MOLECULAR AND CELLULAR BIOLOGY, July 1989, p. 3049–3057
0270-7306/89/073049-09$02.00/0

Vol. 9, No. 7

Comparison of Filler DNA at Immune, Nonimmune, and Oncogenic Rearrangements Suggests Multiple Mechanisms of Formation

DAVID B. ROTH,† XIU-BAO CHANG,‡ AND JOHN H. WILSON*

Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030

Received 19 January 1989/Accepted 19 April 1989

Extra nucleotides (termed filler DNA) are commonly found at the junctions of genetic rearrangements in mammalian cells. The filler DNA at immune system rearrangements, which are called N regions, are generated at VDJ joints primarily by terminal deoxynucleotidyl transferase. However, the origin of filler DNA at genetic rearrangements in nonlymphoid cells is uncertain. In an analysis of more than 200 junctions that arose by circularization of transfected linear DNA (D. B. Roth and J. H. Wilson, Mol. Cell. Biol. 6:4295–4304, 1986), we found 18 junctions with extra nucleotides exactly at the point of circularization. Analysis of these 18 junctions indicated that nonlymphoid cells could add extra nucleotides to the ends of duplex DNA. The characteristics of the extra nucleotides at these junctions and at 31 other rearrangement junctions from nonlymphoid cells were quite similar, suggesting that many genetic rearrangements may pass through a stage with free DNA ends. A comparison of the filler DNA at these 49 nonimmune system rearrangements with 97 N regions derived from immune system rearrangements suggested that lymphoid and nonlymphoid cells use different mechanisms for insertion of filler DNA, as expected from the absence of detectable terminal deoxynucleotidyl transferase in cells from nonlymphoid tissues. The filler DNAs at a smaller group of 22 translocations associated with cancer had features in common with both immune and nonimmune system rearrangements and therefore may represent a mixture of these two processes. Mechanisms that might account for the presence of filler DNA in nonlymphoid cells are discussed.

Extra nucleotides of filler DNA are common at the junctions of genomic rearrangements in mammalian cells but are rare in bacteria and yeasts. Filler DNAs at the junctions between V, D, and J gene segments in immune system rearrangements are known as N regions. The frequency of N regions ranges from less than 5% for immunoglobulin light-chain genes to as much as 90% for T-cell beta-chain genes (9, 31). Filler DNA is also found at about half of the reciprocal translocations associated with cancer, many of which apparently result from mistakes accompanying immune system rearrangements (3, 38, 57). Lymphoid cells, however, are not unique in their ability to incorporate extra nucleotides, since about 10% of the junctions in nonimmune system rearrangements also have segments of filler DNA (49, 50). We will refer to filler DNAs at nonimmune system rearrangements as inserts to distinguish them from N regions at immune system rearrangements.

Rearrangements of immunoglobulin and T-cell receptor genes are thought to occur by a mechanism involving site-specific cleavage to generate DNA ends, followed by end joining to reunite the segments in a new order (1, 23). Terminal deoxynucleotidyl transferase (TdT) plays a major role in the addition of N regions to the broken ends generated by the cleavage event (1, 31). The presence of inserts at nonimmune system rearrangement junctions is also suggested to result from addition of nucleotides to broken ends before joining (49, 63). Since TdT has not been detected in nonlymphoid cells, however, it is likely that a TdT-independent mechanism is responsible for generating inserts (7).

Here we characterize the inserts at 18 nonimmune system junctions that were generated by circularization of linear DNA that was transfected into mammalian cells. The properties of these inserts (along with 31 others from nonlymphoid cells) are contrasted with the properties of N regions at immune system rearrangements and filler DNAs at translocation junctions. These comparisons suggest that multiple mechanisms may be involved in generating filler DNA in nonlymphoid cells.

MATERIALS AND METHODS

Cells and viruses. The CV1 monkey kidney cell line was grown according to standard procedures (63). The simian virus 40 (SV40) mutant su1910 was derived from the Rh911a wild-type strain as described previously (50); it contains a 45-nucleotide-long polylinker in place of 220 nucleotide pairs from the intron of the SV40 T-antigen gene.

DNA transfection. Linear SV40 DNA molecules were produced by digesting su1910 viral DNA with restriction enzymes that cleaved within the polylinker. Substrates containing mismatched ends were produced by digestion with a pair of enzymes. Restriction enzymes were used according to the recommendations of the suppliers. DNA transfections were carried out as described previously (63), with DEAE-dextran as carrier and 0.005 to 0.01 ng of SV40 DNA per 60-mm-diameter dish. This small quantity of DNA ensured that single DNA molecules initiated plaque formation (50). A single plaque was picked from each plate to ensure that each junction arose in an independent recombination event.

DNA sequencing. To prepare viral DNA for sequencing, confluent monolayers of CV1 cells in 24-well plates were

* Corresponding author.
† Present address: Laboratory of Pathology, National Cancer Institute, Bethesda, MD 20892.
‡ Present address: Department of Cell Biology, Baylor College of Medicine, Houston, TX 77030.

