基于系统生物学的精确医学

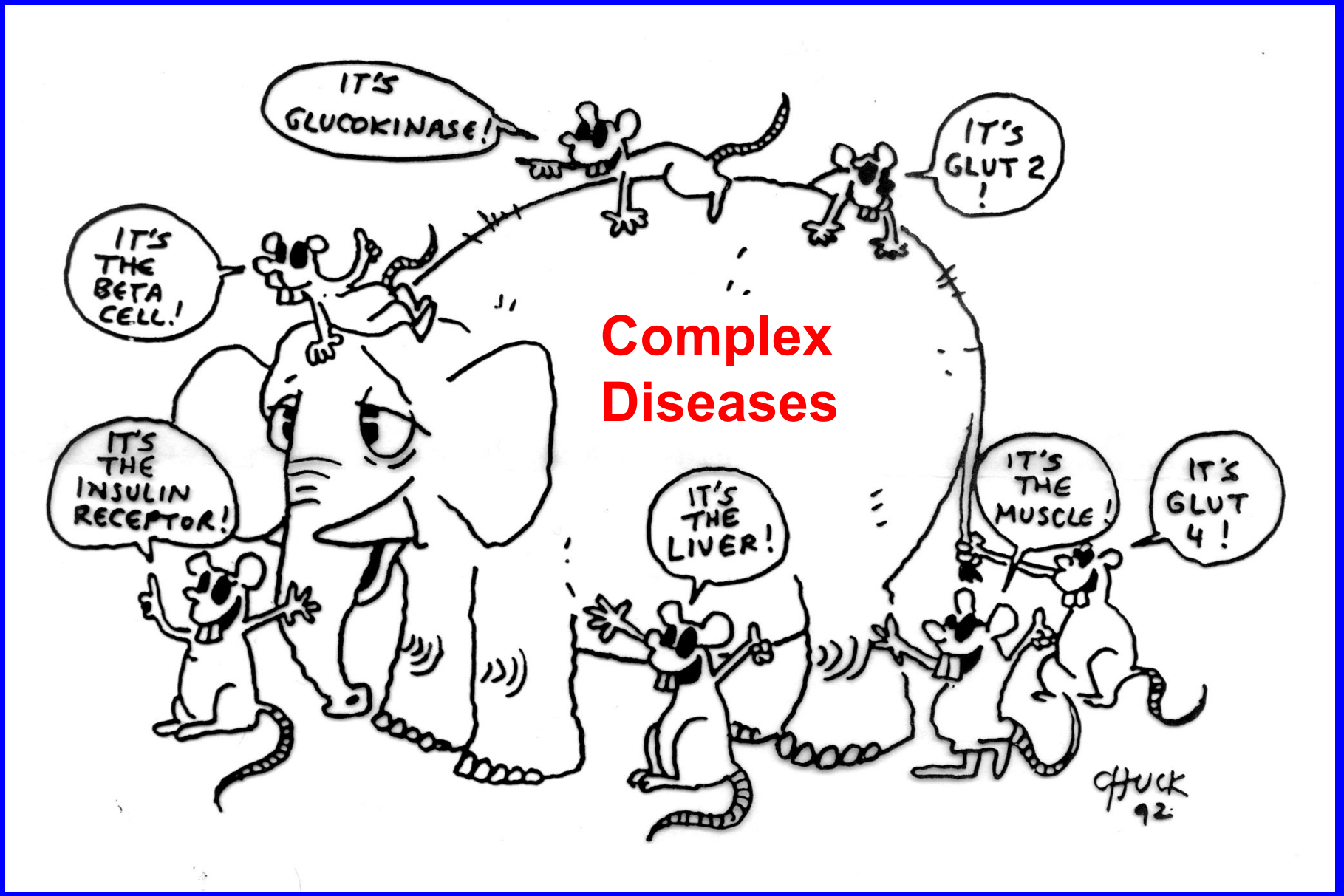
吴家睿

中国科学院系统生物学重点实验室

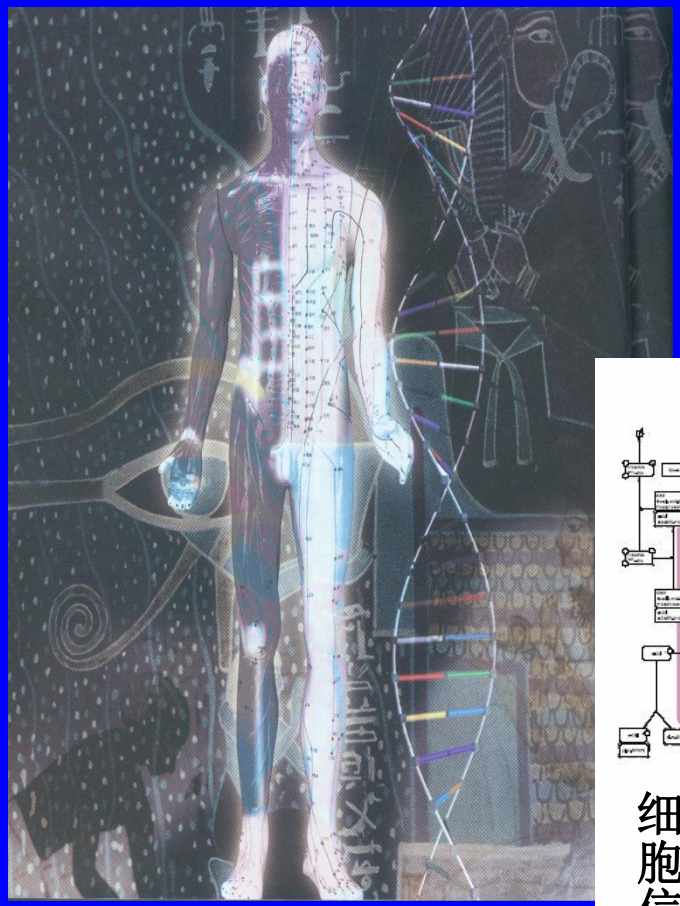
中国科学院生物化学与细胞生物学研究所

个体化的生命复杂系统

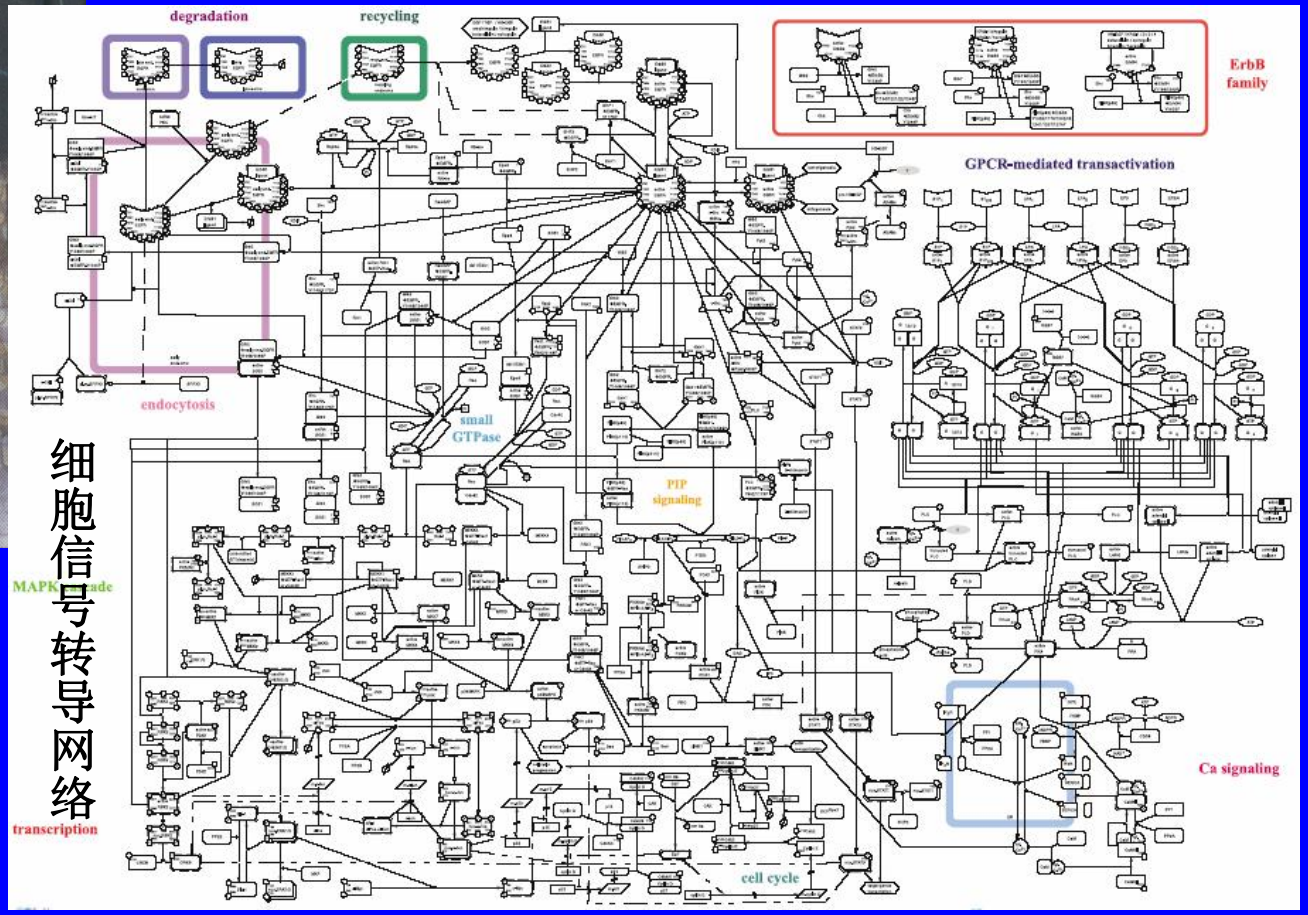
基于分子生物学的疾病观



生命是一个复杂系统

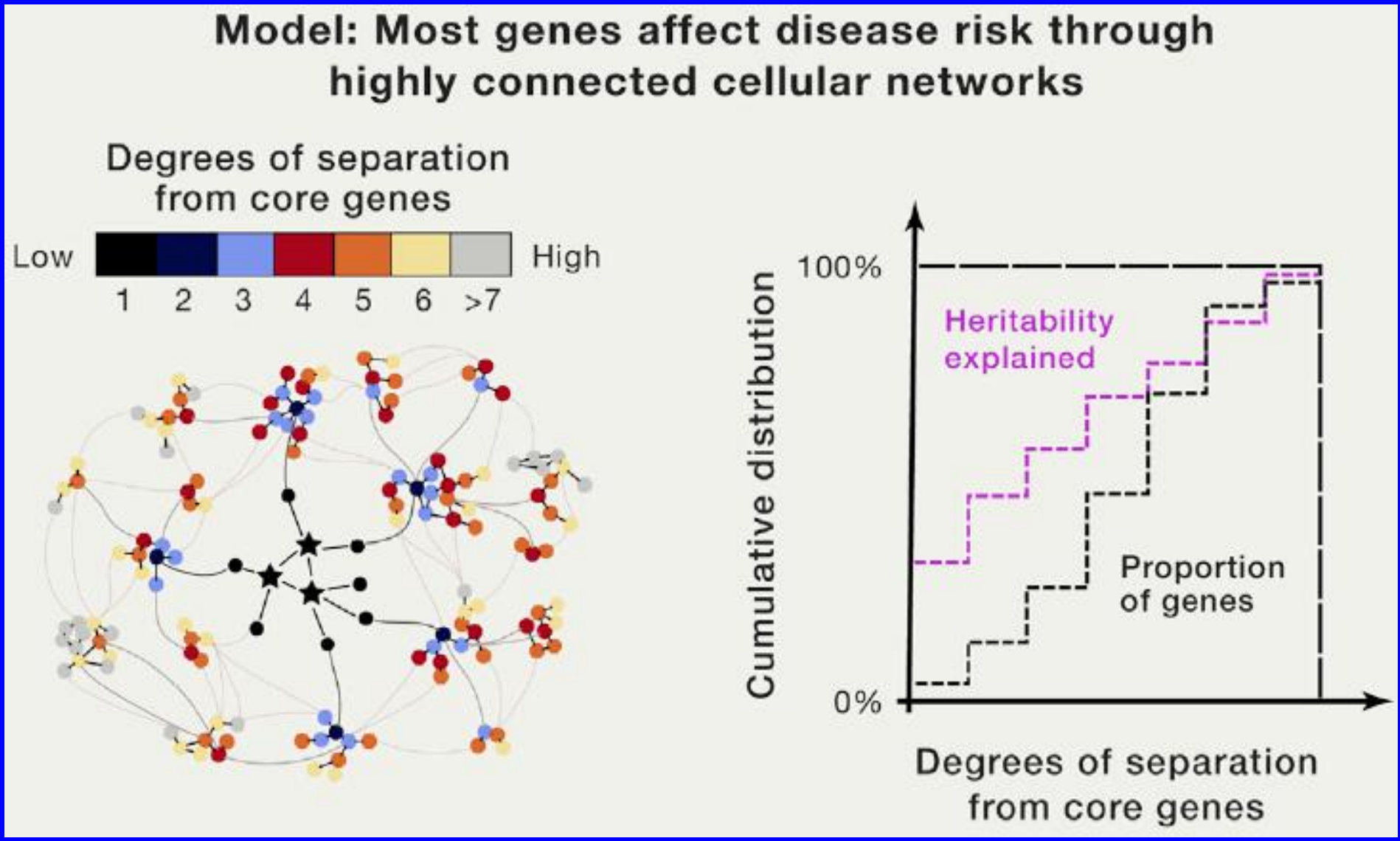


- 基因组： 3 X 10^9 bp
- 基因数： ~25000
- 蛋白数： >200,000



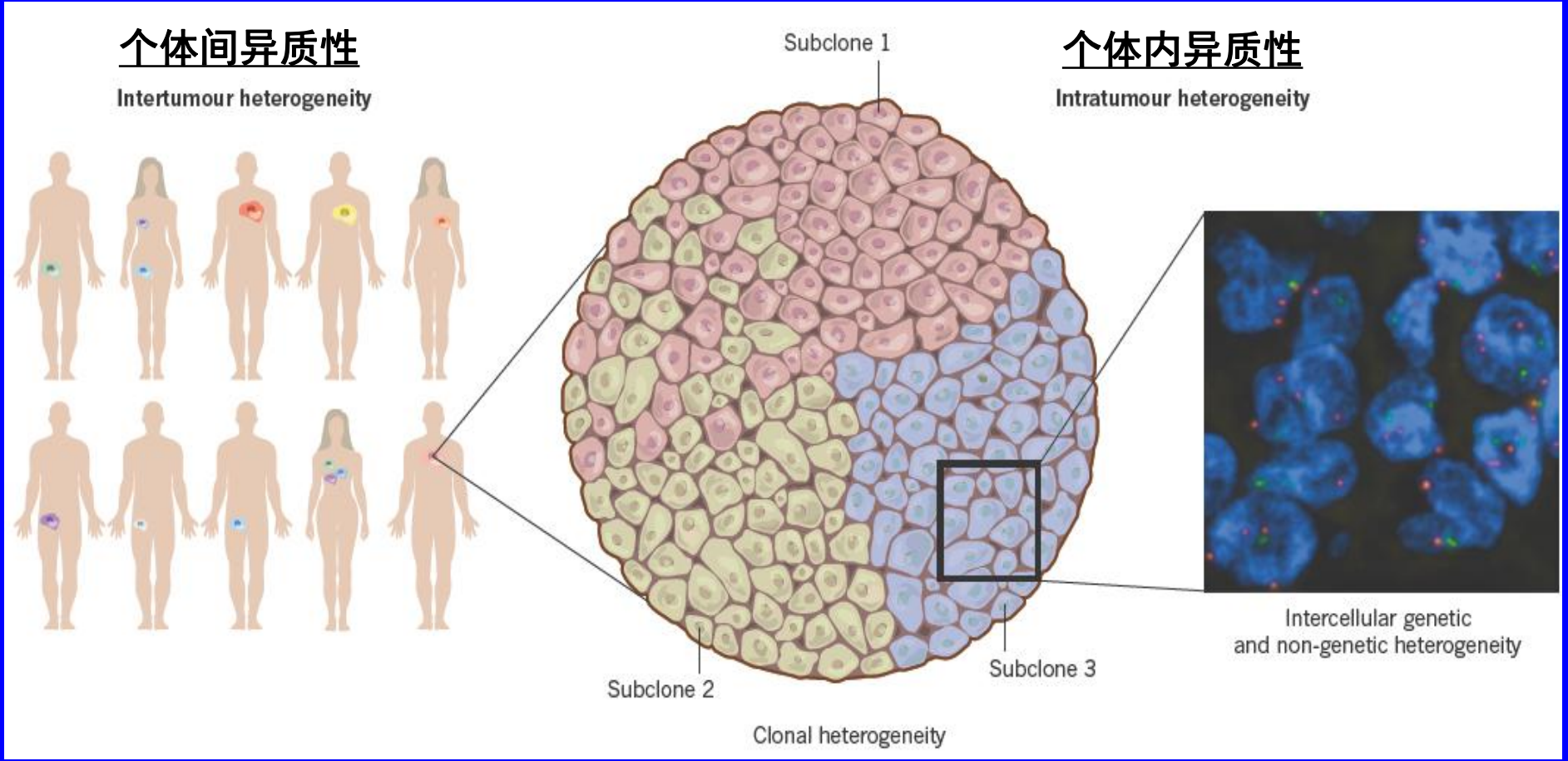
细胞信号转导网络

全基因疾病模型（Omnigenic Model）



Cell 2017 – 06 –15

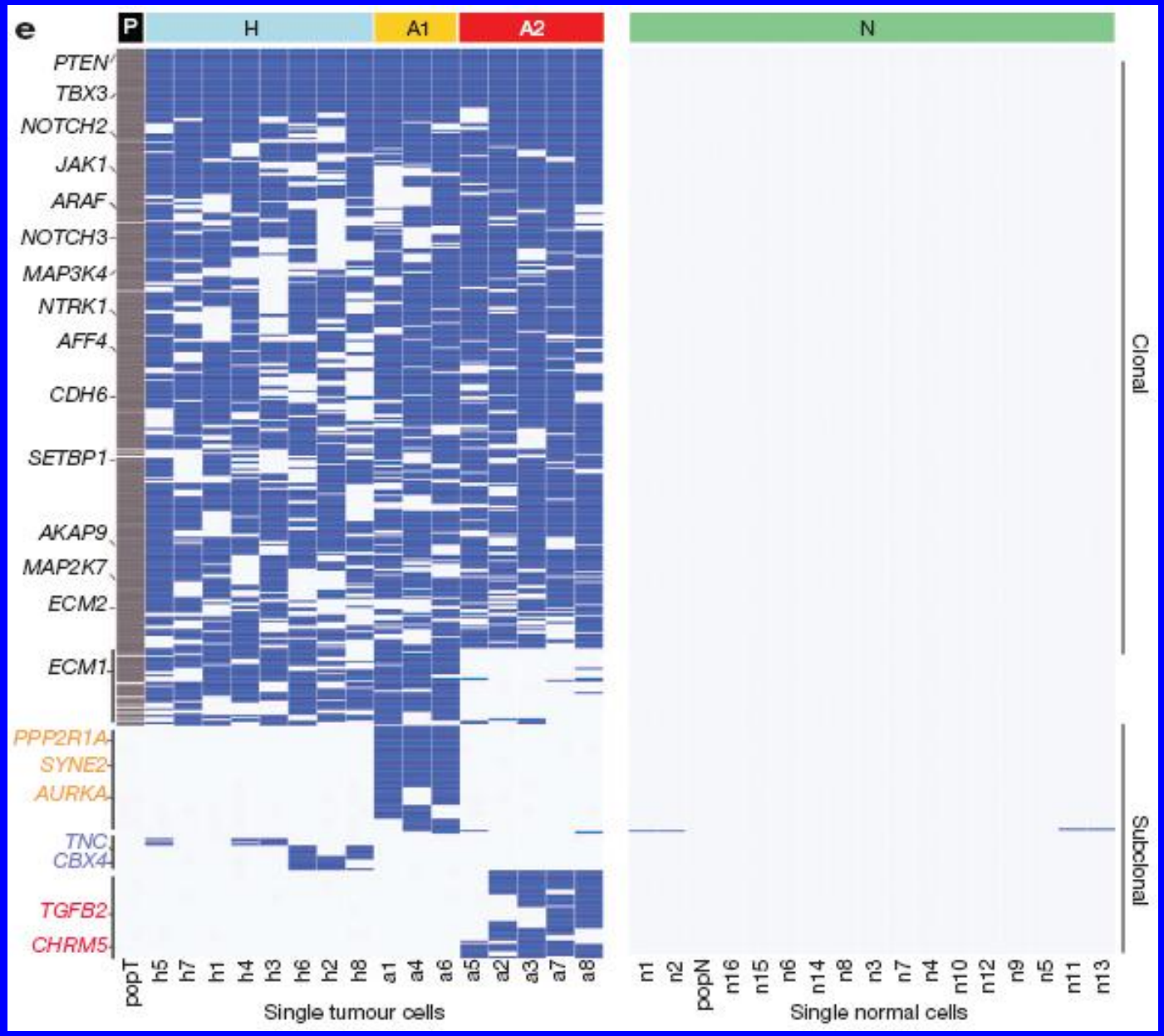
复杂性疾病的挑战：异质性



Nature 2013, 501:338

