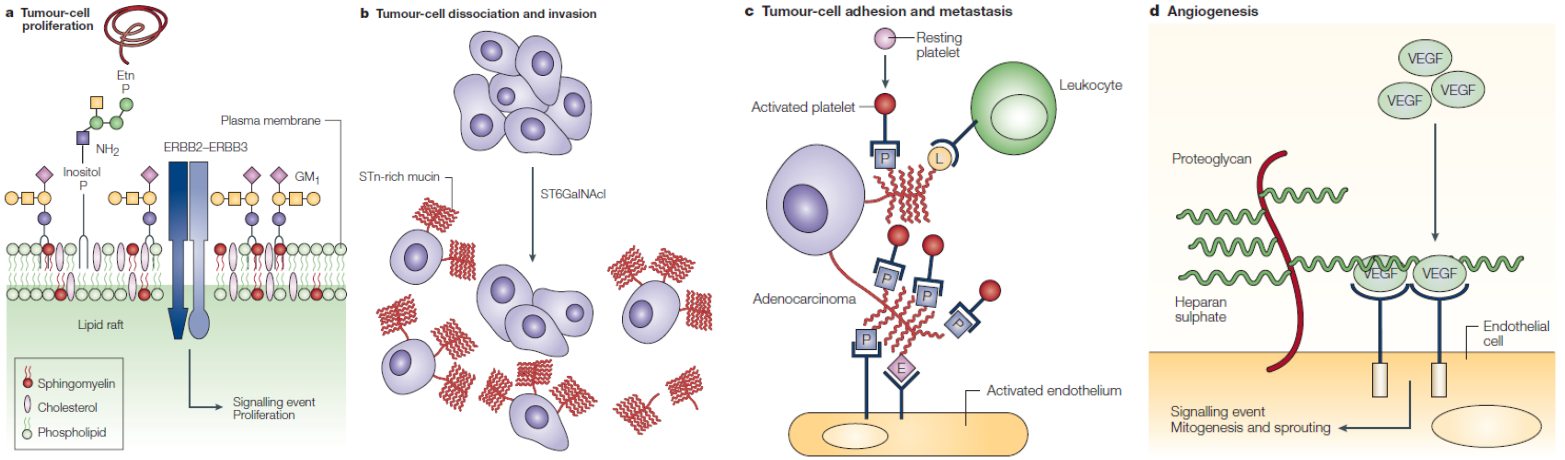
核心岩藻糖化蛋白质特异性发掘的系统解决方案

应万涛 钱小红

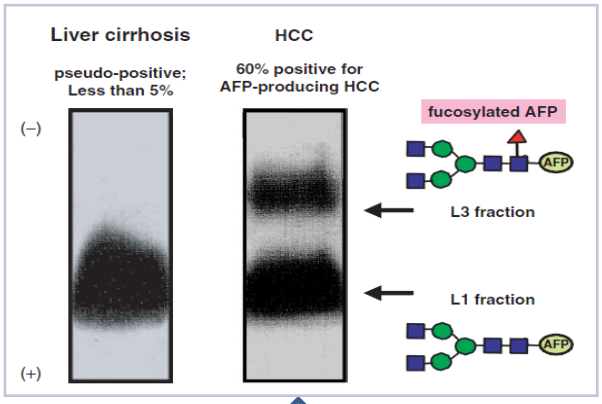
蛋白质组学国家重点实验室
北京蛋白质组研究中心
北京放射医学研究所

2010-11-10 北京 中国科学院计算技术研究所

糖复合物参与复杂的生理病理调控



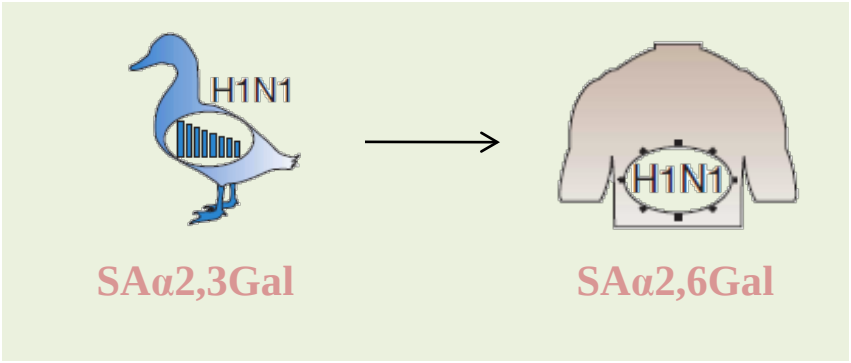
Nature Review Cancer, 2005, 5, 526



AFP-L3 used as a HCC biomarker

PSA, CEA, CA125, CA19-9

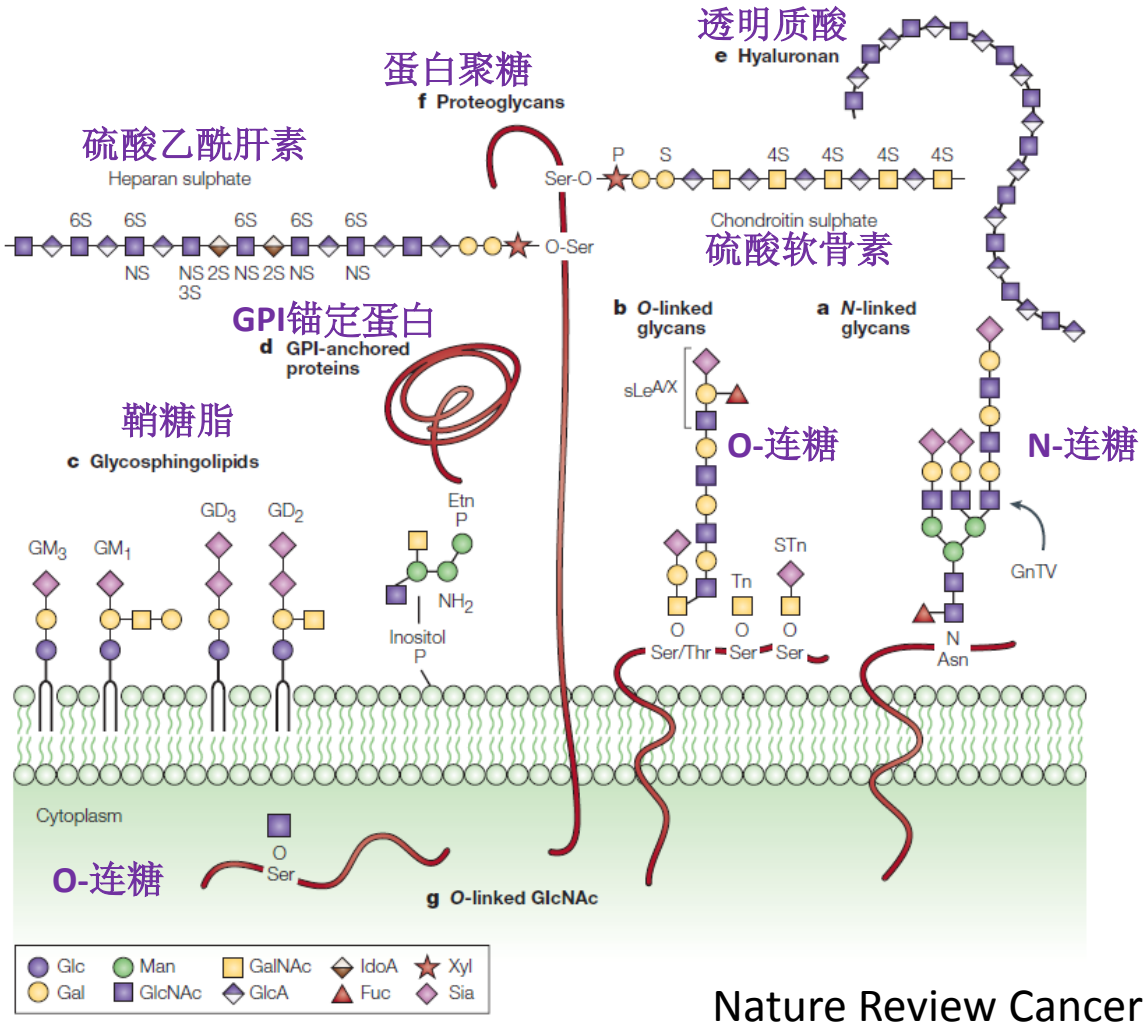
Miyoshi E et al. J. Biochem., 2009, 143: 725–9



Switching of HA receptor specificity promote the transmission of H1N1

Neumann G et al. Nature, 2009, 459: 931-9

糖基化修饰反应类型

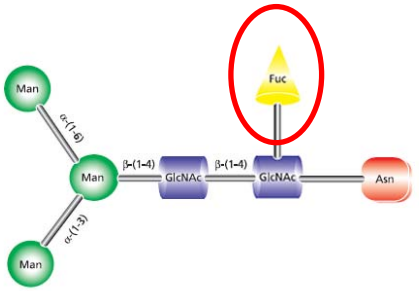


Nature Review Cancer 2005

细胞表面，细菌细胞壁，细胞核/胞浆，
胞外基质，分泌蛋白质

50%蛋白质
为糖蛋白

核心岩藻糖化修饰与癌症发生



肝癌

Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans. ***Proc. Natl. Acad. Sci. U. S. A.*** (2005)

Proteomic Analysis of Serum Associated Fucosylated Glycoproteins in the Development of Primary Hepatocellular Carcinoma. ***J. Proteome. Res.*** (2006)

胰腺癌

Fucosylated haptoglobin is a novel marker for pancreatic cancer: a detailed analysis of the oligosaccharide structure and a possible mechanism for fucosylation. Int. J. Cancer, (2006)

Glycosylation of serum ribonuclease 1 indicates a major endothelial origin and reveals an increase in core fucosylation in pancreatic cancer, ***Glycobiology*** (2007)

肺癌

Dysregulation of TGF-β1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose deficient mice. ***Proc. Natl. Acad. Sci. U. S. A.*** (2005)

The expression of core fucosylated E-cadherin in cancer cells and lung cancer patients: prognostic implications, ***Cell Res.*** (2004)

卵巢癌

Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG, ***Glycobiology,*** (2007)

前列腺癌

Different glycan structures in prostate-specific antigen from prostate cancer sera in relation to seminal plasma PSA. ***Glycobiology*** (2006)

糖蛋白的核心岩藻糖化水平变化与多种癌症相关

甲胎蛋白AFP-肝癌诊断标志物

80年代初发现AFP核心岩藻糖化在肝的不同病理条件升高

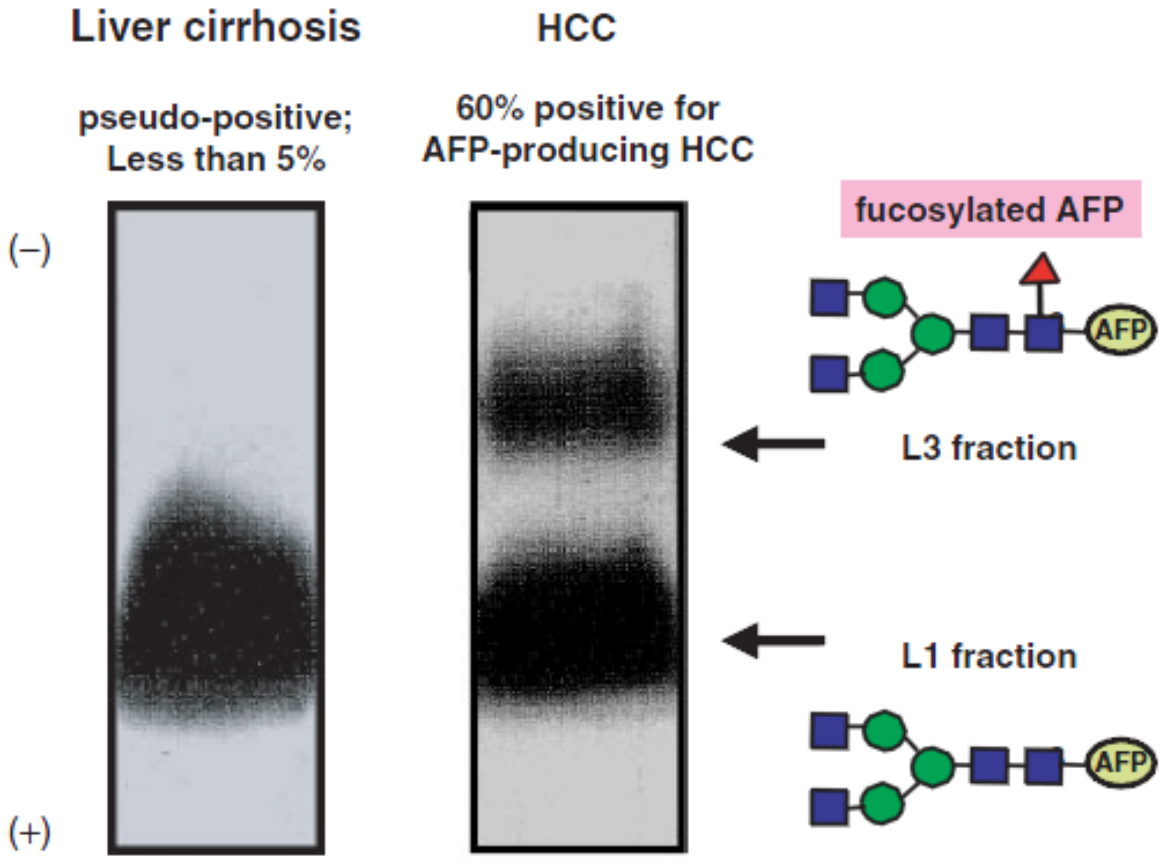
90年代发现AFP-L3具有更好的特异性

2005年FDA批准AFP-L3为HCC的生物标志物

浓度：20 ng/ml

敏感性：50%

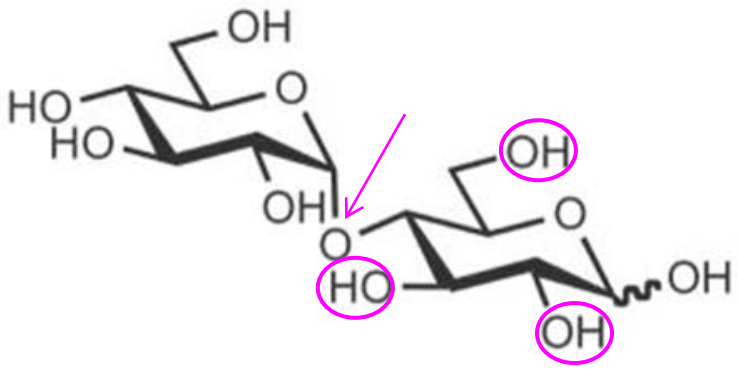
特异性：70%



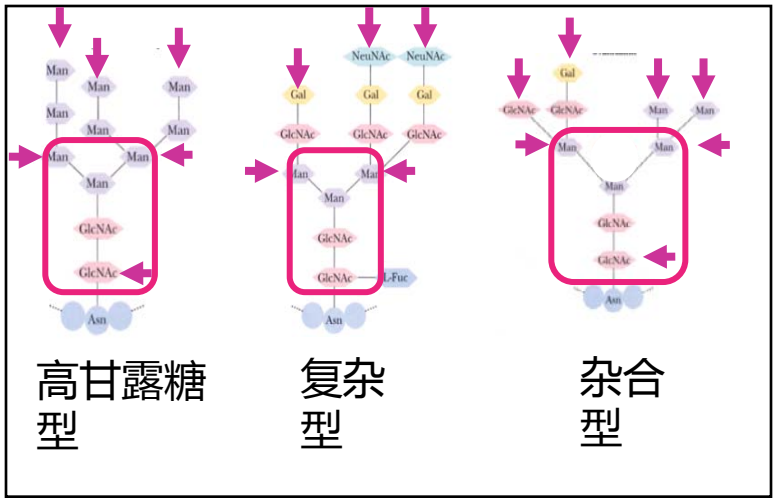
J. Biochem. 143, 725–729 (2008)

