Brief Definitive Mini-Review for *Immunology and Cell Biology* Draft #3 as 21.8.17

Archived at viXra.org on 21.8.17 http://viXra.org/abs/1708.0242

Reverse Transcriptase Mechanism of Somatic Hypermutation: Sixty years of Clonal Selection Theory

Edward J Steele[1]

[1,a] CYO'Connor ERADE Village Foundation Inc. Piara Waters, WA, AUSTRALIA

Running head: *Mechanism of SHM*

Keywords

Somatic Hypermutation
Strand-Biased Mutations
DNA Polymerase–η
A-to-I RNA and DNA editing
RNA Exosome
AID-Deaminase

[a]*Correspondence:*
Associate Professor Edward J Steele PhD
Honorary Research Associate
CYO'Connor ERADE Village Foundation Inc.
24 Genomics Rise, Piara Waters, WA 6112, AUSTRALIA

Tel (+61) (0) 420 863 551
Email: ejsteele@cyo.edu.au

Abstract: The evidence for the reverse transcriptase mechanism of somatic hypermutation is substantial. In this 60th anniversary year of the publication of Burnet's Clonal Selection Theory the evidence is briefly reviewed and updated.

Abbreviations used in this paper:

A>>T , mutations of A exceed mutations of T, e.g. as found in normal murine physiological SHM *in vivo* by 2.9 fold **;ADAR,** Adenosine Deaminase that acts on RNA; **AID**, activation induced cytidine deaminase, a APOBEC family member, initiating via C-to-U lesions at **WRCY/RGYW** C-site motifs in ssDNA of class switch recombination (CSR) and somatic hypermutation (SHM) processes at somatically rearranged Ig V(D)J gene loci; **APOBEC family**, generic abbreviation for the deoxyribonucleic acid, or dC-to-dU, deaminase family of which AID is a member (APOBEC3 A, B, C, D, F, G, H) similar in DNA sequence to the "apolipoprotein B RNA editor" APOBEC1; **AP**, an Abasic, or apurinic/apyrimidinic, site; **APE**, AP endonuclease; **A-to-I**, adenosine-to-inosine RNA editing; **BCR**, B cell Ig receptor; **C**, exons encoding constant regions of Ig molecules; **D**, the small "diversity' element part of the VDJ rearrangement process; **G>>C**, mutations of G exceed mutations of C, e.g. as found in normal murine physiological SHM *in vivo* by 1.7 fold; **I**, Inosine; **Ig**, immunoglobulin; **IGHV, IGLV**, megabase length germline regions encoding unrearranged V element arrays and associated D,J, C regions; J, joining elements, part of the VDJ rearrangement process; **MMR**, mismatch repair; **MSH2-MSH6**, MutSα heterodimer recognizing mispaired bases in DNA duplex; **NTS**, the non-transcribed, or "Top", 5' to 3' strand; **Pol–η** or DNA polymerase-η (eta); **RNA Pol II**, RNA Polymerase II; **R**, Adenosine or Guanine; **RT**, reverse transcriptase; **RT-Pol-η**, reverse transcriptase activity displayed by Pol-**η; SHM**, somatic hypermutation; **TS**, the transcribed, or "Bottom", 3' to 5' strand, in context of a Transcription Bubble; **TSRT**, target site reverse transcription; **U**, uracil; **UNG**, uracyl DNA glycosylase involved in Base Excision Repair at dU sites in DNA resulting in either an Abasic site (AP) or APE-mediated ssDNA nicks (above); **V[D]J**, generic symbol for a rearranged immunoglobulin (or T cell receptor, TCR) variable region gene in the Adaptive Immune System; **W**, weak base pair involving A or U/T; **WA**-site, target motif for ADAR deaminase including DNA Polymerase-eta error prone incorporation *in vitro*; **Y,** pyrimidines T/U or C.

