Comparison of gene expression profiles of lymphoma cell lines from transformed follicular lymphoma, Burkitt's lymphoma and *de novo* diffuse large B-cell lymphoma

Yoshitomo Maesako, Takashi Uchiyama and Hitoshi Ohno[1]

Department of Hematology and Oncology, Graduate School of Medicine, Kyoto University, 54 Shogoin-Kawaramachi, Sakyo-ku, Kyoto 606-8507

(Received February 27, 2003/Revised June 26, 2003/Accepted July 7, 2003)

To determine the specific gene expression in B-cell lymphoma subtypes, we compared expression profiles of cell lines from transformed follicular lymphoma (tFL), Epstein-Barr virus-negative (EBV(−)) Burkitt's lymphoma (BL) and EBV(+)BL. Complementary DNAs were synthesized from these cell lines and hybridized with the Atlas Human 1.2 Array membrane. Hierarchical clustering analysis based upon the levels of 43 genes highlighted characteristic expression patterns of the 3 lymphoma subtypes. Genes expressed at higher levels in tFL than EBV(−)BL and EBV(+)BL included calcium/calmodulin-dependent protein kinase I (*CAMK1*) and mitogen-activated protein kinase 10 (*MAPK10*). EBV(−)BL was characterized by high-level expression of amyloid β precursor protein (*APP*), heat shock 27 kD protein 1 (*HSPB1*) and mothers against decapentaplegic homolog 1 (*MADH1*). Gardner-Rasheed feline sarcoma viral oncogene homolog (*FGR*) was the most significant gene to delineate EBV(+)BL. A subtype prediction algorithm using 34 genes correctly classified 22 (92%) of 24 lymphomas into FL/tFL, EBV(−)BL or EBV(+)BL. By comparison with normal reference B-cell materials, the expression patterns of the selected genes were characteristic of lymphomas. We extended the clustering analysis to cell lines from *de novo* diffuse large B-cell lymphoma (DLBCL). The DLBCL cell lines were either separated from the former 3 lymphoma subtypes or segregated with EBV(+)BL, possibly reflecting variable genetic abnormalities. The associations of *CAMK1* with tFL, *APP* and *MADH1* with EBV(−)BL, *FGR* with EBV(+)BL, and *BCL2* with tFL and DLBCL were confirmed by real-time quantitative reverse transcriptase-mediated polymerase chain reaction assays. This study has provided new molecular markers, expressions of which are closely associated with B-cell lymphoma subtypes. (Cancer Sci 2003; 94: 774–781)

Chromosomal translocations are associated with subtypes of B-cell lymphoma,[1] even though the association is not necessarily specific.[2] The most common translocations are t(14;18)(q32;q21), t(8;14)(q24;q32) and t(3q27), involving *BCL2*, *MYC* and *BCL6* genes, respectively.[1] At present, however, very little is known about the molecular mechanism that connects the chromosomal translocation and clinicopathological features of each lymphoma subtype.

Complementary DNA (cDNA) array is a recently developed technology that simultaneously quantifies the expression levels of a large number of genes. Application of this large-scale gene expression analysis to lymphoma studies has led to the identification of new diagnostic marker genes for particular lymphoma subtypes,[3] previously unknown subclasses within diffuse large B-cell lymphoma (DLBCL)[4, 5] and molecular markers predicting clinical outcome of DLBCL.[6] These molecular classifications of lymphomas using gene expression levels are performed by clustering analysis.[7] The goal of cluster analysis is to separate a collection of complex samples into groups in such a way that the most similar samples are grouped together. Unsupervised clustering methods, which do not use any information from the samples, are used to discover new classes within a generic disease. Supervised clustering methods, in contrast, start from a known classification and find the best methods for distinguishing the classes.[6, 7]

Follicular lymphoma (FL) is characterized by t(14;18)(q32;q21), which leads to molecular fusion of the *BCL2* gene with immunoglobulin heavy chain gene (*IgH*). Although FL initially shows an indolent clinical behavior, the disease is ultimately transformed to aggressive lymphoma (transformed FL, tFL). In contrast, Burkitt's lymphoma (BL) is a high-grade disease and the majority of cases carry t(8;14)(q24;q32). To determine the gene expression characterizing these two contrasting B-cell lymphomas, we compared the gene expression profiles of tFL and BL cell lines in a supervised fashion. Since BL is divided into endemic type BL with Epstein-Barr virus genome (EBV(+)BL) and sporadic type EBV(−)BL, the clustering analysis was performed for tFL-EBV(−)BL-EBV(+)BL subtype distinction. We next extended the clustering algorithm to cell lines from intermediate-grade *de novo* DLBCL carrying various genetic abnormalities. The findings of this study may reveal differential gene expressions that characterize B-cell lymphoma subtypes.

Materials and Methods

Cell lines and clinical materials. tFL cell lines, FL-18 through -518,[8] FL-618, FL-718 and FL-818 lines were established from patients who had had histologically confirmed FL. All the tFL cell lines carried t(14;18)(q24;q32), which is the hallmark of FL. The EBV(−)BL and EBV(+)BL cell lines were described previously.[9] DLBCL cell lines were from patients with *de novo* large cell lymphoma. The B-cell origin of these lymphoma cell lines was determined by immunophenotypic analysis using monoclonal antibodies and flow cytometry. LCL-OHN is an EBV-transformed lymphoblastoid cell line. The cell lines were cultured in RPMI 1640 medium supplemented with 10% fetal calf serum under the standard culture conditions. Clinical materials of lymphoma were stored as frozen viable cell suspensions with 10% dimethyl sulfoxide in liquid nitrogen.

Preparation of peripheral blood B-cells. Peripheral blood B-cells were isolated from buffy coat of multiple healthy donors by using magnetic micro-beads conjugated to monoclonal mouse anti-human CD19 antibody (CD19 MultiSort Kit; Miltenyi Biotec, Bergisch Gladback, Germany). $CD27^-$ naive B-cells were then negatively selected by means of CD27 MicroBeads (Miltenyi Biotec) and the >95% purity of selected cells was con-

[1]To whom correspondence should be addressed.
E-mail: hohno@kuhp.kyoto-u.ac.jp

firmed by flow cytometric analysis. The $CD19^+CD27^-$ B-cells were cultured with sub-lethally irradiated L-cells expressing CD40 ligand (CD40L)[10] for 72 h in the presence of 10 μg/ml anti-IgM antibody (ICN Pharmaceuticals, Inc., Aurora, OH) and 20 ng/ml interleukin-4 (IL-4) (PeproTech EC, Ltd., London, UK). The cells were pelleted and cryopreserved until use for experiments.

