### REPORTS

of all J<sub>H</sub> gene segments (8). Because QM

mice are also homozygous for a gene-targeted deletion of  $J_{\kappa}$  (9), they produce only  $\lambda$ -type

light (L) chains. When combined with any  $\lambda$ 

chain, H chains with an unmodified V exon can be detected by an antibody to the idio-

gene Pms2 in somatic hypermutation, we

first crossed the targeted "knockout" allele  $Pms2^{ko}$  (6) into mice with the QM geno-

type. We will refer to these mice by their configuration at the Pms2 locus: Pms2<sup>ko</sup>/

 $Pms2^{ko}$  (or ko/ko) for homozygous, DNA-

mismatch repair-deficient, QM mice, +/+for standard repair-proficient QM mice, and  $Pms2^{ko}/Pms^+$  (or ko/+) for heterozy-

gous QM mice. The configurations at all Ig

loci are the same for all three types of mice.

must be kept in mind in order to understand

and interpret the experiments to be de-

scribed here. (i) The number of germline

elements contributing to the primary Ig rep-

ertoire-the Ig repertoire developed after

Several properties of the QM mouse

To study the role of the mismatch repair

type (7, 10).

## Mismatch Repair Co-opted by Hypermutation

Marilia Cascalho, Jamie Wong, Charles Steinberg, Matthias Wabl\*

Mice homozygous for a disrupted allele of the mismatch repair gene Pms2 have a mutator phenotype. When this allele is crossed into guasi-monoclonal (QM) mice, which have a very limited B cell repertoire, homozygotes have fewer somatic mutations at the immunoglobulin heavy chain and  $\lambda$  chain loci than do heterozygotes or wild-type QM mice. That is, mismatch repair seems to contribute to somatic hypermutation rather than stifling it. It is suggested that at immunoglobulin loci in hypermutable B cells, mismatched base pairs are "corrected" according to the newly synthesized DNA strand, thereby fixing incipient mutations instead of eliminating them.

Somatic point mutations are introduced at a high rate into the exon encoding the variable (V) region of an immunoglobulin (Ig) and into its flanking sequences (1). The process of somatic hypermutation at the Ig loci is a site-specific, differentiation stagespecific, and lineage-specific phenomenon that contributes to the generation of antibody diversity and affinity maturation (2). It has often been speculated that DNA repair has something to do with hypermutation (3). In one sense, this speculation is tautological; clearly, if all introduced mutations were repaired, no mutations would be seen. One possibility is that the Ig mutator system might deactivate some or all of the error-free repair systems that normally function in eukaryotic cells. A second possibility, however, is that some error-free repair systems might be modified in such a way that they actually introduce mutations rather than correct them. Here we report experiments that demonstrate that the latter hypothesis is true for at least one repair system.

Using the working hypothesis that the Ig mutator system initially introduces mismatches during the replication of one DNA strand, we have studied the role of mismatch repair in hypermutation. The mutSmutL mismatch repair pathway in Escherichia coli has a broad specificity, and its inactivation results in a mutator phenotype (4). Human homologs of mutS and mutL have been implicated in hereditary nonpolyposis colorectal cancer (5). Mice homozygous for a gene-targeted defective allele of Pms2, a mutL homolog, develop tumors more frequently than do wild-type mice, and homozygous males exhibit abnormal chromosome synapsis in meiosis (6).

We developed the quasi-monoclonal (QM) mouse for studying Ig diversification by somatic hypermutation as well as by other

Department of Microbiology and Immunology, University of California, San Francisco, CA 94143-0670, USA

\*To whom correspondence should be addressed. E-mail: mutator@itsa.ucsf.edu

known and unknown mechanisms (7). QM mice are heterozygous at the heavy (H) chain locus: On one allele, the stretch of DNA containing the  $J_{\rm H}$  segments has been replaced by a targeted, rearranged  $V_{\rm H} D J_{\rm H}$ exon encoding an H-chain V region bearing the 17.2.25 idiotype (7); the other allele is nonfunctional because of a targeted deletion

Fig. 1. Flow cytometric analysis of peripheral blood lymphocytes from QM mice heterozygous (ko/+, left) or homozygous (ko/ko, right) for the disrupted Pms2ko allele. Ordinates, fluorescein isothiocyanate-conjugated rat monoclonal antibody to B220 (Pharmingen 01124A); abscissae, biotin-coupled antibody to 17.2.25 V<sub>H</sub>DJ<sub>H</sub> idiotype (monoclonal rat antibody



R2.438.8), followed by phycoerythrin-coupled streptavidin (Pharmingen 13025D). Peripheral blood lymphocytes were prepared from total blood by hemolysis in a standard NH<sub>2</sub>Cl lysis buffer. Incubations with the antibodies were done for 30 min at 4°C before analysis and sorting. Dead cells were excluded by propidium iodide staining. Boxes 1 and 2 indicate the windows used for sorting

+/+

Fig. 2. Somatic mutations in a 21-nt sequence contained in Ju4 alleles derived from B220-positive, idiotype-negative B cells of QM mice (+/+, left) or QM mice homozygous for the disrupted Pms2ko allele (ko/ko, right). The sequence starts with the triplet CCT, which represents the N region of the knocked-in gene. Seguences were determined as described (6, 7).

