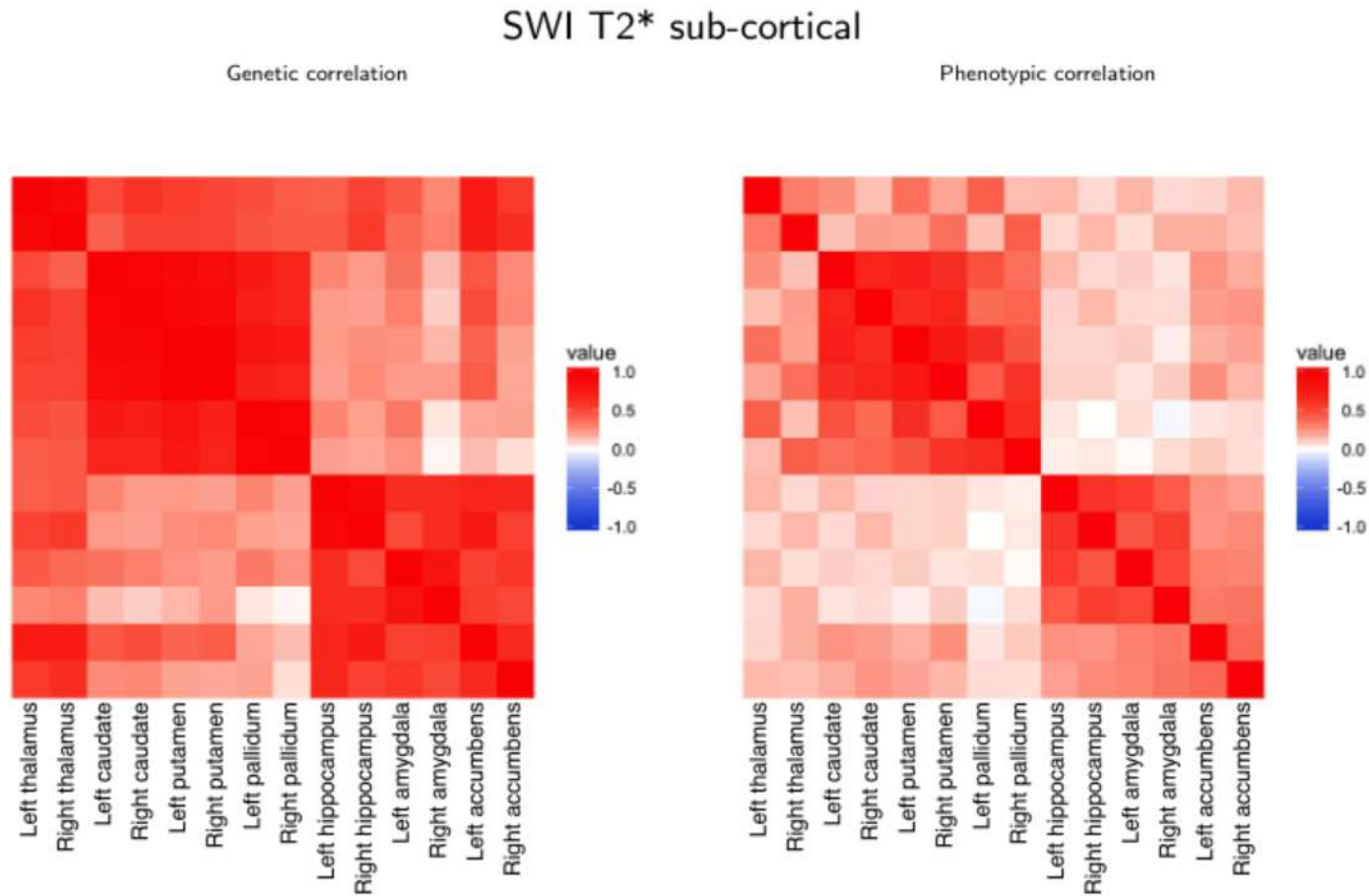
对所有IDPs估计SNP的遗传力。

IDPs被分为三大类：a、结构MRI。b,扩散MRI。c,功能MRI。



使用SBAT软件中的线性混合模型来计算性状的遗传力

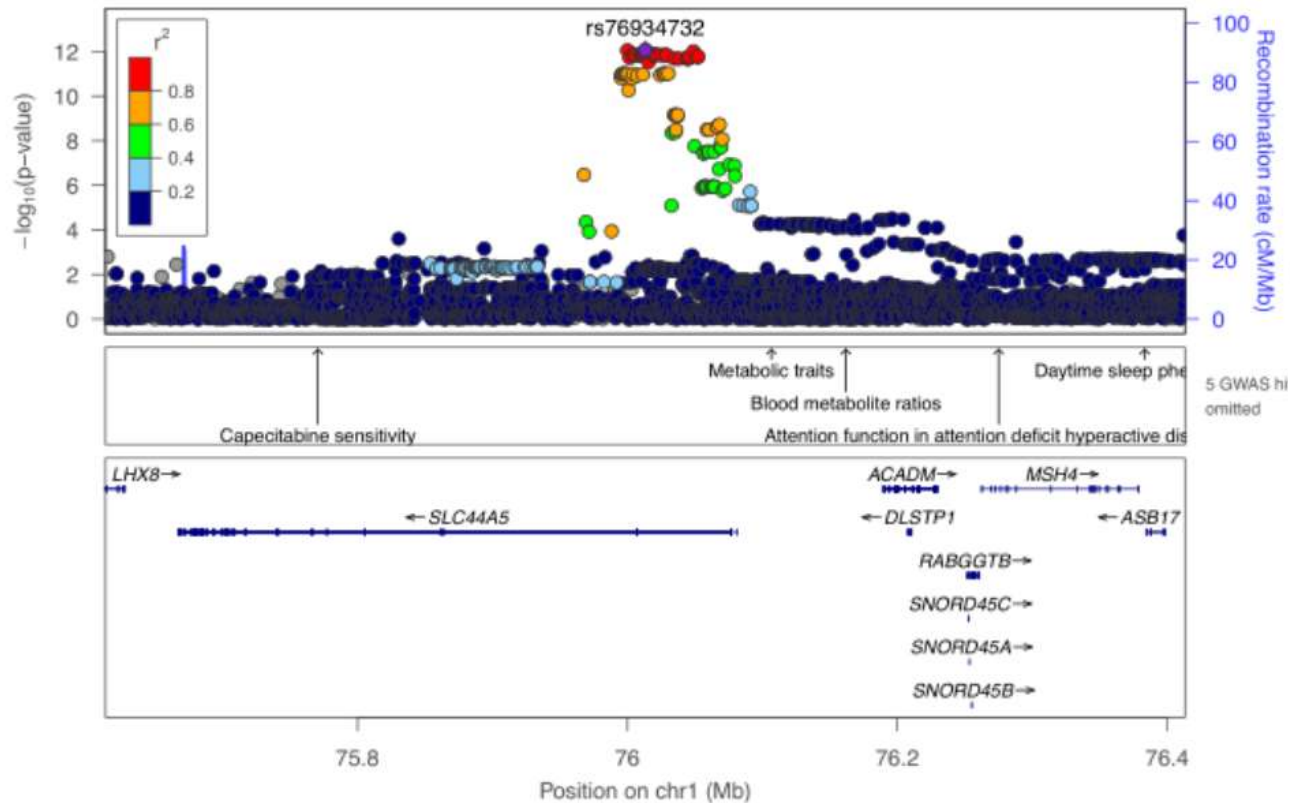
Heritability and genetic correlations of IDPs



存在于不同IDPs之间的表型相关性和遗传相关性
使用SBAT软件中的多性状混合模型估计遗传相关性

Significant associations between IDPs and SNPs

FAST_ROIs_V_cerebellum_VIIa



date: Thu Aug 17 17:58:46 2017

build: hg19

display range: chr1:75613268–76413268 [75613268–76413268]

highlight range: 0 – 0 [0 – 0]

reference SNP: chr1:76013268

number of SNPs plotted: 3686

min P-value: 8.51E–13 [chr1:76013268]

max P-value: 10E–1 [chr1:75725482]

