

基于液相色谱-质谱的 血清糖蛋白质组定量分析

胡丹阳¹, 蒋碧云², 王 献¹, 刘晓慧²

(1. 中南民族大学化学与材料科学学院, 国家民事委员会分析化学重点实验室, 湖北 武汉 430074;

2. 复旦大学生物医学研究院, 上海 200032)

摘要: 糖尿病作为最常见的代谢疾病之一, 是一种受环境和遗传因素影响的多因素疾病。糖基化可导致蛋白功能多样性, 对重大疾病、免疫系统有着深远影响。本实验通过两性离子亲水相互作用色谱(zwitterionic hydrophilic interaction liquid chromatography, ZIC-HILIC)富集技术结合微升级高效液相色谱-四极杆飞行时间质谱法, 对6例糖尿病人和6例正常人的血清中的蛋白和糖蛋白进行了非标记定量分析, 定量了65个差异蛋白和24个差异表达的糖蛋白。通过DAVID软件对这些差异蛋白和差异糖蛋白进行功能富集分析, 进一步了解它们的细胞定位、分子功能和生物过程。该研究揭示了糖尿病发病过程中糖蛋白的变化, 可为疾病的发病机制和临床诊断提供数据参考。

关键词: 糖尿病; 血清糖蛋白; 非标定量; 质谱

中图分类号: O657.63

文献标志码: A

文章编号: 1004-2997(2019)04-0314-11

doi: 10.7538/zpxb.2018.0133

Quantitative and Functional Analysis of Glycoprotein Based on Liquid Chromatography-Mass Spectrometry

HU Dan-yang¹, JIANG Bi-yun², WANG Xian¹, LIU Xiao-hui²

(1. Key Laboratory of Analytical Chemistry of the State Ethnic Affair Commission, College of Chemistry and Materials Science, South-Central University for Nationalities, Wuhan 430074, China;

2. Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China)

Abstract: Diabetes mellitus as one of the most prevalent metabolic disease is a multifactorial disease which is influenced by environmental and genetic factors. It can cause complications in most of organs such as heart, eye, kidney and nervous system which has resulted in high economic cost and burden. Proteomics is a comprehensive strategy to analyze all proteins in a biological system. High performance liquid chromatography-

收稿日期: 2018-10-24; 修回日期: 2019-02-18

基金项目: 国家自然科学基金面上项目(21275167); 国家重大科研计划精准特异灵敏实用临床定量蛋白质组支撑技术研究(2017YFC0906600); 国家重大科研计划蛋白质修饰组的高维度鉴定及其功能网络(2017YFA0505100)资助

作者简介: 胡丹阳(1993—), 女(汉族), 河南人, 硕士研究生, 分析化学专业。E-mail: 1874023723@qq.com

通信作者: 刘晓慧(1981—), 女(汉族), 湖南人, 高级工程师, 从事蛋白质组学研究。E-mail: liuxiaohui@fudan.edu.cn

王 献(1978—), 女(汉族), 武汉人, 教授, 从事质谱分析研究。E-mail: 1437822646@qq.com

网络出版时间: 2019-05-15; 网络出版地址: <http://kns.cnki.net/kcms/detail/11.2979.TH.20190711.1029.014.html>

mass spectrometry (LC-MS/MS) is powerful in peptides and proteins detection. Protein post translational modifications such as *N*-Glycosylation is widely implicated as a common modification in numerous blood proteins and impacts the in-vivo protein functions. Abnormal glycosylation is associated with many human diseases and the study of glycosylation has attracted more and more attention. In this study, serum samples were collected from 6 healthy people and 6 patients with diabetic diseases for comprehensive proteome and *N*-glycosylation modification analysis. All the proteins were digested fully with Trypsin and Zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC) was used to enrich glycosylated peptides. The platform equipped with high performance liquid chromatography and time-of-flight mass spectrometry was applied to acquire the whole proteome as well as the glycopeptide data of the serum samples. Furthermore, the raw data were searched and analyzed by Maxquant software. Student's T-test was used to determine the significantly changed proteins and glycoproteins with *P* less than 0.05, and the fold change ratio was calculated based on the average abundance of the proteins in the diabetic serum and normal serum sample. In order to find more differentially expressed proteins for subsequent analysis, the expression level of $\pm 20\%$ was used as the criteria which means a fold change greater than 1.2 or less than 0.8 as the criteria for reference value. A total of 291 proteins and 181 glycoproteins were identified. 65 altered proteins and 24 differentially expressed glycoproteins in the diabetic serum and normal serum were quantified by Label-free quantitation. Altered expressed proteins and differential expressed glycoproteins were analyzed by Gene Ontology enrichment analysis to further understand their cellular components, molecular functions and biological processes, and compared their differences in these aspects. IgA binding and blood coagulation are more enriched in altered glycoproteins other than proteins. The data reveal changes in glycoproteins during the onset diabetes mellitus, providing a detailed data reference for the occurrence mechanism and clinical diagnosis of the diabetes mellitus.

