tered throughout the interior of the natural RNAs supports the accuracy of the models.

The *Min* 1 RNA may prove more useful than the natural RNase P RNAs for some studies. Because it is smaller than the natural molecules, the *Min* 1 RNA may be more amenable to spectroscopic studies or to crystallization for diffraction analyses. The reduced affinity of the *Min* 1 RNA for the substrate, coupled with its holoenzyme-like maximum velocity, makes *Min* 1 RNA more useful than the natural RNAs for steady-state kinetic analyses that are not complicated by the rate-limiting step of product release. Finally, the *Min* 1 RNA sequence is a new starting point for the elimination of additional sequence elements. We anticipate that further reductions in the size of the active molecule are possible, since many random deletion mutants of the natural RNAs lack conserved features and yet retain some (usually very low) activity (*16, 17*). The inspection of RNase P RNAs from more diverse eubacteria should provide further perspective on the minimal active structure and allow the rational design of RNase P RNAs that are even simpler than the *Min* 1 RNA.

REFERENCES AND NOTES

1. B. C. Stark, R. Kole, E. J. Bowman, S. Altman, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3717 (1978); K. Gardiner and N. R. Pace, *J. Biol. Chem.* **255**, 7507 (1980).
2. C. Guerrier-Takada, K. Gardiner, T. Marsh, N. R. Pace, S. Altman, *Cell* **35**, 849 (1983).
3. C. Reich, G. J. Olsen, B. Pace, N. R. Pace, *Science* **239**, 178 (1988).
4. W. H. McClain, C. Guerrier-Takada, S. Altman, *ibid.* **238**, 527 (1987).
5. J. W. Szostak, *Nature* **322**, 83 (1986); O. C. Uhlenbeck, *ibid.* **328**, 596 (1987).
6. B. D. James *et al.*, *Cell* **52**, 19 (1988).
7. We use the term "homologous" in its strictest sense: homologous sequences have common ancestry and function. Homologous sequences are not necessarily identical; identical sequences are not necessarily homologous.
8. Portions of the *B. megaterium* RNase P RNA sequence were used instead of another Gram-positive RNase P RNA sequence because this would result in the presence of a convenient restriction site in the synthetic gene.
9. Nucleotides 2 to 4 in the *E. coli* sequence were changed from AAG to GGA in the *Min* 1 RNase P RNA in order to accommodate a bacteriophage T7 promoter. Nucleotides 371 to 372 in the *E. coli* sequence were therefore changed from UU to CC in the *Min* 1 RNA in order to maintain Watson-Crick complementarity between the termini. The proposed base pair involving nucleotides 9 and 365 in the *E. coli* sequence was replaced by a different canonical base pair in order to create useful restriction sites (Bcl I and Cla I) near the termini of the synthetic RNase P RNA gene.
10. The two halves of the synthetic gene, defined by the Hind III site in the middle of the designed RNase P RNA, initially were assembled and cloned separately into plasmid vectors. Six synthetic oligodeoxyribonucleotides (40 to 60 nucleotides in length) were annealed to construct each half of the gene. The two halves subsequently were joined at the Hind III site and cloned in a plasmid vector to create pDW133 (*17*). The *Min* 1 RNase P RNA is produced by T7 RNA polymerase-directed transcription of Sma I–cut pDW133 DNA in vitro.
11. C. Reich *et al.*, *J. Biol. Chem.* **261**, 7888 (1986).
12. The *E. coli* RNase P RNA (M1 RNA) was produced by transcription in vitro. An Alu I restriction fragment of *E. coli* DNA that includes all but the first four nucleotides of the RNase P RNA coding sequence (*18*) was joined to a T7 promoter and cloned in pUC19 [C. Yamisch-Perron, J. Vieira, J. Messing, *Gene* **33**, 103 (1985)] to create a template for in vitro transcription (pDW27) (*17*). The T7 RNA polymerase–directed run-off transcript of Sna BI–cut pDW27 DNA has two nucleotides at its 5′ end (GG) and four nucleotides at its 3′ end (UUAC) that are not present in the natural RNase P RNA (Fig. 1). In addition, nucleotides 3 and 4 in the *E. coli* RNase P RNA sequence were changed from AG to UC in the run-off transcript as a consequence of the fusion of the T7 promoter and the truncated RNase P RNA gene.
13. The k_{cat} of the *E. coli* and *Bacillus subtilis* holoenzymes depends upon reaction conditions and the character of the substrate. Examples of values for k_{cat} of the *E. coli* holoenzyme with pre-tRNA substrates are 0.35 min^{-1} [N. Lumelsky and S. Altman, *J. Mol. Biol.* **202**, 443 (1988)], 2 min^{-1} (*2*) 6.6 min^{-1} [C. Guerrier-Takada, A. van Belkum, C. W. A. Pleij, S. Altman, *Cell* **53**, 267 (1988)], and 18.3 min^{-1} (*4*). With the substrate and reaction conditions used in this work, k_{cat} for the holoenzymes of both *E. coli* and *B. subtilis* (*3*) is about 10 min^{-1}.
14. A. Fersht, *Enzyme Structure and Mechanism* (Freeman, New York, 1985), p. 105.
15. K. J. Gardiner, T. L. Marsh, N. R. Pace, *J. Biol. Chem.* **260**, 5415 (1985).
16. C. Guerrier-Takada and S. Altman, *Cell* **45**, 177 (1986); N. P. Lawrence and S. Altman, *J. Mol. Biol.* **191**, 163 (1986).
17. D. S. Waugh, thesis, Indiana University, Bloomington (1989).
18. R. E. Reed, M. F. Baer, C. Guerrier-Takada, H. Donis-Keller, S. Altman, *Cell* **30**, 627 (1982); R. E. Reed and S. Altman, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5359 (1983); H. Sakamoto, N. Kimura, F. Nagawa, Y. Shimura, *Nucleic Acids Res.* **11**, 8237 (1983).

