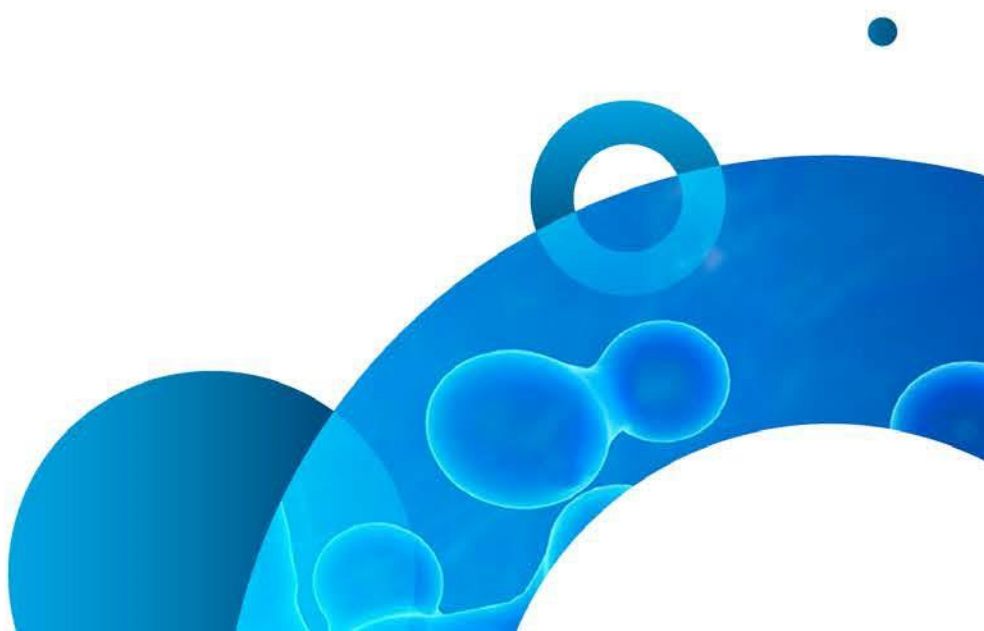
博淼生物
BIOMIAO BIOLOGICAL
-SINCE2009-
Your own Laboratory
您的专属实验室

多组学在分子病因学研究的设计策略

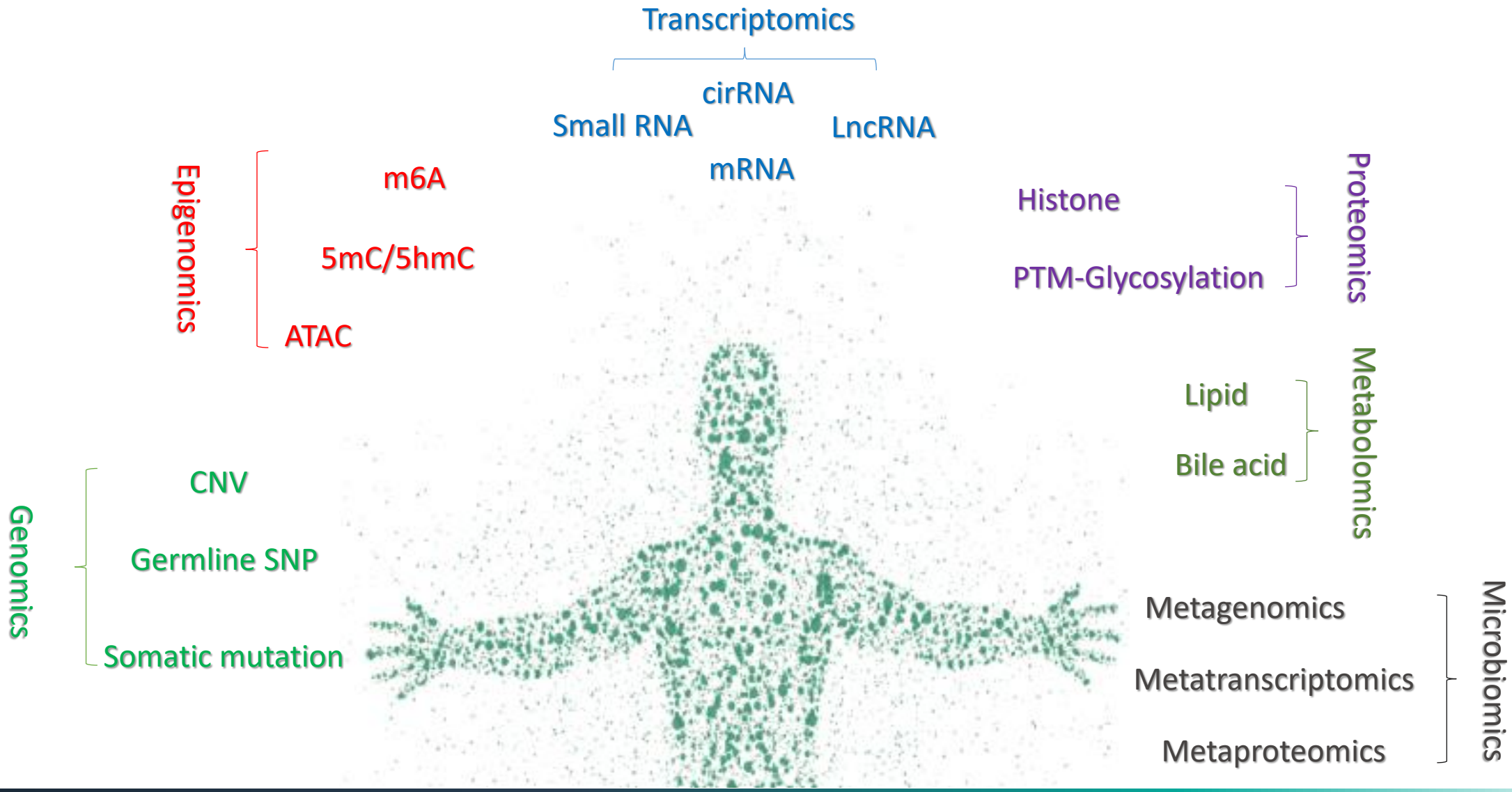
全国统一服务电话：400-6506-908
网址：www.biomiao.com
邮箱：marketing@biomiao.com
地址：北京市丰台区丰管路优橙创新中心B座3012-3015



目录
CONTENTS

Multi-omics研究理论体系概述

生命小宇宙


Transcriptomics
cirRNA
Small RNA
LncRNA
mRNA
Epigenomics
m6A
5mC/5hmC
ATAC
Proteomics
Histone
PTM-Glycosylation
Metabolomics
Lipid
Bile acid
Genomics
CNV
Germline SNP
Somatic mutation
Metagenomics
Metatranscriptomics
Metaproteomics
Microbiomics

Multi-omics Significance

博淼生物
BIO MIAO BIOLOGICAL
SINCE 2009

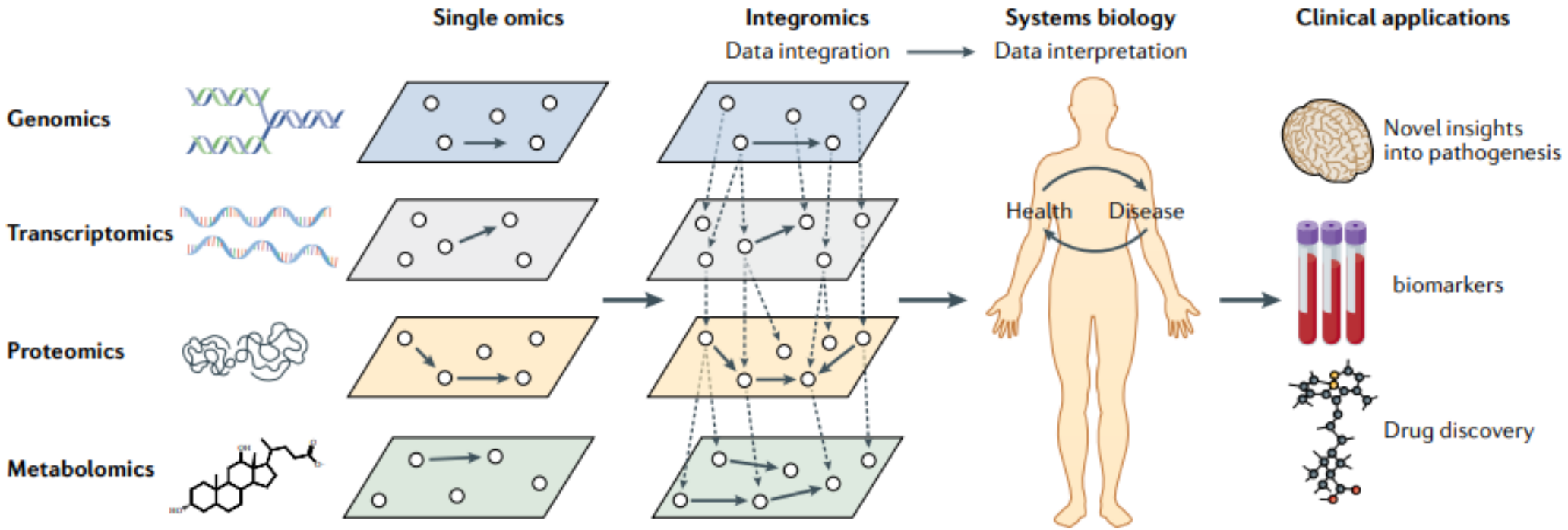
生物世界

- ✓ 基因组、转录组、代谢组…
- ✓ 单细胞多组学
- ✓ 空间多组学
- ✓ 微生物多组学

真实世界

- ✓ 环境暴露组
- ✓ 结局表型组

Multi-omics Significance



Integrate different omics data types Significance

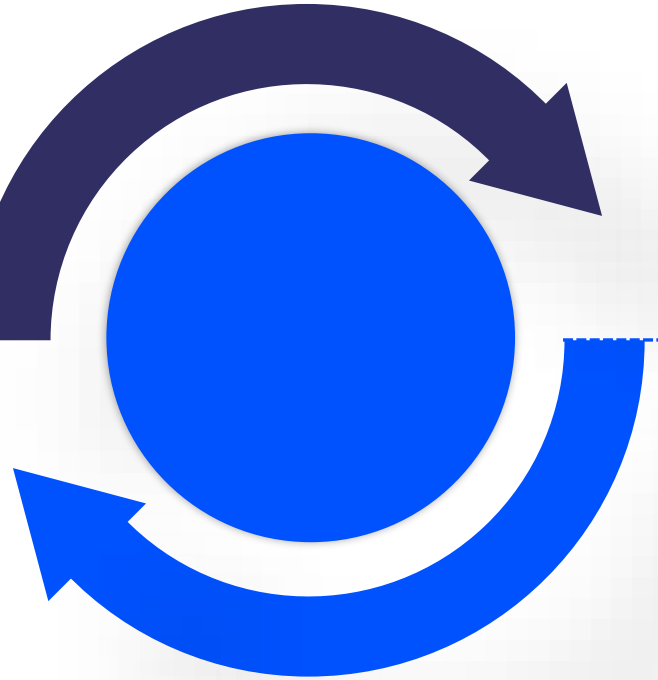
博淼生物 BIO MIAO BIOLOGICAL SINCE 2009

单一组学研究的瓶颈

✓科研结论侧重统计层面

✓生物调控效应微弱

✓标记物模型效果不佳



多组学联合研究的意义

✓构建多组学交互network

✓探讨因果生物学功能机制

✓精准筛选生物通路及关键功能分子

Integration of different omics data types is often used to elucidate potential causative changes that lead to disease, or the treatment targets, that can be then tested in further molecular studies。

Omics roles

基因组学（Genomics）

—What is possible happen

- ✓ 胚系细胞SNP、CNV、InDel等遗传变异
- ✓ 体细胞mutation、InDel等环境突变变异

转录组学（Transcriptomics）

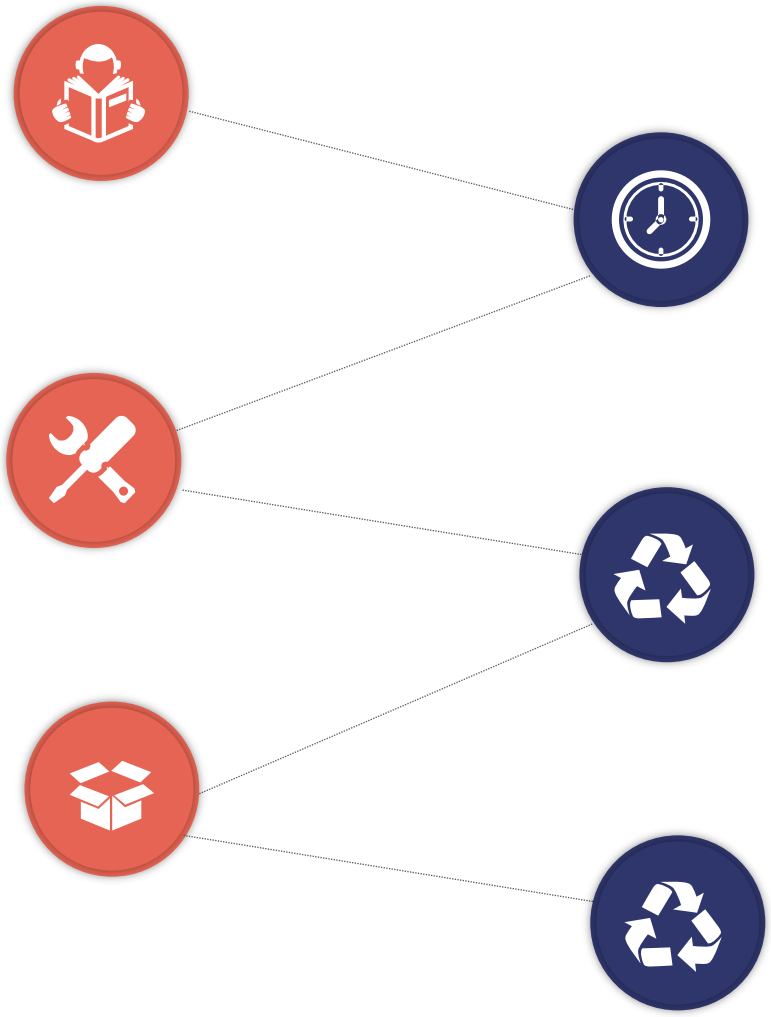
—What appears to be happening

- ✓ 转录本类型：mRNA、miRNA、lncRNA、cirRNA、piRNA等
- ✓ 转录定位：表达、融合、剪切、ceRNA等

代谢组学（Metabolomics）

—What has happened

- ✓ 脂质组、氨基酸类、胆汁酸类、神经递质类等等
- ✓ 宿主代谢物、微生物代谢物



表观基因组学（Epigenomics）

—What appears to be happening

- ✓ 基因组：ATAC、HiC
- ✓ DNA甲基化：5mC、5hmC、6mA等
- ✓ RNA甲基化：m6A等

蛋白质组学（Proteinomics）

—What is happening

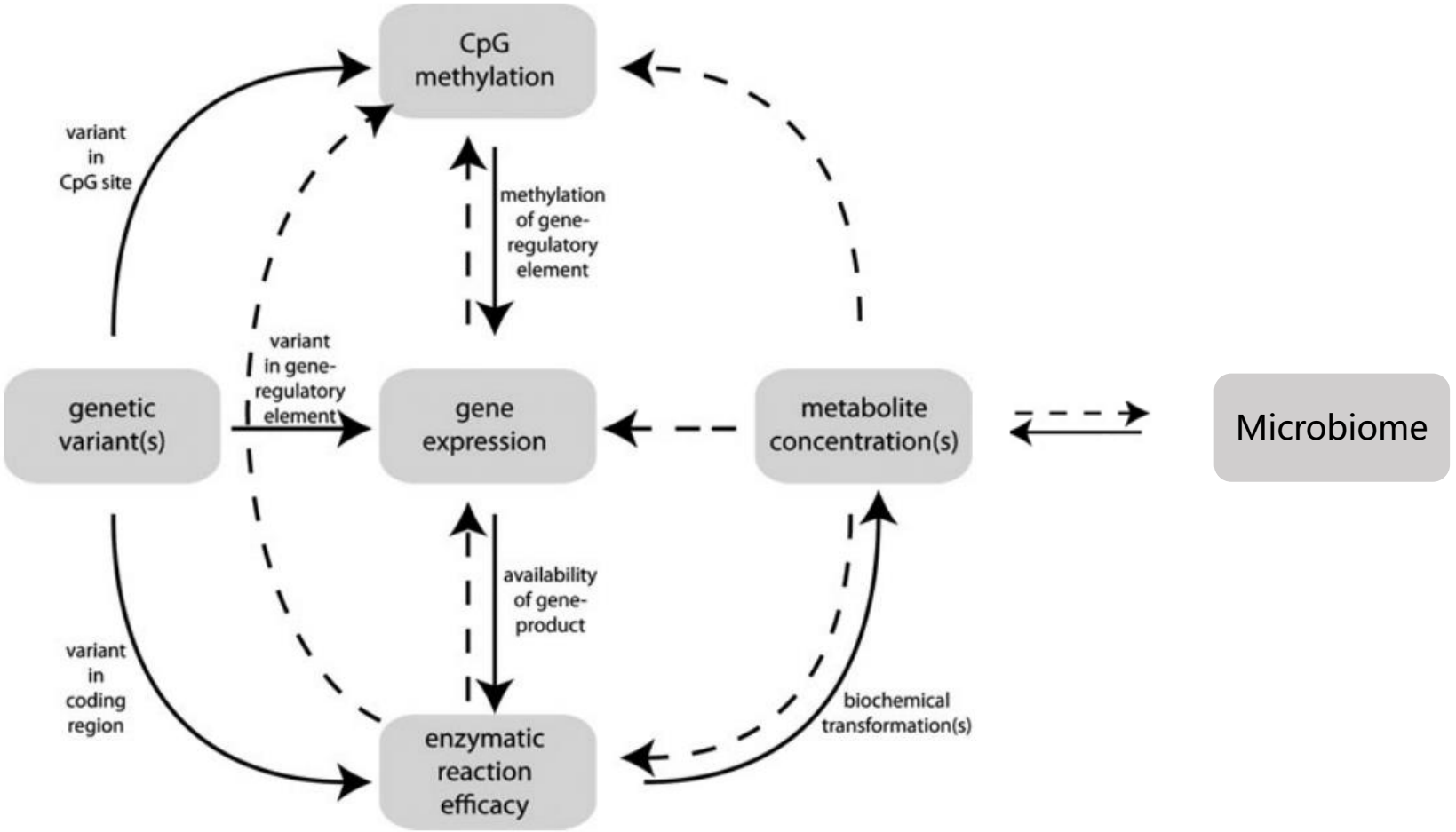
- ✓ 功能蛋白质：酶、转录因子等
- ✓ 修饰蛋白质：糖基化、乙酰化等

微生物组学（Microbiomics）

—The second omics：interact

- ✓ 宏基因组、宏转录组、宏蛋白质组、宏代谢组

Omics network



Omics network

博淼生物 BIO MIAO BIOLOGICAL SINCE 2009

An Expanded View of Complex Traits: From Polygenic to Omnigenic . Cell , 2017

颠 覆 一

多基因调控转向泛基因调控理论

颠 覆 二

核心基因与次要基因协同遗传效应

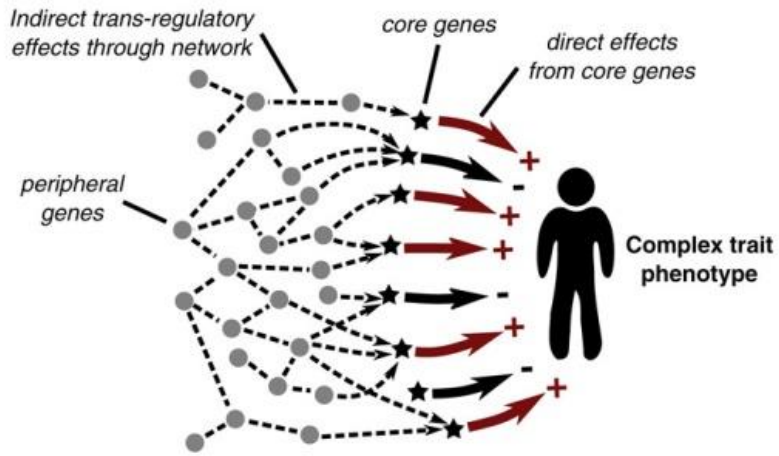
颠 覆 三

泛基因调控模型建立

Omics network

Question: Why is the architecture of complex traits dominated by huge numbers of small effect variants?

Approach: We built a quantitative phenotype model based on core gene expression



$$\mathrm{Var}(Y_i) = \underbrace{\sum_{j=1}^{M} \gamma_j^2 V_{j,\mathrm{cis}}}_{M \text{ core terms}} + \underbrace{\sum_{j=1}^{M} \gamma_j^2 V_{j,\mathrm{trans}}}_{M \text{ trans terms}} + \underbrace{\sum_{j=1}^{M} \sum_{k=1}^{j-1} 2\gamma_j \gamma_k C_{j,k}}_{M^2-M \text{ covariance terms}}$$

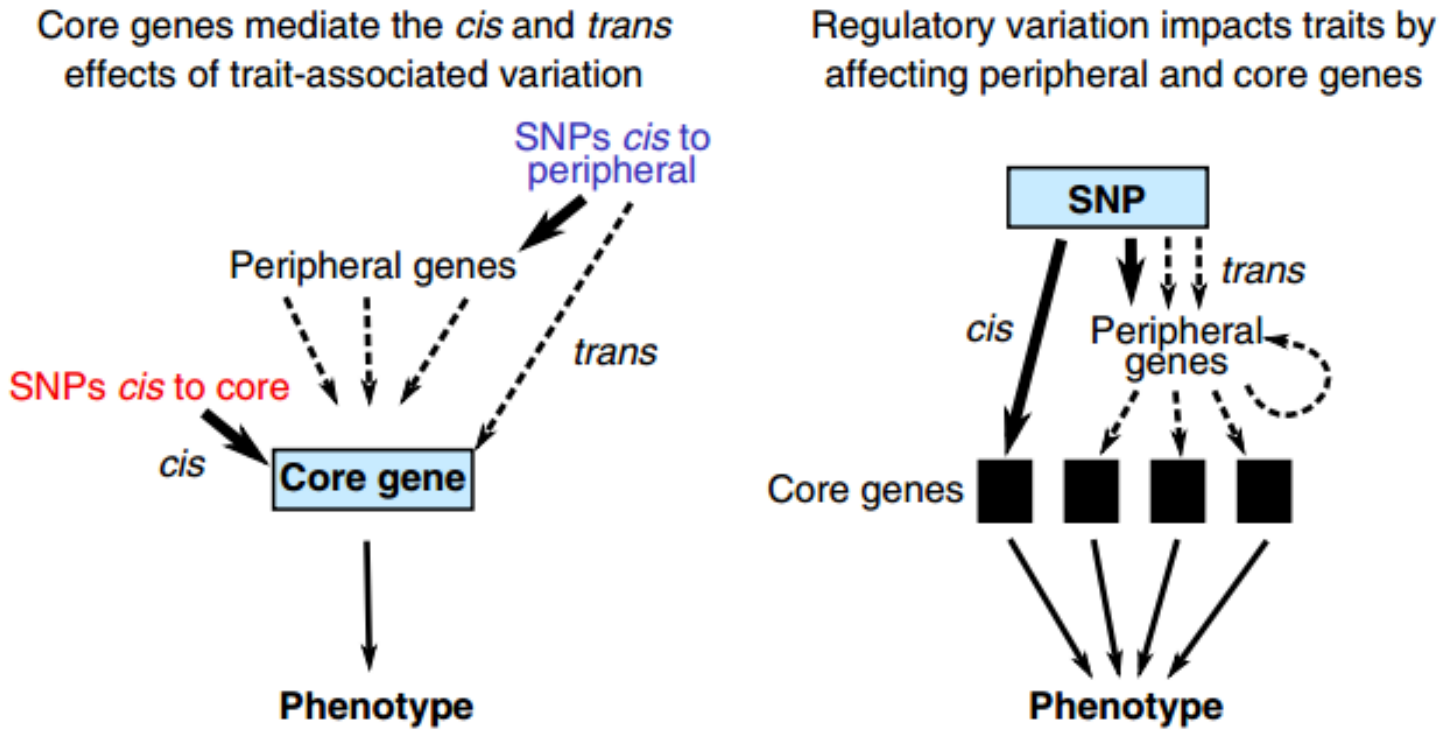
Conclusion: Most of the trait heritability is explained by many small *trans*-regulatory effects from peripheral genes

Highlights

- ✓ Propose a quantitative phenotype model based on core and peripheral genes
- ✓ Model is parameterized using data on cis and trans heritability of gene expression
- ✓ Analysis implies that heritability explained by trans-acting variants is at least 70%
- ✓ Co-regulation of core genes can further amplify the contribution of trans effects

