International Immunology, Vol. 11, No. 5, pp. 825–833
© 1999 The Japanese Society for Immunology

Mutational pattern of the nurse shark antigen receptor gene (NAR) is similar to that of mammalian Ig genes and to spontaneous mutations in evolution: the translesion synthesis model of somatic hypermutation

Marilyn Diaz[1,2], Jovanna Velez[1], Mallika Singh[2], Jan Cerny[2] and **Martin F. Flajnik[1,2]**

[1]University of Miami School of Medicine, Department of Microbiology and Immunology, Miami, FL 33136, USA
[2]University of Maryland at Baltimore, Department of Microbiology and Immunology, 655 West Baltimore Street, Baltimore, MD 21201, USA

Keywords: DNA polymerase ζ, evolution, mutational pattern, nurse shark antigen receptor, *REV3*, somatic hypermutation

Abstract

The pattern of somatic mutations of shark and frog Ig is distinct from somatic hypermutation of Ig in mammals in that there is a bias to mutate GC base pairs and a low frequency of mutations. Previous analysis of the new antigen receptor gene in nurse sharks (NAR), however, revealed no bias to mutate GC base pairs and the frequency of mutation was comparable to that of mammalian IgG. Here, we analyzed 1023 mutations in NAR and found no targeting of the mechanism to any particular nucleotide but did obtain strong evidence for a transition bias and for strand polarity. As seen for all species studied to date, the serine codon AGC/T in NAR was a mutational hotspot. The NAR mutational pattern is most similar to that of mammalian IgG and furthermore both are strikingly akin to mutations acquired during the neutral evolution of nuclear pseudogenes, suggesting that a similar mechanism is at work for both processes. In yeast, most spontaneous mutations are introduced by the translesion synthesis DNA polymerase ζ (*REV3*) and in various DNA repair-deficient backgrounds transitions were more often *REV3*-dependent than were transversions. Therefore, we propose a model of somatic hypermutation where DNA polymerase ζ is recruited to the Ig locus. An excess of DNA glycosylases in germinal center reactions may further enhance the mutation frequency by a *REV3*-dependent mutagenic process known as imbalanced base excision repair.

Introduction

Antigen-driven somatic hypermutation of Ig genes in humans and mice generates high-affinity receptors for memory immune responses (1). Although the hypermutation mechanism has not been elucidated, analysis of passenger transgenes, JC (joining/constant regions) intron sequences, synonymous mutations and even heterologous DNA placed within Ig constructs has revealed intrinsic properties of the mechanism. These include the mutation hotspot sequence AGC/T, a bias for transitions over transversions, a preponderance of base substitutions over deletions and insertions, and strand polarity (2–6). Among ectothermic vertebrates, somatic hypermutation of Ig genes has been shown to occur in the amphibian *Xenopus* and in cartilaginous fish (7–9). Shark and frog Ig hypermutation, however, while displaying the AGC/T

This work is dedicated to the memory of Jovanna Velez

Correspondence to: M. F. Flajnik

Transmitting editor: M . S. Neuberger

Received 18 November 1998, *accepted* 2 February 1999

Downloaded from https://academic.oup.com/intimm/article/11/5/825/710480 by guest on 21 August 2022

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	
Type I	P	R	S	V	T	K	E	T	G	E	S	L	T	I	N	C	V	L	R	**D**	**A**	**S**	**Y**	**A**	**L**	**G**	**S**	
	CCG	AGA	TCA	GTA	ACA	AAG	GAG	ACG	GGC	GAA	TCA	CTG	ACC	ATC	AAC	TGT	GTC	CTA	CGA	**GAT**	**GCG**	**AGC**	**TAT**	**GCA**	**TTG**	**GGC**	**AGC**	81
	---	CA-	A--	A--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	AGT	-A-	-G-	---	---	TC-	---	
Type II	P	Q	T	I	T	K	E	T	G	E	S	L	T	I	N	C	V	L	R	**D**	**S**	**N**	**C**	**A**	**L**	**S**	**S**	

	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	
Type I	T	C	W	Y	R	K	K	S	G	S	**T**	**N**	**E**	**E**	**S**	**I**	**S**	**K**	G	G	R	Y	V	E	T	V	N	
	ACG	TGC	TGG	TAT	CGA	AAA	AAA	TCG	GGC	TCA	**ACA**	**AAC**	**GAG**	**GAG**	**AGC**	**ATA**	**TCG**	**AAA**	GGT	GGA	CGA	TAT	GTT	GAA	ACA	GTT	AAC	162
	---	-A-	---	---	--C	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	
Type II	T	Y	W	Y	R	K	K	S	G	S	**T**	**N**	**E**	**E**	**S**	**I**	**S**	**K**	G	G	R	Y	V	E	T	V	N	

	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	
Type I	S	G	S	K	S	F	S	L	R	I	K	D	L	T	V	E	D	G	G	T	Y	R	C	
	AGC	GGA	TCA	AAG	TCC	TTT	TCT	TTG	AGA	ATT	AAT	GAT	CTA	ACA	GTT	GAA	GAC	GGT	GGC	ACG	TAT	CGT	TGC	231
	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	A--	---	---	---	--A	---	
Type II	S	G	S	K	S	F	S	L	R	I	K	D	L	T	V	E	D	S	G	T	Y	R	C	

Fig. 1. Germline sequences for the V regions of type 1 NAR and type 2 NAR (11), upper and lower sequences respectively. Bold face codons indicate the CDR.

hotspot sequence, a bias for transitions, and the preference for base substitutions, differs from the mammalian mechanism in that mutations usually occur at GC pairs and there is a lower mutation frequency with no concentration of mutations to the complementarity determining region (CDR) (10).

