Iranian J Publ Health, Vol. 36, No.2, 2007, pp.25-37

Original Articles

Nuclear Elongation Factor-1α Gene A Molecular Marker for Iranian Sandfly Identification

**P Parvizi [1], M Assmar [2]*

[1]Molecular Systematics Laboratory, Pasteur Institute of Iran, Tehran 13164, Iran
2 Dept. of Parasitology, Pasteur Institute of Iran

(Received 26 Jun 2006; accepted 4 Mar 2007)

Abstract

Background: Elongation factor-1α, a conserved nuclear protein coding gene was used to identify Iranian sandfly species. The phlebotomine sandfies are the vectors of the parasitic protozoan *Leishmania*, the causative agents of leishmaniasis, in Iran.
Methods: Seven sets of primers were tried. PCR amplification of elongation factor-1α was successfully achieved for all 14 species of Iranian sandflies that we caught, but different primers had to be used.
Results: The aligned DNA sequences of 454 bp (without primers) of the gene had the most similarity to a coding region of the elongation factor-1α genes of *D. melanogaster*, as identified by a BLAST search of GenBank. Each Iranian species, except *Phlebotomus caucasicus* and *P. mongolensis*, had a unique combination of nucleotides, i.e. each had a diagnostic sequence. There were no diagnostic sequences for different geographical populations of the species in Iran. We found only a single copy of Ef-1α gene in most individual sandflies. However EF-1α gene was successfully amplified by PCR but, unfortunately, phylogenetic analysis showed that it might be multicopy in sandflies and so the markers could not be trusted.
Conclusion: More highly polymorphic nuclear loci, like microsatellites, might be needed to distinguish morphologically indistinguishable females of the subgenus *Paraphlebotomus*, e.g. *P. caucasiscus* from *P. mongolensis*, in order to resolve their roles as vectors of *Leishmania* species in gerbils.

Keywords: *Elongation factor-1α (EF-1α), sandflies, Iran*

Introduction

Leishmaniasis is caused by infection of human with parasites of the genus *Leishmania* which is transmitted by phlebotomine sandfies (1, 2).The precise identification of sandfly species is of primary importance, and so difficulties are caused when only the males of related species have good morphological characters for diagnosis.

Lane (3) reviewed new taxonomic methods in sandflies, but of all the methods he described, only the ultrastructure of spermatozoa has been applied at the superaspecific level. Most advanced techniques have been used to split species not to unite them.

A number of molecular-level approaches have been taken to reveal markers of taxonomic value for sandflies (4-7). DNA techniques have been applied to study different aspects of the taxonomy, systematics and evolution of sandflies (8-15), and offer a great potential for accurate species identification.

Elongation factor-1α (EF-1α) is a conserved nuclear protein coding gene that could provide markers for identifying Iranian sandfly species.

EF-1α protein is involved in the GTP-dependent binding of charged transfer (t) RNAs to the acceptor site of the ribosome during translation (16). In *Drosophila*, EF-1α occurs as two copies, EF-1α F1 and EF-1α F2, which are expressed at different times during development (17). EF-1α genes have been characterized in other animals, including brine shrimp, mice, humans and honeybees (17-19). Because of the conserved nature of the amino acid sequence among these disparate organisms, EF-1α has been identified as a potentially useful gene for studies of

***Corresponding author:** Fax: +98 21 66465132, E-mail: parp@pasteur.ac.ir

higher-level phylogenetic relationships, especially in insects (18, 19). Amino acid sequences of EF-1α have been used to resolve evolutionary relationships among early eu-karyotes (20) and among arthropod classes (21).

Contrary to an earlier report (22) of a single copy of EF-1α in honeybees, Danforth & Ji (23) identified and characterized an additional copy present in representatives of all major bee families surveyed.

Esseghir et al. (24) had identified and characterized EF-1α for *Phlebotomus* species of the subgenus *Larroussius*, but PCR amplification was not successful for species of the subgenera *Phlebotomus* and *Paraphlebotomus*.

Testa et al. (13) used EF-1α to identify South American sandflies in the subgeneric group Verrucarum of the genus *Lutzomyia* and identified nine genotypes among 47 sequences. The objective of this part of the study was to determine if EF-1α could provide markers of taxonomic value for Iranian sandflies, both for accurate identifycation and for studying phylogenetic relationships.

Materials and Methods

Collection and identification of sandflies

The collections were carried out from 24 July to 27 August 2001 and from 21 August to 12 September 2002, during the main summer season of activity of adult sandflies in Iran. Collection sites were in villages in Isfahan Province, Turkemen Sahara of Golastan Province and also in Varamin and Karaj in Tehran Province, Ahrom and Bushehr in Bushehr Province and in Tonekabone in Mazanderan Province. Sandflies were collected on sticky papers and funnel traps from gerbil burrows and ruins, in CDC miniature light traps from domestic animal shelters and by aspirator from inside houses.

All sandflies were characterized by their species-specific Cyt b sequences (14, 15, 24). Most were also identified based on morphological characters of the head and abdominal terminalia, which were slide-mounted in Berlese fluid following dissection with sterilized forceps and micro-needles (25, 26).

Extraction of DNA from sandflies for PCR of EF-1α gene Total DNA was extracted from the dissected thorax and attached anterior abdomen of individual sandflies using the method of Ish-Horowicz with minor modifications (27).

PCR amplification of EF-1α Seven sets of primers was tried.

At first, the primer-pair of Testa et al. (13) was tried: EF-F03 (forward) with EF-R04 (reverse). Then, because this did not work, the following new primer pairs were desgined by Dr. Paul Ready and used by one of us (PP) (Table 1).

