Letter to the Editor

The Complete Sequence of the Mitochondrial Genome of the Honeybee Ectoparasite Mite *Varroa destructor* (Acari: Mesostigmata)

Maria Navajas,* Y. Le Conte,† M. Solignac,‡ S. Cros-Arteil,* and J.-M. Cornuet*

*Centre de Biologie et Gestion de Populations, Campus International de Baillarguet, Montferrier, France; †UMR UAPV, Institut National de la Recherche Agronomique, Avignon cedex, France; and ‡Laboratoire Populations, Génétique et Evolution, CNRS, Gif-sur-Yvette, France

Complete mitochondrial genome sequences are now available for 28 Arthropoda (see references at: http:// www.jgi.doe.gov/programs/comparative/Mito_top_level. html), more than for any other invertebrate phylum. As in other metazoans, the mtDNA genome is circular and contains 37 genes: 22 for transfer RNAs (tRNA), 2 for rRNAs (*rrnL* and *rrnS*), and 13 for protein subunits (Wolstenholme 1992). There is also one major noncoding region which is thought to play a role in the initiation of transcription or replication (or both) of the DNA molecule (Goddard and Wolstenholme 1978).

Apart from tRNA genes, whose relative positions in the mitochondrial genome vary (Boore 1999), gene order has been long conserved in some lineages of Arthropoda. The arrangement found in most studied insects, for instance, differs from that of the chelicerate Limulus polyphemus by the location of only one tRNA (Staton, Daehler, and Brown 1997; Lavrov, Boore, and Brown 2000). This latter arrangement is believed to be ancestral for arthropods (Boore et al. 1995). However, variants of this gene order have been described. Among the Acari, complete mtDNA sequences are available for two ticks Ixodes hexagonus and Rhipicephalus sanguineus (Black and Roehrdanz 1998), whereas a broad sampling of gene boundaries is available from another tick Boophilus microplus (Campbell and Barker 1999). Notable rearrangements in the gene order were found for two of these ticks: R. sanguineous (Black and Roehrdanz 1998) and B. microplus (Campbell and Barker 1999). These rearrangements involve a translocation of an eight-gene block bounded by the genes nad1 and trnG (along with other tRNA rearrangements). Such major rearrangements of protein-coding mitochondrial genes are not specific to Acari because other important changes have been found in the insect Heterodoxus macropus (Shao, Campbell, and Barker 2001), the crustacean Pagurus longicarpus (Hickerson and Cunningham 2000), and the millipede Narceus annularus (Lavrov, Boore, and Brown 2002). The three species of Acari for which the complete mtDNA gene arrangement has been studied are all in the family Ixodidae (Parasitiformes: Metastigmata) (Black and Roehrdanz 1998; Campbell and Barker 1999). Nothing is known about gene order

Key words: Varroa destructor, mitochondrial genome, mtDNA, gene order, A $\,+\,$ T content.

Mol. Biol. Evol. 19(12):2313-2317. 2002

in other lineages of Acari. We provide here the complete sequence of the mitochondrial genome of *Varroa destructor* (Parasitiformes: Mesostigmata), a neoparasite of the western honeybee (*Apis mellifera*) and hence a major threat for beekeeping all around the world.

Total genomic DNA was isolated from a single female of V. destructor collected from a honeybee colony maintained in the INRA Zoology station at Avignon, France. Total DNA was isolated by a hexadecyltrimethylammonium bromide procedure (Navajas et al. 1998). The entire mitochondrial genome was amplified in two pieces with the Expand Long Template PCR kit (Roche). PCR primers were designed in the *cox1* gene, using the published partial sequence of V. destructor (positions 239-257 in Anderson and Trueman 2000) and the universal primer N4-J-8944 located in the nad4 (Simon et al. 1994). Using these two PCR primers (5'GCGGTT CCCACTGGTATT3' and 5'GGAGCTTCAACATGAG CTTT3') and their reverse complementary sequences, we produced two fragments: fragment 1 of ca. 6,300 bp and fragment 2 of ca. 10,000 bp. Fragment 1 was sequenced by primer walking. Fragment 2 was digested by the Sau3A restriction enzyme, and the resulting reaction was cloned. Two short clones were sequenced, and their respective sequences were used to define new primers. These were used to produce overlapping PCR products that were cloned and sequenced by primer walking. The control region could be localized but not entirely sequenced due to the presence of large (157 bp) repeated elements. Its size could, however, be estimated after amplification with flanking primers.

