Supplementary table 1: list of microarray probes considered for statistical analysis.

		Chin dataset		van de Vijver dataset		present study	
Pathways	Gene Title	GeneBank	Affy ID	GeneBank	UGID (build #204)	GeneBank	Probe set Name
Initiation of O-glycosylation	GALNT1	NM_020474	A.201722_s_at	NM_020474	Hs.514806	NM_020474	NM_020474.2_psr1_at
	GALNT2	AL525086	A.217787_s_at	NM_004481	Hs.654649	NM_004481	NM_004481.2_psr1_at
	GALNT3	BF063271	A.203397_s_at	NM_004482	Hs.170986	NM_004482	NM_004482.3_psr1_at
	GALNT4	NM_003774	A.220442_at	NM_003774	Hs.25130	NM_003774	NM_003774.3_psr1_at
	GALNT5			AA534406	Hs.269027	NM_014568	NM_014568.1_psr1_at
	GALNT6	NM_007210	A.219956_at	NM_007210	Hs.505575	NM_007210	NM_007210.3_psr1_s_at
	GALNT7	NM_017423	A.218313_s_at	NM_017423	Hs.548088	NM_017423	NM_017423.2_psr1_at
	GALNT8	NM_017417	A.220929_at	NM_017417	Hs.511985	NM_017417	NM_017417.1_psr1_at
	GALNT9					NM_021808	NM_021808.2_psr1_at
	GALNT10	BE906572	A.212256_at	NM_017540	Hs.655011	NM_017540	NM_017540.3_psr1_a_at
	GALNT11	NM_022087	A.219013_at	AI985274	Hs.647109	NM_022087	NM_022087.2_psr1_at
	GALNT12	NM_024642	A.218885_s_at	AI630798	Hs.47099	NM_024642	NM_024642.3_psr1_at
	GALNT13					NM_052917	NM_052917.2_psr1_at
	GALNT14	NM_024572	A.219271_at	AA165698	Hs.468058	NM_024572	NM_024572.2_psr1_s_at
	C1GALT1	NM_020156	A.219439_at	NM_020156	Hs.592180	uc003sqz	uc003sqz.1_psr1_a_at
	GCNT1	NM_001490	A.205505_at	NM_001490	Hs.521568	NM_001097634	NM_001097634.1_psr1_s_at
	GCNT3	NM_004751	A.219508_at	NM_004751	Hs.194710	NM_004751	NM_004751.1_psr1_at
	GCNT4	NM_016591	A.220831_at	NM_016591	Hs.272404	NM_016591	NM_016591.1_psr1_at
	ST3GAL1	NM_003033	A.208322_s_at	NM_003033	Hs.584803	NM_003033	NM_003033.2_psr1_s_at
	ST6GALNAC1			Y11339	Hs.105352	NM_018414	NM_018414.3_psr1_at
	ST6GALNAC2	NM_006456	A.204542_at	NM_006456	Hs.592105	NM_006456	NM_006456.1_psr1_at
Elongation with type 2 lactosamnic chain	B4GALT1	D29805	A.201883_s_at	NM_001497	Hs.651277	NM_001497	NM_001497.2_psr1_at
	B4GALT2	BC002431	A.209413_at	NM_003780	Hs.632403	NM_003780	NM_003780.3_psr1_at
	B4GALT3	AF038661	A.210243_s_at	NM_003779	Hs.321231	NM_003779	NM_003779.2_psr1_a_at
	B3GALT4	BC004523	A.210540_s_at	NM_003778	Hs.13225	NM_003782	NM_003782.3_psr1_at
	B3GNT1	NM_006876	A.203188_at	NM_006876	Hs.8526	NM_006876	NM_006876.2_psr1_at
	B3GNT2	NM_006577	A.219326_s_at	NM_006577	Hs.173203	NM_006577	NM_006577.5_psr1_at
	B3GNT7			AK000770	Hs.299329	NM_145236	NM_145236.1_psr1_s_at
	GCNT2	L19659	A.211020_at	AI654230	Hs.519884	NM_145649	NM_145649.3_psr1_s_at
	FUT3	AW080549	A.214088_s_at	NM_000149	Hs.169238	NM_001097641	NM_001097641.1_psr1_at
	FUT5	U27329	A.211225_at	NM_002034	Hs.631843	NM_002034	NM_002034.2_psr1_at
	FUT6	U27335	A.211465_x_at	NM_000150	Hs.631846	NM_001040701	NM_001040701.1_psr1_x_at
	FUT7	AA767713	A.217696_at	NM_004479	Hs.457	NM_004479	NM_004479.2_psr1_at
