

DNA甲基化研究整体解决方案

全国统一服务电话：400-6506-908

网址：www.biomiao.com

邮箱：marketing@biomiao.com

地址：北京市丰台区西三环优橙创新中心B座3012-3015

CONTENTS
目
录

01 . DNA甲基化理论体系

02 . 科研策略云集剖析

03 . 甲基化注释数据库介绍

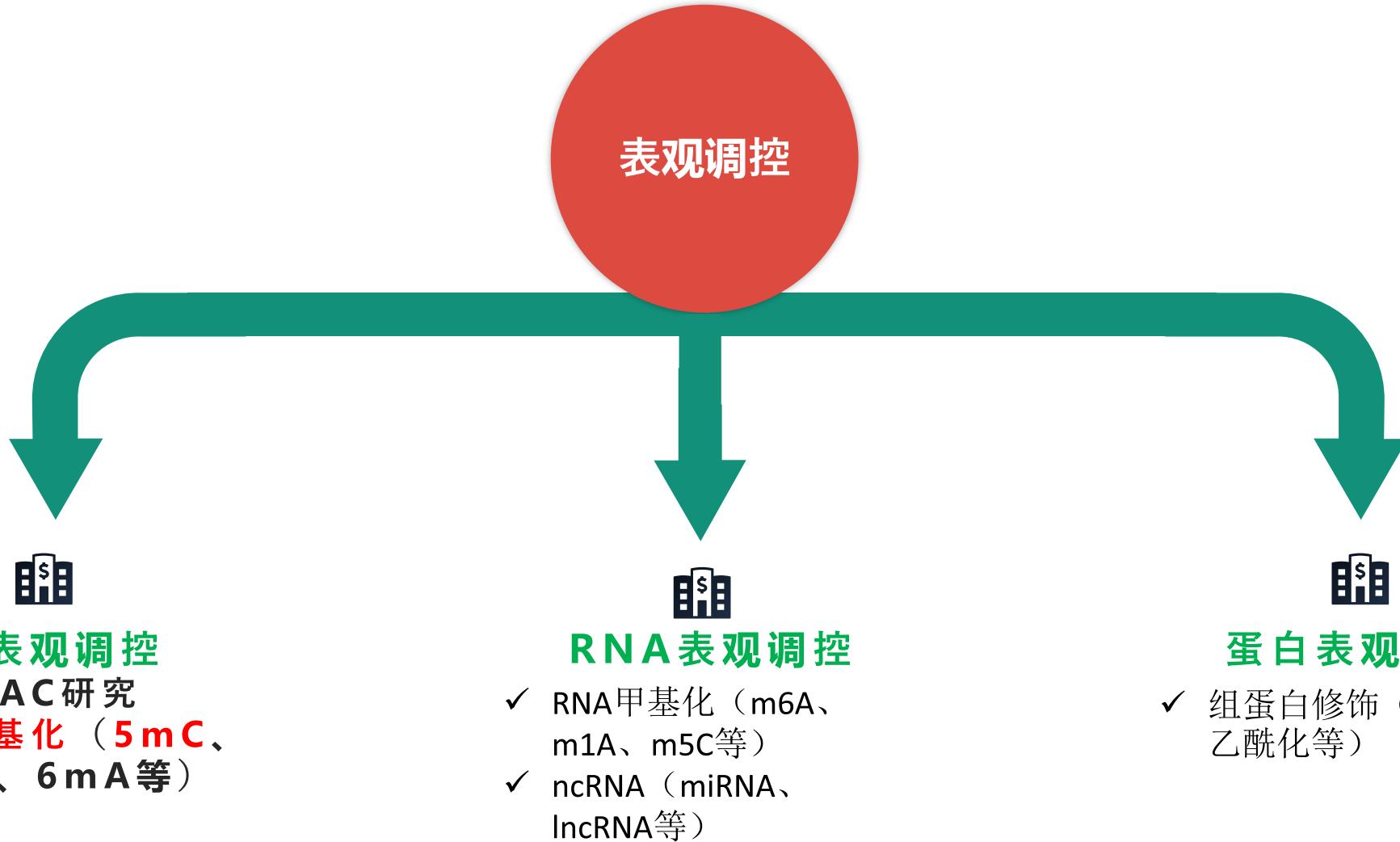
04 . 经典检测技术比拼

05 . 代表文献解读

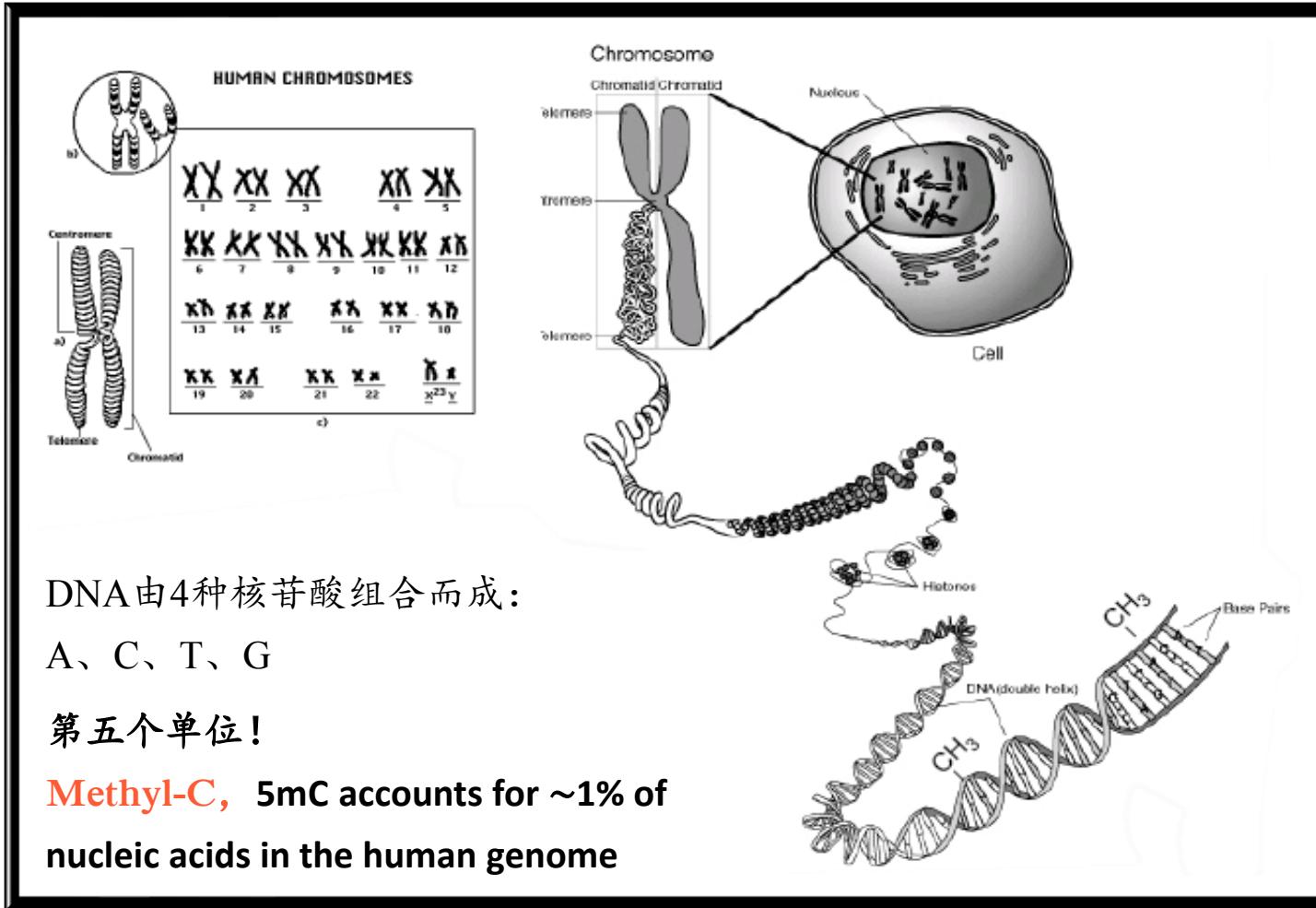
01

DNA甲基化理论体系

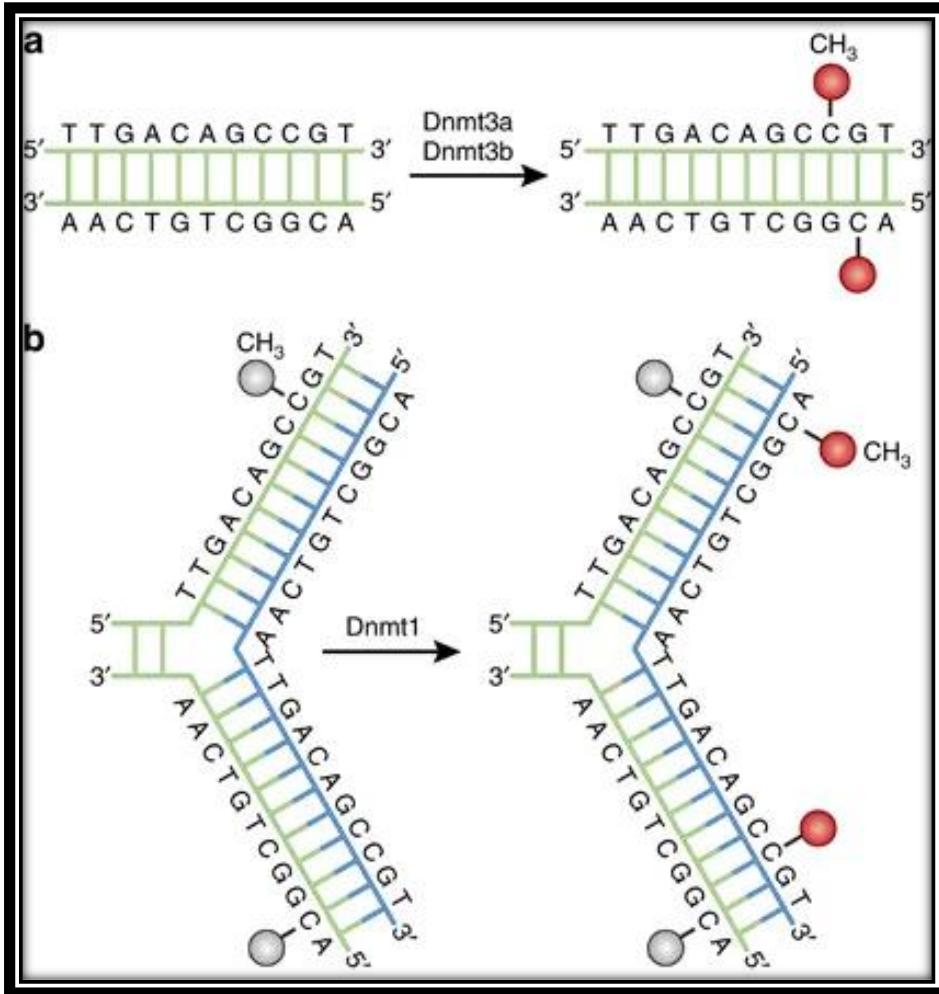
>>> 理论体系



>>> 理论体系



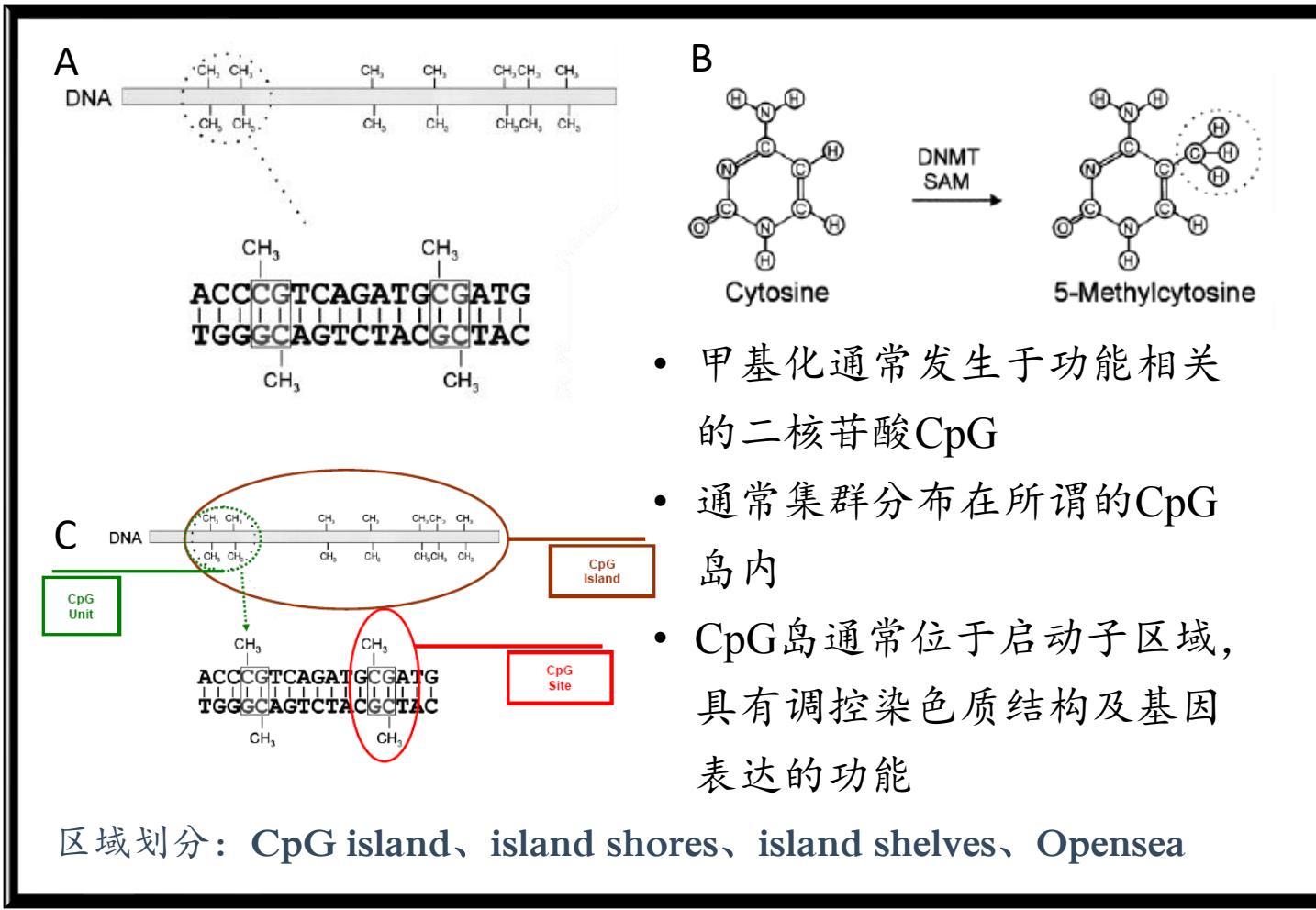
>>> 理论体系



DNA methylation pathways



>>> 理论体系



Intergenic Regions

- ✓ Repress potentially harmful genetic elements expression
- ✓ Repress ncRNA expression

CpG Islands/CpG shores_Promoter&1stExon

- ✓ The majority of gene promoters, roughly 70%, reside within CpG islands, ~50% of CpG islands contain known transcription start sites
- ✓ CpG islands promote gene expression by regulating the chromatin structure and transcription factor binding(rich GC)
- ✓ CpG shores have highly conserved patterns of tissue-specific methylation
- ✓ The methylation of CpG shores is highly correlated with reduced gene expression

Gene Body

- ✓ DNA methylation of the gene body is associated with a higher level of gene expression in dividing cells ,or negatively correlated with gene expression
- ✓ Still unclear

The role of DNA methylation
in different genomic regions

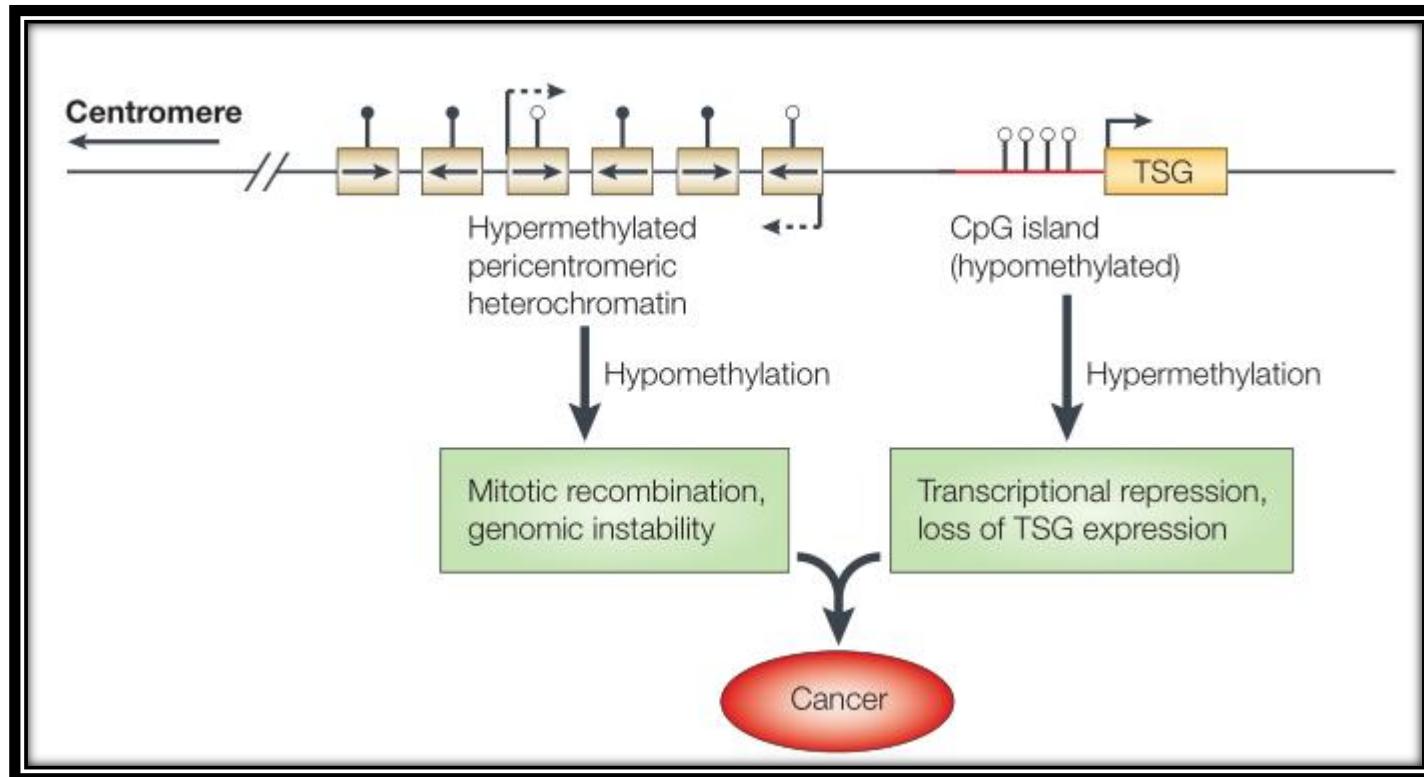


- ✓ DNMTs may recruit histone deacetylase and histone methylase resulting in transcriptional repression
- ✓ DNA methylation can directly decrease expression by preventing transcriptional factors from binding to the DNA
- ✓ DNA methylation can repress transcriptional elongation caused by reduced RNA polymerase II occupancy and chromatin accessibility over the gene body
- ✓ methyl-CpG-binding proteins (MBPs) can identify methylated DNA and recruit corepressors in order to silence the transcription and alter the surrounding chromatin

The mechanisms of hypermethylation Decreased transcription



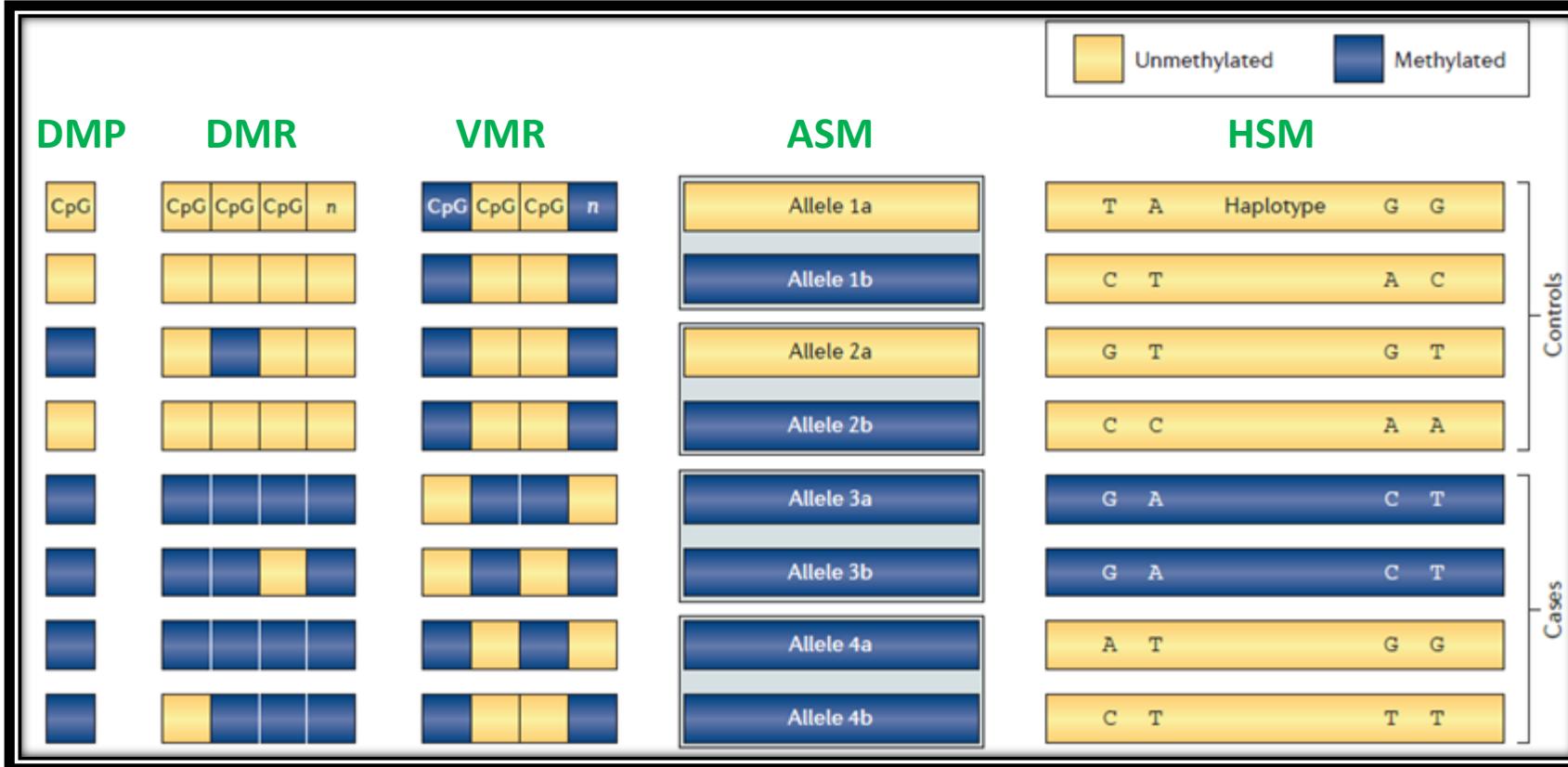
>>> 理论体系



DNA methylation and cancer



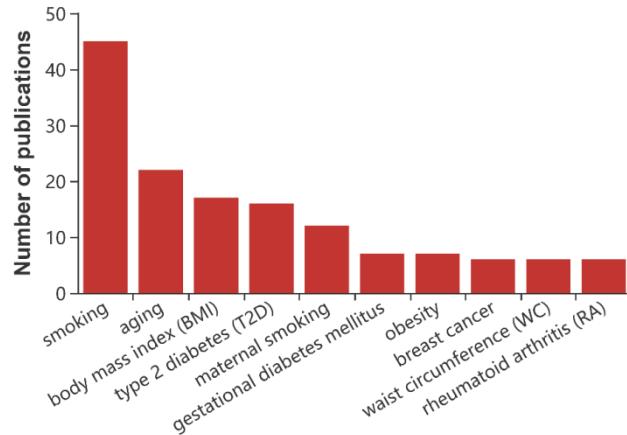
>>> 理论体系



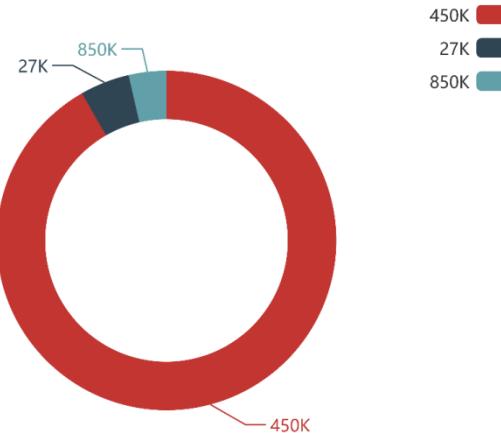
- iDMR — imprinting-specific differentially methylated region
- tDMR — tissue-specific differentially methylated region
- rDMR — reprogramming-specific differentially methylated region
- cDMR — cancer-specific differentially methylated region
- aDMR — ageing-specific differentially methylated region