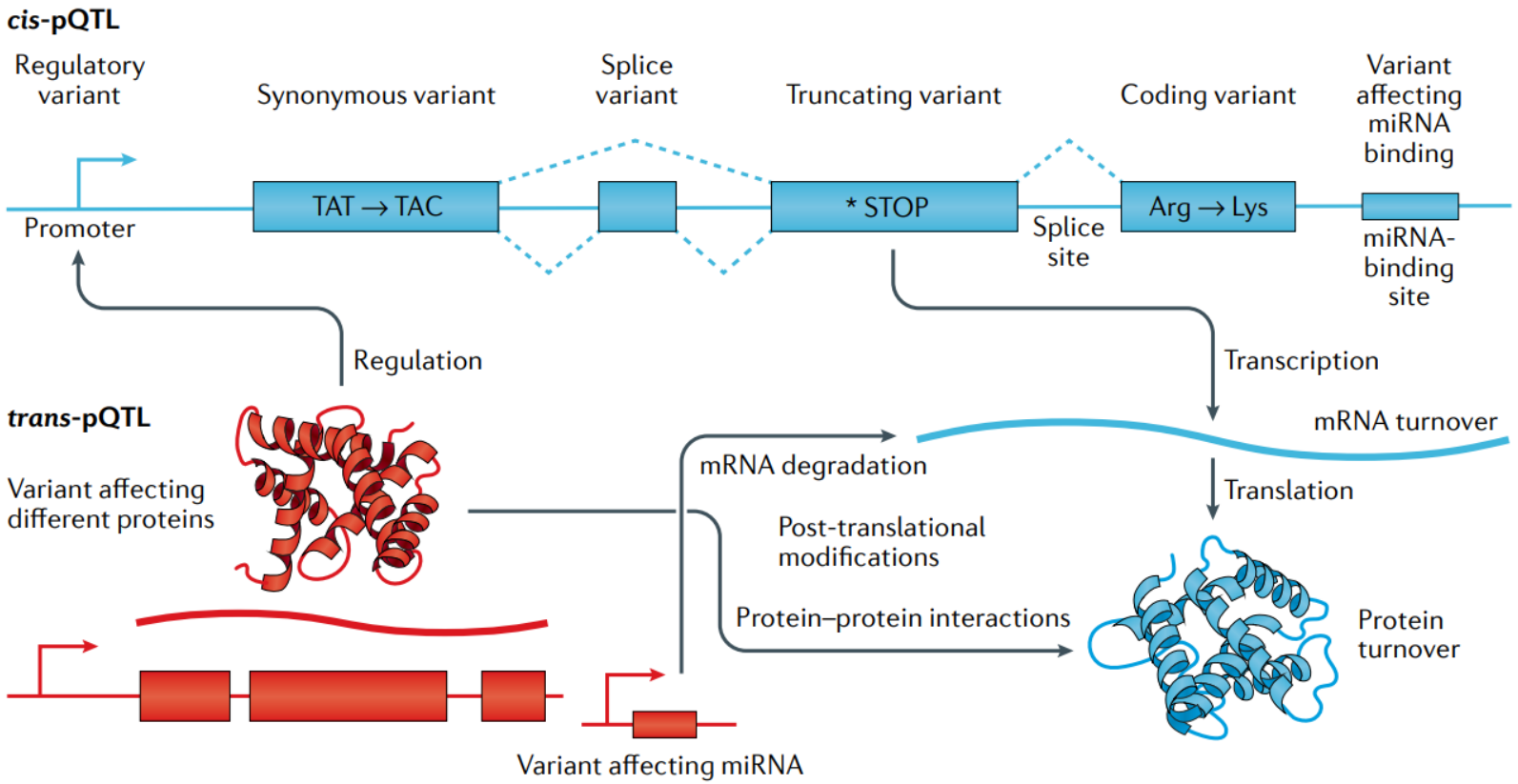
Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. Cell , 2019

Omics network



Causal Pathways for Variants Affecting a Trait through Core Genes

Omics network



Ways a genetic variant can lead to a pQTL

Bulk Omics data type

博淼生物 BIO MIAO BIOLOGICAL SINCE 2009

Epigenomics

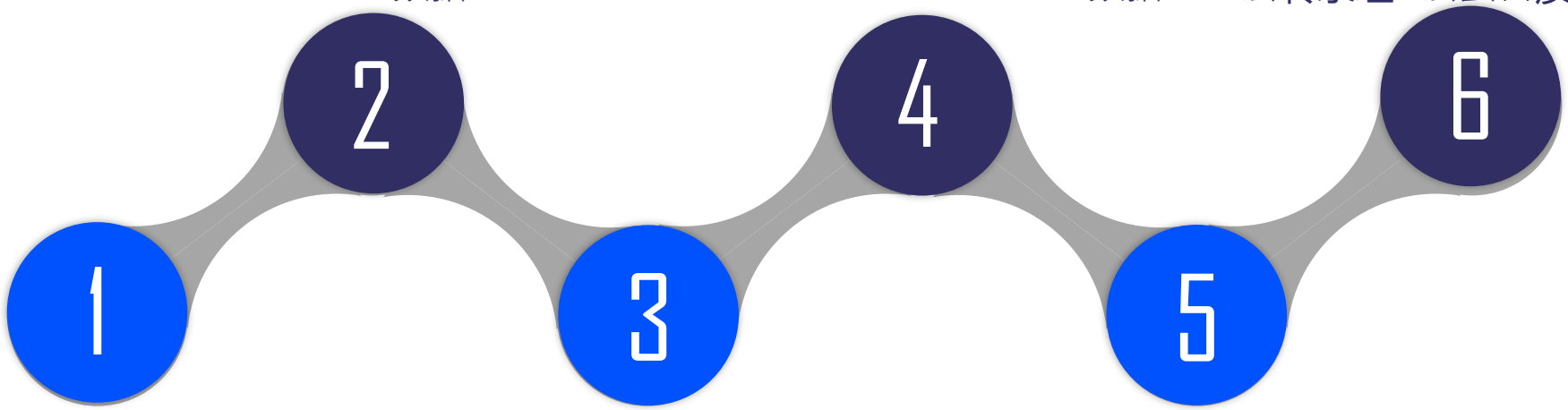
✓EWAS芯片数据
✓ATAC数据
✓HiC数据

Proteomics

✓Global proteome数据
✓Post-translational modifications数据

Microbiomics

✓16S数据
✓宏基因组数据
✓宏转录组&宏蛋白质组&宏代谢组



Genomics

✓GWAS芯片数据
✓WGS数据
✓WES数据

Transcriptomics

✓表达谱芯片数据
✓RNA-seq数据

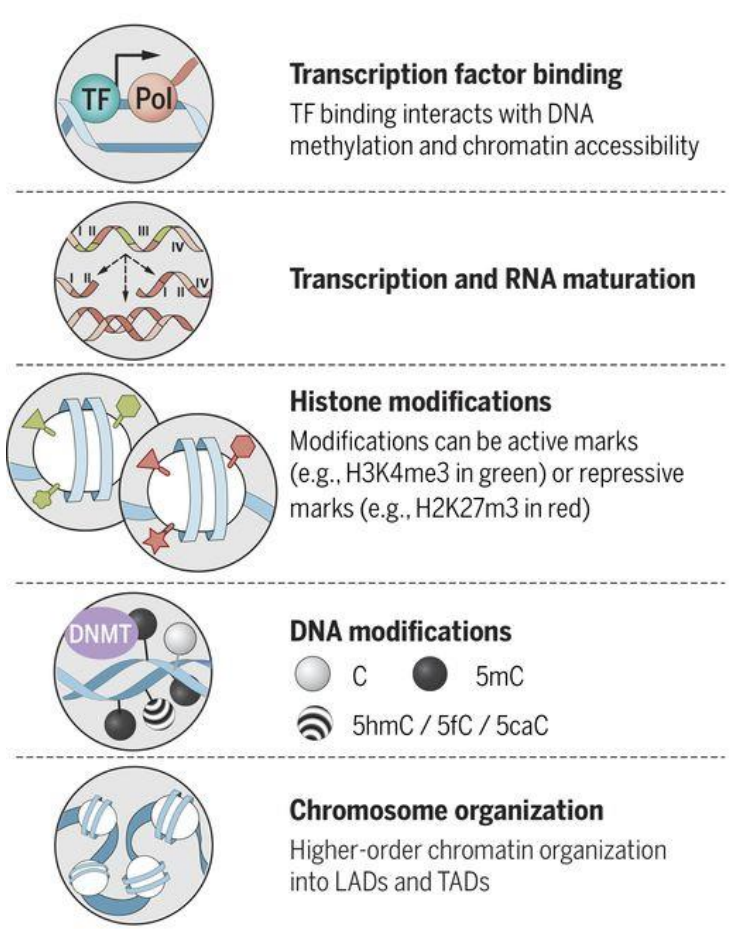
Metabolomics

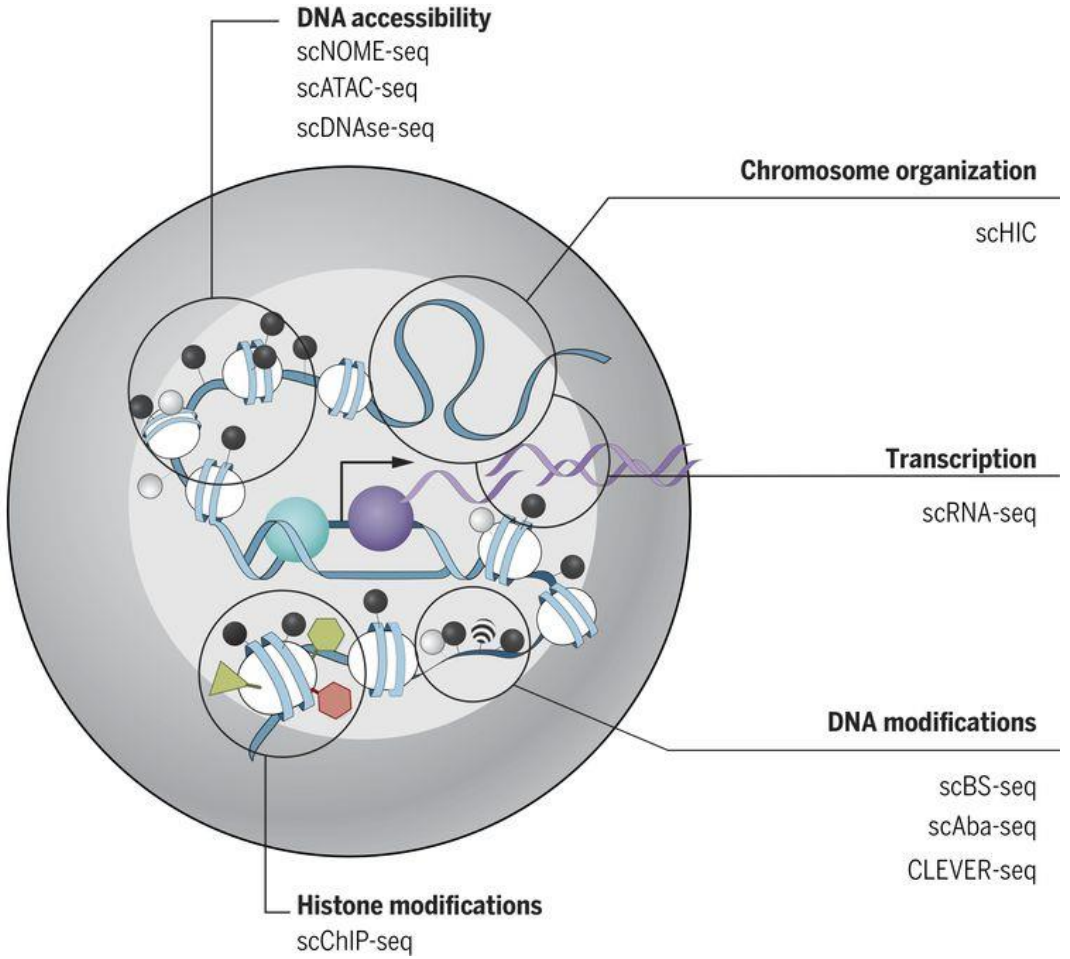
✓非靶向metabolites数据
✓靶向 metabolites数据

Data types for intergrative omics

Data type	Large-scale research efforts	Utility and advantages	Major caveats
Genetic variation	Many GWAS consortia, 1000 Genomes, gnomAD and UK Biobank	Unbiased source of genetic basis of disease and direct inference of causality	At least one step removed from the phenotype
Epigenetics	ENCODE and Roadmap Epigenomics Project	Functional impact and typically easy to infer causality	Not applicable for all phenotypes
Gene expression	GTEx and GEUVADIS	Inexpensive assay for an intermediate step towards the phenotype	Not applicable for all phenotypes
Proteomics and metabolomics	CPTAC, EDRN and Common Fund	Likely to be very close to the phenotype	Expensive and difficult to scale (proteomics)
Microbiome	Human Microbiome Project	Likely to be very close to the phenotype and measures a combination of genetic and environmental influences	Combination of genetic and environmental influences makes it difficult to infer the direction of causality

