RAG2 is regulated differentially in B and T cells by elements 5′ of the promoter

Robert J. Monroe*, Feng Chen*, Roger Ferrini, Laurie Davidson, and Frederick W. Alt†

Howard Hughes Medical Institute, The Children's Hospital, Center for Blood Research and Department of Genetics, Harvard Medical School, Boston, MA 02115

To study *RAG2* gene regulation *in vivo*, we developed a blastocyst complementation method in which *RAG2*-deficient embryonic stem cells were transfected with genomic clones containing *RAG2* and then assessed for their ability to generate lymphocytes. A *RAG2* genomic clone that contained only the *RAG2* promoter sequences rescued V(D)J recombination in *RAG2*-deficient pro-B cell lines, but did not rescue development of *RAG2*-deficient lymphocytes *in vivo*. However, inclusion of varying lengths of sequences 5′ of the *RAG2* promoter generated constructs capable of rescuing only *in vivo* B cell development, as well as other constructs that rescued both B and T cell development. In particular, the 2-kb 5′ region starting just upstream of the RAG2 promoter, as well as the region from 2–7 kb 5′, could independently drive B cell development, but not efficient T cell development. Deletion of the 2-kb 5′ region from the murine germ line demonstrated that this region was not required for *RAG* expression sufficient to generate normal B or T cell numbers, implying redundancy among 5′ elements. We conclude that *RAG2* expression *in vivo* requires elements beyond the core promoter, that such elements contribute to differential regulation in the B vs. T lineages, and that sequences sufficient to direct B cell expression are located in the promoter-proximal 5′ region.

Immunoglobulin and T cell receptor (TCR) variable region genes are assembled in developing lymphocytes from germ-line variable (V), diversity (D), and joining (J) gene segments by a site-specific recombination reaction known as V(D)J recombination (1). The recombination-activating genes *RAG1* and *RAG2* encode the essential lymphocyte-specific components of the reaction, which initiate V(D)J recombination by introducing double-stranded DNA breaks between Ig and TCR V, D, and J coding segments and adjacent recombination signal (RS) sequences (2). To limit V(D)J recombinase activity and the inherent double-stranded DNA breaks to the appropriate lymphocyte subsets, the expression of *RAG* genes is tightly and coordinately regulated during early B and T lymphocyte development.

In the B cell lineage, *RAG* genes are expressed first in pro-B cells undergoing Ig heavy (H) chain gene rearrangements; *RAG* down-regulation follows functional $V_H DJ_H$ rearrangement and pre-B cell receptor expression (3). *RAG* expression is up-regulated in pre-B cells, coincident with the onset of Ig light (L) chain gene rearrangement and persists in surface IgM-positive (sIgM+) immature B cells (3–5). In this context, immature B cells possess the capacity to undergo receptor editing—a process that includes secondary L chain gene rearrangements and specificity changes (6). *RAG* expression is down-regulated again, as immature B cells acquire high levels of sIgM, leave the bone marrow (BM), and enter the splenic transitional B cell compartment (7). In the T cell lineage, *RAG* genes are expressed first in the CD4−CD8− double-negative (DN) thymocyte compartment, the stage during which TCRβ, -γ, and -δ rearrangement commences (8, 9). Productive $V_\beta DJ_\beta$ rearrangement leads to expression of a pre-TCR, cellular expansion and down-regulation of *RAG* gene expression (10). Subsequently, *RAG* expression is up-regulated with the onset of TCRα gene rearrangement and

differentiation of αβ lineage cells into CD4+CD8+ double-positive (DP) thymocytes (8, 11). *RAG* expression is down-regulated a second time following maturation of DP cells into CD4+CD8− or CD4−CD8+ single-positive (SP) thymocytes (12).

Outside of the primary lymphoid tissues, *RAG* expression is limited. In young or immunized mice, *RAG* genes are expressed in a population of splenic B lineage cells with a pre-B cell phenotype (7, 13). *RAG* expression has also been detected in small subsets of peritoneal B1 cells and peripheral T cells subjected to chronic negative selection (14, 15). Only in rare instances, such as the mammalian central nervous system and the chicken bursa, is the expression of one *RAG* gene found in the absence of the other (16, 17). Therefore, it is not surprising that the *RAG1* and *RAG2* genes are tightly linked in a tail-to-tail configuration on the chromosome, lying only a few kilobases apart. This genomic organization has been conserved in all species examined thus far, leading to the hypothesis that the *RAG* genes are derived from a transposon that integrated into the genome of a common vertebrate ancestor (18, 19). This idea has been further supported by findings that the RAG proteins can drive transposition of DNA sequences *in vitro* (20, 21). Moreover, the proximity and coordinate expression of the two genes suggest a common mechanism of transcriptional regulation, possibly by similar promoters or by shared cis-acting elements in the locus, such as enhancers or locus control regions.

Both the *RAG1* and *RAG2* promoters are highly conserved between mice and humans (22, 23). While the human and murine *RAG1* promoters are active in both lymphoid and nonlymphoid cell lines, the murine *RAG2* promoter, which lies within 279 bp of the transcription start site, is lymphoid-specific and differentially regulated in B vs. T cell lines (22–25). In this regard, a conserved sequence critical for *RAG2* promoter activity in B cell lines is bound by the B cell-specific transcription factor BSAP in a tissue-specific manner, while other sequences are required for full promoter activity in T cell lines (23). To identify elements that control *RAG2* gene expression during lymphocyte development *in vivo*, we developed a genetic assay and further characterized potential functions of one such element by means of gene-targeted mutation.

Materials and Methods

Generation of Genomic and Targeting Constructs. An 18-kb *Not*I–*Not*I fragment (13NN) from the murine λ13 genomic phage clone was subcloned into pBluescriptII (Stratagene). Truncation constructs (13RS, 13SN, 13SC, 13HC) were assembled in pBlue-

IMMUNOLOGY

scriptII from 13NN subfragments by using standard cloning procedures. The pLCK-Ras^{V12} construct has been described (26). The *Stu*I–*Hin*dIII (SHΔ) targeting construct (SH-KO) was generated by cloning 4.5-kb *Eco*RI–*Stu*I and 3-kb *Hin*dIII–*Hin*dIII fragments from 13NN into the *Sal*I and *Xho*I sites of the pLNTK vector (27).