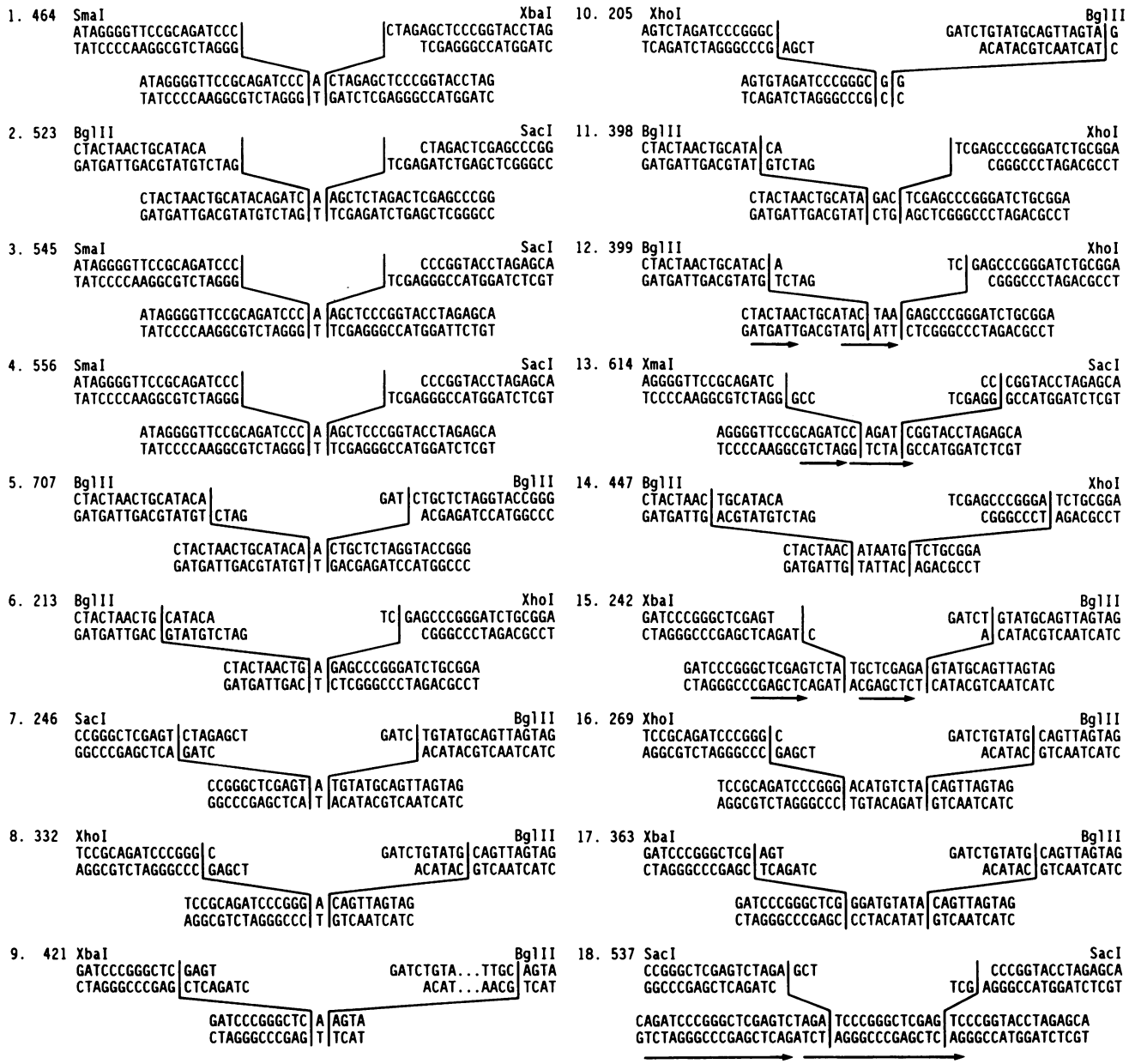

FIG. 1. Sequences across 18 circularization junctions that contain extra nucleotides. Numbers are plaque isolation numbers. The first line shows sequences of the ends of the transfected DNA, along with the restriction enzymes used to create the ends. Sequences of the circularization junctions are shown below. Vertical lines define the position of the junction relative to the parental sequences and flank the inserted nucleotides in the circularization junctions. Symbols: ⟶, directly repeated sequences that include the added nucleotides; • • •, 45 nucleotides omitted to permit the display of the large deletion in junction 9.

infected with picked plaque suspensions. Viral DNA was harvested and treated as described previously (50). Viral DNAs were subjected to double-strand sequence analysis (66), using minor modifications that have been described previously (50).

RESULTS AND DISCUSSION

Characteristics of inserts at sites of circularization. SV40 DNA molecules were linearized by cleavage in a polylinker in the intron of the T-antigen gene and then were introduced into CV1 monkey cells by DEAE-dextran-mediated transfection. Since the viral genome must circularize to initiate lytic infection and plaque formation, transfection with linear molecules provides a selection for cell-mediated end-joining events. We isolated more than 200 independent plaques from several different linear molecules. Viral DNA was prepared from each plaque, and the recombinant junctions were sequenced. In more than 90% of the junctions, one parental sequence was joined directly to another without the addition of extra nucleotides. These junctions (and possible mechanisms for their formation) have been described elsewhere (50). Eighteen of the junctions contained one or more extra nucleotides exactly where the transfected molecule circularized. The junctions with inserts are the focus of this paper.

In Fig. 1, each circularization junction is shown below its

TABLE 1. Filler DNA at genetic rearrangement junctions in mammalian cells: nonimmune system rearrangements[a]

5′ Sequence	Insert	3′ Sequence	Source	Reference
GAAATGTGCG	A	GTCAAACATG	Transfection	49
CCACAGCTGG	A	CAATCCTGAT	SV40 evolutionary variant	52
CCAGTTCCGC	A	GACTAATTTT	Transfection	11
ACTGTGGGAT	A	CCATTTTCTT	Adenovirus-SV40 hybrid	33
CCTGAAGCTG	A	AACATACAGG	Deletion	39
AAAGCCTAGG	G	AAAAAAGCCT	Transfection	11
AGTGAGGAGG	G	TTGGAGGCCT	Transfection	11
TACTAACTGC	AT	ATCTGCTCTA	Transfection	—[b]
CGGGCCAGAC	GT	GCTTGGACAT	Viral integration	12
AGGA	ACA	GGAG	Deletion	36
TAAAAAAAAT	GCA	GGCGGAACTG	Transfection	11
GGAGGCCTGG	ACGG	AGCACCATGG	Transfection	65
TCTAAAACAA	TTAGA	GCCTGAAATA	Transfection	65
AAGCATGAAA	CAATG	TTTAACTGTG	Transfection	63
TTAGGTGTTC	TGTAG	CAGAAAGGAG	Deletion	39
CATACAGATC	AGAGC	AGGTACCGGG	Transfection	—[b]
AGGTACCGGG	GGATC	GGGATCTGCG	Transfection	—[b]
GGCTGAGACC	AGTTGT	CCAAAGTTGC	Deletion	24
GGCAAAACAA	AGCAAA	ACTTTGCTTT	Deletion	14a
AAATTGTAAC	AAGTAGA	GATTCAAGTA	Deletion	54
GAGGAGGCTT	CTGGAGAG	GCAAAGATGG	Transfection	65
CCCACATGTT	TGTGACATGCA	ATCTTACGGA	Plasmid integration	61
AGATTGAGCT	GCCGGTCTGTCGCC	GGCGGGAGCT	Transfection	35
GCAAACAATG	TACCTACAAACAATGTA	GGACAAACTA	Transfection	60
TGGATTGTAG	ATGCGGAGTATGAGTAT	GGCTGCTGGA	Transfection	2
GACGGGGAGA	TCTTGAAACGAGAGAGGA	TGAATACTCA	Transfection	2
TGGCAGAACA	AGTTGGGTCTTGCGGCGGA	GCCGCGGCGG	Transfection	2
CTTTC	GTATAATTAAATTCAGATAATTTAA	TTACA	Transfection	42
TTGGGCTGGA	TGAAGCTGGAAACTATCATTCTCAGCAAACTATCAC	ACCCTGCTTA	Deletion	34
GTCTACCATG	GGTGCAGATGTGGTGGCAAAGAATTTCATAGCACTCA	TTCGTTGTGT	Viral integration	62
TGCACAAGG	ACAGGGAAGCCTGGTATGCTGCAATCTGTGGGATTGCAAA	GATCACAAA	Viral integration	5

[a] For all junctions, the strand containing the filler DNA with the highest fraction of purines is displayed. Where possible, 10 nucleotides on each side of the added nucleotides at the junction are shown. In all cases, the indicated length of filler DNA is a minimum estimate. If one or more nucleotides at the ends of the filler DNA segment matched the parental DNA sequence, they would be counted as parental DNA. Filler DNA is operationally defined as extra nucleotides that cannot be accounted for by the parental DNA sequence.

