

## · 临床诊断 ·

# 血清蛋白质指纹图谱检测诊断早期胃癌临床意义

赵长宏<sup>1</sup>, 陆海波<sup>1</sup>, 马玉彦<sup>2</sup>, 鲁海玲<sup>1</sup>, 唐雅莉<sup>2</sup>, 李鑫<sup>2</sup>

**Detection and Evaluation of Serum Proteomic Patterns by SELDI TOF-MS in Early Gastric Cancer**

ZHAO Chang-hong<sup>1</sup>, LU Hai-bo<sup>1</sup>, MA Yu-yan<sup>2</sup>, LU Hai-ling<sup>1</sup>, TANG Ya-li<sup>2</sup>, LI Xin<sup>2</sup>

1. Department of Oncology, The Affiliated Tumor Hospital of Harbin Medical University, 2. Tumor Institute, Harbin 150040, China

Corresponding Author: LU Hai-bo, E-mail: lu\_haibo200 @126.com

**Abstract** **Objective** To detect the serum proteomic patterns using SELDI TOF-MS ProteinChip array technology in gastric cancer, screen biomarker candidates, build diagnostic models and evaluate its clinical significance in early gastric cancer. **Methods** SELDI TOF-MS ProteinChip was used to detect the serum proteomic patterns of 34 patients with gastric cancer (12 cases of I / II stage and 22 cases of III / IV stage) and 30 healthy people. The diagnostic models were developed and validated by discriminant analysis. **Results** The model composed by 5 protein peaks 2 046m/z、3 179m/z、1 817m/z、1 725m/z and 1 588m/z could do the best in the diagnosis of gastric cancer. The specificity and sensitivity of it were 94.1% (32/34) and 93.3% (28/30) respectively. The single peak 4 665m/z can distinguish stage I / II stage and III / IV gastric cancer, the specificity and sensitivity of it were 91.7%, (11/12) and 95.4%, (21/22). **Conclusion** This method show great potential for the early detection, staging before operation and screening novel and better biomarkers to early gastric cancer.

**Key words**: Early Gastric cancer; SELDI TOF; Diagnosis; Proteomics

**摘要** 目的 应用 SELDI 蛋白质芯片检测胃癌患者血清蛋白质指纹图谱,筛选候选肿瘤标志物以建立诊断模型,并探讨其诊断早期胃癌的临床意义。方法 表面加强激光解吸电离-飞行时间质谱(SELDI TOF-MS)技术及其配套蛋白质芯片检测 34 例胃癌患者(I / II 期 12 例与 III / IV 期 22 例)和 30 例健康人的血清蛋白质组图谱,运用判别分析处理数据筛选标志物并建立诊断模型。结果 2 046m/z、3 179m/z、1 817m/z、1 725m/z 和 1 588m/z 等 5 个蛋白质峰组合所构建的诊断模型能达到鉴别胃癌患者和健康人的最佳诊断效果,特异度 94.1% (32/34),灵敏度 93.3% (28/30)。单个 4 665m/z 蛋白质峰诊断模型可达到鉴别 I / II 期与 III / IV 期胃癌效果,其特异度 91.6% (11/12),灵敏度 95.4% (21/22)。结论 该方法在胃癌的诊断尤其是早期诊断、术前分期及候选肿瘤标志物筛选等方面具有一定价值,值得进一步研究。

**关键词**: 早期胃癌; SELDI TOF; 诊断; 蛋白质组学

中图分类号: R735.2 文献标识码: A 文章编号: 1000-8578 (2006) 10-0750-04

## 0 引言

胃癌是常见恶性肿瘤之一,在世界范围内为第二多发恶性肿瘤,是消化系统最多发肿瘤<sup>[1,2]</sup>。可能是胃癌临床的异质性表现,迄今为止胃癌尚缺乏理想的肿瘤标志物,从而也给其早期诊断、术后疗效评价、预

后监测及胃癌患者的康复等造成了一定的困难。

表面加强激光解吸电离-飞行时间质谱(Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry, SELDI TOF-MS)是近年来发展起来的一种全新的蛋白质组学研究手段,其基本原理是通过特定的化学表面结合蛋白质分子再以激光轰击,使蛋白质成为离子而飞行于电场中并被检测器所测得,进而用不同位置、强度的峰来表示各种不同的蛋白质及其相对含量,最终形成可用于分析的图谱。该方法克服了传统蛋白质组学研究方法所存在的诸多问题,实现了质谱技术用于临床检测的飞跃<sup>[3,4]</sup>。

