[CANCER RESEARCH 64, 4419–4427, July 1, 2004]

High-Throughput Retroviral Tagging for Identification of Genes Involved in Initiation and Progression of Mouse Splenic Marginal Zone Lymphomas

Min Sun Shin,[1] Torgny N. Fredrickson,[1] Janet W. Hartley,[1] Takeshi Suzuki,[2] Keiko Agaki,[2] and Herbert C. Morse III[1]

[1]Laboratory of Immunopathology, National Institute of Allergy and Infectious Diseases, Rockville, Maryland and [2]Mouse Cancer Genetics Program, National Cancer Institute, Frederick, Maryland

ABSTRACT

Human B-cell lymphomas are frequently associated with specific genetic changes caused by chromosomal translocations that activate proto-oncogenes. For lymphomas of mice expressing murine leukemia virus, mutagenic proviral insertions are thought to play a similar role. Here we report studies designed to determine whether specific retroviral integration sites might be associated with a specific subset of mouse B-cell lymphomas and if the genes associated with these sites are regularly altered in expression. We studied splenic marginal zone lymphomas (MZL) of NFS.V$^+$ mice that are unusual in exhibiting frequent progression from low to high grade, potentially allowing assignment of cancer genes to processes of initiation and progression. We used inverse PCR to clone and analyze 212 retroviral integration sites from 43 MZL at different stages of progression. Sixty-two marked common integration sites and included 31 that had been marked previously. Among the new common integration sites, seven were unique to MZL. Using microarrays and real-time quantitative PCR analysis, we defined differential patterns of gene expression in association with disease progression for *Gfi1*, *Sox4*, *Brca2*, *Snf1lk*, *Nfkb1*, *Pou2af1*, *Prdm1*, *Stat6*, and *Blnk*. Heightened expression of *Gfi1* distinguishes MZL from other lymphoma types. The combined use of proviral tagging and analyses of gene expression thus provides a powerful approach to understanding of genes that collaborate in tumorigenesis.

INTRODUCTION

The identification of genes that cause cancer is crucial for elucidation of the molecular mechanisms involved in tumorigenesis. Many retroviruses that cause tumors act by a mechanism of proviral insertional activation of cellular oncogenes or less frequently, by inactivation of tumor suppressor genes (1). Retroviral insertional mutagenesis in BXH2 and AKXD recombinant inbred (RI) mice results in a high incidence of myeloid leukemia and B- or T-cell lymphoma, respectively (2, 3), and the retroviral integration sites (RIS) in these tumors have provided powerful genetic tags for disease gene identification (4, 5). Some of the genes identified by proviral tagging are also associated with human neoplasms (6–9), validating this approach for human disease gene identification.

Although many lymphoma disease genes have been identified by proviral tagging in spontaneous mouse lymphomas (3–6), there has been no concerted effort to relate specific RIS to different diagnostic categories of mouse B-cell lineage lymphomas or to distinguish mutations involved in initiation from those contributing to progression. In this study, we sought to determine whether identification of RIS in mouse splenic marginal zone lymphoma (MZL) could lead to an understanding of the molecular pathogenesis of this disease.

Splenic MZL occurs at a high frequency in NFS/N mice congenic for high-expressing ecotropic murine leukemia virus (MuLV) genes from AKR and C58 mice (NFS.V$^+$; 10, 11) but also occurs in other mice including AKXD RI strains (12) and p53 knockout mice (13). MZL is a clonal disease that develops from B cells of the marginal zone (MZ), a histologically, phenotypically, and functionally distinct population (14), which lies just outside the splenic white pulp. The MZ in most strains of mice, including NFS.V$^+$, is only about 3-cells thick, and MZ B cells represent $\leq$5% of the splenic B-cell population or $\sim$2% of all spleen cells.

A remarkable feature of MZL is the progression from low- to high-grade disease (10, 15). The earliest manifestation of MZL is an expansion of the MZ by $\geq$5-fold. Cytologically, the cells of this extended B-cell subset do not differ from the normal population, and they remain confined to the MZ. Southern blot hybridization analyses of immunoglobulin J_H gene organization in spleens of mice at this stage of disease, designated early MZL (MZL-E), indicate the presence of clonal expansions and sometimes more than one. Densities of clonal bands are consistent with cellular contributions of $\leq$10% to the total number of spleen cells, and the disease is judged to be early lymphoma at this stage.

The transition between MZL-E and tumor cells presumed to be fully transformed is marked by both histological and molecular changes. The first stage of this change, designated MZL, is characterized by cells with increased cytoplasm (size) and nuclei with more open chromatin and obvious nucleoli that are still confined to the MZ (10, 15). Molecular studies of clonality indicate that the MZL tumor population now compromises on the order of 15–30% of spleen cells.

MZL with signs of further progression, designated MZL+, are characterized by a proliferative expansion that results in invasion of the red pulp and, sometimes, early compression of the white pulp. J_H status by Southern blots is indicative of a fully clonal population comprising $\geq\sim$40% of all spleen cells. Mitoses and apoptotic figures seen rarely in MZL are more common at this stage of disease (10).

The evidence for high-grade transition of MZL+ to MZL++ is, again, both morphological/histological and molecular. Cells with increased size, larger nuclei with open chromatin, and prominent nucleoli now occupy not only the MZ but extensively infiltrate the red pulp and compress the white pulp, leaving only a ragged periarteriolar lymphoid sheath around the central arteriole. The cytological features are basically indistinguishable from those of centroblastic lymphomas of follicular B-cell origin (10, 15). It is only the residual affinity for their origins in the MZ and the remnants of white pulp that allow them to be distinguished from diffuse large B-cell lymphomas of centroblastic type that originate within the follicles. By Southern analyses, the transformed cells compromise $\geq$50% of spleen cells.

Two fundamental sets of observations permit MZL-E in NFS.V$^+$ mice to be united in a progression to MZL++. The first is the occurrence in cohorts of mice of lymphomas that provide "bridges," through histological, cytologic, and molecular features, for progression of one stage into the next. This could be considered circumstantial because each putative stage of disease could be envisioned to have distinct cellular and molecular origins. A second approach involves splenic biopsy followed by examination of the same mouse at autopsy.

Received 12/11/03; revised 4/5/04; accepted 4/22/04.

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.

Note: Supplemental data for this article can be found at Cancer Research Online (http://cancerres@aacrjournals.org).

Requests for reprints: Herbert C. Morse III, Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, NIH, 5640 Fishers Lane, Rockville, MD 20852. Phone: (301) 496-6379; Fax: (301) 402-0077; E-mail: hmorse@niaid.nih.gov.

These studies showed that a low-grade tumor, defined at biopsy, can progress to high-grade disease in the same mouse at autopsy with both being demonstrably from the same molecular clone (10). It is noteworthy that whereas all grades of progression were observed in NFS.V$^+$ mice, only-low grade MZL was seen in AKXD RI strains.

Recent studies have demonstrated the power of retroviral tagging in myeloid and lymphoid tumors in an era in which the mouse genome sequence is now known in exquisite detail (4, 5, 16–19). The sequence offers a powerful means for rapid identification of novel oncogenes that flank common integration sites (CIS) of MuLV. To develop an understanding of genes involved in splenic MZL, we applied the inverse PCR method to clone RIS coupled with a comparison to the University of California at Santa Cruz genome sequences to identify RIS in 43-splenic MZL arising in NFS.V$^+$ mice with low numbers of endogenous ecotropic MuLV. Moreover, to determine whether the virus integration alters the expression of surrounding genes thereby contributing to the initiation or progression of splenic MZL, we examined the expression patterns for all genes identified in this study using DNA microarrays and real-time quantitative PCR (qPCR) analysis.

MATERIALS AND METHODS

Mice and Lymphomas. NFS.V$^+$ mice congenic for ecotropic MuLV induction loci from strains AKR and C58 and that develop a high incidence of B-cell lymphomas of different histological types have been described previously (11). Samples obtained at autopsy were frozen for later preparation of DNA and RNA and for histopathology. Histological criteria used for diagnosing splenic MZL at various stages of progression and other classes of mouse B-cell lineage lymphomas are detailed elsewhere (10, 15, 20). All diagnoses were made by a single pathologist to avoid inter-observer variability in classification. C57BL/6 nude mice were purchased from the Jackson Laboratory (Bar Harbor, ME) and maintained in a specific pathogen-free environment until they were killed. Spleens were frozen for later preparation of RNA. Highly enriched (>95% pure) splenic B cells were prepared by negative selection and used for preparation of RNA.

