

# 泛癌细胞类型特异性基因共表达模式

李晶<sup>1,2</sup>, 鞠苇杭<sup>1,2</sup>, 刘珂<sup>1,2</sup>

(1.山东大学齐鲁医学院公共卫生学院医学数据学系,山东 济南 250012;

2.国家健康医疗大数据研究院,山东 济南 250003)

**摘要:**目的 基于单细胞转录组学数据,探究多种癌症组织中不同细胞类型的基因共表达规律。方法 依托大规模公共单细胞转录组学数据,运用CS-CORE方法为6种常见的细胞类型在5种不同的癌症中分别构建了细胞类型特异性的基因共表达图谱。结果 共获得30张高质量的基因共表达图谱。在所研究的细胞类型中,恶性细胞的共表达基因对数量最多,具有相对复杂的基因调控机制;跨癌症类型的分层聚类结果显示转录因子的调控广度在同一细胞类型中倾向于保守;此外,巨噬细胞具有最多的差异共表达基因对,巨噬细胞的基因调控网络可能易受肿瘤微环境的影响。结论 本研究系统构建了癌症组织中细胞类型特异性的基因共表达图谱,并在泛癌视角下揭示了基因共表达的若干规律。

**关键词:**细胞类型特异性;基因共表达;癌症;单细胞转录组;泛癌

中图分类号:R730.2

文献标志码:A

## Pan-cancer cell type-specific gene-gene co-expression pattern

LI Jing<sup>1,2</sup>, JU Weihang<sup>1,2</sup>, LIU Ke<sup>1,2</sup>

(1. Department of Medical Dataology, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250012, Shandong, China; 2. National Institute of Health and Medicine Big Data, Jinan 250003, Shandong, China)

**Abstract: Objective** To investigate the gene co-expression patterns across diverse cell types within multiple cancer tissues using single-cell omics data. **Methods** Utilizing large-scale public single-cell transcriptomics data and the CS-CORE method, cell type-specific gene co-expression profiles were generated for six common cell types across five distinct cancers. **Results** A total of 30 high-quality gene co-expression maps were obtained. Among the analyzed cell types, malignant cells exhibited the highest number of co-expressed gene pairs, suggesting a relatively complex gene regulatory mechanism. Hierarchical clustering across cancer types revealed that the regulatory breadth of transcription factors tended to be conserved within the same cell type. Additionally, macrophages displayed the most differentially co-expressed gene pairs, suggesting that the gene regulatory network of macrophages might be susceptible to the influence of the tumor microenvironment. **Conclusion** This study systematically constructs a cell type-specific gene co-expression map in cancer tissues, uncovering several gene co-expression patterns from a pan-cancer perspective.

**Key words:** Cell type specificity; Gene co-expression; Cancer; Single-cell transcriptome; Pan-cancer

基因共表达模式在多种疾病分子机制的探索以及新型治疗靶点的发现方面发挥着重要作用<sup>[1-4]</sup>。目前,大多数此类研究基于批量样本的基因表达谱来开展。在批量样本中,细胞类型的构成差异是导致基因表达变化的主要因素之一,因此基因共表达值常常会受到细胞类型比例变化的

显著影响<sup>[5-6]</sup>;此外,这种基于批量样本的分析只能揭示样本整体的共表达模式,而无法解释研究样本中某些特定细胞类型的基因共表达规律。因此,利用单一细胞类型的转录组学数据开展基因共表达模式研究能够更准确地揭示细胞类型特异性的基因调控图谱<sup>[7]</sup>。

单细胞 RNA 测序技术能够在单个细胞层面上较为准确地测量基因表达水平, 这为构建细胞类型特异性的基因共表达图谱提供了机遇<sup>[8]</sup>; 然而, 单细胞 RNA 测序数据稀疏性强、噪音大的特点为准确捕获基因共表达网络带来了挑战<sup>[9-10]</sup>。目前, 研究人员已开发出多种方法用于从单细胞 RNA 测序数据中提取基因共表达模式<sup>[11]</sup>。其中, CS-CORE 方法通过显式建模单细胞 RNA 测序数据中的测序深度和测量误差的方式, 显著降低了假阳性率和估计偏差, 在模拟数据集和真实数据集上皆显示出良好的性能<sup>[12]</sup>。

目前, 关于癌症组织中特定细胞类型的基因共表达模式的研究较少。基于此, 本研究应用 CS-CORE 方法, 对乳腺癌、肝癌、卵巢癌、胰腺癌、肾癌 5 种癌症的单细胞 RNA 测序数据集进行分析, 构建 6 种细胞类型(T 细胞、B 细胞、成纤维细胞、内皮细胞、巨噬细胞、上皮细胞)在这 5 种癌症中的基因共表达图谱, 系统解析肿瘤组织中细胞类型特异性的基因共表达规律。

## 1 资料与方法

### 1.1 资料

本研究的单细胞转录组数据来自 Curated Cancer Cell Atlas 数据库 (<https://www.weizmann.ac.il/sites/3CA/>)<sup>[13]</sup>, 从其中纳入了来自乳腺癌、肝癌、卵巢癌、胰腺癌、肾癌 5 种人类原发癌及其癌旁正常组织的样本(所选择的样本数据均由 10×Genomics 测序平台生成并以 read counts 形式提供)。最终, 共筛选获得了 26 个数据集, 共 503 个样本(表 1)。人类转录因子列表来源于 AnimalTFDB v4.0 数据库 (<https://guolab.wchscu.cn/AnimalTFDB4/>)<sup>[14]</sup>。代谢通路和信号通路来源于京都基因与基因组百科全书(Kyoto encyclopedia of genes and genomes, KEGG)<sup>[15]</sup>, 通路基因集下载自 Enrichr 平台 (<https://maayanlab.cloud/Enrichr/>)<sup>[16-18]</sup>。

表 1 不同组织的数据集及样本数量汇总  
Table 1 Summary of datasets and sample sizes of the different tissues