糖基化修饰研究挑战

- 糖链种类复杂
- 糖链存在微不均一性→糖蛋白在胶上呈弥散性分布；一种方法只能富集部分糖蛋白
- 蛋白质酶切后仅2-5%是糖基化肽段→信号被淹没→富集手段的必要性
- 糖肽在质谱中检测困难→糖链优于肽段先断裂；部分单糖呈酸性；糖链质谱图复杂→质谱鉴定需要切去糖链
- 缺乏完善的规模化糖谱解析平台



Glcα1-4Glc
(maltose)

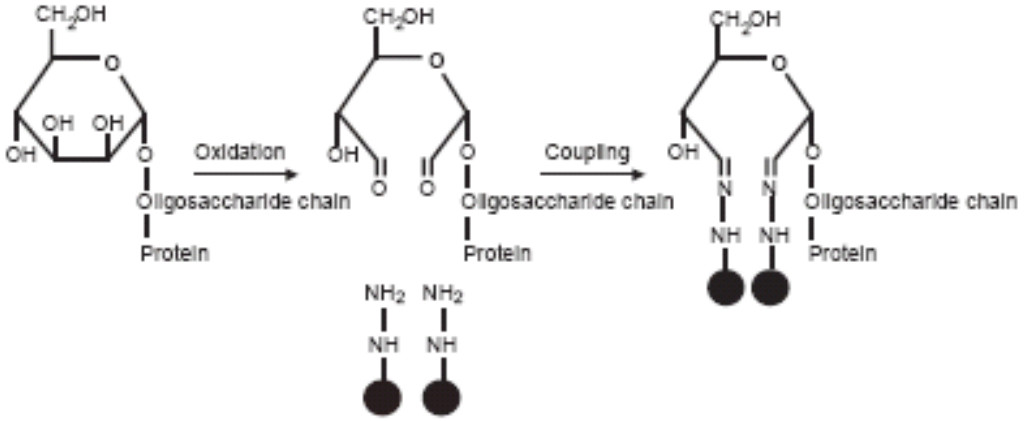


高甘露糖型　复杂型　杂合型

糖基化修饰研究技术

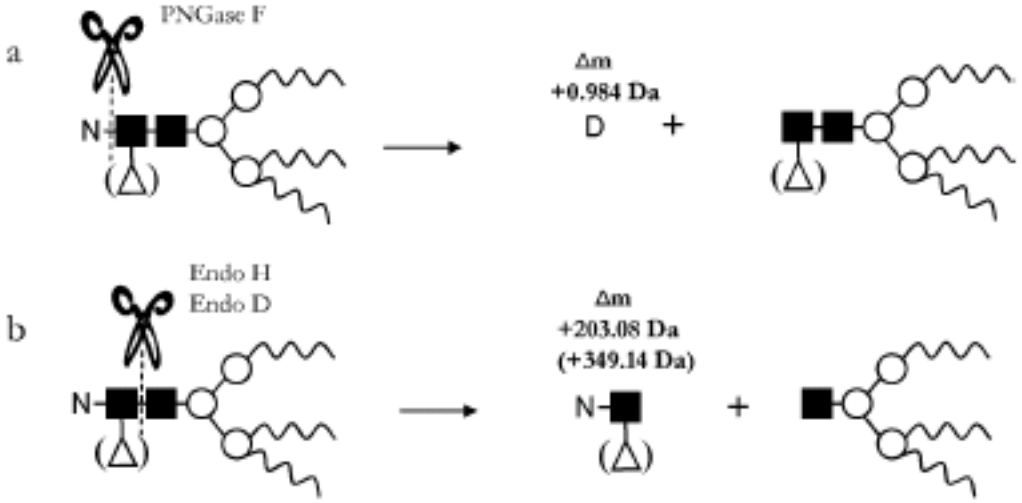
- 富集方法

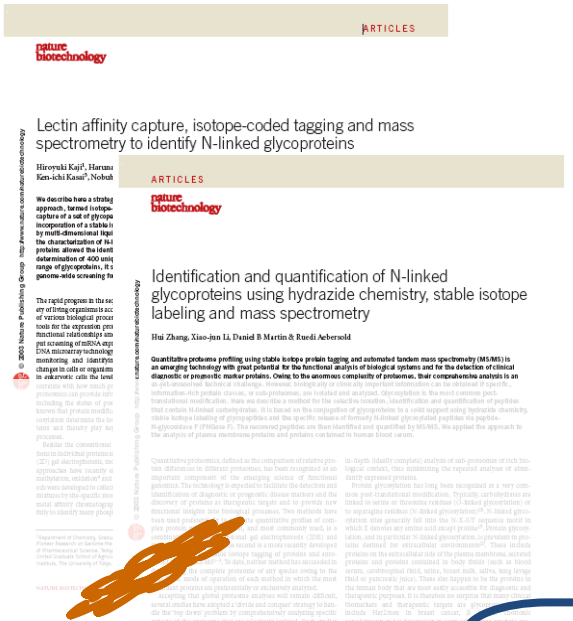
①凝集素亲和层析
②亲水相互作用色谱
③分子排阻层析
④硼酸化合物富集
⑤肼解化学富集方法



- 切糖方法

①PNGase F
②Endo H/D
③Pronase




ARTICLES

nature biotechnology

Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins

ARTICLES

nature biotechnology

Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry

糖基化蛋白质组学研究的经典方法

富集：凝集素、酰肼

糖蛋白富集

蛋白酶切

糖肽富集

PNGase F

完全去除糖链

质谱鉴定

糖蛋白身份信息

糖基化位点信息

核心岩藻糖化修饰研究新思路

Use of targeted glycoproteomics to identify serum glycoproteins that correlate with liver cancer in woodchucks and humans

research articles Journal of proteome research

Proteomic Analysis of Serum Associated Fucosylated Glycoproteins in the Development of Primary Hepatocellular Carcinoma

Mary Ann Comunale†, Melissa Lowman,† Ronald E. Long,† Jonathan Krakover,‡,§ Ramila Philip,‖ Steven Seeholzer,‖ Alison A. Evans,‖ Hie-Won L. Hann,# Timothy M. Block,† and Anand S. Mehta*,†

Drexel Institute for Biotechnology and Virology Research, Department of Microbiology and Immunology, Drexel University College of Medicine, Doylestown, Pennsylvania 18902, The Institute for Virus and Hepatitis Research, Pennsylvania Commonwealth Institute, Doylestown, Pennsylvania 18902, Immunotope Inc., Doylestown, Pennsylvania 18901, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, and Division of Gastroenterology and Hepatology, Thomas Jefferson University, Philadelphia, Pennsylvania 19107

Received September 29, 2005

Changes in N-linked glycosylation are known to occur during the development of cancer. For example, increased branching of oligosaccharides has been associated with metastasis and has been correlated to tumor progression in human cancers of the breast, colon and melanomas. Increases in core fucosylation have also been associated with the development of hepatocellular carcinoma (HCC). Chronic infection with the hepatitis B virus is associated with more than 55% of all cases of hepatocellular carcinoma. We show here that increased levels of core fucosylation can be observed via glycan analysis of total serum and are associated with the development of HCC. In a blinded study, the serum glycoproteins derived from people diagnosed with HBV induced liver cancer were found to possess a dramatically higher level of fucosylation. This change occurs on both immunoglobulin molecules and on other serum glycoproteins. Targeted glycoproteomic analysis was used to identify those glycoproteins that are hyperfucosylated in cancer. In total, 19 proteins were found to be hyperfucosylated in cancer. The potential of these proteins as biomarkers of cancer is discussed.

Keywords: hepatocellular carcinoma • fucosylation • glycomics • biomarkers

Introduction

Infection with hepatitis B virus (HBV) and/or hepatitis C virus (HCV) is the major etiology of hepatocellular cancer (HCC).[1–3] Both HBV and HCV cause acute and chronic liver infections and most chronically infected individuals remain asymptomatic for many years. Clinical disease (HCC and cirrhosis) is not apparent until decades later. 10–40% of all chronic HBV carriers eventually develop liver cancer, and it is estimated that over one million people worldwide die because of HBV/HCV associated liver cancer.[4] Indeed, HBV and HCV infections are associated with over 80% of all HCC cases worldwide and can be as high as 98% in regions where HBV is endemic.[5]

The chronic infection of HBV or HCV is asymptomatic and monitored by regular (usually annual or biannual) physical assessments, serum liver function tests (LFTs) and ultrasound imaging for detection of small masses in the liver (Hepatitis B foundation, 1994). The effectiveness of ultrasound is limited, since the appearance of masses at least 3 cm in size are required for detection, and this often occurs at a stage when the prognosis is poor.[6,7] The correlation between elevated serum concentrations of alpha feto protein (AFP) and the occurrence of HCC has provided a useful surrogate marker for disease.[8] Levels of AFP exceeding 50 ng/mL occur in 30–60% of HCC cases at the time of diagnosis.[9,10] However, AFP levels may fluctuate in chronically infected individuals and AFP is a poor biomarker for small tumors.[5] As early surgical and chemotherapeutic intervention is an afflicted individual's best hope, there is a clear and urgent need for noninvasive, reliable methods of detecting HCC as early as possible.[11–13]

Our previous work has utilized a targeted glycoproteomic approach in an effort to identify serum glycoproteins that correlate with the development of cancer. That is, in an animal model of virally induced liver cancer (woodchuck), total serum glycan analysis revealed a correlation between the development of HCC and an increase in the level of core fucosylation.[14] Proteomic analysis of the fucosylated proteome revealed a protein, GP73, which was increased in animals with HCC. This protein was further analyzed in human samples and has shown to be a better marker of HCC than the currently used marker AFP.[15,16] As this methodology proved successful in an animal model of HCC, it was of great interest to see if even more proteins could be identified in human samples. Thus, in this

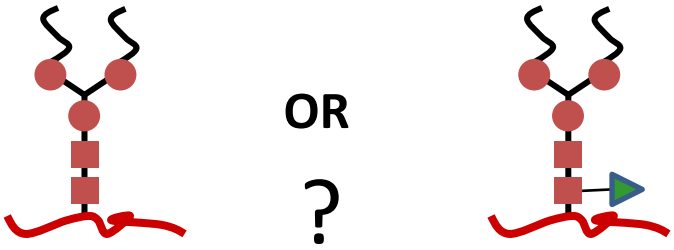
*To whom correspondence should be addressed. E-mail: anand.mehta@drexel.edu.
† Drexel University College of Medicine.
‡ Pennsylvania Commonwealth Institute.
§ Immunotope Inc.
‖ Fox Chase Cancer Center.
Thomas Jefferson University.

