

Action of Repeat-Induced Point Mutation on Both Strands of a Duplex and on Tandem Duplications of Various Sizes in *Neurospora*

Michael K. Watters,^{*,1} Thomas A. Randall,^{*,2} Brian S. Margolin,[†] Eric U. Selker[†]
and David R. Stadler^{*}

^{*}Department of Genetics, University of Washington, Seattle, Washington 98195 and [†]Department of Biology, University of Oregon, Eugene, Oregon 97403

Manuscript received February 8, 1999

Accepted for publication May 28, 1999

ABSTRACT

In *Neurospora crassa*, DNA sequence duplications are detected and altered efficiently during the sexual cycle by a process known as RIP (repeat-induced point mutation). Affected sequences are subjected to multiple GC-to-AT mutations. To explore the pattern in which base changes are laid down by RIP we examined two sets of strains. First, we examined the products of a presumptive spontaneous RIP event at the *mtr* locus. Results of sequencing suggested that a single RIP event produces two distinct patterns of change, descended from the two strands of an affected DNA duplex. Equivalent results were obtained using an exceptional tetrad from a cross with a known duplication flanking the zeta-eta (ζ - η) locus. The *mtr* sequence data were also used to further examine the basis for the differential severity of C-to-T mutations on the coding and noncoding strands in genes. The known bias of RIP toward CpA/TpG sites in conjunction with the sequence bias of *Neurospora* accounts for the differential effect. Finally, we used a collection of tandem repeats (from 16 to 935 bp in length) within the *mtr* gene to examine the length requirement for RIP. No evidence of RIP was found with duplications shorter than 400 bp while all longer tandem duplications were frequently affected. A comparison of these results with vegetative reversion data for the same duplications is consistent with the idea that reversion of long tandem duplications and RIP share a common step.

DUPLICATIONS of portions of the genome can be beneficial in that they provide raw material for evolution. The presence of a duplicate copy of a gene should relax selective pressure allowing sequence divergence. In principle, one copy may acquire a novel function that could improve the fitness of the organism. Duplications may also be hazardous, however. Duplicated regions dispersed throughout the genome can lead to ectopic recombination resulting in deletions or translocations. In addition, mutations can result from duplication of a gene segment or from insertion of duplicated DNA into a gene. To achieve a favorable balance between the benefits and hazards associated with duplications, organisms presumably have developed systems for defending their genetic integrity (Selker 1997; Yoder *et al.* 1997).

In the filamentous fungus *Neurospora crassa*, duplications are efficiently detected and modified during the sexual cycle in the time between fertilization and karyogamy (fusion of two nuclei of opposite mating type to

form a true diploid) by a process known as RIP (repeat induced point mutation; Selker *et al.* 1987; Selker and Garrett 1988; Cambareri *et al.* 1989; Selker 1990). RIP attacks the duplicated regions, heavily mutating both the original and duplicate copy with multiple GC-to-AT transitions. Sequences mutated by RIP are also typically found to be methylated at most of the remaining cytosine residues (Selker *et al.* 1987, 1993a,b; Cambareri *et al.* 1989). The mutations by RIP do not occur randomly; C-to-T mutations occur principally (~75%) at sites 5' of adenines, to a lesser extent at sites 5' of thymines or guanines, and rarely at sites 5' of other cytosines (Cambareri *et al.* 1989; Grayburn and Selker 1989; Selker 1990). RIP presumably reduces the hazards of gene duplications by reducing the homology between duplicated regions, thus reducing the likelihood of exchange between them that might lead to rearrangements (Cambareri *et al.* 1991). In principle, rapid divergence of duplicated sequences by RIP could lead to evolution of new functions, but no evidence for this is yet available.

The timing and effects of RIP are well established but its mechanism is largely unknown. Restriction analyses of sequences affected by RIP almost invariably revealed evidence of identical changes between sister chromosomes in a given meiotic tetrad, indicating that RIP occurred prior to, or during, premeiotic DNA synthesis (Selker 1990). Occasional differences were seen, how-

Corresponding author: Michael K. Watters, Department of Botany, University of British Columbia, 6720 University Blvd., Vancouver, BC V6T 1Z4, Canada. E-mail: mwatters@mail.botany.ubc.ca

¹Present address: Department of Botany, University of British Columbia, Vancouver, BC V6T 1Z4.

²Present address: Department of Plant Pathology, University of California, Riverside, CA 92521.

ever, consistent with the possibility that RIP operates distinctly on both nucleotide chains of a single DNA duplex (Selker *et al.* 1987). Sequences mutated by RIP frequently show only C-to-T or G-to-A changes (not both) on a given chain (Selker *et al.* 1993a,b; Singer *et al.* 1995; Watters and Stadler 1995), suggesting that RIP makes only one type of change (either C to T or G to A). It is attractive to assume that RIP results from methylation and deamination of cytosines in duplicated sequences detected by a mechanism involving DNA:DNA pairing (Selker 1990). We cannot rule out other possible mechanisms, however, such as deamination of unmethylated cytosines. Restriction analyses revealed tetrads with different patterns of RIP among duplication-bearing spores (Selker *et al.* 1987). If this is due to the action of RIP on both chains of the final duplex prior to karyogamy, comparison of the DNA sequences should reveal opposite changes among products of that RIP event. However, if these differences are due to the action of RIP on only a single chain, then only one new RIP pattern should be detected among products of the event (Figure 1).

