

第十二章 生物信息学的应用

第一节 精准医学

徐娟
哈尔滨医科大学

第一节

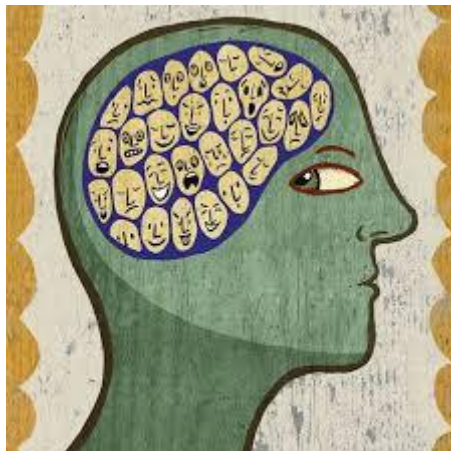
- 精准医学
- 疾病的遗传关联研究
 - 复杂疾病与遗传因素
 - 疾病遗传关联分析
 - 疾病的遗传异质性识别
 - 疾病的遗传机制解析
- 疾病的免疫调控解析
 - 复杂疾病和免疫微环境
 - 免疫调控机制的计算解析
 - 基于免疫微环境实现免疫分型
 - 免疫微环境指导肿瘤精准治疗

- 复杂疾病（**Complex Disease**）具有表型复杂性的特征，在人群中发病率高，是21世纪生物医学重大的挑战之一。
- 人类常见病，包括肿瘤、心脑血管病、代谢系统疾病、神经系统疾病、精神和行为异常等绝大多数都是复杂疾病。
- 中国死因监测数据显示，恶性肿瘤死亡占全部居民死因的近四分之一。

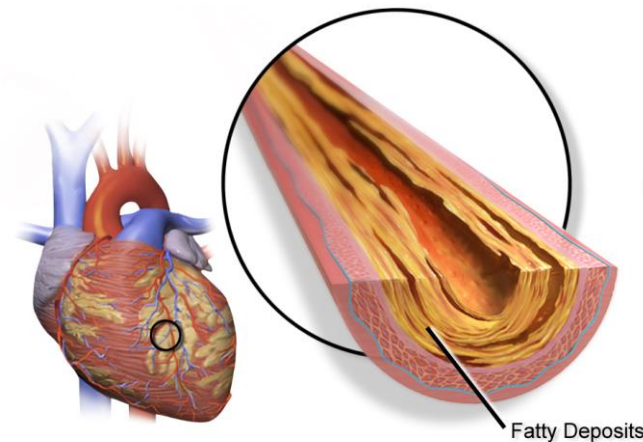
Diabetes



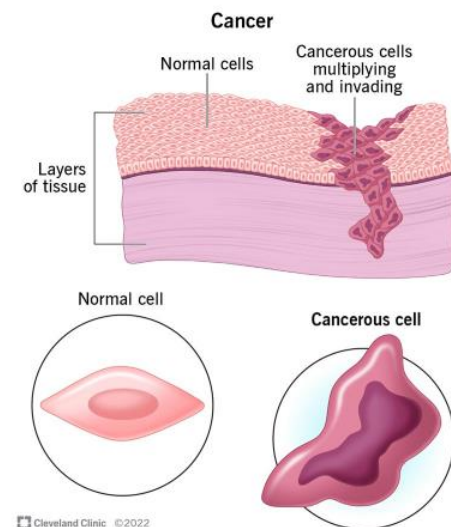
Schizophrenia

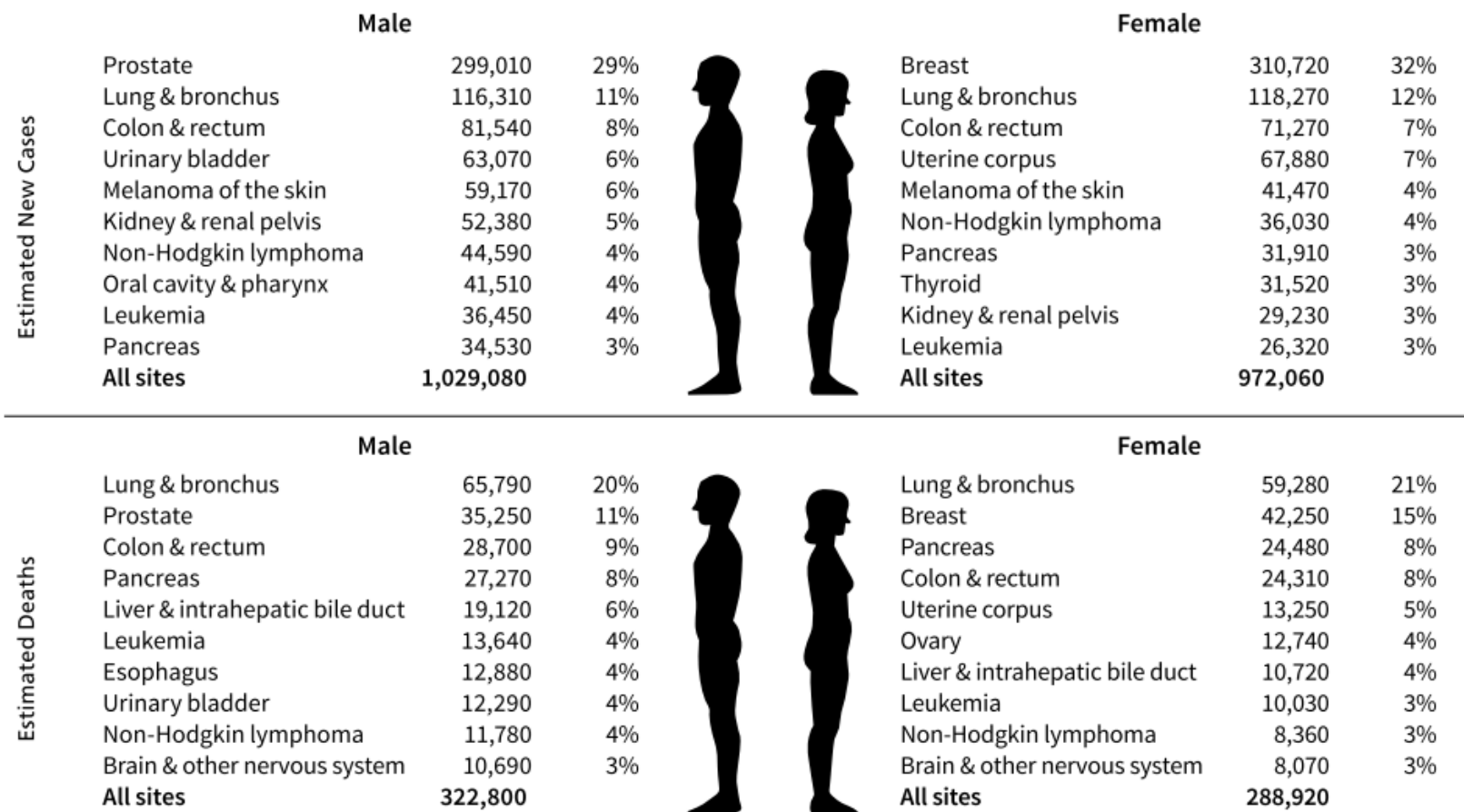


Cardiovascular disease



Cancer





- 美国2024年将有>200万例新发的癌症病例和>611万死亡病例，癌症是美国85岁以下人群的主要死亡原因。
- 男性死亡率排前三的癌症分别是前列腺癌、肺癌和结肠直肠癌
- 女性死亡排名前三的癌症分别是乳腺癌、肺癌和结直肠癌。
- 报告还指出癌症患者越来越年轻，年轻成人中癌症发病率不断增加，特别是宫颈癌和结直肠癌。

2024年美国癌症统计报告

中华肿瘤杂志 2024 年 3 月第 46 卷第 3 期 Chin J Oncol, March 2024, Vol. 46, No. 3



国家癌症中心

2022 年中国恶性肿瘤流行情况分析

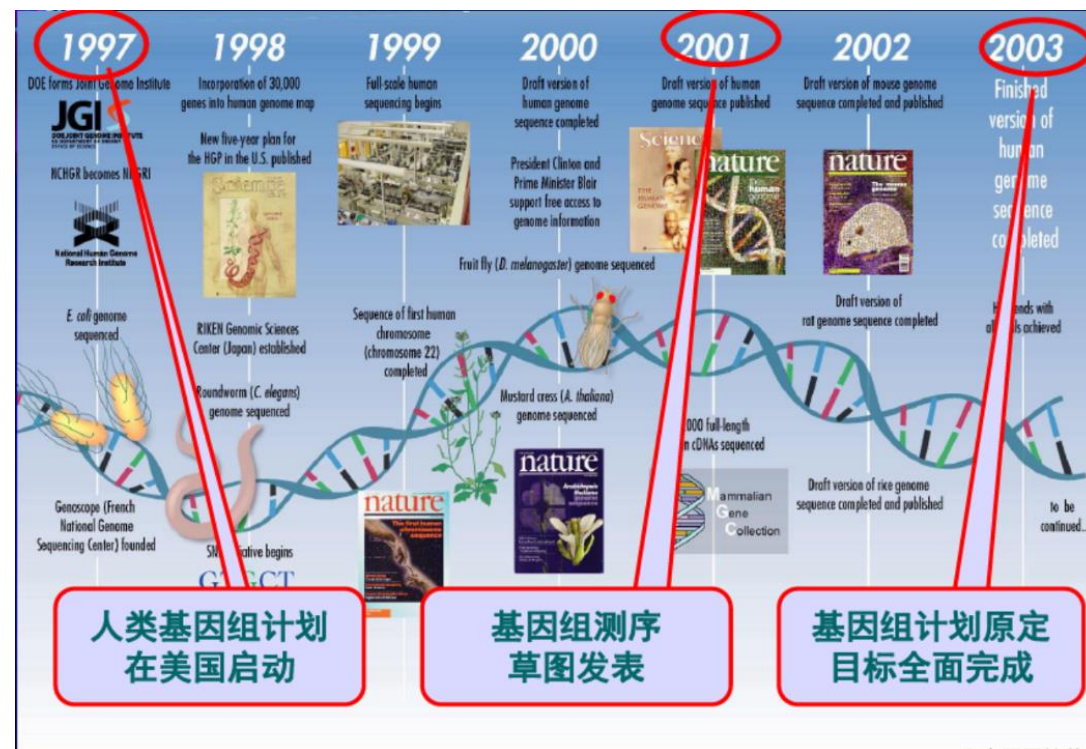
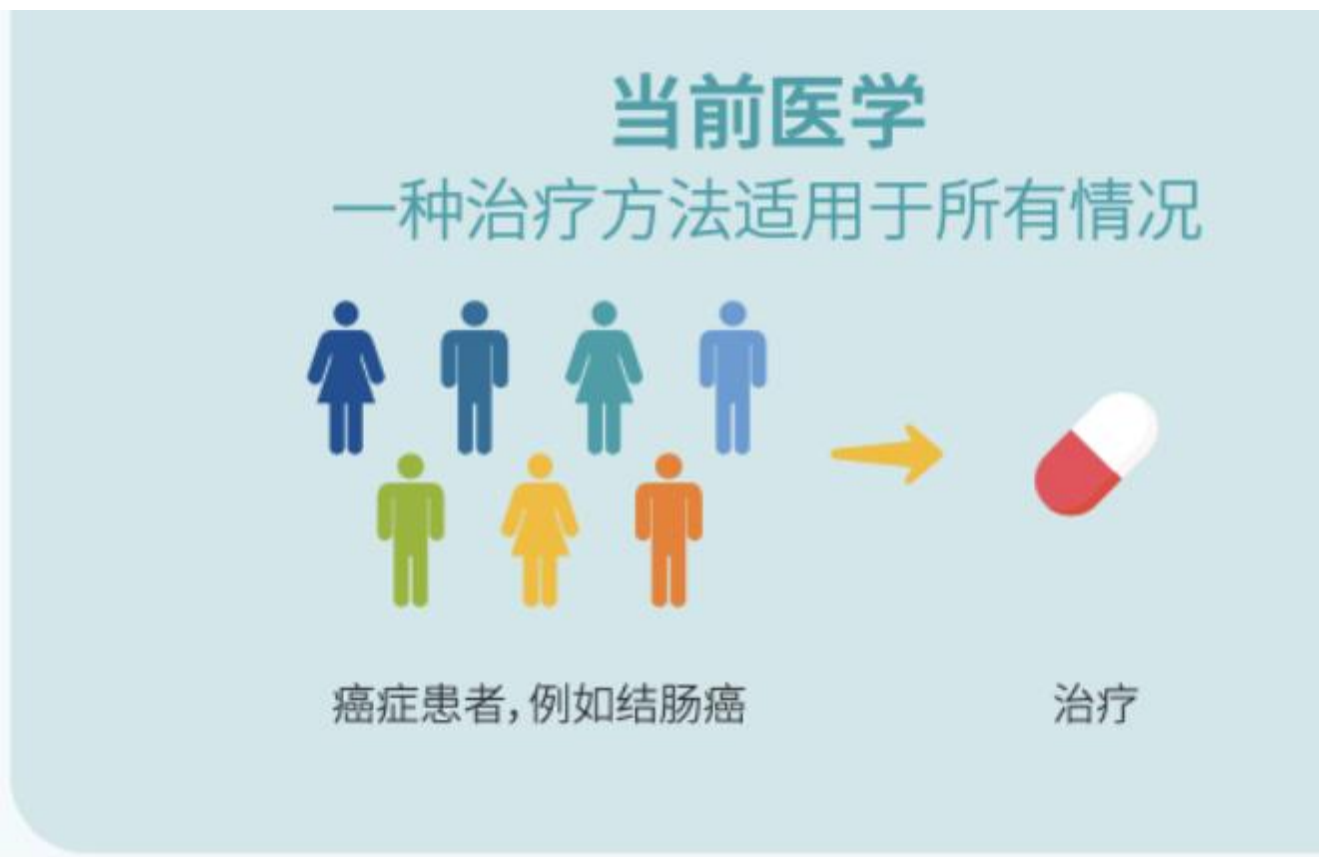
表 2 2022 年中国恶性肿瘤发病情况估计

肿瘤类别	总体			男性			女性		
	发病例数 (万)	中标率 (1/10 万)	世标率 (1/10 万)	发病例数 (万)	中标率 (1/10 万)	世标率 (1/10 万)	发病例数 (万)	中标率 (1/10 万)	世标率 (1/10 万)
全部恶性肿瘤	482.47	208.58	201.61	253.39	212.67	209.61	229.08	208.08	197.03
口腔癌	6.51	2.76	2.72	4.56	3.89	3.87	1.95	1.67	1.60
鼻咽癌	5.10	2.53	2.36	3.67	3.61	3.39	1.44	1.44	1.33
食管癌	22.40	8.24	8.32	16.75	12.90	13.09	5.65	3.81	3.78
胃癌	35.87	13.79	13.72	24.66	19.43	19.47	11.21	8.49	8.29
结直肠癌	51.71	20.29	20.10	30.77	24.83	24.74	20.94	15.97	15.70
肝癌	36.77	15.29	15.03	26.79	23.14	22.72	9.98	7.49	7.42
胆囊癌	3.11	1.14	1.14	1.27	0.97	0.97	1.85	1.30	1.30
胰腺癌	11.87	4.45	4.44	6.71	5.28	5.29	5.15	3.66	3.63
喉癌	2.95	1.14	1.16	2.72	2.15	2.19	0.23	0.17	0.17
肺癌	106.06	40.82	40.78	65.87	51.75	52.03	40.19	30.69	30.34
皮肤黑色素瘤	0.88	0.37	0.37	0.44	0.37	0.37	0.44	0.37	0.36
女性乳腺癌	35.72	35.30	33.04	-	-	-	35.72	35.30	33.04
子宫颈癌	15.07	14.88	13.83	-	-	-	15.07	14.88	13.83
子宫体癌	7.77	7.03	6.84	-	-	-	7.77	7.03	6.84
卵巢癌	6.11	5.98	5.68	-	-	-	6.11	5.98	5.68
前列腺癌	13.42	9.81	9.68	13.42	9.81	9.68	-	-	-
睾丸癌	0.35	0.46	0.41	0.35	0.46	0.41	-	-	-
肾癌	7.37	3.17	3.13	4.73	4.14	4.08	2.64	2.23	2.21
膀胱癌	9.29	3.48	3.44	7.32	5.70	5.67	1.97	1.41	1.39
脑肿瘤	8.75	4.21	4.17	4.24	4.20	4.13	4.51	4.22	4.20
甲状腺癌	46.61	28.85	24.64	12.49	15.81	13.25	34.12	42.48	36.51
淋巴瘤	8.52	3.88	3.77	4.81	4.45	4.34	3.71	3.33	3.21
白血病	8.19	4.35	4.54	4.70	4.95	5.14	3.50	3.77	3.94
其他恶性肿瘤	32.09	13.57	13.48	17.13	14.82	14.76	14.96	12.40	12.28

注：中标率：2000 年中国人口年龄标准化率；世标率：Segi's 世界人口年龄标准化率；- 为无数据

- 2022 年全国新发恶性肿瘤估计为 482.47 万例，恶性肿瘤总体中标发病率为 208.58/10 万，世界恶性肿瘤总体中标发病率为 201.61/10 万，高于世界总体水平。
- 我国恶性肿瘤发病、死亡数持续上升，每年恶性肿瘤所致的医疗花费超过 2200 亿，给社会和家庭带来了沉重的负担。
- 由于人口老龄化的加剧，预计 2040 年相比 2020 年，全球范围内癌症负担将增加 50%，届时新发癌症病例数将达到近 3000 万，这将对正经历社会和经济转型的中国而言影响是显著的。

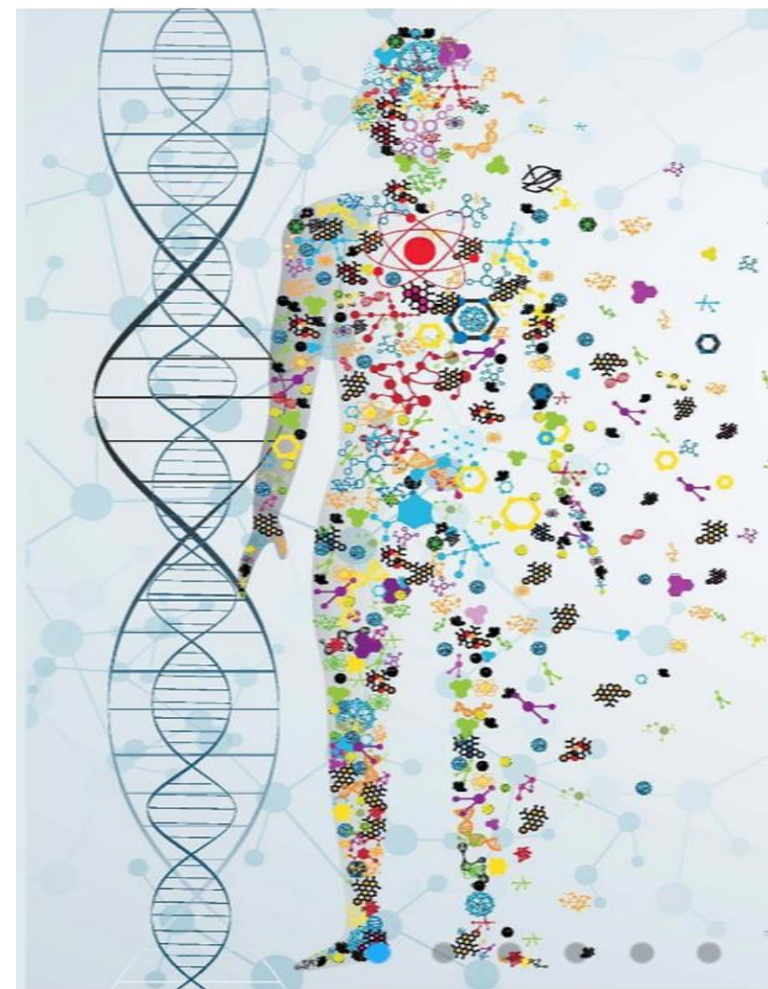
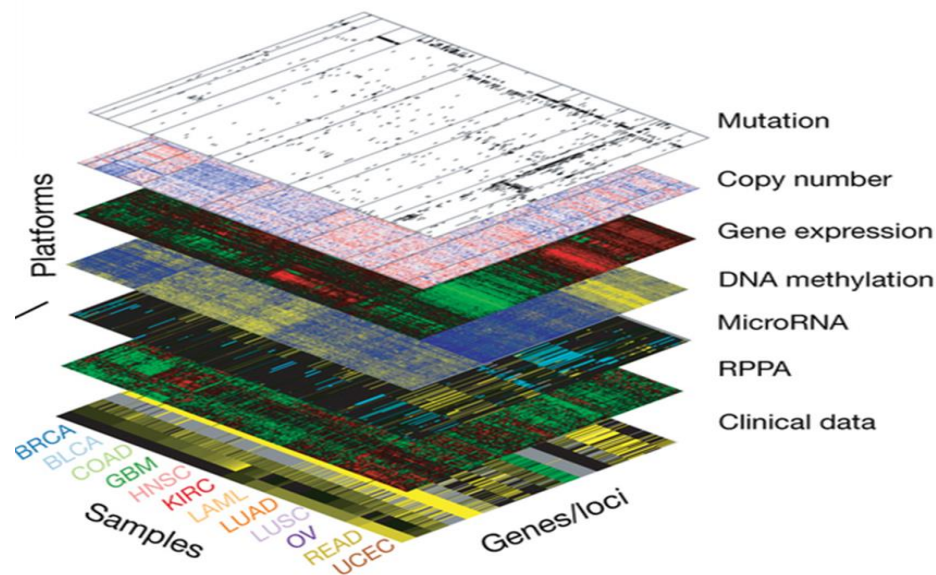
- 患者的个体差异、环境和生活方式等差异考虑较少。



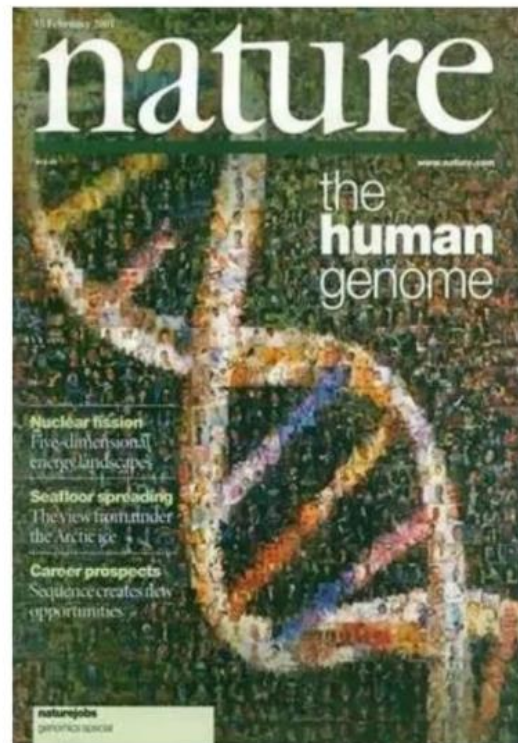
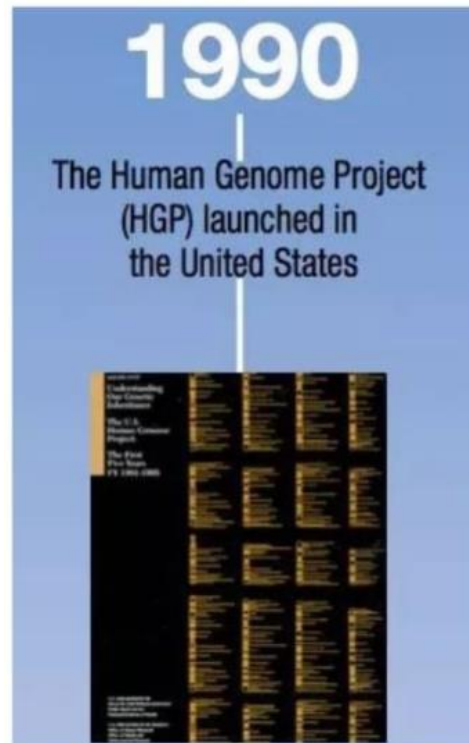
➤ 人类相关的组学计划层出不穷，揭示复杂疾病是遗传和环境因素共同作用的结果。



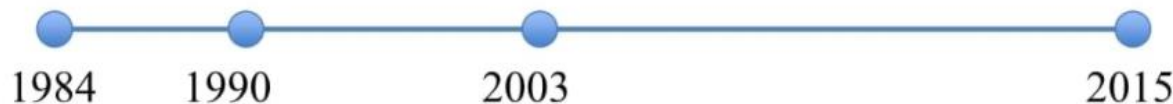
The Human
Phenome Project

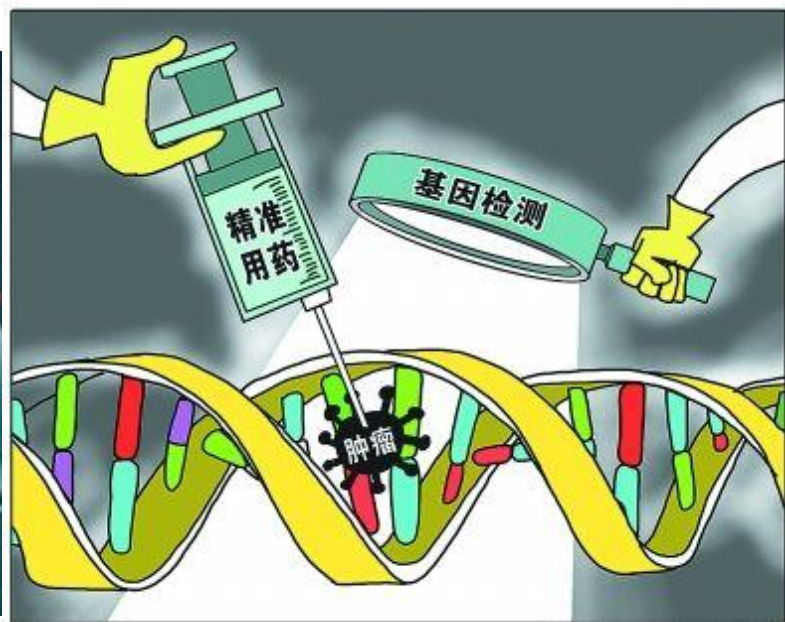


Human Genome Project (HGP)