Key words: diabetes; serum glycoprotein; label-free quantitation; mass spectrometry

糖基化是一种具有重要生物学功能的蛋白质翻译后修饰。与许多其他翻译后修饰不同,蛋白质的 *N* 糖基化的主要结构 *N*-聚糖是由许多单糖单元组成,且呈非线性排列,具有结构多样性,研究蛋白质的糖基化对了解其生物学结构和功能有着重要意义,如对糖蛋白和糖基化位点、对糖型进行鉴定等^[1-3]。研究表明^[4-7],糖蛋白参与了很多重要的生物学过程,包括细胞粘附、分子运输、受体激活和特定的蛋白功能等,许多疾病的病理过程与糖蛋白表达量的变化密切相关。据报道^[8-9],美国食品药品监督管理局批准的 25% 肿瘤标志物都是糖蛋白,例如载脂蛋白 A1、甲胎蛋白(α -Fetoprotein)、癌胚抗原(carcinoembryonic antigen)等。

研究蛋白质糖基化具有一定难度。糖蛋白在复杂样品中的含量较低,糖肽的离子化效率有限,其质谱响应信号强度远低于非糖肽,且具有高度的微观不均一性,若糖蛋白、糖肽不经富集,几乎无法用于质谱直接鉴定^[10-12]。因此,为获得较理想的质谱信息,在样品处理过程中,将糖蛋白、糖肽从复杂样品中富集出来是关键步骤。常见的糖蛋白/糖肽富集方法包括凝集素富集法、肼化学富集法^[13]、两性离子亲水相互作用色谱富集法(zwitterionic hydrophilic interaction liquid chromatography, ZIC-HILIC)等。凝集素富集法是通过凝集素选择性地识别某些特殊结构的单糖或聚糖而实现对糖蛋白的富集,虽然该法具有高度的特异性,但凝集素与糖

蛋白的结合力较弱^[14]。胍化学富集法是通过糖链的顺式二羟基被高碘酸钠氧化形成二醛,然后与固定相上的胍官能团反应形成脎进行富集,其特异性好、效率高,但糖链结构会受到破坏^[15]。糖基化修饰中的聚糖基团具有亲水性,亲水相互作用色谱法(hydrophilic interaction liquid chromatography, HILIC)不需要对样品进行衍生,具有较强的保留能力。在此基础上发展的两性离子亲水相互作用色谱(ZIC-HILIC)则是利用电荷相互作用和亲水相互作用富集糖肽,固定相上存在正电荷和负电荷的中心,可以和被分析物表面基团以及流动相中的电解质产生电荷相互作用和电荷-偶极相互作用,从而对糖肽进行富集^[16]。该方法高效、快速、不易破坏糖链。

糖尿病是一种以高血糖为特征的代谢性疾病,目前糖尿病患者需终生服药,饮食方面有诸多禁忌,严重影响生活质量。糖尿病还伴随多种并发症,长期血糖升高会对心脏、血管、眼睛、肾脏和神经系统造成较大损伤。目前,全球已有超过4亿人患有糖尿病^[17],我国的糖尿病患者约9500万,糖尿病前期患者约1.5亿^[18]。研究发现,患者糖蛋白的表达量发生了变化^[19-21],O-糖基化糖蛋白在糖尿病及其并发症的病因及形成中有重要作用^[22],但关于N-糖基化糖蛋白的相关报道较少。

本工作拟以6例正常人与6例糖尿病人的血清为样品,深入研究N-糖蛋白,对血清进行蛋白质组学和糖蛋白质组学的非标记定量(label-free quantification),以系统、全面地揭示糖尿病发病过程中蛋白质和糖基化后修饰的变化,希望为进一步揭示糖尿病的发病机理提供参考。

1 实验部分

1.1 主要仪器与试剂

Eksigent UPLC 超高效液相色谱仪、Triple-TOFTM 5600 高分辨质谱仪:均为美国 AB Sciex 公司产品;Milli-Q 超纯水系统:美国 Millipore 公司产品;C18 除盐柱:美国 Waters 公司产品;离心机、真空浓缩仪器、恒温混匀仪:德国 Eppendorf 公司产品;两性离子亲水相互作用填料(ZIC-HILIC particles)、-80℃超低温冰箱:

美国 Thermo Fisher 公司产品。

糖尿病人血清、正常人血清:由中山医院提供;碳酸氢铵、二硫苏糖醇(DTT)、碘乙酰胺(IAA)、三氟乙酸(TFA):均为美国 Sigma 公司产品;胰蛋白酶(Trypsin):美国 Promega 公司产品;肽N-糖苷酶F(PNGaseF):美国 New England Biolabs 公司产品;乙腈(ACN):德国 Merck 公司产品;甲酸:美国 Fisher Scientific 公司产品。

1.2 血清收集与处理

将采集的血液样品在不加抗凝剂的情况下室温静置1h凝血,随后以1500 r/min 离心10 min,取上清液得到血清,分装并置于-80℃冰箱保存。实验时取出血清,在冰上进行解冻,低温下以14000 r/min 离心30 min,取中层血清于离心管中,并编号。

1.3 蛋白质酶解与 C18 除盐

各取55 μL 每例血清,分别加入终浓度为1 mol/L 尿素和100 mmol/L 碳酸氢铵溶液稀释后,使血清蛋白质充分变性,然后加入终浓度为10 mmol/L 二硫苏糖醇,于56℃恒温反应1h;避光加入终浓度25 mmol/L 碘乙酰胺,于37℃避光振荡30 min;按照胰蛋白酶与蛋白质量比1:50加入胰蛋白酶,于37℃反应过夜;将酶解后的溶液以14000 r/min 离心15 min 去除悬浮微粒,然后加入200 μL 乙腈活化C18柱,用200 μL 0.1%三氟乙酸平衡柱子;将酶解好的蛋白质溶液重复上样3次,并用0.1%三氟乙酸清洗3次;依次用200 μL 30%乙腈+0.1%三氟乙酸、60%乙腈+0.1%三氟乙酸将肽段从C18柱上洗脱下来,将多次洗脱液混合均匀后,各取每个样品40 μL 洗脱液于离心管中,在冻干机中冻干,直接用于蛋白水平质谱鉴定;将剩余的洗脱液置于冻干机中冻干,待糖肽富集实验。

