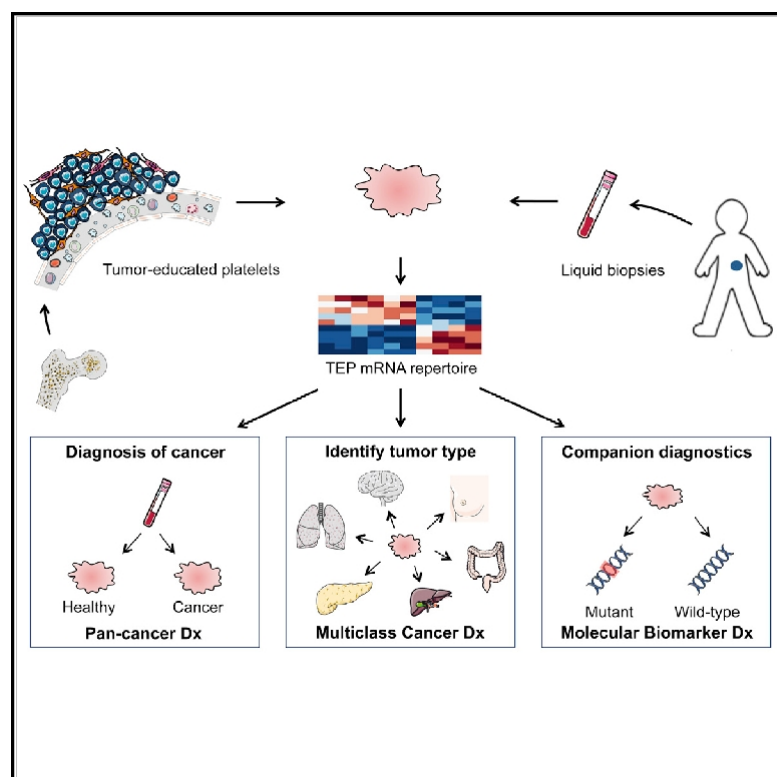


# Cancer Cell

## 肿瘤诱导血小板的 RNA 序列分析实现了基于血液的泛癌、多类和分子途径癌症诊断技术

### 图表摘要



### 作者

Myron G. Best、Nik Sol、Irsan Kooi、  
.....、Bakhos A. Tannous、Pieter  
Wesseling、Thomas Wurdinger

### 通信

[t.wurdinger@vumc.nl](mailto:t.wurdinger@vumc.nl)

### 简介

Best 等人的研究表明，对肿瘤教育血小板的 mRNA 测序能以 96% 的准确率区分癌症患者和健康人，以 71% 的准确率区分患者的六种原发性肿瘤类型，并能识别肿瘤中发现的几种基因改变。

### 亮点

- ▲ 肿瘤通过改变血小板 RNA 图谱 "教育" 血小板 (TEP)
- ▲ TEP 为泛癌症、多类别和辅助诊断提供了 RNA 生物资源
- ▲ 基于 TEP 的液体活检可为临床诊断和治疗选择提供指导
- ▲ 100-500 pg 的血小板总 RNA 足以进行基于 TEP 的诊断

### 编号

GSE68086

2015年11月9日 ©2015 作者

<http://dx.doi.org/10.1016/j.ccell.2015.09.018>



# 肿瘤教育血小板的 RNA 序列分析实现了基于血液的泛癌、多类癌症研究和分子途径癌症诊断技术

Myron G. Best,<sup>1,2</sup> Nik Sol,<sup>3</sup> Irsan Kooi,<sup>4</sup> Jihane Tannous,<sup>5</sup> Bart A. Westerman,<sup>2</sup> Francis Rustenburg,<sup>1,2</sup> Pepijn Schellen,<sup>2,6</sup> Heleen Verschueren,<sup>2,6</sup> Edward Post,<sup>2,6</sup> Jan Koster,<sup>7</sup> Bauke Ylstra,<sup>1</sup> Najim Ameziiane,<sup>4</sup> Josephine Dorsman,<sup>4</sup> Egbert F. Smit,<sup>8</sup> Henk M. Verheul,<sup>9</sup> David P. Noske,<sup>2</sup> Jaap C. Reijneveld,<sup>3</sup> R. Jonas A. Nilsson,<sup>2,6,10</sup> Bakhos A. Tannous,<sup>5,12</sup> Pieter Wesseling,<sup>1,11,12</sup> and Thomas Wurdinger<sup>2,5,6,12,\*</sup>

<sup>1</sup>Department of Pathology, VU University Medical Center, Cancer Center Amsterdam, De Boelelaan 1117, 1081 HV Amsterdam, the Netherlands

<sup>2</sup>Department of Neurosurgery, VU University Medical Center, Cancer Center Amsterdam, De Boelelaan 1117, 1081 HV Amsterdam, the Netherlands

<sup>3</sup>阿姆斯特丹癌症中心 VU 大学医学中心神经病学系, 荷兰阿姆斯特丹 De Boelelaan 1117, 1081 HV

<sup>4</sup>临床遗传学系, 阿姆斯特丹癌症中心 VU 大学医学中心, 荷兰阿姆斯特丹 De Boelelaan 1117, 1081 HV

<sup>5</sup>美国马萨诸塞州查尔斯顿第 13 街 149 号麻省总医院神经内科和哈佛大学医学院神经科学项目 邮编: 02129

<sup>6</sup>thromboDx B.V., 1098 EA Amsterdam, the Netherlands

<sup>7</sup>荷兰阿姆斯特丹 Meibergdreef 9, 1105 AZ 学术医疗中心肿瘤基因组学部

<sup>8</sup>荷兰阿姆斯特丹癌症中心 VU 大学医学中心肺病科, De Boelelaan 1117, 1081 HV Amsterdam, the Netherlands

<sup>9</sup>阿姆斯特丹癌症中心 VU 大学医学中心肿瘤内科, 荷兰阿姆斯特丹 De Boelelaan 1117, 1081 HV

<sup>10</sup>瑞典于默奥大学肿瘤放射科学系, 瑞典于默奥 90185

<sup>11</sup>荷兰奈梅亨 6500 HB 拉德布德大学医学中心病理学系

<sup>12</sup>Co-资深作者

\*通信: [t.wurdinger@vumc.nl](mailto:t.wurdinger@vumc.nl) <http://dx.doi.org/10.1016/j.ccell.2015.09.018>

本文为 CC BY 许可下的开放存取文章 (<http://creativecommons.org/licenses/by/4.0/>)。

## 摘要

肿瘤教育血小板 (TEPs) 被认为是全身和局部肿瘤生长反应的核心参与者, 从而改变了它们的 RNA 图谱。我们通过对 283 份血小板样本进行 mRNA 测序, 确定了 TEPs 的诊断潜力。我们将 228 名局部肿瘤和转移肿瘤患者与 55 名健康人区分开来, 准确率高达 96%。在六种不同的肿瘤类型中, 正确识别原发肿瘤位置的准确率为 71%。此外, 利用代用 TEP mRNA 图谱还能准确区分 MET 或 *HER2* 阳性以及突变 *KRAS*、*EGFR* 或 *PIK3CA* 肿瘤。我们的研究表明, 血小板为泛癌症、多类癌症和辅助诊断提供了一个宝贵的平台, 有可能推动基于血液的 "液体活检" 的临床进展。

## 引言

癌症主要通过临床表现、放射学、生化检验和肿瘤组织的病理分析来诊断。

越来越多的分子诊断测试提供了支持。肿瘤组织样本的分子图谱分析已成为一种潜在的癌症分类方法 (Akbari 等人, 2014 年; Golub 等人, 1999 年; Han 等人, 2014 年; Hoadley 等人, 2014 年; Kandoth 等人, 2014 年)。

## 重要意义

基于血液的 "液体活检" 克服了组织采集的局限性，为微创分子诊断提供了一种手段。癌症早期检测、临床癌症诊断和辅助诊断被认为是液体活检的重要应用。在这里，我们报告了肿瘤教育血小板（TEPs）的 mRNA 图谱可用于泛癌症、多类癌症以及局部和转移性癌症患者的辅助诊断。TEPs 能够精确定位原发肿瘤的位置，这推动了液体活检在癌症诊断中的应用。这项原理验证研究的结果表明，血小板是一种潜在的基于血液的癌症诊断一体化平台，只需使用相当于一滴血的量。

2013; Ramaswamy 等人, 2001; Su 等人, 2001)。为了克服组织采集的局限性, 有人建议使用基于血液的液体活检 (Alix-Panabie`res 等人, 2012 年; Crowley 等人, 2013 年; Haber 和 Velculescu, 2014 年)。目前正在对几种基于血液的生物资源进行液体活检评估, 包括血浆 DNA (Bettegowda 等人, 2014 年; Chan 等人, 2013 年; Diehl 等人, 2008 年; Murtaza 等人, 2013 年; Newman 等人, 2014 年; Thierry 等人, 2014 年) 和循环肿瘤细胞 (Bidard 等人, 2014 年; Dawson 等人, 2013 年; Maheswaran 等人, 2008 年; Rack 等人, 2014 年)。迄今为止, 液体活检在癌症早期检测中的应用一直受到这些生物资源的非特异性影响, 无法确定原发肿瘤的性质 (Alix-Panabie`res 和 Pantel, 2014 年; Bettegowda 等人, 2014 年)。

据报道, 肿瘤教育血小板 (TEPs) 可用于基于血液的癌症诊断 (Calverley 等人, 2010 年; McAllister 和 Weinberg, 2014 年; Nilsson 等人, 2011 年)。血小板是外周血中第二丰富的细胞类型, 是循环中的无核细胞片段, 源自骨髓中的巨核细胞, 传统上因其在止血和伤口愈合中的作用而闻名 (George, 2000; Leslie, 2010)。最近, 血小板已成为全身和局部应对肿瘤生长的核心角色。血小板通过转移肿瘤相关生物分子 ("教育") 与肿瘤细胞对抗是一个新兴概念, 并导致此类生物分子被封存 (Klement 等人, 2009 年; Kuznetsov 等人, 2012 年; McAllister 和 Weinberg, 2014 年; Nilsson 等人, 2011 年; Quail 和 Joyce, 2013 年)。此外, 血小板表面受体活化和脂多糖介导的血小板活化等外部刺激 (Denis 等人, 2005 年; Rondina 等人, 2011 年) 会诱导循环血小板中前 mRNA 的特异性剪接 (Power 等人, 2009 年; Rowley 等人, 2011 年; Schubert 等人, 2014 年)。血小板还可能因癌细胞和肿瘤微环境 (如基质细胞和免疫细胞) 释放的信号而发生队列特异性剪接事件。响应外部信号的特异性剪接事件与血小板直接摄取 (剪接的) 循环 mRNA 的能力相结合, 可为 TEPs 提供高度动态的 mRNA 复制品, 并有可能应用于癌症诊断 (Calverley 等人, 2010 年; Nilsson 等人, 2011 年) (图 1A)。在本研究中, 我们描述了各种癌症患者和健康供体的血小板 mRNA 图谱, 并研究了它们在基于 TEP 的泛癌症、多类癌症和辅助诊断中的应用潜力。