omitted GWAS Hits: chr1:76.383297–Daytime sleep phenotypes, NA

omitted GWAS Hits: NA, NA

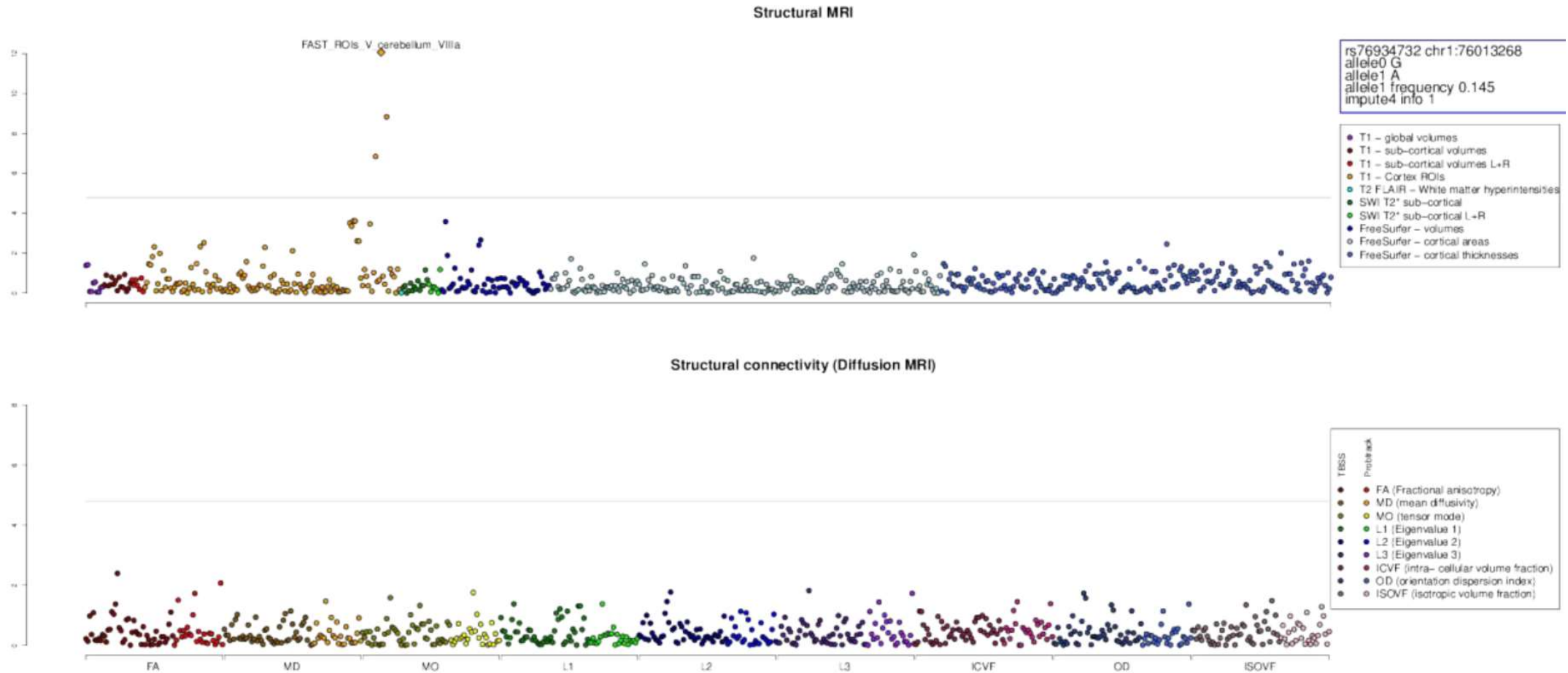
368个基因区域和不同IDPs之间的显著关联

Significant associations between IDPs and SNPs

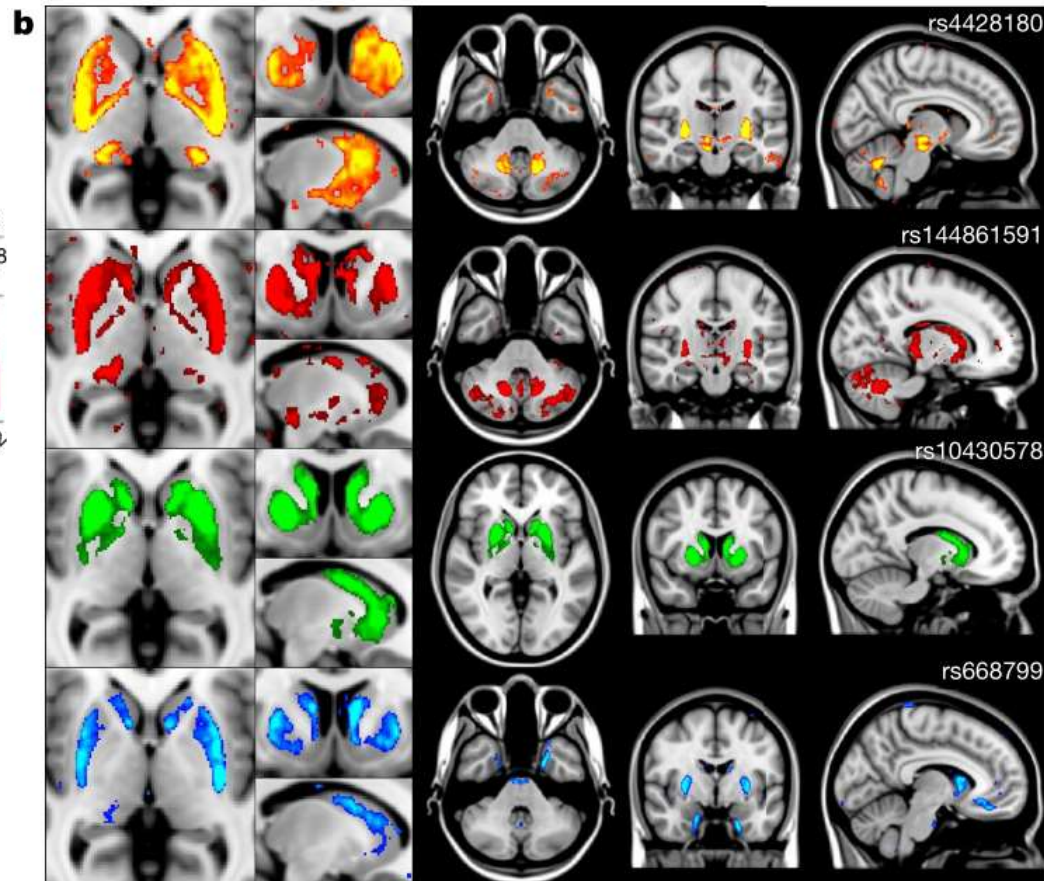
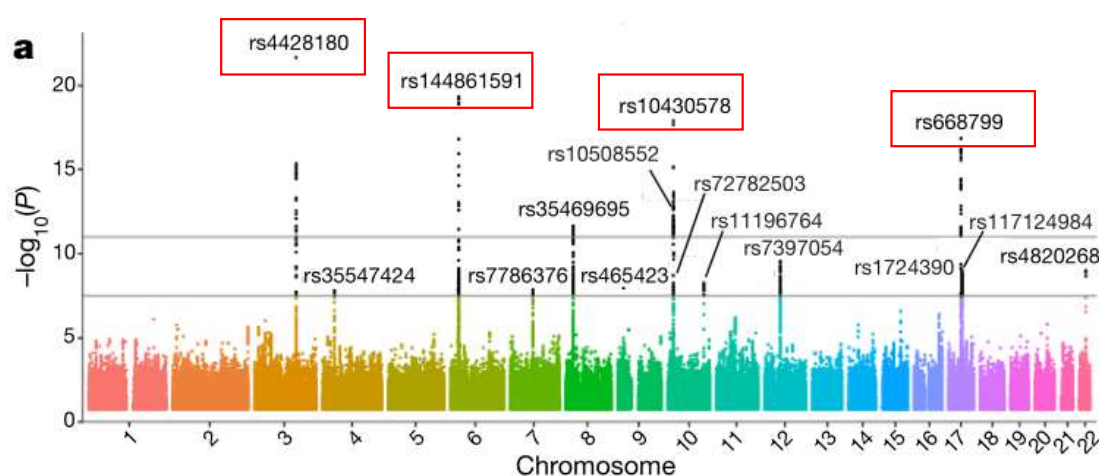
cluster index	cluster name	# IDPs	top IDP	chr	RSID	position	locus	ref allele	nonref allele	nonref AF	p value	replication p-value (N=3456)	replication p-value (N=930)	GTEx eQTL
1	Volume Cerebellum Villa (vermis)	1	T1_FAST_ROIs_V_cerebellum_Villa	1	rs76934732	76013268	SLC44A5	G	A	0.145	8.51E-13	6.10E-04	5.22E-02	SLC44A5 ACADM
2	dMRI Corpus callosum (genu)	1	dMRI_TBSS_ICVF_Genu_of_corpus_callosum	1	rs2365715	156615114	BCAN	A	G	0.388	5.38E-12	4.50E-03	1.33E-02	BCAN, APOA1BP, SYT11
3	Volume WM lesions	1	T2_FLAIR_BIANCA_WMH_volume	2	rs3762515 (5' UTR)	56150864	EFEMP1	C	T	0.0959	4.27E-13	1.18E-02	4.84E-01	
4	rfMRI Cortical and cerebellar motor nodes and edges	2	NODEamps25_0012	2	rs60873293	114092549	intergenic	G	T	0.217	9.86E-15	3.10E-07	9.50E-02	AC016745.3, RP11-480C16.1
5	T2* Pallidum	1	SWI_T2*_pallidum_L+R	2	rs6740926	190326498	WDR75	C	T	0.038	1.31E-14	3.50E-09	3.78E-04	WDR75
6	rfMRI Middle temporal sulcus nodes and edges	2	netmat_ICA_003	3	rs35124509 (missense)	89521693	EPHA3	T	C	0.3853	4.49E-22	3.27E-09	3.73E-03	EPHA3
7	T2* Putamen and pallidum	6	SWI_T2*_putamen_L+R	3	rs4428180	133466374	TF	A	G	0.152	2.23E-22	6.11E-07	1.03E-03	TF
8	rfMRI Prefrontal and parietal edges	1	netmat_ICA_002	3	rs2279829 (3' UTR)	147106319	ZIC4	C	T	0.221	8.34E-12	5.46E-05	2.51E-03	
9	dMRI Superior cerebellar peduncles	8	dMRI_TBSS_ICVF_Superior_cerebellar_peduncle_L	4	rs4697414	23724255	RP11-380P13.2	C	T	0.823	5.83E-24	1.33E-06	4.63E-02	RP13-497K6.1, RP11-380P13.2
10	Volume Putamen, ventral striatum, cerebellum VIIIb, IX, X; T2* Pallidum; dMRI Cerebral peduncles	20	IDP_T1_FAST_ROIs_L_ventral_striatum	4	rs13107325 (missense)	103188709	SLC39A8	C	T	0.073	1.04E-42	6.64E-20	8.97E-06	
11	dMRI Most WM tracts	199	dMRI_ProbtrackX_ICVF_inf_r	5	rs67827860	82860485	VCAN	C	T	0.188	4.06E-37	3.93E-12	2.19E-04	
12	rfMRI Parietal and prefrontal edges	1	netmat_ICA_004	5	rs7442779	92788278	NR2F1-AS1	A	G	0.05	8.18E-15	1.90E-04	4.04E-02	
13	dMRI Corpus callosum (genu, body, splenium)	7	dMRI_TBSS_ICVF_Genu_of_corpus_callosum	5	rs4150221	139719991	HBEGF	T	C	0.264	8.43E-20	1.72E-09	4.06E-02	SRA1

