Available online at www.sciencedirect.com
SCIENCE DIRECT®
Gene 332 (2004) 159–168
GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES
www.elsevier.com/locate/gene

Evolutionary dynamics of satellite DNA in species of the Genus *Formica* (Hymenoptera, Formicidae)

Pedro Lorite[a], José Antonio Carrillo[a], Alberto Tinaut[b], Teresa Palomeque[a,*]

[a] *Departamento de Biología Experimental, Área de Genética, Universidad de Jaén, Jaén 23071, Spain*
[b] *Departamento de Biología Animal y Ecología, Facultad de Ciencias, Universidad de Granada, Granada 18071, Spain*

Received 17 July 2003; received in revised form 29 January 2004; accepted 27 February 2004

Received by B. Dujon

Available online 17 April 2004

Abstract

The satellite DNA has been characterized in eight species of the *Formica* genus. This satellite DNA is organized as tandemly repeated 129-bp monomers in all species and it presents internal inverted repeats. The results of all the analyses performed in the sequences sampled from *Formica cunicularia, F. fusca, F. gerardi, F. rufibarbis, F. selysi, F. frontalis*, and *F. sanguinea* suggest interspecific conservation of satellite DNA. Nevertheless, the results from the comparative analysis of the sequences sampled from *F. subrufa* and the remaining species studied suggest that the mechanisms producing concerted evolution have been efficient in these taxa. A CENP-B-like motif has been found in the satellite DNA from the species analysed, including *F. subrufa.* This satellite DNA is located in the pericentromeric regions of all chromosomes. We suggest that, although the evolution of the DNA satellite in ants could be similar to that in other organisms, there may be some particularities as a result of a haplodiploid system.

Keywords: *Formica*; Formicidae; Satellite DNA; Concerted evolution and CENP-B-like motif

1. Introduction

Comparisons of paralogous repetitive DNA sequences in the genomes of eukaryotes often reveal striking intraspecific similarities but marked interspecific divergence. This apparent evolutionary lack of independence among the members of a sequence family is known as concerted evolution and has been explained by molecular drive (Dover, 1986).

Although satellite sequences have been characterized in a number of species, no sequence-specific role has been definitively established for them. Satellite DNA transcription has been reported in several species of Hymenoptera (Rouleux-Bonnin et al., 1996; Renault et al., 1999; Lorite et al., 2002d) but the functional significance of satellite DNA transcripts currently remains unknown. Some components of centromeric DNA, such as the human alpha satellite and the mouse minor satellite, are thought to be functional components of centromeres, since they contain the CENP-B box, the binding site for centromere protein B. However, gene knockout analysis of CENP-B in mice indicates that this protein is not essential for centromeric activity (for review, Craig et al., 1999).

In the suborder Symphyta (Hymenoptera), the satellite DNA (stDNA) has been described in several species (Rouleux-Bonnin et al., 1996). However, in the suborder Apocrita (Hymenoptera), stDNA sequences have been studied only in some species of parasitoid wasps (Bigot et al., 1990; Renault et al., 1999; Landais et al., 2000) and in the honeybee (Tarès et al., 1993). In ants (Apocrita, Formicidae), these sequences have been investigated only in several species of the genus *Messor*, in *Aphaenogaster subterranea* and in *Monomorium subopacum* (Lorite et al., 2002b,d, in press).

Mantovani et al. (1997) have observed different rates of molecular evolution of stDNA in bisexual and automictic parthenogenetic insect species in the genus *Bacillus* (Insecta, Phasmatodea). In agreement with the authors, the different rate of evolutionary turnover of satellite DNA in

Abbreviations: bp, base pairs; DIG, digoxigenin-11-dUTP; kb, kilobase(s); *F.*, *Formica*; stDNA, satellite DNA.

* Corresponding author. Tel.: +34-953-01-21-56; fax: +34-953-01-21-41.
E-mail address: tpalome@ujaen.es (T. Palomeque).

0378-1119/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2004.02.049

the two types of species has provided experimental support for the idea that meiosis and the chromosome segregation are key in stDNA evolution. The evolutionary dynamics of tandemly repetitive DNAs in sexual organisms lead to a gradual and cohesive spreading of a new repeat unit produced by mutation throughout a particular stDNA family (known as homogenization) and throughout the entire population (fixation). Nevertheless, evidence of biased gene conversion events suggests that, given enough time, sequence homogenization can take place in unisexual species (Luchetti et al., 2003)

Ant male haploids do not undergo meiotic recombination and redistribution of chromosomes to the next generation. Bigot et al. (1990) suggested that in Hymenoptera, the processes of molecular evolution of the satellite DNA would be altered by the special constraints imposed by the haplodiploid system.

In the present paper, we analyse the satellite DNA from eight species of the *Formica* genus. The aim is to contribute to a fuller understanding of the molecular mechanisms that act in the evolution of stDNA and, in particular, to determine the possible influence that the occurrence of meiosis only in one sex, as it happens in haplodiploid insects, could have on these mechanisms.

2. Materials and methods

2.1. Sampling, DNA extraction, cloning, sequencing and Southern analysis

Ant nests were collected for the following species: *Formica cunicularia, F. gerardi,* and *F. subrufa* (Jaén, Spain); *F. frontalis* (Granada, Spain); *F. fusca* and *F. rufibarbis* (Tours, France); *F. selysi* (Morillon, France); and *F. sanguinea* (Tuscany, Italy). These species are belong to the subgenera: *Serviformica* (*F. cunicularia, F. fusca, F. gerardi, F. rufibarbis,* and *F. selysi*), *Formica* (*F. frontalis*), *Raptiformica* (*F. sanguinea*), and *Iberoformica* (*F. subrufa*). Workers were used for genomic DNA extraction using standard procedures.

Genomic DNA was restricted according to the recommendations of the supplier, and fragments were separated by electrophoresis on agarose gels. Fragments smaller than 500 bp produced by digestion of genomic DNA from *F. cunicularia* with *Sau*3A or *Mbo*I were eluted from the agarose gel and inserted into the *Bam*HI site of pUC19 vector. The fragments produced by digesting genomic DNA with *Alu*I were inserted into the *Sma*I site of pUC19 vector. A similar methodology was used with the fragments obtained by *Sau*3A digestion of genomic DNA from *F. fusca F. gerardi*, *F. rufibarbis*, *F. selysi, F. frontalis*, *F. sanguinea* and *F. subrufa*. The fragments produced by digesting genomic DNA from *F. subrufa* with *Alu*I were inserted into the *Sma*I site of pUC19 vector. A portion of the eluted fragment from each species was digoxigenin-labelled by random priming with the DIG system (Roche) and used as hybridization probes. Recombinants yielding positive hybridization signals were directly sequenced on both strands by the dideoxy sequencing method. For Southern analysis, genomic DNA from each species was digested with different restriction enzymes. Agarose gel electrophoresis of digested DNA sampled and processing for transfer onto a nitro-cellulose membranes were performed using standard procedures. Southern hybridization was performed using 20 ng of labelled probe/ml and a final wash in 2 × SSC at 60 °C. Detection of hybridization was performed with a DIG-detection kit (Roche).