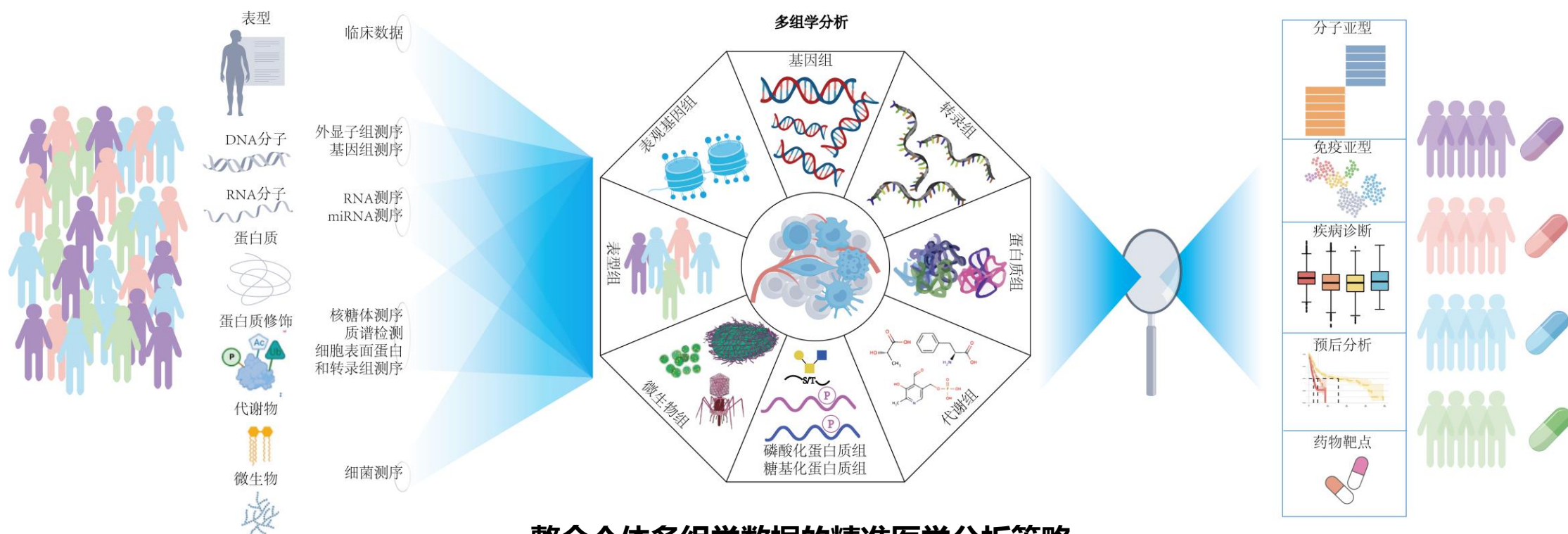


近年来，地中海贫血、白血病等患者受益于基因技术而被成功治愈，展示了基因科技造福人类的美好前景。

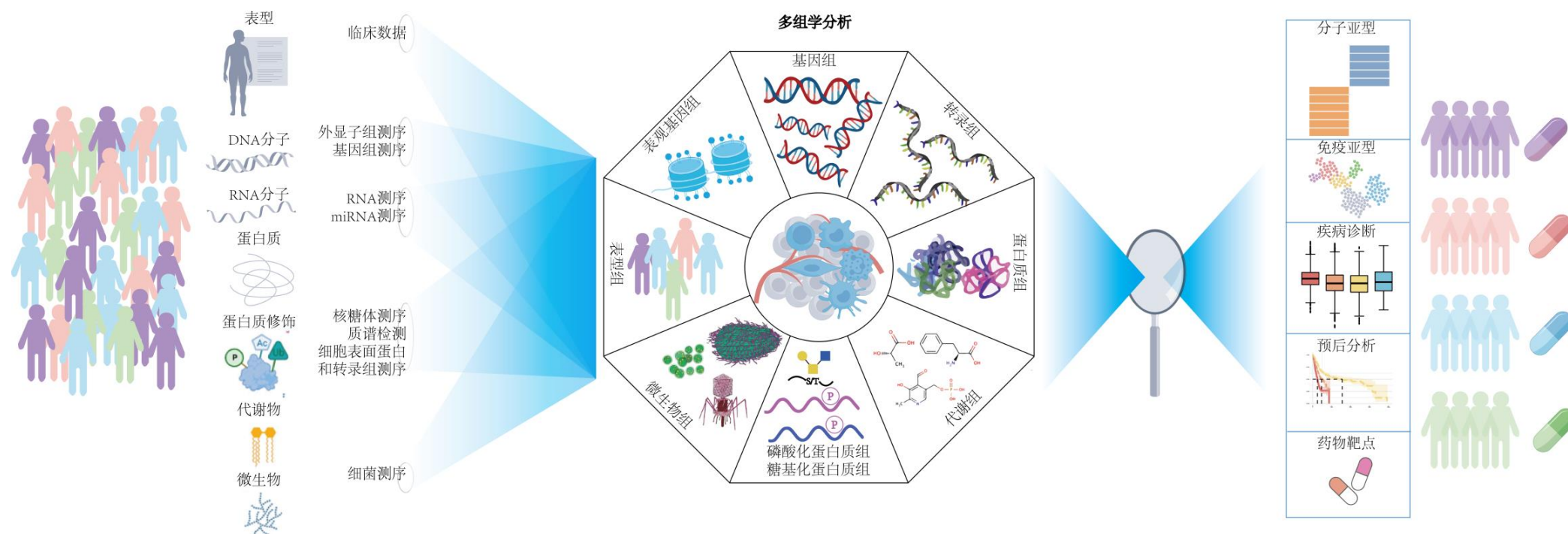




- 精准医学（**precision medicine**）是综合应用现代遗传技术，分子影像技术，生物信息学等技术，将个体的基因差异考虑到疾病的预防诊疗中，结合生活环境和临床数据，实现精准的疾病诊断并制定出具有个性化的预防治疗方案。



整合个体多组学数据的精准医学分析策略



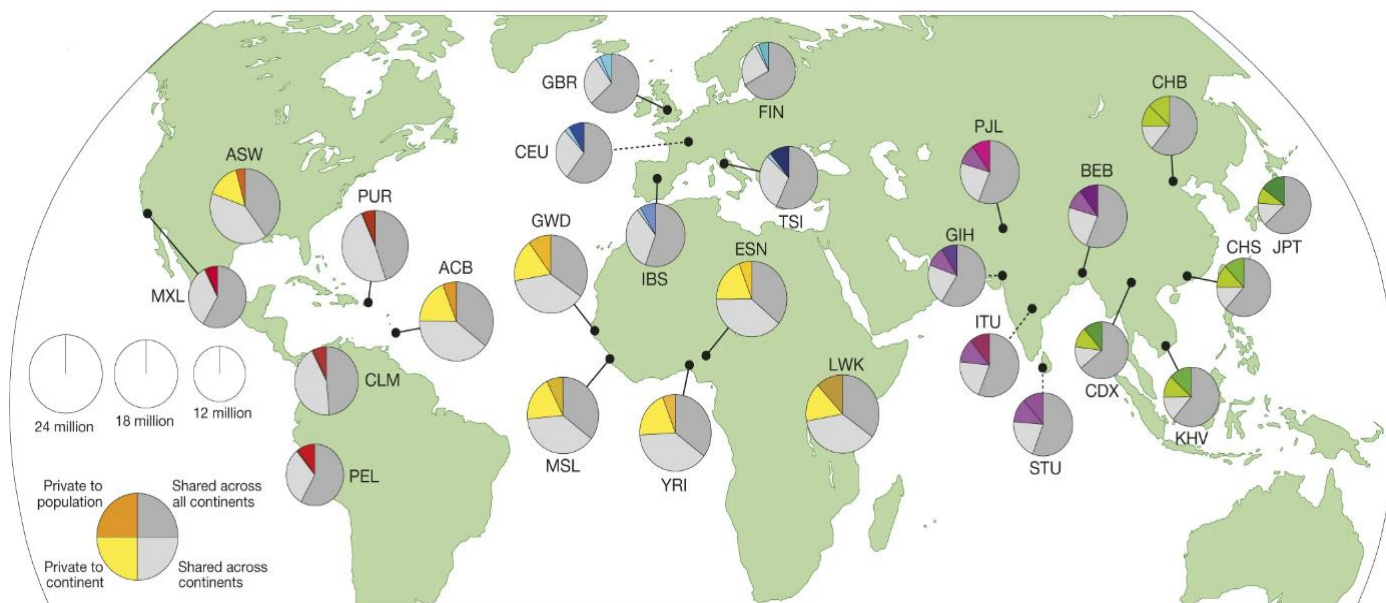
► 聚焦解决如下重大科学问题：

- 疾病发生发展机制的解析，提供疾病发生本质问题的回答；
- 分子诊断（分子分型、分子分期），为个性化治疗和预后判断提供科学依据；
- 生物标志物的发现和疾病早期诊断方法的探索建立，提供疾病治疗的有效时机；
- 靶向药物的研发或个体化治疗方案的给定，特异性地有效治疗疾病，最终实现对疾病和特定人群的个体化精准医疗。

第一节

- 精准医学
- 疾病的遗传关联研究
 - 复杂疾病与遗传因素
 - 疾病遗传关联分析
 - 疾病的遗传异质性识别
 - 疾病的遗传机制解析
- 疾病的免疫调控解析
 - 复杂疾病和免疫微环境
 - 免疫调控机制的计算解析
 - 基于免疫微环境实现免疫分型
 - 免疫微环境指导肿瘤精准治疗

-
- A collage of four images. The top-left image shows a dense crowd of people. The top-right image shows a semi-circular chart with a globe in the center. The bottom-left image shows a DNA double helix. The bottom-right image shows a stylized human figure composed of dots.



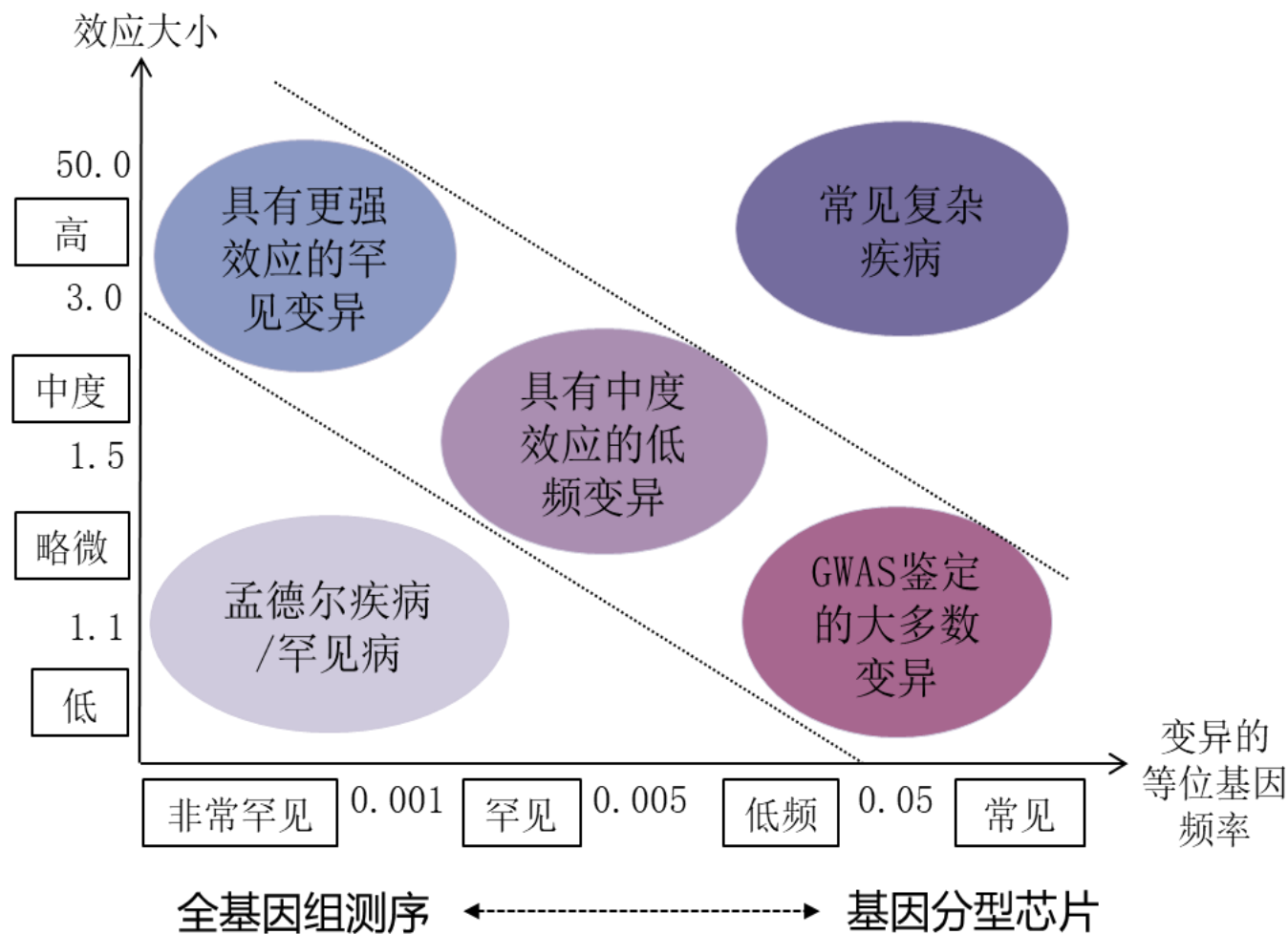
不同地域人群携带共有变异 和 各自不同的独特变异

(The 1000 Genomes Project Consortium. *Nature*. 2015)

- 全基因组关联研究（genome-wide association study, GWAS）是鉴定疾病易感性基因的一套方法体系。
- GWAS catalog 数据库收录了约6,500个GWAS项目，涉及超3,500个表型特征，超过53万对基因-复杂性状关联（截至2023年7月）



- 复杂疾病的遗传风险位点可由多个具有微小效应的常见变异组成，也可包含具有较大表型效应的罕见变异
- 全基因组测序（whole genome sequencing, WGS）可以有效全面检测整个基因组中的低频和罕见变异，有利于鉴定新的、效应强的遗传风险位点

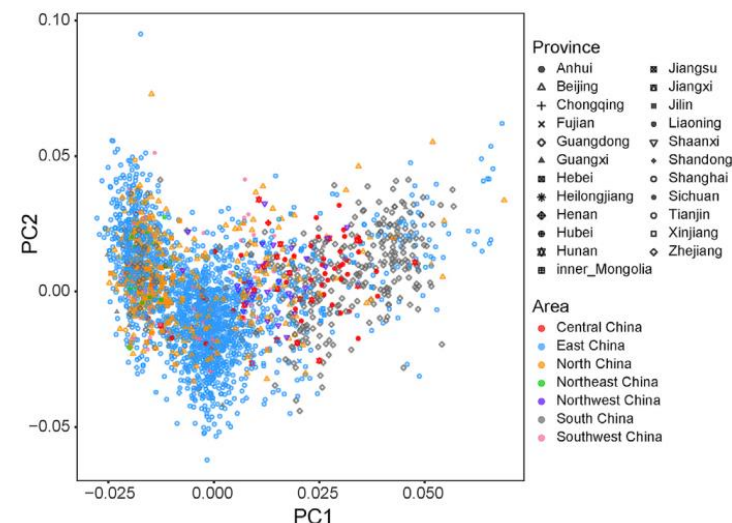
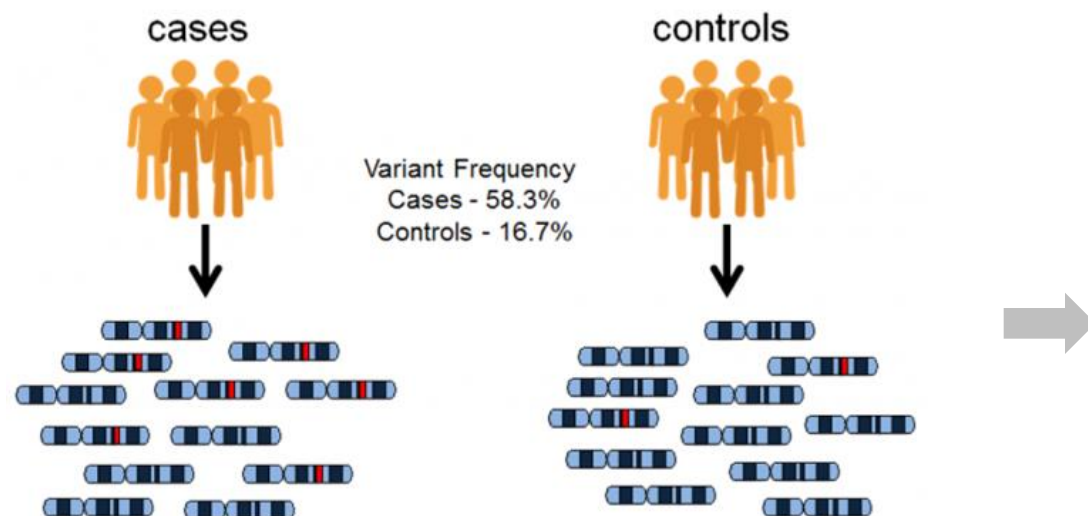


➤ 疾病遗传关联研究路线



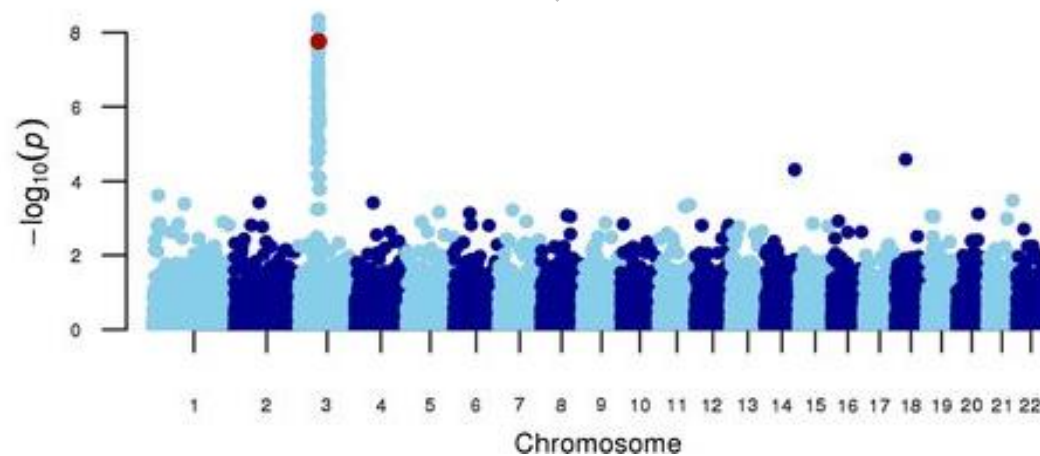
➤ 人群队列收集

通过测定疾病人群队列和对照人群队列的基因组序列，进而比较不同表型人群的遗传变异差异，是准确鉴定致病遗传因素的重要手段。



选取遗传背景相匹配的病例&对照

(Zhang., Cell Reports, 2021)



队列关联研究

➤ 国际代表性的大型全基因组队列研究

项目	发表时间	人数	人群	深度	变异类型	PubMed ID
1KGP	2015/ 2022年	3202人	世界人群	30×	SNP/InDel/SV	26432245/ 36055201
gnomAD	2020年	15708人	世界人群	18×	SNP/InDel/SV/多核 苷酸变异	32461613
冰岛基因组计划	2015年	2636人	冰岛人	20×	SNP/InDel	25807286
UKBB	2022年	150119人	世界人群	32×	SNP/InDel/SV/微卫 星	35859178
TopMed	2021年	53831人	美国	30×	SNP/InDel	33568819
All of us	-	约100万人	世界人群	-	-	-
新加坡 SG10K 计划	2019年	4810人	新加坡人	14×	SNP/InDel	31626772
亚洲 GAsP 计划	2019年	1267人	亚洲人	36×	SNP/InDel	31802016
荷兰 ALS MinE 计划	2021年	6538人	荷兰人	25×	SNP/InDel/微卫星	34873335

➤ 国内代表性的大型全基因组队列研究

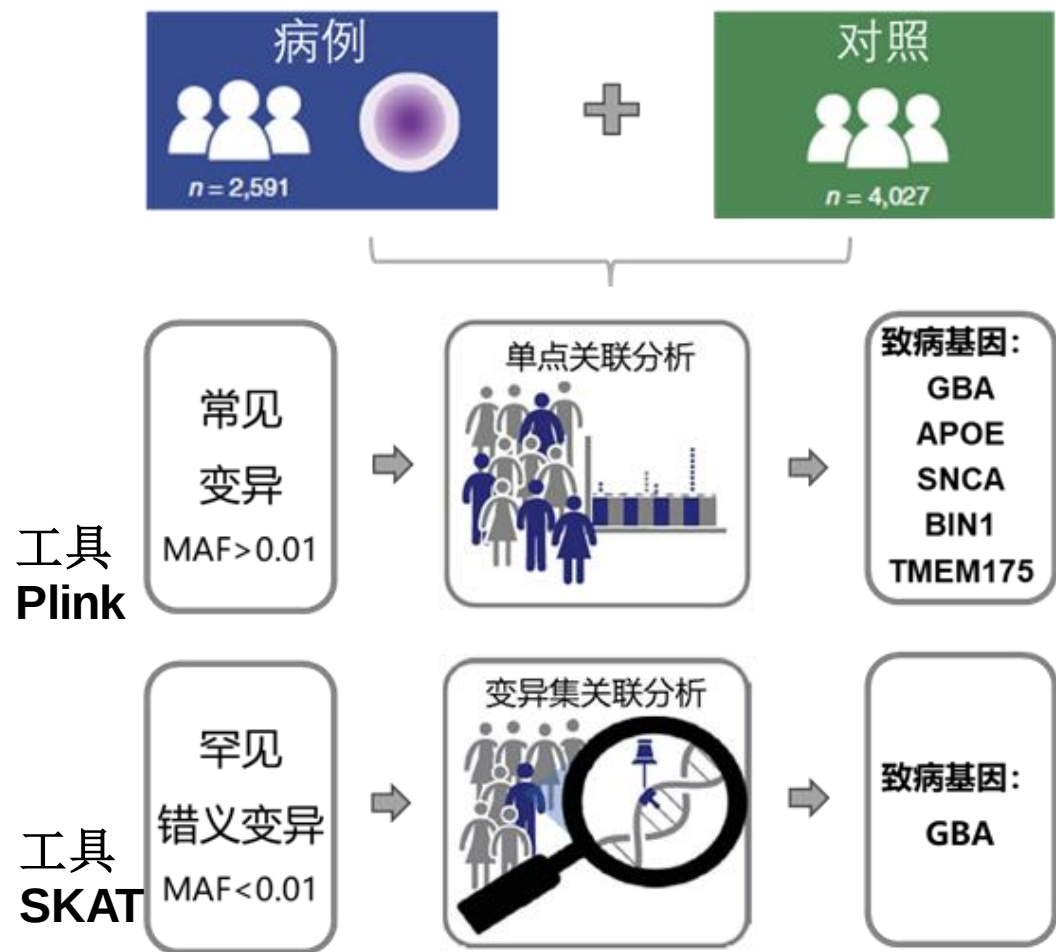
项目	发表时间	人数	人群	深度	变异类型	PubMed ID
中国NyuWa	2021/ 2022/ 2023年	3000-5000人	中国	30×	SNP/InDel/微卫星/转座子插入	34788621/ 35212372/ 37045857
中国ChinaMAP	2020年	10588人	中国	40×	SNP/InDel	32355288
中国WBBC计划	2022年	4480人	中国	14×	SNP/InDel	35618720
南京NSCLC队列	2022年	6004人	中国	30×	SNP/InDel	36113475
中国脑卒中队列	2023年	10241人	中国	30×	SNP/InDel	37479695
湘雅帕金森队列	2023年	4450人	中国	13×	SNP/InDel	36759515
云南吸烟队列	2020年	1058人	中国	34×	SNP/InDel	33033484
西藏藏族队列	2023年	1001人	中国藏族	12×	SNP/InDel/SV	37055782
江苏出生队列	2021年	1137人	中国	30×	SNP/InDel	34108666
广州出生队列	2021年	4053人	中国	4-7×	SNP/InDel	PPR504389

➤ 路易体痴呆症队列

- 基于加性效应模型对路易体痴呆症队列中的常见变异进行单点关联分析，确定了五个显著性基因GBA, APOE, SNCA, BIN1 和TMEM175
- 基于罕见变异位点的聚合检验方法，确定基因GBA在路易体痴呆症发病机制中具有突出作用

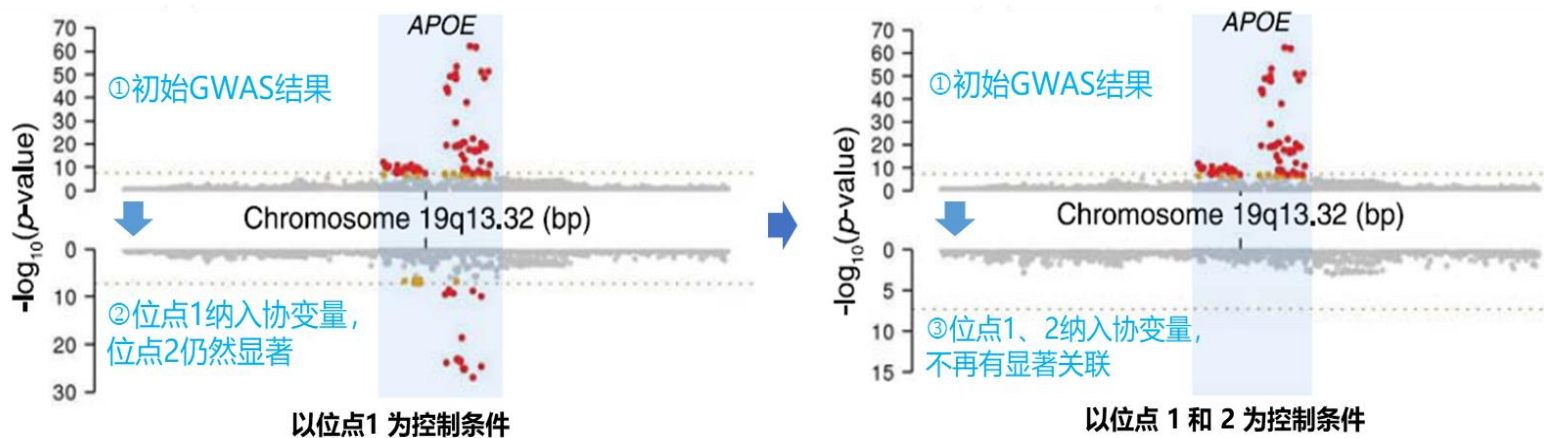
Note: 加性效应 即当两个或多个基因对于某一性状，或是单个基因的不同等位基因对于某一性状的整体作用，等于它们单独作用之和

$$V_G = V_A + V_B$$



路易体痴呆症队列
(Chia R., Nat Genet. 2021)