>>> 理论体系

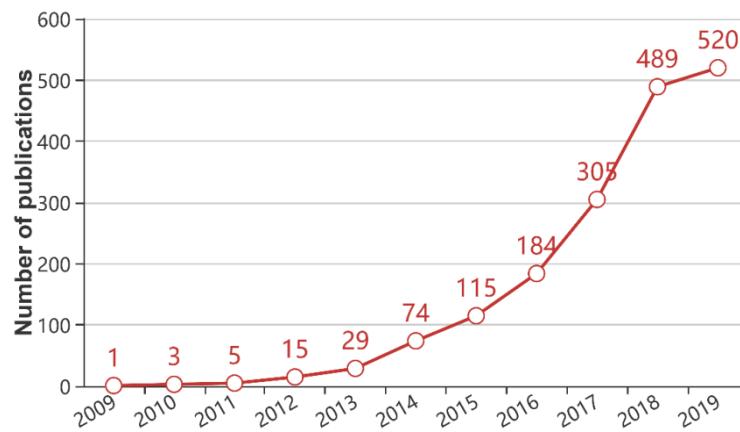
TOP 10 Traits



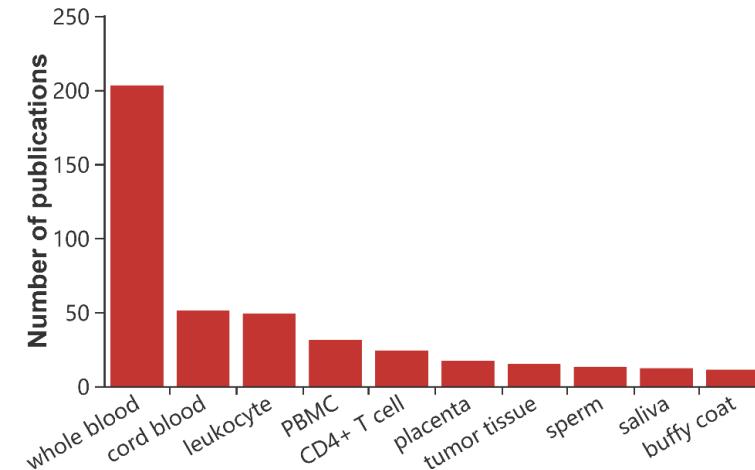
Platforms



Number of Publications



TOP 10 tissues



EWAS研究动态



>>> 理论体系

特异甲基化区域及CpG site定量研究

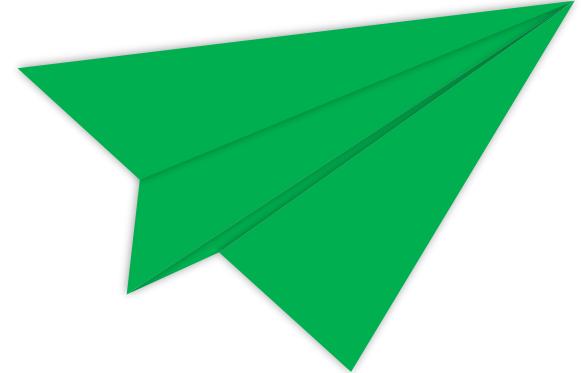
- 焦磷酸测序技术
- Massarray质谱技术
- Multi-PCR NGS技术

甲基化定性研究技术
✓ 数据准确度低
✓ 无法精确到CpG site

全基因组水平CpG site定量研究

- 甲基化芯片——93.5万个CpG site
- 全基因组甲基化NGS测序——WGBS
- 简化全基因组甲基化NGS测序——RRBS

检测技术革新



02

科研策略云集剖析



DNA甲基化科研领域



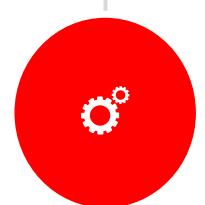
甲基化分子功能机制研究

- ✓ DNA甲基化 write\erase\read机制研究
- ✓ 与基因转录、转录因子结合、组蛋白、miRNA功能机制



甲基化标记物研究

- ✓ 疾病预警、个体化治疗、环境暴露干预等表观基因组标记物筛选及模型建立
- ✓ 表观易感基因发现



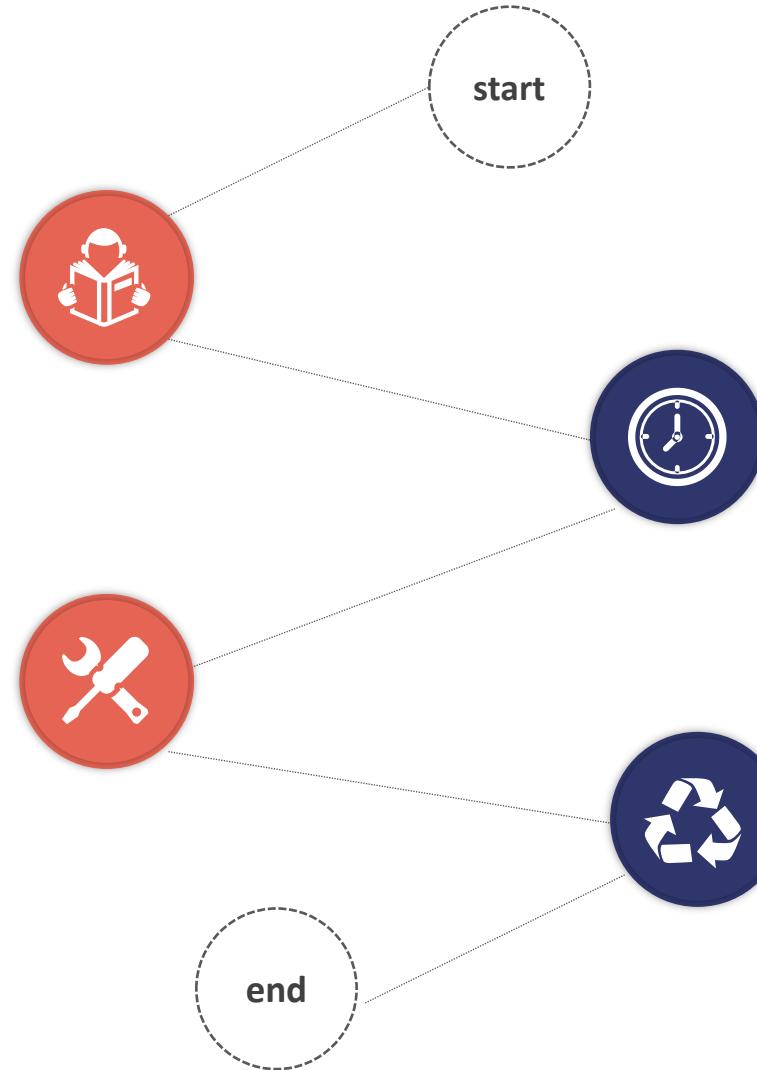
甲基化组学调控机制研究

- ✓ 暴露干预-多组学-临床表型中介调控机制及因果效应
- ✓ 系统调控网络构建

>>> 科研策略—EWAS方案路线

科研设置

- ✓ 疾病类型、药物干预、环境暴露等
- ✓ 样本类型外周血、组织、细胞
- ✓ 技术路线：芯片、测序及对应数据分析策略



EWAS检测

- ✓ EPIC芯片检测
- ✓ EWAS数据分析

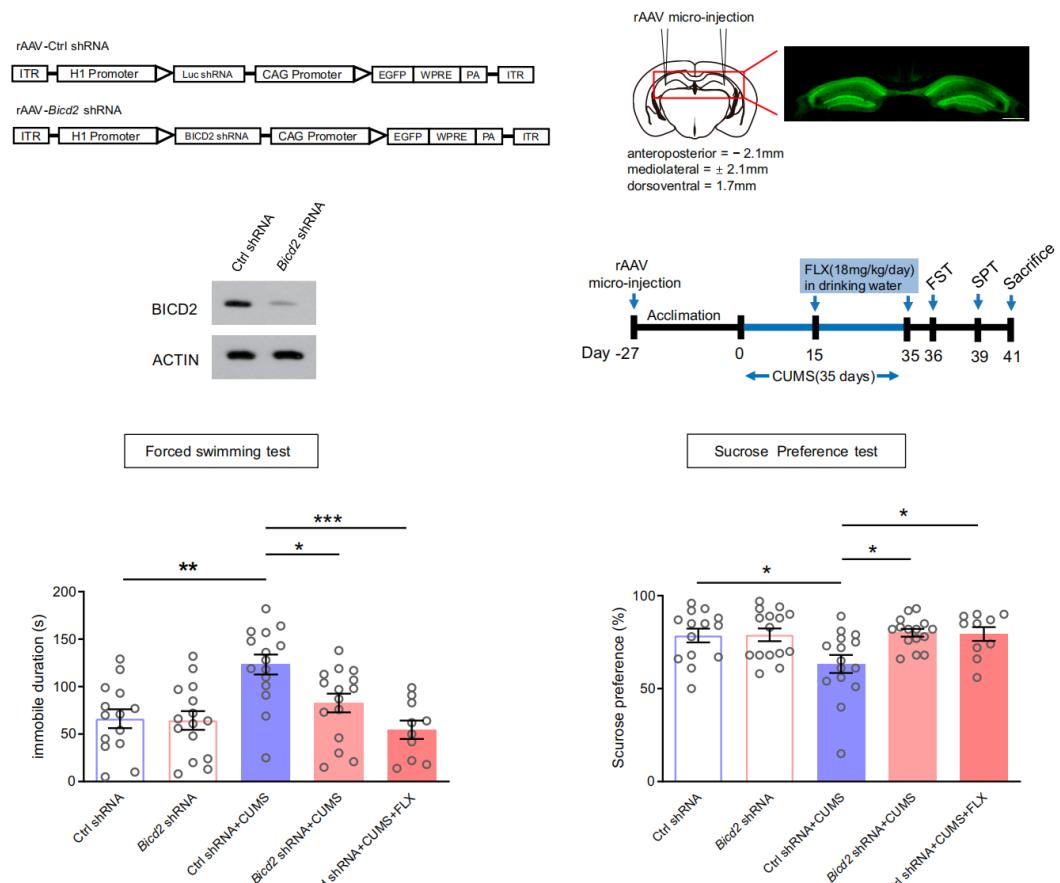
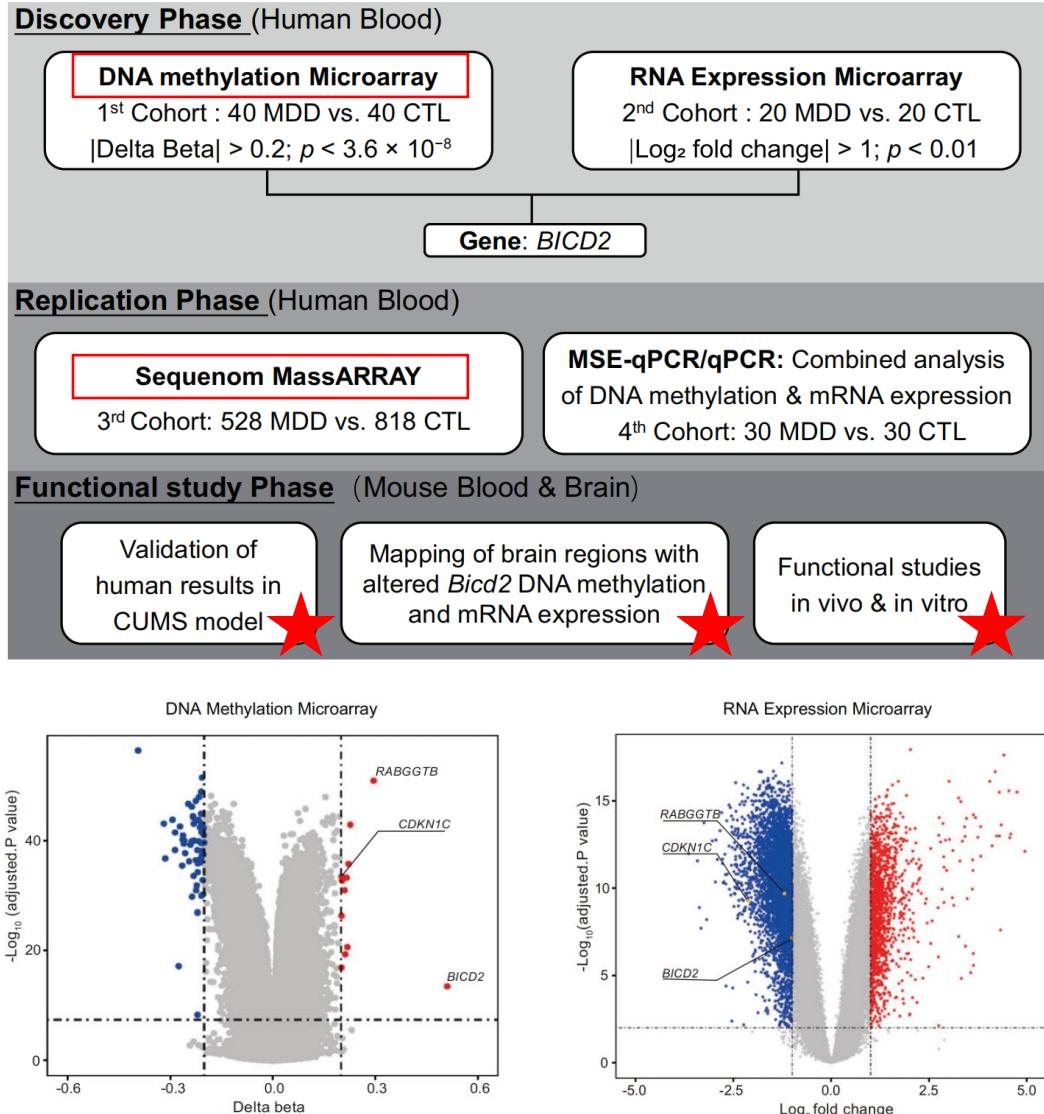
Validation检测

- ✓ 采用二期独立样本进行EWAS靶向位点验证
- ✓ 采用Massarray/MultiPCR NGS技术检测

生物信息分析

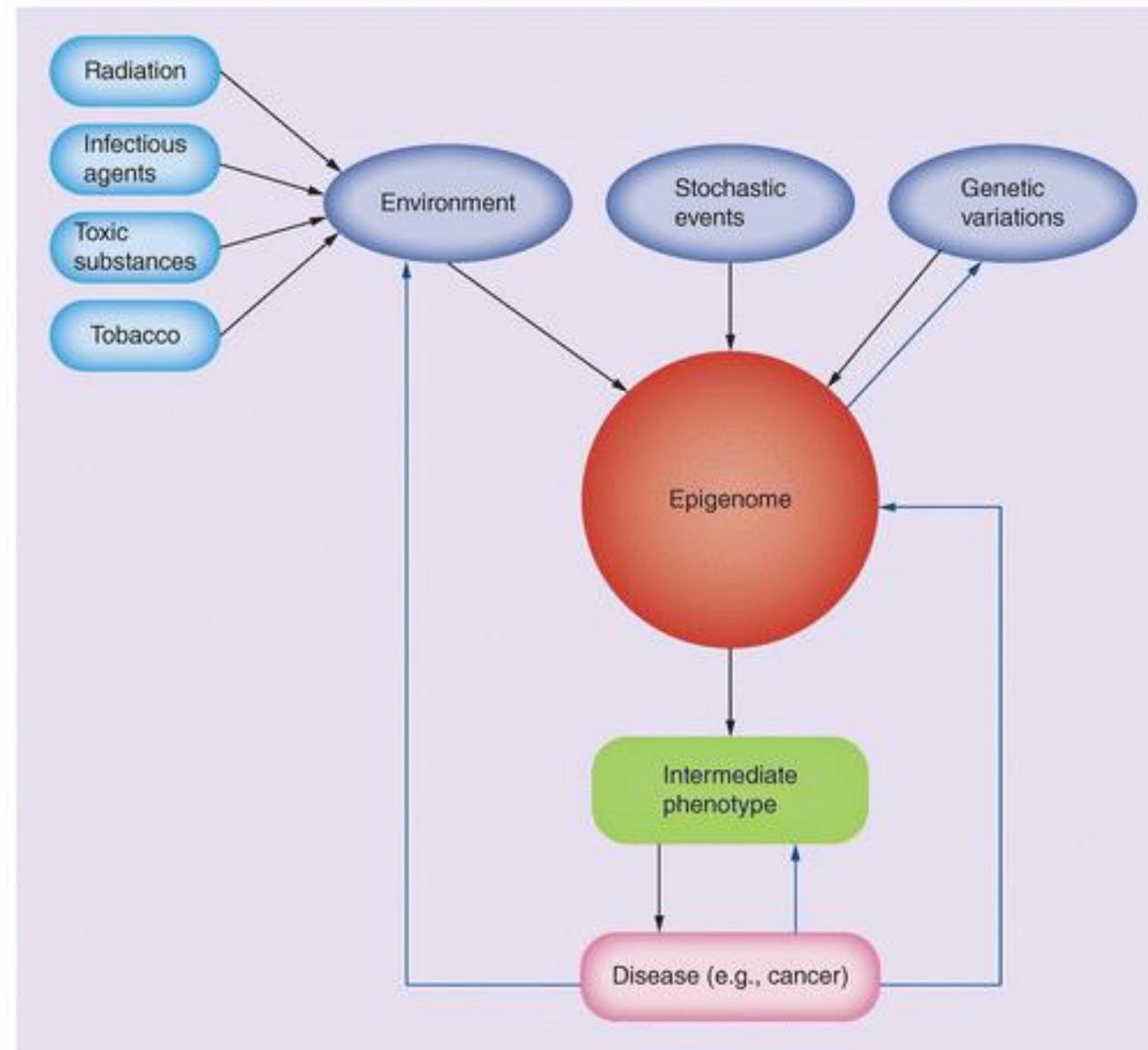
- ✓ 关联分析/meta分析
- ✓ 因果效应分析/中介调控分析
- ✓ 预测模型构建
- ✓ In silico调控功能注释QTM等等

>>> 科研策略—EWAS方案路线

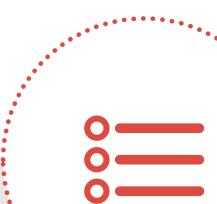


Elevated BICD2 DNA methylation in blood of major depressive disorder patients and reduction of depressive-like behaviors in hippocampal Bicd2-knockdown mice *PNAS* 2022 July

>>> 科研策略—EWAS方案路线之科研设置



>>> 科研策略—EWAS方案路线之科研设置



- ✓ 肿瘤疾病：癌组织/癌旁组织
最佳
- ✓ 脑类/内分泌/感染等类疾病：
全血、唾液
- ✓ 生殖相关疾病：胚胎组织、脐带血

疾病类型



- ✓ 肿瘤靶向药：癌组织
- ✓ 常见疾病用药：全血、PBMC

药物干预



- ✓ 队列样本：组织、全血
- ✓ 机制研究：特定细胞

环境暴露

>>> 科研策略—EWAS方案路线之科研设置



样本类型设置：

- ✓ Peripheral tissues can be used as surrogates for DNA methylation in the brain ?
- ✓ Blood, saliva, buccal, and live brain tissue

01



EWAS芯片数据：

- ✓ HumanMethylation450 (n = 12)
- ✓ HumanMethylationEPIC BeadChip arrays (n= 21)

02



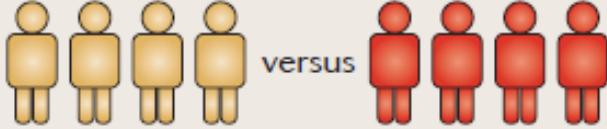
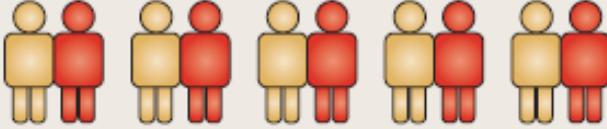
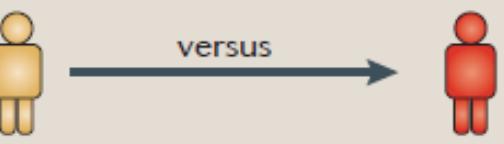
样本间数据相关性比对：

- ✓ saliva–brain correlation ($r= 0.90$) 、 blood–brain ($r =0.86$)、 buccal–brain ($r=0.85$)
- ✓ blood had the highest proportion of CpGs correlated to brain (20.8%)

