64 (1) 45–52

Exon-Intron Structure, Paralogy and Sequenced Regions of Elongation Factor-1 alpha in Hexapoda

Marie Djernæs [1, 2] & Jakob Damgaard [1]

[1] Biological Institute and Zoological Museum, University of Copenhagen, Denmark [JDamgaard@bi.ku.dk]
[2] Department of Biology, University of Alberta, Canada [djernaes@ualberta.ca]

Received 17.iii.2006, accepted 22.viii.2006.
Available online at www.arthropod-systematics.de

> Abstract

Elongation factor-1 alpha (EF-1α) is already widely used and shows even more promise for phylogenetic studies of Hexapoda. However, paralogous copies and the presence of introns pose problems. We survey exon-intron structure, presence of paralogous copies and the number and extent of sequenced regions in all hexapod orders. We assess the phylogenetic utility of the exon-intron structure of EF-1α, which is unexpectedly dynamic with widespread losses and several independent instances of intron gain. Paralogous copies of EF-1α are present in Hemiptera, Thysanoptera, Neuropterida, Coleoptera, Hymenoptera and Diptera. With the presented information about exon-intron structure and paralogous copies, researchers will be able to realise the full phylogenetic potential of EF-1α, including exon-intron structure as this can provide additional characters and help to define clades and paralogous copies. We recommend a suitable focus region of 500 bp for future studies of EF-1α in Hexapoda.

> Key words

Elongation factor-1 alpha, exon-intron structure, intron gain, intron loss, sequenced regions, paralogous copies, Hexapoda.

1. Introduction

Hexapods presently comprise half of the known species on the planet, with an estimated number of 5 million living species (Grimaldi & Engel 2005). Until quite recently, taxonomic and phylogenetic questions in this megadiverse group of organisms could only be addressed using phenotypic characters, but today PCR amplification and automated sequencing of "universal DNA markers" offer a plethora of new characters and character systems. In order to increase the synergistic effects between individual studies, Caterino et al. (2000) argued for focusing on just a small handful of the most widely used molecular markers, comprising the mitochondrial gene encoding *cytochrome c oxidase subunit I* (COI); the *large mitochondrial ribosomal subunit* (16S); the *small nuclear ribosomal subunit* (18S); and the nuclear gene encoding *elongation factor-1 alpha* (EF-1α). Apart from being widely used, these four genes represent different genomes (mitochondrial versus nuclear) and different substitution patterns (protein coding versus ribosomal), whereby incongruence in characters and tree topologies can be examined in detail. However, both mitochondrial and ribosomal genes have characteristics that are problematic in phylogenetic analysis (Swofford et al. 1996; Schuh 2000; Damgaard & Cognato 2003; Funk & Omland 2003).

EF-1α is the only one of these candidates that is linked to sexual reproduction *and* easy to align due to its highly conserved amino acid sequence (Regier & Shultz 1997). It is also widely used due to its ease of amplification (few introns and paralogous copies) and the fact that it works at a wide range of taxonomic levels (Cho et al. 1995). Nuclear protein coding genes have their own set of problems, such as heterozygosity (due to biparental inheritance) and poor survival in museum material, partly due to low copy number (2 per cell in diploid organisms compared to hundreds or even thousands of copies of mitochondrial and ribosomal genes). Other problems include the presence of introns and paralogous copies.

Various papers have surveyed exon-intron structure in EF-1α within Hexapoda (Danforth & Ji 1998; Normark 1999; Carapelli et al. 2000; Jordal 2002). Jordal (2002) is the most comprehensive survey and

indicated length of sequenced regions and position of introns in apterygotes, Hemiptera, Coleoptera C1 and C2 [1], Hymenoptera F1 and F2, Lepidoptera and Diptera F1 and F2. Paralogous copies of EF-1α have been reported in several hexapod orders (HOVEMANN et al. 1988; DANFORTH & JI 1998; NORMARK et al. 1999; JORDAL 2002; MORRIS et al. 2002; DOWNIE & GULLAN 2004; HARING & ASPÖCK 2004), but no complete survey has been published. No survey of the sequenced regions of EF-1α exists. CATERINO et al. (2000) surveyed the number and extent of sequenced regions for several genes used in Hexapoda, but for EF-1α they only noted the total number of sequences and the orders for which it was sequenced.

This paper reviews the extent of sequenced regions in EF-1α, their exon-intron structures, and the presence of paralogous copies. Our purpose is to evaluate the phylogenetic utility of EF-1α and recommend a focal region for future studies of hexapod phylogeny at a variety of taxonomic levels.

