Jpn. J. Cancer Res. **93**, 894–901, August 2002

Microarray Analysis of Gene-expression Profiles in Diffuse Large B-cell Lymphoma: Identification of Genes Related to Disease Progression

Mieko Nishiu,[1] Rempei Yanagawa,[2] Shin-ichi Nakatsuka,[1] Masayuki Yao,[1] Tatsuhiko Tsunoda,[3] Yusuke Nakamura[2] and Katsuyuki Aozasa[1, 4]

[1]Department of Pathology, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita, Osaka 565-0871, [2]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639 and [3]SNP Research Center, Riken Institute of Physical and Chemical Research, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639

To identify genes that are associated with progression of malignant lymphoma, the expression profiles of 18 432 genes were analyzed in diffuse large B-cell lymphomas at early (stages I and II, 6 cases) and advanced stages (stages III and IV, 9 cases) by means of cDNA microarrays. By comparing expression profiles between localized and advanced lymphomas, a number of genes that were differentially expressed were identified: 48 genes with increased expression and 30 genes with reduced expression in advanced-stage diffuse large B-cell lymphomas. Increased expression of *MPHOSPH1*, *RUVBL1*, *CHN2*, *PSA* and *CDC10* genes, and reduced expression of *COL1A2*, *COL4A1*, *FBLN5*, *CLECSF6*, *MIC2*, *CAV1* and *S100A10* genes in the advanced lymphoma group were confirmed by semi-quantitative reverse transcription-PCR. *RUVBL1* and *PSA* expression was further confirmed by real-time quantitative PCR, whose results paralleled the microarray data. The highly expressed genes encode proteins that promote cell proliferation and the genes with reduced expression encode adhesion proteins and target protein for cytotoxic T-lymphocytes. These findings suggested that analysis with cDNA microarrays is a useful approach for identifying genes related to tumor progression and their products could be potential tumor markers or disease-specific targets for anti-tumor therapy.

Key words: Microarray — Expression profile — Diffuse large B-cell lymphoma — Stage — Progression

Diffuse large B-cell lymphoma (DLBCL) is a major constituent of malignant lymphomas in the REAL classification[1)] and comprises 41.5% and 24.6% of non-Hodgkin's lymphomas in Osaka, Japan (unpublished data) and Western countries,[2)] respectively. DLBCL is categorized as one of the aggressive lymphomas. At present, the prognosis of lymphoma cases is well predicted by the International Prognostic Index (IPI),[3)] a widely accepted clinical risk factor model, which incorporates the patient's age, performance status, serum lactate dehydrogenase level, clinical stage, and number of extranodal lesions. The stage of disease is the main prognostic factor for malignant lymphomas, and is divided into early stage (stages I and II) and advanced stage (stages III and IV) based on the Ann Arbor system.[4)]

Information for molecular mechanisms involved in lymphoma development has been accumulating. As for diffuse large B-cell lymphomas, *BCL-6* translocation,[5)] BCL-2 overexpression and *c-MYC* translocation[6)] have been reported. Other biological features reported are expression of cytokine (IL-6),[7)] adhesion molecule (CD44),[8)] and proliferation index (Ki-67).[9)] However, the biological behavior of lymphoma cases could not be accurately predicted through estimation of increased or diminished expression of any one gene or a small number of these genes. Hence, analysis of expression profiles of a large number of genes is an essential step toward understanding in detail the mechanisms of lymphomagenesis and lymphoma progression. cDNA microarrays have been used for investigating gene expression profiles in human cancers.[10)] Recently, Alizadeh *et al.* reported successful molecular classification of DLBCLs by gene expression profiling, which made it possible to identify clinically significant subtypes of tumors.[11)]

In this paper, the expression profiles of 18 432 genes were compared between diffuse large B-cell lymphomas at early and advanced stages, using a cDNA microarray.

MATERIALS AND METHODS

Patients and tissue samples Tumor tissues were obtained with informed consent from 15 patients who underwent open lymph node biopsy or surgical resection at Osaka University Hospital. Histologic specimens were fixed in 10% formalin, and routinely processed for paraffin embedding. Histologic sections, cut at 4 mm, were stained with hematoxylin and eosin and immunoperoxidase procedures

[4] To whom requests for reprints should be addressed.
E-mail: aozasa@molpath.med.osaka-u.ac.jp

(avidin-biotin complex method). Monoclonal antibodies used for immunophenotyping were CD20 (Kyowa Medex, Tokyo), CD79a (Dakopatts, Glostrup, Denmark), CD3 (Dakopatts) and CD45RO (Dakopatts). Through review of the combined histologic and immunohistologic findings, all of the cases were diagnosed as DLBCL. Specimens for RNA extraction were snap-frozen in liquid nitrogen just after resection or stored in "RNA later" reagent (Ambion, Inc., Austin, TX) until use. Patients were staged according to the Ann Arbor system.[4] IPI[3] was also assessed in each patient.

