

Microarray Reveals Differences in Both Tumors and Vascular Specific Gene Expression in *de Novo* CD5⁺ and CD5⁻ Diffuse Large B-Cell Lymphomas¹

Tohru Kobayashi,² Motoko Yamaguchi, Seungchan Kim, Jun Morikawa, Shoko Ogawa, Satoshi Ueno, Edward Suh, Edward Dougherty, Ilya Shmulevich, Hiroshi Shiku,³ and Wei Zhang^{2,3}

The Second Department of Internal Medicine, Mie University School of Medicine, Tsu, 514-8507, Japan [T. K., M. Y., J. M., S. O., S. U., H. S.]; Cancer Genomics Core Lab, Department of Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030 [I. S., W. Z.]; Department of Electrical Engineering, Texas A&M University, College Station, Texas 77840 [S. K., E. D.]; and National Human Genome Research Institute [S. K.] and Division of Computational Biology, Center for Information Technology [E. S.], NIH, Bethesda, Maryland 20892

ABSTRACT

Malignant lymphoma is a heterogeneous disease with different clinical features. Among diffuse large B-cell lymphomas (DLBCLs), a unique subtype has been identified recently based on cell surface marker CD5 and clinicopathological features. These *de novo* CD5⁺ DLBCLs account for ~10% of all of the DLBCLs and have poorer prognosis. To additionally understand this subtype of DLBCLs at the molecular level and to find genes that are differentially expressed in *de novo* CD5⁺ DLBCLs, CD5⁻ DLBCLs, and mantle cell lymphomas, which also have poor prognosis, we performed gene expression profiling using cDNA microarray technology. Data from a total of 9 samples of CD5⁻ DLBCLs, 11 samples of *de novo* CD5⁺ DLBCLs, and 10 samples of mantle cell lymphomas were acquired. A series of genes were identified that distinguish these three types of lymphomas. Among DLBCL cases, integrin $\beta 1$ and/or CD36 adhesion molecules were overexpressed in most cases of CD5⁺ DLBCL. An immunohistochemical confirmation study revealed that integrin $\beta 1$ was expressed on lymphoma cells, which may account for the high extranodal involvement and poor prognosis of CD5⁺ DLBCLs. In contrast, CD36 was overexpressed on vascular endothelia in CD5⁺ DLBCLs, although there was no difference in vascularity detected by von Wilbrand factor antibody between CD5⁺ and CD5⁻ DLBCLs. Those results suggest that CD5⁺ and CD5⁻ DLBCLs have different gene expression signatures in both tumor cells and their vascular systems.

INTRODUCTION

DLBCL⁴ is the most common subtype of B-cell lymphoma and is heterogeneous in the clinical response to current therapy and in survival time. DLBCLs⁴ are also immunophenotypically heterogeneous; ~10% of DLBCLs express CD5 antigen (1–3). CD5 antigen is originally considered a T-cell marker but is also found in a subset of B-cells. Among B-cell malignancies, CD5 is mainly expressed in CLL and MCL (4). It has been shown that the tumor cells from patients with Richter's syndrome resulting from aggressive transformation of CLL express CD5 antigen (5). However, most patients with CD5⁺ DLBCL have no history of lymphoproliferative disease including CLL. Thus, this type of DLBCL is considered to arise "*de novo*" (5). We and others reported previously that *de novo* CD5⁺ DLBCLs have distinct phenotypic, genotypic, and clinical features (1–3, 5, 6). For

example, the *de novo* CD5⁺ DLBCLs are immunohistochemically negative for both CD23 and cyclin D1, whereas CLLs are positive for CD23, and MCLs are positive for cyclin D1 (1–5). The *de novo* CD5⁺ DLBCLs express immunoglobulin heavy chain variable region VH genes with somatic mutations that are different from CLLs and MCLs, indicating different cell origins for CD5⁺ DLBCLs and CLLs/MCLs (1–3). In addition, the prognosis of patients with *de novo* CD5⁺ DLBCL is poorer than that of patients with CD5⁻ DLBCL (6, 7). These findings demonstrate that *de novo* CD5⁺ DLBCLs constitute a distinct subgroup of DLBCLs of poor prognosis.

A recent study showed that gene expression profiling of DLBCLs identified two distinct types of germinal center B-like DLBCL and activated B-like DLBCL (8). In the present study, cDNA microarray analysis was performed on 30 RNA samples of CD5⁻ and CD5⁺ DLBCLs and MCLs (cyclin D1⁺) to identify gene expression activities that are associated with these three groups. Using a robust gene identification approach based on σ -classifier design (9), we identified genes that separate CD5⁺ from CD5⁻ DLBCLs, as well as genes that separate MCLs from DLBCLs. Among the top genes that separate CD5⁺ and CD5⁻ DLBCLs are *integrin $\beta 1$* and *CD36*, which are expressed in tumor cells and vascular endothelia of CD5⁺ DLBCL cases, respectively.

MATERIALS AND METHODS

Patients/Samples. Clinical samples were obtained from 10 patients with CD5⁻ DLBCL, 11 patients with *de novo* CD5⁺ DLBCL, and 10 patients with MCL (Table 1). The diagnoses were made according to the WHO Classification of Tumors of Hematopoietic and Lymphoid Tissues (10). One patient with CD5⁻ DLBCL was excluded, because this patient was diagnosed with mediastinal large B-cell lymphoma, which was already established as a distinct subtype of DLBCL (10).

DNA microarray studies using specimens of patients with hematopoietic malignancies were approved by Institutional Review Committee in Mie University School of Medicine.

Microarray Production. A total of 2142 known human cDNAs were prepared by PCR from the Research Genetics cDNA clone library, purified using MultiScreen PCR plates (Millipore Corp., Bedford, MA) and verified by sequencing at Cancer Genomics Core Lab (M. D. Anderson Cancer Center) before printing (11). The DNA clones, in 384-well plates, were spotted in replicate onto poly-L-lysine-coated microscope slides using an arrayer (Genomic Solutions, Ann Arbor, MI).

