

· 论著 ·

血清蛋白质标记物在神经母细胞瘤中的诊断意义

陈新让¹ 王家祥²

【摘要】 目的 寻找特异的蛋白质标记物,探讨神经母细胞瘤血清蛋白质标记物的检测及初步诊断模型的构建和临床应用。**方法** 收集血清样本 87 例,其中 47 例为神经母细胞瘤患儿,30 例为其它恶性实体肿瘤患儿,10 例为健康儿童;用 ZUCI-Protein Chip Data Analyze System 分析软件进行数据处理;经留一法交叉验证,分类器评价模型的预测效果。**结果** 构建 3 个模型并筛选出 10 个蛋白质标记物,能成功区分神经母细胞瘤和健康儿童蛋白质谱差异表达模型的敏感性为 100.00%,特异性为 100.00%,区分神经母细胞瘤术前和术后蛋白质谱差异表达模型的敏感性为 100.00%,特异性为 100.00%,区分神经母细胞瘤与其它恶性实体肿瘤血清蛋白质指纹图谱模型的敏感性为 88.89%,特异性为 100.00%。**结论** 用 SELDI-TOF-MS 及生物信息学技术,并结合支持向量机(SVM)初步建立的模型可作为神经母细胞瘤的另一种特异性强、敏感性高的辅助检查手段。

【关键词】 血清蛋白类;神经母细胞瘤/诊断

Clinical application of serum protein markers in neuroblastoma. CHEN Xin-rang, WANG Jia-xiang*. The pediatric hospital of Zhengzhou, Zhengzhou, 450052, China

【Abstract】 Objective To search the specific protein biomarkers constructing the serum protein markers in the early diagnosis of Neuroblastoma. **Methods** Eighty-seven serum samples (including 47 Neuroblastoma patients, 30 children malignant solid tumor 10 healthy children) were detected by SELDI-TOF-MS, and proceeded by ZUCI-protein chip Data Analyze System Software to be identified by Jackknife, and evaluate the model's predictive effect. **Results** To analyze all the serum protein icon by cluster analysis can find a proteomics model. The detective model in combined with 10 biomarkers could separate neuroblastoma from the healthy group with sensitivity of 100.00%, specificity of 100.00%; separate preoperative neuroblastoma from postoperative neuroblastoma with sensitivity of 100.00%, specificity of 100.00%; The diagnosis model combined with 5 biomarkers could separate neuroblastoma from other children malignant solid tumors with sensitivity of 88.89%, specificity of 100.00%. **Conclusion** This detective model is a more effective method for the early diagnosis neuroblastoma, and can develop to a new path to research for new tumor markers to study the proteomics in neuroblastoma. The model of neuroblastoma combined by SELDI-TOF-MS with SVM and bioinformation technique can regard as another more specificity and sensitivity auxiliary examination mean.

【Key words】 Blood Proteins; Neuroblastoma/DI

神经母细胞瘤(neuroblastoma, NB)在儿童恶性肿瘤发病率中占第4位,我国每年新发病约3000例,是最常见的外周神经系统恶性肿瘤,1岁以内患儿大部分可自行痊愈,1岁以上患儿恶性程度高,诊断时多已为晚期,预后差。基于 SELDI-TOF-MS 质谱技术在恶性肿瘤诊断中取得的进展,B超、CT等

很难发现微小病变,无法早期发现临床无体征的肿瘤;支持向量机(SVM)具有小样本、良好的推广性能以及全局最优等优点^[1]。SVM算法在膜蛋白类型预测方面效果明显,一致测试对每类预测得到的准确率比交叉校验、独立测试高^[2]。因此,本研究通过 SELDI-TOF-MS 结合 SVM 技术来寻找血清蛋白质标记物,进行神经母细胞瘤的早期诊断。

doi:10.3969/j.issn.1671-6353.2011.05.004

作者单位:1,郑州市儿童医院(河南省郑州市,450052);2,郑州大学第一附属医院(河南省郑州市,450052),通讯作者:陈新让;E-mail:chenxr66608@163.com,本研究为国家自然科学基金资助项目(项目号30571930)