组织类型	数据集数量	样本数
乳腺	8	209
肝脏	4	79
卵巢	5	21
胰腺	4	85
肾脏	5	109

### 1.2 方法

#### 1.2.1 细胞类型特异性基因共表达图谱的构建

任意给出一个癌症 T 的单细胞转录组数据集 D、一种细胞类型 C 以及一个基因对 P, 本研究使用符号  $\text{co-expr}(T, P, C, D)$  来表示使用数据集 D 计算的基因对 P 在细胞类型 C 中的共表达值。如果某种癌症有  $n$  个数据集( $D_1, D_2, \dots, D_n$ ), 那么基因对 P 在癌症 T 的细胞类型 C 中的共表达值  $\text{co-expr}(T, P, C)$  定义为:

$$\text{co-expr}(T, P, C) = \text{median}\{\text{co-expr}(T, P, C, D_i)\} \\ i = 1, 2, \dots, n。$$

使用如下步骤计算  $\text{co-expr}(T, P, C, D)$ 。

步骤 1: 将类型为 C 的细胞从数据集 D 的各个样本中提取出来并进行合并, 随后进行基因过滤; 若某一基因在 C 中的表达比率低于 10%, 则将此基因剔除。

步骤 2: 将 CS-CORE 方法应用于步骤 1 所提取的数据, 计算所有基因对在细胞类型 C 中的共表达值及其对应的 P 值。为了控制错误发现率(false discovery rate, FDR), 对 P 值进行 Benjamini-Hochberg (BH) 校正。如果一个基因对校正后的 P 值  $> 0.05$ , 则认为该基因对中的两个基因不具有共表达关系, 并将其共表达值设为 0<sup>[11]</sup>。

通过以上的计算, 即可以得到在癌症类型 T 的细胞类型 C 中的基因共表达图谱。

任意给出一张基因共表达图谱, 计算其中基因对共表达值的第一四分位数  $Q_1$ , 第三四分位数  $Q_3$ , 以及四分位数间距 IQR。共表达值高于  $Q_3 + 1.5 \times \text{IQR}$  (或低于  $Q_1 - 1.5 \times \text{IQR}$ ) 的基因对被定义为可靠的共表达基因对(图 1)。

#### 1.2.2 转录因子调控广度的计算

任意给出一个转录因子以及一张由“癌症类型-细胞类型”组合所确定的基因共表达谱, 统计该转录因子所参与的共表达基因对的个数, 并将其定义为该转录因子在此“癌症类型-细胞类型”中的调控广度。移除在所有“癌症类型-细胞类型”组合中调控广度均为 0 的转录因子后, 得到一个包含 1 001 个转录因子的调控广度矩阵。

#### 1.2.3 高调控广度转录因子的筛选

本研究将在至少 80% “癌症类型-细胞类型”组合中呈现出较高调控广度[标准化后的  $\log_2$ (调控广度) 值大于 0] 的转录因子定义为高调控广度转录因子, 具体的筛选步骤如下。

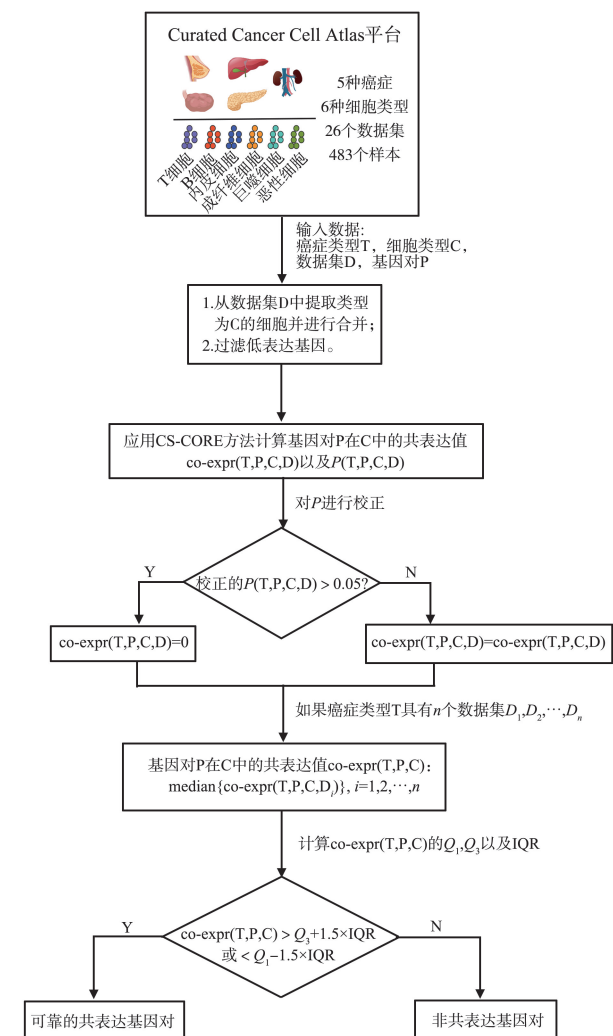


图1 识别可靠的细胞类型特异性共表达基因对的流程图  
Figure 1 Workflow for identifying reliable cell-type-specific co-expressed gene pairs

步骤1:以每个转录因子在30个“癌症类型-细胞类型”组合中的调控广度为特征数据,对1001个转录因子进行层次聚类分析,生成层次聚类树状图,并使用热图可视化这1001个转录因子在30个“癌症类型-细胞类型”组合中的调控广度分布。每一行表示一个转录因子,每一列表示一个“癌症类型-细胞类型”组合。热图的数值为转录因子的调控广度(经过 $\log_2$ 变换,并按列进行标准化)。

步骤2:将层次聚类树状图剪枝为10个类别,观察并对比热图中各类别的调控广度分布,从中选择在至少80%的“癌症类型-细胞类型”组合中均呈现一致性的较高调控广度的类别。