DNA and RNA Analysis. Whole cell RNA was prepared with TRIzol (GIBCO/BRL). Genomic DNA isolation and Southern and Northern blot analyses were performed by using standard protocols as described (27). The 5′ and 3′ SH-KO probes are 500-bp *Bam*HI–*Eco*RI genomic and 1.0-kb *Kpn*I–*Eco*RV RAG2 cDNA fragments, respectively. *RAG2*$^{-/-}$ embryonic stem (ES) cell clones carrying *RAG2* genomic fragments (pLCK-Ras^{V12} construct) were identified by Southern blotting of *Eco*RI- and *Xba*I-digested DNA (*Bam*HI-digested DNA) by using the *Kpn*I–*Eco*RV probe (Ras cDNA probe). Copy number was estimated by comparing relative intensities of endogenous and construct bands. Additional probes (isolated from 13NN) used to confirm integrity of genomic constructs were as follows: 1.5-kb *Not*I–*Eco*RI, 500-bp *Bam*HI–*Eco*RI, 800-bp *Pst*I–*Pst*I, 800-bp *Bgl*II–*Sma*I, and 1.8-kb *Hin*dIII–*Cla*I fragments. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and RAG1 cDNA probes were also used. Densitometry was performed with a Phosphor-Imager and ImageQuant software (Molecular Dynamics). Total signal was calculated by volume integration, and background was defined as the average signal of the object border.

ES Cell Derivation, Transfection, and Analysis. R2$^{BL/5.1}$ ES cell lines were derived by culturing *RAG2*$^{-/-}$ blastocysts on murine embryonic fibroblasts (MEFs). After 3–5 days, inner cell mass (ICM) outgrowths were detached, trypsinized, and replated on MEFs. ES cell colonies appearing after 4–6 days were subcloned and expanded. Several R2$^{BL/5.1}$ ES cell lines were used to generate chimeric mice by RAG2-deficient blastocyst complementation (R2DBC) as described (28); one line, which gave high levels of contribution, was selected for transfection studies. R2$^{BL/5.1}$ ES cells were electroporated with 25 μg of linearized construct DNA and 2.5 μg of linearized pCMV-HygroB:TK, and selected in medium containing 0.5 μg/ml hygromycin B. ES clones carrying appropriate constructs were identified by Southern blot analyses and used to generate chimeric mice by R2DBC.

Generation of SHΔ Mutant ES Cells and Mice. TC1 (wild type) or RAG2:GFP (RAG2$^{GFP/+}$; GFP indicating green fluorescent protein) ES cells (7) were electroporated with 30 μg of *Pvu*I-linearized SH-KO DNA and selected in medium containing G418 and ganciclovir as described (27). SH$^{N/+}$ and SH$^{N(GFP)/+}$ ES cells were identified by Southern blot analysis using the 3′ and 5′ SH-KO probes on *Bam*HI-digested DNA (Figs. 3*B* and 4*B*; and data not shown). SH$^{N/+}$ and SH$^{N(GFP)/+}$ ES cells were transiently transfected with 30 μg of pMC-CreN; clones were isolated and screened by Southern blot analysis for deletion of the phosphoglycerate kinase promoter-driven–neomycin resistance (*PGK-neo*^{r}) gene (SH$^{Δ/+}$ and SH$^{Δ(GFP)/+}$; Figs. 3*B* and 4*B*). SH$^{Δ/+}$ ES cells were subjected to a second round of targeting and Cre-deletion, as above, to isolate SH$^{Δ/Δ}$ ES cells. Two independent SH$^{Δ/Δ}$ and SH$^{Δ(GFP)/+}$ clones were used to generate chimeric mice by R2DBC.

Flow Cytometry. Single cell suspensions were stained with FITC-, phycoerythrin (PE)-, and Cy-Chrome (CyC)-conjugated Abs and analyzed by a FACSCalibur (Becton Dickinson). The following Abs were used (PharMingen): FITC-anti-CD45.1(Ly5.1;A20), -CD45.2(Ly5.2;104); PE-anti-CD4(RM4-5), -CD43(S7), -IgM^a(Igh-6a); and CyC-anti-CD8(53-6.7), -B220 (RA3-6B2). For most FACS plots, ≥20,000 events were collected; dead cells were excluded by size and forward-scatter
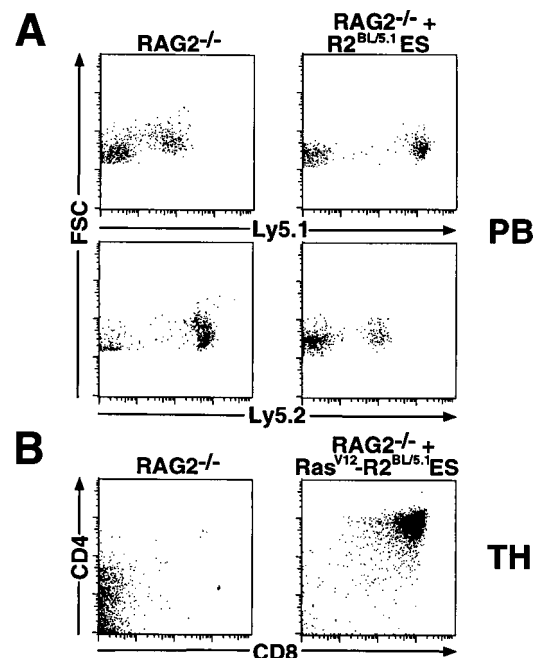

Fig. 1. Characterization of *RAG2*$^{-/-}$ ES cells by means of R2DBC. (*A*) Somatic chimeras were generated by injection of R2$^{BL/5.1}$ ES cells into *RAG2*$^{-/-}$ blastocysts; peripheral blood (PB) isolated from chimeras and *RAG2*$^{-/-}$ controls was analyzed by FACS after staining with FITC-anti-Ly5.1 or FITC-anti-Ly5.2. Ly5.1^{low} cells (*RAG2*$^{-/-}$ panel) and Ly5.2^{low} cells (*RAG2*$^{-/-}$ + R2$^{BL/5.1}$ ES panel) represent a population of leukocytes reproducibly identified in *RAG2*$^{-/-}$ mice, which stain nonspecifically with both FITC-anti-Ly5.1 and anti-Ly5.2. (*B*) R2$^{BL/5.1}$ ES cells transfected with the pLCK-Ras^{V12} construct were assayed by R2DBC. FACS analysis was performed on thymocytes (TH) isolated from chimeras and *RAG2*$^{-/-}$ controls after staining with PE-anti-CD4 and CyC-anti-CD8. Data shown are from representative chimeras.

gating. Data were analyzed with CellQuest (Becton Dickinson) software.

