MOLECULAR AND CELLULAR BIOLOGY, July 1993, p. 4374-4381
0270-7306/93/074374-08$02.00/0
Copyright © 1993, American Society for Microbiology
Vol. 13, No. 7

Unusual Mutation Clusters Provide Insight into Class I Gene Conversion Mechanisms

LARRY R. PEASE,* ROBERT M. HORTON,† JEFFREY K. PULLEN,‡ AND THEODORE J. YUN

Department of Immunology and Department of Biochemistry and Molecular Biology, Mayo Graduate School and Mayo Clinic, Rochester, Minnesota 55905

Received 29 December 1992/Returned for modification 11 March 1993/Accepted 28 April 1993

Genetic diversity among the *K* and *D* alleles of the mouse major histocompatibility complex is generated by gene conversion among members of the class I multigene family. The majority of known class I mutants contain clusters of nucleotide changes that can be traced to linked family members. However, the details of the gene conversion mechanism are not known. The bm3 and bm23 mutations represent exceptions to the usual pattern and provide insight into intermediates generated during the gene conversion process. Both of these variants contain clusters of five nucleotide substitutions, but they differ from the classic conversion mutants in the important respect that no donor gene for either mutation could be identified in the parental genome. Nevertheless, both mutation clusters are composed of individual mutations that do exist within the parent. Therefore, they are not random and appear to be templated. Significantly, the bm3 and bm23 mutation clusters are divided into overlapping regions that match class I genes which have functioned as donor genes in other characterized gene conversion events. The unusual structure of the mutation clusters indicates an underlying gene conversion mechanism that can generate mutation clusters as a result of the interaction of three genes in a single genetic event. The unusual mutation clusters are consistent with a hypothetical gene conversion model involving extrachromosomal intermediates.

The spontaneous mutations in the class I major histocompatibility complex (MHC) genes of mice can be distinguished from random point mutations seen in other genes by two criteria: the mutations generally occur in clusters of multiple nucleotide substitutions, and they contain sequences which also exist in other class I genes (19, 26). In most cases, a donor gene which contains the exact sequence of the mutated region can be identified among cloned members of the class I multigene family isolated from the haplotype that gave rise to the mutant (22). Each of the known single-nucleotide mutations found among class I genes also can be found in other class I genes of the parental haplotype; therefore, this simpler class of variant also appears to arise by the same mechanism (9–11, 16). Together, these observations gave rise to the gene conversion hypothesis, i.e., that sequences from a donor gene are transferred into the parent gene to produce a mutant.

Few details of the mutagenic mechanism are available. Although the observed mutation frequency is high relative to mutation frequencies observed in other genes, the appearance of class I mutants is still a relatively rare event, occurring at a frequency of 10^{-3} to 10^{-5}, depending on the locus and strains of mice considered (20). Some of these gene conversion mutations appear to have occurred mitotically in the germ line (8), possibly at a stage in development that directly precedes the first meiotic cell divisions in the ovary. Little is known about mitotic gene conversion in higher eukaryotic cells in vivo. Some recent evidence has been reported that describes somatic gene conversion within B-cell lineages of rabbits and chickens (14, 17). However, there is no clear link between these somatic events that lead to immunoglobulin sequence diversity and the events that lead to the class I mutations. The low frequency of class I mutations and the fact that the mutants are detectable only by functional immunological assays make direct genetic analysis of the mechanism impractical at present. Furthermore, in vitro approaches have been futile, because no cell lines that undergo conversion of their MHC alleles have been identified.

One approach that can provide insight into the mechanism underlying the generation of this interesting class of variant is to dissect exceptions to the gene conversion model. Here, we describe two class I mutants, K^{bm3} and K^{bm23}, that represent apparent exceptions. Both mutants exhibit all of the hallmarks of the mitotic gene conversion phenomenon. Each is characterized by a cluster of five nucleotide substitutions that presumably arose in a single genetic event. However, there is no apparent donor gene in the parental genome that could serve as the template for the putative gene conversion event. Hence, analysis of the structure of the mutation clusters characteristic of these two mutants provides insight into the mutation mechanism. We also present a hypothetical model consistent with these findings that involves the transfer of sequences from donor to target genes through an RNA intermediate. In addition to accounting for all previously described characteristics of the spontaneous class I mutations, this model also accounts for the structures of the K^{bm3} and K^{bm23} mutation clusters.

MATERIALS AND METHODS

Mice. Pedigreed mice were purchased from the Jackson Laboratories, Bar Harbor, Maine. Mouse strains used in these studies were BALB/cByJ, C57BL/6ByJ, C57BL/10J, C57BL/10SnJ, C57BL/6J-H-2^{bm3}, and B10.D2-H-2^{bm23}/Eg.

Genomic DNA and Southern blots. Genomic DNA was isolated from spleen cells as described previously (23). A total of 40 μg of DNA was digested with endonuclease *Bam*HI or *Hin*dIII (Boehringer Mannheim, Indianapolis,

* Corresponding author.
† Present address: Gortner Laboratory, University of Minnesota, St. Paul, MN 55108.
‡ Present address: Amvax, Inc., Beltsville, MD 20705.

Ind.). Southern blots were prepared by using GeneScreen membranes (Du Pont NEN Research Products, Boston, Mass.) and were hybridized with ^{32}P-labeled oligonucleotide probes as described previously (9). A minor variation of this protocol was used for the blots shown in Fig. 3A and 4 by substituting SSPE (20× stock consists of 3 M NaCl, 0.2 M $NaH_2PO_4 \cdot H_2O$, and 0.02 M EDTA [pH 7.4]) for SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate). Stringency was adjusted by increasing the hybridization and final washing temperatures. High stringency was operationally defined as the conditions under which an oligonucleotide probe hybridized to the mutant gene fragment but failed to hybridize to the homologous parent sequence (internal control). The presence or absence of other hybridizing fragments could then be evaluated. The presented hybridization patterns were observed reproducibly in completely independent experiments. The oligonucleotide probes used in these studies were bm3-77 (GTCCTCAGGCTCACTCGGAAA), bm3-89 (GCCCGCGCTCTGGTT), bm23-63 (TTCTGTGTTTCC CGCTCCC), and bm23-75,77 (CCTCAGGCTCACATGGA AACT).