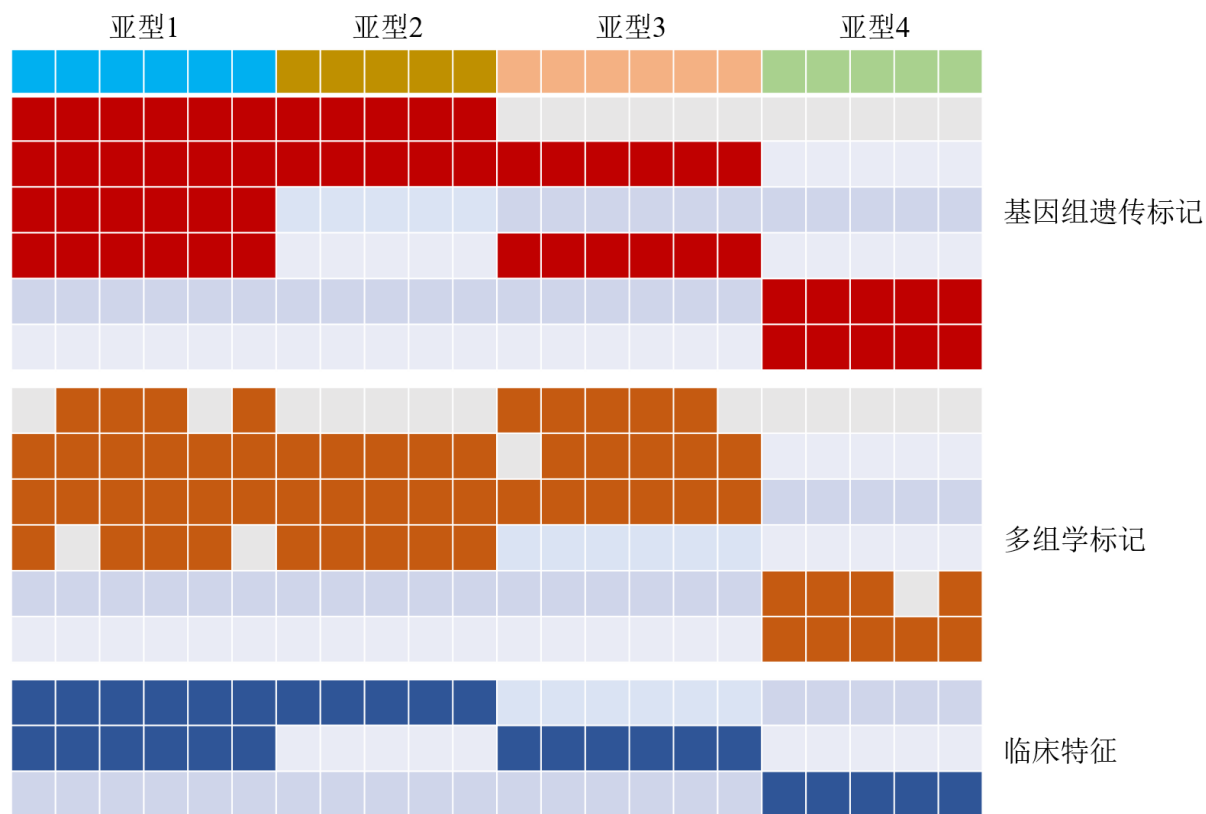
➤ 路易体痴呆症风险基因的因果定位



对路易体痴呆症队列GWAS结果，采用条件回归进行进一步精细定位，在 APOE 基因座检测到两个因果位点

- chr19:44906745G>A
- chr19:44909698A>C

➤ 分子差异

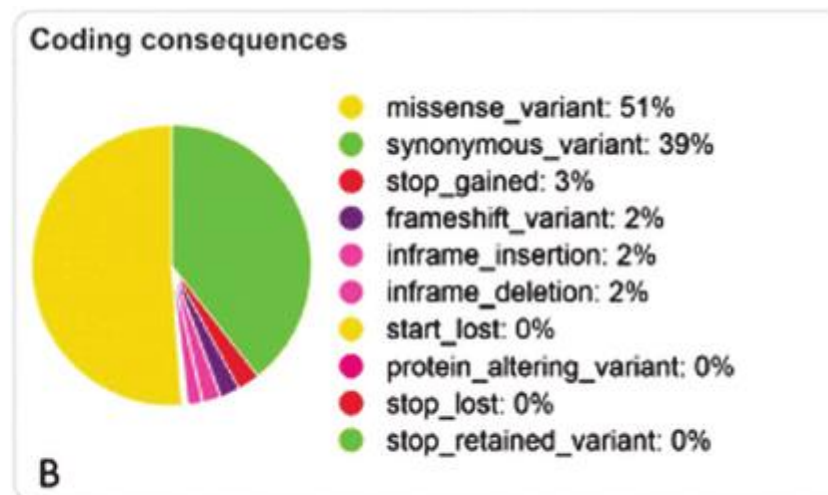
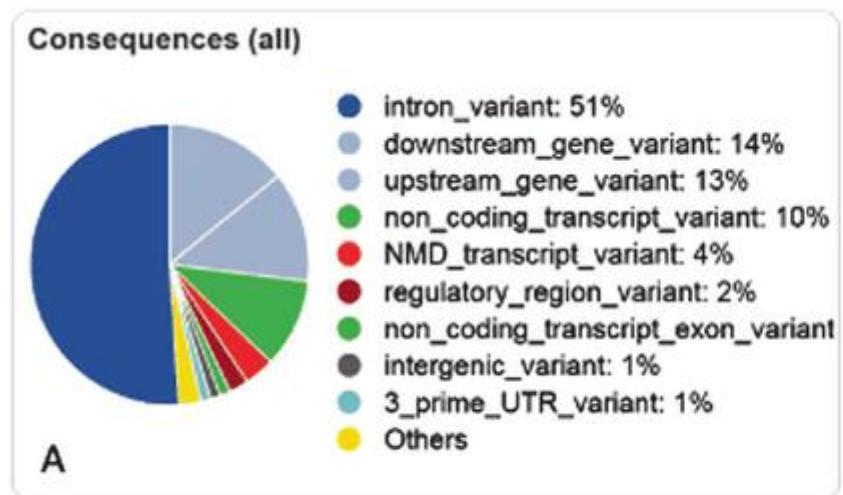
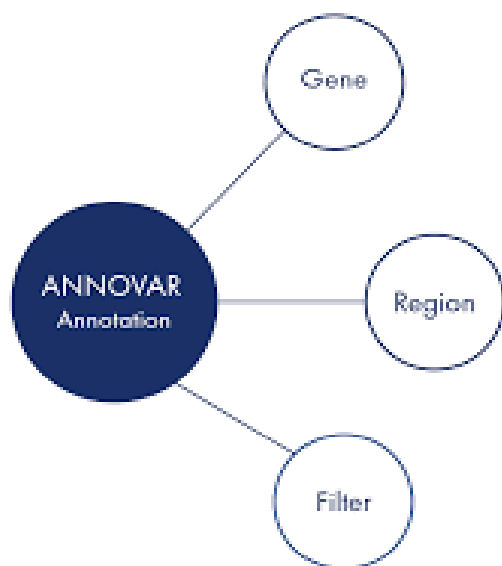


- 复杂疾病在分子角度具有显著的遗传异质性，根据多维分子特征可确定疾病的分子亚型
- 癌症基因组图谱联盟 结合乳腺癌样本的基因组体细胞突变数据、 CNV 芯片、DNA 甲基化、mRNA 芯片、microRNA 测序、反相蛋白微阵列 多个平台的相关数据，通过特征分解和无监督聚类方法确定了乳腺癌的四个分子亚型。

(Koboldt DC., Nature. 2012)

➤ 遗传变异功能注释

ANNOVAR、VEP、SNPEFF、GWAS Catalog



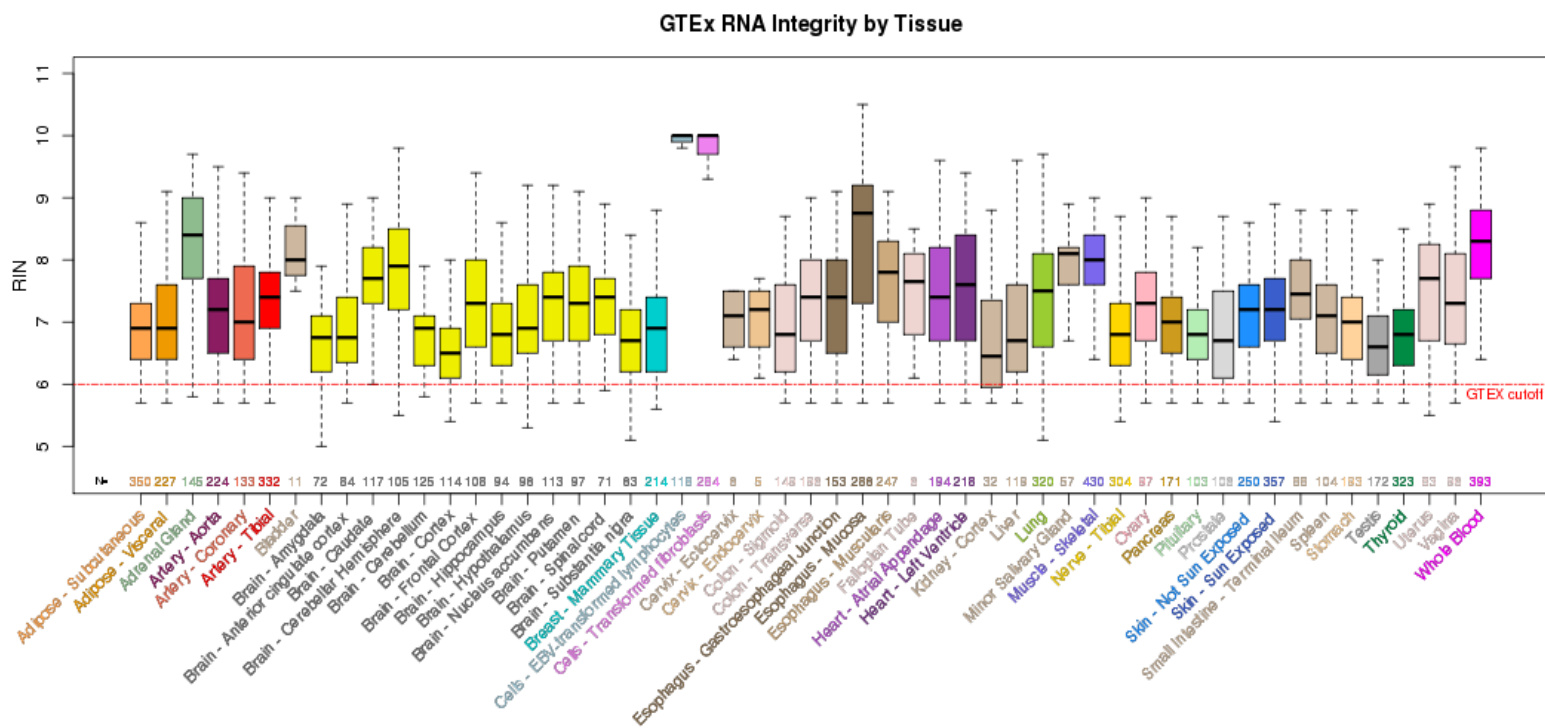
*NMD_transcript_variant: Nonsense mediated decay transcript variant.

VEP注释信息

➤ 多组学共定位

表观基因组、转录组、蛋白质组、代谢组等

人脑皮层发育调控图谱研究中，联合精神分裂症和自闭症谱系障碍的GWAS位点、GTEx xQTL信息，研究人员发掘出一系列潜在的神经系统发育障碍风险位点和相应基因调控机制。



(Walker et al., Cell. 2019)

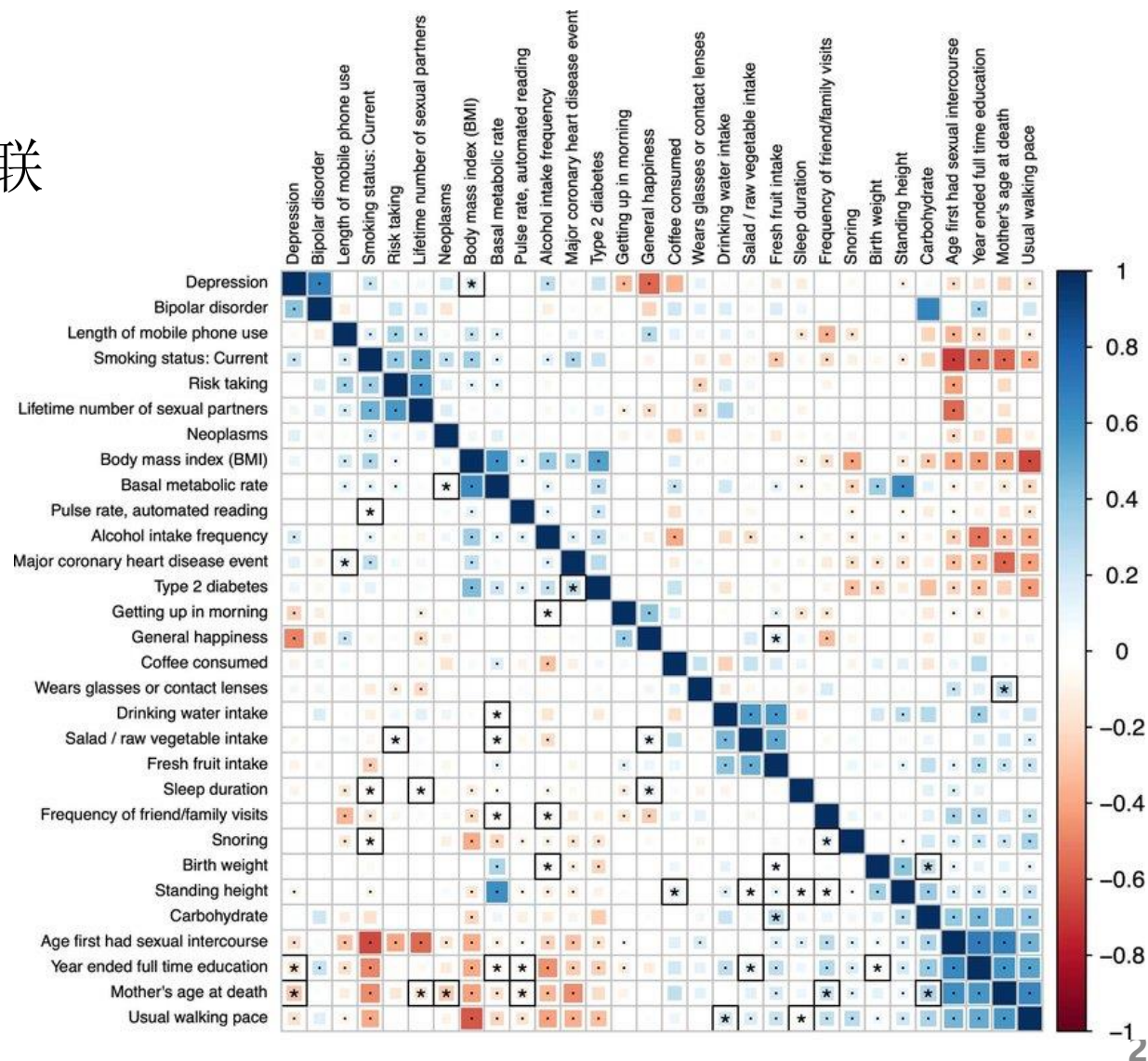
➤ 遗传相关性

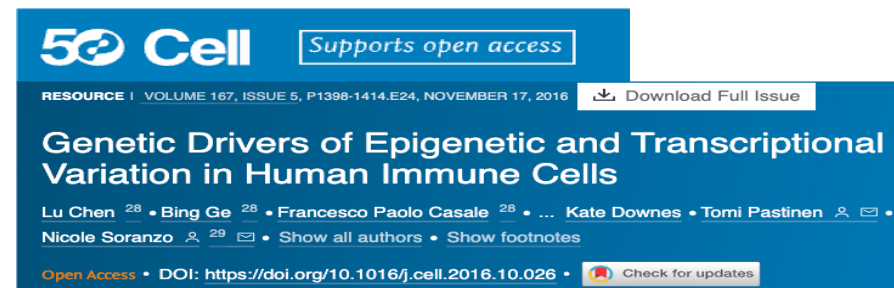
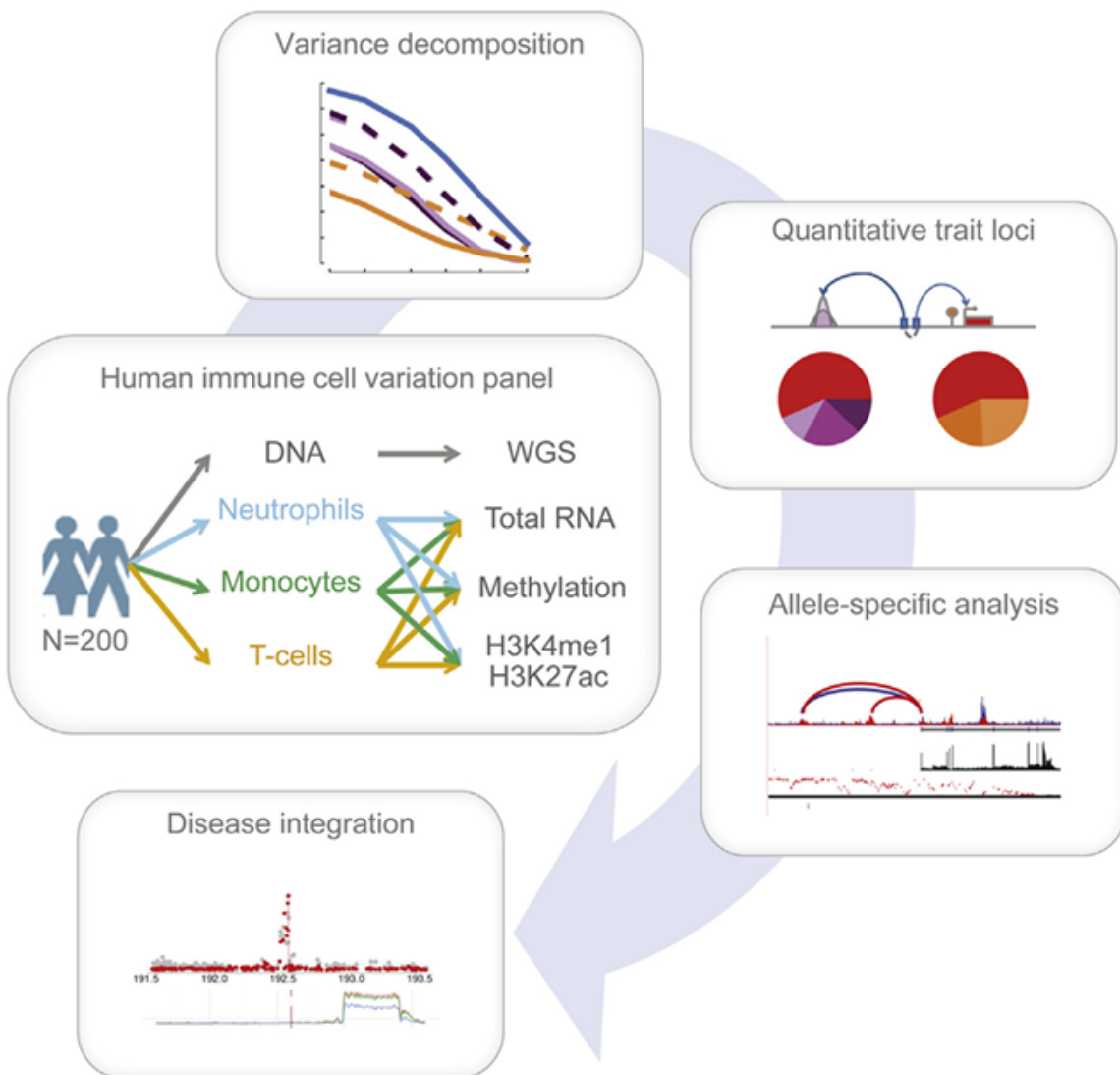
LDSC分析多种表型之间的遗传关联

对人群中24个性状进行遗传相关性分析，
研究结果揭示常见遗传疾病之间存在不同
程度的遗传关联：

- 冠状动脉疾病与成人身高之间显著负相关，身高每增加1个单位，冠状动脉疾病患病风险下降0.17个单位

(Bulik-Sullivan B., Nature Genetics. 2015)





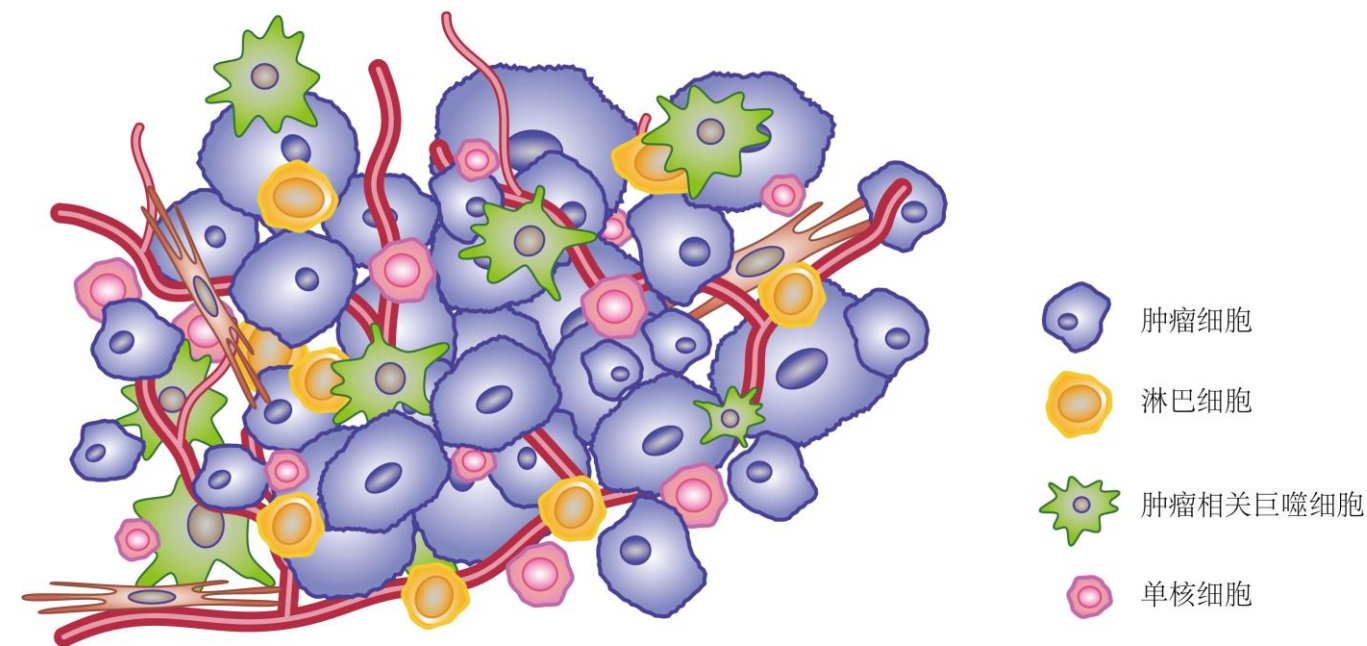
- 遗传因素在癌症、神经系统疾病等多种复杂疾病发生发展中起着关键作用，并对受其影响的患者及其近亲的治疗决策、降低风险的干预措施、疾病筛查和遗传检测产生重要影响。
- 将生物信息方法应用于遗传疾病研究，有助于系统高效地解析复杂疾病遗传发病机制，寻找疾病的原因和治疗的靶点，进而通过更精准和更科学的方法进行预防和干预，提供更为个体化的医疗服务，为改善公共卫生和个体健康水平提供科学支持。

第一节

- 精准医学
- 疾病的遗传关联研究
 - 复杂疾病与遗传因素
 - 疾病遗传关联分析
 - 疾病的遗传异质性识别
 - 疾病的遗传机制解析
- 疾病的免疫调控解析
 - 复杂疾病和免疫微环境
 - 免疫调控机制的计算解析
 - 基于免疫微环境实现免疫分型
 - 免疫微环境指导肿瘤精准治疗

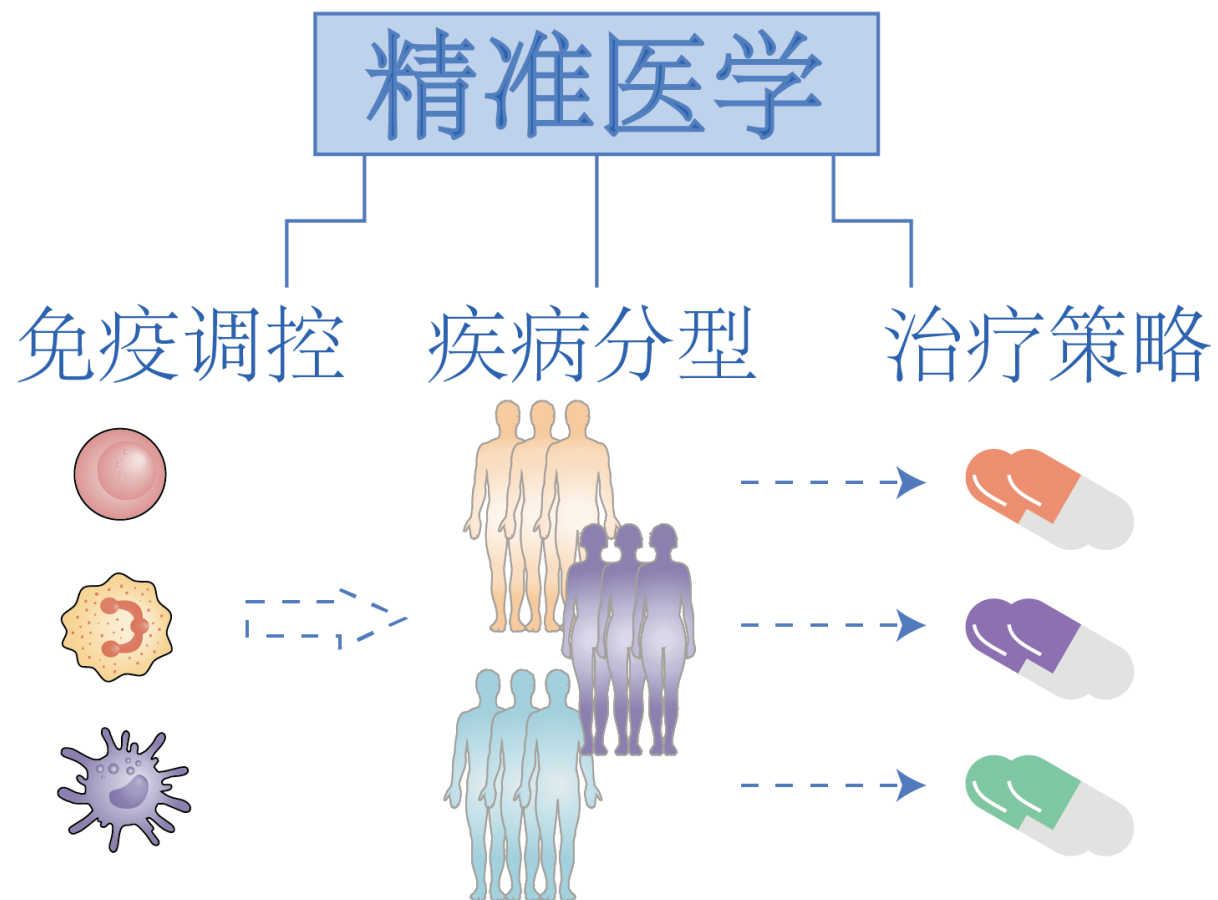
➤ 肿瘤微环境（tumor microenvironment, TME）

- 肿瘤细胞所处的组织环境
- 通常情况下，TME包含肿瘤细胞、多种免疫细胞、基质细胞以及新生血管等。



- 肿瘤微环境中肿瘤相关免疫细胞
- 从宿主免疫系统中逃逸
- 促进肿瘤生长和转移

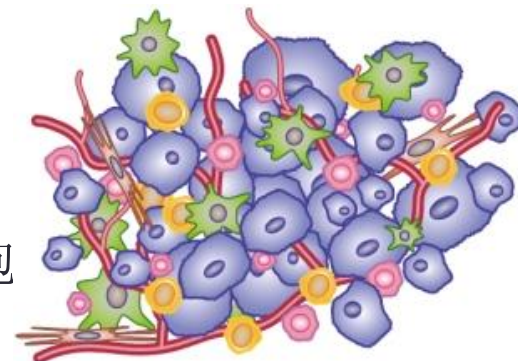
肿瘤微环境示意图



- 整合多组学数据探究复杂疾病发病过程中的受到的免疫调控作用，解析TME中免疫细胞的异质性组成、空间分布以及交互作用，有助于深入解析免疫功能的调控
- 为进一步识别癌症的免疫亚型，挖掘特异性的治疗靶点，推断癌症个性化医疗的发展提供支持。

➤ 肿瘤免疫微环境的计算预测

- 组织水平或单细胞水平的组学数据
- 通过差异分析、关联分析、细胞比例比较和细胞轨迹分析等，探究TME中免疫细胞的组成和浸润比例，识别出与疾病相关的细胞亚群，进一步揭示癌症免疫异质性。

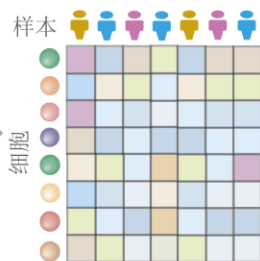


识别疾病相关细胞亚群流程

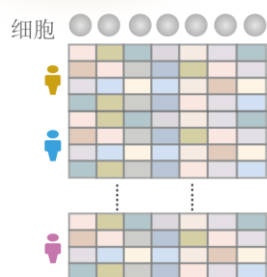
Bulk 数据



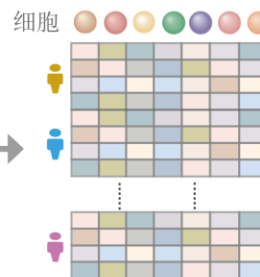
Deconvolution
Ecotyper
BayesPrism
.....