[b] D. B. Roth, L. Stewart, and J. H. Wilson, unpublished data.

input ends, and lines are drawn to indicate the relationship between the two and to highlight the extra nucleotides at the junction. Although the number of junctions is not large, there are some distinctive features. In all cases, the extra nucleotides were at or near the ends of the input DNA. Of the 36 ends that contributed to the 18 junctions, 19 lost either no nucleotides at all or nucleotides only within the terminal single-stranded regions, an additional 15 ends lost less than 10 nucleotide pairs (counting from the duplex portion of the input terminus), and only 2 ends lost more than 10 nucleotide pairs. Therefore, it seems likely that the extra nucleotides are added to free DNA ends (after some strands are trimmed by a nuclease or extended by a polymerase). In addition, there seemed to be a high frequency of one-nucleotide inserts (10 of 18 junctions) and a tendency toward high A+T content (40 of 66 inserted nucleotide pairs were A · T). Curiously, all but one of the single-nucleotide inserts were A · T nucleotide pairs.

Comparison with other nonimmune system junctions containing extra nucleotides. A variety of genetic rearrangements have been characterized in mammalian cells from nonlymphoid tissues. These include deletions and rearrangements in transfected molecules and viral genomes, deletions and rearrangements of chromosomal DNA, and integration of foreign DNA into the chromosome. Junctions with extra nucleotides have been reported in each category (Table 1). In all, 31 junctions (about 10%) contain extra nucleotides; their sequences are shown in Table 1. They are similar to the junctions in Fig. 1 in that one-nucleotide inserts are the most frequent (7 of 31) and the inserts are A+T rich (nearly 60% A+T). The distributions of these two populations of junctions are shown together in Fig. 2A.

The similarities in the frequency and composition of extra nucleotides at circularization junctions and other nonimmune system rearrangement junctions suggested that they might be added by the same mechanism. Since extra nucleotides are added to the free ends of transfected DNA molecules, they may also be added to free ends during other genetic rearrangements in nonlymphoid cells. This notion supports the idea that many genetic rearrangements in mammalian cells may arise in a two-step process in which ends are generated in the first step and then joined in the second (50).

Comparison with junctions at immune system rearrangements. Individually, there is little to distinguish an N region from an insert (Tables 1 and 2). However, a comparison of the populations of N regions and inserts reveals several differences (Fig. 2A and B; Table 4). None of the 97 N regions were longer than 13 nucleotides; 9 of the 49 inserts (18%) were 14 nucleotides or longer. N regions decreased progressively in number with increasing length; inserts showed a distinct peak of 1-nucleotide additions (35%), with a relatively flat distribution from 2 to 40 nucleotides. The overall G+C content of N regions was 57%; the overall G+C content of inserts was 38%. Taken together, these differences support an underlying difference in mechanism of addition.

The generation of N regions by TdT addition of nucleo-

TABLE 2. Filler DNA at genetic rearrangement junctions in mammalian cells: immune system rearrangements[a]

5′ Sequence	N region	3′ Sequence	Gene	Reference
CGAACACCCA	A	AATGTTGCTG	Immunoglobulin lambda light chain	25
CGAACACCCA	A	AATGGTGCTG	Immunoglobulin lambda light chain	25
CGAACACCCA	A	AAGGTTGCTG	Immunoglobulin lambda light chain	25
ATTTCCAGAA	A	CCCCCTGTCC	T-cell beta chain	53
GGGGACAGG	A	TTCTGGAAAT	T-cell beta chain	26
GCTCGGACTA	A	CACAGATACG	T-cell beta chain	9
CTAATCAGGG	A	ACTGAAGCTT	T-cell beta chain	9
GGATCC	A	GACGTTCGGT	Immunoglobulin kappa light chain	31
GGATCCT	G	GTGGACGTTC	Immunoglobulin kappa light chain	31
CAGTAGTCAA	G	GCACTGTGGT	Immunoglobulin heavy chain	29, 30
CCAGTAGTCA	G	GCTACTACCG	Immunoglobulin heavy chain	16
GCTTTTCCGC	G	AGCTGCTGGC	T-cell beta chain	9
CGCCACCTCC	G	CGGCTGCTGG	T-cell beta chain	9
GGTGTCTTGG	G	CCCCCAGTCC	T-cell beta chain	53
TCTGGATGAG	AT	ATAGCTCGGG	T-cell gamma chain	21
GGTACC	AT	ATTCACGTT	Immunoglobulin kappa light chain	32
GGATCCTC	AC	GTGGACGTTC	Immunoglobulin kappa light chain	31
TACGGTAGTA	AA	ACTATGCTAT	Immunoglobulin heavy chain	16
AGACCCTGTC	TG	TTTGCTGGCA	T-cell beta chain	9
GCCCCCATAA	AG	GGTCACTGCA	T-cell alpha chain	64
GTAGTCAAA	AG	GTAGTTACCA	Immunoglobulin heavy chain	29
CCGAACGTCC	AG	GGAGGATCC	Immunoglobulin kappa light chain	31
CCGAACGTCC	AG	GGAGGATCC	Immunoglobulin kappa light chain	31
TTACTGTGCA	GA	CCTACTATAG	Immunoglobulin heavy chain	47
GGATCCTCC	GA	ACGTTCGGTG	Immunoglobulin kappa light chain	31
GGATCCTCC	GA	ACGTTCGGTG	Immunoglobulin kappa light chain	31
ACTGTGCAAG	GG	CTATGGTAAC	Immunoglobulin heavy chain	16
TGTGCAAGAC	GG	ATAAGCCCCC	Immunoglobulin heavy chain	10
GTAGTCAAAG	GG	AGCCGTAGTA	Immunoglobulin heavy chain	29
TGGTAGCTAC	GG	CTATGCTATG	Immunoglobulin heavy chain	48
CAAACCAGGC	GG	GTTG	Immunoglobulin heavy chain	48
ATCCAGGTGT	GG	ACTGCTGGCA	T-cell beta chain	9
CCCCTGTCCC	GG	TCTAGTGGCA	T-cell beta chain	9
CTGGGTCTCT	GG	CCCCTGTCCC	T-cell beta chain	9
AGAACTGCTC	GG	CTGTCCCGAC	T-cell beta chain	9
TGCCAGCAGG	GG	CAGGGGCACA	T-cell beta chain	9
ATTACTGTGC	TGA	TTTATTACTA	Immunoglobulin heavy chain	16
GGACTGGGGG	AGT	TGAGCAGTTC	T-cell beta chain	15
TGCCAGCAGC	GTA	GGACAATCGA	T-cell beta chain	9
TGCCAGCAGT	AGC	ACAAGAGGTG	T-cell beta chain	9
CCCAGGACAG	AGC	AATCAGCCCC	T-cell beta chain	9
ATCGCTAGTC	GAC	TCTCTAGCAC	T-cell beta chain	9
CGGCTGTCCC	GAC	GCTGCTGGCA	T-cell beta chain	9
TAACTATGGT	GAC	TACTATGCTA	Immunoglobulin heavy chain	27
CCAGCAGCTT	CGG	GGGACAGGGT	T-cell beta chain	9
CAGCTCGCGG	AAA	ACCCATTACT	T-cell beta chain	9
AGTAGTCAAA	AGA	AGCTACTACC	Immunoglobulin heavy chain	29
TGTTTTGACT	AGG	CCCCCAGTCC	T-cell beta chain	40
TAGTCAAAGT	GGG	GCTACTACCG	Immunoglobulin heavy chain	16
TCTCCCTCCC	TTGG	GCTGCTGGCA	T-cell beta chain	9
CTGTGCCAGA	CATA	ACTATGGTGA	Immunoglobulin heavy chain	27
AAGTACCAGT	GATA	ACCATCATAG	Immunoglobulin heavy chain	16
CAACTGGG	TGAG	GCTATGGAC	Immunoglobulin heavy chain	48
AAAACAGGGG	ATGG	GAGACCCAGT	T-cell beta chain	9
TAAGCAAACC	GGTA	GGTACTTCGA	Immunoglobulin heavy chain	1
GTGTTAGTCC	GAGC	GCTGCTGGCA	T-cell beta chain	9
TGCTCTGTCC	TGGG	TCACTGGTGG	T-cell beta chain	9
AGTAGTCAAA	AAAG	GCTACTACCG	Immunoglobulin heavy chain	46
GGGGACAGGG	AGAG	GTGCAGAAAC	T-cell beta chain	15
GTGCAAGACA	AAGG	ACGGTAGTAG	Immunoglobulin heavy chain	16
CAACTGGGAC	AAGG	CTACTGGGGC	Immunoglobulin heavy chain	29, 51
CCCCCAGTCC	GGGA	ACTGCTGGCA	T-cell beta chain	17
ATGCCACTAC	GGGG	GCCTGGTTTG	Immunoglobulin heavy chain	10
ACTGGGGGGC	GGGG	ATGAACA	T-cell beta chain	17
GTCGCCCCTC	TATAA	CGGCTACTGG	T-cell beta chain	9
GGCCCCAGAG	CTAGA	CCCCGCTTGG	T-cell beta chain	9
TTATAGAGGG	GCGAC	CACAGATACG	T-cell beta chain	9
GGCTGATTGC	CTGAG	CCCCTCCAGG	T-cell beta chain	9

