**摘要**

癌症相关的成纤维细胞 （CAFs）是肿瘤微环境的主要组成部分，由于明显的异质性，它们在疾病形成过程中的起源、作用、进展以及治疗反应仍不清楚。在这里，我们遵循负向选择策略结合单细胞RNA测序技术，获得了来自基因工程小鼠乳腺癌模型768个间充质细胞的转录组信息，定义了CAFs的三个不同亚群。癌症和人肿瘤的实验样本在转录水平和蛋白质水平揭示了CAFs亚群在空间上的分离可归因于不同的起源，包括血管周围微环境、乳腺脂肪垫以及转化的上皮细胞。每一种CAF亚型的基因谱与其独特的功能过程相关，并且通过与转移性疾病相关联，发现在临床上有独立的预示性能力。因此我们得到结论，广泛定义的CAF群体分辨率的提升为生物标志物驱使的CAFs精确靶向药物开发提供了可能性。

**介绍**

在过去的几十年，人们越来越认识到肿瘤微环对恶性表型的重要性，传统的以肿瘤细胞为中心的癌症观点已经改变。癌细胞与其周围环境相互作用的阐明激发了概念上新型靶向治疗的发展，该治疗目的是阻止不同细胞类型的肿瘤块间旁泌信号的转导。癌症相关的成纤维细胞 （CAF）由许多癌症中肿瘤微环境中最普遍的细胞类型构成，包括乳腺癌、胰腺癌和肝癌，并且已经被证实是许多癌症的标志。在实体瘤里，细胞形态仍然是区分CAFs最可靠的方式，而通常作为细胞标志物的，比如α-平滑肌肌动蛋白（SMA），成纤维细胞特异性蛋白1（FSP-1 /S100A4）或成纤维细胞活化蛋白（FAP），既不全面，也不是完全具体的。标志物一致性的缺乏提高了CAFs由多种亚型组成的多样化细胞群的可能性。最近有研究支持这一观点，例如胰腺导管腺癌，乳腺癌，结肠癌和肺腺癌，其中功能上独特的CAFs亚群能够通过一组有限标志物的表达而鉴定出。另外，CAFs显示出有各种来源，包括常驻成纤维细胞、骨髓来源的间充质干细胞、周细胞以及经历了间充质转变的恶性细胞或内皮细胞，进一步表明了成纤维细胞群体的多样性。

单细胞RNA测序（scRNA-seq）是一种新兴技术，能够克服由于群体RNA测序而掩盖细胞亚群信息这一问题，并且使得探究单个细胞转录组成为可能，这样就能够根据相似的转录过程推断细胞亚群。在肿瘤中，来源于黑色素瘤患者的单个细胞的转录组分析明确定义了恶性与非恶性细胞类型的分类，揭示了肿瘤环境中基质细胞与免疫细胞的生长情况。类似的，最近的另一分析结果显示，通过用scRNA-seq方法，基于细胞标志物的表达，将结直肠癌的细胞归为包括CAFs在内的细胞类型。此外，scRNA-seq技术已经被用于鉴定被Hedgehog刺激后产生的特定应答群体CAFs，而Hedgehog反过来将刺激CAF诱导的癌症干细胞微环境。然而，由于受到分析的细胞数量的限制以及scRNA-seq方法本身的局限性，之前的研究未能将肿瘤内广泛定义的细胞类型，如CAFs，特异的分为不同亚群。

在这，我们采用高灵敏度的Smart-seq2方法来描绘从基因工程乳腺癌模型MMTV-PyMT小鼠中分离的768个CAFs细胞的异质性。我们定义了转录过程不同的三个CAF亚群，。值得注意的是，每个CAF亚群通过不同功能基因的表达情况能够被明确的区分，并且被证实在实体瘤内有独特的空间位置。因此，我们的工作着重于在单细胞水平研究乳腺肿瘤中的CAF群体，并且揭示了肿瘤微环境在以前工作中未被认可的功能多样性，为精准医疗手段的进一步发展开辟了道路。