Results

Derivation and Characterization of *RAG2*$^{-/-}$ ES Cell Lines. We developed an ES cell-based assay to test the ability of the *RAG2* promoter (23) and surrounding sequences to drive developmental stage- and lineage-specific expression of *RAG2 in vivo*. Specifically, we transfected various genomic fragments of the *RAG2* gene into *RAG2*$^{-/-}$ ES cells and then employed the R2DBC assay (28) to assess capacity of the transfected cells to support B and T cell development. For this purpose, we first derived several *RAG2*$^{-/-}$ ES cell lines by culturing blastocysts isolated from *RAG2*$^{-/-}$ mice, which had been bred to carry the *Ly5.1* allele and black coat color (R2$^{BL/5.1}$) (29). When injected into *RAG2*$^{-/-}$ blastocysts isolated from an independent *RAG2*$^{-/-}$ colony carrying the *Ly5.2* allele and agouti coat color (R2$^{AG/5.2}$), these ES cells have the potential to give rise to black fur and Ly5.1-expressing leukocytes.

We tested several R2$^{BL/5.1}$ ES lines for contribution in *RAG2*$^{-/-}$ chimeras generated with R2$^{AG/5.2}$ blastocysts. While the resulting chimeras lacked consistently high levels of black coat color chimerism, they possessed reproducibly high numbers of Ly5.1$^+$ leukocytes, as assessed by FACS analysis of peripheral blood cells (Fig. 1*A*). As a positive control for the ability of the R2$^{BL/5.1}$ ES cells to contribute specifically to lymphocyte lineages, we transfected them with a pLCK-Ras^{V12} expression construct shown previously to promote the differentiation and expansion of RAG-deficient DN thymocytes to DP thymocytes (26). Ras^{V12}-transfected clones gave rise to large numbers of DP

thymocytes in $RAG2^{-/-}$ chimeras (Fig. 1B), demonstrating that the $R2^{BL/5.1}$ ES cells are capable of significant contribution to lymphoid lineages (30).

Sequences Upstream of the *RAG2* Core Promoter Are Required to Generate Specific *RAG2* Expression in Vivo. Having established the capacity of $R2^{BL/5.1}$ ES cells to contribute to lymphoid lineages in $RAG2^{-/-}$ chimeric mice, we employed these cells to test whether a set of genomic DNA fragments encompassing the *RAG2* gene could rescue lymphocyte development *in vivo*. All tested clones contained the complete *RAG2* coding sequences and varying amounts of 5′ and 3′ flanking sequences. The 13HC construct contained the *RAG2* promoter region, which extends to position −279 relative to the start site of transcription and includes the BSAP binding site and all other sequences required for full promoter activity in B and T cell lines (23). The other constructs contained sequences 0 to −8700 (13NN), 0 to −2000 (13SC and 13SN), and 0 to −279 + −2000 to −6700 (13RS). The 13NN and 13SN clones also contained sequences extending 2700 bp 3′ of the *RAG2* gene, while the 13SC, 13HC, and 13RS constructs contained sequences extending 300 bp 3′ of *RAG2*. To simplify discussion, we refer to the sequences extending from the *Hin*dIII site to the *Stu*I site (−279 to −2000 bp) as the 2-kb "SH" region, sequences from the *Stu*I site to the *Eco*RI site (−2000 to −6700 bp) as the 2- to 7-kb "RS" region, and sequences from the *Eco*RI site to the *Not*I site (−6700 to −8700 bp) as the 7- to 9-kb "NR" region. In preliminary studies, we demonstrated that each of these *RAG2* genomic clones (including 13HC) was capable of rescuing the ability of $RAG2^{-/-}$ Abelson murine leukemia virus-transformed pro-B cell lines to undergo rearrangements of their endogenous IgH chain loci when introduced by stable transfection (data not shown).

We first transfected $R2^{BL/5.1}$ ES cells with the 13NN (which contains the most 5′ flanking sequence of tested clones) and the 13HC (which contains the least) genomic fragments and isolated ES clones carrying similar copy numbers ($\approx$2–6) of the constructs (Fig. 2A). We then generated chimeric mice with the 13NN and 13HC clones by R2DBC; at least two independent ES clones were tested for each construct. The 13NN clones restored the development of both B and T cells, while the 13HC clones failed to rescue development of either lineage (Fig. 2B). FACS analysis of BM from 13NN chimeric mice demonstrated significant numbers of $sIgM^-$ pre-B, $sIgM^+B220^{low}$ immature B, and $sIgM^+B220^{high}$ mature B cells. In contrast, 13HC chimeric mice exhibited no detectable $sIgM^+$ B cells in the BM. Analysis of 13NN thymus showed large numbers of DP and SP cells; total thymocyte numbers approached those of wild-type mice (up to 10^8) in several chimeras. In contrast, we observed no rescue of thymocyte development in 13HC mice. In spleen, we found significant numbers of both B and T cells in 13NN mice, but again no mature lymphocytes in 13HC chimeras. As chimeras generated from three independent 13HC clones exhibited significant numbers of $Ly5.1^+$ leukocytes, the observed absence of lymphocyte rescue is not likely to have resulted from poor ES cell contribution.

