Molecular and Cellular Biology, May 1987, p. 1646–1655
0270-7306/87/051646-10$02.00/0
Copyright © 1987, American Society for Microbiology
Vol. 7, No. 5

Analysis of the Integrant in MyK-103 Transgenic Mice in which Males Fail To Transmit the Integrant

THOMAS M. WILKIE and RICHARD D. PALMITER*

Howard Hughes Medical Institute Laboratory, Department of Biochemistry, University of Washington, Seattle, Washington 98195

Received 24 September 1986/Accepted 22 January 1987

Males in the transgenic mouse pedigree MyK-103, although fertile, do not transmit the integrant to offspring. The integrant is on chromosome 6 near the T-cell receptor β-chain locus. It contains four fragments of the plasmid pMK (a metallothionein-thymidine kinase fusion gene) and a 532-base-pair fragment of displaced mouse DNA originating from a previously uncharacterized repetitive DNA family. The integration complex is flanked on either side by a 5-kilobase duplication of mouse DNA normally found in a single copy at this locus. Sequence analysis of the six novel junctions and their donor sequences shows that plasmid-plasmid junctions occurred at patches of limited homology, whereas chromosome-plasmid junctions were nonhomologous.

Transgenic mice have been invaluable in investigation of the expression of foreign genes during development. A number of investigators have also found serendipitous mutations apparently caused by integration of the foreign DNA. In total, 11 recessive mutations in 152 lines of inbred transgenic mice have been reported (reviewed in reference 34). The foreign DNA provides a convenient method for cloning the flanking mouse DNA and may facilitate access to otherwise obscure genetic loci, as illustrated by the analysis of mutations due to retroviral integration (22, 25, 32, 44). Mutations caused by proviral integration are analogous to mutations that are induced by transposon integration in bacteria (12), yeasts (39), maize (14), and fruit flies (51). Many examples of transposon and proviral integration have accumulated which provide a conceptual framework for understanding the consequences of insertional mutagenesis.

The molecular mechanisms which cause the transgenic mouse mutations that cosegregate with integrated foreign DNA have not yet been described. Several transgenic mouse lines carry recessive mutations that cause embryo death, probably due to integration of foreign DNA within or near a gene that is required during embryogenesis (29, 45, 56). In another transgenic line, integration of mouse mammary tumor virus-*myc* genes generated a recessive mutation, causing limb deformity, which is allelic to naturally occurring mutations on mouse chromosome 2 (60). Recently, a second insertional mutation causing limb deformity has been reported (33). In the transgenic pedigree MyK-103, the integrant, which contains four fragments of pMK (a mouse metallothionein-I herpes simplex virus [HSV]-thymidine kinase [TK] fusion gene in pBR322), is transmitted only through the female germ line. MyK-103 males are fertile and sire normal-sized litters but never transmit pMK to offspring (36). We proposed that the integrant disrupts the normal expression of a gene which is required for development of fertile spermatozoa (36).

The MyK-103 pedigree was established following microinjection of several hundred copies of a linearized plasmid into a pronucleus of an F_2 hybrid egg. The founder female was mosaic, and roughly 15% of her cells were transgenic in somatic tissues and in the germ line, which implies that integration occurred after the first zygotic division. The mechanism of DNA integration in cleavage-stage embryos is unknown but appears to be similar to integration in tissue culture cells. The integrants usually contain tandem arrays of direct repeats, and linear molecules integrate more efficiently than circles (8). Only a small fraction of the injected molecules integrate, usually at a single chromosomal site (58). Integration of foreign DNA is presumably dependent on chromosomal (double-strand) breaks. These observations suggest that the number of breaks is limiting in early embryos (8), as has been suggested for integration of foreign DNA in tissue culture cells (18). Perhaps as a consequence of the dependence on double-strand breaks, many of the integration sites which have been characterized in transgenic mice are associated with chromosomal rearrangements, including chromosomal translocations (20, 33), deletion (60), and cointegration with displaced mouse DNA sequences (15). We show here that the MyK-103 integrant also exhibits this last characteristic. In addition, it is flanked on either side by a 5-kilobase (kb) duplication of mouse DNA normally found in a single copy at this locus. Although thousands of transgenic mice have been produced during the past few years, only Southern blot analyses of the integrant structure have been published (refer to the references above). To obtain more detailed information, we have cloned and characterized the MyK-103 integrant and flanking mouse DNA. DNA sequence analysis of this integrant reveals evidence of homologous recombination between fragments of foreign DNA and nonhomologous recombination at the site of chromosomal integration.

MATERIALS AND METHODS

Mouse strains and cell lines. The MyK-103 pedigree was established as described previously (36). Genomic DNA was isolated as described previously (35). DNA from somatic cell hybrids was provided by P. Ashley and D. Cox. DNA from AK×L recombinant inbred lines was provided by the Jackson Laboratory.

Molecular cloning and DNA sequencing. Spleen DNA from a hemizygous MyK-103 female was partially digested with *Mbo*I. Fragments of 15 to 20 kb were collected in fractions from a 10 to 40% sucrose gradient containing 1 M NaCl. EMBL3 λ bacteriophage and packaging extracts were obtained from Vector Systems. Recombinant phage (1.2×10^6 PFU) were propagated in the host bacteria Q359 and LE392

* Corresponding author.

and screened with nick-translated pTK (HSV-TK cloned in pBR322). Subclones were constructed in the M13mp18 and M13mp19 cloning vectors. The normal locus was obtained from a Charon 4A Library of BALB/c DNA (17) by using the *Bam* 1.5 probe (see Fig. 2). DNA sequences were determined by the dideoxy chain termination procedure of Sanger et al. (43). All sequences were determined at least twice in both orientations, except J6, which was obtained from three independent M13 clones all read in the same orientation. Sequencing primers were obtained from the Howard Hughes Medical Institute chemical synthesis facility at the University of Washington. DNA sequences were analyzed on an IBM PC by using the computer programs numseq, DM (30), and D3HOM homology search (19).

DNA quantitation. The copy number of the 532-base-pair (bp) fragment between J6 and J7 was estimated from DNA dot hybridizations. Genomic mouse DNA was sonicated and dispensed in twofold dilution steps into carrier herring sperm DNA (a total of 2 μg of DNA per spot). DNA was baked onto nitrocellulose filters and hybridized with one of three nick-translated probes; two of the probes contained either half of the 532-bp fragment, and the third probe was *Bgl* 1.6, which hybridizes to mouse DNA present in a single copy per haploid genome. The intensity of hybridization to different concentrations of mouse DNA was compared for the three probes. All probes had nearly identical specific activities, since they hybridized equally in a dilution series of a plasmid from which they were isolated.

