

# 应用液体蛋白芯片飞行时间质谱技术研究 原发性肝癌门静脉癌栓形成相关标志物

李长煜 王小林<sup>Δ</sup> 龚高全 程洁敏 陈 颀 刘凌晓 李国平 王成刚 施东华

(复旦大学附属中山医院放射科 上海 200032)

**【摘要】** 目的 应用液体蛋白质芯片飞行时间质谱技术研究肝癌门静脉癌栓(portal vein tumor thrombus, PVTT)形成相关标志物,并建立预测模型。方法 采取非PVTT和PVTT原发性肝癌(hepatocellular carcinoma, HCC)患者血清各20例,采用磁珠分离系统纯化、基质辅助激光解析电离飞行时间质谱(matrix-assisted laser desorption/ionization time of flight mass spectrometry, MALDI-TOF-MS)分析及ClinProTools生物信息学方法研究其血清蛋白表达谱,通过神经网络算法(statistical neural networks, SNN)建立图谱诊断模型并进行验证。结果 经数据对比分析找到39个差异有统计学意义的质荷比( $m/z$ )峰( $P < 0.05$ ),从中筛选出最具划分能力的8个质荷比峰( $m/z = 4\ 420.43, 4\ 055.19, 2\ 988.80, 4\ 091.46, 4\ 072.24, 6\ 935.04, 4\ 109.53$ 和 $5\ 321.17$ ),并以此建立HCC的PVTT诊断模型,模型识别率达到97.5%,而预测率达到92.73%。非PVTT组和PVTT组的准确性分别为100%和95%。结论 筛选得到的蛋白可能成为原发性肝癌PVTT的相关标志物,诊断模型可能对预测PVTT的形成具有重要临床意义。

**【关键词】** 原发性肝癌; 门静脉癌栓; 液体蛋白芯片飞行时间质谱; 蛋白质组

**【中图分类号】** R 735.5 **【文献标志码】** A

## Serum protein biomarkers for portal vein tumor thrombus in hepatocellular carcinoma detected by ClinProt system

LI Chang-yu, WANG Xiao-lin<sup>Δ</sup>, GONG Gao-quan, CHENG Jie-min, CHEN Yi,

LIU Ling-xiao, LI Guo-ping, WANG Cheng-gang, SHI Dong-hua

(Department of Radiology, Zhongshan Hospital, Fudan University, Shanghai 200032, China)

**【Abstract】** **Objective** To screen potential protein biomarkers associated with the formation of portal vein tumor thrombus (PVTT) in hepatocellular carcinoma (HCC) patients, and to build a predictive model. **Methods** The serum samples from 40 HCC patients (20 with PVTT and 20 free of PVTT) were enriched by the magnetic bead and analyze by matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF-MS) and ClinProTools, respectively. Diagnostic model was established and identified by statistical neural networks (SNN). **Results** Thirty-nine peaks with significant statistical differences in mass charge ratio ( $m/z$ ) were found ( $P < 0.05$ ). Eight protein peaks ( $m/z = 4\ 420.43, 4\ 055.19, 2\ 988.80, 4\ 091.46, 4\ 072.24, 6\ 935.04, 4\ 109.53$  and  $5\ 321.17$ ) among them were selected to establish the predictive model of PVTT. Specificity and accuracy of this model was 97.5% and 92.73%, respectively. Predictive values in patients with and without PVTT were 100% and 95%. **Conclusions** These 8 protein peaks may be potential biomarkers of PVTT in HCC patients, and the predictive model may have great clinical significance in predicting the formation of PVTT.