The PCR reaction conditions were: 1 x PCR buffer (Promega), 1.5 mM $MgCl_2$, 0.5 µg of each oligonoucleotide primer, 60 µl of each dNTP (perkin-Elmer), 1.5 µl DNA with the reaction volume completed to 50 µl by distilled water. Following denaturation at 94 ^oC for 5 min and at 80 ^oC for a hot start, PCR consisted of 5 cycles of denaturation at 94 ^oC for 30 sec; annealing for each primer pairs as mentioned above for 30 sec and extension at 72 ^oC for 1 min, followed by 30 similar cycles. Unless stated otherwise, 1.5 µl sandfly DNA was used for PCR amplification of EF-1α.

Direct sequencing of PCR products One hundred nanograms of each purified DNA sample was cycle-sequenced using an ABI Prism® Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit (version 2.0) and AB1 373/377 sequencing systems (ABI, PE Applied Biosystems), with 3.2 pmol of the same primers that were used for PCR.

Aligning and phylogenetic analysis of DNA sequences DNA sequences from both strands were aligned and edited by eye using Sequencher^{Tm} 3.1.1 software (Gene Codes Corporation), based on the amino acid sequences deduced from the universal nuclear genetic code used for *D. melanogaster*. EF-1 α genes were identified by homology to the nu-cleotide and deduced amino acid sequences of *D. melanogaster* and *Larroussius* species (24).

Multiple sequence alignments of both DNA and deduced amino acid sequences were then assembled using SeqPUP (28) followed by the CLUSTAL W Multiple Sequence Alignment Program using default parameters (Clustalw PPC: Clustal W version 1.7; 29).

Phylogenetic analyses were done using PAUP* (30). Relationships were inferred based on genetic distances using the Neighbor Joining (NJ) option with default settings.

Results

Species identification based on morphological charcters Fourteen sandfly species identified based on morphological characters of the head and abdominal terminalia. The study concentrated on taxa found in Iran, which included 4 species of *Phlebotomus (Paraphlebotomus*) *caucasicus, P. (Pa.) mongolensis*, *P. (Pa.) sergenti* and *P. (Pa.) andrejevi;* 2 species of *Phlebotomus (Larroussius) major* and *P.(La.) tobbi;* one specie of *Phlebotomus (Adlerius). halepensis;* one specie of *P. (Synphlebotomus) ansarii;* one specie of *Phlebotomus* (*Phlebotomus) papatasi*; 3 species of *Sergentomyia* (*Sergentomyia) sintoni, S. (Se.) dentate* and *S. (Se) antennata;* and 2 species of *Sergentomyia (Sintonius) clydei* and *S. (Si.) tiberiadis*.

PCR amplification of Elongation factor-1 α haplotypes of* Phlebotomus *and* Sergentomyia *species of different subgenera PCR amplification of EF-1α was successfully achieved for all 14 species of Iranian sandflies that we caught, but different primers had to be used (Table 1, 2). All readable sequences consisted of a single ORF when the sense strand was translated in the second frame using the nuclear genetic code of *D. melanogaster*, and so this suggests a chromosomal origin for the amplified fragments.

Primer pairs EF-F03/ EF-R04 and EF-F05/ EF-R06 were used to try to amplify and sequence fragments from Iranian sandflies that we caught, but they failed to give products or readable sequences.

All primers were used to amplify and sequence fragments from all Iranian sandflies which, we collected and tried are summarized in Table 1. All PCR products produced were of the same size (500 bp without any intron).

Diversity and distribution of Elongation factor-1α haplotypes in* Phlebotomus *and* Sergentomyia *species of different subgenera This analysis was carried out for alignments of the sequences between the primers EF-F05 and EF-R08, but the sequences were sometimes obtained using other pairs of primers.

Genetic distances between pairs of haplotypes were usually low within species (0.00662-0.00680%), but higher between subgenera (0.06689- 0.21658% between five subgenera of *Phlebotomus,* and 0.05959-0.14238 between two subgenera of *Sergentomyia*) and between the two genera (0.08809-0.14238%).

For *P. (Phlebotomus) papatasi* only 6 sequences out of 14 were readable using different pairs of primers and all were unique haplotypes differing by just 1-2 nucleotides. For the other sequences, many nucleotide positions had two bases, suggesting the presence of 2 different alleles in heterozygous flies.

For the subgenus *P.* (*Paraphlebotomus*), genetic distances between species ranged from a low of 0.00220 (*P. caucasicus* / *P. mongolensis* versus *P. (Synphlebotomus) ansarii*) to a high of 0.13928 (*P. caucasicus* versus *P. ansarii*). EF-1α sequences were readable for only 1 out of 7 *P. sergenti*, 2 out of 6 *P. ansarii* and 1 out of 1 *P. andrejevi*. 12 haplotypes were obtained from 26 specimens of *P. caucasicus*/ *P. mongolensis.* Genetic distances were small (0.00220–0.00441, or 1-2 nucleotides), and no haplotype was associated with either of the two male morphotypes or with a geographical region.

Only small numbers of *P. (Larroussius*) species were characterized from only 2 locations (in Tehran and Mazanderan provinces), but there was much diversity: unique haplotypes were found in 7 out of 8 *P. tobbi*; and 3 haplotypes were found in 4 sequences from *P. major*. Some nucleotide positions had two bases, suggesting the presence of 2 different alleles in heterozygous flies. The genetic distance between *P. tobbi* and *P. major* was 0.08391. In contrast, a single hap-

lotype was found in 13 specimens of *P.* (*Adlerius) halepensis* from 1 location in Mazanderan province. Single specimens of 2 unknown species of either the subgenus *Larroussius* or the subgenus *Adlerius* contained unique haplotypes.