RESULTS

General structure of the transgene integrant. The MyK-103 line resulted from the microinjection of plasmid pMK, a metallothionein-HSV-TK fusion gene cloned into pBR322, which was linearized at the unique *Bam*HI site (Fig. 1A). To analyze the fine structure of the integrant, we made a genomic library from MyK-103 DNA that had been partially digested with *Mbo*I. Recombinant λ phage were screened with a nick-translated pBR322 plasmid which contained the HSV-TK gene. We obtained three clones, λ2, λ21, and λ41, which are diagrammed in Fig. 1B. To aid our characterization of the integrant, we cloned the normal locus from a λ Charon 4A genomic library of BALB/c mouse DNA by using hybridization probes of unique mouse DNA that flanked the MyK-103 integrant. We obtained one clone, λ35, which spanned the borders of the integration complex (Fig. 1B). The locations of the restriction endonuclease sites and hybridization probes that were pertinent for deducing the structure of the integrant and the normal locus are shown in Fig. 2. The integration complex is composed of the four fragments of pMK shown in Fig. 1A (labeled w, x, y, and z), which total 18.4 kb in length, and a 532-bp fragment of mouse DNA from a new family of repetitve mouse DNA. In addition, the 18.9-kb insert is flanked on both sides by a 5-kb sequence of mouse DNA which is normally present as a single copy at this locus. Thus, in effect, the MyK-103 chromosome carries an additional 24 kb of DNA (18.4 kb of pMK sequences, 0.5 kb of unknown mouse DNA, and the 5-kb duplication) that is not present on the normal chromosome. The integrant contains six novel junctions, which are indicated as J2 through J7 in Fig. 1B. As a result of the duplication, the DNA sequence S1 and S8 are donor sequences for junctions J7 and J2, respectively. Detailed restriction maps of λ2 and λ35 are identical to the left of the site designated J2, and maps of λ21 and λ35 are identical to the right of J7, indicating that the duplication was exact. Furthermore, the restriction map of λ2, λ21, and λ41 conforms to the genomic map of the integrant which was deduced by Southern analysis (see Fig. 1 in reference 36).

Southern blot analysis verified that a duplication flanks the MyK-103 integrant (Fig. 3). The results of this experiment are interpreted diagrammatically in Fig. 2. Two hybridization probes were used for this analysis, *Bam* 1.5 and *Bgl* 1.6 (both are unique mouse sequences), whose relative positions are shown beneath the normal chromosome in Fig. 2. The novel restriction fragments from the MyK-103 chromosome, which also hybridize to these probes, are shown in the accompanying diagram in Figure 2. The Southern blot in Fig. 3 compares somatic DNA from a hemizygous MyK-103 male with somatic DNA from a non-transgenic C57BL6 × SJL mouse. In the left panel, the *Bgl* 1.6 hybridization probe identifies novel restriction fragments from the MyK-103 chromosome which span J2 (*Pvu*II yields a 4.5-kb band, *Bam*HI yields a 6.4-kb band, and *Eco*RI yields a 9.0-kb band) in addition to the bands derived from the normal chromosome. Each of these restriction fragments has been shown previously to hybridize with pBR322 (36). In the right panel, the *Bam* 1.5 hybridization probe identifies another set of novel restriction fragments from the MyK-103 chromosome which span J7 (*Pvu*II yields an 8.5-kb band, *Eco*RI yields a 3.8-kb band, and *Bgl*II yields a 3.0-kb band), and these fragments also hybridize with probes isolated from pMK. Note that both the *Bgl* 1.6 and *Bam* 1.5 probes hybridized to the same 4.5-kb *Pvu*II fragment that spans J2 as well as the same 4.2-kb *Pvu*II fragment that spans S8. In addition, the DNA restriction fragments which are common to both MyK-103 and nontransgenic mice hybridize with greater intensity than the MyK-103 junction fragments, as would be expected owing to the 5-kb duplication which flanks the MyK-103 integrant.

DNA sequence analysis of the junctions. The precise location of the junctions indicated in Fig. 1 was determined by DNA sequencing from single-strand M13 templates that were generated by subcloning appropriate fragments from the various lambda clones. Only J3 was not cloned. However, Southern analysis clearly indicated that this junction was formed by head-to-head ligation at the *Bam*HI site that was used to linearize pMK (36). This regenerated the *Bam*HI site and created a perfect palindrome 3,450 bases in length. Such palindromes have been shown to be unstable in bacterial hosts commonly used to propagate λ phage (26). It is notable that most of the integrant was cloned on λ41 without rearrangements.

J4 appears to have resulted from recombination between sequences in HSV-TK and pBR322 which are identical at 12 of 13 positions (Fig. 4). The breakpoint in at least one strand of the double helix presumably occurred within the trinucleotide sequence CTT, since position 4 is an adenine in J4 and in pBR322 but is a guanine in HSV-TK. The cytidine insertion several nucleotides to the right of J4 has been verified by sequencing both strands of J4, the comparable regions of J6 (Fig. 5), and a subclone of the pMK plasmid that was used for microinjection. The insertion may have occurred during recombination in the early mouse embryo. It was the only discrepancy found in over 3,800 nucleotides which were sequenced around the junction and donor sequences indicated in Fig. 1. However, we cannot rule out the possibility that this insertion occurred during propagation of the lambda phage.

J5 also occurs within a patch of sequence homology between HSV-TK and pBR322 (Fig. 4). Of 11 nucleotides, 8

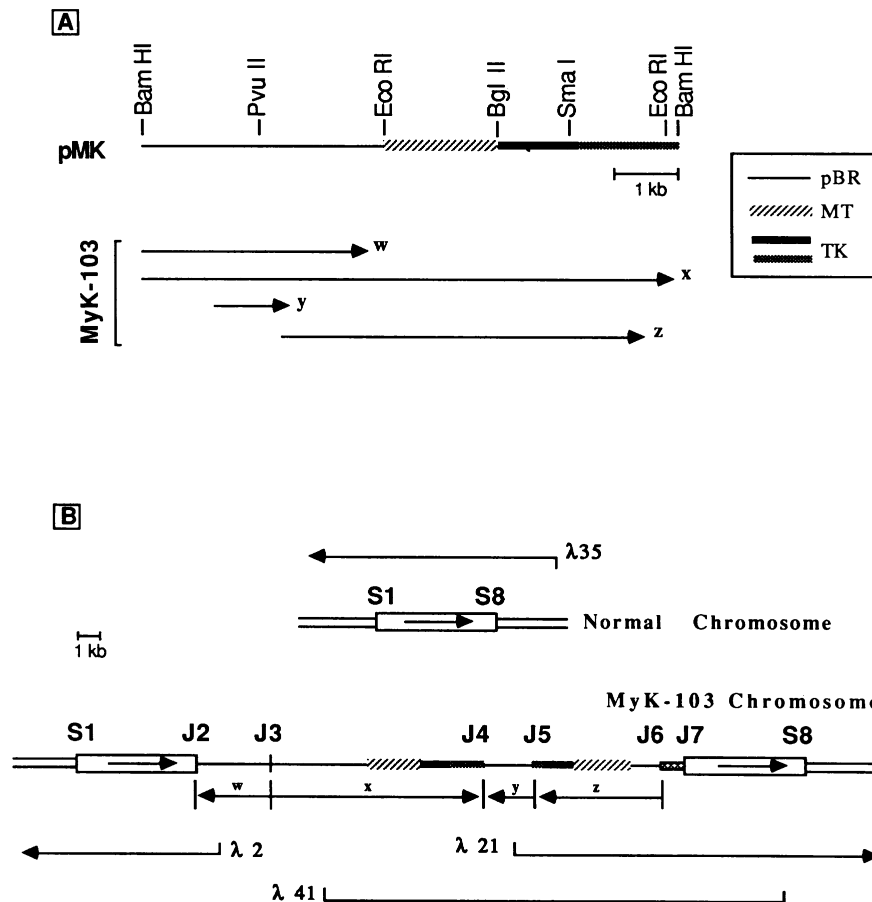

FIG. 1. Structure of the MyK-103 integrant. (A) MyK-103 transgenic mice contain four fragments of pMK, designated w, x, y, and z. *Bam*HI-linearized pMK is 8.5 kb and contains pBR322, the metallothionein-I promoter, and the HSV-TK gene, as indicated: ▬, the HSV-TK coding sequence; ▬, the downstream sequence. (B) The MyK-103 chromosome contains 24 kb of DNA not found in the normal chromosome and six novel junctions, J2 to J7. S1 is a donor sequence of J7, and S8 is a donor sequence of J2. The origin of pMK fragments w, x, y, and z is shown in panel A. ▨, the 532-bp fragment of repetitive mouse DNA that is bounded by J6 and J7. λ2 (23 kb), λ21 (24 kb), and λ41 (17.5 kb) contain DNA from the MyK-103 integrant. The DNA around J3, between λ2 and λ41, was not cloned. λ35 contains DNA normally found at this locus.

are identical around the breakpoint at the cytidine dinucleotide. This homologous region is separated by six nucleotides from another block of nine nucleotides, seven of which are identical between HSV-TK and pBR322. In total, this block of sequence around J5 is homologous between HSV-TK and pBR322 at 16 of 26 positions.