肿瘤组织内在的异质性

单乳腺癌肿瘤的单细胞非同义点突



Nature 2014, 512:155

基于群体观的人体代谢标准

“血糖指数”（glycemic index，GI）：

含有50g的碳水化合物的食物与相当量的葡萄糖相比，在一定时间内（一般为餐后2小时）引起体内血糖应答水平的百分比。

1, 不同的食物有着不同的GI值，GI值是食物的内在属性。

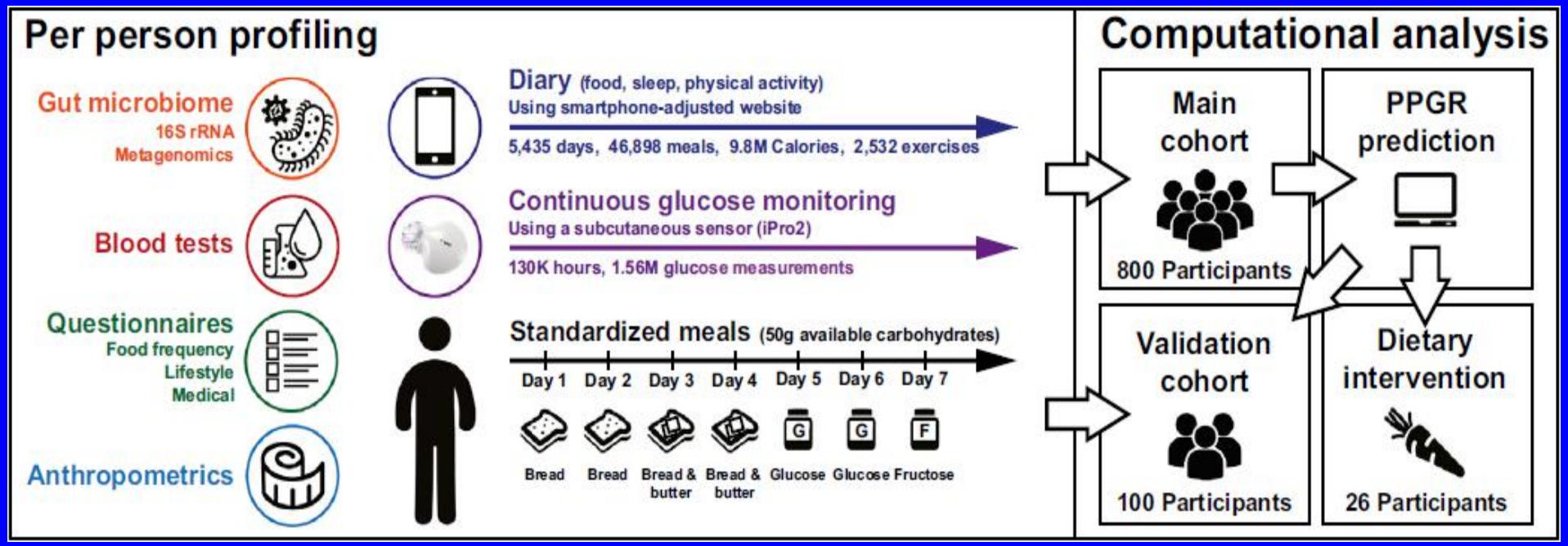
2, 不同的个体对同一种食物有着同样的GI响应。

经典的群体营养观：膳食指南

中国居民膳食指南（2007）

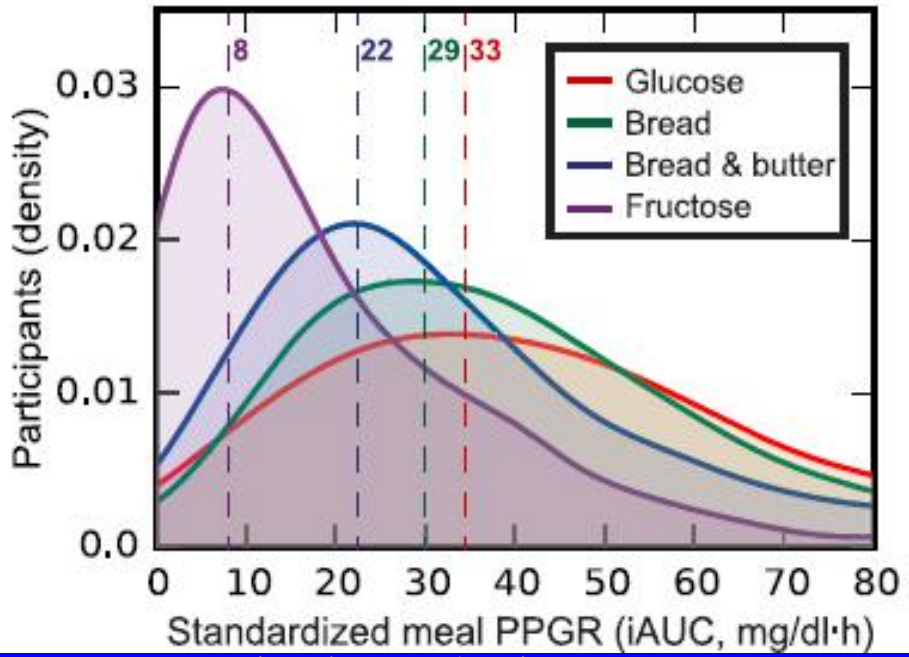
1. 食物多样，谷类为主，粗细搭配
2. 多吃蔬菜水果和薯类
3. 每天吃奶类、大豆或其制品
4. 常吃适量的鱼、禽、蛋和瘦肉
5. 减少烹调油用量，吃清淡少盐膳食
6. 食不过量，天天运动，保持健康体重
7. 三餐分配要合理，零食要适当
8. 每天足量饮水，合理选择饮料
9. 如饮酒应限量
10. 吃新鲜卫生的食物