## 结果

### 肿瘤教育血小板的 mRNA 图谱有别于健康人的血小板

我们前瞻性地收集并分离了健康捐献者 (55 人)、经临床表现和肿瘤组织病理分析及分子诊断测试确诊的早期局部癌症 (39 人) 或晚期转移性癌症 (189 人) 患者的血小板。患者队列包括六种肿瘤类型, 即非小细胞肺癌 (NSCLC,  $n = 60$ )、结直肠癌 (CRC,  $n = 41$ )、胶质母细胞瘤 (GBM,  $n = 39$ )、胰腺癌 (PAAD,  $n = 35$ )、肝胆癌 (HBC,  $n = 14$ ) 和乳腺癌 (BrCa,  $n = 39$ ) (图 1B; 表 1; 表 S1)。该队列中的

健康捐献者的年龄跨度很大 (21-64 岁, 表 1)。

通过对随机选取的新鲜分离血小板样本进行形态学分析确认血小板纯度 (污染度为每 1,000 万个血小板中有 1 至 5 个有核细胞, 见补充实验步骤), 并分离和评估血小板 RNA 的质量和数量 (图 S1A)。总计 100-500 pg 的血小板总 RNA (相当于不到一滴血中纯化的血小板) 被用于 SMARTer mRNA 扩增和测序 (Ramsköld 等人, 2012 年) (图 1C 和 S1A)。血小板 RNA 测序的平均读数为每个样本约 2200 万个读数。在选择跨内含子 (剪接) RNA 读数并排除了覆盖率低的基因后 (见补充实验步骤), 我们在健康供体 ( $n = 55$ ) 和局部及转移性癌症患者 ( $n = 228$ ) 的血小板中检测到了 5,003 种不同的蛋白质编码和非编码 RNA, 并将其用于后续分析。获得的血小板 RNA 图谱与之前报道的血小板 (Bray 等人, 2013 年; Kissopoulou 等人, 2013 年; Rowley 等人, 2011 年; Simon 等人, 2014 年) 和巨核细胞 (Chen 等人, 2014 年) 的 mRNA 图谱相关, 而与各种非相关血细胞的 mRNA 图谱不相关 (Hrdlickova 等人, 2014 年) (图 S1B)。此外, DAVID 基因本体 (GO) 分析显示, 检测到的 RNA 强烈富集了与血小板相关的转录本 (假发现率 [FDR]  $< 10^{-126}$ )。

在这 5,003 条 RNA 中, 我们确定了已知的血小板标记物, 如 B2M、PPBP、TMSB4X、PF4 和几种非编码长 RNA (如 MALAT1)。与健康捐献者的血小板样本相比, TEPs 的 5,003 个 mRNA 中共有 1,453 个增加, 793 个减少 (FDR  $< 0.001$ ), 而这些血小板 mRNA 图谱之间存在很强的相关性 ( $r = 0.90$ , 皮尔逊相关性) (图 1D)。根据差异检测到的血小板 mRNA 进行无监督分层聚类, 可区分出两个样本组, 但有轻微重叠 (图 1E; 表 S2)。DAVID GO 分析显示, 增加的 TEP mRNA 富含囊泡介导的转运和细胞骨架蛋白结合等生物过程, 而减少的 mRNA 则强烈参与 RNA 处理和剪接 (表 S3)。基因组富集 (CAGE) GO 方法学的相关分析表明, TEP mRNA 图谱与癌症组织特征、组蛋白去乙酰化酶调控和血小板有显著的相关性 (表 2)。与健康人的血小板相比, TEP 中 20 种非蛋白编码 RNA 的水平发生了改变, 这些 RNA 显示了肿瘤类型相关的 RNA 特征 (图 S1C)。

接下来, 我们采用支持向量机算法 (SVM/LOOCV, 见实验步骤; 图 S1D 和 S1E), 确定了训练队列 ( $n = 175$ ) 中基于 TEP 的泛癌症分类的诊断准确性 (Ramawamy 等人, 2001 年; Su 等人, 2001 年; Vapnik, 1998 年; Yeang 等人, 2001 年)。简而言之, SVM 算法 (盲法) 通过与所有其他

样本 (175 - 1) 进行比较, 将每个样本分类为癌症或健康样本, 并对所有样本进行 175 次分类和交叉验证。算法

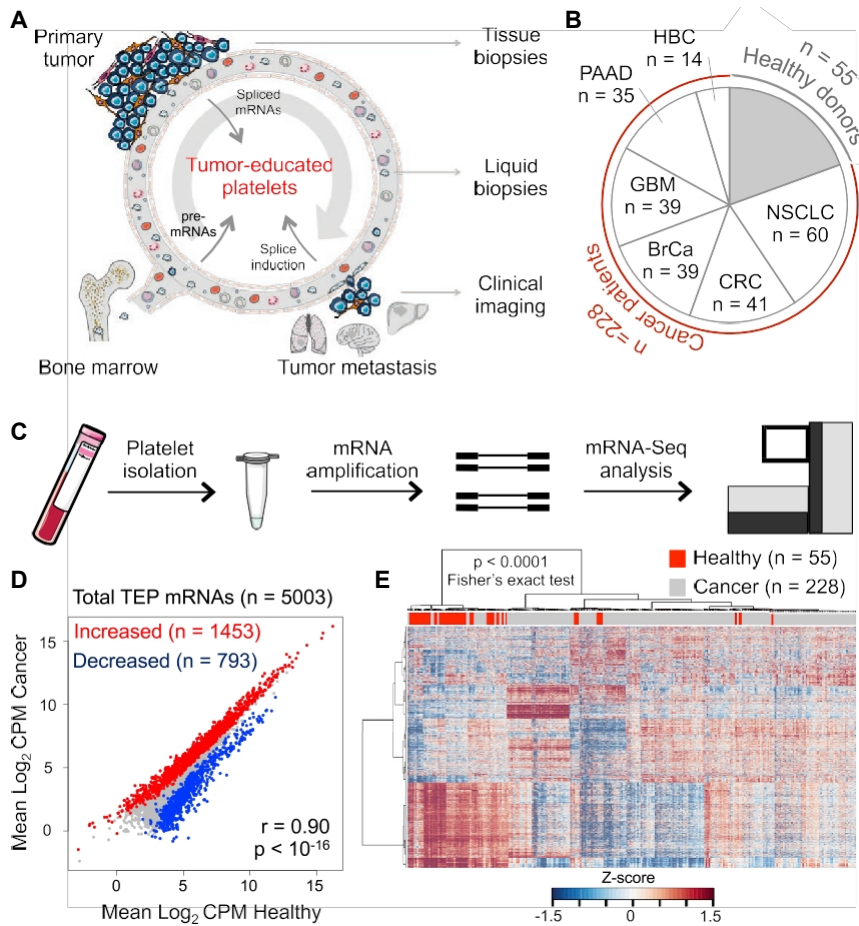


图 1.用于泛癌症诊断的肿瘤诱导血小板 mRNA 图谱分析

(A) 作为液体活检生物源的 肿瘤教育平板 (TEPs) 示意图概览。

(B) 健康捐献者和不同类型癌症患者的血小板样本数量。

(C) TEP mRNA 测序 (mRNA-seq) 工作流程, 从 6 毫升 EDTA 涂层试管开始, 到血小板分离、mRNA 扩增和测序。

(D) 健康捐献者 (HD) 血小板和癌症患者 TEP 中检测到的 mRNA 的相关图, 包括突出显示的增加 (红色) 和减少 (蓝色) 的 TEP mRNA。

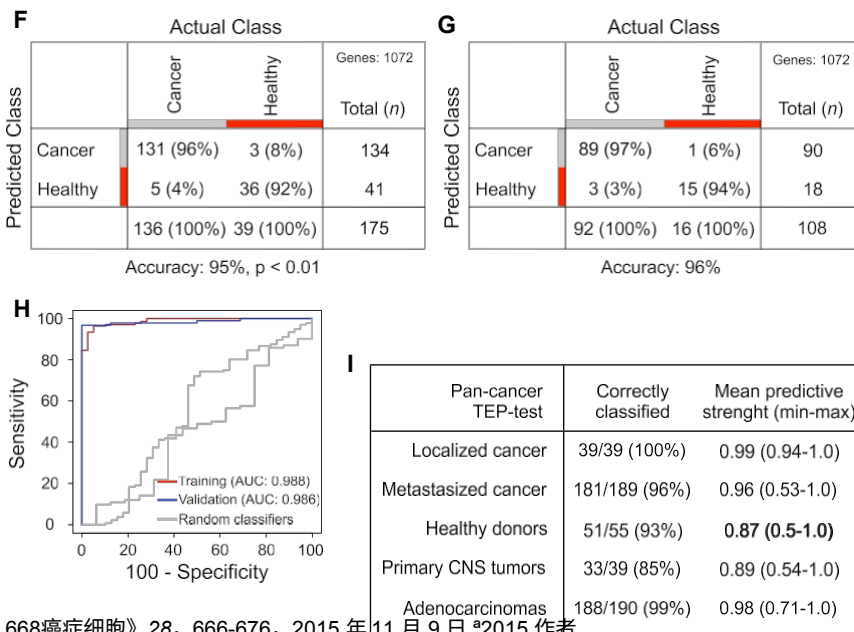
(E) 健康捐献者 (红色) 和癌症患者 (灰色) 的血小板 mRNA 无监督聚类热图。