J6 and J7 define the boundaries of a 532-bp fragment of mouse DNA of unknown origin. The sequence of the junctions and the 532-bp fragment is shown in Fig. 5. A nick-translated probe of this sequence hybridizes to repetitive DNA that is present in about 1,200 copies per haploid genome, mostly within 1.6- and 1.9-kb *Bgl*II fragments which are tandemly arrayed. The DNA sequence of this 532-bp fragment is not homologous to any of the sequences catalogued in the Genebank file of repetitive mouse DNA, nor does this hybridization probe identify a specific transcript by Northern analysis of testis RNA; rather, a smear of randomly sized RNAs is found. Because we have not identified a unique hybridization probe with which to clone the sequences that normally flank the 532-bp sequence, we can only compare J6 with pBR322 and J7 with donor sequence S8.

In contrast to J4 and J5, the junction flanked by mouse DNA at J2 does not occur within a patch of homology between the progenitor sequences. In Fig. 6, J2 is aligned for comparison with pBR322 and with S8, which is nornally found at this locus. Note that 11 novel nucleotides at J2 separate the donor sequences. This exact sequence is not found in pMK nor in any of the λ clones other than at J2. Therefore, it may be the product of de novo synthesis during integration.

Chromosomal location of transgene integrant. We mapped the integration site in the MyK-103 pedigree by using the *Bam* 1.5 and the *Bgl* 1.6 hybridization probes shown in Fig. 2. Liver DNA from several inbred strains (A/J, AKR, C3H, C57BL/6, C57L, DBA, and SJL) was screened for restriction fragment length polymorphisms (RFLPs) by Southern blot after being digested with *Bam*HI, *Ban*II, *Bgl*II, *Stu*I, *Taq*I, or *Xba*I. The only RFLP that we identified was due to a 1.3-kb deletion in AKR mice. The deletion lies entirely within the *Bgl* 1.6 probe, as demonstrated by Southern blots of AKR DNA that were hybridized with the proximal *Bam* 1.5 or the distal *Sal*Hd4.0 probes (Fig. 2). Evidence of this deletion is also shown in Fig. 7, which is a Southern blot that includes liver DNA from C57L and AKR mice. DNA was digested with *Bam*HI and hybridized to an equal mixture of the *Bam* 1.5 and the *Bgl* 1.6 nick-translated probes. The *Bam* 1.5 probe hybridized to a 1.5-kb *Bam* fragment from both

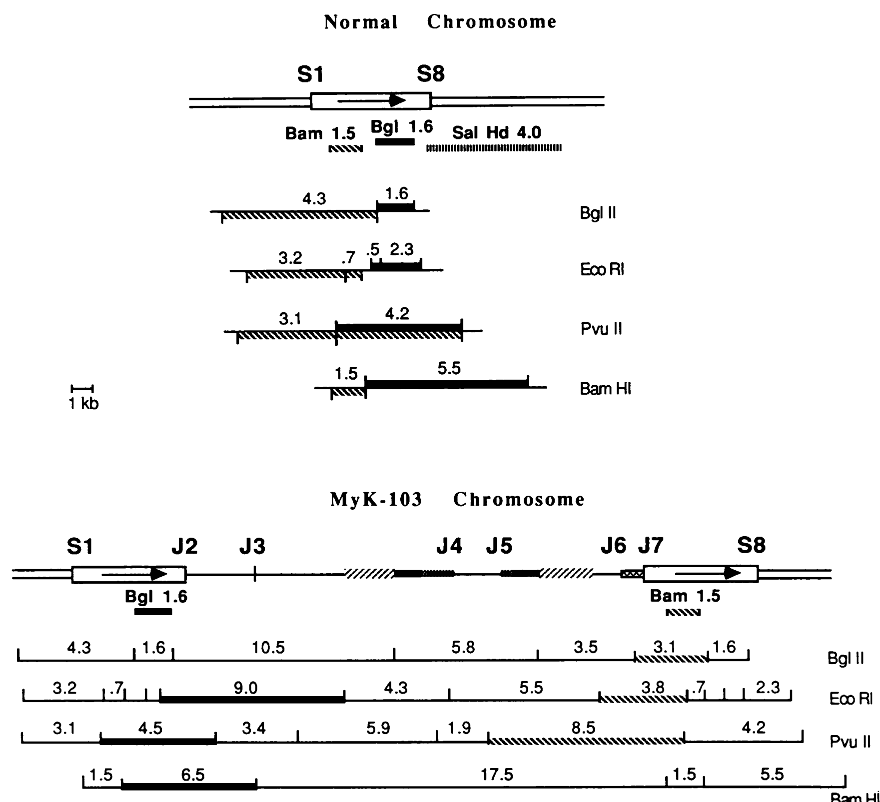

FIG. 2. Restriction maps of the MyK-103 and normal chromosomes. In the upper diagram, restriction fragments from a normal chromosome which hybridize to the *Bgl* 1.6 and the *Bam* 1.5 probes are identified to facilitate the interpretation of Fig. 3. The *Sal*Hd4.0 probe is also referred to in the text. In the lower diagram, the *Bgl* 1.6 and *Bam* 1.5 probes hybridize to novel junction fragments from the MyK-103 chromosome, as indicated. The *Bgl* 1.6 probe (▬) hybridizes to restriction fragments which cover S8, J2, and J3; the *Bam* 1.5 probe (▧) hybridizes to restriction fragments which cover S1, J5, J6, and J7. The *Bgl* 1.6 and *Bam* 1.5 probes both hybridize to the same 4.2-kb *Pvu*II fragment which spans S8, and they hybridize to the same 4.5-kb *Pvu*II fragment that spans J2.

mice. The *Bgl* 1.6 probe hybridized to the adjacent 5.2-kb *Bam* fragment from C57L mice but only faintly to the adjacent 3.9-kb *Bam* fragment in AKR mice that encompasses the deletion.

We exploited this RFLP in AKR mice to screen DNA from a set of 18 AKR × C57L (AK×L) recombinant inbred lines which had been used previously to map other loci. Figure 7 shows the distribution of the parental RFLPs among the recombinant inbred strains, identified as AK×L-5 through AK×L-38. This distribution matches that of the T-cell receptor β-chain locus, with the single exception of line AK×L-19 (40). We confirmed this observation by rehybridizing the blot shown in Fig. 7 with the Vβ8 probe (47) isolated from the T-cell receptor β-chain locus (4). The MyK-103 integration site maps within 11.6 centimorgans of the T-cell receptor β-chain locus with 95% confidence (46) and is positioned on the proximal half of mouse chromosome 6 (11). We also used the *Bam* 1.5 and the *Bgl* 1.6 probes to screen DNA from a panel of Chinese hamster-mouse somatic cell hybrid clones which segregated mouse chromosomes (16). Although this panel of hybrids could not unequivocally localize the integrant of one mouse chromosome, it was consistent with localization on chromosome 6 (P. Ashley and D. Cox, personal communication).

