Homogenization of Tandemly Repeated Nucleotide Sequences by Distance-Dependent Nucleotide Sequence Conversion

J. Dvořák, D. Jue and M. Lassner

Department of Agronomy and Range Science, University of California, Davis, California 95616

Manuscript received December 11, 1986
Revised copy accepted March 27, 1987

ABSTRACT

Previous work revealed that recurrent mutations (=mutation occurring more than once) in the tandemly repeated arrays present in nontranscribed spacers (NTS) of ribosomal RNA genes (rDNA) are clustered, *i.e.*, they most frequently occur in repeats with adjacent or alternate distribution. A possible explanation is that the likelihood of heteroduplex formation, a prerequisite of gene conversion, decreases with the distance between repeats. To test this possibility, evolution of an array of 11 initially homogeneous repeats was computer simulated using three models, two assuming that the likelihood of heteroduplex formation decreases with increasing distance between the repeats and one assuming that it is constant. Patterns of mutation distribution obtained in computer simulations were compared with the distribution of mutations found in the repeated arrays in the NTS of seven rDNA clones. The patterns of mutations generated by the models assuming that the likelihood of heteroduplex formation decreases as distance between the repeats increases agreed with the patterns observed in rDNA; the patterns generated by the model assuming that the likelihood is independent of distance between repeats disagreed with the patterns observed in the rDNA clones. The topology of the heteroduplex formed between DNA in adjacent repeats predicts that the most frequently occurring conversions in the NTS repeated arrays will be shorter than the length of the repeat. The topology of the heteroduplex also predicts that if the heteroduplex leads to crossing over a circular repeat is excised. It is speculated that the circle can transpose or can be amplified via rolling circle replication and subsequently transpose. It is also shown that homogenization of the NTS repeated arrays proceeds at different rates in different species.

SEVERAL mechanisms have been proposed to explain how repeated nucleotide sequences in eukaryotic genomes maintain their homogeneity. BRITTEN and KOHNE (1968) speculated that families of repeated sequences are created by saltatory replication of a sequence. Others proposed a gradual evolutionary turnover of sequences in the genome by unequal crossing over (TARTOF 1975; SMITH 1976). Because many families of repeated sequences are maintained relatively homogeneous, even though they are on two or more chromosomes, some other mechanism in addition to crossing over must be involved in the homogenization of repeated nucleotide sequences. A strong candidate is nucleotide sequence conversion which involves a temporary formation of a heteroduplex between two related sequences. The DNA heteroduplex creates a potential for a transfer of sequence between the parental sequences by repair of mismatched nucleotides. The transfer of sequences between paralogous loci without recombination of flanking sequences has been directly observed in a number of gene systems (JACKSON and FINK 1981; LIEBHABER *et al.* 1981; KLEIN and PETES 1981; AMSTUTZ *et al.* 1985).

Because rare mutations have a low likelihood of surviving in a repeated array if the array is subjected to repeated conversions, homogenization by gene conversion is a conservative evolutionary process (BIRKY and SKAVARIL 1976). The conservative tendency of gene conversion is the strongest if every member of a family has equal potential to form heteroduplex with any other member. Any restriction on the selection of which parental sequences are involved in heteroduplex formation decreases the "effective size" of the gene family. The understanding of the rules by which members of a repeated sequence family are selected for conversion are, therefore, paramount for understanding both the turnover of sequences in the genome and the evolution of new families of repeated genes.

We selected the gene family which codes for the 5.8 S, 18 S, and 26 S (28 S) ribosomal RNA (henceforth rDNA) to study this aspect of the evolution of repeated nucleotide sequences. Several attributes of this gene family make it well suited for this purpose. The gene units are in tandem arrays; since they are found at only one or several loci, they are subject to convenient genetic analysis. Within each gene unit, the coding region is preceded by an "external spacer" which contains a tandem array of short repeated se-

quences, between 100 and 300 bp in length. These arrays are often not transcribed, thus the spacers are often referred to as nontranscribed spacers (NTS).

This gene family is subjected to a hierarchy of homogenizations. Homogenization occurs among the repeats in the repeated array of each NTS, among entire gene units within a specific locus, and among gene units of paralogous loci. The distribution of mutations among repeats within a NTS and between different NTS was used in attempts to find evidence of past homogenization events (KUEHN and ARNHEIM 1983; LASSNER and DVOŘÁK 1986). Mutations that were in two or more repeats (recurrent mutations) showed that the repeated arrays of NTS evolve by a complex pattern of overlapping homogenization events, usually shorter than the length of a repeat (LASSNER and DVOŘÁK 1986). An intriguing finding was that mutations were nonrandomly distributed in the repeated arrays (LASSNER and DVOŘÁK 1986). Recurrent mutations were most frequently present in either neighboring repeats (adjacent distribution) or in repeats separated by one repeat (alternate distribution). It was concluded that the homogenization of repeats in an NTS array occurs chiefly by gene conversion (LASSNER and DVOŘÁK 1986). This nonrandom distribution of recurrent mutations can be explained by assuming that the likelihood of two repeats in a tandem array forming a DNA heteroduplex decreases with their distance from each other. This hypothesis is tested in this paper by following the fate of mutations during computer simulated evolution of initially completely homogeneous arrays of repeats. DNA heteroduplex was allowed either to occur with equal probability among all members of the array or with decreasing probability as the distance between the repeats increased. The distribution of mutations in the computer generated repeated arrays was compared with the distribution of mutations in the rDNA spacers of wheat, rye, maize and mouse.

MATERIALS AND METHODS

Sequenced rDNA spacers: Repeated arrays from NTS in two rDNA clones, pTa71 and pTa250, representing both NTS length variants that exist at the *Nor-B2* locus of bread (*Triticum aestivum* L.) cv. Chinese Spring were sequenced (LASSNER and DVOŘÁK 1986). A repeated array of the spacer in the rDNA variant that is at the *Nor-D3* locus of wheat cv. Cheyenne (pTaNorD3) has also been sequenced (LASSNER, ANDERSON and DVOŘÁK 1987). Repeated arrays of pTa71 and pTaNorD3 were sequenced entirely, but only seven of 11 repeats of pTa250 were sequenced. Determination of the nucleotide sequence has also been reported for rDNA spacer arrays from one rye rDNA clone (APPELS, MORAN and GUSTAFSON 1986), two maize clones (MCMULLEN *et al.* 1986; TOLOCZYKI and FELIX 1986) and one mouse rDNA clone (KUEHN and ARNHEIM 1983). The characteristics of these arrays are listed in Table 1.

To compare the distribution of mutations, several statistics were calculated for each sequenced NTS repeated array. The average number of mutations per polymorphic nucleotide position was calculated by summing the number of mutations of the same kind at all polymorphic nucleotide positions and dividing the sum by the number of polymorphic nucleotide positions in the sequence. If a nucleotide position was polymorphic for mutations of two or more kinds, different mutations were treated as if they occurred at different positions. The distribution of mutations occurring more than once in a repeated array (recurrent mutations) was characterized in the following manner. Percentage of recurrent mutations occurring in adjacent repeats was calculated by summing up all recurrent mutations that occurred in adjacent repeats, dividing the sum by the number of nucleotide positions in the sequence with recurrent mutations and multiplying by 100. Percentages of recurrent mutations with alternate distributions, those in repeats separated by two repeats and those in repeats separated by more than two repeats were each calculated in a similar manner. Additionally, the percentage of positions in the repeated sequence in which all recurrent mutations were in adjacent or alternate repeats was calculated. If there were two different recurrent mutations at a position they were treated separately as if they were at different positions.

