Chromosoma (1990) 99:243–250

CHROMOSOMA

A novel avian W chromosome DNA repeat sequence in the lesser black-backed gull (*Larus fuscus*)

Richard Griffiths and Peter W.H. Holland

Zoology Dept, Oxford University, South Parks Rd, Oxford, OX1 3PS, UK

Received January 18, 1990 / in revised form April 9, 1990
Accepted April 17, 1990 by W. Hennig

Abstract. The phenol emulsion reassociation technique was used to isolate and clone a female specific, repetitive DNA sequence from *Larus fuscus*. The repeat, designated P2000-17, is restricted to the W chromosome, although related sequences occur elsewhere in the genome of *L. fuscus*. Similar sequences were detected in the genome of six other bird species from outside the genus Laridae, but the sequence occurs less frequently and to a similar extent in both sexes. The 298 bp DNA sequence of P2000-17 was determined and found to have extensive sequence identity to the rabbit dihydropyridine (DHP) receptor calcium channel. P2000-17 is represented once within a larger 8.6 kb tandem repeat (LfW-1), which has a complex internal DNA sequence. LfW-1 is highly conserved between repeat motifs and may comprise 3% of the female genome. The possible evolutionary origin of LfW-1 is discussed in relation to the repeat types found on the W and Y chromosomes of other species.

Introduction

Chromosomal sex determination (CSD) has evolved independently many times. It is characterized by specialized, dimorphic sex chromosomes that are presumed to have evolved from a pair of homologous autosomes (Ohno 1967; Bull 1983). The sex chromosomes are termed X and Y in species where the male is the heterogametic sex and Z and W in female heterogamety; the Y and W chromosomes being unique to the heterogametic sex.

The molecular mechanism that underlies the evolution of CSD has been the subject of many theoretical papers (Muller 1914, 1918; Charlesworth 1978; Hamilton 1967; Rice 1987a, b) yet remains unresolved. In fact, work by Ray-Chaudhuri et al. (1970, 1971) on colubrid snakes and Ansari et al. (1988) on ratite birds, suggests that different evolutionary mechanisms have been involved in different taxa.

Despite its multiple evolutionary origins the effects of CSD on the sex chromosomes are apparently analogous. The Y or W chromosome becomes reduced in size, contains a large proportion of heterochromatin and has few expressed genes. It replicates late in S-phase which, coupled with limited homology, restricts the exchange of Y or W genetic material with that of the X or Z chromosome during crossing over, to small, pseudoautosomal regions. In consequence, a major part of the Y or W chromosome is genetically isolated from the rest of the genome.

The presence of constitutive heterochromatin on the Y or W is revealed by C-banding (Stefos and Arrighi 1971) and is usually associated with the presence of highly repetitive DNA (Stefos and Arrighi 1974; John and Miklos 1979). This suggests that once the Y or W chromosome has been partially freed from evolutionary constraint, there is a competitive scramble for survival amongst the DNA sequences present. The selective bias moves from functional genes to those sequences capable of amplification along the chromosome (Doolittle and Sapienza 1980; Orgel and Crick 1980), resulting in a sequence that gives the Y or W its heterochromatic appearance.

A further characteristic of the sequences on the Y or W chromosome may be the capacity to evolve more rapidly than those on the autosomes (Eicher et al. 1989). This may reflect reduced selective pressure for DNA sequence conservation. However, while selection may not oppose sequence rearrangement, the presence of extra DNA in the Y or W chromosome will increase the energetic burden during DNA replication. The removal of Y or W fragments may therefore be of evolutionary advantage (Charlesworth 1978), and could explain the small size of this chromosome. These properties suggest that ordered rounds of repeat sequence multiplication, and concomitant loss of other DNA sequences, will result in the replacement of many single-copy DNA sequences with a few dominant repetitive sequence types. Furthermore, the appearance on the Y or W chromo-

Offprint requests to: R. Griffiths

some of new types of repeat, either through mutation or translocation from the autosomes, could lead to rapid demographic changes in the types of sequence present.

Evidence to support this model has been provided by Kodama et al. (1987) who found that 50% of the DNA of the W chromosome of the chicken (*Gallus gallus domesticus*) consists of two related repeats, designated the XhoI family. Hybridization studies have shown that this family occurs only in the genus *Gallus* and not in other genera of the same suborder (Tone et al. 1984).

Similarily, the 2.1 and 3.4 kb HaeIII repeats described by Cooke (1976) account for 50% of human Y chromosomal DNA; although related sequences do occur in other apes they are not located on the Y chromosome (Cooke et al. 1982; Kunkel and Smith 1982). This may indicate that the HaeIII repeats are of autosomal origin and through translocation have been able to proliferate in the new environment of the human Y chromosome (Deininger and Daniels 1986).

In both chickens and humans the pattern appears similar: certain repeats which can spread rapidly along the Y or W chromosome but which have limited interchromosomal movement, have recently and successfully colonized the W or Y chromosome. Evidence for repeat evolution along similar lines has also been described in mice (Lamar and Palmer 1984; Eicher et al. 1989) although it is not known how general this pattern is.

Other types of repeat that have been isolated from the Y or W chromosome include the simple tandem GATA sequence (Singh et al. 1980) and a variety of transposable elements that are thought to move about the genome via RNA intermediaries (Finnegan 1989). In general these have been isolated from the Y or W chromosomes of a few well-studied species such as mice, humans and Drosophila, hence it is uncertain if these are present on the Y or W chromosome of all the species that exhibit CSD.

In particular, very little is known concerning DNA sequence organization on the avian W chromosome; only in the chicken has it been extensively investigated at the molecular level. Information on the sequence organization of the W chromosome of other bird species would be valuable in discerning the general rules underlying the evolution of the W chromosome since the origin of CSD. It could also reveal differences that may have arisen due to the independent evolution of CSD in birds.

The species used in this work was the lesser black-backed gull (*Larus fuscus*) which can be sexed with some accuracy by head length (Coulson et al. 1983). Two different methodologies were employed to isolate and characterize sex specific DNA. The first was deletion enrichment (Lamar and Palmer 1984), coupled with the Phenol Emulsion Reassociation Technique (PERT; Kohne et al. 1977; Kunkel et al. 1985). Essentially this is a method of competitive hybridization between DNA of males and females that leaves only the unique sequences of females in a clonable form. This technique yielded the W-specific probe designated P2000-17.

