

## 学术前沿专栏:肿瘤早期诊断蛋白质组生物标志物研究

**【特约主持人】**解新芳:西安交通大学第一附属医院教授,获国家级青年人才项目

**【主持人按语】**癌症严重威胁全球公共卫生,肺癌与结直肠癌作为常见恶性肿瘤,其高致死率凸显早期诊断的关键意义.当下迫切需要探寻高灵敏、高特异且无创或微创的诊断标志物.外泌体携带丰富生物分子参与肿瘤进程,在肺癌研究中,发现血浆外泌体蛋白可区分患者与健康人群,部分蛋白与肺癌进展紧密相连;而在结直肠癌领域,非标记定量蛋白质组学揭示血浆蛋白质糖基化修饰差异可有效鉴别两者,特定糖蛋白的糖肽表达与疾病进展显著相关.本专栏聚焦此类研究,剖析肿瘤早期诊断新路径,为癌症防治提供科学依据.

# 结直肠癌的血浆糖蛋白质组学研究

赵洋<sup>1</sup>,曾嘉明<sup>1,2</sup>,赵佳威<sup>3</sup>,孟波<sup>2</sup>

(1.中国计量科学研究院 前沿计量科学中心,北京 100029;2.沈阳化工大学 化学工程学院,沈阳 110142;3.中国计量大学 生命科学学院,杭州 310018)

**摘要:**结直肠癌是全球最常见的恶性肿瘤之一,其致死风险与疾病的早期诊断密切相关.为了提高患者的生存率并降低结直肠癌的致死率,筛查有效的生物标志物用于结直肠癌早期诊断显得尤为重要.研究采用非标记定量蛋白质组学技术,深入分析了结直肠癌患者与健康人群之间血浆样本的蛋白质糖基化修饰组差异.结果鉴定出 152 个与结直肠癌有关的差异表达完整 N-糖肽,能够有效区分健康人群和结直肠癌患者.进一步研究发现,KNG1、THRB、FGB、APOH、ORM2、AHSG、ORM1 和 SERPIND1 等糖蛋白质的 N-糖肽差异表达与结直肠癌的进展显著相关,为结直肠癌的早期诊断与筛查提供了重要数据支撑.

**关键词:**结直肠癌;血浆;蛋白质组;糖基化修饰;LC-MS/MS

**中图分类号:**Q592

**文献标志码:**A

**文章编号:**1000-2367(2025)04-0022-09

结直肠癌(colorectal cancer, CRC)是一种极具威胁的胃肠道恶性肿瘤.在全球范围内, CRC 的发病率位居第三(占发病总数的 10.2%),死亡率排名第二(占死亡总数的 9.4%)<sup>[1]</sup>.据统计,预计到 2030 年,结直肠癌的患病人数将达到 220 万例,死亡人数将达到 110 万例<sup>[2]</sup>.流行病学调查研究表明,结直肠癌的发病率随着年龄的增长而逐渐升高,同时还受到家族史、炎症性肠病、吸烟、过度饮酒、肥胖和糖尿病等多种因素的影响<sup>[3]</sup>.结直肠癌从良性腺瘤病变(息肉)发展成癌症需要十多年的时间<sup>[4]</sup>.在结直肠癌患者中,原位结直肠癌(I 期)的 5 年相对生存率高达 90%,而转移性结直肠癌(IV 期)的 5 年相对生存率仅为 10%<sup>[5]</sup>,因此早期诊断对于提高患者生存率和降低肿瘤转移率至关重要.目前,结直肠癌的诊断方法主要包括结肠镜检查<sup>[6]</sup>、胶囊

**收稿日期:**2023-08-11;**修回日期:**2023-09-13.

**基金项目:**国家重点研发计划项目(2022YFF0608400);中国计量科学研究院基本业务费重点项目(AKYZD2111).

**作者简介(通信作者):**赵洋(1997-),男,四川泸州人,中国计量科学研究院副研究员,研究方向为肿瘤的蛋白质组学, E-mail: zhaoy@nim.ac.cn.

**引用本文:**赵洋,曾嘉明,赵佳威,等.结直肠癌的血浆糖蛋白质组学研究[J].河南师范大学学报(自然科学版),2025, 53(4):22-30.(Zhao Yang, Zeng Jiaming, Zhao Jiawei, et al. Plasma glycoproteomic analysis of colorectal cancer [J]. Journal of Henan Normal University(Natural Science Edition), 2025, 53(4):22-30. DOI: 10.16366/j.cnki.1000-2367.2023.08.11.0002.)

内窥镜<sup>[7]</sup>、CT扫描<sup>[6]</sup>、粪便隐血检测(FOBT)<sup>[8]</sup>以及癌胚抗原(CEA)和碳水化合物抗原199(CA19-9)<sup>[9]</sup>等血清标志物检测。尽管结肠镜检查被认为是诊断结直肠癌的金标准,具备较高的准确性,但由于其侵入性较强,患者的接受度有限<sup>[10]</sup>。目前其余检测方法在灵敏度和特异性等方面仍有不足,不能满足临床诊断需求。因此,迫切需要建立更微创、高灵敏度和高特异性的诊断方法,提高结直肠癌的早期诊断率和治疗成功率。

生物标志物是存在于血液、尿液和脑脊液等体液或组织中的一种生物分子,能够反映肿瘤疾病在特定时间内的生理状态和疾病进程的变化<sup>[11]</sup>。癌症生物标志物涵盖了多种生化分子,包括DNA、RNA、蛋白质以及代谢产物等,在癌症的风险评估、早期诊断、精准预后和治疗效果评价等方面发挥着重要作用<sup>[12]</sup>。糖基化修饰是细胞表面和分泌蛋白上的一种常见且结构复杂的翻译后修饰类型,在多种生物过程中都起到至关重要的调控作用。例如,它不仅帮助蛋白质实现正确的折叠以确保其功能,还涉及细胞的附着、移动、信号传递以及在免疫过程中的识别等关键的生物过程<sup>[13]</sup>。在肿瘤发展过程中,糖基化修饰经常发生异常变化,导致肿瘤恶性特征的出现,如细胞的增殖、迁移和侵袭等<sup>[14]</sup>。DOHERTY等<sup>[15]</sup>研究表明,结直肠癌患者在不同的癌症阶段其血浆N-糖基化修饰具有显著的统计学差异。与健康人群相比,结直肠癌患者血液中核心岩藻糖类的双触角糖F(6)A2G2和F(6)A2G2S(6)1显著降低( $P < 0.0009$ )。PARK等<sup>[16]</sup>已证明膜锚定蛋白质受体上的 $\alpha$ 2,6-联结的唾液酸修饰能够调控结直肠癌细胞的黏附、迁移和侵袭能力。总体而言,蛋白质糖基化修饰的研究对结直肠癌早期诊断具有重要意义,有望为开发更加准确、灵敏的肿瘤早期诊断方法提供新的途径。