Continued on following page

TABLE 2—*Continued*

5′ Sequence	N region	3′ Sequence	Gene	Reference
GACTGGGGGG	AGACA	GCAAACTCCG	T-cell beta chain	40
CTGTGCAAGA	TAAGG	ACTATAAGCC	Immunoglobulin heavy chain	10
CCGAACGTCC	TGGGA	GGAGGATCC	Immunoglobulin kappa light chain	31
CAGCGTTGAA	GGTGA	AGCGGGAGAT	T-cell beta chain	9
AGCCTGTCCC	GGGAC	AAGCTGCTGG	T-cell beta chain	9
AAGCAAACCA	AGAGA	TAGTAACTAT	Immunoglobulin heavy chain	48
CAGTTCCCTG	ATTAGT	GCTGCTGGCG	T-cell beta chain	9
CTGAGCCCCT	CCAGGG	ACTGCTGGCA	T-cell beta chain	9
CCCCCTGTCC	GGTAGG	TGCTGCTGGC	T-cell beta chain	9
GGCAGGGGGC	TAGGGG	ACACTGAAGC	T-cell beta chain	9
TGCCAGCAGT	TCCAAAA	ACAGGGGATG	T-cell beta chain	9
CCCCAGTAAG	GGGGCTT	ATAGTAGGCA	Immunoglobulin heavy chain	10
AAA	CCAGAGA	TAGTAACTAT	Immunoglobulin heavy chain	47
AGTCCCCTGT	AGGACGG	GGCTGCTGGC	T-cell beta chain	40
ACAGACAGGG	TCTTGGG	TCAGCCCCAG	T-cell beta chain	9
TCGACTAGCG	ATCCAAAA	AATGAGCAGT	T-cell beta chain	9
TAGTCAAAGT	TAGTACGA	AGTTACCATA	Immunoglobulin heavy chain	16
ATTTCCAGAA	TCAGGGAC	GTTACATCGT	T-cell beta chain	40
AGTAAGCAAA	CCAAGAGA	TAGTAACTAT	Immunoglobulin heavy chain	46
GCAGTCCACA	CCTGGATGG	TTCACCCCTC	T-cell beta chain	9
GCGTAGGACA	ATCGATGGG	CAATGAGCAG	T-cell beta chain	9
CCGGGACAGG	CTCGGGTGGG	GCAATCAGCC	T-cell beta chain	9
AGCAGCCAAG	CCAAGGGTGG	CAGGGGGCTA	T-cell beta chain	9
GTCCATAGCA	GGGACCTGCGC	CAGTTACCAT	Immunoglobulin heavy chain	16
GCCGCGGAGG	TGGCGTCAATG	CCTACAATGA	T-cell beta chain	9
GGGGACAGGG	TGGGGGTTTGA	CTTTCTTTGG	T-cell beta chain	9
TGTGCCTACT	GGAGGGGAGTAC	GGGCCAAGGG	Immunoglobulin heavy chain	46, 48
AAGAGATACT	GGAGGGGAGTAC	GGGCCAAACA	Immunoglobulin heavy chain	47
GTGTGAGAGA	TAGGAAGTATGGT	AACTGGCGCA	Immunoglobulin heavy chain	16

[a] For all junctions, the strand containing the filler DNA with the highest fraction of purines is displayed. Where possible, 10 nucleotides on each side of the added nucleotides at the junction are shown. In all cases, the indicated length of filler DNA is a minimum estimate. If one or more nucleotides at the ends of the filler DNA segment matched the parental DNA sequence, they would be counted as parental DNA. Filler DNA is operationally defined as extra nucleotides that cannot be accounted for by the parental DNA sequence.

TABLE 3. Filler DNA at genetic rearrangement junctions in mammalian cells: reciprocal translocations[a]

5′ Sequence	Filler DNA	3′ Sequence	Type[b]	Reference
TCTCTCGCTG	A	AGCTGAGCTG	Plasmacytoma (switch)	14
CTGATGTTGG	A	GAGCTGAGGT	Plasmacytoma (switch)	38
TGGTGGGACT	G	TTCGGCAGCT	Burkitt's lymphoma (VDJ)	37
GGTTCCCTGG	G	GCACCACTGC	Follicular lymphoma (VDJ)	57
GCCCTCCTTC	GA	TACTGGGGCC	Follicular lymphoma (VDJ)	8
CTGCCCTCCT	CGA	TACGGTATGG	Follicular lymphoma (VDJ)	57
TTTTACAGAT	GAG	GCCGCGCGCC	Muscular dystrophy (nonimmune)	4
GCTGGTGTGA	AAG	TACGGAGCCT	Plasmacytoma (switch)	14
ATGGTAAGCC	GGTT	CCCCCCCCCC	Plasmacytoma (switch)	14
ATCTTCCAGG	AAGC	CTGGTGTCCT	Burkitt's lymphoma (VDJ)	18
CATAATGAGC	ATAGG	ACCCCAGCCT	Plasmacytoma (switch)	14
GCAAACCTGA	CCAAAT	AAAATTACAG	Burkitt's lymphoma (VDJ)	18
AATTATTTTT	GAAAGT	TAACAATGCC	(VDJ)	13
CAGGGGTCGA	GGAGGGG	GCAAAACACA	Pre-B-cell leukemia (VDJ)	57
TGGCGTGAAC	TACCAGAC	TTGACTACTG	CLL (VDJ)	58
AAGCCAGACC	CCCGAGTGAAG	ACTACTACTA	Follicular lymphoma (VDJ)	57
AGTAGTAGTA	AAGTCGTTATG	AGGTTTCCTG	Follicular lymphoma (VDJ)	3
AATGCAGTGG	GTGCGTGGTTGATGGGGA	TTGACTACTG	Follicular lymphoma (VDJ)	57
CTGCCATTGG	TGTTGGAGGGAACCCGCATC	TGACTACTGG	CLL (VDJ)	58
CCAGACGTCC	CGGGGGGGCCCTAATACTCTAG	GTCCTTGACT	ALL (VDJ)	17
AGCTAATTTT	GTGTATGTTTAGTAGAGACGAGGTTTC	ACAGAAGCTG	CML (nonimmune)	22
GCATGTTTCC	ATTCAGGGATGGCCTACTGAGGGTGCGTGGCATCTAGA	TTCTGACTTA	Plasmacytoma (switch)	41