The second method was to compare the restriction patterns of genomic DNA digests of males and females and to investigate the sequence and organization of these via Southern blotting and hybridization.

Materials and methods

Blood sample collection and DNA preparation. All DNA was prepared from blood. Blood samples from the lesser black-backed gull (*L. fuscus;* Order Charadriiformes) and manx shearwater (*Puffinus puffinus;* Order Procellariiformes) were collected from Skomer Island (Dyfed, UK), mixed 1:1 with BLB buffer (250 mM EDTA, 5% SDS, 50 mM Tris pH 8) to prevent DNA degradation and stored frozen at −20° C (samples were routinely stored at 4° C for up to a week before storage).

Herring gull (*Larus argentatus*) samples were collected by Dr. M. Hall from captive birds, previously sexed by laparotomy. The erythrocyte fraction of whole blood was isolated and resuspended in an equal volume of S100ET (100 mM NaCl, 100 mM EDTA, 100 mM Tris-HCl pH 8) and frozen at −20° C.

Whole blood samples from the jackdaw (*Corvus monedula*), great tit (*Parus major*), house sparrow, (*Passer domesticus;* all Order Passeriformes) and the mute swan (*Cygnus olor;* Order Anseriformes) were collected in Oxfordshire (UK), mixed with an equal volume of S100ET and stored at −20° C. Blood samples from the razorbill (*Alca torda;* Order Charadiiformes) were collected on Skomer Island and kindly donated by Mr. R.H. Wagner.

To isolate genomic DNA, 30 µl of the blood/buffer mixture was dispersed in 300 µl of SET (100 mM NaCl, 1 mM EDTA, 100 mM Tris-HCl pH 8) and 30 µl of 10% SDS and 10 µl of 10 mg/ml proteinase K added, before incubation overnight at 55° C. The sample was extracted twice each with equal volumes of phenol, phenol:chloroform and chloroform (24:1 chloroform:isoamyl alcohol) before being ethanol precipitated, washed in 70% ethanol and resuspended in TE (1 mM EDTA, 10 mM Tris-HCl pH 8).

Isolation and cloning of a sex specific DNA. Deletion enrichment of female specific DNA via PERT was performed as described by Kunkel et al. (1985) with modifications suggested by our own preliminary experiments. Fifty nanograms of DNA of *L. fuscus* females was cut to completion with Sau3A. Ten micrograms of DNA of males in 300 µl of TE was sheared by 45 s sonication, using an Ultrasonic Processor (Heat Systems, Ultrasonics, NY), to a mean size of 900 bp and exhaustively dephosphorylated with calf intestinal phosphatase (BCL, UK). The DNA samples of males and females were combined, boiled (10 min) and cooled rapidly on ice. Reassociation (Kohne et al. 1977) was carried out in a final volume of 2 ml, in a 10 ml volumetric flask, in 2 M potassium thiocyanate (KSCN), 7.5% phenol (Ultrapure, BRL) adjusted to pH 8. The flask was rotated constantly at 18 rpm on a Luckham (UK) Tumbler, at room temperature, to maintain the phenol emulsion, for 5 days. The reaction mixture was chloroform extracted, ethanol precipitated and resuspended in 100 µl TE.

Five microlitres of the PERT product was cloned into 10 ng BamHI cut, dephosphorylated pUC13 (Pharmacia, UK) using T4 DNA ligase in a total volume of 10 µl, under conditions specified by the supplier (Gibco BRL, UK).

After transformation into competent DH5α F′ *Escherichia coli* (prepared by the frozen storage method of Hanahan 1985) recombinants were identified by blue/white screening using isopropyl-1-thio-β-D-galactopyranoside (IPTG) and Bluogal (Gibco BRL) followed by restriction analysis of plasmid DNA extracted from positive clones (Serghini et al. 1989). PERT recombinants were then screened by their ability to hybridize to genomic dot blots of DNA from *L. argentatus* and *L. fuscus* of known sex and unsexed individuals of *P. puffinus, C. monedula, P. major, P. domesticus, C. olor* and *A. torda*. Dot blots were prepared from blood treated with proteinase K, as above, but extracted only once with phenol:chloroform. A sub-sample of this, equivalent to 100 nl of whole blood, was immobilized on a Genescreen (Dupont, NEN) membrane by the NaI method (Costanzi and Gillespie 1987).

Recombinant plasmids were labelled by random-primer synthesis (Feinberg and Vogelstein 1984) and hybridized to dot blots in 1.5 × SSC (1 × SSC is 0.15 M NaCl, 15 mM sodium citrate, pH 7.5) 0.1% SDS, 100 μg/ml yeast RNA, and 5 × Denhardt's (0.1% Ficoll 400, 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin) at 48° C (low stringency) or at 58° C (high stringency). After overnight hybridization, filters were washed twice for 20 min at either 50° C (low stringency) or 62° C (high stringency) in 1.5 × SSC, 0.1% SDS, before exposure to X-OMAT film (Kodak) for 30 min to 4 days, at −70° C with two intensifying screens.

DNA sequencing and analyses. Both strands of the P2000-17 insert were sequenced using T7 DNA polymerase and deaza-GTP sequencing mixes following the supplier's instructions (Pharmacia).

The nucleotide sequence was analysed using the Analyseq package (Staden 1984a) to search for open reading frames (ORFs), hairpin loops and inverted and tandem repeats. The method of Fickett (1982) was employed to search for possible protein coding regions and the methods of Shepherd (1981) and Staden (1984b) were used to identify the reading frame most likely to be coding. The FastN algorithm was used to compare the nucleotide sequences of all six possible reading frames of P2000-17 with those in the EMBL database. Using the FastP algorithm (Lipman and Pearson 1985) the amino acid sequences predicted from all reading frames of P2000-17 were compared with protein sequences stored in the National Biomedical Research Foundation Protein Identification Resource.