血液的动态变化能够敏锐地反映患者疾病的病理生理状态,提供从时间和空间上监测疾病进展的独特视角。血液检测已经融入常规生化检测,其独特之处在于稳定性良好、无创性、易获取性和长期保存等优点。然而,CEA和CA19-9虽已被用作结直肠癌检测和治疗监测的血液标志物,但它们的敏感性仅为40%~70%,特异性也仅为73%~90%<sup>[17]</sup>。因此,为了提升结直肠癌诊断的准确性,亟须探索更加精确的候选生物标志物。综上,本研究基于超高分辨质谱分析技术,对结直肠癌患者的血浆样本进行了系统性的蛋白质糖基化修饰组研究,深入探索结直肠癌患者血浆中糖基化修饰的异常改变,筛选具有高灵敏度、强特异性且易于筛查检测的候选糖基化修饰生物标志物,为结直肠癌的早期诊断和预后筛查提供分子病理学支持。

## 1 材料与方法

### 1.1 实验材料、仪器和研究对象

结直肠癌血浆糖蛋白质组学研究所用实验材料和仪器见附录表S1。结直肠癌患者和健康人的血浆样本均由中国医学科学院肿瘤医院提供,本研究经中国医学科学院肿瘤医院伦理委员会审批,所有参与本项研究的受试者均签署书面知情同意书。

### 1.2 糖蛋白质组分析的样品制备

使用Pierce<sup>TM</sup> TOP 14试剂盒(赛默飞世尔科技)去除血浆高丰度蛋白,取10  $\mu$ L血浆样品,采用试剂盒去除其前14种高丰度蛋白质。用过滤器辅助样品制备方案(filter-aided sample preparation, FASP)酶切样品蛋白<sup>[18-19]</sup>。首先,将已去除高丰度蛋白样品转移到30 kDa超滤管中并离心15 min,加入尿素溶液(8 mol/L尿素溶解在0.1 mol/L Tris-HCl中, pH 8.5)离心洗涤,加入10 mmol/L二硫苏糖醇(DTT)混匀后置于37  $^{\circ}$ C恒温箱中孵育4 h。12 000 r/min离心15 min除废液,加入50 mmol/L碘乙酰胺(IAA)混匀后室温避光孵育0.5 h。离心后再加入10 mmol/L二硫苏糖醇(DTT)室温避光放置15 min,12 000 r/min离心。分别加入尿素溶液和50 mmol/L碳酸氢铵(ABC)洗涤,重复3次。更换超滤管套管,加入50 mmol/L ABC和20  $\mu$ g胰蛋白酶进行蛋白酶切,37  $^{\circ}$ C恒温箱中孵育16 h。反应结束后离心收集肽段,依次加入水、50 mmol/L ABC洗涤超滤管,离心收集洗脱液。使用SpeedVac真空干燥离心机(Eppendorf)完全干燥肽段,回收肽段储存在-80  $^{\circ}$ C,用于糖肽富集。

利用已开发的HILIC方法<sup>[20]</sup>富集完整糖肽。首先,加入质量分数0.1%三氟乙酸(TFA)溶液活化Venusil HILIC(5  $\mu$ m, 10 nm)填料,重复3次。加入体积分数80%乙腈/体积分数0.2% TFA溶液平衡填料,重复3次。回收肽段用体积分数80%乙腈/0.2% TFA溶液溶解,加入已活化平衡的HILIC填料,置于室温旋转混合仪上孵育2 h。将混合物转移到装有C8膜的Tip柱中,离心后再吸尿管底离心液加入Tip柱中离心。

加入体积分数 80% 乙腈/体积分数 0.2% TFA 溶液洗涤除去未结合肽段. 更换离心管, 加入体积分数 0.1% TFA 溶液洗脱收集糖肽, 重复操作 3 次. 使用 SpeedVac 真空干燥离心机(Eppendorf)烘干糖肽, 储存在  $-80^{\circ}\text{C}$  用于 LC-MS/MS 分析.

### 1.3 LC-MS/MS 质谱检测分析

糖蛋白质组学分析, 使用液质联用(LC-MS/MS)仪器, 包括纳米电喷雾离子源的 EASY-nLC 1200 液相色谱系统和 Orbitrap Fusion Lumos Tribrid 质谱仪(赛默飞世尔科技). 加入 A 液溶解糖肽(A 液: 体积分数 0.1% 甲酸; B 液: 体积分数 80% 乙腈/体积分数 0.1% 甲酸), 使用 NanoDrop OneC 分光光度计(赛默飞世尔科技)在 205 nm 的吸光度下测定糖肽浓度, 如前所述<sup>[21]</sup>.

取  $1\ \mu\text{g}$  的糖肽加载到直径  $100\ \mu\text{m}$ , 长度 30 cm 色谱柱中( $1.9\ \mu\text{m}$  ReproSil-Pur C18-AQ; Dr. Maisch). 采用 90 min 梯度洗脱糖肽, 流速为 600 nL/min. 梯度组成: 0~8 min, 5%~10% B; 8~68 min, 10%~22% B; 68~82 min, 22%~32% B; 82~88 min, 32%~90% B 和 88~90 min, 90% B. 质谱设置参数如下: Orbitrap Fusion Lumos Tribrid 质谱仪设置采用 OT-OT 模式; 一级质谱扫描的 AGC 目标值为 400 000; 扫描的质量范围 8 002 000 m/z; 分辨率为 120 000; 离子最大注入时间为 100 ms. 二级质谱扫描 AGC 目标值为 50 000; 分辨率为 15 000, 离子最大注入时间为 250 ms, 使用高能碰撞解离(HCD)对 26 电荷的母离子进行碎裂, 阶梯碰撞能量为 20%, 30% 和 40%; 动态排除时间为 15 s, 质量数据由 XCalibur 进行实时采集.

### 1.4 质谱数据分析

使用人类 Swiss-Prot 蛋白库(版本 20220329, 20 376 个序列)通过 pGlyco<sup>[22]</sup> 和 PANDA 软件分析完整糖肽原始质谱文件. 修饰包括 Carbamidomethylation(C)、Acetyl(Protein N-term)、Deamidation(N)和 Oxidation(M). 通过将“N”改为“J”来修饰 N-糖基化序列. 母离子质量偏差(mass tolerance)为  $\pm 0.0004\%$ , 子离子质量偏差(peptide tolerance)为  $\pm 0.0020\%$ , 允许 2 个漏切位点. 完整糖肽鉴定的质量控制方法设置为 FDR 小于 1%, 其他设置参数选择默认值.

### 1.5 统计学方法

数据分析使用 Perseus<sup>[23]</sup> 完成. N-糖肽的显著性阈值设置为  $P < 0.05$ , 两者的倍数变化  $\geq |1.2|$ . 采用 Benjamini-Hochberg(BH)程序进行多次测试校正. 筛选出的差异基因, 使用 ClusterProfiler 4.0<sup>[24]</sup> 做 GO (gene ontology) 和 KEGG(kyoto encyclopedia of genes and genomes)分析. 使用 String 进行蛋白互作分析(PPI)<sup>[25]</sup>, PPI 网络使用 Cytoscape 优化, 关键基因使用 cytoHubba 插件筛选.

## 2 结果

### 2.1 蛋白质糖基化修饰组质谱数据采集与数据解析

本研究运用超高分辨质谱技术, 针对 5 例结直肠癌患者和 5 例健康人的血浆样本, 进行了糖基化修饰组的系统性质谱数据采集, 研究技术路线如图 1 所示.