**结果**

**单细胞RNA测序揭示了乳腺CAFs亚群分类。**

为了在细胞水平和功能水平提高乳腺癌中CAFs的分类程度，我们采用单细胞RNA测序技术来检测从乳腺癌模型MMTV-PyMT小鼠的肿瘤中分离出的间质细胞。由于CAF缺乏统一标志物，也因为有望发现以前未知的CAF亚群，我们采用负向筛选策略，使用荧光激活细胞分选技术（FACS）来分离除上皮细胞、免疫细胞、内皮细胞和周细胞之外的其他细胞：EpCAM−/CD45−/CD31−/NG2−（图1a，b；补充图1a-e）。分离出的细胞包含2.5%来自14周龄MMTV-PyMT小鼠晚期癌的单个活细胞。分离出的细胞的CAF标志物PDGFRα和α-SMA蛋白的免疫染色证实了所获得的这一群体的纯度，因为大部分细胞对一种或两种标记物染色呈阳性（补充图1f）。每个CAFs样本的scRNA-seq文库在两块384孔板上进行建库，使用Smart-seq2方法并且在细胞裂解液中加入外源性RNA控制联盟（ERCC）来质控。

基于五种质控指标，768个库中有52个库由于低质量而被过滤掉（补充图1g-k）。所有细胞中基因平均计数少于一次的则被去除，最终的得到10835个内源性基因和53个质控基因用于进一步下游分析。每个细胞含有平均大约4600不同基因的转录本（图1c）。为了研究这一群分离出的CAFs细胞是否能够代表不同细胞亚群，我们用不同方法进行了降维。与ERCC相比确定有明显差异的557个基因的主成分分析（PCA）的结果（补充图2a）显示，有两个主要的群和一个小的群（补充图2b）。接着，还用同一群基因，通过t分布随机领域嵌入（t-SNE）方法的二维投射将细胞明确分为由DBSCAN定义的四群，这1-4群表明了CAFs亚型的存在，且具有离散的基因表达谱（图1d）。由于来自两个肿瘤的细胞用PCA以及t-SNE这两种相似方法分群，且通常呈现相似的质控指标，因此我们接下来的分析不考虑细胞的起源（补充图2b，c）。重要的是，在任何细胞中都没有明显检测到阴性选择标志物：Epcam，Pecam1和Ptprc，排除了观察到的细胞群有间充质细胞和非间充质细胞类型混合的可能性（补充图2d-f）。然而，尽管选择了抗GN2蛋白表达的细胞，但在1和3群中仍检测到来自编码蛋白NG2的基因Cspg4的转录本有一定量的表达（补充图2g）。表明GN2蛋白表面暴露可以忽略不计或其mRNA有翻译后修饰。

**CAFs亚群有不同的基因程序**

接下来，我们研究了细胞亚型中典型CAF标志物的表达。虽然我们在每个细胞中检测到了至少一个CAF标志物的表达，但是大多数细胞仅表达非特异性间充质标志物转录物Vim和Sparc，说明需要一个更好的能够描绘CAF群体一般性和独特性的分子（图1e）。值得注意的是，Pdgfra由群体2中的细胞特异性表达，而Pdgfrb由除群体4之外的所有细胞表达，Fap，S100a4（编码FSP-1）和 Acta2 （编码α-SMA）在所有的四群CAF群体中都有表达。另外，如FACS数据所示，CAF的两个主要亚型，即群体1和群体2，细胞大小也不同。进一步表明细胞亚群具有独立生物物理特性（图1f）。为了确认能聚类，我们使用了一个叫SC3的R包，这个包是为单细胞转录组开发的，结果得到了与之前t-SNE方法得到的相似的分群（补充图3a）。

由于细胞外基质（ECM）的产生和修饰是成纤维细胞的关键功能，我们研究了包含在matrisome中的编码ECM蛋白的基因的转录情况，以寻求CAF亚群的生物学验证。实际上，基于matrisome基因集的无监督层次聚类，我们观察到716个细胞根据之前定义的CAF群体聚类，而群体1和群体3是混在一起的（补充图3b）。所有群体都显示出matrisome基因的独特表达特征，说明每个CAF群体都具有特定生物学功能的独特基质。群体2有最强的ECM特征，通常高表达基因matrisome。为了能检测差异表达基因特异区分每一个CAF亚型，我们对定义的群体进行了再生性优化检验统计（ROTS）。 将每个群体与其他聚集的群体进行比较以发现独特的基因特征，并且将FDR小于0.001的上调基因认为是显著差异表达（SDE）。我们在群体2中检测到了1999个SDE基因，而群体1，3和4分别具有522590和859个SDE基因。每个群体的前18个SDE基因在热图（图2a）中显示。我们用常用算法证实了ROTS的结果，例如SCDE20，edgeR21， DESeq222，和Wilcoxon rank-sum test，都获得了相似的能够相互覆盖基因列表（补充图4a）。