308 Journal of Proteome Research 2006, 5, 305–315
Published on Web 01/24/2006

10.1021/pr050320n CCC: $33.50 © 2006 American Chemical Society

采用经典糖蛋白质组学研究方法。Block 等人在肝癌研究中（PNAS，2005）发现了5种核心岩藻糖化上调蛋白。

采用糖蛋白质组学方法，Comunale 等人（JPR, 2006）发现了19种CF上调蛋白。并认为这些CF蛋白联合作为诊断标志物较AFP更为可信。

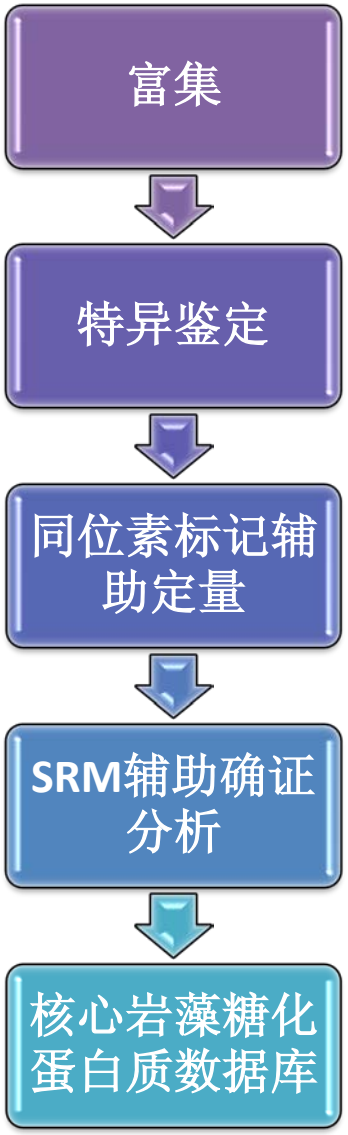


- 保留核心岩藻糖化修饰的直接证据
- 针对上述结果设计特异性质谱采集方案
- 糖肽谱图的自动化筛选及鉴定的软件系统

目标：

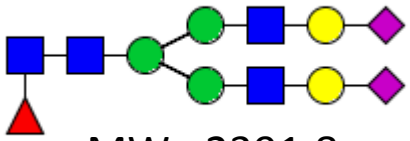
精确地、规模化定性及定量鉴定与确认。

方法：



一、糖肽富集方法优化

策略：利用凝集素在蛋白水平具有较好的特异性，与糖肽本身具有较大分子量的特点，采用**两步富集**策略相对特异富集核心岩藻糖化肽段。



MW：2391.8

蛋白混合物

凝集素富集

核心岩藻糖化蛋白

胰蛋白酶切

肽段混合物

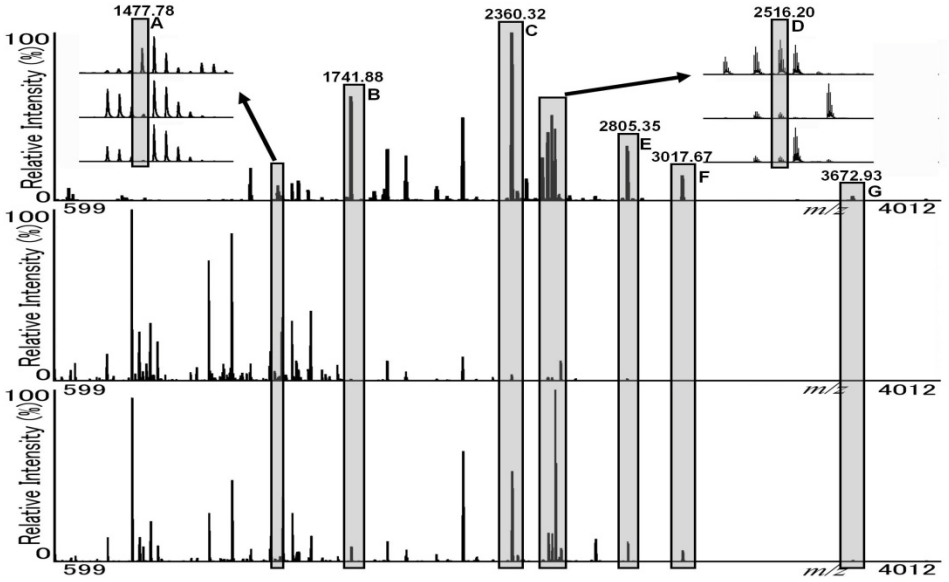
分子量超滤

糖肽混合物

滞留组分
MW > 3000Da

滤过组分
MW < 3000Da

未处理样品



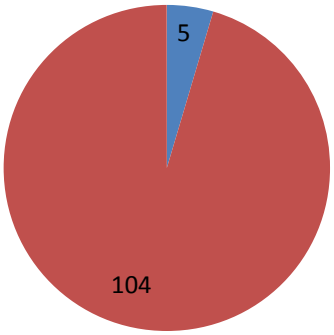
优点：操作简单，使实验容易重复。

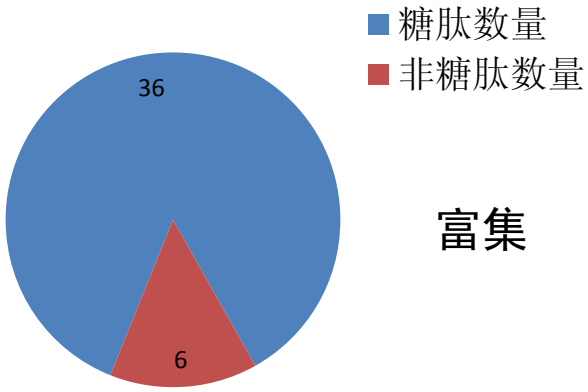
血浆复杂样本测试

正常人血浆去除IgG和HSA后，用胰蛋白酶进行全酶切，酶切的肽段混合物采用SCX进行分离。选取其中4-8分钟的一个馏分，取一半样品用膜超滤方法富集，另一半样本则不做任何处理，PNGase F切除糖链后用RPLC-ESI-LTQ-FT-MS鉴定肽段。

采用SEQUEST鉴定，swissprot人库，肽段假阳性率<1%

不富集

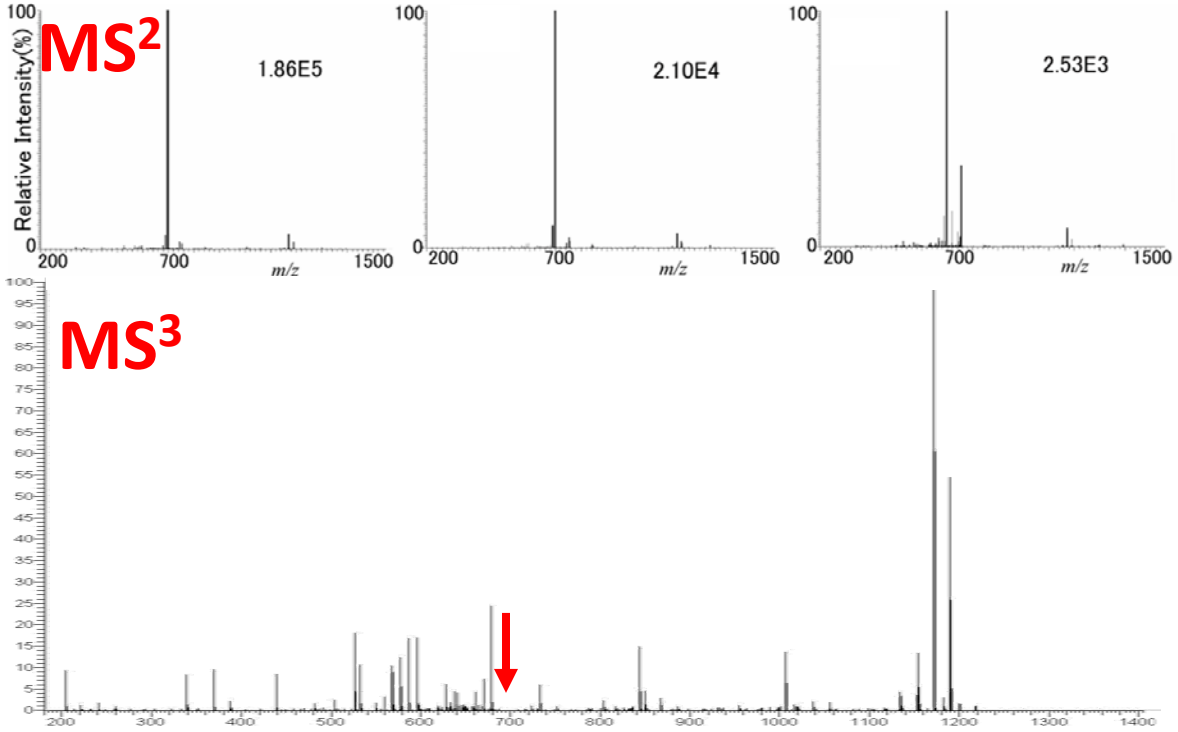




富集

糖肽选择性>85%

二、中性丢失依赖的三级质谱方法



提高了碎片信息质量，而且减少了所需搜索图谱数，因此提高数据分析效率。

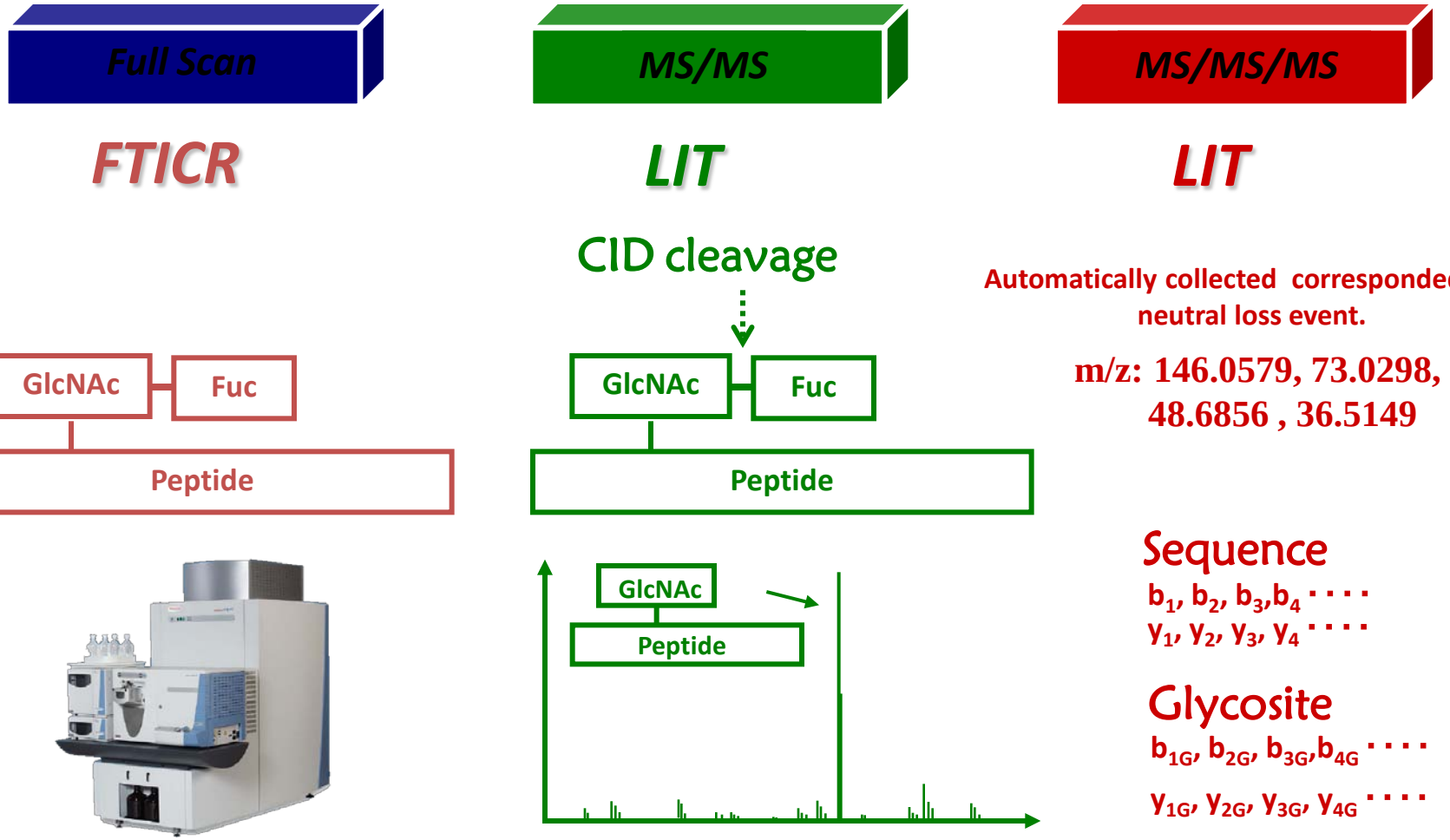
MS^3较MS^2的优点：

一、图谱质量高：峰分布大为均一，峰信噪比好。

二、图谱信息简单：不需要考虑岩藻糖残基产生的碎片信息。

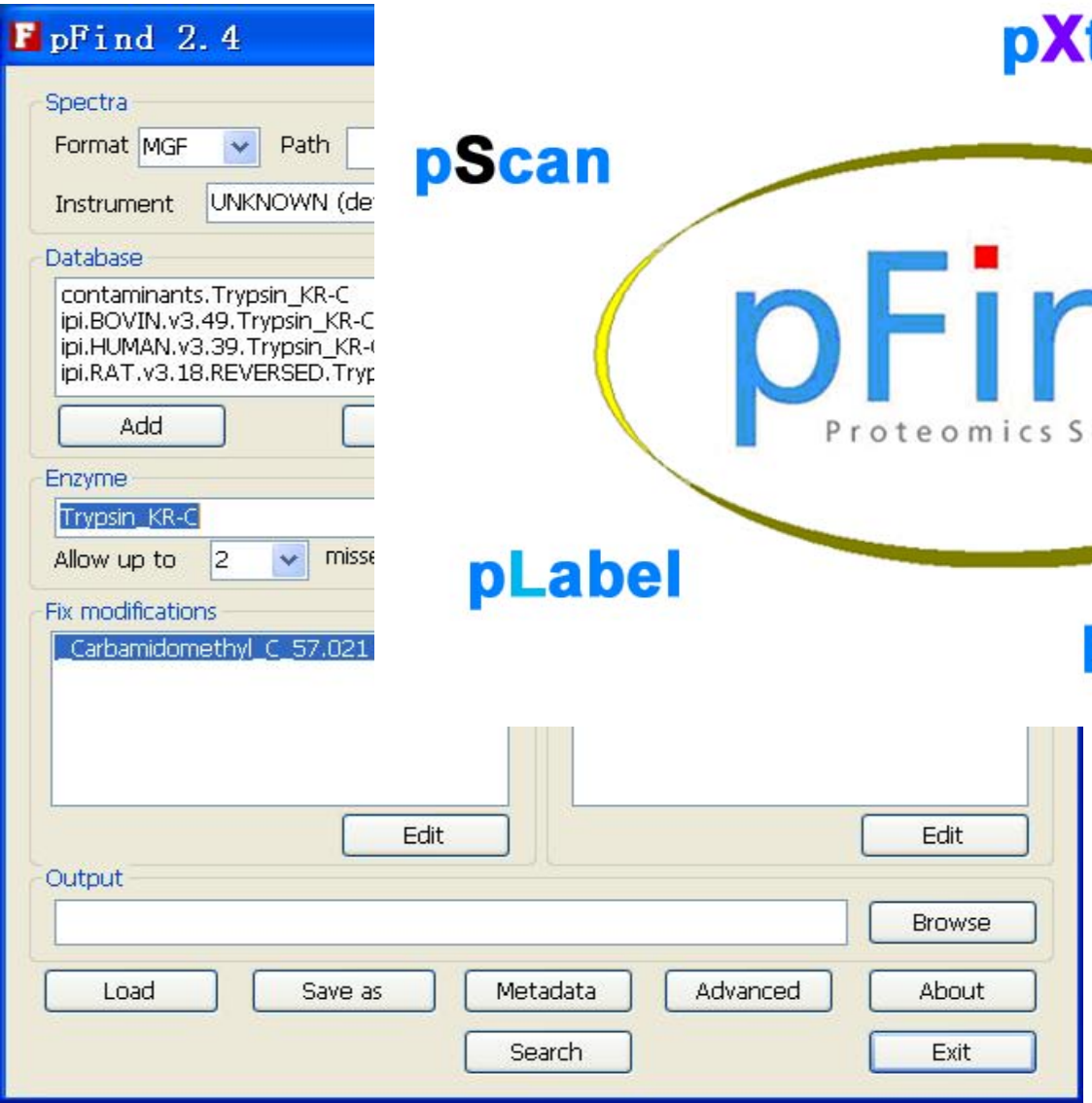
三、干扰峰出现概率小：通过两次母离子选择，一级中相邻干扰峰在三级中产生碎片的可能性极小。