RNA extraction and T7-based RNA amplification Total RNA was extracted from each specimen using TRIzol reagent (Life Technologies, Inc., Rockville, MD) according to the manufacturer's instructions. As a common control, we used Human Lymph Node poly(A)$^+$ RNA, from a pool of 12 normal, whole lymph nodes (Clontech, Palo Alto, CA). Total extracted RNAs and control poly(A)$^+$ RNA were treated with 10 units of DNase I (Roche, Basel, Switzerland) at 37°C for 1 h in the presence of 1 unit of RNase inhibitor (Toyobo, Osaka), to remove any contaminating genomic DNA. Aliquots of the DNase I-treated RNAs were subjected to T7-based RNA amplification, because analysis on a microarray requires several micrograms of mRNA for synthesizing probes. T7-based RNA amplification was carried out as described previously,[12] using Ampliscribe T7 Transcription Kit (Epicentre Technologies, Madison, WI). The remainder of un-amplified total RNA was stored for semi-quantitative reverse transcription-PCR (RT-PCR) and real-time quantitative PCR analysis.

Preparation of microarrays Microarrays containing 18 432 human cDNAs were designed on the basis of DNA sequence data in the Unigene database of the National Center for Biotechnology Information (NCBI). The cDNAs were amplified by RT-PCR using poly(A)$^+$ RNA isolated from various human organs as templates; the length of the amplicons ranged from 200 to 1100 bp without repetitive or poly(A)$^+$ sequences. The PCR products were spotted in duplicate on type-7 glass slides (Amersham Pharmacia Biotech, Little Chalfont, UK) using an Array Spotter Generation III (Amersham Pharmacia Biotech). Each slide contained 52 housekeeping genes, used to normalize the signal intensities of the different fluorescent dyes.

Labeling, hybridization and scanning Fluorescent probe, labeled with Cy3 dye (Amersham Pharmacia Biotech), was prepared from each experimental RNA sample. A common reference cDNA probe, labeled with Cy5 dye (Amersham Pharmacia Biotech), was prepared from RNA from a pool of normal lymph nodes. The use of a common control cDNA probe allows comparison of the relative expression of each gene to be made across all samples possible. Each Cy3-labeled experimental cDNA probe was mixed with the Cy5-labeled reference probe and was used for hybridization to the microarray. After hybridization, the slides were washed, and then scanned using the Array Scanner Generation III (Amersham Pharmacia Biotech).

Quantitation of data The intensity of each hybridization signal was evaluated photometrically by the Array Vision computer program (Amersham Pharmacia Biotech) and normalized to the averaged signals of housekeeping genes. The Cy3:Cy5 ratio for each sample was calculated by averaging spots. A cutoff value for each expression level was automatically calculated according to the background fluctuation. The fluctuation can be estimated as the variance of the log ratio of Cy3:Cy5 minus the variance of the log ratio of Cy3:Cy5 of highly expressed genes (the upper 30%, where the background fluctuation is so small that it could be ignored). We used genes with an expression level of more than the cutoff value. Genes with low signal intensities were excluded from further investigation.

Selection of genes with different expression The fluorescence ratio (Cy3:Cy5) was quantified for each gene to reflect the relative abundance of the gene in each experimental sample compared with the reference sample. To detect genes that were expressed differently between the two groups, Mann-Whitney tests was used with the criterion of significance set at $P<0.05$. Among the genes selected, we used the genes that showed a 2-fold or greater difference in the median intensity of expression between the two groups.

Semi-quantitative RT-PCR Total RNAs were treated with DNase I (Roche) and reverse-transcribed using oligo(dT)$_{12-18}$ primer with Superscript II reverse transcriptase (Life Technologies, Inc.). Each PCR with gene-specific primers was carried out for 20–35 cycles at 94°C for denaturing, 57°C for annealing, and 72°C for extension. Amplified products were analyzed by electrophoresis on 2% agarose gels. The sequences of the primers used were as follows: CHN2 forward, 5′-GCCGTCCATGAAGTGCTGA-3′; CHN2 reverse, 5′-TGGTACCGCATATCATGCAG-3′; MPHOSPH1 forward, 5′-CATTTGTGGGTACATAGTACA-3′; MPHOSPH1 reverse, 5′-CTGATAGCACAATAGAGTGA-3′; RUVBL1 forward, 5′-GCATGTCGAAGAGATCAGTGA-3′; RUVBL1 reverse, 5′-CTGAACTGACAGCGCTGCA-3′; PSA forward, 5′-GTGATTGTCCGTGATGACCT-3′; PSA reverse, 5′-CAGACACGTAGAATCCTTGAG-3′; CDC10 forward, 5′-GCTTACTGTGCACCTAGAGC-3′; CDC10 reverse, 5′-TTTCCTTGTTGGTTCCTACCA-3′; Hs.49221 forward, 5′-TCCAGTGTGGTTTTTAACAGAT-3′; Hs.49221 reverse, 5′-CTCCTGGTAATAGCACTCAAC-3′; Hs.54699 forward, 5′-GTAGAATGAAAAAGGCACCTG-3′; Hs.54699 reverse, 5′-AAAACACTCACATGTCTCTGAT-3′; CAV1 forward, 5′-ACGTCCACACCGTCTGTGA-3′; CAV1 reverse, 5′-ACTTATGAAAACTGAGAAAGTGA-3′; S100A10 forward, 5′-GGCAAAGTGGGCTTCCAGA-3′; S100A10 reverse, 5′-GCTAAGTGTCCTGATCTGCT-3′; COL4A1 forward, 5′-

