

Cancer Biology

In-depth *N*-glycome profiling of paired colorectal cancer and non-tumorigenic tissues reveals cancer-, stage- and EGFR-specific protein *N*-glycosylation

Manveen K Sethi², Hoguen Kim⁴, Cheol Keun Park⁴, Mark S Baker³, Young-Ki Paik⁵, Nicolle H Packer², William S Hancock^{3,6}, Susan Fanayan^{3,†}, and Morten Thaysen-Andersen^{2,†,1}

²Department of Chemistry and Biomolecular Sciences, ³Department of Biomedical Sciences, Macquarie University, North Ryde NSW 2109, Australia, ⁴Department of Pathology, Yonsei University College of Medicine, Seoul 120-752, Korea, ⁵Yonsei Proteome Research Center, Yonsei University, Seoul 120-749, Korea, and ⁶Barnett Institute and Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA 02115, USA

[†]To whom correspondence should be addressed: Tel: +61-2-9850-7487; e-mail: morten.andersen@mq.edu.au

[†]Co-senior authors.

Received 4 May 2015; Revised 3 June 2015; Accepted 11 June 2015

Abstract

Glycomics may assist in uncovering the structure–function relationships of protein glycosylation and identify glycoprotein markers in colorectal cancer (CRC) research. Herein, we performed label-free quantitative glycomics on a carbon-liquid chromatography–tandem mass spectrometry-based analytical platform to accurately profile the *N*-glycosylation changes associated with CRC malignancy. *N*-Glycome profiling was performed on isolated membrane proteomes of paired tumorigenic and adjacent non-tumorigenic colon tissues from a cohort of five males (62.6 ± 13.1 y.o.) suffering from colorectal adenocarcinoma. The CRC tissues were typed according to their epidermal growth factor receptor (EGFR) status by western blotting and immunohistochemistry. Detailed *N*-glycan characterization and relative quantitation identified an extensive structural heterogeneity with a total of 91 *N*-glycans. CRC-specific *N*-glycosylation phenotypes were observed including an overrepresentation of high mannose, hybrid and paucimannosidic type *N*-glycans and an under-representation of complex *N*-glycans ($P < 0.05$). Sialylation, in particular $\alpha 2,6$ -sialylation, was significantly higher in CRC tumors relative to non-tumorigenic tissues, whereas $\alpha 2,3$ -sialylation was down-regulated ($P < 0.05$). CRC stage-specific *N*-glycosylation was detected by high $\alpha 2,3$ -sialylation and low bisecting $\beta 1,4$ -GlcNAcylation and Lewis-type fucosylation in mid-late relative to early stage CRC. Interestingly, a novel link between the EGFR status and the *N*-glycosylation was identified using hierarchical clustering of the *N*-glycome profiles. EGFR-specific *N*-glycan signatures included high bisecting $\beta 1,4$ -GlcNAcylation and low $\alpha 2,3$ -sialylation (both $P < 0.05$) relative to EGFR-negative CRC tissues. This is the first study to correlate CRC stage and EGFR status with specific *N*-glycan features, thus advancing our understanding of the mechanisms causing the biomolecular deregulation associated with CRC.

Key words: Colorectal cancer, EGFR, glycome, glycomics, *N*-glycosylation

Introduction

Colorectal cancer (CRC), which is the leading cause of cancer-related deaths in developed countries, follows a gradual progression from a small polyp through early stage cancer to late and metastatic stages where the cancer has spread to various body parts (Jemal et al. 2011). Unfortunately, most CRC cases are detected at a late or metastatic stage when the 5-year survival is <10% (Etzioni et al. 2003). There is a lack of robust, accurate and non-invasive diagnostic tools for the early detection of CRC. At present, the screening methods used in the clinic, e.g., barium enema, colonoscopy, sigmoidoscopy and fecal occult blood testing have either low sensitivity/specificity and/or are costly, invasive and/or uncomfortable with a low compliance (Davies et al. 2005; Terdiman 2005). Biomarkers from biological specimens (e.g., blood, bodily fluids or tissues) can provide a more sensitive, specific and less invasive alternative for early CRC detection (Wulfkühle et al. 2003). Currently, carcinoembryonic antigen is used to monitor CRC recurrence, but has, similar to other CRC candidate biomarkers (e.g., carbohydrate antigen 19-9) poor sensitivity (~30–40%) and specificity (85–90%) for the early detection of the disease (Duffy et al. 2003; Tanaka et al. 2010). Thus, there is still an unmet need for the development of informative and accurate biomarkers for the early detection and surveillance of the disease. Importantly, enhanced insights into critical molecular mechanisms driving the onset of CRC may facilitate such development of more targeted approaches.

Modern proteomics and genomics investigations have identified several candidate CRC markers including ezrin, epidermal growth factor receptor (EGFR), spectrins, junction plakoglobin and cadherin 17 (Wang et al. 2009; Luque-Garcia et al. 2010; Fanayan et al. 2013; de Wit et al. 2014; Gan et al. 2014). Driven by technology advancements in the field of analytical glycosciences (Jensen et al. 2012; Parker et al. 2013; Adamczyk et al. 2014; Frost and Li 2014; Huffman et al. 2014; Thaysen-Andersen and Packer 2014; Zhou et al. 2015), glycomics and glycoproteomics are also gaining momentum in cancer research and hold considerable promise to deliver new candidate glycan markers for disease. Protein glycans play crucial roles in various biological processes including cell adhesion, proliferation, cellular signaling and immune response processes (Christiansen et al. 2014). It is thus not surprising that changes in protein glycosylation are recognized as a cause and/or result of malignant processes including tumor development, metastasis and invasion (Tanahashi et al. 1990; Chen et al. 2007). High levels of *N*-glycan sialylation, sulfation and paucimannosylation, and a concomitant decrease in bisecting β 1,4-*N*-acetylglucosamine (GlcNAc) containing *N*-glycans were previously reported in CRC tumor tissues relative to non-tumorigenic tissues (Balog et al. 2012). Alterations in the *N*-glycan types (e.g., high mannosylation and paucimannosylation) and specific substructural features (e.g., bisecting GlcNAcylation, bi-/tri-antennary branching, fucosylation and sialylation) have also been observed in other cancers (Chrostek and Cylwik 2011; de Leoz et al. 2011; Liu, Nie, et al. 2013; Lee, Thaysen-Andersen et al. 2014; Zhang et al. 2014). However, still relative little is known of the exact structural regulation of the glycoepitopes and their protein carriers in CRC and the functional consequences for the pathogenesis.

CRC represents a multi-gene disease affecting the expression of many glycoproteins amongst other biomolecules. EGFR is one such glycoprotein, which plays a critical role in the majority of human cancers (Yarden 2001). EGFR is a highly *N*-glycosylated cell-surface glycoprotein (Carpenter and Cohen 1990) and a member of the ErbB family of receptor tyrosine kinases (Schlessinger 2002). Aberrant *EGFR* gene expression has been observed in various cancer types

(Nicholson et al. 2001; Ellina et al. 2014) correlating well with increased cell growth, proliferation, invasion, differentiation and metastasis (Mendelsohn and Baselga 2006; Ellina et al. 2014). Elevated levels of EGFR protein in tumors have also been associated with reduced overall survival in colon and breast cancer patients (Harris et al. 1992; Baas et al. 2011). Previous studies have shown that EGFR signaling can be modulated by *N*-glycosylation changes such as sialylation, fucosylation and bisecting GlcNAcylation (Sato et al. 2001; Liu et al. 2011). Thus, studying the regulation of EGFR and how it affects or is being affected by protein *N*-glycosylation may provide novel insights into the molecular mechanisms associated with CRC.

Recently, we performed a comprehensive analysis of the membrane protein *N*-glycome in a panel of CRC cell lines representing pathologically different types of the disease (Sethi et al. 2014). *N*-glycosylation signatures common to all the CRC cell lines included the dominance of high mannosylation and α 2,6-sialylation. In addition, other CRC type-specific *N*-glycosylation features were observed, which correlated with the pathological features of the individual cell lines. Finally, correlation analysis resulted in the hypothesis that the *N*-glycosylation signatures of the CRC cells in culture were associated with the EGFR expression status. In support of this association, quantitative proteomics of paired CRC and non-tumorigenic tissues from the same cohort of CRC patients, as are investigated in this study, identified several CRC-, stage- and EGFR-specific proteins and deregulated pathways, which provided a possible mechanistic link to the cancer-induced molecular perturbation (Sethi et al. 2015). Importantly, this and other studies suggested that the altered glycoproteins observed in CRC may, in part, originate from glycosylated proteins emanating from other sources than the CRC cells, e.g., liver or inflammatory cells. Taken together, this indicates that multiple separate mechanisms may cause the aberrant glycosylation signatures associated with the tumor microenvironment.

Herein, quantitative glycomics was utilized to test the association between protein glycosylation and CRC by accurately profiling the *N*-glycomes of paired tumor and non-tumorigenic tissue samples from five male patients suffering from adenocarcinoma CRC. Several CRC-, stage- and indeed EGFR-dependent *N*-glycosylation signatures were identified, which offer important new insights into the molecular mechanisms of the disease.

Results

N-glycome profiling of CRC tumor and adjacent non-tumorigenic colon tissue

Paired primary tumor (T1–T5) and adjacent non-tumorigenic tissue (N1–N5) samples were obtained from five CRC patients (P1–P5) undergoing surgery for adenocarcinomas at the Severance Hospital, Seoul, Korea. Patient-specific information and details of the investigated tissues are summarized in Table I. *N*-Glycans from membrane-enriched proteins extracted from these tissues were enzymatically released to completion, purified and analyzed in their reduced but otherwise underivatized form using porous graphitized carbon (PGC)-LC–electrospray ionization (ESI)-negative ion-MS/MS-based quantitative glycomics. In total, 61 monosaccharide compositions were confidently identified from the 10 CRC tumors and non-tumor tissues. This set of monosaccharide compositions corresponded to a total of 91 individual glycan structural isomers, Table II and Supplementary data, Table S1. The *N*-glycomes included the presence of the conventional human *N*-glycan types, i.e., high mannose, complex and hybrid

Table I. Patient-specific characteristics of the CRC patients and their paired (T-N) tissues used in this study

Patient (P) (paired tumor (T) and normal (N) tissue)	Pathological diagnosis	Type	Age ^a (y.o.)	Gender	Location	Differentiation status	EGFR expression (+/-) ^c	Distance of tumor (T) to normal (N) tissue (cm)	Tumor stage ^b	Patient status
P1 (T1 and N1)	Adeno-carcinoma	Polypoid type	64	Male	Sigmoid	Well	+	5	I	Alive
P2 (T2 and N2)	Adeno-carcinoma	Ulcerative	69	Male	Rectum	Moderate	-	22	IIIB	Alive
P3 (T3 and N3)	Adeno-carcinoma	Ulcer-fungating	67	Male	Cecum	Moderate	-	15	IIIB	Alive
P4 (T4 and N4)	Adeno-carcinoma	Annular constrictive	40	Male	Rectum	Well	+	5	IV	Deceased
P5 (T5 and N5)	Adeno-carcinoma	Ulcer-fungating	73	Male	Transverse	Poor	-	25	IV	Alive

^aAt time of diagnosis.

^bThe clinical staging was based on the tumor, node and metastasis (TNM) staging system.

^c+ indicates presence and - indicates absence of EGFR based on western blotting and IHC.

type structures as well as the non-conventional paucimannosidic class, which we recently validated, were carried by intact proteins produced by human neutrophils (Thaysen-Andersen et al. 2015). The complex and hybrid type *N*-glycans carried several substructural glyco-determinants including β 1,4-(bisecting)GlcNAcylation, α 2,3-/ α 2,6-sialylation, *N*-acetylglucosamine (LacNAc) repeats and α 1,3/4-(Lewis)- and α 1,6-(core)-type fucosylation. The *N*-glycan structures were characterized using accurate monoisotopic precursor mass, PGC LC relative and absolute retention time patterns and collision-induced dissociation (CID)-based MS/MS fragmentation, Supplementary data, Figure S1. The mass profiles of the five CRC tumor tissues and their corresponding paired non-tumor tissues appeared qualitatively similar at a glance, Supplementary data, Figure S2. The majority of the identified *N*-glycans (74 structures, 81.3%) were indeed common between the two tissue types, Figure 1. However, 15 *N*-glycans (16.5%) were observed to be CRC tissue specific, encompassing mostly hybrid, bisecting GlcNAcylation complex and paucimannosidic *N*-glycan structures. Only two complex type structures (2.2%) were uniquely observed in the non-tumor tissues.