2.2. Sequence analysis

Multiple-sequence alignment was performed using the CLUSTALW program. Nucleotide diversity and pairwise sequence divergences were estimated using DnaSP program (Rozas and Rozas, 1999). The nucleotide diversity Pi (JC), in each species was calculated as the average number of nucleotide substitutions per site between two sequences with Jukes and Cantor's (1969) correction. Sequence divergences were calculated as the average of nucleotide substitution per site between species (D_{xy} value from DnaSP, Nei, 1987, Eqs. 10–20). Distance analysis was conducted with the program MEGA (Kumar et al., 1993). Trees were constructed by the neighbour-joining method (Saitou and Nei, 1987) with bootstrap values on 500 replications.

The pattern of variation at each nucleotide position of satellite DNA shared between two species was analysed by the method of Strachan et al. (1985). This method compares the sequences from pairs of species at each nucleotide position independently. The positions are divided into six stages or ''classes'' (see Fig. 4 in Strachan et al., 1985). Class 1 represents an absence of mutations (or their spread) in the presumed progenitor bases, which are still shared by the two species. Class 2 represents rare mutations (or low levels of subsequent spread) resulting in the appearance of a minority of clones with a new mutation in a given position in one species whilst the other species remains homogenous for the progenitor base at the corresponding position. Class 3 covers cases in which no distinction can be made between minority and majority frequencies, in that a mutation and the progenitor base have approximately equal frequencies, while the other species is homogeneous for one of the two bases. Class 4 includes positions in which a mutation, which is apparently absent in one of the species, has replaced the progenitor base in the majority of members in the other species. Class 5 represents positions in which the two species are internally homogeneous for bases that are diagnostically different for each of the species. In classes 2–5, only two types of bases are found at a given position across all clones of a pair of species. Class 6 represents the first stage of subsequent replacement of the mutation fixed in stage 5.

2.3. Chromosome preparation, and in situ hybridization

Chromosome spreads were obtained from worker prepupae cerebral ganglion cells. In situ hybridization was performed by using FOCUA-15 insert as probe. Chromosome spreads, hybridization and signal detection were performed as previously described by Lorite et al. (2002a).

3. Results

3.1. Detection, cloning, and sequencing of satellite DNA

Digestion of genomic DNA from *F. cunicularia* with restriction enzymes *AluI* or *Sau*3A showed the presence of characteristic ladders of stDNA (Fig. 1a). A similar ladder with bands differing by about 130 bp was also found when *Mbo*I (isoschizomere of *Sau*3A) was used (data not shown). Thirteen clones from the *Sau*3A digest (FOCU-1, 5, 10, 15, 16, 27A, 27B, 30, 33, 34, 35, 46A, and 46B), three from the *Mbo*I digest (FOCUM-3, 5, and 22) and three from the *Alu*I digest (FOCUA-5, 7, and 15) were sequenced (EMBL accession numbers from AJ238724 to AJ238728, from AJ238730 to AJ238732, from AJ308973 to AJ308981, AJ308985, and AJ308988).

Digestion of genomic DNA from *F. fusca F. gerardi*, *F. rufibarbis*, *F. selysi*, *F. frontalis*, and *F. sanguinea* with *Sau3A* also showed the presence of characteristic ladders of stDNA having bands with differences of about 130 bp. Genomic DNA from *F. subrufa* digested with *Sau*3A did not reveal visible bands in stained agarose gel. Only a weak band of less of 100 bp was visible. However, genomic DNA from *F. subrufa* digested with *Alu*I revealed a typical ladder in stained agarose gel (data not shown). Six positive clones from *F. fusca* (FOFU-2, 7, 11, 15, 16, and 22), eight from *F. gerardi* (FOGE-5, 6, 30, 34, 41, 114, 116, and 136), nine from *F. rufibarbis* (FORU-1, 2, 3, 7, 8, 10, 12, 38, and 39) five from *F. selysi* (FOSE-2, 3, 4, 14, and 15), 10 from *F. frontalis* (FOFR-1, 4, 6, 7, 10, 11, 18, 22, 26, and 27), and 11 from *F. sanguinea* (FOSA-1, 2, 12, 26, 30, 31, 33, 34, 37, 40, and 41) were sequenced (EMBL accession numbers from AJ508813 to AJ508861). Three positive clones obtained by *Sau*3A and 14 by *Alu*I digestion of genomic DNA from *F. subrufa* (FOSU-1, 5 and 8, and FOSUA-16, 18, 41, 44, 120, 124, 129, 135, 160, 182, 189, 195, 210, and 214, respectively) were sequenced (EMBL accession numbers from AJ508862 to AJ508878).

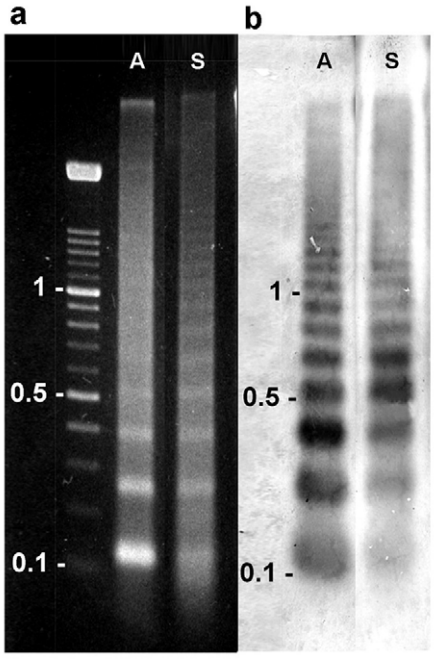


Fig. 1. (a) Electrophoretic separation of restricted *F. cunicularia* genomic DNA, (b) Southern blot of restricted *F. cunicularia* genomic DNA using FOCUA-15 fragment as probe. For each restriction enzyme (A = *Alu*I, S = *Sau*3A), 4 μg of total genomic DNA was digested, fractionated in 2% agarose gel, transferred into nitrocellulose membrane and hybridized with the digoxigenin-labelled fragment (FOCUA-15). The numbers on the left indicate the size of DNA fragments in kb.