19. Supported by NIH grant GM34527 to N.R.P. C.J.G. was supported by NIH grant GM29231 to B. Vold (SRI International). We are grateful to M. Baer and S. Altman for providing *E. coli* RNase P protein, to B. Pace for providing *B. subtilis* RNase P protein, and to G. Olsen for assistance with artwork and many fruitful discussions regarding the design of the *Min* 1 RNase P RNA. We are also indebted to J. Brown, A. Burgin, S. Darr, E. DeLong, D. Hunt, T. Schmidt, D. Smith, and J. Thomas for helpful comments on the manuscript and to S. Irvine for typing.

6 March 1989; accepted 21 April 1989

Repeat-Induced G-C to A-T Mutations in *Neurospora*

EDWARD B. CAMBARERI, BRYAN C. JENSEN, ERIC SCHABTACH, ERIC U. SELKER*

In the *Neurospora* genome duplicate sequences are detected and altered in the sexual phase. Both copies of duplicate genes are inactivated at high frequency, whether or not they are linked. Restriction sites change, and affected sequences typically become heavily methylated. To characterize the alterations of the DNA, duplicated sequences were isolated before and after one or more sexual cycles. DNA sequencing and heteroduplex analyses demonstrated that the process (termed RIP) produces exclusively G-C to A-T mutations. Changes occur principally at sites where adenine is 3′ of the changed cytosine. A sequence duplicated at a distant site in the genome lost approximately 10 percent of its G-C pairs in one passage through a cross. A closely linked duplication of the same sequence that was passed twice through a cross lost about half of its G-C pairs. The results suggest a mechanism for the RIP process.

DNA SEQUENCE DUPLICATIONS provide the critical first step for gene amplification and provide the raw material for evolution of new genes. While vital for evolution, sequence duplications can also have negative consequences. Dispersed repeated genes can mediate exchanges resulting in deletions, inversions, or translocations. In addition, altered gene dosage can result in a detrimental imbalance of gene products. In many organisms, such as fungi and bacteria, virtually all genes are present in one copy per haploid genome.

In the multicellular fungus *Neurospora crassa*, the paucity of duplicated genes may not simply be due to natural selection. Duplications are efficiently detected and altered in a specialized (dikaryotic) tissue formed by fertilization (*1, 2*). The process affects both copies of a duplicated sequence as revealed by gene inactivation, changes in the position of restriction sites, and de novo methylation of cytosines in the repeated DNA. Because the process is limited to the stage between fertilization and nuclear fusion, susceptible cells have a nucleus from each parent. Thus a cell should survive a duplication and inactivation, even of an essential gene, so long as the duplication were only in one of the parents. Both nuclei should deliver their genetic material to meiosis. Standard recombination processes would produce meiotic products having different combinations of altered and unaltered copies of the duplicated sequences (Fig. 1). Just the class of cells

Institute of Molecular Biology and Department of Biology, University of Oregon, Eugene, OR 97403.

*To whom correspondence should be addressed.

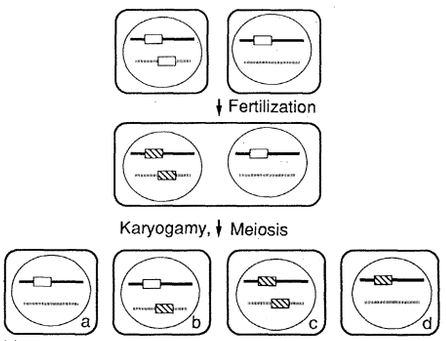

Fig. 1. Fate of duplicated sequences altered by the RIP process. A cross between a wild-type strain (top right) and a strain harboring an unlinked duplication of a chromosomal segment (top left) is illustrated. For clarity only two of the seven chromosomes of *Neurospora* are indicated. Both copies of the duplicated segment (boxed) are altered by RIP (cross-hatching) in the dikaryon resulting from fertilization (*1*). The homologous segment in the wild-type nucleus (white box) is unaltered, and in general, should complement mutations resulting from RIP. Nuclear fusion and meiosis immediately follow. The four possible combinations of chromosomes harboring the altered or unaltered homologs are shown. If the duplicated segment included essential genes, two of the meiotic products (c and d) should be inviable. One of the remaining meiotic products (b) would receive both altered and unaltered sequences.

