

DIR: a novel DNA rearrangement associated with inverted repeats

David J. Pinder, Catherine E. Blake and David R.F. Leach*

Institute of Cell and Molecular Biology, University of Edinburgh, Kings Buildings, Edinburgh EH9 3JR, UK

Received October 16, 1996; Revised and Accepted November 18, 1996

ABSTRACT

A novel DNA rearrangement has been characterized that is both a direct and inverted repeat. This rearrangement involves the 2-fold duplication of a plasmid sequence adjacent to the site of insertion of a long palindrome. The sequence of this rearrangement suggests that it has arisen by strand slippage from the leading to the lagging strand of the replication fork as a consequence of the presence of the long palindrome.

INTRODUCTION

Genomes are not static, they have both the necessary levels of fidelity for stable inheritance and the required flexibility for genetic change. A primary mechanism of genetic change is recombination. Genome rearrangements in both prokaryotes and eukaryotes may result from 'homologous recombination', between long homologous sequences (1–3), as well as from 'illegitimate recombination', between sequences of little or no homology (2,4).

Illegitimate recombination mechanisms are important in evolution (5,6), biotechnology (7,8) and have been suggested to play a role in repeated sequence instability associated with certain human genetic disorders (9–11) and colorectal carcinomas (12,13). Illegitimate recombination may be divided into two broad classes: (i) end joining, which is mediated by enzymes which cut and join DNA, such as topoisomerases, site-specific DNases and proteins which initiate rolling circle replication; (ii) strand slippage, where, after pausing at the replication fork, the nascent strand can dissociate from one template and pair with another. Strand slippage can occur in association with short homologous sequences, but, unlike homologous recombination, this process is RecA independent both in plasmids (14–16) and the chromosome of *Escherichia coli* (15).

Sequences which have the potential to form unusual secondary structures are known to promote pausing of DNA replication and can stimulate replication slippage (17). Such sequences include: (i) oligopurine–oligopyrimidine sequences, able to adopt a triple helical structure in supercoiled DNA, both *in vivo* (18) and *in vitro* (19); (ii) tri- and tetranucleotide repeats, thought to form pseudo-hairpins when single-stranded; (iii) alternating GC or GT dinucleotide repeats, capable of forming Z-DNA; (iv) DNA palindromes (inverted repeats), able to form hairpins, when single-stranded, or cruciforms, when double-stranded.

Long DNA palindromes are unstable in both eukaryotes and prokaryotes (20,21) and have been shown to halt the progress of the replication fork *in vitro* (22,23). DNA palindromes longer than 150–200 bp cannot be cloned in wild-type *E.coli*. Either the replicon containing the palindrome is so poorly replicated that it is inviable (24,25) or the palindrome is so unstable that it suffers partial or complete deletion (26,27). Many studies of deletion stimulated by palindromic sequences have shown that their endpoints tend to occur in short direct repeats (21,32–35). *Escherichia coli sbcCD* mutants are known to increase the viability of replicons containing long palindromes (28–31) and this results in the stabilization of these sequences within a population.

In this work we have isolated a novel mutation composed of both direct and inverted repeats that we call DIR (direct and inverted repeat). This rearrangement is associated with loss of a 571 bp palindrome inserted at the *EcoRI* site in the multicloning region of plasmid pUC18. Its structure and origin suggest that it may have arisen by strand slippage across the replication fork.

MATERIALS AND METHODS

Medium and bacterial strains

AXI medium is LB agar supplemented with 100 µg/ml ampicillin, 64 µg/ml X-gal and 30 µg/ml IPTG. Plasmids were maintained in *E.coli* strains DL733 (JM83 $\Delta sbcCD::Km^R$) or JM83 (36) for sequencing of the DIR mutation. The 571 bp palindrome used in this work is a perfect inverted repeat of 228 bp separated by 15 bp (37). It is composed of non-essential bacteriophage λ DNA and was obtained from DRL 116 ($\lambda \Delta B pal spi6 cI857$).

Plasmid DNA

The plasmid vector used was pUC18 (36). The construction of pAC2, pMS7 and pDIR1 is described in Results. Plasmid DNA was isolated using the Qiagen Plasmid Midi Kit (Qiagen Inc.) except for the preparation of multimers, when DNA was isolated using the alkaline SDS method of Birnboim and Doly (38).