Inverse PCR Cloning. Inverse PCR was performed as described previously (4) with only slight modification. Briefly, 5 μg of tumor DNA was digested to completion with *Sac*II and *Bam*HI (New England BioLabs, Beverly, MA). Digested DNA was then self-circularized by dilution and ligation using T4 DNA ligase (3200 units; New England BioLabs, Beverly, MA) in a total volume of 600 μl at 16°C for 16 h. Circular DNA was precipitated with ethanol and dissolved in 40 μl of 10 mM Tris (pH 8)-1 mM EDTA buffer. Two μl were used in the primary PCR in a 50-μl PCR reaction volume containing 20 nmol each of deoxynucleoside triphosphate, 10 pmol each forward and reverse primer, 1× buffer 2, and 2.5 units enzyme mix in the Expand Long Template PCR System (Roche, Indianapolis, IN). We used a Perkin-Elmer thermocycler programmed at 94°C for 2 min, followed by 30 cycles of 94°C for 25 s, 60°C for 60 s, 68°C for 15 min, and a final extension step at 68°C for 30 min. Using electrophoresis on a 1% agarose gel, we quantified the primary PCR product and used 0.1–1.0 μl primary PCR product as the template in the secondary PCR reaction. The secondary PCR used the same conditions as the primary, except that secondary amplication primers were used. We separated the secondary PCR product on a 1% agarose gel, purified it using the QIAquick Gel Extraction kit (Qiagen, Valencia, CA), and directly cloned it using the Topo TA Cloning kit (Invitrogen, Carlsbad, CA). The primer sequences for inverse PCR have been described previously (4).

DNA Sequencing. Clones were sequenced using the Thermo Sequenase Cycle Sequencing kit (USB, Cleveland, OH) and LI-COR Global Edition IR^2 System Sequencer (LI-COR, Lincoln, NE) with the M13 forward and reverse primers (LI-COR).

Sequence Comparison. We compared the RIS sequences against the publically available mouse genome database October 2003 draft assembly[3] and identified annotated candidate genes located near each RIS. We compared the RIS with previously identified RIS in the web-accessible Mouse Retroviral Tagged Cancer Gene Database.[4] To define common integration sites, we have adopted the statistical approach used in the publication by Suzuki *et al.* (18). In this algorithm, two or more integrations falling within a 30-kb window are considered as a common site. In the data presented, we identify the gene most proximal to that site, even if it is much >30-kb away from the common site.

Oligonucleotide Microarrays. Microarray chips, printed by Microarray Research Facility, National Institute of Allergy and Infectious Diseases,[5] comprised approximately 16,000 mouse gene targets represented by 70mer oligonucleotides purchased from Compugen (Jamesburg, NJ). These oligonucleotides were arrayed on poly-L lysine-coated slides, stored in a dry dark area, and used within 60 days. Total RNA was isolated from tumor using TRIzol Reagent (Invitrogen). Samples were processed using a modified version of the National Human Genome Research Institute RNA extraction protocol.[6] A reference sample was created by extracting total RNA from ten well characterized hematopoetic cell lines and making a pool comprising equal amounts of RNA from each line. Tumor cell lines included T cell, macrophage, pre-B cell, mature B cell, and plasmacytoma. RNA was labeled using a modified version of a National Human Genome Research Institute-labeling protocol by direct incorporation of dUTP nucleotides labeled with fluorescent dyes (Cy3 and Cy5) during cDNA generation using reverse transcriptase. Tumor RNA was labeled with Cy3, and reference RNA was labeled with Cy5 to form a sample pair. Members of a sample pair were combined and applied to the prehybridized microarray chip. Chips were prehybridized in a sealed chamber immersed in a 42°C waterbath using established procedures.[5] Chips were then washed twice for 2 min with agitation, first in ultra-pure water, then in isopropyl alcohol. After the last isopropyl alcohol wash, samples were quickly placed in a room-temperature centrifuge and spun dry (3000 rmp, 5 min). The chips were hybridized overnight in a sealed humidified slide chamber immersed in a light-protected 42°C waterbath, then washed, dried, and stored in the dark until scanned.

Chips were scanned using an Axon 4000B scanner and GenePix 4.0 software. Photomultiplier tube voltage settings were adjusted to allow 1% saturation of signal in both channels. Image data were extracted, and extracted data were analyzed using National Institute of Allergy and Infectious Diseases mAdb data storage and analysis programs.[5]

Real-Time qPCR. Total RNA was isolated from tumor cells using TRIzol Reagent (Invitrogen). Reverse transcription was performed using 1 μg of RNA, random hexamer primer (Invitrogen), and Superscript II (Invitrogen). The primers for real-time PCR were designed using the Primer Express software (Applied Biosystems, Foster City, CA) and synthesized at MWG-Biotec (High Point, NC; Supplemental Table 1). Each real-time PCR reaction was performed in a mix of 10-μl reaction mixture containing 50 ng of cDNA, 2× SYBR Green PCR Master Mix (Applied Biosystems), and 3 μM of each primer. The reaction mixture was denatured for 10 min at 94°C and incubated for 40 cycles (denaturing for 15 s at 95°C and annealing and extending for 1 min at 60°C) using ABI PRISM 7900HT (Applied Biosystems). All samples were tested in triplicate, and average values were used for quantification. Analysis was performed using SDS v2.1 software (Applied Biosystems) according to the manufacturer's instruction. *Gapdh* was used as an endogenous reference. The comparative C_T method ($\Delta\Delta C_T$) was used for quantification of gene expression.

RESULTS

RIS Identified by Inverse PCR. We isolated, cloned, and sequenced a total of 243 RIS from 43 splenic MZL (Supplemental Table 2). These cases comprised 3 of MZL-E, 20 of MZL, 12 of MZL+, and 8 of MZL++. We then carried out homology searches of the mouse genome database[3] and identified 212 unambiguous RIS. This corresponded to an average of approximately four integrations per tumor.

Several viral integrations were found to target the same locus in independent tumors, thus defining a CIS and inferring the presence of a cancer-associated gene (1, 19). We compared the 212 integrations

[3] http://genome.ucsc.edu/.

[4] http://genome2.ncifcrf.gov/RTCGD.

[5] http://madb.niaid.nih.gov.

[6] http://research.nhgri.gov/microarray.