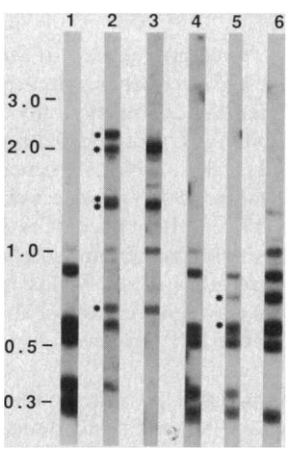

Fig. 2. Restriction analysis of duplicated sequences exposed to the RIP process. Samples of genomic DNA (~0.5 μg) from transformant T-ES74-1 (lane 1), the second generation offspring of T-ES174-1 ($L\text{-}G_2$, lane 2), transformant T-ES174-9 (lane 4), and the first generation offspring of T-ES174-9 ($U\text{-}G_1$, lane 5) were digested with Mbo I. For comparisons, samples of plasmid DNA (~0.2 ng) of pEC24 (cloned from $L\text{-}G_2$; lane 3) and pBJ20 (cloned from $U\text{-}G_1$, lane 6) were digested with Sau 3A. Mbo I was used for the genomic DNA since this enzyme is insensitive to cytosine methylation, typical of DNA affected by RIP. Its isoschizomer, Sau 3A, was used for the plasmid DNA since Sau 3A is insensitive to adenine methylation, contributed by the bacterial host. Restriction digests were fractionated through a 1% agarose gel and probed for the duplicated anonymous ("flank") sequences (*13*). Bands resulting from RIP are marked with asterisks.

receiving only altered copies may be inviable. Because of the timing of the inactivation and the alteration of restriction site patterns, this process was designated "rearrangement induced premeiotically" (RIP).

As a step toward understanding the consequences of RIP, and to learn about its mechanism, we have examined the nature of the sequence alterations that result from it. RIP operates on both genetically linked and unlinked duplicated sequences, although to different extents. A 6-kb direct duplication separated by approximately 7 kb of single-copy sequences never survived a cross unaltered, whereas an unlinked duplication of the same sequences escaped RIP at a frequency of approximately 50% (*1*). Judging from the changes in restriction patterns observed for linked versus unlinked duplicated sequences, we concluded that proximity of the duplicated elements influences the severity of alterations by the RIP process. We therefore studied a segment of DNA from an unlinked duplication that had been relatively lightly altered by RIP, as well as a segment from a linked duplication that may have suffered the maximum possible damage.

We chose a linked duplication of an arbitrary 6-kb sequence that had been exposed to two generations of RIP as an example of a severely altered sequence. DNA blot hybridizations indicated that the size of virtually every restriction fragment in the duplicated DNA of the strain $L\text{-}G_2$ had been altered (Fig. 2, lanes 1 and 2). Subsequent crosses of this strain did not result in further changes in the restriction pattern, suggesting that the duplicated sequences may have suffered the maximum possible damage (*3, 4*). As an example of a mildly affected sequence, we selected a first generation isolate, $U\text{-}G_1$, from a cross of an unlinked duplication of the same arbitrary sequence. Most of the restriction fragments in the duplicated sequences of $U\text{-}G_1$ appeared unaltered (Fig. 2, lanes 4 and 5). Although only minor alterations were observed with restriction enzymes insensitive to cytosine methylation, such as Mbo I and Taq I, digests with enzymes sensitive to cytosine methylation, such as Sau 3A, demonstrated that cytosines in the duplicated sequences of $U\text{-}G_1$ were methylated by RIP.

Our investigation of the alterations resulting from RIP was facilitated by cloning the affected sequences. To do this, we took advantage of bacterial plasmid (pUC8) se-