Polymerase chain reaction (PCR) and reverse transcriptase-mediated (RT-) PCR. Sequences of the oligonucleotide primers used to detect oncogene/immunoglobulin (Ig) fusion genes by long-distance PCR were described previously.[9] Long-distance inverse PCR was developed to detect non-Ig/*BCL6* translocation.[11] Each PCR reaction mixture (50 μl) contained 100 ng of genomic DNA, reaction buffer, dNTP mixture, 20 pmol of each primer and 2.5 U *LA Taq* DNA polymerase (TaKaRa Shuzo, Kyoto). PCR cycling variables for long DNA targets were previously described in detail.[9] The presence of EBV genome was determined by PCR amplification of the region of EBV-encoded small RNA (EBER). The primer sequences were 5′-GTGGTCCGCATGTTTTGATC-3′ and 5′-GCAACGGCTGTCCTGTTTGA-3′.[2] For RT-PCR, randomly primed cDNA was synthesized from 1 μg of total cellular RNA (First-Strand cDNA Synthesis Kit; Pharmacia, Piscataway, NJ), and PCR assays were performed in 50 μl reaction volumes containing 2 μl of cDNA, reaction buffer, 0.25 mmol each dNTP, 20 pmol of each primer, and 0.5 U *Taq* polymerase (TaKaRa Shuzo). Aliquots of the PCR products were analyzed by ethidium bromide-stained gel electrophoresis.

cDNA array hybridization. Total cellular RNA was prepared using an "RNeasy" Total RNA Kit (Qiagen, Hilden, Germany). The RNA was incubated with RNase-free DNase to eliminate potential genomic DNA contamination. For an array hybridization, 40 μg of total RNA was reverse-transcribed in the presence of [α-^{32}P]dATP using the Atlas Pure Total RNA Labeling System (Clontech, Palo Alto, CA). The Atlas Human 1.2 Array (Clontech) contained a total of 1176 cDNA fragments of known human genes and several housekeeping genes. Hybridization of the ^{32}P-labeled cDNA with the array membrane was performed according to the manufacturer's instructions. The membrane was exposed to an imaging analyzer (BAS2000; Fuji Photo Film, Tokyo) and the hybridization signals were quantified by the ArrayGauge version 1.3 software (Fuji Photo Film).

Analysis of cDNA array data. Data from each array were normalized by the median value to eliminate the variability due to the sample labeling or to the exposure duration. To select differentially expressed genes, we used a 2-sample *t* test for distinction of 2 subtypes or a one-way analysis of variance (ANOVA) for distinction of 3 subtypes. We next performed hierarchical clustering analysis of the array data using the GeneSpring version 4.1.5 software (Silicon Genetics, Redwood City, CA) to generate tree structures based upon the degree of similarity, as well as matrices comparing the levels of expression of individual genes in each sample. The Class Predictor program was used to predict the lymphoma subtype of a set of samples. The program first examines all genes in the training set individually, and then uses the most predictive genes to classify the test set. The predictive strength represents the power of the gene to discriminate each subtype from all the others.

Real-time quantitative RT-PCR. Real-time RT-PCR analysis based on the *Taq*Man methodology was performed using an ABI Prism 7700 Sequence Detection System (Applied Biosystems, Foster City, CA). cDNA was synthesized from 5 μg of total cellular RNA isolated from clinical materials using random primers. The DNA standard template containing a cDNA sequence of interest was generated by cloning into pGEM-T EASY vector (Promega, Madison, WI). An aliquot of cDNA or standard template DNA (1 μl) was added to the PCR reaction mixture containing 1× *Taq*Man Universal PCR Master Mix (Applied Biosystems), each primer at 400 n*M* and 100 n*M* probe in a total reaction volume of 50 μl. After initial incubation for 2 min at 50°C and 10 min at 95°C to activate the *Taq* polymerase, 50 cycles of 15 s at 95°C and 1 min at 60°C were carried out. The C_T (threshold cycle) parameter was defined as the fractional cycle number at which the fluorescence generated by cleavage of the probe passed a preset threshold. The standard curve, in which C_T decreased in linear proportion to the log of the template copy number, was established using serially diluted recombinant pGEM-T plasmid DNA. The C_T values of test materials were plotted on the standard curve and the corresponding copy number was calculated using the Sequence Detecter version 1.6 software (Applied Biosystems). The cDNA copy number of a gene of interest in the test materials was divided by that of the endogenous reference, 18S ribosomal RNA (rRNA) (*Taq*Man Ribosomal RNA Control Reagents; Applied

Table 1. Characteristics of tFL, EBV(+)BL and EBV(−)BL cell lines subjected to Atlas Array analysis

Subtype	Designation	Translocation	Genetic abnormality	EBV	Origin
tFL	FL-18	t(14;18)(q32;q21)	*BCL2*-MBR/*IgH*-J_H	−	Japanese
	FL-18EB	t(14;18)(q32;q21)	*BCL2*-MBR/*IgH*-J_H	+	Japanese
	FL-218	t(14;18)(q32;q21)	*BCL2*-MBR/*IgH*-J_H	−	Japanese
	FL-318	t(14;18)(q32;q21)	*BCL2*-MBR/*IgH*-J_H	−	Japanese
	FL-418	t(14;18)(q32;q21)	*BCL2*-MBR/*IgH*-J_H	−	Japanese
	FL-518	t(14;18)(q32;q21)	*BCL2*-MBR/*IgH*-J_H	−	Japanese
	FL-618	t(14;18)(q32;q21)	*BCL2*-MBR/*IgH*-J_H	−	Japanese
	FL-718	t(14;18)(q32;q21)	*BCL2*-MBR/*IgH*-J_H	−	Japanese
	FL-818	t(14;18)(q32;q21), another 14q+	*BCL2*-MBR/*IgH*-J_H	−	Japanese
EBV(−)BL	Kobayashi	t(8;14)(q24;q32)	*MYC*/*IgH*-Cγ	−	Japanese
	KS-Bu3	t(8;14)(q24;q32)	*MYC*/*IgH*-Cγ	−	Japanese
	Tree92	t(8;14)(q24;q32)	*MYC*/*IgH*-Cμ	−	Japanese
	KPB-L2	t(8;14)(q24;q32)	*MYC*/*IgH*-Cγ	−	Japanese
	Kimura	t(8;14)(q24;q32)	*MYC*/*IgH*-Cγ	−	Japanese
	Ramos	t(8;14)(q24;q32)	*MYC*/*IgH*-Cμ	−	American
EBV(+)BL	Daudi	t(8;14)(q24;q32)	No *MYC* rearrangement	+	African
	Raji	t(8;14)(q24;q32)	*MYC*/*IgH*-Cγ	+	African
	Middle91	t(8;14)(q24;q32)	No *MYC* rearrangement	+	Japanese
	P3HR-1	t(8;14)(q24;q32)	No *MYC* rearrangement	+	African
	HBL-5	t(8;14)(q24;q32)	*MYC*/*IgH*-Cγ	+	Japanese

Biosystems), the level of which in Raji cells was defined as 1. All assays were performed in triplicate.

Statistical analysis. Data were analyzed with the StatView version 5.0 statistical software package (Abacus Concept, Berkeley, CA). To evaluate the significance of differences between lymphoma subtypes, the data of real-time PCR were log-transformed and a 2-sample *t* test was applied. *P* values <0.05 were considered to be statistically significant.

