

[文章编号] 1000-4718(2009)02-0389-03

# SELDI-TOF-MS 技术检测心绞痛患者血浆蛋白指纹图谱的研究\*

何文娟<sup>1</sup>, 胡豫<sup>1△</sup>, 张小平<sup>1</sup>, 余捷凯<sup>2</sup>, 洪梅<sup>1</sup>, 郭涛<sup>1</sup>, 宋善俊<sup>1</sup>

(<sup>1</sup>华中科技大学同济医学院协和医院血液病研究所,湖北省生物靶向治疗重点实验室,湖北 武汉 430022;<sup>2</sup>浙江大学附属第二医院肿瘤研究所,浙江 杭州 310009)

**[摘要]** 目的:采用表面增强激光解析/电离飞行时间质谱(SELDI-TOF-MS)技术检测心绞痛患者血浆蛋白质指纹图谱,从中筛选出特异的分子标志物。方法:应用 SELDI-TOF-MS 技术及金属离子螯合(IMAC-3)蛋白质芯片对 20 例心绞痛患者和 29 例正常对照者血浆样本进行检测,借助生物信息学工具(非线性的支持向量机,SVM)提出心绞痛的诊断模型,并运用留一交叉验证法来评估该模型的判别效能。结果:用 SELDI-TOF-MS 技术筛选出由 3 个有显著差异的蛋白质峰[质荷比( $m/z$ )分别为 2 667.3、5 914.0 和 6 890.5]组合构建的诊断模型,可将 20 例心绞痛患者和 29 例正常人全部正确分组,诊断特异性和灵敏度均为 100%。结论:SELDI-TOF-MS 技术在心绞痛的诊断中具有较高灵敏度和特异性,发现的蛋白质峰可能在心绞痛的发病中起一定作用,血浆中分子标志物的发现有助于心绞痛的早期诊断。

**[关键词]** 心绞痛;表面增强激光解吸/电离飞行时间质谱;蛋白质组

**[KEY WORDS]** Angina pectoris; Surface-enhanced laser desorption/ionization time-of-flight; Proteome

**[中图分类号]** R541.4 **[文献标识码]** A

表面增强激光解析/电离飞行时间质谱(surface-enhanced laser desorption/ionization time-of-flight mass spectrometry, SELDI-TOF-MS)技术是一种新兴的蛋白组学研究手段,它具有检测时间短、灵敏度高等特点<sup>[1]</sup>,分析蛋白质的改变不仅可直接获得功能相关的标志物,而且还适用于标准化和临床应用<sup>[2]</sup>,现已应用于对疾病生物标志物的分析中<sup>[3-5]</sup>,然而,这一技术很少用于对心血管疾病特异性蛋白质的检测。本研究采用 SELDI 仪及金属离子螯合(metal-ion affinity chromatography, IMAC-3)蛋白质芯片检测心绞痛患者和正常对照者血浆蛋白质指纹图谱,筛选出蛋白标志物并建立诊断模型。

## 材料和方法

### 1 材料

**1.1 仪器和芯片** PBS-II 型 SELDI-TOF-MS 系统和 IMAC-3 蛋白芯片为 CIPHERGEN Biosystems 产品;芥子酸为 Fluka 产品;其它试剂均为 Sigma 产品。

**1.2 对象** 病例组共 20 例,均为 2006 年 1-3 月在武汉协和医院心内科住院的心绞痛患者,年龄 41-78 岁,平均 59 岁,男 14 例,女 6 例。诊断均符合 2000 年中华医学会心血管病学分会心绞痛诊断和指南的诊断标准。对照组共 29 例,均为同期在该院体检的健康人,年龄 43-78 岁,平均 56 岁,男 16 例,女 13 例。病例组与对照组在采集时间上相匹配。在血液标本采集前均未接受抗凝剂及纤维蛋白溶解药物治疗。

**1.3 标本收集** 在患者入院后 24 h 内采集外周血标本于

3.2% 枸橼酸三钠抗凝管(抗凝剂:全血为 1:9)中并混匀,采集后 30 min 内将标本在 4℃ 下以 3 500 r/min 离心 6 min,将上层血浆分装于已消毒的 EP 管中,保存于 -80℃ 冰箱,待测。