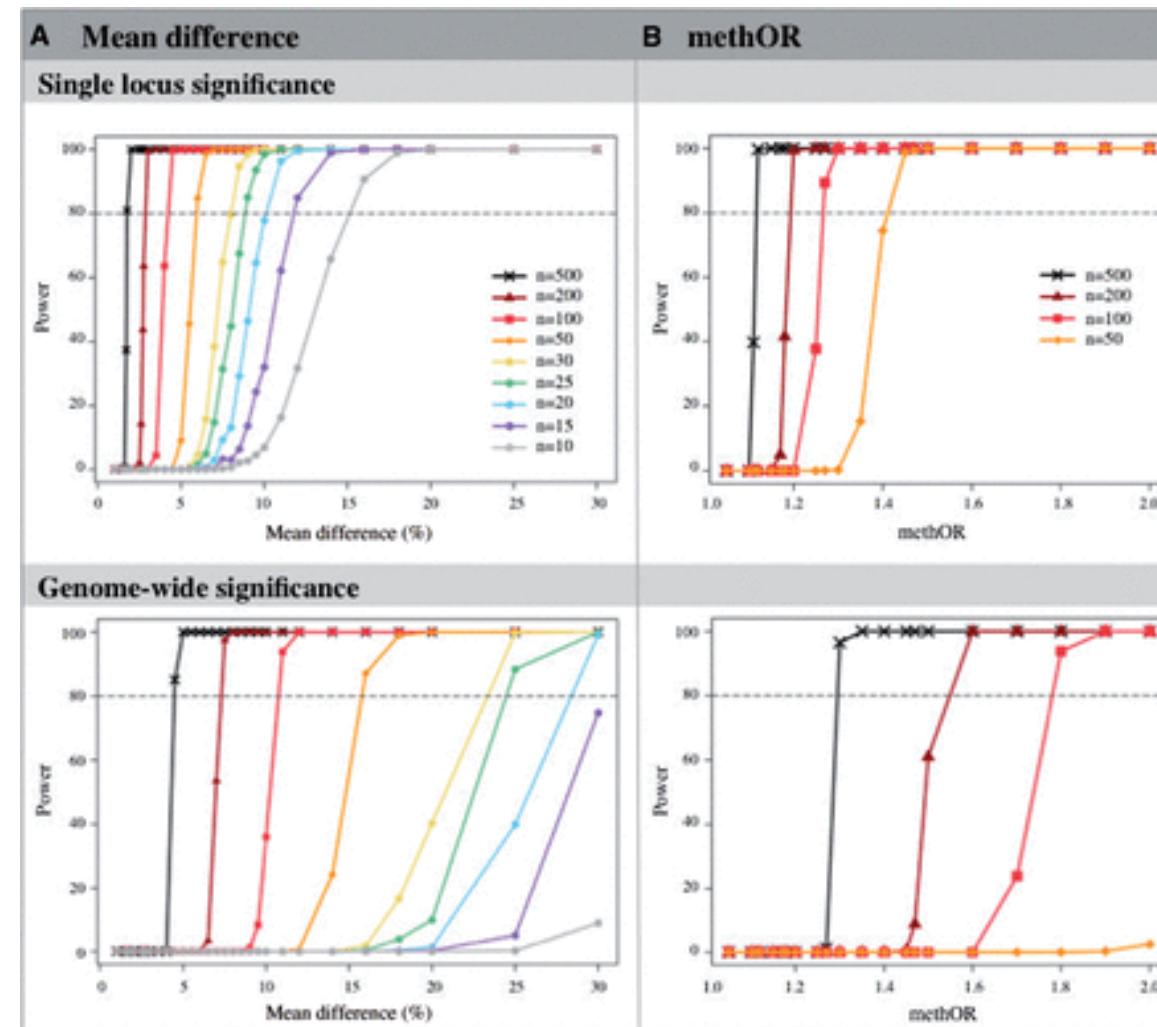
03

Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals 2019

>>> 科研策略—EWAS方案路线之科研设置

	Key advantage	Key disadvantage
Case versus control (singletons) 	Many cohorts exist	Cannot easily control for environmental and genetic confounders
Families 	Can study potential inheritance	Few large cohorts of this type exist
Disease-discordant monozygotic twins 	Can control for genetics	Few large cohorts of this type exist
Prospectively sampled, longitudinal 	Can establish causality	Slow and difficult to establish

>>> 科研策略—EWAS方案路线之科研设置



Methylation odds ratio (methOR)

methOR

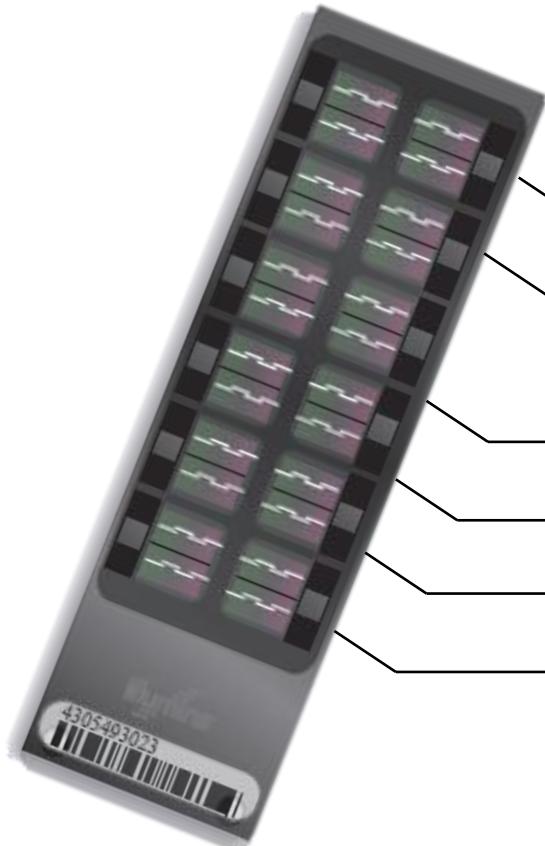
$$= \frac{\text{Mean Methylation}_{\text{Case}} \times (1 - \text{Mean Methylation}_{\text{Control}})}{(1 - \text{Mean Methylation}_{\text{Case}}) \times \text{Mean Methylation}_{\text{Control}}}$$

满足0.8的检验效能
推荐样本量>200

Tsai, et al. International Journal of Epidemiology, 2015

>>> 科研策略—EWAS方案路线之技术检测

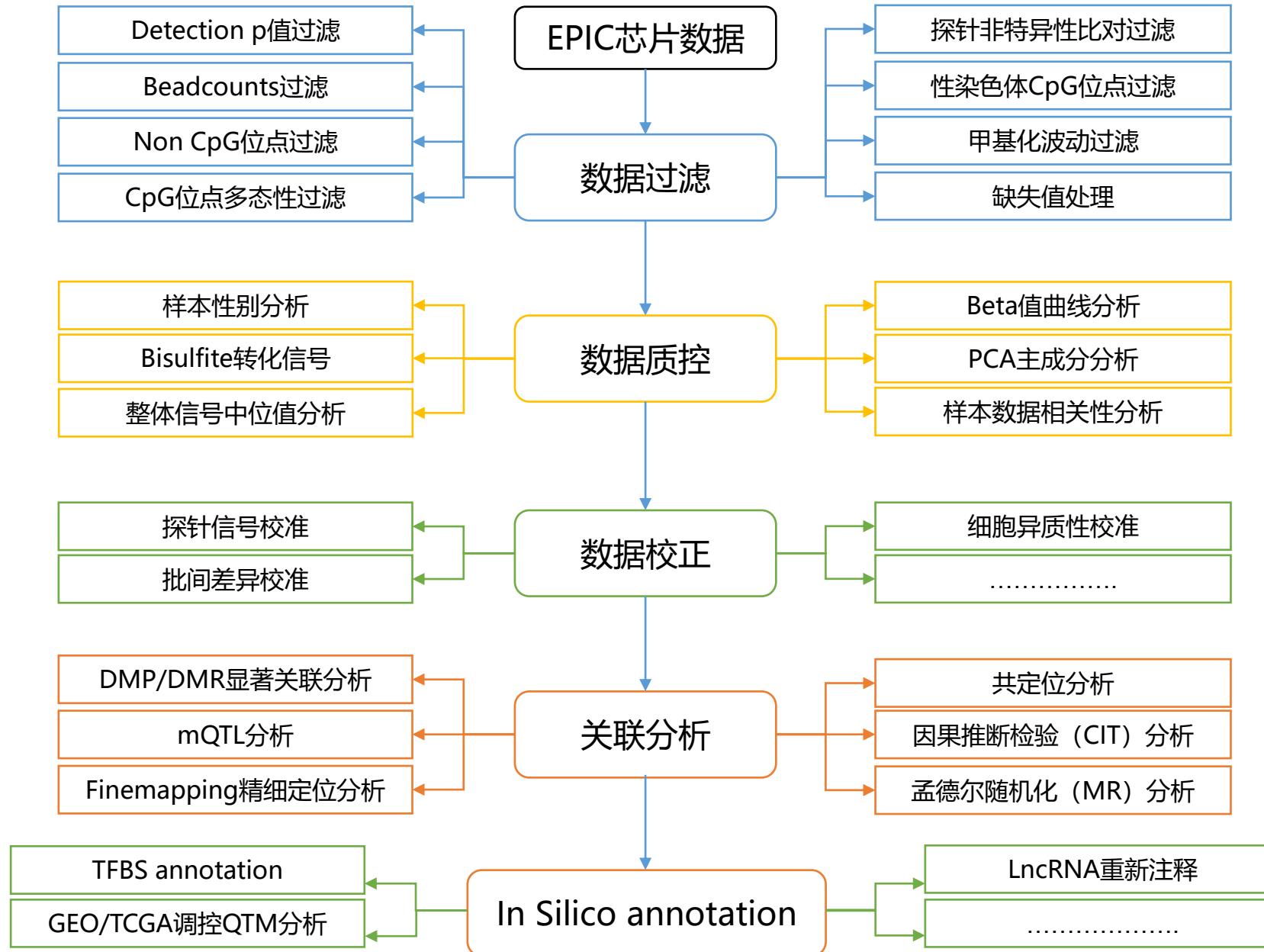
Infinium MethylationEPICv2.0 BeadChip



- Over 935,000 methylation sites
- Enhancers and super-enhancers identified by ChIP-Seq in cancer and cell line samples
- Expanded coverage of CpG islands
- miRNA promoter regions
- DNase hypersensitivity sites
- Compatible with FFPE Sample

... ...

>>> 科研策略—EWAS方案路线之EWAS分析流程



>>> 科研策略—EWAS方案路线之EWAS分析流程

根据detection P值进行过滤

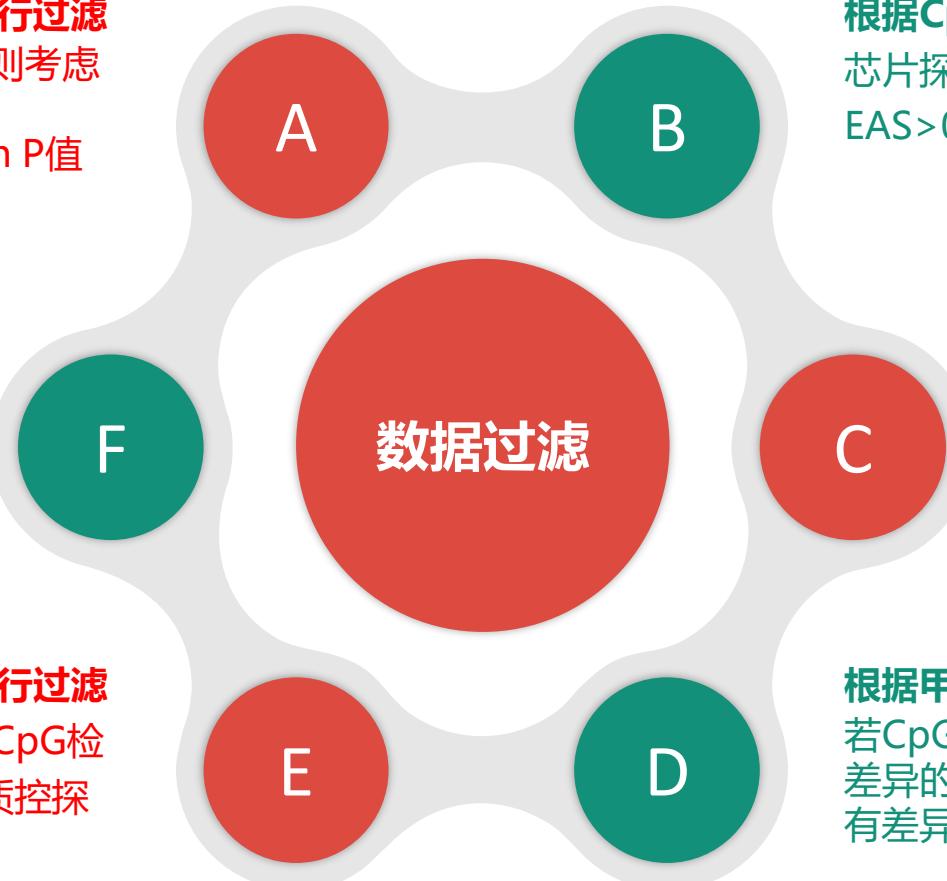
- i) 10%的CpG位点detection P值大于0.01，则考虑滤除该样本；
- ii) 若过滤样本后，某CpG位点仍有detection P值大于0.01，则滤除该位点；

根据beadcounts进行过滤

在超过5%的样本中，某CpG位点NBeads数小于3，则滤除该位点

根据Non CpG位点进行过滤

在850k芯片中，包含大量质控探针等非CpG检测探针，如59个SNP位点、635个各类质控探针等等，在EWAS分析时应予以滤除



根据CpG位点多态性进行过滤

芯片探针靶向区域中包含SNP位点，删除MAF IN EAS>0.05的位点，保障甲基化数据准确性

过滤性染色体CpG位点

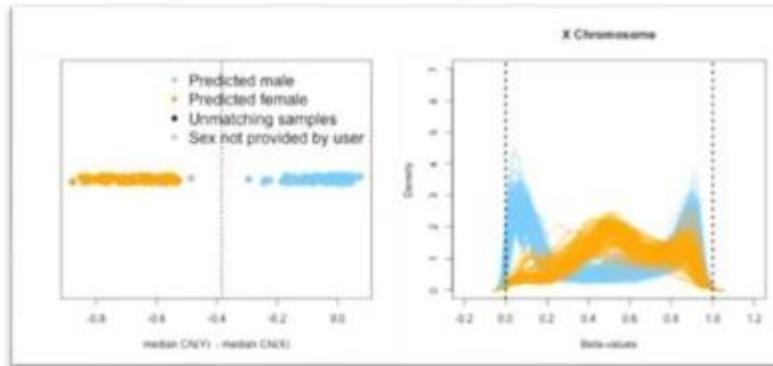
在进行EWAS分析时，如果所关注的样本性状与性别无关，则需要滤除性染色体上的CpG位点。因为这些位点的甲基化分布与性别有明显关联，会影响后续EWAS分析

根据甲基化水平波动程度进行过滤

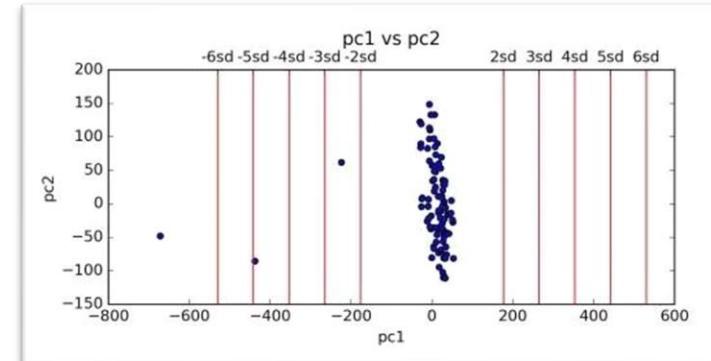
若CpG一般认为标准差小于0.01的，即可视为没有差异的CpG位点。意味着该位点在各样本间几乎没有差异，在预处理时便可予以滤除。

>>> 科研策略—EWAS方案路线之EWAS分析流程

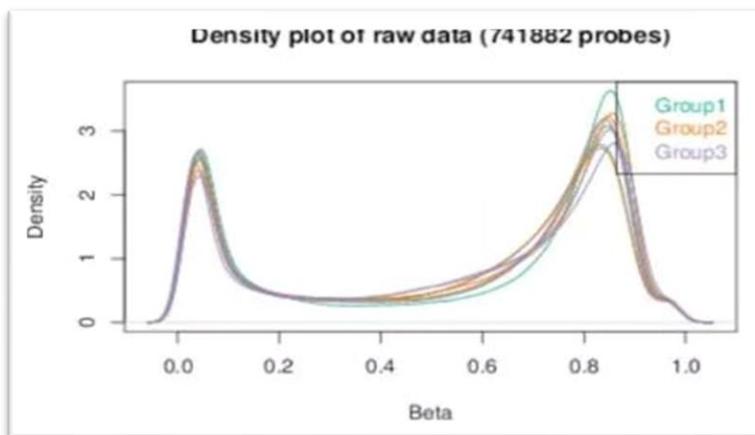
样本性别分析



PCA主成分分析



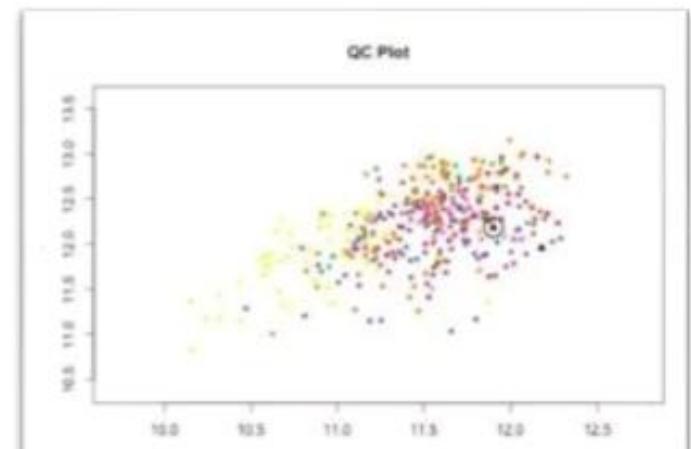
Beta值曲线分析



数据质控

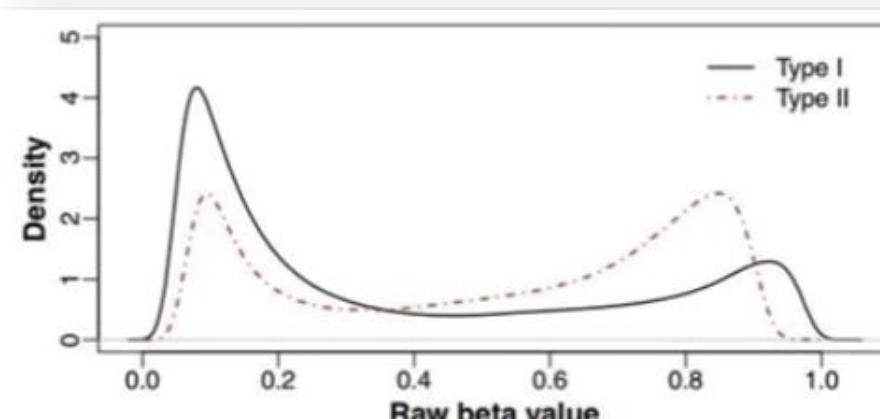


整体信号中位数分析

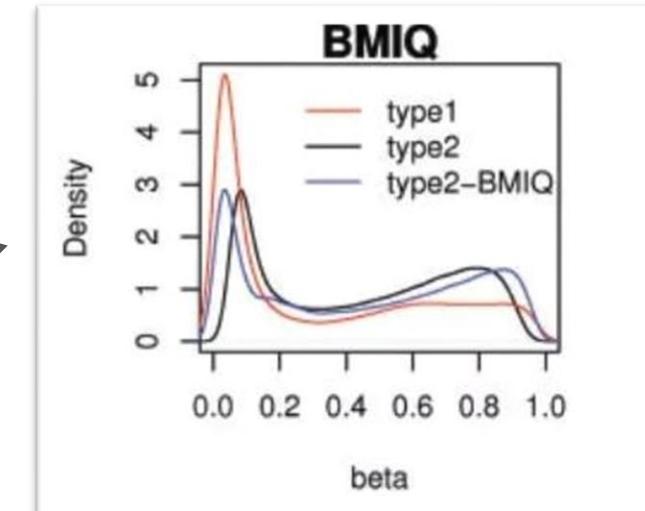


>>> 科研策略—EWAS方案路线之EWAS分析流程

探针信号校准



批间差校准



数据校正



细胞异质性校准

	A	B	C	D	E	F	G
1		sample0	sample1	sample2	sample3	sample4	sample5
2	cg0	0.29433	0.28877	0.29933	0.34376	0.49576	0.18989
3	cg1	0.62762	0.49305	0.064414	0.73766	0.45359	0.74551
4	cg2	0.24081	0.10491	0.75838	0.62884	0.91626	0.13941
5	cg3	0.78783	0.45985	0.045439	0.526	0.04393	0.2247
6	cg4	0.2742	0.74456	0.83392	0.57176	0.60745	0.27522
7	cg5	0.35986	0.52509	0.92509	0.36041	0.89432	0.67379
8	cg6	0.56192	0.29362	0.34886	0.22754	0.80577	0.50971
9	cg7	0.13997	0.99088	0.13815	0.17224	0.4895	0.026827
10	cg8	0.30218	0.90771	0.92043	0.84602	0.55574	0.36847
11	cg9	0.37909	0.11527	0.16601	0.68918	0.90051	0.12646

科研策略—EWAS方案路线之EWAS分析流程

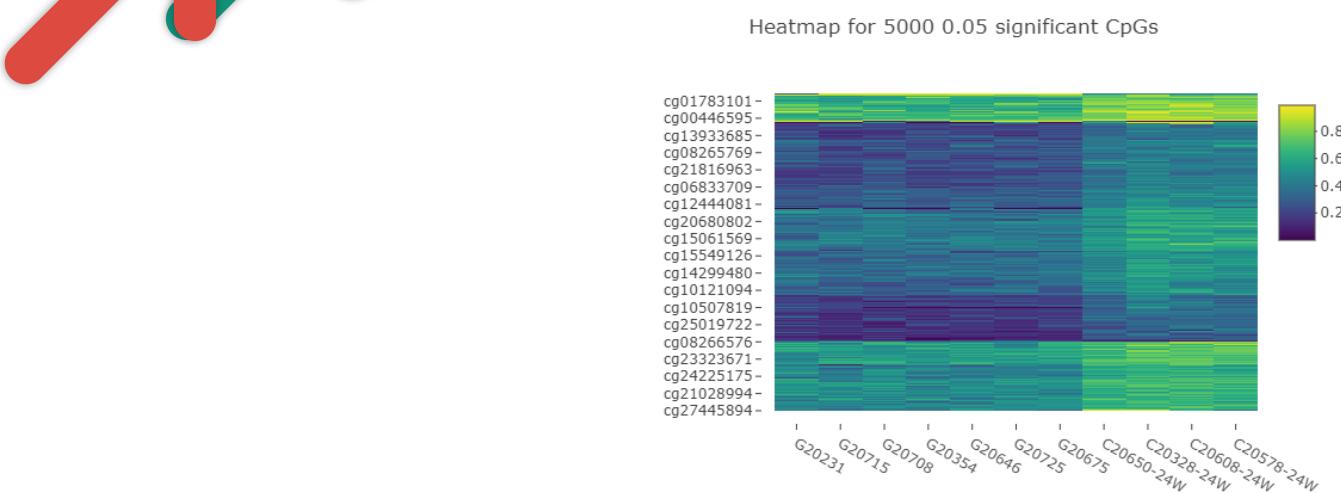
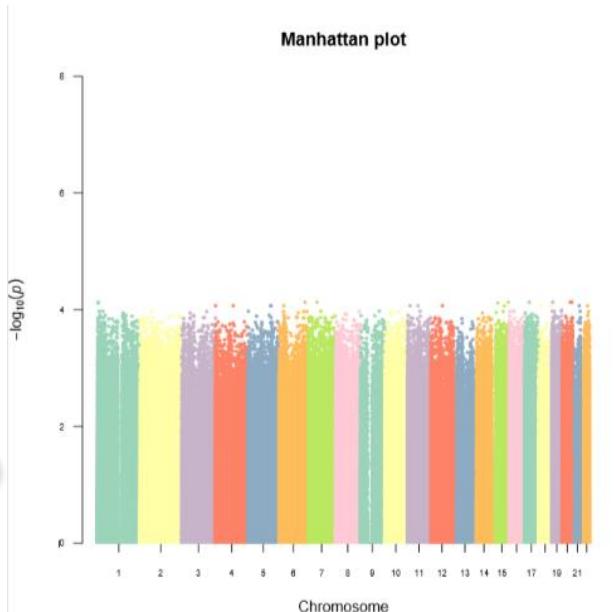
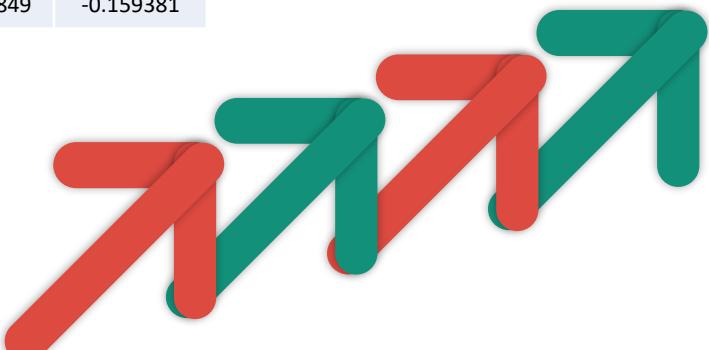


