Coexpression of β1,6-*N*-Acetylglucosaminyltransferase V Glycoprotein Substrates Defines Aggressive Breast Cancers with Poor Outcome

Summar F. Siddiqui,[1,3] John Pawelek,[2,3] Tamara Handerson,[4] Chen-Yong Lin,[5] Robert B. Dickson,[5] David L. Rimm,[1,3] and Robert L. Camp[1,3]

Departments of [1]Pathology and [2]Dermatology, and the [3]Yale Cancer Center, Yale University School of Medicine, New Haven, Connecticut; [4]Department of Pathology, University of Massachusetts School of Medicine, Worcester, Massachusetts; and [5]Departments of Oncology and Pathology and the Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, District of Columbia

Abstract

β1,6-*N*-Acetylglucosaminyltransferase-V (GnT-V) catalyzes the addition of complex oligosaccharide side chains to glycoproteins, regulating the expression and function of several proteins involved in tumor metastasis. We analyzed the expression of five cell-surface glycoprotein substrates of GnT-V, matriptase, β1-integrin, epidermal growth factor receptor, lamp-1, and N-cadherin, on a tissue microarray cohort of 670 breast carcinomas with 30-year follow-up. *Phaseolus vulgaris* leukocytic phytohemagglutinin (LPHA), a lectin specific for β1,6-branched oligosaccharides, was used to assay GnT-V activity. Our results show a high degree of correlation of the LPHA staining with matriptase, lamp-1, and N-cadherin expressions, but not with epidermal growth factor receptor or β1-integrin expressions. In addition, many of the GnT-V substrate proteins exhibited strong coassociations. Elevated levels of GnT-V substrates were correlated with various markers of tumor progression, including positive node status, large tumor size, estrogen receptor negativity, HER2/neu overexpression, and high nuclear grade. Furthermore, LPHA and matriptase showed significant association with disease-related survival. Unsupervised hierarchical clustering of the GnT-V substrate protein expression and LPHA revealed two distinct clusters: one with higher expression of all markers and poor patient outcome and one with lower expression and good outcome. These clusters showed independent prognostic value for disease-related survival when compared with traditional markers of tumor progression. Our results indicate that GnT-V substrate proteins represent a unique subset of coexpressed tumor markers associated with aggressive disease. (Cancer Epidemiol Biomarkers Prev 2005;14(11):2517–23)

Introduction

Previous studies have shown that malignant transformation is associated with alterations in protein glycosylation. β1,6-*N*-Acetylglucosaminyltransferase V (GnT-V), an enzyme regulated by the ras/raf pathway, catalyzes the transfer of β1,6 branches onto the complex oligosaccharide side chains of cell-surface proteins (1-4). Increased expression of GnT-V alters sugar residues on specific glycoproteins, which affects their expression and/or function, leading to enhanced tumor invasion and cancer progression (2, 5-8).

There are several known GnT-V substrates with diverse roles in cell adhesion, migration, and proliferation, including matriptase, lamp-1, N-cadherin, β1-integrin, and epidermal growth factor receptor (EGFR). Matriptase is an epithelial-derived serine protease that regulates cell motility and growth via the hepatocyte growth factor pathway (9, 10). Lamp-1 is expressed on cell and lysosome membranes and is important in lysosomal trafficking, matrix degradation, and cell adhesion (11, 12). N-Cadherin is a calcium-dependent glycoprotein that mediates cell-cell adhesion and cell motility (10, 13). The β1-integrin subunit associates with several different α-subunits forming a heterodimeric protein involved in cell-cell and cell-matrix binding (14). EGFR is a tyrosine kinase receptor which binds to several EGF-related growth factors promoting cell proliferation (15, 16).

Studies using GnT-V–transfected cell lines have analyzed the effect of GnT-V overexpression on cell-surface glycoprotein expression and function. The expression of both matriptase and lamp-1 is enhanced by GnT-V glycosylation where the addition of β1,6 branches prevents the degradation of both proteins (12, 17). In contrast, overexpression of GnT-V has not been shown to alter the expression of EGFR, β1-integrin, or N-cadherin (18). However, GnT-V–mediated glycosylation is also known to alter the activities of GnT-V substrates in various ways as detailed in Discussion.

The plant lectin leukocytic phytohemagglutinin (LPHA) exhibits high specificity for β1,6 branching on *N*-glycans and can be used to assess GnT-V functionality on formalin-fixed, paraffin-embedded tumors (4, 19). Several prior studies have suggested a role for increased β1,6 branched oligosaccharides in cancer progression using LPHA (19-21). We have recently completed a study on LPHA, which showed that breast cancer metastases exhibited significantly higher LPHA staining than matched primary tumors ($P < 0.0001$), and found that LPHA expression in primary tumors was an independent predictor of poor outcome (22).

In this study, we evaluate whether the relationships between GnT-V and its substrates, as observed in cell lines, are seen in primary patient samples. We assess the ability of these associations to predict tumor progression in a cohort of 670 breast cancer samples with 30-year patient follow-up. Expressions of LPHA, matriptase, lamp-1, N-cadherin, β1-integrin, and EGFR were assessed using the automated quantitative analysis (AQUA) system, which provides a quantitative assessment of immunohistochemical stains on a

Received 6/23/05; revised 8/2/05; accepted 9/13/05.

