J Mol Evol (1984) 20:322–329

Journal of
Molecular Evolution

Unexpectedly Slow Homogenisation within a Repetitive DNA Family Shared Between Two Subspecies of Tsetse Fly

Martin Trick and Gabriel A. Dover

Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK

Summary. Repetitive DNA families in sexual species are subject to a variety of turnover mechanisms capable of homogenising newly arising mutations. Very high levels of homogeneity in DNA families in some species of *Drosophila* indicate that the rate of turnover is fast relative to that of mutation. To gauge the generality of such phenomena, we cloned and sequenced individual members of homologous repetitive DNA families from two subspecies of tsetse fly, *Glossina morsitans centralis* and *G. morsitans morsitans*. Unexpectedly high levels of variation were found within each subspecies, averaging 24% and 31%, respectively. Contiguous repeats and repeats cloned at random were comparably divergent. Nevertheless, it was possible to identify three instances of apparent homogenisation, each being, remarkably, of an insertion/deletion nature. We conclude that the rate of turnover in the tsetse families is comparable to that of most mutations, and discuss the possible parameters affecting flux in these families.

Key words: Repetitive DNA families — Homogenisation — Sequence variation — Tsetse fly species

Introduction

It has been recognised that the multigene and noncoding repetitive DNA families of eukaryotic genomes are in a state of turnover due to novel genetic exchanges (Hood et al. 1975; Kedes 1979; Ohta 1980; Dover 1982; Dover et al. 1982, 1984; Jones and Kafatos 1982; Arnheim 1983; Ohta and Dover 1984). These stem mainly from the apparent propensity of member genes to communicate non-reciprocally, and thereby homogenise, species-specific variants throughout the majority of constituent members and in most individuals of a sexual population (molecular drive). The mechanism of homogenisation is likely to vary with the genomic organisation of the family; unequal exchange, gene conversion and transposition have each been implicated in different systems of family homogeneity (for references, see Dover 1982).

The unequal exchange model (Smith 1976) seems to fit in an economical manner most of the available data on the evolution of tandemly repeated sequences. Unequal mitotic and meiotic exchanges have been directly demonstrated in yeast and *Drosophila melanogaster* rDNA arrays (Petes 1980; Szostak and Wu 1980; Coen and Dover 1983). Sister chromatid exchanges have been shown to occur in the heterochromatin of *D. melanogaster* (Tartof 1974; Yamamoto and Miklos 1978), which is characterised by a concentration of highly repeated DNA families (Peacock et al. 1977; Brutlag 1981).

Experimental investigation of variation amongst members of noncoding, tandem DNA families can be used to assay the results of such homogenising processes in the probable absence of selective constraints; the resolution offered by direct DNA sequencing techniques allows the fine details of both modes and rates of change to be elucidated.

Evidence of a pervasive, low level of variation has been obtained in the African Green Monkey α-satellite DNA (Thayer et al. 1981). In these studies, sequencing of cloned dimeric fragments indicated that adjacent repeats are not necessarily more similar than are repeats selected at random. Extensive sequence analysis in repetitive DNAs has been car-

Offprint requests to: G.A. Dover

ried out both within and between species of Hawaiian *Drosophila* (Miklos and Gill 1981) and in the *melanogaster* species subgroup of *Drosophila* (Strachan et al. 1982; T. Strachan, D.A. Webb, and G.A. Dover, submitted for publication). The latter studies, in particular, have shown the within-species variation in two DNA families to be an order of magnitude lower than most of the between-species divergences, indicating that the primary rate of turnover (i.e., rate of homogenisation) is considerably faster than that of mutation in these genomes. This disparity in rates allows the classic observation of concerted evolution to be made.

To gauge the wider biological significance of homogenisation, we set out to test whether the same principles applied to repetitive DNA families in insect genera other than *Drosophila*. The present studies relate to the evolution of a family shared between subspecies of the tsetse fly *Glossina morsitans*. Some aspects of repeated sequence evolution within this species group have already been described (Amos and Dover 1981). In this paper we present an examination of nucleotide variation amongst contiguous repeats from a cloned region of a DNA family in *G. morsitans centralis* and between their counterparts, cloned at random, from the subspecies *G. morsitans morsitans*. These two subspecies are morphologically (Potts 1970) and cytologically (Southern and Pell 1973) very similar and are capable of interbreeding in the laboratory to produce some fertile offspring (Curtis 1972). Limited gene flow may still continue between the two forms in the wild, in that a sympatric zone of overlap (Curtis 1972) and the occurrence of natural hybrids (Vanderplank 1944, 1947) have been reported. A study of restriction site differences within a defined segment of the mitochondrial genomes of the flies has tentatively estimated the separation time of the two lineages as 1 million years ago (Trick and Dover 1984).

