

糖型特异性的糖蛋白质组学研究

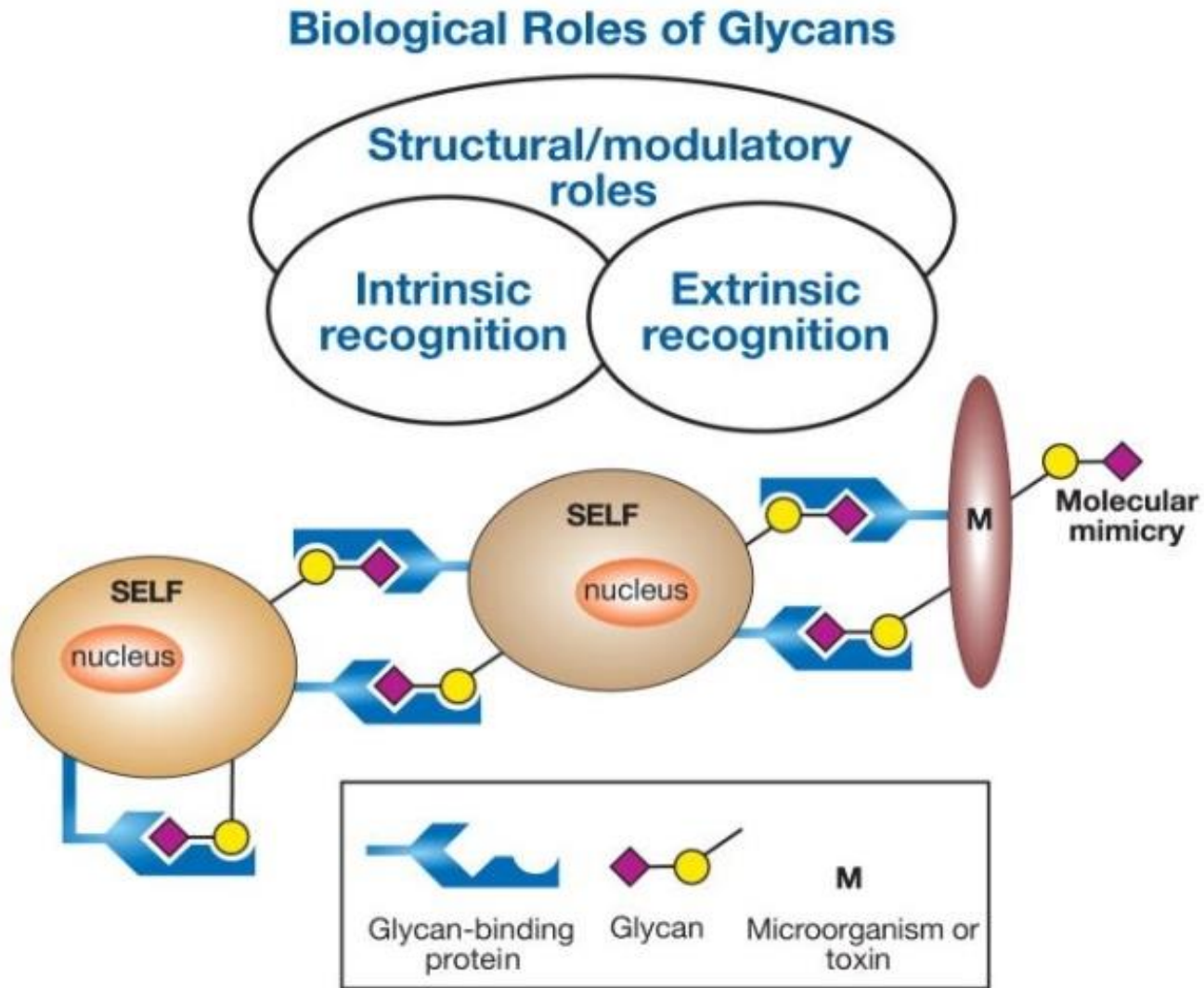
应万涛

蛋白质组学国家重点实验室

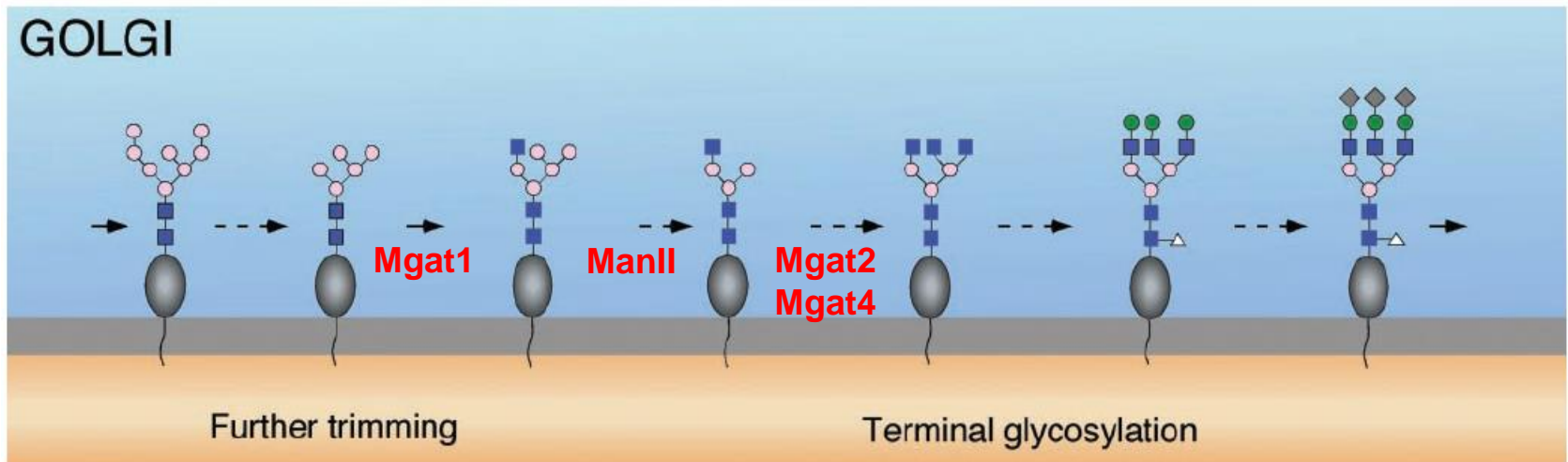
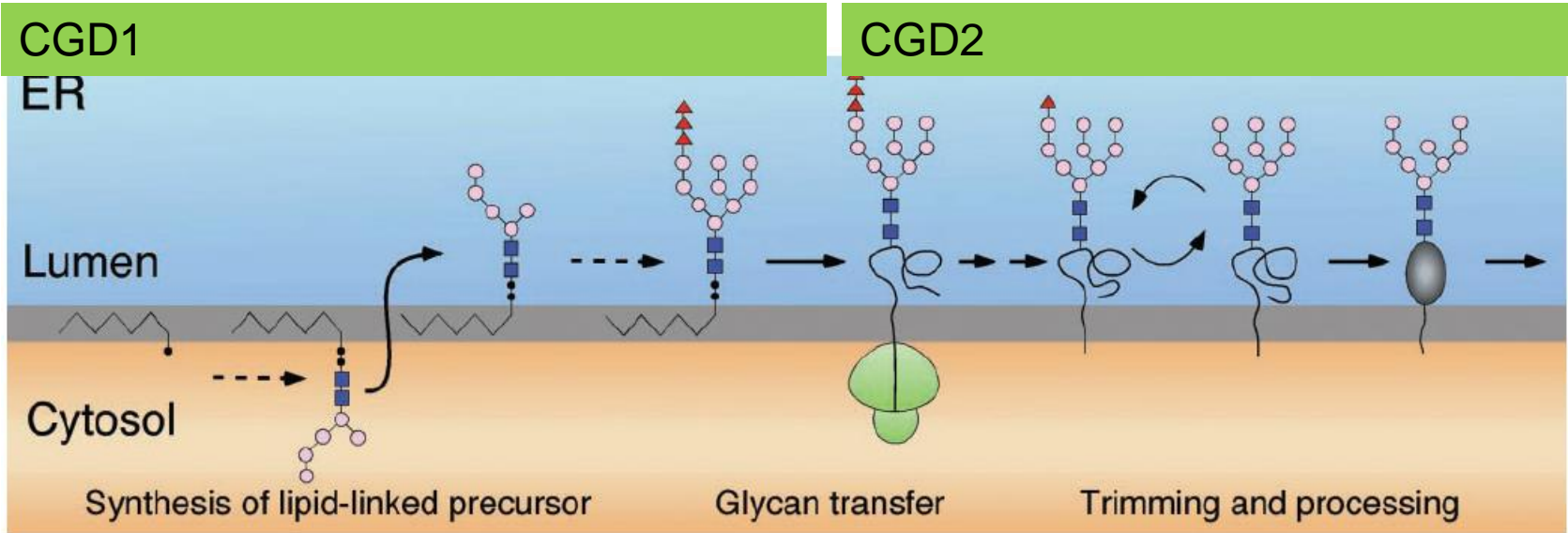
北京蛋白质组研究中心

