



ELSEVIER

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng



Research

New Technology of Tumor Diagnosis and Treatment—Article

从结直肠癌细胞和组织样本释放的外泌体中检测肿瘤相关抗原自身抗体的多路生物传感诊断平台展示了对结直肠癌的高诊断能力

Ana Montero-Calle ^a, Itziar Aranguren-Abeigon ^a, María Garranzo-Asensio ^{a,b}, Carmen Poves ^c,
 María Jesús Fernández-Aceñero ^d, Javier Martínez-Useros ^e, Rodrigo Sanz ^c, Jana Džiaková ^c,
 Javier Rodríguez-Cobos ^f, Guillermo Solís-Fernández ^a, Eloy Povedano ^g, María Gamella ^g,
 Rebeca Magnolia Torrente-Rodríguez ^g, Miren Alonso-Navarro ^a, Vivian de los Ríos ^h, J. Ignacio Casal ^h,
 Gemma Domínguez ^f, Ana Guzman-Aranguez ^b, Alberto Peláez-García ⁱ, José Manuel Pingarrón ^g,
 Susana Campuzano ^g, Rodrigo Barderas ^{a,*}

^a Chronic Disease Program (UFIEC), Carlos III Health Institute, Madrid E-28220, Spain

^b Department of Biochemistry and Molecular Biology, Faculty of Optics and Optometry, Complutense University of Madrid, Madrid E-28040, Spain

^c Gastroenterology Unit, San Carlos Clinical Hospital, Madrid E-28040, Spain

^d Surgical Pathology Department, San Carlos Clinical Hospital, Madrid E-28040, Spain

^e Translational Oncology Division, OncoHealth Institute, Jimenez Diaz Foundation University Hospital, Madrid E-28040, Spain

^f Department of Biochemistry, Faculty of Medicine, Alberto Sols Institute of Biomedical Research, CSIC-UAM, Madrid E-28029, Spain

^g Department of Analytical Chemistry, Faculty of Chemical Sciences, Complutense University of Madrid, Madrid E-28040, Spain

^h Center for Biological Research, Spanish National Research Council (CSIC), Madrid, E-28040, Spain

ⁱ Molecular Pathology and Therapeutic Targets Group, La Paz University Hospital (IdiPAZ), Madrid E-28046, Spain

ARTICLE INFO

Article history:

Received 30 September 2020

Revised 5 February 2021

Accepted 27 April 2021

Available online 14 August 2021

关键词

自身抗体
 诊断
 结直肠癌
 外泌体
 肿瘤微环境
 体液免疫应答
 即时医疗
 生物传感器

摘要

结直肠癌(CRC)是全球导致癌症相关死亡的第二大原因。CRC患者的5年生存率与诊断分期息息相关。CRC早期诊断患者的5年生存率高于80%,CRC晚期诊断患者的5年生存率低于10%。大量研究表明,患者血清中特定CRC自身抗原[肿瘤相关抗原(TAA)]对应的自身抗体有助于早期诊断。因此,本文旨在识别CRC特异性自身抗体的自身抗原靶点,借助液体活检技术有效筛查CRC患者和健康个体。为此,我们从CRC患者体内提取5个CRC细胞系和组织样本,通过免疫沉淀-质谱分析法测量其分泌的外泌体的蛋白质含量,共鉴定出103种蛋白质为潜在的CRC特异性自身抗原。采用生物信息学技术和荟萃分析,我们选定了15种与实际CRC自身抗原类似度更高的蛋白质,以便后续借助蛋白质印迹法(WB)和免疫组织化学技术(IHC)评估它们在CRC预后中的作用。结果发现,在患者的组织和血浆外泌体样本中,有11种蛋白质发生蛋白质水平失调,有9种蛋白质与CRC预后有关。经验证发现,除一例外,所有研究均显示出具有统计学意义的高诊断能力,采用荧光Halotag磁珠,或者借助多路生物传感平台(磁性微载体为固相载体,由用于CRC细胞检测的共价固定Halotag融合蛋白修饰),能够有效筛查区分CRC患者、癌前病变个体与健康个体。综上所述,本文的研究结果突出了此研究方法在识别慢性疾病特征TAA时的有效性;此外,本文采用的测量血浆中TAA对应的自身抗体水平的方法可以应用到具有高诊断能力的CRC检测即时医疗(POC)设备中。

© 2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author.

E-mail address: r.barderas@isciii.es (R. Barderas).

2095-8099/© 2021 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

英文原文: *Engineering* 2021, 7(10): 1393–1412

引用本文: Ana Montero-Calle, Itziar Aranguren-Abeigon, María Garranzo-Asensio, Carmen Poves, María Jesús Fernández-Aceñero, Javier Martínez-Useros, Rodrigo Sanz, Jana Džiaková, Javier Rodríguez-Cobos, Guillermo Solís-Fernández, Eloy Povedano, María Gamella, Rebeca Magnolia Torrente-Rodríguez, Miren Alonso-Navarro, Vivian de los Ríos, J. Ignacio Casal, Gemma Domínguez, Ana Guzman-Aranguez, Alberto Peláez-García, José Manuel Pingarrón, Susana Campuzano, Rodrigo Barderas. Multiplexed Biosensing Diagnostic Platforms Detecting Autoantibodies to Tumor-Associated Antigens from Exosomes Released by CRC Cells and Tissue Samples Showed High Diagnostic Ability for Colorectal Cancer. *Engineering*, <https://doi.org/10.1016/j.eng.2021.04.026>

1. 引言

结直肠癌 (colorectal cancer, CRC) 是全球第三大高发癌症, 也是第二大癌症导致的死亡原因[1-2]。60% 以上的 CRC 患者都确诊于晚期, 而此时肿瘤已扩散到附近的淋巴结 (III 期) 或远端器官 (IV 期), 主要包括肝和肺。在 III 期和 IV 期确诊 CRC 的患者的 5 年生存率分别低于 70% 和 10%, 而如果在早期阶段确诊, CRC 患者的 5 年生存率可提高到 80% 以上[3]。目前 CRC 的诊断技术是有创的, 或对疾病的特异性不够, 如结肠镜检查或粪便潜血试验。此外, 目前的生物标志物对预后有帮助, 但不能用于诊断, 例如, 在血浆样本中检测癌胚抗原 (CEA) [4-5]。因此, 有必要使用新的生物标志物对该疾病进行早期、特异性和无创性诊断。CRC 是由局限于肠黏膜的息肉 (腺瘤) 发展成腺癌的。首先, 肿瘤发展到肠道黏膜下层 (I 期), 然后进入固有肌层 (II 期), 最后, 它扩散到其他器官 (III 期和 IV 期) [4,6]。在 CRC 的不同阶段, 突变、表观遗传和基因改变会导致蛋白质改变和失调。其中一些改变的自身蛋白成为肿瘤相关抗原 (TAA), 在 CRC 患者体内诱导体液免疫反应[7-8]。因此, 抗 TAA 的自身抗体[免疫球蛋白 G (IgG)] 可以作为有用的生物标志物, 通过液体活检在疾病早期进行特异性检测[9-10]。

近年来, 多种蛋白质组学技术被用于鉴定 TAA, 重点是微阵列的应用, 如噬菌体或天然重组蛋白微阵列[11-13]。然而, 这些技术受到同时分析的样本数量或打印在阵列上的蛋白质数量的限制。蛋白质阵列只包括人类所有蛋白质的 50%~66%, 不包括翻译后修饰或蛋白质组。最近, 我们开发了一种新的方法来解决这些问题, 即使用从 CRC 细胞中提取的蛋白质鉴定 TAA。该方法包括免疫沉淀和液相色谱-串联质谱 (LC-MS/MS) 程序的组合, 然后是血清反应性试验[14]。

外泌体是细胞分泌的 30~150 nm 大小的细胞外囊泡。它们通过血管或淋巴管介导周围细胞之间以及与远端器官细胞之间的细胞信息传递[15-16]。据报道, 肿瘤源性外泌体 (TDE) 参与癌症发展和转移的各个环节, 包括诱导上皮-间充质转化的激活, 转移前生态位的形成, 以及受体细胞免疫反应的调节[17-19]。因 TED 组成、功能及其在体液中的分布, TED 是一种可用于特异性识别 CRC 的令人感兴趣的、几乎未被探索的 TAA 来源。

肿瘤微环境 (TME) 由癌细胞、各种驻留和浸润的宿主细胞、分泌因子和细胞外基质蛋白组成。基质滋养癌细胞, 促进肿瘤的生长、侵袭和转移。此外, TME 有助于调控 TAA 表达和免疫细胞浸润[20]。因此, 对包含癌细

胞和肿瘤基质的 CRC 肿瘤组织的分析, 应该识别出在培养的癌细胞分析中遗漏的肿瘤间质新抗原[14]。

由于近年来电化学生物传感器和生物传感方法取得了巨大进步, 这些技术作为非常合适的高性价比工具, 正在稳步发展。这些方法易于操作, 几乎不需要培训, 它们可以在现场 (如在非实验室环境或资源有限的地区) 进行多重和 (或) 多组分析, 使用少量的生物样本和最少的可操作时间, 提供准确可靠的定量结果[21-22]。这些设备最近被用于探索癌症生物标志物, 特别关注流行和高死亡率的肿瘤 (如 CRC), 以协助传统方法的早期诊断和分期。仅在过去两年中, 基于电化学亲和性的生物传感器就展示了独特的机会来确定 CRC 相关转录组生物标志物 (micro RNA)、表观遗传生物标志物[即 DNA 的全局和区域水平上的 5-甲基胞嘧啶 (5-mC) 和 5-羟甲基胞嘧啶 (5-hmC) 以及 RNA 的全局 m^6 -甲基腺苷 (m^6A), 即使具有单基敏感性], 以及蛋白质组学生物标志物 (如与 CRC 存在相关的蛋白质、受体或蛋白酶, 以及与肿瘤侵袭和进展相关的缺氧和转移事件) [23]。在这里, 我们的目标是通过使用免疫沉淀串联质谱的方法, 研究两个互补且几乎未被报道过的能产生 CRC 特征 TAA 的蛋白质, 扩大具有诊断价值的 CRC 特异性 TAA 的数量: ① CRC 患者早期 (I 期和 II 期) 的 TME 组织样本; ② CRC 患者晚期 (III 期和 IV 期) 的 TME 组织标本; ③ 由具有不同转移能力的 CRC 细胞系分泌的外泌体。从 CRC 患者中共鉴定出 103 种不同的蛋白质, 这些蛋白质可被 CRC 患者的 IgG 特异性识别。进一步选择其中 15 种进行临床验证: ① 生物信息学和荟萃分析, 以确定其潜在的失调和与 CRC 的关联; ② 以 CRC 不同分期患者血清、CRC 癌前病变患者血清和健康人群血清为对照, 检测其自身抗体, 确定其诊断价值; ③ Western blot (WB) 和免疫组化 (IHC) 分析其作为 CRC 预后标志物的作用。最后, 我们开发了一种电化学生物传感多路平台, 作为一种能够检测血浆中所有这些自身抗体的即时医疗 (POC) 设备。

2. 材料和方法

2.1. 结直肠癌细胞系

等基因 CRC 细胞系 SW480 和 SW620 取自美国模式培养物保藏所 (ATCC)。等基因细胞系 KM12C、KM12SM 和 KM12L4a 取自费德勒实验室 (MD 安德森癌症中心) [24-25]。根据既定方案, 在杜尔贝科改良的 Eagle 培养基 (DMEM; 龙沙, 瑞士) 中添加 10% 胎牛血清 (FBS; 西格玛奥德里奇, 美国)、1×L-谷氨酰胺 (龙沙) 和 1×青霉

素/链霉素（龙沙），以培养CRC细胞。

2.2. 组织和血浆样本

圣卡洛斯临床医院（马德里）卡洛斯三世健康研究所的机构伦理审查委员会批准了这项关于生物标志物发现和验证的研究（CEI PI 45）。组织和血浆样本来自圣卡洛斯临床医院（IdISSC）生物库，该生物库隶属于与欧洲区域发展基金 FEDER 基金共同成立的西班牙国家生物库网络（ISCII），并获得机构伦理审查委员会的批准。我们获得了所有患者的书面知情同意书。血浆和组织样本通过标准化的样本收集方案收集，并储存在 -80°C 下直至使用[9, 13,26]。

IP使用的是术前获得的12份来自III期和IV期CRC患者的血浆样本和12份结肠镜检查阴性且病理学未知的健康个体的血浆样本（表1和附录A中的表S1），以及来自I~IV期CRC患者的13个最佳切割温度（OCT）复合物包埋冷冻组织样本（表1和附录A中的表S2）。发光和生物传感血清反应性测定的样本，来自手术前（ $n=38$ ）和癌前病变（低级别或高级别腺瘤）（ $n=36$ ）的I~IV期CRC患者，以及来自结肠镜检查阴性且病理未知的健康个体（ $n=38$ ），总共112份血浆样本（表1和表S1）。

由具有胃肠道病理学专业知识的病理学家按照国际推荐标准对癌前病变和恶性结直肠癌变进行评估。根据异型增生的结构和细胞学可将癌前病变分为低级别或高级别腺瘤。恶性转变也被关注并按照Dukes或Kikuchi系统进行。

WB分析（表1和表S2）的样本来自17名I~IV期CRC患者的OCT包埋的冷冻肿瘤和配对非肿瘤组织、I~IV期CRC患者血浆中分离的17个外泌体（ $n=4$ ）、具有癌前病变的个体（ $n=4$ ），以及作为对照组的健康个体（ $n=9$ ）（表1和表S1）。

2.3. 外泌体分离和纯化

为了分离和纯化CRC细胞分泌的外泌体，每个细胞系接种8个175 cm²细胞培养瓶（康宁，美国）。细胞在 37°C 和5%二氧化碳（CO₂）下生长至90%融合。然后将细胞用1×磷酸盐缓冲液（PBS）洗涤3次，并在 37°C 和5% CO₂下与20 mL不含FBS的DMEM孵育1 h，以去除FBS和外源性外泌体。接下来将细胞用1×PBS洗涤3次，并与20 mL不含FBS的DMEM在 37°C 和5% CO₂下孵育48 h。然后将每个细胞系的160 mL条件培养基在 4°C 下以500g离心5 min以去除细胞碎片（沉淀），然后在 4°C 下再次以2000g离心10 min以去除大于1 μm的囊泡（颗粒）。之后收集澄清和条件培养基，并在 -80°C 下储存直至使用。

然后通过差速离心进行外泌体的纯化。条件培养基在 4°C 下以10 000g离心（Beckman-Coulter超速离心机，XL-100K，美国）30 min，以去除大小为1 μm的细胞外微泡（MV）。接下来，将条件培养基在 4°C 下以100 000g进一步离心70 min以进行外泌体沉淀。弃去上清液，将外泌体（沉淀）重悬于最终体积为1 mL 1×PBS的每个细

表1 研究中使用的CRC患者和健康个体的血浆样本信息

Samples	Number (<i>n</i>)	Age average ± SD (years)	Age range (years)	Gender (<i>n</i>)		Stage				
				Male	Female	I	II	III	IV	
Plasma samples										
IP	Healthy individuals	12	63.1 ± 20.8	35–83	4	8	—	—	—	—
	CRC patients	12	69.1 ± 8.7	56–85	7	5	—	—	9	3
Seroreactivity assays	Healthy individuals	38	60.1 ± 15.3	23–88	14	24	—	—	—	—
	Low-grade adenoma individuals	19	61.8 ± 7.2	40–73	11	6	—	—	—	—
	High-grade adenoma individuals	17	60.7 ± 8.2	49–79	13	6	—	—	—	—
	CRC patients	38	70.3 ± 12.3	39–81	22	16	3	4	15	16
WB (exosomes)	Healthy individuals	4	61.8 ± 8.3	52–70	3	1	—	—	—	—
	Premalignant lesion individuals	4	59.8 ± 14.7	38–70	2	2	—	—	—	—
	CRC patients	9	68.3 ± 7.3	61–85	5	2	4	1	2	2
Tissue samples										
IP stages I–II	CRC patients	6	77 ± 9.0	60–86	2	4	5	1	—	—
IP stages III–IV	CRC patients	7	78 ± 8.2	65–87	4	3	—	—	6	1
WB	CRC patients	17	74.3 ± 13.1	39–88	9	8	5	6	5	1
IHC	CRC patients	72	65.7 ± 8.7	41–82	47	25	—	4	3	65

SD: standard deviation.

