J Mol Evol (1994) 38:352–362

Journal of Molecular Evolution

© Springer-Verlag New York Inc. 1994

Junctions Between Repetitive DNAs on the PSR Chromosome of *Nasonia vitripennis*: Association of Palindromes with Recombination

Kent M. Reed,* Leo W. Beukeboom, Danna G. Eickbush, John H. Werren**

Department of Biology, University of Rochester, Rochester, NY 14627, USA

Received: 14 May 1993 / Revised: 23 July 1993

Abstract. The Paternal-Sex-Ratio (PSR) chromosome of *Nasonia vitripennis* contains several families of repetitive DNAs that show significant sequence divergence but share two palindromic regions. This study reports on the analysis of junctions between two of these repetitive DNA families (psr2 and psr18). Three lambda clones that hybridized to both repeat families were isolated from PSR-genomic DNA libraries through multiple screenings and analyzed by Southern blots. Analysis of clones showed a region in which the two repeat types are interspersed, flanked by uniform blocks of each repeat type. PCR amplification of genomic DNA confirmed the contiguous arrangement of psr2 and psr18 on PSR and identified an additional junction region between these repeats that was not present in the lambda inserts. We isolated and sequenced 41 clones from the lambda inserts and genomic PCR products containing junction sequences. Sequence analysis showed that all transitions between psr2 and psr18 repeats occurred near one of the two palindromes. Based on the inheritance pattern of PSR, recombination between repeats on this chromosome must be mitotic (rather than meiotic) in origin. The occurrence of exchanges near the palindromes suggests that these sequences enhance recombination between repeat units. Rapid amplification of repetitive DNA may have been an important factor in the evolution of the PSR chromosome.

Key words: Paternal-Sex-Ratio — *Nasonia* — Repetitive DNA — Junctions — Recombination — Palindromes

Introduction

Repetitive DNAs are found in nearly all multicellular eukaryotes; however, their sequences are highly divergent between species. The length of individual repeat units varies and is dependant on the size of the region of recombination (Smith 1976; Stephan 1989) and on physical constraints such as nucleosome phasing (Musich et al. 1977; Brown et al. 1979). For example, a wide range of organisms show tandem arrays comprised of individual repeats 150–200 nucleotides in length. This size is similar to the number of nucleotides contained per core nucleosome (Richmond et al. 1984).

Large arrays of repetitive DNA are commonly believed to be generated by unequal crossing over between sister chromatids (Southern 1975; Smith 1976). However, controversy exists about the importance of other mechanisms, such as gene amplification (Walsch 1987), replication slippage (Stephan 1989), and sudden large-scale amplification (Britten and Kohne 1986; Sutton and McCallum 1972). Bacteriophage λ clone inserts containing repetitive DNAs are known to be unstable, suggesting an ongoing dynamic process (Arnheim and Kuehn 1979; Collins and Rubin 1983; Levis and Rubin 1982; Kiyama et al. 1987). Once amplified, homogeneity of repetitive arrays becomes a balance between divergence of individual repeat units through mutation and homogenization through recombination (Dod et al. 1989).

One approach to studying the mechanisms of repet-

* *Present address:* Department of Biology, University of Wisconsin—Milwaukee, Milwaukee, WI 53201, USA
** *Present address:* Max-Planck-Institut für Verhaltensphysiologie, D-82319 Seewiesen (Post Starnberg), Germany
Correspondence to: John H. Werren

itive DNA evolution is to examine the structure of junction sites between different tandem arrays or between tandem arrays and single-copy DNA. The structure of junctions between two short (5–10 bp) repetitive elements has been examined in *Drosophila* (Lohe and Brutlag 1987). These authors found two classes of junctions. One showed an abrupt switch between repeat types with maintenance of repeat periodicity. The second class was more complex, involving interspersion of repeat types and the possible insertion of a mobile element.

We have been studying the structure and evolution of repetitive DNA with reference to the Paternal-Sex-Ratio (PSR) chromosome of *Nasonia vitripennis.* PSR is a supernumerary (B) chromosome that causes destruction of paternal chromosomes in fertilized eggs. PSR contains several families of repetitive DNAs that appear to be unique to this chromosome and not present on the autosomes (Eickbush et al. 1992). This chromosome is particularly interesting because it does not normally experience meiosis. PSR occurs only in males which have mitotic, not meiotic, gametogenesis. Since homologous chromosomal recombination during meiosis is not plausible, mechanisms of amplification and maintenance of repetitive DNAs must be strictly mitotic.

Paternal-Sex-Ratio (PSR) causes all-male families in the parasitoid wasp, *Nasonia vitripennis.* As in other Hymenoptera, *Nasonia* has haplodiploid sex determination—males usually develop from unfertilized (haploid) eggs whereas females develop from fertilized (diploid) eggs. Male carriers of PSR produce motile sperm which effectively fertilize eggs. However, in the first mitotic division of the fertilized egg, the paternal chromosomes (except PSR) condense into a single chromatin mass and are eventually lost (Werren et al. 1987; Reed and Werren in prep). Although the mechanism of PSR action is still poorly understood, the net effect is that PSR converts fertilized (diploid) eggs, which would have developed into females, into haploid eggs that develop into PSR-bearing males. PSR is an extreme example of "selfish" or parasitic DNA for two reasons (Werren et al. 1988; Nur et al. 1988; Werren 1990). First, PSR enhances its own transmission by haploidizing its host genome. Transmission is enhanced because as an unpaired chromosome PSR has high transmission through mitosis (spermatogenesis), whereas it would have low transmission through meiosis (oogenesis). Second, PSR is "selfish" because it completely eliminates its host's genome each generation.

Until recently, no unique (single copy) sequences had been isolated from PSR and we believe the chromosome is primarily comprised of repetitive DNA. The chromosome contains several families of repetitive DNAs that can be distinguished based upon sequence divergence and hybridization criteria (Nur et al. 1988; Eickbush et al. 1992). Three of the families (psr2, psr18, and psr22) are specific to PSR, based on Southern hybridization. The basic repeat size of these families ranges from 154 to 214 bp with levels of similarity between the repeat families ranging from 45 to 71% (Eickbush et al. 1992). Despite considerable divergence, these families share two small (10 and 14 bp) palindromic DNA sequences which have apparently been evolutionarily conserved among the repeat families. Although the sequences are actually dyadic and not true palindromes, herein we use the term palindrome for consistency with the initial description (Eickbush et al. 1992). A fourth family (psr79), present on the autosomes but apparently amplified on PSR (Eickbush et al. 1992), does not contain the palindromes. Restriction-enzyme mapping of lambda clones and genomic Southerns indicates that the repeat families are organized in long, tandem arrays. Deletion analysis corroborates this interpretation and further suggests that most of the families are present in single arrays (Beukeboom and Werren 1993). From analysis of the original library (Nur et al. 1988), one clone was known to hybridize to two repeat families (psr2 and psr18). Here we report the molecular characterization of junctions between these repetitive DNA families on the PSR chromosome and show that recombination has preferentially occurred near palindromic regions in the repeats.