Table 1 *CIS*[a] *in mouse splenic MZL*

Candidate gene[b]	Protein	No. hits[c]	RIS	Mouse chromosome[d]	Human chromosome	Human gene
New CIS (splenic MZL specific)						
Stk10	Serine/threonine kinase 10	4		11 A4	5q35.1	*STK10*
Abcc5	ATP-binding cassette, sub-family C (CFTR/MRP), member 5	3		16 A3	3q27	*ABCC5*
Alas1	Aminolevulinic acid synthase 1	2		9 F1	3p21.1	*ALAS1*
Brca2	Breast cancer 2	2		5 G3	13q12.3	*BRCA2*
Grb2	Growth factor receptor-bound protein 2	2		11 E2	17q24-q25	*GRB2*
Ube2v1	Ubiquitin-conjugating enzyme E2 variant 1	2		2	20q13.2	*UBE2V1*
Slc12a8	Solute carrier family 12, member 8	2		16 B3	ND	ND
Known CIS						
Sox4	SRY-box containing gene 4	8	Sox4	13 A3-A5	6p22.3	*SOX4*
Mela	Melanoma antigen	7		8	ND	ND
Rarg	Retinoic acid receptor, γ	2	Evi22	15 E-F3	12q13	*RARG*
Bach2	BTB and CNC homology 2	2	Evi59	4	6q15	*BACH2*
Gfi1	Growth factor independent 1	2	Gfi1	5 E5	1p22	*GFI1*
4933427L07Rik	4933427L07Rik protein	1	Evi153	18 E2	ND	ND
Bcl11a	B-cell CLL/lymphoma 11A (zinc finger protein)	1	Evi9	11 A3.2	2p16.1	*BCL11A*
Ccnd1	Cyclin D1	1	Si19	7 F5	11q13	*CCND1*
Ccnd2	Cyclin D2	1	Ccnd2	6 F3	12p13	*CCND2*
Dusp10	Dual specificity phosphatase 10	1		1 H5	1q41	*DUSP10*
Edaradd	EDAR (ectodysplasin-A receptor)-associated death domain	1		13 A1	ND	ND
Eps15-rs	Epidermal growth factor receptor pathway substrate 15	1	Evi90	8 B3.3	19p13.12	*EPS15R*
Gnb1l	G protein, β polypeptide 1-like	1	NKI-13	16 A3	22q11.2	*GNB1L*
Hcph	Hemopoietic cell phosphatase	1	Evi77	6 F3	ND	ND
Idb3	Inhibitor of DNA binding 3	1	Evi62	4 D3	1p36.13-p36.12	*ID3*
Junb	Jun-B oncogene	1	Evi93	8 C2-D1	19p13.2	*JUNB*
Kif9	Kinesin family member 9	1		9 F2	3p21.31	*KIF9*
Lasp1	LIM and SH3 protein 1	1	Evi149	11 C-D	17q11-q21.3	*LASP1*
Mcrs1	Microspherule protein 1	1	Dkmi22	15 F1	12q13.12	*MCRS1*
Nfkb1	Nuclear factor κ-B, subunit 1	1	Evi157	3 G3	4q24	*NFKB1*
Prdm1	PR domain containing 1, with ZNF domain	1	Evi101	10 B2	6q21-q22.1	*PRDM1*
Ptbp1	Polypyrimidine tract binding protein 1	1	Evi103	10 C1	19p13.3	*PTBP1*
Rasgrp1	RAS guanyl releasing protein 1	1	Evi18	2 E5	15q15	*RASGRP1*
Rel	Reticuloendotheliosis oncogene	1		11A3.2	1q21	*REL*
Runx3	Runt related transcription factor 3	1	NKI-11	4 D3	1p36	*RUNX3*
Sfpi1	SFFV proviral integration 1	1	Sfpi1	2 E3	11p11.2	*SPI1*
Snai1	Snail homolog 1 (*Drosophila*)	1	Evi154	2 H3	20q13.1-q13.2	*SNAI1*
Sypl	Synaptophysin-like protein	1	Cm12	12 B2	7q22.2	*SYPL*
Tnnt3	Troponin T3, skeletal, fast	1	Dkmi13	7 F5	11p15.5	*TNNT3*
Zfhx1b	Zinc finger homeobox 1b	1	Nki3	2 C1	2q22	*ZFHX1B*
Zfp36l2	Zinc finger protein 36, C3H type-like 2	1	Evi134	17 E4	2p22.3-p21	*ZFP36L2*
New CIS (shared)						
1500010B24Rik	RIKEN cDNA 1500010B24	3		X F4	Yq11.222	*EIF1AY*
Ehd2	EH-domain containing 2	3		9 E4	19q13.3	*EHD2*
Klf5	Kruppel-like factor 5	2		14 E2.1	13q21.33	*KLF5*
Snf1lk	SNF1-like kinase	2		17 A3.3	21q22.3	*SNF1LK*
Bapx1	Bagpipe homeobox gene 1 homolog	1		5 B3	4p16.1	*BAPX1*
Bat1a	HLA-B-associated transcript 1A	1		17 B1	6p21.3	*BAT1*
Camk2b	Calcium/calmodulin-dependent protein kinase II, β	1		11 A1	7p14.3-p14.1	*CAMK2B*
Cd37	CD37 antigen	1		7 B2	19p13-q13.4	*CD37*
Cirbp	Cold inducible RNA binding protein	1		10 C1	19p13.3	*CIRBP*
Cmah	Cytidine monophospho-*N*-acetylneuraminic acid hydroxylase	1		13 A3.2	6p21.32	*CMAH*
E330036L07Rik	RIKEN cDNA E330036L07	1		15 F1	ND	ND
Emilin2	Elastin microfibril interface located protein 2	1		17 E1.3	ND	ND
G7e-pending	G7e protein	1		17 B1	ND	ND
Ifnar2	Interferon (α and β) receptor 2	1		16 C3.3	21q22.11	*IFNAR2*
Itpr1	Inositol 1,4,5-triphosphate receptor 1	1		6 E1-E2	3p26-p25	*ITPR1*
Oazin	Ornithine decarboxylase antizyme inhibitor	1		15 C	8q22.3	*OAZIN*
Pdlim3	PDZ and LIM domain 3	1		8 B1.1	ND	ND
Pitx1	Paired-like homeodomain transcription factor 1	1		13 B1	5q31	*PITX1*
Ptk6	PTK6 protein tyrosine kinase 6	1		2 H4	20q13.3	*PTK6*
Rfx5	Regulatory factor X, 5 (influences HLA class II expression)	1		3 F2	1q21	*RFX5*
Rgs10	Regulator of G-protein signalling 10	1		7 F3	10q25	*RGS10*
Slfn2	Schlafen 2	1		11 C	ND	ND
Ssr1	Signal sequence receptor, α	1		13 A3.3	ND	ND
Txnrd1	Thioredoxin reductase 1	1		10 C1	12q23-q24.1	*TXNRD1*

[a] CIS, common integration sites; MZL, marginal zone lymphoma; RIS, retroviral integration sites; CFTR, cystic fibrosis transmembrane conductance regulator; MRP, multidrug resistance-associated protein; SRY, sex-determining region Y; ND, not determined; BTB, blood tumor barrier; CLL, chronic lymphocytic leukemia; ZNF, zinc-finger protein; RAS, renin angiotensin; SFFV, spleen focus-forming virus; EH, epoxide hydrolase; SNF1, SNF1 kinase; HLA-B, human leukocyte antigen-B; HLA, human leukocyte antigen.

[b] Numbers of proviral insertion in MZL detected by inverse polymerase chain reaction.

[c] Primary mapping data are from available from the Mouse Genome Informatics database.

[d] Candidate genes are based on their proximity to integrations. The gene nomenclature used is from the Mouse Genome Informatics database (http://www.informatics.jax.org) or the National Center for Biotechnology Information.

with each other and with previously identified RIS in the Mouse Retroviral Tagged Cancer Gene Database[4] yielding a total of 62 CIS (Table 1).

Homology searches in GenBank allowed us to chromosomally map each of 62 CIS (Supplemental Fig. 1). Among these, 31 had been identified as CIS in earlier studies. However, 31 were identified as new. These included 24 that were shared with other lymphoma/leukemia types and 7 that were specific to MZL (Table 1). The 24 CIS that were not MZL specific were defined by one or more RIS identified in our study with an additional site defined in published work. The 7 new CIS that were MZL specific are unique to this study. Sixteen CIS were identified as targets of more than two viral integra-

