

动物所提出空间细胞类型组分解析新算法

2023-08-11 来源：动物研究所

【字体：大 中 小】

 语音播报

8月7日，中国科学院动物研究所翟巍/马亮团队在《自然-通讯》(Nature Communications)上，发表了题为SONAR enables cell type deconvolution with spatially weighted Poisson-Gamma model for spatial transcriptomics的研究论文，提出了新颖的基于空间转录组数据解析空间细胞类型组分的算法模型，并命名为SONAR。

生物体由各种类型的细胞组成。生物系统的结构功能和动态变化很大程度上取决于细胞的空间分布，揭示和利用其规律对探索生命现象和探究疾病进化机制具有重要意义。当前，空间转录组测序 (spatial transcriptomics) 提供了强有力的手段，可在保留空间信息的同时测量组织中的基因表达，从而为解析细胞的空间分布和功能提供了可能。由于受到空间分辨率的限制，此类技术的每个空间位点 (spot) 上的基因表达通常来自于多个细胞的混合。因此，需要通过有效的解卷积算法 (deconvolution) 才能有效获得细胞类型空间组成。目前，已发表的对空间转录组数据解卷积的算法，或未能充分考虑转录组的高度稀疏、高噪声等特性，或未能在解卷积过程中充分利用空间近邻的相似性信息，在实际的应用中常出现推断错误或不稳健等问题。

该研究提出的SONAR算法是基于空间加权回归框架的概率模型，采用Poisson-Gamma分布对空间转录组数据的原始计数进行建模，并可以根据空间转录组数据的特点对位置特异性偏移 (location shift) 和表达计数的过度离散 (overdispersion) 进行综合考虑 (图1)。为了防止在高异质性的组织区域中过度使用空间信息 (例如，在跨结构边界或肿瘤中，其细胞的空间构成可能发生剧烈变化)，SONAR同时引入三个模块 (空间核函数、预聚类、弹性加权) 对空间信息加以筛选并进行有效利用。

该研究在大量具有不同局部特征 (如主导类型细胞丰度/类型数量等) 和不同全局特征 (如空间分布/区域转变模式等) 的模拟集上，以及基于真实的单细胞精度空间转录组数据集 (小鼠大脑/人类心脏数据集) 上验证了SONAR相较于其他算法在细胞组分解析准确性上的优势。在诸多方法中，仅SONAR能够解析心脏流出道区域上神经嵴细胞 (cNCC) 和雪旺祖细胞 (SPC) 的细微分布 (图2)。

该工作将SONAR应用于具有高异质性的胰腺导管癌 (PDAC) 和人类肝细胞癌 (HCC) 数据中，刻画了具有区域特异性细胞类型的空间分布。在HCC数据中，SONAR精细地揭示了在肿瘤/正常组织的过渡区域上，肿瘤微环境中的免疫细胞和成纤维细胞的共定位变化趋势 (图3)。

综上，该研究开发了精准利用空间信息的解析空间细胞类型组分解析新算法——SONAR，设计、应用在具有不同空间模式的模拟集上，并在各类真实数据集上进行了解析和探索。随着空间转录组技术的普及和细胞图谱数据的不断累积，SONAR将为解析生命过程和疾病进展、精确探索大规模空间转录组数据提供助力。研究工作得到国家重点研发计划和国家自然科学基金面上项目等的支持。