Materials and Methods

Identification and Characterization of Lambda Clones. Construction of the *N. vitripennis* library and screening for PSR-specific clones are described in Nur et al. (1988) and Eickbush et al. (1992). A second PSR total genomic library was constructed to look for additional PSR-specific repeats and for clones containing more than one repeat family. DNA of PSR males was isolated and cloned into lambda vector (EMBL3, Stratagene) after partial digestion with *Sau3*A. Approximately 60,000 clones, representing about three *Nasonia* genomes (Rasch et al. 1975), were incubated overnight at 37°C on agar plates (NCZY medium) with the bacterial strain KH402 as host. Replicate nitrocellulose filters were lifted from each plate and hybridized in parallel at moderate stringency (65°C, 4 × SSC, for details see Beukeboom and Werren 1993) to P^{32}-labeled plasmid (pUC) vectors containing psr2, psr18, psr22, or psr79 repeats. Clones which hybridized to more than one repeat type were isolated and then rescreened with the appropriate probes. This method identified four clones that hybridized to more than one repeat family. Phages were subsequently amplified as described by Maniatis et al. (1982).

Purified phage DNA was digested with excess enzyme under conditions described by the supplier (Gibco BRL, New England Biolabs). Digested DNA was subjected to electrophoresis on horizontal 1% agarose gels and transferred to Duralose (Stratagene) filters for Southern analysis. Filters were hybridized to PSR-specific probes under moderate stringency (65°C, 4 × SSC). For multiple screenings, either filters were stripped and reprobed, or duplicate filters were analyzed. Southern blots identified restriction enzymes that cut the lambda inserts into small fragments for subcloning into the sequencing vector M13mp19. DNA templates were prepared and sequenced as described in Maniatis et al. (1982). Sequences were analyzed with ESEE (Cabot and Beckenbach 1989) and the matrix analysis programs of Macvector (Biotechnologies, Inc.).

PCR Amplification and Analysis of Junctions. As an alternative method to detect junctions between the major repeat families (2, 18, 22, and 79), polymerase chain reaction (PCR; Innis et al. 1990) was used to directly examine the PSR chromosome. The following oligonucleotide primers, specific for each repeat family, were used: 2F (5′-GAACGCCGTCCTGAGGCC-3′), 2R (5′-GGCCTCAGGACGGCGTTC-3′), 18F (5′-GAGGAAAAGATGTCGCTCC-3′), 18R (5′-CGACATCTTTTCCTCAGCG-3′), 22F (5′-CGTCTTAAATGAAGAAGC-3′), 22R (5′-GCTTCTTCATTTAAGACG-3′), 79F (5′-GGAGGAAGGCGAGCGAGGG-3′), and 79R (5′-CCTCGCTCGCCTTCCTCCG-3′). PCR primers were synthesized with an Applied Biosystems 392 DNA/RNA synthesizer, deblocked by incubation for 8 h at 55°C, and purified by the method of Sawadogo and Van Dyke (1991). Target sequences were amplified by PCR using an Ericomp thermal cycler. Each reaction contained 1 μl DNA sample, 2.5 μl 10× *Taq* polymerase buffer, 1 μl nucleotide mix (10 mM each), 0.5 μl 10 μM primer (0.25 μl each), 0.5 μl (2.5 units) *Taq* DNA polymerase, and ddH_2O to final volume of 25 μl. The PCR reaction cocktail was prepared in one batch and then added to each sample separately. Replicate reactions, lacking template DNA, were run to test for contamination of reagents. Reactions were incubated for one cycle consisting of 1 min at 94°C, 1 min at 36°C, and 2 min at 72°C followed by 35 cycles consisting of 30 s at 92°C, 30 s at 55°C, and 1 min at 72°C. All two-primer combinations were tried successively with total PSR genomic DNA as the target. Amplified products (5 μl) were resolved on 1.0% agarose gels and run on 1× TAE buffer. In the initial amplifications, both lambda clones and subclones known to contain junctions were used as templates for the PCR. This allowed for direct comparisons of the organization of the various clones, positive controls for amplification, and the ability to identify products amplified from the PSR chromosome but not represented in the lambda clones. For cloning, whole PCR products were purified (GeneClean, Bio101) or specific bands were cut from agarose gels and the DNA was eluted by spin-column centrifugation. DNA was directly cloned into a T-tailed M13mp18 derivative (W. Burke, unpublished) and sequenced.

Results

In addition to the clone (λG) reported in Nur et al. (1988), screening of the new library found two additional clones that hybridized to both psr2 and psr18 (clones λad9 and λad19) and one clone that hybridized to psr18 and psr79 (clone λ18-79). However, sequence analysis (Beukeboom 1992) suggested that the 18-79 junction was a cloning artifact. PCR with both combinations of the 18 and 79 primers failed to amplify this putative junction from the PSR chromosome, again supporting the contention that the association of psr18 and psr79 was a cloning artifact. No clones contained junctions involving the psr22 repeat family.

Characterization of Lambda Clones

DNA organization of the three psr2-18 clones (λad9, λad19, and λG) was investigated by Southern analysis with 14 restriction enzymes. Several digestion patterns were distinguished. Some enzymes cut the DNA rarely, resulting in large fragments that hybridized to both psr2 and psr18. However, most enzymes produced small fragments (<10 kilobases, kb) which hybridized to either one or both repeats. Digestion patterns for three enzymes are presented in Fig. 1. Southern blots indicated three regions or blocks of DNA: One block of psr2, one of psr18, and one of interspersed (or intermediate) sequence.

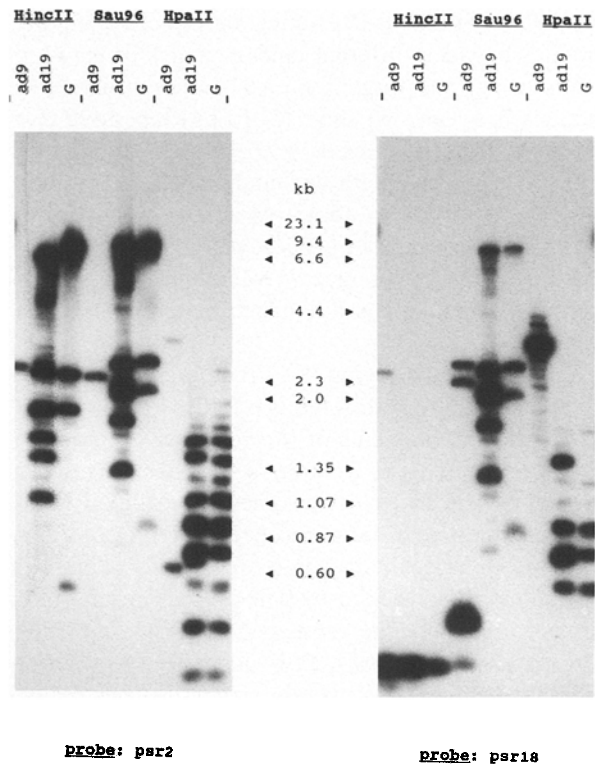

Fig. 1. Southern blots of 2-18 clones λad9, λad19, and λG digested with three restriction enzymes. Blots were hybridized to a psr2 (*left*) or psr 18 (*right*) probe (P^{32}-labeled subclones in the plasmid vector pUC). *Numbered arrowheads* indicate fragment sizes (in kilobases).