材料与方法

一、对象

收集血清样本 77 例,来自神经母细胞瘤患儿

37 例(1 例已化疗 4 个疗程),其它恶性实体肿瘤患儿 30 例(包括肾母细胞瘤 21 例,肝母细胞瘤 3 例,肾横纹肌肉瘤 3 例,恶性畸胎瘤 2 例,胰母细胞瘤 1 例),正常对照组 10 例,均为幼儿园体检健康的儿童。所有恶性肿瘤患儿均经免疫组化证实后由 2 名病理学教授证实。37 例神经母细胞瘤患儿中,男 22 例,女 15 例,平均年龄(1.9 ± 0.21)岁。30 例其它恶性实体肿瘤患儿中,男 21 例,女 9 例,平均年龄(3.7 ± 0.15)岁。正常对照组与肾母细胞瘤组患儿年龄性别相配对。所有外周静脉血标本均于手术日清晨空腹抽取,于 4°C 下静置 1 h 后 3KG000 rpm 离心 10 min,分离出的血清储存于 -80°C 低温冰箱中。CT 投照方法:西门子 SCT 机,扫描范围自膈肌上方 1 cm 至肾下极,间隔 8 mm,层厚 8 mm。感兴趣间隔 5 mm,层厚 5 mm,扫描时间 2 s。造影剂采用欧乃派克,剂量为 $1.5 \sim 2.0 \text{ mL/kg}$ 。

二、主要仪器和试剂

Ross Temp Inc, CIPHERGEN PBS II + SELDI - TOF - MS 和 WCX2 蛋白芯片, ORBIT300, Nuaire, Bioprocessor, Wallac 1296 - 003 DELFIAPLATESHAKE, 均购自美国 CIPHERGEN 公司; Milli-Q Plus 购自 Millipore inc 公司; CHAPS, Urea, DTT, NaAc, SPA (Sinapinic acid) 均购自美国 Promega 公司。

三、SELDI 操作的技术路线

样本冰融, 4°C 10 000 rpm 离心 2 min; 取 96 孔板, 放冰盒上, 每孔加 U9 (9M Urea, 2% CHAPS, 1% DTT) 10 μL , 血清 5 μL , 4°C 层析柜 600 rpm 振荡 30 min; 振荡结束前 15 min 做芯片预处理, 芯片装入 Bioprocessor 中并记下芯片号; 每孔加 NaAC (100 Mm, pH 4) 200 μL , 层析柜中 600 rpm, 5 min, 重复 1 次; U9 处理后 96 孔板放冰上, 快速加入 NaAC 185 μL 在层析柜中 6 000 rpm 振荡 2 min; 加已处理好的样本 100 μL 到芯片, 层析柜 4°C 600 rpm 振荡 1 h, 甩干, 重复 3 次; 用 hplc 水 200 μL 洗 2 次快速甩干; 用 50% 饱和的 SPA 1 μL 重复 2 次; 上机检测; 将芯片放入 CIPHERGEN 读谱仪中检测。

四、数据收集与处理

用 protein chip software 3.1 做校正处理, 应用 ZUCI-Protein Chip Data Analyze System 软件包进行分析。SVM 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建立判别模型, 用留一法交叉验证作为评估模型判别效果。

五、临床分期标准

临床分期按国际神经母细胞瘤分期标准(International Neuroblastoma Staging System, INSS) [5]。

六、统计学处理

所有初步筛选出的质荷比做 Wilcoxon 秩和检验, 检验标准设为 $\alpha = 0.05$ 。

结 果

一、神经母细胞瘤组与正常对照组

神经母细胞瘤组和正常对照组的质谱经过初步过滤筛选得到 535 个 M/Z 峰, 对其相对强度做 Wilcoxon 秩和检验分析, 得到 P 值 < 0.01 的 M/Z 峰 2 个, 从差异显著蛋白质峰的任意组合中, 采用 SVM 筛选出预测值的约登(youden)指数最高的组合模型, 筛选出 M/Z 位于 4 450.2、6 439.4 和 7 693.2 的标志物 3 个(表 1), 显示神经母细胞瘤组低表达, 正常对照组高表达(图 1)。联合 2 个潜在标记物作为输入值, 用留一法交叉检验, 在测试集上判别模型的敏感性为 100.00%, 特异性为 100.00%。