1.4 糖肽富集

将上述剩余肽段用200 μL 80%乙腈+1%三氟乙酸溶液溶解,并用200 μL 80%乙腈+1%三氟乙酸溶液平衡活化ZIC-HILIC柱3次,重复上样3次,用80%乙腈+1%三氟乙酸淋洗6次除去未结合的肽段;然后将保留于ZIC-HILIC柱上的糖肽用0.1%TFA洗脱2次,收集洗脱液,冻干后置于-20℃冰箱中

备用。将富集的糖肽用 100 mmol/L 碳酸氢铵溶解,向其中加入 PNGaseF 酶,并在 37 ℃ 孵育振荡过夜处理;将酶解后的肽段冻干,并用 0.1% 甲酸溶解,待 LC-MS/MS 分析,所有样品处理过程一致。

1.5 实验条件

用于做全蛋白质组分析的肽段以及经富集后的糖肽,均以相同的液相色谱和质谱条件进行分析。采用 Eksigent UPLC 色谱仪与 AB Sciex TripleTOF® 5600 质谱仪联用系统进行 LC-MS/MS 实验。

1.5.1 色谱条件 C18 反相色谱柱(0.3 mm×150 mm×3 μm);流动相:A 为 0.1%FA 水溶液,B 为 0.1%FA 的乙腈溶液;流速 5 μL/min;洗脱程序:0~1 min(5%~7%B),1~94 min(7%~24%B),94~109 min(24%~38%B),109~109.5 min(38%~80%B),109.5~114 min(80%B),114~114.5 min(80%~5%B),114.5~119 min(5%B)。

1.5.2 质谱条件 正离子模式检测;气帘气压强 172.4 kPa;喷雾电压 2 300 V;一级质谱采集的质量扫描范围为 m/z 350~1 250;采集模式为数据依赖采集,并选取丰度最高的 20 个母离子进行二级扫描;动态排除时间为 14 s,排除已检测到的离子;检测的肽段所带电荷数为 2~4。

1.6 数据库检索

质谱数据使用 Maxquant(1.6.0.1 版本)软件进行分析,以 Uniprot_human(20 394 条肽段)为蛋白数据库进行搜库。搜索参数:Trypsin 蛋白酶,糖肽/糖蛋白鉴定漏切上限是 2 个(全蛋白组漏切位点参数设置为 1);可变修饰(variable modification):蛋白 N 端乙酰化(N-terminal acetylation(+42.01 u)),甲硫氨酸上的氧化修饰(oxidation[M](+15.99 u)),天冬酰胺脱酰胺基(deamidated-N(+0.98 u))(仅限于糖蛋白);固定修饰(fixed modification):半胱氨酸的还原烷基化修饰(carbamidomethyl[C],(+57.02 u));MS 的质量误差为 40×10^{-6} ,MS/MS 的质量误差设为 0.1 u,蛋白假阳性率(false discovery rate, FDR)设为 1%;组间匹配(match between runs):匹配时间窗口为 0.7;对齐时间窗口为 20;非标定量参数:LFQ

label-free 定量,设定选择独有肽段(unique peptide)的强度信息进行定量。各组数据依据总强度信息进行强度归一化后,组间通过肽段强度进行定量分析。

2 结果与讨论

2.1 实验方法

作为最广泛的蛋白后修饰,糖基化修饰在疾病的发生和发展过程中受到较多关注。糖蛋白天然丰度较低,非糖基化肽会严重干扰糖肽的信号,因此蛋白富集是糖蛋白组学分析中的重要环节之一。ZIC-HILIC 是一种简单、快速且高通量的糖蛋白富集方法,处理过程不需要衍生,且不破坏糖链结构,因此被广泛应用于糖基化位点、完整糖肽和位点特异性的糖蛋白分析。本研究采用 ZIC-HILIC 对糖肽进行富集和分析,为了更准确地找到糖基化水平发生变化的蛋白质,选用未经富集的肽段进行定量分析,以蛋白质表达水平作为参考,同时找出差异蛋白质和糖基化水平发生变化的差异糖蛋白质信息,作为研究糖尿病发生发展的参考。

常规蛋白质组学研究采用的纳升级液相色谱-质谱法(nano-LC-MS/MS)具有检测灵敏度高、但是也存在较多缺点。首先,nano-LC-MS/MS 的负载容量低,导致生物基质中的浓度敏感度较低;其次,操作稳定性相对较差,需要耗费大量的人力、物力用来维护仪器,且分析通量相对较低,限制了其在大型队列分析中的应用。为了增加分析通量,且保证定量分析的稳定性,本研究选择微升级液相色谱-质谱法(micro-LC-MS/MS),同时采用 label-free 方法进行定量分析,一次分析 12 个样品的蛋白质组差异和糖蛋白质组差异。

血清样品取材侵入性小、易于保存,且包含了机体丰富的生理病理信息,是疾病早期诊断和预后研究中重要的生物样品材料。因此,本研究以 6 例正常人与 6 例糖尿病人的血清为样品,对其进行 Trypsin 酶解后分成 2 份,其中 1 份进行 LC-MS/MS 分析,得到全蛋白质组的定性和定量结果;另一份经 HILIC 富集糖肽与 PNGaseF 酶切糖肽后进行 LC-MS/MS 鉴定,得到糖肽的鉴定和差异表达结果,实验流程示于图 1。