The amino acid sequences of the protein-coding genes were inferred from the Drosophila mtDNA genetic code. Protein gene sequences were identified by their similarity to published arthropod mtDNA and by alignment with sequences from *R. sanguineus* and *I. hexagonus*. We identified tRNA genes by using their potential to be folded into tRNA-like secondary structures with the tRNscan-SE program-modifying parameters to identify unusual tRNAs (Lowe and Eddy 1997). The tRNA secondary structures were detected in this way for 20 out of the 22 tRNAs. The remaining *trnC* and *trnS*(agn) were detected by inspecting noncoding sequences for the tRNA-like secondary structure by eye.

The entire mitochondrial genome of *V. destructor* has an estimated size of 16,477 bp and belongs to the Korean mitochondrial haplotype described by De Guzman et al. (1998) that has been identified on *A. mellifera* in Europe, the Middle East, Africa, Asia, and the Americas (Anderson and Trueman 2000). The mitochondrial genome size of *V. destructor* is larger than those of the three Acari previously studied, which are ca. 15 kb and

Address for correspondence and reprints: Maria Navajas, Centre de Biologie et Gestion de Populations, Campus International de Bailarguet, CS 30 016, 34988 Montferrier-sur-Lez cedex, France. E-mail: navajas@ensam.inra.fr.

 $[\]ensuremath{\mathbb C}$ 2002 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

						A + T CONTENT	٨T	
				I		Large		
SUBPHYLUM	ORDER	Family	SPECIES	SIZE MT GENOME (bp)	mt Genome	Noncoding Region	Difference	ACCESSION NUMBER
Chelicerata	Dermanyssina	Varroidae	Varroa destructor	16,476	80.0	7.97	0.3	AJ493124
Chelicerata	Ixodida	Ixodidae	Ixodes hexagonus	14,539	72.6	71.8	0.8	AF081828
Chelicerata	Ixodida	Ixodidae	Rhipicephalus sanguineus 1 ^a	14,710	<i>77.9</i>	66.6	11.3	AF081829
Chelicerata	Ixodida	Ixodidae	Rhipicephalus sanguineus 2ª	14,710	77.9	67.3	10.6	AF081829
Chelicerata	Xyphosura	Limulidae	Limulus polyphemus	14,985	67.6	81.3	-13.7	AF002644
Crustacea	Anostraca	Artemiidae	Artemia franciscana	15,770	64.5	68.0	-3.5	X69067
Crustacea	Cladocera	Daphniidae	Daphnia pulex	15,333	62.3	67.0	-4.7	AF117817
Crustacea	Decapoda	Panaeidae	Penaeus monodon	15,984	70.6	81.5	-10.9	AF217843
Hexapoda	Diptera	Culicidae	Anopheles gambiae	15,363	77.6	94.2	-16.6	L20934
Hexapoda	Diptera	Tephritidae	Ceratitis capitata	15,980	77.5	91.1	-13.6	CCA242872
Hexapoda	Diptera	Calliphoridae	Cochliomyia hominivorax	16,022	77.0	90.8	-13.8	AF260826
Hexapoda	Diptera	Culicidae	Anopheles quadrimaculatus	15,455	77.4	93.5	-16.1	L04272
Hexapoda	Diptera	Drosophilidae	Drosophila yakuba	16,019	78.6	92.9	-14.3	X03240
Hexapoda	Diptera	Drosophilidae	Drosophila melanogaster	16,517	82.2	96.0	-13.8	U37541
Hexapoda	Hemiptera	Reduvidae	Triatoma dimidiata	17,019	69.5	66.0	3.5	NC_002609
Hexapoda	Hymenoptera	Apidae	Apis mellifera	16,343	84.9	96.0	-11.1	L06178
Hexapoda	Lepidoptera	Bombycidae	Bombyx mori	15,643	81.3	95.4	-14.1	AF149768
Hexapoda	Orthoptera	Acrididae	Locusta migratoria	15,772	75.3	86.0	-10.7	X80245
Hexapoda	Phthiraptera	Boopidae	Heterodoxus macropus	14,760	79.3	93.0	-13.7	NC_002651
Myriapoda	Helminthomorpha	Spirobolidae	Narceus annularus	14,868	63.7	70.8	-7.1	AY055727
Myriapoda	Lithobiomorpha	Lithobiidae	Lithobius forficatus	15,695	67.9	77.1	-9.2	AF309492
Myriapoda	Spirostrepetida	Hargapagophoridae	Thyropygus sp.	15,133	67.8	64.7	3.1	AY055728
0011 EF 0	-	- - -						