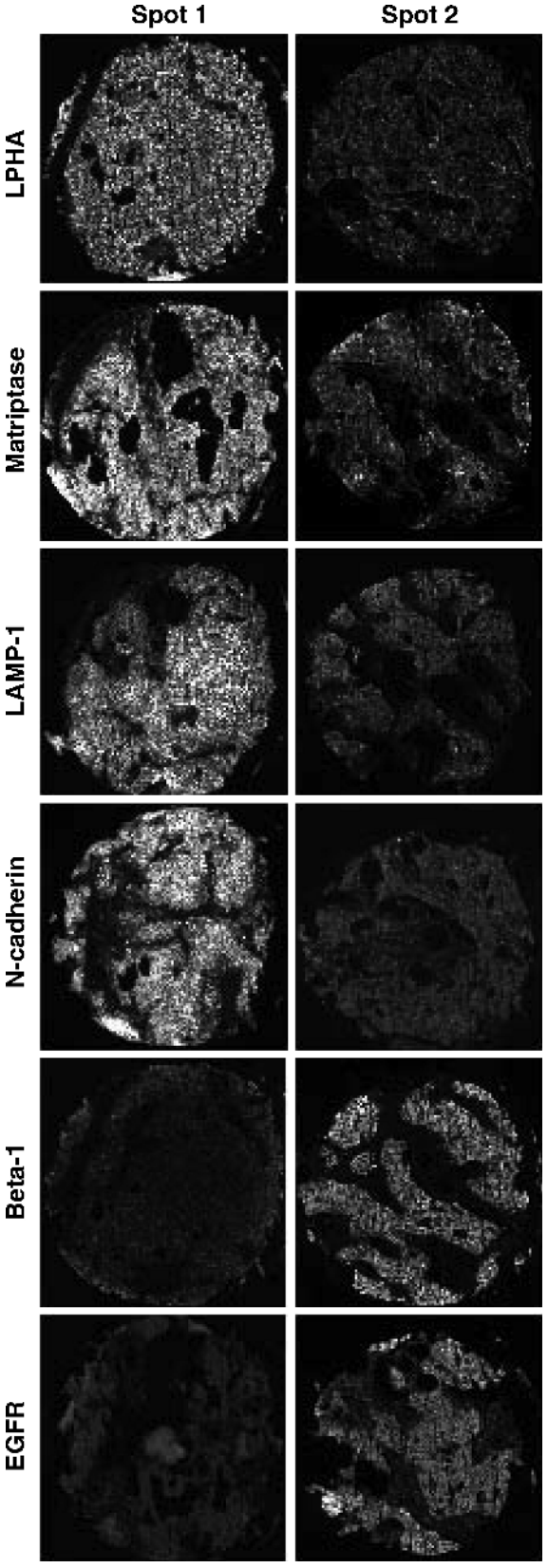
Grant support: Breast Cancer Alliance (R.L. Camp and D.L. Rimm), NIH grant K08 ES11571 (R.L. Camp), NIH grant R21 CA 100825 (D.L. Rimm), NIH grant R33 CA 106709 (D.L. Rimm), NIH grant R01 CA 096851 (R.B. Dickson). U.S. Army grants DAMD-17-02-0463 and DAMD17-02-1-0634 (D.L. Rimm), and Vion Pharmaceuticals (J. Pawelek).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Robert L. Camp, Department of Pathology, Yale University School of Medicine, 310 Cedar Street, P.O. Box 208023, New Haven, CT 06520-8023. Phone: 203-785-6340; 203-737-5089. E-mail: robert.camp@yale.edu

doi:10.1158/1055-9965.EPI-05-0464

continuous scale (23, 24). Our study shows significant correlations between LPHA staining and the expression of several GnT-V substrate proteins. Combinatorial analysis using unsupervised hierarchical clustering reveals subsets where all GnT-V substrates and LPHA trend higher or where they trend lower, defining tumors with significantly different outcomes.



Materials and Methods

Tissue Microarray Cohort. Tissue microarrays consisting of 670 separate breast carcinomas (50% node positive) were constructed using previously described techniques (25, 26). Tumors consisted of 74% ductal and 26% lobular carcinomas. Mean patient follow-up time was 152 months. Formalin-fixed, paraffin-embedded blocks were obtained from the Yale University Department of Pathology archives and patient follow-up information was obtained from the Yale Tumor Registry under protocols outlined by the Yale Human Investigation Committee. Complete treatment information was not available for the entire cohort; however, most patients were treated with postsurgical local irradiation. None of the node-negative patients were given adjuvant systemic therapy. Among the node-positive patients, ~15% were given chemotherapy and ~27% were given tamoxifen (post-1978). Slides from all blocks were reviewed by a pathologist (R.L.C.) to select representative areas of invasive tumor—away from normal tissue or *in situ* carcinoma—to be cored and placed on the tissue microarray using a Tissue Microarrayer (Beecher Instruments, Silver Spring, MD). After construction, the tissue microarrays were cut into 5-μm sections, placed on glass slides using adhesive tape-transfer system (Instumedics, Inc., Hackensack, NJ), and UV cross-linked. Antibody and lectin staining was done on two separate microarrays and the results were averaged.

Immunohistochemistry. Slides were deparaffinized in xylene and rehydrated through graded ethanol and endogenous peroxidases were blocked with 3% hydrogen peroxide in methanol. Heat-induced antigen retrieval was done by pressure cooking in citrate buffer (pH 6.0) for 15 minutes. Slides were blocked with 0.3% bovine serum albumin in TBS. To create a tumor mask for all the markers, the slides were incubated for 1 hour at room temperature with rabbit anti–pan-cytokeratin (1:100; DakoCytomation, Carpinteria, CA) for N-cadherin and EGFR or with mouse anti–pan-cytokeratin (AE1/AE3, 1:100; DakoCytomation) for lamp-1, β1-integrin, matriptase, and LPHA. The slides were washed with TBS containing 0.05% Tween 20 and incubated with corresponding secondary antibodies, Envision antimouse or antirabbit (DakoCytomation). Slides were then visualized with Cy3-Tyramide (Perkin-Elmer, Wellesley, MA) for 10 minutes. Slides were then transferred to tap water for 10 minutes and re-pressure cooked in citrate buffer (pH 6) for 15 minutes to remove previously deposited antibody. Slides were reblocked with 0.3% bovine serum albumin in TBS and new primary antibodies were applied to all the slides and incubated overnight with the exception of LPHA, which was incubated at room temperature for 1 hour. Monoclonal anti-matriptase antibody was prepared as previously described (10) and used at 1:1,000 dilution. The rest of the antibodies were commercially obtained and used at the following dilutions: biotinylated LPHA (1:8,000; Vector Laboratories, Burlingame, CA), mouse anti–N-cadherin (1:500; BD Biosciences, San Jose, CA), mouse anti-human β1-integrin (1:300; Chemicon, Temecula, CA), rabbit anti-EGFR (1:200; Cell Signaling Technology, Beverly, MA), and mouse anti-human lamp-1 (1:200; BD Biosciences). The specificities of the antibodies were established