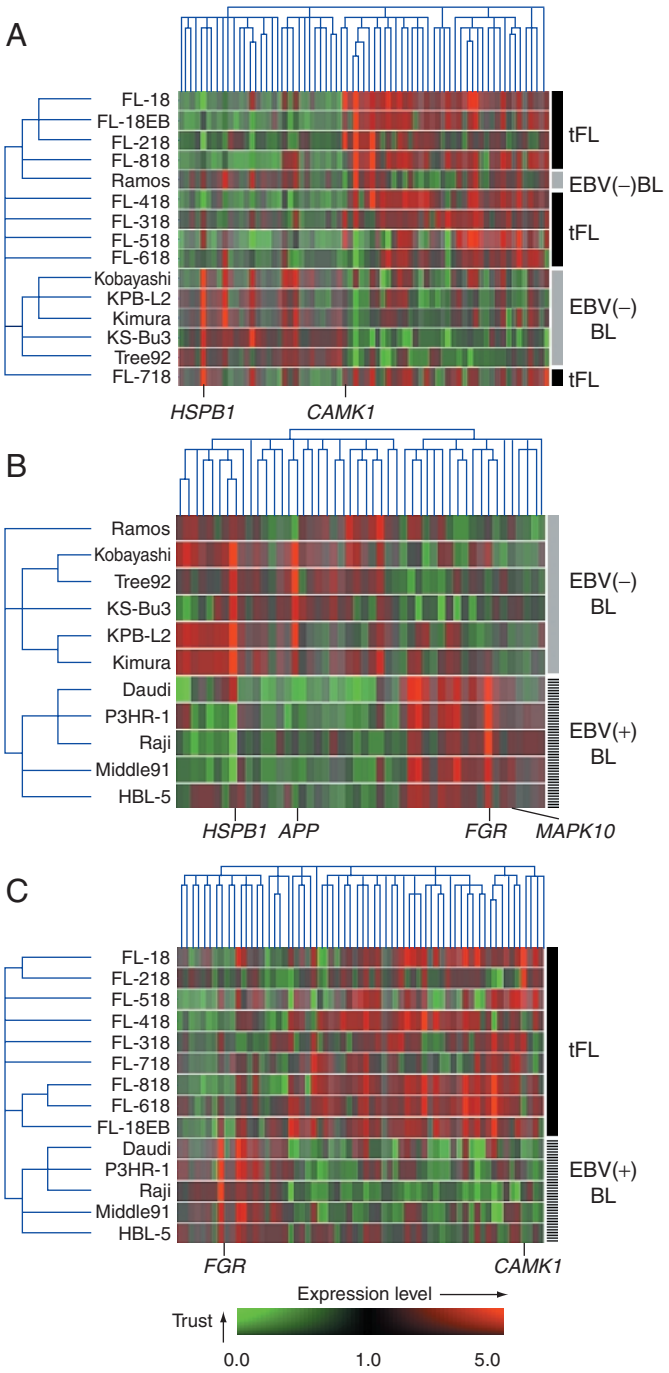

Fig. 1. Two-dimensional hierarchical clustering analyses for (A) tFL-EBV(–)BL, (B) EBV(–)BL-EBV(+)BL and (C) tFL-EBV(+)BL subtype distinction. The dendrogram of each axis, where the length of the branches represents a similarity distance of the expression profiles, shows the relation of the samples (left) and the genes (top). The gene expression values and trust are color-coded as indicated by the scale at the bottom. The gene symbols are described in Table 2.

Results

Characteristics of tFL, EBV(–)BL and EBV(+)BL cell lines. Table 1 summarizes the characteristics of tFL, EBV(–)BL and EBV(+)BL cell lines subjected to the cDNA array analysis. Although these cell lines may not be perfect replicates of primary tumors, the cells retained a majority of immunophenotypic features, as well as genetic abnormalities, possessed by the parental primary tumors.

All tFL cell lines were established and characterized in our laboratory. The patients initially developed FL, which was finally transformed to aggressive lymphoma following variable periods of low-grade disease. The 9 tFL lines carried a t(14;18)(q32;q21) on cytogenetic analysis, and the rearrangement occurred within the major breakpoint cluster region (MBR) of *BCL2* and *J* segments of the *IgH*. FL-18EB, which is a derivative of FL-18, was artificially infected by EBV at the initiation of the culture.[12] t(8;14)(q24;q32), involving *MYC* and *IgH* genes, was detected by long-distance PCR using primer sets for the *MYC* exon 2 and the constant region genes of *IgH*.[9]

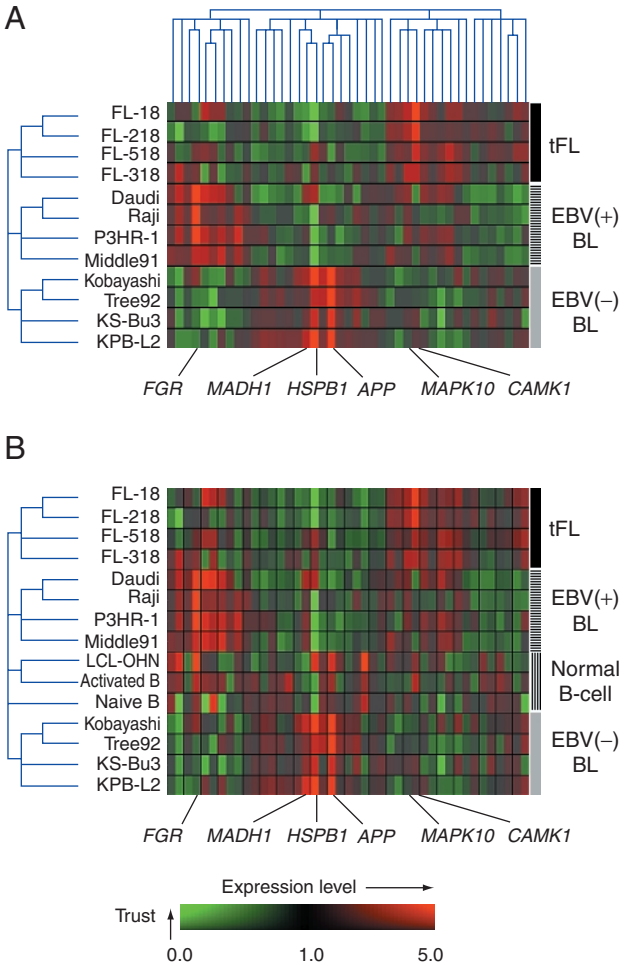

Fig. 2. Clustering dendrograms based upon expression profiles of 43 genes selected by the ANOVA *t* test. (A) Two-dimensional hierarchical clustering analysis for tFL-EBV(–)BL-EBV(+)BL subtype distinction. Four representative cell lines were selected from each subtype. The gene symbols are described in Table 2. (B) Comparison of gene expression profiles of the 3 lymphoma subtypes with normal reference B-cell materials, including EBV-transformed lymphoblastoid cell line LCL-OHN, $CD19^{+}CD27^{-}$ blood naive B-cells and activated B-cells by CD40L, IgM-crosslinking and IL-4. The order of 43 genes was maintained as determined in A, while the clustering algorithm was applied to the sample axis.