个体化血糖指数的研究

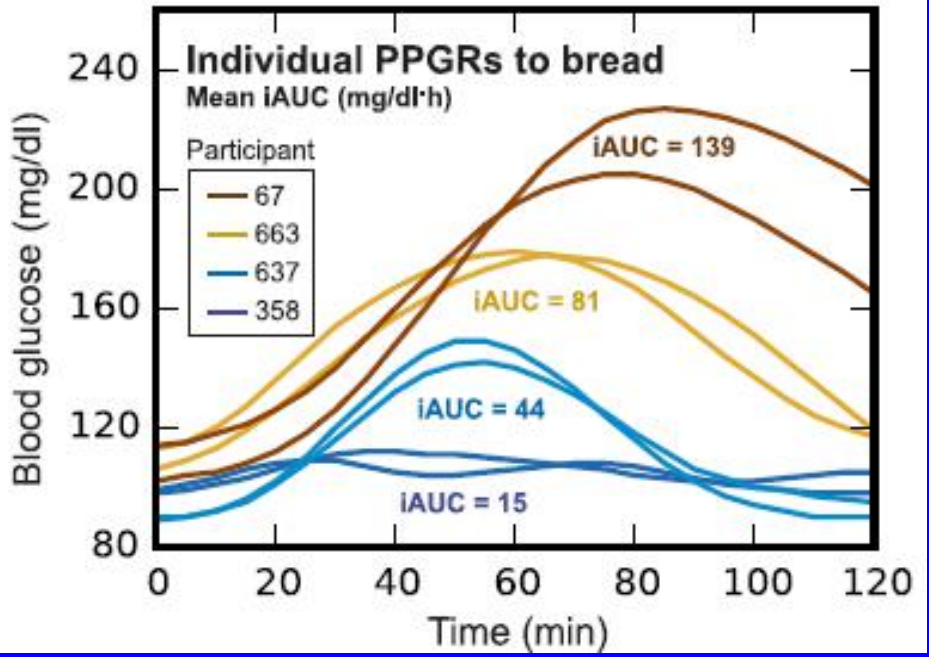


Cell 2015, 163: 1079

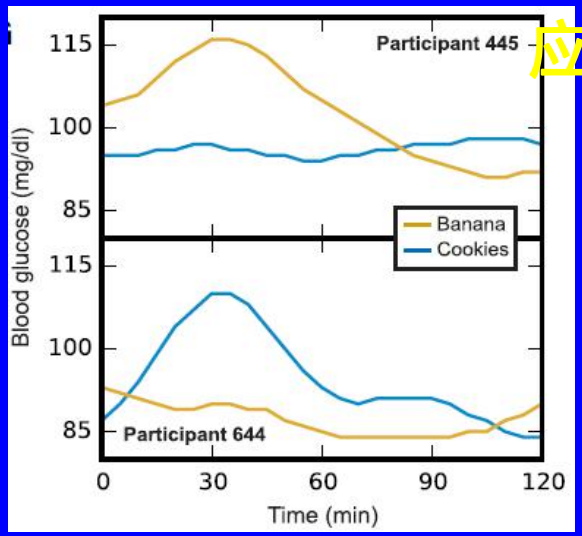
不同个体对食物的血糖应答水平不一样



对不同食物的响应



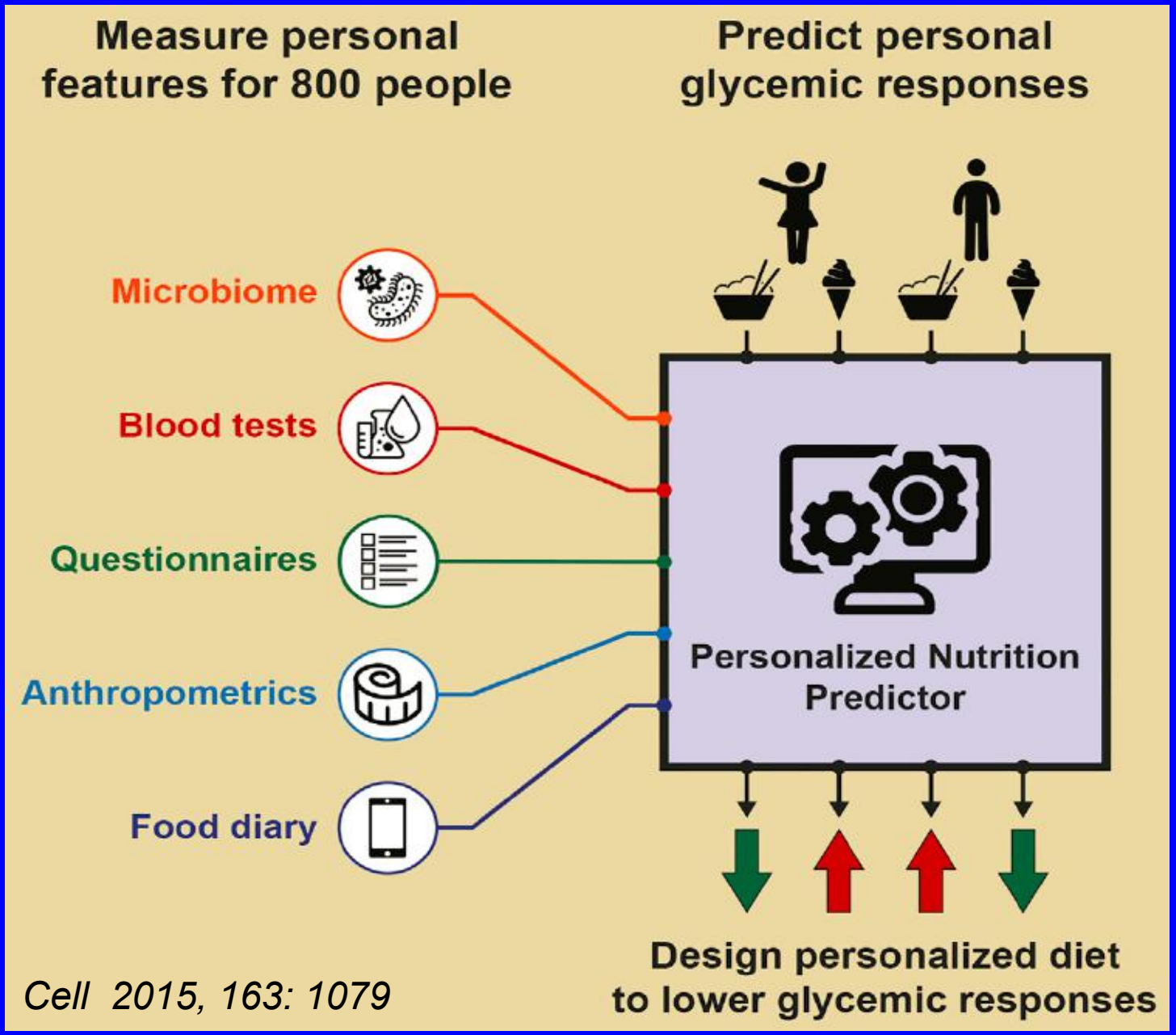
对不同时间的响应



应 两个体的食物响

Cell 2015, 163: 1079

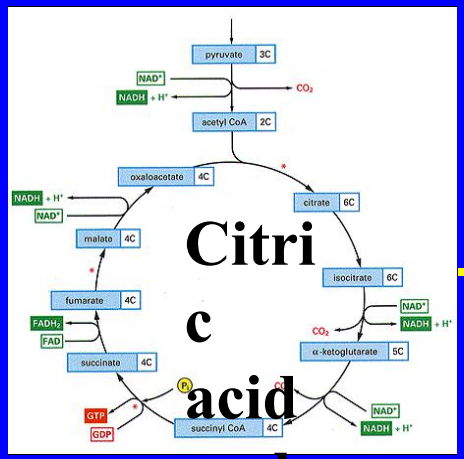
开展个体化营养干预



Cell 2015, 163: 1079

基于系统生物学的精确医学

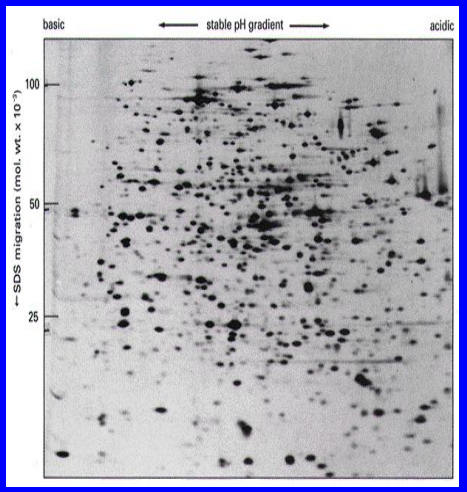
系统生物学：整合各种生物分子数据



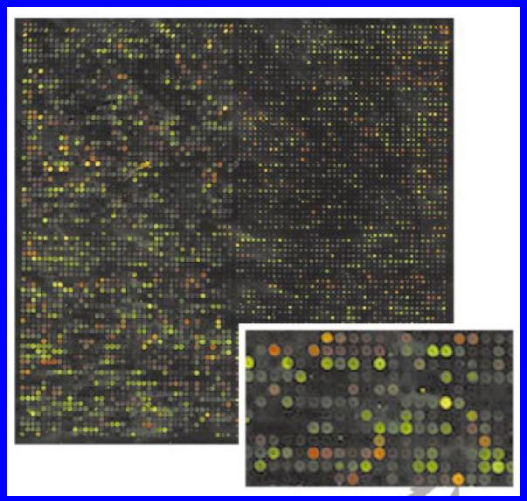
Citric acid cycle

Metabolome

Proteome



Transcriptome