NAR is an antigen receptor in the nurse shark that diversifies somatically by hypermutation (11). We have recently shown that NAR hypermutation does not generate the primary repertoire like in sheep (12), but is probably antigen driven (13). In a previous study, NAR clones isolated from the spleen of an adult shark showed a high frequency of mutation, similar to that observed in human/mouse Ig (11). Given this high mutation frequency one might expect that the mechanism in NAR is ancestrally related to the mechanism in mammals. Conversely, NAR hypermutation may have evolved convergently and displays properties distinct from the mammalian mechanism. In this study, we analyze the mutational pattern from 150 NAR clones to further ascertain whether it resembles hypermutation in mammals. In addition, we compare the types of mutations found in NAR genes with spontaneous mutations generated over evolutionary time. Finally, the data are synthesized to generate a refined model for the somatic hypermutation mechanism based on translesion synthesis by the DNA polymerase ζ (10).

Methods

NAR mutation analysis: RNA isolation, first-strand synthesis, cloning and sequencing

Reaction conditions and the oligonucleotide primers used for first-strand cDNA synthesis and PCR amplifications have been described in detail elsewhere (13). Briefly, RNA was isolated from peripheral blood lymphocytes (PBL) of three adult sharks, and cDNA was synthesized from each shark using primers specific to the transmembrane region (Tm) and secretory (Sec) tail of NAR. cDNA templates for each set of samples were PCR amplified with forward primers to the variable (V) region of NAR (Fig. 1) and reverse primers to the first constant (C) domain. Genomic DNA was isolated from erythrocytes as described previously (13) and PCR amplified with the same forward primer used for cDNA PCR amplification and reverse primers complementing the joining region. The PCR products from both cDNA and DNA amplifications were subcloned into the PCR2.1 vector; ligation and transformation reactions were performed following the protocol for the Original TA cloning kit (Invitrogen, Carlsbad, CA). Clones were sequenced with Sequenase (US Biochemical, Cleveland, OH) and sequences derived from each set of cDNA samples were compared to the germline using the Intelligenetics DNA analysis package (Intelligenetics, Campbell, CA). To establish a background error rate, cDNA for the J chain locus of the nurse shark was synthesized, PCR amplified, subcloned and sequenced as done for the NAR clones (as previously described) (13). Conditions and experimental procedures for the generation of a nurse shark spleen cDNA library and screening of NAR^+ clones have been described in detail elsewhere (11).

Quantitative analysis

The expected frequency for each type of mutation was estimated by determining how often a given substitution type can generate a synonymous mutation. The expected frequency of mutations impacting each of the nucleotides was determined by the frequency of each nucleotide at degenerate sites. Finally, the expected number of transitions and transversions was estimated by multiplying the expected

Downloaded from https://academic.oup.com/intimm/article/11/5/825/710480 by guest on 21 August 2022

Table 1. NAR hypermutation displays a bias for transitions and strand polarity but no targeting to GC bases

To	From				Total
	G	A	T	C	
G	–	37 (0.41/0.40)	8 (0.20/0.19)	5 (0.6/0.19)	50
A	14 (0.56/0.41)[a]	–	3 (0.07/0.27)	9 (0.12/0.25)	26
T	6 (0.24/0.27)	25 (0.28/0.29)	–	63 (0.82/0.56)[b]	94
C	5 (0.20/0.32)	27 (0.30/0.30)	29 (0.72/0.54)	–	61
Total	25	89	40	77	231
Observed	11%	38%	17%	33%	
Expected	18%	35%	17%	26%	
	Observed[c]	Expected			
Transitions	143 (115)	111 (95)			
Transversions	88	120 (107)			

[a]Observed/expected frequencies in parenthesis.
[b]Includes c to t mutations to the agc hotspot. An excess of c to t transitions is still detected when the hotspot c to t mutations are excluded (0.71/0.56).
[c]Numbers in parenthesis exclude c to t transitions to the agc hotspot.

frequency of each class (i.e. transitions) to the total number of observed mutations. The DNASIS software for genetic analysis (Hitachi Software, South San Francisco, CA) was used to determine whether potential secondary structures could be formed within the NAR V region.

Results

Analysis of NAR synonymous mutations

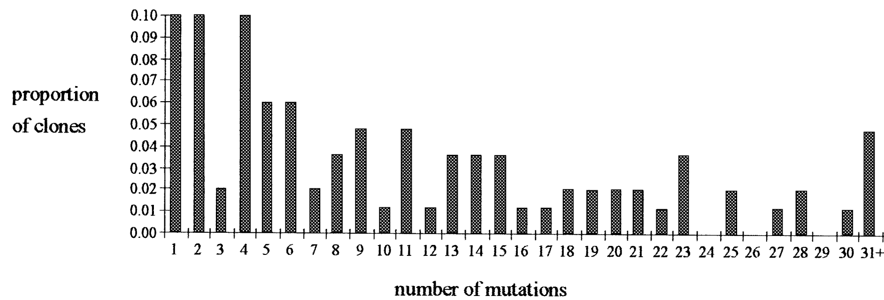
To determine the intrinsic properties of the hypermutation mechanism, synonymous mutations found in NAR cDNA clones were enumerated and compared to expected values. No targeting to any particular nucleotide was detected, such as the GC bias found in shark and frog Ig mutations (13) (Table 1). Replacement substitutions did not reflect a GC bias either (data not shown). A significant bias towards transitions was observed, even when the bias in the universal code for synonymous changes to be generated by transitions in 2-fold degenerate codons was considered (χ^2, $P < 0.01$; Table 1).

The transition bias remained significant when transitions in the 'hotspot' AGC codons in the CDR were removed (χ^2, $P < 0.05$) and was still apparent when the analysis was limited to third-base mutations of 4-fold degenerate codons in clones with <15 mutations. Previously, we found that the majority of NAR transmembrane (Tm) clones, unlike secretory (Sec) clones, had few or no mutations and that the few Tm mutations displayed no evidence of antigen-driven selection (13). We concluded in that study that NAR hypermutation in the nurse shark does not generate the repertoire and suggested that the few Tm mutations were a result of a leaky mutational mechanism, perhaps during lymphocyte development. Consistent with the idea that Tm mutations are not under antigen-driven selection, a bias for transitions among replacement substitutions was detected in Tm clones. This finding further supports the premise that the transition bias is a property of the NAR hypermutation mechanism as replacements are not prone to generate transitions—any bias from the universal code favors transversions since third base-generated replacements occur only via transversions. Finally, there was a somewhat higher frequency of mutations from C than from G and the transversion T to A but not A to T was significantly lower than expected (*G*-test, $P < 0.05$; Table 1) suggesting strand bias.