(F) 对健康捐献者和癌症患者进行泛癌症 SVM/LOOCV 诊断的交叉表 (n = 175)。所示为样本数和检出率 (以百分比表示)。

(G) 泛癌症 SVM 算法在验证队列 (n = 108) 中的表现。以百分比表示样本数和检出率。

(H) SVM 诊断训练组 (红色)、验证组 (蓝色) 和随机分类器的 ROC 曲线, 显示训练组和验证组 (灰色) 偶然获得的分类准确率。

(I) 五个亚组 SVM 分类的总准确率, 包括相应的预测强度。基因, SVM 算法训练中包含的 mRNA 数量。另见图 S1 和表 S1、S2、S3 和 S4。



的泛癌症 TEP 分类器  
我们选择了 1,072 条 RNA (表 S4)

在 n = 175 的训练队列中, 灵敏度为 96%, 特异度为 92%, 准确率为 95% (图 1F)。使用未参与输入基因列表选择和算法训练的单独验证队列 (n = 108) 进行后续验证, 结果灵敏度为 97%, 特异度为 94%, 准确度为 95% (图 1F)。良品率为 96% (图 1G), 而非

图 1H) 和高预测强度 (图 1I)。与此相反, 在 SVM 训练过程中通过多轮随机洗牌标签 (permutation) 确定的随机分类器 (见实验步骤) 没有预测能力 (平均总准



确率：78%，SD  $\pm$  0.3%， $p < 0.01$ ），  
从而显示出研究队列中两组的不平衡代表性、

为了确定分类算法的特定输入基因列表，我们对差异进行了方差分析（如 R 软件包 edgeR 所实现的）。为了确定分类算法的特定输入基因列表，我们进行了方差分析差异测试（在 R 软件包 edgeR 中实现），得出了分类器特定基因列表（表 S4）。对于特定算法

我们的程序具有特异性。在训练集和验证集（比例为 60:40）中对整个数据集进行了 100 次随机分类比例子采样，得到了相似的准确率（平均总准确率：96%，标度： $\pm$  2%），证实了该数据集的分类准确性具有可重复性。值得注意的是，所有 39 名

表 1 患者特征摘要

患者小	总计 (n)		性别 男 (%) <sup>a</sup>		年龄 (SD) <sup>b</sup>		转移 (%) 培训验证		突变	存在率 (%) 培训验证	
高清	39	16	21 (54)	6 (38)	41 (13)	38 (16)	-	-	-	-	-
GBM	23	16	18 (78)	10 (63)	59 (16)	62 (14)	0 (0)	0 (0)	-	-	-
NSCLC	36	24	14 (39)	14 (58)	60 (11)	59 (12)	33 (92)	23 (96)	KRAS	15 (42)	11 (46)
									表皮生长因子受体	14 (39)	7 (29)
									MET-过表达	5 (14)	3 (13)
儿童权利委员会	25	16	13 (52)	9 (56)	59 (13)	63 (16)	20 (80)	15 (94)	KRAS	7 (28)	8 (50)
PAAD	21	14	12 (57)	7 (50)	66 (9)	66 (10)	15 (71)	9 (64)	KRAS	13 (62)	9 (64)
BrCa	23	16	0 (0)	0 (0)	59 (11)	59 (11)	16 (70)	9 (56)	HER2+	7 (30)	5 (31)
									PIK3CA	6 (26)	2 (13)
									三阴性	5 (22)	3 (19)
HBC	8	6	6 (75)	2 (33)	68 (13)	62 (16)	6 (75)	4 (67)	KRAS	3 (38)	1 (17)

<sup>b</sup>表示的是平均年龄（岁）。

在 39 名中枢神经系统原发性肿瘤患者中，有 33 人被正确分类为癌症患者（图 1I）。对之前在各种非癌症患者血小板中发现的 RNA 水平不同的 22 个基因（Gnatenko 等人，2010 年；Healy 等人，2006 年；Lood 等人，2010 年；Raghavachari 等人，2007 年）进行可视化后发现，在我们的 TEP 数据集中，这些基因的水平参差不齐（图 S1F），这表明非癌症患者的血小板 RNA 再序列与癌症患者不同。

血小板的肿瘤特异性教育程序可进行多类癌症诊断

除了泛癌症诊断外，TEP mRNA 图谱还能区分健康供体和特定类型的癌症患者，健康供体和所有六种肿瘤类型（即 NSCLC、CRC、GBM、PAAD、BrCa 和 HBC）的血小板 mRNA 水平差异的无监督分层聚类就证明了这一点。NSCLC、CRC、GBM、PAAD、BrCa 和 HBC（图 2A，所有  $p < 0.0001$ ，Fisher's exact 检验和 S2A；表 S5），并由此产生了肿瘤特异性基因列表，作为肿瘤特异性算法的训练和验证输入（表 S4）。在对全女性 BrCa 患者组进行无监督聚类时，排除了男性健康供体，以避免因性别特异性血小板 mRNA 图谱造成的样本偏差（图 S2B）。基于 SVM 对健康供体的所有单个肿瘤类别进行分类的结果是，在训练和验证队列中，两组均有明确的区分，灵敏度和特异性都很高，38/39（97%）名局部疾病的癌症患者被正确分类（图 2B

和 S2C）。笼式 GO 分析表明，各肿瘤类型的 TEPs 之间存在生物过程差异，这表明存在肿瘤特异性 "教育" 计划（表 S6）。我们没有检测到足够的 mRNA 水平差异来区分非转移性肿瘤患者和转移性肿瘤患者，这表明血小板谱的改变主要受分子肿瘤类型的影响，其次受肿瘤进展和转移的影响。

接下来，我们确定了能否通过分析 TEP 档案来区分胃肠道中三种不同类型的腺癌，即 CRC、PAAD 和 HBC。我们开发了一种 CRC/PAAD/HBC 算法，该算法对混合 TEP 样本（ $n = 90$ ）进行了正确分类，总体准确率为 76%（随机分类器的平均总体准确率为 42%，SD:  $\pm 5\%$ ，HBC:  $\pm 5\%$ ）：42%，SD:  $\pm 5\%$ ， $p < 0.01$ ，图 2C）。为了确定 TEP mRNA 图谱是否允许对所有肿瘤类型和健康供体进行多类癌症诊断，我们扩展了 SVM/ LOOCV 分类测试，使用多种算法将训练队列中的每个样本（ $n = 175$ ）分类为健康供体或六种肿瘤类型之一（图 S2D 和 S2E）。多类癌症诊断测试结果的平均准确率为 71%（随机分类器的平均总体准确率：19%，SD:  $\pm 2\%$ ， $p < 0.01$ ，图 2D），表明血小板 mRNA 图谱具有显著的多类癌症鉴别能力。基于 SVM 的多类分类器的分类能力在 108 个样本的验证队列中得到了证实，总体准确率为 71%（图 2E）。71% 的总体准确率可能不足以用于癌症诊断。不过，在最初分类错误的样本中，根据 SVM 算法选择的最强分类强度排名第二的分类在 60% 的情况下是正确的。这样，综合第一和第二级分类的总体准确率为 89%。HBC 样本的验证得分较低，可能是由于样本数量较少，也可能是由于这类癌症（肝细胞癌和胆管癌）的异质性。

代用 TEP mRNA 肿瘤标志物可反映肿瘤组织生物标志物 血液为检测用于治疗选择的伴随诊断生物标志物提供了一种前景广阔的生物资源（Bette-gowda 等，2014 年；Crowley 等，2013 年；Papadopoulos 等，2006 年）。我们选择了在匹配的肿瘤组织中证实有明显治疗指导标记物的患者血小板样本。

表 2.泛癌症 CAGE 基因本体

	25 大 GO 相关性		
	#	最低 <sup>a</sup>	最高 <sup>a</sup>
下降			
翻译	10	-0.865	-0.890
免疫, T 细胞	5	-0.853	-0.883
癌症相关	2	-0.875	-0.887
病毒复制	2	-0.875	-0.878
IL 1信号	2	-0.869	-0.874
RNA 处理	1	-0.886	
Ago2-Dicer-沉默	1	-0.882	
蛋白质代谢	1	-0.879	
受体处理	1	-0.869	
向上			
癌症相关	6	-0.783	-0.906
感染	3	-0.798	-0.853
HDAC	3	-0.795	-0.852
血小板	3	-0.837	-0.906
细胞骨架	2	-0.801	-0.886
低氧	2	-0.763	-0.937
蛋白酶	1	-0.854	
免疫缺陷	1	-0.812	
差异化	1	-0.810	
免疫分化	1	-0.801	
甲基化	1	-0.778	
新陈代谢	1	-0.768	

血小板-mRNA图谱与3,875个Broad研究所策划的基因组的顶级相关性。  
CAGE, 基因组富集相关性分析; GO, 基因本体论; #, 每个注释的命中数; IL, ； HDAC, 组蛋白去乙酰化酶。  
<sup>a</sup>所示为每个注释的最低和最高相关性。

虽然血小板mRNA图谱中含有检测不到或含量较低的这些突变生物标记物, 但TEP mRNA图谱确实能将PAAD、CRC、NSCLC和HBC患者中的KRAS突变肿瘤与KRAS野生型肿瘤以及NSCLC患者中的表皮生长因子受体 (EGFR) 突变肿瘤区分开来 (与随机分类器相比, 所有P < 0.01, 图3A- 3E; 表S4)。尽管分析的样本数量相对较少, 而且需要考虑算法过拟合的风险, 但TEP 图谱仍能区分 HER2-扩增、PIK3CA 突变或三阴性 BrCa患者以及 MET 过表达的 NSCLC 患者 (与随机分类器相比, 所

有 p < 0.01, 图 3F-3I)。

随后, 我们在 Illumina Miseq 平台上使用前瞻性收集的局部或转移性癌症患者血液样本, 比较了 TEP mRNA 分类方法与 KRAS (外显子 12 和 13) 和表皮生长因子受体 (外显子 20 和 21) 扩增子深度测序策略 (约 5,0003 个覆盖率) 的诊断准确性。这种方法确实能在血浆 DNA 和血小板 RNA 中检测到单个突变 KRAS 和表皮生长因子受体序列 (表 S7), 这表明突变肿瘤 RNA 生物标记物在 TEPs 中被封存并具有潜在的教育能力。突变 KRAS 被去