It is now 60 years since Sir MacFarlane Burnet first published *The Clonal Selection Theory of Acquired Immunity,*[1] the foundation stone of modern immunology. Somatic mutation of the immunoglobulin variable region genes has been part and parcel of his clonal selection concept since its inception and is central to a rational understanding of immunological diversification, self-tolerance and the emergence of cancer. We now have a very good idea of the molecular mechanism of somatic hypermutation. I have chosen to fit this scientific progress within 60 key publications since the late 1950s (Table 1). The *most likely* and *plausible* central molecular mechanism of Ig SHM, that fits with and explains all the evidence [58] is based on "Reverse Transcription" of the base-modified Ig pre-mRNA (Figure 1). That is, error-prone reverse transcription, by DNA Polymerase-η, of the Ig pre-mRNA template intermediate at rearranged V(D)J gene somatic loci. The Ig pre-mRNA encoding the V(D)J region is copied off the transcribed DNA strand carrying prior AID C-to-U deamination lesions (Uracils and Abasic sites), and it also accumulates ADAR-deaminase mediated RNA editing A-to-I modifications. This already base-modified pre-mRNA sequence is then copied back to the B lymphocyte genomic DNA and integrated at the rearranged VDJ site (concurrent with antigen-mediated selection of BCR bearing B lymphocytes, Centrocytes, in the Germinal Centre). This is essentially the "Reverse Transcriptase Mechanism" which Jeff Pollard and I first published 30 years ago.[21] The mechanistic steps, many logical, are clearly outlined in Figure 1 which shows that the A>>T and G>>C strand bias-generating mutagenic activity is firmly focused on the nascent RNA intermediate in the context of the Transcription Bubble. [52,52,54,58,60] Recent publications should be consulted for further recent definitive ADAR A-to-I editing of RNA and DNA moieties at RNA:DNA hybrids within Transcription Bubbles.[58-60] Not only is it important to understand the correct molecular mechanism of SHM for cancer diagnosis and detection [57,64] but also to the current efforts to better understand [31,65] the origin of Ig diversity involving the mechanism of evolution of the sets germline V segments and the long IGHV and

IGLV haplotypes in individual human beings. [66,67] The author welcomes discussion by email.

Table 1 History of Somatic Hypermutation : Developments Relevant to The Reverse Transcriptase Mechanism