**【Key words】** hepatocellular carcinoma; portal vein tumor thrombus; ClinProt system; proteome

门静脉癌栓(portal vein tumor thrombus, PVTT)是原发性肝癌(hepatocellular carcinoma, HCC)的重

要生物学行为,在加重门脉高压、加速肿瘤转移等病理过程中起到关键作用,是影响HCC患者治疗方案

<sup>Δ</sup>Corresponding author E-mail: xlwang-shmu@vip.sina.com

及预后的重要因素<sup>[1]</sup>。HCC 患者 PVTT 的发生率为 44.0%~62.2%,如不治疗,合并 PVTT 的 HCC 患者中位生存期为 3~6 个月,同时 PVTT 也是 HCC 的治疗难点之一<sup>[2]</sup>。提高 HCC 患者的生存率、降低死亡率关键在于早期发现与诊断,准确判断病情阶段,选择合理治疗方案。积极寻找敏感性和特异性较高的 PVTT 相关标志物可以帮助临床进一步了解肝癌发展规律、准确判断病情阶段、合理选择治疗方案。近年来蛋白质组学技术特别是质谱技术的发展为肿瘤诊断开创了新领域,有望以复杂的蛋白质表达模式代替单一生物标志物,整体解析肿瘤的蛋白表达谱诊断成为可能。目前一种新的应用功能磁珠、ANCHORCHIP 技术、基质辅助激光解吸电离飞行时间质谱(matrix-assisted laser desorption/ionization time of flight mass spectrometry, MALDI-TOF-MS) 和 ClinProTools (CPT) 的生物信息学方法组成的 ClinProt 系统以其高敏感性和可重复性逐渐被用于肿瘤血清蛋白质表达谱的研究<sup>[3-6]</sup>,我们应用此技术平台对非 PVTT 和 PVTT 肝癌患者血清蛋白质组进行比较研究,探索具有潜在诊断意义的血清肿瘤标志物。

## 资料和方法

**资料** 本实验所用血清标本来自复旦大学附属中山医院介入科 2008 年 11 月至 2009 年 6 月间收治的 HCC 患者。入选标准为新发现的、符合 1999 年全国肝癌学术会议诊断标准的肝癌患者,无远处转移,未经手术、介入及放、化疗等治疗。选取非 PVTT 和 PVTT 的患者各 20 例。PVTT 组为通过影像学检查(CT 或 MR)可以明确诊断的门静脉主干或一级分支癌栓;非 PVTT 组为无影像学可见的门静脉任何分支癌栓,且临床随访 3 个月影像学未见门静脉任何分支癌栓生成。40 例肝癌患者中男性 32 例,女性 8 例,年龄(52.2 ± 11.1)岁,AFP 阳性(≥20 ng/mL)者 33 例,AFP 阴性(<20 ng/mL)者 7 例(表 1)。所有研究对象均签署知情同意书。患者在清晨空腹状态下采集静脉血 4 mL,置于 BD 管中,室温垂直放置 1 h,5 000 r/min 离心 5 min,取血清按 20 μL 分装, -80 °C 保存。

**试剂和仪器** 为去除血清中高丰度蛋白的影响,本研究利用液体蛋白芯片(又称磁珠,表面经预处理,捕获蛋白的表面积远大于固体蛋白芯片)进行血清蛋白提纯<sup>[7-8]</sup>。选用 SPE-CM 弱阳离子磁珠进行分析,磁珠直径小于 1 μm。分离用成套试剂(德国 Bruker Daltonics 公司),分析用三氟乙酸、乙腈、乙醇、丙酮(色谱纯,美国 Sigma 公司),标准血清

(美国 Sigma 公司),标准品为肽段和蛋白的混合物(德国 Bruker Daltonics 公司),质谱仪(Autoflex III 型,德国 Bruker Daltonics 公司)。

表 1 HCC 患者的临床资料  
Tab 1 Clinical data of HCC patients

Characters	Sample No.	
	PVTT group	PVTT group
Gender		
Male	14	18
Female	6	2
Average age (y)	52.3(31 - 74)	51.5(27 - 65)
Clinical index (ng/mL)		
AFP	3 876.11	3 426.6
CEA	13.435	15.423
CA199	36.503	30.287
Bilubin	15.281	11.321
Albumin	37.294	39.346