接下来，我们用每个群体的前150个SDE基因通过基因本体论（GO）来定义基因特征，在功能上描述每个亚群（图2b）。群体1的SDE基因显著性富集在血管发育和血管生成的GO集（图2b），因此我们将这种亚型称之为血管CAFs。群体2的SDE基因富集在与CEM和EMT相关的GO集上（值得注意的是，该基因集主要包含基质相关基因），证实了我们之前用matrisome基因集得到的层次聚类结果（图2b和补充图3b）。由于明显的ECM特征，群体2中的细胞被命名为基质CAF（mCAF）。细胞周期相关基因集占群体3 SDE基因的主要部分。与GO分类一致，训练的细胞周期分类器Cyclone将群体3中的大部分细胞鉴定为处于细胞周期G2，M或S期，而其他群中的细胞主要处于G1期（补充图4b）。因此，群体3细胞被称为周期CAF（cCAF）。基于SDE基因，群体4检测到的基因集与细胞分化相关，也与发育及组织形态形成有关。因此，我们将这一亚型标记为发育CAFs（dCAFs）。

**vCAF起源于血管周围**

由于vCAF中的SDE基因与参与血管发育的基因密切相关，我们研究了内皮细胞（如Cdh5，Pecam1，Cd34和Tie1）典型标志基因的表达，以排除vCAF群体的污染。令人欣慰的是，在我们分析的细胞中没有找到内皮细胞标志物有意义表达的任何证据（图2c）。相反，vCAFs的SDE基因集包括了血管调节基因，例如Notch3，Epas1， Col18a1和Nr2f2（图3a）。除了控制血管生成的基因以外，转录因子和参与细胞连接的基因也在vCAF特点转录组中显著表达（图3a）。我们用desmin作为标志物，证实了MMTV-PyMT小鼠肿瘤基质部分中有高丰度vCAFs（图3b）。值得注意的是，与肿瘤前缘相比，肿瘤核心的间充质基质细胞vCAF阳性标志物的比例明显更高（图3b）。与它们的功能明显一致，vCAF主要位于脉管系统附近，如vCAF标志物Nidogen-2和内皮细胞标记物CD31的免疫染色所示（图3c）。值得注意的是，我们观察到Nidogen-2阳性细胞在肿瘤发展的早期阶段（8周龄MMTV-PyMT小鼠）与血管紧密相关。在肿瘤发展过程中，发现越来越多的Nidogen-2阳性细胞从血管中分离，显示来自12周龄和15周龄小鼠的肿瘤基质被浸润。由于Nidogen-2是一种分泌蛋白，我们用RNAscope原位杂交（RNA-ISH）方法将vCAF转录本Kcnj8和Notch3与Nidogen-2免疫染色，证实了其作为免疫染色中vCAF标志物的有效性（补充图4c）。令人欣慰的是，Nidogen-2在人乳腺组织中的免疫染色结果显示出其基质表达模式（图3d），并且在人蛋白质图谱中包含的人乳腺癌中观察到了类似的模式，这一结果提供了独立证据（补充图4d）。我们进一步在来自原位移植模型的肿瘤中检测到了Nidogen-2阳性基质细胞，包括鼠细胞系4T1和EO771以及人乳腺癌细胞系MDA-MB-231（补充图4e）。因此，基于它们的组织学定位，我们得出结论，vCAF亚群源自血管周细胞群体，其随后在肿瘤进展过程中侵入肿瘤基质。

