THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 7, pp. 4803–4811, February 16, 2007

Tumor-specific and Proliferation-specific Gene Expression Typifies Murine Transgenic B Cell Lymphomagenesis*[S]

Received for publication, June 19, 2006, and in revised form, October 27, 2006 Published, JBC Papers in Press, December 13, 2006, DOI 10.1074/jbc.M605870200

Marc E. Lenburg[‡], Anupama Sinha[§], Douglas V. Faller[§], and Gerald V. Denis[§1]

From the [‡]Department of Genetics and Genomics, and [§]Cancer Research Center, Boston University School of Medicine, Boston, Massachusetts 02118

The dual bromodomain protein Brd2 is closely related to the basal transcription factor $TAF_{II}250$, which is essential for cyclin A transactivation and mammalian cell cycle progression. In transgenic mice, constitutive lymphoid expression of Brd2 causes a malignancy most similar to human diffuse large B cell lymphoma. We compare the genome-wide transcriptional expression profiles of these lymphomas with those of proliferating and resting normal B cells. Transgenic tumors reproducibly show differential expression of a large number of genes important for cell cycle control and lymphocyte biology; expression patterns are either tumor-specific or proliferation-specific. Several of their human orthologs have been implicated in human lymphomagenesis. Others correlate with human disease survival time. *BRD2* is underexpressed in some subtypes of human lymphoma and these subtypes display a number of similarities to the *BRD2*-mediated murine tumors. We illustrate with a high degree of detail that cancer is more than rampant cellular proliferation, but involves the additional transcriptional mobilization of many genes, some of them poorly characterized, which show a tumor-specific pattern of gene expression.

Diffuse large cell lymphoma (DLCL)[2] is an aggressive B cell malignancy that accounts for 40% of non-Hodgkin's lymphomas (NHL) and is the most common type of lymphoma in adults (1). NHL incidence rates have roughly doubled since 1970 (2), and 56,000 new cases will be diagnosed in 2006, with 19,000 deaths, making NHL the fifth ranking cause of cancer deaths in the United States (1). Key features of DLCL are its extremely heterogeneous presentation (3–5), which gives rise to significant individual-level differences in responsiveness to therapy (4), and 60% mortality (5), making treatment decisions difficult (3, 5) and making a molecular understanding of the origin of this B cell malignancy a distant goal.

We have published a transgenic (Tg) mouse model for DLCL, based on constitutive (Eμ), B lineage-restricted expression of a newly described oncogene, the double bromodomain protein Brd2 (6), which is related to the basal transcription factor $TAF_{II}250$ (7), a key participant in the transcriptional control of cyclin A and cell cycle progression (8). The expression of cyclin A is up-regulated in Tg B cells (6), and the cell cycle is destabilized. Tg models make it possible to examine the consequences of DLCL in an inbred animal strain with a single initiating genetic lesion. This feature provides a model for DLCL with the least possible intrinsic diversity. Murine lymphoma models have been well established for some time, such as Eμ-*Myc* mice (9). However, the Eμ-*Myc* phenotype includes malignancies at several stages of lymphocyte development and greater cellular diversity than is seen in Eμ-*BRD2* (Tg) mice. Unlike Eμ-*Myc* mice, which frequently develop fatal pre-B or B-cell lymphomas after 6 weeks of age (10), Tg mice develop mature B cell lymphomas with monoclonal immunoglobulin (Ig) genes after 28 weeks (6). The clinical relevance of this model is partly justified by robust *in vitro* evidence that *BRD2* is an oncogene situated within the class II major histocompatibility complex at human chromosome 6p21 (11). Unlike all other genes in this region, *BRD2* alone appears not to be involved directly in antigen processing, but has a signal transduction function (11).

Brd2 is a nuclear-localized transcription factor kinase (12) and its *Drosophila* ortholog is female sterile homeotic (*fsh*) (7, 13, 14), which encodes a homeotic protein and probable transcription regulator that is an upstream activator of the *trithorax* locus (15). *MLL*, the human homolog of *trithorax*, is disrupted in mixed-lineage (myeloid and lymphoid) human leukemias associated with 11q23 translocations (16, 17). Recent studies suggest that bromodomain-containing proteins provide a scaffold or platform for transcription or chromatin-remodeling complexes, anchoring them to nucleosomes (18–20). Furthermore, oncogenic fusion proteins that contain bromodomains participate in neoplasia, such as the BET protein Brd4/MCAP (21, 22), which is closely related to Brd2, sharing its dual bromodomain and extraterminal (ET) domain structure (21). *BRD4* is rearranged in t (15, 19) translocations associated with aggressive carcinomas of the respiratory tract and its fusion partners encode likely oncoproteins (23). Also, translocations that fuse the bromodomain and histone acetyltransferase (HAT) domain of CREB-binding protein (CBP) to MLL (24–26) or to the monocytic leukemia zinc finger protein (27), or that fuse p300 to MLL, create oncoproteins (28) associated with

* This work was supported in part by United States Public Health Service Grants CA75107 and CA102889 (to G. V. D.) from the NCI, National Institutes of Health and RSG05072-LIB (to G. V. D.) from the American Cancer Society. 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.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Table S1 and File 1.

[1] To whom correspondence should be addressed: Cancer Research Center, Rm. K521, Boston University School of Medicine, 80 East Concord St., Boston, MA 02118. Tel.: 617-414-1371; Fax: 617-638-5673; E-mail: gdenis@bu.edu.

[2] The abbreviations used are: DLCL, diffuse large cell lymphoma; FACS, flow cytometry; FDR, false discovery rate; FITC, fluorescein isothiocyanate; FL, follicular lymphoma; GC, germinal center; HAT, histone acetyltransferase; HBSS, Hank's buffered salt solution; Ig, immunoglobulin; NHL, non-Hodgkin's lymphoma; PBS, phosphate-buffered saline; PCA, principal component analysis; PE, phycoerythrin; Tg, transgenic.

acute leukemias. The bromodomain is required for full transforming activity in at least one of these cases (25). Therefore, inappropriate targeting of HATs to certain promoters can cause lymphoid malignancy.

Our study was motivated by the genome-wide transcriptional profiling of human lymphomas that has revealed previously unsuspected subclasses of DLCL, the transcriptional "fingerprint" of which can be used to classify patients into the most appropriate therapeutic group (3). Some of these human profiles are correlated with the fingerprint of mitogenically activated B cells, and are associated with a more aggressive course and higher mortality rates (a five-year median survival rate of 16%), whereas others are correlated with the fingerprint of germinal center (GC) B cells and are associated with more indolent disease and a 5-year median survival rate of 76% (4). As a result of the DLCL phenotype of Tg mice and the lack of studies suggesting a direct role for alterations in *BRD2* expression in human lymphomagenesis or leukemogenesis, we chose to study the gene expression patterns of this malignancy in greater detail. An advantage of the relatively late onset of full blown malignant lymphomas in Tg mice is the ability to collect lymphoid tissue from mice at progressive stages of lymphomagenesis, from very early, "pre-malignant" splenic B cells through to aggressive, end-stage splenic lymphomas associated with peripheral leukemia and organ failure. We compared the genome-wide transcriptional signatures we obtained from malignant cells to normal resting B cells and mitogenically activated B cells from syngeneic mice.

