

# Phylogeny of the infraorder Culicomorpha (Diptera: Nematocera) based on 28S RNA gene sequences

JAN PAWLOWSKI,<sup>1</sup> RYSZARD SZADZIEWSKI,<sup>2</sup> DARIUSZ KMIĘCIAK,<sup>3</sup> JOSÉ FAHRNI<sup>1</sup> and GABRIEL BITTAR<sup>4</sup> <sup>1</sup>Department of Zoology and Animal Biology, University of Geneva, Switzerland; <sup>2</sup>Department of Invertebrate Zoology, University of Gdansk, Poland; <sup>3</sup>Institute of Human Genetics, Poznan, Poland; <sup>4</sup>Department of Physical Chemistry, University of Geneva, Switzerland

**Abstract.** Phylogenetic relationships between the families of the infraorder Culicomorpha were investigated by using partial 28S ribosomal RNA gene sequences. All families traditionally placed in this infraorder were investigated and confirmed as clades. On the other hand, some of the morphological relationships between these families were found to be in disagreement with phylogenies based on molecular characters. Our results did not support the generally accepted division of the Culicomorpha into two superfamilies, the Culicoidea (Culicidae + Corethrellidae + Chaoboridae + Dixidae) and the Chironomoidea (Chironomidae + Ceratopogonidae + Simuliidae + Thaumaleidae). Precisely, if the sister-group relationship between Culicidae, Chaoboridae and Corethrellidae was clearly confirmed, the Dixidae, traditionally considered as closely related to these two families, were not placed close to them on our trees. On the other hand, strong evidence was found for grouping together the Simuliidae and the Thaumaleidae, in spite of the cytological and morphological differences between these two families. The position of the Ceratopogonidae was uncertain, and the Chironomidae appeared as a possible sister group to the rest of Culicomorpha. The phylogenetic positions of the groups characterized by feeding on vertebrate blood or insect haemolymph (the Culicidae, Chaoboridae, Ceratopogonidae and Simuliidae) suggest that haematophagy has appeared at least twice in the evolution of Culicomorpha.

## Introduction

The infraorder Culicomorpha includes mosquitoes, black flies and various biting and non-biting midges. The adult females are nectar-feeders and/or blood-feeders, the latter feeding mostly on blood of mammals and birds. Some biting Culicomorpha are well-known vectors of pathogens and parasites responsible for several major human diseases, including malaria, filariasis and onchocerciasis (Lane & Crosskey, 1993). They also play an important role as biting pests affecting economy in many regions of the world. It has been assumed that feeding on vertebrate blood is primitive relative to the other feeding habits in Culicomorpha (Grogan & Szadziewski, 1988).

According to the most widely accepted phylogenetic classification (Hennig, 1973), based on the morphology of all

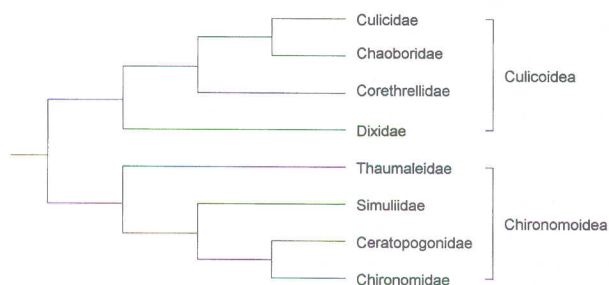
development stages, the Culicomorpha are divided into two superfamilies, the Culicoidea and Chironomoidea. The Culicoidea include three families, the Culicidae, Chaoboridae and Dixidae, while the Chironomoidea include four families, the Chironomidae, Ceratopogonidae, Thaumaleidae and Simuliidae. There were few attempts to modify Hennig's classification. Hackman & Väisänen (1982), on the basis of cytology and wing venation, suggested a close relationship of the Thaumaleidae with the anisopodoid branch and placed that family within the infraorder Bibiomorpha. Subsequently Kovalev (1989, 1990) and Krzeminska *et al.* (1993), solely on the wing venation of the Thaumaleidae, removed that family from the Culicomorpha and placed it in the infraorder Axymiomorpha. Kalugina (1991) removed the family Simuliidae from the Culicomorpha and placed it in the infraorder Ptychopteromorpha. Finally, Wood & Borkent (1989) as well as Borkent & McKeever (1990), on the basis of larval and pupal morphology, proposed a new family Corethrellidae for the single genus *Corethrella*, which was traditionally included in the Chaoboridae (Fig. 1).