该表汇总了38组SNP-IDP关联的聚类(n = 8428)。如果SNP在该基因中，基因座列将详细说明该基因。

Significant associations between IDPs and SNPs



Significant associations between IDPs and SNPs



硬膜T2*与4个SNPs关联的曼哈顿图和体素SNP空间关联图。

a.与双侧硬膜内IDP T2*的GWAS有关的曼哈顿图。

b.空间图显示，硬膜中与T2*联系最紧密的四个SNPs(每行一个)在整个大脑中有不同的体素模式影响：

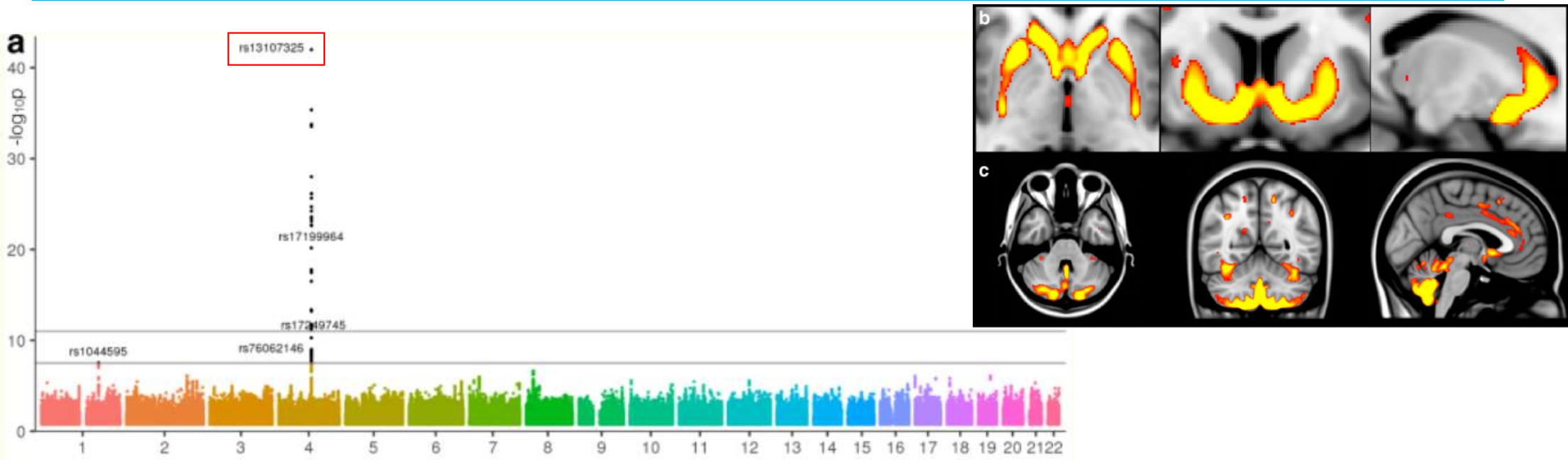
rs4428180 (TF)的作用存在于尾状核

rs144861591 (HFE)位于小脑的背纹状体、丘脑底核

rs10430578 (SLC39A12)在整个背侧纹状体和苍白球

rs668799 (COASY)在整个背侧纹状体和皮层。

Significant associations between IDPs and SNPs

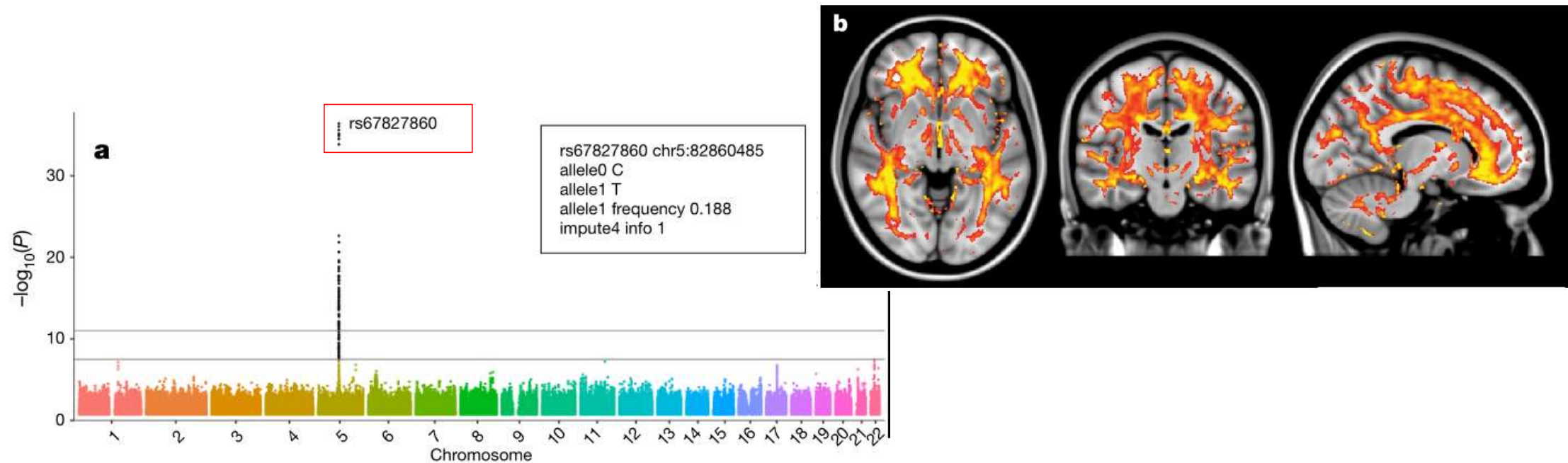


灰质体积与rs13107325 (SLC39A8)之间关联的曼哈顿图和体素SNP空间关联图。

a.与左腹侧纹状体灰质体积IDP的GWAS有关的曼哈顿图。

b, c.rs13107325对体素局部灰质体积的空间映射。该图谱显示rs13107325的作用更广泛地见于双侧的腹尾状核、壳核、腹侧纹状体、前扣带皮层。

Significant associations between IDPs and SNPs

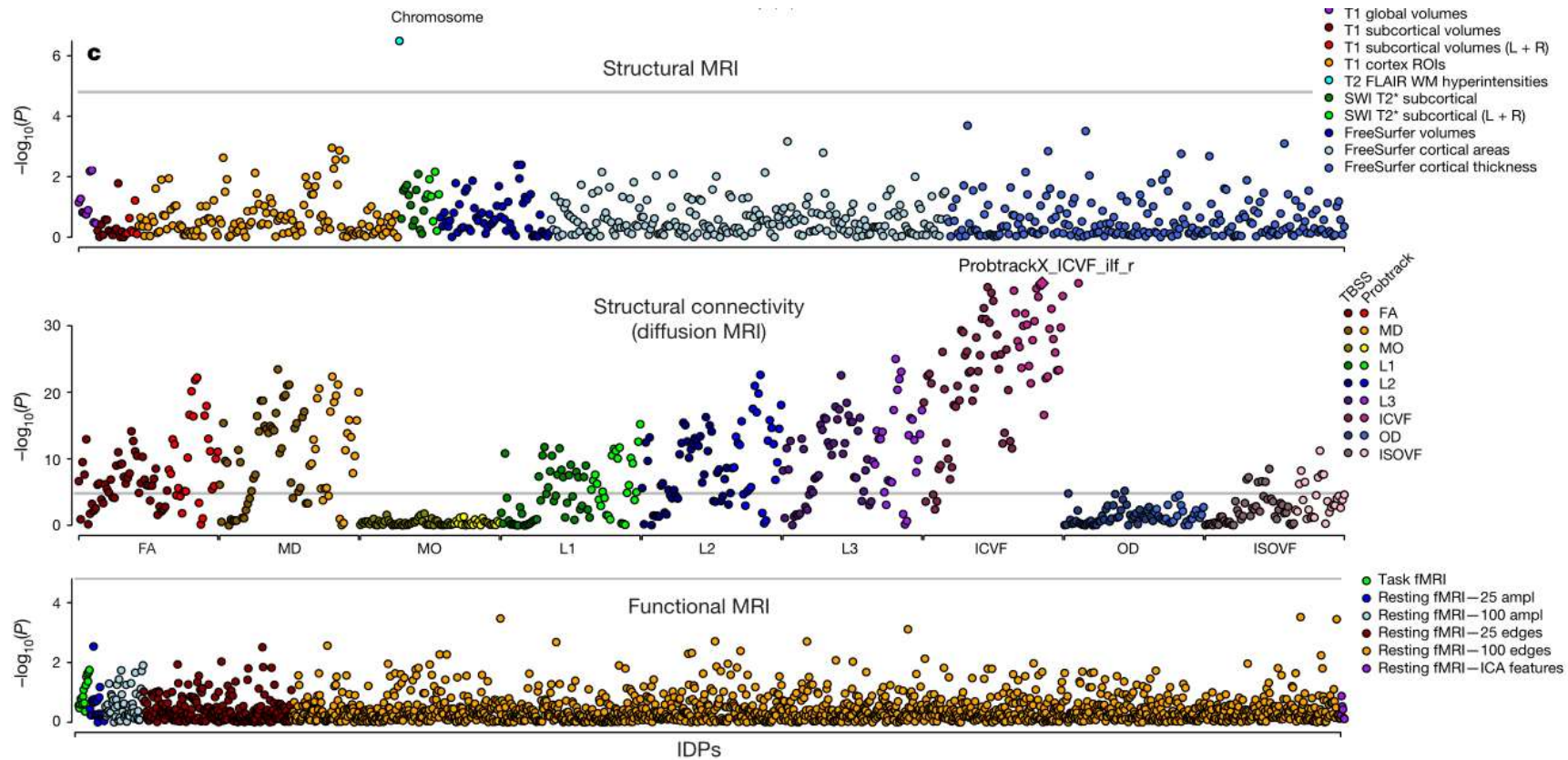


dMRI 影像中ICVF指标GWAS分析的曼哈顿图以及相关SNP (rs67827860) 的相关脑区分布图。

a. ICVF的GWAS分析的曼哈顿图，ICVF与rs67827860显著相关。

b. rs67827860与体素级别白质ICVF关联的脑区分布图。

Significant associations between IDPs and SNPs



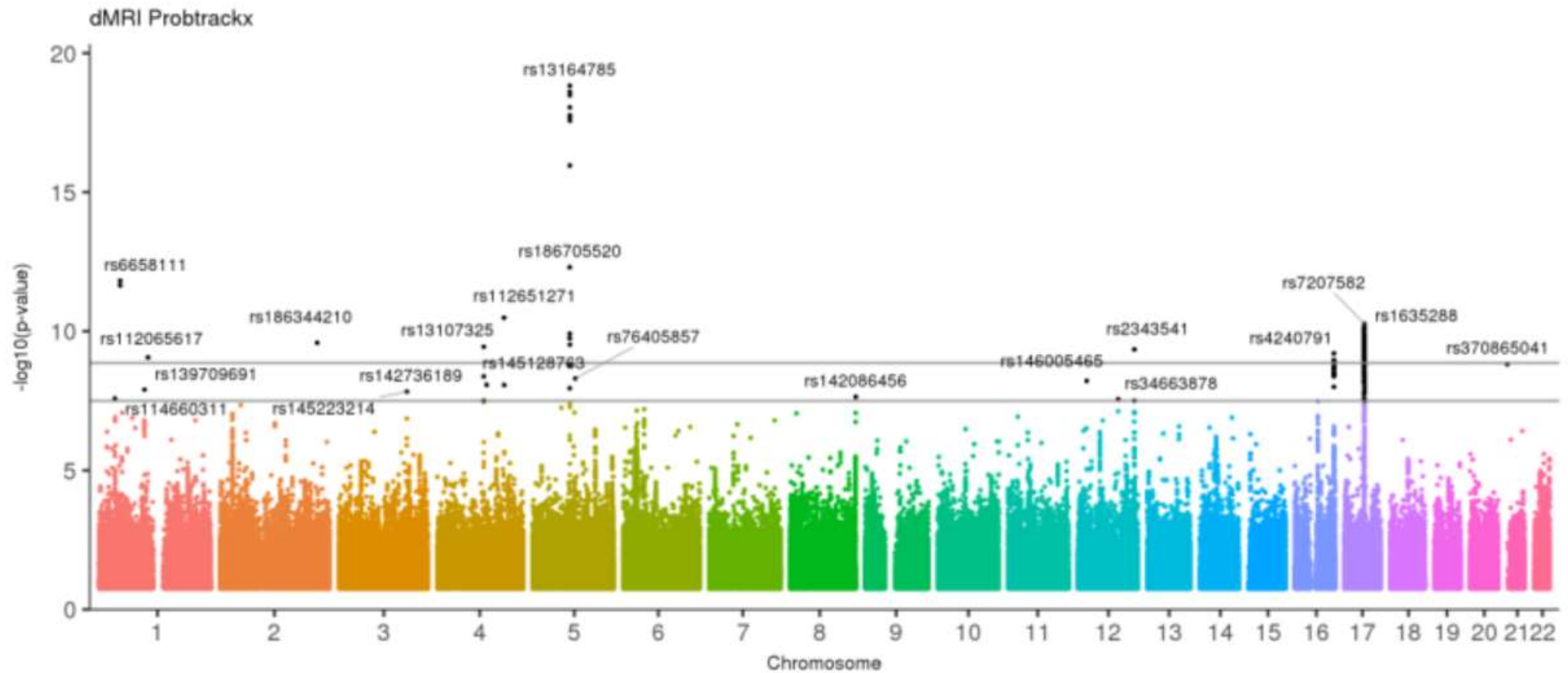