Based on our previous transcriptional profiling in this lymphoma model (6), we hypothesized that we would identify several patterns of differential gene expression, with one group representing a "proliferation signature," similar to mitogenically activated normal B cells, another group representing a "cancer signature," comprising genes specific to lymphomagenesis and malignancy. We used Affymetrix arrays that detect the expression of 22,500 murine transcripts and found over 3,000 that are differentially expressed between resting B cells and aggressive Tg lymphomas. Examination of the expression of these genes in earlier stage lymphomas and activated B cells resulted in the identification of clusters of co-expressed genes, and reveals a number of genes that are specifically altered by lymphomagenesis, and not B cell activation, which may be important in human DLCL lymphomagenesis.

EXPERIMENTAL PROCEDURES

Mice—Animals were handled humanely in accordance with Institutional Animal Care and Use Committee (IACUC), state and Federal regulatory requirements; the Boston University Medical Campus IACUC reviewed and authorized this study. Tg mice developed lymphomas over the course of the experiment and exhibited clinical signs, including failure to nest, hunched posture, loss of appetite, sleepiness, inactivity, and ruffled fur (6).

Flow Cytometry—For analysis of splenocytes from normal or lymphoma mice, single cell suspensions of spleen were dispersed with a 70-μm nylon cell strainer (BD Falcon) in HBSS and stained in PBS supplemented with 0.1% bovine serum albumin. For surface cytofluorimetric analyses, FITC-coupled antibody against the pan B cell marker B220 and PE-coupled antibodies against mouse CD5, IgM, and IgD were from eBioscience (San Diego, CA). Antibodies against CD25, CD69, B7-1, B7-2, and rat isotype controls were from BD Pharmingen. Cell staining was performed in the presence of F_c receptor blocking antibody (clone 2.4G2, eBioscience), and cells were detected by flow cytometry with a FACSCalibur system (Becton Dickinson, San Jose, CA). Signals were analyzed with Cell Quest software. Viability (>99%) was determined by trypan blue. Cells were stained with B220-FITC antibodies and compared with PE-coupled antibodies for other antigens (6).

B Cell Activation in Vitro—Spleens were isolated and, after erythrocyte lysis, B cells were purified with anti-CD43 negative selection using magnetic beads (Miltenyi Biotec, Auburn, CA) in PBS supplemented with 2 mM EDTA and 0.5% bovine serum albumin as previously described (6). B cells were cultured at 37 °C, 5% CO_2 and 100% humidity in Na_2HCO_3-buffered RPMI1640 medium, supplemented with 10% heat-inactivated fetal bovine serum (Whittaker Bioproducts, Walkersville, MD), 50 μM 2-mercaptoethanol, glutamine, penicillin, and streptomycin, stimulated for 48 h with 10 μg/ml each of goat anti-mouse IgM $F(ab')_2$ antibody, μ chain-specific (Jackson ImmunoResearch, West Grove, PA), rat anti-mouse CD40 (BD Pharmingen, San Diego, CA) and 10 ng/ml murine recombinant interleukin-4 (eBioscience). Cell surface markers for B cell activation were measured with flow cytometry as above.

Microarray Analysis—RNA was isolated from purified B cells as above, extracted with acid phenol and precipitated from 2-propyl alcohol, quantified, examined by nucleic acid electrophoresis for purity and integrity, labeled and hybridized in accordance with standard Affymetrix protocols to Murine Genome 430A Arrays (Affymetrix, Santa Clara, CA). These arrays contained representative sequences for ~22,500 transcripts that were identified from the UniGene data base (Build 107, June 2002).

Whole genome expression profiles were generated for 26 samples. Probe hybridization intensities were linearly scaled to a mean probe hybridization intensity of 500 units for each array using Microarray Suite 5.0 (MAS 5.0) software (Affymetrix, Santa Clara, CA). The signal intensities from the twenty-two probes in each probeset were then used to determine an overall expression level for each transcript and a measure of sequence-specific hybridization. These data and the probe level data have been submitted to the NCBI Gene Expression Omnibus (29) under accession number GSE6136. To eliminate expression measurements of genes that are not expressed in B cells, we identified 5,561 probesets among the 22,690 tested that failed to exhibit significant sequence-specific hybridization in any of the twenty-six samples using the detection calls generated by MAS 5. The hybridization intensities of the remaining 17,129 probesets were $\log_{10}$ transformed and used in subsequent analyses.

For some analyses, hybridization intensities were further Z-Score normalized such that each probeset had a mean intensity of 0 with a standard deviation of 1. Principal component analysis (PCA) was performed with these normalized data in DecisionSite for Functional Genomics (Spotfire, Somerville, MA). Student's *t*-tests were performed using Excel (Microsoft,

Redmond, WA). A false discovery rate (FDR) for each *t* test *p* value was calculated (30), and the smallest *p* value that gave an FDR > 0.05 ($p = \sim 0.01$) was used as the threshold *p* value for evidence of significant differential expression. Hierarchical clustering was performed with normalized data and Ward's method (31) in DecisionSite. Fold changes for each probeset are expressed as $\log_2$ relative to the average hybridization intensity across all samples.

Comparison with Human Lymphoma Datasets—Mouse probesets for transcripts orthologous to human genes known to be involved in lymphomagenesis were identified in a multi-step procedure (see supplemental Methods). We used the Monti *et al.* dataset (32) to determine the expression of human genes orthologous to mouse genes that are differentially expressed during Tg lymphomagenesis in a collection of human clinical lymphoma samples. We first processed the dataset to remove probesets with sequence-specific hybridization in less than 9 of the 176 human lymphoma samples (<5% of samples) and $\log_{10}$-transformed the remaining hybridization intensities. The resulting dataset was then searched with NetAffx (33) to identify probesets that detect expression of transcripts orthologous to various mouse genes of interest.

Analyses of variance were used to identify genes that vary between the three DLCL lymphoma classes identified by Monti *et al.* (32) and were performed in PartekPro (Partek). Pearson correlation between gene expression and patient survival time was performed using the cor.test function in the stats package of R 2.2.1.3.[3]

RESULTS

We characterized B cell lymphomas as they occurred sporadically in breeding lines of Tg mice over 2 years. We performed phenotypic analyses of splenic pathology, using flow cytometry and immunohistochemistry, and classified each tumor according to its stage of advancement, and confirmed our previously published histopathology of spleen (6). After we had assembled a well characterized tumor bank, we performed microarray analysis of RNA extracted from purified B cells of tumors or controls. Microarray analysis generated a dataset to compare cell types and draw conclusions about molecular characteristics of the tumors.

Splenocytes were isolated from normal mice and mice with advanced lymphoma, then subjected to flow cytometry analysis. Forward *versus* side scatter profiles show a shift to larger cell size in lymphoma (Fig. 1*A*, *FSC*). Immunophenotype was consistent with a B-1 cell identity (Fig. 1*A*, *CD5*, *sIgM*), as previously reported (6). B cell surface markers also show that the lymphoma cells were activated (Fig. 1*A*, *panels B7-1* and *B7-2*). B cell surface markers defined two *in vitro* control conditions for interpretation of transcriptional signatures, independent of the microarray measurements themselves: "resting" cells from spleens of normal, non-Tg female (6–8 weeks) mice, mitogenically unstimulated (Fig. 1*B*, *open histograms*) or stimulated to become "activated" (*closed histograms*). Activated cells showed increased CD25, CD69, B7-1, and B7-2 expression, and decreased IgM and IgD (Fig. 1*B*), as expected (6). We later used total RNA from these B cells to define the transcriptional signatures "resting" and "activated," against which we compared the signatures of B cell lymphomas by microarray analysis, described below.