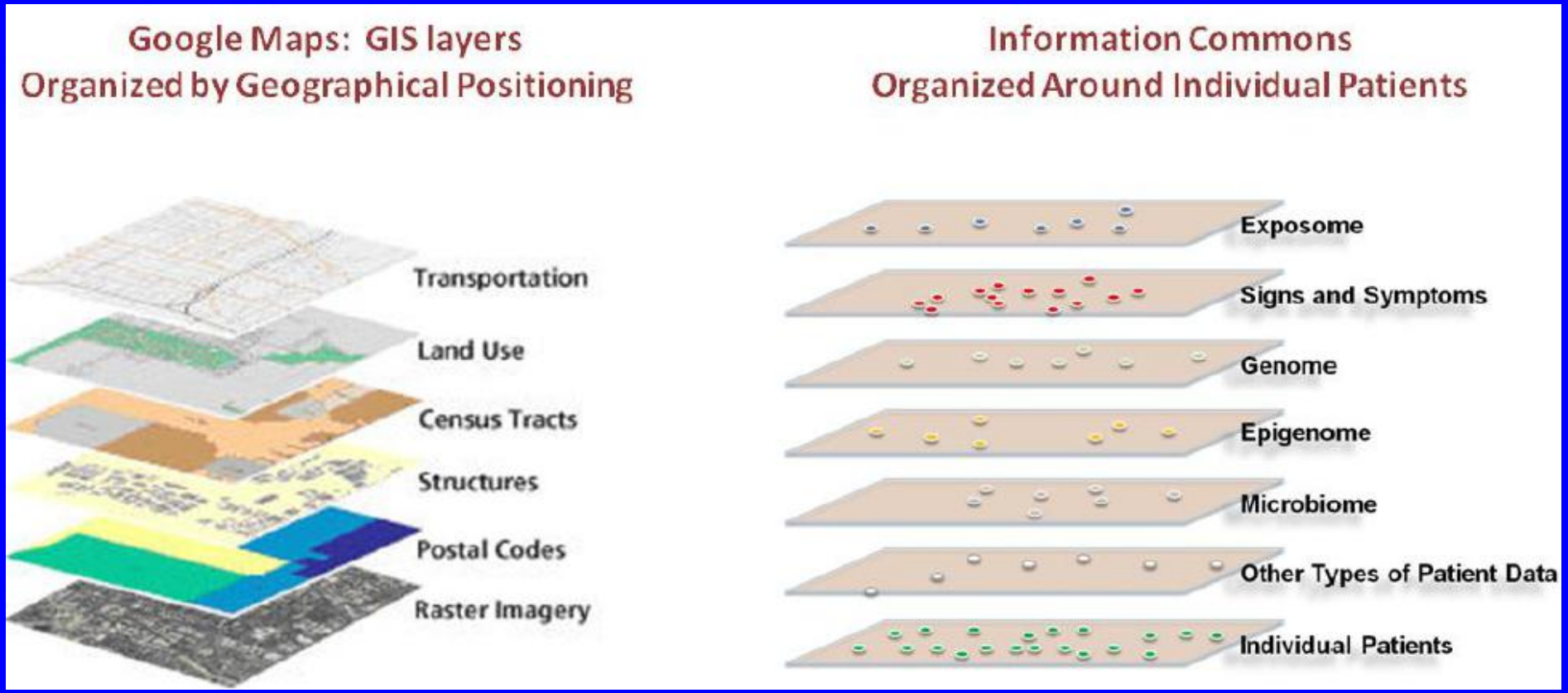


Genome

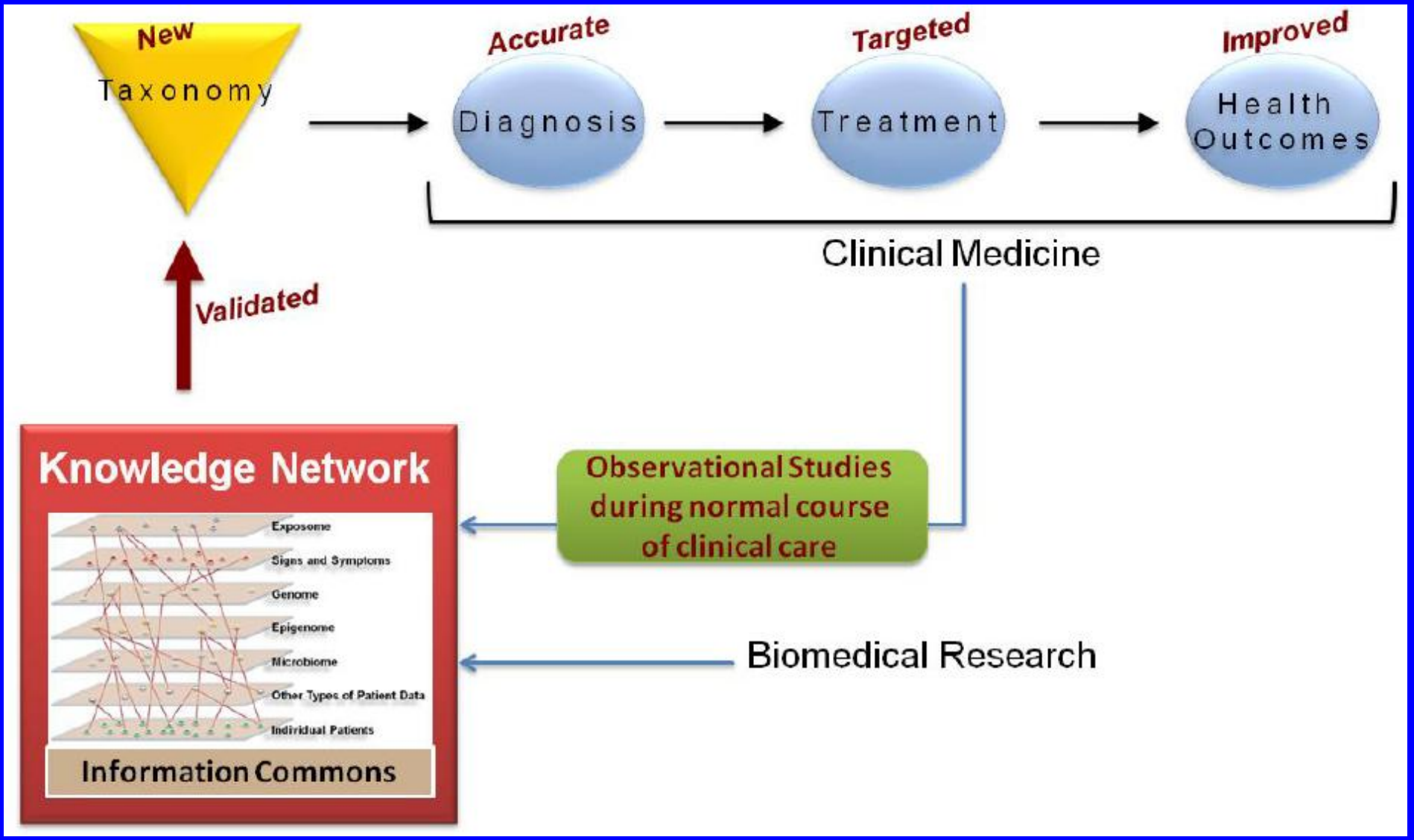
Precision Medicine

National Research Council (2011-11)

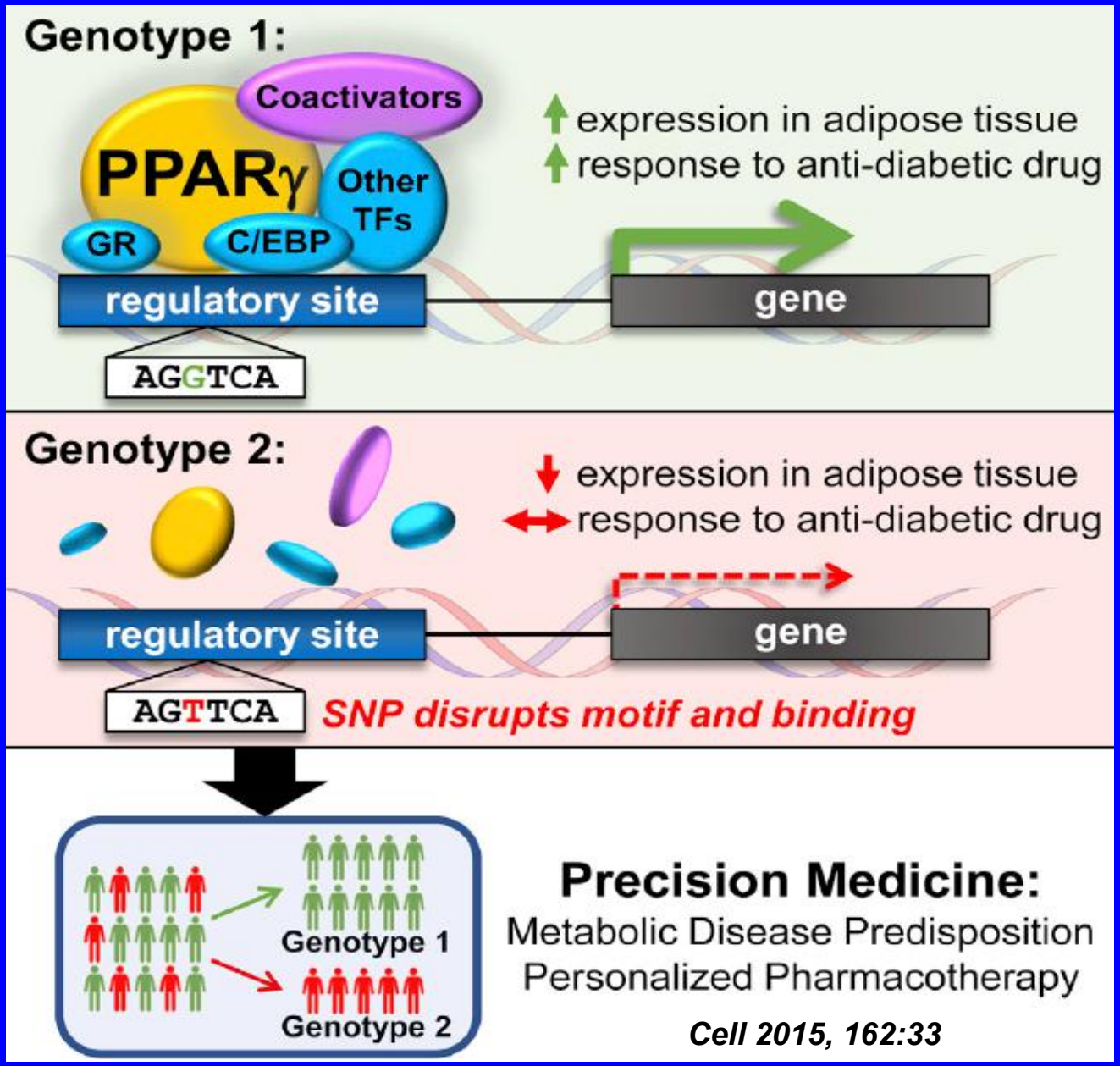
Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease



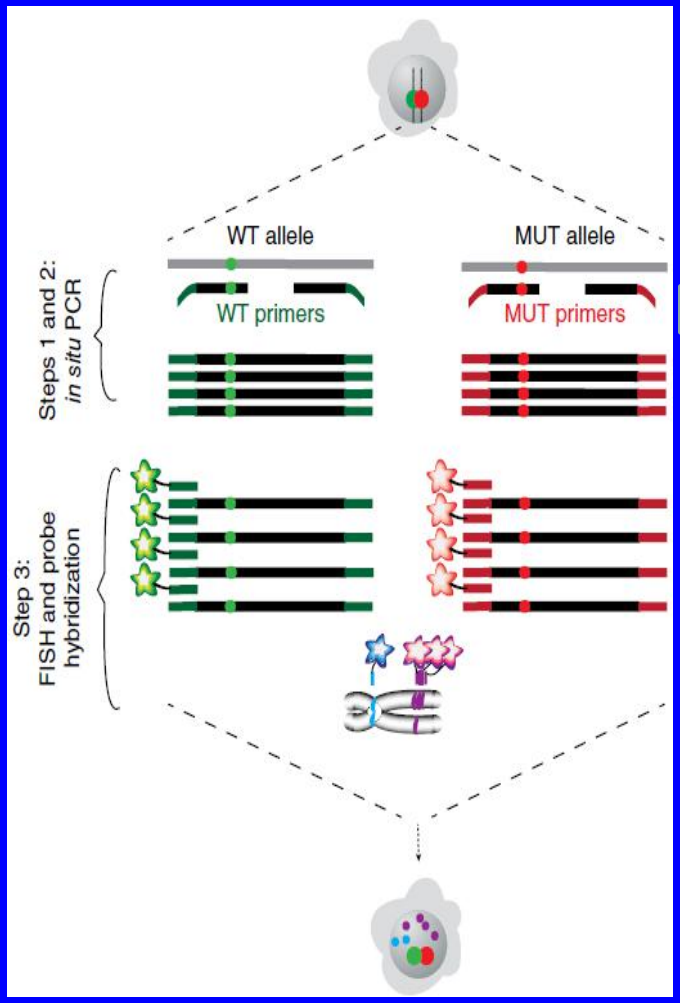
基于生物组学数据整合的精确医学路径



精准医学: 找到一把正确的“钥匙”？



“精确”医学不是“精准”医学

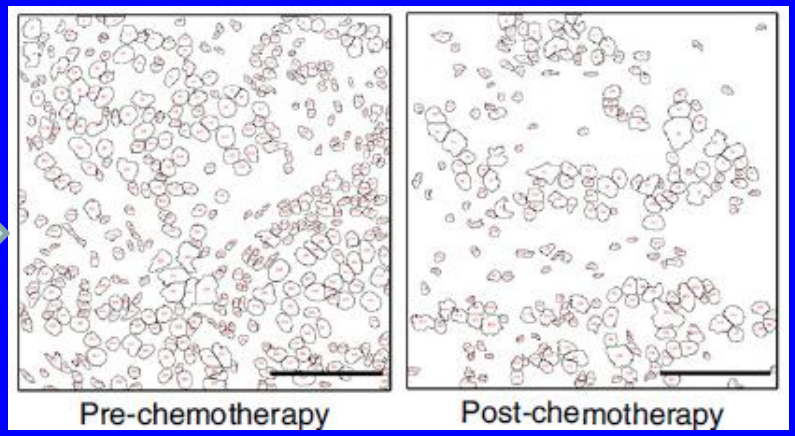


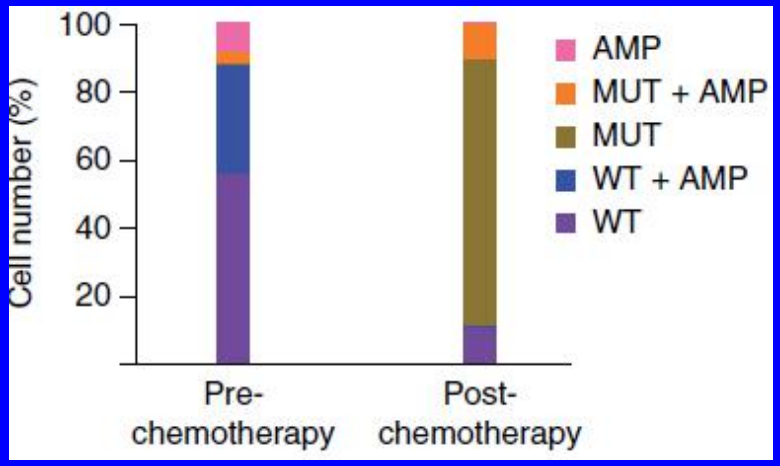
单细胞原位分析技术
(specific-to-allele PCR–FISH)