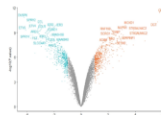
单细胞数据



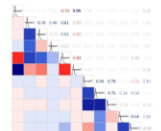
手动注释
自动注释



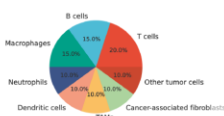
差异分析



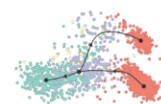
相关性分析



比例分析

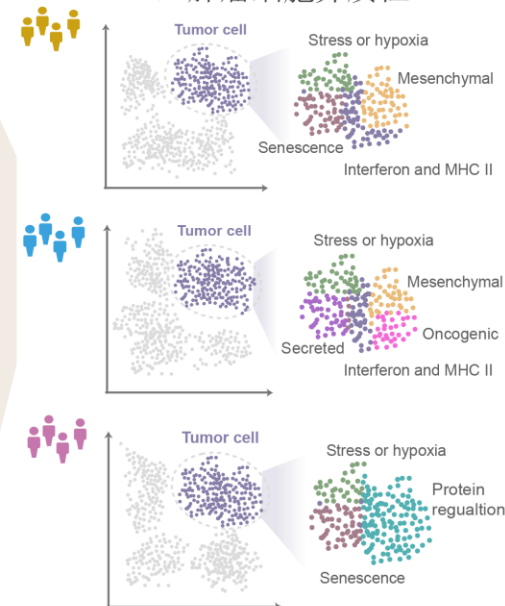


伪时序分析

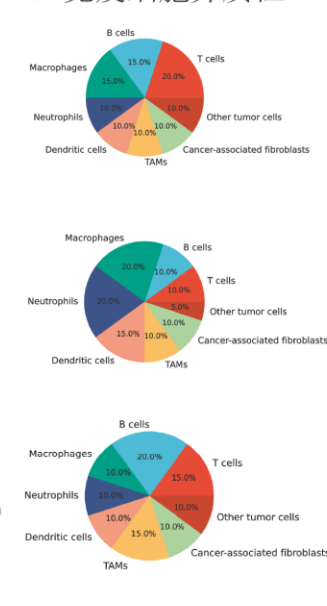


揭示疾病相关细胞亚群的异质性

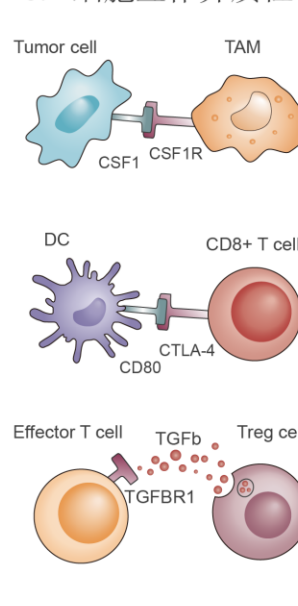
1. 肿瘤细胞异质性



2. 免疫细胞异质性



3. 细胞互作异质性



➤ 肿瘤免疫微环境的计算预测

- bulk RNA测序检测的是细胞混合物，其转录组数据是多种不同细胞混合的表达水平。

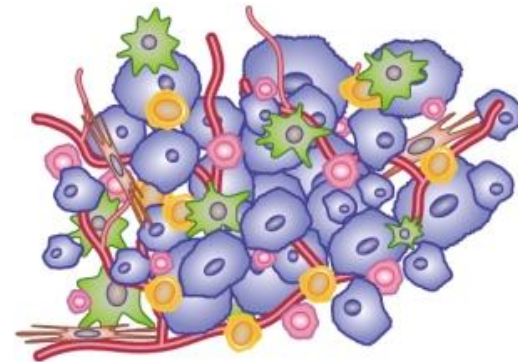


表 12-5 bulk 转录组数据预测细胞类型组成工具

方法	原理	作用
CIBERSORT	线性支持向量回归	22 种常见免疫细胞类型含量的估算；结合给定单细胞基因表达图谱进行细胞类型含量的估算。
BayesPrism	基于统计边缘化方法的 贝叶斯细胞比例重建	结合给定单细胞表达数据及细胞类型图谱进行细胞类型含量的估算。
EcoTyper	机器学习	细胞状态及多细胞生态系统的估算。

➤ CIBERSORT

- 从相关RNA转录本的表达中估计细胞浸润比例

	样本表达		细胞1	细胞2	细胞3	细胞4
基因1	10	=	2	3	5	3
基因2	20		5	1	2	5
基因3	15		7	8	7	8

组织测序数据的反卷积示意图

nature methods

Article | Published: 30 March 2015

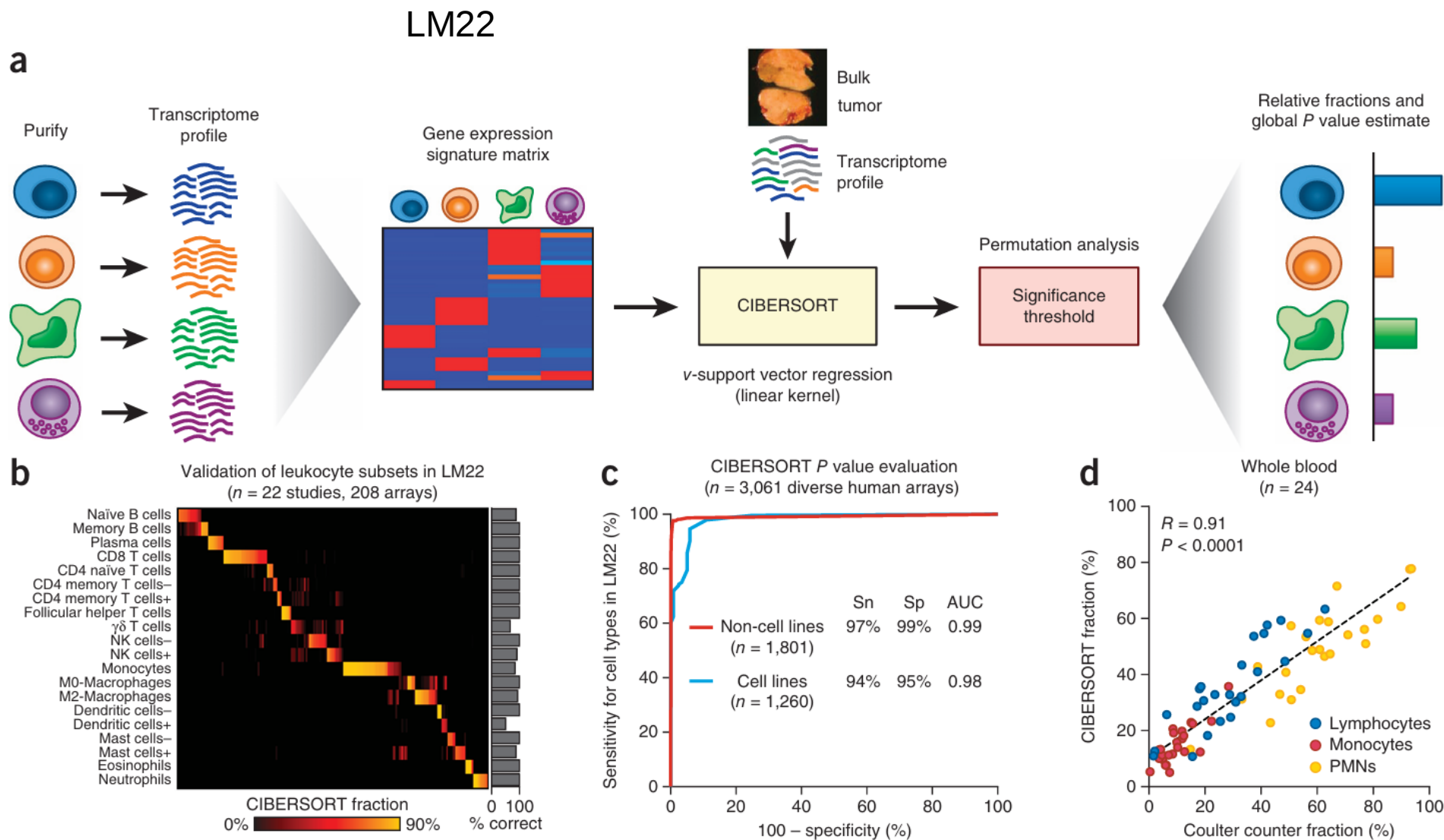
Robust enumeration of cell subsets from tissue expression profiles

Aaron M Newman, Chih Long Liu, Michael R Green, Andrew J Gentles, Weiguo Feng, Yue Xu, Chuong D Hoang, Maximilian Diehn & Ash A Alizadeh

	比例
基因1	a
基因2	b
基因3	c
基因4	d

leukocyte gene
signature matrix
(LM22) : 22个
功能性定义的人类
免疫亚群

<https://cibersortx.stanford.edu/>



CIBERSORTx

CIBERSORT

CIBERSORTx is an analytical tool from the **Alizadeh Lab** and **Newman Lab** to impute gene expression profiles and provide an estimation of the abundances of member cell types in a mixed cell population, using gene expression data.

EcoTyper, a machine learning framework for the identification of cell states and ecosystems from bulk, single-cell, and spatially-resolved expression data, is now available. EcoTyper extends CIBERSORTx for large-scale profiling of cellular ecosystems.

The CIBERSORT legacy website has been incorporated into the CIBERSORTx website under the menu item **CS Archives**.

nature
biotechnology

ARTICLES

<https://doi.org/10.1038/s41587-019-0114-2>

Determining cell type abundance and expression from bulk tissues with digital cytometry

Aaron M. Newman^{1,2*}, Chloé B. Steen^{3,4}, Chih Long Liu^{1,3}, Andrew J. Gentles^{2,3,5,6},
Aadel A. Chaudhuri^{7,8}, Florian Scherer^{3,9}, Michael S. Khodadoust³, Mohammad S. Esfahani^{3,5,8},
Bogdan A. Luca⁶, David Steiner³, Maximilian Diehn^{1,6,8} and Ash A. Alizadeh^{1,3,5,8,9*}

CIBERSORTx: Digital Cytometry

Select Analysis Module:

1. Create Signature Matrix

2. Impute Cell Fractions

3. Impute Cell Expression

Currently selected module: 2. Impute Cell Fractions

Select Analysis Mode:

Example

Custom

Configure Custom Input:

Signature matrix file*

LM22 (22 immune cell types)

More Information...

Mixture file*

Melanoma (Van Allen et al.) Mixture

More Information...

☐ Enable batch correction

☒ Disable quantile normalization (disabling is recommended for RNA-Seq data)

☐ Run in absolute mode

More Information...

Permutations for significance analysis

1000

More Information...

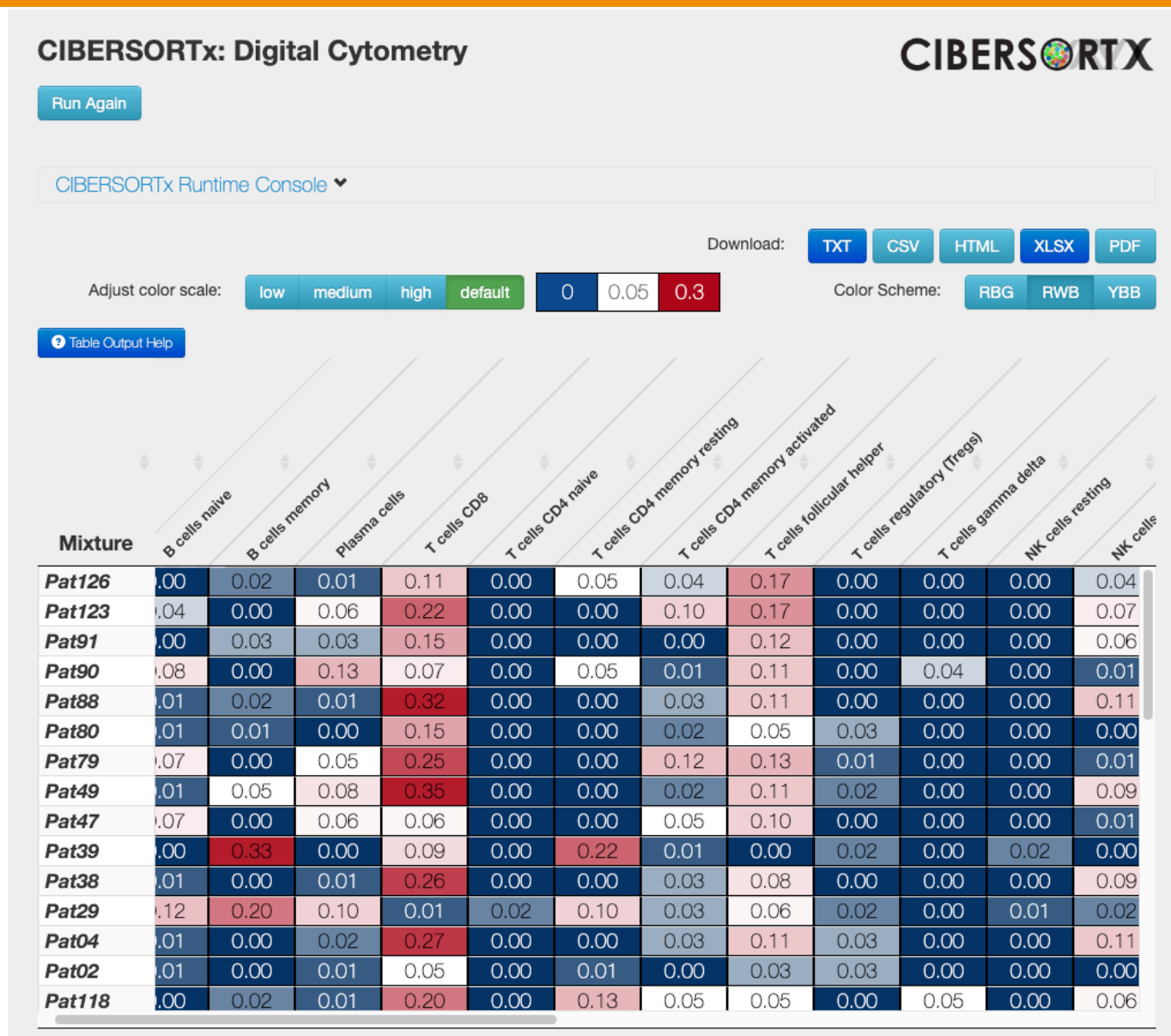
Run

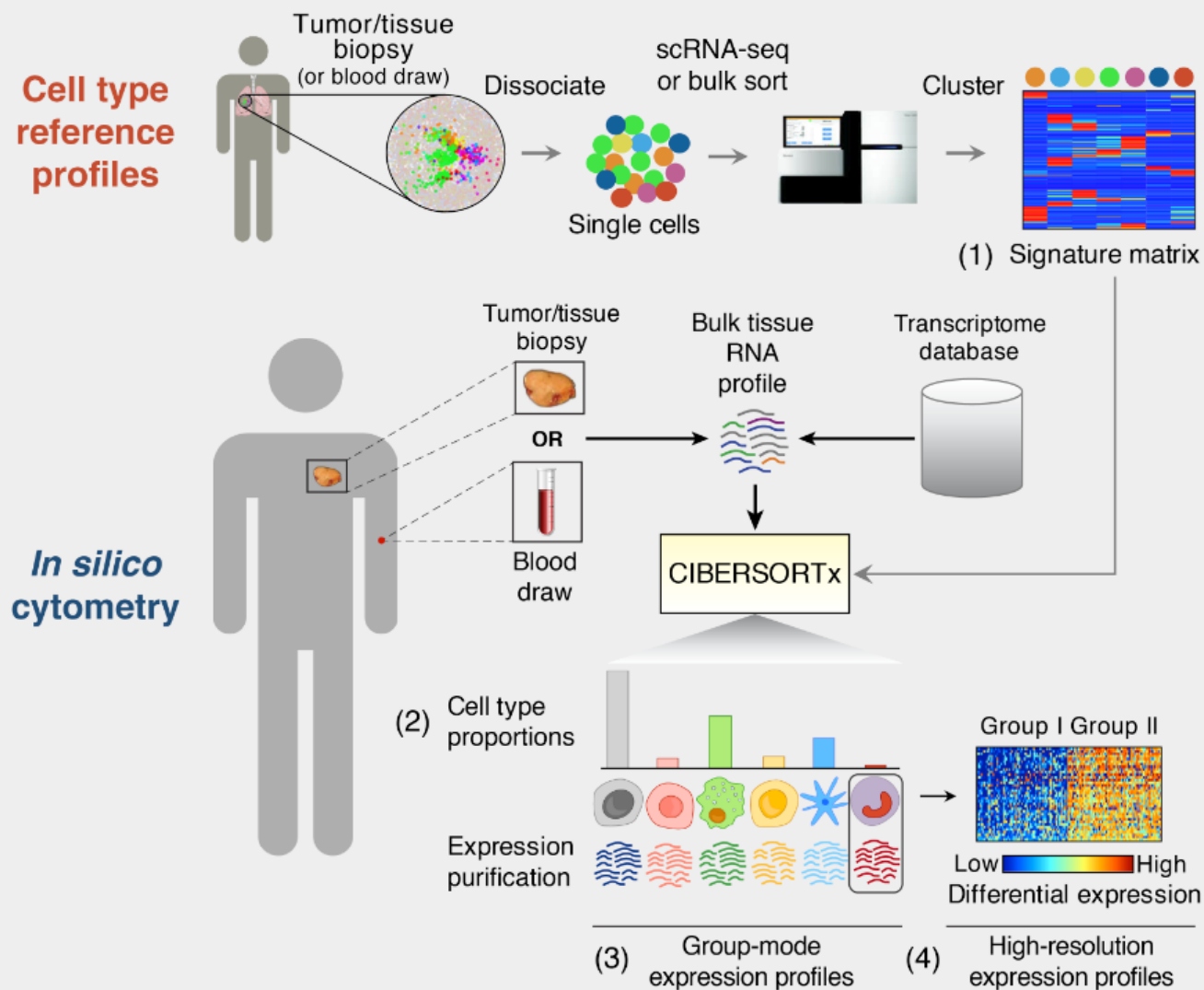
	1	2	3	4	5	6	7	8	9	10	11
1	GENE	B cells nai	B cells me	Plasma ce	T cells CD	T cells CD	T cells CD	T cells CD	T cells foli	T cells reg	T cells gai
2	ABCB4	555.71	10.74	7.23	4.31	4.61	7.41	8.04	6.47	7.83	9.31
3	ABCB9	15.6	22.09	653.39	24.22	35.67	30.05	38.46	17.6	46.07	19.72
4	ACAP1	215.31	321.62	38.62	1055.61	1790.1	922.15	340.88	1107.8	1995.48	280.08
5	ACHE	15.12	16.65	22.12	13.43	27.19	18.44	13.44	14.81	24.65	33.66
6	ACP5	605.9	1935.2	1120.1	306.31	744.66	557.82	248.55	711.95	958.92	493.97
7	ADAM28	1943.74	1148.12	324.78	22.69	40.06	21.23	15.21	24.28	74.77	24.37
8	ADAMDEC	371.03	318.48	127.97	44.62	80.96	47.43	30.55	68.74	156.71	231.83
9	ADAMTS3	146.2	106.05	74.34	42.39	79.24	57.39	53.41	35.41	76.28	65.27
10	ADRB2	486.34	510.08	289.8	899.65	77.15	595.21	142.83	97.17	217.06	2378.45
11	AIF1	24.07	20.32	21.97	742.82	718.61	35.17	39.51	19.08	29.37	1086.96
12	AIM2	327.96	4538.99	1429.67	294.48	57.08	250.3	3311.98	1115.68	315.65	407.59
13	ALOX15	12.23	13.75	9.07	9.77	18.56	12.01	5.24	13.15	22.38	14.45
14	ALOX5	1799.68	2425.68	190.33	78.68	74.93	132.01	59.09	35	143.01	441.56
15	AMPD1	263.75	217.76	3012.77	66.11	68.66	85.32	68.46	53.96	79.26	114.05
16	ANGPT4	23.12	21.31	230.68	44.68	22.06	44.48	99.59	101.45	29.47	29.26
17	ANKRD55	284.41	218.92	117.2	141.7	1853.34	436.42	79.77	1077.33	123.88	176.24
18	APOBEC3	102.43	89.6	99.59	64.36	107.24	52.74	57.78	71.38	36.39	723.14
19	APOBEC3	1283.45	1135.29	2034.56	2090.85	336.65	2340.67	1540.77	386.05	979.72	5415.05
20	APOL3	222.63	530.4	754.87	1209.12	950.79	1615.03	1119.49	351.13	993.46	1173.27
21	APOL6	17.31	11.61	7.83	8.45	11.96	8.44	57.52	7.18	13.37	9.72
22	AQP9	21.34	15.11	15.66	13.68	14.19	14.89	13.12	18.67	64.26	188.87
23	ARHGAP2	34.64	48.31	23.83	45.01	71.86	22.63	15.45	25.41	32.57	12.49
24	ARRB1	15.93	15.2	18.99	19.09	18.71	25.18	23.06	18.44	7.9	15.11
25	ASGR1	40.87	42.5	42.61	12.19	10.83	13.27	8.2	5.12	19.56	304.79
26	ASGR2	30.26	33.91	33.91	29.78	36.87	41.82	34.46	31.38	113.23	167.1
27	ATHL1	367.96	535.78	782.07	806.1	5333.94	794.95	231.18	1765.55	250.52	216.98
28	ATPRR4	123.61	86.81	83.73	69.79	59.89	51.01	77.35	29.05	58.63	106.47

signature
在细胞中的
表达谱

	TCGA-2V-A95S	TCGA-2Y-A9GS	TCGA-2Y-A9GT	TCGA-2Y-A9GU
UBE2Q2P2	-0.2717	-0.5324	-0.0775	0.0824
SSX9	-0.3675	-0.3675	-0.1933	-0.3675
CXORF67	-0.1215	-0.0387	-0.1215	-0.1215
EFCAB8	0.5946	-0.5944	-0.0489	-0.5944
EFCAB12	-0.5564	-0.5065	0.3281	-0.6531
A1BG	-0.5826	-0.573	0.9068	-0.9311
A1CF	-1.3638	-0.0581	-0.4438	-0.4812
RBFOX1	-0.1618	0.6202	-0.1618	-0.1618
GGACT	0.1972	-0.7036	-0.3268	8.4759
A2ML1	-0.3072	-0.3072	0.0145	0.2913
A2M	1.9399	-0.5751	-0.69	1.5222
A4GALT	-0.5411	-0.0379	-0.308	-0.5923
A4GNT	12.1408	-0.1617	-0.1862	-0.2066
AAAS	0.5402	-0.0366	-0.6154	-0.8451
AACSP1	-0.1131	-0.1131	-0.1131	-0.1131
AACS	0.4444	0.0069	-1.0287	0.9569
AADACL2	-0.2368	-0.2368	-0.2368	-0.2368
AADACL3	-0.2097	0.2509	-0.2097	-0.2097
AADACL4	-0.163	-0.163	-0.163	-0.163
AADAC	-0.9304	0.3545	-0.1318	0.157
AAGAB	1.3743	1.3264	-1.1741	-1.7773

基因在
样本中的
表达谱





Run CIBERSORTx »

➤ 识别肿瘤细胞

- 肿瘤细胞表面标记物的高表达可用于识别肿瘤细胞。当这些肿瘤细胞表面与其它细胞类群相比，如果某个细胞类群呈现出高的肿瘤标记基因表达，那么这个细胞亚群可以认为是潜在的肿瘤细胞亚群

识别肿瘤细胞的已知癌症标记基因

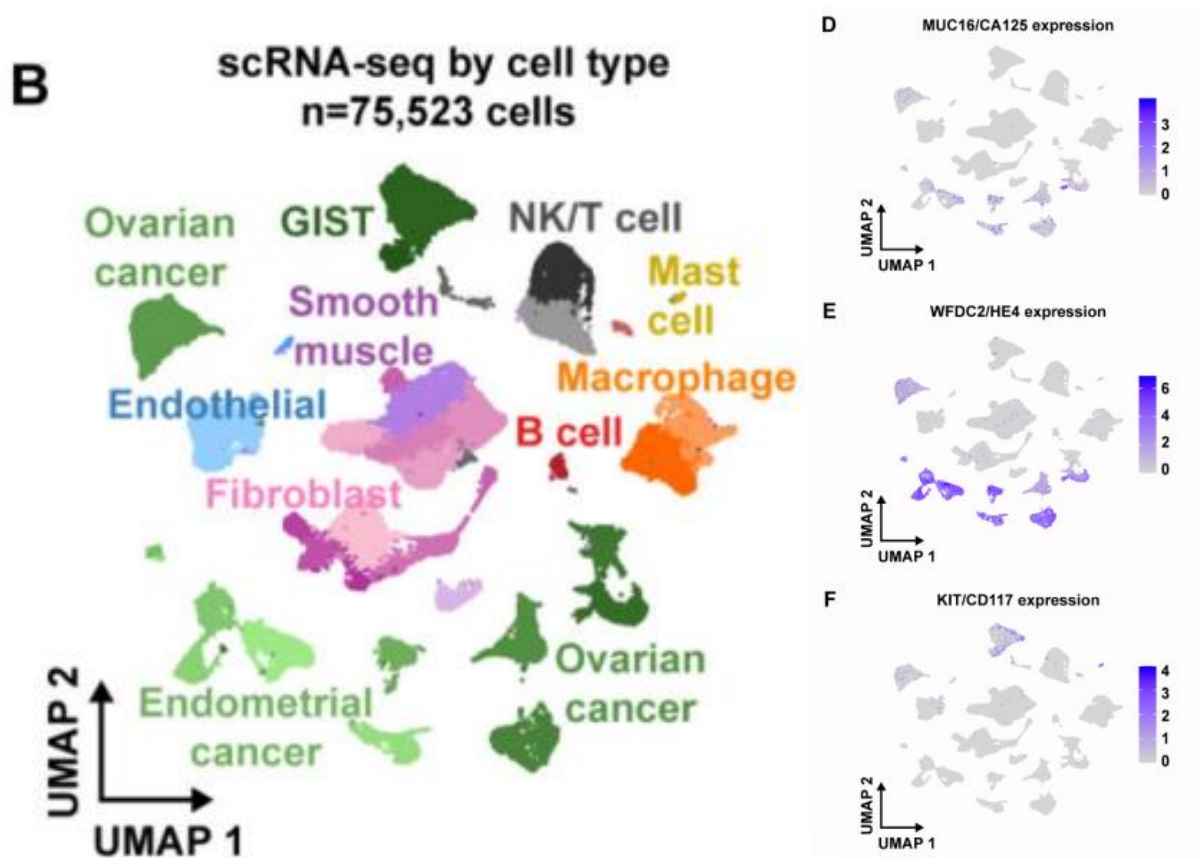
癌症类型	常用标记基因	参考文献
胶质母细胞瘤	SOD2、SCD5、SOX2、OLIG1	PMID: 36689332
葡萄膜黑色素瘤	MLANA、MITF、DCT	PMID: 31980621
卵巢癌	MUC16	PMID: 34739872
子宫内膜癌	WFDC2	PMID: 34739872
胃肠道间质瘤	KIT	PMID: 34739872