最终,第1,2,4,5和7类转录因子表现出较高的调控广度和一致性,将其定义为高调控广度转录因子。

1.2.4 “癌症类型-细胞类型”组合中代谢和信号通路的平均基因共表达水平的计算及其可视化  
任意给出一条代谢通路或信号通路以及一张由

“癌症类型-细胞类型”组合所确定的基因共表达谱,计算该通路上所有基因对的平均共表达值,并将此定义为该通路在此“癌症类型-细胞类型”中的基因共表达水平。使用热图可视化这些通路在30个“癌症类型-细胞类型”组合中的基因共表达水平。每一行表示一个代谢(或信号)通路,每一列表示一个“癌症类型-细胞类型”组合。热图的数值为代谢通路上的基因对的平均共表达值。

1.2.5 癌症组织和正常组织中细胞类型特异性差异共表达基因对的获得

本研究为乳腺癌、肝癌、卵巢癌、肾癌相应的正常组织(癌旁对照)生成了细胞类型特异性的基因共表达谱。

任意给出一种细胞类型C和一种癌症T,按照下列步骤获取在癌症组织和癌旁正常组织的细胞类型C中的差异共表达基因对。

步骤1:参照1.2.1部分,构建与癌症T匹配的癌旁正常组织N的细胞类型C的基因共表达图谱。

步骤2:计算在癌症组织T和癌旁正常组织N的细胞类型C中每个基因对共表达值的差值。

步骤3:对于细胞类型C,计算所有基因对的差异共表达值的第一四分位数 $Q_1$ ,第三四分位数 $Q_3$ 以及四分位数间距IQR。差异共表达值高于 $Q_3+1.5\times IQR$ 或低于 $Q_1-1.5\times IQR$ 的基因对,被定义为在癌症组织T和癌旁正常组织N的细胞类型C中具有显著差异共表达的基因对。

对于每一种细胞类型,进一步将至少在两个癌症中可以被鉴定的差异共表达基因对定义为该细胞类型的高度可信差异共表达基因对。

1.2.6 软件工具

所有的分析都在R 4.2.3中进行。Seurat 4.3.0包用于处理单细胞RNA测序数据<sup>[19]</sup>;ggplot2 3.5.0包用于绘制散点图<sup>[20]</sup>;ComplexHeatmap 2.14.0包用于绘制热图<sup>[21]</sup>;clusterProfiler 4.6.2包用于KEGG通路富集分析<sup>[22]</sup>。数据分析所用代码保存于[https://github.com/bioklab/proj\\_coexpression](https://github.com/bioklab/proj_coexpression)。

1.3 统计学处理

采用R 4.2.3软件进行统计分析,采用两因素方差分析研究癌症类型和细胞类型对共表达基因对数的影响,检验水准 $\alpha=0.05$ 。

## 2 结果

### 2.1 泛癌细胞类型特异性基因共表达图谱的构建

6种细胞(T细胞、B细胞、成纤维细胞、内皮细

胞、巨噬细胞和恶性细胞)在5种癌症中共构建了30张基因共表达图谱,即有30个不同的“癌症类型-细胞类型”组合。

共表达基因对数在不同的情境下存在较大的变异:最低值为3 025(肝癌-B细胞),最高值为14 163 504(卵巢癌-恶性细胞)。双因素方差分析结果显示,癌症类型( $P=0.002$ )和细胞类型( $P=0.008$ )均对共表达基因对数有影响。对所得到的基因共表达图谱按照共表达基因对数进行排序,发现卵巢癌、胰腺癌、肾癌、乳腺癌的恶性细胞排在前4位(图2);与排序结果一致,在所有细胞类型中,恶性细胞拥有最高的平均共表达基因对数,提示恶性细胞相较于其他微环境细胞类型存在着更为复杂的基因调控关系。

在B细胞、成纤维细胞和巨噬细胞中,胰腺癌共表达基因对的数量最高;在T细胞和内皮细胞中,乳腺癌共表达基因对的数量最高;在恶性细胞中,卵巢癌共表达基因对的数量最高。

### 2.2 转录因子的调控广度分析

将1 001个转录因子的调控广度作为特征,对30个“癌症类型-细胞类型”组合进行层次聚类分

析,结果显示,T细胞、巨噬细胞、内皮细胞的转录因子调控广度存在着较强的跨疾病保守性,受癌症类型影响较小。例如,在层次聚类中,来自不同癌症的T细胞可以形成一个单独的类;此外,来自乳腺癌、卵巢癌、胰腺癌、肾癌的内皮细胞和巨噬细胞也可形成单独聚类。见图3。