Genomic DNA digests and the characterization of the female specific DNA repeat. L. fuscus and *L. argentatus* samples were digested with a range of restriction enzymes under conditions recommended by the manufacturers (BRL, BCL). Digests were electrophoresed on 0.4% to 0.75% agarose gels and examined for the presence of sex specific satellite bands. Subsequently, the DNA was transferred (Maniatis et al. 1982) by Southern blotting to Genescreen and probed at both low and high stringency with isolated P2000-17 insert, prepared by a simplified isolation method. Five micrograms of the P2000-17 recombinant plasmid was digested with EcoRI + HindIII and electrophoresed on a 0.75% agarose gel. The insert was cut out of the gel and broken up with a Gilson tip. An approximately equal volume of phenol (equilibrated with TBE) was added and the sample incubated at 60° C for 20–30 min. The sample was then centrifuged for 15 min (16,000 *g*) and the supernatant transferred to a fresh tube, where it was extracted with 1 vol. 24:1 chloroform:isoamyl alcohol at room temperature and recentrifuged for 5 min (16,000 *g*). The supernatant was again removed, precipitated, washed in 70% ethanol and resuspended in distilled water.

Partial digests of *L. fuscus* and *L. argentatus* genomic DNA were performed, blotted and hybridized at high stringency to determine the internal arrangement of the large repeat unit. This was further investigated through a series of single and double restriction enzyme digests of genomic DNA from both *Larus* species which were electrophoresed on agarose gels of 0.4% to 0.75%. Densitometric tracings of photographic negatives of such gels were made with an LKB Ultroscan XL laser densitometer. The tracings were computer analysed (GelScan XL) to estimate the percentage of the female genome of *L. fuscus* that is made up of the large repeat unit.

Results

Cloning of female specific DNA from L. fuscus

Transformation of *E. coli* DH5α F′ with the product of the PERT ligation yielded 62 colonies, of which 25 were identified as recombinants by blue/white screening. The recombinant plasmids were isolated and the potential inserts separated by agarose gel electrophoresis after restriction digested with EcoRI + HindIII. One plasmid, designated P2000-17, contained an insert of about 300 bp which subsequently showed female specific hybridization, at high stringency, to genomic dot blots of *L. fuscus* and *L. argentatus*. Figure 1 illustrates the *L. argentatus* dot blot: of the 33 individuals probed with P2000-17, all 12 females showed a similar, high degree of hybridization while in the 21 males a signal was visible but much reduced.

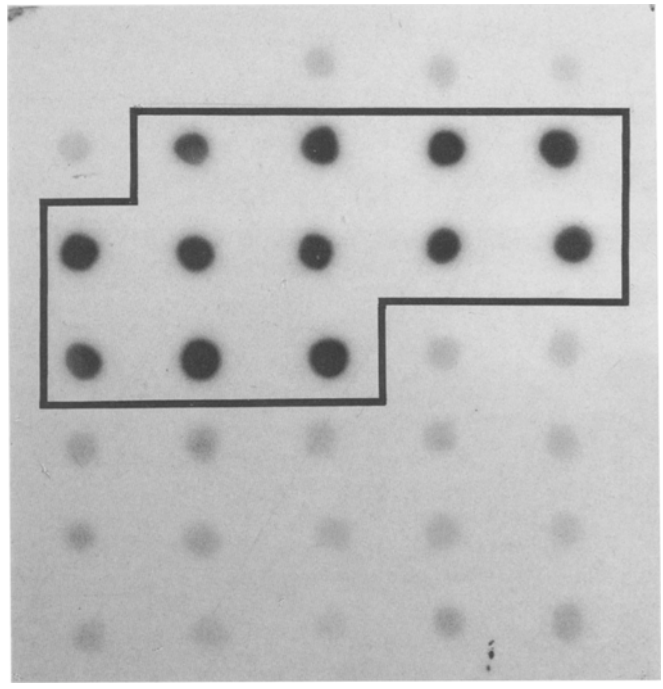

Fig. 1. Dot blots of genomic DNA from 33 *L. argentatus* individuals hybridized to ^{32}P-labelled P2000-17 at high stringency. The dots enclosed by a black line are females the rest males

Genomic dot blots of the DNA from several unsexed *P. puffinus, P. major, P. domesticus, C. monedula, C. olor* (all species from outside the order Charadiiformes) and *A. torda* (a Charadiiforme) were also probed (results not shown). At high stringency significant hybridization was detected only in *A. torda* where it occurred to a similar degree in all 12 individuals probed. The level of hybridization in *A. torda* is approximately five-fold less than that of male Laridae. In the five non-Charadiiforme species hybridization occurred only at low stringency and to a much lesser extent than that of *A. torda;* again there was no evidence of sex specific hybridization. It may be inferred from the taxonomic distribution of the six species that sequences related to P2000-17 occur widely, though apparently not sex specifically, in the class Aves. Moreover the evolutionary distance from the Laridae appears to be reflected in a divergence or relative scarcity of the complementary sequence.

Sequence analysis of P2000-17

Nucleotide sequence determination revealed that the P2000-17 insert is 298 bp which includes the two flanking Sau3A sites. Analysis of the DNA sequence gave

```
1    GATCAAAGCATAGATAAAGTCGAACCTGACTCAACTCCACAGACT   AGTCT    50
3032 CCA-CTTCG-GG---GGCCC-AGCTGCTGTA--GGG---T-----CCA-CGAG  3084

51   GGTTCCACTGCCACTATCTGAAACCCCCGA     CAATGACACCTTATTGAG   98
3085 -AGGA--TG-G-C-CG-T-AC---AA----GTGGAG-TG-C--T---C--C-T 3137

99   C  GGACTTAAGAATC  TGCC CCTTGATGATGAGCTTTGGGATAATGCACT  146
3138 -ATCT--A-C-TCC--AT----TT---C-------A-A-CTTTG- GG--TT- 3189

147  GTCACGATACACTGCAAGCACTGAGACATTAGAAAACATGAGGGATTTAAATA  199
3190 ----TCG-CAC--T-C--    --- --GGG-G-G--- ---TACAAG--C-G 3236

200  CACCTGCCCTCCTCTCCCCAGCGCCAGTGTGTGCAGTATGCGCTGAAGGCTCG  254
3237 -  GA--TGGA-AAGAA-----------------------C--------C-- 3287

255  CCCGCTGCGGCGCTACATCCCCAAGAACCCCTACCAGTACCAGATC         298
3288 ---A--T---T-------------------A------------G-G       3333
```