Pathology Groups—As expected in intermediate stages of splenic lymphoid malignancy, B cell Ig genes no longer showed polyclonal phenotypes, but an oligoclonal phenotype emerged, which was replaced by a monoclonal phenotype in many advanced cases (6). The features of histopathology, Ig gene clonality, and surface markers allowed us to define five pathology groups for microarray analysis (Table 1). Group 1 (P1) comprised normal resting B cells from asymptomatic Tg or non-Tg mice (6–8 weeks) without a premalignant phenotype. These samples were tentatively grouped together in P1, because histology, FACS, and Ig analysis indicated that there were no detectable differences between asymptomatic Tg and non-Tg B cells. This supposition was later borne out with genome-wide transcriptome analysis, wherein PCA, shown below, confirmed that asymptomatic Tg B cells had a very similar signature to non-Tg B cells. Group 2 (P2) comprised normal activated B cells from asymptomatic non-Tg mice. Group 3 (P3) comprised B cells of Tg mice that showed early clinical signs and spleen enlargement of 1.25–2.0 times normal weight, premalignant expansion of peritoneal B cells and B cell hyperplasia by flow cytometry and histology (Fig. 1 and Ref. 6). About 90% of the samples in P3 showed oligoclonal Ig genes. Group 4 (P4) comprised Tg mice with more severe clinical signs and pathology, including severe splenomegaly (2.0–10.0+ times normal weight), B cell lymphoma and disruption of splenic architecture (6) as well lymphadenopathy and B cell infiltration of bone marrow. B cells of P4 and P5 were frequently monoclonal for Ig genes. Group 5 (P5) included Tg mice with end-stage pathologies and always showed additional leukemic infiltrates in liver, lung, and kidney. Heart and brain were the only organs that did not show infiltrates, the appearance of which infiltrates coincided with peripheral B cell leukemia and severe anemia (Results not shown).

Ig Gene Sequences Discarded—To gain insight into the molecular mechanisms associated with lymphomagenesis in Tg mice, we compared the genome-wide gene expression patterns of the P1–P5 categories of B cells from Table 1. As we have previously reported for this Tg B cell lymphoma (6), the majority of gene expression changes reflected a decrease in mouse immune cell markers, particularly Igκ light chain sequences. Many of these markers were decreased between 50- and 100-fold, which we interpreted to be a hallmark of the emergence of clonality in the B cell lymphomas. Given this reason for differential expression of Ig genes, we removed probesets for detecting the expression of these genes from the dataset.

Principal Component Analysis—First, we sought to understand the general similarities and differences in gene expression between P1 and P5 by PCA of expression of all genes in the filtered dataset (Fig. 2*A*). The most important observation was that aggressive lymphomas (P5) have a pattern of gene expression more similar to activated B cells (P2) than resting B cells (P1). Intermediate-stage lymphomas (P3 and P4) have a pattern

[3] R Development Core Team (2005) *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3–900051-07-0.

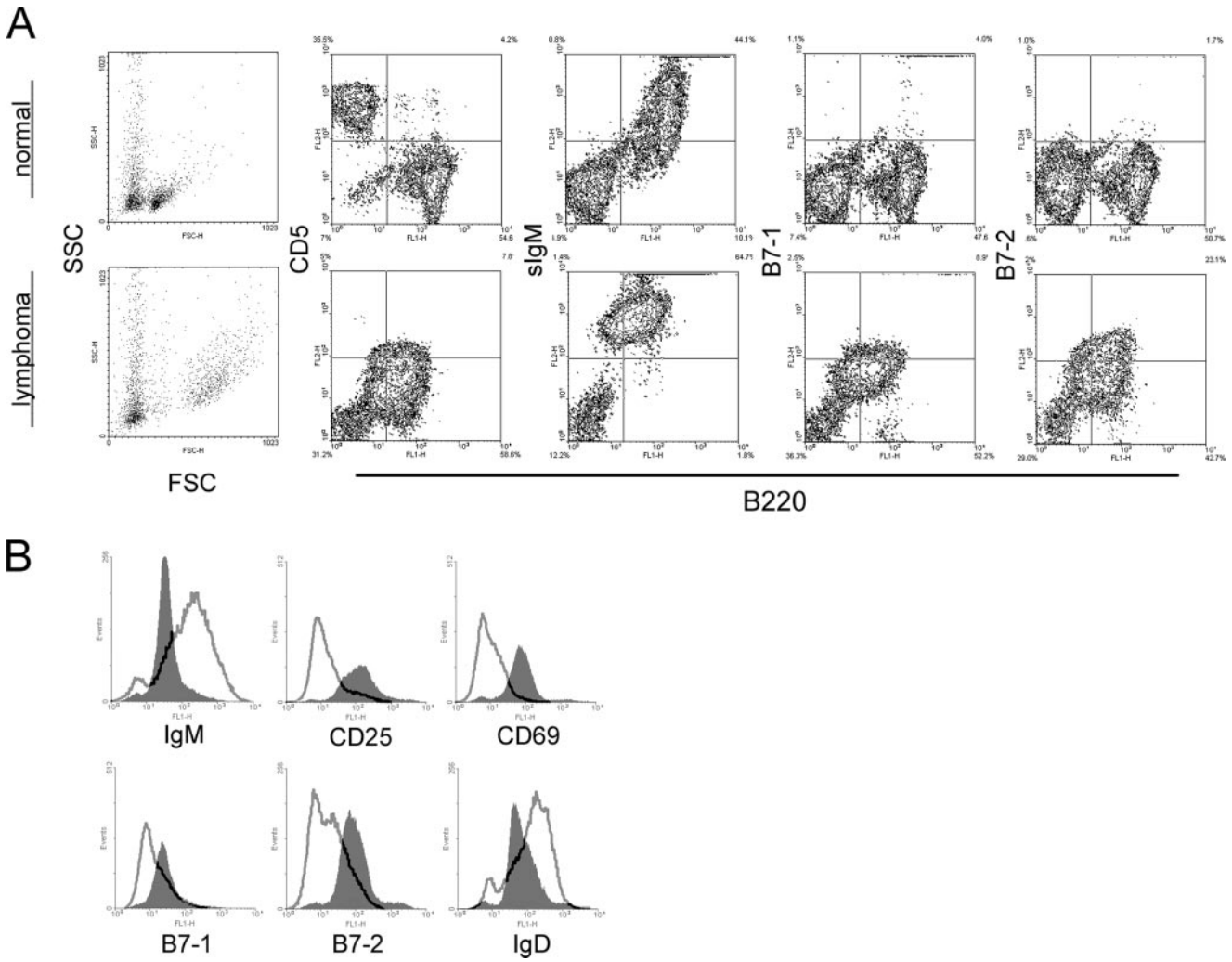

FIGURE 1. **Analysis of malignant B cells; normal, primary, resting, and activated B-2 cells.** *A*, flow cytometry of splenocytes from normal and lymphoma mice. Forward (*FSC*) *versus* side (*SSC*) scatter reveals larger lymphoid cells in spleens of lymphoma mice. Markers for B cell identity (*CD5, sIgM*) and activation (*B7-1, B7-2*), each coupled to PE (*y-axes; FL2 channel, logarithmic scale*), were used in two-color flow cytometry analyses against the pan B cell marker B220, which was coupled to FITC (*x-axes; FL1 channel, logarithmic scale*). *B*, purified normal splenic B-2 cells from non-Tg mice were stimulated *in vitro* with anti-IgM, anti-CD40, and interleukin-4 as described, then assayed for B cell activation markers (*CD25, CD69, B7-1, B7-2*). (*Open histogram*, resting cells; *closed histogram*, cells after 48 h stimulation). Isotype controls have been previously published (6).