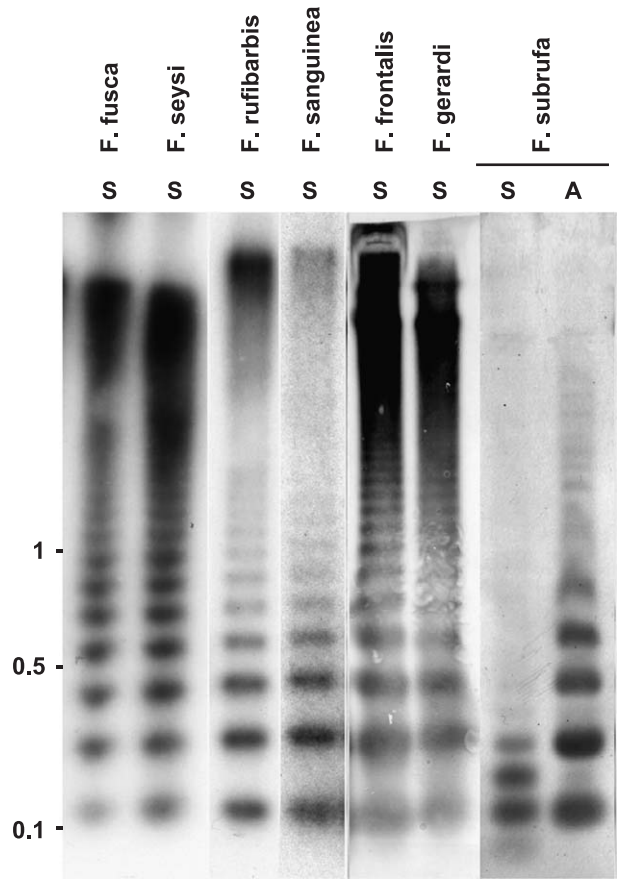
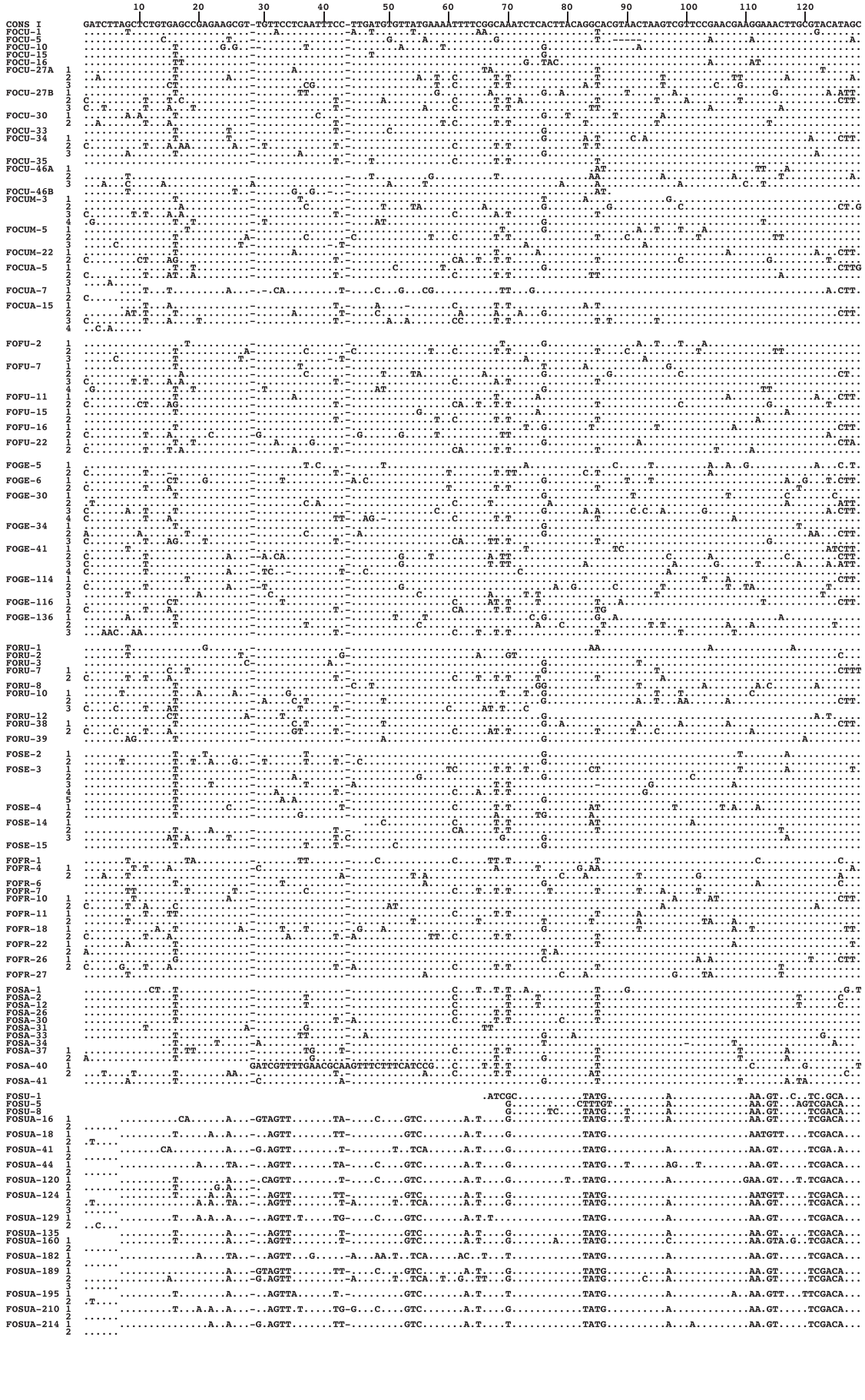


Fig. 2. Southern blot of restricted *F. fusca, F. selysi, F. rufibarbis, F. sanguinea, F. frontalis F. gerardi*, and *F. subrufa* genomic DNA using FOCUA-15 fragments as probe. For each restriction enzyme (A = *Alu*I, S = *Sau*3A), 4 μg of total genomic DNA was digested, fractionated in 2% agarose gel, transferred onto nitrocellulose membrane and hybridized with the digoxigenin-labelled fragment (FOCUA-15).The numbers on the left indicate the size of DNA fragments in kb.


10 20 30 40 50 60 70 80 90 100 110 120
CONS I GATCTTAGCTCTGTGAGCCGAGAAGCGT-TGTTCCTCAATTTCC-TTGATGTGTTATGAAAATTTTCGGCAAATCTCACTTACAGGCACGTAACTAAGTCGTTCCGAACGAAGGAAACTTGCGTACATAGC
FOCU-1
FOCU-5
FOCU-10
FOCU-15
FOCU-16
FOCU-27A 1
2
3
FOCU-27B 1
2
3
FOCU-30 1
2
FOCU-33
FOCU-34 1
2
3
FOCU-35
FOCU-46A 1
2
3
FOCU-46B
FOCUM-3 1
2
3
4
FOCUM-5 1
2
3
FOCUM-22 1
2
FOCUA-5 1
2
3
FOCUA-7 1
2
FOCUA-15 1
2
3
4
FOFU-2 1
2
3
FOFU-7 1
2
3
4
FOFU-11 1
2
FOFU-15 1
2
FOFU-16 1
2
FOFU-22 1
2
FOGE-5 1
2
FOGE-6 1
2
FOGE-30 1
2
3
4
FOGE-34 1
2
3
FOGE-41 1
2
3
4
FOGE-114 1
2
3
FOGE-116 1
2
FOGE-136 1
2
3
FORU-1
FORU-2
FORU-3
FORU-7 1
2
FORU-8
FORU-10 1
2
3
FORU-12
FORU-38 1
2
FORU-39
FOSE-2 1
2
FOSE-3 1
2
3
4
5
FOSE-4 1
2
FOSE-14 1
2
3
FOSE-15
FOFR-1
FOFR-4 1
2
FOFR-6
FOFR-7
FOFR-10 1
2
FOFR-11 1
2
FOFR-18 1
2
FOFR-22 1
2
FOFR-26 1
2
FOFR-27
FOSA-1
FOSA-2
FOSA-12
FOSA-26
FOSA-30
FOSA-31
FOSA-33
FOSA-34
FOSA-37 1
2
FOSA-40 1
2
FOSA-41
FOSU-1
FOSU-5
FOSU-8
FOSUA-16 1
2
FOSUA-18 1
2
FOSUA-41 1
2
FOSUA-44 1
2
FOSUA-120 1
2
FOSUA-124 1
2
3
FOSUA-129 1
2
FOSUA-135
FOSUA-160 1
2
FOSUA-182 1
2
FOSUA-189 1
2
3
FOSUA-195 1
2
FOSUA-210 1
2
FOSUA-214 1
2