Year	Author	Main Development-Discovery-Concept	Ref.
1957-59	FM Burnet	Somatic mutation concept for Ab mainly in context "forbidden" anti-self clones	1
1959	J Lederberg	Somatic mutation explicit in lymphocyte development and Ab diversity	2
1962	JB Fleishman et al	Amino acid variation in N -terminal regions of V or antigen binding regions	3
1966	S Brenner, C Milstein	Model: V region specific nicking and error prone repair - "somatic hypermutation"	4
1967	O Smithies	Somatic "Master->Slave" Gene Recombination model Ab diversity	5
1967	GM Edeleman, JA Gally	Somatic recombination between duplicated V gene model Ab diversity	6
1968	M Cohn	Molecular biology of expectation - rational for SHM and response to unexpected	7
1970	MG Weigert et al	Somatic variability in Lambda light chain V region protein sequences	8
1970	TT Wu, EA Kabat	Hypervariable regions coincide with and define antigen contact regions	9
1974	AJ Cunningham	The generation of antibody diversity after antigen	10
1974	M Cohn	Somatic mutation explanation for Ab diversity clearly laid out	11
1976	S Tonegawa, C Steinberg	DNA V gene counting confirms somatic mutation at molecular level in V lambda	12
1977	S Tonegawa et al	DNA V gene counting confirms somatic mutation at molecular level in V lambda	13
1981	PJ Gearhart et al	Somatic hypermutation of the TEPC15 VH rearranged gene *in vivo*	14
1981	ALM Bothwell et al	Somatic hypermutation to the VH186.2 VH rearranged gene *in vivo*	15
1981	E Selsing, U Storb	Somatic hypermutation of the MOPC167 VK rearranged gene *in vivo*	16
1982	PJ Gearhart	Somatic Hypermutation in Rearranged (VDJ) Variable Region Genes *In Vivo*	17
1983	PJ Gearhart,DF Bogenhagen	Somatic mutations occur in the 5' and 3' non-ding regions around VDJ genes	18
1985	C Berek, C Milstein	Use of hybridoma technique to sample somatic V[D]J mutant generation *in vivo*	19
1986	A Cumano, K Rajewsky	Further use hybridoma technique to sample somatic VDJ mutants *in vivo*	20
1987	EJ Steele, JW Pollard	Model : The reverse transcriptase mechanism of somatic hypermutation	21
1987	Golding et al	First hint of strand biases in somatic hypermutation patterns viz. A>G *versus* T>C	22
1990	GW Both et al	Defining the 5' and 3' boundaries of somatic hypermutation at VDJ genes	23
1990	SG Lebecque, PJ Gearhart	Defining 5' and 3' boundaries of somatic hypermutation at VDJ genes	24
1991-96	IB Rogozin et al	Identification RGYW/WRCY and WA hotspots in somatic hypermutation data	25,26
1992	EJ Steele et al	Defining the asymmetrical 5' to 3' somatic mutation distribution around V[D]J genes	27
1993	AG Betz et al	Defining the mutational hot spots across mutated V[D]J transgenes genes	28
1995	J Yelamos et al	Any non-Ig sequences parked between Promotor and J-C intron somatically mutate	29
1996	A Peters, U Storb	Strong evidence that transcription of VDJ target regions allows somatic mutation	30
1995-98	GF Weiller, RV Blanden et al	The SHM signature is written into the germline V segment array	31
1998	C Milstein et al	Both DNA strands targeted for G:C and A:T mutations in somatic hypermutation	32
1998	Y Fukita et al	Strong correlative evidence that transcription of VDJ allows somatic mutation	33
1998	C Rada et al	In MSH2-deficient mice mutations are G:C focused suggesting two stages SHM	34
1999	C Masutani et al	Discovery of DNA Polymerase -eta and Y family translesion polymerases	35
2000	M Muramatsu et al	AID discovered - required to intiate SHM and Ig Class Switch Recombination	36
2001-2	IB Rogozin, Y Pavlov et al	Error-prone DNA Polymerase eta SHM spectrum correlates with WA hotspots	37,38
2001	X Zeng et al	DNA Polymerase eta is the A:T mutator in somatic hypermutation in humans	39
2002-4	MS Neuberger et al	Definitive evidence that AID is a direct DNA C-to-U deaminase of the APOBEC family	40
2003	R Bransteitter et al	AID deaminates C>U on ssDNA - targets displaced strand Transcription Bubble	41
2003	J Chaudhuri et al	AID deaminates C>U on ssDNA - targets displaced strand Transcription Bubble	42
2003	SK Dickerson et al	AID deaminates C>U on ssDNA - targets displaced strand Transcription Bubble	43
2004	J Chaudhuri et al	AID deaminates C>U on ssDNA - targets displaced strand Transcription Bubble	44
2004	HM Shen, U Storb	AID targets both strands at Transcription Bubbles during transcription VDJ	45
2004	C Rada et al	MSH2-MSH6 -/-and Uracil DNA Glycosylase -/-define G:C and A:T mutation phases	46
2004	A Franklin et al	Human DNA Polymerase eta is an efficient reverse transcriptase, as are kapp, iota	47
2004	EJ Steele et al	First hint that A>G versus T>C strand bias involves an A>I RNA edited intermediate	48
2005	TM Wilson et al	MSH2-MSH6 stimulates DNA polymerase eta, suggesting a role for A:T mutations	49
2006	EJ Steele, RA Lindley et al	Evidence WA>WG mutations correlate with the number nascent WA RNA stem loops	50
2007	F Delbos et al	Evidence that DNA Polymerase eta is the sole error-prone A:T SHM mutator *in vivo*	51
2009	EJ Steele	SHM data 1984-2008 shows A>>T, G>>C strand biases explained by RNA/RT-model	52
2010-13	EJ Steele, RA Lindley	A>>T, G>>T SHM strand biases evident in non-Ig genes across all cancer exomes	53, 54
2011	U Basu et al	RNA exosome exposes ssDNA for AID on transcribed strand at Transcription Bubbles	55
2011	RW Maul et al	AID generated Uracils physically located in the DNA of VDJ & Ig class switch regions	56
2013	RA Lindley	Codon-context targeted somatic mutation (TSM) in cancer exomes	57
2016	EJ Steele	Extant evidence supports the RNA/RT-based model and not the DNA-based model	58
2017	Zheng et al	ADAR can directly edit both RNA and DNA A-sites in RNA:DNA hybrids	59
2017	EJ Steele, RA Lindley	ADAR A>I Editing at RNA:DNA Hybrids is strong support for RNA/RT-based model	60

