Brief report

IgH partner breakpoint sequences provide evidence that AID initiates t(11;14) and t(8;14) chromosomal breaks in mantle cell and Burkitt lymphomas

Harvey A. Greisman,[1] Zhengfei Lu,[2-5] Albert G. Tsai,[2-5] Timothy C. Greiner,[6] Hye Son Yi,[1] and Michael R. Lieber[2-5]

[1]Department of Laboratory Medicine, University of Washington, Seattle, WA; Departments of [2]Pathology, [3]Biochemistry & Molecular Biology, [4]Molecular Microbiology & Immunology, and [5]Molecular & Computational Biology, USC Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA; and [6]Department of Pathology and Microbiology, University of Nebraska, Omaha, NE

Previous studies have implicated activation-induced cytidine deaminase (AID) in B-cell translocations but have failed to identify any association between their chromosomal breakpoints and known AID target sequences. Analysis of 56 unclustered IgH-CCND1 translocations in mantle cell lymphoma across the ∼ 344-kb bcl-1 breakpoint locus demonstrates that half of the *CCND1* breaks are near CpG dinucleotides. Most of these CpG breaks are at CGC motifs, and half of the remaining breaks are near WGCW, both known AID targets. These findings provide the strongest evidence to date that AID initiates chromosomal breaks in translocations that occur in human bone marrow B-cell progenitors. We also identify WGCW breaks at the *MYC* locus in Burkitt lymphoma translocations and murine IgH-MYC translocations, both of which arise in mature germinal center B cells. Finally, we propose a developmental model to explain the transition from CpG breaks in early human B-cell progenitors to WGCW breaks in later stage B cells. (*Blood*. 2012;120(14):2864-2867)

Introduction

Most human lymphoma translocations, including the t(14;18)/IgH-BCL2 in follicular lymphoma and the t(11;14)/IgH-CCND1 in mantle cell lymphoma (MCL), appear to arise in B-cell progenitors in the bone marrow and to depend on the RAG complex for chromosome breaks at J_H and D_H segments in the immunoglobulin heavy chain (*IgH*) locus. Almost nothing is known about the chromosomal breakage process at *BCL2*, *CCND1*, or other *IgH* partner loci because no murine or other animal models currently exist for this type of translocation. Recently, we showed that many such breakpoints are specifically targeted to the dinucleotide sequence CpG,[1,2] and we proposed a chromosome breakage model in which activation-induced cytidine deaminase (AID) creates stable T:G mismatches at methylated CpG sites, followed by RAG cutting of repair intermediates.[1] Here we analyze a large set of "unclustered" t(11;14) breakpoints (ie, those that fall outside of the *CCND1* major translocation cluster [MTC]), and we show that most of these breaks occur at or near known AID hotspots.

Methods

Samples were obtained from the University of Washington Hematopathology Laboratory (22 cases) and the University of Nebraska (10 cases) after approval by the University of Washington Institutional Review Board. t(11;14) breakpoints were mapped using translocation comparative genomic hybridization (translocation CGH).[3] Breakpoints were amplified and Sanger sequenced using patient-specific primers (available on request). Statistical methods used to analyze proximity to CpG, and WGCW motifs were described by Tsai et al[1] and in supplemental Methods and supplemental Results (available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article).

Results and discussion

We used a modification of array-based CGH called translocation CGH[3] to identify and clone 56 unclustered t(11;14) breakpoints in 32 cases of MCL. A total of 55 of the 56 breakpoints (98%) at the *IgH* locus on chromosome 14q32 mapped to a J_H, D_H, or V_H gene segment and showed features consistent with RAG-dependent breakage during V(D)J recombination (see Table 1 and supplemental Tables 1-4, and supplemental Figure 2). The breakpoints on chromosome 11q13 span 329 kb of the 344-kb bcl-1 breakpoint region between *CCND1* and *MYEOV*, the closest gene centromeric to *CCND1* (Figure 1A). The most centromeric breakpoint was located 12 kb from *MYEOV*, 331 kb from *CCND1*, and > 200 kb from the most distant known breakpoint (Table 1).[4,5] The most telomeric breakpoint was only 2 kb from the *CCND1* gene. Previous studies[7-9] have suggested that the majority of t(11;14) breakpoints are telomeric to the MTC, but 12 (37%) of our 32 cases had breakpoints centromeric to the MTC. Nevertheless, the other 20 cases (63%) had breakpoints telomeric to the MTC, significantly more than would be expected if the breakpoints were equally distributed across the entire bcl-1 region ($P = .0003$, binomial test).