胞系中。然后,使用20 mL 1×PBS洗涤每个外泌体制剂,并将样品在4 °C下再次以100 000g离心70 min。最后,弃去上清液,将含有外泌体的沉淀物重新悬浮在500 μL 1×PBS中,并在-80 °C下储存直至使用。在IP之前分析了两份10 μL的等分试样,以避免样品的连续冷冻和解冻。

为了分离和纯化人血浆外泌体,首先将每个血浆样品3 mL在4 °C下以17 000g离心30 min,以去除细胞外MV。然后将上清液通过0.22 μm过滤器过滤,随后以120 000g超速离心90 min以收集外泌体(附录A中的图S1)。最后,将细胞培养外泌体样品在1×PBS中稀释至1:50~1:100,人血浆外泌体样品在1×PBS中稀释至1:1000,并使用NanoSight NS300 (Malvern Panalytical, 英国)进行分析,以验证发生了适当的外泌体分离和纯化。

2.4. 电子显微镜和冷冻电子显微镜

通过带有负染色的电子显微镜(EM)和冷冻电子显微镜(cryoEM)分析外泌体颗粒。首先,将外泌体样品在1×PBS和4%多聚甲醛中固定5 min;把它们加到辉光放电的碳涂层网格上5 min,并用2%的醋酸双氧钨水溶液进行负染色。然后在配备了LaB6灯丝的FEI Tecnai 12 EM上分析样品,同时该灯丝在120 kV下运行。图像是用FEI Ceta数码相机记录的。

或者,通过cryoEM分析指定的外泌体制剂。为此,在预先亲水化的铑网格上收集外泌体制剂,并将样品快速冷冻在乙烷中。然后用在120 kV下运行的FEI Tecnai 12显微镜对样品进行分析。图像记录在电荷耦合器件(CCD)1k Gatan相机上。

2.5. 蛋白质提取和定量

CRC细胞系在37 °C和5% CO₂下生长至90%融合,用含有4 mmol·L⁻¹乙二胺四乙酸(EDTA)的1×PBS重悬,并在室温下以1200 r·min⁻¹离心5 min。然后,将细胞重新悬浮在500 μL裂解缓冲液(RIPA, Sigma Aldrich, 美国)中,并补充1×蛋白酶和磷酸酶抑制剂(MedChem-Express, 美国),并使用16G和18G针头注射器通过机械分解进行裂解。最后,将样品在4 °C下以10 000g离心10 min,并将蛋白质提取物(上清液)储存在-80 °C直至使用。

本研究使用了来自OCT嵌入的CRC组织样品的蛋白质提取物。首先,将组织样品在干冰上切成小块,用500 μL 1×PBS洗涤两次,并在4 °C下以10 000g离心5 min。然后将组织样品重新悬浮在300 μL RIPA中,并补充1×蛋白酶和磷酸酶抑制剂,同时用TissueLyser II (Qiagen, Hilden, 德国)机械裂解(两次循环,30 Hz每次2 min)。接下来,

将200 μL RIPA补充1×蛋白酶和磷酸酶抑制剂添加到每个样品中,并用TissueLyser II再次机械分解样品(两个循环,30 Hz每次2 min)。最后,将样品在4 °C下以10 000g离心10 min,并将蛋白质提取物(上清液)储存在-80 °C直至使用。

通过Trp定量方法[27]确定蛋白质浓度,并在10%十二烷基硫酸钠-聚丙烯酰胺凝胶电泳(SDS-PAGE)还原条件下通过考马斯蓝染色确认。

使用MicroBCA蛋白质检测试剂盒(Thermo Fisher Scientific, 美国)定量外泌体样品的蛋白质浓度,并在10% SDS-PAGE后通过考马斯蓝染色和WB确认。对于SDS-PAGE,外泌体用补充有0.15% β-巯基乙醇的上样缓冲液裂解(5个循环,冰上5 min,95 °C下5 min)。对于IP,通过机械分解,外泌体用RIPA裂解同时辅以1×蛋白酶和磷酸酶抑制剂。然后将样品在4 °C下以10 000g离心10 min,收集并储存蛋白质提取物(上清液)直至使用。

2.6. SDS-PAGE 和蛋白质印迹

通过考马斯蓝染色和WB分析蛋白质表达。在还原条件下使用10% SDS-PAGE分离5 μg的每种蛋白质提取物,然后在100 V下转移到硝酸纤维素膜上停留1 h。在室温下,用0.1%吐温1×PBS和3%脱脂牛奶(封闭缓冲液)封闭膜1 h,然后在封闭缓冲液中用优化稀释度(表S2)的一抗一起在4 °C下振荡孵育过夜(O/N)。接下来,将膜用0.1%吐温1×PBS洗涤3次,并在室温下与适当的指定辣根过氧化物酶(HRP)偶联的二抗(附录A中的表S3)一起在封闭缓冲液中振荡孵育1 h。随后用0.1%吐温PBS洗涤膜3次。使用ECL Pico Plus化学发光试剂(Thermo Fisher Scientific, 美国)产生信号,在Amersham Imager 680 (GE Healthcare, 美国)上检测。使用Image J软件量化蛋白质条带强度,并根据相应泳道的总蛋白质含量进行标准化(丽春红染色-Sigma Aldrich, 美国)。

2.7. 免疫沉淀

使用共200 μg的每种外泌体蛋白质提取物(1 mg总蛋白质)进行外泌体中所含蛋白质的IP,然而2 mg总等分子蛋白质提取物用于研究CRC TME组织(分别在I-II阶段和III-IV阶段)。蛋白质池通过与20 μL蛋白G琼脂糖珠(SCBT, 事先用RIPA平衡)一起在4 °C下旋转孵育1 h进行预净化。

首先,从III期和IV期CRC患者的血浆样本库(n=12)和健康个体的血浆样本库(n=12)中分离出IgG(表1和表S1)。将200 μL蛋白G琼脂糖珠分别与每个50 μL

血清池和350 μL 200 $\text{mmol}\cdot\text{L}^{-1}$ 磷酸钠在pH 7.0下孵育90 min。然后用1 mL的200 $\text{mmol}\cdot\text{L}^{-1}$ 磷酸钠在pH 7.0下将珠子洗涤两次,持续5 min,并再次与另外50 μL 的血清池和350 μL 的200 $\text{mmol}\cdot\text{L}^{-1}$ 磷酸钠在pH 7.0下孵育90 min。然后,将珠子用1 mL的200 $\text{mmol}\cdot\text{L}^{-1}$ 磷酸钠在pH 7.0下洗涤四次,持续5 min。最后,将结合的IgG用300 μL 0.1 $\text{mol}\cdot\text{L}^{-1}$ 甘氨酸在pH 2.7下进行两次4 min的洗脱。洗脱的蛋白质用30 μL 1 $\text{mol}\cdot\text{L}^{-1}$ Tris缓冲液(pH 8.8)中和,分离的IgG浓度用Nanodrop 2000C (Thermo Fisher Scientific)定量。

接下来,如前所述[14],将从CRC患者和作为对照的健康个体中分离出的IgG与蛋白G琼脂糖珠共价结合。简而言之,将先前在pH 7.4 (1 \times TBS)中含有150 $\text{mmol}\cdot\text{L}^{-1}$ NaCl的100 $\text{mmol}\cdot\text{L}^{-1}$ Tris-HCl中平衡的100 μL 糖珠与200 μg 从CRC患者或对照中分离的IgG分别在1 \times TBS中孵育2 h。接下来,用1 \times TBS洗涤珠子,并与200 $\text{mmol}\cdot\text{L}^{-1}$ 三乙醇胺在pH 8.9 (Sigma Aldrich)下孵育3次,持续5 min,以将先前的缓冲液中去除所有痕量Tris。之后,将珠子与含有50 $\text{mmol}\cdot\text{L}^{-1}$ 庚二酸二甲酯(DMP, Sigma Aldrich, 美国)的200 $\text{mmol}\cdot\text{L}^{-1}$ 三乙醇胺在pH 8.9下孵育1 h,以确保有效的IgG-蛋白G交联。然后用pH 8.9的200 $\text{mmol}\cdot\text{L}^{-1}$ 三乙醇胺洗涤珠子,并用乙醇胺(Sigma Aldrich)封闭15 min。最后,非交联的IgG珠用含有2 $\text{mol}\cdot\text{L}^{-1}$ 尿素的0.1 $\text{mol}\cdot\text{L}^{-1}$ 甘氨酸在pH 2.7下洗脱5 min,同时在与蛋白质提取物一起孵育之前,珠子先用RIPA平衡。

最后,将预清洗的蛋白质提取物与对照O/N的IgG共价结合的蛋白G珠在4 $^{\circ}\text{C}$ 下旋转孵育,随后与CRC患者O/N的IgG共价结合的蛋白G珠在4 $^{\circ}\text{C}$ 下旋转孵育。与蛋白质提取物孵育后,用1 \times TBS洗涤珠子11次,并用100 μL 0.1 $\text{mol}\cdot\text{L}^{-1}$ 甘氨酸(pH 2.7)洗脱对照和CRC患者的IgG识别的蛋白质两次(5 min)。洗脱的蛋白质用pH 8.0的2 $\text{mol}\cdot\text{L}^{-1}$ 碳酸氢铵(ABC)缓冲液中和,并使用Nanodrop 2000C (Thermo Fisher Scientific, 美国)定量蛋白质浓度。

所有孵育均在室温下旋转进行,所有上清液通过在13 200 $\text{r}\cdot\text{min}^{-1}$ 和4 $^{\circ}\text{C}$ 下离心5 min去除。所有步骤的等分试样在10% SDS-PAGE后通过考马斯蓝或银染色进行分析(附录A中的图S2)。

2.8. 液相色谱-串联质谱

在免疫沉淀的洗脱蛋白中,50%通过10% SDS-PAGE分离并用考马斯亮蓝G-250染色。然后,将整条泳道切成小块,分别用50 $\text{mol}\cdot\text{L}^{-1}$ ABC/50%乙腈(ACN)不染色,用ACN脱水并干燥。然后用25 $\text{mol}\cdot\text{L}^{-1}$ ABC中的10 $\text{mol}\cdot\text{L}^{-1}$

二硫苏糖醇(DTT)还原样品,并用碘乙酰胺烷基化至50 $\text{mol}\cdot\text{L}^{-1}$ 的最终浓度。接下来,将凝胶块干燥,用12.5 $\text{ng}\cdot\text{mL}^{-1}$ 猪胰蛋白酶(Thermo Fisher Scientific, 美国)在50 $\text{mol}\cdot\text{L}^{-1}$ ABC中再水化,并在37 $^{\circ}\text{C}$ 下孵育过夜。使用100% ACN和0.5% TFA提取肽,同时使用带有0.6 μL C18树脂(Millipore, Sigma Aldrich Química SL, 西班牙)的Zip Tip纯化并干燥。最后,在生物研究中心(CIB-CSIC, 西班牙)的蛋白质组学和基因组学设施中通过纳米LC-MS/MS进行分析之前,将样品溶解在5 μL 0.1%甲酸/2% ACN中。

在Easy-nLC 1000纳米系统(Thermo Fisher Scientific, 美国)上进行肽分离。为了进行分析,将样品装入预柱Acclaim PepMap 100 (Thermo Fisher Scientific),并在长50 cm、内径75 μm 、粒径2 μm (Thermo Fisher Scientific)的RSLC PepMap C18中洗脱。流动相流速为300 $\text{nL}\cdot\text{min}^{-1}$,使用0.1%甲酸水溶液(溶剂A)、0.1%甲酸和100% ACN(溶剂B)。梯度曲线设置如下:5%~35%溶剂B持续90 min,35%~100%溶剂B持续4 min和100%溶剂B持续8 min。每个样品注射4 μL 。

使用Q-Exactive质谱仪(Thermo Fisher Scientific)进行MS分析。对于电离,使用1900 V的液接电压和300 $^{\circ}\text{C}$ 的毛细管温度。全扫描方法采用 m/z 400~1500质量范围,Orbitrap分辨率为70 000(在 m/z 200处),目标自动增益控制(AGC)值为 3×10^6 ,最大进样时间为100 ms。调查扫描后,选择了15个最强的前体离子进行MS/MS碎裂。使用27 eV的归一化碰撞能量进行碎裂,并以起始质量为 m/z 200、AGC目标为 2×10^5 、分辨率为17 500(在 m/z 200处)、强度阈值为 8×10^3 、隔离窗口为2.0 m/z 单位以及最大IT为100 ms获得MS/MS扫描。电荷态筛选可以拒绝未指定的、单电荷的、等于或大于7个质子化的离子。20 s的动态排除时间用于区分先前选择的离子。

2.9. 质谱数据分析

使用Proteome Discoverer(版本1.4.1.14, Thermo Fisher Scientific)标准化工作流程分析MS数据。使用Mascot搜索引擎(2.6版, Matrix Science, 英国)针对SwissProt_2016_10.fasta智人(人类)数据库(20 121个序列蛋白质条目)搜索质谱*.raw文件。前体和片段的质量容差分别设置为10 ppm和0.02 Da,允许两个缺失的裂解,半胱氨酸的氨基甲酰甲基化作为固定修饰,而蛋氨酸的氧化以及N端的乙酰化作为可变修饰。使用 q 值阈值为0.01的Percolator算法[28]过滤已识别的肽。

2.10. 生物信息学和荟萃分析

使用亲和纯化 MS 数据库的污染物存储库 (CRAPome), 对鉴定为 CRC 特异性潜在自身抗原的蛋白质进行了第一次筛选[29]。在该数据库中报道的超过 15% 的 IP 实验中出现的蛋白质在进一步分析后被去除。然后, 使用 STRING (版本 11.0) 和 DAVID (版本 6.8) 数据库对蛋白质进行分析, 以研究蛋白质富集并确定这些蛋白质所涉及的改变的网络和通路[30–32]。STRING 设置固定为马尔可夫聚类算法 (MCL) 聚类富集 2, 置信度分数为 0.4。

通过分析与 CRC 相关的遗传改变是否与这些蛋白质相关, 对所选蛋白质进行了广泛的荟萃分析。为此, 使用了 UALCAN [33] 和 cBioPortal [34] 数据库, 因为它们分别提供了有关 CRC 和配对健康组织样本之间基因表达失调以及肿瘤组织样本中鉴定的基因改变的信息。此外, 研究人员利用来自人类蛋白质图谱[35–36] 和 GEPIA 2 [37–38] 的信息, 分别使用最佳截止值或四分位数检查鉴定的蛋白质以前是否与 CRC 患者的预后或生存相关。

2.11. 免疫组织化学

在免疫组织化学 (IHC) 分析中, 本研究使用 MTA-1 组织阵列仪 (Beecher Instruments, 美国) 构建了组织微阵列, 该组织微阵列包含来自 72 名 CRC 患者的 88 个核心样本 (表 1 和附录 A 中的表 S4)。组织阵列仪在石蜡包埋组织中预先选定的肿瘤区域上打孔完成每个核心样本 (直径: 1 mm) 的采样。在 2 μm 切片中进行染色。此外, 在 60 $^{\circ}\text{C}$ 下孵育 15 min 使载玻片脱蜡, 然后在 PT-Link (丹麦) 上 95 $^{\circ}\text{C}$ 的低 pH 缓冲溶液 (EnVision FLEX Target Retrieval Solution, Dako, 丹麦) 中孵育 20 min。本实验将支架与过氧化物酶阻断剂 (Dako, 丹麦) 一起孵育以阻断内源性过氧化物酶。使用指定抗体的优化稀释液, 使活检在 4 $^{\circ}\text{C}$ 下孵育过夜 (表 S3)。组织与适当的抗 Ig 辣根过氧化物酶偶联聚合物 (EnVision FLEX-HRP, Dako, 丹麦) 一起孵育, 以检测抗原抗体反应。所有抗体和抗-Ig HRP-偶联抗体均表现出较高的特异性, 并且这些抗体均未单独呈阳性。然后用 3,3'-二氨基联苯胺作为显色液使切片显色 7 min, 并用苏木精进行复染。本实验按照既定的方案进行显色操作和免疫反应[14,39]。

2.12. Gateway 克隆和蛋白质表达

序列验证的全长互补 DNA (cDNA) 质粒是从公开的 DNASU 质粒库 (<https://dnasu.org/DNASU/Home.do>) [40] 中获得的, 质粒包含 *TP53*、*TGM3*、*ACTR3*、*PSPH*、*MT-*

CO2、*SPDL1*、*HMGCS2*、*CYB5R1*、*RAB25*、*FABP5*、*SDF2L1*、*S100A8*、*SERPINB12*、*SERPINB3*、*RAB2A* 和 *DEFA1* 基因, 这些质粒位于灵活的 pDONR221、pDONR223 或 pENTR223 质粒载体系统中。随后, 按照制造商的说明, 将可读框 (ORF) 转移到 pANT7_cHalo 载体中, 通过 LR Clonase 反应 (Invitrogen, 美国) 进行体外表达, 以获得其 C 端表达为 Halotag 融合蛋白的全长蛋白[41–42]。接下来, 用 LR 反应的产物转化 TOP10 大肠杆菌细胞以分离所需要的表达质粒。根据制造商的说明, 使用 Miniprep (Neobiotech, 美国) 纯化供体和表达质粒, 并在使用前对序列进行验证。

根据制造商的发光免疫分析说明, 使用 1-Step Human Coupled IVT Kit HeLa 细胞裂解物 (Thermo Fisher Scientific, 美国) 表达融合蛋白。如前所述, 通过 WB 对蛋白表达进行分析, 使用每孔 0.67 μL 的蛋白表达、抗-Halotag 小鼠单克隆抗体 (1/2500, Promega, 美国) 作为一抗, 并且使用 HRP 标记的山羊抗-小鼠 (GAM, Sigma Aldrich) 作为二抗。

