Volume 16 Number 9 1988 Nucleic Acids Research

Accuracy of intrachromosomal gene conversion in mouse cells

Janet L.Stachelek and R.Michael Liskay

Departments of Therapeutic Radiology and Human Genetics, Yale School of Medicine, 333 Cedar St, New Haven, CT 06510, USA

Received February 9, 1988; Revised and Accepted April 6, 1988

ABSTRACT

Results of several recent studies suggest that homologous recombination and related processes in mammalian cells are highly mutagenic. We have examined the products of intrachromosomal gene conversion events that encompassed the last intron of the chicken thymidine kinase gene. Following plasmid rescue and DNA sequencing, we find no mutations associated with twenty conversion events representing 5380 total base pairs of which 2414 base pairs are intron sequence. Based on these studies we conclude that intrachromosomal gene conversion in mouse cells is not a highly mutagenic process but rather it operates with fidelity.

INTRODUCTION

Several recent studies in mammalian cells suggest that homologous recombination and related processes might be more error-prone than previously believed. Of particular interest is a study by Thomas and Capecchi (1) indicating that gene targeting in mammalian cells can be highly mutagenic. In this study, about half of the events involved an error-prone process that the authors termed heteroduplex-induced mutagenesis. A more recent study suggests that not all gene targeting interactions are subject to such a high error rate (2). However, in this latter report the nature of the mutations to be corrected and the fact that targeting demanded gene function precluded detection of many error-prone events. Examination of heteroduplex repair in mammalian cells has revealed a highly error-prone process, at least in the pathways acting on extrachromosomal molecules (3-5). When preformed heteroduplexes are introduced by transfection or microinjection, repair is associated with a higher than expected mutation frequency, i.e., 2-10%. In another extrachromosomal study even a nick, a presumed early step in heteroduplex repair, induced a high frequency of nearby errors (6). In studies involving extrachromosomal sequences it is not entirely clear what fraction, if any, of these "high-frequency" errors are due to the

preparation of the input DNA, e.g. UV light exposure during gel isolation of the DNA, or the transfection process itself (7). An examination of the fidelity of chromosomal homologous recombination with a system that should allow detection of many errors has not been reported.

Our laboratory has been engaged in the study of homologous recombination between closely-linked gene pairs in the mammalian genome. In light of the above mentioned studies, we have examined products of intrachromosomal gene conversion like events at the nucleotide level to determine whether such events are subject to high rates of mutation. The studies to be described involve the rescue of converted chicken thymidine kinase (tk) genes and subsequent DNA sequencing in the vicinity of the conversion events. We have used an intrachromosomal recombination system designed to study conversion events that include an intron to enhance the recovery of any mutations that might accompany gene conversion.

MATERIALS AND METHODS

Enzymes and buffers

Restriction endonucleases and T4 DNA ligase (New England Biolabs), and the Sequenase kit (US Biochemicals) used for DNA sequence analyses were used according to the supplier's recommended conditions.

Preparation of plasmid DNA

Plasmids containing the four Xho I linker insertion mutations (XH34, XA98, XR14 and XA21) of the chicken tk gene were a generous gift from Dr. Jesse Kwoh (8). A double mutant chicken tk gene was constructed from the two mutants XH34 and XA98 and designated DX3498. This double mutant gene intended to serve as the recipient in gene conversion events was inserted into the Bam HI site of plasmid pJS-1 (9). Plasmids XR14 and XA21 were used to derive a double mutant gene DX1421. The 800 bp Xho I fragment from DX1421, containing wild-type chicken tk information that overlaps the two Xho I mutations in the recipient gene, was inserted into the unique Xho I site of the plasmid pV200 so as to position the fragment between LTR sequences from Moloney murine leukemia virus. The plasmid pV200 was kindly provided by Dr. Steven Goff. A 4.1 kb Eco RI fragment containing the LTR sequences flanking the 800 bp Xho I fragment was inserted into the Hind III site of pJS-1 using synthetic Hind III linkers (New England Biolabs). The final substrate can be seen in Figure 1.

Derivation of the experimental cell lines

The plasmid substrate was linearized by cleavage with Cla I and microinjected at a concentration of 0.5 ug/ml so that each mouse L cell

injected should receive 0.5 - 1 molecule per injection (10). Stable transformants were isolated following selection in G418 as previously described (11). Identification of cell lines containing a single copy of the recombination substrate, determination of conversion rates and isolation of tk^+ convertants were performed as previously described (13).

Rescue and sequencing of chicken tk sequences

The corrected gene from tk^+ convertants and the mutant recipient gene from two of the three tk^- parent lines were recovered by plasmid rescue as previously described (1). Briefly, 6 ug of genomic DNA were digested with Sac I, ligated at a concentration of 5ug/ml and used to transform the bacterial strain MH1 to neomycin resistance.