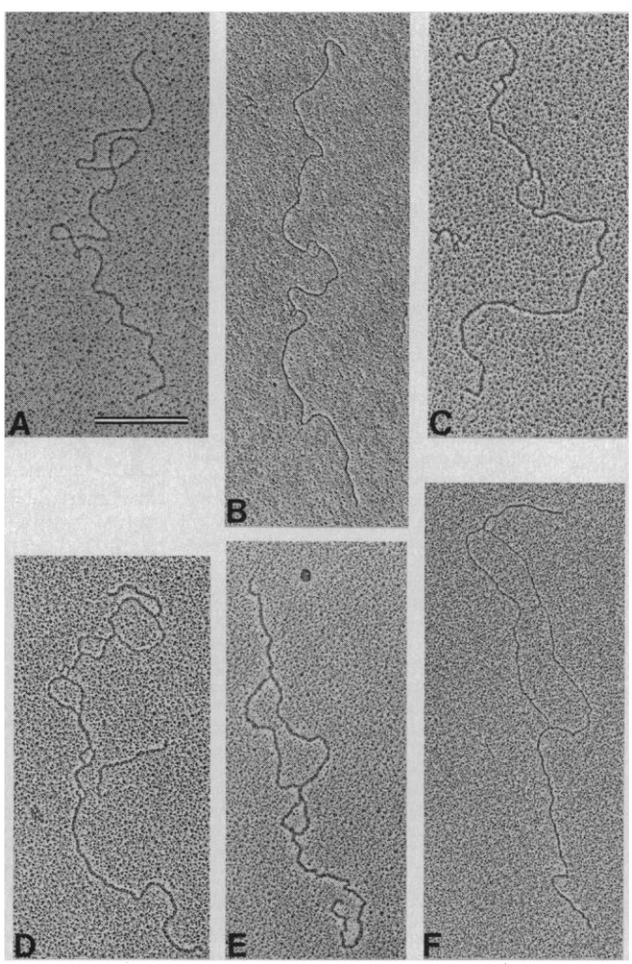

Fig. 3. Heteroduplex analysis of a sequence severely altered by the RIP process. Plasmid DNA of pEC24 and pEC25 (containing the altered sequences of $L\text{-}G_2$ and the unaltered sequences of T-ES174-1, respectively) were linearized by digestion with Bgl II, denatured, and allowed to renature together for 20 to 30 min at 25° to 30°C by standard techniques (*14*). Samples of the heteroduplex solution were diluted to give 0.1*M* tris (*p*H 8.5), 0.01*M* sodium EDTA containing 35% (**A**), 40% (**B**), 50% (**C**), 60% (**D**), or 70% (**E**) formamide, or 0.05*M* tris (*p*H 8.5), 0.005*M* sodium EDTA containing 80% (**F**) formamide. The samples were then spread for electron microscopy at room temperature (~22°C) under isodenaturing conditions. A 6-kb segment of these ~12.5-kb plasmids was duplicated in the transformant and thus presumably susceptible to RIP. Representative heteroduplexes are shown. The bar corresponds to 0.5 μm.

quences adjacent to one copy of the duplicated DNA. Restriction mapping showed that, in each case, a single Bgl II fragment contained the desired sequences along with the bacterial sequences that allow for selection and propagation of the DNA in *Escherichia coli*. Genomic DNA was digested with Bgl II, treated with DNA ligase, and transformed into a strain of *E. coli* tolerant of DNA that contains methylcytosine (*5*). As expected, colonies arising on medium containing ampicillin contained the desired plasmids. The altered sequences of L-G_2, U-G_1, and the parental sequences from transformant T-ES174-1, were isolated in this way. Restriction analysis of the cloned sequences that had been exposed to RIP revealed the novel fragments detected in the genomic DNA, strengthening the conclusion that the alterations were not simply due to some form of DNA modification such as methylation (Fig. 2, lanes 3 and 6).

We had previously noted that the alterations resulting from RIP do not, at least in general, result in large changes in the overall length of the affected sequences (*1*). It was not clear from the Southern hybridizations, however, whether the alterations were classical rearrangements (that is, insertions, deletions, inversions, or substitutions), point mutations, or a combination of these. To determine whether the cloned sequences, altered by RIP, had suffered any gross rearrangements, we prepared heteroduplexes between the altered sequences and their native counterparts and examined them by electron microscopy. Heteroduplexes between pBJ20, the plasmid containing the mildly altered sequence from isolate U-G_1, and pES174, the parental plasmid, appeared completely paired. Rearrangements involving about 200 bp or more would have been detected. We conclude that the changed restriction fragments in the duplicated sequence of U-G_1 must result from small changes such as point mutations.

In contrast to the results with pBJ20, electron microscopy of heteroduplexes between pEC24 (having the severely altered sequences of L-G_2) and pEC25 (having the corresponding unaltered sequences) revealed structural differences between these plasmids. Under standard heteroduplexing conditions, small single-stranded bubbles, diagnostic of substantial differences in base sequence or substitutions of DNA segments were seen (Fig. 3B). To distinguish between these alternatives, we examined heteroduplexes spread under a variety of conditions. The standard conditions (40% formamide, 0.1*M* tris, 22°C) are expected to allow pairing of molecules having up to 40% mismatch (*6*). If the observed bubbles reflected sequence divergence, it seemed likely that they would shrink, or disappear, if the stringency of the hybridization were further relaxed. Heteroduplexes prepared under more stringent conditions, however, might reveal additional regions of difference. If the bubbles were due to substitutions, their size should not change with the hybridization conditions. In fact, the size and number of bubbles did vary with the conditions (Fig. 3). The heteroduplexed molecules became increasingly single-stranded as the formamide concentration was increased. At 70% formamide, virtually the entire duplicated region was single-stranded.