In 16 of 32 cases (50%), one or both *CCND1* breakpoints were ≤ 4 nt from a CpG dinucleotide (referred to below as CpG breakpoints), including 10 cases (31%) with breakpoints directly at a CpG. The majority of the CpG breakpoints (10 of 16, 63%) were located at CGC (Table 1), a known AID hotspot that is distinctive for not conforming to the consensus WRC motif for AID-dependent DNA deamination.[10,11] In the remaining 16 cases (50%), the *CCND1* breakpoints were ≥ 10 nucleotides (nt) from the

Figure 1. Distribution and features of unclustered *CCND1* breakpoints and a model for CpG and WGCW breaks in *IgH* translocations. (A) A 400-kb segment of chromosome 11q13 (chr11:68 800 000-69 200 000 in NCBI Build 36) is shown. The *MYEOV* and *CCND1* genes and MTC are shown for reference. CpG breakpoints are shown above the axis, and non-CpG breakpoints are shown below. Blue represents κ-restricted cases; red, λ-restricted cases; and ○, cases for which light chain restriction is unknown. (B) Our results implicate AID-initiated chromosomal breaks in the genesis of IgH-CCND1 translocations in MCL, IgH-MYC translocations in Burkitt lymphoma, and IgH-BCL6 translocations in follicular and diffuse large B-cell lymphomas (see supplemental Results). AID expression is highest in antigen-stimulated mature B cells undergoing somatic hypermutation and CSR but is expressed at lower levels in immature and transitional 1 (T1) B cells, and at even lower levels in pro-B and pre-B cells. RAG activity is associated with V(D)J recombination in pro-B cells during *IgH* rearrangement, in pre-B cells during light chain rearrangement, and in immature B cells during receptor editing. Lightning bolts indicate the stage at which each translocation is proposed to occur. Our model proposes that low levels of AID favor CpG breaks (blue font), whereas intermediate or high levels of AID favor WGCW breaks (red). The staggered cytosine bases on opposite strands of the CpG and WGCW motifs are highlighted to indicate the potential of this configuration to promote DSBs.
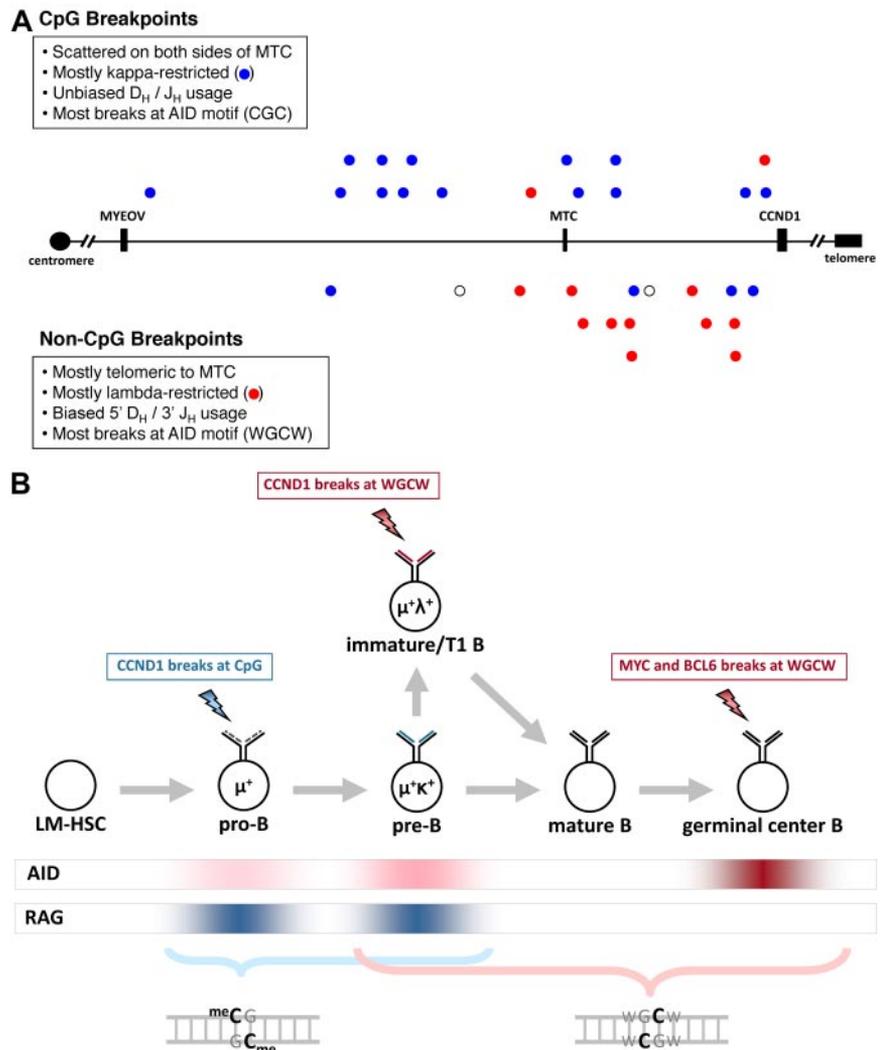
closest CpG ("non-CpG" breakpoints), including 5 cases for which der(11) breakpoint sequences were unavailable. *CCND1* breakpoints were not associated with any other dinucleotide or RAG motifs (supplemental Table 5A).