Radiolabelling of plasmid DNA fragments

Aliquots of 4 µg plasmid DNA were digested to completion with 20 U *EcoRI* (Boehringer Mannheim) and radiolabelled using 10 µCi [α - ^{35}S]dATP (~600 Ci/mmol) and 1 U Klenow enzyme (Boehringer Mannheim). The DNA was purified using the

* To whom correspondence should be addressed. Tel: +44 131 650 5373; Fax: +44 131 650 8650; Email: d.leach@ed.ac.uk

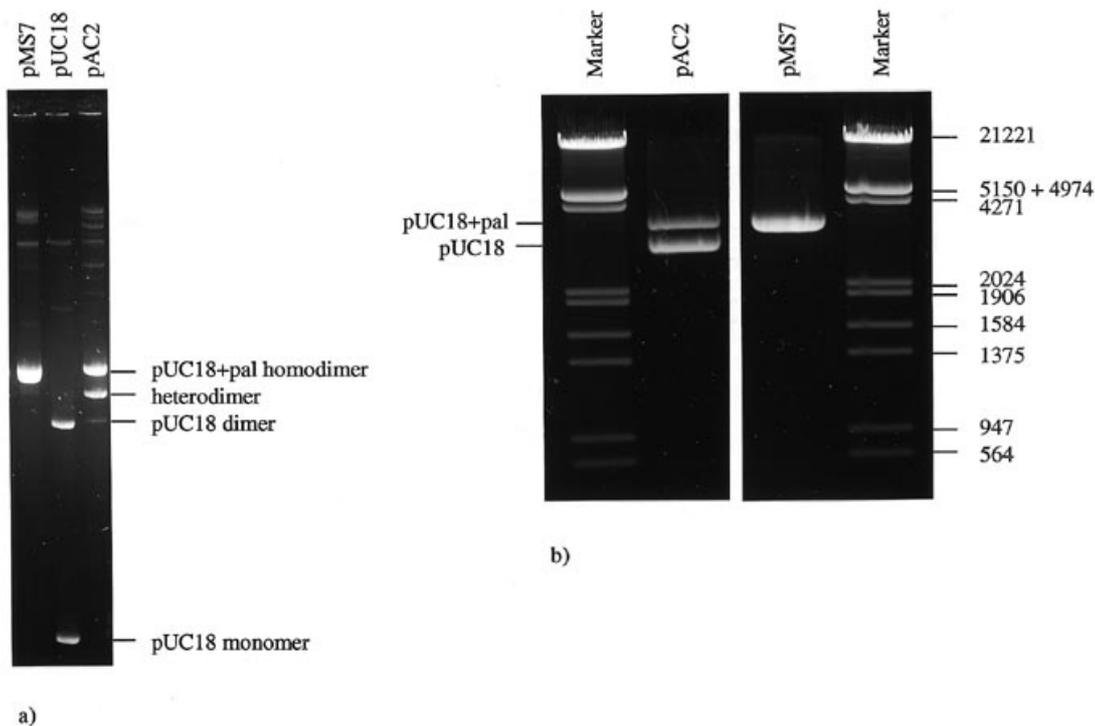


Figure 1. (a) Agarose gel electrophoresis of plasmids pMS7, pUC18 and pAC2. The ratio of heterodimer to homodimer present in the population of pAC2 DNA was variable from one preparation to another. (b) *Hind*III restriction digests of plasmids pAC2 and pMS7.

QIAquick™ Nucleotide Removal Kit (Qiagen Inc.). Samples of 2.5 µg purified DNA were digested with 10 U *Bam*HI (Boehringer Mannheim) and purified once more using a QIAquick™ Nucleotide Removal Kit. ³²P-End-labelled marker V (Boehringer Mannheim) and marker 8-32 (Pharmacia) were used as size markers.

Polyacrylamide gel electrophoresis

Radiolabelled samples were denatured by boiling and electrophoresed on a 10% Longranger™ polyacrylamide gel (AT Biochem) containing 7 M urea. The gel was run at 55°C to prevent hairpin formation. After drying, the gel was autoradiographed.

DNA sequencing

Sequencing was carried out using the Sanger technique (39). Two primers were used: primer 5'-GACTGGAAAGCGGGCA-3' (596L) was manufactured by Perkin Elmer Econopure™ and primer 5'-GTTTTCCAGTCACGAC-3' (-40) was supplied with the Sequenase® v2.0 Sequencing Kit (US Biochemical). DNA was sequenced using the Sequenase® v2.0 Sequencing Kit. DMSO sequencing was carried out with the Sequenase® v2.0 Sequencing Kit with the addition of 10% DMSO (Sigma Chemical Co.) to the reaction mix and to the GATC termination mixes. Samples were denatured by boiling and electrophoresed on a 6% Longranger™ polyacrylamide gel (AT Biochem) containing 7 M urea. The gel was run at 50°C and autoradiographed after drying.