TABLE 1

Summary of pathology groups

Based on clinical presentation, pathology, and flow cytometry, we grouped 26 samples into five major classes. *P1*, normal resting B cells; *P2*, normal activated B cells; *P3*, marginal cases; *P4*, transitional cases; *P5*, aggressive cases (*aggressive* 8, 9, 19, 20, 21, 22, 23; *transitional* 13, 14, 15, 16, 17; *marginal* 5, 6, 7, 10, 11, 12; *activated* 2, 25, 26; *resting* 1, 3, 4, 18, 24).

Group	*n*	Description	Splenomegaly	Other pathology	FACS	Ig gene clonality
P1	5	Resting	None	None	Normal resting	Polyclonal
P2	3	Activated	None	None	Normal activated	Polyclonal
P3	6	Marginal	+/−	+/−	Abnormal	Oligoclonal
P4	5	Transitional	+	+	Mixed	Oligoclonal
P5	7	Aggressive	+++	+++	B-1 clones	Monoclonal

of gene expression intermediate between resting B cells (P1) and aggressive lymphomas (P5), consistent with pathology. Second, to identify specific genes that are differentially expressed in aggressive lymphomas, we used the Student's *t* test to identify genes differentially expressed between P1 and P5. This analysis identified 3,342 significantly differentially expressed probesets, using a p value threshold of ~0.01 to maintain an overall FDR of 5%.

We then repeated the PCA using only these probesets to determine if the genes that are affected by lymphomagenesis are changed similarly in B cell activation (Fig. 2*B*). Two patterns might be expected: variation between P1, P2, and P5 could be *collinear*, implying that lymphomagenesis is essentially a mode of proliferation; or variation between P1 and P5 could be *orthogonal* to variation between P1 and P2, implying that genes affected by lymphomagenesis are tumor-specific. In fact, *both* patterns were observed: PCA of genes differentially expressed between resting B cells and aggressive lymphomas indicates that this group contains both tumor-specific and proliferation-specific genes, as indicated by the similar location of P5 and P2

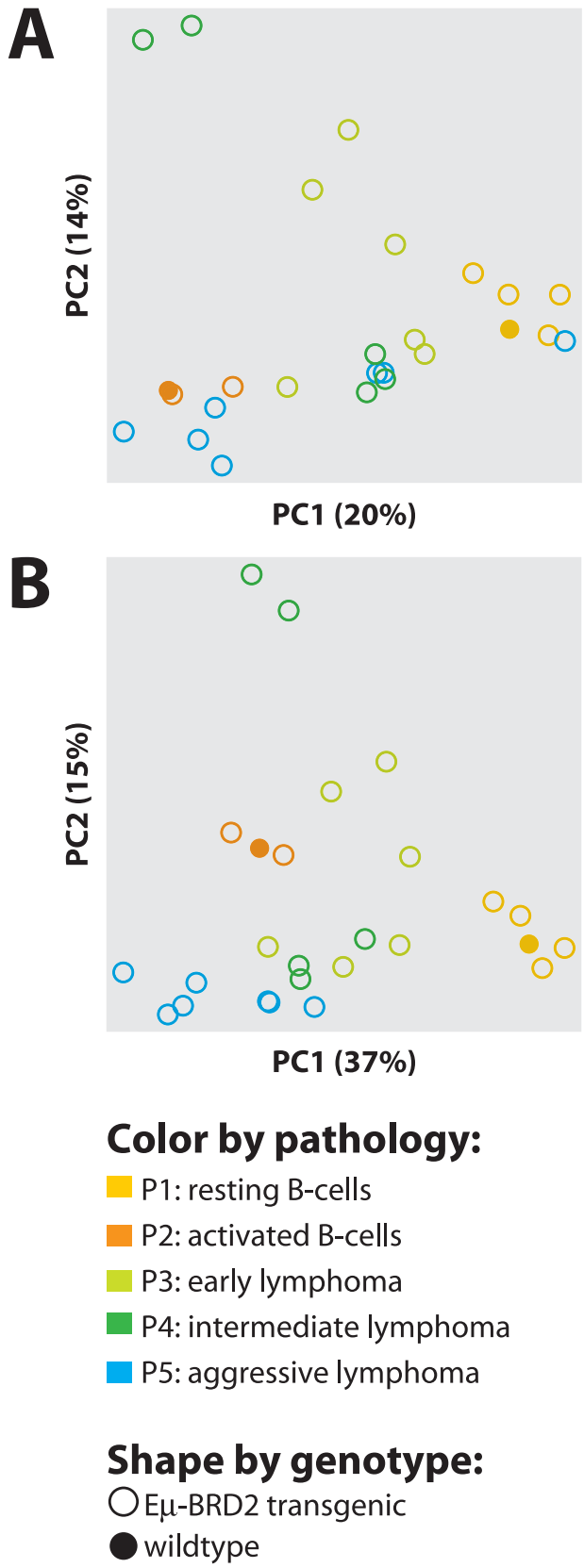

FIGURE 2. **Global differences in gene expression during Tg lymphomagenesis.** *A*, PCA of gene expression measurements from all genes with sequence-specific hybridization intensity in Tg and non-Tg mice, showing that the Tg does not dramatically perturb gene expression in resting (*P1*) and activated (*P2*) B cells; but that B cell activation and Tg-mediated lymphomagenesis result in a similar pattern of gene expression changes. *B*, PCA of gene expression measurements from 3,342 probesets that vary significantly between resting B cells (*P1*) and aggressive Tg lymphomas (*P5*) shows that the patterns of gene expression that are activated upon lymphomagenesis and upon B cell activation have both similarities and differences. The number in *parentheses* on each axis is the fraction of the total variation in gene expression that is accounted for by each principal component.

relative to P1 along the first principal component, as well as the distinct location of P5 relative to P2 along the second principal component. PCA also indicates that the intermediate expression phenotype of intermediate grade lymphomas (P3 and P4) is also preserved for this gene list.

Clusters of Gene Expression—Next, we used hierarchical clustering to identify distinct patterns of gene expression across the different B cell states for genes that vary between P1 and P5. We defined seven clusters (*C1–C7*; Fig. 3*A*). As expected from the similar pattern of gene expression seen in P2 and P5 (Fig. 2*B*), many clusters (C1, C3–C6) contain genes that are similarly differentially expressed in P2 and P5, relative to P1, and are thus "proliferation-specific" in this context (see supplemental Table S1 and supplemental File 1). C1 contains genes that are highly expressed in P2 and P5 relative to P1 and intermediate lymphomas. C1 should be considered in parallel with the down-regulated genes of C6, which share a roughly inverse pattern. C2 genes were highly expressed in P5, but not in P1 or P2; C2 therefore identifies genes that are "lymphoma specific" in this context. C7 genes that were expressed at lower levels in lymphoma and thus represent another pattern of lymphoma-specific gene expression; C7 is roughly the inverse of C2. The genes in clusters C2 and C7 are potential tumor-specific diagnostic markers and potential targets for therapeutic intervention.

