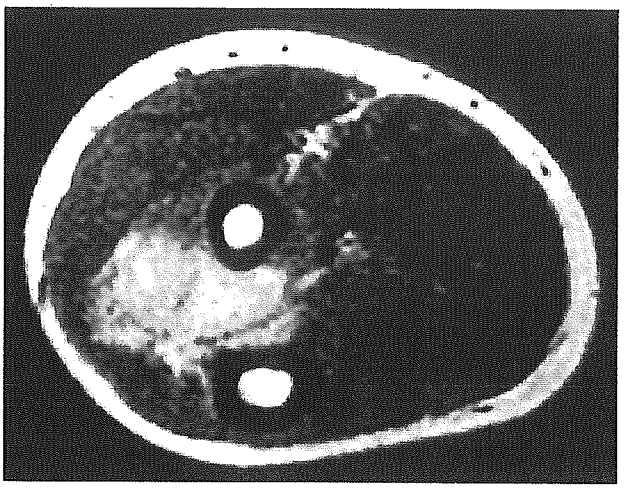

Figure 1 T2-weighted magnetic resonance imaging showing a deep soft-tissue mass.

tion, 1:30), muscle-specific actin (clone HHF-35, Enzo Biochem, Farmingdale, NY, USA, 1:50), glial fibrillary acidic protein (GFAP; clone GA5, Novocastra, Newcastle upon Tyne, UK, 1:100), low-molecular-weight cytokeratin (clone CAM5.2, Becton Dickinson, Franklin Lakes, NJ, USA, 1:2), high-molecular-weight cytokeratin (HMWCK; clone 34betaE12, DakoCytomation, 1:40), epithelial membrane antigen (EMA; clone GP1.4, Novocastra, 1:100), S-100 protein (polyclonal, DakoCytomation, 1:400), and Ki-67 antigen (clone MIB-1, DakoCytomation, 1:40).

Tumor cells were diffuse and strongly positive for S-100 protein (Fig. 3a). In addition, EMA and low-molecular-weight cytokeratin showed distinct positivity in most tumor cells (Fig. 3b,c). In contrast, only focal, weak staining was observed for SMA and vimentin, and no cells were positive for desmin, muscle actin, GFAP, and HMWCK. The Ki-67 proliferating cell marker showed wide range positivity in the tumor, from <1% to 10%. Lung biopsy specimens obtained 6 months later showed essentially the same morphological and immunohistochemical findings as the primary tumor.

Electronmicroscopic examination using a sample retrieved from the formalin-fixed tissue (primary tumor) revealed intercellular desmosome-like junctions and basal laminae (Fig. 4). Taking all these findings into consideration, we finally diagnosed a primary malignant myoepithelioma (myoepithelial carcinoma) of the soft tissue.

DISCUSSION

Myoepithelial tumors of the soft tissue are very rare,[3–5] and are considered to be morphologically the same spectrum as their skin[6] and salivary gland[2] counterparts. Their incidence is not accurately known because of their rarity, and the difference, if any, between the sexes is also obscure. In the limited number of studies available, the patients' average age was 35 years, with a significant number being children under 10 years old.[1] The commonest sites for the tumors are the subcutaneous or deep subfascial soft tissues of the extremities. However, the head and neck and the trunk region are also possible locations for primary sites.[1,4]

Although most myoepithelial tumors follow a benign clinical course, malignant cases do exist. Such cases (malignant myoepithelial tumors) are extremely rare; indeed, only a small number have ever been reported.[3,4] It is quite difficult to predict the biological behavior of myoepithelial tumors. In 2003, Hornick and Fletcher reported a large-scale study of myoepithelial tumors of the soft tissue.[4] They reported that the cytologically malignant tumors seemed to have a tendency to behave aggressively because approximately one-third of cytologically malignant cases metastasized in their series, whereas none of the 33 cases diagnosed as cytologically benign did. In the present case, many tumor cells at the primary site showed obvious cellular atypia with some mitoses, and the tumor later metastasized. These findings and course are consistent with their data. Although the precise criteria for malignancy remain unclear, cytologically malignant cases might have to be radically treated.

For the purposes of differential diagnosis, myoepithelial tumors of the soft tissue need to be considered along with metastatic carcinoma, mixed tumors of the skin, parachordoma, extraskeletal myxoid chondrosarcoma, myxoid leiomyosarcoma, epithelioid sarcoma, and nerve sheath myxoma.[3–5]

Of these, metastatic carcinomas are easily excluded in the present case because of its long clinical history and the patient's age. Mixed tumors of the skin (cutaneous mixed tumors) are histologically identical to myoepithelial tumors of the soft tissue, with the location of the tumor being the key to distinguishing between them.

Parachordomas are important candidates for differential diagnosis. They are very rare, soft-tissue tumors of unknown lineage, usually arising in deep soft tissue, characteristics that imply a similarity to chordoma as a definition.[7] They are slowly growing tumors, occasionally with late recurrence, and exhibit a variety of histological patterns and cytological features. In recent years, however, they have come to be considered as identical to myoepithelial tumors of soft tissue because of their close resemblance (as revealed by ultrastructural and immunohistochemical analyses) in such features as a constant, strong S-100 protein expression and the existence of primitive cell junctions and basal laminae.[1,8]

Extraskeletal myxoid chondrosarcomas (EMC), which commonly arise in the extremities, are also important candidates for differential diagnosis because their histological features are very similar to those of myoepithelial tumors in some cases.[3–5] However, EMC differ from myoepithelial

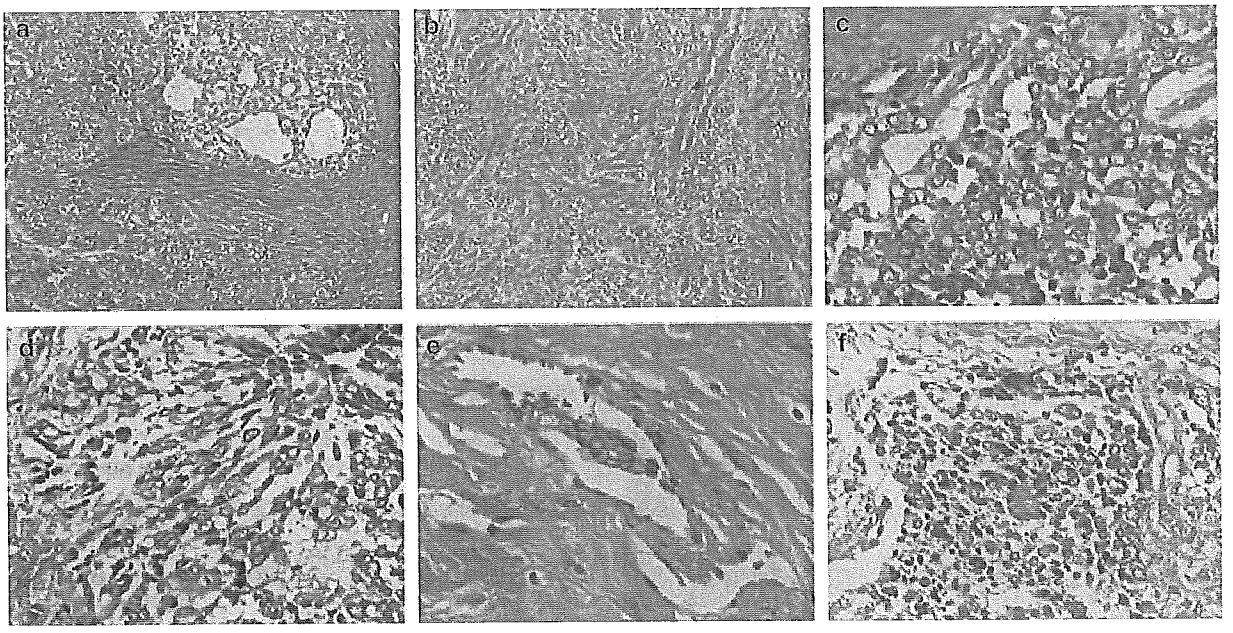

Figure 2 Histology of the tumor. It exhibited (**a**) a lobular growth pattern with (**b**) a widely infiltrative zone. (**c**,**d**) Both epithelioid and spindle cells were observed. (**e**,**f**) In high-power view, vascular invasion and small foci of necrosis were observed.

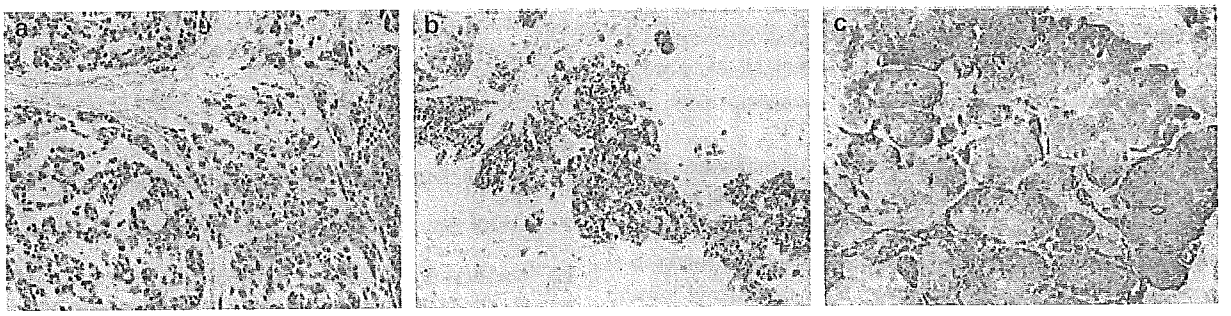

Figure 3 Immunohistochemical analysis of (**a**) S-100 protein, (**b**) low-molecular-weight cytokeratin, and (**c**) epithelial membrane antigen. Myoepithelial nature is evident.

tumors of soft tissue immunohistochemically. Vimentin is the only marker consistently expressed in EMC. Although S-100 protein, cytokeratin, and EMA are focally positive in some cases, their positivity is limited to a minority of tumor cells.[1,4] Ultrastructural examination might also help to distinguish EMC from myoepithelial tumors by revealing the lack of epithelial characters in EMC. Further, a recent genetic study has revealed characteristic fusion transcripts generated by chromosomal aberrations in EMC, including EWS/NR4A3 with t(9;22)(q22;q12) and RBP56/NR4A3 with t(9;17)(q22;q11).[1] These aberrations can be detected by reverse transcription–polymerase chain reaction (RT-PCR) and have not been recorded in myoepithelial tumors of the soft tissue.

In relatively rare cases, a myoepithelial tumor may be rich in spindle cells,[1,4] and may thereby resemble other soft-tissue tumors such as myxoid leiomyosarcoma, epithelioid sarcoma, and nerve sheath myxoma,[4] especially in small biopsy specimens. In such cases, immunoreactivity for both epithelial markers and S-100 protein would support a diagnosis of myoepithelial tumor, and reliance on morphological findings alone could be diagnostically misleading on occasions.

The present case is especially interesting in that the tumor had preexisted as a deep soft-tissue mass since the patient was 3 years old. This history might suggest the possibility of a malignant transformation of a preexisting benign mixed tumor. However, his preexisting tumor had not been histologically examined, and the resected tumors did not include a morphologically benign mixed tumor. Among carcinomas ex mixed tumor, myoepithelial carcinomas ex mixed tumor have occasionally been reported in salivary glands, although they

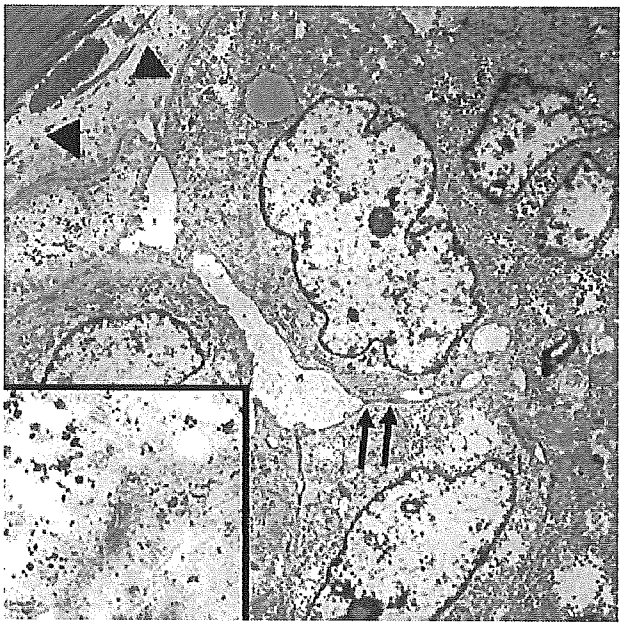

Figure 4 Ultrastructural appearance of tumor cells showing basal lamina (arrowhead) and desmosome-like junction (arrow, inset).

are very rare.[2] Among soft-tissue cases, Hornick and Fletcher reported that two of 35 myoepithelial carcinomas of the soft tissue contained areas of morphologically benign mixed tumor.[4] The clinical history, however, was not described in their report.

In summary, we have reported a case of malignant myoepithelioma of the soft tissue in which multiple lung metastases developed. Pathologists must be aware that a malignant myoepithelial tumor should be included among the candidates for differential diagnosis when epithelioid tumor cells with a myxoid background are identified in deep soft tissue. Although cases with cytologically malignant features seem to have a tendency to show a malignant clinical course, the precise criteria for malignancy remain unknown.

ACKNOWLEDGMENTS

This work was supported in part by Grant-in-Aid for Scientific Research B-15390115 from the Japan Society for the Promotion of Science. The authors thank Mrs A. Ishida for excellent technical assistance.