Sequences 5′ of the *RAG2* Gene Sufficient to Rescue B Cell, but Not T Cell, Development. Having identified a *RAG2* genomic clone that could rescue development of both B and T cell lineages, we sought to define elements within this clone potentially involved in lineage-specific regulation of *RAG2*. For this purpose, we transfected $R2^{BL/5.1}$ ES cells with the 13SN, 13SC, and 13RS genomic fragments (Fig. 2A); isolated ES clones carrying similar copy numbers ($\approx$2–6) of the constructs; and generated chimeric mice with these clones by R2DBC. At least two independent ES clones for each construct were tested. The transfected 13SN, 13SC, and 13RS clones gave rise to significant numbers of $sIgM^+$ BM and splenic B cells in $RAG2^{-/-}$ chimeras, but yielded few
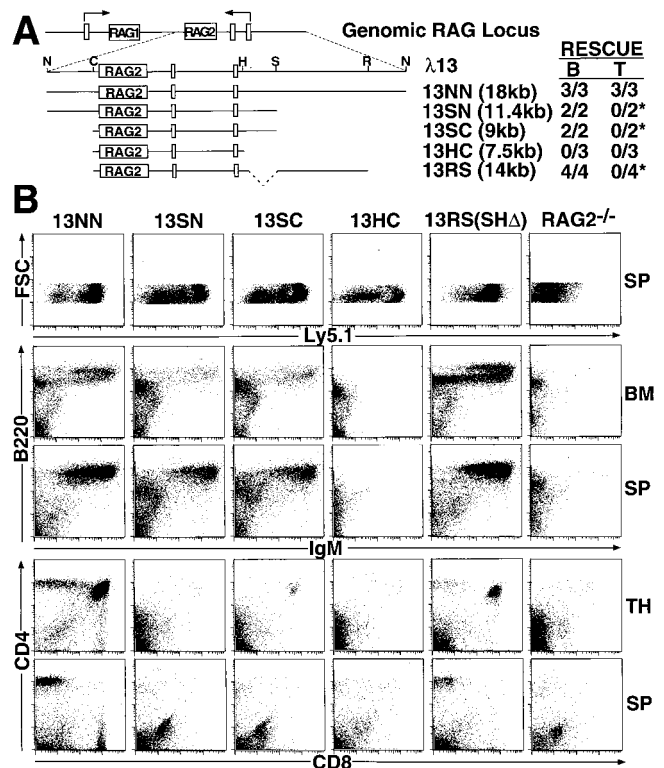

Fig. 2. Differential rescue of B and T cell development in $RAG2^{-/-}$ chimeric mice with *RAG2* genomic fragments. (A) Schematic of the λ13 genomic clone containing the *RAG2* gene in relation to the genomic *RAG* locus. Restriction sites used for the generation of truncations are shown (N, *Not*I; C, *Cla*I; H, *Hin*dIII; S, *Stu*I; R, *Eco*RI). Indicated below are the complete genomic fragment (13NN) and various truncations (13SN, 13SC, 13HC, and 13RS) tested in this study (see text for details). A summary table is shown to the right of the constructs to indicate the number of independent $R2^{BL/5.1}$ ES clones that rescued B and T cell development for each construct; "*" indicates the presence of a small number of DP/SP thymocytes or peripheral T cells in these chimeras, but minimal overall T cell rescue. (B) $R2^{BL/5.1}$ ES cells transfected with the 13NN, 13SN, 13SC, 13HC, or 13RS genomic fragments were assayed by R2DBC. Bone marrow (BM), thymus (TH), and spleen (SP) cells isolated from the various chimeras and $RAG2^{-/-}$ controls were stained with (*i*) FITC-anti-Ly5.1, (*ii*) PE-anti-IgM and CyC-anti-B220, or (*iii*) PE-anti-CD4 and CyC-anti-CD8 and analyzed by FACS. Data shown are from representative chimeric mice.

thymocytes that had progressed beyond the DN stage and few peripheral T cells (Fig. 2B). In some of these chimeras, we observed a small number of DP/SP thymocytes or SP splenic or lymph node T cells, suggesting that the 13SN, 13SC, and 13RS fragments may allow for very low or leaky levels of *RAG2* expression in developing T cells.