RNA Amplification and Target Labeling. Lymphoma tissues were ground to powder under frozen conditions and lysed in the lysis buffer TRI reagent (MRC, Cincinnati, OH). Cell pellets of six hematopoietic cell lines (K562, HL60, NB4, BV173, KBM7, and Jurkat) were lysed in the TRI reagent. Control RNA was prepared by mixing the same amount of total RNA extracted from these six cell lines. The labeling reaction was performed as described previously (12, 13).

Microarray Hybridization and Image Scanning. To hybridize the slides, purified and labeled cDNA targets were dissolved in 130 μ l of total volume of ExpressHyb solution (Clontech, Palo Alto, CA) containing 8 μ g of polyA_{40–60} (Amersham Pharmacia), 2 μ g of yeast tRNA (Life Technologies, Inc.), and 10 μ g of human Cot I DNA (Life Technologies, Inc.). The mixture was heated to 95°C for 10 min then applied to the slides and covered by a

Received 8/5/02; accepted 10/30/02.

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.

¹ Supported in part by grants (No.13671059, No.12217064, and No.12217062) from Japanese Ministry of Education, Culture, Sports, Science and Technology; Tobacco Settlement Fund to M. D. Anderson Cancer Center as appropriated by the Texas Legislature; and a grant from Kadoorie Foundation to M. D. Anderson Cancer Center.

² To whom requests for reprints should be addressed, at The Second Department of Internal Medicine, Mie University School of Medicine, 2-174, Edobashi, Tsu, 514-8507, Japan. Phone: 81-59-231-5016; Fax: 81-59-231-5200; E-mail: tohruk@clin.medic.mie-u.ac.jp (T. K.) or Cancer Genomics Core Laboratory, Department of Pathology, Box 85, The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030. Phone: (713) 745-1103; Fax: (713) 792-5549; E-mail: wzhang@mdanderson.org (W. Z.).

³ These authors contributed equally to this work.

⁴ The abbreviation used are: DLBCL, diffuse large B-cell lymphoma; CD, cluster of differentiation; CLL, chronic lymphocytic leukemia; MCL, mantle cell lymphoma; EST, expressed sequence tag; vWF, von Wilbrand factor.

Table 1 Summary of patient samples for cDNA microarray study

Case	Diagnosis	Biopsied site	Age	Gender	CS	IPI	CD20	CD5	Cyclin D1
1	CD5-DLBCL	LN	71	M	II	low	+	-	N.T.
2	CD5-DLBCL	LN	75	M	III	high	+	-	N.T.
3	CD5-DLBCL	LN	70	F	III	high-int	+	-	N.T.
4	CD5-DLBCL	LN	44	M	III	low-int	+	-	N.T.
5	CD5-DLBCL	LN	64	F	II	high-int	+	-	N.T.
6	CD5-DLBCL	LN	70	F	II	low-int	+	-	N.T.
7	CD5-DLBCL	LN	77	M	III	low-int	+	-	N.T.
8	CD5-DLBCL	LN	58	M	II	low	+	-	N.T.
9	CD5-DLBCL	LN	42	M	III	low-int	+	-	N.T.
10	CD5+DLBCL	LN	70	F	IV	high	+	+	-
11	CD5+DLBCL	LN	66	M	III	high-int	+	+	-
12	CD5+DLBCL	LN	79	M	III	high	+	+	-
13	CD5+DLBCL	LN	74	F	I	low	+	+	-
14 ^b	CD5+DLBCL	LN	59	F	II	low	+	+	-
15	CD5+DLBCL	LN	59	M	IV	high-int	+	+	-
16	CD5+DLBCL	LN	54	M	IV	low-int	+	+	-
17	CD5+DLBCL	Breast	64	F	IV	high	+	+	-
18 ^b	CD5+DLBCL	Stomach	62	F	I	low	+	+	-
19	CD5+DLBCL	Small Bowel	65	F	II	low	+	+	-
20	CD5+DLBCL	Waldeyer Ring	57	F	II	low	+	+	-
21	MCL	LN	79	F	II	low-int	+	+	+
22	MCL	LN	72	M	IV	high	+	+	+
23	MCL	LN	81	M	III	low-int	+	+	+
24	MCL	LN	80	M	III	low-int	+	+	+
25	MCL	PB	57	M	IV	low-int	+	+	+
26	MCL	LN	62	M	III	high-int	+	+	+
27	MCL	LN	69	M	IV	high	+	+	+
28	MCL	LN	63	M	IV	high-int	+	+	+
29	MCL	LN	63	M	IV	high	+	+	+
30	MCL	LN	62	M	IV	high	+	+	+

^a LN, lymph node; PB, peripheral blood; CS, clinical stage; IPI, international prognostic index; int, intermediate; N.T., not tested.

^b Somatic mutation of immunoglobulin heavy chain gene was analyzed in Ref. 5.

coverslip. Hybridization was carried out at 60°C for 14–16 h in a moisturized box. Slides were washed at 37°C in 1× SSC (3 M sodium chloride and 0.3 M sodium citrate), 0.01% SDS, 0.2× SSC, 0.01% SDS, and twice in 0.1× SSC sequentially for 2 min each. Hybridized arrays were scanned at 10 μm resolution on a GeneTAC LS-IV scanner (Genomic Solution), and the obtained signal intensities were quantified with ArrayVision (Imaging Research Inc., St. Catharines, Ontario, Canada).

Assessment of Replicability of the Data. As already mentioned above, each gene is duplicated on the array. Thus, the variation between the replicate spots can be used to assess the reliability of the measurement of that gene expression. Informally, if the two replicate measurements are close to each other, then the estimate of that gene expression can be obtained by a combination (*e.g.*, average) of the two replicates. If, on the other hand, the two replicates are quite different, then that gene should be flagged as unreliable, and the measurements should not be used in subsequent analysis. We use the following simple method to flag unreliable measurements. For each array (patient), we compute the SD of the absolute values of the differences between the corresponding replicates. Then, any absolute value of the difference that exceeds three times the SD is flagged as being unreliable. Using our procedure, ~3% of the genes, on the average, are deemed unreliable.

Algorithm for Finding Strong Feature (Gene) Sets. We desire classifiers that categorize sample tissues based on the expression values of a set of genes. Because the number of samples for clinical studies is often small, we use a simple classifier and a few genes (at most three in this study) to form the classifier. This helps mitigate the likelihood that a classifier that does well on the sample data, but is not good relative to the populations. As for a specific algorithm to design the classifier, we used a recently developed σ -classifier algorithm. σ -classifier is designed from a probability distribution resulting from spreading the mass of the sample points via a circular distribution to make classification more difficult, whereas maintaining sample geometry (9).