In this study, we report unexpectedly high levels of sequence variation, averaging 30% divergence between repeats, both within and between the subspecies, although certain sequence motifs of an insertion/deletion nature appear to have been homogenised between the two. We discuss possible reasons for this apparent low rate of turnover (relative to mutation) in the tsetse family of sequences.

Materials and Methods

DNA Extraction. Newly deposited pupae of *G. morsitans centralis* and *G. morsitans morsitans* were supplied by the Tsetse Research Laboratory, Bristol, U.K. from their breeding colonies (Mews et al. 1977). DNA was extracted from 20–30 pupae at a time essentially according to the method of Coen et al. (1982) except that an additional phenol extraction step was included.

Restriction Analysis. All restriction enzymes used were purchased from commercial suppliers and used in accordance with the manufacturers' recommendations. Digests were electrophoresed on horizontal agarose gels submerged in TBE buffer (89 mM Tris, 2.5 mM disodium ethylenediaminetetraacetate (Na_2EDTA), 89 mM boric acid) and transferred to nitrocellulose filters according to Southern (1975), with an initial depurination step (Wahl et al. 1979). Plasmid DNA was nick translated (Rigby et al. 1977) with [^{32}P]dATP (3000 Ci mmol^{-1}; New England Nuclear). Hybridisations were conducted at 42°C overnight with approximately 10^7 cpm activity of probe in 50% formamide, 5× SSC (1 × SSC = 0.15 M NaCl and 0.015 M Na citrate), 0.5% sodium dodecyl sulphate (SDS). Filters were washed extensively at room temperature in 3 mM Tris (unneutralised) after hybridisation. Washed filters were dried and autoradiographed at −70°C using pre-flashed Fuji X-ray film backed by Mach II intensifying screens.

Cloning Procedures. Plasmid pGc22 was isolated by colony hybridisation (Grunstein and Hogness 1975) from a bank of recombinants of *G. morsitans centralis* genomic DNA constructed by ligating Eco RI-digested DNA into phosphatase-treated pAT153 vector (Twigg and Sherratt 1980). The constituent 160-bp repeats, liberated from the vector by Rsa I digestion, were prepared by electroelution (McDonell et al. 1977) and recloned into phosphatase-treated, Sma I-digested M13mp9 vector (Messing 1981). Homologous Rsa I repeats from *G. morsitans morsitans* DNA were directly cloned into the Sma I site of the M13 vector. Ligated DNAs were used to transform competent JM101 cells and single-stranded DNA templates were prepared from phage giving white plaques on indicator plates.

DNA Sequencing. M13 clones were sequenced by the dideoxynucleotide method (Sanger et al. 1977) using a 15-bp universal primer (New England Biolabs). Under the conditions employed, 200 bases could be read from the cloning site. Computer programs to analyse and compare sequence data were written in FORTRAN and executed on the University of Cambridge IBM 3081 computer. Matrix plots were generated by implementing appropriate routines from the Cambridge University Computing Service's graphics package.

Results

Isolation and Characterisation of pGc22

The recombinant plasmid pGc22 was isolated from an Eco RI-generated bank of *G. morsitans centralis* sequences by colony hybridisation with a total DNA probe. The strength of the hybridisation signal was commensurate with a high genomic copy number for the cloned sequence. Accordingly, pGc22 plasmid DNA was radioactively labelled and hybridised to a filter carrying various restriction digests of DNAs prepared from *G.m. centralis* and *G.m. morsitans* (Fig. 1).

Alu I and Rsa I cleave the majority of those genomic sequences homologous to the probe into an approximately 160-bp ladder of fragments in each subspecies (Fig. 1, lanes C and E). In contrast, Eco RI and Eco RV cleave only a minority of the *G.m. centralis* sequences into higher multiples of 160 bp (Fig. 1, lanes A and D). The probe also identifies two weak bands corresponding to Eco RI fragments of anomalous lengths: approximately 370 bp and 530 bp, (indicated by dots in Fig. 1).