Figure 1. Immunohistochemical staining of LPHA and GnT-V substrates for two typical histospots are shown (×10). The first series of histospot (*left*) is typical of a tumor that coexpresses LPHA, matriptase, lamp-1, and N-cadherin. The second (*right*) shows a tumor with positive staining for β1-integrin and EGFR but weak staining for the remaining proteins. Images were analyzed using AQUA and only the signal that colocalizes with the tumor membrane and cytoplasm was used for further analysis.

Table 1. Pairwise correlations of GnT-V substrate proteins and LPHA

Marker	LPHA	Matriptase	EGFR	Lamp-1	β1-Integrin	N-Cadherin
LPHA		**<0.0001**	0.4032	**0.0037**	0.4798	**<0.0001**
Matriptase	**<0.0001**		**0.0006**	**<0.0001**	**<0.0001**	**<0.0001**
EGFR	0.4032	**0.0006**		**0.0003**	**<0.0001**	**<0.0001**
Lamp-1	**0.0037**	**<0.0001**	**0.0003**		**<0.0001**	0.5273
β1-Integrin	0.4798	**<0.0001**	**<0.0001**	**<0.0001**		0.9789
N-Cadherin	**<0.0001**	**<0.0001**	**<0.0001**	0.5273	0.9789	

NOTE: *P* values are computed from the Pearson correlation. Statistically significant values are in boldface. The Bonferroni adjustment corrects for α inflation due to the simultaneous analysis of multiple tests. Using this correction, biomarkers with $P < 0.0033$ are considered to be statistically significant ($\alpha = 0.05$).

Table 3. Continuous univariate and optimal cutpoint analysis

Marker	Continuous univariate *P*	Relative risk
LPHA	**0.016**	**1.7**
Matriptase	**0.035**	**1.3**
EGFR	0.340	1.5
Lamp-1	0.062	1.2
β1-Integrin	0.270	1.3
N-Cadherin	0.088	1.2

NOTE: Univariate *P* values are based on the Cox model using each marker independently as a continuous marker. Relative risks were calculated at the optimal AQUA score (LPHA, >45.7; matriptase, >52.4; EGFR, >49.5; lamp-1, >99.6; β1-integrin, >65.4; and N-cadherin, >27.6). Statistically significant values are in boldface.

by individual vendors using Western blots and/or immunoprecipitation. As a control for staining specificity, LPHA was preincubated with porcine thyroglobulin (7.5 μmol/L; Sigma, St. Louis, MO) before addition to the tissue microarray. Mouse anti-matriptase monoclonal antibody has been previously characterized (10, 27). Corresponding secondary antibodies were applied for 1 hour at room temperature, Envision antimouse or antirabbit (DakoCytomation). The LPHA slide was incubated with streptavadin-horseradish peroxidase (1:100; Perkin-Elmer). To visualize the nuclei, 4′,6-diamidino-2-phenylindole was included in the secondary antibody mixture (Molecular Probes, Eugene, OR). Tumor cells were distinguished from stroma using Alexa 488–tagged anti-cytokeratin antibodies. The fluorescent chromogen Cy5-Tyramide (Perkin-Elmer) was used for target identification because its far red emission spectra are outside the range of tissue autofluorescence.

Automated Quantitative Analysis. The PM-1000 system consisting of Olympus AX-51 epifluorescence microscope with a Prior microscope stage and digital camera using IPLabs (Scanalytics, Inc., Fairfax, VA) was used to acquire high resolution (1,024 × 1,024 pixel; 0.5 μm resolution) monochromactic images of each histospot. Images were analyzed using the AQUA system as previously described (23, 24, 28-30). In brief, AQUA reconstructs pseudocolor images of histospots based on imaging of the fluorescent signals (i.e., 4′,6-diamidino-2-phenylindole–positive nuclei, Alexa 488-cytokerain, and Cy5 target marker). AQUA automatically distinguishes tumor from stroma based on cytokeratin expression and separates tumor cytoplasm/membrane from nuclei based on 4′,6-diamidino-2-phenylindole expression in cytokeratin-positive cells. Histocores were taken from tumor areas with no contaminating (cytokeratin-positive) normal epithelium. For this study, target signals from the membrane and cytoplasm of keratin-positive tumor cells were used for analysis. Target intensities are expressed as AQUA scores (total target intensity divided by the total area of tumor cytoplasm or membrane). This process is fully automated with no required user interface to define regions of interest. Histospots with <5% of tumor were excluded from further analysis. Image results were reviewed by one researcher (S.S.) to ensure the validity of the automated process.