C3 contains genes that are transcriptionally up-regulated in all pathology classes except P1. C3 should be considered an inverse of C4 and C5. C4 and C5 comprise the two largest groups of differentially expressed genes: genes that are down-regulated in all classes relative to P1, suggesting that C4 and C5 include genes that maintain quiescence. Consistent with this idea, several C5 genes involved in signal transduction are negative regulators of mitogenic signaling such as *Rhoh*, *Pten*, and *Crk*. For many clusters, lymphomas of intermediate pathology (P3 and P4) exhibit a level of gene expression that is intermediate between P1 and P5. Some probesets were expressed at levels very similar to P1 in intermediate lymphomas (*e.g.* probesets in C1 and C6), suggesting that differential expression of these genes occurs specifically in P2 and P5. The lymphoma samples with intermediate pathology exhibit a more heterogeneous pattern of gene expression, suggesting that lymphomagenesis might occur along several pathways that converge on a more uniform terminal phenotype. A complete list of all the genes in each cluster is supplied in supplemental File 1.

Clinical Relevance—We compared our results with three large-scale gene expression profiling studies of clinical lymphoma samples that have defined gene expression signatures for human lymphoma subtypes: Shipp *et al.* (3), Alizadeh *et al.* (5), and Monti *et al.* (32). We expected to find that differentially expressed genes in the Tg mouse model are orthologous to human genes known to be important in lymphomagenesis. First, for the Shipp *et al.* (3) dataset that compares activated B cell-like and GC cell-like DLCL, 40% of probesets (27/68) for

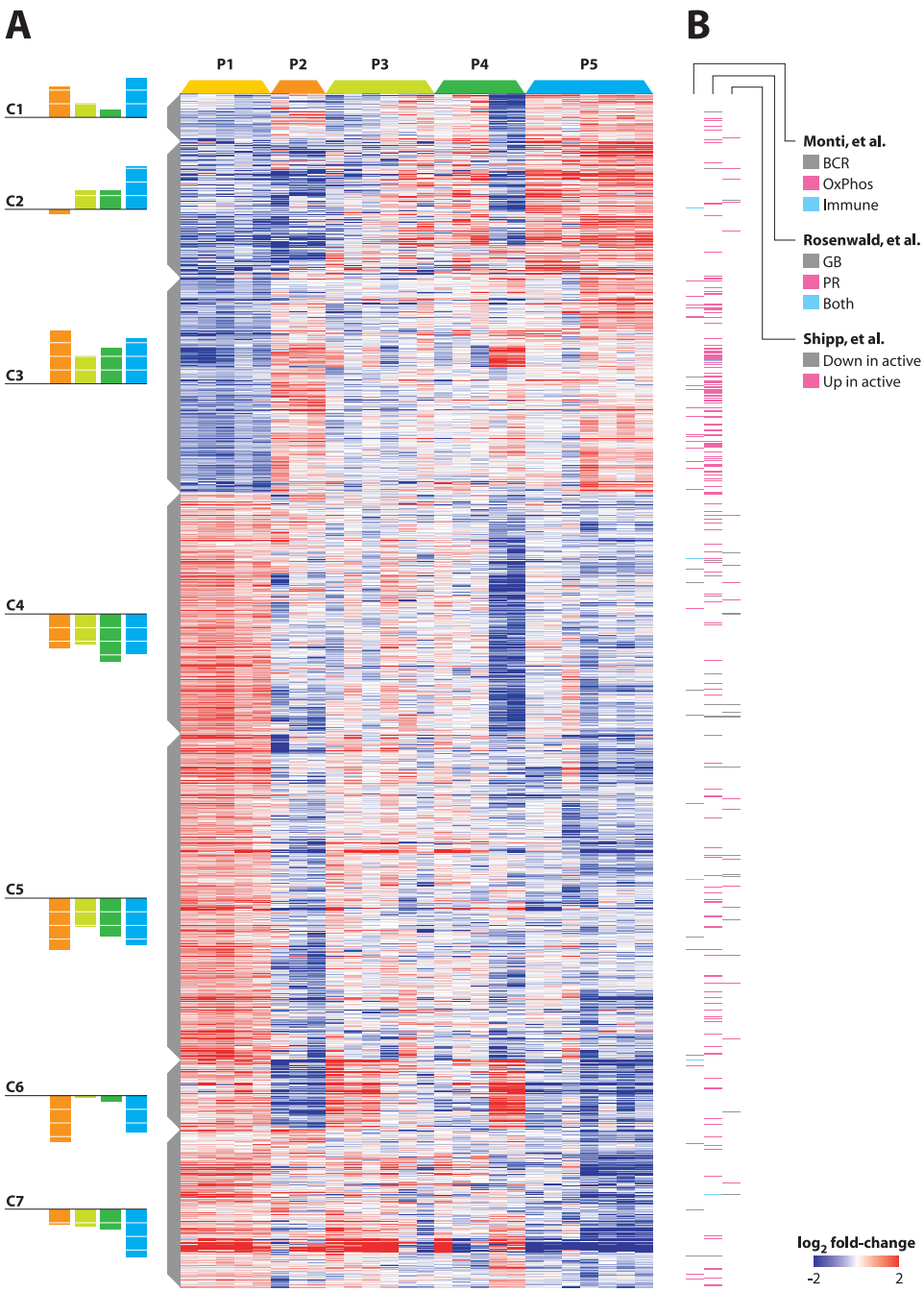

FIGURE 3. **Gene expression patterns.** *A*, distinct patterns of expression for genes that are differentially expressed in aggressive Tg lymphomas. The 3,342 probesets that vary in hybridization intensity between P1 and P5 are organized from *top* to *bottom* by hierarchical clustering. The seven groups of probesets with the most homogenous intragroup pattern of gene expression are labeled C1–C7. Fold-change relative to the mean hybridization intensity of each probeset is indicated on a continuous scale from *blue* to *red*. The *bar graph* to the *left* of each cluster represents the average fold-change for all of the genes in that cluster relative to resting B cells for each of the other four pathology groups. As expected from the PCA (*Fig. 2B*), the expression of the genes in most clusters is similar between P2 and P5. Two notable exceptions are C2 and C7 where the level of gene expression observed in P2 is more similar to P1 than P5. *B*, human orthologs of some genes that are differentially expressed in aggressive Tg lymphomas have been implicated in human lymphomagenesis. Mouse orthologs of genes that have been reported to distinguish different classes of human lymphoma were examined for overlap with 3,342 probesets that vary between resting B cells and aggressive Tg lymphomas. Genes that vary in aggressive Tg lymphomas and have been implicated in human lymphomagenesis are color-coded according to the human lymphoma subtype they distinguish. The highest density of genes that have been implicated in human lymphomagenesis among the genes that vary in the Tg lymphomas is in C3 and largely consists of genes that distinguish highly proliferative human lymphomas (roughly *OxPhos* in Monti *et al.* (32) and *Proliferative* in Alizadeh *et al.* (5)). Genes in the Monti *et al.* (32) *BCR* category and the Alizadeh *et al.* (5) *GB* category roughly distinguish germinal center (*GC*) B cell-like human lymphomas. The genes in the Monti *et al.* (32) category labeled *Immune* distinguish human lymphomas that have a Host Response phenotype. The Shipp *et al.* (3) categories refer to genes that are either expressed either at a lower level (*Down in active*) or higher level (*Up in active*) in highly proliferative human lymphomas.

mouse genes orthologous to human genes that distinguish these human lymphoma subtypes were differentially expressed in P5 relative to P1 (Fig. 3*B*). Second, for the Alizadeh *et al.* (5) dataset that compares normal B cells, DLCL, FL and chronic lymphocytic leukemia (CLL) cells, 28% of probesets (139/495) in their proliferation signature and 59% (17/29) in their GC signature were differentially expressed in P5 relative to P1 (Fig. 3*B*). Third, for the Monti *et al.* (32) dataset of DLCL cells, 22% of probesets (26/118) that distinguish Oxidative Phosphorylation (*OP*), B Cell Receptor/proliferation (*BCR*) and Host Response/immune (*HR*) subtypes were differentially expressed in P5 relative to P1 (Fig. 3*B*).