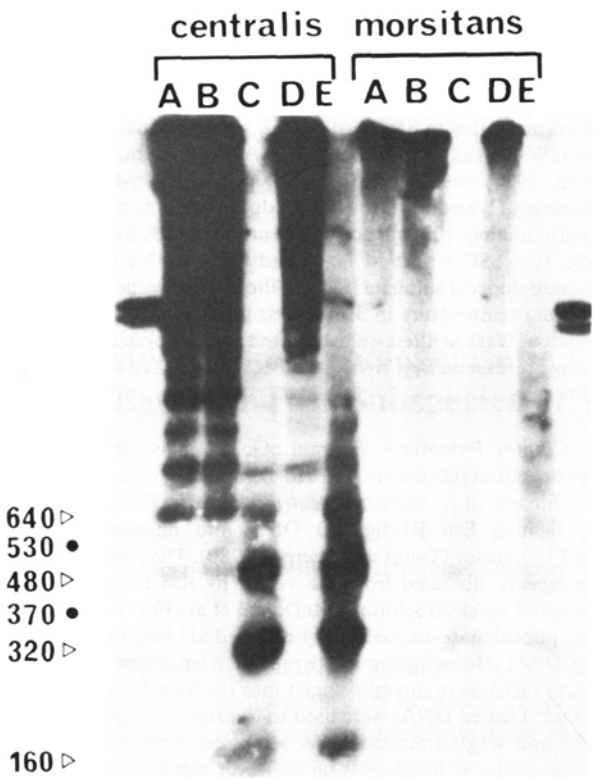

Fig. 1. Restriction digests of total genomic DNAs from *G.m. centralis* and *G.m. morsitans* electrophoresed on 1.2% agarose, blotted onto nitrocellulose and hybridised with pGc22. Restriction enzymes used in lanes were as follows: A, Eco RI; B, Eco RI*; C, Alu I; D, Eco RV; E, Rsa I. Flanking marker tracks are Hinf I digests of pGc22. Figures are in base pairs

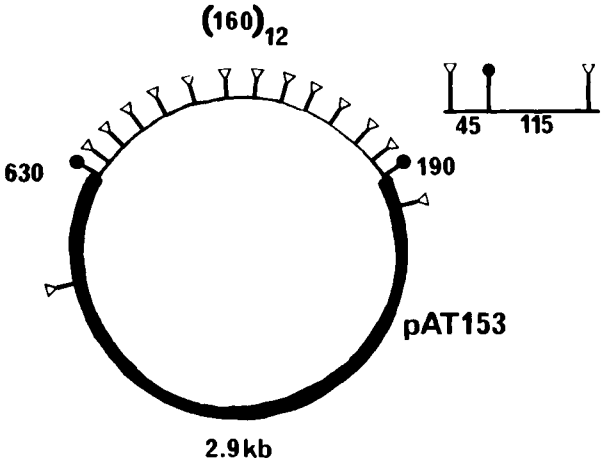

Fig. 2. Restriction map of plasmid pGc22. The thick lines denote vector sequences; the thin line, inserted *G.m. centralis* DNA. Eco RI (●) and Rsa I (▽) restriction sites are shown; the figures refer to the sizes (in base pairs) of the Rsa I fragments obtained. The inset illustrates the inferred structure of the terminal repeats

The conclusion is that pGc22 contains a set of contiguous repeated sequences originating from within a subfamily of a repetitive DNA family in the *G.m. centralis* genome. There appear to be related and similarly organised, yet substantially diverged, sequences in the *G.m. morsitans* genome.

To confirm the proposed internally repetitive structure of pGc22, we digested the plasmid DNA with Eco RI, yielding a 2.0-kb insert. Subsequent digestion with Rsa I gave results consistent with this Eco RI insert comprising approximately 12 Rsa I-sensitive 160-bp repeats (data not shown). The sizes of the hybrid fragments, 190 bp and 630 bp, generated by the combination of the Rsa I sites in the left and rightmost repeats in the array and sites located in the flanking vector sequences, indicate that the distance between Rsa I and Eco RI sites in the terminal repeats is approximately 115 bp (Fig. 2).

Sequencing of the Internal Rsa I Repeats

To assay the variation between members of this set of contiguous repeats, pGc22 plasmid DNA was digested with Rsa I and the 160-bp band, produced by the co-migration of the constituent monomers, was isolated and subcloned into an M13 vector for nucleotide sequencing. Of the four different clones sequenced, there were two in each orientation.

Figure 3 summarises the data obtained. The average length of the clones sequenced is 152 bp. An examination of the consensus sequence reveals an Alu I site at position 63 (Fig. 3) that is represented in each clone, reflecting the majority Alu I cleavage of the genomic family. In addition, each clone is characterised by the split halves of an Rsa I site at its left and right ends.

No clone contains an intact Eco RI site, although mutable cleavage sites, capable of conversion to the full Eco RI sequence by a single nucleotide substitution, are located at positions 112 and 119 in the consensus sequence (GATTTC and GAATTT, respectively). The 115-bp interval between Rsa I and Eco RI sites, derived from restriction mapping of the outermost repeats in the cloned array, strongly suggests that authentic Eco RI cleavage sites have indeed arisen at one of these positions in these particular repeats. The presence of an additional mutable Eco RI site (GAATGC) at position 70 in the sequence of clone GMC3 may be significant, in that Eco RI fragments of anomalous lengths could arise from repeats with Eco RI sites in alternative locations.