DISCUSSION

The most intriguing aspect of the MyK-103 pedigree is that only females transmit the integrant to offspring (16). Heterozygous males are fertile and sire litters of normal size, yet they never transmit the foreign DNA. Our efforts to understand the mechanism of this pattern of germ line transmission in MyK-103 males compelled us to clone the integrant and flanking mouse DNA. Unique flanking sequences were then used to isolate the normal DNA at the site of integration. The integrant has an unexpectedly complex structure, which is the subject of this report.

Analysis of the integrant. Tandem arrays of foreign DNA are frequently observed in tissue culture cells, whether DNA has been introduced by coprecipitation with $CaPO_4$ (38) or by microinjection (18). In transgenic mice, the foreign DNA also integrates as tandem head-to-tail arrays, as deduced by Southern blotting experiments after digestion with enzymes that cut only once within the injected DNA. Minor hybridizing bands are often observed as well and are usually assumed to be junction fragments; however, they might represent unusual arrangements of DNA in addition to the long tandem arrays (7). In contrast, Southern blot analysis of the MyK-103 integrant revealed that it was unusual (36), and this was confirmed by cloning out the integrant and studying its structure in detail. It is composed of four fragments derived from the injected pMK plasmid integrated in an unusual inverted configuration along with a 532-bp fragment of mouse DNA of unknown origin. The integrant is flanked on both sides by a 5-kb duplication of mouse DNA which is normally present as a single copy at this locus (Fig. 1B). In total, the MyK-103 integrant is distinguished by six novel

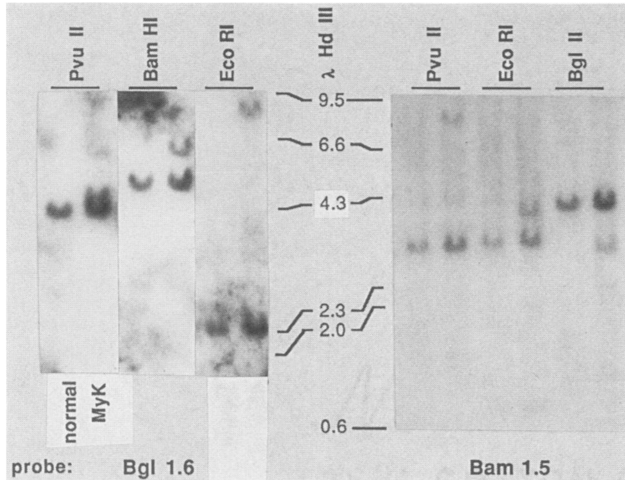

FIG. 3. Analysis of the 5-kb duplication flanking the MyK-103 integrant. Spleen DNA (2.5 μg) from a nontransgenic C57BL6 × SJL mouse and a hemizygous MyK-103 mouse was digested with the indicated enzymes, electrophoresed on 0.8% agarose gels side by side, transferred to nitrocellulose, and hybridized with either the *Bgl* 1.6 probe or the *Bam* 1.5 probe shown in Fig. 2. The sizes, in kilobases, of λ *Hin*dIII markers are shown between the panels.

junctions. The DNA sequences of these juctions were compared when possible to their progenitor sequences. As a result of the duplication, the DNA sequences at positions S1 and S8 are donor sequences for J7 and J2, respectively. Only J3 was not cloned, but Southern blot analysis clearly indicated that it was formed by ligation between two pBR322 *Bam*HI sites; this created a perfect palindrome 3,450 bp in length that has been stable in somatic tissues and the female germ line through 13 generations.

J4 and J5, shown in Fig. 4 between HSV-TK and pBR322,

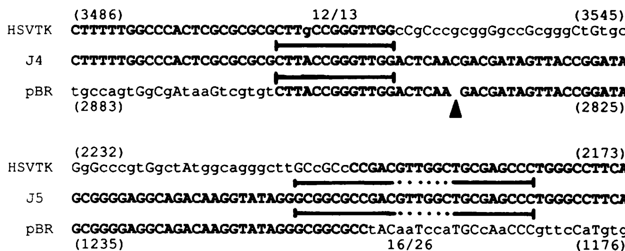

FIG. 4. Sequence of the donor molecules and novel junctions J4 and J5. The sequence shown is of the strand that reads from left to right in Fig. 1B. In each case, the top donor sequence is from HSV-TK and the bottom donor sequence is from pBR322. Boldface capital letters indicate a donor sequance which is homologous to the novel junction sequence. Capital letters indicate single nucleotides which are identical in the donor and the novel junction sequences at that position. Lowercase letters indicate nonidentical nucleotides. The first and last nucleotide of the donor sequence shown is numbered; the numbering for pBR322 is by the method of Sutcliffe (54), and that of HSV-TK is as described by D. McGeoch (personal communication) and our unpublished observations. The HSV-TK sequence is numbered from the upstream *Bam*HI site; the *Bgl*II site in the pMK fusion gene is at nucleotide 780 in this numbering system but at nucleotide 460 in that of Wagner et al. (57). Regions of partial homology between both donors have been bracketed around the novel junctions; 12 of 13 nucleotides are identical between the donors at J4, and 16 of 26 nucleotides are identical between the donors at J5. The arrowhead indicates a C insertion in J4 not found in pMK or at J6 (see Fig. 6).

are clearly different from J2 and J7 between flanking mouse DNA and the integrant. J4 and J5 occur within patches of sequence identity between the two progenitor molecules. Recombination within patches of homology between HSV-TK and pBR322 has been observed previously in cultured cells following DNA transfection (2). In fact, the breakpoint in pBR322 at J5 falls at the same nucleotide as, or one removed from, the breakpoint in a similar junction (J1) that was characterized by Anderson et al. (2). There is no obvious structural feature of the pBR322 and HSV-TK sequences which correlates with the location of every breakpoint in these two studies. However, the sequence immediately to the right of J5 (in HSV-TK; Fig. 8) bears a striking resemblance to a consensus sequence from minisatellite DNA which has been associated with several recombination events (23). This sequence also resembles the Eβ recombination hotspot (52) and the *Escherichia coli* Chi sequence (50).

Homologous recombination between extrachromosomal DNAs has been studied for a variety of eucaryotic systems (6, 13, 28, 41, 42, 55). Homologous recombination is greatly enhanced by linearizing the DNA at an appropriate site such that complementary single strands might be exposed by exonucleolytic activity (28). It is conceivable that pMK donor sequences at J4 and J5 were exposed by exonucleolytic activity and that these ends recombined with each other at patches of homology. Exonucleolytic degradation of the pMK donor molecules might have proceeded from the *Bam*HI site or from another double-strand break that occurred after microinjection. Alternatively, a single-stranded end may have provided the substrate for strand invasion at a patch of homology in a circular molecule.