In the opposite direction, heteroduplexes prepared in 35% formamide and spread under isodenaturing conditions exhibited just one or two very small bubbles (Fig. 3A). When the formamide concentration was reduced 5% further, complete pairing was observed, indicating that the original bubbles were not due to substitutions. We conclude, therefore, that the RIP process does not result in gross rearrangements.

On the basis of heteroduplex analyses, it seemed likely that RIP results in point mutations. We therefore sequenced segments of the native and altered duplicated regions. We determined the sequence of ~0.6-kb segments of the native and altered regions of the linked duplicated sequence and ~1-kb segments of the native and altered regions of the unlinked duplicated sequence (Fig. 4). In both cases, all the differences were point mutations, and they were exclusively polarized transition mutations of one type; G-C base pairs in the native sequence were replaced by A-T base pairs in the altered sequence. Thus, the RIP process results in point mutations, consistent with the results from the heteroduplex analyses. On the basis of this information, we suggest changing the name of the phenomenon from "rearrangement induced premeiotically" to "repeat-induced point mutation."

In a segment of 925 bp of the unlinked duplication, which started out with a G+C content of 53%, 10% of the G-C base pairs were converted to A-T base pairs. On the strand shown, there are 34 changes of G to A, but only 13 changes of C to T (Fig. 4B), which is significantly different from a random distribution. In the case of the linked duplicated sequences exposed to RIP through two generations, 31% of the G-C base pairs were converted in the first 563 bp of the duplicated segment. As expected from the heteroduplex analysis, the distribution of mutations is not uniform. Only six changes occurred in the first 120 bp of the duplicated segment, and these were all within a 16-bp stretch (Fig. 5A). In a 300-bp

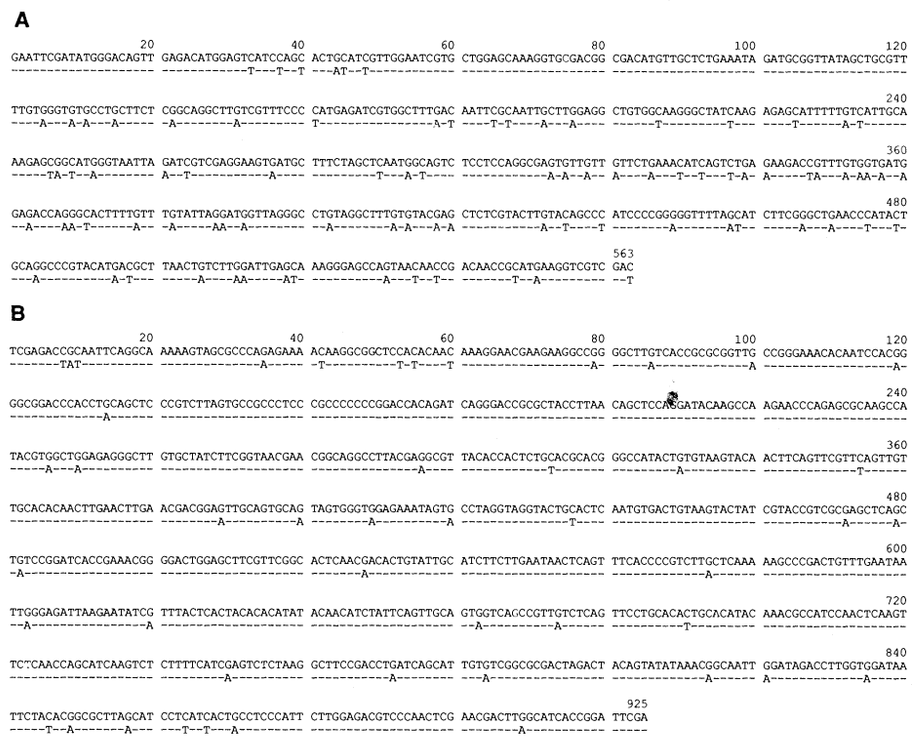

Fig. 4. Sequence comparisons of DNA segments before and after exposure to the RIP process. Paired clones were sequenced on both strands and compared. The complete nucleotide sequence of the unaltered segment is shown (above) and differences resulting from the RIP process are indicated underneath. The sequences compared in (**A**) came from the linked duplication (segment extends to the right from the Eco RI site at the edge of the duplication). The sequences compared in (**B**) came from the unlinked duplication (segment extends to the right starting approximately 1350 bp from the Eco RI site at the edge of the duplication).