Quantitative *N*-glycoprofiling reveals significant CRC-specific *N*-glycosylation

Deeper analysis of the CRC-specific *N*-glycosylation was performed by quantitatively comparing the glycoprofiles of the CRC (T1–T5) and non-tumor (N1–N5) tissues according to their relative distribution of *N*-glycan types (i.e., high mannose, hybrid, complex and paucimannosidic) and substructural determinants (i.e., sialylation), Figure 2. Complex and high mannose type *N*-glycans were abundant structures in both CRC (51.6% \pm 7.8% and 39.7% \pm 7.4%, respectively) and non-tumor (65.9% \pm 6.8% and 29.7% \pm 7.2%, respectively) tissues. Paucimannosidic and hybrid type *N*-glycans were of relative low abundance in CRC (5.8% \pm 2.4% and 2.9% \pm 1.0%, respectively) and non-tumor (3.1% \pm 0.7% and 1.3% \pm 0.3%, respectively) tissues. Upon statistical evaluation of these distributions, it became clear that the percentage of high mannose, paucimannosidic and hybrid type *N*-glycans were significantly overrepresented in CRC tissues (all $P < 0.05$, $n = 5$, paired), Figure 2A. Conversely, complex type *N*-glycans were significantly under-represented in the CRC tissues ($P = 0.01$, $n = 5$, paired) compared with the matched non-tumor tissues.

The degree of sialylation was also altered in the CRC tissues, Figure 2B. The relative level of total sialylation (not accounting for the sialyl linkage type) of complex and hybrid type *N*-glycans was significantly up-regulated in CRC tissues (80.1% \pm 9.1%) compared with non-tumor tissues (73.3% \pm 7.7%, $P = 0.01$, $n = 5$, paired);

however, since the level of the sialic acid-containing complex type *N*-glycans were simultaneously found to be lower in tumor tissues, little or negligible oversialylation per cell may in fact be displayed by CRC tissues. Sialyl linkage analysis demonstrated that the relative oversialylation in CRC tissues arose from a higher α 2,6-sialylation of the CRC tissues (55.2% \pm 15.1%) compared with non-tumor tissues (37.1% \pm 7.5%, $P = 0.003$, $n = 5$, paired), Figure 2C. A concomitant decrease in the α 2,3-sialylation in CRC (26.3% \pm 16.3%) relative to the non-tumor (37.4% \pm 6.1%, $P = 0.03$, $n = 5$, paired) tissues was observed. The relative level of total fucosylation and the fucosyl linkage type-specific distribution on complex, hybrid and paucimannosidic type *N*-glycans was statistically similar between the CRC and the non-tumor tissues (all $P > 0.05$), Supplementary data, Figure S3A and B. Similarly, there was no detectable CRC-specific regulation of the bisecting GlcNAcylation and LacNAcylation showing similar relative levels in all tissues (both $P > 0.05$), Supplementary data, Figure S3C and D. Hierarchical clustering of the glycoprofiles accurately separated the EGFR⁺ tumors (T1 and T4) from the EGFR⁻ tumors (T2, T3 and T5) and crudely clustered the CRC stages of the patients, Figure 3. This suggested not only EGFR dependency of the CRC tissue *N*-glycosylation but also some CRC stage-dependency. These two potential associations between the molecular profile and the cancer phenotype were then investigated further.

CRC stage-specific *N*-glycosylation signatures

CRC stage-dependent *N*-glycosylation was investigated by grouping the CRC tissues into early (T1, stage I), mid (T2, T3, stage IIIB) and late (T4, T5, stage IV) disease stages. Several individual glycan structures were identified to be stage specific, Supplementary data, Figure S4. In total, seven *N*-glycans, mainly bisecting GlcNAc-containing structures, were uniquely identified in the early CRC stage tissue samples. In contrast, 24 *N*-glycans consisting of paucimannosidic and various complex type *N*-glycans containing LacNAc and LacdiNAc repeats and α 2,3-sialylation were observed exclusively in the mid-late CRC stage. Quantitative glycoprofile comparison revealed that although no CRC stage-specific regulation of the distribution of *N*-glycan types was observed (data not shown), several substructural glycosylation features were altered with the disease progression, Table III. High levels of total sialylation for the mid-late stage (80.2–87.6%) relative to the early stage levels (64.8%) were observed, which was driven by a significant (3- to 10-fold) increase specifically in the α 2,3-sialylation on the *N*-glycans in the mid-late CRC stages relative to early stage. High bisecting GlcNAcylation (38.2%) and Lewis-type fucosylation (30.2%) relative to mid-late stage tumors

Table II. Overview of the 61 N-glycan monosaccharide compositions and 91 individual glycan isomers identified in the ten CRC tissue samples

Glycan no.	Theo. <i>m/z</i>	Obs. <i>m/z</i>	Charge (Z)	Monosaccharide composition				Isoform	Relative abundance (mean ± SD, %)	
				Hex	HexNAc	Fuc	NeuAc		CRC tumor (T1–T5)	Non-tumor (N1–N5)
1	571.2	571.2	−1		2	1		a	*	ND
2	587.2	587.2	−1	1	2			a	*	ND
3	698.2	698.3	−2	6	2			a	8.0 ± 1.8	5.2 ± 1.1
4	731.3	731.3	−2	3	4	1		a	0.3 ± 0.1	0.1 ± 0.2
5	733.2	733.3	−1	1	2	1		a	0.3 ± 0.7	ND
6	749.2	749.3	−1	2	2			a	0.4 ± 0.2	0.4 ± 0.4
7	759.8	759.8	−2	3	5			a	0.1 ± 0.1	0.2 ± 0.0
8a	779.3	779.3	−2	7	2			a	4.0 ± 1.8	3.4 ± 1.4
8b	779.3	779.3	−2	7	2			b	1.6 ± 0.5	0.6 ± 0.4
9	783.3	783.3	−2	4	3		1	a	0.2 ± 0.1	ND
10	799.8	799.8	−2	6	3			a	0.3 ± 0.2	0.2 ± 0.2
11a	812.3	812.3	−2	4	4	1		a	0.2 ± 0.3	0.1 ± 0.2
11b	812.3	812.3	−2	4	4	1		b	0.3 ± 0.4	0.6 ± 0.5
12	820.3	820.3	−2	5	4			a	0.4 ± 0.3	0.4 ± 0.2
13	832.8	832.8	−2	3	5	1		a	0.5 ± 0.7	1.2 ± 0.7
14	840.8	840.8	−2	4	5			a	0.1 ± 0.2	*
15	856.3	856.3	−2	4	3	1	1	a	0.5 ± 0.6	0.5 ± 0.3
16a	860.3	860.3	−2	8	2			a	7.0 ± 2.1	5.5 ± 1.1
16b	860.3	860.3	−2	8	2			b	0.4 ± 0.5	ND
17a	864.3	864.4	−2	5	3		1	a	0.4 ± 0.3	0.1 ± 0.1
17b	864.3	864.3	−2	5	3		1	b	0.1 ± 0.1	0.1 ± 0.1
18	872.8	872.9	−2	6	3	1		a	0.2 ± 0.3	0.0 ± 0.1
19	884.8	884.9	−2	4	4		1	a	0.1 ± 0.3	0.0 ± 0.0
20a	893.3	893.3	−2	5	4	1		a	0.1 ± 0.2	*
20b	893.3	893.3	−2	5	4	1		b	2.5 ± 2.0	4.3 ± 1.0
21	895.4	895.4	−1	2	2	1		a	1.9 ± 1.6	0.7 ± 0.3
22	911.4	911.3	−1	3	2			a	1.1 ± 0.2	0.8 ± 0.2
23	913.9	913.3	−2	4	5	1		a	0.5 ± 0.1	1.1 ± 0.5
24	921.8	921.9	−2	5	5			a	0.1 ± 0.1	0.1 ± 0.1
25	934.3	934.4	−2	3	6	1		a	0.2 ± 0.2	*
26a	937.3	936.9	−2	5	3	1	1	a	0.2 ± 0.2	0.1 ± 0.1
26b	937.3	937.4	−2	5	3	1	1	b	0.1 ± 0.2	0.1 ± 0.0
27	941.3	941.4	−2	9	2			a	9.6 ± 2.5	8.5 ± 3.0
28a	945.3	945.4	2	6	3		1	a	0.8 ± 0.7	0.4 ± 0.2
28b	945.3	945.4	−2	6	3		1	b	0.2 ± 0.2	0.2 ± 0.1
29a	957.8	957.9	−2	4	4	1	1	a	0.3 ± 0.1	0.1 ± 0.1
29b	957.8	957.9	−1	4	4	1	1	b	0.3 ± 0.4	0.2 ± 0.1
30a	965.8	965.9	−1	5	4		1	a	2.6 ± 0.6	2.5 ± 1.3
30b	965.8	965.9	−1	5	4		1	b	0.8 ± 0.5	0.8 ± 0.6
31a	974.4	974.4	−2	6	4	1		a	0.1 ± 0.3	ND
31b	974.4	974.4	−2	6	4	1		b	*	ND
31c	974.4	974.4	−2	6	4	1		c	0.1 ± 0.3	ND
32a	986.9	986.9	−2	4	5	2		a	0.1 ± 0.1	0.1 ± 0.2
32b	986.9	986.9	−2	4	5	2		b	*	0.1 ± 0.2
32c	986.9	986.9	−2	4	5	2		c	ND	0.1 ± 0.2
33a	994.9	994.9	−2	5	5	1		a	0.1 ± 0.3	0.3 ± 0.1
33b	994.9	994.9	−2	5	5	1		b	1.1 ± 0.7	1.8 ± 1.2
34	1018.4	1018.5	−2	6	3	1	1	a	0.1 ± 0.2	ND
35	1022.3	1022.4	−2	10	2			a	0.6 ± 0.3	0.4 ± 0.3
36a	1038.9	1038.9	−2	5	4	1	1	a	0.5 ± 0.4	*
36b	1038.9	1038.9	−2	5	4	1	1	b	3.0 ± 0.4	2.3 ± 0.6
36c	1038.9	1038.9	−2	5	4	1	1	c	4.4 ± 3.4	9.8 ± 1.3
37	1057.4	1057.5	−1	3	2	1		a	2.1 ± 0.3	1.3 ± 0.4
38	1059.4	1059.4	−2	4	5	1	1	a	0.8 ± 0.5	0.4 ± 0.2
39a	1067.9	1067.9	−2	5	5	2		a	0.1 ± 0.3	0.1 ± 0.1
39b	1067.9	1067.9	−2	5	5	2		b	0.6 ± 1.2	2.5 ± 1.9
40a	1073.4	1073.4	−1	4	2			a	0.1 ± 0.1	0.1 ± 0.2
40b	1073.4	1073.4	−1	4	2			b	0.5 ± 0.4	0.4 ± 0.1
40c	1073.4	1073.5	−1	4	2			c	0.5 ± 0.3	0.4 ± 0.1
41a	1075.9	1075.9	−2	6	5	1		a	0.8 ± 0.4	0.8 ± 0.1

Continued

Table II. Continued

Glycan no.	Theo. <i>m/z</i>	Obs. <i>m/z</i>	Charge (Z)	Monosaccharide composition				Isoform	Relative abundance (mean ± SD, %)	
				Hex	HexNAc	Fuc	NeuAc		CRC tumor (T1–T5)	Non-tumor (N1–N5)
41b	1075.9	1075.9	–2	6	5	1		b	0.2 ± 0.3	0.2 ± 0.3
42	1080.4	1080.4	–2	3	4	3		a	0.4 ± 0.6	1.3 ± 0.5
43a	1111.4	1111.4	–2	5	4		2	a	4.8 ± 1.3	5.3 ± 3.7
43b	1111.4	1111.4	–2	5	4		2	b	0.9 ± 1.1	0.9 ± 0.6
44a	1111.9	1111.9	–2	5	4	2	1	a	4.4 ± 0.9	3.8 ± 2.3
44b	1111.9	1111.9	–2	5	4	2	1	b	0.6 ± 0.3	0.4 ± 0.2
45	1132.4	1132.4	–2	4	5	2	1	a	0.2 ± 0.2	ND
46	1140.4	1140.4	–2	5	5	1	1	a	0.9 ± 0.4	1.3 ± 0.8
47	1140.9	1141.0	–2	5	5	3		a	0.5 ± 0.8	1.4 ± 0.7
48	1148.4	1148.4	–2	6	5		1	a	*	0.3 ± 0.3
49	1160.9	1160.9	–2	4	6	1	1	a	0.1 ± 0.2	ND
50	1177.4	1177.4	–2	6	6	1		a	ND	0.2 ± 0.4
51a	1184.4	1184.5	–2	5	4	1	2	a	2.7 ± 2.1	1.3 ± 0.4
51b	1184.4	1184.5	–2	5	4	1	2	b	1.7 ± 1.1	1.2 ± 0.2
51c	1184.4	1184.5	–2	5	4	1	2	c	5.8 ± 5.1	11.1 ± 4.6
52	1213.5	1213.5	–2	5	5	2	1	a	0.8 ± 0.7	1.0 ± 0.7
53a	1221.4	1221.5	–2	6	5	1	1	a	0.2 ± 0.2	0.1 ± 0.1
53b	1221.4	1221.5	–2	6	5	1	1	b	0.9 ± 0.5	1.2 ± 0.8
53c	1221.4	1221.5	–2	6	5	1	1	c	1.0 ± 0.8	1.1 ± 0.4
53d	1221.4	1221.4	–2	6	5	1	1	d	0.7 ± 0.4	0.4 ± 0.2
54a	1235.4	1235.5	–1	5	2			a	0.4 ± 0.4	0.2 ± 0.2
54b	1235.4	1235.4	–1	5	2			b	0.6 ± 0.1	0.3 ± 0.3
54c	1235.4	1235.4	–1	5	2			c	6.5 ± 4.5	4.7 ± 2.0
55	1258.5	1258.5	–2	7	6	1		a	0.4 ± 0.3	0.4 ± 0.4
56a	1260.5	1260.6	–1	3	3	1		a	0.2 ± 0.2	*
56b	1260.5	1260.5	–1	3	3	1		b	0.1 ± 0.1	*
57	1285.9	1285.6	–2	5	5	1	2	a	0.7 ± 0.7	0.9 ± 0.3
58	1294.5	1294.5	–2	6	5	2	1	a	0.1 ± 0.1	ND
59	1315.0	1315.1	–2	5	6	2	1	a	0.4 ± 0.5	0.1 ± 0.2
60	1367.0	1367.0	–2	6	5	1	2	a	1.3 ± 0.7	1.3 ± 0.8
61	1367.5	1367.5	–2	6	5	2	1	a	0.1 ± 0.1	ND

The relative abundances (mean ± SD) of the observed tumor (T1–T5) and non-tumor (N1–N5) N-glycan isomers are summarized. *, trace amounts detected (<0.1% relative abundance), which were not included in the quantitation; ND, not detected. See Supplementary data, Table S1 for more structural information.