Table 2 *RIS[a] from MZL in relation to genes involved in human cancer*

Gene	Human disease	Human chromosome	Alteration
CIS genes known or related to genes involved in human cancer[b]			
Brca2	Breast, pancreas, AML, CLL, MDS	13q12-q13	Mutation
Bach2	CML	6q15	Down-regulated expression
Bcl11a	CLL	t(2;14)(p13;q32.3)	BCL11A - IGH
Ccnd1[c]	B-CLL, MM, MCL	t(11;14)(q13;q32)	BCL1/CCND1 - IgH
	SMZL, DLBCL, LPL, breast	t(11;19)(q13;p13)	BCL1/CCND1 - FLRG
Lasp1	AMML-M4	t(11;17)(q23;q21)	ALL1/MLL - LASP1
Runx3	AML, stomach	1p36	Mutation
Klf5	Breast	13q21.33	Somatic deletion
Rnf34	Esophageal	12q24.31	Overexpression
Genes identified once by IPCR known or related to genes involved in human cancer			
Cdk6	SMZL, B-ALL	t(2;7)(p12;q21)	Igκ - CDK6 fusion
	NHL	t(2;7)(p12;q21–22)	Overexpression
Etv6	CML, APL, AMML, B-ALL, AML	t(12;22) (p13;q11)	ETV6/TEL fusion
Hspca	DLBCL	t(3;?)(q27;?)	BCL6/LAZ3 - HSP89A
Pou2af1	ALL	t(3;11)(q27;q23)	BCL6/LAZ3 - POU2AF1/OBF-1
Tcf7	Colon, T-ALL	5q31	Activation
Akt3	Breast, prostate, glioblastoma	1q44	Overexpression
Eif3s10	Lung	10q26	Overexpression
Etv1	Ewing's sarcoma	t(7;22)(p22;q12)	ETV1-EWS fusion
Lgals8	Prostate	1q42-q43	Tumor antigen
Mdm4	Glioma	1q32	Amplification, overexpression
Nr4a3	Chondrosarcoma	t(9;22) (q22;q12)	TEC-EWS fusion
Pik3r1	Colon, ovarian	5q12-q13	Mutation
Prkce	Thyroid	2p21	Amplification
Tgfbr2	Colon, stomach, esophageal	3p22	Mutation

[a] RIS, retroviral integration sites; MZL, marginal zone lymphoma; B-ALL, acute lymphoblastic leukemia/lymphoma, B cell lineage; T-ALL, acute lymphoblastic leukemia/lymphoma, T cell lineage; AML, acute myelogenous leukemia; IPCR, inverse polymerase chain reaction; CLL, chronic lymphocytic leukemia; MDS, myelodysplastic syndrome; CML, chronic myelogenous leukemia; MCL, mantle cell lymphoma; LPL, lymphoplasmacytic lymphoma; SMZL, splenic MZL; DLBCL, diffuse large B-cell lymphoma; MM, multiple myeloma; AMML, acute myelomonocytic leukemia; B-NHL, B cell non-Hodgkin lymphomas; APL, acute promyelocytic leukemia; MZL, marginal zone lymphoma; IGH, immunoglobulin heavy chain; BCL1, B-cell leukemia cyclin D1; CDK6, cyclin-dependent kinase 6; BCL, B-cell leukemia; CCND1, cyclin D1; FLRG, follistatin-related gene; ALL, acute lymphoblastic leukemia; MLL, myeloid-lympoid leukemia; ETV6/TEL, ETS-variant gene6/translocated ets leukemia; OBF1, Oct binding factor 1.

[b] The information of CIS genes that are known or related to human cancer genes is from Atlas of Genetics and Cytogenetics in Oncology and Haematology (http://www.infobiogen.fr/services/chromcancer/Genes/Geneliste.html).

[c] Genes in bold show alteration in human SMZL.

tions and of these, seven were specific to splenic MZL (Table 1). *Sox4* (*Evi16*) was the most frequently targeted CIS (eight tumors).

Most CIS genes encode proteins that fall into classes commonly associated with cancer, including transcription factors, cell cycle regulatory proteins, or proteins involved in signal transduction (Supplemental Table 3). The largest class in this study was transcription factor (*Sox4*, *Bach2*, *Gfi1*, *Prdm1*, *Zfp36l1*, *Nfkb1*, *Bcl11a*, *Junb*, and *Runx3*).

Six CIS genes (*Brca2*, *Bach2*, *Bcl11a*, *Ccnd1*, *Lasp1*, and *Runx3*) are known to be altered by translocation or point mutation in human leukemia or lymphoma (Table 2), indicating again the power of retroviral tagging for identifying candidate disease genes. Furthermore, two CIS genes (*Klf5* and *Rnf34*) are deleted or improperly expressed in other human cancers (Table 2). In addition, the five following oncogenes were identified among the CIS: *Grb2*, *Junb*, *Rel*, *Ccnd2*, and *Gfi1*.

We also identified 150 mouse genes targeted only once by retroviral integration that comprised 116 known genes and 34 uncharacterized genes. Many belong to gene classes associated with cancer (Supplemental Table 3) and are mutated by a variety of mechanisms including translocation, point mutation, and amplification in human cancer (Table 2). Most notably, *Cdk6*, *Etv6*, and *Pou2af1* are dysregulated by translocation in human lymphoma and leukemia (Table 2). In addition, three putative tumor suppressor genes (*Blnk*, *Cdkn1a*, and *Limd1*) and one proto-oncogene (*Rab3a*) were identified in this panel. Because these genes are not located at CIS, however, their potential role in splenic MZL should be interpreted with caution (21).

Differential Patterns of Gene Expression during MZL Progression Determined by DNA Microarray and qPCR Analysis. Previous studies of RIS have examined the expression patterns of only one or a few linked genes. This study takes advantage of an extensive archive of MZL at different stages of progression that can be used for analyzing global gene expression patterns using microarrays and independent, quantitative analyses of expression using qPCR. With microarrays, we studied two of MZL-E, eight of MZL, eight of MZL+, and seven of MZL++ (Fig. 1*A*; Supplemental Fig. 2). As controls, comparisons were made with data from total nude spleen cells that contain on the order of 80% B cells. The qPCR studies examined five cases each of MZL, MZL+, and MZL++ with negatively selected normal splenic B cells serving as controls (Figs. 1*B* and 2).

We first evaluated 16 CIS identified as sites of multiple viral integrations (Fig. 2; Supplemental Table 4). As noted above, *Sox4* (*Evi16*) was the most common target with 8 insertions. Sox4 is a transcriptional activator that is restricted to immature B and T lymphocytes in adult mice and is critical for B-cell differentiation beyond the pro-B stage (22). All eight insertions interrupted 3′ noncoding sequences from 0.9 to 9 kb downstream and might be suspected to enhance message stability (Supplemental Table 4). Of note, six of the eight insertions were recovered from MZL+. The *Sox4* oligonucleotide on the microarrays failed to generate good data. qPCR was used to compare *Sox4* transcript levels in the various stages of progression with levels in purified normal splenic B cells (Fig. 1*B*). *Sox4* transcripts were greatly increased over normal only for MZL+. This suggests that heightened expression of *Sox4* may contribute to the changes that distinguish MZL from MZL+.

Gfi1 encodes a nuclear zinc finger transcriptional repressor that collaborates with *Myc* for transformation of pre-B cells (23) and with *Myc* and *Pim1* for transformation of intrathymic T cells (24, 25). In T-cell lymphomagenesis, heightened expression appears to reduce the requirement for extrinsic growth factors (24) while enhancing proliferation and inhibiting activation-induced cell death (26). Mechanisms of *Gfi1* activity that might contribute to transformation of B lineage cells have not been defined in mouse and human lymphomas. In the present study, we identified two RIS in different tumors located within 1 kb of the transcriptional start site. Both integrations were recovered