The pattern of synonymous mutations in NAR and the frequency of mutation (Figs 2 and 3) is similar to that observed in mouse Ig passenger transgenes (4), suggesting that the mutational mechanism of NAR shares many, if not all, the components of Ig hypermutation. Furthermore, mutations in NAR and Ig passenger transgenes are similar to the pattern of mutation seen in the evolution of nuclear non-Ig pseudogenes (Fig. 3). The bias for transitions in both somatic hypermutation and pseudogene evolution had been noted previously (4,14). In addition, mutations from A, both in evolution and in somatic hypermutation, were more evenly distributed among transitions and transversions while transitions predominated in mutations from the other nucleotides (Fig. 3). However, in contrast to mitochondrial genes where >90% of the mutations were transitions (15,16), there was still a significant number of transversions in both nuclear pseudogene mutations and immune gene somatic hypermutation (Table 2). Also, in one pseudogene study of seven sequences, evidence for strand bias was found (17).

The high frequency of transitions in one group of pseudogenes (18) had been attributed to the deamination of methylated C to form T (2). This possibility, however, was adjusted for in the pseudogene study and a bias for transitions was still obtained (54.4%) (18). Furthermore, this biochemical modification cannot account for the transitional bias from A and T in pseudogenes. In summary, these results strongly suggest that the mechanism responsible for the formation of mutations in pseudogene evolution has been recruited to somatically mutate antigen receptor genes.

Downloaded from https://academic.oup.com/intimm/article/11/5/825/710480 by guest on 21 August 2022



mutation frequency (among mutated clones) = 38.8 x 10^{-3}

Fig. 2. A majority of Sec clones (83 of 97) were mutated with most having four mutations and more, yielding a mutation frequency similar to mammalian IgG.

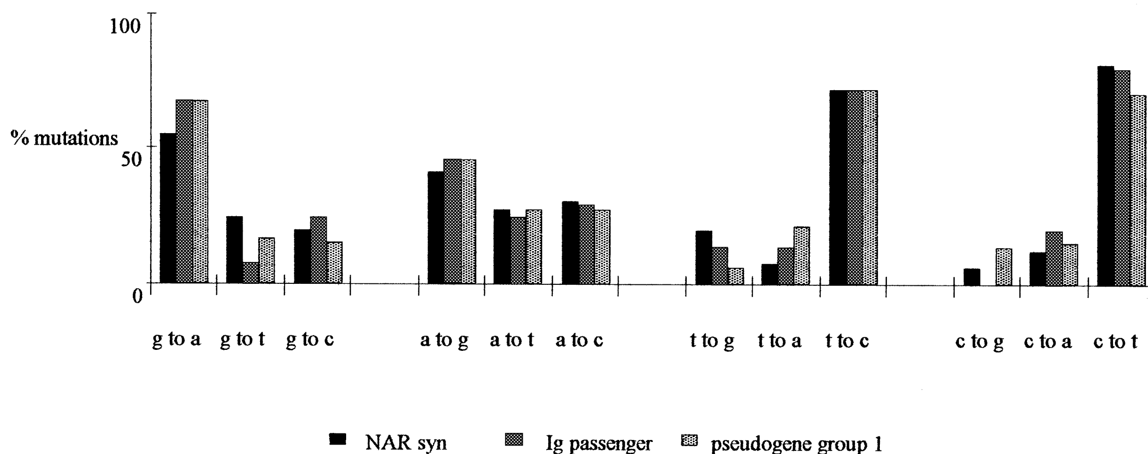


Fig. 3. The pattern of mutations of NAR synonymous changes (n = 231) and passenger Ig transgenes (n = 71) (4) is similar to that of pseudogenes in evolution (goat $\varphi\beta^x$globin, goat $\varphi\beta^z$globin and human tubulin pseudogenes were included in the analysis) (n = 265) (18). Pseudogene data from Wu and Maeda (17) are not shown in this analysis, but revealed a similar pattern. Changes from each nucleotide were treated as independent groups.

Table 2. The bias for transitions in somatic hypermutation of NAR and mammalian Ig is similar in magnitude to that detected in the evolution of nuclear pseudogenes

Sequence analyzed	Transitions (%)	References
Nar Syn	62	this paper
Nar Syn (4-fold degeneracy)	56	this paper
Ig passenger transgene	63	Betz *et al.* (4)
JC intron	60	Gonzalez-Fernandez *et al.* (22)
Ig clones out of frame	59	Gonzalez-Fernandez *et al.* (22)
Pseudogenes (8)	59.2	Li *et al.* (18)
Pseudogenes (7)	66	Wu and Maeda (17)
Mitochondria	96	Brown *et al.* (15)

Insertions, deletions and hotspots

Of 1023 mutations analyzed from four different sharks, 1018 were single base substitutions. The few deletions and insertions are displayed in Fig. 4. Consistent with a previous study of mutations in human tonsillar cells (19), deletions and insertions in NAR occurred in triplet or triplet multiples, presumably to preserve the reading frame. It is likely that in non-coding sequences single base pair insertions and deletions may be more common.

Interestingly, three of the five insertions/deletions occurred near the end of CDR2. This region has a high propensity to form a hairpin loop (codons 42–50 in Fig. 1), with a stem 9 bases long, in which 8 bp can form, and a loop consisting of three codons (Fig. 4). At the base of the stem is an AGC hotspot. An 18 bp duplication begins with the cga codon after the loop, one of the insertions occurred in the middle of the loop and the sole deletion removed the agc codon from the base of the stem.