Here we report an examination of these possibilities. First, we characterized the presumptive products of a tetrad that resulted from an apparent RIP event at the *mtr* locus. The *mtr* strains were found in a study of spontaneous mutation in the sexual cycle and appear to have resulted from the action of RIP on a proposed transient spontaneous duplication (Watters and Stadler 1995). We then confirmed these results with a definite tetrad in which restriction analysis had revealed spores with different RIP patterns (Selker *et al.* 1987). We also report analyses of the distribution of mutations by RIP in *mtr* and *am* that shed additional light on observations that RIP sometimes generates mostly silent mutations, sometimes generates relatively more nonsense and missense mutations (Singer *et al.* 1995; Watters and Stadler 1995), and occasionally generates functional mutant alleles (Fincham 1990). Finally, we exploited a collection of spontaneous duplications of various lengths in the *mtr* gene to examine the relationship between repeat length, frequency of RIP, and the rate of vegetative reversion resulting from homologous recombination between the duplicate elements. This was motivated by prior evidence that RIP, like recombination, involves direct pairing of repeated sequences: RIP attacks sequences in pairs (Selker and Garrett 1988), regardless of the number of repeats present (Fincham *et al.* 1989), and is associated with a high rate of premeiotic loss of duplicated DNA (Selker *et al.* 1987).

MATERIALS AND METHODS

Strains and media: *mtr* mutant strains pA24.1, pA24.2, and pA24.3 were recovered in a survey of spontaneous mutation in the sexual cycle of *N. crassa*. A full account of the origins of these strains has been described (Watters and Stadler

1995) and all *mtr* duplication strains used in this study have been described previously (Stadler *et al.* 1991; Dillon and Stadler 1994; Watters and Stadler 1995). Strains C46-1 and C46-2, which contain a linked duplication modified by RIP, were also previously described (Selker *et al.* 1987).

Vegetative propagation and crosses were carried out on standard growth media (Davis and Deserres 1970) supplemented as described in Stadler *et al.* (1991).

Molecular analysis: Neurospora DNA was prepared by the methods of Zolan and Pukkila (1986) or Oakley *et al.* (1987). Restriction enzymes were used according to manufacturers' instructions. Various portions of the *mtr* gene were amplified by PCR using a variety of 20-mer oligonucleotides. Fragments of ζ - η flank sequences were amplified from genomic DNA by PCR using the following primers: 5' TTCACACAG GAAACAG 3' (primer 1) and 5' CGGGATCCGTCGACGACC TTYATRCGG 3' (primer 2). Two ambiguities were added to primer 2 at the most RIP-mutable dinucleotides (CpA/TpG) so as to be able to amplify sequences that had been modified by RIP. Sequencing of pA24 and SR22 RIP mutants was accomplished using either Sequenase (USB) or fmol DNA sequencing (Promega, Madison, WI).

Reversion analysis: High-frequency reversion rates were monitored by a fluctuation test, using the method of the median (Vonborstel 1978). Each *mtr* duplication strain was grown in a series of parallel tube cultures. Reversion was monitored by streaking a drop of conidial suspension containing about 30,000 cells on selective medium. Larger aliquots were used to measure low-frequency reversion. The mutation rates graphed in Figure 4 are means from 3 to 14 trials.

RIP analysis: All *mtr* duplication strains were crossed to *mtr*⁺ strains. Progeny bearing the duplication are *Mtr*⁻, whether or not they have experienced RIP. Changes in reversion frequencies provide evidence of RIP, however. Unaffected progeny revert at a high rate like their parents, while those that have been mutated by RIP are not revertible, presumably because of the base changes by RIP. This change in reversion frequencies was used to score the progeny of these crosses for RIP.

RESULTS

Operation of RIP on both strands of a duplex: We examined two sets of strains to gain insight into whether RIP can cause different mutations on two strands of a duplex. First, we sequenced all of the recovered descendants of an apparently simple RIP event collected as part of a survey of spontaneous mutation during the sexual cycle of Neurospora (Watters and Stadler 1995). In that study, only 52 spontaneous mutant events were detected in a total of 1500 cross-plates. This low rate provided confidence that all mutant spores recovered from a given plate originated from a single mutational event. Mutants in this survey were collected by plating all recoverable spores from a cross in order to maximize the recovery of sibling mutants (descendants of a single mutational event). The efficiency of recovering other mutants in the study indicated that recovery of siblings was very efficient. Ascospores recovered in groups of two, three, or four were consistent with origination from single asci, as judged by examination of independent markers in the cross (Watters and Stadler 1995). The strain that showed mutations char-