Plasmid DNA was prepared from minilysates as previously described (14) with the following modifications. Lysates were treated with one-fifth volume 5M lithium acetate, 0.1% SDS, 1mM EDTA and 100 ug/ml Proteinase K followed by phenol extraction and ethanol precipitation. Plasmid DNA was sequenced by a double-stranded DNA sequencing protocol and a U. S. Biochemicals Sequenase kit. The oligonucleotides used for sequencing were 18mers and complementary to nucleotides 1900-1917 and 2167-2184 according to the numbering system of Kwoh and Engler (15).

RESULTS AND DISCUSSION

We derived three lines of mouse L cells that each contain a single copy of the recombination substrate shown in Figure 1. The gene duplication consists of two chicken tk sequences oriented as direct repeats. One is a full-length gene containing two 8 bp oligonucleotide insertion mutations that create Xho I restriction sites. One mutation is located in the fifth exon and one in the sixth and last exon, at positions 1932 and 2163, respectively, according to the numbering of Kwoh and Engler (15). The other chicken tk sequence represents an 800 bp 3' portion of the wild-type chicken tk gene. This chicken tk sequence is flanked by LTR sequences of the Moloney murine leukemia virus. This truncated 3' portion of the chicken tk gene can act as a donor of information to convert the recipient, the gene containing the two Xho I mutations, to a functional (tk^+) sequence. Gene convertants were specifically recovered by selection in HAT medium which selects for tk^+ cells, and 400 ug/ml G418 to select for retention of the neo sequence (11). That these products are actually the result of gene conversion, that is, nonreciprocal transfer of information, rather than double reciprocal exchange has been reinforced by recent studies from our laboratory (12). Single reciprocal events between the two

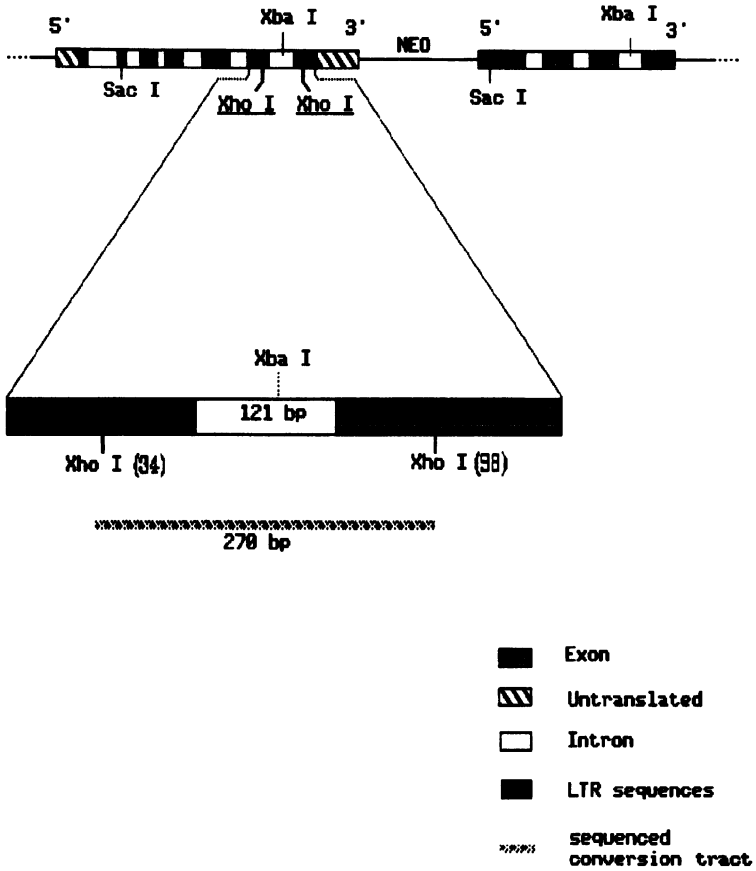

Figure 1 Gene conversion substrate and sequenced conversion tract.
The above duplication of chicken tk gene sequences when present in tk⁻ mouse L cells provides a substrate for intrachromosomal gene conversion events that span the last intron of the chicken tk gene. The chicken tk sequences are divided into intron and exon sequences as indicated. The chicken tk sequence on the left, intended to serve as the recipient in gene conversion, is a complete chicken tk gene containing two 8bp Xho I linker insertion mutations, 34 and 98, located in the fifth and sixth exon, respectively. Each mutation disrupts gene function and both must be removed to give a functional gene. The sequence on the right serves as donor in gene conversion and consists of a 3' portion of the chicken tk gene including most of the third exon and extending through the last exon. The donor acts to remove the two Xho I insertion mutations in the recipient, thus creating a functional tk gene and allowing growth in selective medium. The donor is bounded by LTR sequences from the Moloney murine leukemia virus. The portion of the substrate that was rescued lies between the two Sac I sites as shown. The enlargement of part of the recipient gene shows the region of sequenced conversion tract (270bp) that includes the last intron. The unique Xba I site marks the last intron of chicken tk sequences.

chicken tk genes can also result in the reconstruction of a functional tk gene, but these events result in the loss of the neo sequence (11, and unpublished data) and therefore were selected against by growth in G418-containing medium.