Agarose gel electrophoresis

A sample of 0.5–1 µg plasmid DNA was digested to completion with 10 U *Hind*III or *Pvu*II (Boehringer Mannheim). Electrophoresis of undigested plasmid DNA was carried out on 1%

agarose gels (Flowgen) at 2 V/cm overnight. *Hind*III digests were carried out on 1% agarose gels (Flowgen) at 5 V/cm. Electrophoresis of *Pvu*II digests was performed on 2% NuSieve® 3:1 (Flowgen) agarose at 4 V/cm. Gels were stained with ethidium bromide (0.5 µg/ml) and photographed using GRAB-IT™ (UVP Inc.) or Polaroid 667 film. A *Hind*III/*Eco*RI digest of λ DNA or Marker VI (Boehringer Mannheim) was used as a size standard. In preparation for sequencing, the *Pvu*II fragment containing the DIR insert was purified by electrophoresis, excised from the gel and purified using a QIAquick™ Gel Extraction Kit (Qiagen Inc.).

RESULTS

Construction and characterization of pAC2 and pMS7

Plasmid pAC2 was constructed by ligating a 571 bp DNA palindrome (37) into the *Eco*RI site of pUC18. When the *E. coli* strain DL733 ($\Delta sbcCD::Km^R$) was transformed with pAC2 DNA the colonies obtained were observed to have an unusual phenotype on AXI medium (white with blue sectors). This phenotype suggested some genetic instability of pAC2, blue sectors being composed of cells able to hydrolyse X-gal and white sectors being composed of cells unable to do so. Plasmid pMS7 was isolated by plating DL733 containing pAC2 on AXI medium and looking for rare white clones.

Plasmid DNA of pAC2 and pMS7 were compared by agarose gel electrophoresis. The analysis of uncut pAC2 DNA revealed a population of different multimeric forms, including a species whose migration suggested the existence of a heterodimer in which only one copy of the palindrome was present (Fig. 1a). Uncut pMS7 DNA was primarily a dimer of the palindrome-containing DNA (Fig. 1a). This interpretation was supported by the DNA restriction pattern, after digestion with *Hind*III (Fig. 1b).

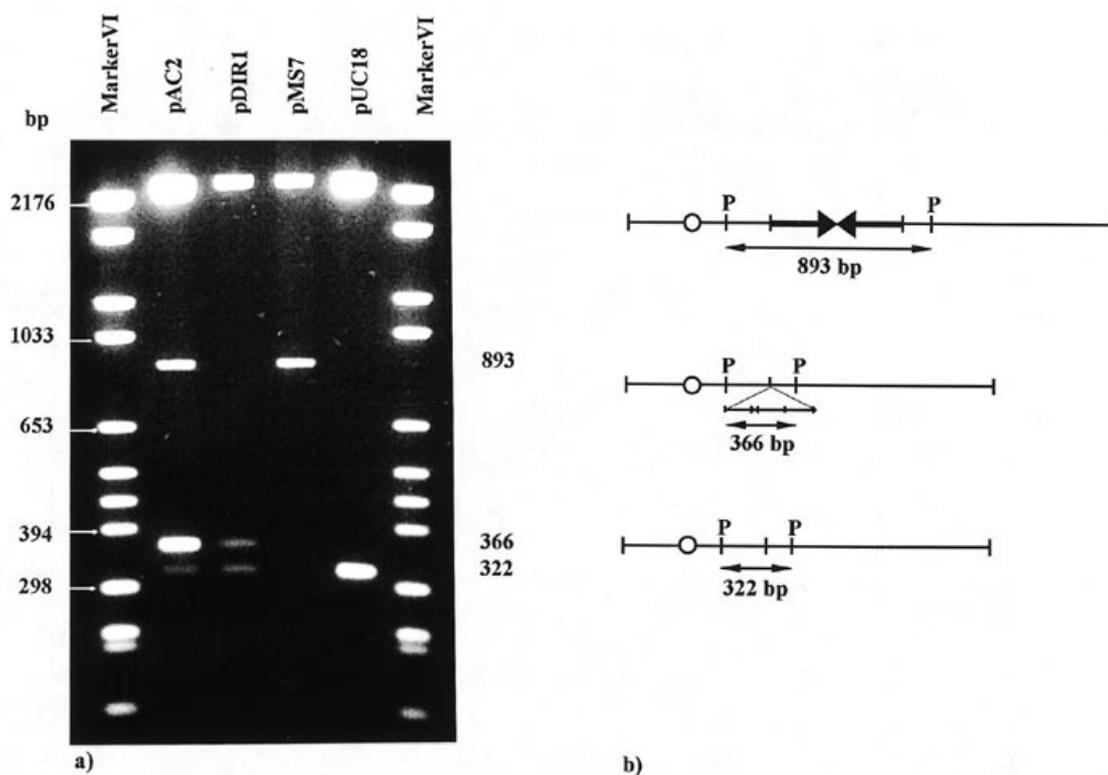


Figure 2. (a) *PvuII* restriction digests of pAC2, pDIR1, pMS7 and pUC18. (b) Maps of the monomeric components of the plasmids analysed by restriction with *PvuII*. pAC2 contains major proportions of the 893 bp fragment containing the 571 bp palindrome and the 366 bp fragment containing the DIR. It also contains a minor proportion of the 322 bp fragment characteristic of pUC18. pDIR1 contains equal proportions of the 366 bp DIR fragment and the 322 bp pUC18 fragment. pMS7 contains only the 893 bp palindrome-containing fragment. pUC18 contains only the 322 bp *PvuII* fragment.