REFERENCES

1 Fletcher CDM, Unni KK, Mertens F, eds. *World Health Organization Classification Tumours Pathol Genet Tumours Soft Tissue Bone.* Lyon: IARC Press, 2002.
2 Ellis GL, Auclair PL. *Tumors of the Salivary Gland. Atlas of Tumor Pathology*, 3rd series, fascicle 17. Washington, DC: Armed Forces Institute Pathology, 1996.
3 Kilpatrick SE, Hitchcock MG, Kraus MD, Calonje E, Fletcher CD. Mixed tumors and myoepitheliomas of soft tissue: a clinicopathological study of 19 cases with a unifying concept. *Am J Surg Pathol* 1997; **21**: 13–22.
4 Hornick JL, Fletcher CD. Myoepithelial tumors of soft tissue: a clinicopathological and immunohistochemical study of 101 cases with evaluation of prognostic parameters. *Am J Surg Pathol* 2003; **27**: 1183–96.
5 Adachi T, Oda Y, Sakamono A *et al.* Mixed tumor of deep soft tissue. *Pathol Int* 2003; **53**: 35–9.
6 Michal M, Miettinen M. Myoepitheliomas of the skin and soft tissues. Report of 12 cases. *Virchows Arch* 1999; **434**: 393–400.
7 Weiss SW, Goldblum JR. *Enzinger and Weiss's Soft Tissue Tumors*, 4th edn. St Louis: Mosby, 2001.
8 Karabela-Bouropoulou V, Skourtas C, Liapi-Avgeri G, Mahaira H. Parachordoma. A case report of a very rare soft tissue tumor. *Pathol Res Pract* 1996; **192**: 972–8; discussion 979–81.

Glycobiology vol. 15 no. 10 pp. 1016–1024, 2005
doi:10.1093/glycob/cwi086
Advance Access publication on June 2, 2005

Expression of core 2 β1,6-*N*-acetylglucosaminyltransferase facilitates prostate cancer progression

Shigeru Hagisawa[2], Chikara Ohyama[3], Toshiko Takahashi[2], Mareyuki Endoh[4], Takuya Moriya[2], Jun Nakayama[5], Yoichi Arai[2], and Minoru Fukuda[1,6]

[2]Department of Urology, Tohoku University School of Medicine, Sendai, Japan;[3]Department of Urology, Hirosaki University School of Medicine, Hirosaki, Japan;[4]Department of Pathology, Tohoku University Hospital, Sendai, Japan;[5]Department of Pathology, Shinshu University School of Medicine, Matsumoto, Japan; and [6]Glycobiology Program, Cancer Research Center, The Burnham Institute, La Jolla, CA 92037

Received on April 15, 2005; revised on May 23, 2005; accepted on May 24, 2005

Cell surface carbohydrates expressed on epithelial cells are thought to play an important role in tumor progression. Previously, we have shown that expression of core 2-branched *O*-glycans is closely correlated with vessel invasion and depth of invasion in colon and lung carcinomas. In this study, we found that expression of core 2 β1,6-*N*-acetylglucosaminyltransferase-1, Core2GnT, is positively correlated with the progression of prostate cancer in human patients. Statistical analysis demonstrated that Core2GnT is an independent predictor for progressed pathological stage (pT3) and for prostate-specific antigen (PSA) relapse. To determine directly the roles of Core2GnT in prostate cancer progression, we set up an experimental tumor model using the LNCaP prostate cancer cell line. Because this line does not express Core2GnT, we established an LNCaP line stably expressing Core2GnT, LNCap-Core2GnT, by transfecting cDNA encoding Core2GnT. When mock-transfected LNCaP cells and LNCaP-Core2GnT were inoculated in the prostate of nude mice, LNCaP-Core2GnT cells produced three times heavier prostate tumors than mock-transfected LNCaP cells. Furthermore, we found that LNCaP-Core2GnT cells adhered more strongly to prostate stromal cells, type IV collagen and laminin than did LNCaP-mock cells, but LNCaP and LNCaP-Core2GnT cells grew almost at the same rate on plates coated with type IV collagen or laminin. These results indicate that Core2GnT is an extremely useful prognostic marker for prostate cancer progression. The results also suggest that acquiring Core2GnT in prostate carcinoma cells facilitates adhesion to type IV collagen and laminin, and this increased adhesion may be a cause for aggressive tumor formation by prostate cancer cells expressing Core2GnT.

Key words: core 2 β1,6-*N*-acetylglucosaminyltransferase/core 2-branched *O*-glycans/prostate cancer/prostate stromal cell/tumor metastasis

[1]To whom correspondence should be addressed; e-mail: minoru@burnham.org

Introduction

In the United States, prostate cancer is the most common malignancy affecting men and the second leading cause of cancer death (American Cancer Society, 2003). The incidence of prostate cancer varies worldwide, with the highest rates found in the United States, Canada, and Scandinavia and the lowest rates in China and other parts of Asia (Quinn and Babb, 2002). However, increase in prostate cancer incidence has occurred in countries with relatively low incidence rates of prostate cancer (Gronberg, 2003).

Several treatments are currently available for early prostate cancer, such as watchful waiting, radical retropubic prostatectomy, laparoscopic prostatectomy, transperineal prostatectomy, extrabeam irradiation, and transperineal implantation of radioisotopes (brachytherapy). After radical prostatectomy, the disease recurs in an estimated 15–30% of patients, suggesting that undetected cancer cells may have spread beyond the prostate gland before surgery (Han *et al.*, 2001; Roberts *et al.*, 2001). Several clinical parameters including tumor stage, tumor grade as measured by the Gleason score (GS), and the serum level of prostate-specific antigen (PSA) are typically used to assess the risk of disease progression at the time of diagnosis (Partin *et al.*, 1997). Kattan *et al.* (2001) also have developed useful nomograms to help evaluate the likelihood of disease-free survival after radical prostatectomy or brachytherapy for localized prostate cancer. However, these and other models have limitations as demonstrated by their good, but not excellent, association with outcome, as reviewed by Ross *et al.*, (2001). It is thus imperative to develop a novel biomarker to accurately assess the risk of disease progression in patients with clinically localized prostate cancer so that appropriate treatment can be selected.

Neoplastic transformation of epithelial cells is accompanied by alterations in expression of cell surface carbohydrates (Ohyama *et al.*, 1995; Fukuda, 1996; Watanabe *et al.*, 2002). For example, the expression of sialyl Lewis A or sialyl Lewis X on the cell surface of colorectal cancer is positively correlated with poor patient outcome (Nakamori *et al.*, 1993). In mucin-type glycoproteins, sialyl Lewis X and sialyl Lewis A are often present as a capping structure on core 2-branched *O*-glycans, and these structures have been shown to serve as selectin ligands (Lowe, 1994). Following the formation of a core 2-branch, sequential addition of sialic acid and α1,3-linked or α1,4-linked fucose results in formation of sialyl Lewis X or sialyl Lewis A (Lowe, 1994). It was also reported that a soluble form of *N*-acetylglucosaminyltransferase-V (GnT-V) might facilitate angiogenesis (Saito *et al.*, 2002). By contrast, forced expression of

core 3 *O*-glycans, which apparently down-regulates core 1 *O*-glycans, reduced tumor formation by human colonic carcinoma cells (Iwai *et al.*, 2005).

Core 2 β1,6-*N*-acetylglucosaminyltransferase-I (Core2GnT-I) (Bierhuizen and Fukuda, 1992; Schachter and Brockhausen, 1992) is a key enzyme forming core 2-branched *O*-glycans (Galβ1→3(GlcNAcβ1→6)GalNAcα→Serine/Threonine) by catalyzing the transfer of N-acetylglucosamine (GlcNAc) to core 1 *O*-glycan (Galβ1→3GalNAcα→Serine/Threonine). It was reported earlier that increases in Core2GnT activity are seen in human leukemia cells and metastatic murine tumor cell lines (Brockhausen *et al.*, 1991; Saitoh *et al.*, 1991; Yousefi *et al.*, 1991). Similarly, the expression of sialyl Lewis A on core 2-branched *O*-glycans, detected by CA19-9 antibody, is associated with the progression of colorectal carcinoma (Shimono *et al.*, 1994).

We have cloned a cDNA encoding Core2GnT from human promyelocytic leukemia HL-60 cells by expression cloning (Bierhuizen and Fukuda, 1992) and raised antibodies against Core2GnT (Skrincosky *et al.*, 1997). We showed that expression of core 2-branched *O*-glycans is closely correlated with the malignant potential of colorectal cancer and pulmonary adenocarcinoma by analyzing the expression of *Core2GnT* mRNA (Shimodaira *et al.*, 1997; Machida *et al.*, 2001). Notably, we found that the expression of Core2GnT gene is better correlated with the progression of tumors in both colon and lung cancer than expression of sialyl Lewis X or sialyl Lewis A itself (Shimodaira *et al.*, 1997; Machida *et al.*, 2001). These observations suggest that expression of Core2GnT and resultant core 2-branched *O*-glycans is highly correlated with the progression of various tumors. However, there have been no attempts to elucidate the clinicopathological significance of Core2GnT status in prostate cancer.

It has been reported that overexpression of fibroblast growth factor 8-b in human prostate cancer cell line, LNCaP cells, allowed them to evade the growth inhibitory effect of stromal cells (Song *et al.*, 2000). It has also been shown that prostate cancer cells grew better after forming capillary-like stroma when cocultured in collagen gels and endothelial cells in fibrin (Janvier *et al.*, 1997). Similarly, coculturing of prostate cancer cells with stromal cells facilitates angiogenesis surrounding prostate cancer cells and leads to increased size of prostate tumors (Tuxhorn *et al.*, 2002), whereas the invasion of DU145 prostate cancer cell line was facilitated by stromal-derived hepatocyte growth factor (Nishimura *et al.*, 1999). These results indicate the importance of prostate stromal cells in contact with prostate cancer cells in xenograft tumor formation. However, no studies have been undertaken to elucidate the roles of core 2-branched *O*-glycans in xenograft tumor formation by prostate cancer cells.

Here, we immunohistochemically examined Core2GnT status in prostate needle biopsy specimens using the anti-Core2GnT antibodies previously prepared (Skrincosky *et al.*, 1997). We found that expression of Core2GnT in preoperative prostate tissue could be a highly useful predictor of the pathological stage. Moreover, we demonstrated that LNCaP prostate cancer cells form larger tumors in nude mice after LNCaP cells were transfected with a Core2GnT expression vector, and those transfected cells adhere more efficiently to prostate stromal cells than mock-transfected LNCaP cells. These results as a whole indicate that Core2GnT and core 2-branched *O*-glycans synthesized play a critical role in prostate cancer progression, most likely through increased interaction with prostate stromal cells.

Results

Expression of Core2GnT is positively correlated with progression of prostate cancer

To evaluate the role of Core2GnT in prostate cancer progression, we utilized antibodies specifically raised against Core2GnT protein (Skrincosky *et al.*, 1997). Because the specimens examined in this study were embedded in paraffin, immunohistochemical detection of Core2GnT was more efficient than detection of either *Core2GnT* transcripts by in situ hybridization or core 2-branched *O*-glycans. The detection of core 2-branched *O*-glycans by specific antibodies is severely limited because available antibodies can detect core 2 *O*-glycans only when they are capped with sialyl Lewis X (Berg *et al.*, 1991; Kumamoto *et al.*, 1998; Kobayashi *et al.*, 2004) or core 2-branched *O*-glycans attached to CD43 (leukosialin) (Piller *et al.*, 1991).

The results demonstrate that normal prostate gland barely express Core2GnT, whereas prostate cancer cells express significant levels of Core2GnT as detected by rabbit anti-Core2GnT antibodies (Figure 1A and B). Using these criteria, we then examined specimens from 69 patients exhibiting different clinical parameters. First, we found that Core2GnT expression in biopsy specimens was positively correlated with serum PSA level (Table I). Prostate cancer with a low Gleason sum (five and six) and a high Gleason sum (eight and nine) express Core2GnT in 11 and 51% of the patients, respectively. We found that 4% of Core2GnT-negative patients are at pT3 disease, whereas 53% of core 2-positive patients are with pT3 disease (Table I).

The most important clinical implication of preoperative parameter is whether it can predict PSA relapse, thus the necessity of secondary treatment. As shown in Figure 2, Core2GnT-positive patients had significantly higher risk for PSA relapse after prostatectomy. This result indicates that expression of Core2GnT alone or in combination with PSA is an excellent predictor of progression of prostate cancer.

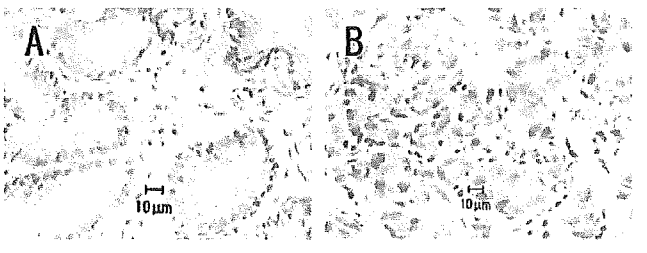

Fig. 1. Immunohistochemistry of human prostate cancer biopsy specimens. Prostate biopsy specimens were stained with anti-core 2 β1,6-*N*-acetylglucosaminyltransferase-1 (anti-Core2GnT) antibodies followed by horseradish peroxidase-conjugated secondary antibodies. A normal specimen (**A**) and a specimen with a Gleason sum of eight (**B**) were examined. Counterstaining was performed using hematoxylin. Core2GnT-positive cancer cells are stained in red.