ACAGCCAGACCATTCAGATC-3′; COL4A1 reverse, 5′-GCTGTAAGCGTTTGCGTAGT-3′; COL1A2 forward, 5′-GTGGCTTTTGAATATCTTCCAC-3′; COL1A2 reverse, 5′-TCTTCTCAAATTCTTCTAGCAC-3′; FBLN5 forward, 5′-CCTCTCATTGGCACCAAGG-3′; FBLN5 reverse, 5′-CACCAACAATCTTCTATCAGG-3′; CLECSF6 forward, 5′-ATGTTAATTGTCTTGGTCCTCA-3′; CLECSF6 reverse, 5′-GTGAATCAGTCAATGTATAGAC-3′; MIC2 forward, 5′-TGCAGAACAAGGGGAGGTG-3′; MIC2 reverse, 5′-CAGAAGGCCTCCATCTCTG-3′; G3PDH forward, 5′-GACAACAGCCTCAAGATCATCA-3′; G3PDH reverse, 5′-GGTCCACCACTGACACTGTG-3′.

Real-time quantitative PCR The third fraction of the original total RNA was used as a template in the real-time PCR (TaqMan 5′ nuclease fluorogenic quantitative PCR assay) analysis. PCR was performed with a PE Applied Biosystems 7700 Prism (PE Applied Biosystems, Foster City, CA). Matching primers and fluorescence probes (see below) were designed for *PSA* and *RUVBL1* genes according to the Primer Express program provided by PE Applied Biosystems. β-Actin primers and probes were purchased from PE Applied Biosystems. The PCR reaction was performed in a total volume of 25 μl containing 1× TaqMan buffer, 0.2 m*M* dATP, dCTP, dGTP and 0.4 m*M* dUTP, 0.625 unit of *AmpliTaq* Gold (PE Applied Biosystems), 0.25 unit of Amperase uracil-N-glycosylase (PE Applied Biosystems), 5 m*M* $MgCl_2$, and 2 μl of each appropriately diluted sample. Both of the forward and reverse primers were used at 900 n*M* for *PSA* and *RUVBL1* and 300 n*M* for β-actin. The final probe concentration was adjusted at 150 n*M*. The sequences of the primers used were as follows: PSA-TaqMan forward, 5′-CGTGGTGATTGTCCGTGATG-3′; PSA-TaqMan reverse, 5′-GTTTCCAGCCTGCACCTTGT-3′; PSA-TaqMan probe, 5′-TTTGCCCTCCGAGAGTGCCCCT-3′; RUVBL1-TaqMan forward, 5′-TGGCACCAAGACCACACTGA-3′; RUVBL1-TaqMan reverse, 5′-TCCCGTTGATTTTAGCAAGCA-3′; RUVBL1-TaqMan probe, 5′-CAGCTGCTGACCCCGGCCA-3′. Each sample was normalized with respect to its β-actin content. The pool of normal human lymph node tissues was used as a calibrator, and the relative amounts of each gene were also normalized with respect to the calibrator. For example, the levels of PSA expression relative to the *β-actin* gene and the calibrator are given as follows:

$$\frac{\text{PSA}^{\text{sample}}/\beta\text{-actin}^{\text{sample}}}{\text{PSA}^{\text{calibrator}}/\beta\text{-actin}^{\text{calibrator}}}$$

The reproducibility of the quantitative measurements was evaluated by conducting triplicate PCR assessments.

RESULTS

Clinical features of the 15 patients are summarized in Table I. Six patients were diagnosed as having early stage

Table I. Clinical Characteristics of 15 Patients with Diffuse Large B-cell Lymphomas

Case no.	Age/Sex	Primary site	Stage[a]	IPI[b]	B-symptoms[c]	LDH[d]	Chemotherapy[e]	Effect[f]
38	51/F	Thyroid	I	0 (Low)	No	Normal	CHOP	CR
27	24/M	Lymph node	I	0 (Low)	No	Normal	THP-CHOP	CR
13	54/F	Mammary gland	II	0 (Low)	No	Normal	VACOP-B	CR
15	82/M	Lymph node	II	1 (Low)	No	Normal	CHOP	CR
26	53/M	Thyroid	II	1 (Low)	No	High	VACOP-B	CR
22	57/M	Lymph node	II	1 (Low)	No	High	VACOP-B	CR
6	59/M	Lymph node	III	2 (Low intermediate)	No	Normal	CHOP	CR
32	58/F	Lymph node	III	3 (High intermediate)	No	High	VACOP-B	NR
30	71/M	Lymph node	III	4 (High)	No	High	CHOP	PR
29	62/M	Lymph node	III	4 (High)	No	High	VACOP-B	PD
20	85/F	Lymph node	III	4 (High)	Yes	High	CHOP	CR
23	77/F	Lymph node	III	4 (High)	No	High	THP-CHOP	PD
11	55/F	Mammary gland	IV	4 (High)	No	High	CHOP	PD
12	67/F	Lymph node	IV	5 (High)	Yes	High	CHOP	PD
18	60/F	Lymph node	IV	5 (High)	Yes	High	CHOP	NR