J6 and J7 define the boundaries of a 532-bp fragment of repetitive mouse DNA that separates pMK from the right-flanking host DNA. The origin of this fragment is unknown; it is not homologous to pMK, to the mouse DNA flanking the integrant in λ2 or λ21, or to any mouse repetitive DNA family which is catalogued in the Genebank file of DNA sequences. We have partially characterized this newly discovered repetitive family; it is present in approximately 1,200 copies per haploid genome, mostly within 1.6- and 1.9-kb *Bgl*II fragments which are tandemly reiterated. We suggest that one or more of these repeats may have existed as an extrachromosomal piece of DNA which was ligated to a pMK fragment; chromosomal integration left only the

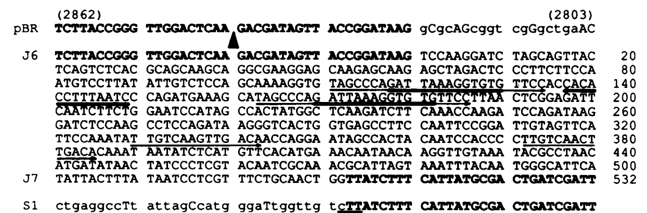

FIG. 5. Sequence of donor molecules and the 532-bp fragment of mouse DNA between J6 and J7. The sequence shown is of the strand reading from left to right in Fig. 1B. The upper donor sequence is of pBR322, as described in Fig. 4; the lower donor sequence is from S1 (Fig. 1B). The sequence of the 532-bp fragment of mouse DNA is given in capital letters and is numbered to the right of the last nucleotide in each line. The subtended arrows indicate reiterated sequence motifs, and the thick line identifies a core sequence common to the repeats between nucleotides 111 to 186. The arrowhead indicates the position of the C insertion unique to J4. The underline at CTT in donor S1 identifies a potential site for preferential topoisomerase 1 nicking.

```
S8   CAATCCCTGGGACCCACATGTTgGaGggAgagAAcgcacacaAaatgtccACtG(ac)15
J2   CAATCCCTGGGACCCACATGTTtGtGacAtgcAATCTTACGGATGGCATGACAGTAAGAG
pBR  actTggtTGaGtaCtcaccagTcacagaAaagcATCTTACGGATGGCATGACAGTAAGAG
     (3858)                                                 (3799)
```

FIG. 6. Sequence of donor molecules and J2. The significance of capital and lowercase letters are as described in the legend to Fig. 4. The filler DNA at J2 is bracketed. The overline in S8 identifies the proximal region in a stretch of alternating dPuPy, 61 nucleotides in length, which is interrupted by six nucleotides (five interruptions are shown here). The last dAC pair shown is repeated 15 times. The underline at GTT in donor S8 identifies a potential site for preferential topoisomerase 1 nicking.

532-bp remnant at the MyK-103 locus. It is also possible that pMK first integrated within a cluster of the tandem *Bgl*II repeats, then was deleted by intrachromosomal recombination between flanking *Bgl*II repeats, and finally integrated at the MyK-103 locus. Repetitive sequences have been proposed to be favorable sites for the integration of foreign DNA (24), but apart from this 532-bp fragment, we have found only one short stretch of repetitive DNA in the middle of the 5-kb duplicated sequence which flanks the integrant on either side. The flanking sequence upstream of S1 is devoid of repeat DNA for at least 2 kb, and the unique sequence extends at least 4 kb downstream of S8.

J2 and J7, between the flanking mouse DNA and the integrant, have several interesting features that distinguish them from the junctions between extrachromosomal pieces of DNA. The donor sequences from pMK, S1 and S8, are nonhomologous at junctions J2 and J7 (also, S1 and S8 do not share any patches of homology). Presumably, plasmid integration occurred at double-strand breaks within the chromosome, but we do not know the mechanism which introduced these breaks. It is curious that a recognition sequence of topoisomerase I, G/CTT (5), lies at the breakpoint in both J2 and J7. Topoisomerase I has been implicated in the nonhomologous chromosomal excision of simian virus 40 (5). Topoisomerase I has single-strand nicking activity but may cause a double-strand break if it cuts near another nick or a single-strand gap (10). In addition, the J2 breakpoint falls next to the first nucleotide in a poly(dPuPy) tract in donor sequence S8 (Fig. 6). Several other chromosomal rearrangements and integration sites have been documented near or within poly(dPuPy) tracts (9, 24, 48, 49, 53). Poly(dPuPy) tracts can form Z-DNA structures (3); DNA in transition between the B- to Z-DNA form has been postulated to exist in open strands that might facilitate DNA recombination and genome rearrangements (31).

J2 has an 11-nucleotide sequence that was not contributed by either donor sequence. Similar insertion sequences have been documented for several eucaryotic recombinant junctions. Immunoglobulin splice junctions have incorporated up to nine novel nucleotides as the product of terminal transferase action (1), and a rearranged human T cell γ gene also contains nine novel nucleotides (37). Longer novel sequences have been reported at immunoglobulin translocation breakpoints (see 13a for a review) and at viral recombination junctions, but their origins are obscure in all but one case. A 37-nucleotide stretch of filler DNA at a virus-host junction has been identified to be an inverted duplication of a single-copy host sequence found 650 bp upstream (59). Anderson et al. (2) reported three cases in which 17 to 19 nucleotides were inserted between the junctions of HSV-TK and pBR322 donor sequences that were formed in transfected tissue culture cells. On the basis of sequence homology, the filler DNA was suggested to arise from the insertion of fragments of pBR322 carrier DNA into the junction. The origin of the 11 novel nucleotides at J2 is unknown; they may have been synthesized de novo by a terminal transferaselike enzyme. Alternatively, like the 532-bp fragment, the 11 nucleotides may be a remnant of the *Bgl*II repeat family. If a copy of the *Bgl*II repeat was part of a circular extrachromosomal pMK complex, and if chromosomal integration occurred at a breakpoint within the *Bgl*II repeat DNA, then 11 nucleotides might have distributed to the left junction (J2) and 532 bp may have distributed to the right junction (J7).

Formation of the chromosomal integration complex. Although the dynamics of the integration events cannot be deduced from this analysis, the following scenario is consistent with the structure of the integrant. After microinjection,

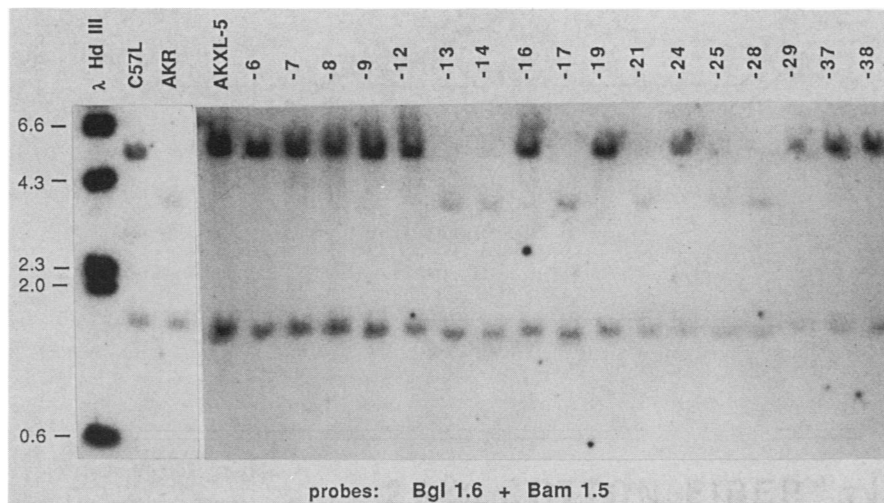

FIG. 7. RFLP segregation in AK×L recombinant inbred lines. Liver DNA (2.5 μg) was digested with *Bam*HI, Southern blotted, and hybridized with a mixture of the *Bgl* 1.6 and *Bam* 1.5 probes. DNA was isolated from C57L mice, AKR/J mice, or AK×L recombinant inbred mice, as indicated in the figure. The 5.5-kb band is inherited from the C57L parental strain; the 3.9-kb band is inherited from the AKR parental strain.