With the aid of several computer programs, we searched the consensus sequence for the presence of internal subrepeats or regions of dyad symmetry. No significant instances of either type of feature emerged from these studies.

Sequence Variation Between Repeats

Despite having originated from within a chromosomal segment just 2 kb long, the sequenced repeats

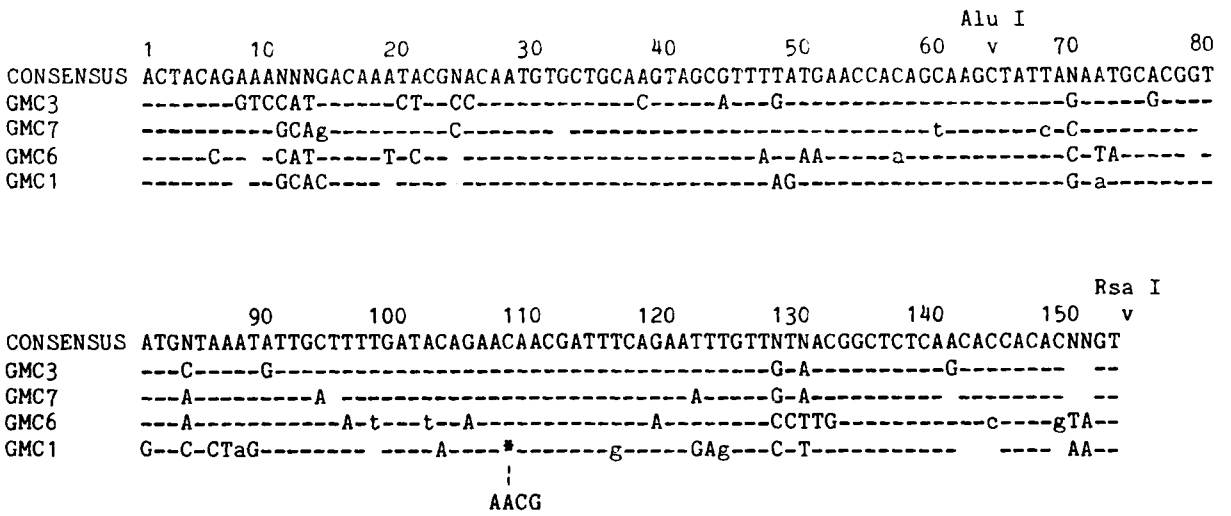

Fig. 3. Nucleotide sequences of the cloned *G.m. centralis* Rsa I repeats. The consensus sequence is derived from the most common nucleotide at each position, ambivalent positions being marked "N." Upper-case letters refer to substitutions; lower-case letters to single insertions after the corresponding consensus nucleotide; and asterisks to longer insertions. Spaces in the lines drawn below the consensus sequence indicate deletions

Table 1. Matrix of percentage nucleotide divergence between all *Glossina morsitans* clones sequenced

	G. m. centralis				*G. m. morsitans*			
	GMC3	GMC7	GMC6	GMC1	GMM1	GMM2	GMM3	GMM6
GMC3	0	16	24	27	35	31	28	29
GMC7		0	23	24	32	35	30	34
GMC6			0	30	35	38	36	37
GMC1				0	31	34	27	34
GMM1					0	35	19	34
GMM2						0	33	37
GMM3							0	29
GMM6								0

display significant nucleotide variation; there is an average between-clone divergence of 24% (Table 1). Clones GMC3 and GMC7 are the most alike, with a sequence divergence of 16%, whilst GMC6 and GMC1 are the most dissimilar, with a divergence of 30%. Nevertheless, there are extensive tracts of perfect homology (up to 25 bp long) between different pairs of clones (Fig. 3); in addition, the mapping data of pGc22 show that the 4-bp Rsa I sites have remained intact in all 12 of the cloned repeats.

Most of the observed variation results from single-nucleotide substitutions, although there are several instances of contiguous nucleotides being replaced. Insertion/deletion differences are, on the whole, less frequent, accounting for 40% of the total variation. Clone GMC1 is characterised at position 108 by an extra sequence that forms a 4-bp direct repeat of the succeeding sequence.

A closer examination of Fig. 3 reveals that certain sequence motifs are shared between individual repeats. Most of these single-nucleotide commonalities are not statistically significant, given the inherent level of variation and the small sample size. However, the CAT substitution at positions 11–13 in clones GMC3 and GMC6 and the 2-bp deletions at positions 150–151 in clones GMC3 and GMC7 are significant; the probability of their occurring independently would be less than 10^{-5}. These two motifs are found in all possible combinations amongst the four repeats sequenced.

