

· 论 著 ·

食管癌血清差异蛋白质组测定结果分析

谭小林¹, 任兴军², 王开正^{1△}, 杨成虎², 冯 峰², 胡琼英¹, 丁银环¹, 戴天阳¹

(1. 泸州医学院附属医院检验科, 四川泸州 646000; 2. 四川省射洪县人民医院 629200)

摘要:目的 用表面增强激光解析电离飞行时间质谱技术(SELDI-TOF-MS)寻找食管癌(EC)患者血清中差异蛋白质组,并建立诊断模型,为其广泛用于临床诊断积累数据资料。方法 用 SELDI 芯片检测食管癌及其相关人群的血清, Biomarker Wizard Software 软件筛选出差异蛋白,选择具有显著差异的蛋白质组建立人工神经网络诊断模型,经盲法验证后使用 SPSS 软件分析各个差异蛋白以及联合人工神经网络技术后的诊断效率,并鉴定相关蛋白。结果 筛选出显著差异($P < 0.01$)的蛋白质组如下:在食管癌患者血清中表达增高蛋白组[(5 017.6±4.89)、(7 458.5±6.49)、(7 908.1±7.80)、(8 111±8.45)、(8 577±7.80) kD],表达降低蛋白组[(7 748.3±9.10)、(5 890.9±7.32)、(4 213.8±5.93) kD]。差异蛋白质组建立食管癌人工神经网络筛查模型和诊断模型,其灵敏度和特异性超过 90%,盲法验证的效果好;人工神经网络预测输出值构成的 ROC 曲线下面积分别为 81.6%和 82.1%,经查阅数据库鉴定发现两种蛋白质。结论 血清蛋白质谱指纹图结合人工神经网络技术进行蛋白质组学数据挖掘所建模型诊断价值较大,与临床诊断符合程度良好。

关键词:食管肿瘤;质谱分析法;蛋白质组学;神经网络计算机;诊断模型

doi:10.3969/j.issn.1671-8348.2011.07.007

文献标识码:A

文章编号:1671-8348(2011)07-0641-03

Analysis of detection results of different serum proteomes in esophageal cancer

Tan Xiaolin¹, Ren Xingjun², Wang Kaizheng^{1△}, Yang Chenghu², Feng Feng²,
Hu Qiongying¹, Ding Yinhan¹, Dai Tianyang¹

(1. Department of Laboratory Medicine, Affiliated Hospital of Luzhou Medical College, Luzhou, Sichuan 646000, China;

2. Shehong County People's Hospital, Shehong, Sichuan 629200, China)

Abstract: **Objective** To find differential serum proteomes in esophageal cancer by enhanced laser desorption ionization time of flight mass spectrometry (SELDI-TOF-MS) to establish diagnostic models for accumulating the data information for wide use in clinical diagnostic. **Methods** To detect the serum of the patients with esophageal cancer and related groups by SELDI chip, filter out different proteomes with the Biomarker Wizard Software. Then to develop the artificial neural network diagnostic models using significant different proteomes and use SPSS software to analyze diagnostic efficiency of significant difference in proteomes and the combined artificial neural network after verified by the blind. Finally, to identify related proteins. **Results** The selected different proteomes ($P < 0.01$) were as follows. In esophageal cancer, the intensity increasing proteomes were 5 017.6±4.89, 7 458.5±6.49, 7 908.1±7.80, 8 111±8.45, 8 577±7.80 kD, lowerly expressed proteomes were 7 748.3±9.10, 5 890.9±7.32, 4 213.8±5.93 kD. Built the screening and diagnosis models of artificial neural network by different proteomes of esophageal cancer. Sensitivity and specificity were over 90%. After blinding test, they were effective. The area under the ROC curve consisting by the output value of artificial neural network were 81.6% and 82.1%, searching the related databases, two proteins were found. **Conclusion** The value of diagnostic models are useful by associating serum protein fingerprint spectrums with artificial neural network technology in proteomics data mining and are accord with the clinical diagnosis.

