Author's personal copy

Breast Cancer Res Treat (2014) 145:295–305
DOI 10.1007/s10549-014-2949-z

PRECLINICAL STUDY

Prognostic relevance of glycosylation-associated genes in breast cancer

Karin Milde-Langosch · Thomas Karn · Marcus Schmidt · Christine zu Eulenburg · Leticia Oliveira-Ferrer · Ralph M. Wirtz · Udo Schumacher · Isabell Witzel · Dina Schütze · Volkmar Müller

Received: 13 November 2013 / Accepted: 2 April 2014 / Published online: 16 April 2014
© Springer Science+Business Media New York 2014

Abstract Glycosylation of cellular proteins has important impact on their stability and functional properties, and glycan structures strongly influence cell adhesion. Many enzymes are involved in glycoconjugate synthesis and degradation, but there is only limited information about their role in breast cancer progression. Therefore, we retrieved RNA expression data of 202 glycosylation genes generated by microarray analysis (Affymetrix HG-U133A) in a cohort of 194 mammary carcinomas with long-term follow-up information. After univariate and multivariate Cox regression analysis, genes with independent prognostic value were identified. These were further analysed by Kaplan–Meier analysis and log-rank tests, and their prognostic value was validated in a second cohort of 200 tumour samples from patients without systemic therapy. In our first cohort, we identified 24 genes with independent prognostic value, coding for sixteen anabolic and eight catabolic enzymes. Functionally, these genes are involved in all important glycosylation pathways, namely O-glycosylation, N-glycosylation, O-fucosylation, synthesis of glycosaminoglycans and glycolipids. Eighteen genes also showed prognostic significance in chemotherapy-treated patients. In the second cohort, six of the 24 relevant genes were of prognostic significance (FUT1, FUCA1, POFUT1, MAN1A1, RPN1 and DPM1), whereas a trend was observed for three additional probesets (GCNT4, ST3GAL6 and UGCG). In a stratified analysis of molecular subtypes combining both cohorts, great differences appeared suggesting a predominant role of N-glycosylation in luminal cancers and O-glycosylation in triple-negative ones. Correlations of gene expression with metastases of various localizations point to a role of glycan structures in organ-specific metastatic spread. Our results indicate that various glycosylation reactions influence progression and metastasis of breast cancer and might thus represent potential therapeutic targets.

Keywords Glycosylation · Glycan genes · Breast cancer · Prognosis · Microarray analysis

Electronic supplementary material The online version of this article (doi:10.1007/s10549-014-2949-z) contains supplementary material, which is available to authorized users.

K. Milde-Langosch (✉) · L. Oliveira-Ferrer · I. Witzel · D. Schütze · V. Müller
Department of Gynecology, University Hospital Hamburg-Eppendorf, Hamburg, Germany
e-mail: milde@uke.de

T. Karn
Department of Obstetrics and Gynecology, Goethe-University Frankfurt, Frankfurt am Main, Germany

M. Schmidt
Department of Obstetrics and Gynecology, Johannes Gutenberg-University Mainz, Mainz, Germany

C. zu Eulenburg
Department of Medical Biometry and Epidemiology, University Hospital Hamburg-Eppendorf, Hamburg, Germany

R. M. Wirtz
STRATIFYER Molecular Pathology GmbH, Cologne, Germany

U. Schumacher
Department of Anatomy and Experimental Morphology, University Hospital Hamburg-Eppendorf, Hamburg, Germany

Introduction

After extensive research on the genome, transcriptome and proteome of human cells, the study of glycans has gradually come into focus in the last years. The term glycan defines the carbohydrate moiety of a glycoprotein,

Author's personal copy

glycolipid or a proteoglycan. These glycoconjugates derive from glycosylation of proteins or lipids by stepwise covalent attachment of oligosaccharides. This co- or post-translational modification has deep impact on protein folding and stability, transport and secretion, enzymatic activity, binding properties and antigenicity.

Glycosylation has been reported to change significantly during embryogenesis, and aberrant glycans due to defects in single glycosylation genes frequently lead to embryonic lethality [27, 38]. Moreover, changes in glycan structures have been early detected in various malignancies including breast cancer [7]. Glycoproteins are used as tumour markers for diagnosis and tumour monitoring, and their use as diagnostic tools or therapeutic targets has been discussed [8, 31, 36].

The importance of glycosylation processes is reflected by the high number of about 700 genes involved in synthesis, modification, binding and degradation of glycan structures. Based on the glycosylated molecules and the biochemical process, four main groups of glycans are differentiated: the O-glycans, N-glycans, glycosaminoglycans (GAG) and glycolipids. In addition, domain-specific O- or C-linked glycosylation of various proteins takes place, i.e. O-fucosylation of Ser/Thr residues in EGF repeats [26].

O-glycosylation involves post-translational attachment of *N*-acetylgalactosamine (GalNAc) to OH-groups of serine or threonine residues in mucin-type glycoproteins. This is followed by a stepwise attachment of a limited number of sugar monomers (GlcNac, Gal, GalNAc, Fuc and *N*-acetylneuraminic acid) by various glycosyltransferases leading to specific oligosaccharide structures [5, 29]. Some of these structures are well-known tumour-associated epitopes in carcinomas, i.e. sialyl-Lewis X or sialy-Lewis A, Tn antigen, or serve as tumour markers for monitoring of cancer development, i.e. CA19-9.

During N-glycosylation, an endoplasmatic reticulum-bound oligosaccharide is first built by stepwise transfer of 2 GlcNAc, 9 mannose and 3 terminal Glc units to a membrane-bound dolichol anchor. This structure is then transferred to amino groups of asparagine residues of nascent proteins and further modified by branching, trimming, fucosylation, etc., leading to a wide variety of protein-bound oligosaccharide structures [29]. Both O- and N-glycans play an important role in adhesion of leukocytes to the endothelium during inflammation. According to a widely accepted model, the same glycan groups are involved in adhesion of tumour cells to endothelia during metastasis [25]. Here, the binding to selectins plays an important role.

Glycolipids are formed by covalent attachment of sugar residues to a ceramide in the external side of the outer cell membrane. According to their structure, glycolipids of the lacto-, ganglio-, neolacto- and globo-series can be differentiated. Several of these lipids are highly sialylated and are able to act as selectin ligands [33].

Glycosaminoglycans (GAGs) are characterized by long linear disaccharide chains attached to a protein backbone via a short linker oligosaccharide. They constitute a large fraction of the extracellular matrix and are crucial for cell motility and adhesion [1]. Depending on the nature of the disaccharide chains, the GAGs are divided into chondroitin, heparin, keratan, dermatan and hyaluronan, the latter lacking a protein backbone. GAGs are modified by addition of sulphate by various sulfotransferases (CHST's), which are frequently important for their proper function. Chondroitin sulphate has been reported to influence metastasis due to its binding affinity to P- and L-selectins [9, 22].

Although various reports suggest that glycosylation has an important impact on tumour growth and progression, the knowledge about the prognostic and/or predictive role of glycan structures and glycosylation enzymes in human carcinomas is still very limited. Therefore, we studied mRNA gene expression of all anabolic or catabolic glycosylation genes which are represented on HG-U133A microarray chips in two well-characterized breast cancer cohorts. These results add novel information about the glycosylation reactions in mammary carcinomas which influence progression and metastasis and might thus represent potential therapeutic targets.