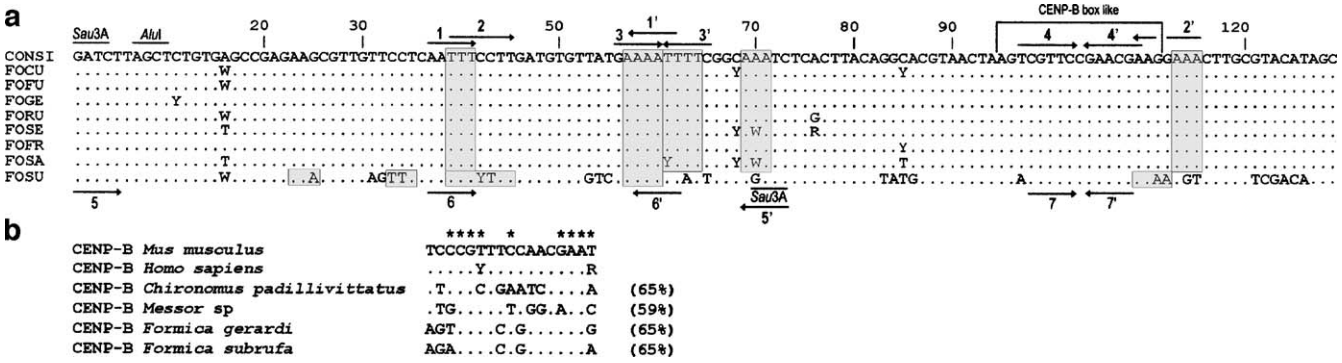


Fig. 4. (a) Consensus sequence from each species and consensus sequence derived from them (CONS I). The recognition sites discussed in the text are indicated. CENP-B-like motifs is also indicated. The A or T tracts ≥3 correspond to the shaded boxes. The inverse repetitions (1:1′, 2:2′, 3:3′, 4:4′) are indicated by arrows above the consensus sequence (CONS I). The inverse repetitions (5:5′, 6:6′, 7:7′) from *F. subrufa* are indicated by arrows under their consensus sequence. (b) Comparison of published CENP-B boxes (Wong and Rattner, 1988; Masumoto et al., 1989; López and Edström, 1998; Lorite et al., 2002b) with potential CENP-B box of the *F. gerardi* and *F. subrufa*. The degree of similarity with mammalian CENP-B box is represented. Asterisks mark the positions that are believed to be essential for binding the CENP-B protein in vitro (Masumoto et al., 1989).

Genomic DNA was examined by Southern-blot hybridization. FOCUA-15 fragments as probe hybridized to multimers of a basic 129-bp unit obtained by *Alu*I or *Sau*3A digestion of genomic DNA from *F. cunicularia* (Fig. 1b), indicating that the cloned fragments belonged to tandemly repetitive DNA sequences. A similar ladder pattern was also detected using FOCUA-15 fragments as probe and genomic DNA from *F. fusca*, *F. selysi*, *F. rufibarbis*, *F. sanguinea*, *F. frontalis* and *F. gerardi*, digested with *Sau*3A, and genomic DNA from *F. subrufa* digested with *Alu*I (Fig. 2). The Southern analysis of genomic DNA of this last species digested with *Sau*3A revealed the presence of hybridization bands outside the repeat ladder of the 129-bp monomers. These bands were approximately 190 and 70 bp long. Also, the hybridizations were visible only from the corresponding dimers downwards (Fig. 2). The sequence analysis indicated the presence of two internal recognition sites for *Sau*3A, one at the beginning of the monomer sequence, as in the other species and another in nucleotide position 70, which could have given rise to the observed bands. In addition, these recognition sites are found in almost all of the sequences sampled from *F. subrufa* (Fig. 3).

The program CLUSTALW was used to establish the consensus sequence of the fragments isolated from these eight species. The basic monomeric repeat was 129 bp long in all species. The comparison of the sequencing results is summarized in Fig. 3, which shows the alignment of repeating units and the consensus sequence derived all of them (consensus I). The 5-prime end of the clone FOSA-40 (with two monomers, 40-1 and 40-2) possesses a small (31 bp) sequence that shows similarity with the inverted 3-prime end of the monomeric unit. Fig. 3 shows some incomplete monomeric units of the species analysed. The shorter length could be a cloning artefact caused by the repetitive nature of the DNA under study.

3.2. Organization and sequence comparisons

The content in A + T in DNA satellite sampled in each species varies from 54% in *F. selysi* to a 58% in *F. cunicularia*. The consensus sequence of each of the species contains several A or T rich runs, as well as several internal inverted subrepeats (Fig. 4a).

Analysis of the consensus sequences from *F. cunicularia, F fusca F. gerardi*, *F. rufibarbis F. selysi* (subgenus *Serviformica*) *F. frontalis* (subgenus *Formica*), and *F. sanguinea* (subgenus *Raptiformica*) revealed only three positions where some of these sequences vary (positions 16, 76 and 85). On the contrary, the consensus sequence of *F. subrufa* (subgenus *Iberoformica*) shows 27 nucleotide positions that differ in relation to the sequence consensus from most of the species. Between the sequence consensus of *F. rufibarbis* with respect to that of *F. subrufa*, there were 28 different nucleotide positions (Fig. 4a).

Transition stages in this satellite DNA were studied by the method of Strachan et al. (1985). From Table 1, we draw the following conclusions: (i) a high percentage of nucleotide positions cannot be included in any of the six categories proposed by Strachan et al. (1985) because there are more than two pairs of bases at a given position. Although this percentage varies widely (20.2–74.4%), in general, it exceeds those in other organisms (Mravinac et al., 2002; Pons et al., 2002). Some of the nucleotide positions (N_1) have exclusively the same type of mutation in the clones from the two species. These shared mutations could be considered ancestral ones. Nevertheless, most of

Fig. 3. Multiple sequence alignments of all sequenced clones from *F. cunicularia* (FOCU, FOCUM and FOCUA), *F. fusca* (FOFU), *F. gerardi* (FOGE), *F. rufibarbis* (FORU), *F. selysi* (FOSE), *F. frontalis* (FOFR) *F. sanguinea* (FOSA), and *F. subrufa* (FOSU and FOSUA). The sequences designated 1, 2… are monomeric units from the same clone. Monomer consensus sequence derived from them is also shown (CONS I).