Computer simulated evolution of a repeated array: The evolution of an initially homogeneous repeated array was simulated by computer. An array of 11 tandem repeats was used because the actual numbers of repeats in rDNA spacers vary around this number. A mutation was placed in a specific repeat. The first (source) repeat involved in conversion was selected randomly from the 11 repeats. Three different models were used to select a second (target) repeat for DNA heteroduplex formation: (1) The target was selected randomly (henceforth, random model). (2) The target was selected from the remaining 10 repeats with exponentially decreasing probability given by the equation $p = (1/2)^x$, where p is the probability that a repeat is going to be selected for heteroduplex formation with the source and x is the distance of the target repeat from the source repeat. For an adjacent repeat $x = 1$, for alternate $x = 2$, etc. This model will be referred to as model 1. (3) Another nonrandom model of heteroduplex formation was given by an equation of

$$p = \frac{1}{x} \times \frac{1}{\sum_{i=1}^{n} \frac{1}{x_i}},$$

where p is the probability that a repeat is going to be selected as the target, n is the number of repeats in the array (11), and x is the distance of the target repeat from the source repeat. For an adjacent repeat $x = 1$, for alternate $x = 2$, etc. The second term of the probability function is a factor to normalize the sum of probabilities for all target repeats to equal one. The probability of heteroduplex formation in this model also decreases with the distance between repeats in the array but less precipitously than in model 1. This model will henceforth be referred to as model 2. Conversion was performed without a directional bias, *i.e.*, the likelihood of a source being converted by the target was the same as the reverse. Conversions were repeated until either the mutation was eliminated from the array or fixed. This whole process was repeated 50 times for each of the 11 repeats and constituted a single "run." Ten such runs were performed. The first two runs were used to calculate the mean number of mutations per position after 10, 20, 50, 100 and 150 cycles of conversion in the array and to characterize the distribution of mutations. The distribution of mutations among the 11 repeats was characterized by calculating the

TABLE 1

Number of repeats and their modal lengths in base pairs (bp) in the nontranscribed spacers of rDNA gene units

Clone designation	Source	Modal length in bp	No. repeats per array	Reference
pTa71	Wheat *B* genome	133	12	LASSNER and DVOŘÁK (1986)
pTa250	Wheat *B* genome	133	11	LASSNER and DVOŘÁK (1986)
pTaNorD3	Wheat *D* genome	120	7	LASSNER, ANDERSON and DVOŘÁK (1987)
pScR4	Rye	134	10	APPELS, MORAN and GUSTAFSON (1986)
pZmr1	Maize	200	10	MCMULLEN *et al.* (1986)
λMr1	Maize	200	9	TOLOCZYKI and FEIX (1986)
—	Mouse	$116+(T)_n$	13	KUEHN and ARNHEIM (1983)

percentages of mutations in adjacent and alternate repeats, those separated by two repeats, and those separated by more than two repeats after 10, 20, 50, 100 and 150 cycles of conversion. Finally, the percentages of arrays with all mutations in adjacent or alternate repeats after 10, 20, 50, 100 and 150 conversions were calculated. All ten runs were used to calculate the mean number of conversion cycles leading to elimination or fixation of mutations in the repeated array, the percentage of mutations that become fixed relative to those that became eliminated by the process, and the ratios of recurrent to nonrecurrent mutations in individual repeats. All recurrent mutations of one kind at a nucleotide position are considered a single mutational event.

Statistical analysis: Analysis of variance for completely randomized design with unequal sample size was performed for each characteristic of the sequenced rDNA and computer generated repeated arrays reported in Tables 3–5. The level of a characteristic after the 10, 20, 50, 100 and 150 conversion cycles in the computer generated arrays and in the rDNA were the "treatments." If *F*-test was significant at the 5% level for a characteristic, pairwise comparisons were made by calculating least significant difference (LSD) values for the 5% probability level taking into account unequal sample size.

The statistical significances of the mean numbers of conversions needed to eliminate or fix a mutation and the percentage of mutations that became fixed in the computer generated arrays were tested by analysis of variance for completely randomized design considering the repeat in which a mutation originated as a replication and model as a treatment (Tables 7–9). Pairwise comparisons were made by calculating LSD at the 5% probability level.

The ratios of recurrent to nonrecurrent mutational events were statistically tested (Table 6) in the following manner. Because the mean frequency of nonrecurrent mutations was 1 in all ratios, only the relative proportions of recurrent mutational events at each repeat were analyzed by analysis of variance for completely randomized design with equal sample size. If *F*-test was significant at the 5% probability level, LSD at 5% probability level was calculated to make pairwise comparisons.

RESULTS

Sequenced rDNA spacer repeated arrays: A mutation is defined as a departure of a repeat from a consensus sequence. Because consensus sequence is defined as a sequence of nucleotides most frequently occurring at each nucleotide position in the array, the mean number of mutations of one kind per nucleotide position cannot exceed half the number of repeats. Additionally, more than one kind of mutation occasionally occurs at a position and since these were treated separately, as if they were at different positions, the maximum number of mutations of one kind per position may be additionally reduced. If no homogenization occurred, the mean number of mutations per polymorphic nucleotide position is expected to be 1.0. Because the same mutation can by chance originate more than once, some recurrent mutations will be encountered even if no homogenization occurred, and the mean would be higher than 1.0.

The mean number of mutations of one kind per polymorphic position varied from 1.2 in the rye repeated array to 2.8 in the mouse repeated array (Table 2). In the repeated arrays from rye and the *Nor-D3* locus of wheat (pTaNorD3) most mutations were unique, whereas in the mouse rDNA repeated array most mutations were recurrent. This is clearly apparent from the ratio of the recurrent to nonrecurrent mutational events (all recurrent mutations at a nucleotide position are considered as a single event). This ratio is 0.2:1 for the rye and pTaNorD3 repeated arrays, whereas it is 1.8:1 for the mouse array (Table 2). This ratio is similar for two rDNA clones from the wheat *Nor-B2* locus (pTa71 and pTa250) and identical for two rDNA arrays from maize (pZmr1 and λMr1) (Table 2) suggesting that this trait may be genome specific. The correspondence of the ratios for the maize clones is remarkable considering the fact that the repeated array in clone λMr1 contains about twice as many mutations than that of clone pZmr1.