2. Materials and methods

2.1. DNA sequences and protocols

We searched GenBank (www.ncbi.nih.gov) for EF-1α sequences across all hexapod orders, by searching 'Nucleotide' for 'elongation factor 1 alpha (desired taxon)', e.g. 'elongation factor 1 alpha coleoptera'. In hexapod orders with 20 or fewer EF-1α sequences the number was ascertained by searching 'Nucleotide' and in orders with more than 20 sequences a combination of the approach described above and blasting (nBlast on the NCBI homepage) was used. Blast searches were performed by blasting the coding part of the *Drosophila melanogaster* L. F1 copy[2], as defined by HOVEMANN et al. (1988) and corresponding to positions 2063–3454 in the F1 sequence (GenBank acc.no. X06869), and limiting the blast to the desired insect order. As nBlast does not show all the possible hits if there are several hundreds of them, blasting was performed against subordinate taxa of the relevant order to ascertain the number of sequences.

Due to differences in exon-intron structure among various Hemiptera, unpublished sequences from semi-aquatic bugs (Heteroptera-Gerromorpha) were included, and representatives from other infraorders and families of true bugs were sequenced. The primers Prowler (5'CAG GCT GAT TGY GCT GTA CTT ATY CTT GC 3') and Shirley (5'GCY TCG TGG TGC ATY TCS AC 3') designed by DAMGAARD et al. (2000) were used to PCR amplify and sequence a 570 bp long region in EF-1α corresponding to position 322–892 in the *D. melanogaster* F1 copy. To include other hexapod groups, a newly designed primer pair Manto (5'GGA ACB TCW CAG GCT GAY TGT GC 3') and Phasma (5'GGC GCR AAD GTN ACN ACC ATD CCR GG 3') were used for a 532 bp long region corresponding to positions 313–845 in the *D. melanogaster* F1 copy. DNA extraction, amplification and sequencing follow DAMGAARD et al. (2005) with the following modification: heteropterans were run for 30–35 cycles at an annealing temperature of 50°C, while other hexapods were run for 30 cycles at an annealing temperature of 60°C. Single band amplicons were thereby obtained for Isoptera, Mantophasmatodea, Dermaptera and Grylloblattodea. Several bands were obtained for Archaeognatha. Extraction of DNA from the bands allowed single band amplicons to be amplified in Archaeognatha and upon sequencing one proved to be EF-1α, the rest were short pieces of non-sense DNA. No evidence for multiple copies was found in the material.

Both of the primers Manto and Phasma were shown to be situated across a possible intron site, 324/325 [3] and 823/824 respectively. Consequently taxa which were amplified and sequenced using these primers presumably do not have introns in either of the two positions (JORDAL 2002).

The survey was done in January 2006, and a list of all sequenced individuals is shown in Tab. 1.

2.2. Exon-intron structure

Following JORDAL (2002), introns were identified by misalignment, by AT rich regions and by GT (5') and AG (3') intron terminals, which are canonical for spliceosomal introns (LEWIN 2000; QIU et al. 2004). The exceptions are I 753/754 in two megalopterans and one neuropteran (HARING & ASPÖCK 2004) and I 144/145 in *Pantala flavescens* (F.) (Odonata), which have GC and AG intron terminals. To pinpoint the exact position of each intron, sequences from all taxa possessing the intron were aligned against the *D. melanogaster* F1. Introns in the same position were assumed to be homologous.

[1] C1 and C2, as well as F1 and F2, denotes two distinct copies of EF-1α that occur in the same genome.
[2] Hereafter we use the term '*D. melanogaster* F1 copy' to refer to the coding part of the *Drosophila melanogaster* F1 copy.
[3] Introns are denominated by their position relative to the *D. melanogaster* F1 copy, thus an intron placed between position 144 and 145 is referred to as I 144/145.

Tab. 1. Individuals sequenced in this study.

Order/Infraorder	Species	Locality	Acc. no.
Archaeognatha	*Petrobius* sp.	Denmark	DQ531741
Grylloblattodea	*Galloisiana* sp.	Japan	DQ531727
Dermaptera	*Forficula auricularia* L.	Denmark	DQ531726
Mantophasmatodea	*Austrophasma gansbaaiensis* Klass et al.	South Africa	DQ531740
Isoptera	Hodotermitidae sp.	South Africa	DQ531738
Hemiptera:			
Gerromorpha	*Hebrus pusillus* (Fallén)	Denmark	DQ531730
	Hydrometra gracilenta Horvath	Denmark	DQ531729
	Hermatobates djiboutensis Coutière & Martin	Maldives	DQ531728
	Perittopus asiaticus Zettel	Thailand	DQ531734
	Euvelia sp.	Venezuela	DQ531725
	Microvelia beameri McKinstry	U.S.A.	DQ531731
	Microvelia buenoi Drake	Denmark	DQ531732
	Steinovelia stagnalis (Burmeister)	U.S.A.	DQ531736
	Striduvelia sp.	Nicaragua	DQ531737
	Velia affinis Kolenati	Cyprus	DQ531739
	Neogerris hesione (Kirkaldy)	U.S.A.	DQ531733
Nepomorpha	*Diaprepocoris zealandiae* Hale	New Zealand	DQ531724
Pentatomorpha	*Dolycoris baccarum* (L.)	Denmark	DQ525838
Cimicomorpha	*Reduvius personatus* (L.)	Denmark	DQ531735