	Pro CCT	Tyr TAC	Tyr TAT	Ala GCT	Met ATG	Asp GAC	Tyr TAC
1		C		T			
2			C				
3				A			
4				A			
5				-G-			
6				-T-			
7	C	-TT		T			-T-
8							
9					T-T		
10							
11				-G-			
12			C				
13	C		-T-	-G-	T		
14			-				
15			-				
16							

#### ko/ko

Pro	Tyr	Tyr	Ala	Met	Asp	Tyr
CCT	TAC	TAT	GCT	AIG	GAC	TAC
		C				
C		-T-	-TC			

standard V(D)J rearrangement—is highly limited in QM mice; only one VH, one D, one  $J_H$ , three  $V_{\lambda}$ , and three  $J_{\lambda}$  are involved. This property is an advantage for identifying

mutations, because the germline sequences are known. (ii) Because of the low diversity of the primary repertoire, there is an enormous selective pressure for Ig diversification.

Thus, V-gene replacements, which would ordinarily be of little consequence, become frequent. This property is an advantage for studying V-gene replacement, but it negates

	107										137										167						
	Glu	Thr	Val	Thr	Leu	Thr	Cys	Arg	Ser	Ser	Thr	Gly	Ala	Val	Thr	Thr	Ser	Asn	Tyr	Ala	Asn	Trp	Val	Gln	Glu	Lys	Pro
	GAA	ACA	GIC	ACA	CIC	ACT	TGT	CGC	TCA	AGT	ACT	666	GCT	GIT	ACA	ACT	AGT	AAC	TAT	GUU	AAC	TGG	GIC	CAA	GAA	AAA	CCA
ko/+																											
091937																											
091957		G													G												
081546																	-A-										
090317	G																G										
100247		G			G					-A-							-A-						A				
100203																						C					
100215	G																G										
100832																											
100827																											
100822																											
100840																											
ko/ko																											
092358																											
090342																											
090350																											
081532																											
081536																											
081529																											
081528																											
100141																											
100124																											
100158																											
100726																											
100103																											
100727																											
				197										227										257			
	Asp	His	Leu	197 Phe	Thr	Gly	Leu	Ile	Gly	Gly	Thr	Asn	Asn	227 Arg	Ala	Pro	Gly	Val	Pro	Ala	Arg	Phe	Ser	257 Gly	Ser	Leu	Ile
	Asp GAT	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA	Gly GGT	Gly GGT	Thr ACC	Asn AAC	Asn AAC	227 Arg CGA	Ala GCT	Pro CCA	Gly GGT	Val GTT	Pro CCT	Ala GCC	Arg AGA	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT
ko/+	Asp GAT	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA	Gly GGT	Gly GGT	Thr ACC	Asn AAC	Asn AAC	227 Arg CGA	Ala GCT	Pro CCA	Gly GGT	Val GTT	Pro CCT	Ala GCC	Arg AGA	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT
<b>ko/+</b> 091937	Asp GAT	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA	Gly GGT	Gly GGT	Thr ACC	Asn AAC	Asn AAC	227 Arg CGA	Ala GCT	Pro CCA	Gly GGT	Val GTT	Pro CCT	Ala GCC	Arg AGA	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT
<b>ko/+</b> 091937 091957	Asp GAT  A	His CAT	Leu TTA 	197 Phe TTC	Thr ACT 	Gly GGT 	Leu CTA 	Ile ATA 	Gly GGT  A	Gly GGT	Thr ACC	Asn AAC  -G-	Asn AAC 	227 Arg CGA	Ala GCT  -G-	Pro CCA	Gly GGT 	Val GTT  A	Pro CCT	Ala GCC 	Arg AGA  -A-	Phe TTC 	Ser TCA 	257 Gly GGC	Ser TCC	Leu CTG 	Ile ATT C -G-
<b>ko/+</b> 091937 091957 081546	Asp GAT  A	His CAT 	Leu TTA 	197 Phe TTC	Thr ACT 	Gly GGT	Leu CTA 	Ile ATA  C	Gly GGT  A	Gly GGT	Thr ACC	Asn AAC  -G-	Asn AAC	227 Arg CGA 	Ala GCT  -G-	Pro CCA	Gly GGT 	Val GTT A	Pro CCT	Ala GCC 	Arg AGA  -A-	Phe TTC	Ser TCA 	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-
<b>ko/+</b> 091937 091957 081546 081547 090317	Asp GAT  	His CAT	Leu TTA  	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA  	Ile ATA C 	Gly GGT A 	Gly GGT	Thr ACC	Asn AAC  	Asn AAC	227 Arg CGA  C 	Ala GCT -G- 	Pro CCA	Gly GGT	Val GTT A 	Pro CCT	Ala GCC	Arg AGA  -A- 	Phe TTC	Ser TCA 	257 Gly GGC	Ser TCC	Leu CTG  	Ile ATT C -G- 
<b>ko/+</b> 091937 091957 081546 081547 090317 100247	Asp GAT A 	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA C 	Gly GGT A 	Gly GGT	Thr ACC	Asn AAC   G 	Asn AAC	227 Arg CGA  C 	Ala GCT -G-  A	Pro CCA	Gly GGT	Val GTT A A A A	Pro CCT	Ala GCC	Arg AGA  -A-  	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG   	Ile ATT C -G- 