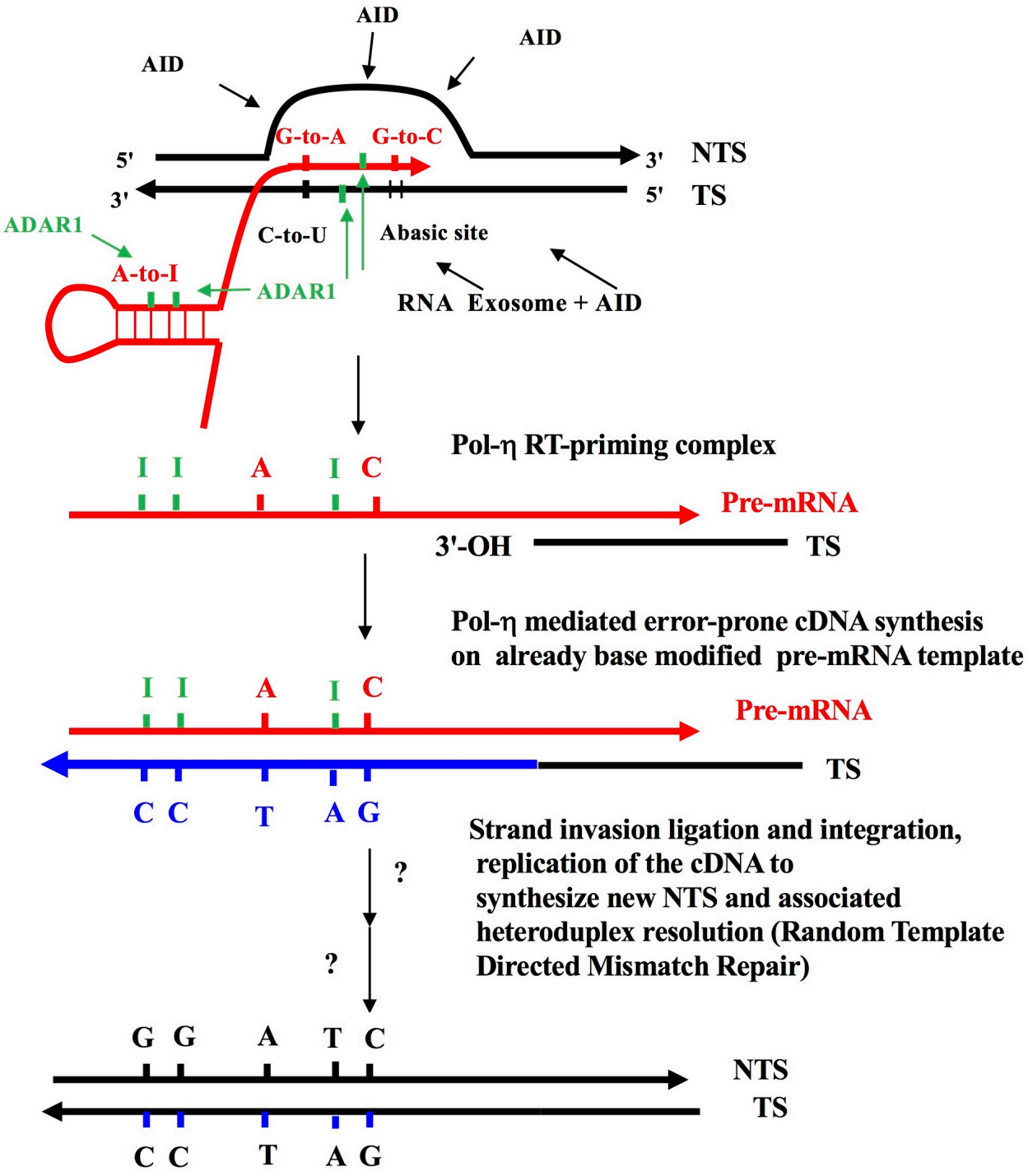

Figure 1 Legend - The Reverse Transcriptase Mechanism of Somatic Hypermutation
Modified in part from Figure 1 in Lindley and Steele.[54] This is an adaptation of the target site reverse transcription (TSRT) process of Luan et al. [61] Shown is an RNA Polymerase II-generated Transcription Bubble with C and A substrate deamination by AID and ADAR proteins; and the sequelae showing some key hypothesized DNA and RNA intermediates highlighted for the generation of the main strand-biased mutation signatures involving A-to-G, G-to-A, G-to-T and G-to-C.[52,58,60] Black lines are DNA strands, red lines are pre-mRNA, blue thick lines are cDNA strands copied off pre-mRNA by reverse transcription via DNA polymerase η. [47] Green bars are Inosines. Shown also is the action of the RNA exosome [55] allowing access of AID deaminase to

cytosines on the Transcribed Strand (TS). The ssDNA regions on the displaced Non-Transcribed Strand (NTS) are established targets of AID action.[41-44] With respect to the RNA intermediary step in this process mutations are first introduced at the DNA level by AID-mediated C-to-U deaminations, and then uracil DNA glycosylase (UNG)-generated excision create Abasic sites in the TS (which can further mature into single strand nicks via the action of AP endonuclease generating the 3'-OH in the TS). These template Uracil and Abasic sites are transcribed into pre-mRNA by RNA Pol II generating G-to-A and G-to-C modifications respectively in the pre-mRNA as shown in Kuraoka et al. [62] which on TSRT-mediated reverse transcription, [61] integration and DNA replication result in G-to-A and G-to-C mutations in the NTS, in a strand biased manner.[52,54,58,60] Separately, adenosine-to-inosine (A-to-I) RNA editing events at WA targets, mediated by ADAR1 deaminase, in the nascent pre-mRNA emerging from Transcription Bubble-proximal dsRNA stem loops may be copied back into DNA by reverse transcription via Pol-η. [50] In theory ADARs can also deaminate the RNA and DNA moieties in the RNA: DNA hybrid.