Phylogenetic analysis. We performed a phylogenetic analysis of intron positions with present/absent as character states using Dollo (Dollo.up) parsimony in PAUP* 4.0b10 (Swofford 1998) with 1000 random addition replicates and TBR branch swapping. We used well-established superordinal groupings as terminal taxa in some cases: Dictyoptera: combined sequence of Mantodea and Isoptera; Neuropterida: Neuroptera, Megaloptera and Raphidioptera; Amphiesmenoptera: Trichoptera and Lepidoptera. We included only informative characters and terminal taxa with known states for at least $^2/_3$ of the characters as analyses including more characters and taxa were completely unresolved due to unknown character states. We scored introns that were both present and absent within a single order as present as we assumed there were no independent intron gains. The only exception is I 823/824 in Siphonaptera, present only in a single species, as it is unclear whether this sequence represents a paralogous copy. *Artemia salina* (L.) (Crustacea-Branchiopoda) was chosen as outgroup as numerous recent studies have placed Crustacea or part of Crustacea as the sistergroup of Hexapoda (Richter 2002) and Branchiopoda have been supported as the sistergroup by comprehensive and very recent studies (Regier et al. 2005; Mallatt & Giribet 2006). Futhermore, the complete sequence of EF-1α is known for *Artemia salina*. The character matrix is shown in Tab. 2.

3. Results and discussion

3.1. Survey of sequenced regions

Fig. 1 summarises our survey of the sequenced regions of EF-1α of different hexapod groups. No distinction has been made between different copies. Although such a distinction would be relevant and useful, it was deemed too difficult to make, especially for mRNA sequences.

Most hexapod sequences cover a region of 400 bp that corresponds to position 400–800 in the *D. melanogaster* F1 copy. Of the 29 orders for which data is available, only Protura, Ephemeroptera, Mantodea, Thysanoptera, Psocoptera and Phthiraptera lack this region. Endopterygotan orders generally have sequence for a larger region, from position 200 to 1000. Apart from the Endopterygota, long stretches (more than 600 bp) are also available in quantity from Odonata and Hemiptera. The most diverse orders have the highest number of available sequences, with more than 200 sequences deposited in Genbank for the five most speciose insect orders.

Tab. 2. Character matrix for intron presence/absence. 1: presence, 0: absence, ?: unknown. Introns are specified by their position (first line) and by their number in the tree in Fig. 3 (second line).

Taxon	**I 144/145** Intron 1	**I 324/325** Intron 2	**I 492/493** Intron 3	**I 753/754** Intron 4	**I 823/824** Intron 5	**I 1029/1030** Intron 6
Artemia salina	1	1	0	0	0	1
Archaeognatha	?	0	0	0	0	?
Odonata	1	1	0	1	0	1
Grylloblattodea	?	0	1	1	0	?
Dermaptera	?	0	0	0	0	?
Mantophasmatodea	?	0	1	0	0	?
Dictyoptera	1	0	1	1	0	?
Hemiptera	1	1	1	1	0	1
Neuropterida	?	0	0	1	0	?
Coleoptera C1	0	0	0	1	0	0
Coleoptera C2	?	0	0	1	0	1
Hymenoptera F1	0	0	0	0	1	0
Hymenoptera F2	1	0	0	1	0	1
Amphiesmenoptera	0	0	0	0	0	0
Siphonaptera	0	0	0	0	0	0
Mecoptera	0	0	0	0	0	0
Diptera F1	0	0	0	0	0	0
Diptera F2	0	0	0	0	1	1

3.2. Exon-intron structure

Fig. 2 shows our survey of intron positions. In some instances our results disagree with earlier reports. Thysanoptera possess I 1150/1151, also found in Cicadidae (Hemiptera) and Psocoptera, not I 1029/1030 as indicated by MORRIS et al. (2002). COGNATO & VOGLER (2001) reported an intron prior to I 753/754 in some ipine beetles (Curculionoidea). However, this extra intron is an artefact; the only intron found in the ipine sequences is I 753/754. The new alignment does not change the results or conclusions of COGNATO & VOGLER (2001) (A. Cognato pers. comm.). MOULTON (2000) noted the presence of a single intron in *Ectemnia* sp. (Diptera), purportedly not homologous to any of the introns found in the Diptera F2 copy. However, the intron is in position 823/824, the same position as the first intron in the F2 copy.