### 2 方法

**2.1 标本准备** 血浆标本在冰上解冻 30 min 后,以 10 000 r/min 转速于 4℃ 离心 2 min,取 5 μL 样本分别加入 96 孔细胞培养板上,用 10 μL 尿素缓冲液(9 mol/L 尿素,2% chaps,1% DTT)稀释后,4℃ 600 r/min 摇床 30 min;快速加入 185 μL 醋酸钠(pH 4.0),4℃ 600 r/min 摇床 2 min。

**2.2 蛋白芯片的预处理** 将 100 mmol/L 硫酸铜 50 μL 加到 8 孔 IMAC-3 芯片池内,室温 600 r/min 摇床 10 min,甩干后加 200 μL 去离子水,室温 600 r/min 摇床 1 min,甩干后加 100 mmol/L 醋酸钠(pH 4.0) 200 μL,室温 600 r/min 摇床 5 min,甩干后加 200 μL 去离子水室温 600 r/min 摇床 1 min,甩干;加 200 μL 磷酸盐缓冲液(PBS, pH 7.2),室温 600 r/min 摇床 5 min,甩干后重复 1 次。

**2.3 上样** 向已置入 Bioprocessor (CIPHERGEN Biosystems Inc.)的蛋白质芯片池内加入 100 μL 血浆稀释样本,4℃ 600 r/min 摇床 1 h;甩掉未与芯片结合的血浆混合物后,用 100 mmol/L 醋酸钠(pH 4.0) 200 μL 漂洗 3 次,去离子水 200 μL 漂洗 2 次;最后,从 Bioprocessor 中取出芯片,室温下晾干。在 SELDI-TOF-MS 分析之前,向蛋白质芯片各孔加入 1 μL 半饱和的 SPA(饱和 SPA 溶于 50% 乙腈和 0.5% 三氟乙酸, Sigma),5 min 后重复 1 次,室温下晾干。

**2.4 芯片检测** 由 Protein Biology System (PBS, CIPHERGEN Biosystems Inc.) IIc 型质谱仪处理,相应参数设定如下:激光

[收稿日期] 2007-11-12 [修回日期] 2008-10-15

\* [基金项目] 国家自然科学基金资助项目(No. 30700332);卫生部重点资助项目(No. 2004468)

△ 通讯作者 Tel: 027-85726307; E-mail: hwjsun@163.com

强度为 210,检测器灵敏度为 9,最大分子量为 100 kD,最佳范围是 2 - 20 kD。收集数据前用标准蛋白质芯片校正仪器。所有芯片的检测和分析均在同 1 d 完成。

3 生物信息学和生物统计学分析

所有实验数据(数据预处理和模型建立)均由浙江大学肿瘤中心蛋白芯片数据分析系统(ZUCI - PDAS)进行分析。将原始数据用平移不变小波变换(undecimated discrete wavelet transform, UDWT)去除噪音,并减掉基线。用局部极值的方法找出样本各自的蛋白峰,过滤掉  $m/z$  在 2 000 以下及信噪比(signal - to - noise ratio, S/N)小于 3 的峰。这时从每个样本中找出来的峰的个数及其  $m/z$  都不同。将各样本中  $m/z$  的差异小于 0.3% 的峰聚为一类,并去掉聚类后只出现在少于 10% 样本中的峰。对找出的峰在各个样本中的强度做均一化处理。分别用非线性支持向量机(nonlinear support vector machine, SVM)方法建立判别模型<sup>[6]</sup>,用留一交叉验证法(Leave - one - out cross validation)作为评估模型判别效能的方法。SVM 采用径向基核函数(radial based kernel),Gamma 值设为 0.6,惩罚函数(C)设为 19。特征向量的选取采用统计过滤结合模型依赖性筛选的方法。对每个质荷比峰做 Wilcoxon 秩和检验,选出  $P$  值最小的 10 个峰作进一步分析。将 10 个峰的任意组合用于 SVM 模型的输入,用留一法评估各模型的预测效果,选出建立 SVM 模型预测的约登指数(Youden's index)最高的组合作为最终的候选标志物,建立的模型和留一法交叉验证的结果作为最终的结果。