EF-1α sequences were obtained for three species of the subgenus *Sergentomyia* (*Sergentomyia*), and inter-specific genetic distances were 0.02413-0.13110. 5 similar EF-1α haplotypes (genetic distances 0.00220–0.00222) were obtained from 9 specimens of *S. sintoni*. There was some evidence for regional associations of haplotypes.

The other two species of subgenus *Sergentomyia* all came from the province of Golastan and showed haplotype diversity, although the genetic distances were small: 0.00220 between the haplotypes from 2 specimens of *S. dentata*, and 0.00328 between the haplotypes from 2 specimens of *S. antennata*.

Only two species of subgenus *S. (Sintonius)* were characterized, one from Bushehr Province and another from the province of Golastan. Genetic distances were small (0.00220-0.00222) within species and larger (0.05959-0.06388) between species. Two haplotypes were found in 2 specimens of *S. tiberiadis* from the province of Golastan, and three haplotypes were found in 4 specimens of *S. clydei* from Bushehr province.

Phylogenetic analysis of Elongation factor–1α gene haplotypes The aligned DNA sequences of 454 bp (without primers) had most similarity to a coding region of the EF-1α genes of *D. melanogaster*, as identified by a BLAST search of GenBank. There was mostly only third position, synonymous substitutions among the sequences.

There was no deduced amino acid sequence polymorphism within taxa or populations of the same species except between IRN234 & IRN362 of *P. ansarii*, IRN377 & IRN412 of *P. papatasi*, and IRN468 & IRN469 of *S. antennata* (Table 3).

The input data matrix of variant characters among 151 amino acids of EF-1α was used for parsimony analysis of all sandfly species captured in Iran and some sandflies and insects from GenBank (Table 2).

The consensus of the thirty most parsimonious trees given by a heuristic search with equal character weighting gave strong support for most sandfly sequences being monophyletic with the gene EF-1α, not EF-2α, except for those from *P. papatasi* and unknown IRN387. Among the EF-1α sequences, most from *P. (Larroussius*) species were in one clade and most for *P. (Paraphlebotomus)* species were in another clade. The genus *Sergentomya* was not resolved as a monophyletic clade. The amino acid sequences for *P. caucasicus* and *P. mongolensis* were the same, providing no diagnostic markers for separating the two morphologically close females.

In total, pairwise genetic distances were 0.00676-0.16000, with the highest between a male of *S. antennata* from Turkemen Sahara (IRN468) and Apis F1, and the lowest between IRN377 and IRN412, the males of *P. papatasi* from Isfahan and Bushehr provinces.

Analysis of DNA haplotypes The input data matrix of variant characters among 454 nucleotides of EF-1α was used for parsimony analysis of most haplotypes of all sandfly species captured in Iran and some sandflies and insects from GenBank. Among the 175 variant characters, 157 were informative and 18 uninformative in maximum parsimony analyses. Nine most parsimonious trees were given by an heuristic search with equal character weighting. The bootstrap 50% majority rule consensus tree (Fig. 1) was congruent with a Neighbor-joining phylogenetic tree (Fig. 1) These trees gave strong support for all sandfly sequences being monophyletic, but did not resolve if EF-1α (Dros F1) or EF-1α (Dros F2) of *D. melanogaster* was the sister gene. Most sequences from flies of the same subgenus were in the same clade, but these clades were not always monophyletic: *P. papatasi* was grouped with *P. (Paraphlebotomus)* species; and *S. (Sintonius)* species with *P. (Larroussius*) species; and again the genus *Sergentomyia* was not resolved as a monophyletic clade. All sequences for *P. mongolensis* were of one haplotype (IRN242) shared with *P. cau-*

casicus, which had 2 other haplotypes, not shared, and so this DNA fragment did not provide diagnostic markers for separating the females of these two species (Fig. 1, 2).

The lowest pairwise genetic distance was 0.00220 between a male of *P. caucasicus* (IRN307) and a female of *P. caucasicus* / *P. mongolensis* (IRN242), both from Isfahan. In total, pairwise genetic distances were 0.00220-0.25691, with the highest for a male of *P. ansarii* from Isfahan (IRN234) and Apis F1.

Amino acids and nucleotides of EF-1α as markers for Iranian sandflies Alignments of 151 amino acids were analysed for all our Iranian sandflies, to look for diagnostic characters for genera, subgenera, species and geographical populations.

There were no fixed, diagnostic amino acid positions for genera and subgenera with more than one species sampled, because their sequences were not monophyletic. Some species had diagnostic amino acids. For other species no diagnostic amino acids were found.

Diagnostic amino acids were found for the species within some subgenera. Each Iranian species, except *P. caucasicus* and *P. mongolensis*, had a unique combination of nucleotides, i.e. each had a diagnostic sequence. There were no EF-1α nucleotides diagnostic for populations of the species that were captured in different regions of Iran.

Note: Nucleotide sequence data reported in this paper are available in GenBank, EMBL and DDBJ databases under accession numbers from *EF416833* to *EF416855*.