Table 2. Differentially expressed genes between tFL, EBV(−)BL and EBV(+)BL

Symbol (HUGO)	Designation	Subtype	High or low	Predictive strength
CAMK1	calcium/calmodulin-dependent protein kinase I	tFL	High	6.2046
CBLB	Cas-Br-M (murine) ectropic retroviral transforming sequence b	tFL	High	4.5951
MAPK10	mitogen-activated protein kinase 10	tFL	High	4.5951
FER	fer (fps/fes related) tyrosine kinase (phosphoprotein NCP94)	tFL	High	4.5951
GRB2	growth factor receptor-bound protein 2	tFL	High	4.5951
PRKDC	protein kinase, DNA-activated, catalytic polypeptide	tFL	High	4.5951
TADA3L[1)]	ADA3-like protein	tFL	High	4.5951
PAH	phenylalanine hydroxylase	tFL	High	4.5951
MST1R	macrophage stimulating 1 receptor (c-met-related tyrosine kinase)	tFL	Low	6.2046
TIMP1	tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)	tFL	Low	4.5951
CDKL1	cyclin-dependent kinase-like 1 (CDC2-related kinase)	EBV(−)BL	High	6.2046
APP	amyloid β (A4) precursor protein (protease nexin-II, Alzheimer disease)	EBV(−)BL	High	6.2046
HSPB1	heat shock 27 kD protein 1	EBV(−)BL	High	6.2046
MADH1	MAD, mothers against decapentaplegic homolog 1 (*Drosophila*)	EBV(−)BL	High	4.5951
WT1	Wilms tumor 1	EBV(−)BL	High	4.5951
GTF2H1	general transcription factor IIH, polypeptide 1 (62 kD subunit)	EBV(−)BL	High	4.5951
CNTN2	contactin 2 (axonal)	EBV(−)BL	High	4.5951
INSR	insulin receptor	EBV(−)BL	High	4.5951
FGFR1	fibroblast growth factor receptor 1 (fms-related tyrosine kinase 2, Pfeiffer syndrome)	EBV(−)BL	High	4.5951
SCYA1	small inducible cytokine A1, I-309	EBV(−)BL	High	4.5951
FGR	Gardner-Rasheed feline sarcoma viral (v-fgr) oncogene homolog	EBV(+)BL	High	6.2046
PIM1	pim-1 oncogene	EBV(+)BL	High	4.5951
APC	adenomatosis polyposis coli	EBV(+)BL	High	4.5951
VIL2	villin 2 (ezrin)	EBV(+)BL	High	4.5951
ETV6	ets variant gene 6 (TEL oncogene)	EBV(+)BL	High	4.5951
CASK	calcium/calmodulin-dependent serine protein kinase (MAGUK family)	EBV(+)BL	High	4.5951
CD4	CD4 antigen (p55)	EBV(+)BL	High	4.5951
ICAM1	intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	EBV(+)BL	High	4.5951
NMB	Neuromedin B	EBV(+)BL	Low	6.2046
MMP7	matrix metalloproteinase 7 (matrilysin, uterine)	EBV(+)BL	Low	6.2046
CDK2	cyclin-dependent kinase 2	EBV(+)BL	Low	4.5951
FZD2	frizzled homolog 2 (*Drosophila*)	EBV(+)BL	Low	4.5951
BCL2L1	BCL2-like 1	EBV(+)BL	Low	4.5951
IRF2	interferon regulatory factor 2	EBV(+)BL	Low	4.5951

1) Not included in the HUGO nomenclature database.

The presence of the EBV genome was determined by PCR of the region of EBER.

Identification of differentially expressed genes between tFL, EBV(−)BL and EBV(+)BL. Total cellular RNA extracted from the cell lines was reverse-transcribed and labeled with ^{32}P. The cDNA was hybridized with the Atlas Human 1.2 Array membrane and the expression profiles were analyzed using the GeneSpring software. Comparison between tFL and EBV(−)BL cell lines revealed that 68 genes were differentially expressed between the 2 subtypes at a $P<0.05$ significance level by 2-sample *t* test. Fig. 1A shows the clustering dendrogram based upon the expression of 68 genes. The EBV(−)BL cell lines, except for Ramos, generated a cluster, highlighting a gene expression pattern of this lymphoma subtype. On the other hand, 5 of 9 tFL cell lines were segregated independently from the remaining 4 and FL-718 shared some expressions with EBV(−)BL. We applied the identical comparison approach to EBV(−)BL versus EBV(+)BL and tFL versus EBV(+)BL. The results showed that expression patterns of 48 ($P<0.05$) and 63 ($P<0.05$) genes, respectively, most effectively separated the contrasted lymphoma subtypes (Fig. 1, B and C). However, Ramos was again separated from the other EBV(−)BL cell lines on the EBV(−)BL-EBV(+)BL subtype distinction and tFL cell lines were not clustered into a single branch on the tFL-EVB(+)BL distinction. Since 29 of the total of 179 genes selected were overlapped, these 2-subtype distinction studies discovered 150 genes, whose expressions are potentially correlated to one of the 3 lymphoma subtypes.

Table 3. Subtype prediction of FL/tFL, EBV(−)BL or EBV(+)BL

		Prediction			Total
		FL/tFL	EBV(−)BL	EBV(+)BL	
True subtype	FL[1)]/tFL	12	1[2)]	0	13
	EBV(−)BL	1[3)]	5	0	6
	EBV(+)BL	0	0	5	5
Total		13	6	5	24

1) Four FL clinical materials.
2) FL-418.
3) Ramos.

Identification of differentially expressed genes between tFL, EBV(−)BL and EBV(+)BL. We next identified 43 genes that distinguish the tFL, EBV(−)BL or EBV(+)BL subtype with a $P<0.05$ significance level on ANOVA. Of the 43 genes selected, 27 genes were included in the 2-subtype distinction studies. To optimally separate the 3 lymphoma subtypes by hierarchical clustering analysis, we selected 4 representative cell lines from each subtype, i.e., tFL (FL-18, FL-218, FL-318 and FL-518), EBV(−)BL (Kobayashi, KS-Bu3, Tree92 and KPB-L2) and EBV(+)BL (Daudi, Raji, Middle91 and P3HR-1). As shown in Fig. 2A, the analysis based upon the expression profiles of 43 genes readily separated the 12 cell lines into 3 sample clusters. In contrast, the gene tree diverged into 4 gene clusters; 3 clusters were correlated with each lymphoma subtype, while the genes of the remaining 1 were expressed at a lower level in EBV(+)BL than in the other 2 subtypes.

Table 2 lists the most significant 34 genes with ≥4.5951 predictive strength to delineate the 3 lymphoma subtypes. Genes expressed at higher levels in tFL than EBV(−)BL and EBV(+)BL included calcium/calmodulin-dependent protein kinase I (CaMKI; the HUGO designation, *CAMK1*) and mitogen-activated protein kinase 10 (*MAPK10*). EBV(−)BL was characterized by higher-level expression of amyloid β precursor protein (*APP*), heat shock 27 kD protein 1 (*HSPB1*) and mothers against decapentaplegic homolog 1 (*MADH1*) than the other 2 subtypes. On the other hand, Gardner-Rasheed feline sarcoma viral oncogene homolog (*FGR*) was the most significant gene to delineate EBV(+)BL.

To determine whether the gene-expression patterns of the 34 genes listed in Table 2 can predict the lymphoma subtype, we used the Class Predictor voting algorithm. Of 20 cell lines listed in Table 1, as well as 4 additional clinical materials with FL, 22 (92%) were correctly classified into FL/tFL, EBV(−)BL or EBV(+)BL (Table 3). Thus, this study provided useful distinction marker genes of the 3 lymphoma subtypes.

Comparison with normal reference B-cells. We compared the expression patterns of the 43 genes between lymphoma cells and normal reference B-cells. Hierarchical clustering was performed for the samples axis composed of the 12 lymphoma cell lines in addition to a lymphoblastoid cell line LCL-OHN, blood naive B-cells and activated B-cells, while maintaining the order of the genes as determined in Fig. 2A. As shown in Fig. 2B, the 3 normal reference B-cell materials were segregated separately from the 3 lymphoma clusters. LCL-OHN shared high-level expression of *FGR*, and of *APP* and *HSPB1* with EBV(+)BL and EBV(−)BL, respectively, but lacked expression of *MADH1*. Expressions of *APP* and *CAMK1* were both induced in B cells stimulated by CD40L, IgM-crosslinking and IL-4, while *MAPK10* was down-regulated.