中性丢失依赖的三级质谱采集方法



使用LTQ-FT型质谱仪，一级扫描在FT中进行，使得到高精度母离子质量信息，而串联质谱扫描在离子阱中完成，利用了其高扫描速度和灵敏度的特点。

pFind软件方案



pXtract

pScan

pBuild

pFind

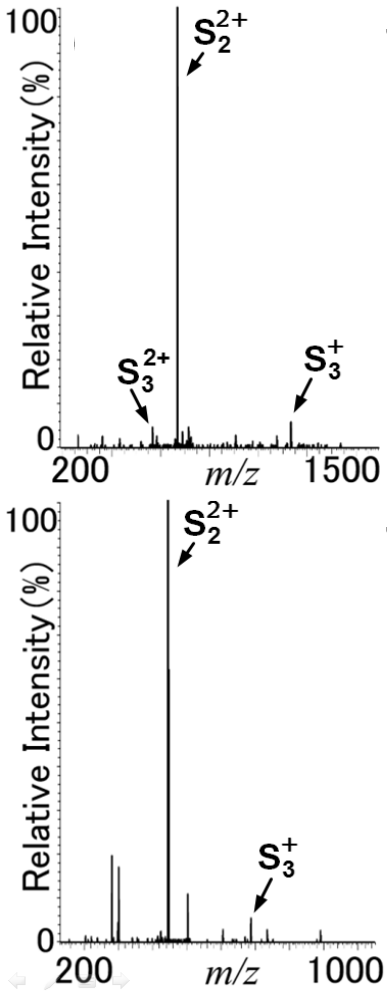
Proteomics Solutions

pLabel

pCluster

不依赖数据库的候选图谱挑选

手工分析了9800张图谱后，建立不依赖数据库的核心岩藻糖化糖肽图谱挑选规则。



一、计算标志离子质量

S_1 = Peptide + 349.14 Da = Parent ion

S_2 = Peptide + 203.08 Da

S_3 = Peptide

二、初步筛选

MS^2中最强峰必须为S_2离子(0.5m/z误差内)，且与母离子电荷相同。

三、根据标志离子信号进行打分

∑≥9的MS^2对应的MS^3图谱作为候选图谱

Ratio of SSP to S_2	Error of S_2	Ratio of SSP ' to S_3
⩽ 20% (+5)	⩽ 0.2m/z (+5)	⩽ 50% (+5)
⩽ 30% (+3)	⩽ 0.3m/z (+3)	⩽ 70% (+3)
⩽ 40% (+2)	⩽ 0.4m/z (+2)	⩽ 150% (+2)
⩽ 50% (+1)	⩽ 0.5m/z (+1)	> 150% (+1)
> 50% (-1)		Disappear (-1)

优　点

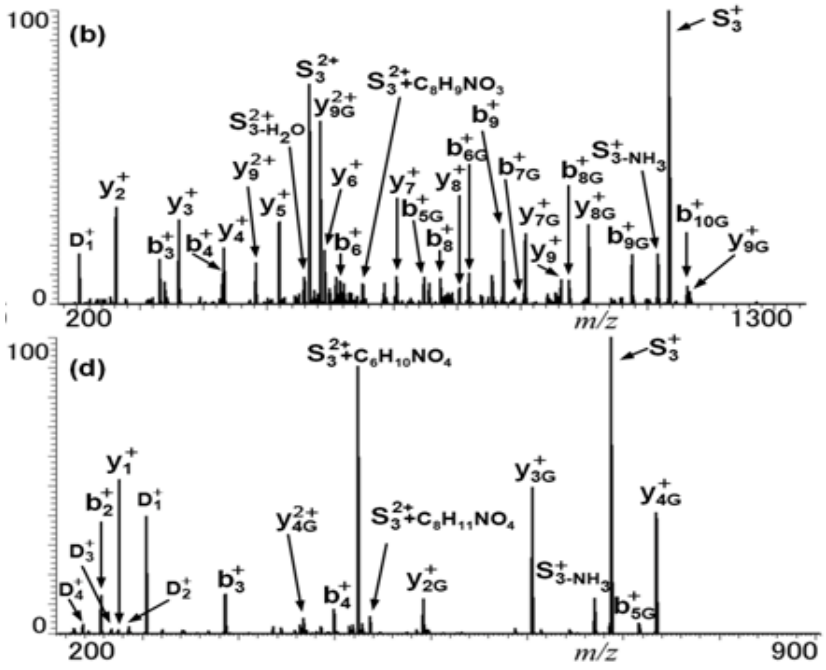
A：提高鉴定准确度和检索速度。

B：可用于统计分析，发现图谱特点。

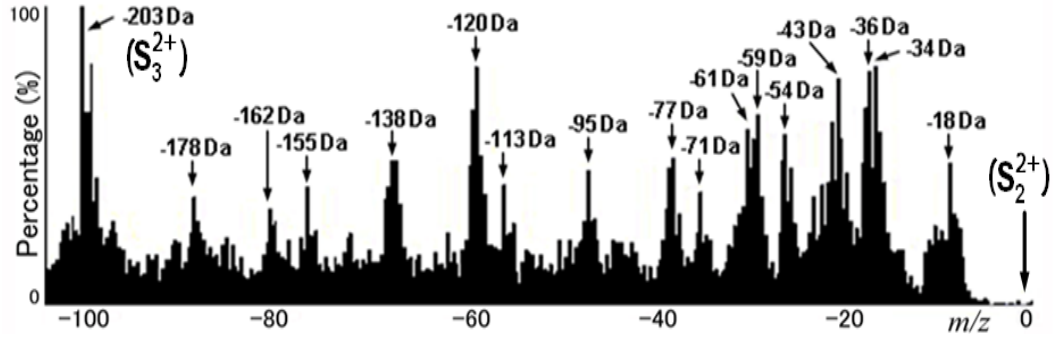
C：对数据库中不包含序列的发现。

图谱优化

糖肽的串联质谱行为与普通的肽段明显不同，正确识别其特殊的碎片离子，是其大规模精确鉴定必要基础。



S_2 Loss	203	178	162	155	138	120	113	95	77
	$C_8H_{13}NO_5$	$C_6H_{12}NO_5$	$C_6H_{10}O_5$	138+NH_3	120+H_2O	$C_4H_8O_4$	59+3H_2O	59+2H_2O	59+H_2O
S_3 Add	-	C_2H	C_2H_3N			C_4H_5NO			$C_6H_6O_3$
S_2 Loss	71	61	59	54	43	36	34	18	0
	54+NH3	43+H_2O	C_2H_5NO	3H_2O	C_2H_3O	2H_2O	2NH_3	H_2O	-
S_3 Add		$C_6H_8NO_3$	$C_6H_8O_4$		$C_6H_{10}NO_4$	$C_8H_9NO_3$		$C_8H_{11}NO_4$	$C_8H_{13}NO_5$



在MS³图谱中，去除GlcNAc碎裂峰。

对大量候选图谱中的糖残基碎片信号区域进行统计分析，以发现非肽信号来源。

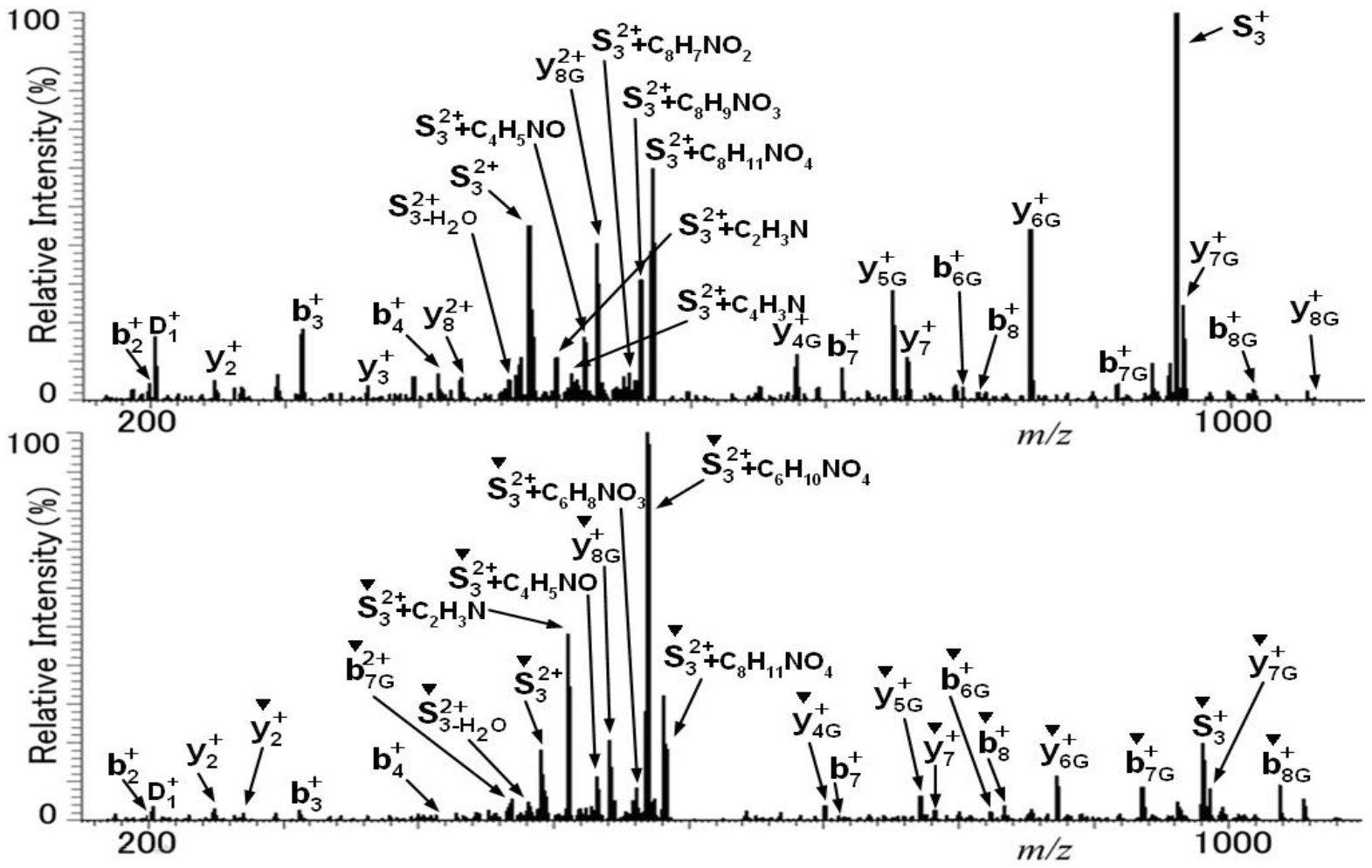
图谱优化方法：将搜索引擎无法考虑的糖残基碎片信号从三级图谱中去除，之后再进行数据库检索（提高10%鉴定）。

	鉴定肽段	正确肽段	鉴定谱图
优化后	112	108	2258
未处理	104	98	2163

(正反库， 1%假阳性率， 肝癌血浆样本）

经过图谱优化，不仅提高了图谱鉴定成功率，也增加了搜索结果的可靠性。

pFind发现Na^+离子峰



上方的图谱$[M+2H]^{2+}$，下方图谱$[M+H+Na]^{2+}$。

下面图谱中大部分碎片离子峰的相对强度较低，此外**Na离子**倾向于结合在含有**N乙酰氨基葡萄糖**的碎片上。图中三角形表示碎片含有Na离子。

核心岩藻糖化蛋白质识别结果

正常人血浆：**105**个位点，**73**个蛋白。

肝癌病人血浆：**106**个位点，**79**个蛋白。

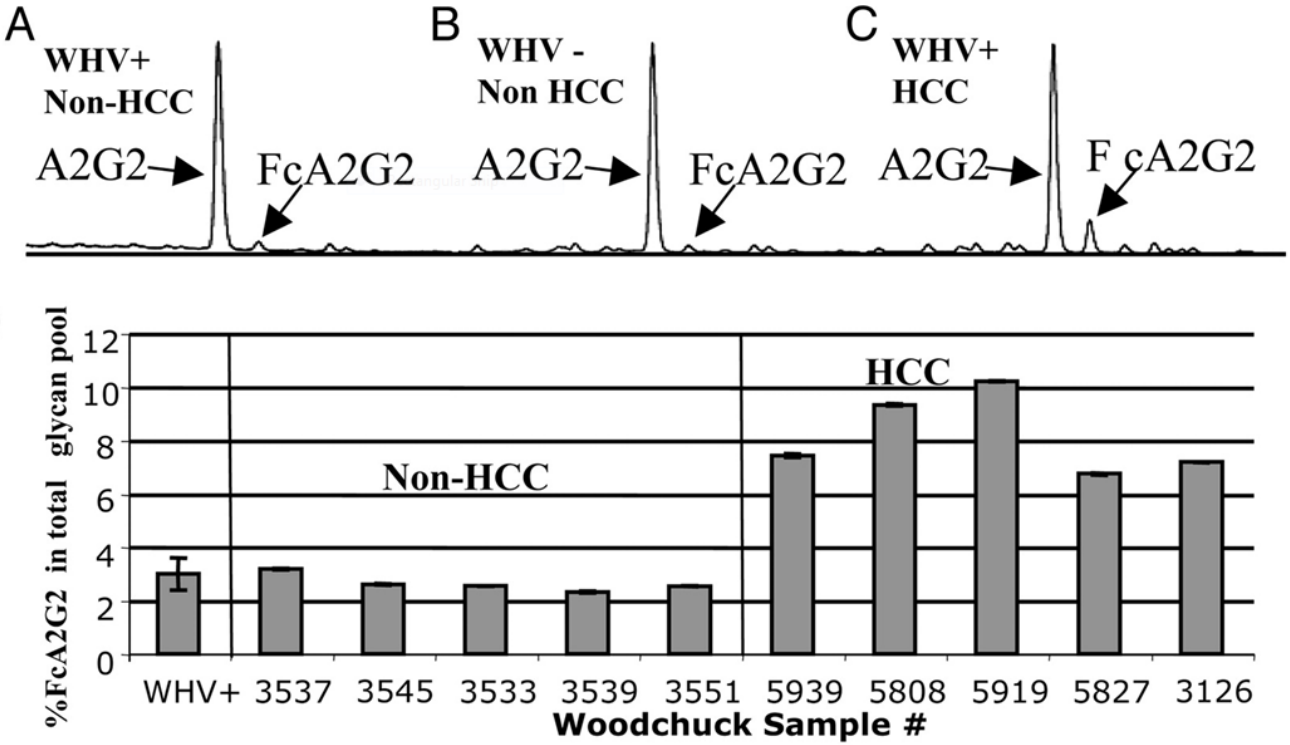
肝癌/胰腺癌/正常人血浆中鉴定核心岩藻糖化蛋白**100+**，位点超过**200+**。