Table 1: The primers of Elongation factor-1α gene used in this study

Sets Primers Forward/ Reverse	Primer Forward Sequences	Primer Reverse Sequences	First Annealing Temperature	Second Annealing Temperature
EF-F03 / EF-R04	5´ GCTCCTGGACATCGTGA (T/C)TT 3´	5´ AGTGCTTCGTGGTG TAT(C/T)TC 3´	38 °C	44 °C
EF-F05 / EF-R08	5´ CCTGGACATCGTGATTTCAT 3´	5´ CCACCAATCTTGTA GACATCCTG 3´	44 °C	48 °C
EF-F05 / EF-R10	5´ CCTGGACATCGTGATTTCAT 3´	5´ CCACCAATCTTGTA GACATCTTG 3´	48 °C	52 °C
EF-F07 / EF-R12	5´ GCTCCTGGACATCGTGATTT (C/T) AT 3´	5´ CCACCAATCTTGTA GACATCTTGCAG 3´	50 °C	54 °C
EF-F07 / EF-R14	5´ GCTCCTGGACATCGTGATTT (C/T)AT 3´	5´ CCACCAATCTTGTA GACATCTTGAAG 3´	50 °C	54 °C
EF-F09 / EF-R12	5´ GCTCCTGGACATCGTGATTT (C/T)AT(C/T)AA 3´	5´ CCACCAATCTTGTA GACATCTTGCAG 3´	52 °C	56 °C
EF-F05 / EF-R06	5´ CCTGGACATCGTGATTTCAT 3´	5´ TTACCTTCAGCGTT ACCTTC 3´	44 °C	48 °C

Table 2: Number of successful amplifications of EF-1alpha fragment for different primers and Iranian sandfly species

Primers	*P. (Pa) cauc/P. mong*				*P. (Pa) sergenti**				*P. (Sy) ansarii**				*P. (Pa.) andrejevi**				*P. (La.) tobbi*			
	Total	**P+**	**BS**	**WP**	**Total**	**P+**	**BS**	**WP**	**Total**	**P+**	**BS**	**WP**	**Total**	**P+**	**BS**	**WP**	**Total**	**P+**	**BS**	**WP**
EF-F05/ EF-R06	6	No			4	No			2	No			1	1			7		7	
EF-F05/ EF-R08	30	27	2	1	5	1	4		6	2	4		1		1		7	7		
EF-F05/ EF-R10	6	6			6		6		2		2		1		1		2	2		
EF-F07/ EF-R12	4	4			6		6		2		2		1	1			2	2		
EF-F09/ EF-R12	6	4		2	4			4	3			3	1			1	7	7		
EF-F07/ EF-R14	4	4			6	1	5		6	2	4		1		1		ND			

Table 2: Continued…

Primers	*P. (La.) major*				*P. (Ad.) halepensis*				*P. (Ph.) papatasi**				*S. (Se.) sintoni*			
	Total	**P+**	**BS**	**WP**	**Total**	**P+**	**BS**	**WP**	**Total**	**P+**	**BS**	**WP**	**Total**	**P+**	**BS**	**WP**
EF-F05/ EF-R06	ND								2		2					
EF-F05/ EF-R08	4	4			14	12	1	1	6	3	3		19	9	5	5
EF-F05/ EF-R10	ND				4	3		1	6	4	2	4	4			4
EF-F07/ EF-R12	ND				2	2			6	1	5		2			2
EF-F09/ EF-R12	4	4			11	11			2			2	2			2
EF-F07/ EF-R14	ND				ND				6	3	3		ND			

* These species had very bad and mixed sequnces No= negative ND= not done P+ = positive BS= Bad sequences , WP = Weak PCR, *P. (Pa) cauc/P. mong = P. (Pa) caucasicus/P. (Pa), mongolensis*
For *S. (Se) antennata, S. (Se) dentata, S. (Si) tiberiadis* and *S. (Si) clydei* only primer set EF-F05/EF-R08 was used

Table 3: Input data matrix of variant characters among 151 amino acids of Elongation Factor- 1α, used for parsimony analysis of all sandfly species captured in Iran and some sandflies and insects from GenBank

Morphological Identification of specimen	Species code	haplotype	Amino acid position 111111111111111 166668000111122222344 73468857806783457810 6	DNA haplotypes of the same or other species with shared amino acid haplotype	GenBank accession number
Unknown	unkn	IRN387	VFSSFAPEKPAVENNNKCESA	–	EF416845
P. papatasi	papa	IRN377	VYHAYAVEKPSIDKVERTDSP	–	EF416843
P. papatasi	papa	IRN412	VYHAYAAEKPSIDKVERTDSP	–	EF416844
P. ansarii	ansa	IRN234	VFQAFAATNPSIEKADKTDSP	–	EF416838
P. ansarii	ansa	IRN362	VFHAFAATNPSIEKADKTDSP	–	EF416839
P. andrejevi	andr	IRN333	VENNFAATNPSIDKADNTDSP	–	EF416840
P. caucasicus/	cauc/	IRN242	VFSNFAATNPAIEKADKTDAP	IRN258, IRN307	EF416837/36/35
P. mongolensis	mong				
P. sergenti	serg	IRN195	VFSAFAATNPSIDKADKTDSP	–	EF416841
P. halepensis	hale	IRN277	VFSAYAATNPAVEKADKTDAA	–	EF416842
P. major	majo	IRN287	VFSPFAASNGAVEKADKTEAP	–	EF416834
P. tobbi	tobb	IRN336	VFSAFAASNGAIEKADKTEAP	tobb_Esse[1], pern_Esse (*P. perniciosus*)	EF416833
S. clydei	clyd	IRN403	VFSNFAPTNPAIEKADKTDAP	IRN347 & IRN458 (*S. tiberiadis*)	EF416849/50/51
S. antennata	ante	IRN468	?FSAFAAN?MAIEKADKTEAP	–	EF416852
S. antennata	ante	IRN469	VFSAFAANNLAIEKADKTEAP	–	EF416853
S. dentata	dent	IRN460	VFSAFAASNGAIE????????	IRN462	EF416854/55
S. sintoni	sint	IRN341	VFSNFAPTNPAIEKADKCDAP	IRN398, IRN426	EF416846/47/48
A. mellifera	Apis	Apis F1	VYSAFSPPKPKVENADKTESA	–	
A. mellifera	Apis	Apis F2	VYSTFAVSKPTVEKVEKCETA	–	
D. melanogaster	Dros	Dros F1	DYSAYAPTNPEVGNADKTDAA	–	
D. melanogaster	Dros	Dros F2	DYSAYSPEKPSVEKAEKCDQP	–	
P. ariasi	aria	aria_Esse	VFSPFAASNGAIEKADKTEAP	–	
P. langeroni	lang	lang_Esse	VFSAFAASNGAIEKADKTEAP	long_Esse	
P. neglectus	negl	negl_Esse	VFSPFAASNGAVEKADKTEAP	negl2_Esse, orie_Esse	
P. perfiliewi	perf	perf_Esse	VFSEFAASNGAIEKADKTEAP	–	