Nucleotide Sequences of Homologous Repeats from G.m. morsitans

To compare these results with those for the related DNA family, Rsa I-generated monomer restriction fragments from total genomic *G.m. morsitans* DNA were gel purified and cloned directly into the M13 vector. Four different clones were sequenced, two in each orientation. These sequences are shown in Fig. 4. Since the population of restriction fragments from which these clones were derived accounts for approximately 80% of the genomic family, we ex-

```
                                                                    Alu I
          1        10        20        30        40        50        60        v 70        80
CONSENSUS ACTACAGNAACCGATGATAAATTCGCNNAAATGTGCTGCAANTAGCNTTTTNGCAACCACAGCAAGCTATTAGAATGCTA
GMM1      -----GAATT--*------------  AC-G----------A----G----T--------------------------
                      ¦
                      AC
GMM2      -------A------------G-----  A-------------G--A-A--G-AT-----G---G---T---G-t-*-----
                                                                                  ¦
                                                                                  GGGCGTTG
GMM3      -------G----  --T----------AC--T----------A----A----AA-----------------------
GMM6      -------G---------C--------GCT------------G----G---AT- --------g----------------


                                                                          Rsa I
                  90        100       110       120       130       140       v
CONSENSUS AATATTGCTTTTTATAAAGAAAANCAACGATTTCAGAAAATTTGTTTTTANNGCTCTCAGAACCCNACGT
GMM1      cTG------------G-------G--G----------------------CT-------------Aa---
GMM2      ------g-----G---C------CA--------------------ATGTATATTATGACAGAAGT
GMM3      -G---a-------  ---------G-------G--G---------------CA-----------AC----
GMM6      ------------A----------CT-aT-------------  G----G--TG---G-G--CATAAGTGGGAATTGT
```

Fig. 4. Nucleotide sequences of the cloned *G.m. morsitans* Rsa I repeats. Conventions are as in Fig. 3. The distal portion of the consensus sequence is by necessity derived from clones GMM1 and GMM3 only

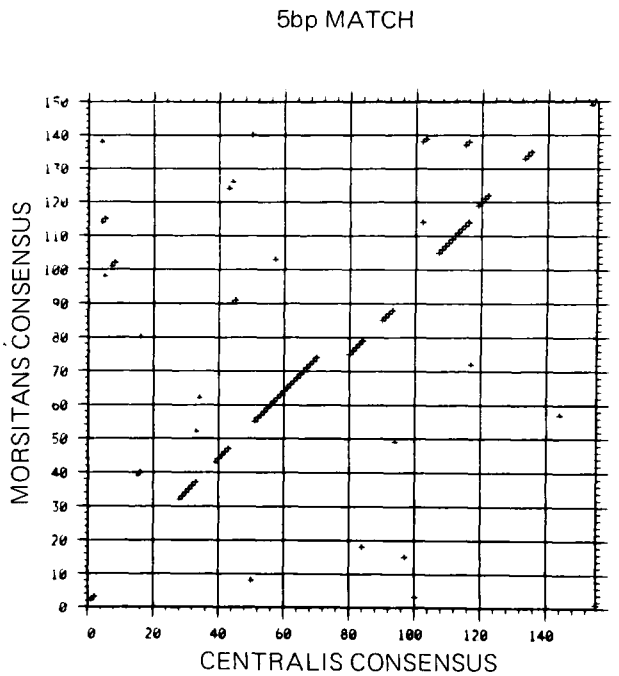

Fig. 5. Computer-generated matrix plot of homology between the *G.m. centralis* and *G.m. morsitans* consensus sequences. A cross is placed at coordinates at which five consecutive nucleotides are identical; thus sequence homology appears as lines coincident with or parallel to the diagonal

pected this sample to be more representative than were the *G.m. centralis* clones. Indeed, the distal sequences of clones GMM2 and GMM6 are irreconcilable (Fig. 4). Clone GMM2 is characterised at position 75 by an 8-bp GC-rich extra sequence. This marked heterogeneity may indicate a sampling from divergent subfamilies. In accord with the physically dispersed origins of these repeats, there are no significant examples of shared and recombined sequence motifs.

The divergences between the sequenced *G.m. morsitans* repeats are generally higher than those between the contiguous *G.m. centralis* repeats; the average divergence is 31%, although clones GMM1 and GMM3 display a level of sequence divergence (19%) comparable to that between the most similar *G.m. centralis* clones, GMC3 and GMC7 (Table 1).