Most of these taxonomic modifications resulted from

Correspondence: Jan Pawlowski, Département de Zoologie et Biologie Animale, Université de Genève, 154, route de Malagnou, 1224 Chêne-Bougeries, Switzerland. E-mail: pawlow@sc2a.unige.ch

**Table 1.** List of 11 species of the Culicomorpha examined in this study.

Family	Species	Locality	Date	Stage/sex
Culicidae	<i>Culex pipiens</i>	Wyskok, Poland	1994	Male
	<i>Anopheles maculipennis</i>	Wyskok, Poland	1994	Female
	<i>Aedes</i> sp.	Wyskok, Poland	1994	Male
Chaoboridae	<i>Chaoborus crystallinus</i>	Wyskok, Poland	1994	Female
	<i>Corethrella</i> sp.	Tarcoles, Costa Rica	1993	Female
Ceratopogonidae	<i>Culicoides punctatus</i>	Wyskok, Poland	1993	Female
	<i>Dasyhelea saxicola</i>	Geneva	1994	Male
Dixidae	<i>Paradixa</i> sp.	Ojcow nr. Krakow, Poland	1988	Larva
Simuliidae	<i>Simulium</i> sp.	Tuchola, Poland	1993	Female
Thaumaleidae	<i>Androprosopa larvata</i>	Pieniny Mts, Poland	1988	Larva
Chironomidae	<i>Camptochironomus</i> sp.	Geneva	1994	Male

**Fig. 1.** Cladogram showing the phylogenetic relationships between the families of the Culicomorpha based on morphological analysis (after Borkent & McKeever, 1990).

different interpretations of well studied morphological characters. Until now, the use of molecular data in systematics of Culicomorpha was limited to the study of intraspecific variation and species identification in mosquitoes and biting midges (Paskewitz *et al.*, 1993; Raich *et al.*, 1993). Molecular systematic studies, based on nuclear 18S and 28S DNA sequence data, provided evidence for resolving the phylogeny of different groups of insects (Wheeler, 1989; Carmean *et al.*, 1992; Martin & Pashley, 1992; Weller *et al.*, 1992; Wheeler *et al.*, 1993; Pelandakis & Solignac, 1993; Campbell *et al.*, 1995).

In the present study we investigate the evolutionary history of Culicomorpha using DNA sequence data obtained from the 28S RNA gene from eleven genera representing all families included in this infraorder. Corresponding sequences of *Aedes albopictus* (EMBL L22060) and *Drosophila melanogaster* (EMBL M29800) were added to the alignment, the latter dipteran being used as an outgroup to the Culicomorpha.

## Methods

### Sources of material

A list of genera and species used in this study, as well as their systematic position and collection data, are given in Table 1.

### DNA extraction

DNA was isolated using Chelex 100 (Bio-Rad, Richmond, CA, U.S.A.). This method is very efficient in extracting DNA from different types of samples, including fossil and museum specimens (Walsh *et al.*, 1991). Fresh, alcohol-preserved and dried insects were suspended in 500 µL of 5% Chelex 100 in 1.5-mL microtubes, briefly macerated and shaken overnight at 56 °C. The samples were then vortexed, placed in a dry bath set at 95 °C for 5 min and centrifuged for 5 min. The supernatant, containing the extracted DNA, was transferred to sterile microtubes and stored at -20 °C until needed.