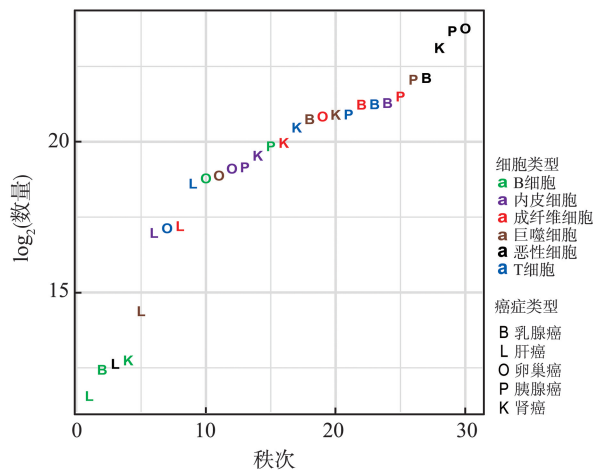


图2 30张基因共表达图谱排序  
Figure 2 Thirty gene co-expression maps were sorted

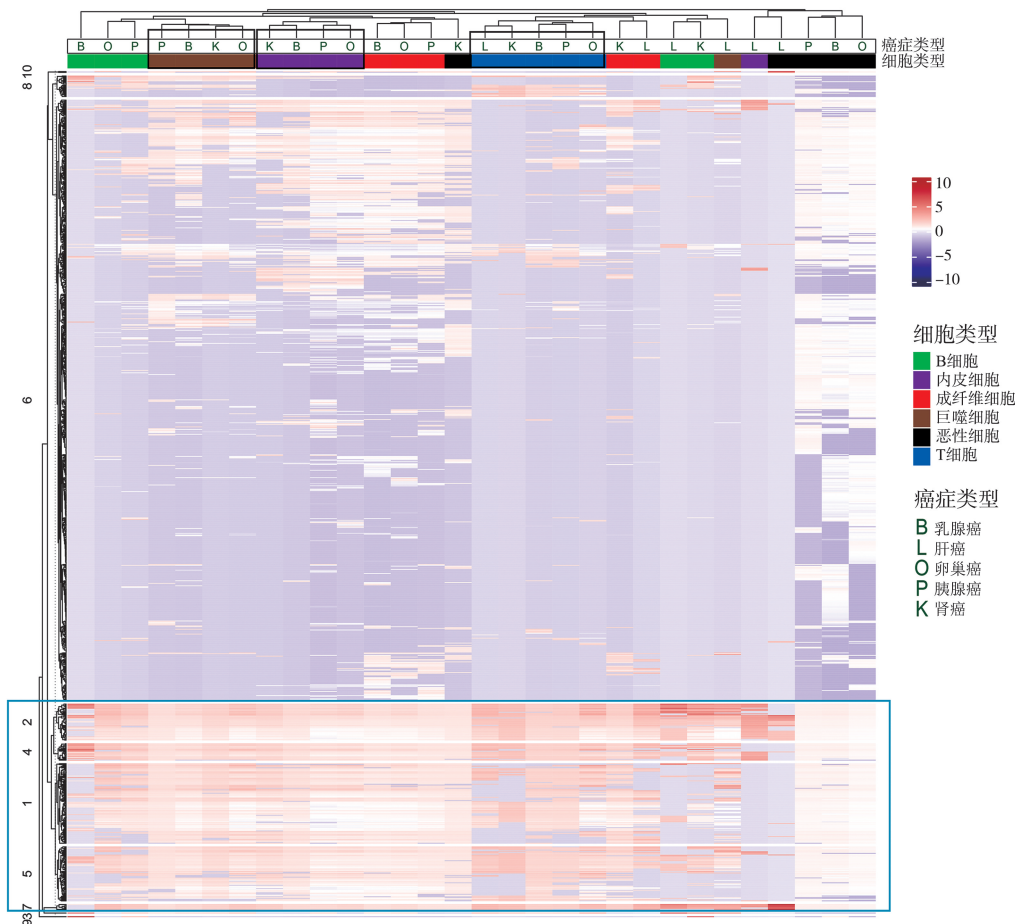


图3 在30个“癌症类型-细胞类型”组合中转录因子的调控广度  
Figure 3 The regulatory breadth of transcription factors for the 30 cancer type-cell type combinations

部分转录因子在多个细胞类型和癌症中均展现出较高的调控广度,其所参与的生物学过程主要分为以下6类:细胞生长和分化调控(如 *FOXO1*、*RUNX1*、*SMAD2* 等基因);信号转导(如 *NFKB1*、*NFKB2*、*RELA* 等基因);染色质重塑和基因表达调控(如 *ARID1A*、*ARID2*、*CTCF*、*SMARCC1* 等基因);细胞周期和凋亡调控(如 *BCLAF1*、*DDIT3*、*EGR1* 等基因);应激响应(如 *HIF1A*、*HSF1* 等基因);免疫调控(如 *IRF1*、*IRF2*、*STAT1*、*STAT6* 等基因)。

### 2.3 KEGG 通路基因的共表达特性

将46条代谢通路的平均基因共表达水平作为特征,对30个“癌症类型-细胞类型”组合进行层次聚类分析,结果显示乳腺癌和卵巢癌的恶性细胞可以形成一个单独的类,并在丙酮酸代谢、嘧啶代谢等关键能量代谢通路上显示出高度的基因共表达水平;乳腺癌和胰腺癌的巨噬细胞也可以形成一个单独的类,并在丙酮酸代谢、核黄素代谢、胆固醇代谢等通路的基因上表现出高度共表达。见图4。