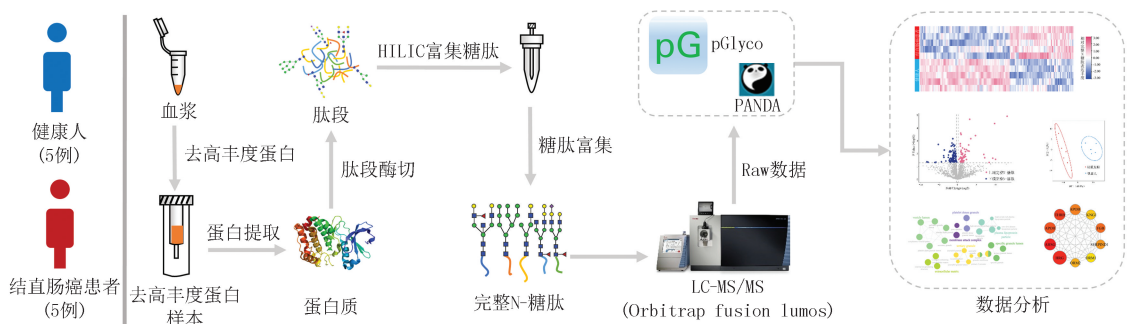


图1 基于质谱的结直肠癌血浆糖蛋白质组学技术路线

Fig.1 Technical route of plasma glycoproteomics for colorectal cancer based on mass spectrometry

首先, 借助 Pierce<sup>TM</sup> TOP 14 试剂盒, 有效去除了血浆中 14 种高丰度蛋白质, 减少可能对后续分析造成

干扰的因素.然后,采用 FASP 肽段酶切方案,成功获取了肽段序列.随后,利用 HILIC 糖肽富集方案,完成糖基化修饰肽段富集,有效提高了糖基化修饰肽段的检测灵敏度.最后,基于 Orbitrap Fusion Lumos 质谱仪器,完成了肽段信息的质谱数据采集,并通过 pGlyco 和 Panda 软件完成了质谱数据的定性与定量分析.通过数据解析,本研究在 10 例样本中共鉴定出 2 175 条完整 N-糖肽,平均鉴定数为 1 959 条,涉及 161 个糖蛋白质(附录图 S1(a))、475 种 N-聚糖和 245 个 N-糖基化修饰位点(附录图 S1(b)).

## 2.2 蛋白质糖基化修饰组的细胞定位与功能

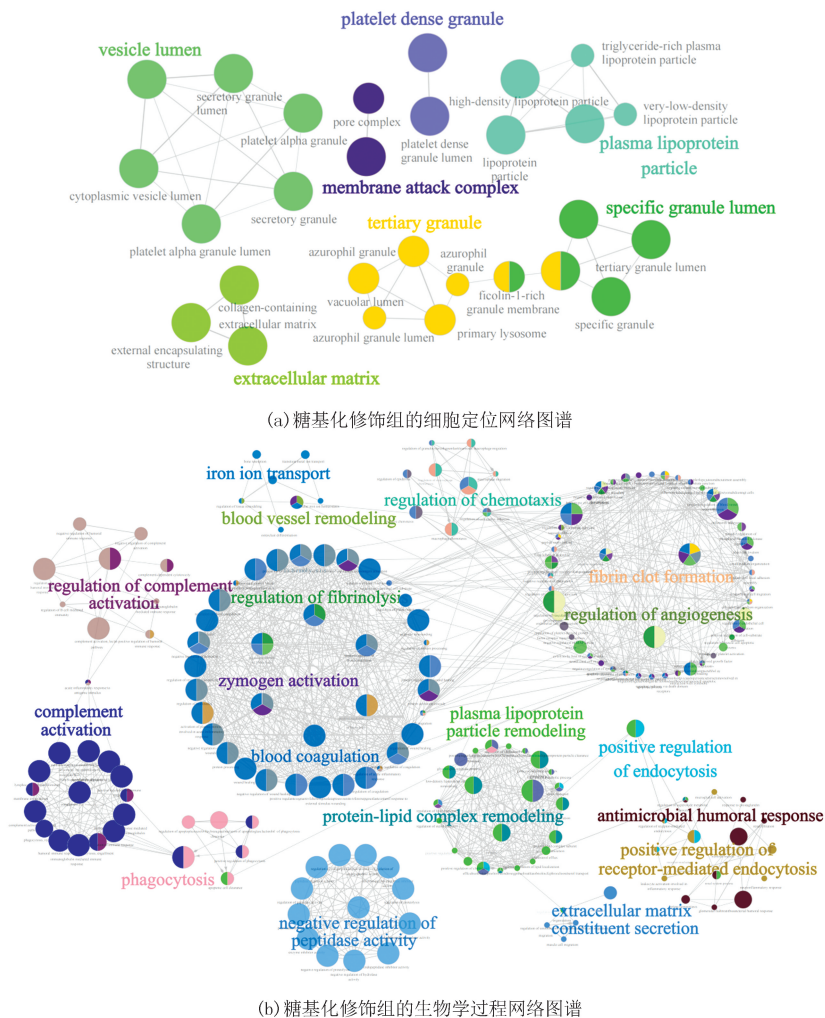
通过细胞定位的富集分析,本研究鉴定到具有糖基化修饰的血浆蛋白质主要聚焦在 28 个细胞组成模块内.这些模块进一步归纳为 7 个网络模块,它们主要定位于囊泡腔、膜攻击复合物、血小板致密颗粒、血浆脂蛋白颗粒、细胞外基质、三级颗粒、特定颗粒腔等细胞结构,如图 2(a)所示.生物学过程的富集分析结果表明,所鉴定的血浆糖蛋白质主要集中于 209 个信号通路中,并被整合为 18 个相似的网络模块.这些模块主要与补体激活、血管重塑、趋化反应、纤维蛋白凝块生成、纤维蛋白溶解、血管新生、酶原的激活、血液凝固、吞噬机制、肽酶活性、血浆脂蛋白颗粒的重塑以及细胞外基质成分的分泌等生物学过程有关,如图 2(b)所示.

## 2.3 结直肠癌血浆的差异表达糖基化肽段

为了系统地筛查结直肠癌血浆中差异表达的糖基化肽段,对结直肠癌患者与健康人的血浆蛋白质糖基化修饰进行了对比分析,并采用 *t*-检验作为主要评估工具.与健康人群相比,结直肠癌患者的血浆中鉴定出 152 条差异表达的完整 N-糖肽,其中 63 条呈现上调表达趋势和 88 条呈现下调表达趋势,如图 3(a)和 3(c)所示.利用无监督的主成分分析(PCA),第一主成分表征了 60.3%的完整 N-糖肽变量,第二主成分表征了 9.2%的完整 N-糖肽变量,如图 3(b)所示.基于差异表达的完整 N-糖肽,本研究在第一主成分和第二主成分上能够有效区分结直肠癌患者与健康人群.结果表明,血浆中的糖基化修饰水平变化可以作为结直肠癌早期诊断的有效指标.

## 2.4 差异 N-糖肽的功能注释

利用 ClusterProfiler 软件,针对差异表达的完整 N-糖肽展开功能富集分析,其中主要参照基因本体(GO)和 KEGG 的信号通路.分析结果表明,差异表达的完整 N-糖肽在细胞结构上,显著富集在细胞外基质(collagen-containing extracellular matrix)、囊泡腔(vesicle lumen)、分泌颗粒腔(secretory granule lumen)、



注:不同颜色表示不同网络模块,各网络模块最显著的信号通路用加粗颜色字体标注.

