

## • 研究论文 •

## 基于液质联用蛋白质组学技术的阿尔茨海默病血清生物标志物及生物信息学研究

许婷婷<sup>1</sup>, 郭鹏<sup>2</sup>, 张巍<sup>2\*</sup>, 王晓良<sup>1\*</sup>

(1. 中国医学科学院、北京协和医学院药物研究所, 北京 100050; 2. 首都医科大学附属北京天坛医院, 神经病学中心, 北京 100070)

**摘要:** 为了寻找有效的阿尔茨海默病 (Alzheimer's disease, AD) 血液生物标志物, 本研究采用非标记液相色谱-串联质谱技术对 AD 患者和健康对照受试者 (healthy control subjects, HC) 的血清进行比较蛋白质组学分析, 成功鉴定到 387 个蛋白, 其中 61 个蛋白是显著变化的。基因本体 (gene ontology, GO) 富集分析结果显示, 一些与 AD 发病有关的 GO 功能条目在差异蛋白中显著富集, 包括胆固醇和脂质代谢、炎症反应、凝血和止血过程以及免疫反应等生物学过程。因此, 根据上述结果和蛋白在 8 个比较组中含量变化的一致性程度, 共选取了 18 个蛋白作为 AD 候选生物标志物。蛋白-蛋白相互作用结果提示这 18 个蛋白与淀粉样  $\beta$  (A4) 前体蛋白 [amyloid beta (A4) precursor protein, APP] 都具有直接或间接的作用, 因此, 推测这些蛋白表达水平或功能的改变可能会影响淀粉样蛋白 (amyloid- $\beta$ , A $\beta$ ) 代谢和参与 AD 的发生, 具有成为 AD 血液生物标志物的潜能。

**关键词:** 阿尔茨海默病; 生物标志物; 血清; 非标记液相色谱-串联质谱法; 蛋白质组学  
中图分类号: R963 文献标识码: A 文章编号: 0513-4870(2020)07-1604-10

## Serum biomarkers and bioinformatics of Alzheimer's disease based on LC-MS/MS proteomics

XU Ting-ting<sup>1</sup>, GUO Peng<sup>2</sup>, ZHANG Wei<sup>2\*</sup>, WANG Xiao-liang<sup>1\*</sup>

(1. Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China; 2. Neurology Center, Beijing Tian Tan Hospital, Capital Medical University, Beijing 100070, China)

**Abstract:** To identify potential serum proteins that might serve as biomarkers for Alzheimer's disease (AD), we performed comparative proteomic profiling of sera from AD and healthy control subjects using label-free LC-MS/MS. Our study identified 387 proteins, 61 of which showed significant changes in the serum of AD patients compared to healthy controls. Gene ontology (GO) enrichment analysis showed that some GO terms related to the pathogenesis of AD were significantly enriched in differentially expressed proteins, including cholesterol and lipid metabolism, inflammation, coagulation and hemostasis processes, and immune responses. Therefore, based on the above results and the consistency of protein content changes in the 8 comparison groups, 18 proteins were selected as candidate biomarkers. Protein-protein interaction results suggest that these 18 proteins can directly or indirectly interact with APP. Therefore, changes in the levels or functions of these proteins may affect A $\beta$  metabolism and participate in the occurrence of AD, and have the potential to become AD blood biomarkers.

**Key words:** Alzheimer's disease; biomarker; serum; label free LC-MS/MS; proteomics

收稿日期: 2020-05-18; 修回日期: 2020-06-18.

基金项目: 国家自然科学基金资助项目 (81573417).

\*通讯作者 Tel / Fax: 86-10-63165330, E-mail: wangxl@imm.ac.cn;

Tel: 13911996107, Fax: 86-10-59975128, E-mail: ttyyzw@163.com

DOI: 10.16438/j.0513-4870.2020-0770

阿尔茨海默病 (Alzheimer's disease, AD) 作为一种神经退行性疾病, 现已成为威胁老年人身心健康的重大疾病。据统计, 目前全球大约有 4 400 万人患有痴呆症<sup>[1]</sup>。AD 是痴呆中最主要的类型。在美国, AD 已成为 65 岁以上老年人的第五大常见死因<sup>[2]</sup>。目前, 我国也有 700 多万老年人正在遭受 AD 的困扰, 其中 65 岁以上人群 AD 患病率为 3.21%<sup>[3,4]</sup>。AD 发病率的持续增加, 必将会给全球带来沉重的经济负担<sup>[5]</sup>。因此, AD 被认为是当代全球公共卫生的又一大挑战。

AD 从发现至今已有 100 多年的历史, 但是发病机制仍不明确。目前, 临床用于治疗 AD 的药物均是对症治疗, 无法延缓或逆转疾病的进程, 因此, AD 尚无有效的治疗药物<sup>[6]</sup>。同时药物研发状况也并不理想, 大多数临床试验均以失败告终<sup>[7-9]</sup>。鉴于 AD 的疾病特点和研究现状, 寻找到可用于 AD 早期诊断的生物标志物和新的药物靶点尤为重要。目前 AD 的生物标志物主要包括功能神经影像学 and 生物化学的标志物, 其中神经影像学 and 脑脊液 (CSF) 的生物标志物已应用于 AD 发病前期轻度认知功能障碍 (mild cognitive impairment, MCI) 和临床前阶段的诊断<sup>[10]</sup>。CSF 的生物标志物包括: 低水平的  $A\beta_{1-42}$  或  $A\beta_{1-40}$  和高水平的总 Tau 蛋白及磷酸化的 Tau 蛋白; 功能神经影像学的标志物包括: 正电子发射断层扫描 (PET) 技术检测大脑中淀粉样蛋白沉积的增加, 磁共振成像 (MRI) 技术检测内侧颞叶萎缩和氟脱氧葡萄糖正电子发射断层扫描 (FDG-PET) 技术检测脑内葡萄糖代谢改变<sup>[11,12]</sup>。但是由于脑脊液样本的获取具有创伤性、神经影像学检查的费用较高, 限制了二者在临床进行早期筛查中的广泛应用。相比之下, 血液样本在样品获取、检查费用以及快速批量筛查上具有很大优势, 因此, 血液生物标志物已成为 AD 生物标志物的研究热点。

近年来, 随着液相和质谱技术的发展, 基于质谱的蛋白质组学技术已被广泛应用于临床和基础研究中, 其中同位素标记相对和绝对定量 (isobaric tags for relative and absolute quantification, iTRAQ) 和非标记定量蛋白质组学技术在 AD 血液和 CSF 生物标志物的研究中已被使用<sup>[13-16]</sup>。相比于 iTRAQ, 非标记液相-质谱蛋白质组学技术 (label free LC-MS/MS) 因其具有样品制备简单、重现性好、成本较低等优点, 得到了广泛应用<sup>[17,18]</sup>。为了进一步研究鉴定到的差异蛋白在 AD 中的作用, 本文尝试将蛋白质组学研究结果进行生物信息学分析, 鉴定了 32 个 AD 患者和 32 个健康对照 (HC) 血清样本中存在的差异蛋白。然后, 根据基因本体富集和蛋白相互作用分析发现, 18 个蛋白的含量或功能的变化可能与 AD 的发病有不同程度的相关性, 它们