Key words: esophageal neoplasms; mass spectrometry; proteomics; neural networks (computer); diagnosis model

食管癌(esophageal carcinoma, EC)是我国常见恶性肿瘤之一,本研究使用表面增强激光解析电离飞行时间质谱技术(SELDI-TOF-MS)结合蛋白芯片筛选差异蛋白,建立诊断模型,并进行盲法验证,结合 SPSS 分析其诊断效果和临床应用价值。

1 资料与方法

1.1 一般资料 本研究分为 4 组:食管癌组 302 例,年龄 20~82 岁,男女比例为 1.92:1;健康对照组 203 例,年龄 18~74 岁,男女比例为 1.73:1;其他疾病组 274 例,年龄 17~76 岁,男女比例为 1.81:1;其他恶性肿瘤组 338 例,年龄 15~84 岁,男女比例为 1.77:1。其他疾病组为胃、肠、肺良性病变,脑损伤、淋巴结炎、精神分裂症等患者。其他恶性肿瘤组包括肝癌、肠癌、肺癌、胃癌、胰腺癌等患者。所有患者均为泸州医学院附

属医院和射洪县人民医院 2008 年 7 月至 2009 年 11 月门诊或住院病例;所有癌症患者均于术前空腹无菌采取静脉血,均于术后病理学确诊且在采血之前未经任何放、化疗治疗,其他疾病组均于治疗前采集标本并排除了肝炎、糖尿病、结缔组织病等易使血清中产生干扰蛋白的疾病。健康对照组(为两个医院健康体检者),每例标本均做常规生化检测,肝、肾功能,血脂、血糖均在正常范围内。所有实验样本无溶血。所有入选对象在性别和年龄等方面比较差异无统计学意义。

1.2 方法

1.2.1 主要仪器、样品准备和实验步骤 蛋白指纹图谱分析仪(PBS II/C, 美国),水(HPLC 级),三氟乙酸,SPA,乙腈(HPLC 级),PBS(磷酸盐缓冲液)均购自美国 Sigma 公司。96

表 1 标志性差异蛋白峰相对高度($\bar{x} \pm s$, kD)

| 组别 | 4 213.8 | 5 017.6 | 5 890.9 | 7 458.5 | 7 748.3 | 7 908.1 | 8 111.9 | 8 577.8 |
|---------|------------|-------------|-------------|------------|--------------|--------------|--------------|--------------|
| 食管癌* | 1.44±2.48▼ | 10.59±8.94▲ | 0.40±0.82▼ | 9.52±7.72▲ | 14.29±10.00▼ | 19.98±12.17▲ | 34.42±12.08▲ | 24.78±13.12▲ |
| 其他恶性肿瘤组 | 5.21±3.31 | 1.53±2.10 | 28.02±14.69 | 0.89±1.03 | 26.10±11.37 | 2.94±2.01 | 9.74±7.76 | 11.51±10.88 |
| 其他疾病组 | 8.53±5.10 | 2.70±3.13 | 24.62±17.39 | 1.70±1.38 | 28.88±12.67 | 2.21±1.29 | 8.49±7.41 | 9.90±10.89 |
| 健康对照组 | 11.06±5.23 | 2.45±2.69 | 31.67±15.41 | 1.79±1.27 | 30.70±9.68 | 2.94±1.64 | 15.21±8.71 | 13.14±9.81 |

*:对4组疾病每个差异蛋白分别进行方差分析,结果显示各个差异蛋白均 $P<0.01$ 。▼:在食管癌组表达下降的差异蛋白;▲:在食管癌组表达增高的差异蛋白。

孔蛋白质芯片工作平台、Au蛋白芯片等。血清样品准备:抽取清晨空腹前臂静脉全血3 mL。采血前2 h内禁止输入蛋白类药物,禁止大量食入蛋白食品,禁止剧烈运动。采血于普通试管(不抗凝),避免溶血,置4℃冰箱中1~2 h,然后4℃、3 000 r/min离心10 min。取上清液加入Eppendorf管。再次3 000 r/min离心5 min,Eppendorf管分装,放入-80℃保存。实验步骤:(1)预处理芯片;(2)血清样品处理;(3)点样;(4)点加能量吸收分子,用蛋白指纹图谱分析仪(PBS II/C)检测。