[59,60] The strand invasion and integration of newly synthesized cDNA TS, as well as random-template mismatch repair [63] are hypothesized additional steps (not shown here). In more detail: RNA Pol II introduces mutations in the Ig pre-mRNA as it copies the AID lesions in TS DNA, concurrently A-to-I RNA edited sites appear in RNA stem(-loops) forming in nascent pre-mRNA near the transcription bubble [50] as well as in RNA:DNA hybrids within the bubble.[59,60] Next, the RT-priming substrates are formed by annealing the nicked TS strand with an exposed 3'-OH end (for Y Family translesion DNA polymerase-η,[35] now acting in its reverse transcriptase mode.[47] These could arise due to excisions at previous AID-mediated Abasic sites, or due to an excision introduced by endonuclease activity associated with the MSH2-MSH6 heterodimer engaging a U:G mispaired lesion.[49] This allows extension of a new TS by cDNA synthesis from the 3'-OH end copying the already base modified pre-mRNA template (with Inosine base pairing preferentially, like G, with C). Shown is an A>T transversion generated at the RT step at a template Inosine.

References

1. Burnet FM. *The Clonal Selection Theory of Acquired Immunity*. Cambridge University Press, 1959

2. Lederberg J 1959 Genes and Antibodies. *Science*. 1959; **129**: 1649-1653

3. Fleischman JB, Pain RH, RR Porter RR 1962 Reduction of gamma-globulins . *Arch.Biochem.Biophys* 1962; ;Suppl **1**:174-180.

4. Brenner S, Milstein C. Origin of antibody variation. *Nature* 1966; **211** : 242-246

5. Smithies O.Antibody Variability. Somatic recombination between the elements if "antibody gene pairs" may explain antibody variability. *Science* 1967; **157**: 267-273

6. Edelman GM, Gally JA. Somatic recombination of duplicated genes: an hypothesis on the origin Ab diversity. *Proc. Natl. Acad. Sci. U.S.A* 1967; **57** : 353- 358

7. Cohn M. The molecular biology of expectation. In *Nucleic Acids in Immunology* Ed Plescia OJ, Brain W Springer-Verlag NY pp 671 -715, 1968.