Among tested clones, efficient T cell rescue in our system was observed only with the complete 13NN genomic fragment. On the other hand, both the 2-kb SH and 2- to 7-kb RS regions upstream of the minimal promoter independently drove sufficient levels of *RAG2* expression to rescue development of substantial B cell numbers. In our experience with the R2DBC assay, we have found that, when there is a difference, the vast majority of ES cells are able to reconstitute T cell development more efficiently than B cell development (31). In the absence of T lineage-specific mutations, we have rarely observed chimeras generated by R2DBC that had significant numbers of B cells in the absence of T cells. For this reason, the ability of certain genomic *RAG2* fragments to reconstitute B cells, in the absence of significant T cell reconstitution, strongly argues for the existence of elements within these fragments that play an important role in *RAG2* gene regulation.
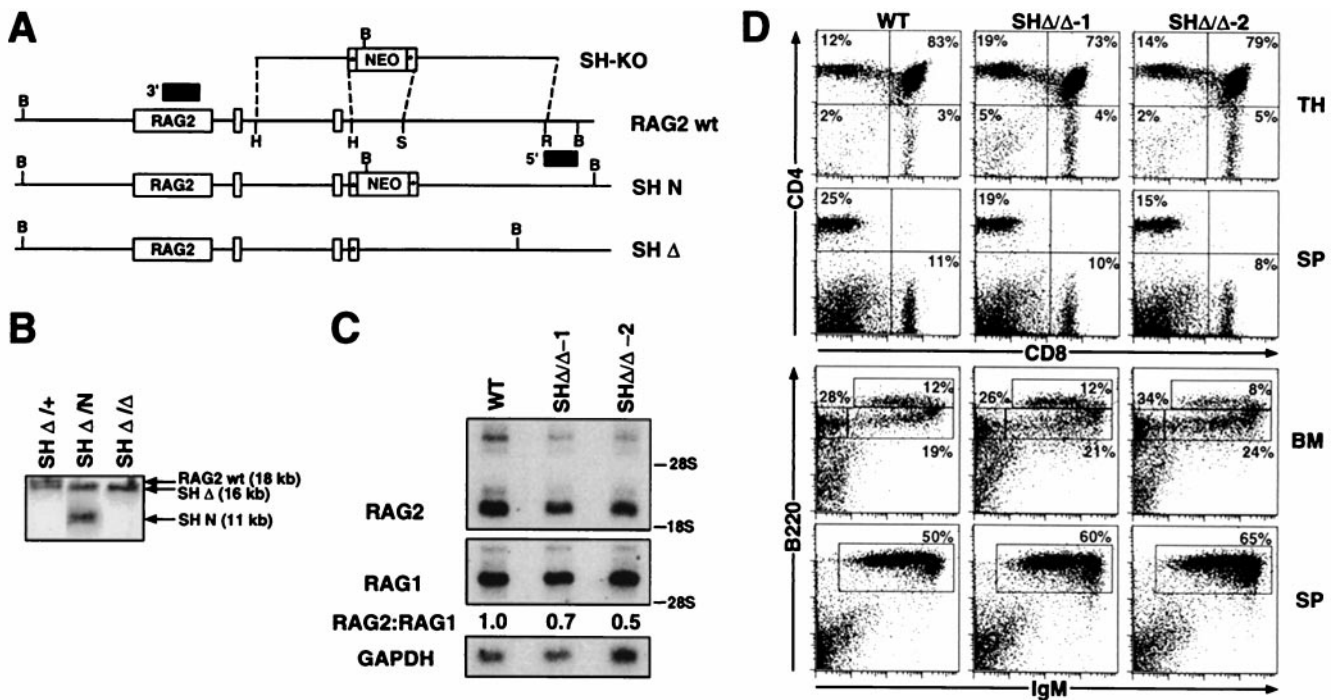
Fig. 3. SHΔ/Δ chimeric mice exhibit normal B and T cell development and a minor reduction in thymic expression of *RAG2*. (*A*) A targeting vector (SH-KO) was designed to delete the endogenous *Stu*I–*Hind*III (2-kb SH) region located 5′ of the *RAG2* gene, employing sequences 5′ and 3′ of the *Stu*I and *Hind*III sites to direct homologous recombination and a *loxP*-flanked *PGK-neo*[r] gene. Also indicated are *Bam*HI (B), *Eco*RI (R), *Stu*I (S), and *Hind*III (H) restriction sites; *loxP* sites (starred boxes); and probes used for analysis of 5′ and 3′ homologous recombination (shaded boxes). The endogenous *RAG2* allele (RAG2 wt) is depicted with the SH-KO vector and alleles in which the SH region has been replaced with either the *PGK-neo*[r] gene (SH N) or a single *loxP* site (SH Δ), following Cre-mediated deletion. (*B*) Southern blot analysis with the 3′ probe on *Bam*HI-digested DNA isolated from (*i*) ES cells in which the SH region had been targeted and replaced on one allele with a *loxP* site (SH Δ/+), (*ii*) SH Δ/+ ES cells targeted on the second allele with the *PGK-neo*[r] gene (SH Δ/N), and (*iii*) SH Δ/N ES cells following Cre-mediated deletion of the *PGK-neo*[r] gene (SH Δ/Δ). (C) Northern blot analysis of thymocyte RNA isolated from wild-type mice and SHΔ/Δ:RAG2[−/−] chimeric mice for the presence of RAG2 transcripts. Following hybridization with a RAG2 cDNA probe, the blot was stripped and rehybridized with a RAG1 cDNA probe to normalize levels of thymocyte RNA. The relative ratios of RAG2 to RAG1 hybridizing transcripts are shown. Hybridization with a GAPDH probe is included as an RNA loading control. (*D*) Thymus (TH), BM, and splenic (SP) cells from wild-type (WT) and SHΔ/Δ:RAG2[−/−] chimeric mice (generated from two independent SHΔ/Δ ES clones) were stained with (*i*) PE-anti-CD4 and CyC-anti-CD8, (*ii*) PE-anti-CD43 and CyC-anti-B220, or (*iii*) PE-anti-IgM and CyC-anti-B220 and analyzed by FACS. The percentages of total lymphocytes in various gated populations are shown from analysis of representative chimeras and WT mice.

Targeted Mutation of the 2-kb SH Region. In light of our findings that the 2-kb SH region, in combination with the *RAG2* promoter, drove sufficient levels of *RAG2* to rescue B cell development, we sought to determine its role in the endogenous regulation of *RAG2* by gene-targeted mutation. We designed a targeting construct (SH-KO) to replace the 2-kb SH region with a *loxP*-flanked *PGK-neo*[r] gene, allowing for subsequent deletion of the marker gene from targeted loci by transient expression of the Cre recombinase ("Cre-deletion"). We first generated a double knock-out ES cell line in which the SH region was deleted and replaced with a single *loxP* site on both alleles by means of sequential targeting of the SH region and Cre-deletion of marker sequences on each allele (Fig. 3 *A* and *B*). In parallel, we targeted an ES cell line in which *RAG2* had been replaced by the *RAG2:GFP* fusion gene on one allele (Fig. 4 *A* and *B*; RAG2[GFP/+]) (7). ES clones in which the SH region had been replaced with the *PGK-neo*[r] gene on the *RAG2:GFP* allele were identified (SH[N(GFP)/+]; Fig. 4*B*) and subjected to Cre-deletion. The resulting ES cells (SH[Δ(GFP)/+]; Fig. 4*B*) carried one wild-type *RAG2* allele and one *RAG2* allele in which the SH region was deleted upstream of the previously knocked-in *RAG2:GFP* fusion gene.