(6.7–11.0% for bisecting GlcNAc and 11.8–13.2% for Lewis-type fucosylation) were important early stage CRC features. Alterations in the same N-glycosylation features were seen, but were less pronounced, in the matching non-tumorigenic tissue.

EGFR-specific N-glycosylation of CRC tissues

The potential influence of EGFR on the CRC tissue N-glycosylation was investigated by grouping the CRC tissues based on their EGFR expression as determined by immunohistochemical (IHC) and western blotting; two CRC tissues (T1 and T4) were EGFR⁺ while three CRC tissues (T2, T3 and T5) exhibited very low or negligible EGFR expression, Table IV. Surprisingly, the adjacent non-tumor tissues did not display the same EGFR expression patterns as the corresponding tumor tissues (i.e., EGFR⁺: N1, N3 and N5; EGFR[–]: N2 and N4), Supplementary data, Figure S5A. While the majority of EGFR⁺ and EGFR[–] tumor N-glycans were similar, some unique structures were present in these two molecular subtypes of CRC, Supplementary data, Figure S6. In total, 12 N-glycans, mainly bisecting GlcNAc-containing complex and hybrid type N-glycans, were detected exclusively in the EGFR⁺ CRC tissues. In contrast, 16 N-glycans encompassing paucimannosidic and α2,3-sialylated complex type N-glycans were unique to the EGFR[–] tumors. N-Glycome comparisons showed an overrepresentation of high mannose (47.6% ± 2.1%) and hybrid type N-glycans (3.8% ± 0.2%) in

EGFR⁺ relative to EGFR[–] CRC tissues (high mannose: 34.5% ± 1.9%, $P = 0.02$; hybrid: 2.2% ± 0.7%, $P = 0.04$), Table V. In contrast, complex type N-glycans were significantly increased ($P = 0.02$), whereas paucimannosidic N-glycans were unchanged ($P = 0.50$) in EGFR[–] tumors relative to EGFR⁺ tumors.

At the substructural glycan level, bisecting GlcNAcylation appeared to be associated with EGFR expression as evidenced by the significantly higher level of bisecting GlcNAc-containing N-glycans in EGFR⁺ tumors (26.6% ± 11.7%) relative to EGFR[–] tumors (6.8% ± 0.9%, $P = 0.03$), Table V. Although no EGFR-dependent alteration in the total sialylation was detected, significantly lower levels of α2,3-sialylation were observed for the EGFR⁺ tumor (9.3% ± 3.9%) when compared with EGFR[–] tumor (37.7% ± 9.1%, $P = 0.02$) tissues. This was accompanied by a non-significant ($P = 0.22$) increase in the α2,6-sialylation for EGFR⁺ tumors (66.9% ± 12.8%) relative to EGFR[–] tumors (47.3% ± 13.7%). The EGFR-dependent sialyl linkage specificity was illustrated by the extracted ion chromatograms (EICs) of two prominent mono- and di-sialylated biantennary complex type N-glycans (*m/z* 1038.9 and 1184.4), where altered ratios of α2,3 : α2,6-sialylation were clearly observed for EGFR⁺ relative to EGFR[–] CRC tissues, Figure 4. No EGFR-dependent regulation was observed for the total fucosylation and LacNAcylation. Interestingly, no significant EGFR-dependent modulation of the protein N-glycosylation was observed in the non-tumorigenic tissues, Supplementary data, Figure S5B, implying that

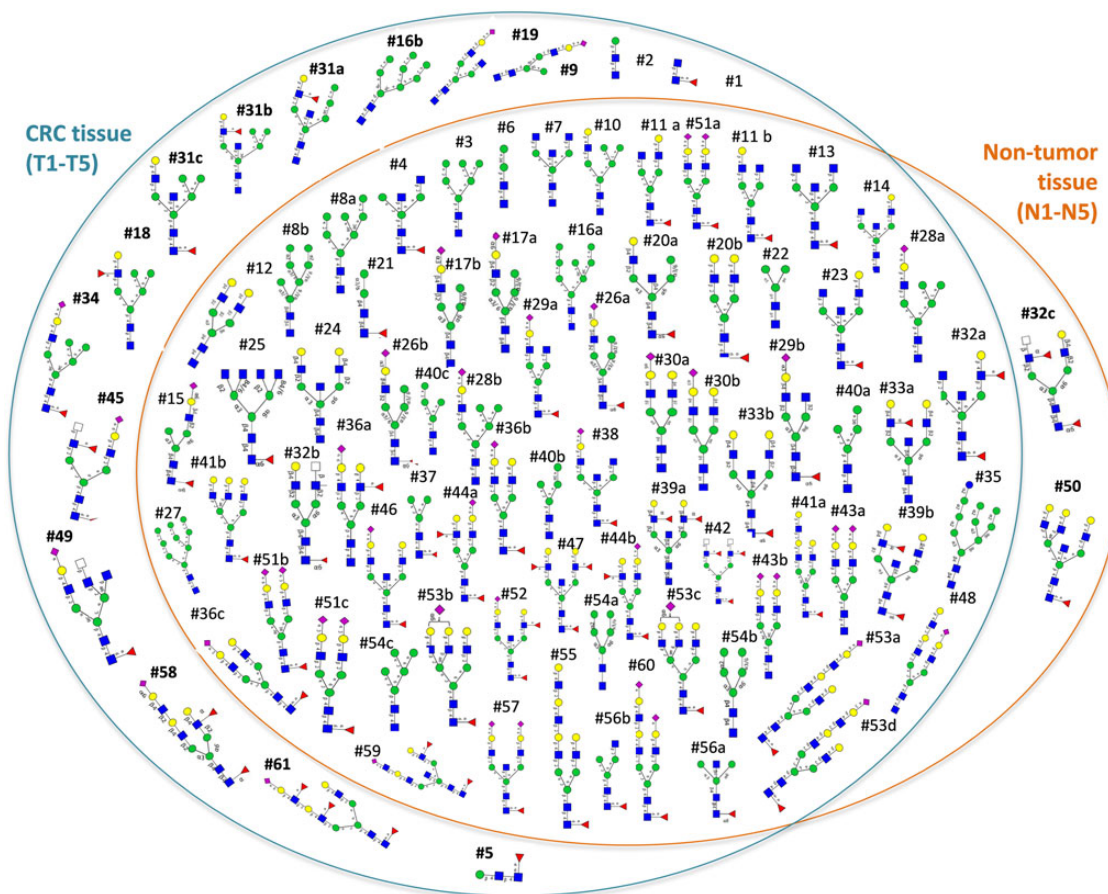


Fig. 1. CRC tissue-specific N-glycosylation signatures as shown by the presence of 15 unique N-glycans in the CRC tissues (T1–T5), which did not appear in the adjacent non-tumor tissues (N1–N5). Monosaccharide symbols according to the Consortium for Functional Glycomics (CFG) nomenclature. Key: fucose (red triangle), mannose (green circle), GlcNAc (blue square), sialic acid (purple diamond), galactose (yellow circle) and HexNAc (GlcNAc or GalNAc) (open square). The N-glycans are accompanied by their designated numbers as listed in Table II and Supplementary data, Table S1 and on the MS profiles spectra in Supplementary data, Figure S2. This figure is available in black and white in print and in color at *Glycobiology* online.

the potential EGFR-based regulation of the N-glycosylation may be dependent on other malignancy related cellular processes or factors specific to the CRC tissues. The CRC disease-, stage- and EGFR-specific N-glycan alterations are summarized in Table VI. It is important to stress that these interesting molecular trends associating the CRC stage development and the EGFR expression with the N-glycosylation phenotype need further validation using a larger cohort of CRC patients covering the individual disease stages and variations in EGFR expression.

Investigating the causes of aberrant N-glycomes in CRC tissues

The causes responsible for the aberrant N-glycomes observed in CRC tissues were then assessed in order to more mechanistically understand the molecular regulation associated with the cancer pathogenesis, Figure 5A. The significant representation of relatively unique neutrophil molecular signatures in the CRC-specific N-glycome [six N-glycans (#1, #5, #9, #16b, #58 and #61) of 15 N-glycan monosaccharide compositions, 40.0%] (Babu et al. 2009; Thaysen-Andersen et al. 2015) and the previously acquired CRC tissue proteome (Sethi et al., 2015) (60 of 174 unique CRC tissue proteins, 34.5%) (Rorvig et al. 2013) indicated a strong contribution of (glyco)proteins from neutrophils, which localize to the inflamed and cancerous CRC tumor microenvironment, Figure 5B. In addition, it appeared that the total N-glycome and

membrane proteome (Sethi et al., 2015) of the CRC tumor tissue in addition to being able to separate the CRC tissues from the non-tumor tissues (data not shown) using hierarchical clustering also separated the five CRC patients into stage types in a similar pattern, Figure 5C. These similarities strongly indicated that alterations in protein expression are significant contributors to the aberrant N-glycome observed in CRC tumor tissues. Finally, the aberrant glyco-features observed in the CRC tumor tissues, which largely recapitulate glyco-epitope alternations in CRC reported elsewhere (de Leoz et al. 2011; Liu, Nie, et al. 2013; Anugraham et al. 2014; Sethi et al. 2014) correlated with the aberrant expression of the responsible glycosylation enzymes as detected by others (Harvey et al. 1992; Pousset et al. 1997; Recchi et al. 1998; Fernandez-Rodriguez et al. 2000; Wang et al. 2003; Park and Lee 2013; Anugraham et al. 2014; Sethi et al. 2014), Figure 5D. Taken together, these lines of evidence suggest that the glycosylation machinery of one or multiple cell types in the tumor microenvironment is significantly altered, which together with the aberrant protein expression and cellular distribution (neutrophil infiltration) will jointly contribute to the CRC-specific glycosylation detected in the tissue N-glycomes.