图2 蛋白质糖基化修饰组的细胞定位与功能注释

Fig.2 Cellular localization and functional annotation of glycosylated modified proteins

细胞质囊泡腔(cytoplasmic vesicle lumen)和内质网腔(endoplasmic reticulum lumen)等结构上.差异表达 N-糖肽主要参与了硫化物结合(sulfur compound binding)、糖胺聚糖结合(glycosaminoglycan binding)、肝素结合(heparin binding)、酶抑制剂活性(enzyme inhibitor activity)和内肽酶抑制剂活性(endopeptidase inhibitor activity)等分子功能.此外,差异表达的完整 N-糖肽还参与了体液免疫反应(humoral immune response)、凝血的负调节(negative regulation of coagulation)、内肽酶活性的负调节(negative regulation of endopeptidase activity)、纤溶(fibrinolysis)和补体激活(complement activation)等生物学过程(图 4(a)).在 KEGG 信号通路中,差异表达的完整 N-糖肽主要参与了补体和凝血级联(complement and coagulation cascades)、肌动蛋白细胞骨架的调节(regulation of actin cytoskeleton)、神经活性配体-受体相互作用(neuroactive ligand-receptor interaction)、胆固醇代谢(cholesterol metabolism)和 ECM 受体相互作用(ECM-receptor interaction)等信号通路(图 4(b)).其中,胆固醇代谢和 ECM 受体等信号通路已明确报道与肿瘤的发生发展具有密切关系<sup>[26]</sup>.

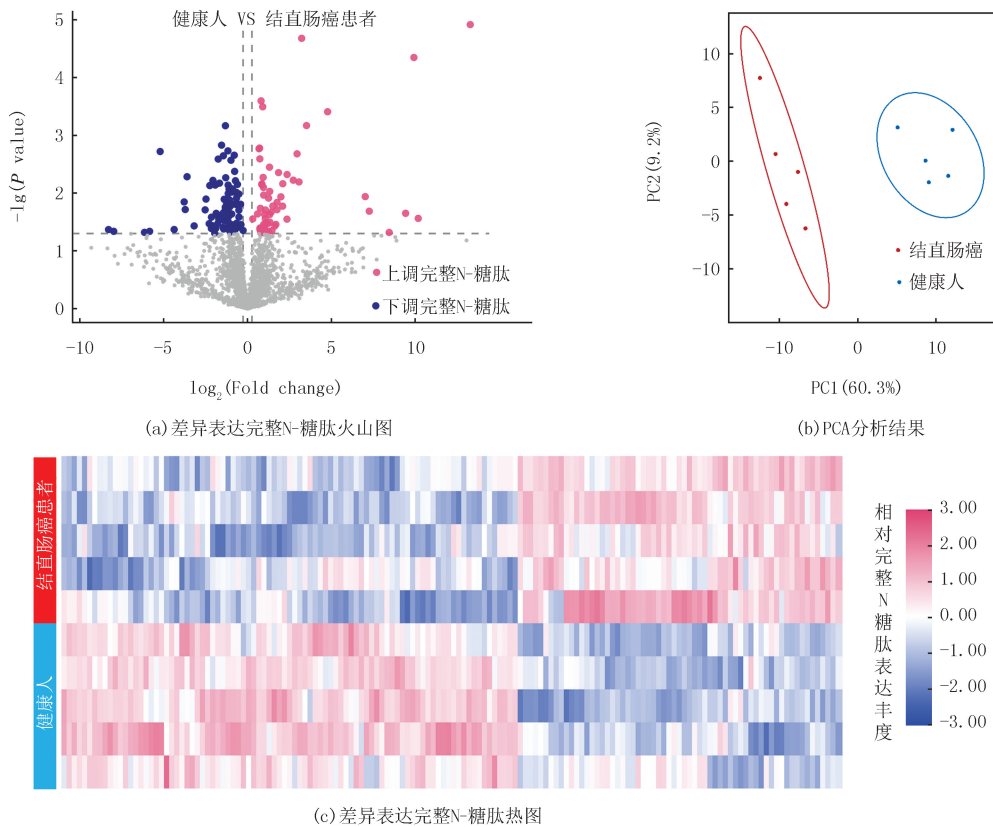


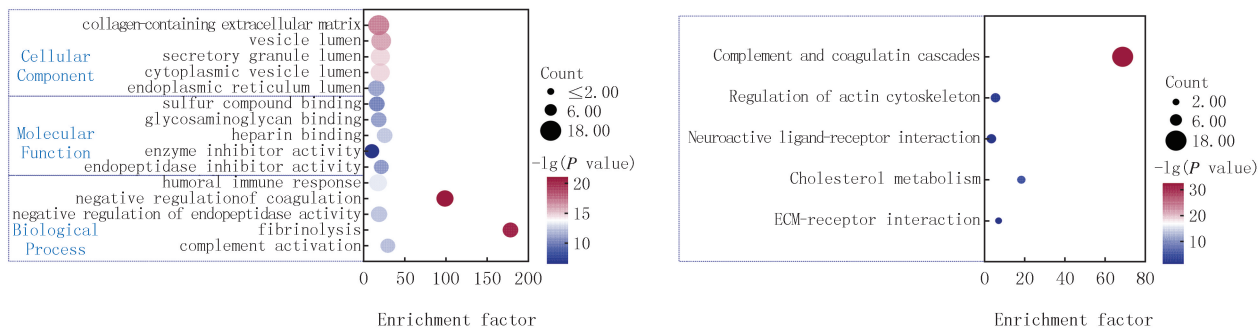
图3 结直肠癌的血浆差异表达完整N-糖肽

Fig.3 Differentially expressed intact N-glycopeptides in the plasma of colorectal cancer

## 2.5 蛋白质相互作用网络分析及候选标志物筛选

采用蛋白质相互作用网络(PPI)分析方法,进一步对差异表达完整 N-糖肽的蛋白相关性进行深入分析,拟获得具有代表性的糖基化修饰候选生物标志物.基于 Cytoscape 软件,采用 cytoHubba 插件对蛋白质相互作用网络(PPI)进行了系统性分析,筛查获得了 10 个具有代表性的候选生物标志物,主要包括 KNG1、F2、FGB、HRG、APOH、APOB、ORM2、AHSG、ORM1 和 SERPIND1 等糖蛋白质(图 5(a,b)).相对于健康人群而言,结直肠癌患者 FGB 蛋白质的完整 N-糖肽 GTAGNALMDGASQLMGGEN394R\_HexNAc5Hex4, AHSG 蛋白质的完整 N-糖肽 KVCQDCPLLAPLN156DTR\_HexNAc3Hex7Fuc2, ORM1 蛋白质的完整 N-糖肽 NEEYN56K\_HexNAc6Hex5NeuAc2Fuc3 等表达量显著上调;KNG1 蛋白质的完整 N-糖肽 LNAENN294ATFYFK\_HexNAc6Hex10, ORM2 蛋白质的完整 N-糖肽 QN88QCFYJSSYLVNQR\_HexNAc6Hex5NeuAc3, APOH 蛋白质的 N-糖肽 LGN253WSAMPSCK\_HexNAc7Hex4NeuAc1 等表达量显著下调,如图 5(c)所示.此外,研究系统分析了 FGB 蛋白质的完整 N-糖肽 GTAGNALMDGASQLMGGEN394R\_Hex