A classifier that has a small error for a large variance is desirable, because its performance is more likely to be robust relative to new data. If a classifier has an extremely small error (≈ 0) for a small σ but a large error for large σ , then we do not consider it sufficiently strong. Intuitively, this method mitigates overfitting the data, because it favors solutions for which little changes in the data lead to little changes in the classifier.

The second feature of the algorithm is that it searches for gene combinations that separate the classes. This feature is especially attractive in biological

settings where heterogeneity is the norm and no single gene can dictate classifications.

All of the computations that search and analyze classifiers in this study were done on a Beowulf-based supercomputer at the Center for Information Technology at NIH. This system is a distributed memory parallel computer consisting of a total of 780 XP/Athlon and Pentium III processors interconnected through a high-speed network.

Histopathology and Immunophenotypic Study. Lymphoma tissues were fixed in 10% formalin and embedded in paraffin, followed by the staining of sections (5-μm thick) with H&E and Giemsa. The immunophenotypic study of tumor cells used a labeled avidin-biotin method for frozen sections as described previously (14). The monoclonal antibodies used in the present study were Leu4(CD3), Leu1(CD5), Leu12 (CD19), and CR2 (CD21; Becton Dickinson, Mountain View, CA); CALLA(CD10), L26 (CD20), MHM6(CD23), anti-IgG, anti-IgA, anti-IgM, anti-IgD, anti- κ , anti- λ , and F8/86 (vWF; DAKO, Carpinteria, CA); 4B7R (integrin β 1) and SMO (CD36; Santa Cruz Biotechnology, Inc., Santa Cruz, CA). In all of the cases, the tumor cells expressed CD19 and/or CD20, but not CD3. Cyclin D1 (IBL, Gunma, Japan) expression was examined by immunohistochemistry using paraffin sections. More than 20% positivity of the tumor cells was assumed to indicate positivity for the purposes of this study.

RESULTS

Identification of Genes Differentially Expressed in CD5⁺ and CD5⁻ DLBCLs and MCLs. Total RNAs were isolated from tumor samples of 11 CD5⁺ DLBCLs, 9 CD5⁻ DLBCLs, and 10 MCLs, and subjected to microarray analysis. After obtaining the quantified gene expression data for the 30 lymphoma samples, we set out to identify strong feature genes that robustly distinguish the three subtypes of lymphomas using the σ -classifier algorithm (9). The advantage of this algorithm is that it not only identifies genes that by themselves separate the groups, it also identifies combination of genes with improved discriminating power. This algorithm also generates a conservative classifier gene list that avoids over-fitting of data, thus producing less error for future classification (9).

Using this algorithm, we identified a number of gene combinations that robustly separate CD5⁻ DLBCL *versus de novo* CD5⁺ DLBCL, as well as MCL *versus* DLBCL. We first examined the genes that were selected repeatedly in the combinations (Tables 2 and 3). To have a global evaluation of how well those strong feature genes separate the three groups of lymphomas when they were merged together, we carried out a hierarchical clustering analysis using 25 genes from the MCL *versus* DLBCL classifications and 42 genes from CD5⁺ *versus* CD5⁻ classifications (Fig. 1, A and B). MCLs were clustered correctly (Fig. 1A), and *de novo* CD5⁺ and CD5⁻ DLBCL cases were completely separated (Fig. 1B) as expected.

A CD5 gene was spotted on the microarray used in the present study, but in most cases, the expression level of CD5 detected by the microarray was weak. Therefore, the detectable expression levels of CD5 did not reflect the immunohistochemical outcome of CD5-positive DLBCL and CD5-negative DLBCL. On the other hand, consistent with the immunohistological staining, cyclin D1 levels

Table 2 Partial list of strong classifier genes separating MCLs and DLBCLs

Accession no.	Gene name	Expression in DLBCLs
AA485036	Human mRNA for KIAA0201 gene, complete cds	+ ^a
NM.005084	<i>Homo sapiens</i> phospholipase A2, group VII (platelet-activating factor acetylhydrolase, plasma; PLA2G7), mRNA	+
NM.004125	<i>Homo sapiens</i> guanine nucleotide binding protein 10 (GNG10) mRNA	+
NM.000917	<i>Homo sapiens</i> procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), α polypeptide 1 (P4HA1)	+
N70349	<i>H. sapiens</i> ART4 gene	+
AC004662	<i>Homo sapiens</i> chromosome 17, clone hRPC.842.A_23, complete sequence	+
AA449289	<i>H. sapiens</i> mRNA for smoothelin	+
AA487700	Cyclin D1 (PRAD1; parathyroid adenomatosis 1)	-
N48137	Glycophorin E	-
U37251	Human KRAB zinc finger protein (ZNF177) mRNA, splicing variant, complete cds	-
AA418564	Human (clone 8B1) Br-cadherin mRNA, complete cds	-
AA465460	Human mRNA for RNA helicase (HRH1), complete cds	-
N91416	ESTs, Weakly similar to coded for by <i>C. elegans</i> cDNA cm04e9 (<i>C. elegans</i>)	-
W72473	PHOSPHATIDYLINOSITOL 3-KINASE CATALYTIC SUBUNIT, α ISOFORM	-
AJ271207	<i>Homo sapiens</i> mRNA for epithelial calcium channel (ECAC1 gene)	-
AA598572	Spleen tyrosine kinase	-
N77183	Catalase	-
R80779	Human protein kinase (MLK-3) mRNA, complete cds	-
H57830	Human gene for histone H1(0)	-
AA425602	ZONA PELLUCIDA SPERM-BINDING PROTEIN 3A PRECURSOR	-
W86776	Histidine ammonia-lyase	-
NM.005123	<i>Homo sapiens</i> nuclear receptor subfamily 1, group H, member 4 (NR1H4), mRNA	-
D50683	<i>Homo sapiens</i> mRNA for TGF- β IIIR α , complete cds	-
AA457027	GTPase-activating protein ras p21 (RASA)	-
N62179	Human methylmalonate semialdehyde dehydrogenase gene, complete cds	-
R73545	Human surface antigen mRNA, complete cds	-
R56552	INTERLEUKIN ENHANCER-BINDING FACTOR	-
AA478647	Intercellular adhesion molecule 3	-
AA443000	Colony stimulating factor 3 receptor (granulocyte)	-
AA479771	Human mRNA for KIAA0076 gene, complete cds	-
AA682293	<i>Homo sapiens</i> phenylalanine hydroxylase (PAH) mutant Q20stop mRNA	-
NM.000873	<i>Homo sapiens</i> intercellular adhesion molecule 2 (ICAM2) mRNA	-
N80713	<i>H. sapiens</i> mRNA for unknown protein expressed in macrophages	-
AA425556	Cadherin 3 (P-cadherin)	-
AA406185	Human mRNA for KIAA0335 gene, complete cds	-
W24076	Active BCR-related gene	-
AA436871	Human syntaxin 3 mRNA, complete cds	-