有望成为 AD 血液生物标志物。希望本文工作对发现 AD 血液生物标志物、揭示它们在 AD 发生中的作用以及寻找治疗 AD 药物的新靶点带来一些新的思路。

## 材料与方法

**血清样本** AD 患者血清样品来自于北京天坛医院, 该实验所有的参与者均签署了知情同意书。AD 患者的诊断是依据美国老龄研究所和阿尔茨海默症协会 (NIA-AA) 的诊断标准。健康对照受试者的纳入标准包括健康状况良好, 既往无神经、精神相关的疾病史, 空腹血糖、血压和血脂正常。AD 患者和健康对照受试者的人口统计学和临床特征详见表 1。

**Table 1** The demographics and clinical characteristics of Alzheimer's disease (AD) patients and healthy control subjects. The detailed clinical data of 2 AD patients were lost. Data are displayed as the mean  $\pm$  SD

Item	Subject	
	Health control	AD
Numbers	32	32
Mean age, years (SD)	66.4 $\pm$ 4.2	66.8 $\pm$ 4.6
Gender, male/female	10/22	12/20
Mini mental state exam score (SD)	-	19.0 $\pm$ 9.0
Medication history (Yes/No)	-	10/20
Mean duration of disease, years (SD)	-	3.6 $\pm$ 4.0
Glu / mmol·L <sup>-1</sup>	5.11 $\pm$ 0.89	5.40 $\pm$ 1.50
TC / mmol·L <sup>-1</sup>	4.00 $\pm$ 1.00	3.90 $\pm$ 0.85
TG / mmol·L <sup>-1</sup>	1.10 $\pm$ 0.65	1.15 $\pm$ 0.79

**血清样本制备** 将血液样品置于真空负压的不添加任何抗凝剂采血管中, 室温静置 2 h 后, 将其离心 (3 000 r·min<sup>-1</sup>, 5 min, 4 °C), 得到的上清液即为血清, 收集血清样本于 -80 °C 储存备用。取 32 例 AD 患者和 32 例 HC 的血清样本, 按照等体积原则分别将 4 例 AD 或 HC 血清样本混合成 1 个血清样本池进行实验, 因此, 本研究使用了 8 个 AD 患者血清样本池和 8 个 HC 血清样本池, 进行蛋白质组学分析。利用 Agilent 公司的抗体亲和柱 (multiple affinity removal column HSA/IgG, 4.6 mm $\times$ 50 mm, Agilent, Palo Alto, CA, USA) 去除人血清中 2 种高丰度蛋白 (白蛋白和 IgG)。按照低丰度蛋白的出峰时间收集低丰度蛋白。随后, 使用 Amicon Ultra-4 离心过滤器 (截留值为 3 kDa; Millipore, Billerica, MA, USA) 在 12 000  $\times g$ 、4 °C 条件下离心 30 min, 以进行脱盐, 交换缓冲液和浓缩。在处理后的样品中加入 10 mmol·L<sup>-1</sup> 二硫苏糖醇 (DTT, Genview, USA) 在 37 °C 下还原 1 h, 然后再加入 50 mmol·L<sup>-1</sup> 碘乙酰胺 (IAA, Sigma-Aldrich Co., St. Louis, USA) 在室温、黑暗环境下放置 1 h 进行烷基化。最后, 将蛋白质用胰蛋白酶 (Promega, Madison, WI, USA) 以 1:50 的比例在 37 °C

下消化过夜。收集消化的样品, 无需进一步处理即可保存在 $-80^{\circ}\text{C}$ 下备用。

**Label free LC-MS/MS 分析** 血清样品采用纳升流速 HPLC 液相系统 Easy nLC 1000 进行分离。缓冲液: A 液为 0.1% 甲酸水溶液, B 液为 0.08% 甲酸-乙腈水溶液。色谱柱以 95% 的 A 液平衡。样品由自动进样器上样到质谱预柱 (C18, 0.10 mm×20 mm, 3 mm), 再经分析柱 (C18, 0.15 mm×120 mm, 1.9 mm) 分离, 流速为  $600\text{ nL}\cdot\text{min}^{-1}$ 。血清样品经毛细管高效液相色谱分离后用 Q-Exactive HF 质谱仪 (Thermo Scientific) 进行质谱分析。

**质谱数据搜库、蛋白定量和差异蛋白筛选** 质谱分析原始数据是用软件 Sequest HT 和 Proteome Discoverer (Thermo Scientific) 进行搜库定性及定量分析。本次使用数据库: uniprot-human\_170221, 搜库软件: Proteome Discoverer。搜库时将原始质谱数据文件通过 Proteome Discoverer 提交至 Sequest HT 服务器, 选择已经建立好的数据库, 然后进行数据库搜索, 结果以 Peptide FDR  $\leq 0.01$  筛选过滤。在 1.2 差异倍数 (FC = fold change) 和  $P < 0.05$  阈值条件下, 进行差异蛋白筛选。FC  $\geq 1.20$  和  $P < 0.05$  为上调 (up), FC  $\leq 0.83$  和  $P < 0.05$  为下调 (down),  $0.83 < \text{FC} < 1.20$  认为表达量无明显变化。

**生物信息学分析** 基因本体 (gene ontology, GO) 富集分析是把所有差异蛋白向 GO 数据库 (<http://www.geneontology.org/>) 的各个条目映射, 计算每个条目的蛋白数目, 然后应用超几何检验, 找出与所有蛋白背景相比, 在差异蛋白中显著富集的 GO 条目。其计算公式为:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}} \quad (1)$$

式 (1) 中  $N$  为所有蛋白中具有 GO 注释信息的蛋白数目,  $n$  为  $N$  中差异蛋白的数目,  $M$  为所有蛋白中注释到某个 GO 条目的蛋白数目,  $m$  为注释到某个 GO 条目的差异蛋白数目。计算得到  $P$  值, 以  $P \leq 0.05$  为阈值, 满足此条件的 GO 条目定义为在差异蛋白中显著富集的 GO 条目。

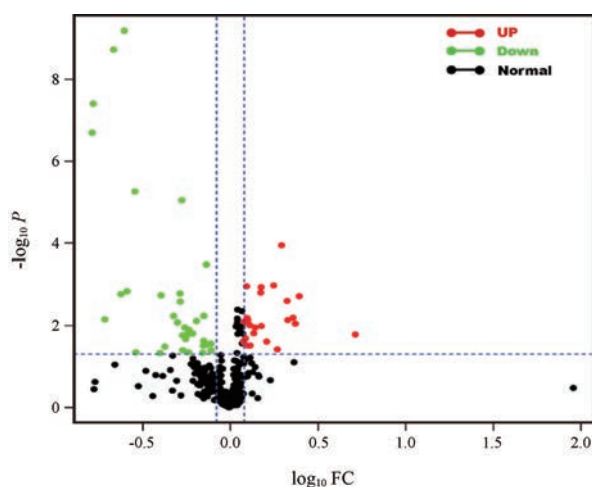
蛋白-蛋白相互作用分析使用在线生物信息网站 String (<https://string-db.org/>) 进行蛋白互用分析。将蛋白全称上传到 String 10.5 以获取蛋白互作信息。