NAc5Hex4 的质谱图,并进行标注,证明了质谱分析结果的可靠性,如附录图 S2 所示。

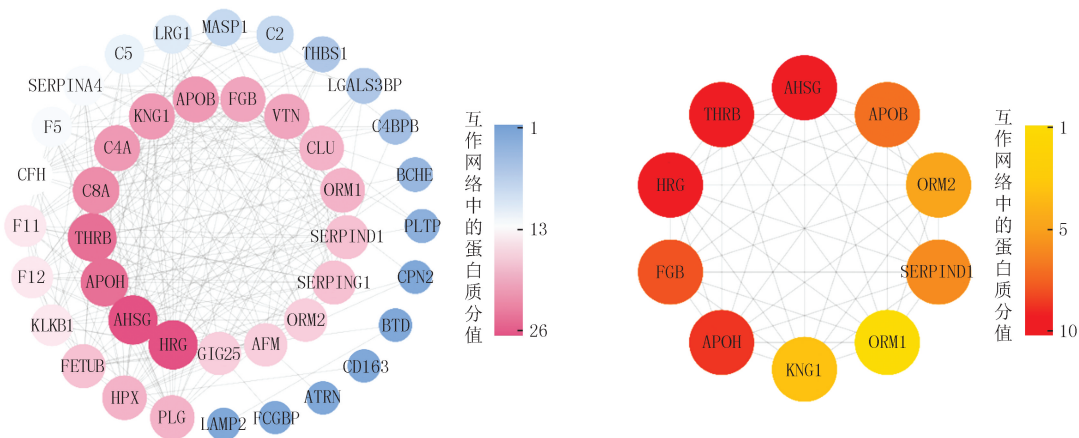


(a) 差异表达N-糖肽的GO分析

(b) 差异表达N-糖肽的KEGG分析

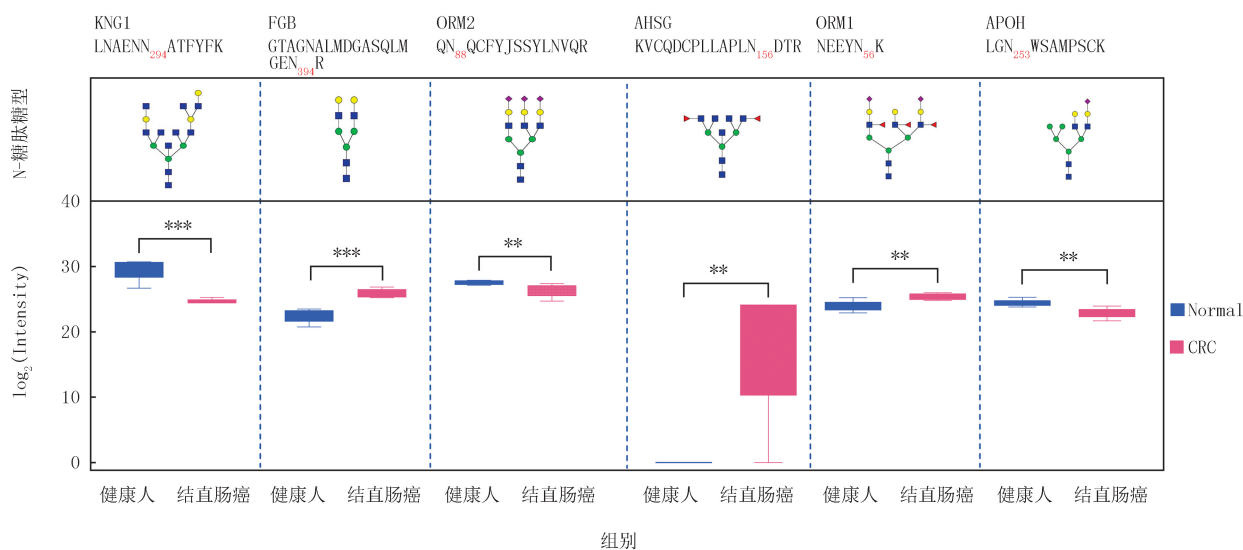
图4 差异表达N-糖肽的功能注释分析结果

Fig. 4 Results of functional annotation analysis of differentially expressed N-glycopeptides



(a) 差异表达糖蛋白质相互作用网络图谱

(b) 核心蛋白质网络图谱



(c) 糖基化修饰候选生物标志物

注:图中\*\*表示 $P < 0.01$ , \*\*\*表示 $P < 0.001$ 。

图5 蛋白质相互作用网络分析及糖基化候选生物标志物筛查

Fig. 5 Protein interaction network analysis and glycosylation candidate biomarker screening

### 3 讨论

目前,在结直肠癌的早期诊断中,CEA 是应用最广泛的血浆生物标志物.此外,还有 CA19-9、CA242、CA50、CA74-2 和 TIMP-1 等生物标志物,但其敏感性和特异性表现都不佳<sup>[9]</sup>.此外,现行的临床筛查手段包括粪便隐血检测、结肠镜、胶囊内窥镜及 CT 扫描等<sup>[6-7]</sup>,但由于某些筛查方法可能带来的风险或不适,寻找无创或更为敏感和特异的生物标志物显得尤为重要.本研究利用非标记的糖基化修饰蛋白质组分析技术成功鉴定了 152 条与结直肠癌有关的差异表达 N-糖肽,涉及 47 种糖蛋白.其中 63 条 N-糖肽呈现上调,而 88 条呈现下调.基于以上发现,筛选出了 10 种可能作为结直肠癌血浆生物标志物的候选蛋白质多肽,包括 KNG1、F2、FGB、HRG、APOH、APOB、ORM2、AHSG、ORM1 和 SERPIND1 等. APOB 和 HRG 蛋白质及其糖基化修饰的差异表达已被文献报道与结直肠癌发生发展有关. APOB 主要负责脂质的运输,研究发现处于不同阶段的结直肠癌患者,其血浆中的 APOB 蛋白表达量均有显著增高,且其糖基化修饰异常与结直肠癌的异常增生相关<sup>[27]</sup>, APOB 可能是结直肠癌术后预后评估的有效生物标志物<sup>[28]</sup>. 另外,富含组氨酸的糖蛋白(histidine-rich glycoprotein, HRG)已被确认为一种有效抑制肿瘤血管生成的物质,它可以显著减少肿瘤细胞的迁移、血管新生和生长<sup>[29]</sup>. 研究显示,结直肠癌或腺癌患者的糖基化修饰存在异常变化,其中唾液酸和岩藻糖基化水平显著上升,这暗示血液 HRG 蛋白的 N-糖肽有望成为结直肠癌的潜在生物标志物<sup>[30]</sup>.