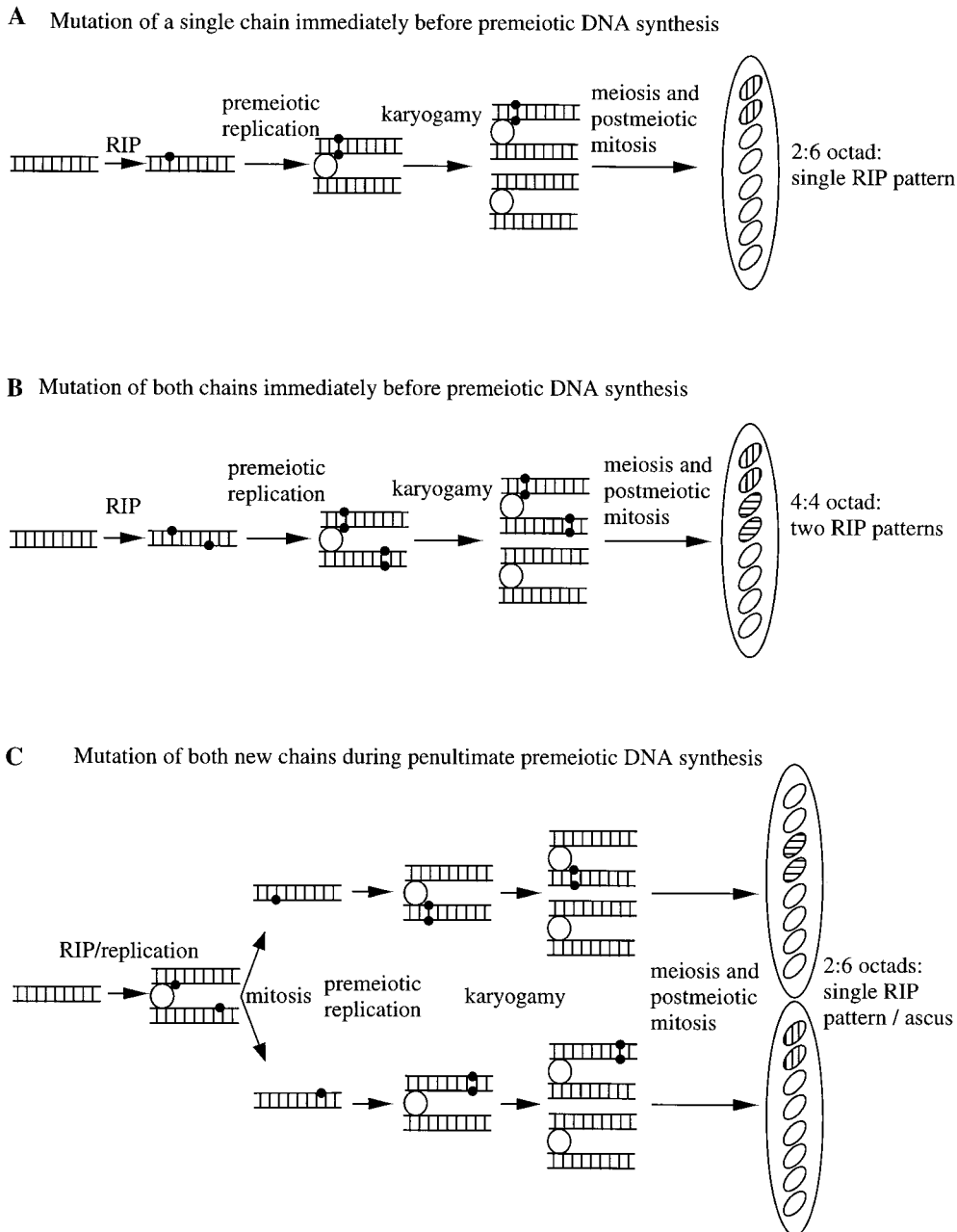


Figure 1.—Potential patterns of mutation by RIP immediately before karyogamy. Each horizontal line represents a single polynucleotide chain of DNA. Open circles represent centromeres. Solid circles represent mutations caused by RIP. Complete octads of spores are indicated on the right. Spores containing DNA unaltered by RIP are shown as open ovals. Spores containing DNA altered by RIP are shown as marked ovals. Two cross-hatching patterns are shown to indicate the two possible types of changes (C-to-T or G-to-A) on a given strand. Meiosis and the postmeiotic division are not detailed. (A) The single-strand model. Under this model, the RIP machinery damages only one strand and generates heteroduplex DNA. Model A would result in one RIP pattern found in two of the eight spores of an ascus. (B) The two-strand model. In the basic two-strand model, the RIP mechanism affects both strands of a duplex. RIP is assumed to result in only one type of primary mutation (either C-to-T or G-to-A). Thus, the two strands of the duplex must show different RIP patterns. Model B would result in two RIP patterns represented by two spores each. These two patterns would show opposite types of mutations on a given chain. (C) Replication-dependent two strand model. In this derivative of the basic two-strand model, RIP occurs with replication, producing two heteroduplexes with mutations in the newly synthesized chains. As with model B, opposite RIP patterns would be produced, but unlike model B, they would be found in separate asci.

acteristic of RIP, pA24, was collected in a cluster of three mutant spores. These three spores most likely descended from a single mutational event occurring just prior to meiosis and were shot from a single ascus. Sequencing of one of the three isolates revealed 60 GC-to-AT transitions, providing evidence that the mutant had resulted from RIP. In this isolate, all 60 base substitutions were C to T on the sense strand, suggesting that pA24 might have been subjected to only a single round of RIP. Models proposed previously (Watters and Stadler 1995) make different predictions with regard to mutational events occurring just prior to meiosis.

Under model A (Figure 1A), a single strand of the duplex is affected during one cell cycle. The resulting heteroduplex is then resolved either by replication or repair. If the heteroduplex were resolved by replication just prior to entering meiosis, it would produce an ascus in which only one of the meiotic products (two spores in the octad) contained mutations. If the heteroduplex were acted upon by a repair mechanism before replication, it would produce an ascus in which two meiotic products (four spores) contained identical mutations (not shown).