Statistical Analysis. The JMP v5.0.1 software (SAS Institute, Inc., Cary, NC) was used for statistical and survival analyses. Correlations between continuous variables were measured using the Pearson correlation; associations between mixed (continuous versus nominal data) were assessed using *t* tests; and those between nominal variables were analyzed using the χ^2 test. Kaplan-Meier curves were used to show patient survival over time with statistical significance determined using the log-rank test. Univariate and multivariate survival analyses were done using the Cox proportional hazards model. Unsupervised hierarchical clustering was done using Cluster and Treeview (Eisen Laboratory, Stanford University, Palo Alto, CA; ref. 31). In brief, cases were randomly assigned into ''training'' and ''validation'' cohorts ($n = 314$ and $n = 315$, respectively). Clustering was done on tumors in the training set which had a value for at least five of six biomarkers ($n = 290$). Values were converted into *z* scores before clustering to normalize between markers (23, 32). The centroid for each cluster was calculated based on the average biomarker values of each tumor in the cluster. The geometric distance from each tumor in the validation set to the centroid of each training cluster was determined and tumors were assigned to the closest cluster. Reassignment of the training set resulted in 85.2% of these tumors being reclassified into the same cluster.

Results

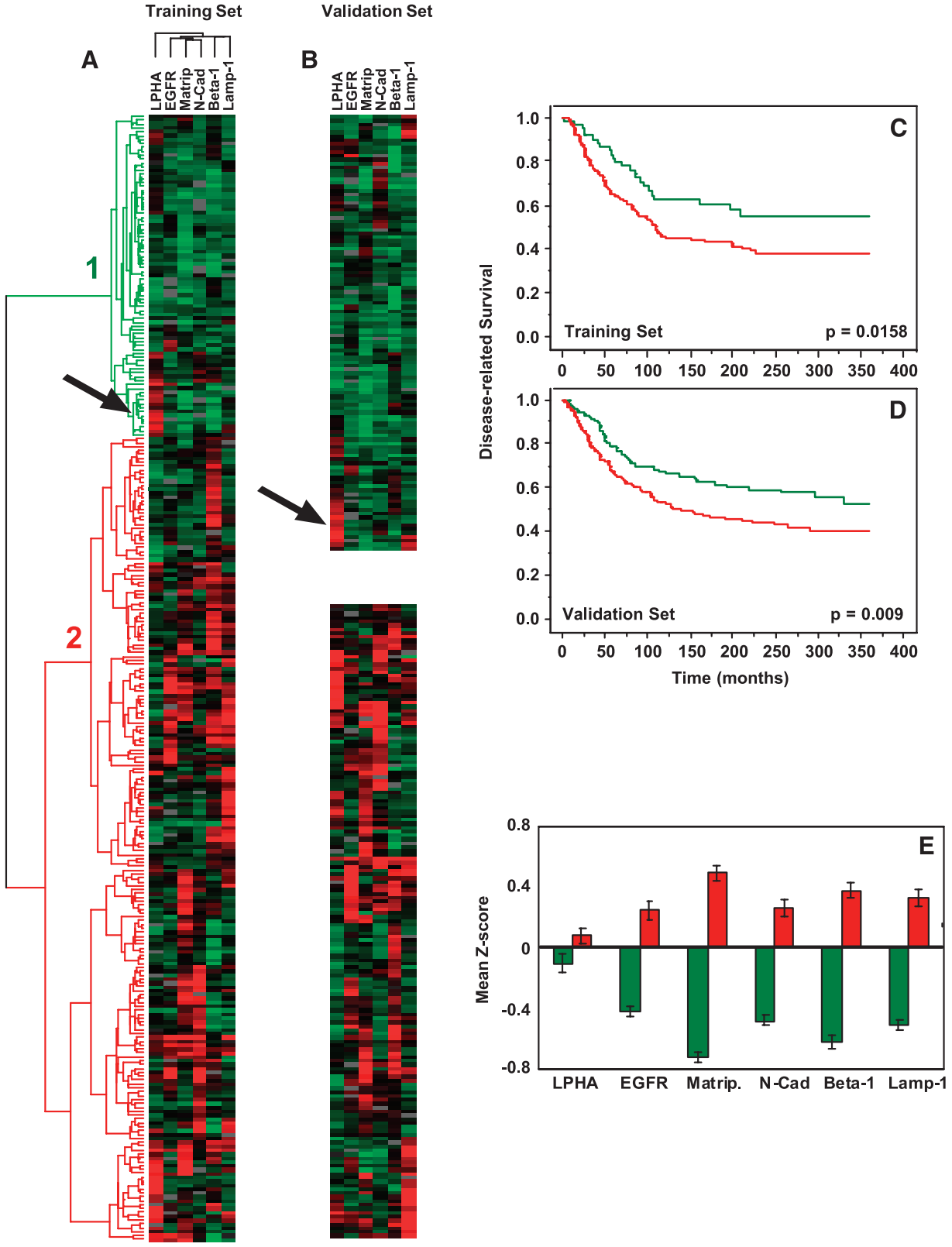
GnT-V Substrate Proteins Are Coassociated. We did a quantitative automated analysis of the expression of several known GnT-V substrate proteins (Fig. 1). The LPHA lectin was

Table 2. *T* tests of GnT-V substrate proteins, LPHA, and traditional breast cancer markers

Marker	LPHA		Matriptase		EGFR		Lamp-1		β1-Integrin		N-Cadherin	
	Δ	*P*	Δ	*P*	Δ	*P*	Δ	*P*	Δ	*P*	Δ	*P*
Node positive	↓ 4	0.8015	↑ 56	0.055	↑ 27	0.1438	↑ **156**	**<0.0001**	↑ 55	**0.0488**	↑ 31	0.0939
Size > 2 cm	↑ 7	0.6526	↑ **86**	**0.0033**	↑ 32	0.0814	↑ **113**	**0.0026**	↓ 31	0.2281	↑ **38**	**0.0451**
High nuclear grade	↑ 7	0.7043	↑ **156**	**<0.0001**	↑ 1	0.9502	↓ 10	0.8170	↓ 28	0.3857	↑ **68**	**0.0009**
HER2 3+	↑ 42	0.1135	↑ **181**	**0.0001**	↓ 34	0.2484	↓ 21	0.7219	↑ **147**	**0.0001**	↑ 38	0.2045
Estrogen receptor negative	↓ 1	0.9324	↑ **70**	**0.0180**	↓ 29	0.1206	↓ 4	0.9222	↓ 24	0.4032	↑ 26	0.1562
Age > 50 y	↑ 6	0.7284	↓ 29	0.3808	↑ 5	0.7987	↑ 58	0.1546	↑ 34	0.2716	↓ 8	0.6952
Lobular	↓ 40	0.1417	↓ **182**	**<0.0001**	↓ **72**	**0.0145**	↓ 93	0.1417	↓ 58	0.1818	↓ **67**	**0.0217**