### PCR amplification, cloning and sequencing of PCR products

Polymerase chain reactions (PCRs) were performed in a total volume of 50 µL consisting of 1 × TAQ buffer, 100 µM of dNTPs, 50 µM each of the two primers, 2.5 U Taq DNA polymerase (Boehringer) and 1 µL of the DNA extract. The following primer pairs were used: 10i (5' GTGCAAA-TCGATTGTCAGA) – 15i (5' TCTATCGACTAGAGACTC), 0ic (5' GAAGTTTC(CT)CTCAGGATAGC) – 11 (5' GTTAC-(AG)CACTCCTTA(AG)C(AG)G) and 12i (5' GTAAGCAG-AACTGGTGCT) – 21 (5' GTCAA(AG)CTCAA(AC)AGG-GTC), with their 5' ends situated, respectively, at positions 1053, 2129, 1125, 1462, 1316, and 2728 (starting from the 5' terminal end) of the 28S RNA gene of *Drosophila melanogaster* (Tautz *et al.*, 1988). The PCR products were purified using Spin-Bind DNA extraction units (FMC) and cloned in the pGEM-T Vector System (Promega) using Supercompetent Cells XL2-Blue MRF (Stratagene). Both strands of each fragment were sequenced following the manufacturers' instructions using the *fmol* DNA Sequencing System (Promega), and the [ $\gamma$ -<sup>32</sup>P] ATP end-labelled primers. Sequencing was performed using the amplification primers as well as the internal primers: 12r (5' CCAGTTCTGCTTACCAA), 11r (5' C(CT)G(CT)TAAGG-AGTG(CT)GTAAC) and 14i (5' GGATG(AG)CTGAAGTGGGA), situated at the respective positions 1329, 1444, 1690 of the 28S RNA gene of *D. melanogaster*.

The sequences have been deposited in the EMBL under accession numbers: *Culicoides punctatus* (X89639), *Dasyhelea*

*saxicola* (X89640), *Aedes* sp. (X89641), *Culex pipiens* (X89642), *Anopheles maculipennis* (X89643), *Chaoborus crystallinus* (X89644), *Corethrella* sp. (X89645), *Simulium* sp. (X9646), *Androposopa larvata* (X89647), *Paradixa* sp. (X89648), *Camptochironomus* sp. (X89649).

### Sequence analysis

Gel sequences were aligned with CLUSTAL V (Higgins *et al.*, 1992) as well as by hand, using the Genetic Data Environment (GDE) software, version 2.2, distributed by the Ribosomal Database Project (RDP) (Larson *et al.*, 1993).

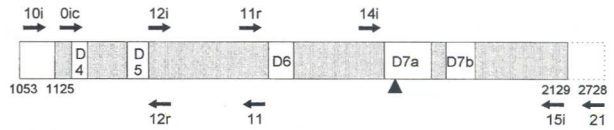
Phylogenetic trees were inferred using four different methods: (i) the neighbour-joining (NJ) method (Saitou & Nei, 1987), as implemented in the GDE and CLUSTAL V software; (ii) a recently developed phenetic compatibility method, Anataxis, found in the *Vivaras* package (Bittar & Carter, 1994; Bittar, 1995; Nadot *et al.*, 1995) (iii) a Wagner maximum parsimony (MP) method, using PAUP 3.1.1. (Swofford, 1993), with gaps counted as 'fifth base'; (iv) a maximum-likelihood (ML) method, as implemented in the DNAML program in the PHYLIP 3.5 package (Felsenstein, 1993) also available in the GDE 2.2.

To evaluate the phylogenetic information content of our sequences, the statistical skewness test described by Hillis (1991) was performed. The statistical robustness of the proposed phylogenies was also analysed, by measuring the frequency of appearance of proposed nodes with, respectively, 500 (MP) and 10 000 (NJ) bootstrap replications (Felsenstein, 1985), and by estimating the decay index of the parsimony branches (Donoghue *et al.*, 1992).

### Results

DNA sequences obtained in this study correspond to the fragment of the large subunit (LSU) 28S RNA gene of *D. melanogaster* starting at position 1125 and ending at position 2136 (Annex 1). This fragment has been chosen because it is composed of a mosaic of variable and conserved regions, and thus is suitable for simultaneous analysis of both closely and distinctly related taxonomic groups. It contains expansion segments D4, D5, D6, D7a and D7b (Hancock *et al.*, 1988) excluding the region of 44 nucleotides processed out of the mature 28S RNA, cleaving the 28S molecule into two similarly sized fragments (28S  $\alpha$  and 28S  $\beta$ ) (Fig. 2). The total number of characters in alignment, including gaps, was 1155. Of these, 287 sites situated in the expansion segments D4, D5, D6, D7a and D7b, of which alignment was difficult or ambiguous, were excluded from phylogenetic analysis (sites containing gaps were only included in the data set if they could be interpreted as a single event). Relative to the three other bases, the base C was clearly under-represented, the average frequencies for bases A, C, G and T being, respectively, 29, 17, 26 and 28%.