Model B (Figure 1B) proposes that the RIP machinery

pa24.1/3TT ...	
mtr	AAG AAG TCC ATC GTT GCT CTC GGC TTG ATT GAA ATC TTC ATC TAC ACC GTT -	867
pa24.2	
pa24.1/3	
mtr	ACT GGT GGC GTC GTT TAC GCT TTC GTC GGC CCC GAG GTC CAG TCT CCT GCC -	918
pa24.2	
pa24.1/3T	
mtr	TTG CTC TCT GCT GGC CCT CTT CTC GCC AAG GTT GCT TTC GGC ATT GCC CTC -	969
pa24.2 A.. ..	
pa24.1/3T	
mtr	CCC GTC ATC TTC ATC TCT GGC AGT ATC AAC ACT GTT GTC GTC AGC AGG TAT -	1020
pa24.2 A.. ..	
pa24.1/3T ..T	
mtr	CTG ATT GAG CGC ATC TGG CCC AAC AAC GTC ATT CGC TAT GTC AAC ACC CCA -	1071
pa24.2 A.. ..	
pa24.1/3 T.TT	
mtr	GCG GGT TGG ATG GTT TGG CTT GGT TTT GAC TTT GGC ATT ACC CTC ATT GCC -	1122
pa24.2 A.. ..	
pa24.1/3TT	
mtr	TGG GTT ATT GCT GAG GCC ATC CCT TTC TTC TCT GAT CTG TTG GCC ATC TGC -	1173
pa24.2	..A. A.. ..	
pa24.1/3	
mtr	TCG GCT CTC TTC ATT TCC GGT TTT AGC TTC TAT TTC CCT GCC TTG ATG TAT -	1224
pa24.2 A.. ..A ..	
pa24.1/3TTTT	
mtr	TTC AAG ATC ACC AGG AAC GAT GCC AAG AGC CAG GGC AAG AAG TAC TTC TTG -	1275
pa24.2 A.. ..	
pa24.1/3T ..TTT	
mtr	GAT GCC CTC AAC ATG CTC TGC TTC GTC ATC GGC ATG GGC ATT CTT GGT ATT -	1326
pa24.2 A.. ..A .. A.. ..	
pa24.1/3 T.. ..T	
mtr	GGT ACC TAC GCC GCT ATT CAG GAC ATTGTAAGTTTGGCCCGCTTTTCTGTTACTCTTTT -	1385
pa24.2	
pa24.1/3 T.. ..T	
mtr	GCACACAAATGCTAACCTGCTTCTCAG ATG GAC CGT TAC GAC CAT GGC AAG GTT TCG -	1442
pa24.2 A.. ..	
pa24.1/3T	
mtr	AAG CCT TAT AGC TGT GCT CCC TTG GCT TAA	1472
pa24.2 A.. ..	

Figure 2.—(Continued)

diately before, or during, the final round of premeiotic DNA synthesis. To examine this possibility, we sequenced a section of the flank region from the two meiotic products showing different restriction fragment polymorphisms, C46-1 and C46-2. Both common and spore-specific mutations were found. In a 564-bp segment there were 26 mutations present in both of the progeny (Figure 3). These included both C-to-T and G-to-A changes on the chain shown and presumably occurred before the last premeiotic mitotic division. An additional 23 mutations were present in only one of the two mutated progeny. Eleven of these were only in strain C46-1 and were all G-to-A changes on the chain shown. The other 12 mutations were only in strain C46-2 and were all C-to-T changes on the same chain. These data confirm the finding with the pa24 mutants that both chains of the DNA duplex can be modified by RIP. However, the observation of two opposing RIP patterns recovered from a single ascus is consistent with Model B only. This does not eliminate the possibility that other cases of RIP could result from pathways similar to Models A or C.

Additional informative tetrads were sought using an

mtr duplication mutant. Mutant allele SR22, containing a 446-bp tandem duplication in the *mtr* gene, was crossed to *mtr*⁺. Among tetrads examined, none showed differences between spore pairs. Some showed single-type mutation patterns (C-to-T or G-to-A) that may reflect single rounds of RIP before the penultimate premeiotic round of DNA synthesis. If we examine the distribution of mutations in the two elements of the duplication (data not shown), we find significantly different distributions of RIP sites ($\chi^2 = 40.2$, $P < 0.1\%$). This is chiefly due to significantly lower levels of RIP near the outer borders of the duplication, as found previously (*e.g.*, Grayburn and Selker 1989).

Specificity of RIP: One striking difference between pa24 (with C-to-T mutations on the coding strand) and pa24.2 (with G-to-A mutations on the coding strand) was that the damage in pa24, but not pa24.2, was mostly in the third position of codons. All possible third-position C-to-T changes are silent, having no effect on protein sequence. Singer *et al.* (1995) previously noted that most C-to-T mutations on the coding strand in a set of *am*^{RIP} alleles were silent, whereas most G-to-A mutations on the coding strand were missense, and