Class I genes. A panel of cosmid clones containing complete overlaps of the *K*, *D*, *Q*, and *T* regions of the *H-2^b* haplotype (34) was provided by Gerald Waneck (Massachusetts General Hospital, Boston, Mass.). This panel of cosmids has been analyzed previously with appropriate control probes that cross-react with all of the known class I genes, demonstrating the presence of known sequences in each of the lanes (2). Oligonucleotides bm3-77, bm3-89, and bm23-75,77 were used in the analysis.

Isolation and sequence analysis of *K^{bm3}* and *K^{bm23}* mutation clusters. The *K^{bm3}* gene was isolated from genomic spleen DNA by a size-selected lambda phage-cloning strategy. Genomic DNA was digested with the endonuclease *Xba*I, which is known to generate a 1.8-kb fragment of the *K^b* gene containing a major segment of the promoter region and exons 1 through 3. DNA fragments of the appropriate sizes were cloned by using the lambda-zap vector system (Stratagene, La Jolla, Calif.). Phage containing the desired fragments were recovered by using the oligonucleotide probe K^b-152 (GAGTCTCTCTGCTTCAC) that was shown previously (25) to hybridize uniquely to the *K^b* gene in the *H-2^b* cosmid library of class I genes. Double-stranded plasmids containing the appropriate fragments were excised from the phage and sequenced by a dideoxy-chain termination procedure. The *K^{bm23}* mutation cluster was isolated by two independent procedures. The first was by polymerase chain reaction amplification of the second exon of the *K* locus with primers that are homologous to intron 1 (GCGTCCTCGG GTCGACCACCGGACCC) and intron 2 (CTGTTCCCGAA GCTTTGGACTT) sequences. The resulting fragments were digested with the endonucleases *Sal*I and *Hin*dIII, whose recognition sites were included in the primers. The digested fragments were cloned into the vector pUC 18, sequenced, and authenticated as *K* locus products by sequence homology to *K^b* in highly polymorphic sequences in exon 2. The second approach used *K* locus-specific primers to isolate full-length cDNA from spleen RNA. This procedure has been described previously (13). Plasmids containing the appropriate cDNA as determined by using oligonucleotide K^b-152 were sequenced to establish the structure of the mutation cluster. The overlapping segments of the two sequences, including the entire mutation cluster, were identical.

RESULTS

The structure of *K^{bm3}*. The mutant *K^{bm3}* was reported by Igor Egorov in the former Soviet Union in 1964 (3). Because this variant was discovered segregating in a colony of inbred mice, the precise generation that gave rise to the mutation cannot be established with certainty. Nonetheless, the *K^{bm3}* mutation has many of the characteristics of a MHC gene conversion mutation. Protein sequence analysis revealed two amino acid substitutions, documenting changes at positions 77 and 89 (5, 26). Subsequently, RNA sequence analysis confirmed these findings and established that four nucleotide substitutions account for the two amino acid changes (7).

A departure from the pattern established by previously characterized class I mutants was evident when efforts failed to identify a donor gene for *K^{bm3}* among the other members of the class I multigene family of the *H-2^b* haplotype (7). The search was extended to the rest of the genome by using oligonucleotides to probe genomic Southern blots. No genes containing the precise sequences matching the stretch encoding amino acids 77 through 89 of the mutant were identified (7). However, a homologous stretch of sequence in the *D^b* gene was found to differ from the mutant sequence by a single nucleotide, suggesting to Geliebter and Nathenson the possibility that a patchy copy mechanism may have introduced an imperfect copy of the donor sequence and hence explaining this single exception to the gene conversion model.

The *K^{bm3}* mutation is of additional interest because it is adjacent to an exon-intron boundary. Since all of the characterized gene conversion mutations as well as most of the variability among alleles are found in exons, it had been suggested that the conversion mechanism may operate through an mRNA (spliced) intermediate (33). Consequently, we reexamined the structure of the *K^{bm3}* mutation. An *Xba*I fragment containing the first three exons of the *K^{bm3}* gene was isolated from genomic DNA by using a size-selected, lambda phage-cloning strategy (as described in Materials and Methods). The gene was isolated from the library with an oligonucleotide probe, K^b-152, which is known to be specific for the *K^b* gene (25). The sequence of the *K^{bm3}* mutation is shown in Fig. 1A. In addition to the four previously described nucleotide mutations, the cluster of substitutions included an additional nucleotide change distal to the codon for amino acid 89, 12 nucleotides inside intron 2. The remainder of the sequence was identical to that of the *K^b* parent gene. The *K^{bm3}* mutation, therefore, is composed of a cluster of five nucleotide changes, four in the coding region of exon 2 and one inside the adjacent second intron.

Three important insights are revealed by the *K^{bm3}* sequence. First, the mechanism that gave rise to the mutation affected intron sequences as well as exon sequences. This implies that fully spliced mRNA species are not necessary intermediates in this process. Second, as shown in Fig. 1A, the complete *K^{bm3}* mutation cluster differs from the homologous *D^b* sequence not by one but by three nucleotides, substantially weakening the argument that *D^b* might be an imperfect donor gene for the conversion event (7). Copy or editing errors would now have to be evoked at three different positions along the template. We have confirmed the findings of Geliebter and Nathenson that there was no exact donor sequence for the *K^{bm3}* mutation in the *H-2^b* haplotype by using mutation-specific oligonucleotides similar to the oligonucleotides that they used to probe genomic and cosmid