2.13. 血清反应性发光测定

如前所述[42–45], 使用按照制造商说明制备的 Magne Halotag 磁珠 (MB, Promega, 美国), 通过发光 Halotag 磁珠免疫分析法分析 p53、TGM3、ACTR3、PSPH、MT-CO2、SPDL1、HMGCS2、CYB5R1、RAB25、FABP5、SDF2L1、S100A8、SERPINB12、SERPINB3、RAB2A 和 DEFA1 蛋白的血清反应性。首先, 将 MB 在 500 μL 洗涤缓冲液 (0.1% 吐温-20、0.05% Triton X-100, 在 1 \times PBS 中) 中平衡 3 次, 在室温下旋转 5 min。然后, 将体外表达蛋白与 MB O/N 在 4 $^{\circ}\text{C}$ 下旋转孵育, 每次反应的蛋白量为 0.67 μL , MB 悬浮液为 0.5 μL , 最终体积为 600 μL 1 \times PBS。此外, 96 孔板 (Bio-Plex Pro Flat 底板, Bio-Rad, 美国) 用 300 μL 的 0.1% 的吐温 1 \times PBS 封闭, 并在 4 $^{\circ}\text{C}$ 和 120 $\text{r}\cdot\text{min}^{-1}$ 的条件下添加 3% 的牛血清白蛋白 (BSA) O/N。蛋白质固定后, 将 MB 用洗涤缓冲液在室温下旋转洗涤 3 次, 每次 5 min, 然后用 SuperBlock 封闭缓冲液 (Thermo Fisher Scientific) 在室温下旋转封闭 1 h。之后, 将 MB 转移到 96 孔板 (Bio-Plex Pro Flat 底板, Bio-Rad) 上并与 112 个血浆样品分别孵育, 每个血浆样品为 50 μL , 使用 1/400 稀释的 0.1% 吐温 1 \times PBS 进行孵育, 并在 4 $^{\circ}\text{C}$ 和 150 $\text{r}\cdot\text{min}^{-1}$ 的条件下补充 3% 的 BSA O/N (表 1 和表 S1)。接下来, 将孔洗涤 3 次, 并与 50 μL HRP 标记的抗-人 IgG 二抗 (Dako) 一起孵育, 在添加 3% BSA 的 0.1% 吐温 1 \times PBS 中稀释到 1/3000, 同时在室温和 150 $\text{r}\cdot\text{min}^{-1}$ 的条件下

孵育 1 h。最后, 将孔洗涤 3 次, 并用 ECL Pico Plus 化学发光试剂 (Thermo Fisher Scientific) 生成信号并记录到 Spark 多模式酶标仪 (Tecan Trading AG, Switzerland) 上。作为测定的对照, 实验使用融合蛋白共价固定的 MB 和抗-HaloTag 抗体以及适当的 HRP 偶联二抗一起孵育 (表 S3)。

2.14. 血清反应生物传感分析

由于研究人员已经报道且优化了用于血清反应性生物传感测定的实验方案[41–42], 因此本实验在此遵循此方案。蛋白质在体外表达, 锚定在 MB 上, 并与血浆样品一起孵育[41–42]。为了验证正确的蛋白质固定, 用抗-HaloTagmAb (Promega) 检测 HaloTag 融合蛋白, 然后与稀释到 1/3000 的 HRP 偶联抗鼠 IgG (Sigma Aldrich) 孵育 1 h。安培测量是在氢醌 (HQ) /H₂O₂ 系统存在的情况下进行的, 该系统具有一次性丝网印刷碳电极, 并通过磁性捕获工作电极上带有免疫复合物的 MB [41–42]。生物识别事件通过 HQ 介导的 H₂O₂ 酶促 (HRP) 还原产生的阴极电流的变化进行监测, 并在 -0.20 V 电压 (相对于 Ag 伪参考电极) 下进行测量[41–42]。

2.15. 统计分析

本实验使用 Microsoft Excel 2019 和 GraphPad Prism 5 程序获得点图、平均值、均数标准误和 *t* 检验。使用 R (v3.6.2, Austria) 计算非参数 Mann-Whitney *U* 检验值。*p* 值小于 0.05 被认为具有统计学意义。用 R (v3.6.2) 中的 ModelGood 和 Epi 包构建受试者操作特征 (ROC) 曲线, 用于确定每种蛋白质单独或组合的诊断能力。

3. 结果

在这项工作中, 我们分析了 CRC TDE 和 TME 作为识别可能有助于早期检测这种疾病的 CRC 特异性 TAA 的补充和稀有分析来源。对于 TAA 的鉴定, 我们使用了最近描述的方法, 该方法包括 IP 和 LC-MS/MS 的组合, 使用外泌体和 CRC 肿瘤组织, 然后进行血清反应性测定[14]。用于 TAA 鉴定的外泌体, 分别从 SW480 和 KM12C CRC 细胞的分泌组中分离出来, 作为原发性肿瘤的模式; 从 SW620、KM12SM 和 KM12L4a 细胞的分泌组中分离出来, 作为 CRC 转移到淋巴结、肝脏、肝和肺的模式。实验使用从不同阶段患者的血浆样本中分离的 CRC 组织样本和外泌体用于验证疾病过程中 TAA 的失调。最后, 我们将 TAA 集成到基于 HaloTag 技术的电化学免疫传感平台中, 以构建 POC 设备, 用于区分 CRC 患者和携带癌前病

变的个体以及健康个体[41–42]。图 1 描述了这项研究的工作流程。

3.1. CRC 细胞系分泌的外泌体的表征

将培养细胞的培养液离心, 分离 CRC 细胞以释放细胞外囊泡。CRC 细胞在不含 FBS 的 DMEM 中孵育 48 h 后收集条件培养基, 从而避免外源外泌体的分离。在将外泌体蛋白质含量用到 IP 分析中之前, 对纯化的样品进行表征以确认外泌体的适当分离和纯化已经发生 (图 2)。

使用 EM [图 2 (a)], 我们观察到不大于 150 nm 的囊泡, 由于样品制备需要经历脱水过程, 这些囊泡与具有扭曲结构的外泌体囊泡尺寸相容。未检测到污染物, 如先前未收获的 MV 或蛋白质聚集体。为了进一步确认, 我们使用 cryoEM 分析来自 SW480 细胞的纯化囊泡 [图 2 (a)]。我们观察到不大于 150 nm 的囊泡被脂质双层包围, 并具有天然的外泌体结构。最后, 使用 NS300, 我们观察了悬浮液中的分离囊泡, 并获得所有样品的平均大小和浓度 [图 2 (b)]。从非转移性 KM12C 和 SW480 细胞中分离出的囊泡, 比它们的同基因转移性对应物 KM12SM 和 KM12L4a 细胞以及 SW620 细胞分泌的囊泡在大小上更不均匀, 且数量更多。此外, 所有样品中囊泡的平均尺寸在 140–200 nm 之间, 这与外泌体样品的尺寸一致。

我们还使用考马斯蓝染色法和 WB 分析了外泌体样品的蛋白质提取物 [图 2 (c)]。这些分析表明, 相同 CRC 细胞的外泌体和细胞裂解物之间的蛋白质含量存在差异。重要的是, Alix 是细胞外囊泡特异性的蛋白质标志物, 仅能在外泌体样品中被观察到。总的来说, 这些结果证实了从 CRC 细胞中适当分离和纯化外泌体的情况已经发生。

3.2. 通过免疫沉淀结合 LC-MS/MS 鉴定潜在的 TAA

本研究通过 IP 和 LC-MS/MS 以检查由 CRC 细胞分泌的外泌体和来自 I–IV 期患者的 CRC 组织样本, 以确定在疾病过程中失调的 CRC 特异性 TAA。首先使用考马斯蓝染色法 [图 2 (c) 和附录 A 中的图 S2 (a)] 来分析来自外泌体和组织样本的蛋白质提取物, 以评估其质量。

其次, 分别从来自 CRC 患者 (*n*=12) 和作为对照的健康个体 (*n*=12) 的血浆样品中分离 IgG (表 1 和表 S1), 并与蛋白 G 琼脂糖珠共价结合。然后将蛋白质提取物与来自对照的 IgG 一起孵育, 以去除对对照具有血清反应性的蛋白质。然后, 将澄清的蛋白质提取物与来自 CRC 患者的 IgG 一起孵育, 以鉴定对该疾病具有特异性的 TAA。在 LC-MS/MS 分析之前, 本实验将对对照的 IgG 具有血清反

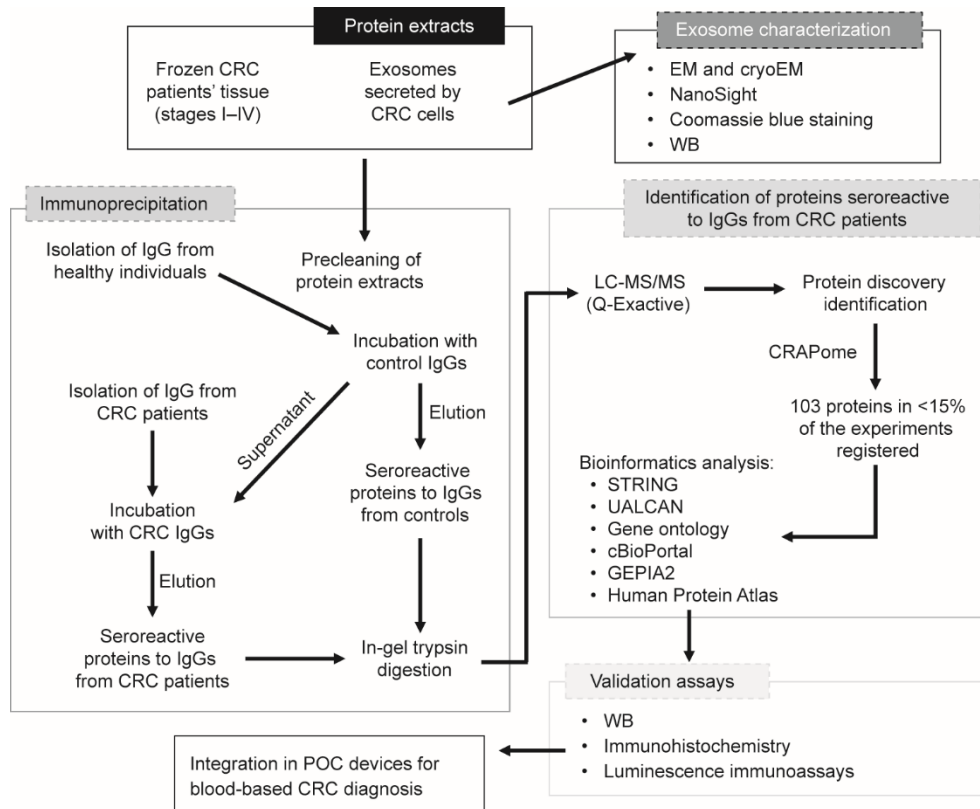


图1. 本研究中使用的工作流程方法，用于识别、验证和整合来自CRC患者的外泌体和组织样本中TAA的POC诊断设备。使用IP结合质谱法研究了来自五种具有不同转移能力的CRC细胞系分泌的外泌体的蛋白质提取物，和来自I-IV期CRC患者的组织样本，以鉴定在疾病过程中失调的CRC特异性潜在TAA。本研究对来自CRC患者鉴定出IgG具有特异性血清反应性的103种蛋白质进行了荟萃分析，并确定了15种蛋白质更有可能被实际用于验证TAA。WB、IHC和荟萃分析被用于确定它们与CRC发展、进展和预后的关联。最后，在将蛋白质整合到类似的POC设备上以诊断CRC和结肠直肠癌前病变个体之前，通过血清反应性发光测定验证了这些蛋白质的血清反应性。

应的蛋白质和对CRC患者IgG有血清反应的蛋白质，在10%的SDS-PAGE操作后进行镀银染色洗脱和分析[图3(a)]。

接下来，将对两种IgG均具有血清反应性的蛋白质在10% SDS-PAGE中浓缩，用胰蛋白酶进行凝胶内消化，然后使用MS进行鉴定。三个IP测定总共鉴定了919种蛋白质（来自I-II期患者组织样本的404种蛋白质，来自III-IV期患者组织样本的270种蛋白质和来自外泌体的245种蛋白质）。其中235种蛋白质仅对来自CRC患者的IgG呈血清反应，194种蛋白质仅对来自健康个体的IgG呈血清反应，并且有490种蛋白质出现在了两种分析中[图3(b)和附录A中的表S5-S7]。

在生物信息学和荟萃分析之前，在CRAPome数据库中筛选了235种被专门鉴定为对来自CRC患者的IgG具有血清反应的蛋白质，在数据库中注册的IP实验中，这些蛋白质的出现率低于15%。然后总共选择了103种不同的蛋白质，因为它们更有可能是实际针对CRC的TAA。在这些蛋白质中，研究中进行的两个或三个IP测定鉴定出了其中13种蛋白质[表2和图3(c)]。值得注意的是，

103种蛋白质中的一些在之前的研究中被鉴定为CRC TAA [14]，如TALDO1、MT-CO2或SPCS2蛋白，这为用于鉴定TAA方法的准确性提供了支持。

3.3. 生物信息学、荟萃分析和验证

随后，本研究对鉴定为CRC潜在TAA的103种蛋白质进行了广泛的荟萃分析(图4)。并使用STRING和DAVID数据库来检查这些蛋白质所涉及的蛋白质网络和途径。通过DAVID数据库，本文发现这些蛋白质涉及多个关键的生物过程和功能，其中一些在癌症中发挥着关键作用[图4(a)]。关于它们的细胞定位，蛋白质主要位于细胞质或位于细胞外泌体中，这表明外泌体富含该疾病特异性的TAA。此外，使用STRING数据库，蛋白质通过直接和间接相互作用分为了10个簇，包括与代谢过程、免疫反应、细胞骨架组织、信号转导、蛋白质转运、应激反应、钙结合和细胞分化相关的蛋白质[图4(b)]，这些都是与癌症密切相关的过程。

接下来，由于自身抗体的蛋白质靶标在肿瘤组织中失调，并且可能发生点突变、移码突变或过表达[7-8]以及

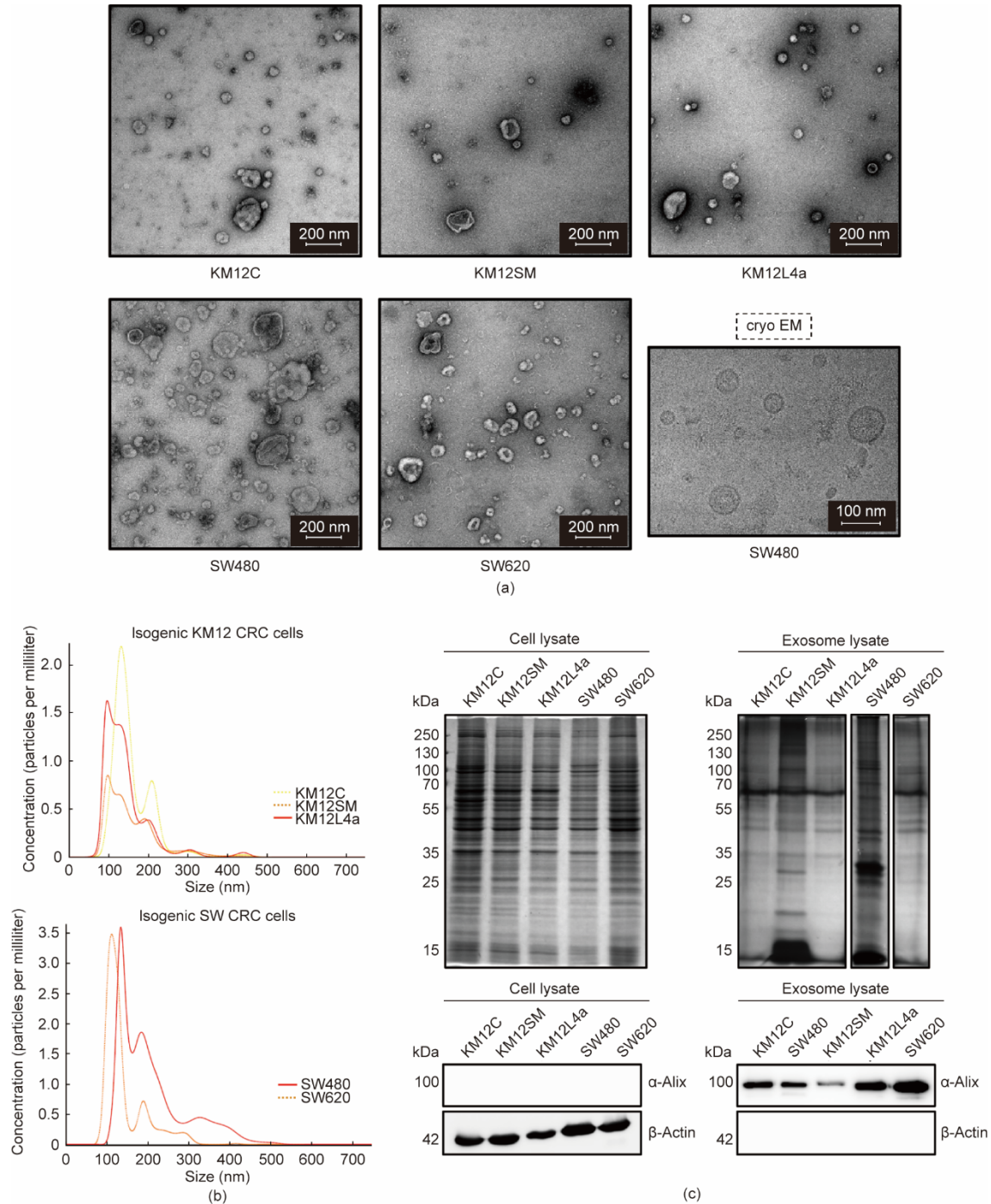


图2. CRC细胞系释放的外泌体的表征。(a) 通过差速离心分离和纯化的CRC细胞系分泌的囊泡，其电子显微镜和冷冻电镜图像。能观察到小于150 nm的囊泡，囊泡被与外泌体相容的脂质双层包围。(b) 用NS300获得的CRC细胞分泌的外泌体的重叠光谱显示，它们在不同的等基因CRC细胞中的大小和浓度存在差异。与等基因转移性SW620、KM12SM和KM12L4a细胞分别分泌的外泌体相比，非转移性SW480和KM12C细胞分泌的外泌体在大小上更不均匀，且体积更大。此外，SW同基因细胞比KM12细胞分泌的外泌体更多。(c) 外泌体和细胞裂解物的考马斯亮蓝染色和WB（上图）揭示了外泌体和细胞之间蛋白质模式的差异。Alix是一种细胞外囊泡特异性蛋白质，仅能在外泌体中观察到。