全表型组关联分析结果

c. rs67827860与3144个影像指标的全表型组分析结果图。点的颜色用来表示不同的影像指标类型。该snp除了同白质高信号相关，还同其他多种dMRI影像指标显著相关。

多表型联合测试：
用来对大规模影像指标进行关联分析，来检验每个SNP与不同性状分组的关联。

利用多表型关联分析方法对23组IDP进行分析，在 $-\log_{10}(P) > 7.5$ 的阈值下共发现278个SNP与影像指标显著相关，其中有170个SNP在多次矫正后仍是显著的。这170个SNP中有138个可以在另一个由3456人组成的独立数据集上满足 $p < 0.05$ 。

IDP group	Number of IDPs per group
T1_brain_vol_all	10
T1_Subcortical_all	15
T1_Subcortical_L_plus_R	7
T1_FAST_ROIs	139
Freesurfer_volume	59
Freesurfer_area	212
Freesurfer_thickness	212
T2_star	14
T2_star_L_plus_R	7
dMRI_Probtrackx	243
dMRI_FA	75
dMRI_MD	75
dMRI_MO	75
dMRI_L1	75
dMRI_L2	75
dMRI_L3	75
dMRI_ICVF	75
dMRI_OD	75
dMRI_ISOVF	75
rfMRI_ICA_Features	6
tfMRI	16
rfMRI_25	21
rfMRI_100	55



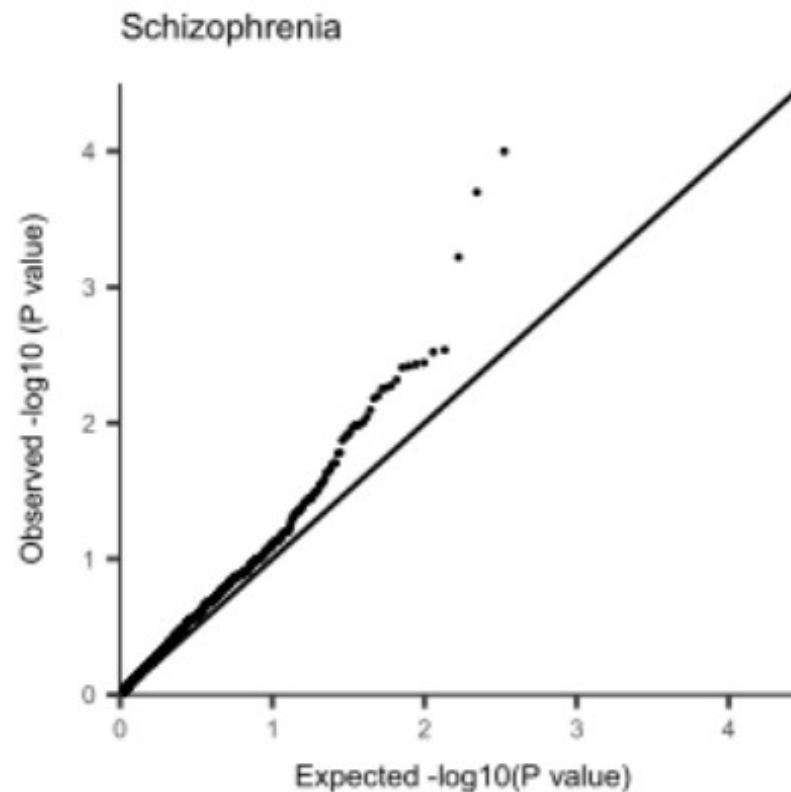
23组IDPs进行多性状GWAS的GWAS曼哈顿图。

多表型联合测试优点：

- 1.通过对一组影像指标拟合相关联合模型，利用基因相关性来提高分析的可靠程度。
- 2.在一个GWAS分析中同时分析多个表型可以避免由于进行多次GWAS分析带来的多重矫正问题。

使用LD Score回归估计IDPs与神经质人格特征、自闭症谱系和睡眠时间，以及7个疾病特征:注意缺陷多动障碍、精神分裂症、重度抑郁症和双相情感障碍、阿尔茨海默病、中风和肌萎缩性侧索硬化症的GWAS数据。