The Mecopterida (Trichoptera, Lepidoptera, Siphonaptera, Mecoptera and Diptera – no sequences are available for Strepsiptera) have sequences with few or no introns, and seem to have lost introns predominantly in the 5' half of the gene. This contradicts current theory as introns are thought to be lost through reverse transcription which starts at the 3' end and often terminates before reaching the 5' end of the gene (SAKURAI et al. 2002; MOURIER & JEFFARES 2003). Some non-endopterygotan groups, such as Campodeina, Japygina, Archaeognatha and Dermaptera, have no or only one known intron, and thus must have lost introns. Furthermore, as can be seen from Fig. 2, widespread introns have been lost repeatedly and independently in the Hexapoda. Intron gains seem to be the case for I 313/314, present only in Colletidae (Hymenoptera), I 430/431, present only in the Coleoptera C2 copy, I 519/520, only found in Collembola and Japygina, and I 823/824, only present in some Mecopterida and the Hymenoptera F1 copy. As our knowledge of EF-1α expands, additional instances of intron loss and gain will probably be discovered.

Phylogenetic utility of exon-intron structure. CARAPELLI et al. (2000) thought exon-intron structure of EF-1α showed promise as a phylogenetic marker, but other investigators considered exon-intron structure unsuitable due to homoplasy (DANFORTH & JI 1998; WADA et al. 2002).

As an intron is highly unlikely to be gained twice (ROKAS et al. 1999; WADA et al. 2002), but can be lost many times, a phylogenetic analysis of intron position was performed using Dollo parsimony. A 50% majority rule consensus tree (Fig. 3) of the resulting 786 trees is relatively resolved. The presence of I 492/493 and I 753/754 hold the Hexapoda together and I 823/824 holds Hymenoptera F1 and Diptera F2 together. However, most clades are united by intron losses, and these must be regarded as highly homoplasious characters. Intronless sequences thus pose a special problem as can be seen from the erroneous grouping of Archaeognatha and Dermaptera with the endopterygotan orders.

Taking into account that the analysis is based on only six informative characters, it performs remarkably well.