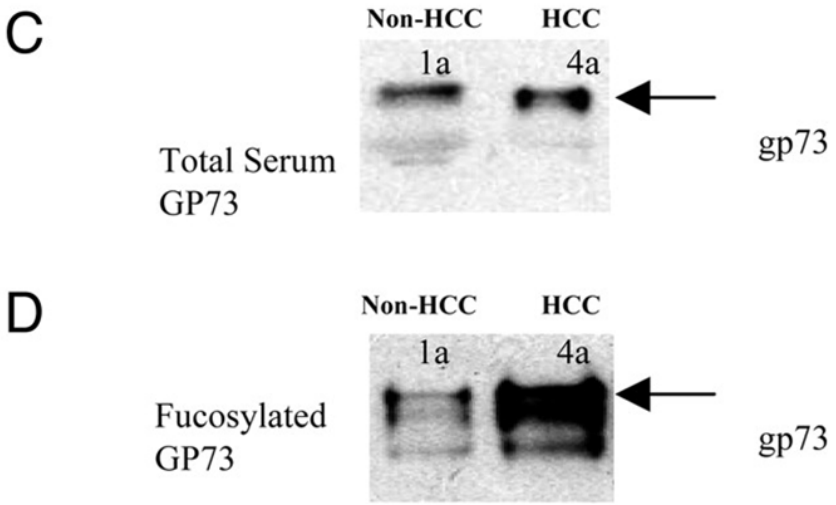
同期国际进展

- **5 CF蛋白质**
  - ❖凝集素提取，双向电泳与免疫印迹，MALDI-TOF/MS．PNAS，2005
- **19 CF蛋白质**
  - ❖凝集素提取，双向电泳与免疫印迹，LC-MS/MS．JPR，2006
- **56 CF蛋白质**
  - ❖高丰度去除，凝集素提取，LC-MS/MS．JPR，2009

靶向CF分析发现：

Gp73-潜在肝癌诊断标志物





Block T M et al. PNAS 2005;102:779-784

论文发表于

Molecular & Cellular Proteomics, 2009, 8 (5), 913-923.

Research

A Strategy for Precise and Large Scale Identification of Core Fucosylated Glycoproteins*

Wei Jia‡§¶, Zhuang Lu‡¶‖, Yan Fu¶**, Hai-Peng Wang**, Le-Heng Wang**, Hao Chi**, Zuo-Fei Yuan**, Zhao-Bin Zheng‡, Li-Na Song‡, Huan-Huan Han‡, Yi-Min Liang‡, Jing-Lan Wang‡, Yun Cai‡, Yu-Kui Zhang‖, Yu-Lin Deng‖, Wan-Tao Ying‡ ‡‡, Si-Min He**§§, and Xiao-Hong Qian‡¶¶

Core fucosylation (CF) patterns of some glycoproteins are more sensitive and specific than evaluation of their total respective protein levels for diagnosis of many diseases, such as cancers. Global profiling and quantitative characterization of CF glycoproteins may reveal potent biomarkers for clinical applications. However, current techniques are unable to reveal CF glycoproteins precisely on a large scale. Here we developed a robust strategy that integrates molecular weight cutoff, neutral loss-dependent MS^3, database-independent candidate spectrum filtering, and optimization to effectively identify CF glycoproteins. The rationale for spectrum treatment was innovatively based on computation of the mass distribution in spectra of CF glycopeptides. The efficacy of this strategy was demonstrated by implementation for plasma from healthy subjects and subjects with hepatocellular carcinoma. Over 100 CF glycoproteins and CF sites were identified, and over 10,000 mass spectra of CF glycopeptide were found. The scale of identification results indicates great progress for finding biomarkers with a particular and attractive prospect, and the candidate spectra will be a useful resource for the improvement of database searching methods for glycopeptides. *Molecular & Cellular Proteomics* 8:913-923, 2009.

Glycoproteins are implicated in a wide range of biological processes such as fertilization, development, the immune response, cell signaling, and apoptosis. Altered glycosylation patterns can affect the conformations of glycoproteins and their functions and interactions with other molecules (1, 2).

From the ‡State Key Laboratory of Proteomics-Beijing Proteome Research Center-Beijing Institute of Radiation Medicine, No. 33 Life Science Park Road, Changping District, Beijing 102206, China, §Institute of Biophysics, Chinese Academy of Sciences, No. 15 Datun Road, Chaoyang District, Beijing 100101, China, ‖Beijing Institute of Technology, No. 5 South Zhongguancun Street, Haidian District, Beijing 100081, China, and **Institute of Computing Technology, Chinese Academy of Sciences, No. 6 Kexueyuan South Road, Beijing 100190, China

Received, November 5, 2008, and in revised form, January 7, 2009
Published, MCP Papers in Press, January 12, 2009, DOI 10.1074/mcp.M800504-MCP200

Abnormal glycosylation has been demonstrated in many pathological processes. Targeted glycosylation research is considered increasingly important as a way to find novel therapeutic approaches (2, 3), and core fucosylation (CF)[1] glycoproteomics has attracted particularly great attention (4, 5). Previous reports show that CF glycoproteins are involved in many important physiological processes, such as transforming growth factor-β1 (6) and epidermal growth factor signaling pathways (7). They also play key roles in many pathological processes, such as hepatocellular carcinoma (HCC) (8, 9), pancreatic cancer (10, 11), lung cancer (6, 12), ovarian cancer (13), and prostate cancer (14). Moreover the CF patterns of several glycoproteins have been reported to serve as more sensitive and specific biomarkers than their total respective protein levels (6, 9, 15, 16). The combination of a biomarker panel of CF glycoproteins is expected to serve as a more reliable diagnostic standard (13).

Glycoproteomics research has been conducted for several years and has led to the generation of many effective evaluation methods. Most of these methods use lectin or the chemical reagent hydrazide to enrich glycopeptides. The oligosaccharide chains are then completely released by treatment of the glycopeptides with peptide-*N*-glycosidase F. Finally the deglycosylated peptides and the deglycosylation sites are identified by tandem mass spectrometric analysis (17, 18). Although impressive results have been attained, this commonly used strategy is not an ideal choice for CF glycoproteins research. First, the enrichment specificity of lectin is not satisfactory (19) as hydrazide chemical reactions irreversibly destroy glycan structures, particularly fucose tags. Second, the deglycosylation site is determined by the 0.9840-Da mass shift caused by the asparagine to aspartic acid transfer; its confidence can be compromised by deamination of the Asn. Besides that, the CF site can no longer be distinguished from other glycosylation sites in the same glycoprotein. Thus, the ideal way to precisely identify CF glycoproteins on a large

[1] The abbreviations used are: CF, core fucosylation; HCC, hepatocellular carcinoma; rhEPO, recombinant human erythropoietin; RP, reversed phase; S_n, symbol ion 2; S_n, symbol ion 3; HS, Hersman-Schmid; SCX, strong cation exchange; LTQ, linear trap quadrupole.

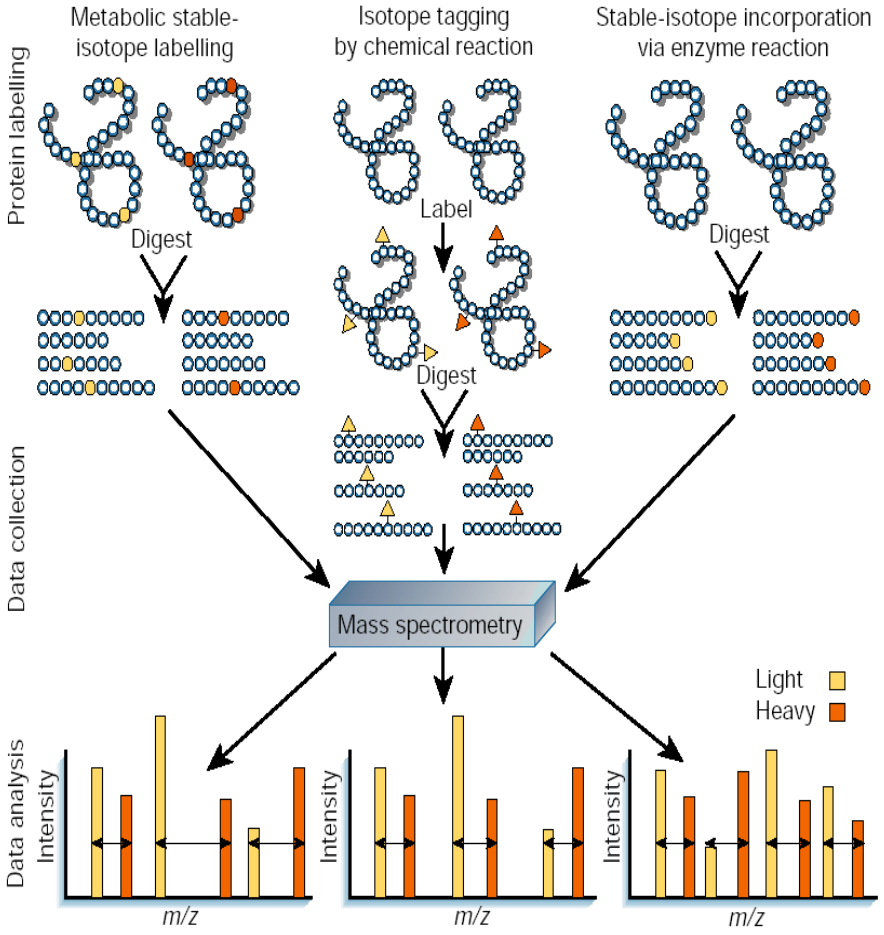
审稿人评价

This paper proposes a strategy for identifying core fucosylated glycoproteins which involves an imaginative integration of methodologies …… Indeed, their identification of over a hundred core fucosylated glycoproteins in plasma is very impressive and this merits publication.

三、基于18O稳定同位素标记的相对与绝对定量蛋白质组研究技术

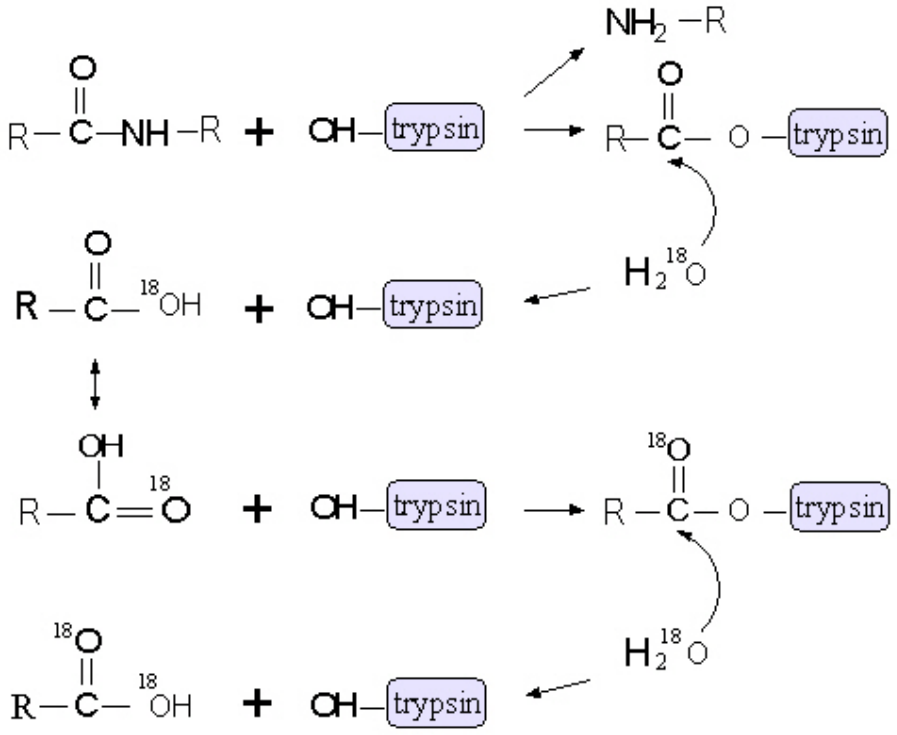
稳定同位素标记辅助的定量蛋白质组研究技术

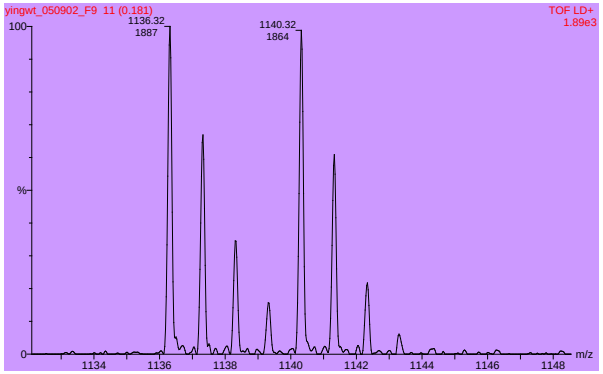
1. 化学同位素标记
   - 同位素编码亲和标签
   - 化学同位素标记
2. 酶催化同位素标记
3. 代谢同位素标记
   - Stable isotope labeling by amino acids in cell culture
   - Elemental labeling (N15)

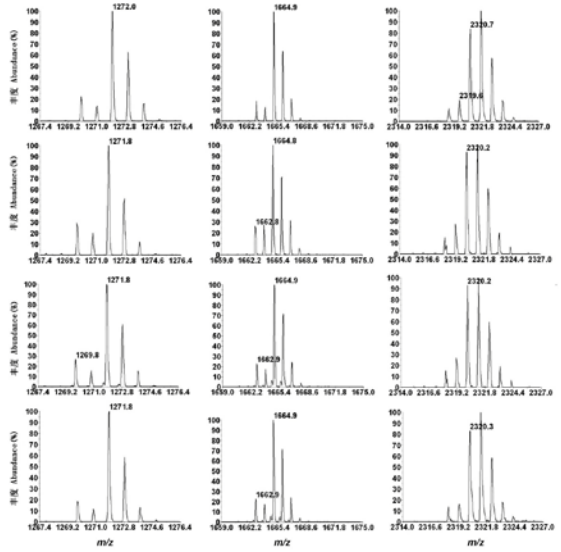


Aebersold, R.; Mann, M. Nature 2003, 422, 198-207.