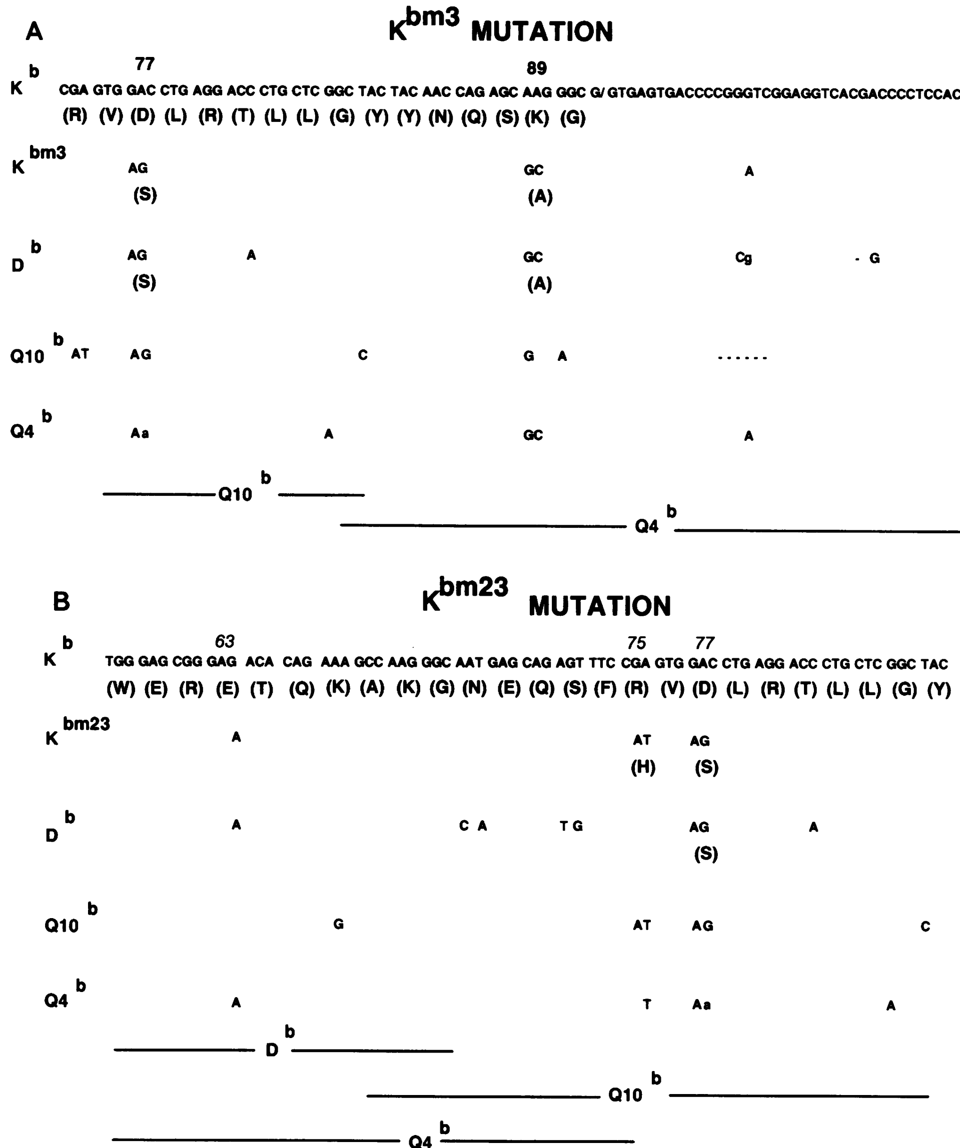


FIG. 1. The K^{bm3} and K^{bm23} mutations and their relationship to other *H-2*^{b} class I genes. (A) The sequence of K^{bm3} was determined from a genomic clone encompassing exons 1, 2, and 3. The five nucleotide changes relative to the K^b sequence (33) are indicated. The nonrandom nature of the changes is emphasized by comparison of the mutation cluster to the D^b (1), $Q10^b$ (19), and $Q4^b$ (30) sequences. Regions of identity between the mutation cluster and the putative *Q4* and *Q10* donor sequences are indicated. (B) The sequence of K^{bm23} was determined from two independent polymerase chain reaction-amplified fragments, one from cDNA and one from genomic DNA. The five nucleotide changes relative to the K^b sequence are indicated. Comparisons with putative donor sequences are as described for panel A above.

DNA sequences (data not shown). No gene containing all five mutations could be found in either the C57BL/6Sn or the C57BL/10J line, two strains that diverged from each other in the 1930s, well before the splitting off of the Egorov C57BL/6JY subline.

Third, although no donor gene was present, the nucleotide changes in the cluster of K^{bm3} mutations did not appear to be the result of random editing or copying errors by DNA repair or recombinatory machinery, because each substitution represented known structural diversity within the class I genes of the *H-2*^{b} haplotype (e.g., D^b, *Q4*, and *Q10* in Fig. 1A). Consequently, this observation implies that a mechanism responsible for the mutation cluster has introduced a complex array of nucleotide changes involving sequences present within the haplotype without use of a contiguous donor sequence from the genome to serve as a template. Unfortunately, the pedigree data available for the K^{bm3} mutation do not precisely identify the exact mating within the inbred lineage that gave rise to the mutant. Although there is no compelling reason to assume that this mutation is any different from the other mutations of K^b that are known to have occurred in a single generation, one might still argue that the accumulation of the bm3 mutation cluster represents a rare example of the multiple independent mutations.

Structure of K^{bm23}. The argument that the K^{bm3} mutation represents an important exception among the class I mutations was strengthened recently when a second class I mutation, K^{bm23}, with similar properties was identified by Egorov and Egorov (4). The pedigree records in this case are more precise. The bm23 mutation arose in a (C57BL/10SnEg