军事医学科学院放射与辐射医学研究所

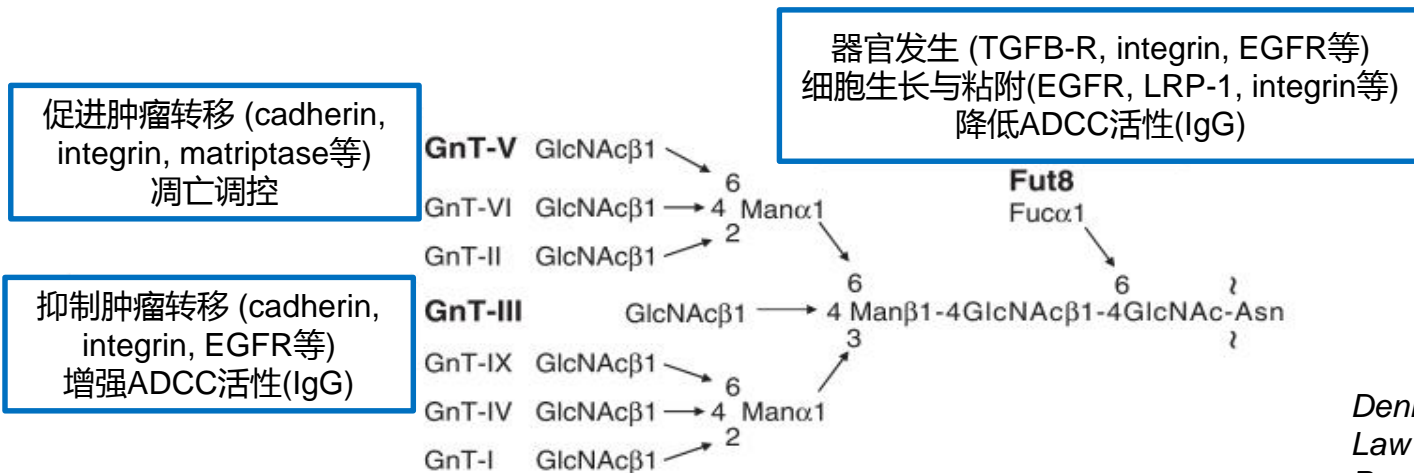
糖基化修饰执行重要生物学功能



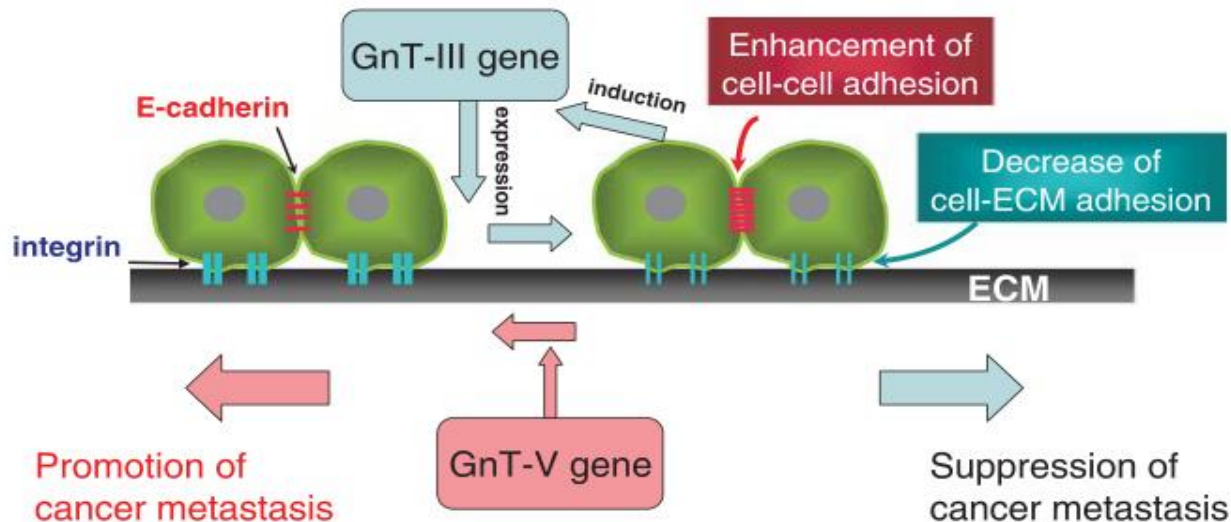
N-糖链加工过程



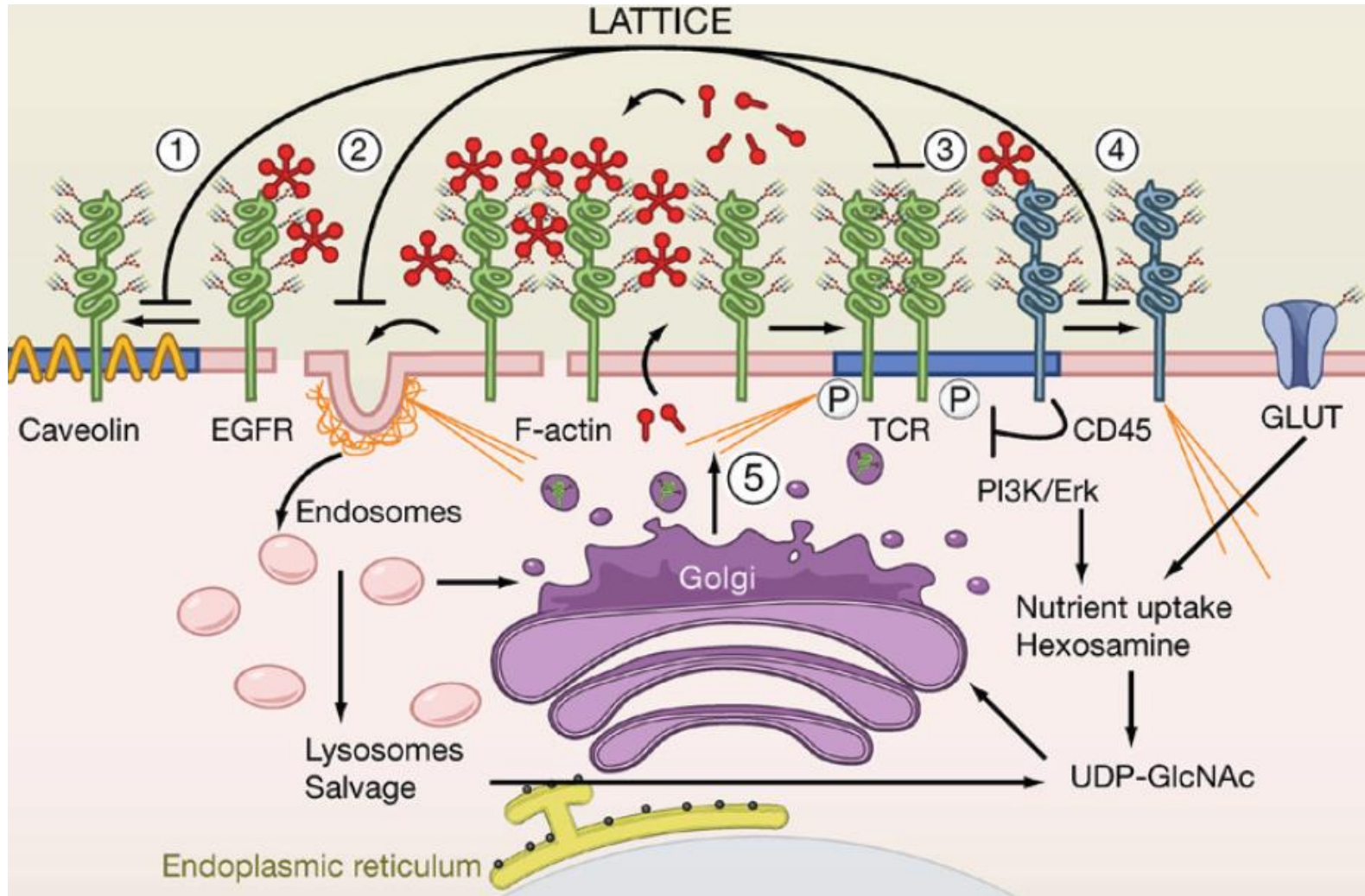
糖型修饰与重要生理与病理进程相关



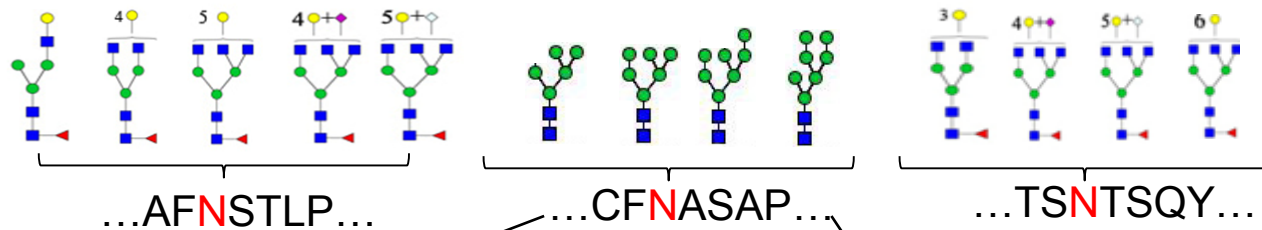
Dennis JW, Science, 1987
 Law KS, Cell, 2007
 Pang PC, Science, 2011
 Liu YC, PNAS, 2011



Galectins通过与复合糖作用调控信号传递



位点特异的糖基化修饰信息有限



收集文献中的信息

数据有限

近40年，117篇文献

474个位点，169个糖蛋白

来源不同

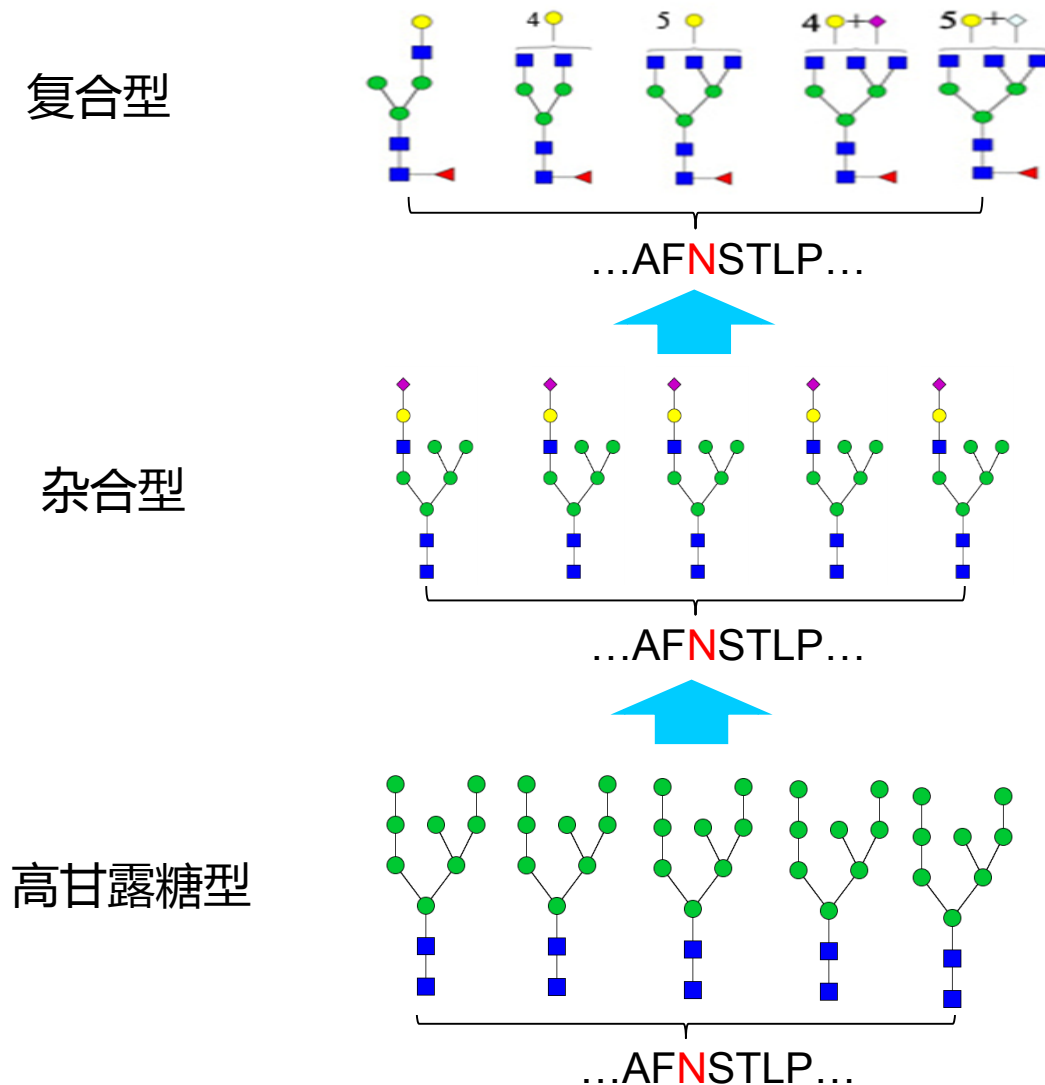
种属，组织，状态~~

缺乏实验数据

?

基于 α -甘露糖苷酶抑制剂的 糖蛋白糖型研究

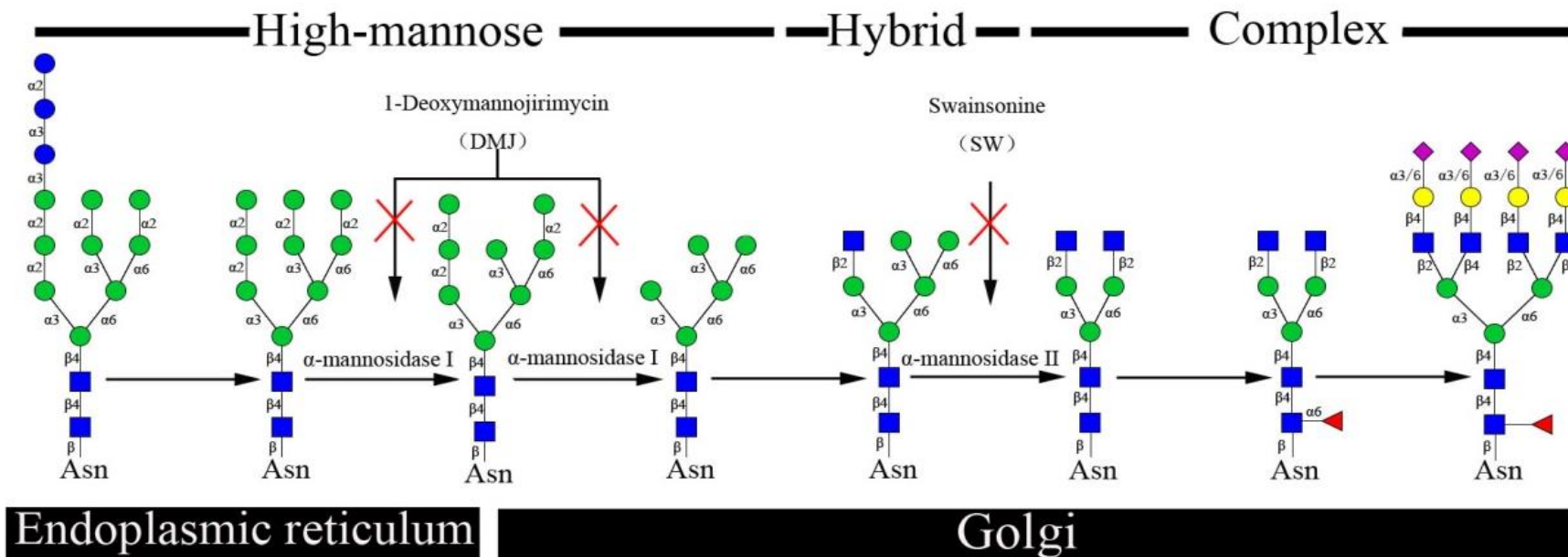
糖型简化与富集



糖链逆向
生物富集

- ◆ 简化糖型
- ◆ 提高单一糖型丰度
- ◆ Con A响应值提高
- ◆ 定量变化提示糖型信息

N-糖基化修饰通路干扰作用机制



● Mannose

● Galactose

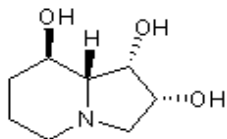
● Glucose

▲ Fucose

■ GlcNAc

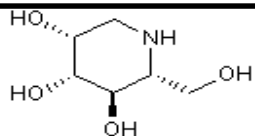
◆ Sialic acid

SW



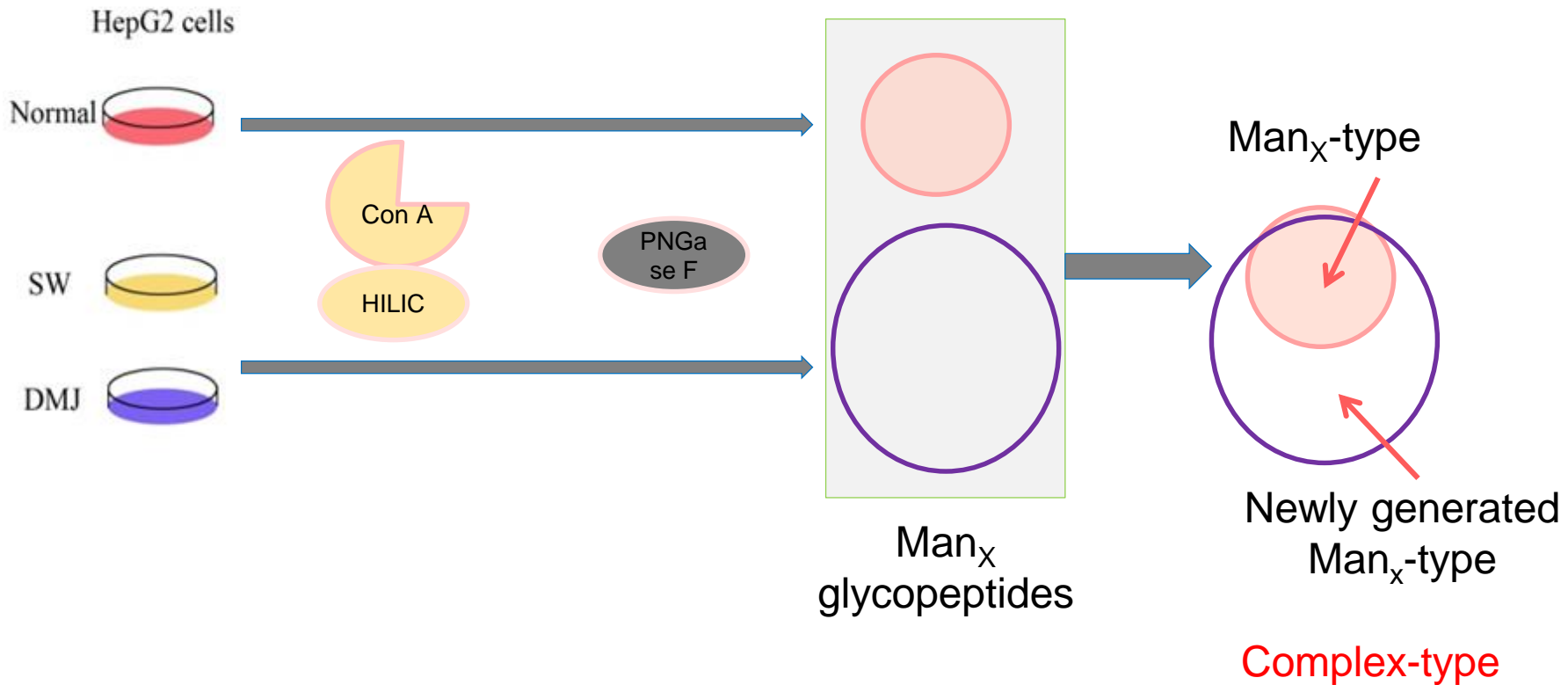
- α -甘露糖苷酶 II 的抑制剂。
- 抑制肺癌、胃癌、肝癌、黑色素瘤、淋巴瘤等多种肿瘤的生长。

DMJ



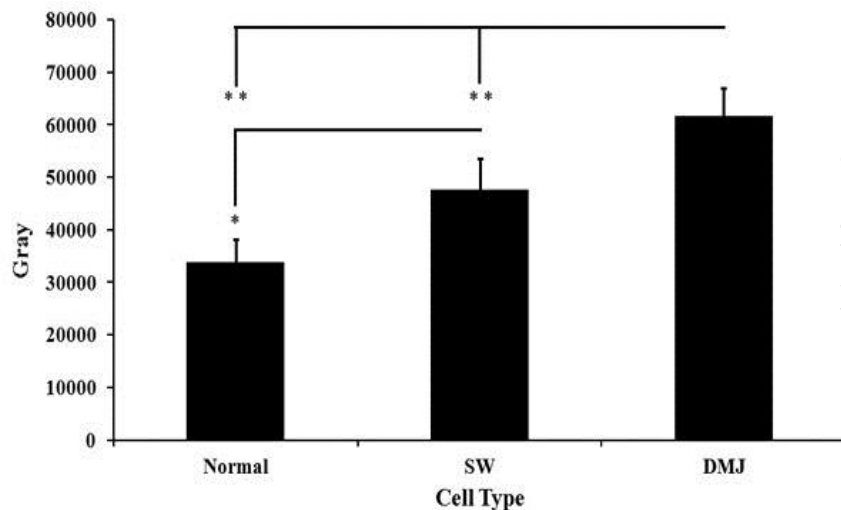
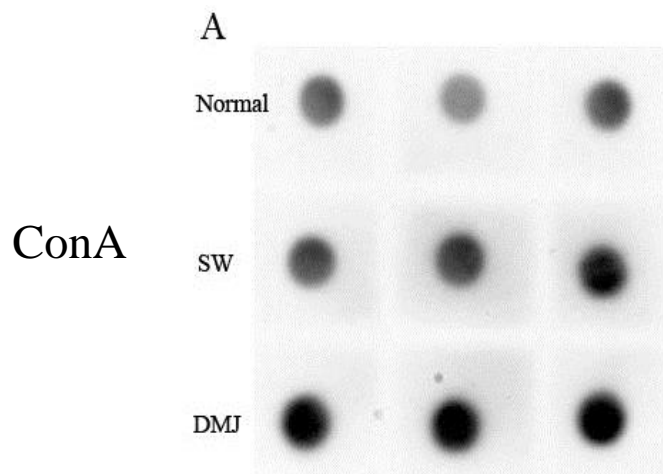
- 人工合成，是 α -甘露糖苷酶 I 的抑制剂。