Recircularization experiments

pAC2 and pMS7 were recircularized by digesting the plasmids with *BsrFI*, religating the ends and transforming DL733 ($\Delta sbcCD$). The results are shown in Table 1. These results indicate a significant increase in the yield of highly sectoring colonies derived from pMS7. Plasmid DNA was isolated from a number of these sectoring clones (derived from both pAC2 and pMS7) and restriction analysis revealed that although they had the sectoring phenotype of pAC2, they did not contain the 571 bp DNA palindrome. One such plasmid, denoted pDIR1 (derived from the monomerization of pAC2), was chosen for further detailed analysis.

Analysis of pDIR1 by agarose gel electrophoresis

Agarose gel electrophoresis of uncut pDIR1 DNA demonstrated that the simplest form co-migrated with dimers of pUC18 (data not shown). However, restriction analysis with *PvuII* suggested that the plasmid was in fact predominantly a heterodimer of pUC18 and pUC18 with an insert of ~40 nt. Furthermore, this analysis suggested that the heterodimer present in pAC2 DNA contained the 571 bp palindrome on one side and this same ~40 nt insert on the other (see Fig. 2). This also confirmed that pMS7 was primarily present as homomultimers of palindrome-containing DNA.

Table 1. Percentage total colony type, after digestion with *BsrFI* and ligation

Plasmid	%Total		
	Blue colonies	Sectoring colonies 'pAC2 like'	White colonies
pAC2 cut with <i>BsrFI</i> +ligase	5.5	86	8.5
pMS7 cut with <i>BsrFI</i> +ligase	0	8.5	91.5
pAC2 cut with <i>BsrFI</i> -ligase	9	75	15
pMS7 cut with <i>BsrFI</i> -ligase	0	0	0
pAC2 uncut	10.5	80	9.5
pMS7 uncut	0	1	99

Fine structure restriction analysis of pDIR1 by polyacrylamide gel electrophoresis

pAC2, pDIR1, pMS7 and pUC18 were digested with *EcoRI* and radiolabelled using [α - ^{35}S]dATP and Klenow enzyme. A sample of the purified DNA was digested with *BamHI*, run on a 10% LongrangerTM polyacrylamide gel and the bands visualized by autoradiography after the gel had dried (Fig. 3). The restriction pattern obtained confirmed the presence of a second *EcoRI* site in the multicloning region, giving a 46 bp fragment (42 bp plus

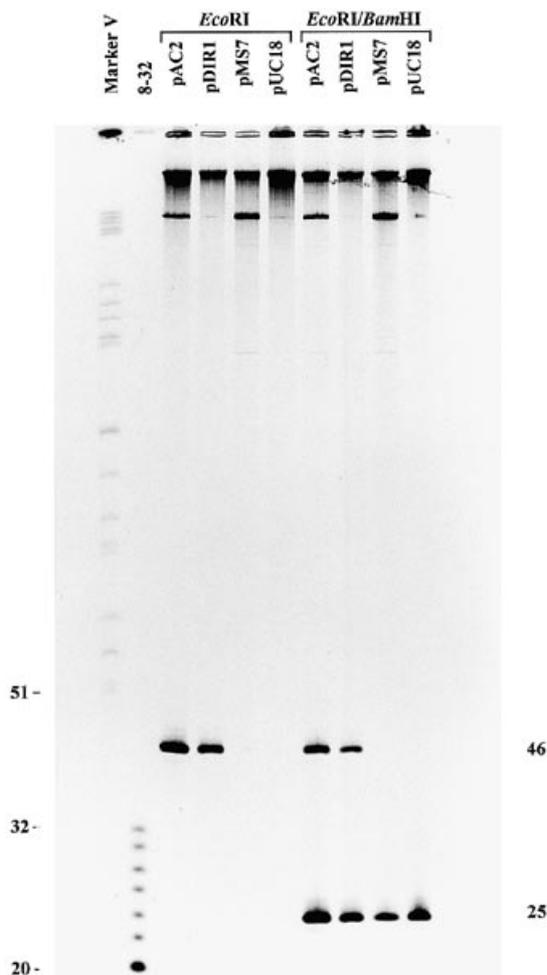


Figure 3. Autoradiograph of *EcoRI* single digests and *EcoRI/BamHI* double digests of plasmids pAC2, pDIR1, pMS7 and pUC18. The bands at 46 nt in the *EcoRI/BamHI* double digests are products of partial digestion by *BamHI*.