1.2.2 SELDI质谱仪参数设置、校准、原始数据采集和输出以及质量控制 用已知相对分子质量的标准蛋白质芯片做好内标。将结合好蛋白质的芯片放入质谱仪中进行检测。原始数据由ProteinChip Software3.2软件采集,所有样本均使用相同的参数设置,严格按照预定操作步骤进行。同时在各个批次内和各批间检测预先收集的混合血清,控制批内和批间差异。

1.2.3 数据处理、建立模型 用ProteinChip Software3.2软件收集质谱信号,Biomarker Wizard Software软件筛选差异蛋白。然后用神经网络软件office插件和SPSS11.5软件对处理后的数据进行分析。建立模型时随机选取123例食管癌和62例健康者建立筛查模型;随机选取食管癌123例、其它肿瘤27例和其他疾病60例建立诊断模型。未参与建模病例作为验证组标本对所建筛查模型和诊断模型进行盲法验证。运用共轭梯度学习方法改善的反向传播算法建立模型。筛查模型采用3层结构,输入层含8个神经元,隐含层含5个神经元输出层含1个神经元;临床诊断值:食管癌设为1,健康者设为0。诊断模型输入层含8个神经元,隐含层含5个神经元,输出层含1个神经元;临床诊断值:食管癌设为1,其他疾病设为0,作为输出层的1个节点。训练次数设定为1 000,学习率为0.02,加入动量计算法。

1.2.4 模型盲法验证、评估诊断效果以及鉴定差异蛋白 用未参与建模的标本对食管癌筛查模型和诊断模型进行盲法验证做出预测结果。分析单个差异蛋白诊断效果和用神经网络软件联合各个差异蛋白后的诊断效果。结合患者临床和病理诊断并使用SPSS11.5软件分析盲法验证诊断效果。查阅相关数据库鉴定蛋白。

2 结 果

2.1 实验的重复性 本研究建立了SELDI-TOF-MS二维内标校准方法^[1],采用合成多肽分子作为内标物,建立内标校准法以校准SELDI-TOF-MS质谱图的横坐标和纵坐标,以提高该方法检测蛋白质相对分子质量和丰度的重复性。

2.2 差异蛋白的检出 用Biomarker Wizard软件对所做血清图谱进行分析,差异蛋白分布在如下范围之内:(5 017.6±4.89)、(7 458.5±6.49)、(7 908.1±7.80)、(8 111±8.45)、(8 577±7.80)、(7 748.3±9.10)、(5 890.9±7.32)、(4 213.8±5.93)kD,排除批内和批间差异较大的蛋白后,选出其中的8个蛋白(P 值均小于0.01)。8个标志性差异蛋白在

各组中的相对强度见表1。

2.3 筛查模型和诊断模型的诊断效率 使用神经网络软件结合前述差异蛋白建立的筛查模型可以将健康者与食管癌患者准确的分组,灵敏度为93.2%,特异度为95.6%。诊断模型能够明确的将食管癌从各组疾病中区分出,其灵敏度为96.3%,特异度为97.2%。

2.4 两个模型经盲法验证后各个差异蛋白诊断效率 用未参与建模标本对食管癌筛查和诊断模型进行验证。盲法验证后差异蛋白强度使用SPSS软件作图。降低组各个差异蛋白(A:峰4 213,B:峰5 890,C:峰7 748)强度ROC曲线,见插图1、2;增高组各个差异蛋白(D:峰5 017,E:峰7 458,F:峰7 908,G:峰8 111,H:峰8 577)强度ROC曲线,见插图3、4,各个差异蛋白在筛查和诊断模型的ROC曲线下面积(AUC)见表2。

表 2 筛查模型和诊断模型盲法验证后的 ROC AUC 表

| 项目 | 峰值 | 筛查模型 AUC | 诊断模型 AUC |
|---------|----------|----------|----------|
| 食管癌降低蛋白 | 4 213(A) | 0.621 | 0.608 |
| | 5 890(B) | 0.701 | 0.757 |
| | 7 748(C) | 0.538 | 0.673 |
| 食管癌增高蛋白 | 5 017(D) | 0.580 | 0.576 |
| | 7 458(E) | 0.715 | 0.700 |
| | 7 908(F) | 0.695 | 0.667 |
| | 8 111(G) | 0.803 | 0.766 |
| | 8 577(H) | 0.638 | 0.611 |