**结果统计** 临床数据结果以  $\text{mean} \pm \text{SD}$  表示。采用独立样本  $t$  检验,  $*P < 0.05$  认为存在显著性差异。

## 结果

### 1 AD 患者和健康受试者血清差异蛋白质组学分析

将 8 个 AD 和 8 个 HC 血清样品池除去 2 种主要的高丰度蛋白 (白蛋白和 IgG) 后, 使用 label free LC-MS/MS 蛋白质组学技术分析和鉴定了血清中存在的差异表达的蛋白。结果显示, 成功鉴定了 387 个蛋白, 然后按照差异倍数 (FC  $\geq 1.20$  或  $\leq 0.83$ ) 和统计学显著性 ( $P < 0.05$ ) 标准进行筛选, 共得到 61 个差异蛋白。与 HC 相比, 在 AD 组中表达上调的共 26 个 (红色圆点数), 下调的共 35 个 (绿色圆点数), 如图 1 所示, 具体的蛋白信息见表 2 和表 3。



**Figure 1** Volcano plot of the differentially expressed proteins. FC: Fold change. The red dot represents the number of up-regulated protein, green represents the number of down-regulated protein, and black represents no difference

### 2 血清差异蛋白的表达模式聚类分析

为了更加直接地观察本研究鉴定到的差异蛋白在不同样品间表达水平的相近程度和变化情况, 本文采用了表达模式聚类分析方法进行研究。该方法是模式识别和数据挖掘中普遍使用的一种方法, 根据数学特征提取分类标准, 对数据进行分类, 从而观察数据相近程度<sup>[19]</sup>。本研究将差异蛋白的表达水平作为分类标准, 将每个样品池单独作为一类, 然后根据差异蛋白的表达量将不同类之间距离最近的进行合并, 合并后重新计算类间距离, 这个过程一直持续到将所有样本归为一类为止。距离越近说明两组数据性质越相近, 距离越远说明数据关联较远。

表达模式聚类分析结果表明, 根据差异蛋白在不同血清样本中的表达量可以很好地将 16 个血清样本池分别聚类于 AD 和 HC 组中, 不同组别样本间没有交叉 (图 2), 并且它们在同一组的不同血清样本池间的表达也具有较好的一致性, 其中在 AD 或 HC 组 8 个样品

**Table 2** Identification of proteins significantly increased in the serum from AD patients compared with health controls. Accession means accession number, provided by the uniprot-human\_170221.fasta. pI is the theoretical calculated from the amino acid sequence of the predicted mature protein. MW refers to the theoretical molecular mass is kDa calculated from the amino acid sequence

Accession	Protein	Coverage	Peptide	MW	pI	Score	Average FC (AD/HC)	P-value
P07996	Thrombospondin-1	20.9	20	129.3	4.94	2 057	5.25	0.02
P02776	Platelet factor 4	35.6	4	10.8	8.62	1 421	2.51	< 0.01
B4DTH2	cDNA FLJ61165, highly similar to fibronectin	50.4	14	37.9	5.52	12 377	2.39	< 0.01
A0A1B2FJA7	Apolipoprotein L1 (fragment)	35.9	9	27.1	7.61	3 199	2.32	< 0.01
C9IZP8	Complement C1s subcomponent (fragment)	16.5	1	11.8	4.2	376	2.16	< 0.01
A0A0S2Z3Y1	Lectin galactoside-binding soluble 3 binding protein isoform 1 (fragment)	23.4	10	65.3	5.27	1 457	2.14	< 0.01
P02775	Platelet basic protein	38.3	6	13.9	8.79	2 320	1.99	< 0.01
P01766	Immunoglobulin heavy variable 3-13	19.0	2	12.5	7.08	369	1.88	0.04
B4DPQ0	Complement C1r subcomponent	45.6	20	81.8	6.37	7 629	1.80	< 0.01
Q5NV69	V1-13 protein (fragment)	17.0	1	10.3	5.17	113	1.64	0.02
P08709	Coagulation factor VII	9.4	3	51.6	7.23	143	1.53	0.01
B4DPQ3	cDNA FLJ51034, highly similar to vitamin K-dependent protein C (EC 3.4.21.69)	24.8	10	57.4	7.53	1 302	1.52	< 0.01
A8K2N0	cDNA FLJ77835, highly similar to Homo sapiens complement component 1, s subcomponent (C1S), transcript variant 2, mRNA	33.4	19	76.6	4.98	6 543	1.52	< 0.01
P20851	C4b-binding protein beta chain	25.4	5	28.3	5.14	837	1.41	0.01
P08253	72 kDa type IV collagenase	5.8	3	73.8	5.47	271	1.38	0.02
Q6UX71	Plexin domain-containing protein 2	2.3	1	59.5	6.46	400	1.32	0.03
B0YIW2	Apolipoprotein C-III	41.0	4	12.8	8.18	5 342	1.31	0.01
P00748	Coagulation factor XII	27.2	13	67.7	7.74	7 976	1.29	0.03
P80108	Phosphatidylinositol-glycan-specific phospholipase D	23.1	17	92.3	6.37	4 092	1.28	< 0.01
Q13790	Apolipoprotein F	11.7	3	35.4	5.64	1 777	1.27	< 0.01
A8K6C9	cDNA FLJ78037, highly similar to Homo sapiens insulin-like growth factor 2 (somatomedin A), mRNA	13.9	2	20.2	9.32	506	1.26	< 0.01
P02753	Retinol-binding protein 4	71.6	10	23	6.07	12 363	1.24	0.02
P49908	Selenoprotein P	14.4	6	43.2	7.87	1 386	1.23	< 0.01
P43251	Biotinidase	21.2	9	61.1	6.25	3 638	1.21	0.03
O00391	Sulfhydryl oxidase 1	15.5	9	82.5	8.92	2 181	1.21	0.02
A0A0S2Z4L3	ProteinS isoform 2 (fragment)	36.9	20	78.8	5.74	5 933	1.21	< 0.01

池中的5个及以上中表达量一致且在组间存在差异的蛋白共35个,即一半以上的差异蛋白在组内的表达存在较好的重复性,而在组间具有明显差异,因此,本研究所用的非标记液质联用蛋白定量方法所得到的结果是相对稳定和可靠的,为后续研究奠定了基础。

### 3 血清差异蛋白的GO富集分析

GO分析已成为生物信息领域中一个重要的方法和工具,它通过利用本体所代表的生物学信息,从而实现了生物体中各种基因组产物(例如基因、蛋白质等)生物学功能的描述<sup>[20]</sup>。结果表明,60个GO功能条目显著富集在差异蛋白中,其中一些GO功能条目支持了目前研究认为的可能与AD发病机制相关的分子生物学过程,包括胆固醇和脂质代谢、炎症反应、凝血和止血过程和免疫反应等(图3)。因此,结合表达模式聚类分析结果,从中选取了15个参与上述生物学过程