Table 1
Number of fixed differences between species and percentage of nucleotide positions falling into one of the six classes of transition stages proposed by Strachan et al. (1985)

		Fixed differences	Class 1 (%)	Class 2 (%)	Class 3 (%)	Class 4 (%)	Class 5 (%)	Class 6 (%)	N_1[a] (%)	N_2[b] (%)
FOCU	FOFU	–	18.6	22.5	–	–	–	–	24.8	34.1
	FOGE	–	11.6	14.0	–	–	–	–	17.8	56.6
	FORU	–	14.7	27.1	–	–	–	–	13.2	45.0
	FOSE	–	14.0	38.8	–	–	–	–	6.2	41.1
	FOFR	–	16.3	26.4	–	–	–	–	9.3	48.1
	FOSA	–	14.0	35.7	–	–	–	–	7.0	43.4
FOFU	FOGE	–	18.6	31.0	–	–	–	–	15.5	34.9
	FORU	–	31.8	34.1	–	–	–	–	13.2	20.9
	FOSE	–	37.2	33.3	–	–	–	–	11.6	17.8
	FOFR	–	31.0	31.8	–	–	–	–	16.3	20.9
	FOSA	–	31.8	39.5	–	–	–	–	11.6	17.1
FOGE	FORU	–	21.7	29.5	–	–	–	–	14.0	34.9
	FOSE	–	24.8	29.5	–	–	–	–	7.8	38.0
	FOFR	–	20.9	28.7	–	–	–	–	14.0	36.4
	FOSA	–	22.5	37.2	–	–	–	–	7.0	33.3
FORU	FOSE	–	40.3	31.0	0.8	–	–	–	6.2	21.7
	FOFR	–	35.7	29.5	–	–	–	–	9.3	25.6
	FOSA	–	38.6	35.6	–	–	–	–	8.5	20.2
FOSE	FOFR	–	40.3	31.0	–	–	–	–	6.2	22.5
	FOSA	–	43.4	36.4	–	–	–	–	7.0	13.2
FOFR	FOSA	–	40.3	28.7	–	–	–	–	10.1	20.9
FOSU	FOCU	10	10.0	26.4	0.8	2.3	2.3	3.9	7.0	47.3
	FOFU	13	23.3	30.2	–	6.8	3.9	3.9	6.2	25.6
	FOGE	11	12.4	31.0	2.3	3.1	2.3	4.7	9.3	34.9
	FORU	15	20.1	36.4	–	5.4	6.2	4.7	3.9	23.3
	FOSE	16	30.2	28.7	1.6	3.9	8.5	3.1	3.1	20.9
	FOFR	12	24.0	29.5	–	6.2	4.7	4.7	7.8	23.3
	FOSA	15	27.1	31.0	0.8	3.1	10.0	1.6	8.5	17.8

N_1 and N_2 represent the positions that cannot be assigned unequivocally to a class.
[a] Nucleotide positions that have exclusively the same type of mutation in the clones from the two species.
[b] Nucleotide positions with different mutations in the two species.

the positions not included in this analysis present mutations of different types in the clones of the two species (N_2). These new mutations probably occurred after the species split. In addition, we observed with a certain frequency the existence of the four bases in a same position, for example, in nucleotide position 16 from *F. frontalis*, in the 43 from *F. subrufa*, in the 68 from *F. cunicularia*, and in the 85 from *F. gerardi*. In this table, the second category (N_2) clearly predominates over the first (N_1) in all the comparisons. This second category also includes positions that correspond to differences fixed between the two species, as discussed below. (ii) The result of the analysis in all the species considered in this paper with exception of *F. subrufa* demonstrates that the nucleotide positions considered in this analysis correspond to classes 1 and 2. Only there is a nucleotide position including in class 3, concretely the nucleotide position 16 in the comparison between *F. rufibarbis* and *F. selysi.* Such a distribution of mutation indicated interspecific conservation of stDNA and the absence of any species-specific mutation fixed in all clones of a single species, which would fall into class 5. (iii) The results from the analysis between *F. subrufa* and the other species are different. In all the comparisons, there were nucleotide positions in-

Table 2
Nucleotide diversity in each species (diagonal) and pairwise divergence values between species

	FOCU	FOFU	FOGE	FORU	FOSE	FOFR	FOSA	FOSU
FOCU	0.15 ± 0.01							
FOFU	0.14 ± 0.01	0.13 ± 0.01						
FOGE	0.16 ± 0.01	0.15 ± 0.01	0.16 ± 0.01					
FORU	0.15 ± 0.01	0.13 ± 0.02	0.15 ± 0.02	0.14 ± 0.01				
FOSE	0.13 ± 0.01	0.11 ± 0.01	0.14 ± 0.01	0.12 ± 0.02	0.10 ± 0.01			
FOFR	0.16 ± 0.01	0.13 ± 0.01	0.15 ± 0.01	0.13 ± 0.02	0.12 ± 0.01	0.13 ± 0.01		
FOSA	0.13 ± 0.01	0.13 ± 0.02	0.14 ± 0.02	0.13 ± 0.02	0.10 ± 0.01	0.12 ± 0.01	0.10 ± 0.01	
FOSU	0.40 ± 0.03	0.40 ± 0.04	0.39 ± 0.03	0.39 ± 0.04	0.36 ± 0.04	0.38 ± 0.04	0.40 ± 0.04	0.10 ± 0.01

cluded in classes 5 and 6. Table 1 also shows the number of fixed differences between species (nucleotide sites at which all of the sequences in one population differ from all of the sequences in the second population) determined using the DnaSP program. In agreement with Strachan et al. (1985), the nucleotide position corresponding to differences fixed between species would be included in class 5 or 6. We have observed that in some comparisons, the number of fixed differences is greater than the sum of the nucleotide positions included in classes 5 and 6. For example, in *F. subrufa–F. cunicularia* pair, the number of fixed differences is 10 while the indicated sum is 8. Between these two species, there are two other fixed nucleotide differences, namely in 32 and 122 nucleotide position. In position 32 in the different monomers from *F. cunicularia*, there is mainly T, although some have C or A in this position. On the contrary, all the monomers of *F. subrufa* present G in this position. The comparison corresponding to position 32 is included in the N_2 category defined above. The same is true for position 122.

Nucleotide diversity was estimated for the sequences corresponding to each species. Also, DNA divergence between species was studied (Table 2). As can be seen in this table, the divergences (between sequences of two different species) are similar or even smaller that their respective nucleotide diversities (within sequences of the same species), independently of which they belong to the same or different subgenus, except for the comparison made with *F. subrufa.* In this latter species, the nucleotide diversity was smaller than the sequence divergence between repeats units from different species (Table 2).

The phylogenetic tree constructed by the neighbour-joining method showed a topology in accordance with the aforementioned data. Two main branches are visible and statistically supported by significant bootstrap values (Fig. 5). One branch leads to sequences from all species, with the exception of the sequences from *F. subrufa*. In this branch, the sequences appear intermixed without reflecting taxonomic affinity. Another branch exclusively leads to sequences from *F. subrufa*.