2.5 盲法验证后联合神经网络软件的两个模型诊断效率 筛查和诊断模型经盲法验证并联合神经网络软件后的预测输出值构成的ROC AUC分别为0.816和0.821。

2.6 查阅蛋白数据库鉴定结果 经查阅蛋白数据库,找到两种与之对应的蛋白质:7 908 kD对应的为子宫珠蛋白(utero-globin,UG),相对分子质量为7 915kD;8 577 kD对应的为淀粉样蛋白-A,相对分子质量为8 575 kD。

3 讨 论

食管癌是中国常见恶性肿瘤,以鳞癌多见,本研究仅以此为研究对象。对食管癌采用传统方法镜下染色^[2]和CT检查等有一定的诊断价值,但诊断率较低,待患者出现典型症状后多为中、晚期。早期手术后的5年生存率高达90%,中、晚期手术切除的生存率一直较低(20%左右),因此早诊断、早治疗是防治该病的关键^[3]。质谱技术已经广泛用于肿瘤的临床研究。本实验室建立的内标校准法提高了SELDI-TOF-MS检测中的重复性。食管癌的差异蛋白的研究一直是国内外研究的热点。国内冯笑山等^[4]、王士杰等^[5]、徐淑永等^[6]报道了与本研究筛选结果相近相对分子质量的食管癌差异蛋白。许多国

内外学者采用 SELDI 技术筛选出差异蛋白,并建立了效果较好的模型。黄颖峰^[7]和程真珍^[8]使用该技术建立了灵敏度、特异度、盲筛准确度较高的模型。胡琼英等^[9]使用 SELDI 技术结合人工神经网络,发现其用于胰腺癌的差异蛋白筛选,建立癌症的早期诊断模型方面有很好的效果。然而目前大多数研究所建立的诊断模型中的样本例数不大,且使用数据挖掘软件不统一,造成模型可信度不一致。本研究将盲法验证后的各个差异蛋白经过 SPSS11.5 软件分析其诊断效率后,发现各个差异蛋白强度经过 ANN 联合处理后的诊断效果(AUC 分别为 0.816 和 0.821)明显高于单独使用各个差异蛋白的诊断效果。这可能与神经网络本身的数据挖掘的优越性有关。神经网络根据数据本身的内在联系进行建模,其抗干扰能力、适应性、自学习能力很好。然而经过筛选的差异蛋白必须经过鉴定并验证后才可能真正用于临床,因此各个实验室研究人员都在积极寻找食管癌的差异蛋白变化和差异蛋白适合的鉴定方法。Zhang 等^[10]发现 clusterin 在食管鳞癌术前血清中的表达显著降低。张红蕾等^[11]亦正寻找食管鳞癌特异的生物标志物。本研究将发现的差异蛋白查阅蛋白数据库,发现 7 908 kD 与之对应的蛋白可能为 UG,其定位于 11 号染色体的长臂,其是由类固醇激素诱导分泌的,进化上高度保守,UG 具有抗炎、肿瘤抑制及免疫调节多种生物学功能。据报道 UG 通过受体介导,自分泌和旁分泌途径,从而逆转癌细胞一些改变的表型,如侵入性。Yoon 等^[12]用腺病毒为载体的 UG-cDNA 转染肺癌细胞,这通过抑制核因子 κ B 的(NF- κ B)的活性减弱了环氧化酶-2 的基因表达。而该蛋白在食管癌组中表达强度升高可能为对癌组织的反应性增高。8 577.8 kD 查阅后对应的为淀粉样蛋白-A,人血清淀粉样蛋白 A、血清淀粉样蛋白基因均位于 11 号染色体短臂,属急性时相反应蛋白;当机体受到各种损伤刺激(烧伤、炎症、肿瘤等)时,血清中其水平即可能迅速而显著升高。目前传统的食管癌诊断方法已经不能满足临床的需要,通过 SELDI 技术结合人工神经网络和联合蛋白组学数据库鉴定差异蛋白的方式建立诊断模型,经过大样本的盲法验证模型效果发现与临床诊断吻合程度良好。鉴于很多蛋白质组学研究者已经完成了大量的基于质谱技术的蛋白质鉴定^[13],同时人们对于某些物质参与食管癌的形成机制已经有了一定的了解^[14-16];如果能够进一步根据数据库查询结果使用相关方法如 ELISA 等验证查询结果后所建的诊断模型的效果势必明显优于单一使用差异蛋白的结果。因此这可能是以后广泛用于临床的一个新思路。