4 bp derived from filling in the *EcoRI* site overhang with Klenow enzyme) in both pAC2 and pDIR1. pMS7 and pUC18 do not have this second *EcoRI* site. In contrast, all four plasmids have the correct 25 bp (21 plus 4 bp) *EcoRI-BamHI* fragment.

Sequencing of the 42 bp insert present in pDIR1

DNA sequencing of the polylinker region of pDIR1, pUC18 and pAC2 was carried out on both strands using two primers: 596L and -40 (see Materials and Methods). Since pDIR1 and pAC2 were found to contain heteromultimers, sequencing of the 42 bp insert required its separation from the other component of the multimer. This was done in two ways. Firstly, the *PvuII* fragment containing the insert was purified from the other parts of the multimer by agarose gel electrophoresis and sequencing was performed on the DNA fragment isolated from the gel. Secondly, pDIR1 DNA was used to transform JM83 and a homomultimer containing the 42 bp insert was isolated that could be sequenced directly. Normal 'Sanger' DNA sequencing revealed a region in the multicloning site where the DNA polymerase seemed to stall.

This suggested the presence of secondary structure. Sequencing in the presence of 10% DMSO improved the ability of the polymerase to sequence through this region and revealed the presence of a novel type of mutation (see Fig. 4). The sequence shows the presence of a duplicated *EcoRI-BamHI* multicloning site fragment with the same region inverted in between the duplication. We call this a 'DIR' structure, since it is composed of both direct and inverted repeats.

DISCUSSION

A 571 bp imperfect palindrome of bacteriophage λ DNA was introduced into the cloning vector pUC18 to form plasmid pAC2 and an unstable phenotype was observed. It was initially thought that the unstable phenotype of pAC2 might be due to deletion of the 571 bp palindrome. Deletion of the palindrome would allow production of a functional α subunit of β -galactosidase. *Escherichia coli* cells harbouring the plasmid would have a dark blue colour (the result of hydrolysis of X-gal, a chromogenic substrate) and this would generate a sectoring clone. However, this proved not to be the case. A derivative of pAC2 was isolated that showed almost no sectoring and this plasmid (pMS7) still contained the palindrome. Furthermore, when pAC2 and pMS7 were monomerized sectoring clones were discovered that did not contain the 571 bp DNA palindrome. One of these clones was analysed in detail and was found to contain an inverted and directly duplicated region of the plasmid polylinker adjacent to the site of insertion of the palindrome, but the palindrome itself was no longer present. We have called this structure a DIR for 'direct and inverted repeat'. The same DIR structure was also found to be present in pAC2, which is a population of DNAs including a heteromultimer composed of one monomeric subunit containing the palindrome and the other containing the DIR structure. The stabilized plasmid pMS7 was found to be primarily a homomultimer of the palindrome-containing subunit. The unstable (sectoring) derivatives of pAC2 and pMS7 that had lost the 571 bp palindrome were all found to be heteromultimers with one subunit containing the DIR structure and the other containing wild-type pUC18 sequence. Small amounts of a DNA fragment of the predicted size of pUC18 were also seen in the DNA of pAC2. Since all these plasmids were maintained in Rec⁺ cells, interconversion between various forms was occurring to generate subpopulations of plasmids. The results shown in Table 1 are consistent with heterogeneous populations of plasmids in the DNA preparations of pAC2 and pMS7 and the isolation of subpopulations by forcing monomerization. Alternatively, monomerization may stimulate rearrangement and future work will be needed to distinguish between these two possibilities. The plasmids investigated here are shown diagrammatically in Figure 5.

Attention was focused on the nature of the DIR mutation. This DNA sequence was found to be composed of a direct repeat of the *EcoRI-BamHI* portion of the pUC18 polylinker separated by an inverted repeat of the same DNA sequence. This structure can be viewed as a palindrome and a half, with the inverted *BamHI-EcoRI* fragment possessing the potential to form a hairpin structure with one of the direct repeats on either side. The ability to form such a structure would in theory make the sequence unstable and able to be deleted by illegitimate recombination using the direct repeats.

These observations support the hypothesis that the sectoring phenotype is due to the presence of the DIR mutation in both