Fig. 2. The nucleotide sequence of P2000-17 (upper sequence) and that of the DHP receptor cDNA (lower sequence; Tanabe et al. 1987). Dashes (−) represent nucleotide identity and gaps have been introduced to maximize sequence alignment. The underlined portion of the DHP receptor cDNA is the S6 putative membrane spanning coding region

no evidence for internal tandem or inverted repeats, or for unusual base composition. The sequence has one continuous ORF on one strand with potential to code for an amino acid sequence rich in proline and acidic residues. Fickett's (1982) method produced a testcode suggesting that the strand containing the ORF was a protein coding region. Analysis of reading frame by the RNY preference method of Shepherd (1981) and the positional base preference method of Staden (1984b) indicated that the ORF had the nucleotide sequence most characteristic of a protein coding region.

When the two strands of the sequence were used to search the EMBL nucleotide sequence database, a significant match was found to the rabbit mRNA for dihydropyridine (DHP) receptor (Tanabe et al. 1987). This mRNA contains a 6,083 bp region that encodes a voltage sensitive calcium channel. It consists of four homologous regions (I–IV) that each contain six putative membrane spanning sequences (S1–6). The region of sequence identity spans the last part of unit III which includes S6, the interunit region and the first 18 bp of unit IV.

When gaps are introduced to maximize alignment, but counted as mismatches, P2000-17 and the DHP receptor share 55.6% identity in a 311 bp overlap (Fig. 2). The region of best fit occurs mainly over the interunit region where the sequences share 92.4% identity over 79 nucleotides.

The alignment shown interrupts the ORF of P2000-17 indicating that the two sequences cannot produce similar protein products. However, the high level of nucleotide sequence similarity may indicate evolutionary links between P2000-17 and the DHP receptor.

Molecular organization of the female specific repeat

The experimental results described in this section are indistinguishable between the two gull species *L. fuscus* and *L. argentatus*. This indicates a high degree of sequence conservation and accordingly the figures and results presented are representative of both species.

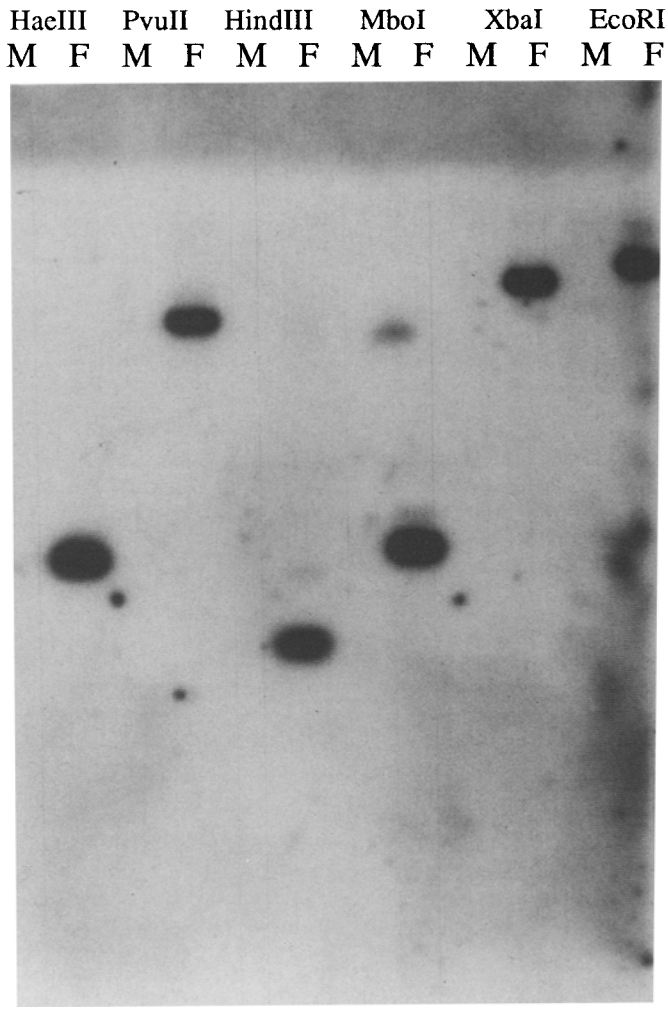

Fig. 3. Southern blots of an *L. argentatus* male (M) and female (F) genomic DNA cut with a variety of restriction enzymes and hybridized to ^{32}P-labelled P2000-17 at high stringency. The sizes of the bands are: HaeIII, 0.85 kb; PvuII, 3.9 kb; HindIII, 0.42 kb; MboI, 0.78 kb; XbaI, 7 kb; EcoRI, 8.6 kb

Figure 3 illustrates a high stringency hybridization of P2000-17 to restriction digests (HaeIII, PvuII, HindIII, MboI, XbaI and EcoRI) of *L. argentatus* genomic DNA. Hybridization is exclusive to single bands in lanes containing DNA of females which corroborates the evidence from the dot blots to suggest that the sequence is primarily sex linked and so associated with the W chromosome. The hybridization of P2000-17 to dot blots of genomic DNA of males does indicate that sequences with some complementarity exist elsewhere in the gull genome. However, low stringency hybridizations to digests of DNA of males yield inconsistent results, some of which suggest that the related sequences are located on the Z chromosome, although this requires further elucidation.

When restriction digests of genomic DNA of males and females are run on agarose gels, staining with ethidium bromide reveals female specific satellites (Fig. 4).