^a + sign represents the overexpression of gene in DLBCL and - sign the underexpression.

Table 3 Partial list of strong classifier genes separating CD5⁺ and CD5⁻ DLBCLs

Accession No.	Gene Name	Expression in CD5 ⁺ DLBCLs
AA040170	MONOCYTE CHEMOTACTIC PROTEIN 3 PRECURSOR	+ ^a
NML006759	<i>Homo sapiens</i> UDP-glucose pyrophosphorylase 2 (UGP2), mRNA	+
W67174	Integrin, β 1	+
N39161	CD36 antigen (collagen type I receptor, thrombospondin receptor)	+
H82706	Inhibitor of DNA binding 2, dominant negative helix-loop-helix protein	+
AA450205	Human mRNA for translocation protein-1, complete cds	+
AA702422	MJD1	-
AA427561	Human heparan sulfate proteoglycan (HSPG2) mRNA, complete cds	-
AA488336	Keratin 17	-
AA233809	Transforming growth factor, β 2	-
AA069414	Glial fibrillary acidic protein	-
AA022561	Special AT-rich sequence binding protein 1	-
AA490721	Human splicing factor SRp30c mRNA, complete cds	-
T80232	Human autotaxin mRNA, complete cds	-
AA626012	Zinc finger protein 76	-
AA055242	S100 calcium-binding protein A3 (formerly S100E)	-
AA453969	Testis-specific lactate dehydrogenase (LDHC4, LDHX) mRNA	-
AA480035	<i>H. sapiens</i> mRNA for ITBA2 protein	-
NML005835	<i>Homo sapiens</i> sodium phosphate transporter 3 (NPT3) mRNA	-
AF040706	<i>Homo sapiens</i> putative tumor suppressor protein (Fus-2) mRNA	-
AA489234	Human putative serine/threonine protein kinase PRK (prk) mRNA	-
T81764	Cell division cycle 27	-
NM.000016	<i>Homo sapiens</i> acyl-Coenzyme A dehydrogenase	-
AA428859	<i>Human glutathione transferase ζ 1 (GSTZ1) mRNA, complete cds</i>	-
NM.000116	<i>Homo sapiens</i> tafazzin (cardiomyopathy, dilated 3A (X-linked);	-
R33030	PROBABLE PROTEIN DISULFIDE ISOMERASE ER-60 PRECURSOR	-
AA252169	<i>Homo sapiens</i> zinc finger protein mRNA, complete cds	-
AA430751	<i>H. sapiens</i> mRNA for mediator of receptor-induced toxicity	-
H78483	Human huntingtin interacting protein (HIP2) mRNA, complete cds	-
H56595	Human synaptobrevin 2 (SYB2) gene	-
AA480826	Human mRNA for proton-ATPase-like protein, complete cds	-
H59757	V-raf murine sarcoma 3611 viral oncogene homolog 1	-
AA456882	Integrin cytoplasmic domain associated protein (Icap-1a) mRNA	-
AA598815	PROTEASOME ζ CHAIN	-
N40841	PROCHOLECYSTOKININ PRECURSOR	-
AA448866	<i>Homo sapiens</i> dishevelled 2 (DVL2) mRNA, complete cds	-
H79778	Human histone deacetylase 3 (HD3) mRNA, complete cds	-
AJ271207	<i>Homo sapiens</i> mRNA for epithelial calcium channel (ECAC1 gene)	-
AA069596	THYROLIBERIN PRECURSOR	-
H73241	VASOACTIVE INTESTINAL POLYPEPTIDE RECEPTOR 1 PRECURSOR	-
NM.006268	<i>Homo sapiens</i> requiem, apoptosis response zinc finger gene (REQ) mRNA	-

^a + sign represents the overexpression of gene in CD5⁺ DLBCLs and - sign the underexpression.

were high in most cases of MCL, and cyclin D1 is the top classifier for MCL *versus* DLBCL (Table 2).

Confirmation of Integrin β 1 and CD36 Overexpression in CD5⁺ DLBCLs by Immunohistochemistry. Among the genes that separate *de novo* CD5⁺ and CD5⁻ DLBCL, six genes, *integrin β 1*, *CD36*, *mRNA for translocation protein-1*, *monocyte chemotactic protein 3 precursor*, *UDP-glucose pyrophosphorylase 2*, and *inhibitor of DNA binding 2* were overexpressed in *de novo* CD5⁺ DLBCL (Fig. 1B). We focused on two genes, *integrin β 1* and *CD36*, for confirmation studies by immunohistochemistry, because both of these two genes were overexpressed in most cases of CD5⁺ DLBCL. Furthermore, overexpression of *integrin β 1* and *CD36*, adhesion molecules,