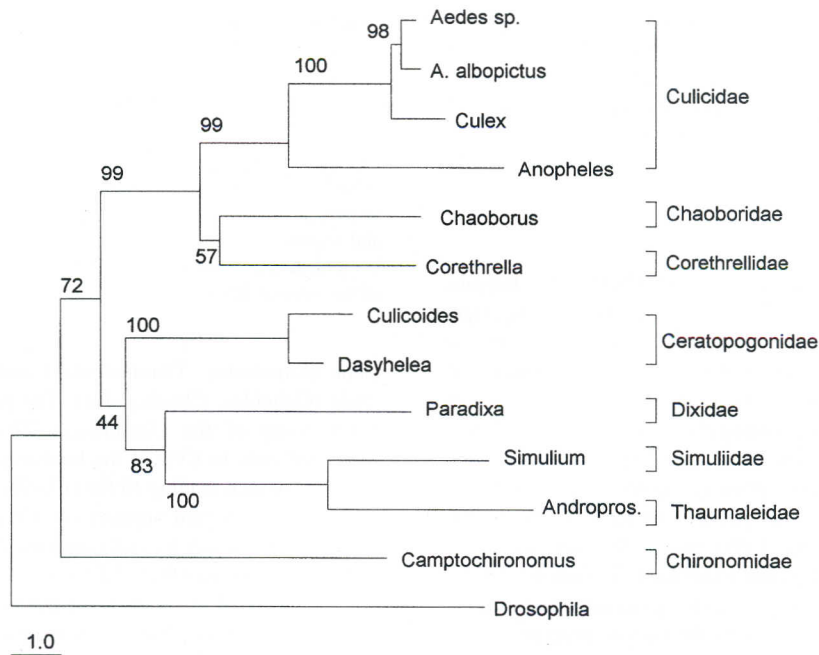
The neighbour-joining phenogram (Fig. 3) established the Culicidae and Ceratopogonidae families as two strong clades, with 99–100% support. There was also 100% support for the



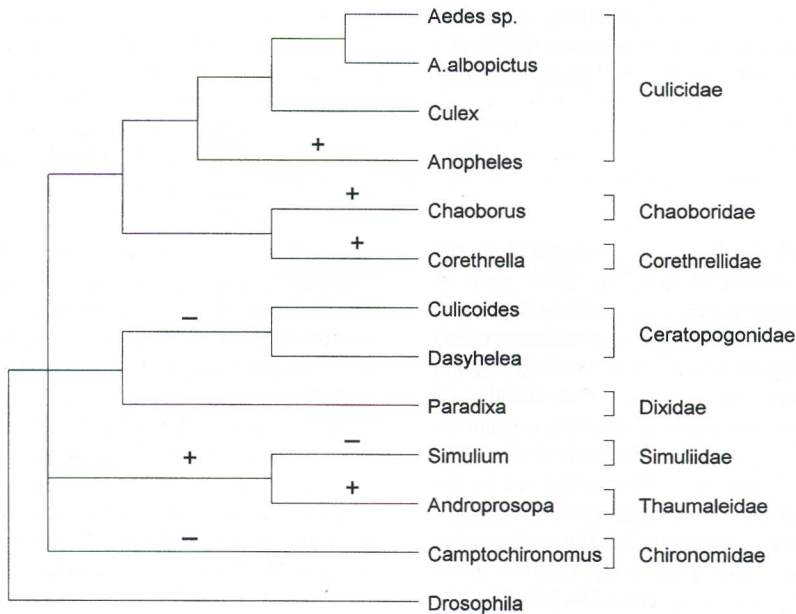
**Fig. 2.** Map of the part of the 28S DNA of *Drosophila melanogaster* (EMBL M29800) corresponding to the fragment amplified and sequenced in 11 species of Culicomorpha. Positions of amplification and sequencing primers are marked. The expansion segments D4–D7 are indicated as unfilled. Black arrow under D7a indicates the position of the internal RNA processing site.