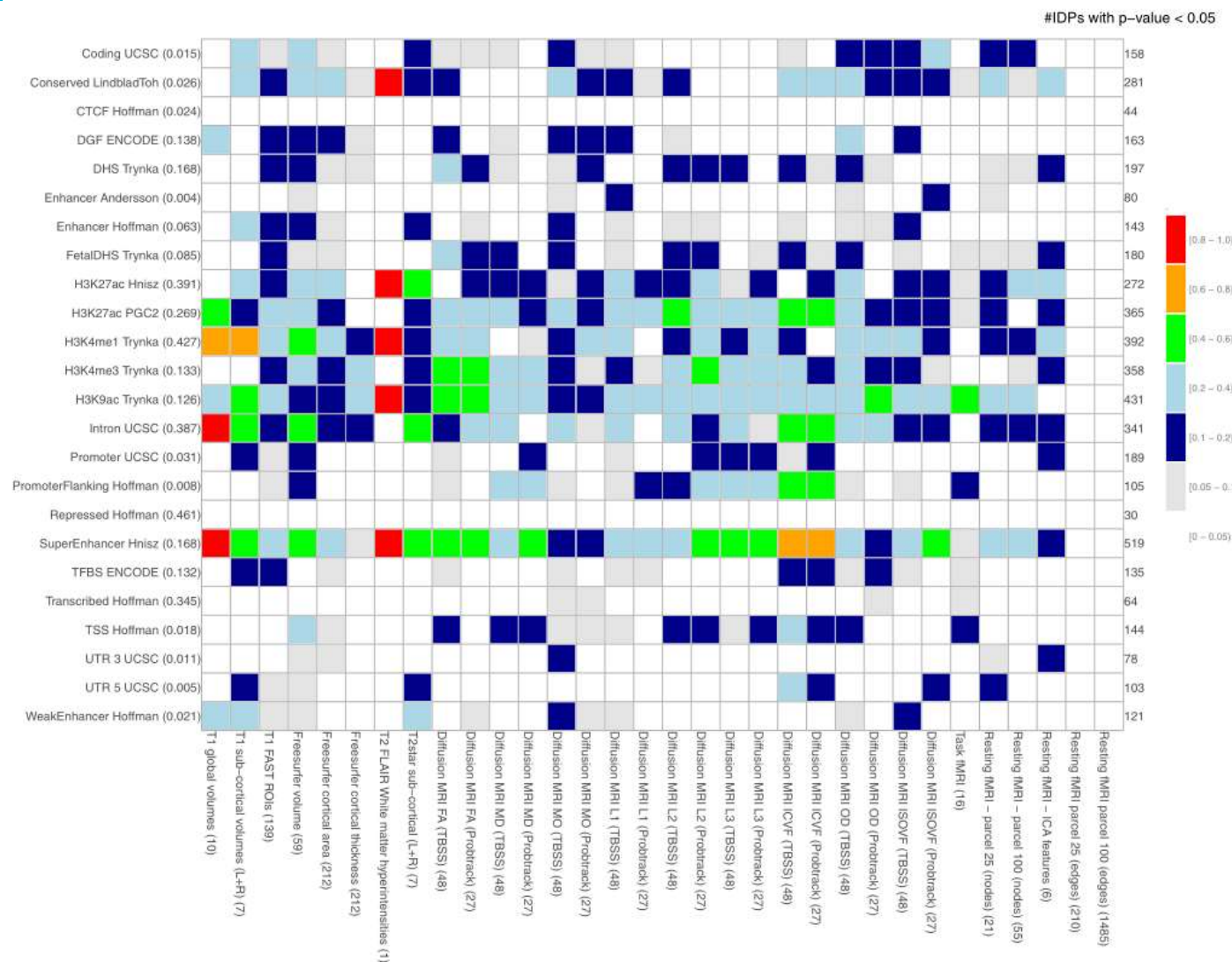
发现肌萎缩侧索硬化症（ALS）,精神分裂症以及中风主要和白质纤维束的dMRI影像指标相关。



Genetic correlation with clinically relevant traits


78个突变根据其在基因组中的不同功能分为24种功能类别。
超级增强子及组蛋白修饰相关的区域在多个功能和结构影像指标组中富集,结果与其具有较高的遗传力结果一致。

X轴为23组IDPs
Y轴为24种功能类别的突变



Letter | [Published: 25 June 2018](#)

Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence

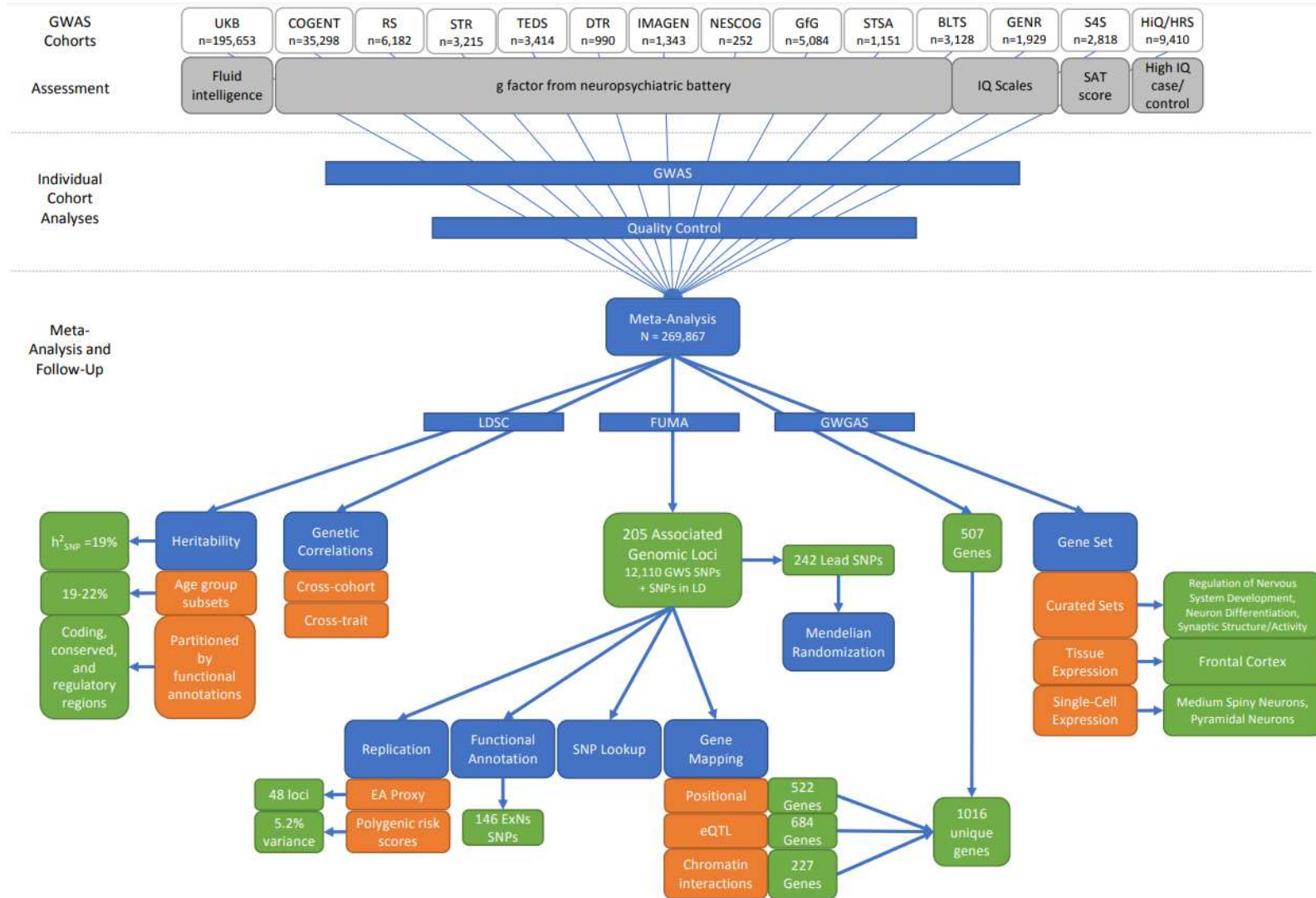
[Jeanne E. Savage](#), [Philip R. Jansen](#), [...] [Danielle Posthuma](#) 

[Nature Genetics](#) **50**, 912–919 (2018) | [Cite this article](#)

24k Accesses | **285** Citations | **489** Altmetric | [Metrics](#)