的差异蛋白作为AD候选相关蛋白。此外,根据蛋白质在8个比较组(例如AD-1/HC-1、AD-2/HC-2等)中含量变化的一致性程度,又选取了3个参与图3所示生物学过程的蛋白作为AD候选生物标志物。尽管这3个蛋白在蛋白质组学的结果中不满足 $P < 0.05$ 要求(满足差异倍数要求),仍选取它们的原因是因为这3个蛋白在8个比较组中观察到较好的一致性变化,提示它们可能在AD疾病进展中具有相对稳定和持续的功能。因此,作者认为这18个蛋白可作为AD候选的血液生物标志物进行后续研究,这些蛋白的具体信息见表4<sup>[13,15,21-32]</sup>。

### 4 血清差异蛋白的相互作用分析

使用String数据库(Version10.5)探索18个蛋白之间以及它们与已知的AD相关蛋白之间是否具有相关性,包括淀粉样蛋白(A4)前体蛋白[amyloid beta (A4)

**Table 3** Identification of proteins significantly decreased in the serum from AD patients compared with health controls

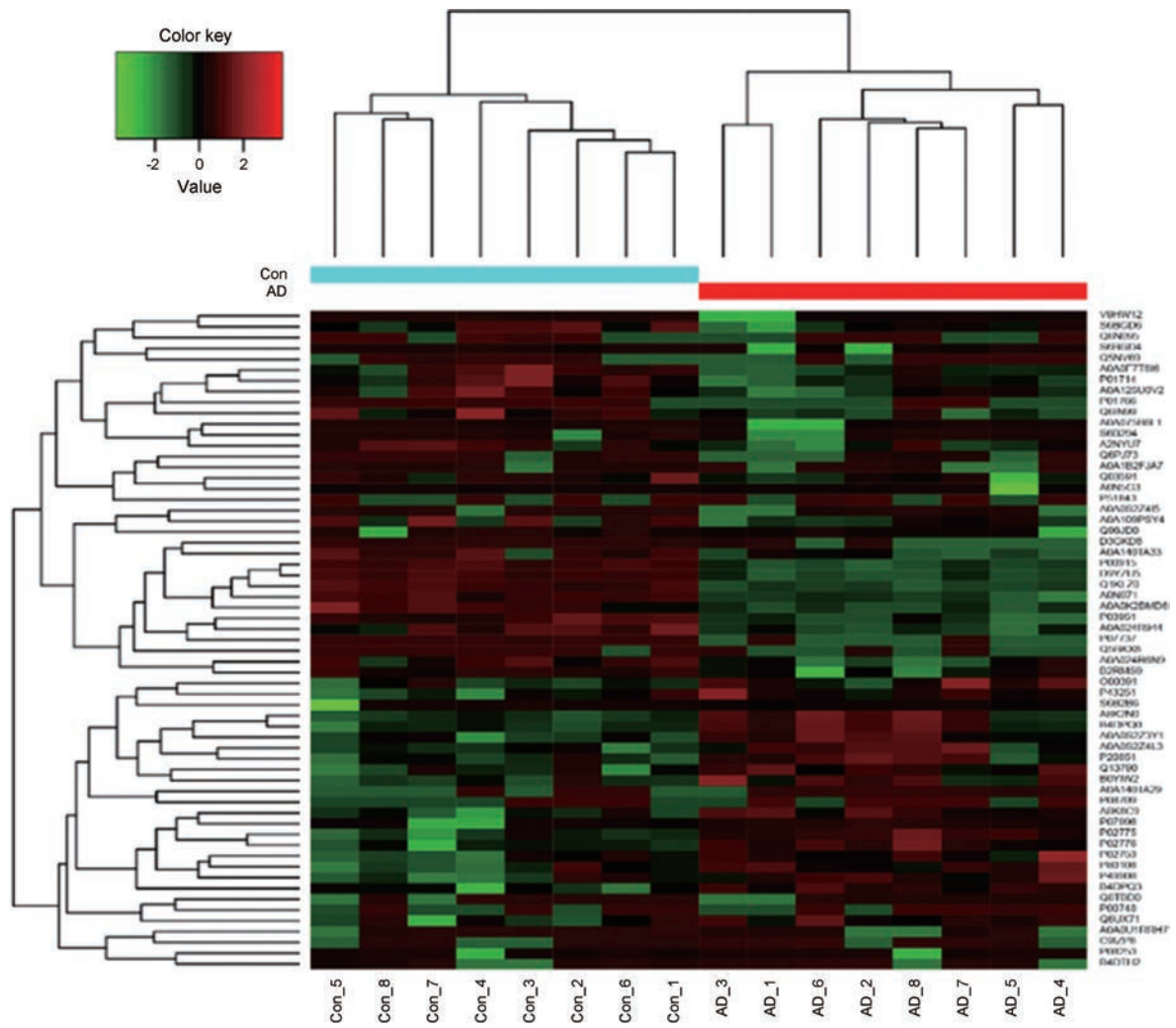
Accession	Protein	Coverage	Peptide	MW	pI	Score	Average FC (AD/HC)	P-value
S6BGD6	IgG L chain	49.4	8	24.8	7.24	18 763	0.79	0.04
A0N5G3	Rheumatoid factor G9 light chain (fragment)	42.1	4	12.9	6.52	3 983	0.79	0.03
P51843	Nuclear receptor subfamily 0 group B member 1	1.5	1	51.7	8.07	223	0.79	0.04
Q03591	Complement factor H-related protein 1	47.9	13	37.6	7.39	10 465	0.75	0.03
A0A024R944	Serpin peptidase inhibitor, clade C (antithrombin), member 1, isoform CRA_a	50.2	23	52.6	6.71	30 745	0.74	< 0.01
A0A125U0V2	GCT-A1 heavy chain variable region (fragment)	19.7	2	13.7	8.78	2 650	0.72	0.02
P01714	Immunoglobulin lambda variable 3-19	17.0	2	12	4.96	837	0.72	< 0.01
S6B2B6	IgG H chain	8.4	2	29.3	8.68	2 078	0.71	0.03
A0A109PSY4	MS-A1 light chain variable region (fragment)	32.1	2	11.4	8.48	1 235	0.71	0.05
A0A0F7T8I6	IGHV4-31 protein (fragment)	9.0	1	10.8	8.44	481	0.65	< 0.01
Q6IN99	IGL@ protein	40.4	7	24.9	6.74	18 340	0.62	0.02
A2NYU7	Heavy chain Fab (fragment)	7.1	1	14	9.25	844	0.59	0.01
S6BGD4	IgG H chain	12.0	2	25	9.48	2 109	0.59	0.04
P07737	Profilin-1	35.7	4	15	8.27	239	0.58	0.02
A0A0S2Z4I5	Complement factor properdin isoform 1 (fragment)	20.3	7	51.2	7.9	280	0.57	0.02
Q8TBD0	Uncharacterized protein (fragment)	7.6	1	14	7.9	282	0.56	0.01
S6B294	IgG L chain	27.6	4	20.1	8.72	13 044	0.54	0.04
A0A140TA29	Complement C4-B	59.8	81	187.5	7.33	110 690	0.54	0.02
P03951	Coagulation factor XI	29.9	15	70.1	8.1	3 227	0.54	< 0.01
A0A140TA33	Tenascin-X	1.3	5	458	5.17	289	0.53	< 0.01
B2RMS9	Inter-alpha (globulin) inhibitor H4 (plasma kallikrein-sensitive glycoprotein)	47.6	34	103.3	6.98	47 641	0.53	< 0.01
A0A0U1RRH7	Histone H2A	17.1	1	18.5	11.52	583	0.51	< 0.01
A0A024R6N9	Serpin peptidase inhibitor, clade A (alpha-1 antitrypsin, antitrypsin), member 5, isoform CRA_a	20.0	8	45.6	9.26	1 770	0.48	< 0.01
Q6N095	Uncharacterized protein	12.6	5	52.3	8.57	9 599	0.43	0.03
A0A075B6L1	Ig lambda-7 chain C region (fragment)	41.5	3	11.3	8.29	13 453	0.41	< 0.01
D3GKD8	A-gamma globin Osilo variant	59.9	7	16.2	7.2	1 937	0.40	0.05
Q6PJ73	DPF1 protein (fragment)	5.1	1	47.5	8.1	19	0.29	0.04
V9HW12	Epididymis secretory sperm binding protein Li 2a	32.8	7	21.9	5.97	639	0.29	< 0.01
P00915	Carbonic anhydrase 1	62.8	12	28.9	7.12	2 520	0.22	< 0.01
Q5BKX8	Muscle-related coiled-coil protein	2.7	1	41.9	8.28	32	0.26	< 0.01
D9YZU5	Beta-globin	95.2	13	16	7.28	32 360	0.25	< 0.01
A0A0K2BMD8	Mutant hemoglobin alpha 2 globin chain	91.5	10	15.2	8.68	20 255	0.24	< 0.01
Q96JD0	Amyloid lambda 6 light chain variable region SAR (fragment)	20.7	2	12.3	5.01	3 245	0.20	< 0.01
Q1KLZ0	HCG15971, isoform CRA_a	38.9	11	41.7	5.48	2 915	0.17	< 0.01
A0N071	Delta globin	83.7	13	16	8.05	18 860	0.17	< 0.01