(Esse= Esseghir *et al*., 2000), (background black= diagnostic for species; background grey= diagnostic for species within same subgenus)

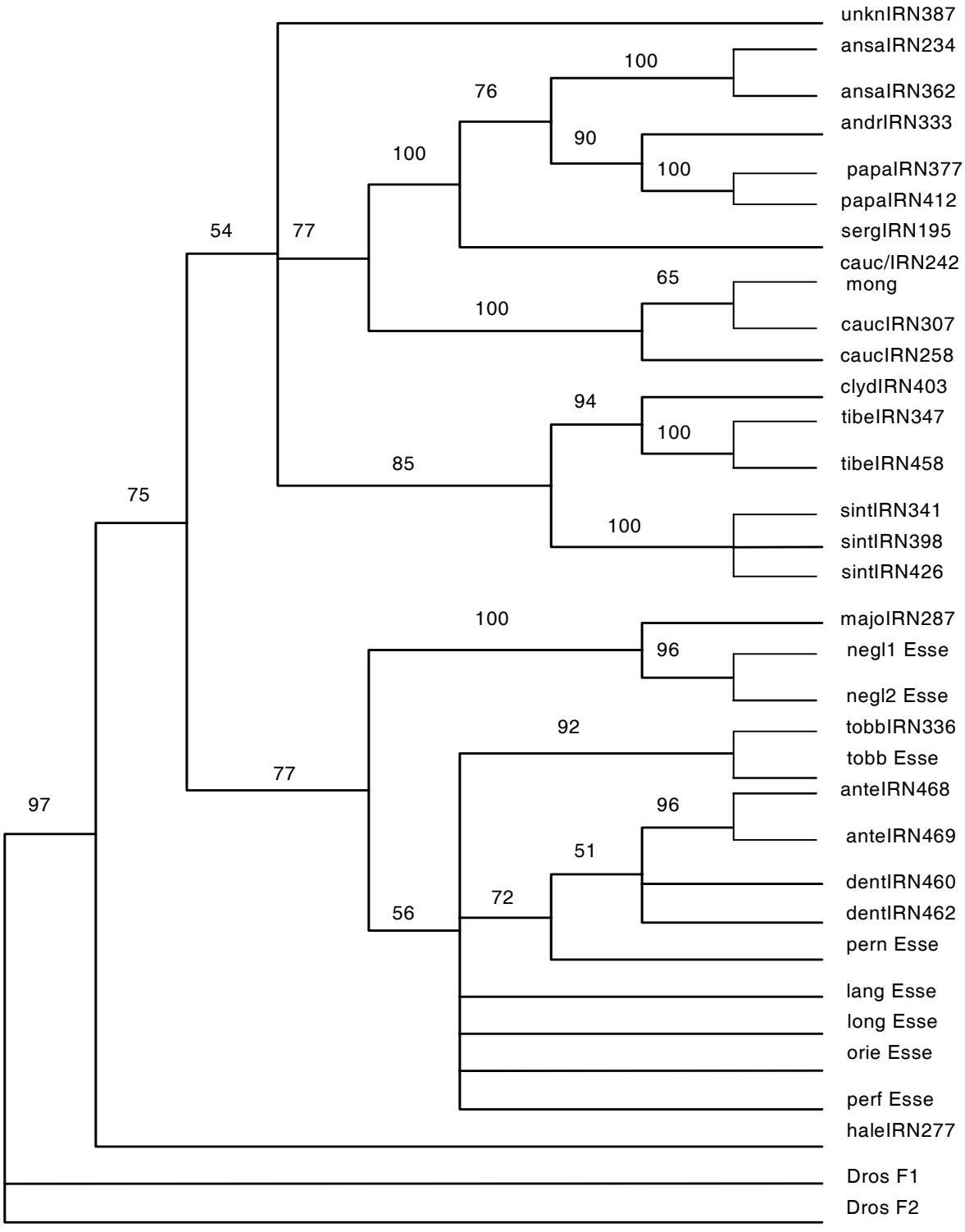

Fig. 1: Bootstrap 50% majority-rule consensus tree for 454 nucleotides of EF- 1α, based on unweighted parsimony analysis with an heuristic search using PAUP*.