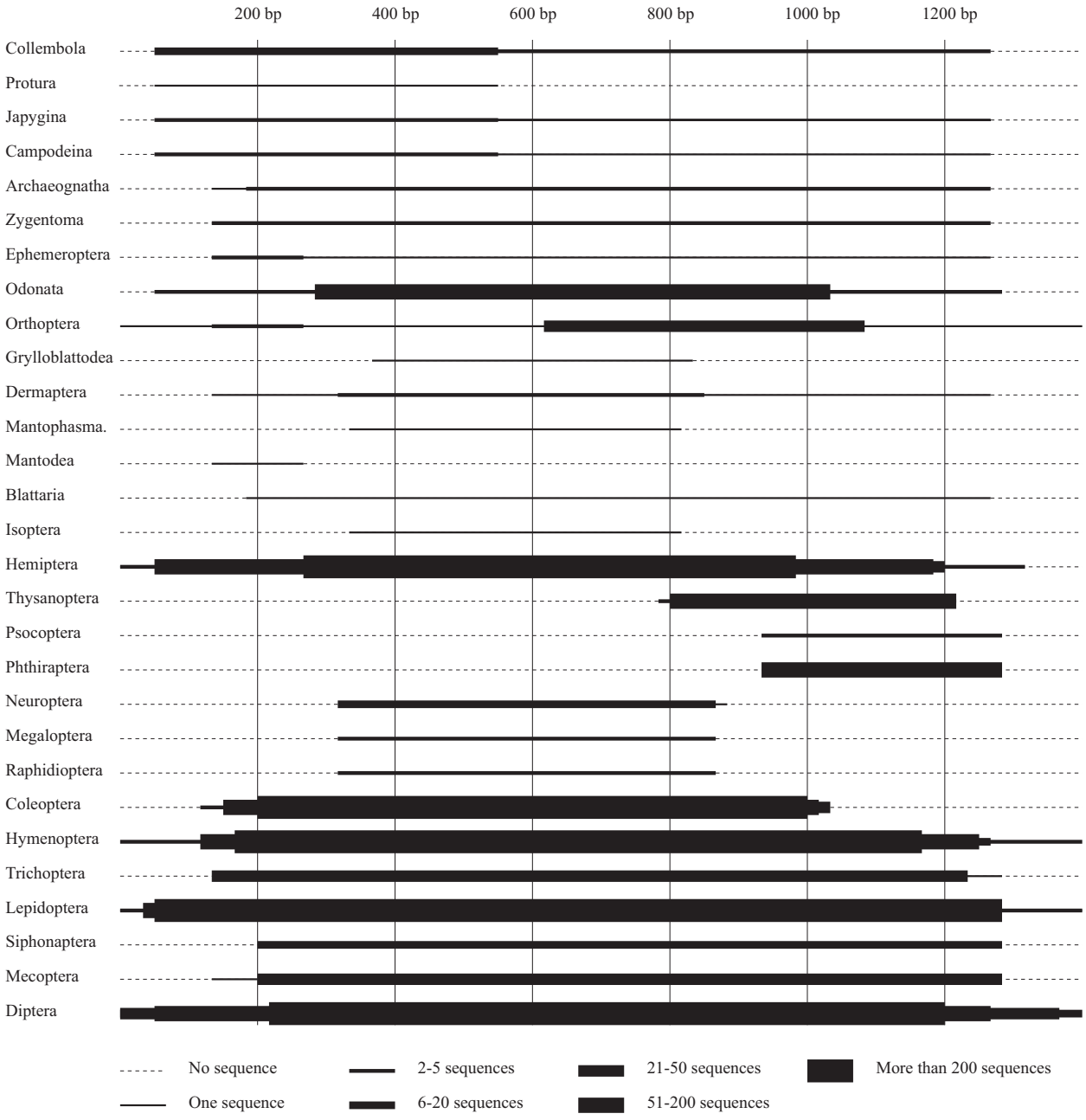

Fig. 1. Survey of sequenced regions and approximate number of sequences available including sequences obtained in this study. Sequences are aligned to the *D. melanogaster* F1 copy.

If intron positions are to be used as a major source of characters, data should be collected from several genes in order to get enough characters. Alternatively, intron positions can provide a few additional characters to a DNA sequence data set.

3.3. Paralogous gene copies

Hovemann et al. (1988) found two copies of EF-1α in Diptera with different expression patterns during ontogenesis; F1 is expressed continuously, while F2 has a stage specific expression. Paralogous copies are now known from several insect orders, including Hymenoptera (Danforth & Ji 1998), Coleoptera (Jordal 2002), Thysanoptera (Morris et al. 2002), Hemiptera-Coccoidea (Downie & Gullan 2004) and members of the Neuropterida (Haring & Aspöck 2004). For Coleoptera and Hymenoptera different functions for the two gene copies are conceivable as the copies appear well differentiated and stable. In other cases extra copies of EF-1α are most likely pseudogenes. In Coleoptera, Hymenoptera and Diptera