NOTE: Δ represents the mean difference between tumors fulfilling the traditional marker phenotype (left column) and those that do not (e.g., node-positive tumors have a lamp-1 AQUA score that is 156 points higher than node-negative tumors; $P < 0.0001$). Statistically significant values are in boldface.

used to assess the degree of β1,6-branched glycosylation and served as a marker for GnT-V activity. We found a high degree of correlation between GnT-V substrate proteins and LPHA, as assessed by comparing the continuous expression of each marker using the Pearson correlation (Table 1). In particular, LPHA expression was correlated with lamp-1 ($P = 0.0037$), matriptase ($P < 0.0001$), and N-cadherin ($P < 0.0001$). There was no association between LPHA and β1-integrin ($P = 0.4798$) or EGFR ($P = 0.4032$; Table 1). Despite the lack of association between LPHA and EGFR or β1-integrin, these proteins were highly correlated with the expression of most of other GnT-V substrates. Taken together, these findings suggest a high degree of correlation between GnT-V substrates and LPHA in human breast carcinoma.



GnT-V Substrate Protein Expression Correlates with Traditional Breast Cancer Markers. We compared the expression of LPHA and the GnT-V substrates with traditional markers of breast cancer using a *t*-test analysis (Table 2). In general, each of the proteins correlated with at least one marker of aggressive breast cancer. Matriptase, in particular, correlated with most aggressive markers (large tumor size, high nuclear grade, HER2 overexpression, and estrogen receptor negativity; Table 2). N-Cadherin, β1-integrin, lamp-1, EGFR, and LPHA associated with one or two of the known markers of aggressive breast cancer (Table 2). Lamp-1 was the best predictor of node status ($P < 0.0001$), exhibiting high level expression on node-positive versus node-negative tumors (mean AQUA score difference of 156). In all cases, overexpression of a particular GnT-V substrate protein associated with aggressive disease. In addition, the expression levels of matriptase, EGFR, and N-cadherin were all reduced on lobular versus ductal carcinomas. Previous studies have shown reduced EGFR expression on lobular carcinomas (33) but we believe that this is the first time it has been seen with matriptase and N-cadherin.

LPHA and Matriptase Are Predictive of Patient Outcome. Next, we determined if LPHA and the GnT-V substrate proteins were prognostic for disease-related survival. Using continuous univariate analysis, we found that high levels of LPHA and matriptase were predictive of poor outcome ($P = 0.016$ and 0.035, respectively; Table 3). Previously, we have shown that both LPHA and matriptase are independent prognostic indicators in breast cancer when assessed using chromogenic stain and manual (visual) inspection (22, 34). Although our current study validates these findings, when assessed in an automated, quantitative fashion, neither LPHA nor matriptase exhibited a robust relative risk (1.7 and 1.3, respectively; Table 3). Furthermore, neither LPHA nor matriptase retained independent prognostic value when analyzed as a continuous variable and compared with traditional histopathologic markers in a multivariate analysis (data not shown). The reason for this discrepancy is unclear but may represent differences in the cohorts analyzed or bias differences between manual and automated scoring. All of the GnT-V substrates trended in the same direction (overexpression equating with poor outcome) and lamp-1 and N-cadherin were close to achieving statistical significance ($P = 0.062$ and $P = 0.088$, respectively; Table 3). These data suggest that GnT-V substrate proteins and GnT-V activity (as measured by LPHA) tend to be overexpressed on poor-prognosis tumors; however, this association is generally weak.

Clustering Shows the Coexpression of GnT-V Substrates Is Associated with Aggressive Disease. Although we showed above that, as a group, breast cancers tend to coexpress pairs of GnT-V substrate proteins, we were interested in determining how individual tumors would self-associate using unsupervised hierarchical clustering (Fig. 2). A randomly selected training set was clustered and two distinct clusters were identified: one with low expression of all GnT-V substrates and LPHA and one with higher and more heterogeneous expression (Fig. 2, *clusters 1* and *2*, respectively). When the 30-year disease-related survivals of these subsets were compared, the low-expressing population (cluster 1) exhibited a significantly better outcome ($P = 0.0158$). Additional subdivisions of these clusters did not define tumors with significantly different survival rates. Tumors in a separate validation set were then assigned to clusters based on their distance to the centroid of the preestablished clusters. As with the training set, the low-expressing tumors exhibited significantly better outcome ($P = 0.009$). *T* tests of marker expression between the clusters show that all markers contribute significantly to the clustering analysis (Fig. 2, $P < 0.0001$). Interestingly, a small subset of cluster 1 tumors exhibits high LPHA expression in both the training and validation cohorts (Fig. 2, *arrow*). However, this subset has no significant difference in outcome when compared with other cluster 1 tumors ($P = 0.6674$).

We then assessed potential associations between the clusters and traditional markers of aggressive tumors. χ^2 correlations showed that cluster 2 tumors were statistically more likely to be node positive, larger than 2 cm, and have a high nuclear grade (Table 4).