**液体蛋白芯片提取血清蛋白** 血清冰浴下解冻,取 5 μL 加入已装好结合缓冲液(BS)和 SPE-CM 磁珠的样品管中,置磁架上分离 5 min;仔细吸去悬浮的液体,用磁珠清洗缓冲液(WS)清洗 3 次,每次清洗在磁珠分离器前后相邻两孔间反复移动样品管 10 次(注意磁珠在管中的运动),直至液体清澈,完全吸去悬浮的液体并避免吸走磁珠;加入 5 μL 磁珠洗脱缓冲液(ES),溶解贴壁的磁珠;放入磁珠分离器使磁珠贴壁 2 min;将上清液移入干净的 0.5 mL 样品管,加入 5 μL 稳定缓冲液(SS),小心吹打、混匀,即可用于质谱分析<sup>[7-8]</sup>。

**MALDI-TOF-MS 检测的质量控制及标准** 每个标准品可对其周围的 4 个样品蛋白分子量进行校正,采用同一标准品结晶靶点、不同结晶点多点采集,累加后获得标准品图谱,校正平均分子量偏差应小于 0.01%<sup>[9]</sup>。采用标准血清对整个实验流程进行质量控制,每 7 个样本与 1 个标准血清匹配,同时进行实验,采集的标准血清图谱与数据库中已有的对应 WCX 磁珠相同采集方法的标准血清数据进行比较,计算变异系数。

**MALDI-TOF-MS 检测肝癌患者血清蛋白指纹图谱** 所有样品随机点样于 384 普通靶,先点 1 μL 样品,室温放干后,再点 1 μL 基质(3 mg/mL CHCA、50% ACN、2% TFA),室温放干后上机检测。质谱数据采集方法采用 ClinProt 线性阳离子模式,参数设置如下:第一离子源 20.00 kV,第二离子源 18.90 kV,检测范围为相对分子质量 1 000~12 000。每个样品重复点 3 个靶点,同一靶点的不同结晶点进行多点采集,每个点采集 50 shots,选择 8 个不同位置,累计 400 shots。为获得更好的质谱图,先选择高能量(8%~10%)激光轰击,然后换成

低能量激光采集图谱<sup>[8]</sup>。

**ClinProTools 分析蛋白表达谱和统计学处理** 首先,对图谱进行预处理,包括去噪平滑,过滤信噪比小于5的多肽峰,并进行校正和归一化。随后,选择软件内嵌的统计算法进行统计分析,对每个蛋白质荷比峰( $m/z$ )作  $T$  检验和 Wilconxon 秩和检验,正态分布数据选用  $T$  检验方法,非正态分布数据选用 Wilconxon 秩和检验, $P < 0.05$  为差异有统计学意义。通过神经网络算法(statistical neural networks, SNN)建立图谱诊断模型<sup>[10]</sup>。所有样本作为训练集,构建分类模型并验证,计算得出灵敏度和特异度。交叉验证用来预测今后验证数据的能力,所有样本随机选取70%作为训练集构建分类器,再用剩下的30%作为测试集进行验证,得出验证的正确率,过程重复10次,得出平均验证正确率,即为预测率。

## 结 果

**标准品和标准血清保证系统稳定性和实验重复性** 本研究分别采用标准品和标准血清作为质控标准,作为外标校正的标准品含有11个肽(表2),平均分子量偏差小于100 ppm,每8个样品数据采集前先做外标校正;同时加入4个标准血清作为质控。

表2 标准品多肽峰

Tab 2 Protein peaks of substance

Substance	$m/z$
Angiotensin II	1 047.18
Angiotensin I	1 297.48
Substance P	1 348.64
Bombesin	1 620.86
ACTH clip 1-17	2 094.42
ACTH clip 18-39	2 466.68
Somatostatin 28	3 149.57
Ubiquitin	4 283.45
Insulin	5 734.56
Cytochrome c	6 181.05
Ubiquitin	8 565.89