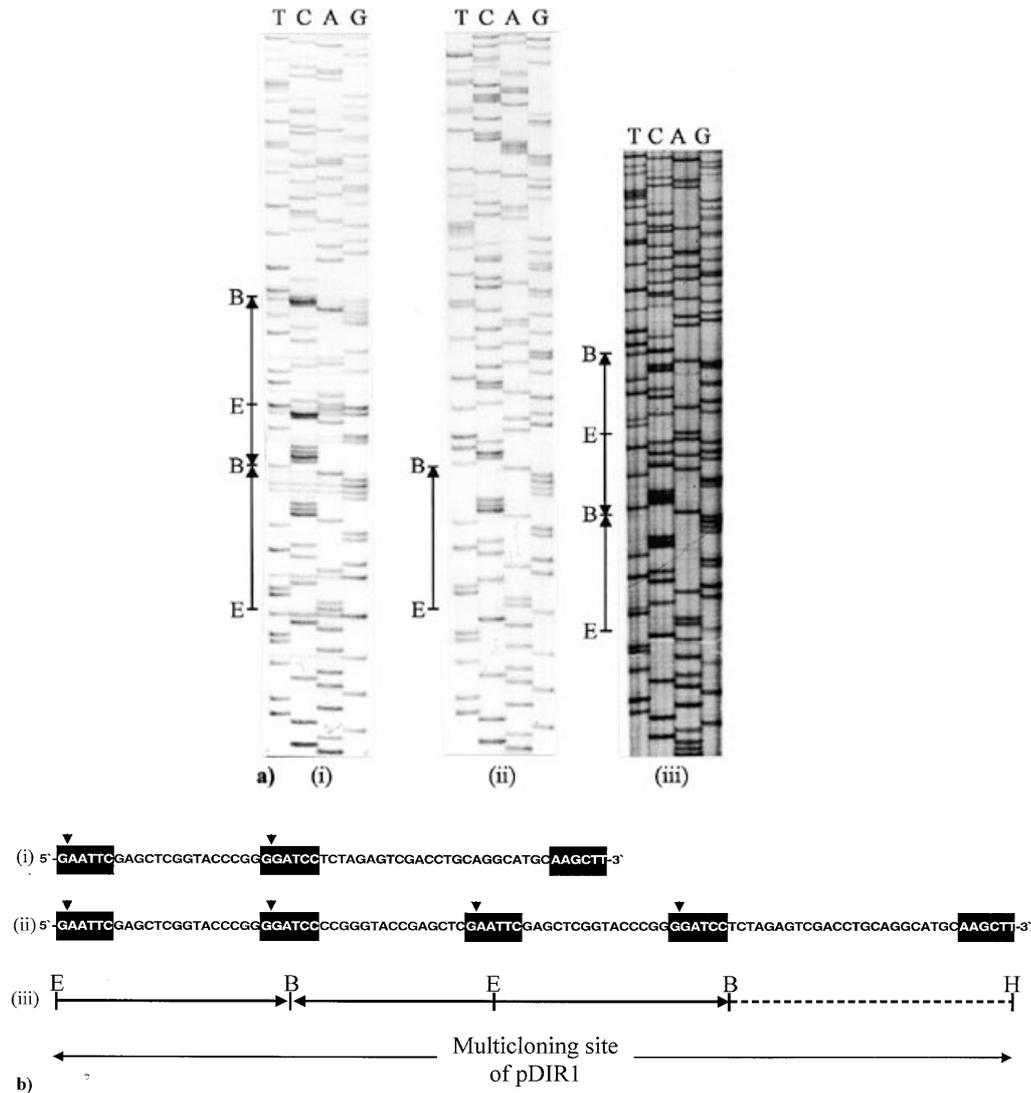


Figure 4. (a) Autoradiograph of the DIR sequence. The data shown were obtained using primer 596L on pUC18 and on the homomultimer substrate derived from pDIR1 after transformation of JM83. (i) DIR sequence using the standard sequencing method. (ii) pUC18 sequence using the standard method. (iii) DIR sequence using the DMSO method to minimize pausing of Sequenase[®] at regions of secondary structure. The complete sequence of DIR could not be obtained on any individual gel but all ambiguities could be resolved using the DNA sequence obtained from the -40 primer (data not shown). (b) Sequence of the pUC18 multicloning site and DIR. (i) Sequence of the multicloning site of pUC18. (ii) Sequence of the multicloning site in pDIR1. (iii) Diagrammatic representation of DIR.

pAC2 and pDIR1. It is possible that deletion of the 571 bp DNA palindrome is responsible for some of the sectors in clones containing pAC2, but this deletion event appears to occur less frequently than deletion of the DIR structure.

Restriction analysis of the products of the monomerization experiment showed that it is possible to generate the DIR structure not only from pAC2 but also from pMS7. This argues that the DIR mutation arises from a precursor containing the 571 bp palindrome. This is also consistent with the interpretation that the origin of the DIR mutation in pAC2 was a consequence of the original cloning of the 571 bp palindrome in pUC18. The structure and origin of the DIR mutation are most easily explained if the 571 bp palindrome facilitates strand slippage across the replication fork from the leading to the lagging strand, as shown in Figure 6. We suggest that the first step is the rare extrusion of the 571 bp palindrome to form a stable cruciform by intrastrand base pairing. This causes the DNA polymerase to pause and

facilitates replication across the fork from the leading to the lagging strand. A new palindromic strand is thus formed which can fold on itself to form a hairpin and leading strand synthesis can resume by recopying the already duplicated region. A further strand slippage event across the base of the extruded 571 bp palindrome using the *EcoRI* sites bounding it as direct repeats leads to loss of the original palindrome and completion of a DNA strand containing only the DIR sequence.