➤ 识别肿瘤细胞 - 妇科肿瘤

Molecular Cell

Resource

A multi-omic single-cell landscape of human gynecologic malignancies



- 研究对5名子宫内膜癌患者、5名卵巢癌患者以及一名由胃肠道间质癌转移至卵巢的患者进行了单细胞测序。
- 使用FDA批准的典型标记基因识别子宫内膜癌、卵巢癌以及胃肠道间质癌中的肿瘤细胞亚群
(MUC16/CA125: EC; WFDC2/HE4: OC; KIT: GIST)

➤ 识别疾病相关免疫细胞亚群

部分疾病相关免疫细胞及其功能亚群

细胞类型	细胞亚型	细胞功能	相关疾病
Cancer associated fibroblasts, CAF癌症相关成纤维细胞	myCAFs,肌成纤维细胞型癌症相关成纤维细胞	沉积、修改和降解细胞外基质，促进肿瘤生长和转移。高表达 α SMA，与平滑肌收缩、细胞外基质（ECM）组织和胶原蛋白形成相关。	胰腺癌、乳腺癌、肺癌
	iCAFs, 炎症型癌症相关成纤维细胞	表达多种炎症介质如IL-6, IL-11等，影响肿瘤免疫环境。	胰腺癌
	apCAFs, 抗原提呈相关成纤维细胞	表达MHC II相关基因，能够在体外部分激活CD4+ T细胞。	胰腺癌、乳腺癌
	immune-regulatory CAFs, 免疫调节型癌症相关成纤维细胞	参与调节免疫细胞的招募和活性，通过分泌细胞因子和化学因子来调控抗肿瘤免疫。	乳腺癌、胰腺癌

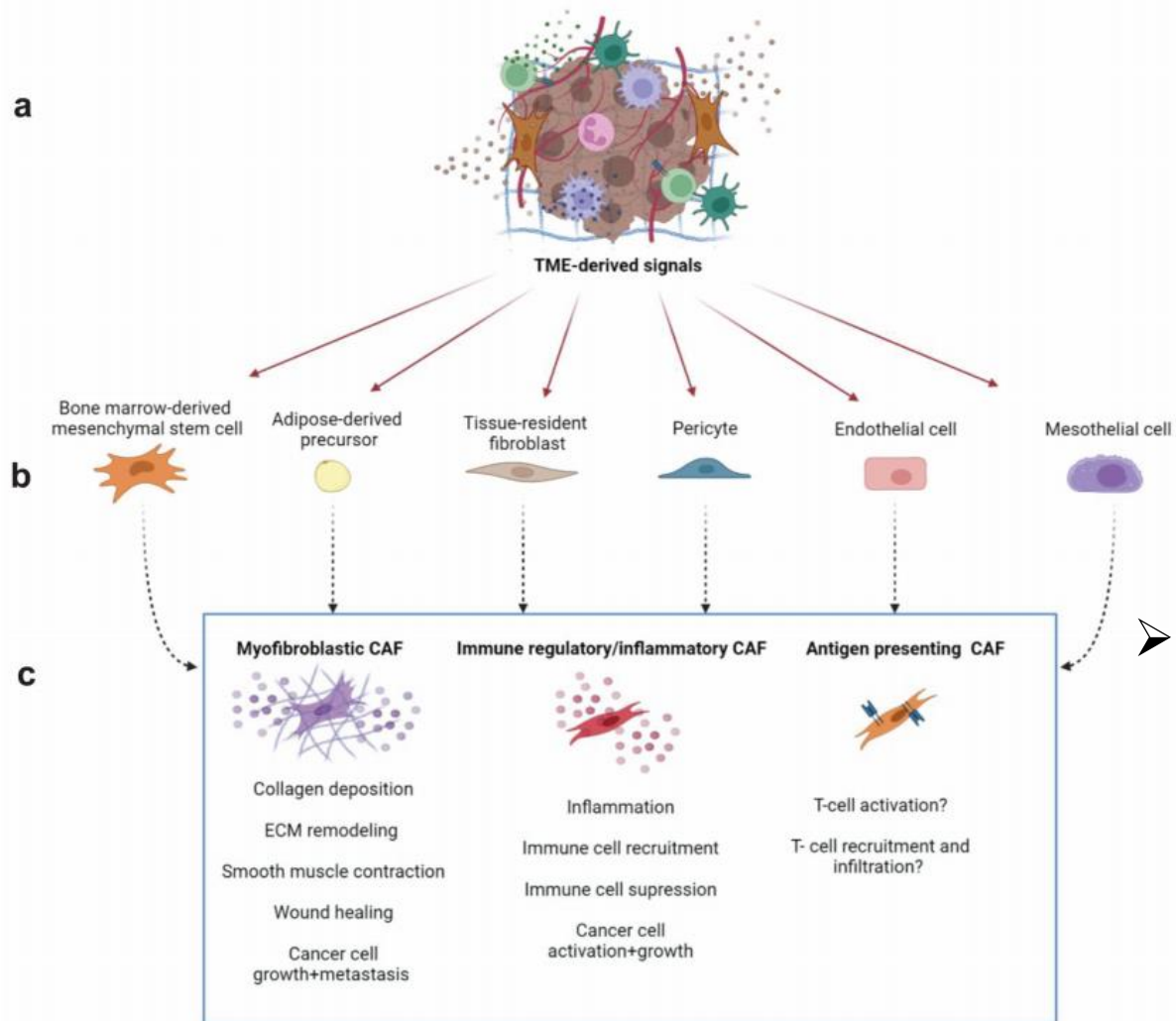
➤ 识别疾病相关细胞亚群 - 成纤维细胞

[nature](#) > [nature cancer](#) > [review articles](#) > [article](#)

Review Article | Published: 26 July 2022

Cancer-associated fibroblasts in the single-cell era

[Dor Lavie](#), [Aviad Ben-Shmuel](#), [Neta Erez](#) & [Ruth Scherz-Shouval](#)



➤ 在特定条件下，其他基质细胞类型可发生类别的转换：

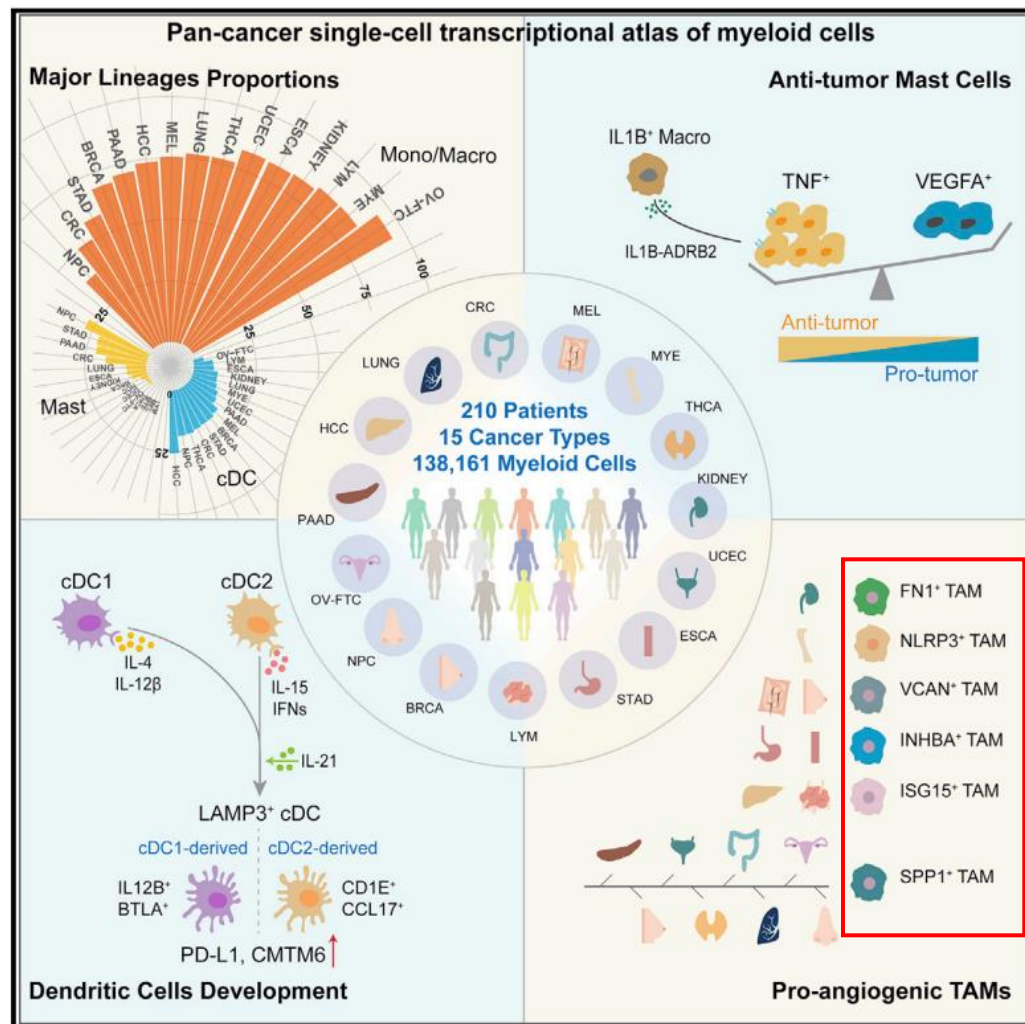
- TGF- β 1可诱导内皮细胞向成纤维表型转换；
- KLF4能够诱导血管周围细胞向细胞外基质表型的成纤维细胞转换。

➤ 识别疾病相关免疫细胞亚群

部分疾病相关免疫细胞及其功能亚群

细胞类型	细胞亚型	细胞功能	相关疾病
Macrophages, 巨噬细胞	M1 macrophages, 促炎相关巨噬细胞	产生促炎症因子（如TNF、IL-1和IL-6），具有强大的微生物杀伤能力和抗肿瘤功能。	乳腺癌、结直肠癌、自身免疫疾病
	M2 macrophages, 抗炎相关巨噬细胞	产生抗炎因子如IL-10和TGF-β帮助伤口愈合、组织重建和调节免疫反应。	过敏反应、哮喘、乳腺癌、肺癌
	TAMs, 肿瘤相关巨噬细胞	常常表现出M2型巨噬细胞特征，通过促进肿瘤微环境中的血管生成和细胞增殖等功能来支持肿瘤生长和转移。	乳腺癌、肺癌、黑色素瘤、胰腺癌
	Regulatory macrophages, 调节型巨噬细胞	在特定条件下产生大量的IL-10，参与免疫抑制过程，调节免疫系统反应，减少组织损伤。	炎症性疾病、器官移植、黑色素瘤

➤ 识别疾病相关细胞亚群 - 巨噬细胞



50 Cell

Supports open access

A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells

Sijin Cheng¹³ • Ziyi Li¹³ • Ranran Gao • ... Zhaode Bu • Jiafu Ji • Zemin Zhang¹⁴

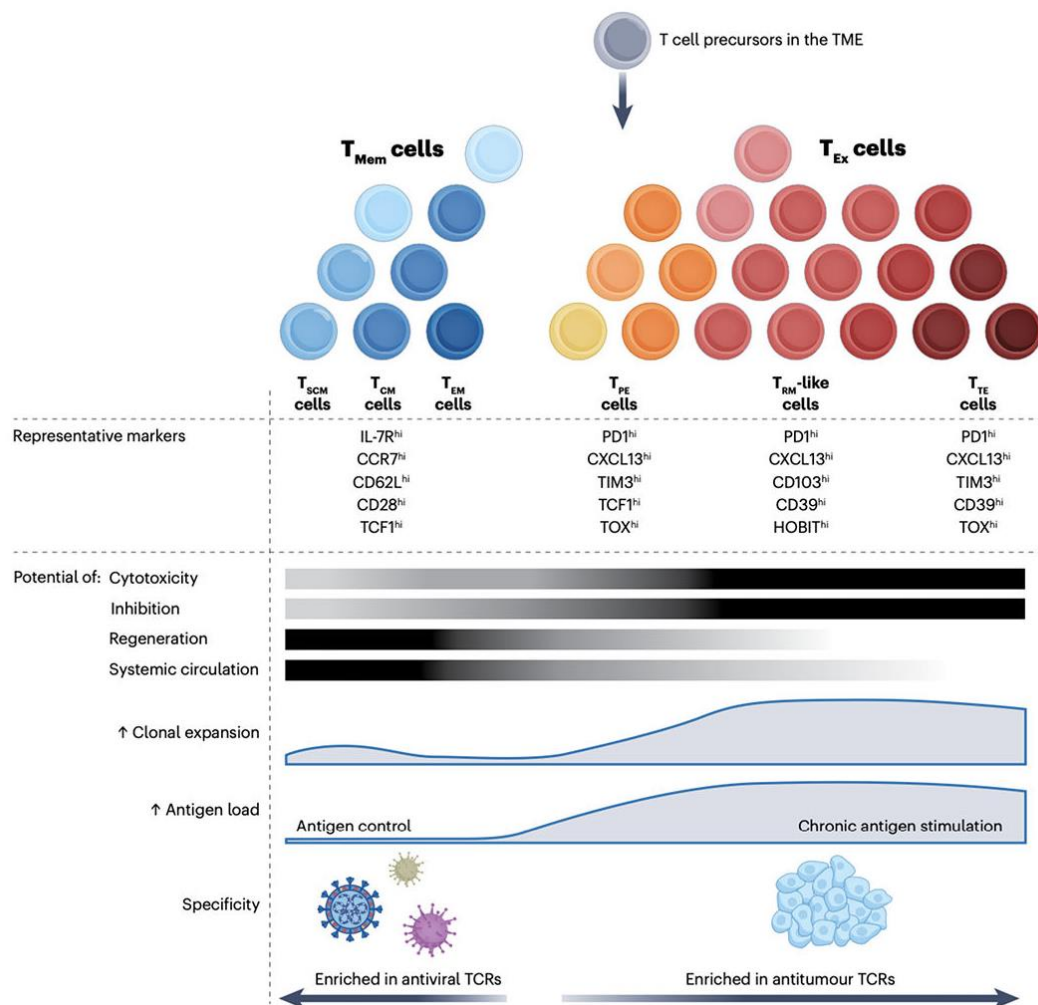
- 通过对15种癌症类型中210名患者的380个骨髓样本进行scRNA-seq数据分析，基于典型细胞标记识别出四种常见的主要髓系细胞类型(肥大细胞、pDCs、cDC、单核细胞/巨噬细胞)。
- 使用标记基因的表达模式来区分不同的巨噬细胞亚群。在TAMs中，特别是SPP1⁺ TAMs，在多种癌症类型中被鉴定为与血管生成相关。
- TAMs在肿瘤微环境中可以表现出M1（抗肿瘤）和M2（促肿瘤）表型。ISG15⁺ TAMs表现出M1表型，表达多种干扰素诱导基因，具有促炎症和抗肿瘤的特征。而C1QC⁺ TAMs和其他一些TAM亚群则表现出M2表型，与肿瘤的免疫抑制和进展相关。

➤ 识别疾病相关免疫细胞亚群

部分疾病相关免疫细胞及其功能亚群

细胞类型	细胞亚型	细胞功能	相关疾病
T cell, T细胞	Regulatory T cell, 调节性T细胞	专门用于免疫抑制的T细胞亚群，在维持免疫自我耐受和体内平衡方面起着关键作用	肺癌、黑色素瘤、自身免疫疾病
	Exhausted T cells, 耗竭T细胞	随着疾病的进展以分层方式丧失效应功能。	黑色素瘤、非小细胞肺癌
	cytotoxic T cell, 毒性T细胞	通过分泌炎性细胞因子以及穿孔素和颗粒酶等细胞裂解分子来杀死病原体感染的细胞或恶性细胞	胃癌、乳腺癌
	memory T cell, 记忆性T细胞	在再次感染时迅速响应，提供长期免疫保护。	乳腺癌、胃癌、头颈部鳞状细胞癌

➤ 识别疾病相关细胞亚群 - T细胞



[nature](#) > [nature reviews cancer](#) > [review articles](#) > [article](#)

Review Article | Published: 12 April 2023

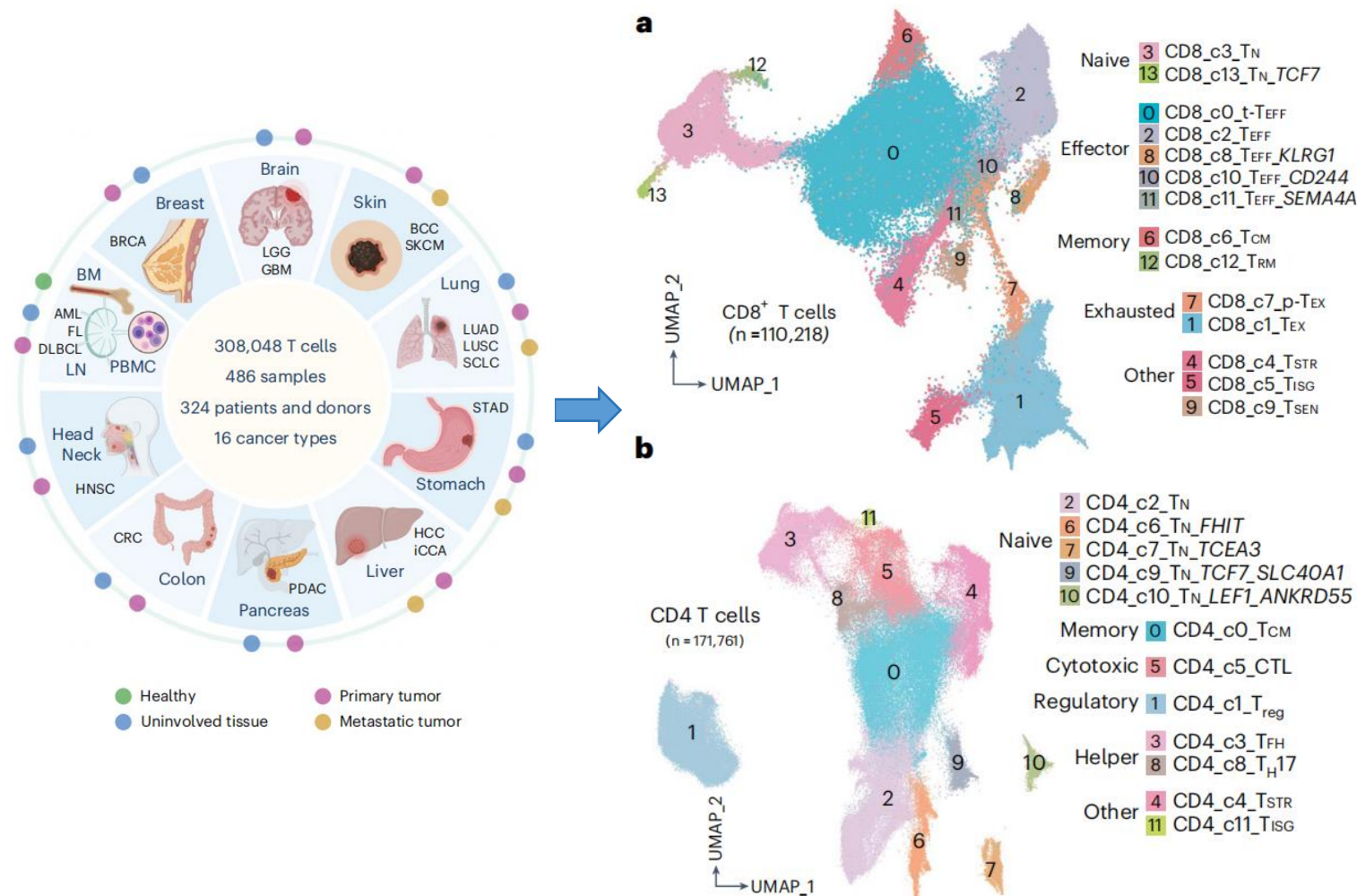
Dynamics and specificities of T cells in cancer immunotherapy

[Giacomo Oliveira](#) & [Catherine J. Wu](#)

➤ 肿瘤浸润T淋巴细胞可以简化的分类为T记忆细胞与T耗尽细胞:

- T记忆细胞细胞通常在生理免疫反应期间产生，有助于建立一个持久的抗原特异性 T 细胞库。
- 在分化后期的T细胞通常表达PD1、CTLA4、LAG3、TIM3等与T细胞功能衰竭的基因，这是由于 TME 内的持续性肿瘤抗原刺激提供的长期抗原刺激，细胞毒性功能受到不同程度的损害。

➤ 识别疾病相关细胞亚群 - T细胞



nature medicine

Resource | Published: 29 May 2023

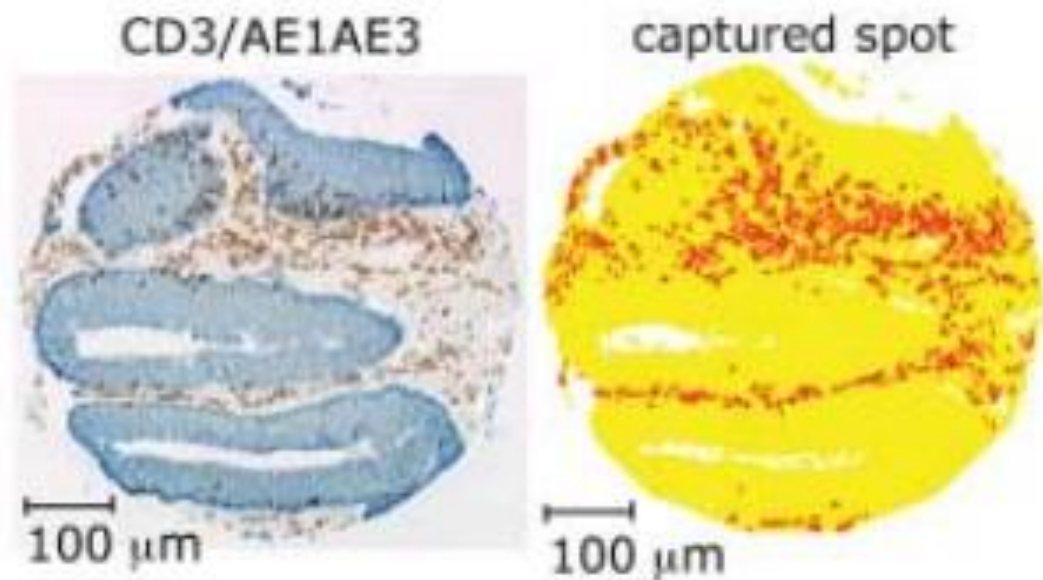
Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance

Yanshuo Chu, Enyu Dai, Yating Li, Guangchun Han, Guangsheng Pei, Davis R. Ingram, Krupa Thakkar, Jiang-Jiang Qin, Minghao Dang, Xiuning Le, Can Hu, Qing Deng, Ansam Sinjab, Pravesh Gupta, Ruiping Wang, Dapeng Hao, Fuduan Peng, Xinmiao Yan, Yunhe Liu, Shumei Song, Shaojun Zhang, John V. Heymach, Alexandre Reuben, Yasir Y. Elamin, ... Linghua Wang + Show authors

- 该研究构建了基于16种癌症类型，包含27个scRNA-seq数据集的308,048个肿瘤浸润T细胞的单细胞图谱，揭示了以前未描述的T细胞状态、滤泡辅助T细胞、调节性和增殖性T细胞的异质亚群，并发现了一种独特的应激反应状态—TSTR，以热休克基因表达为特征。
- 共表征了32种T细胞状态，并进一步确定了CD4调节亚群中的七个亚群、CD4滤泡辅助性T细胞群中的五个亚群以及增殖T细胞中的8个状态。这些发现强调了肿瘤微环境中T细胞状态的广泛异质性，以及进一步了解这些状态如何促进疾病进展和免疫治疗反应的必要性。

➤ 免疫分型

- 肿瘤免疫疗法已经彻底改变了癌症治疗的范式，免疫治疗效果与患者的免疫微环境状态密切相关。



- 样本的**T细胞密度**是免疫细胞浸润程度的直接证据，利用免疫染色技术对T淋巴细胞进行定量，进而利用T细胞百分比表示肿瘤样本的T细胞密度：
 - “热”肿瘤：表现出T淋巴细胞高浸润特点
 - “冷”肿瘤：表现出T淋巴细胞的缺失或排斥

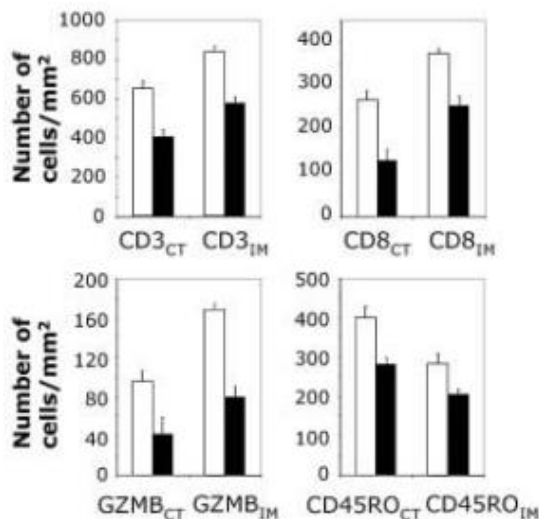
Science

Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome

JÉRÔME GALON, ANNE COSTES, FATIMA SANCHEZ-CABO, AMOS KIRILOVSKY, BERNHARD MLECNIK, CHRISTINE LAGORCE-PAGÈS, MARIE TOSOLINI, MATTHIEU CAMUS,

ANNE BERGER, [...] AND FRANCK PAGÈS +6 authors [Authors Info & Affiliations](#)

SCIENCE • 29 Sep 2006 • Vol 313, Issue 5795 • pp. 1960-1964 • DOI: 10.1126/science.1129139



Science

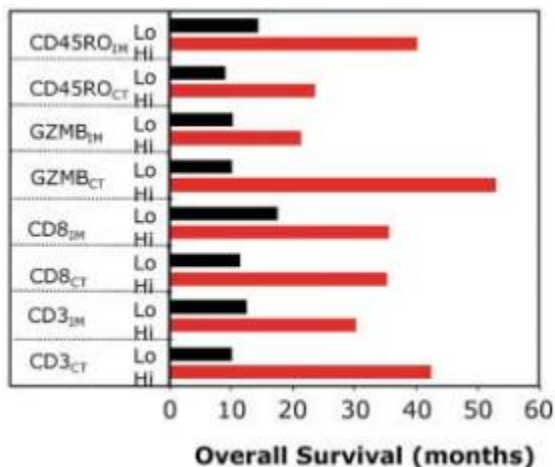
Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome

JÉRÔME GALON, ANNE COSTES, FATIMA SANCHEZ-CABO, AMOS KIRILOVSKY, BERNHARD MLECNIK, CHRISTINE LAGORCE-PAGÈS, MARIE TOSOLINI, MATTHIEU CAMUS,

ANNE BERGER, [...] AND FRANCK PAGÈS +6 authors [Authors Info & Affiliations](#)