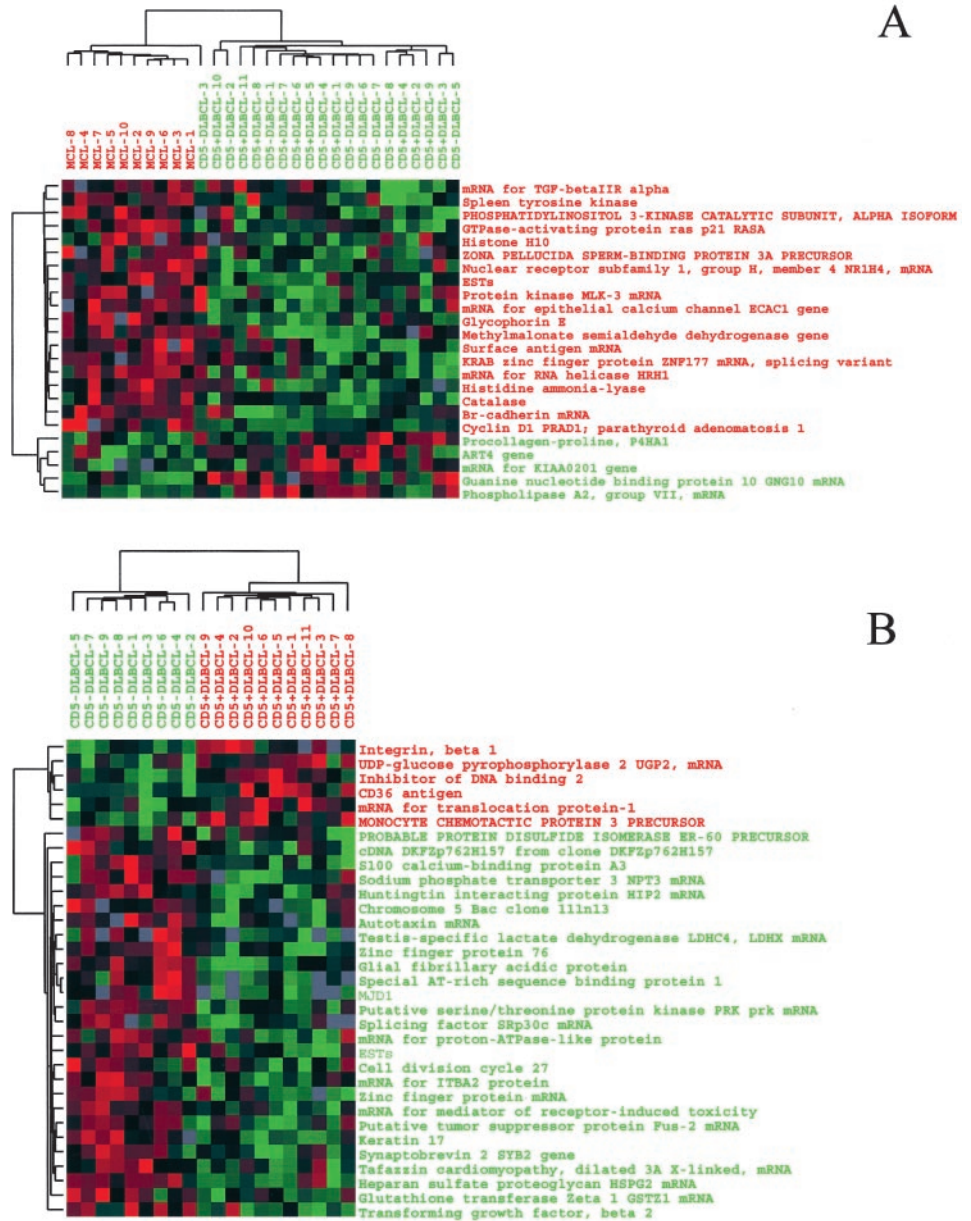


Fig. 1. Clustering of MCL and DLBCL patients. A, MCL versus DLBCL; B, CD5⁺ versus CD5⁻ DLBCL.

may account for the aggressive feature of CD5⁺ DLBCL. Immunohistochemistry assay was selected for confirmation because results of this assay will extend the microarray data to protein level and to cellular localization level. The second benefit is important because microarray measure a population average effect, and tumor tissues are of a mixture of different cell types.

Immunohistochemical staining of integrin $\beta 1$ showed that integrin $\beta 1$ was expressed on lymphoma cells of CD5⁺ DLBCLs (Fig. 2B). However, immunohistochemical staining of CD36 revealed that CD36 was overexpressed on vascular endothelia of CD5⁺ DLBCLs and not expressed on lymphoma cells (Fig. 2F). In all cases of CD5⁻ nodal DLBCLs tested immunohistochemically, CD36 was expressed neither on vascular endothelia nor on lymphoma cells (Table 4; Fig. 2D). To evaluate the vascularity of lymphoma tissues, vWF staining was performed simultaneously using close sections from the same lymphoma tissues used in CD36 staining (Fig. 2, C, E, and G). There were no significant differences in vascularity among CD5⁺ DLBCLs, CD5⁻ DLBCLs, and MCLs (Fig. 2, C, E, and G), suggesting that CD36 overexpression on vascular endothelial cells of CD5⁺ DLBCLs

reflects a different molecular signature in the vessels of CD5⁺ DLBCLs. Among 8 cases of CD5⁻ extranodal DLBCLs, CD36 was expressed on vascular endothelia in 5 cases, but the expression levels of CD36 in CD5⁻ extranodal DLBCLs were weaker than those in CD5⁺ DLBCL cases (Table 2). In 5 patients with MCLs including 3 new patients not analyzed by microarray, integrin $\beta 1$ was weakly expressed in 1 patient, and CD36 was expressed on a few vessels in 3 patients (Fig. 2H).

Gene Combinations as Stronger Classifiers. Cancers are highly heterogeneous, and multiple molecular changes occur during cancer progression. Even the best markers do not cover all cancers of the same type. In the present study, integrin $\beta 1$ and CD36 are good markers in a sense that they are overexpressed in most of CD5⁺ DLBCLs. However, they do not necessarily co-overexpress. As mentioned previously, classifiers constructed from combinations of genes provide more accurate classification than single-gene classifiers. On the basis of the mathematical model used in this study, classification by the two-gene sets of integrin $\beta 1$ and CD36 for CD5⁺ DLBCLs and CD5⁻ DLBCLs (Fig. 2I) provides a more robust classifier. In other