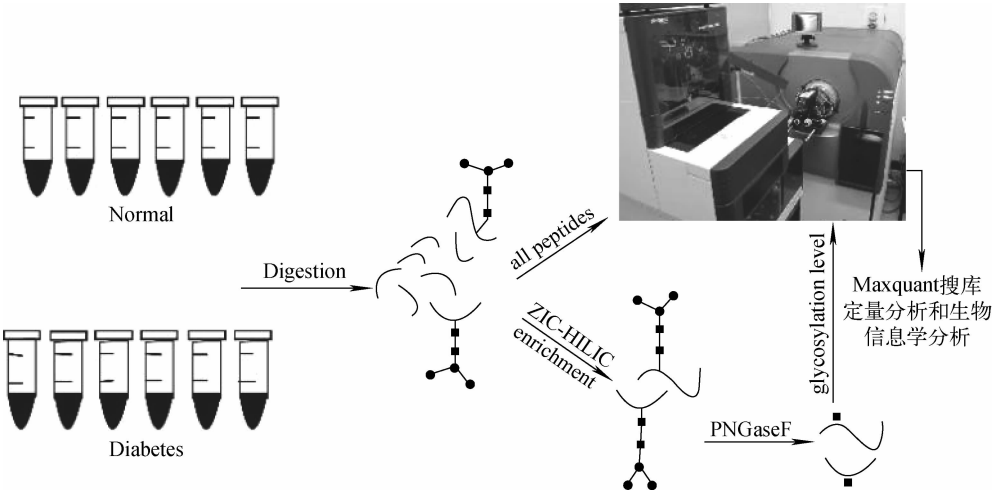


图 1 实验流程图
Fig. 1 Flow chart of analysis of serum

2.2 血清蛋白与糖蛋白定量结果分析

通过 Maxquant 搜库得到 6 例正常人与 6 例糖尿病人血清样品的质谱数据信息,在蛋白质水平上,选取 2 种样品均含有的肽段进行去冗余,共鉴定出 291 个非冗余蛋白质和2 500条非冗余肽段;在糖肽水平上,首先以 Deamidated 为条件筛选出糖基化的肽段,选取 2 种样品均含有的糖基化肽段去冗余,并找出具有糖基化位点序列特征(N-! P-S/T/C,其中! P不是脯氨酸)的肽段,结果共鉴定到 181 个非冗余糖蛋白和 637 条非冗余糖基化肽段。

采用 Student’s T-test,以蛋白质在糖尿病

患者血清和正常人血清中表达含量作为参考,找出具有显著性差异的蛋白($P<0.05$),并分别以两组样品中的蛋白强度平均值计算含量比值。为找到更多的差异蛋白质,采用表达量 $\pm 20\%$ 为差异,即蛋白在糖尿病患者血清中相比正常人血清升高 1.2 倍或者降低 0.8 倍为参考值^[23-25]。在蛋白水平上共分析得到 65 种差异蛋白质,其中有 29 种显著下调,36 种显著上调,结果列于表 1。在糖蛋白水平上,分析得到 24 种正常人血清与糖尿病患者血清的差异糖蛋白,其中有 17 种显著上调,7 种显著下调,并对糖蛋白的位点进行标注,结果列于表 2。

表 1 蛋白水平上差异蛋白信息
Table 1 Information of differential protein on protein level

蛋白名称 Protein name	序列号 Uniprot accession	差异比 Fold change	P 值 P-value	输入名 Entry name
血清淀粉样蛋白 P 组分	P02743	0	0.0286	SAMP
补充因子 I	P05156	0	0.0336	CFAI
血红蛋白亚单位 β	P68871	0	0.0379	HBB
维生素 D 结合蛋白	P02774	0	0.0099	VTDB
人基膜聚糖/内腔蛋白	P51884	0	0.0283	LUM
羧肽酶 N 亚基 2	P22792	0	0.0073	CPN2
纤维连接蛋白	P02751	0	0.0445	FINC
补体 C4-A	P0C0L4	0	0.0371	CO4A
吸引素重组蛋白	O75882	0	0.0289	ATRN
C4b 结合蛋白 β 链	P20851	0	0.0258	C4BPB
生物素酶	P43251	0	0.0331	BTD

续表 1

蛋白名称 Protein name	序列号 Uniprot accession	差异比 Fold change	P 值 P-value	输入名 Entry name
α-1B-糖蛋白	P04217	0.09	0.0278	A1BG
免疫球蛋白 κ 轻链	P0DOX7	0.10	0.0002	IGK
α-2-巨球蛋白	P01023	0.11	0.0144	A2MG
血红蛋白亚单位 α	P69905	0.12	0.0414	HBA
载脂蛋白 L1	O14791	0.15	3.8552×10 ^{−5}	APOL1
免疫球蛋白 κ 变量 3-20	P01619	0.22	0.0032	KV320
α-1-酸性糖蛋白 2	P19652	0.26	0.0435	A1AG2
凝血素	P00734	0.27	0.0058	THRB
载脂蛋白 C-I	P02654	0.31	0.0284	APOC1
免疫球蛋白重常数 γ2	P01859	0.32	0.0476	IGHG2
透明质酸结合蛋白 2	Q14520	0.35	0.0385	HABP2
补体成分 C8 γ 链	P07360	0.35	0.0376	CO8G
富含亮氨酸的 α-2-糖蛋白	P02750	0.38	0.0134	A2GL
补体 C1s 子成分	P09871	0.42	0.0192	C1S
免疫球蛋白 J 链	P01591	0.45	0.0431	IGJ
皮质类固醇结合球蛋白	P08185	0.59	0.0242	CBG
α-1-酸性糖蛋白 1	P02763	0.66	0.0470	A1AG1
免疫球蛋白 κ 常数	P01834	0.78	0.0094	IGKC
载脂蛋白 A-IV	P06727	1.60	0.0334	APOA4
载脂蛋白 A-II	P02652	1.69	0.0363	APOA2
补体因子 B	P00751	1.78	0.0211	CFAB
间 α-胰蛋白酶抑制剂重链 H2	P19823	1.86	0.0422	ITI1H2
血凝素	P02790	1.95	0.0496	HEMO
血清转铁蛋白	P02787	2.02	0.0082	TRFE
α-2-HS 糖蛋白	P02765	2.04	0.0118	FETUA
血纤维蛋白溶酶原	P00747	2.04	0.0431	PLMN
载脂蛋白 C-III	P02656	2.11	0.0452	APOC3
载脂蛋白 B-100	P04114	2.39	0.0004	APOB
血清对氧酶/芳酯酶 1	P27169	2.40	0.0135	PON1
色素上皮衍生因子	P36955	2.48	0.0400	PEDF
CD5 抗原样	O43866	2.49	0.0351	CD5L
α-1 抗胰蛋白酶	P01009	2.51	0.0024	A1AT
间 α-胰蛋白酶抑制剂重链 H4	Q14624	2.61	0.0089	ITI1H4
血清白蛋白	P02768	2.66	0.0498	ALBU
富含组氨酸的糖蛋白	P04196	2.73	0.0005	HRG
补体 C1r 子成分	P00736	2.92	0.0437	C1R
Afamin 蛋白	P43652	2.93	0.0302	AFAM
补体因子 H	P08603	3.21	0.0449	CFAH
激肽原-1	P01042	3.31	0.0128	KNG1
补体成分 C6	P13671	3.54	0.0092	CO6
抗凝血酶 III	P01008	3.75	0.0007	ANT3
C4b 结合蛋白 α 链	P04003	3.89	0.0047	C4BPA
补体 C3	P01024	4.01	0.0156	CO3
纤维蛋白原 α 链	P02671	4.66	0.0202	FIBA