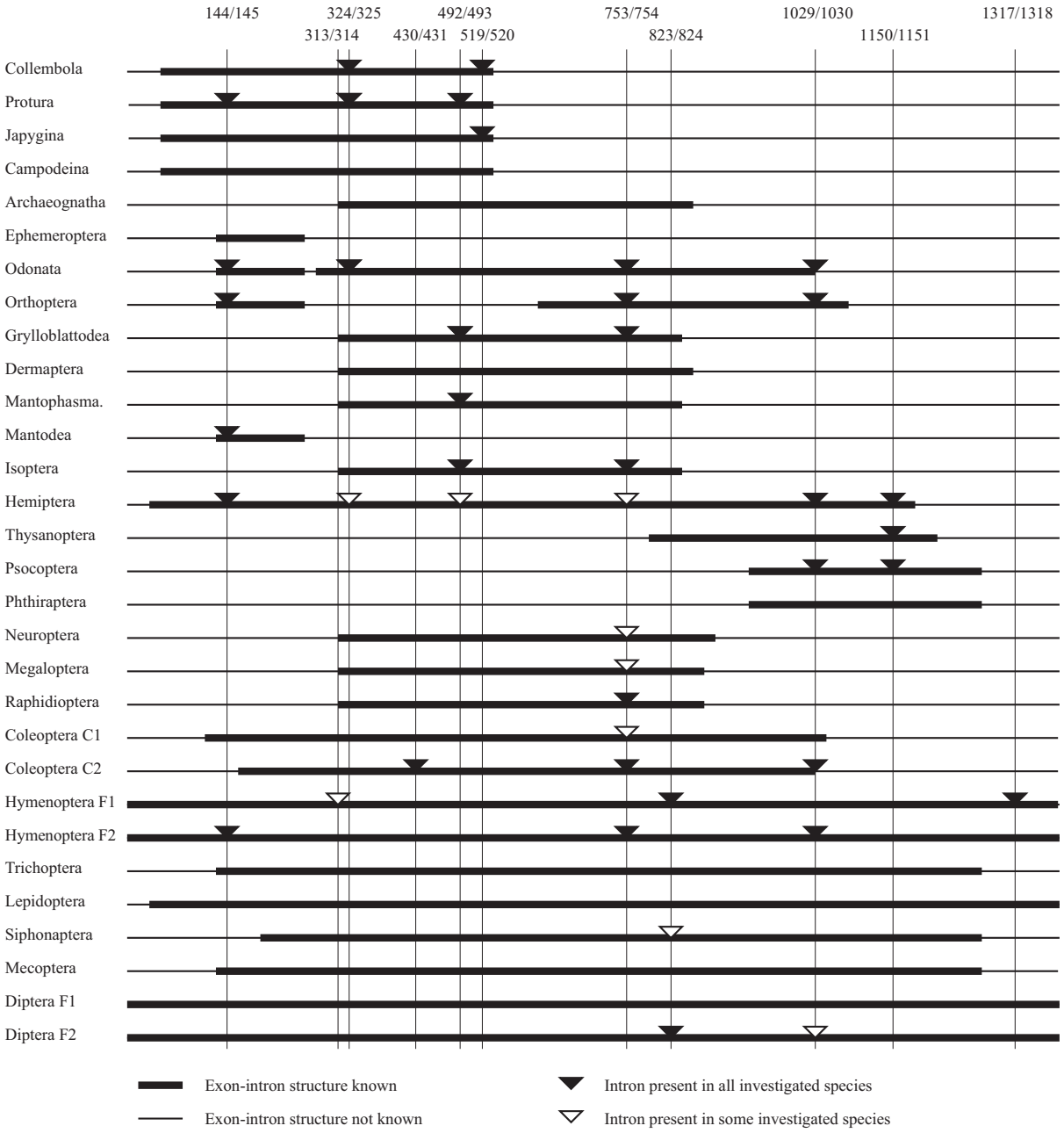

Fig. 2. Known intron positions in Hexapoda. Intraordinal variation in exon-intron structure includes: **Hemiptera**: I 324/325 present in Cicadidae, Coccoidea and Gerromorpha except three closely related *Gerris* species (Damgaard & Djernæs unpubl. data), absent in Aphidoidea; I 492/493 present in Cicadidae, Coccoidea, Aphidoidea, Reduviidae, Corixidae and Pentatomidae, absent in Gerromorpha; I 753/754 present in Cicadidae, Coccoidea, Aphidoidea, Corixidae, Pentatomidae, absent in Gerromorpha except Hebridae and Hydrometridae. **Neuroptera** and **Megaloptera**: I 753/754 absent in some sequences, both types of sequence have been obtained from a single individual (HARING & ASPÖCK 2004). **Coleoptera C1**: I 753/754 absent in some *Curculio* sequences, both types of sequence have been found within a single species (HUGHES & VOGLER 2004; J. Hughes pers. comm.). **Hymenoptera F1**: I 313/314 present only in Colletidae, an apparently novel intron (BRADY & DANFORTH 2004). **Siphonaptera**: I 823/824 present in *Opthalmopsylla volgensis* Smit, but not in any of the other siphonapteran sequences or any of the mecopteran sequences studied. **Diptera F2**: I 823/824 absent in *Ectemnia sp*. The survey was done in January 2006.

the copies can be distinguished based on exon-intron structure and differences in the coding sequence. In Neuropterida differences in exon-intron structure were found, but no clear distinction was evident based on coding sequence (HARING & ASPÖCK 2004). Coccoidea and Thysanoptera show no differences in exon-intron structure for their different copies, but the sequence, including the coding sequence, is different (MORRIS et al. 2002; DOWNIE & GULLAN 2004). Apart from the question of copy homology, JORDAL (2002) found that the Coleoptera C1 copy had higher phylogenetic utility than the C2 copy.

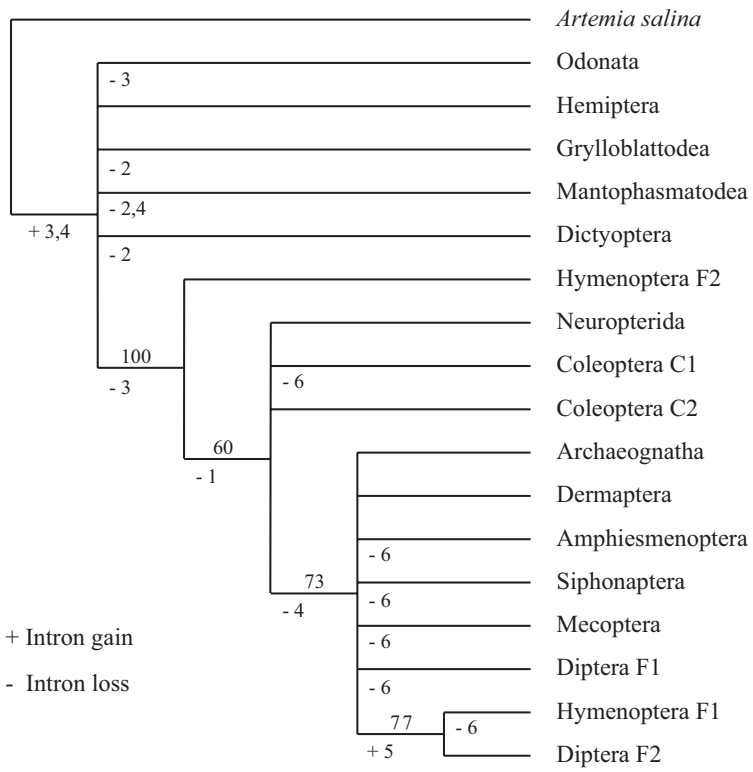

Fig. 3. Phylogenetic tree based on intron positions. Majority rule tree based on 786 most parsimonious trees. 11 steps, CI 0.5455, RI 0.9194. Numbers above branches denotes percentage of most parsimonious trees in which the relevant clade is found. Characters: **1**: I 144/145, **2**: I 324/325, **3**: I 492/493, **4**: 753/754, **5**: I 823/824, **6**: I 1029/1030. Character changes indicated below clades. I 144/145, I 324/325 and I 1029/1030 occur in both *Artemia salina* and various hexapods and might be present throughout Pancrustacea.