	1				50
C46-1
parent	GAATTCGATA	TGGGACAGTT	GAGACATGGA	GTCATCCAGC	ACTGCATCGT
C46-2	A.....
	51				100
C46-1
parent	TGGAATCGTG	CTGGAGCAAA	GGTGCACGG	CGACATGTTG	CTCTGAAATA
C46-2	A.....
	101				150
C46-1
parent	GATGCGGTTA	TAGCTGCGTT	TTGTGGGTGT	GCCTGCTTCT	CGGCAGGCTT
C46-2
	151				200
C46-1
parent	..T.....	T.....T.
C46-2	GTCGTTTCCC	CATGAGATCG	TGGCTTTGAC	AATTCGCAAT	TGCTTGGAGG
	201				250
C46-1
parent	CTGTGGCAAG	GGCTATCAAG	AGAGCATTTT	TGTCATTGCA	AAGAGCGGCA
C46-2	A.T.....
	251				300
C46-1
parent	.A.....	A.....
C46-2	TGGGTAATTA	GATCGTCCGA	GGAAGTGATG	CTTCTAGCT	CAATGGCAGT
	301				350
C46-1
parent	CTCCTCCAGG	CGAGTGTGT	TGTTCTGAAA	CATCAGTCTG	AGAAGACCGT
C46-2
	351				400
C46-1
parent	TTGTGGTGAT	GGAGACCAGG	GCACTTTTGT	TTGTATTAGG	ATGGTTAGGG
C46-2
	401				450
C46-1
parent	CCTGTAGGCT	TTGTGTACGA	GCTCTCGTAC	TTGTACAGCC	CATCCCCGGG
C46-2	A.....	A.A.....
	451				500
C46-1
parent	GGTTTTAGCA	TCTTCGGGCT	GAACCCATAC	TGCAGGCCCG	TACATGACGC
C46-2	A.A.....
	501				550
C46-1
parent	TAACTGTCT	TGGATTGAGC	AAAGGGAGCC	AGTAACAACC	GACAACCGCA
C46-2	A.....
	551	564			
C46-1			
parent	TGAAGGTCGT	CGAC			
C46-2			

Figure 3.—Sequence from the flank tetrad. The sequences of strains C46-1 and C46-2 are shown. These strains originated from a single tetrad as described previously (Selker *et al.* 1987). Wild-type and mutant sequences are shown in a manner similar to Figure 2. Mutations common to both progeny are indicated in black text. Mutations unique to only one of the progeny are indicated by white text on a black background.

pointed out that this is attributable to codon usage (*e.g.*, G is common in the first codon position in *Neurospora*, while C is common in the third). In the *mtr* data set of 42 mutational sites in the open reading frame (ORF), 6 were in the first position, 1 was in the second position, and 33 were in the third position of codons. The first base change to cause loss of function is a stop codon introduced at position 1255, as determined by deletion mapping (data not shown). In an attempt to fully account for the lack of mutational impact in this example, we carried out a statistical analysis.

First we examined the distribution of bases in the *mtr* gene by comparing the distribution of mutated C residues in pA24 to the distribution of all C's in the *mtr* ORF (Table 1). The difference between these distributions is significant as determined by χ^2 ($P = 0.1\%$). We

then expanded the scope of the examination to account for the known context bias of RIP. RIP strongly favors 5' CpA 3' sites (5' TpG 3' sites on the complementary strand). We determined the distribution of CpN sites at each codon position in the *mtr* ORF (Table 2). Dividing the number of mutated CpA sites in pA24 (33) by the total number of CpA sites in the *mtr* ORF (110) gives the expected probability that any given CpA site in this mutant would be altered (30%). This is based on the assumption that the probability of a particular C residue mutating via RIP depends solely on its 3' neighboring base. By carrying out similar calculations for the remaining CpN combinations and positions and then applying these probabilities to the chart of CpN sites at various codon positions (Table 2) we derived the predicted distribution of mutations for this mutant

TABLE 1
Codon position bias of RIP in pA24:
Distribution of cytosines

Position ^a	C ^b	RIP ^c
1st	84	6
2nd	104	1
3rd	239	35
Total	427	42

^a Codon position of the cytosine.

^b Number of cytosines in the wild-type sequence.

^c Number of cytosines modified by RIP.

(Table 3a). The expected distribution predicts both the strong bias of mutations in third positions of codons and an apparent shift in CpN bias between first, second, and third codon positions. A comparison of the expected distribution with the observed distribution (Table 3b) revealed no significant difference. A similar analysis (not shown) of the expected distribution of G-to-A mutations by RIP at *mtr* matches the observations for mutant pA24.2 and predicts the observed high frequency of first position changes. This in turn predicts the high frequency of missense mutations caused by G-to-A mutations on the sense strand. Thus the observed difference in mutational impact on the two strands is attributable to the specificity of RIP within the confines of codon usage in this gene.

We carried out an equivalent analysis on mutations in the progeny of the *mtr* duplication mutant SR22 (not shown). Comparison of the observed and expected distributions indicated that they are different at a probability of 3.9%. The significance of this difference is weak and is based primarily on a paucity of RIP at second position CpA sites. The second position CpA sites in the target sequence are all in the extreme 5' region of the duplicated segment, a region that was only lightly mutated in all mutants analyzed.

Length requirements for RIP and mitotic recombination: The dependence of RIP on duplications led to the suggestion that pairing of duplicate sequences is

TABLE 2
Codon position bias of RIP in pA24:
Distribution of CpNs in *mtr*

Position ^a	CpA	CpT	CpG	CpC
1st	19	35	9	21
2nd	5	35	11	53
3rd	86	56	61	36
Total	110	126	81	110
RIP ^b	33	7	1	1
% RIP	30	5.6	1.2	0.9

^a Codon position of cytosine.

^b Number of cytosines modified by RIP.