二、神经母细胞瘤术前组与术后组

神经母细胞瘤组和正常对照组的质谱经过初步过滤筛选得到 508 个 M/Z 峰, 对其相对强度做 Wilcoxon 秩和检验分析, 得到 P 值 < 0.01 的 M/Z 峰 3 个, 从差异显著蛋白质峰的任意组合中, 采用 SVM 筛选出预测值的约登(youden)指数最高的组合模型, 筛选出 M/Z 位于 4 244.3 和 7 619.4 的标志物 2 个(表 2), 显示神经母细胞瘤术前组低表达, 术后组高表达(图 2)。联合 2 个潜在标记物作为输入值, 用留一法交叉检验, 在测试集上判别模型的敏感性为 100.00%, 特异性为 100.00%。

三、神经母细胞瘤组与其它恶性实体瘤组(盲法验证)

神经母细胞瘤组和其它恶性实体瘤组质谱经过初步过滤筛选得到 520 个 M/Z 峰, 对其相对强度做 Wilcoxon 秩和检验分析, 得到 P 值 < 0.01 的 M/Z 峰 4 个, 从差异显著蛋白质峰的任意组合中, 采用 SVM 筛选出预测值的约登(youden)指数最高的组合模型, 筛选出 M/Z 位于 3 414.50、4 634.14、4 715.97、6 616.45 和 9 309.98 的标志物 5 个(表 3), 显示神经母细胞瘤组低表达, 其它恶性实体瘤组高表达(图 3)。联合 2 个潜在标记物作为输入值, 用留一法分类器评价模型的预测效果, 在测试集上判别模型的敏感性为 88.89%, 特异性为 100.00%。

表 1 神经母细胞瘤组与正常对照组 M/Z 位于 4 450.2、6 439.4、7 693.2 的统计资料(N-神经母细胞瘤,D-正常小儿)

M/Z	Mean-N	SD-N	Mean-D	SD-D	P 值
4 450.213	1 427.84	119.58	850.67	86.65	0.000 207 596 71
6 439.426	4 961.89	990.86	1 518.03	801.35	0.000 383 504 08
7 693.237	2 760.66	221.31	876.23	91.35	0.000 497 852 31

表 2 神经母细胞瘤术前组与术后组 M/Z 位于 4 244.3、7 619.4 的统计资料(N-神经母细胞瘤术前组,D-术后组)

M/Z	Mean-N	SD-N	Mean-D	SD-D	P 值
4 244.343	1 427.84	119.58	950.67	76.65	0.000 120 758 67
7 619.412	4 961.89	1 060.66	1 518.03	301.35	0.004 535 143 92

表 3 神经母细胞瘤组与其它恶性实体肿瘤组 M/Z 位于 3 414.50、4 634.14、4 715.97、6 616.45、9 309.98 的统计资料(N-神经母细胞瘤,D-其它小儿恶性实体肿瘤)

M/Z	Mean-N	SD-N	Mean-D	SD-D	P 值
3 414.497 6	1 093.94	71.78	1 389.41	99.31	0.001 124 396 95
4 634.143 2	849.78	50.23	1 118.43	76.95	0.001 215 354 60
4 715.968 2	222.54	51.95	306.31	71.34	0.001 361 314 96
6 616.445 7	927.91	62.77	1 147.96	99.91	0.001 506 600 05
9 309.984 3	1 191.52	78.93	2 479.14	81.42	0.001 831 067 29

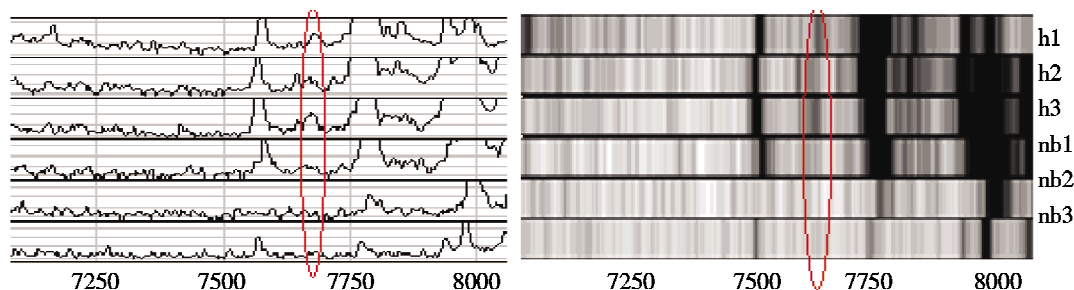


图 1 M/Z 位于 7 693.2 的蛋白质谱差别(nb 是神经母细胞瘤,h 是正常小儿)