Materials and methods

Patient cohorts

In a first screening approach, mRNA expression data obtained from primary breast cancer tissue samples from our hospital ($n = 194$) were analyzed ('Hamburg Cohort'). Patients were treated between 1991 and 2002 and selected on the basis of tissue availability. One hundred six patients received anthracyclin-based adjuvant chemotherapy regimens (mainly epirubicin/cyclophosphamide (EC) or cyclophosphamide/methotrexate/fluorouracil (CMF)). Seventy-four patients received endocrine therapy only, 8 patients were treated by radiation without any systemic therapy and 4 patients remained untreated after surgery (no information: 2 patients).

Informed consent for the scientific use of tissue materials, which was approved by the local ethics committees (for Hamburg: Ethik-Kommission der Ärztekammer Hamburg, #OB/V/03; for Mainz: Ethik-Kommission der Landesärztekammer Rheinland-Pfalz, #837.139.05 (4797)), was obtained from all patients. The study was performed in accordance to the principles of the declaration of Helsinki and REMARK criteria [23]. No radiotherapy, neoadjuvant chemotherapy or endocrine therapy had been administered before surgery.

Author's personal copy

Table 1 Glycosylation enzymes with independent prognostic impact in breast cancer patients

Gene symbol	Affymetrix no.	Gene product	Mean expression value (range)	Catabolic/ anabolic	Involvement in metabolism of glycan class	Hamburg cohort					Mainz cohort
						Multivariate Cox regression analysis (*p* value)		Cut-off for Kaplan–Meier analysis	Log-rank test (*p* value)		Cox regression analysis: DFS (*p* value)
						DFS	OAS		RFS	OAS	
FUT1	206109_at	Fucosyltransferase 1	167 (25–429)	Anabolic	O glycans/glycolipids	0.025	0.034	Q3	0.005	0.004	0.033
FUT6	210399_x_at	Fucosyltransferase 6	221 (24–705)	Anabolic	O glycans/glycolipids	0.029	0.064	Median	0.066	ns	ns
FUCA1	*202838_at*	*Fucosidase 1*	*1505 (381–5032)*	Catabolic	O glycans/glycolipids	*0.026*	*0.017*	Median	*0.007*	*0.005*	*0.006*
GCNT1	205505_at	Glucosaminyl (*N*-Acetyl) transferase 1	127 (12–776)	Anabolic	O glycans/glycolipids	0.035	0.038	Median	0.010	0.002	ns
GCNT4	220831_at	Glucosaminyl (*N*-Acetyl) Transferase 4	67 (7–207)	Anabolic	O glycans/glycolipids	0.043	ns	Q3	<0.001	<0.001	0.062
B3GNT4	221240_s_at	UDP-GlcNAc:betaGal beta-1,3-*N*-acetylglucosaminyltransferase	120 (23–469)	Anabolic	O glycans/glycolipids	0.013	0.004	Median	0.001	<0.001	ns
ST3GAL6	213355_at	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	94 (7–435)	Anabolic	O glycans/glycolipids	<0.001	0.008	Q3	0.037	0.074	ns
ST3GAL6	210942_s_at	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	140 (21–549)	Anabolic	O glycans/glycolipids	0.061	0.007	Q3	0.026	0.001	0.085
NAGA	202944_at	*N*-acetylgalactosaminidase, alpha	410 (108–948)	Catabolic	O glycans/glycolipids	0.050	0.021	Q1	0.004	0.004	ns
UGCG	221765_at	UDP-glucose ceramide glucosyltransferase	479 (22–2677)	Anabolic	glycolipids	ns	0.046	Q1	0.005	<0.001	0.065
POFUT1	*212349_at*	*Protein O-fucosyltransferase 1*	*261 (32–599)*	Anabolic	core fucosylation	*0.042*	*0.006*	Q1	*0.020*	*0.001*	*0.049*
GCS1	210627_s_at	Mannosyl-oligosaccharide glucosidase 1 (MOGS)	401 (175–799)	*Catabolic*	N-glycans	0.001	<0.001	Q3	0.003	<0.001	ns
GANAB	214626_s_at	Glucosidase, alpha	870 (330–1798)	Catabolic	N-glycans etc.	0.011	0.010	Median	0.006	0.002	ns
MAN1A1	*221760_at*	*Mannosidase 1A1*	*1095 (91–5117)*	*Catabolic*	N-glycans	*0.012*	*0.024*	Q3	*0.004*	*0.025*	*<0.001*
RPN1	201011_at	Ribophorin 1	1574 (552–4344)	Anabolic	N-glycans	0.024	0.053	Q1	0.013	0.018	0.003
DPM1	202673_at	Dolichyl-phosphate mannosyltransferase polypeptide 1, catalytic subunit	1523 (761–5424)	Anabolic	N-glycans	0.055	0.001	Q3	ns	0.031	<0.001
HAS2	206432_at	Hyaluronan synthase 2	71 (2–285)	Anabolic	GAG	ns	0.050	Q3	0.049	ns	ns
XYLT2	*219401_at*	*Xylosyltransferase 2*	*225 (32–888)*	*Anabolic*	*GAG*	*0.015*	*0.016*	*Q1*	*0.001*	*0.003*	*ns*
EXT1	201995_at	Exostosin glycosyltransferase 1	599 (245–1718)	Anabolic	GAG	ns	0.043	Q3	0.006	0.006	ns
GNS	*212334_at*	*Glucosamine (N-acetyl)-6-sulfatase*	*1756 (629–7759)*	*Catabolic*	GAG	*ns*	*0.047*	Q1	*0.094*	*0.006*	ns
HYAL1	210619_s_at	Hyaluronoglucosaminidase 1	76 (7–487)	Catabolic	GAG	0.016	0.002	Q1	ns	0.015	ns
CSGALNACT2	218871_x_at	Chondroitin sulphate *N*-acetylgalactosaminyltransferase 2	278 (36–786)	Anabolic	GAG	0.005	0.020	Q3	0.005	0.076	ns
CHST3	209834_at	Carbohydrate (chondroitin 6) sulfotransferase 3	351 (11–2075)	Anabolic	GAG	0.075	0.011	Q3	ns	0.020	ns
EXTL3	211051_s_at	Alpha-*N*-acetylglucosaminyltransferase II	113 (16–557)	Anabolic	GAG	0.049	ns	Median	0.032	0.089	ns
GALNS	206335_at	Galactosamine (*N*-acetyl)-6-sulphate sulfatase	473 (134–1630)	Catabolic	GAG	0.001	0.002	Q3	0.070	0.013	ns

P values <0.100 are given; genes which are associated with longer RFS/OAS are shown in italics

Author's personal copy

The second cohort consisted of 200 consecutive lymph node-negative breast cancer patients, treated at the Department of Obstetrics and Gynecology of the Johannes Gutenberg University, Mainz between 1988 and 1998 ('Mainz Cohort'; [32]). Patients did not receive any systemic therapy in the adjuvant or neoadjuvant setting, but treated either with modified radical mastectomy ($n = 75$) or with breast-conserving surgery followed by irradiation ($n = 125$), and did not show evidence of distant metastases at the time of surgery. The clinical and histological characteristics of both cohorts are given in Supplementary Table S1.