结 果

1 IMAC - 3 蛋白质芯片的重复性

将来自同一健康者的 10 份血浆稀释样本以随机的方式加到已预处理的 IMAC - 3 芯片中。所选蛋白峰( $m/z$ : 4 583、6 640 和 11 734)的峰值及标准化强度的变异系数(coefficient of variation, CV)分别是 0.03% 和 18%。

2 差异蛋白质质荷比峰的筛选

对 20 例心绞痛患者和 29 例正常对照者血浆蛋白质指纹图谱数据进行比较,经分析后检测到 289 个蛋白质荷比峰,其中 27 个蛋白质荷比峰具有极显著差异,见表 1。根据 Wilcoxon 秩和检验结果,将  $P$  值最小的 10 个蛋白峰任意组合用于 SVM 模型的输入后,得出了检测效率最高的由其中 3 个蛋白质峰(质荷比分别为 2 667.3、5 914.0 和 6 890.5)组合构建的诊断模型。由留一交叉验证法评估此模型的判别效果,其区分心绞痛患者与健康对照者的敏感性和特异性均为 100% (即准确率为 100%)。与正常人比较,这 3 个峰在心绞痛患者中均为低表达,见图 1。

3 ROC 曲线(receiver operator characteristic curves)下面积

诊断模型中质荷比分别为 5 914.0、2 667.3 和 6 890.5 的蛋白质的 ROC 曲线下面积分别为 1.000(95% CI 1.000 - 1.000)、0.947(95% CI 0.889 - 1.004)和 0.929(95% CI 0.857 - 1.002),见图 2。

表 1 心绞痛组和对照组差异蛋白质峰的强度比较

Tab 1 Descriptive statistics of peaks making up a differential pattern( $\bar{x} \pm s$ )

| Mass - charge ratio of protein( $m/z$ ) | The average intensity of protein peak in AP cases( $n = 20$ ) | The average intensity of protein peak in controls( $n = 29$ ) | $P$ value    |
|---|---|---|--------------|
| 5 914                                   | 4 541.66 ± 2 008.33   | 15 789.44 ± 3 501.28  | 0.0000000039 |
| 2 667                                   | 1 761.87 ± 973.98   | 3 499.08 ± 1 221.85   | 0.0000040765 |
| 6 891                                   | 1 879.65 ± 1 054.55   | 5 520.44 ± 3321.69  | 0.0000040765 |
| 6 121                                   | 879.50 ± 257.85   | 1 985.72 ± 869.30   | 0.0000000743 |
| 5 931                                   | 1 989.38 ± 757.61   | 3 958.53 ± 1 139.23   | 0.0000002021 |
| 5 957                                   | 1 233.31 ± 286.37   | 2 038.49 ± 512.06   | 0.0000002802 |
| 8 939                                   | 1 518.91 ± 856.39   | 3 703.70 ± 1 658.27   | 0.0000004788 |
| 6 136                                   | 862.66 ± 316.89   | 1 383.45 ± 399.05   | 0.0000033497 |
| 5 262                                   | 1 025.99 ± 221.24   | 1 638.32 ± 473.97   | 0.0000049533 |
| 6 099                                   | 856.20 ± 217.90   | 1 275.31 ± 317.71   | 0.0000049533 |
| 8 654                                   | 1 315.86 ± 865.25   | 2 915.11 ± 1 158.85   | 0.0000088012 |
| 8 629                                   | 3 521.79 ± 2 660.43   | 8 586.11 ± 4 409.23   | 0.0000116691 |
| 8 610                                   | 10 105.63 ± 7 717.38  | 20 894.25 ± 4 966.23  | 0.0000128092 |
| 6 184                                   | 847.35 ± 272.50   | 2 238.60 ± 2 067.22   | 0.0000202942 |
| 4 310                                   | 2 613.64 ± 2 115.53   | 5 806.00 ± 2 646.58   | 0.0000588123 |
| 4 751                                   | 1 888.17 ± 788.65   | 1 147.86 ± 352.11   | 0.0000760557 |
| 6 820                                   | 1 925.40 ± 559.59   | 4 035.59 ± 2 349.18   | 0.0000900987 |
| 4 677                                   | 1 581.95 ± 428.31   | 1 044.70 ± 402.30   | 0.0001158285 |
| 8 817                                   | 1 277.84 ± 903.10   | 2 464.02 ± 1 239.07   | 0.0001258460 |
| 8 566                                   | 888.39 ± 603.13   | 1 574.40 ± 759.02   | 0.0001483799 |
| 8 713                                   | 1 081.10 ± 496.60   | 1 761.91 ± 613.29   | 0.0001483799 |
| 6 905                                   | 1 586.19 ± 489.70   | 2 649.59 ± 1 112.70   | 0.0001746740 |
| 5 274                                   | 1 046.01 ± 253.81   | 1 431.43 ± 359.47   | 0.0002608431 |
| 9 664                                   | 1 281.76 ± 347.78   | 958.36 ± 210.69   | 0.0002822923 |
| 4 702                                   | 1 463.58 ± 497.83   | 1 032.43 ± 310.32   | 0.0006561640 |
| 9 696                                   | 1 363.53 ± 412.90   | 1 069.08 ± 237.98   | 0.0006561640 |
| 6 558                                   | 1 044.15 ± 323.74   | 1 927.56 ± 1 305.41   | 0.0007068014 |