其他改变，本文调查了与CRC相关的这些蛋白质的遗传失调和（或）任何其他改变是否在之前已经被描述过。我们通过使用UALCAN发现CRC组织中的基因表达与相邻配对健康组织中103种蛋白质中的79种相比，存在统计学上显著的失调 ($p < 0.05$)，其中39种蛋白质的基因表达

在CRC中上调[图4 (c)]。此外，本文通过使用cBioPortal在癌症基因组图谱 (TCGA) 程序 (PanCancer Atlas 和 Nature 2012) 的两个CRC数据集中调查了先前针对这些蛋白质描述的基因改变。研究发现，除了一种蛋白质外，其他所有被鉴定的蛋白质都有不同的基因改变（主要是突

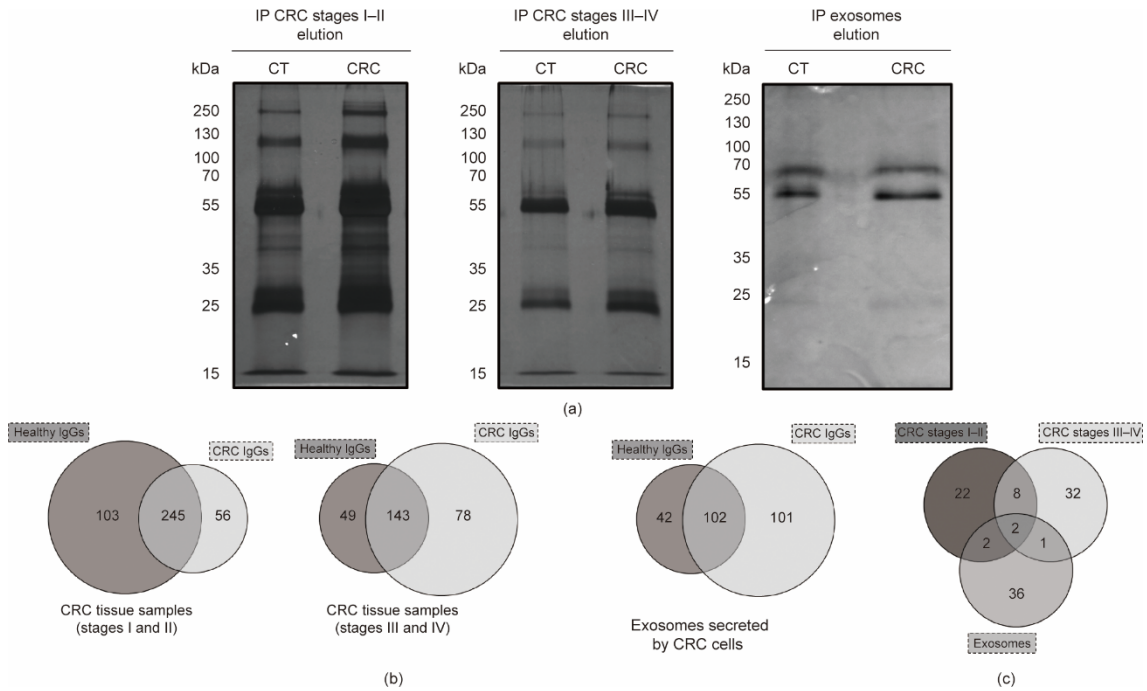


图 3. 来自 CRC 细胞衍生的外泌体、肿瘤组织和自身抗体靶向的肿瘤基质外泌体的蛋白质在 LC-MS/MS 后的 IP。(a) 在本研究中的三种 IP 后，来自健康个体和来自 CRC 患者的、对 Ig 具有血清反应的蛋白质洗脱的银染。(b) 对健康个体的 IgG 有血清反应的蛋白质、对 CRC 患者的 IgG 有血清反应的蛋白质或对两种 IgG 都有血清反应的蛋白质的维恩图，在每个 IP 中通过 LC-MS/MS 进行鉴定和定量。从三种免疫沉淀物中总共鉴定出 919 种蛋白质，其中 235 种蛋白质对来自 CRC 患者的 IgG 具有特异性血清反应。(c) 103 种蛋白质的维恩图被确定为每个 IP 中 CRC 特异性的潜在 TAA，这些蛋白质出现在 CRAPome 数据库中登记的小于 15% 的免疫沉淀实验中。本项研究在多个 IP 中鉴定了其中 13 种蛋白质。CT：对照。

变) [图 4 (d)]。此外，在这些 TCGA 数据集中注册的 75% 以上的 CRC 组织样本中，发现了基因改变和异常表达。这些数据证实，这些蛋白质发生了变化，可能诱导 CRC 患者的体液发生免疫反应。

然后，本文重点分析了根据以下条件选择的 103 种蛋

白质 (表 2) 中的 15 种：①先前与 CRC 相关的基因改变和 (或) 失调；②关于其与 CRC 或其他癌症相关性的现有信息；③了解与这些蛋白质相关的生物过程和细胞定位；④抗体和 cDNA 的可用性。

首先，本文使用 WB 来分析它们在 CRC 组织或来自

表 2 在三种免疫沉淀试验中鉴定出的具有 CRC 潜在诊断价值的蛋白质，在 CRAPome 数据库中的免疫沉淀试验出现率小于 15%

Accession number	Description	Gene	Σ coverage (%)	Σ Peptides	Σ PSMS	CRAPome Number of experiments (found/total)	Identification
A4D1P6	WD repeat-containing protein 91	<i>WDR91</i>	0.94	1	2	0	IP3
O15143	Actin-related protein 2/3 complex subunit 1B	<i>ARPC1B</i>	8.33	2	3	9/411	IP1, IP2
O15144	Actin-related protein 2/3 complex subunit 2	<i>ARPC2</i>	3.67	1	1	18/411	IP1, IP2
O15145	Actin-related protein 2/3 complex subunit 3	<i>ARPC3</i>	19.66	3	3	15/411	IP1
O15511	Actin-related protein 2/3 complex subunit 5	<i>ARPC5</i>	7.95	1	1	12/411	IP1
O75083	WD repeat-containing protein 1	<i>WDR1</i>	1.32	1	1	24/411	IP3
O75223	Gamma-glutamylcyclotransferase	<i>GGCT</i>	11.7	2	2	30/411	IP3
O94832	Myosin-IId	<i>MYO1D</i>	5.86	7	7	33/411	IP1, IP2
O95994	Anterior gradient protein 2 homolog	<i>AGR2</i>	10.29	2	2	0	IP1
P00352	Retinal dehydrogenase 1	<i>ALDH1A1</i>	1.6	1	1	4/411	IP3
P00403	Cytochrome c oxidase subunit 2	<i>MT-CO2</i>	4.41	1	1	20/411	IP1, IP3
P00488	Coagulation factor XIII A chain	<i>F13A1</i>	3.28	3	4	0	IP2
P01023	Alpha-2-macroglobulin	<i>A2M</i>	0.81	1	1	16/411	IP1
P01040	Cystatin-A	<i>CSTA</i>	19.39	2	2	45/411	IP1

Accession number	Description	Gene	Σ coverage (%)	Σ Peptides	Σ PSMs	CRAPome Number of experiments (found/total)	Identification
P02008	Hemoglobin subunit zeta	<i>HBZ</i>	4.93	1	1	0	IP2
P02671	Fibrinogen alpha chain	<i>FGA</i>	4.27	3	3	2/411	IP2
P02675	Fibrinogen beta chain	<i>FGB</i>	13.24	6	6	9/411	IP2
P02751	Fibronectin	<i>FN1</i>	0.88	2	2	20/411	IP2
P02788	Lactotransferrin	<i>TRFL</i>	1.41	1	1	28/411	IP3
P04040	Catalase	<i>CAT</i>	4.55	2	3	37/411	IP3
P05089	Arginase-1	<i>ARG1</i>	3.42	1	1	50/411	IP3
P05109	Protein S100-A8	<i>S100A8</i>	7.53	1	1	57/411	IP3
P06396	Gelsolin	<i>GSN</i>	6.01	4	5	26/411	IP2
P06703	Protein S100-A6	<i>S100A6</i>	8.89	1	1	29/411	IP1
P07741	Adenine phosphoribosyltransferase	<i>APRT</i>	6.11	1	1	25/411	IP3
P08134	Rho-related GTP-binding protein RhoC	<i>RHOC</i>	5.7	1	1	35/411	IP1
P08246	Neutrophil elastase	<i>ELANE</i>	13.48	2	2	2/411	IP2
P08519	Apolipoprotein(a)	<i>APOA</i>	6.82	1	1	0	IP1, IP2
P08559	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	<i>PDHA1</i>	9.74	4	4	47/411	IP2
P09327	Villin-1	<i>VILI</i>	13.18	9	11	0	IP2
P10620	Microsomal glutathione S-transferase 1	<i>MGST1</i>	10.32	1	1	0	IP1
P16520	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-3	<i>GNB3</i>	2.94	1	1	39/411	IP2
P18754	Regulator of chromosome condensation	<i>RCC1</i>	2.38	1	1	61/411	IP3
P20339	Ras-related protein Rab-5A	<i>RAB5A</i>	5.12	1	1	42/411	IP2
P20930	Filaggrin	<i>FLG</i>	1.77	3	4	43/411	IP3
P21796	Voltage-dependent anion-selective channel protein 1	<i>VDAC1</i>	3.89	1	1	58/411	IP3
P22531	Small proline-rich protein 2E	<i>SPRR2E</i>	25	1	1	24/411	IP3
P22735	Protein-glutamine gamma-glutamyltransferase K	<i>TGM1</i>	0.86	1	1	34/411	IP3
P23490	Loricrin	<i>LOR</i>	2.56	1	1	5/411	IP3
P27338	Amine oxidase [flavin-containing] B	<i>MAOB</i>	2.12	1	1	0	IP2
P28331	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial	<i>NDUFS1</i>	4.13	2	2	19/411	IP1, IP2
P29508	Serpin B3	<i>SERPINB3</i>	3.08	1	1	36/411	IP3
P30044	Peroxiredoxin-5, mitochondrial	<i>PRDX5</i>	8.88	2	2	26/411	IP1
P31040	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	<i>SDHA</i>	2.86	2	3	44/411	IP2
P31930	Cytochrome b-c1 complex subunit 1, mitochondrial	<i>UQCRC1</i>	2.5	1	1	30/411	IP2
P37837	Transaldolase	<i>TALDO</i>	2.37	1	1	41/411	IP3
P42357	Histidine ammonia-lyase	<i>HAL</i>	1.22	1	1	10/411	IP3
P45880	Voltage-dependent anion-selective channel protein 2	<i>VDAC2</i>	9.52	2	2	58/411	IP2
P49773	Histidine triad nucleotide-binding protein 1	<i>HINT1</i>	7.14	1	1	27/411	IP3
P50552	Vasodilator-stimulated phosphoprotein	<i>VASP</i>	3.42	1	1	6/411	IP2
P51148	Ras-related protein Rab-5C	<i>RAB5C</i>	5.56	1	1	61/411	IP3
P51149	Ras-related protein Rab-7a	<i>RAB7A</i>	5.31	1	1	63/411	IP3
P54868	Hydroxymethylglutaryl-CoA synthase, mitochondrial	<i>HMGCS2</i>	8.07	4	4	2/411	IP1

Accession number	Description	Gene	Σ coverage (%)	Σ Peptides	Σ PSMs	CRAPome Number of experiments (found/total)	Identification
P56134	ATP synthase subunit f, mitochondrial	<i>ATP5J2</i>	13.83	1	1	28/411	IP1
P57735	Ras-related protein Rab-25	<i>RAB25</i>	3.76	1	1	1/411	IP2
P59665	Neutrophil defensin 1	<i>DEFA1</i>	19.15	2	2	23/411	IP3
P59998	Actin-related protein 2/3 complex subunit 4	<i>ARPC4</i>	6.55	1	1	39/411	IP1, IP2
P61019	Ras-related protein Rab-2A	<i>RAB2A</i>	11.32	2	2	35/411	IP2
P61088	Ubiquitin-conjugating enzyme E2 N	<i>UBE2N</i>	7.24	1	1	46/411	IP3
P61158	Actin-related protein 3	<i>ACTR3</i>	2.15	1	1	41/411	IP1, IP2, IP3
P61160	Actin-related protein 2	<i>ARP2</i>	3.55	1	1	51/411	IP1, IP2
P61626	Lysozyme C	<i>LYZ</i>	4.73	1	1	58/411	IP2
P62491	Ras-related protein Rab-11A	<i>RAB11A</i>	3.7	1	1	47/411	IP3
P78330	Phosphoserine phosphatase	<i>PSPH</i>	64	11	18	1/411	IP1, IP2, IP3
Q01469	Fatty acid-binding protein, epidermal	<i>FABP5</i>	6.67	1	2	38/411	IP2
Q08188	Protein-glutamine gamma-glutamyltransferase E	<i>TGM3</i>	1.88	1	1	37/411	IP2, IP3
Q08380	Galectin-3-binding protein	<i>LGALS3BP</i>	1.88	1	1	25/411	IP3
Q08554	Desmocollin-1	<i>DSC1</i>	2.57	2	2	51/411	IP3
Q09666	Neuroblast differentiation-associated protein AHNAK	<i>AHNAK</i>	1.68	1	2	62/411	IP3
Q12792	Twinfilin-1	<i>TWF1</i>	2	1	1	28/411	IP1
Q13835	Plakophilin-1	<i>PKP1</i>	3.08	2	2	34/411	IP1, IP3
Q14019	Coactosin-like protein	<i>COTL1</i>	4.93	1	1	1/411	IP2
Q15005	Signal peptidase complex subunit 2	<i>SPCS2</i>	3.98	1	1	7/411	IP3
Q16666	Gamma-interferon-inducible protein 16	<i>IFI16</i>	1.66	1	1	9/411	IP2
Q16851	UTP--glucose-1-phosphate uridylyltransferase	<i>UGP2</i>	1.77	1	1	11/411	IP3
Q5VU13	V-set and immunoglobulin domain-containing protein 8	<i>VSIG8</i>	2.42	1	1	3/411	IP1
Q86SJ6	Desmoglein-4	<i>DSG4</i>	0.77	1	1	1/411	IP1
Q8IXJ6	NAD-dependent protein deacetylase sirtuin-2	<i>SIRT2</i>	2.06	1	1	2/411	IP2
Q8N7P1	Inactive phospholipase D5	<i>PLD5</i>	1.31	1	1	0	IP2
Q8WVV4	Protein POF1B	<i>POF1B</i>	4.41	2	2	25/411	IP3
Q92954	Proteoglycan 4	<i>PRG4</i>	3.42	1	1	0	IP3
Q969V3	Nicalin	<i>NCLN</i>	2.49	1	1	6/411	IP1
Q96CW1	AP-2 complex subunit mu	<i>AP2M1</i>	1.61	1	1	21/411	IP1, IP2
Q96EA4	Protein Spindly	<i>SPDL1</i>	1.16	1	1	19/411	IP1
Q96IU4	Abhydrolase domain-containing protein 14B	<i>ABHD14B</i>	3.33	1	1	1/411	IP1
Q96P63	Serpin B12	<i>SERPINB12</i>	1.98	1	2	37/411	IP3
Q96PK6	RNA-binding protein 14	<i>RBM14</i>	1.79	1	1	52/411	IP3
Q96QA5	Gasdermin-A	<i>GSDMA</i>	1.8	1	1	22/411	IP3
Q99698	Lysosomal-trafficking regulator	<i>LYST</i>	0.21	1	1	1/411	IP2
Q9BPW8	Protein NipSnap homolog 1	<i>NIPSNAP1</i>	3.17	1	1	14/411	IP1
Q9BPX5	Actin-related protein 2/3 complex subunit 5- like protein	<i>ARPC5L</i>	7.84	1	1	3/411	IP1
Q9HCN8	Stromal cell-derived factor 2-like protein 1	<i>SDF2L1</i>	19.46	3	4	5/411	IP2
Q9NWT1	p21-activated protein kinase-interacting protein 1	<i>PAK1IP1</i>	2.3	1	1	22/411	IP2
Q9NYL9	Tropomodulin-3	<i>TMOD3</i>	5.68	2	2	63/411	IP1

Accession number	Description	Gene	Σ coverage (%)	Σ Peptides	Σ PSMs	CRAPome Number of experiments (found/total)	Identification
Q9NNT1	Calmodulin-like protein 5	<i>CALML5</i>	9.59	1	1	53/411	IP3
Q9UBG3	Cornulin	<i>CRNN</i>	2.83	1	1	1/411	IP3
Q9UGM3	Deleted in malignant brain tumors 1 protein	<i>DMBT1</i>	6.96	1	1	8/411	IP1
Q9UHQ9	NADH-cytochrome b5 reductase 1	<i>CYB5R1</i>	2.62	1	1	0	IP2
Q9Y277	Voltage-dependent anion-selective channel protein 3	<i>VDAC3</i>	8.48	2	2	46/411	IP2
Q9Y2Y8	Proteoglycan 3	<i>PRG3</i>	4.89	1	1	0	IP2
Q9Y3A5	Ribosome maturation protein SBDS	<i>SBDS</i>	3.2	1	1	13/411	IP2
Q9Y4L1	Hypoxia upregulated protein 1	<i>HYOU1</i>	1.2	1	1	50/411	IP1
Q9Y6N5	Sulfide:quinone oxidoreductase, mitochondrial	<i>SQRDL</i>	1.78	1	1	2/411	IP2

IP1: immunoprecipitation with tissue samples from CRC patients at stages I and II; IP2: immunoprecipitation with tissue samples from CRC patients at stages III and IV; IP3: immunoprecipitation with exosomes secreted by CRC cells.