**磁珠筛选差异多肽峰** 将 PVTT 和非 PVTT 患者血清用磁珠分离系统纯化,然后用 MALDI-TOF-MS 分析。用 FlexAnalysis 3.0 软件对图谱进行分析,共检测到113个多肽峰。采用 ClinProTools 2.1 软件对数据进行统计学处理,差异有统计学意义( $P < 0.05$ )的多肽峰共有39个(表3)。

在 PVTT 患者中, $m/z$  值为 1 208.34、2 988.80、839.18、831.61、807.17、1 619.6 等24个蛋白稳定上调, $m/z$  值为 5 291.31、5 321.17、5 265.6、5 823.1 等15个蛋白稳定下调。对软件自动筛选出的最具划分能力的

两个蛋白( $m/z$  为 4 091.46 和 5 337.19)进行分析,代表峰图以及在样品间的分布图见图1、2,非 PVTT 组与 PVTT 组样品分布重叠区域较少,说明这两个蛋白的分组能力良好。

表3 PVTT 组和非 PVTT 组间表达量有统计学差异的多肽峰

Tab 3 Protein peaks with statistical difference between PVTT group and free of PVTT group

$m/z$	PWKW ( $P$ value) <sup>(1)</sup>	Average PI <sup>(2)</sup>		SD	
		FP <sup>(3)</sup>	P <sup>(4)</sup>	FP	P
1 208.34	<0.001	7.14	3.42	1.53	1.61
2 988.80	<0.001	17.86	12.34	3.45	2.90
5 291.31	<0.001	11.53	20.69	3.50	7.04
839.18	<0.001	6.77	2.64	4.28	1.46
5 321.17	<0.001	19.13	40.02	7.37	17.50
831.61	<0.001	8.95	3.14	7.86	1.70
807.17	<0.001	6.75	2.24	6.98	1.04
5 265.60	<0.001	11.73	17.57	2.42	4.75
5 823.10	<0.001	64.04	129.41	47.99	71.40
1 619.60	<0.001	10.89	6.23	2.58	3.27
1 466.52	<0.001	13.38	7.11	4.70	4.19
6 563.03	<0.001	11.50	8.27	2.56	1.97
4 055.19	<0.001	21.25	37.81	7.79	14.19
5 353.95	<0.001	44.94	88.84	29.96	40.34
4 420.43	<0.001	20.80	11.37	9.84	4.13
1 450.51	<0.001	8.80	5.51	2.06	2.87
5 806.01	0.001 11	113.78	229.15	99.75	174.41
6 242.33	0.001 19	6.81	4.90	2.05	3.22
1 262.69	0.001 45	8.23	5.27	2.10	2.49
1 969.87	0.003 57	12.87	9.01	3.08	3.20
8 934.09	0.003 81	136.24	335.19	114.16	202.52
4 091.46	0.004 06	46.56	75.91	18.08	28.46
1 058.51	0.004 34	3.53	2.28	1.27	1.48
6 110.06	0.005 76	21.92	11.11	17.39	4.58
2 972.39	0.006 14	27.44	20.03	7.47	5.95
4 072.24	0.006 56	21.83	34.48	5.73	16.82
2 211.02	0.008 83	34.01	22.74	34.72	48.32
4 109.53	0.008 83	16.42	23.34	5.67	7.67
845.80	0.008 83	13.40	7.89	10.03	6.32
4 194.55	0.008 84	37.64	65.22	19.81	30.44
7 062.32	0.008 84	11.25	8.49	3.16	2.77
1 331.38	0.008 84	8.08	6.15	1.91	3.41
5 337.19	0.009 45	106.32	190.30	55.25	112.43
5 248.97	0.010 10	8.68	11.65	1.81	3.66
9 287.76	0.017 20	111.68	248.16	83.93	178.93
2 863.94	0.020 00	19.97	13.35	7.87	7.40
7 432.68	0.021 20	8.80	6.85	2.48	2.48
854.66	0.022 50	7.70	4.84	3.89	3.45
5 597.50	0.023 90	9.70	7.54	2.89	2.85

(<sup>1</sup>) The peaks of lowest value in PWKW have the most significant statistical difference and the highest discrimination between these two groups; (<sup>2</sup>) Average PI: Average peak intensity; (<sup>3</sup>) FP: Free of PVTT group; (<sup>4</sup>) P: PVTT group.