Downloaded from https://academic.oup.com/intimm/article/11/5/825/710480 by guest on 21 August 2022

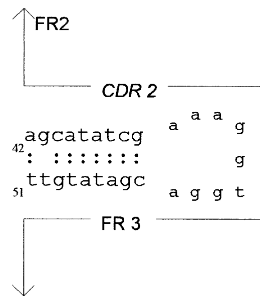
Deletions **Insertions** **Duplication**

$_{40}$gaggagagcata* $_{20}$gatgcgagcgcatatgca $_{47}$ggacgatatgttgaaacagttcgatatgttgaaacagttaac*

$_{68}$acagttgaaagagac

$_{44}$tcgaaagttggtgga*

* mutation at CDR2 loop :



Secondary structure energy = -9.4 kcal/mol

Fig. 4. Insertions and deletions in the NAR locus tended to occur in a region of potential secondary structure in CDR2. Codon numbers (indicated in Fig. 1) are noted in the 5′ position. Underlined residues depict the inserted, deleted or duplicated codons. The secondary structure energy was calculated by the DNASIS program (Hitachi Software).

Among silent mutations, the most commonly modified sequence in the CDR of NAR is the AGC motif, known to be a hotspot for somatic hypermutation (shown here for the most common types of NAR genes, types 1 and 2; Fig. 5A and B respectively). The TGC and GCA triplets were also mutated frequently. The GCA triplet had not been previously identified as a hotspot in any species. However, GCA falls within the short palindrome, *atgcat* (see tandem mutation discussion below). When replacements were included, other potential hotspot motifs such as AAC, AGT, TAC and AAA were detected, although a role of selection preserving mutations at these codons cannot be ruled out (AAC and TAC also appear to be hotspots in mammalian Ig, (20). In framework (FR) 3, we detected another hotspot in a GTT codon which is embedded within a hotspot and a short palindrome: *aacagtt*. In fact, some of the small FR peaks were embedded within this sequence.

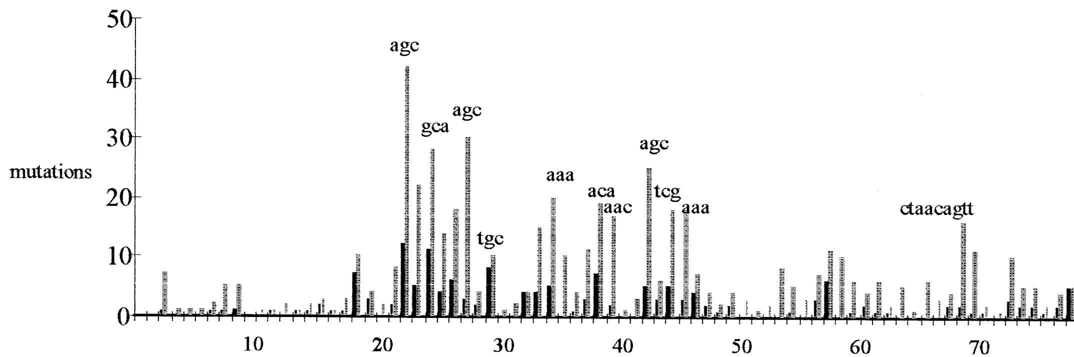
Tandem mutations

Unlike mutations found in Ig genes in all other species, many of the mutations in NAR occurred in doublets and triplets. It is not clear, however, whether a high frequency of tandem mutations is a unique property of the NAR mutation mechanism or a result of NAR's high mutation frequency combined with heavy CDR clustering. To differentiate between these two possibilities, we analyzed tandem mutations in the FR of Sec clones with <20 mutations and in the entire NAR V region of the poorly mutated Tm clones. Two observations suggested that the generation of tandem mutations is not entirely (if at all) an intrinsic property of the mechanism (Table 3): (i) multiple changes tended to occur at hotspot motifs or within palindromic sequences (in some cases, tandem mutations were embedded in *both* a hotspot motif and a palindrome) and (ii) a string of adenines in FR2 (codons 33–34; Fig. 1) was a hotspot for mutations, particularly for mutations in tandem; mononucleotide runs have been implicated in a hypermutable allelic form of the human APC gene associated with an inherited form of colorectal cancer (21). The *aacagtt* sequence, which is found twice in the NAR V sequence, was the most common palindrome associated with multiple tandem mutations, accounting for 34 of 115 mutations in this analysis. Of the 48 tandem events examined, 46% occurred at or next to palindromes while only 17.5% of all single mutations (40 mutations) and 17% of tandem events (six events) within FR1, a region not expected to be a 'good' substrate for mutations, were associated with a palindrome (*G*-test, $P < 0.05$). Potential secondary structures and hotspot association with palindromes have been observed in mammalian Ig (2,19,22–24) and implicated as potential players in the mutational mechanism (2). Recently, a mouse transgenic 'mutational substrate' model, designed to contain recognition sites for restriction enzymes to facilitate mutation analysis, unexpectedly resulted in a dramatic increase in mutations in and around the repeated palindromic sequences (25). The authors found a positive correlation between mutation peaks and the energy required to melt the potential secondary structures formed by these palindromes, although they argue that increased mutation is displaced by ~40 nucleotides 5′ from the hairpin structures.

Potential secondary structure formation during the reverse transcription of the RNA to generate the cDNA clones used in this study could potentially explain the increased frequency of tandem events, insertions and deletions near palindromic sequences. However, the insertions and deletions involved

Downloaded from https://academic.oup.com/intimm/article/11/5/825/710480 by guest on 21 August 2022

A.



B.