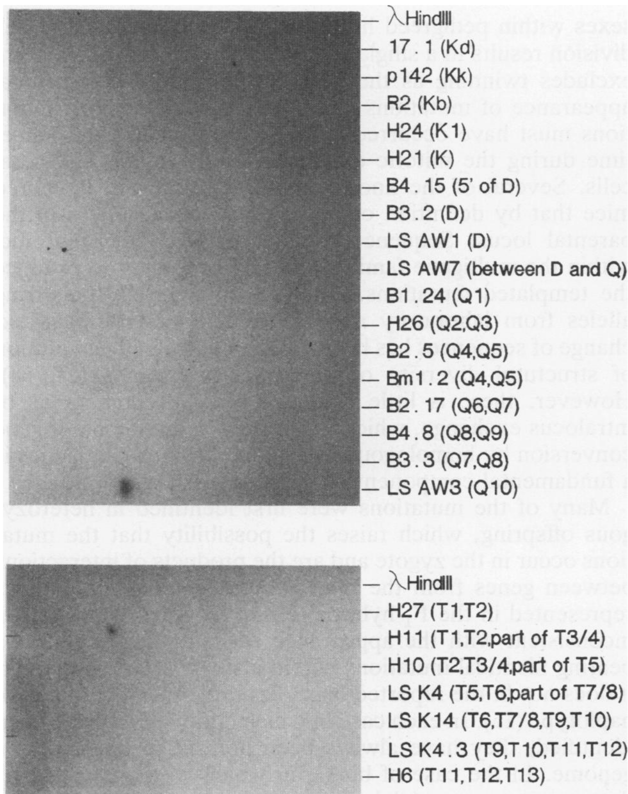


FIG. 2. The $Q10^b$ gene shares the 3′ portion of the mutation cluster with K^{bm23}. The panel of cosmids containing members of the $H\text{-}2^b$ class I genes was described previously (34). Recombinant clones containing the K^d (c17.1), K^k (p142), and K^b (pR2) sequences were also included in the analysis. Each of the DNA samples was digested with *Bam*HI endonuclease. The resulting Southern blot was probed with the oligonucleotide bm23-75,77 (representing sequences in the 3′ end of the mutation cluster) under stringent hybridization and washing conditions.

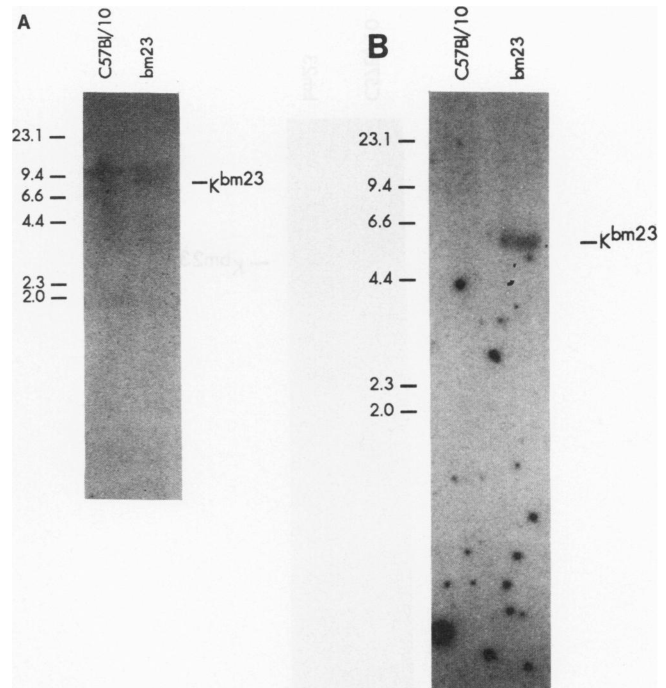


FIG. 3. Only a single sequence is identified in the parental genome with the 3′ end oligonucleotide bm23-75,77 to probe genomic Southern blots. Forty-microgram samples of spleen DNA isolated from C57BL/10SnJ or B10.D2-H-2^{bm23}/Eg mice and digested with either the endonuclease *Bam*HI (A) or *Hin*dIII (B) were used to prepare Southern blots. The blots were probed with ^{32}P-end-labeled oligonucleotide bm23-75,77 under stringent hybridization and washing conditions as described in Materials and Methods. The hybridization and washing temperatures for the blots shown were 49°C (A) and 60°C (B). Higher stringency was used in the second blot to visualize the closely spaced bands individually. Lambda *Hin*dIII molecular size markers were used as standards.

× B10.D2/nSnEg)F_1 hybrid and is one of the three known examples of the concordant appearance of identical mutants in a single mating, suggesting that this mutation is the result of a mitotic mechanism that gives rise to complex, nonrandom arrays of nucleotide substitutions that appear to be the products of gene conversion. The sequence of the K^{bm23} mutation was established by two independent polymerase chain reaction experiments. Two sets of oligonucleotide primers were used to isolate amplified DNA from genomic DNA and from cDNA. Each of the sequences agreed with the other, and the sequences are shown in Fig. 1B. The K^{bm23} mutation consisted of five nucleotide substitutions spanning the codons for amino acids 63 through 77. Two amino acid substitutions resulted from these changes, an R→H interchange at position 75 and a D→S interchange at position 77. A silent nucleotide substitution was present in the codon for amino acid 63. The remainder of the sequences, including the adjacent sequences in the second intron, were identical to that for K^b.

An oligonucleotide, bm23-75,77, that spans the four base changes resulting in the amino acid substitutions at positions 75 and 77 was used to screen a panel of cosmid clones representing the known $H\text{-}2^b$ class I genes in an attempt to identify a donor gene for K^{bm23}. Only a single potential donor gene was identified, $Q10^b$ (Fig. 2). The sequence of the $Q10^b$ gene has been reported elsewhere (18) but differs by two nucleotides from the K^{bm23} mutation cluster sequence (Fig. 1B). Therefore, none of the class I sequences in the cosmid bank represent a potential donor gene for the K^{bm23} mutation.

The oligonucleotide probe was also used to search for potential donor sequences in genomic DNA from C57BL/10SnJ mice. *Bam*HI and *Hin*dIII endonucleases were used to fragment DNA for Southern blot analyses. The probe specifically recognized the *K* locus in the mutant bm23 DNA but not in the parental genomic C57BL/10 DNA, illustrating the specificity of the hybridization and washing conditions (Fig. 3A and B). While no other hybridizing *Bam*HI fragment was observed in either DNA, a single *Hin*dIII fragment of approximately 10 kb was identified in both parental and mutant DNAs. Because we found that the cloned $Q10^b$ gene is on a 10-kb *Hin*dIII fragment, the observed pattern is consistent with the hybridization of this oligonucleotide to the $Q10^b$ gene. The $Q10^b$ sequence recognized by the probe is on a 0.8-kb *Bam*HI fragment (Fig. 2) (18), a size that is difficult to visualize by this oligonucleotide-based Southern blot technique on genomic DNA. We conclude from this analysis that there is no perfect donor sequence in the genome of the parental mouse and that the closest known match is $Q10^b$, a gene that differs by two nucleotides from the K^{bm23} mutation cluster.