The sequences described here are not significantly similar to previously described sequences deposited in databases. However, a visual inspection of centromere-related sequences from the literature uncovered some noteworthy similarities. The strongest similarity was found in the comparison with mammalian CENP-B box. Fig. 4b shows CENP-B boxes from *Mus musculus* (Wong and Rattner, 1988), from human (Masumoto et al., 1989), and CENP-B-like boxes from the insect *Chironomus padillivittatus* (López and Edström, 1998), from the ant genus *Messor*

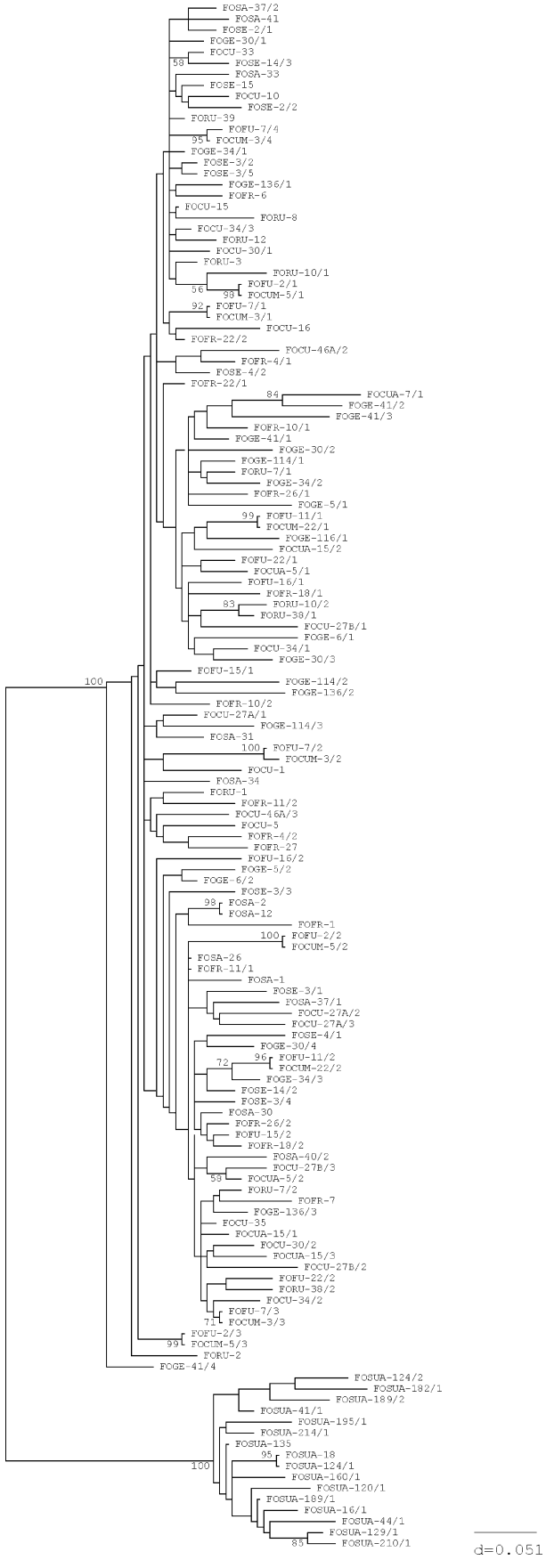


Fig. 5. Dendrogram showing neighbour-joining tree of monomers from the eight species of the *Formica* genus. The bar represents genetic distance $d = 0.051$ of pairwise comparisons calculated according to Kimura's "Two-Parameter Method" (Kimura, 1980). The numbers above and under branches indicate bootstrap values for 500 replicates.

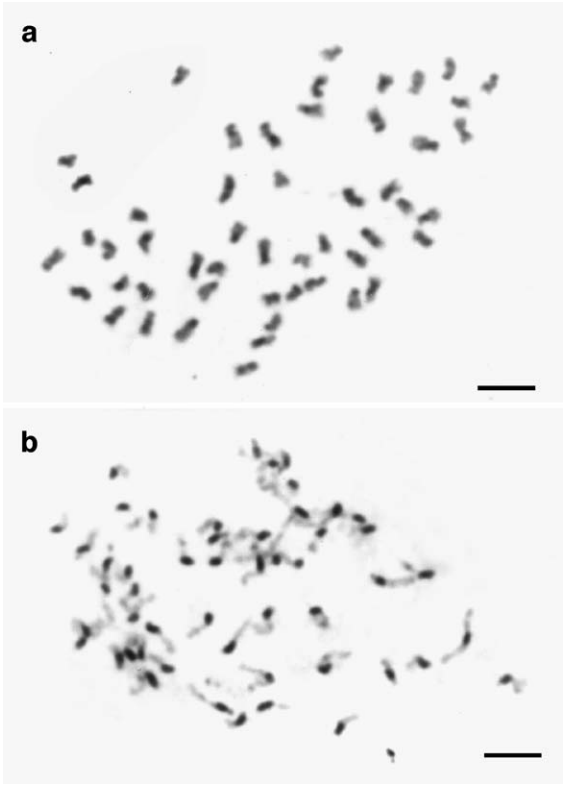


Fig. 6. (a) Metaphase chromosomes from worker prepupae cerebral ganglion cell of *F. frontalis*, $2n = 52$. (b) In situ hybridization using FOCUA-15 insert as probe. Bar = 5 μm.

(Lorite et al., 2002b), and CENP-B-like boxes from *F. gerardi* and *F. subrufa*. In this figure the inverse and complementary sequence CENP-B boxes from *M. musculus*, human and *C. padillivittatus* are represented. We note that in this region (95–111), all consensus sequences from the sampled *Formica* genus have the same nucleotide composition, except the corresponding to *F. subrufa* (Fig. 4a). Nevertheless, these nucleotide differences do not affect the degree of similarity with the mammalian CENP-B box (Fig. 4b).

3.3. Chromosome localization of satellite DNA

Although chromosomal analyses have been made only in some species of the *Formica* genus, the karyotypes from *Formica* sp. are considered very similar and with low variation in relation to the chromosome number (Lorite et al., 2002c). In situ hybridization with a FOCUA-15 insert as the probe showed a pericentromeric location of these sequences in all the mitotic chromosomes from *F. frontalis* (Fig. 6).

4. Discussion

The analysis of the satellite DNA from eight species of the *Formica* genus revealed that this satellite DNA is organised in tandem repeats of 129-bp monomers. The consensus sequence in each species had several A- or T-rich runs as well as some short inverted subrepeats. The presence of internal repeats is common in the stDNA from other Hymenoptera (Renault et al., 1999; Lorite et al., 2002b,d, in press).

All the results in the present study of the stDNA sampled from *F. gerardi*, *F. frontalis*, *F. rufibarbis, F. fusca, F. selysi, F. sanguinea,* and *F. cunicularia* suggest interspecific conservation of these types of sequences and that all the sampled sequenced monomers belong to the same stDNA family. The consensus sequences from these species revealed only three positions where there is variation in some of these sequences. Nevertheless, the observed nucleotide substitutions are not species-specific, since the same nucleotide substitution is present in some monomers of other species, pointing to their probable presence in ancestral repeat variants, as suggested by the analysis performed with the method of Strachan et al. (1985). The results of this analysis show some peculiar characteristics, as discussed below.