^a Two different control regions have been detected in *R. sanguineus*; long noncoding regions have been detected in *R. sanguineus*.

 Table 1

 Genome Size and A + T Content (in percentage) of the Entire Mitochondrial (mt) DNA Genome and of the Large Noncoding Region of Several Arthropoda

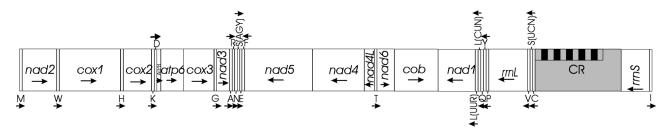


FIG. 1.—Mitochondrial gene arrangement in V. destructor. Arrows indicate the direction of gene transcription. Abbreviations of proteincoding genes and rRNA genes: atp6 and atp8 = ATP synthase subunits 6 and 8; cox1-cox3 = cytochrome oxidase subunits 1–3; cob = cytochrome b; nad1-nad6 = NADH subunits 1–6; nad4L = NADH subunit 4L; rrnL and rrnS = large and small ribosomal subunits RNA. The tRNA genes are denoted by the one-letter amino acid abbreviation. The control region is shaded, and the repeats are represented by boxes.

lies in the upper size range for arthropods (table 1). The gene organization map is given in figure 1. The proteinencoding genes and the rRNA genes are all at the same relative position as in *I. hexagonus* and *L. polyphemus* and are considered ancestral in arthropods, but six and nine of the 22 tRNA genes, respectively, lie in different relative positions. In addition, the control region is located between the rrnL and rrnS ribosomal genes (see discussion below). The position of two tRNA clusters is conserved in V. destructor when compared with insects (Boore 1999): (1) trnA, trnR, trnN, trnS(agn), trnE, and trnF between the nad3 and nad5 genes, and (2) trnK and *trnD* between the *cox2* and *atp8* genes. Translation, initiation, and termination signals as well as the codon usage of the V. destructor mitochondrial genome do not display any unusual characteristics.

In V. destructor, we found the 22 tRNAs that are generally found in metazoan mitochondrial genomes. Identified anticodons were all identical to those of Drosophila. So are those of the ticks R. sanguineus and I. hexagonus, with the exception of the anticodon of the *trnS*(agn), which is UCU instead of GCU in Drosophila and V. destructor. Although several mismatches (seven TT) were found in the arms of the tRNA genes, their putative secondary structure is consistent with those previously described in Acari. In two cases trnC and *trnS*(agn), the structure appeared to be truncated and to lack the dihydrouridine (DHU)-arm found in the canonical tRNA. Instead, these arms are replaced by a DHUarm-replacement loop. These modified structures in V. destructor tRNAs also have been described for two other Acari in the case of the trnS(agn) and for R. sanguineous in the case of trnC (Black and Roehrdanz 1998). Thus, cloverleaf structures seem unlikely for both tRNAs, although RNA analysis is required to confirm the proposed secondary structure. Among tRNAs encoded by metazoan mtDNAs, there is a remarkable diversity of structure. The modified secondary structure of the *trnS*(agn) is relatively common in Metazoa and has been suggested to have occurred early in metazoan evolution (Wolstenholme 1992). The modification of the standard secondary structure of trnC, although less common, also has been reported previously in several animal mtDNAs.