Single–cell Multi–omics





Single-cell Omics data types

博淼生物
BIO MIAO BIOLOGICAL
SINCE 2009

Genomics
✓scCNV-seq数据

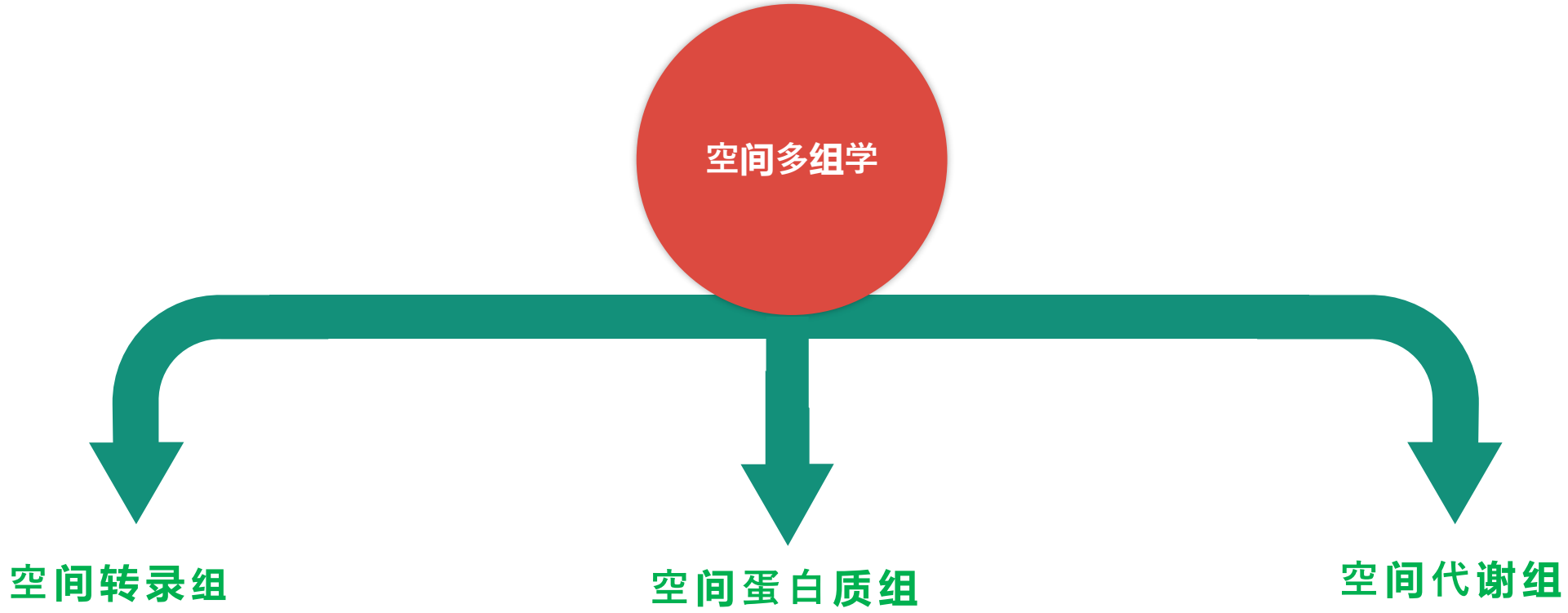
Epigenomics
✓ scATAC-seq数据

Transcriptomics
✓scRNA-seq数据

Proteomics
✓Sc-surface protein-seq数据

Immune Repertoire
✓scVDJ-seq数据

Spatial multi-omics data types



Multi-omics方案设计策略概述

The Genome first approach—样本遗传基线数据

Epigenomics
meGWAS-meQTL

Phenotype
GWAS-asscoiation

Transcriptomics
TGWAS-eQTL、sQTL、roQTL
mRNA/miRNA
/LncRNA/CirRNA/Splice

Single cell RNA-seq
Sc-eQTL

Protein&Glycomics
pGWAS-pQTL
glycGWAS-glycQTL

Metabolome&Microbiomics
mGWAS-mQTL、miGWAS-miQTL

The Genome first approach

科研主旨

关联分析研究

✓QTL分析Cis_与trans分析遗传调控网络network构建

✓精准筛选重要驱动功能基因

✓以PRS为基础的多组学标记物联合模型优化构建

因果效应分析研究

✓MR分析、COLOC分析、Mediation分析、Condition分析

✓探讨不同组学与表型组、暴露因素与表型组、不同表型组间因果关系

The Genome first approach

博淼生物 BIO MIAO BIOLOGICAL SINCE 2009

研究目的

研究手段

研究定位的转变

单一GWAS研究的局限性

✓侧重相关性结论的挖掘

✓遗传模型对于复杂表型的片面性

✓高质量队列样本的长周期

✓非功能区域易感SNP位点的分子调控功能机制挑战

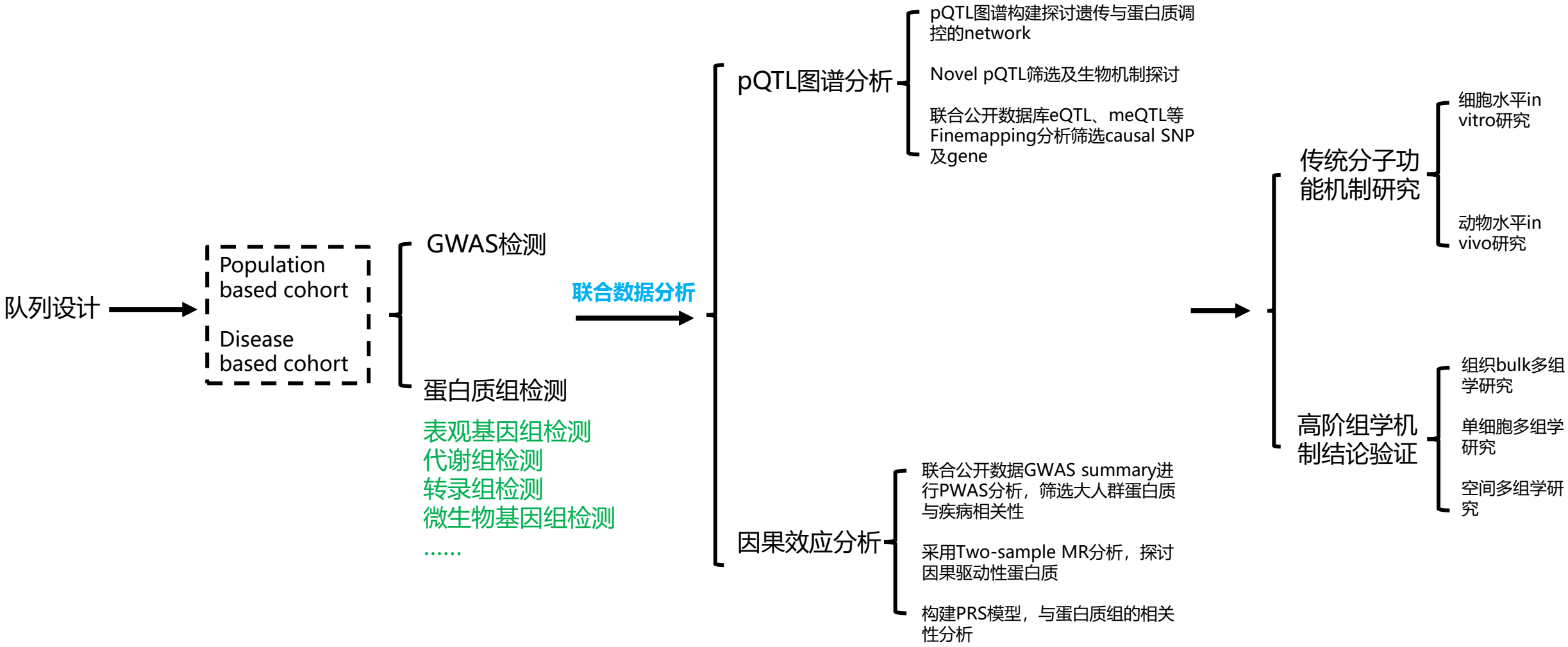
GWAS多组学研究的广阔性

✓ 构建不同纬度的QTL调控network以及易感SNP causal结论的挖掘

✓ 不同多组学TWAS、MeWAS、PWAS等融合分析及MR随机化因果推断分析

✓ 筛选与表型有关的重要生物学通路及其中关键调控功能分子

GWAS&Multi-omics approach研究策略——研究路径概述



GWAS&Multi-omics approach研究策略——研究路径概述

pQTL科研应用场景

发现新的药物靶点

确定与特定蛋白质相关的遗传变异，这些蛋白质可能为药物研发提供新的靶标

发现药物的新治疗适应症

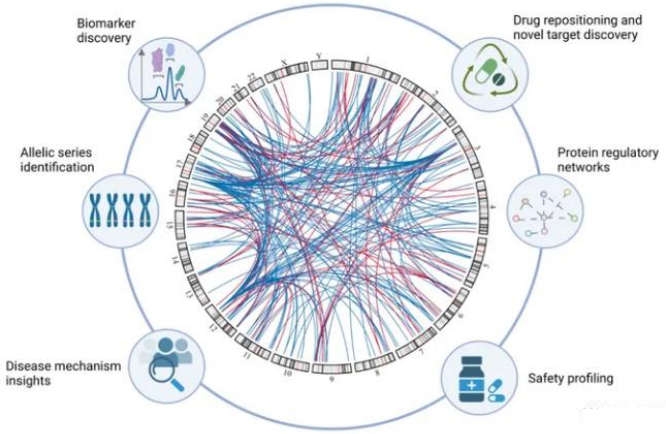
通过观察特定药物对蛋白质水平的影响，探索药物在不同疾病领域的潜在治疗适应症

识别临床相关的生物标志物

揭示与蛋白质表达水平相关的遗传变异，这些变异可能与特定疾病的风险或疾病的诊断、预测和治疗响应相关

发现疾病的致病机制

通过识别与蛋白质表达水平变化相关的遗传变异，来揭示与疾病发病机制相关的生物学过程



Multi-omics strategy to epidemiology

动物模型验证性研究

研究优势：环境因素可控、组织容易获取、精准表型设置等等

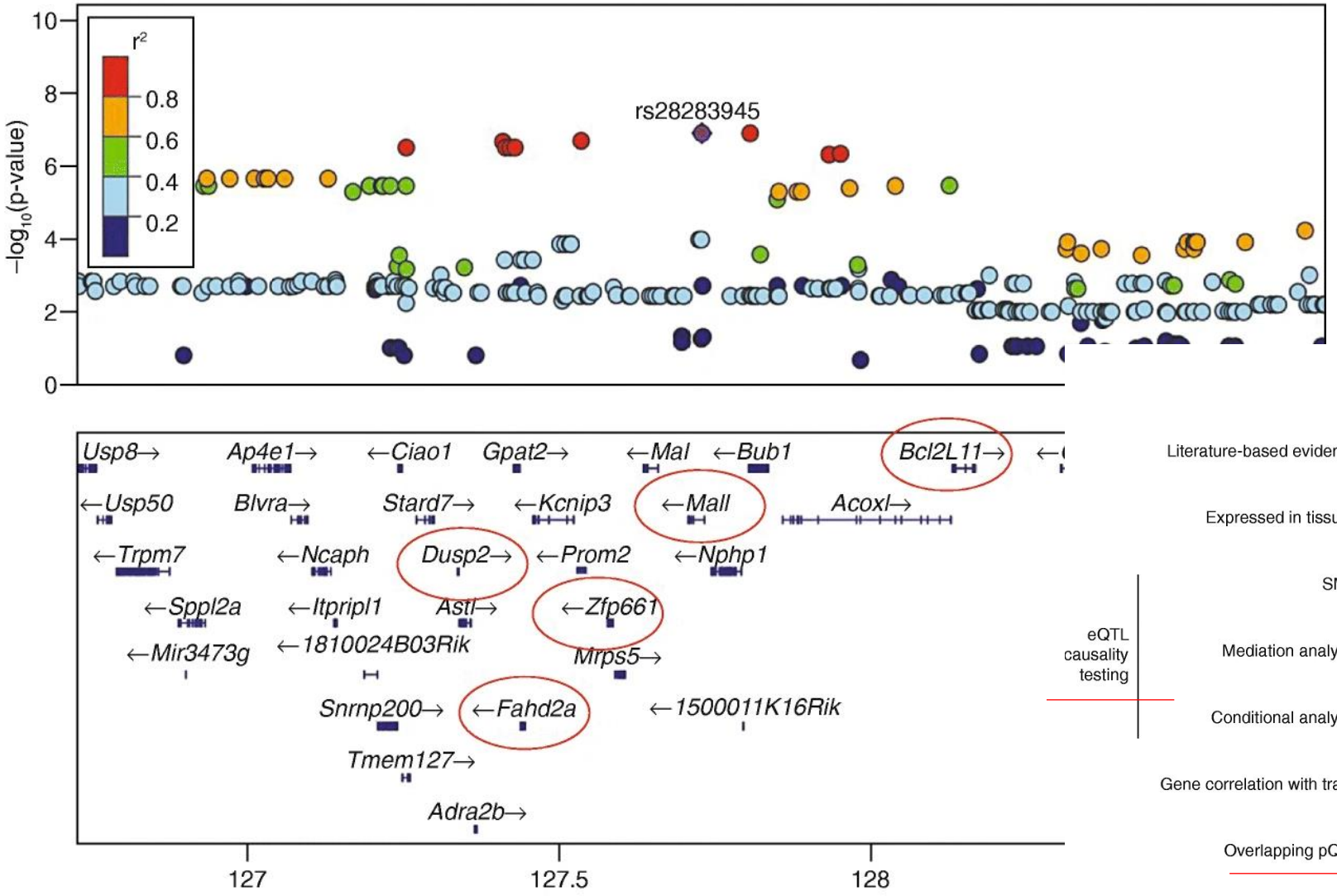
研究劣势：无法准确反应人体生物学通路、疾病表型不一致等等

单细胞组学验证性研究

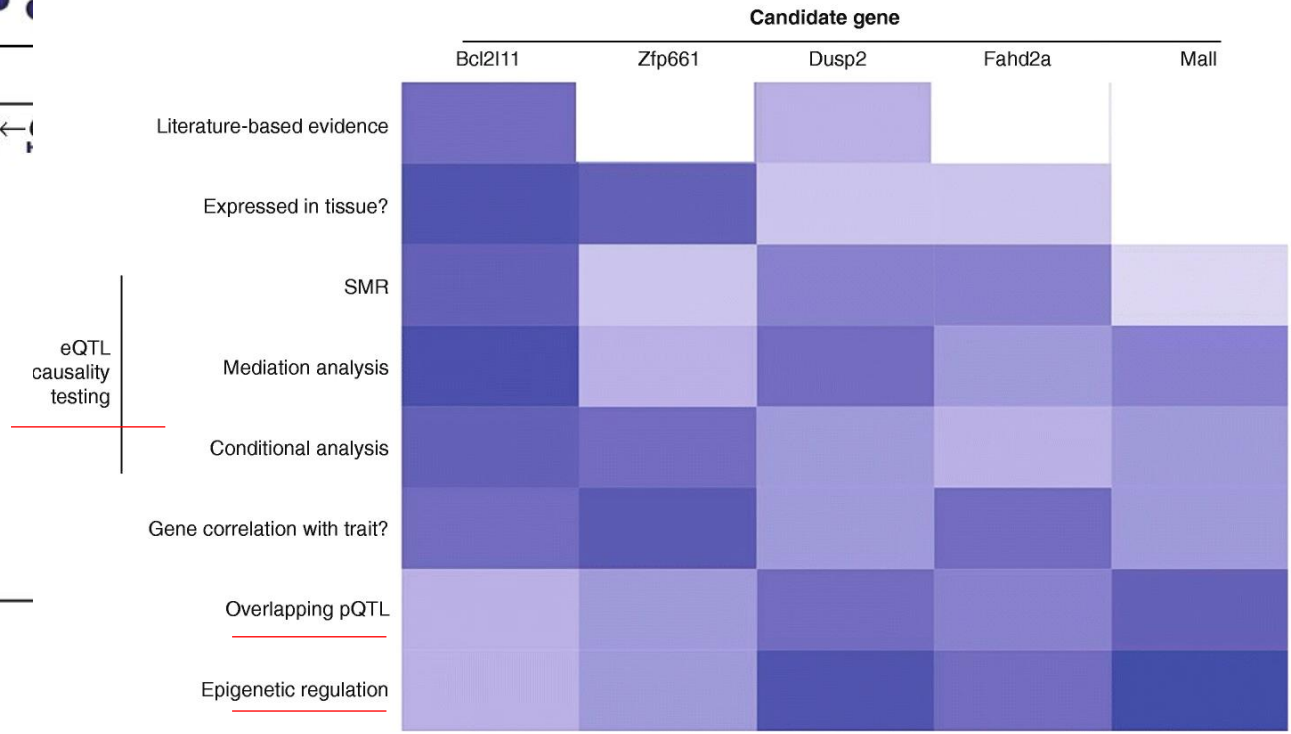
研究优势：在单细胞水平验证原位细胞图谱及对应的关键功能基因、调控通路等

研究劣势：无法系统验证对于细胞的分子功能调控机制

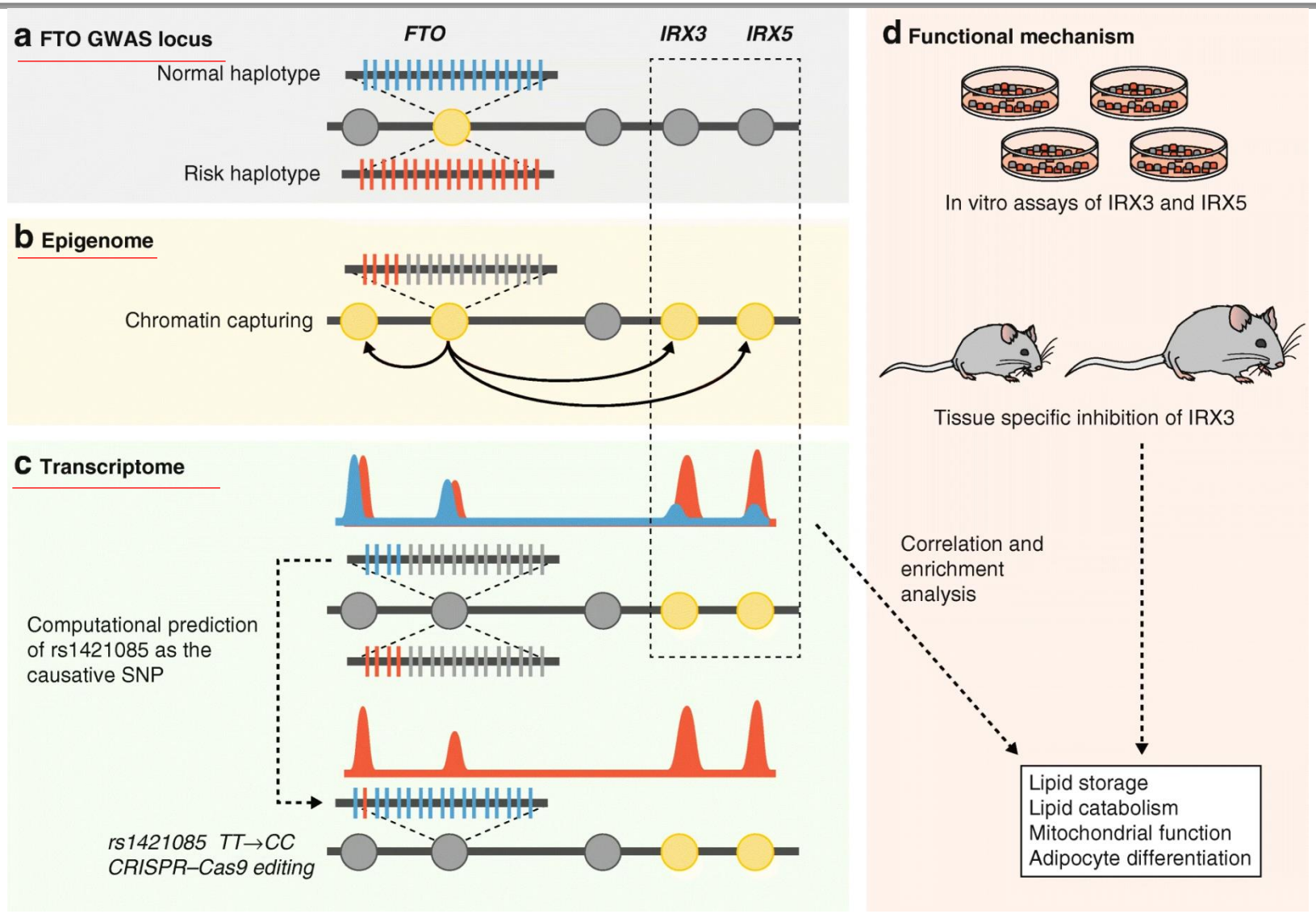
The Genome first approach—优选易感位点及因果机制分析



Narrowing causal mechanisms



The Genome first approach—筛选功能基因及功能验证



FTO obesity variant circuitry and adipocyte browning in humans. N Engl J Med.

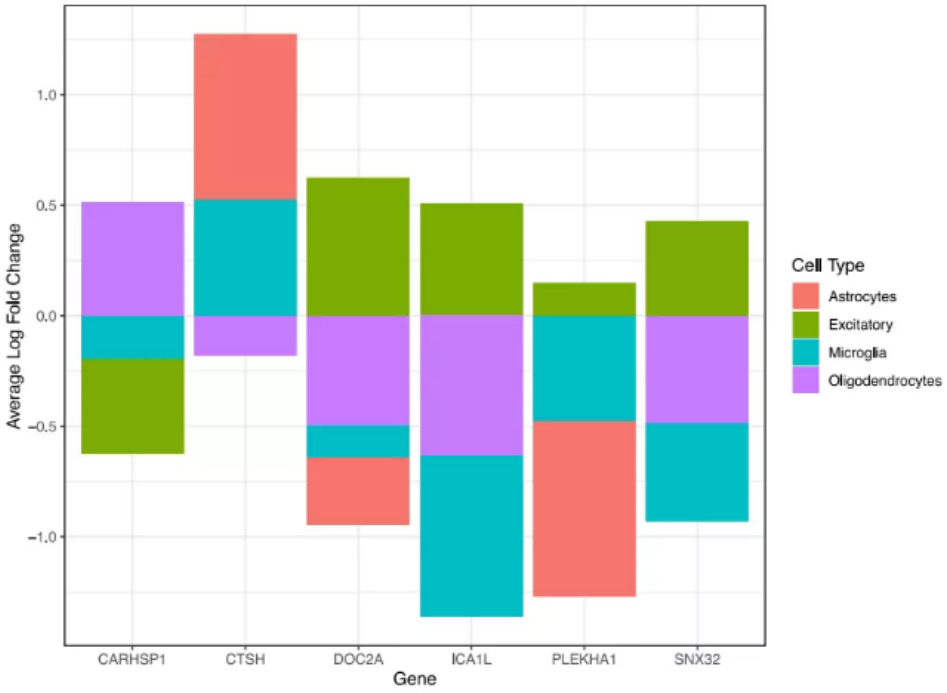
The Genome first approach—联合蛋白质组PWAS及MR因果效应分析

博淼生物 BIO MIAO BIOLOGICAL SINCE 2009

Table 3 | Summary of the 11 AD PWAS-significant genes with evidence for being consistent with a causal role in AD

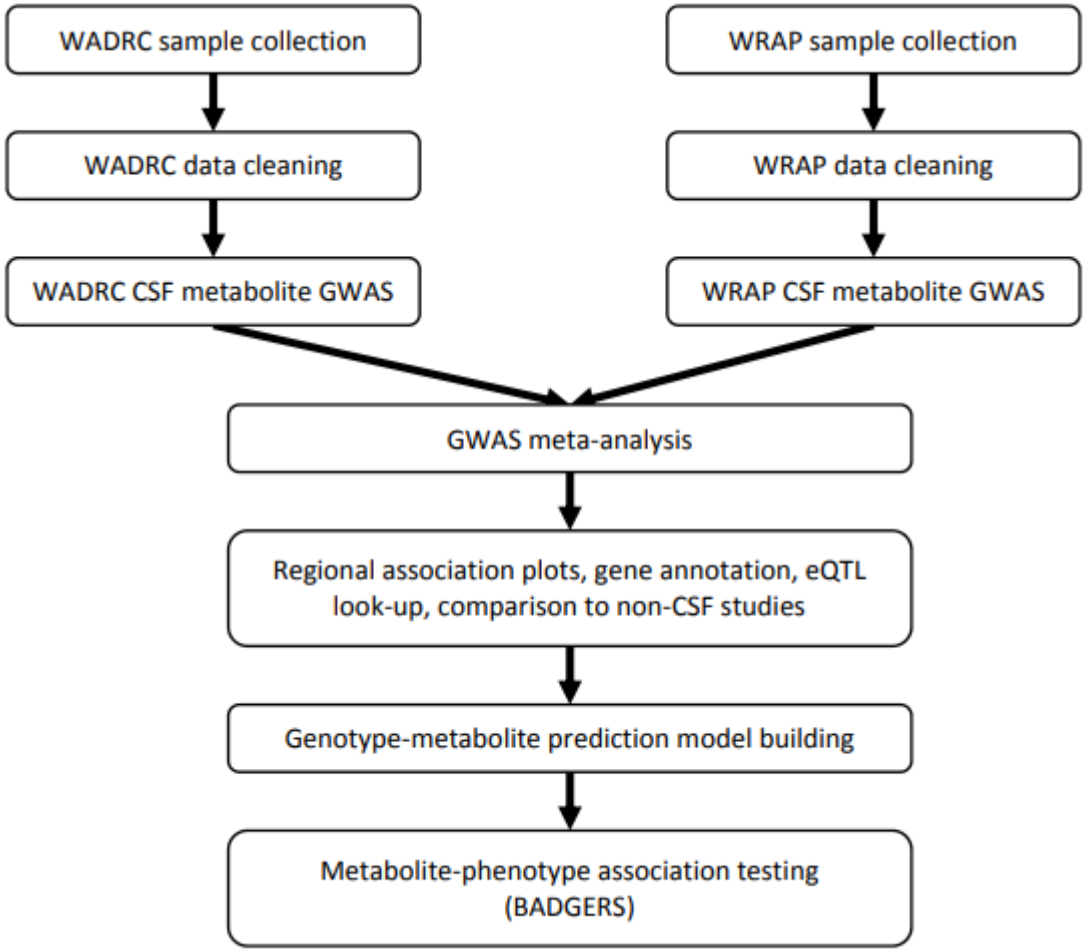
			Discovery PWAS	Confirmation PWAS	Evidence for causality		TWAS significant	New gene
	Gene	Chromosome			COLOC	SMR		
1	*CTSH*	15	Significant	Replicated	Yes	Yes	Suggestive	Yes
2	*DOC2A*	16	Significant	Replicated	Yes	Yes	N/A	Yes
3	*ICA1L*	2	Significant	Replicated	Yes	Yes	No	Yes
4	*LACTB*	15	Significant	Replicated	Yes	Yes	Suggestive	No
5	*SNX32*	11	Significant	Replicated	Yes	Yes	Yes	Yes
6	*ACE*	17	Significant	Replicated	Yes	No	Yes	Yes
7	*RTFDC1*	20	Significant	Replicated	Yes	No	Suggestive	No
8	*CARHSP1*	16	Significant	Replicated	No	Yes	Yes	Yes
9	*STX6*	1	Significant	Replicated	No	Yes	Yes	Yes
10	*STX4*[a]	16	Significant	-	Yes	Yes	Yes	No
11	*PLEKHA1*[a]	10	Significant	-	Yes	Yes	N/A	Yes

[a]Proteins not found in the confirmation PWAS. N/A refers to genes that did not have significant heritability estimates to be included in the TWAS of AD. The full results for the TWAS are shown in Supplementary Tables 17 and 18. 'Suggestive' in the 'TWAS significant' column refers to genes with 0.05 < TWAS nominal $P < 0.1$. New gene refers to genes not within a 1-Mb window of SNPs with $P < 5 \times 10^{-8}$ identified in Jansen et al. AD GWAS'.

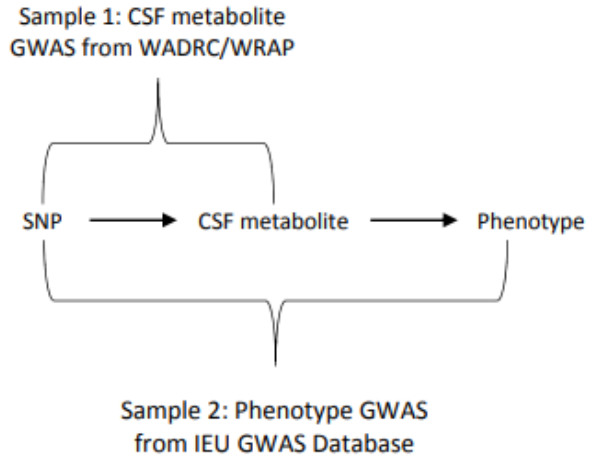


Integrating human brain proteomes with genome-wide association data implicates new proteins in Alzheimer's disease pathogenesis *Nature Genetics*: February 2021

The Genome first approach—联合代谢组MWAS分析



Supplementary Figure 1. Overview of MWAS

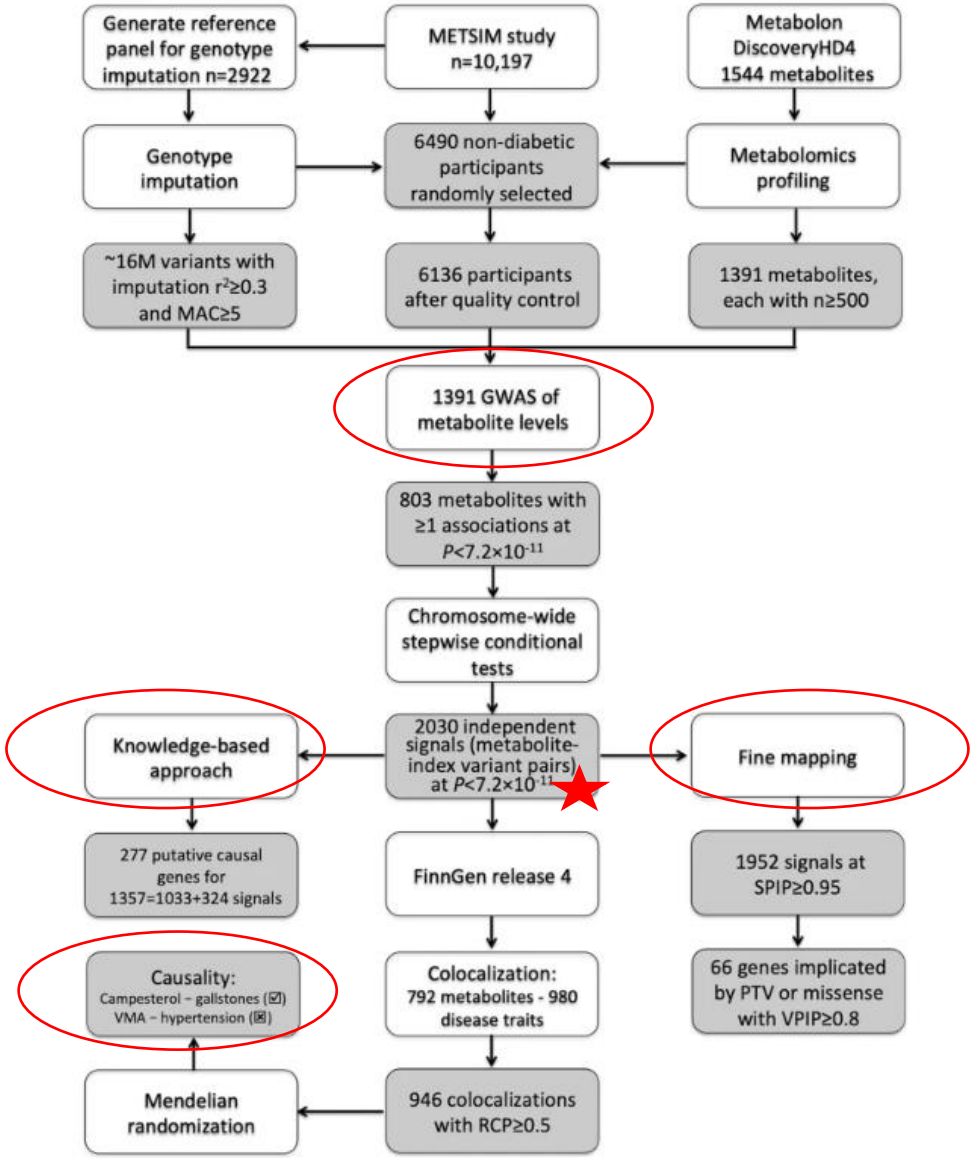


Supplementary Figure 29. Two-sample Mendelian Randomization model set-up

The general outline of our MWAS approach is as follows: (1) identify single-nucleotide polymorphism (SNP)-metabolite associations; (2) build metabolite prediction models using genotypes; (3) test metabolite–phenotype associations with publicly available GWAS summary statistics. Step 1 is used to demonstrate that SNP-metabolite associations do indeed exist and thus justify the building of metabolite prediction models in step 2 on a cohort where both genotype and metabolite data are present. Step 3 uses the prediction models in conjunction with publicly available GWAS summary statistics on neurological and psychiatric phenotypes to test metabolite–phenotype associations. The advantage of MWAS is that it allows for this metabolite–phenotype association testing to occur in GWAS datasets where only genotypes and phenotypes (not metabolites) were originally measured.

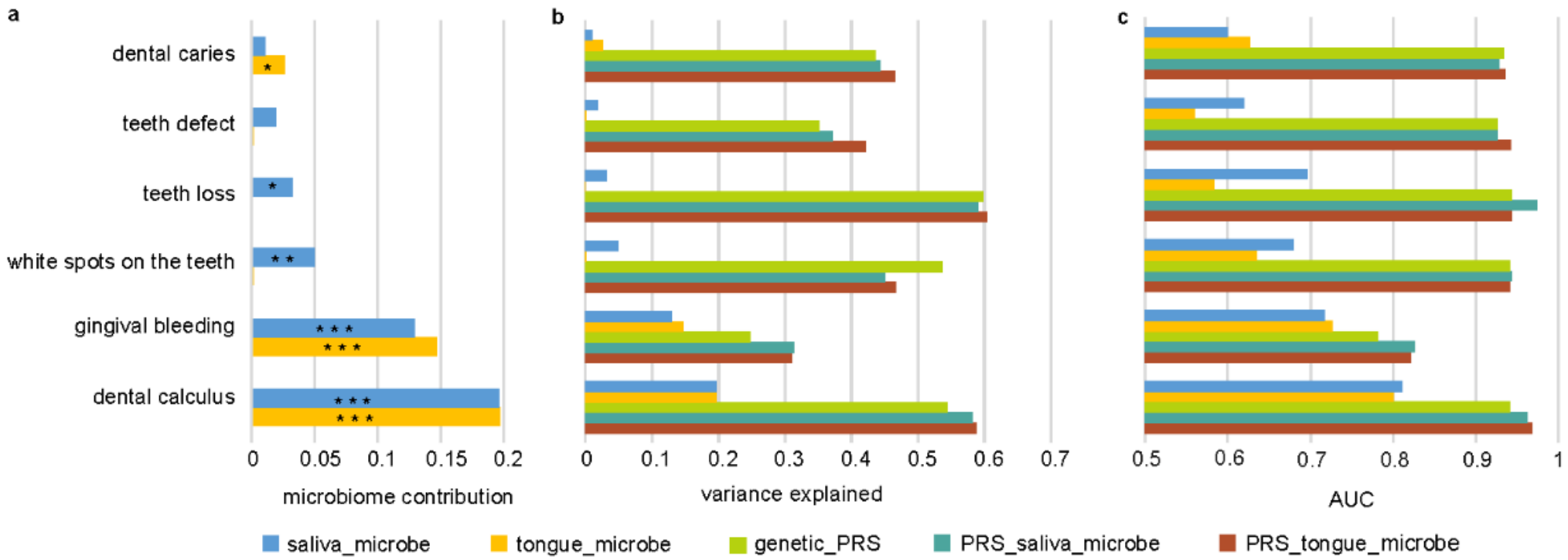
Cerebrospinal fluid metabolomics identifies 19 brain-related phenotype associations *Communications biology* 2021

The Genome first approach—mQTL图谱构建及分子调控机制挖掘



Genome-wide association studies of metabolites in Finnish men identify disease-relevant loci. *Nature communications* 2022

The Genome first approach—PRS联合多组学模型构建



图a：微生物对6种口腔问题的解释度；图b：微生物，基因PRS以及两者结合对6种口腔问题的解释度；图c：微生物，基因PRS以及两者结合对6种口腔问题风险的预测效能（AUC）

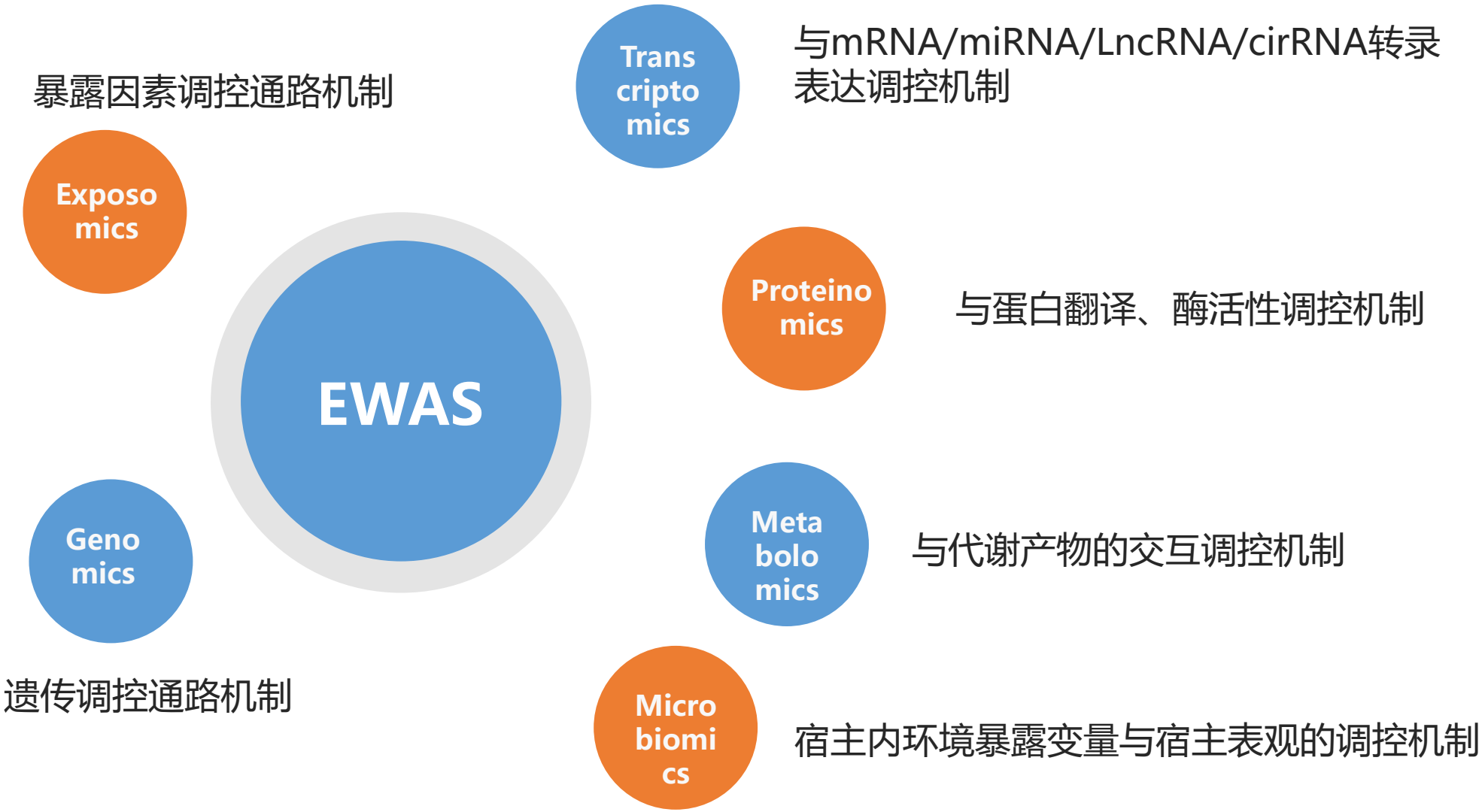
Metagenome-genome-wide association studies reveal human genetic impact on the oral microbiome. Cell Discovery (2021)