TABLE 3
Codon position bias of RIP in pA24: Expected and
observed distribution of RIP at cytosines

Position ^a	CpA	CpT	CpG	CpC	Total
a. Expected distribution					
1st	5.7	2	0.1	0.2	8
2nd	1.5	2	0.1	0.5	4.1
3rd	25.8	3.1	0.7	0.3	29.9
Total	33	7.1	0.9	1	42
b. Observed distribution					
1st	3	3	0	0	6
2nd	0	1	0	0	1
3rd	30	3	1	1	35
Total	33	7	1	1	42

^a Codon position of the cytosine.

normally required for RIP to occur. Pairing of duplications is also believed to be required for reversion of tandem duplications via mitotic recombination. Indeed, the linked duplication that led to the discovery of RIP was subject to an unusually high frequency of premeiotic recombination, resulting in elimination of the duplication (Selker *et al.* 1987). A subsequent study using different constructs strengthened the correlation between RIP and recombination (Irelan *et al.* 1994). In the present study, we examined the possible connection between RIP and mitotic recombination using a collection of spontaneous tandem duplications within the *mtr* gene, ranging from 16 to 935 bp long. In particular, we used this set of strains to determine whether vegetative reversion and RIP would show similar dependencies on length of the duplications. The results of the tests on vegetative reversion are presented in Figure 4. For the short duplications, the relationship between length and reversion rate appears exponential. As shown in Figure 4, this relationship extends to ~400–450 bp. Repeats between 446 and 935 bp long all showed similar reversion rates (~10⁻⁴). Tests for RIP revealed that the 446- to 935-bp duplications showed similar sensitivities to RIP, 70–100% of spores containing duplications being affected. No evidence of RIP was detected for duplications shorter than 380 bp long.

DISCUSSION

Mechanism of RIP: Any chromosomal duplication above a threshold size appears susceptible to RIP. The frequency and severity of damage by RIP varies, however, and depends only partly on obvious variables such as length and position of the duplicate elements. By chance, a duplication may (1) slip through a cross unscathed, (2) be lightly mutated, perhaps by one cycle of RIP, or (3) be heavily mutated and show evidence of more than one round of RIP. Furthermore, products of RIP are sometimes methylated and sometimes not.

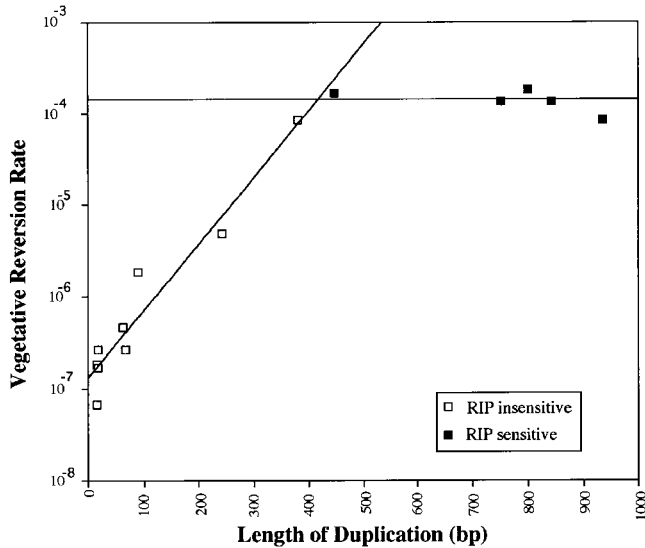


Figure 4.—Vegetative instability and RIP vs. length of duplications in *mtr*. The curve shown for the shorter duplications was determined by regression analysis. The line for the larger duplications represents the mean rate of reversion for these duplications. Several shorter (16, 18, 63, and 68 bp) duplications that were not tested are assumed to be not vulnerable to RIP.

Nevertheless, within a single meiotic tetrad, DNA subjected to RIP typically shows the same alterations, indicating that RIP occurs before, or during, premeiotic DNA synthesis (Selker *et al.* 1987). In a fraction of tetrads (4/25 in the study of Selker *et al.* 1987), two distinct patterns of RIP are seen, consistent with the operation of RIP on individual strands of the duplex entering into meiosis. A less likely possibility is that RIP occasionally operates in meiosis. The fact that RIP commonly leaves just C-to-T but not G-to-A changes, or vice versa, on a given strand strongly suggests that RIP initially causes a single type of mutation (Singer *et al.* 1995; Watters and Stadler 1995). We took advantage of this to address the question of whether RIP can attack both strands of a duplex. A single episode of RIP on both strands of the duplex in the final premeiotic cell (assuming the mutations are not repaired) should give rise to one pair of spores with C-to-T changes on a given strand and another pair with G-to-A changes on the equivalent strand. In contrast, if only one strand were affected in a single episode of RIP, only one pair of spores would show changes (Figure 1).

Here we report two observations consistent with the operation of RIP on both strands of a duplex (Figure 1B). The first evidence came from the discovery of opposite changes (C-to-T vs. G-to-A) in the *mtr* gene of a presumptive tetrad that appeared to have resulted from a single episode of RIP. Supporting evidence came from a definite tetrad showing restriction fragment polymorphisms in the duplicated “flank” region (Selker *et al.* 1987). Sequence analysis of flank DNA from this exam-

ple showed opposite changes (C-to-T vs. G-to-A) between pairs of affected spores superimposed on a mixed background of common mutations reflecting somewhat earlier RIP events. Thus evidence from both a probable RIP event in the *mtr* gene and a definite RIP event in the flank DNA strongly suggests that both strands of a DNA duplex can be modified by RIP during one cell cycle. It remains possible, of course, that RIP does not always affect both strands, *i.e.*, more than one model may be correct. Our data are inconsistent, however, with models in which the mutations must be limited to one strand. It is worth noting that Model B is consistent with several molecular models. One possibility is that a RIP “machine” proceeds down the target sequence, affecting both strands of both copies in one pass. An alternative is that the machine affects only one strand at a time so that multiple passes and/or multiple machines are required to affect both strands. Finally, it is possible that RIP results from a nonprocessive process in which a sequence is first tagged (*e.g.*, by pairing of the duplicate sequences) and then this region is subjected to multiple independent mutagenic events.