In the analysis of the phylogenetic tree of the sampled monomers all of these seven species, no species specific clustering was found. The sequences appear intermixed without reflecting taxonomic affinity. This arrangement reflects that many repeat variants are shared by different species and that genetic distances between monomers of different species are similar to or even shorter than genetic distances between monomers within species. Also, in all pairwise comparisons, the divergences between sequences of two different species proved similar or even smaller than their respective nucleotide diversities within sequences of the same species. Similar sequence conservation has also been described in the satellite DNA family from ant genus *Messor* (Lorite et al., 2002b). There are several examples of conservation of stDNA sequences over long evolutionary periods (for review, see Ugarković and Plohl, 2002).

The results from the comparison between the stDNA from *F. subrufa* and that corresponding to the other species analysed are different. First, the study of the transition states reveals nucleotide positions corresponding to classes 5 and 6 in all pair comparisons. Class 5 represents the classical observation of concerted evolution (Dover, 1986). Class 6 represents the first stage of subsequent replacement of the mutation fixed at stage 5. Secondly, there are fixed nucleotide differences in all the comparisons between *F. subrufa* and the remaining species and the sequence divergences are higher than the nucleotide diversity observed from stDNA of *F. subrufa*. Finally, in agreement with all the data presented, the phylogenetic analysis reveals the existence of two separate clades, one composed of the monomer repeats from *F. subrufa*, and other composed by the rest of the species studied in this paper.

It is not clear why some satellite DNA sequences remained conserved for such long evolutionary periods, while others underwent dynamic nucleotide changes. This contrasting behaviour has been reported even between

closely related species. The evolutionary dynamics of the PIM357 satDNA sequence from the Coleoptera genus *Pimelia* differs in species endemic to the Canary Islands with respect to that of the Iberian *Pimelia* species (Pons et al., 1997, 2002). In the stDNA from Iberian *Pimelia* species, the divergences (between sequences of two different species) are higher than their respective nucleotide diversities (within sequences of the same species). This does not apply to the stDNA of Canary Islands species in which the nucleotide diversities are also higher than the corresponding ones to the Iberian species. These facts point to weaker homogenization processes in the stDNA from Canary Island *Pimelia*. All the data in the present work suggest that the mechanism behind concerted evolution has worked effectively in *F. subrufa.* These data also suggest that the repetitive sequences from *F. subrufa* could be considered to be a subfamily of the stDNA family from the other *Formica* species studied.

The genus *Formica* is a member of the subfamily Formicinae. Unfortunately, the evolutionary relationships between the different species of the genus are currently understood and there are no data regarding the time of divergence of the species from each other. The species analysed belong, respectively, to the subgenus *Serviformica* (*F. cunicularia, F. fusca, F. gerardi, F. rufibarbis* and *F. selysi*), *Formica* (*F. frontalis*), *Raptiformica* (*F. sanguinea*) and the subgenus *Iberoformica* (*F. subrufa*). All these subgenera include other species in addition to those indicated, with exception of the subgenus *Iberoformica*, defined by Tinaut (1990), which only includes *F. subrufa*, an endemic Iberian species (Bolton, 1995). *F. subrufa* was including in the *Serviformica* subgenus, in agreement with the classification of Wheeler (1913). Morphologic differences (Tinaut, 1990) and cytogenetic data (Lorite et al., 2002c) supported the separation of *F. subrufa* from the *Serviformica* subgenus. The results presented here concerning satellite DNA also support its separation of the *Serviformica* subgenus.

A study was made of the different transition stages in the fixation of randomly produced variant repeats in this satellite DNA by the method of Strachan et al. (1985) as we mentioned before. In agreement with Strachan et al. (1985), the rate of production of new variant repeats is slower than their rate of spread. This hypothesis is supported by the fact that most of the nucleotide positions can be included in one of the six proposed classes. Also according to such authors, classes 3 and 4 have no more than two nucleotides showing polymorphism in one species, when the other species is invariant in these positions. If mutation and spreading were operating at similar rates, it would expect consistently high levels of within-species variation, possibly for variant repeats representing all four bases, at any one position, each of which might have spread to varying extents. In addition, the general paucity of transition stages (with most of the family replacement included in classes 3 and 4) would also indicate that the replacement is relatively fast. Nevertheless, our results are different. We found a high percentage of nucleotide positions that cannot be included in any of these six categories. We also detected the four bases in a same position with a certain frequency, and the family replacement does not appear to be fast. We suggested that the results could be related at least partly to the absence of meiosis in one of sexes, a situation that could change the rate at which new variants spread. If a mutation occurs in a repetitive unit of a male, it will be inherited by all their daughters. The family homogenization and the fixation in the population of this new mutation, or its loss, will depend of the processes of molecular drive in the females.

The spreading of the new variants throughout the repeated family leads to variant homogenization and takes place by means of a variety of genomic turnover mechanisms, involving nonreciprocal DNA transfer within and between chromosomes (gene conversion, unequal crossing-over, slippage replication, transposition, RNA-mediated exchange) (Dover, 1986). Computer analysis suggests a positive correlation between the rate of unequal crossing-over and sequence homogeneity, the former acting as a long-range homogenization force. Clearly, exchange rates cannot be too low relative to the mutation rate in order to maintain the commonly observed high degree of intraspecific sequence homogeneity (Stephan, 1997). Fixation is achieved through the spreading of new variants in the populations as a consequence of chromosome redistribution into new combinations at each generation by the sexual process. In ants, this process can happen only in the females. We suggest that, although the evolution of the satellite DNA in ants could be similar to the rate in other organisms, some particularities could result from haplodiploid system. It is also possible, as Luchetti et al. (2003) have suggested, that ‘‘satellite dynamic appears to be the outcome of both general molecular processes and specific organismal traits’’.

Another important characteristic of the satellite DNA studied in the present work is the existence a CENP-B-like motif in the analysed species, including to *F. subrufa*. A CENP-B-like motif has also been described in the satDNA of ants from the genus *Messor* (Lorite et al., 2002b). The CENP-B-like motif from the genera *Formica* and *Messor* show 65% and 59%, respectively, of sequence similarity with the mammalian CENP-B box. In the present work, by means of in situ hybridization, we showed that the stDNA from *Formica* is located in the centromere region of all chromosomes. The CENP-B-like motif is present in taxonomically different groups of ants, since the genus *Messor* is included in the subfamily Myrmicinae, and the genus *Formica* in the subfamily Formicinae. Based on the similarity of the CENP-B motif to a family of transposases, Kipling and Warburton (1997) argue that CENP-B boxes might be present at centromeres, not because of any role related to centromere function per se, but because of its possible role in modulating genome evolution by inducing recombination hostspots. These new data suggested that the CENP-B-like motif might be involved in the evolutionary

dynamics of satellite DNA, as suggested by other authors (Kipling and Warburton, 1997; López and Edström, 1998; Lorite et al., 2002b).