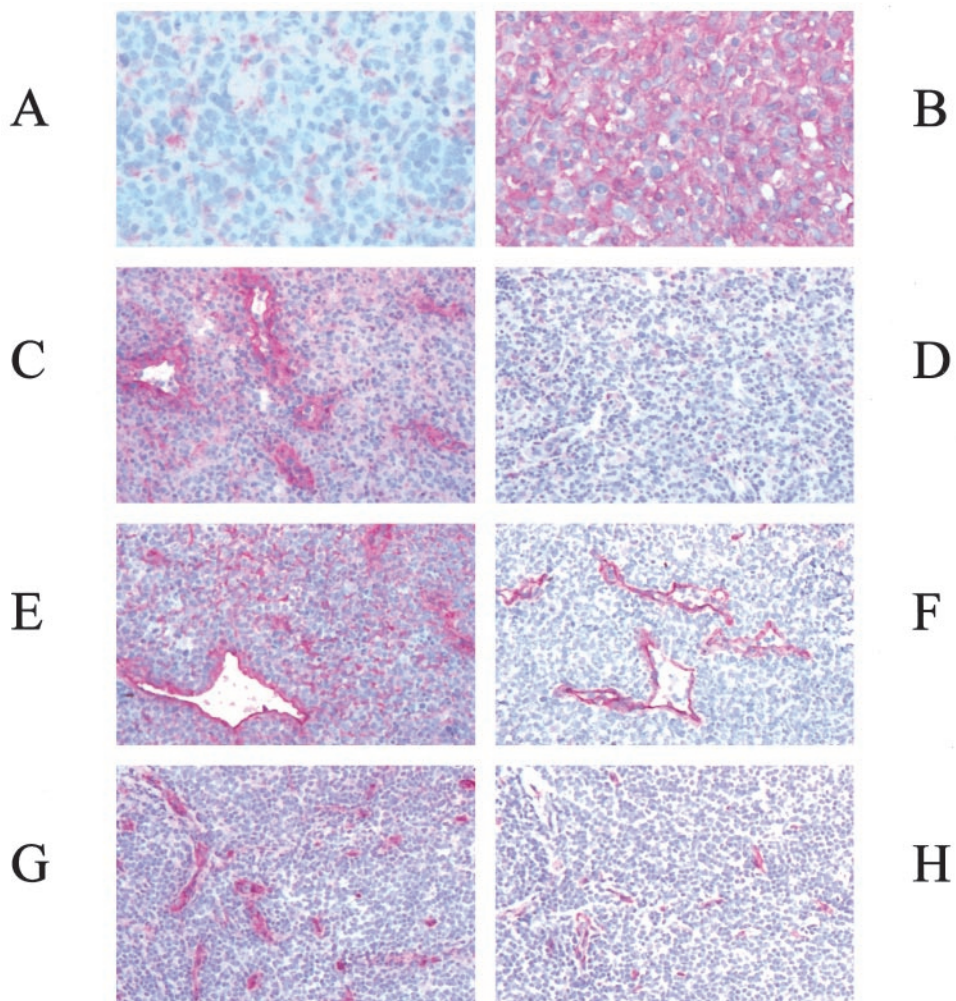
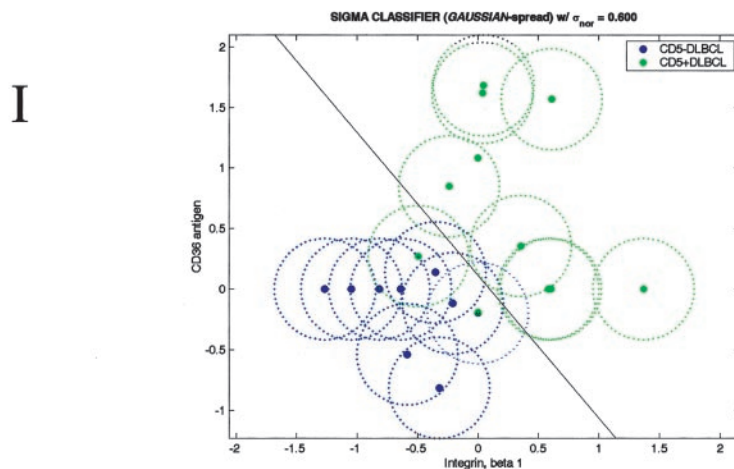


Fig. 2. Immunohistochemical staining of lymphoma tissues for integrin $\beta 1$, vWF, and CD36. *A*, integrin $\beta 1$ staining of CD5⁻ DLBCL tissue; integrin $\beta 1$ is not expressed (case 7); *B*, integrin $\beta 1$ staining of CD5⁺ DLBCL tissue; integrin $\beta 1$ is expressed on lymphoma cells (case 12); *C*, vWF staining of CD5⁻ DLBCL tissue; vWF is expressed on vascular endothelia; *D*, CD36 staining of CD5⁻ DLBCL tissue; CD36 is not expressed. *C* and *D* are close sections from a same tissue (case 7); *E*, vWF staining of CD5⁺ DLBCL tissue; vWF is expressed on vascular endothelia; *F*, CD36 staining of CD5⁺ DLBCL tissue; CD36 is expressed on vascular endothelia. *E* and *F* are close sections from a same tissue (case 12); *G*, vWF staining of MCL tissue; vWF is expressed on vascular endothelia; *H*, CD36 staining of MCL tissue; CD36 is expressed on a few vessels. *G* and *H* are close sections from a same tissue (case 22); *I*, examples of two-gene classifier for separation of CD5⁺ and CD5⁻ DLBCL by integrin $\beta 1$ and CD36; \circ (blue) for CD5⁻ DLBCL, \circ (green) for CD5⁺ DLBCL.



words, the combination of these two molecular events is more reflective of the molecular activities in the CD5⁺ group of DLBCLs.