博淼基因组核心技术服务项目

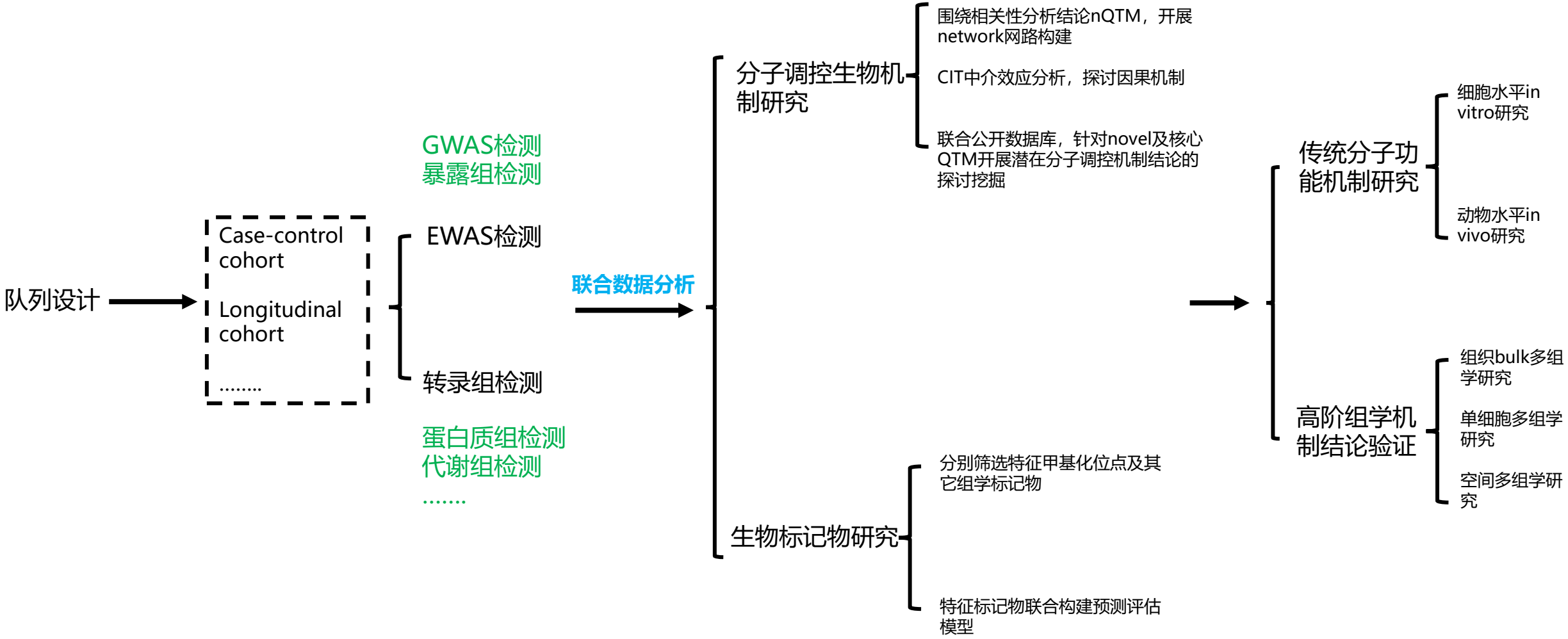
博淼生物
BIO MIAO BIOLOGICAL
SINCE 2009

GWAS芯片全基因组分型	Illumina 中华\ASAMD\CGA等
Multi-PCR NGS 靶向SNP分型	适合100-1000个SNP位点高通量靶向SNP分型
Massarray 靶向SNP分型	适合10-100个SNP位点中通量靶向SNP分型
KASP/Taqman靶向SNP分型	适合1-10个SNP位点低通量靶向SNP分型

The Epigenomics first approach—功能调控中介数据



The Epigenomics first approach研究策略—研究路径概述

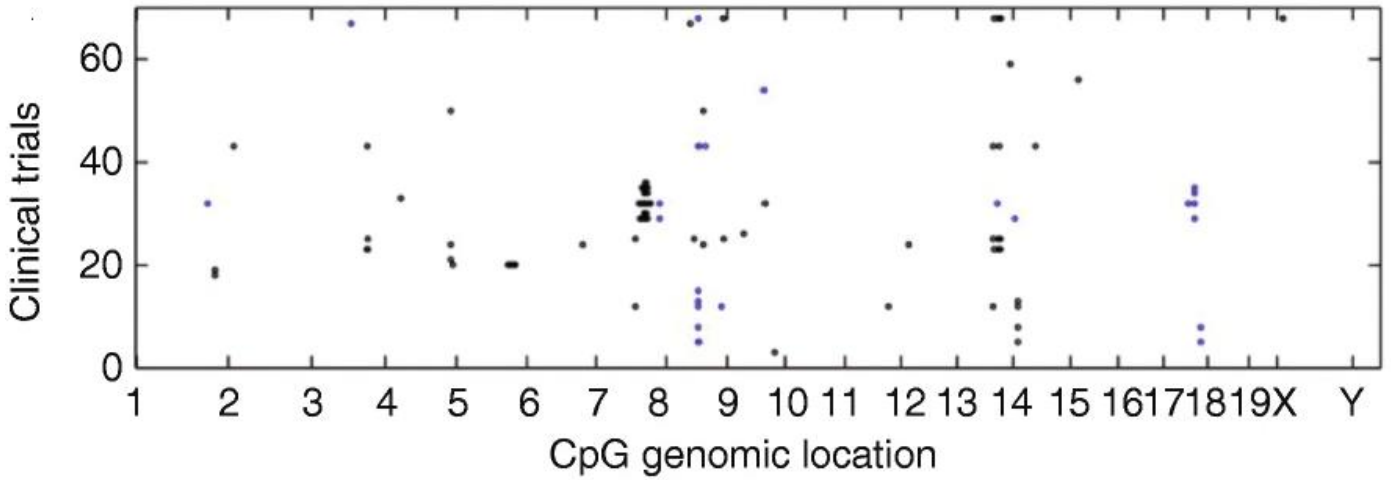


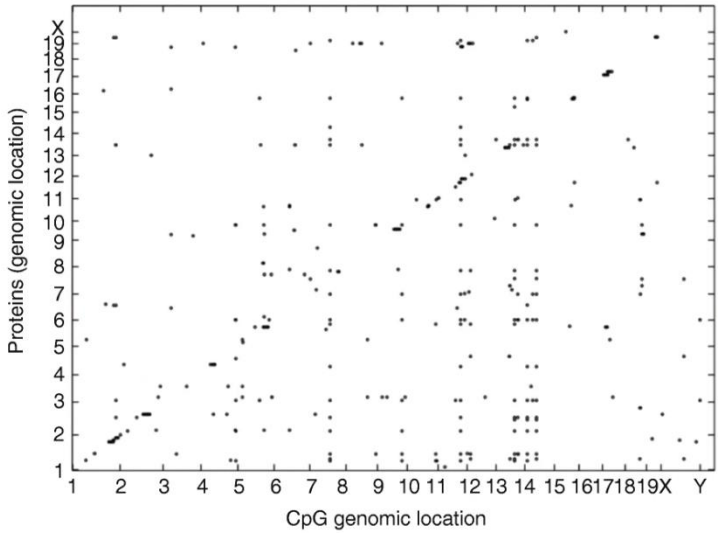
The Epigenomics first approach—功能调控中介数据

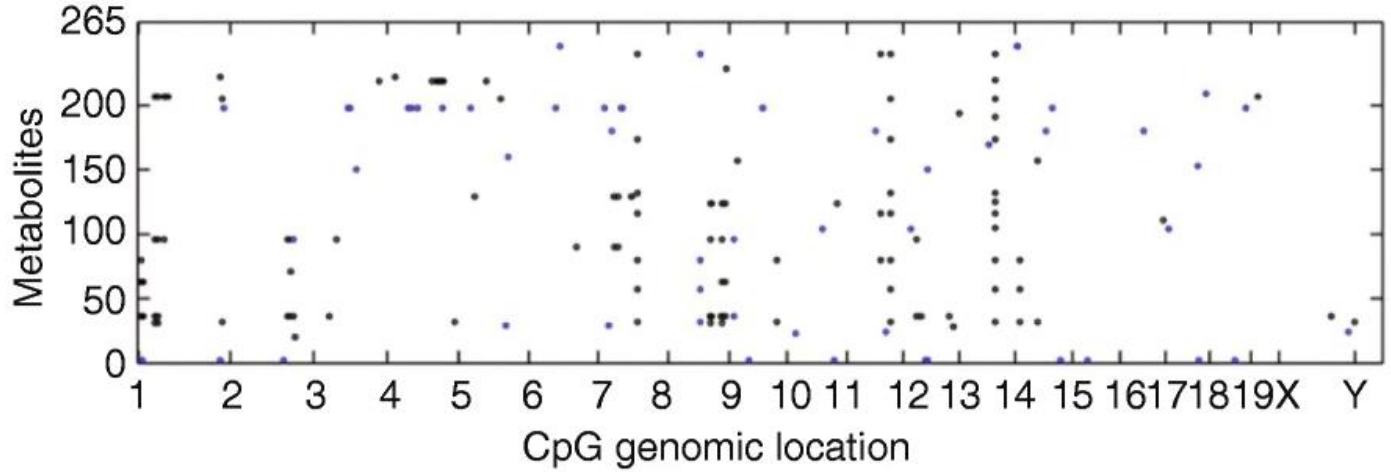
87% of cis-eQTL SNPs, 82% of cis-pQTL SNPs, and 59% of GWAS Catalog SNPs are also cis-meQTL variants

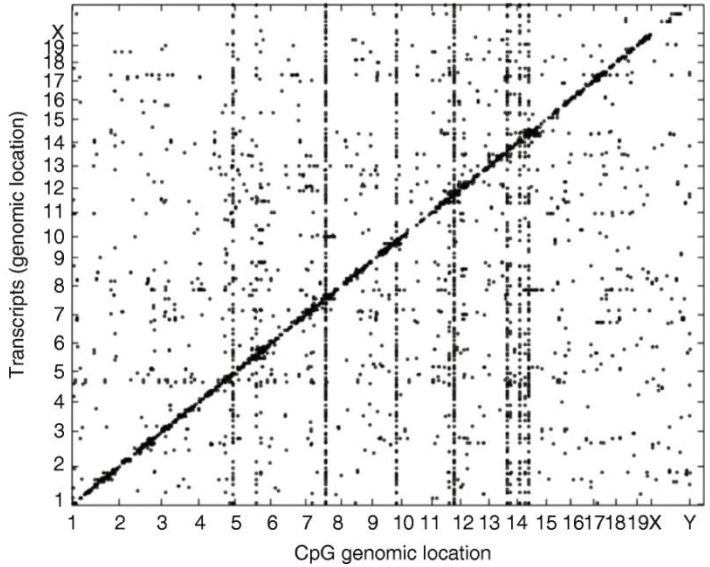
		cis-meQTL			*trans*-meQTL		
	Total SNPs	N of the overlap	Fold Change	P-value	N of the overlap	Fold Change	P-value
cis-eSNPs FHS (1e-7)	460,536	399,670	1.66	0	84,591	2.33	0
cis-eSNPs Eur-Meta	398,524	356,500	1.71	0	70,810	2.25	0
cis-pQTLs FHS-1000g	12,401	10,176	1.57	0	2,045	2.09	0
pQTLs FHS-1000g	19,942	16,660	1.6	0	4,428	2.81	0
cis-pQTLs NC2016	376	318	1.62	0	53	1.78	1.63E-05
pQTLs NC2016	456	385	1.61	0	68	1.89	1.79E-07
metabolism QTLs	135	111	1.57	8.38E-14	26	2.44	6.52E-06
GWAS Catalog	32,260	19,000	1.13	0	3078	1.21	0

The Epigenomics first approach—多组学相关性分析QTM

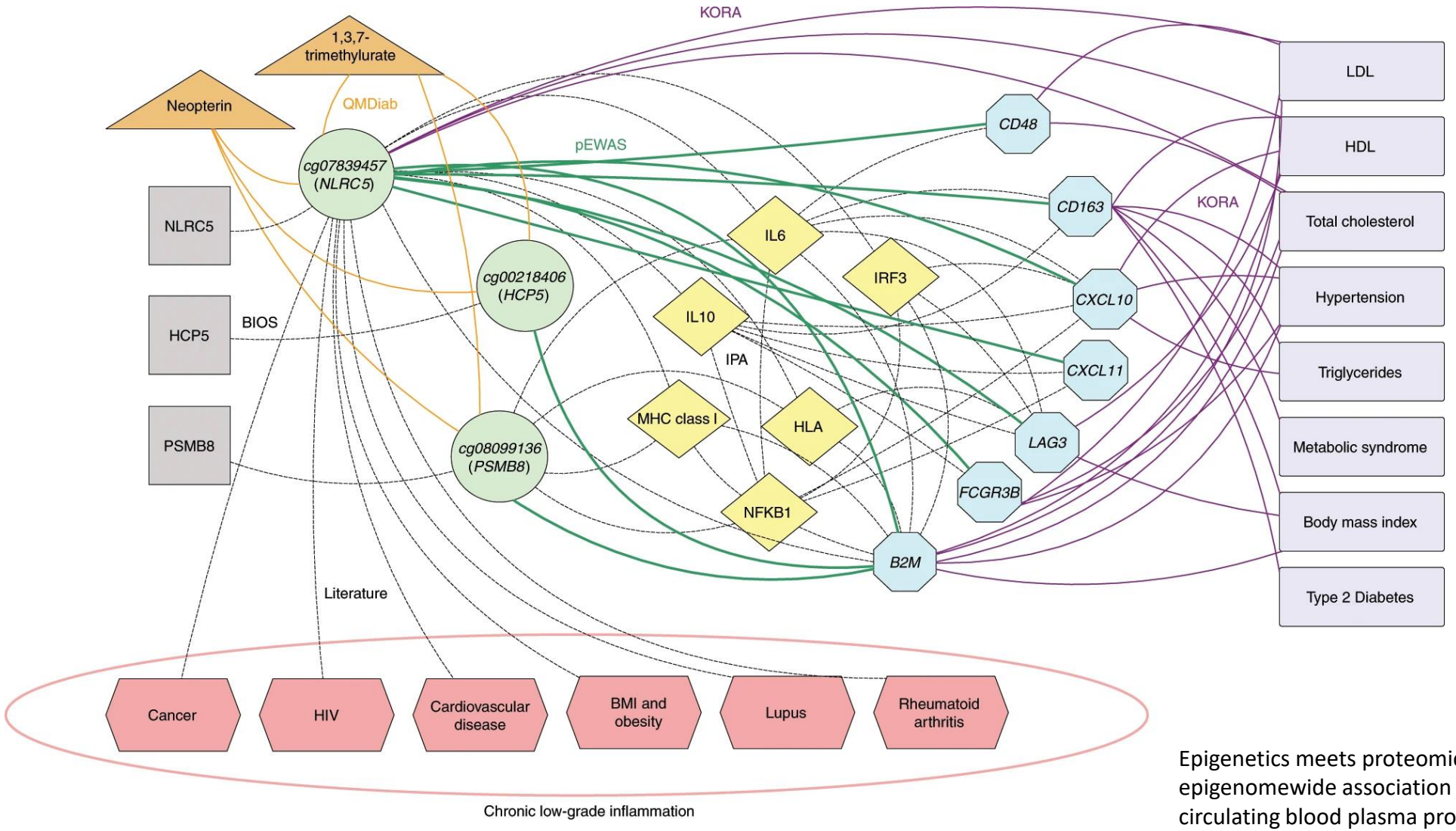






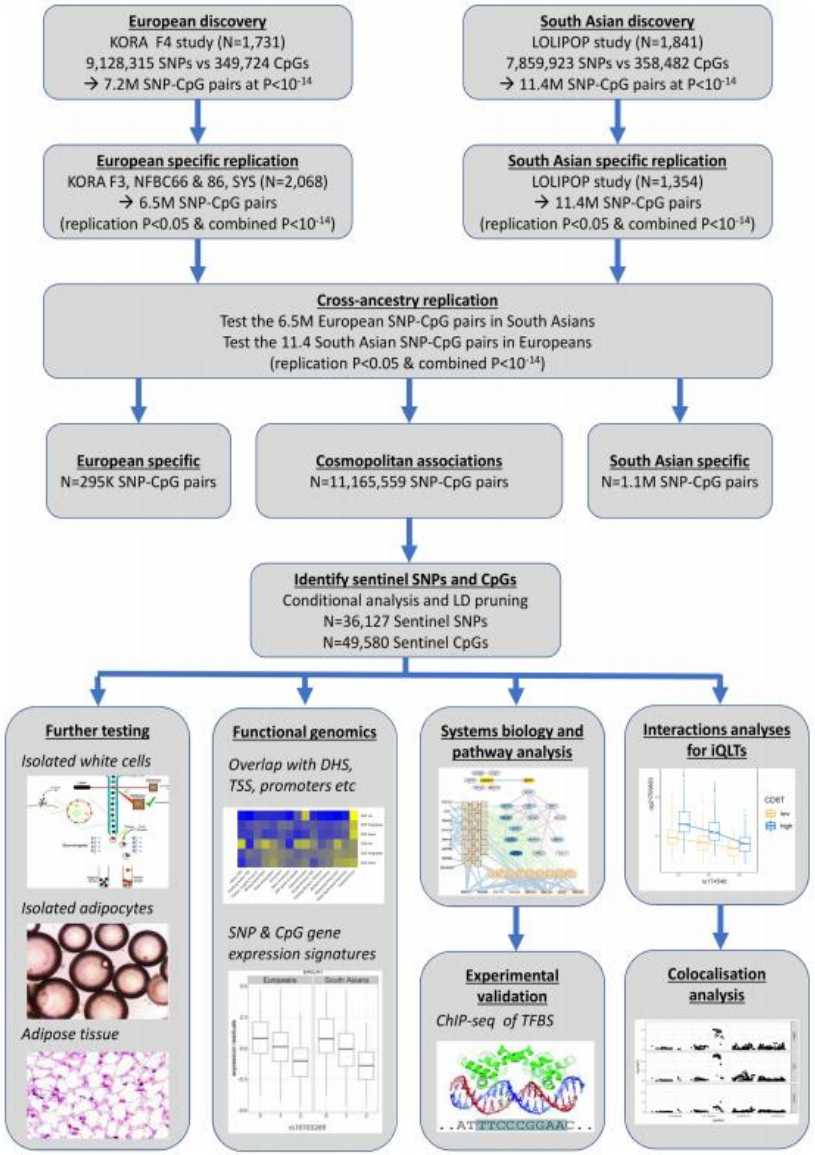


The Epigenomics first approach—多组学调控网络构建



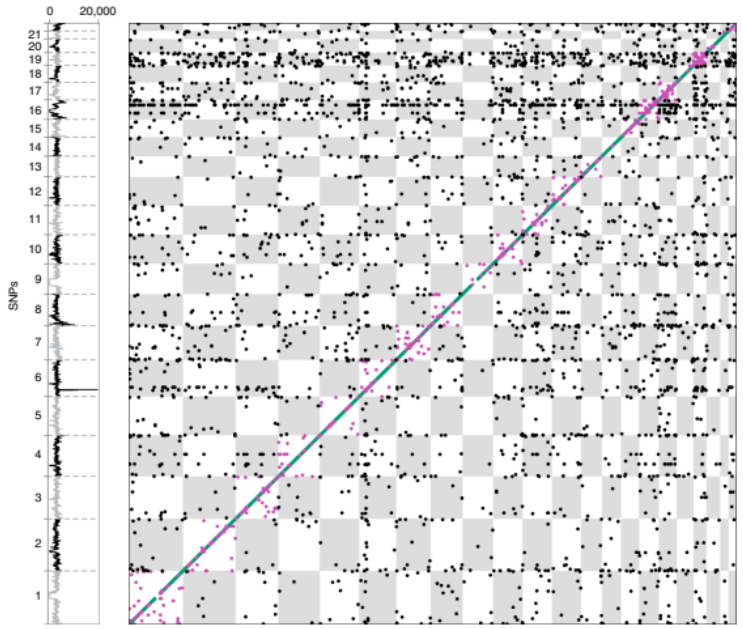
Epigenetics meets proteomics in an epigenomewide association study with circulating blood plasma protein traits *Nature communications* 2020

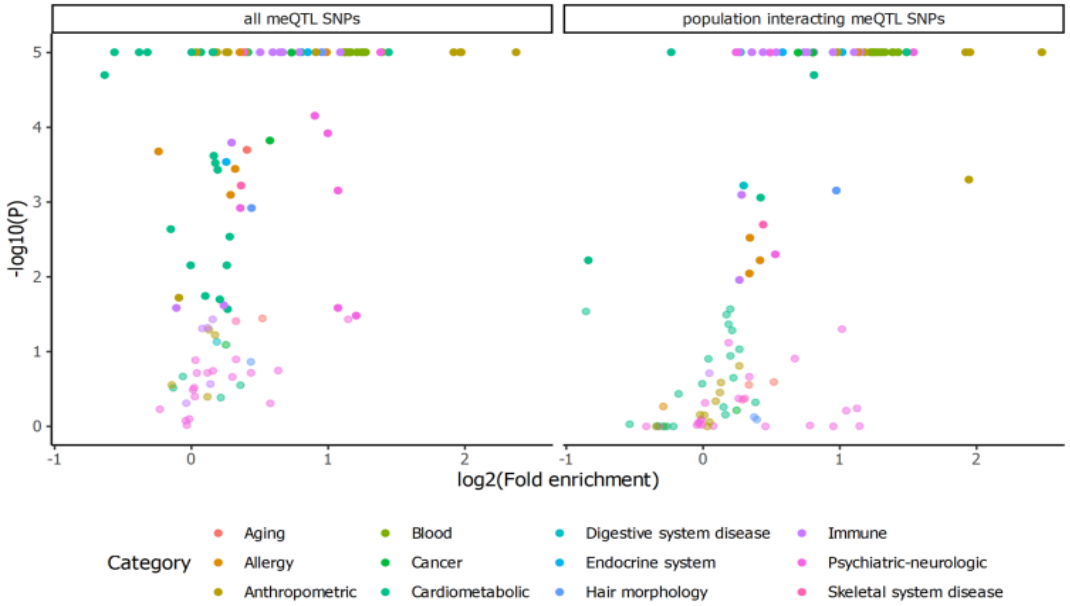
The Epigenomics first approach—meQTL图谱分析

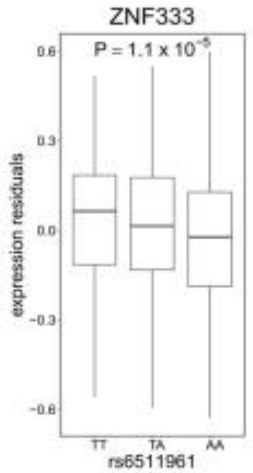


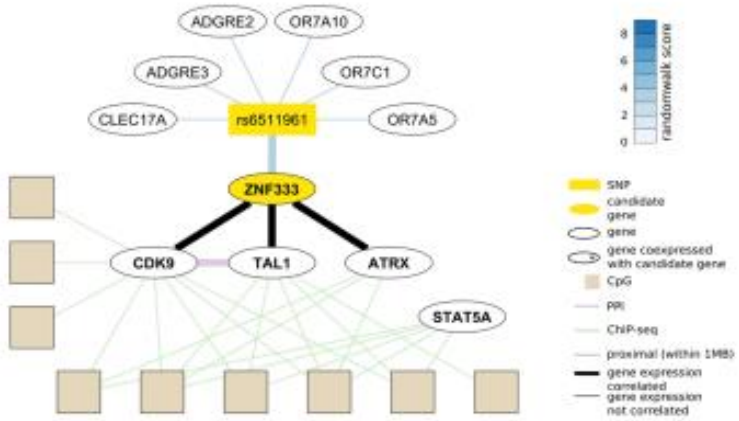
Genetic variation influencing DNA methylation provides insights into molecular mechanisms regulating genomic function *Nature Genetics* 2022 Jan

The Epigenomics first approach—meQTL图谱分析

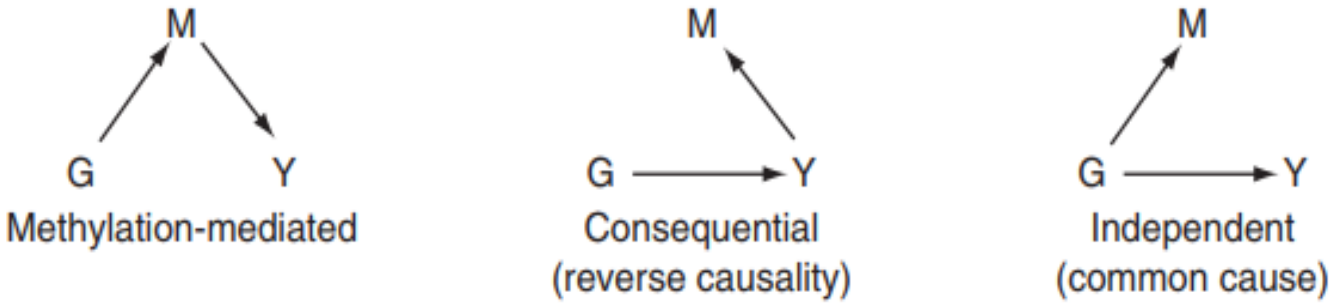


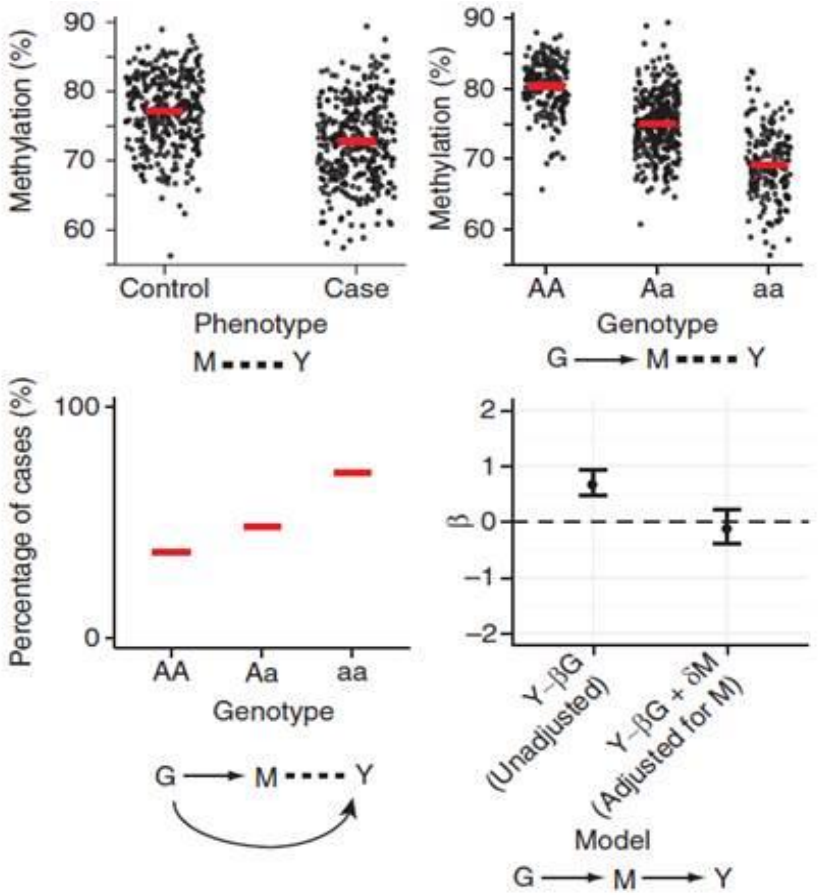






The Epigenomics first approach—CIT（mediation)分析





Two-step epigenetic mendelian randomization: A strategy for establishing the causal role of epigenetic processes in pathways to disease

The Epigenomics first approach—eQTM分析及分子调控功能机制验证

博淼生物 BIO MIAO BIOLOGICAL — SINCE 2009 —

Discovery Phase (Human Blood)

DNA methylation Microarray
1st Cohort : 40 MDD vs. 40 CTL
|Delta Beta| > 0.2; $p < 3.6 \times 10^{-8}$

RNA Expression Microarray
2nd Cohort : 20 MDD vs. 20 CTL
|Log_2 fold change| > 1; $p < 0.01$