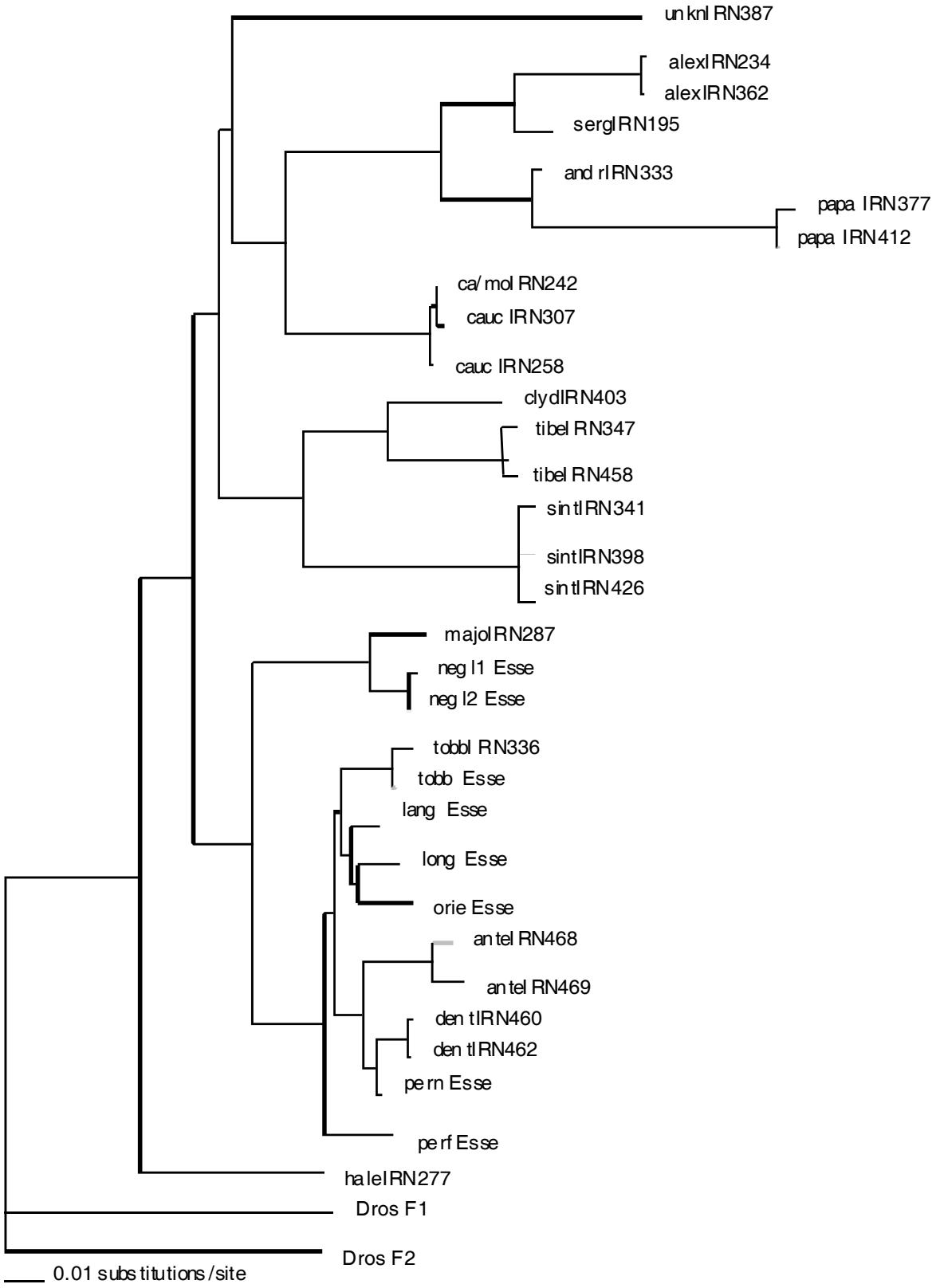

Fig. 2: Neighbor-joining phylogenetic tree for 454 nucleotides sequences of EF- 1α, produced using PAUP*.

Discussion

We have characterized EF-1α DNA sequences of Iranian sandflies in a region of the gene without introns, which was selected based on the primers and relationships of EF-1α sequences reported for other sandflies (13, 24) and insects (23). Additional sandfly-specific primers were designed based on our preliminary sequencing results. These primers worked effectively for PCR amplification of DNA in all species of sandflies tested and for DNA sequencing in most of these species. There might have been difficulties with some sandfly species because of the presence of multiple loci. We translated the nucleotide sequences into amino acid sequences and analyzed the resulting data set of 151 amino acid positions. Some parts of the phylogeny of the amino acid sequences were the same as the species tree. However, other amino acid clades were not monophyletic for species of the same subgenus or genus, and really it is possible that our trees contain a mixture of paraphyletic gene fragments from two gene loci.

Of 454 nucleotide sites, most but not all of the phylogenetic information occured in the third codon positions, and the most substitutions were synonymous. Compared with the amino acid trees, the nucleotide trees were more concordant with the species trees, but not completely. Again it is possible that our trees contain a mixture of paraphyletic gene fragments from two gene loci. This is different from the results of Esseghir *et al.* (24), who found that EF-1α gene trees were almost completely concordant with sandfly species trees, and this showed the phylogenetic informativeness of EF-1α. However, Esseghir et al. (24) only studied *P. (Larroussius)* species, because their primers did not work for *P. (Paraphlebotomus)* species and *P. (Phlebotomus)* species. Unlike the primers of Esseghir et al. (24), our new primers may not have had the specificity for one EF-1α locus.