This analysis indicates many tumor-specific and proliferation-specific genes we identified have not yet been implicated in human lymphomagenesis. Given the preponderance of mouse-specific genes, we next sought to determine if Tg lymphomas are generally analogous to clinically relevant human lymphomas. We found that *BRD2* is expressed in the OP class of DLCL at approximately one-half the level observed in BCR and HR classes (32) (Fig. 4*A*; $p < 0.001$). This result suggests that *BRD2* function might be modulated in the OP class. Therefore, we tested genes that vary between different human DLCL classes for overlap with genes that vary between P1 and P5. Analysis of variance identified 4,120 probesets in the Monti *et al.* (32) dataset that vary between human DLCL classes with $p < 10^{-5}$. 827 probesets to orthologous murine genes are significantly differentially expressed between P1 and P5 (Fig. 4*B*). The majority of genes that are more highly expressed in the OP class of DLCL compared with BCR or HR classes tend to be more highly expressed in P5 compared with P1 (Fig. 4*C*). The converse is also true: genes expressed at lower levels in the OP class tend to be expressed at lower levels in P5 compared with P1. Both of these trends are statistically significant (χ^2 $p \ll 0.001$).

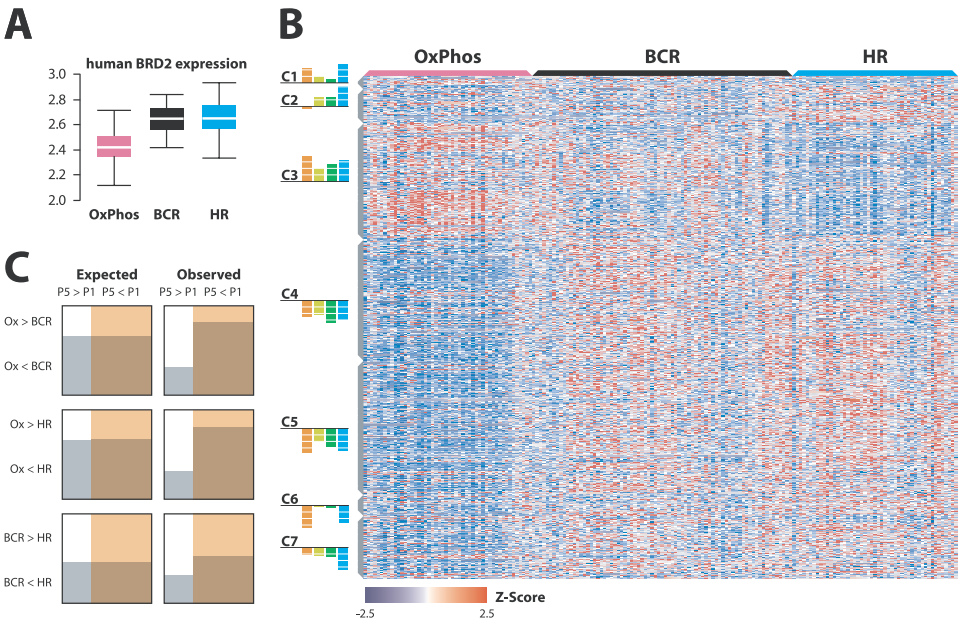

FIGURE 4. **Differential expression of Tg lymphomagenesis genes in human lymphomas.** A dataset from Monti *et al.* (32) with gene expression measurements from 176 clinical human lymphoma samples that had been classified into three different DLCL classes by gene expression pattern (32) was examined for differential expression of Tg lymphomagenesis genes. *A*, box plot of human $\log_{10}$ *BRD2* expression levels in OP, BCR, and HR lymphoma classes. *B*, expression levels in human lymphomas of 827 probesets that vary between human lymphoma subtypes and interrogate genes orthologous to those that vary in aggressive Tg lymphomas are organized from top to bottom by hierarchical clustering of the expression patterns observed in the Tg lymphomas (*Fig. 4A*). *C*, many more genes than would be expected by chance are either more highly expressed in the OP subtype and more highly expressed in the aggressive Tg lymphomas or expressed at lower levels in the OP subtype and expressed at lower levels in the aggressive Tg lymphomas.

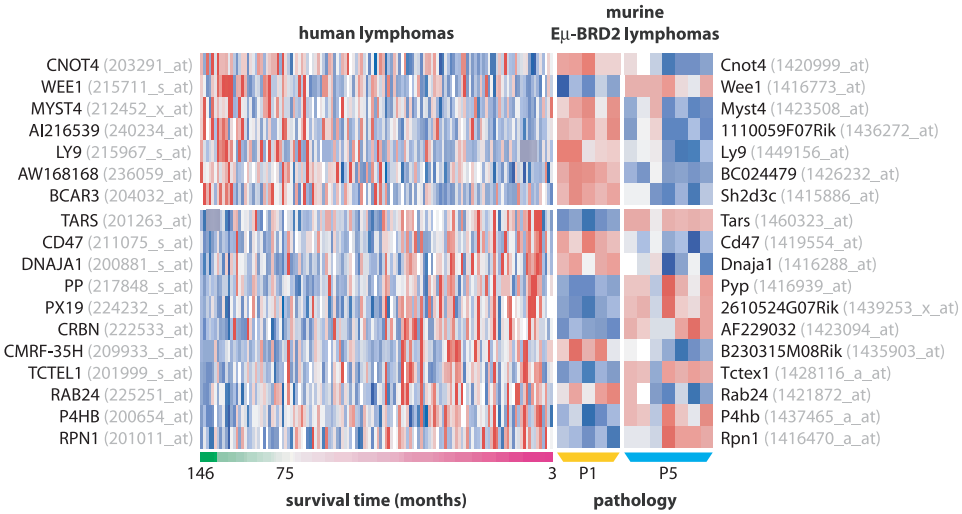

FIGURE 5. **Differential expression of Tg lymphomagenesis genes that correlate with human survival time.** We identified 120 probesets in the Monti *et al.* (32) dataset that were significantly positively or negatively correlated with lymphoma survival time and found that 20 of these interrogate the expression of genes that are orthologous to those that vary between resting B cells and aggressive Tg lymphomas. Human lymphoma expression is shown on the *left* with samples organized from *left to right* by survival time. Murine resting B cell (*P1*) and aggressive Tg lymphoma (*P5*) expression is shown on the *right*. The probesets are organized from *top to bottom* by hierarchical clustering of the human gene expression pattern.

These results suggest that the human lymphomagenesis pathway that gives rise to the OP class of DLCL has a pattern of gene expression that overlaps significantly with *BRD2*-driven murine lymphomagenesis.

Tg lymphomagenesis causes aggressive, transplantable disease with median survival time of 28 days (6); we therefore investigated whether Tg lymphomagenesis affects the expression of genes correlated with human DLCL survival time. We identified 120 probesets in the Monti *et al.* (32) dataset that were significantly positively or negatively correlated with patient survival time (Pearson correlation; $p < 0.005$; corresponding to FDR < 0.10). Of these, 20 probesets were significantly differentially expressed between P1 and P5 (Fig. 5). Genes that are negatively correlated with survival tend to be more highly expressed in P5 compared with P1. The converse is also true: genes positively correlated with survival tend to be expressed at lower levels in P5 compared with P1. This trend is borderline statistically significant (Fisher-Exact $p = 0.066$). These parallels in survival-time gene expression may be independent of the similarities between OP and Tg lymphomas, because there is no apparent difference in the mean survival time of OP class DLCL patients (results not shown). These results suggest that similar molecular mechanisms may determine DLCL survival time in both humans and mice, and also that *BRD2*-mediated murine lymphomas may be an appropriate model system for studying factors that influence human DLCL outcomes.