Finally, we looked at the independent prognostic ability of clustering. We compared the relative risk of patients in cluster 1 (lower expression of all markers) and cluster 2 (higher expression of all markers). Compared with traditional histopathologic markers of aggressive disease in a multivariate analysis, clustering retained its significance (relative risk, 1.36; 95% confidence interval, 1.02-1.83; $P = 0.0392$; Table 5). Among the other traditional breast cancer markers, only node status and tumor size were independent predictors of poor outcome (Table 5). Taken together, these results suggest hierarchical clustering using GnT-V substrate protein expression and LPHA can subdivide tumors into clusters with a more or less aggressive phenotype and independent prognostic value.

Discussion

GnT-V catalyzes the transfer of *N*-acetylglucosamine to α1,6-mannose in the pentasaccharide core of acceptor glycans, resulting in the production of complex tri- or tetra-antennary glycans with multiple β1,6 branches (35-37). The GnT-V–catalyzed addition of β1,6-branched oligosaccharides has been associated with progression in a variety of tumors including breast (22, 38), colon (20), non–small-cell lung (39), and melanoma (19).

This study uses a cohort of archival, formalin-fixed paraffin-embedded breast cancers to assess the association between β1,6 branching (as assayed using LPHA), the expression of GnT-V substrates, and markers of aggressive disease. This was done using a new method for automated, quantitative analysis for immunohistochemically stained tissue microarrays, AQUA (24). This system permits an objective assessment of protein

Figure 2. Unsupervised hierarchical clustering of LPHA and GnT-V substrate proteins reveals distinct tumor subsets. **A** and **B.** The heat maps are colored such that low expression is green and high expression is red (*clusters 1* and *2*). The distances between clusters, as calculated using average linkage clustering, are represented in the dendogram. Clustering was initially done on a "training" set of tumors ($n = 282$), which divided the tumors into two separate clusters (*red* and *green*, **A**). These clusters exhibited a significant 30-year disease-related survival difference (log-rank $P = 0.0158$; **C**). Additional subdivisions of the training set failed to reveal additional subsets with significantly different survival. Tumors in a separate "validation" cohort were then clustered based on their proximity to the centroids of the clusters identified in the training set ($n = 282$; **B**). Again, the two clusters defined different populations of tumors with significantly different survival (log-rank $P = 0.009$; **D**). Expression of each GnT-V substrate as well as LPHA in the "high" (*red*) versus "low" (*green*) clusters was significantly different ($P < 0.0001$; **E**). Because clustering was done using *z* scores (see Material and Methods), a score of 0 represents the mean of the population, where positive values represent relative overexpression and negative values represent relative underexpression (**E**). A subpopulation of tumors exhibits high level of LPHA expression in the absence of other substrates (**A** and **B**, *arrows*). This subset, however, has no significant difference in survival ($P = 0.6674$).

biomarker expression over a broad, continuous dynamic range. One limitation of using fixed archival specimens is that we could only assess coassociations between global levels of β1,6 branching and the expression of GnT-V substrates, rather than assessing the level of β1,6 branching on specific substrates. Consequently, we focused on five proteins, matriptase, lamp-1, N-cadherin, β1-integrin, and EGFR, which previously have been shown to be glycosylated by GnT-V in *in vitro* studies.

GnT-V–catalyzed glycosylation on each of these substrates induces functional changes that promote tumor aggression. In particular, GnT-V glycosylation of matriptase and lamp-1 inhibits their degradation and prolongs their localization and activity at the cell membrane (12, 17, 40). Matriptase potentiates the binding of hepatocyte growth factor to the Met receptor, which promotes tumor proliferation and motility (9, 17, 34). The stabilization of lamp-1 results in increased extracellular matrix degradation (12). Lamp-1 can also be highly decorated with sialyl Lewis X antigens carried by the polylactosamine antennae created by GnT-V–catalyzed β1,6 branching (41). This, in turn, promotes metastasis by promoting tumor-endothelial binding (41-44). β1,6 branching on β1-integrin inhibits the formation of the fibronectin receptor, α5β1, resulting in decreased extracellular matrix adhesion and increased cell motility (14, 45-47). GnT-V glycosylation of N-cadherin reduces N-cadherin clustering at the cell membrane, resulting in reduced adhesion (18). GnT-V–catalyzed glycosylation of EGFR enhances ligand binding, autophosphorylation, and subsequent downstream signaling, thus promoting cell proliferation (15, 16, 48).

In this study, we find that the expression of individual GnT-V substrates is associated with markers of tumor aggression (e.g., node status, tumor size, and nuclear grade). We also note that the expression of many of these substrates (as well as β1,6 branching) is frequently coassociated on tumors. Although these associations do not imply an underlying mechanistic link for coexpression, they do show that tumors tend to express globally higher or lower levels of GnT-V substrates and β1,6 branching. Of these markers, only matriptase and LPHA were able to significantly predict outcome; however, their predictive power was weak and exhibited no independent prognostic value. We used unsupervised hierarchical clustering to determine if the levels of β1,6 branching and GnT-V substrate expression would classify tumors into more and less aggressive phenotypes. Our study shows that tumors that coexpress higher levels of β1,6 branching and appropriate GnT-V substrates have a significantly worse outcome. In contrast to the relatively weak predictive power of assessing each substrate individually, clustering of tumors based on expression of all of the substrates together provided a model that could predict outcome independent of traditional markers of aggressive disease. The statistical significance of this model was validated using a training-set validation-set approach. We note that there is a subset of LPHA-high, GnT-V-substrate-low tumors (Fig. 2, arrow). The outcome of these tumors is identical to LPHA-low, GnT-V-substrate-low tumors, suggesting that high levels of β1,6 branching (LPHA staining) in the absence of the appropriate GnT-V substrates do not affect outcome.

Our results suggest that whereas there is limited prognostic benefit in assessing individual GnT-V substrates or LPHA-binding, coexpression of these substrates in the presence of β1,6 branching provides strong, independent prognostic information. These findings validate prior *in vitro* findings on the importance of GnT-V glycosylation in tumor progression on a cohort of human breast cancers. To our knowledge, this is the first report of glycosylation-associated coexpression of glycoprotein substrates in primary tumor samples. That these substrates together help define tumor aggression provides a novel perspective on the role of aberrant glycosylation in metastasis.