Discussion

In this study, CRC-, disease stage- and EGFR-specific N-glycosylations were identified by performing detailed N-glycome profiling of

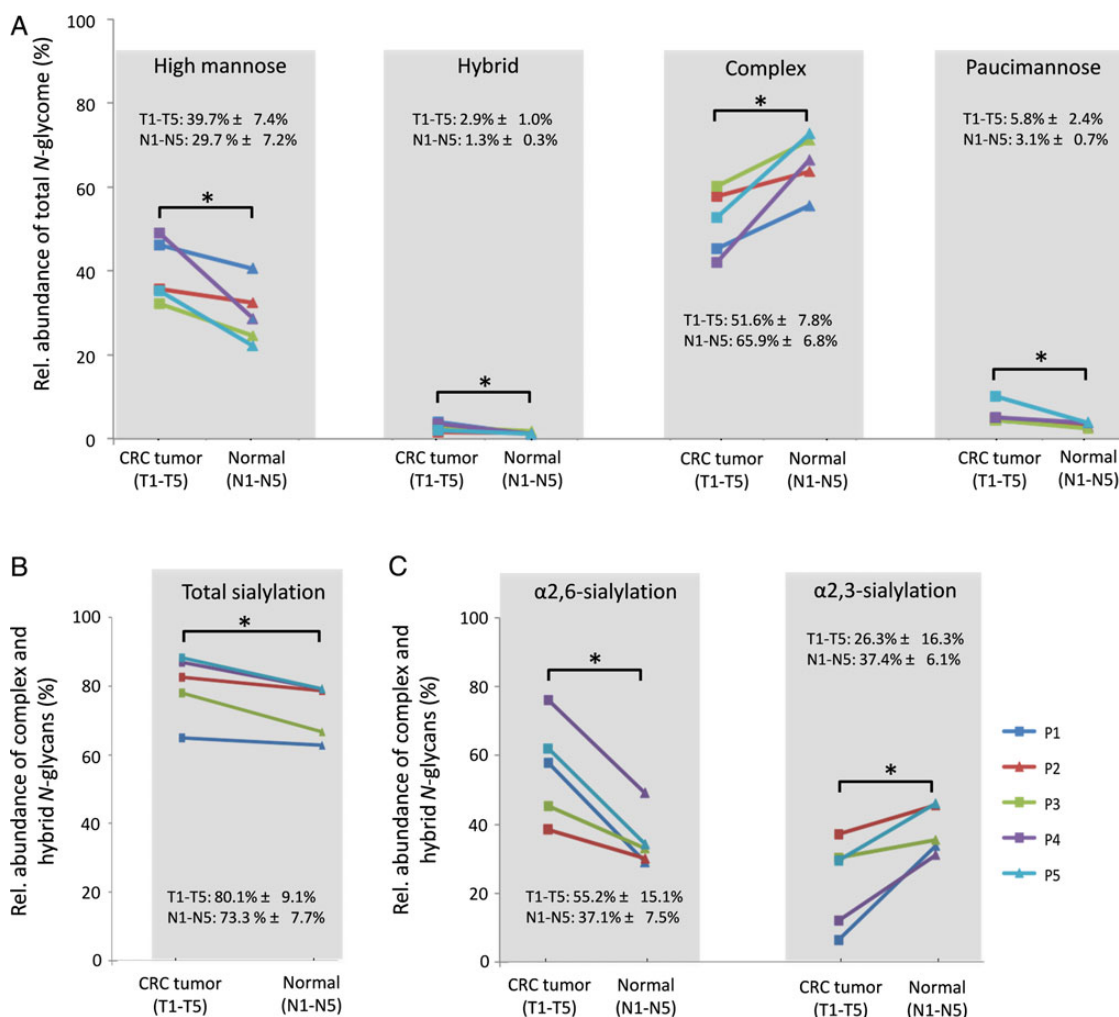


Fig. 2. (A) Relative *N*-glycan type distribution in the profiled *N*-glycomes of CRC (T1–T5) and non-tumor (N1–N5) tissues. The *N*-glycans were classified into the four major human types, i.e., high mannose, hybrid, complex and paucimannosidic *N*-glycans. (B) CRC-specific regulation of the total level of sialylation, irrespective of the sialyl linkage, of the complex and hybrid type *N*-glycans was observed as shown by the difference in sialylation of *N*-glycans derived from tumor (T1–T5) relative to non-tumorigenic (N1–N5) tissues. (C) The linkage-specific (α 2,6- vs α 2,3-) distribution of sialylated *N*-glycans was also unique in tumor (T1–T5) compared with non-tumorigenic (N1–N5) tissues. Statistical analysis was performed for all glycan features from tumor and their corresponding non-tumorigenic tissues to test for CRC-specific differences. Significance was evaluated based on paired *t*-tests ($n=5$). * $P < 0.05$. T, tumor, N, non-tumor. The mean \pm SD is indicated for all measurements. This figure is available in black and white in print and in color at *Glycobiology* online.

five paired CRC tissues and adjacent non-tumorigenic colon tissues. The CRC-specific *N*-glycosylation signature showing elevated levels of high mannose and reduced levels of complex type *N*-glycans is consistent with our previous report of a panel of high mannose-rich CRC cell lines (Sethi et al. 2014). Elevated high mannose type *N*-glycosylation has also been associated with various other cancers including ovarian (Anugraham et al. 2014) and breast cancer (Liu, Nie, et al. 2013), where high levels of the immature $\text{GlcNAc}_2\text{Man}_9$ have been reported (de Leoz et al. 2011). Taken together, these observations suggest that elevated levels of high mannose are associated with malignant and/or inflammatory processes in some manner. The overrepresentation of high mannose *N*-glycans in CRC tissues from either the tumor or resident inflammatory cells may indicate an altered *N*-glycosylation biosynthesis in the endoplasmic reticulum towards less *N*-glycan processing, resulting in an intracellular accumulation of high mannose type *N*-glycoproteins. High mannose structures, in particular the immature $\text{GlcNAc}_2\text{Man}_9\text{Glc}_{0-1}$ moieties, were recently shown to be highly enriched in ER-rich microsomes of a

panel of breast cancer cells (Lee, Lin et al. 2014), where it was suggested that the incomplete maturation of *N*-glycosylation in the secretory pathway could arise from higher protein secretion (higher trafficking rate) and/or lower abundance/activity of the glycosylation processing enzymes.

The elevation of the unusual paucimannosidic *N*-glycans in CRC tumor tissues is also consistent with a recent study showing an increase in paucimannosylation in CRC tumors relative to non-tumorigenic tissues (Balog et al. 2012). Paucimannosylation of proteins has been considered to be absent in humans (Schachter 2009), but recent studies have indicated the presence of paucimannosidic *N*-glycans in human specimen (Balog et al. 2012; Everest-Dass et al. 2012; Sethi et al. 2014; Dahmen et al. 2015; Venkatakrishnan et al. 2015). The ability of humans to produce and secrete significant levels of paucimannosylation by neutrophils was recently validated, which suggests that paucimannosidic structures arise from extensive *N*-glycan processing, i.e., truncation (Thaysen-Andersen et al. 2015). In this light, the CRC-specific reduction of the level of complex *N*-glycans and the

elevation of paucimannosylation may, in part, arise from a diverted biosynthetic route in the CRC tissues resulting in N-glycan truncation rather than extension of the glycan antennas. However, the resemblance of a substantial proportion of the molecular features observed in the CRC tissues to characteristic glycan (e.g., paucimannose) and protein (e.g., azurocidin and neutrophil elastase, Sethi et al. 2015) features of human neutrophils points towards a significant contribution of glycoproteins secreted from infiltrating neutrophils to the observed CRC tissue glycome profile. Neutrophils are being increasingly recognized as important and abundant elements in the tumor

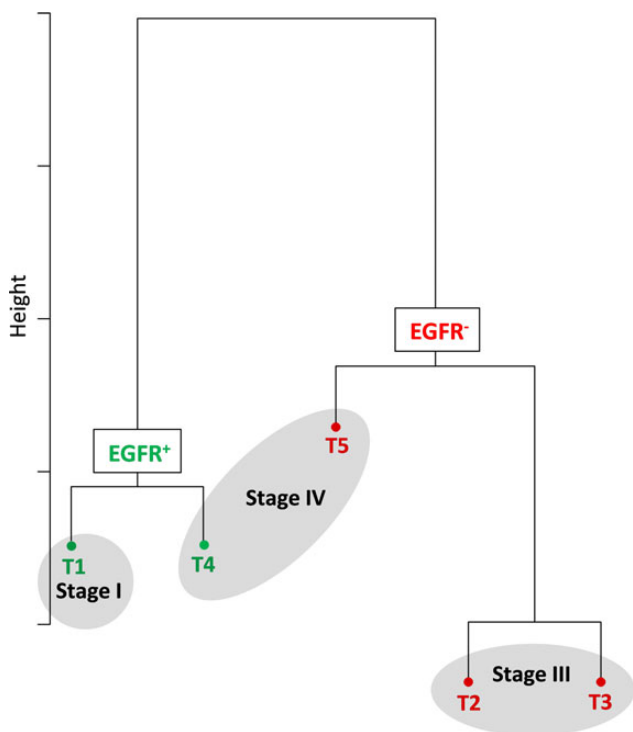


Fig. 3. EGFR- and stage-dependent separation of the CRC tumor tissues (T1–T5) as shown by a hierarchical clustering based on the total membrane protein N-glycomes. This separation suggested that the CRC tissue N-glycosylation profile is associated with the EGFR expression status and the CRC disease stage. This figure is available in black and white in print and in color at *Glycobiology* online.

microenvironment (Cools-Lartigue et al. 2014). Finally, the abilities of the CRC tissue N-glycome and proteome to separate the patients into similar subgroups based on hierarchical clustering indicate that altered protein expression from one or multiple cell types in the tumor microenvironment may also contribute to the aberrant glycosignatures observed in the heterogeneous cellular and molecular make-up of the CRC tissues. The precise contribution of these three potential sources of regulation that create the aberrant protein N-glycosylation in CRC tissues (see Figure 5A for summary) needs to be dissected further in order to understand the functional impact on the disease mechanisms.

Over-sialylation was found to be another CRC-specific N-glycosylation feature, which corroborated recent findings showing elevated sialylation in CRC tissues relative to non-tumorigenic tissues (Balog et al. 2012). Here we show by linkage-specific analysis that elevated α 2,6-sialylation was the major contributor to the oversialylation, while α 2,3-sialylation was significantly reduced in CRC tissues. We derived similar conclusions from our recent profiling of several CRC cell lines, which documented that the α 2,6-sialyl linkage is the major form of sialylation in CRC-derived cells (Sethi et al. 2014). In support of this, transcriptomics by RNA-Seq analysis identified relative high *ST6GAL1* and low *ST3GAL3* gene expression (that encode for the sialyltransferases responsible for α 2,6- and α 2,3-sialylation, respectively) in the same panel of CRC cell lines (Sethi et al. 2014). The sialyl linkage-specific signatures of CRC tissues have also been confirmed by IHC showing weak staining of the α 2,3-sialic acid recognizing *Maackia amurensis* lectin and strong staining of the α 2,6-sialic acid recognizing *Sambucus nigra* lectin (Fernandez-Rodriguez et al. 2000). Similar alterations in the sialylation have been reported in other cancers such as breast, liver and ovarian cancers (Pousset et al. 1997; Recchi et al. 1998; Anugraham et al. 2014) as well as squamous cell carcinoma of the cervix (Wang et al. 2003). Moreover, recent clinical reports have associated the regulation of *ST6GAL1* gene expression with CRC progression, invasion and metastasis (Harvey et al. 1992; Park and Lee 2013). Thus, several lines of evidence associate aberrant *ST6GAL1* expression and elevated levels of α 2,6-sialylation with the functional involvement of these terminal glyco-determinants in malignant processes in CRC. This glycosylation signature may as a consequence act as a potential marker and/or as a target for therapeutic intervention against CRC.

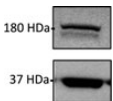
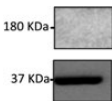
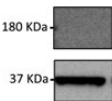
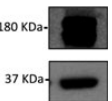
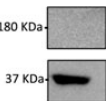
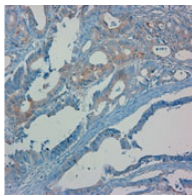
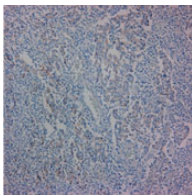
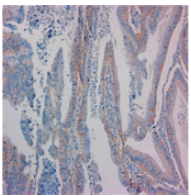
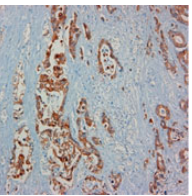
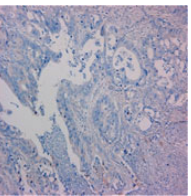
Our data indicated no regulation of N-glycoprotein fucosylation (core and Lewis type) in CRC tissues, which is in good agreement with a previous study reporting no differential IHC staining

Table III. CRC stage-specific distribution of N-glycan substructural features of CRC (T1–T5) and non-tumor (N1–N5) tissues as divided into their respective stages, i.e., early stage I (T1 and N1), mid stage III (T2–T3 and N2–N3) and late stage IV (T4–T5 and N4–N5)

Substructural features	Relative abundance (mean, %) of CRC and non-tumorigenic tissues					
	Early stage (Stage I—T1)	Early stage (Stage I—N1)	Mid stage (Stage III—T2–T3)	Mid stage (Stage III—N2–N3)	Late stage (Stage IV—T4–T5)	Late stage (Stage IV—N4–N5)
Total sialylation	64.8	62.8	80.2	72.6	87.6	79.1
α 2,6-sialylation	57.8	32.0	40.1	36.5	69.0	41.2
α 2,3-sialylation	6.5	30.7	41.8	37.9	20.7	38.1
Total fucosylation	81.1	85.1	81.4	83.6	74.0	78.4
Core fucosylation	75.4	84.3	79.1	82.8	72.6	77.7
Lewis-fucosylation	30.2	21.5	11.8	17.0	13.2	16.2
Bisecting GlcNAc	38.2	24.3	6.7	20.3	11.0	17.3
LacNAcylation	3.2	5.8	8.2	4.5	6.1	5.3

The N-glycan types are not included in this table since no CRC stage-dependent regulation was observed on this level (data not shown).