SH[Δ/Δ] and SH[Δ(GFP)/+] ES clones were both used to generate chimeric mice by R2DBC. We analyzed lymphocyte development in SH[Δ/Δ] chimeras generated from two independent ES clones. Both SH[Δ/Δ]-1 and SH[Δ/Δ]-2 chimeras exhibited completely normal numbers of B and T cells as assessed by FACS analysis

of BM, thymus, and spleen (Fig. 3*D*). Specifically, SH[Δ/Δ] chimeras showed normal thymic cellularity and distribution of DP and SP thymocyte subsets, in addition to normal numbers of CD4[+] and CD8[+] splenic T cells (Fig. 3*D*). B cell development in SH[Δ/Δ] chimeras was also equivalent to wild-type controls, as demonstrated by full reconstitution of CD43[low]B220[low]/sIgM[−]B220[low]pre-B, sIgM[+]B220[low] immature B, and sIgM[+]B220[high] mature B cell subsets in the BM, and normal numbers of splenic B cells (Fig. 3*D*). We next analyzed total thymus RNA from SH[Δ/Δ] and wild-type mice for expression of *RAG2* (Fig. 3*C*). When normalized for the levels of RAG1- or GAPDH-hybridizing transcripts, the levels of RAG2 transcripts in SH[Δ/Δ] thymocytes were only slightly reduced relative to those in wild-type thymocytes (≈1.5- to 2-fold).

SHΔ-RAG2:GFP chimeric mice exhibited normal B and T cell development (Fig. 4). We analyzed RAG2:GFP expression by FACS in B and T cell subsets isolated from SHΔ-RAG2:GFP chimeras and similar chimeras generated from the parental "wild-type" RAG2:GFP ES cell line. Expression of RAG2:GFP was slightly reduced (≈1.5- to 2-fold) but not absent in DP thymocytes from SHΔ-RAG2:GFP chimeras, as compared with those from wild-type RAG2:GFP chimeras (Fig. 4*C*). This modest decrease in RAG2:GFP protein expression closely parallels the decrease in RAG2 mRNA levels found in SH[Δ/Δ] thymocytes. Significantly, deletion of the SH region did not lead to prolonged or disregulated RAG2:GFP expression in SP thymocytes or mature T cells. In B lineage cells, RAG2:GFP
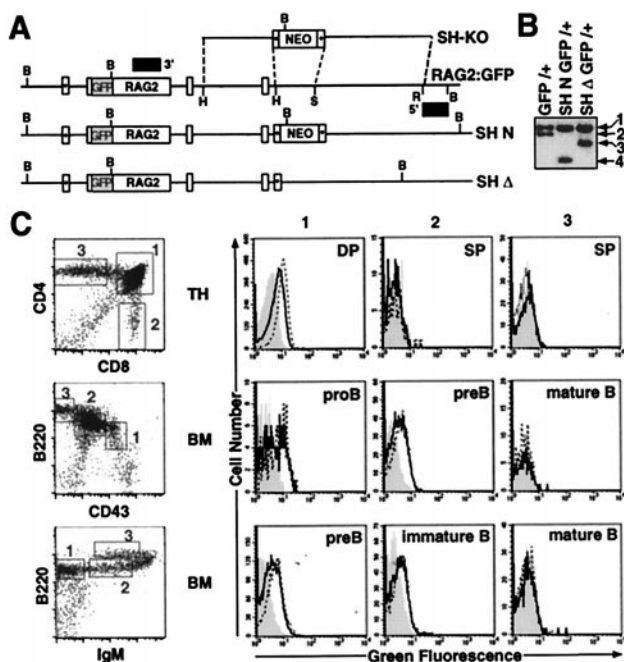
Fig. 4. RAG2:GFP expression in SHΔ-RAG2:GFP chimeric mice. (*A*) Targeting of the *RAG2:GFP* ES cell line with the SH-KO vector (as described in the legend of Fig. 3*A*). The *RAG2:GFP* knock-in allele (RAG2:GFP) is depicted with the SH-KO vector and alleles in which the SH region has been replaced with either the *PGK-neo*[r] gene (SH N) or a single *loxP* site (SH Δ), following Cre-mediated deletion. (*B*) Southern blot analysis with the 3′ probe on *Bam*HI-digested DNA isolated from (*i*) *RAG2:GFP* knock-in ES cells (GFP/+), (*ii*) GFP/+ ES cells in which the SH region was targeted and replaced on the *GFP* allele with the *PGK-neo*[r] gene (SH N GFP/+), and (iii) SH N GFP/+ ES cells following Cre-mediated deletion of the *PGK-neo*[r] gene (SH Δ GFP/+). 1, *RAG2* wt allele (18 kb); 2, *RAG2:GFP* allele (15 kb); 3, SHΔ GFP allele (11 kb); 4, SH N GFP allele (7 kb). (*C*) Thymocytes (TH) and BM from wild-type mice and from RAG2:GFP and SHΔ-RAG2:GFP chimeric mice were stained with (*i*) PE-anti-CD4 and CyC-anti-CD8, (*ii*) PE-anti-CD43 and CyC-anti-B220, or (*iii*) PE-anti-IgM and CyC-anti-B220 and analyzed by FACS. Representative FACS plots are shown to illustrate gated populations, as there were no differences in development between wild-type and chimeric mice. For each of the gated TH and BM populations (as numbered), histograms for green fluorescence were generated. WT histograms (shaded) were overlayed with RAG2:GFP histograms (open dashed-line) and SHΔ-RAG2:GFP histograms (open solid line). Developmental T and B cell stages are indicated for each set of histograms.

expression was similar in CD43^high B220^low pro-B cells isolated from SHΔ and wild-type RAG2:GFP chimeras, and only slightly reduced, on average, in sIgM^− B220^low pre-B cells (Fig. 4*C*). RAG2:GFP levels were essentially equivalent in SHΔ and wild-type RAG2:GFP sIgM^+ B220^low immature B cells in the BM, which may reflect the relative stability of the RAG2 protein in this population (3). Finally, the SH region deletion also did not result in aberrant RAG2:GFP expression in mature B cell subsets. We conclude that the SH region is not required for endogenous expression of *RAG2* in developing B or T cells or for apparently normal lymphocyte development, but that the region may be necessary for maximal expression of the gene in both lineages.