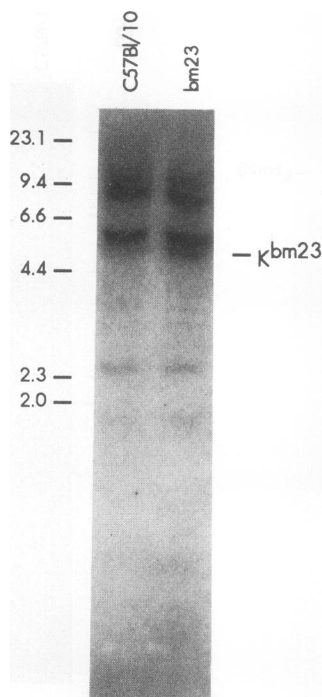


FIG. 4. Multiple parental sequences hybridize to an oligonucleotide probe matching the 5′ end sequences of the K^{bm23} mutation cluster. Forty-microgram samples of genomic spleen DNA as described in the legend to Fig. 3 were digested with the endonuclease *Bam*HI and used along with lambda *Hin*dIII markers to prepare a Southern blot. The blot was probed with the ^{32}P-end-labeled oligonucleotide bm23-63. Hybridization and washing were performed at 45°C.

Several genes that hybridized to a second probe, bm23-63, representing the 5′ region of the mutation cluster were detected in a *Bam*HI digest of both parental and mutant DNAs (Fig. 4). Two genes known to share this sequence are D^b and $Q4^b$ (Fig. 1B). Both of these genes match the 5′ end of the mutation cluster, encompassing the entire region mismatched by the $Q10^b$ sequence. Therefore, as in the K^{bm3} mutation, each of the mutations represents known diversity within the class I gene family of the $H\text{-}2^b$ haplotype, but no single member of the family could have served as the donor gene for the transfer of the genetic information to the parental K^b gene.

DISCUSSION

The details of the mechanism that results in the transfer of sequences between different members of the class I multigene family are not known. A straightforward scheme would involve a classic gene conversion model based on the resolution of heteroduplexes (12) formed during the pairing of sister loci as part of meiotic recombinatory events. Several facts about the class I mutants argue that this view is not correct.

The best estimate of the developmental stage at which the class I mutants occur is during the formation of the ovary, beginning at day 12, when the first signs of sexual differentiation can be identified, and ending around day 19, when most of the oocytes have entered meiosis. The most compelling argument supporting this estimate is the sporadic appearance, as in bm23, of identical mutants of different sexes within pedigreed litters (4, 21). Because meiotic cell division results in a single gamete in the female, this pattern excludes twinning as the explanation for the coincidental appearance of mutations among siblings. Therefore, mutations must have occurred before meiotic reduction, sometime during the mitotic expansion of the primordial germ cells. Several of the known mutations occurred in inbred mice that by definition contain two identical alleles at the parental locus. Sequence transfer between different loci within the multigene family would be necessary to produce the templated mutations. Analyses of naturally occurring alleles from laboratory mice revealed that interlocus exchange of sequences has been commonplace in the evolution of structural diversity of mouse class I genes (24, 28). However, there is little evidence for a predominance of intralocus exchange, which is the expected outcome if gene conversion by homologous pairing of chromosomal genes is a fundamental component of the process.

Many of the mutations were first identified in heterozygous offspring, which raises the possibility that the mutations occur in the zygote and are the products of interactions between genes from the two different parental haplotypes represented in the F_1 hybrids. However, this hypothesis is inconsistent with the appearance of multiple variant pups bearing identical mutations within a single litter. Moreover, it does not fit the reported bias of mutations in the maternal haplotype (15). In each case in which donor genes have been identified, they have always been present in the maternal genome. In the case of bm23, in which we have found no donor gene, a potential donor sequence was also not found in the paternal $H\text{-}2^d$ haplotype, by the same strategies reported here (data not shown). Because the bm3 mutation arose in an inbred mouse, by definition the donor gene is missing from both the maternal and the paternal genomes. The idea that the gene conversion events occur premiotically, early in the development of the maternal gametes, is most consistent with the observed profile of mutations.

Each of the nucleotide substitutions, as well as the combinations of nucleotide changes that constitute individual codons within the bm3 and bm23 mutations, represents naturally occurring diversity that exists within the class I genes of the maternal $H\text{-}2^b$ haplotype. It may be significant that both of the mutation clusters appear to be neatly divided into two regions, i.e., a 5′ segment that corresponds exactly to one member of the multigene family and a 3′ segment that corresponds exactly to one member of the multigene family and a 3′ segment that corresponds exactly to another member of the class I gene family (Fig. 1). Importantly, each of these gene family members, namely, D^b, $Q4^b$, and $Q10^b$, has been identified previously as a donor gene involved in the generation of a K^b mutation (7, 8, 19). We postulate that the bm3 and bm23 variants are templated mutants and that they differ from other defined mutants because two donor genes, instead of the more typical single donor gene, have participated in the transfer of sequence into the parental target gene. The unusual aspect of these two conversion mutations is that three genes appear to have interacted in a single genetic-recombination event.