续表 1

蛋白名称	序列号	差异比	<i>P</i> 值	输入名
Protein name	Uniprot accession	Fold change	<i>P</i> -value	Entry name
血小板因子 4	P02776	4. 89	0. 0105	PLF4
载脂蛋白 A-I	P02647	4. 98	0. 0386	APOA1
玻璃粘连蛋白	P04004	5. 21	0. 0117	VTNC
补体成分 C7	P10643	6. 51	0. 0315	CO7
维生素 K 依赖的蛋白 C	P04070	6. 97	0. 0478	PROC
补体 C2	P06681	6. 99	0. 0237	CO2
视黄醇结合蛋白 4	P02753	8. 68	0. 0484	RET4
AMBP 蛋白	P02760	11. 09	0. 0220	AMBP
间 α -胰蛋白酶抑制剂重链 H1	P19827	13. 44	0. 0065	ITIH1
载脂蛋白 M	O95445	# DIV/0!	0. 3632	APOM

注：# DIV/0! 代表正常血清 6 次重复中有 0 次有定量值的情况

表 2 糖尿病人与正常人血清的差异糖蛋白信息

Table 2 Differential glycoprotein information between diabetic and normal human serum

肽段序列	蛋白名称	序列号	差异比	<i>P</i> 值	输入名
Peptide sequence	Protein name	Uniprot accession	Fold change	<i>P</i> -value	Entry name
LDAPTNLQFVN * ETDSTVLVR	纤维连接蛋白	P02751	0. 17	0. 0265	FINC
EHYN * LSAATCSPGQMCGHYTQVVWAK	肽酶抑制剂 16	Q6UXB8	0. 17	0. 0328	PI16
EN * ETEIHK	胆碱酯酶	P06276	0. 24	0. 0157	CHLE
TN * STFVQALVEHVK	鞘脂激活蛋白原	P07602	0. 31	0. 0009	SAP
LAN * LTQGEDQYYLR	丛生蛋白	P10909	0. 53	0. 0007	CLUS
LSLHRPALEDLLLGSEAN * LTCTLTGLR	免疫球蛋白重常数 α 1	P01876	0. 62	0. 0120	IGHA1
SVQEIQATFFYFTPN * KTEDTIFLR	α -1-酸性糖蛋白 1	P02763	0. 71	0. 0132	A1AG1
TMFPN * LTDVR	补体 C2	P06681	1. 47	0. 0154	CO2
VTISGVYDLGDVLEEMGIADLFTNQAN * FSR	皮质类固醇结合球蛋白	P08185	1. 57	0. 0343	CBG
N * LSMPLLPADPHK	肝素辅因子 2	P05546	1. 66	0. 0341	HEP2
EGYSN * ISYIVNVNHQGISSR	硒蛋白 P	P49908	1. 70	0. 0448	SEPP1
KEHETCLAPELYNGN * YSTTQK	凝血因子 XIII B 链	P05160	1. 72	0. 0473	F13B
YNSQN * QSNNQFVLYR	激肽原-1	P01042	1. 93	0. 0374	KNG1
GINYN * SSVAK	凝血因子 XI	P03951	2. 02	0. 0111	FA11
SRPAN * HCVYFYGDEISFSCHETSR	C4b 结合蛋白 α 链	P04003	2. 28	0. 0170	C4BPA
VLYLAAYN * CTLRPVSK	胎球蛋白-B	Q9UGM5	2. 444	0. 0131	FETUB
N * GTGHGNSTHHGPEYMR	血凝素	P02790	2. 45	0. 0057	HEMO
N * QSVNVFLGHTAIDEMLK	补体 C1r 亚组分样蛋白	Q9NZP8	2. 51	0. 0298	C1RL
FVEGSHN * STVSLTTK	载脂蛋白 B-100	P04114	2. 83	0. 0101	APOB
SKWN * ITMESYVVHTNYDEYAIFLTK	AMBP 蛋白	P02760	3. 83	0. 0136	AMBP
N * HSCEPCQTLAVR	凝血因子 XII	P00748	4. 30	0. 0295	FA12
LVSGN * LTDR	ADAMTS 样蛋白 4	Q6UY14	4. 53	0. 0019	ATL4
EVSFLN * CSLDNGGCTHYCLEEVGWR	维生素 K 依赖的蛋白 C	P04070	5. 44	0. 0229	PROC
EDNAVQPN * SSYTYVWHATER	凝血因子 V	P12259	6. 87	0. 0181	FA5