In all arrays except that of rye, most recurrent mutations were in adjacent repeats, fewer were in alternate repeats, and still fewer were separated by two repeats which were about equal to those separated by more than two intervening repeats. A recurrent mutation at a nucleotide position could be in more than two repeats. Such a mutation could simultaneously be in, *e.g.*, adjacent and alternate repeats and would figure twice in the data causing the sum of the percentages to exceed 100. The number of nucleotide positions in which all recurrent mutations were in adjacent or alternate repeats exceeded 40% in all

TABLE 2

Mean number of mutations per polymorphic nucleotide position in the repeated sequences of wheat, maize, rye and mouse rDNA spacers and characterization of the distribution of mutations present in more than one repeat (recurrent mutations) among the repeats

Characteristic	Cloned sequence						
	Wheat *Nor-B2* pTa71	Wheat *Nor-B2* pTa250	Wheat *Nor-D3* pTa NorD3	Maize pZmr1	Maize λMr1	Rye pScR4	Mouse —
Mean no. mutations of one kind per polymorphic nucleotide position	2.4	2.4	1.4	2.1	1.5	1.2	2.8
Ratio of recurrent to nonrecurrent mutational events	0.6:1	0.8:1	0.2:1	0.6:1	0.6:1	0.2:1	1.8:1
Percent of recurrent mutations in adjacent repeats	78	70	60	62	61	0	69
Percent of recurrent mutations in alternate repeats	21	37	20	25	13	100	36
Percent of recurrent mutations in repeats separated by two repeats	4	0	20	7	20	0	10
Percent of recurrent mutations in repeats separated by more than two repeats	8	6	10	16	25	0	15
Percent of nucleotide positions in which all recurrent mutations are in adjacent or alternate repeats	70	80	40	73	44	100	52

arrays. In the rye repeated array all recurrent mutations were in alternate repeats. In this context the rye repeated array is clearly exceptional and will not be included into comparisons of the distribution of mutations in the NTS repeated arrays and those generated by computer.

Computer simulated evolution of repeated arrays: *Random model.* In this computer simulation any member of the array of repeats had equal chance to form a DNA heteroduplex with the randomly selected source repeat. Repeated cycles of DNA heteroduplex formation and conversion resulted in elimination of most mutations from the array. An average of 90.9% of mutations were eliminated, while 9.1% of mutations were fixed in the array (Table 9). The ratio of recurrent to nonrecurrent mutational events steadily increased from 0.6:1 after 10 cycles of conversion in the array to 9.2:1 after 150 cycles (Table 3). Reproduction of a mutation due to repeated cycles of DNA heteroduplex formation and conversion was associated with progressive increase in the number of mutations occurring in the adjacent repeats and decrease in the number of mutations in repeats separated by two or more repeats (Table 3).

In the repeated arrays of rDNA spacers listed in Table 2 there was an average of 2.1 mutations of one kind per polymorphic nucleotide position. This number of mutations per polymorphic nucleotide position is close to 1.9 mutations of one kind per polymorphic nucleotide position after 20 conversion cycles in the computer generated repeated arrays. Although mutations in an rDNA NTS differ in their age they will be treated as if they originated at the same time in comparisons with the repeated arrays generated by the computer. After 20 conversion cycles the computer generated repeated arrays were quite different from those of the rDNA spacers. In the computer generated arrays 33.9% of recurrent mutations were in adjacent repeats, whereas 66.7% were observed in the rDNA repeated arrays (Table 3). Conversely, 38.4% recurrent mutations were separated by more than two repeats whereas only 13.3% of such recurrent mutations were observed in the rDNA repeated arrays. Recurrent mutations were clearly more clustered in the repeated arrays of rDNA spacers than is predicted by this model; 59.8% of nucleotide positions had all recurrent mutations in adjacent or alternate repeats in the rDNA repeated arrays but only 30.0% in the computer generated arrays. The comparison of the distribution of recurrent mutations in the repeated arrays of the rDNA with that in the computer generated array is unsatisfactory and, therefore, the hypothesis that heteroduplex formation is independent of the distance between the repeats is rejected.

Model 1. In this computer simulated evolution of repeated arrays the likelihood of DNA heteroduplex formation between two repeats decreased exponentially as the distance between the repeats increased. The probability of DNA heteroduplex formation was 0.500 between adjacent repeats, 0.250 between alternate repeats, 0.125 between those separated by two repeats, etc. As in the previous simulation, most mutations were eliminated from the repeated array; 91.4% mutations followed this fate whereas 8.6% were fixed (Table 9). The ratio of recurrent to nonrecurrent mutational events increased from 1:1 after 10 cycles of conversion to 9.9:1 after 150 cycles of conversion. Mutations that by chance survived the

TABLE 3

Mean number of mutations per nucleotide position in computer generated repeated arrays after specified number of conversions in an array of eleven repeats and characterization of the distribution of recurrent mutations among the repeats

Characteristic	No. conversion cycles since the origin of mutation					Observed mean of rDNA arrays
	10	20	50	100	150	
Mean no. mutations per polymorphic nucleotide position	1.5a*	1.9a	3.0b	4.7c	5.3d	2.1a
Ratio of recurrent to nonrecurrent mutational events	0.6:1a	1.1:1b	2.0:1c	6.2:1d	9.2:1e	0.7:1a
Percent of recurrent mutations in adjacent repeats	26.9a	33.9b	60.4c	69.8d	77.0e	66.7cd
Percent of recurrent mutations in alternate repeats	26.2a	27.1ab	28.9ab	31.2ab	33.3b	25.3ab
Percent of recurrent mutations in repeats separated by two repeats	21.0a	17.5a	19.1ab	14.1b	10.2b	10.2b
Percent of recurrent mutations in repeats separated by more than two repeats	42.2a	38.4a	22.2b	14.4c	11.1c	13.3bc
Percent of nucleotide positions in which all recurrent mutations are in adjacent or alternate repeats	26.9a	30.0a	26.0a	42.5b	47.4b	59.8b

The likelihood for a repeat to form a heteroduplex with the target repeat is equal for all members of the array. Means of two runs of 50 simulations each.

* Numbers in rows followed by the same letter are not statistically significant at the 5% probability level according to least significant difference (LSD).

repeated cycles of conversion became progressively more abundant (Table 4). The observed 2.1 mutations of one kind per polymorphic position in the rDNA was linearly extrapolated to correspond to 14 cycles of conversion. It was linearly extrapolated from data in Table 4 that after 14 conversion cycles 69.9% and 27.8% of the recurrent mutations would be in the adjacent and alternate repeats, respectively, and 78.5% of the nucleotide positions with recurrent mutations would have all mutations in adjacent or alternate repeats. These numbers are close to the observed percentages of recurrent mutations in adjacent and alternate repeats in the rDNA repeated arrays: 66.7% and 25.3%, respectively. The observed percentage of positions with recurrent mutations with all mutations in adjacent or alternate repeats, 59.8%, was lower,

TABLE 4

Mean number of mutations per nucleotide position in computer generated repeated arrays after specified numbers of conversions in the array of eleven repeats and characterization of the distribution of recurrent mutations among the repeats

Characteristic	No. conversions since the origin of mutation					Mean of rDNA arrays
	10	20	50	100	150	
Mean no. mutations per polymorphic nucleotide position	1.8a*	2.5a	4.0b	5.2c	5.1c	2.1a
Ratio of recurrent to nonrecurrent mutational events	1.0:1a	1.9:1b	3.6:1c	10.4:1d	9.9:1d	0.7:1a
Percent of recurrent mutations in adjacent repeats	65.5a	76.4bc	82.8c	89.6c	85.3c	66.7ab
Percent of recurrent mutations in alternate repeats	29.7a	26.6a	19.4b	14.4c	17.2c	25.3ab
Percent of recurrent mutations in repeats separated by two repeats	10.6a	8.4ab	7.0ab	6.0b	2.0c	10.2a
Percent of recurrent mutations in repeats separated by more than two repeats	7.1a	5.9a	6.2a	4.9a	5.4a	13.3b
Percent of nucleotide positions in which all recurrent mutations are in adjacent or alternate repeats	78.5a	78.5a	69.1a	70.1a	73.1a	59.8a

The likelihood for a repeat to form a DNA heteroduplex with the target repeat decreases with the distance from the target repeat according to model 1. Means of two runs of 50 simulations each.