mini-satellite DNA	GGAGGTGGGCAGGaXG
Eβ recomb hotspot	GGAGGTaGGCAGGcAG
J5 HSVTK	cGAcGTtGGCTGcgAG
Chi	GCTGGtgG

FIG. 8. Comparison of J5 HSV-TK sequence with recombinational hotspots. The sequence in HSV-TK near J5 is shown, which shares homology with the human minisatellite core sequence (23), the Eβ recombinational hotspot (52), and the λ phage Chi sequence (50). Boldface capital letters indicate identity between each sequence, capital letters indicate the predominant base in that position, and lowercase letters indicate nucleotides which are unique to that sequence.

most of the linear pMK molecules probably recircularized; a few molecules undoubtedly followed other pathways, such as w and x, which ligated head-to-head to generate J3 (Fig. 1). Many of the circularized plasmids may have undergone homologous recombination to form the usual tandem arrays, but they probably never integrated and were lost. The fragments labeled y and z in Fig. 1 may have been circular or linear monomers which recombined, along with the w-x dimer, at patches of homology to generate J4 and J5. The mechanism which generated J6 between the 532-bp fragment of repetitive mouse DNA and pMK molecule z is unclear. An extrachromosomal copy of the repetitive mouse DNA could have been generated by looping out as a result of homologous intrachromosomal recombination within the tandemly arrayed *Bgl*II repeat family. Recombination between this DNA and pMK molecule z may have generated J6. Alternatively, the pMK aggregate (w, x, y, and z) may have transiently integrated into a tandem array of the *Bgl*II repeat family to generate J6. In this case, the complex subsequently deleted, still harboring a portion of the repetitive DNA. Then the complex (w, x, y, z, and the repetitive DNA) presumably integrated in a single cell of the cleavage-stage embryo or inner cell mass. The flanking 5-kb duplication of mouse DNA at the integration locus is most easily explained by postulating that integration occurred during replication. Since homologs do not pair during mitosis, recombination probably occurred between sister chromatids, possibly within a replication bubble. The integrant would have served as a linker by which the sister chromatids recombined (Fig. 9). This crossover might have left a 5-kb deletion or an unrepaired double-strand break in the other chromatid and, following segregation might have had deleterious consequences for the resulting cells some time later in development. Perhaps because the MyK-103 founder was mosaic, this was not of serious consequence to the embryo.

The MyK-103 founder was a female that carried foreign DNA in about 15% of the somatic cells and the germ line (36). Mosaicism is found in roughly 30% of all founder transgenic mice and probably occurs as a consequence of plasmid integration after one or more rounds of chromosomal DNA replication (58). The structure of the pMK integrant may have been formed during the first few cellular divisions in early embryogenesis, followed by chromosomal integration in a cleavage-stage embryo, which resulted in mosaicism of the founder female. This is consistent with the model in which integration occurs during chromosomal DNA replication. Alternatively, integration may have occurred prior to the first round of replication but may have been deleted by homologous recombination between the

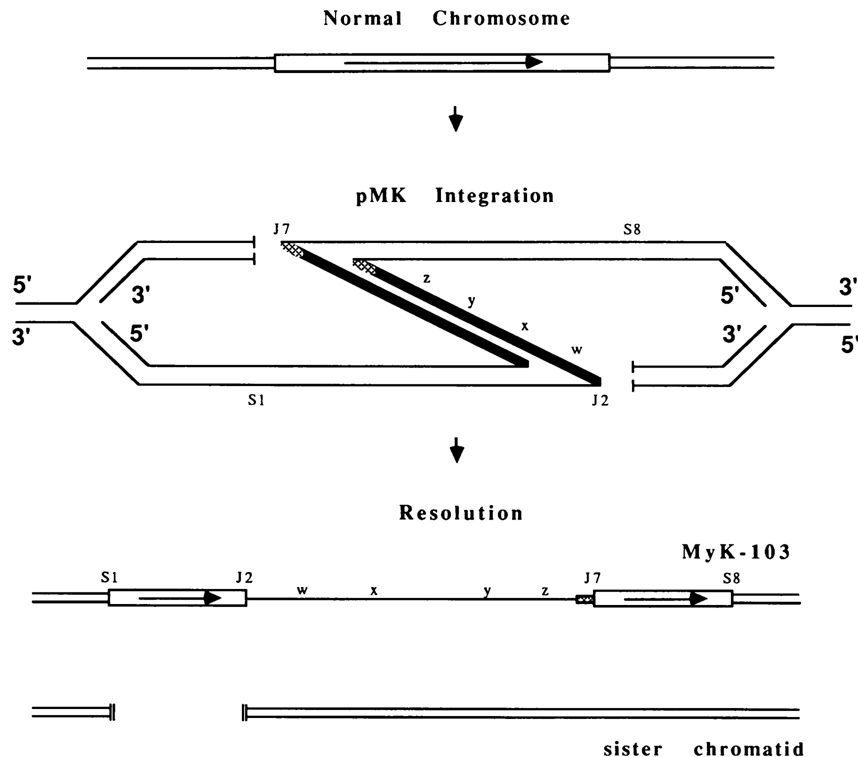

FIG. 9. Integration of pMK generated a 5-kb duplication. The top diagram shows the normal chromosome and the position of the 5-kb sequence, which was duplicated in MyK-103 (Fig. 1B), before DNA replication begins. The middle diagram shows the two strands of DNA during replication. Integration of pMK (▬; w, x, y, and z) at breakpoints in sister chromatids may have caused the 5-kb duplication. The bottom diagram shows the MyK-103 and sister chromatid following separation. Four fragments of pMK (w, x, y, and z) and the 532-bp fragment of repetitive mouse DNA are identified as in Fig. 1.

5-kb repeats in the cleavage-stage embryo. Clonal expansion of the two cell types in the inner cell mass could have subsequently resulted in mosaicism in the adult MyK-103 founder.

Insertional mutation and transmission ratio distortion. The MyK-103 pedigree is maintained in the heterozygous state because males do not transmit the foreign DNA to offspring. We have assayed germ line transmission from 17 different males, yet not one of their 352 pups has inherited the foreign DNA. There is no apparent reduction in litter size associated with paternal transmission, since outbred males and females have an average of eight pups per litter. It therefore seems likely that the block to paternal transmission occurs prior to fertilization.

One explanation is that the MyK-103 integrant inactivates a gene that must function during spermatogenesis or activates a gene that should be quiescent. As a consequence of either defect, sperm bearing the MyK-103 integrant are postulated to be infertile. The structure of the integration complex suggests that infertility might be relieved by deletion of the integrant via intrachromosomal recombination within the flanking duplicated sequences or by unequal crossing over between sister chromatids. To ascertain whether recombination occurs, we took advantage of the deletion that lies within the *Bgl* 1.6 fragment in the 5-kb duplication that flanks the integrant in the AKR strain of mice (Fig. 2). MyK-103 females were mated to AKR males, and G_1 males were backcrossed to AKR females. The G_2 progeny were scored for the transmission of the AKR chromosome 6 bearing the deletion or the C57/SJL chromosome 6 bearing the transgene by using a probe that lies within the deletion in the AKR strain. Both chromosomes were transmitted with equal frequency, and Southern blots indicated that the chromosome that bore the MyK-103 integrant resembled wild-type C57/SJL chromosome 6 in the G_2 progeny (T. Wilkie, R. Hammer, and R. Palmiter, unpublished observations). Mature sperm carrying the MyK-103 integrant are present in the cauda epididymis, albeit at somewhat reduced abundance relative to sperm without the integrant (36; unpublished observations). Therefore, we assume that deletion of the integrant in male germ cells restores function and leads to clones of wild-type sperm that develop in an otherwise sterile mouse. This model and the data supporting it will be developed more fully elsewhere.