precursor protein, APP]、载脂蛋白 E (apolipoprotein E, ApoE)、Tau 蛋白 (Tau)、早老素 1 (presenilin 1, PSEN1)、早老素 2 (presenilin 2, PSEN2) 和  $\beta$ -淀粉样前体蛋白裂解酶 1 ( $\beta$ -site APP-cleaving enzyme 1, BAEC1)。结果如图 4 所示, 所有蛋白都富集在蛋白相互作用的网络图中, 主要集中在 5 个功能模块, 包括 APP 代谢 (APP、PSEN1、PSEN2、BACE1 和 Tau)、胆固醇和脂类代谢 (ApoE、ApoM、ApoL1、RBP4 和 GPLD1)、补体 (C1R、C4B、C4BPB、C1QA 和 CFHR1)、凝血和止血过程 (F7、F12、F11、KNG1、AT-III 和 THBS1) 以及免疫和炎症反应 (MMP2、PPBP 和 PF4), 它们之间形成了紧密相连, 其中凝血和止血功能模块与 APP 结合更为紧密。这

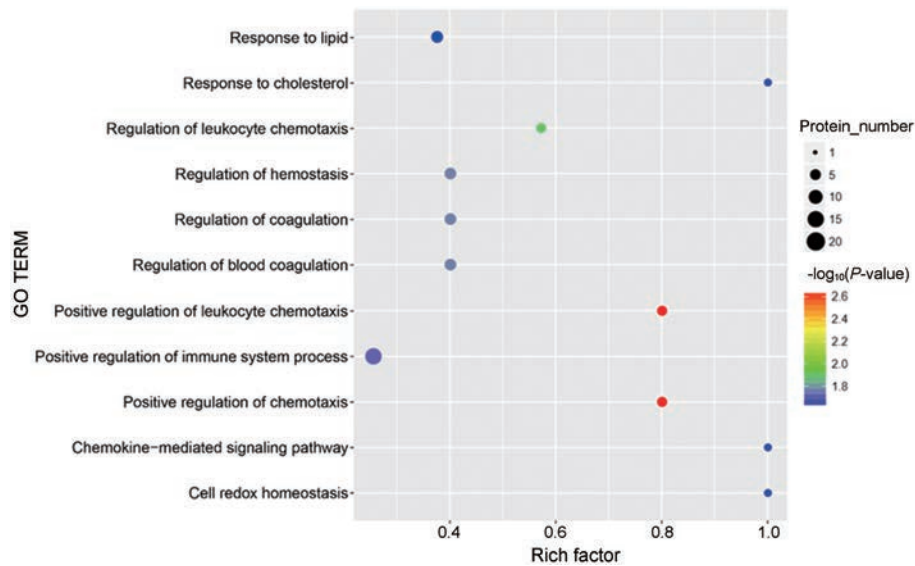
一结果表明选取的 18 个 AD 候选生物标志物与 APP 都具有直接或间接的作用, 它们可能通过影响  $A\beta$  代谢参与 AD 中来, 因此, 可推测这 18 个蛋白的表达水平或功能的改变可能与 AD 的发生有关, 具有成为 AD 血液生物标志物的潜能。

## 讨论

AD 是作为常见的神经退行性疾病, 其主要病理学特征为脑内细胞外  $A\beta$  沉积形成的老年斑、细胞内 Tau 蛋白过度磷酸化构成的神经元纤维缠结、胆碱能缺乏以及与认知、记忆密切相关脑区神经元的丢失和突触的改变等<sup>[33]</sup>。关于 AD 的病因及发病机制至今尚



**Figure 2** Cluster analysis of differential protein expression patterns. Each row in the figure represents a protein, each column is a sample/repetition, and different colors indicate different expression levels (quantity values are  $\log_2$  taken during the drawing, and median corrected)



**Figure 3** Gene ontology (GO) enrichment analysis of differentially expressed proteins. The x-axis rich factor represents the a multiple of the change in the total proportion of differentially expressed proteins in the GO classification relative to the proportion of identified proteins. The y-axis represents the enrichment significance of GO biological processes may be associated with pathogenesis of AD