A separate block of psr2 (repeat length 171 bp) is indicated in the *Hpa*II digest (Fig. 1) which show small (<600 bp) fragments that are absent in the psr18 hybridization. These smaller fragments represent monomers and dimers of psr2. A separate block of psr18 (repeat length 214 bp) is indicated in the *Hin*cII digest, where very small (monomer size) fragments hybridized to psr18, but not to psr2. Hybridization of similar-sized fragments (600–2,000 bp) to both probes in the *Hpa*II and *Sau*96A digests indicated interspersion of psr2 and psr18 repeats and/or intermediate sequences larger than a single repeat unit. The ladderlike pattern in these digests is indicative of higher-order periodicity within this interspersed region.

Southern blots showed that the λad9 clone was very different in organization from λad19 and λG. Its weak hybridization to psr2 but strong hybridization to psr18 suggests that it contains many psr18 but few psr2 repeats. Again, presence of a block of psr18 was indicated by the small fragments in the *Hin*cII digest. Digestion patterns of clones λad19 and λG were very similar, although not identical. For example, the uppermost band

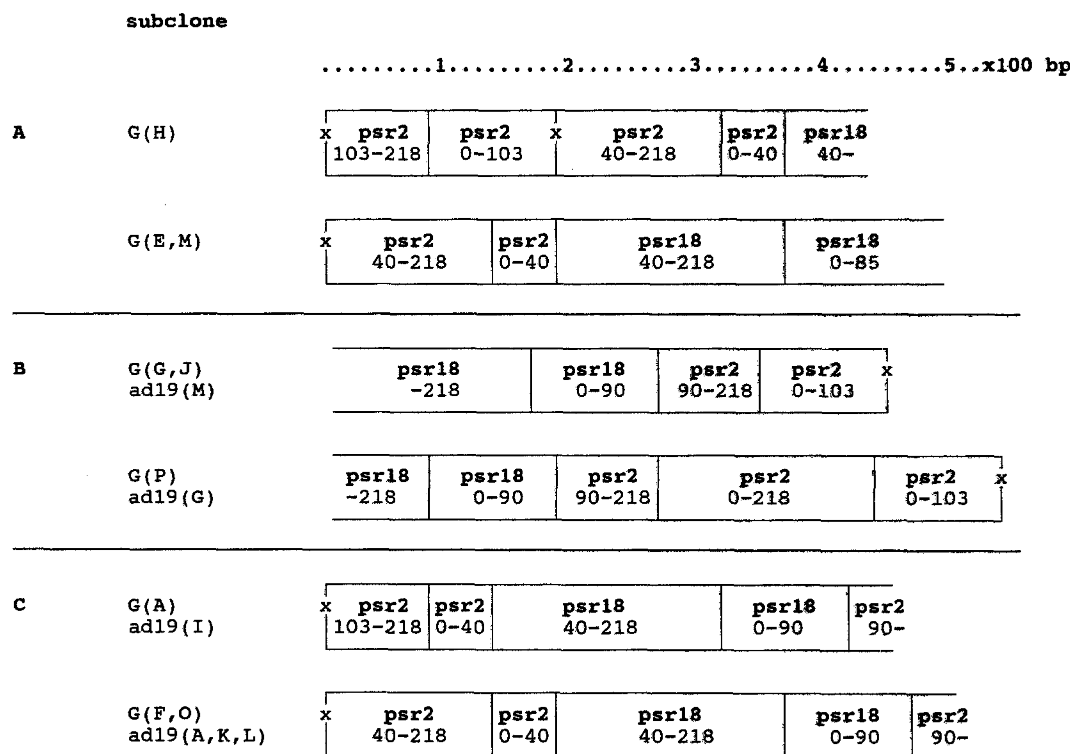

Fig. 2. Summary of DNA organization of subclones containing junctions between PSR repeat families, subclones from 2-18 clones λG and λad19. Sequence similarity to one of the two repeats (see Fig. 3) is shown in the block. The approximate length of each sequence (sub)unit can be determined from the *numbered line* on top (one point is 10 bp). Presence of restriction sites of enzymes used in the cloning procedures is indicated at some repeat borders. The symbol *x* denotes restriction sites of the subcloning enzyme (*Hin*dIII, A↓AGCTT). *Numbers within blocks* refer to approximate positions in the 218-bp sequence "frame" (cf. Fig. 3).

in the *Hpa*II digest of λG is not present in λad19. Most likely these two clones contain overlapping regions of the PSR chromosome.

Subcloning and Sequencing of Lambda Clones

Two of the three psr2-18 clones (λG and λad19) were subcloned for sequencing with *Hin*dIII. A total of 10 subclones from λG and nine from λad19 were sequenced. Subclones with similar DNA organization were found among both lambda clones, again suggesting they arose from an overlapping region of the PSR chromosome. The subclones fell into two groups: Those (n = 15) that contained both psr2 and psr18 sequences and those (n = 4) containing only psr2. Sequences other than psr2 and psr18 were not isolated.

Several junctions between psr2 and psr18 repeats were identified and these are summarized in Fig. 2. Roughly half of the subclones (n = 8) contained one to several psr2 repeats adjacent to one to several psr18 repeats (Fig. 2a,b). The remaining subclones (n = 7) included a stretch of psr18 flanked on both sides by psr2 (Fig. 2c). In most cases, only the first 400–500 nucleotides were sequenced, which did not always include the entire insert.

Sequences are aligned in Fig. 2 for comparison to the psr2 and psr18 sequences as published in Eickbush et al (1992). All junctions between psr2 and psr18 repeats were of one of two types. The first type (psr2 → psr18; ad19(A) in Fig. 3) occurred between positions 18 and 40 in the 218-bp frame. This region coincides with the palindromic sequence (palin I) found between positions 19 and 32. A junction of similar structure was sequenced from the DNA amplified from λad9 by PCR.

The precise boundaries between repeats (bp number) could not be determined because psr2 and psr18 have similar sequences in this region. Boundaries were determined by the first position discriminative for both repeats. Thus, the adenosine at position 18 is typical for psr2, and the two adenosine residues at positions 41 and 42 are typical for psr18. Because of the presence of *Hin*dIII restriction sites at positions 40 and 103 in some psr2 repeats (see Fig. 2), we initially viewed these junctions with some skepticism. However, presence of these junction types in the PSR genome was verified through direct amplification using PCR. (See below.)

The second type of junction (psr18 → psr2; ad19(G) in Fig. 3) occurred at position 90 of the 218-bp frame, in the second half of the repeat unit. Clones containing this junction included a deletion of the AT-rich region (positions 104–122). This deleted region included the other known palindromic sequence (palin II). Except for four positions (134, 135, 185, and 190) the remaining sequence after this deletion was homologous to psr2.