a) Clinical stage by Ann Arbor system.
b) Risk assessment by International Prognostic Index.
c) Presence of at least one of these: unexplained weight loss, unexplained fever, night sweats.
d) Serum lactate dehydrogenase level.
e) Variations of initial chemotherapeutic treatment. ADM, adriamycin; EDX, cyclophosphamide; VCR, vincristine; PSL, predonisolone; VP16, etoposide; BLE, bleomycin. CHOP: EDX, ADM, VCR, PSL. THP-CHOP: therarubicin, EDX, VCR, PSL. VACOP-B: VP16, EDX, ADM, VCR, PSL, BLE.
f) CR, complete response; PR, partial response; NR, no response; PD, progressive disease (World Health Organization).

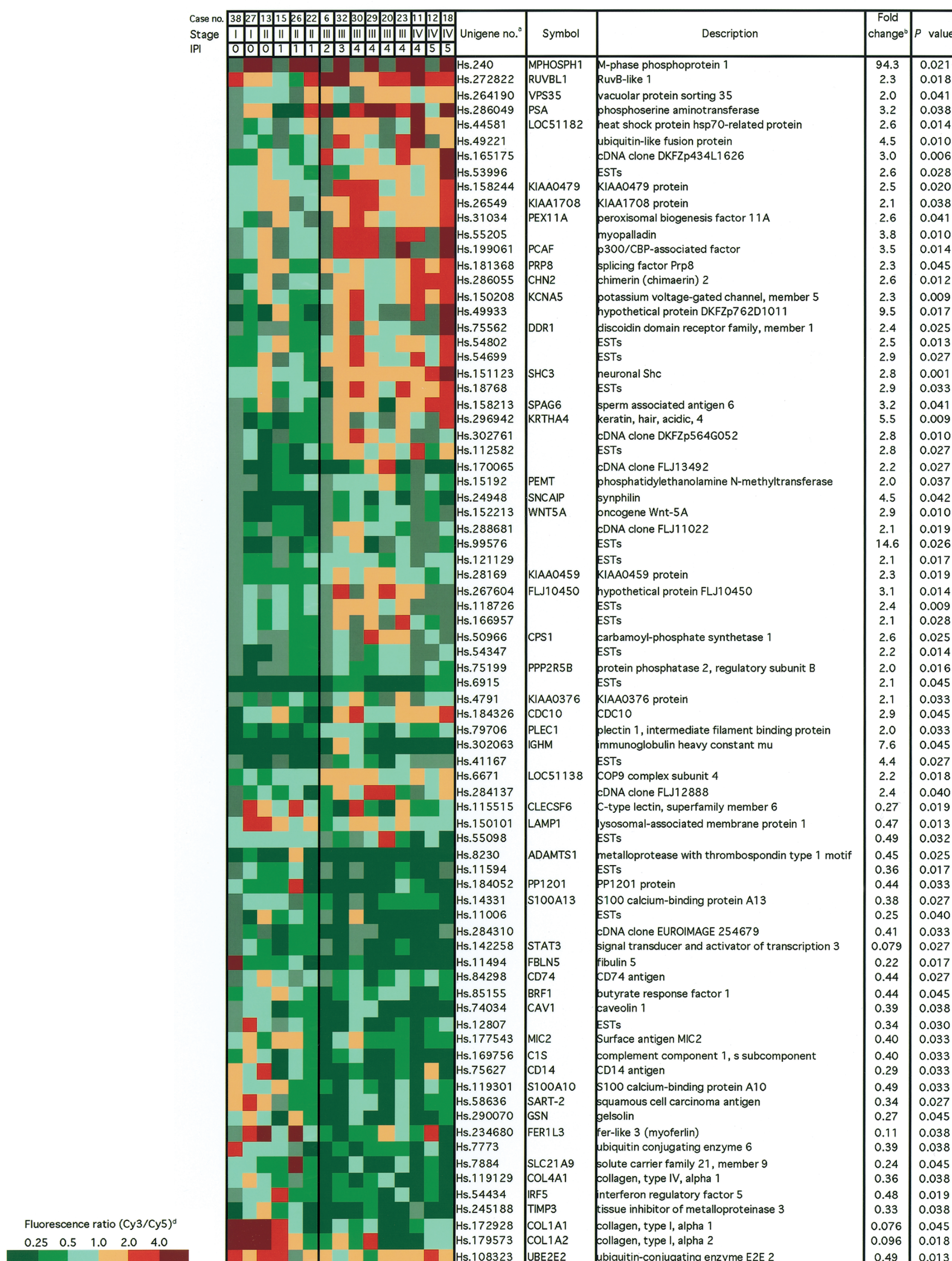

Fig. 1. Genes that were expressed differentially between localized and advanced diffuse large B-cell lymphomas. (a) Unigene accession number in the database of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). (b) Fold change represents the ratio of the median intensity of expression between the two groups. (c) P value of the Mann-Whitney test. (d) Color patches represent the relative expression levels (Cy3/Cy5) from dark green (lowest) to dark red (highest). Missing or excluded values are coded as gray.