注：* 表示糖蛋白的位点

2.3 蛋白与糖蛋白水平的比较分析

通过比较蛋白和糖基化水平的蛋白质,发现有 21 种蛋白质在蛋白水平上无差异而糖基化水平表达有差异,结果列于表 3。选取其中 2 种代表性的糖蛋白进行分析。

硒蛋白 P 是一种特殊的细胞外糖蛋白,有研究表明^[26-28],无论硒蛋白过量或缺乏,均会引起体内葡萄糖稳态的异常调节。本研究中,糖尿病患者血清中的硒蛋白 P 有 1.7 倍上调,根据文献报道^[27],硒蛋白 P 通过抑制肝脏中 AMP 活化蛋白激酶和乙酰 CoA 羧化酶的磷酸

化水平抑制胰岛素信号转导,从而导致胰岛素抵抗和糖尿病的发生和发展。

载脂蛋白 B-100 的主要功能是:1) 将胆固醇运送到外周组织;2) 识别、调节低密度脂蛋白受体与低密度脂蛋白的结合;3) 刺激血管平滑肌增生,与内膜基质结合^[29]。结果表明,糖尿病患者血清中的载脂蛋白 B-100 的含量有 2.83 倍上调,可能是由于糖尿病患者血浆中低密度脂蛋白胆固醇浓度升高,或高血糖下低密度脂蛋白修饰使其无法进入正常的代谢途径,从而引起病变^[30]。

表 3 蛋白水平无差异而糖基化水平有差异的蛋白质信息

Table 3 Protein information with no difference in protein levels and difference in glycosylation levels				
蛋白名称 Protein name	序列号 Uniprot accession	差异比 Fold change	P 值 P-value	输入名 Entry name
纤维连接蛋白	P02751	0.17	0.0265	FINC
胆碱酯酶	P06276	0.24	0.0157	CHLE
丛生蛋白	P10909	0.53	0.0007	CLUS
免疫球蛋白重常数 α 1	P01876	0.62	0.0120	IGHA1
α 1-酸性糖蛋白 1	P02763	0.71	0.0132	A1AG1
补体 C2	P06681	1.47	0.0154	CO2
皮质类固醇结合球蛋白	P08185	1.57	0.0343	CBG
肝素辅因子 2	P05546	1.66	0.0341	HEP2
硒蛋白 P	P49908	1.70	0.0448	SEPP1
凝血因子 XIII B 链	P05160	1.72	0.0473	F13B
激肽原-1	P01042	1.93	0.0374	KNG1
凝血因子 XI	P03951	2.02	0.0111	FA11
C4b 结合蛋白 α 链	P04003	2.28	0.0170	C4BPA
胎球蛋白-B	Q9UGM5	2.44	0.0131	FETUB
血凝素	P02790	2.45	0.0057	HEMO
补体 C1r 亚组分样蛋白	Q9NZP8	2.51	0.0298	C1RL
载脂蛋白 B-100	P04114	2.83	0.0101	APOB
AMBP 蛋白	P02760	3.83	0.0136	AMBP
凝血因子 XII	P00748	4.30	0.0295	FA12
维生素 K 依赖的蛋白 C	P04070	5.44	0.0229	PROC
凝血因子 V	P12259	6.87	0.0181	FA5

2.4 差异蛋白与糖蛋白的功能富集分析及比较

为了深入了解鉴定到的差异蛋白与差异糖蛋白所属的细胞组分、对应的分子功能和参与

的生物学过程,本实验使用 DAVID 软件(<https://david.ncifcrf.gov/>)对鉴定到的差异蛋白与差异糖蛋白进行基因本体功能富集分析

(gene ontology, GO),结果示于图 2。

从图 2 可见,差异蛋白质主要集中在血微粒、细胞外区域、胞外分泌体、细胞外间隙等区域;差异糖蛋白质主要集中于血微粒、细胞外区域、血小板、胞外分泌体等区域。在分子功能富集分析中,差异蛋白质行使的分子功能包括丝氨酸型肽链内切酶活性、丝氨酸型肽链内切酶抑制剂活性、肽链内切酶抑制剂活性、肝素结合等;差异糖蛋白的分子功能包括丝氨酸型肽链内切酶活性、肝素结合、丝氨酸型肽链内切酶抑制剂活性、免疫球蛋白 A 结合等。差异蛋白参与的主要生物学过程包

括补体激活(经典途径)、血小板解聚、补体激活的调节、补体激活等;差异糖蛋白参与的生物进程有血小板解聚、凝血(内在途径)、受体介导内存作用、凝血等。

比较差异蛋白质与差异糖蛋白质功能富集结果,发现二者细胞定位接近;在分子功能上,差异蛋白质中丝氨酸型肽链内切酶活性、丝氨酸型肽链内切酶抑制剂活性所占比例与差异糖蛋白中相应比例相差较大;生物过程中,差异蛋白参与的生物过程较多,且分布相对均匀,差异糖蛋白中血小板解聚以及凝血相关过程所占比例较大。