Discussion

The *RAG2* Promoter Does Not Support Expression of *RAG2 in Vivo*. Recent transient transfection experiments indicated that the *RAG2* promoter is lymphoid-specific and differentially regulated in T and B cell lines (23). However, with respect to our *in vivo* developmental assay, a genomic fragment containing the *RAG2* promoter but no further 5′ sequences was not capable of driving

RAG2 expression in B or T cell precursors sufficient to rescue lymphocyte development. In contrast, the identical genomic fragment did support rescue of IgH chain gene rearrangement in *RAG2*-deficient pro-B cell lines (data not shown). We conclude that the *RAG2* promoter alone is not sufficient to drive lineage- or developmental stage-specific regulation of *RAG2* gene expression *in vivo*. Rather, the promoter appears to require cooperative interactions with other cis-elements in the locus to achieve this function. Previous studies have also indicated that gene regulation *in vivo* may require higher orders of control than expression after transfection into cell lines *in vitro* (32).

Redundant Elements 5′ of *RAG2* Direct B Cell Lineage Expression. Significant reconstitution of precursor and mature B cell, but not T cell, compartments was achieved by addition of either the 2-kb SH or the more distal 2- to 7-kb RS *RAG2* upstream regions. These findings suggest the existence of at least two upstream elements, with potentially overlapping or redundant functions, that contribute to *RAG2* regulation in developing B cells. While the precise function of these putative elements remains to be determined, one of the possible explanations for B cell-specificity would be that elements within the 2-kb SH and 2- to 7-kb RS regions might cooperate with B cell-specific elements in the core promoter, such as the BSAP binding site (23). Although the construct complementation analyses clearly suggested that the 2-kb SH region harbors a B lineage-specific element, its germ-line deletion led to only a very modest reduction of *RAG2* expression, in T lineage as well as B lineage cells. A likely explanation for the difference between the transfection and knock-out results is the existence of an element that *in vivo* is functionally redundant to an element within the 2-kb SH region (e.g., within the 2- to 7-kb RS region). Similar findings have been obtained regarding potential redundant functions of several other enhancers, including the Ig κ intronic and 3′ enhancers, 3′ IgH locus enhancers, and 5′ β-globin locus enhancers (27, 33–36).

Different Sequences 5′ of *RAG2* May Be Required for Expression of *RAG2* in B and T Cell Lineages. It is striking that sequences upstream of the *RAG2* gene can drive expression sufficient for development of normal B cell numbers in the absence of significant T cell numbers, inferring the potential existence of T cell-specific *RAG2* regulatory sequences. The nature of these putative elements, as well as the mechanism by which they function, remain to be determined. Rescue of T cell development might be accomplished by elements that raise RAG2 levels in both B and T cells, with a higher level required for T vs. B cell antigen receptor gene rearrangement. Alternatively, T cell rescue might be accomplished by lineage-specific *RAG2* regulatory elements. Regulation of gene expression by distinct elements in different lineages previously has been observed in the context of CD8α gene regulation by different enhancer elements in thymus-dependent vs. thymus-independent T cell lineages (37, 38). We speculate that regulation of *RAG* expression in B and T lineage cells involves a set of common elements that function in a lineage-independent fashion, as well as a set of distinct elements that function in a lineage-dependent fashion. In this context, our data suggest a more 5′ location of *RAG2* regulatory elements required to support T cell development, as addition of sequences just upstream (e.g., 7- to 9-kb NR region) of those which drive development of normal B cell numbers leads to development of normal T cell numbers. We note, however, that other elements (for example, locus control regions or enhancers) also may exist elsewhere in the endogenous *RAG* locus and contribute to the overall regulation of *RAG2* expression.

Our findings are generally in agreement with those of a parallel study of *RAG* gene regulation in transgenic mice carrying bacterial artificial chromosomes (BACs) engineered to

express GFP in place of RAG2 and yellow fluorescent protein (YFP) in place of RAG1 (39). In those mice, a 10-kb region 5′ of *RAG2* drove developmentally regulated expression of GFP and YFP in B lineage precursors, supporting the notion that sequences 5′ of *RAG2* are involved in the regulation of *RAG2* and *RAG1*. The BAC studies also suggested B vs. T cell-specific regulatory regions 5′ of *RAG2*. In particular, they showed that sequences sufficient for high level *RAG2* expression in DN thymocytes were located within the 10-kb region 5′ of *RAG2*, but that those required for high level *RAG2* expression in DP thymocytes were located further upstream (39). There are several possible explanations for the apparent discrepancy between the latter result and our finding that a construct containing 9 kb of sequences 5′ of *RAG2* supports normal T cell development. First, lower levels of *RAG2* expression may be sufficient to drive essentially normal levels of TCRα rearrangement in DP T cells. Alternatively, there could be a high level of *RAG2* expression in a few DP cells, allowing for TCRα rearrangement and subsequent differentiation of these cells into SP thymocytes and peripheral T cells. In the latter context, it will be important to examine both the B and T cell repertoires generated by the various *RAG2* constructs.

Potential Influence of *RAG* Expression Levels on Development and Disease. As the precise levels of *RAG* expression required for normal V(D)J recombination during lymphocyte development are not known, it is possible that the lymphocyte reconstitution observed with some of our genomic fragments does not reflect wild-type levels of *RAG2* expression. In this regard, certain V(D)J recombination deficiencies lead to imbalances in T vs. B cell numbers. For example, Omenn syndrome (OS) patients with mutations in *RAG1* and *RAG2* coding sequences that compromise recombinase function have variable numbers of mature T cells in the absence of B cells (40). In this context, our studies show that mutations in *RAG2* regulatory regions also have the potential to lead to altered ratios of B vs. T cells and theoretically could contribute to diseases such as OS. A precedent for regulatory mutations resulting in clinical disease is provided by four naturally occurring deletions spanning the β-globin locus control region, which produce the classical hematologic features of β-thalassemia in the absence of mutations within the β-globin gene itself (41).