[a] For all junctions, the strand containing the filler DNA with the highest fraction of purines is displayed. Where possible, 10 nucleotides on each side of the added nucleotides at the junction are shown. In all cases, the indicated length of filler DNA is a minimum estimate. If one or more nucleotides at the ends of the filler DNA segment matched the parental DNA sequence, they would be counted as parental DNA. Filler DNA is operationally defined as extra nucleotides that cannot be accounted for by the parental DNA sequence.

[b] CLL, Chronic lymphoblastic leukemia; ALL, acute lymphoblastic leukemia; CML, Chronic myeloid leukemia.

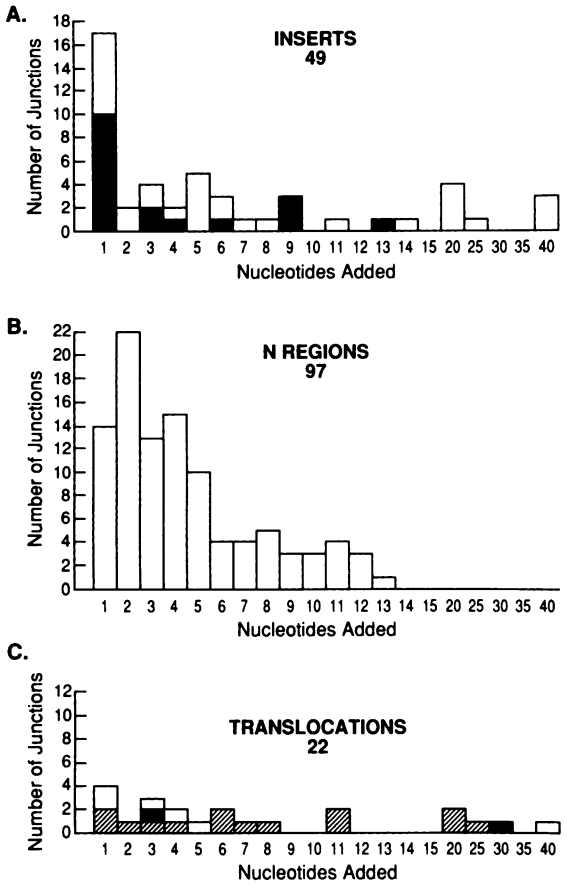

FIG. 2. Length distributions of filler DNAs at nonimmune rearrangements (A), at immune system rearrangements (B), and at reciprocal translocations (C). Symbols: (A) ■, junctions shown in Fig. 1; □, junctions from Table 1; (C) ▨, junctions that involve a break near the recombination signals at V, D, and J immunoglobulin gene segments; □, junctions that involve a break at heavy-chain switch signals; ■, junctions that do not involve a break at an immunoglobulin locus.

tides to the ends of DNA molecules is supported by several observations. The length distribution of N regions is consistent with the notion that nucleotides are added one by one to ends. (Since most sequenced immune system rearrangements arose from productive events leading to expressed genes, however, the distribution may be altered somewhat by selection within the immune system.) In addition, the high G+C content of N regions is consistent with the marked in vitro preference of TdT for addition of G nucleotides (3; Table 4). Finally, the presence of N regions in immunoglobulin gene rearrangements correlates with the expression of TdT (10, 31). Since TdT has not been detected in nonlymphoid cells (7), it seems unlikely to account for the extra nucleotides at nonimmune system junctions.

TABLE 4. Characteristics of filler DNA at mammalian rearrangement junctions

Junctions	%G+C	% with length (nucleotides) of:	
		>13	1
Immune system rearrangements	57	0	13
Nonimmune system rearrangements	38	18	35
Reciprocal translocations	53	23	18

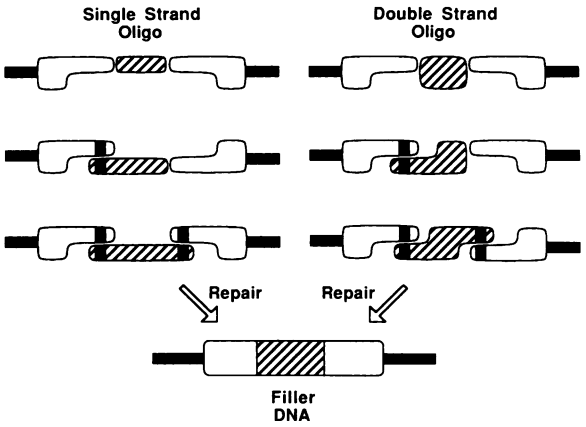

FIG. 3. Oligonucleotide capture during the joining of DNA ends. Symbols: ▨, oligonucleotides; ■, pairing of complementary nucleotides between the oligonucleotides and the DNA ends. Repair of the intermediate structures to generate a junction with filler DNA presumably involves filling in gaps and ligation of nicks.

Comparison with junctions at reciprocal chromosome translocations. Extra nucleotides were found at 22 of the reciprocal translocations associated with lymphoid neoplasms (Table 3 and Fig. 2C). Of these junctions, 20 were located near immune system rearrangement signals; 14 evidently resulted from errors in VDJ joining, and 6 resulted from errors in class switching. Two rearrangements did not involve a break at an immune locus (Fig. 2C). The length distribution of these filler DNAs resembled that of the inserts (Fig. 2A and C) but with a less prominent class of one-nucleotide additions. Although the overall G+C content was 53% (Table 4), the extra nucleotides at translocation junctions involving VDJ recombination signals had a high G+C content (60%) similar to that of N regions (57%). By contrast, the extra nucleotides at translocation junctions that did not involve VDJ recombination signals had a low G+C content (43%) similar to that of inserts (38%). Filler DNA at these rearrangements, therefore, may arise by a combination of the processes that lead to N regions and inserts.

Potential sources of inserts in nonlymphoid cells. In principle, inserts in nonlymphoid cells could be created by an enzyme with an activity similar to that of TdT; however, no such enzymatic activity has been reported. In addition, the distribution of insert lengths suggests that different mechanisms might account for the flat part of the distribution and for the relatively high frequency of one-nucleotide inserts.