Lastly, we focused our attention on tumor specific genes that are of known importance in lymphoid biology. Targeted therapies for lymphoma are most likely to be successful if they are directed at genes that are differentially expressed along the tumor specific axis described above, orthogonal to the proliferation specific axis. The proliferation signature genes (supplemental Table S1, Clusters C3 and C6) are therefore of interest to studies of proliferation, but the tumor signature genes (supplemental Table S1, Clusters C2 and C7) are of deeper interest for the development of therapeutic targets, which we intend next to pursue. The most strongly differentially expressed genes of C2 and C7 that have important functions in lymphoid biology are shown in Fig. 6. Many of the genes that are induced specifically in B cell malignancies are not surprising, such as *Il5ra* (interleukin 5 receptor α, which is involved in Socs4 activation), *Slamf9* (SLAM family member 9, a *CD2* family member and co-receptor for lympho-

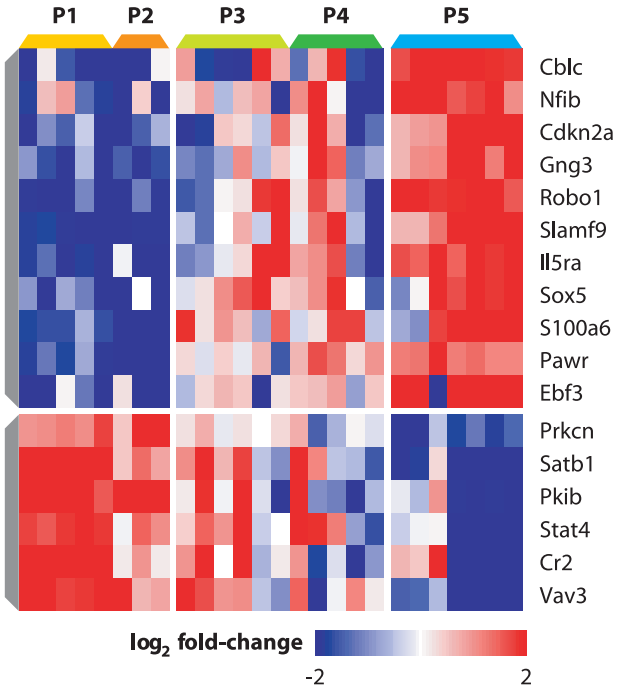

FIGURE 6. **Tumor-specific genes of particular importance in lymphocyte biology.** Genes of particular importance in lymphocyte biology that are up-regulated in aggressive lymphomas (*P5*) relative to both resting and activated normal B cells (*P1, P2*) from cluster C2 are identified in the upper block. Genes that are downregulated in aggressive lymphomas relative to both resting and activated normal B cells from cluster C7 are identified in the lower block. Samples from lymphomas with intermediate pathology (*P3, P4*) showed intermediate levels of expression for these genes.

cyte activation), *Gng3* (the γ3 subunit of a G protein), and *Cdkn2a* (cyclin-dependent kinase inhibitor *2A*; $p16^{INK4a}$, which shows increased expression in certain tumors). C7 genes included *Prkcn*, *Satb1*, *Pkib*, *Stat4*, *Vav3*, and *Cr2* (*Cd21*, also called complement receptor 2).

DISCUSSION

The cellular phenotype of Tg lymphoma is more closely related to the activated B cell DLCL phenotype of NHL than to the GC or FL phenotype, consistent with the aggressive course of the malignancy (6). Dual bromodomain proteins such as Brd2 contribute to cyclin A transcriptional control; Brd2 multiprotein complexes contain E2F proteins (35), histone modification enzymes (36) and Swi/Snf-dependent nucleosome remodeling machinery (37). Apart from cyclin A transactivation, constitutive expression of a double bromodomain protein such as Brd2 might conceivably affect many loci, given the reported binding of Brd2 bromodomains to acetylated histone H4 (38). Furthermore, mutations in chromatin-related transcriptional control proteins often have broad epigenetic effects; *swi/snf* mutations in yeast transcriptionally repress many genes but activate many others (34).

We determined that the biology of Tg lymphomagenesis bears some gene expression features of mitogenically activated, primary splenic B-2 cells, as well as changes in gene expression that are tumor-specific. We found that a number of genes previously reported to be involved in human lymphomagenesis are differentially expressed in Tg lymphomas, and that these genes were primarily those that we found to be differentially expressed in both activated B cells and Tg lymphomas (Cluster C3 in Fig. 2*A*).

To explore other parallels between Tg lymphomagenesis and human disease, we examined the expression of *BRD2* in a dataset of 176 DLCL samples that had been categorized into three classes by genome-wide expression profiling (32) and found that *BRD2* is expressed at ~2-fold lower levels in a human DLCL class (termed *OP*) that is characterized by the increased expression of genes involved in oxidative phosphorylation, mitochondrial function, and negative regulators of apoptosis. Gene expression differences in both Tg lymphomas and the OP class of DLCL are similar suggesting that *BRD2* constitutive expression in mice triggers a pathway that at least partially mimics human OP-like lymphomagenesis.

It is surprising that the gene expression patterns of OP-DLCL and Tg lymphomas are similar given the decreased expression of *BRD2* in OP relative to other classes of human DLCL. There are several possible scenarios to account for these results. 1) Constitutive expression of *BRD2* in the Tg model may cause a dominant negative phenotype. 2) Because there is no evidence that the OP phenotype is directly caused by lowered expression of *BRD2*, an OP-modulated regulator of lymphomagenesis and constitutive expression of *BRD2* might similarly activate an OP/*BRD2*-responsive pathway that results in down-regulation of *BRD2* expression. 3) Both constitutive- and underexpression of *BRD2* might cause a similar phenotype. 4) A defect in the cell cycle regulation of *BRD2* in OP tumors could cause an OP/*BRD2* lymphoma phenotype and also overall lower steady-state levels of *BRD2* transcript. Although we are unaware of measurements of *BRD2* expression in the Alizadeh *et al.* dataset (5), examination of the Shipp *et al.* dataset (3) revealed that *BRD2* expression is slightly (but significantly) lower in DLCL compared with FL (analysis not shown). This decreased expression of *BRD2* in DLCL is particularly intriguing because the Tg lymphomas have a phenotype that is more DLCL- than FL-like and this pattern is similar to the pattern observed for the OP class of DLCL.

BRD2-mediated lymphomagenesis results in the increased expression of genes associated with poor prognosis and the decreased expression of genes associated with long-term survival in humans, which is consistent with the aggressive nature of the Tg tumors. Experiments that modulate the expression of these genes in combination with E*μ-BRD2* may help to determine if any of these genes are critical regulators of lymphoma severity or survival. Our analysis also identifies clusters of genes (supplemental Table S1) that may be biomarkers of aggressive NHL, which have functions throughout signal transduction, including cytokine function, kinase cascades, and transcription factors, and several poorly characterized proteins with *Drosophila* homologs.