SCIENCE • 29 Sep 2006 • Vol 313, Issue 5795 • pp. 1960-1964 • DOI: 10.1126/science.1129139

- 无论是在肿瘤区域(CT)还是在肿瘤侵袭边缘(IM), 未复发患者肿瘤样本中(white bars)各种免疫细胞密度(CD3, CD8, GZMB和CD45RO)均高于复发患者(black bars)。



- 生存分析显示, 免疫细胞浸润水平较高的患者拥有更好的生存预后。

- 除了使用免疫细胞密度表征免疫分型外，bulk转录组数据以及相关免疫分型方法的应用进一步揭示了样本的免疫特点和多样的免疫分型。

Cell

Volume 186, Issue 18, 31 August 2023, Pages 3945-3967.e26

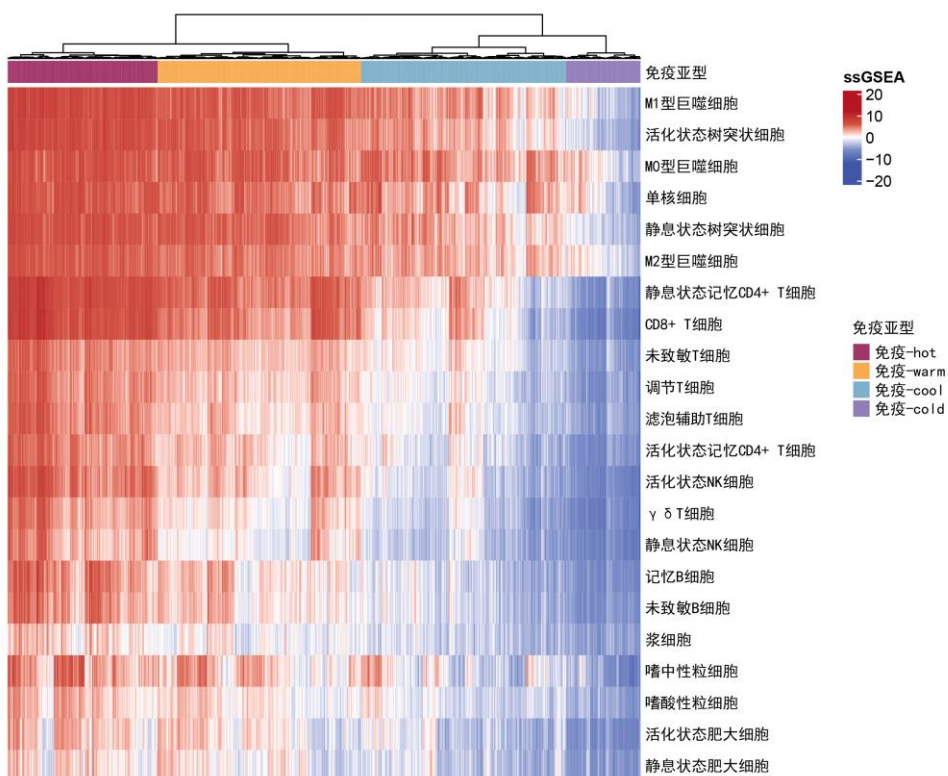
Article

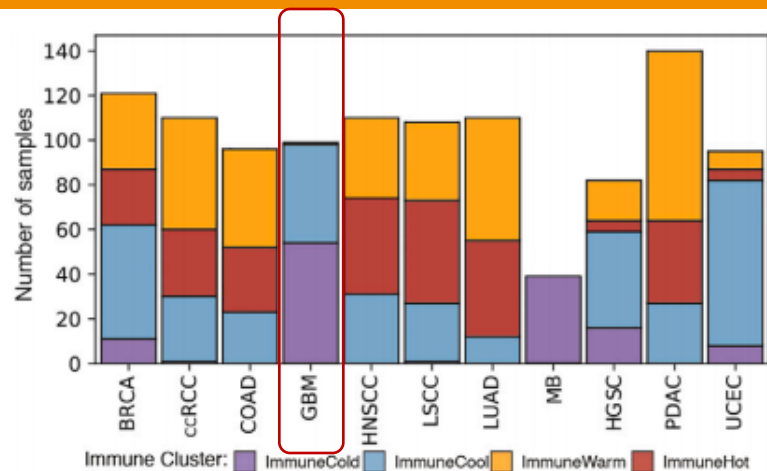
Pan-cancer analysis of post-translational modifications reveals shared patterns of protein regulation

Yifat Geffen^{1,2,15}, Shankara Anand^{1,15}, Yo Akiyama^{1,15}, Tomer M. Yaron^{3,15}, Yizhe Song^{4,15}, Jared L. Johnson^{3,15}, Akshay Govindan^{4,15}, Özgün Babur^{5,15}, Yize Li^{4,15}, Emily Huntsman³, Liang-Bao Wang⁴, Chet Birger¹, David I. Heiman¹, Qing Zhang¹, Mendy Miller¹, Yosef E. Maruvka⁶, Nicholas J. Haradhvala¹, Anna Calinawan⁷, Saveliy Belkin¹, Alexander Kerelsky³,...Daniel Cui Zhou

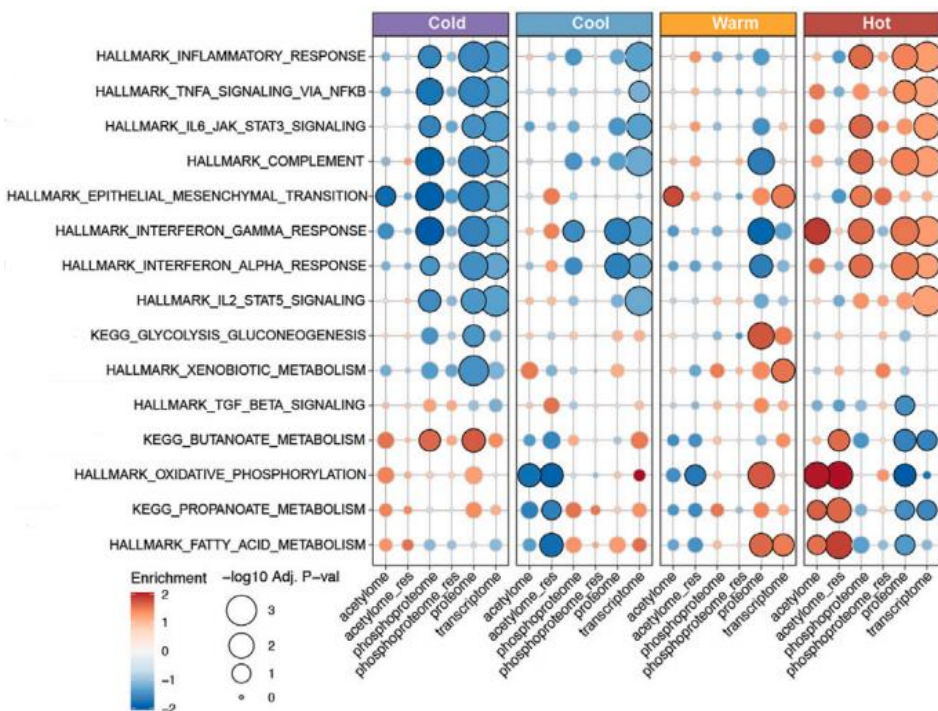


- 利用CIBERSORT及其精选基因集，计算bulk转录组数据的22种免疫细胞浸润比例，基于细胞浸润比例并使用无监督聚类方法对样本进行免疫分型：
- immune-hot: 表现出免疫相关通路的显著富
 - immune-warm: 同一结构域的磷酸化位点显著增加
 - immune-cool: 多种代谢途径的乙酰化水平显著差异，而这些途径的乙酰化对免疫抑制起着重要作用
 - immune-cold: 表现出耗竭相关免疫通路的显著富集

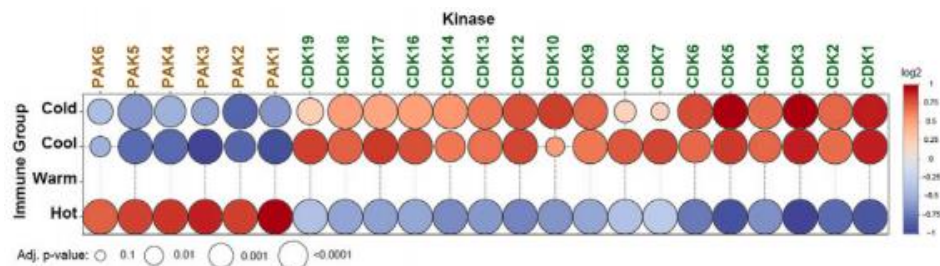




- 通过对这些免疫亚型进一步研究可以发现，脑部肿瘤全部由immune-cold肿瘤组成，造成这一结果的可能原因是血脑屏障，而其他肿瘤类型中免疫亚型的分布比较多样。

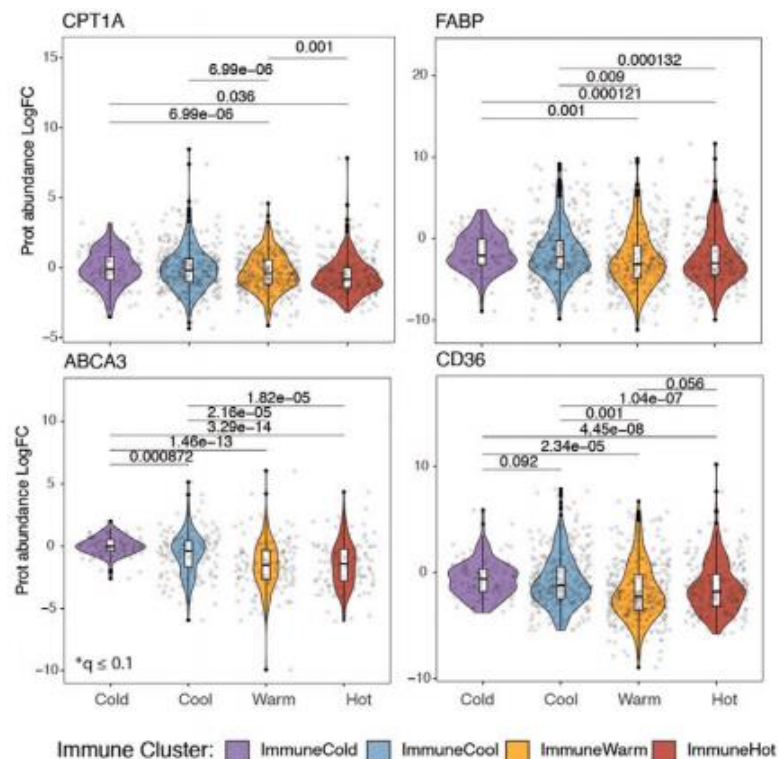


- immune-hot肿瘤表现出免疫功能通路的增加
- immune-cold肿瘤的免疫功能显著较低；
- immune-cool肿瘤中部分通路的乙酰化水平较低，表明低乙酰化水平可能有助于该亚型抵消免疫通路的低激活；
- immune-warm肿瘤表现出通路激活的多样性，提示该类肿瘤有较高的可塑性。



- 蛋白激酶分析显示，周期蛋白依赖性激酶CDK家族在immune-cold和immune-cool肿瘤中高富集，提示该类肿瘤的高增殖；

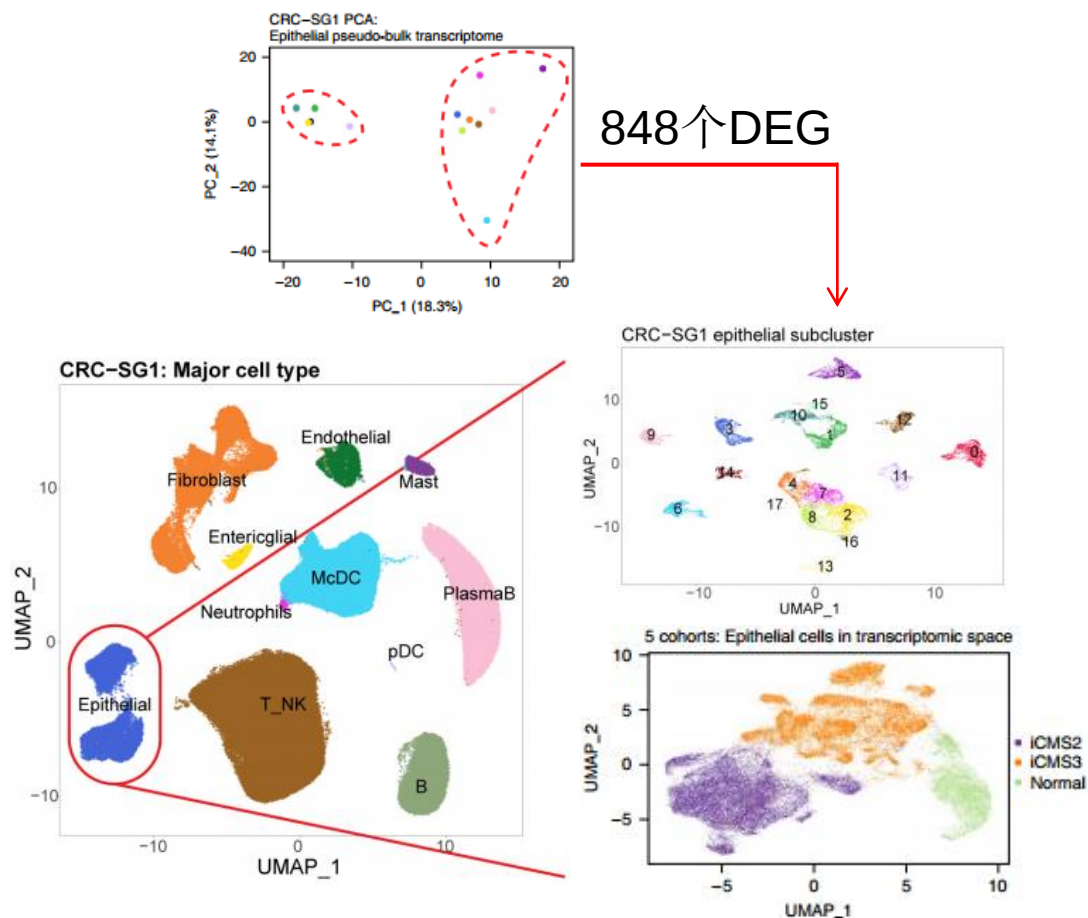
- 而immune-hot肿瘤表达较高水平的丝氨酸/苏氨酸蛋白激酶。



- 此外，immune-cold肿瘤具有较高表达水平的细胞内转运体(CPT1A)和细胞表面转运体(FABP4、ABCA和CD36)，提示immune-cold肿瘤摄取长链脂肪酸(质膜的组成部分和有效的燃料来源)的增加。

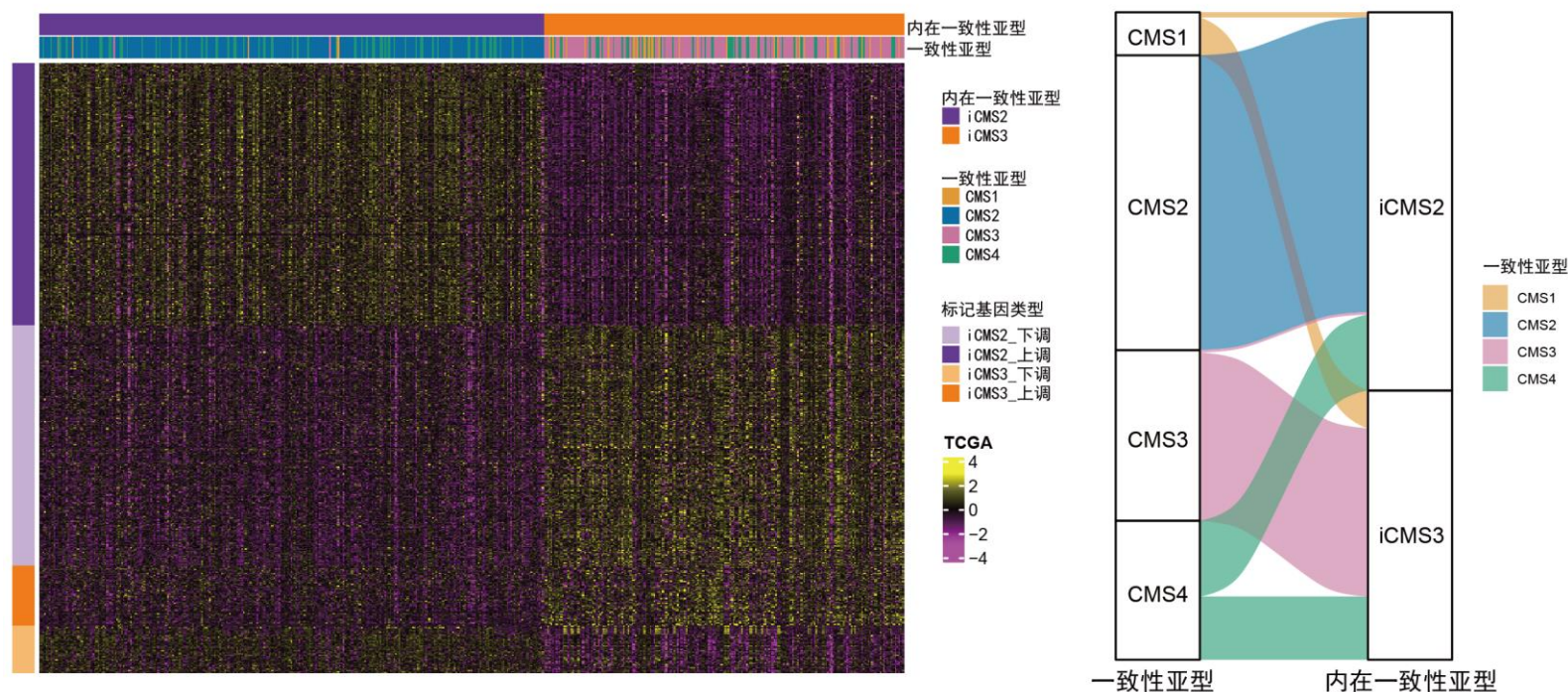
➤ 基于单细胞转录组数据

- bulk转录组测量异质组织中的总基因表达，因此组成细胞的转录组、它们的比例和肿瘤微环境的相互作用是模糊的。单细胞RNA测序(scRNA-seq)在细胞分辨率上表征转录组，能够识别细胞类型及其表达谱，能够更加细致的刻画肿瘤微环境。

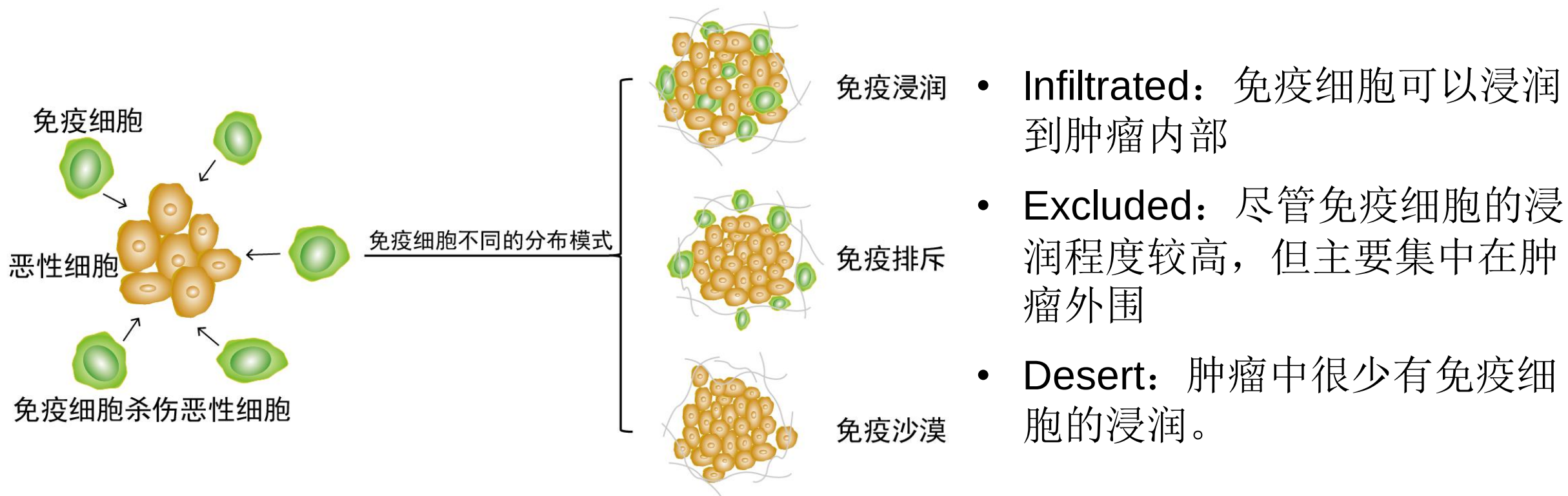


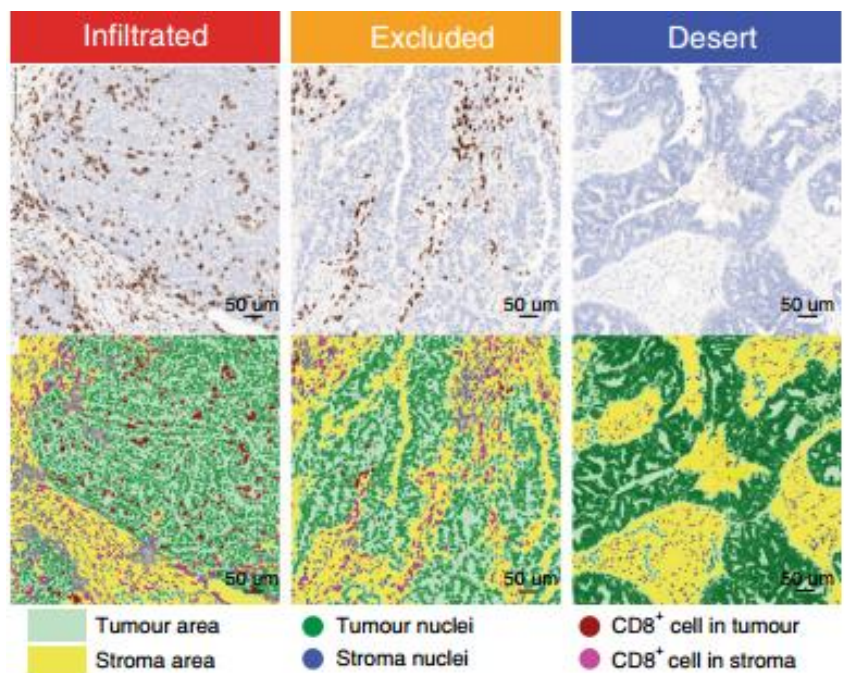
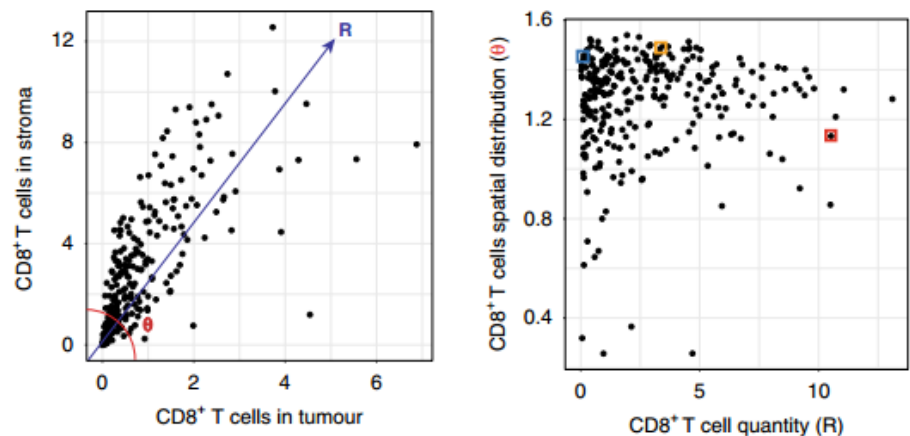
- 通过使用R包Seurat等生信工具，对结直肠癌(CRC)单细胞数据进行处理，将细胞进行降维聚类并注释每个簇的细胞类型。
- 主成分分析(PCA)将CRC的bulk样本划分了两个不同的上皮亚群，通过利用这两组的差异基因(DEG)对注释出的上皮细胞进行聚类，得到了两个主要的肿瘤亚型(iCMS2和iCMS3)和一个正常上皮细胞簇。

- 基于单细胞转录组测序量化细胞内在特征，对样本进行分型：
- iCMS2: 具有较强的免疫排斥信号，以排除或逃避细胞毒性宿主免疫系统。
 - iCMS3: 具有显著上调的癌症成纤维细胞信号，形成肿瘤免疫抑制生态位。



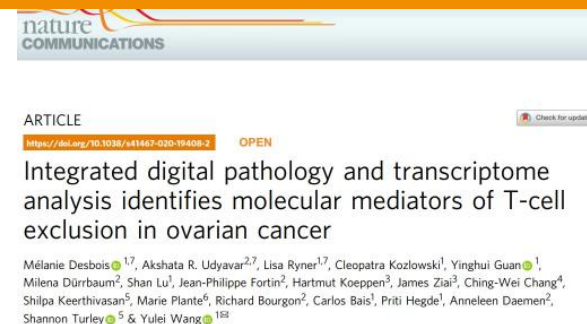
- 癌症免疫微环境中T细胞的空间分布可进一步划分癌症免疫分型，比较经典的亚型划分包括“浸润型”（“Infiltrated”）、“排斥型”（“Excluded”）和“沙漠型”（“Desert”）。

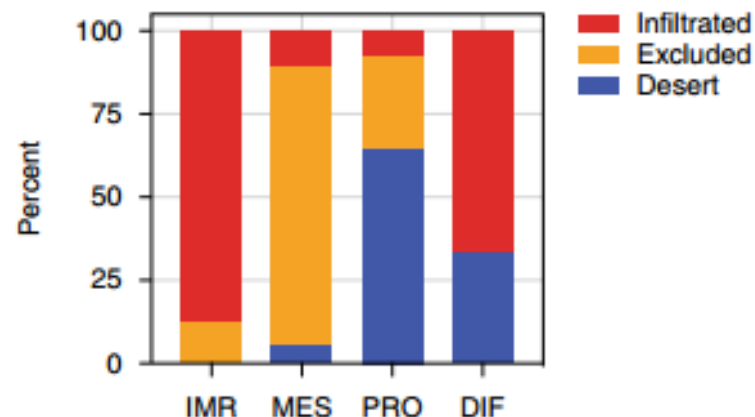




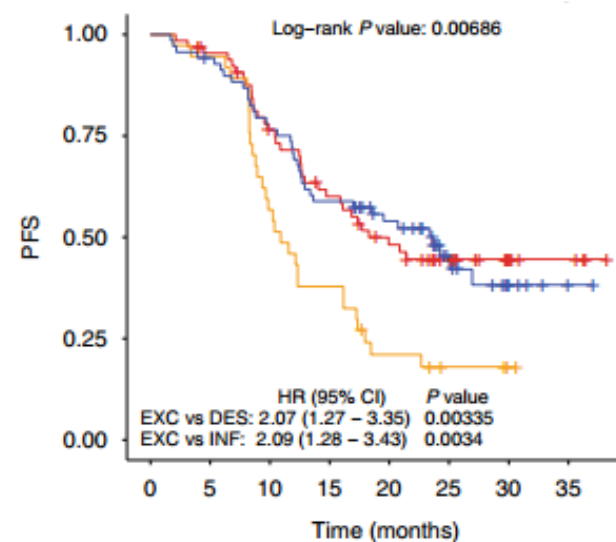
- 对病理切片进行染色，对卵巢肿瘤的上皮细胞和基质细胞进行区分，并**量化CD8+ T细胞计数和空间分布**。
- 为了更好地捕获和量化CD8+ T细胞浸润模式，将这些CD8指标转换为极坐标。