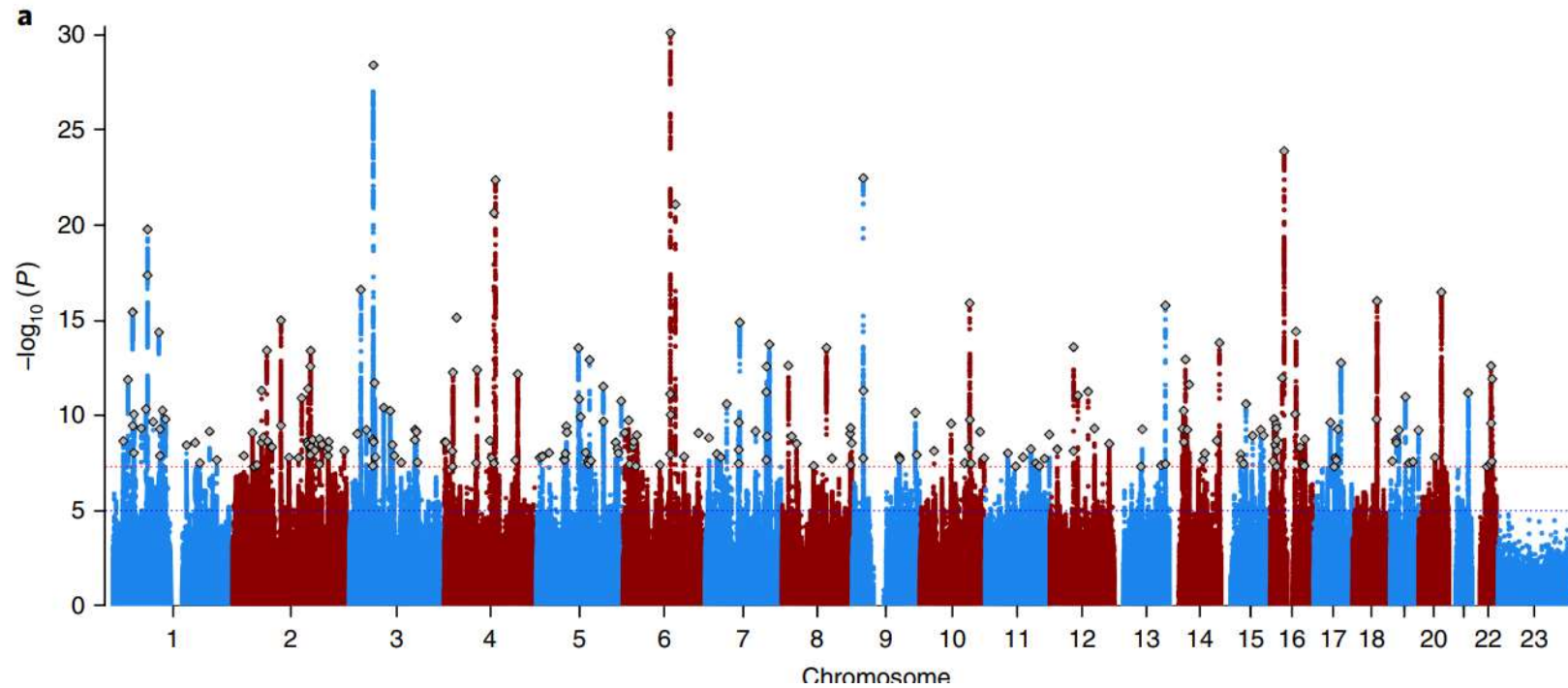
- 智力在过往的研究中已经被证明是高度遗传的，同时与健康 and 正常行为也有很高的关联性，本文发表之前仅有24个基因位点被发现与智力有关
- 本文使用了269,867例基因数据和认知测试结果进行了全基因组关联分析，发现了190个全新的基因位点和使用eQTL方法、染色体互作和gene-based关联测试发现的939个与智力相关的基因。
- 本文还使用数据证明了智力与健康之间的关联性，还使用孟德尔随机化（MR）发现了智力对于阿尔兹海默症和多动症有保护作用，但对精神分裂有着多样的双向因果关系。

meta-analysis in 269,867 individuals



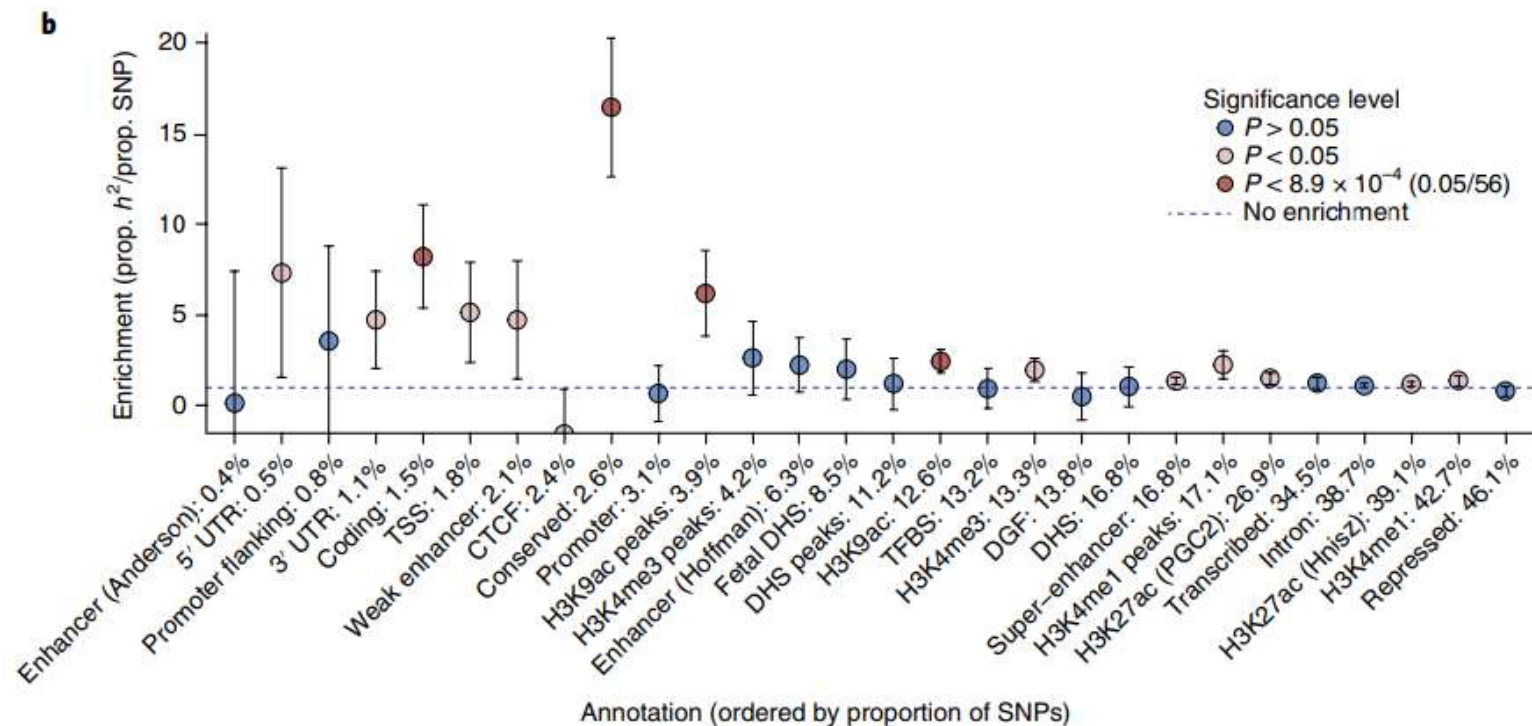
- 文章主要数据来源和研究依据
- 文章使用了来自14个欧洲人群，总共包含9,295,118个变异的数据
- 智力数据来自各式各样的认知测试结果，基本都是测试认知功能的流体部分 (fluid domains of cognitive functioning)
- 衡量智力主要依靠g值，同时在不同的测试方法中都能发现较好的一致性
- 各个样本之间的基因一致性较好 (mean $r=0.67$)，智力遗传力同之前的工作也是一致的 ($h^2=0.22$)

- 关联分析结果
- 共有12,110个和智力显著关联的变异被发现，包含242个lead SNP，205个基因



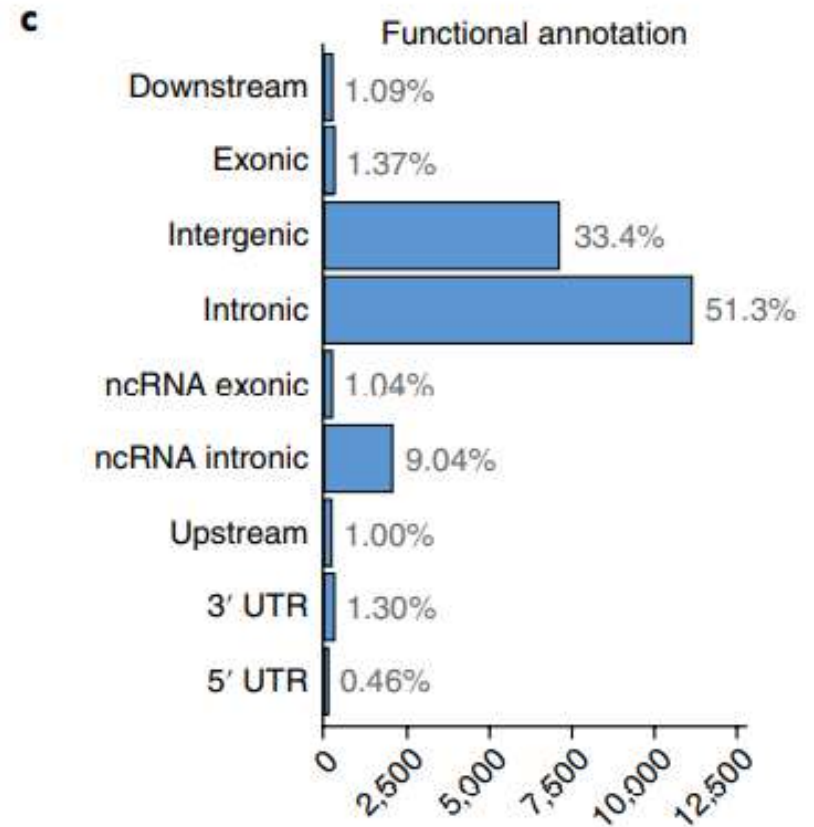
- 关联分析结果

- 在关联分析结果中还发现遗传力高的SNP集中在保守区间、编码区间和组蛋白Lys9乙酰化区间 (H3K9ac)