Twenty independent convertants were isolated by HAT-G418 selection. The full length tk gene was, as expected, resistant to Xho I digestion as confirmed by Southern blot hybridization analysis (data not shown). The convertants arose at a rate of 5×10^{-8}, a rate that is consistent with our previous studies for coconversion of two markers separated by 360 bp in the same gene (9). In that study we observed that a silent restriction site polymorphism between two mutant sites was consistently coconverted (23 of 23 events) with the two mutations, strongly suggesting that conversion involves transfer of contiguous blocks of genetic information (9). Based on these previous observations we argue that the conversion events responsible for the correction of the two Xho I mutations in the chicken tk gene must also include the last intron of the chicken tk gene (see Figure 1). Examination of the nucleotide sequence within the last intron of the converted gene should be most informative concerning the mutation rate associated with intrachromosomal gene conversion in mouse L cells, because, in general, introns are tolerant of many mutations (16). Also, two different 8bp oligonucleotide insertions into the last intron of the chicken tk gene have no effect on gene expression as measured by gene transfer efficiency (8). In addition, there do not appear to be any critical sequences within the sixth intron of the chicken tk gene because deletion of this intron has no effect on gene expression in mouse L cells (17).

As a note of interest, all recombinants examined by Southern blot hybridization and/or sequencing did contain the last intron of the recipient gene as indicated by sensitivity of the recipient gene to digestion with Xba I which cleaves uniquely in the last intron (data not shown). Because the "donor" sequence located within the LTRs (see Figure 1) is transcribed and the last intron is spliced in the parent lines 10 and 61 (unpublished data), we suggest that these conversion events did not involve a spliced RNA intermediate. Had any of the events proceeded via a spliced RNA, or cDNA, and assuming that the conversion events involved transfer of a contiguous block of information, as discussed above, the last intron of the recipient gene should have been removed by conversion.

The corrected gene from twenty tk^+ convertants and the double mutant recipient gene from two of the three tk^- parent lines (10 and 61) were

recovered by plasmid rescue. For each convertant, 270 bp of sequence which included the two sites of the Xho I mutations and 121 bp representing the last intron of the chicken tk gene, was determined. The sequence in this region for two of the three parent lines, numbers 10 and 61, was also determined following plasmid rescue. The results of the sequence analysis indicated that all sequences were in perfect agreement with the previously published sequence of the normal chicken tk gene (see Table 1). Therefore, no errors were introduced during these twenty conversion events. Assuming that a variety of mutations could be tolerated in the intron region of the chicken tk gene and that certain base substitutions in the exons, e.g. third base positions, should also allow gene function our results indicate that intrachromosomal gene conversion in mouse cells is <u>not</u> highly mutagenic.

Our results suggest accuracy in intrachromosomal recombination and are in sharp contrast with the gene targeting results of Thomas and Capecchi in which 50% of the homology-dependent interactions appeared to have resulted in mutations (1). Based upon DNA sequence analysis, the authors argued that the process specifically required a site of heteroduplex and the presence of direct repeats in close proximity to the region of heteroduplex. These two criteria for heteroduplex-induced mutagenesis are at least partially fulfilled in our studies. First, two potential

Table 1
Sequencing results

Parent line	# of conversions rescued	# of intron nt sequenced	# of total nt sequenced	# of mutations
10	12	1452	3240	0
61	6	720	1620	0
62	2	242	540	0
Totals	20	2414	5380	0

A summary of the results of sequenced gene conversions is shown. The three parent lines used to derive convertants, the number of conversions rescued from each parent line, the number of intron and total nucleotides sequenced and the number of mutations observed is presented. Note: Rescued genes from two parent lines, 10 and 61, were also sequenced in the relevant region and the sequences agreed with the previously published chicken tk sequence (15) with the exception that they contained as expected the two 8bp Xho I linker-insertion mutations 34 and 98. nt = nucleotides.

heteroduplex sites, the two Xho I mutations in the full-length gene, are present in our recombination substrate. Although it is difficult to rigorously demonstrate that heteroduplex repair was responsible for the gene conversion events studied here, other results from our laboratory are consistent with heteroduplex repair mediating intrachromosomal gene conversion in mouse cells (13). Secondly, inspection of the sequence within the intron of the chicken tk gene does reveal the presence of several short (5-6bp) direct repeats. However, the distances between the regions of heteroduplex and the direct repeats are greater in our study than was the case for Thomas and Capecchi. That this distance might be critical for heteroduplex-induced mutagenesis was suggested by their finding that mutations were observed only in the repeat closest to the mismatch (1). A second difference between our study and that of Thomas and Capecchi is that theirs is a targeting system involving one chromosomal and one extrachromosomal sequence, while our study involves conversion events between two sequences located on the same chromosome.