Gene: *BICD2*

Replication Phase (Human Blood)

Sequenom MassARRAY
3rd Cohort: 528 MDD vs. 818 CTL

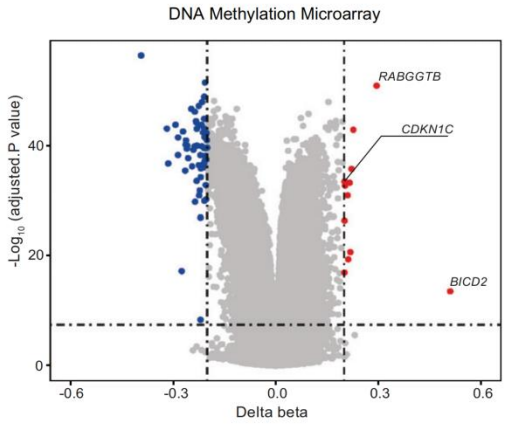
MSE-qPCR/qPCR: Combined analysis of DNA methylation & mRNA expression
4th Cohort: 30 MDD vs. 30 CTL

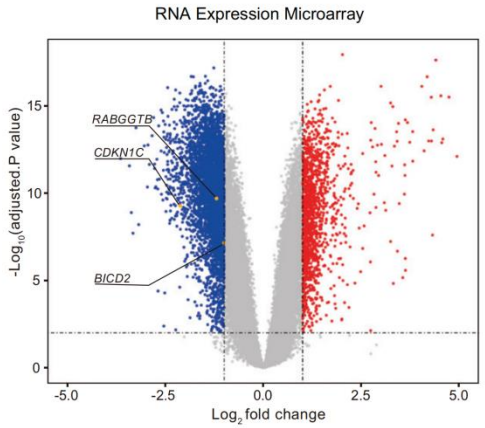
Functional study Phase (Mouse Blood & Brain)

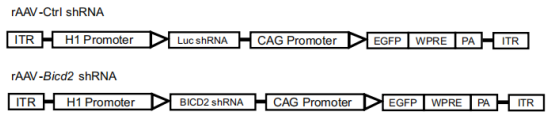
Validation of human results in CUMS model

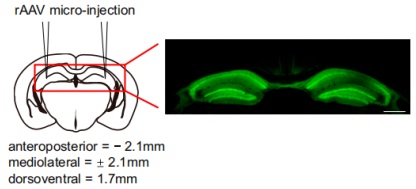
Mapping of brain regions with altered *Bicd2* DNA methylation and mRNA expression

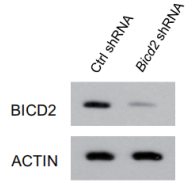
Functional studies in vivo & in vitro

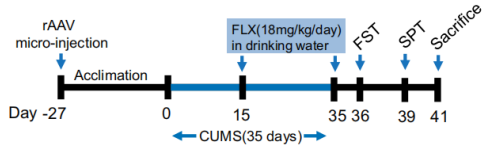




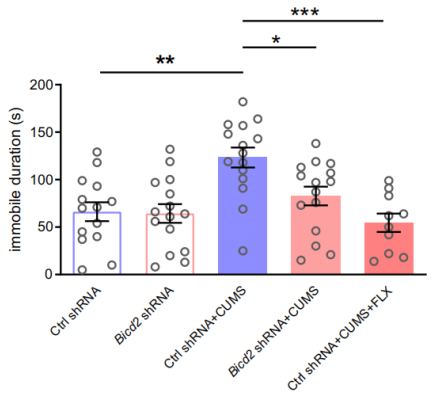




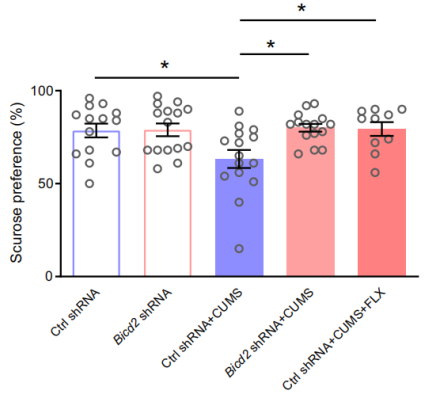




Forced swimming test



Sucrose Preference test



Elevated BICD2 DNA methylation in blood of major depressive disorder patients and reduction of depressive-like behaviors in hippocampal Bicd2-knockdown mice *PNAS 2022 July*

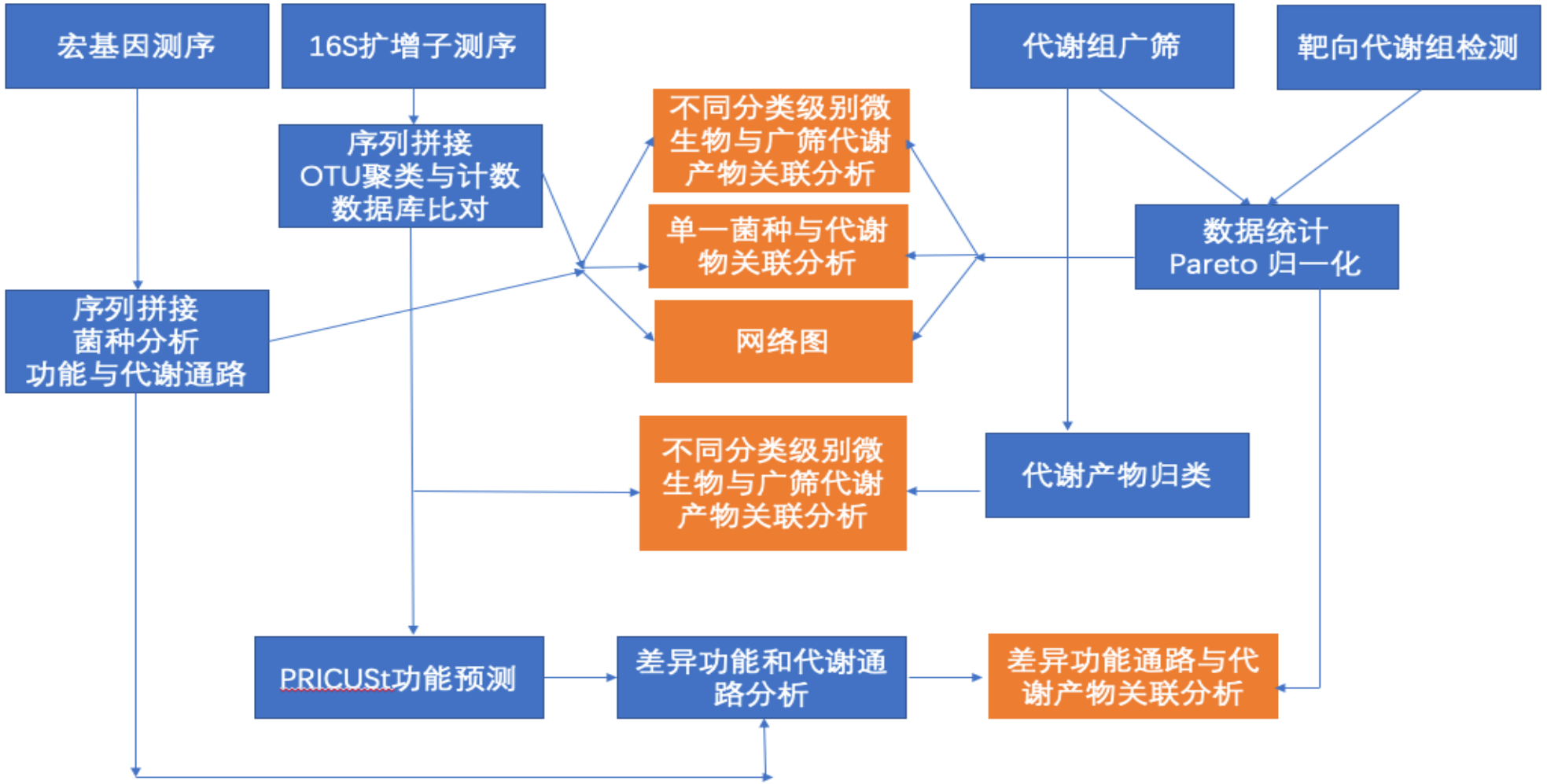
博淼表观基因组核心技术服务项目

EWAS芯片 Illumina EPIC V2.0—935K芯片

Multi-PCR NGS 靶向甲基化定量 适合20-200个高通量靶向甲基化区域定量检测

Massarray 靶向甲基化定量 适合1-20个中通量靶向甲基化区域定量检测

The Microbiomics first approach



The Microbiomics first approach

博淼生物
BIO MIAO BIOLOGICAL
SINCE 2009

粪便/肠道/口腔内容物样本
(16S/宏基因组)

血清/血浆样本
(广筛、氨基酸、脂肪酸、短链脂肪酸、胆汁酸)

侧重研究微生物的代谢产物对人体的影响

粪便/肠道内容物样本
(16S/宏基因组)

粪便样本
(广筛、氨基酸、脂肪酸、短链脂肪酸、胆汁酸)

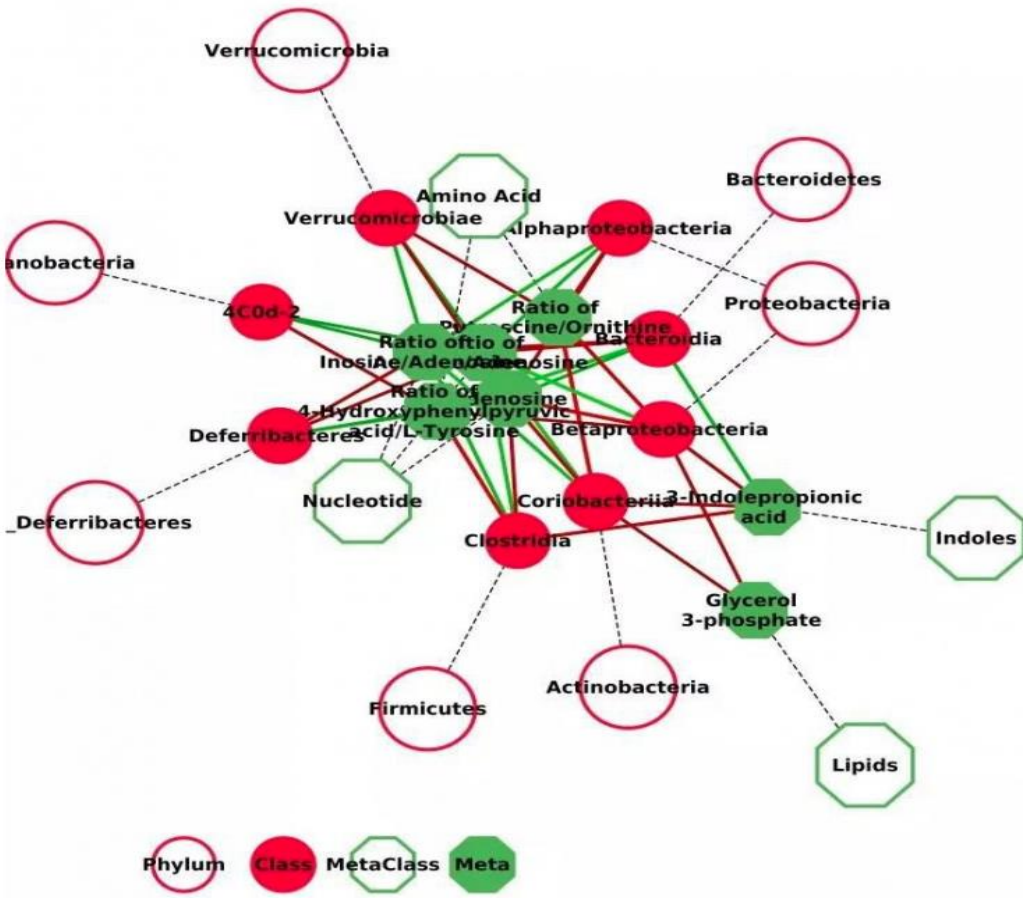
侧重于研究微生物本身分解或产生代谢物的情况

粪便/肠道内容物样本
(16S/宏基因组)

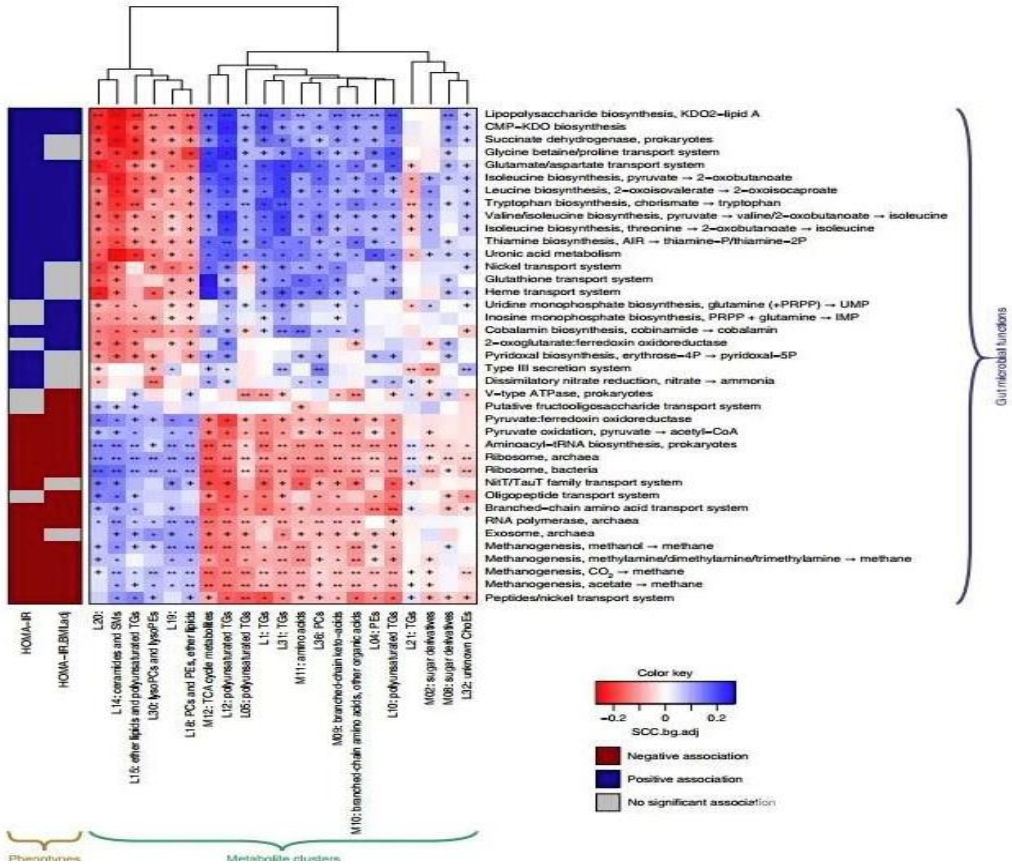
尿液/血清、血浆样本
(神经递质)

侧重于“肠脑轴”研究

The Microbiomics first approach —关联网络及代谢通路分析

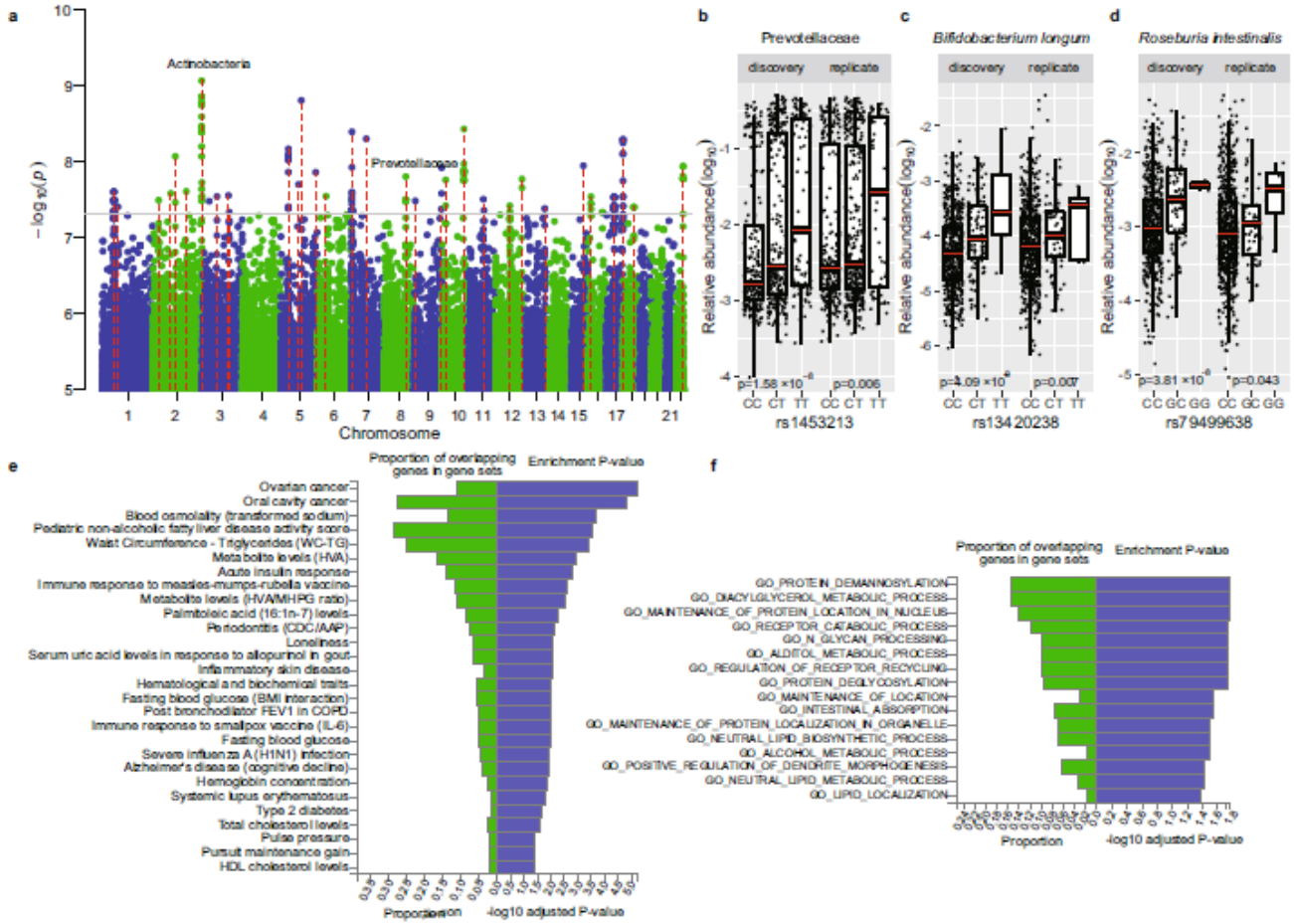


微生物与代谢产物关联网路图



差异菌种代谢功能与代谢产物关联分析

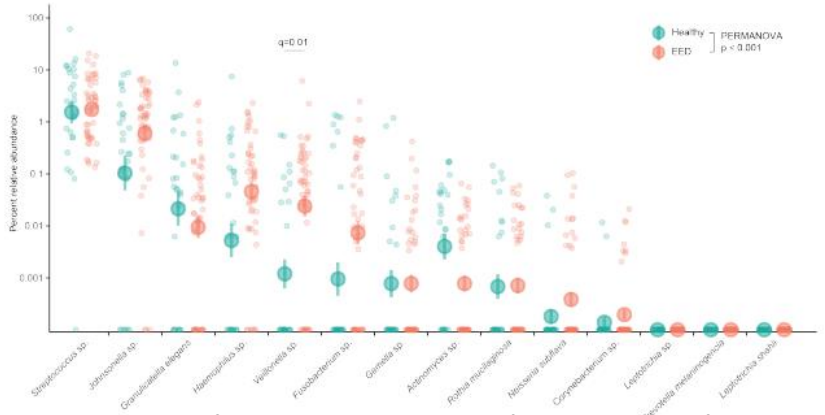
The Microbiomics first approach —M-GWAS分析



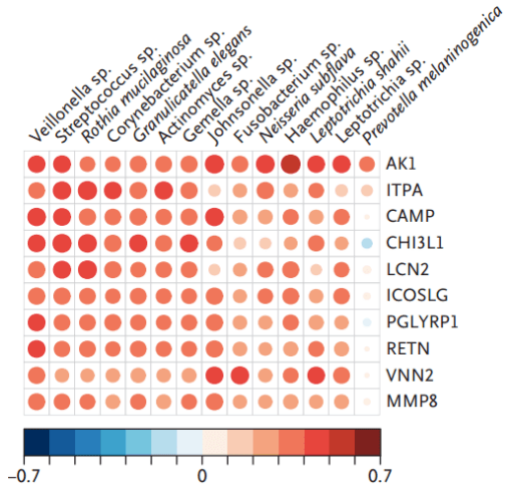
图a:Manhattan图展示了37个常见变异的基因座与特定肠道细菌分类群显著相关；b-d:3个宿主基因-特定肠道菌的关联在发现和验证集中均显著；e-f:KEGG等通路分析显示与肠道菌显著相关的遗传信号主要集中在代谢、神经和免疫功能等方面。

A genome-wide association study for gut metagenome in Chinese adults illuminates complex diseases. Cell Discov 7, 9 (2021)

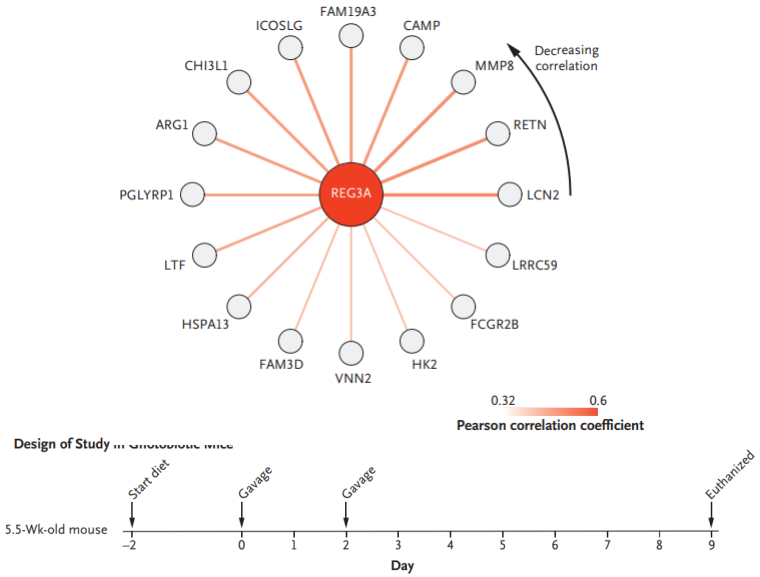
The Microbiomics first approach —宏基因组&蛋白质组调控机制验证



Relative abundances of core duodenal taxa in the fecal microbiota of children living in Mirpur who have healthy growth phenotypes (n=27) and those with EED (n=48)



The top 10 positive correlations between members of the 14 core taxa and duodenal proteins



A shows the design of the experiment in which gnotobiotic mice

分析 36 名环境肠功能障碍（EED）患儿的十二指肠微生物群，鉴定出 14 个不属于典型肠道病原体的 EED 核心细菌类群

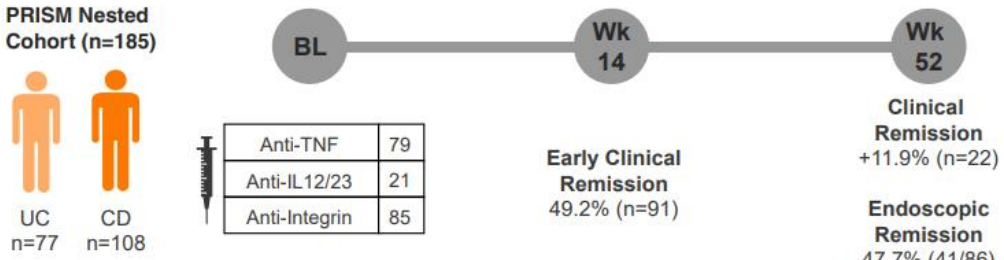
特征菌群的绝对水平，与患儿生长负相关，与参与免疫炎症应答的十二指肠蛋白（如 LCN2）正相关，且在患儿粪便中的含量不同于健康儿童

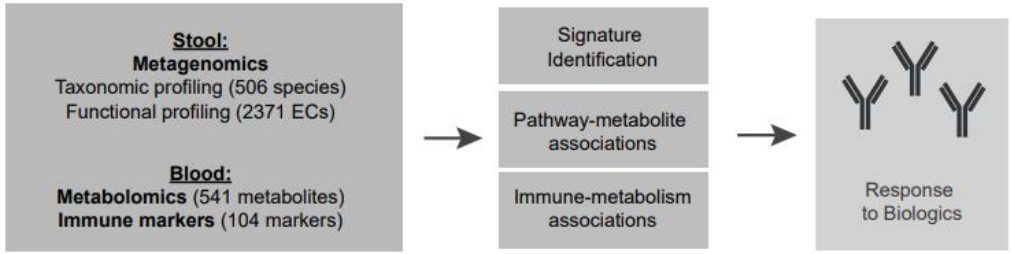
和 EED 核心菌相关的十二指肠蛋白，与血浆 REG3A 等显著相关

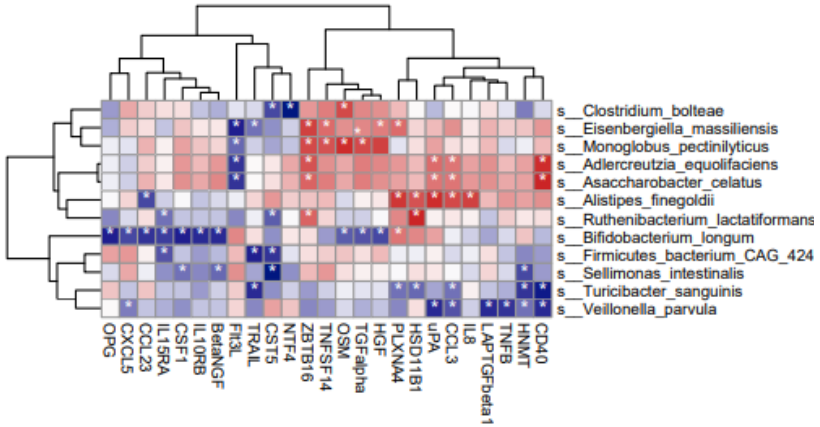
定植 EED 十二指肠菌群的部分分离菌（含大部分 EED 核心菌）可能通过促进十二指肠炎症应答，参与 EED 的发生发展，从而导致发育不良