Comparisons Between Subspecies

A survey of all possible pairwise comparisons amongst the eight clones sequenced (Table 1) reveals that the average between-subspecies variation (33%) is not much greater than either of the two respective within-subspecies variations (24% and 31%), although the small sample sizes involved necessitate caution in assuming the significance of these results. A result of this is that the clones tend to display homologies that transcend the subspecies of origin. For example, the sequence divergence between *G.m. morsitans* clone GMM3 and *G.m. centralis* clone GMC3 is 28%, less than the average *G.m. morsitans* within-family variation. This effect can be further illustrated at the sequence level. Clone GMM3 is characterised at position 26 by the sequence CA-CAAT (Fig. 4), which is found in its entirety in clone GMC7 and in partial form in the other *G.m. centralis* clones. In addition, it bears the terminal sequence ACCACACGT, which characterises both GMC3 and GMC7 of the *G.m. centralis* clones.

Despite the general absence of homogenisation, some features of the cloned sequences do show species-diagnostic variants that appear to have be-

come fixed within each family. A comparison of the two aligned consensus sequences (Fig. 5) reveals long tracts of strong homology, with a major discontinuity at coordinates (70, 75). This is a manifestation of a 9-bp insertion/deletion difference that is diagnostic for the two subspecies. Each of the *G.m. centralis* clones examined is characterised at position 75 by a variable block of nucleotides (consensus ACGGTATGC) that has no counterpart in the *G.m. morsitans* sequences. Two other features also cluster the clones according to genome of origin. These are insertion/deletion differences of 3 bp and 2 bp at positions 102 and 119, respectively, in the *G.m. morsitans* consensus sequence.

Discussion

We isolated and sequenced individual members of related repetitive DNA families from two subspecies of tsetse fly. Estimates of the within- and between-subspecies variations were made as a metric for the overall rates of turnover and subsequent homogenisation. Our results show levels of within-family heterogeneity (24% and 31%) that are approximately an order of magnitude higher than those previously reported for DNA families in other genera (Hsieh and Brutlag 1979; Hörz and Altenburger 1981; Thayer et al. 1981; Strachan et al. 1982; T. Strachan, D.A. Webb, and G.A. Dover, submitted for publication). Additionally, we find that the mean between-family divergence is not significantly greater than the observed within-family variation in each subspecies.

Interestingly, there is a degree of overlap between the observed levels of variation detected amongst the contiguous repeats cloned from within a subfamily in the *G.m. centralis* genome and those detected amongst the repeats cloned essentially at random from the whole *G.m. morsitans* family. This is similar to the finding that repeats in mouse satellite DNA subfamilies are no more divergent than other, seemingly more representative repeats (Hörz and Altenburger 1981). In contrast, repeats located at the ends of arrays in African Green Monkey α-satellite DNA and isolated repeats of the dispersed MIF-1 family in the mouse can be more divergent than other repeats, accumulating significant sequence variation (McCutchan et al. 1982; Brown and Piechaczyk 1983).

From the observation of certain diagnostic sequence motifs that are in linked and unlinked configurations in separate repeats from the cloned *G.m. centralis* array, we conclude that there have been genetic exchanges between repeats in this small region of the chromosome, possibly by unequal crossing-over. What is apparent is that since this putative ancestral exchange event there has arisen a mass of nucleotide variation that has not been homogenised to neighbouring repeats. Although this variation appears to be distributed randomly over the lengths of the four clones sequenced, it is striking that the restriction mapping data show that each of the 12 Rsa I sites contained within the cloned DNA has remained intact. On the basis of 24% random and independent divergence occurring amongst the 12 repeats, only 4 such sites would be expected to be conserved. It would appear, therefore, that sequences considerably shorter than the repeat length may be homogenised without concomitant unequal exchange. It is possible that domains of gene conversion shorter than the repeat length could be responsible, similarly to those observed in the human immunoglobulin V_K gene family (Bentley and Rabbitts 1983).

It is clear from the high within-family variation that the overall rate of homogenisation is lower in the tsetse families than in the 360 and 500 families in sibling species of the *melanogaster* subgroup of *Drosophila* (Strachan et al. 1982; T. Strachan, D.A. Webb, and G.A. Dover, submitted for publication), in which variation averages 3%, compared with an average between-species variation of 30%. Several parameters would be expected to contribute to this rate (Dover 1982; Dover et al. 1982, 1984; Ohta and Dover 1984), including the rate of mutation and the family copy number and karyotypic distribution. Karyotypic distribution may be different between the two genera. For example, the *Drosophila* 360 family seems to be confined to the X-chromosome (Peacock et al. 1977; Brutlag 1981) and, although the locations of the tsetse sequences are unknown, there is the possibility of their being karyotypically dispersed to both the regular and the supernumerary complements that characterise many tsetse species (Itard 1973; Southern and Pell 1973; Southern 1980). In situ hybridisation studies have revealed the presence of other repetitive families on both regular centromeres and supernumerary telomeres in *G. morsitans morsitans* (Amos and Dover 1981). However, since the *Drosophila* 500 family is dispersed to all chromosomes and yet still displays a mere 3% variation (Strachan et al. 1982; T. Strachan, D.A. Webb, and G.A. Dover, submitted for publication), these considerations may be inadequate to explain the large discrepancy in homogenisation rates between the two genera. We may therefore have to look for an explanation in terms of the primary rates of the homogenising processes themselves or possibly in terms of the parameters, such as generation time and effective population size, that influence homogenisation in the chromosome pool (Dover 1982; Ohta and Dover 1983, 1984).