Application of the marker genes to *de novo* DLBCL cell lines. We next investigated whether the marker genes for the tFL-EBV(−)-BL-EBV(+)BL subtype distinction were applicable to cell lines from *de novo* DLBCL. Table 4 lists 5 DLBCL cell lines subjected to Atlas Array analysis. The patients had no evidence of preceding low-grade lymphoma. YM,[13] Shk[14] and OYB carried a variety of genetic abnormalities involving *MYC*, *BCL2*, *BCL6* and *Ig* loci, while DLBCL-1 and DLBCL-2 lacked the known rearrangements of the 3 oncogenes. All the DLBCL samples were negative for the EBV genome.

We first compared the mRNA levels of *MYC*, *BCL2* and *BCL6* genes in the tFL, EBV(−)BL, EBV(+)BL and DLBCL subtypes in addition to the normal reference B-cells to determine whether the translocations influence the expression levels of relevant oncogenes. Since the cDNA clone of *BCL2* included in the Atlas Human 1.2 Array corresponded to the area downstream of the breakpoint clusters of t(14;18)(q32;q21), we developed an RT-PCR using primer combinations for 5′ of MBR to detect truncated *BCL2* messages of t(14;18)-carrying cells. As shown in Fig. 3A, the levels of *BCL2* mRNA in tFL were apparently higher than those of EBV(−)BL and EBV(+)BL; however, the levels in DLBCL cell lines were comparable with or even higher than those of tFL. On the other hand, the expression levels of *BCL6* were low in YM and OYB carrying genetic alterations of the gene. *BCL6* was not induced in B cells activated by CD40L, anti-IgM and IL-4, as described previously.[4] There was no clear correlation of the levels of *MYC* mRNA with particular lymphoma subtypes.

We next extended the clustering analysis to tFL-EBV(−)BL-EBV(+)BL-DLBCL subtype distinction based upon the expression levels of the 34 genes listed in Table 2. To determine whether the DLBCL cell lines constitute a single branch, the clustering algorithm was applied to both the gene and sample

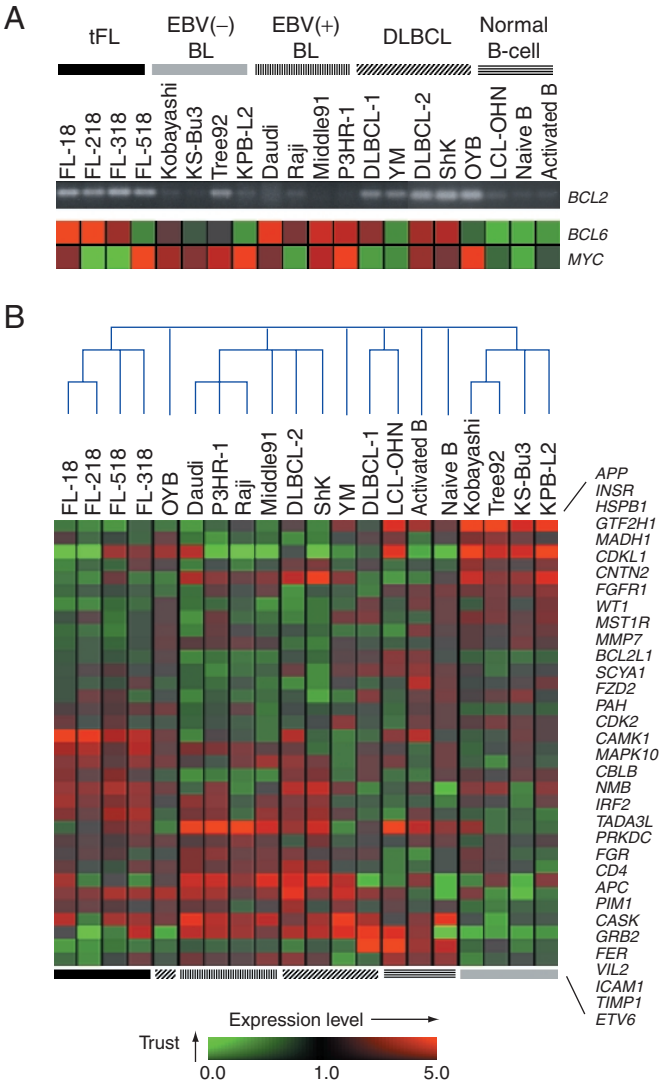

Fig. 3. Gene expression profiles of tFL, EBV(−)BL, EBV(+)BL and DLBCL subtypes, and normal reference B-cells. (A) Comparison of the mRNA levels of *BCL2* gene by RT-PCR, and those of *MYC* and *BCL6* genes by Atlas Array analysis. (B) Expression levels of 34 tFL-EBV(−)BL-EBV(+)BL distinction genes listed in Table 2. In this matrix, each row represents a gene and each column represents an individual cell material. Hierarchical clustering analysis was applied to both the gene and sample axes, although the dendrogram representing the relation between the genes is omitted for simplicity. The color scale is shown at the bottom.

Table 4. Characteristics of *de novo* DLBCL cell lines subjected to Atlas Array analysis

Designation	Chromosome abnormalities	Genetic abnormalities
OYB	?inv(3q), 14q+	Deletion within 5′ region of *BCL6*
ShK	t(8;14)(q24;q32), 18q+	5′-*BCL2*/*IgLλ*, *MYC*/*IgH*-Cα
YM	t(2;18)(p11;q21), t(3;16)(q27;p11)	5′-*BCL2*/*IgLκ*, *IL-21R*/*BCL6*
DLBCL-1	Complex	No rearrangement of *MYC*, *BCL2* and *BCL6*
DLBCL-2	Tetraploid, 3q+	No rearrangement of *MYC*, *BCL2* and *BCL6*

axes. As shown in Fig. 3B, the matrix again highlighted characteristic expression patterns of tFL, EBV(−)BL and EBV(+)BL. In contrast, DLBCL cell lines showed variable gene expression profiles. ShK, which carried a *MYC/IgH*-Cα translocation, and DLBCL-2 were segregated into the EBV(+)BL branch, sharing expressions of calcium/calmodulin-dependent serine protein kinase (*CASK*), growth factor receptor-bound protein 2 (*GRAB2*), intercellular adhesion molecule 1 (*ICAM1*) and villin 2 (*VIL2*). DLBCL-1 showed similarities with LCL-OHN, including high level expression of tissue inhibitor of metalloproteinase 1 (*TIMP1*) and ICAM1. The remaining 2 had no apparent similarities with tFL, EBV(−)BL or EBV(+)BL. These results may be a reflection of the variable genetic abnormalities of the DLBCL cell lines studied. Furthermore, separation of DLBCL from tFL suggests a difference between large cell lymphoma arising from FL and that lacking such a background.