在原发性肿瘤组织中存在 *KRAS* 突变的患者中，分别有 62% 和 39% 的血浆 DNA ( $n = 103$ , kappa 统计量 = 0.370,  $p < 0.05$ ) 和血小板 RNA ( $n = 144$ , kappa 统计量 = 0.213,  $p < 0.05$ ) 检测到了 *KRAS* 突变。血浆DNA检测的灵敏度相对较低，其他报道 (Bettegowda等人, 2014年; Thierry等人, 2014年) 也提到了这一点，其部分原因可能是血样储存时间相对较长导致血浆DNA质量下降 (EDTA血样在血浆分离前在室温下储存长达48小时)。与血浆 DNA 和血小板 RNA 的 *KRAS* 扩增片段测序分析相比，要区分血液中的 *KRAS* 突变型和野生型肿瘤，TEP mRNA 图谱与组织分子状态的一致性更高 (kappa 统计 = 0.795-0.895,  $p < 0.05$ ) (表 S7)。因此，TEP mRNA 图谱可以利用潜在的基于血液的肿瘤组织生物标记物替代标志物，对癌症患者进行分层和治疗选择。

TEP 图谱为癌症患者的血基液体活检提供了一体化的生物资源明确区分肿瘤的原发性和转移性可能很困难，而且会妨碍适当的治疗选择。由于 TEP 图谱与根据原发器官确定的不同肿瘤类型非常相似，与全身扩散无关，因此有可能用于器官特异性癌症诊断。因此，我们选择了所有健康供体和肺部 (154 例)、脑部 (114 例) 或肝脏 (127 例) 原发性或转移性肿瘤患者。我们进行了 "器官检查"，并指导 SVM/LOOCV 算法确定肺部、脑部和肝脏是否存在癌症 (准确率分别为 96%、91% 和 96%)，将癌症细分为原发性或转移性肿瘤 (准确率分别为 84%、93% 和 90%)，并在出现转移的情况下确定潜在的起源器官 (准确率分别为 64%、70% 和 64%)。血小板 mRNA 原文件能准确地将癌症分配到不同的器官 (图 4)。此外，使用相同的 TEP mRNA 图谱，我们还能再次确定肿瘤组织的生物标记物状态 (准确率分别为 90%、82% 和 93%) (图 4)。

## 讨论

使用血基液体活检来检测、诊断和监测癌症，可以更早地诊断癌症，通过定制分子靶向治疗来降低成本，为癌症患者提供更多便利，并最终辅助临床肿瘤决策。目前正在评估的基于血液的生物资源在癌症诊断中表现出不理想的灵敏度，尤其是在局部疾病患者中。迄今为止，包括血浆 DNA、外泌体和 CTCs 在内的所有血液生物资源都未被用于多类癌症诊断 (Alix-Panabieres 和 Pantel, 2014 年; Bettegowda 等人, 2014 年; Skog 等人, 2008 年)，这阻碍了其在早期癌症检测中的应用。在此，我们报告了血小板 mRNA 分子检测可为所有癌症患者

提供有价值的诊断信息，这些患者横跨六种不同的肿瘤类型。我们的研究表明，血小板可作为一种

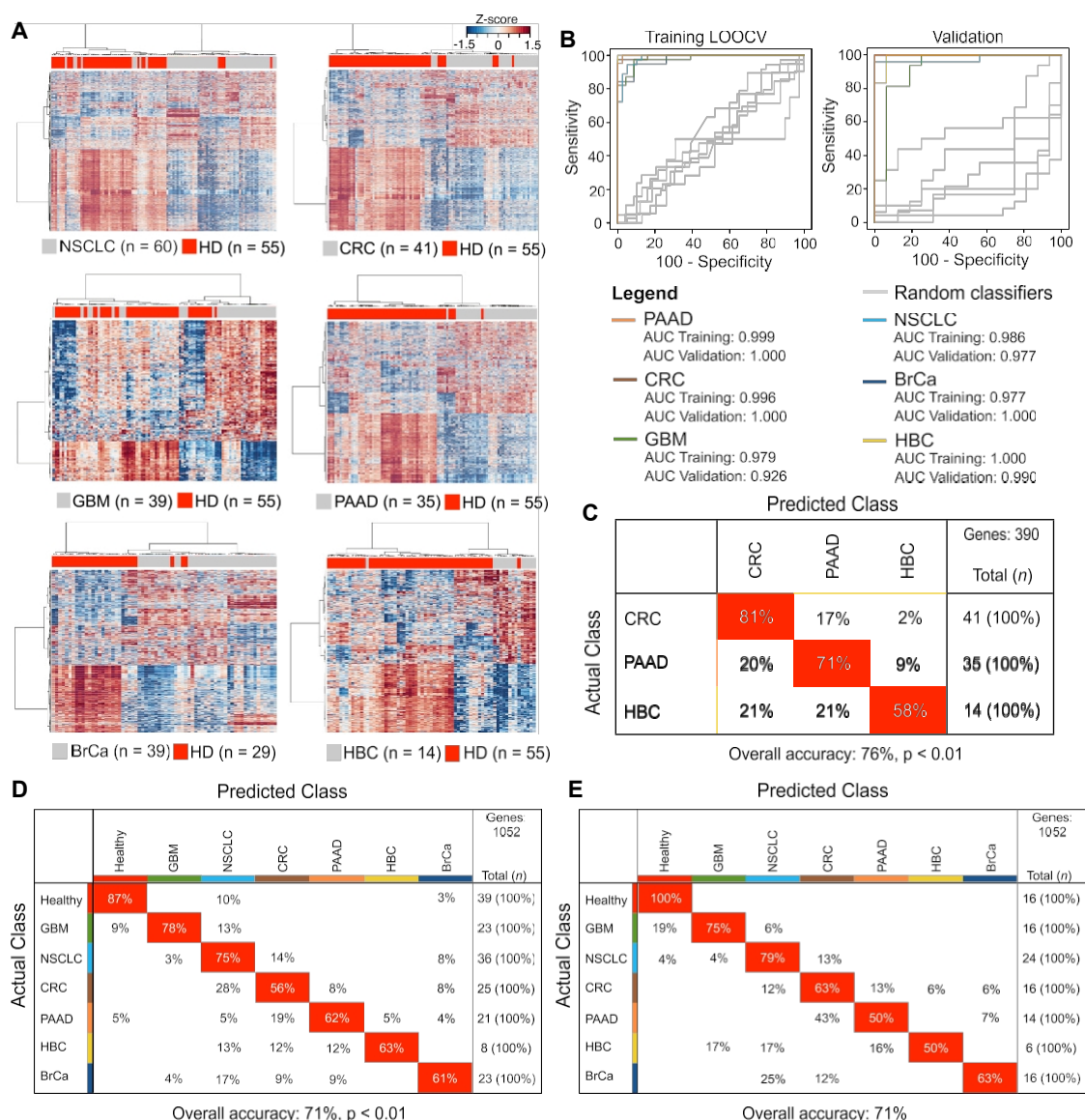


图 2.用于多类癌症诊断的肿瘤诱导血小板 mRNA 图谱

(A) 健康供体 (HD;  $n = 55$ ) (红色) 和非小细胞肺癌 (NSCLC;  $n = 60$ )、结直肠癌 (CRC;  $n = 41$ )、胶质母细胞瘤 (GBM;  $n = 39$ )、胰腺癌 (PAAD,  $n = 35$ )、乳腺癌 (BrCa;  $n = 39$ ; 女性 HD;  $n = 29$ ) 和肝癌 (HBC;  $n = 14$ ) 患者的血小板 mRNA 无监督聚类热图。

(B) SVM 诊断健康供体和单个肿瘤类别的训练 (左) 和验证 (右) 队列的 ROC 曲线。灰色显示的是 随机分类器, 表示偶然获得的分类准确率。

(C) 对 CRC、PAAD 和 HBC 患者进行多类 SVM/LOOCV 诊断的混淆矩阵。所示为与实际类别相比的检出率 (百分比)。

(D) 由健康捐献者 (健康) 和 GBM、NSCLC、PAAD、CRC、BrCa 和 HBC 患者组成的训练队列的多类 SVM/LOOCV 诊断的混淆矩阵。所示为与实际类别相比的检出率 (百分比)。

(E) 多分类 SVM 算法在验证队列 ( $n = 108$ ) 中的混淆矩阵。所示为样本数和检出率 (百分比)。基因, SVM 算法训练中包含的 mRNA 数量。

另见图 S2 和表 S4、S5 和 S6。

一体化生物资源可广泛扫描一般癌症的分子踪迹, 并对肿瘤类型  
和分子亚类提供强有力的指示。这包括患有局部疾病的患者, 可  
利用常规临床诊断对每种特定肿瘤类型进行有针对性的诊断确认

自从在癌症患者的血液中发现循环肿瘤物质 (Leon 等人, 1977 年) 并认识到基于血液的液体活检的临床效用以来, 大量研究对血液在癌症诊断、预后和治疗监测中的应用进行了评估 (Alix-Panabières 等人, 2012 年; Bidard 等人, 2014 年; Crowley 等人, 2013 年; Haber 等人, 2014 年)。