根据LNCipedia数据库重新注释ncRNA信息

>>> 科研策略—EWAS方案路线之EWAS分析流程

关联分析--DMP

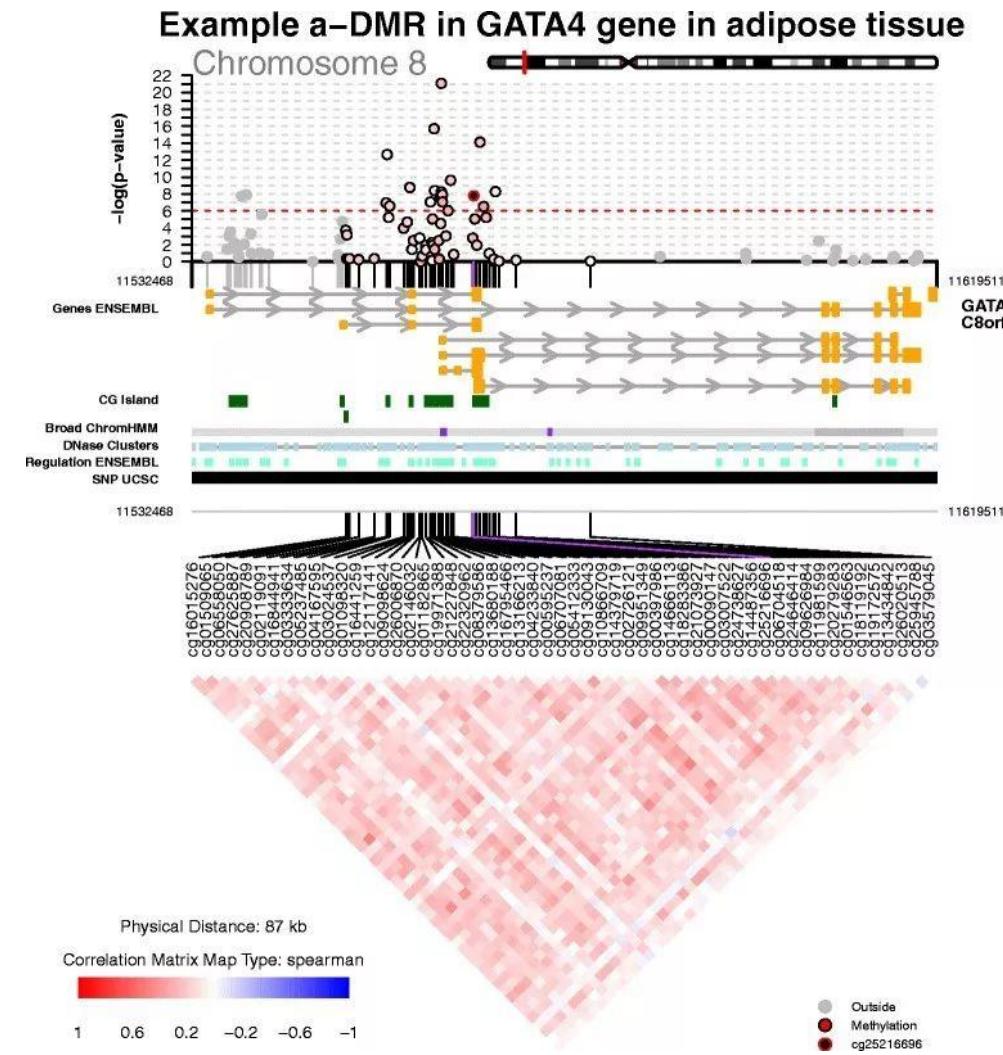
ID	CHR	MAP INFO	gene	P.Value	Adj.P. Value	Gcase_AVG	control_AVG	deltaBeta
cg27236629	15	96877720	NR2F2	3.405543e-10	7.394090e-05	0.14867684	0.29878089	-0.150104
cg23023285	7	50850635	GRB10	3.795204e-10	7.394090e-05	0.33209521	0.60151018	-0.269415
cg08576184	6	151814892	C6orf97	4.159995e-10	7.394090e-05	0.13352196	0.2844085	-0.150886
cg23159236	20	57464002	GNAS	5.653920e-10	7.394090e-05	0.08202661	0.24140849	-0.159381



科研策略—EWAS方案路线之EWAS分析流程

关联分析--DMR

ID	DMRindex	CHR	MAPINFO	Strand	Type	gene	feature	cgi	feat.cgi	adj.P.	ValdeltaBeta
cg1554344DMR_1		6	33386036	R	II	CUTA	1stExon	island	1stExon-i	0.013385	0.028329326
cg0285422DMR_1		6	33385965	F	II	CUTA	1stExon	island	1stExon-i	0.029339	0.032856908
cg0904505DMR_1		6	33385967	F	II	CUTA	1stExon	island	1stExon-i	0.032301	0.028526937
cg0283314DMR_1		6	33385344	R	I	CUTA	Body	shore	Body-shor	0.65866	0.082499274
cg0997786DMR_1		6	33385582	F	II	CUTA	Body	shore	Body-shor	0.660699	0.088980087
cg2307345DMR_1		6	33385236	F	I	CUTA	Body	shore	Body-shor	0.677043	0.079701109
cg0008772DMR_1		6	33385262	F	II	CUTA	Body	shore	Body-shor	0.678473	0.059245709
cg1162885DMR_1		6	33385246	F	I	CUTA	Body	shore	Body-shor	0.679073	0.095595376
cg2356582DMR_1		6	33385056	R	II	CUTA	Body	shore	Body-shor	0.75155	0.058304409
cg1301844DMR_1		6	33385440	F	I	CUTA	Body	shore	Body-shor	0.762975	0.017147783
cg1396294DMR_1		6	33385698	R	I	CUTA	Body	island	Body-isla	0.843354	0.004863854
cg2158369DMR_1		6	33385779	F	II	CUTA	Body	island	Body-isla	0.868997	0.002681564
cg2558965DMR_2		6	28891121	R	I	TRIM27	1stExon	island	1stExon-i	0.003152	0.098998432
cg2721966DMR_2		6	28891728	R	I	TRIM27	5' UTR	island	5' UTR-isl	0.003314	0.075410629
cg2356971DMR_2		6	28891915	R	I	TRIM27	TSS200	island	TSS200-i	0.003429	0.079615105
cg0367144DMR_2		6	28891045	F	I	TRIM27	1stExon	island	1stExon-i	0.004202	0.056052559
cg2074107DMR_2		6	28891917	F	II	TRIM27	TSS200	island	TSS200-i	0.004872	0.042326551
cg077914CDMR_2		6	28891716	R	I	TRIM27	5' UTR	island	5' UTR-isl	0.00714	0.058870813
cg0364732DMR_2		6	28891109	R	II	TRIM27	1stExon	island	1stExon-i	0.00714	0.029170848
cg0087520DMR_2		6	28891298	F	I	TRIM27	1stExon	island	1stExon-i	0.007518	0.07028657
cg0980531DMR_2		6	28891289	F	I	TRIM27	1stExon	island	1stExon-i	0.010096	0.069654145
cg0327034DMR_2		6	28891204	R	I	TRIM27	1stExon	island	1stExon-i	0.014551	0.053188134
cg0639569DMR_2		6	28891945	R	II	TRIM27	TSS200	island	TSS200-i	0.016278	0.03076428
cg1162944DMR_2		6	28891340	R	I	TRIM27	1stExon	island	1stExon-i	0.019955	0.044307781



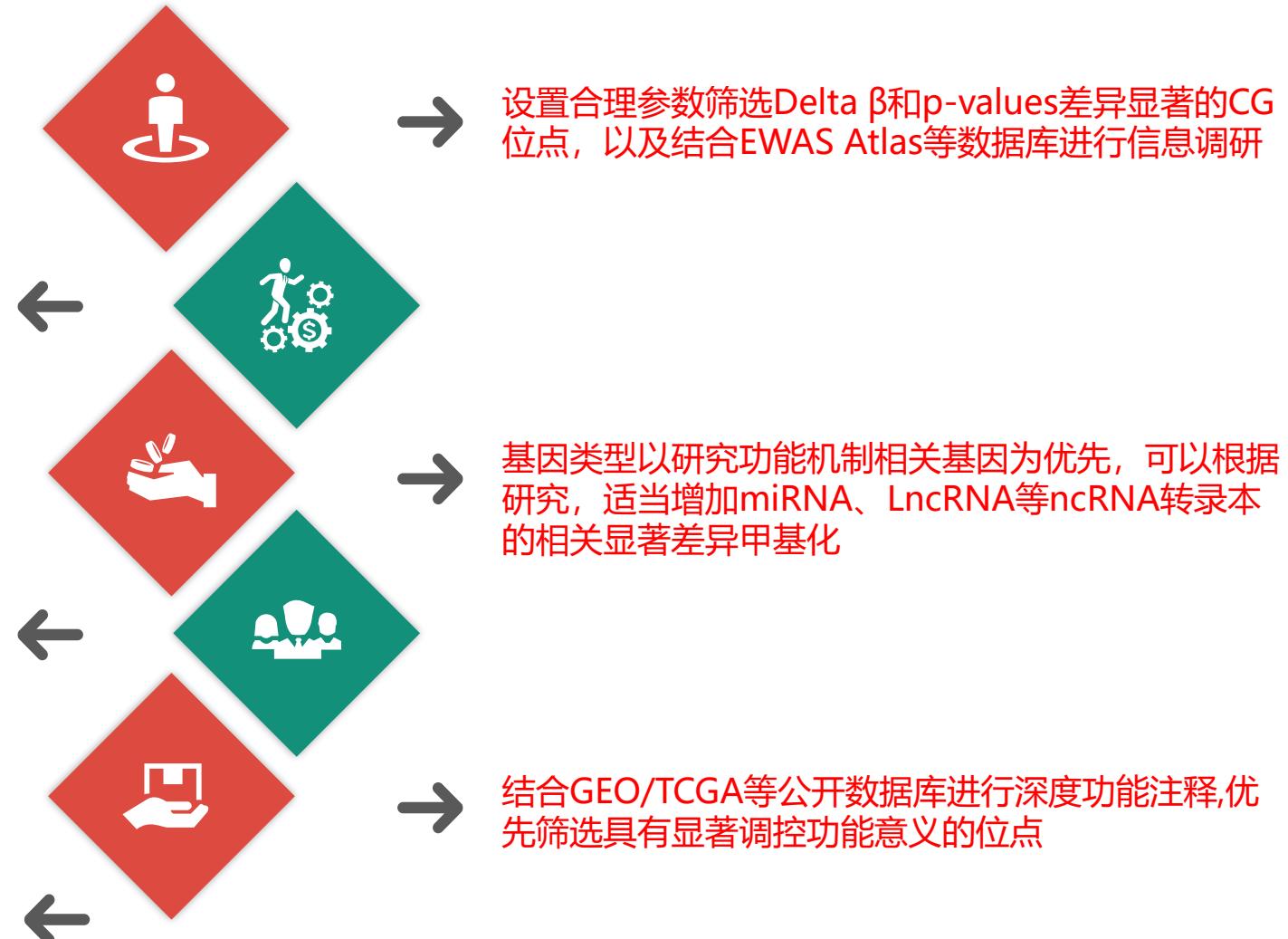
>>> 科研策略—EWAS方案路线之EWAS靶向验证位点优化策略

博淼专属服务

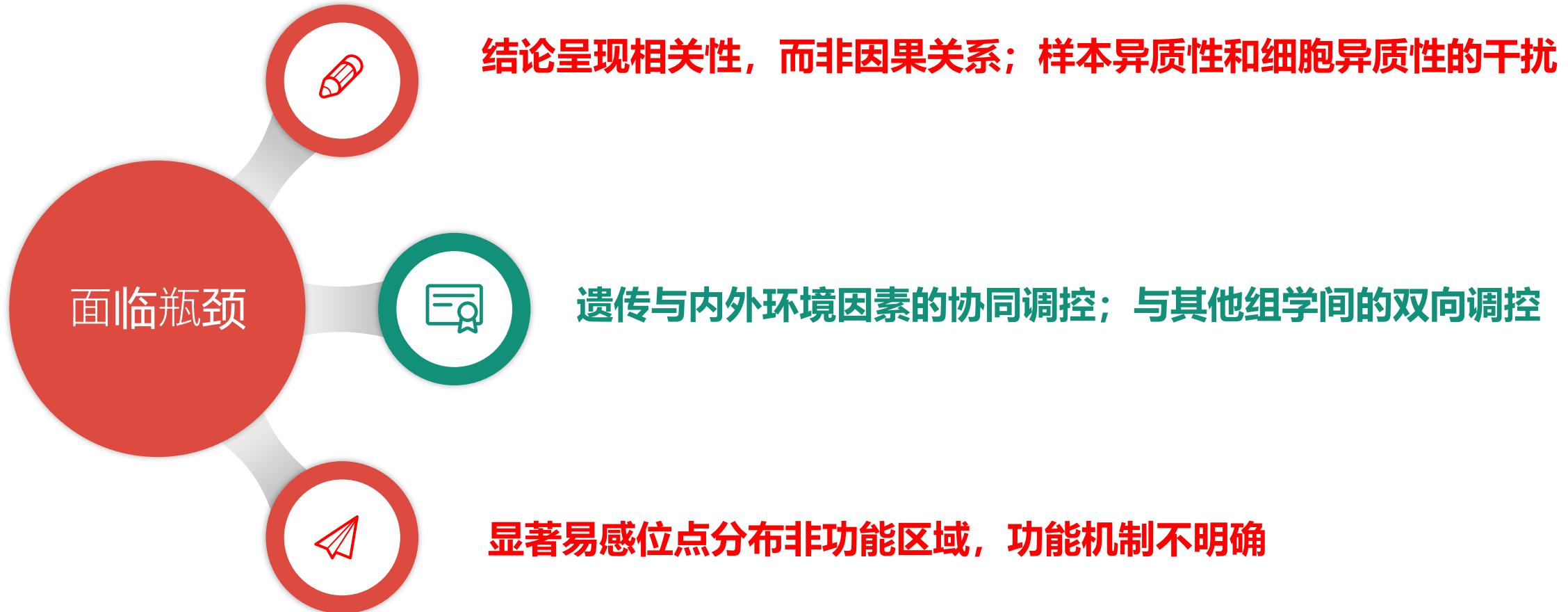
针对芯片中部分高密度覆盖区域进行DMR分析，优先选择差异显著的调控区域DMR，如island、shore、shelf区域

基因位置选择，优先选择CpG位点位于基因Promoter区（包括TSS1500, TSS200），以及5' UTR区的芯片位点，且以其中CpG island为优先

根据初步选择的候选位点，计算ROC及AUC，评估预测的灵敏度及特异性



>>> 科研策略—传统EWAS研究面临的瓶颈



>>> 科研策略—靶向候选基因甲基化方案路线

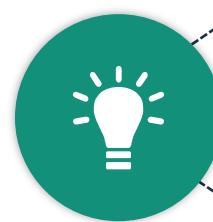
候选基因筛选

- ✓ 数据来源：自主数据RNA-seq、WES、GWAS等；公开数据库TCGA、GEO、GTEx等；文献报道
- ✓ 基因类型：mRNA、miRNA、LncRNA、cirRNA等



甲基化区域筛选

- ✓ 评估Promoter/1stExon区域
- ✓ 预测CpG island/shore/shelf区域
- ✓ 评估TFBS、SNP信息



靶向CpG site定量检测

- ✓ 中通量Massarray技术
- ✓ 高通量Multi-PCR NGS技术



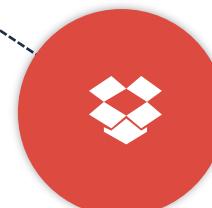
生物信息分析

- ✓ 分组性状及数量性状关联分析
- ✓ 因果效应分析
- ✓ QTM调控分析等



样本设置

- ✓ 外周血、组织、细胞、唾液等等
- ✓ 样本数量

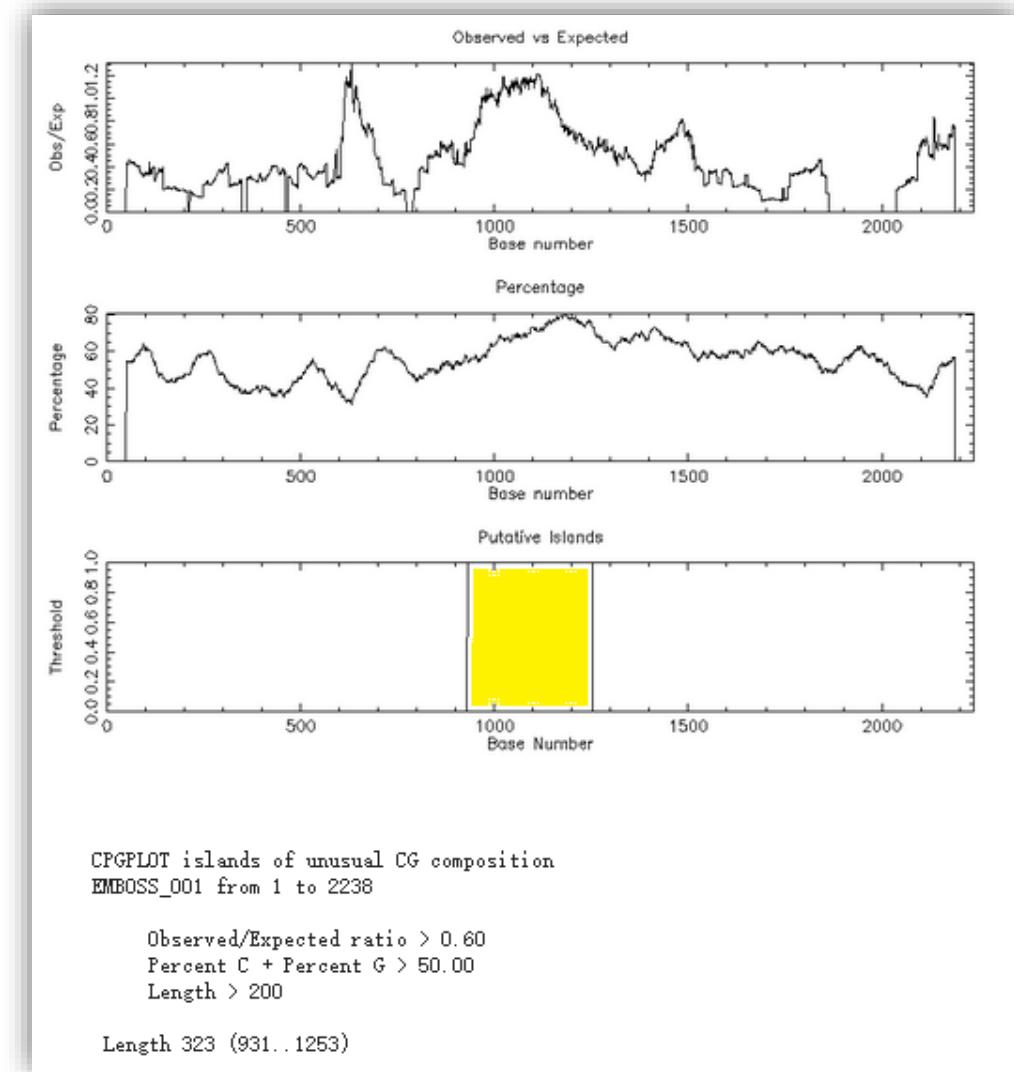


>>> 科研策略—靶向候选基因甲基化方案路线

CpG island筛选

通过http://www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/
寻找符合特定参数要求的CpG island基因序列区域

WINDOW SIZE	MINIMUM LENGTH	MINIMUM OBSERVED	MINIMUM PERCENTAGE
100	200	0.6	50



>>> 科研策略—靶向候选基因甲基化方案路线

TFBS结合区域预测

通过jaspar.genereg.net/search针对候选区域进行转录因子预测

entry	score
M00048 ADR1	95.4
M00175 AP-4	93.8
M00176 AP-4	91.9
M00175 AP-4	91.7
M00032 c-Ets-	90.2
M00176 AP-4	89.2
M00263 StuAp	87.7
M00253 cap	86.7

>>> 科研策略—靶向候选基因甲基化方案路线

待检测区域调控元件

-2400 TACTTTATAATTTCATGCCTGTCTTGCTTAATCTCTTAATCCTGTTATCTTCAAGCTAAGGATGTACGTCACCTCAGGACCACTGTGATAATT
CdxA CdxA NIT2

-2300 GTGTTAACTGTACAGATTGATTGCAAAACATGTGTGTTGAACAATATGAAATCAGTGCACCTGAAAAAGAGCAGAATAACAGCAATTAGGGAAACA
CDP CR Sox-5 HSF HSF

-2200 AGGGAAGACAACATAAGGTCTGACTGCCTGCGGGTCCGGCAAAGGGAGCCATATTTCTTGCAGAGAGCCTATAAATAGACCTGCAAGTAGGAG
ADR1 TATA

-2100 AGATATTGCTAATTCTTGCTAGCATGGAATATTAATATTAACACCCTGGAAAGGAATGCATTCTGGGGGAGGTCTATAAATGGCCGCTCTGGGA
CdxA ADR1 TATA

-2000 ATGTCTATCCTACGCAATGGAGATAAGGACTGAGATA CGCCCTGGTCTCCTGCAGTACCCCTCAGGCTTACTAGGGTGGTAAAAACTCCGCCCTGGTAAA
GATA-X

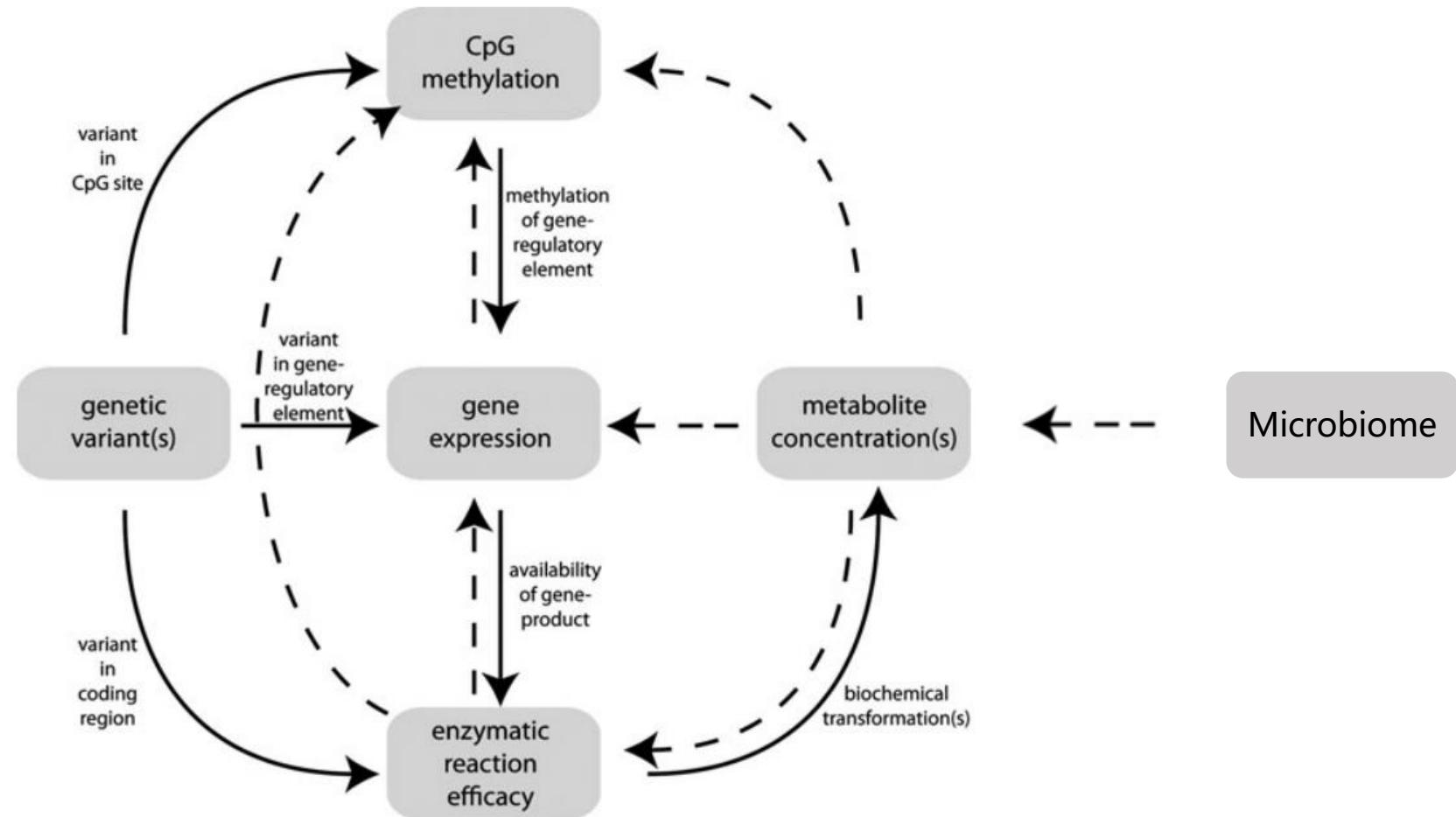
C(rs4728096) T(rs12536535)