It is important to note that variation in turnover

rate and other factors affecting the rate of homogenisation would lead to different expectations as to the extent to which concerted evolution can be observed between a pair of species. Similar within- and between-species levels of divergence have been found for *Alu* family repeats in higher apes (Grimaldi et al. 1981). However, examination across greater phylogenetic distances reveals the distinctive species-specific effects of homogenisation (Jelinek and Schmid 1982; Daniels et al. 1983).

It is of interest that despite the apparently low rates of turnover in tsetse flies, we have identified a few instances of homogenisation between the repeats that we sampled, and remarkably, each is an insertion/deletion difference. It is possible that such a class of mutations may be more readily homogenised by some mechanism such as mismatch-repair even though a low overall homogenisation rate and a short separation time have conspired to impede the accumulation of diagnostic nucleotide substitutions.

Acknowledgments. We would like to express our gratitude to Dr. A.M. Jordan, Dr. Ian Maudlin and the staff of the Tsetse Research Laboratory for their excellent assistance. We also thank our colleagues Enrico Coen, Tom Strachan and Gerald Franz for their help and discussion. M.T. acknowledges the receipt of an SERC CASE studentship. These studies were also supported by SERC grant no. GR/B/68107.

References

Amos CA, Dover GA (1981) The distribution of repetitive DNAs between regular and supernumerary chromosomes in species of *Glossina*: a two-step process in the origin of supernumeraries. Chromosoma 81:673–690

Arnheim N (1983) Concerted evolution of multigene families. In: Koehn R, Nei M (eds) Evolution of genes and proteins. Sinauer Associates, Sunderland, Massachusetts, pp 38–61

Bentley DL, Rabbitts TH (1983) Evolution of immunoglobulin V genes: evidence indicating that recently duplicated human V_K sequences have diverged by gene conversion. Cell 32:181–189

Brown SDM, Piechaczyk M (1983) Insertion sequences and tandem repetitions as sources of variation in a dispersed repeat family. J Mol Biol 165:249–256

Brutlag DL (1981) Molecular arrangements and evolution of heterochromatic DNA. Annu Rev Genet 14:121–144

Coen ES, Dover GA (1983) Unequal exchanges and the coevolution of X and Y rDNA arrays in *D. melanogaster*. Cell 33:849–855

Coen ES, Thoday JM, Dover GA (1982) Rate of turnover of structural variants in the rDNA gene family of *Drosophila melanogaster*. Nature 295:564–568

Curtis CF (1972) Sterility from crosses between sub-species of the tsetse fly *Glossina morsitans*. Acta Trop (Basel) 29:250–268

Daniels GR, Fox GM, Loewensteiner D, Schmid CW, Deininger PL (1983) Species-specific homogeneity of the primate Alu family of repeated DNA sequences. Nucleic Acids Res 11: 7579–7593

Dover, GA (1982) Molecular drive: a cohesive mode of species evolution. Nature 299:111–117

Dover GA, Brown SDM, Coen ES, Dallas J, Strachan T, Trick M (1982) The dynamics of genome evolution and species differentiation. In: Dover GA, Flavell RB (eds) Genome evolution. Academic Press, London, pp 343–374

Dover GA, Trick M, Strachan T, Coen ES, Brown SDM (1984) DNA family turnover and the coevolution of chromosomes. In: Bennett MD, Gropp A and Wolf U (eds) Chromosomes Today, vol 8. Allen & Unwin, London, pp 229–240

Grimaldi G, Queen C, Singer MF (1981) Interspersed repeated sequences in the African Green Monkey genome that are homologous to the human Alu family. Nucleic Acids Res 9:5553–5568

Grunstein M, Hogness DS (1975) Colony hybridisation: a method for the isolation of cloned DNAs that contain a specific gene. Proc Natl Acad Sci USA 72:3961–3965

Hood L, Campbell JH, Elgin SCR (1975) The organisation, expression and evolution of antibody genes and other multigene families. Annu Rev Genet 9:305–353

Hörz W, Altenburger W (1981) Nucleotide sequence of mouse satellite DNA. Nucleic Acids Res 9:683–696

Hsieh T-S, Brutlag D (1979) Sequence and sequence variation within the 1.688 g/cm^3 satellite DNA of *Drosophila melanogaster*. J Mol Biol 135:465–481