Most nuclear genes are present in just one or a few copies, simplifying identification of orthologous loci for diagnostics and phylogenetics. While normally viewed as a single copy gene in insects (23), the presence of two copies of EF-1α in distantly related holometabolous orders raises the possibility that two copies are widespread in true flies (Diptera), including sandflies. This could result either from an ancient gene duplication that occurred before the divergence of flies or from parallel gene duplications in the ancestors of different groups of Diptera. Phylogenetic analyses suggest that the gene duplication events occurred independently and in parallel in bees and flies (22, 31). The fly *D. melanogaster* has two copies of EF-1α, called F1 and F2, which are approximately 10% divergent in amino acid sequence (17). The F1 copy has no introns interrupting its coding region, although the 5´ untranslated region has one, whereas the F2 copy has two introns each in its 5´ untranslated and coding regions. Like other Hymenoptera (32), the honey bee (*Apis mellifera*) has two copies of EF-1α, called F1 and F2 as well. The copy F1 has two introns and copy F2 has three introns. In order to determine if the two copies of EF-1α in *Apis* could be homologues of the two copies of EF-1α in *Drosophila*, the authors examined intron position as a criterion of similarity (23). Alignment of coding regions revealed that some intron positions are shared between the *Drosophila* and *Apis* copies, but the positions are uninformative as to the historical relationships among copies (23). Within Hymenoptera, the presence of introns should be advantageous for systematics by permitting convenient separation of gene specific PCR amplification products (32).

We found only a single copy of EF-1α gene in most individual sandflies. This conclusion is based on: 1) finding only one readable DNA haplotype in each specimen, and 2) all the haplotypes of each species being monophyletic (in a unique clade). However, some individual sandflies of some species often gave sequences that were not readable, and this could be because two alleles of EF-1α were often being am-

plified and directly sequenced. The species involved were: *P. (P.) papatasi*, *P. (Pa.) sergenti*, *P. (Pa.) andrejevi*, and *P. (Pa.) ansarii*.

There are two likely explanations for finding a single clade of very similar haplotypes in some species and mixed sequences in other species. The first group of species could contain only one EF-1α gene locus, and the second group could have two loci. An alternative explanation is that there are two loci in all species, but some of the PCR primers may show greater homology to one of the loci in some species. These alternatives could be investigated in the future when cloned sequences from species are available, by considering the genetic distances between sequences and by considering whether the phylogenetic relationships of the gene sequences are congruent with those of the species.

In conclusion the males of *P. mongolensis* shared a haplotype with those of *P. caucasicus* and so, like Cyt b, EF-1α locus/loci could not be used to distinguish the cryptic females of these two species. Otherwise, all species had diagnostic nucleotide sequences.

Acknowledgements

We thank Dr PD. Ready, the leader of sandflies and leishmaniasis programme at the Natural History Museum (NHM), London, for his help, including the design of the new primers for amplifying EF-1α gene fragments and supervising in his laboratory the work reported in this paper, which was first published in the PhD thesis of PP. In the NHM, Julia Llewellyn-Hughes and Johann Testa helped with DNA sequencing.

Refferences

1. Cook GC (1996). *Manson's tropical diseases*. 20th ed. WB Saunders Company Ltd. London, pp. 112-144.
2. Killick-Kendrick R (1990). Phlebotomine vectors of the leishmaniases: A review. *Med Vet Entomol*, 4 (1): 1-24.
3. Lane RP (1993). Sandflies (Phlebotominae). *Medical insects and arachids*. Chapman and Hall, London, pp. 78-119.
4. Ryan L (1986). *Flebótomus do estado do Pará, Brasil. (Diptera: Psychodidae: Phlebotominae).* Documento Tecnico Instituto Evandro Chagas, Belém. Ministerio de Saúde, Belém, pp. 95-112.
5. Phillips A, Le Pont F, Desjeux P, Broomfield G, Molyneux DH (1990). Separation of *Psychodopygus carrerai carrerai* and *P. yucumensis* (Diptera: Psychodidae) by gas chromatography of cuticular hydrocabons. *Act Trop*, 47 (2): 145-49.
6. Pesson B, Wallon M, Floer MT, Kristensen AR (1991). Isoenzymatic studies of Mediterranean populations of sandflies of the subgenus *Larroussius*. *Parassitologia*, 33: (Supplement) 471-6 (in French).
7. Dujardin JC, Le Pont F, Cruz M, Tarrieu LF, Guderian R, Echeverria R, Tibayrenc M (1996). Cryptic speciation in *Lutzomyia* (*Nyssomyia*) *trapidoi* (Fairchild & Hertig) (Diptera: Psychodidae) detected by multilocus enzyme electrophoresis. *Am J Trop Med Hyg*, 54(1): 42-5.
8. Ready PD, Lainson R, Shaw JJ, Ward RD (1986). The ecology of *Lutzomyia umbratilis* Ward and Fraiha (Diptera: Psychodidae), the major vector to man of *Leishmania braziliensis guyanensis* in north-eastern Amazonian Brazil. *Bull Entomol Res*, 76 (1): 21-40.
9. Ready PD, Smith DF, Killick-Kendrick R (1988). DNA hybridizations on squash-blotted sandflies to identify both *Phlebotomus papatasi* and infecting *Leishmania major*. *Med Vet Entomol*, 2(1): 109-16.
10. Ready PD, Day JC, de Souza AA, Rangel EF, Davies CR (1997). Mitochondrial DNA characterization of populations of *Lutzomyia whitmani* (Diptera: Psychodidae) incriminated in the peri-domestic and silvatic transmission of *Leishmania* species in Brazil. *Bull Entomol Res*, 87(2): 187-195.