^{18}O标记引进稳定同位素







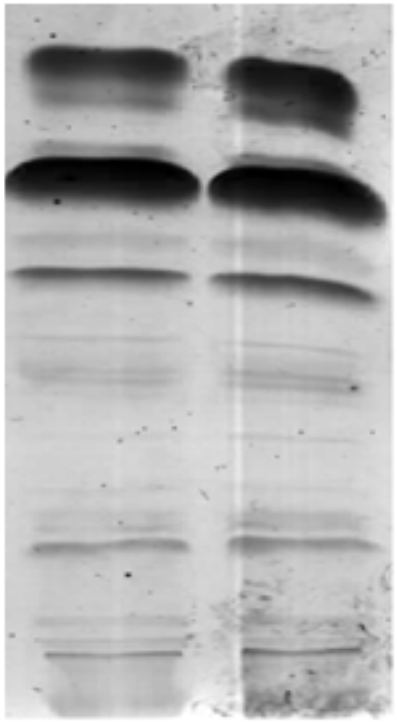
减少标记时间的同时提高了标记效率

有效地抑制了回交反应，保证了标记产物的稳定性

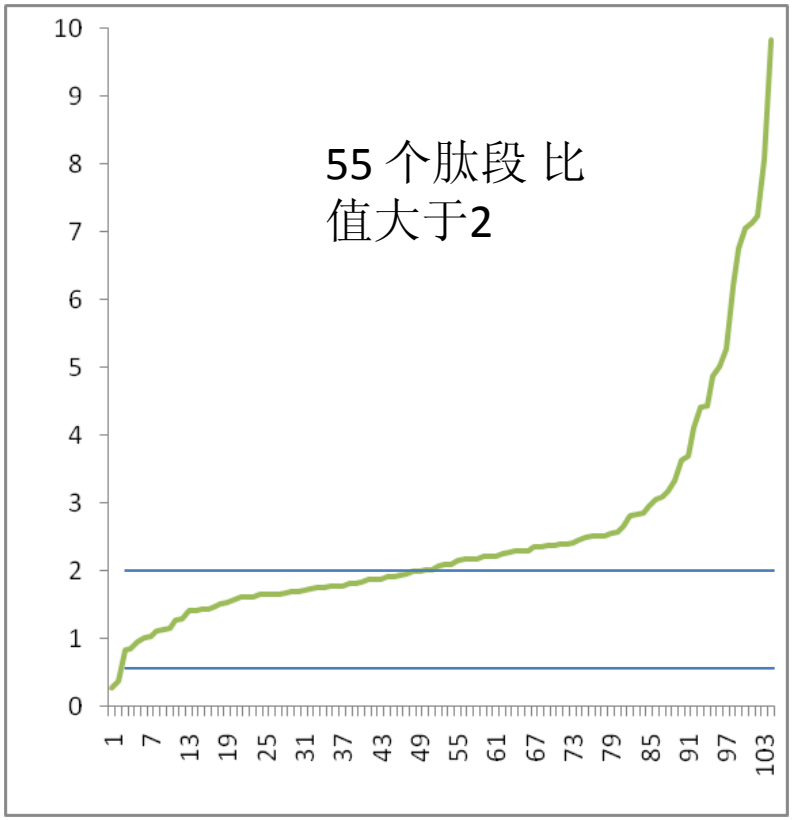
赵焱，钱小红等. 分析化学，2010，38（1）：91-4

pFindQ 软件包

胰腺癌核心岩藻化糖蛋白定量比较



胰腺癌/ 正常 混合血清SDS-PAGE



胰腺癌/ 正常 混合血清核心岩藻糖化肽段差异表达

待发表数据

定量结果验证（clusterin）

- Clusterin，参与介导细胞聚集，稳定细胞膜等，并在细胞凋亡、细胞周期调节、DNA损伤、细胞黏附中起着重要的作用。其改变发生于多种肿瘤发生过程中

蛋白号	蛋白分子量	蛋白等电点				蛋白名		
肽段No.	肽段电荷	肽段序列	肽段打分	肽段理论值	误差	肽段修饰	比例（胰腺癌/正常）	肽段信噪比
P10909|CLUS_HUMAN	49500.43	5.98				Clusterin precursor - Homo sapiens (Human)		
1	2	R.LANLTQGEDQYYLR.V	74.81	1885.9057	0.0106	3 N (HexNAc_N) ;	1.6579	177.3822
2	2	K.EDALNETR.E	49.13	1149.5149	0.0022	5 N (HexNAc_N) ;	1.7197	259.1406
3	2	K.KKEDALNETR.E	60.72	1405.7048	0.0039	7 N (HexNAc_N) ;	2.2517	182.0917
4	3	K.MLNTSSLLEQLNEQFNWVSR.L	29.12	2611.2588	0.0119	3 N (HexNAc_N) ;	2.8228	7.0407

定量结果验证

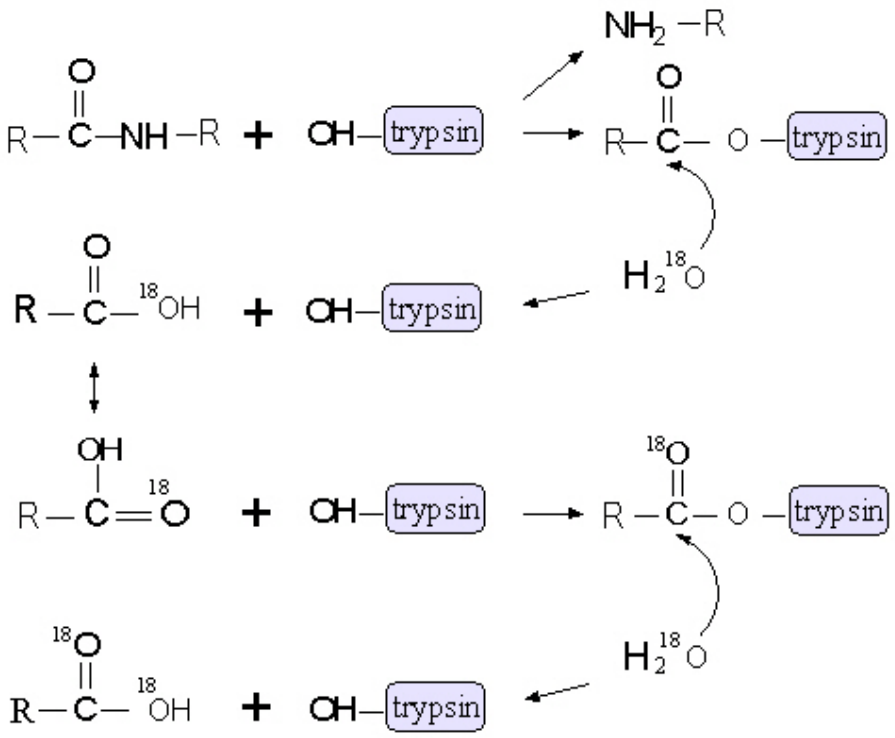
- 同时在蛋白水平（蛋白抗体western blotting）和核心岩藻糖基化水平 （LCH western blotting）量的改变进行验证

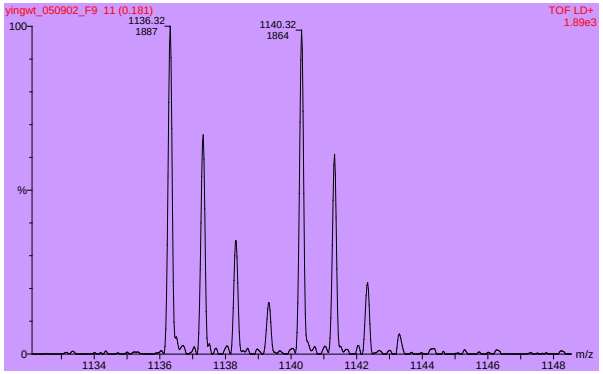
Control PC　　Control PC

clusterin蛋白抗体 wb　　LCH wb

蛋白抗体wb结果显示，正常人和胰腺癌病人血清中clusterin蛋白表达量没有显著差异，而LCH wb结果表明，该蛋白岩藻糖基化水平显著上调，与之前相对定量结果一致。

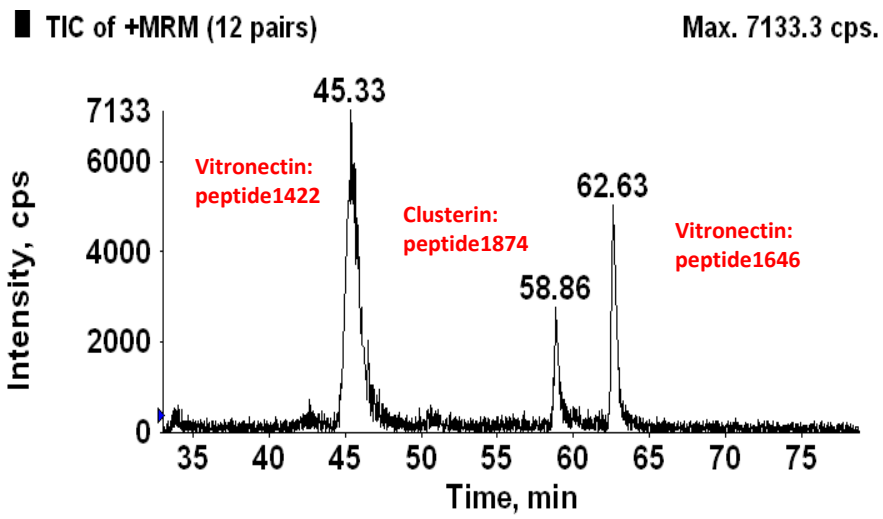
四、基于^{18}O稳定同位素标记的绝对定量蛋白质组研究技术



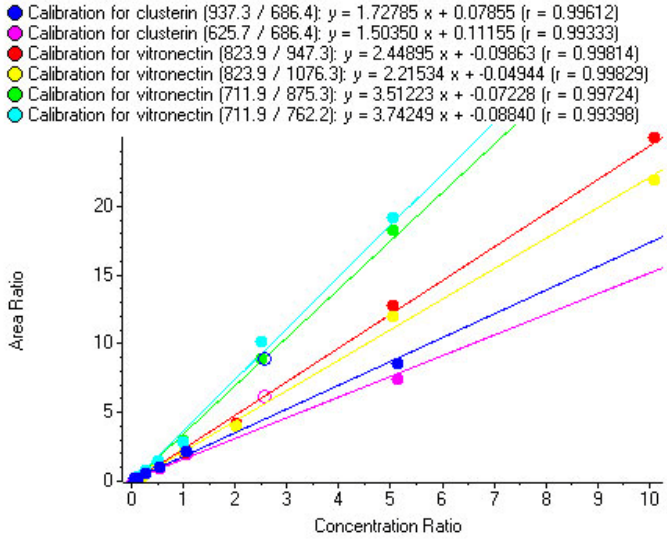


MRM定量指标评价

特异性



标准曲线和线性



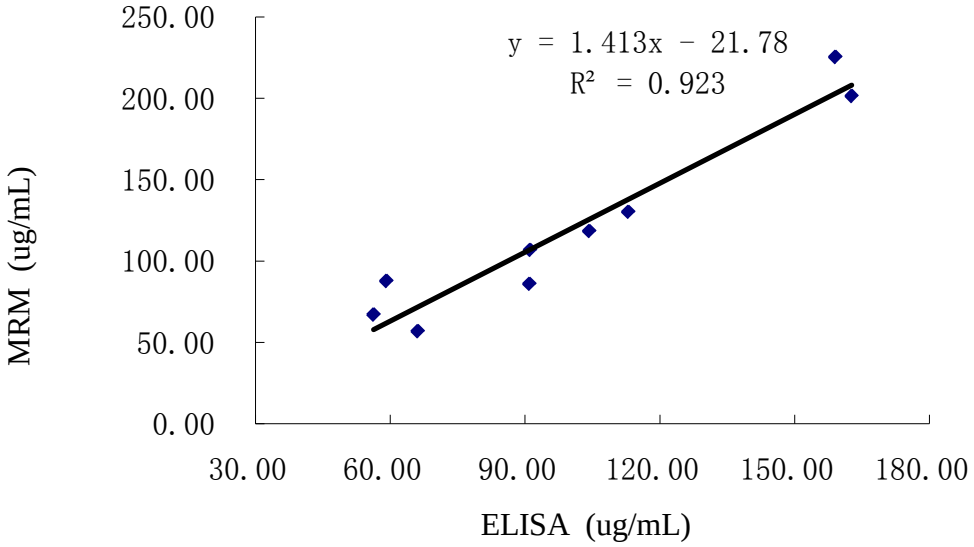
精密度和准确度

		VIT-1		VIT-2		CLU-1	
	Tansition	711.9/875.3	711.9/762.2	823.9/947.3	823.9/1076.3	937.3/686.4	625.7/686.4
LQC(0.8fmol/ul)	%RSD	1.99	5.14	6.65	11.96	/	/
	%RE	-0.17	4.85	0.22	5.06	-8.15	14.65
MQC(4fmol/ul)	%RSD	3.70	2.47	2.60	4.93	/	/
	%RE	-13.30	-20.01	-10.55	-6.91	1.94	4.15
HQC(20fmol/ul)	%RSD	3.92	8.69	2.47	2.97	/	/
	%RE	1.31	2.53	5.52	8.41	-3.84	-4.90

检测重现性和预处理回收率

	VIT-1		VIT-2		CLU-1	
Sample NO.	711.9/875.3	711.9/762.2	823.9/947.3	823.9/1076.3	937.3/686.4	625.7/686.4
Preparation recovey (Average%)	92.50%	93.88%	88.91%	91.93%	87.49%	82.22%
Detection repeatability RSD(%)	6.87	28.87	4.27	5.19	6.72	3.32