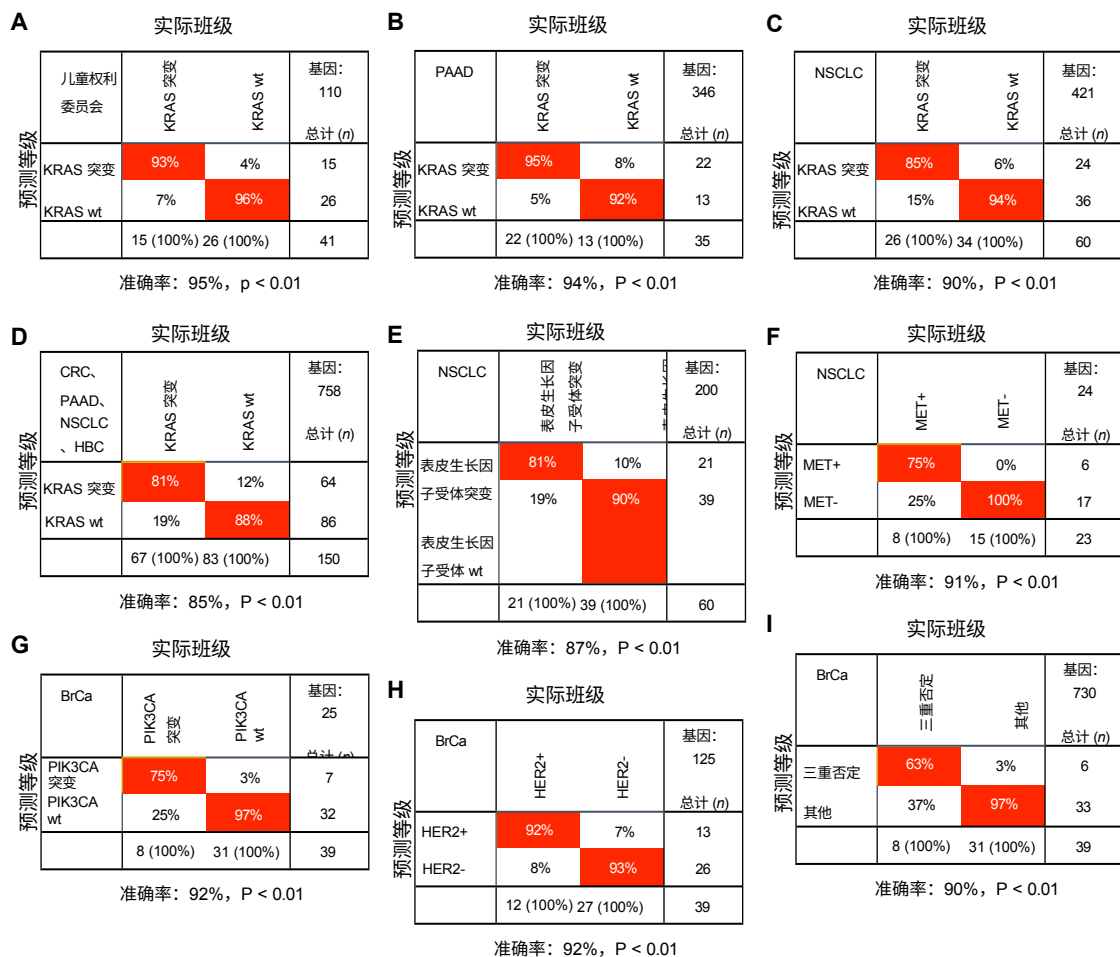


图 3.用于分子途径诊断的肿瘤诱导血小板 mRNA 图谱

SVM/LOOCV 诊断与 (A) CRC、(B) PAAD 和 (C) NSCLC 患者的分子标记 KRAS, (D) CRC、PAAD、NSCLC 或 HBC 患者联合队列中的 KRAS, (E) NSCLC 患者中的 EGFR 和 (F) MET, (G) BrCa 患者中的 PIK3CA 突变、(H) HER2-扩增和 (I) 三阴性状态的交叉表。基因, SVM 算法训练中包含的 mRNA 数量。另见表 S4 和 S7。

和 Velculescu, 2014 年)。随着高灵敏度靶向检测方法的发展, 如靶向深度测序 (Newman 等人, 2014 年)、液滴数字 PCR (Bettegowda 等人, 2014 年) 和等位基因特异性 PCR (Maheswaran 等人, 2008 年; Thierry 等人, 2014 年), 液体活检在临床应用中的效用和适用性也在加速提高。此前, 这些进展允许对各种生物资源进行泛癌症比较, 并发现在超过 75% 的癌症中, 包括晚期胰腺癌、结直肠癌、乳腺癌和卵巢癌, 都能检测到无细胞 DNA, 尽管检测率取决于肿瘤的等级和分析深度 (Bettegowda 等人, 2014 年)。在这里, 我们发现几乎所有癌症患者的血小板 RNA 图谱都会受到影响, 无论肿瘤类型如何, 尽管肿瘤相关 RNA 的丰度在癌症患者中似乎各不相同。此外, 在 676 癌症细胞》28, 666-676, 2015 年 11 月 9 日 \*2015 作者

88% 的局部 KRAS 突变癌症患者中, 组织生物标记物的替代 RNA 共轭标志物也是通过肿瘤特定和泛癌症 SVM/LOOCV 程序测定的, 这些标志物可从微量 (100-500 pg) 血小板 RNA 中轻易获得。由于在分离血小板颗粒之前, 全血可在室温下保存 48 小时, 同时还能保持高质量的 RNA 和主要的癌症 RNA 特征, 因此 TEPs 更易于在日常临床实验室实践中使用, 并有可能在进一步处理血液样本之前运送。



血小板广泛参与肿瘤生长和癌症进展 (Gay 和 Felding-Habermann, 2011 年)。血小板能整合溶解的肿瘤相关蛋白 (Klement 等人, 2009 年) 以及剪接和未剪接的 mRNA (Calverley 等人, 2010 年; Nilsson 等人, 2011 年), 同时血小板还能与肿瘤细胞 (Labelle 等人, 2011 年)、中性粒细胞 (Sreeramkumar 等人, 2014 年) 直接相互作用 (2011)、中性粒细胞 (Sreeramkumar 等人, 2014)、循环 NK 细胞 (Palumbo 等人, 2005; Placke 等人, 2012) 和循环肿瘤细胞 (Ting 等人, 2014; Yu 等人, 2013)。有趣的是, 体内实验显示, 乳腺癌通过向循环血小板提供促炎症和促血管生成蛋白, 支持休眠转移灶的生长, 从而介导全身性感染 (Kuznetsov 等人, 2012 年)。利用基因本体论方法 (CAGE), 我们将 TEP 癌症特征与可公开获得的数据集进行了关联。事实上, 我们发现了与癌症组织、缺氧、血小板特征和细胞骨架的广泛相关性, 这可能反映了 TEPs 的 "警戒" 和促癌状态。我们观察到与 RNA 翻译、T 细胞免疫和白细胞间信号转导相关的 RNA 呈强负相关, 这意味着 TEPs 对参与这些生物过程的 RNA 或这些 RNA 向蛋白质的协调转导的需求减少 (Denis 等人, 2005 年)。我们观察到, TEPs 中的肿瘤特异性教育程序主要受肿瘤类型的影响, 其次受肿瘤进展和转移的影响。虽然我们无法

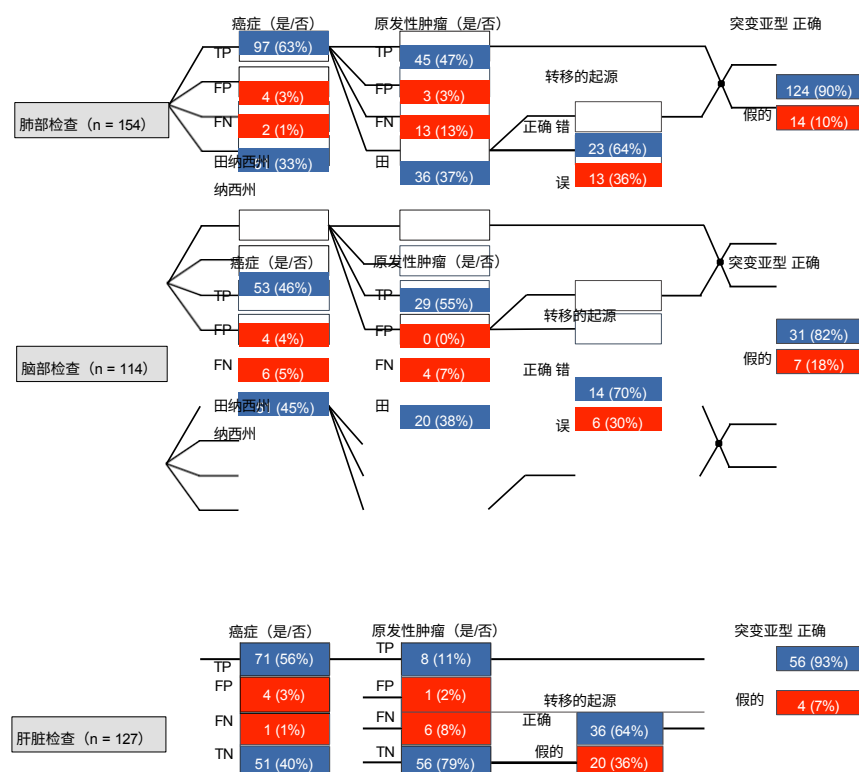


图 4.以器官为重点的 TEP Cancer 诊断对健康捐献者的 SVM/LOOCV 诊断 (n = 55)

以及原发性或转移性肿瘤患者肺部 (99 人; 共 154 项检测)、脑部 (62 人; 共 114 项检测) 或肝脏 (72 人; 共 127 项检测) 中的负荷, 以确定是否存在癌症, 将癌症细分为原发性肿瘤或转移性肿瘤, 如果是转移性肿瘤, 则确定原发器官, 并正确确定分子标记。值得注意的是, 在一些样本的突变亚型被纳入多个分类器 (即 *KRAS*、表皮生长因子受体、*PIK3CA*、*HER2*-扩增、*MET*-过表达或三重突变), 而一些样本的突变亚型被纳入多个分类器 (即 *KRAS*、表皮生长因子受体、*PIK3CA*、*HER2*-扩增、*MET*-过表达或三重突变)。

阴性), 说明突变检测的数量高于纳入样本的总数。TP, 真阳性; FP, 假阳性; FN, 假阴性; TN, 真阴性。所示为样本数量和检出率 (以百分比表示)。

我们不排除使用更大的样本集可以生成 SVM 算法, 该算法确实有能力区分某些阶段的癌症, 包括原位癌甚至恶性前病变。此外, 不同的分子肿瘤亚型 (如 *HER2*-扩增型 BrCa 与野生型 BrCa) 会对血小板谱产生不同的影响, 这可能是由不同分子肿瘤亚型产生的不同 "教育" 刺激造成的 (Koboldt 等人, 2012 年)。总之, 癌症患者血小板中的 RNA 含量取决于骨髓巨核细胞的转录状态 (Calverley 等人, 2010 年; McAllister 和 Weinberg, 2014 年), 并由剪接 RNA 的封存 (Nilsson 等人, 2011 年)、RNA 的释放 (Calverley 等人, 2010 年; McAllister 和 Weinberg, 2014 年) 完成、RNA 的释放 (Clancy 和 Freedman, 2014; Kirschbaum 等人, 2015; Rak 和 Guha, 2012; Risitano 等人, 2012), 可能还有血小板循环过程中的队列特异性 pre-mRNA 剪接。在成功治疗肿瘤后, 血小板前体部分或完全正常化将使基于 TEP 的疾病复发监测成为可能, 这需要对后续血小板样本进行分析。未来的研究将需要在 (小非编码) RNA (Laffont 等人, 2013 年; Landry 等人, 2009 年; Leidinger 等人, 2014 年; Lu 等人, 2005 年) 和蛋白质 (Burkhart 等人, 2014 年; Geiger 等人, 2013 年; Klement 等人