clade (Simuliidae, Thaumaleidae) and 99% support for the clade (Culicidae, Chaoboridae). The position of Dixidae as a sister group of the (Simuliidae, Thaumaleidae) clade was supported only in 83% of the bootstrap replicates, support for the short branch leading to the (*Chaoborus*, *Corethrella*) clade was low (57%), and support for Ceratopogonidae as sister-group to the (*Paradixa* (*Androposopa*, *Simulium*)) clade had an even lower bootstrap probability (44%). Finally, the chironomid species appeared at the base of the tree, as a sister group to other Culicomorpha, but this was only weakly supported, by only 72% of the bootstrap replicates. Maximum likelihood analysis (tree not shown) gave similar results, differing only in position of Ceratopogonidae which appear as sister-group to (Culicidae, Chaoboridae) clade.

The Anataxis phenetic trees-compatibility method (Bittar & Carter, 1994; Bittar, 1995) was performed to test these preliminary results. The advantages of this method can be described in the following way: 'when a data set is large, a numerical phenetics analysis with a distance matrix method can be very fast. However, strongly differing rates of molecular divergence within the different lineages artificially join taxa from slowly evolving lineages, whereas taxa from quickly evolving lineages are pushed out to a basal position in the phenogram. Moreover, the homoplastic characters are not distinguished from homologous ones, with the result that convergences and reversions are not taken into account. The absence of distinction between synapomorphies and symplesiomorphies results in reversible lineages being treated as slowly evolving and thus being artificially grouped. Neither are autapomorphies distinguished from synapomorphies and symplesiomorphies, with thus quickly evolving lineages considered, as already said, as leading to out-group taxa. Therefore a net result of classical phenetical methods is to group taxa sharing plesiomorphies on one hand, and to group together convergent taxa (e.g. sharing common adaptive responses) on the other hand, rather than to group them according to common ancestry. Anataxis was developed as an alternative to such quick but often biased approaches. It is a trees-compatibility distance method that groups taxa while taking into account both the possibility of homoplasy and of different rates of evolution within different branches. It allows a quick analysis of a large set of data by using all possible information without having to analyse the possible evolutionary story of each site (as cladistic parsimony methods do), while nevertheless avoiding the phenetics pitfalls of widely differing evolutionary rates in the different lineages. It should be noted that the Anataxis trees



**Fig. 3.** Neighbour-joining tree of Culicomorpha. Negative distances were freely allowed, sites with gaps were taken into account and distances were corrected according to Kimura's 2-parameters method. The bootstrap probabilities of the nodes for 10 000 replicates are given along the internal branches.

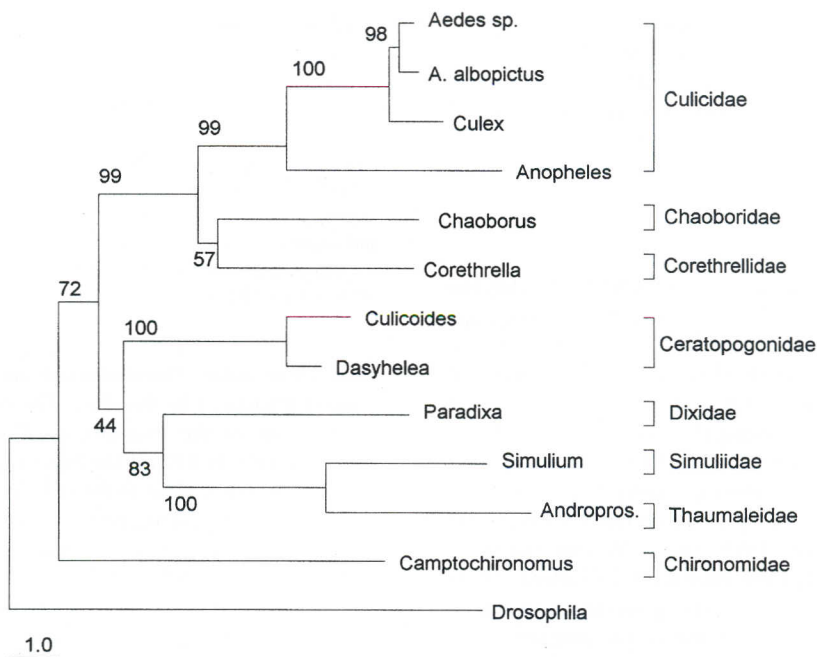


**Fig. 4.** Anataxis tree, obtained by setting *Drosophila* as outgroup to all other species. + over a branch indicates relatively rapid evolution in the lineage, - indicates relatively slow evolution.