Table I. Core 2 β1,6-*N*-acetylglucosaminyltransferase-1 (Core2GnT) status and clinicopathological parameters

	Core2GnT negative	Core2GnT positive	*p* value
Number of patients	24	45	
Age (years old) (mean ± standard deviation)	64.5 ± 6.6	68.2 ± 4.8	0.0102
Prostate-specific antigen (ng/mL) (mean ± standard deviation)	7.4 ± 3.7	11.9 ± 8.9	0.0196
Biopsy Gleason score			0.0002
≤6	13%	5%	
=7	9%	17%	
≥8	2%	23%	
Final pathological stage (%)			<0.0001
pT2	23 (96%)	21 (47%)	
pT3	1 (4%)	24 (53%)	

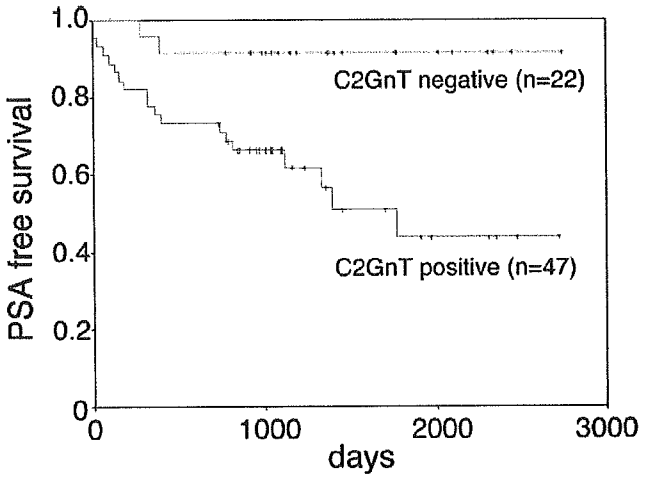

Fig. 2. Core 2 β1,6-*N*-acetylglucosaminyltransferase-1 (Core2GnT) expression is highly correlated with poor prognosis. Sixty-nine patients with T1 or T2 were divided in two groups depending on whether they were Core2GnT positive or negative, and the Kaplan-Meier survival analysis was calculated.

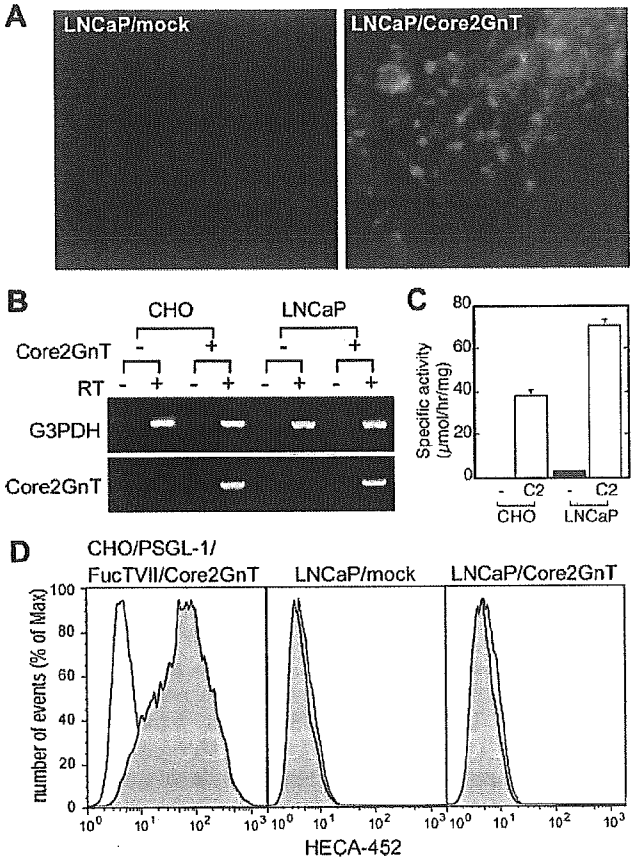

Fig. 3. Expression profile of sialyl Lewis X and core 2 1^b,6-N-acetylglucosaminyltransferase-1 (Core2GnT) on LNCaP cells before and after forced expression of Core2GnT. (A) LNCaP-mock and LNCaP-Core2GnT cells were stained with rabbit anti-Core2GnT antibodies followed by horseradish peroxidase-conjugated secondary antibody. The magnification is the same as that in Figure 1. (B) RT–PCR was performed on total RNA isolated from LNCaP-mock and LNCaP-core 2 β1,6-*N*-acetylglucosaminyltransferase-1 (LNCaP-Core2GnT) cells. RT–PCR for glyceraldehyde 3-phosphate dehydrogenase (G3PDH) is shown for positive control. Chinese hamster ovary (CHO) cells and CHO cells-expressing Core2GnT served as a negative and positive control for Core2GnT, respectively. (C) Core2GnT activity was assayed for LNCaP-mock and LNCaP-Core2GnT cells. (D) LNCaP-mock and LNCaP-Core2GnT cells were negative for sialyl Lewis X and sialyl Lewis A as assessed by FACS and HECA-452 antibody (Berg et al., 1991). CHO-PSGL-1/F7/C2 cells were used as a positive control for sialyl Lewis X expression.

Expression of Core2GnT cDNA into the LNCaP prostate cancer cell line

To determine the role of Core2GnT in prostate cancer progression, we stably transfected the LNCaP prostate cancer cell line with cDNA encoding *Core2GnT*. As shown in Figure 3A, the parent LNCaP cells do not express Core2GnT. After the transfection of *Core2GnT* cDNA, five clones of transfected cells expressing Core2GnT as assessed by staining with anti-Core2GnT antibodies (Figure 3A) were identified, and one was designated LNCaP-Core2GnT. Representative results using this clone are presented hereafter. The transcript (Figure 3B) and the enzymatic activity (Figure 3C) of Core2GnT were detected in LNCaP-Core2GnT but not in the mock-transfected LNCaP cells. The parent LNCaP cells were negative for HECA-452 even after transfection of Core2GnT, indicating the absence of sialyl Lewis X and sialyl Lewis A (Figure 3D).

LNCaP-Core2GnT and LNCaP-mock cells grow in vitro at a similar rate

We determined the cell number of LNCaP-Core2GnT and mock-transfected LNCaP cells at 2 days interval during 9-day culture. The results show no significant differences in cell numbers during in vitro culture between LNCaP-Core2GnT and mock-transfected LNCaP cells (data not shown), indicating that acquisition of Core2GnT does not lead to increased proliferation of LNCaP-Core2GnT cells cultured in vitro.

LNCaP-Core2GnT cells form larger tumors upon orthotopic inoculation into mouse prostate

To determine if expression of Core2GnT and core 2-branched *O*-glycans alters tumor formation in vivo, we inoculated LNCaP-Core2GnT and LNCaP-mock into the mouse prostate. Four weeks after inoculation, mice were killed, and the prostates were examined. The results shown in Figure 4A illustrate that LNCaP-Core2GnT produced much larger tumors than did LNCaP-mock cells, and the weight of prostates derived from LNCaP-Core2GnT inoculation was more than three times than that of prostate derived from mock-transfected LNCaP cells (Figure 4B). Almost identical results were obtained in a repeated experiment and the experiments using another LNCaP-Core2GnT cell line. These results indicate that expression of core 2-branched *O*-glycans, even in the absence of sialyl Lewis X or sialyl Lewis A on LNCaP cells, leads to significantly increased tumorigenicity.

Histological examination on tumor in nude mice

Histological examination revealed massive tumor formation of LNCaP-Core2GnT cells in the prostate of the nude mice (Figure 5A and C). The tumors formed by LNCaP-Core2GnT were positive for Core2GnT even several weeks after inoculation, whereas the mouse prostate without inoculation of LNCaP-Core2GnT cells was negative for Core2GnT (Figure 5B). Tumors formed by LNCaP-Core2GnT cells were characterized by a larger nuclei and a prominent nucleolus (Figure 5A and C). Papillary structures seen in the normal mouse prostate (Figure 5B) disappear in the tumors, further indicating the aggressiveness of these tumors.

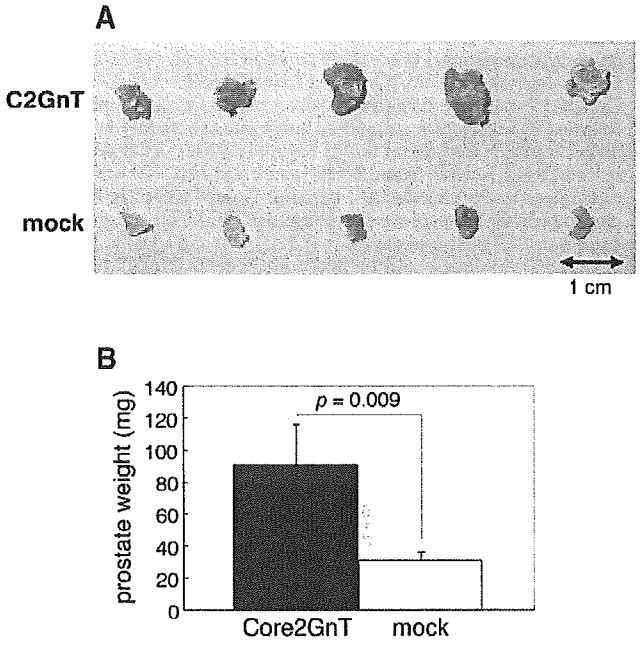

Fig. 4. Tumor formation by LNCaP cells in nude mice. LNCaP-mock and LNCaP-core 2 β1,6-*N*-acetylglucosaminyltransferase-1 (LNCaP-Core2GnT) cells were inoculated in the prostate of the nude mice. **(A)** After 4 weeks, extirpated prostate specimens were obtained and photographed. **(B)** The wet weight of prostate was measured. LNCaP-Core2GnT cells produced tumor-bearing prostate as much as three times heavier than did mock-transfected LNCaP cells.

Adhesion of LNCaP-Core2GnT cells to prostate stroma cells

We then determined whether LNCaP-Core2GnT cells adhere better to prostate stroma cells than LNCaP-mock cells. Figure 6A and B illustrates that more LNCaP-Core2GnT cells bound to stromal cells than LNCaP-mock cells with statistical significance. We then tested the adhesion of LNCaP-Core2GnT and LNCaP-mock cells to various molecules present in the extracellular matrix, which were coated on plates. As shown in Figure 6C, LNCaP-Core2GnT cells adhered more efficiently to type IV collagen and laminin than did LNCaP-mock cells. Statistical differences between two cell types were more significant for adhesion to type IV collagen and laminin than that for the others. These results combined indicate that LNCaP cells expressing core 2-branched *O*-glycans adhere more efficiently to prostate stroma cells than do LNCaP cells lacking core 2-branched *O*-glycans, most likely because of the increased adhesion to type IV collagen and laminin.

We next measured the cell growth after the cells were attached to collagen IV or laminin. The initial number of cells added was adjusted so that the numbers of initially attached cells were almost the same between LNCaP-Core2GnT and LNCaP cells. The results show that the adhesion to type IV collagen or laminin did not increase the cell growth compared with the adhesion to control plates, and that there is essentially no difference in growth of LNCaP and LNCaP-Core2GnT cells (Figure 6D).

Discussion

This study demonstrates that expression of Core2GnT on human prostate cancer cells correlates positively with the aggressive potential of prostate cancer. First, expression of Core2GnT in preoperative biopsy specimens is highly correlated with advanced stages of prostate cancer and is more prevalent in patients with higher GS. Second, the expression of Core2GnT in biopsy specimen predicts advanced disease. Although all of the clinical cases examined were organ confined, the final pathological stage was pT3 for 53% of Core2GnT-positive patients, in contrast with only 4% pT3 for Core2GnT-negative patients (Table I). Moreover, Core2GnT-positive patients exhibited a much less promising prognosis than Core2GnT-negative patients, suggesting that expression of Core2GnT is an independent predictor for recurrence of prostate cancer (Figure 2). These results indicate that expression of Core2GnT is highly correlated with tumor progression.

In this study, we also demonstrated that forced expression of Core2GnT in LNCaP prostate cancer cells led to significantly larger tumors upon inoculation to the mouse prostate. Significantly, this increase in tumor formation was achieved without expression of sialyl Lewis X or sialyl Lewis A capping structures. These results indicate that the

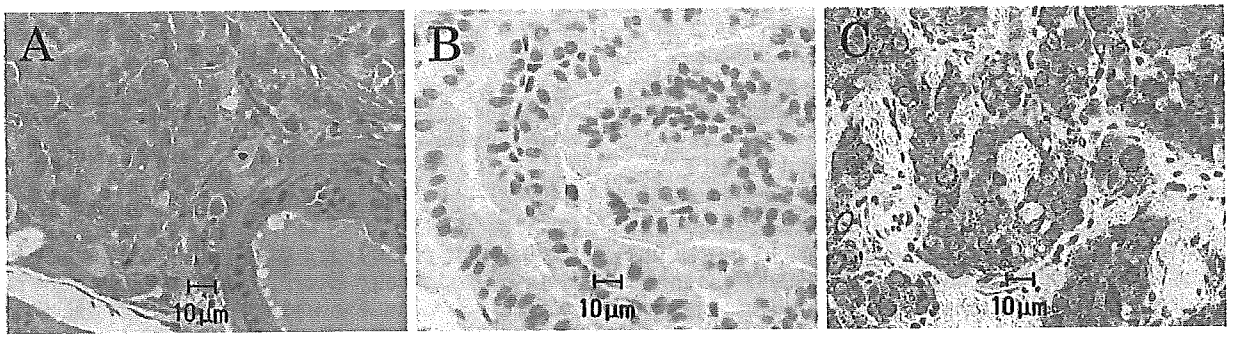

Fig. 5. Microscopic findings of the tumor formed by LNCaP cells in the prostate of mice. **(A)** Hematoxylin and eosin staining revealed the existence of tumor cells in the prostate of nude mice inoculated with LNCaP-core 2 β1,6-*N*-acetylglucosaminyltransferase-1 (LNCaP-Core2GnT). **(B)** Immunostaining using anti-Core2GnT antibodies was negative for mouse prostate. **(C)** Positive staining of Core2GnT was found in the tumor formed by LNCaP-Core2GnT. Scale bar represents 10 μm.