此外, AHSG、APOH、FGB、KNG1、ORM1、ORM2、THRB 蛋白质差异表达被发现与结直肠癌有关.  $\alpha$ -2-HS-糖蛋白(alpha-2-HS-glycoprotein, AHSG)已被报道参与人类的大脑发育、骨代谢调节<sup>[31]</sup>、胰岛素抑制<sup>[32]</sup>、结直肠癌的迁移和侵袭<sup>[33]</sup>等许多正常和病理过程.通过对健康人和结直肠癌患者的血清蛋白质组学分析, AHSG 可能参与结直肠癌的发病机制,并可作为潜在的血清学诊断的生物标志物<sup>[34]</sup>.同时, CHOI 等<sup>[35]</sup>使用同种分析方法发现 AHSG 具有预测腺瘤向癌进展的诊断潜力. MA 等<sup>[36]</sup>分析了血清样本的蛋白质和代谢物谱,发现 APOH 可作为生物标志物,并验证了它对结直肠癌诊断的潜力.纤维蛋白原由三对多肽链组成,分别命名为  $\alpha$ 、 $\beta$  和  $\gamma$ ,这些链分别由 FGA、FGB(fibrinogen beta chain)和 FGG 编码.纤维蛋白原主要参与凝血、炎症和血管生成.研究发现 FGB 在结直肠癌血浆中的蛋白表达量显著上调,提示 FGB 可作为结直肠癌患者早期诊断和了解肿瘤发生的血浆潜在生物标志物<sup>[37]</sup>.另外, YANG 等<sup>[38]</sup>发现 FGB 作为肝转移性结直肠癌的诊断和治疗生物标志物起着关键作用.激肽原 1(kininogen-1)在凝血过程中发挥重要作用,可抑制内皮细胞增殖和血管生成<sup>[39]</sup>.已有研究发现 KNG1 在早期结直肠癌患者血清中的异常表达,表明该蛋白可能是早期检测的潜在生物标志物<sup>[40]</sup>. ORM1(alpha-1-acid glycoprotein)和 ORM2(alpha-2-acid glycoprotein)都属于急性期蛋白,炎症反应时含量增加. KASAHARA 等<sup>[41]</sup>发现 ORM1 对结直肠癌患者总生存期(OS)有显著贡献,提示 ORM1 可能是与 OS 密切相关的因素之一. GAO 等<sup>[42]</sup>采用酶联免疫吸附法测定血浆 ORM2 水平,多因素分析表明,血浆 ORM2 是 II 期结直肠癌患者总体和肿瘤特异性生存的独立预后因素( $P < 0.05$ ).在此之前, ZHANG 等<sup>[43]</sup>通过定量蛋白质组学分析,证明了 ORM2 可作为结直肠癌诊断的潜在生物标志物.凝血酶原(THRB)是凝血机制的重要组成部分.已明确报道凝血酶原在结肠腺癌的进展中发挥着广泛的作用, THRB 既促进原发肿瘤的生长,也促进循环肿瘤细胞形成远处转移的潜力<sup>[44]</sup>.

综上,本研究系统地鉴定了蛋白质糖基化修饰在结直肠癌患者血液中的差异表达模式,与早期关于结直肠癌的研究结果高度吻合,进一步验证了血浆蛋白质糖基化修饰组学方法的稳定性和准确性.除了之前报道的生物标志物外,我们还鉴定了 SERPIND1 作为新的结直肠癌潜在生物标志物. SERPIND1 是一种糖蛋白及蛋白酶抑制剂,具有抑制凝血酶的活性,并与肝素、硫酸皮肤素及其他内源性糖胺聚糖发生相互作用<sup>[45]</sup>.鉴于本研究的样本量有限,为进一步确定 AHSG、APOH、FGB、KNG1、ORM1、ORM2、THRB 和 SERPIND1 与结直肠癌的相关性,未来将在更大的样本队列中进行验证研究.同时,为确保所筛选生物标志物的精确性和稳健性,后续研究还需进行相关生物学功能验证实验,进一步揭示其在结直肠癌发展过程中的分子角色和生物学意义.

## 参 考 文 献

- [1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020; GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA*, 2021, 71(3): 209-249.
- [2] ARNOLD M, SIERRA M S, LAVERSANNE M, et al. Global patterns and trends in colorectal cancer incidence and mortality[J]. *Gut*, 2017, 66(4): 683-691.
- [3] KUIPERS E J, GRADY W M, LIEBERMAN D, et al. Colorectal cancer[J]. *Nature Reviews Disease Primers*, 2015, 1: 15065.
- [4] MUTO T, BUSSEY H J, MORSON B C. The evolution of cancer of the colon and rectum[J]. *Cancer*, 1975, 36(6): 2251-2270.
- [5] YANG X, WEI W G, TAN S S, et al. Identification and verification of HCAR3 and INSL5 as new potential therapeutic targets of colorectal cancer[J]. *World Journal of Surgical Oncology*, 2021, 19(1): 248.
- [6] PICKHARDT P J, HASSAN C, HALLIGAN S, et al. Colorectal cancer: CT colonography and colonoscopy for detection; systematic review and meta-analysis[J]. *Radiology*, 2011, 259(2): 393-405.
- [7] SPADA C, HASSAN C, GALMICHE J P, et al. Colon capsule endoscopy; European society of gastrointestinal endoscopy (ESGE) guideline[J]. *Endoscopy*, 2012, 44(5): 527-536.
- [8] HEWITSON P, GLASZIOU P, WATSON E, et al. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemocult): an update[J]. *The American Journal of Gastroenterology*, 2008, 103(6): 1541-1549.
- [9] VUKOBRAT-BIJEDIC Z, HUSIC-SELIMOVIC A, SOFIC A, et al. Cancer antigens (CEA and CA 19-9) as markers of advanced stage of colorectal carcinoma[J]. *Medical Archives*, 2013, 67(6): 397-401.
- [10] JENKINSON F, STEELE R J C. Colorectal cancer screening-methodology[J]. *The Surgeon*, 2010, 8(3): 164-171.
- [11] LANGAN R C, MULLINAX J E, RAIJI M T, et al. Colorectal cancer biomarkers and the potential role of cancer stem cells[J]. *Journal of Cancer*, 2013, 4(3): 241-250.
- [12] WU L, QU X G. Cancer biomarker detection; recent achievements and challenges[J]. *Chemical Society Reviews*, 2015, 44(10): 2963-2997.
- [13] PINHO S S, REIS C A. Glycosylation in cancer; mechanisms and clinical implications[J]. *Nature Reviews Cancer*, 2015, 15(9): 540-555.
- [14] OHTSUBO K, MARTH J D. Glycosylation in cellular mechanisms of health and disease[J]. *Cell*, 2006, 126(5): 855-867.
- [15] DOHERTY M, THEODORATOU E, WALSH I, et al. Plasma N-glycans in colorectal cancer risk[J]. *Scientific Reports*, 2018, 8(1): 8655.
- [16] PARK J J, LEE M. Increasing the  $\alpha$ 2,6 sialylation of glycoproteins may contribute to metastatic spread and therapeutic resistance in colorectal cancer[J]. *Gut and Liver*, 2013, 7(6): 629-641.
- [17] PLEBANI M, BANFI G, BERNARDINI S, et al. Serum or plasma? An old question looking for new answers[J]. *Clinical Chemistry and Laboratory Medicine*, 2020, 58(2): 178-187.
- [18] ZHAO Y, WANG M, MENG B, et al. Identification of dysregulated complement activation pathways driven by N-glycosylation alterations in T2D patients[J]. *Frontiers in Chemistry*, 2021, 9: 677621.
- [19] WIŚNIEWSKI J R, ZOUGMAN A, NAGARAJ N, et al. Universal sample preparation method for proteome analysis[J]. *Nature Methods*, 2009, 6(5): 359-362.
- [20] ZHANG Y, ZHAO W J, ZHAO Y, et al. Comparative glycoproteomic profiling of human body fluid between healthy controls and patients with papillary thyroid carcinoma[J]. *Journal of Proteome Research*, 2020, 19(7): 2539-2552.
- [21] JIANG Y, SUN A H, ZHAO Y, et al. Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma[J]. *Nature*, 2019, 567(7747): 257-261.
- [22] LIU M Q, ZENG W F, FANG P, et al. PGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification[J]. *Nature Communications*, 2017, 8(1): 438.
- [23] TYANOVA S, TEMU T, SINITYCYN P, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data[J]. *Nature Methods*, 2016, 13(9): 731-740.
- [24] WU T Z, HU E Q, XU S B, et al. ClusterProfiler 4.0: a universal enrichment tool for interpreting omics data[J]. *The Innovation*, 2021, 2(3): 100141.
- [25] SZKLARCZYK D, FRANCESCHINI A, WYDER S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life[J]. *Nucleic Acids Research*, 2015, 43(Database issue): D447-D452.
- [26] BAO Y L, WANG L, SHI L, et al. Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer[J]. *Cellular & Molecular Biology Letters*, 2019, 24: 38.
- [27] KOPYLOV A T, STEPANOV A A, MALSAGOVA K A, et al. Revelation of proteomic indicators for colorectal cancer in initial stages of development[J]. *Molecules*, 2020, 25(3): 619.
- [28] CHEN X Q, WU P W, LIU D H, et al. Prognostic significance of high triglyceride and apolipoprotein B levels in patients with stage III and high-risk stage II colorectal cancer undergoing curative surgery[J]. *Oncology Letters*, 2020, 20(1): 705-714.
- [29] JONES A L, HULETT M D, PARISH C R. Histidine-rich glycoprotein; a novel adaptor protein in plasma that modulates the immune,