<b>Ko/+</b> 091937 091957 081546 081547 090317 100247 100203	Asp GAT A  	His CAT	Leu TTA    	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA C  	Gly GGT A  A A	Gly GGT	Thr ACC  T G	Asn AAC  G G  G	Asn AAC	227 Arg CGA  C  	Ala GCT -G-  A -T-	Pro CCA	Gly GGT   -A- 	Val GTT A A A A	Pro CCT	Ala GCC	Arg AGA    	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-  
<b>ko/+</b> 091937 091957 081546 081547 090317 100247 100203 100215	Asp GAT    	His CAT	Leu TTA     	197 Phe TTC	Thr ACT	Gly GGT    	Leu CTA	Ile ATA C 	Gly GGT A  A A	Gly GGT	Thr ACC  T G T T	Asn AAC  G G G G G	Asn AAC	227 Arg CGA  C  	Ala GCT  A -T- -T-	Pro CCA	Gly GGT   -A-  -A-	Val GTT A A A A A	Pro CCT	Ala GCC	Arg AGA     	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-   
<b>ko/+</b> 091937 081546 081547 090317 100247 100203 100215 100833	Asp GAT A   	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA C   	Gly GGT A  A A A 	Gly GGT	Thr ACC  T G T T T	Asn AAC  G G G G G	Asn AAC	227 Arg CGA    	Ala GCT -G-  A -T- 	Pro CCA	Gly GGT  -A-  -A- 	Val GTT A A A A A A	Pro CCT	Ala GCC	Arg AGA     	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-    
<b>ko/+</b> 091937 091957 081546 081547 090317 100247 100203 100215 100833 100832	Asp GAT A    	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA C    	Gly GGT A  A A  A 	G1y GGT	Thr ACC	Asn AAC  G G G G G G 	Asn AAC	227 Arg CGA     	Ala GCT -G-  A -T-  	Pro CCA	Gly GGT	Val GTT A A A A A A 	Pro CCT	Ala GCC	Arg AGA      	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-     
<b>ko/+</b> 091937 091957 081546 081547 090317 100247 100203 100215 100833 100832 100827	Asp GAT A     	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT       	Leu CTA	Ile ATA	Gly GGT A  A A  A 	Gly GGT	Thr ACC	Asn AAC  G G G G  G  	Asn AAC	227 Arg CGA  C     	Ala GCT  A -T-   	Pro CCA	Gly GGT  -A-  -A-  -A-  	Val GTT A A A A A A A	Pro CCT	Ala GCC	Arg AGA       	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-      
<b>ko/+</b> 091937 091957 081546 081547 090317 100203 100215 100833 100832 100822 100822	Asp GAT       	His CAT       	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA C      	Gly GGT A  A A   	Gly GGT       	Thr ACC	Asn AAC  G  G  G  -	Asn AAC	227 Arg CGA C C     	Ala GCT -G-  A -T-    	Pro CCA	Gly GGT  -A- -A-  -A-  -A-  	Val GTT A A A A A A A	Pro CCT	Ala GCC	Arg AGA 	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-      
<b>ko/+</b> 091937 091957 081546 081547 090317 100247 100203 100215 100833 100832 100827 100822 100840 <b>ko/ko</b>	Asp GAT       	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT       	Leu CTA	Ile ATA C      	Gly GGT A  A A A  	Gly GGT       	Thr ACC	Asn AAC	Asn AAC	2227 Arg CGA 	Ala GCT -G-  A -T-    	Pro CCA	Gly GGT  -A-  -A-  -A-  	Val GTT A A A A A  A  A  	Pro CCT	Ala GCC	Arg AGA A-      	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-        
<b>ko/+</b> 091937 091957 081546 081547 090317 100203 100215 100833 100832 100822 100822 100840 <b>ko/ko</b> 092358	Asp GAT A      	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT      	Leu CTA	Ile ATA C      	Gly GGT A  A A   	Gly GGT      	Thr ACC	Asn AAC	Asn AAC	2227 Arg CGA	Ala GCT -G-  A -T-    	Pro CCA	Gly GGT  -A-  -A-   	Val GTT A A A A A  A 	Pro CCT	Ala GCC	Arg AGA A-      	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT -G-      
<b>ko/+</b> 091937 081546 081547 090317 100203 100215 100833 100832 100822 100840 <b>ko/ko</b> 092358 090342	Asp GAT       	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA C      	Gly GGT A  A    	Gly GGT       	Thr ACC	Asn AAC	Asn AAC	2227 Arg CGA 	Ala GCT -G-  A      	Pro CCA	Gly GGT  -A-  -A-    	Val GTT A  A A  A  -	Pro CCT	Ala GCC	Arg AGA       	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-      
<b>ko/+</b> 091937 081546 081547 090317 100247 100203 100812 100832 100822 100822 100840 <b>ko/ko</b> 092358 090342 090350	Asp GAT       	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA C      	Gly GGT A  A    	Gly GGT	Thr ACC	Asn AAC	Asn AAC	2227 Arg CGA 	Ala GCT -G- -T- -T- 	Pro CCA	Gly GGT       	Val GTT A  A A  A  -	Pro CCT	Ala GCC	Arg AGA       	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-      