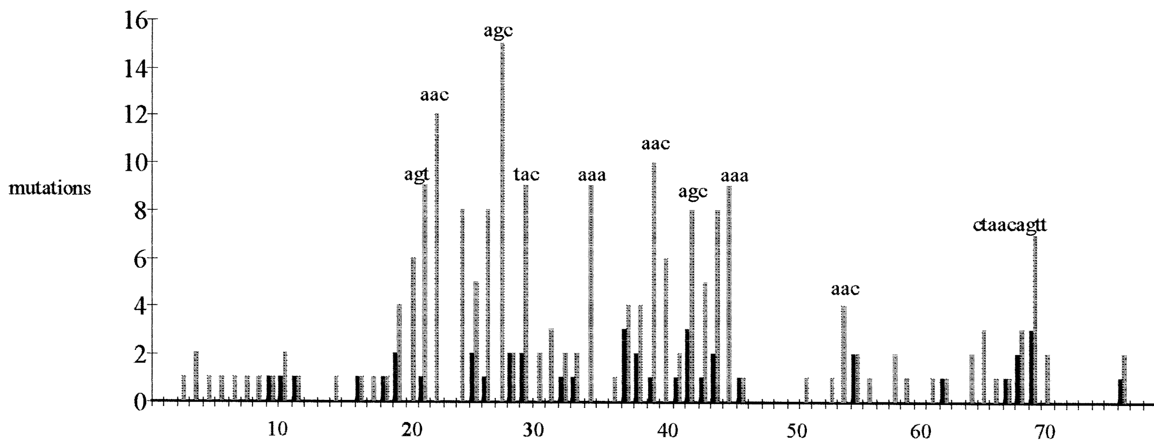


Fig. 5. NAR mutational hotspots. Complex sequence (across-codons including the last base of codon 67–69) represents a short palindrome. Total mutations at each codon (see Fig. 1) are depicted by dark gray bars and synonymous changes alone are depicted in black bars.

triplets suggesting selection to maintain reading frame, and a similar mutational analysis of cDNA NAR clones generated from nurse shark pups revealed low levels of mutation (62 mutations out of 127 clones) in which none of the mutations were in tandem and no insertions nor deletions were detected (M. Diaz and M. F. Flajnik, unpublished data).

Discussion

Like Ig of all species studied to date, the NAR hypermutation mechanism generates more transitions than transversions, the AGC/T motif is a mutation hotspot and most changes are base substitutions. Like mammalian but not *Xenopus* or shark Ig, the frequency of hypermutation of NAR genes is high (average of 5.14 mutations/clone, which includes unmutated clones) and GC base pairs are not targeted (13). Strand bias is also a property of NAR hypermutation but it is not known whether mutation in *Xenopus* and shark Ig targets only one strand. It is clear from these findings that hypermutation of NAR in the nurse shark is most similar in its pattern and frequency of mutations to that of mammalian Ig, and it almost certainly shares many, if not all, of the components of the mammalian mechanism. NAR hypermutation apparently displays one unique characteristic: a significant proportion of the mutations occurred in tandem. However, a high proportion of tandem mutations were embedded within hotspot motifs or palindromes, suggesting that at least some of the frequent doublets and triplets were sequence driven. Interestingly, Ig genes in mice deficient to the mismatch repair gene PMS2 had an increased frequency of tandem mutations in Ig genes (26), suggesting a role for PMS2 in repairing tandem mutations; perhaps access of PMS2 to the NAR locus during hypermutation is blocked. The mechanism by which tandem mutations are introduced near palindromes is unclear at this time, but we propose that palindromes either form small ‘bubbles’ of secondary structure in the path of a polymerase or they serve as signal sequences for accessory molecules to introduce nicks along the V region. The correlation between the stability of potential secondary structure and mutation frequency found in the mouse study (25) suggests that hairpin

Downloaded from https://academic.oup.com/intimm/article/11/5/825/710480 by guest on 21 August 2022

Table 3. Tandem mutations in NAR tended to occur at hotspots and/or at short palindromic sequences

Doublets		Triplets	Quartet	Quintet
$_{5}$acaaaggag	$_{10}$gaatca	$_{33}$aaaaaaa	$_{53}$**gttaac**agc	$_{67}$ct**aacagtt**gaa
$_{30}$tggtatcgc	$_{51}$ga**aacagtt**	$_{74}$acgta**tcga**		
$_{53}$**gttaac**agc	$_{28}$ac**gtac**	$_{71}$gacagt		
$_{28}$ac**gtac**	$_{6}$aaggagacg	$_{67}$ct**aacagt**t		
$_{67}$ct**aacagtt**	$_{23}$t**atgcat**tg	$_{33}$aaaaaa		
$_{60}$ttttctttg	$_{33}$aaaaaa	$_{71}$gacggtggc		
$_{67}$ct**aacagtt**ga	$_{29}$tactggtat	$_{33}$aaaaaa		
$_{44}$tcgaaaggt*	$_{33}$aaaaaa	$_{33}$aaaaaa		
$_{60}$ttttctttg	$_{56}$ggatcaaag	$_{72}$ggtggcacg		
$_{67}$ct**aacagtt**	$_{7}$gagacggc	$_{60}$ttttctttg		
$_{2}$a**gatc**a	$_{17}$gtcctacga			
$_{30}$tggta**tcga**	$_{6}$aaggagacg			
$_{51}$ga**aacagtt**	$_{51}$ga**aacagtt**			
$_{56}$g**gatc**aaag	$_{67}$ct**aacagtt**			
$_{57}$tcaaag	$_{54}$aacagcgga			
$_{67}$c*t***aacagtt**	$_{51}$ga**aacagtt**			
$_{67}$ct**aacagtt**	$_{67}$ct**aacagtt**			
$_{29}$tactggtat				
$_{60}$ttttcttg				

Total tandem mutations (underlined); germline sequence shown:	111
Mutations in or next to palindromes(bold):	29
Mutations in hotspots (agc,tgc,aac,tac):	8
Mutations in palindromes with hotspots:	22
Mutations in mononucleotide runs (all are aaaaaa):	16
Mutations in other:	34
*In CDR2 hairpin loop (see Fig. 3):	2

structures enhance mutation by presenting obstacles to either an RNA polymerase as previously suggested (25) or to an error-prone DNA polymerase. Inverted repeats previously were shown to be associated with base substitutions, deletions and insertions in human Ig (19). The authors of that study proposed that a polymerase with limited proofreading abilities might associate loosely with the DNA template and be susceptible to areas of secondary structure.