**mCAF是驻留成纤维细胞的后代**

肿瘤基质的mCAF子集特异性表达多种ECM相关基因的转录本，例如糖蛋白（Dcn，Lum和Vcan），结构蛋白（Col14a1），基质细胞蛋白（Fbln1，Fbln2和Smoc），和基质修饰酶（Lox和Loxl1）（图3e）。另外，mCAF大量表达免疫细胞吸引因子CXCL14，提示其对调节肿瘤免疫应答有一定作用。mCAF标志物Fibulin-1和PDGFRα的免疫染色显示，在肿瘤侵袭前沿中，Fibulin-1和PDGFRα阳性细胞普遍存在，而肿瘤核心中mCAF的丰度相对较低（图3f）。两种mCAF标记物Fibulin-1和PDGFRα鉴定出人乳腺癌组织的肿瘤基质中有mCAF大量浸润（图3g，h和补充图4d，f）。与vCAF相反，在MMTV-PyMT小鼠模型的肿瘤进展期间mCAF数量相对减少（图3c）。在原位移植模型中，我们主要在共移植中观察到mCAF，其在异移植模型中稀疏存在（补充图4e）。有趣的是，与与12周龄和14周龄小鼠的恶性组织相反，其中mCAF分别占总CAF群体的40.3％和20.0％，89.0％的成纤维细胞通过FACS检测非转基因FVB / N小鼠乳腺表达的mCAF标志物来分离（图1d和3i，j）。基于mCAFs和正常乳腺中主要的成纤维细胞群有相似标志物表达，我们得出结论，mCAF可能源自肿瘤衍生的常驻成纤维细胞。

**cCAF是vCAF的增殖部分。**

使用matrisome基因集的SC3聚类（补充图3b）证明了cCAF和vCAF聚集在一起。实际上，仅仅发现细胞周期基因在cCAFs和vCAFs之间有差异表达（图3k）。此外，对增殖标志物Ki-67的免疫染色证明，分裂的基质成纤维细胞主要存在于vCAF的巢中，而不是mCAF，因此确定了cCAFs 的位置并认认定cCAF确实是目前参与细胞分裂的vCAF。因此，我们得出结论，cCAF代表了vCAF的增殖片段；基于它们在分离时的相对丰度，vCAF群体中7.7％的细胞被分离了。

**dCAF与肿瘤上皮细胞有共同的表达模式**

除了具有独特的ECM基因谱外，dCAF还根据各种干细胞相关的基因（Scrg1，Sox9和Sox10等）的表达来区分（图4a），与组织发展中推测的功能保持一致。由特异性标志物SCRG1分离出的dCAF在MMTV-PyMT小鼠的肿瘤组织中是比较少的，与从原始肿瘤中分离的该亚型的细胞数量较少的结果一致（图1d和4b）。有趣的是，dCAFs在肿瘤发展的早期阶段与恶性上皮细胞混合，而SCRG1阳性细胞在上皮细胞和晚期肿瘤的基质带中被发现（图4b）。SCRG1的表达在人类其他组织中有类似的分布（补充图4d）。有趣的是，在dCAF中强烈检测到转基因PyMT致癌基因的表达，表明了该细胞亚群的恶性细胞来源（图4c）。

我们接下来用RNA-ISH来检测有各种标志物的肿瘤组织中的dCAF特异性转录本。（补充图5a，b）。主要在肿瘤上皮中检测到低水平Mia转录本，伴随少量表达增加的热点。Spint2转录本在肿瘤上皮中均匀表达，但不在基质中表达，就像通过mCAF特异性转录本Mfap5鉴定出的（补充图5b）。人组织的免疫染色证实了肿瘤上皮中的MIA表达，在基质带中几乎没有离散的MIA阳性CAFs。总之，dCAF SDE基因在肿瘤上皮细胞和基质间充质细胞中的重叠表达表明dCAsF可能源自经历过上皮-间质转化（EMT）的肿瘤细胞。

**CAF亚群表示组织学上不同的实体**

为了最终证明在恶性病变中存在空间上不同的CAF亚群，我们通过使用荧光报道分子或免疫染色直观显现所有亚组。实际上，同时检测vCAF标志物Nidogen-2，mCAF标志物PDGFRα和dCAF标志物SCRG1，鉴定了三种不同的基质群体，这三种群体具有与肿瘤细胞巢相关的不同生长模式和定位（图4d）。为了更好地显示CAF亚群和肿瘤的其他组成成分，我们使用了双光子共聚焦显微镜。得到的图像再次显示了由特异性标志物鉴定到的三个不同CAF群（图4e）。重要的是，我们发现PDGFRα阳性mCAFs主要存在于富含胶原蛋白的部分中，这与其作为ECM来源的推断作用一致。另外，发现SCRG1阳性dCAFs位于肿瘤-基质边界上，表明它们可能是恶性上皮细胞的来源。最后，发现Nidogen-2阳性vCAFs沿着血管以及基质部分分布。为了使用更多标志物确认检测到的CAF亚群是不同的，我们用vCAF和mCAF标志物转录本做RNA-ISH，观察到几对标志物的表达没有重叠（补充图5d-e）。相反，mCAF标志物Svep1与常用但混杂有CAF标记物Acta2的表达（补充图5f）之间检测到之前预期到的部分重叠。最后，我们通过免疫染色证实了mCAF和vCAF标志物Fibulin-1和Nidogen-2在人体组织切片中的互斥性（补充图5g）。