The CpG and non-CpG breaks were differentially distributed across the bcl-1 breakpoint region (Figure 1A). The breakpoints in 13 (81%) of the 16 non-CpG cases were telomeric to the MTC, significantly more than the 5 of 16 (32%) expected if equally distributed across the bcl-1 region ($P = .00006$, binomial test). In contrast, the CpG breakpoints appear to be randomly distributed across both centromeric and telomeric regions ($P = .2$, binomial). Thus, the overall predominance of telomeric bcl-1 breakpoints in our series can be explained by the non-CpG breaks alone. This differential distribution of CpG and non-CpG breakpoints suggested that these 2 types of breaks are biologically distinct.

Most B-cell lymphomas show a slight preference for immunoglobulin κ over λ light chain restriction, presumably reflecting the sequential rearrangement of the κ locus before the λ locus during normal B-cell development.[12,13] To explore the unusual λ restriction bias that has been observed for MCL,[14] we compared light chain restriction status to *CCND1* breakpoint type and location. Ten (71%) of the 14 non-CpG cases for which light chain data were available (Table 1) were λ-restricted compared with only 2 of 16 CpG cases (12%), indicating a strong association between

CpG status and light chain restriction ($P = .0022$, 2-tailed Fisher exact test).

Non-CpG breakpoints also were significantly more likely than CpG breakpoints to involve the most 3′ J_H segment (J_H6) and the most 5′ D_H segment (D_H2-2; supplemental Tables 2 and 4). J_H6 was rearranged in 8 of 14 non-CpG cases (57%) but in only 2 of 16 (11%) CpG cases ($P = .019$ by Fisher exact test). Similarly, D_H2-2 was rearranged in 4 of 10 non-CpG cases (40%) compared with 0 of 12 CpG cases ($P = .029$). Three of the 4 D_H2-2 cases had J_H6 rearranged on the der(14) chromosome and the remaining case had a D_H2-2 signal joint at the der(14) breakpoint (Table 1). Biased usage of 3′ J_H segments and 5′ D_H segments has been previously identified at t(14;18) breakpoints[15] but not for t(11;14) breakpoints at the MTC.[16] Taken together, the preferential use of λ light chain and the biased usage of 3′ J_H and 5′ D_H segments in non-CpG cases suggest that these rearrangements occur at a later stage of B-cell development than CpG breakpoints.

To explore whether AID is involved in generating the non-CpG breaks, we examined the proximity of these breaks to the sequence motif WGCW, where W = A or T. WGCW was recently shown to facilitate AID-dependent double-strand breaks (DSBs) during physiologic class-switch recombination (CSR) by providing 2 antiparallel AID hotspot motifs (WRC, where R = A or G).[17] Strikingly, 12 of the 27 non-CpG breaks at