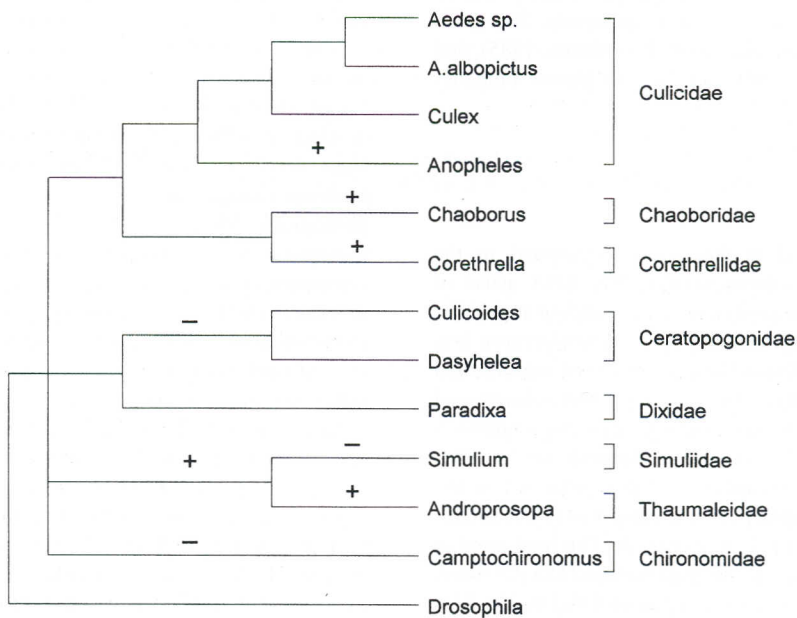
indicate polychotomies when information is not judged sufficient, in order to avoid the unnecessary display of uncertain or unsupported phylogenetic relationships.

The Anataxis tree (Fig. 4) differs from the NJ tree in

proposing *Paradixa* as sister group to the Ceratopogonidae rather than to the (Simuliidae, Thaumaleidae) clade, and in being unable, with the available data, to position *Camptochironomus* as basal to all other Culicomorpha. It must be



**Fig. 3.** Neighbour-joining tree of Culicomorpha. Negative distances were freely allowed, sites with gaps were taken into account and distances were corrected according to Kimura's 2-parameters method. The bootstrap probabilities of the nodes for 10 000 replicates are given along the internal branches.

