扩展功能

本文信息

- Supporting info
- PDF(1116KB)
- [HTML全文]
- 参考文献[PDF]
- 参考文献

服务与反馈

- 把本文推荐给朋友
- 加入我的书架
- 加入引用管理器
- 引用本文
- Email Alert
- 文章反馈
- 浏览反馈信息

本文关键词相关文章

- 蛋白鉴定
- 胰蛋白酶肽段
- 酶切/漏切肽段
- 马尔科夫链
- 概率预测

本文作者相关文章

PubMed

研究论文

基于马尔科夫链的胰蛋白酶肽段蛋白酶切位点概率预测

刘辉[1], 张纪阳[1], 孙汉昌[1], 徐长明[1], 张伟[1], 马海滨[1], 朱云平[2], 谢红卫[1]

1. 国防科学技术大学机电工程与自动化学院，长沙 410073;
2. 军事医学科学院放射与辐射医学研究所，北京蛋白质组研究中心，蛋白质组学国家重点实验室，北京 102206

摘要：

在基于质谱技术的蛋白质鉴定过程中，数据库搜索是主要的方法。漏切位点和酶切规则决定了图谱候选肽段的范围，是数据库搜索算法的重要参数。对于常用的胰蛋白酶切来说，除了局部构象、三维结构、实验条件，以及其它偶然因素会影响赖氨酸K或者精氨酸R后的位点能否被酶切外，该位点附近的其它氨基酸也会影响蛋白水解酶的酶切效果。从质谱图谱中时常会鉴定出包含漏切位点的肽段，因此，预测蛋白质的酶切位点能够为数据库搜索算法提供更为可靠的模型，也能够为了解和分析蛋白质的酶切规律提供依据。本文提出了一种基于马尔科夫(Markov)链的预测方法，能够利用蛋白质的序列信息来预测候选酶切位点的酶切概率，在蛋白酶切过程中，预测肽段的覆盖率可以达到85%以上。

关键词： 蛋白鉴定 胰蛋白酶肽段 酶切/漏切肽段 马尔科夫链 概率预测

The Probabilistic Prediction of Protein Cleavage Site for Tryptic Peptides Based on Markov Chain

LIU Hui[1], ZHANG Jiyang[1], SUN Hanchang[1], XU Changming[1], ZHANG Wei[1], MA Haibin[1], ZHU Yunping[2], XIE Hongwei[1]

1. College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha 410073, China;
2. State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Radiation Medicine, Beijing 102206, China

Abstract:

Database search is the main method in protein identification based on mass spectrometry, and missed cleavage sites and the rules of digestion are important parameters of database search algorithms that determine the scope of the candidate peptides. Besides many factors such as local conformation, tertiary structure, experimental conditions and causal interference during the trypsin digestion, the effect of trypsin is also likely to be influenced by other residues around the cleavage site whether or not the sites after lysine or arginine could be digested. The tryptic peptides including missed cleavage sites are often identified from mass spectrum, so the prediction of cleaved sites can provide a more reliable model for database search algorithms and for further understanding and analyzing the regular pattern of the protein digestion for the trypsin. A predictive method based on Markov chain is presented to be able to predict the cleaved probability of candidate cleavage sites for tryptic peptides from the amino acid sequence of protein with the accuracy up to 85%.

Keywords: Protein identification Tryptic peptides Cleaved/missed peptides Markov chain Probabilistic prediction

参考文献：

1. Delahunty C, Yates JR 3rd. Protein identication using 2D-LC-MS/MS. Methods, 2005, 35: 248~255

2. Yates JR 3rd, Ruse CI, Nakorchevsky A. Proteomics by mass spectrometry: Approaches, advances, and applications. Annu Rev Biomed Eng, 2009, 11: 49~79

3. Yates JR 3rd, McCormack AL, Link AJ, Schieltz D, Eng J, Hays L. Future prospects for the analysis of complex biological systems using micro-column liquid chromatography-electrospray tandem mass spectrometry. Analyst, 1996, 121(7): 65~76

4. Gevaert K, Vandekerckhove J. Protein identification methods in proteomics. Electrophoresis, 2000, 21(6): 1145~1154

5. Enga JK, McCormacka AL, Yates JR 3rd. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectr, 1994, 5(11): 976~989

6. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis, 1999, 20(18): 3551~3567

7. Siepen JA, Keevil EJ, Knight D, Hubbard SJ. Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. J Proteome Res, 2007, 6(1): 399~408

8. Olsen JV, Ong SE, Mann M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. Mol Cell Proteomics, 2004, 3(6): 608~614

9. Hubbard SJ. The structural aspects of limited proteolysis of native proteins. Biochim Biophys Acta, 1998, 1382: 191~206

10. Schechter I, Berger A. On the size of the active site in proteases. I. Papain. Biochem Biophys Res Commun, 1967, 27(2): 157~162

11. Monigatti F, Berndt P. Algorithm for accurate similarity measurements of peptide mass fingerprints and its application. J Am Soc Mass Spectrom, 2005, 16(1): 13~21

12. Yen CY, Russell S, Mendoza AM, Meyer-Arendt K, Sun S, Cios KJ, Ahn NG, Resing KA. Improving sensitivity in shotgun proteomics using a peptide-centric database with reduced complexity: Protease cleavage and SCX elution rules from data mining of MS/MS spectra. Anal Chem, 2006, 78(4): 1071~1084

13. Thiede B, Lamer S, Mattow J, Siejak F, Dimmler C, Rudel T, Jungblut PR. Analysis of missed cleavage sites, tryptophan oxidation and N-terminal pyroglutamylation after in-gel tryptic digestion. Rapid Commun Mass Spectrom, 2000, 14(6): 496~502

14. McLaughlin T, Siepen JA, Selley J, Lynch JA, Lau KW, Yin H, Gaskell SJ, Hubbard SJ. PepSeeker: A database of proteome peptide identifications for investigating fragmentation patterns. Nucleic Acid Res, 2006, 34(Database issue): D649~654

15. Gattiker A, Bienvenut WV, Bairoch A, Gasteiger E. FindPept, a tool to identify unmatched masses in peptide mass fingerprinting protein identification. Proteomics, 2002, 2(10): 1435~1444

16. Falkner JA, Kachman M, Veine DM, Walker A, Strahler JR, Andrews PC. Validated MALDI-TOF/TOF mass spectra for protein standards. J Am Soc Mass Spectrom, 2007, 18(5): 850~854

17. Craig R, Beavis RC. TANDEM: Matching proteins with tandem mass spectra. Bioinformatics, 2004, 20(9): 1466~1467

本刊中的类似文章

文章评论

反馈人		邮箱地址	
反馈标题		验证码	7921