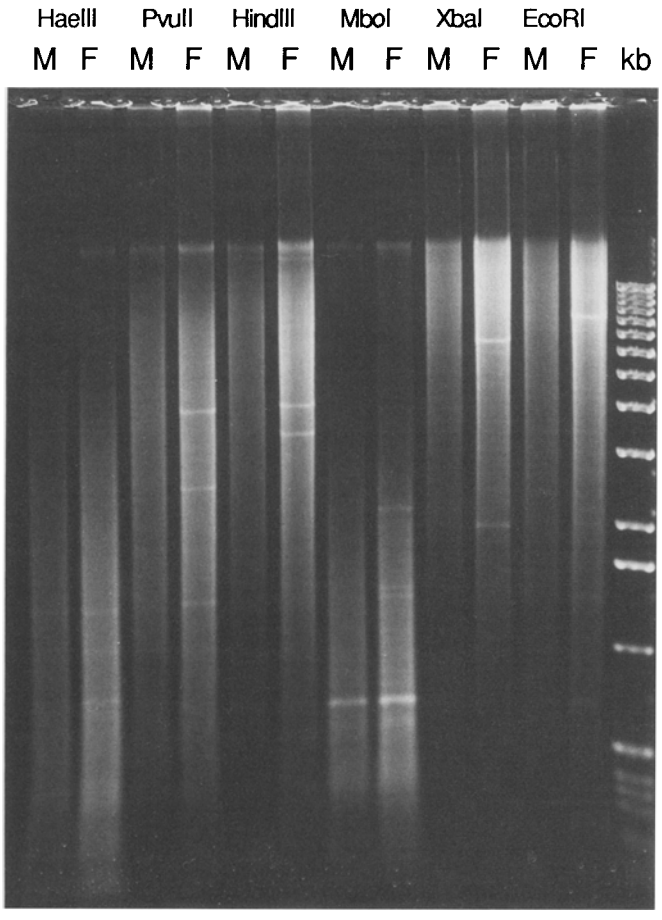

Fig. 4. An ethidium bromide stained, 0.5% agarose gel of *L. fuscus* genomic digests showing female specific repeats. The lanes contain male (M) and female (F) DNA digested with the enzymes shown except the final lane which contains molecular size markers descending in 1 kb intervals from 12 to 2 kb, then 1.6, 1, 0.5, 0.4, 0.35 and 0.3 kb

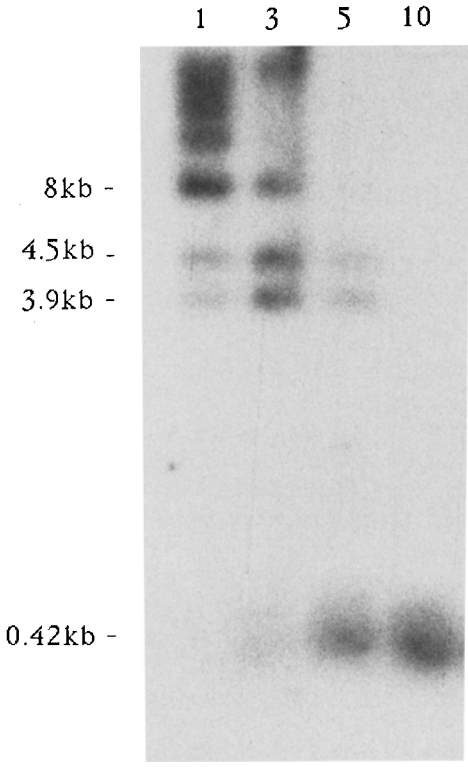

Fig. 5. A Southern blot of a series of HindIII digests of *L. fuscus* genomic DNA2000-17 stopped after the specified times (min) to give partial digestion and hybridized to ^{32}P-labelled P2000-17 at high stringency

The sizes of these visible satellites were compared with those of the bands obtained by Southern blot hybridization with the P2000-17 probe. The visible satellites produced by digestion with HaeIII (0.85 kb), PvuII (3.9 kb), XbaI (7.0 kb) and EcoRI (8.6 kb) correspond in size to the single bands that occur on autoradiographs after the gel is blotted and probed with P2000-17 (Fig. 3). Such is also the case with BamHI digests, where the fragment is of 8.6 kb, and PstI (8.6 kb), XhoI (4.3 kb) and HincII (4 kb; results not shown) digests. In contrast there are no visible satellites corresponding to the 0.78 kb MboI and 0.42 kb HindIII bands that hybridize to P2000-17 (Fig. 4). It is likely that the former is obscured by another satellite of this size common to both sexes, while the latter being of small size and thus low effective concentration is also present but not discernable. In conclusion, it appears that P2000-17 forms part of a larger repeated sequence.

Autoradiographs show a single cognate band for each different restriction digest (with the exception of Sau3A, which will be discussed later; results not shown). The size of this band varies from 0.42 kb with HindIII to the large EcoRI fragment of 8.6 kb. This may indicate that HindIII internally cleaves a repetitive tandem array within the EcoRI fragment or alternatively that the EcoRI fragment is more complex and HindIII digestion releases a single 0.42 kb sequence from each repeat that is recognized by P2000-17.

We favoured the latter explanation as digests with some enzymes, such as HindIII and XbaI, produced additional sex specific satellites that did not hybridize to P2000-17 (Figs. 3 and 4). In contrast, EcoRI-cut, DNA of female gulls produced a single band of similar size on both gel and autoradiograph. We hypothesized that the visible satellites that did not correspond to bands on the autoradiograph were other fragments of the same large repeat unit that had a sequence unrelated to that of P2000-17.

This hypothesis was confirmed by the Southern blot hybridization of P2000-17 to partial HindIII digests of *L. fuscus* DNA (Fig. 5). After a 3 min digestion there are four major bands that hybridize to P2000-17. The largest is undigested genomic DNA while the two smallest (4.5 and 3.9 kb) correspond in size to those visible satellites seen in HindIII digests (4.1 and 3.5 kb; Fig. 4) plus the 0.42 kb band that labels with P2000-17 in blots of completed HindIII digests (Fig. 3). The remaining band, of 8 kb, demonstrates that the 4.1 and 3.5 kb fragments flank the 0.42 kb unit as it represents the three before internal cleavage. As the reaction progresses the 0.42 kb unit is progressively removed from these flanking sequences to give a single cognate band of 0.42 kb.