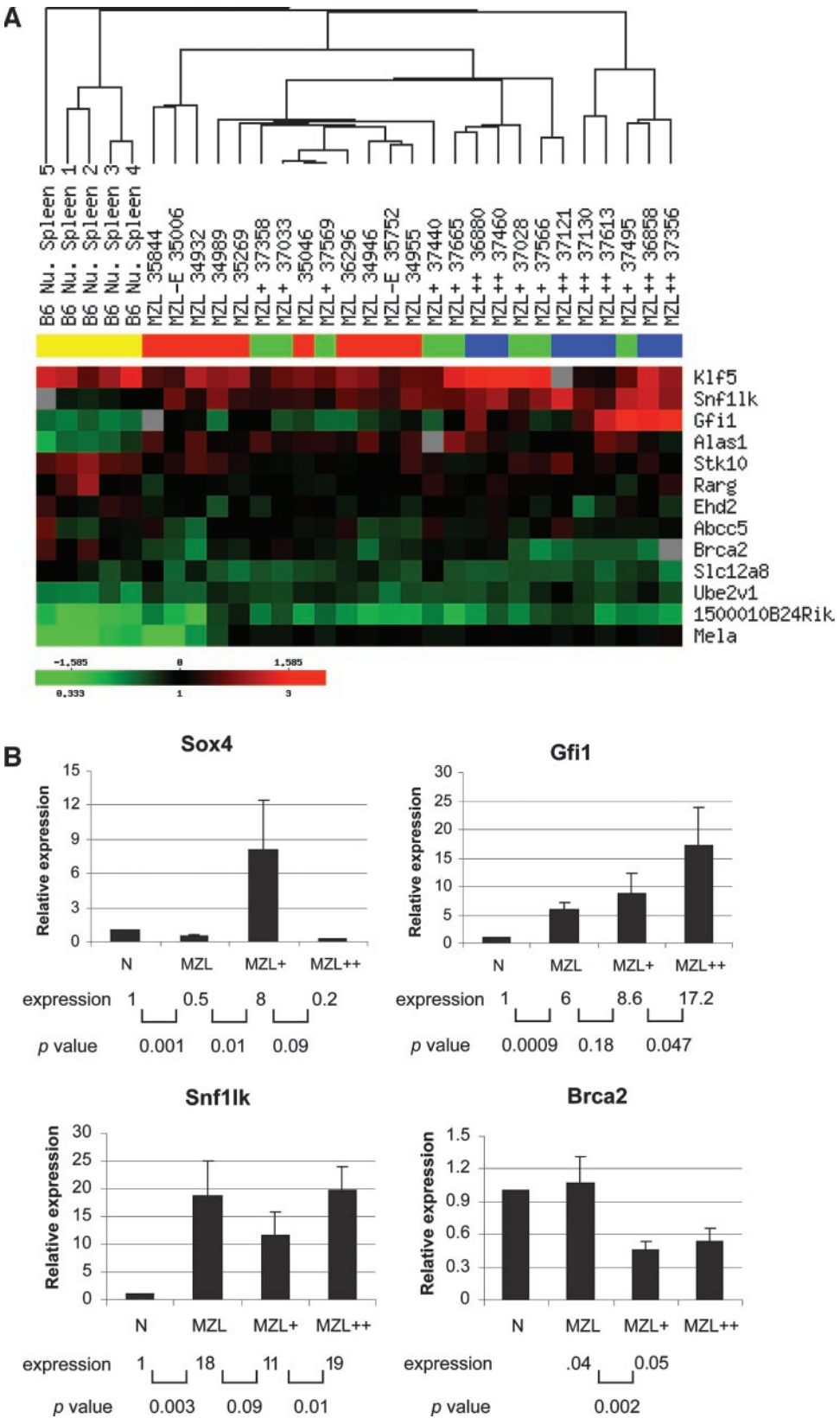

Fig. 1. Expression patterns of multi-hit common integration site (CIS) genes. *A*, hierarchical clustering of genes and arrays results represent the ratio of hybridization of the test samples to the reference mRNA sample. Ratios are depicted according to the *color scale shown at the bottom* that extends from fluorescence ratios of 0.25 to 4. *Gray spots* indicate missing or excluded data. *MZL*, marginal zone lymphoma. *B*, qPCR analysis of *Gfi1*, *Sox4*, *Slc12a8*, and *Brca2* in MZL ($n = 5$), MZL+ ($n = 5$), and MZL++ ($n = 5$). Data are normalized to the mean expression in B cells from normal spleen ($n = 2$). MZL data represent the mean ± 1SD of three independent experiments. *N*, B cells from normal spleen; *qPCR*, quantitative PCR. Mean expression levels are given under *each bar* and *P*s (*t* test) denote significance of differences between values for normal *versus* MZL, MZL *versus* MZL+, and MZL+ *versus* MZL++.

from MZL++ (Supplemental Table 4). Although RNA from these particular cases of MZL++ was not available for testing, eight others tested on the arrays exhibited substantially higher levels of expression than the controls for these studies prepared from normal nude spleen cells (Fig. 1*A*). In qPCR analyses of the five lymphomas representing each stage of progression, transcripts were significantly higher than in normal B cells for each category and increased in level with disease progression (Fig. 1*B*). Increasing expression of *Gfi1* with progression most likely reflects increases in the proportions of tumor cells in spleen that occur with advancing disease rather than increased expres-

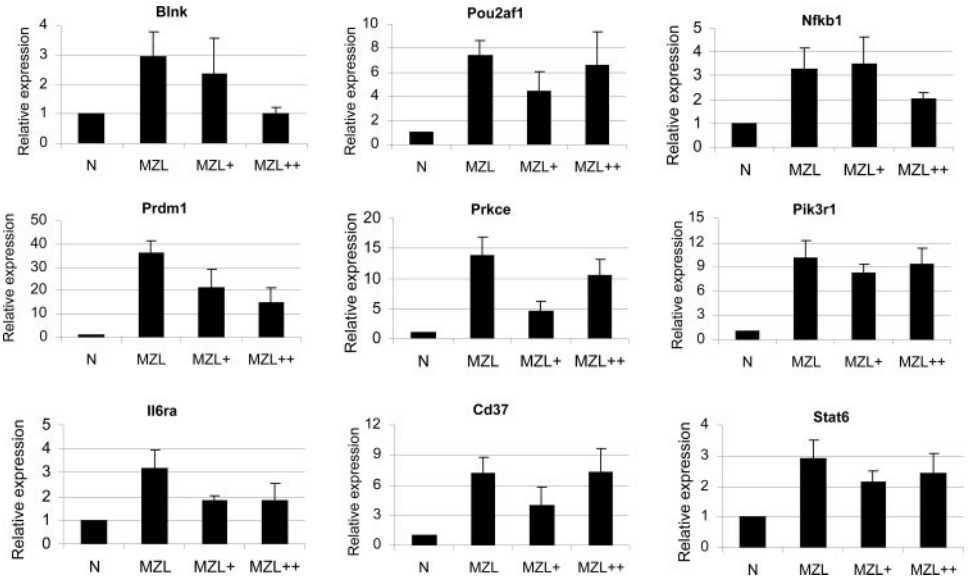

Fig. 2. Expression pattern of retroviral integration site (RIS) genes by qPCR analysis in MZL ($n = 5$), MZL+ ($n = 5$), and MZL++ ($n = 5$). Data are relative to the mean expression in B cells from normal spleen ($n = 2$). MZL data represent the mean of three independent experiments. *N*, B cells from normal spleen; *MZL*, marginal zone lymphoma; *qPCR*, quantitative PCR.

sion per cell. If true this would suggest a role for deregulated *Gfi1* expression in MZL initiation rather than progression. To determine whether this relationship of *Gfi1* to MZL was unique to this lymphoma type or was common to other distinct lymphoma classes, we used qPCR to compare *Gfi1* levels in the three stages of MZL progression to levels in other B lineage lymphomas including centroblastic and immunoblastic diffuse large B-cell lymphomas, Burkitt-like lymphomas and Burkitt lymphoma (Fig. 3). The results of these studies showed that high-level expression of *Gfi1* was unique to MZL.

Snf1lk is the murine homologue of the yeast *Snf1* gene. *Snf1* encodes a histone H3 Ser-10 kinase involved in transcriptional regulation. Activation of specific promoters requires the linked and sequential modifications of histone H3 by Ser-10 phosphorylation that leads to GCN5-dependent acetylation (27). The *Snf1lk* gene is expressed in a variety of normal and transformed mouse and human tissues, but its function in mammalian cells is not known. One insertion near *Snf1lk* was ~90 kb 5′ whereas the second was ~11 kb 3′ (Supplemental Table 4). By both array and qPCR analyses, *Snf1lk* transcripts were expressed at high levels in MZL regardless of the stage of progression (Fig. 1).

Brca2 encodes a nuclear protein implicated in DNA repair, and *Brca2* mutations resulting in inactivation are found in breast and ovarian cancers. It is suggested that Brca2 inactivation may be a key event leading to genomic instability and tumorigenesis (28). Both RIS lay 3′ of the gene within 1 kb of each other. One was recovered from a MZL and the second from a MZL+ (Supplemental Table 4). Analyses of transcripts by both array and qPCR showed that *Brca2* levels were in the range of normal for MZL but were significantly reduced for both MZL+ and MZL++ (Fig. 1). Down-regulation of *Brca2* could potentially increase the frequency of unrepaired DNA breaks that would otherwise activate p53-dependent and other apoptotic pathways. Progression to MZL+ and MZL++ is clearly associated with increased mitotic and apoptotic activity as compared with MZL (10), and down-regulation of *Brca2* may provide a salvage signal for cells destined to die without this change.