Analysis of mRNA expression data

Tissue samples were snap-frozen after surgery and stored in liquid nitrogen until use. RNA extraction, cDNA synthesis and microarray analysis were performed in the same laboratory for all samples as described [15]. Tumour cell content exceeded 40 % in all the samples, as shown by H&E staining of cryo-cut sections. The Affymetrix (Santa Clara, CA, USA) HG-U133A array and GeneChip System were used to quantify the relative transcript abundance in the breast cancer tissues. Hybridization intensity data were automatically acquired and processed by Affymetrix Microarray Suite 5.0 software. Arrays were analyzed using MAS5 algorithm. All microarray data have been submitted to Gene Expression Omnibus (GEO) under the following accession numbers. 'Hamburg-Cohort': GSE26971 (samples GSM663775-GSM663852), GSE31519 (samples GSM 782523-GSM782529), GSE31519 (samples GSM782554-GSM782568), GSE46184 (samples GSM1125783-GSM112 856); 'Mainz-Cohort': GSE11121.

Using previous publications and the GlycoGene database (http://riodb.ibase.aist.go.jp/rcmg/ggdb/), we first created a list of 241 anabolic and 60 catabolic glycosylation enzymes. Glycan-binding proteins or glycan transporters were not included in this study. Of these 301 genes, 202 (316 probesets) were represented on the Affymetrix HG-U133A array. A list of all relevant genes and probesets is given in Supplementary Table S2.

Statistics

In a first screening approach, we used univariate Cox regression analysis using continuous expression data and overall survival (OAS) or recurrence-free survival (RFS). For selected genes, the cohort was divided into quartiles according to their expression values, and Kaplan–Meier analyses with log-rank tests were performed with the most suitable cut-off (Q1 = lowest 25 % vs. higher 75 %, Q2 = median, Q3 = lowest 75 % vs. higher 25 %).

A multivariate Cox proportional hazards regression model was used to examine simultaneously the effects of multiple covariates on survival including nodal involvement (positive vs. negative), tumour stage (T1, T2, T3, T4), histological grading (G1/G2 vs. G3), oestrogen receptor (ER) status and the tested glycosylation enzyme. Associations with clinical and histological variables were calculated by χ^2 tests using the groups given above, and correlations with age by Pearson test using continuous data. All p values are two sided, and p values of <0.05 were considered to indicate a significant result. Due to the explorative character of our analysis, no multiple testing adjustment was considered. All analyses were performed using the SPSS 21 software.

Results

Enzymes of different glycosylation pathways are independent prognostic markers in breast cancer patients

In a first screening approach, the expression data of 316 probesets which are present on the HG-U133A genechip representing 155 anabolic and 47 catabolic glycosylation enzymes were extracted from the microarray data of 194 breast cancer samples from the Hamburg cohort. Probesets without or with only minimal expression in our samples (mean expression values below 60; $n = 64$) were excluded from further analysis (Supplementary Table S2).

With the remaining 252 probesets, a univariate Cox regression analysis was performed, using the continuous expression values. Interestingly, 71 probesets representing 64 genes were significantly associated with RFS, OAS or both. The respective p values are given in Supplementary Table S3. Since we were mainly interested in genes with independent prognostic value, we further performed multivariate Cox regression analysis including clinical stage, histological grading, nodal involvement, ER status and the respective expression values. Here, mRNA expression of 24 genes (25 probesets) turned out to be of independent prognostic significance in the Hamburg cohort. Among them were 16 anabolic and 8 catabolic glycosylation enzymes (Table 1).

In order to further evaluate the significance of these 24 genes, Kaplan–Meier analysis was performed. Again, log-rank tests resulted in significant differences in RFS or OAS or both (Table 1). For five genes (POFUT1, FUCA1, XYLT2, GNS and MAN1A1), high expression was associated with longer RFS and OAS, whereas expression of the other 19 genes correlated with a poor prognosis (Fig. 1).

Regarding the glycan classes, most of the prognostic genes in the Hamburg cohort are involved in synthesis,