Nature Genetics 2015-8-24

一个乳腺癌组织样本上各单细胞*PIK3CA*和*HER2*

单细胞原位分析信号

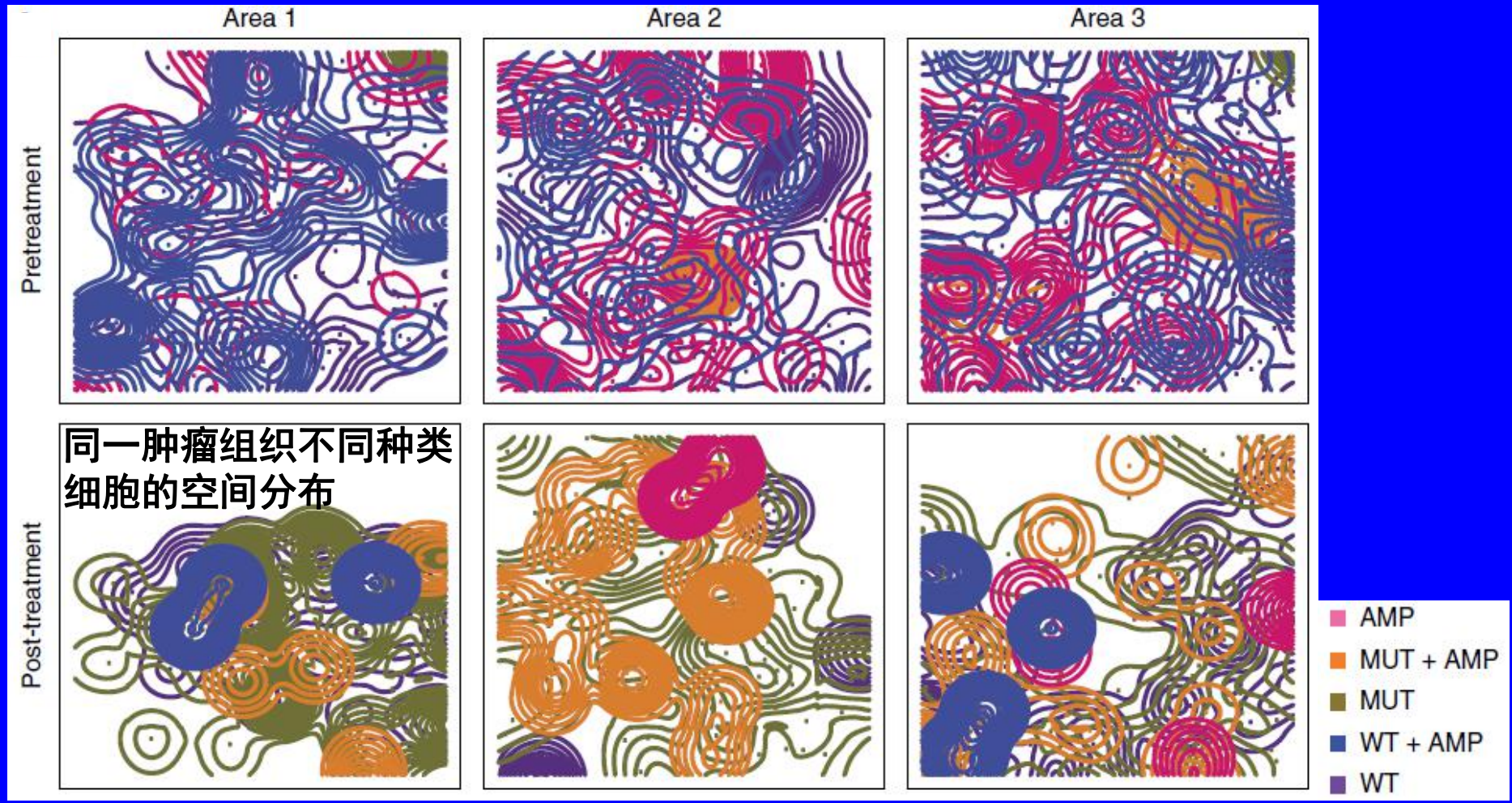




***PIK3CA*突变和*HER2*扩增发生在同一组织不同细胞**

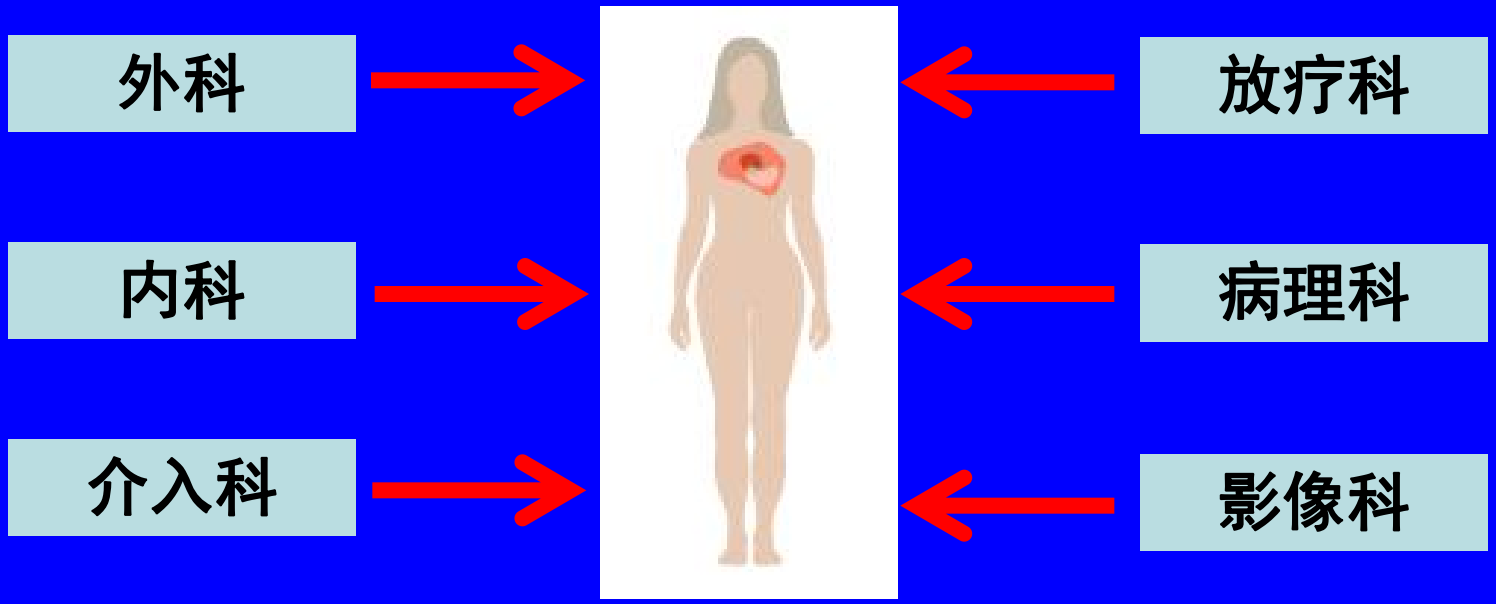
精确医学需要进行全局性分析和判断

- HER2靶向治疗会导致肿瘤组织的细胞类型分布的变化而引起抗药性

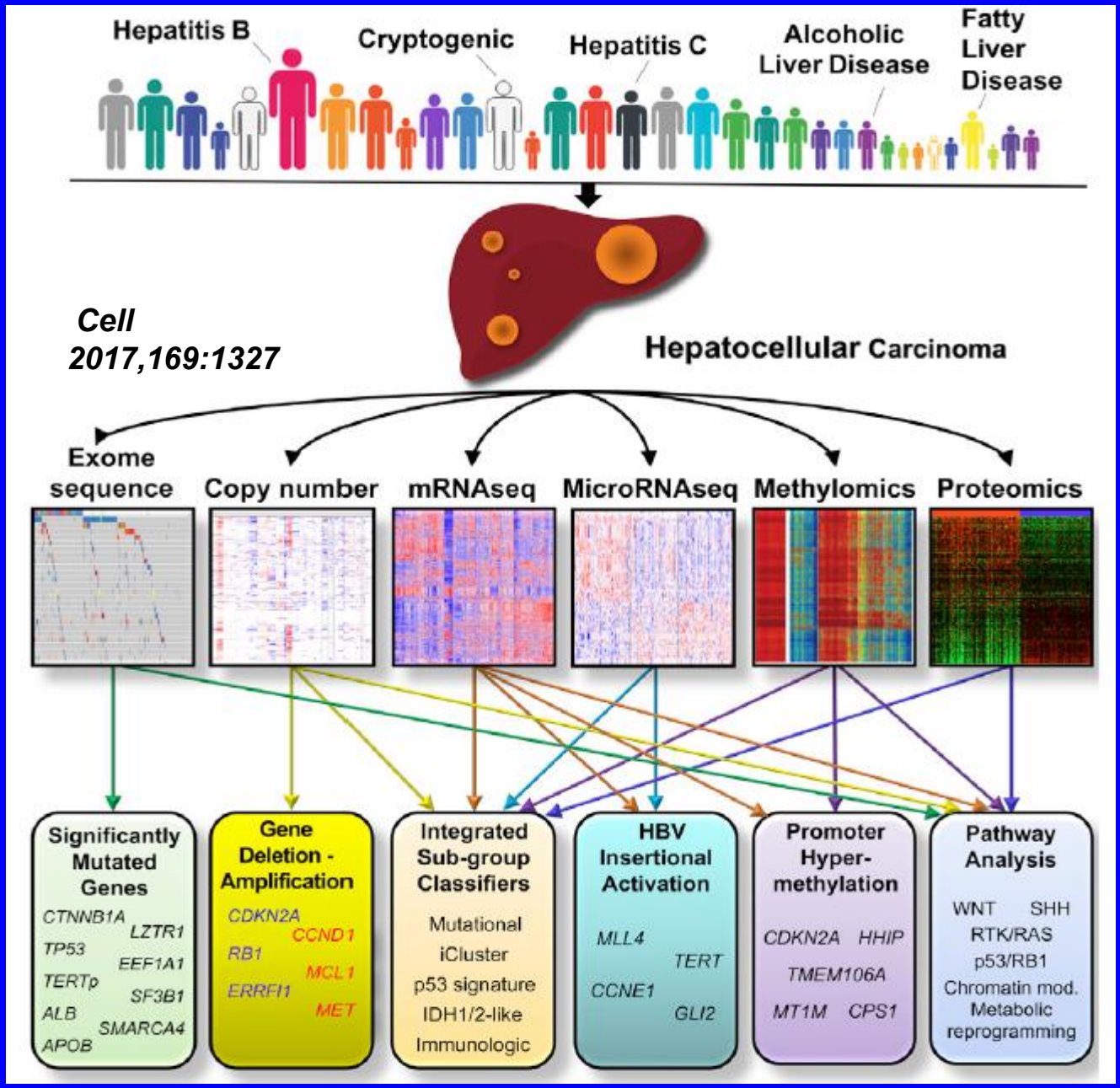


精确医学的起源

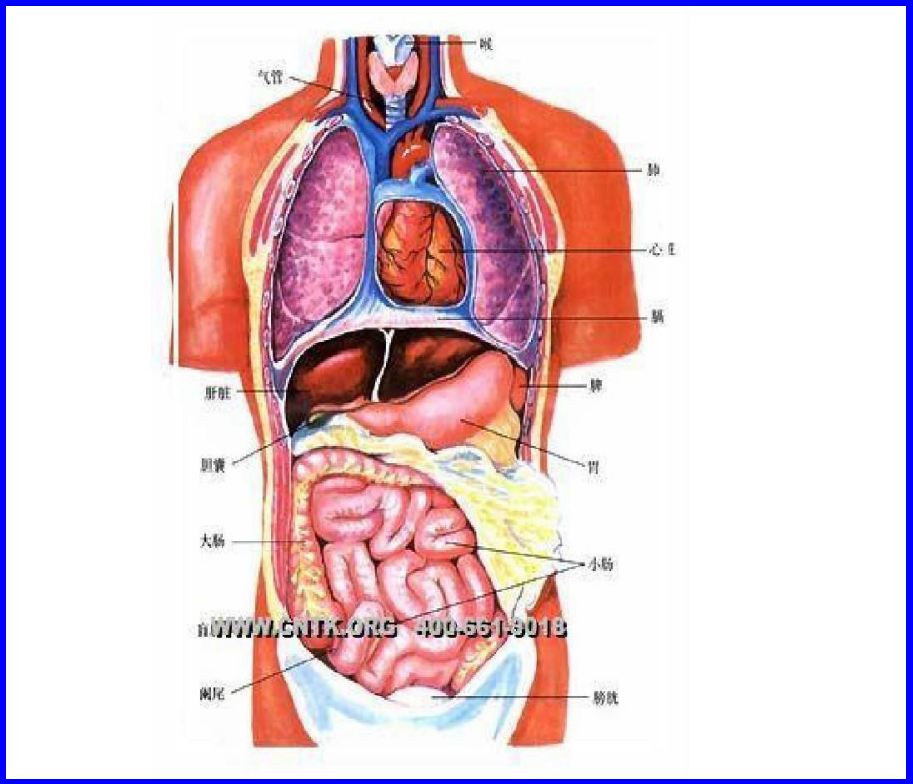
多学科团队诊断模式（Multidisciplinary team, MDT)