Sequence bias of RIP: It is not rare for RIP to result in only C-to-T changes on a given strand. Singer *et al.* (1995) noted that because of codon bias in *Neurospora*, C-to-T changes on the coding strand should have less impact than G-to-A changes. We have expanded and clarified this observation. We demonstrate that the distribution of RIP sites among codon positions can be predicted by the CpN bias of RIP in combination with the sequence of the gene. Calculations (not shown) related to those described above for the *am* and *mtr* genes, but based on the general codon bias of *N. crassa* (Gurr *et al.* 1987), suggest that this conclusion is generalizable. Most (56%) of the highly vulnerable CpA dinucleotides are expected to place the mutable C in the third position of codons, where mutations by RIP are always silent. For the less vulnerable CpT and CpG dinucleotides, the third codon position is also expected to be the most affected. Only for the rarely affected CpC dinucleotides is the third position not the most common location of the first C. Thus C-to-T mutations due to RIP should be concentrated in the third position of codons where they do not affect the protein sequence. These estimates are based on the assumption that the juxtaposition of codons is random.

Similar calculations (not shown) using the codon bias based on available *Neurospora* sequences can be used to predict probable protein changes resulting from RIP. Approximately 82% of C-to-T mutations on the coding strand should result in either silent or conservative changes (75 and 7%, respectively). The rest are divided between nonconservative missense and nonsense mutations. About 62% of G-to-A changes on the coding strand should be silent or conservative (23 and 39%, respectively). Approximately 5% of the sites should be nonsense mutations with the rest being nonconservative

missense mutations. This is similar to the distribution of mutational types observed following RIP at *am* (Singer *et al.* 1995).

In spite of general trends, the consequences of RIP acting on any given sequence will depend on its set and juxtaposition of codons. The location of vulnerable sites relative to regions coding for active sites or sensitive protein domains, for example, will strongly influence a gene's sensitivity to RIP. If changes are largely confined to silent or conservative mutations, protein function may not be abolished. Indeed, RIP does not always result in null alleles (Fincham 1990). In addition, recombination of null alleles can lead to functional alleles that still retain large tracts of RIP mutations (M. K. Watters and D. R. Stadler, unpublished observations).

Minimal sequence length for RIP and recombination:

There are several suggestions that RIP and recombination share common elements. First, both involve pairwise interactions. Second, several correlations between RIP and recombination have been observed (Selker *et al.* 1987; Butler and Metzberg 1989; Selker 1990; Cambareri *et al.* 1991; Foss and Selker 1991; Irelan *et al.* 1994). The association between RIP and recombination remains unclear, however. We took advantage of a set of tandem duplications to examine the length dependence of RIP in relation to the mitotic reversion rates of these repeats. Reversion of tandem duplications is thought to occur by two basic mechanisms: (1) slippage (Streisinger *et al.* 1966; Kunkel 1990; Ripley 1990), in which the repeats mispair during replication, resulting in loss of one copy of the repeat, and (2) recombination, in which the repeats pair and recombine, removing the intervening DNA. Slippage is seen commonly as a source of spontaneous mutation in *Neurospora* at runs of bases and short tandem repeat sequences (Dillon and Stadler 1994; Watters and Stadler 1995). Recombinational mechanisms are constrained by the length of DNA required to pair effectively, leading to a minimum length of repeat required for efficient recombination (Jinks-Robertson *et al.* 1993).

Examination of the dependence of vegetative reversion on length of duplication shows a clear breakpoint at ~400–450 bp. Duplications shorter than this show an exponential relationship between length and vegetative reversion rates. Longer duplications show a roughly constant rate of reversion ($\sim 10^{-4}$). This breakpoint is coincident with the minimum length of repeat that is required for vulnerability to RIP in the sexual cycle and may reflect a shift in the nature of the reversion mechanism at this point (*i.e.*, a shift from slippage to a recombinational mechanism). The coincidence between the minimum lengths for RIP and efficient reversion is consistent with the possibility that reversion of long tandem duplications and RIP share a length-dependent step. This step may be DNA pairing, which is presumed to be required for both RIP and long-repeat reversion.

Although the strains with tandem duplications showing high reversion frequencies were also those subject to RIP, the 380-bp duplication appeared possibly exceptional. This duplication reverts at frequencies similar to those of the longer duplications but does not appear to be subject to RIP. The reason for this is unclear. The length of the duplication may be just below the minimum length required for RIP. It remains possible, however, that lack of RIP in this mutant is due to some undetermined context effect. Examination of a larger sample of duplications in the critical 300–600-bp range would help resolve this question.

In conclusion, we have demonstrated that RIP can affect both strands of a duplex in a single cell cycle. This is a new clue to the mechanism of the RIP process. We have also demonstrated that the sequence bias of *Neurospora* in combination with the context bias of RIP can predict the distribution of RIP sites relative to codon position. This accounts for the potentially low impact of C-to-T mutations on the coding strand of genes and raises the possibility that some genes or portions of genes could be protected from RIP by virtue of their choice of codons and codon combinations. Finally, we compared the frequencies of reversion and RIP for tandem duplications of various sizes and showed that these processes show similar dependencies on length of the duplications.

We thank Helen Macleod for her contributions to preliminary work that led to this study. The work at *mtr* was supported by a grant from the National Science Foundation to D. Stadler (MCB-922057) and a grant from the National Institutes of Health to E. Selker (GM-08995) provided support for work on the flank tetrad.