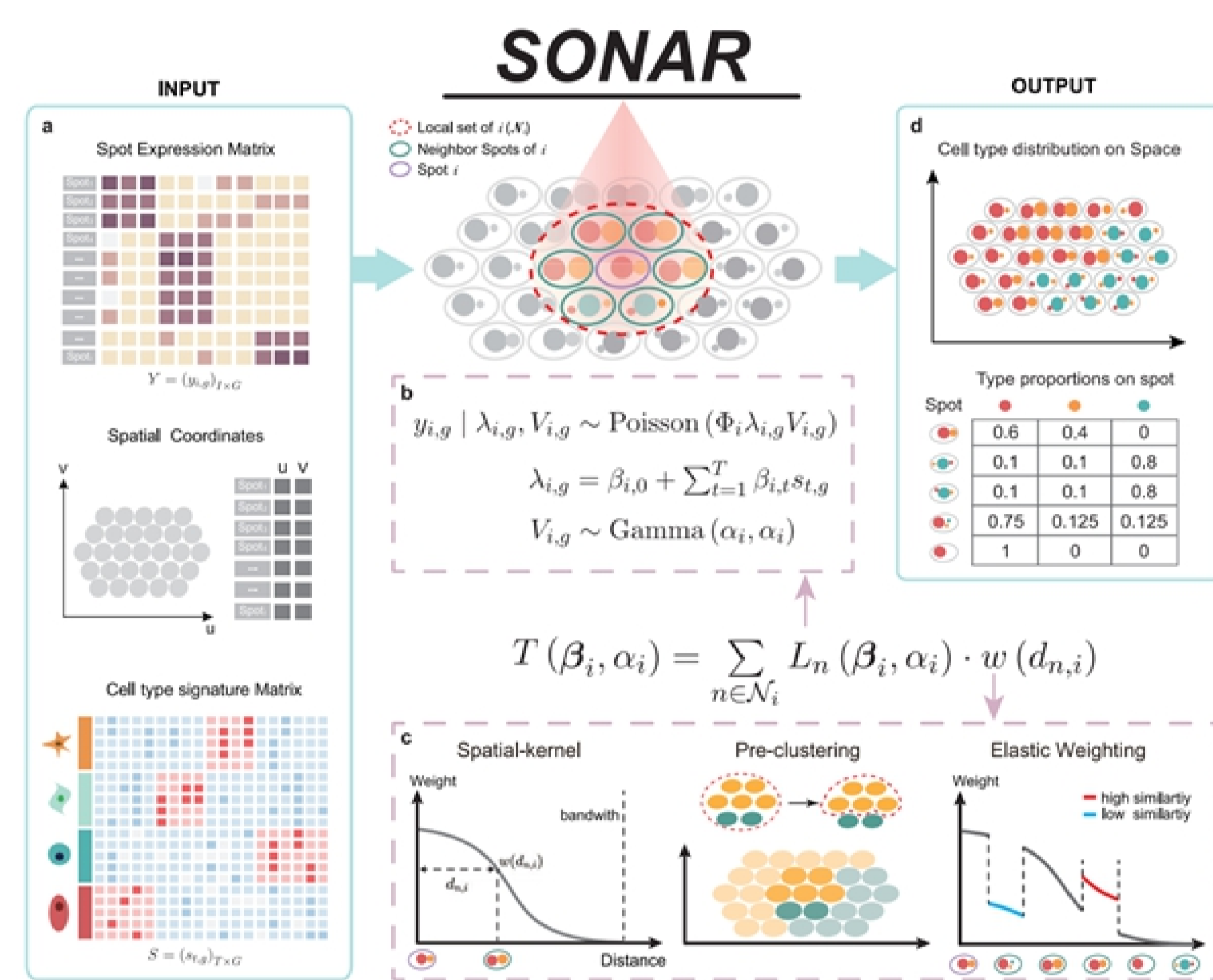
[论文链接](#)