<b>ko/+</b> 091937 091957 081546 081547 090317 100247 100203 100832 100832 100822 100840 <b>ko/ko</b> 092358 090342 090350 090347	Asp GAT       	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATA C      	Gly GGT A  A    	Gly GGT       	Thr ACC	Asn AAC  G G G   	Asn AAC	2277 Arg CGA	Ala GCT -G-  A -T-    	Pro CCA	Gly GGT  -A-  -A-    	Val GTT A A A    	Pro CCT	Ala GCC	Arg AGA A-      	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C       -
<b>ko/+</b> 091937 081546 081547 090317 100247 100203 100812 100832 100822 100840 <b>ko/ko</b> 092358 090342 090350 090347 081532 081536	Asp GAT       	His CAT	Leu TTA	197 Phe Trc	Thr ACT	Gly GGT	Leu CTA	Ile ATA C      	Gly GGT A  A    	Gly GGT       	Thr ACC	Asn AAC  G G G   	Asn AAC	2277 Arg CGA	Ala GCT -G-  A -T-    	Pro CCA	Gly GGT  -A-      -	Val GTT A A A    	Pro CCT	Ala GCC	Arg AGA       	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C G-        
<b>ko/+</b> 091937 091957 081546 081547 090317 100247 100203 100215 100832 100822 100822 100840 <b>ko/ko</b> 092358 090342 090350 090347 081532 081536 081529	Asp GAT	His CAT	Leu TTA	197 Phe Trc	Thr ACT       	Gly GGT       	Leu CTA	Ile ATACC	Gly GGT A      	Gly GGT       	Thr ACC	Asn AAC	Asn AAC	2277 Argg CGA	Ala GCT -G- -T- -T- -T- 	Pro CCA	Gly GGT  -A-  -A-    	Val GTT A A A A A A A	Pro CCT	Ala GCC	Arg AGA A-      	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C G-        
<b>ko/+</b> 091937 091957 081546 081547 090317 100247 100203 100832 100832 100822 100840 <b>ko/ko</b> 092358 090342 090350 090347 081532 081529 081528	Asp GAT       	His CAT	Leu TTA	197 Phe TTC	Thr ACT       	Gly GGT       	Leu CTA	Ile ATACC	Gly GGT A      	Gly GGT       	Thr ACC  G T T    	Asn AAC	Asn AAC	2277 Argg CGA	Ala GCT -G- -T- -T- -T- 	Pro CCA	Gly GGT  -A-  -A-    	Val GTT A A A A A A  A        	Pro CCT	Ala GCC	Arg AGA	Phe TTC	Ser TCA	257 Gly GGC       	Ser TCC	Leu CTG	Ile ATT C -G-        
<i>ko/+</i> 091937 091957 081546 081547 090317 100247 100203 100832 100822 100822 100840 <i>ko/ko</i> 092358 090342 090350 090342 090350 090347 081532 081528 091904	Asp GAT       	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT       	Leu CTA	Ile ATAC	Gly GGT A  A  A  	Gly GGT 	Thr ACC	Asn -G-   G  G  	Asm AAC	2277 Argg CGA C C       	Ala GCT       	Pro CCA	Gly GGT       	Val GTT A A A  A  	Pro CCT	Ala GCC	Arg AGA A-      	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-        
<b>ko/+</b> 091937 091957 081546 081547 090317 100203 100215 100833 100832 100827 100822 100840 <b>ko/ko</b> 092358 090342 090350 090347 081532 081532 081528 091904 100141 100124	Asp GAT       	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT       	Leu CTA	Ile ATAC	Gly GGT A  A    	Gly GGT       	Thr ACC	Asn -G-   G    	Asm AAC	2277 Argg CGA C C C 	Ala GCT G- T-     	Pro CCA	Gly GGT       	Val GTT   A  A  	Pro CCT	Ala GCC	Arg AGA       	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-      
<b>ko/+</b> 091937 091957 081546 081547 090317 100203 100823 100832 100822 100840 <b>ko/ko</b> 092358 090342 090350 090347 081532 081532 081528 091904 100141 100124	Asp GAT       	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT       	Leu CTA	Ile ATAC	Gly GGT A  A    	Gly GGT       	Thr ACC	Asn AAC	Asn AAC	2277 Arg CGA C C       	Ala GCT -G- -T- -T- -T- -T- -T- -T- -T- -T- -T	Pro CCA	Gly GGT       	Val GTT A A A    	Pro CCT	Ala GCC	Arg AGA	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-      
<b>ko/+</b> 091937 081546 081547 090317 100247 100203 100822 100822 100822 100822 100820 <b>ko/ko</b> 090340 090342 090350 090347 081532 081536 081529 081528 091904 100124 1001258 100726	Asp GAT       	His CAT	Leu TTA	197 Phe TTC	Thr ACT       	Gly GGT       	Leu CTA	Ile ATAC	Gly GGT A      	Gly GGT       	Thr ACC	Asn AAC	Asn AAC	227 Arg CGA	Ala GCT -G- -T- -T- -T-    	Pro CCA	Gly GGT  -A-      -	Val GTT A A A A  A  -	Pro CCT	Ala GCC	Arg AGA	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-      
<b>ko/+</b> 091937 091957 081546 081547 090317 100247 100203 100832 100832 100822 100840 <b>ko/ko</b> 092358 090342 090350 090347 081532 081536 081529 081528 091904 100141 100124 1001258 100726 100922	Asp GAT       	His CAT	Leu TTA	197 Phe TTC	Thr ACT	Gly GGT	Leu CTA	Ile ATAC	Gly GGT A      	Gly GGT       	Thr ACC	Asn AAC	Asn AAC	227 Arg CGA	Ala GCT G- T- T- T-   	Pro CCA	Gly GGT  -A-  -A-    	Val GTT A A A A A A A	Pro CCT	Ala GCC	Arg AGA	Phe TTC	Ser TCA	257 Gly GGC	Ser TCC	Leu CTG	Ile ATT C -G-      