复合型糖修饰蛋白质发现策略

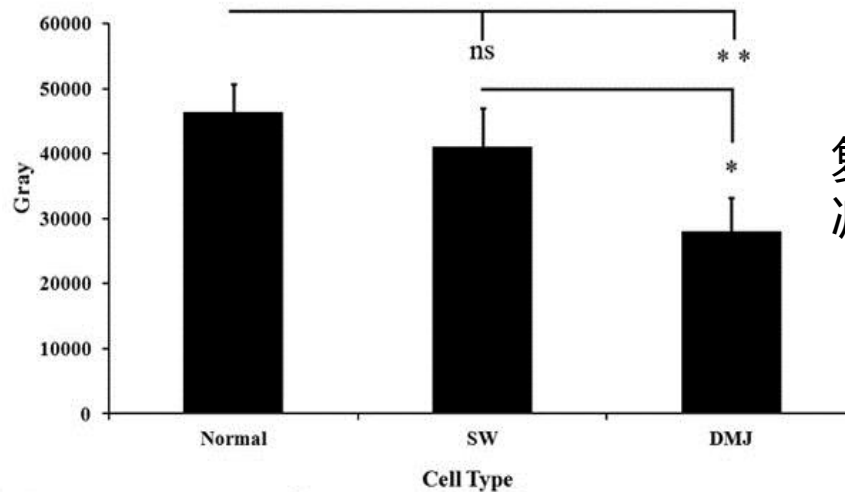
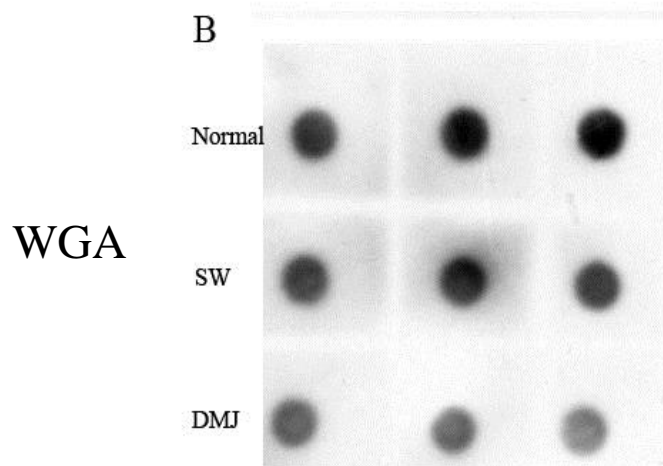


结果

抑制效果的评价-凝集素点杂交

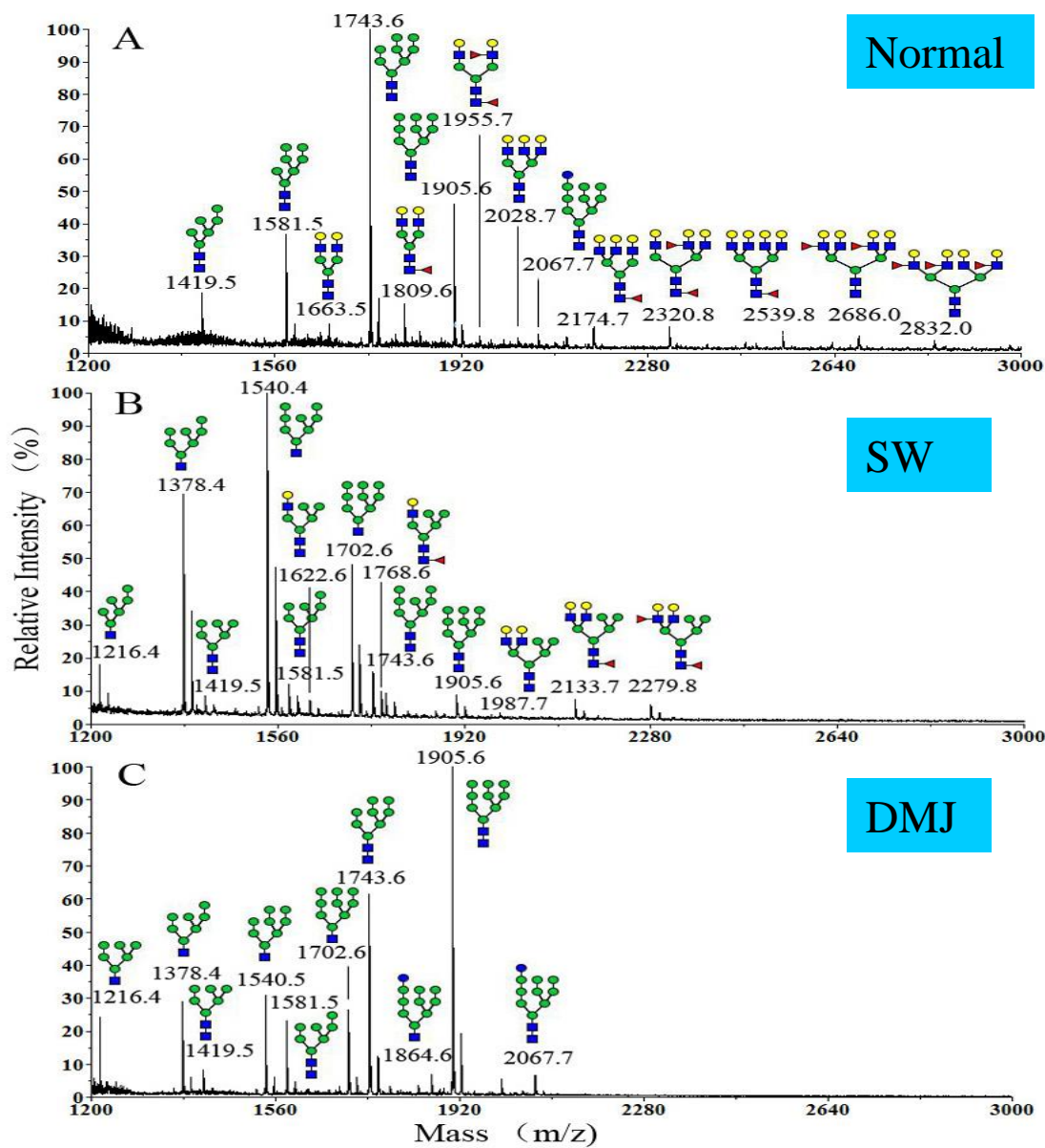


高甘露糖型
和/或杂合型
结构增多



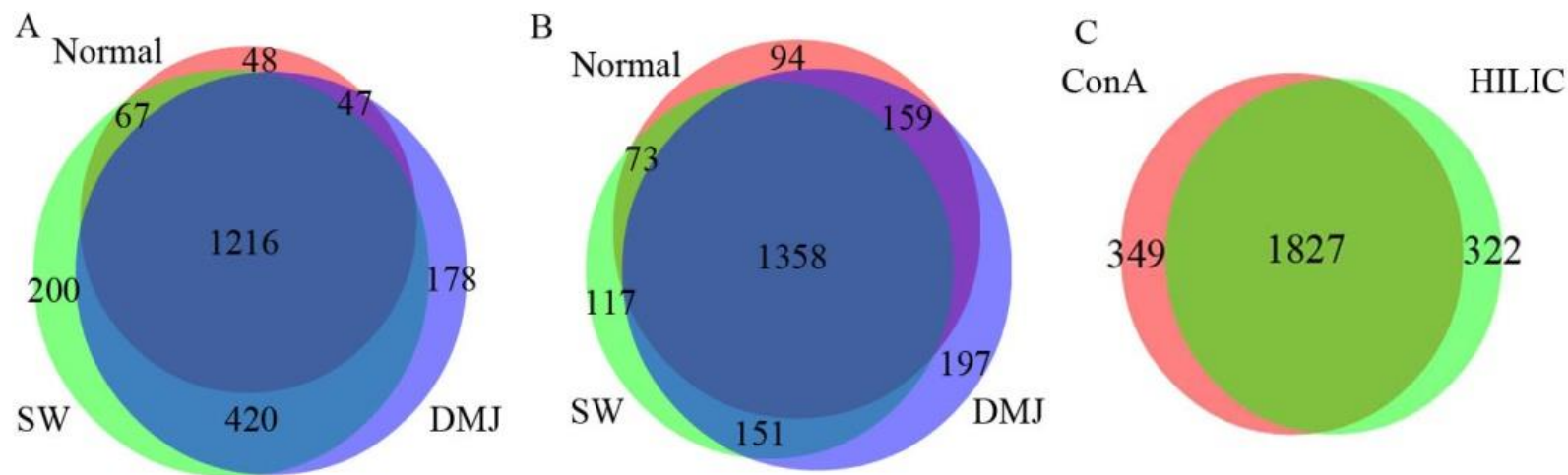
复合型结构
减少

抑制效果的评价- 糖型分析



SW和DMJ成功地
将N-糖抑制为高
甘露糖型/杂合
型或高甘露糖型。

糖基化位点鉴定结果



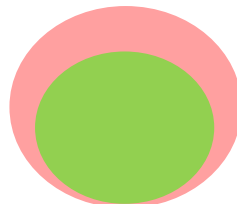
共鉴定到N-糖基化位点2498个，来自898个糖蛋白

	Unique N-glycosites
Known site of known glycoprotein	906
Novel site of known glycoprotein	1209
Novel site of novel glycoprotein	383

N-糖位点所带糖型的分析

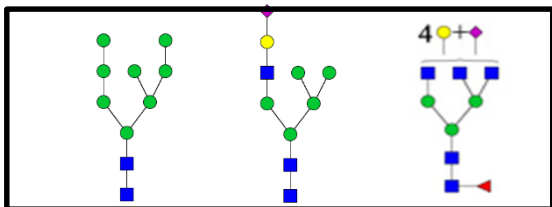
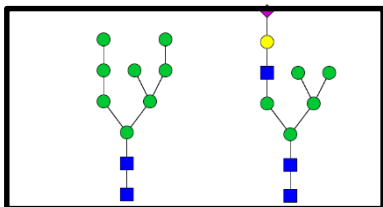
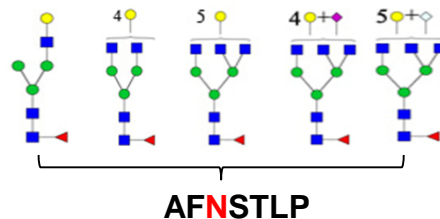
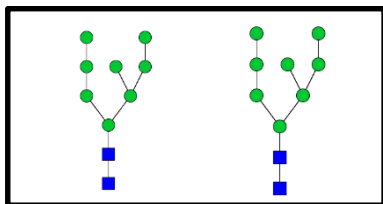
SW+DMJ: 2192个
Normal: 1389个

差减



其他?

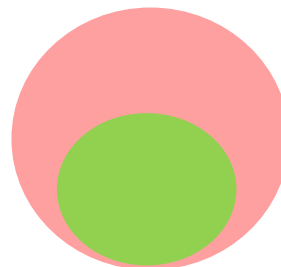
复合型糖链的位点803个



此类位点所在肽段的信号在抑制剂处理后增强。

N-糖位点所带糖型的分析

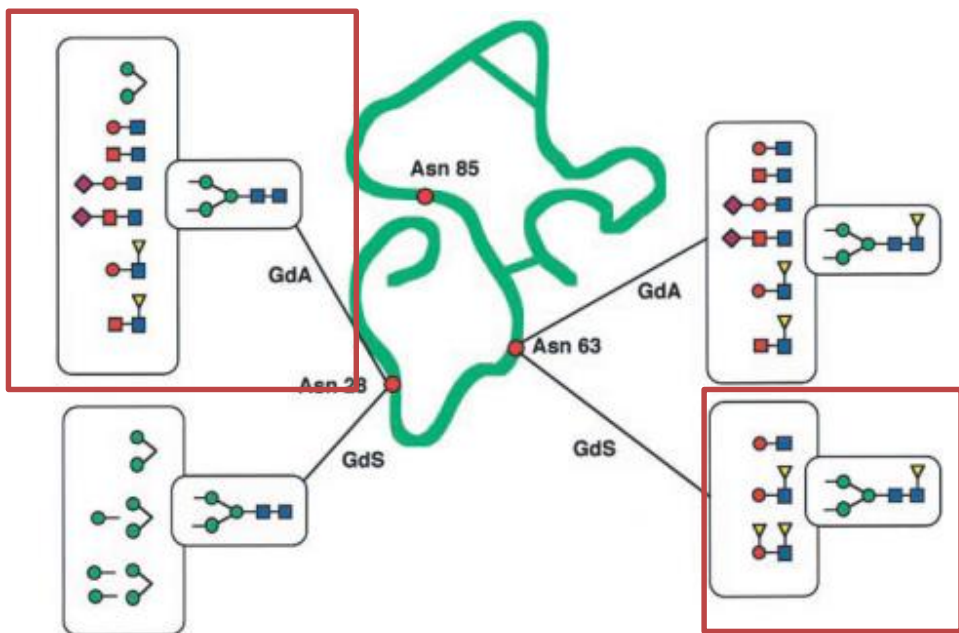
SW+DMJ: 2192个
Normal: 1389个



复合型糖链的位点803个

带有高甘露糖型和/或杂合型糖链?

不一定!

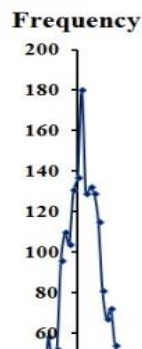


既有高甘露糖型或杂合型，以及复合糖的位点，抑制前后均能鉴定到。此类位点所在肽段的信号在抑制剂处理后会增强。

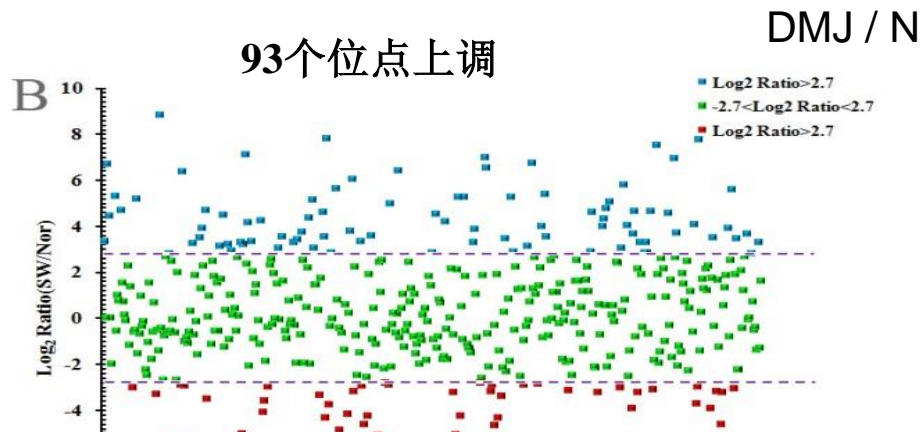
N-糖位点所带糖型的分析

95% 置信区间 $[-2.7, 2.7]$

A



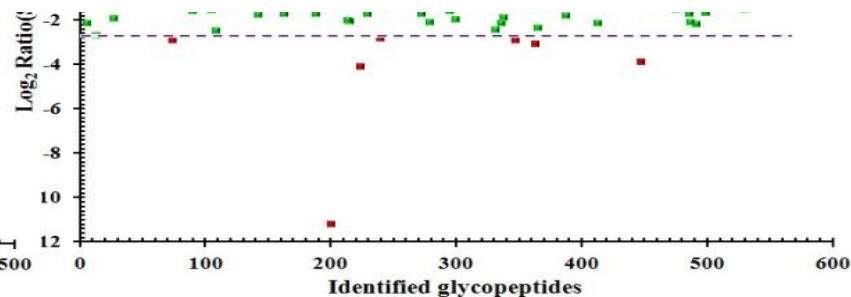
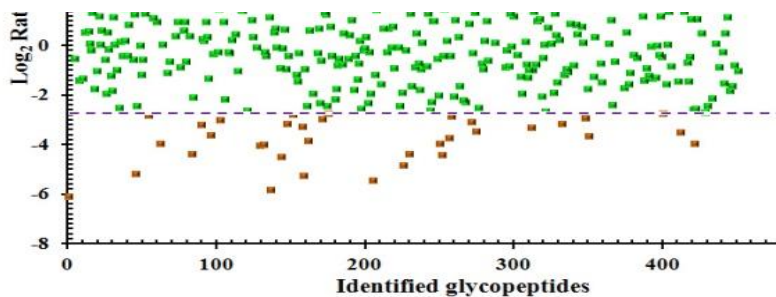
B