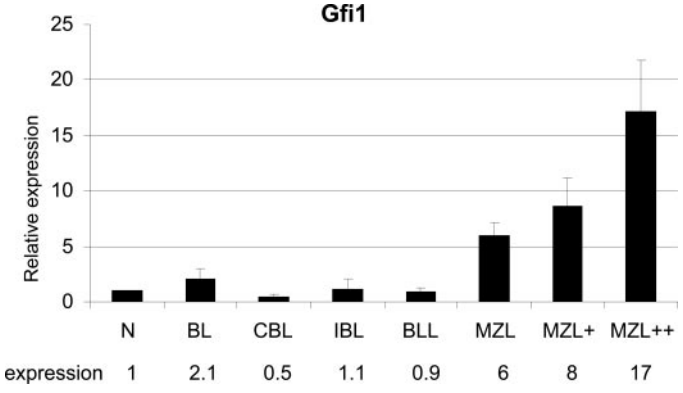

Fig. 3. Quantitation of *Gfi1* transcripts in different classes of B-cell lymphomas. RNA prepared from five cases each of centroblastic (CBL) and immunoblastic (IBL) diffuse large B-cell lymphomas, Burkitt-like (BLL) and Burkitt lymphomas (BL), MZL, MZL+, and MZL++ and two samples of negatively selected normal splenic B cells (*N*) were evaluated for levels of *Gfi1* transcripts by qPCR and normalized to the mean levels in normal B cells. *MZL*, marginal zone lymphoma; *qPCR*, quantitative PCR.

Analyses of other genes located at multi-hit CIS in MZL by array analyses (Fig. 1*A*) and qPCR (data not shown) showed, when taken together, that levels of expression for *Alas1*, *Stk10*, *Rarg*, *Ehd2*, and *Scl12a8* were not significantly altered with disease progression.

We also studied expression of genes at CIS and single-hit RIS and found substantial changes in gene expression patterns for 9 of 35 that were examined by both qPCR and arrays. Two general patterns were seen. First, expression of *Blnk*, *Prdm1*, and *Il6ra* was significantly increased in MZL over the levels found for normal splenic B cells and tended to decrease with disease progression (Fig. 2; Supplemental Fig. 2). None of these genes are directly implicated in human cancers although all contribute in important ways to normal B-cell signaling and differentiation. By qPCR, transcripts for *Pou2af1*, *Nfkb1*, *Prkce*, *Pik3r1*, *Cd37*, and *Stat6* were all substantially increased in MZL over normal B cells, and these expression levels were sustained to varying extents in MZL+ and MZL++ (Fig. 2). *Pou2af1*, through translocation (29), and *Pik3r1*, by mutation of regulatory sequences (30), have been directly implicated in cancer development whereas aberrant activation of the protein encoded by *Nfkb1* is connected with multiple aspects of oncogenesis (31). To determine whether, like *Gfi1*, altered expression of *Stat6*, *Nfkb1*, *Blnk*, *Snf1lk*, *Brca2*, and *Pou2af1* could be specially associated with this lymphoma type, we performed qPCR for these genes using lymphomas of different origin in relation to expression in normal B cells (Supplemental Fig. 3). The results showed that expression of these genes was variable among the various lymphoma types with no unique association with MZL being evident.

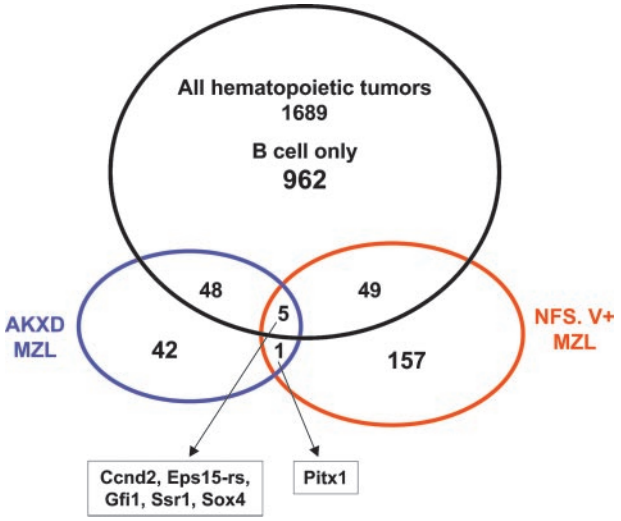

Fig. 4. Ven diagram of relations among retroviral integration sites (RIS) identified in all hematopoietic tumors, B lymphomas alone and MZL from AKXD RI and NFS.V$^+$ mice. *RI*, recombinant inbred; *MZL*, marginal zone lymphoma.

Relations of RIS Found among All Hematopoietic Tumors, B Lymphomas, and MZL from AKXD RI and NFS.V$^+$ Mice. All but one previous study of RIS in mouse lymphomas classified the tumors as T-cell or B-cell origin without attempting to identify subtypes of B-cell lymphomas. In a singular study of lymphomas occurring in AKXD RI strains, 96 RIS were cloned from a series of MZL (18). Importantly, the histological diagnoses in the latter study were made by the same pathologist as for the present study, and the targets were cloned using techniques identical to those used in this study. This minimizes variabilities that could arise from inter-observer bias in making histological diagnoses or from different technical approaches to identifying RIS. We reasoned that the same disease in the two genetic backgrounds (AKXD and NFS), with Akv-type ecotropic viruses as the common mutator, would share a substantial proportion of the sites identified in either setting. Remarkably, only 6 of 300 RIS were shared by the two sets of MZL (Fig. 4; Supplemental Table 5).

To determine whether other analysis of a single disease induced on different genetic backgrounds might show similar dissociations of candidate disease genes, we turned to data from two separate studies. The first included two sets of mice deficient in p27 that were infected neonatally with Moloney MuLV (32, 33). Both sets developed almost exclusively thymic T-cell lymphomas. Among 62 RIS identified in one set by inverse PCR (32) and 30 identified in the second using splinkerettes (33), only five were shared. These included *Myc/Pvt1*, *Gfi1*, *Rras*, *Jundp1*, and an unknown gene. The second study (18) examined RIS is two sets of Burkitt-like lymphoma, one appearing spontaneously in AKXD RI mice and the second induced in NIH Swiss mice by inoculation of MuLV recovered from B lymphomas of CFW mice (34). Among 208 RIS cloned from AKXD tumors and 87 from the NIH Swiss, only eight were shared. Taken together, these findings raise the possibility that strain background differences may markedly effect the types of genes contributing to the pathogenesis of a single disease. Precedent for this comes from studies of Eμ-Myc transgenic mice that demonstrated the development of T-cell lymphomas on one genetic background and B-cell lymphomas on another (35). It seems less likely that histologically indistinguishable diseases in genetically different mice represent dissimilar disorders.

DISCUSSION

Splenic MZL of mice present a major challenge for understanding the molecular events involved not only in the initiation but also the progression of a single disease. In the present study, we used the approach of proviral tagging of candidate disease genes to generate insights into the pathogenesis of this disorder. By examining primary tissues from MZL at all stages of progression, we identified 62 CIS, 31 of which were newly identified when paired with single hits from other studies and 7 of which were specific for MZL. Transcriptional profiling using DNA microarrays and qPCR made it possible to relate these sites to gene expression patterns. Two genes of particular interest, *Gfi1* and *Sox4*, were among the set that was found to be targeted in both this and an earlier survey of RIS in MZL of AKXD RI mice (18). *Gfi1* was expressed at increased levels throughout all stages of MZL progression whereas heightened expression of *Sox4* was unique to high-grade MZL+.