- vascular and coagulation systems[J]. *Immunology and Cell Biology*, 2005, 83(2): 106-118.
- [30] QIU Y H, PATWA T H, XU L, et al. Plasma glycoprotein profiling for colorectal cancer biomarker identification by lectin glycoarray and lectin blot[J]. *Journal of Proteome Research*, 2008, 7(4): 1693-1703.
- [31] REYNOLDS J L, SKEPPER J N, MCNAIR R, et al. Multifunctional roles for serum protein fetuin-a in inhibition of human vascular smooth muscle cell calcification[J]. *Journal of the American Society of Nephrology*, 2005, 16(10): 2920-2930.
- [32] SRINIVAS P R, WAGNER A S, REDDY L V, et al. Serum alpha 2-HS-glycoprotein is an inhibitor of the human insulin receptor at the tyrosine kinase level[J]. *Molecular Endocrinology*, 1993, 7(11): 1445-1455.
- [33] MORI K, EMOTO M, INABA M. Fetuin-A: a multifunctional protein[J]. *Recent Patents on Endocrine, Metabolic & Immune Drug Discovery*, 2011, 5(2): 124-146.
- [34] FAN N J, KANG R, GE X Y, et al. Identification alpha-2-HS-glycoprotein precursor and tubulin beta chain as serology diagnosis biomarker of colorectal cancer[J]. *Diagnostic Pathology*, 2014, 9: 53.
- [35] CHOI J W, LIU H, SHIN D H, et al. Proteomic and cytokine plasma biomarkers for predicting progression from colorectal adenoma to carcinoma in human patients[J]. *Proteomics*, 2013, 13(15): 2361-2374.
- [36] MA Y L, ZHANG P, WANG F, et al. An integrated proteomics and metabolomics approach for defining oncofetal biomarkers in the colorectal cancer[J]. *Annals of Surgery*, 2012, 255(4): 720-730.
- [37] SUN Z J, JI S R, WU J L, et al. Proteomics-based identification of candidate exosomal glycoprotein biomarkers and their value for diagnosing colorectal cancer[J]. *Frontiers in Oncology*, 2021, 11: 725211.
- [38] YANG W, SHI J, ZHOU Y, et al. Co-expression network analysis identified key proteins in association with hepatic metastatic colorectal cancer[J]. *Proteomics Clinical Applications*, 2019, 13(6): e1900017.
- [39] MOTTA G, ROJKJAER R, HASAN A A K, et al. High molecular weight kiningogen regulates prekallikrein assembly and activation on endothelial cells; a novel mechanism for contact activation[J]. *Blood*, 1998, 91(2): 516-528.
- [40] QUESADA-CALVO F, MASSOT C, BERTRAND V, et al. OLFM4, KNG1 and Sec24C identified by proteomics and immunohistochemistry as potential markers of early colorectal cancer stages[J]. *Clinical Proteomics*, 2017, 14: 9.
- [41] KASAHARA K, NARUMI R, NAGAYAMA S, et al. A large-scale targeted proteomics of plasma extracellular vesicles shows utility for prognosis prediction subtyping in colorectal cancer[J]. *Cancer Medicine*, 2023, 12(6): 7616-7626.
- [42] GAO F, ZHANG X H, WHANG S, et al. Prognostic impact of plasma ORM2 levels in patients with stage II colorectal cancer[J]. *Annals of Clinical and Laboratory Science*, 2014, 44(4): 388-393.
- [43] ZHANG X H, XIAO Z Y, LIU X Y, et al. The potential role of ORM2 in the development of colorectal cancer[J]. *PLoS One*, 2012, 7(2): e31868.
- [44] BEITIA M, ROMANO P, LARRINAGA G, et al. The activation of prothrombin seems to play an earlier role than the complement system in the progression of colorectal cancer: a mass spectrometry evaluation[J]. *Diagnostics*, 2020, 10(12): 1077.
- [45] BARTSCH R, WENZEL C, PLUSCHNIG U, et al. Prognostic value of monitoring tumour markers CA 15-3 and CEA during fulvestrant treatment[J]. *BMC Cancer*, 2006, 6: 81.

## Plasma glycoproteomic analysis of colorectal cancer

Zhao Yang<sup>1</sup>, Zeng Jiaming<sup>1,2</sup>, Zhao Jiawei<sup>3</sup>, Meng Bo<sup>2</sup>

(1. Technology Innovation Center of Mass Spectrometry for State Market Regulation, Center for Advanced Measurement Science, National Institute of Metrology, Beijing 100029, China; 2. College of Chemical Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China; 3. College of Life Sciences, China Jiliang University, Hangzhou 310018, China)

**Abstract:** Colorectal cancer is one of the most common malignant tumors worldwide, and its mortality risk is closely related to the stage of early diagnosis. To improve patient prognosis and reduce mortality, identifying effective biomarkers for early diagnosis of colorectal cancer is crucial. In this study, label-free quantitative glycoproteomics technology was applied to conduct a comprehensive analysis of the glycosylation patterns in plasma samples from colorectal cancer patients and healthy individuals. As a result, 152 differential integrated N-glycopeptides with upregulated or downregulated expression specific to colorectal cancer patients were identified, effectively distinguishing between healthy individuals and patients. Further analysis revealed that the specific expression of integrated N-glycopeptides from proteins like KNG1, THRB, FGB, APOH, ORM2, AHSG, ORM1, and SERPIND1 significantly correlates with the progression of colorectal cancer, offering substantial data support for early disease diagnosis and screening.