(I or II) and nine as having advanced stage (III or IV) disease. IPI was determined for each patient based on the clinical information. Except for one case (no. 6), patients in the localized stage group were enrolled in the low-risk group by IPI, and those in the advanced stage group in the high-risk group.

Differentially expressed genes are shown in Fig. 1, in which 48 genes with increased expression and 30 genes with decreased expression in the advanced lymphomas are listed.

To confirm and validate the results obtained with the cDNA microarray, we analyzed expression of some of the differentially expressed genes by semi-quantitative RT-PCR. Semi-quantitative RT-PCR analyses (Fig. 2) verified the increased expression of *MPHOSPH1*, *PUVBL1*, *PSA*, *CHN2*, *CDC10*, two sets of ESTs (expressed sequence tags) numbered Hs.49221 and Hs.54699, and decreased expression of *CLECSF6*, *FBLN5*, *CAV1*, *MIC2*, *S100A10*, *COL4A1* and *COL1A2* in the advanced lymphoma group.

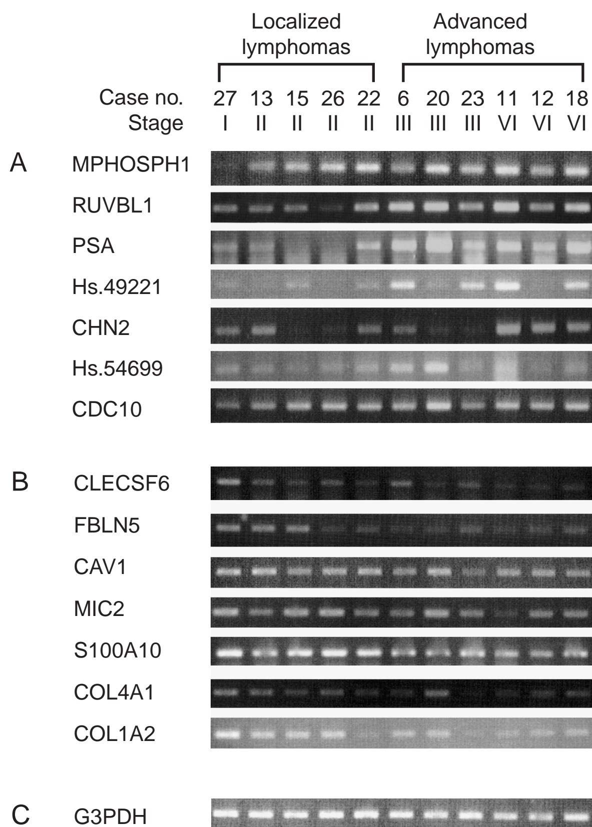

Fig. 2. Semi-quantitative RT-PCR analysis of total RNAs from diffuse large B-cell lymphomas. Genes with increased (A) and reduced (B) expression in advanced lymphomas identified by the microarray. (C) The integrity of each RNA template was controlled through amplification of G3PDH.

Expression patterns found at semi-quantitative RT-PCR analysis were similar to those on the microarrays. For example, *CHN2* was highly expressed in stage IV samples (cases 11, 12 and 18) by microarray hybridization and also by semi-quantitative RT-PCR analysis. Overexpression of the *PSA* gene was observed in cases 6, 11, 18 and 20 by both the microarray and semi-quantitative RT-PCR analyses. Among the down-regulated genes on the microarray, *FBLN5* was shown to be decreased in advanced lymphoma samples by semi-quantitative RT-PCR analysis.

To further investigate the reliability of our array data, we measured the expression levels of *PSA* and *RUVBL1* genes using real-time quantitative PCR. Fig. 3, A and C demonstrate that the relative expression levels of each gene determined by real-time quantitative PCR were similar to those observed with cDNA microarrays. Linear regression lines (correlation coefficient=0.89, *P* value<0.01 for *PSA*, and correlation coefficient=1.32, *P* value<0.01, for *RUVBL1*) were fitted (Fig. 3, B and D).