**CAF亚群是独立的预后生物标志物。**

我们接下来要开始确定是否可以在来自人类患者样品的大量RNA-seq数据中鉴定出之前观察到的CAF亚型。我们推断通过使用由高度相关基因组成的特征性基因表达谱可以最优地检测出细胞亚群，因为大量数据中转录过程共同调节的特征将指示常见细胞起源。因此，使用来自癌症基因组图谱（TCGA）数据库乳腺癌的大量RNA-seq数据，我们在每个CAF亚型的SDE基因中鉴定出了高度相关的基因，统计了vCAFs的7个基因和mCAFs的30个基因的精简基因集。与其他细胞亚群相比，该基因集是特异性的，因此进一步表明在人乳腺肿瘤中也存在CAF亚型（图5a）。此外，基因集在其他癌症（如胰腺癌，肺癌和肾细胞癌）的大量RNA-seq数据中也具有高度和特异性相关性，表明在不同恶性肿瘤中CAF发展有一定程度的共性（图5b-d）。当对dCAFs使用相同的方法时，只有2个基因保持相关性系数> 0.7，表明与其他CAF群体相比，许多代表dCAFs的基因也可能由肿瘤组织内的其他细胞类型表达，或者dCAFs非常稀少，这与它们可能以瞬时EMT为起源这一推论相一致（补充表1）。我们没有尝试获得cCAFs的基因谱，因为这将不可避免地与增殖相关基因的总体表达相混淆。。

接下来，为了确定vCAFs和mCAFs功能上不同的基因程序是否也可以从TCGA数据库的大量数据分析中被辨别出来，我们研究了每种细胞亚型的精练基因谱与其推断功能的meta基因之间的相关性，即血管生成和ECM生成的调节。与来自小鼠肿瘤的数据一致，vCAF特征与乳腺肿瘤中的内皮细胞meta基因高度相关，而mCAF特征与ECM meta基因非常相关（图5e-h）。值得注意的是，未观察到相反的关系，强烈表明了vCAFs和mCAFs在小鼠和人肿瘤之间功能上不同的基因程序的保守性。此外，使用胰腺癌的转录数据进行了类似的研究（补充图6a）

CAFs被认为是确定癌症患者预后关键参数的重要调节物，包括肿瘤进展，转移性接种和对治疗的反应。将CAF集的平均组合中心表达作为细胞丰度的代表，我们接下来将确定细胞CAF亚型是否与人群中的转移性传播相关。我们首先用病理对照研究中基于人群的嵌套病例转录数据进行研究，该研究包括768名受试者。190名患有远端转移性疾病（病例）的乳腺癌患者从连续的一系列个体中选择，并且每个患者在诊断时与辅助治疗、年龄和日历期密切匹配，其中3名患者无转移（对照组）。来自病例对照研究数据集的研究显示，vCAFs的基因特征与内皮细胞meta基因及微血管特征强烈相关，而mCAFs则与基质来源的侵袭特征和基质相关的治疗预测特征高度相关（图5i）。一个强烈支持CAF作为恶性表型修饰因子的概念中，vCAF基因特征被发现是一个独立的预后指标，该指标与转移性疾病的风险增加相关，涉及常见的条件逻辑回归模型中的单变量和多变量分析，例如淋巴结状态，肿瘤大小，HER2状态和增殖指数等风险因素（表1）。同样，尽管程度较轻，mCAF与播散性疾病的风险也相关（表1）。病例对照研究中的所有相关性均与肿瘤的分子亚型无关（补充图6b，c），并且vCAFs和mCAFs都与PAM50增殖meta基因弱相关，作为CAF丰度和转移性传播之间的联系，排除了癌细胞分裂的一般影响，（图5i）。我们观察到vCAF和mCAF与其他微环境基因特征的相关性，分子亚型的独立性，以及来自METABRIC群体中1875例患者第二个临床注释良好的基因数据集证实了复发的相关性（补充图6d，补充表2）。最后，为了得到分析结果显示出的vCAFs和mCAFs侵袭作用的实验相关性，我们在细胞迁移系统上室接种了从MMTV-PyMT小鼠分离出的PeRo-Bas1乳腺癌细胞。恶性细胞从下室被分离，其中FACS分离的CAF群体用涂有Matrigel的膜接种。支持我们之前在人类肿瘤中的发现，与单独细胞培养基相比，vCAFs和mCAFs均显著增加了通过基质侵入下室的癌细胞数量（图5j）。总之，vCAFs和mCAFs的基因谱在大量RNA测序数据中很容易检测到，并且对人类肿瘤具有生物学和临床意义。