Examples of mechanisms that could account for the flat portion of the distribution are illustrated in Fig. 3, which depicts the addition of preformed blocks of nucleotides to broken ends. The oligonucleotides could be single stranded or double stranded and could be joined to ends by direct ligation or after pairing through short-sequence homologies. Direct ligation of single strands and pairing through short-sequence homologies have both been implicated in the efficient joining of DNA ends in transfected DNA molecules (50). The finding that mammalian cells contain a pool of oligonucleotides with a distribution of lengths similar to that of inserts also is consistent with addition of oligonucleotides as a mechanism of insert formation (43). The cellular pool of oligonucleotides is mainly RNA rather than DNA, raising the possibility that inserts may arise from addition of RNA fragments. Addition of oligonucleotides is supported by the

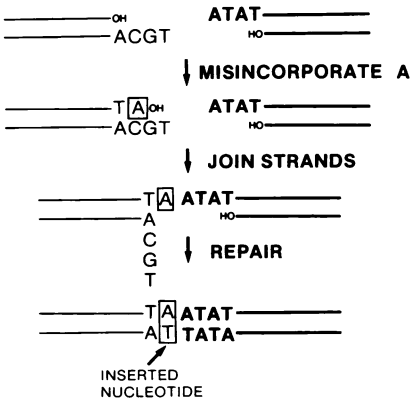

FIG. 4. Misincorporation of A residues during repair of a 5' single-stranded extension. Misincorporation of A residues, followed by strand joining and repair, can generate a junction with a one-nucleotide insertion.

observation that half of the inserts longer than 10 base pairs (Table 1) were composed of blocks of nucleotides that seemed to be derived from other DNA molecules in the cell (2, 34, 62) such as pBR322 DNA, which was present in one series of experiments (2).

Another mechanism that may contribute to the presence of extra nucleotides is suggested by examination of the junctions shown in Fig. 1. Four of the eight inserts with more than a single nucleotide contained sequences that were also present near one of the ends of the parental DNA molecule, forming direct repeats (Fig. 1, junctions 12, 13, 15, and 18). The lengths of these repeated sequences, which ranged from 6 to 18 nucleotides, suggest strongly that the inserted nucleotides are somehow copied by use of nearby sequences in the parental DNA as a template. We have previously described a potential mechanism for the formation of such inserts; it involves breathing of the duplex ends, followed by slipped mispairing and repair synthesis (49). Both slipped mispairing and direct addition of cellular oligonucleotides would also account for the base composition of nonimmune system inserts, which is very similar to the overall base composition of mammalian DNA.

The striking peak of one-nucleotide inserts may be accounted for by a different mechanism. One possibility involves misincorporation rather than actual nucleotide addition. In 38% of the nonimmune system junctions with inserts, only one base pair was added; in 14 of these 17 junctions (82%), the added base pair was an A · T. These single-nucleotide insertions could arise if DNA polymerase incorporates noncomplementary nucleotides during polymerization reactions near DNA ends, thereby creating base changes that appear to be one-nucleotide additions. Misincorporation could occur in the filling in of an end or during repair synthesis after slipped mispairing. A simple illustration (without the complications of nuclease action) of how misincorporation at an end might lead to an added nucleotide is shown in Fig. 4. Although the frequency of misincorporation at ends is unknown, the correct, complete filling of ends occurs at low frequency in DNA molecules transfected into monkey cells (49, 50). Experiments using purified enzymes indicate that DNA polymerase beta normally is relatively error prone, making one mistake per 1,500 bases (28). Even so, the error frequency would need to be considerably enhanced to account for the overall frequency of one-nucleotide inserts. It is possible that DNA polymerase has lowered fidelity near ends or, stated alternatively, that ends represent an ambiguous coding situation for DNA polymerase. The high frequency of A · T inserts is consistent with the preferential addition of A residues when DNA polymerases are faced with ambiguity (20, 44).

The relatively high frequency of single-nucleotide additions in inserts as compared with N regions and chromosome translocations (Fig. 2) might also reflect the extremely high frequency of point mutations that occurs in DNA molecules transfected into mammalian cells (6, 45). Indeed, in transfected DNA, the observed mutations almost always occur at G · C base pairs and result in transition to A · T or transversion to T · A (19). Mutation during transfection may explain why the inserts at nonimmune system junctions, which were derived largely from transfection experiments, contained a much higher frequency of single-nucleotide additions than did N regions and chromosome translocations. Yet mutation cannot be the whole story, since the first four junctions in Fig. 1 retained all of the parental nucleotides; none were mutated.

Conclusions. DNA polymerases and DNA joining activities are likely to be present in all cells and are presumably active during immunoglobulin and T-cell receptor gene rearrangements. Thus, alternative mechanisms, such as those shown in Fig. 3 and 4 and others that were discussed, may also generate N regions in lymphoid cells. Mechanisms that do not depend on TdT have been suggested to account for N regions that arise in cells that do not contain detectable TdT (31). Also, the relatively high A+T content (57%) of one-nucleotide N regions, despite the marked preference of TdT for insertion of G residues, might indicate some contribution by the events that give rise to A+T-rich one-nucleotide inserts in nonlymphoid cells. These alternative mechanisms of formation of filler DNA also presumably apply to translocations. Thus, the mere presence of filler DNA at a translocation junction does not necessarily implicate the immune system recombination machinery (4, 22).

The variety of ways to generate filler DNA underscores the versatility of end-joining processes in mammalian cells. End joining may represent the normal pathway for repairing double-strand breaks in mammalian cells (59). Since somatic cells are not the germ line, small nucleotide losses or gains at sites of repair represent no cost to future generations, nor do they represent a cost to the cell except in the unlikely event that they occur at a critical point in an expressed gene. By contrast, bacteria and yeasts repair double-strand breaks predominantly by homologous recombination (55, 56). Perhaps the sloppiness that accompanies random end joining is intolerable in an organism that is its own germ line.

ACKNOWLEDGMENTS

We thank De Dieu and Kathleen Marburger for expert technical assistance. We thank Fred Alt, Lee Hood, Pat Concannon, and Hamisch Young for providing us with information in advance of publication. We thank David Baltimore, Susan Berget, and Fred Bollum for critical comments on earlier versions of the manuscript and the Molecular Biology Group in the Department of Biochemistry at Baylor College of Medicine for helpful criticism.

This work was supported by grants from the Public Health Service (CA15743, GM33405, and GM38219, from the National Institutes of Health) and the Robert A. Welch Foundation (Q-977). D.B.R. was supported by the Edward J. and Josephine Hudson Scholars Fund during his training in the Medical Scientist Training Program.