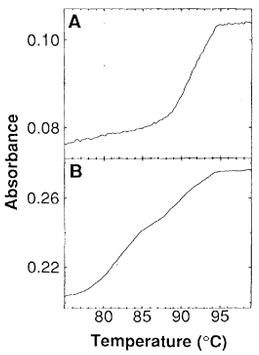

Fig. 5. Reduction of melting temperature of sequence altered by RIP. Melting spectra were obtained with sheared DNA of pEC25 (**A**) and pEC24 (**B**) in 0.15*M* NaCl, 0.015*M* sodium citrate *p*H 7.0. The ~12.5-kb plasmids pEC25 and pEC24 are identical except pEC25 has a native copy of the duplicated 6-kb segment, whereas pEC24 has the altered form of this segment. Absorbance was monitored at 260 nm twice per minute; temperature was increased 0.5°C per minute. Whereas pEC25 melted unimodally around 91°C, pEC24 showed an early melting component around 80°C. These melting temperatures correspond to G+C contents of 53 and 26%, respectively (7). The melting temperature of *N. crassa* genomic DNA in the same buffer was determined as a control. The value obtained (91° to 92°C) corresponds to 53 to 55% G+C, consistent with previous determinations (*15*).

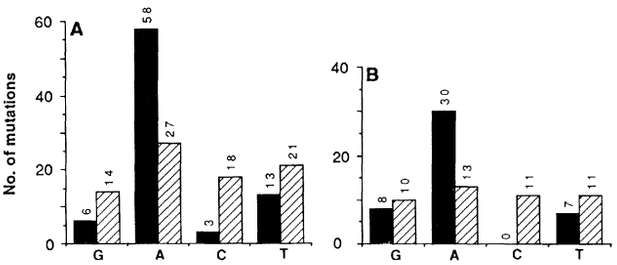

Fig. 6. Sequence-specificity of RIP. Site-specificity was investigated by analysis of the sequence context of the mutations. Each mutation in the linked (**A**) and unlinked (**B**) sequences determined in this study was considered on the strand showing the C to T change. The solid bars indicate the number of mutations that occurred immediately 5′ of the base categories shown. The striped bars indicate the number of mutations that would be expected on a random basis, taking into account the dinucleotide composition of each of the original sequences (Fig. 4). A relatively small number of mutations (8 of 280 for the linked duplication, and 2 of 47 for the unlinked duplication), which occurred adjacent to another mutation, could not be unambiguously assigned to dinucleotide categories, and were therefore not considered in the analysis. The χ^2 values calculated from the distribution of mutations illustrated in (A) and (B) are ~55 and ~33, respectively. Thus, the probability that G-C to A-T mutations occurred at random is <0.001.

segment starting 240 bp from the edge of the duplicated segment, 41% of the G-C base pairs were changed to A-T base pairs, resulting in approximately 21% sequence divergence. The melting behavior of the altered sequences and limited sequencing data from other portions of the duplicated segment make it clear that other segments of the duplicated DNA suffered even more extensive mutation.

The numerous G-C to A-T mutations detected by sequence analysis presumably account for the changed restriction fragments in the regions exposed to RIP and for the single-stranded bubbles observed in heteroduplexes. To get an overall measure of the change in base composition due to RIP, we compared the denaturation profile of plasmids containing the native sequence (pEC25) or severely altered sequence (pEC24). The plasmid pEC25 denatured unimodally around 91°C (Fig. 5A), corresponding to an average G+C content of 53% (7). In striking contrast, the denaturation profile of pEC24 showed a broad early melting component centered around 80°C (Fig. 5B), corresponding to a G+C content of 26%, down 27% relative to the native sequence. We conclude that approximately half of the G-C base pairs in the affected sequence were converted to A-T base pairs as a consequence of the RIP process. The frequency and specific nature of RIP suggest that we reexamine the meaning of "spontaneous" mutation; cells can "cause" mutations to occur.

From the very specific nature of the sequence alterations resulting from RIP we can rule out a number of hypothetical mechanisms for the process and focus attention on likely mechanisms. The RIP process does not result from any of a variety of possible mechanisms that would scramble duplicated sequences. Indeed, there is no reason to believe that RIP involves cleavage of the DNA backbone. However, a high level of intramolecular recombination between closely linked direct repeats is temporally correlated with RIP (*1, 8*).

One of the remarkable features of the RIP process is its efficiency. Duplicated sequences are detected by the cell at high frequency, whether or not they are genetically linked, and are then riddled with polarized transition mutations. The same type of mutation occurs nonspecifically at lower frequencies in bacterial cells that are deficient for uracil DNA glycosylase (*9*), apparently because of spontaneous deamination of cytosines, to give uracils. In normal cells, uracil glycosylase presumably removes uracils from DNA before DNA replication to avert potential mutations. Deamination of 5-methylcytosine produces thymine (5-methyluracil), which is not a substrate of uracil glycosylase. Thus, deamination of 5-methylcytosines in DNA would produce G-T mismatches, which if "repaired" to A-T, or resolved by DNA replication, would establish a polarized transition mutation. Indeed, it has been shown in *E. coli* that 5-methylcytosines can be mutational hot spots (*10*). The frequency of transition mutations in RIP is too high to be accounted for by spontaneous deamination of cytosines or 5-methylcytosines. Nevertheless, RIP may occur by a related mechanism. Specifically, cytosines or methylcytosines in duplicated sequences may be enzymatically deaminated and simply left unrepaired. Mismatches resulting from such a mechanism would be resolved by DNA replication. The fact that more G to A than C to T changes were observed on one strand in the case of the sequence mildly altered by RIP (Fig. 4B), may reflect this proposed mechanism. No asymmetry was observed in the case of the more severely altered sequence, but this is not surprising since this sequence probably underwent multiple cycles of RIP. The process operates before meiosis in the period between fertilization and nuclear fusion, a stage thought to consist of roughly ten cell divisions. According to our hypothesis, the occasional differences in "RIP patterns" observed between sister meiotic products reflect deamination events in the final DNA duplex before premeiotic DNA synthesis.