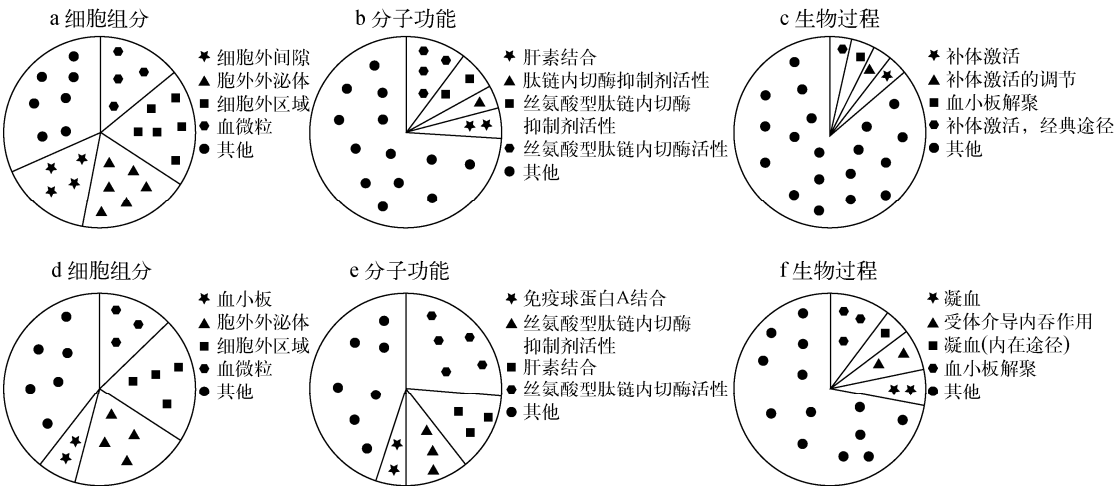


图 2 差异蛋白(a,b,c)与差异糖蛋白(d,e,f)的 GO 分析图

Fig.2 GO analysis of differentially expressed proteins (a, b, c) and differentially expressed glycoproteins (d, e, f)

3 结论

为揭示糖尿病发病过程中糖蛋白的变化,本研究根据糖尿病的特点以及分析要求,从样品选择、糖肽富集方法、LC-MS/MS 分析和数据处理等方面进行了探讨。选择血清作为研究对象,并选用稳定性好的 micro-LC-MS/MS 方法进行 label-free 检测,建立了基于所有肽段与 ZIC-HILIC 富集后肽段双重定量的方法以确定蛋白质糖基化的差异。结果表明,实验得到 65 个蛋白水平的差异蛋白,鉴定得到 24 种糖基化水平的显著差异糖蛋白,该实验可为复杂体系样品分析、差异蛋白质的定性定量分析提供可靠准确的方法,同时为糖尿病的研究提供详实的数据参考。

参考文献:

[1] RAJU T S, LANG S E. Diversity in structure and functions of antibody sialylation in the Fc [J]. Current Opinion in Biotechnology, 2014, 30: 147-152.

[2] MOREMEN K W, TIEMEYER M, NAIM A V. Vertebrate protein glycosylation: diversity, synthesis and function[J]. Nature Reviews Molecular Cell Biology, 2012, 13(7): 448-462.

[3] STADLMANN J, HOID M, TAUBENSCHMID J, MECHTLER K, PENNINGER J M. Analysis of PNGase F-resistant N-Glycopeptides using SugarQb for proteome discoverer 2.1 Reveals Cryptic Substrate Specificities[J]. Proteomics, 2018, 18(13): 1700 436-1700 440.

- [4] THAYSEN-ANDERSEN M, PACKER N H, SCHULZ B L. Maturing glycoproteomics technologies provide unique structural insights into the *N*-glycoproteome and its regulation in health and disease[J]. *Molecular & Cellular Proteomics*, 2016, 15(6): 1 773-1 790.
- [5] SCHEYS F, van DAMME E J M, de SCHUTTER K, STAES K, GEVAERT K, SMAGGHE G. Evolutionarily conserved and species-specific glycoproteins in the *N*-glycoproteomes of diverse insect species[J]. *Insect Biochemistry & Molecular Biology*, 2018, 100: 22-29.
- [6] WRIGHT R D, COOPER D. Glycobiology of leukocyte trafficking in inflammation[J]. *Glycobiology*, 2014, 24(12): 1 242-1 251.
- [7] CHEN W, SMEEKENS J M, WU R. A universal chemical enrichment method for mapping the yeast *N*-glycoproteome by mass spectrometry (MS)[J]. *Molecular & Cellular Proteomics*, 2014, 13(6): 1 563-1 572.
- [8] RUHAAK L R, TAYLOR S L, STROBLE C, NGUYEN U T, PARKER E A, SONG T, LEBRILLA C B, ROM W N, PASS H, KIM K, KELLY K, MIYAMOTO S. Differential *N*-glycosylation patterns in lung adenocarcinoma tissue [J]. *Journal of Proteome Research*, 2015, 14 (11): 4 538-4 549.
- [9] ROSA P, SILVIA B, ARIADNA S, PAULINE M R, de RAFAEL L. Altered glycosylation in tumours focused to cancer diagnosis[J]. *Disease Markers*, 2008, 25(4/5): 207-218.
- [10] HE X M, LIANG X C, CHEN X, YUAN B F, ZHOU P, ZHANG L N, FENG Y Q. High strength and hydrophilic chitosan microspheres for the selective enrichment of *N*-glycopeptides [J]. *Analytical Chemistry*, 2017, 89(18): 9 712-9 721.
- [11] ZHANG Y, JIAO J, YANG P Y, LU H J. Mass spectrometry-based *N*-glycoproteomics for cancer biomarker discovery[J]. *Clinical Proteomics*, 2014, 11(1): 18-31.
- [12] DESAIRE H. Glycopeptide analysis, recent developments and applications[J]. *Molecular & Cellular Proteomics*, 2013, 12(4): 893-901.
- [13] CARAGATA M, SHAH A K, SCHULZ B L, HILL M M, PUNYADEERA C. Enrichment and identification of glycoproteins in human saliva using lectin magnetic bead arrays[J]. *Analytical Biochemistry*, 2016, 497: 76-82.
- [14] MANN B F, MANN A K P, SKRABALAK S E, NOVOTNY M V. Macroporous silica particles derivatized for enhanced lectin affinity enrichment of glycoproteins[J]. *Analytical Chemistry*, 2013, 85(3): 1 905-1 912.
- [15] SAJID M S, JABEEN F, HUSSAIN D, ASHIQ M N, NAJAM-UI-HAQ M. Hydrazide-functionalized affinity on conventional support materials for glycopeptide enrichment[J]. *Analytical & Bioanalytical Chemistry*, 2017, 409(12): 3 135-3 143.
- [16] ALAGESAN K, KHILJI S K, KOLARICH D. It is all about the solvent: on the importance of the mobile phase for ZIC-HILIC glycopeptide enrichment[J]. *Analytical & Bioanalytical Chemistry*, 2017, 409(2): 529-538.
- [17] CHO N H, SHAW J E, KARURANGA S, HUANG Y, DA ROCHA FERNANDES J D, OHIROGGE A W, MALANDA B. IDF Diabetes Atlas: global estimates of diabetes prevalence for 2017 and projections for 2045[J]. *Diabetes Research and Clinical Practice*, 2018, 138: 271-281.
- [18] 王永胜, 杨丽霞, 程涛, 王振华, 孟祥云, 周思彤. 糖尿病肾病的炎症致病机制与中药防治[J]. *中国实验方剂学杂志*, 2018, 24(2): 200-207.
- WANG Yongsheng, YANG Lixia, CHENG Tao, WANG Zhenhua, MENG Xiangyun, ZHOU Sitong. Inflammatory pathogenesis of diabetic nephropathy and prevention and treatment of traditional Chinese medicine[J]. *Chinese Journal of Experimental Traditional Medical Formulae*, 2018, 24(2): 200-207 (in Chinese).
- [19] SYNGELAKI A, VISSER G H, KRITHINAKIS K, WRIGHT A, NICOLAIDES K H. First trimester screening for gestational diabetes mellitus by maternal factors and markers of inflammation[J]. *Metabolism-clinical & Experimental*, 2016, 65(3): 131-137.
- [20] COCELLI L P, DIKENSOY E, CICEK H, IBAR Y, KUL S, BALAT O. Pseudocholinesterase in gestational diabetes: positive correlation with LDL and negative correlation with triglyceride [J]. *Archives of Gynecology & Obstetrics*, 2012, 286(1): 43-49.