Table 1. Characteristics of t(11;14) breakpoints

	Case	κ/λ	der(14) breakpoints						der(11) breakpoints	
			IgH	CCND1	Distance to CCND1	Nearest CpG	CGC?	Nearest WGCW	CCND1	IgH
CpG breakpoints	W7	κ	J_H5	68 833 760	331 294	1	Y	—	68 833 763	D_H4-17
	W8	κ	J_H6	68 935 608	229 446	0	N	—	68 935 595	D_H3-3
	W18	κ	J_H5	68 940 344	224 710	0	Y	—	68 940 341	D_H3-10
	W15	κ	J_H6	68 957 805*	207 249	0	Y	—	68 957 803	D_H3-3
	W19	κ	J_H4	68 957 805*	207 249	0	Y	—	—	—
	W17	κ	J_H4	68 969 141	195 913	0	Y	—	68 969 139	D_H1−26
	N1	κ	J_H4	68 973 670	191 384	0	Y	—	68 973 668	D_H3-9
	W13	κ	J_H5	68 989 831	175 223	2	Y	—	—	—
	N8	λ	J_H1	69 037 471†	127 583	0	N	—	69 037 466	D_H3-22
	W10	κ	J_H4	69 056 460	108 594	2	N	—	69 056 458	D_H2-21
	N6	κ	J_H4	69 062 733	102 321	1	N	—	69 062 741	D_H3-3
	W11	κ	J_H4	69 082 644	82 410	0	N	—	—	—
	W14	κ	J_H5	69 082 854‡	82 200	2	Y	—	69 082 840	D_H5-5
	W16	κ	J_H4	69 152 043	13 011	4	Y	—	69 152 034	V_H4-34
	N2	λ	J_H4	69 162 371	2 683	0	Y	—	69 162 368	D_H5-12
	W22	κ	J_H1	69 163 029	2 025	0	N	—	69 163 022	D_H3-10
Non-CpG breakpoints	W3	κ	J_H6	68 930 336	234 718	18	—	3	—	—
	N10	—	J_H6	68 999 340	165 714	29	—	14	68 999 340	D_H3-3
	N3	λ	J_H4	69 031 511	133 543	68	—	1	69 031 503	J_H4s§
	W6	λ	J_H6	69 059 199	105 855	98	—	4	—	—
	W4	λ	J_H5	69 065 258	99 796	60	—	16	69 065 237	D_H2-21
	W2	λ	J_H4	69 080 506	85 548	25	—	1	69 080 980	D_H4-23
	W12	λ	J_H6	69 090 218	74 836	167	—	7	69 090 208	D_H2-2
	N4	λ	J_H4	69 091 320	73 734	173	—	10	—	—
	N5	κ	J_H6	69 092 350	72 704	33	—	3	69 092 347	D_H3-3
	N9	—	D_H2-2s§	69 100 683	64 371	10	—	0	69 100 678	D_H2-2
	W5	λ	J_H6	69 123 555	41 499	15	—	12	—	—
	W1	λ	J_H4	69 131 130	33 924	207	—	41	69 131 129	D_H5-18
	W20	κ	J_H4/J_H5‖	69 144 576	20 478	86	—	28	—	—
	W9	λ	J_H4	69 146 211	18 843	106	—	3	69 146 211	D_H4-17
	N7	λ	J_H6	69 146 814	18 240	29	—	36	69 146 895	D_H2-2
	W21	κ	J_H6	69 156 220	8 834	242	—	0	69 156 210	D_H2-2

— indicates not applicable.
*CCND1 breaks at same CpG.
†CCND1 break at same CpG as patient HO.[4]
‡CCND1 break at same CpG as cell line MO1094.[6]
§V(D)J signal join (see supplemental Figures 1 and 2).
‖Non-V(D)J break between J_H4 and J_H5 (see supplemental Figure 1).

CCND1 are within 4 nt of a WGCW motif, making it the second-most frequent break site after CpG ($P = .02$; supplemental Figure 4; supplemental Table 5B).

Similarly, we analyzed a large database of 156 human and 299 murine IgH-MYC breakpoints and found a striking proximity of MYC breaks to WGCW in both human Burkitt lymphomas ($P < .0001$) and murine plasmacytomas ($P < .000001$). Because of the large degree of breakpoint heterogeneity around WGCW at both CCND1 and MYC loci, we included breaks occurring within 4 bp of WGCW (see supplemental Figure 4A-C). Neither CCND1 nor MYC breakpoints were associated with other 4-nt sequence motifs having equivalent degeneracy to WGCW ($P \geq .19$), including GWWC, CWWG, SATS, and STAS (S = G or C), each of which contains 2 G:C and 2 A:T base pairs and a central or peripheral palindromic dinucleotide pair (supplemental Table 5B-D).