**Fig. 3** (above and on facing page). Comparison of the  $\lambda_1$  alleles derived from B220-negative, idiotype-positive B cells (window 2 in Fig. 1) of QM mice heterozygous (*ko*/+, upper set) or homozygous (*ko*/*ko*, lower set) for the disrupted *Pms2<sup>ko</sup>* allele. The 317-nt sequences contain most of V $\lambda_1$ . The contiguous cDNA sequence is numbered starting with 1 at the first nucleotide of the ATG start codon. Nucleotide identity is indicated by a dash (-); a blank indicates not read. PCR amplification was done from cDNA with 5 pmol of the V $\lambda_1$  5' primer 5'-GGAATTCCTGCACTCACCACATCACCTGG and C $\lambda_1$ 

3' primer 3'-GGATCCTACCTTCCAGTCCACTGTCACC. The reactions were done in a total volume of 50  $\mu$ l with 2.5 U of Taq polymerase and 1× PCR buffer with 1.5 mM MgCl<sub>2</sub> and 200  $\mu$ M dNTPs. The program consisted of 40 cycles of 1 min at 94°C, 1 min at 59°C, and 2 min at 72°C. The last cycle included a prolonged extension step (10 min) to favor the A addition necessary for cloning the PCR fragments into the Invitrogen pCR 2.1 vector. Double-stranded DNA was prepared from clones and sequenced.