收稿日期: 2006-03-27; 修回日期: 2006-06-14

基金项目: 2003 年省自然科学基金资助项目(D0349),哈尔滨市科学技术计划项目(2005AA9cs116-15)

作者单位: 1. 150040 哈尔滨医科大学附属肿瘤医院内二科; 2. 黑龙江省肿瘤研究所

通讯作者: 陆海波, E-mail: lu\_haibo200 @126.com

作者简介: 赵长宏(1962-),男,博士,主任医师,主要从事消化系肿瘤的基础研究以及诊断治疗方面的临床研究

本研究应用 SELDI 质谱仪及其配套蛋白质芯片检测了 34 例胃癌患者和 30 例健康人的血清蛋白质指纹图谱,以筛选可用于临床胃癌早期诊断的特异性生物标记物。

## 1 资料和方法

1.1 资料 2004 年 1 月 ~ 8 月间我院收治的初诊胃癌患者 34 例,中位年龄 52 岁(34 ~ 74 岁)。常规根治手术后按 UICC 1997 年胃癌 TNM 分期法进行分期,其中 I 期患者 6 例, II 期 6 例, III 期 5 例,

IV 期 17 例,诊断均经术后病理证实。健康人群来自我院经胃镜取材并经病理证实的无胃部疾病的 30 例,中位年龄 49 岁(25 ~ 72 岁)。所有血样均于清晨空腹采集,静置离心分离血清,-80 ℃ 低温冰箱保存。

1.2 主要仪器、软件及试剂 PBS- 表面加强激光解吸电离-飞行时间质谱仪(SELDI TOF-MS)、能量吸收分子 SPA、CM10 型蛋白质芯片及相应分析软件 ProteinChip Software 3.2.1(美国 Ciphergen 公司);HEPES 缓冲盐(Sigma 公司);CHAPS 缓冲盐(Sigma 公司)。

### 1.3 方法 CM-10 芯片(弱阳离子 + 疏水膜)

1.3.1 步骤 用 HPLC<sub>2</sub>O 洗涤芯片,400r/min,震荡 5min 三次。先将 10mM/L HCL 倒入带盖试管中盖好后。震荡 5min。在用去离子水冲洗数次,再将芯片装入 HPLC<sub>2</sub>O 试管中震荡 5min。将芯片架子,胶垫超声清洗干净,用去离子水冲洗数次晾干后,将芯片装入处理器上。每孔加入 200μl 0.1M NaAC 结合 buffer,室温振荡 250r/min,5min 重复一次。甩干后再上样品。

1.3.2 样品处理 血清样品冰上溶解,4 度离心 10 000r/min,5min。取上清 20μl 加入 30μl U9(含 DTT)混匀,4 度震荡 250r/min,20min,每孔用封口膜封好。U1 配法:100μl U9(含 DTT)+900μl 50 mM Hepes pH 7.0。每孔加入 100μl U1 缓冲液混匀,盖严振荡 250r/min,30min。从上述 150μl 变形样品中取出 50μl,加入到 200μl 0.1M NaAC 结合 buffer 中混匀,取出 100μl 上样,室温振荡 250r/min,60min,取出样品。每孔加入 150μl 0.1M NaAC buffer,室温振荡 250r/min,5min,倒掉后加入 150μl NaAC buffer,共 3 次。再用 1mM Hepes pH 4.0 淋洗芯片 30s,重复一次。除去芯片晾干。

将 SPA 高速离心 30s,在 SPA 管中,加入 200μl 乙腈,200μl 1% TFA 充分振荡 5min,静止 5min。离心 1 000r/min 1min。每孔加入 SPA 0.5 ~ 1.0μl/点,重复一次。两次之间各点风干。

### 1.3.3 芯片检测 用加有 All-in-one 标准蛋白

的 NP20 芯片校正质谱仪,在 Ciphergen Protein-Chip 软件中设定读片程序读取芯片数据。计算机以每秒  $1 \times 10^9$  Hz 的速度获得原始数据并快速精确地绘制出蛋白质质谱图。其中纵坐标为蛋白质相对含量,横坐标为蛋白质质荷比。

1.3.4 蛋白质质荷比 结合 SPA 后的蛋白质在 SELDI TOF-MS 氦氖激光器的激光轰击下电离,带有电荷的蛋白质在加速电场的作用下,不同质荷比的蛋白质在长度一定的真空管中飞行所需的时间不同,蛋白质的质荷比(M/Z)与电离飞行时间的平方成正比。由  $E = UZ = 1/2mv^2$ ;  $t = L/v$  推出  $M/Z = Kt^2 = (2U/L^2)t^2$ 。其中 Z 为离子所带电荷数,U 为电压,v 为飞行速度,L 为加速飞行电场电压,K 为常数。