The proximity of the CCND1 breaks to the known AID motifs CGC and WGCW provides strong and direct evidence that AID cooperates with RAG to initiate IgH-CCND1 translocations in B-cell progenitors. The palindromic nature of CpG and WGCW facilitates chromosome breakage at CCND1 by providing cytosine bases on both strands of the DNA duplex in almost direct opposition to one another (Figure 1B), thereby increasing the likelihood of a DSB.[17]

To explain why MYC breaks occur at WGCW but not at CpG[1] whereas CCND1 breaks can occur at either CpG or WGCW, we propose the following mechanistic and developmental model for IgH translocations (Figure 1B). CpG breaks are favored at the earliest stages of B-cell development when AID is expressed at very low levels[18-22] because deamination of methylated CpG gives rise to T:G mismatches that are repaired 2500-fold less efficiently than U:G mismatches at nonmethylated cytosines.[1] The T:G mismatches preferentially predispose to chromosome breaks because they persist for much longer than U:G mismatches, even though it is likely that both T:G and U:G lesions arise at very low levels.

MYC breaks occur at WGCW because this motif is favored at the very high level of AID found in germinal center B cells, as indicated by the strong preference for DSBs at WGCW motifs in the IgH locus during physiologic CSR.[17] We hypothesize that the non-CpG breaks in MCL might occur in immature/transitional 1 (T1) stage B cells, where AID is expressed at an intermediate level[18-22] (Figure 1B), presumably in conjunction with receptor editing.[22-24] This would explain their preference for WGCW motifs

as well as their D_H/J_H bias and λ light chain preference, in contrast to the unbiased D_H/J_H usage and κ predominance in CpG cases (Figure 1A).

Given the strong evidence linking AID to DNA lesions within actively transcribed genes,[25,26] it might seem surprising that AID is targeting CpG and WGCW sites that are so far from the *CCND1* gene (Figure 1A). However, several recent genome-wide studies of AID-induced mutations and chromosomal breaks in mature murine B cells[26-29] clearly show that AID lesions are not limited to genic regions. For example, Klein et al reported that 10%-20% of their AID-dependent translocation hotspots are intergenic.[28] Murine models[30] and other experimental systems to study chromosomal translocations in B-cell precursors are needed to better understand the mechanisms underlying these most common *IgH* translocations.

Authorship

Contribution: H.A.G. designed research; H.S.Y. collected data; H.A.G., A.G.T., and M.R.L. analyzed and interpreted data; Z.L. and A.G.T. performed statistical analyses; T.C.G. contributed critical reagents; and H.A.G. and M.R.L. wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

The current affiliation for A.G.T. is Department of Pathology, Stanford University, Stanford, CA.

Correspondence: Harvey A. Greisman, Department of Laboratory Medicine, University of Washington Medical Center, Box 357110, Seattle, WA 98195; e-mail: greisman@uw.edu.

References

1. Tsai AG, Lu H, Raghavan SC, Muschen M, Hsieh C-L, Lieber MR. Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity. *Cell*. 2008;135(6):1130-1142.

2. Tsai A, Lieber M. Mechanisms of chromosomal rearrangement in the human genome. *BMC Genom*. 2010;11(Suppl 1):S1.

3. Greisman HA, Hoffman NG, Yi HS. Rapid high-resolution mapping of balanced chromosomal rearrangements on tiling CGH arrays. *J Mol Diagn*. 2011;13(6):621-633.

4. Galiegue-Zouitina S, Collyn-d'Hooghe M, Denis C, et al. Molecular cloning of a t(11;14)(q13;q32) translocation breakpoint centromeric to the BCL1-MTC. *Genes Chromosomes Cancer*. 1994;11(4):246-255.

5. Degan M, Doliana R, Gloghini A, et al. A novel bcl-1/JH breakpoint from a patient affected by mantle cell lymphoma extends the major translocation cluster. *J Pathol*. 2002;197(2):256-263.

6. Meeker TC, Sellers W, Harvey R, et al. Cloning of the t(11;14)(q13;q32) translocation breakpoints from two human leukemia cell lines. *Leukemia*. 1991;5(9):733-737.

7. de Boer CJ, Loyson S, Kluin PM, Kluin-Nelemans HC, Schuuring E, van Krieken JH. Multiple breakpoints within the BCL-1 locus in B-cell lymphoma: rearrangements of the cyclin D1 gene. *Cancer Res*. 1993;53(18):4148-4152.

8. Williams ME, Swerdlow SH, Rosenberg CL, Arnold A. Characterization of chromosome 11 translocation breakpoints at the bcl-1 and PRAD1 loci in centrocytic lymphoma. *Cancer Res*. 1992; 52(19 Suppl):5541s-5544s.