It is likely that the activated B cell expression signature is partially dependent on the mitogenic stimulus. For example, our choice of anti-IgM, anti-CD40 and interleukin-4 as mitogens is likely to create a different transcriptional outcome than polyclonal mitogenic activation with lipopolysaccharide, or endogenous B cell receptor engagement through the immunological synapse in the spleen. Gene expression profiles of normal and Tg B cells with different mitogenic stimuli may there-

fore permit further insight into the mechanisms by which lymphoma cells become highly proliferative. Likewise, the B-1 identity of the aggressive Tg tumors (6) may subsume gene expression patterns that depart intrinsically from the identity of normal, activated B-2 cells.

Given the genetically well controlled nature of this Tg model for DLCL, we are able to conclude that lymphomagenesis involves changes in the expression of genes that are important for proliferation, such as cell cycle genes, and a different set of genes that are specific to malignancy. The tumor-specific class is most likely to include useful therapeutic targets and should be examined further within the context of human lymphomagenesis.

Acknowledgments—We thank J. Tumang and T. Rothstein for flow cytometry reagents and discussion; N. Gerry and G. Frampton for assistance with microarrays; P. Sebastiani, B. Nikolajczyk; and members of the Cancer Research Center at Boston University School of Medicine for helpful criticism.

REFERENCES

1. Edwards, B. K., Brown, M. L., Wingo, P. A., Howe, H. L., Ward, E., Ries, L. A. G., Schrag, D., Jamison, P. M., Jemal, A., Wu, X. C., Friedman, C., Harlan, L., Warren, J., Anderson, R. N., and Pickle, L. W. (2005) *J. Nat. Cancer Inst.* **97,** 1407–1427
2. Hennessy, B. T., Hanrahan, E. O., and Daly, P. A. (2004) *Lancet Oncol.* **5,** 341–353
3. Shipp, M. A., Ross, K. N., Tamayo, P., Weng, A. P., Kutok, J. L., Aguiar, R. C., Gaasenbeek, M., Angelo, M., Reich, M., Pinkus, G. S., Ray, T. S., Koval, M. A., Last, K. W., Norton, A., Lister, T. A., Mesirov, J., Neuberg, D. S., Lander, E. S., Aster, J. C., and Golub, T. R. (2002) *Nat. Med.* **8,** 68–74
4. Godwin, J. E., and Fisher, R. I. (2001) *Clin. Lymphoma* **2,** 155–163
5. 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. (2000) *Nature* **403,** 503–511
6. Greenwald, R. J., Tumang, J. R., Sinha, A., Currier, N., Cardiff, R. D., Rothstein, T. L., Faller, D. V., and Denis, G. V. (2004) *Blood* **103,** 1475–1484
7. Beck, S., Hanson, I., Kelly, A., Pappin, D. J. C., and Trowsdale, J. (1992) *DNA Seq.* **2,** 203–210
8. Wang, E. H., Zou, S., and Tjian, R. (1997) *Genes Dev.* **11,** 2658–2669
9. Adams, J. M., Harris, A. W., Pinkert, C. A., Corcoran, L. M., Alexander, W. S., Cory, S., Palmiter, R. D., and Brinster, R. L. (1985) *Nature* **318,** 533–538
10. Harris, A. W., Pinkert, C. A., Crawford, M., Langdon, W. Y., Brinster, R. L., and Adams, J. M. (1988) *J. Exp. Med.* **167,** 353–571
11. Ting, J. P., and Trowsdale, J. (2002) *Cell* **109,** Suppl. S21–S33
12. Denis, G. V., and Green, M. R. (1996) *Genes Dev.* **10,** 261–271
13. Forquignon, F. (1981) *Wilhelm Roux's Arch. Dev. Biol.* **190,** 132–138
14. Gans, M., Forquignon, F., and Masson, M. (1980) *Genetics* **96,** 887–902
15. Mozer, B. A., and Dawid, I. B. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 3738–3742
16. Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C. M., and Canaani, E. (1992) *Cell* **71,** 701–708
17. Tkachuk, D. C., Kohler, S., and Cleary, M. L. (1992) *Cell* **71,** 691–700
18. Denis, G. V. (2001) *Front. Biosci.* **6,** D1065–1068
19. Horn, P. J., and Peterson, C. L. (2001) *Front. Biosci.* **6,** D1019–1023
20. Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. -M. (1999) *Nature* **399,** 491–496
21. Florence, B., and Faller, D. V. (2001) *Front. Biosci.* **6,** D1008–D1018
22. Dey, A., Ellenberg, J., Farina, A., Coleman, A. E., Maruyama, T., Sciortino, S., Lippincott-Schwartz, J., and Ozato, K. (2000) *Mol. Cell. Biol.* **20,** 6537–6549
23. French, C. A., Miyoshi, I., Kubonishi, I., Grier, H. E., Perez-Atayde, A. R., and Fletcher, J. A. (2003) *Cancer Res.* **63,** 304–307
24. Sobulo, O. M., Borrow, J., Tomek, R., Reshmi, S., Harden, A., Schlegelberger, B., Housman, D., Doggett, N. A., Rowley, J. D., and Zeleznik-Le, N. J. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 8732–8737
25. Lavau, C., Du, C., Thirman, M., and Zeleznik-Le, N. (2000) *EMBO J.* **19,** 4655–4664
26. Liedman, D., and Zeleznik-Le, N. (2001) *Curr. Opin. Hematol.* **8,** 218–223
27. Borrow, J., Stanton, V. P., Jr., and Andresen, J. M. (1996) *Nat. Genet.* **14,** 33–41
28. Ida, K., Kitabayashi, I., Taki, T., Taniwaki, M., Noro, K., Yamamoto, M., Ohki, M., and Hayashi, Y. (1997) *Blood* **90,** 4699–4704
29. Barrett, T., Suzek, T. O., Troup, D. B., Wilhite, S. E., Ngau, W.-C., Ledoux, P., Rudnev, D., Lash, A. E., Fujibuchi, W., and Edgar, R. (2005) *Nucleic Acids Res.* **33,** D562–D566
30. Benjamini, Y., and Hochberg, Y. (1995) *J. Roy. Statist. Soc. Ser. B* **57,** 289–300
31. Ward, J. H., Jr. (1963) *J. Am. Stat. Assoc.* **38,** 236–244
32. Monti, S., Savage, K. J., Kutok, J. L., Feuerhake, F., Kurtin, P., Mihm, M., Wu, B., Pasqualucci, L., Neuberg, D., Aguiar, R. C., Dal Cin, P., Ladd, C., Pinkus, G. S., Salles, G., Harris, N. L., Dalla-Favera, R., Habermann, T. M., Aster, J. C., Golub, T. R., and Shipp, M. A. (2005) *Blood* **105,** 1851–1861
33. Liu, G., Loraine, A. E., Shigeta, R., Cline, M., Cheng, J., Valmeekam, V., Sun, S., Kulp, D., and Siani-Rose, M. A. (2003) *Nucleic Acids Res.* **31,** 82–86
34. Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. (1998) *Cell* **95,** 717–728
35. Denis, G. V., Vaziri, C., Guo, N., and Faller, D. V. (2000) *Cell Growth Diff.* **11,** 417–424
36. Sinha, A., Faller, D. V., and Denis, G. V. (2005) *Biochem. J.* **387,** 257–269
37. Denis, G. V., McComb, M. E., Faller, D. V., Sinha, A., Romesser, P. B., and Costello, C. E. (2006) *J. Proteome Res.* **5,** 502–511
38. Kanno, T., Kano, Y., Siegel, R. M., Jang, M. K., Lenardo, M. J., and Ozato, K. (2004) *Cell* **13,** 33–43