As in all other mitochondrial genomes sequenced so far, two genes for rRNAs were present in *V. destructor*, one for the small (*rrnS*) and one for the large (*rrnL*) ribosomal subunit. We identified the 5'- and 3'-ends of *rrnS* and *rrnL* by comparing their nucleotide sequences with those of the two other Acari sequenced, *I. hexagonus* and *R. sanguineus*. The *trnV* gene, which is located between the two rRNA genes in all arthropods studied so far, except in *Heterodoxus macropus* (Shao, Campbell, and Barker 2001), also lies next to the *rrnL* in *V. destructor*, despite the insertion of two other tRNA genes, *trnS*(ugn) and *trnC*, and of the control region (see below).

There are 2,561 nt unassigned to genes in the sequenced mitochondrial molecule of V. destructor. The largest noncoding region is estimated to be 2,173 nt long and includes several repetitions of a 157-bp motif. This length is well in the range of other arthropods, which shows remarkable variability, from 263 nt for R. sanguineus to 4,601 nt for D. melanogaster, with an extreme size of 9-13 kb in some weevils (Zhang and Hewitt 1997). Outside this region, there are 388 bp unassigned to genes, scattered in short runs (1-93 bp). The largest noncoding region is believed to be involved in the regulation of transcription and replication (Goddard and Wolstenholme 1978); thus, it also is called the control region. This region is known to form secondary structures that are proposed to be involved in the generation of repeats in mtDNA during replication (Broughton and Dowling 1994). Eleven potential sites for the formation of 4- to 26-bp stem-loop structures with loops 6-23 bp were identified in the 460-nt region adjacent to the tandem repeat region. No mismatches were identified in the stems which were supported by 1–7 GC pairs. The 157-bp tandem repeat also has potential for forming two hairpins of 90 and 49 nt with a stable structure supported by 8 and 2 GC matches. These different features strongly suggest that this noncoding region plays a role in DNA replication or transcription initiation (or both) (Saccone, Attimonelli, and Sbisa 1985).

In invertebrates, the relative location of the control region in the genome shows great diversity. Even within the insects, regions flanking the control region may differ among taxa due to tRNA transposition (rearrangement) during evolution. In the case of *V. destructor*, the putative control region lies between the *trnC* and the *rrnS* genes which is unique to this species. The gene cluster *rrnS*—*trnV-rrnL* is present in almost all other Arthropoda, and this arrangement is also common in vertebrates and others (Boore 1999). However, in some other metazoa, as in the Echinoidea Paracentrotus lividus, the two ribosomal genes are separated by a stretch

of about 3.3 kbp which contains the *nad1* and *nad2* genes and a cluster of 15 tRNA genes (Cantatore et al. 1989). Other representatives from the Echinoidea also display two nonadjacent rRNA genes (with the *trnV* gene inserted or not inserted) (Boore 1999).

At the 3'-end of the control region, three nucleotides apart from the 3'-end of the trnC gene, we found an array of tandem repeats extending to a length estimated at 10 repeats of 157 bp plus a truncated repeat of 140 bp. This 1,710-bp region could not be sequenced fully due to the presence of the repeats (no unique internal sequencing primers could be designed), and only the sequence of three repeats at each end of the region could be obtained. All the six sequenced repeats were identical in sequence, suggesting that the remaining repeats also had the same sequence. According to this, the putative control region has the potential of forming 10 identical secondary structures plus a remaining shorter hairpin. Each of the 157-bp tandem repeat has the potential for forming two hairpins of 90 and 47 paired nucleotides with a stable structure supported by 8 and 2 GC matches.

Repeat sequences are common in control regions of many metazoans, and length variation due to variable number of repeats are observed frequently within and among closely related species (e.g., Moritz and Brown 1987; Campbell and Barker 1999; Dotson and Beard 2001). Variation of size and copy number of a repeat unit is responsible to a large degree for the size variation of the control region and also of the whole mitochondrial genome. Copy number variation of the tandem repeat also has been observed at the individual level, leading to the existence of mtDNA molecules of different sizes in an individual, a phenomenon known as length heteroplasmy. Mitochondrial DNA length heteroplasmy has been observed widely in animals (for reviews see Rand 1993; Zhang and Hewitt 1997). Preliminary results obtained through PCR suggest that heteroplasmy is present in V. destructor. However, this has to be confirmed through Southern blot hybridization because PCR is prone to artifacts in such situations (Fumagalli et al. 1996; Campbell, Sturm, and Barker 2001). In any case, the existence of long hairpin secondary structures, such as those found in the 157-bp motif, generally increase the probability of slipped-strand mispairing (see Levinson and Gutman 1987) by bringing closer the motif sequences to be mispaired.