It is unlikely that the structures of the bm3 and bm23 mutations are both products of two independent gene conversion mutations. The low observed frequency of single events (10^{-3} to 10^{-4}) argues against the possibility that 2 of less than 30 characterized mutants could be derived from two independent mutation events (expected frequency of 10^{-6} to 10^{-8}). Furthermore, the distribution of the nucleotide mutations is such that they are clustered in the 3′ 100

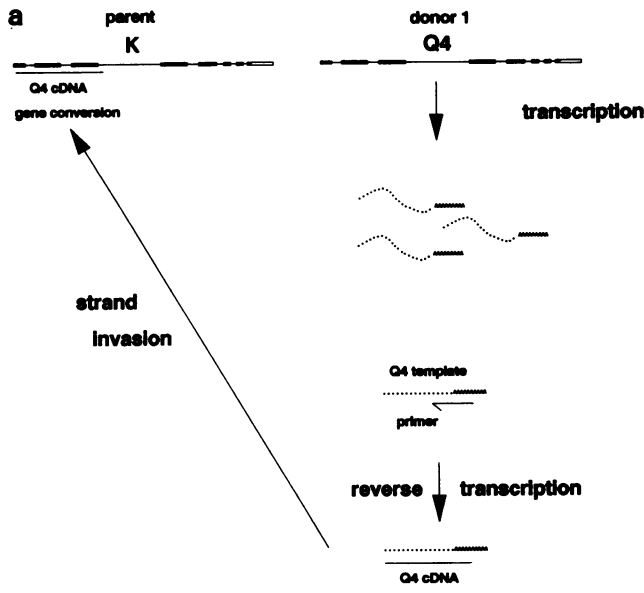
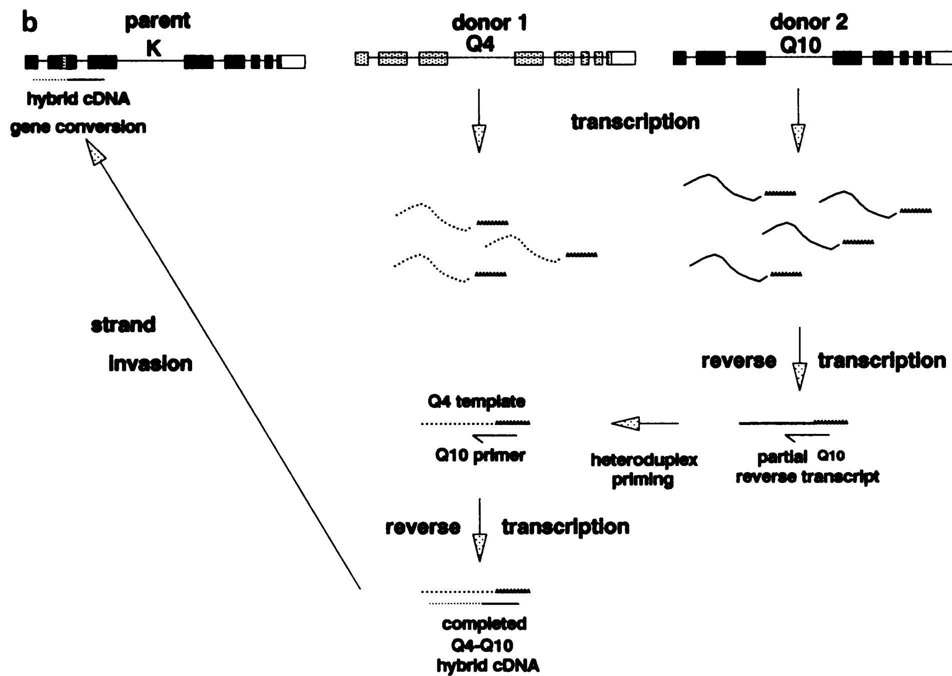


FIG. 5. Hypothetical mutation mechanism involving cDNA intermediates. Mutations arise in this scheme from sequences anywhere in the genome. (a) Donor sequences are transcribed, and the resulting RNA is used as a template in the synthesis of cDNA. The cDNA serves as the source of strand-invading DNA that ultimately results in heteroduplex formation and gene conversion as a consequence of heteroduplex resolution. (b) This process also provides a mechanism that can give rise to recombinant templates that would produce mutation clusters as seen in the K^{bm3} and K^{bm23} mutants. In this scheme, some of the forming cDNA molecules will not be full length. A small percentage of the cDNA fragments could then serve as primers in a second round of cDNA synthesis. If the cDNA fragments cross-prime RNA from another class I gene, hybrid cDNA molecules would be formed. Such molecules might compose a small portion of the pool of cDNA that invades the target gene, generates heteroduplexes, and ultimately resolves into class I gene conversion mutants. If the frequency of recombinant cDNA molecules generated approximated 5%, the observed frequency of donorless gene conversion mutants (1 in 20) would result.



nucleotides of exon 2 in both bm3 and bm23, providing further evidence against the possibility that the 5′ and 3′ sequences within the mutation clusters were transferred in completely independent mutation events. The low expected frequency of double mutants also argues compellingly against the possibility that the bm3 and bm23 variants arose by successive mutations in different generations. Even allowing the mutation clusters to accumulate changes over 30 generations (approximately 10 years) would not substantially change these conclusions on the basis of the large differences between the expected and observed frequencies of the exceptional variants.

The alternative view is that the mutations occurred coordinately as the result of the interactions among three genes tied together in a single genetic event. There are no currently understood genetic processes that involve the direct interactions of three genes in a recombination scheme. However, three genes could interact sequentially in a linked series of events to yield the unusual mutation clusters identified in the bm3 and bm23 mutants. Each recombination event would involve the pairing of two genes or their products. Since the bm3 and bm23 mutations are clustered in one of several regions of the *K* gene known to undergo conversion events, it is unlikely that two donor genes sequentially interact with the target sequence. However, if the slow step in the series of interactions were the gene conversion of the target locus, the opportunity would exist for recombination between donor sequences prior to the conversion event.

Donor gene recombination could occur either at the level of the chromosome or extrachromosomally. Chromosomal recombination between members of the class I multigene family is not generally observed, particularly not at a frequency approaching 7% (two recombinant donor sequences of 30 mutants). Furthermore, genes of the *Q* and *T* regions that have been identified previously as donors are among the most conserved class I genes in the mouse (6). Consequently, chromosomal recombination between donor genes does not appear to be a significant factor in the generation of class I mutations.

The hypothesis that extrachromosomal intermediates may be important components of the class I gene conversion process, however, is consistent with all the known facts about the class I mutations. The essential steps in gene conversion are strand invasion, heteroduplex formation, and heteroduplex resolution. Strand invasion and heteroduplex formation can be initiated by either chromosomal or extrachromosomal DNA (29). Strand invasion implies gene pairing between donor and target sequences. Although there is considerable variation in homology among members of the class I gene family, the donor genes have come from the group of genes most closely related to the target loci (22). The involvement of a diffusible extrachromosomal intermediate would free donor and target genes from spatial constraints, allowing genes that map hundreds of thousands of nucleotides apart to interact. This fits well with the observation that genes mapping to distant loci in the *D*, *Q*, and *T* regions of the mouse MHC appear to have served as donor genes for the conversion of the *K* locus (22).