DISCUSSION

We identified as many as 78 genes whose expression was definitely increased or decreased in diffuse large B-cell lymphoma at the advanced stage. Some of the highly

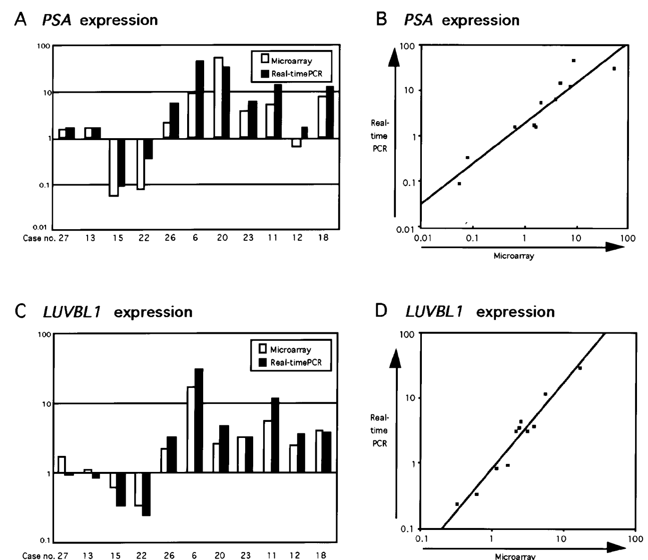

Fig. 3. A comparative ratio analysis by real-time quantitative PCR and cDNA microarray hybridization of the *PSA* gene (A, B) and the *RUVBL1* gene (C, D). Expression level is shown as the abundance of the transcript relative to its abundance in the common reference RNA (a pool of normal human lymph node tissues). Linear regression lines (correlation coefficient=0.89, *P* value<0.01 for *PSA*, and correlation coefficient=1.32, *P* value<0.01 for *RUVBL1*) were fitted (B, D).

expressed genes in the advanced lymphoma group were supposed to contribute to tumorigenesis or tumor progression. As a cell-cycle modulator, MPHOSPH1 is reported to be phosphorylated in the M phase, and has thus been suggested to play a role in the mitotic process.[13] *CDC10* encodes a protein homologous to the *Saccharomyces cerevisiae* cell division protein.[14] PRP8, with a homology to yeast protein, is expected to be involved in pre-mRNA splicing and cell cycle progression.[15] Genes associated with cell proliferation signaling or other signaling pathways were also included. Nuclear protein RUVBL1 is reported to be an essential cofactor of oncogenic transformation by c-Myc.[16] It has also been isolated as a binding partner of β-catenin and the TATA-box binding protein, suggesting a role of RUVBL1 in the nuclear function of β-catenin.[17] *CHN2* encodes a GTPase-activating protein for p21Rac.[18] Because the small-GTP binding protein Rac plays important roles in controlling actin polymerization and cellular morphogenesis,[19] alterations of its signaling pathway could change the attachment and migratory capacity of tumor cells. Because *CHN2* expression was exclusively high in the stage IV lymphomas, it must be elucidated whether so-far unknown functions of this gene correlate with lymphoma dissemination or not. Phosphoserine aminotransferase (PSA) is required for the major phosphorylation pathway of serine biosynthesis, whose activities are known to be increased in neoplastic tissues.[20] These results, together with our previous findings, suggest that PSA expression could be a useful marker for proliferation status of malignant lymphomas.

There were many genes whose expression was reduced in the advanced lymphomas. Among them, down-regulation of the three collagen genes, *COL1A1*, *COL1A2* and *COL4A1*, was the most notable. Collagens participate in a variety of cellular processes such as differentiation, tumorigenesis and apoptosis. Overexpression of COL1A2 (type I collagen a2) suppresses tumorigenesis in RAS-transformed NIH3T3 cells.[21] Loss of type IV collagen has been reported in mesenchymal tumors such as fibrosarcomas.[22] Our data suggest that the decrease in collagens may contribute to lymphomagenesis and/or dissemination of lymphoma cells. *TIMP3* expression was reduced in the advanced lymphomas. TIMP3 belongs to the metalloproteinase(s) inhibitor family and is known to act on matrix metalloproteinase(s) (MMP)-1, 2, 3, 7, 9, 13, 14 and 15, which affect remodeling of the extracellular matrix.[23] Interestingly, Andreu *et al.* identified TIMP3 and COL1A2 as inhibitors of oncogenic transformation induced by epidermal growth factor (EGF); mRNA expression of these genes is specifically repressed in EGF-transformed cells.[24]

As for cell adhesion molecules, *FBLN5* was recently reported to encode a secreted RGD (Arg-Gly-Asp) motif-containing protein that promotes cell adhesion and may play a role in vascular development and remodeling.[25] *FBLN5* was clearly shown to be down-regulated in advanced lymphomas (Fig. 2), so further investigation to elucidate the association of this gene with lymphoma development and/or progression is needed. *CLECSF6* encodes a C-type lectin surface receptor and is predominantly expressed in hematopoietic tissues.[26] Among B-lymphocytes, CLECSF6 is mostly expressed in resting or circulating B-cells and appears to be down-regulated in activated lymphoid cells.[26] Reduced expression of this gene in the advanced lymphomas might result in activation of lymphoid cells. Caveolin-1 (CAV1) is a caveolar membrane protein that might act as scaffolding for organizing certain signaling molecules such as G-proteins and tyrosine kinases.[27] Previous studies suggested that caveolin-1 acts as a tumor suppressor gene and a negative regulator of the Ras-p42/44 MAP kinase cascade.[27] *SART-2* encodes a tumor antigen recognized by cytotoxic T lymphocytes,[28] and thus down-regulation of *SART-2* might favor escape of tumor cells from immune surveillance of host cytotoxic T lymphocytes.