In bold: The 15 identified proteins more likely to be TAAs specific to CRC among the 103 identified proteins and selected for seroreactivity validation. NADH: nicotinamide adenine dinucleotide; UTP: uridine-5'-triphosphate; ATP: adenosine triphosphate; NAD: nicotinamide adenine dinucleotide. Σ coverage, total percentage of protein sequence coverage sequenced by LC-MS/MS; Σ Peptides: total number of peptides; Σ PSMs: total number of peptide spectrum matches.

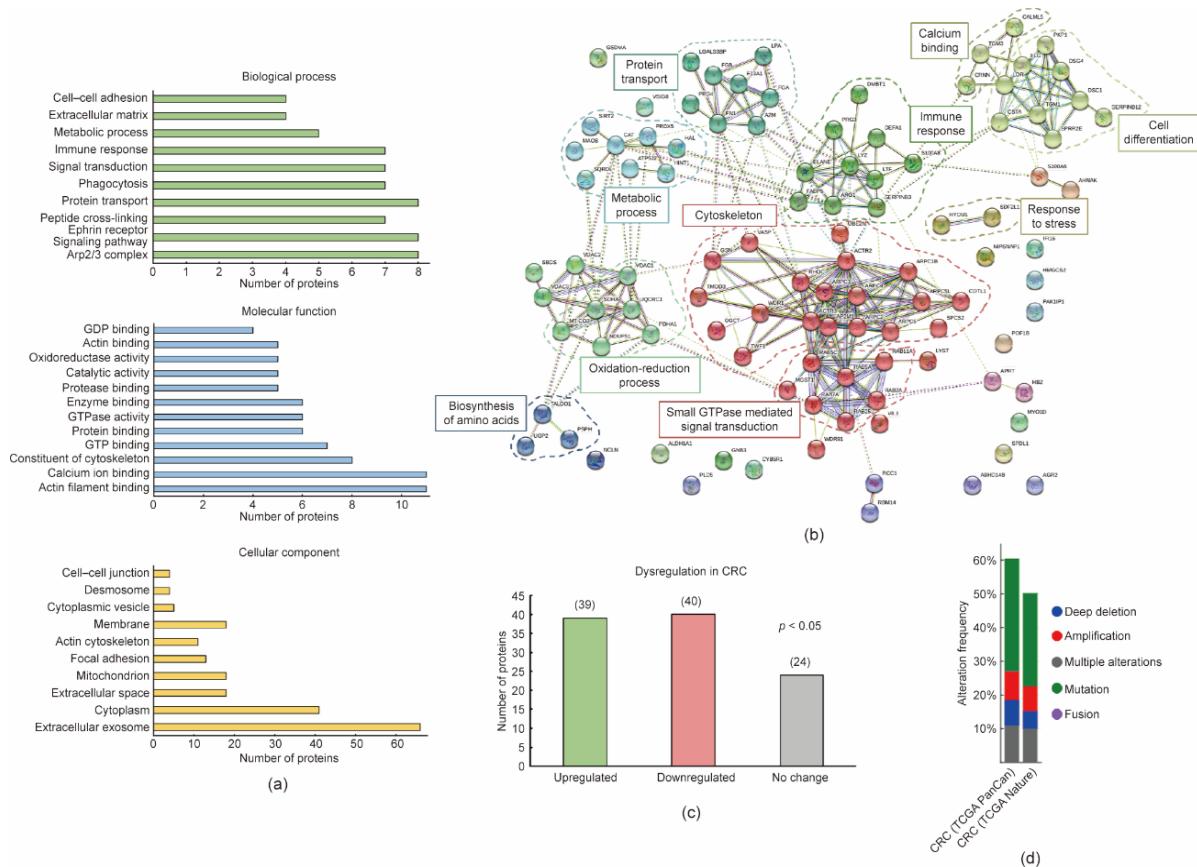


图4. 选择作为CRC特异性潜在TAA的蛋白质的生物信息学和荟萃分析。(a) DAVID数据库确定了103种中已鉴定蛋白质中许多蛋白质涉及的生物过程、分子功能和细胞组分过程。其中一些过程是在癌症发展中发生改变的相关过程和功能，其中一些在癌症中至关重要，如细胞间黏附、代谢过程、蛋白质转运或信号转导。此外，其中一些蛋白质位于外泌体和细胞外空间，突出了它们被免疫系统识别的敏感性。(b) STRING揭示了与CRC发展密切相关的蛋白质之间10个不同的相互作用簇，如细胞通信、蛋白质转运或免疫反应。(c) UALCAN数据库显示，与103种已鉴定的蛋白质中的79种蛋白质的配对健康组织样本相比，CRC组织样本中基因表达的统计学显著 ($p < 0.05$) 失调。(d) cBioPortal揭示了这些蛋白质中以前与CRC相关的102种基因改变 (TCGA Nature 2012 和 PanCancer Atlas)。在这些数据库中登记的组织样本中，有超过50%描述了这些基因的改变。GDP: 鸟嘌呤二核苷酸磷酸; GTP: 鸟苷三磷酸。

CRC患者、具有癌前病变的个体和健康个体的外泌体中蛋白质水平的潜在失调。与非肿瘤配对组织样本 ($n = 17$) 相比, 在肿瘤组织样本中分析的12种蛋白质中的8种被发现具有统计学意义的失调, 其中, SERPINB12、HMGCS2 和 FABP5 下调, ACTR3、TGM3、SPDL1、SDF2L1 和 RAB25 在肿瘤样品中上调[图5 (a) 和 (b)]。有趣的是, 在比较CRC不同阶段的蛋白质水平时, 12种分析蛋白质中有9种存在统计学上的显著差异。SDF2L1、RAB25、PSPH 和 SPDL1 在I和III期(即早期)上调。此外, 我们观察到TGM3 和 ACTR3 在疾病的所有阶段都有所上调[图5 (c)]。相比之下, HMGCS2、FABP5 和 SERPINB12 在疾病的不同阶段发生下调[图5 (d)], 并且它们的水平(连同SERPINB3的水平)随着疾病的进展而降低[图5 (e)]。

针对人血浆外泌体 ($n = 17$), 在分析的12种蛋白质中, 有10种的蛋白质水平在病理样本和对照样本之间存在差异, 其中, SPDL1 和 SERPINB12 上调, 而SERPINB3、ACTR3、MT-CO2、RAB25、HMGCS2、PSPH、TGM3 和 SDF2L1 在病理样本中下调, 倍数变化分别高于1.5和小于0.67 [图6 (a) 和 (b)]。此外, 从具有癌前病变的个体到CRC患者, TGM3、ACTR3、RAB25 和 HMGCS2 的水平随着疾病的进展而降低[图6 (c)]。此外, 在早期(CRC I~II期)和晚期(CRC III~IV期)阶段的CRC患者中发现SERPINB3、SDF2L1、PSPH 和 MT-CO₂ 水平降低[图6 (c)]。最后, SPDL1 在具有癌前病变的个体中下调, 在CRC患者中上调, 而SERPINB12 的水平在疾病进展期间增加[图6 (c)]。

3.4. CRC TAA 在CRC预后中的作用

接下来, 我们使用IHC和荟萃分析来研究所选的15种蛋白质作为CRC预后标志物的作用。由于抗体的可获得性, 我们使用IHC在蛋白质水平上通过与CRC转移或复发相关的四种不同的组织微阵列(TMA)来分析SPDL1、ACTR3 和 SDF2L1 与CRC预后的关系[图7 (a) 和 (b)]。ACTR3 和 SPDL1 显示细胞质染色, 而SDF2L1 显示细胞质和核染色[图7 (a)]。ACTR3 与CRC淋巴结转移显著相关。此外, SPDL1 还可以在统计学上区分有淋巴结转移的CRC患者和无淋巴结转移的CRC患者。有肺或无肝转移的患者与无肺或肝转移的患者之间的差异无统计学意义, 但SPDL1 水平升高与肝、肺转移有关。有趣的是, 核SDF2L1 可以根据复发状态和存活率在统计学上区分CRC患者, 较低的SDF2L1 和最差的预后和复发情况相关[图7 (b)]。

为了将我们的研究扩展到其他12种蛋白质, 我们对CRC患者的总存活率进行了Kaplan-Meier分析, 比较了mRNA低表达和高表达水平之间的差异。有趣的是, long-rank 检验发现S100A8、MT-CO₂、CYB5R1、HMGCS2、FABP5、SDF2L1、ACTR3、SPDL1 和 TGM3 与CRC患者的预后显著相关。研究发现, S100A8、MT-CO₂ 和 CYB5R1 的高mRNA水平与CRC患者的预后相关性较差, 而FABP5、ACTR3、TGM3、SDF2L1、HMGCS2 和 SPDL1 的过表达与较高的存活率相关(附录A中的图S3)。最后, 使用GEPIA2工具研究基因标记与CRC患者和其他癌症患者预后的关系。值得注意的是, 通过用log-rank检验比较高表达和低表达的CRC组之间的差异, 发现mRNA水平与较高的存活率有显著相关性 ($p = 0.0069$) [图7 (c)]。重要的是, 在相同条件下, 还在其他癌症中调查了该基因标记, 以评估其与CRC的相关性。从TCGA项目的其他31个癌症数据集中检索数据后, 标记显示与胃肠道(食管癌、肝癌、胃癌和胰腺癌)和最常见癌症(前列腺癌、乳腺癌和肺癌)的总体存活率之间没有显著相关性[图7 (d) 和 (e)]。另外, 恶性胶质瘤、脑低级别胶质瘤、宫颈鳞状细胞癌、皮肤黑色素瘤、甲状腺癌、子宫癌肉瘤和葡萄膜黑色素瘤与低存活率显著相关。没有发现与其他癌症有关联(附录A中的图S4)。这些结果表明, 该标记与CRC患者较高的总体存活率有关。

因为自身抗原是在疾病时变化的自体蛋白, 许多此类蛋白质在先前被描述为发生改变, 或在CRC中被WB和IHC在mRNA水平和蛋白质水平上观察到失调, 该事实说明了它们作为特定疾病实际TAA的潜在作用。

3.5. 潜在TAA的血清反应性发光分析

接下来, 为了识别自身抗体的靶蛋白, 并具有随后整合到POC设备中的诊断能力, 我们研究了选定的15种蛋白质作为融合蛋白对Halotag的血清反应性(表2)。

我们用基于Halotag的发光微球免疫分析法检测了这些蛋白质对I~IV期CRC患者 ($n=38$)、癌前病变患者 ($n=36$) 和结肠镜检查阴性健康个体 ($n=38$) 血浆标本的血清反应性, 其中, p53 和 cHalo 蛋白分别作为阳性和阴性对照。克隆到pANT7_cHalo载体中的蛋白质与HeLa提取物在体外表达, 并作为其C端与Halotag融合的蛋白质, 在整个Halotag中共价地固定在MB上, 分别与每个血浆样本孵育, 随后与次级HRP-抗人IgG抗体孵育[图8 (a)]。首先, 我们观察到病理组(CRC患者和癌前病变者)与健康组相比, 针对所有这些蛋白质的自身抗体水平

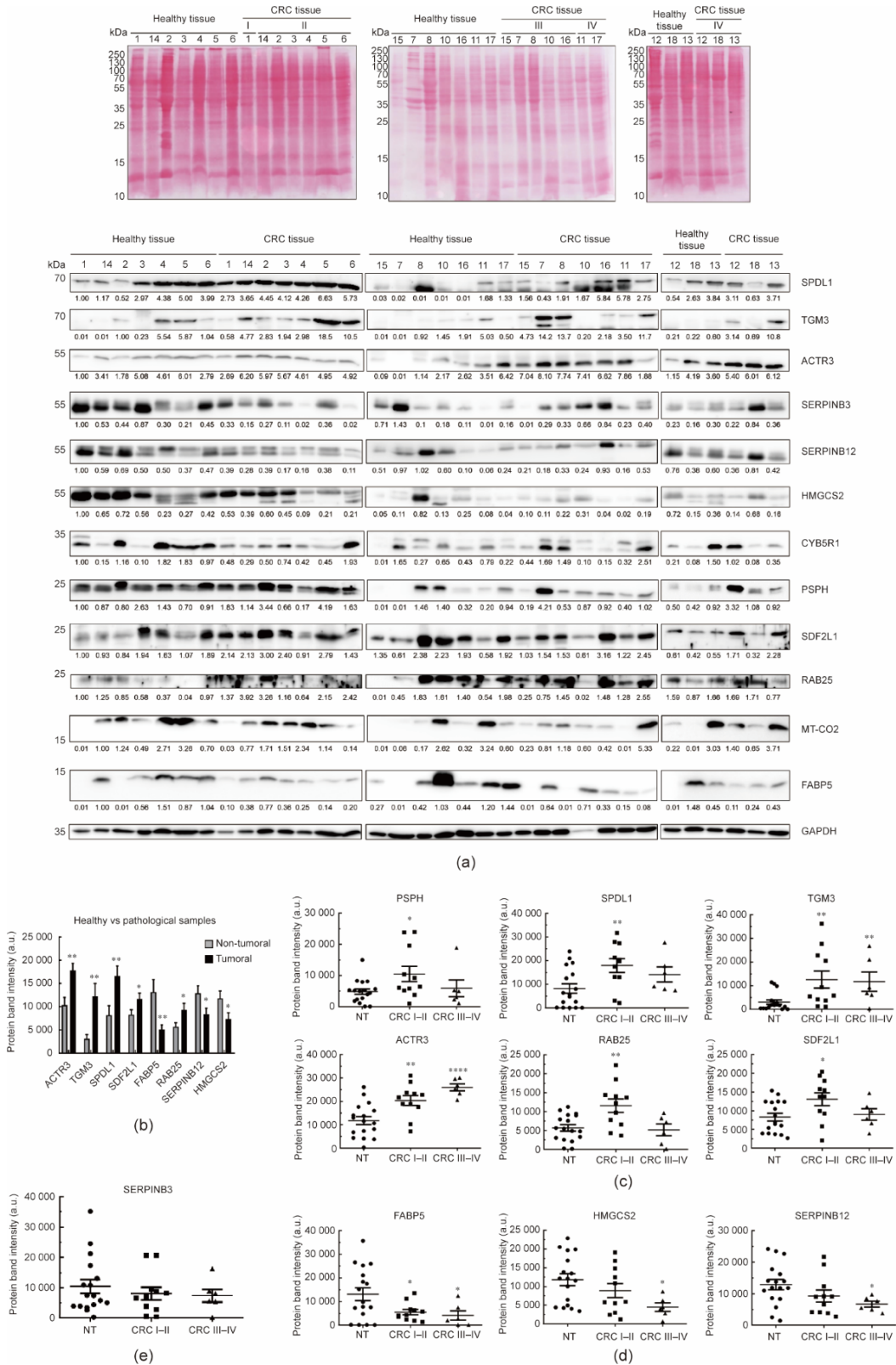


图5. CRC患者组织样本中潜在TAA的蛋白质水平失调分析。(a)分析了硝酸纤维素膜的丽春红染色和来自CRC有12种潜在的TAA患者的配对肿瘤和非肿瘤组织样本的免疫印迹。GAPDH用作测定中的对照。根据它们各自的丽春红染色的总蛋白质泳道含量，使用ImageJ量化的蛋白质条带进行标准化。(b) 12种蛋白质中的5种发现非肿瘤和肿瘤组织样品之间蛋白质水平显著失调，肿瘤样品中SERPINB12和FABP5下调，ACTR3、TGM3和SPDL1上调。(c) 在阶段I和II中，发现蛋白质SDF2L1、RAB25、PSPH和SPDL1的过表达，而TGM3和ACTR3在CRC的不同阶段过表达。(d) 发现蛋白质HMGCS2、FABP5和SERPINB12的蛋白质水平下调，观察到它们随着疾病的进展而减少，SERPINB3的水平也是如此。(e) CYB5R1和MT-CO2的蛋白质水平在肿瘤和非肿瘤组织样本中相似。*： $p < 0.05$ ；**： $p < 0.01$ ；****： $p < 0.0001$ 。I期CRC患者的配对肿瘤/健康正常组织：1、2、3、4、5和14；在第二阶段：6、7、10、17和18；第三阶段：11、12、13、15和16；第四阶段：使用8个。NT：非肿瘤性。