RIS at *Gfi1* fall into a ~50-kb region that comprises a series of CIS including *Evi5*, *Eis1*, and *Pal1*. Insertions at any of these sites result in enhanced transcription of *Gfi1* (23). The insertions identified in this study fell into the *Gfi1* CIS. Although studies involving selection for factor independence of lymphomas *in vitro* identified a role for *Gfi1* in T-cell lymphoma progression (36), the gene is also involved in the initiation of tumorigenesis (23). *Gfi1* is a transcriptional repressor that is thought to contribute to transformation by restricting the transcription of genes normally inhibitory for cell survival, such as *Bax* and *Bak* (37) and proliferation, including p27^{Kip1} (38). Our study suggests a role for deregulated *Gfi1* expression in lymphoma initiation rather than progression.

Gfi1 collaborates with *Myc* in accelerating the development of pre-B lymphomas in Eμ-*Myc* transgenic mice infected with Moloney MuLV (39). In these tumors, integrations at *Gfi1* are mutually exclusive with insertions at the *Bmi1* locus. This indicates that *Gfi1* and *Bmi1* belong to the same complementation group (19, 21) and, by extension, may act on similar downstream targets for lymphoid transformation. Insertions at *Bmi1* result in enhanced expression of an unmutated protein (37, 38). *Bmi1*, like *Gfi1*, is a transcriptional repressor but belongs to the polycomb family of genes. Eμ-*Bmi1*-transgenic mice overexpressing the wild-type protein exhibit perturbed lymphoid development and succumb almost uniformly to B and T cell neoplasms (40, 41) demonstrating activity as a directly transforming oncogene. The tumor suppressor genes *p16* and *p19ARF* are critical downstream targets of *Bmi1*-mediated repression (42) thereby providing separate, cooperative functions with Myc for transformation. Although the relevance for *Bmi1*-mediated transformation is not clear, recent studies have demonstrated markedly different profiles for B-cell expression of *Bmi1* in association with distinct stages of the human germinal cell reaction (43); features of MZ B-cell expression were not reported.

Sox4 was the most frequent CIS identified in earlier screens of mouse hematopoietic tumors, with B-cell lymphomas being targeted more frequently than T-cell or myeloid neoplasms.[7] It was also the most common target found in our study. The great majority of insertions lie within 3-kb 3′ to the gene. *Sox4* belongs to a large family of transcription factors related by homology in their DNA-binding domains to the HMG-box region of the testis-determining gene, *SRY*. Expression in adult mice is limited to immature T and B lymphocytes, and B-cell development in mice lacking *Sox4* is blocked at the pro-B-cell stage (22). Interestingly, *Sox4* is a downstream target of the interleukin-5Rα chain with syntenin functioning as an intermediate (44). The genes regulated by this specific pathway are not known.

The finding that heightened expression of *Sox4* was restricted to MZL+, the tumor grade in which it was most frequently targeted, is intriguing. This could suggest that activation of *Sox4* is required for

[7] http://rtcgd.ncifcrf.gov/mm4/index.html.

progression of MZL to MZL+, but that expression need not be sustained for progression to MZL++. Possibly other genetic changes may substitute for *Sox4* in maintaining these signaling requirements. A similar suggestion has been made for *MYC*-independent tumors that develop after *MYC* inactivation in a conditional transgenic model for *MYC* regulation (45). Alternatively, not all MZL+ may develop by progression from MZL nor MZL++ by from progression from MZL+. Human diffuse large B-cell lymphomas can appear *de novo* or by progression from pre-existing follicular lymphoma. The possibility that *Sox4* insertions might distinguish *de novo* MZL+ from cases that progress from MZL will require additional study.

It is quite striking that the RIS identified in MZL of NFS.V$^+$ mice showed so little overlap with RIS in MZL of AKXD RI strains, only 6 of 300 or 2% of sites were held in common. Remarkably, little concordance also was seen between data from two sets of mice with Moloney-induced thymic T-cell lymphomas (5 of 92 sites) and another two other sets of mice with spontaneous Burkitt-like lymphomas (8/295 sites). The differences cannot be ascribed to discrepancies in histological classification as all diagnoses of MZL and Burkitt-like lymphomas were made by the same pathologist, and thymic T lymphomas are uniformly lymphoblastic. The differences between the two sets of MZL and Burkitt-like lymphomas are also not attributable to technical differences because the same technique was used throughout both studies. RIS in the two sets of mice with T lymphomas were identified by different PCR techniques (32). It seems unlikely that strain-specific polymorphisms could bias the integrations in a systematic way or select among a remarkably large range of pathways that can lead to the same pathological phenotype.

These considerations can be extended to question the degree to which the integrations in any study deviate from a pattern that could be considered random. Algorithms have been developed to assign parameters of significance to the definition of a CIS, but they are known to be compromised by the inability to account for preferential insertions determined by sequence, chromatin conformation, or gene expression. A recent publication provides unusual insights into this problem. In a study of over 900 different MuLV RIS in human HeLa cells, it was shown that there was preferential integration in proximity to the start sites of transcriptionally active genes (46); 34.2% of integrations were within genes, 11.2% were within 5-kb upstream of genes, and 3.4% were within 5-kb downstream of genes. The frequencies of integrations within and 5′ to genes were significantly different from random whereas that for downstream hits was indistinguishable from random. The fact that these sites were recovered 48 h postinfection should exclude any bias based on cell growth or viability. In addition, they were likely to reflect only primary integrations because the data were mostly generated using recombinant retroviruses.

The data from the present study showed that 34.2% of the RIS in NFS.V$^+$ MZL were within genes, 20.6% were within 5-kb upstream of genes, and 6.6% were within 5-kb downstream of genes. The frequency of RIS within genes does not distinguish this study from that of human cells, but the frequencies of both 5′ and 3′ integrations were significantly different from those described in HeLa ($P = 0.004$ for 5′ and $P = 0.016$ for 3′ integrations; Fisher exact test). Consequently, the 3′ integrations also differed significant from the calculations for random placements in HeLa cells. Although extrapolation of data from a human carcinoma cell line to spontaneous MZL of mice is likely to have its limitations, we feel the differential features of RIS in NFS.V$^+$ MZL reflect a bias toward survival and proliferative advantages conferred by these mutagenic events.

These findings place several demands on future studies designed to identify genes involved in specific lymphoid diseases such as MZL. First, it will be critical to devise means for distinguishing RIS that reflect "noise" from those truly involved in pathogenesis. This need is quite distinct from that required of sensitized screens designed to define complementation groups. Second, it will be necessary to develop mechanisms for expressing candidate disease genes or aborting their expression in restricted lymphoid subsets that differ but little from other cells of the same lineage in stage of differentiation or anatomical location. As an example, there is currently no promoter/enhancer cassette that could be used to reliably drive expression of a gene in MZ B cells. Our studies of MZL have established a firm foundation for approaching these issues.

ACKNOWLEDGMENTS

We thank Jim Neil, Anton Berns, and Ewan Cameron for thoughtful comments on the manuscript and members of the Morse laboratory for helpful discussions.