	ST3GAL3			NM_006279	Hs.597915	NM_006279	NM_006279.2_psr1_a_at
	ST3GAL4	NM_006278	A.203759_at	NM_006278	Hs.591947	NM_006278	NM_006278.1_psr1_at
	CHST2	NM_004267	A.203921_at	NM_004267	Hs.8786	NM_004267	NM_004267.3_psr1_at
	CHST4	NM_005769	A.220446_s_at	NM_005769	Hs.251383	NM_005769	NM_005769.1_psr1_s_at
Synthesis of sphyngolipids	UGCG	NM_003358	A.204881_s_at	NM_003358	Hs.304249	NM_003358	NM_003358.1_psr1_at
	B4GALT5	NM_004775	A.206232_s_at	NM_004776	Hs.370487	NM_004776	NM_004776.3_psr1_at
	B4GALT6	NM_004776	A.221485_at	NM_004775	Hs.591063	NM_004775	NM_004775.2_psr1_at
	B3GNT5			AI825936	Hs.208267	NM_032047	NM_032047.4_psr1_at
	FUT4	AF305083	A.209892_at	NM_002033	Hs.390420	NM_002033	NM_002033.2_psr1_at
	ST3GAL6	AB022918	A.210942_s_at	NM_006100	Hs.148716	NM_006100	NM_006100.2_psr1_s_at
Synthesis of Heparan sulfate	XYLT1	AI693140	A.213725_x_at	AW009370	Hs.610023	NM_022166	NM_022166.3_psr1_x_at
	XYLT2	NM_022167	A.219401_at			NM_022167	NM_022167.2_psr1_at
	B4GALT7	AK022566	A.222191_s_at	NM_007255	Hs.455109	NM_007255	NM_007255.1_psr1_s_at
	B3GALT6			AI201722	Hs.284284	NM_080605	NM_080605.3_psr1_at
	B3GAT1	NM_018644	A.219521_at	AB029396	Hs.381050	NM_018644	NM_018644.3_psr1_at
	B3GAT2	NM_012200	A.203452_at	AI954745	Hs.653102	NM_080742	NM_080742.2_psr1_at
	B3GAT3	AB009598	A.35179_at	NM_012200	Hs.502759	NM_012200	NM_012200.2_psr1_a_at
	EXTL1	NM_004455	A.206329_at	NM_004455	Hs.150956	NM_004455	NM_004455.2_psr1_at
	EXTL2	AF000416	A.209537_at	AF000416	Hs.357637	NM_001033025	NM_001033025.1_psr1_at
	EXTL3	AF001690	A.209202_s_at	NM_001440	Hs.491354	NM_001440	NM_001440.2_psr1_at
	EXT1	NM_000127	A.201995_at	NM_000127	Hs.492618	NM_000127	NM_000127.2_psr1_at
	EXT2	AA196245	A.202012_s_at	NM_000401	Hs.368404	NM_207122	NM_207122.1_psr1_at
	NDST1	AL526632	A.202607_at	NM_001543	Hs.222055	NM_001543	NM_001543.3_psr1_at
	NDST2	NM_003635	A.203916_at	NM_003635	Hs.654758	NM_003635	NM_003635.2_psr1_a_at
	NDST3	NM_004784	A.220429_at	NM_004784	Hs.480596	NM_004784	NM_004784.1_psr1_at
	NDST4	NM_022569	A.208334_at			NM_022569	NM_022569.1_psr1_at
	HS2ST1	NM_012262	A.203284_s_at	NM_012262	Hs.48823	NM_012262	NM_012262.2_psr1_at
	HS3ST1	NM_005114	A.205466_s_at	NM_005114	Hs.507348	NM_005114	NM_005114.2_psr1_at
	HS3ST2	NM_006043	A.219697_at	NM_006043	Hs.115830	NM_006043	NM_006043.1_psr1_at
	HS3ST3A1	NM_006042	A.219985_at	NM_006042	Hs.462270	NM_006042	NM_006042.1_psr1_at
	HS3ST3B1	NM_006041	A.221062_at	AI042497	Hs.48384	NM_006041	NM_006041.1_psr1_at
	HS3ST4			AW204561	Hs.655275	NM_006040	NM_006040.2_psr1_at
	HS3ST5					NM_153612	NM_153612.2_psr1_a_at
	HS3ST6					NM_001009606	NM_001009606.2_psr1_at
	HS6ST1	NM_004807	A.206997_s_at	AI632014	Hs.512841	NM_004807	NM_004807.2_psr1_at
	HS6ST2			AI767756	Hs.385956	NM_147175	NM_147175.3_psr1_at
	HS6ST3			H15259	Hs.171001	NM_153456	NM_153456.2_psr1_at