DISCUSSION

In the present study, we used a 4800-feature cDNA microarray including >2000 known cellular genes to analyze 30 different lymphoma samples. One-third of the cases represent a well-recognized subtype of lymphoma, MCL. The present study selected strong feature genes and classified MCLs from DLBCLs correctly. The fact that

cyclin D1 is selected as a top strong feature gene validates the reliability of the microarray experiments in the present study and supports the previous reports in the literature (14, 15). The other two-thirds are DLBCLs consisting of two recently identified subgroups represented by the surface marker CD5. The *de novo* CD5⁺ DLBCL has been reported to be a unique subgroup of DLBCL clinicopathologically, which we expect to have unique gene expression events reflecting their pathophysiological status. Indeed, among the gene expression events associated with CD5⁺ DLBCLs are the

Table 4 Immunohistochemistry of integrin $\beta 1$, CD36, and vWF

Cases of No. 1–3, 10–14, 18–20, 22, and 27 correspond to cases of Table 1. Integrin $\beta 1$ was expressed on lymphoma cells, and CD36 and vWF were expressed on vascular endothelia. The density of integrin $\beta 1$ was as follows: –, not expressed; +/-, weekly positive; +, positive; ++, strongly positive. CD36⁺ lymphoma cells were not seen in the specimens analyzed in the present study. The frequency of CD36⁺ vessels was calculated as follows: –, CD36⁺ vessels were not or rarely seen; +, CD36 was expressed in a few vessels detected by vWF antibody; ++, CD36 was expressed in more than half vessels detected by vWF antibody.

Case	Dx	Biopsied site	Integrin $\beta 1$	CD36/vWF
1	CD5-DLBCL	LN	+/-	-
2	CD5-DLBCL	LN	+/-	-
3	CD5-DLBCL	LN	-	-
A	CD5-DLBCL	LN	-	-
B	CD5-DLBCL	LN	+/-	-
C	CD5-DLBCL	LN	-	-
D	CD5-DLBCL	LN	-	-
E	CD5-DLBCL	LN	-	-
F	CD5-DLBCL	LN	-	-
G	CD5-DLBCL	Waldeyer	-	-
H	CD5-DLBCL	Waldeyer	-	-
I	CD5-DLBCL	Waldeyer	-	+
J	CD5-DLBCL	Waldeyer	-	+
K	CD5-DLBCL	Waldeyer	-	+
L	CD5-DLBCL	Maxillary sinus	+	+
M	CD5-DLBCL	Stomach	-	-
N	CD5-DLBCL	Stomach	-	+
10	CD5+DLBCL	LN	+/-	+
12	CD5+DLBCL	LN	+	++
13	CD5+DLBCL	LN	+	++
14	CD5+DLBCL	LN	-	++
P	CD5+DLBCL	LN	+	++
Q	CD5+DLBCL	LN	++	+
18	CD5+DLBCL	Waldeyer	+	++
19	CD5+DLBCL	Stomach	+	++
20	CD5+DLBCL	Ileocecal mass	+	+
22	MCL	LN	-	+
27	MCL	LN	-	+
R	MCL	LN	-	+
S	MCL	LN	-	-
T	MCL	Waldeyer	+/-	-

overexpression of some interesting genes. Integrin $\beta 1$, one of adhesion molecules, plays an important role in B-cell lymphoma adhesion and chemotaxis on fibronectin (16); disruption of integrin $\beta 1$ gene in a lymphoma cell line reduced its metastatic potential (17). It has been reported that protein expression of β -integrin adhesion molecules in non-Hodgkin's lymphoma correlated with extranodal involvement (18) and that negative or low expression of β -integrin is associated with favorable prognosis (18).

A particular case is CD36 antigen, which is a thrombospondin receptor. CD36 is transcriptionally regulated by Oct-2, which is a regulator of B-cell differentiation (19). CD36 expression in chronic B-cell lymphoproliferative disorders is related to tumor metastasis (20), and its expression in B-CLLs is reported to be an indicator of tumor cell dissemination (21). Because of the known information from literature, it was first anticipated that CD36 expresses in lymphoma cells. Our immunohistochemistry staining revealed otherwise. CD36 in the CD5⁺ DLBCL tissues is expressed in the vascular endothelial cells rather than the lymphoma cells. This finding has offered several insights. First, endothelial cells in different types of cancers have different molecular events. Thus, the "normal" endothelial cells in cancer are part of the pathophysiological system of the disease. This finding supports the use of "unpurified" tumor tissues for genomic and molecular study rather than using "purified" tumor cells using microdissection. The fact that vascular endothelium is an integral part of cancer is also supported by a recent report that endothelial cells of tumor and normal tissues have different gene expression profiles (22). The second implication is that tumor cells have close communication with endothelial cells residing in the tumor. It has been reported that intravascular or intrasinusoidal infiltration was observed in 19% of *de novo* CD5⁺ DLBCLs (7). It is

conceivable that CD36 expressing endothelial cells are susceptible to contribute to intravascular lymphoma, which is occasionally seen in CD5⁺ DLBCLs.

This hypothesis is supported by our analysis, which showed that combination of integrin $\beta 1$ and CD36 better associates with CD5⁺ DLBCL than each of the two genes alone. It is conceivable that the CD5⁺ lymphoma cells with high integrin $\beta 1$ expression communicate with the endothelial cells, which express CD36. This notion is additionally supported by a recent report that integrin $\beta 1$ physically interacts with CD36 (23). Thus, CD36 may serve as a target for intervention to improve the prognosis of CD5⁺ DLBCLs. It would also be of special interest to investigate the mechanism through which CD36 is activated in the endothelial cells of CD5⁺ DLBCLs.

ACKNOWLEDGMENTS

We thank Dr. Richard Ford for his valuable discussion and critical reading of this manuscript.