part of the advantage of the limited primary repertoire for studying hypermutation. (iii) With flow cytometry, we can identify several B cell populations in the peripheral blood of QM mice (7, 11). A naïve IgM-positive population has never been activated, and the Ig loci harbor few if any somatic mutations. Another population, which falls within window 1 (Fig. 1), consists of idiotype-negative cells, most of which have modified their  $V_H DJ_H$  exon by  $V_H$  gene replacement and somatic hypermutation; this population contains a substantial proportion of cells expressing IgM (7, 11). A third population, which falls within window 2 (Fig. 1), consists of idiotype-positive cells, most of which (>95%) have switched their H-chain isotypes and show numerous somatic mutations in their  $V_H DJ_H$  exons (12). (iv) These mutations have been selected in response to environmental antigens. As with the V-gene replacements, they result from a strong selective pressure due to the limited primary repertoire-a phenomenon that has been termed hyperselection (11). This means that Ig genes encoding antibodies against a whole panoply of antigens are studied, rather than those against a single antigen as in conventional immune response experiments.

With flow cytometry, we compared the cell population profiles of four ko/ko and four ko/+ QM mice with fluorescence-labeled antibodies specific for the following pairs of markers: B220 versus the idiotype,  $\mu^{a}$  versus the idiotype, B220 versus  $\lambda$ , and  $\mu^{a}$  versus  $\lambda$ . In addition, another four ko/ko and four ko/+ QM mice were analyzed only with the  $\mu^{a}$ -versus-idiotype pair of reagents.

No appreciable qualitative or quantitative differences among the two genotypes were apparent. Figure 1 shows a typical pattern with the B220-versus-idiotype pair.

In a pilot experiment, we compared the H-chain alleles from sorted idiotype-negative B cells of ko/ko and +/+ QM mice. From the sorted cells we isolated mRNA and reverse-transcribed it to get cDNA, then amplified H-chain sequences by polymerase chain reaction (PCR) and cloned and sequenced them. Because the  $V_{\rm H}$  segments from gate 1 cells all differ, and the D segments are too near to the recombination breakpoints resulting from the  $\mathrm{V}_\mathrm{H}$  gene replacement process, we analyzed a short 21nucleotide (nt) stretch of  $J_H^4$  sequence, which is shared by all alleles (Fig. 2). In the +/+ mice, there were 20 mutations in a total of 306 nt, yielding a frequency of 6.5%; in the ko/ko mice, there were 5 mutations in 478 nt, yielding a frequency of 1.0%. That is, mice deficient in the mismatch repair enzyme actually had a lower mutation frequency than did wild-type mice. This seemingly paradoxical result suggested to us that mismatch repair is a component rather than an antagonist of the Ig hypermutation system.

In a similar experiment, we analyzed the preformed  $V_H DJ_H$  exon in idiotype-positive cells. Because this population has not undergone  $V_H$  replacement, we could analyze a longer sequence per allele—310 nt covering part of  $V_H$ , all of D, and part of  $J_H$  (7). Alleles from the +/+ mice had 112 mutations in a total of 3086 nt, yielding a frequency of 3.6%; the *ko/ko* mice had 2 mutations in 1240 nt, or 0.16%.

317

287

We also investigated the  $\lambda_1$  L-chain locus of ko/ko and ko/+ QM mice in the idiotype-positive cell population defined by window 2 (Fig. 1). We sorted peripheral blood B cells from pools of six mice of each genotype, isolated mRNA, reverse-transcribed it into cDNA, then amplified  $\lambda_1$ L-chain sequences, which were cloned and sequenced (Fig. 3). The sequences cover a stretch of 237 nt covering most of  $V\lambda_1$ ; they have been truncated well short of the VJ junction to avoid any chance of confusion with junctional diversity (13). In ko/+ OM mice, there were 49 mutations in 3056 nt, yielding a frequency of 1.6%. In ko/ko QM littermates, there were only 7 mutations in a total of 3675 nt, yielding a frequency of 0.2%. Thus, the absence of mismatch repair also leads to a decrease in hypermutation at the  $\lambda$  L-chain locus.

In other similar experiments we also examined  $\lambda_1$  sequences obtained from idiotype-negative cells. In the same 237 nt stretch of V $\lambda_1$  used above, there were 21 mutations in 4977 nt (a frequency of 0.4%) in ko/+ QM mice and only 2 mutations in 4392 nt (a frequency of 0.04%) in ko/ko QM mice. The lower frequency of mutations in the idiotype-negative cells was not unexpected. Because the idiotype-positive cells have retained the gene-targeted V<sub>H</sub>DJ<sub>H</sub> exon, antigen selection pressure ought to be stronger on the  $\lambda$  gene and thus result in a higher mutation frequency; indeed, this seems to be the case.