From a mechanistic perspective, each of the essential components to generate extrachromosomal substrates from class I donor genes is known to exist within the mammalian germ line. The model we propose is that extrachromosomal intermediates are generated from the class I donor genes by transcription. Species of heterogeneous nuclear RNA or mRNA may serve as substrates for reverse transcriptase activity. The presence of process pseudogenes in mammalian genomes not only demonstrates the existence of reverse transcriptase activity in mammalian cells but also illustrates the fact that cDNA can be present in a form that can integrate into germ line DNA (32). cDNA could also function as a substrate for homologous recombination, a process that may lead to the gene conversion phenotype characteristic of the MHC mutants (Fig. 5). A precedent for an RNA intermediate in gene conversion between two related loci recently has been reported for yeast cells (1). Also important to the model is the fact that the *D*, *Q4*, and *Q10* genes proposed as donors for the bm3 and bm23 mutations are all known to be transcriptionally active, as are the *K*, *K1*, and *T5* loci, three other genes believed to have served as donor gene templates for other mutants (7).

A salient feature of this hypothesis is that reverse transcription of RNA transcribed from class I genes also provides a mechanism for the recombination of donor gene sequences. Successive rounds of reverse transcription could produce chimeric intermediates that result in recombinant mutation clusters (Fig. 5B). Chimeric cDNA molecules could form when partial cDNA fragments serve to cross-prime DNA synthesis with heterologous RNA as a template. The resulting molecules would contain 5′ sequences derived from one donor gene and 3′ sequences from another. Gene conversion events resulting from the resolution of heteroduplex formed between these chimeras and the target gene would lead to mutation clusters of the kind described here for the bm3 and bm23 mutations. However, since there is no specific information about the frequency of reverse transcription in the generation of donor fragments, this model does not address the frequency with which chimeric donor cDNA would be generated.

Although the hypothesis that cDNA can act as intermediates in class I gene conversion is consistent with the observed mutation clusters, there are several other features of the structure of the class I multigene family not addressed by the model. Among these is the absence of class I process pseudogenes in the mouse genome. The reasons why some multigene families have numerous process pseudogenes and others do not are not known and may be related to more complex issues beyond the generation of cDNA. Another possible consequence of the role of cDNA extrachromosomal intermediates in gene conversion is that introns could be deleted as a result of conversion events across introns of genomic DNA by cDNA strands derived from processed mRNA. Intron deletions of this kind in the class I gene family have not been observed. In the case of the bm3 mutation, the fact that the mutation cluster extends into the intron suggests that spliced mRNA is not necessarily the source of the proposed cDNA. If heterogeneous nuclear RNA were the source of cDNA, the presence of introns would be conserved and mutations would extend into the introns, as seen in the case of bm3.

The proposed mechanism of cDNA-mediated recombination between class I donor genes could have implications for other multigene families as well. Individually, the steps of the proposed mechanism are known to exist in some form in mammalian cells, and there is no reason to believe that any of the steps would be uniquely suited to class I sequences. Interchanges between discrete loci have been important in the evolution of immunoglobulin (35) and probably also T-cell receptor gene segments in which structural differences between the expressed members of a multigene family are known to have important functional consequences. This concept can be extended to gene families that encode differentially regulated isomers, as in the globin or tubulin families. Clear examples of gene conversion have been cited for each of these groups (27, 31), but the mechanism responsible for these recombination events remains undefined, as it is for the class I genes. Consequently, understanding the mechanism underlying the generation of the class I mutations could provide broader insight into the forces shaping eukaryotic genomes.

REFERENCES

1. **Derr, L. K., and J. N. Strathern.** 1993. A role for reverse transcripts in gene conversion. Nature (London) **361:**170–173.
2. **Duran, L. W., and L. R. Pease.** 1988. Tracing the evolution of H-2 D region genes using sequences associated with a repetitive element. J. Immunol. **141:**295–301.
3. **Egorov, I. K.** 1967. A mutation of the histocompatibility-2 locus in the mouse. Genetika **3:**136–144.
4. **Egorov, O. S., and I. K. Egorov.** 1984. H-2^{bm23}, a new K^b mutant similar to but not identical with H-2^{bm3}. Immunogenetics **20:**83–87.
5. **Ewenstein, B. M., H. Uehara, T. Nisizawa, R. W. Melvold, H. I. Kohn, and S. G. Nathenson.** 1980. Biochemical studies on the H-2K antigens of the MHC mutants bm3 and bm11. Immunogenetics **11:**383–395.
6. **Flaherty, L., E. Elliott, J. A. Tine, A. C. Walsh, and J. B. Waters.** 1990. Immunogenetics of the Q and TL regions of the mouse. Crit. Rev. Immunol. **10:**131–175.
7. **Geliebter, J., and S. G. Nathenson.** 1988. Microrecombinations generate sequence diversity in the murine major histocompatibility complex: analysis of the K^{bm3}, K^{bm4}, K^{bm10}, and K^{bm11} mutants. Mol. Cell. Biol. **8:**4342–4352.