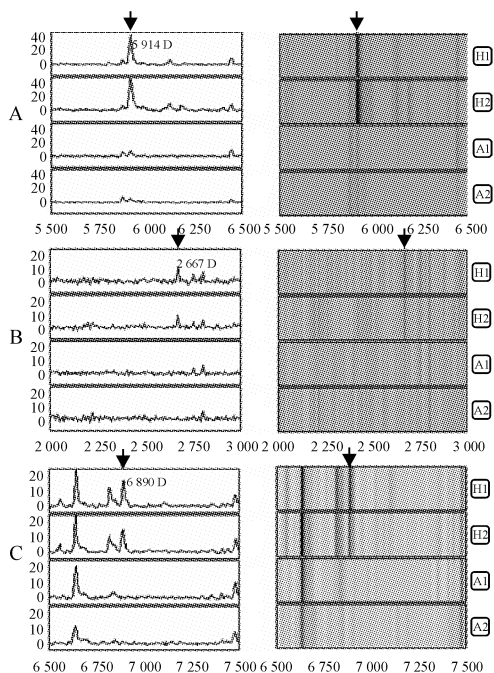


Fig 1 Representative spectra (left panel) and pseudo - gel views (right panel) of the candidate biomarkers from AP (A1 - 2) vs controls (H1 - 2). X - axis represents the molecular mass calculation ( $m/z$  values), while y - axis represents relative intensity. A, B, and C indicate different expressions of the plasma markers with  $m/z$  of 5 914. 0D, 2 667. 3D and 6 890. 5D, respectively.

图 1 心绞痛组和对照组蛋白质谱(左侧)与模拟凝胶图(右侧)

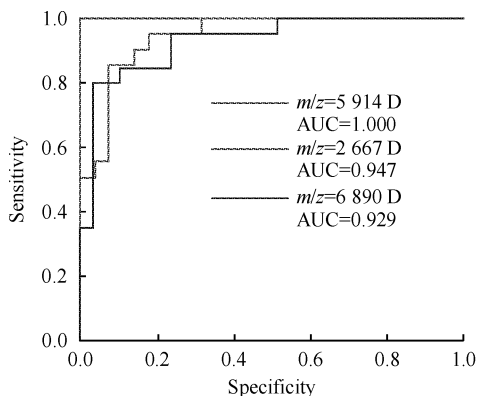


Fig 2 ROC curves of proteomic features at  $m/z$  of 5 914. 0 D (red), 2 667. 3 D (green) and 6 890. 5 D (blue) for differentiating patients with AP from controls. AUC; area under the curve.