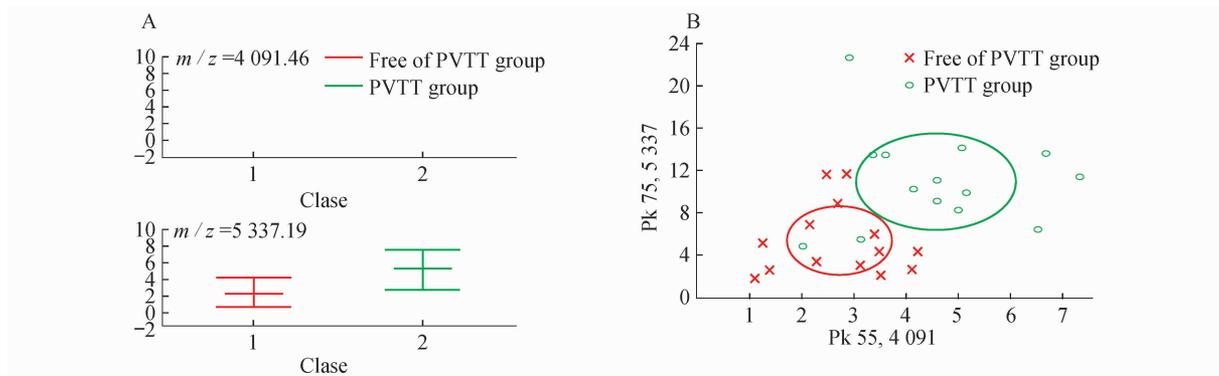


图1 两组中  $m/z$  为 4 091.46 和 5 337.19 的标准差分布图和样品分布图

Fig 1 The distribution of SD and sample for the two protein peaks ( $m/z=4\ 091.46$  and  $5\ 337.19$ ) in PVTT group and free of PVTT group

A: The distribution of SD; B: The distribution of sample; Y axis: 5 337.19; X axis: 4 091.46; The values showed the intensity of each sample.

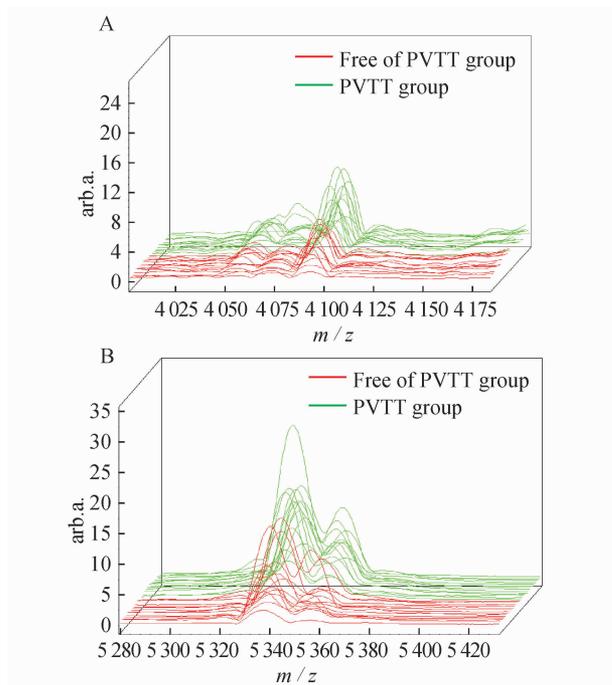


图2 两组样本中  $m/z$  为 4 091.46 和 5 337.19 的多肽峰的质谱图

Fig 2 Mass spectrum of the two protein peaks ( $m/z=4\ 091.46$  and  $5\ 337.19$ ) in PVTT group and free of PVTT group

A:  $m/z=4\ 091.46$ ; B:  $m/z=5\ 337.19$ . The peak intensity are much higher in PVTT group than in free of PVTT group for most samples.