Furthermore we identified approximately 30 ESTs or cDNA sequences of unknown function that were increased or decreased in advanced lymphoma, suggesting that a large number of genes of unproven function are also involved in lymphomagenesis and progression of DLBCL.

Intragroup heterogeneity of expression level was observed in most genes. This may be due to tumor characteristics other than clinical stage. For *SHC3*, *CDC10*, *CHN2*, *KCNA5*, there seems to be a tendency that the relative expression level was high in cases resistant to adriamycin-based chemotherapeutic treatment (NR and PD cases in Table I). Because the kinds of chemotherapeutic agents used varied from case to case, a clear conclusion would require analysis of larger numbers of cases. As for the *PSA* gene, which we discussed as a new marker, there was no correlation between expression levels and chemotherapeutic response.

In summary, the present study on DLBCL supports the view that analysis with cDNA microarrays is a useful approach for investigating tumor-related genes whose altered expression is linked to tumor progression, and also for identifying genes whose products could be potential tumor markers or disease-specific targets for anti-tumor therapy.

ACKNOWLEDGMENTS

This work was supported in part by grants (11470353, 11670212, 11680546, 12576004, 12670159, 12770087) from the Ministry of Education, Culture, Sports, Science and Technology Japan.

(Received January 29, 2002/Revised May 9, 2002/Accepted May 17, 2002)

REFERENCES

1) Harris, N. L., Jaffe, E. S., Stein, H., Banks, P. M., Chan, J. K. C., Cleary, M. L., Delsol, G., Wolf-Peeters, C. D., Falini, B., Gatter, K. C., Grogan, T. M., Isaacson, P. G., Knowles, D. M., Mason, D. Y., Muller-Hermelink, H. K., Pileri, S. A., Piris, M. A., Ralfkiaer, E. and Warnke, R. A. A revised European-American classification of lymphoid neoplasms: a proposal from the international lymphoma study group. *Blood*, **84**, 1361–1392 (1994).
2) Pittaluga, S., Bijnens, L., Teodorovic, I., Hagenbeek, A., Meerwaldt, J. H., Somers, R., Thomas, J., Noordijk, E. M. and De Wolf-Peeters, C. Clinical analysis of 670 cases in two trials of the European Organization for the Research and Treatment of Cancer Lymphoma Cooperative Group subtyped according to the Revised European-American Classification of Lymphoid Neoplasms: a comparison with the Working Formulation. *Blood*, **87**, 4358–4367 (1996).
3) The International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. *N. Engl. J. Med.*, **329**, 987–994 (1993).
4) Carbone, P. P., Kaplan, H. S., Musshoff, K., Smithers, D. W. and Tubiana, M. Report of the committee on Hodgkin's disease staging classification. *Cancer Res.*, **31**, 1860–1861 (1971).
5) Offit, K., Lo Coco, F., Louie, D. C., Parsa, N. Z., Leung, D., Portlock, C., Ye, B. H., Lista, F., Filippa, D. A., Rosenbaum, A., Ladanyi, M., Jhanwar, S., Dalla-Favera, R. and Chaganti, R. S. K. Rearrangement of the bcl-6 gene as a prognostic marker in diffuse large-cell lymphoma. *N. Engl. J. Med.*, **331**, 74–80 (1994).
6) Kramer, M. H., Hermans, J., Wijburg, E., Philippo, K., Geelen, E., van Krieken, J. H., de Jong, D., Maartense, E., Schuuring, E. and Kluin, P. M. Clinical relevance of BCL2, BCL6, and MYC rearrangements in diffuse large B-cell lymphoma. *Blood*, **92**, 3152–3162 (1998).
7) Seymour, J. F., Talpaz, M., Cabanillas, F., Wetzler, M. and Kurzrock, R. Serum interleukin-6 levels correlate with prognosis in diffuse large-cell lymphoma. *J. Clin. Oncol.*, **13**, 575–582 (1995).
8) Stauder, R., Eisterer, W., Thaler, J. and Gunthert, U. CD44 variant isoforms in non-Hodgkin's lymphoma: a new independent prognostic factor. *Blood*, **85**, 2885–2899 (1995).
9) Miller, T. P., Grogan, T. M., Dahlberg, S., Spier, C. M., Braziel, R. M., Banks, P. M., Foucar, K., Kjeldsberg, C. R., Levy, N., Nathwani, B. N., Schnitzer, B., Tubbs, R. R., Gaynor, E. R. and Fisher, R. I. Prognostic significance of the Ki-67-associated proliferative antigen in aggressive non-Hodgkin's lymphomas: a prospective Southwest Oncology Group trial. *Blood*, **83**, 1460–1466 (1994).
10) DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A. and Trent, J. M. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.*, **14**, 457–460 (1996).
11) 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*, **403**, 503–511 (2000).
12) Luo, L., Salunga, R. C., Guo, H., Bittner, A., Joy, K. C., Galindo, J. E., Xiao, H., Rogers, K. E., Wan, J. S., Jackson, M. R. and Erlander, M. G. Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat. Med.*, **5**, 117–122 (1995).
13) Westendorf, J. M., Rao, P. N. and Gerace, L. Cloning of cDNA for M-phase phosphoproteins recognized by the MPM2 monoclonal antibody and determination of the phosphorylated epitope. *Proc. Natl. Acad. Sci. USA*, **91**, 714–718 (1994).
14) Nakatsuru, S., Sudo, K. and Nakamura, Y. Molecular cloning of a human cDNA homologous to CDC10 in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.*, **202**, 82–87 (1994).
15) Shea, J. E., Toyn, J. H. and Johnston, L. H. The budding yeast U5 snRNP Prp8 is a highly conserved protein which links RNA splicing with cell cycle progression. *Nucleic Acids Res.*, **22**, 5555–5564 (1994).
16) Wood, M. A., McMahon, S. B. and Cole, M. D. An ATPase/helicase complex is an essential cofactor for oncogenic transformation by c-Myc. *Mol. Cell*, **5**, 321–330 (2000).
17) Bauer, A., Huber, O. and Kemler, R. Pontin 52, an interaction partner of beta-catenin, binds to the TATA box binding protein. *Proc. Natl. Acad. Sci. USA*, **95**, 14787–14792 (1998).
18) Leung, T., How, B. E., Manser, E. and Lim, L. Cerebellar bata 2-chimaerin, a GTPase-activating protein for p21 ras-related rac is specifically expressed in granule cells and has a unique N-terminal SH2 domain. *J. Biol. Chem.*, **269**, 12888–12892 (1994).
19) Joneson, T., McDonough, M., Bar-Sagi, D. and Van Aelst, L. RAC regulation of actin polymerization and proliferation by a pathway distinct from Jun kinase. *Science*, **274**, 1374–1376 (1996).
20) Snell, K. Enzymes of serine metabolism in normal, developing and neoplastic rat tissues. *Adv. Enzyme Regul.*, **22**, 325–400 (1984).
21) Travers, H., French, N. S. and Norton, J. D. Suppression of tumorigenicity in Ras-transformed fibroblasts by alpha 2(I) collagen. *Cell Growth Differ.*, **7**, 1353–1360 (1996).
22) Trueb, B. and Odermatt, B. F. Loss of type VI collagen in experimental and most spontaneous human fibrosarcomas.