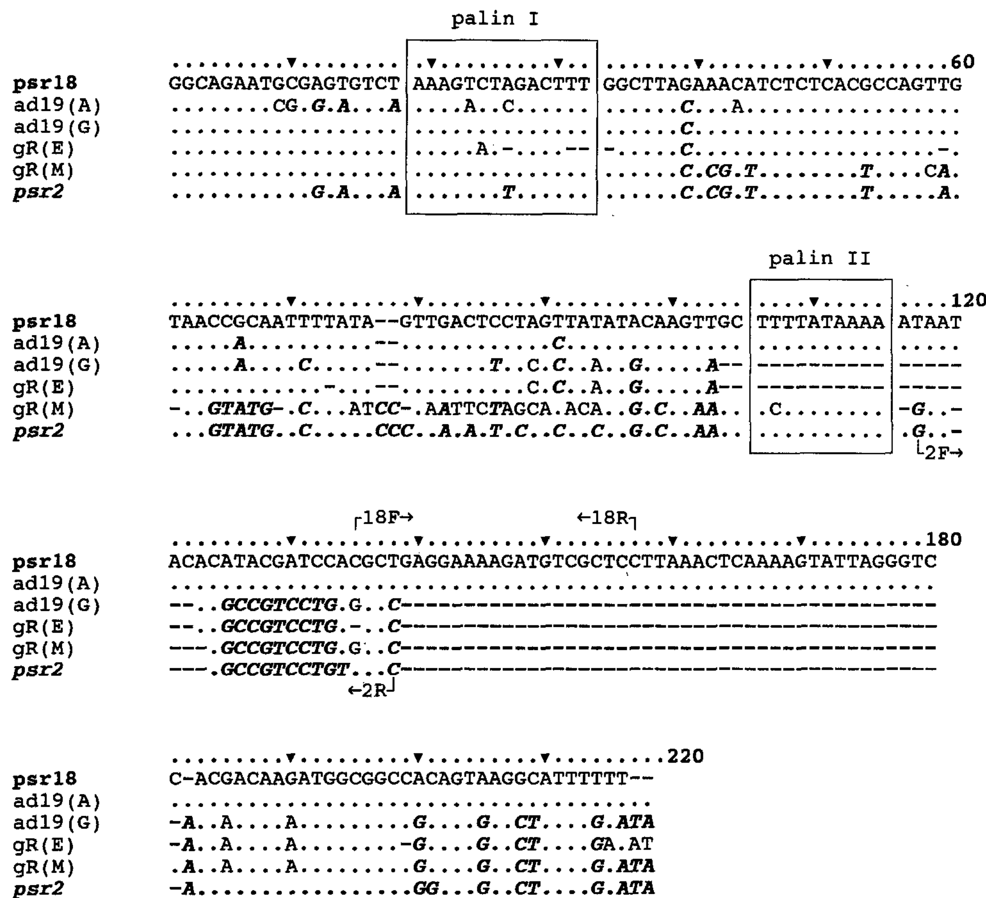

Fig. 3. Partial sequences representing the three junction types between the psr2 and psr18 repeat families. Included are sequences obtained from lambda subclones ad19(A) (psr2 ⟶ psr18) and ad19(G) (psr18 ⟶ psr2) and from cloned genomic PCR products gR(E) (psr18 ⟶ psr2) and gR(M) (psr18 ⟶ psr2). Included for comparison are sequences of psr2 (psr2-1, EMBL X64095) and psr18 (psr18-1, EMBL X64094) repeats (from Eickbush et al. 1992). Sequences are aligned within a 218-bp frame; actual repeat lengths are 213 bp (psr18) and 171 bp (psr2). The palindromic regions (I and II) are enclosed in *boxes; dashes* denote deletions; and the position and orientation of PCR primers for psr2 and psr18 are indicated. Positions diagnostic for the psr2 repeat are in *bold italics*.

PCR Amplification and Sequencing from Genomic DNA

We next used PCR to verify the presence of the psr2-18 junctions in the PSR genome and look for additional junction regions. Use of lambda clones as templates in the PCR reactions allowed us to make further comparisons of the organization of their inserts. PCR successfully amplified junctions between psr2 and psr18 from genomic DNA, confirming the contiguous distribution of these repeats on the PSR chromosome. Junctions involving the other repetitive DNA families (psr22 and psr79) were not detected by PCR amplification using various combinations (forward and reverse) of primers.

The presence of bands equal in size to those amplified from two previously sequenced lambda subclones (ad19(G) and ad19(A)) confirmed that the primers were amplifying psr2-18 junctions (Fig. 4). Electrophoresis produced a 200-bp ladder in the products amplified with the primer combination 2R/18F from the λad19 and λG templates (Fig. 4). In both cases, the smallest fragment ~400 bp) in the ladder corresponded in size to the product from the lambda subclone ad19(G) (lane 1). Results from the PCR amplification are consistent with the Southern analysis, which suggests that λad9 is distinctly different from the other lambda clones. Electrophoresis of the λad9 products produced a ladder with minimum fragment length of approximately 300 bp (Fig. 4, lane 4). A ladder also was displayed by the products amplified directly from PSR genomic DNA. However, in the case of the genomic DNA, the smallest fragment was approximately 200 bp in length (one repeat unit). This indicates that the PSR genome contained at least one junction between psr2 and psr18 that was not included in the lambda inserts.

As was the case for the 2R/18F primers, electrophoresis showed ladders in PCR products from the primer combination 2F/18R. The minimum fragment size in all three lambda clones and PSR genomic DNA (Fig. 4b) was approximately 200 bp and corresponded to the minimum size predicted from the ad19(A) sequence (lane 7). Fragments in the ladder increased in approximately 200-bp intervals (one repeat unit length).

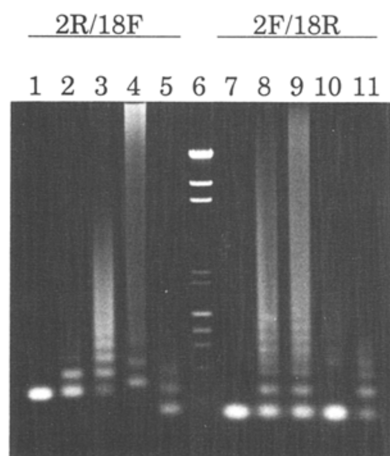

Fig. 4. Electrophoretic analysis of PCR products amplified with PSR-specific oligonucleotides. Included are products amplified with primer combination 2R/18F (lanes 1–5) and 2F/18R (lanes 7–11). Lane 6 contains a λ *Hin*dIII/øX174 *Hae*III standard. Template DNAs are as follows: Lambda subclone ad19(G) (lane 1), λad19 (lanes 2 and 8), λG (lanes 3 and 9), λad9 (lanes 4 and 10), PSR total genomic (lanes 5 and 11), and lambda subclone ad19(A) (lane 7).

The size of the amplified products, including both the smallest fragment and the distance between fragments, was similar for the PSR genomic DNA and lambda clone templates. These results indicated that psr2 and psr18 repeats are interspersed with each other over the junction regions.

A total of 25 clones containing junctions between psr2 and psr18 were obtained from the PCR products amplified directly from genomic DNA (Fig. 5). For clarity we denote clones obtained by PCR of genomic DNA with the prefix "g." Sequence analysis found three types of junction regions. Organization of the first two types corresponds to the junctions sequenced previously from the lambda inserts. The majority of these sequences (n = 12; Fig. 5a), amplified with the 2F/18R primer combination, were analogous to the psr2 → psr18 junctions occurring near palindrome I which were sequenced from the lambda inserts. Again, the exact position where psr2 crossed into psr18 could not be determined because of the degree of similarity between the repeats in this region.