Table 4. χ^2 correlation of traditional markers and clusters

Marker	High cluster *n* (%)	Low cluster *n* (%)	*P*
Node positive	**217 (57)**	**94 (40)**	**<0.0001**
Size > 2 cm	**348 (58)**	**94 (43)**	**0.0004**
High nuclear grade	**120 (33)**	**45 (21)**	**0.0019**
HER2 3+	78 (21)	29 (14)	0.0529
Estrogen receptor negative	181 (48)	98 (46)	0.6989
Age > 50 y	168 (72)	271 (71)	0.8082
Lobular	**27 (12)**	**26 (25)**	**0.0017**

NOTE: Values represent the number (%) of tumors with the given traditional marker phenotype (left column) in the given cluster. Statistically significant values are in boldface.

Table 5. Multivariate analysis of cluster 1 versus cluster 2 and traditional markers

Marker	Relative risk	95% confidence interval	*P*
Cluster	**1.36**	**1.02-1.82**	**0.0392**
Node positive	**2.26**	**1.68-3.03**	**<0.0001**
Tumor size	**1.05**	**1.03-1.07**	**<0.0001**
High nuclear grade	1.10	0.83-1.42	0.4949
HER2 3+	1.38	0.98-1.95	0.0670
Estrogen receptor negative	1.22	0.93-1.62	0.1526

NOTE: Statistically significant values are in boldface.

References

1. Kaneko M, Alvarez-Manilla G, Kamar M, et al. A novel β(1,6)-*N*-acetylglucosaminyltransferase V (GnT-VB)(1). FEBS Lett 2003;554:515–9.
2. Dennis JW, Granovsky M, Warren CE. Glycoprotein glycosylation and cancer progression. Biochim Biophys Acta 1999;1473:21–34.
3. Couldrey C, Green JE. Metastases: the glycan connection. Breast Cancer Res 2000;2:321–3.
4. Mitchell BS, Brooks SA, Leathem AJ, Schumacher U. Do HPA and PHA-L have the same binding pattern in metastasizing human breast and colon cancers? Cancer Lett 1998;123:113–9.
5. Granovsky M, Fata J, Pawling J, Muller WJ, Khokha R, Dennis JW. Suppression of tumor growth and metastasis in Mgat5-deficient mice. Nat Med 2000;6:306–12.
6. Taniguchi N, Ihara S, Saito T, Miyoshi E, Ikeda Y, Honke K. Implication of GnT-V in cancer metastasis: a glycomic approach for identification of a target protein and its unique function as an angiogenic cofactor. Glycoconj J 2001;18:859–65.
7. Ono M, Hakomori S. Glycosylation defining cancer cell motility and invasiveness. Glycoconj J 2004;20:71–8.
8. Brooks SA, Hall DM. Investigations into the potential role of aberrant *N*-acetylgalactosamine glycans in tumour cell interactions with basement membrane components. Clin Exp Metastasis 2002;19:487–93.
9. Ihara S, Miyoshi E, Nakahara S, et al. Addition of β1-6 GlcNAc branching to the oligosaccharide attached to Asn 772 in the serine protease domain of matriptase plays a pivotal role in its stability and resistance against trypsin. Glycobiology 2004;14:139–46.
10. Lin CY, Anders J, Johnson M, Dickson RB. Purification and characterization of a complex containing matriptase and a Kunitz-type serine protease inhibitor from human milk. J Biol Chem 1999;274:18237–42.
11. Carlsson SR, Fukuda M. The polylactosaminoglycans of human lysosomal membrane glycoproteins lamp-1 and lamp-2. Localization on the peptide backbones. J Biol Chem 1990;265:20488–95.
12. Fukuda M. Lysosomal membrane glycoproteins. Structure, biosynthesis, and intracellular trafficking. J Biol Chem 1991;266:21327–30.
13. Nieman MT, Prudoff RS, Johnson KR, Wheelock MJ. N-Cadherin promotes motility in human breast cancer cells regardless of their E-cadherin expression. J Cell Biol 1999;147:631–44.
14. Guo HB, Lee I, Kamar M, Akiyama SK, Pierce M. Aberrant N-glycosylation of β1 integrin causes reduced α5β1 integrin clustering and stimulates cell migration. Cancer Res 2002;62:6837–45.
15. Guo P, Wang QY, Guo HB, Shen ZH, Chen HL. *N*-Acetylglucosaminyltransferase V modifies the signaling pathway of epidermal growth factor receptor. Cell Mol Life Sci 2004;61:1795–804.
16. Pierce M, Buckhaults P, Chen L, Fregien N. Regulation of *N*-acetylglucosaminyltransferase V and Asn-linked oligosaccharide β(1,6) branching by a