图1. SONAR算法工作流程图

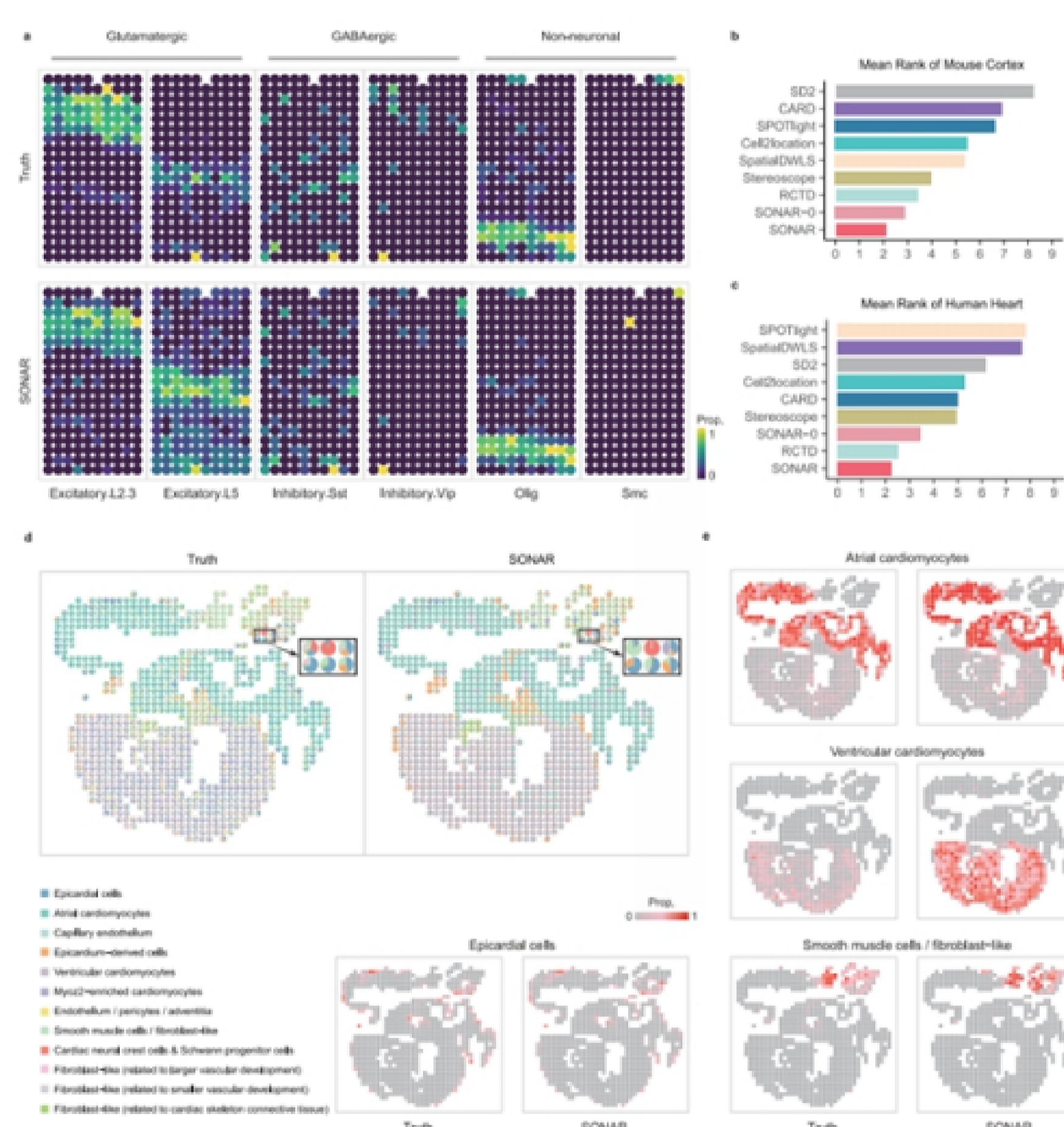


图2. SONAR显著提升了对大脑皮层的结构解析以及稳定识别出心脏的细微结构

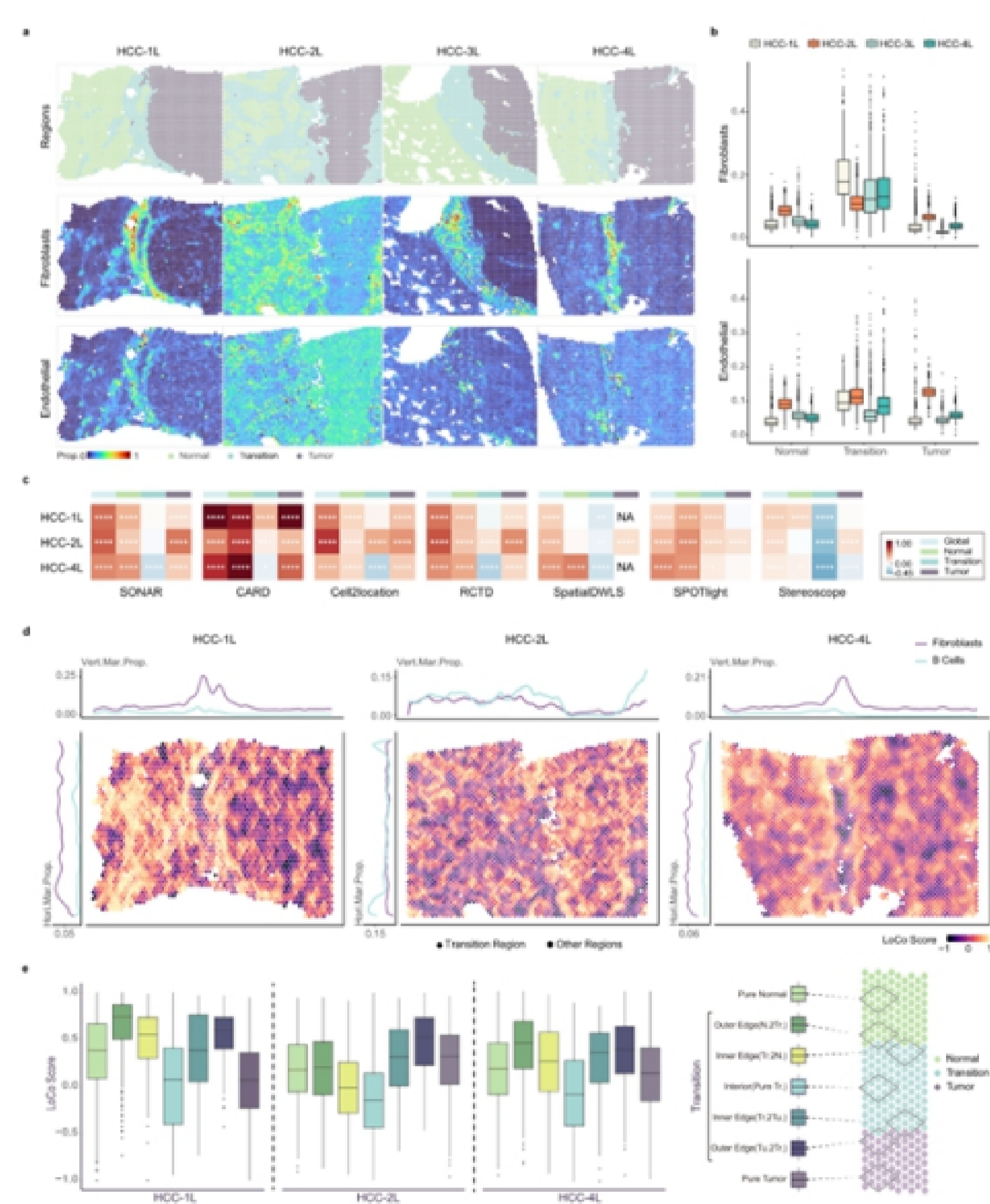


图3. SONAR肝脏肿瘤与正常组织过渡区域上免疫细胞和成纤维细胞的共定位趋势

责任编辑：侯茜 打印    更多分享

» 上一篇：国家纳米中心在单分子层COF膜反常输运机制研究中获进展
 » 下一篇：物理所非绝热全量子化方法研究获进展



扫一扫在手机打开当前页