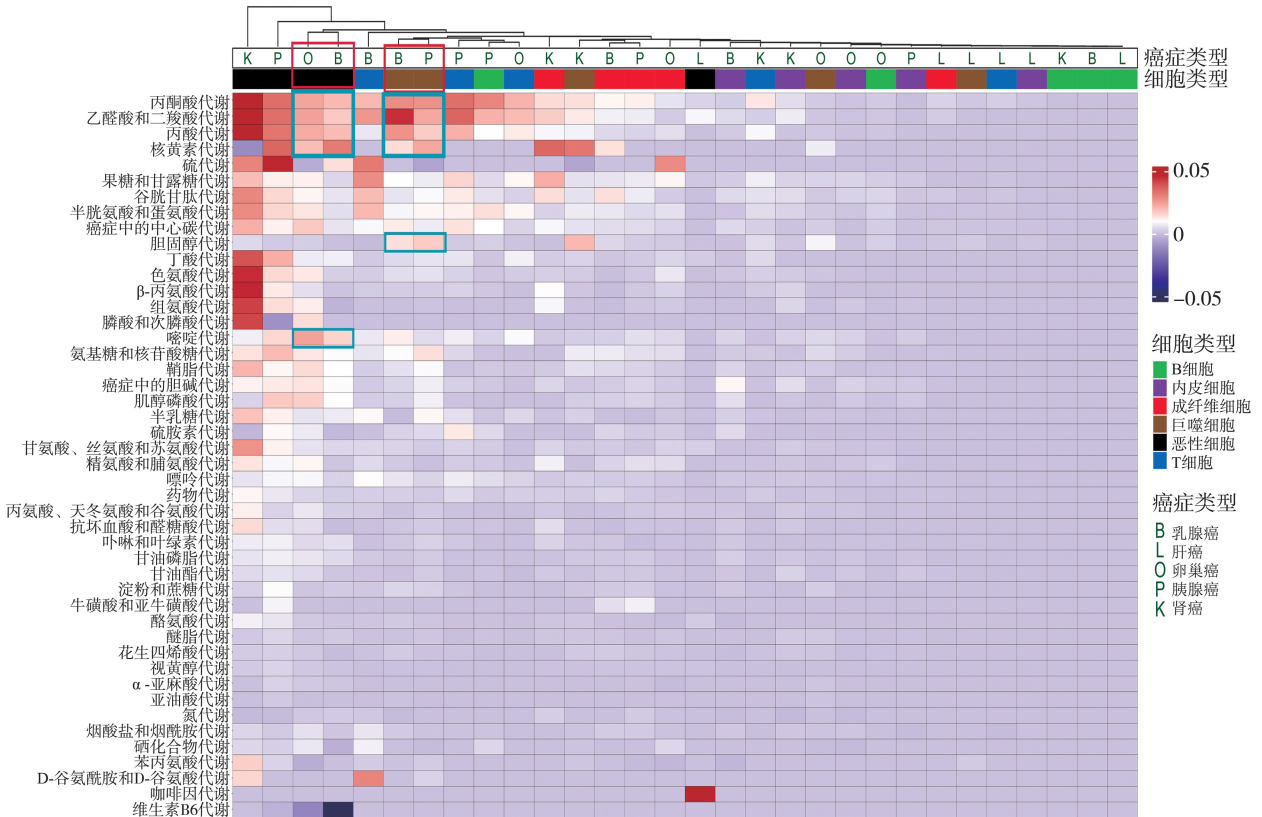


图4 在30个“癌症类型-细胞类型”组合中代谢通路的平均基因共表达水平  
Figure 4 The average gene co-expression levels of metabolic pathways for the 30 cancer type-cell type combinations

将57条信号通路的平均基因共表达水平作为特征,对30个“癌症类型-细胞类型”组合进行层次聚类分析,结果表明,来自胰腺癌、卵巢癌和肾癌的恶性细胞形成了一个单独的聚类,且部分信号通路在这3种恶性细胞中均展现出高度的基因共表达水平,它们所参与的生物学过程主要分为以下5类:细胞分化、增殖、凋亡与发育(如 *Notch*、*ERBB*、*FoxO*、*Hedgehog* 信号通路等);代谢与能量调节(如磷酸戊糖途径、鞘脂信号通路等);激素调控(如胰岛素、催乳素、甲状腺激素信号通路等);神经系统功能(如神经营养素信号通路、内源性大麻素逆向信号传导等);基因表达调控(如 mRNA 监测通路等)。见图5。

### 2.4 癌症组织与正常组织基因共表达谱的比较

表2比较了6种细胞类型在四种癌症组织(vs. 癌旁正常组织)中的差异共表达基因对数量,同时统计了每种细胞类型中的高度可信差异共表达基因对数,发现在内皮细胞和巨噬细胞中分别存在116和526对高度可信差异共表达基因对。图6展示了一个典型的巨噬细胞的高度可信差异共表达基因对 *NDUFB2-PFM1*。这两个基因在卵巢癌和肾癌的癌症组织的巨噬细胞中均有着显著的正相关关系,而在癌旁对照组织的巨噬细胞中,这一正相关关系则丧失。