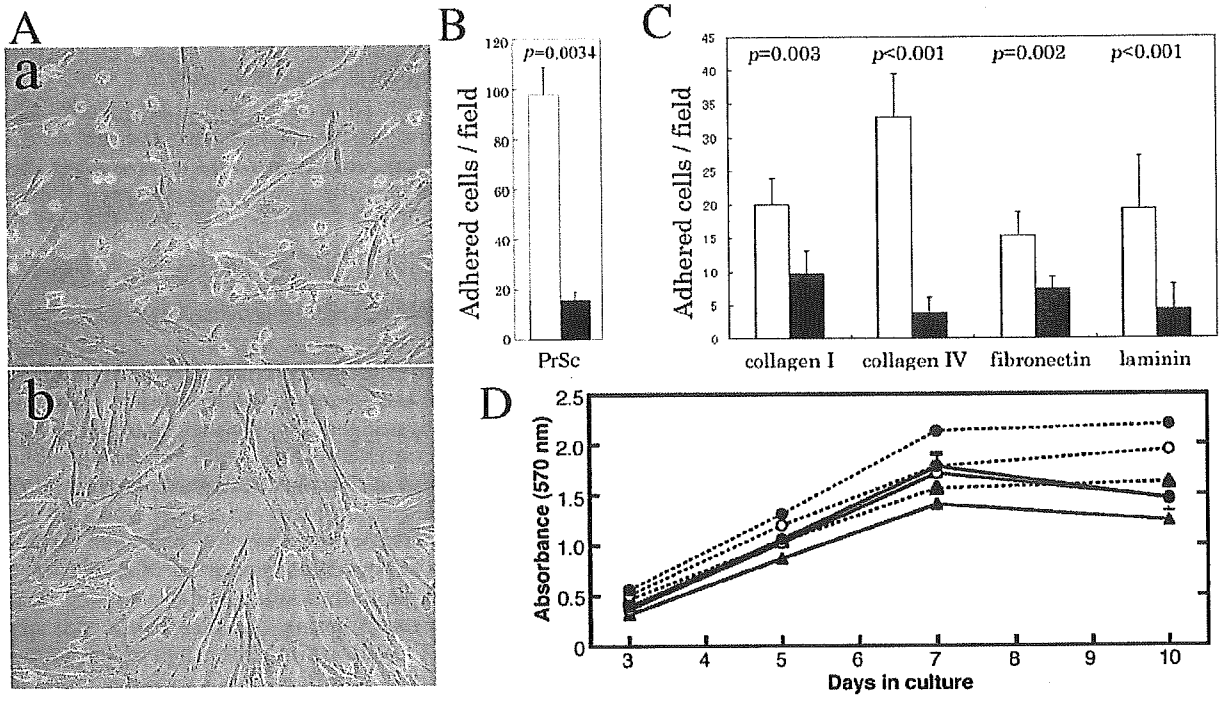

Fig. 6. Adhesion of LNCaP-core 2 β1,6-*N*-acetylglucosaminyltransferase-1 (LNCaP-Core2GnT) and LNCaP-mock cells to stromal cells and basement membrane components and its effect on in vitro growth. **(A)** Representative of photomicrographs of LNCaP-Core2GnT cells (a) and LNCaP-mock cells (b), adhered to prostate stromal cells. The long narrow cells are stromal cells, and the round small cells are LNCaP cells. **(B)** Number of LNCaP-Core2GnT (open column) and LNCaP-mock cells (closed column) adhered to prostate stromal cells (PrSc). **(C)** Adhesion of LNCaP-Core2GnT (open column) and LNCaP-mock (closed column) to various components of the basement membrane. All of these assays were carried out in triplicate. **(D)** Cell growth exhibited by mock-transfected LNCaP (solid lines) and LNCaP-Core2GnT (dotted lines) cells are shown. The cells were seed on control plates (closed circles), plates coated on laminin (open circles), and type IV collagen (closed triangles).

expression of core 2-branched *O*-glycans alone led to the increased tumor formation. Previously, it was shown in colon and lung carcinoma patients that expression of Core2GnT has a higher correlation with tumor progression than expression of sialyl Lewis X (Shimodaira *et al.*, 1997; Machida *et al.*, 2001). However, in those studies, expression of sialyl Lewis X was almost always associated with tumors, and the effect of core 2-branched oligosaccharides on tumor progression could not be separated from the effect of sialyl Lewis X capping structures on core 2-branched oligosaccharides. In this study employing LNCaP cells that lack α1,3/4-fucosyltransferase, the role of core 2-branched oligosaccharides was revealed in the absence of sialyl Lewis X and sialyl Lewis A.

Previously, it was suggested that core 2-branched *O*-glycans without sialyl Lewis X or sialyl Lewis A capping

structures can be recognized by carbohydrate-binding proteins distinct from selectins, such as galectins (Leffler and Barondes, 1986; Bourne *et al.*, 1994; Stowell *et al.*, 2004). Similarly, attachment of colonic carcinoma cells to liver sections was found to be dependent on core 2-branched *O*-glycans, but not on sialyl Lewis X (Ota *et al.*, 2000). It has been reported also that *N*-acetyllactosamine–galectin interaction may be impaired in *N*-acetylglucosaminyltransferase-V (GnT-V)-deficient mice and that such impairment causes a hyperimmune response in GnT-V knockout mice (Demetriou *et al.*, 2001). On the other hand, overexpression of core 2-branched *O*-glycans on T-lymphocytes leads to apoptosis, probably through binding to carbohydrate-binding proteins on opposing cells (Perillo *et al.*, 1995; Galvan *et al.*, 2000; Priatel *et al.*, 2000).

Based on the findings described above, it is possible to hypothesize that core 2-branched oligosaccharides on prostate epithelial cells are recognized by carbohydrate-binding proteins on stromal cells, and such binding may stimulate adhesion to stromal cells and thus tumor growth. Because galectins are shown to bind core 2-branched *O*-glycans, one of these carbohydrates-binding proteins is likely galectin. Indeed, the expression of galectins on the cell surface has been implicated in tumor metastasis (Raz and Lotan, 1987). As an alternative but not necessarily mutually exclusive possibility, proteins carrying core 2-branched oligosaccharides on prostate cancer cells may stimulate growth of prostate cancer cells after they acquire core 2-branched oligosaccharides. In relation to this hypothesis, it has been shown that overexpression of Core2GnT in PC12 cells results in increased phosphorylation of mitogen-activated protein kinase and c-fos promoter activation (Koya *et al.*, 1999). A recent study showed that CD43, a major glycoprotein carrying mucin-type *O*-glycans, regulates interleukin-2 production through its cytoplasmic tail, suggesting that intracellular signaling is a critical factor in the activity of CD43 (Tong *et al.*, 2004). Further studies are necessary to determine if any of these mechanisms operate in stimulating tumor formation after acquisition of Core2GnT by prostate cancer cells.

The present studies demonstrated that LNCaP cells expressing core 2-branched *O*-glycans adhere more efficiently to type IV collagen and laminin. However, cell growth of LNCaP cells on plates coated with collagen IV or laminin was not increased over control plates. Previously, it has been reported that the expression of the transcripts for type IV collagen α1 chain, laminin β1 chain, and S-laminin are increased in metastatic prostate cancer, compared with benign prostate glands (Pfohler *et al.*, 1998). It has also been reported that laminin-1 and α6β1 integrin were shown to facilitate normal morphogenesis of the prostate gland, acinal formation, whereas type IV collagen rather inhibits the normal morphogenesis (Bello-DeOcampo *et al.*, 2001). Because type IV collagen and laminin are components of basement membrane, they are likely deposited between prostate epithelial cells and stromal cells. These findings, together with the findings in this study, indicate that the adhesion to type IV collagen works more dominantly than the adhesion to laminin, preventing normal morphogenesis by prostate cancer cells. It is tempting to speculate that such effect may render those cancer cells undifferentiated, thus allowing them to be more neoplastic in nature.

Materials and methods

Patients

Between January 1994 and May 1999, 93 patients underwent staging pelvic lymphadenectomy and radical prostatectomy for clinically localized prostate cancer by a single surgeon (Y.A.) at Kurashiki Central Hospital, Kurashiki, Japan. Of these, 22 patients who had T3 disease or received neoadjuvant hormone therapy were excluded from the study, and two patients were excluded due to scarcity of biopsy specimens; the study sample thus consisted of 69 patients with clinically organ-confined (T1/T2) disease. The surgery was performed via the anatomical approach originally described by Walsh, (1992). No patients received adjuvant endocrine therapy until postoperative PSA failure. The ethical committee of Kurashiki Central Hospital approved the protocols for this study.

Preoperatively, all patients were evaluated by digital rectal examination, serum PSA, bone scan, pelvic computed tomography scan, and transrectal ultrasonography. Six to twelve prostate needle biopsy samples were obtained via ultrasound guidance by using an 18 G needle and evaluated for the presence of cancer lesion.

Pathologic diagnosis and Gleason scoring of prostate cancer

The 1997 American Joint Committee on Cancer (AJCC) staging system (Flemming *et al.*, 1997) was used to assign the stage, and the Gleason grading system (Gleason, 1966) was used for tumor grading. According to the latter grading system, the progression of prostate cancer was divided into five stages (fifth is the most advanced stage), and the score of the most frequently observed cancer cells and that of the second most frequently observed cells are obtained, for example, 4 + 3. The Gleason scoring of each specimen was made by a single pathologist (M.E.) who was unaware of the clinical data. Staining for Core2GnT was also compared with pT (representing the stage of pathological primary lesion); pT2 is defined as the primary lesion confined to inside the prostate, whereas pT3 corresponds to the primary lesion breaking through the capsule to infiltrate outside the prostate (Inaba *et al.*, 2003).

Measurement of PSA and follow-up of patients

Serum PSA levels were determined with IMx (Abbott Laboratories, Abbott Park, IL). Postoperative PSA values were considered elevated (PSA failure) if values of 0.1 ng/mL or greater were obtained on two consecutive visits 1 month apart. Time zero was defined as the day of surgery. Patients whose PSA level never became undetectable (<0.1 ng/mL) postoperatively were scored as failures at time zero.

Immunohistochemistry

Biopsy specimens were fixed with 10% buffered formalin for 1 h and embedded in paraffin. The samples were cut at 3 μm thickness and subjected to immunohistochemistry.

Deparaffinized specimens were incubated with rabbit anti-human Core2GnT polyclonal antibody (Skrincosky *et al.*, 1997) followed by goat anti-rabbit IgG antibodies conjugated with horseradish peroxidase (Nichirei, Tokyo, Japan). A control experiment was done by omitting the primary antibody from the staining procedure. The results of immunostaining were evaluated by persons unaware of the clinical data. Based on staining status of Golgi apparatus, specimens possessing 10% and more positive glands were judged Core2GnT positive.

Statistical analysis

The Chi-square test was used to assess the association of Core2GnT status with clinical and pathological parameters. PSA-free survival was evaluated by Kaplan-Meier curves. Differences between groups were evaluated using the log-rank test.

Stable transfectants

The human prostate cancer cell line LNCaP, purchased from the American Type Culture Collection (Rockville, MD), was maintained in RPMI 1640 medium containing 10% fetal calf serum. The cells were transfected with pcDNA3-Core2GnT using LipofectAMINE (Invitrogen, Carlsbad, CA), as described previously (Mitoma *et al.*, 2003). After G418 selection (200 μg/mL; Invitrogen), 20 single colonies were examined for immunocytochemical detection of Core2GnT using anti-Core2GnT antibody, and five stable transfectants expressing Core2GnT were isolated (LNCaP-Core2GnT). Among the stable transfectants, two clones (LNCaP-Core2GnT-1 and -2) were used in tumor assays. Because LNCaP-Core2GnT-1 and -2 yielded identical results in the following experiments, we designated LNCaP-Core2GnT-1 cells as LNCaP-Core2GnT cells, and the results obtained by LNCaP-Core2GnT cells were shown. As a control, LNCaP cells were transfected with pcDNA3 empty vector, and the transfected cells were designated LNCaP mock.

RT–PCR of Core2GlcNAcT-I

Chinese hamster ovary (CHO) cells were transfected with mammalian expression vectors encoding P-selectin glycoprotein ligand-1 (PSGL-1), fucosyltransferase-VII (FucT-VII) and Core2GnT establishing CHO-PSGL-1/F7/C2 and CHO-PSGL-1/F7 cells, respectively (Mitoma *et al.*, 2003; Kobayashi *et al.*, 2004). Total RNA was prepared using the TRIzol reagent (Invitrogen). Reverse transcriptase-polymerase chain reaction (RT–PCR) of glyceraldehyde-3-phosphate dehydrogenase and Core2GnT was performed, as described previously (Mitoma *et al.*, 2003). The enzymatic activity of Core2GnT was measured, as described previously (Bierhuizen and Fukuda, 1992; Skrincosky *et al.*, 1997), using Galβ1→3GalNAcα1→p-nitrophenol (Toronto Chemicals, Toronto, Canada) as a substance.

Fluorescence-activated cell sorting analysis

Cells grown to semiconfluency were dissociated into monodispersed cells using an enzyme-free cell dissociation solution (Hank's based), incubated with HECA-452 monoclonal antibody, followed by the second antibody and were subjected to fluorescence-activated cell sorting (FACS) analysis, as described previously (Mitoma *et al.*, 2003; Kobayashi *et al.*, 2004). HECA-452 antibody was used to detect both sialyl Lewis X and sialyl Lewis A antigen (Berg *et al.*, 1991; Kobayashi *et al.*, 2004).

In vitro cell proliferation

LNCaP-mock and LNCaP-Core2GnT cells were seeded onto 96-well plates at 10^5 cells/mL in RPMI 1640 containing 10% fetal calf serum and 200 μg/mL of G418 and cultured for various times. The number of living cells was measured every other day using the Cell Counting Kit (Wako Pure Chemical Industries, Tokyo, Japan). The cells (2×10^4 cells) were also seeded on 24-well plates coated with 10 μg of type IV collagen or laminin (Sigma, St. Louis, MO), and cell proliferation was measured as described (Mitoma *et al.*, 1998). Triplicate cultures were used for each sample.

Orthotopic tumor cell inoculation

Balb/c nude (nu/nu) mice (6- to 8-week-old males obtained from Clea Japan [Tokyo, Japan]) were used for orthotopic tumor cell injection after anesthetizing with pentobarbital. About 2×10^6 cells exhibiting >90% viability as judged by trypan blue staining were suspended in 20 μL of serum-free RPMI 1640 medium and inoculated into the posterior lobe of the prostate with pentobar, as described previously (Inaba *et al.*, 2003). After tumor cell inoculation, the wound was closed with surgical clips. Four weeks after inoculation, mice were killed, and prostates were removed followed by fixation with buffered formalin solution.