-1900 TTTGTGGTCAGACCAGTTCTGCTCTGAACACTGTTCTGTTAAGATGTTATCAAGACAATA CGTGCACCGCTGAACACAGACCCCTATCAG
AML-1a NIT2 cap

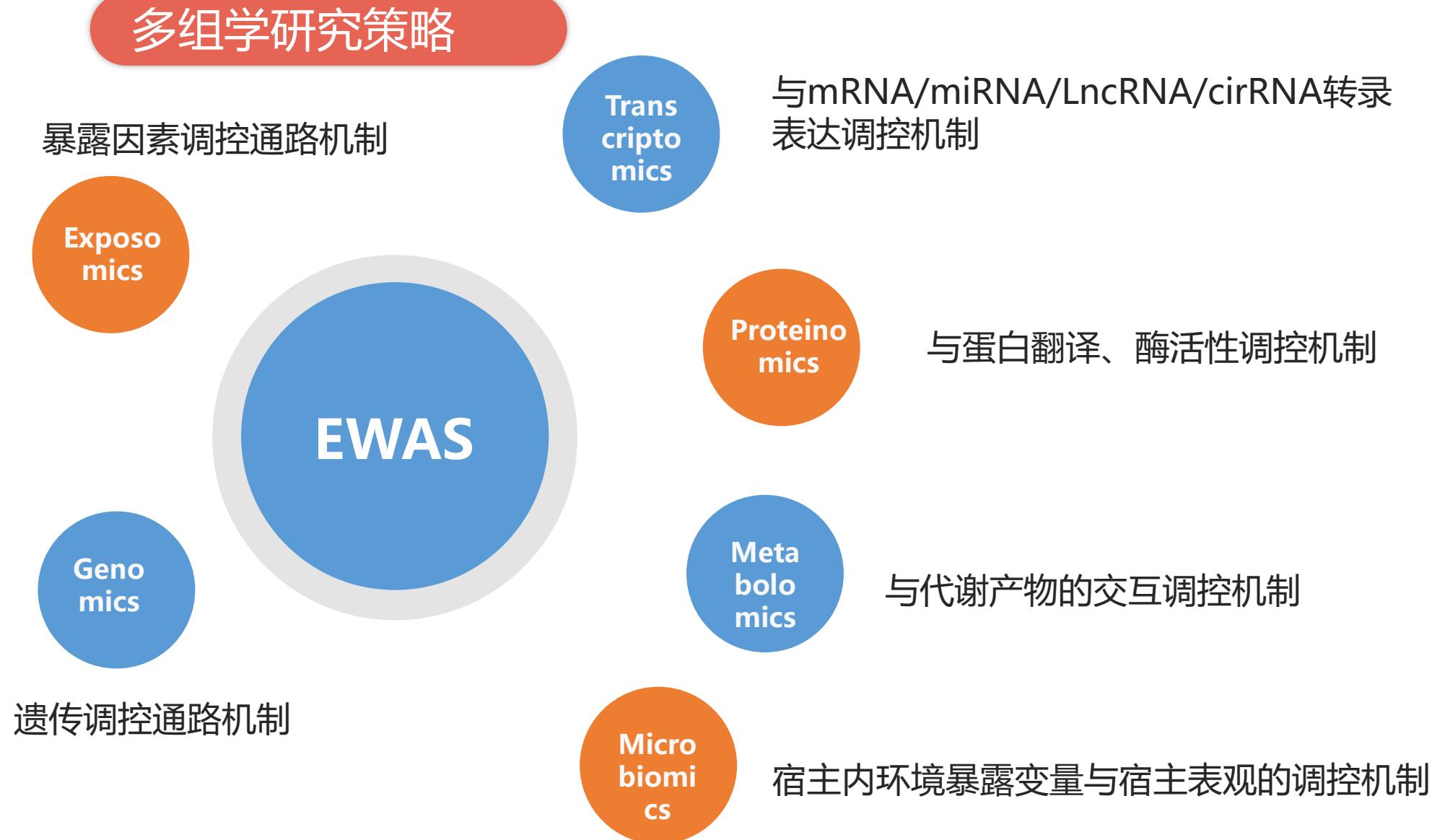
-1800 TAGTTCTCCTTTGCCCTTGAAGCATGTGATCTACTCCCTGTTACACCCCTCACCTTGAAACCCCTAATAAAAAACTTGCTGGTTGAGGCTC
Hb

>>> 科研策略—表观基因组联合多组学方案路线

多组学调控路线图



>>> 科研策略—表观基因组联合多组学方案路线





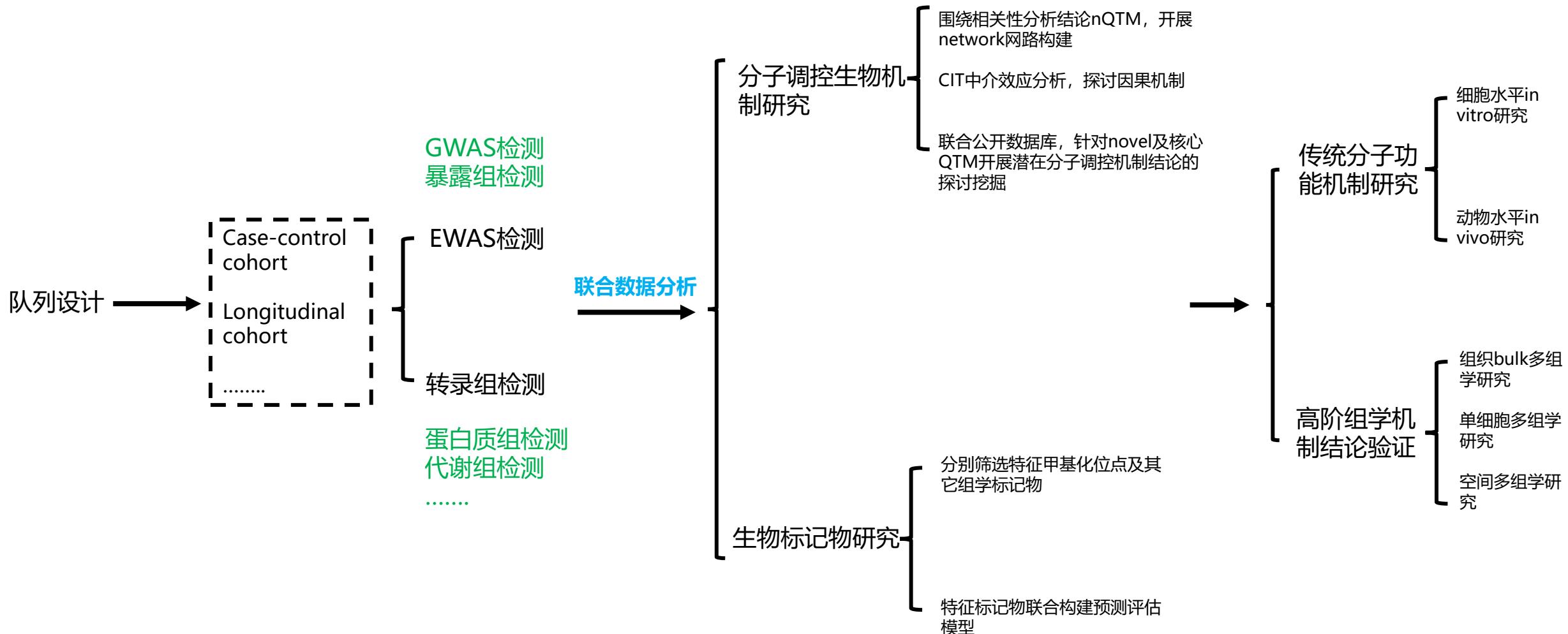
The Epigenomics first approach—功能调控中介组学

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

	Total SNPs	cis-meQTL			trans-meQTL		
		N of the overlap	Fold Change	P-value	N of the overlap	Fold Change	P-value
<i>cis</i> -eSNPs FHS (1e-7)	460,536	399,670	1.66	0	84,591	2.33	0
<i>cis</i> -eSNPs Eur-Meta	398,524	356,500	1.71	0	70,810	2.25	0
<i>cis</i> -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
<i>cis</i> -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
<i>metabolism</i> 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

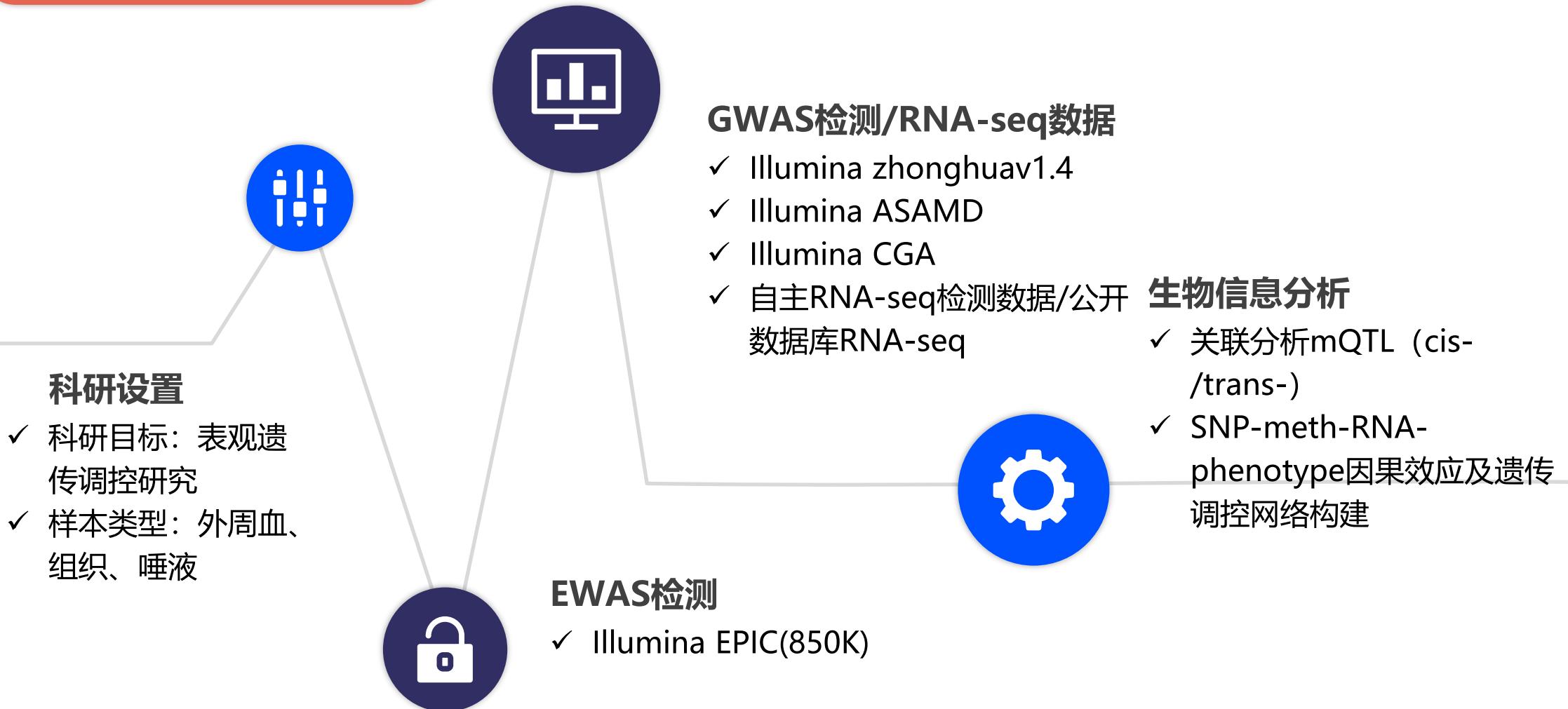


The Epigenomics first approach 研究策略—研究路径概述



>>> 科研策略—表观基因组联合多组学方案路线

EWAS&GWAS联合方案



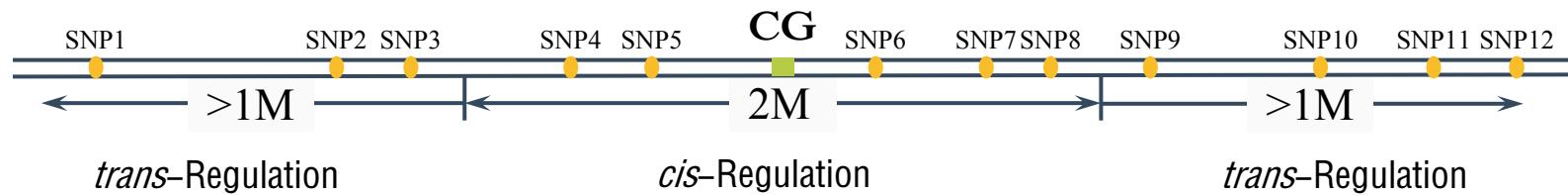
>>> 科研策略—表观基因组联合多组学方案路线

SNP-meth遗传调控



Allele-specific methylation (ASM) { Associated with methylation of a nearby CpG site
SNP itself destroyed a CpG site by changing the C or G

Haplotype-specific methylation (HSM) { Associated with methylation of a nearby CpG site

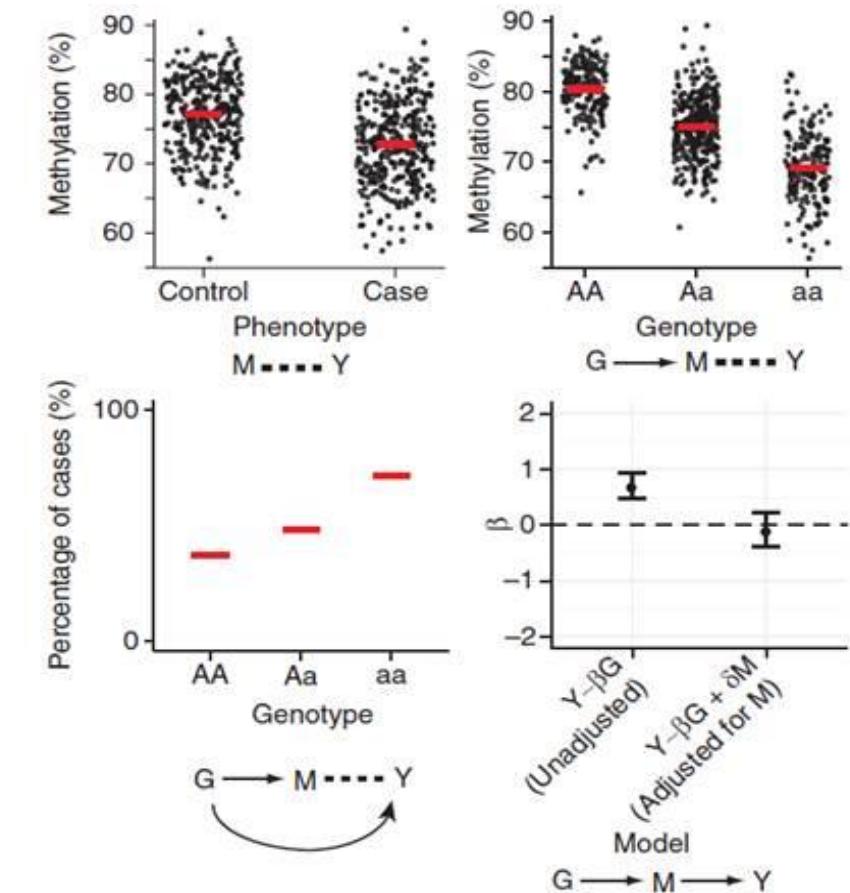
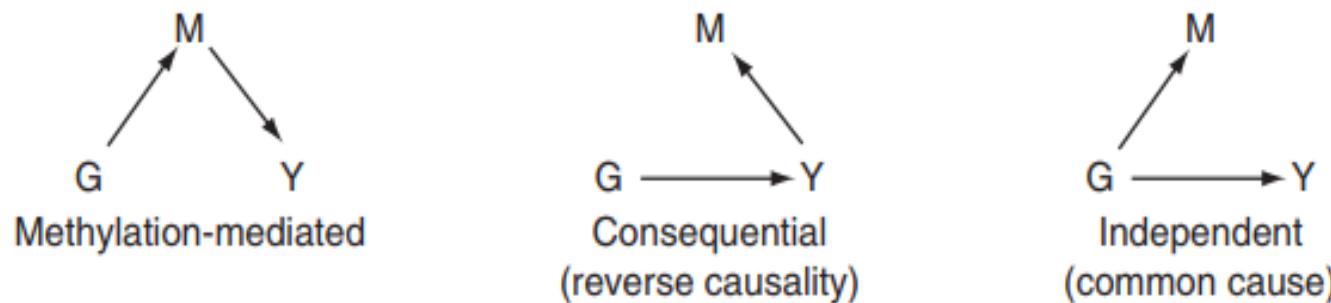


cis-Regulation: cisacting regulation by DNA elements in or adjacent to each CpG site

trans-Regulation: trans-acting regulation by factors from the genomic regions distant from the CpG sites, including from different chromosomes.

>>> 科研策略—表观基因组联合多组学方案路线

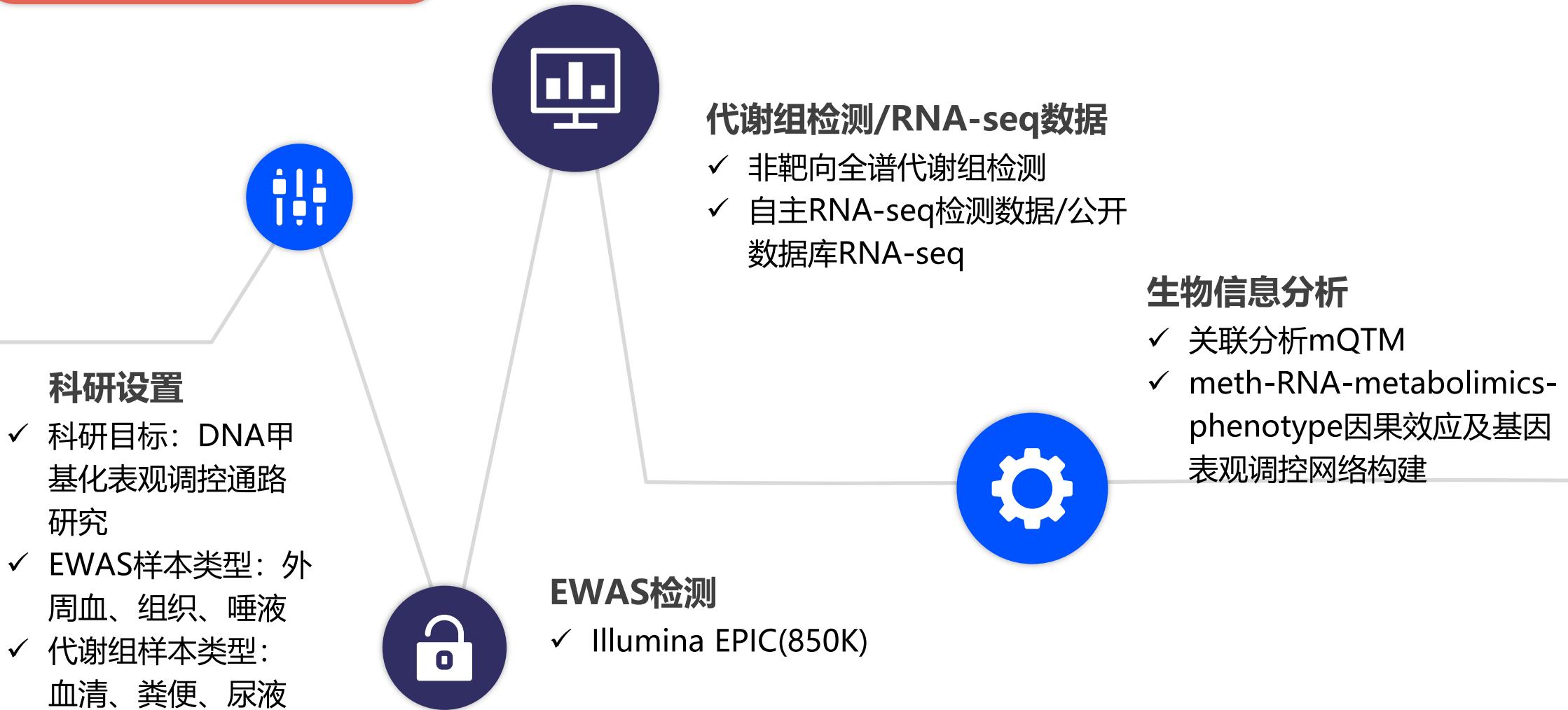
SNP-meth-Phenotype因果效应



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

>>> 科研策略—表观基因组联合多组学方案路线

EWAS&代谢组联合方案

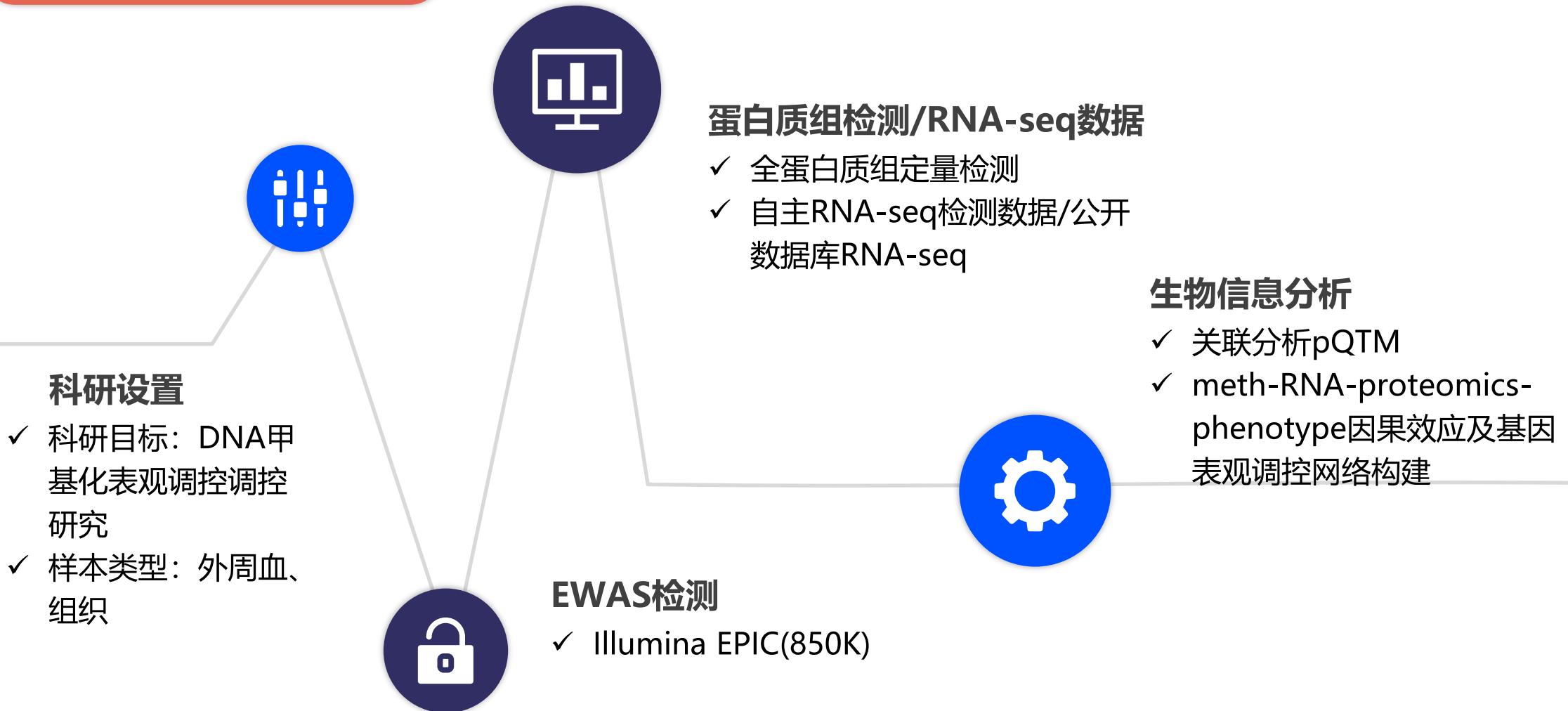


EWAS&代谢组关联分析

Locus name	CpG	Chr	Pos	Metabolic trait	Beta'	r^2	P-value
ACADS	cg24768164	12	121 163 261	Butyrylcarnitine ^a	-0.998	0.221	2.0×10^{-108}
PYROXD2	cg26690318	10	<u>100</u> <u>167</u> <u>465</u>	X-12092 ^b	2.171	0.138	2.2×10^{-60}
NAT8	cg13584399	2	<u>73</u> <u>907</u> <u>327</u>	N-acetylpornithine ^a	-0.950	0.120	8.9×10^{-52}
ACADM	cg10523679	1	76 189 770	Hexanoylcarnitine ^a	-0.456	0.065	1.8×10^{-30}