1. Willerford, D. M., Swat, W. & Alt, F. W. (1996) *Curr. Opin. Genet. Dev.* **6,** 603–609.
2. Gellert, M. (1997) *Adv. Immunol.* **64,** 39–64.
3. Grawunder, U., Leu, T. M., Schatz, D. G., Werner, A., Rolink, A. G., Melchers, F. & Winkler, T. H. (1995) *Immunity* **3,** 601–608.
4. Ma, A., Fisher, P., Dildrop, R., Oltz, E., Rathbun, G., Achacoso, P., Stall, A. & Alt, F. W. (1992) *EMBO J.* **11,** 2727–2734.
5. Li, Y. S., Hayakawa, K. & Hardy, R. R. (1993) *J. Exp. Med.* **178,** 951–960.
6. Hertz, M. & Nemazee, D. (1998) *Curr. Opin. Immunol.* **10,** 208–213.
7. Monroe, R. J., Seidl, K. J., Gaertner, F., Han, S., Chen, F., Sekiguchi, J., Wang, J., Ferrini, R., Davidson, L., Kelsoe, G. & Alt, F. W. (1999) *Immunity* **11,** 201–212.
8. Wilson, A., Held, W. & MacDonald, H. R. (1994) *J. Exp. Med.* **179,** 1355–1360.
9. Ismaili, J., Antica, M. & Wu, L. (1996) *Eur. J. Immunol.* **26,** 731–737.
10. Fehling, H. J. & von Boehmer, H. (1997) *Curr. Opin. Immunol.* **9,** 263–275.
11. Guy Grand, D., Vanden Broecke, C., Briottet, C., Malassis-Seris, M., Selz, F. & Vassalli, P. (1992) *Eur. J. Immunol.* **22,** 505–510.
12. Turka, L. A., Schatz, D. G., Oettinger, M. A., Chun, J. J., Gorka, C., Lee, K., McCormack, W. T. & Thompson, C. B. (1991) *Science* **253,** 778–781.
13. Han, S., Zheng, B., Takahashi, Y. & Kelsoe, G. (1997) *Semin. Immunol.* **9,** 255–260.
14. Qin, X. F., Schwers, S., Yu, W., Papavasiliou, F., Suh, H., Nussenzweig, A., Rajewsky, K. & Nussenzweig, M. C. (1999) *Nature (London)* **397,** 355–359.
15. McMahan, C. J. & Fink, P. J. (1998) *Immunity* **9,** 637–647.
16. Chun, J. J., Schatz, D. G., Oettinger, M. A., Jaenisch, R. & Baltimore, D. (1991) *Cell* **64,** 189–200.
17. Carlson, L. M., Oettinger, M. A., Schatz, D. G., Masteller, E. L., Hurley, E. A., McCormack, W. T., Baltimore, D. & Thompson, C. B. (1991) *Cell* **64,** 201–208.
18. Oettinger, M. A., Schatz, D. G., Gorka, C. & Baltimore, D. (1990) *Science* **248,** 1517–1523.
19. Thompson, C. B. (1995) *Immunity* **3,** 531–539.
20. Agrawal, A., Eastman, Q. M. & Schatz, D. G. (1998) *Nature (London)* **394,** 744–751.
21. Hiom, K., Melek, M. & Gellert, M. (1998) *Cell* **94,** 463–470.
22. Zarrin, A. A., Fong, I., Malkin, L., Marsden, P. A. & Berinstein, N. L. (1997) *J. Immunol.* **159,** 4382–4394.
23. Lauring, J. & Schlissel, M. S. (1999) *Mol. Cell. Biol.* **19,** 2601–2612.
24. Brown, S. T., Miranda, G. A., Galic, Z., Hartman, I. Z., Lyon, C. J. & Aguilera, R. J. (1997) *J. Immunol.* **158,** 5071–5074.
25. Fuller, K. & Storb, U. (1997) *Mol. Immunol.* **34,** 939–954.
26. Swat, W., Shinkai, Y., Cheng, H. L., Davidson, L. & Alt, F. W. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 4683–4687.
27. Gorman, J. R., van der Stoep, N., Monroe, R., Cogne, M., Davidson, L. & Alt, F. W. (1996) *Immunity* **5,** 241–252.
28. Chen, J., Lansford, R., Stewart, V., Young, F. & Alt, F. W. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 4528–4532.
29. Chen, J. (1996) *Adv. Immunol.* **62,** 31–59.
30. Gartner, F., Alt, F. W., Monroe, R., Chu, M., Sleckman, B. P., Davidson, L. & Swat, W. (1999) *Immunity* **10,** 537–546.
31. Chen, J. & Alt, F. W. (1994) in *Trangenesis and Targeted Mutagenesis in Immunology*, eds. Bluethmann, H. & Ohashi, P. S. (Academic, San Diego), pp. 35–49.
32. Zimmerman, K., Legouy, E., Stewart, V., Depinho, R. & Alt, F. W. (1990) *Mol. Cell. Biol.* **10,** 2096–2103.
33. Xu, Y., Davidson, L., Alt, F. W. & Baltimore, D. (1996) *Immunity* **4,** 377–385.
34. Manis, J. P., van der Stoep, N., Tian, M., Ferrini, R., Davidson, L., Bottaro, A. & Alt, F. W. (1998) *J. Exp. Med.* **188,** 1421–1431.
35. Fiering, S., Epner, E., Robinson, K., Zhuang, Y., Telling, A., Hu, M., Martin, D. I., Enver, T., Ley, T. J. & Groudine, M. (1995) *Genes Dev.* **9,** 2203–2213.
36. Epner, E., Reik, A., Cimbora, D., Telling, A., Bender, M. A., Fiering, S., Enver, T., Martin, D. I., Kennedy, M., Keller, G. & Groudine, M. (1998) *Mol. Cell* **2,** 447–455.
37. Hostert, A., Garefalaki, A., Mavria, G., Tolaini, M., Roderick, K., Norton, T., Mee, P. J., Tybulewicz, V. L., Coles, M. & Kioussis, D. (1998) *Immunity* **9,** 497–508.
38. Ellmeier, W., Sunshine, M. J., Losos, K. & Littman, D. R. (1998) *Immunity* **9,** 485–496.
39. Yu, W., Misulovin, Z., Suh, H., Hardy, R. R., Jankovic, M., Yannoutsos, N. & Nussenzweig, M. C. (1999) *Science* **285,** 1080–1084.
40. Villa, A., Santagata, S., Bozzi, F., Giliani, S., Frattini, A., Imberti, L., Gatta, L. B., Ochs, H. D., Schwarz, K., Notarangelo, L. D., *et al.* (1998) *Cell* **93,** 885–896.
41. Higgs, D. R. (1998) *Cell* **95,** 299–302.