Author's personal copy

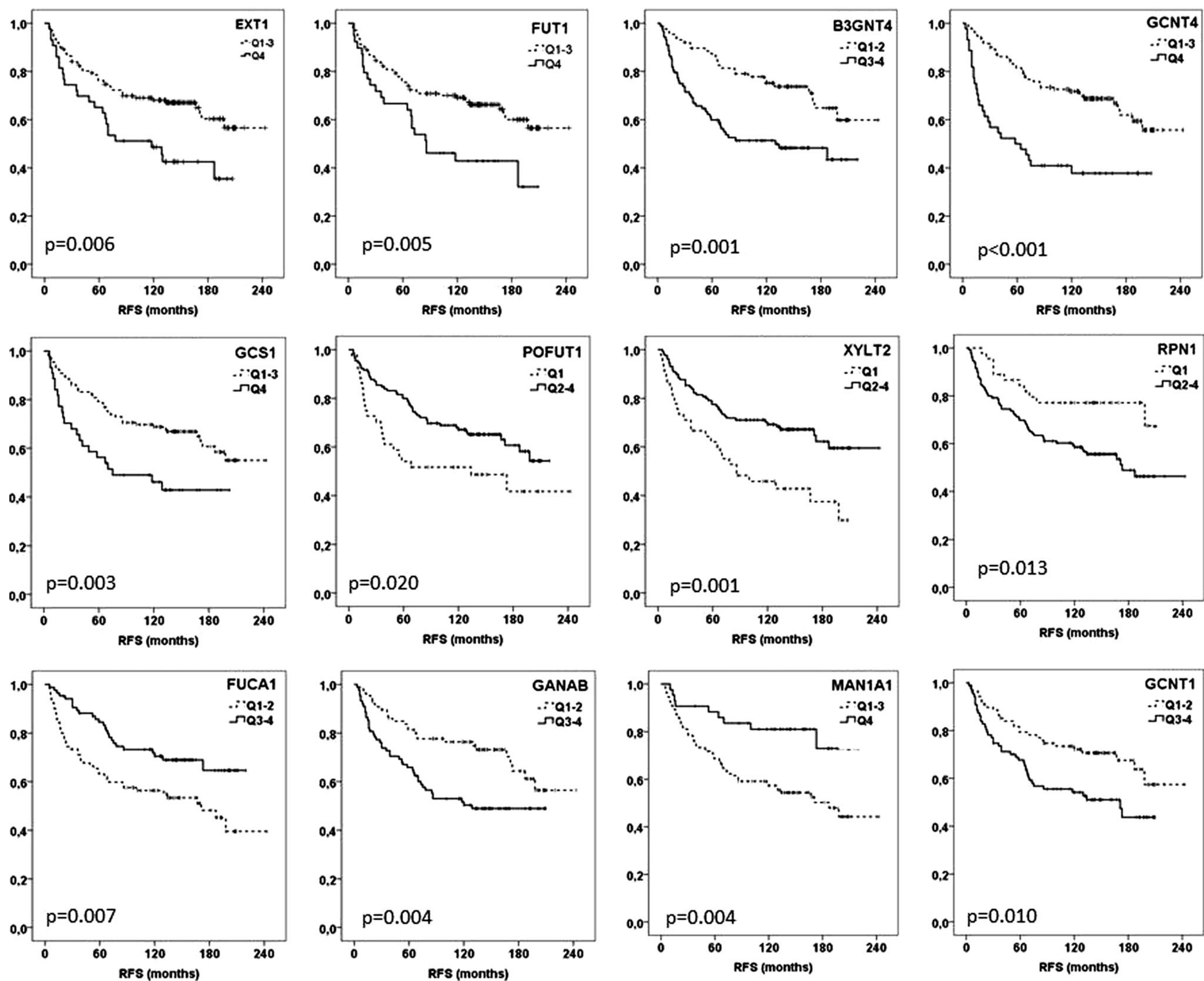

Fig. 1 Representative Kaplan–Meier curves showing the prognostic value of twelve glycosylation enzymes in 194 patients from our Hamburg cohort (see Table 1). *p* Values after log-rank tests are given. *Broken line* lower expression of the gene of interest; *solid line* higher expression

modification or degradation of GAGs ($n = 9$), seven genes are mainly involved in O-glycosylation, and five genes take part in N-glycosylation processes. The rest is involved in glycolipid synthesis ($n = 2$) or other pathways.

For further examination of the prognostic impact, expression data of the 24 relevant genes of our Hamburg cohort were obtained from an independent dataset of 200 breast cancer patients treated at the University of Mainz. In contrast to the first cohort, this group included only node-negative, mostly ER-positive patients which were treated by surgery and radiation only without any systemic treatment. Since only <10 % of the patients had died at the end of the observation period, we only included RFS in our analysis (Table 1). Here, six of the 24 relevant genes turned out as significant prognostic markers in Cox regression analysis (FUT1, FUCA1, POFUT1, MAN1A1, RPN1 and DPM1), whereas a trend was observed for three additional probesets (GCNT4, ST3GAL6 and UGCG). As observed in the Hamburg cohort, high expression of the fucosidase FUCA1, the mannosidase MAN1A1 and the protein O-fucosyltransferase POFUT1 was associated with longer RFS, whereas the other genes correlated with a worse prognosis.

Predictive value of glycosylation enzymes in chemotherapy-treated patients

In order to analyze, if the glycosylation enzymes were also predictive for therapy response, we performed Cox regression analyses in the subgroup of 100 patients who were treated with adjuvant chemotherapy. Out of the 24 prognostic genes, eighteen also showed significant associations

Author's personal copy

Table 2 Stratified Cox regression analysis for associations of glycosylation enzymes with RFS or OAS in patient subgroups

Affymetrix no.	Gene symbol[a]	Molecular subgroup (Hamburg + Mainz)			Chemotherapy-treated patients (Hamburg) ($n = 100$)		Involvement in metabolism of glycan class
		Luminal ($n = 278$) RFS (p value)	Her2 ($n = 50$) RFS (p value)	TNT ($n = 63$) RFS (p value)	RFS (p value)	OAS (p value)	
206109_at	FUT1	0.068	0.305	0.234	**0.032**	**0.025**	O glycans and glycolipids
210399_x_at	FUT6	0.816	0.454	0.692	0.082	0.055	O glycans and glycolipids
202838_at	*FUCA1*	*0.212*	*0.147*	***0.001***	***0.007***	***0.003***	*O glycans and glycolipids*
205505_at	GCNT1	0.964	0.902	**0.025**	0.070	0.068	O glycans and glycolipids
220831_at	GCNT4	0.798	0.458	0.798	**0.006**	**0.038**	O glycans and glycolipids
221240_s_at	B3GNT4	0.088	**0.042**	0.289	**0.008**	**0.001**	O glycans and glycolipids
213355_at	ST3GAL6	0.997	0.411	**0.036**	**0.002**	**0.005**	O glycans and glycolipids
210942_s_at	ST3GAL6	0.204	0.505	**0.033**	**0.001**	**<0.001**	O glycans and glycolipids
202944_at	NAGA	0.303	0.393	0.979	0.639	0.307	O glycans and glycolipids
221765_at	UGCG	0.849	**0.024**	0.497	0.551	**0.043**	Glycolipids
212349_at	*POFUT1*	***0.026***	*0.823*	*0.938*	***0.041***	***0.001***	*Core fucosylation*
210627_s_at	GCS1	**0.020**	0.150	0.090	**<0.001**	**<0.001**	N-glycans
211934_x_at	GANAB	0.065	0.077	0.233	**0.006**	**0.022**	N-glycans etc.
221760_at	*MAN1A1*	***<0.001***	*0.943*	*0.054*	***0.006***	***0.023***	*N-glycans*
201011_at	RPN1	**0.006**	0.474	0.156	**0.024**	**0.013**	N-glycans
202673_at	DPM1	**0.042**	0.121	0.418	0.095	**0.019**	N-glycans
206432_at	HAS2	0.527	0.239	0.761	0.613	0.119	GAG
219401_at	*XYLT2*	*0.845*	*0.420*	*0.586*	*0.164*	*0.278*	*GAG*
201995_at	EXT1	0.172	0.771	0.392	**0.008**	**0.001**	GAG
212334_at	GNS	0.937	**0.003**	0.083	0.061	0.063	GAG
210619_s_at	HYAL1	0.455	0.079	**0.032**	0.191	**0.037**	GAG
218871_x_at	CSGALNACT2	0.841	0.179	0.275	**0.026**	**0.009**	GAG
209834_at	CHST3	**0.022**	0.091	0.173	0.052	**0.012**	GAG
211051_s_at	EXTL3	0.784	0.403	0.204	**0.020**	**0.015**	GAG
206335_at	GALNS	0.584	0.241	0.698	**<0.001**	**0.004**	GAG

Significant correlations ($p < 0.050$) are shown in bold

[a] Genes which are associated with longer RFS/OAS are shown in italics

with RFS and/or OAS in this subcohort (Table 2). For some genes, i.e. EXT1, GCNT4, FUCA1, POFUT1 and EXTL3, the association with outcome was even more pronounced compared to the total cohort, suggesting a specific role in chemoresistance.

Stratified analysis of glycosylation enzymes in molecular subgroups of breast cancer

We further investigated if the expression and prognostic value of these enzymes varied between different molecular subgroups. First, we performed Pearson correlations in order to investigate associations with oestrogen receptor (ESR1), progesterone receptor (PGR1) and HER2 (ERBB2) mRNA expression levels. High ESR1 levels clearly correlated (correlation coefficient $\rho > 0.4$ or < -0.4) with low expression of the carbohydrate sulfotransferases CHST3, CHST6 and CHST11, and with high expression of the phosphatidyl GlcNAc transferase subunits PIGG, PIGH and PIGT. In addition, a positive correlation with the GalNAc transferase GALNT7 was found, which controls the initial step of O-glycosylation (Table 3). Using the same criteria for PGR1, only one correlation ($\rho < -0.4$ and >0.4) with the fucosyl transferase FUT3 was found, whereas ERBB2 expression did not correlate with the analyzed glycosylation genes. When ER-positive and ER-negative carcinomas were compared for expression of the analyzed glycosylation genes by t-test, 61 genes showed a positive association and 28 genes exhibited an inverse correlation with ER status as determined by immunohistochemistry (Supplementary Table S4).

