

## 摘要

蛋白质翻译后修饰是蛋白质组学的重要研究内容，目前，基于串联质谱的蛋白质序列数据库搜索是鉴定翻译后修饰的主要方法。然而，随着修饰研究的深入，对多修饰鉴定、低丰度修饰检测与未知修饰发现等提出了更高的要求。本文在分析现有修饰检测及鉴定算法的不足以及修饰谱图特点的基础上，设计并实现了一种有效的修饰鉴定算法，对于修饰研究提供了基础性的技术支持。

本文的核心是设计了一种基于点积法的用于修饰、非修饰谱之间的打分方法，针对其他方法在高电荷谱图的处理上的缺陷本文给出了解决方法。这种方法充分考虑到了高电荷谱图中碎片离子峰电荷数目的多种情况，给出最优的谱峰匹配结果。这种方法能够从谱图数据中获得更多的修饰非修饰谱图对，从而发现更多的修饰类型，对微量修饰的发现具有重要的意义。这种方法还可以发现氨基酸突变以及非特异酶切肽段等情况。本文在两个标准蛋白数据和一个岩藻糖蛋白实际样品数据中对方法进行实验，发现了多种修饰类型，其中包括一些微量修饰和一些鲜有报道的修饰类型。还发现了多种氨基酸突变的情况。在鉴定了修饰之后，所有数据的谱图解析率均提高了 50% 以上。

本文开发了相应的修饰发现及鉴定软件 pCluster。pCluster 同时集成了利用串联质谱谱图聚类进行修饰发现的功能和利用一级谱进行修饰发现的算法 DeltAMT。它可以辅助用户使用这两种不同类型的方法做修饰研究。

另外，本文还开发了一款串联质谱谱图批量标注软件 pLabel。谱图标注软件是基于串联质谱技术的蛋白质组学研究中的常用的软件。pLabel 满足了研究人员对于自动标图的需要，为研究带来便利。目前，pLabel 已经在国内外一些主要蛋白质组学实验室发挥着辅助质谱分析的作用。

**关键词：** 串联质谱技术，谱图聚类，翻译后修饰，高精度质谱数据