最终只在内皮细胞和巨噬细胞中鉴定到了高度可信差异共表达基因对,其主要与能量代谢、蛋白质合成与加工等生物学功能相关。

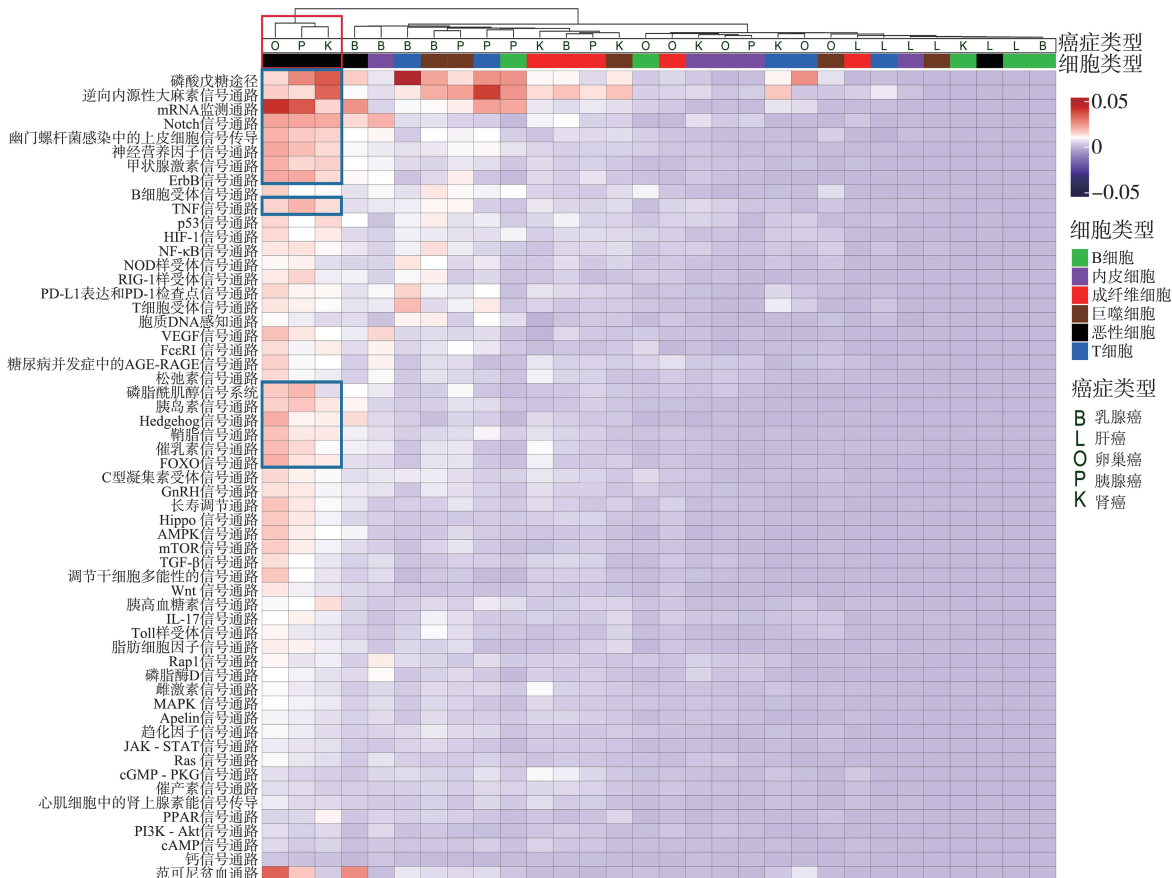


图5 在30个“癌症类型-细胞类型”组合中信号通路的平均基因共表达水平

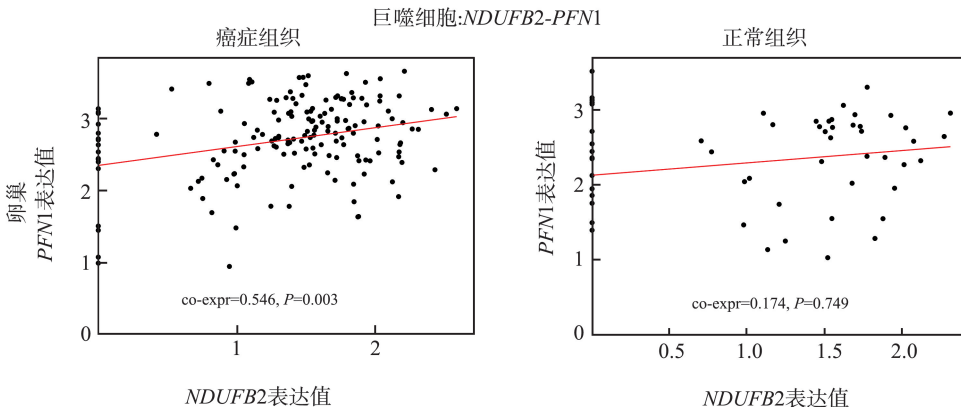
Figure 5 The average gene co-expression levels of signaling pathways for the 30 cancer type-cell type combinations

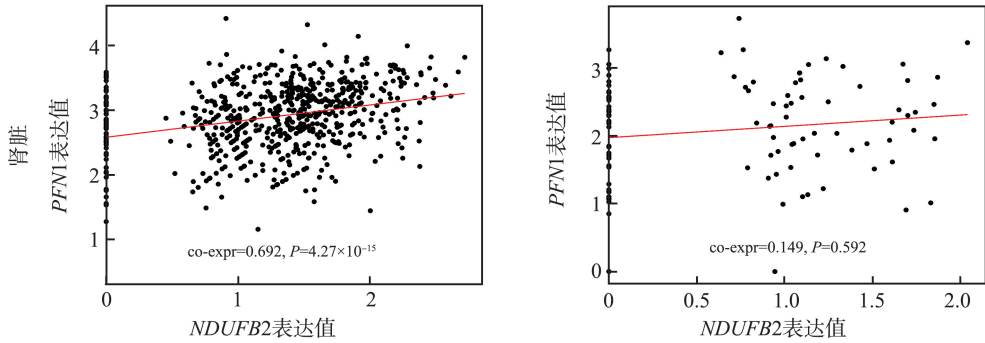
表2 各“组织-细胞类型”组合中差异共表达基因对数量

Table 2 Number of differentially co-expressed gene pairs in each tissue-cell type combination

细胞类型	乳腺	肝脏	卵巢	肾脏	平均差异共表达	高度可信差异共表达
T 细胞	—	27 672	0	35 502	21 058	0
B 细胞	—	234	—	9 843	5 039	0
成纤维细胞	51 137	1 067	—	4 013	18 739	0
内皮细胞	—	6 171	32 579	4 053	14 268	116
巨噬细胞	—	1	45 739	27 619	24 453	526
上皮细胞	246	0	—	91 561	30 602	0

注:表格中“—”表示在正常组织中因缺少该种细胞类型的数据,从而导致差异共表达基因对数据的缺失。





红线表示线性拟合趋势,co-expr 为通过 CS-CORE 方法计算的该基因对的共表达值。 $P$  值经过 BH 校正。

图6 巨噬细胞的高度可信差异共表达基因对 *NDUFB2-PFN1* 在不同情境下的共表达情况

Figure 6 The co-expression of the highly reliable differentially co-expressed gene pair *NDUFB2-PFN1* in macrophages under various conditions

### 3 讨论

本研究利用公共领域单细胞 RNA 测序数据,系统构建了癌症组织中的细胞类型特异性基因共表达图谱。结果表明,对于同一种细胞类型,不同的癌症会显著影响细胞内部的转录调控。例如,胰腺癌中 B 细胞和成纤维细胞组的共表达基因对数量明显高于其他癌症类型;而乳腺癌中的 T 细胞和内皮细胞组则具有更高的共表达基因对数量。这种共表达基因对数量的差异可能反映了肿瘤微环境的多样性和异质性,为理解癌症的发生与进展提供了新的视角。

众多研究表明,转录因子在介导基因共表达中发挥重要作用<sup>[23-25]</sup>。因此进一步研究转录因子在介导基因共表达中的作用时发现,T 细胞、巨噬细胞等在不同癌症类型中的转录因子调控广度具有相似性,表明转录因子在这些细胞类型中的基因调控功能较为稳定。此外,部分转录因子在较多的癌症类型和细胞类型中表现出较高的调控广度,提示它们可能在细胞的基因调控中发挥核心作用。若干转录因子已经被报道会影响癌症的发生和发展过程。例如,*FOXO1* 可以通过抑制凋亡和上调基质金属蛋白酶(如 *MMP-1* 和 *MMP-9*) 促进肿瘤的进展与转移<sup>[26-28]</sup>。*HIF-1* 的激活在肿瘤细胞对氧环境变化的适应性反应中起着关键作用,促进肿瘤的存活和进展<sup>[29-32]</sup>。