Adhesion assay

A prostate stromal cell line (Nishimura *et al.*, 1999) was cultured to 50% confluence in 24-well culture plates, as described previously (Krill *et al.*, 1997; Nishimura *et al.*, 1999). LNCaP-mock and LNCaP-Core2GnT cells were washed three times with phosphate buffered saline (PBS), harvested, resuspended in RPMI 1640 medium containing 1% fetal calf serum, and added on prostate stromal cells at the density of 10^5 cells/mL. The cells were incubated for 10 min at room temperature with continuous rotation at $1 \times g$. After washing with cold PBS, bound cells were counted under a light microscope obtaining mean cell numbers in 10 different fields. Alternatively, plates were precoated with various molecules present in the basement membrane (collagen I, collagen IV, fibronectin, and laminin, purchased from Sigma), which were dissolved at 100 μg/mL of PBS. After incubation for 30 min at room temperature and washing with PBS, the wells were then treated with 0.1% bovine serum albumin solution in PBS to prevent nonspecific adhesion, before addition of LNCaP cells.

Acknowledgments

The authors thank Drs. Michiko Fukuda and Junya Mitoma for useful discussion, Dr. Elise Lamar for critical reading of the manuscript, and Ms. Aleli Morse for organizing

the manuscript. This work was supported by NIH grant R37 CA 33000 (to M.F.), CREST from Japan Science and Technology Agency, grant B-16390459 from the Japan Society for the Promotion of Science (to C.O.), and grant B-15390115 from the Japan Society for the Promotion of Science (to J.N.).

Abbreviations

Core2GnT, core 2 β1,6-*N*-acetylglucosaminyltransferase-1; CHO, Chinese hamster ovary; FACS, fluorescence-activated cell sorting; GlcNAc, N-acetylglucosamine; PBS, phosphate buffered saline; PSA, prostate-specific antigen; PSGL-1, P-selectin glycoprotein ligand-1; RT-PCR, reverse transcriptase-polymerase chain reaction.

References

American Cancer Society. (2003) *Cancer Facts & Figures*. American Cancer Society, Atlanta, GA, pp. 1–21.

Bello-DeOcampo, D., Kleinman, H.K., Deocampo, N.D., and Webber, M.M. (2001) Laminin-1 and 61 integrin regulate acinar morphogenesis of normal and malignant human prostate epithelial cells. *Prostate*, **46**, 142–153.

Berg, E.L., Robinson, M.K., Mansson, O., Butcher, E.C., and Magnani, J.L. (1991) A carbohydrate domain common to both sialyl Le (a) and sialyl Le (X) is recognized by the endothelial cell leukocyte adhesion molecule ELAM-1. *J. Biol. Chem.*, **266**, 14869–14872.

Bierhuizen, M.F. and Fukuda, M. (1992) Expression cloning of a cDNA encoding UDP-GlcNAc: Gal β, 1–3-GalNAc-R. (GlcNAc to GalNAc) β1–6GlcNAc transferase by gene transfer into CHO cells expressing polyoma large tumor antigen. *Proc. Natl. Acad. Sci. U. S. A.*, **89**, 9326–9330.

Bourne, Y., Bolgiano, B., Liao, D.I., Strecker, G., Cantau, P., Herzberg, O., Feizi, T., and Cambillau, C. (1994) Crosslinking of mammalian lectin (galectin-1) by complex biantennary saccharides. *Nat. Struct. Biol.*, **1**, 863–870.

Brockhausen, I., Kuhns, W., Schachter, H., Matta, K.L., Sutherland, D.R., and Baker, M.A. (1991) Biosynthesis of *O*-glycans in leukocytes from normal donors and from patients with leukemia: increase in *O*-glycan core, 2 UDP-GlcNAc: Gal β 3GalNAc -R. (GlcNAc to GalNAc) (1–6) -*N*-acetylglucosaminyltransferase in leukemic cells. *Cancer Res.*, **51**, 1257–1263.

Demetriou, M., Granovsky, M., Quaggin, S., and Dennis, J.W. (2001) Negative regulation of T-cell activation and autoimmunity by Mgat5 *N*-glycosylation. *Nature*, **409**, 733–739.

Flemming, I.D., Copper, J.S., Henson, D.E., Hutter, R.V.P., Kennedy, B.J., and Murphy, G.P. (eds) (1997) *Joint Committee on Cancer Staging Manual*. JP Lippincott, Philadelphia, PA, pp. 219–222.

Fukuda, M. (1996) Possible roles of tumor-associated carbohydrate antigens. *Cancer Res.*, **56**, 2237–2244.

Galvan, M., Tsuboi, S., Fukuda, M., and Baum, L.G. (2000) Expression of a specific glycosyltransferase enzyme regulates T cell death mediated by galectin-1. *J. Biol. Chem.*, **275**, 16730–16737.

Gleason, D.F. (1966) Classification of prostatic carcinomas. *Cancer Chemother. Rep.*, **50**, 125–128.

Gronberg, H. (2003) Prostate cancer epidemiology. *Lancet*, **361**, 859–864.

Han, M., Partin, A.W., Pound, C.R., Epstein, J.I., and Walsh, P.C. (2001) Long-term biochemical disease-free and cancer-specific survival following anatomic radical retropubic prostatectomy. The 15-year Johns Hopkins experience. *Urol. Clin. North Am.*, **28**, 555–565.

Inaba, Y., Ohyama, C., Kato, T., Satoh, M., Saito, H., Hagisawa, S., Takahashi, T., Endoh, M., Fukuda, M.N., Arai, Y., and Fukuda, M. (2003) Gene transfer of 1,3-fucosyltransferase increases tumor growth of the PC-3 human prostate cancer cell line through enhanced adhesion to prostatic stromal cells. *Int. J. Cancer*, **107**, 949–957.

Iwai, T., Kudo, T., Kawamoto, R., Kubota, T., Togayachi, A., Hiruma, T., Okada, T., Kawamoto, T., Morozumi, K., and Narimatsu, H. (2005) Core 3 synthase is down-regulated in colon carcinoma and profoundly suppresses the metastatic potential of carcinoma cells. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 4572–4577.

Janvier, R., Sourla, A., Koutsilieris, M., and Doillon, C.J. (1997) Stromal fibroblasts are required for PC-3 human prostate cancer cells to produce capillary-like formation of endothelial cells in a three-dimensional co-culture system. *Anticancer Res.*, **17**, 1551–1557.

Kattan, M.W., Potters, L., Blasko, J.C., Beyer, D.C., Fearn, P., Cavanagh, W., Leibel, S., and Scardino, P.T. (2001) Pretreatment nomogram for predicting freedom from recurrence after permanent prostate brachytherapy in prostate cancer. *Urology*, **58**, 393–399.

Kobayashi, M., Mitoma, J., Nakamura, N., Katsuyama, T., Nakayama, J., and Fukuda, M. (2004) Induction of peripheral lymph node addressin in human gastric mucosa infected by Helicobacter pylori. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 17807–17812.

Koya, D., Dennis, J.W., Warren, C.E., Takahara, N., Schoen, F.J., Nishio, Y., Nakajima, T., Lipes, M.A., and King, G.L. (1999) Overexpression of core 2 *N*-acetylglycosaminyltransferase enhances cytokine actions and induces hypertrophic myocardium in transgenic mice. *FASEB J.*, **13**, 2329–2337.

Krill, D., Shuman, M., Thompson, M.T., Becich, M.J., and Strom, S.C. (1997) A simple method for the isolation and culture of epithelial and stromal cells from benign and neoplastic prostates. *Urology*, **49**, 981–988.

Kumamoto, K., Mitsuoka, C., Izawa, M., Kimura, N., Otsubo, N., Ishida, H., Kiso, M., Yamada, T., Hirohashi, S., and Kannagi, R. (1998) Specific detection of sialyl Lewis X determinant carried on the mucin GlcNAcβ1 – >6GalNAcα core structure as a tumor-associated antigen. *Biochem. Biophys. Res. Commun.*, **247**, 514–517.

Leffler, H. and Barondes, S.H. (1986) Specificity of binding of three soluble rat lung lectins to substituted and unsubstituted mammalian β-galactosides. *J. Biol. Chem.*, **261**, 10119–10126.

Lowe, B.J. (1994) Carbohydrate recognition in cell-cell interaction. In Fukuda, M. and Hindsgaul, O. (eds), Molecular Biology. Oxford Unversity Press, Oxford, pp. 163–205.

Machida, E., Nakayama, J., Amano, J., and Fukuda, M. (2001) Clinicopathological significance of core 2 β1,6-*N*-acetylglucosaminyltransferase messenger RNA expressed in the pulmonary adenocarcinoma determined by in situ hybridization. *Cancer Res.*, **61**, 2226–2231.

Mitoma, J., Furuya, S., and Hirabayashi, Y. (1998) A novel metabolic communication between neurons and astrocytes: non-essential amino acid L-serine released from astrocytes is essential for developing hippocampal neurons. *Neurosci. Res.*, **30**, 195–199.

Mitoma, J., Petryniak, B., Hiraoka, N., Yeh, J.C., Lowe, J.B., and Fukuda, M. (2003) Extended core 1 and core 2 branched *O*-glycans differentially modulate sialyl Lewis X-type L-selectin ligand activity. *J. Biol. Chem.*, **278**, 9953–9961.

Nakamori, S., Kameyama, M., Imaoka, S., Furukawa, H., Ishikawa, O., Sasaki, Y., Kabuto, T., Iwanaga, T., Matsushita, Y., and Irimura, T. (1993) Increased expression of sialyl Lewisx antigen correlates with poor survival in patients with colorectal carcinoma: clinicopathological and immunohistochemical study. *Cancer Res.*, **53**, 3632–3637.

Nishimura, K., Kitamura, M., Miura, H., Nonomura, N., Takada, S., Takahara, S., Matsumoto, K., Nakamura, T., and Matsumiya, K. (1999) Prostate stromal cell-derived hepatocyte growth factor induces invasion of prostate cancer cell line DU145 through tumor–stromal interaction. *Prostate*, **41**, 145–153.

Ohyama, C., Orikasa, S., Kawamura, S., Satoh, M., Saito, S., Fukushi, Y., Levery, S.B., and Hakomori, S. (1995) Galactosylgloboside expression in seminoma. Inverse correlation with metastatic potential. *Cancer*, **76**, 1043–1050.

Ota, M., Takamura, N., and Irimura, T. (2000) Involvement of cell surface glycans in adhesion of human colon carcinoma cells to liver tissue in a frozen section assay: role of endo-β-galactosidase-sensitive structures. *Cancer Res.*, **60**, 5261–5268.

Partin, A.W., Kattan, M.W., Subong, E.N., Walsh, P.C., Wojno, K.J., Oesterling, J.E., Scardino, P.T., and Pearson, J.D. (1997) Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. *JAMA*, **277**, 1445–1451.

Perillo, N.L., Pace, K.E., Seilhamer, J.J., and Baum, L.G. (1995) Apoptosis of T cells mediated by galectin-1. *Nature*, **378**, 736–739.

Pfohler, C., Fixemer, T., Jung, V., Dooley, S., Remberger, K., and Bonkhoff, H. (1998) In situ hybridization analysis of genes coding collagen IV α1 chain, laminin, 1 chain, and S-laminin in prostate tissue and prostate cancer: increased basement membrane gene expression in high-grade and metastatic lesions. *Prostate*, **36**, 143–150.

Piller, F., Le Deist, F., Weinberg, K.I., Parkman, R., and Fukuda, M. (1991) Altered *O*-glycan synthesis in lymphocytes from patients with Wiskott–Aldrich syndrome. *J. Exp. Med.*, **173**, 1501–1510.

Priatel, J.J., Chui, D., Hiraoka, N., Simmons, C.J., Richardson, K.B., Page, D.M., Fukuda, M., Varki, N.M., and Marth, J.D. (2000) The ST3Gal-I sialyltransferase controls CD8+ T lymphocyte homeostasis by modulating *O*-glycan biosynthesis. *Immunity*, **12**, 273–283.

Quinn, M. and Babb, P. (2002) Patterns and trends in prostate cancer incidence, survival, prevalence and mortality. Part I: international comparisons. *BJU Int.*, **90**, 162–173.

Raz, A. and Lotan, R. (1987) Endogenous galactoside-binding lectins: a new class of functional tumor cell surface molecules related to metastasis. *Cancer Metastasis Rev.*, **6**, 433–452.

Roberts, S.G., Blute, M.L., Bergstralh, E.J., Slezak, J.M., and Zincke, H. (2001) PSA doubling time as a predictor of clinical progression after biochemical failure following radical prostatectomy for prostate cancer. *Mayo Clin. Proc.*, **76**, 576–581.

Ross, P.L., Scardino, P.T., and Kattan, M.W. (2001) A catalog of prostate cancer nomograms. *J. Urol.*, **165**, 1562–1568.

Saito, T., Miyoshi, E., Sasai, K., Nakano, N., Eguchi, H., Honke, K., and Taniguchi, N. (2002) A secreted type of β1,6-*N*-acetylglucosaminyltransferase V (GnT-V) induces tumor angiogenesis without mediation of glycosylation: a novel function of GnT-V distinct from the original glycosyltransferase activity. *J. Biol. Chem.*, **277**, 17002–17008.

Saitoh, O., Piller, F., Fox, R.I., and Fukuda, M. (1991) T-lymphocytic leukemia expresses complex, branched *O*-linked oligosaccharides on a major sialoglycoprotein, leukosialin. *Blood*, **77**, 1491–1499.

Schachter, H. and Brockhausen, I. (1992) The biosynthesis of serine (threonine)-*N*-acetylgalactosamine-linked carbohydrate moieties. In Allen, H.J. and Kisailus, E.C. (eds), *Glycoconjugates: Composition, Structure and Function*. Marcel Dekker, New York, pp. 263–332.