**模型建立和盲筛** 通过 SNN 建立图谱诊断模型,选取非 PVTT 组和 PVTT 组各 14 例标本的结果用于建模。建立模型选用特征峰  $m/z$  为 4 420.43、4 055.19、2 988.80、4 091.46、4 072.24、6 935.04、4 109.53 和 5 321.17 的 8 个蛋白峰。模型识别率达到了 97.5%,准确性在非 PVTT 组和 PVTT 组分别为 100% 和 95%,而预测率达到了 92.73%。对于模型验证,我们使用其余各 6 例标本进行盲筛验证,采用交叉

验证来预测今后验证数据的能力,所有样本随机选取 70% 作为训练集构建分类器,其余 30% 作为测试集进行验证,模型识别率达到了 92.86%,准确性在非 PVTT 组和 PVTT 组分别为 92.86% 和 92.86%。

## 讨 论

HCC 容易侵犯肝内血管,特别是门静脉,PVTT 是导致转移复发和影响疗效的最重要因素。现有各种单一疗法无法取得满意疗效,肝癌合并 PVTT 的治疗应采取积极的综合治疗方案。目前 PVTT 的诊断主要依靠 CT 及 MR 等影像学手段,仍然难以早期发现或者预测 PVTT 的形成。早期发现 PVTT、探寻肝癌 PVTT 形成机制及寻找关键防治靶点是肝癌研究中的重大课题,对提高患者生存期具有重要意义。目前研究认为,肝癌 PVTT 的形成与门静脉系统解剖特点,肿瘤微环境、抑癌基因的失活、细胞外基质功能紊乱、黏附因子变化及肿瘤血管生成因子等多方面因素协同作用相关。

蛋白质组学是一个迅速发展的领域,为肝癌早期诊断和标志物鉴定提供了新思路<sup>[11]</sup>。蛋白质组学技术的发展经历了二维凝胶电泳结合质谱技术、激光捕获显微切割技术以及近些年发展起来的飞行时间质谱技术、质谱成像技术等<sup>[12-13]</sup>。SELDI 蛋白质芯片技术是一项非常具有潜力的适合于差异蛋白质组学研究的新技术。SELDI 蛋白质芯片能通过不同的表面修饰选择性地捕获蛋白质,因此有效地降低了待测样品蛋白质成分的复杂性,并具有快速、操作简单、样品用量少的特点。应用该技术已经成功地筛选出包括卵巢癌、前列腺癌、膀胱癌、喉癌等癌症的蛋白标志物。但是,SELDI 技术重复性较差,且不能对筛选出来的差异蛋白进行鉴定。本研

究采用了 ClinProt 系统。与 SELDI-TOF-MS 相比, ClinProt 系统可用 MALDI-TOF/TOF-MS 技术对差异表达的多肽/蛋白进行序列鉴定,从而明确获得的差异蛋白是否已知,加深人们对肿瘤发生、发展机制的认识。因此, ClinProt 系统是一种快速、简便易行、高通量且重复性高的蛋白质组学研究方法,其综合了液体蛋白芯片-磁珠分离系统, AnchorChip 技术与质谱技术的优点,大大提高了对蛋白质的鉴定能力,可用于生物标志物发现、鉴定和蛋白质谱的分析<sup>[11]</sup>。