growth factor signaling pathway and effects on cell adhesion and metastatic potential. Glycoconj J 1997;14:623–30.
17. Ihara S, Miyoshi E, Ko JH, et al. Prometastatic effect of *N*-acetylglucosaminyltransferase V is due to modification and stabilization of active matriptase by adding β1–6 GlcNAc branching. J Biol Chem 2002;277:16960–7.
18. Guo HB, Lee I, Kamar M, Pierce M. *N*-Acetylglucosaminyltransferase V expression levels regulate cadherin-associated homotypic cell-cell adhesion and intracellular signaling pathways. J Biol Chem 2003;278:52412–24.
19. Handerson T, Pawelek JM. β1,6-Branched oligosaccharides and coarse vesicles: a common, pervasive phenotype in melanoma and other human cancers. Cancer Res 2003;63:5363–9.
20. Murata K, Miyoshi E, Kameyama M, et al. Expression of *N*-acetylglucosaminyltransferase V in colorectal cancer correlates with metastasis and poor prognosis. Clin Cancer Res 2000;6:1772–7.
21. Jin XL, Liu HB, Zhang Y, Wang BS, Chen HL. Alteration in *N*-acetylglucosaminyltransferase activities and glycan structure in tissue and bile glycoproteins from extrahepatic bile duct carcinoma. Glycoconj J 2004; 20:399–406.
22. Handerson T, Camp R, Harigopal M, Rimm D, Pawelek J. β1,6-Branched oligosaccharides are increased in lymph node metastases and predict poor outcome in breast carcinoma. Clin Cancer Res 2005;11:2969–73.
23. Rubin MA, Zerkowski MP, Camp RL, et al. Quantitative determination of expression of the prostate cancer protein α-methylacyl-CoA racemase using automated quantitative analysis (AQUA): a novel paradigm for automated and continuous biomarker measurements. Am J Pathol 2004;164:831–40.
24. Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. Nat Med 2002;8: 1323–7.
25. Rimm DL, Camp RL, Charette LA, Olsen DA, Provost E. Amplification of tissue by construction of tissue microarrays. Exp Mol Pathol 2001;70:255–64.
26. Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844–7.
27. Lin CY, Wang JK, Torri J, Dou L, Sang QA, Dickson RB. Characterization of a novel, membrane-bound, 80-kDa matrix-degrading protease from human breast cancer cells. Monoclonal antibody production, isolation, and localization. J Biol Chem 1997;272:9147–52.
28. Harigopal M, Berger AJ, Camp RL, Rimm DL, Kluger HM. Automated quantitative analysis of E-cadherin expression in lymph node metastases is predictive of survival in invasive ductal breast cancer. Clin Cancer Res 2005; 11:4083–9.
29. Berger AJ, Camp RL, Divito KA, Kluger HM, Halaban R, Rimm DL. Automated quantitative analysis of HDM2 expression in malignant melanoma shows association with early-stage disease and improved outcome. Cancer Res 2004;64:8767–72.
30. Divito KA, Berger AJ, Camp RL, Dolled-Filhart M, Rimm DL, Kluger HM. Automated quantitative analysis of tissue microarrays reveals an association between high Bcl-2 expression and improved outcome in melanoma. Cancer Res 2004;64:8773–7.
31. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95: 14863–8.
32. Cheadle C, Vawter MP, Freed WJ, Becker KG. Analysis of microarray data using *Z* score transformation. J Mol Diagn 2003;5:73–81.
33. Arpino G, Bardou VJ, Clark GM, Elledge RM. Infiltrating lobular carcinoma of the breast: tumor characteristics and clinical outcome. Breast Cancer Res 2004;6:R149–56.
34. Kang JY, Dolled-Filhart M, Ocal IT, et al. Tissue microarray analysis of hepatocyte growth factor/Met pathway components reveals a role for Met, matriptase, and hepatocyte growth factor activator inhibitor 1 in the progression of node-negative breast cancer. Cancer Res 2003;63:1101–5.
35. Yamashita K, Ohkura T, Tachibana Y, Takasaki S, Kobata A. Comparative study of the oligosaccharides released from baby hamster kidney cells and their polyoma transformant by hydrazinolysis. J Biol Chem 1984;259: 10834–40.
36. Hakomori S. Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. Cancer Res 1996;56:5309–18.
37. Brockhausen I, Narasimhan S, Schachter H. The biosynthesis of highly branched *N*-glycans: studies on the sequential pathway and functional role of *N*-acetylglucosaminyltransferases I, II, III, IV, V and VI. Biochimie 1988; 70:1521–33.
38. Fernandes B, Sagman U, Auger M, Demetrio M, Dennis JW. β1-6 branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia. Cancer Res 1991;51:718–23.
39. Dosaka-Akita H, Miyoshi E, Suzuki O, Itoh T, Katoh H, Taniguchi N. Expression of *N*-acetylglucosaminyltransferase v is associated with prognosis and histology in non-small cell lung cancers. Clin Cancer Res 2004;10: 1773–9.
40. Kundra R, Kornfeld S. Asparagine-linked oligosaccharides protect Lamp-1 and Lamp-2 from intracellular proteolysis. J Biol Chem 1999;274: 31039–46.
41. Phillips ML, Nudelman E, Gaeta FC, et al. ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Lex. Science 1990;250:1130–2.
42. Fukushima K, Hirota M, Terasaki PI, et al. Characterization of sialosylated Lewisx as a new tumor-associated antigen. Cancer Res 1984;44:5279–85.
43. Walz G, Aruffo A, Kolanus W, Bevilacqua M, Seed B. Recognition by ELAM-1 of the sialyl-Lex determinant on myeloid and tumor cells. Science 1990;250: 1132–5.
44. Sackstein R, Dimitroff CJ. A hematopoietic cell L-selectin ligand that is distinct from PSGL-1 and displays *N*-glycan-dependent binding activity. Blood 2000;96:2765–74.
45. Zheng M, Fang H, Hakomori S. Functional role of N-glycosylation in α5β1 integrin receptor. De-N-glycosylation induces dissociation or altered association of α5 and β1 subunits and concomitant loss of fibronectin binding activity. J Biol Chem 1994;269:12325–31.
46. Howlett AR, Bailey N, Damsky C, Petersen OW, Bissell MJ. Cellular growth and survival are mediated by β1 integrins in normal human breast epithelium but not in breast carcinoma. J Cell Sci 1995;108:1945–57.
47. Isaji T, Gu J, Nishiuchi R, et al. Introduction of bisecting GlcNAc into integrin α5β1 reduces ligand binding and down-regulates cell adhesion and cell migration. J Biol Chem 2004;279:19747–54.
48. Bishayee S. Role of conformational alteration in the epidermal growth factor receptor (EGFR) function. Biochem Pharmacol 2000;60:1217–23.