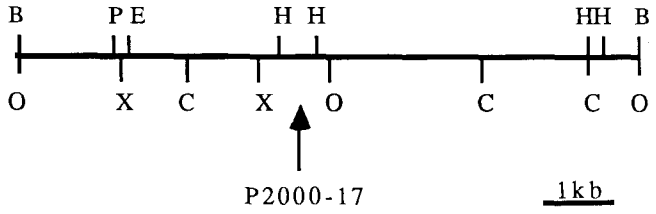

Fig. 6. The restriction map of the 8.6 kb tandem repeat LfW-1. B, BamHI; C, HincII; E, EcoRI; H, HindIII; O, XhoI, P, PstI; X, XbaI

The sequence recognized by P2000-17 is thus represented once in each 8.6 kb repeat unit.

A series of double restriction digests, with enzymes found to cut the 8.6 kb motif relatively infrequently, were run on agarose gels. These demonstrated that the repeat is tandemly arranged and yielded a restriction map (Fig. 6) of the basic repeat unit.

We estimate, from densitometric analysis of the photographic negative for Figure 4, that the repeat accounts for approximately 3% of the total female genome. Given that the size of the diploid genome of *L. argentatus* is 8.24 pg (Venturini et al. 1987) then the repeat is present as 25,500 copies.

Hybridization of P2000-17 to a Sau3A digest of gull DNA reveals two bands. One is of the same size as that produced by the enzyme's isoschizomer, MboI, although both bands are fainter. One explanation would be the partial methylation of cytosine residues which would prevent complete digestion by Sau3A (but not MboI whose action is unaffected by the presence of 5-methylcytosine) and so yield a second band.

The origin of the 298 bp P2000-17 insert itself is unclear. It was derived from a Sau3A *L. fuscus* digest but recognizes Sau3A bands of 750 and 900 bp rather than 298 bp. One possibility is that the 780 bp fragment possesses an internal Sau3A restriction site that is present in only a few of the repeat units; P2000-17 would then represent a minor fraction.

Discussion

To clone the female specific P2000-17 repeat sequence we used a method similar to that of Kunkel et al. (1985) which combines deletion enrichment (Lamar and Palmer 1984) with PERT (Kohne et al. 1977). In theory the strategy should give a high yield of recombinant clones enriched for the sequences specific to the trace DNA.

In practice, Kunkel et al. (1985) found DNA yields were about 5% of input, which they attributed to loss of partially single-stranded DNA during purification. Kohne et al. (1977) encountered a similar problem which they associated with adsorption of DNA to the walls of the reaction vessel. In the present study where the amount of DNA available was small this did present a technical problem. To maintain a high driver to trace ratio we used 50 ng of female *L. fuscus* DNA, much less than used by Kunkel et al. (1985). Assuming a 95% loss (which may be an underestimate, as adsorption may be related to the surface area of the reaction vessel, not DNA concentration) only 2.5 ng of DNA would be recovered. We calculate that as only a small proportion of this originates from the W chromosome and still less would have reassociated correctly, that yields of clonable material would be in the low picogram range. This is a serious limitation of the technique; hence our success in cloning female specific DNA should not be taken as an indication of the general applicability of PERT.

A further problem is that this method will discriminate against the isolation of repeats with a short, conserved basic unit. Discrimination would occur during reassociation where the chance of a misalignment between the repeat units is high; hence the specific cohesive termini are less likely to reform to allow cloning. Since this type of repeat is thought to be abundant and sometimes sex linked it may constitute another shortcoming of the technique. However, where more complex sequences are the subject of investigation this may well be an advantage.

The W chromosomes of all the snakes with CSD, so far studied, are characterized by the presence of the simple GATA or Bkm repeat (Singh et al. 1976, 1980). Similarly the only sex specific repeat previously reported from birds is the chicken XhoI family which is based on a consensus 21 bp motif (Kodama et al. 1987). In contrast, the results described in this paper provide evidence for a very different type of W chromosome specific repeat which is present in the gulls *L. fuscus* and *L. argentatus*. In this study, PERT resulted in the isolation of a 298 bp female specific clone, P2000-17, with no internal repeats. We have shown that the clone represents an internal fragment of an 8.6 kb, tandemly repeated sequence, which we shall refer to as LfW-1 (*L. fuscus* W chromosome repeat 1).

The size of LfW-1 is greater than many repeat sequences, although Steele et al. (1986) have described a human repeat that is larger. The latter is derived from both retroviral and flanking cellular sequences that have been amplified as a single unit through an unknown mechanism. However, this human sequence is located on several different chromosomes and probably does not occur as a tandem array.

Like the repeat described by Steele et al. (1986), it has often been found that larger repeat units are based on sequences that have arisen as a result of selection for function. With this in mind, we investigated the possible origin of LfW-1 by determining the nucleotide sequence of its internal fragment P2000-17. The base sequence demonstrates that LfW-1 has not evolved via simple serial duplication of small subunits and so may indeed have evolved from a complex DNA sequence, prior to any tandem duplication events. Analysis of the P2000-17 sequence revealed several unusual features, and allowed the formation of at least two alternative hypotheses.

First, P2000-17 has an ORF throughout its length which has a nucleotide sequence typical of a protein coding region. This suggests that part of LfW-1 may either have protein coding potential, or have recently evolved from a functional protein coding gene.

A second possibility is suggested by the extensive sequence identity between P2000-17 and the rabbit DHP receptor. This raises the intriguing possibility that they

have evolved from a common ancestral ion channel sequence. It is important to note that the optimum sequence alignment does not maintain the ORF of P2000-17, so implying that the two hypotheses are mutually exclusive.

It is unlikely that LfW-1 is a functional gene because of its repetitive organization and its presence on the W chromosome (which by analogy with the mammalian Y chromosome may encode few functional genes). Consequently we have explored the possibility that LfW-1 is based on a DHP receptor pseudogene or processed pseudogene, and there is some circumstantial evidence to support this. First, the voltage sensitive channels are a highly conserved group (Baumann et al. 1988; Campbell et al. 1988) so the similarity between the rabbit DHP receptor and LfW-1 may reflect a similar relationship between LfW-1 and an extant avian voltage sensitive channel. Second, the DHP receptor is part of a multigene family which is thought to have evolved from a single ancestral gene (Ramaswami and Tanouye 1989). The amplification required to produce the contemporary gene family is also likely to produce pseudogenes, a course of events known to have given rise to several mammalian gene families (Wagner 1986; Vanin 1985). Third, the DHP receptor gene is large (the mRNA is 6.5 kb; Ellis et al. 1988) so it could form a large proportion of LfW-1. Finally, if tandem duplication on the W chromosome occurs through unequal crossover (Smith 1976) then the four homologous regions of the DHP receptor may facilitate this mechanism.