Obviously, we would like to find the gene that is presumably disrupted in germ cells of MyK-103 males. We have isolated and tested several unique hybridization probes derived from λ2 and λ21, which span more than 7 kb of mouse DNA, but we have not yet detected any discrete transcripts in Northern blots of testis RNA from either heterozygous MyK-103 males or control mice; however, we may not have tested a probe corresponding to exonic sequences, or the transcripts might be rare. It is possible that HSV-TK expression is responsible for the infertility, although this seems unlikely because several other lines of transgenic mice that contain pMK and express HSV-TK show normal male germ line transmission.

Several reports of recessive mutations that cosegregate with integrated foreign DNA have generated interest in exploiting insertional mutagenesis in transgenic mice. In these lines, integration of the foreign DNA presumably causes a mutation in a gene that is crucial for development and may provide a convenient tag for isolating the normal genetic locus. However, in contrast to precise integration that is characteristic of retroviruses, all of the characterized integrants in transgenic mice that were generated by microinjection are associated with deletions (60), duplications (as in MyK-103), or more complex chromosomal rearrangements (15) and translocations (20, 33). Such anomalies could seriously hinder the analysis of the mutation and cloning of the normal gene. We have suggested that the duplication in MyK-103 is generated by integration of foreign DNA between sister chromatids. A similar mechanism could generate deletions of indeterminate length. Chromosomal translocations might also be generated by an analogous mechanism if the ends of the foreign DNA integrated into two different chromosomes. Thus, in general, integration of foreign DNA may potentially affect many genes within large regions of a chromosome.

ACKNOLWEDGMENTS

We are grateful to Ralph Brinster and Bob Hammer for establishing and maintaining the MyK-103 line. We thank Eric Mercer for his help with sequence analysis, Patricia Ashley and David Cox for providing DNA from hamster-mouse somatic cell hybrids, Lee Hood for the Vβ8 probe, and Roger Perlmutter, Bob Braun, and Jacques Peschon for their helpful suggestions during the preparation of this manuscript.

This work was supported in part by Public Health Service grant HD-09172 from the National Institutes of Health.

LITERATURE CITED

1. **Alt, F. W., N. Rosenberg, V. Erea, E. Siden, and D. Baltimore.** 1982. Multiple immunoglobulin heavy-chain gene transcripts in Abelson murine leukemia virus-transformed lymphoid cell lines. Mol. Cell. Biol. **2**:386–400.
2. **Anderson, R. A., S. Kato, and R. D. Camerini-Otero.** 1984. A pattern of partially homologous recombination in mouse L-cells. Proc. Natl. Acad. Sci. USA **81**:206–210.
3. **Arnott, S., R. Chandrasekaran, and D. L. Birdsall.** 1980. Left handed DNA helices. Nature (London) **283**:743–745.
4. **Barth, R. K., R. S. Kim, N. C. Lan, T. Hunkapiller, N. Sobieck, A. Winoto, H. Gershenfeld, C. Okada, D. Hansburg, I. L. Weissman, and L. Hood.** 1985. The murine T-cell receptor uses a limited repertoire of expressed Vβ gene segments. Nature (London) **316**:517–523.
5. **Been, M. D., R. R. Burgess, and J. J. Champoux.** 1984. Nucleotide sequence preference at rat liver and wheat germ type 1 DNA topoisomerase breakage sites in duplex SV40 DNA. Nucleic Acids Res. **12**:3097–3114.
6. **Brenner, D. A., S. Kato, R. A. Anderson, A. C. Smigocki, and R. D. Camerini-Otero.** 1984. The recombination and integration of DNAs introduced into mouse L-cells. Cold Spring Harbor Symp. Quant. Biol. **49**:151–160.
7. **Brinster, R. L., H. Y. Chen, M. Trumbauer, A. W. Senear, R. Warren, and R. D. Palmiter.** 1981. Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. Cell **27**:223–231.
8. **Brinster, R. L., H. Y. Chen, M. E. Trumbauer, M. K. Yagle, and R. D. Palmiter.** 1985. Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs. Proc. Natl. Acad. Sci. USA **82**:4438–4442.
9. **Bullock, P., J. J. Champoux, and M. Botchen.** 1985. Association of crossover points with topoisomerase I cleavage sites: a model for non-homologous recombination. Science **230**:954–958.
10. **Bullock, P., J. Miller, and M. Botchan.** 1986. Studies of the effects of poly d(pGpT) · d(pApC) and poly d(pCpG) · (pCpG) repeats on homologous recombination in somatiic cells. Mol. Cell. Biol. **6**:3948–3953.
11. **Caccia, N., M. Kronenberg, D. Saxe, R. Haars, G. A. P. Bruns, J. Goverman, M. Malissen, H. Willard, Y. Yoshikai, M. Simon, L. Hood, and T. W. Mak.** 1984. The T cell receptor β chain genes are located on chromosome 6 in mice and chromosome 7 in humans. Cell **37**:1091–1099.
12. **Calos, M. P., and J. H. Miller.** 1980. Transposable elements. Cell **20**:579–595.

13. **Carroll, D., S. H. Wright, R. K. Wolff, E. Grzbieta, and G. B. Maryon.** 1986. Efficient homologous recombination of linear DNA substrates after injection into *Xenopus laevis* oocytes. Mol. Cell. Biol. **6**:2053–2061.

13a. **Cory, S.** 1986. Activation of cellular oncogenes in hematopoietic cells by chromosomal translocation. Adv. Cancer Res. **47**:189–234.

14. **Courage-Tebbe, V., H. P. Doring, N. Fedoroff, and P. Starlinger.** 1983. The controlling element DS at the shrunken locus in Zea mays: structure of the unstable sh-m 5933 allele and several revertants. Cell **34**:383–393.

15. **Covarrubias, L., Y. Nishida, and B. Mintz.** 1985. Early developmental mutations due to DNA rearrangements in transgenic mouse embryos. Cold Spring Harbor Symp. Quant. Biol. **50**:447–452.

16. **Cox, D. R., J. A. Sawicki, D. Yee, E. Appella, and C. J. Epstein.** 1982. Assignment of the gene for β_2-microglobulin (Bzm) to mouse chromosome 2. Proc. Natl. Acad. Sci. USA **79**:1930–1934.

17. **Davis, M. M., K. Calame, P. W. Early, D. L. Livant, R. Joho, I. L. Weismann, and L. Hood.** 1980. An immunoglobulin heavy-chain gene is formed by at least two recombinational events. Nature (London) **283**:733–739.

18. **Folger, K. R., E. A. Wong, G. Wahl, and M. R. Capecchi.** 1982. Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. Mol. Cell. Biol. **2**:1372–1387.

19. **Fristensky, B.** 1986. Improving the efficiency of dot-matrix similarity searches through use of an oligomer table. Nucleic Acids Res. **14**:597–610.

20. **Gordon, J. W., and F. H. Ruddle.** 1985. DNA-mediated genetic transformation of mouse embryos and bone marrow—a review. Gene **33**:121–136.

21. **Gropp, A., and H. Winking.** 1981. Robertsonian translocations: cytology, meiosis, segregation patterns and biological consequence of heterozygosity. Symp. Zool. Soc. Lond. **47**:141–181.

22. **Hutchison, K. W., N. G. Copeland, and N. A. Jenkins.** 1984. Dilute-coat-color locus of mice: nucleotide sequence analysis of the d^{+2J} and d^{dta} revertant alleles. Mol. Cell. Biol. **4**:2899–2904.