Supplementary table 2: primers used in qRT-PCR

Gene	Forward Primer (5'-3')	Reverse primer (5'-3')
ACTB *	CCCTCCATCGTCCACCGCAAATGCTTC	CGACTGCTGTCACCTTCACCGTTCCAG
RPLP0†	CAGCATCTACAACCCTGAAGTG	GTGTAATCCGTCTCCACAGACA
PUM1†	GATTATTCAGGCACGCAGGT	AGCAGCGCTGATGATGTATG
GCNT1 †	AGAAGGATACACAAAACGTACC	ACCTTTCTAGCTAACTGTGCTC
ST6GALNAC2 †	GGCTTCACCTCCGTGCCACAAG	CTAGGCCCTCAGGGACAGGCAC
FUT3 †	CCTGCTGGAGTCCTTTGTGGCC	GCAGGCAAGTCTTCTGGAGGGG
FUT4 *	CTCTTTGAGCCTCTGTCACCTGAAC	CTTTCTCTGATTCCTGGTTTTCTGCC
FUT5 †	AATCTAGGTACCAGACGGTGCGC	AGGTAGGTGAGGCCCTGGGAAAG
FUT6 ‡	CAAAGCCACATCGCATTGAA	ATCCCCGTTGCAGAACCA
FUT7 *	GAGAGCCGATACCAACGCTTCTTTG	ACCAACCCTCAAGGTCCTCATAGAC
ST3GAL3 *	TACCCTTGGCAGTGTGGCAGTGACC	AGACTCCTGTGTAGGTGGTGCTGGC
ST3GAL4 *	TCTCTTCTACCCTGAATCTGCCCAC	ACCCTTTCGCACCCGCTTCTTATCAC
ST3GAL6 †	AGAGTCCTTTGCACTACTATGG	CACTGTTAGCATCATCTTCTGAG
EXT1 †	TGGCTACATGCCGCTGATCCAC	CCCCCACTCAGCCGGATTCCTC
HS3ST1 †	CCGCAGACCATCATCATCGGCG	AGCCCAAGCCGTGGCTGTAATG

* Previously used in Julien et al. (1)
† Designed for this study using Primers3Plus opensource software (2)
‡ Previously used in Barthel et al. (3)

1. Julien S, Grimshaw MJ, Sutton-Smith M, et al. Sialyl-Lewisx on P-selectin glycoprotein ligand-1 is regulated during differentiation and maturation of dendritic cells: A mechanism involving the glycosyltransferases C2GnT1 and ST3Gal I. Journal of Immunology 2007;179:5701-10.
2. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JA Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 2007;35:W71-4.
3. Barthel SR, Gavino JD, Wiese GK, Jaynes JM, Siddiqui J, Dimitroff CJ Analysis of glycosyltransferase expression in metastatic prostate cancer cells capable of rolling activity on microvascular endothelial (E)-selectin. Glycobiology 2008;18:806-17.