>>> 科研策略—表观基因组联合多组学方案路线

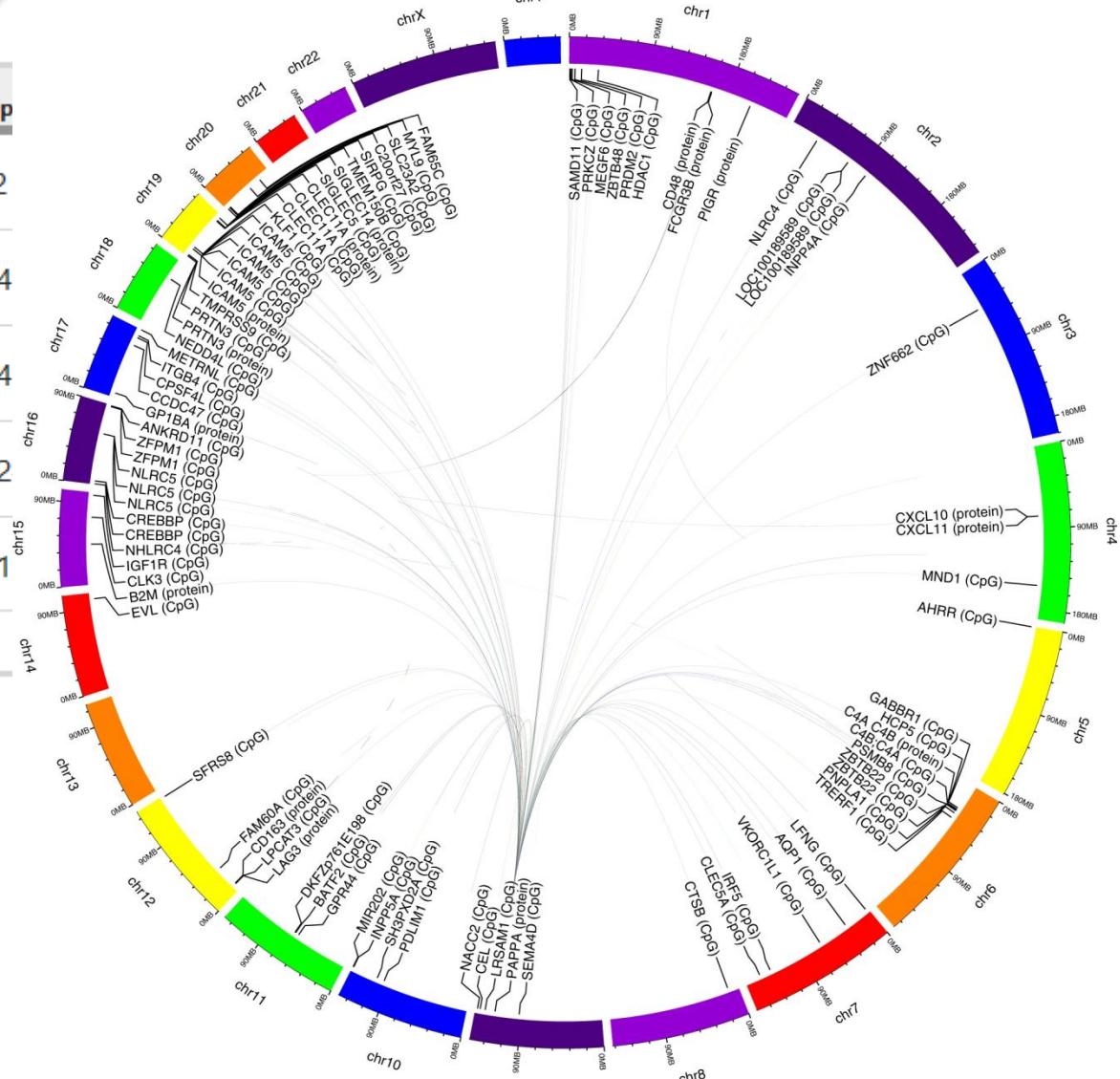
EWAS&蛋白质组联合方案



科研策略—表观基因组联合多组学方案路线

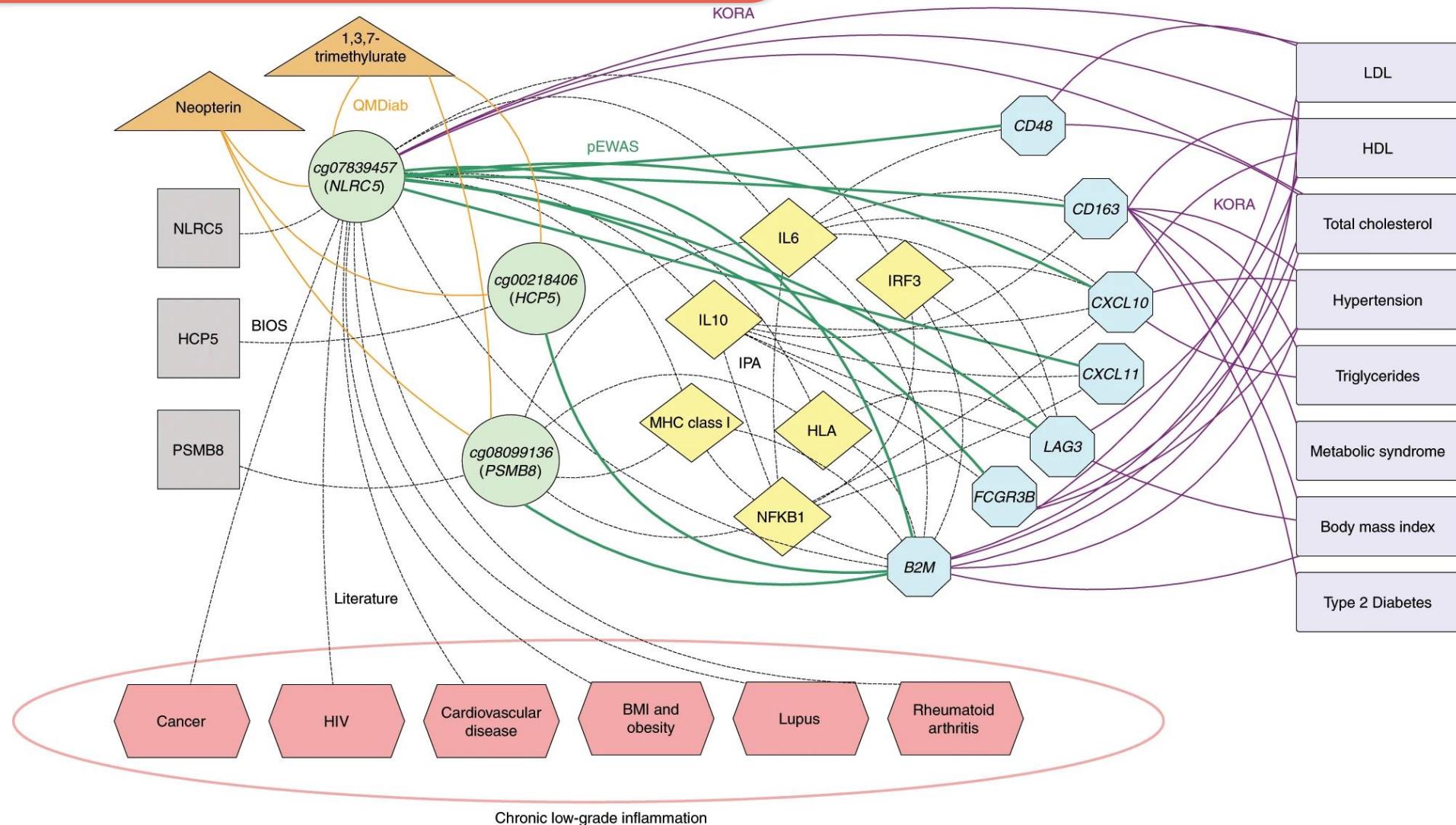
EWAS&蛋白质组关联分析-cis/trans-pQTM

CpG	pQTM (this study)	P _{pQTM}	Beta _p
cg07708453 (<i>PRDM2</i>) chr1:14,032,034	PAPPA (4148-49_2) chr9:118,916,083-119,164,601	3.10×10^{-16}	0.262
cg19393755 (<i>CPSF4L</i>) chr17:71,258,101	PAPPA	2.03×10^{-14}	-0.24
cg10831642 (<i>SH3PXD2A</i>) chr10:105,378,344	PAPPA	8.19×10^{-12}	-0.24
cg26272069 (<i>GABBR1</i>) chr6:29,591,706	PAPPA	9.25×10^{-12}	-0.22
cg20290167 (<i>METRNL</i>) chr17:81,040,724	PAPPA	5.58×10^{-11}	-0.21
(Total: 72 PAPPA pQTMs)		-	-



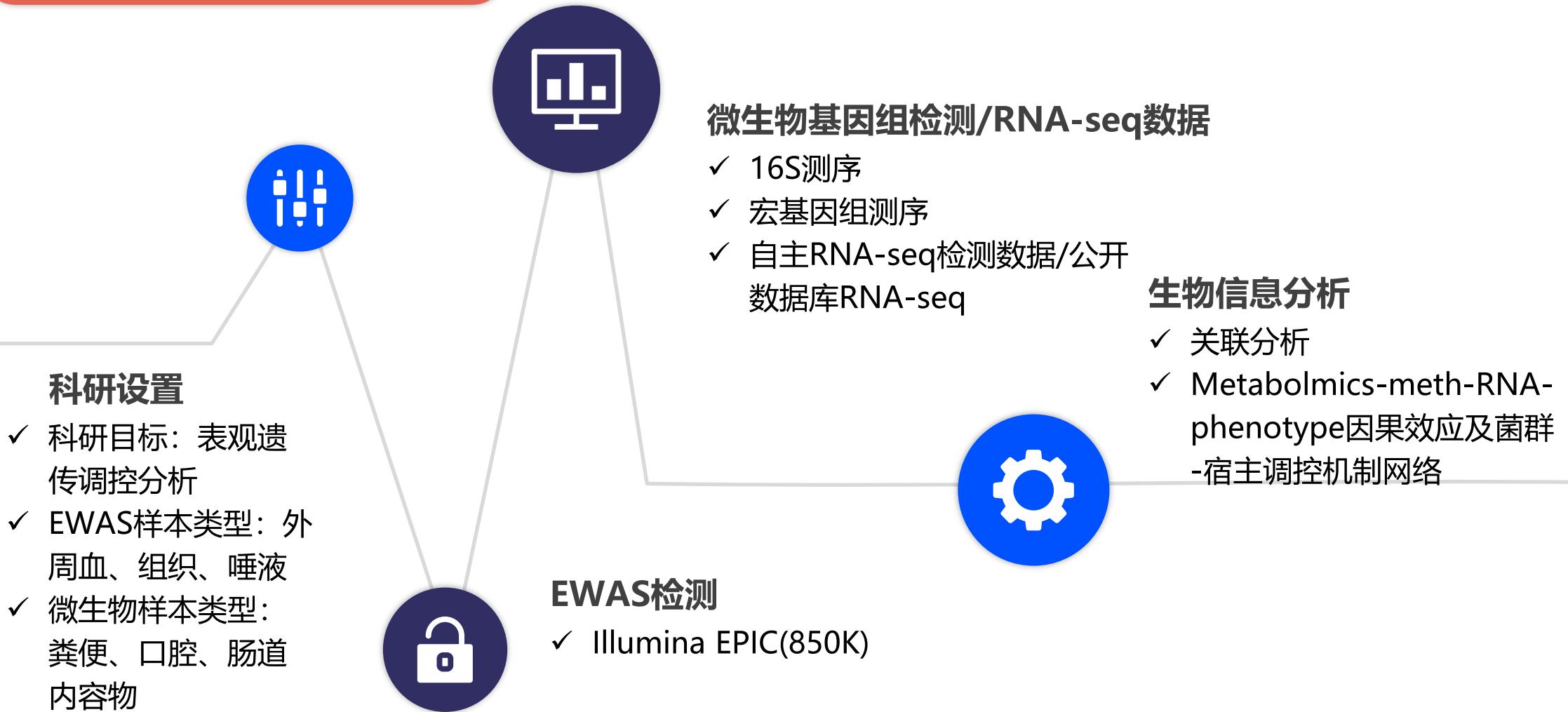
>>> 科研策略—表观基因组联合多组学方案路线

DNA甲基化-蛋白质-表型调控通路构建



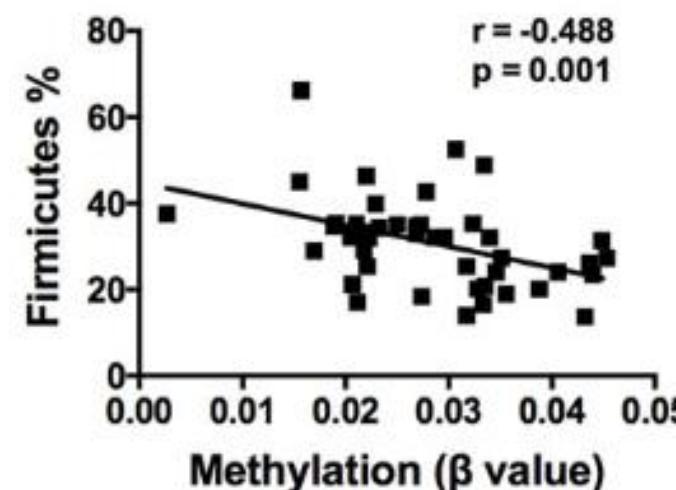
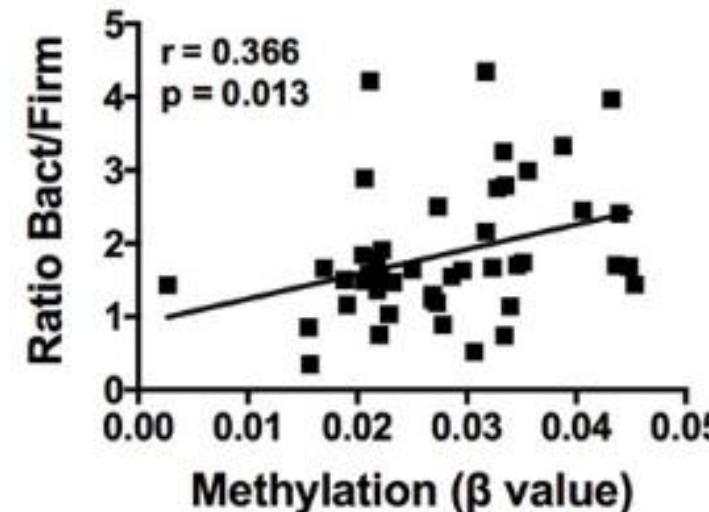
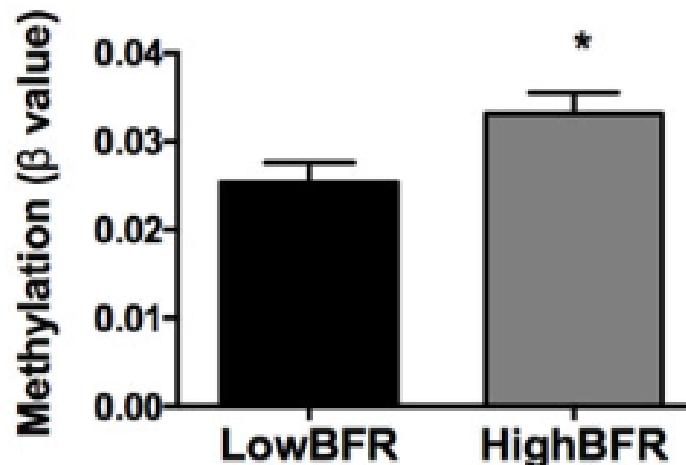
>>> 科研策略—表观基因组联合多组学方案路线

EWAS&微生物基因组联合方案



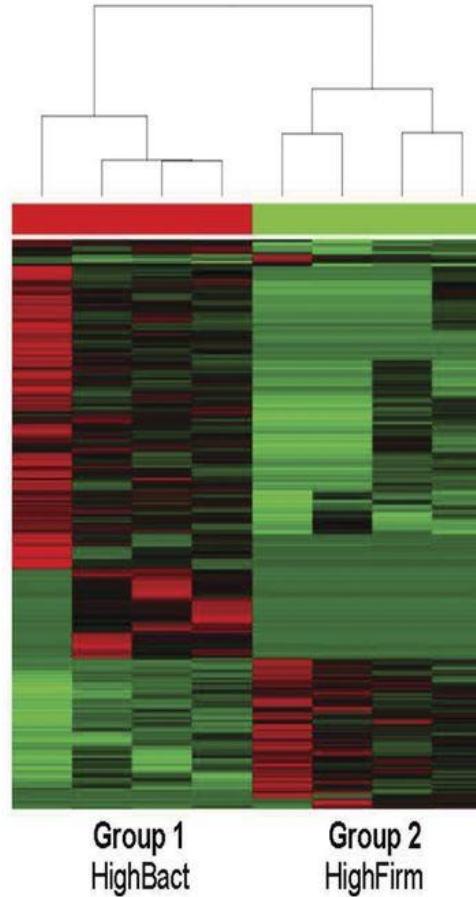
>>> 科研策略—表观基因组联合多组学方案路线

EWAS&微生物关联分析

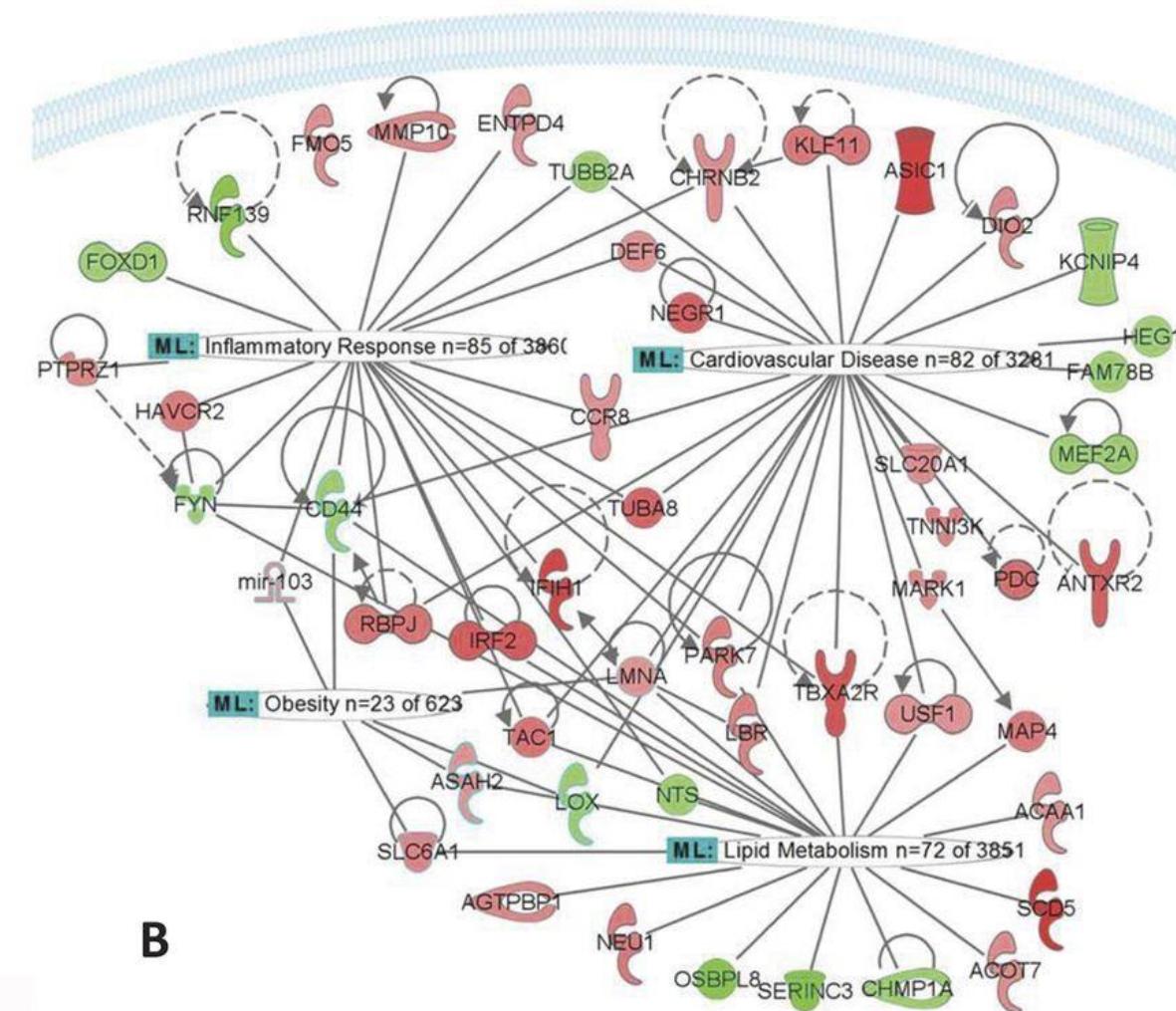


>>> 科研策略—表观基因组联合多组学方案路线

微生物-甲基化-表型调控网络构建



A



B

03

甲基化注释数据库介绍

>>> 数据库介绍

EWAS Atlas: EWAS研究数据注释

EWAS Atlas Browse EWAS Toolkit Downloads Statistics API Help EWAS Data Hub

Filter Hide Traits (426) Probes (248079) Genes (34188) Studies (1062) Publications (631)

Trait(s)
Gene symbol
Promoter/Body
Probe ID
P value rank from to Go!