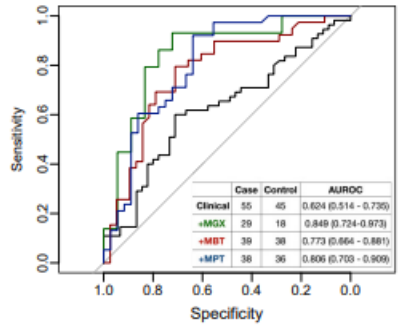
Duodenal microbiota in stunted undernourished children with enteropathy NEJM 2020

The Microbiomics first approach —宏基因组&血清代谢组及蛋白质组









多组学揭示了与不同IBD疗法缓解相关的微生物组机制

一级到二级胆汁酸的7α/β-脱羟作用决定了抗细胞因子的反应

27种特有的血清免疫蛋白与肠道微生物组相关联

多组学联合预测疗效模型相对单一组学模型AUC显著提升

Heatmap of 25 serum proteins in IBD patients with at least one significant ($q < 0.1$) association with metagenomic taxonomic features

black, clinical variables only; green, clinical and metagenomic features; brown, clinical and metabolomic features; blue, clinical and proteomics; orange, clinical and all three 'omics

Multi-omics reveal microbial determinants impacting responses to biologic therapies in inflammatory bowel disease Cell Host Microbe 2021 Aug

博淼微生物基因组&代谢组核心技术服务项目

博淼生物
BIO MIAO BIOLOGICAL
SINCE 2009

16S/18S/ITS扩增子测序 价格低廉、适合大队列Discover研究

宏基因组测序 微生物全基因组检测、功能机制深度注释

微生物简化基因组测序 精确菌种鉴定、疑难样本检测（肿瘤组织等）

非靶向代谢组检测 全景式呈现代谢物定性和相对定量的检测

高通量靶向代谢组 兼具非靶向技术的种类高通量及靶向技术的绝对定量

靶向代谢组检测 肠道菌群代谢物、胆汁酸、甾体激素等特定种类代谢物绝对定量

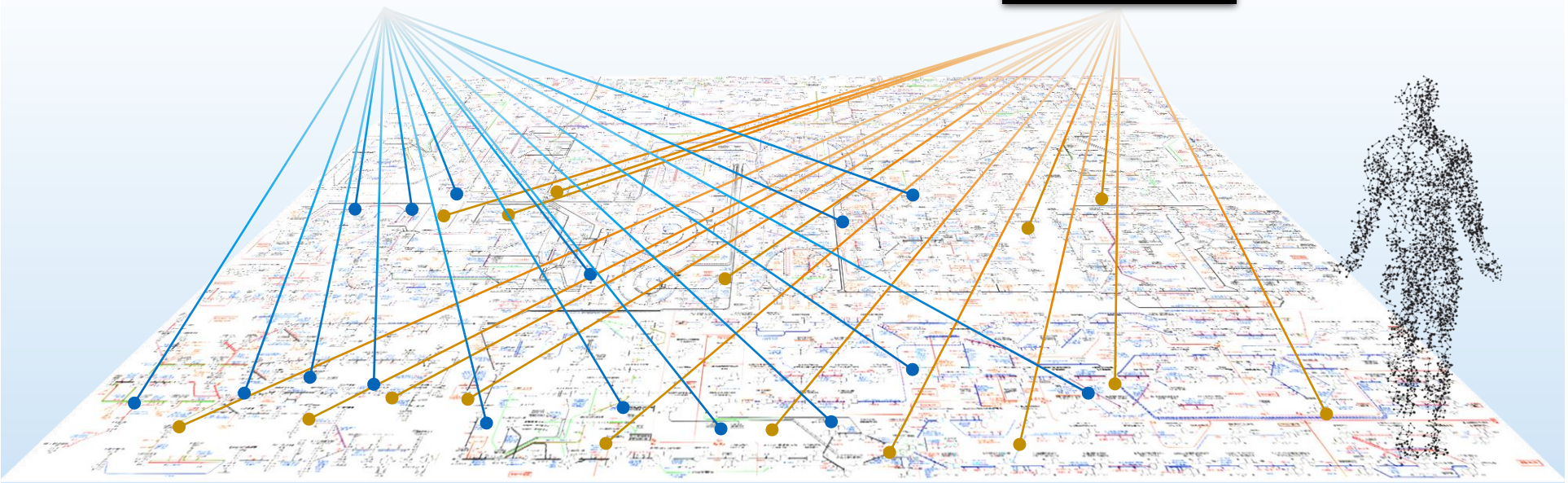
The Proteomics first approach —功能分子表型数据

遗传因素+环境暴露

↓

*Molecular Phenomics*分子表型组

Metabolomics　　***Proteomics***



↓

Marker panel, personalized medicine, bio-insight→ better health

The Proteomics first approach

调控功能机制的验证

✓显著差异蛋白及代谢物的相关性分析、KEGG通路联合分析、PCA比较分析等

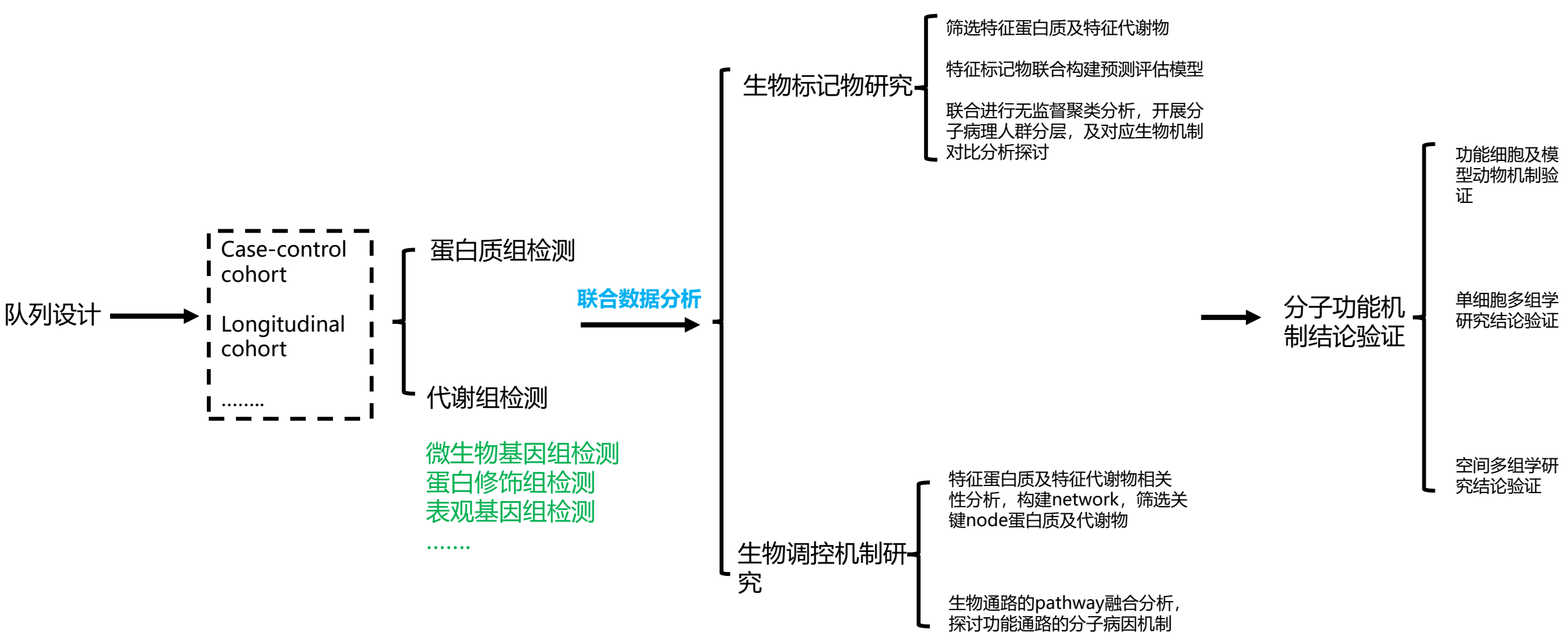
✓系统描绘蛋白至代谢的调控过程，挖掘关键蛋白与代谢物上下游调控通路

多组学联合预测模型构建

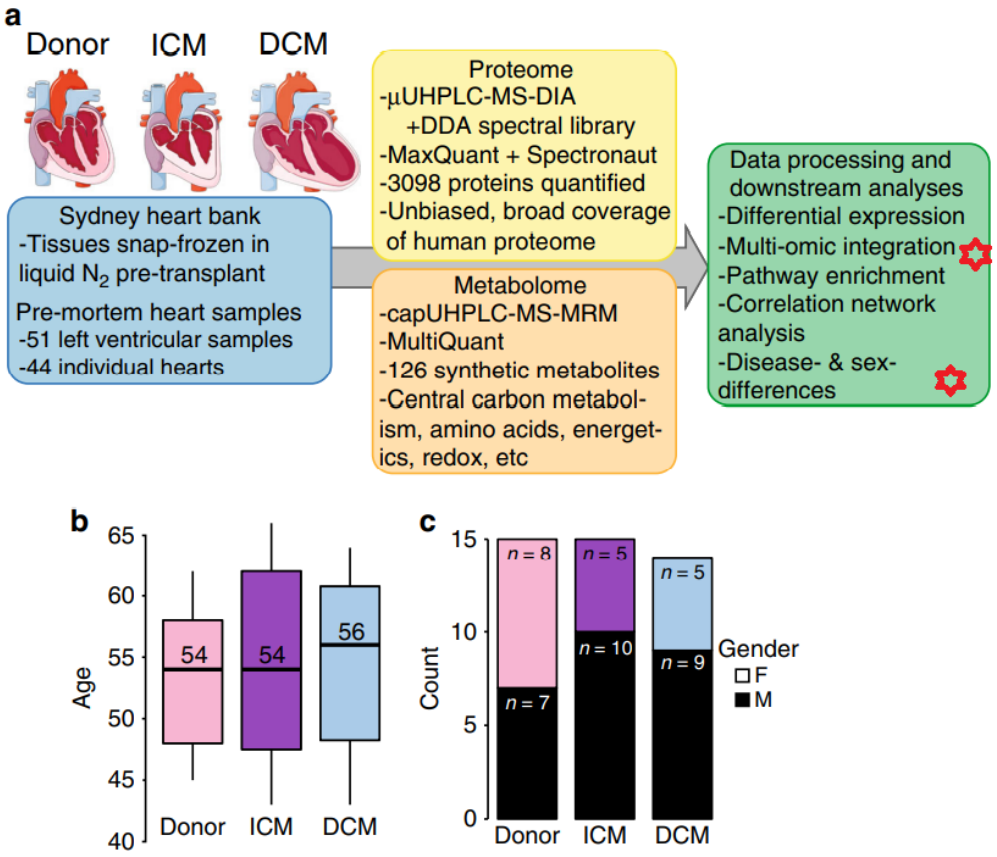
✓**单一组学标记物、多组学标记物联合，基于机器学习算法进行预测模型构建及AUC效率比较**

✓**多维度组学标记物联合模型构建预测效率更佳**

Proteomics&Metabolomics approach研究策略—研究路径概述



The Proteomics first approach —调控功能机制研究



研究设计

疾病表型：不同病因心肌病（缺血性心肌病、扩张性心肌病）的生物调控机制差异	特色：围绕不同性别对应疾病组织进行了蛋白质组学及代谢组差异分析

单一组学差异分析

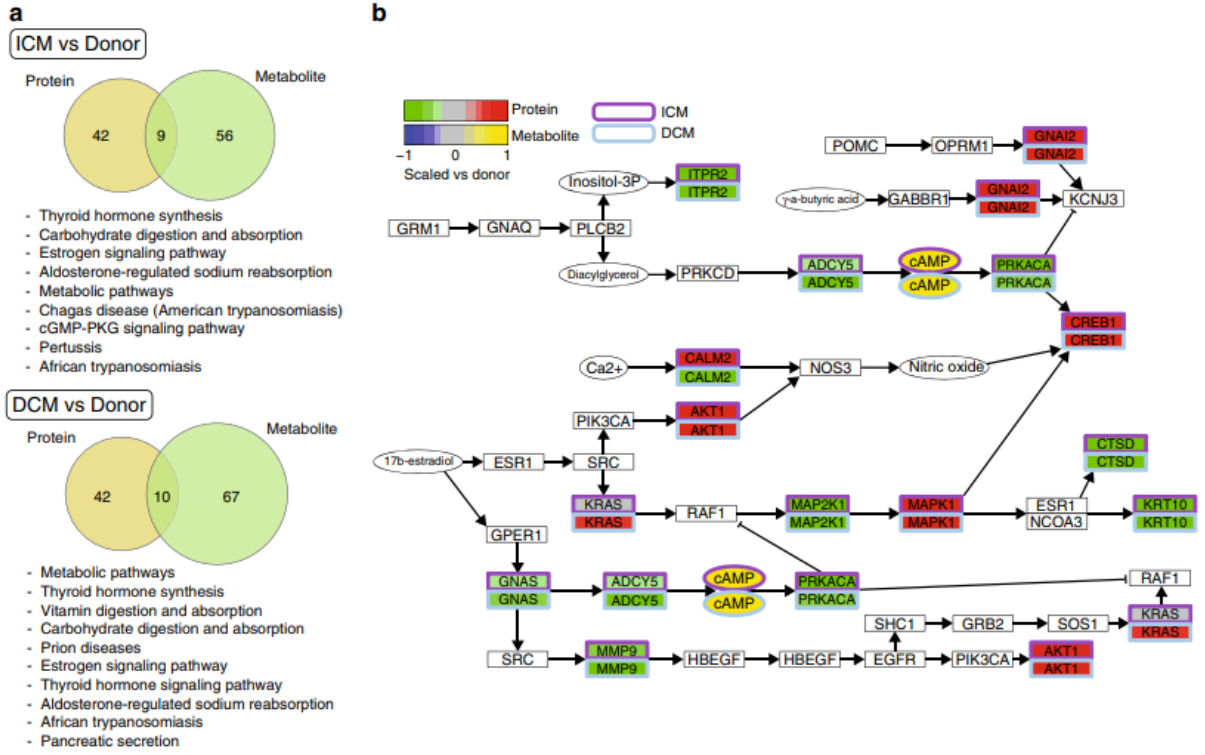
蛋白质组差异分析：根据ICM\DCM分别与Donor组间分析，筛选差异蛋白质及相关生物学特征	代谢组差异分析：根据ICM\DCM分别与Donor组间分析，筛选差异代谢物及相关生物学特征

蛋白质组&代谢组联合分析

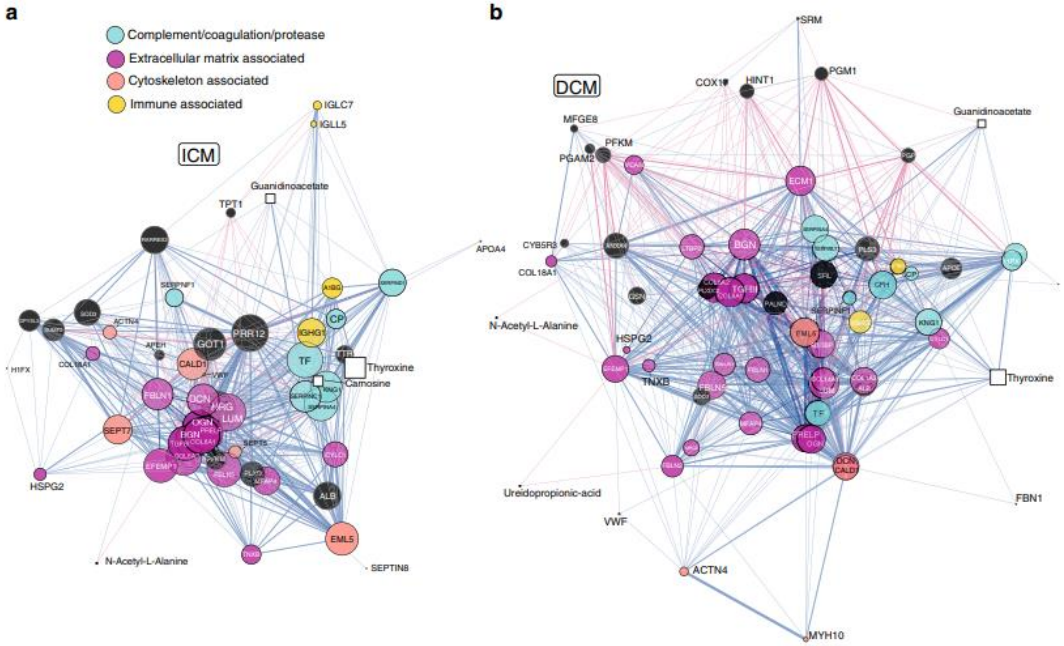
KEGG通路联合分析：对 ICM 和 DCM 的蛋白质和代谢物水平上进行KEGG通路分析,显示了两者在蛋白质和代谢物水平上调节的共同途径	显著差异蛋白质与代谢物相关性分析：在相关性分析基础上构建网络调控图，筛选关键节点蛋白质及代谢物

Core functional nodes and sex-specific pathways in human ischaemic and dilated cardiomyopathy *NATURE COMMUNICATIONS | (2020) Article*

The Proteomics first approach —调控功能机制研究



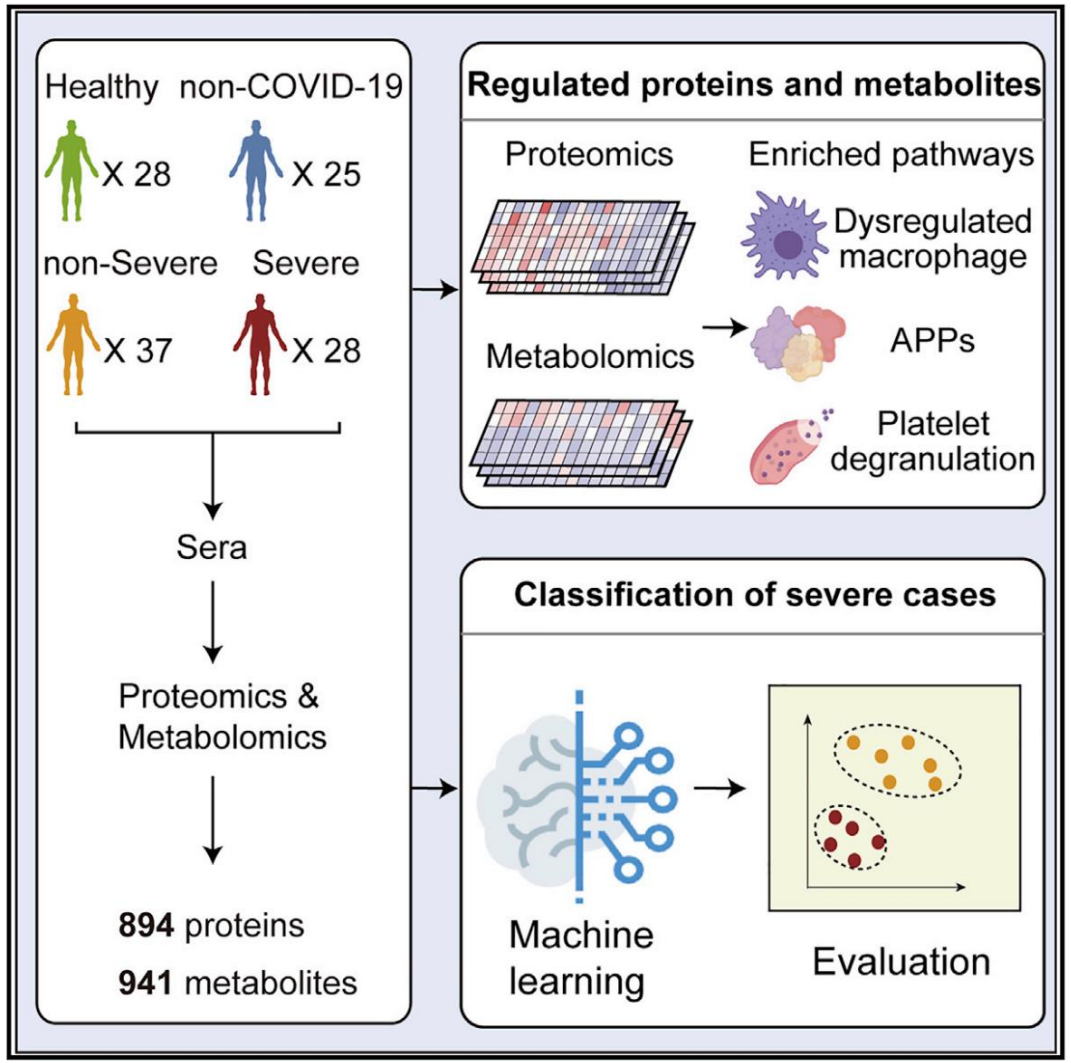
Pathways analysis



Network analysis

Core functional nodes and sex-specific pathways in human ischaemic and dilated cardiomyopathy *NATURE COMMUNICATIONS | (2020) Article*

The Proteomics first approach —联合预测模型构建研究



研究设计

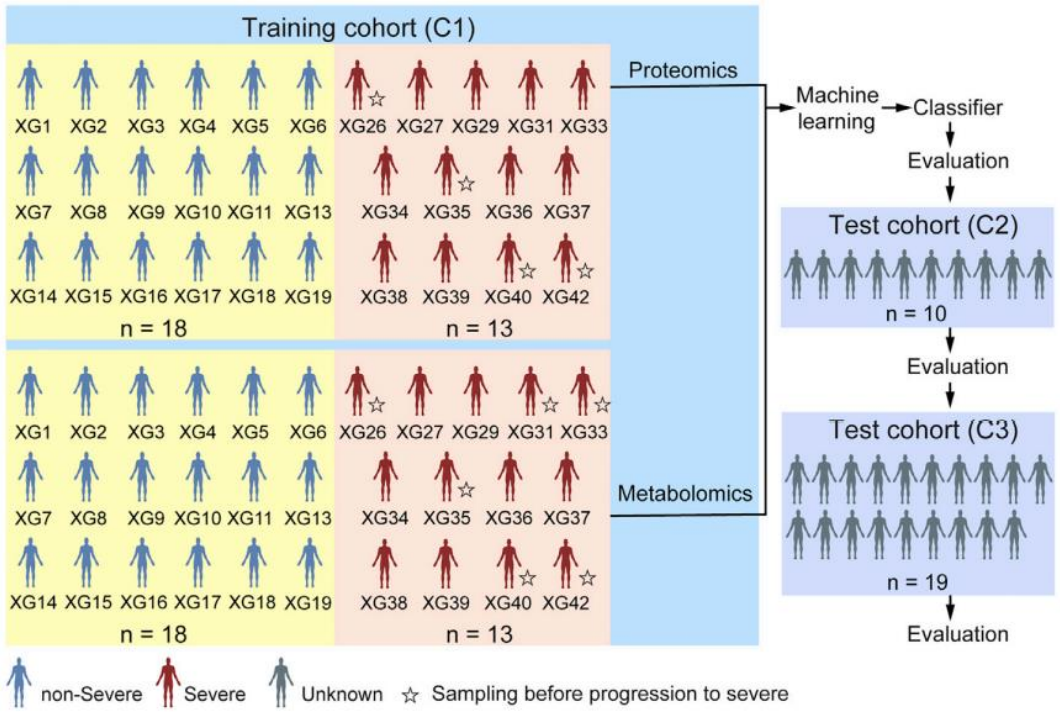
疾病表型：COVID-19感染患者的严重程度

Highlights

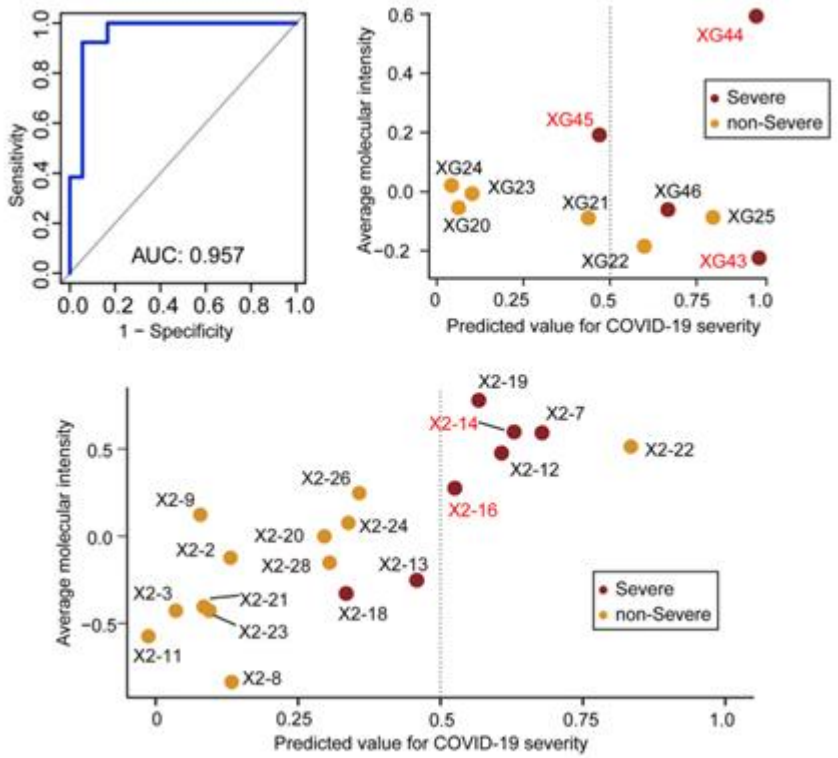
93 proteins show differential expression in severe COVID-19 patient sera	204 metabolites in COVID-19 patient sera correlate with disease severity	A model composed of 29 serum factors shows patient stratification potential(22 proteins+9metabolites)	Pathway analysis highlights metabolic and immune dysregulation in COVID-19 patients

Proteomic and metabolomic characterization of COVID-19 patient sera *CELL | (2020) Article*

The Proteomics first approach —联合预测模型构建研究



Study design for machine-learning-based classifier development for severe COVID-19 patients



Separation of Severe and Non-severe COVID-19 Patients by Machine Learning of Proteomic and Metabolomic Features

博淼蛋白质组核心技术服务项目

博淼生物
BIO MIAO BIOLOGICAL
SINCE 2009

4D-DIA/3D-DIA蛋白质技术 非同位素标记、全谱精准定量检测、适合大队列研究

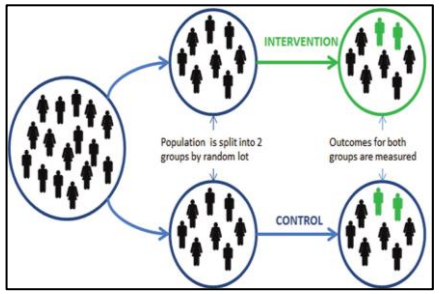
iTRAQ/TMT蛋白质组技术 同位素标记、定量准确性高、适合小样本研究

Olink蛋白质组技术 采用创新PEA技术、96-3072个靶向蛋白检测、适合队列研究

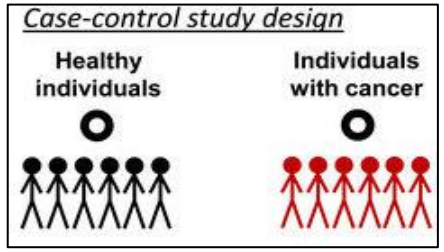
PRM靶向定量蛋白质技术 针对10-50个靶向蛋白相对/绝对定量、适合与DIA技术联合开展队列研究