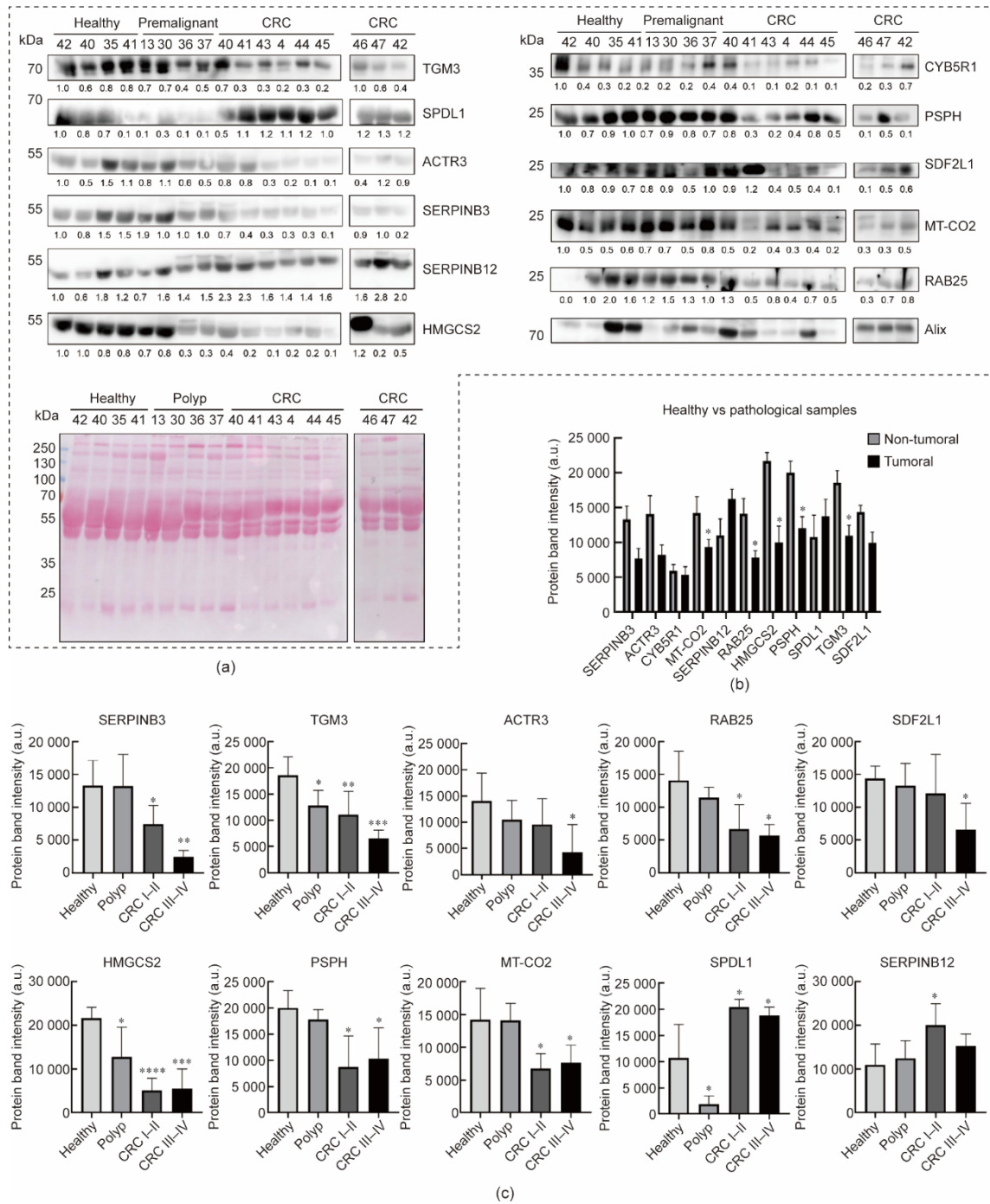


图6. 人血浆样本外泌体中潜在TAA的失调分析。(a) 分析CRC患者、癌前病变个体和具有12种潜在TAA的健康个体的外泌体的WB，并对硝酸纤维素膜进行丽春红染色。Alix在测定中用作细胞外囊泡的对照。使用ImageJ量化的蛋白质条带，根据它们各自的丽春红染色的总蛋白质泳道含量进行标准化。(b) 在分析的12种蛋白质中，有10种发现病理个体（CRC患者和具有癌前病变的个体）和对照（倍数变化 ≥ 1.5 ）之间的蛋白质水平差异，其中，SPDL1和SERPINB12上调，而SERPINB3、ACTR3、MT-CO2、RAB25、HMGCS2、PSPH、TGM3和SDF2L1在病理样本中下调。(c) 在这些蛋白质中，有10种发现了疾病不同阶段的蛋白质调节异常。TGM3、ACTR3、RAB25和HMGCS2蛋白质水平随着疾病的进展而下降。SERPINB3、SDF2L1、PSPH和MT-CO2在早期（CRC I~II）和晚期（CRC III~IV）阶段的CRC患者中过表达。SPDL1在癌前病变个体中下调，在CRC患者中过表达，并且SERPINB12蛋白质水平在疾病进展过程中增加。CYB5R1的蛋白质水平在健康个体和患者中相似，并且在在外显子组样本中未检测到FABP5。分析中不考虑异常值。*： $p < 0.05$ ；**： $p < 0.01$ ；***： $p < 0.001$ ；****： $p < 0.0001$ 。第一阶段CRC样本的外泌体：40、41、46和47；第二阶段：43；III期：4和42；IV期：44和45用于与来自健康和癌前个体的外泌体进行比较。

更高，15种蛋白质中有14种具有统计学意义 ($p < 0.05$) [图8 (b)]。而且，除了SERPINB12、FABP5、SERPINB3和DEFA1之外，所有基因均能在健康对照组和癌

前病变者之间进行统计学区分 ($p < 0.04$)，p53也是如此 [图8 (c)]。此外，与p53不同，除了SERPINB12、RAB2A和FABP5之外，所有基因均能显著区分 ($p <$

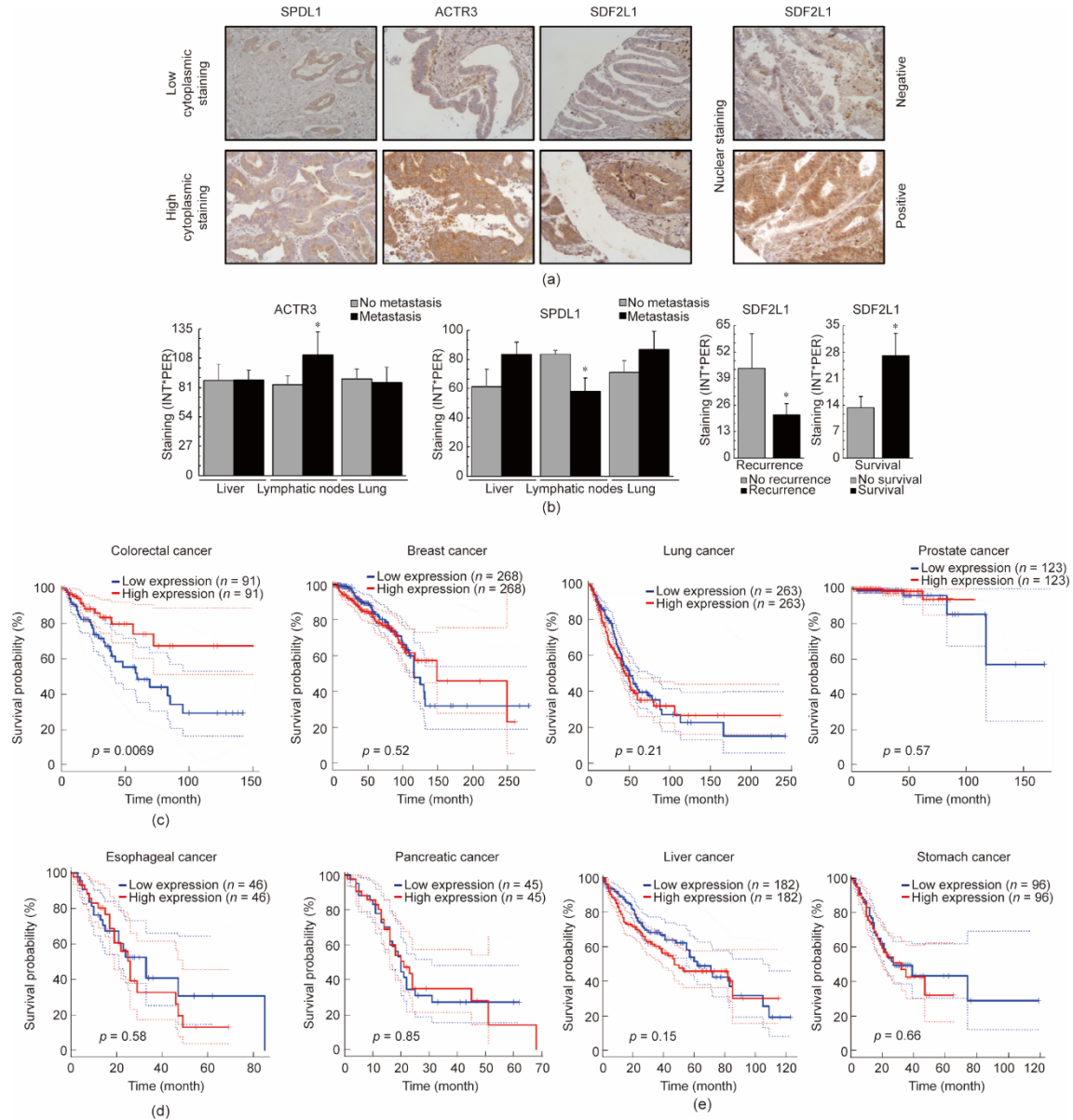


图7. CRC中已识别的自身抗体TAA靶点与预后相关性的评估。(a) 具有代表性的SPDL1、ACTR3和SDF2L1免疫组化染色图像。(b) IHC分析显示ACTR3与淋巴结转移、SDF2L1与复发和生存状态、核SPDL1与淋巴结转移相关(* $p < 0.05$)。(c) ~ (e) 9种蛋白(S100A8、MT-CO2、CYB5R1、FABP5、ACTR3、TGM3、SDF2L1、HMGCS2和SPDL1)的组合与CRC患者存活率的相关性。用GEPIA2研究该标记与CRC和其他癌症预后的关系。我们调查了大肠腺癌(COAD和READ)TCGA数据集,并与乳腺癌(BRCA)、前列腺癌(PRAD)、胃腺癌(STAD)、肺腺癌(LUAD)、食管癌(ESCA)、肝腺癌(LIAD)和胰腺癌(PAAD)作为对照,以研究九个基因特征与CRC预后的特异性关联。

0.04) 健康对照组和CRC患者($p = 0.09$) [图8 (d)]。

接下来,为了进一步研究这些自身抗体的诊断价值,将癌前病变组分为低级别和高级别腺瘤组,以比较它们与健康个体和CRC组的血清反应性[附录A中的图S5 (a)]。首先,低级别和高级别腺瘤组的血清反应性均高于健康个体。 $p53$ 、PSPH、RAB2A、TGM3、ACTR3、CYB5R1、MT-CO2、S100A8、SDF2L1和SPDL1能从统计学上区分低级别恶性腺瘤和健康个体($p < 0.009$), $p53$ 、TGM3、S100A8、SDF2L1和RAB25能够区分高级别恶性腺瘤和健康个体($p < 0.05$)。然而,尽管在低级别和高级别腺瘤

和CRC组之间发现了血清反应强度的差异,但这些TAA都不能从统计学上区分CRC患者和癌前病变者。

总而言之,这些结果表明,SPDL1、FABP5、SERPINB3、DEFA1、CYB5R1、TGM3、ACTR3、SDF2L1、S100A8、RAB25、RAB2A、PSPH、MT-CO2和HMGCS2是自身抗原,其血清反应性可以区分CRC患者和(或)癌前病变者与健康个体。