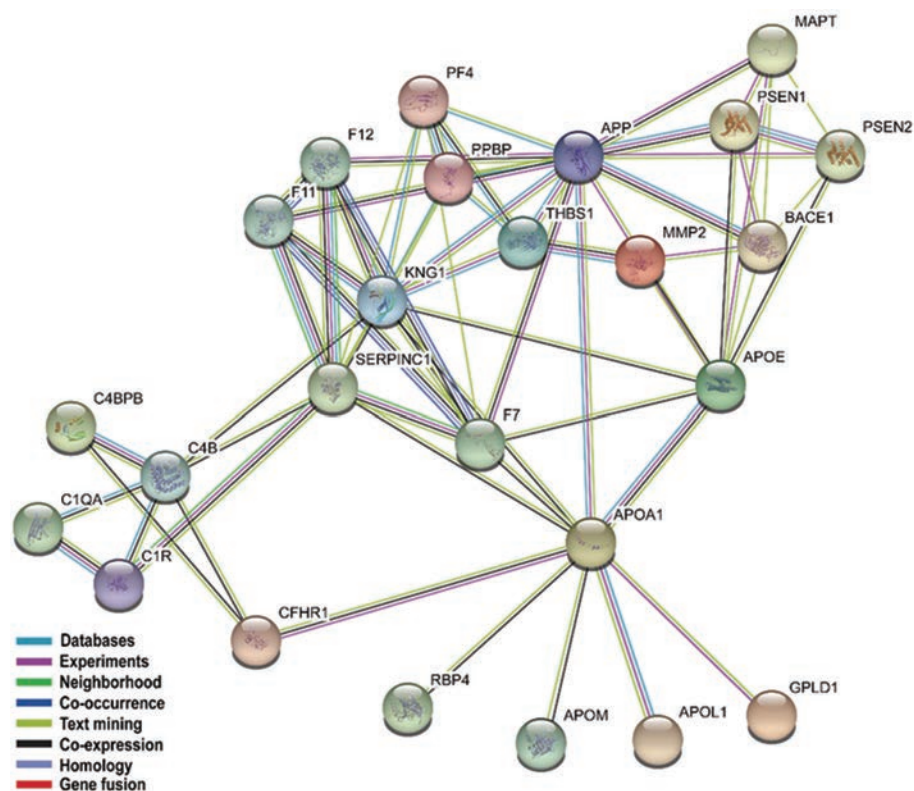
**Table 4** The function of 18 proteins selected as candidate biomarkers for AD and their related research in AD blood biomarkers. 15 differentially expressed proteins meet the fold difference criteria (ratios  $\geq 1.20$  or  $\leq 0.83$ ) and  $P < 0.05$ . ApoM, KNG1 and C1QA meet the fold difference criteria, but don't meet the statistical significance criteria. (+), protein increased in abundant; (-), protein decreased in abundant; (=), no significant difference; No, not been reported

Accession	Protein	Gene name	MW	FC	P-value	Function	References in AD blood-based biomarkers
P03951	Coagulation factor XI	F11	70.1	0.54	< 0.01	Coagulation, inflammation	(-) <sup>[21]</sup>
P02775	Platelet basic protein	PPBP	13.9	2.00	< 0.01	Coagulation, inflammation, immune response	(-) <sup>[15]</sup>
A0A024R944	Antithrombin-III	SERPINC1	52.6	0.74	< 0.01	Anti-coagulation, anti-inflammation	(+) <sup>[22]</sup>
B4DPQ0	Complement C1r subcomponent	C1R	81.8	1.80	< 0.01	Complement system	No
P02776	Platelet factor 4	PF4	10.8	2.51	< 0.01	Coagulation, inflammation response, wound repair	No
A0A1B2FJA7	Apolipoprotein L1 (fragment)	APOL1	27.1	2.32	< 0.01	Cholesterol metabolism, autophagic cell death, inflammatory response, innate immunity	No
P80108	Phosphatidylinositol-glycan-specific phospholipase D	GPLD1	92.3	1.28	< 0.01	Hydrolyze the inositol phosphate linkage in proteins anchored by phosphatidylinositol glycans	No
P08709	Coagulation factor VII	F7	51.6	1.53	0.01	Coagulation	No
P20851	C4b-binding protein beta chain	C4BPB	28.3	1.41	0.01	Accelerate decay of classical and lectin C3 convertase	(-) <sup>[23]</sup> ; (+) <sup>[24]</sup>
P02753	Retinol-binding protein 4	RBP4	23	1.24	0.02	Retinol transporter	No
P07996	Thrombospondin-1	THBS1	129.3	5.25	0.02	Hemostasis	(-) <sup>[15]</sup>
A0A1407A29	Complement C4-B	C4B	187.5	0.54	0.02	Complement system	(+) <sup>[25]</sup> ; (-) <sup>[26]</sup>
P08253	72 kDa type IV collagenase	MMP2	73.8	1.38	0.02	Degradation of extracellular matrix, regulation of vascularization, inflammatory response	(+) <sup>[27,28]</sup> ; (-) <sup>[29]</sup>
Q03591	Complement factor H-related protein 1	CFHR1	37.6	0.75	0.03	Complement system	(-) <sup>[30]</sup>
P00748	Coagulation factor XII	F12	67.7	1.29	0.03	Coagulation	(-) <sup>[31]</sup>
O95445	Apolipoprotein M	APOM	21.2	0.62	0.06	Lipid transporter activity, antioxidant activity, binding	(-) <sup>[13]</sup>
B4E1C2	Kininogen 1, isoform CRA_b	KNG1	71.9	0.78	0.28	Receptor binding, zinc ion binding, heparin binding, cysteine-type endopeptidase inhibitor activity	(-) <sup>[32]</sup>
P02745	Complement C1q subcomponent subunit A	C1QA	26	0.69	0.24	Complement system, regulation of immune system process	No

无明确的定论, 存在诸多学说, 包括淀粉样蛋白假说<sup>[34]</sup>、胆碱能假说<sup>[35]</sup>、谷氨酸假说<sup>[36]</sup>、氧化应激假说<sup>[37]</sup>和金属假说<sup>[38]</sup>等。随着AD发病机制和药物研发的深入, 部分AD发病假说受到了质疑, 但是又提出一些新的观点且在研究中也得到证实包括神经炎症<sup>[39]</sup>、补体异常激活<sup>[40]</sup>、脂质代谢紊乱<sup>[41]</sup>以及凝血系统异常<sup>[42,43]</sup>等。此外, AD发病隐匿, 在疾病早期难以发现, 且目前没有有效的治疗药物, 因此, 找到可用于AD早期诊断的生物标志物和新的药物靶点已成为当务之急。目前,

CSF和神经影像学的生物标志物已应用于AD的早期诊断<sup>[10]</sup>, 但是由于CSF获取具有损伤性和神经影像学检查的高费用, 限制了二者在临床的广泛应用。相比之下, 血液样品较易获取, 并且每天大约有500 mL CSF被吸收入血, 因此血液是寻找AD生物标记物的理想样本。此外, 在AD病理改变中存在血脑屏障破坏<sup>[44]</sup>, 因此, 在AD患者脑组织与血液之间的蛋白具有一定的相关性。

基于上述研究背景, 本研究使用血清样本进行



**Figure 4** Network analysis diagram of protein-protein interaction analysis. The protein-protein interaction networks of the 18 proteins listed in table 4 and well-known proteins related to AD were created using STRING database (Version10.5). Each node represents a protein and the links between nodes represent known or predicted functional protein-protein interactions