* Numbers in rows followed by the same letter are not statistically significant at the 5% probability level (LSD).

TABLE 5

Mean number of mutations per nucleotide position in computer generated repeated arrays after specified numbers of conversions in an array of eleven repeats and characterization of the distribution of recurrent mutations among the repeats

Characteristic	No. conversions since the origin of mutation					Mean of rDNA arrays
	10	20	50	100	150	
Mean no. mutations per polymorphic nucleotide position	1.7a*	2.4b	3.7c	4.8d	5.4e	2.1ab
Ratio of recurrent to nonrecurrent mutational events	0.7:1a	1.3:1b	4.0:1e	5.6:1d	6.7:1e	0.7:1a
Percent of recurrent mutations in adjacent repeats	47.1a	55.4b	73.2c	74.2cd	81.9d	66.7bc
Percent of recurrent mutations in alternate repeats	26.7a	27.8a	26.6a	25.7a	26.1a	25.3a
Percent of recurrent mutations in repeats separated by two repeats	19.4a	16.7ab	12.9bc	10.8bc	8.8c	10.2bc
Percent of recurrent mutations in repeats separated by more than two repeats	23.5a	23.7a	14.2a	16.0ab	11.1b	13.3ab
Percent of nucleotide positions in which all recurrent mutations are in adjacent or alternate repeats	50.1ab	44.1a	46.4a	43.6a	57.2ab	59.8b

The likelihood for a repeat to form a DNA heteroduplex with the target repeat is decreasing with the distance from the target repeat according to model 2. Means of two runs of 50 simulations each.

* Numbers in rows followed by the same letter are not significantly different from each other at the 5% probability level (LSD).

but not significantly, than the 78.5% extrapolated. The extrapolated percent of recurrent mutations in repeats separated by two repeats, 9.3%, agreed with the 10.2% in rDNA NTS. The percent of mutations in repeats separated by more than two repeats was somewhat lower in repeated arrays generated by computer than in those of rDNA. The ratio of recurrent to nonrecurrent mutational events was linearly extrapolated to be 1.3:1 which is slightly but significantly higher than observed 0.7:1 in rDNA NTS.

Model 2. In this model the likelihood for a target repeat to be involved in the formation of DNA heteroduplex with the source repeat also decreased with distance from the source repeat. The likelihood of DNA heteroduplex formation was 0.331 between adjacent repeats, 0.167 between alternate repeats, 0.110 between those separated by two repeats, etc. Most mutations were eliminated from the arrays; 90.1% were eliminated and 9.9% were fixed (Table 9). As in the two preceding models, the numbers of mutations per polymorphic position increased with the number of conversion cycles. The mean of 2.1 mutations of the same kind per polymorphic position in the rDNA repeated arrays was linearly extrapolated to correspond to 16 conversion cycles. It was linearly extrapolated from data in Table 5 that after this number of conversions 52.1% recurrent mutations would be in adjacent repeats, 27.3% in alternate repeats, 18.3% in repeats separated by two repeats and 23.6% in repeats separated by more than two repeats. The percentages observed in rDNA, 66.7, 25.3, 10.2 and 13.3%, respectively, did not significantly (LSD) differ from the above data. The extrapolated percentage of nucleotide positions at which all mutations were in adjacent or alternate repeats, 47.7%, was significantly lower (LSD) than the observed 59.8%. The ratio of recurrent to nonrecurrent mutational events was linearly extrapolated to be 1.1:1 after 16 conversion cycles which is not significantly differ from the observed ratio of 0.7:1.

Ratios of recurrent to nonrecurrent mutational events, elimination and fixation of mutations: The ratios of recurrent to nonrecurrent mutational events increased with the number of conversion cycles in all three models (Tables 3–5). The position in the array of a repeat in which a mutation originated had an effect on the propagation of the mutation only if relatively few conversions had occurred in the array since the origin of the mutation (data not shown). While there was no difference in the ratio among the 11 repeats under the random model, in the repeated arrays generated by computer using model 1 mutations at the edges of the array, particularly in the end repeats (1 and 11) required more conversion cycles to initiate propagation than internal repeats (Table 6). A similar tendency was apparent in the computer simulation according to model 2 (Table 6). The location of the repeat in which the mutation originated had, however, no effect on the mean number of conversions needed to occur in the array to eliminate or fix the mutation, regardless of the model (Tables 7–9). More conversions were needed to eliminate or fix a mutation in the array under the random model than under the distance-dependent conversion models

TABLE 6

Mean ratios of recurrent to nonrecurrent mutational events in the array of 11 repeats after 10 cycles of conversion

	Repeat in which the mutation originated											
Model	1	2	3	4	5	6	7	8	9	10	11	Mean
Random	0.6:1a*	0.5:1a	0.6:1a	0.7:1a	0.7:1a	0.5:1a	0.6:1a	0.5:1a	0.6:1a	0.6:1a	0.6:1a	0.6:1**
Model 2	0.6:1a	0.9:1b	0.9:1b	0.9:1b	0.9:1b	0.9:1b	1.0:1b	1.0:1b	0.7:1ab	0.7:1ab	0.5:1a	0.8:1
Model 1	0.5:1a	1.0:1b	1.0:1b	1.1:1b	1.2:1b	1.6:1bc	1.0:1b	1.2:1b	1.1:1b	0.9:1ab	0.6:1a	1.0:1

The likelihood for the repeats to form a DNA heteroduplex is equal for all repeats (random model) or decreases with distance (models 1 and 2). Mutation originated in the specified repeat and each ratio was calculated from 10 runs of 50 simulations leading either to elimination or fixation of mutation.

* Ratios in the rows followed by the same a, b or c letter are not significantly different at the 5% probability level (LSD).

** Mean ratios in the last column are all significantly different from each other at the 5% probability level (LSD).

TABLE 7

Mean number of conversions in the array of 11 repeats needed to eliminate a mutation from the array

	Repeat in which the mutation originated											
Model	1	2	3	4	5	6	7	8	9	10	11	Mean
Random	45.2	39.5	38.2	36.8	38.5	38.9	38.9	41.1	40.0	39.1	37.3	39.6a*
Model 2	38.1	38.9	32.8	33.6	33.2	35.6	34.6	34.9	32.5	32.6	34.3	34.6b
Model 1	45.6	32.4	26.8	24.8	29.4	24.9	30.9	37.4	29.1	30.8	33.4	31.4b

The mutation originated in the specified repeat. The likelihood for two repeats to form DNA heteroduplex is either independent of distance (random model) or decreases with distance (models 1 and 2).

* Means followed by the same letter are not significantly different at the 5% probability level (LSD).

TABLE 8

Mean number of conversions in the array of 11 repeats needed for the fixation of a mutation in the array

	Repeat in which the mutation originated											
Model	1	2	3	4	5	6	7	8	9	10	11	Mean
Random	181	173	181	186	186	188	163	201	190	184	179	183a*
Model 2	160	157	125	150	169	170	155	165	148	159	177	171a
Model 1	168	134	142	153	145	120	147	117	169	138	165	145b

The mutation originated in the specified repeat. The likelihood for two repeats to form DNA heteroduplex is either independent of the distance (random model) or decreases with distance (models 1 and 2).