REFERENCES

1. Rosenberg N, Jolicoeur P. Retroviral pathogenesis. In: Coffin JM, Hughes SH, Varmus HE, editors, Retroviruses, Cold Spring Harbor, NY: Cold Sprimg Harbor Laboratory Press: 1997. p. 475–586.
2. Bedigian HG, Johnson DA, Jenkins NA, Copeland NG, Evans R. Spontaneous and induced leukemias of myeloid origin in recombinant inbred BXH mice. J Virol 1984;51:586–94.
3. Gilbert DJ, Neumann PE, Taylor BA, Jenkins NA, Copeland NG. Susceptibility of AKXD recombinant inbred mouse strains to lymphomas. J Virol 1993;67:2083–90.
4. Li J, Shen H, Himmel KL, et al. Leukaemia disease genes: large-scale cloning and pathway predictions. Nat Genet 1999;23:348–53.
5. Hansen GM, Skapura D, Justice MJ. Genetic profile of insertion mutagenesis in mouse leukemias and lymphomas. Genome Res 2000;10:237–43.
6. Copeland NG, Jenkins NA. Myeloid leukemia: disease genes and mouse models. In: Hiai H, Hino O, editors. Animal models of cancer predisposition syndromes. Basel: Karger; 1999. p. 53–63.
7. Ogawa S, Kurokawa M, Tanaka T, et al. Structurally altered Evi-1 protein generated in the 3q21q26 syndrome. Oncogene 1996;13:183–91.
8. Roberts T, Chernova O, Cowell JK. NB45, a member of the TBC1 domain family of genes, is truncated as a result of a constitutional t(1;10)(p22;q21) chromosome translocation in a patient with stage 45 neuroblastoma. Hum Mol Genet 1998;7:1169–78.
9. Look AT. Oncogenic transcription factors in the human acute leukemias. Science (Wash D C) 1997;278:1059–64.
10. Fredrickson TN, Lennert K, Chattopadhyay SK, Morse HC III, Hartley JW. Splenic marginal zone lymphomas of mice. Am J Pathol 1999;154:805–12.
11. Hartley JW, Chattopadhyay SK, Lander MR, et al. Accelerated appearance of multiple B cell lymphoma types in NFS/N mice congenic for ecotropic murine leukemia viruses. Lab Investig 2000;80:159–69.
12. Morse HC III, Qi CF, Chattopadhyay SK, et al. Combined histologic and molecular features reveal previously unappreciated subsets of lymphoma in AKXD recombinant inbred mice. Leuk Res 2001;8:719–33.
13. Ward JM, Tadesse-Heath L, Perkins SN, Chattopadhyay SK, Hursting SD, Morse HC III. Splenic marginal zone B-cell and thymic T-cell lymphomas in p53-deficient mice. Lab Investig 1999;1:3–14.
14. Lopes-Carvalho T, Kearney JF. Development and selection of marginal zone B cells. Immunol Rev 2004;197:192–205.
15. Fredrickson TN, Harris AW. Atlas of mouse hematopathology. Amsterdam, Netherlands: Harwood Academic Publishers; 2000.
16. Lund AH, Turner G, Trubetskoy A, et al. Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. Nat Genet 2002;32:160–5.
17. Mikkers H, Allen J, Knipscheer P, Romeyn L, Hart A, Berns A. High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. Nat Genet 2002;32:153–9.
18. Suzuki T, Shen H, Akagi K, et al. New genes involved in cancer identified by retroviral tagging. Nat Genet 2002;32:166–74.
19. Mikkers H, Berns A. Retroviral insertional mutagenesis: tagging cancer pathways. Adv Cancer Res 2003;88:53–99.
20. Morse HC, Anver MR, Fredrickson TN, et al. Bethesda proposals for classification of lymphoid neoplasms in mice. Blood 2002;100:246–58.
21. Neil JC, Cameron ER. Retroviral insertion sites and cancer: fountain of all knowledge? Cancer Cell 2002;2:253–5.
22. Schilham MW, Oosterwegel MA, Moerer P, et al. Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. Nature (Lond) 1996;380:711–4.
23. Scheijen B, Jonkers J, Acton D, Berns A. Characterization of pal-1, a common proviral insertion site in murine leukemia virus-induced lymphomas of c-myc and Pim-1 transgenic mice. J Virol 1997;71:9–16.
24. Zornig M, Schmidt T, Karsunky H, Grzeschiczek A, Moroy T. Zinc finger protein GFI-1 cooperates with myc and pim-1 in T-cell lymphomagenesis by reducing the requirements for IL-2. Oncogene 1996;12:1789–801.
25. Schmidt T, Karsunky H, Gau E, Zevnik B, Elsasser HP, Moroy T. Zinc finger protein GFI-1 has low oncogenic potential but cooperates strongly with pim and myc genes in T-cell lymphomagenesis. Oncogene 1998;17:2661–77.

26. Karsunky H, Mende I, Schmidt T, Moroy T. High levels of the onco-protein Gfi-1 accelerate T-cell proliferation and inhibit activation induced T-cell death in Jurkat T-cells. Oncogene 2002;21:1571–9.
27. Lo WS, Duggan L, Tolga NC, et al. Snf1–a histone kinase that works in concert with the histone acetyltransferase Gcn5 to regulate transcription. Science (Wash D C) 2001;293:1142–6.
28. Davies AA, Masson JY, McIlwraith MJ, et al. Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. Mol Cell 2001;7:273–82.
29. Galieque Zouitina S, Quief S, Hildebrand MP, et al. The B cell transcriptional coactivator BOB1/OBF1 gene fuses to the LAZ3/BCL6 gene by t(3;11)(q27;q23.1) chromosomal translocation in a B cell leukemia line (Karpas 231). Leukemia 1996; 10:579–87.
30. Philp AJ, Campbell IG, Leet C, et al. The phosphatidylinositol 3′-kinase p85α gene is an oncogene in human ovarian and colon tumors. Cancer Res 2001;61: 7426–39.
31. Baldwin AS. Control of oncogenesis and cancer therapy resistance by the transcription factor NF-kappaB. J Clin Investig 2001;107:241–6.
32. Hwang HC, Martins CP, Bronkhorst Y, et al. Identification of oncogenes collaborating with p27Kip1 loss by insertional mutagenesis and high-throughput insertion site analysis. Proc Natl Acad Sci USA 2002;99:11293–8.
33. Martins CP, Berns A. Loss of P27(Kip1) but Not P21(Cip1) decreases survival and synergizes with Myc in murine lymphomagenesis. EMBO J 2002;21:3739–48.
34. Taddesse-Heath L, Chattopadhyay SK, Dillehay DL, et al. Lymphomas and high-level expression of murine leukemia viruses in CFW mice. J Virol 2000;74:6832–47.
35. Yukawa K, Kikutani H, Inomoto T, et al. Strain dependency of B and T lymphoma development in immunoglobulin heavy chain enhancer (Eμ)-myc transgenic mice. J Exp Med 1989;170:711–26.
36. Gilks CB, Bear SE, Grimes HL, Tsichlis PN. Progression of interleukin-2 (Il-2)-dependent rat T-cell lymphoma lines to Il-2-independent growth following activation of a gene (Gfi-1) encoding a novel zinc finger protein. Mol Cell Biol 1993;13: 1759–68.
37. Grimes HL, Gilks CB, Chan TO, Porter S, Tsichlis PN. The Gfi-1 protooncoprotein represses Bax expression and inhibits T-cell death. Proc Natl Acad Sci USA 1996; 93:14569–73.
38. Zhu J, Guo L, Min B, et al. Growth factor independent-1 induced by IL-4 regulates Th2 cell proliferation. Immunity 2002;16:733–44.
39. van Lohuizen M, Verbeek S, Scheijen B, Wientjens E, van der Gulden H, Berns A. Identification of cooperating oncogenes in Eμ-myc transgenic mice by provirus tagging. Cell 1991;65:737–52.
40. Haupt Y, Bath M. L, Harris AW, Adams JM. Bmi-1 transgene induces lymphomas and collaborates with Myc in tumorigenesis. Oncogene 1993;8:3161–4.
41. Alkema MJ, Jacobs H, Vanlohuizen M, Berns A. Perturbation of B and T cell development and predisposition to lymphomagenesis in E Mu Bmi1 transgenic mice require the Bmi1 ring finger. Oncogene 1997;15:899–910.
42. Jacobs JJL, Kieboom K, Marino S, Depinho RA, Van Lohuizen M. The Oncogene and polycomb-group gene Bmi-1 regulates cell proliferation and senescence through the Ink4a locus. Nature (Lond) 1999;397:164–8.
43. Raaphorst FM, Van Kemenade FJ, Fieret E, et al. Cutting edge: polycomb gene expression patterns reflect distinct B cell differentiation stages in human germinal centers. J Immunol 2000;164:1–4.
44. Geijsen N, Uings IJ, Pals C, et al. Cytokine-specific transcriptional regulation through an IL-5Ralpha interacting protein. Science (Wash D C) 2001;293:1136–48.
45. Karlsson A, Giuriato S, Tang F, Fung-Weier J, Levan G, Felsher DW. Genomically complex lymphomas undergo sustained tumor regression upon Myc inactivation unless they acquire novel chromosomal translocations. Blood 2003;101:2797–803.
46. Wu X, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. Science (Wash D C) 2003;300:1749–51.