Although our data reveal that RIP can occur in various sequence contexts, quantitative analysis of the sequences flanking the mutations shows that RIP occurs preferentially at certain G-C base pairs. The mutations of the DNA strand showing C to T changes occurred primarily where there was an adenine 3′ of the changed cytosine, but rarely where there was a cytosine 3′ of the changed cytosine (Fig. 6). For example, in the segment of the linked duplication, ~64% of CpA dinucleotides changed, whereas ~18%, ~13%, and ~5%, respectively, of CpT, CpG, and CpC dinucleotides, changed. The identity of the sec ond base pair on the 3′ side of the mutations, and of the first and second base pairs on the 5′ side of the mutations, did not appear to deviate significantly from a random distribution. Interestingly, the low sequence specificity observed, which is unusual for nucleic acid–enzyme interactions, is characteristic of eukaryotic DNA methyltransferases. The rat enzyme, for example, methylates ~99%, ~14%, ~6%, and ~0%, respectively, of CpG, CpA, CpT, and CpC dinucleotides, in vitro (*11*). Although the distribution of methylation in *N. crassa* has not been examined in detail, we know that 5-methylcytosine is not limited to a particu-

lar dinucleotide in this organism (*12*). Limited sequence specificity may reflect either low precision of an enzyme or else specific recognition of a structural feature of DNA, such as local helical twist, curvature, or mechanical flexibility.

That sequences altered by RIP are typically methylated suggests that methylation itself may play an active role in the mutation process, but this need not be so. The *N. crassa* genome contains "constitutive" methylated sequences that appear to go through crosses unaltered. Thus, methylation alone is not sufficient to cause the frequent transition mutations characteristic of RIP. Conversely, once rendered methylated by RIP, a sequence need not stay duplicated to retain methylation. When the elements of an unlinked duplication, methylated by RIP, segregate to separate cells, the resulting solitary elements continue to be methylated (*2*). Persistent methylation is not simply a result of passive maintenance methylation (*4*). Thus, the point mutations resulting from RIP must themselves be responsible for rendering the sequence a good substrate for DNA methylation. This finding is a potentially important clue for solving the mystery of why certain sequences are methylated in eukaryotes.

REFERENCES AND NOTES

1. E. U. Selker, E. B. Cambareri, B. C. Jensen, K. R. Haack, *Cell* **51**, 741 (1987).
2. E. U. Selker and P. W. Garrett, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6870 (1988).
3. The linked duplication of the parental strain T-ES174-1 resulted from integration of pES174 into genomic DNA by general recombination between the homologous segments [referred to as "flank" in (*1*) and (*13*)] of the plasmid and recipient cell. The second generation isolate used in this study, $G_{2(T-1)}$:20-5, is abbreviated L-G_2 for linked second generation. The unlinked duplication of the parental strain T-ES174-9 resulted by nonhomologous integration of a single copy of the same plasmid. The first generation isolate $G_{1(T-9)}$:3 is abbreviated U-G_1 for unlinked first generation.
4. E. B. Cambareri, M. J. Singer, E. U. Selker, unpublished results.
5. Strain K802; E. A. Raleigh and G. Wilson, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 9070 (1986).
6. Homoduplexes of both plasmids were stable in spreading solutions containing as much as 80% formamide. Above this concentration, regions of pEC24 began to melt. The melting temperature of DNA is reduced by ~0.7°C for each 1% increase in formamide, and by essentially the same amount for each 1% increase in mismatch [see R. W. Davis and R. W. Hyman, *J. Mol. Biol.* **62**, 287 (1971)]. If we assume that the bubbles observed in pEC24/pEC25 hybrids at 40% formamide reflect sequence divergence, the sequences must have diverged at least ~40% in these regions.
7. J. Marmur and P. Doty, *J. Mol. Biol.* **2**, 109 (1962).
8. D. K. Butler and R. L. Metzenberg, *Genetics*, in press.
9. B. K. Duncan and J. H. Miller, *Nature* **287**, 560 (1980).
10. C. Coulondre, J. H. Miller, P. J. Farabaugh, W. Gilbert, *Nature* **274**, 755 (1978).
11. D. Simon *et al.*, *ibid.* **304**, 275 (1983).
12. E. U. Selker and J. N. Stevens, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8114 (1985).
13. E. U. Selker, B. C. Jensen, G. A. Richardson, *Science* **238**, 48 (1987).
14. R. W. Davis, D. Botstein, J. R. Roth, *A Manual for Genetic Engineering: Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1980), pp. 188–191.
15. G. D. Fasman, Ed., *Handbook of Biochemistry and Molecular Biology*, vol. 2, *Nucleic Acids* (CRC Press, Cleveland, OH, ed. 3, 1976), p. 208.