Table IV. Western blotting and IHC analysis of EGFR (×100 magnification) of tumor (T1–T5)

Tumor tissues	T1	T2	T3	T4	T5
Western blotting of EGFR					
Immunohistochemistry (IHC) stain for EGFR (×100 magnification)					
EGFR intensity score from IHC	+1	+1	+1	+2	0

EGFR was observed at 180 kDa; GAPDH (37 kDa) was used as a protein loading control. EGFR expression was classified into three grades (0 = negative, 1+ = focal weak positive, 2+ = diffuse strong positive). EGFR⁺ tumor (T1 and T4) are labeled green and EGFR⁻ tumor (T2, T3 and T5) are presented in red. The figure in the table is available in black and white in print and in color at *Glycobiology* online.

Table V. Distribution of the *N*-glycan types and substructural glyco-features (mean ± SD) of EGFR⁺ (T1, T4) and EGFR⁻ (T2, T3 and T5) CRC tumor tissues

	Relative abundance (mean ± SD, %)		Significance (EGFR ⁺ Vs EGFR ⁻ CRC tissues)
	EGFR ⁺ CRC tissues (T1, T4)	EGFR ⁻ CRC tissues (T2, T3, T5)	
<i>N</i>-Glycan type			
High mannose	47.6 ± 2.1	34.5 ± 1.9	*
Hybrid	3.8 ± 0.2	2.2 ± 0.7	*
Complex	43.7 ± 2.3	56.9 ± 3.8	*
Paucimannose	4.8 ± 0.5	6.5 ± 3.1	ns
Substructural feature			
Total fucosylation	74.4 ± 9.5	81.1 ± 0.6	ns
Total sialylation	75.9 ± 11.6	82.9 ± 5.1	ns
α2,6-Sialylation	66.9 ± 12.8	47.4 ± 13.7	ns
α2,3-Sialylation	9.3 ± 3.9	37.7 ± 9.1	*
Bisecting	26.6 ± 11.7	6.8 ± 0.9	*
GlcNAcylation			
LacNAcylation	3.8 ± 0.8	8.1 ± 0.3	ns

Expression differences of these glycoepitopes in the EGFR⁺ relative to EGFR⁻ tumors were investigated using student *t*-test (two tailed, type 3); significance ($P < 0.05$) is indicated by “*” and no significance ($P \geq 0.05$) by ns.

by the fucose-recognizing *Aleuria aurantia* lectin of CRC and non-tumorigenic colon tissues (Fernandez-Rodriguez et al. 2000). However, another study showed a CRC-specific increase in the enzyme activity and level of the α1,6-fucosyltransferase (Fut 8), which is responsible for the addition of core fucosylation (Muinelo-Romay et al. 2008). Hyper-fucosylation was also reported in other cancer types such as hepatocellular carcinoma (Chrostek and Cylwik 2011). Conversely, studies on ovarian cancer and gastric cancer have shown an association between malignancy and decreased levels of core fucosylation of N-linked glycoproteins (Anugraham et al. 2014; Zhao et al. 2014). Based on this contradicting literature and

the non-regulated fucosylation observed here, the role of terminal fucosylation in CRC remains unclear.

Increased sulfation of protein *N*-glycans was previously reported in CRC tissues relative to non-tumorigenic tissues (Balog et al. 2012). Sulfated *N*-glycans were not observed in this study most likely due to an alkaline NaBH₄-based elimination of these relative labile glycan modifications prior to the PGC-LC-MS/MS analysis (unpublished observation). The levels of *O*-linked sulfomucins have, in contrast, been reported to be decreased in the mucosa derived from colon cancer and inflammatory bowel diseases such as ulcerative colitis relative to normal mucosa (Yamori et al. 1987, 1989; Raouf et al. 1992). Another more recent study showed a unique expression of 3'-sulfo Lewis(x), the 3-*O*-sulfo-Globo unit, in colon cancer cell lines (Chandrasekaran et al. 2006). Taken together, these studies indicate that sulfation of protein *N*-glycans and, in particular, *O*-glycans is regulated during CRC, but the limited literature also highlights the need for establishing the exact molecular regulation and the functional involvement of the glycan sulfation in the CRC pathogenesis.

Our *N*-glycome data suggested that the CRC tissue *N*-glycosylation partially reflected disease progression and stage. Interestingly, this disease stage specificity of the glycome landscape in CRC tissues was largely reflected in the proteome profiles of the same paired CRC and non-tumorigenic tissues, but from a slightly larger cohort ($n = 8$) of CRC patients (Sethi et al. 2015), Figure 5C. This implies that the stage-specific molecular glyco-signatures are caused primarily by regulation of the protein expression and less by direct changes in the glycosylation machinery. High total sialylation, predominantly α2,3-sialylation, and lower bisecting GlcNAcylation and Lewis-type fucosylation were identified in the mid-late stages of CRC relative to the early stage. This is in agreement with a recent report demonstrating decreased bisecting GlcNAcylation in the metastatic ovarian cell line (SKOV-3) relative to the less metastatic ovarian counterparts (Zhang et al. 2014). In further support, the relationship between lower bisecting GlcNAcylation and higher metastatic potential has been reported (Yoshimura et al. 1996; Takahashi et al. 2009; Xu et al. 2012). Elevated α2,3-sialylation in the mid-late CRC stages corroborates previous observations, which associated increased sialylation to malignant characteristics, e.g., cancer invasion, progression

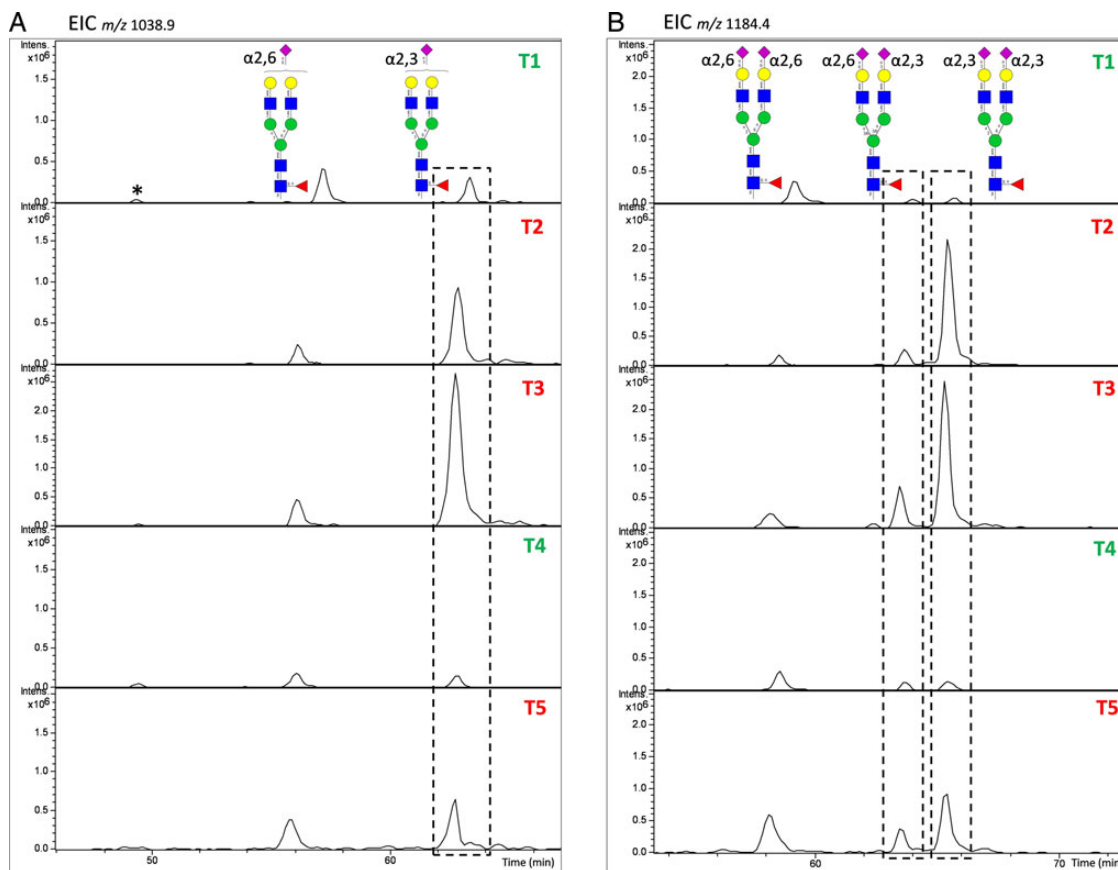


Fig. 4. EICs of two mono- and di-sialylated biantennary complex type *N*-glycans (A) m/z 1038.9⁻ and (B) m/z 1184.4⁻ containing both α 2,6- and α 2,3-sialylation. The ratio of α 2,3- : α 2,6-sialylation was found to be clearly EGFR dependent as evaluated by the lower level of α 2,3-sialylation in EGFR⁺ CRC (green, T1,T4) relative to EGFR⁻ CRC tissues (red, T2, T3 and T5). "*" represents a low abundant glycan isomer (glycan #36a). This figure is available in black and white in print and in color at *Glycobiology* online.

and metastasis (Harvey et al. 1992; Kyselova et al. 2008; Cui et al. 2011; Ferreira et al. 2013). Finally, the dominance of Lewis-type fucosylation in the early CRC stages was in agreement with a study reporting a correlation between the overexpression of α 1,2-fucosyltransferase 1 (Fut 1), which catalyzes the fucose transfer to galactose residues, and a reduced metastatic potential of pancreatic cancer cells (Aubert et al. 2000). These observed stage-specific alterations were also reflected in the matching non-tumorigenic tissues indicating more systemic aberration of the protein expression and the glycosylation machinery during disease progression. Taken together, even when considering the caveat of the limited sample set, the CRC stage-specific *N*-glycosylation opens interesting avenues for further investigation of stage-based markers for CRC.

EGFR is overexpressed in over one-third of all CRC tumors and is a recognized predictor of decreased survival and increased metastasis (Nicholson et al. 2001; Ellina et al. 2014). Accumulating evidence suggests that EGFR regulates the expression of multiple *N*-glycosylation epitopes, e.g., sialylation, fucosylation and bisecting GlcNAcylation, as well as glycosylation enzymes, e.g., *N*-acetylgalactosaminyltransferase 2 (*GALNT2*) and *N*-acetylglucosaminyltransferase V (*MGAT5*) (Sato et al. 2001; Matsumoto et al. 2008; Takahashi et al. 2009; Liu et al. 2011; Liu, Liu, et al. 2013; Lin et al. 2014). In line with this, several EGFR-dependent *N*-glycosylation signatures have been identified in the CRC tissues investigated here. Most importantly, high levels of bisecting GlcNAcylation were identified in EGFR⁺

relative to EGFR⁻ CRC tissues. High levels of bisecting GlcNAcylation may be related to the enhanced activity of *N*-acetylglucosaminyltransferase III (GnT-III) (*MGAT3*) that catalyzes the addition of β 1,4-linked (bisecting) GlcNAc residues to the trimannosylchitobiose core. In relation to this, a recent study demonstrated that WNT/ β -Catenin signalling regulates GnT-III expression (Kurimoto et al. 2014). Equally, EGFR is a known activator of the WNT/ β -Catenin signalling pathway (Hu and Li 2010). The potential crosstalk between EGFR and GnT-III suggests that elevated levels of bisecting GlcNAcylation in EGFR⁺ CRC tissues may result from an activated WNT/ β -Catenin signalling pathway. Furthermore, bisecting GlcNAc determinants on *N*-glycans displayed directly on EGFR were reported to diminish the formation of LacNAc repeats, which, in turn, decreased binding of EGFR to galectin-3 present on cell surfaces (and consequent endocytosis) (Takahashi et al. 2009). Our data showed a trend towards lower levels of LacNAc repeat-containing *N*-glycans in EGFR⁺ relative to EGFR⁻ CRC tissues suggesting an EGFR-regulated inverse correlation between bisecting GlcNAcylation and LacNAcylation. The elevated α 2,6-sialylation in EGFR⁺ CRC tumors is supported by an EGFR-centric study reporting that α 2,6-sialylation is directly displayed by EGFR-attached glycans in a panel of EGFR⁺ (but not EGFR⁻) CRC cell lines including HCT116, SW480, HT-29 and Lovo (Park et al. 2012). In addition, the EGFR-dependent regulation of α 2,3-sialylation and bisecting GlcNAcylation in opposite directions in the CRC tissues agrees well with our previous observations of

Table VI. Summary of the identified CRC-, disease stage- and EGFR-specific protein N-glycosylation signatures

	CRC relative to non-tumor tissue	EGFR+ relative to EGFR- CRC tissue	Late relative to early stage CRC tissue
N-Glycan type			
High mannose	↑	↑	No change
Hybrid	↑	↑	No change
Complex type	↓	↓	No change
Paucimannose	↑	No change	No change
Substructural feature			
Total fucosylation	No change	No change	No change
Core type	No change	No change	No change
fucosylation			
Lewis-type	No change	No change	↓
fucosylation			
Total sialylation	↑	No change	↑
α2,6-Sialylation	↑	No change	No change
α2,3-Sialylation	↓	↓	↑
Bisecting	No change	↑	↓
GlcNAcylation			
LacNAcylation	No change	No change	No change

The N-glycan types and their individual substructural glycosylation features are represented with ↑ and ↓ to indicate their significant ($P < 0.05$) positive or negative regulation, respectively. “No change” represents unaltered ($P \geq 0.05$) N-glycosylation features. The stage-specific regulation is based on the molecular trends observed rather than statistical significance due to the limited sample set available.

high bisecting GlcNAcylation and low α2,3-sialylation in the EGFR⁺ LIM1215 and EGFR⁻ LIM2405 cell line, respectively (Fanayan et al. 2013; Sethi et al. 2014). Together, these findings provide evidence supporting an EGFR-based regulation of terminal N-glycosylation determinants and their involvement in central processes involved in CRC pathogenesis.