关于肝癌 PVTT 的研究国内外报道较少。Qiu 等<sup>[14]</sup>利用 2-DE 技术平台联合 MALDI-TOF MS/MS 对伴有 PVTT 和非 PVTT 的人肝癌原发瘤组织中的差异蛋白质表达谱进行对比分析,最终鉴定出 11 种相对分子量低于  $2 \times 10^4$  的蛋白质,其中在 PVTT 组中表达上调的有 5 种,包括钙结合蛋白 S100A11 等。本研究采用 ClinPro 系统分析了 20 例非 PVTT 肝癌患者和 20 例伴有 PVTT 肝癌患者的血清。结果显示, PVTT 患者有 24 个差异蛋白高表达, 15 个差异蛋白低表达。其中,软件自动筛选出了最具划分能力的 2 个蛋白,  $m/z$  分别是 4 091.46 和 5 337.19。诊断模型识别率和预测能力都较为满意,如果能进一步验证与完善,利用该模型通过血清标本早期判断 PVTT 的形成并及早干预对于提高肝癌治疗效果具有重要意义,显示出 ClinProt 系统在肿瘤特异标志物研究上有广泛的应用前景。通过以上研究奠定的基础,下一步我们将增加临床样本数对实验数据加以验证,对诊断模型中的差异多肽峰进行分离纯化和鉴定,寻找出敏感性高、特异性强的标志物应用于临床诊断。

#### 参 考 文 献

[1] Pirisi M, Avellini C, Fabris C, et al. Portal vein thrombosis in hepatocellular carcinoma: age and sex distribution in an autopsy study[J]. *J Cancer Res Clin Oncol*, 1998, 124(7): 397-400.

- [2] Minagawa M, Makuuchi M. Treatment of hepatocellular carcinoma accompanied by portal vein tumor thrombus[J]. *World J Gastroenterol*, 2006, 12: 7 561-7 567.
- [3] Adam BL, Qu Y, Davis JW, et al. Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men[J]. *Cancer Res*, 2002, 62(13): 3 609-3 614.
- [4] Richter R, Schulz-Knappe P, Schrader M, et al. Composition of the peptide fraction in human blood plasma: database of circulating human peptides[J]. *J Chromatogr B Biomed Sci Appl*, 1999, 726(1-2): 25-35.
- [5] Diamandis EP. Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems[J]. *J Natl Cancer Inst*, 2004, 96(5): 353-356.
- [6] Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool opportunities and potential limitations[J]. *Mol Cell Proteomics*, 2004, 3(4): 367-378.
- [7] 王静, 张小为, 袁杨, 等. 应用液体蛋白芯片飞行时间质谱系统研究卵巢癌的血清蛋白表达谱[J]. *现代妇产科进展*, 2008, 17(2): 234-236.
- [8] 李甜, 谌宏鸣. 应用液体蛋白芯片-飞行时间质谱技术从血清中筛选胃癌肿瘤标志蛋白[J]. *现代肿瘤医学*, 2009, 17(8): 221-223.
- [9] 王坚, 李尧华, 郑晓立, 等. 弱阳离子交换纳米磁珠介导的帕金森病血清蛋白质组学的初步研究[J]. *中华神经科杂志*, 2009, 42(2): 18-20.
- [10] Ketterlinus R, Hsieh SY, Teng SH, et al. Fishing for biomarkers: analyzing mass spectrometry data with the new ClinProTools software[J]. *Biotechniques*, 2005, Suppl: 37-40.
- [11] Wong SC, Chan CM, Ma BB, et al. Advanced proteomic technologies for cancer biomarker discovery[J]. *Expert Rev Proteomics*, 2009, 6(2): 123-134.
- [12] 黄成, 樊嘉, 周俭, 等. 肝细胞癌门静脉癌栓形成相关的血清蛋白分子标记物研究[J]. *中华医学杂志*, 2005, 85: 781-785.
- [13] 黄成, 樊嘉, 周俭, 等. 转移与无转移肝细胞癌的血清蛋白质指纹图谱的比较[J]. *中华实验外科杂志*, 2005, 22: 102-104.
- [14] Qiu JG, Fan J, Liu YK, et al. Screening and detection of portal vein tumor thrombi-associated serum low molecular weight protein biomarkers in human hepatocellular carcinoma[J]. *J Cancer Res Clin Oncol*, 2008, 134(3): 299-305.

(收稿日期: 2009-12-15; 编辑: 段佳)