**讨论**

总之，我们通过使用scRNA-seq技术，为至少三个空间和功能上不同的乳腺CAF亚群的存在提供了令人信服的证据，我们已经显着改善了CAF研究的细胞分辨率。组织学表征提示了CAF不同的细胞来源，vCAFs，mCAFs和dCAFs似乎分别源自血管周围位置，驻留成纤维细胞和经受EMT的恶性细胞。通过用在小鼠和人肿瘤之间，以及在不同的恶性疾病之间保守的精炼的基因特征来看，CAF亚型在大量转录数据中是可区分的。值得注意的是，不同CAF亚型基因的表达，作为细胞丰度的特征，可用作人乳腺癌中转移性传播的独立预测因子，表明了观察到的CAF亚群具有生物学相关性。

通过使用716个单独高质量转录组scRNA-seq数据和无偏聚类，我们以高分辨率剖析肿瘤微环境中最突出的细胞成分，从而在广泛定义的CAF群体中鉴定出三个不同的细胞亚群。基于对有限的标志物表达的分析，发现肿瘤微环境包含各种细胞类型的亚群，包括巨噬细胞，内皮细胞和CAFs。最近的一项研究通过scRNA-seq定义了两个CAFs在结直肠癌中的可能的亚群。先前使用scRNA-seq分析肿瘤的工作尝试将癌症中所有组成细胞分类，从而影响了每种细胞类型研究的分辨率。我们用力高灵敏度的Smart-seq2方法特异性分析了大量无偏好性CAFs群体，该方法能够罗列出数百到数千个基因的综合列表，这些基因区分了每个CAF子集，从而描述了肿瘤块特定元素的完整复杂性。是否更深入的分析会揭示CAFs更低丰度的子集，或者更加细分我们观察到的聚类，仍有待测试。此外需要细节分析，以便比较和对比我们的研究定义的CAF亚群和文献中最近描述的CAF亚群。例如，基于六种常用间充质细胞标志物差异表达，用FACS分离出的乳腺CAFs亚群被证明具有免疫抑制基因程序。免疫调节功能未被鉴定为我们研究中观察到的任何CAF群体的特殊特征，需要更详细的分析以进一步研究，因为确实检测到重要免疫调节元件的差异表达。此外，最近的两项研究表明，CAF亚群分别表达CD10 / Gpr77和Hedgehog的靶基因如Fgf5能促进癌症干细胞特征。有趣的是，CD10和Gpr77都是由dCAsF特异性表达的，这提升了恶性干细胞通过EMT维持自身利基的可能性。随着越来越多的研究通过各种技术详细介绍了CAF亚群，开发跨平台比较计算工具来理解各种CAF亚群之间功能关系将变得越来越重要。