Itard J (1973) Revue des connaissances actuelles sur la cytogenetique des Glossines (Diptera). Rev Elev Med Vet Pays Trop 26:151–167

Jelinek WR, Schmid CW (1982) Repetitive sequences in eukaryotic DNA and their expression. Annu Rev Biochem 51: 813–844

Jones CW, Kafatos FC (1982) Accepted mutations in a gene family: evolutionary diversification of duplicated DNA. J Mol Evol 19:87–103

Kedes LH (1979) Histone genes and histone messengers. Annu Rev Biochem 48:837–870

McCutchan T, Hsu H, Thayer RE, Singer MF (1982) Organisation of African Green Monkey DNA at junctions between α-satellite and other DNA sequences. J Mol Biol 157:195–211

McDonell MW, Simon MN, Studier FW (1977) Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J Mol Biol 110:119–146

Messing J (1981) M13mp2 and derivatives: a molecular cloning system for DNA sequencing, strand-specific hybridisation and *in vitro* mutagenesis. In: Walton A (ed) Third Cleveland symposium on macromolecules, recombinant DNA. Elsevier, Amsterdam, pp 143–153

Mews AR, Langley PA, Pimley RW, Flood MET (1977) Large-scale rearing of tsetse flies (*Glossina* spp.) in the absence of a living host. Bull Entomol Res 67:119–128

Miklos GLG, Gill AC (1981) The DNA sequences of cloned complex satellite DNAs from Hawaiian *Drosophila* and their bearing on satellite DNA sequence conservation. Chromosoma 82:409–427

Ohta T (1980) Evolution and variation in multigene families. Springer-Verlag, New York (Lecture notes in biomathematics, vol 37)

Ohta T, Dover GA (1983) Population genetics of multigene families that are dispersed into two or more chromosomes. Proc Natl Acad Sci USA 80:4079–4083

Ohta T, Dover GA (1984) The cohesive population genetics of molecular drive. Genetics (in press)

Peacock WJ, Lohe AR, Gerlach WL, Dunsmuir P, Dennis ES, Appels R (1977) Fine structure and evolution of DNA in heterochromatin. Cold Spring Harbor Symp Quant Biol 43: 1121–1135

Petes TD (1980) Unequal meiotic recombination within tandem arrays of yeast ribosomal RNA genes. Cell 19:765–774

Potts WH (1970) Systematics and identification of *Glossina*. In: Mulligan HW (ed) The African trypanosomiases. Allen and Unwin, London, pp 243–273

Rigby PWJ, Dickmann M, Rhodes C, Berg P (1977) Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J Mol Biol 113:237–251

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467

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

Southern DI (1980) Chromosome diversity in tsetse flies. In: Blackman RL, Hewitt GM, Ashburner M (eds) Insect cytogenetics. Blackwell, London, pp 225–243

Southern DI, Pell PE (1973) Chromosome relationships and meiotic mechanisms of certain *morsitans* group tsetse flies and their hybrids. Chromosoma 44:319–334

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:505–517

Strachan T, Coen ES, Webb DA, Dover GA (1982) Modes and rates of change of complex DNA families of *Drosophila*. J Mol Biol 158:37–54

Szostak JW, Wu R (1980) Unequal crossing over in the ribosomal DNA of *Saccharomyces cerevisiae*. Nature 284:426–430

Tartof K (1974) Unequal mitotic sister chromatid exchange and disproportionate replication as mechanisms regulating ribosomal RNA gene redundancies. Cold Spring Harbor Symp Quant Biol 38:491–500

Thayer RE, Singer MF, McCutchan TF (1981) Sequence relationships between single repeat units of highly reiterated African Green Monkey DNA. Nucleic Acids Res 9:169–181

Trick M (1983) Aspects of genome evolution in tsetse fly species. Ph.D. thesis, University of Cambridge, Cambridge, England

Trick M, Dover GA (1984) Genetic relationships between subspecies of the tsetse fly *Glossina morsitans* inferred from variation in mitochondrial sequences. Can J Gen Cytol (in press)

Twigg AJ, Sherratt D (1980) Trans-complementable copy-number mutants of plasmid ColE1. Nature 283:216–218

Vanderplank FL (1944) Hybridisation between *Glossina* species and suggested new method for control of certain species of tsetse. Nature 154:607–608

Vanderplank FL (1947) Experiments on the hybridisation of tsetse flies (*Glossina*, Diptera) and the possibility of a new method of control. Trans R Entomol Soc Lond 98:1–18

Wahl GM, Stern M, Stark G (1979) Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridisation by using dextran sulphate. Proc Natl Acad Sci USA 76:3683–3687

Yamamoto M, Miklos GLG (1978) Genetic studies on the heterochromatin in *D. melanogaster* and their implication for the function of satellite DNA. Chromosoma 66:71–98