As discussed above, three studies have suggested that extrachromosomal repair of preformed heteroduplexes transfected into mammlian cells is highly error-prone (3-5). In one study a minimum of 2% of the heteroduplexes suffered error-prone repair manifested mainly by deletions and duplications (4), whereas in a second study 5-10% of the repair events appeared to have been error-prone (3). It seems likely that intrachromosomal recombination is more accurate than processes involving extrachromosomal DNA. That differences exist in terms of the rates (11, 18) and the homology requirements (10) for intrachromosomal versus extrachromosomal recombination is clear.

In conclusion, we report the first examination at the nucleotide level of the products of a gene conversion-like process acting on chromosomal sequences in mammalian cells. The inclusion of an intron in the conversion events should have allowed the recovery of many types of mutations within the conversion tract, including base substitutions, insertions, and deletions (of at least smaller size). The DNA sequence analysis of twenty independent conversion events representing 5380 bp of total sequence and 2414 bp of intron sequence did <u>not</u> reveal any mutations associated with the conversion events. This indicates an error rate per nucleotide of $< 1.8 \times 10^{-4}$ (or $< 4.0 \times 10^{-4}$ if only intron sequences are considered). Therefore, we conclude that intrachromosomal gene conversion in mouse cells operates with greater fidelity than previous related studies would have predicted.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Jesse Kwoh for supplying us with plasmids and for useful discussions on plasmid constructions. We would also like to thank Dr. Kirk Thomas for supplying the bacterial strain MH-1 and for his valuable advice on plasmid rescue procedures, and Dr. Michael Seidman for his advice on the preparation of DNA for sequencing. We would also like to thank Dr. Alan Waldman, Roni Bollag and Alan Godwin for their critical reading of the manuscript. This work was supported by a Public Health Service Grant RO1 GM32741 to RML. RML is a recipient of a Leukemia Society of America Scholar Award.

REFERENCES

1. Thomas, K.R. and Capecchi, M.R. (1986) Nature 324, 34-38.
2. Song, K.-Y. Schwartz, F., Maeda, N., Smithies, O. and Kucherlapati, R. (1987) Proc. Natl. Acad. Sci. USA 84, 6820-6824.
3. Folger, K.R., Thomas, K.R. and Capecchi, M.R. (1985) Mol. Cell. Biol. 5, 70-74.
4. Weiss, U. and Wilson, J.H. (1988) Nucl. Acids Res. (in press).
5. Ayares, D., Ganea, D., Chekuri, L., Campbell, C. R. and Kucherlapati, R. (1987) Mol. Cell. Biol. 7, 1656-1662.
6. Seidman, M. Bredberg, A., Seetharam, S. and Kraemer, K. H. (1987) Proc. Natl. Acad. Sci. USA 84, 4944-4948.
7. Calos, M. P., Lebkowski, J. S. and Botchan, M. R. (1983) Proc. Natl. Acad. Sci., USA 80, 3015-3019.
8. Kwoh, T.J., Zipser, D. and Wigler, M. (1983) J. Molec. Appl. Genet. 2, 191-200.
9. Liskay, R.M. and Stachelek, J.L. (1986) Proc. Natl. Acad. Sci. USA 83, 1802-1806.
10. Waldman, A.S. and Liskay, R.M. (1987) Proc. Natl. Acad. Sci. USA 84 5340-5344.
11. Letsou, A. and Liskay, R.M. (1986) In: Gene Transfer, 383-409. (ed., R. Kucherlapati) Plenum Publishing, New York.
12. Bollag, R.J. and Liskay, R.M. (1988) Genetics 119 (in press).
13. Letsou, A. and Liskay, R.M. (1987) Genetics 117, 759-769.
14. Birnbaum, H.C. and Doly, J.C. (1979) Nucl. Acids Res. 7,1513-1523.
15. Kwoh, T. J. and Engler, J. (1984) Nucl. Acids Res. 12, 3959-3971.
16. Sharp, P. A. (1987) Science 235, 766-771.
17. Gross, M. K., Kainz, M. S. and Merrill, G. F. (1987) Mol. Cell. Biol. 7:4576-4581.
18. Kucherlapati, R. (1986) In: Gene Transfer, 363-381. (ed., R. Kucherlapati) Plenum Publishing, New York.