CAF的起源仍然存在争议。可想而知，因为来自不同谱系的CAF均在实体肿瘤中出现。实际上，我们的分析表明不同的CAF子集可能具有不同的来源。有趣的是，基于标志物表达的相似性，以及与周围正常组织接近的外周位置，mCAF似乎起源于正常小鼠乳腺中主要的驻留成纤维细胞。相反，来自MMTV-PyMT小鼠的恶性乳腺组织以vCAFs为主，可能是由于它们的增殖能力表现为cCAF特性。vCAF亚群与周细胞共享许多标记基因，包括Cspg4，Rgs5，Pdgfrb和Des，尽管水平相对较低。有报道称内皮唾液酸蛋白（Tem1或Cd248）的表达是活化的间充质细胞的标志物，包括CAFs和肿瘤周细胞。实际上，内皮唾液酸蛋白由vCAFs高表达。我们观察到vCAFs在早期肿瘤中与血管相关紧密，随后在进展期间脱离和侵入肿瘤组织，表明vCAFs与血管周细胞之间存在密切关系。总之，基因表达数据和定位有利于为vCAFs推测周细胞来源，尽管周细胞-成纤维细胞转变的概念一直处于最近的争论之中。vCAFs在肿瘤核心中富集的事实可能表明缺氧促使vCAF从其血管周围环境中脱离；它们表达Epas1（HIF2-α）进一步支持这一概念。最后，发现表示dCAF的SDE基因也由肿瘤上皮细胞表达，包括转基因致癌基因PyMT。尽管与其他CAF群体相比，dCAFs不能更好地表达EMT的经典标志物如Slug，Snail和twist1（补充图7），但它们的转录特征和组织学定位表明了上皮起源。可以想象，我们的阴性选择策略可能没有排除已经经历EMT过程的恶性细胞，因为据报道EpCAM在间充质转变期间被下调。有趣的是，检测到的dCAF SDE基因提供了新的区别特征的独特来源，在非EMT恶性细胞和组织内间充质基质细胞的混合群体中，其有助于原位EMT细胞的功能和组织学定义。

众所周知，CAF是许多癌症的标志。然而，最近对胰腺导管腺癌的研究与肿瘤支持性CAF的观点相悖，因为在根除α-SMA 阳性的 CAFs或针对Hedgehog诱导的促纤维化反应后，观察到生长和侵袭性的增加。CAFs对恶性表型的矛盾影响可以通过具有相反功能细胞亚群的存在来解释。我们基于全局性基因表达模式将CAF分为三种不同的亚型。基于已知CAF功能（例如ECM产生）重要基因集的表达，细胞的独立无监督聚类强烈支持CAF的功能分组。另外，通过对肿瘤组织内细胞独特标记物的免疫染色作图，验证了CAF亚群的空间分离。值得注意的是，我们观察到vCAFs，mCAFs和dCAFs之间表达的ECM基因谱有着显着差异。虽然mCAFs大量产生多种基质成分，但vCAFs和dCAFs的表达模式受到更多限制，分别专门生产基底膜产物和旁分泌信号分子。基于特异性基因的表达，可以对该CAF亚群的功能进行推断。作为一个案例，我们最近在另一项研究中确定了旁泌性PDGFCC向CAFs信号传导，PDGFCC是作为乳腺癌基底样分子亚型的调节因子。PDGFCC的受体，即PDGFRα，仅由mCAFs表达，该亚群负责乳腺肿瘤中ERα-阴性特性的鉴定。最近的另一项研究证实了CAF与ERα状态之间的联系，证明CD146阳性CAF这一亚群增加了乳腺肿瘤中激素受体的表达。实际上，CD196不存在于mCAFs中，并且在我们的分类中仅由vCAF表达，这说明特异性靶向DGFRα+ / CD146- CAFs是一项有吸引力的策略，该策略通过使基质样乳腺肿瘤转变为ERα+表型变得对内分泌治疗手段敏感。因此，除了提供关于CAFs不同亚群功能特性的信息之外，我们的方法还提供了用于进一步开发的可能药物靶标。在翻译工作中，我们发现vCAF和mCAF特征在乳腺肿瘤患者样本中高度保守，表明代表不同生物学功能的成纤维细胞亚型也是人类恶性肿瘤的特征。一系列不同恶性疾病的交叉比较表明，我们定义的CAF亚群存在于许多但不是所有其他肿瘤类型中。特定肿瘤中不同CAF亚群的谱似乎反应了生物学特性的不同方面，包括细胞起源和分子激活状态。从大量RNA序列中分辨基质细胞亚群的能力使得利用微环境中的定量手段来作为推测或预后的生物标志物成为可能。支持这一推论的一个例子就是，在由超过2600例乳腺癌患者组成的两类病人vCAFs 或者 mCAFs中，考虑了转移性散播的vCAFs 或者 mCAFs特征具有预后能力。另外，mCAF特征与治疗性预后的基质特征具有相关性。综上，我们呈现了一种更好的细胞、分子及功能性乳腺癌CAFs的分类方法，开启了新型靶向药物开发及更高精确性的具有临床意义的生物标志物。