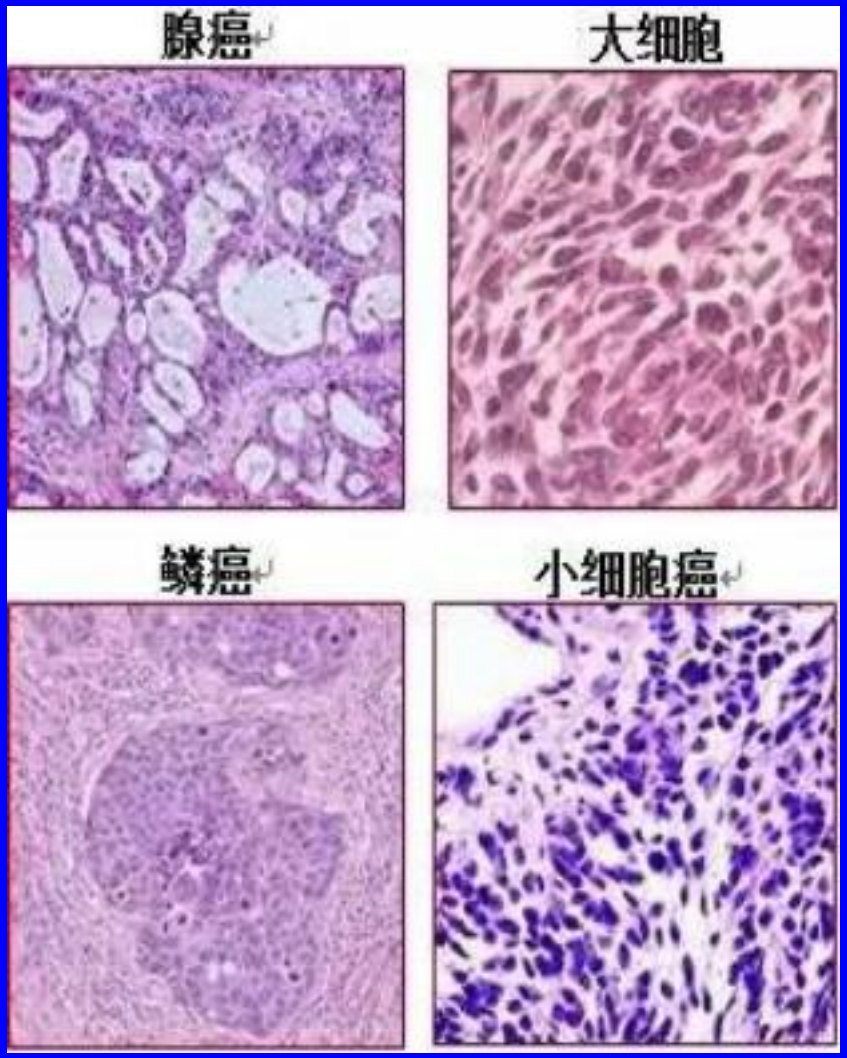
肿瘤的多组学整合型研究



经典的疾病分类标准



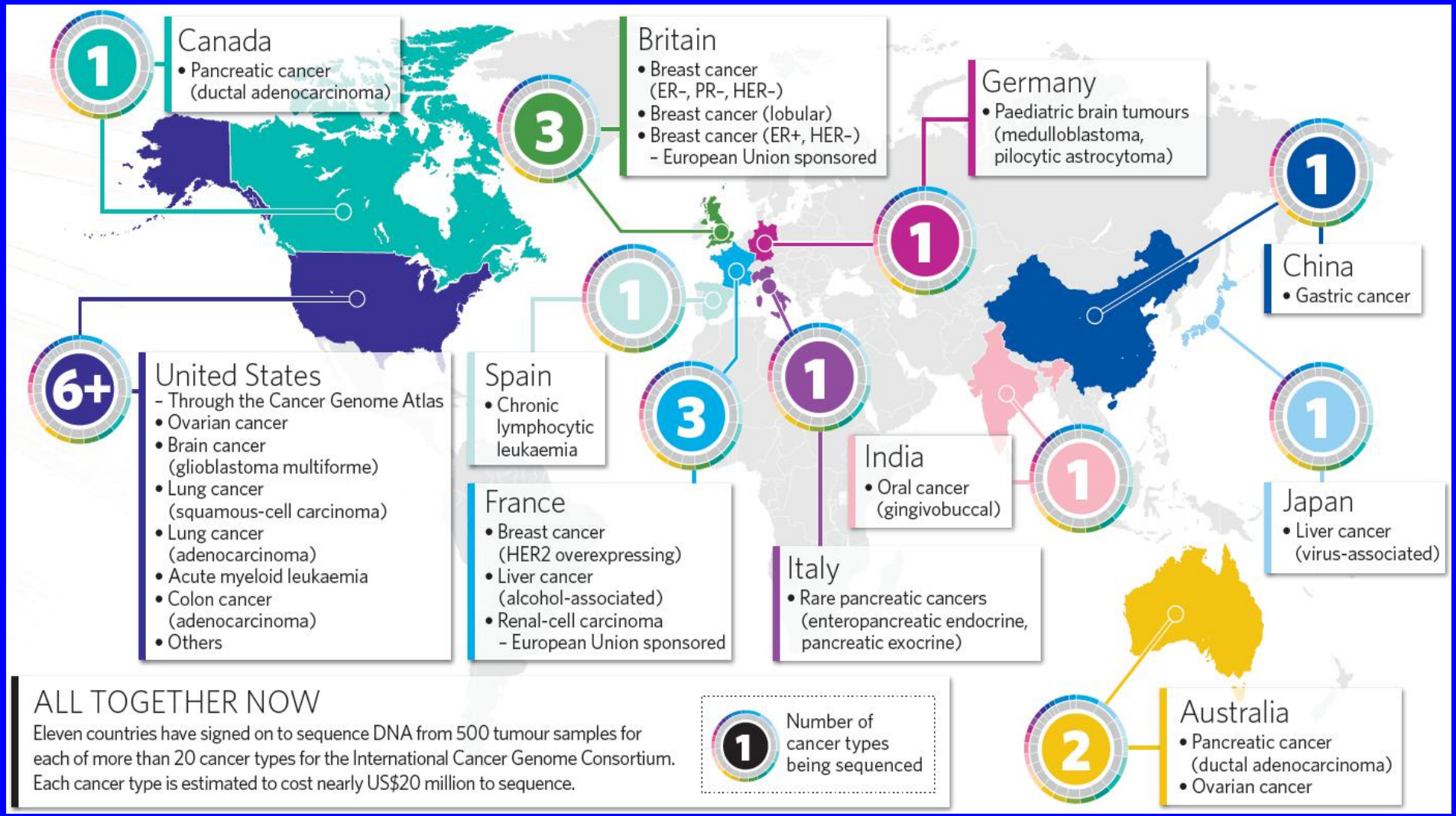
+



解剖部位

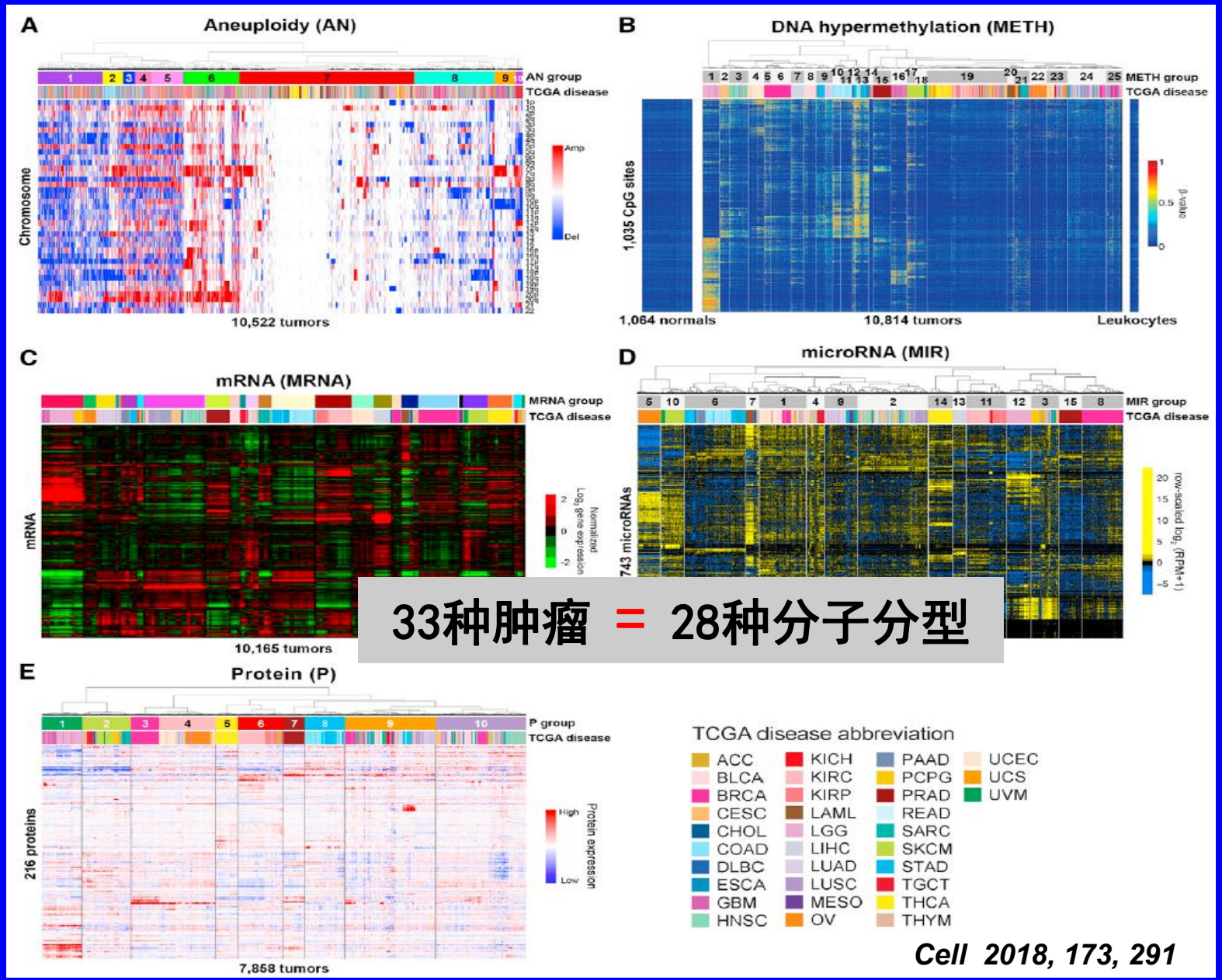
病理症状

The Cancer Genome Atlas (TCGA)



- 目前的肿瘤基因组数据：20 PB (10^{15} bytes)