3.6. 经鉴定的CRC TTA 的诊断价值分析

为了评估15个TAA的诊断价值,我们获得了单独的

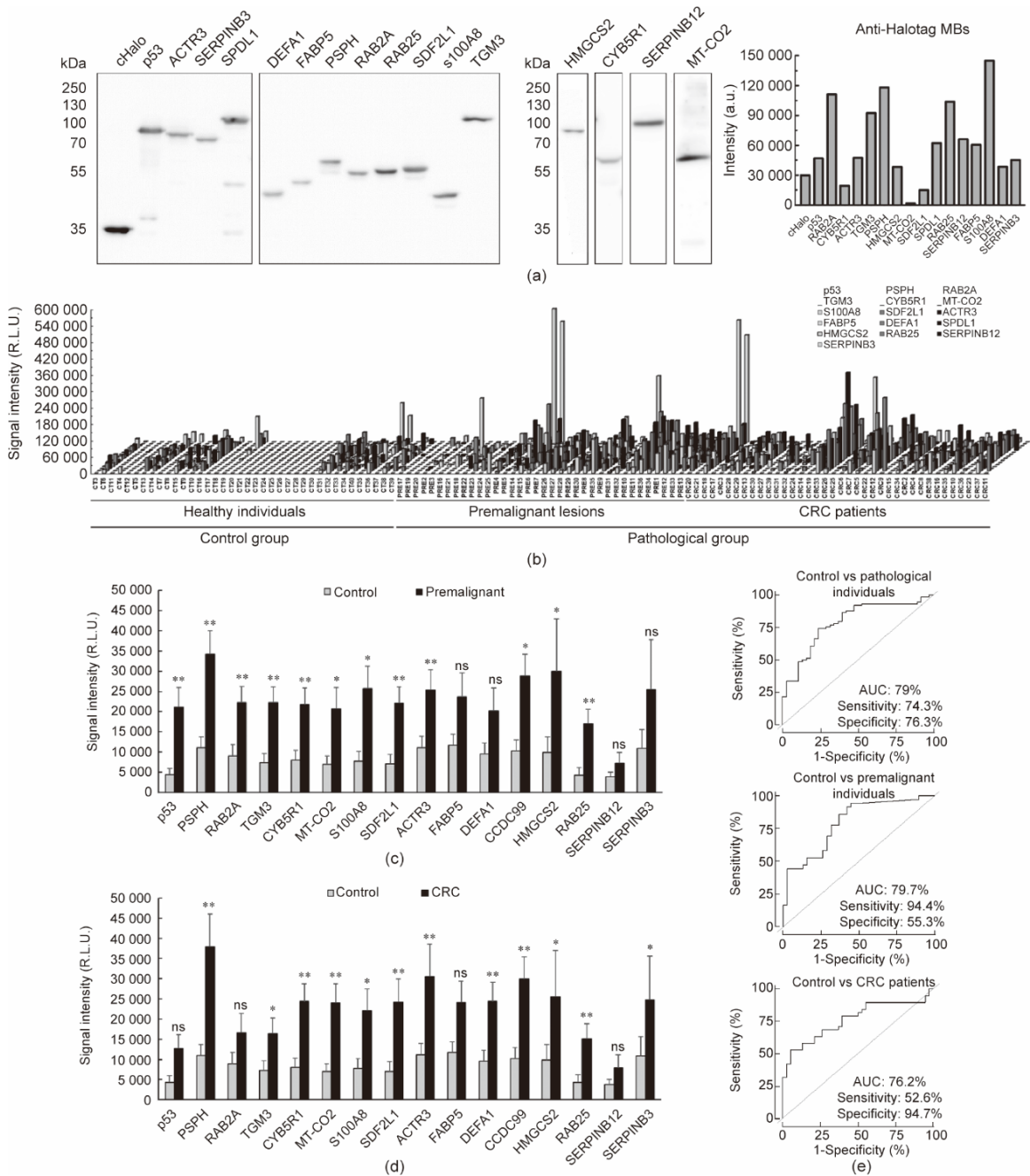


图8. 基于Halotag标记的发光微珠分析与潜在的TAA。(a)经WB鉴定克隆正确,为Halotag融合蛋白,并通过荧光分析检测15种蛋白质的体外表达。所有蛋白质均可被抗Halotag抗体识别。(b)来自CRC患者、癌前病变患者和健康个体的112份个体血浆样本中自身抗体水平的差异(表1和表S1)。除了SERPINB12之外,所有入选的TAA在病理组和对照组之间均有统计学差异($p < 0.05$),P53也有统计学差异($p < 0.05$)。(c)抗PSPH($p = 0.002$)、RAB2A($p = 0.008$)、TGM3($p = 0.005$)、CYB5R1($p = 0.007$)、MT-CO2($p = 0.019$)、S100A8($p = 0.018$)、SDF2L1($p = 0.002$)、ACTR3($p = 0.009$)、SPDL1($p = 0.010$)、HMGCS2($p = 0.039$)、RAB25($p = 0.009$)和p53($p = 0.00$)的血清反应性可以在统计上区分癌前病变患者和健康个体。(d)抗PSPH($p = 0.001$)、TGM3($p = 0.040$)、CYB5R1($p = 0.003$)、MT-CO2($p = 0.004$)、S100A8($p = 0.040$)、SDF2L1($p = 0.004$)、ACTR3($p = 0.007$)、DEFA1($p = 0.008$)、SPDL1($p = 0.003$)、HMGCS2($p = 0.043$)、RAB25($p = 0.003$)和SERPERPERP的自身抗体水平可以在统计上区分健康个体和CRC患者。(e)对14例对潜在TAA的血清反应性不同的患者与健康个体进行诊断价值评估。p53与除SERPINB12外的15种蛋白质联合检测,其AUC为79%,敏感性为74.3%,特异性为76.3%,区分病理个体(癌前病变个体和CRC患者)与健康个体的敏感性和特异性分别为74.3%和76.3%(上图)。联合检测抗PSPH、RAB2A、RAB2A、TGM3、CYB5R1、MT-CO2、S100A8、SDF2L1、ACTR3、SPDL1、HMGCS2、RAB25和p53自身抗体的AUC、敏感性和特异性分别为79.7%、94.4%和55.3%(中图)。PSPH、TGM3、CYB5R1、MT-CO2、S100A8、SDF2L1、ACTR3、DEFA1、SPDL1、HMGCS2、RAB25和SERPINB3联合检测的AUC为76.2%,敏感性为52.6%,特异性为94.7%(下图)。ns:不显著;R.L.U.:相对光单位;*: $p < 0.05$; **: $p < 0.01$ 。

ROC曲线和组合的ROC曲线[图8(e)和附录A中的表S8]。在区分癌前病变者和对照时,所有TAA显示曲线下

面积(AUC)均高于60%,PSPH和SDF2L1显示AUC高于70%,灵敏度和特异性分别高达72.2%和78.9%。此

外,除了p53之外,所有TAA的自身抗体在统计学上区分了CRC患者和健康个体,显示PSPH和SPDL1的最高AUC高达70.5%,敏感性和特异性分别高达76.3%和60.5%。重要的是,具有统计学意义的血清反应性的TAA还显示:在区分病理性CRC患者和对照组方面AUC分别为79%和74.3%,敏感性和特异性分别为74.3%和76.3%;在区分癌前病变者和对照组方面AUC为79.7%,敏感性和特异性分别为94.4%和55.3%;在区分CRC患者和健康个体方面AUC为76.2%,敏感性和特异性分别为52.6%和94.7%。

为进一步分析这些TAA的早期诊断价值,本文获得了低级和高级腺瘤蛋白质的单独和组合ROC曲线,其蛋白质和健康个体对比具有不同血清反应性[附录A中的表S8和图S5(b)]。在区分低级或高级腺瘤个体与健康个体时,TAA显示个体AUC介于55.7%~79.6%之间。综合而言,TAA在鉴别低级腺瘤个体和健康个体时的AUC为84.1%,敏感性和特异性分别为73.7%和81.6%;在区分高级腺瘤和健康个体时,AUC为75%,敏感度和特异性分别为76.5%和78.9%。

这些结果突出了此处确定的CRC TAA作为血液生物标志物在早期结肠直肠病变或CRC检测中的有效性;因此,这些TAA应成为CRC检测诊断组的一部分。相应地,我

们在多路电化学免疫传感平台上分析了TAA,以确定哪些TAA可以包含在POC设备中,并用于临床环境中的CRC检测。

3.7. 鉴定TAA作为CRC自身抗体靶点在类POC设备中的诊断价值分析

在先前的优化条件下[41–42],我们对所开发的生物传感平台检测人血浆中针对已识别TAA的自身抗体存在的能力进行了测试。我们随机分析了两名健康个体、两名癌前病变患者和两名CRC患者的血浆样本。值得注意的是,与癌前病变和CRC患者相比,对照组个体的血清反应性显著降低。更重要的是,正如之前通过荧光观察到的那样,除了p53之外,大多数CRC和癌前病变患者对大于8个TAA有反应。所有癌前病变和CRC个体都具有超过8个CRC自身抗原的自身抗体,并且与健康个体相比显示出清晰的信号,并且高于背景(图9)。总之,这些结果证明了该方法作为检测CRC的潜在临床工具的有效性。更重要的是,我们的结果表明,这些TAA应包含在标记面板中,以便通过简单的血液测试进行CRC检测。

总而言之,这些结果强调了本研究遵循的方法十分有效,其不仅可以用于鉴定具有显著诊断能力的TAA,而且可以找到与CRC进展有关的自身抗体的蛋白靶点,这些蛋白靶点可作为CRC预后标志物。此外,我们的结

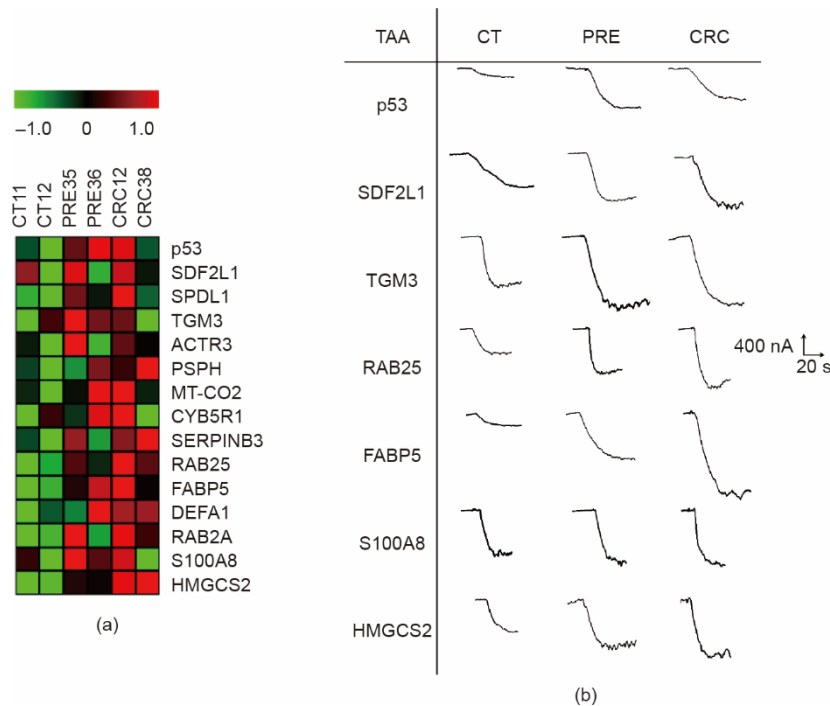


图9. 多路电化学免疫传感平台测量血浆样本中的自身抗体。(a)使用多实验查看器(v4.9.0, 马萨诸塞州, 美国)对随机选择的血浆样本获得的指示蛋白质的单个安培响应进行可视化。CRC患者(CRC)、癌前病变者(PRE)和健康对照者(CT)的安培响应明显不同。绿色信号: 无或低安培响应; 黑色信号: 中度响应; 红色信号: 高安培响应。(b)癌前病变和CRC患者获得的TAA实际安培响应比健康个体大得多, 后者的安培响应显著较低或检测不到。

果表明,开发的发光或生物传感方法克服了蛋白质的纯化、降解和固定化方面的缺陷,因此在通过POC设备检测CRC方面具有很好的潜力。

4. 讨论

全球人口的寿命延长和逐渐老龄化正在导致慢性病显著增加。根据世界卫生组织(World Health Organization)的数据,随着人口老龄化,慢性病在现在和将来都是卫生系统的主要经济负担。慢性病是致残的主要原因,全球63%的死亡是由慢性病造成的。CRC是发病率最高的非传染性慢性病之一。为了减少其对社会和医疗成本的影响,必须对CRC采取全面的预防和控制方法。从这个意义上说,最好的癌症治疗策略包括在癌症预防的早期阶段加强对CRC的检测,同时确定新的特异性诊断生物标志物,因为这一策略可以在病情恶化之前对个体进行治疗,从而提高他们的存活率和生活质量。这样的策略还将降低CRC相关的医疗成本,因为在疾病可以治愈的早期阶段治疗患者,比在治愈可能性严重降低的晚期阶段的治疗费用要便宜得多。为此,有必要在开发新的诊断工具的同时,识别和验证用于早期检测CRC的有效生物标志物,以允许对这些生物标志物进行多项同时检测。这是这项工作的主要目标,重点是基于体液反应的蛋白质生物标志物(可用于CRC的早期诊断)的识别和验证,以及开发新的诊断平台来整合这些生物标志物。

在这项工作中,我们采用了一种多层次集成方法,范围从外泌体和CRC组织样本中TAA的识别和验证,到构建一种能够将识别和验证的TAA整合到一个标记中的POC设备,并用于检测CRC患者和癌前病变者。首先,我们使用了最近开发的一种方法,包括直接与LC-MS/MS和CRC细胞蛋白提取物偶联的IP[14]。其次,这种方法使我们能够通过分析两种研究较少的自身抗原来源来识别与CRC相关的新的自身抗体及其蛋白质靶点。我们分析了外泌体和冷冻组织样本,以研究肿瘤上皮细胞以外的TME。除了之前通过LC-MS/MS[14]或蛋白质微阵列筛选[11]从CRC细胞中鉴定的TAA外,我们还获得了CRC TAA的全局图像。再次,我们用正交技术和荟萃分析验证了这些结果,并证明除了一个已识别的自身抗体靶蛋白外,所有的蛋白质在mRNA和(或)蛋白质水平上都存在失调。此外,这种失调与CRC患者的预后有关。从这个意义上说,我们已经证明了CRC患者体液免疫反应的靶蛋白可以用于诊断和预后。值得注意的是,我们已经确定了一个与良好的CRC预后相关的9个基因特征,这与任

何其他胃肠道肿瘤或高发癌症无关。重要的是,本文中确定的针对TAA的自身抗体能够检测到癌前病变和CRC。因此,这些TAA符合整合到诊断面板中的标准,可以用来检测早期和晚期的CRC病变。这一信息将使临床医生能够监测患有CRC的高危人群。然后,临床医生可以通过液体活检发现潜在的病变,并通过结肠镜检查移除癌前病变,以及在疾病的任何阶段检测CRC。

最后,作为这项工作的主要目标,我们不仅通过最近开发的自组装发光免疫分析验证了15个TAA的差异血清反应性[41-42],并进一步证明了它们在CRC诊断中的有效性,而且还将它们集成到POC生物传感设备平台中。该平台可整合到临床环境中,用于检测CRC和癌前个体。

为了识别TAA,我们检测了TAA的两个互补来源: CRC外泌体和冷冻肿瘤组织中的CRC TME。我们假设,由于在血液和淋巴管中循环的细胞外囊泡是与位于远端器官中进行细胞间通信的关键媒介,肿瘤来源的外泌体应该是发现CRC患者自身抗体靶标TAA的一个有趣的来源。此外,由于TME有助于调节新生抗原的表达和免疫细胞的浸润[20],我们假设对包含癌细胞和肿瘤间质的CRC组织的分析应该是鉴定培养中癌细胞分析遗漏的肿瘤间质新生抗原的一个有趣的来源[14]。重要的是,这些来源在很大程度上还未被用于识别CRC特异性TAA。在本研究中,15个有效的CRC TAA中有8个是从外泌体中鉴定出来的,11个是从CRC肿瘤组织中鉴定出来的(有4个共同的TAA)。这项研究的一个有趣结果是观察到转移能力较低的KM12C和SW480细胞比相应的同基因转移细胞分泌更多的囊泡。尽管从相同体积的条件培养基中分离而出,SW480细胞的外泌体分泌量最高,而KM12SM和KM12L4a细胞的外泌体浓度最低。这些发现可能与这些细胞的转移作用有关,与非转移细胞系的情况相反,这些细胞必须为定植准备好转移生态位[46]。

这项研究的目标之一是确定CRC中在mRNA或蛋白质水平上表现出失调的蛋白质的TAA。我们发现,在CRC患者中,除了一种自身抗体外,所有自身抗体靶蛋白都是失调的。重要的是,我们观察到SERPINB12、SDF2L1、RAB25、ACTR3、TGM3和PSPH在组织和外泌体中表现出相反的失调。这些结果表明这些蛋白质在肿瘤扩散中可能起不同的作用。从这个意义上说,在外泌体中下调而在组织样本中上调的蛋白质,如SDF2L1、ACTR3或TGM3,可能会保留在原发肿瘤中,因为它们是维持肿瘤的必需蛋白质。相反,由于外泌体在远端细胞之间介导细胞通信中发挥作用,在外泌体中上调而在组织样本中下调的蛋白质可能参与转移生态位的建立或维持。这些

发现还表明，外泌体携带的蛋白质的失调可以诱导CRC患者的体液免疫反应，因为这些蛋白质比肿瘤组织的蛋白质更容易暴露于CRC和癌前个体的免疫系统中。

关于经验证的CRC TAAS数据集，SPDL1、SDF2L1、SERPINB3和MT-CO2以前没有被描述为癌症自身抗原，也没有被鉴定为与CRC相关。相反，尽管PSPH之前没有被描述为癌症自身抗原，但它之前被认为与CRC有关，因为它被描述为染色体7p上的一个潜在驱动基因[47]。FABP5、S100A8、RAB2A、ACTR3、CYB5R1、TGM3、HMGCS2和DEFA1中也存在类似的发现：这些蛋白质被描述为在CRC中发生了改变，而不是被描述为自身抗体的癌症自身抗原靶标。FABP5已被描述为具有促进细胞生长和CRC细胞转移的潜能[48]。S100A8参与了CRC形成[49]，RAB25和RAB2A分别被描述为CRC的抑癌基因和新的癌基因[50–51]。ACTR3已被发现在CRC和转移中发生改变[52–53]，而HMGCS2已被描述为可增强CRC的侵袭和转移[54]。CYB5R1已被描述为与上皮细胞向间充质转化（EMT）相关，并被认为是CRC的不良预后[55]。据报道，TGM3在CRC的发展过程中作为肿瘤的抑制因子[56]。此外，还发现DEFA1在CRC中过表达[57]。值得注意的是，上述所有蛋白质都是CRC自身抗原，可作为自身抗体的靶点，并可以在疾病的早期和晚期将CRC患者与健康个体区分开来，包括癌前病变（低级别和高级别腺瘤）患者。

最后，本研究的主要目标是识别和验证用于CRC诊断的CRC TAA，并将其集成到对临床常规有用的POC设备中。需要检测多种自身抗体，以便更特异、更敏感地诊断CRC。本文展示的血清反应发光和生物传感方法可以成功用于区分CRC患者和癌前病变患者与健康个体对照组。据我们所知，这是第一份研究特定自身抗体标记的报告，该标记包括从细胞系和（或）直接从患者冷冻组织中分离的外泌体中抗TAA的自身抗体，并已集成到电化学生物平台中进行自身抗体检测，目的是实现患者护理。

关于该标记的性能比较，与血液中循环肿瘤DNA（ctDNA）标志物（即新兴标志物）或CEA（已建立的临床标志物）的测定相比，我们观察到，尽管ctDNA标志物和CEA可以区分具有高AUC、高特异性和高敏感性的CRC（特别是在疾病的晚期），但它们不足以发现癌前或癌前期病变，其AUC和敏感性分别低于60%和50% [58–59]。这些数据进一步证明，除了简化最终方法外，免疫系统的持续生物信号放大功能使自身抗体筛查在肿瘤的早期诊断中具有独特的优势。从这个意义上说，TAA联合检测可以将CRC或癌前病变个体与健康个体区分开来，

两者的AUC都接近80%。然而，应该强调的是，我们的目标是将此处确定的TAA与以前确定的CRC TAA [11,13–14,60]进行整合，以开发能够整合所有CRC特异性TAA的诊断平台，从而获得对癌前病变患者和CRC患者的高度诊断能力。这将使通过一种简单、非侵入性、快速和廉价的血液检测来检测CRC成为可能。