AD生物标志物的研究。由于血清样本成分复杂,因此能够对其进行有效的分离是血液蛋白质组学研究的关键。近年来,随着高效液相色谱和质谱技术的发展,基于质谱的蛋白质组学技术(包括iTRAQ和非标记蛋白定量)已被广泛应用于复杂临床样本的研究,例如血浆/血清、尿液和组织<sup>[45-47]</sup>。本研究使用抗体亲和柱去除血清中白蛋白和IgG后,进行非标记蛋白质组学定量分析,成功鉴定到387个蛋白,其中61个蛋白是差异性表达的,而且表达模式聚类分析结果显示这些蛋白在组内有很好的重复,组间差异明显,因此,证明了本研究所用的非标记液质联用蛋白定量方法得到的结果是相对稳定和可靠的。然而,遗憾的是,在本研究和其他类似的研究中均没有鉴别到或观察到已知的AD相关蛋白的含量改变<sup>[15]</sup>,如A $\beta$ 、Tau和ApoE等。出现这一结果的原因主要有以下两个方面:① 本研究仅去除了人血清中的两种高丰度蛋白(白蛋白和IgG),虽然可以满足label-free LC-MS/MS实验要求,但是还是会有多种低丰度蛋白被掩盖,特别是在血清中含量很低的蛋白,如A $\beta$ 和Tau蛋白等;② 由于iTRAQ和非标记蛋白定量方法本身的局限性,研究发现这两种方法对于监测微小的变化不敏感<sup>[24]</sup>。当然,在本研究中鉴

定到的部分差异蛋白在AD患者血清中的变化结果也得到了相关研究的支持<sup>[13,15,32]</sup>,因此,本研究所用的label free LC-MS/MS技术所得的阳性结果是可靠的,也从另一个方面提示AD相关的病理学改变涉及范围很广,不仅仅与A $\beta$ 和Tau蛋白有关。

本研究将血清差异蛋白质组学和生物信息学分析的结果相结合来寻找潜在的AD血液生物标志物。GO作为一种重要的生物信息学分析手段,其在基于蛋白质组学技术的AD生物标志物的研究中也得到了应用<sup>[15]</sup>。本研究采用GO富集分析来描述在AD患者血清中存在的差异蛋白中显著富集的功能条目,进而分析这些差异蛋白与AD之间是否具有相关性。研究发现,有60个GO功能条目显著富集在差异蛋白中,这60个GO条目中包括一些目前研究认为的可能与AD发病有关的生物学过程<sup>[15]</sup>,如胆固醇和脂质代谢、炎症反应、凝血和止血过程以及免疫反应等。根据GO分析结果和蛋白在8个比较组中含量变化的一致性程度,从中选取了18个参与上述生物学过程的蛋白作为AD候选生物标志物。然后将它们和已知的AD相关蛋白进行蛋白互作分析,结果显示所有蛋白均富集在蛋白互作的网络图中,且这18个蛋白都与APP相关分

子具有相互作用关系,因此,这18个蛋白可能通过与A $\beta$ 生物学过程的相互作用参与了AD发生发展。有趣的是,作者发现凝血和止血功能模块与APP连接更为紧密,有多个蛋白与APP具有直接作用的关系,如KNG1、F12、F7和THBS1等,其中F12过去已被证实可以与A $\beta$ 直接结合并被激活引起凝血系统的异常而导致AD的发生<sup>[21]</sup>。而且,本实验中KNG1和F11在AD患者血液中含量降低的结果与文献<sup>[21,32]</sup>报道是一致的,因此,上述结果提示参与凝血和止血过程的相关蛋白具有很大的潜能成为AD血液生物标志物。

但是,这18个蛋白能否成为AD有效的生物标志物仍需要使用更多的样本量和更深入的研究进行验证。此外,基于AD发病机制的复杂性,需要更多的生物标志物来反映AD的病理变化,不仅仅局限于围绕A $\beta$ 和tau蛋白的病理样改变的标志物。因此,本研究为寻找有效的AD血液生物标志物和新的药物靶点提供了一定的理论基础。

**作者贡献:** 许婷婷负责实验设计、实验实施、数据收集、结果分析和文章撰写等工作;郭鹏和张巍负责临床样本采集和临床数据收集等工作;王晓良负责研究思路的提出、实验设计、实验结果分析和文章修改等工作,为该实验的主要负责人。

**利益冲突:** 本文无利益冲突。

## References

- [1] Latypova X, Martin L. 2015: which new directions for Alzheimer's disease? [J]. *Front Cell Neurosci*, 2014, 8: 417.
- [2] Association Alzheimer's. 2016 Alzheimer's disease facts and figures [J]. *Alzheimers Dement*, 2016, 12: 459-509.
- [3] Jia J, Wang F, Wei C, et al. The prevalence of dementia in urban and rural areas of China [J]. *Alzheimers Dement*, 2014, 10: 1-9.
- [4] Jia J, Wei C, Chen S, et al. The cost of Alzheimer's disease in China and re-estimation of costs worldwide [J]. *Alzheimers Dement*, 2018, 14: 483-491.
- [5] Cummings J, Aisen PS, Dubois B, et al. Drug development in Alzheimer's disease: the path to 2025 [J]. *Alzheimer's Res Ther*, 2016, 8: 39.
- [6] Schneider LS, Sano M. Current Alzheimer's disease clinical trials: methods and placebo outcomes [J]. *Alzheimer's Dement*, 2009, 5: 388-397.
- [7] Schneider LS, Mangialasche F, Andreasen N, et al. Clinical trials and late-stage drug development for Alzheimer's disease: an appraisal from 1984 to 2014 [J]. *J Intern Med*, 2014, 275: 251-283.
- [8] Cummings JL, Morstorf T, Zhong K. Alzheimer's disease drug-development pipeline: few candidates, frequent failures [J]. *Alzheimer's Res Ther*, 2014, 6: 37.
- [9] Mapstone M, Cheema AK, Fiandaca MS, et al. Plasma phospholipids identify antecedent memory impairment in older adults [J]. *Nat Med*, 2014, 20: 415.
- [10] Jack CR, Bennett DA, Blennow K, et al. NIA-AA research framework: toward a biological definition of Alzheimer's disease [J]. *Alzheimer's Dement*, 2018, 14: 535-562.
- [11] Albert MS, Dekosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease [J]. *Alzheimer's Dement*, 2011, 7: 270-279.
- [12] Herukka SK, Simonsen AH, Andreasen N, et al. Recommendations for cerebrospinal fluid Alzheimer's disease biomarkers in the diagnostic evaluation of mild cognitive impairment [J]. *Alzheimer's Dement*, 2017, 13: 285-295.
- [13] Muenchhoff J, Poljak A, Song F, et al. Plasma protein profiling of mild cognitive impairment and Alzheimer's disease across two independent cohorts [J]. *J Alzheimer's Dis*, 2015, 43: 1355-1373.
- [14] Song F, Poljak A, Kochan NA, et al. Plasma protein profiling of mild cognitive impairment and Alzheimer's disease using iTRAQ quantitative proteomics [J]. *Proteome Sci*, 2014, 12: 5.
- [15] Shen L, Liao L, Chen C, et al. Proteomics analysis of blood serums from Alzheimer's disease patients using iTRAQ labeling technology [J]. *J Alzheimer's Dis*, 2017, 56: 361-378.
- [16] Heywood WE, Galimberti D, Bliss E, et al. Identification of novel CSF biomarkers for neurodegeneration and their validation by a high-throughput multiplexed targeted proteomic assay [J]. *Mol Neurodegener*, 2015, 10: 64.
- [17] Levin Y, Schwarz E, Wang L, et al. Label-free LC-MS/MS quantitative proteomics for large-scale biomarker discovery in complex samples [J]. *J Sep Sci*, 2007, 30: 2198-2203.
- [18] Zhu W, Smith JW, Huang CM. Mass spectrometry-based label-free quantitative proteomics [J]. *BioMed Res Int*, 2010, 2010: 840518.
- [19] Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? [J]. *J Classif*, 2014, 31: 274-295.
- [20] Gene Ontology Consortium. Gene ontology consortium: going forward [J]. *Nucleic Acids Res*, 2015, 43 (Database issue): D1049-D1056.
- [21] Zamolodchikov D, Renné T, Strickland S. The Alzheimer's disease peptide  $\beta$ -amyloid promotes thrombin generation through activation of coagulation factor XII [J]. *J Thromb Haemost*, 2016, 14: 995-1007.
- [22] Yang MH, Yang YH, Lu CY, et al. Activity-dependent neuroprotector homeobox protein: a candidate protein identified in serum as diagnostic biomarker for Alzheimer's disease [J]. *J Proteomics*, 2012, 75: 3617-3629.
- [23] Song F, Poljak A, Kochan NA, et al. Plasma protein profiling of