It is not clear how important this type of strand slippage event has been in genome evolution. However, we have observed that several intergenic repeated units (IRUs; 40), also known as enterobacterial repeated intergenic consensus sequences (ERICs; 41), have extended structures that include sequences duplicated in direct and inverted orientations (G.Cromie, J.Collins and D.Leach, unpublished data). The organization of these extended IRU/ERIC sequences is more complex than can be accounted for by the simple mechanism envisaged for the formation of the DIR

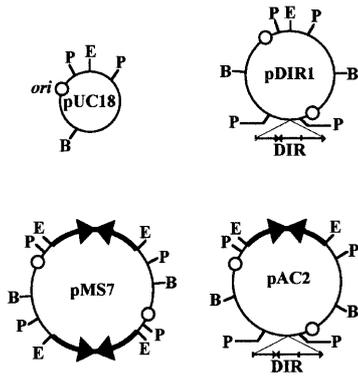


Figure 5. Diagrammatic representation of plasmids pUC18, pDIR1, pMS7 and pAC2. The origin of DNA replication is represented by an open circle; restriction sites for *EcoRI*, *BsrFI* and *PvuII* are represented by E, B and P respectively; the 571 bp palindrome is represented by arrows in bold.

structure, but could have arisen from more complex strand slippage events. Computational searching for DIR structures in genomes will be needed to determine their frequency, distribution and association with other DNA sequences.

ACKNOWLEDGEMENTS

We thank Alison Chalker for construction of plasmid pAC2, Morgan Shaw for isolation of pMS7 and Thorsten Allers for photography. Funding for this work has come from the BBSRC, the MRC and NBL Gene Sciences Ltd.

REFERENCES

- 1 Kowalczykowski, S.C., Dixon, D.D., Eggleston, A.K., Lauder, S.D. and Rehrauer, W.M. (1994) *Microbiol. Rev.*, **58**, 401–465.
- 2 Leach, D.R.F. (1996) *Genetic Recombination*. Blackwell Science, Oxford, UK.
- 3 West, S.C. (1992) *Annu. Rev. Biochem.*, **61**, 603–640.
- 4 Ehrlich, S.D., Bierne, H., d'Alenon, E., Villette, D., Petranovic, M., Noirot, P. and Michel, B. (1993) *Gene*, **135**, 161–166.
- 5 Anderson, R.P. and Roth, J.R. (1977) *Annu. Rev. Microbiol.*, **31**, 473–505.
- 6 Ohno, S. (1970) *Evolution by Gene Duplication*. Springer-Verlag, New York, NY.
- 7 Ehrlich, S.D., Noirot, P., Petit, M.A., Janniere, L., Michel, B. and te Riele, H. (1986) In Setlow, J.K. and Hollander, A. (eds), *Genetic Engineering*. Plenum, New York, NY, Vol. 8, pp. 71–83.
- 8 Summers, D.K. (1994) *Dev. Biol. Stand.*, **83**, 7–11.
- 9 Richards, R.I. and Sutherland, G.R. (1992) *Cell*, **70**, 709–712.
- 10 Wells, R.D. and Sinden, R.R. (1993) In Davies, D. and Warren, S. (eds), *Genome Analysis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, Vol. 7, pp. 107–138.
- 11 Wells, R.D. (1996) *J. Biol. Chem.*, **271**, 2875–2878.
- 12 Thibodeau, S.N., Bren, G. and Schaid, D. (1993) *Science*, **260**, 816–819.
- 13 Ionov, Y., Peinado, M.A., Malkhosyan, S., Shibata, D. and Perucho, M. (1993) *Nature*, **363**, 558–561.
- 14 Matfield, M., Badawi, R. and Brammer, W.J. (1985) *Mol. Gen. Genet.*, **199**, 518–523.
- 15 Lovett, S.T., Drapkin, P.T., Sutera, V.A.J. and Gluckman-Peskind, T.J. (1993) *Genetics*, **135**, 631–642.
- 16 Bi, X. and Liu, L.F. (1994) *J. Mol. Biol.*, **235**, 414–423.
- 17 Sinden, R.R. and Wells, R.D. (1992) *Curr. Opin. Biotechnol.*, **3**, 612–622.
- 18 Rao, B.S., Manor, H. and Martin, R.G. (1988) *Nucleic Acids Res.*, **16**, 8077–8094.
- 19 Dayn, A., Samadashwily, G.M. and Mirkin, S.M. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 11406–11410.
- 20 Leach, D.R.F. (1994) *BioEssays*, **16**, 893–900.
- 21 Collick, A., Drew, J., Penberth, J., Bois, P., Luckett, J., Scaerou, F., Jeffreys, A. and Reik, W. (1996) *EMBO J.*, **15**, 1163–1171.