Validation of differential gene expression by real-time quantitative RT-PCR. To confirm the gene expression profiles characterizing each lymphoma subtype, we determined the mRNA levels of 5 representative marker genes, i.e., *CAMK1*, *APP*, *MADH1*, *FGR* and *BCL2*, by real-time quantitative RT-PCR. Table 5 represents the sequences of oligonucleotide primers and a fluorogenic probe for each gene. The cell lines analyzed were the 25 cell lines listed in Tables 1 and 4, and 4 additional cell lines, i.e., Black93[9] (EBV(−)BL subtype), and KIS-1,[15] Manca[16] and DL-4[2] (DLBCL subtype). Black93, Manca and DL-4 carry the

Table 5. Sequences of oligonucleotide primers and fluorogenic probes for *CAMK1*, *APP*, *MADH1*, *FGR* and *BCL2* genes

Gene	Accession no.		Sequence (5′ to 3′)
CAMK1	L41816	Forward	AGAATGAGATTGCTGTCCTGCA
		Reverse	CCCCCACTCTCATAGATGTCATC
		Probe	AAGATCAAGCACCCCAACATTGTAGCCCT
APP	Y00264	Forward	ACTGACCACTCGACCAGGTTCT
		Reverse	TTTGAACCCACATCTTCTGCAA
		Probe	TCCTGAGTCATGTCGGAATTCTGCA
MADH1	U57456	Forward	GCCGATGGACACAAACATGA
		Reverse	AAGCAACCGCCTGAACATCT
		Probe	TCCCCTGCCCTCAGAAATCAACAGAG
FGR	M19722	Forward	GCATCCTCATTTGCCCACAT
		Reverse	AATCCCTGACACACCCCTGAT
		Probe	CAACTACAGCAACTTCTCCTCTCAGGCCA
BCL2	M14745	Forward	GGAGGCTGGGATGCCTTT
		Reverse	CCAGATAGGCACCCAGGGT
		Probe	CCAGCATGCGGCCTCTGTTTGATTT

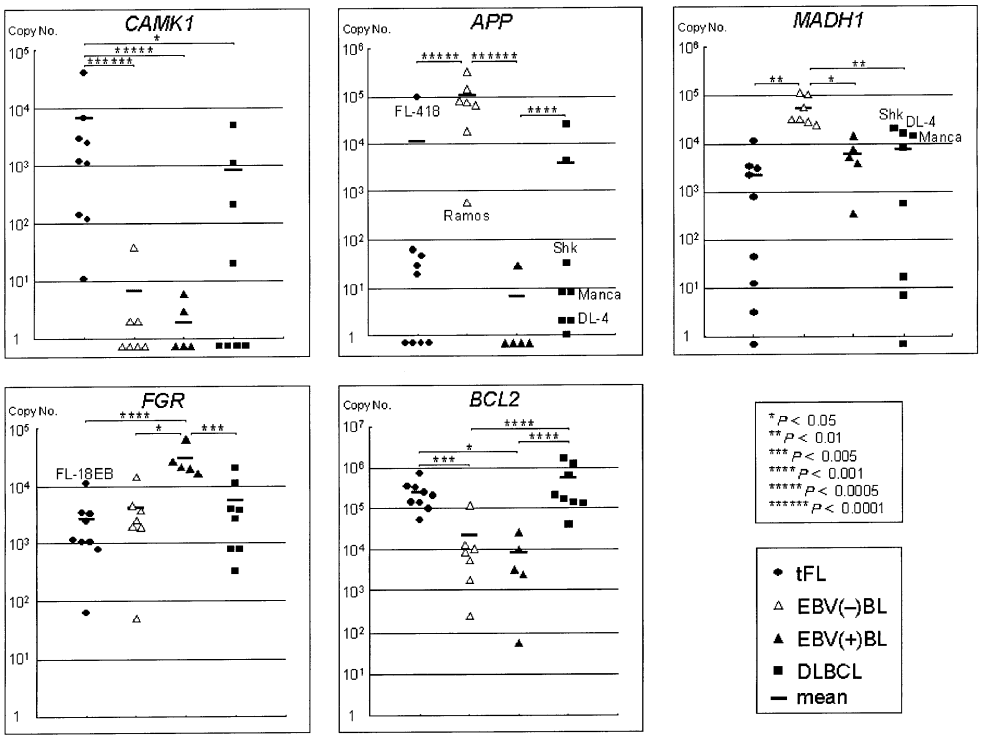

Fig. 4. Real-time quantitative RT-PCR assay to measure the mRNA levels of *CAMK1*, *APP*, *MADH1*, *FGR* and *BCL2* genes in tFL (n=9), EBV(−)BL (n=7), EBV(+)BL (n=5) and DLBCL (n=8) cell lines. The calculated copy number of mRNA of each gene was normalized by the level of 18S rRNA, the level of which of Raji cells was defined as 1. Copy numbers less than 1 were plotted under the baseline. The additional cell lines analyzed that are not included in Tables 1 and 4 were: Black93 (EBV(−)BL), KIS-1, Manca and DL-4 (DLBCL).

t(8;14)(q24;q32) and *MYC/IgH* fusion gene; however, the primary tumors of the latter 2 lines showed non-BL histopathology. KIS-1 had a t(9;14)(p13;q32), leading to *PAX5/IgH* gene fusion.

Fig. 4 shows scattered plot diagrams comparing the expression levels of these 5 marker genes. The calculated mRNA copy number was normalized by the level of 18S rRNA of each sample. Although the levels of CAMK1 in tFL cell lines were considerably variable, the differences from those of EBV(−)BL, EBV(+)BL and DLBCL were significant. The *APP* and *MADH1* expression levels were significantly higher in EBV(−)BL than in the other 3 subtypes. High- and low-level expression of the *APP* gene in FL-418 and Ramos, respectively, lead to the failure of the subtype prediction of these 2 cell lines by the Class Predictor algorithm (Table 3). ShK, Manca and DL-4 carrying t(8;14)(q24;q32) expressed *MADH1* at higher levels than the remaining DLBCL cell lines, while *APP* was at low levels. *FGR* was apparently a distinctive marker gene to discriminate EBV(+)BL from the other 3 subtypes. The level of *FGR* mRNA of FL-18EB cells was 4-fold higher than that of FL-18; however, the level was lower than the range of EBV(+)BL cells. DLBCL cell lines expressed *BCL2* at high levels irrespective of the presence or absence of rearrangement of the gene. These results, obtained by real-time PCR assays, confirmed the significant association of high-level expression of CAMKI with tFL, APP and MADH1 with EBV(−)BL, FGR with EBV(+)BL, and BCL2 with tFL and DLBCL.

Discussion

Here, we applied cDNA array technology to compare the expression profiles of B-cell lymphoma cell lines, and identified molecular markers that characterize each lymphoma subtype. The disadvantage of studies using cell lines is that additional transforming events may be required to immortalize tumor cells *in vitro*; therefore cell lines may not be true representatives of tumor cells *in vivo*. Nevertheless, the observed difference is potentially a reflection of fundamental genetic and biological characteristics of each lymphoma subtype, since the cell lines used presumably share gene expression programs associated with the properties of B cells, resistance to chemotherapeutic agents and rapid proliferation in the tissue culture environment. The selection of *FGR*, the expression of which was previously shown to be induced by EBV,[17, 18] with the highest predictive strength as the marker for EBV(+)BL supports the validity of this approach. Another advantage of the study of cell lines over that of clinical materials is the elimination of the profile of non-malignant components contaminating lymphoma specimens, although this may conversely lead to missing information on tumor cell-reactive cell interaction.[6]

The tFL-EBV(−)BL-EBV(+)BL distinction study by a hierarchical clustering algorithm revealed that 43 of 1176 genes were differentially expressed between the 3 subtypes, and a class prediction test based upon the expression patterns of the most significant 34 genes was able to classify the subtypes of 92% of the samples tested. By comparison with normal reference B-cell materials, the expression patterns of the selected genes were characteristic of lymphoma cells. We next found that DLBCL cell lines exhibited variable expression patterns, and were either separated from the former 3 subtypes or segregated with EBV(+)BL by clustering analysis. Finally, we confirmed the association of representative marker genes with each lymphoma subtype by real-time quantitative RT-PCR.