, 2009 年) 水平上解决肿瘤特异性 "有教养的" 特征问题, 并确定基因本体、基于血液的癌症分类的能力。

总之, 我们提供了强有力的证据, 证明血小板对多种癌症患者进行基于液体活检的分子诊断具有临床意义。还需要进一步验证, 以确定替代 TEP 图谱在基于血液的辅助诊断、治疗选择、纵向监测和疾病复发监测方面的潜力。此外, 我们还希望通过加入更多样本来进一步改进自学算法。对于这种方法, 应在抽血后 48 小时内从全血中分离出血小板部分, 然后冷冻血小板部分用于癌症诊断。此外, 未来的研究还应该探讨血小板减少的原因和抗血小板减少的作用。

本研究中发现的离群样本（如被归类为癌症患者的健康捐献者）可能会增加风险。慢性或一过性炎症性疾病、心血管事件和其他非癌症疾病等全身性因素也可能影响血小板 mRNA 图谱，需要在后续研究中进行评估，可能还包括预测为癌症的个体。

## 实验程序

### 样本采集和研究监督

所有患者和健康献血者的血液均在荷兰阿姆斯特丹 VU 大学医学中心或波士顿马萨诸塞州总医院（MGH）抽取，装在含有抗凝血剂 EDTA 的 6 毫升紫盖 BD Vacutainers 中。为尽量减少血小板在室温下长期储存的影响以及血小板 RNA 质量和数量的损失，样本在采血后 48 小时内处理。患者的血液样本是在术前采集的（GBM）或在门诊随访时采集的（CRC、NSCLC、PAAD、BrCa、HBC）。其中 9 例癌症患者样本是在首次采血后数月内采集的同一患者的随访样本（NSCLC 患者 5 例，PAAD 患者 2 例，BrCa 和 HBC 患者各 1 例）。局部疾病癌症患者是指医生或其他影像学或/病理学检查发现原发肿瘤没有向远处器官转移的癌症患者。胶质母细胞瘤是一种很少转移的肿瘤，其患者被视为晚期（高级别）癌症患者。训练队列和验证队列的样本是同时采集和处理的。通过组织 DNA 测序（包括新一代测序 SNaPShot）对患者的肿瘤组织进行分析，以确定是否存在基因改变。

>500，包括 *KRAS*、表皮生长因子受体和 *PIK3CA*（Dias-Santagata 等人，2010 年）。

非小细胞肺癌 FFPE 切片中 MET 过表达的评估采用免疫组化法（抗总 cMET SP44 Rabbit 单克隆抗体（mAb），Ventana 公司或 A2H2-3 抗人 MET mAb（Gruver 等人，2014 年））。BrCa 肿瘤组织的雌激素和孕激素状态以及 BrCa 肿瘤组织的 *HER2* 扩增情况分别采用免疫组织化学和荧光原位杂交法进行测定，并根据波士顿 MGH 的常规临床诊断方案进行评分。健康献血者在采血时或之前未被诊断出患有癌症。本研究按照《赫尔辛基宣言》的原则进行。研究获得了各医院审查委员会和伦理委员会的批准，并获得了所有受试者的知情同意。健康供体的临床随访

由于根据医院的伦理规定对这些样本进行了匿名处理，因此无法提供这些数据。

### 支持向量机分类器

对于二元（泛癌症）和多类样本分类，使用了由 e1071 R 软件包实现的支持向量机（SVM）算法。SVM 算法主要是确定所有样本在高维空间中的位置，其中每个轴代表一个转录本，特定转录本的样本表达水平决定了其在轴上的位置。在训练过程中，SVM 算法会根据每个类别中最接近超平面的样本与超平面的距离，绘制一个最能分隔两个类别的超平面。不同的样本类别必须分别位于超平面的两侧。然后，在高维空间中定位一个带有掩蔽类别标识的测试样本，并根据特定样本到所构建的超平面的距离“预判”其类别。对于多类 SVM 分类算法，采用的是单对单（OVO）方法。在这里，每个类别都与所有其他单独类别进行比较，因此 SVM 算法定义了多个超平面。为了对训练队列中的所有样本进行交叉验证，SVM 算法由训练队列中的所有样本减去一个进行训练，剩余样本用于（盲）分类。这一过程在所有样本中重复进行，直到每个样本都预测一次为止（leave-one-out cross-validation [LOOCV] 程序）。预测正确率作为分类器的准确率进行报告。为了评估 SVM 算法在独立数据集上的预测价值（该数据集之前未参与 SVM 训练过程，因此对该算法来说是全新的），该算法在训练数据集上进行了训练，所有 SVM 参数均固定不变，并对属于验证队列的样本进行了预测。此外，还执行了一个迭代（1003）过程，即在训练集和验证集中随机抽取数据集的样本（训练：验证=60:40（所有癌症类别）或 70:30（健康个体）的比例，每个样本类别的样本根据各个类别的总规模按此比例抽取（类别比例，分层抽样）），并报告所有单个分类的平均准确率。SVM 算法的内部性能可以通过启用 SVM 调整功能来提高，这意味着可以通过对用于算法训练的数据集进行随机子采样（“校内交叉验证”）来优化 SVM 算法的参数（伽马、成本）。在构建 SVM 算法之前，排除了低表达的转录本（在所有样本中均小于 5 个读数），并按照“补充实验-心理程序（转录本的差异表达）”中的描述对读数计数进行了归一化处理。对于每个预测个体，通过方差分析检验差异，进行特征选择（识别对预测性能有显著影响的转录本），得出分类器特异性输入基因列表（表 S4）。

<0.95（泛癌 KRAS），<0.90（CRC、PAAD 和 NSCLC KRAS 和 HER2-扩增 BrCa），<0.80（PIK3CA BrCa），<0.70（NSCLC 表皮生长因子受体），<0.50

（三阴性状态 BrCa）、<0.30（MET-过表达 NSCLC）、<0.10（CRC/PAAD/HBC）、<0.0001（多类肿瘤类型和单个肿瘤类别-健康）和 <0.00005（泛癌/健康癌）。为提高预测性能，对 SVM 进行了内部调整。所有单个肿瘤类别与健康供体和分子通路 SVM 算法都通过（默认）10 倍内部交叉验证进行了调整。泛癌症/健康癌症、多类肿瘤类型和胃肠道 CRC/PAAD/HBC SVM 算法通过 2 倍内部交叉验证进行调整。使用 LOOCV 方法分析了泛癌症和多类肿瘤类型、单个肿瘤类别与健康供体检测、胃肠道 CRC/PAAD/HBC 检测以及所有分子通路检测的训练队列。为了提高多类肿瘤类型检验的分类特异性，开发了额外的二元和多类分类算法，即泛癌检验（图 1F 和 1G）、HBC-CRC、HBC-PAAD、BrCa-CRC、BrCa-CRC-NSCLC 和 680 癌症细胞》28，666-676，2015 年 11 月 9 日 \*2015 作者

BrCa-HD-GBM-NSCLC 检验、

分别对训练组和验证组进行评估，结果如下

依次应用于多类肿瘤类型测试。所有预测为补充分类器任一条件的样本都将使用过滤器重新评估。后一种肿瘤分类被视为后续分类。此外，根据性别特异性 RNA（图 S2B），预测为全女性乳腺癌的样本为男性样本、

而在泛癌症测试中,被预测为健康的样本和被预测为癌症的样本会被自动分配到预测强度第二的类别,并由 SVM 输出作为补充。为了确定分类器偶然获得的准确率,对 SVM 算法用于训练的样本的类别标签进行了随机排列("随机分类器")。对 100 个 LOOCV 分类程序进行了这一处理。P 值是通过计算与观察到的总准确率相比总准确率相似或更高的整体随机分类器 LOOCV 分类准确率来确定的。预测强度还被用作生成接收者操作曲线(ROC)的输入,该曲线在 R 软件包 pROC (1.7.3 版)中实现。器官检查是根据训练队列的 SVM/LOOCV 结果汇编和验证队列的后续预测结果计算得出的,共涉及 283 个样本。对泛癌症二元 SVM、多类 SVM 和所有分子通路 SVM 算法进行了单独处理。每个器官检查(所有健康捐献者、在特定器官中患有原发性肿瘤的所有样本以及已知转移到特定器官的所有样本)所包含的样本都被选中。只有在器官检查的某一级预测正确的样本才会进入下一级评估。在"突变亚型"阶段的正确预测和错误预测计数是根据所有分子通路 SVM 算法确定的,所选样本都包含在这些算法中。

**基因组富集相关性分析** 基因组富集相关性分析(CAGE)是在在线平台 R2 (R2.amc.nl)上进行的。为了能在基于微阵列的统计平台上分析 RNA 测序读数计数,以测序批次和样本组为变量,对每百万归一化读数计数进行了空格转换,并上传到 R2 环境中。与其他类别相比,某一肿瘤类型或所有肿瘤类别(泛癌)的高度相关 mRNA (FDR < 0.01) 被用来生成类别特异性基因特征。随后,将这些单个特征与布罗德研究所(Broad Institute)提供的 3,875 个策划基因集进行相关性分析(<http://www.broadinstitute.org/gsea>)。排名前 25 位的相关性由两名独立研究人员(M.G.B.和 B.A.W.)手动注释,经双方研究人员同意后报告共享注释术语。

## 编号

本文报告的原始测序数据的登录号为 GEO: GSE68086。

## 补充资料

补充信息包括补充实验程序、两幅图和七张表,可与本文一起在 <http://dx.doi.org/10.1016/j.ccell.2015.09.018> 上在线查阅。

## 作者贡献

M.G.B.、B.A.T.、P.W.和 T.W. 设计了这项研究并撰写了手稿。E.F.S.、D.P.N.、H.M.V.、J.C.R.和 B.A.T. 提供了临床样本。M.G.B.、N.S.、J.T.、F.R.、P.S.、J.D.、B.Y.、H.V.和 E.P. 进行了 mRNA-seq 的样本处理。R.J.A.N.、P.S.、H.V.、E.P.和 T.W. 设计并执行扩增子测序测定。M.G.B.、N.S.、I.K.、J.D.、B.A.W.、J.K.、N.A.、E.P.和 T.W. 进行了数据分析和解释。所有作者都对手稿提出了重要意见。

## 利益冲突

P.S.、H.V.、E.P.、R.J.A.N. 和 T.W. 是 thromboDx BV 的员工。R.J.A.N. 和 T.W. 是 thromboDx BV 公司的股东和创始人。

## 致谢

欧洲研究理事会 E8626 (R.J.A.N.、E.F.S.和 T.W.) 和 336540 (T.W.)、荷兰科学与技术组织 (Dutch Organisation of Technology and Technology, D.C., D.C. and T.W.)