The similarities in the pattern of synonymous mutations in NAR and mammalian Ig to that of non-coding sequences and pseudogenes in evolution, particularly the bias for transitions, may provide clues to the molecular basis for hypermutation. These data suggest that an ancient mechanism was recruited to generate somatic mutations in Ig, implicating replication and/or DNA repair (both processes have been proposed to be involved in somatic hypermutation of Ig by others, 27–31). The intrinsic bias for the genetic code to generate synonymous changes by way of transitions may have been driven by a propensity of DNA polymerases to commit such errors more often than transversions. In fact, a bias for transitions is characteristic of the error-prone DNA polymerase β (32). Furthermore, the similar bias for transitions seen in somatic hypermutation and pseudogenes in evolution cannot be explained by an active role of repair molecules without the involvement of a polymerase; repair molecules normally correct mutations rather than induce them. Errors during replication and mutagenic bypass of lesions in a process known as translesion synthesis (33) are thought to be significant sources of spontaneous base substitutions in most genes. Previously, we hypothesized that a translesion synthesis polymerase homologous to DNA polymerase ζ is responsible for

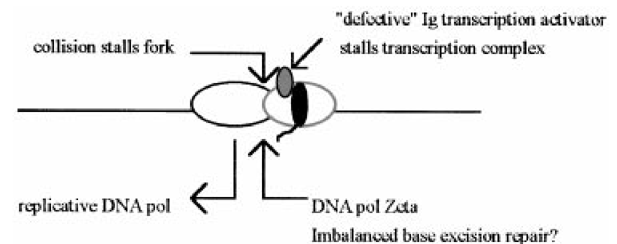


Fig. 6. Translesion synthesis model of somatic hypermutation. A 'defective' Ig transcription activator stalls the transcription machinery causing a collision with the replication fork (10). The stalled replication fork sends a lesion signal and DNA polymerase ζ continues synthesis of the daughter strand. Mutations may be further enhanced by imbalanced BER where the generation of multiple abasic sites by DNA glycosylases are then processed or bypassed in an error-prone fashion by pol ζ.

somatic hypermutation of antigen receptor genes (10). In the yeast, DNA polymerase ζ, in which the catalytic subunit is encoded by *REV3*, is a non-replicative polymerase that carries out most if not all of the translesion synthesis (34). DNA polymerase ζ also engages non-damaged templates in sites near double-strand breaks, increasing the frequency of base substitutions in the vicinity of the break by at least 100-fold (35). In *REV3* mutants, spontaneous mutations are reduced to one-fifth of wild-type levels (36). Thus DNA polymerase ζ is responsible for the majority of spontaneous mutations in the yeast, where analysis of mutations in non-coding regions revealed a bias for transitions over transversions (59% transitions) (37). Furthermore, in various DNA repair-deficient back-

Downloaded from https://academic.oup.com/intimm/article/11/5/825/710480 by guest on 21 August 2022

grounds, transitions were more often *REV3*-dependent than transversions (38). In our model, we envisioned that a 'defective' Ig transcription activator stalls the transcription machinery in the Ig locus during germinal center reactions, causing the replication fork to stall. The stalled replication fork alerts the cell with a 'lesion' signal and the error-prone translesion synthesis polymerase is recruited into the Ig locus (Fig. 6). Recently, it was shown in yeast that imbalanced base excision repair (BER) can increase the mutation frequency in a *REV3*-dependent fashion by up to 600-fold (39). High levels of a DNA glycosylase (3MeA), relative to the apurinic/apyrimidinic endonuclease, dramatically increased the level of spontaneous mutations, presumably by generating multiple abasic sites which were then processed or bypassed by DNA polymerase ζ. We propose that somatic hypermutation of antigen receptor loci is enhanced by a similar mechanism, where the lesion signal recruits glycosylases to the Ig locus to generate abasic sites, that are processed by the translesion synthesis polymerase ζ. Alternatively, DNA polymerase ζ is simply recruited to bypass the 'lesion' and continue synthesis of the daughter strand for a 'short patch' (34). Human and mouse homologues to DNA polymerase ζ have recently been cloned (40,41), and the human homolog appears to be involved in mutagenic lesion bypass (41).

Most BER in mammals is carried out by DNA polymerase β (42), but somatic hypermutation is unaffected in mice deficient for this polymerase (43). It is still possible that a DNA polymerase β homolog has evolved to somatically mutate Ig genes (29). However, the recently elucidated *REV3*-dependent mutagenic BER pathway in yeast may represent an ancient alternative pathway in mammals. Interestingly, a DNA glycosylase (OGG) was recently found to be overexpressed in germinal centers, which led the researchers to propose that this enzyme is involved in somatic hypermutation (44).

The translesion synthesis DNA polymerase ζ (*REV3*) is a good candidate; it is mutagenic, in yeast it mostly generates base substitutions, it appears to have a bias for transitions and as it is not a processive polymerase (34) it may be susceptible to DNA secondary structure. The recent cloning of this enzyme in mice and humans will allow us to test the model with knockout or antisense technology. A role for mismatch repair, however, is likely if hypermutation occurs during or after replication of the Ig locus; regardless of the method of mutation introduction, post-replication repair must be down-regulated or blocked from the Ig locus for perpetuation of mutations in B cell clones. However, mice deficient in the mismatch repair gene, MSH2, display an altered pattern of mutations (29,43,45,46), suggesting that some level of mismatch repair does occur in the mutated Ig locus template. The targeting of mutation to GC bases in shark IgM but not in NAR does in fact suggest that somatic hypermutation may in fact be a multi-layered process, as previously suggested (10,46). It will be interesting to determine whether other locus-specific differences in the pattern of hypermutation occur in the sharks, where the cluster type genomic arrangement (8) of antigen receptor loci may have enabled the disassociation of the different layers of the mechanism in different loci.