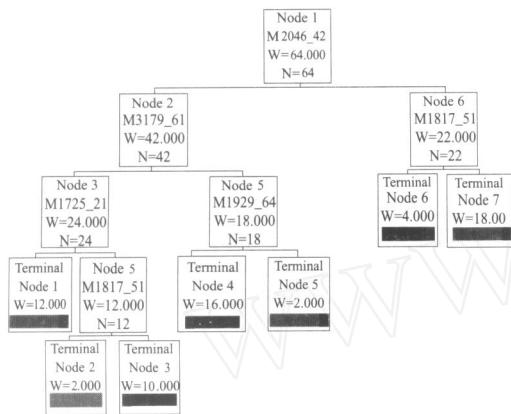
1.3.5 蛋白质相对含量 带有正电荷的蛋白质离子束在到达检测器的一瞬间,电子倍增器将产生的瞬时电流(瞬时电流  $I_t = Q/t$ ,其中 Q 为 t 时刻检测器检测到的电荷数)转换成蛋白质的相对含量。

1.4 数据收集及统计学分析 质谱仪参数设定:激光强度 190,灵敏度 9,数据收集范围 1 500 ~ 20 000m/z(蛋白质质量和电荷比值,以下简称质荷比)。原始数据先以 Proteinchip Software 3.2.1 软件校正,使总离子强度及分子量达到均一,并过滤噪音,初始噪音过滤值 5,二次噪音过滤值 2,以 8% 为最小阈值进行聚类。经上述数据预处理后,比较胃癌患者与健康人血清蛋白质质谱数据(由 Proteinchip Software 3.2.1 自带的 Biomarker Wizard 5.0 软件完成),用数据树方法对目标对象进行分类,对数据分组及相关性进行分析。寻找两组之间表达有差异的蛋白质峰,结合各个峰的权重比较不同蛋白质峰排列组合的判别分析结果,选出最佳排列组合,最后同时输出原始判别和交叉验证的结果。

## 2 结果

2.1 生物标记物的发现 用 Biomarker Wizard 软件和 Proteinchip Software 软件分析 34 例胃癌患者和 30 例健康人血清中蛋白质的组成,结果发现 2 046m/z、3 179m/z、1 817m/z、1 725m/z 和 1 588m/z 等 5 个蛋白质组成的生物标记物可将正常人与胃癌患者准确的分组。用这个生物标记物分析数据共产生 7 个终节点,见图 1:其中 1、2、5、7 均为只包含胃癌患者,共 34 例,满足 1.475(M2046)、2.239(M3179)、1.321(M1817)(括号代表相应蛋白质相对含量,下同)为终节点 1,满足 1.475(M2046)、2.239(M3179)、>1.321(M1752)m/z、-0.856(1817)为终节点 2,满足 1.475(M2046)、>2.239(M3179)、>

1.798(M1929) 为终节点 5 , 满足  $> 1.475$  (M2046) 、  $> -0.697$  (M1817) 为终节点 7 ; 终节点 3 、 4 、 6 为正常人 , 共 30 例 , 满足  $1.475$  (M2046) 、  $2.239$  (M3179) 、  $> 1.321$  (M1752) 、  $> -0.856$  (1817) 为终节点 3 , 满足  $1.475$  (M2046) 、  $> 2.2397$  (M3179) 、  $1.798$  (M1929) 为终节点 4 , 满足  $> 1.475$  (M2046) 、  $> -0.697$  (M1817) 为终节点 7 。错误分组率为 1.67 %。



峰值  $2046\text{m/z}$  、  $3179\text{m/z}$  、  $1817\text{m/z}$  、  $1725\text{m/z}$  和  $1588\text{m/z}$  截点分别为  $1.475$  、  $2.239$  、  $-0.856$  、  $1.321$