Although the LfW-1 repeat seems to be specific to the W chromosome, there are related sequences that hybridize to P2000-17 elsewhere in the genome. In humans it has recently been shown that transposition between the non-homologous regions of the X and Y chromosomes does occur (Cooke et al. 1984; Page et al. 1984). Such a mechanism may have resulted in the transposition of an ancestral LfW-1 sequence between the W and Z chromosomes, to be followed by independent diversification and amplification. A possible alternative is the retroposition of different processed pseudogenes, either from the same or related functional genes, throughout the genome followed by amplification on the W chromosome. Processed pseudogenes of the human argininosuccinate synthetase gene have become associated with autosomes and X and Y chromosomes through retroposition (Su et al. 1984), although no amplification is apparent.

We estimate that approximately 3% of the genome of the female gull is made up of LfW-1 repeats and suggest that a large proportion of these are situated on the W chromosome. This implies that the gull W chromosome is at least twice the estimated size of that of the chicken (Tone et al. 1982). Comparison with other species is difficult because of a paucity of information; similarly we have been unable to find any karyotypic data on the size of the W chromosome in the Laridae.

P2000-17 related sequences were also detected in the genomes of all the avian species examined. This may be analogous to the widespread occurrence of the GATA repeat in a variety of vertebrates and invertebrates (Singh et al. 1981). However, the phylogenetic distribution of GATA may owe more to convergent evolution than sequence conservation (Levinson et al. 1985) because of the simple repetitious nature of GATA. This would be unlikely with P2000-17.

One possibility is that the related sequences detected in other species represent an avian DHP receptor gene. If the sequence of this were similar to that of the rabbit it would hybridize to P2000-17, irrespective of whether the sequence identity of P2000-17 and the rabbit DHP receptor is a result of divergent or convergent evolution.

The restriction patterns of LfW-1 from *L. fuscus* and *L. argentatus* are identical. Whilst this indicates a high level of sequence conservation it must be realized that they are a very closely related species and able to produce fertile hybrids (Tinbergen 1960). Sequence conservation and occurrence in one, or a few, related species is characteristic of other Y or W chromosome specific repeats which include the human HaeIII family (Cooke et al. 1982), the chicken XhoI family (Tone et al. 1984) and a range of mouse Y specific repeats (Bishop et al. 1985; Lamar and Palmer 1984; Eicher et al. 1989). Together these observations suggest that recent evolutionary origin and amplification is a characteristic of many Y or W chromosomal repeats. We therefore suggest that Y or W chromosomal DNA evolves in a cyclical manner, where sequences of widely varying complexity, which either originate on or are transposed to the sex chromosomes, undergo amplification, diversification and degeneration.

Acknowledgements. Thanks are due to Dr. M. Hall for blood samples of *L. fuscus*, Mr. R.H. Wagner for those of *A. torda*, Dr. M.E. Bramwell for his help with laser densitometry, Dr. S.E. Kearsey and Dr. W.R. Brown for their comments on the manuscript and Dr. C.M. Perrins and Ms N.A. Williams for their help and advice. RG is supported by an SERC grant.

References

Ansari HA, Takagi N, Sasaki M (1988) Morphological differentiation of sex chromosomes in three species of ratite birds. Cytogenet Cell Genet 47:185–188

Baumann A, Grupe A, Ackermann A, Pongs O (1988) Structure of the voltage-dependent potassium channel is highly conserved from Drosophila to vertebrate central nervous systems. EMBO J 7:2457–2463

Bishop CE, Boursot B, Bonhomme F, Hatat D (1985) Most classical *Mus musculus domesticus* laboratory mouse strains carry a *Mus musculus musculus* Y chromosome. Nature 315:70–72

Bull JJ (1983) Evolution of sex determining mechanisms. Benjamin/Cummings, Menlo Park, Calif

Campbell KP, Leung AT, Sharp AH (1988) The biochemistry and molecular biology of the dihydropyridine-sensitive calcium channel. Trends Neurol Sci 11:425–430

Charlesworth B (1978) Model for the evolution of Y chromosomes and dosage compensation. Proc Natl Acad Sci USA 75:5618–5622

Cooke HJ (1976) Repeated sequence specific to human males. Nature 262:182–186

Cooke HJ, Schmidke J, Gosden JR (1982) Characterization of a human Y chromosome repeated sequence and related sequences in higher primates. Chromosoma 87:491–502

Cooke HJ, Brown WAR, Rappold GA (1984) Closely related sequences on human X and Y chromosomes outside the pairing region. Nature 311:259–261

Costanzi C, Gillespie D (1987) Fast blot immobilization of DNA and RNA from cells. Methods in Enzymol 152:582–588

Coulson JC, Thomas CS, Butterfield JEL, Duncan N, Monaghan P, Sledder C (1983) The use of head and bill length to sex live gulls, Laridae. Ibis 125:549–557

Deininger PL, Daniels GR (1986) The recent evolution of mammalian repetitive DNA elements. Trends Genet 2:76–80

Dootlittle WF, Sapienza C (1980) Selfish genes, the phenotype paradigm and genome evolution. Nature 284:601–603

Eicher EM, Hutchinson KW, Philips SJ, Tucker PK, Lee BK (1989) A repeated segment on the mouse Y chromosome is composed of retroviral-related, Y enriched and Y specific sequences. Genetics 122:181–192

Ellis SB, Williams ME, Ways NR, Brenner R, Sharp AH, Leung AT, Campbell KP, McKenna E, Koch WJ, Hui A, Schwartz A, Harpold MM (1988) Sequence and expression of mRNA's encoding the 1 and 2 subunits of a DHP-sensitive calcium channel. Science 241:1661–1664

Feinberg AP, Vogelstein B (1984) Addendum: a technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 137:266–227

Fickett JW (1982) Recognition of protein coding regions in a DNA sequence. Nucleic Acids Res 10:5303–5318