23. **Jefferys, A. J., V. Wilson, and S. L. Thein.** 1985. Hypervariable minisatellite regions in human DNA. Nature (London) **314**: 67–73.

24. **Kato, S., R. A. Anderson, and R. O. Camerini-Otero.** 1986. Foreign DNA introduced by calcium phosphate is integrated into repetitive DNA elements of the mouse L cell genome. Mol. Cell. Biol. **6**:1787–1795.

25. **King, W., M. D. Patel, L. I. Lobel, S. P. Goff, and M. C. Nguyen-Huu.** 1985. Insertion mutagenesis of embryonal carcinoma cells by retroviruses. Science **228**:554–558.

26. **Leach, P. R. F., and F. W. Stahl.** 1983. Viability of λ phages carrying a perfect palindrome in the absence of recombination nucleases. Nature (London) **305**:448–451.

27. **Lee, N. E., P. D'Eustachio, D. Praftcheva, F. H. Ruddle, S. M. Hedrick, and M. M. Davis.** 1984. Murine T cell receptor beta chain is encoded on chromosome 6. J. Exp. Med. **160**:905–913.

28. **Lin, F. L., K. Spearle, and N. Sternberg.** 1984. Model for homologous recombination during transfer of DNA into mouse L cells: role for DNA ends in the recombination process. Mol. Cell. Biol. **4**:1020–1034.

29. **Mark, W. H., K. Signorelli, and E. Lacy.** 1985. An insertional mutation in a transgenic mouse line results in developmental arrest at day 5 of gestation. Cold Spring Harbor Symp. Quant. Biol. **50**:453–464.

30. **Mount, D. W., and B. Conrad.** 1986. Improved programs for DNA and protein sequence analysis on the IBM personal computer and other standard computer systems. Nucleic Acids Res. **14**:443–454.

31. **Nordheim, A., and A. Rich.** 1983. The sequence (dC · dA)n · (dG-dT)n forms left-handed Z-DNA in negatively supercoiled plasmids. Proc. Natl. Acad. Sci. USA **80**:1821–1825.

32. **Nusse, R., and H. E. Varmus.** 1982. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. Cell **31**:99–109.

33. **Overbeek, P. A., S. P. Lai, K. R. Van Quill, and H. Westphal.** 1986. Tissue-specific expression in transgenic mice of a fused gene containing RSV terminal sequences. Science **231**:1574–1576.

34. **Palmiter, R. D., and R. L. Brinster.** 1986. Germline transformation of mice. Annu. Rev. Genet. **20**:465–499.

35. **Palmiter, R. D., H. Y. Chen, and R. L. Brinster.** 1982. Differential regulation of metallothionein-thymidine kinase fusion genes in transgenic mice and their offspring. Cell. **29**:701–710.

36. **Palmiter, R. D., T. M. Wilkie, H. Y. Chen, and R. L. Brinster.** 1984. Transmission distortion and mosaicism in an unusual transgenic mouse pedigree. Cell **36**:869–877.

37. **Quertermous, T., W. Strauss, C. Murre, D. P. Dialynas, J. L. Strominger, and J. G. Seidman.** 1986. Human T-cell γ genes contain N segments and have marked junctional variability. Nature (London) **322**:184–187.

38. **Robins, D. M., S. Ripley, A. S. Henderson, and R. Axel.** 1981. Transforming DNA integrated into the host chromosome. Cell **23**:29–39.

39. **Roeder, G. S., P. J. Farabaugh, D. T. Ghaleff, and G. R. Fink.** 1980. The origins of gene instability in yeast. Science **209**:1375–1380.

40. **Roehm, N. W., A. Carbone, E. Kushnir, B. A. Taylor, R. J. Riblet, P. Marrach, and J. W. Kappler.** 1985. The major histocompatibility complex-restricted antigen receptor on T cells: the genetics of expression of an allotype. J. Immunol. **135**:2176–2182.

41. **Roth, D. B., T. N. Porter, and J. H. Wilson.** 1985. Mechanisms of nonhomologous recombination in mammalian cells. Mol. Cell. Biol. **5**:2599–2607.

42. **Rubnitz, J., and S. Subramani.** 1984. The minimum amount of homology required for homologous recombination in mammalian cells. Mol. Cell. Biol. **4**:2253–2258.

43. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

44. **Schnieke, A., A. Harbers, and R. Jaenisch.** 1983. Embryonic lethal mutation in mice induced by retrovirus insertion into the α1(I) collagen gene. Nature (London) **304**:315–320.

45. **Shani, M.** 1986. Tissue-specific and developmentally regulated expression of a chimeric actin-globin gene in transgenic mice. Mol. Cell. Biol. **6**:2624–2631.

46. **Silver, J., and C. E. Buckler.** 1986. Statistical consideration for linage analysis using recombinant inbred strains and back crosses. Proc. Natl. Acad. Sci. USA **83**:1423–1427.

47. **Sim, G. K., and A. A. Augustin.** 1985. Vβ gene polymorphism and a major polyclonal T cell receptor idiotype. Cell **42**:89–92.

48. **Skowronski, J., A. Pwcienniczak, A. Bednarek, and J. Jaworski.** 1984. Bovine 1.709 satellite: recombination hotspots and dispersed repeated sequences. J. Mol. Biol. **177**:399–416.

49. **Slightom, J. L., A. E. Blechl, and O. Smithies.** 1980. Human fetal G_γ and A_γ-globin genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell **21**:627–638.

50. **Smith, G. R., S. M. Kimes, D. W. Schultz, A. Taylor, and K. L. Trinan.** 1981. Structure of Chi hotspots of generalized recombination. Cell **24**:429–436.

51. **Spradling, A. C., and G. M. Rubin.** 1981. Drosophila genome organization: conserved and dynamic aspects. Annu. Rev. Genet. **15**:219–264.

52. **Steinmetz, M., D. Stephan, and K. F. Lindahl.** 1986. Gene organization and recombinational hotspots in the murine major histocompatibility complex. Cell **44**:895–904.

53. **Stringer, J. R.** 1985. Recombination between poly d(GT) · d(CA) sequences in simian virus 40-infected culture cells. Mol. Cell. Biol. **5**:1247–1259.

54. **Sutcliff, J. G.** 1978. Complete nucleotide sequence of the Escherichia coli plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. **43**:77–90.

55. **Thomas, V. K., K. R. Folger, and M. R. Capecchi.** 1986. High

frequency targeting of genes to specific sites in the mammalian genome. Cell **44**:419–428.

56. **Wagner, E. F., L. Covarrubias, T. A. Stewart, and B. Mintz.** 1983. Prenatal lethalities in mice homozygous for human growth hormone gene sequences integrated in the germ line. Cell **35**:647–655.
57. **Wagner, M. J., J. A. Sharp, and W. C. Summers.** 1981. Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **78**:1441–1445.
58. **Wilkie, T. M., R. L. Brinster, and R. D. Palmiter.** 1986. Germline and somatic mosaicism in transgenic mice. Dev. Biol. **118**:9–18.
59. **Williams, T. J., and M. Fried.** 1986. Inverted duplication-transposition event in mammalian cells at an illegitimate recombination join. Mol. Cell. Biol. **6**:2179–2184.
60. **Woychik, R. P., T. A. Stewart, L. G. Davis, P. D'Eustachio, and P. Leder.** 1985. An inherited limb deformity created by insertional mutagenesis in a transgenic mouse. Nature (London) **318**:36–40.