- mild cognitive impairment and Alzheimer's disease using iTRAQ quantitative proteomics [J]. *Proteome Sci*, 2014, 12: 5.
- [24] Muenchhoff J, Poljak A, Thalamuthu A, et al. Changes in the plasma proteome at asymptomatic and symptomatic stages of autosomal dominant Alzheimer's disease [J]. *Sci Rep*, 2016, 6: 29078.
- [25] Giometto B, Argentiero V, Sanson F, et al. Acute-phase proteins in Alzheimer's disease [J]. *Eur Neurol*, 1988, 28: 30-33.
- [26] Hye A, Lynham S, Thambisetty M, et al. Proteome-based plasma biomarkers for Alzheimer's disease [J]. *Brain*, 2006, 129: 3042-3050.
- [27] Doecke JD, Laws SM, Faux NG, et al. Blood-based protein biomarkers for diagnosis of Alzheimer's disease [J]. *Arch Neurol*, 2012, 69: 1318-1325.
- [28] Duits FH, Hernandez-Guillamon M, Montaner J, et al. Matrix metalloproteinases in Alzheimer's disease and concurrent cerebral microbleeds [J]. *J Alzheimer's Dis*, 2015, 48: 711-720.
- [29] Lim NK, Villemagne VL, Soon CP, et al. Investigation of matrix metalloproteinases, MMP-2 and MMP-9, in plasma reveals a decrease of MMP-2 in Alzheimer's disease [J]. *J Alzheimer's Dis*, 2011, 26: 779-786.
- [30] Ashton NJ, Kiddle SJ, Graf J, et al. Blood protein predictors of brain amyloid for enrichment in clinical trials? [J]. *Alzheimer's Dement (Amst)*, 2015, 1: 48-60.
- [31] Zamolodchikov D, Chen ZL, Conti BA, et al. Activation of the factor XII-driven contact system in Alzheimer's disease patient and mouse model plasma [J]. *Proc Natl Acad Sci U S A*, 2015, 112: 4068-4073.
- [32] Kitamura Y, Usami R, Ichihara S, et al. Plasma protein profiling for potential biomarkers in the early diagnosis of Alzheimer's disease [J]. *Neurol Res*, 2017, 39: 231-238.
- [33] Parihar MS, Hemnani T. Alzheimer's disease pathogenesis and therapeutic interventions [J]. *J Clin Neurosci*, 2004, 11: 456-467.
- [34] Goedert M, Spillantini MG. A century of Alzheimer's disease [J]. *Science*, 2006, 314: 777-781.
- [35] Craig LA, Hong NS, McDonald RJ. Revisiting the cholinergic hypothesis in the development of Alzheimer's disease [J]. *Neurosci Biobehav Rev*, 2011, 35: 1397-1409.
- [36] Bezprozvanny I, Mattson MP. Neuronal calcium mishandling and the pathogenesis of Alzheimer's disease [J]. *Trends Neurosci*, 2008, 31: 454-463.
- [37] Pratico D. Oxidative stress hypothesis in Alzheimer's disease: a reappraisal [J]. *Trends Pharmacol Sci*, 2008, 29: 609-615.
- [38] Bonda DJ, Lee HG, Blair JA, et al. Role of metal dyshomeostasis in Alzheimer's disease [J]. *Metallomics*, 2011, 3: 267-270.
- [39] van Eldik LJ, Carrillo MC, Cole PE, et al. The roles of inflammation and immune mechanisms in Alzheimer's disease [J]. *Alzheimer's Dement (N Y)*, 2016, 2: 99-109.
- [40] Eikelenboom P, Van Exel E, Veerhuis R, et al. Innate immunity and the etiology of late-onset Alzheimer's disease [J]. *Neurodegener Dis*, 2012, 10: 271-273.
- [41] Wong MW, Braidy N, Poljak A, et al. Dysregulation of lipids in Alzheimer's disease and their role as potential biomarkers [J]. *Alzheimer's Dement*, 2017, 13: 810-827.
- [42] Cortes-Canteli M, Zamolodchikov D, Ahn HJ, et al. Fibrinogen and altered hemostasis in Alzheimer's disease [J]. *J Alzheimer's Dis*, 2012, 32: 599-608.
- [43] Jarre A, Gowert NS, Donner L, et al. Pre-activated blood platelets and a pro-thrombotic phenotype in APP23 mice modeling Alzheimer's disease [J]. *Cell Signal*, 2014, 26: 2040-2050.
- [44] Yamazaki Y, Kanekiyo T. Blood-brain barrier dysfunction and the pathogenesis of Alzheimer's disease [J]. *Int J Mol Sci*, 2017, 18: 1965.
- [45] Choudhary C, Mann M. Decoding signalling networks by mass spectrometry-based proteomics [J]. *Nat Rev Mol Cell Biol*, 2010, 11: 427.
- [46] Xie F, Liu T, Qian WJ, et al. Liquid chromatography-mass spectrometry-based quantitative proteomics [J]. *J Biol Chem*, 2011, 286: 25443-25449.
- [47] Wang H, Shi T, Qian WJ, et al. The clinical impact of recent advances in LC-MS for cancer biomarker discovery and verification [J]. *Expert Rev Proteomics*, 2016, 13: 99-114.