Acknowledgements

We are grateful to Michael Neuberger for helpful suggestions. This work was supported by NIH grant RR06603. M. D. is a Burroughs Wellcome Fund Fellow of The Life Sciences Research Foundation.

Abbreviations

BER	base excision repair
C	constant region
CDR	complementarity determining region
FR	framework region
JC	intron between the joining and the constant regions
NAR	nurse shark or new antigen receptor
Sec	secretory
Tm	transmembrane
V	variable region

References

1 Tsiagbe, V. K., Inghirami, G. and Thorbecke, G. J. 1996. The physiology of germinal centers. *Crit. Rev. Immunol.* 16:381.
2 Golding, G. B., Gearhart, P. S. and Glickman, B. W. 1987. Pattern of somatic mutations in immunoglobulin variable genes. *Genetics* 115:169.
3 Betz, A. G., Rada, C., Pannell, R., Milstein, C. and Neuberger, M. S. 1993. Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hotspots. *Proc. Natl Acad. Sci. USA* 90:2385.
4 Betz, A. G., Neuberger, M. S. and Milstein, C. 1993. Discriminating intrinsic and antigen-selected mutational hotspots in immunoglobulin V genes. *Immunol. Today* 14:405.
5 Dorner, T., Brezinscheck, H.-P., Brezinscheck, R. I., Foster, S. J., Domiati-Saad, R. and Lipsky, P. E. 1997. Analysis of the frequency and pattern of somatic mutations with nonproductively rearranged human variable heavy chain genes. *J. Immunol.* 158:2779.
6 Insel, R. A. and Varade, W. S. 1994. Bias in somatic hypermutation of human Vh genes. *Int. J. Immunol.* 5:1437.
7 Wilson, M., Hsu, E., Marcuz, A., Courtet, M., DuPasquier, L. and Steinberg, C. 1992. What limits affinity maturation of antibodies in *Xenopus*—the rate of somatic mutation or the ability to select mutants? *EMBO J.* 11:4337.
8 Hinds-Frey, K. R., Nishikata, H., Litman, R. T. and Litman, G. W. 1993. Somatic variation precedes extensive diversification of germline sequences by combinatorial joining in the evolution of immunoglobulin heavy chain diversity. *J. Exp. Med.* 178:815.
9 Anderson, M. K., Shamblott, M. J., Litman, R. T. and Litman, G. W. 1995. Generation of immunoglobulin light chain diversity in *Raja erinacea* is not associated with somatic rearrangement, an exception to a central paradigm of B cell immunity. *J. Exp. Med.* 182:109.
10 Diaz, M. and Flajnik, M. F. 1998. Evolution of somatic hypermutation and gene conversion in adaptive immunity. *Immunol. Rev.* 162:13.
11 Greenberg, A. S., Avila, D., Hughes, M., Hughes, A., McKinney, E. C. and Flajnik, M. F. 1995. A new antigen receptor gene family that undergoes rearrangement and extensive somatic diversification in sharks. *Nature* 374:168.
12 Reynaud, C.-A., Mackay, C. R., Muller, R. G. and Weill, J.-C. 1991. Somatic generation of diversity in a mammalian primary lymphoid organ: the sheep ileal Peyer's patches. *Cell* 64:995.
13 Diaz, M., Greenberg, A. S. and Flajnik, M. F. 1998. Somatic hypermutation of the new antigen receptor gene (NAR) in the nurse shark does not generate the repertoire: possible role in antigen-driven reactions in the absence of germinal centers. *Proc. Natl Acad. Sci. USA* 95: in press.
14 Wagner, S. D., Milstein, C. and Neuberger, M. S. 1995. Codon bias targets mutation. *Nature* 376:732.
15 Brown, W. M., Prager, E. M., Wang, A. and Wilson, A. C. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. *J. Mol. Evol.* 18:225.

Downloaded from https://academic.oup.com/intimm/article/11/5/825/710480 by guest on 21 August 2022

16 Aquadro, C. F. and Greenberg, B. D. 1983. Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics* 103:287.

17 Wu, C.-I. and Maeda, N. 1987. Inequality in mutation rates of the two strands of DNA. *Nature* 327:169.

18 Li, W.-H., Wu, C.-I. and Luo C.-C. 1984. Nonrandomness of point mutation as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. *J. Mol. Evol.* 21:58.

19 Wilson, P. C., de Bouteiller, O., Liu, Y. J., Potter, K., Banchereau, J., Capra, J. D. and Pascual, V. 1998. Somatic hypermutation introduces insertions and deletions into immunoglobulin V genes. *J. Exp. Med.* 187:59.

20 Smith, D. S., Creadow, G., Jena, P. K., Portanova, J. P., Kotzin, B. L. and Wysocki, L. J. 1996. Di- and trinucleotide target pReferences of somatic mutagenesis in normal and autoreactive B cells. *J. Immunol.* 156:2642.

21 Laken, S. J., Petersen, G. M., Gruber, S. B., Oddoux, C., Ostrer, H., Giardiello, F. M., Hamilton, S. R., Hampel, H., Markowitz, A., Klimstra, D., Jhanwar, S., Winawer, S., Offit, K., Luce, M. C., Kinzler, K. W. and Vogelstein, B. 1997. Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. *Nat. Genet.* 17:79.

22 Gonzalez-Fernandez, A., Gupta, S. K., Pannell, R., Neuberger, M. S. and Milstein, C. 1994. Somatic mutation of immunoglobulin λ chains: A segment of the major intron hypermutates as much as the complementarity-determining regions. *Proc. Natl Acad. Sci. USA* 91:12614.