值得注意的是，尽管ctDNA标志物在诊断CRC患者中的敏感性和特异性已被证明高于本研究中通过自身抗体检测获得的敏感性和特异性，但应该强调的是，比基于Halotag的电化学生物传感平台相比，这些ctDNA的分离、纯化和检测程序更复杂、更长、更昂贵。此外，由于抗原-抗体相互作用的特异性和亲和力较高，定量ctDNA所需的血浆样本量（每10~15 ng ctDNA需要1.5 mL血浆）比检测自身抗体所需的血浆样本量（1~5 μ L的血浆样本稀释1:400）要大得多。还应该提及的是，电化学生物平台最近已经通过有限数量样本初步显示出区分CRC患者的健康组织和肿瘤组织的潜力，其敏感性和特异性分别为89%和67%，通过在全球水平上靶向5-mC和5-hmC来区分CRC患者的健康组织和肿瘤组织，且不需要亚硫酸氢盐治疗和聚合酶链反应（PCR）扩增，但需要基因组DNA提取和变性[61]。因此，这里使用的生物传感平台可以整合不同特征的标志物，以提高检测方法对CRC早期和晚期诊断的敏感性和特异性。

5. 结论

在本研究中，我们通过IP和LC-MS/MS分析了从CRC患者血浆中分离出IgG，通过鉴定未被研究的CRC来源（即外泌体和CRC冷冻组织中的TME），确定了一种基于血液的CRC诊断特征。在这项研究中，我们通过IP和LC-MS/MS分析，确定了一个基于血液的CRC诊断特征。这些候选TAA可用于CRC基于血液诊断的多重生物标志物平台的一部分，并集成在本文所述的生物传感器上。由此产生的平台将是CRC诊断的理想平台，因为它将允许使用最小的血浆样本体积同时进行自身抗体检测。总之，本文提出的新型多功能电化学生物传感策略满足了临床常规和POC检测的要求，大大减少了检测时间和生产成本。

虽然我们通过比较CRC和癌前病变患者与健康人血浆，无偏地发现了自身抗体标志物，但这些生物标志物的诊断能力应该进一步通过使用大量的样本和独立的患者队列来验证。进一步的研究必定会通过复用本文描述的TAA和先前验证的CRC TTA来获得诊断平台[14,41–42,44]。这

种多路生物（传感）平台（如本文所述，集成了15个CRC TAA）对于基于血液的CRC临床诊断应该具有足够高的灵敏度、特异性和AUC，从而显示出极大的提高患者存活率和降低卫生系统治疗CRC相关成本的潜力。

Acknowledgements

This work was supported by the financial support of the PI17CIII/00045 and PI20CIII/00019 grants from the AES-ISCIII program to R.B. The financial support of the PID2019-103899RB-I00 (Ministerio de Ciencia e Innovación) Research Project and the TRANSNANOAVANSENS-CM Program from the Comunidad de Madrid (S2018/NMT-4349) to S.C. are gratefully acknowledged. G.D. acknowledges the financial support of PI15/00246 grant of the FIS and Cátedra UAM-Roche en Medicina de Innovación. The FPU predoctoral contract to A.M.-C. is supported by the Spanish Ministerio de Educación, Cultura y Deporte. G.S.-F. is recipient of a predoctoral contract (1193818 N) supported by The Flanders Research Foundation (FWO). M.A.-N. was supported by a contract of the Programa Operativo de Empleo Juvenil la Iniciativa de Empleo Juvenil (YEI) with the participation of the Consejería de Educación, Juventud y Deporte de la Comunidad de Madrid y del Fondo Social Europeo. The predoctoral contract from the Spanish Ministerio de Economía y Competitividad (BES-2016-076606, E.P.) and Talento-Contract from Comunidad de Madrid (2019-T2/IND-15965, R.M.T.-R.) are also gratefully acknowledged.

Compliance with ethics guidelines

Ana Montero-Calle, Itziar Aranguren-Abeigón, María Garranzo-Asensio, Carmen Poves, María Jesús Fernández-Aceñero, Javier Martínez-Useros, Rodrigo Sanz, Jana Dziková, Javier Rodríguez-Cobos, Guillermo Solís-Fernández, Eloy Povedano, María Gamella, Rebeca Magnolia Torrente-Rodríguez, Miren Alonso-Navarro, Vivian de los Ríos, J. Ignacio Casal, Gemma Domínguez-Muñoz, Ana Guzmán-Aránguez, Alberto Peláez-García, José Manuel Pingarrón, Susana Campuzano, and Rodrigo Barderas declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.eng.2021.04.026>.

References

- [1] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68(6):394–424.
- [2] Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal cancer statistics, 2020. *CA Cancer J Clin* 2020;70(3):145–64.
- [3] www.cancer.org [Internet]. Atlanta: American Cancer Society; 2020[cited Feb 5, 2021]. Available from: <https://www.cancer.org/cancer/colon-rectal-cancer>.
- [4] Balchen V, Simon K. Colorectal cancer development and advances in screening. *Clin Interv Aging* 2016;11:967–76.
- [5] Lech G, Slotwiński R, Słodkowski M, Krasnodębski IW. Colorectal cancer tumour markers and biomarkers: recent therapeutic advances. *World J Gastroenterol* 2016;22(5):1745–55.
- [6] Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, et al. Colorectal cancer. *Nat Rev Dis Primers* 2015;1(1):15065.
- [7] Ignacio Casal J, Barderas R. Identification of cancer autoantigens in serum: toward diagnostic/prognostic testing? *Mol Diagn Ther* 2010;14(3):149–54.
- [8] Anderson KS, LaBaer J. The sentinel within: exploiting the immune system for cancer biomarkers. *J Proteome Res* 2005;4(4):1123–33.
- [9] Barderas R, Villar-Vázquez R, Fernández-Aceñero MJ, Babel I, Peláez-García A, Torres S, et al. Sporadic colon cancer murine models demonstrate the value of autoantibody detection for preclinical cancer diagnosis. *Sci Rep* 2013;3(1):2938.
- [10] Macdonald IK, Parsy-Kowalska CB, Chapman CJ. Autoantibodies: opportunities for early cancer detection. *Trends Cancer* 2017;3(3):198–213.
- [11] Barderas R, Babel I, Casal JI. Colorectal cancer proteomics, molecular characterization and biomarker discovery. *Proteomics Clin Appl* 2010;4(2):159–78.
- [12] Rodrigo B, Roi VV. Colorectal Cancer Circulating Biomarkers. In: Preedy V, Patel V, editors. *Biomarkers in Cancer*. Amsterdam: Springer Netherlands; 2015. p. 1–21.
- [13] Babel I, Barderas R, Díaz-Uriarte R, Martínez-Torrecuadrada JL, Sánchez-Carbayo M, Casal JI. Identification of tumor-associated autoantigens for the diagnosis of colorectal cancer in serum using high density protein microarrays. *Mol Cell Proteomics* 2009;8(10):2382–95.
- [14] Garranzo-Asensio M, San Segundo-Acosta P, Poves C, Fernández-Aceñero MJ, Martínez-Useros J, Montero-Calle A, et al. Identification of tumor-associated antigens with diagnostic ability of colorectal cancer by in-depth immunomic and seroproteomic analysis. *J Proteomics* 2020;214:103635.
- [15] Colombo M, Raposo G, Théry C. Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. *Annu Rev Cell Dev Biol* 2014;30(1):255–89.
- [16] Vader P, Breakefield XO, Wood MJ. Extracellular vesicles: emerging targets for cancer therapy. *Trends Mol Med* 2014;20(7):385–93.
- [17] Wee I, Syn N, Sethi G, Goh BC, Wang L. Role of tumor-derived exosomes in cancer metastasis. *Biochim Biophys Acta Rev Cancer* 2019;1871(1):12–9.
- [18] Steinbichler TB, Dudás J, Riechelmann H, Skvortsova II. The role of exosomes in cancer metastasis. *Semin Cancer Biol* 2017;44:170–81.
- [19] Syn N, Wang L, Sethi G, Thiery JP, Goh BC. Exosome-mediated metastasis: from epithelial-mesenchymal transition to escape from immunosurveillance. *Trends Pharmacol Sci* 2016;37(7):606–17.
- [20] Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med* 2018;24(8):1277–89.
- [21] Vogiazzi V, de la Cruz A, Mishra S, Shanov V, Heineman WR, Dionysiou DD. A comprehensive review: development of electrochemical biosensors for detection of cyanotoxins in freshwater. *ACS Sens* 2019;4(5):1151–73.
- [22] Cui F, Zhou Z, Zhou HS. Measurement and analysis of cancer biomarkers based on electrochemical biosensors. *J Electrochem Soc* 2020;167(3):037525.
- [23] Campuzano S, Barderas R, Pedrero M, Yáñez-Sedeño P, Pingarrón JM.

- Electrochemical biosensing to move forward in cancer epigenetics and metastasis: a review. *Anal Chim Acta* 2020;1109:169–90.
- [24] Morikawa K, Walker SM, Nakajima M, Pathak S, Jessup JM, Fidler IJ. Influence of organ environment on the growth, selection, and metastasis of human colon carcinoma cells in nude mice. *Cancer Res* 1988;48(23):6863–71.
- [25] Morikawa K, Walker SM, Jessup JM, Fidler IJ. *In vivo* selection of highly metastatic cells from surgical specimens of different primary human colon carcinomas implanted into nude mice. *Cancer Res* 1988;48(7):1943–8.
- [26] Barderas R, Babel I, Díaz-Uriarte R, Moreno V, Suárez A, Bonilla F, et al. An optimized predictor panel for colorectal cancer diagnosis based on the combination of tumor-associated antigens obtained from protein and phage microarrays. *J Proteomics* 2012;75(15):4647–55.
- [27] Wiśniewski JR, Gaugaz FZ. Fast and sensitive total protein and peptide assays for proteomic analysis. *Anal Chem* 2015;87(8):4110–6.
- [28] Käll L, Canterbury JD, Weston J, Noble WS, MacCoss MJ. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat Methods* 2007;4(11):923–5.
- [29] Mellacheruvu D, Wright X, Couzens AL, Lambert JP, St-Denis NA, Li T, et al. The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat Methods* 2013;10(8):730–6.
- [30] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;47(D1):D607–13.
- [31] Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, et al. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* 2007;35(suppl_2):W169–75.
- [32] Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* 2003;4(9):R60.
- [33] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, BvskChakravathi, et al. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* 2017;19(8):649–58.
- [34] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012;2(5):401–4.
- [35] Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, et al. Towards a knowledge-based Human Protein Atlas. *Nat Biotechnol* 2010;28(12):1248–50.
- [36] Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. *Science* 2015;347(6220):1260419.
- [37] Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res* 2019;47(W1):W556–60.
- [38] Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res* 2017;45(W1):W98–102.
- [39] Garranzo-Asensio M, San Segundo-Acosta P, Martínez-Useros J, Montero-Calle A, Fernández-Aceñero MJ, Häggmark-Månberg A, et al. Identification of prefrontal cortex protein alterations in Alzheimer’s disease. *Oncotarget* 2018;9(13):10847–67.
- [40] Seiler CY, Park JG, Sharma A, Hunter P, Surapaneni P, Sedillo C, et al. DNASU plasmid and PSI: biology-materials repositories: resources to accelerate biological research. *Nucleic Acids Res* 2014;42(D1):D1253–60.
- [41] Garranzo-Asensio M, Guzman-Arangué A, Povés C, Fernández-Aceñero MJ, Torrente-Rodríguez RM, Ruiz-Valdepeñas Montiel V, et al. Toward liquid biopsy: determination of the humoral immune response in cancer patients using haloTag fusion protein-modified electrochemical bioplatfroms. *Anal Chem* 2016;88(24):12339–45.
- [42] Garranzo-Asensio M, Guzmán-Arangué A, Povedano E, Ruiz-Valdepeñas Montiel V, Povés C, Fernández-Aceñero MJ, et al. Multiplexed monitoring of a novel autoantibody diagnostic signature of colorectal cancer using HaloTag technology-based electrochemical immunosensing platform. *Theranostics* 2020;10(7):3022–34.
- [43] Montero-Calle A, San Segundo-Acosta P, Garranzo-Asensio M, Rábano A, Barderas R. The molecular misreading of APP and UBB induces a humoral immune response in Alzheimer’s disease patients with diagnostic ability. *Mol Neurobiol* 2020;57(2):1009–20.
- [44] Garranzo-Asensio M, Guzmán-Arangué A, Povés C, Fernández-Aceñero MJ, Montero-Calle A, Ceron MA, et al. The specific seroreactivity to Δ Np73 isoforms shows higher diagnostic ability in colorectal cancer patients than the canonical p73 protein. *Sci Rep* 2019;9(1):13547.
- [45] San Segundo-Acosta P, Montero-Calle A, Fuentes M, Rábano A, Villalba M, Barderas R. Identification of Alzheimer’s disease autoantibodies and their target biomarkers by phage microarrays. *J Proteome Res* 2019;18(7):2940–53.
- [46] Guo Y, Ji X, Liu J, Fan D, Zhou Q, Chen C, et al. Effects of exosomes on pre-metastatic niche formation in tumors. *Mol Cancer* 2019;18(1):39.
- [47] Sato K, Masuda T, Hu Q, Tobo T, Kidogami S, Ogawa Y, et al. Phosphoserine phosphatase is a novel prognostic biomarker on chromosome 7 in colorectal cancer. *Anticancer Res* 2017;37(5):2365–71.
- [48] Kawaguchi K, Senga S, Kubota C, Kawamura Y, Ke Y, Fujii H. High expression of fatty acid-binding protein 5 promotes cell growth and metastatic potential of colorectal cancer cells. *FEBS Open Bio* 2016;6(3):190–9.
- [49] Duan L, Wu R, Ye L, Wang H, Yang X, Zhang Y, et al. S100A8 and S100A9 are associated with colorectal carcinoma progression and contribute to colorectal carcinoma cell survival and migration via Wnt/ β -catenin pathway. *PLoS One* 2013;8(4):e62092.
- [50] Zheng XF, Liu KX, Wang XM, Zhang R, Li X. MicroRNA192 acts as a tumor suppressor in colon cancer and simvastatin activates miR192 to inhibit cancer cell growth. *Mol Med Rep* 2019;19(3):1753–60.
- [51] Goldenring JR, Nam KT. Rab25 as a tumour suppressor in colon carcinogenesis. *Br J Cancer* 2011;104(1):33–6.
- [52] Otsubo T, Iwaya K, Mukai Y, Mizokami Y, Serizawa H, Matsuoka T, et al. Involvement of Arp2/3 complex in the process of colorectal carcinogenesis. *Mod Pathol* 2004;17(4):461–7.
- [53] Mendes M, Peláez-García A, López-Lucendo M, Bartolomé RA, Calviño E, Barderas R, et al. Mapping the spatial proteome of metastatic cells in colorectal cancer. *Proteomics* 2017;17(19):1700094.
- [54] Chen SW, Chou CT, Chang CC, Li YJ, Chen ST, Lin IC, et al. HMGCS2 enhances invasion and metastasis via direct interaction with PPAR α to activate Src signaling in colorectal cancer and oral cancer. *Oncotarget* 2017;8(14):22460–76.
- [55] Woischke C, Blaj C, Schmidt EM, Lamprecht S, Engel J, Hermeking H, et al. CYB5R1 links epithelial-mesenchymal transition and poor prognosis in colorectal cancer. *Oncotarget* 2016;7(21):31350–60.
- [56] Feng Y, Ji D, Huang Y, Ji B, Zhang Y, Li J, et al. TGM3 functions as a tumor suppressor by repressing epithelial-mesenchymal transition and the PI3K/AKT signaling pathway in colorectal cancer. *Oncol Rep* 2020;43(3):864–76.
- [57] Radeva MY, Jahns F, Wilhelm A, Gleis M, Settmacher U, Greulich KO, et al. Defensin alpha 6 (DEFA 6) overexpression threshold of over 60 fold can distinguish between adenoma and fully blown colon carcinoma in individual patients. *BMC Cancer* 2010;10(1):588.
- [58] Luo H, Zhao Q, Wei W, Zheng L, Yi S, Li G, et al. Circulating tumor DNA methylation profiles enable early diagnosis, prognosis prediction, and screening for colorectal cancer. *Sci Transl Med* 2020;12(524):eaax7533.
- [59] Palmqvist R, Engarås B, Lindmark G, Hallmans G, Tavelin B, Nilsson O, et al. Prediagnostic levels of carcinoembryonic antigen and CA 242 in colorectal cancer: a matched case-control study. *Dis Colon Rectum* 2003;46(11):1538–44.
- [60] Babel I, Barderas R, Díaz-Uriarte R, Moreno V, Suarez A, Salazar R, et al. Identification of MST1/STK4 and SULF1 proteins as autoantibody targets for the diagnosis of colorectal cancer by using phage microarrays. *Mol Cell Proteomics* 2011;10(3):M110.001784.
- [61] Povedano E, Montiel VR, Valverde A, Navarro-Villoslada F, Yáñez-Sedeño P, Pedrero M, et al. Versatile electroanalytical bioplatfroms for simultaneous determination of cancer-related DNA 5-methyl- and 5-hydroxymethyl-cytosines at global and gene-specific levels in human serum and tissues. *ACS Sens* 2019;4(1):227–34.