Two of the 25 genomic clones contained a junction similar to the psr18 → psr2 junction (position 90) sequenced from the lambda subclones (Fig. 5b). However, these genomic clones (gR(J) and gR(E)) differed from the lambda subclones with respect to the sequence at palindrome I. (Compare ad19(G) with gR(E) in Fig 3.) Both sequences showed a substitution at position 24 and two deletions in the region between positions 26 and 34. Like ad19(G), sequences from these genomic clones contain a deletion (positions 104–122) that includes palindrome II.

Among the sequenced PCR products amplified from genomic DNA was one junction type not represented in the lambda inserts. This junction was amplified with the primer combination 2R/18F and had a minimum repeat unit of approximately 200 bp. This 200-bp junction sequence corresponds to the "extra" 200-bp band amplified from PSR genomic DNA but absent from the products amplified from the lambda templates. A total of 11 clones containing this junction were sequenced (Fig. 5c). Only one clone (gR(M)) contained more than one repeat unit.

Sequence analysis of these clones revealed a junction between psr18 and psr2 near position 40 (Fig. 3). This junction type differed from the psr2 → psr18 junctions sequenced from the lambda subclones in that the first portion of the sequence corresponded to psr18 rather than psr2. However, in these clones the cross-over site was still in the region of palindrome I. One clone (gR(M)) showed sequence divergence from the standard psr2 in positions 80–95 (just prior to palindrome II) and may represent a sequence variant in this family.

Comparison of Junction Regions

Given the different sizes and sources of the sequenced clones, and the opportunity for polymerase error in amplification and sequencing, base-by-base comparisons are not practical. Indeed, with the exception of four clones (gR(D, G, H, and I)) all clones of similar size and region differed by at least one nucleotide substitution. Variation at the actual junctions between repeats is of greater interest in the present study and, therefore, we emphasize only these regions. The junction regions of all sequenced clones are presented in aligned form in Fig. 6.

psr2 → psr18 (Palindrome I)

The differences among the junction regions of clones containing the psr2 → psr18 junction type can be summarized over three segments. First, in the 18 bases preceding palindrome I (Fig. 6a), 11 clones have C and G nucleotides at positions 9 and 10, respectively. This combination was not represented in psr2 clones or psr18 clones reported by Eickbush et al. (1992). Interestingly, most of the sequences obtained from the lambda inserts contained this variant, the majority amplified directly from the genome did not. Within palindrome I, most of the clones carry the C variant at position 26. This nucleotide substitution is known to occur in psr2 (Eickbush et al. 1992) and was present in the "non-junction" psr2 variants sequenced from the lambda inserts (G(P) Fig. 6a). In the region after the palindrome, most differences were single nucleotide changes except for the expansion in several clones between positions 40 and 50. In all cases the shift from psr2 to psr18 in this group occurred in the region just after palindrome I.

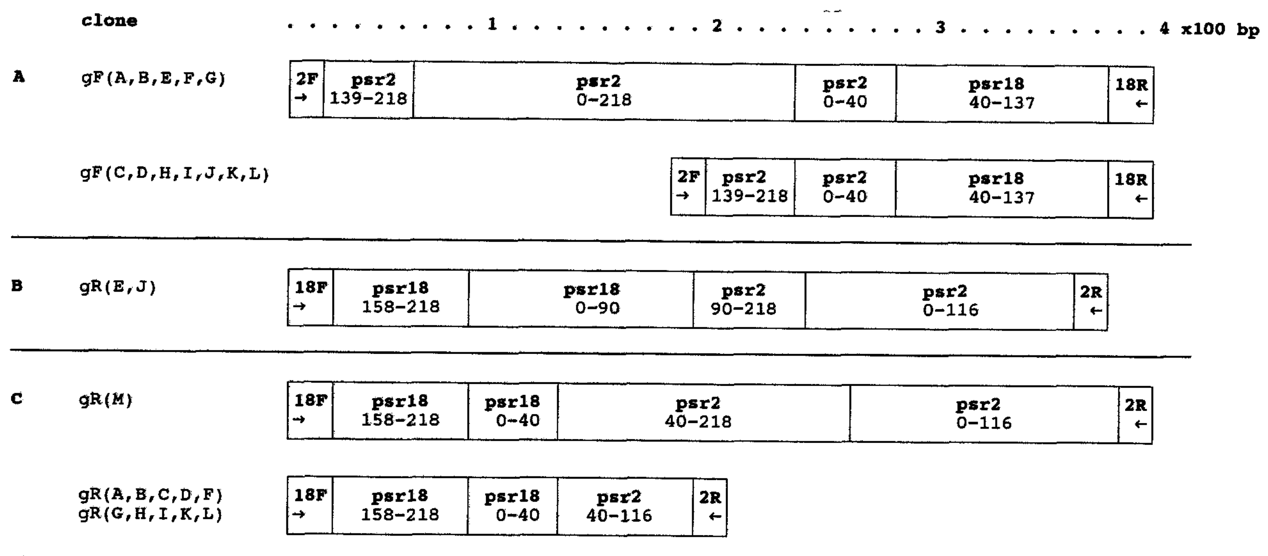

Fig. 5. Summary of DNA organization of junction sequences amplified from genomic DNA by PCR. A Sequences amplified with the 2F/18R primer combination. **B** and **C** Sequences amplified with the 2R/18F primer combination. Sequence similarity to the two repeats (psr2 and psr18) is shown in the block. The approximate length of each sequence (sub)unit can be determined from the *numbered line* on top (one point is 10 bp). *Numbers within blocks* refer to positions in 218 bp sequence frame (cf. Fig. 3).

psr18 → psr2 (Palindrome II)

The junction occurring near palindrome II differs from those involving palindrome I in that these junctions involve the complete deletion of the palindrome. The principal difference among these clones is the deletion of four bases from six of the clones at positions 135–138 (Fig. 6b). One major distinction between psr2 and psr18 is the large (43 bp) deletion which begins at position 139. Thus, the deletion of an additional four bases in these clones extends this deletion to 47 bp. The significance of the deletion of palindrome II in these sequences is not clear. If the palindromic regions are important for recombination, as the data suggest, these variants may be less likely to participate in future recombination. These repeat units still retain palindrome I and may represent a truncated form of the larger repeat unit or perhaps even a predecessor. The shared autosomal repeat (psr79) has a basic repeat unit of only 94 bases and lacks both palindromes, although it does include a 28-bp interval which is A/T rich (90%).

psr18 → psr2 (Palindrome I)

Examination of the sequences containing junctions where psr18 crosses into psr2 near palindrome I shows that these clones and their junctions fall into two groups (Fig. 6c). The first group includes three clones (gR(A, K, and L)) in which the transition from psr18 to psr2 occurs prior to the palindrome (before position 12). In the second group (n = 8 clones), the transition occurs after the palindrome (prior to position 40). The first group of clones also contains nucleotide variants at positions 25 and 26 within palindrome I. Again, this junction region was not included in the lambda inserts but was amplified directly from the genome by PCR.