LITERATURE CITED

- Butler, D. K., and R. L. Metzberg, 1989 Premeiotic change of nucleolus organizer size in *Neurospora*. *Genetics* **122**: 783–791.
- Cambareri, E. B., B. C. Jenson, E. Schabtach and E. U. Selker, 1989 Repeat-induced G-C to A-T mutations in *Neurospora*. *Science* **244**: 1571–1575.
- Cambareri, E. B., M. J. Singer and E. U. Selker, 1991 Recurrence of repeat-induced point mutation (RIP) in *Neurospora crassa*. *Genetics* **127**: 699–710.
- Davis, R. H., and F. J. Deserres, 1970 Genetic and microbiological research techniques for *Neurospora crassa*. *Methods Enzymol.* **17**: 79–143.
- Dillon, D., and D. R. Stadler, 1994 Spontaneous mutation at the *mtr* locus of *Neurospora*: the molecular spectrum in wild-type and mutator strains. *Genetics* **138**: 61–74.
- Fincham, J. R. S., 1990 Generation of new functional mutant alleles by premeiotic disruption of the *Neurospora crassa am* gene. *Curr. Genet.* **18**: 441–445.
- Fincham, J. R. S., I. F. Connerton, E. Notarianni and K. Harrington, 1989 Premeiotic disruption of duplicated and triplicated copies of the *Neurospora crassa am* (glutamate dehydrogenase) gene. *Curr. Genet.* **15**: 327–334.
- Foss, H. M., and E. U. Selker, 1991 Efficient DNA pairing in a *Neurospora* mutant defective in chromosome pairing. *Mol. Gen. Genet.* **231**: 49–52.
- Grayburn, W. S., and E. U. Selker, 1989 A natural case of RIP: degeneration of the DNA sequence in an ancestral tandem duplication. *Mol. Cell. Biol.* **9**: 4416–4421.
- Gurr, S. J., S. E. Unkles and J. R. Kinghorn, 1987 The structure and organization of nuclear genes of filamentous fungi, in *Gene*

- Structure in Eukaryotic Microbes*, Vol. 22, edited by J. R. Kinghorn. Special Publications of the Society for General Microbiology, Washington, DC.
- Irelan, J. T., A. T. Hagemann and E. U. Selker, 1994 High frequency repeat-induced point mutation (RIP) is not associated with efficient recombination in *Neurospora*. *Genetics* **138**: 1093–1103.
- Jinks-Robertson, S., M. Michelitch and S. Ramcharan, 1993 Substrate length requirements for efficient mitotic recombination in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **13**: 3937–3950.
- Kunkel, T. A., 1990 Misalignment mediated DNA synthesis errors. *Biochemistry* **29**: 8003–8010.
- Oakley, C., C. Weil, P. Kretz and B. Oakley, 1987 Cloning of the *riboB* locus of *Aspergillus nidulans*. *Gene* **53**: 293–298.
- Ripley, L. S., 1990 Frameshift mutations: determinants of specificity. *Annu. Rev. Genet.* **24**: 189–213.
- Selker, E. U., 1990 Premeiotic instability of repeated sequences in *Neurospora crassa*. *Annu. Rev. Genet.* **24**: 579–613.
- Selker, E. U., 1997 Epigenetic phenomena in filamentous fungi: useful paradigms or repeat-induced confusion? *Trends Genet.* **13**: 296–301.
- Selker, E. U., and P. W. Garrett, 1988 DNA sequence duplications trigger gene inactivation in *Neurospora crassa*. *Proc. Natl. Acad. Sci. USA* **85**: 6870–6874.
- Selker, E. U., E. B. Cambareri, B. C. Jenson and K. R. Haack, 1987 Rearrangement of duplicated DNA in specialized cells of *Neurospora*. *Cell* **51**: 741–752.
- Selker, E. U., D. Y. Fritz and M. J. Singer, 1993a Dense nonsymmetrical DNA methylation resulting from repeat-induced point mutation in *Neurospora*. *Science* **262**: 1724–1728.
- Selker, E. U., G. A. Richardson, P. W. Garrett-Engele, M. J. Singer and V. Miao, 1993b Dissection of the signal for DNA methylation in the zeta-eta region of *Neurospora*. *Cold Spring Harbor Symp. Quant. Biol.* **58**: 323–329.
- Singer, M. J., B. A. Marcotte and E. U. Selker, 1995 DNA methylation associated with repeat-induced point mutation in *Neurospora crassa*. *Mol. Cell. Biol.* **15**: 5586–5597.
- Stadler, D. R., H. Macleod and D. Dillon, 1991 Spontaneous mutation at the *mtr* locus of *Neurospora*. The spectrum of mutant types. *Genetics* **129**: 39–45.
- Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita *et al.*, 1966 Frameshift mutations and the genetic code. *Cold Spring Harbor Symp. Quant. Biol.* **31**: 77–84.
- Vonborstel, R. C., 1978 Measuring spontaneous mutation rates in yeast. *Methods Cell Biol.* **20**: 1–24.
- Watters, M. K., and D. R. Stadler, 1995 Spontaneous mutation during the sexual cycle of *Neurospora crassa*. *Genetics* **139**: 137–145.
- Yoder, J. A., C. P. Walsh and T. H. Bestor, 1997 Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**: 335–340.
- Zolan, M. E., and P. J. Pukkila, 1986 Inheritance of DNA methylation in *Coprinus cinereus*. *Mol. Cell. Biol.* **6**: 195–200.

Communicating editor: R. H. Davis