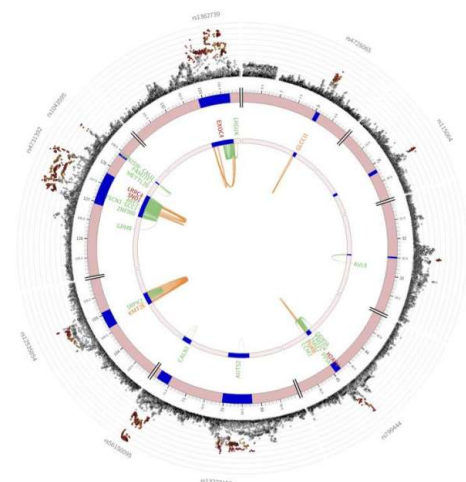
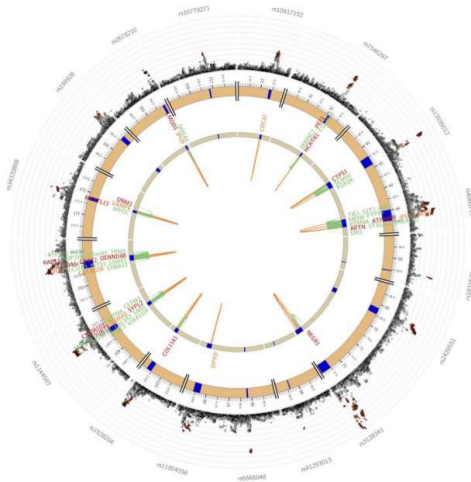
- 关联分析结果

- 对变异结果的功能注释显示主要的变异是在内含子 (intron) 和基因间区域内 (intergenic), 但仍有部分SNP是外显子内的非同义突变
- 发现了rs13107325, 一个与智力有强烈关联的错义突变。

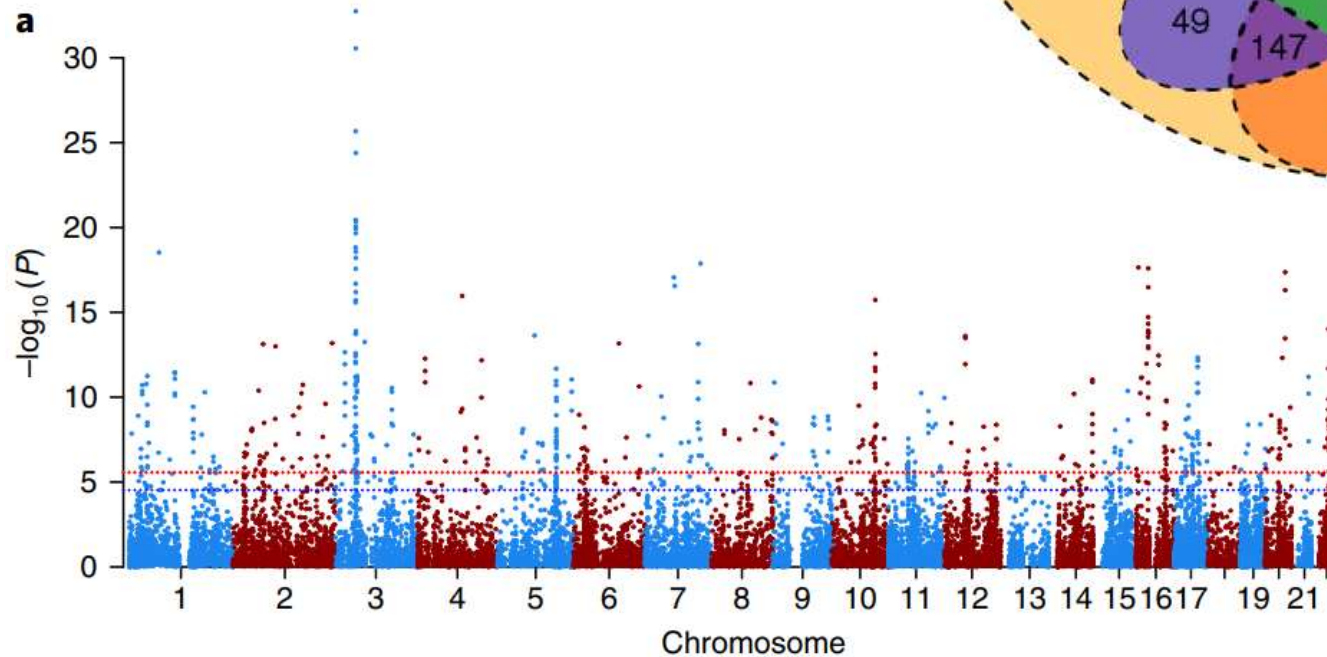
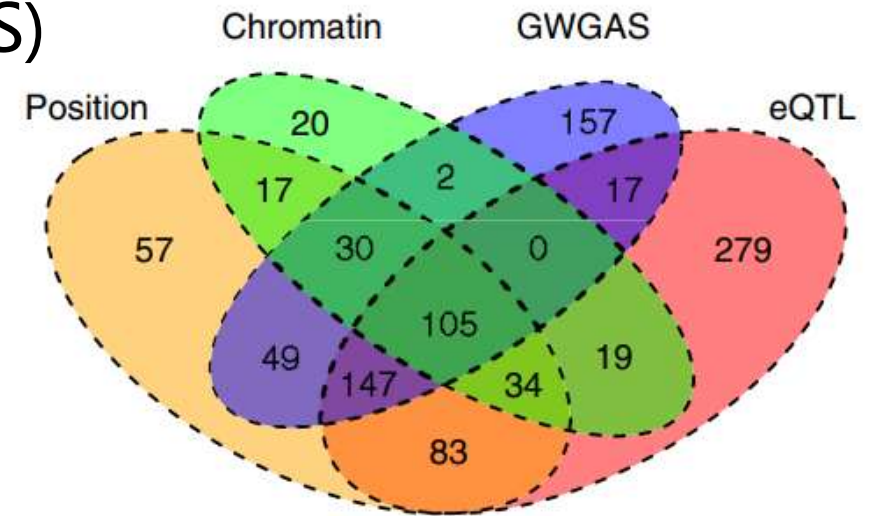


- 基因定位 (gene-mapping)
- 文章使用了三种FUMA软件包含的基因定位方法，positional gene-mapping将SNP比对到了522个基因上，eQTL比对到了684个基因，染色质相互作用图谱（chromatin interaction mapping）比对到了227个基因

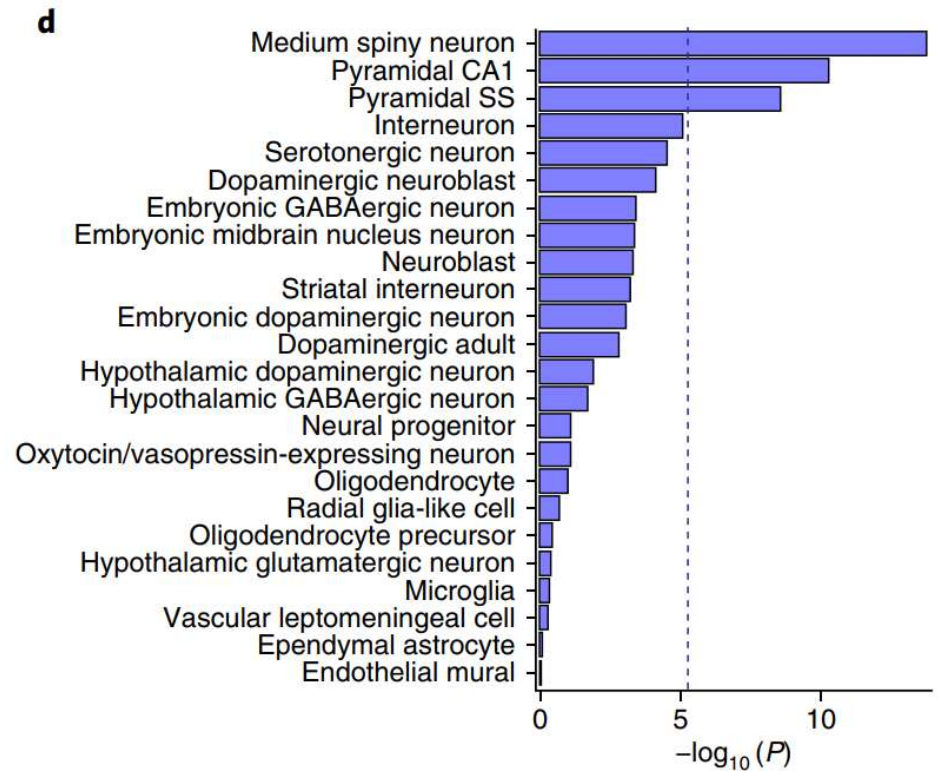
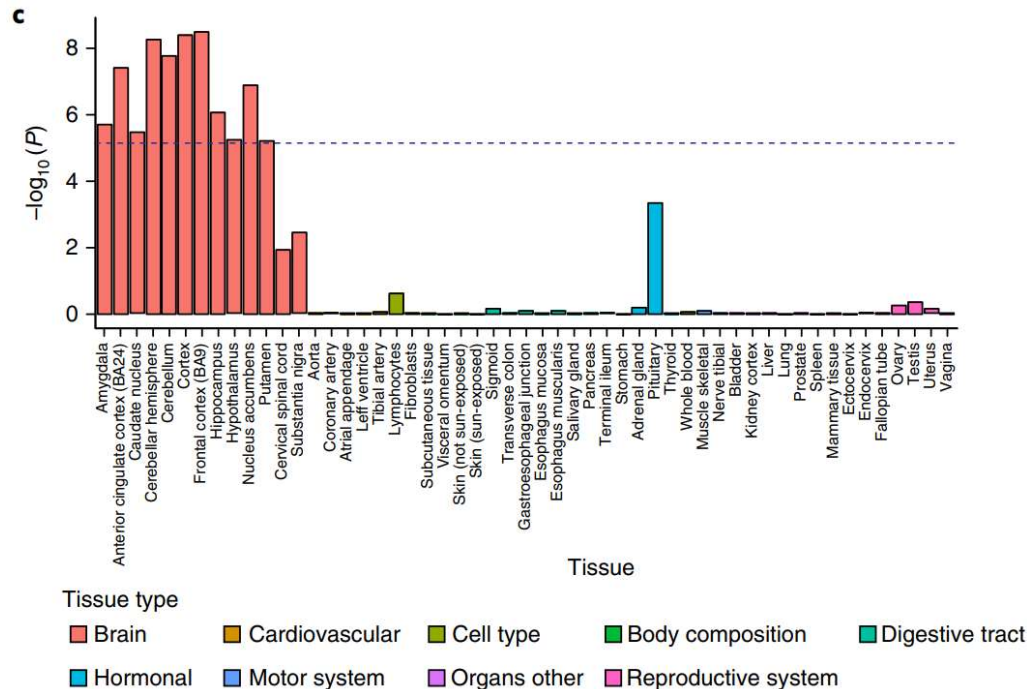
定位方法	positional gene-mapping	eQTL	染色质相互作用图谱
比对到的基因数量	522	684	227