参考文献:

- [1] 丁银环,胡琼英,梁双花,等. 运用内标校准法提高表面增强激光解吸电离飞行时间质谱检测中的重复性[J]. 中华检验医学杂志,2009,32(3):337-339.
- [2] 方晓松,周攀,汪宇鸿,等. 内镜下卢戈氏液染色对早期食管癌及癌前病变的诊断研究[J]. 重庆医学,2009,38

(11):1365-1366.

- [3] 熊刚,邱阳,杨康,等. 1 924 例食管癌和贲门癌的外科治疗经验[J]. 重庆医学,2008,37(4):342-344.
- [4] 冯笑山,王立东,单探幽,等. 食管癌淋巴结转移的蛋白质组学[J]. 郑州大学学报:医学版,2007,42(3):397-399.
- [5] 王士杰,张立玮,于卫芳,等. 高发区筛查人群食管鳞癌血清蛋白指纹图谱诊断模型的建立及临床价值[J]. 中华肿瘤杂志,2007,29(6):441-443.
- [6] 徐淑永,张琼,张朝霞,等. 新疆维吾尔族食管癌患者血清蛋白指纹图谱诊断模型的建立[J]. 新疆医科大学学报,2008,31(2):154-159.
- [7] 黄颖峰. 食管鳞癌血清 CM10 蛋白芯片诊断模型在早期食管癌筛查中的应用[D]. 北京:中国协和医科大学,2006.
- [8] 程真珍. 新疆维吾尔族食管癌 SELDI-TOF-MS 蛋白质组图谱的分析[D]. 乌鲁木齐:新疆医科大学,2008.
- [9] 胡琼英,王开正,丁银环,等. 胰腺癌血清差异蛋白的筛选及其临床应用[J]. 中华消化杂志,2009,29(11):767-768.
- [10] Zhang LY, Ying WT, Mao YS, et al. Loss of clusterin both in serum and tissue correlates with the tumorigenesis of esophageal squamous cell carcinoma via proteomics approaches[J]. World J Gastroenterol, 2003, 9(4): 650-654.
- [11] 张红蕾,高春芳,王秀丽,等. 飞行时间质谱技术在食管鳞癌血清学诊断中的应用[J]. 临床检验杂志,2007,25(4):280-282.
- [12] Yoon JM, Lim JJ, Yoo CG, et al. Adenovirus-uteroglobin suppresses COX-2 expression via inhibition of NF- κ B activity in lung cancer cells[J]. Lung Cancer, 2005, 48(2): 201-209.
- [13] Meuwis MA, Fillet M, Geurts P, et al. Biomarker discovery for inflammatory bowel disease, using proteomic serum profiling[J]. Biochem Pharmacol, 2007, 73(9): 1422-1433.
- [14] 童强,侯晓华,金曙. P27kip1 表达在食管癌放射抗拒中的意义[J]. 华中科技大学学报:医学版,2009,38(6):808-811.
- [15] Fukuoka T, Hibi K, Nakao A. Aberrant methylation is frequently observed in advanced esophageal squamous cell carcinoma[J]. Anticancer Res, 2006, 26(5): 3333-3335.
- [16] 黄青远,陈东义,赵志龙,等. Livin mRNA 和 Livin 蛋白在食管癌组织中的表达[J]. 中国现代医学,2007,17(3):298-304.

(收稿日期:2010-05-10 修回日期:2010-09-15)

更正

本刊 2010 年第 39 卷第 22 期刊登的《事件相关电位 P300 对系统性红斑狼疮患者认知功能障碍的临床研究》一文的基金项目应为四川省卫生厅科研课题[090150],特此更正。