In order to compare the prognostic value of the 24 relevant genes in luminal, HER2-positive and triple-negative carcinomas, both cohorts were combined and divided into

Author's personal copy

three molecular subgroups based on their ESR1 and HER2 mRNA expression as described [18, 24]: luminal tumours ($n = 278$), Her2-positive tumours ($n = 50$) and triple-negative carcinomas ($n = 63$). In a stratified Cox regression analysis, we found striking differences regarding the correlations with RFS. Whereas four enzymes of the N-glycosylation pathway (RPN1, DPM1, MAN1A1 and GCS1) as well as POFUT1 and CHST3 were highly prognostic in luminal tumours only, three genes mainly involved in metabolism of O-glycans and glycolipids (FUCA1, GCNT1 and ST3GAL6) and the hyaluronidase HYAL1 turned out to be exclusively prognostic in triple-negative tumours (Table 2). In HER2-positive cases, only the sulfatase GNS and the enzymes B3GNT4 and UGCG which are mainly involved in glycolipid synthesis were associated with RFS.

Correlations with distant metastasis

Several oligosaccharide structures play a role as ligands for selectins which represent important adhesion proteins of the endothelial cell layers, and thus facilitate extravasation of circulating tumour cells at the site of future metastasis. These selectin ligands are generally formed by O-glycosylation (i.e. sialyl Lewis X and sialyl Lewis A), but can also be attached to glycolipids, N-glycans or GAGs.

Therefore, we analyzed if the expression of the 24 prognostic genes correlated with distant metastases in general or metastasis to specific organ sites in the Hamburg cohort (Table 4). By χ^2 tests using the quartiles (see above), eleven genes were significantly associated with distant metastasis, three additional genes showed a non-significant association ($p =$ between 0.050 and 0.100). Regarding the sites of distant metastases, there were striking differences: cancers with relapse to bone were characterized by significantly higher GCNT4 and B3GNT4 and lower XYLT2 expression levels as compared to carcinomas metastasising to other sites (Table 4; Fig. 2). In contrast, tumours with lung metastases showed lower FUCA1 and UGCG expression and higher GANAB, CSGALNACT2 and B3GNT4 levels compared to the rest. Interestingly, tumours metastasizing to the brain showed significantly more hyaluronan synthase (HAS2), but also hyaluronidase (HYAL1) expression indicating an active hyaluronan metabolism in these cases. Significant associations are shown in Table 4 and representative correlations in Fig. 2.

Discussion

In the last years, an increasing interest in the biological role of glycosylation in general and especially in malignant progression of human cancer has developed. Most studies focused on the role of O-glycosylation and the formation of selectin ligands, due to the essential role of selectin binding in hematogenic metastasis [4, 16]. Although these studies suggested an impact of glycosylation on tumour growth and progression, the knowledge about the prognostic and/or predictive role of glycan structures in human carcinomas is still limited. This is mainly due to technical problems which hamper a sensitive and specific analysis of sugar residues in cancer tissues. The analysis of expression levels of glycosylation enzymes might be a suitable indirect approach to at least partially overcome these limitations.

Table 3 Correlations of glycosylation enzymes with ESR1 and PGR mRNA expression

		Affymetrix no.	$\rho =$ (Pearson test)
ESR1 (205225_at)	CHST3	32094_at	−0.449
	CHST3	209834_at	−0.460
	CHST6	221059_s_at	−0.463
	CHST11	219634_at	−0.535
	PIGG	218652_s_at	0.458
	PIGH	209625_at	0.412
	PIGT	217770_at	0.571
	GALNT7	218313_s_at	0.513
PGR (208305_at)	FUT3	214088_s_at	0.470

Only correlations with $\rho = <-0.40$ or >0.40 after Pearson correlation were shown

A recent study analyzed the mRNA expression of 419 genes involved in synthesis, degradation, modification and binding of glycans in breast cancer samples and breast tissue from healthy women [29] without reporting on the outcome of patients. The authors found a large number of genes which were deregulated in cancer specimens, which were involved not only in O-glycosylation, but also in the metabolism of N-glycans, GAGs and glycolipids.

In our investigation, we concentrated on the prognostic or predictive effect of glycan genes and their possible involvement in metastatic spread. Among the genes with independent prognostic significance, we found genes from all important glycosylation pathways, which is similar to prior results concerning genes with deregulated expression in breast cancer [29]. Due to technical reasons, it cannot be ruled out that the detected RNA partly originated from stromal cells. Yet, cellular models have previously shown the relevance of glycosylation enzymes in epithelial cancer cells themselves [7], and it is commonly acknowledged that the interaction between tumour and stroma in vivo is of relevance.

Among the 202 glycosylation genes included in our analysis, 64 (32 %) showed a significant prognostic impact in univariate Cox regression, and 24 thereof (12 %) retained prognostic significance in multivariate analysis. This high percentage of prognostic genes underlines the

Author's personal copy

Table 4 Correlations of glycosylation enzymes with metastasis to different organs (*p* values)

Gene symbol[a]	Affymetrix no.	Lymph node involvement ($n = 60$)	Distant metastasis ($n = 56$)	Bone metastasis ($n = 34$)	Lung metastasis ($n = 28$)	Visceral/ hepatis metastasis ($n = 30$)	Brain metastasis ($n = 15$)
FUT1	206109_at	ns	**0.001**	ns	ns	**0.001**	ns
FUCA1	*202838_at*	*ns*	***0.014***	*ns*	***<0.001***	*ns*	*ns*
GCNT1	205505_at	ns	**0.020**	ns	ns	ns	ns
GCNT4	220831_at	ns	**0.025**	**0.016**	ns	**0.036**	ns
B3GNT4	221240_s_at	**0.002**	**0.003**	**0.048**	**0.021**	ns	0.054
ST3GAL6	210942_s_at	ns	0.056	ns	ns	ns	ns
NAGA	202944_at	ns	**0.008**	ns	ns	ns	ns
UGCG	*221765_at*	*ns*	*0.077*	*ns*	***0.001***	*ns*	*0.092*
POFUT1	*212349_at*	*ns*	*ns*	*ns*	*ns*	***0.007***	***0.026***
GCS1	210627_s_at	ns	**0.007**	ns	ns	**0.008**	ns
GANAB	214626_s_at	ns	**0.022**	ns	**0.019**	**0.029**	**0.042**
RPN1	201011_at	ns	0.055	ns	ns	ns	ns
HAS2	206432_at	ns	ns	ns	ns	ns	**0.044**
XYLT2	*219401_at*	*ns*	***0.038***	***0.005***	*ns*	*ns*	*ns*
GNS	*212334_at*	***0.030***	ns	ns	ns	ns	ns
HYAL1	210619_s_at	ns	**0.019**	ns	ns	0.080	**0.003**
CSGALNACT2	218871_x_at	ns	**0.008**	ns	**0.035**	ns	ns
EXTL3	211051_s_at	**0.004**	ns	0.053	ns	ns	ns

Significant correlations ($p < 0.050$) are shown in bold; inverse correlations are given in italics

ns no significant association ($p > 0.10$)

[a] Only genes with p values <0.100 are shown

important role of glycosylation processes in tumour progression. The list of 24 prognostic genes in the first cohort included not only enzymes of glycan synthesis, but also catabolic genes involved in glycan trimming and degradation. For example, high expression of the fucosyltransferases FUT1 and FUT6 is associated with poor prognosis, whereas overexpression of the fucosidase FUCA1 correlated with relatively good outcome.