The *APP* gene is expressed in EBV(−)BL at significantly higher levels than tFL, EBV(+)BL and DLBCL. The gene, located on chromosome band 21q21.3, encodes a membrane-associated glycoprotein that is synthesized in a variety of tissues.[19, 20] A4-amyloid, which is the major component of senile plaques and neurofibrillary tangles found in the brain of patients with Alzheimer's disease, is derived from the larger precursor APP protein.[19] On the other hand, *APP* expression is detectable in peripheral blood mononuclear cells (PBMCs) at both the mRNA and protein levels.[21] When PBMCs are stimulated by phytohemagglutinin, anti-CD3 antibody for T cells and anti-IgM antibody for B cells, respectively, the cells are induced not only to express *APP* on their surface,[22, 23] but also to secrete the soluble form of APP in the culture medium.[24] Thus, it is possible that APP is an activation antigen of lymphocytes and that high-level *APP* expression may be related to the rapid proliferation rate of EBV(−)BL cells. APP has various functions, including roles in cell-cell and cell-extracellular matrix adhesion[25] and protease inhibition.[26] These findings are consistent with the idea that high-level *APP* expression may be related to the rapid proliferation rate of EBV(−)BL cells, as well as the histopathology, which is characterized by prominent macrophage infiltration.

APP transcripts consist of several alternatively spliced mRNAs (corresponding to APP_{770}, APP_{751}, APP_{714}, APP_{695} and APP_{563}) that are differentially regulated in a cell type-specific and developmentally determined manner.[21] APP_{695}, which lacks the protease inhibitor domain, is the most abundantly expressed in neurons.[20, 21] In contrast, the most abundant mRNA species in PBMCs is APP_{751}, while APP_{770} mRNA is the predominant transcript in a neuroblastoma cell line.[21] We performed RT-PCR to detect all the species of *APP* mRNA and found that the relative amounts of *APP* mRNA in EBV(−)BL were comparable with those of PBMCs (data not shown).

MADH1 is another candidate gene to delineate EBV(−)BL. The gene encodes Smad1, which belongs to a class of proteins that function as intracellular signaling effectors for the transforming growth factor-β (TGF-β) superfamily, including TGF-βs, activins and bone morphogenic proteins (BMPs).[27] Smad2 and Smad3 act downstream of TGF-β and activin, whereas Smad1 in addition to Smad5 and Smad8 are activated by BMP. The TGF-β-Smad2/3 signaling pathway induces not only growth inhibition, but also apoptosis in several cell types; treatment of EBV(−)BL cells with TGF-β leads to apoptosis or growth arrest in the G1 phase of the cell cycle.[28, 29] In contrast, EBV(+)BL cells are resistant to these anti-proliferative and pro-apoptotic effects of TGF-β treatment.[29] It is possible that EBV-encoded proteins interact with the Smads to escape the tumor suppressor effects, although a down-regulation of TGF-β type II receptor was demonstrated to block the TGF-β response in EBV(+)BL cells.[29] In this study, we found that *MADH1* mRNA is expressed at significantly higher levels in EBV(−)BL than tFL and EBV(+)BL, and t(8;14)(q24;q32)-positive DLBCL cell lines expressed this gene at high levels. Currently, no data are available as to whether the BMPs-Smad1 signaling pathway functions in normal or neoplastic B cells.

The levels of *CAMK1* expression were considerably variable among the tFL cell lines studied and some of the DLBCL cell lines expressed the gene at high levels. Nevertheless, the statistical analysis indicated that *CAMK1* is a marker gene characterizing tFL. An increase in intracellular Ca^{2+} is a ubiquitous signal that induces numerous cellular responses. Ca^{2+} associates with calmodulin (CaM) to form a Ca^{2+}/CaM complex that binds to many signal transducers, including CaMKs.[30] Induction of CaMKII in Jurkat T cells leads to inhibition of transcription of *interleukin-2* gene,[31] suggesting that the Ca^{2+} signal through CaMKII affects T-cell regulation. CaMKI and CaMKIV are strongly expressed in the brain, whereas the former is ubiquitously expressed in many tissues.[32] Both enzymes share a common requirement for phosphorylation by an upstream protein kinase (CaMKK).[32] Although, at present, little is known about the substrates of CaMKI, cAMP response element-binding protein (CREB) was reported to be a target that

is phosphorylated and activated by CaMKI.[30, 32] Thus, it is likely that high-level expression of CaMKI affects transcription of a variety of genes in neoplastic B cells.

Chromosomal translocation in B-cell lymphoma leads to deregulated expression of the oncogene product located at the breakpoint.[33] However, there is no clear difference in the level of expression of a particular oncogene between lymphoma tissues carrying the relevant translocation and those lacking the gene rearrangement. We showed here that the *BCL2* mRNA levels of tFL with t(14;18)(q32;q21) measured by real-time RT-PCR were significantly higher than those of EBV(−)BL and EBV(+)BL, whereas DLBCL cell lines expressed *BCL2* at high levels independently of rearrangement of the gene. By analogy, *MYC* was not selected in the tFL-EBV(−)BL-EBV(+)BL subtype distinction. On the other hand, expression levels of *BCL6* were not necessarily associated with genetic abnormalities of the gene. These findings suggest that the mRNA level of the oncogene affected by the corresponding translocation cannot be a marker of B-cell lymphoma subtype. The present study using cDNA array analysis revealed previously unknown marker genes and potential signal transduction pathways. These differential gene expressions might be related to the clinicopathological features of each B-cell lymphoma subtype.

The authors wish to thank Drs. N. Kadowaki and T. Kitawaki for their help in isolation and stimulation of peripheral blood B-cells. This work was supported by a Grant-in-Aid (13470204) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