带有复合型糖链的位点：803个

带有部分复合型糖链的位点：99个

带有高甘露糖型和/或杂合型糖链的位点：1290个（1389-99）



位点/蛋白的糖型类型验证

仅含高甘露糖和杂合型N-糖的蛋白：393

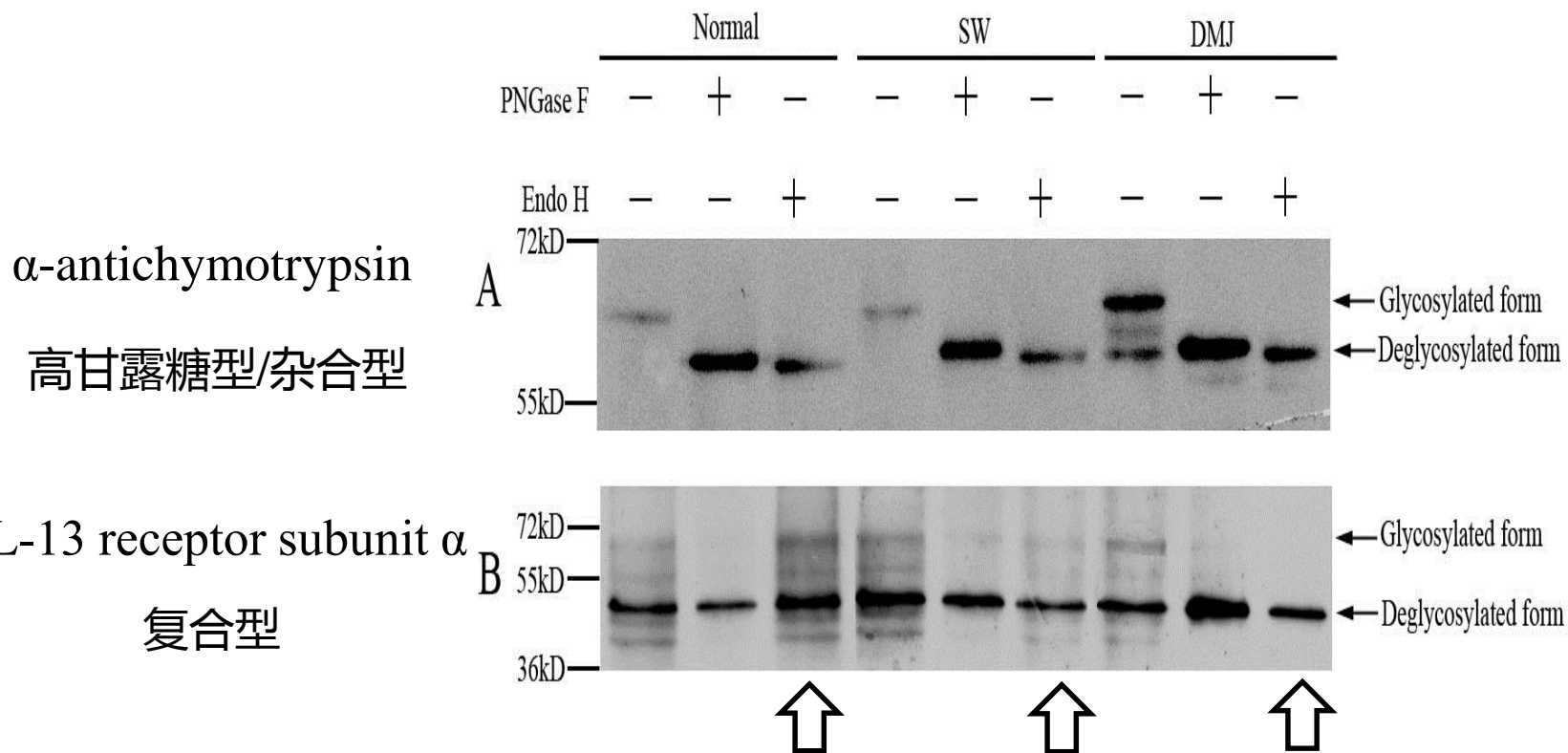
仅含复合型N-糖的蛋白：262

各种糖型都有的蛋白：243

仅含高甘露糖和杂合型N-糖的位点：313

含部分复合型N-糖的位点：45

仅含复合型N-糖的位点：193



与文献报道蛋白质的一致性

高级结构决定糖链结构？

EGFR

Wu S, et al. MCP, 2006

List of glycopeptides of EGFR identified upon stimulation of an A 431 cell line with EGF

Start	End	Peptide sequences ^a	[M + H] ⁺ ^b	Observed charge states ^c	Glycosylation sites ^d
14	56	LTQLGTFEDHFLSLQRMFN*NCEVVLGNLEITYVQRNYDLSFLK	5165.5863	4+, 5+	Asn ³² -p (complex)
57	105	TIQEVAGYVLIALNTVERIPILENLQIIRGNMYENSALAVLSNYDAN*K	5559.8831	4+, 5+	Asn ¹⁰⁴ -n
110	165	ELPMRNLQEILHGAVRFSNNPALCNVESIQWRDIVSSDFLSN*MS-MDFQNHLGSCQK	6533.1104	6+, 7+	Asn ¹⁵¹ -f (complex)
166	185	CDPSCPNG*GSCWGAGEENCQK	2312.8423	2+, 3+	Asn ¹⁷² -n
323	333	DSLSIN*ATNIK	1175.6266	2+, 3+	Asn ³²⁸ -f (high mannose)
337	372	N*CTSISGDLHILPVAFRGDSFTHTPPLDPQELDILK	4004.0221	4+, 5+	Asn ³³⁷ -f (high mannose)
376	407	EITGFLLIQAWPEN*RTDLHAFENLEIIRGRTK	3781.0183	3+, 4+	Asn ³⁸⁹ -p (complex)
408	430	QHGQFSLAVVSLN*ITSLGLRSLK	2468.3983	3+, 4+	Asn ⁴²⁰ -f (complex)
477	514	ATGQVCHALCSPEGCWGPEPRDCVSCRN*VSRGRECVDK	4446.9423	4+, 5+, 6+, 7+	Asn ⁵⁰⁴ -f (complex)
515	569	CNLEGEPRFEVENSECIQCHPECLPQAMN*ITCTGRGPDNCIQCAHYIDGPHCVK	6570.8425	5+, 6+, 7+, 8+	Asn ⁵⁴⁴ -p (complex)
570	585	TCPAGVMGEN*NTLWVK	1776.8407	2+, 3+	Asn ⁵⁷⁹ -p (complex)
586	618	YADAGHVCHLCHPN*CTYGCTGPGLEGCPNGPK	3657.5184	3+, 4+, 5+	Asn ⁵⁹⁹ -f (high mannose)

High mannose/
hybrid

complex

Thaysen-Andersen M, et al. Anal Chem, 2009

P01857 Ig gamma-1 chain C region EEQYnSTYR 180 Complex

complex

Rudd PM, et al. Biochemistry, 1999

P80188 Neutrophil gelatinase-associated lipocalin LCN2 SYnVTSVLFR 85 Complex

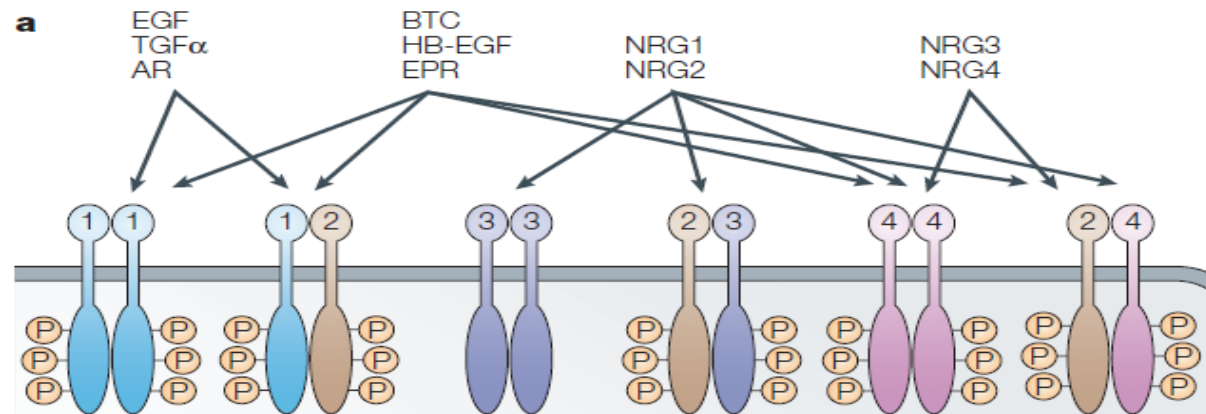
complex

部分新发现糖型信息

受体酪氨酸蛋白激酶erbB家族

蛋白质名	基因	匹配肽段	位点	推测糖型
Receptor tyrosine-protein kinase erbB-2	ERBB2	NNQLALTLIDTnR (S)	187	Complex
Receptor tyrosine-protein kinase erbB-2	ERBB2	GPGPTQCVnCSQF	530	Complex
Receptor tyrosine-protein kinase erbB-2	ERBB2	HFnHSGICELHCPALVTY	259	Complex
Receptor tyrosine-protein kinase erbB-2	ERBB2	CQPQnGSVTCFGPE	571	Complex
Receptor tyrosine-protein kinase erbB-3	ERBB3	NTnSSHALRQL	126	Complex
Receptor tyrosine-protein kinase erbB-3	ERBB3	HFnDSGACVPR	250	High-mannose and/or hybrid
Receptor tyrosine-protein kinase erbB-3	ERBB3	FQTVDSSNIDGFVnCTK	353	High-mannose and/or hybrid
Receptor tyrosine-protein kinase erbB-3	ERBB3	SnLTTIGGRSLY	414	High-mannose and/or hybrid
Receptor tyrosine-protein kinase erbB-3	ERBB3	NLnVTSLGFR	437	High-mannose and/or hybrid
Receptor tyrosine-protein kinase erbB-3	ERBB3	QLCYHHSLnWTK	469	High-mannose and/or hybrid

None of the ligands bind ERBB2, but ERBB2 is the preferred dimerization partner for all the other ERBB receptors.- Nat Rev Cancer 2005



部分新发现糖型信息

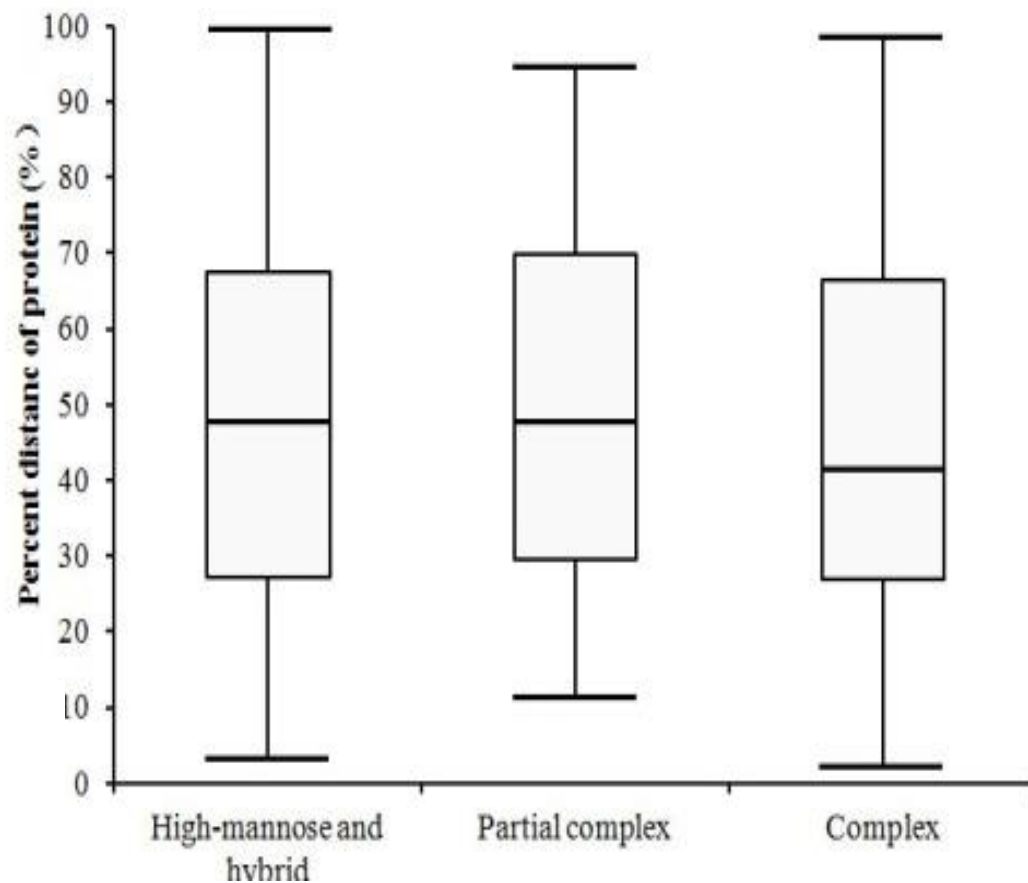
成纤维生长因子受体家族蛋白

蛋白质名	基因	匹配肽段	位点	推测糖型
Fibroblast growth factor receptor 1	FGFR1	SPHRPILQAGLPAnK(T)	264	Complex
Fibroblast growth factor receptor 1	FGFR1	TAGVnTTDKEMEVHLHR	317	Complex
Fibroblast growth factor receptor 1	FGFR1	DGVQLAESnR(T)	77	Complex
Fibroblast growth factor receptor-like 1	FGFRL1	AGAI nATYK	231	Complex
Fibroblast growth factor receptor-like 1	FGFRL1	GAEGRHnSTIDVGGQKF	293	Complex
Fibroblast growth factor receptor 3	FGFR3	GnYTCVVENK	225	Complex
Fibroblast growth factor receptor 3	FGFR3	SLHnVTFEDAGEY	328	Complex
Fibroblast growth factor receptor 4	FGFR4	SPHRPILQAGLPAnTTAVVGS DVELLCK	258	Complex
Fibroblast growth factor receptor 4	FGFR4	HIVInGSSFSGADGFPYVQVLK	290	Complex
Fibroblast growth factor receptor 4	FGFR4	TADInSSEVEVLYLR	311	High-mannose and/or hybrid
Fibroblast growth factor receptor 4	FGFR4	LRnVSAEDAGEY	322	High-mannose and/or hybrid
Fibroblast growth factor receptor 3	FGFR3	LQVLnASHEDSGAYSCR	98	High-mannose and/or hybrid