In conclusion, this is the first study we are aware of that correlates CRC disease stage/progression and EGFR status with protein N-glycosylation signatures. These correlations were enabled by the accurate mapping of the N-glycosylation landscape of CRC tumor tissues with their paired adjacent non-tumor colon tissues. Mechanistically, the unique CRC tissue N-glycomes appear to arise from several sources including a deregulated glycosylation machinery and protein expression of the multiple cell types making up the complex tumor microenvironment and a significant contribution of glycoproteins secreted from infiltrating neutrophils. This study contributes to other system-wide initiatives aimed at advancing our understanding of the biomolecular deregulation associated with CRC. This study opens up new avenues for further investigations targeting the identification of novel glyco-signature-based markers for the early diagnosis and treatment of CRC and other diseases.

Materials and methods

Materials

Bradford dye reagent, AG 50W-X8 cation exchange resin, quick start™ and Tris base (Tris (hydroxymethyl) aminomethane or trimethamine) were purchased from Bio-Rad (Hercules, CA). Protease inhibitor cocktail tablets (EDTA free) and peptide-N-Glycosidase F (PNGase F) (*Flavobacterium meningospeticum*) were purchased from Roche Diagnostics (Mannheim, Germany). Polyvinylidene

fluoride membranes were purchased from Millipore. Empty top and C-18 stage tips were purchased from Glygen (Columbia, MD). Bovine serum albumin stock solution (2 mg/mL) and sodium chloride were purchased from GE Healthcare and Amersco, respectively. Triton X-114, sodium borohydride and all other chemicals (unless otherwise stated) were purchased from Sigma-Aldrich (St. Louis, MO).

Tissue samples

Paired primary CRC tissue samples (T1–T5) and their corresponding adjacent non-tumorigenic tissues (N1–N5) were obtained from the archives of the Department of Pathology, Yonsei University, Seoul, Korea and the Liver Cancer Specimen Bank of the National Research Resource Bank program of the Korea Science and Engineering Foundation of the Ministry of Science and Technology (Seoul, Korea). The CRC diagnosis of all tissue samples was carried out by a pathologist at the Severance Hospital, Yonsei University, Seoul, Korea. These samples were obtained for research purposes only with authorization from the Institutional Review Board of the College of Medicine at Yonsei University with informed consent of the subjects. All CRC primary tumors were identified as adenocarcinomas by a pathologist. The tissues were obtained from various sites (sigmoid, transverse and rectum) of the colon from five male patients (62.6 ± 13.1 y.o. at the time of sampling). The CRC tissues represented different disease stages (I–IV) and all tissues were typed according to their EGFR status [EGFR^{+/−} = 2/3] as determined by western blotting and IHC analyses.

Assessment of EGFR expression by western blotting and IHC

For western blotting analysis, whole tissue lysates were prepared using a passive lysis buffer (Promega). Samples were run on SDS-PAGE, transferred to nitrocellulose membranes and probed with primary anti-EGFR antibody (sc-03; Santa Cruz Biotechnology, Santa Cruz, CA) (200 μg/mL stock, used at 1 : 200 (v/v) dilution) for 1 h at room temperature. Membranes were washed and then incubated with secondary antibodies (sc-2301, goat anti-rabbit IgG-HRP, Santa Cruz Biotechnology) (400 μg/mL stock, used at 1 : 5000 (v/v) dilution) for 1 h at room temperature. Western blot images were recorded with an LAS-4000 Mini camera (Fujifilm). For IHC analysis, 3 μm thick tissue sections were obtained from formalin-fixed, paraffin-embedded CRC tissues of the five study subjects. The anti-EGFR primary antibody (monoclonal, ab52894, Abcam, Cambridge, UK) was used at 1 : 100 (v/v) dilution. IHC was performed following the manufacturer's protocol with the Ventana Discovery XT autoimmuno-stainer (Ventana Medical Systems, Tucson, AZ). EGFR expression was evaluated based on the anti-EGFR antibody intensity upon CRC tissue staining by a three-grade classification (0 = negative, 1+ = focal weak positive, 2+ = diffuse strong positive).

Tissue lysis and enrichment of membrane proteins by Triton X-114 phase-partitioning

Lysis of tissues and membrane protein enrichment were carried out as described in our previous study (Sethi et al. 2014). Briefly, tissue samples were lysed using a lysis buffer containing 200 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA·2H₂O, and a protease inhibitor cocktail (Roche Diagnostics, Germany). Samples were incubated on ice for 30 min followed by ultra-sonication (Sonicator ultrasonic processor XL, Misonix Inc, NY) at 15 s intervals for 2 min, with 15 s pause between each treatment. Samples were further centrifuged at $17,000 \times g$ for 1 h (Beckmann coulter Avanti™ J-25 centrifuge) and the supernatant obtained was diluted with a buffer containing 20 mM

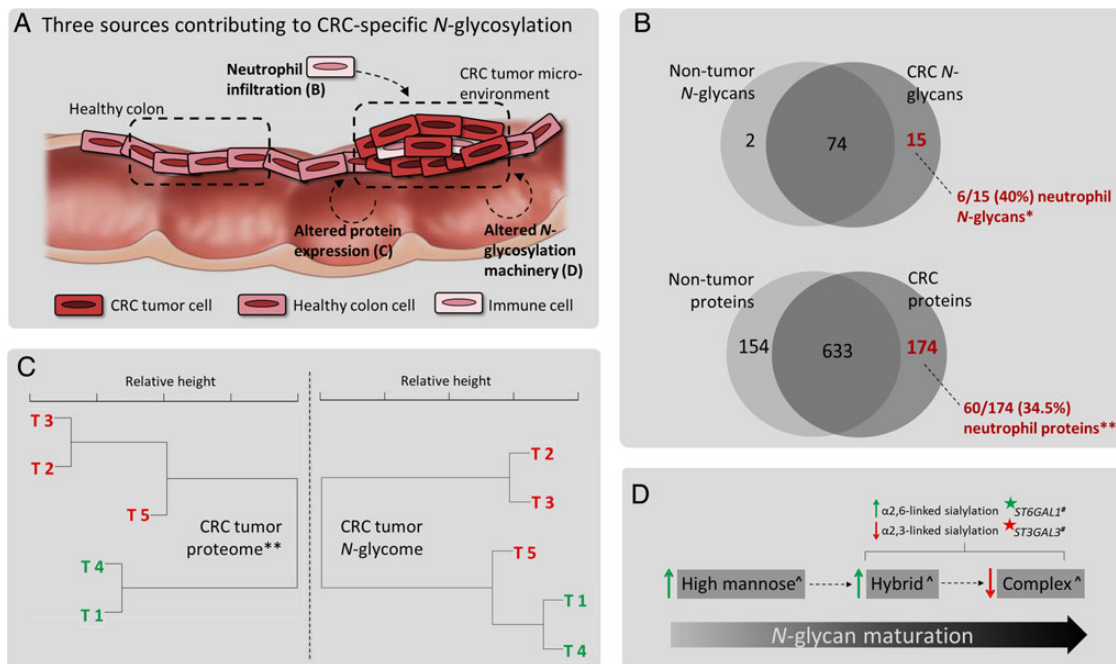


Fig. 5. (A) Schematic presentation of the regulation associated with the cancer pathogenesis including cellular and molecular regulation of the CRC tumor microenvironment leading to the aberrant *N*-glycomes observed in CRC tumor tissues. (B) The CRC-specific *N*-glycans (top) and proteins (bottom) are rich in molecular signatures from human neutrophils indicating cellular infiltration in the CRC tumor microenvironment. (*) The CRC tissue *N*-glycomes were compared with previously reported *N*-glycomes of human neutrophils (Babu et al. 2009; Thaysen-Andersen et al. 2015). (**) The CRC tissue proteomes ($n=8$) (Sethi et al. 2015) were compared with previously reported proteome data of human neutrophils (Rorvig et al. 2013). (C) The CRC tissue *N*-glycomes (right) and membrane proteomes (left) (Sethi et al. 2015) clustered the five CRC patients similarly according to their CRC stage (color-coded) and separately from non-tumor tissues (data not shown), indicating that alterations in protein expression is a significant contributor to the aberrant *N*-glycosylation observed in CRC tumors. (D) The aberrant glyco-features, which were supported by similar observations by others in CRC tissues or cells ^ indicates de Leoz et al. (2011), Liu, Nie, et al. (2013), Anugraham et al. (2014) and Sethi et al. (2014), correlated with the aberrant expression of the responsible glycosylation enzymes as previously reported # indicates Harvey et al. (1992), Pousset et al. (1997), Recchi et al. (1998), Fernandez-Rodriguez et al. (2000), Wang et al. (2003), Park and Lee (2013), Anugraham et al. (2014) and Sethi et al. (2014). This indicated that an altered glycosylation machinery of one or multiple cell types in the tumor microenvironment also contributed to the unique *N*-glycosylation signature of CRC tissues. Green arrows indicate positive regulation and red arrows represent a negative regulation of the individual glyco-features in CRC tissues relative to non-tumorigenic colon tissues. Increased and decreased gene expression in CRC tissues are represented by green and red stars, respectively. This figure is available in black and white in print and in color at *Glycobiology* online.

Tris-HCl, 100 mM NaCl buffer followed by ultra-centrifugation at 120,000 $\times g$ for 80 min (Beckmann Coulter optima™ L-100 XP ultra-centrifuge). The supernatant was removed and the pellet was resuspended in 20 mM Tris-HCl, 100 mM NaCl and 1% (v/v) Triton X-114 and chilled on ice with intermittent vortexing for 20 min. Samples were heated at 37°C for 20 min and phase-partitioned by centrifugation at 1000 $\times g$ for 3 min. The upper aqueous phase was removed and stored. The detergent phase was further diluted with four volumes of 1% (v/v) Triton X-114 and the phase-partitioning was repeated. The combined aqueous phases were mixed with nine volumes of ice-cold acetone overnight at -20°C to precipitate proteins and remove any detergent. Precipitated membrane proteins were solubilized with 8 M urea and protein concentrations determined by Bradford protein assay. The enriched membrane proteins were stored at -80°C if not used immediately.

Release of *N*-glycans from membrane glycoproteins

N-Glycans were released from isolated membrane proteins using an established protocol (Jensen et al. 2012). Briefly, 20 μg of membrane proteins were spotted on polyvinylidene fluoride membranes and stained with Direct Blue (Sigma-Aldrich). Membrane spots were excised and washed in separate wells in a flat bottom polypropylene 96-well plate (Corning Incorporated, Corning, NY). *N*-Linked glycans

were released from the membranes using 3 U/well of PNGase F (*Flavobacterium meningospeticum*) with overnight incubation at 37°C. Released *N*-glycans were reduced with 20 μL of 1 M NaBH₄ in 50 mM KOH for 3 h at 50°C. The glycan reduction was quenched using 2 μL of glacial acetic acid, desalted using AG 50W-X8 cation exchange resins (Bio-Rad) and packed in empty top zip-tips. Released and reduced *N*-glycans were dried by vacuum centrifugation and further purified by repeated cycles of washing with 100 μL of methanol and drying by vacuum centrifugation after every wash. The *N*-glycans were further cleaned up with PGC SPE where the retained and desalted *N*-glycans were eluted by 0.1% (v/v) trifluoroacetic acid in 40% acetonitrile followed by drying by vacuum centrifugation. The purified *N*-glycans were resuspended in 10 μL of water and subjected to PGC-LC-ESI-MS/MS in two technical replicate injections.