Finnegan DJ (1989) Eukaryotic transposable elements and genome evolution. Trends Genet 5:103–107

Hamilton WD (1967) Extraordinary sex ratios. Science 156:477–488

Hanahan D (1985) Techniques for transformation of *E. coli*. In: Glover DM (ed) DNA cloning vol I: a practical approach. IRL press Oxford, p 109–135

John B, Miklos BLB (1979) Functional aspects of satellite DNA and heterochromatin. Int Rev Cytol 58:1–23

Kodama H, Saitoh H, Tone M, Kuhara S, Sasaki Y, Mizuno S (1987) Nucleotide sequences and unusual electrophoretic behavior of the W chromosome-specific repeating DNA units of the domestic fowl, Gallus gallus domesticus. Chromosoma 96:18–25

Kohne DF, Levinson SA, Byers MJ (1977) Room temperature method for increasing the rate of DNA reassociation by many thousand fold: the phenol emulsion reassociation technique. Biochemistry 16:5329–5341

Kunkel LM, Smith KD (1982) Evolution of human Y chromosome DNA. Chromosoma 86:209–228

Kunkel LM, Monaco AP, Middlesworth W, Ochs HD, Latts SA (1985) Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. Proc Natl Acad Sci USA 82:4778–4782

Lamar EE, Palmer E (1984) Y-encoded, species specific DNA in mice: evidence that the Y chromosome exists in two polymorphic forms in inbred strains. Cell 37:171–177

Levinson G, Marsh JL, Epplen JT, Gutman GA (1985) Cross hybridizing snake minisatellite, *Drosophila* and mouse DNA sequences may have arisen independently. Mol Biol Evol 2:494–504

Lipman DJ, Pearson WR (1985) Rapid and sensitive protein similarity searches. Science 227:1435–1441

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

Muller HJ (1914) A gene for the fourth chromosome of *Drosophila*. J Exp Zool 17:324–337

Muller HJ (1918) Genetic variability, twin hybrids and constant hybrids, in a case of balanced lethan factors. Genetics 3:422–499

Ohno S (1967) Sex chromosomes and sex linked genes. Springer, Berlin Heidelberg New York

Orgel LE, Crick FH (1980) Selfish DNA: the ultimate parasite. Nature 284:604–607

Page DC, Harper ME, Love J, Botstein D (1984) Occurrence of a transposition from the X-chromosome long arm to the Y-chromosome short arm during human evolution. Nature 311:119–123

Ramaswami M, Tanouye MA (1989) Two sodium-channel genes in *Drosophila:* implications for channel diversity. Proc Natl Acad Sci USA 86:2079–2082

Ray-Chaudhuri SP, Singh L, Sharma T (1970) Sexual dimorphism in somatic interphase nuclei of snakes. Cytogenetics 91:140–423

Ray-Chaudhuri SP, Singh L, Sharma T (1971) Evolution of sex chromosomes and formation of W chromatin in snakes. Chromosoma 33:239–251

Rice WR (1987a) Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. Genetics 116:161–167

Rice WR (1987b) The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. Evolution 41:911–914

Serghini MA, Ritzenthaler C, Pinck L (1989) A rapid and efficient 'miniprep' for the isolation of plasmid DNA. Nucleic Acids Res 17:3604

Shepherd JCW (1981) Method to determine the reading frame of a protein from the purine/pyrimidine genome sequence and its possible evolutionary justification. Proc Natl Acad Sci USA 78:1596–1600

Singh L, Purdom IF, Jones KW (1976) Satellite DNA and evolution of sex chromosomes. Chromosoma 59:43–62

Singh L, Purdom IF, Jones KW (1980) Sex chromosome associated satellite DNA: evolution and conservation. Chromosoma 79:137–157

Singh L, Purdom IF, Jones KW (1981) Conserved sex chromosome associated nucleotide sequences in eukaryotes. Cold Spring Harbor Symp Quant Biol 45:805–814

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

Staden R (1984a) Graphic methods to determine the function of nucleic acid sequences. Nucleic Acids Res 12:521–538

Staden R (1984b) Measurement of the effects that coding for a protein has on a DNA sequence and their use for finding genes. Nucleic Acids Res 12:551–567

Steele PE, Rabson AB, Bryan T, O'Brien SJ (1986) Amplification and chromosomal dispersion of human endogenous retroviral sequences. J Virol 59:545–550

Stefos K, Arrighi FE (1971) The heterochromatic nature of the W chromosome in birds. Exp Cell Res 68:228–231

Stefos K, Arrighi FE (1974) Repetitive DNA of Gallus domesticus and its cytological locations. Exp Cell Res 83:9–14

Su TS, Nussbaum RL, Airhart S, Ledbetter DH, Mohandas T, O'Brien WE, Beaudet AL (1984) Human chromosomal assignments for 14 argenininosuccinate synthetase pseudogenes: cloned DNA's as reagents for cytogenetic analysis. Am J Hum Genet 36:954–964

Tanabe T, Takeshima H, Mikami A, Flockerzi V, Takahashi H, Kangowa K, Kojima M, Matsuo H, Hirose T, Numa S (1987) Primary structure of the receptor for calcium channel blockers from skeletal muscle. Nature 328:313–328

Tinbergen N (1960) The herring gull's world. Harper and Row, Boston

Tone M, Nakano N, Takao E, Narisawa S, Mizuno S (1982) Demonstration of W chromosome-specific repetitive DNA sequences in the domestic fowl *Gallus g. domesticus*. Chromosoma 86:551–569

Tone M, Sakaki Y, Hashiguchi T, Mizuno S (1984) Genus specificity and extensive methylation of the W chromosome-specific repetitive DNA sequences from the domestic fowl, *Gallus gallus domesticus*. Chromosoma 89:228–237

Vanin EF (1985) Processed pseudogenes: characteristics and evolution. Annu Rev Genet 19:253–272

Venturini G, Capanna E, Fontana B (1987) Size and structure of the bird genome II. Repetitive DNA and sequence organization. Comp Biochem Physiol 87B:975–979

Wagner M (1986) A consideration of the origin of processed pseudogenes. Trends Genet 2:134–137