The base composition of the strand that encodes the majority of genes in the mtDNA of *V. destructor* has 39% of adenine, 41% of thymine, 8% of cytosine, and 12% of guanine, which makes an A + T content of 80.0%. Significant variation in nucleotide composition exists among mitochondrial genomes (see table 1 for Arthropoda). The overall A + T content of *V. destructor* is similar to that of two other Acari (*R. sanguineus* 77.9%, *B. microplus* 82.4%) and to those of several arthropods but higher than that of the third acari previously studied, *I. hexagonus* (72.6%), or of the extreme low value detected in the crustacean *Artemia franciscana* (64.5%) (Valverde et al. 1994). The A + T content of the coding strand for protein genes, calculated for

each gene individually, was 80.9%, which is comparable with the value calculated for the entire mtDNA (80.0%). Different subsets of the mitochondrial genome of V. destructor, including protein-coding genes, ribosomal subunits, and control region, were examined for A + Tcontent, and all display comparable values around 80%. The most noticeable result is the A + T content of the noncoding region (79.7%) that is very close to the value estimated for the entire mtDNA (80.0%). This region is usually AT-rich by comparison with the rest of the mtDNA genome. This bias is extreme in the mosquito Anopheles gambiae for which the A + T content of the AT-rich region is 16% higher than that of the entire genome (table 1). This situation applies to most arthropods examined, with the noticeable exception of all mites sequenced so far, together with the hexapod Triatoma dim*idiata* (Dotson and Beard 2001) and the myriapod Thyropygus sp. (Lavrov, Boore, and Brown 2002). These species all have a lower content of A + T in the noncoding region than in the entire genome (table 1). The higher A + T content generally found in the control region could be interpreted as a characteristic favoring the melting of DNA strands, required to initiate replication as well as transcription. The absence of such a higher A + T content in Acari (or even a lower content) may imply a higher efficiency of the enzymatic complex involved in the DNA denaturation before replicationtranscription.

The annotated sequence of the entire mitochondrial genome of *V. destructor* is available from EMBL under the accession number AJ493124.

Acknowledgments

We thank M. Monnerot and M. Caillaud for helpful comments on the manuscript, S. Gimenez for technical assistance, and N. Fradet for his contribution in the identification of some tRNA genes. We also thank J.-Y. Rasplus for help with arthropod systematics. This work was supported by an EC grant, FEOGA B 02712, for beekeeping to Y.L.C.

LITERATURE CITED

- ANDERSON, D. L., and J. W. H. TRUEMAN. 2000. Varroa jacobsoni (Acari: Varroidae) is more than one species. Exp. Appl. Acarol. 24:165–189.
- BLACK, W. C. IV, and R. L. ROEHRDANZ. 1998. Mitochondrial gene order in not conserved in Arthropods: Prostriate and metastriate tick mitochondrial genomes. Mol. Biol. Evol. 15:1772–1785.
- BOORE, J. L. 1999. Animal mitochondrial genomes. Nucleic Acids Res. 27:1767–1780.
- BOORE, J. L., T. M. COLLINS, D. STANTON, and L. L. DAEHLER. 1995. Deducing the patterns of arthopod phylogeny from mitochondrial DNA rearrangements. Nature 376:163–165.
- BROUGHTON, R. E., and T. E. DOWLING. 1994. Length variation in mitochondrial DNA of the minnow *Cyprinella spiloptera*. Genetics **138**:179–190.
- CAMPBELL, N. J., R. A. STURM, and S. C. BARKER. 2001. Large mitochondrial repeats multiplied during the polymerase chain reaction. Mol. Ecol. Notes 1:336–340.