复合糖位点更靠近蛋白质的N端

距N-端的距离

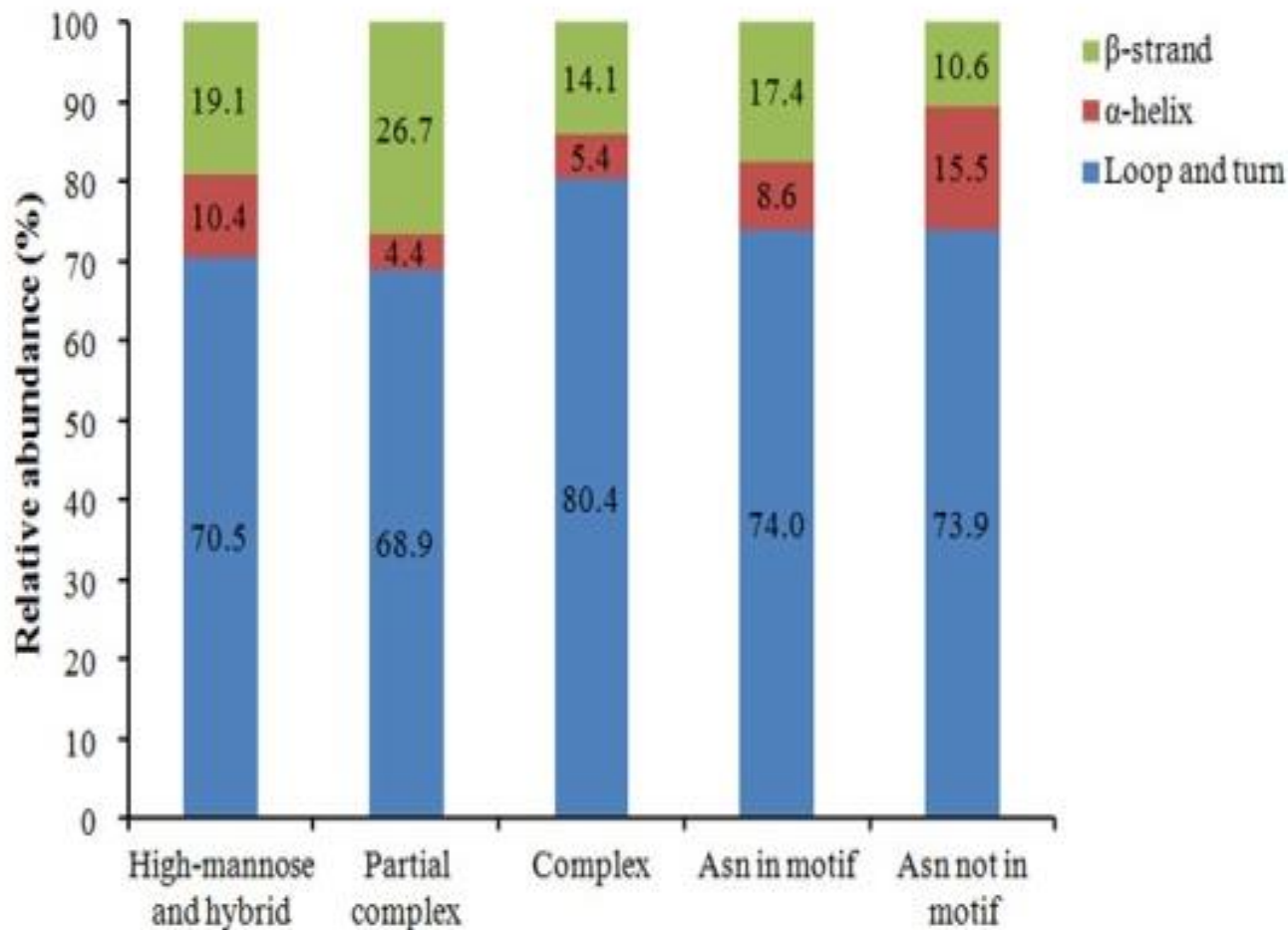
一级结构



复合型糖位点位于倾向于环和转角位置

二级结构分布

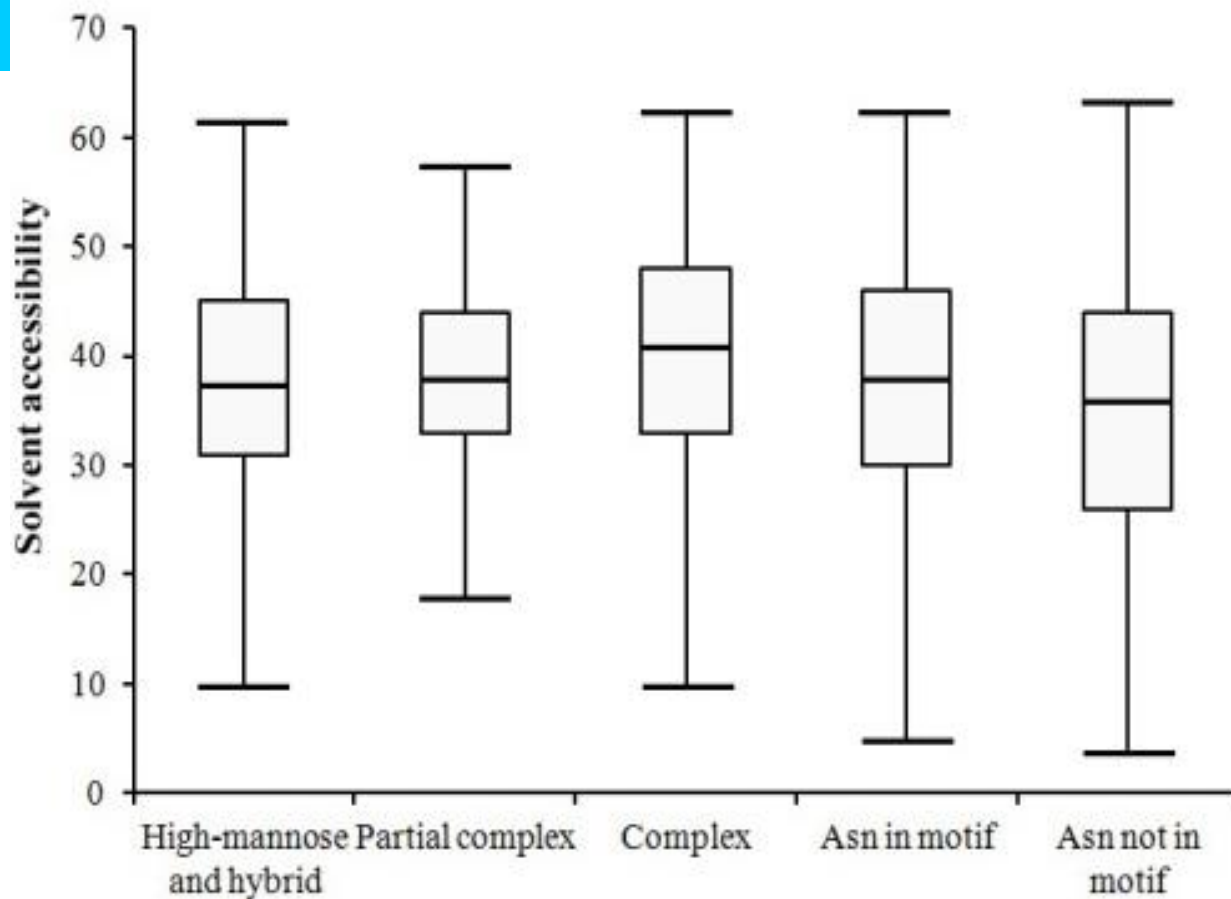
二级结构



不同糖型N-糖蛋白的溶液可接近性

溶液可接近性

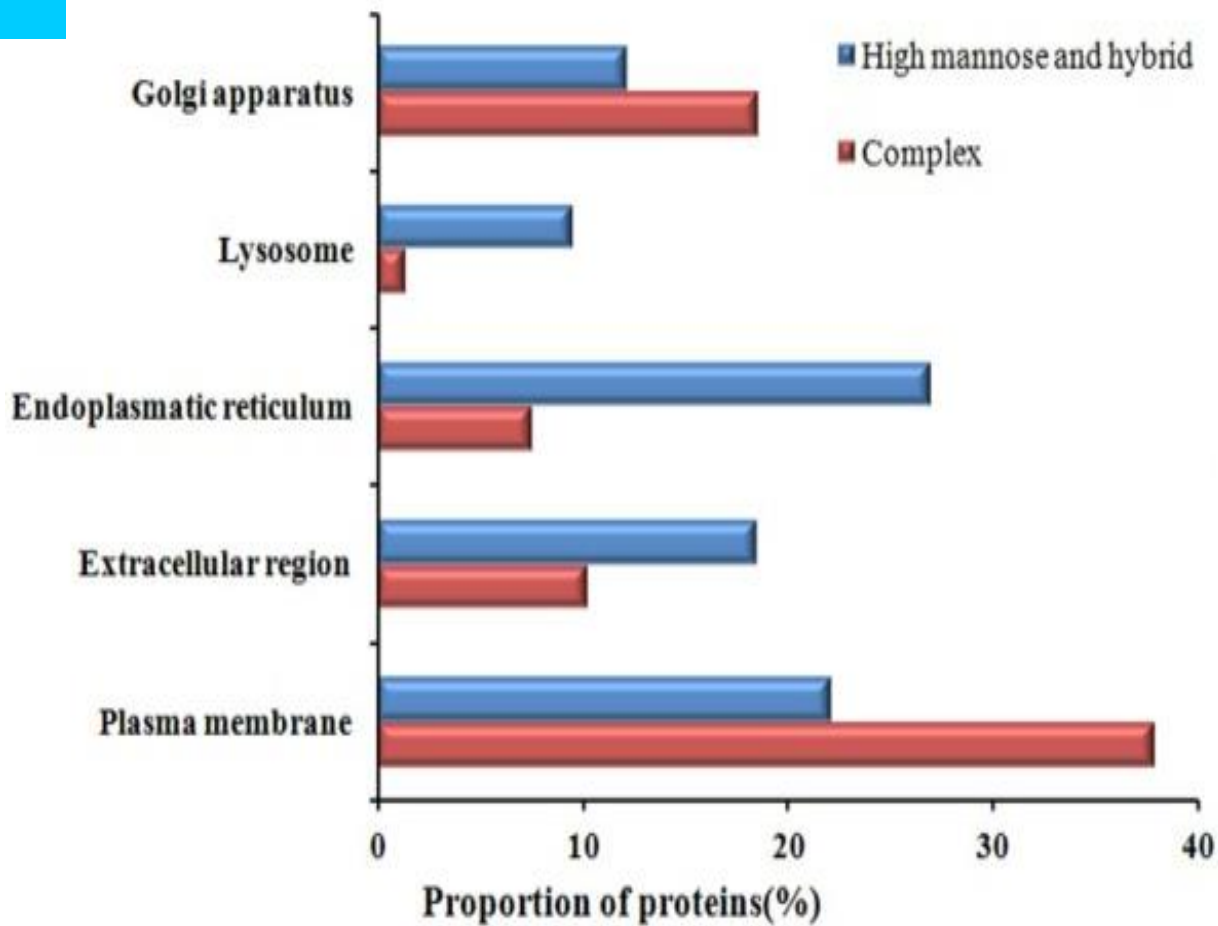
三级结构



带有复合型糖链的糖基化位点的溶液可接近性略大于带有高甘露糖型和/或杂合型糖链的位点。

糖型参与蛋白质的定位过程

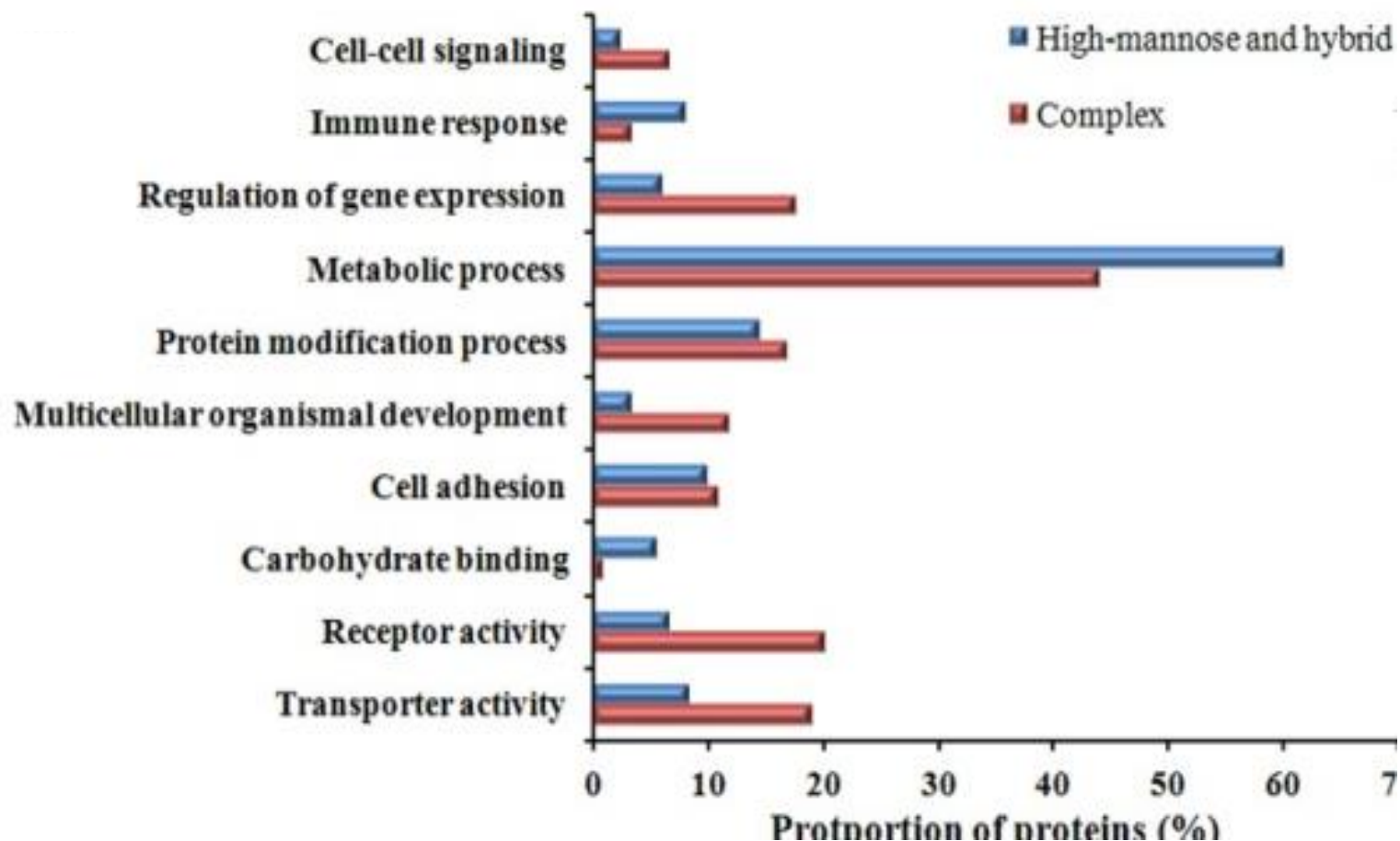
定位



复合糖参与信号传导与调控

高甘露糖参与保守生物学过程

功能



小结

发展了一种N-糖蛋白质组学研究策略，在位点鉴定的同时确定N-糖基化类型。

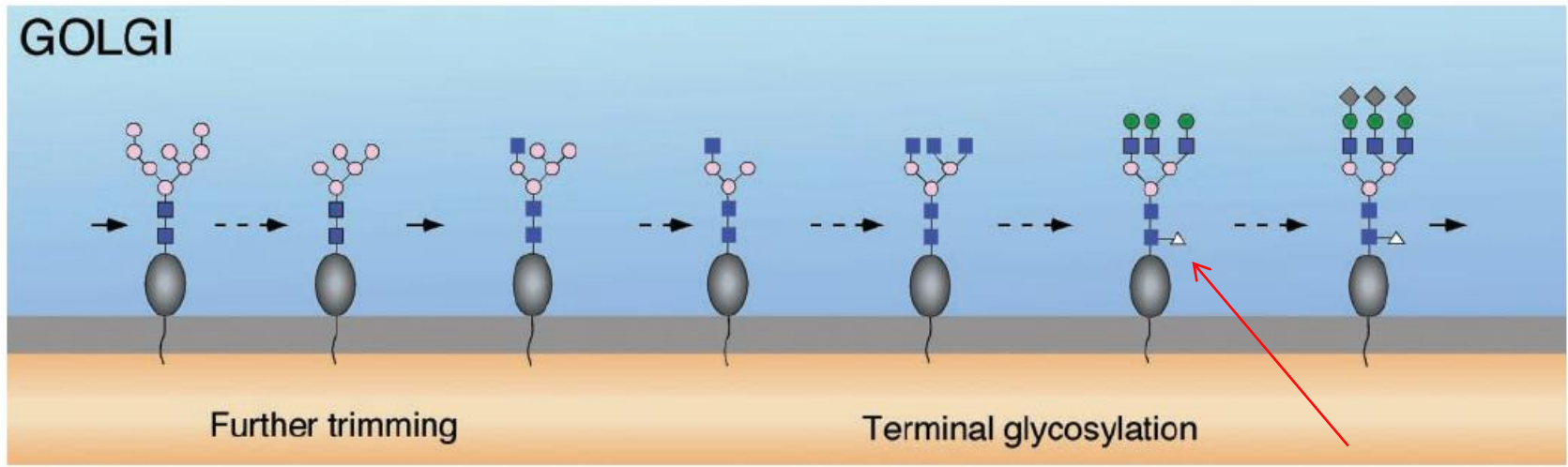
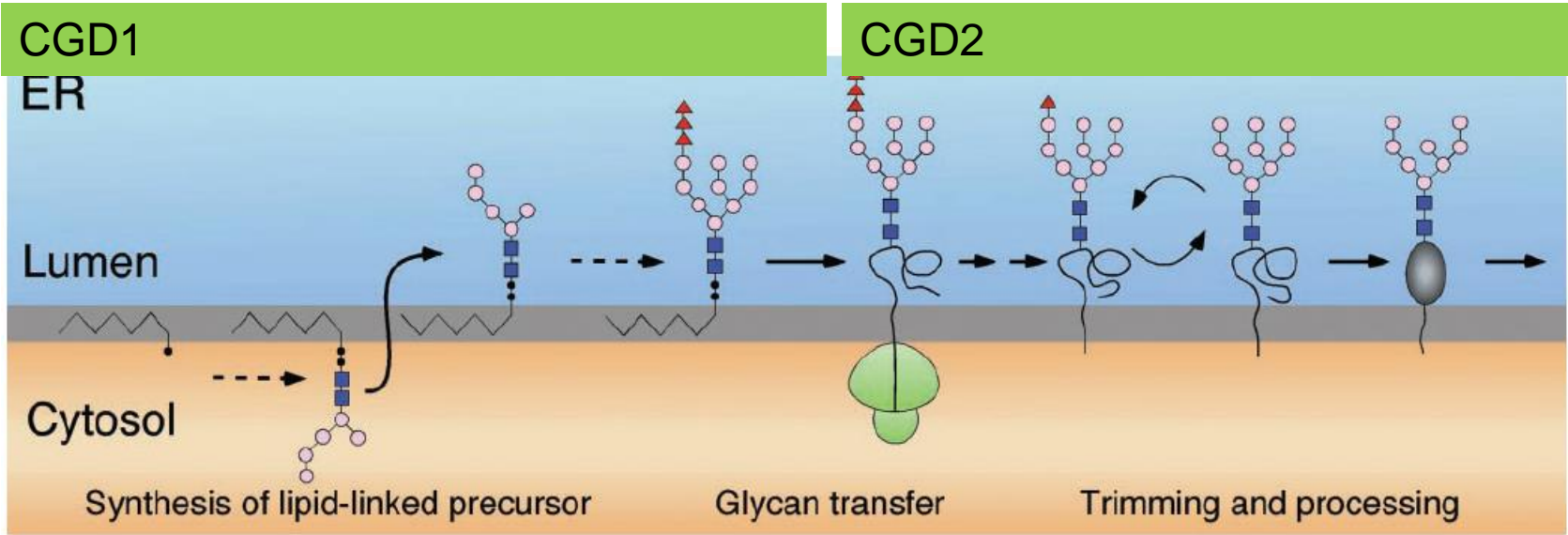
获得了相同来源的具有糖型信息的N-糖基化位点数据集。

糖基化位点所处的一级和二级结构对该位点上所带的糖型类型的影响非常轻微，而受到蛋白质与糖链之间的相互作用影响的蛋白质的三维结构可能对糖链的加工过程起着决定性的作用。

带有不同类型N-糖链的糖蛋白分别在不同的亚细胞器中得到了富集，糖链结构可能是重要的蛋白分选信号。