- [21] LEE L Y, THAYSEN ANDERSEN M, BAKER M S, PACKER N H, HANCOCK W S, FANAYAN S. Comprehensive *N*-glycome profiling of cultured human epithelial breast cells identifies unique secretome *N*-glycosylation signatures enabling tumorigenic subtype classification[J]. *Journal of Proteome Research*, 2014, 13(11): 4 783-4 795.
- [22] MA J, HART G W. Protein *O*-GlcNAcylation in diabetes and diabetic complications[J]. *Expert Review of Proteomics*, 2013, 10(4): 365-380.
- [23] MENG Y Y, LIU F, PANG C Y, FANG S, SONG M Z, WANG D, LI W H, YU S X. Label-free quantitative proteomics analysis of cotton leaf response to nitric oxide[J]. *Journal of Proteome Research*, 2011, 10(12): 5 416-5 432.
- [24] LIU X H, ZHENG W M, WANG W S, SHEN H L, LIU L X, LOU W H, WANG X L, YANG P Y. A new panel of pancreatic cancer biomarkers discovered using a mass spectrometry-based pipeline[J]. *British Journal of Cancer*, 2017, 117: 1 846-1 854.
- [25] MIYAMOTO-MIKAMI E, TSUJI K, HORII N, HASEGAWA N, FUJIE S, HOMMA T, UCHIDA M, HAMAOKA T, KANEHISA H, TABATA I, LEMITSU M. Gene expression profile of muscle adaptation to high-intensity intermittent exercise training in young men[J]. *Scientific Reports*, 2018, 8(1): 16 811-16 824.
- [26] MITA Y, NAKAYAMA K, INARI S, NISHITO Y, YOSHIOKA Y, SAKAI N, SOTANI K, NAGAMURA T, KUZUHARA Y, INAGAKI K, IWASAKI M, MISU H, IKEGAWA M, TAKAMURA T, NOGUCHI N, SAITO Y. Selenoprotein P-neutralizing antibodies improve insulin secretion and glucose sensitivity in type 2 diabetes mouse models[J]. *Nature Communications*, 2017, 8(1): 1 658-1 674.
- [27] 周军, 白兆帅, 徐辉碧, 黄开勋. 硒蛋白与糖尿病——硒的两面性[J]. *化学进展*, 2013, 25(4): 488-494.
- ZHOU Jun, BAI Zhaoshuai, XU Huibi, HUANG Kaixun. Selenium and diabetes-two-sidedness of selenium[J]. *Progress in Chemistry*, 2013, 25(4): 488-494(in Chinese).
- [28] LABUNSKYY V M, LEE B C, HANDY D E, LOSCALZO J, HATFIELD D L, GLADY-SHEV V N. Both maximal expression of selenoproteins and selenoprotein deficiency can promote development of type 2 diabetes-like phenotype in mice[J]. *Antioxid Redox Signal*, 2011, 14(12): 2 327-2 336.
- [29] 陈勇, 金权膺, 蒋莹. 载脂蛋白 A1, B100 与糖尿病[J]. *江苏医药*, 1990, 16(4): 205-208.
- CHEN Yong, JIN Quanying, JIANG Ying. Apolipoprotein A1, B100 and diabetes[J]. *Jiangsu Medical Journal*, 1990, 16(4): 205-208(in Chinese).
- [30] 陈霞. 低密度脂蛋白与糖尿病的关系[J]. *内科*, 2012, 7(3): 304-306.
- CHEN Xia. The relationship between low density lipoprotein and diabetes[J]. *Internal Medicine of China*, 2012, 7(3): 304-306(in Chinese).