LITERATURE CITED

1. **Alt, F. W., and D. Baltimore.** 1982. Joining of immunoglobulin heavy chain gene segments: implications from a chromosome

with evidence of three D-J_H fusions. Proc. Natl. Acad. Sci. USA **79**:4118–4122.
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. **Bakhashi, A., J. J. Wright, W. Graninger, M. Seto, J. Owens, J. Cossman, J. P. Jensen, P. Goldman, and S. J. Korsmeyer.** 1987. Mechanism of the t(14:18) chromosomal translocation: structural analysis of both derivative 14 and 18 reciprocal partners. Proc. Natl. Acad. Sci. USA **84**:2396–2400.
4. **Bodrug, S. E., P. N. Ray, I. L. Gonzalez, R. D. Schmickel, J. E. Sylvester, and R. G. Worton.** 1987. Molecular analysis of a constitutional X-autosome translocation in a female with muscular dystrophy. Science **237**:1620–1624.
5. **Botchan, M., J. Stringer, T. Mitchison, and J. Sambrook.** 1980. Integration and excision of SV40 DNA from the chromosome of a transformed cell. Cell **20**:143–152.
6. **Calos, M. P., J. S. Lebkowski, and M. R. Botchan.** 1983. High mutation frequency in DNA transfected into mammalian cells. Proc. Natl. Acad. Sci. USA **80**:3015–3019.
7. **Chang, L. M. S.** 1971. Development of terminal deoxynucleotidyl transferase activity in embryonic calf thymus gland. Biochem. Biophy. Res. Commun. **44**:124–131.
8. **Cleary, M. L., and J. Sklar.** 1985. Nucleotide sequence of a t(14;18) chromosomal breakpoint in follicular lymphoma and demonstration of a breakpoint-cluster region near a transcriptionally active locus on chromosome 18. Proc. Natl. Acad. Sci. USA **82**:7439–7443.
9. **Concannon, P., L. A. Pickering, P. Kung, and L. Hood.** 1986. Diversity and structure of human T-cell receptor V_{beta} genes. Proc. Natl. Acad. Sci. USA **83**:6598–6602.
10. **Desiderio, S. V., G. D. Yancopoulos, M. Paskind, E. Thomas, M. A. Boss, N. Landau, F. W. Alt, and D. Baltimore.** 1984. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. Nature (London) **311**:752–755.
11. **DiMaio, D., and D. Nathans.** 1980. Cold-sensitive regulatory mutants of simian virus 40. J. Mol. Biol. **140**:129–142.
12. **Doerfler, W., W. Gahlmann, S. Stabel, R. Deuring, U. Lichtenberg, M. Schulz, D. Eick, and R. Leisten.** 1983. On the mechanism of recombination between adenoviral and cellular DNAs: the structure of junction sites. Curr. Top. Microbiol. Immunol. **109**:193–228.
13. **Finger, L. R., R. C. Harvey, R. C. A. Moore, L. C. Showe, and C. M. Croce.** 1986. A common mechanism of chromosomal translocation in T- and B-cell neoplasia. Science **234**:982–985.
14. **Gerondakis, S., S. Cory, and J. Adams.** 1984. Translocation of the myc cellular oncogene to the immunoglobulin heavy chain locus in murine plasmacytomas is an imprecise reciprocal exchange. Cell **36**:973–982.

14a. **Gibbs, R. A., P.-N. Nguyen, L. J. McBride, S. M. Koepf, and C. T. Caskey.** 1989. Identification of mutations leading to the Lesch-Nyhan syndrome by automated direct DNA sequencing of in vitro amplified cDNA. Proc. Natl. Acad. Sci. USA **86**:1919–1923.

15. **Hagiya, M., D. D. Davis, L. D. Shultz, and H. Sakano.** 1986. Non-germline elements (NGE) are present in the T cell receptor beta-chain genes isolated from the mutant mouse motheaten (me/me). J. Immunol. **136**:2697–2700.
16. **Hagiya, M., D. D. Davis, T. Takahashi, K. Okuda, W. C. Raschke, and H. Sakano.** 1986. Two types of immunoglobulin-negative Abelson murine leukemia virus-transformed cells: implications for B-lymphocyte differentiation. Proc. Natl. Acad. Sci. USA **83**:145–149.
17. **Haluska, F. G., S. Finver, Y. Tsujimoto, and C. M. Croce.** 1986. The t(8:14) chromosomal translocation occurring in B-cell malignancies results from mistakes in V-D-J joining. Nature (London) **324**:158–161.
18. **Haluska, F. G., Y. Tsujimoto, and C. M. Croce.** 1987. The t(8;14) chromosome translocation of the Burkitt lymphoma cell line Daudi occurred during immunoglobulin gene rearrangement and involved the heavy chain diversity region. Proc. Natl. Acad. Sci. USA **84**:6835–6839.
19. **Hauser, J., A. S. Levine, and K. Dixon.** 1987. Unique pattern of point mutations arising after gene transfer into mammalian cells. EMBO J. **6**:63–67.
20. **Hauser, J., M. M. Seidman, K. Sidur, and K. Dixon.** 1986. Sequence specificity of point mutations induced during passage of a UV-irradiated shuttle vector plasmid in monkey cells. Mol. Cell. Biol. **6**:277–285.
21. **Hayday, A. C., H. Saito, S. D. Gillies, D. M. Kranz, G. Tanigawa, H. N. Eisen, and S. Tonegawa.** 1985. Structure, organization, and somatic rearrangement of T cell gamma genes. Cell **40**:259–269.
22. **Heisterkamp, N., K. Stam, J. Groffen, A. de Klein, and G. Grosveld.** 1985. Structural organization of the bcr gene and its role in the Ph' translocation. Nature (London) **315**:758–761.
23. **Hope, T. J., R. J. Aguilera, M. E. Minie, and H. Sakano.** 1986. Endonucleolytic activity that cleaves immunoglobulin recombination sequences. Science **231**:1141–1145.
24. **Jennings, M. W., R. W. Jones, W. G. Wood, and D. J. Weatherall.** 1985. Analysis of an inversion within the human beta globin cluster. Nucleic Acids Res. **13**:2897–2906.
25. **Karjalainen, K., and C. Coleclough.** 1985. An unusual type of V-J joining diversifies the primary repertoire of mouse lambda light chains. Nature (London) **314**:544–546.
26. **Kavaler, J., M. M. Davis, and Y.-H. Chien.** 1984. Localization of a T-cell receptor diversity-region element. Nature (London) **310**:421–423.
27. **Keinfield, R., R. R. Hardy, D. Tarlington, J. Dangl, L. A. Herzenberg, and M. Weigert.** 1986. Recombination between an expressed immunoglobulin heavy-chain gene and a germline variable gene segment in a Ly 1^+ B-cell lymphoma. Nature (London) **322**:843–846.
28. **Kunkel, T. A., and P. S. Alexander.** 1986. The base substitution fidelity of eucaryotic DNA polymerase: mispairing frequencies, site preferences, and base substitution by dislocation. J. Biol. Chem. **261**:160–166.
29. **Kurosawa, Y., and S. Tonegawa.** 1982. Organization, structure, and assembly of immunoglobulin heavy chain diversity segments. J. Exp. Med. **155**:201–218.
30. **Kurosawa, Y., H. von Boehmer, W. Haas, H. Sakano, A. Trauneker, and S. Tonegawa.** 1981. Identification of D segments of immunoglobulin heavy-chain genes and their rearrangement in T lymphocytes. Nature (London) **290**:565–570.
31. **Landau, N. R., D. G. Schatz, M. Rosa, and D. Baltimore.** 1987. Increased frequency of N-region insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase retroviral expression vector. Mol. Cell. Biol. **7**:3237–3243.
32. **Lewis, S., A. Gifford, and D. Baltimore.** 1985. DNA elements are asymmetrically joined during the site-specific recombination of kappa immunoglobulin genes. Science **228**:677–685.
33. **Ling, L. E., M. M. Manos, and Y. Gluzman.** 1982. Sequences of the junction in adenovirus 2-SV40 hybrids: example of illegitimate recombination. Nucleic Acids Res. **10**:8099–8112.
34. **Mager, D. L., P. S. Henthorn, and O. Smithies.** 1985. A Chinese $G_{gamma} + A_{(gamma\ delta\ beta)}^{0}$ thalassemia deletion: comparison to other deletions in the human beta-globin gene cluster and sequence analysis of the breakpoints. Nucleic Acids Res. **13**: 6559–6575.
35. **Munz, P. L., and C. S. H. Young.** 1987. The creation of adenovirus genomes with viable, stable, internal redundancies centered about the E2b region. Virology **158**:52–60.
36. **Nalbantoglu, J., G. Phear, and M. Meuth.** 1987. DNA sequence analysis of spontaneous mutations at the *aprt* locus of hamster cells. Mol. Cell. Biol. **7**:1445–1449.
37. **Neri, A., F. Barriga, D. M. Knowles, I. T. Magrath, and R. Dalla-Favera.** 1988. Different regions of the immunoglobulin heavy-chain locus are involved in chromosomal translocations in distinct pathogenetic forms of Burkitt lymphoma. Proc. Natl. Acad. Sci. USA **85**:2748–2752.
38. **Neuberger, M. S., and F. Calabi.** 1983. Reciprocal chromosome translocation between c-myc and immunoglobulin gamma 2b genes. Nature (London) **305**:240–243.
39. **Nicholls, R. D., N. Fischel-Ghodsian, and D. R. Higgs.** 1987. Recombination at the human alpha-globin gene cluster: se-