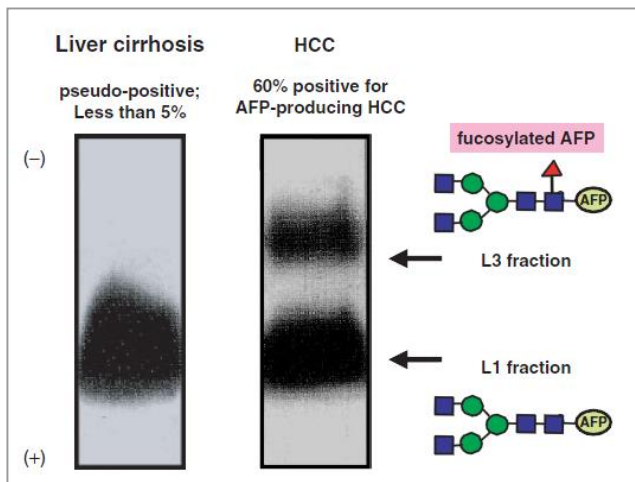
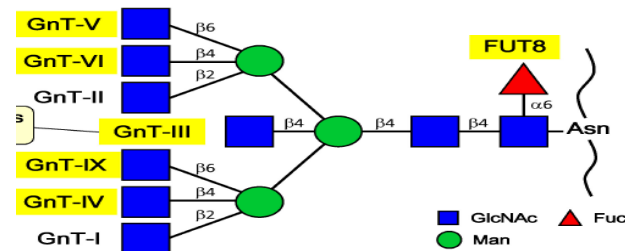
核心岩藻糖化肽段特异性定性定量分析

N-糖链加工过程



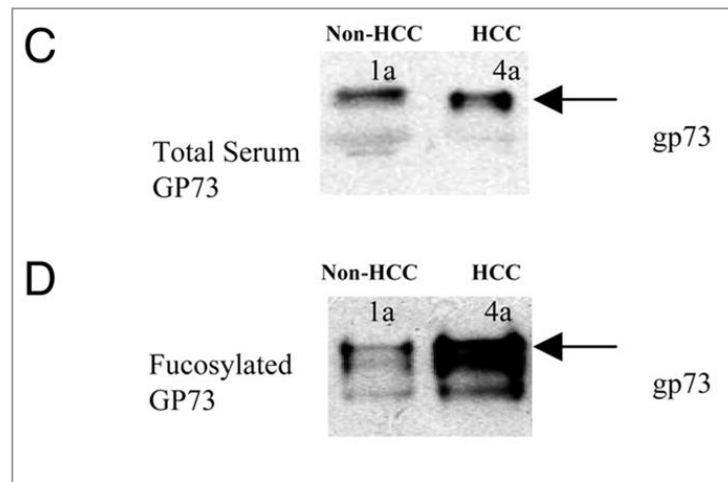
核心岩藻糖化蛋白质与肿瘤诊断

已报道的核心岩藻糖化修饰与多种肿瘤发生密切相关



AFP-L3 used as a HCC biomarker

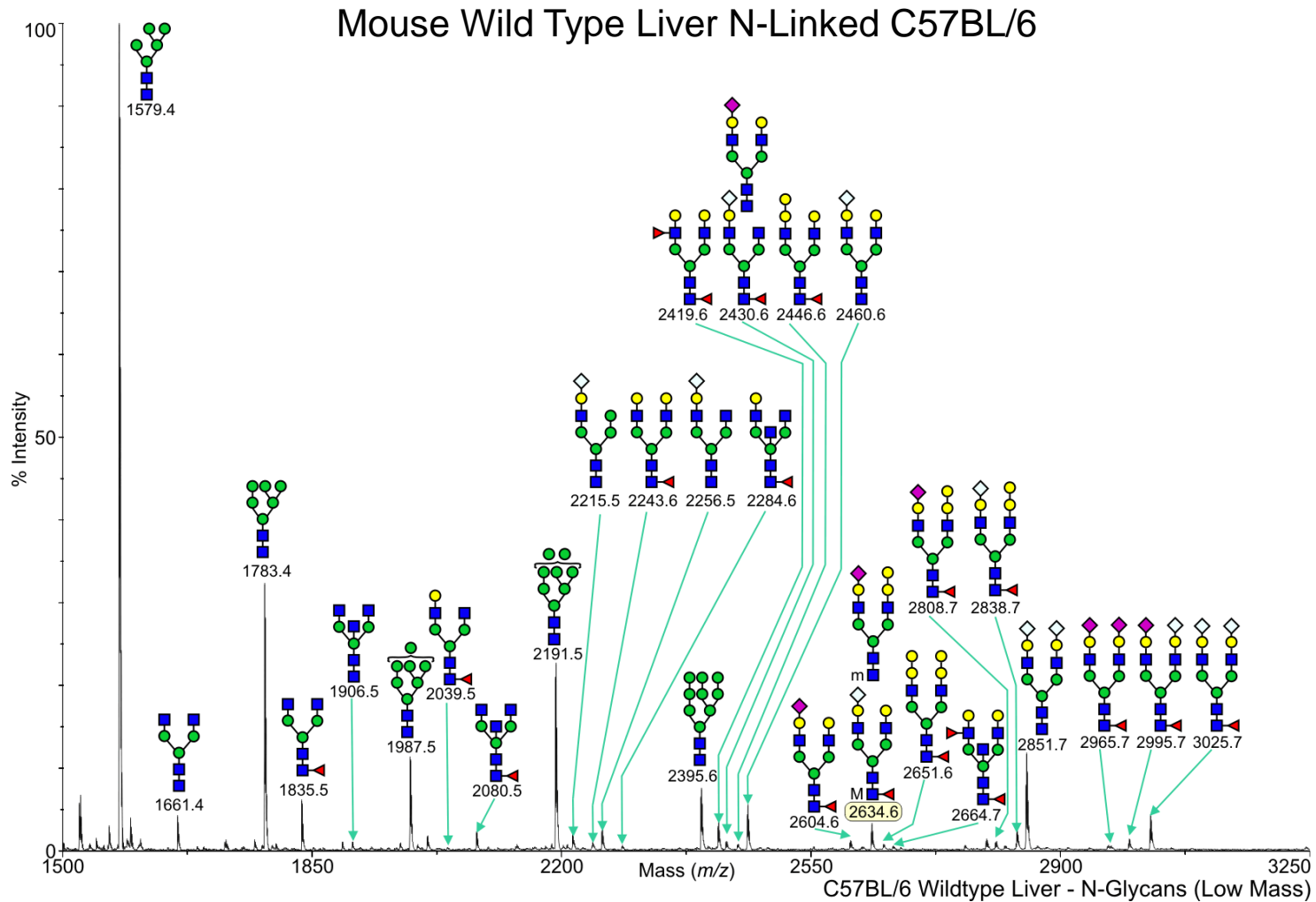
Miyoshi E et al. *J. Biochem.*, 2009, 143: 725-9



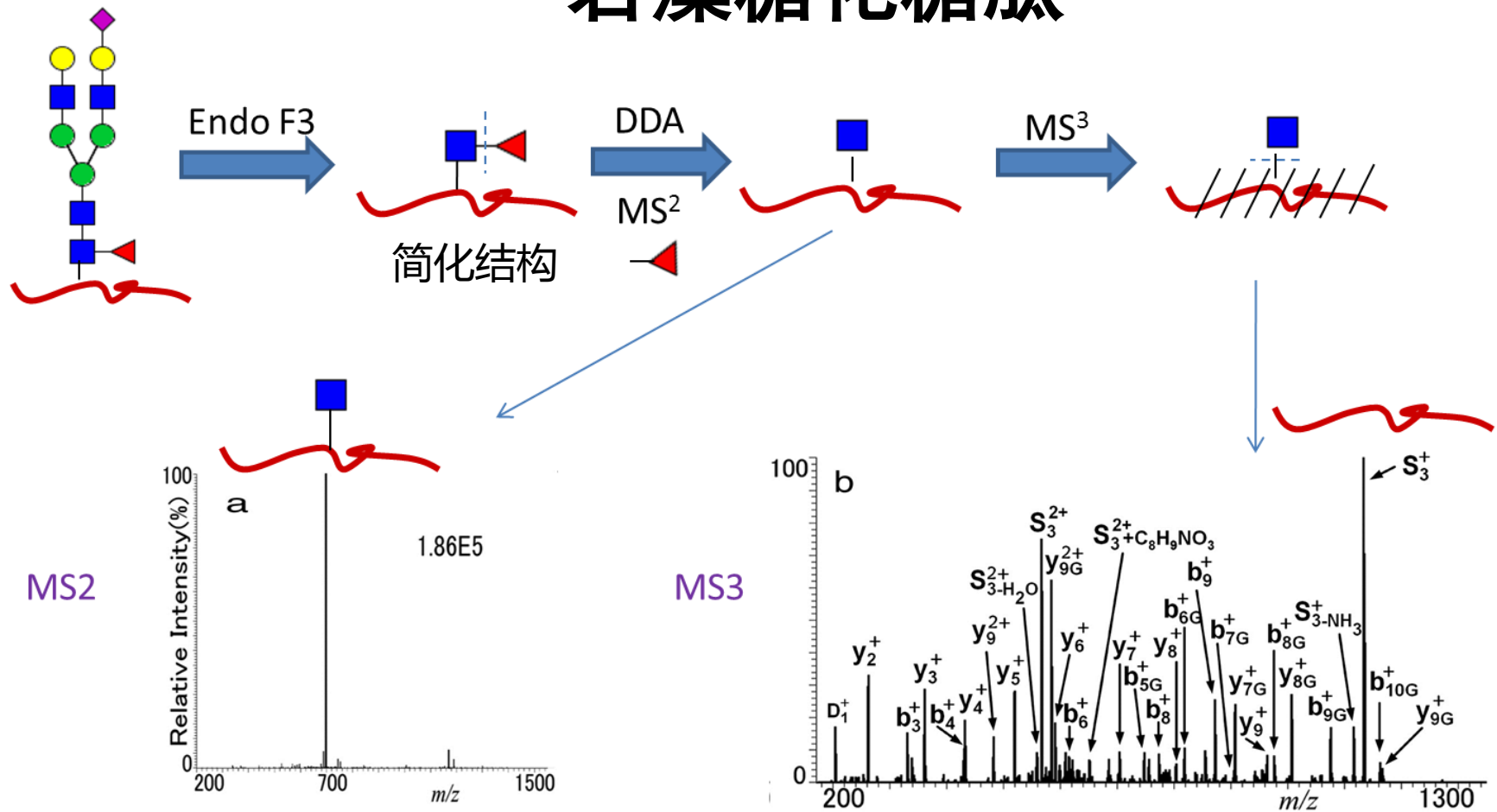
GP73 as a potential HCC biomarker

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

核心岩藻糖化修饰提示复合或杂合糖修饰的存在



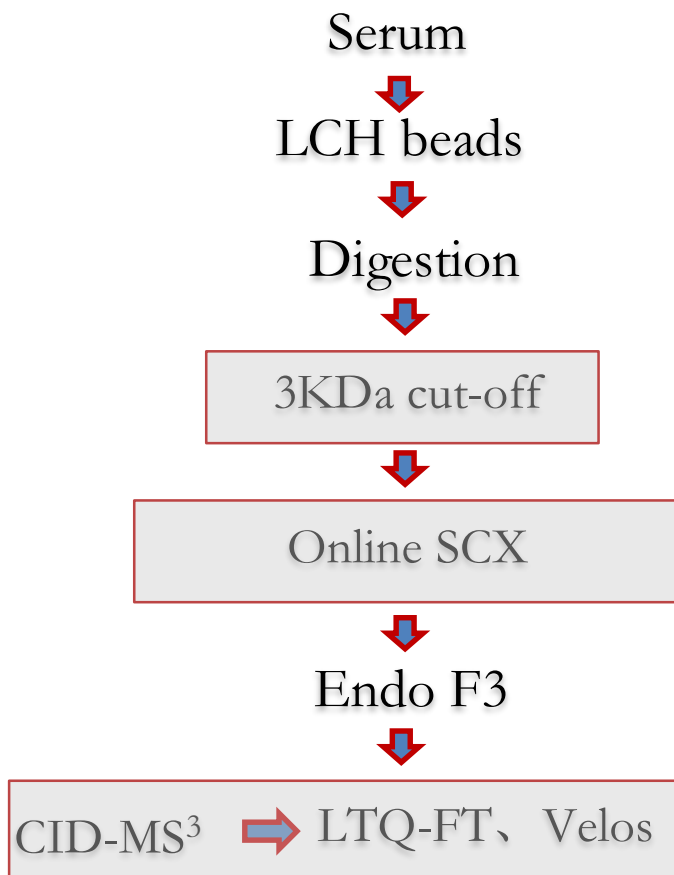
基于中性丢失的三级质谱鉴定核心 岩藻糖化糖肽



Jia W, Ying W et al. *Mol Cell Proteomics*, 2009
 Zhao Yan, Ying W et al. *Anal Chem*, 2011