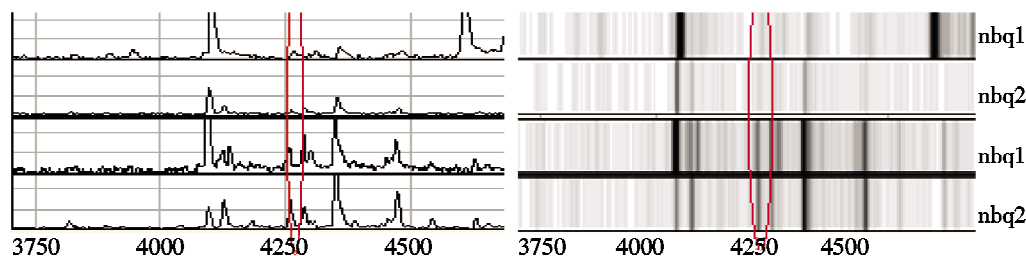


图 2 M/Z 位于 4 244.3 的蛋白质谱差别(nbq 是神经母细胞瘤,nbh 是神经母细胞瘤术后组)

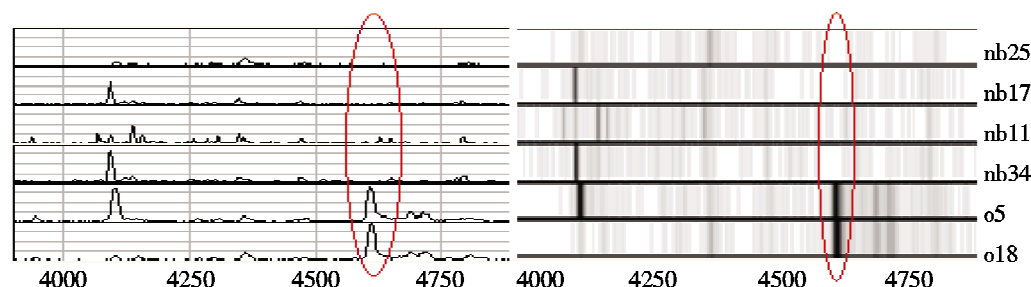


图 3 M/Z 位于 4 634.14 的蛋白质谱差别(nb 是神经母细胞瘤,o 是其它小儿恶性实体肿瘤肿瘤)

讨 论

神经母细胞瘤是小儿时期胚胎性交感神经系统的恶性肿瘤,在儿童肿瘤死亡率中高居第 1 位^[3]。婴儿最多见,约 60%~80% 的病例见于 5 岁以内小儿,其中 11%~41% 见于 1 岁以内婴儿。因肿瘤原发部位隐蔽,早期无特殊症状,易误诊误治,至确诊时多已属晚期。VMA 浓度测定是早期诊断本病的可靠依据之一,B 超与 CT 检查有较好的相关性,确诊须依靠病理检查。早期(I 期)病例 5 年生存率超过 90%,许多患儿仅通过外科手术便可治愈。因此,提高 I 期诊断率,对于降低患儿死亡率具有重要意义。