**Keywords:** colorectal cancer; plasma; proteome; glycosylation modification; LC-MS/MS

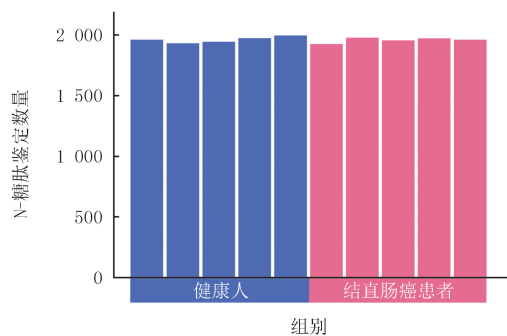
[责任编辑 刘洋 赵晓华]

附录

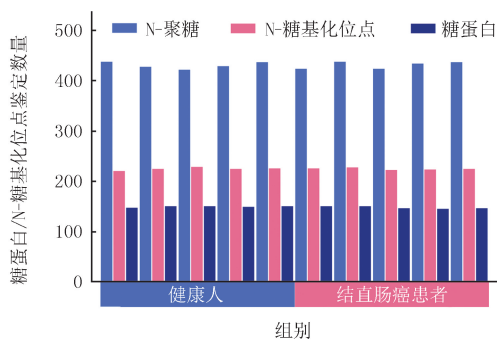
表 S1 实验材料与仪器

Tab. S1 Experimental materials and instruments

材料/仪器名称	货号/批号	生产厂家
三氟乙酸(TFA)	WH328711	Thermo Fisher
尿素(UA)	U5128-1KG	Sigma
二硫苏糖醇(DTT)	SLCF2685	Sigma
碘乙酰胺(IAA)	SLCC6164	Sigma
氯化钙(CaCl <sub>2</sub> )	BCCC3561	Sigma
碳酸氢铵(NH <sub>4</sub> HCO <sub>3</sub> )	BCBV0122	Sigma
30 kDa 超滤管	R7CA92317	Sigma
测序级胰蛋白酶(Trypsin)	P01001	酶知源
HILIC 填料	P22-00336	Agela Technologies
纳升级液相	Easy nLC1200	Thermo Fisher
轨道离子阱高分辨质谱仪	Orbitrap Fusion Lumos	Thermo Fisher
微量紫外分光光度计	NanoDrop OneC	Thermo Fisher
全波长酶标仪	Multiskan sky	Thermo Fisher



(a) 完整N-糖肽鉴定数

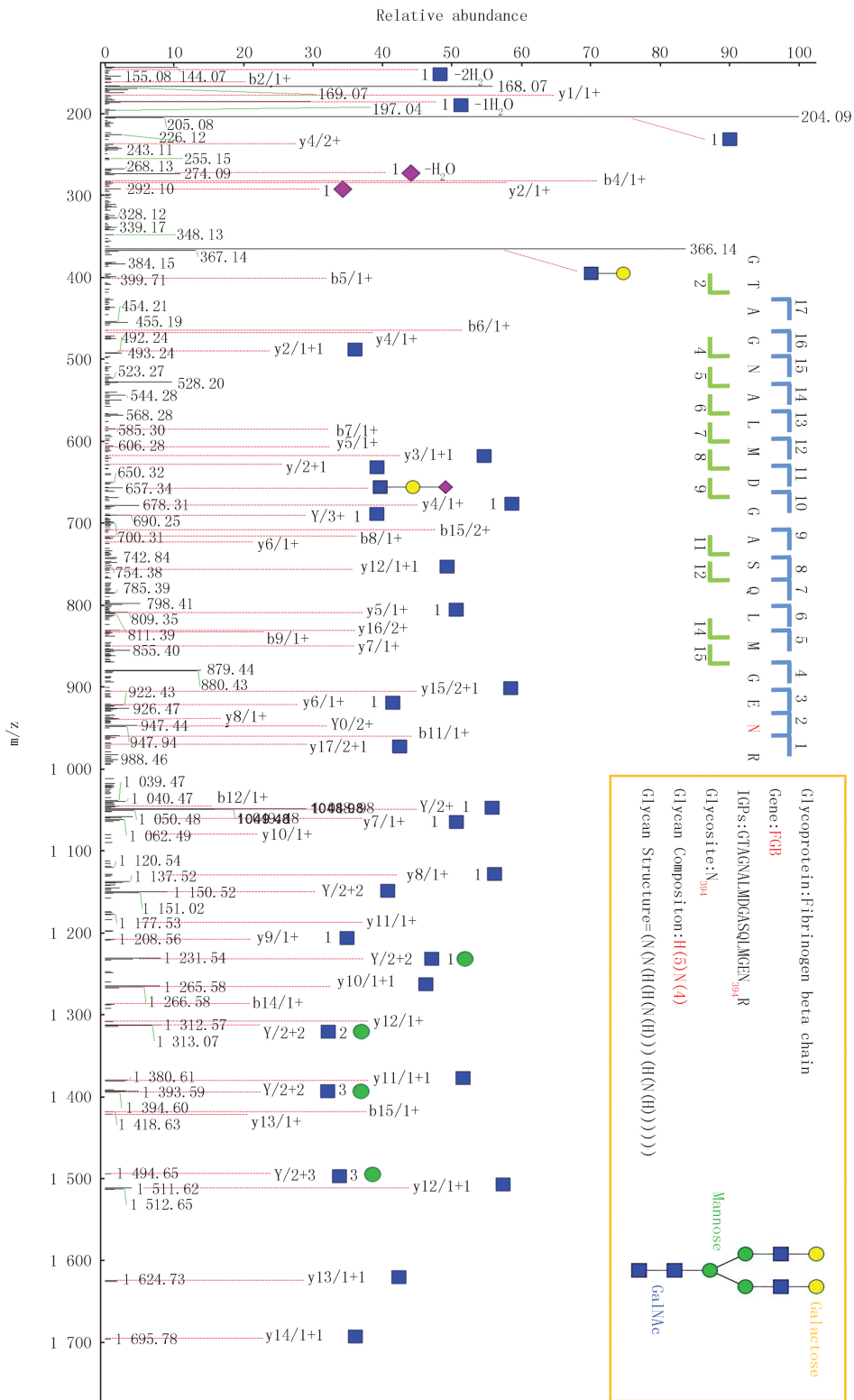


(b) 糖蛋白/N-糖基化位点/N-聚糖鉴定数

图S1 结肠直肠癌血浆糖蛋白质组学技术流程及糖肽鉴定结果

Fig.S1 Workflow for colorectal cancer plasma glycoproteome profiling and results of N-glycopeptide identification

C1890#25168 RT:62.46 AV:1 NL:2.38E5  
 T:FTMS+CSI D Full ms2 880.1193@hcd30.00[42.0000-1771.0000]



图S2 FGB蛋白质的N-糖肽谱图解析

Fig.S2 Analysis of N-glycopeptide profile of FGB protein