Supplementary table 3: Correlation of glyco-genes expression with ER status or distant metastasis (extended version).

		ER status *			Distant metastasis †	
	Datasets ‡	**A**	**B**	**C**	**A**	**B**
	ERpos n=	75	226	8	DMpos n= 87	194
	ERneg n=	43	69	10	DMneg n= 31	101
Pathways §	Gene Title	p values ||			p values	
Initiation of O-glycosylation and Core synthesis	GALNT2**		0.0087			*0.089*
	GALNT3**	0.0045	0.0013			
	GALNT4			*0.0866*		
	GALNT5	NA ‡‡	*0.0196*			
	GALNT6	<0.0001	<0.0001			
	GALNT7	0.0009	<0.0001			<0.01
	GALNT10	*0.0687*	<0.0001			
	GALNT11				<0.05	<0.05
	GALNT12**		*0.0505*			
	GALNT14		<0.0001			<0.01
	C1GALT1**	0.0009	0.0005			
	GCNT1**		<0.0001			<0.05
	ST3GAL1				*0.069*	
	ST6GALNAC			*0.0542*		
Elongation with type 2 lactosamine chain	B4GALT3		*0.1255*			
	B3GALT4					
	B3GNT1**	0.0162	0.0004			*0.078*
	B3GNT7	NA	<0.0001			
	GCNT2**		<0.0001			
	FUT3**		<0.0001			<0.05
	FUT5		0.0031			
	FUT6				*0.059*	
	FUT7		<0.0001			
	CHST2	<0.0001	<0.0001			
	CHST4		*0.0969*			
Synthesis of sphyngolipids	UGCG**	<0.0001	<0.0001		<0.01	*0.063*
	B4GALT5	0.0216	0.0027	0.0158		<0.01
	B3GNT5**	NA	<0.0001	0.0415		
	FUT4	0.0019	<0.0001			*0.064*
	ST3GAL6**	*0.1275*	0.0186			
Synthesis of Heparan sulfate	B3GAT1**		0.0793			<0.05
	B3GAT3				<0.05	<0.05
	EXTL1				<0.05	*0.077*
	EXT1	<0.0001	0.0343	0.0221		<0.05
	HS3ST1		0.0002			*0.070*
	HS3ST3B1		*0.0674*			
	HS6ST3	NA	0.0402			

* ER: estrogen receptor alpha. Glyco-genes significantly more expressed in ER-negative tumours than ER-positive ones are shaded. Glycogenes significantly less expressed are left on white background. Blank: no statistical significance found.
† Glyco-genes significantly more expressed in metastatic tumours (DMpos) than in non metastatic ones (DMneg) are shaded. Glycogenes significantly less expressed are left on white background. Blank: no statistical significance found.
‡ A: Chin et al (24), B: van de Viver et al. (25), C: present study. Number of cases for ER and DM groups are indicated accordingly.
§Pathways are referring to Figure 1
|| statistical significance was tested with the unpaired Student's t-test. All p values were two tailed. For ER status, multiple testing correction was applied using the Westfall and Young permutation is a step-down procedure (1993, Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment, Volume 1). For distant metastasis, only the original p values obtained by simple Student test are reported, and can be considered as a trend. p<0.05 was considered as statistically significant. Borderline p values are indicated in italic.

** Glycogenes varying accordingly to ER status in each dataset but only reaching statistical significance in one or two of them.
‡‡ NA: Not available, there was no probe targeting this glycogene on the microarray chip used (supplementary table 1).