1. Ong ST, Le Beau MM. Chromosomal abnormalities and molecular genetics of non-Hodgkin's lymphoma. *Semin Oncol* 1998; **25**: 447–60.
2. Akasaka T, Akasaka H, Ueda C, Yonetani N, Maesako Y, Shimizu A, Yamabe H, Fukuhara S, Uchiyama T, Ohno H. Molecular and clinical features of non-Burkitt's, diffuse large-cell lymphoma of B-cell type associated with the c-*MYC*/immunoglobulin heavy-chain fusion gene. *J Clin Oncol* 2000; **18**: 510–8.
3. Wellmann A, Thieblemont C, Pittaluga S, Sakai A, Jaffe ES, Siebert P, Raffeld M. Detection of differentially expressed genes in lymphomas using cDNA arrays: identification of clusterin as a new diagnostic marker for anaplastic large-cell lymphomas. *Blood* 2000; **96**: 398–404.
4. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000; **403**: 503–11.
5. Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltnane JM, Hurt EM, Zhao H, Averett L, Yang L, Wilson WH, Jaffe ES, Simon R, Klausner RD, Powell J, Duffey PL, Longo DL, Greiner TC, Weisenburger DD, Sanger WG, Dave BJ, Lynch JC, Vose J, Armitage JO, Montserrat E, Lopez-Guillermo A, Grogan TM, Miller TP, LeBlanc M, Ott G, Kvaloy S, Delabie J, Holte H, Krajci P, Stokke T, Staudt LM. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* 2002; **346**: 1937–47.
6. Shipp MA, Ross KN, Tamayo P, Weng AP, Kutok JL, Aguiar RC, Gaasenbeek M, Angelo M, Reich M, Pinkus GS, Ray TS, Koval MA, Last KW, Norton A, Lister TA, Mesirov J, Neuberg DS, Lander ES, Aster JC, Golub TR. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* 2002; **8**: 68–74.
7. Zhang W, Laborde PM, Coombes KR, Berry DA, Hamilton SR. Cancer genomics: promises and complexities. *Clin Cancer Res* 2001; **7**: 2159–67.
8. Akasaka T, Muramatsu M, Kadowaki N, Ohno H, Ishizaki K, Yamabe H, Fukuhara S, Okuma M. *p53* mutation in B-cell lymphoid neoplasms with reference to oncogene rearrangements associated with chromosomal translocations. *Jpn J Cancer Res* 1996; **87**: 930–7.
9. Akasaka T, Akasaka H, Ohno H. Polymerase chain reaction amplification of long DNA targets: application to analysis of chromosomal translocations in human B-cell tumors (Review). *Int J Oncol* 1998; **12**: 113–21.
10. Arpin C, Dechanet J, Van Kooten C, Merville P, Grouard G, Briere F, Banchereau J, Liu YJ. Generation of memory B cells and plasma cells *in vitro*. *Science* 1995; **268**: 720–2.
11. Akasaka H, Akasaka T, Kurata M, Ueda C, Shimizu A, Uchiyama T, Ohno H. Molecular anatomy of *BCL6* translocations revealed by long-distance polymerase chain reaction-based assays. *Cancer Res* 2000; **60**: 2335–41.
12. Doi S, Ohno H, Tatsumi E, Arita Y, Kamesaki H, Fukuhara S, Nishikori M, Miwa H, Kita K, Hatanaka M, Honjo T, Fujii H, Uchino H. Lymphoma cell line (FL-18) and Epstein-Barr virus-carrying cell line (FL-18-EB) obtained from a patient with follicular lymphoma: monoclonal derivation and different properties. *Blood* 1987; **70**: 1619–23.
13. Ueda C, Akasaka T, Kurata M, Maesako Y, Nishikori M, Ichinohasama R, Imada K, Uchiyama T, Ohno H. The gene for interleukin-21 receptor is the partner of *BCL6* in t(3;16)(q27;p11), which is recurrently observed in diffuse large B-cell lymphoma. *Oncogene* 2002; **21**: 368–76.
14. Yonetani N, Ueda C, Akasaka T, Nishikori M, Uchiyama T, Ohno H. Heterogeneous breakpoints on the immunoglobulin genes are involved in fusion with the 5′ region of *BCL2* in B-cell tumors. *Jpn J Cancer Res* 2001; **92**: 933–40.
15. Morrison AM, Jager U, Chott A, Schebesta M, Haas OA, Busslinger M. Deregulated *PAX-5* transcription from a translocated IgH promoter in marginal zone lymphoma. *Blood* 1998; **92**: 3865–78.
16. Nishikori M, Hansen H, Jhanwar S, Fried J, Sordillo P, Koziner B, Lloyd K, Clarkson B. Establishment of a near-tetraploid B-cell lymphoma line with duplication of the 8;14 translocation. *Cancer Genet Cytogenet* 1984; **12**: 39–50.
17. Cheah MS, Ley TJ, Tronick SR, Robbins KC. *fgr* proto-oncogene mRNA induced in B lymphocytes by Epstein-Barr virus infection. *Nature* 1986; **319**: 238–40.
18. Klein C, Busson P, Tursz T, Young LS, Raab-Traub N. Expression of the c-*fgr* related transcripts in Epstein-Barr virus-associated malignancies. *Int J Cancer* 1988; **42**: 29–35.
19. Kang J, Lemaire HG, Unterbeck A, Salbaum JM, Masters CL, Grzeschik KH, Multhaup G, Beyreuther K, Muller-Hill B. The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature* 1987; **325**: 733–6.
20. Li QX, Fuller SJ, Beyreuther K, Masters CL. The amyloid precursor protein of Alzheimer disease in human brain and blood. *J Leukoc Biol* 1999; **66**: 567–74.
21. Ledoux S, Rebai N, Dagenais A, Shaw IT, Nalbantoglu J, Sekaly RP, Cashman NR. Amyloid precursor protein in peripheral mononuclear cells is up-regulated with cell activation. *J Immunol* 1993; **150**: 5566–75.
22. Monning U, Konig G, Banati RB, Mechler H, Czech C, Gehrmann J, Schreiter-Gasser U, Masters CL, Beyreuther K. Alzheimer βA4-amyloid protein precursor in immunocompetent cells. *J Biol Chem* 1992; **267**: 23950–6.
23. Bullido MJ, Munoz-Fernandez MA, Recuero M, Fresno M, Valdivieso F. Alzheimer's amyloid precursor protein is expressed on the surface of hematopoietic cells upon activation. *Biochim Biophys Acta* 1996; **1313**: 54–62.
24. Monning U, Konig G, Prior R, Mechler H, Schreiter-Gasser U, Masters CL, Beyreuther K. Synthesis and secretion of Alzheimer amyloid βA4 precursor protein by stimulated human peripheral blood leucocytes. *FEBS Lett* 1990; **277**: 261–6.
25. Schubert D, Jin LW, Saitoh T, Cole G. The regulation of amyloid beta protein precursor secretion and its modulatory role in cell adhesion. *Neuron* 1989; **3**: 689–94.
26. Miyazaki K, Hasegawa M, Funahashi K, Umeda M. A metalloproteinase inhibitor domain in Alzheimer amyloid protein precursor. *Natur*e 1993; **362**: 839–41.
27. Derynck R, Zhang Y, Feng XH. Smads: transcriptional activators of TGF-β responses. *Cell* 1998; **95**: 737–40.
28. Saltzman A, Munro R, Searfoss G, Franks C, Jaye M, Ivashchenko Y. Transforming growth factor-β-mediated apoptosis in the Ramos B-lymphoma cell line is accompanied by caspase activation and Bcl-XL downregulation. *Exp Cell Res* 1998; **242**: 244–54.
29. Inman GJ, Allday MJ. Resistance to TGF-β1 correlates with a reduction of TGF-β type II receptor expression in Burkitt's lymphoma and Epstein-Barr virus-transformed B lymphoblastoid cell lines. *J Gen Virol* 2000; **81**: 1567–78.
30. Means AR. Regulatory cascades involving calmodulin-dependent protein kinases. *Mol Endocrinol* 2000; **14**: 4–13.
31. Nghiem P, Ollick T, Gardner P, Schulman H. Interleukin-2 transcriptional block by multifunctional Ca^{2+}/calmodulin kinase. *Nature* 1994; **371**: 347–50.
32. Corcoran EE, Means AR. Defining Ca^{2+}/calmodulin-dependent protein kinase cascades in transcriptional regulation. *J Biol Chem* 2001; **276**: 2975–8.
33. Cleary ML. Oncogenic conversion of transcription factors by chromosomal translocations. *Cell* 1991; **66**: 619–22.