8. Weigert MG, Cesari IM, Yonkovich SJ, Cohn M 1970 variability in the Lambda light chain sequences. *Nature* 1970; **228** : 1045 -1047.

9. Wu TT, Kabat EA. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complimentarity. *J. Exp. Med*. 1970; **132** : 211 - 250

10. Cunningham AJ 1974. The generation of antibody diversity: its dependence on antigenic stimulation. *Contemp. Top. Molec. Immunol*. 1974; **3** : 1 - 26.

11. Cohn M. A rationale for ordering the data on antibody diversity. In *Progress Immunol. II Biol Aspects I* Brent L, Holborow J (Eds) .Nth Holl Publ Co. Amsterdam-Oxford-New York pp 261 - 284, 1974.

12. Tonegawa S, Steinberg C. Too many chains - Too few genes. In *Generation of Antibody Diversity : A New Look*. Cunningham AJ (Ed) Acad. Press NY pp 175 - 182. 1976.

13. Tonegawa S, Hozumai N, Matthyssens GH, Schuller R 1977 Somatic changes in the context and content of Immunoglobulin Genes*. Cold Spring Harb. Symp. Quant. Biol*. **41**. Pt 2 pp 872-889.

14. Gearhart PJ, Johnson ND, Douglas R, Hood L. IgG Antibodies to phosphorylcholine exhibits more diversity than their IgM counterparts. *Nature* 1981; **291** : 29-34.

15. Bothwell ALM, Pashkind M, Reth M, Imanishi-Kari T, Rajewsky, Baltimore D. Heavy chain variable region contribution to the NPb family of antibodies: somatic mutation evident in a IgG2a variable region. *Cell*. 1981; **24** : 625 - 637.

16. Selsing E, Storb U. Somatic mutation of immunoglobulin light-chain variable region genes. *Cell* 1981; **25** : 47 - 58.

17. Gearhart PJ 1982 Generation of Immunoglobulin Variable Gene diversity. *Immunol. Today* 1982; **3** : 107-112.

18. Gearhart PJ, Bogenhagen DF. Clusters of point mutations are found exclusively around rearranged antibody variable gene. *Proc. Natl. Acad. Sci. U.S.A* 1983; **80** : 3439-3443.

19. Berek C, Griffith GM, Milstein C. Molecular events during maturation of the immune response to oxazolone. *Nature* 1985; **316** : 412-418.

20. Cumano A, Rajewsky K. Clonal recruitment and somatic mutation in the generation of immunological memory to the hapten NP. *EMBO J*. 1986; **5** : 2459-2468

21. Steele EJ, Pollard JW. Hypothesis: Somatic hypermutation by gene conversion via the error-prone DNA->RNA->DNA information loop. *Molec. Immunol.* 1987; **24** : 667-673.

22. Golding, GB, Gearhart PJ, Glickman RW. Patterns of somatic mutations in immunoglobulin variable genes. *Genetics* 1987; **115** : 169-176.

23. Both GW, Taylor L, Pollard JW, Steele EJ. Distribution of mutations around rearranged heavy-chain antibody variable-region genes. *Mol.Cell.Biol*. 1990; **10** : 5187 - 5196.

24. Lebecque SG, Gearhart PJ. Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary is near the promoter, 3' boundary is approximately 1 kb from V-D-J gene. *J. Exp. Med*. 1990; **172** : 1717–1727

25. Rogozin IB, Solovyov VV, Kolchanov NA. Somatic hypermutagenesis in immunoglobulin genes. I. Correlation between somatic mutations and repeats. Somatic mutation properties and clonal selection. *Biochim. Biophys. Acta*. 1991; **1089** : 175-182.

26. Rogozin IB, Sredneva NE, Kolchanov NA.Somatic hypermutagenesis in immunoglobulin genes. III. Somatic mutations in the chicken light chain locus. *Biochim. Biophys. Acta*. 1996; **1306** : 171–178.

27. Steele EJ, Rothenfluh HS, Both GW. Defining the nucleic acid substrate for somatic hypermutation. *Immunol. Cell Biol*. 1992; **70** : 129-144

28. Betz AG, Neuberger MS, Milstein C. Discriminating intrinsic and antigen-selected mutational hotspots in immunoglobulin V genes. *Immunol. Today* 1993; **14** : 405-411.