* Means followed by the same letter are not significantly different at the 5% probability level (LSD).

TABLE 9

Percent of mutations that become fixed in the array of 11 repeats by repeated cycles of conversion

	Repeat in which the mutation originated											
Model	1	2	3	4	5	6	7	8	9	10	11	Mean
Random	9.6	9.6	7.1	9.0	14.2	7.2	7.0	9.5	8.0	9.0	10.0	9.1a*
Model 2	7.2	11.0	9.5	9.1	8.6	11.6	12.6	11.2	9.6	8.9	9.5	9.9a
Model 1	8.8	6.3	7.5	8.0	8.6	8.2	8.3	8.6	10.0	8.4	10.3	8.6a

The likelihood for two repeats to form a DNA heteroduplex is equal for all repeats (random model) or decreases with distance (model 1 and 2). Mutation originated in the specified repeat.

* Means followed by the same letter are not significantly different at the 5% probability level (LSD).

(Tables 7 and 8) because the distance-dependent models were more efficient in propagation of mutations (Table 6). There was no statistical difference among the three models in the percentages of mutations that survived the repeated conversion cycles (Table 8); in all three models less than 10% of mutations that originated in the repeated array were ultimately fixed (Table 9).

Conversion events in the rDNA repeated arrays: The distribution of mutations in three rDNA spacers were analyzed in an attempt to determine the approximate lengths of each conversion-like event and whether their frequency is distance-dependent.

Clone pTa71. Positions at which recurrent mutations occur in the repeated array are listed at Figure 1. Mutations are designated 1, whereas the consensus is

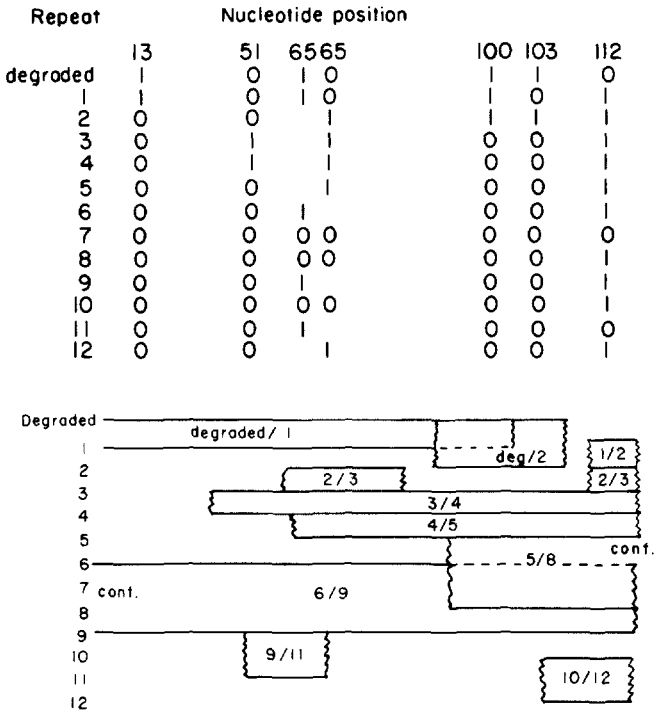

FIGURE 1.—The upper part shows the distribution of mutations (1) relative to the consensus sequence (0) at the nucleotide positions in which recurrent mutations occur in pTa71. The position of the repeat in the array is specified on the left. The most upstream repeat contains part of a repeat of which homology with other repeats is degraded by many mutations. This repeat is designated degraded. At position 65 two different recurrent mutations occur. The position is, therefore, present twice. The lower part shows the putative conversions apparent from the distribution of the recurrent mutations. The length of each conversion is specified by the horizontal lines for both involved repeats. The repeats involved in a conversion are designated by a fraction, *e.g.*, 2/3 indicates that repeats 2 and 3 are involved. Note that it is not clear in most cases exactly where each conversion begins and ends.

designated 0. Comparison of pTa71 and pTa250 indicated that a mutation at position 112 has spread through the array except for repeats 7 and 11 which retain the original nucleotide (LASSNER and DVOŘÁK 1986). The conversion-like events suggested by recurrent mutations are shown in Figure 1. In some cases it was impossible to decide what sequence of events lead to the present distribution of mutations among repeats. It is, nevertheless, clear that most conversion-like events are shorter than the length of a repeat (133 bp) and majority involved repeats near each other. Of 10 putative conversions 6 involved adjacent repeats, 3 alternate repeats, and 1 involved repeats separated by 2 repeats.

Clone pTa250. Comparison of pTa250 with pTa71 indicated that a mutation at position 33 spread through the array except for repeat 11 which contains the original nucleotide (LASSNER and DVOŘÁK 1986). No conversion-like event was longer than the length of the repeat, 133 bp. The data suggested 3 conversions involving adjacent repeats, 3 involving alternate

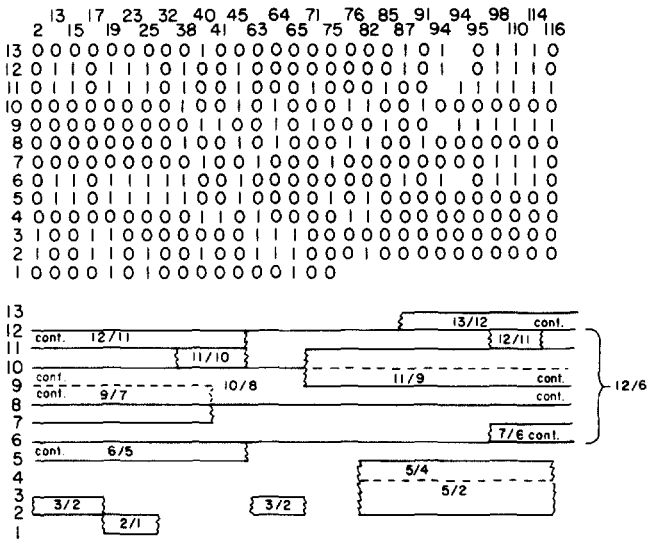

FIGURE 2.—The distribution of recurrent mutations (1) relative to the consensus sequence (0) in the rDNA from a mouse clone and putative conversions in the repeated array. Two different recurrent mutations occur at position 94, the position is, therefore, present twice. For further explanation see Figure 1.

repeats and 1 involving repeats separated by four repeats (data not shown).

Mouse rDNA clone. Because of a high number of recurrent mutations in the mouse spacer array many conversion-like events can be traced with a high degree of confidence. For example, a conversion-like event between repeats 12 and 13 begins around position 84 and extends beyond position 46 in repeats 11 and 12, spanning almost the entire length of a repeat. There is a total of 12 mutations in this region of the array and all are faithfully duplicated in this interval (Fig. 2). Another long conversion-like event occurs between repeats 9 and 11 which continues through repeats 8 and 10 into repeats 7 and 9. The conversion spans more than a length of a repeat. All 15 recurrent mutations in this interval are duplicated. The distribution of recurrent mutations in the array suggests 8 conversions between adjacent repeats, 2 between alternate repeats, and 1 between repeats separated by five repeats. Because conversions resulting in reversions of mutations into the consensus nucleotide are hard to identify, no attempt was made to account for those. The absolute number of conversions in this array and those of the wheat clones must, therefore, be greater. The present assignment of conversion-like events into the mouse array resembles the geneology among the repeats reported by KUEHN and ARNHEIM (1983) but they differ in several places. KUEHN and ARNHEIM considered the repeat as a unit of homogenization, which is probably artificial because heteroduplex can be initiated and terminated anywhere within the array.