11. Aransay AM, Scoulica E, Tselentis Y (2000). Detection and identification of *Leishmania* DNA within naturally infected sandflies by Semi-nested PCR on minicircle kinetoplastic DNA. *Appl Environ Microb,* 66: 1933-38.
12. Aransay AM, Ready PD, Morillas-Marquez F (2003). Population differentiation of *Phlebotomus perniciosus* in Spain following post-glacial dispersal. *Heredity,* 90(4): 316-25.
13. Testa JM, Montoya-Lerma J, Cadena H, Oviedo M, Ready PD (2002). Molecular identification of vectors of *Leishmania* in Colombia: mitochondrial introgression in the *Lutzomyia townsendi* series. *Act Trop,* 84 (2): 205-18.
14. Parvizi P, Benlarbi M, Ready PD (2003). Mitochondrial and *Wolbachia* markers for the sandfly *Phlebotomus papatasi*: little population differentiation between peridomestic sites and gerbil burrows in Isfahan province, Iran. *Med Vet Entomol,* 17(4): 351-62.
15. Parvizi P, Mauricio I, Aransay AM, Miles MA, Ready PD (2005). First detection of *Leishmania major* in peridomestic Iranian sandflies: comparison of nested PCR of nuclear ITS ribosomal DNA and semi-nested PCR of minicircle kinetoplast DNA. *Act Trop,* 93 (1): 75-83.
16. Maroni G (1993). *An atlas of* Drosophila *genes*. Oxford University Press, Oxford, pp.22-9
17. Hovemann B, Richter S, Walldorf U, Cziepluch C (1988). Two genes encode related cytoplasmic elongation factors 1 alpha (EF-1 alpha) in *Drosophila melanogaster* with continuous and stage specific expression. *Nucleic Acids Res,* 16: 3175-94.
18. Mitchell A, Cho S, Regier JC, Mitter C, Poole RW, Mathews M (1997). Phylogenetic utility of elongation factor-1α in Noctuoidea (Insecta: Lepidoptera): the limits of synonymous substitution. *Mol Biol Evol,* 14 (4): 381-90.
19. Belshaw R, Quicke DLJ (1997). A molecular phylogeny of the Aphidiinae (Hymenoptera: Braconidae). *Mol Phyl Evol,* 7(3): 281-93.
20. Kamaishi T, Hashimoto T, Nakamura Y, Nakamura F, Murata S, Okada N, et al. (1996). Protein phylogeny of translation elongation factor EF-1α suggests microsporidians are extremely ancient eukaryotes. *J Mol Evol,* 42 (3): 257-63.
21. Regier JC, Shultz JW (1997). Molecular phylogeny of the major arthropod groups indicates polyphyly of crustaceans and a new hypothesis for the origin of hexapods. *Mol Biol Evol,* 14: 902-13.
22. Walldorf U, Hovemann BT (1990). *Apis mellifera* cytoplasmic elongation factor 1 alpha (EF-1 alpha) is closely related to *Drosophila melanogaster* EF-1 alpha. *Federation of European Biochemical Societies Letter* (*FEBS Lett),* 267(2): 245-49.
23. Danforth BN, Ji S (1998). Elongation factor-1α occurs as two copies in bees: Implications for phylogenetic analysis of EF-1α sequences in insects. *Mol Biol Evol,* 15 (2): 225-35.
24. Esseghir S, Ready PD, Ben-Ismail R (2000). Speciation of *Phlebotomus* sandflies of the subgenus *Larroussius* coincided with the late Miocene-Pliocene aridification of the Mediterranean subregion. *Biol J Linn Soc, 70* (2): 189-219.
25. Lewis DJ (1982). A taxonomic review of the genus *Phlebotomus* (Diptera: Psychodidae). *Bull Brit Mus,* 45: 121-209.
26. Nadim A, Javadian E (1976). Key for species identification of sandflies (Phlebotominae; Diptera) of Iran. *Iranian J Publ Health,* 5 (1): 35-44
27. Ready PD, Lainson R, Shaw JJ, Souza AA (1991). DNA probes for distinguishing *Psychodopygus wellcomei* from Psychodopygus complexus (Diptera: Psy-

chodidae). *Mem Inst Oswaldo Cruz*, 86 (1): 41-9.

28. Gilbert DG, (1995). *SeqPUP a biosequence editor and analysis application program*. Biology department, Indiana University, New Albany, USA, Available from: http://ftp.sunet.se/pub/molbio/seqpup/c++/optional/SerqPup-help/.
29. Higgins DG, Thomson JD & Gibson TJ (1997). CLUSTAL W Mutiple Sequence Alignment Program, Version 1.7. Available from: www.sgi.com/chembio/resources/clustalw/
30. Swofford DL (2002). *PAUP*: Phylogenetic Analysis Using Parsimony* (* and other methods) version 4.0. Sinauer Associates, Sunderland, Massachusetts.
31. Whiting MF, Carpenter JM, Wheeler QD, Wheeler WC (1997). The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Syst Biol*, 46(1): 1-68.
32. Rokas A, Nylander JAA, Ronquist F, Stone GN (2002). Maximum-Likelihood analysis of eight phylogenetic markers in Gallwasps (Hymenoptera: Cynipidae): Implications for insect phylogenetic studies. *Mol Phyl Evol*, 22 (2): 206-19.