9. Vaandrager JW, Schuuring E, Zwikstra E, et al. Direct visualization of dispersed 11q13 chromosomal translocations in mantle cell lymphoma by multicolor DNA fiber fluorescence in situ hybridization. *Blood*. 1996;88(4):1177-1182.

10. Pham P, Bransteitter R, Petruska J, Goodman MF. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature*. 2003;424(6944):103-107.

11. Yu K, Huang F-T, Lieber MR. DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine. *J Biol Chem*. 2004;279(8):6496-6500.

12. Korsmeyer SJ, Hieter PA, Sharrow SO, Goldman CK, Leder P, Waldmann TA. Normal human B cells display ordered light chain gene rearrangements and deletions. *J Exp Med*. 1982;156(4):975-985.

13. Bräuninger A, Goossens T, Rajewsky K, Küppers R. Regulation of immunoglobulin light chain gene rearrangements during early B cell development in the human. *Eur J Immunol*. 2001;31(12):3631-3637.

14. Swerdlow SH, Campo E, Harris NL, et al. *WHO Classification of Tumours of the Haematopoietic and Lymphoid Tissues* (4th Ed). Geneva, Switzerland: WHO Press; 2008.

15. Jäger U, Böcskör S, Le T, et al. Follicular lymphomas' BCL-2/IgH junctions contain templated nucleotide insertions: novel insights into the mechanism of t(14;18) translocation. *Blood*. 2000;95(11):3520-3529.

16. Welzel N, Le T, Marculescu R, et al. Templated nucleotide addition and immunoglobulin JH-gene utilization in t(11;14) junctions: implications for the mechanism of translocation and the origin of mantle cell lymphoma. *Cancer Res*. 2001;61(4):1629-1636.

17. Han L, Masani S, Yu K. Overlapping activation-induced cytidine deaminase hotspot motifs in Ig class-switch recombination. *Proc Natl Acad Sci U S A*. 2011;108(28):11584-11589.

18. Mao C, Jiang L, Melo-Jorge M, et al. T cell-independent somatic hypermutation in murine B cells with an immature phenotype. *Immunity*. 2004;20(2):133-144.

19. Han J-H, Akira S, Calame K, Beutler B, Selsing E, Imanishi-Kari T. Class switch recombination and somatic hypermutation in early mouse B cells are mediated by B cell and Toll-like receptors. *Immunity*. 2007;27(1):64-75.

20. Ueda Y, Liao D, Yang K, Patel A, Kelsoe G. T-independent activation-induced cytidine deaminase expression, class-switch recombination, and antibody production by immature/transitional 1 B cells. *J Immunol*. 2007;178(6):3593-3601.

21. Kuraoka M, Liao D, Yang K, et al. Activation-induced cytidine deaminase expression and activity in the absence of germinal centers: insights into hyper-IgM syndrome. *J Immunol*. 2009; 183(5):3237-3248.

22. Kuraoka M, Holl TM, Liao D, et al. Activation-induced cytidine deaminase mediates central tolerance in B cells. *Proc Natl Acad Sci U S A*. 2011; 108(28):11560-11565.

23. Meyers G, Ng Y-S, Bannock JM, et al. Activation-induced cytidine deaminase (AID) is required for B-cell tolerance in humans. *Proc Natl Acad Sci U S A*. 2011;108(28):11554-11559.

24. Nemazee D. Receptor editing in lymphocyte development and central tolerance. *Nat Rev Immunol*. 2006;6(10):728-740.

25. Liu M, Duke JL, Richter DJ, et al. Two levels of protection for the B cell genome during somatic hypermutation. *Nature*. 2008;451(7180):841.

26. Yamane A, Resch W, Kuo N, et al. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat Immunol*. 2011;12(1):62-69.

27. Staszewski O, Baker RE, Ucher AJ, Martier R, Stavnezer J, Guikema JEJ. Activation-induced cytidine deaminase induces reproducible DNA breaks at many non-Ig loci in activated B cells. *Mol Cell*. 2011;41(2):232-242.

28. Klein IA, Resch W, Jankovic M, et al. Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell*. 2011;147(1):95-106.

29. Chiarle R, Zhang Y, Frock RL, et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell*. 2011;147(1):107-119.

30. Zhang Y, McCord RP, Ho Y-J, et al. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell*. 2012; 148(5):908-921.