图 1 胃癌患者血清与健康人血清决定树

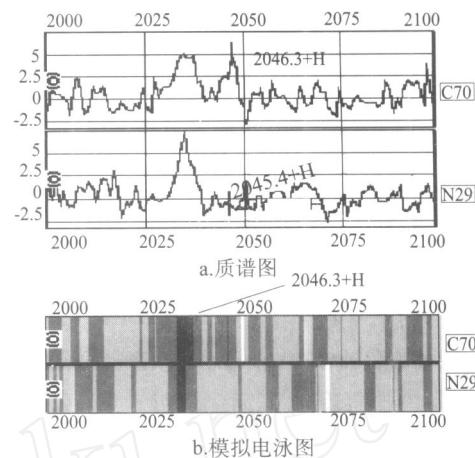
**2.2 胃癌患者与健康人的比较** 通过对 34 例胃癌患者和 30 例健康人血清蛋白质质谱数据的比较 , 判别分析显示其中 5 个蛋白质峰组合所构建的诊断模型 ( 模型 ) 能达到最佳诊断效果 :  $2046\text{m/z}$  、  $3179\text{m/z}$  、  $1817\text{m/z}$  、  $1725\text{m/z}$  和  $1588\text{m/z}$  , 其中质荷比为  $2046\text{m/z}$  的蛋白质峰如图 1 所示 , 可见正常人质荷比为  $2046\text{m/z}$  的蛋白质含量明显低于胃癌患者 , 经方差分析 , 差异有显著性 ( $P < 0.05$ ) 。原始判别的准确率为 93.75 % (60/64) , 特异度 94.1 % (32/34) , 灵敏度 93.3 % (28/30) , 见表 1 ; 交叉验证的准确率为 76.56 % (49/64) , 特异度 73.5 % (25/34) , 灵敏度 80.0 % (24/30) , 见表 2 。交叉验证显示模型 对 34 例胃癌患者中的 25 例做出了准确预测。

表 1 模型 原始判别的结果(例)

分组	胃癌患者	健康人	总计
胃癌患者	32	2	34
健康人	2	28	30
总计	34	30	64

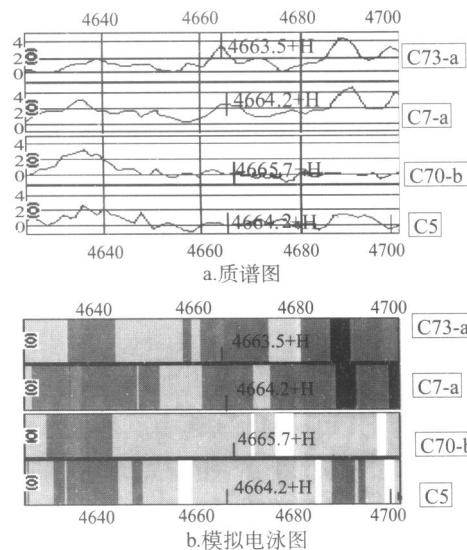
表 2 模型 交叉验证的结果(例)

分组	胃癌患者	健康人	总计
胃癌患者	25	9	34
健康人	6	24	30
总计	31	33	64



a, b 均为 1 例胃癌患者和 1 例健康人的对照。可见该蛋白质在胃癌患者中表达高于健康人

图 2 质核比为  $2046\text{m/z}$  的蛋白质峰



a, b 均为 2 例早期患者 ( / ) 和 2 例晚期胃癌 ( / ) 患者。可见该蛋白质在早期胃癌患者中表达高于晚期患者

图 3 质核比为  $4066\text{m/z}$  的蛋白质峰

**2.3 / 期与 / 期胃癌患者的比较** 比较 12 例 / 期患者与 22 例 / 期患者的质谱数据 , 其中  $4665\text{m/z}$  的单个蛋白质峰的诊断模型 ( 模型 ) 可达到鉴别 / 期与 / 期患者的最佳效果。原始判别 , 总准确率 94.1 % (32/34) , 特异度 91.7 % (11/12) , 灵敏度 95.4 % (21/22) ; 交叉验证结果 , 特异度 91.7 % (11/12) , 灵敏度 86.4 % (19/22) , 总准确率 88.2 % (30/34) 。

### 3 讨论

胃癌的诊断手段正从症状、临床体检、胃镜检查等经典手段向基因诊断、分子生物学检测手段发展<sup>[1,2]</sup> , 但作为后者重要组成部分之一的胃癌血清肿瘤标志物 , 近年来发展较缓 , 不能满足临床实际应用的需要。运用 SELDI TOF-MS 技术来分析鉴定

肿瘤不同于非肿瘤的蛋白质表达谱的全貌,从而获得内因、外因共同作用下肿瘤发生发展的较完整信息<sup>[5]</sup>,正是针对目前研究现状的不足之处做出的有益尝试,该方法已成功应用于卵巢癌以及前列腺癌<sup>[6,7]</sup>。在胃癌方面的研究也偶见报道<sup>[8,9]</sup>。