Several, but not all of the 24 prognostic genes have been implicated in tumour progression in varying experimental settings: FUT1 and FUT6 are involved in the synthesis of Lewis antigens which are important ligands for selectins [21]. In several cancer entities, these enzymes were shown to increase resistance to apoptosis [11], adhesion [17], tumour growth [12] and metastasis [16]. The fucosidase FUCA1 cleaves off fucose residues, thus degrading Lewis antigens. In experimental systems, it reduces the invasive potential of breast cancer cells and their adhesion to selectins and endothelial cells under flow conditions [39, 40].

GCNT1 (C2GNT1) and GCNT4 (C2GNT3) are glucosaminyl (*N*-acetyl) transferases involved in formation of core 2 structures which serve as carriers of Lewis antigens. In bladder cancer cells, these branched glycans have an influence on immune evasion and metastasis [35]. Additionally, GCNT1 promotes aggressiveness of germ cell tumours [13]. In leukaemia and oral squamous carcinoma cells, GCNT1 contributes to synthesis of selectin ligands [19, 30]. This is also true for the sialyl transferase ST3GAL6 which is of prognostic relevance in our cohort. Similar to other sialyl transferases, ST3GAL6 has been shown to promote migration and metastasis of various cancer cells [6, 28].

In addition to enzymes related to the formation of selectin ligands, we found an independent prognostic value of key enzymes of N-glycosylation. This was most prominent in luminal (hormone-receptor positive) carcinomas, whereas the former genes were more important in triple-negative tumours. N-glycosylation is important for various signalling processes and influences cancer progression in complex ways [20]. DPM1 synthesizes the mannose donor molecule, dolichol-phosphat mannose (Dol-P-Man). Ribophorin (RPN1), as part of the oligosaccharidyltransferase (OST) complex, catalyzes transfer of a mannose-containing oligosaccharide from the dolichol carrier to nascent polypeptides, and mannosidase (MAN1A1), glucosidase alpha (GANAB) and glucosidase 1 (GCS1) are involved in trimming and maturation of N-linked oligosaccharides. Interestingly, GCS1 and GANAB turned out as unfavourable prognostic indicators in breast cancer, whereas high MAN1A1 expression was associated with a better prognosis. The reason underlying this effect remains to be

Author's personal copy

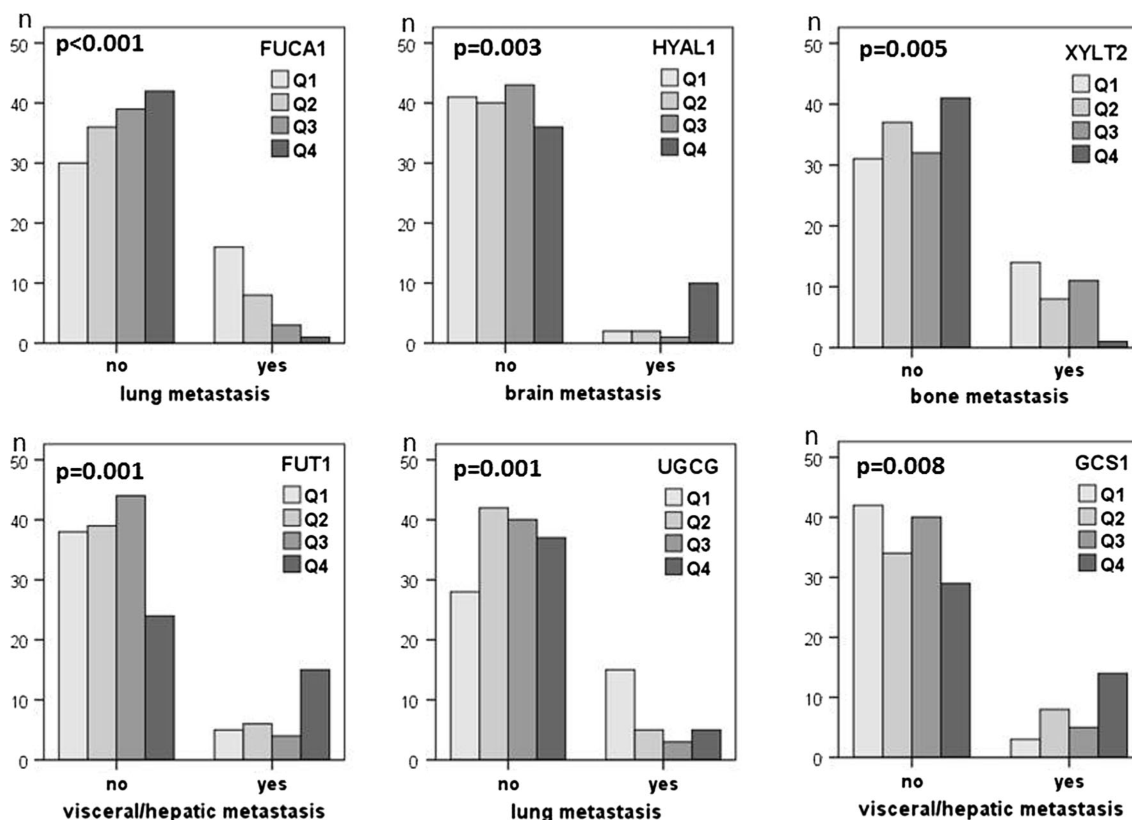

Fig. 2 Significant associations of selected glycosylation genes with the localization of metastasis (see Table 4). *p* values after χ^2 tests are given. Cases were divided into four groups of equal size (quartiles Q1–Q4) depending on expression values of the respective gene, and correlated with the presence of metastasis at specific sites. *Y-axis* number of tumours

analyzed. In vitro, mannosidase treatment of MCF7 breast cancer cells inhibited the formation of multicellular foci [2].

Interestingly, several enzymes involved in the formation of proteoglycans were also relevant in the Hamburg cohort. Strikingly, high expressions of the hyaluronan synthase HAS2 and the hyaluronidase HYAL1 were both associated with poor prognosis and metastasis to the brain. Hyaluronan is an independent prognostic indicator in breast cancer [3], and inhibition of HAS2 blocks metastases in animal models [37]. For HYAL1, both tumour suppressor and oncogenic activities were reported [10, 34]. In our study, a significant association of HYAL1 expression with early recurrence was only observed in triple-negative carcinomas which points to molecular-subtype specific differences.

Another enzyme involved in the metabolism of proteoglycans is CHST3 which catalyzes sulfation of chondroitin, resulting in the formation of selectin ligands [9]. In contrast, the sulfatase GNS which leads to breakdown of heparan and keratan sulphate is associated with a good prognosis in our cohort.