Recovery of this junction distinct from the lambda inserts suggests the presence of at least two junction regions between these repeats on the PSR chromosome. It is currently not known whether junctions involving the other repeat families exist on PSR, although none were detected in either the library screen or through PCR. Failure to detect additional junction regions in the library could have been the result of selective elimination of junction-site clones during its generation and screening. Results of the PCR screening suggest that blocks of the remaining repetitive elements (psr22 and psr79) are not directly adjacent to each other or to psr2 and psr18 on the chromosome but are separated by other DNAs. However, we cannot dismiss the possibility that the lack of amplification was due to sequence divergence at the primer sites.

Discussion

The importance of palindromic sequences in recombination is still poorly understood (Higgins et al. 1988). They have been shown to increase recombination (Anderson and Roth 1978; Warren and Green 1985; Higgins et al. 1988) and DNA amplification (Cohen and Kopecko 1976). They are known to occur at sites of DNA breakage associated with gene conversion, nonhomologous recombination, and sister chromatid exchange (Israelewski 1983; Henthorn et al. 1986; Krawinkel et al. 1986; Hyrien et al. 1987; Vogel et al. 1990) and may affect chromosome structure (Gilson et al. 1986; Ohno 1990) or pairing (Sobell 1972; Doyle 1978).

Several studies have found palindromic sequences as-

```
A      psr2 → psr18              1.........▾.........▾.........▾.........▾.............▾...........60

                                                      Palin I
              psr18              GG-CAGAATGCGAGTGTCT |AAAG-TCTAGACTTT| GGCTTA-GAAAC-----ATCT-CTCACGCCA-GTT--G
              psr18 ad19(I)      ................... |......A.G....--| .......C..............................

λsubclones    ad9(A)             .........-..G.A...A |........C......| .......C..-.....-.......................
              G(A)               .........CG.ggA...A |...AA...C......| .......C..-....A........................
              G(E,F,M) ad19(L)   .........CG.G.A...A |........C......| .......C..-....A........................
              G(O)               ...........--A...A  |..-.C...C......| .......C..-....A........................
              ad19(A)            ..G......CG.G.A...A |.....A..C......| .......C..-....A........................
              ad19(I)            .........CG.ggA...A |...A....C......| .......C..-....A........................
              ad19(K)            .........-G.G.A...A |....C...C......| .......C..-....A........................
              G(H)               .........CG.g.A...A |........C......| .......C..-...AA.........T..........

PSR genome    gF(J)              .........-..G.A...A |........C......| .......C................................
              gF(K)              ............GgA...A |........C......| .-.....C..-.....................-...
              gF(A)              ............G.A...A |........C......| .......C..-....A........................
              gF(C,D,L)          .........CGggA...A  |........C......| .......C..-....A........................
              gF(H)              ............G.A...A |........C......| .......C...AAGCGC.---.-.................
              gF(E)              ............G.A...A |....C...-......| ......AC..-....A....T.............-...
              gF(F,G)            ............G.A...A |..-............| .....T.C..-....A........................
              gF(I)              ............G.A...A |...............| .......C..-..AAA..................-...
              gF(B)              -......-....gg....A |..............-| ..G....C----.TGGC...............A..ATA.

              psr2 G(P)          ............G.A...A |C.......C......| .......C.CG......T.........T.....CA...
              psr2               ............G.A...A |........T......| .......C.CG......T.........T......A...


B      psr18 → psr2              70.........▾.........▾.........▾.........▾.........▾.........▾.........140

                                                                    Palin II
              psr18              TTTTATA--GTTGACTCCTAGTTATATACAAGTTGC |TTTTATAAAA| ATAATACACATACGATCCACGCTGA

λsubclones    ad19(K)            .................-.-.c..-..cg....a-- |----------| --------.gc-gtc-tg.------
              ad19(L)            .................-.-.c..-..cgc...a-- |----------| --------.cg-gtc-tg.------
              G(A)               ...................C.C..A..G.....a-- |----------| -------c.gccgtcctg.------
              G(F)               .................-.c.c..-..g.....a-- |----------| -------..gc-gtc-tg.------
              ad19(A)            .................-.-.c..-..cg....a-- |----------| --------.gc-gt--tg.------
              ad19(I)            ...................C.C..A..G.....A-- |----------| -------.-Gc-gctctg.------
              G(G,J)             ...................C.C..A..G.....A-- |----------| -------..GCCGTCCTG.G..C--
              G(P)               ..............c....C.C..A..G.....A-- |----------| -------C.GCCGTCCTG.G..C--
              ad19(G,M)          .C.............T..C.C..A..G.....A-- |----------| -------..GCCGTCCTG.G..C--

PSR genome    gR(E,J)            A..................C.C..A..G......-- |----------| -------..GCCGTCCTG.-..C--

              psr2               .C.....CCC..A.A.T.C..C..C..G.C..AA.. |..........| .G..----.GCCGTCCTGT...C--


C      psr18 → psr2              ...1.......▾.........▾.........▾.........▾.........▾.........60

                                                      Palin I
              psr18              T--GGCAGAATGCGAGTGTCT |AAAGTCTAGACTTT| GGCTTAGAAACATCTCTCACGCCAGTTG

PSR genome    gR(A,K,L)          T--.T.........G.A...A |......GC......| ......C.CG.T........T....CA.
              gR(B,C,D,F)        T--.................. |..............| ......C.CG.T........T....CA.
              gR(G,H,I,M)
              psr2               ATA...........G.A...A |.......T......| ......C.CG.T........T.....A.
```

Fig. 6. Partial sequences of all sequenced junctions between the psr2 and psr18 repeat families. **A** psr2 ⟶ psr18. **B** and **C** psr18 ⟶ psr2. Included are sequences of psr2 (psr2-1, EMBL X64095) and psr18 (psr18-1, EMBL X64094) repeats (from Eickbush et al. 1992) and partial sequences of psr2 and psr18 from nonjunction regions of two lambda subclones (G(P) and ad19(I), respectively). *Numbered line on top* (one point is 1 bp) refers to positions in 218-bp sequence frame. The palindromic regions (I and II) are enclosed in *boxes* and *dashes* denote deletions. Positions diagnostic for the psr2 repeat are in *bold italics*.

sociated with repetitive DNAs. For example, the common highly repetitive DNA of cetaceans contains a large 72-bp inverted repeat (Gretarsdottir and Árnason 1992), and Ekker et al. (1992) found a 15-bp palindrome in repetitive DNA from the zebrafish. Likewise, the highly repeated DNA sequences in the Japanese hagfish contain two inverted dyadic sequences (Kubota et al. 1993). Bigot et al. (1990) suggested a functional role for palindromes in the amplification of repetitive DNA in two wasp species. These authors hypothesized that the palindromes could form secondary dyadic structures, serving as recognition sites for binding of recombinational enzymes. Chalker et al. (1993) provided further evidence of cruciform extrusion of palindromes in vivo. Results from the present study are also suggestive for a causal relationship between palindromic sequences and recombination events. We are currently investigating whether the PSR palindromic sequences are also present on the normal (A) chromosomes of *N. vitripennis* and closely related species.