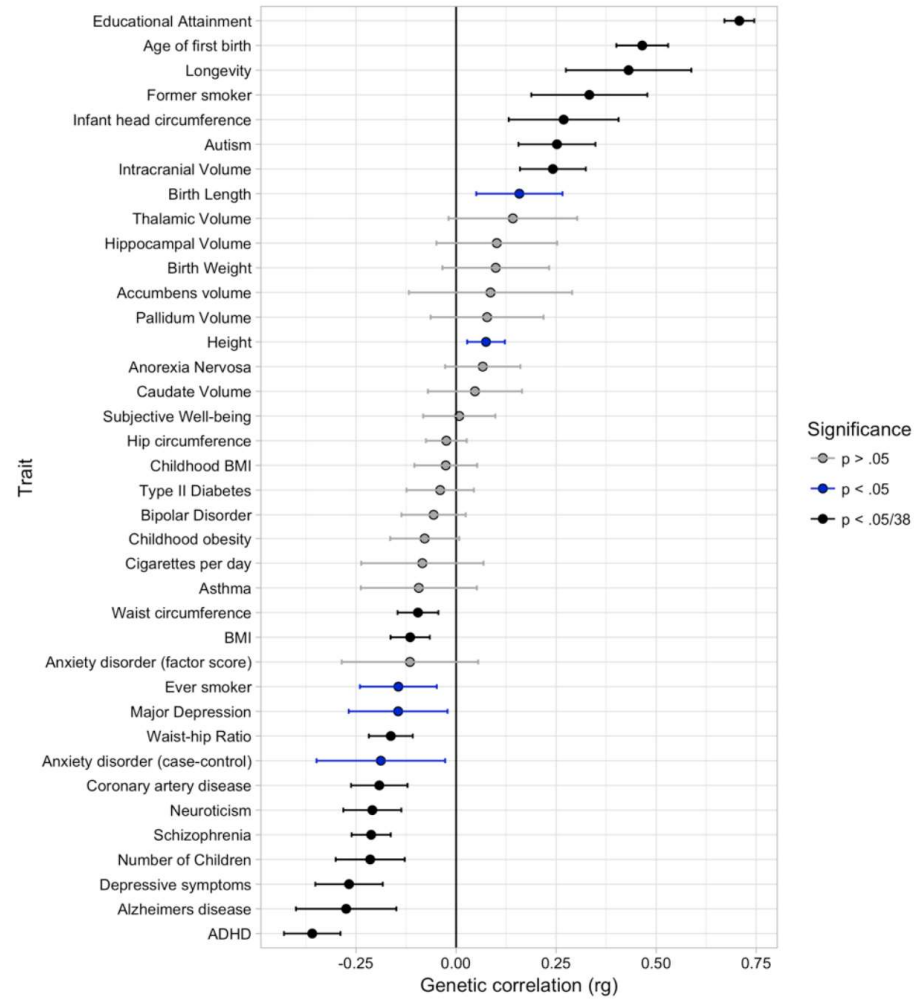
- 全基因组gene-based分析 (GWGAS)
- GWGAS找到了507个相关的基因



- SNP在各个脑区以及组织的表现

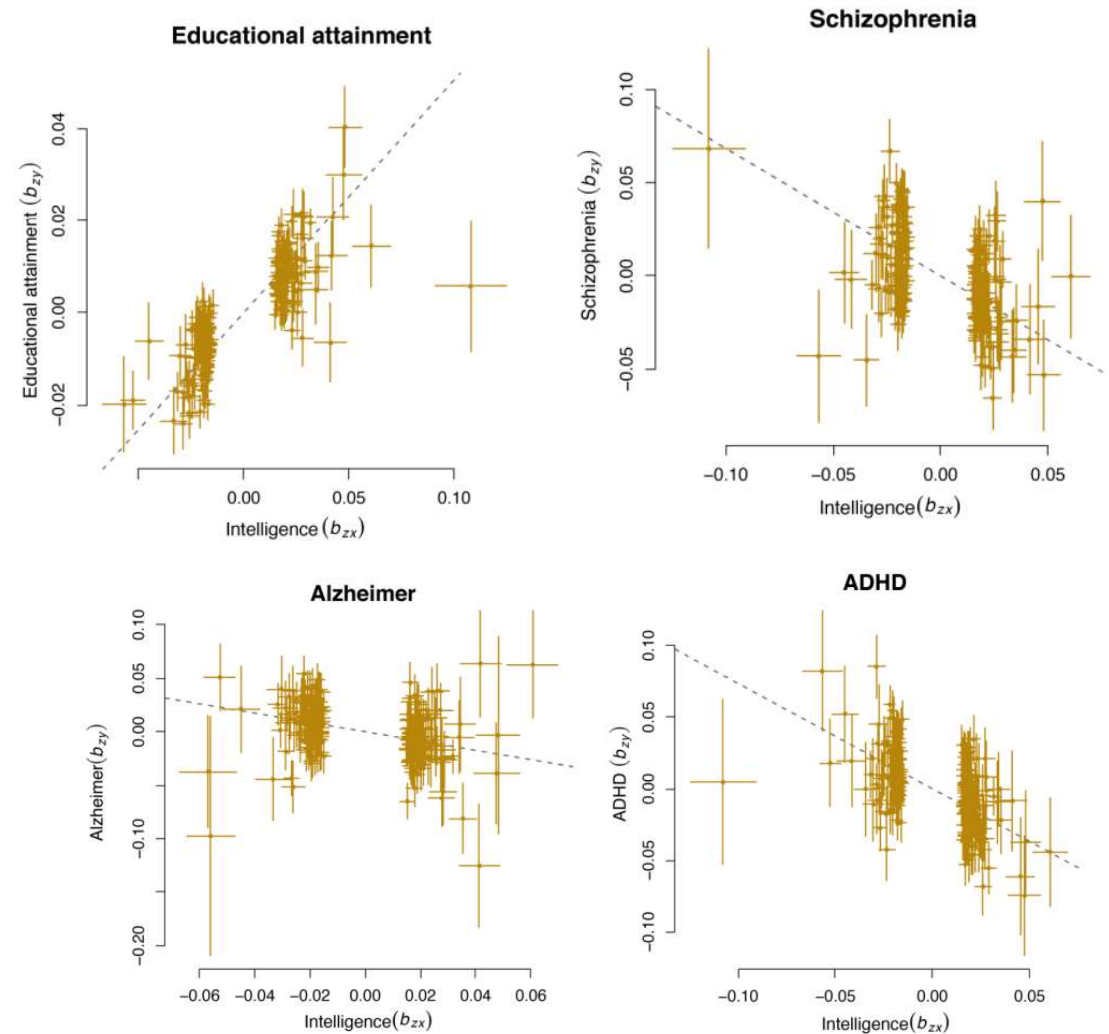


- 智力与其他性状的联系
- 在将智力和其他性状进行回归分析中，发现智力同ADHD、抑郁现象、阿尔兹海默症、精神分裂症以及寿命的关系



Supplementary Figure 10. Genetic correlations between intelligence and other traits

- 孟德尔随机化分析
- 文章还发现教育因素和认知能力具有双向因果性
- 发现智力对精神分裂症是一个保护因素，精神分裂症对智力也有不好的影响；高智力不易得ADHD和阿尔兹海默症，但更容易得自闭症



BGI 华大

THANKS

OMICS FOR ALL
基因科技造福人类