图 2 质荷比为 5 914.0D (红)、2 667.3D (绿) 和 6 890.5D (蓝) 的蛋白质 ROC 曲线

### 讨 论

急性心肌梗死是发病率和死亡率较高的疾病之一,且近年来发病呈现年轻化趋势<sup>[7]</sup>,在急性心肌梗死发生之前,绝大多数患者已经存在心绞痛症状,因此,心绞痛的及时诊断和治疗对防止心肌梗死的发生和疾病相关死亡率有着至关重要的作用。目前,心绞痛的诊断主要依靠心肌缺血的症状和体征、心电图检查以及冠状动脉造影等<sup>[8]</sup>,后者是诊断冠心病的重要方法,然而,由于其侵入性、费用高、禁忌症多以

及复杂的操作过程等缺点,不能作为常规的检查方法,因此,寻找兼备高敏感、高特异性的疾病分子标志物的工作显得尤为重要。SELDI - TOF - MS 通过检测患者和正常人血浆中的蛋白质指纹图谱并进行比较,有望从中找出具有差异的分子标志物,建立相应的诊断模型,并且有可能找到对疾病发生具有促进或延缓作用的蛋白质分子。

本研究运用 SELDI - TOF - MS 及蛋白质芯片技术比较健康者和心绞痛患者血浆蛋白质表达之间的差异,发现具有极显著差异的蛋白质峰共 27 个,并确立了能够完全区分健康对照和心绞痛患者的蛋白质峰组合建立的诊断模型,具有一定的诊断价值。同时,我们还发现,建立诊断模型的 3 种蛋白质峰在心绞痛患者中均为低表达,我们初步推测,这 3 种蛋白质可能对冠状动脉具有保护作用或能够延缓粥样斑块的形成和进展。

此外,我们还发现,在 3 种蛋白质峰的 ROC 曲线下面积中, $m/z$  为 5 914. 0 的蛋白质峰达到了 1,提示这种蛋白质可能对心绞痛最具诊断价值。但是,由于建立诊断模型的基本思路是选择具有最佳诊断效力的蛋白质峰组合而不是评价单个蛋白质峰的诊断效力<sup>[9]</sup>,因此,我们更应该重视蛋白质峰组合的诊断价值。

由于本研究中所纳入的病例数较少,因此尚未在新的样本中验证这一诊断模型,这正是本研究中的一个不足之处。即使该模型能够完全区分前述 49 例样本,我们仍将扩大样本量,在不同人群中对模型的有效性进行进一步验证。

综上所述,通过 SELDI - TOF - MS 技术建立的诊断模型对于心绞痛的诊断具有一定的价值,我们将扩大样本量对这一诊断模型的可靠性进行进一步的验证,对差异蛋白进行鉴定,并对它们在正常人和患者血液中的作用进行深入的探索和研究。

### [参 考 文 献]

- [1] 廖秋林,陈晓东,丁彦青. 血清蛋白质组学技术及研究进展[J]. 中国病理生理杂志,2008,24(2):402 - 407.
- [2] Rodland KD. Proteomics and cancer diagnosis: the potential of mass spectrometry [J]. Clin Biochem, 2004, 37(7): 579 - 583.
- [3] Poon TC, Yip TT, Chan AT, et al. Comprehensive proteomic profiling identifies serum proteomic signatures for detection of hepatocellular carcinoma and its subtypes[J]. Clin Chem, 2003, 49(5): 752 - 760.
- [4] Grus FH, Podust VN, Bruns K, et al. SELDI - TOF - MS proteinchip array profiling of tears from patients with dry eye [J]. Invest Ophthalmol Vis Sci, 2005, 46(3): 863 - 876.
- [5] 王全晖,高春芳,王秀丽,等. 利用 SELDI - TOF 质谱技术分析大肠癌患者血清蛋白质谱的变化[J]. 中国病理生理杂志,2005,21(10):1896 - 1900.
- [6] Liu Y. Active learning with support vector machine applied to gene expression data for cancer classification[J]. J Chem Inf Comput Sci, 2004, 44(6): 1936 - 1941.
- [7] 杨旭玲,裴晓冬. 心肌肌钙蛋白 T 在急性冠脉综合症诊断中的意义[J]. 中国实验诊断学,2007,11(1):23 - 24.
- [8] 陈灏珠. 冠状动脉粥样硬化性心脏病[A]. 见:叶任高,陆再英 主编. 内科学[M]. 第 6 版. 北京:人民卫生出版社,2004. 272 - 299.
- [9] 胡 跃,张苏展,余捷凯,等. 乳腺癌患者血清蛋白质指纹图谱检测及其临床意义[J]. 中华检验医学杂志, 2004,27(10):646 - 648.