每种疾病在症状出现之前及不同的发病阶段,都有其独特的蛋白质变化,分析这些早期发生的蛋白质变化有望获得早期诊断的生物标志物^[4]。SELDI-TOF-MS 技术已在肝癌、肺癌、前列腺癌、胰腺癌、食管癌、胃癌、卵巢癌、子宫内膜癌、结(直)肠癌、乳腺癌、甲状腺癌、肾母细胞瘤和泌尿系肿瘤生物标志物筛选方面取得突破性进展^[5-10]。作者通过 SELDI-TOF-MS 和分子生物学相结合应用支持向量机(SVM)分析方法对神经母细胞瘤、小儿实体肿瘤和健康儿童的血清蛋白质组进行筛查,筛选出特异性强、敏感性高的血清肿瘤标记物,在建立的 2 个蛋白指纹图诊断模型中,均有高表达和低表达 2 个候选标志物群,其相对应的功能群可能为癌基因产物的高表达及抑癌基因产物的低表达,从而提示神经母细胞瘤发生发展过程是癌基因的激活、过度表达和抑癌基因的失活、功能受抑制相互作用的结果。由于神经母细胞瘤的基因已经发生了某种改变,必然导致蛋白质表达异常,筛选出异常表达的蛋白质标记物为神经母细胞瘤的早期诊断提供了一种理论依据。本试验通过运用 SELDI-TOF-MS 技术寻找血清标记物,然后用 Wilcoxon 秩和检验分析根据 *P* 值评价各个峰对区分样本的相对重要性。将差异显著的质荷峰随机组合输入 SVM,筛选特异性强、敏感性高的血清标记物,并进行盲法测试。本实验建立的血清蛋白指纹图谱模型成功区分开神经细胞瘤和正常小儿、术前神经母细胞瘤和术后神经母细胞瘤以及神经母细胞瘤与其它恶性实体瘤。因其高效、快速和敏感等特点及特异性(100.00%)优于传统的辅助检查,为神经母细胞瘤的早期诊断和准确分期提供了一种新的途径与方法;还对神经母

细胞瘤早期定性诊断有一定的辅助诊断价值,可弥补神经母细胞瘤辅助检查的不足,指导临床并制定出合理的治疗方案。SELDI-TOF-MS 技术可以满足临床上无创、特异、早期诊断神经母细胞瘤的要求,既克服了患儿在肿瘤很小时难以发现的缺陷,又弥补了临床上 VAM 测定及 IVP 等辅助检查的不足。

支持向量机(SVM)方法是 Vapnik 等^[11]提出的一种分类技术,具有识别小样本模型的推广性,模型选择、过拟和、维数灾难等问题在 SVM 中可得到很好解决^[12]。SVM 预测算法只是基于氨基酸组成预测膜蛋白类型,没有考虑蛋白质多肽链上氨基酸残基的序列特性。在后继工作中,除了可以考虑膜蛋白氨基酸组成外,还可以引入其他组成膜蛋白序列的氨基酸残基之间的特征,如疏水性、带电荷性、分子间相互作用力等特性;但因样本收集有限,仅检测出能够反映早期血清蛋白质谱变化的模型,以后需继续扩大样本量以臻完善。

参 考 文 献

- Guo Y, Yu L. Using support vector machine combined with auto covariance to predict protein-protein interactions from protein sequences[J]. *Nucleic Acids Res*, 2008, 1-15.
- Jorissen RN, Gilson MK. Virtual screening of molecular databases using a support vector machine[J]. *Chem Inf Model*, 2005, 45(3): 549-561.
- 陈新让,王家祥,余捷凯,等. 基于支持向量机神经母细胞瘤血清蛋白质标记物的检测及临床应用[J]. *临床小儿外科杂志*, 2008, 12(2): 17-22.
- Liotta LA, Ferrari M, Petricoin E. Clinical proteomics: Written in blood[J]. *Nature*, Oct, 2003, 425 (6961): 905.
- 王忠荣,刘胜勇. 儿童神经母细胞瘤的治疗[J]. *临床小儿外科杂志*, 2009, 8(4): 62.
- Zhang Bo, Suyongping, Wangfengchao, et al. Identification of differentially expresses proteins of gamma ray irradiated rat intestinal epithelial IEC6 cells by two dimension algal electrophoresis and matrix assisted laser sorption/ionization time of flight mass spectrometry[J]. *Proteomics*, 2005, 5(2): 426-432.
- Bjorhall K, Miliotis T, Davidsson P. Comparison of different depletion strategies for improved resolution in proteomic analysis of human serum samples[J]. *Proteomics*, 2005, 5(1): 307-317.
- Dai J, Zhang Y, Wang J, et al. Identification of degradation products formed during per formic oxidation of peptides and proteins by high-performance (下转第 336 页)