12个组分，24小时，100个糖基化位点

核心岩藻糖化蛋白研究策略存在问题

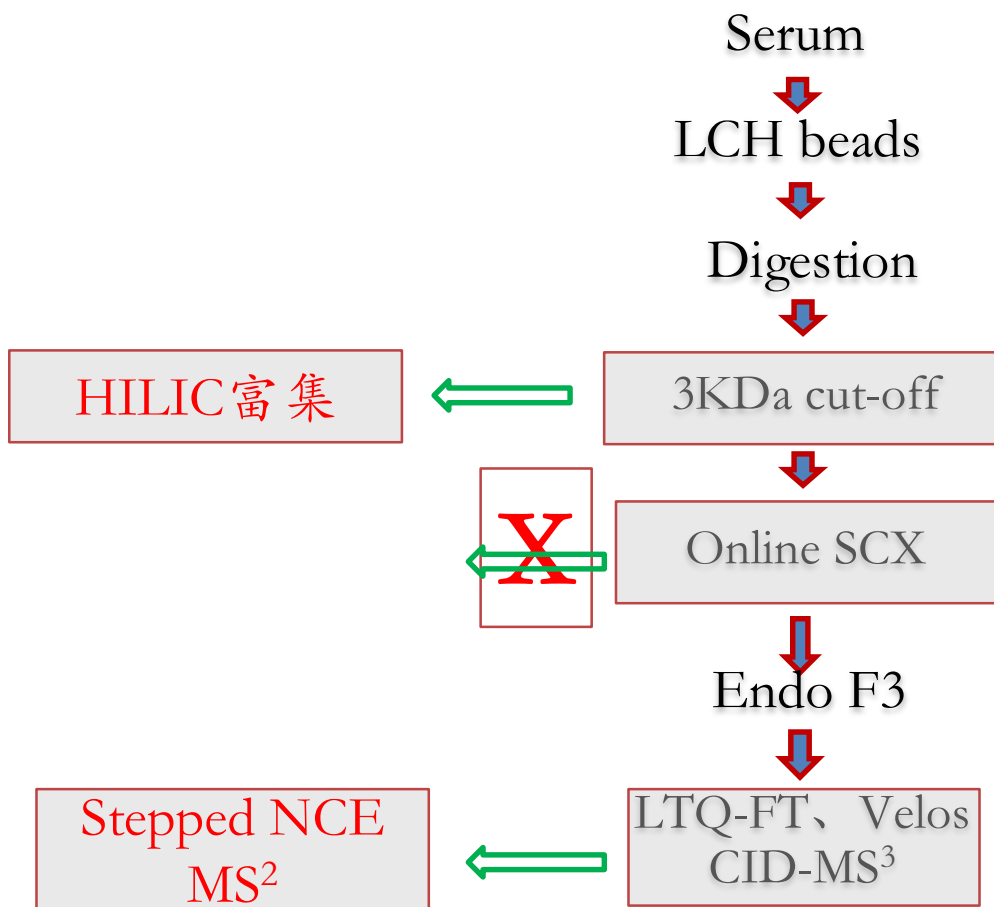


选择性差

12个组分，分析周期长，
定量重复性差

优点：保留NL，利于位点确证
缺点：扫描速度慢，通量受限

策略改进



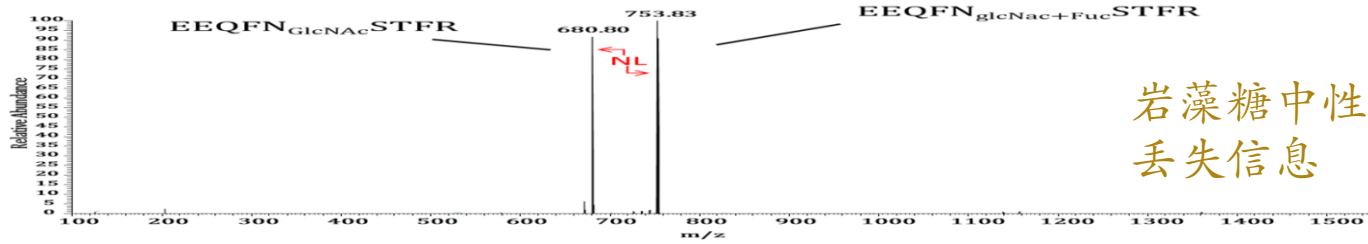
选择性差

12个组分，分析周期长，
定量重复性差

优点：保留NL，利于位点确证
缺点：扫描速度慢，通量受限

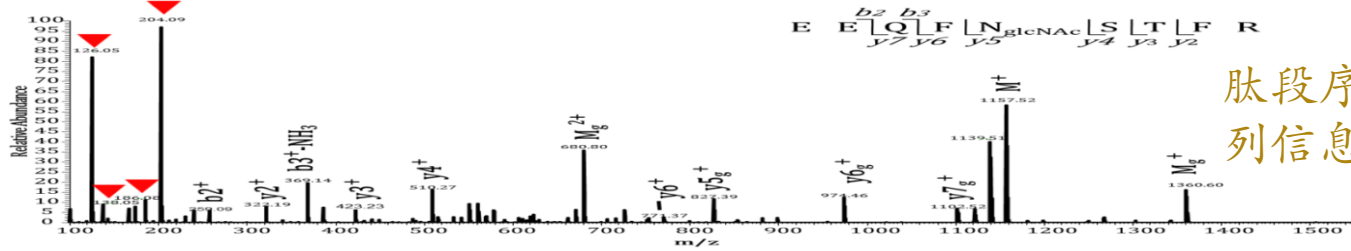
简化核心岩藻糖化肽段的步进（二级）碎裂规律

NCE=10



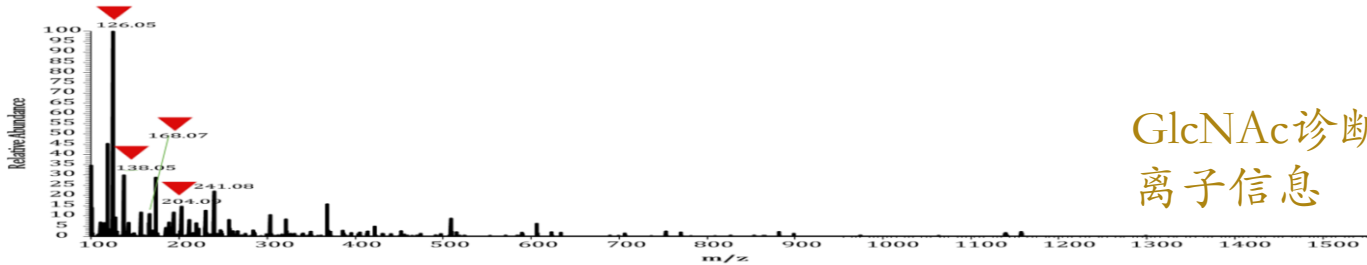
岩藻糖中性
丢失信息

NCE=22



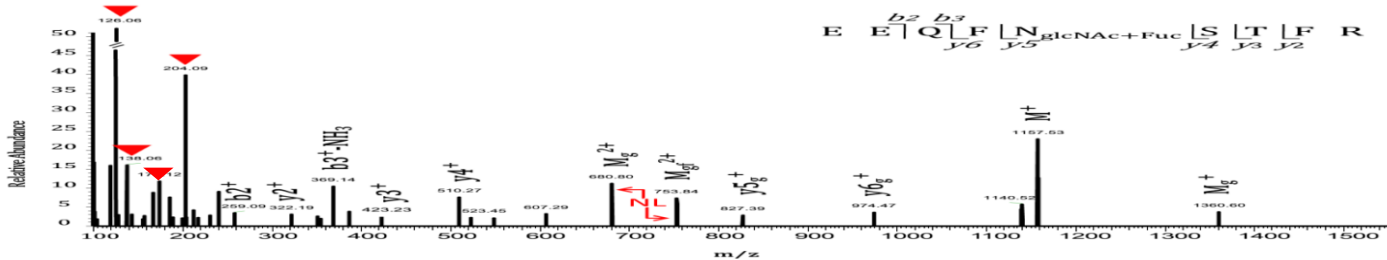
肽段序
列信息

NCE=34

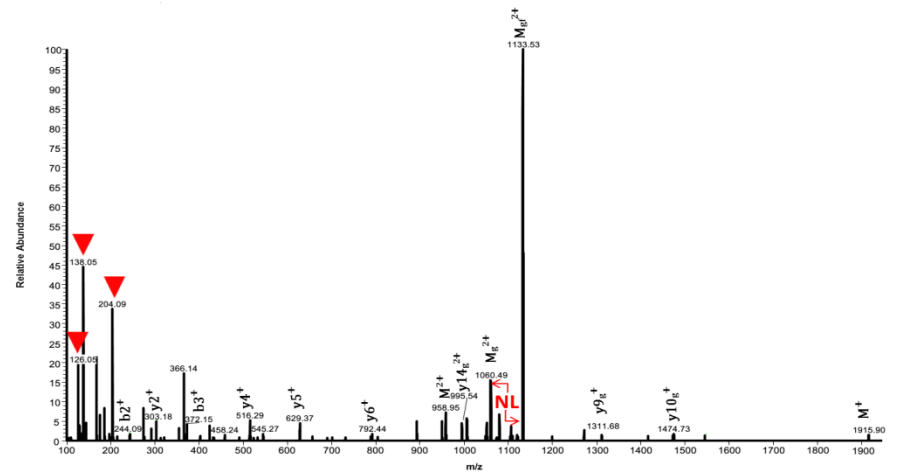
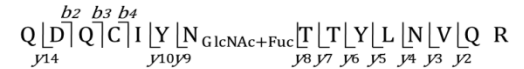
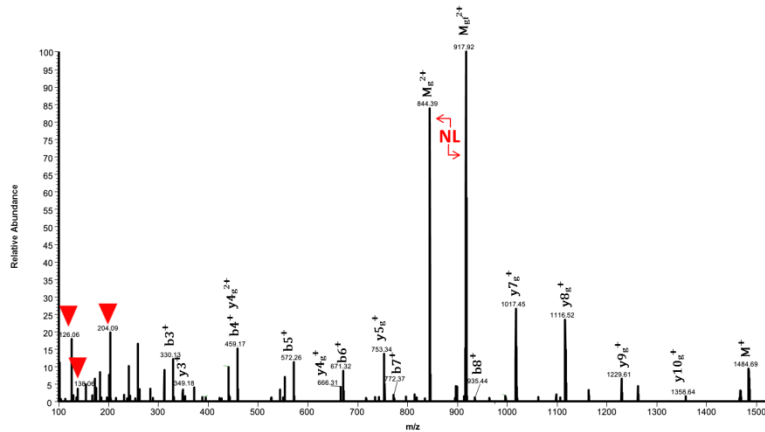
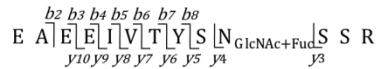
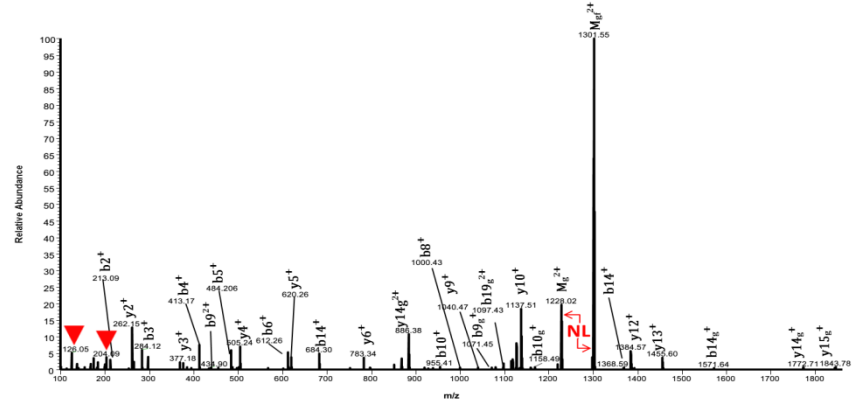
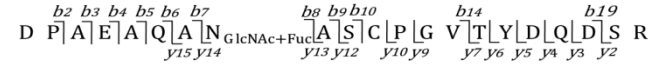
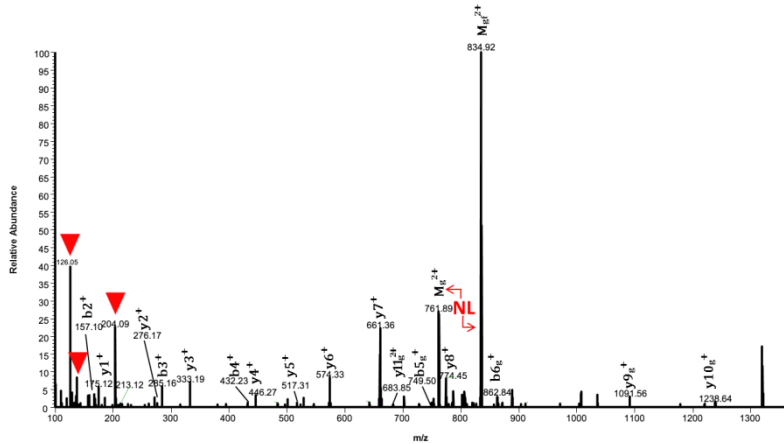
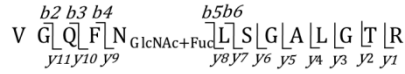