Sequence conservation of the PSR palindromes among repeat families and the occurrence of all junctions near the palindrome (palin I or II) advocate a functional role for these sequences in exchange. Recent studies suggest that short stretches of near sequence identity are more important than general homology in homologous recombination (Abastado et al. 1987; Waldman and Liskay 1987; Metzenberg et al. 1991; Cabot et al. 1993). However, Kiyama et al. (1987) found evidence that palindromic structures rather than extensive sequence homology were important for recombination in a *Sau*3A alphoidlike human repetitive DNA. In addition, Trinh and Sinden (1993) found evidence that secondary structures, when present during DNA replication, can facilitate and direct specific mutagenic events. In their studies, certain palindromic sequences with the potential to form hairpin structures showed a high frequency of deletion. In the present study, all junctions near palin II included the complete deletion of this palindrome. Recently, Wevrick et al. (1992) isolated a junction between α-satellite DNA and a novel low-copy repetitive element in humans. Sequence analysis found an abrupt junction located at a region containing stretches of A's and T's. Similarly, the junction involving palindrome II in the present study occurred near a region containing an A/T stretch. This junction was wider and more complicated with greater sequence variation and deletions than the junction at palindrome I.

All PSR-specific repeats (psr2, psr18, and psr22 families) possess the palindromes (palin I and II) and show partial sequence homology, suggesting evolution from a common ancestor (Eickbush et al. 1992). The original junction between the psr2 and psr18 repeats likely consisted of a single pair of adjacent repeats. Subsequent rearrangements/amplifications appear to have involved differently sized repeating units. Evidence for this is the interspersion of repeat types and the higher-order periodicity (multiple repeat unit) in the junction region. Similar patterns of dimeric and tetrameric periodicity are reported from human α-satellite DNA (Shmookler Reis et al. 1985; Willard and Waye 1987).

The palindromic regions on PSR are present in repeats that number in the tens of thousands and in repetitive DNA families which have otherwise diverged significantly from each other. It is difficult to explain this conservation unless the palindromes have some selective function. As discussed, they could play a role in amplification of repeats. A second possibility is that the palindromes and A/T-rich areas play an active role in heterochromatin condensation by enhancing the formation of secondary structure (Rojas-Rousse et al. 1993). Eickbush et al. (1992) proposed that the palindromes could be involved in PSR expression by acting as a "sink" to sequester a protein necessary for condensation or division of the paternal chromosomes in fertilized eggs. Based on gel-shift assays, a nuclear protein has been detected that specifically binds to PSR repeats containing palindrome I (Liang, Benjayati, and Werren, unpublished). However, it is currently unknown whether this protein has any functional role or whether binding to the repeats occurs in vivo.

The existence of large regions of repetitive DNA on PSR that are not present on the autosomes of *Nasonia vitripennis* or its sibling species indicates that these repeats amplified after the origin of this chromosome. Our current belief is that PSR arose from an interspecies hybridization between *N. vitripennis* and its sibling species *N. longicornis* (Werren 1991). Evidence supporting this view is that interspecies crosses in *Nasonia* are known to result in paternal chromosome fragmentation due to cytoplasmic incompatibility (Ryan and Saul 1968; Breeuwer and Werren 1990). It has been shown that centromeric fragments occasionally survive incompatible crosses and are stably transmitted through males, although poorly transmitted through females. The distribution of PSR in natural populations corresponds to a region of sympatry between *N. vitripennis* and *N. longicornis,* two species showing partial cytoplasmic incompatibility (Breeuwer 1992). Thus, PSR may be the result of a megamutation, having originated as a centromeric fragment (or several fused fragments) from an incompatible mating.

Rapid amplification of repetitive DNA may have been an important factor in the evolution of PSR. Walker (1971) suggested that an increase in the amount of centromeric heterochromatin might confer a selective advantage to a chromosome. We have found that PSR chromosomes with induced deletions show a positive correlation between the total abundance of repetitive DNA and transmission of the chromosome (Beukeboom 1992). Reduced transmission of PSR-deletion chromosomes was found to be the result of somatic and germline mosaicism (Beukeboom et al. 1992), indicating that mitotic stability may be correlated with overall size of the chromosome, and perhaps specifically with abundance of repetitive DNA. The presence of palindromic sequences may have enhanced amplification of the repeats on this chromosome, increasing its stability.

Acknowledgments. We thank W. Burke and H. Liang for technical assistance, T. Eickbush for providing use of laboratory facilities, and B. McAllister and two anonymous reviewers for comments on the manuscript. This research was supported by a grant from the National Institutes of Health (R01-HD24310) to J.H. Werren.

References

Abastado J-P, Darche F, Godeau B, Cami B, Kourilsky P (1987) Intramolecular recombination between partially homologous sequences in *Eschericia coli* and *Xenopus laevis* oocytes. Proc Natl Acad Sci USA 84:6496–6500

Anderson RP, Roth J (1978) Tandem chromosomal duplications in

Salmonella typhimurium: fusion of histidine genes to novel promoters. J Mol Biol 119:147–166

Arnheim N, Kuehn M (1979) The genetic behavior of a cloned mouse ribosomal DNA segment mimics ribosomal gene evolution. J Mol Biol 134:743–763

Beukeboom LW (1992) Genetics and population biology of the paternal sex ratio (PSR) chromosome. PhD thesis, University of Rochester, Rochester, NY

Beukeboom LW, Reed KM, Werren JH (1992) Effects of deletions on mitotic stability of the Paternal-Sex-Ratio (PSR) chromosome from *Nasonia.* Chromosoma 102:20–26

Beukeboom LW, Werren JH (1993) Deletion analysis of the parasitic paternal-sex-ratio (PSR) chromosome from *Nasonia vitripennis.* Genetics 133:637–648

Bigot Y, Hamelin M-H, Periquet G (1990) Heterochromatin condensation and evolution of unique satellite-DNA families in two parasitic wasp species: *Diadromus pulchellus* and *Eupelmus vuilleti* (Hymenoptera). Mol Biol Evol 7:351–364

Breeuwer JAJ (1992) Postzygotic reproductive isolation in the genus *Nasonia* (Hymenoptera, Pteromalidae): the role of cytoplasmic microorganisms and other heritable cytoplasmic factors. PhD thesis, University of Rochester, Rochester, NY

Breeuwer JAJ, Werren JH (1990) Microorganisms associated with chromosome destruction and reproductive isolation between two insect species. Nature 346:558–560

Britten RJ, Kohne DE (1986) Repeated sequences in DNA. Science 161:529–540

Brown FL, Musich PR, Maio JJ (1979) The repetitive sequence structure of component alpha DNA and its relationship to the nucleosomes of African green monkey. J Mol Biol 131:777–799

Cabot EL, Beckenbach AT (1989) Simultaneous editing of multiple nucleic acid and protein sequences using ESEE. Comp Appl Biosci 5:233–234

Cabot EL, Doshi P, Wu M-L, Wu C-I (1993) Population genetics of tandem repeats in centromeric heterochromatin: evidence for unequal crossovers at the *Responder* locus of *Drosophila melanogaster.* Genetics 135:477–487

Chalker AF, Okeley EA, Davison A, Leach DRF (1993) The effects of central asymmetry on the propagation of palindromic DNA in bacteriophage λ are consistent with cruciform extrusion in vivo. Genetics 133:143–148

Cohen SN, Kopecko DJ (1976) Structural evolution of bacterial plasmids: role of translocating genetic elements and DNA sequence insertions. Fed Proc 35:2031–2036

Collins M, Rubin GM (1983) High frequency precise excision of the *Drosophila* foldback transposable element. Nature 303:259–260