基于这两个指标将样本划分为“浸润型”、“排斥型”和“沙漠型”。





- 卵巢癌的四种临床相关的分子亚型，具有不同的免疫亚型的富集程度：**IMR**分子亚型在浸润型中高度富集；而**MES**肿瘤在排斥型中高度富集；沙漠型肿瘤主要为**PRO**或**DIF**分子亚型。

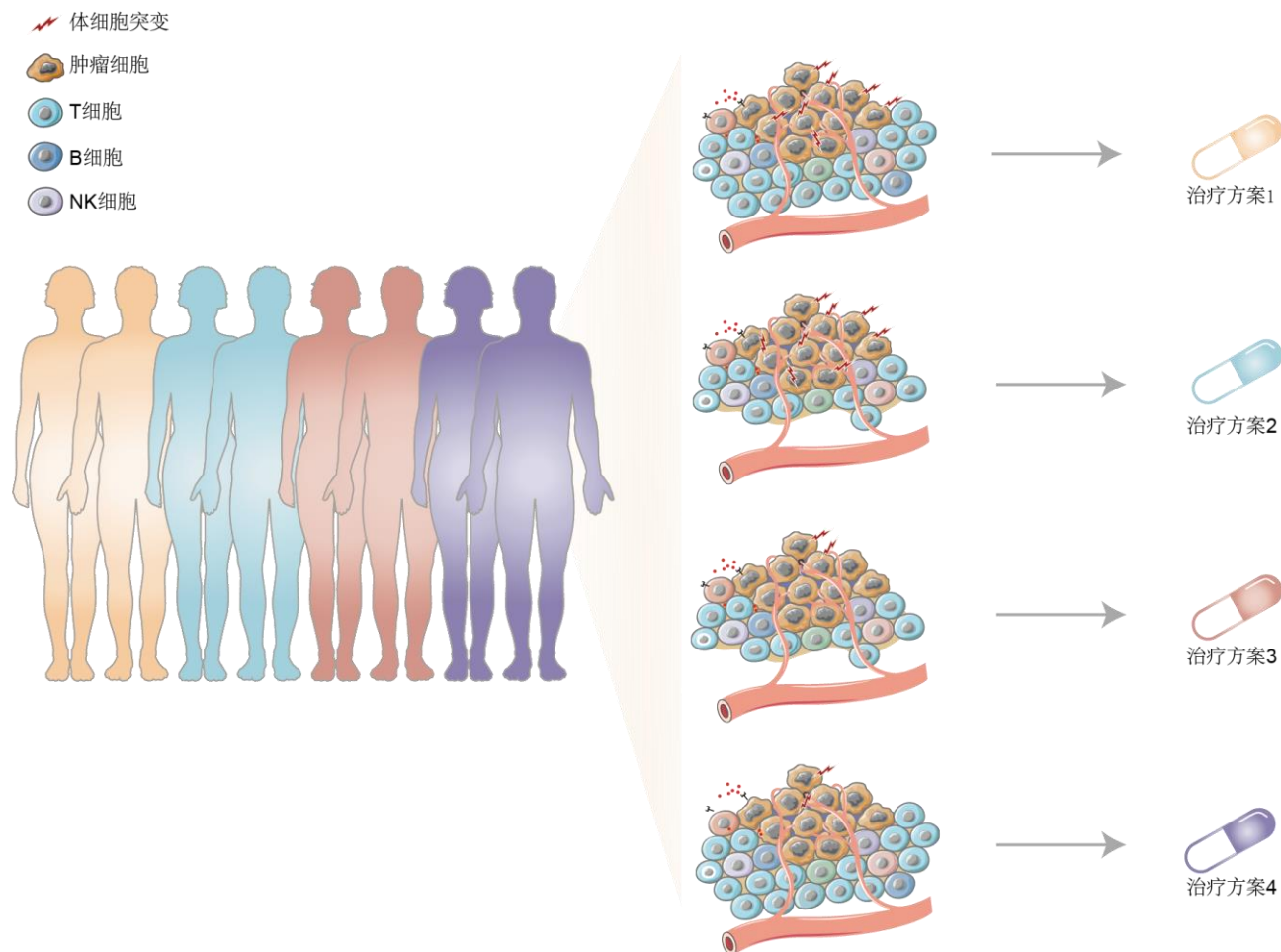


- 生存分析结果显示，与浸润型或沙漠型肿瘤相比，排除型肿瘤的无进展生存期(PFS)明显较短。

Infiltrated	66	61	47	36	27	18	7	4
Excluded	37	35	21	14	7	4	1	0
Desert	69	64	52	40	31	14	5	1

Desbois et al, Nat Commun, 2020

- 肿瘤代表着复杂的生态系统，不同肿瘤微环境中通常具有不同类型和比例的免疫和基质细胞，在治疗时充分考虑肿瘤微环境的异质性是实现精准医疗的必要措施。



- 常见的治疗干预包括手术、化学药物治疗、放射治疗和免疫治疗，可能会影响TME
- 治疗干预也会受到TME和患者全身变化的影响。
- ✓ 因此，精确刻画治疗过程中患者体内肿瘤微环境的变化，这极大地促进了精准医学的发展。

肿瘤免疫治疗相关数据资源

数据库名称	内容与特点	网址
TIGER	包含了 1508 个有临床免疫治疗信息的肿瘤样本（ 8 种癌症）和 11057 个没有免疫治疗信息的肿瘤/正常样本（ 33 种癌症）的 bulk RNA-seq 数据，以及 655 个样本（ 25 种癌症）中 2116945 个免疫细胞的单细胞转录组数据。还从已发表的肿瘤免疫研究中收集了 31 个免疫治疗相关的 CRISPR screening 数据集。	http://tiger.canceromics.org/
TISIDB	包含基于接受免疫治疗患者队列的转录组和外显子组数据所识别的治疗响应组与非响应组差异表达基因或差异突变基因共 2178 个。	http://cis.hku.hk/TISIDB/
CAMOIP	泛癌免疫治疗多组学综合分析的 Web 服务器，可以筛选预后标志物，并分析与标志物和免疫疗法相关的机制（ 4000 多名患者）。	http://www.camoip.net/
TCIA	主要包含了 TCGA 的 20 种实体瘤的综合免疫基因组分析结果，包含部分黑色素瘤的免疫治疗数据	https://www.tcia.at/home
TICH	收集了人类和小鼠的肿瘤相关 单细胞转录组数据，并进行了分析。除了未经治疗的患者数据集外，还包括那些接受过治疗的患者数据集，包含一些免疫治疗、化疗和靶向治疗的数据。	http://tisch.comp-genomics.org/home/
ICBatlas	是一个为 Immune Checkpoint Blockade (ICB) 治疗患者提供全面表达资源和功能分析的数据库。它还比较了有反应组与无反应组，以及治疗前与治疗组之间的比较分析结果。	https://guolab.wchscu.cn/ICBatlas/
ICBcomb	通过比较 ICB 治疗组、其他药物治疗组和联合治疗组的差异，分析接受 ICB 联合治疗的人类和小鼠转录组学特征并对其进行全面描述。 ICBcomb 包含 1399 个样本，涉及 29 种癌症类型和 52 种药物。	https://guolab.wchscu.cn/ICBcomb/

 Tumor Immunotherapy Gene Expression Resource[Home](#) [Single-cell Immunity](#) [Immunotherapy Response](#) [Response Signature](#) [Immune Screening](#) [Search](#) [Help](#) [Download](#) [Statistics](#) [Contact](#)

Welcome to TIGER

A Web Portal of Tumor Immunotherapy Gene Expression Resource

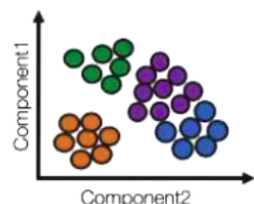
Please Input A Gene Symbol (e.g. PDCD1 or CXCL13)

Search

TIGER is a web-accessible portal for integrative analysis of the gene expression data related to tumor immunology. TIGER contains bulk transcriptome data for 1508 tumor samples with immunotherapy clinical outcome and 11,057 tumor/normal samples from TCGA, and single cell transcriptome data for 2,116,945 cells of 655 samples, among which 119,039 cells of 63 samples have immunotherapy clinical data.

TIGER provides the following four modules to access the resource

Single Cell Immunity



Immunotherapy Response



Response Signature



Immune Screening



TIGER将收集的数据整合到四个模块中：单细胞免疫（Single-cell Immunity）、免疫治疗响应（Immunotherapy Response）、响应标签（Response Signature）和免疫筛选（Immune Screening）。用户可以通过这四个模块对数据进行深入的挖掘和结果展示。

TIGER提供的可下载的四种数据

带有临床信息的免疫治疗bulk数据，免疫治疗单细胞数据的marker矩阵以及差异基因矩阵（未提供原始表达矩阵），免疫响应特征，免疫筛选数据

[Home](#) [Single-cell Immunity](#) [Immunotherapy Response](#) [Response Signature](#) [Immune Screening](#) [Search](#) [Help](#) [Download](#) [Statistics](#) [Contact](#)

Immunotherapy Dataset

Dataset ID	Cancer Type	PMID	Therapy	Dataset Size	Responder Number	Non-Responder Number	Expression Data	Clinical Data
GBM-PRJNA482620	Glioblastoma	30996326	anti-PD-1	34	17	17	Download Rds Download Tsv	Download Rds Download Tsv
HNSC-GSE93157	Squamous head and neck	28487385	anti-PD-1	5	2	3	Download Rds Download Tsv	Download Rds Download Tsv

Single-Cell Datasets

[Contact us](#) to obtain the processed scRNA-seq data.

Dataset Name	Cancer Type	Differential Expression Analysis	PMID	Sample Size	Cell Number	Cell Marker Matrix	Differential Expression Matrix
Qian J, et al. Cell Res. 2020 Jun 21	Breast cancer (BC)	Tumor	32561859	14	41,292	Download Rds Download Tsv	Download Rds Download Tsv

Immunotherapy Response Signature

Signature Name	Signature Cite	Journal Info	PMID	Download
T cell-inflamed GEP	T cell-inflamed GEP, Ayers M, 2017	J Clin Invest. 2017	28650338	Download Tsv

Immune Screening Datasets

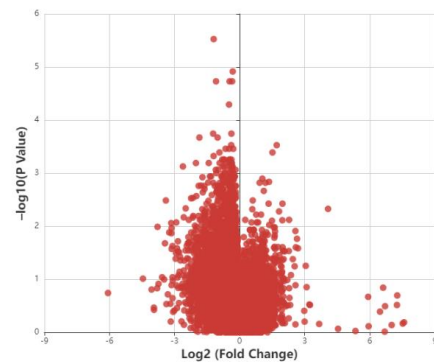
Dataset ID	Dataset Type	Cancer Type	PMID	Download
JEKO-1_T_CELL_KILLING	CRISPR-Cas9	Mantle Cell Lymphoma	29301958	Download Rds Download Tsv

免疫治疗响应 (Immunotherapy Response)

该模块提供了对bulk RNA基因表达数据与临床免疫治疗信息的分析功能，包括差异表达分析、生存分析、基因集查询。在交互式热图中，用户可以点击图中的单元格了解详细信息，如生存分析以及所选基因在应答者和非应答者之间或治疗前后的表达差异

以GBM-PRJNA482620_anti-PD-1数据集为例

Differential Expression Between Responder And Non-Responder



GBM-PRJNA482620_anti-PD-1
Gene:5S_rRNA

Responder Group Non-Responder

CR PR SD

Normalized By
☒ None ☐ Gene

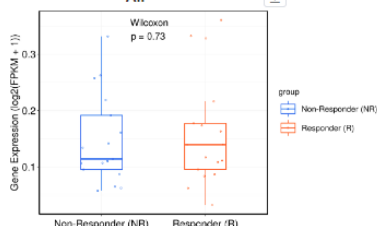
Log Scale
☒ Yes ☐ No

Clinical Attributes
☒ None ☐ Gender ☐ Age

Plot

Differential Expression between R and NR

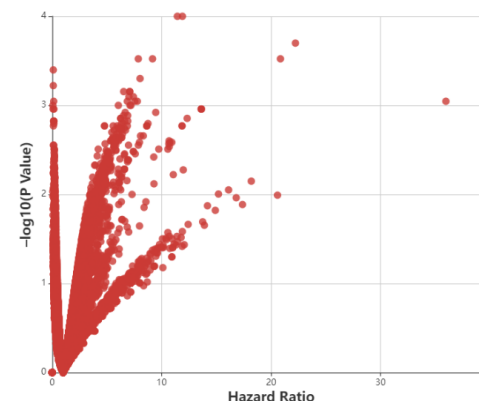
All



Performance comparison with known immune response signatures

Signature Name	Description	AUC
5S_rRNA	5S_rRNA	0.5363
T cell-inflamed GEP	CCL5, CD27, CD274, CD276, CD8A, CMKLR1, CXCL9, CXCR6, HLA-DQA1, HLA-DRB1, HLA-E, IDO1, LAG3, NKG7, PDCD1LG2, PSMB10, STAT1	0.5709
CAF	Cancer-associated fibroblast	0.5225

Survival Analysis



GBM-PRJNA482620_anti-PD-1
Gene:5S_rRNA

Normalized By

☒ None ☐ Gene

Group Cutoff

Low Expression High Expression

0.50 0.50

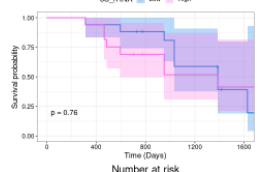
Cox Proportional-Hazards Model

☒ All ☐ Gender ☐ Age

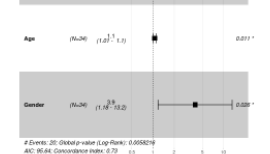
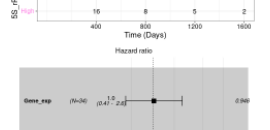
Plot

Survival Analysis (Overall Survival)

5S_rRNA Low High

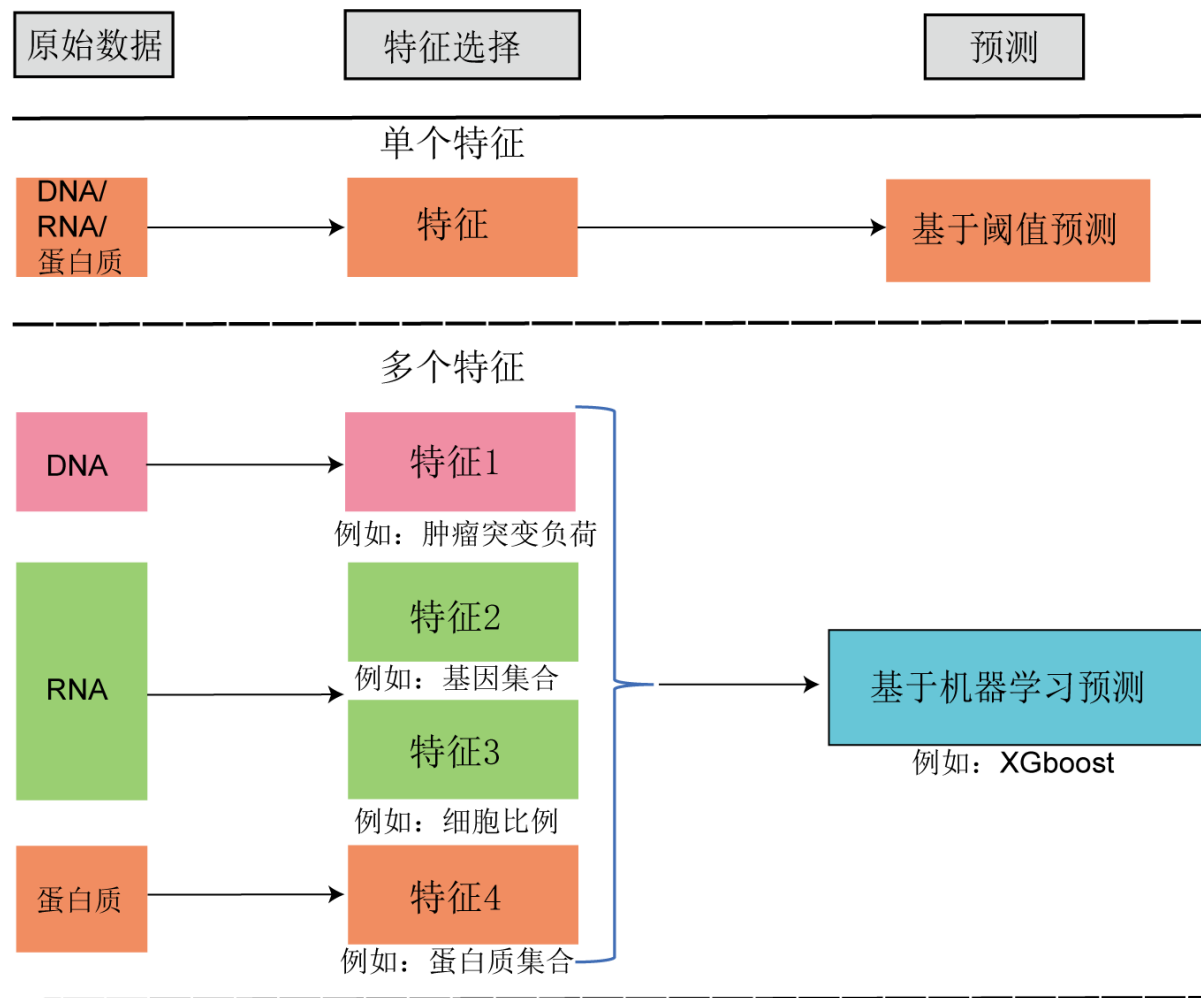


5S_rRNA Low High



Performance comparison with known immune response signatures

Signature Name	Description	HR	95% CI	P Value
5S_rRNA	5S_rRNA	1.123	0.4603-2.74	0.7558
T cell-inflamed GEP	CCL5, CD27, CD274, CD276, CD8A, CMKLR1, CXCL9, CXCR6, HLA-DQA1, HLA-DRB1, HLA-E, IDO1, LAG3, NKG7, PDCD1LG2, PSMB10, STAT1	1.478	0.5495-3.975	0.4223
IFNG	CXCL10, CXCL9, HLA-DQA1, IDO1, IFNG, STAT1	1.004	0.3765-2.679	0.9895
CD8	CD8A, CD8B	0.9456	0.3538-2.528	0.9039
CD274	CD274	1.171	0.4393-3.124	0.768
TLS	CCL19, CCL21, CXCL13, CCR7, SELL, LAMP3, CXCR4, CD86, BCL6	1.337	0.5123-3.49	0.547
CD79B, CD1D, CCR				

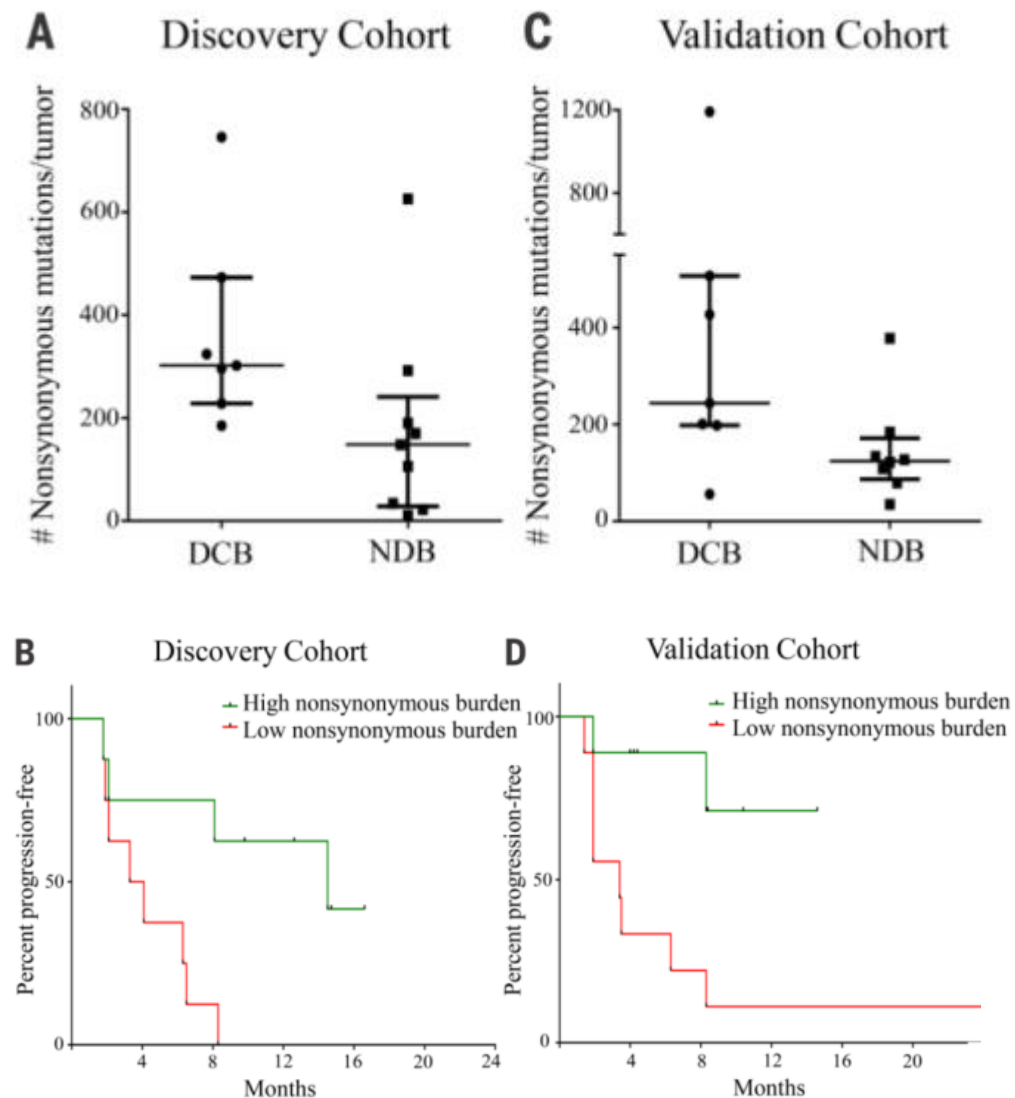


预测免疫治疗响应的策略

➤ 利用生物医学大数据以及生物信息学方法预测免疫治疗响应的策略总体可分为两类：

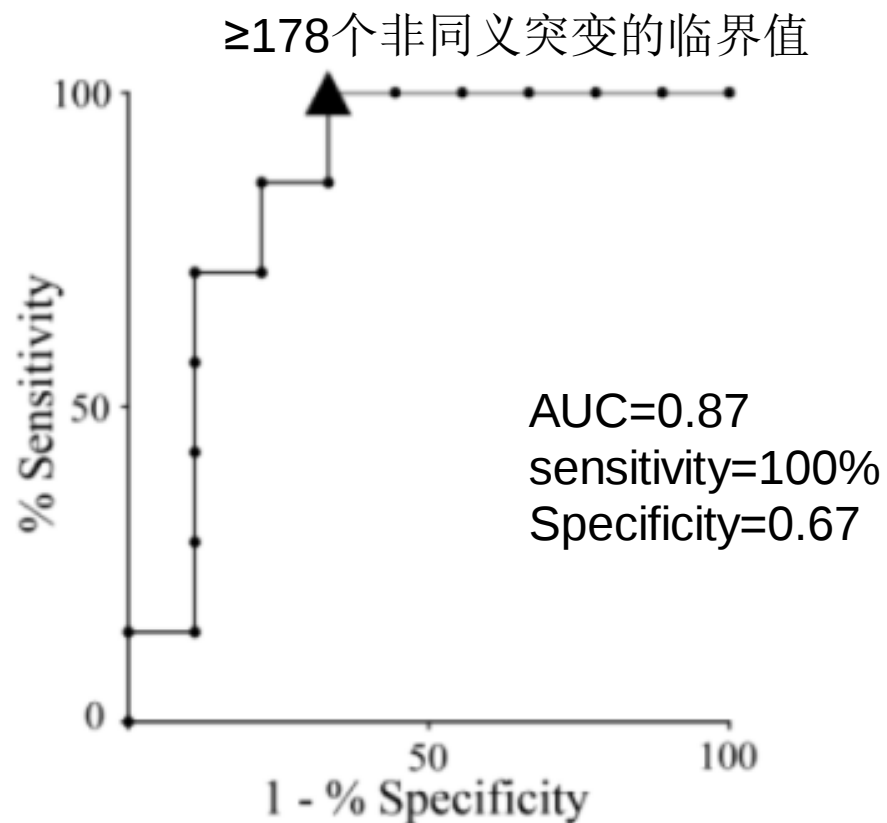
- 基于单个特征预测
- 整合多个特征建立预测模型。

这两类方法都可以评估患者是否会从免疫治疗中获益，从而更准确地识别将受益的患者，对免疫治疗的优化具有重大意义



- 对来自用抗PD-1治疗的两个独立患者队列的非小细胞肺癌的外显子组(样本数分别为 $n = 16$ 和 $n = 18$)及其匹配的正常DNA进行了测序，以此来探究非小细胞肺癌的突变景观可能影响对抗PD-1治疗的反应。
- 文章发现抗PD-1的临床疗效与较高的体细胞非同义突变负荷相关，在测试队列($n = 16$)中，DCB的患者非同义突变的中位数为302，而NDB为148（两组间Whitney $P = 0.02$ ）。
- 在验证队列($n = 18$)中，DCB的患者非同义突变的中位数为244，而NDB的为125（两组间Whitney $P = 0.04$ ）。

DCB: 持久临床获益
NDB: 无持久临床获益



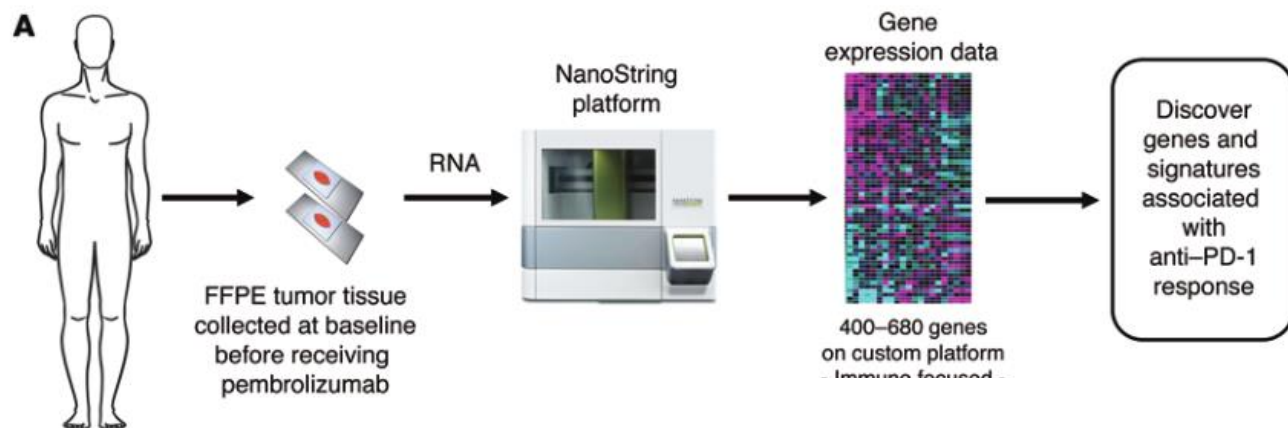
非同义突变负荷与DCB的ROC曲线

- 测试队列中，非同义突变负荷和DCB之间有很高的一致性，AUC为87%。
- 非同义突变负荷≥178(结合最大敏感性和最佳特异性的分界点)的患者DCB的可能性比为3.0。
- 将这一分界点应用于验证队列，肿瘤患者中含有≥178个突变的DCB发生率为75%,而<178个突变的DCB发生率为14%,相应的敏感性为86%,特异性为75%。