判别分析是一种根据观察或测量到的若干变量值,来判断研究对象如何分类的常用统计分析方法,是应用计算机进行疾病辅助诊断的主要统计学基础。其基本原理是根据已知的分类(如属于胃癌组或是属于健康人组)和表明特征的变量值(即各蛋白质峰的峰值数据)建立判别函数,再将某个个体的自变量值代入到判别函数,得出判别分数或计算属于各类的概率,从而判断该个体属于哪一类。本研究过程中的判别分析主要分两个步骤进行,即原始判别和交叉验证。前者是将所有病例的峰值数据建立函数,再逐一将每个病例代入函数进行判别;后者则是利用除了某一病例外的其余所有病例的数据建立函数,再把这一未参与函数构建的病例数据代入函数进行判别,如此判别每一个病例。这种交叉验证的方法在统计学上称为“留一法”,与原始判别相比更为真实、稳定、可靠,所以我们选择了交叉验证的结果作为最终结果。

本研究运用 SELDI TOF-MS 蛋白质芯片技术,得出了胃癌患者不同于健康人的血清蛋白质指纹图谱,应用判别分析筛选出能达到最佳诊断效果的蛋白质峰组合并建立诊断模型,交叉验证结果较满意,其灵敏度和特异度均远高于现有的各种肿瘤标志物,具有一定的诊断价值,筛选出的差异表达蛋白质将有可能成为新的候选肿瘤标志物。通过比较其对不同分期胃癌患者的诊断准确率可见,该模型对Ⅰ、Ⅱ期患者的诊断能力不亚于对Ⅲ、Ⅳ期的诊断,提示本方法在胃癌早期诊断方面有一定价值,所筛选出的候选肿瘤标志物也将有可能用于早期检测。而上文所提及的几种现有标志物,目前仅认为其在胃癌复发和转移的监测中有一定参考价值,而对早期诊断的价值则较低。比较Ⅰ/Ⅱ期患者和Ⅲ/Ⅳ期患者质谱数据构建的模型(单个蛋白质峰 4 665 m/z),能够较为准确的区分某一个胃癌患者是Ⅰ/Ⅱ期还是Ⅲ/Ⅳ期,有助于早期诊断、改善预

后,同时对术前治疗及手术方式的合理选择也具有一定的参考价值。

需要指出的是,运用判别分析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;所得到的最佳组合也并非诊断效力由高到低排列的各个峰的简单叠加。因此,利用这一方法筛选候选肿瘤标志物时,不能仅局限于最佳诊断组合中所列出的几个蛋白质,而应该着眼于所有表达差异具有统计学意义的蛋白质。

综上所述,SELDI TOF-MS 技术在胃癌的诊断尤其是早期诊断、候选肿瘤标志物的筛选及术前分期等方面具有一定的价值。进一步的研究仍在进行中。

#### 参考文献:

- [1] 金懋林. 胃癌[A]. 见:孙燕. 内科肿瘤学[M]. 北京:人民卫生出版社,2001. 549.
- [2] Howard W, John C, Paul. Neoplasms of the stomach, Cancer Medicine, E5th[J]. American Cancer Society, 2001, 1355.
- [3] Merchant M, Weinberger SR. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry[J]. Electrophoresis, 2000, 21: 1164-1177.
- [4] Issaq HJ, Veenstra TD, Conrads TP, et al. The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification[J]. Biochem Biophys Res Commun, 2002, 292: 587-592.
- [5] 蒋知俭. 医学统计学[M]. 北京:人民卫生出版社,1997. 281-299.
- [6] Petricoin EF, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to identify ovarian cancer[J]. Lancet, 2002, 359: 572-577.
- [7] Vlahou A, Schellhammer PF, Mendrinos S, et al. Development of a novel proteomic approach for the detection of transitional cell carcinoma of the bladder in urine[J]. Am J Pathol, 2001, 158: 1491-1502.
- [8] Ebert MP, Meuer J, Wiemer JC, et al. Identification of gastric cancer patients by serum protein profiling[J]. J Proteome Res, 2004, 3(6): 1261-1266.
- [9] Melle C, Ernst G, Schimmel B, et al. Characterization of pepsinogen C as a potential biomarker for gastric cancer using a histoproteomic approach[J]. J Proteome Res, 2005, 4(5): 1799-1804.

[编辑:安 凤]