DISCUSSION

It was assumed in all three models that conversions occur without a directional bias. In a system in which

each repeat has an equal chance to be involved in a conversion with any other repeat, a directional bias in conversions was shown to result in reduction of the average number of conversion cycles needed to fix or eliminate polymorphism (BIRKY and SKAVARIL 1976). The present computer simulations showed that large numbers of conversions were needed to spread a unique mutation through the repeated array. The average number of conversion cycles needed to fix a mutation in the array of 11 repeats ranged from 145 for model 1 to 183 for the random model. If directional bias does occur in at least some of the conversions, these numbers would be unrealistic.

Both the computer simulation and the sequence data for the NTS spacers showed that homogenization does not involve all repeats in array equally. The end repeats at both 3′ and 5′ ends of NTS repeated arrays almost invariably contain a large number of unique mutations. While nothing like that was suggested by the simulation based on the random model, the other two models suggested the same phenomenon. Mutations that originated in repeats 1 and 11 had reduced likelihood to be reproduced or eliminated from the repeats relative to mutations that originated elsewhere. Vice versa, mutations originating inside of the array have reduced likelihood of proliferating into the end repeats. As a result, the ratio of recurrent to nonrecurrent mutational events are the lowest for repeats 1 and 11. If repeats most frequently form DNA heteroduplex with their immediate neighbors, the end repeats will be less frequently involved in a DNA heteroduplex and conversion because they have a neighbor only on one side. Once an end repeat accumulates a substantial number of unique mutations, or fails to accumulate mutations that become prevalent inside the array, its ability to form a DNA heteroduplex with the internal repeats may be even further reduced as a result of reduced nucleotide sequence homology. Thus the end repeats of an array are expected to become degraded by mutations. Since this tendency is most prominent if the likelihood of the DNA heteroduplex decreases rapidly with distance between repeats (model 1), the ubiquity of degraded repeats flanking the NTS repeated arrays is one more line of evidence that distance-dependent homogenization is more realistic than the model assuming that DNA heteroduplex formation occurs randomly.

The repeat in which a mutation originated had no apparent effect on the average number of conversions in the array needed to fix or eliminate the mutation from the array. Since it takes a large number of conversions to a fix a mutation, the initially slow start for the proliferation of mutations originating in repeats 1 and 11 will ultimately become negligible.

The central assumption, common to all three models, was that homogenization in the rDNA is mediated by conversions. However, the observed distribution of recurrent mutations in NTS would be also obtained if homogenization proceeded by unequal crossing over (KIMURA and OHTA 1979). The agreement of the observed distribution of recurrent mutations in NTS with that predicted by computer simulations speaks, therefore, neither for nor against either mechanism of homogenization; the preference for conversions over crossing over as the principal means of homogenization of the NTS spacer arrays is based on other arguments (LASSNER and DVOŘÁK 1986). Of them the most important is that *Nor* loci in wheat and related species usually show only one or two NTS length variants. If unequal crossing over were the principal means of spreading a mutation through the NTS array, many NTS lengths would be expected in each allele.

The repeated arrays generated by the computer resembled the actual repeated arrays in the NTS of rDNA only when the likelihood of DNA heteroduplex formation decreased with the distance between repeats. Although the actual relationship between the distance between repeats and the likelihood of heteroduplex formation was not determined, the simulation based on model 1 concurred with the observed data somewhat better than that based on model 2, indicating that the likelihood of DNA heteroduplex formation is decreasing rapidly with the increase in the distance between repeats.

That conversions play a significant role for homogenization of repeated gene families within a chromosome and on nonhomologous chromosomes has been shown in fungi (KLEIN and PETES 1981; JINKS-ROBERTSON and PETES 1985; KOHLI *et al.* 1984; ROEDER, SMITH and LAMBIE 1984). However, only meager experimental data are available to show that conversions are distance-dependent, as hypothesized here. The best line of evidence comes from the work of ROEDER, SMITH and LAMBIE (1984) who followed the movement, by gene conversion, of a yeast TY element which was marked by insertion of the *URA 3* gene. The movement was clearly distance dependent, the conversion of the marked TY element most frequently occurred with TY elements residing in the same chromosome. Within the chromosome, conversions occurred more frequently with nearby elements. On rare occasions the marked TY element duplicated prior to its insertion at another TY element, resulting in a tandem triplication. Remarkably, in one of the two cases this happened one of the two duplicated TY elements was converted by the unmarked one. The conversion of a TY element placed in tandem was, thus, three orders of magnitude higher than a conversion of a TY element elsewhere on the same or other chromosomes. On the other hand, conversions of

three serine t-RNA genes, two on chromosome *I* and one on chromosome *III*, of *Schizosaccharomyces pombe* did not reveal preference for genes on the same chromosome and all occurred rarely (KOHLI *et al.* 1984). Whether this is caused by the genes on the same chromosome being too far apart to show a "distance effect" or by other factors is not clear.

Other experimental data consistant with the hypothesis that conversions are distance dependent come from the 5 S rRNA gene family of *Neurospora crassa*. Genes coding for the 5 S rRNA are on at least six of the seven chromosomes and are dispersed among unrelated nucleotide sequences (METZENBERG *et al.* 1985). The family appears to be subjected to homogenization by conversions restricted to the coding regions (MORZYCKA-WROBLEWSKA *et al.* 1985). Genes sharing common sequence show a tendency to be near each other (METZENBERGER *et al.* 1985). If the relative homogeneity of the 5 S family is maintained by gene conversion, the likelihood of conversions must decrease with the distance between the genes.

If homogenization by conversions is distance-independent, a repeated family would be progressing slowly from one state to another more or less uniformly throughout the entire genome. If homogenization is, however, distance-dependent, an entirely different pattern is expected. This pattern would be characterized by a tendency toward clustering of variants in a repeated gene family. Translocation or transposition of a group of repeated sequences into a distant site is expected to initiate stochastic evolution of a new gene family. Repeated genes belonging to a single gene family present at several loci almost always show clustering of variants. For example, Chinese Spring wheat rDNA loci *Nor-B1* and *Nor-B2* residue on chromosomes *1B* and *6B*, respectively; each locus contains two NTS-length variants. Those that are at *Nor-B1* are different from those at *Nor-B2*. In addition to this clustering the two variants within the *Nor-B2* locus are also clustered (DVOŘÁK and APPELS 1986). A clustering of rDNA variants was also shown for *Drosophila melanogaster* (SHARP, GHANDI and PROCUNIER 1983) and *Elytrigia elongata* (DVOŘÁK *et al.* 1984). The 5 S rRNA genes in wheat are on the short (p) arms of chromosomes *1B*, *1D*, *5A*, *5B* and *5D*. The gene units on chromosomes of group 1 differ greatly in spacer length and nucleotide sequence from those on chromosomes of group 5 (LASSNER and DVOŘÁK 1985). Clustering of the repeated genes coding for wheat gliadins is another example. These genes belong to several closely related subfamilies occurring on chromosomes of homoeologous groups 1 and 6. Subfamilies of β- and γ-gliadins occur in both homoeologous groups whereas α-gliadins occur only on chromosomes of groups 6 and ω-gliadins occur only on chromosomes of group 1 (PAYNE *et al.* 1984).