科学研究 93612003 和 91711366 (T.W.)、荷兰癌症协会 (J.C.R.、H.M.V. 和 T.W.)、Stichting STOPHersentumoren.nl (M.G.B. 和 P.W.)、NIH/NCI CA176359 和 CA069246 (B.A.T.)、CFF Norrland (R.J.A.N.)和瑞典研究理事会 (R.J.A.N.)。我们感谢 Esther Drees、Magda Grabowska、Danijela Koppers-Lalic、Michiel Pegtel、Wessel van Wieringen、Phillip de Witt Hamer 和 W. Peter Vandertop。

收到: 2015 年 3 月 23 日

已修订: 2015 年 7 月 2 日

已接受: 2015年9月25日

发表于出版日期: 2015 年 10 月 29 日

# 参考文献

Akbani, R., Ng, P.K.S., Werner, H.M.J., Shahmoradgol, M., Zhang, F., Ju, Z., Liu, W., Yang, J.-Y., Yoshihara, K., Li, J., et al. (2014).泛癌症蛋白质组 透视癌症基因组图谱。Nat.Nat.5, 3887.

Alix-Panabie` res, C. and Pantel, K. (2014)。循环肿瘤细胞 研究的挑战。Nat.Rev. Cancer 14, 623-631.

Alix-Panabie` res, C., Schwarzenbach, H., and Pantel, K. (2012)。循环 肿瘤细胞和循环肿瘤 DNA。Annu.Rev. Med.63, 199-215.

Bettegowda, C., Sausen, M., Leary, R.J., Kinde, I., Wang, Y., Agrawal, N., Bartlett, B.R., Wang, H., Luber, B., Alani, R.M., et al. (2014)。早期和晚期人类恶性肿瘤中循环肿瘤 DNA 的 检测。Sci.Med.6, 224ra24.

Bidard, F.-C., Peeters, D.J., Fehm, T., Nole` , F., Gisbert-Criado, R., Mavroudis, D., Grisanti, S., Generali, D., Garcia-Saenz, J.A., Stebbing, J., et al. (2014)。转移性乳腺癌患者循环肿瘤细胞的临床有效性: 患者个体数据的汇总分析》(Lancet Oncol.Lancet Oncol.15, 406-414。

Bray, P.F., McKenzie, S.E., Edelstein, L.C., Nagalla, S., Delgrosso, K., Ertel, A., Kupper, J., Jing, Y., London, E., Loher, P., et al. (2013)。无核人血小板的复杂转录- tional landscape。BMC Genomics 14, 1.

Burkhardt, J.M., Gambaryan, S., Watson, S.P., Jurk, K., Walter, U., Sickmann, A., Heemskerck, J.W.M., and Zahedi, R.P. (2014)。蛋白质组学能告诉我们 关于血小板的什么? Circ.114, 1204-1219。

Calverley, D.C., Phang, T.L., Choudhury, Q.G., Gao, B., Oton, A.B., Weyant, M.J., and Geraci, M.W. (2010)。转移性肺癌中血小板基因表达的显著下调。临床. 3, 227-232.

Chan, K.C.A., Jiang, P., Chan, C.W.M., Sun, K., Wong, J., Hui, E.P., Chan, S.L..、

Chan, W.C., Hui, D.S.C., Ng, S.S.M., et al. (2013)。通过血浆 DNA 亚硫酸氢盐测序无创检测癌症相关的全基因组低甲基化和拷贝数畸 变。Proc.Natl.USA 110, 18761-18768.

Chen, L., Kostadima, M., Martens, J.H.A., Canu, G., Garcia, S.P., Turro, E., Downes, K., Macaulay, I.C., Bielczyk-Maczynska, E., Coe, S., et al. (2014)。人类血液原生体血统传承过程中的转录多样性 tors.Science 345, 1251033.

Clancy, L. 和 Freedman, J.E. (2014)。血栓形成的新范例: 血小板 RNA 转移的新型介质和生物标志物。J. Thromb.Thrombolysis 37, 12-16.

Crowley, E., Di Nicolantonio, F., Loupakis, F., and Bardelli, A. (2013)。液体活检 : 监测血液中的癌症遗传学。Nat.Rev. Clin.Oncol. 10, 472-484.

Dawson, S.-J., Tsui, D.W.Y., Murtaza, M., Biggs, H., Rueda, O.M., Chin, S.-F., Dunning, M.J., Gale, D., Forshew, T., Mahler-Araujo, B., et al. (2013)。分析 循环肿瘤 DNA 以监测转移性乳腺癌。N. Engl.368, 1199-1209.

Denis, M.M., Tolley, N.D., Bunting, M., Schwartz, H., Jiang, H., Lindemann, S., Yost, C.C., Rubner, F.J., Albertine, K.H., Swoboda, K.J., et al. (2005).Escaping the nuclear confines: signal-dependent pre-mRNA splicing in anucleate platelets.Cell 122, 379-391.

Dias-Santagata, D., Akhavanfard, S., David, S.S., Vernovsky, K., Kuhlmann, G., Boisvert, S.L., Stubbs, H., McDermott, U., Settleman, J., Kwak, E.L..、

等人 (2010 年)。人类肿瘤的快速靶向突变分析：指导个性化癌症医疗的临床平台。EMBO Mol.Med.2, 146-158.

Diehl, F., Schmidt, K., Choti, M.A., Romans, K., Goodman, S., Li, M., Thornton, K., Agrawal, N., Sokoll, L., Szabo, S.A., et al. (2008)。循环突变 DNA 评估肿瘤动态。Nat.Med.14, 985-990.

Gay, L.J. and Felding-Habermann, B. (2011)。血小板对肿瘤转移的贡献。Nat.Rev. Cancer 11, 123-134.

Geiger, J., Burkhart, J.M., Gambaryan, S., Walter, U., Sickmann, A., and Zahedi, R.P. (2013)。反应：血小板转录组与蛋白质组相关而非相关。血液 121, 5257-5258.

George, J.N. (2000)。血小板。柳叶刀》355 期, 1531-1539 页。

Gnatenko, D.V., Zhu, W., Xu, X., Samuel, E.T., Monaghan, M., Zarrabi, M.H., Kim, C., Dhundale, A., and Bahou, W.F. (2010)。使用遗传生物标志物的血小板增多类预测模型。血液 115, 7-14.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., et al. (1999)。癌症的分子分类：通过基因表达监测发现类别和预测类别。Science 286, 531-537.

Gruver, A.M., Liu, L., Vaillancourt, P., Yan, S.-C.B., Cook, J.D., Roseberry Baker, J.A., Felke, E.M., Lacy, M.E., Marchal, C.C., Szpurka, H., et al. (2014)。识别人肝细胞生长因子受体 (MET) 胞外结构域的高灵敏度和特异性鼠单克隆抗体的免疫组化应用。组织病理学》65, 879-896.

Haber, D.A. and Velculescu, V.E. (2014)。基于血液的癌症分析：循环肿瘤细胞和循环肿瘤 DNA。Cancer Discov.4, 650-661.

Han, L., Yuan, Y., Zheng, S., Yang, Y., Li, J., Edgerton, M.E., Diao, L., Xu, Y., Verhaak, R.G.W., and Liang, H. (2014)。泛癌症伪基因表达分析揭示了生物学和临床相关的肿瘤亚型。Nat.Comm.5, 3963.

Healy, A.M., Pickard, M.D., Pradhan, A.D., Wang, Y., Chen, Z., Croce, K., Sakuma, M., Shi, C., Zago, A.C., Garasic, J., et al. (2006)。血小板表达谱分析和临床验证髓系相关蛋白-14 是心血管事件的新型判定因素。Circulation 113, 2278-2284.

Hoadley, K.A., Yau, C., Wolf, D.M., Cherniack, A.D., Tamborero, D., Ng, S., Leiserson, M.D.M., Niu, B., McLellan, M.D., Uzunangelov, V., et al.; Cancer Genome Atlas Research Network (2014)。12种癌症类型的多平台分析揭示了原发组织内和原发组织间的分子分类。细胞》158 期, 929-944 页。

Hrdlickova, B., Kumar, V., Kanduri, K., Zhernakova, D.V., Tripathi, S., Karjalainen, J., Lund, R.J., Li, Y., Ullah, U., Modderman, R., et al. (2014)。位于自身免疫性疾病相关区域的长非编码 RNA 的表达谱揭示了免疫细胞类型的特异性。Genome Med.6, 88.

Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., et al. (2013)。12种主要癌症类型的突变景观和意义。自然》502, 333-339.

Kirschbaum, M., Karimian, G., Adelmeijer, J., Giepmans, B.N.G., Porte, R.J., and Lisman, T. (2015)。水平 RNA 转移介导血小板诱导的肝细胞增殖。Blood 126, 798-806.

Kissopoulou, A., Jonasson, J., Lindahl, T.L., and Osman, A. (2013)。人体血小板PolyA+ mRNAs和rRNA- depleted total RNA的下一代测序分析。PLoS ONE 8, e81809.

Klement, G.L., Yip, T.-T., Cassiola, F., Kikuchi, L., Cervi, D., Podust, V.,

Italiano, J.E., Wheatley, E., Abou-Slaybi, A., Bender, E., et al. (2009)。血小板主动封存血管生成调节因子。血液 113, 2835-2842.

Koboldt, D.C., Fulton, R.S., McLellan, M.D., Schmidt, H., Kalicki-Veizer, J., McMichael, J.F., Fulton, L.L., Dooling, D.J., Ding, L., Mardis, E.R., et al.; Cancer Genome Atlas Network (2012)。人类乳腺肿瘤的综合分子画像。自然》490 期, 61-70 页。

Kuznetsov, H.S., Marsh, T., Markens, B.A., Castanˆ o, Z., Greene-Colozzi, A., Hay, S.A., Brown, V.E., Richardson, A.L., Signoretti, S., Battinelli, E.M., and McAllister, S.S. (2012)。鉴定建立由促血管生成血小板和骨髓衍生细胞定义的肿瘤支持性大环境的管腔型乳腺癌。Cancer Discov.2, 1150-1165.