REFERENCES

- Kume, M., Suzuki, R., Yatabe, Y., Kagami, Y., Miura, A. B., Morishima, Y., Nakamura, S., and Seto, M. Somatic hypermutations in the V_H segment of immunoglobulin genes of CD5-positive diffuse large B-cell lymphoma. *Jpn. J. Cancer Res.*, 88: 1087–1093, 1997.
- Taniguchi, M., Oka, K., Hiasa, A., Yamaguchi, M., Ohno, T., Kita, K., and Shiku, H. De novo CD5⁺ diffuse large B-cell lymphomas express VH genes with somatic mutation. *Blood*, 91: 1145–1151, 1998.
- Nakamura, N., Hashimoto, Y., Kuze, T., Tasaki, K., Sasaki, Y., Sato, M., and Abe, M. Analysis of the immunoglobulin heavy chain gene variable region of CD5-positive diffuse large B-cell lymphoma. *Lab. Invest.*, 79: 925–933, 1999.
- Burns, B. F., Warnke, R. A., Doggett, R. S., and Rouse, R. V. Expression of a T-cell antigen (Leu1) by B-cell lymphomas. *Am. J. Pathol.*, 113: 165–171, 1983.
- Matolcsy, A., Chadburn, A., and Knowles, D. M. De novo CD5-positive and Richter's syndrome-associated diffuse large B cell lymphomas are genotypically distinct. *Am. J. Pathol.*, 147: 207–216, 1995.
- Yamaguchi, M., Ohno, T., Oka, K., Taniguchi, M., Ito, M., Kita, K., and Shiku, H. De novo CD5-positive diffuse large B-cell lymphoma: clinical characteristics and therapeutic outcome. *Br. J. Haematol.*, 105: 1133–1139, 1999.
- Yamaguchi, M., Seto, M., Okamoto, M., Ichinohasama, R., Nakamura, N., Yoshino, T., Suzumiya, J., Murase, T., Miura, I., Akasaka, T., Tamaru, J., Suzuki, R., Kagami, Y., Hirano, M., Morishima, Y., Ueda, R., Shiku, H., and Nakamura, S. De novo CD5+ diffuse large B-cell lymphoma: a clinicopathologic study of 109 patients. *Blood*, 99: 815–821, 2002.
- Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., Powell, J. I., Yang, L., Marti, G. E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D. B., Tibshirani, R., Sherlock, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., Levy, R., Wilson, W., Grever, M. R., Byrd, J. C., Botstein, D., Brown, P. O., and Staudt, L. M. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature (Lond.)*, 403: 503–511, 2000.
- Kim, S., Dougherty, E. R., Barrera, J., Chen, Y., Bittner, M. L., and Trent, J. M. Strong feature sets from small samples. *J. Comput. Biol.*, 9: 129–148, 2002.
- Jaffe, E. S., Harris, N. L., Stein, H., Vardiman, J. W., (Eds.). *Pathology and Genetics, WHO Classification of Tumours, Tumours of Haematopoietic and Lymphoid Tissues*, pp. 119–187. Lyon: IARC Press, 2001.
- Taylor, E., Cogdell, D., Coombes, K., Hu, L., Ramdas, L., Tabor, A., Hamilton, S., and Zhang, W. Sequence verification as quality control step for production of cDNA microarray. *Biotechniques*, 31: 62–65, 2001.
- Hu, L., Wang, J., Baggerly, K., Wang, H., Fuller, G. N., Hamilton, S. R., Coombes, K. R., and Zhang, W. Obtaining reliable information from minute amounts of RNA using cDNA microarrays. *BMC Genomics*, 3: 16, 2002.
- Shmulevich, I., Hunt, K., El-Naggar, A., Taylor, E., Ramdas, L., Laborde, P., Hess, K. R., Pollock, R., and Zhang, W. Tumor specific gene expression profiles in human leiomyosarcoma: an evaluation of intr-tumor heterogeneity. *Cancer (Phila.)*, 94: 2069–2075, 2002.
- Oka, K., Ohno, T., Kita, K., Yamaguchi, M., Takakura, N., Nishii, K., Miwa, H., and Shirakawa, S. PRAD1 gene over-expression in mantle-cell lymphoma but not in other low-grade B-cell lymphomas, including extranodal lymphoma. *Br. J. Haematol.*, 86: 786–791, 1994.
- Suzuki, R., Takemura, K., Tsutsumi, M., Nakamura, S., Hamajima, N., and Seto, M. Detection of cyclin D1 overexpression by real-time reverse-transcriptase-mediated quantitative polymerase chain reaction for the diagnosis of mantle cell lymphoma. *Am. J. Pathol.*, 159: 425–429, 1995.
- Yin, Z., Giacomello, E., Gabriele, E., Zardi, L., Aota, S., Yamada, K. M., Skerlavaji, B., Doliana, R., Colombatti, A., and Perris, R. Cooperative activity of $\alpha 4\beta 1$ and $\alpha 4\beta 7$ integrins in mediating human B-cell lymphoma adhesion and chemotaxis on fibronectin.

- tin through recognition of multiple synergizing binding sites within the central cell-binding domain. *Blood*, *93*: 1221–1230, 1999.
17. Stroeken, P. J., van Rijnthoven, E. A., van der Valk, M. A., and Roos, E. Targeted disruption of the $\beta 1$ integrin gene in a lymphoma cell line greatly reduces metastatic capacity. *Cancer Res.*, *58*: 1569–1577, 1998.
 18. Terol, M. J., Lopez-Guillermo, A., Bosch, F., Villamor, N., Cid, M. C., Campo, E., and Montserrat, E. Expression of β -integrin adhesion molecule in non-Hodgkin's lymphoma: correlation with clinical and evolutive features. *J. Clin. Oncol.*, *17*: 1869–1875, 1999.
 19. Corcoran, L. M., Karvelas, M., Nossal, G. J. V., Ye, Z-S., Jacks, T., and Baltimore, D. Oct-2, although not required for early B-cell development, is critical for later B-cell maturation and for postnatal survival. *Genes Dev.*, *7*: 570–582, 1993.
 20. Ruttella, S., Rumi, C., Di Mario, A., and Leone, G. Expression of thrombospondin receptor (CD36) in chronic B-cell lymphoproliferative disorders: a role of tumor metastasis. *Eur. J. Histochem.*, *41*(Suppl. 2): 53–54, 1997.
 21. Rutella, S., Rumi, C., Puggioni, P., Barberi, T., Di Mario, A., Larocca, L. M., and Leone, G. Expression of thrombospondin receptor (CD36) in B-cell chronic lymphocytic leukemia as an indicator of tumor cell dissemination. *Haematologica*, *84*: 419–424, 1999.
 22. St. Croix, B., Rago, C., Velculescu, V., Traverso, G., Romans, K. E., Montgomery, E., Lal, A., Riggins, G. J., Lengauer, C., Vogelstein, B., and Kinzler, K. W. Genes expressed in human tumor endothelium. *Science* (Wash. DC), *289*: 1121–1122, 2000.
 23. Thorne, R. F., Marshall, J. F., Shafren, D. R., Gibson, P. G., Hart, I. R., and Burns, G. F. The integrins $\alpha 3\beta 1$ and $\alpha 6\beta 1$ physically and functionally associate with CD36 in human melanoma cells. *J. Biol. Chem.*, *275*: 35264–35275, 2000.