- CAMPBELL, N. J. H., and S. C. BARKER. 1999. The novel mitochondrial gene arrangement of the cattle tick, *Boophilus microplus*: fivefold tandem repetition of a coding region. Mol. Biol. Evol. 16:732–740.
- CANTATORE, P., M. ROBERTI, G. RAINALDI, M. N. GADALETA, and C. SACCONE. 1989. The complete nucleotide sequence, gene organization, and genetic code of the mitochondrial genome of *Paracentrotus lividus*. J. Biol. Chem. 264: 10965–10975.
- DE GUZMAN, L. I., T. E. RINDERER, J. A. STELZER, and D. L. ANDERSON. 1998. Congruence of RAPD and mitochondrial DNA markers in assessing *Varroa jacobsoni* genotypes. J. Apic. Res. **37**:49–51.
- DOTSON, E. M., and C. B. BEARD. 2001. Sequence and organization of the mitochondrial genome of the Chagas disease vector, *Triatoma dimidiata*. Insect Mol. Biol. 10:205–215.
- FUMAGALLI, L., P. TABERLET, L. FAVRE, and J. HAUSSER. 1996. Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. Mol. Biol. Evol. 13:31–46.
- GODDARD, J. M., and D. R. WOLSTENHOLME. 1978. Origin and direction of replication in mitochondrial DNA molecules from *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA 75:3886–3890.
- HICKERSON, M. J., and C. W. CUNNINGHAM. 2000. Dramatic mitochondrial gene rearrangements in the hermit crab *Pagurus longicarpus* (Crustacea, Anomura). Mol. Biol. Evol. 17:639–644.
- LAVROV, D. V., J. L. BOORE, and W. M. BROWN. 2000. The complete mitochondrial DNA sequence of the horseshoe crab *Limulus poluphemus*. Mol. Biol. Evol. **17**:813–824.
- ——. 2002. Complete mtDNA sequences of two millipedes suggest a new model for mitochondrial gene rearrangements: duplication and nonrandom loss. Mol. Biol. Evol. **19**:163–169.
- LEVINSON, G., and G. A. GUTMAN. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Mol. Biol. Evol. **4**:203–221.
- LOWE, T. M., and S. R. EDDY. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequences. Nucleic Acids Res. **25**:955–964.

- MORITZ, C., and W. M. BROWN. 1987. Tandem duplications in aminal mitochondrial DNA: variation in incidence and gene content. Proc. Natl. Acad. Sci. USA 84:7183–7187.
- NAVAJAS, M., J. LAGNEL, J. GUTIERREZ, and P. BOURSOT. 1998. Species wide homogeneity of nuclear ribosomal ITS2 sequences in the spider mite *Tetranychus urticae* contrasts with extensive mitochondrial COI polymorphism. Heredity 80:742–752.
- RAND, D. M. 1993. Endotherms, ectotherms and mitochondrial genome-size variation. J. Mol. Evol. 37:281–295.
- SACCONE, C., M. ATTIMONELLI, and E. SBISA. 1985. Primary and higher order structural analysis of animal mitochondrial DNA. Pp. 37–47 in E. QUAGLIARIELLO et al., eds. Achievements and perspectives of mitochondrial research, Vol. 2. Biogenesis. Elsevier, New York.
- SHAO, R., N. J. H. CAMPBELL, and S. C. BARKER. 2001. Numerous gene rearrangements in the mitochondrial genome of the wallaby louse, *Heterodoxus macropus* (Phthiraptera). Mol. Biol. Evol. 18:858–865.
- SIMON, C., F. FRATI, A. BECKENBACH, B. CRESPI, H. LIU, and P. FLOOK. 1994. Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87:651–701.
- STATON, J. L., L. L. DAEHLER, and W. N. BROWN. 1997. Mitochondrial gene arrangement of the horseshoe crab *Limulus polyphemus* L.: conservation of major features among arthopod classes. Mol. Biol. Evol. 14:867–874.
- VALVERDE, J. R., B. BATUECAS, C. MORATILLA, R. MARCO, and R. GARESSE. 1994. The complete mitochondrial DNA sequence of the crustacean *Artemia franciscana*. 39:400– 408.
- WOLSTENHOLME, D. R. 1992. Animal mitochondrial DNA: structure and evolution. Int. Rev. Cytol. **141**:173–216.
- ZHANG, D. X., and G. M. HEWITT. 1997. Insect mitochondrial control region: a review of its structure, evolution and use-fulness in evolutionary studies. Biochem. Syst. Ecol. 25: 99–120.

Ross CROZIER, reviewing editor

Accepted July 17, 2002