- Labelle, M., Begum, S., and Hynes, R.O. (2011)。血小板与癌细胞之间的直接信号传导诱导上皮-间充质样转化并促进转移》(Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis.*Cancer Cell* 20, 576-590。
- Laffont, B., Corduan, A., Ple´, H., Duchez, A.-C., Cloutier, N., Boilard, E., and Provost, P. (2013)。活化血小板可通过微颗粒将 mRNA 调控 Ago2 microRNA 复合物输送到内皮细胞。血液 122, 253-261。
- Landry, P., Plante, I., Ouellet, D.L., Perron, M.P., Rousseau, G., and Provost, P. (2009)。无核血小板中存在 microRNA 通路。Nat.Struct.Mol.16, 961-966。
- Leidinger, P., Backes, C., Dahmke, I.N., Galata, V., Huwer, H., Stehle, I., Bals, R., Keller, A., and Meese, E. (2014)。肺癌患者和健康对照组血细胞亚组的 miRNome 分析。Oncotarget 5, 9484-9497。
- Leon, S.A., Shapiro, B., Sklaroff, D.M., and Yaros, M.J. (1977)。癌症患者血清中的游离 DNA 与治疗效果》(Free DNA in the serum of cancer patients and the effect of therapy.*Cancer Res.* 37, 646-650。
- Leslie, M. (2010)。细胞生物学。凝血之外：血小板的力量》。科学》328 期，562-564 页。
- Lood, C., Amisten, S., Gullstrand, B., Jo´n sen, A., Allhorn, M., Truedsson, L., Sturfelt, G., Erlinge, D., and Bengtsson, A.A. (2010)。系统性红斑狼疮患者的血小板转录谱和蛋白质表达：I 型干扰素系统的上调与血管疾病密切相关。血液 116, 1951-1957。
- Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., et al. (2005)。MicroRNA 表达谱可对人类癌症进行分类。自然》435 期, 834-838 页。
- Maheswaran, S., Sequist, L.V., Nagrath, S., Ulkus, L., Brannigan, B., Collura, C.V., Inserra, E., Diederichs, S., Iafrate, A.J., Bell, D.W., et al. (2008)。循环肺癌细胞中表皮生长因子受体突变的检测》(Detection of mutations in EGFR in circulating lung-cancer cells.*N. Engl.J.* 359, 366-377。
- McAllister, S.S. and Weinberg, R.A. (2014)。肿瘤诱导的全身环境是癌症进展和转移的关键调节因素。Nat.16, 717-727。
- Murtaza, M., Dawson, S.-J., Tsui, D.W.Y., Gale, D., Forshew, T., Piskorz, A.M., 、
- Parkinson, C., Chin, S.-F., Kingsbury, Z., Wong, A.S.C., et al. (2013)。通过血浆 DNA 测序对癌症治疗获得性抗药性的非侵入性 分析。Nature 497, 108-112。
- Newman, A.M., Bratman, S.V., To, J., Wynne, J.F., Eclöv, N.C.W., Modlin, L.A., Liu, C.L., Neal, J.W., Wakelee, H.A., Merritt, R.E., et al. (2014)。一种可广泛覆盖患者的超灵敏循环肿瘤 DNA 定量方法。Nat.Med.20, 548-554。
- Nilsson, R.J.A., Balaj, L., Hulleman, E., van Rijn, S., Pegtel, D.M., Walraven, M., Widmark, A., Gerrijsen, W.R., Verheul, H.M., Vandertop, W.P., et al. (2011)。血小板含有肿瘤衍生的 RNA 生物标记物。血液 118, 3680-3683。
- Palumbo, J.S., Talmage, K.E., Massari, J.V., La Jeunesse, C.M., Flick, M.J., Kombrinck, K.W., Jirouskova´, M., and Degen, J.L. (2005)。血小板和纤维蛋白 (ogen) 通过阻碍自然杀伤细胞介导的肿瘤细胞清除增加转移潜力。血液 105, 178-185。
- Papadopoulos, N., Kinzler, K.W., and Vogelstein, B. (2006)。同伴诊断在开发和
- 和使用突变靶向癌症疗法中的作用。Nat.Biotechnol.24, 985-995。
- Placke, T., O´r gel, M., Schaller, M., Jung, G., Rammensee, H.-G., Kopp, H.-G., and Salih, H.R. (2012)。血小板衍生的 MHC I 类赋予癌细胞一种假正常表型，从而颠覆了自然杀伤免疫细胞的抗肿瘤反应性。Cancer Res. 72, 440-448。
- Power, K.A., McRedmond, J.P., de Stefani, A., Gallagher, W.M., and Gaora, P.O. (2009)。高通量蛋白质组学检测人血小板中的新型剪接异构体。PLoS ONE 4, e5001。
- Quail, D.F. and Joyce, J.A. (2013)。肿瘤进展和转移的微环境调控。Nat.Med.19, 1423-1437。
- Rack, B., Schindlbeck, C., Ju´ckstock, J., Andergassen, U., Hepp, P., Zwingers, T., Friedl, T.W.P., Lorenz, R., Tesch, H., Fasching, P.A., et al;



SUCCESS 研究小组 (2014 年)。循环肿瘤细胞可预测早期中高风险乳腺癌患者的生存期。J. Natl. Cancer Inst. 106, dju066.

Raghavachari, N., Xu, X., Harris, A., Villagra, J., Logun, C., Barb, J., Solomon, M.A., Suffredini, A.F., Danner, R.L., Kato, G., et al. (2007)。血小板转录组扩增表达谱分析揭示镰状细胞病患者精氨酸代谢途径的变化。Circulation 115, 1551-1562.

Rak, J. 和 Guha, A. (2012)。细胞外囊泡--传播癌症基因的载体。生物论文 34, 489-497。

Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, C.H., Angelo, M., Ladd, C., Reich, M., Latulippe, E., Mesirov, J.P., et al. (2001)。使用肿瘤基因表达特征进行多类癌症诊断。Proc.Natl.美国 98, 15149-15154.

Ramskoˆ Id, D., Luo, S., Wang, Y.-C., Li, R., Deng, Q., Faridani, O.R., Daniels, G.A., Khrebtukova, I., Loring, J.F., Laurent, L.C., et al. (2012)。从单细胞水平的 RNA 和单个循环肿瘤细胞中获取全长 mRNA-Seq。Nat.Biotechnol.30, 777-782.

Risitano, A., Beaulieu, L.M., Vitseva, O., and Freedman, J.E. (2012)。血小板和类血小板颗粒介导细胞间 RNA 转移。血液 119, 6288-6295。

Rondina, M.T., Schwertz, H., Harris, E.S., Kraemer, B.F., Campbell, R.A., Mackman, N., Grissom, C.K., Weyrich, A.S., and Zimmerman, G.A. (2011)。败血症环境触发人血小板中拼接组织因子 mRNA 的表达。J. Thromb.Haemost.9, 748-758.

Rowley, J.W., Oler, A.J., Tolley, N.D., Hunter, B.N., Low, E.N., Nix, D.A., Yost, C.C., Zimmerman, G.A., and Weyrich, A.S. (2011)。人类和小鼠血小板转录组的全基因组 RNA-seq 分析。血液 118, e101-e111。

Schubert, S., Weyrich, A.S., and Rowley, J.W. (2014)。A tour through the transcriptional landscape of platelets.血液 124, 493-502。

Simon, L.M., Edelstein, L.C., Nagalla, S., Woodley, A.B., Chen, E.S., Kong, X., Ma, L., Fortina, P., Kunapuli, S., Holinstat, M., et al. (2014)。血小板组学揭示的与年龄和性别相关的人类血小板 microRNA-mRNA 网络。Blood 123, e37-e45.

Skog, J., Wuˆ rding, T., van Rijn, S., Meijer, D.H., Gainche, L., Sena-Esteves, M., Curry, W.T., Jr., Carter, B.S., Krichevsky, A.M., and Breakefield, X.O. (2008)。胶质母细胞瘤微囊运输 RNA 和蛋白质, 促进肿瘤生长并提供诊断生物标记物。Nat.10, 1470- 1476.

Sreeramkumar, V., Adrover, J.M., Ballesteros, I., Cuartero, M.I., Rossaint, J., Bilbao, I., Nacher, M., Pitaval, C., Radovanovic, I., Fukui, Y., et al. (2014)。中性粒细胞扫描活化血小板以引发炎症。Science 346, 1234-1238.

Su, A.I., Welsh, J.B., Sapinoso, L.M., Kern, S.G., Dimitrov, P., Lapp, H., Schultz, P.G., Powell, S.M., Moskaluk, C.A., Frierson, H.F., et al. (2001)。使用基因表达符号对人类癌症进行分子分类。Cancer Res. 61, 7388-7393.

Thierry, A.R., Mouliere, F., El Messaoudi, S., Mollevi, C., Lopez-Crapez, E., Rolet, F., Gillet, B., Gongora, C., Dechelotte, P., Robert, B., et al. (2014)。从循环肿瘤 DNA 中检测 KRAS 和 BRAF 突变的临床验证 lating tumor DNA.Nat.Med.20, 430-435.

Ting, D.T., Wittner, B.S., Ligorio, M., Vincent Jordan, N., Shah, A.M., Miyamoto, D.T., Aceto, N., Bersani, F., Brannigan, B.W., Xega, K., et al. (2014)。单细胞 RNA 测序确定了胰腺循环肿瘤细胞的细胞外基质基因表达。Cell Rep. 8, 1905-1918.

Vapnik, V.N. (1998).Statistical Learning Theory (Wiley).

Yeang, C.H., Ramaswamy, S., Tamayo, P., Mukherjee, S., Rifkin, R.M., Angelo, M., Reich, M., Lander, E., Mesirov, J., and Golub, T. (2001)。多种

肿瘤类型的分子分类。Bioinformatics 17 (Suppl 1), S316-S322.

Yu, M., Bardia, A., Wittner, B.S., Stott, S.L., Smas, M.E., Ting, D.T., Isakoff, S.J., Ciciliano, J.C., Wells, M.N., Shah, A.M., et al. (2013)。循环中的乳腺肿瘤细胞表现出上皮和间质组成的动态变化。Science 339, 580-584.