29. Yelamos J, Klix N, Goyenechea B, Lozano, F, Chui YL, Gonzalez Fernandez A,Pannell R, Neuberger MS, Milstein C. Targeting of non-Ig sequences in place of the V segment by somatic hypermutation. *Nature* 1995; **376** : 225-229

30. Peters A, Storb U. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity* 1996; 4 : 57-65.

31. Blanden RV, Rothenfluh HS, Zylstra P, Weiller GF, Steele EJ. The signature of somatic hypermutation appears to be written into the germline IgV segment repertoire. *Immunol. Rev*. 1998; **162** : 117-132.

32. Milstein C, Neuberger MS, Staden. R. Both DNA strands of antibody genes are hypermutation targets. *Proc. Natl. Acad. Sci. U.S.A*. 1998; **95** : 8791- 8794.

33. Fukita Y, Jacobs H, Rajewsky K. Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity* 1998; **9** : 105 – 114.

34. Rada C, Ehrenstein MR, Neuberger MS, Milstein C. Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. *Immunity* 1998; **9** : 135 - 141.

35. Goodman MF. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem*. 2002; **71** : 17-50.

36. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T.Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a

potential RNA editing enzyme. *Cell* 2000; **102** : 553–563.

37. Rogozin IB, Pavlov YI, Bebenek K, Matsuda, T, Kunkel TA. Somatic mutation hotspots correlate with DNA polymerase eta error spectrum. *Nat. Immunol*. 2001; **2** : 530 -536

38. Pavlov YI, Rogozin IB, Galkin AP, Aksenova AV,Hanaoka F, Rada C, Kunkel TA. Correlation of somatic hypermutation specificity and A-T base pair substitution errors by DNA polymerase-eta during copying of a mouse immunoglobulin kappa light chain transgene. *Proc. Natl. Acad. Sci. U.S.A* 2002; **99** : 9954 - 9959.

39. Zeng X, Winter DB, Kasmer C, Kraemer KH, Lehmann AR, Gearhart PJ. DNA polymerase-eta is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat. Immunol*. 2001; **2** : 537–541

40. Noia JM, Neuberger MS.Molecular mechanisms of somatic hypermutation. *Annu. Rev. Biochem*. 2007; **76** : 1-22

41. Bransteitter R, Pham P, Scharff MD, Goodman MF. Activation-induced cytidine deaminase deaminates deoxcytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci. U.S.A* 2003; **100** : 4102-4107.

42. Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E, Alt FW.Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* 2003; **422** : 726-730.

43. Dickerson SK, Market E,Besmer E, Papvasiliou FN. AID mediates hypermutation by deaminating single stranded DNA. *J. Exp. Med*. 2003; **197** : 1291-1296.

44. Chaudhuri J, Khuong C, Alt FW. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature* 2004; **430** : 992–998.

45. Shen HM, Storb U. Activation-induced cytidine deaminase (AID) can target both DNA strands when the DNA is supercoiled. *Proc. Natl. Acad. Sci. U.S.A.* 2004; **101**: 12997-13002.

46. Rada C, Di Noia JM,Neuberger MS. Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. *Mol. Cell* 2004; **16** : 163-171.

47. Franklin A, Milburn PJ, Blanden RV, Steele EJ.Human DNA polymerase-eta an A-T mutator in somatic hypermutation of rearranged immunoglobulin genes, is a reverse transcriptase. *Immunol. Cell Biol*. 2004; **82** : 219-225.

48. Steele EJ, Franklin A, Blanden RV.Genesis of the strand biased signature in somatic hypermutation of rearranged immunoglobulin variable genes. *Immunol. Cell Biol.* 2004; **82** : 208-218.

49. Wilson TM, Vaisman A, Martomo SA, Sullivan P, Lan L,Hanaoka F, Yasui A, Woodgate R, Gearhart PJ. MSH2-MSH6 stimulates DNA polymerase eta, suggesting a role for A:T mutations in antibody genes. *J. Exp. Med*. 2005; **201** : 637- 645.