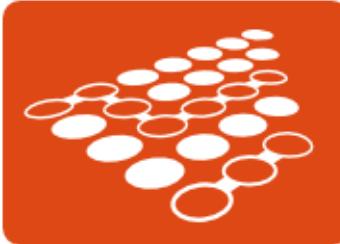
Correlation: Hypermethylation Hypomethylation Not report
Trait type: Cancer Non-cancer disease Phenotype Environmental factor Behavior
+ Show detail information or genome browser
Click on a hyperlink will use it as a search condition

Probe ID	Studies	Correlations	Location	Related genes (transcript: location)	CpG islands	Related traits
cg05575921	72 +		chr5: 373378 +	AHRR (ENST00000505113: body) AHRR (ENST00000316418: body) AHRR (ENST00000512529: body) AHRR (ENST00000514523: body) AHRR (ENST00000510400: body)	Shelf	lung carcinoma lung cancer cardiovascular risk HIV f... atopy mortality cognitive funct... metabolic trait blood protein biomarker lev... IgG glycosylation lung f...

>>> 数据库介绍

Wanderer: DNA甲基化&RNA转录调控注释

maplab tools ▾ Wanderer Documentation Support



Wanderer
An interactive viewer to explore DNA methylation
and gene expression data in human cancer

Gene Symbol or Ensembl Gene Id
BRCA1
Examples: BRCA1 or ENSG00000141510
Important: Press refresh after entering a new gene name
Refresh

41,095,000
Select a probe:
cg12984107
cg16029534
cg16919093
cg18830083
cg07054526
cg13782816
cg04582861
cg12984107
cg19531713
cg00000000

Zoom
BRCA1 Normal
Breast invasive carcinoma
n=84 Pearson r=-0.124

$\log_2(\text{normalized rsem} + 1) \text{ of BRCA1}$

methylation beta value of cg12984107

Zoom
BRCA1 Tumor
Breast invasive carcinoma
n=720 Pearson r=-0.421

$\log_2(\text{normalized rsem} + 1) \text{ of BRCA1}$

methylation beta value of cg12984107

This figure displays the Wanderer interface and its functionality. It includes a navigation bar at the top with links to maplab tools, Wanderer, Documentation, and Support. Below the navigation bar is a main content area featuring the Wanderer logo and a brief description: "An interactive viewer to explore DNA methylation and gene expression data in human cancer". On the left, there is a search input field for "Gene Symbol or Ensembl Gene Id" with "BRCA1" entered. Below the input field are examples and a note to refresh after entering a new gene name. A "Refresh" button is located at the bottom left. In the center, there is a dropdown menu titled "Select a probe:" showing a list of CpG probes. The probe "cg12984107" is selected and highlighted with a blue background. To the right, two scatter plots are shown under the heading "Zoom". The first plot shows the relationship between the methylation beta value of probe cg12984107 (x-axis, 0.0 to 1.0) and the $\log_2(\text{normalized rsem} + 1)$ of BRCA1 (y-axis, 4 to 10) for normal breast tissue samples (n=84). The Pearson correlation coefficient is -0.124. The second plot shows the same relationship for tumor samples (n=720), with a Pearson correlation coefficient of -0.421.

>>> 数据库介绍



mqtldb: SNP&meth调控注释

mQTL Database Home Search Download Help



Search

Perform a quick search for mQTLs across the ARIES mQTL database.

[Take the tour »](#)

SNPs/CpGs

rs498045
cg24851651

>>> 数据库介绍

iMETHYL: DNA甲基化联合SNP及RNA转录



>>> 数据库介绍

Lnc2Meth: LncRNA&meth调控注释

Lnc2Meth
regulatory relationships between human long non-coding RNAs and DNA methylation

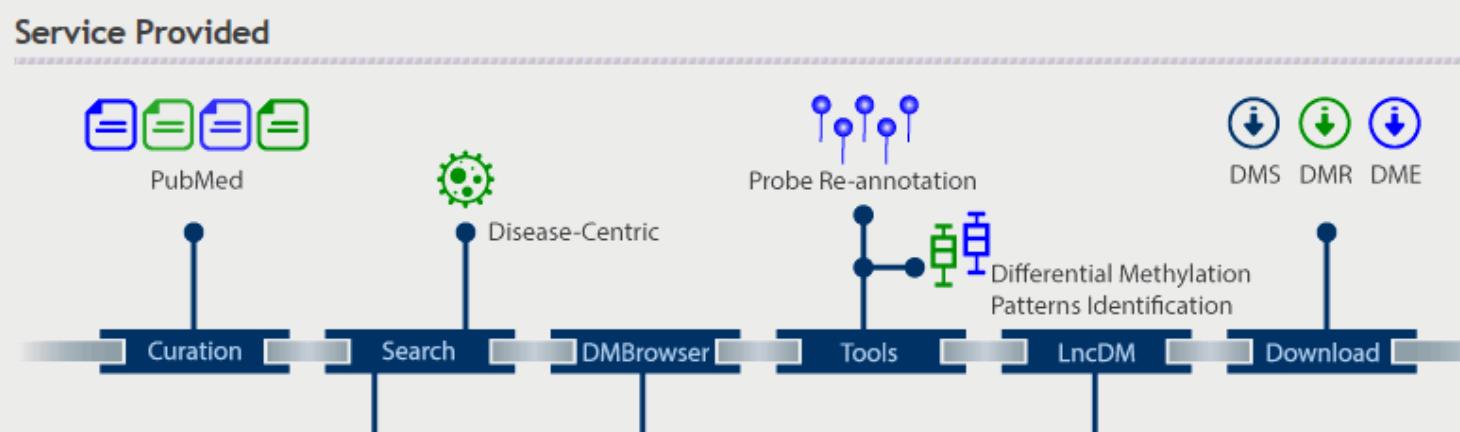
Home Curation Search DMBrowser Tools LncDM Download Help



Welcome to Lnc2Meth

Long noncoding RNAs (lncRNAs) play an important role in modulating gene expression or participating in some essential epigenetic regulation processes, including chromatin modification or DNA methylation. DNA methylation is a fundamental feature of epigenomes that can affect the expression of protein-coding or non-coding transcripts. Here, we constructed **Lnc2Meth**, aimed to provide a comprehensive resource and web tool for clarifying the regulatory relationships between human lncRNAs and associated DNA methylation in diverse diseases.

Service Provided



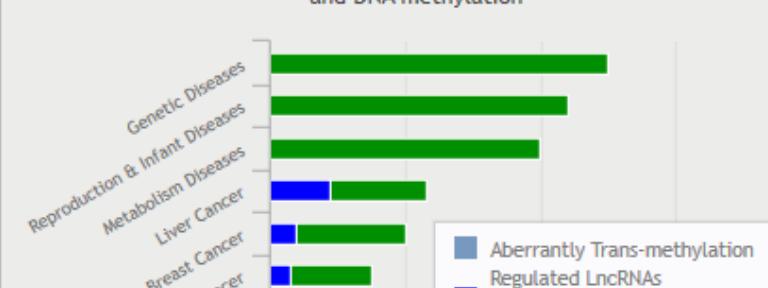
Quick Search

ex:MEG3, Breast Cancer

e.g. IncRNA "MEG3", disease "Breast Cancer"

Lnc2Meth Statistics

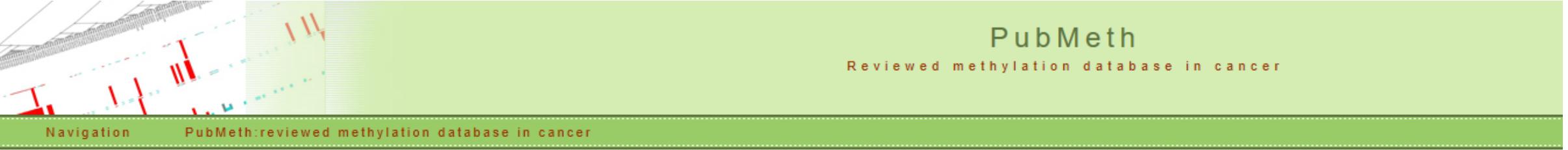
Statistics of curated regulatory relationship between human lncRNAs and DNA methylation



Disease Category	Aberrantly Trans-methylation	Regulated lncRNAs
Genetic Diseases	Low	High
Reproduction & Infant Diseases	Low	High
Metabolism Diseases	Low	High
Liver Cancer	High	Low
Breast Cancer	High	Low

>>> 数据库介绍

PubMeth: 癌症相关的甲基化基因数据库



The screenshot shows the homepage of the PubMeth database. At the top right, the logo "PubMeth" is displayed above the text "Reviewed methylation database in cancer". On the left, there is a navigation menu with links to Home, Search PubMeth, Tutorials, PubMeth creation, Contact & disclaimer, and Submit data to PubMeth. Below the menu, a section titled "PubMeth: start searching" provides information on how to search the database, mentioning two ways: gene-centric and genome-wide. It includes a "Gene-centric:" section with a detailed description of the search process and a "Genome-wide:" section. The main content area features a 3D visualization of DNA methylation data, showing red and green tracks on a grid.

Navigation **PubMeth: reviewed methylation database in cancer**

Home **PubMeth: start searching**

Search PubMeth **reviewed methylation database in cancer**

Tutorials **There are two ways of searching the database:**

Gene-centric:

which cancer types (and subtypes) are reported as being methylated in the genes that are searched?

> **browse through genes**

- i** Browse through the genes in PubMeth: select a gene and discover in which cancer types it is described as methylated. Then continue browsing to see the full details for the chosen gene in a certain cancer type.
- +** Fast browsing as everything is precomputed, simply browse to discover
- No summarised overview of the genes of your interest, user could get lost in too much (levels of) information

> **start a gene-centric search**

- i** Specify the genes of your interest, check if they are present in PubMeth and get a summarized overview with only the genes you selected
- +** Strong focus: only the genes of interest are included; summarisation views are good guides

>>> 数据库介绍

DiseaseMeth: 人类疾病相关的甲基化基因数据库

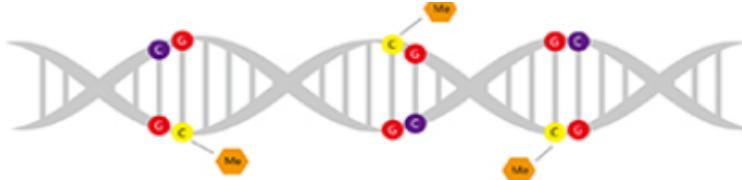
 **DiseaseMeth version 2.0**
The human disease methylation database

[Home](#) | [Search](#) | [Analyze](#) | [DisMethBrowser](#) | [Download](#) | [Help](#)

DiseaseSearch

[**< back**](#)

Disease:	<ul style="list-style-type: none">▶ <input checked="" type="checkbox"/> Cancer▶ <input type="checkbox"/> Genetic Disorder▶ <input type="checkbox"/> Metabolic Disorder▶ <input type="checkbox"/> Autoimmune Disease▶ <input type="checkbox"/> Neurological Disease	*
Gene Symbol: <input type="text"/> Example:TP53		<input type="button" value="search"/>



04

经典检测技术比拼

>>> 检测技术概述

Technique	Characteristics
Whole-genome bisulfite sequencing (BS-Seq or WGBS)	In bisulfite-treated DNA, unmethylated cytosines are converted into thymidines. ⁵⁰ Next-generation sequencing provides a complete overview of CpG methylation level at base-pair resolution.
Reduced-representation bisulfite sequencing (RRBS) or restriction enzyme-enriched sequencing (rrBS-Seq)	RRBS involves digesting DNA with a methylation-insensitive enzyme to enrich the sample for CpG islands. The CpG-enriched sample is then bisulfite-treated and sequenced. RRBS is an efficient technique that is suitable for obtaining information from most CpG islands and information about sequences outside CpG-rich regions. ⁵¹⁻⁵²
Affinity-enrichment-based sequencing techniques (MBD-Seq or MeDIP-Seq)	MBD-Seq ⁵³ and MeDIP-seq ⁵⁴ combine the advantages of next-generation sequencing and enrichment of methylated regions by immunoprecipitation.
DNA methylation arrays	CpG-specific array technology is an alternative option for determining a genome-wide DNA methylation profile. The Human Methylation 450 beadchip assay (Illumina) covers 99% of all human RefSeq ⁵⁵ genes and approximately 450,000 CpGs overall. ⁵⁶
Locus-specific DNA methylation analysis	In addition to genome-wide technologies, locus-specific identification of the DNA methylation level is a cost-effective strategy, especially if single genes are already established as biomarkers for diagnosis or prognosis.

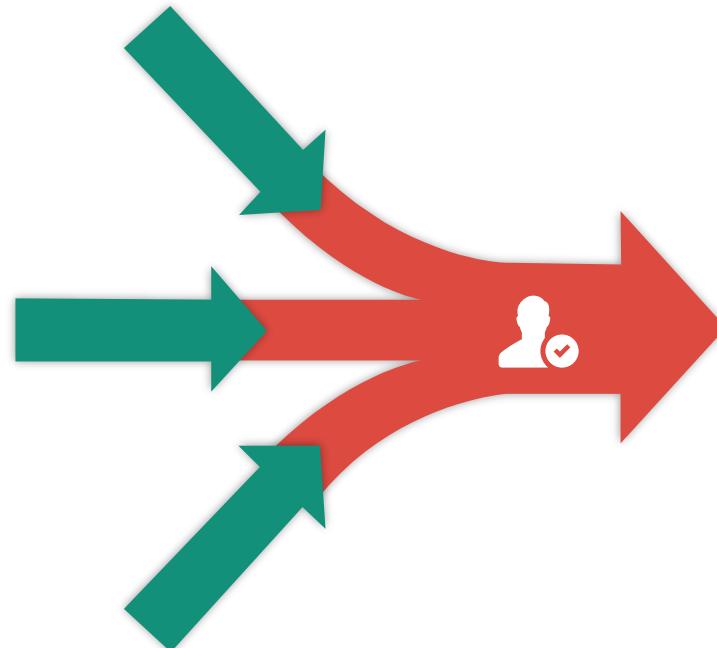


>>> 检测技术对比

Quantitative comparison of DNA methylation assays for biomarker development and clinical applications JULY 2016 *Nature biotechnology*

思路设计

- ✓ 选取6种global assays技术 & 21种locus-specific assays
- ✓ 32个样本，来自7个不同城市的18个实验室独立检测
- ✓ 数据灵敏度、重复性、准确性、通量参数对比



三个性能最佳的检测技术：

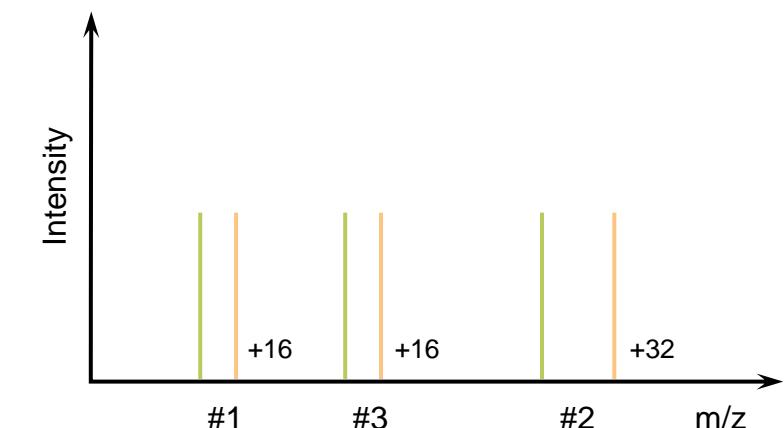
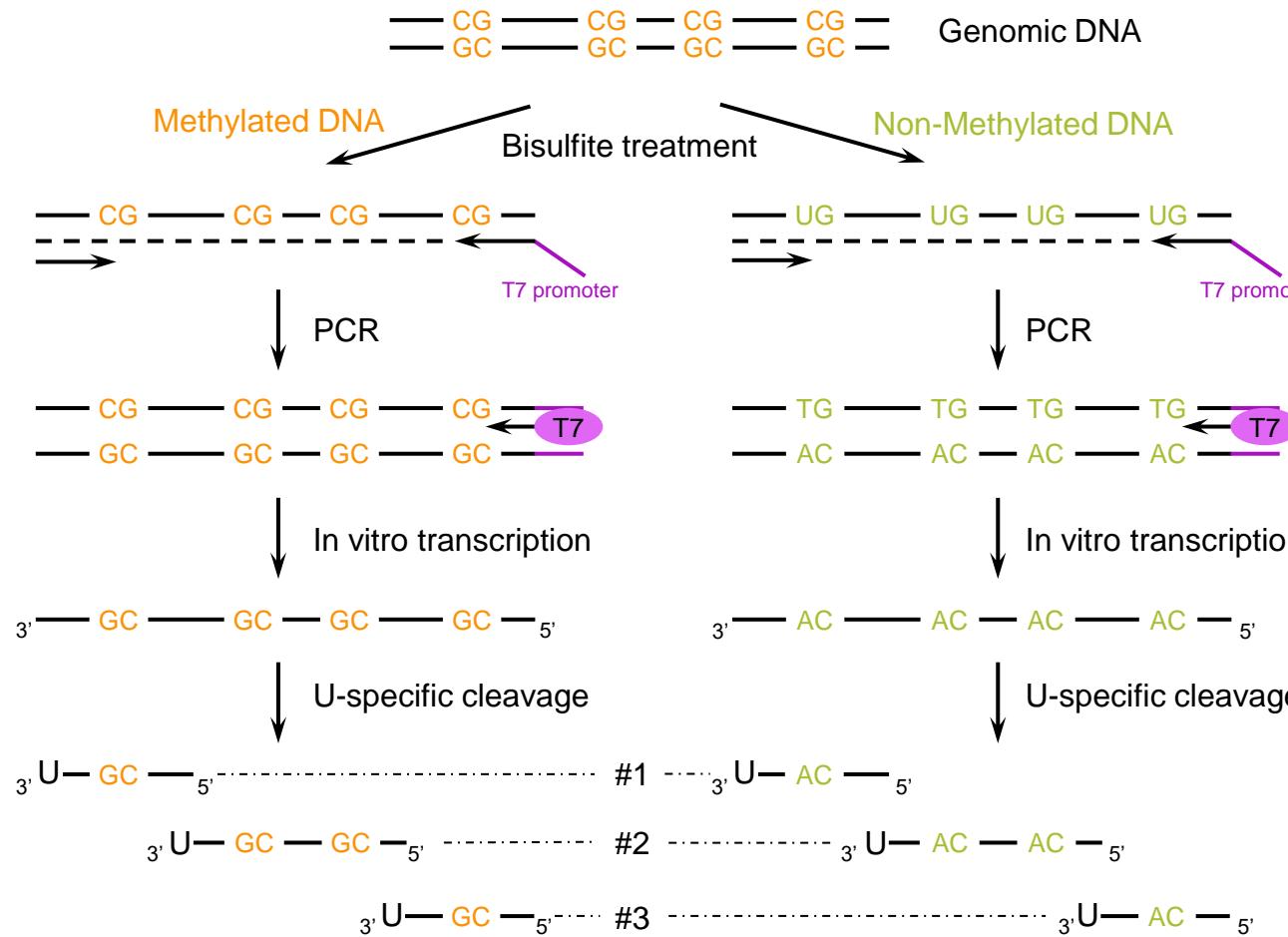
- ✓ Among the absolute DNA methylation assays, AmpliconBS and Pyroseq showed the best all-round performance, closely followed by EpiTyper
- ✓ Global assays present lower accuracy

最佳技术对比

- ✓ Pyroseq can work well even on minute amounts of highly fragmented DNA
- ✓ EpiTyper provides the highest sample throughput
- ✓ AmpliconBS is the best choice for assaying dozens of genomic regions in parallel