PGC-LC-ESI-MS/MS and glycan data analysis

Released *N*-glycans were separated on a Hypercarb PGC column with the following specifications: Pore size: 250 Å; particle size: 5 μm ; inner diameter: 320 μm ; length: 10 cm (Thermo Scientific). Glycan separation was performed on an Agilent 1100 capillary LC (Agilent Technologies, Santa Clara, CA) and detection was performed using an Agilent MSD three-dimensional ion-trap XCT Plus mass spectrometer

coupled directly to the LC. The instrument parameter settings were similar to those described previously (Sethi et al. 2014). Briefly, separation was carried out at a constant flow rate of 2 $\mu\text{L}/\text{min}$ using a linear gradient from 0 to 16% (v/v) acetonitrile in 10 mM NH_4HCO_3 for 45 min, followed by a gradient from 16–45% (v/v) acetonitrile in 10 mM NH_4HCO_3 over 20 min and finally washing the column with 45% (v/v) acetonitrile in 10 mM NH_4HCO_3 for 6 min and re-equilibration in 10 mM NH_4HCO_3 (aq). ESI-MS was performed in negative ion mode with two scan events: MS full scan (m/z 100–2200) with scan speed of 8200 $m/z/s$ and data-dependent MS/MS scan after CID of the top two most intense precursor ions with an absolute intensity threshold of 30,000. No dynamic exclusion was activated. Raw MS data were viewed using ESI-compass 1.3 software (Bruker Daltonics) and the glycan molecular masses were derived from manual interpretation. The glycan monosaccharide compositions for the glycan precursors were predicted using the GlycoMod tool (<http://www.expasy.ch/tools/glycomod>) with mass accuracies generally better than 0.2 Da. The glycan structures were characterized using (i) molecular monoisotopic mass; (ii) CID-MS/MS de novo sequencing; and (iii) PGC-LC relative retention time (Pabst et al. 2007). For further validation, MS/MS fragmentation profiles were matched to in silico calculated glycosidic fragments (Everest-Dass et al. 2013), CID-MS/MS fragment spectra of structures deposited in UniCarb-DB (<http://www.unicarb-db.org>) (Hayes et al. 2011) and to spectra from recently published in-house datasets (Parker et al. 2013; Sethi et al. 2014). GlycoWorkbench (<http://www.eurocarb-db.org/applications/ms-tools>) (Ceroni et al. 2008) was additionally employed to assist in topology and branching assignment of the structures and for the generation of in silico generated glycosidic (B/Y- and C/Z-type) fragments. The characterized *N*-glycans were assigned with non-metric confidence scores (low/medium/high) based on the above criteria as well as their structural similarities (Supplementary data, Table S1). Some structural ambiguities, in particular with regards to the linkage types, remained and such structures were left without complete assignment. Partial linkage and topology assignment was performed for some structural aspects using the knowledge of the biosynthetic machinery. The annotated MS/MS fragment spectra (Supplementary data, Figure S1) were deposited in the UniCarb-DB MS/MS glycan database (<http://www.unicarb-db.org>). The relative abundance of each *N*-glycan isomer was manually determined from the relative area of the EICs of *N*-glycan precursors ($\pm m/z$ 0.5) in all observed charged states ($Z = -1$ to -3). This quantitative approach has been evaluated against other *N*-glycan workflows and found to be accurate when glycoprofiles from multiple samples are being compared (Leymarie et al. 2013).

Statistical analysis

Student's *t*-tests were performed with two-tailed distribution using Microsoft Excel to test for alterations in the obtained *N*-glycome profiles (i.e., the relative abundance of the *N*-glycan types, individual *N*-glycans or *N*-glycan substructural features). Paired *t*-test (type 1) was performed to test for such differences of the CRC tumor and their paired non-tumorigenic colon tissues and separately for the EGFR^{+/−} dependency using an unpaired heteroscedastic *t*-test (type 2). Both analyses were performed using the normalized relative abundance (in %) of the individual *N*-glycan structures or common features on multiple glycans. $P < 0.05$ was regarded as significant. The relative abundance of the glycan structures and substructures are provided as mean \pm SD in all bar graph representations. Hierarchical cluster analysis was performed on the CRC tissues *N*-glycomes (T1–T5) and

previously published proteomes (Sethi et al. 2015) to assess the ability of the two types of biomolecules to separate the tissues into individual groups. Hierarchical clustering was performed using the “hclust” implementation from R statistical package with Euclidean distance and complete linkage.

Supplementary data

Supplementary data for this article are available online at <http://glycob.oxfordjournals.org/>.

Acknowledgements

This research was facilitated through access to the Australian Proteomics Analysis Facility (APAF) and funds from 2013 Northern translational cancer research unit (NTCRU) research scholar award and a Macquarie University postgraduate student research scholarship (MQRES). We thank Dr. Dana Pascovici for assistance with statistical analysis. M.T.-A. was supported by an Early Career Fellowship from the Cancer Institute, NSW, Australia.

Conflict of interest statement

None declared.

Abbreviations

CID, collision-induced dissociation; CRC, colorectal cancer; EGFR, epidermal growth factor receptor; EIC, extracted ion chromatogram; ESI, electrospray ionization; GlcNAc, *N*-acetylglucosamine; GnT-III, *N*-acetylglucosaminyltransferase III; IHC, immunohistochemistry; LacNAc, *N*-acetylglucosamine; LC, liquid chromatography; MS/MS, tandem mass spectrometry; PGC, porous graphitized carbon; PNGase F, peptide-*N*-glycosidase F.

References

- Adamczyk B, Tharmalingam-Jaikaran T, Schomberg M, Szekrenyes A, Kelly RM, Karlsson NG, Guttman A, Rudd PM. 2014. Comparison of separation techniques for the elucidation of IgG *N*-glycans pooled from healthy mammalian species. *Carbohydr Res*. 389:174–185.
- Anugraham M, Jacob F, Nixdorf S, Everest-Dass AV, Heinzelmann-Schwarz V, Packer NH. 2014. Specific glycosylation of membrane proteins in epithelial ovarian cancer cell lines: Glycan structures reflect gene expression and DNA methylation status. *Mol Cell Proteomics*. 13(9):2213–2232.
- Aubert M, Panicot L, Crotte C, Gibier P, Lombardo D, Sadoulet MO, Mas E. 2000. Restoration of alpha(1,2) fucosyltransferase activity decreases adhesive and metastatic properties of human pancreatic cancer cells. *Cancer Res*. 60(5):1449–1456.
- Baas JM, Krens LL, Guchelaar HJ, Morreau H, Gelderblom H. 2011. Concordance of predictive markers for EGFR inhibitors in primary tumors and metastases in colorectal cancer: A review. *Oncologist*. 16(9):1239–1249.
- Babu P, North SJ, Jang-Lee J, Chalabi S, Mackerness K, Stowell SR, Cummings RD, Rankin S, Dell A, Haslam SM. 2009. Structural characterisation of neutrophil glycans by ultra sensitive mass spectrometric glycomics methodology. *Glycoconj J*. 26(8):975–986.
- Balog CI, Stavenhagen K, Fung WL, Koeleman CA, McDonnell LA, Verhoeven A, Mesker WE, Tollenaar RA, Deelder AM, Wuhrer M. 2012. *N*-glycosylation of colorectal cancer tissues: A liquid chromatography and mass spectrometry-based investigation. *Mol Cell Proteomics*. 11(9):571–585.
- Carpenter G, Cohen S. 1990. Epidermal growth factor. *J Biol Chem*. 265(14):7709–7712.
- Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM. 2008. GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res*. 7(4):1650–1659.

- Chandrasekaran EV, Xue J, Neelamegham S, Matta KL. 2006. The pattern of glycosyl- and sulfotransferase activities in cancer cell lines: A predictor of individual cancer-associated distinct carbohydrate structures for the structural identification of signature glycans. *Carbohydr Res*. 341(8):983–994.
- Chen S, LaRoche T, Hamelinck D, Bergsma D, Brenner D, Simeone D, Brand RE, Haab BB. 2007. Multiplexed analysis of glycan variation on native proteins captured by antibody microarrays. *Nat Methods*. 4(5):437–444.
- Christiansen MN, Chik J, Lee L, Anugraham M, Abrahams JL, Packer NH. 2014. Cell surface protein glycosylation in cancer. *Proteomics*. 14(4–5): 525–546.
- Chrostek L, Cylwik B. 2011. [The alteration of proteins glycosylation in liver diseases]. *Pol Merkur Lekarski*. 31(181):60–64.
- Cools-Lartigue J, Spicer J, Najmeh S, Ferri L. 2014. Neutrophil extracellular traps in cancer progression. *Cell Mol Life Sci*. 71(21):4179–4194.
- Cui H, Lin Y, Yue L, Zhao X, Liu J. 2011. Differential expression of the alpha2,3-sialic acid residues in breast cancer is associated with metastatic potential. *Oncol Rep*. 25(5):1365–1371.
- Dahmen AC, Fergen MT, Laurini C, Schmitz B, Loke I, Thaysen-Andersen M, Diestel S. 2015. Paucimannosidic glycoepitopes are functionally involved in proliferation of neural progenitor cells in the subventricular zone. *Glycobiology* 25(8):869–880.
- Davies RJ, Miller R, Coleman N. 2005. Colorectal cancer screening: Prospects for molecular stool analysis. *Nat Rev Cancer*. 5(3):199–209.
- de Leoz ML, Young LJ, An HJ, Kronewitter SR, Kim J, Miyamoto S, Borowsky AD, Chew HK, Lebrilla CB. 2011. High-mannose glycans are elevated during breast cancer progression. *Mol Cell Proteomics*. 10(1): M110 002717.
- de Wit M, Kant H, Piersma SR, Pham TV, Mongera S, van MBerkel P, Boven E, Ponten F, Meijer GA, Jimenez CR, et al. 2014. Colorectal cancer candidate biomarkers identified by tissue secretome proteome profiling. *J Proteomics*. 99:26–39.
- Duffy MJ, van Dalen A, Haglund C, Hansson L, Klapdor R, Lamerz R, Nilsson O, Sturgeon C, Topolcan O. 2003. Clinical utility of biochemical markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines. *Eur J Cancer*. 39(6):718–727.
- Ellina MI, Bouris P, Aletas AJ, Theocharis AD, Kletsas D, Karamanos NK. 2014. EGFR and HER2 exert distinct roles on colon cancer cell functional properties and expression of matrix macromolecules. *Biochim Biophys Acta*. 1840(8):2651–2661.
- Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, Reid B, Radich J, Anderson G, Hartwell L. 2003. The case for early detection. *Nat Rev Cancer*. 3(4):243–252.
- Everest-Dass AV, Abrahams JL, Kolarich D, Packer NH, Campbell MP. 2013. Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS. *J Am Soc Mass Spectrom*. 24(6):895–906.
- Everest-Dass AV, Jin D, Thaysen-Andersen M, Nevalainen H, Kolarich D, Packer NH. 2012. Comparative structural analysis of the glycosylation of salivary and buccal cell proteins: Innate protection against infection by *Candida albicans*. *Glycobiology*. 22(11):1465–1479.
- Fanayan S, Smith JT, Lee LY, Yan F, Snyder M, Hancock WS, Nice E. 2013. Proteogenomic analysis of human colon carcinoma cell lines LIM1215, LIM1899, and LIM2405. *J Proteome Res*. 12(4):1732–1742.
- Fernandez-Rodriguez J, Feijoo-Carnero C, Merino-Trigo A, Paez de la Cadena M, Rodriguez-Berocal FJ, de Carlos A, Butron M, Martinez-Zorzano VS. 2000. Immunohistochemical analysis of sialic acid and fucose composition in human colorectal adenocarcinoma. *Tumour Biol*. 21(3): 153–164.
- Ferreira SA, Vasconcelos JL, Silva RC, Cavalcanti CL, Bezerra CL, Rego MJ, Beltrao EI. 2013. Expression patterns of alpha2,3-sialyltransferase I and alpha2,6-sialyltransferase I in human cutaneous epithelial lesions. *Eur J Histochem*. 57(1):e7.
- Frost DC, Li L. 2014. Recent advances in mass spectrometry-based glycoproteomics. *Adv Protein Chem Struct Biol*. 95:71–123.
- Gan Y, Chen D, Li X. 2014. Proteomic analysis reveals novel proteins associated with progression and differentiation of colorectal carcinoma. *J Cancer Res Ther*. 10(1):89–96.
- Harris AL, Nicholson S, Sainsbury R, Wright C, Farnon J. 1992. Epidermal growth factor receptor and other oncogenes as prognostic markers. *J Natl Cancer Inst Monogr*. (11):181–187.
- Harvey BE, Toth CA, Wagner HE, Steele GD Jr, Thomas P. 1992. Sialyltransferase activity and hepatic tumor growth in a nude mouse model of colorectal cancer metastases. *Cancer Res*. 52(7):1775–1779.
- Hayes CA, Karlsson NG, Struwe WB, Lisacek F, Rudd PM, Packer NH, Campbell MP. 2011. UniCarb-DB: A database resource for glycomic discovery. *Bioinformatics*. 27(9):1343–1344.
- Hu T, Li C. 2010. Convergence between Wnt-beta-catenin and EGFR signaling in cancer. *Mol Cancer*. 9:236.
- Huffman JE, Pucic-Bakovic M, Klaric L, Hennig R, Selman MH, Vuckovic F, Novokmet M, Kristic J, Borowiak M, Muth T, et al. 2014. Comparative performance of four methods for high-throughput glycosylation analysis of immunoglobulin G in genetic and epidemiological research. *Mol Cell Proteomics*. 13(6):1598–1610.
- Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. 2011. Global cancer statistics. *CA Cancer J Clin*. 61(2):69–90.
- Jensen PH, Karlsson NG, Kolarich D, Packer NH. 2012. Structural analysis of N- and O-glycans released from glycoproteins. *Nat Protoc*. 7(7):1299–1310.
- Kurimoto A, Kitazume S, Kizuka Y, Nakajima K, Oka R, Fujinawa R, Korekane H, Yamaguchi Y, Wada Y, Taniguchi N. 2014. The absence of core fucose up-regulates GnT-III and Wnt target genes: A possible mechanism for an adaptive response in terms of glycan function. *J Biol Chem*. 289(17):11704–11714.
- Kyselova Z, Mechref Y, Kang P, Goetz JA, Dobrolecki LE, Sledge GW, Schnaper L, Hickey RJ, Malkas LH, Novotny MV. 2008. Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles. *Clin Chem*. 54(7):1166–1175.
- Lee LY, Lin CH, Fanayan S, Packer NH, Thaysen-Andersen M. 2014. Differential site accessibility mechanistically explains subcellular-specific N-glycosylation determinants. *Front Immunol*. 5:404.
- Lee LY, Thaysen-Andersen M, Baker MS, Packer NH, Hancock WS, Fanayan S. 2014. Comprehensive N-glycome profiling of cultured human epithelial breast cells identifies unique secretome N-glycosylation signatures enabling tumorigenic subtype classification. *J Proteome Res*. 13(11):4783–4795.
- Leymarie N, Griffin PJ, Jonscher K, Kolarich D, Orlando R, McComb M, Zaia J, Aguilan J, Alley WR, Altmann F, et al. 2013. Interlaboratory study on differential analysis of protein glycosylation by mass spectrometry: The ABRF glycoprotein research multi-institutional study 2012. *Mol Cell Proteomics*. 12(10):2935–2951.
- Lin MC, Huang MJ, Liu CH, Yang TL, Huang MC. 2014. GALNT2 enhances migration and invasion of oral squamous cell carcinoma by regulating EGFR glycosylation and activity. *Oral Oncol*. 50(5):478–484.
- Liu J, Liu H, Zhang W, Wu Q, Liu W, Liu Y, Pan D, Xu J, Gu J. 2013. N-acetylglucosaminyltransferase V confers hepatoma cells with resistance to anoikis through EGFR/PAK1 activation. *Glycobiology*. 23(9):1097–1109.
- Liu X, Nie H, Zhang Y, Yao Y, Maitikabili A, Qu Y, Shi S, Chen C, Li Y. 2013. Cell surface-specific N-glycan profiling in breast cancer. *PLoS ONE*. 8(8): e72704.
- Liu YC, Yen HY, Chen CY, Chen CH, Cheng PF, Juan YH, Khoo KH, Yu CJ, Yang PC, Hsu TL, et al. 2011. Sialylation and fucosylation of epidermal growth factor receptor suppress its dimerization and activation in lung cancer cells. *Proc Natl Acad Sci USA*. 108(28):11332–11337.
- Luque-Garcia JL, Martinez-Torrecuadrada JL, Epifano C, Canamero M, Babel I, Casal JI. 2010. Differential protein expression on the cell surface of colorectal cancer cells associated to tumor metastasis. *Proteomics*. 10(5):940–952.
- Matsumoto K, Yokote H, Arao T, Maegawa M, Tanaka K, Fujita Y, Shimizu C, Hanafusa T, Fujiwara Y, Nishio K. 2008. N-Glycan fucosylation of epidermal growth factor receptor modulates receptor activity and sensitivity to epidermal growth factor receptor tyrosine kinase inhibitor. *Cancer Sci*. 99(8):1611–1617.
- Mendelsohn J, Baselga J. 2006. Epidermal growth factor receptor targeting in cancer. *Semin Oncol*. 33(4):369–385.
- Muinelo-Romay L, Vazquez-Martin C, Villar-Portela S, Cuevas E, Gil-Martin E, Fernandez-Briera A. 2008. Expression and enzyme activity