50. Steele EJ, Lindley RA, Wen J, Weiler GF. Computational analyses show A-to-G mutations correlate with nascent mRNA hairpins at somatic hypermutation hotspots. *DNA Repair* 2006; **5** : 1346 - 1363.

51.] Delbos F, Aoufouchi S, Faili A, Weill J-C, Reynaud C-A. DNA polymerase eta is the sole contributor of A/T modifications during immunoglobulin gene hypermutation in the mouse. *J. Exp. Med*. 2007; **204** : 17 - 23.

52. Steele EJ. Mechanism of somatic hypermutation: Critical analysis of strand biased mutation signatures at A:T and G:C base pairs. *Molec. Immunol* 2009; **46** : 305-320.

53. Steele EJ, Lindley RA. Somatic mutation patterns in non-lymphoid cancers resemble the strand biased somatic hypermutation spectra of antibody genes. *DNA Repair* 2010; **9** : 600-603.

54. Lindley RA, Steele EJ. Critical analysis of strand-biased somatic mutation signatures in TP53 versus Ig genes, in genome -wide data and the etiology of cancer *ISRN Genomics*. Vol 2013 Article ID 921418, 18 pages.

55. Basu U, Meng FL, Keim C, Grinstein V, Pefanis E,Eccleston J, et al.The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* 2011; **144** : 353-363.

56. Maul RW, Saribasak H, Martomo SA,McClure RL Yang, W, Vaisman, A, Gramlich, HS, Schatz DG, Woodgate R, Wilson DM, Gearhart PJ. Uracil residues dependent on the deaminase AID in immunoglobulin gene variable and switch regions. *Nat. Immunol*. 2011; **12** : 70-76.

57. Lindley R. The importance of condon context for understanding the Ig-like somatic hypermutation starnd-biased patterns in TP53 mutations in breast cancer. *Cancer Genetics* 2013; **206** : 222-226.

58. Steele EJ. Somatic hypermutation in immunity and cancer: Critical analysis of strand-biased and codon-context mutation signatures. *DNA Repair*. 2016; **45** : 1-24.

59. Zheng YC,Lorenzo C, Beal PA. DNA Editing in DNA/RNA hybrids by adenosine deaminases that act on RNA. *Nucleic Acids Res*. 2017; **45** : 3369-3377.

60. Steele EJ, Lindley RA. ADAR deaminase A-to-I editing of DNA and RNA moieties of RNA:DNA hybrids has implications for the mechanism of Ig somatic hypermutation. *DNA Repair* 2017; **55** : 1-6.

61. Luan DD, Korman MH, Jakubczak JL, Eichbush TH.Reverse transcription of R2B mRNA is primed by a nick at the chromosomal target site; A mechanism for non-LTR retrotransposition. *Cell* 1993; **72** : 595-605.

62. Kuraoka I, Endou M, Yamaguchi Y, Wada Y, Handa H, Tanaka K. Effects of endogenous DNA base lesions on transcription elongation by mammalian RNA polymerase II. *J. Biol. Chem*.2003; **278** : 7294-7299.

63. MacPhee DG. Mismatch repair, somatic mutations and the origins of cancer. *Cancer Res*.1995; **55** : 5499-5492.

64. Lindley RA, Humbert P, Larmer C, Akmeemana EH, Pendlebury CRR. Association between Targeted Somatic Mutation (TSM) signatures and HGS-OvCa progression. *Cancer Med*. 2016; **5** : 2629 - 2640

65. Steele EJ, Lloyd SS. Soma-to-germline feedback is implied by the extreme polymorphism at IGHV relative to MHC. *BioEssays* 2015; **37**: 557-569.

66. Kidd MJ, Chen Z, Wang Y, Jackson KJ, et al. The inference of phased haplotypes for the immunoglobulin H chain V region gene loci by analysis of VDJ rearrangements. *J Immunol* 2012; **188**: 1333-13340.

67. Watson CT, Steinberg KM, Huddleston J, Warren RL, et al. Complete haplotype sequence of the human immunoglobulin heavy- chain variable, diversity, and joining genes and characterization of allelic and copy-number variation. *Amer J Hum Genet* 2013; **92**: 530-546.