DCB: 持久临床获益

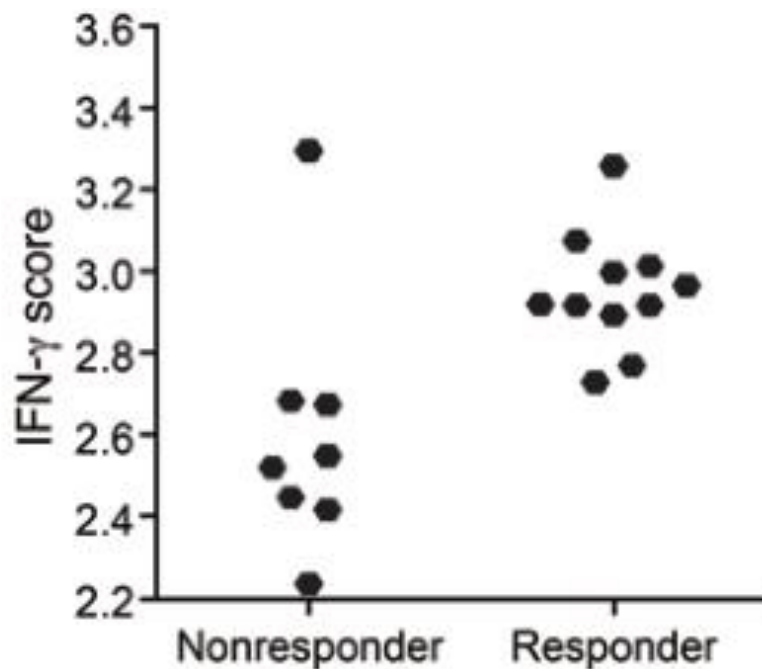
NDB: 无持久临床获益

样本获取



- 19名转移性黑色素瘤患者活检样本（8个响应，11个非响应）
- 检测了680个肿瘤免疫相关基因的表达量
- 分析应答者和非应答者之间的差异基因进而构建基因集来对接受免疫治疗的患者结局进行预测。

特征提取---与IFN- γ 信号直接相关



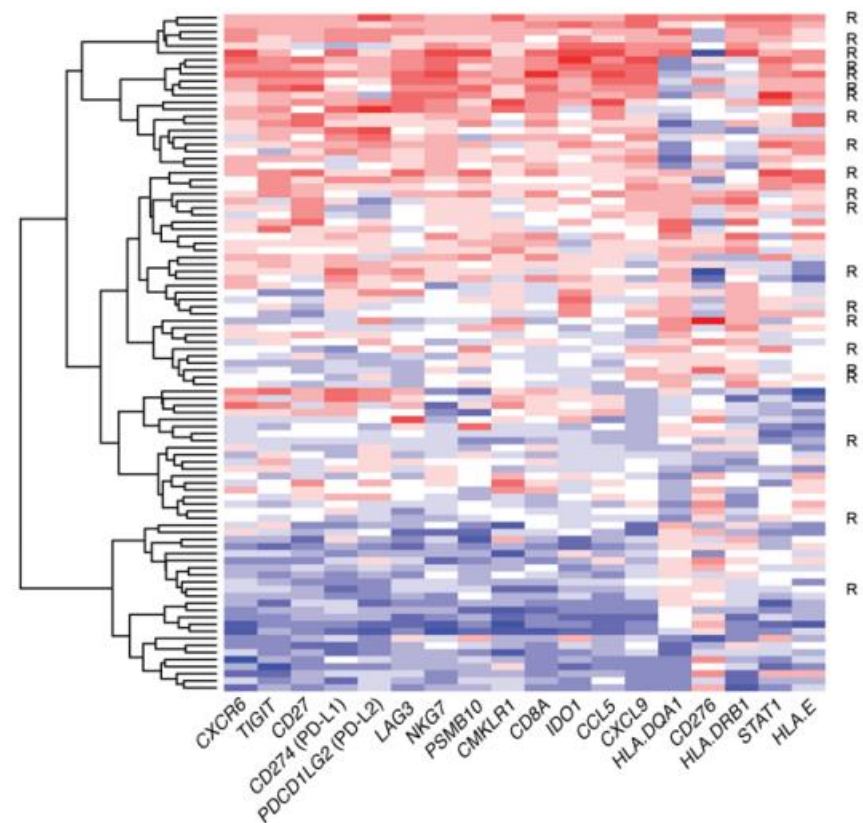
- 发现应答者和非应答者之间的差异TOP基因与IFN- γ 信号直接相关，并与IFN- γ 的表达相关，基于此构建了一个包含10个基因 “preliminary IFN- γ ” signature (IFNG、STAT1、CCR5、CXCL9、CXCL10、CXCL11、IDO1、PRF1、GZMA和MHCII·HLA-DRA),能够在19名黑色素瘤试点数据患者中区分抗PD-1治疗的应答者和无应答者。

特征提取

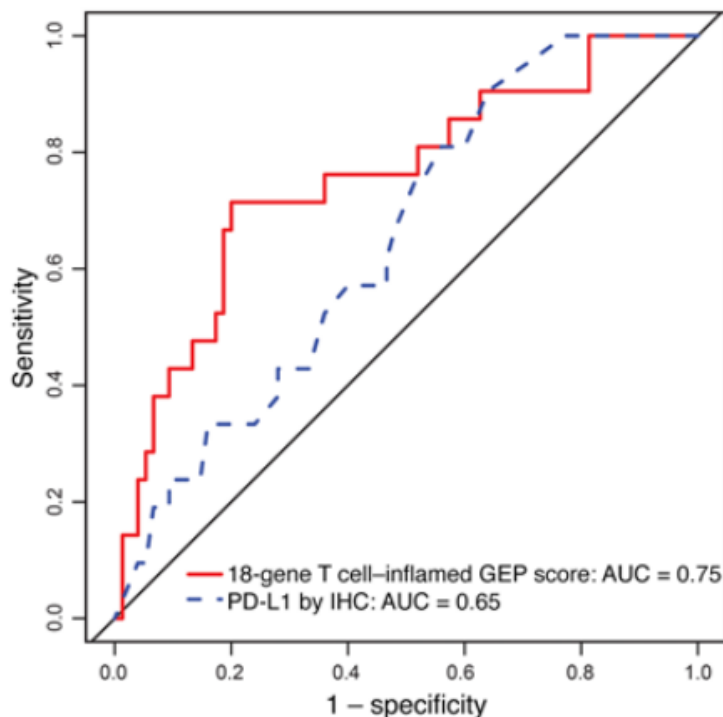


- 基于上一步识别到的“preliminary IFN- γ ” signature的基础上纳入了与19个黑素瘤肿瘤患者预后正相关的基因，进而形成了包含28个基因在内的“Preliminary expanded ” signature。
- 这28个基因包括与细胞溶解活性相关的基因(例如, granzyme A/B/K、PRF1)、引发炎症的细胞因子/趋化因子(CXCR6、CXCL9、CCL5和CCR5)、T细胞标记(CD3D、CD3E、CD2、IL2RG)、NK细胞活性(NKG7、HLA-E)、抗原呈递(CIITA、HLADRA)和其他免疫调节因子(LAG3、IDO1、SLAMF6)。
- 测试这2个signature发现，其均与接受免疫治疗的临床结果之间显著关联。

构建模型

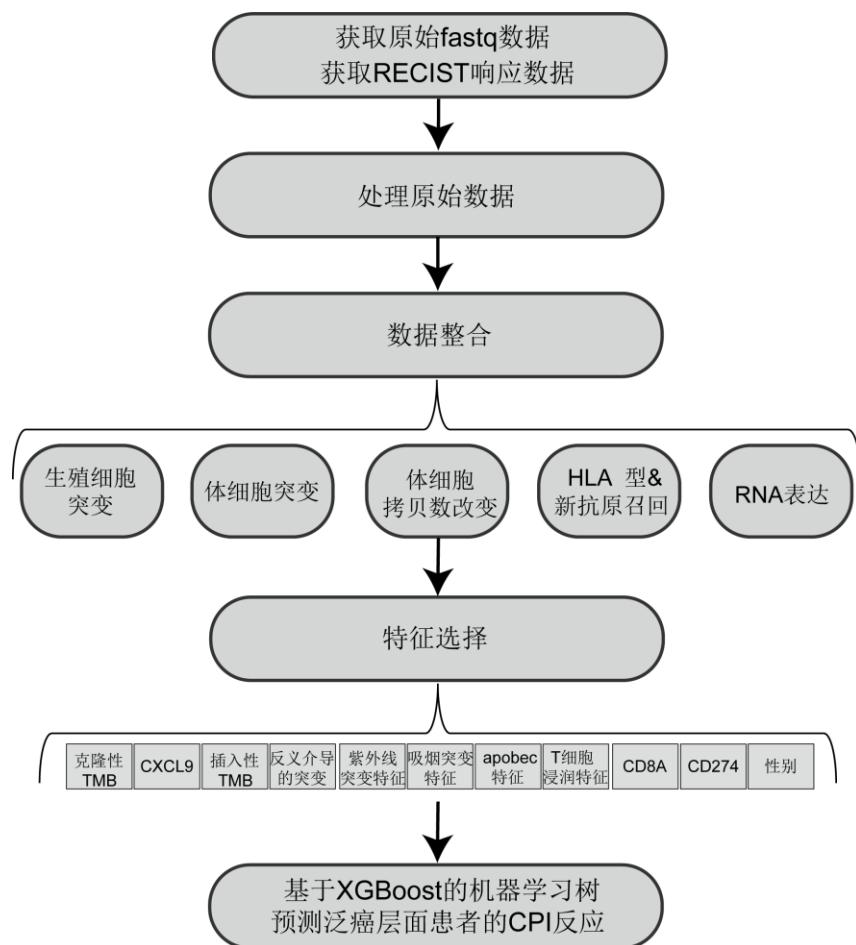


最终的18个signature gene表达热图
行为样本，列为基因



- 最后 在另外一组包含62个黑色素瘤肿瘤队列中，删除与预后关联不显著的基因，最终形成了包含18个基因在内的 **T cell-inflamed signature**，其在独立的HNSCC患者队列中的预测效能为0.75，高于PD-L1(0.65)。

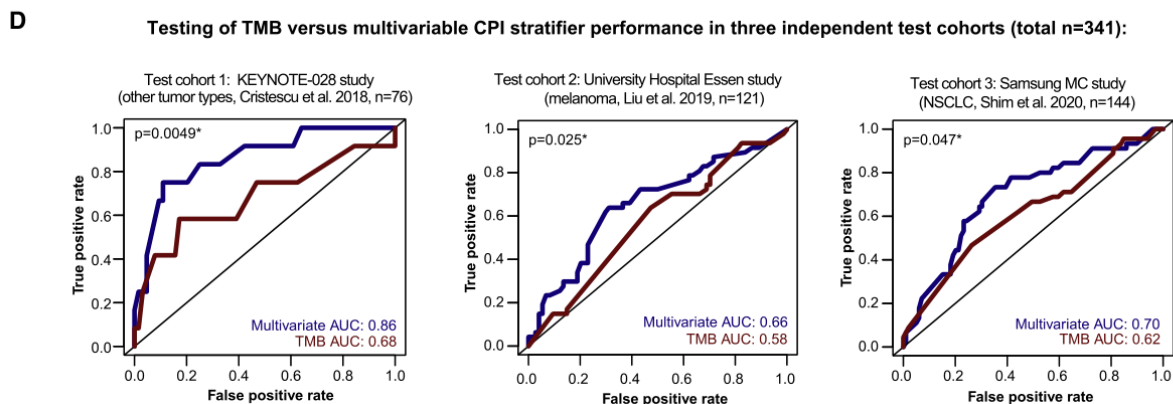
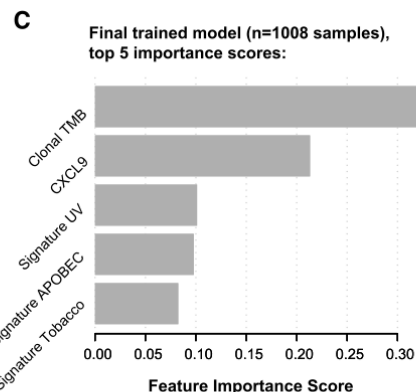
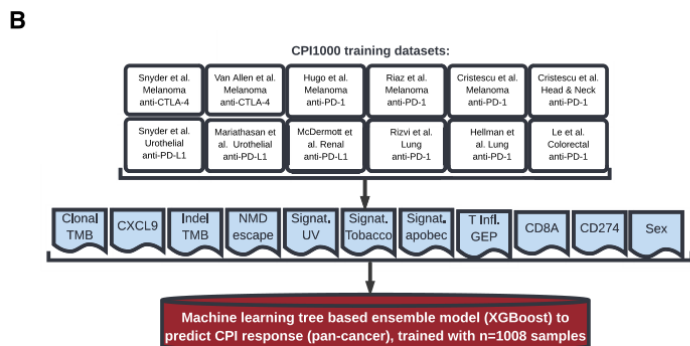
- 随着免疫治疗相关组学数据越来越丰富，利用生物信息学方法我们可以破译不同标志物之间的相对权重和相关性。



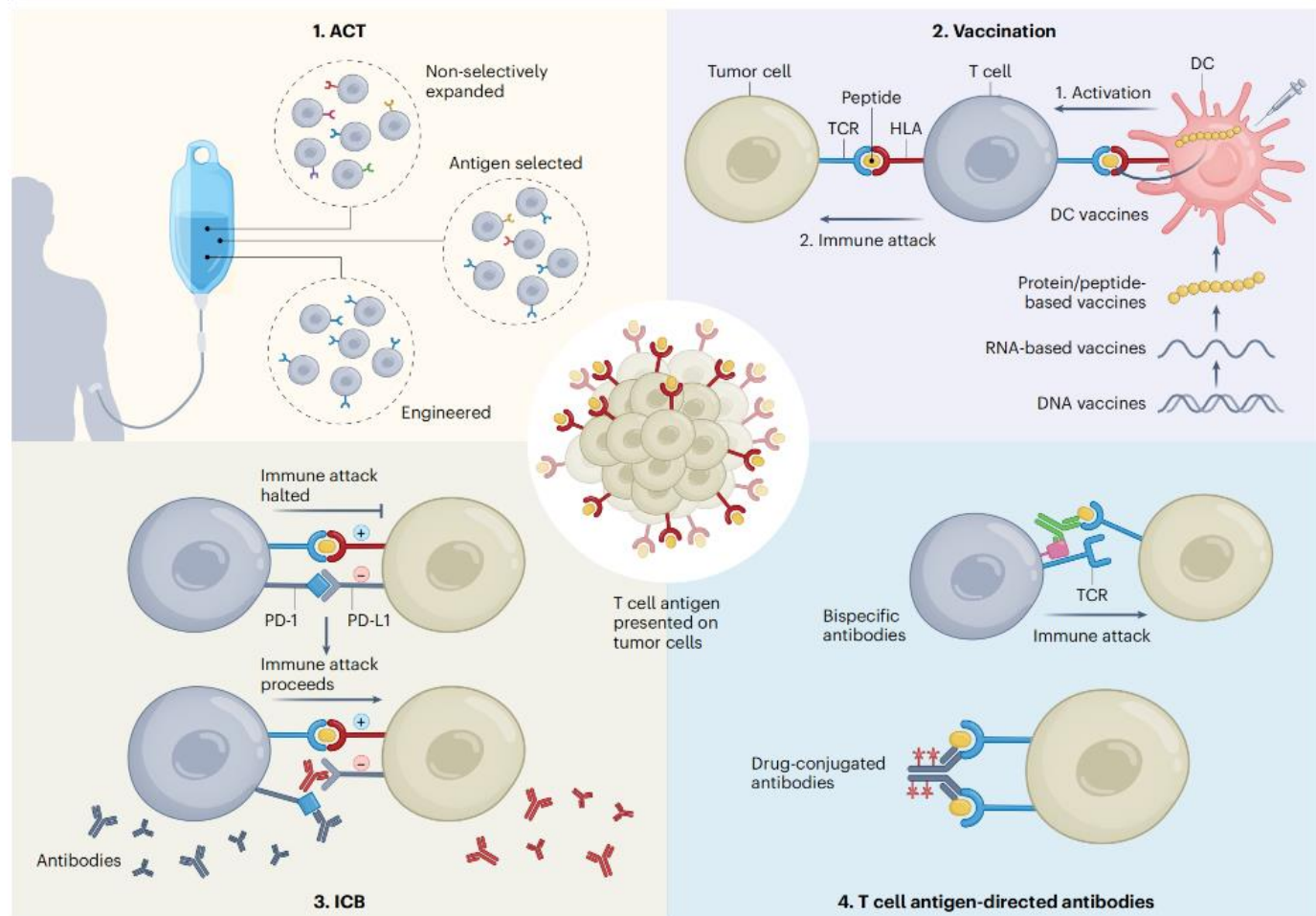
整合多个特征预测模型流程图

- 收集整理了多种类型免疫检查点抑制剂的应答标志物，
- 通过meta分析将效应量整合起来
- 基于机器学习算法XGBoost，针对每种癌症类型构建多变量预测模型。

Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition



- 从12篇研究中收集了1000多个接受免疫治疗的样本训练多变量泛癌模型
- 最终发现预测效应最强的生物标志物是克隆性TMB
- 多变量的预测模型的预测效能可在3个独立的数据集中均优于TMB。



nature cancer

Review article

<https://doi.org/10.1038/s43018-023-00588-x>

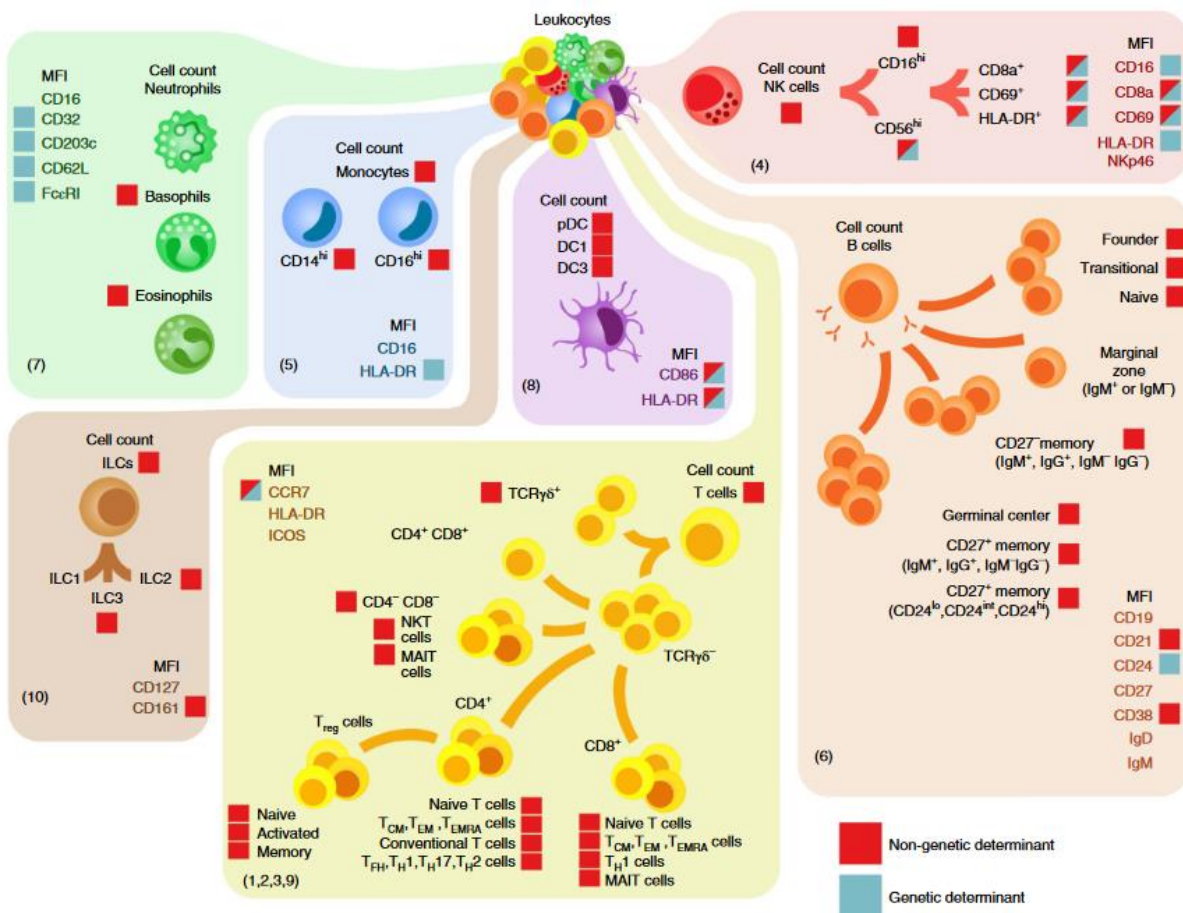
The landscape of T cell antigens for cancer immunotherapy

Received: 21 December 2021

Aviyah Peri¹, Nadja Salomon², Yochai Wolf^{3,4}, Sebastian Kreiter²,
Mustafa Diken² & Yardena Samuels¹

Accepted: 18 May 2023

- 免疫检查点抑制剂、基于新抗原的个体化癌症疫苗和过继T细胞疗法等免疫疗法正在逐步走向临床，使癌症患者的精准化治疗成为可能。



nature immunology

Resource | Published: 23 February 2018

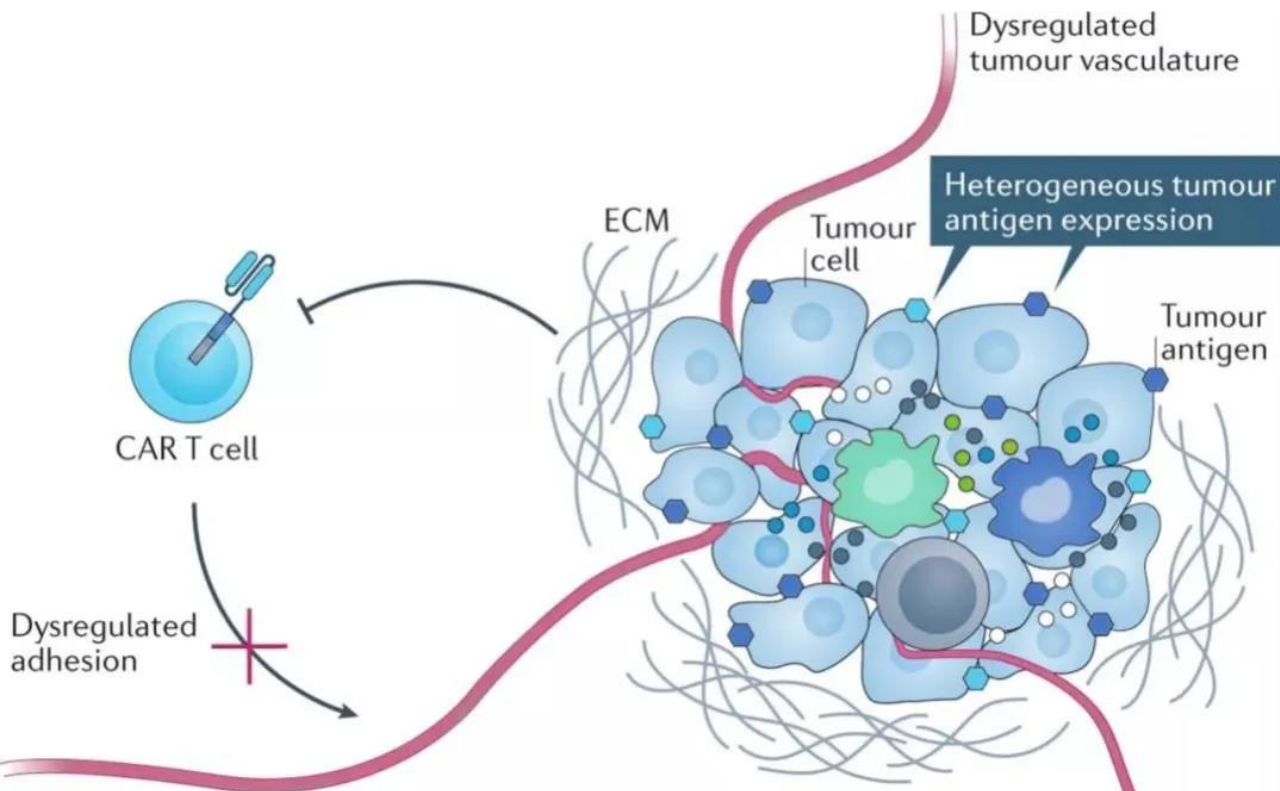
Natural variation in the parameters of innate immune cells is preferentially driven by genetic factors

Etienne Patin , Milena Hasan, Jacob Bergstedt, Vincent Rouilly, Valentina Libri, Alejandra Urrutia, Cécile Alanio, Petar Scepanovic, Christian Hammer, Friederike Jönsson, Benoît Beitz, Hélène Quach, Yoong Wearn Lim, Julie Hunkapiller, Magge Zepeda, Cherie Green, Barbara Piasecka, Claire Leloup, Lars Rogge, François Huetz, Isabelle Peguillet, Olivier Lantz, Magnus Fontes, James P. Di Santo, The Milieu Intérieur Consortium  + Show authors

➤ 但是仍然存在一些问题和挑战。

- 首先，肿瘤浸润淋巴细胞的分类需要进一步细分，不同类型的TIL浸润对精准治疗的预后指导意义是不一样的。
- 肿瘤细胞可能存在免疫逃逸，因此需要探索更好的联合治疗策略，从而激活机体的固有免疫和改善肿瘤免疫微环境。
- 进入实体瘤的免疫细胞可能会面临不一样的营养和代谢微环境。

Andrew J. Hou, Laurence C. Chen & Yvonne Y. Chen



- 但是仍然存在一些问题和挑战。
- 实体瘤的物理屏障（如胶原纤维等基质屏障，血脑屏障）可能会阻碍免疫细胞进入实体瘤组织并且降低免疫细胞的疗效。

1. **版权声明：**本PPT及其所有内容（以下简称“本PPT”）仅用于教育和教学用途，版权归属于本PPT作者。
2. **使用要求：**任何使用本PPT的行为均须遵守以下条件：
 - 1) **致谢和标注：**若部分或全部使用本PPT的内容，请在使用内容的适当位置标注出处，并致谢本PPT作者。
 - 2) **修改和再分发：**未经作者书面许可，不得对本PPT进行修改或再分发。
3. **禁止商业化使用：**严禁将本PPT用于任何形式的商业化用途，包括但不限于：
 - 1) 通过网络或其他途径进行付费使用或分发；
 - 2) 在商业培训、广告或其他商业活动中使用本PPT的内容。
4. **法律责任：**任何违反上述条款的行为，作者保留追究法律责任的权利，包括但不限于：
 - 1) 要求停止侵权行为；
 - 2) 追究侵权使用者的经济赔偿责任。
5. **其他规定：**
 - 1) 本使用条款的解释权归本PPT作者所有。
 - 2) 作者保留随时更新本使用条款的权利，更新后的条款将即时生效。

谢谢大家!