Several functions have been suggested for satellite DNA but none has yet been demonstrated. We considered that the accumulation of data and specially the comparison between groups with different reproductive strategies can be very useful for the knowledge of the function and evolution of these repeated sequences.

Acknowledgements

We thank Dra. Patrizia D'Ettorre from University of Regensburg (Germany) for providing us samples for the species *F. fusca*, *F. selysi* and *F. sanguinea*. The Spanish PGCDGESIC BOS2002-03217 and the PAI CECJA CVI 220 supported this work. We are grateful for the valuable comments and suggestions on the manuscript given by two anonymous referees.

References

Bigot, Y., Hamelin, M., 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.

Bolton, B., 1995. A New General Catalogue of the Ants of the World. Harvard Univ. Press, Cambridge, MA.

Craig, J.M., Earnshaw, W.C., Vagnarelli, P., 1999. Mammalian centromeres: DNA sequence, protein composition, and role in cell cycle progression. Exp. Cell Res. 246, 249–262.

Dover, G.A., 1986. Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. Trends Genet. 2, 159–165.

Jukes, T.H., Cantor, C.R., 1969. Evolution of protein molecules. In: Munro, H.N. (Ed.), Mammalian Protein Metabolism. Academic Press, New York, pp. 21–132.

Kimura, M., 1980. A simple method for estimating evolutionary rates of bases substitutions. J. Mol. Evol. 16, 111–120.

Kipling, D., Warburton, P.B., 1997. Centromeres, CENP-B and *Tigger* too. Trends Genet. 13, 141–144.

Kumar, S., Tamura, K., Nei, M., 1993. MEGA: Molecular Evolutionary Genetics Analysis. Version 1.01. Pennsylvania State Univesity, University Park, PA.

Landais, I., Chavigny, P., Castagnone, C., Pizzol, J., Abad, P., Vanlerberghe-Masutti, F., 2000. Characterization of a highly conserved satellite DNA from the parasitoid wasp *Trichogramma brassicae*. Gene 255, 65–73.

López, C.C., Edström, J.E., 1998. Interdispersed centromeric element with a CEP-B box-like motif in *Chironomus pallidivittatus*. Nucleic Acids Res. 26, 4168–4172.

Lorite, P., Carrillo, J.A., Palomeque, T., 2002a. Conservation of $(TTAGG)_n$ telomeric sequences among ants (Hymenoptera, Formicidae). J. Heredity 93, 282–285.

Lorite, P., Carrillo, J.A., Tinaut, A., Palomeque, T., 2002b. Comparative study of satellite DNA in ants of the *Messor* genus (Hymenoptera, Formicidae). Gene 297, 113–122.

Lorite, P., Carrillo, J.A., Tinaut, A., Palomeque, T., 2002c. Chromosome numbers in Spanish Formicidae: IV. New data of species from the genus *Camponotus*, *Formica*, *Lasius*, *Messor* and *Monomorium*. Sociobiology 40, 331–341.

Lorite, P., Renault, S., Rouleux-Bonnin, F., Bigot, S., Periquet, G., Palomeque, T., 2002d. Genomic organization and transcription of satellite DNA in the ant *Aphaenogaster subterranea* (Hymenoptera, Formicidae). Genome 45, 609–616.

Lorite, P., Carrillo, J.A., Palomeque, T., 2004. Isolation and characterization of two families of satellite DNA with repetitive units of 135 bp and 2.5 kb in the ant *Monomorium subopacum* (Hymenoptera, Formicidae). Cytogenet. Genome Res. (in press).

Luchetti, A., Cesar, M., Carrara, G., Cavichi, S., Passamonti, M., Scali, V., Mantovani, M., 2003. Unisexuality and molecular drive: Bag320 sequence diversity in *Bacillus* taxa (Insecta Phasmatoidea). J. Mol. Evol. 56, 587–596.

Mantovani, B., Tinti, F., Bachmann, L., Scalli, V., 1997. The Bag320 satellite DNA family in *Bacillus* stick insects (Phasmatodea): different rates of molecular evolution of highly repetitive DNA in bisexual and parthenogenetic taxa. Mol. Biol. Evol. 14, 1197–1205.

Masumoto, H., Masukata, H., Muro, Y., Nozaki, N., Okazaki, T., 1989. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. J. Cell Biol. 109, 1963–1973.

Mravinac, B., Plohl, M., Meštrović, N., Ugarković, Đ., 2002. Sequence of PRAT satellite "frozen" in some coleopteran species. J. Mol. Evol. 54, 774–783.

Nei, M., 1987. Molecular Evolutionary Genetics. Columbia Univ. Press, New York.

Pons, J., Bruvo, B., Juan, C., Petitpierre, C., Plohl, M., Ugarković, Đ., 1997. Conservation of satellite DNA in species of the genus *Pimelia* (Tenebrionidae, Coleoptera). Gene 255, 183–190.

Pons, J., Petitpierre, E., Juan, C., 2002. Evolutionary dynamic of satellite DNA family PIM357 in species of the genus *Pimelia* (Tenebrionidae, Coleoptera). Mol. Biol. Evol. 19, 1329–1340.

Renault, S., Rouleux-Bonnin, F., Periquet, G., Bigot, Y., 1999. Satellite DNA transcription in *Diadromus pulchellus* (Hymenoptera). Insect Biochem. Mol. Biol. 29, 103–111.

Rouleux-Bonnin, F., Renault, S., Bigot, Y., Periquet, G., 1996. Transcription of four satellite DNA subfamilies in *Diprion pini* (Hymenoptera, Symphyta, Diprionidae). Eur. J. Biochem. 238, 752–759.

Rozas, J., Rozas, R., 1999. DnaSP version 3: and integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15, 174–175.

Saitou, N., Nei, M., 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Stephan, W., 1997. Tandemly repeated non-coding DNA sequences. In: Meyers, R.A. (Ed.), The Encyclopedia of Molecular Biology and Molecular Medicine, vol. 6. Verlagsgesellschaft, Weinheim, pp. 1–10.

Strachan, T., Webb, D., Dover, G.A., 1985. Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. EMBO J. 4, 1701–1708.

Tarès, S., Cornuet, J., Abad, P., 1993. Characterization of an unusually conserved *Alu*I highly reiterated DNA sequence family from the honeybee, *Apis mellifera*. Genetics 134, 1195–1204.

Tinaut, A., 1990. Descripción del macho de *Formica subrufa* Roger, 1859 y creación de un nuevo subgénero (Hymenoptera, Formicidae). EOS 65, 281–291.

Ugarković, Đ., Plohl, M., 2002. Variation in satellite DNA profiles—causes and effects. EMBO J. 21, 5955–5959.

Wheeler, W.M., 1913. A revision of the ants of the genus *Formica* (Linné) Mayr. Bulletin of the Museum of Comparative Zoology at Harvard College 53, 379–565.

Wong, A.K.C., Rattner, J.B., 1988. Sequence organization and cytological localization of the minor satellite of mouse. Nucleic Acids Res. 16, 11645–11661.