Int. J. Cancer, **86**, 331–336 (2000).
23) Apte, S. S., Olsen, B. R. and Murphy, G. The gene structure of tissue inhibitor of metalloproteinases (TIMP)-3 and its inhibitory activities define the distinct TIMP gene family. *J. Biol. Chem.*, **270**, 14313–14318 (1995).
24) Andreu, T., Beckers, T., Thoenes, E., Hilgard, P. and von Melchner, H. Gene trapping identifies inhibitors of oncogenic transformation. The tissue inhibitor of metalloproteinases-3 (TIMP3) and collagen type I alpha2 (COL1A2) are epidermal growth factor-regulated growth repressors. *J. Biol. Chem.*, **273**, 13848–13854 (1998).
25) Nakamura, T., Ruiz-Lozano, P., Lindner, V., Yabe, D., Taniwaki, M., Furukawa, Y., Kobuke, K., Tashiro, K., Lu, Z., Andon, N. L., Schaub, R., Matsumori, A., Sasayama, S., Chien, K. R. and Honjo, T. DANCE, a novel secreted RGD protein expressed in developing, atherosclerotic, and balloon-injured arteries. *J. Biol. Chem.*, **274**, 22476–22483 (1999).
26) Bates, E. E., Fournier, N., Garcia, E., Valladeau, J., Durand, I., Pin, J. J., Zurawski, S. M., Patel, S., Abrams, J. S., Lebecque, S., Garrone, P. and Saeland, S. APCs express DCIR, a novel C-type lectin surface receptor containing an immunoreceptor tyrosine-based inhibitory motif. *J. Immunol.*, **163**, 1973–1983 (1999).
27) Engelman, J. A., Zhang, X., Galbiati, F., Volonte, D., Sotgia, F., Pestell, R. G., Minetti, C., Scherer, P. E., Okamoto, T. and Lisanti, M. P. Molecular genetics of the caveolin gene family: implications for human cancers, diabetes, Alzheimer disease, and muscular dystrophy. *Am. J. Hum. Genet.*, **63**, 1578–1587 (1998).
28) Nakao, M., Shichijo, S., Imaizumi, T., Inoue, Y., Matsunaga, K., Yamada, A., Kikuchi, M., Tsuda, N., Ohta, K., Takamori, S., Yamana, H., Fujita, H. and Itoh, K. Identification of a gene coding for a new squamous cell carcinoma antigen recognized by the CTL. *J. Immunol.*, **164**, 2565–2574 (2000).