

研究报告

# 由质谱分析数据建立肝癌病人与健康人血清的分类模型

徐 琨<sup>1</sup>; 朱尔一<sup>1</sup>; 杨芑原<sup>2, 4</sup>; 刘银坤<sup>3, 4</sup>

1. 厦门大学化学化工学院, 现代分析科学教育部重点实验室, 福建 厦门 361005

2. 复旦大学化学系, 上海 200433

3. 复旦大学附属中山医院, 肝癌研究所, 上海 200032

4. 复旦大学生物医学研究院, 上海 200032

收稿日期 2008-3-13 修回日期 2008-5-23 网络版发布日期:

**摘要** 通过对质谱数据处理和分析, 建立肝细胞癌(HCC)病人与健康人的分类模型, 可以用来判别样本是否来源于肝癌病人。运用表面加强激光解析电离-飞行时间质谱(SELDI-TOF-MS)获取肝癌病人和健康人血清蛋白指纹图谱数据, 并运用偏最小二乘(PLS)变量筛选法建立分类模型, 最终得到分类模型的交叉检验相关系数达0.96以上, 判别准确率大大提高。同时对模型进行分析, 找出对肝癌病人和健康人的差异有重要影响的因素或变量。这些变量为30个质荷比区间内特定蛋白的峰强度值, 反映这些质荷比区间内蛋白量的增加或减少与肝癌的形成有密切关系, 可作为重要的生物标志物进一步加以研究。并且采用所得模型的拟合值等一些信息做分类图, 能较好地表达回归模型的分类效果。

**关键词** [肝细胞癌](#) [表面加强激光解析电离-飞行时间质谱](#) [偏最小二乘变量筛选](#) [分类模型](#) [海量数据建模](#)

分类号 [O 657.63](#)

## Classification Model of HCC Patients and Healthy People Established from Mass Spectrometry Data

XU Kun<sup>1</sup>; ZHU Er-yi<sup>1</sup>; YANG Peng-yuan<sup>2, 4</sup>; LIU Yi n-kun<sup>3, 4</sup>

1. Key Laboratory of Analytical Sciences of the Ministry of Education,

College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, China; 2. Department of Chemistry, Fudan University, Shanghai 200433, China;

3. Liver Cancer Institute, Zhongshan Hospital, Shanghai 200032, China;

4. Institute of Biomedical Sciences, Fudan University, Shanghai 200032, China

**Abstract** Classification models of hepatocellular carcinoma (HCC) patients and healthy people were built from mass spectrometry data, which could be used for the detection of HCC. Surface-enhanced laser desorption and ionization time-of-flight mass spectrometry (SELDI-TOF-MS) technique was applied to get the data of serum protein from HCC patients and healthy people. Then PLS variable selection method was used to deal with the data and to establish the classification model. The cross validation relativity coefficients of the model comes to over 0.96. Furthermore the important factors or variables that discriminated HCC patients and healthy people were found by analyzing the model. The 30 variables were several peak intensities of protein from some m/z sections, which could express the up-regulation or down-regulation of protein in the sections. As potential biomarkers, the proteins may be closely related to the formation of HCC, which can be deeply studied. The classification figures constructed by the fitting value of the model in the article are clear and intuitive, and can express the discrimination effect of the model well.

### 扩展功能

#### 本文信息

▶ [Supporting info](#)

▶ [\[PDF全文\]\(154KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [参考文献](#)

#### 服务与反馈

▶ [把本文推荐给朋友](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

#### 相关信息

▶ [本刊中 包含“肝细胞癌”的 相关文章](#)

▶ 本文作者相关文章

· [徐 琨](#)

· [朱尔一](#)

· [杨芑原](#)

·

· [刘银坤](#)

·

**Key words** [HCC](#) [SELDI-TOF-MS](#) [PLS](#) [variable selection](#) [classification](#) [model](#)  
[el](#) [huge](#) [data](#) [modeling](#)

DOI

---

通讯作者 朱尔一 [ryzhu@xmu.edu.cn](mailto:ryzhu@xmu.edu.cn)