- of alpha(1,6)fucosyltransferase in human colorectal cancer. *Int J Cancer*. 123(3):641–646.
- Nicholson RI, Gee JM, Harper ME. 2001. EGFR and cancer prognosis. *Eur J Cancer*. 37 Suppl 4:S9–S15.
- Pabst M, Bondili JS, Stadlmann J, Mach L, Altmann F. 2007. Mass + retention time = structure: A strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans. *Anal Chem*. 79(13):5051–5057.
- Park JJ, Lee M. 2013. Increasing the alpha 2, 6 sialylation of glycoproteins may contribute to metastatic spread and therapeutic resistance in colorectal cancer. *Gut Liver*. 7(6):629–641.
- Park JJ, Yi JY, Jin YB, Lee YJ, Lee JS, Lee YS, Ko YG, Lee M. 2012. Sialylation of epidermal growth factor receptor regulates receptor activity and chemosensitivity to gefitinib in colon cancer cells. *Biochem Pharmacol*. 83(7):849–857.
- Parker BL, Thaysen-Andersen M, Solis N, Scott NE, Larsen MR, Graham ME, Packer NH, Cordwell SJ. 2013. Site-specific glycan-peptide analysis for determination of N-glycoproteome heterogeneity. *J Proteome Res*. 12(12):5791–5800.
- Pousset D, Piller V, Bureaud N, Monsigny M, Piller F. 1997. Increased alpha2,6 sialylation of N-glycans in a transgenic mouse model of hepatocellular carcinoma. *Cancer Res*. 57(19):4249–4256.
- Raouf AH, Tsai HH, Parker N, Hoffman J, Walker RJ, Rhodes JM. 1992. Sulphation of colonic and rectal mucin in inflammatory bowel disease: Reduced sulphation of rectal mucus in ulcerative colitis. *Clin Sci (Lond)*. 83(5):623–626.
- Recchi MA, Hebbbar M, Hornez L, Harduin-Lepers A, Peyrat JP, Delannoy P. 1998. Multiplex reverse transcription polymerase chain reaction assessment of sialyltransferase expression in human breast cancer. *Cancer Res*. 58(18):4066–4070.
- Rorvig S, Ostergaard O, Heegaard NH, Borregaard N. 2013. Proteome profiling of human neutrophil granule subsets, secretory vesicles, and cell membrane: Correlation with transcriptome profiling of neutrophil precursors. *J Leukoc Biol*. 94(4):711–721.
- Sato Y, Takahashi M, Shibukawa Y, Jain SK, Hamaoka R, Miyagawa J, Yaginuma Y, Honke K, Ishikawa M, Taniguchi N. 2001. Overexpression of N-acetylglucosaminyltransferase III enhances the epidermal growth factor-induced phosphorylation of ERK in HeLaS3 cells by up-regulation of the internalization rate of the receptors. *J Biol Chem*. 276(15):11956–11962.
- Schachter H. 2009. Paucimannose N-glycans in *Caenorhabditis elegans* and *Drosophila melanogaster*. *Carbohydr Res*. 344(12):1391–1396.
- Schlessinger J. 2002. Ligand-induced, receptor-mediated dimerization and activation of EGF receptor. *Cell*. 110(6):669–672.
- Sethi MK, Thaysen-Andersen M, Kim H, Park CK, Baker MS, Packer NH, Paik YK, Hancock WS, Fanayan S, et al. 2015. Quantitative proteomic analysis of paired colorectal cancer and non-tumorigenic tissues reveals signature proteins and perturbed pathways involved in CRC progression and metastasis. *J Proteomics*. 126:54–67.
- Sethi MK, Thaysen-Andersen M, Smith JT, Baker MS, Packer NH, Hancock WS, Fanayan S. 2014. Comparative N-glycan profiling of colorectal cancer cell lines reveals unique bisecting GlcNAc and alpha-2,3-linked sialic acid determinants are associated with membrane proteins of the more metastatic/aggressive cell lines. *J Proteome Res*. 13(1):277–288.
- Takahashi M, Kuroki Y, Ohtsubo K, Taniguchi N. 2009. Core fucose and bisecting GlcNAc, the direct modifiers of the N-glycan core: Their functions and target proteins. *Carbohydr Res*. 344(12):1387–1390.
- Tanahashi T, Namba K, Murao T. 1990. [Studies on relationship between histology, tumor markers (prostatic acid phosphatase, prostate specific antigen, gamma-seminoprotein, leu-7) and clinical course in prostate cancer]. *Nihon Hinyokika Gakkai Zasshi*. 81(5):680–685.
- Tanaka T, Tanaka M, Ishigamori R. 2010. Biomarkers for colorectal cancer. *Int J Mol Sci*. 11(9):3209–3225.
- Terdiman JP. 2005. Colonoscopy is superior to flexible sigmoidoscopy for colorectal cancer screening: Now beyond a reasonable doubt? *Gastroenterology*. 129(5):1793–1794.
- Thaysen-Andersen M, Packer NH. 2014. Advances in LC-MS/MS-based glycoproteomics: Getting closer to system-wide site-specific mapping of the N- and O-glycoproteome. *Biochim Biophys Acta*. 1844(9):1437–1452.
- Thaysen-Andersen M, Venkatakrishnan V, Loke I, Laurini C, Diestel S, Parker BL, Packer NH. 2015. Human neutrophils secrete bioactive paucimannosidic proteins from azurophilic granules into pathogen-infected sputum. *J Biol Chem*. 290(14):8789–8802.
- Venkatakrishnan V, Thaysen-Andersen M, Chen SC, Nevalainen H, Packer NH. 2015. Cystic fibrosis and bacterial colonization define the sputum N-glycosylation phenotype. *Glycobiology*. 25(1):88–100.
- Wang PH, Lee WL, Lee YR, Juang CM, Chen YJ, Chao HT, Tsai YC, Yuan CC. 2003. Enhanced expression of alpha 2,6-sialyltransferase ST6Gal I in cervical squamous cell carcinoma. *Gynecol Oncol*. 89(3):395–401.
- Wang HJ, Zhu JS, Zhang Q, Sun Q, Guo H. 2009. High level of ezrin expression in colorectal cancer tissues is closely related to tumor malignancy. *World J Gastroenterol*. 15(16):2016–2019.
- Wulfkuehl JD, Liotta LA, Petricoin EF. 2003. Proteomic applications for the early detection of cancer. *Nat Rev Cancer*. 3(4):267–275.
- Xu Q, Isaji T, Lu Y, Gu W, Kondo M, Fukuda T, Du Y, Gu J. 2012. Roles of N-acetylglucosaminyltransferase III in epithelial-to-mesenchymal transition induced by transforming growth factor beta1 (TGF-beta1) in epithelial cell lines. *J Biol Chem*. 287(20):16563–16574.
- Yamori T, Kimura H, Stewart K, Ota DM, Cleary KR, Irimura T. 1987. Differential production of high molecular weight sulfated glycoproteins in normal colonic mucosa, primary colon carcinoma, and metastases. *Cancer Res*. 47(10):2741–2747.
- Yamori T, Ota DM, Cleary KR, Hoff S, Hager LG, Irimura T. 1989. Monoclonal antibody against human colonic sulfomucin: Immunochemical detection of its binding sites in colonic mucosa, colorectal primary carcinoma, and metastases. *Cancer Res*. 49(4):887–894.
- Yarden Y. 2001. The EGFR family and its ligands in human cancer: signalling mechanisms and therapeutic opportunities. *Eur J Cancer*. 37(Suppl 4):S3–S8.
- Yoshimura M, Ihara Y, Matsuzawa Y, Taniguchi N. 1996. Aberrant glycosylation of E-cadherin enhances cell-cell binding to suppress metastasis. *J Biol Chem*. 271(23):13811–13815.
- Zhang X, Wang Y, Qian Y, Wu X, Zhang Z, Liu X, Zhao R, Zhou L, Ruan Y, Xu J, et al. 2014. Discovery of specific metastasis-related N-glycan alterations in epithelial ovarian cancer based on quantitative glycomics. *PLoS ONE*. 9(2):e87978.
- Zhao YP, Xu XY, Fang M, Wang H, You Q, Yi CH, Ji J, Gu X, Zhou PT, Cheng C, et al. 2014. Decreased core-fucosylation contributes to malignancy in gastric cancer. *PLoS ONE*. 9(4):e94536.
- Zhou S, Hu Y, DeSantos-Garcia JL, Mechref Y. 2015. Quantitation of Per-methylated N-Glycans through Multiple-Reaction Monitoring (MRM) LC-MS/MS. *J Am Soc Mass Spectrom*. 26(4):596–603.