Shimodaira, K., Nakayama, J., Nakamura, N., Hasebe, O., Katsuyama, T., and Fukuda, M. (1997) Carcinoma-associated expression of core 2 β-1,6-*N*-acetylglucosaminyltransferase gene in human colorectal cancer: role of *O*-glycans in tumor progression. *Cancer Res.*, **57**, 5201–5206.

Shimono, R., Mori, M., Akazawa, K., Adachi, Y., and Sgimachi, K. (1994) Immunohistochemical expression of carbohydrate antigen 19–9 in colorectal carcinoma. *Am. J. Gastroenterol.*, **89**, 101–105.

Skrincosky, D., Kain, R., El-Battari, A., Exner, M., Kerjaschki, D., and Fukuda, M. (1997) Altered Golgi localization of core 2 β-1,6-*N*-acetylglucosaminyltransferase leads to decreased synthesis of branched *O*-glycans. *J. Biol. Chem.*, **272**, 22695–22702.

Song, Z., Powell, W.C., Kasahara, N., van Bokhoven, A., Miller, G.J., and Roy-Burman, P. (2000) The effect of fibroblast growth factor, 8, isoform b, on the biology of prostate carcinoma cells and their interaction with stromal cells. *Cancer Res.*, **60**, 6730–6736.

Stowell, S.R., Dias-Baruffi, M., Penttila, L., Renkonen, O., Nyame, A.K., and Cummings, R.D. (2004) Human galectin-1 recognition of poly-*N*-acetyllactosamine and chimeric polysaccharides. *Glycobiology*, **14**, 157–167.

Tong, J., Allenspach, E.J., Takahashi, S.M., Mody, P.D., Park, C., Burkhardt, J.K., and Sperling, A.I. (2004) CD43 regulation of T cell activation is not through steric inhibition of T cell–APC interactions but through an intracellular mechanism. *J. Exp. Med.*, **199**, 1277–1283.

Tuxhorn, J.A., McAlhany, S.J., Dang, T.D., Ayala, G.E., and Rowley, D.R. (2002) Stromal cells promote angiogenesis and growth of human prostate tumors in a differential reactive stroma (DRS) xenograft model. *Cancer Res.*, **62**, 3298–3307.

Walsh, P.C. (1992) *Campbell's Urology*. W.B. Saunders, Philadelphia, PA.

Watanabe, R., Ohyama, C., Aoki, H., Takahashi, T., Satoh, M., Saito, S., Hoshi, S., Ishii, A., Saito, M., and Arai, Y. (2002) Ganglioside G (M3) overexpression induces apoptosis and reduces malignant potential in murine bladder cancer. *Cancer Res.*, **62**, 3850–3854.

Yousefi, S., Higgins, E., Daoling, Z., Pollex-Kruger, A., Hindsgaul, O., and Dennis, J.W. (1991) Increased UDP-GlcNAc: Gal β1–3GaLNAc-R. (GlcNAc to GaLNAc) β-1, 6-*N*-acetylglucosaminyltransferase activity in metastatic murine tumor cell lines. Control of polylactosamine synthesis. *J. Biol. Chem.*, **266**, 1772–1782.

 DOI 10.1002/pmic.200401286

REGULAR ARTICLE

Protein expression pattern distinguishes different lymphoid neoplasms

Kazuyasu Fujii[1, 2], Tadashi Kondo[1], Hideki Yokoo[1], Tesshi Yamada[1], Yoshihiro Matsuno[3], Keiji Iwatsuki[2] and Setsuo Hirohashi[1]

[1] Cancer Proteomics Project, National Cancer Center Research Institute, Tokyo, Japan
[2] Department of Dermatology, Okayama University Graduate School of Medicine and Dentistry, Okayama, Japan
[3] Pathology Division, National Cancer Center Research Institute, Tokyo, Japan

To identify proteins associated with the histological subtypes of lymphoid neoplasms, we studied the proteomes of 42 cell lines from human lymphoid neoplasms including Hodgkin's lymphoma (HL; four cell lines), B cell malignancies (19 cell lines), T cell malignancies (16 cell lines), and natural killer (NK) cell lymphoma (three cell lines). The protein spots were sequentially selected by (i) Wilcoxon or Kruskal–Wallis tests to find the spots whose intensity was significantly ($p < 0.05$) different among the cell line groups, (ii) by statistical-learning methods to prioritize the spots according to their contribution to the classification, and (iii) by unsupervised classification methods to validate the classification robustness by the selected spots. The selected spots discriminated (i) between HL cells and other cells, (ii) between the cells from B cell malignancies, T cell malignancies, and NK cell lymphoma cells, and (iii) between HL cells and anaplastic large cell lymphoma cells. Among the 31 informative protein spots, MS identified 24 proteins corresponding to 23 spots. Previous reports did not correlate these proteins to lymphocyte differentiation, suggesting that a proteomic study would identify the novel mechanisms responsible for the histogenesis of lymphoid neoplasms. These proteins may have potential as differential diagnostic markers for lymphoid neoplasms.

Received: May 9, 2004
Revised: February 9, 2005
Accepted: February 11, 2005

Keywords:
2D-DIGE / Lymphoid neoplasm / Multivariate analysis / Statistical-learning method

1 Introduction

The recent development of proteomic technologies allows us to describe more than 1000 features of protein expression in a quantitative and reproducible way. The proteomic approach has been used to address many biological questions, and has wide practical utility in the field of medicine [1]. Because the final outcome of genetic and epigenetic events is the content of proteins in the cell, proteomics is conceptually the best and the most accurate and valuable modality for studying the mechanisms of biological phenotypes. In the post-genome era, genome databases coupled with modern high-throughput screening systems, including DNA microarrays and SAGE, have enabled a comprehensive understanding of mRNA expression [2]. However, many lines of evidence suggest that mRNA expression reflects protein expression for only a small proportion of genes [3–6]. Further, PTMs such as phosphorylation and glycation cannot be predicted by measuring the amount of mRNA or by studying nucleic acid sequences. These difficulties underline the potential advantages of proteomic over transcriptomic approaches. However, comprehensive profiling of all proteins expressed in cells has not yet been achieved, and proteomic technologies equivalent to DNA microarray and SAGE are not presently available. For example, many key proteins present in low amounts, such as kinases and oncogene products, are difficult to observe by

Correspondence: Dr. Tadashi Kondo, Cancer Proteomics Project, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
E-mail: takondo@gan2.res.ncc.go.jp
Fax: +81-3-3547-5298

Abbreviations: ALCL, anaplastic large cell lymphoma; **HL**, Hodgkin's lymphoma; **NK**, natural killer

conventional 2-D PAGE. Antibody array may facilitate expression studies, but its utility is limited by the availability of antibodies suitable for the experiments. Therefore, intense efforts are now underway to develop proteomic technologies able to reveal a greater fraction of the proteome, both by modifying present protocols and by developing novel approaches [7–9]. Moreover, as is the case for transcriptomic studies, data-mining techniques, including multivariate analysis and statistical-learning methods, have been used to analyze proteomic data from malignant tumors to find the proteins and the protein networks responsible for the clinicopathological features of cancer [10]. Such use of proteomics will be expanded by the development of new methods.

The differential diagnosis of many lymphoma types has been achieved by morphological findings and immunophenotyping assisted by karyotyping or molecular analysis of specific gene rearrangements [11–14]. However, the subtypes of some lymphomas resemble more than one disease entity [15], and biomarkers for better differential diagnosis have been developed using DNA microarray technologies [16]. Although such studies have successfully identified candidate gene sets, we are still far from a complete description of the biology of lymphoma and therefore other approaches are needed to overcome some of the limitations of present studies.

Here, to demonstrate the utility of the proteomic approach in identifying tumor-specific markers, we used 2-D DIGE and multivariate methods to study a series of lymphoma cell lines. 2-D DIGE captured 83 977 features of the proteome across 42 cell lines from lymphoid neoplasms. This is the first systematic large-scale proteomic study to identify proteins associated with lymphoid neoplasms. Statistical-learning methods were used to distinguish a small number of candidate protein spots specific for the original histology, and subsequent MS study identified the proteins corresponding to the spots.

2 Materials and methods

2.1 Cell lines and protein extraction

The cell lines from human lymphoid neoplasms and their characteristics are summarized in Table 1 [17–55]. They were maintained until use with the culture medium recommended by the distributors. For protein extraction, the cells (5×10^7–1×10^8) were pelleted by centrifugation at 3000 rpm for 5 min and washed three times with PBS. The cells were treated with 10% TCA for 30 min, and then resuspended in urea lysis buffer containing 7 M urea, 2 M thiourea, 3% CHAPS, 1% Triton X-100 for 30 min (1 mL of urea lysis buffer per 100 mg wet weight cells). The cell lysate was centrifuged at 15 000 rpm for 30 min and the supernatant (cellular protein fraction) was recovered. The protein concentration was measured with a Protein Assay Kit (Bio-Rad) and adjusted to 1 mg/mL with urea lysis buffer. All procedures were performed on ice.

2.2 Fluorescence labeling and sample preparation for 2-D DIGE

First, portions of the protein samples from all cell lines were mixed to generate a reference sample. Protein samples from each cell line and the reference sample were then labeled with Cy5 and Cy3, respectively. In brief, the protein sample was adjusted to pH 8.5 with 30 mM Tris. Then, the samples were mixed with 200 pmol of fluorescent dye for 30 min. The labeling reaction was terminated by incubation with 10 mM lysine for 10 min. Urea lysis buffer containing 130 mM DTT and 2% ampholyte was added and incubated for 15 min. Equal amounts of Cy3- and Cy5-labeled protein samples were mixed, and the total volume was adjusted to 1680 µL with urea lysis buffer containing 65 mM DTT and 1% ampholyte. The labeled protein sample was applied to four separate 2-D PAGE gels. All labeling procedures were performed on ice and in the dark.

2.3 2-D PAGE

Proteins were separated by IEF and subsequently by SDS-PAGE, according to our previous report [56]. IEF was performed using linear IPGs (Immobiline Dry Strip, length 24 cm, p*I* range between 4.0 and 7.0;. Amersham Biosciences). The IPGs were rehydrated with 420 µL of protein sample at 20°C for 12 h. Proteins were separated on the basis of the p*I* with a total of 96 kVh on an IPGphor (Amersham Biosciences). The IPGs were equilibrated for 20 min in equilibration buffer containing 3 M urea, 50 mM Tris (pH 8.8), 30% glycerol, 1.0% SDS, and 16 mM DTT, and then for 20 min in the same buffer containing 122 mM iodoacetamide instead of DTT. After equilibration, the IPGs were placed on 9–15% gradient polyacrylamide gels and the proteins were separated at 20°C for 15 h at a constant wattage of 17 W per 12 gels. All electrophoresis procedures were performed in the dark.

2.4 Image analysis

After the second-dimension electrophoresis, the gel was scanned at appropriate wavelengths for Cy3 and Cy5 with a MasterImager 2640 (Amersham Biosciences). A pair of Cy3 and Cy5 images was stored as a single apf file and transformed to a pair of tiff files. The relative spot intensity was calculated as the ratio between the absolute intensity of an individual spot and the total intensity of the gel. The relative spot intensity of the Cy5 image was normalized by that of the Cy3 image obtained from the same gel. Since the Cy3 image was generated from the reference sample, a common mixture of all cell lines, this normalization procedure compensates for gel-to-gel variations. Each sample was run on four gels, and average spot intensities were calculated for quantitative comparison.

Table 1. Characters of cell lines, culture conditions, and number of analyzed spots

Cell line		Cell line		Culture medium[d]	Number of spots[e]	Spots in 80%	
number[a]	Cell origin[b]	name	Source[c]			Cy3-image[f]	References
	HL						
1	HL	KM-H2	DSMZ	1	2006	1597	17
2	HL	HDLM2	DSMZ	2	1925	1538	18
3	HL	L-428	DSMZ	1	1944	1571	19
4	HL	HD-MY-Z	DSMZ	1	2018	1579	20
	B cell						
5	BL	Namalwa	Human-Science	1	2016	1595	21
6	BL	EB-3	Human-Science	1	1986	1543	22
7	BL	RAMOS	Human-Science	1	2044	1611	23
8	BL	Raji	Tohoku	1	1850	1474	24
9	BL	TL-1	Tohoku	1	2002	1574	25
10	BL	Daudi	Tohoku	1	1942	1509	26
11	BL	HS-sultan	Human-Science	1	1998	1579	27, 28
12	B-ALL	RS4;11	DSMZ	3	2013	1587	29
13	B-ALL	REH	DSMZ	1	2066	1597	30
14	B-ALL	NALM-6	DSMZ	1	1989	1568	31
15	DLBCL	DB	ATCC	4	1984	1575	32
16	DLBCL	Toledo	ATCC	4	1994	1598	33
17	DLBCL	Pfeiffer	ATCC	4	2042	1605	33
18	DLBCL	KARPAS-422	DSMZ	1	1930	1538	34
19	DLBCL	OCI-LY-19	DSMZ	3	1958	1578	35
20	FL	DOHH-2	DSMZ	1	2485	1650	36
21	FL	SU-DHL-4	DSMZ	1	1979	1577	37
22	PCL	KARPAS-620	DSMZ	2	2029	1570	38
23	PCL	SK-MM2	DSMZ	1	2068	1592	39
	T cell						
24	T-ALL	JURKAT	Tohoku	1	1913	1493	40
25	T-ALL	PEER	Tohoku	1	2018	1591	41
26	T-ALL	CCRF-CEM	Tohoku	1	2012	1589	42
27	T-ALL	Molt3	Human-Science	1	1948	1570	43
28	T-ALL	Molt4	Human-Science	1	1937	1556	43
29	T-ALL	TALL-1	Human-Science	1	1996	1609	44
30	T-ALL	CCRF-HSB2	Human-Science	1	2000	1609	45
31	ALCL	SU-DHL-1	DSMZ	1	1952	1563	46
32	ALCL	Karpas299	DSMZ	1	1938	1568	47
33	ALCL	SR-786	DSMZ	5	1982	1595	48
34	ALCL	SUP-M2	DSMZ	1	2014	1593	49
35	ATL	ILT-Mat	Tohoku	6	1983	1565	50
36	ATL	TL-SU	Tohoku	1	2018	1550	40
37	ATL	TL-Hir	Tohoku	1	2006	1594	51
38	CTCL	Hut78	Tohoku	1	1912	1537	52
39	CTCL	Hut102	Tohoku	6	1978	1597	52
	NK cell						
40	NK cell lymphoma	KHYG-1	Human-Science	6	2085	1592	53
41	NK cell lymphoma	NK92	ATCC	7	1977	1569	54
42	NK cell lymphoma	KAI3	Human-Science	6	2040	1598	55