MRM分析结果与ELISA定量结果比较



technical notes Journal of proteome research

Combination of Improved ^{18}O Incorporation and Multiple Reaction Monitoring: A Universal Strategy for Absolute Quantitative Verification of Serum Candidate Biomarkers of Liver Cancer

Yan Zhao,[§] Wei Jia,[§] Wei Sun,[§] Wenhai Jin,[†] Lihai Guo,[†] Junying Wei,[§] Wantao Ying,[§] Yangjun Zhang,[§] Yongming Xie,[†] Ying Jiang,[§] Fuchu He,[§] and Xiaohong Qian*[,§]

State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Radiation Medicine, No.33 Life Science Park Road, Changping District, Beijing 102206, P. R. China, and AB SCIEX Asia Pacific Application Support Center, Shanghai, 200233, P. R. China

Received December 24, 2009

Abstract: Stable isotope dilution-multiple reaction monitoring-mass spectrometry (SID-MRM-MS), which is an alternative to immunoassay methods such as ELISA and Western blotting, has been used to alleviate the bottlenecks of high-throughput verification of biomarker candidates recently. However, the inconvenience and high isotope consumption required to obtain stably labeled peptide impedes the broad application of this method. In our study, the ^{18}O-labeling method was introduced to generate stable isotope-labeled peptides instead of the Fmoc chemical synthesis and Qconcat recombinant protein synthesis methods. To make ^{18}O-labeling suitable for absolute quantification, we have added the following procedures: (1) RapiGest SF and microwave heating were added to increase the labeling efficiency; (2) trypsin was deactivated completely by chemical modification using tris(2-carboxyethyl)phosphine (TCEP) and iodoacetamide (IAA) to prevent back-exchange of ^{18}O to ^{16}O, and (3) MRM parameters were optimized to maximize specificity and better distinguish between ^{18}O-labeled and unlabeled peptides. As a result, the ^{18}O-labeled peptides can be prepared in less than 1 h with satisfactory efficiency (>97%) and remained stable for 1 week, compared to traditional protocols that require 5 h for labeling with poor stability. Excellent separation of ^{18}O-labeled and unlabeled peptides was achieved by the MRM-MS spectrum. Finally, through the combined improvement in ^{18}O-labeling with multiple reaction monitoring, an absolute quantification strategy was developed to quantitatively verify hepatocellular carcinoma-related biomarker candidates, namely, vitronectin and clusterin, in undepleted serum samples. Sample preparation and capillary-HPLC analysis were optimized for high-throughput applications. The reliability of this strategy was further evaluated by method validation, with accuracy (%RE) and precision (%RSD) of less than 20% and good linearity ($r^2 > 0.99$), and clinical validation, which were consistent with previously reported results. In summary, our strategy can promote broader application of SID-MRM-MS for biomarkers from discovery to verification regarding the significant advantages of the convenient and flexible generation of internal standards, the reduction in the sample labeling steps, and the simple transition.

Keywords: MRM • ^{18}O-Labeling • Absolute quantification • Biomarker verification • Serum • Liver cancer • Vitronectin • Clusterin

Introduction

Biomarkers play important roles in diagnosis, treatment, progression, and prognosis of cancer,[1] the discovery of which have attracted extensive attentions in the medical field.[2-4] The advent of proteomics technology including two-dimensional polyacrylamide gel electrophoresis, mass spectrometry, and protein microarray have led to a proliferation of numerous biomarker candidates by global detection and quantitation of proteins, which greatly accelerate the discovery process.[5] However, before the results can be applied to clinical management, these biomarker candidates need further validation to address their sensitivity, specificity, reproducibility, and accuracy.[6] A general pipeline to obtain ultimate biomarkers for clinical use includes four phases: discovery, qualification, verification, and clinical validation.[7,8] Although hundreds of candidates can be discovered, further verification of these candidates in more cases of samples has become the rate-limiting step in a biomarker pipeline,[9] because the conventional verification methods, such as Western blotting and enzyme-linked immunosorbent assay (ELISA), are not suitable for large-scale analysis due to poor throughput and problems associated with obtaining monoclonal antibodies.[10] Take the study of Pietsma et al.[11] as an example, among the 136 secretome candidate biomarkers yielded by comparative analysis of MEF/Toff cells expressing IGF1R and stimulated with IGF1, only 2 of these were able to be further verified by ELISA assays.

In recent years, stable isotope dilution-multiple reaction monitoring-mass spectrometry (SID-MRM-MS) was introduced to alleviate the bottleneck in biomarker development.[12-16] In

* To whom correspondence may be addressed: Xiaohong Qian, e-mail: qianxh1@yahoo.com.cn; phone: 8610-80705055; fax: 8610-80705155.
§ State Key Laboratory of Proteomics, Beijing Proteome Research Center.
† AB SCIEX Asia Pacific Application Support Center.

10.1021/pr901068

Published on Web 04/21/2010

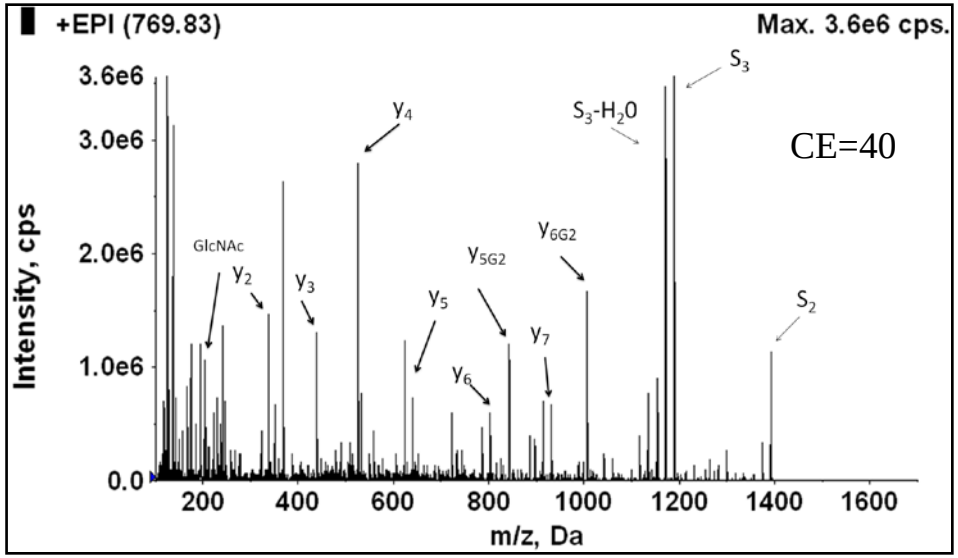
血清样品Vitronecin浓度测定值比较（ug/ml）

检测方法	N400	N430	N407	L21	L15	L13	Y64	Y41	Y72
ELISA	113.00	159.08	162.69	66.12	56.30	59.17	104.30	90.98	91.11
MRM	130.23	225.40	201.40	56.80	67.00	87.72	118.44	85.98	106.82

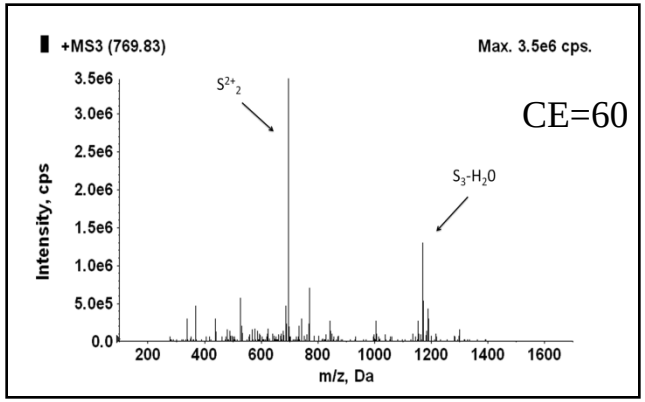
Vitronectin为例，随机选择了9例血清样品。分别用MRM的定量方法和ELISA定量方法进行检测，对测得的浓度进行相关性曲线分析。0.923的良好相关性验证了定量策略的可靠性。