The spread of a mutation through a repeated gene family, if it is not guided by natural selection or favored by a bias in conversion, depends on the frequency of the conversion and/or unequal crossing over events occurring in the family. The number of generations needed for a repeated gene family to fix a mutation depends, therefore, to a large extent on the number of these events per generation. An increase in the frequency of a rare variant in the family can potentially increase its effect on the phenotype, and, thus, expose the variant to natural selection; this could aid its fixation if it provides an organism with a selective advantage. Hence, the frequency of conversions per generation can potentially affect the rate of evolution of a population. It is, therefore, intriguing that even closely related species differ in the ratio of recurrent to nonrecurrent mutations in the NTS arrays. If mutation rates are more-or-less constant in the NTS of different genomes, the differences in the ratios of recurrent to nonrecurrent mutations would indicate that the homogenization rates are different. In *T. aestivum* the *B*-genome rDNA spacers may be subjected to a rate of homogenization four times faster than those of the *D*-genome. Maize rDNA is homogenizing with a rate similar to the wheat *B* genome whereas rye rDNA is homogenizing with a rate similar to that of the *D* genome. However, the rye rDNA spacer may be exceptional since the large deletion that occurs in alternate repeats of the array (APPELS, MORAN and GUSTAFSON 1986) probably causes the repeated unit to correspond no more to the 133 bp of the wheat unit but to 1.61 of it. This could account for the unusual distribution of recurrent mutations which all are in alternate repeats. If this assumption is correct, then all recurrent mutations in alternate repeats would actually be in adjacent repeats. Based on this assumption the data for rye rDNA can be included into the calculation of rDNA means from data in Table 2, yielding 71.4, 21.7, 8.7 and 11.4% of recurrent mutations in adjacent and alternate repeats, in repeats separated by two repeats and repeats separated by more than two repeats, respectively. These data are even closer to those obtained by computer simulation using the exponential model (model 1), 69.9, 27.8, 10.2 and 6.4%, respectively, than those used earlier. Also the observed number of nucleotide positions in which all recurrent mutations were in adjacent or alternate repeats, 65.5%, is closer to 78.5% obtained from the model 1 computer simulation.

The ratio of recurrent to nonrecurrent mutational events indicates a higher rate of homogenization in the mouse rDNA NTS than any of the plant rDNA spacers. Whether the differences among the genomes reflect differences in the overall rate of heteroduplex

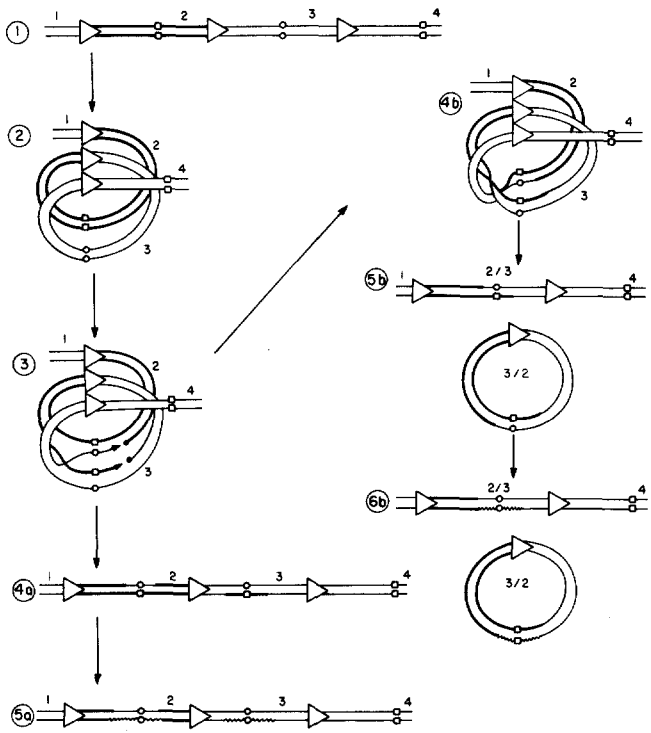

FIGURE 3.—Topology of heteroduplex formation leading to conversion and crossing over between adjacent repeats in a tandem array of four repeats. (1) Four repeats are shown of which repeat 3 differs by a mutation (○) from the consensus (□). (2) Alignment of repeats 2 and 3. (3) Strand invasion introducing a base mismatch (4a) into heteroduplex formed between repeats 2 and 3. (5a) Excision and repair shown by wavy line results in conversion of repeat 2 by repeat 3. Isomerization (4b) results in an excission of a circular hybrid repeat composed of a part of repeat 3 and part of repeat 2 (5b), followed by excision and repair of single strands (6b).

formation in the entire genome or only in rDNA spacers is not clear. It is possible that they are related to presence or absence of recombination stimulating sequences in rDNA. In that case the rDNA spacer of mouse should be more recombinogenic than that of the wheat *B* genome or maize, which in turn should be more recombinogenic than that of the *D* genome. A dodecamer sequence has been identified in wheat rDNA which is a source of hypervariation in NTS and was suggested that the dodecamer may stimulate recombination (LASSNER, ANDERSON and DVOŘÁK 1987). It is intriguing that the repeats of the *B* genome which appear to homogenize faster than those of the *D* genome have two of these sequences in tandem whereas the *D* genome has only one (LASSNER, ANDERSON and DVOŘÁK 1987). It is also intriguing that some murine rDNA alleles show a presence of a large number of spacer length variants (KUEHN and ARNHEIM 1983), whereas only one or a few variants have always been detected in wheat and related species (APPELS and DVOŘÁK 1982; DVOŘÁK and APPELS 1982). The wheat rDNA *Nor-B2* locus was indeed shown to be subjected to low levels of homologous crossing-over (DVOŘÁK and APPELS 1986). Regardless of the cause, it appears that even closely related genomes, such as the wheat *B* and *D* genomes, may potentially show significant differences in the rates of evolution of the same repeated sequence family.

The preponderance of conversions in the wheat and mouse NTS were shorter than the length of a single repeat. Although there is no *a priori* reason for this pattern to occur if heteroduplex was occurring among distantly placed repeats, this is an inevitable consequence if the heteroduplex formation occurred between adjacent repeats. Figure 3 shows how we envision such a DNA heteroduplex occurring. If branch migration in the heteroduplex proceeds further than the length of a single repeat, it begins displacing the already formed DNA heteroduplex. If this does happen the bridge migration and heteroduplex movement can proceed through the array but at any given time the heteroduplex is a maximum of one repeat long. A homologous stretch longer than a repeat was detected in the mouse rDNA repeated array, but it was between alternate repeats (Fig. 2) which can generate a heteroduplex of a maximum length of two repeats. In wheat clones the only stretch of homology indicating a conversion longer than a single repeat involved repeats separated by two repeats (Fig. 1). Thus, the preponderance of conversion-like events shorter than the length of a repeat is also compatible with the proposal that the likelihood of heteroduplex is highest between neighboring repeats and rapidly declines with the distance.