As stated above, it is obvious that if all pre-mutations introduced by the Ig mutator system were repaired, there would be no

	Gly GGA	Asp GAC	Lys AAG	Ala GCT	Ala GCC	Leu CTC	Thr ACC	Ile ATC	Thr ACA	Gly GGG	Ala GCA	Gln CAG	Thr ACT	Glu GAG	Asp GAT	Glu GAG	Ala GCA	Ile ATA	Tyr TAT	Phe TTC	Cys TGT	Ala GCT	Leu CTA	Trp TGG	Tyr TAC
ko/+																									
091937																									
091957												-G-						C-T							
081546																									
081547																				-C-					
090317		-G-	G																						
100247						-A-																			
100203									G		T														
100215															A										
100833																									
100832																									
100827																									
100822																									
100840																									
ko/ko																									
092358					A			G																	
090342												A									C				
090350																									
090347																									
081532									-																
081536																									
081529																									
081528																									
091904																									~
100141																									
100124																									
100158																									
100726																									
100922																									
100103				<u>`</u>																					
100727	-A-																								

somatic hypermutation. Perhaps the simplest way for the Ig mutator system to avoid that scenario would be to overwhelm the DNA repair system with so many mutations that they could not all be repaired before becoming fixed. But if that were the case, then Pms2ko/Pms2ko mice should have a higher frequency of mutations at the Ig loci, as they have at other loci (6, 14); in fact, the frequency is lower at the Ig loci. Another simple strategy for the Ig mutator system would be to throw a monkey wrench into the works to turn off or otherwise ensure that DNA repair was ineffective in hypermutable B cells. But if that were the case, then Pms2<sup>ko</sup>/Pms2<sup>ko</sup> mice should have the same frequency of mutations at the Ig loci; in fact, the frequency is lower. Yet another strategy would be for the Ig mutator system to co-opt the DNA repair system to subvert it to create rather than prevent mutations. The third strategy would seem to be the only one that would explain the results of the experiments described here. Of course, the above argument applies only to a mismatch repair system requiring *Pms2*. How the Ig mutator system deals with other DNA repair pathways can only be discovered by experimentation.

The prototypic mismatch repair system in E. coli corrects the newly synthesized DNA strand, which is transiently unmethylated (15), using the old methylated DNA strand as a template. In eukaryotic cells, the basis for strand repair bias is not well understood, although it may involve singlestrand breaks (16). We propose the following mechanism for the action of Pms2 at the Ig loci: After mismatches have been introduced at an Ig locus in hypermutable cells by an unknown mechanism, the mismatch repair system identifies the "wrong," mutated strand as a template and thus fixes the mutations. In mice without mismatch repair, this model, in its simplest version, predicts that at the next replication the old strand will give rise to a nonmutant allele, whereas the new strand will give rise to an allele with one or more mutations. Thus, the frequency of mutations would be reduced by one-half. Because we found rather lower mutation frequencies in the absence of repair than would be predicted by this basic model, it might require some elaboration. For example, other repair mechanisms might correct mismatches in the absence of the repair system involving Pms2, and this would reduce the mutation frequency to less than one-half.

It has been reported that many human tumors exhibit high mutation rates (17). We envisage that the co-option of mismatch repair that we have described here for Ig hypermutation may also play a role in some of these tumor phenotypes. That would require, however, that the co-option not be limited to Ig genes but have a broad scope of action.

#### **REFERENCES AND NOTES**

- M. G. Weigert, I. M. Cesari, S. J. Yonkovich, M. Cohn, *Nature* **228**, 1045 (1970); D. McKean et al., *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3180 (1984); C. A. Reynaud, L. Quint, B. Bertocci, J.-C. Weill, *Sem. Immunol.* **8**, 125 (1996).
- 2. M. Wabl and C. Steinberg, Curr. Opin. Immunol. 8, 89 (1996).
- S. Brenner and C. Milstein, *Nature* **211**, 242 (1966);
   M. Wabl *et al.*, *Immunol. Rev.* **96**, 91 (1986); N. Kim *et al.*, *J. Exp. Med.* **186**, 413 (1997).
- P. Nevers and H. Spatz, *Mol. Gen. Genet.* **139**, 233 (1975); P. J. Pułkkila, J. Peterson, G. Herman, P. Modrich, M. Meselson, *Genetics* **104**, 571 (1983); A.-L. Lu, S. Clark, P. Modrich, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4639 (1983).
- R. A. Fishel *et al.*, *Cell* **75**, 1027 (1993); F. S. Leach *et al.*, *ibid.*, p. 1215; C. E. Bronner *et al.*, *Nature* **368**, 258 (1994); N. C. Nicolaides *et al.*, *ibid.* **371**, 75 (1994); N. Papadopoulos *et al.*, *Science* **263**, 1625 (1994).
- 6. S. M. Baker et al., Cell 82, 309 (1995); L. Narayanan
- et al., Proc. Natl. Acad. Sci. U.S.A. 94, 3122 (1997).
- M. Cascalho, A. Ma, S. Lee, L. Masat, M. Wabl, Science 272, 1649 (1996).