GlcNAc诊断
离子信息

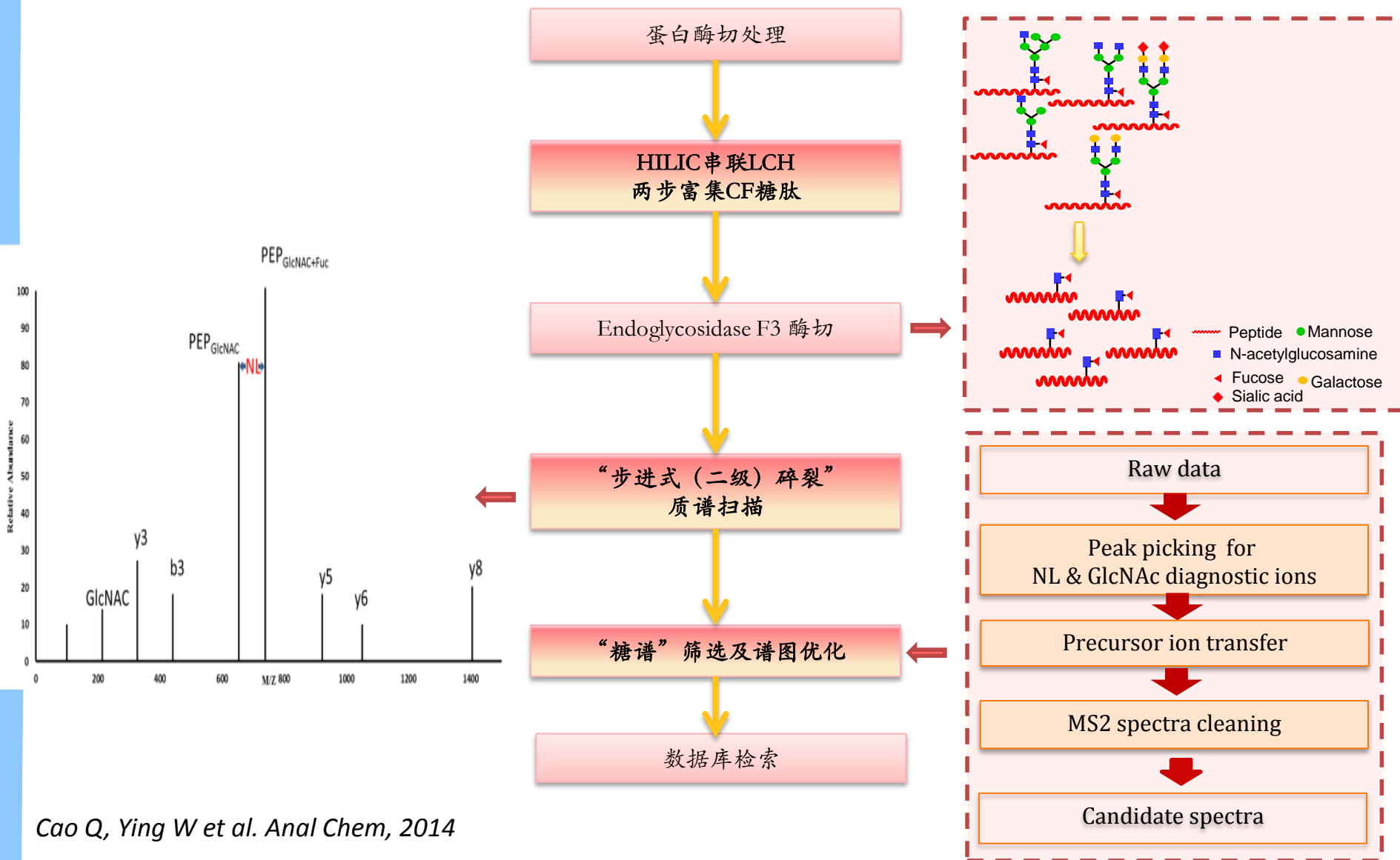
Stepped NCE



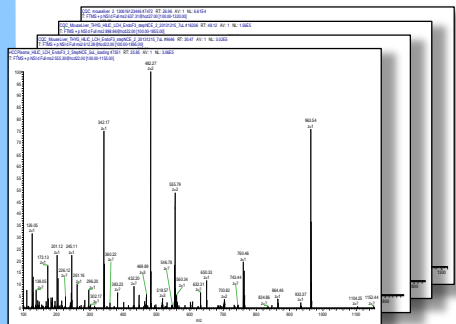
简化核心岩藻糖化肽段的步进碎裂规律



基于“步进式碎裂”扫描及“糖谱”筛选处理的简化糖肽鉴定策略

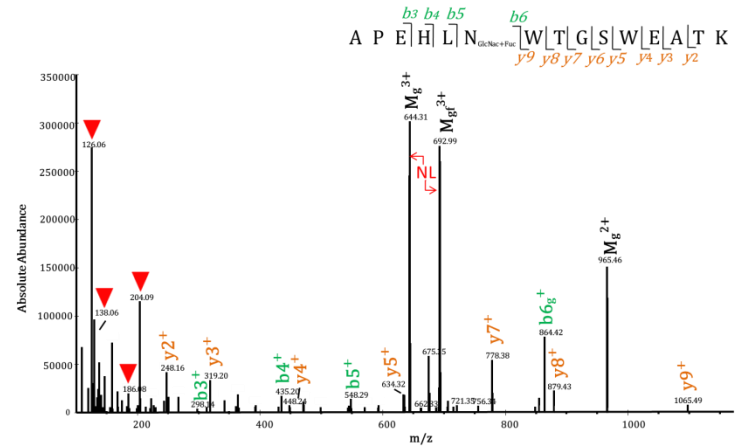


谱图挑选及简化处理

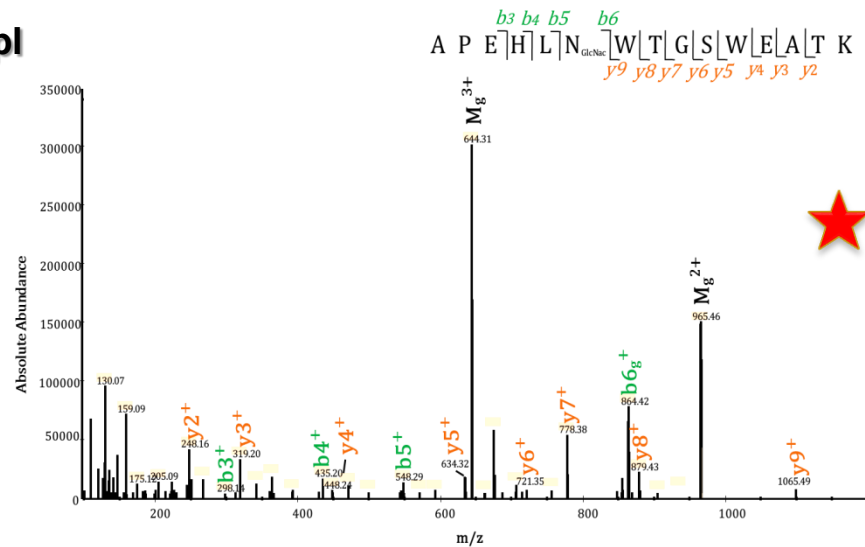
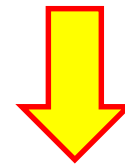


CF_StepNCE.pl

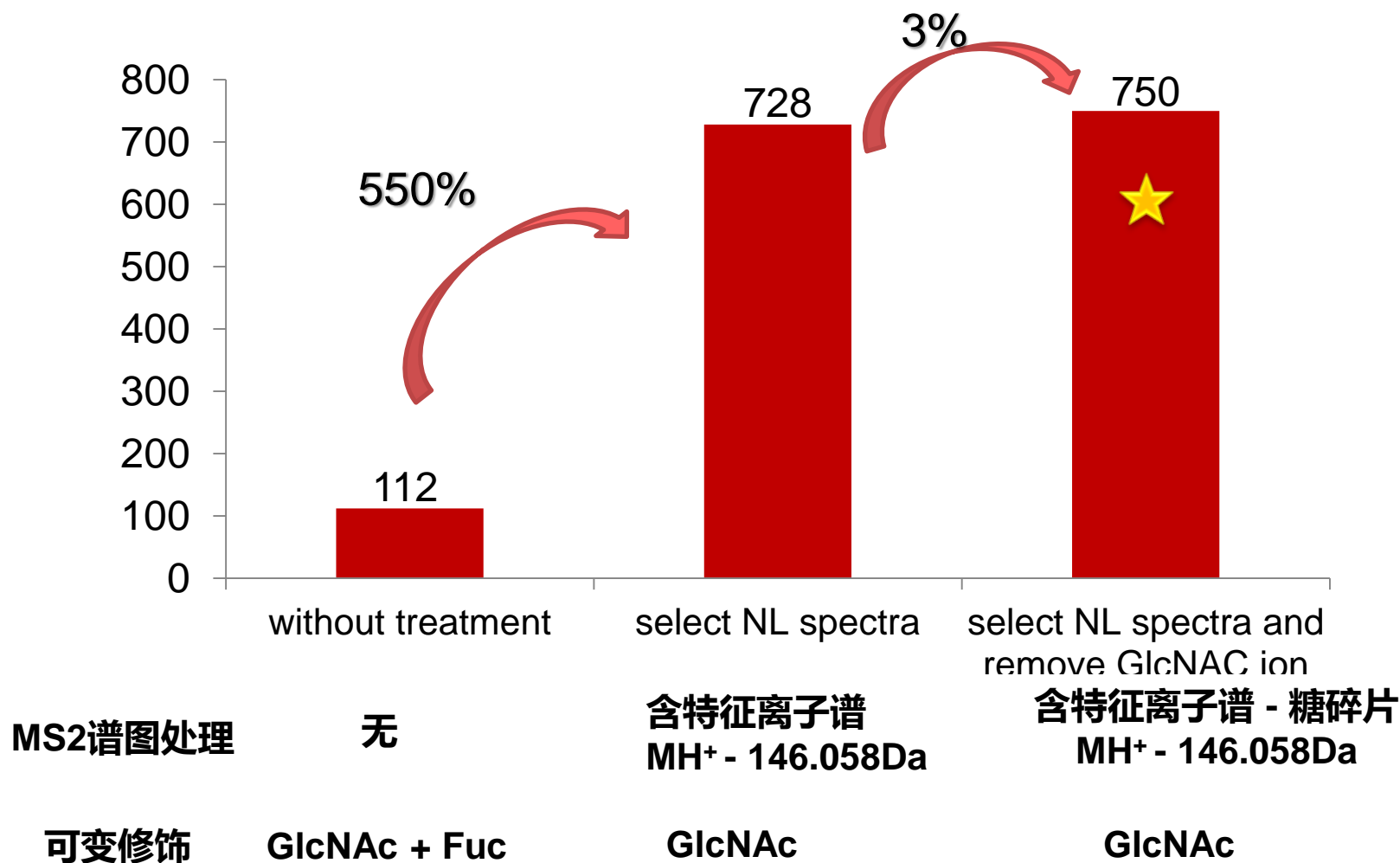
糖谱
挑选



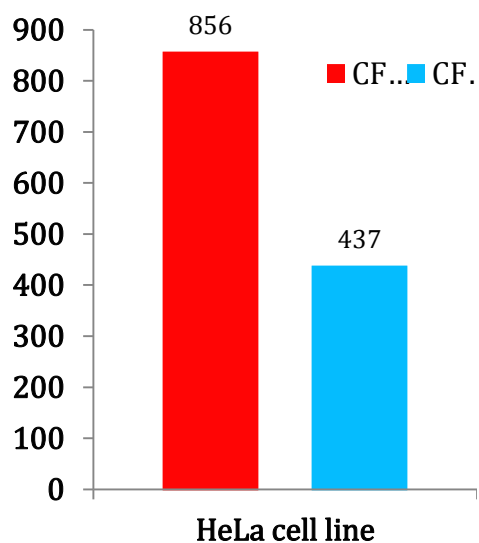
谱图优化
去除糖碎片



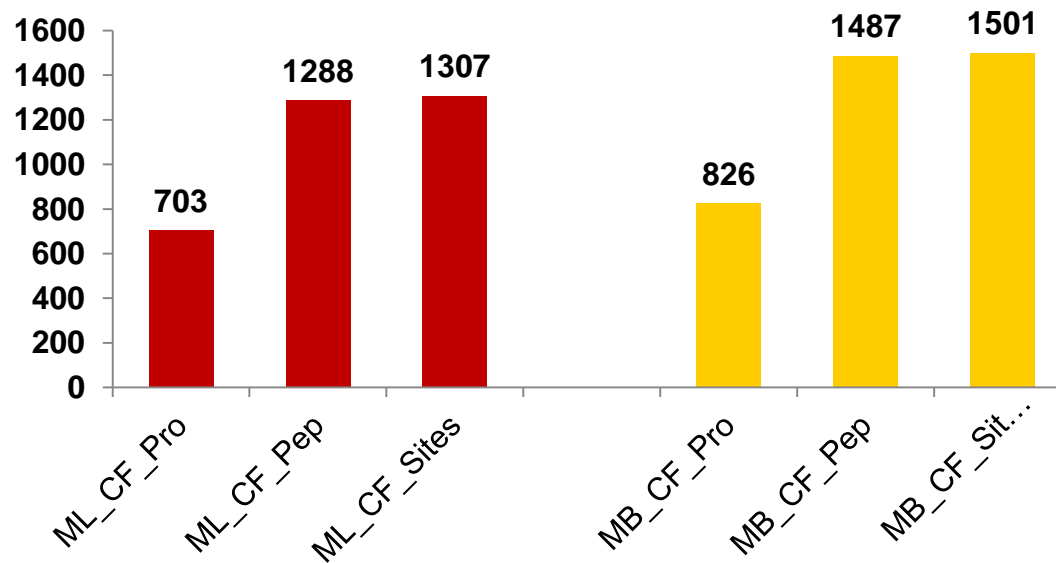
质谱数据经“糖谱”筛选及谱图优化处理后，CF鉴定效率显著提升



核心岩藻糖化蛋白质大规模鉴定



HeLa 细胞系

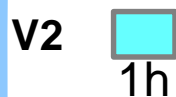


小鼠肝脏组织

小鼠脑组织

新 (V2) 老 (V1) 方法效率比较

消耗质谱机时



位点覆盖规模



Cao Q et al. Anal Chem, 2014

下一步



肿瘤组
织、体液



CF糖肽
富集



MS非标定
量分析



候选靶
标验证

核心岩藻糖化蛋白质组定量研究

0.0 2.0 ∞

HCC/Healthy



AFP, E-select...

Apolipoprotein B (14 fold), A1AT(3.6fold), GP73(3.5fold)...

定量指标:
PSMs by
pFind

肝癌患者血浆样本中鉴定的差异CF蛋白质列表 (部分)

GI Number	CF glycoprotein	HCC up-regulation fold
105990532	Apolipoprotein B-100	14.0
21071030	Alpha-1B-glycoprotein	7.5
189163528	Alpha-1-antitrypsin	3.6
29550838	Golgi membrane protein 1 (GP 73)	3.5
154146262	IgGfC-binding protein	3.1
223718260	ADAMTS-like protein 2	3.0
4502505	C4b-binding protein beta	3.0
4502337	Zinc-alpha-2-glycoprotein	2.8
89191868	Von Willebrand factor preproprotein	2.7
21489959	Immunoglobulin J chain	2.6
4557871	Serotransferrin	2.5
50659080	Alpha-1-antichymotrypsin	2.5
167857790	Alpha-1-acid glycoprotein 1	2.3
156523970	Alpha-2-HS-glycoprotein preproprotein	2.1
56237029	Integrin alpha-5	2.0
189458817	Transferrin receptor protein 1	2.0
7705308	Growth/differentiation factor 2	2.0
262206315	L-selectin	2.0
110735433	CD276 antigen isoform b	特异表达
148762980	Sialic acid-binding Ig-like lectin 14	特异表达
187960042	E-selectin	特异表达
301172750	Mucin-5B	特异表达
4501989	Alpha-fetoprotein (AFP)	特异表达

肝癌病人血浆/正常人血浆

小结

- 整合了糖肽富集、糖肽步进能量碎裂和“糖谱筛选”为核心的糖肽谱图优化处理方法，发展了一种新的核心岩藻糖肽大规模精确鉴定策略。

- 解决了先前方法中存在的富集效率、鉴定通量、分析周期等层次的问题。为CF定性与定量蛋白质组研究奠定基础。

总结

- 位点特异性糖型参与并调控蛋白质功能
- 完整糖肽分离技术、质谱分析技术、信息解析技术的发展，仍任重道远

致谢

Also groups :

Pro. Xiaohong Qian



Pro. Simin He



Pro. Yan Fu



NCSF, MOST