a) Cell line numbers refer to those in Figs. 3 and 4

b) Cell origin as described in the indicated references. HL, Hodgkin's lymphoma; BL, Burkitt's lymphoma; B-ALL, B cell acute lymphoblastic leukemia; DLBCL, diffuse large B cell lymphoma; FL, follicular lymphoma; PCL, plasmacytoma; T-ALL, T cell acute lymphoblastic leukemia; ALCL, anaplastic large cell lymphoma; ATL, adult T cell leukemia; CTCL, cutaneous T cell lymphoma

c) Source of cell lines: DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Braunschweig, Germany); Human Science (Osaka, Japan); Tohoku, Institute of Development, Aging and Cancer, Tohoku University (Miyagi, Japan); ATCC, American Type Culture Collection (Manassas, VA)

d) Culture media are those recommended by suppliers. Culture media: 1, RPMI 1640 90% + FBS 10%; 2, RPMI 1640 80% + FBS 20%; 3, alpha-MEM 90% + FBS 10%; 4, RPMI 1640 90%, FBS 10%, 10 mM HEPES, 1 mM sodium pyruvate, 4.5 g/L glucose, 1.5 g/L sodium bicarbonate; 5, RPMI 1640 85% + FBS 15%; 6, RPMI 1640 90%, FBS 10%, IL-2 100 U/mL; 7, alpha-MEM 75%, FBS 12.5%, horse serum 12.5%, IL-2 100 U/L, 0.2 mM inositol, 0.1 mM 2-ME, 0.02 mM folic acid

e) Total number of spots detected by DeCyder software

f) Number of spots present in more than 80% of Cy3 images

2.5 Data-mining process

For scatter plot analysis, proteins were extracted three times from KAI3 cells and labeled with Cy5 fluorescent dye as described in Section 2.2. They were mixed with the Cy3-labeled reference sample and separated on quadruplicate 2-D PAGE gels. The normalized and averaged spot intensities were transformed logarithmically and subjected to scatter plot analysis. Overall correlation of the expression patterns across the 42 cell lines was monitored by a correlation matrix.

The Wilcoxon or Kruskal–Wallis test was employed to identify the protein spots that discriminated between the cell line groups with high confidence. The class prediction system, having learned the expression features of the groups, was then used to classify samples. Classification accuracy of various spot sets was assessed by leave-one-out cross-validation. In the leave-one-out cross-validation, each cell line sample was left out in turn, and the prediction model with the discriminator spot set was constructed using the remaining cell line samples. The cell line sample left out was then used as the test case to evaluate the accuracy of the class prediction. This process was repeated by reducing lower-ranked spots for all samples. For spot ranking, various algorithms were used. The lower-ranked spots were removed and the classification accuracy was calculated again. This process was repeated and the cross-validation error rate was plotted as a function of spot numbers. These analyses were performed using Impressionist (Gene Data, Basel, Switzerland).

GeneMaths (Applied Maths, Sint-Martens-Latem, Belgium) was used for hierarchical clustering analysis by measuring Euclidian distance as a similarity coefficient and by Ward's algorithm as a tree-building method.

2.6 In-gel digestion and MS study

Protein identification was performed as described previously [57]. In brief, preparative gels were made by loading 500 μg of non-labeled protein sample and the gels were stained with SYPRO Ruby. The gel image was acquired by scanning the gels at the appropriate wavelength for SYPRO Ruby with a MasterImager 2640 (Amersham Biosciences). Spots were marked with the DIA mode of DeCyder software and recovered with an automated spot collector (Spot Picker, Amersham Biosciences). The recovered gel plugs were washed extensively with 50 mM ammonium bicarbonate and air-dried. The protein in the gel was treated with 200 ng of TPCK-treated trypsin at 37°C overnight. The trypsin-digested peptides were recovered by incubation with 50% ACN/0.1% TFA and mixed with an equal volume of matrix solution, dihydroxybenzoic acid. PMF and MS/MS analysis were performed on a Q-Star Pulser-i equipped with the oMALDI ion source (Applied Biosystems, Foster City, CA, USA). The results of identification by PMF and MS/MS were scored with the Analyst QS and Mascot programs, respectively, and the top-scoring gene products were considered to be the corresponding proteins.

3 Results

3.1 Protein expression profiles of lymphoma cell lines

We designed the experiment so that the 2-D image of each cell line was normalized with respect to the common image in the same gel to avoid gel-to-gel variations resulting from electrophoresis. The reference sample and the experimental samples were labeled with Cy3 and Cy5, respectively. The Cy3-labeled reference sample and each Cy5-labeled sample were mixed together and co-separated by 2-D PAGE. As all gels produce a common Cy3 image of the reference sample, standardization of the spot intensity of the Cy5 image to that of the Cy3 image can compensate for gel-to-gel variation, allowing quantitative and reproducible study of protein expression. This approach enables semi-quantitative comparison of proteins between the samples. Based on the relative intensity of the spots, we estimated the degrees of both up- or down-regulation of proteins among samples. Because the absolute amounts of proteins could be measured from the fluorescence intensity using the standard curve, quantitative inter-protein comparison would have been possible. However, we did not attempt this in the present study. An example of a two-channel 2-D image is shown in Fig. 1A. The blue Cy5 image of KAI3 cells is merged with the red Cy3 image of the reference sample in the same gel. As the reference sample contains protein extracts from all cell lines, including KAI3, the spots on the Cy3 image include those of KAI3 cells. The location of the spots is almost the same in the two images because the two samples were co-separated in the same gel and the electrophoretic properties of Cy3- or Cy5-labeled proteins are designed to be almost identical.

The reproducibility of the protein expression profile was evaluated by scatter plot analysis. Three independent 2-D DIGE separations of the protein extract of KAI3 cells were performed, and the correlation of spot intensity was examined. The scattergram in Fig. 1B shows the high reproducibility of spot intensity: in all pairs of experiments, the intensity of at least 98.23% of spots was distributed within a two-fold difference and the correlation coefficient was at least 0.7176.

The overall correlation of spot intensities across the 42 cell lines is shown in Fig. 1C. The degree of correlation of protein expression patterns was demonstrated by measuring correlation coefficients, indicating that cell lines derived from lymphoid neoplasms with similar phenotypes had common protein expression profiles. In particular, the cell lines in the B cell group showed more homogeneous protein expression profiles, perhaps reflecting the biological variations of lymphocyte lineages: B cells differentiate into plasmacytes and are recognized by CD19 throughout the B cell lineage until plasma cell differentiation. On the other hand, a correlation matrix revealed that cell lines of the T cell group were relatively heterogeneous in terms of their protein expression profiles, perhaps reflecting the clinical features of

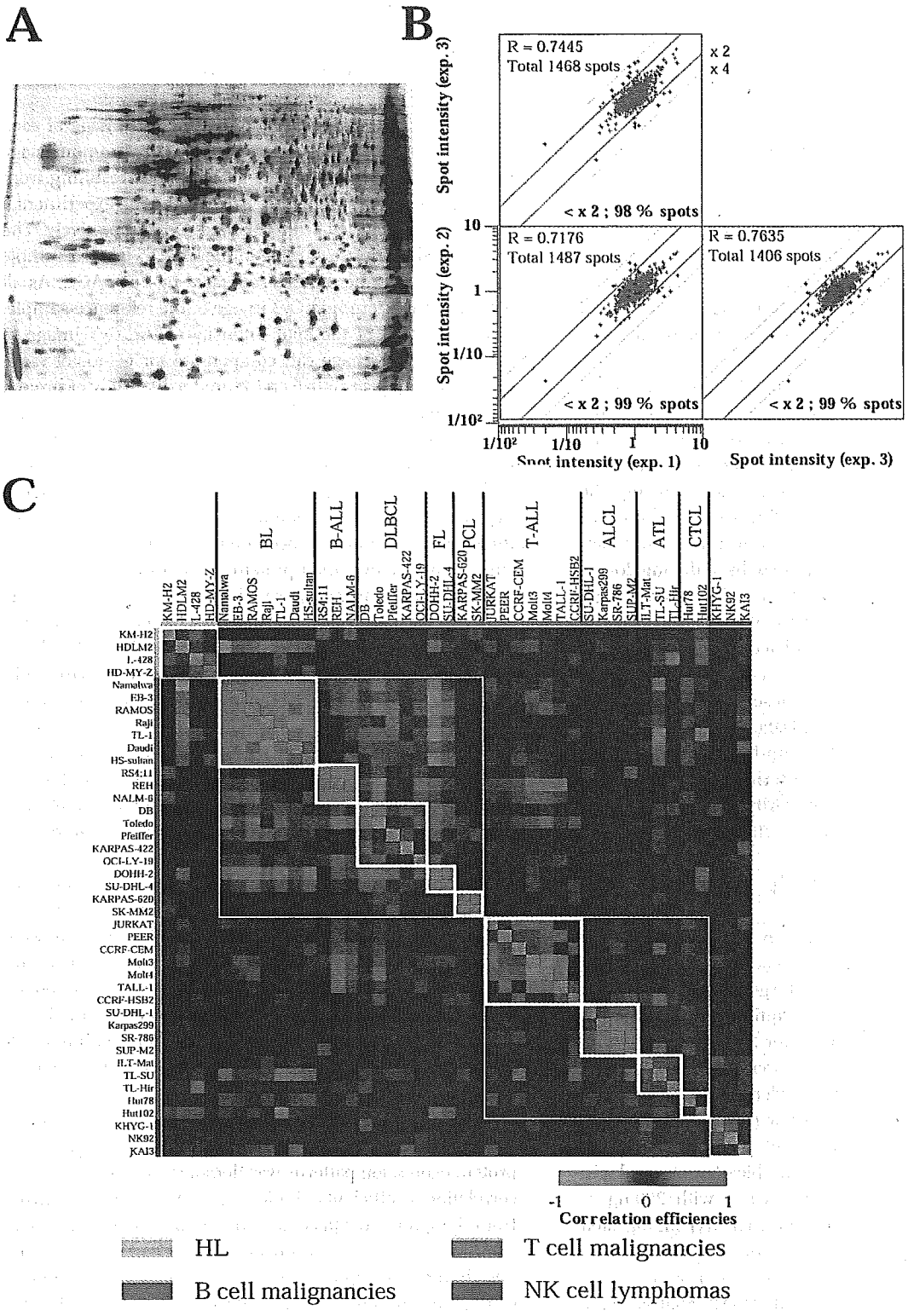

Figure 1. Quantitative and reproducible protein expression profiles. (A) Two-color image of a 2-D profile. A Cy5 image of KAI3 cells (blue) was merged with a Cy3 image of the reference sample (red). (B) Scattergram of expression profile of KAI3 cells. Comparison of data from three independent experiments revealed the high reproducibility of protein expression profiling. (C) Correlation matrix of cell lines based on protein expression profiles. The similarity of profiles is shown by color. The boxed matrixes indicate the combinations of cells having the same histological subtype.

T cell lymphoma; one-third of T cell lymphoma arise, spread to, or relapse at extranodal sites, and the site of disease is important for disease definition [58].

3.2 Strategy of proteomic analysis

The DeCyder software merged and quantified 83 977 protein spots in 168 gels representing the 42 cell lines (Fig. 2). Among them, 66 143 spots having valid values in at least 80% of Cy3 images were retained for further analysis. The spots with an expression level statistically and significantly different between particular cell line groups were selected with the Wilcoxon or Kruskal–Wallis tests ($p < 0.05$). We grouped and compared the 42 cell lines as follows: (i) Hodgkin's lymphoma (HL) cells *versus* other cells, (ii) cells from B cell malignancies *versus* cells from T cell malignancies *versus* cells from natural killer (NK) cell lymphoma, (iii) HL-cells *versus* anaplastic large cell lymphoma (ALCL) cells. Various numbers of spots were selected depending on the sample sets: 197 spots for discrimination between HL cells and other cells, 251 spots for discrimination between the cells from B cell malignancies, T cell malignancies and NK cell lymphoma, and 84 spots for discrimination between HL cells and ALCL cells. On the basis of the leave-one-out cross-validation error rate, we prioritized the protein spots according to their contribution to the classification: the protein spot sets that minimized the classification error rate were chosen as the best. The robustness of the discrimination potential of these spot sets was further evaluated by unsupervised classification methods, including principal component analysis and hierarchical clustering analysis. Among the selected spot sets, those that clearly categorized the cell lines into groups by these methods were finally considered as the most informative for the classification. The proteins corresponding to these spots were identified by MS.