If the intermediate structure of heteroduplex shown in Figure 3 is resolved toward crossing over, the outcome is a deletion of a double-stranded circle from the locus. The circle provides a potential for transposition of a repeat elsewhere in the genome and for amplification of a single or a few repeats by rolling circle replication (HOURCADE, DRESSLER and WOLFSON 1973) and then transposition to other places in the genome. It may, therefore, not be surprising that in specific cases sequences homologous with the NTS repeats are found interspersed in the genome as, *e.g.*, in mouse (ARNHEIM *et al.* 1980). This could be one of the general mechanisms how some interspersed repeated sequences are generated and move around the genome.

The authors express their appreciation to JAN DVOŘÁK, JR., for his assistance with the statistical analyses of the data and to MARCIA CARY for drawing the figures. This work was supported by NSF grant DCB 8316480 to J.D.

LITERATURE CITED

AMSTUTZ, H., P. MUNZ, W.-D. HEYER, U. LEUPOLD and J. KOHLI, 1985 Concerted evolution of tRNA genes: intergenic conversion among three unlinked serine tRNA genes in *S. pombe*. Cell **40:** 879–886.

APPELS, R. and J. DVOŘÁK, 1982 The wheat ribosomal DNA spacer region: its structure and variation in populations and among species. Theor. Appl. Genet. **63:** 337–348.

APPELS, R., L. B. MORAN and J. P. GUSTAFSON, 1986 The structure of DNA from the rye (*Secale cerale*) *Nor-R1* locus and its behavior in wheat backgrounds. Can. J. Genet. Cytol. **28:** 673–685.

ARNHEIM, L., P. SEPERACK, J. BANERJI, R. B. LANG, R. MIESFELD and K. B. MARCU, 1980 Mouse rDNA nontranscribed spacer sequences are found flanking immunoglobulin C_H genes and elsewhere throughout the genome. Cell **22:** 179–185.

BIRKY, W. C. and R. V. SKAVARIL, 1976 Maintenance of genetic homogeneity in systems with multiple genomes. Genet. Res. **27:** 249–265.

BRITTEN, R. J. and D. E. KOHNE, 1968 Repeated sequences in DNA. Science **161:** 529–540.

DVOŘÁK, J. and R. APPELS, 1982 Chromosome and nucleotide sequence differentiation in genomes of polyploid *Triticum* species. Theor. Appl. Genet. **63:** 349–360.

DVOŘÁK, J. and R. APPELS, 1986 Investigation of homologous crossing over and sister chromatid exchange in the wheat *Nor-B2* locus coding for rRNA and *Gli-B2* locus coding for gliadins. Genetics **115:** 1037–1056.

DVOŘÁK, J., M. W. LASSNER, R. S. KOTA and K.-C. CHEN, 1984 The distribution of the ribosomal RNA genes in the *Triticum speltoides* and *Elytrigia elongata* genomes. Can. J. Genet. Cytol. **26:** 628–632.

HOURCADE, D., D. DRESSLER and J. WOLFSON, 1973 The amplification of ribosomal RNA genes involves a rolling circle intermediate. Proc. Natl. Acad. Sci. USA **70:** 2926–2930.

JACKSON, J. A. and G. R. FINK, 1981 Gene conversion between duplicated genetic elements in yeast. Nature **292:** 306–311.

JINKS-ROBERTSON, S. and T. D. PETES, 1985 High-frequency meiotic gene conversion between repeated genes on nonhomologous chromosomes in yeast. Proc. Natl. Acad. Sci. USA **82:** 3350–3354.

KIMURA, M. and T. OHTA, 1979 Population genetics of multigene family with special reference to decrease of genetic correlation with distance between gene members on a chromosome. Proc. Natl. Acad. Sci. USA **76:** 4001–4005.

KLEIN, H. L. and T. D. PETES, 1981 Intrachromosomal gene conversion in yeast. Nature **289:** 144–148.

KOHLI, J., P. MUNZ, R. AEBI, H. AMSTUTZ, G. GYSLER, W.-D. HEYER, L. LEHMANN, P. SCHUCHERT, P. SZANKASI, P. THURIAUX, U. LEUPOLD, J. BELL, V. GAMULIN, H. HOTTINGER, D. PEARSON and D. SOLL, 1984 Interallelic and intergenic conversion in three serine tRNA genes of *Schizosaccharomyces pombe*. Cold Spring Harbor Symp. Quant. Biol. **49:** 31–40.

KUEHN, M. and N. ARNHEIM, 1983 Nucleotide sequence of the genetically labile repeated elements 5′ to the origin of mouse rRNA transcription. Nucleic Acid Res. **11:** 211–224.

LASSNER, M., O. ANDERSON and J. DVOŘÁK, 1987 Inferences on evolution and intergenomic homogenization of ribosomal RNA gene spacers based on the molecular structure of a clone from the D-genome of bread wheat. Genome. In press.

LASSNER, M. W. and J. DVOŘÁK, 1985 Organization of the 5S rRNA family in wheat. J. Cell. Biochem. Suppl. 9C, p. 219.

LASSNER, M. and J. DVOŘÁK, 1986 Preferential homogenization between adjacent and alternate subrepeats in wheat rDNA. Nucleic Acid Res. **14:** 5499–5512.

LIEBHABER, S. A., M. GOOSSENS and Y. W. KAN, 1981 Homology and concerted evolution at the $\alpha 1$ and $\alpha 2$ loci of human α-globin. Nature **290:** 26–29.

MCMULLEN, M. D., B. HUNTER, R. L. PHILLIPS and I. RUBENSTEIN, 1986 The structure of the maize ribosomal DNA spacer region. Nucleic Acid Res. **14:** 4953–4968.

METZENBERG, R. L., J. N. STEVENS, E. V. SELKER and E. MORZYCKA-WROBLEWSKA, 1985 Identification and chromosomal distribution of 5S rRNA genes in *Neurospora crassa*. Proc. Natl. Acad. Sci. USA **82:** 2067–2071.

MORZYCKA-WROBLEWSKA, E., E. V. SELKER, J. N. STEVENS and R. L. METZENBERG, 1985 Concerted evolution of dispersed *Neurospora crassa* 5S RNA genes: pattern of sequence conservation between allelic and nonallelic genes. Mol. Cell. Biol. **5:** 46–51.

PAYNE, P. I., L. M. HOLT, E. A. JACKSON and C. N. LAW, 1984 Wheat storage proteins: their genetics and their potential for manipulation by plant breeding. Philos. Trans. R. Soc. Lond. B **304:** 359–371.

ROEDER, G. S., M. SMITH and E. J. LAMBIE, 1984 Intrachromosomal movement of genetically marked *Saccharomyces cerevisiae* transposons by gene conversion. Mol. Cell. Biol. **4:** 703–711.

SHARP, Z. D., V. V. GHANDI and J. D. PROCUNIER, 1983 X chromosome nucleolus organizer mutants which alter major Type I repeat multiplicity in *Drosophila melanogaster*. Mol. Gen. Genet. **190:** 438–443.

SMITH, G. P., 1976 Evolution of repeated DNA sequences by unequal crossover. Science **191:** 528–535.

TARTOF, K. D., 1975 Redundant genes. Annu. Rev. Genet. **9:** 355–385.

TOLOCZYKI, C. and G. FEIX, 1986 Occurrence of a homologous repeat units in the external spacer region of a nuclear maize rRNA gene unit. Nucleic Acid Res. **14:** 4969–4986.

Communicating editor: M. T. CLEGG