Multi-omics分子病因学领域研究展望

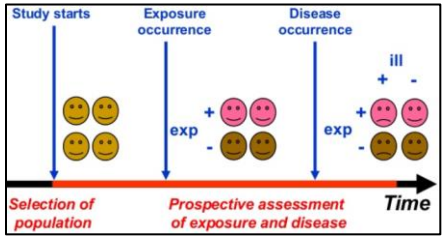
Population sample types



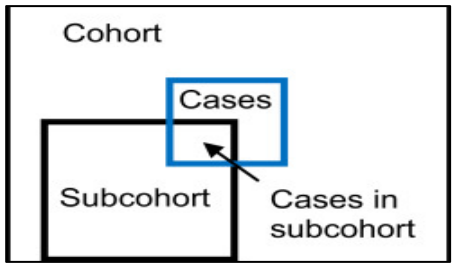
Randomized controlled trial(RCT)
随机对照研究



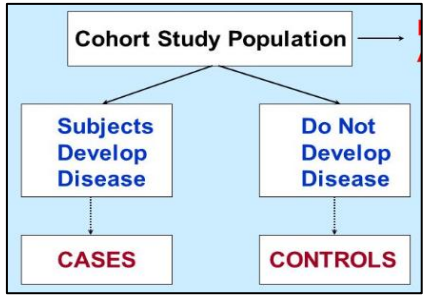
Case-control study
病例对照研究——**精细化表型亚分组**



Prospective cohort study
前瞻性队列研究

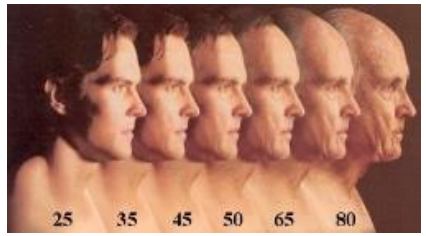


Case-cohort study
病例-队列研究
病例参比式研究



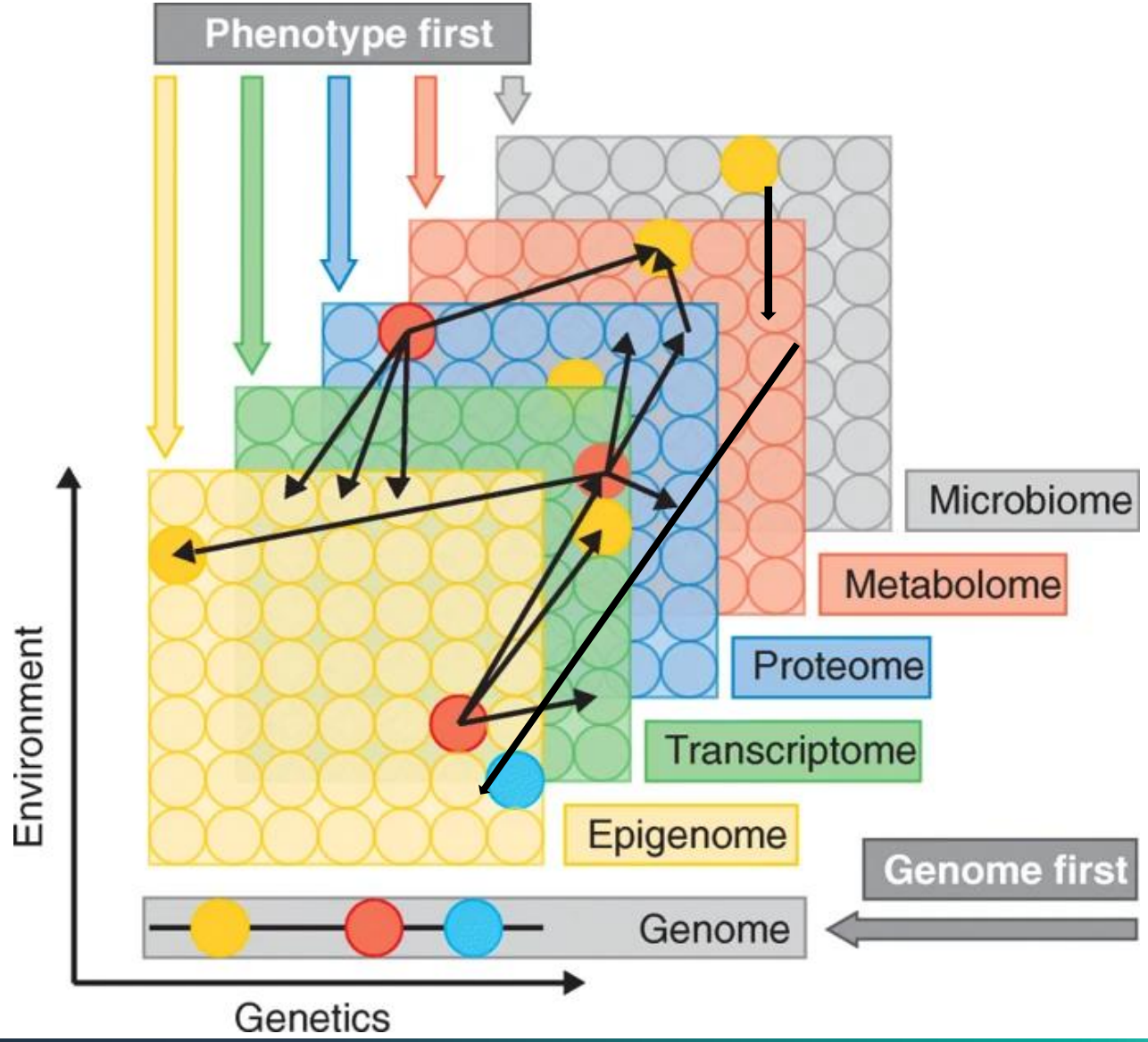
Nested case-control study
巢式病例对照研究
队列内病例对照研究

Cross-sectional study
横断面研究



Longitudinal study
纵贯研究，纵向研究

Multi-omics approaches to disease research



Multi-omics trajectory to disease research

博淼生物
BIO MIAO BIOLOGICAL
SINCE 2009

多组学调控网络构建及因果效应分析

生物学功能通路挖掘及关键调控分子筛选

细胞体外及动物体内分子调控功能机制验证

Multi-omics strategy to epidemiology

动物模型验证性研究

研究优势：环境因素可控、组织容易获取、精准表型设置等等

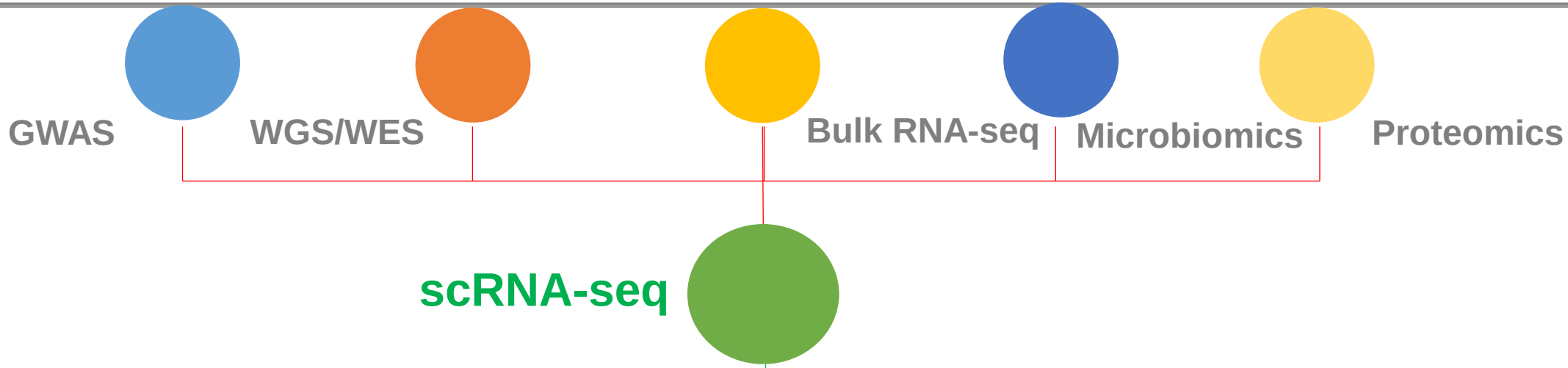
研究劣势：无法准确反应人体生物学通路、疾病表型不一致等等

单细胞组学验证性研究

研究优势：在单细胞水平验证原位细胞图谱及对应的关键功能基因、调控通路等

研究劣势：无法系统验证对于细胞的分子功能调控机制

Bulk omics&Single cell omics



- ✓ 遗传易感基因在单细胞水平调控机制解析
- ✓ 通过单细胞eQTL精准筛选关键调控基因

- ✓ 功能基因在单细胞水平转录及通路调控分析
- ✓ 探讨基因变异对于单细胞水平转录的潜在调控机制

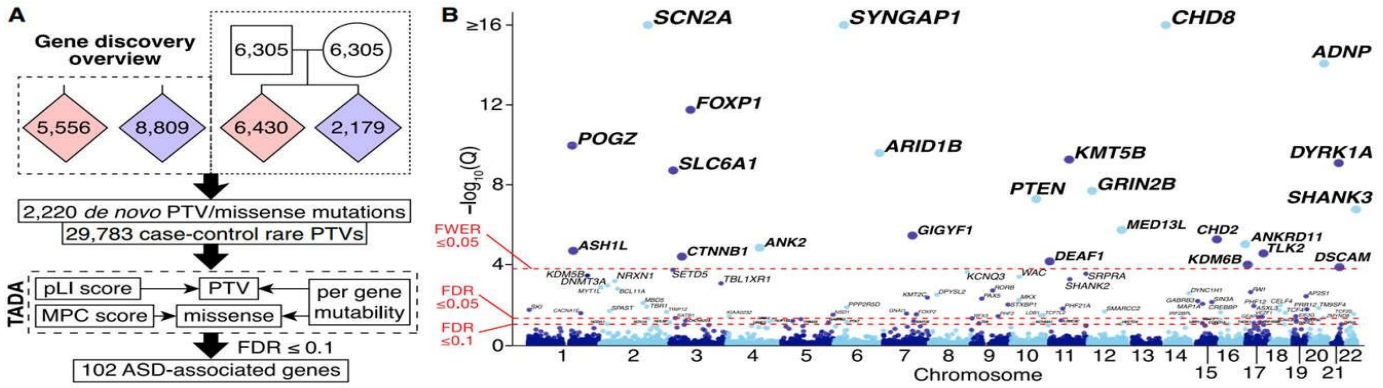
- ✓群体样本RNA-seq数据对于scRNA科研结论的佐证
- ✓scRNA关键细胞功能基因验证

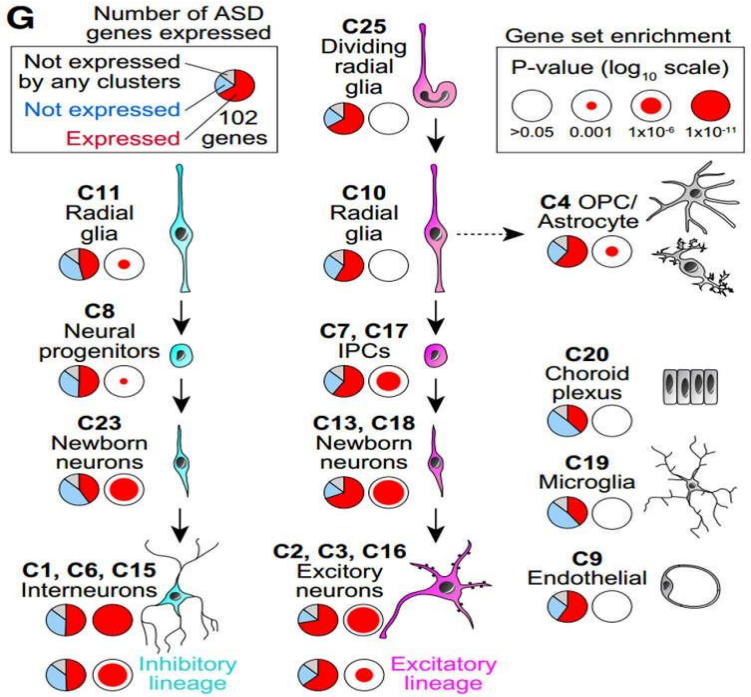
- ✓微生物基因组对于宿主单细胞水平关键基因的转录调控
- ✓精准构建微生物基因组—代谢组—scRNA调控体系

- ✓预测评估调节性蛋白的细胞来源
- ✓开展细胞与表型的功能机制研究

WES&ScRNA-seq联合研究

探究遗传致病突变所在的基因在细胞类型特异性表达



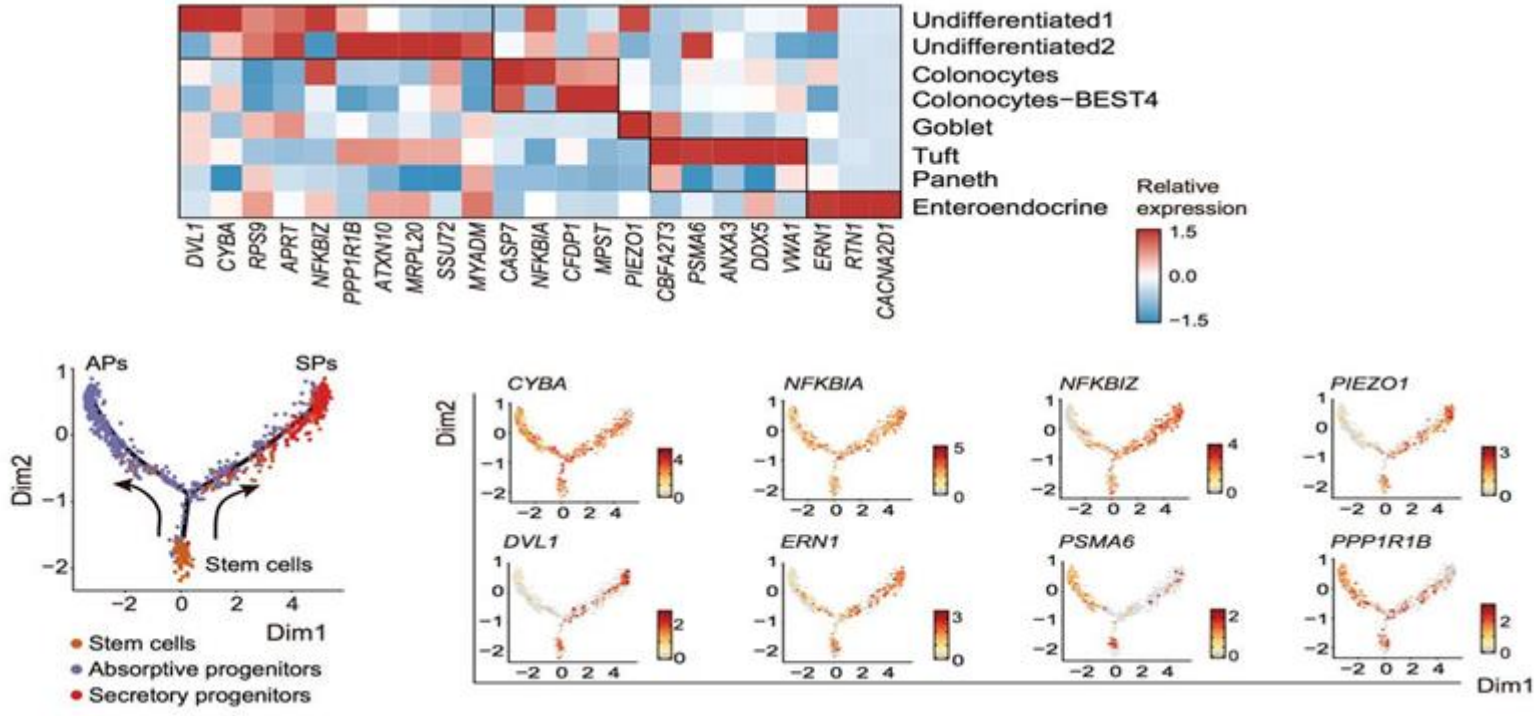


- 在自闭症谱系障碍（ASD）的研究中，分析了35,584例样本的全外显子数据，共识别 出102个与ASD相关的高风险基因。
- 单细胞测序结果表明，这102个基因主要在成熟的抑制和兴奋神经元中高表达。
- 本研究更深入了解了自闭症高风险基因的分子机制，也有助于发现治疗和减轻自闭症某些严重症状的方法。

F. Kyle Satterstrom et al., Cell, 2020

GWAS&ScRNA-seq联合研究

探究致病风险基因差异表达对疾病的促进作用



- 本研究用GWAS分析鉴定出244个PIBD风险相关的基因，在上皮效应细胞和干或祖细胞亚型中叠加候选PIBD风险基因，例如，*ERN1*与CD和UC风险相关，并介导未折叠蛋白反应，在干细胞、祖细胞和肠内分泌细胞中高度表达；*CASP7*，一种细胞消亡蛋白酶，介导感染诱导的上皮细胞死亡，在结肠细胞（colonocytes）和colonocytes-BEST4 细胞中表达。
- PIBD风险基因的差异表达可能导致上皮完整性受损和应激反应。

Bing Huang et al., Cell, 2019

GWAS&ScRNA-seq联合研究

Single-cell eQTLGen
Consortium

Experimental setup

Population cohort: Single cells:

Genotypes scRNA-seq

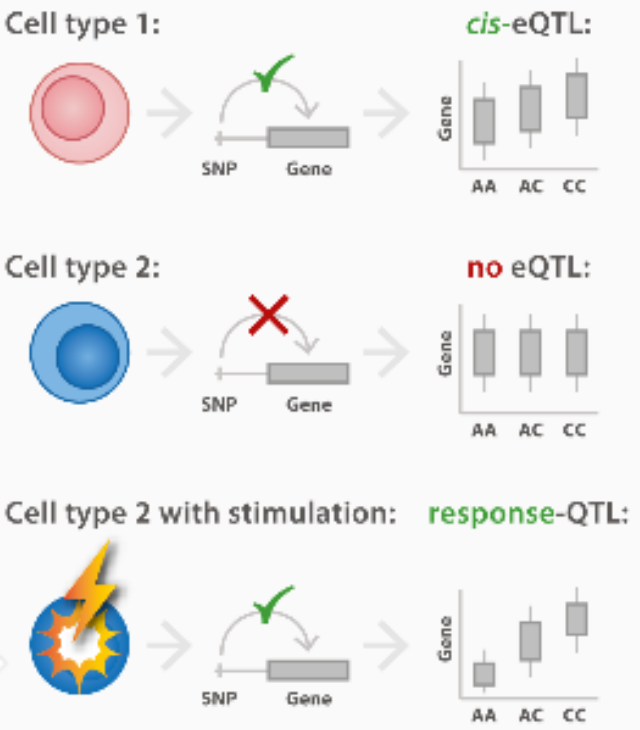
Data harmonization

QC + Normalization, cell type assign-ment, gene expression imputation

Federated eQTL analysis

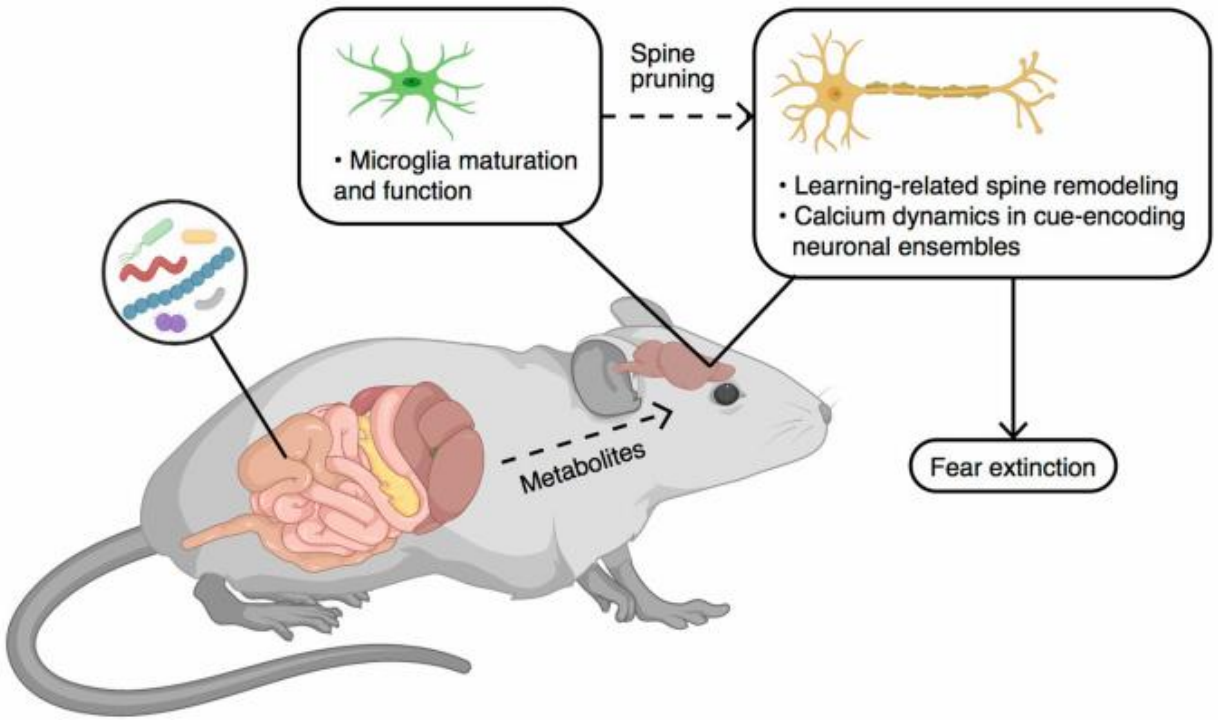
Consistent eQTL analysis per cohort, efficient meta-analysis procedure

Single-cell eQTLs (cis, trans, response)



Personalized Gene Regulatory Networks

Microbiomics&ScRNA-seq联合研究



单细胞测序

小鼠大脑的大脑前额叶皮质snRNA-seq，其中小胶质细胞基因表达的改变或许在重塑大脑在学习过程中细胞之间的连接上扮演着非常关键的角色，而这些改变并未在健康小鼠的大脑小胶质细胞中发现

微生物基因组测序

通过16S测序技术检测无菌小鼠和正常小鼠模型菌群结构

A

B

C

D

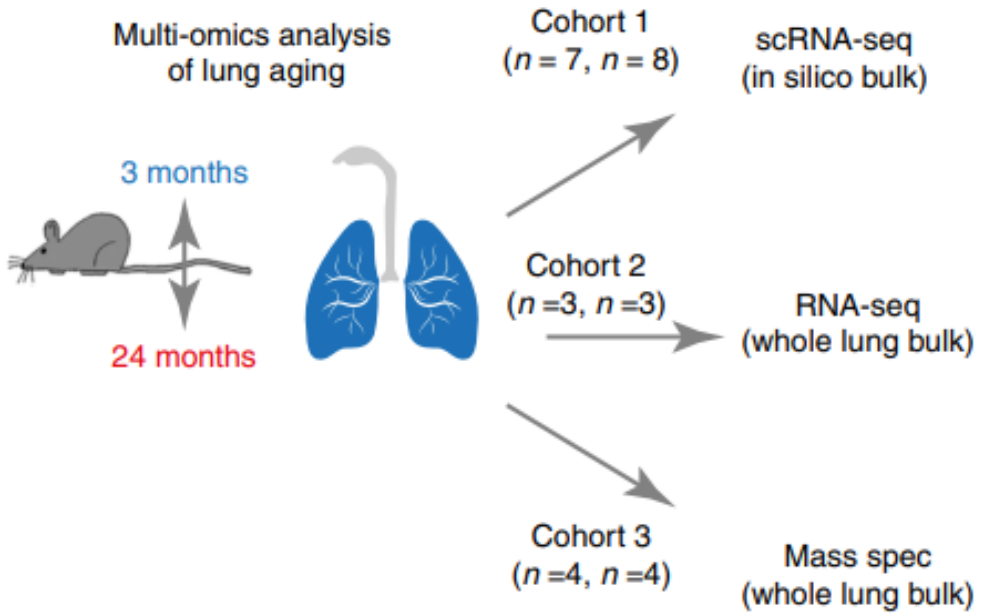
代谢组检测

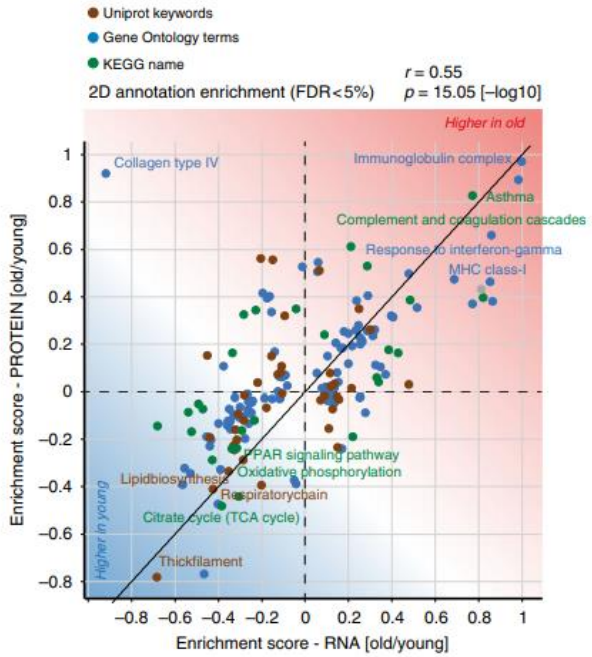
筛选四种代谢物在缺乏微生物菌群的小鼠体内含量明显低于对照组，推测菌群通过微生物源代谢物影响大脑中的神经元和小胶质细胞