quence features and topological constraints. Cell **49**:369–378.
40. **Okazaki, K., D. D. Davis, and H. Sakano.** 1987. T cell receptor beta gene sequences in circular DNA of thymocyte nuclei: direct evidence for intramolecular DNA deletion in V-D-J joining. Cell **49**:477–485.
41. **Piccoli, S. P., P. G. Caimi, and M. D. Cole.** 1984. A conserved sequence at c-myc oncogene chromosomal translocation breakpoints in plasmacytomas. Nature (London) **310**:327–330.
42. **Pipas, J. M., K. W. Peden, and D. Nathans.** 1983. Mutational analysis of simian virus 40 T antigen: isolation and characterization of mutants with deletions in the T antigen gene. Mol. Cell. Biol. **3**:203–213.
43. **Plesner, P., J. Goodchild, H. M. Kalckar, and P. C. Zamecnik.** 1987. Oligonucleotides with rapid turnover of the phosphate groups occur endogenously in eukaryotic cells. Proc. Natl. Acad. Sci. USA **84**:1936–1939.
44. **Rabkin, S. D., P. D. Moore, and B. S. Strauss.** 1983. In vitro bypass of UV-induced lesions by Escherichia coli DNA polymerase I: specificity of nucleotide incorporations. Proc. Natl. Acad. Sci. USA **80**:1541–1545.
45. **Razzaque, A., S. Chakrabarti, S. Joffe, and M. Seidman.** 1984. Mutagenesis of a shuttle vector plasmid in mammalian cells. Mol. Cell. Biol. **4**:435–441.
46. **Reth, M. G., and F. W. Alt.** 1984. Novel immunoglobulin heavy chains are produced from DJ_H gene segment rearrangements in lymphoid cells. Nature (London) **312**:418–423.
47. **Reth, M., P. Gehrmann, E. Petrac, and P. Wiese.** 1986. A novel V_H to V_HDJ_H joining mechanism in heavy-chain-negative (null) pre-B cells results in heavy-chain production. Nature (London) **322**:840–842.
48. **Reth, M. G., S. Jackson, and F. W. Alt.** 1986. V_HDJ_H formation and DJ_H replacement during pre-B differentiation: nonrandom usage of gene segments. EMBO J. **5**:2131–2138.
49. **Roth, D. B., T. N. Porter, and J. H. Wilson.** 1985. Mechanisms of nonhomologous recombination in mammalian cells. Mol. Cell. Biol. **5**:2599–2607.
50. **Roth, D. B., and J. H. Wilson.** 1986. Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. Mol. Cell. Biol. **6**:4295–4304.
51. **Sakano, H., Y. Kurosawa, M. Weigert, and S. Tonegawa.** 1981. Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes. Nature (London) **290**:562–565.
52. **Sheflin, L., A. Celeste, and M. Woodworth-Gutai.** 1983. Recombination in simian virus 40-infected cells. J. Biol. Chem. **258**: 14315–14321.
53. **Siu, G., M. Kronenberg, E. Strauss, R. Haars, T. W. Mak, and L. Hood.** 1984. The structure, rearrangement and expression of D_{beta} gene segments of the murine T-cell antigen receptor. Nature (London) **311**:344–350.
54. **Spritz, R. A., and S. H. Orkin.** 1982. Duplication followed by deletion accounts for the structure of an Indian deletion $beta^0$ thalassemia gene. Nucleic Acids Res. **10**:8025–8029.
55. **Stahl, F. W.** 1986. Roles of double-strand breaks in generalized genetic recombination. Prog. Nucleic Acid Res. Mol. Biol. **33**:169–194.
56. **Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein, and F. S. Stahl.** 1983. The double-strand-break repair model for recombination. Cell **33**:25–35.
57. **Tsujimoto, Y., J. Gorham, J. Cossman, E. Jaffe, and C. M. Croce.** 1985. The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. Science **229**:1390–1393.
58. **Tsujimoto, Y., E. Jaffe, J. Cossman, J. Gorham, P. C. Nowell, and C. M. Croce.** 1985. Clustering of breakpoints on chromosome 11 in human B-cell neoplasms with the t(11;14) chromosome translocation. Nature (London) **315**:340–343.
59. **Wake, C. T., F. Vernaleone, and J. H. Wilson.** 1985. Topological requirements for homologous recombination among DNA molecules transfected into mammalian cells. Mol. Cell. Biol. **5**: 2080–2089.
60. **Weiss, U., and J. H. Wilson.** 1988. Heteroduplex-induced mutagenesis in mammalian cells. Nucleic Acids Res. **16**:2313–2321.
61. **Wilkie, T. M., and R. D. Palmiter.** 1987. Analysis of the integrant in MyK-103 transgenic mice in which males fail to transmit the integrant. Mol. Cell. Biol. **7**:1646–1655.
62. **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.
63. **Wilson, J. H., P. B. Berget, and J. M. Pipas.** 1982. Somatic cells efficiently join unrelated DNA segments end-to-end. Mol. Cell. Biol. **2**:1258–1269.
64. **Winoto, A., S. Mjolsness, and L. Hood.** 1985. Genomic organization of the genes encoding mouse T-cell receptor alpha-chain. Nature (London) **316**:832–836.
65. **Woodworth-Gutai, M.** 1981. Recombination in SV40-infected cells: viral DNA sequences at sites of circularization of transfecting linear DNA. Virology **109**:353–365.
66. **Zagursky, R. J., K. Baumeister, N. Lomax, and M. L. Berman.** 1985. Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. Gene Anal. Techn. **2**: 89–94.