- 8. J. Chen et al., EMBO J. 12, 821 (1993).
- 9. J. Chen et al., Int. Immunol. 5, 647 (1993).
- 10. This antibody was produced by T. Imanishi-Kari. Its
- detailed characterization has not yet been published.
  11. M. Cascalho, J. Wong, M. Wabl, *J. Immunol.* 159, 5795 (1997).
- M. Cascalho, J. Wong, M. Wabl, unpublished data.
   There is also some evidence for a polymorphism at the 3' end of Vλ<sub>1</sub>.
- N. de Wind *et al.*, *Cell* 82, 321 (1995); W. Edelmann *et al.*, *ibid.* 85, 1125 (1996).
- M. Meselson, in *Recombination of the Genetic Material*, K. B. Low, Ed. (Academic Press, San Diego, CA, 1988), pp. 91–113.
- J. Holmes et al., Proc. Natl. Acad. Sci. U.S.A. 87, 5837 (1990); D. C. Thomas et al., J. Biol. Chem. 266, 3744 (1991); P. Modrich and R. Lahue, Annu. Rev. Biochem 65, 101 (1996).
- 17. L. A. Loeb, Science 277, 1449 (1997).
- 18. We thank T. Imanishi-Kari for providing antibodies to the idiotype VH17.2.25; G. Wu and M. Liskay for comments on the manuscript; and S. Baker and M. Liskay for providing *Pms2* knockout mice. This work was supported by NIH grant 1R01 GM37699; by the Engalitcheff Award of the Arthritis Foundation; by funds from the Markey Trust; by a Howard Hughes grant for transgenic mice; and by a grant from the Junta Nacional de Investigação Científica e Tecnológica, Praxis XXI–BD 3763/94 to M.C.

1 December 1997; accepted 21 January 1998

# Pleiotropy and the Preservation of Perfection

David Waxman and Joel R. Peck\*

A mathematical model is presented in which a single mutation can affect multiple phenotypic characters, each of which is subject to stabilizing selection. A wide range of mutations is allowed, including ones that produce extremely small phenotypic changes. The analysis shows that, when three or more characters are affected by each mutation, a single optimal genetic sequence may become common. This result provides a hypothesis to explain the low levels of variation and low rates of substitution that are observed at some loci.

**M**any continuously varying phenotypic characters are subject to stabilizing selection, so that the optimal phenotypic value lies between the minimum and maximum possible values (1-7). These phenotypic characters can be anything from the circumference of the stem of a plant to the distance between two subunits within a protein. Models of stabilizing selection often allow for a continuous range of mutations, so that some mutations have very small effects, whereas others have substantial effects (2, 3, 8-14). We follow this approach in the present study. Analyses of stabilizing selection have concentrated on models for which any given mutation affects only one phenotypic character. Nevertheless, mutations that affect multiple characters are well known and are commonly regarded as ubiquitous (5, 14– 25). Here, we show that, when three or more characters are affected by each mutation, a single optimal genetic sequence may become predominant. This finding contrasts sharply with the usual finding that, in equilibrium, the optimal sequence is rare and many slightly suboptimal sequences are present.

Consider a simple nonpleiotropic model of viability selection in a very large population of haploid and asexual organisms (the results are expected to generalize to sex and diploidy). Parents produce offspring and then die, so that generations are discrete. After birth, offspring undergo viability selection, and the probability that an individual will survive depends on phenotype, which is described by measurements on k different characters. An individual's measurement on the *i*th character is denoted by  $z_i$  (where  $-\infty < z_i < \infty$ ). These characters are chosen so that they affect fitness independently, and  $z_i$ = 0 is the optimal value for each character.

D. Waxman, Centre for the Study of Evolution and Centre for Theoretical Physics, University of Sussex, Brighton, BN1 UK.

J. R. Peck, Centre for the Study of Evolution and School of Biological Sciences, University of Sussex, Brighton, BN1 9QG UK.

<sup>\*</sup>To whom correspondence should be addressed.