23 Wysocki, L. J., Zhang, X., Smith, D. J., Snyder, C. M. and Bonocino, C. 1998. Somatic origin of T-cell epitopes within antibody variable regions: significance to monoclonal therapy and genesis of systemic autoimmune disease. *Immunol. Rev.* 162:233.

24 Dorner, T., Foster, S. S., Brezinscheck, H.-P. and Lipsky, P. E. 1998. Analysis of the targeting of the hypermutational machinery and the impact of subsequent selection on the distribution of nucleotide changes in human VhDJh rearrangements. *Immunol. Rev.* 162:161.

25 Storb, U., Klotz, E. L., Hackett, J., Kage, K., Bozek, G. and Martin, T. E. 1998. A hypermutable insert in an immunoglobulin transgene contains hotspots of somatic mutation and sequences predicting highly stable structures in the RNA transcript. *J. Exp. Med.* 188:689.

26 Winter, D. B., Phung, Q. H., Umar, A., Baker, S. M., Tarone, R. E., Tanaka, K., Liskay, R. M., Kunkel, T. A., Bohr, V. A. and Gearhart, P. J. 1998. Altered spectra of hypermutation in antibodies from mice deficient for the DNA mismatch repair protein PMS2. *Proc. Natl Acad. Sci. USA* 85:6953.

27 Brenner, S. and Milstein, C. 1966. Origin of antibody variation. *Nature* 211:242.

28 Peters, A. and Storb, U. 1996. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity* 4:57.

29 Bertocci, B., Quint, L., Delbos, F., Garcia, C., Reynaud, C.-A. and Weill, J.-C. 1998. Probing immunoglobulin genes hypermutation with microsatellites suggests a nonreplicative short patch DNA synthesis process. *Immunity* 9:257.

30 Cascalho, M., Wong, J., Steinberg, C. and Wabl, M. 1998. Mismatch repair co-opted by hypermutation. *Science* 279:1207.

31 Winter, D. B., Sattar, N. and Gearhart, P. J. 1998. The role of promoter-intron interactions in directing hypermutation. *Curr. Top. Microbiol. Immunol.* 229:1–10.

32 Kunkel, T. A. 1985. The mutational specificity of DNA polymerase-β during *in vitro* DNA synthesis. *The J. Biochem.* 260:5787.

33 Napolitano, R. L., Lambert, I. B. and Fuchs, R. P. P. 1997. SOS factors involved in translesion synthesis. *Proc. Natl Acad. Sci. USA* 94:5733

34 Nelson, J. R., Lawrence, C. W. and Hinckle, D. C. 1996. Thymine–thymine dimer bypass by yeast DNA polymerase ζ. *Science* 272:1646.

35 Holbeck, S. L. and Strathern, J. N. 1997. A role for REV3 in mutagenesis during double-strand break repair in *Saccharomyces cerevisiae*. *Genetics* 147:1017.

36 Quah, S.-K., von Borstel, R. C. and Hastings, P. J. 1980. The origin of spontaneous mutations in *Saccharomyces cerevisiae*. *Genetics* 96:819.

37 Hohman, S. and Gozalbo, D. 1989. Comparison of the nucleotide sequence of a yeast gene family I. Distribution and spectrum of spontaneous base substitutions. *Mutat. Res.* 215:79.

38 Roche, H., Gietz, R. D. and Kunz, B. A. 1995. Specificities of the *Saccharomyces cerevisiae rad6*, *rad18*, and *rad52* mutators exhibit different degrees of dependence on the *rev3* gene product, a putative nonessential DNA polymerase. *Genetics* 140:443.

39 Glassner, B. J., Rasmussen, L. J., Najarian, M. T., Psnick, L. M. and Samson, L. D. 1998. Generation of a strong mutator phenotype in yeast by imbalanced base excision repair. *Proc. Natl Acad. Sci. USA* 95:9997.

40 Kajiwara, K., Nagawawa, H., Shimizu-Nishikawa, S., Ookuri, T., Kimura, M. and Sugaya, E. 1996. Molecular characterization of seizure-related genes isolated by differential screening. *Biochem. Biophys. Res. Commun.* 219:795.

41 Gibbs, P. E. M., McGregor, W. G., Maher, V. M., Nisson, P. and Lawrence, C. W. 1998. A human homolog of the *Saccharomyces cerevisiae* REV3 gene, which encodes the catalytic subunits of DNA polymerase ζ. *Proc. Natl Acad. Sci. USA* 95:6876.

42 Sobol, R. W., Horton, J. K., Kuhnr, R., Gu, H., Singhal, R. K., Prasad, R., Rajewsky, K. and Wilson, S. H. 1996. Requirement of mammalian DNA polymerase beta in base excision repair. *Nature* 379:183.

43 Jacobs, H., Fukita, Y., van der Host, G. T. L., deBoer, J., Weeda, G., Esser, J., deWind, N., Engelward, B. P., Samson, L., Verbeck, S., Menissier de Murcia, J., de Murcia, G., te Riele, H. and Rajewsky, K. 1998. Hypermutation of immunoglobulin genes in memory B-cells of DNA repair-deficient mice. *J. Exp. Med.* 187:1735.

44 Kuo, F. C. and Sklar, J. 1997. Augmented expression of a human gene for 8-oxoguanine DNA glycosylases (*MutM*) in B lymphocytes of the dark zone in lymph node germinal centers. *J. Exp. Med.* 186:1547.

45 Phung, Q. H., Winter, D. B., Cranston, A., Tarone, R. E., Bohr, V. A., Fishel, R. and Gearhart, P. J. 1998. Increase hypermutation at G and C nucleotides in immunoglobulin variable genes from mice defficient for the MSH2 mismatch repair protein. *J. Exp. Med.* 187:1745.

46 Rada, C., Ehrenstein, M. R., Neuberger, M. S. and Milstein, C. 1998. Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. *Immunity* 9:135.

Downloaded from https://academic.oup.com/intimm/article/11/5/825/710480 by guest on 21 August 2022