16. We thank P. von Hippel and W. Rees for advice and assistance with the spectrophotometry, H. Foss for stimulating discussions, and L. Bruhn, H. Foss, D. Hawley, J. Selker, and W. Sistrom for comments on the manuscript. Supported by NIH gramt GM 35690 and NSF grant DCB 8718163.

28 December 1988; accepted 6 April 1989

Cloning of Breakpoints of a Chromosome Translocation Identifies the AN2 Locus

MANFRED GESSLER, KALLE O. J. SIMOLA, GAIL A. P. BRUNS

Chromosome translocations involving 11p13 have been associated with familial aniridia in two kindreds highlighting the chromosomal localization of the AN2 locus. This locus is also part of the WAGR complex (Wilms tumor, aniridia, genitourinary abnormalities, and mental retardation). In one kindred, the translocation is associated with a deletion, and probes for this region were used to identify and clone the breakpoints of the translocation in the second kindred. Comparison of phage restriction maps exclude the presence of any sizable deletion in this case. Sequences at the chromosome 11 breakpoint are conserved in multiple species, suggesting that the translocation falls within the AN2 gene.

THE PHENOTYPIC FEATURES OF THE WAGR syndrome (Wilms tumor, aniridia, genitourinary abnormalities, and mental retardation) indicate the presence of a constellation of genes important in human developmental processes within a small region of chromosome 11p13. Aniridia is a congenital developmental disorder of the eye, characterized by complete or partial absence of the iris, cataracts, lens anomalies, early onset glaucoma, and progressive loss of vision (*1*). It occurs in sporadic or autosomal dominant familial form without other phenotypic features at a frequency of 1:64,000 to 1:96,000. Many of the sporadic cases likely represent new mutations as aniridia is frequently inherited as an autosomal dominant in subsequent generations. Two genetic loci have been identified. In one large kindred, Ferrell *et al.* (*2*) observed linkage of aniridia (AN1) and the erythrocyte acid phosphatase (ACP1) locus on chromosome 2p. The presence of a second aniridia locus, the AN2 locus, on chromosome 11p is supported both by the consistent association of aniridia with the constitutional 11p13 deletions of the WAGR syndrome (*3*) and by the different inherited translocations with 11p13 breakpoints associated with familial aniridia in two kindreds (*4, 5*). The breakpoints of both translocations were localized by somatic cell hybrid analysis between the genes encoding the beta subunit of follicle-stimulating hormone (FSHB) and catalase (CAT) within the critical region of overlap of WAGR deletions (*6*). Mapping and cloning of chromosomal translocation breakpoints has been a powerful tool for characterization of *myc*-

M. Gessler and G. A. P. Bruns, Genetics Division, The Children's Hospital, and Department of Pediatrics, Harvard Medical School, Boston, MA 02115.
K. O. J. Simola, Department of Medical Genetics, University of Helsinki, Helsinki, Finland.

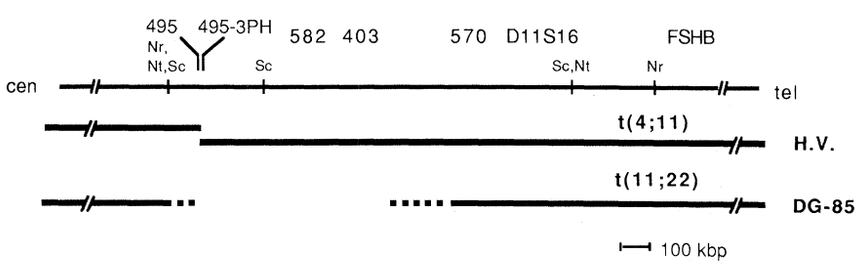

Fig. 1. Long-range chromosomal map around the aniridia locus. The location of probes used in this study and cutting sites for the enzymes Not I (Nt), Nru I (Nr) and Sac II (Sc) were determined by pulsed-field gel analysis (*12*). The position of the breakpoints for the t(4;11) translocation (H.V.) and the approximate size of the deletion associated with the t(11;22) translocation (DG-85) are indicated. The centromeric breakpoint of t(11;22) is likely between the Not I site and 495 as a rearrangement of the adjacent Not I fragment was not detected by Compton *et al.* in the DG-85 cell line (*16*). Probe 495 is located on the centromeric and probe p495-3PH on the telomeric side of the H.V. translocation breakpoint.