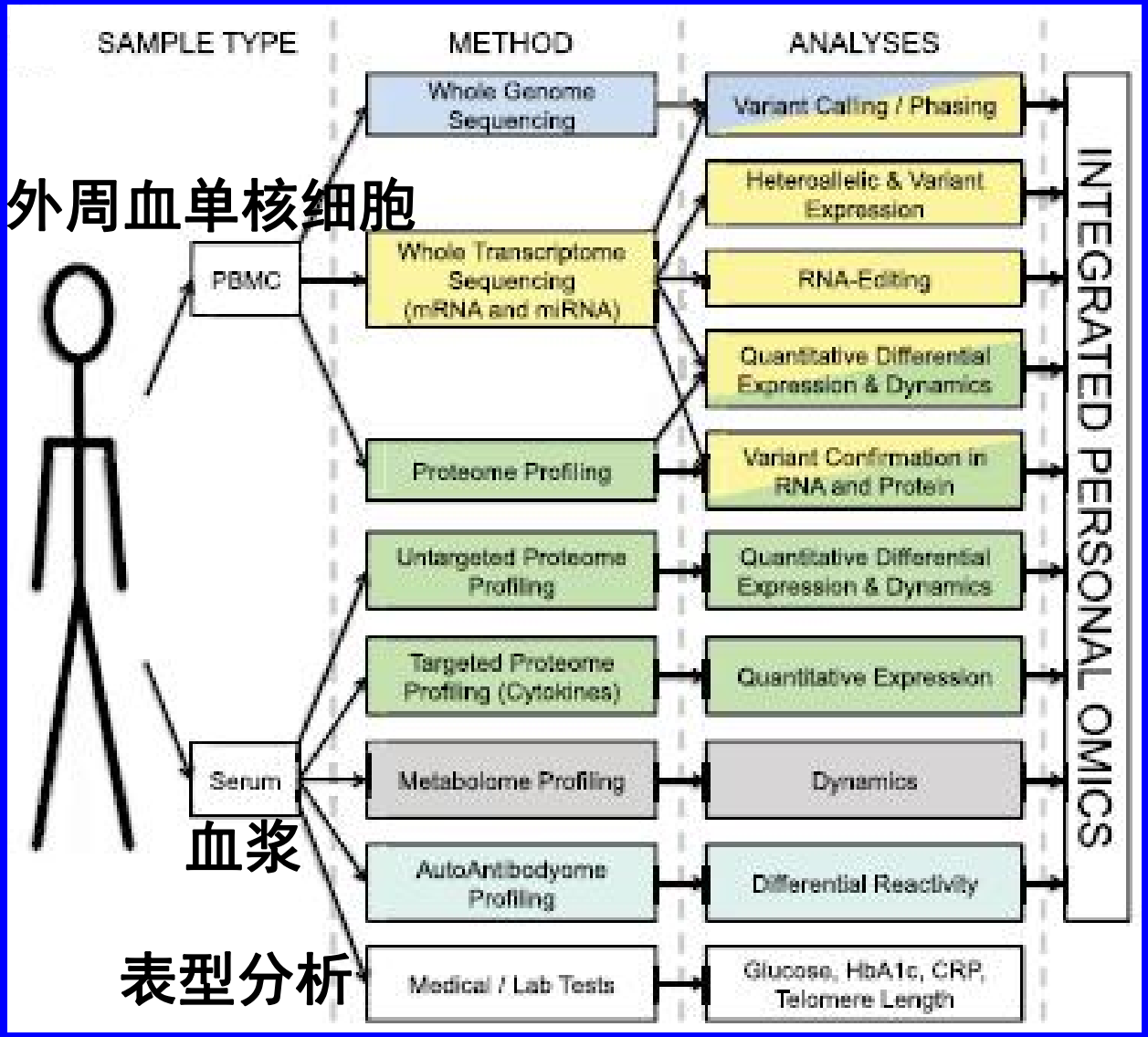
基于多组学整合的“跨肿瘤”的分子分型



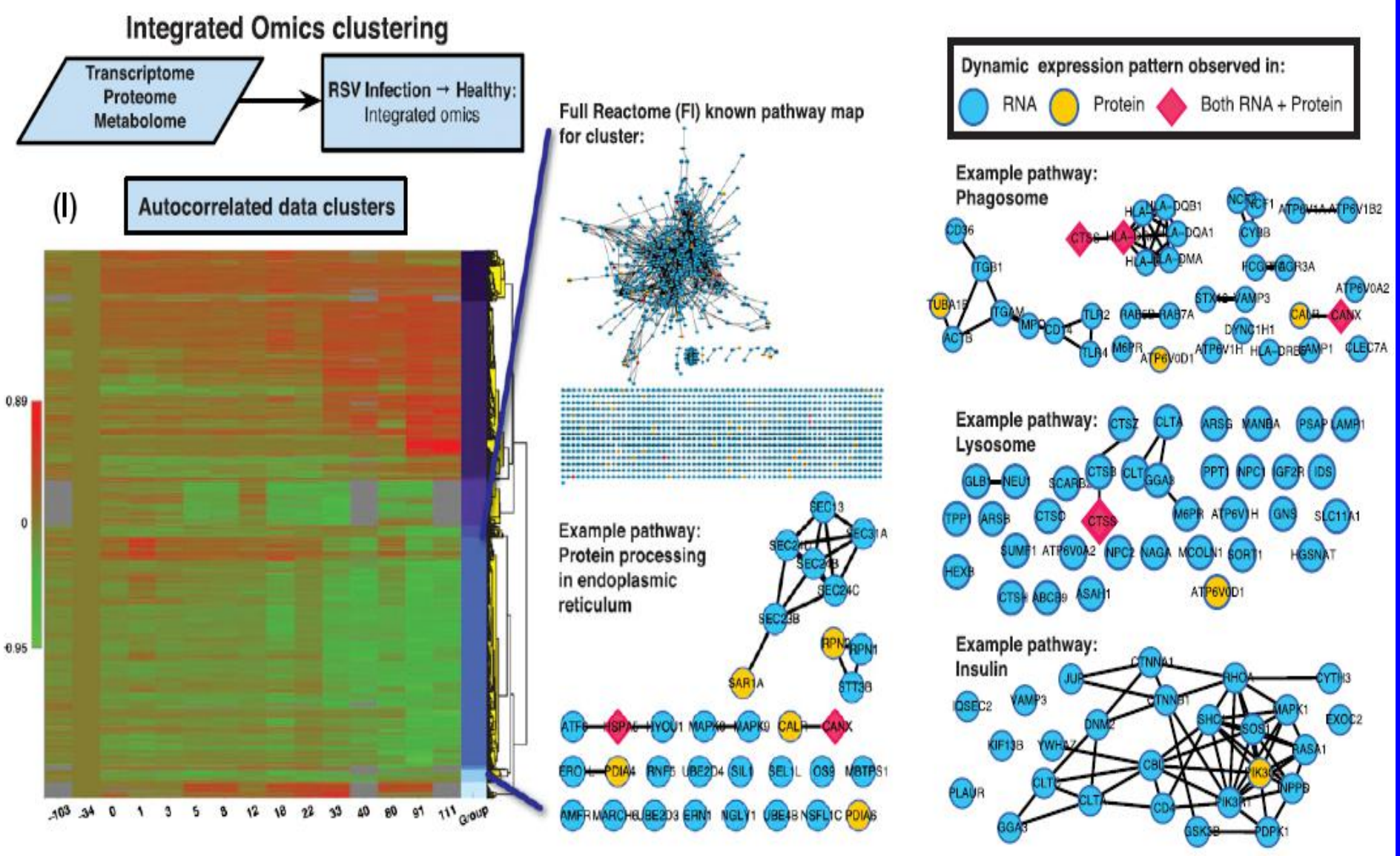
Cell 2018, 173, 291

Personal Omics Profiling Reveals Dynamic Molecular and Medical Phenotypes

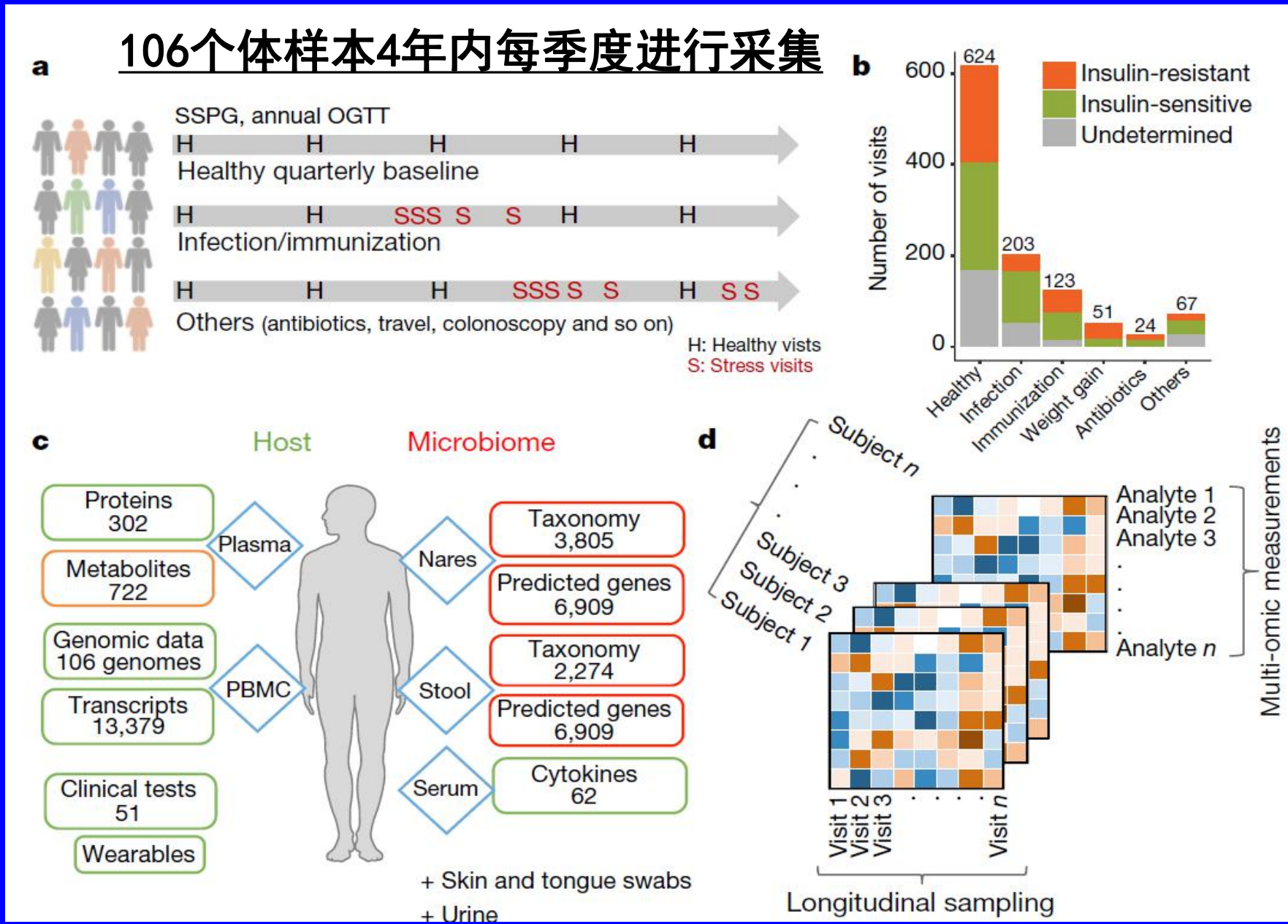
Cell 2012, 148:1293



Integrating Dynamical Data for 400 Days

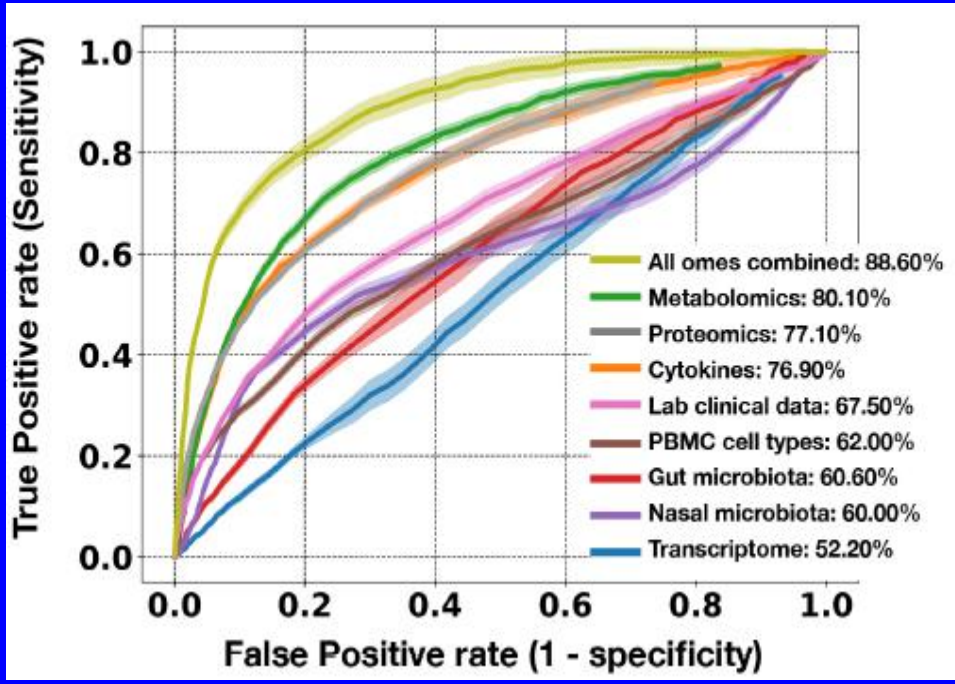


个体的多组学研究策略

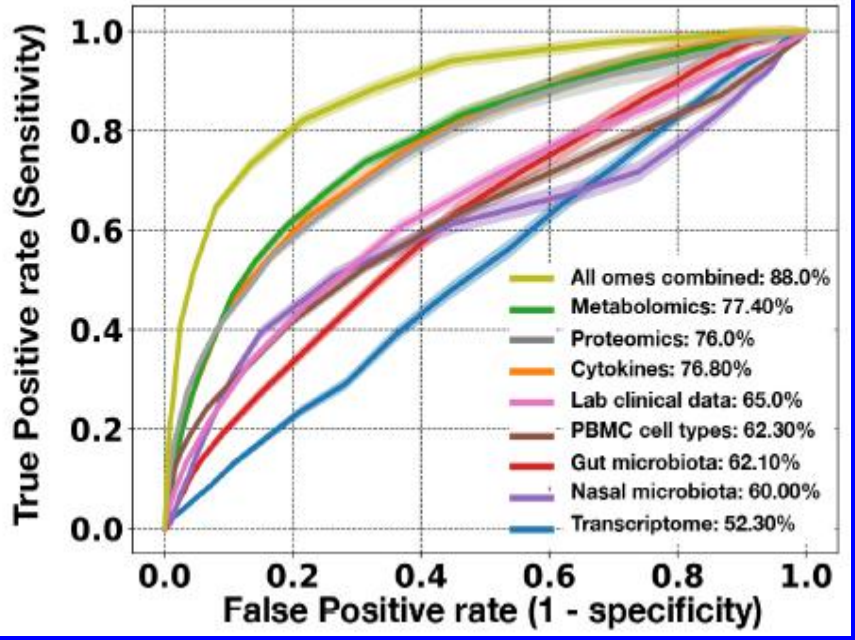

a
106个体样本4年内每季度进行采集
SSPG, annual OGTT
H H H H H
Healthy quarterly baseline
H H SSS S S H H
Infection/immunization
H H H SSS S S H S S
Others (antibiotics, travel, colonoscopy and so on)
H: Healthy vists
S: Stress visits
b
Number of visits
600
400
200
0
624
203
123
51
24
67
Insulin-resistant
Insulin-sensitive
Undetermined
Healthy
Infection
Immunization
Weight gain
Antibiotics
Others
c
Host
Microbiome
Proteins 302
Metabolites 722
Genomic data 106 genomes
Transcripts 13,379
Clinical tests 51
Wearables
Plasma
PBMC
Nares
Stool
Serum
Taxonomy 3,805
Predicted genes 6,909
Taxonomy 2,274
Predicted genes 6,909
Cytokines 62
+ Skin and tongue swabs
+ Urine
d
Subject n
Subject 3
Subject 2
Subject 1
Analyte 1
Analyte 2
Analyte 3
Analyte n
Multi-omic measurements
Visit 1
Visit 2
Visit 3
Visit n
Longitudinal sampling

利用个体的多组学数据进行疾病预测

区别健康状态和呼吸道病毒感染状态



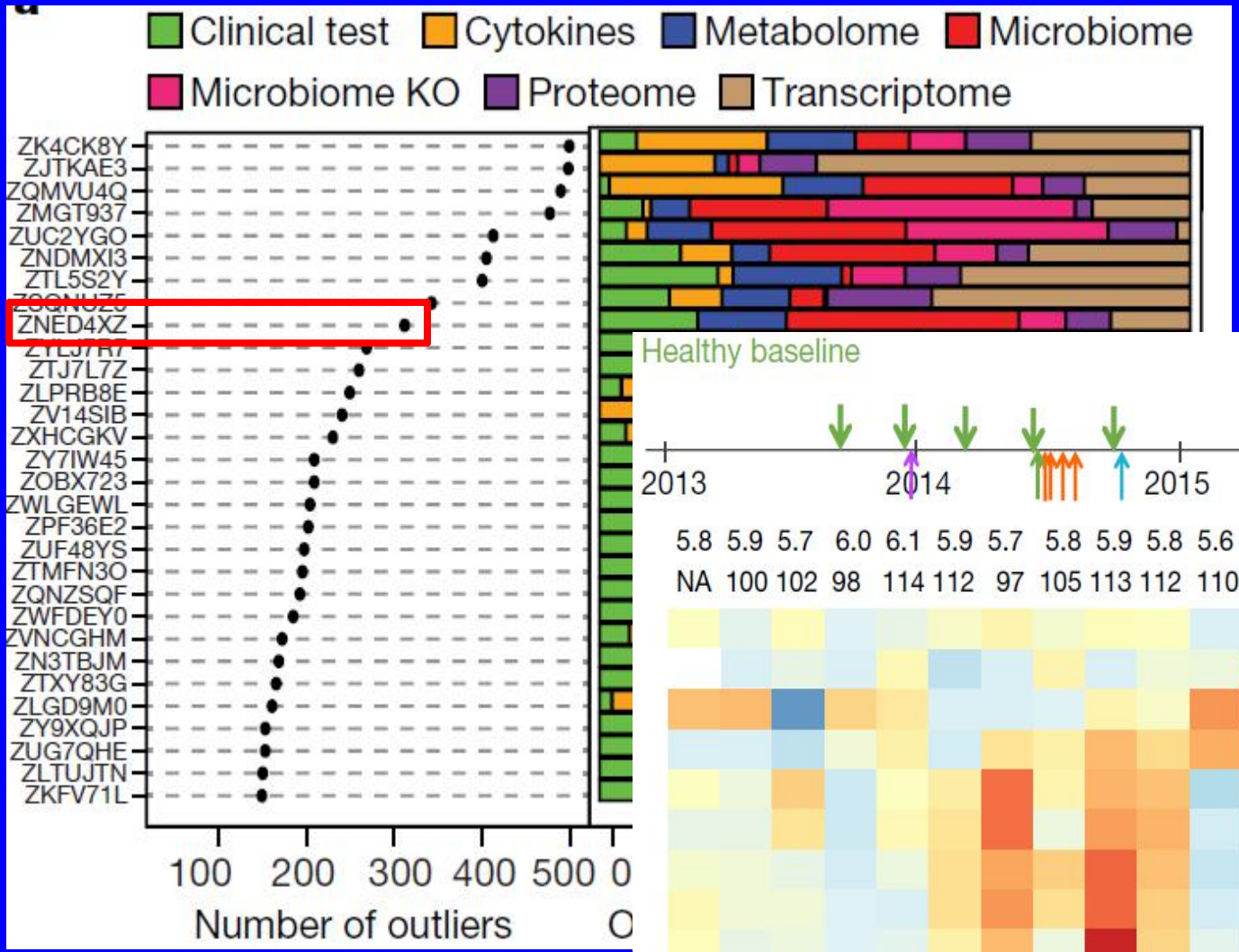
Logistic Regression Model