核心岩藻糖肽质谱裂解条件优化

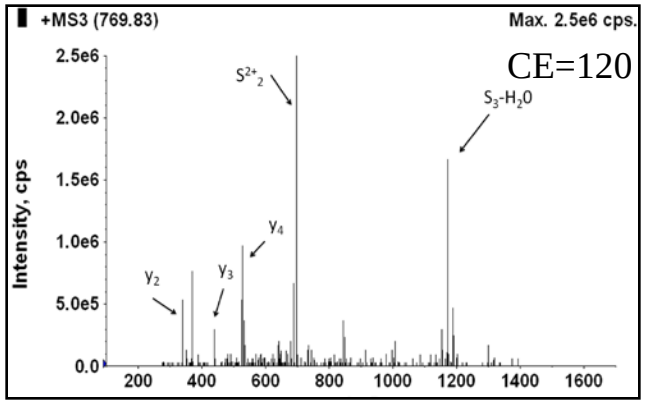
三级四极杆 VS 离子阱



三级四极杆的二级碎裂



离子阱的二级碎裂



5类离子主要为：

丢失一个糖基的母离子；两个糖基均丢失的母离子；带糖基的b或y离子；不带糖基的b或y离子； 糖基的碎片离子

如何选择灵敏度高，特异性强的子离子？
并且保证这些离子的稳定出现？

碎片离子峰强度 vs 碰撞能量

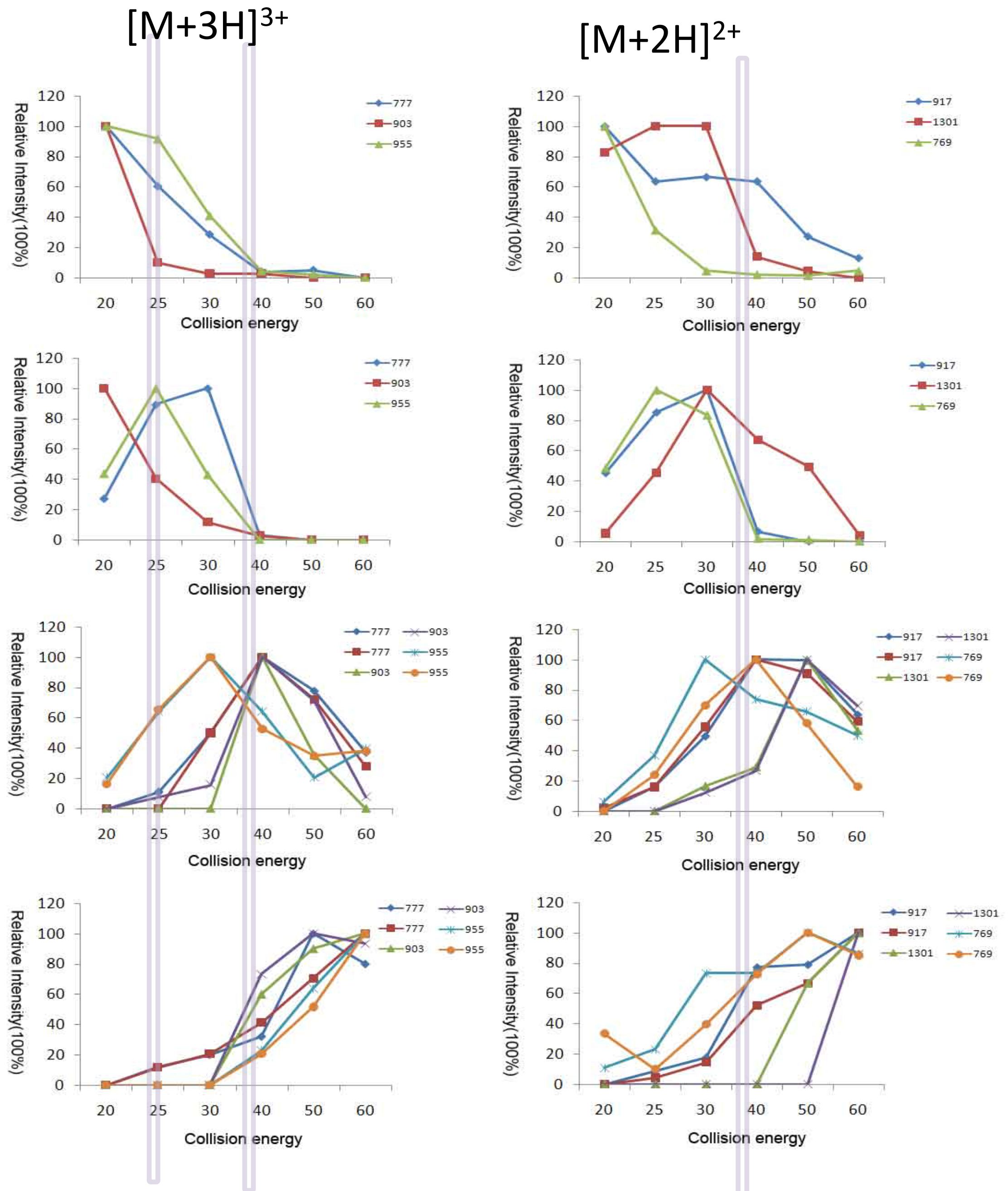
来源

母离子

中型丢失

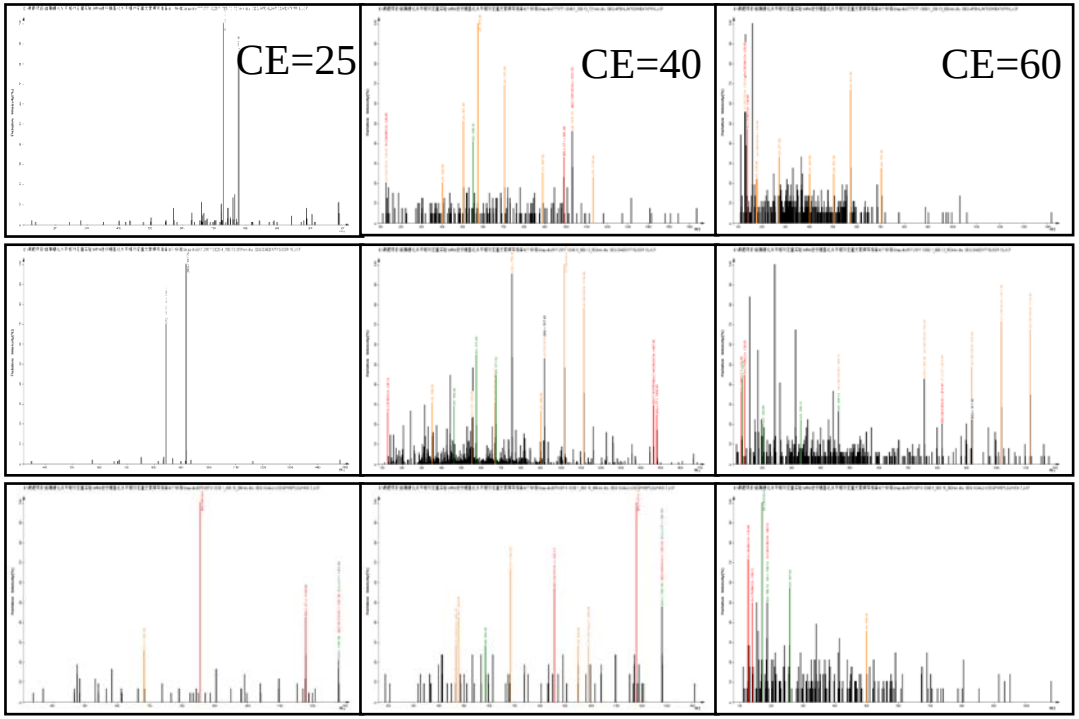
Y系离子

糖碎片

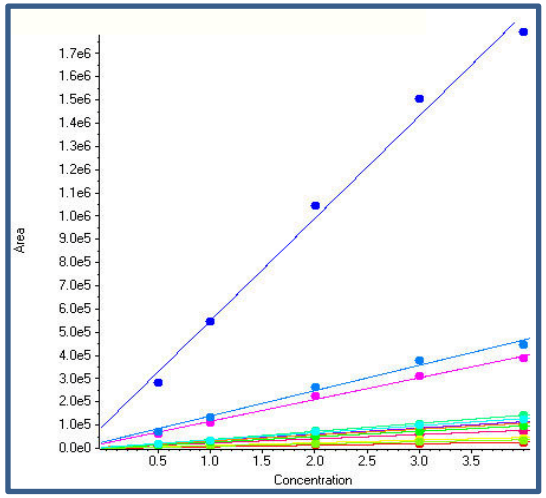


MRM灵敏度 vs 特异性

二级离子与碰撞能量



二级离子与定量



769.8/696.8	y=4.4e5x+1.1e5(r=0.9946)
769.8/1392.6	y=2.7e4x+5.5e3(r=0.9963)
769.8/1189.5	y=9.4e4x+2.3e4(r=0.9964)
769.8/843.4	y=1.8e4x+4.6e3(r=0.9935)
769.8/803.4	y=6.5e3x+1.5e2(r=0.9970)
769.8/1006.4	y=2.5e4x+6.8e3(r=0.9906)
769.8/338.2	y=1.1e4x+1.2e3(r=0.9973)
769.8/439.2	y=9.1e3x+7.9e2(r=0.9940)
769.8/526.3	y=2.3e4x+3.2e3(r=0.9968)
769.8/204.1	y=3.5e4x+3.2e3(r=0.9977)
769.8/138.1	y=3.1e4x+5.8e3(r=0.9971)
769.8/126.1	y=1.1e5x+3.1e4(r=0.9927)

发生中性丢失的母离子——CE=25——**增强检测灵敏度**

肽段碎裂产生的y离子——CE=40——**增强检测特异性**

核心岩藻糖化肽段MRM规模化相对定量确证分析

血清目的蛋白核心岩藻糖肽段、非糖基化肽段相对定量

糖基化水平相对定量：

血清 + 甲状腺球蛋白糖肽

↓

LCA凝集素富集

↓

还原、烷基化、酶切

↓

超滤糖基化肽段富集

↓

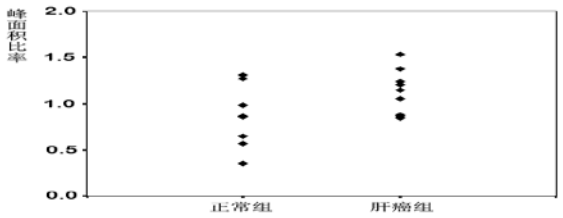
Endo F3切糖

↓

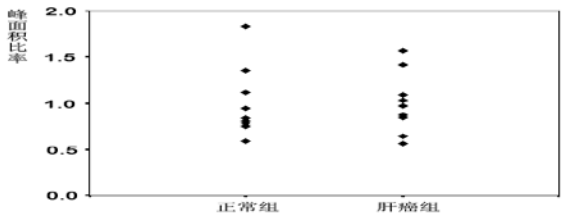
MRM

Protein Name	CF-peptide	Transition	CE
Ig gamma-1 chain C region	EEQYJSTYR	769.8 / 696.8	25
		769.8 / 1189.5	40
		769.8 / 1006.5	40
Hemopexin precursor	SWPAVGJCSSALR	877.4 / 804.4	25
		877.4 / 1067.5	40
		877.4 / 1334.6	40
Ig gamma-2 chain C region	EEQFJSTFR	753.8 / 1360.4	25
		753.8 / 1157.3	40
		753.8 / 974.4	40
Ig alpha-1 chain C region	TPLTAJITK	654.3 / 581.3	25
		654.3 / 647.4	40
		654.3 / 760.3	40
Alpha-2-macroglobulin precursor	GCVLLSYLJETVTVSASLESVR	916.1 / 867.4	25
		916.1 / 848.5	40
		916.1 / 1048.6	40
	VSJQTLSLFFTVLQDVPVR	1256.7 / 1183.6	25
		1256.7 / 826.5	40
		1256.7 / 1320.6	40
Ceruloplasmin precursor	EJLTAPGSDSAVFFEQGTTR	1238.6 / 1165.7	25
		1238.6 / 1598.5	40
		1238.6 / 985.2	40
Bovine thyroglobulin	APEHLJWTGSWEATKPR	777.4 / 728.6	25
		777.4 / 572.4	40
		777.4 / 1132.6	40
	VGQFJLSGALGTR	834.9 / 761.9	25
		834.9 / 661.6	40
		834.9 / 1319.6	40
	EAEEIVTYSJSSR	917.4 / 844.5	25
		917.4 / 1017.5	40
		917.4 / 1116.4	40
	FLVNVGQFJLSGALGTR	1072.1 / 998.6	25
		1072.1 / 661.4	40
		1072.1 / 1423.7	40

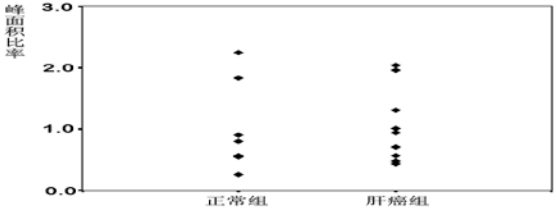
肝癌核心岩藻糖化肽段SRM定量验证



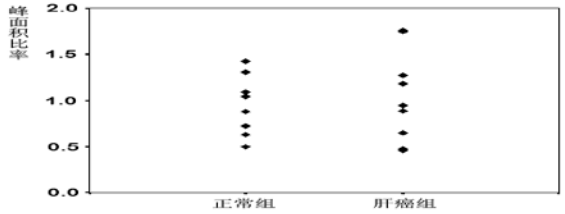
Alpha-2-macroglobulin 1256（p=0.122）



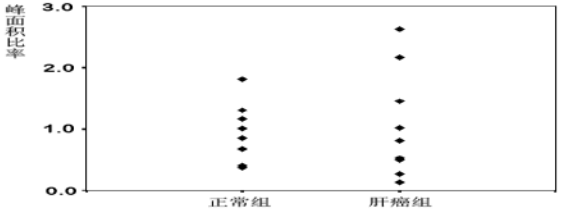
Alpha-2-macroglobulin 916（p=0.691）



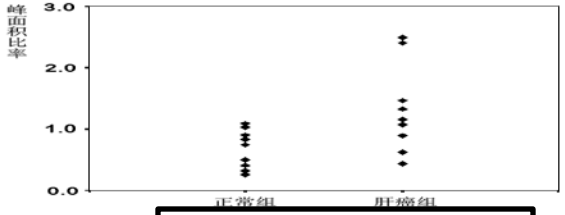
Hemopexin（p=0.566）



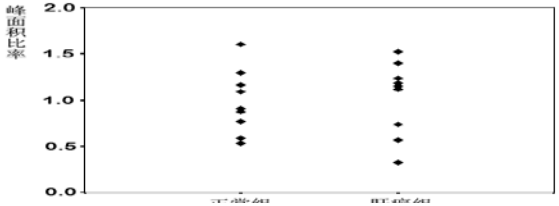
IgG-1（p=0.895）



IgA-2（p=0.895）



Ceruloplasmin（p=0.024）



IgG-2（p=0.691）

核心岩藻糖化肽段的相对定量结果及统计分析

七个糖基化肽段的相对定量结果显示：

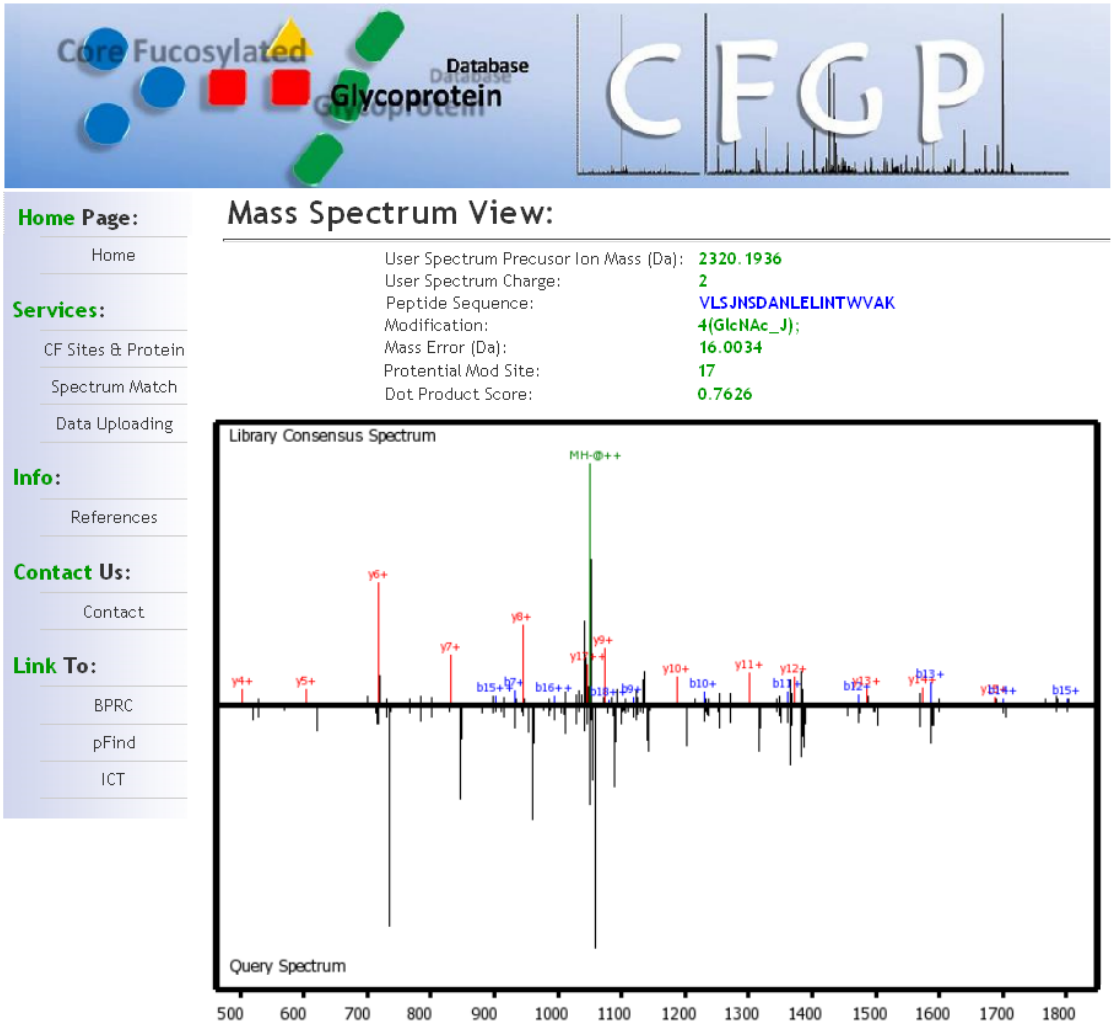
Ceruloplasmin蛋白397位点发生的核心岩藻糖化修饰在肝癌和正常对照组对比中具有显著差异，在肝癌组呈上调趋势。

待发表数据

五、核心岩藻糖化蛋白质数据库

1. 100+ 鉴定的核心岩藻糖蛋白质及谱图收集及搜库功能
2. 文献同行鉴定的核心岩藻糖蛋白质数据
3. 核心岩藻糖基化位点
4. 经典生化及分子生物学研究中核心岩藻糖化修饰蛋白质收录
5. 定量确证分析结果
6. 相关文献索引
7. 组织/ 体液、生理/ 病理相关核心岩藻糖数据集

CF蛋白质作用机制及重要功能蛋白质发掘



Wei Jia, Ding Ye, Yan Fu, Ji-FengWang, Wan-Tao Ying, Xiao-Hong Qian, Si-Min He In preparation

结论

- 对糖肽富集、质谱采集、图谱软件解析进行优化，实现对核心岩藻糖化蛋白的大规模精确鉴定，建立同期最大规模核心岩藻糖化蛋白及位点数据集。
- 通过统计分析，揭示了核心岩藻糖化糖肽在离子阱型质谱中的行为，为其碎裂规律研究提供基础数据集，并为在其它类型质谱中的研究提供了方法学及结果解析参照。
- 核心岩藻糖化蛋白质的规模化分析，归功于化学、生物学及信息学的有机整合。

致谢

中国科学院计算所
Yan Fu, Leheng Wang
Hao Chi, Simin He

北京蛋白质组研究中心
Wei Jia, Zhuang Lu,
Yan Zhao, Jiangyang Zhang, Xiaohong Qian

National Natural Science Foundation of China
National Key Program for Basic Research of China
National Hi-Tech Research and Development Program of China