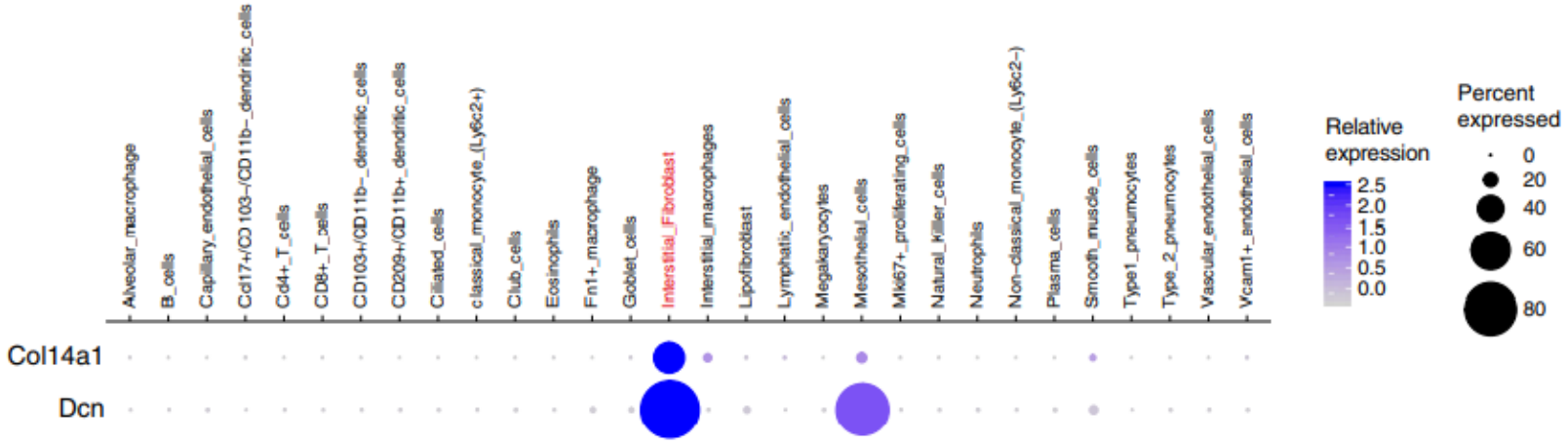
实验结论

A model in which alterations in the microbiota and their metabolites influence neuronal function and learningrelated plasticity, which may be due to altered microglia-mediated synaptic pruning, and subsequently regulate fear extinction behaviour

Proteomics&ScRNA-seq联合研究

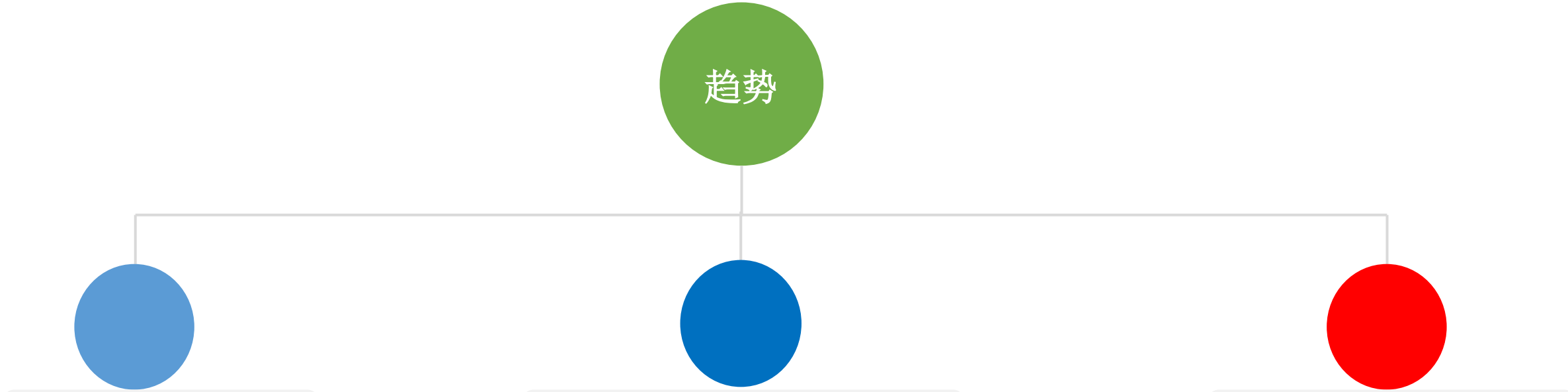






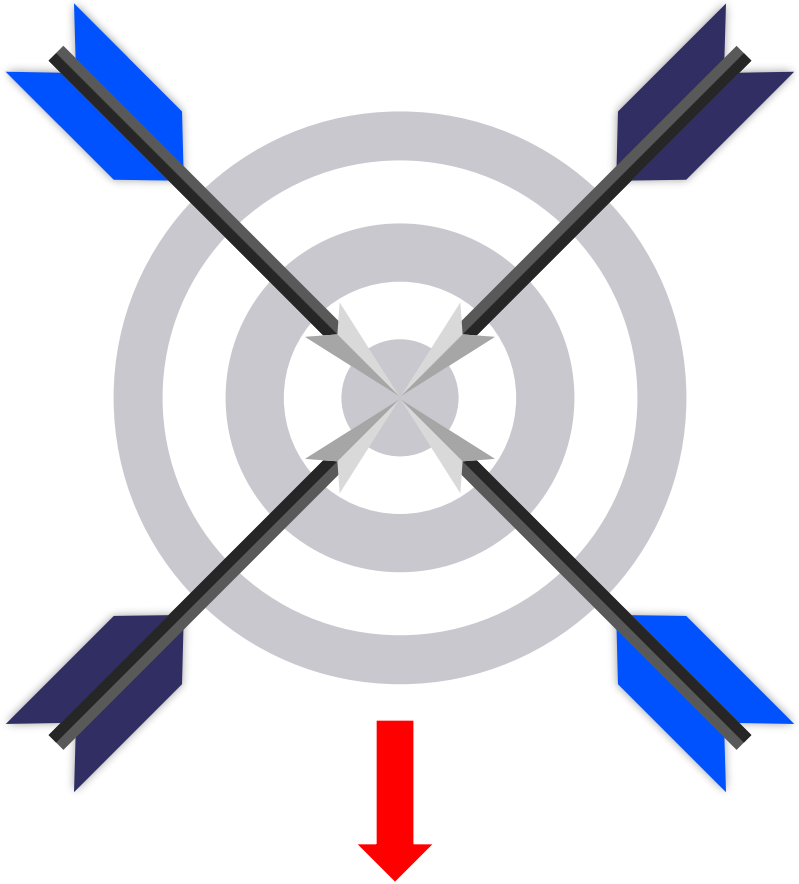
- 运用单细胞转录组和蛋白质谱分析量化了30种细胞类型的细胞活性状态变化，通过整合分析，预测了调节性蛋白的细胞来源，衰老小鼠中2型肺泡壁细胞和脂肪成纤维细胞中的胆固醇合成增加；
- 创建了公开的肺衰老细胞参考图谱；
- 揭示了肺部老化的几大标志：2型肺泡壁细胞和脂肪成纤维细胞胆固醇合成的增加，以及呼吸道上皮细胞的改变。

多维度研究体系的发展趋势



Multi-omics research elements

Skilled design



Large numbers of sample

Tailored statistical analyses

Enough Money

DATA INTERPRETATION

Cell Discov. 2021 Oct 31;7(1):103. doi: 10.1038/s41421-021-00341-7.

Discovery of new genetic loci for male sexual orientation in Han population

15 BioMiao Biological Technology (Beijing) Co., Ltd, Beijing, China.

Abstract

Epidemiological studies have demonstrated that the genetic factors partly influence the development of same-sex sexual behavior, but most genetic studies have focused on people of primarily European ancestry, potentially missing important biological insights. Here, we performed a two-stage genome-wide association study (GWAS) with a total sample of 1478 homosexual males and 3313 heterosexual males in Han Chinese populations and identified two genetic loci (rs17320865, Xq27.3, FMR1NB, $P_{meta} = 8.36 \times 10^{-8}$, OR = 1.29; rs7259428, 19q12, ZNF536, $P_{meta} = 7.58 \times 10^{-8}$, OR = 0.75) showing consistent association with male sexual orientation. A fixed-effect meta-analysis including individuals of Han Chinese (n = 4791) and European ancestries (n = 408,995) revealed 3 genome-wide significant loci of same-sex sexual behavior (rs9677294, 2p22.1, SLC8A1, $P_{meta} = 1.95 \times 10^{-8}$; rs2414487, 15q21.3, LOC145783, $P_{meta} = 4.53 \times 10^{-9}$; rs2106525, 7q31.1, MDFIC, $P_{meta} = 6.24 \times 10^{-9}$). These findings may provide new insights into the genetic basis of male sexual orientation from a wider population scope. Furthermore, we defined the average ZNF536-immunoreactivity (ZNF536-ir) concentration in the suprachiasmatic nucleus (SCN) as lower in homosexual individuals than in heterosexual individuals (0.011 ± 0.001 vs 0.021 ± 0.004, P = 0.013) in a postmortem study. In addition, compared with heterosexuals, the percentage of ZNF536 stained area in the SCN was also smaller in the homosexuals (0.075 ± 0.040 vs 0.137 ± 0.103, P = 0.043). More homosexual preference was observed in FMR1NB-knockout mice and we also found significant differences in the expression of serotonin, dopamine, and inflammation pathways that were reported to be related to sexual orientation when comparing CRISPR-mediated FMR1NB knockout mice to matched wild-type target C57 male mice.

npj | parkinson's disease

www.nature.com/npjparkd

ARTICLE OPEN

Check for updates

Genome-wide association study using whole-genome sequencing identifies risk loci for Parkinson's disease in Chinese population

Genome-wide association studies (GWASs) have identified numerous susceptibility loci for Parkinson's disease (PD), but its genetic architecture remains underexplored in populations of non-European ancestry. To identify genetic variants associated with PD in the Chinese population, we performed a GWAS using whole-genome sequencing (WGS) in 1,972 cases and 2,478 controls, and a replication study in a total of 8209 cases and 9454 controls. We identified one new risk variant rs61204179 ($P_{combined} = 1.47 \times 10^{-9}$) with low allele frequency, four previously reported risk variants (*NUCKS1/RAB29*-rs11557080, *SNCA*-rs356182, *FYN*-rs997368, and *VPS13C*-rs2251086), as well as three risk variants in *LRRK2* coding region (A419V, R1628P, and G2385R) with genome-wide significance ($P < 5 \times 10^{-8}$) for PD in Chinese population. Moreover, of the reported genome-wide significant risk variants found mostly in European ancestry populations, the correlation coefficient (r_b) of effect size accounting for sampling errors was 0.91 between datasets and 63.6% attained $P < 0.05$ in Chinese population. Accordingly, we estimated a heritability of 0.14–0.18 for PD, and a moderate genetic correlation between European ancestry and Chinese populations ($r_g = 0.47$, se $= 0.21$). Polygenic risk score (PRS) analysis revealed that individuals with PRS values in the highest quartile had a 3.9-fold higher risk of developing PD than the lowest quartile. In conclusion, the present GWAS identified PD-associated variants in Chinese population, as well as genetic factors shared among distant populations. Our findings shed light on the genetic homogeneity and heterogeneity of PD in different ethnic groups and suggested WGS might continue to improve our understanding of the genetic architecture of PD.

npj Parkinson's Disease (2023)9:22 ; https://doi.org/10.1038/s41531-023-00456-6

博淼部分组学代表文献（2021-2023）——DNA甲基化表观调控机制研究

PNAS RESEARCH ARTICLE | MEDICAL SCIENCES OPEN ACCESS

Elevated *BICD2* DNA methylation in blood of major depressive disorder patients and reduction of depressive-like behaviors in hippocampal *Bicd2*-knockdown mice

Jianbo Xiu[a,b,1], Jiayu Li[a,b,1], Zeyue Liu[a,b,1], Hui Wei[a,b], Caiyun Zhu[a,b], Rongrong Han[a,b], Zijing Liu[a,b], Wanwan Zhu[a,b], Yan Shen[a,b], and Qi Xu[a,b,2]

Edited by Joseph Takahashi, University of Texas Southwestern Medical Center, Dallas, TX; received February 3, 2022; accepted June 7, 2022

Major depressive disorder (MDD) is a prevalent and devastating mental illness. To date, the diagnosis of MDD is largely dependent on clinical interviews and questionnaires and still lacks a reliable biomarker. DNA methylation has a stable and reversible nature and is likely associated with the course and therapeutic efficacy of complex diseases, which may play an important role in the etiology of a disease. Here, we identified and validated a DNA methylation biomarker for MDD from four independent cohorts of the Chinese Han population. First, we integrated the analysis of the DNA methylation microarray ($n = 80$) and RNA expression microarray data ($n = 40$) and identified *BICD2* as the top-ranked gene. In the replication phase, we employed the Sequenom MassARRAY method to confirm the DNA hypermethylation change in a large sample size ($n = 1{,}346$) and used the methylation-sensitive restriction enzymes and a quantitative PCR approach (MSE-qPCR) and qPCR method to confirm the correlation between DNA hypermethylation and mRNA down-regulation of *BICD2* ($n = 60$). The results were replicated in the peripheral blood of mice with depressive-like behaviors, while in the hippocampus of mice, *Bicd2* showed DNA hypomethylation and mRNA/protein up-regulation. Hippocampal *Bicd2* knockdown demonstrates antidepressant action in the chronic unpredictable mild stress (CUMS) mouse model of depression, which may be mediated by increased BDNF expression. Our study identified a potential DNA methylation biomarker and investigated its functional implications, which could be exploited to improve the diagnosis and treatment of MDD.

Sequenom MassARRAY Methylation Analysis. The Sequenom MassARRAY platform (BioMiao Biological Technology) was used to quantitatively examine methylation according to the protocol recommended by the manufacturer (38). The genomic DNA was bisulfite-converted using the EZ-96 DNA Methylation-Gold Kit (D5008, Zymo Research). The primers were designed by EpiDesigner software (www.epidesigner.com). The sequences of the primers of the target sequence were 5′-aggaagagagAGTTTGGTAGTTAGGGAAGAAGGGT-3′ (forward primer) and 5′-cagtaatacgactcactatagggagaaggctACTTTACAAACACCCAAACCACTAA-3′ (reverse primer). The T7 promoter sequence was added to the PCR products, and in vitro RNA transcription was then performed and processed by base-specific cleavage. Small RNA fragments with CpG sites were obtained. The time of flight mass spectrometry (MALDI-TOF MS) was used to determine the molecular weights of each fragment. The methylation level of individual units was measured by quantitative methylation analysis (Sequenom). A linear regression model was used on the Sequenom MassARRAY methylation data to analyze the differences in methylation levels between groups (MDD and control) with adjustment for sex (male and female) and age.

Received: 4 December 2020 | Revised: 24 August 2021 | Accepted: 30 August 2021

DOI: 10.1002/alz.12484

RESEARCH ARTICLE

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

A metabolite panel that differentiates Alzheimer's disease from other dementia types

Abstract

Introduction: Alzheimer's disease (AD) is associated with altered metabolites. This study aimed to determine the validity of using circulating metabolites to differentiate AD from other dementias.

Methods: Blood metabolites were measured in three data sets. Data set 1 (controls, 27; AD, 28) was used for analyzing differential metabolites. Data set 2 (controls, 93; AD, 92) was used to establish a diagnostic AD model with use of a metabolite panel. The model was applied to Data set 3 (controls, 76; AD, 76; other dementias, 205) to verify its capacity for differentiating AD from other dementias.

Results: Data set 1 revealed 7 upregulated and 77 downregulated metabolites. In Data set 2, a panel of 11 metabolites was included in a model that could distinguish AD from controls. In Data set 3, this panel was used to successfully differentiate AD from other dementias.

Discussion: This study revealed an AD-specific panel of 11 metabolites that may be used for AD diagnosis.

ACKNOWLEDGMENTS

The authors thank Zeming Wu from iPhenome (Yunpukang) Biotechnology Inc. and BioMiao Biological Technology (Beijing) Company for technically supporting the experiments. This study was financially supported by Beijing Brain Initiative from Beijing Municipal Science & Technology Commission (Z201100005520016); National Natural Science Foundation of China (81870825, 82071194); and Beijing Municipal Natural Science Foundation (7202061).

博淼部分组学代表文献（2021-2023）——外泌体蛋白质组标记物模型研究

Cai et al. Alzheimer's Research & Therapy (2022) 14:181
https://doi.org/10.1186/s13195-022-01133-1

Alzheimer's
Research & Therapy

RESEARCH Open Access

Proteomic profiling of circulating plasma exosomes reveals novel biomarkers of Alzheimer's disease

Abstract

Background: Neuronal- and astrocyte-derived exosomes have been identified as an optimal source for screening biomarkers for Alzheimer's disease (AD). However, few studies focus on the bulk exosome population isolated from plasma of AD. This study investigated whether proteins in bulk exosomes can aid in the diagnosis of AD.

Methods: The plasma exosomes were collected by ultracentrifuge. Protein samples were extracted from exosomes. Cerebrospinal fluid levels of amyloid β (Aβ)42 and phosphorylated tau (P-tau)181 were measured for diagnostic purposes. A pilot study (controls, 20; AD, 20) followed by a second dataset (controls, 56; AD, 58) was used to establish a diagnostic model of AD. Mass spectrometry-based proteomics was performed to profile the plasma exosomal proteome. Parallel reaction monitoring was used to further confirm the differentially expressed proteins.

Results: In total, 328 proteins in plasma exosomes were quantified. Among them, 31 proteins were altered in AD patients, and 12 were validated. The receiver operating characteristic curve analysis revealed a combination of six proteins (upregulated: Ig-like domain-containing protein (A0A0G2JRQ6), complement C1q subcomponent subunit C (C1QC), complement component C9 (CO9), platelet glycoprotein Ib beta chain (GP1BB), Ras suppressor protein 1 (RSU1); downregulated: disintegrin and metalloproteinase domain 10 (ADA10)) has the capacity to differentiate AD patients from healthy controls with high accuracy. Linear correlation analysis showed that the combination was significantly correlated with cognitive performance.

Conclusions: The combination of plasma exosomal proteins A0A0G2JRQ6, C1QC, CO9, GP1BB, RSU1, and ADA10 acts as a novel candidate biomarker to differentiate AD patients from healthy individuals.

Keywords: Alzheimer's disease, Exosome, Proteomics, Biomarker, Diagnosis

Acknowledgements

We thank BioMiao Biological Technology (Beijing) Company for technically supporting the experiments.

博淼部分组学代表文献（2021-2023）——单细胞转录组分子病因机制研究

Exp Mol Med. 2023 Mar; 55(3): 597–611.
Published online 2023 Mar 6. doi: 10.1038/s12276-023-00957-7
PMCID: PMC10073150
PMID: 36879115

Increased retinoic acid signaling decreases lung metastasis in salivary adenoid cystic carcinoma by inhibiting the noncanonical Notch1 pathway

MYB-NFIB fusion and *NOTCH1* mutation are common hallmark genetic events in salivary gland adenoid cystic carcinoma (SACC). However, abnormal expression of MYB and NOTCH1 is also observed in patients without *MYB-NFIB* fusion and *NOTCH1* mutation. Here, we explore in-depth the molecular mechanisms of lung metastasis through single-cell RNA sequencing (scRNA-seq) and exome target capture sequencing in two SACC patients without *MYB-NFIB* fusion and *NOTCH1* mutation. Twenty-five types of cells in primary and metastatic tissues were identified via Seurat clustering and categorized into four main stages ranging from near-normal to cancer-based on the abundance of each cell cluster in normal tissue. In this context, we identified the Notch signaling pathway enrichment in almost all cancer cells; RNA velocity, trajectory, and sub-clustering analyses were performed to deeply investigate cancer progenitor-like cell clusters in primary tumor-associated lung metastases, and signature genes of progenitor-like cells were enriched in the "MYC_TARGETS_V2" gene set. In vitro, we detected the NICD1-MYB-MYC complex by co-immunoprecipitation (Co-IP) and incidentally identified retinoic acid (RA) as an endogenous antagonist of genes in the "MYC_TARGETS_V2" gene set. Following this, we confirmed that all-trans retinoic acid (ATRA) suppresses the lung metastasis of SACC by correcting erroneous cell differentiation mainly caused by aberrant NOTCH1 or MYB expression. Bioinformatic, RNA-seq, and immunohistochemical (IHC) analyses of primary tissues and metastatic lung tissues from patients with SACC suggested that RA system insufficiency partially promotes lung metastasis. These findings imply the value of the RA system in diagnosis and treatment.

Single-cell sorting, t-distributed stochastic neighbor embedding (t-SNE), and cell annotation

Cells from fresh tumor tissues were isolated for the preparation of single-cell suspensions via the Chromium™ Single Cell 3' Solution technique, and then analyzed by BioMiao Biological Technology Co., Ltd. (Beijing). Raw data (150–200 Gb) were obtained from six samples. The following numbers

cancers

MDPI

Article

Expansion of Colorectal Cancer Biomarkers Based on Gut Bacteria and Viruses

Abstract: The alterations in gut bacteria are closely related to colorectal cancer. However, studies on adenoma are still scarce. Besides, the associations of gut viruses with colorectal tumor, and the interactions of bacteria with viruses in colorectal tumors are still under exploration. Therefore, a metagenomic sequencing of stool samples from patients with colorectal adenoma (CRA), colorectal cancer (CRC), and healthy controls was performed to identify changes in gut microbiome in patients with colorectal tumors. Five CRC-enriched bacteria (*Peptostreptococcus stomatis*, *Clostridium symbiosum*, *Hungatella hathewayi*, *Parvimonas micra*, and *Gemella morbillorum*) were identified as a diagnostic model to identify CRC patients, and the efficacy of the diagnostic model was verifiable in 1523 metagenomic samples from ten cohorts of eight different countries. We identified the positive association of *Bacteroides fragilis* with PD-L1 expression and PD-1 checkpoint pathway, providing a possible direction for studying bacterial carcinogenesis mechanisms. Furthermore, the increased interactions within the microbiome in patients may play roles in the development of CRC. In conclusion, this study identified novel microbiota combinations with discrimination for colorectal tumor, and revealed the potential interactions of gut bacteria with viruses in the adenoma-carcinoma sequence, which implies that the microbiome, but not only bacteria, should be paid more attention in further studies.

博淼技术服务项目一览图

博淼生物
BIO MIAO BIOLOGICAL
SINCE 2009

基因组学服务

- GWAS芯片/WES-seq/Target NGS-seq
- Massarray /Multi-PCR NGS/Taqman /KASP SNP分型
- 16S扩增子测序/宏基因组测序/微生物简化基因组测序
- HLA-seq/TCR&BCR-seq

表观基因组学服务

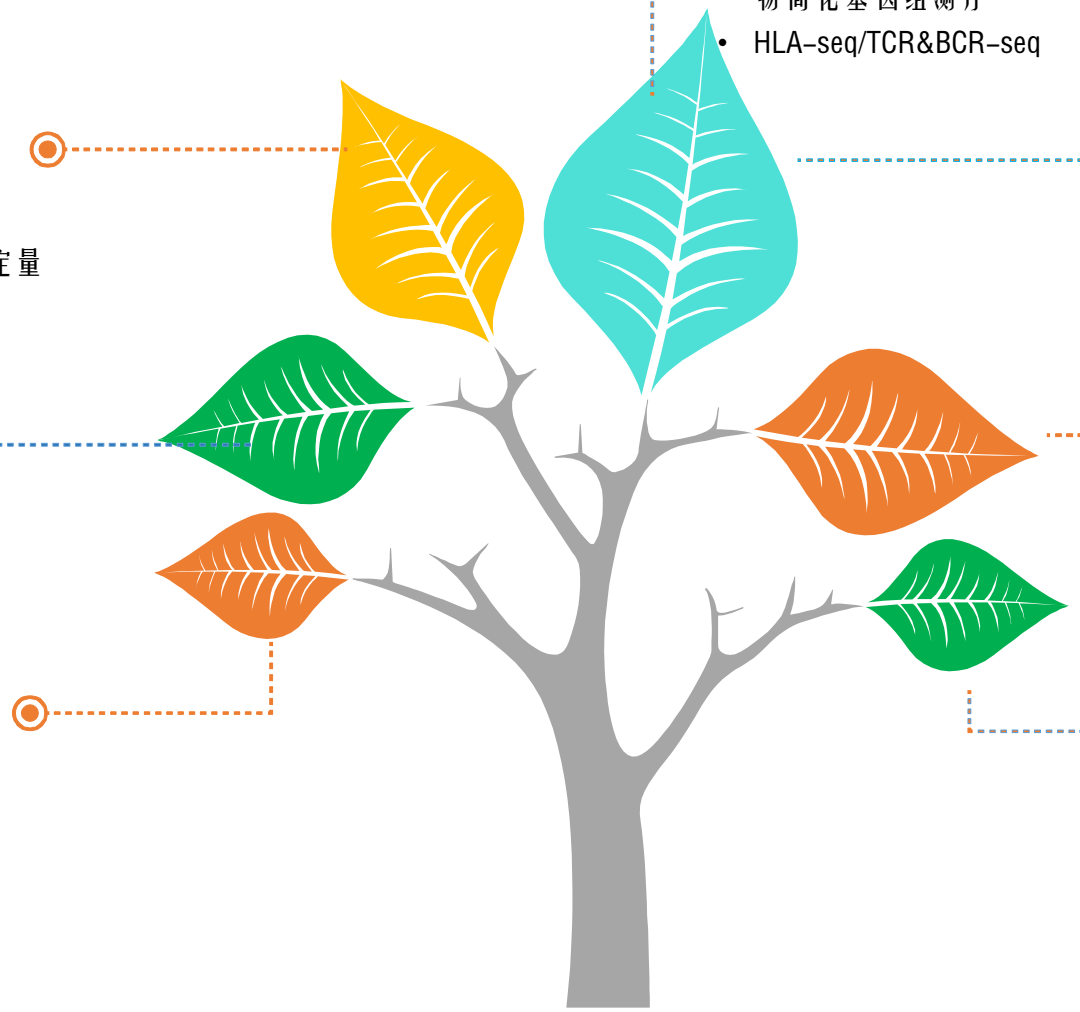
- EWAS芯片
- Multi-PCR NGS靶向DNA甲基化定量
- Massarray 靶向DNA甲基化定量

转录组学服务

- 转录组NGS
- 表达谱芯片
- RT-qPCR靶向转录本定量

单细胞组学服务

- 单细胞转录组测序
- 单细胞免疫组库测序
- 单细胞ATAC测序&转录组测序
- 空间转录组测序



代谢组学服务

- 非靶向代谢组
- 非靶向脂质组
- 高通量靶向代谢组
- 靶向代谢组项目系列

蛋白质组学服务

- 4D-DIA/DIA定量蛋白质组/TMT定量蛋白质组/Label free定量蛋白质组
- Olink定量蛋白质组/PRM靶向蛋白/ELISA
- 修饰蛋白质组
- 高密度自身抗体蛋白芯片

多组学联合研究服务

- GWAS&多组学技术服务
- EWAS&多组学技术服务
- 微生物基因组&代谢组技术服务
- 蛋白质组&代谢组技术服务

感谢各位的聆听

Your own Laboratory

——您的专属实验室