3.4. Conclusion

EF-1α is suitable at wide range of taxonomic levels due to its combination of variable third positions and conserved amino acid sequence (Cho et al. 1995). It has been widely used in phylogenetic studies across the Hexapoda despite problems that have prevented full exploitation. These problems stem from features common to all nuclear protein coding genes: low copy number compared to mitochondrial and ribosomal genes, the degenerate third codon position, the presence of introns and often also paralogous copies. Low copy number complicates collection, storage and extraction of suitable template material for amplification. The degenerate third codon position and presence of introns complicates the development of universal primers, and long introns can prevent amplification. Finally, paralogous copies need to be taken into account as they can seriously affect the results of a phylogenetic analysis.

Our survey of intron positions in EF-1α will allow investigators to avoid placing primers across introns as well as to choose intron poor regions for amplification. Based on our surveys of intron positions and sequenced regions, the region between I 492/493 and I 1029/1030 is recommended as a suitable focus region for future studies of EF-1α in Hexapoda. This region contains no more than one intron in any hexapod order and is already available for a great variety and number of hexapods. The exon-intron structure of EF-1α in Hexapoda is dynamic and can provide additional characters in phylogenetic analyses as well as diagnostic markers for different clades and gene copies.

4. Acknowledgements

Thanks to Martin B. Hebsgaard for supplying us with DNA from *D. zealandiae*. Thanks to Dan Jeffares for a useful discussion of the subject of intron loss and gain. Thanks to Doug A. Downie and Joseph Hughes for supplying us with additional information on exon-intron structure in their respective groups. Thanks to Anthony I. Cognato for supplying us with an alignment of EF-1α for all the species included in Cognato & Vogler (2001). We are grateful to Thomas J. Simonsen and Felix A. Sperling for valuable comments on this manuscript. The study was primarily funded by grants from the VILLUM KANN RASMUSSEN Foundation and partially by Felix Sperling's NSERC grant.

5. References

Brady, S.G. & B.N. Danforth 2004. Recent introns gain in elongation factor-1α of colletid bees (Hymenoptera: Colletidae). – Molecular Biology and Evolution **21**: 691–696.

CARAPELLI, A., F. FRATI, F. NARDI, R. DALLAI & C. SIMON 2000. Molecular phylogeny of the apterygotan insects based on nuclear and mitochondrial genes. – Pedobiologia **44**: 361–373.

CATERINO, M.S., S. CHO & F.A.H. SPERLING 2000. The current state of insect systematics: a thriving tower of Babel. – Annual Review of Entomology **45**: 1–54.

CHO, S., A. MITCHELL, J.C. REGIER, C. MITTER, R.W. POOLE, T.P. FRIEDLANDER & S. ZHAO 1995. A highly conserved nuclear gene for low-level phylogenetics: elongation factor-1α recovers morphology-based tree for heliothine moths. – Molecular Biology and Evolution **12**: 650–656.

COGNATO, A.I. & A.P. VOGLER 2001. Exploring data interaction and nucleotide alignment in a multiple gene analysis of *Ips* (Coleoptera: Scolytinae). – Systematic Biology **50**: 758–780.

DAMGAARD, J., N.M. ANDERSEN & F.A.H. SPERLING 2000. Phylogeny of the water strider genus *Aquarius* Schellenberg (Heteroptera: Gerridae) based on nuclear and mitochondrial DNA sequences and morphology. – Insect Systematics and Evolution **31**: 71–90.

DAMGAARD, J., N.M. ANDERSEN & R. MEIER 2005. Combining molecular and morphological analyses of water strider phylogeny (Hemiptera–Heteroptera, Gerromorpha): effects of alignment and taxon sampling. – Systematic Entomology **30**: 289–309.

DAMGAARD, J. & A.I. COGNATO 2003. Sources of character conflict in a clade of water striders (Heteroptera: Gerridae). – Cladistics **19**: 512–526.

DANFORTH, B.N. & S. JI 1998. Elongation Factor-1α occurs as two copies in bees: implications for phylogenetic analysis of EF-1α sequences in insects. – Molecular Biology and Evolution **15**: 225–235.

DOWNIE, D.A. & P.J. GULLAN 2004. Phylogenetic analysis of mealybugs (Hemiptera: Coccoidea: Pseudococcidae) based on DNA sequences from three nuclear genes, and a review of the higher classification. – Systematic Entomology **29**: 238–259.