An unexpected result of our investigation is the association of high POFUT1 expression with a good prognosis in both cohorts. POFUT1 attaches fucose through an O-glycosidic linkage to conserved serine or threonine residues in EGF domains. This fucosylation has been reported to play a crucial role in Notch signalling [21], which has been implicated in cancer metastasis [14]. Yet, EGF domains are also found in other proteins, and the role of O-fucosylation in these cases is poorly understood. How POFUT1 and some other genes (UGCG, NAGA, GANAB, EXTL2, etc.) affect the malignant behaviour of breast cancer cells should be further analyzed in experimental models.

Compared to our results within the first (Hamburg) cohort, not all of the 24 glycosylation enzymes of interest were also prognostic in the Mainz cohort. This might result from discrepancies in patients' characteristics, since all tumours from the latter cohort were node-negative, mostly of early stage and not treated with adjuvant therapy. In contrast, the Hamburg cohort included chemotherapy-treated tumours, and some of the observed associations might be also explained by a predictive role of the examined enzymes and a role of these in therapy response.

A strength of our study is the information about the anatomical sites of metastasis and the availability of a cohort of untreated patients that allows to examine the

Author's personal copy

prognostic role of the enzymes. A drawback is the cohort size that does not allow to examine accurately all subgroups of patients of potential clinical relevance and the effect of different systemic therapies in detail. In addition, only glycosylation genes which were represented on the HG-U133A microarrays were included in this study, and the prognostic significance of various additional genes remains to be investigated. Nevertheless, the high number of 202 anabolic or catabolic genes (316 probesets) which were analyzed in this study included most important enzymes of all glycosylation pathways.

Taken together, we found that enzymes from all glycosylation pathways might play an independent prognostic role in breast cancer, with differences between molecular subtypes. This indicates an impact of those pathways on breast cancer progression and metastasis and thus might be interesting in search for suitable therapeutic targets. How changes in glycosylation enzymes influence the glycan structures and which protein or lipid structures are affected on the surfaces of cancer cells remains to be investigated.

Acknowledgments We are grateful to Katrin Beck, Sylke Krenkel, Maila Rossberg and Kathrin Eylmann for excellent technical assistance.

Conflict of interest R.M. Wirtz is an employee of a diagnostic company. The other authors have no competing interests.

References

1. Afratis N, Gialeli C, Nikitovic D, Tsegenidis T, Karousou E, Theocharis AD, Pavao MS, Tzanakakis GN, Karamanos NK (2012) Glycosaminoglycans: key players in cancer cell biology and treatment. FEBS J 279:1177–1197. doi:10.1111/j.1742-4658.2012.08529.x
2. Arcaro KF, Wang J, Otis CN, Zuckerman BM (2004) Beta-galactosidase and alpha-mannosidase inhibit formation of multicellular nodules in breast cancer cell cultures. Anticancer Res 24:139–144
3. Auvinen P, Tammi R, Parkkinen J, Tammi M, Agren U, Johansson R, Hirvikoski P, Eskelinen M, Kosma VM (2000) Hyaluronan in peritumoral stroma and malignant cells associates with breast cancer spreading and predicts survival. Am J Pathol 156:529–536. doi:10.1016/S0002-9440(10)64757-8
4. Brockhausen I (2006) Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions. EMBO Rep 7: 599–604
5. Brockhausen I (1999) Pathways of O-glycan biosynthesis in cancer cells. Biochim Biophys Acta 1473:67–95
6. Carvalho AS, Harduin-Lepers A, Magalhaes A, Machado E, Mendes N, Costa LT, Matthiesen R, Almeida R, Costa J, Reis CA (2010) Differential expression of alpha-2,3-sialyltransferases and alpha-1,3/4-fucosyltransferases regulates the levels of sialyl Lewis a and sialyl Lewis x in gastrointestinal carcinoma cells. Int J Biochem Cell Biol 42:80–89. doi:10.1016/j.biocel.2009.09.010
7. Cazet A, Julien S, Bobowski M, Burchell J, Delannoy P (2010) Tumour-associated carbohydrate antigens in breast cancer. Breast Cancer Res BCR 12:204. doi:10.1186/bcr2577
8. Contessa JN, Bhojani MS, Freeze HH, Ross BD, Rehemtulla A, Lawrence TS (2010) Molecular imaging of N-linked glycosylation suggests glycan biosynthesis is a novel target for cancer therapy. Clin Cancer Res 16:3205–3214. doi:10.1158/1078-0432.CCR-09-3331
9. Cooney CA, Jousheghany F, Yao-Borengasser A, Phanavanh B, Gomes T, Kieber-Emmons AM, Siegel ER, Suva LJ, Ferrone S, Kieber-Emmons T, Monzavi-Karbassi B (2011) Chondroitin sulfates play a major role in breast cancer metastasis: a role for CSPG4 and CHST11 gene expression in forming surface P-selectin ligands in aggressive breast cancer cells. Breast Cancer Res BCR 13:R58. doi:10.1186/bcr2895
10. Gotte M, Yip GW (2006) Heparanase, hyaluronan, and CD44 in cancers: a breast carcinoma perspective. Cancer Res 66: 10233–10237. doi:10.1158/0008-5472.CAN-06-1464
11. Goupille C, Marionneau S, Bureau V, Hallouin F, Meichenin M, Rocher J, Le Pendu J (2000) Alpha1,2fucosyltransferase increases resistance to apoptosis of rat colon carcinoma cells. Glycobiology 10:375–382
12. Guo Q, Guo B, Wang Y, Wu J, Jiang W, Zhao S, Qiao S, Wu Y (2012) Functional analysis of alpha1,3/4-fucosyltransferase VI in human hepatocellular carcinoma cells. Biochem Biophys Res Commun 417:311–317. doi:10.1016/j.bbrc.2011.11.106
13. Hatakeyama S, Kyan A, Yamamoto H, Okamoto A, Sugiyama N, Suzuki Y, Yoneyama T, Hashimoto Y, Koie T, Yamada S, Saito H, Arai Y, Fukuda M, Ohyama C (2010) Core 2 *N*-acetylglucosaminyltransferase-1 expression induces aggressive potential of testicular germ cell tumor. Int J Cancer 127:1052–1059. doi:10.1002/ijc.25117
14. Hu YY, Zheng MH, Zhang R, Liang YM, Han H (2012) Notch signaling pathway and cancer metastasis. Adv Exp Med Biol 727:186–198. doi:10.1007/978-1-4614-0899-4_14
15. Ihnen M, Muller V, Wirtz RM, Schroder C, Krenkel S, Witzel I, Lisboa BW, Janicke F, Milde-Langosch K (2008) Predictive impact of activated leukocyte cell adhesion molecule (ALCAM/CD166) in breast cancer. Breast Cancer Res Treat 112(3):419–427
16. Julien S, Ivetic A, Grigoriadis A, QiZe D, Burford B, Sproviero D, Picco G, Gillett C, Papp SL, Schaffer L, Tutt A, Taylor-Papadimitriou J, Pinder SE, Burchell JM (2011) Selectin ligand sialyl-Lewis x antigen drives metastasis of hormone-dependent breast cancers. Cancer Res 71:7683–7693. doi:10.1158/0008-5472.CAN-11-1139
17. Kanoh A, Ota M, Narimatsu H, Irimura T (2003) Expression levels of FUT6 gene transfected into human colon carcinoma cells switch two sialyl-Lewis X-related carbohydrate antigens with distinct properties in cell adhesion. Biochem Biophys Res Commun 303:896–901
18. Karn T, Metzler D, Ruckhaberle E, Hanker L, Gatje R, Solbach C, Ahr A, Schmidt M, Holtrich U, Kaufmann M, Rody A (2010) Data-driven derivation of cutoffs from a pool of 3,030 Affymetrix arrays to stratify distinct clinical types of breast cancer. Breast Cancer Res Treat 120:567–579. doi:10.1007/s10549-009-0416-z
19. Kikuchi J, Ozaki H, Nonomura C, Shinohara H, Iguchi S, Nojiri H, Hamada H, Kiuchi A, Nakamura M (2005) Transfection of antisense core 2 beta1,6-*N*-acetylglucosaminyltransferase-1 cDNA suppresses selectin ligand expression and tissue infiltration of B-cell precursor leukemia cells. Leukemia 19:1934–1940. doi:10.1038/sj.leu.2403951
20. Lau KS, Dennis JW (2008) N-Glycans in cancer progression. Glycobiology 18:750–760. doi:10.1093/glycob/cwn071
21. Listinsky JJ, Siegal GP, Listinsky CM (2011) The emerging importance of alpha-L-fucose in human breast cancer: a review. Am J Transl Res 3:292–322
22. Martinez P, Vergoten G, Colomb F, Bobowski M, Steenackers A, Carpentier M, Allain F, Delannoy P, Julien S (2013) Over-sulfated glycosaminoglycans are alternative selectin ligands:

Author's personal copy

insights into molecular interactions and possible role in breast cancer metastasis. Clin Exp Metastasis. doi:10.1007/s10585-013-9592-7

23. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM (2006) REporting recommendations for tumor MARKer prognostic studies (REMARK). Breast Cancer Res Treat 100:229–235
24. Milde-Langosch K, Karn T, Muller V, Witzel I, Rody A, Schmidt M, Wirtz RM (2013) Validity of the proliferation markers Ki67, TOP2A, and RacGAP1 in molecular subgroups of breast cancer. Breast Cancer Res Treat 137:57–67. doi:10.1007/s10549-012-2296-x
25. Miles FL, Pruitt FL, van Golen KL, Cooper CR (2008) Stepping out of the flow: capillary extravasation in cancer metastasis. Clin Exp Metastasis 25:305–324
26. Moremen KW, Tiemeyer M, Nairn AV (2012) Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol 13:448–462. doi:10.1038/nrm3383
27. Nairn AV, Aoki K, dela Rosa M, Porterfield M, Lim JM, Kulik M, Pierce JM, Wells L, Dalton S, Tiemeyer M, Moremen KW (2012) Regulation of glycan structures in murine embryonic stem cells: combined transcript profiling of glycan-related genes and glycan structural analysis. J Biol Chem 287:37835–37856. doi:10.1074/jbc.M112.405233
28. Perez-Garay M, Arteta B, Llop E, Cobler L, Pages L, Ortiz R, Ferri MJ, de Bolos C, Figueras J, de Llorens R, Vidal-Vanaclocha F, Peracaula R (2013) Alpha2,3-Sialyltransferase ST3Gal IV promotes migration and metastasis in pancreatic adenocarcinoma cells and tends to be highly expressed in pancreatic adenocarcinoma tissues. Int J Biochem Cell Biol 45:1748–1757. doi:10.1016/j.biocel.2013.05.015
29. Potapenko IO, Haakensen VD, Luders T, Helland A, Bukholm I, Sorlie T, Kristensen VN, Lingjaerde OC, Borresen-Dale AL (2010) Glycan gene expression signatures distinguish normal and malignant breast tissue; possible role in diagnosis and progression. EJC Suppl 8:102
30. Renkonen J, Rabina J, Mattila P, Grenman R, Renkonen R (2001) Core 2 beta1,6-*N*-acetylglucosaminyltransferases and alpha1,3-fucosyltransferases regulate the synthesis of O-glycans on selectin ligands on oral cavity carcinoma cells. APMIS Acta Pathol Microbiol Immunol Scand 109:500–506
31. Roberts JD, Klein JL, Palmantier R, Dhume ST, George MD, Olden K (1998) The role of protein glycosylation inhibitors in the prevention of metastasis and therapy of cancer. Cancer Detect Prev 22:455–462
32. Schmidt M, Bohm D, von Torne C, Steiner E, Puhl A, Pilch H, Lehr HA, Hengstler JG, Kolbl H, Gehrmann M (2008) The humoral immune system has a key prognostic impact in node-negative breast cancer. Cancer Res 68:5405–5413
33. Shirure VS, Henson KA, Schnaar RL, Nimrichter L, Burdick MM (2011) Gangliosides expressed on breast cancer cells are E-selectin ligands. Biochem Biophys Res Commun 406:423–429. doi:10.1016/j.bbrc.2011.02.061
34. Stern R (2005) Hyaluronan metabolism: a major paradox in cancer biology. Pathol Biol 53:372–382. doi:10.1016/j.patbio.2004.12.021
35. Suzuki Y, Sutoh M, Hatakeyama S, Mori K, Yamamoto H, Koie T, Saitoh H, Yamaya K, Funyu T, Habuchi T, Arai Y, Fukuda M, Ohyama C, Tsuboi S (2012) MUC1 carrying core 2 O-glycans functions as a molecular shield against NK cell attack, promoting bladder tumor metastasis. Int J Oncol 40:1831–1838. doi:10.3892/ijo.2012.1411
36. Taylor-Papadimitriou J, Epenetos AA (1994) Exploiting altered glycosylation patterns in cancer: progress and challenges in diagnosis and therapy. Trends Biotechnol 12:227–233. doi:10.1016/0167-7799(94)90121-X
37. Udabage L, Brownlee GR, Waltham M, Blick T, Walker EC, Heldin P, Nilsson SK, Thompson EW, Brown TJ (2005) Antisense-mediated suppression of hyaluronan synthase 2 inhibits the tumorigenesis and progression of breast cancer. Cancer Res 65:6139–6150. doi:10.1158/0008-5472.CAN-04-1622
38. Varki A, Lowe JB (2009) Biological roles of glycans. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (eds) Essentials of glycobiology. Cold Spring Harbor, New York
39. Yuan K, Kucik D, Singh RK, Listinsky CM, Listinsky JJ, Siegal GP (2008) Alterations in human breast cancer adhesion-motility in response to changes in cell surface glycoproteins displaying alpha-L-fucose moieties. Int J Oncol 32:797–807
40. Yuan K, Listinsky CM, Singh RK, Listinsky JJ, Siegal GP (2008) Cell surface associated alpha-L-fucose moieties modulate human breast cancer neoplastic progression. Pathol Oncol Res POR 14:145–156. doi:10.1007/s12253-008-9036-x