Dod B, Mottez E, Desmarais E, Bonhomme F, Roizés G (1989) Concerted evolution of light satellite DNA in genus *Mus* implies amplification and homogenization of large blocks of repeats. Mol Biol Evol 6:478–491

Doyle GG (1978) A general theory of chromosome pairing based on the palindromic DNA model of Sobell with modifications and amplifications. J Theor Biol 70:171–184

Eickbush DG, Eickbush TH, Werren JH (1992) Molecular characterization of repetitive DNA sequences from a B chromosome. Chromosoma 101:575–583

Ekker M, Andreas F, Westerfield M (1992) Identification of two families of satellite-like repetitive DNA sequences from zebra fish (*Brachydanio rerio*). Genomics 13:1169–1173

Gilson E, Perrin D, Clement J-M, Szmelcman S, Dassa E, Hofnung M (1986) Palindromic units from *E. coli* as binding sites for a chromoid-associated protein. FEBS 206:323–328

Gretarsdottir S, Árnason U (1992) Evolution of the common cetacean highly repetitive DNA component and the systematic position of *Orcaella brevirostris.* J Mol Evol 34:201–208

Henthorn PS, Mager DL, Huisman THJ, Smithies O (1986) A gene deletion ending within a complex array of repeated sequences 3′ to the human β-globin gene cluster. Proc Natl Acad Sci USA 83: 5194–5198

Higgins CF, McLaren RS, Newbury SF (1988) Repetitive extragenic palindromic sequences, mRNA stability and gene expression: evolution by gene conversion? Gene 72:3–14

Hyrien O, Debatisse M, Buttin G, de Saint Vincent BR (1987) A hotspot for normal amplification joints in a mosaic of Alu-like repeats and palindromic A+T-rich DNA. EMBO J 6:2401–2408

Innis MA, Gelfand DH, Sninsky JJ, White TJ (1990) PCR protocols: a guide to methods and applications. Academic Press, San Diego, CA

Israelewski N (1983) Structure and function of an AT-rich, interspersed repetitive sequence from *Chironomus thummi* solenoid DNA, 142bp palindrome-frame and homologies with the sequence for site-specific recombination of bacterial transposons. Nucleic Acids Res 11:6985–6996

Kiyama R, Okumura K, Matsui H, Bruns GAP, Kanda N, Oishi M (1987) Nature of recombination involved in excision and rearrangements of human repetitive DNA. J Mol Biol 198:589–598

Kubota S, Kuro-o M, Mizuno S, Kohno S (1993) Germ line-restricted, highly repeated DNA sequences and their chromosomal localization in a Japanese hagfish (*Eptatretus okinoseanus*). Chromosoma 102:163–173

Krawinkel V, Zoebelein G, Bothwell ALM (1986) Palindromic sequences are associated with sites of DNA breakage during gene conversion. Nucleic Acids Res 14:3871–3882

Levis R, Rubin GM (1982) The unstable W^{DZL} mutation of *Drosophila* is caused by a 13 kilobase insertion that is precisely excised in phenotypic revertants. Cell 30:543–550

Lohe AR, Brutlag M (1987) Adjacent satellite DNA segments in *Drosophila:* structure of junctions. J Mol Biol 194:171–179

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Metzenberg AB, Wurzer G, Huisman THJ, Smithies O (1991) Homology requirements for unequal crossing over in humans. Genetics 128:143–161

Musich PR, Brown FL, Maio JJ (1977) Mammalian repetitive DNA and the subunit structure of chromatin. Cold Spring Harbor Symp Quant Biol 42:1147–1160

Nur U, Werren JH, Eickbush D, Burke W, Eickbush T (1988) A "selfish" B chromosome that enhances its transmission by eliminating the paternal chromosomes. Science 240:512–514

Ohno S (1990) Grammatical analysis of DNA sequences provides a rationale for the regulatory control of an entire chromosome. Genet Res, Cambridge 56:115–120

Rasch EM, Cassidy JD, King RC (1975) Estimates of genome size in haploid-diploid species of parasitoid wasps. Abstr Histochem Soc, p 317

Richmond TJ, Finch JT, Rushton B, Rhodes D, Klug A (1984) Structure of the nucleosome core particle at 7Å resolution. Nature 311: 532–537

Rojas-Rousse D, Bigot Y, Periquet G (1993) DNA insertions as a component of the evolution of unique satellite DNA families in two genera of parasitoid wasps; *Diadromus* and *Eupelmus* (Hymenoptera). Mol Biol Evol 10:383–396

Ryan SL, Saul GB (1968) Post-fertilization effect of incompatibility factors in *Mormoniella.* Mol Gen Genet 103:29–36

Sawadogo M, Van Dyke MW (1991) A rapid method for the purification of deprotected oligodeoxynucleotides. Nucleic Acids Res 19:674

Shmookler Reis RJ, Srivastava A, Beranek DT, Goldstein S (1985) Human alphoid family of tandemly repeated DNA. Sequence of cloned tetrameric fragments and analysis of familial divergence. J Mol Biol 186:31–41

Smith GP (1976) Evolution of repeated DNA sequences by unequal crossover. Science 191:528–535

Sobell HM (1972) Molecular mechanism for genetic recombination. Proc Natl Acad Sci USA 69:2483–2487

Southern EM (1975) Long range periodicities in mouse satellite DNA. J Mol Biol 94:51–69

Stephan W (1989) Tandem-repetitive noncoding DNA: forms and forces. Mol Biol Evol 6:198–212

Sutton WD, McCallum M (1972) Related satellite DNAs in the genus *Mus.* J Mol Biol 71:633–656

Trinh TQ, Sinden RR (1993) The influence of primary and secondary DNA structure in deletion and duplication between direct repeats in *Escherichia coli.* Genetics 134:409–422

Vogel JM, Nieto MC, Fischer A, Goodenow RS (1990) Overlapping palindromic sequences associated with somatic deletion and meiotic recombination of MHC class I genes. Mol Immunol 27:875–886

Waldman AS, Liskay RM (1987) Differential effects of base-pair mismatch on intrachromosomal versus extrachromosomal recombination in mouse cells. Proc Natl Acad Sci USA 84:5340–5344

Walker PBM (1971) Origin of satellite DNA. Nature 229:306–308

Walsch JB (1987) Persistence of tandem arrays: implications for satellite and simple-sequence DNAs. Genetics 115:553–567

Warren GJ, Green RL (1985) Comparison of physical and genetic properties of palindromic DNA sequences. J Bacteriol 161:1103–1111

Werren JH (1990) The paternal-sex-ratio chromosome of *Nasonia.* Am Nat 137:392–402

Werren JH, Nur U, Eickbush D (1987) An extrachromosomal factor causing loss of paternal chromosomes. Nature 327:75–76

Werren JH, Nur U, Wu C-I (1988) Selfish genetic elements. Trends Ecol Evol 3:297–302

Wevrick R, Willard VP, Willard HF (1992) Structure of DNA near long tandem arrays of alpha satellite DNA at the centromere of human chromosome 7. Genomics 14:912–923

Willard HF, Waye JS (1987) Hierarchical order in chromosome-specific human alpha satellite DNA. Trends Genet 3:192–198