FUNK, D.J. & K.E. OMLAND 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. – Annual Review of Ecology, Evolution and Systematics **34**: 397–423.

GRIMALDI, D. & M.S. ENGEL 2005. The Evolution of the Insects. – Cambridge University Press, Cambridge. 772 pp.

HARING, E. & U. ASPÖCK 2004. Phylogeny of the Neuropterida: a first molecular approach. – Systematic Entomology **29**: 415–430.

HOVEMANN, B., S. RICHTER, U. WALLDORF & C. CZIEPLUCH 1988. Two genes encode related cytoplasmic elongation factors 1α (EF-1α) in *Drosophila melanogaster* with continuous and stage specific expression. – Nucleic Acids Research **16**: 3175–3194.

HUGHES, J. & A.P. VOGLER 2004. The phylogeny of acorn weevils (genus *Curculio*) from mitochondrial and nuclear DNA sequences: the problem of incomplete data. – Molecular Phylogenetics and Evolution **32**: 601–615.

JORDAL, B.H. 2002. Elongation Factor 1 α resolves the monophyly of the haplodiploid ambrosia beetles Xyleborini (Coleoptera: Curculionidae). – Insect Molecular Biology **11**: 453–465.

LEWIN, B. 2000. Genes VII. – Oxford University Press, Oxford. 990 pp.

MALLATT, J. & G. GIRIBET 2006. Further use of nearly complete 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. – Molecular Phylogenetics and Evolution **40**: 772–794.

MORRIS, D.C., M.P. SCHWARZ, S.J.B. COOPER & L.A. MOUND 2002. Phylogenetics of Australian *Acacia* thrips: the evolution of behaviour and ecology. – Molecular Phylogenetics and Evolution **25**: 278–292.

MOULTON, J.K. 2000. Molecular sequence data resolves basal divergences within Simuliidae (Diptera). – Systematic Entomology **25**: 95–113.

MOURIER, T. & D.C. JEFFARES 2003. Eukaryotic intron loss. – Science **300**: 1393.

NORMARK, B.B. 1999. Evolution in a putatively ancient asexual aphid lineage: recombination and rapid karyotype change. – Evolution **53**: 1458–1469.

NORMARK, B.B., B.H. JORDAL & B.D. FARRELL 1999. Origin of a haplodiploid beetle lineage. – Proceedings of the Royal Society of London Series B **266**: 2253–2259.

QUI, W.-G., N. SCHISLER & A. STOLTZFUS 2004. The evolutionary gain of spliceosomal introns: sequence and phase preferences. – Molecular Biology and Evolution **21**: 1252–1263.

REGIER, J.C. & J.W. SHULTZ 1997. Molecular phylogeny of the major arthropod groups indicates polyphyly of crustaceans and a new hypothesis for the origin of hexapods. – Molecular Biology and Evolution **14**: 902–913.

REGIER, J.C., J.W. SHULTZ & R.E. KAMBIC 2005. Pancrustacean phylogeny: hexapods are terrestrial crustaceans and maxillopods are not monophyletic. – Proceedings of the Royal Society of London Series B **272**: 395–401.

RICHTER, S. 2002. The Tetraconata concept: hexapod-crustacean relationships and the phylogeny of Crustacea. – Organisms Diversity & Evolution **2**: 217–237.

ROKAS, A., J. KATHIRITHAMBY & P.W.H. HOLLAND 1999. Intron insertion as a phylogenetic character: the engrailed homeobox of Strepsiptera does not indicate affinity with Diptera. – Insect Molecular Biology **8**: 527–530.

SAKURAI, A., S. FUJIMORI, H. KOCHIWA, S. KITAMURA-ABE, T. WASHIO, R. SAITO, P. CARNINCI, Y. HAYASHIZAKI & M. TOMITA 2002. On biased distribution of introns in various eukaryotes. – Gene **300**: 89–95.

SCHUH, R.T. 2000. Biological Systematics: Principles and Application. – Cornell University Press, Ithaca. 236 pp.

SWOFFORD, D.L. 1998. PAUP*. Phylogenetic Analysis Using Parsimony (* and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

SWOFFORD, D.L., G.J. OLSEN, P.J. WADDELL & D.M. HILLIS 1996. Phylogenetic inference. Pp. 407–514 *in*: D.M. HILLIS, C. MORITZ & B.K. MABLE. (eds.), Molecular Systematics. 2nd edition. – Sinauer Associates, Sunderland, Massachusetts.

WADA, H., M. KOBAYASHI, R. SATO, N. SATOH, H. MIYASAKA & Y. SHIRAYAMA 2002. Dynamic insertion-deletion of introns in deutorostome EF-1α genes. – Journal of Molecular Evolution **54**: 118–128.