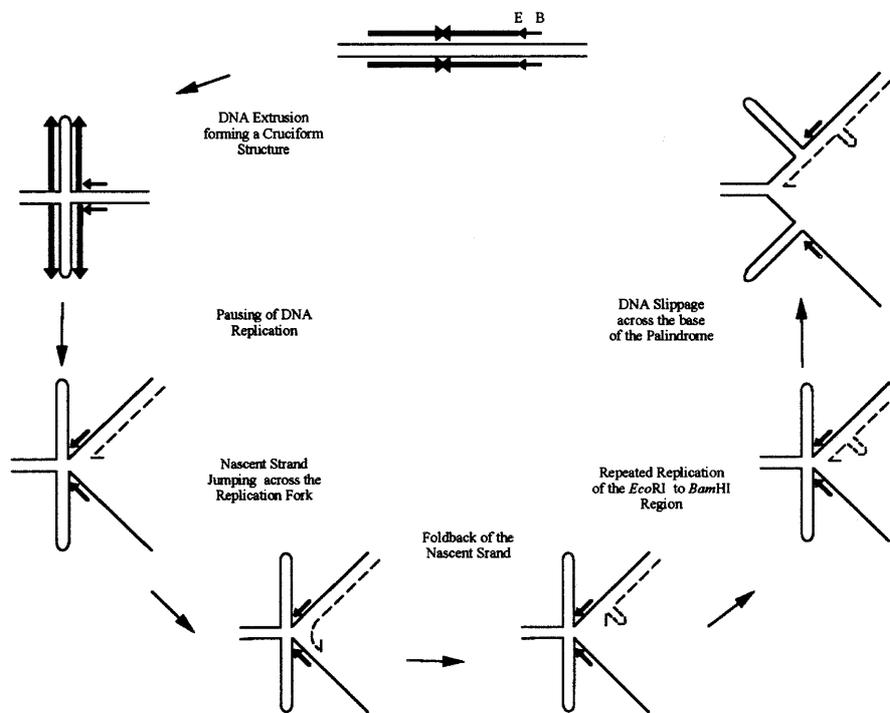


Figure 6. Model for the formation of DIR. The 571 bp palindrome is represented by long bold arrows; the *EcoRI*–*BamHI* fragment that is duplicated in DIR is represented by a short arrow; the nascent leading strand is represented by a dotted line. See text for further details.

- 22 LaDuca,R.J., Fay,P.J., Chuang,C., McHenry,C.S. and Bambara,R.A. (1983) *Biochemistry*, **22**, 5177–5188.
- 23 Bedinger,P., Munn,M. and Alberts,B.M. (1989) *J. Biol. Chem.*, **264**, 16880–16886.
- 24 Warren,G.J. and Green,R.L. (1985) *J. Bacteriol.*, **161**, 1103–1111.
- 25 Yoshimura,H., Yoshino,T., Hirose,T., Nakamura,Y., Higashi,M., Hase,T., Yamaguchi,K., Hirokawa,H. and Masamune,Y. (1986) *J. Gen. Appl. Microbiol.*, **32**, 393–404.
- 26 Collins,J. (1981) *Cold Spring Harbor Symp. Quant. Biol.*, **45**, 409–416.
- 27 Hagan,C.E. and Warren,G.J. (1983) *Gene*, **24**, 317–326.
- 28 Collins,J., Volkaert,G. and Nevers,P. (1982) *Gene*, **19**, 139–146.
- 29 Leach,D.R.F. and Stahl,F.W. (1983) *Nature*, **305**, 448–451.
- 30 Chalker,A.F., Leach,D.R.F. and Lloyd,R.G. (1988) *Gene*, **71**, 201–205.
- 31 Gibson,F.P., Leach,D.R.F. and Lloyd,R.G. (1992) *J. Bacteriol.* **174**, 1222–1228.
- 32 Weston-Hafer,C. and Berg,D.E. (1991) *Genetics*, **127**, 649–655.
- 33 Peeters,B.P.H., deBoer,J.H., Bron,S. and Venema,G. (1988) *Mol. Gen. Genet.*, **212**, 450–458.
- 34 Bichara,M., Schumacher,S. and Fuchs,R.P.P. (1995) *Genetics*, **140**, 897–907.
- 35 Das Gupta,U., Weston-Hafer,K. and Berg,D.E. (1987) *Genetics*, **115**, 41–49.
- 36 Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103–119.
- 37 Lindsey,J.C. and Leach,D.R.F. (1989) *J. Mol. Biol.*, **206**, 779–782.
- 38 Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513–1523.
- 39 Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
- 40 Sharples,G.J. and Lloyd,R.G. (1990) *Nucleic Acids Res.*, **18**, 6503–6508.
- 41 Houlton,C.S.J., Higgins,C.F. and Sharp,P.M. (1991) *Mol. Microbiol.*, **5**, 825–834.