8. **Geliebter, J., R. A. Zeff, R. W. Melvold, and S. G. Nathenson.** 1986. Mitotic recombination in germ cells generated two MHC mutants determined to be identical by RNA sequence analysis: K^{bm9} and K^{bm6}. Proc. Natl. Acad. Sci. USA **83:**3371–3375.
9. **Geliebter, J., R. A. Zeff, D. H. Schulze, L. R. Pease, E. H. Weiss, A. L. Mellor, R. A. Flavell, and S. G. Nathenson.** 1986. Interaction between *K^b* and *Q4* gene sequences generates the K^{bm6} mutation. Mol. Cell. Biol. **6:**645–652.
10. **Hemmi, S., J. Geliebter, R. A. Zeff, R. W. Melvold, and S. G. Nathenson.** 1988. Three spontaneous H-$2D^b$ mutants are generated by genetic micro-recombination (gene conversion) events: impact on the H-2 restricted immune responsiveness. J. Exp. Med. **168:**2319–2335.
11. **Hildebrand, W. H., R. M. Horton, L. R. Pease, and J. M. Martinko.** 1992. Nucleotide sequence analysis of H-$2D^f$ and the spontaneous in vivo H-$2D^{fm2}$ mutation. Mol. Immunol. **29:**61–69.
12. **Holliday, R.** 1974. Molecular aspects of genetic exchange and gene conversion. Genetics **78:**273–287.
13. **Horton, R. M., W. H. Hildebrand, J. M. Martinko, and L. R. Pease.** 1990. Structural analysis of H-$2K^f$ and H-$2K^{fm1}$ by using H-2K locus-specific sequences. J. Immunol. **145:**1782–1787.
14. **Knight, K. L.** 1992. Restricted V_H gene usage and generation of antibody diversity in rabbit. Annu. Rev. Immunol. **10:**593–616.
15. **Loh, D. Y., and D. Baltimore.** 1984. Sexual preference of apparent gene conversion events in MHC genes of mice. Nature (London) **309:**639–640.
16. **Martinko, J. M., J. C. Solheim, and J. Geliebter.** 1988. The H-$2K^{km1}$ mutation: a single nucleotide substitution is responsible for multiple functional differences in a class I MHC molecule. Mol. Immunol. **25:**267–274.
17. **McCormack, W. T., L. W. Tjoelker, and C. B. Thompson.** 1991. Avian B-cell development: generation of an immunoglobulin repertoire by gene conversion. Annu. Rev. Immunol. **9:**219–241.
18. **Mellor, A. L., E. H. Weiss, M. Kress, G. Jay, and R. A. Flavell.** 1984. A nonpolymorphic class I gene in the murine major histocompatibility complex. Cell **36:**139–144.
19. **Mellor, A. L., E. H. Weiss, K. Ramachandran, and R. A. Flavell.** 1983. A potential donor gene for the bm1 gene conversion event in the C57BL mouse. Nature (London) **306:**792–795.
20. **Melvold, R. W., and H. I. Kohn.** 1990. Spontaneous frequency of H-2 mutations, p. 3–13. *In* I. K. Egorov and C. S. David (ed.), Transgenic mice and mutants in MHC research. Springer-Verlag KG, Berlin.
21. **Melvold, R. W., H. I. Kohn, and G. R. Dunn.** 1982. History and genealogy of the H-$2K^b$ mutants from the C57BL/6Kh colony. Immunogenetics **15:**177–185.
22. **Nathenson, S. G., J. Geliebter, G. M. Pfaffenbach, and R. A. Zeff.** 1986. Murine major histocompatibility complex class-I mutants: molecular analysis and structure-function implications. Annu. Rev. Immunol. **4:**471–502.
23. **Pease, L. R., S. G. Nathenson, and L. A. Leinwand.** 1982. Mapping class I gene sequences in the major histocompatibility complex. Nature (London) **298:**382–385.
24. **Pease, L. R., J. K. Pullen, Z. Cai, and R. M. Horton.** 1991. Contributions of interlocus exchange to the structural diversity of the H-2K, D, and L alleles, p. 163–170. *In* J. Klein and D. Klein (ed.), Molecular evolution of the major histocompatibility complex, vol. H 59. Springer-Verlag KG, Berlin.
25. **Pease, L. R., D. E. Regnerus, B. K. Huntley, and L. W. Duran.** 1987. The relative contributions of gene conversion-like substitutions and unique mutations toward polymorphism among MHC class I genes, p. 77–88. *In* C. S. David (ed.), H-2 antigens. Plenum Publishing Corporation.
26. **Pease, L. R., D. H. Schulze, G. M. Pfaffenbach, and S. G. Nathenson.** 1983. Spontaneous H-2 mutants provide evidence that a copy mechanism analogous to gene conversion generates polymorphism in the major histocompatibility complex. Proc. Natl. Acad. Sci. USA **80:**242–246.
27. **Powers, P. A., and O. Smithies.** 1986. Short gene conversions in the human fetal globin gene region: a by-product of chromosome pairing during meiosis? Genetics **112:**343–358.
28. **Pullen, J. K., R. M. Horton, Z. Cai, and L. R. Pease.** 1992. Structural diversity of the classical H-2 genes: K, D, and L. J. Immunol. **148:**953–967.
29. **Radding, C. M.** 1982. Homologous pairing and strand exchange in genetic recombination. Annu. Rev. Genet. **16:**405–437.
30. **Robinson, P. J., D. Bevec, A. L. Mellor, and E. H. Weiss.** 1988. Sequence of the mouse Q4 class I gene and characterization of the gene product. Immunogenetics **27:**79–86.
31. **Sullivan, K. F., J. T. Y. Lau, and D. W. Cleveland.** 1985. Apparent gene conversion between β-tubulin genes yields multiple regulatory pathways for a single β-tubulin polypeptide isotype. Mol. Cell. Biol. **5:**2454–2465.
32. **Weiner, A. M., P. L. Deininger, and A. Efstratiadis.** 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu. Rev. Biochem. **55:**631–661.
33. **Weiss, E., L. Golden, R. Zakut, A. Mellor, K. Fahrner, S. Kvist, and R. A. Flavell.** 1983. The DNA sequence of the H-$2K^b$ gene: evidence for gene conversion as a mechanism for the generation of polymorphism in histocompatibility antigens. EMBO J. **2:**453–462.
34. **Weiss, E. H., L. Golden, K. Fahrner, A. L. Mellor, J. J. Devlin, H. Bullman, H. Tiddens, H. Bud, and R. A. Flavell.** 1984. Organization and evolution of the class I gene family in the major histocompatibility complex of the C57BL/10 mouse. Nature (London) **310:**650–655.
35. **Wysocki, L. J., and M. L. Gefter.** 1989. Gene conversion and the generation of antibody diversity. Annu. Rev. Biochem. **58:**509–531.