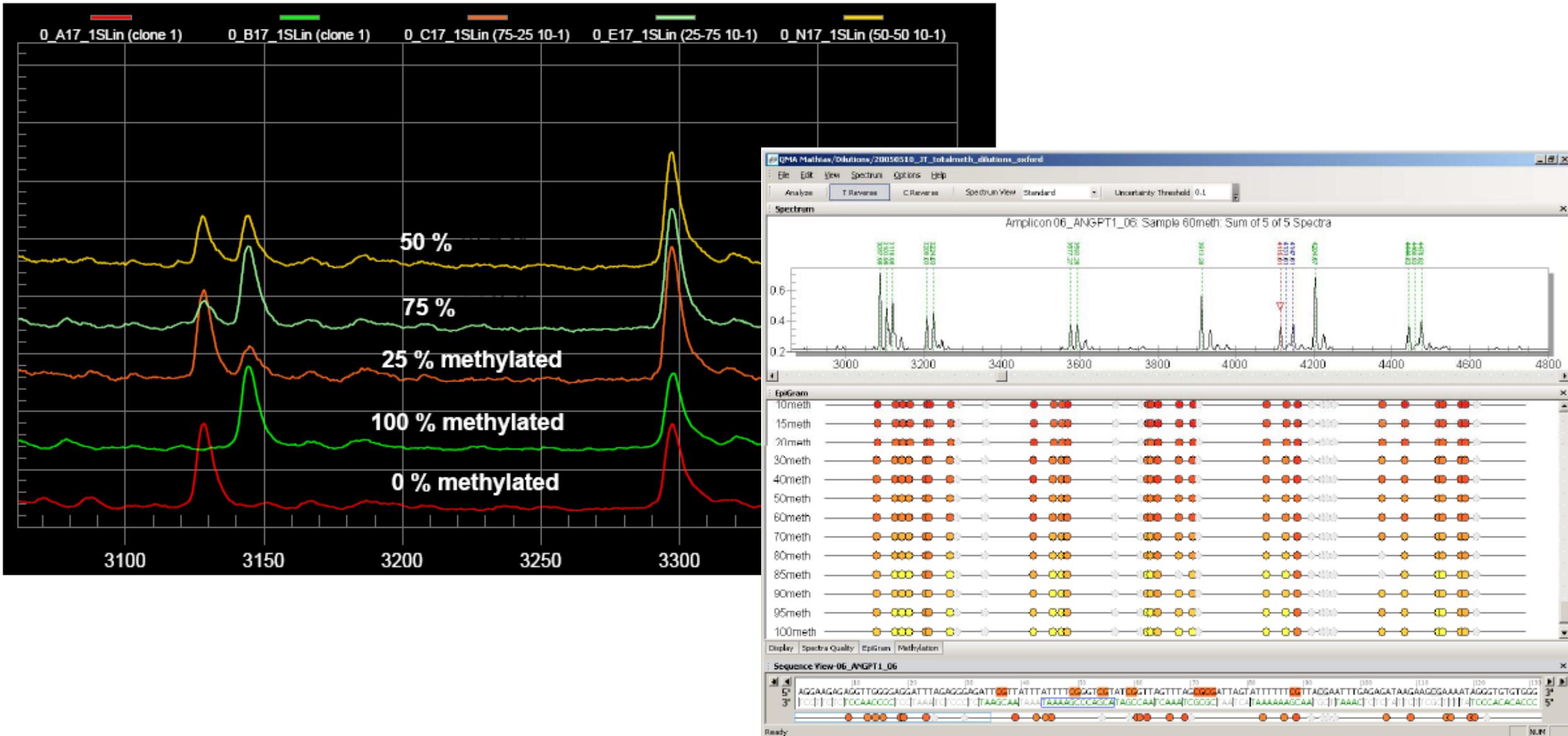
>>> 检测技术对比

Massarry(EpiTyper)甲基化技术原理



>>> 检测技术对比

Massarry(EpiTyper)甲基化数据展示



>>> 检测技术对比

Massarry(EpiTyper)甲基化技术参数



中通量：适用1-20个片段的中通量验证，样本数量不受限

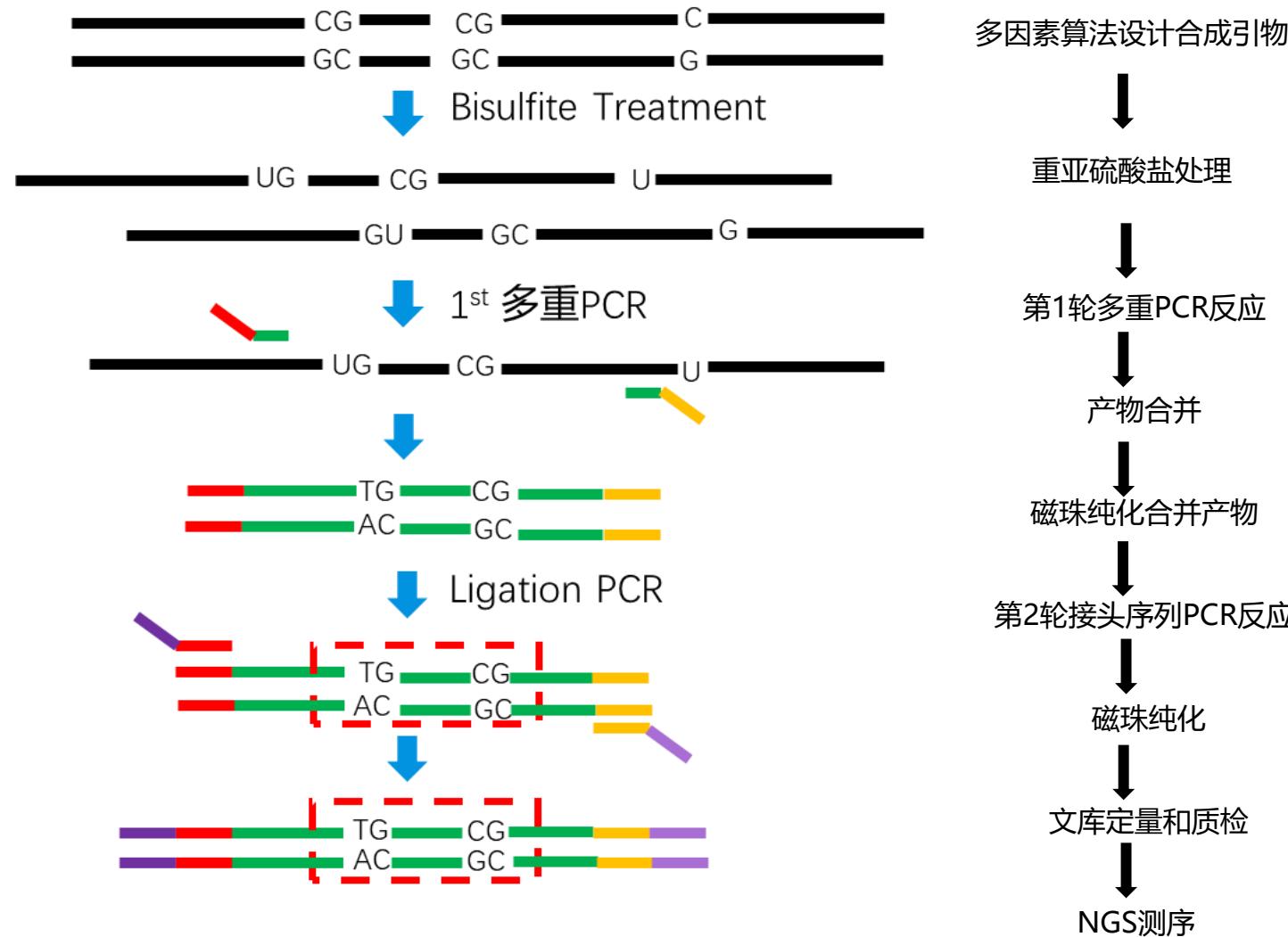
长片段：覆盖约400-600bp长度的区域位点，属于最长的locus-specific assays

高准确：扩增子内单个CpG sites进行准确定量

数据简便：可视化数据产出

>>> 检测技术对比

Multi-PCR NGS(AmpliconBS) 甲基化技术原理



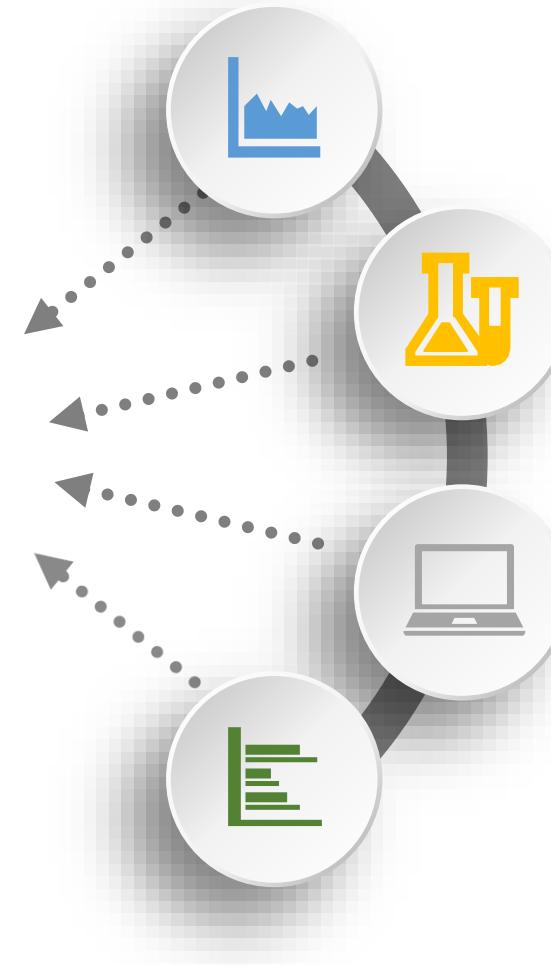
>>> 检测技术对比

Multi-PCR NGS(AmpliconBS)甲基化数据展示

Chr	Start	End	Methy_Level	Methy_CpG_Num	unMethy_CpG_Num
chr1	9352405	9352406	44	5171	6417
chr1	9352407	9352408	39	4590	7003
chr1	9352421	9352422	10	1174	10427
chr1	9352428	9352429	16	1956	9648
chr1	9352437	9352438	32	3780	7819
chr1	9352443	9352444	15	1768	9831
chr1	9352452	9352453	53	6188	5403
chr1	9352457	9352458	23	2687	8913
chr1	9352459	9352460	19	2314	9288
chr1	9352463	9352464	64	7490	4101
chr1	9352467	9352468	77	9030	2570
chr1	9352475	9352476	64	7512	4087
chr1	9352500	9352501	24	2819	8772
chr1	9352503	9352504	54	6328	5264
chr1	9352515	9352516	24	2866	8725
chr1	9352521	9352522	50	5806	5781
chr1	150293836	150293837	1	40	2218
chr1	150293844	150293845	1	24	2237
chr1	150293847	150293848	1	28	2231
chr1	150293853	150293854	0	16	2242
chr1	150293881	150293882	1	28	2234
chr1	150293887	150293888	0	16	2246
chr1	150293889	150293890	2	51	2210

>>> 检测技术对比

Multi-PCR NGS(AmpliconBS)甲基化技术参数



高通量：适用20-200个片段的中通量验证

短片段：覆盖约100-250bp长度的区域位点

高准确：扩增子内单个CpG sites进行准确定量，平均测序深度>1000

一箭双雕：可同时对于该区域的SNP分型，进行ASM数据分析

05

代表文献解读

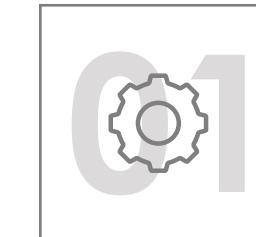
>>> 代表文献解读

经典EWAS表观基因组关联研究



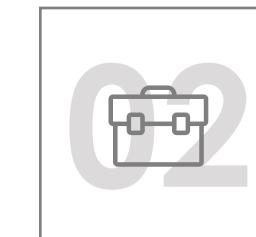
Parkinson' s disease is associated with DNA methylation levels in human blood and saliva

Genome Medicine (2017)



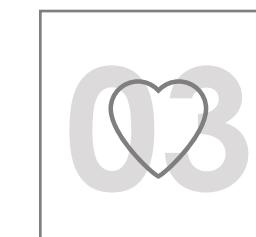
方案设计

- ✓ 样本表型：帕金森 样本类型：外周血、唾液
- ✓ 科研目标：发现外周血及唾液中与PD显著相关的甲基化位点及基因；血液中细胞类型差异对于显著位点的影响；易感基因的功能通路分析



检测技术

- ✓ Global assay: EWAS-450K芯片



结果讨论

- ✓ 在外周血及唾液两个不同样本类型来源，均发现显著与PD关联的甲基化位点；
- ✓ 最显著关联甲基化位点程度变化可有效反应血液中细胞类型的变化
- ✓ 关键易感基因通路分析与免疫调控系统有关

基于EWAS建立疾病预警模型研究



DNA methylation array analyses identified breastcancer-associated HYAL2 methylation in peripheral blood

Int. J. Cancer 2015 Apr



方案设计

- ✓ 样本表型：乳腺癌 样本类型：外周血
- ✓ 科研目标：通过外周血中特定甲基化位点进行早期BC诊断



检测技术

- ✓ Global assay: EWAS-27K芯片
- ✓ Locus-Specific assay: Massarray技术



结果讨论

- ✓ 早期BC/健康人群分组EWAS芯片关联分析，筛选显著相关甲基化位点及基因。
- ✓ 独立样本验证显著甲基化位点，并与RNA表达进行eQTM分析，存在负相关调控关系
- ✓ 采用显著位点建立预测模型，有效区分早期BC及健康人群(AUC=0.89)，且能够有效识别年轻女性早期BC (AUC=0.87)

EWAS&RNA转录组联合研究

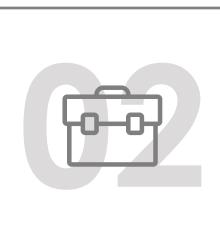


TElevated BICD2 DNA methylation in blood of major depressive disorder patients and reduction of depressive-like behaviors in hippocampal Bicd2-knockdown mice PNAS
PNAS 2022



方案设计

- ✓ 样本表型：重度抑郁症 样本类型：外周血
- ✓ 科研目标：通过全基因组甲基化与RNA转录组的联合分析，筛选关键调控靶基因，并在in vivo及in vitro水平进行分子调控功能机制验证



检测技术

- ✓ 甲基化：EWAS-450K芯片
- ✓ RNA转录组：全转录组表达谱芯片



结果讨论

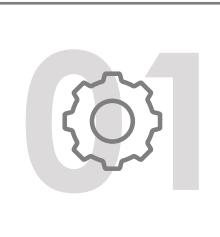
- ✓ 通过EWAS芯片和RNA表达谱芯片的联合分析发现，BICD2基因是通过DNA甲基化芯片分析得到的甲基化差异最为显著的基因，位于其3UTR区的cg14341177位点在重度抑郁症患者中表现为高甲基化，而RNA表达谱芯片分析显示在重度抑郁症患者外周血中BICD2的mRNA表达水平下降。
- ✓ 采用Agena MassARRAY方法开展独立大样本靶向位点验证，进行独立样本的组学数据统计性验证。
- ✓ 基于抑郁模型小鼠，通过腺相关病毒转基因方法、基因表达分析和行为学检测等方法进行了BICD2的功能和机制研究。

EWAS&GWAS联合研究



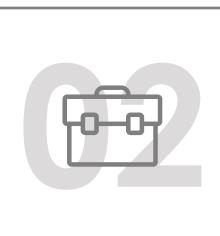
Trans-ancestry genome-wide association study identifies 12 genetic loci influencing blood pressure and implicates a role for DNA methylation.

Nature Genet. 2015



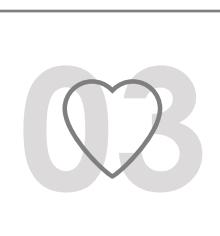
方案设计

- ✓ 样本表型：高血压 样本类型：外周血
- ✓ 科研目标：分析SNP-meth-blood pressure的相关性及构建调控机制通路



检测技术

- ✓ 甲基化：EWAS-450K芯片
- ✓ SNP分型：GWAS芯片



结果讨论

- ✓ 与血压最显著相关的SNP位点与甲基化位点存在显著的cis-/trans-mQTL调控关系
- ✓ 发现关键易感基因AMH的一个SNP位点直接参与调控不同基因29个CpG sites的甲基化变化
- ✓ 建立遗传变异-表观基因甲基化-血压变化的因果调控效应

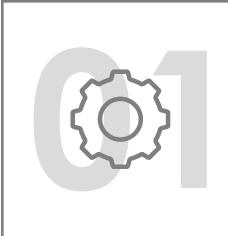
>>> 代表文献解读

EWAS&Metabolomics联合研究



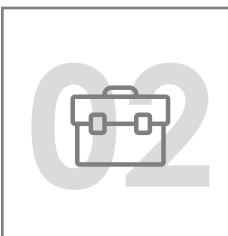
Untargeted metabolomics reveals multiple metabolites influencing smoking-related DNA methylation

Epigenomics. Mar 2018



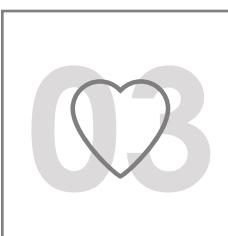
方案设计

- ✓ 样本表型：男性双胞胎 样本类型：外周血、血浆
- ✓ 科研目标：分别分析吸烟与代谢物的关联性，吸烟与甲基化的关联性，以及与吸烟显著相关代谢物的EWAS分析。



检测技术

- ✓ 甲基化：EWAS-450K芯片
- ✓ 代谢组：LC-MS/MS非靶向代谢组技术



结果讨论

- ✓ 代谢组关联分析筛选了12个与吸烟显著相关的代谢组
- ✓ 发现与上述代谢组显著负相关的甲基化位点
- ✓ 构建吸烟-甲基化-代谢物的调控关联通路

>>> 代表文献解读

EWAS&Proteomics联合研究



Epigenetics meets proteomics in an epigenome-wide association study with circulating blood plasma protein traits

NATURE COMMUNICATIONS (2020)



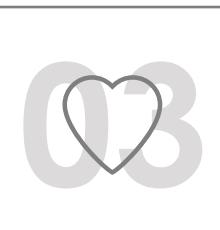
方案设计

- ✓ 样本表型：随机正常人群 样本类型：外周血、血浆
- ✓ 科研目标：探讨外周血中甲基化与蛋白质的关联分析，从表观基因层面挖掘新的生物功能通路



检测技术

- ✓ 甲基化：EWAS-450K芯片
- ✓ 蛋白组：SOMAscan



结果讨论

- ✓ 通过pQTM分析，筛选显著的cis-及trans-pQTM的甲基化位点及蛋白质，构建正常人群的关联调控网络图
- ✓ 通过功能注释及分析，发现存在DNA甲基化-免疫蛋白-炎症感染的调控通路，为人群慢性初级炎症提供了数据支撑

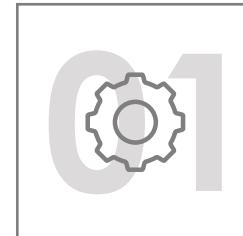
>>> 代表文献解读

EWAS&Microbiota联合研究



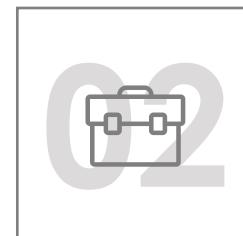
Gut Microbiota Composition Is Associated With the Global DNA Methylation Pattern in Obesity

Front. Genet., July 2019



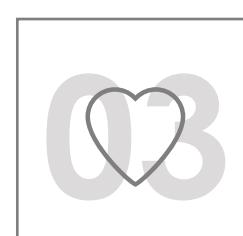
方案设计

- ✓ 样本表型：肥胖症 样本类型：外周血、内脏脂肪组织、粪便
- ✓ 科研目标：肥胖人群肠道微生物菌群与宿主表观基因甲基化是否关联，及网络调控机制预估



检测技术

- ✓ 甲基化：EWAS-850K芯片
- ✓ 微生物基因组：16S 测序 靶基因RNA定量：RT-qPCR



结果讨论

- ✓ 通过16S技术进行肥胖人群的肠道微生物基因组检测，并最终根据BFR菌群含量进行有效分组；
- ✓ 通过血液及组织EWAS数据的菌群人群差异组间关联分析，均分析获得显著甲基化位点；并对其中部分功能基因在组织中进行RNA定量检测，与甲基化存在负相关
- ✓ 肠道微生物可能通过影响宿主基因甲基化改变，并最终导致肥胖发生

>>> 代表文献解读

靶向基因SNP&meth联合研究



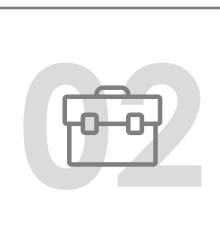
Association of
norepinephrine
transporter methylation
with in vivo
NET expression and
hyperactivity–impulsivity
symptoms in ADHD
measured with PET

MOLECULAR PSYCHIATRY 2019



方案设计

- ✓ 样本表型：注意力缺陷多动症 样本类型：外周血
- ✓ 样本分组：23个ADHD患者及23个正常对照；18个包含PET扫描定量脑组织不同部位NET的患者及对照



检测技术

- ✓ 甲基化：Massarray—SLC6A2 promoter
- ✓ SNP分型：Massarray—4个SNP



科研结论

- ✓ ADHD患者中，发现一个甲基化位点与脑组织不同部位NET浓度呈现负相关
- ✓ 一个甲基化区域与ADHD严重程度负相关
- ✓ 未发现SNP与DNA甲基化的交互作用及关联调控

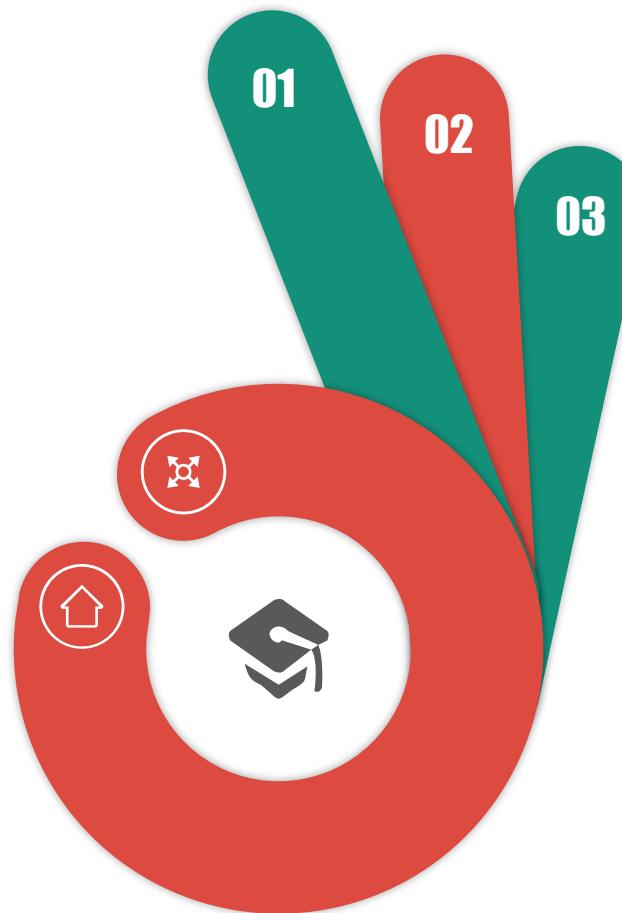
>>> DNA甲基化研究要素归纳

创新的方案策略

缜密的样本设置

细致的数据分析

准确的检测技术



>>> 博森技术服务项目一览图

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

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
- HLA-seq/TCR&BCR-seq

代谢组学服务

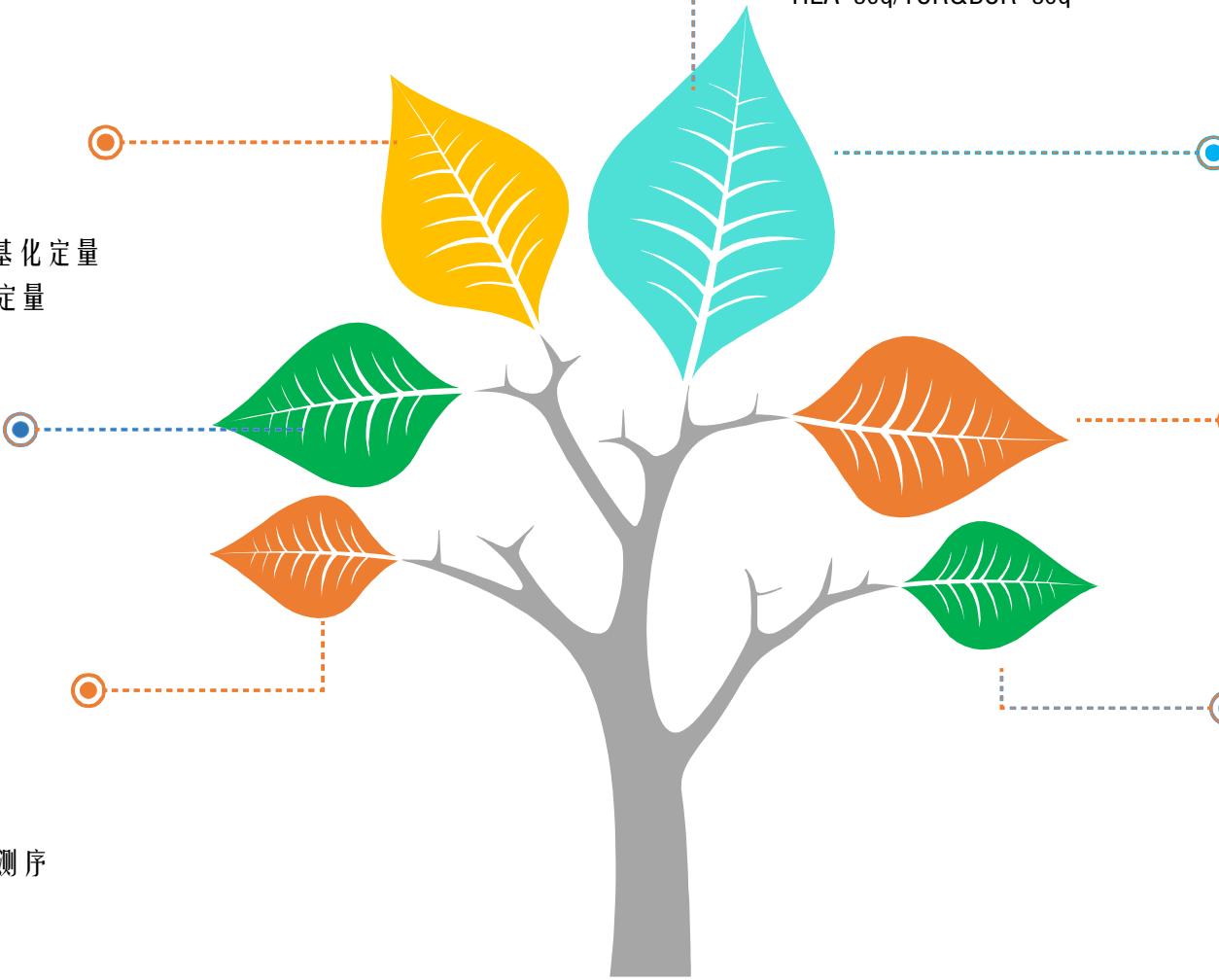
- 非靶向代谢组
- 非靶向脂质组
- 高通量靶向代谢组
- 靶向代谢组项目系列

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



更多技术服务：基因组学、微生物基因组学、单细胞组学、转录组学、
蛋白质组学、代谢组学等请访问公司网站www.biomiao.com
或与本公司区域销售索要相关资料