对 KEGG 通路基因的共表达特性的分析结果显示,乳腺癌和卵巢癌的恶性细胞在丙酮酸代谢、嘧啶代谢等关键代谢通路中显示出高度的基因共表达水平。这一现象表明,尽管不同癌症类型的分子背景和临床表现存在差异,但某些基础代谢通路可能是肿瘤细胞生长与存活的核心需求。例如,丙酮酸

代谢是细胞能量代谢的关键过程,且在肿瘤细胞中常常表现出异常调节,以适应快速增殖的需求<sup>[33]</sup>。本研究还发现,乳腺癌和胰腺癌的巨噬细胞均在丙酮酸代谢、核黄素代谢、胆固醇代谢等通路中显示出高度的基因共表达水平,提示这些代谢通路可能与巨噬细胞的免疫功能以及在肿瘤中的适应性变化密切相关。例如,癌症细胞可以通过改变巨噬细胞的胆固醇代谢来促进巨噬细胞分化,进而促进肿瘤微环境的建立<sup>[34]</sup>。

类似地,信号通路基因调控特征在恶性细胞中表现出一定的保守性。此外,本研究发现部分通路基因在卵巢癌、肾癌和胰腺癌的恶性细胞中具有高度共表达水平,说明这些通路可能在这 3 种癌症的发生和发展中起到重要作用,如 Notch 信号通路的失调促进恶性肿瘤的上皮-间质转化和血管生成,并与癌症增殖、侵袭和转移密切相关<sup>[35]</sup>。

与正常组织相比,癌症组织中的细胞类型特异性共表达基因对表现出显著差异,特别是在巨噬细胞中,表明疾病可能通过重塑细胞内的基因共表达网络来影响细胞功能。巨噬细胞在肿瘤微环境中常扮演促肿瘤或抗肿瘤双重作用,这种显著的共表达变化可能反映了肿瘤相关巨噬细胞的分化和功能改变<sup>[36]</sup>。这些差异基因为诊断癌症组织中异常细胞类型提供了潜在的生物标志物,同时也可能成为靶向治疗的切入点。

在本研究的基础上,未来工作可以从两方向展开深入探索:①基于共表达的结果筛选更多已知信号通路的新成员,并通过实验进行验证;②探索基因对之间的共表达在预测癌症患者对治疗的响应和生存结局方面的潜力。

综上所述,恶性细胞具有最复杂的调控网络,各类细胞类型的转录因子调控在跨癌种层面表现出保守性,且肿瘤微环境会重塑巨噬细胞的代谢通路。

该发现为靶向细胞类型特异性调控机制提供了新理论参考。

## 参考文献:

- [1] Lee HK, Hsu AK, Sajdak J, et al. Coexpression analysis of human genes across many microarray data sets [J]. *Genome Res*, 2004, 14(6): 1085-1094.
- [2] Torkamani A, Dean B, Schork NJ, et al. Coexpression network analysis of neural tissue reveals perturbations in developmental processes in schizophrenia [J]. *Genome Res*, 2010, 20(4): 403-412.
- [3] Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis [J]. *Stat Appl Genet Mol Biol*, 2005, 4: Article17. doi: 10.2202/1544-6115.1128
- [4] Mostafavi S, Gaiteri C, Sullivan SE, et al. A molecular network of the aging human brain provides insights into the pathology and cognitive decline of Alzheimer's disease [J]. *Nat Neurosci*, 2018, 21(6): 811-819.
- [5] Hao RH, Zhang TP, Jiang F, et al. Revealing brain cell-stratified causality through dissecting causal variants according to their cell-type-specific effects on gene expression [J]. *Nat Commun*, 2024, 15(1): 4890. doi: 10.1038/s41467-024-49263-4
- [6] Su C, Zhang JF, Zhao HY. Estimating cell-type-specific gene co-expression networks from bulk gene expression data with an application to Alzheimer's disease [J]. *J Am Stat Assoc*, 2024, 119(546): 811-824.
- [7] Harris BD, Crow M, Fischer S, et al. Single-cell co-expression analysis reveals that transcriptional modules are shared across cell types in the brain [J]. *Cell Syst*, 2021, 12(7): 748-756.
- [8] Hao YH, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data [J]. *Cell*, 2021, 184(13): 3573-3587.
- [9] Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression [J]. *Genome Biol*, 2019, 20(1): 296. doi: 10.1186/s13059-019-1874-1
- [10] Sarkar A, Stephens M. Separating measurement and expression models clarifies confusion in single-cell RNA sequencing analysis [J]. *Nat Genet*, 2021, 53(6): 770-777.
- [11] Kumar N, Mishra B, Athar M, et al. Inference of gene regulatory network from single-cell transcriptomic data using pySCENIC [J]. *Methods Mol Biol*, 2021, 2328: 171-182. doi: 10.1007/978-1-0716-1534-8\_10
- [12] Su C, Xu ZC, Shan XN, et al. Cell-type-specific co-expression inference from single cell RNA-sequencing data [J]. *Nat Commun*, 2023, 14(1): 4846. doi: 10.1038/s41467-023-40503-7
- [13] Gavish A, Tyler M, Greenwald AC, et al. Hallmarks of transcriptional intratumour heterogeneity across a thousand tumours [J]. *Nature*, 2023, 618(7965): 598-606.
- [14] Shen WK, Chen SY, Gan ZQ, et al. AnimalTFDB 4.0: a comprehensive animal transcription factor database updated with variation and expression annotations [J]. *Nucleic Acids Res*, 2023, 51(D1): D39-D45.
- [15] Kanehisa M, Goto S. KEGG Kyoto encyclopedia of genes and genomes [J]. *Nucleic Acids Res*, 2000, 28(1): 27-30.
- [16] Chen EY, Tan CM, Kou Y, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool [J]. *BMC Bioinformatics*, 2013, 14: 128. doi: 10.1186/1471-2105-14-128
- [17] Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update [J]. *Nucleic Acids Res*, 2016, 44(W1): W90-W97.
- [18] Xie ZR, Bailey A, Kuleshov MV, et al. Gene set knowledge discovery with enrichr [J]. *Curr Protoc*, 2021, 1(3): e90. doi: 10.1002/cpz1.90
- [19] Hao YH, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data [J]. *Cell*, 2021, 184(13): 3573-3587.
- [20] Wickham H. *ggplot2: elegant graphics for data analysis* [M]. New York: Springer-Verlag, 2016.
- [21] Gu ZG, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data [J]. *Bioinformatics*, 2016, 32(18): 2847-2849.
- [22] Yu GC, Wang LG, Han YY, et al. clusterProfiler: an R package for comparing biological themes among gene clusters [J]. *OMICS*, 2012, 16(5): 284-287.
- [23] Hu XL, Hu YH, Wu FJ, et al. Integration of single-cell multi-omics for gene regulatory network inference [J]. *Comput Struct Biotechnol J*, 2020, 18: 1925-1938. doi: 10.1016/j.csbj.2020.06.033
- [24] Doughty BR, Hinks MM, Schaepe JM, et al. Single-molecule states link transcription factor binding to gene expression [J]. *Nature*, 2024, 636(8043): 745-754.
- [25] Kciuk M, Gielecińska A, Kołat D, et al. Cancer-associated transcription factors in DNA damage response [J]. *Biochim Biophys Acta Rev Cancer*, 2022, 1877(4): 188757. doi: 10.1016/j.bbcan.2022.188757
- [26] Farhan M, Silva M, Xing XG, et al. Role of FOXO transcription factors in cancer metabolism and angiogenesis [J]. *Cells*, 2020, 9(7): 1586. doi: 10.3390/cells9071586
- [27] Jiramongkol Y, Lam EW. FOXO transcription factor family in cancer and metastasis [J]. *Cancer Metastasis Rev*, 2020,