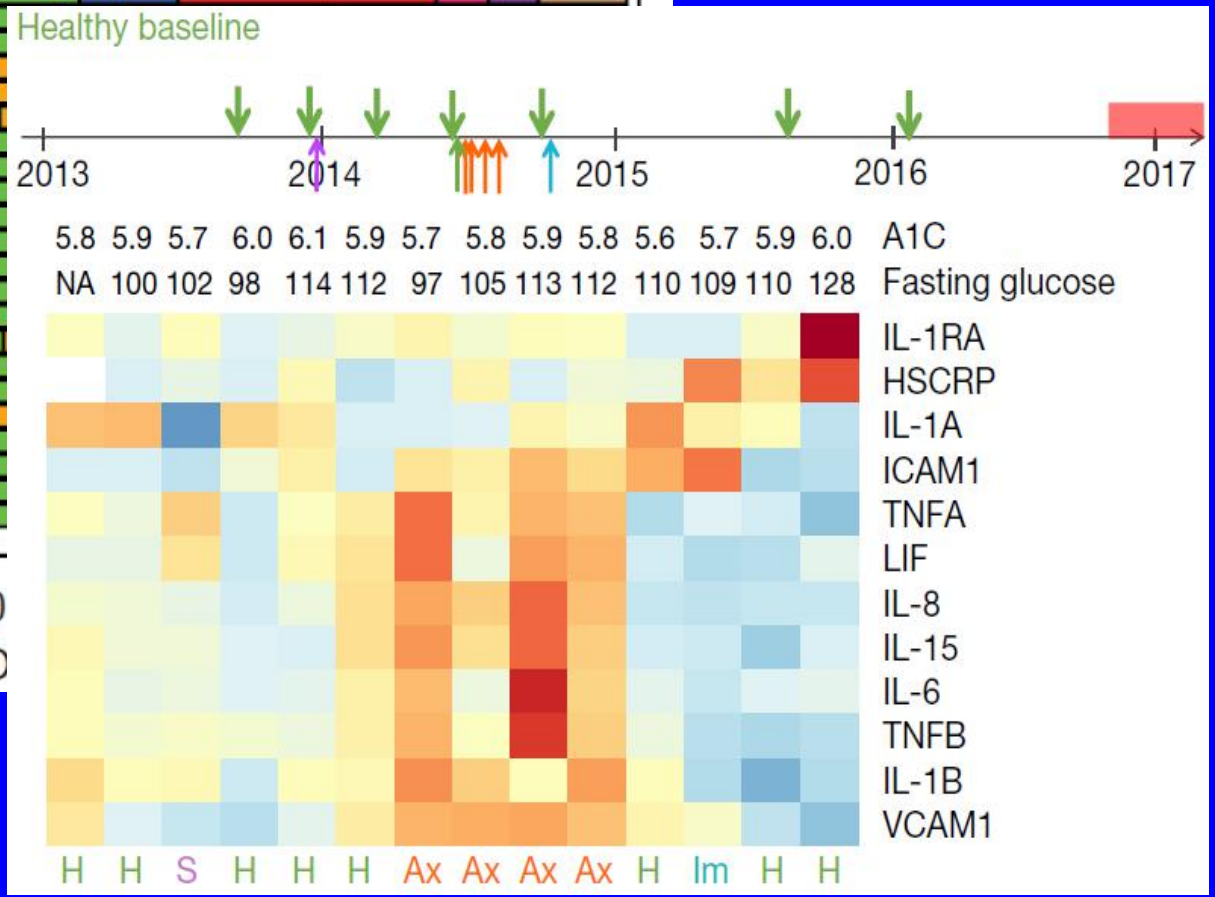
Support Vector Machine Model

- 多组学的准确度相较于单组学稳定，因为对于不同的疾病模型，不同的组学预测能力参差不齐，单组学表现极不稳定，也许在A疾病中代谢组准确度高，在B疾病中蛋白组准确度高。

利用个体的多组学数据进行早期监测



- 该个体被查出多组学指标异常数高达300多个，10个月后被临床诊断为糖尿病



- purple for one stress visit
- red for four antibiotics visits
- blue for a immunization visit

美国精确医学先导队列项目（PMI）

2015年9月17日发布

The Precision Medicine Initiative Cohort Program – Building a Research Foundation for 21st Century Medicine

百万美国志愿者人群

- 个人基本情况与基线调查数据
- 个人健康状况监测数据

（基于可穿戴设备/家用设备）

- 结构化的个人电子健康档案
- 非结构化的个人电子健康档案
- 个人的生物学样本数据
- 个人的社会化数据

美国精确医学先导队列项目（PMI）

Table 2.1: Timeline when expected PMI cohort capabilities will be realized. The estimated timeline for focused research for each type of investigation is indicated by the number of "+" characters in each cell.

基于百万正常人群队列的研究

		Time in years			
		0-2	3-5	5-10	>10
Cohort Capabilities	1. Discovery of disease risk factors	+	+++	+++	++++
	2. Pharmacogenomics	+	+++	+++	+++
	3. Discovery of disease biomarkers	+	++	+++	+++
	4. mHealth connections with disease outcomes		+	++	++++
	5. Impact of loss-of-function mutations		+	+++	+++
	6. New classifications of diseases		+	+++	++++
	7. Empowering participants	+++	+++	+++	+++
	8. Clinical trials of targeted therapies		+	+++	+++

基于系统生物学的精确医学研究

- 多组学
- 大数据
- 长时程

微信公众号：吾家睿见

- 精确医学面面观

谢谢！