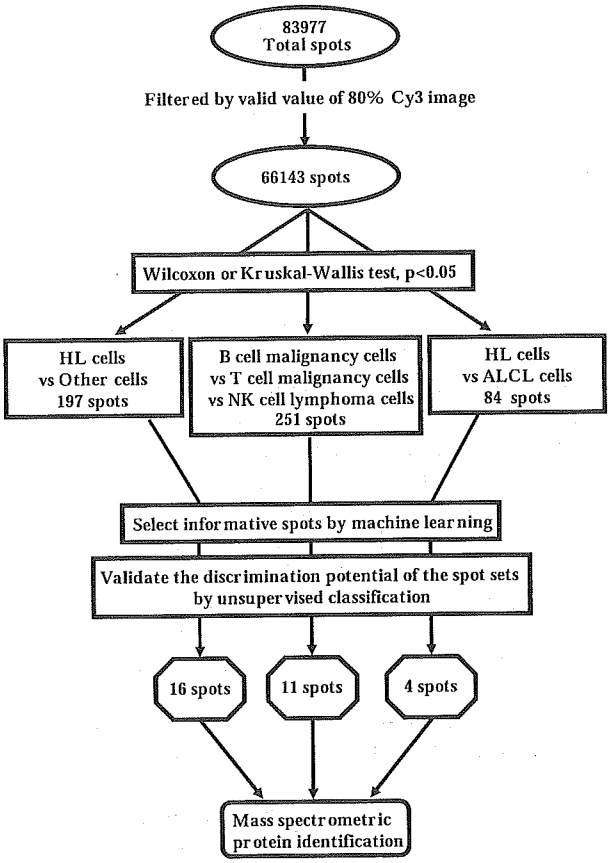

Figure 2. Schematic diagram of the data analysis strategy in this study. The total number of spots recognized and quantified by the DeCyder software was initially 83 977. Spots were filtered out if they appeared in less than 80% of Cy3 images. The Wilcoxon or Kruskal–Wallis test ($p < 0.05$) was used to identify the differentially expressed spots, and machine-learning methods were used to select the informative spot sets. Unsupervised classification methods were used to evaluate the discrimination potential of the spots. The proteins corresponding to those spots were identified by MS.

3.3 Multivariate analysis of cell lines

We illustrate in Fig. 3 the process of spot selection using multivariate analysis and statistical-learning methods for discrimination between HL cells and other cells. The Wilcoxon test selected 197 spots whose intensity was significantly different ($p < 0.05$) between the two groups. Among the spot ranking methods, the plot of the leave-one-out cross-validation error rate revealed that Fisher linear discriminant analysis identified minimal spot sets including 16 or 32 spots by which HL cells could be discriminated from other cells with the lowest cross-validation error rates (Fig. 3A). Although the other three algorithms also yielded spot sets that could segregate the cell line groups, more spots were required to obtain the lowest cross-validation error rate or the minimal error rate with those algorithms was higher than that by sparse linear discriminant analysis (Fig. 3A). The robustness of the discrimination potential of the two spot sets was examined with unsupervised methods. Using the 16 spots, principal component analysis (Fig. 3B) and hierarchical clustering analysis (Fig. 3C) clearly divided the cells into two groups. On the other hand, the unsupervised study using 32 spots resulted in ambiguous classification (data not shown). Therefore, we concluded that the set containing 16 spots was most informative for discrimination between HL cells and other cells.

We performed the same procedure for the other sets of the cell line groups, comparing protein expression profiles between (i) B cell malignancy cells, T cell malignancy cells and NK cell lymphoma cells (Fig. 4A) and (ii) HL cells *versus* ALCL cells (Fig. 4B). The Wilcoxon test was used for discrimination between two groups and the Kruskal–Wallis test was used for three groups. The results of unsupervised classification analysis of the cell lines using the selected spots are illustrated in Fig. 4. In principal component analysis, all cell lines were categorized into the expected groups. In hierarchical clustering analysis, T cell leukemia cell lines PEER and CCRF-HSB2 were localized in the group of B cell

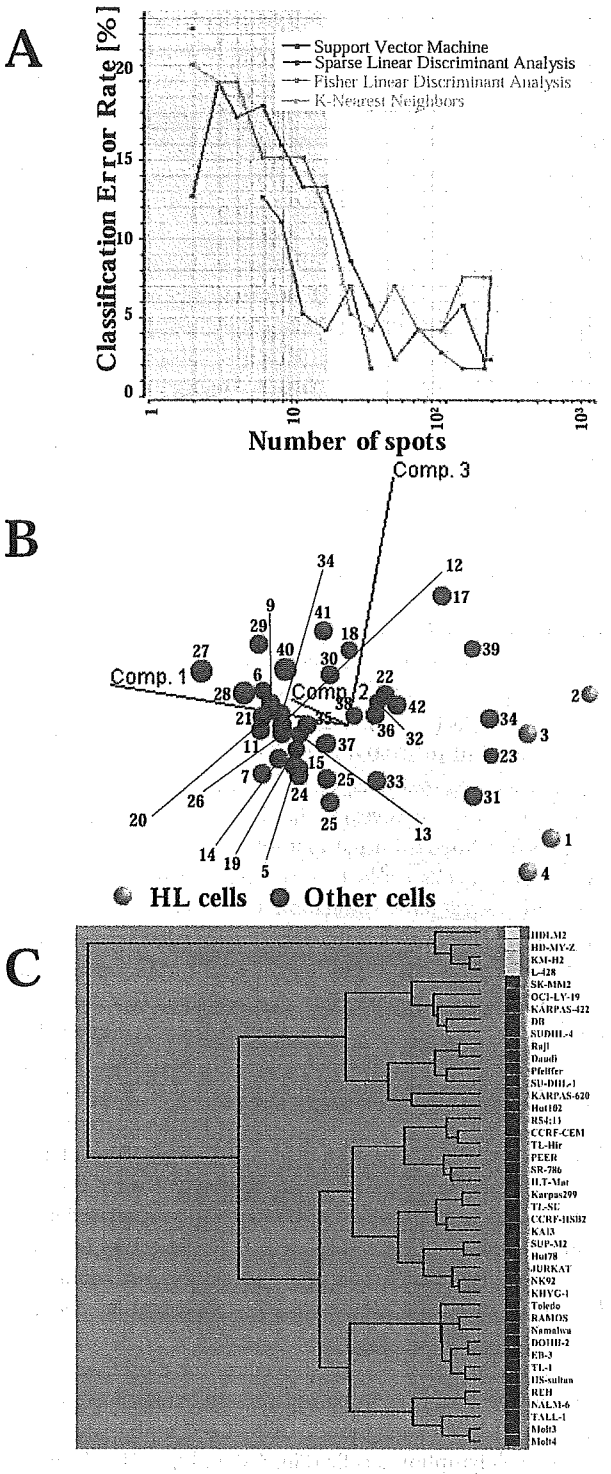

Figure 3. Multivariate analysis of protein expression profiles of HL cells and other cells. (A) Spot ranking methods were used to select the spot set that could discriminate the HL cell group from other cell group with minimal cross-validation error rate. Cross-validation error rate is plotted as a function of the number of spots that best discriminate the two groups. (B) Principal component analysis of the cell lines using the selected 16 spots. The cell lines are numbered according to Table 1. (C) Hierarchical clustering analysis of the cell lines using the selected 16 spots.

malignancy cells and B cell lymphoma cell line KARPAS-422 was clustered with T cell malignancy cells, but the other cell lines were grouped according to their origin (Fig. 4A).

3.4 Localization of the informative spots on 2-D gels and identification of proteins corresponding to the spots

The location of the informative spots on the 2-D images is shown in Fig. 5. The proteins corresponding to the numbered spots were later identified by MS.

Figure 6A illustrates the process of protein identification, using spot 21 as an example. The mass spectrogram of the tryptic digest of spot 21 was subjected to PMF analysis, resulting in the identification of galectin-1 with an identification score of 1009 and protein coverage of 58.5% by amino acid count. The identification was also confirmed by MS/MS data for the tryptic peptide with an m/z value of 1800.6786 (Fig. 6B–D), identifying spot 21 as galectin-1 with a MASCOT score of 86. Figure 6E shows the amino acid sequence of galectin-1, indicating that seven peptides were used for the identification. Similar procedures were performed for the other protein spots. Of 31 protein spots, 23 protein spots were identified. The results of identification are summarized in Table 2.

4 Discussion

HL is characterized by the presence of Hodgkin and Reed-Sternberg (HRS) cells, which usually comprise less than 1% of the cellular infiltrate in the lymphoma tissues [59]. We found that HL cells had a distinct overall proteomic profile: the HL cells were more similar to each other than to any other cell groups (Fig. 1C). It is now generally accepted that HL cells in most cases are derived from germinal center B cells (or, rarely, T cells), although the expression of many B cell markers is lost in HL cells [60]. Our findings suggest that HL cells develop their own protein expression pattern during the course of transformation and that the loss of B cell markers reflects this overall alteration. These concepts are consistent with recent transcriptomic studies on HRS cells, where on the basis of mRNA expression profiles the cells clustered as a distinct entity irrespective of their B or T cell origin [61].

In the comparison among B cell malignancies, T cell malignancies and NK cell lymphoma, both principal component analysis and hierarchical clustering analysis divided these cells according to their original phenotypic groups. In addition, the protein expression pattern of NK lymphoma cells was closer to that of T cell malignancies than that of B cell malignancies. These findings are consistent with recent studies showing that although B cells, T cells, and NK cells are derived from a common lymphoid progenitor [62], NK cells are biologically more related to T cells than to B cells [58]. Although NK cells and T cells can be dis-

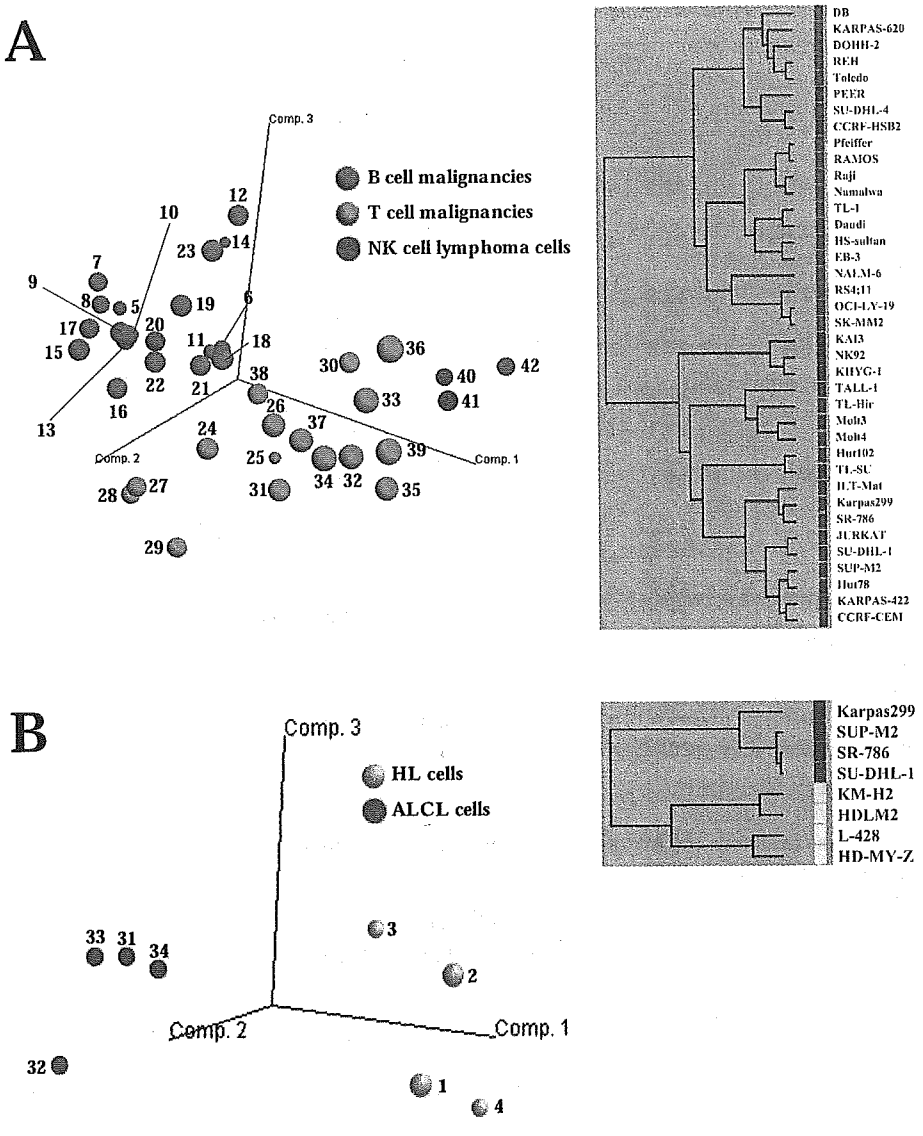

Figure 4. Validation of the selected spots for classification. The cell lines were grouped by principal component analysis and hierarchical clustering analysis using the selected spots. (A) Cell lines from B cell malignancies, T cell malignancies, and NK cell lymphoma. (B) Cell lines from HL and ALCL. The color code in the right panel of hierarchical clustering analysis corresponds to that in principal component analysis. The cell lines are numbered according to Table 1.

tinguished by immunophenotype and molecular genetic studies, there is overlap in NK cell and T cell antigen expression, function, and patterns of disease [58]. We identified seven proteins that were considered to distinguish these three histological subtypes on the basis of their expression pattern. However, linkage of these proteins to histological differentiation was not found by literature validation. Further studies of these proteins may provide unique findings suggestive of novel concepts in lymphocyte ontogeny.

We also examined whether a proteomic approach could contribute to the development of biomarkers for the diagnostically gray area at the interface between HL and other lymphoid neoplasms such as ALCL. ALCL and HL are biologically distinct entities: the majority of ALCL and HL cells are of T cell and B cell origin, respectively [63]. Therefore, the expression of T cell or B cell markers suggests an appropriate diagnosis. However, in some cases, the morphological and immunological features of these neoplasms may overlap considerably, and tumors in this gray area may present clinical problems [64]. Recently, transcriptomic studies, based on cDNA array technology, have revealed that HL cells and ALCL cells can be distinguished from each other by the expression of a small number of genes, including clusterin [65], c-*jun*, calmodulin, growth factor receptor-bound protein (GRB2), and S100A4 [66]. Our study included the protein spots corresponding to Jun-activated domain binding protein 1, calmodulin, and GRB2 (data not shown). However, they were not detected as important proteins for the classification in our study. There are four possible explanations for this discrepancy. First, many lines of evidence suggest that the expression of mRNA does not necessarily reflect that of protein [3–6]. Therefore, transcriptomic and proteomic studies may produce discrepant results by their nature. Second, 2-D DIGE is less sensitive than DNA microarray and does