- 39(3): 681-709.
- [28] Liu Y, Ao X, Jia Y, et al. The FOXO family of transcription factors: key molecular players in gastric cancer [J]. *J Mol Med (Berl)*, 2022, 100(7): 997-1015.
- [29] Taneja N, Chauhan A, Kulshreshtha R, et al. HIF-1 mediated metabolic reprogramming in cancer: mechanisms and therapeutic implications [J]. *Life Sci*, 2024, 352: 122890. doi: 10.1016/j.lfs.2024.122890
- [30] Lee SH, Golinska M, Griffiths JR. HIF-1-independent mechanisms regulating metabolic adaptation in hypoxic cancer cells [J]. *Cells*, 2021, 10(9): 2371. doi: 10.3390/cells10092371
- [31] Rani S, Roy S, Singh M, et al. Regulation of transactivation at C-TAD domain of HIF-1  $\alpha$  by factor-inhibiting HIF-1  $\alpha$  (FIH-1): a potential target for therapeutic intervention in cancer [J]. *Oxid Med Cell Longev*, 2022, 2022: 2407223. doi: 10.1155/2022/2407223
- [32] Infantino V, Santarsiero A, Convertini P, et al. Cancer cell metabolism in hypoxia: role of HIF-1 as key regulator and therapeutic target [J]. *Int J Mol Sci*, 2021, 22(11): 5703. doi: 10.3390/ijms22115703
- [33] Kiesel VA, Sheeley MP, Coleman MF, et al. Pyruvate carboxylase and cancer progression [J]. *Cancer Metab*, 2021, 9(1): 20. doi: 10.1186/s40170-021-00256-7
- [34] Huang BL, Song BL, Xu CQ. Cholesterol metabolism in cancer: mechanisms and therapeutic opportunities [J]. *Nat Metab*, 2020, 2(2): 132-141.
- [35] Shi QM, Xue C, Zeng YF, et al. Notch signaling pathway in cancer: from mechanistic insights to targeted therapies [J]. *Signal Transduct Target Ther*, 2024, 9(1): 128. doi: 10.1038/s41392-024-01828-x
- [36] Li LX, Tian Y. The role of metabolic reprogramming of tumor-associated macrophages in shaping the immunosuppressive tumor microenvironment [J]. *Biomed Pharmacother*, 2023, 161: 114504. doi: 10.1016/j.biopha.2023.114504

(编辑:相峰)

(上接第30页)

- [40] Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression [J]. *Int J Epidemiol*, 2015, 44(2): 512-525.
- [41] Verbanck M, Chen CY, Nesle B, et al. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases [J]. *Nat Genet*, 2018, 50(5): 693-698.
- [42] Moskalev AA, Shaposhnikov MV, Plyusnina EN, et al. The role of DNA damage and repair in aging through the prism of Koch-like criteria [J]. *Ageing Res Rev*, 2013, 12(2): 661-684.
- [43] Cawthon RM, Smith KR, O'Brien E, et al. Association between telomere length in blood and mortality in people aged 60 years or older [J]. *Lancet*, 2003, 361(9355): 393-395.
- [44] Hannum G, Guinney J, Zhao L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates [J]. *Mol Cell*, 2013, 49(2): 359-367.
- [45] Hashmi SA, Sachdeva S, Sindhu U, et al. The implications of frailty in older adults with epilepsy [J]. *Epilepsia Open*, 2024, 9(6): 2128-2143.
- [46] Das C, Tyler JK. Histone exchange and histone modifications during transcription and aging [J]. *Biochim Biophys Acta*, 2013, 1819(3-4): 332-342.
- [47] Tan CL, Plotkin JL, Venø MT, et al. MicroRNA-128 governs neuronal excitability and motor behavior in mice [J]. *Science*, 2013, 342(6163): 1254-1258.
- [48] Eller MM, Zuberi AR, Fu X, et al. Valine and inflammation drive epilepsy in a mouse model of ECHS1 deficiency [J]. *bioRxiv*, 2024. doi: 10.1101/2024.06.13.598697
- [49] Jiang R, Noble S, Sui J, et al. Associations of physical frailty with health outcomes and brain structure in 483 033 middle-aged and older adults: a population-based study from the UK Biobank [J]. *Lancet Digit Health*, 2023, 5(6): e350-e359.

(编辑:相峰)