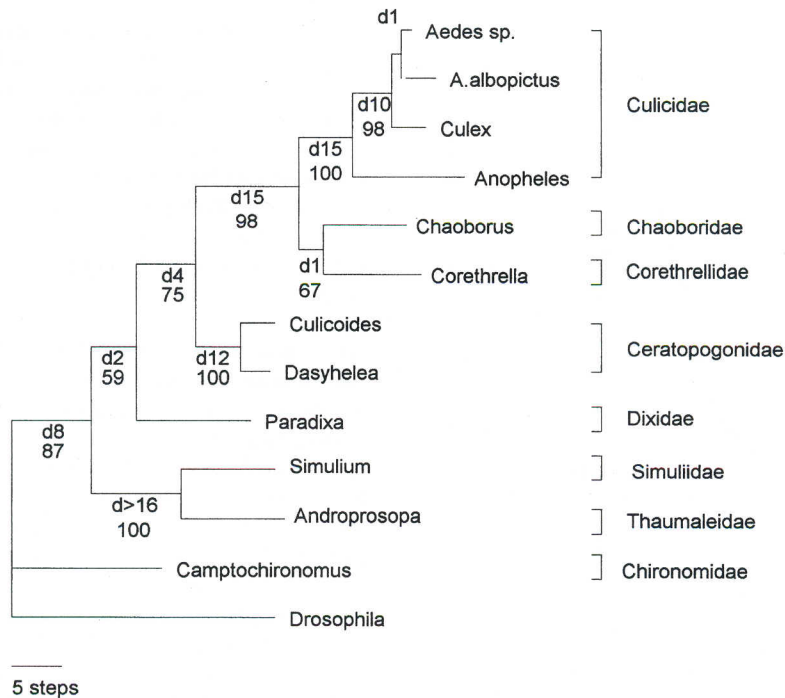


**Fig. 4.** Anataxis tree, obtained by setting *Drosophila* as outgroup to all other species. + over a branch indicates relatively rapid evolution in the lineage, - indicates relatively slow evolution.

indicate polychotomies when information is not judged sufficient, in order to avoid the unnecessary display of uncertain or unsupported phylogenetic relationships.

The Anataxis tree (Fig. 4) differs from the NJ tree in

proposing *Paradixa* as sister group to the Ceratopogonidae rather than to the (Simuliidae, Thaumaleidae) clade, and in being unable, with the available data, to position *Camptochironomus* as basal to all other Culicomorpha. It must be



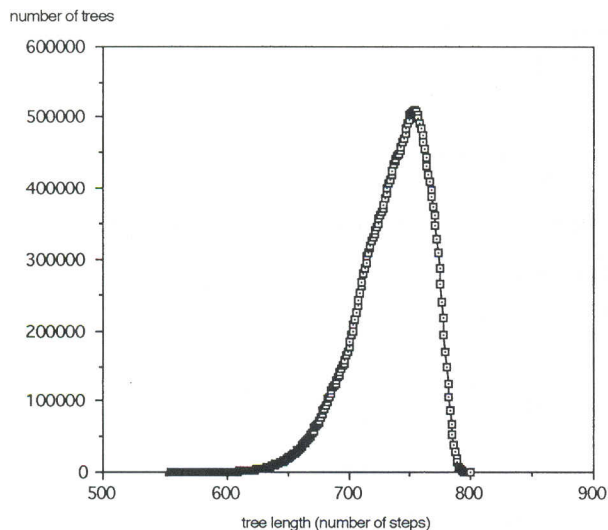
**Fig. 5.** Maximum parsimony phylogram. The branch lengths, branch decay values for strict consensus, and percentage of retention of the node for 500 bootstrap replications, are all indicated. Each gap was treated as a '5th base': minimal tree total length = 674, c.i. = 0.656, c.i.x = 0.570, r.c.i. = 0.386, r.c.i.x = 0.335 (if gaps are not treated as '5th base', length = 603, length excluding uninformative characters = 489, c.i. = 0.643, c.i.x = 0.560, r.c.i. = 0.385, r.c.i.x = 0.335). The decay = 1 tree is the strict consensus of 4 trees, with length = 682 and r.c.i.x = 0.323. The decay = 2 (and decay = 3) tree is the strict consensus of 7 (9) trees, with length = 688 and r.c.i.x = 0.313. The decay = 4 (to 7) tree is the strict consensus of 12 (to 34) trees, with length = 704 and r.c.i.x = 0.289. The characters were considered as unordered, transitions and transversions had the same weight, all invariant characters were ignored, and no character was considered as multistate in any taxon. The tree is unrooted, *Drosophila* is used as an outgroup and the Culicomorpha group was not *a priori* defined as monophyletic. All parsimony trees were constructed with character-state optimization such that reversions (rather than convergences or autapomorphies) were privileged (Farris accelerated transformation, 'acctrans' option). Branches with length zero were collapsed to yield polychotomies. Branch-and-bound compute was via stepwise, addition sequence was furthest, and the MULPARS option was enabled so that all minimal trees could be saved during the search.

noted that the program issued during its calculations a warning of possible homoplasy between the branches leading to *Camptochironomus*, *Paradixa* and the Ceratopogonidae node, and that if the chironomid species is forced to be an outgroup to the other Culicomorpha, *Paradixa* and the Ceratopogonidae cannot any more be joined together; hence this proposed association necessitates more data to be confirmed. Anataxis also indicated that the lineages had evolved at strongly different rates (as expressed by the + and - signs over the branches on Fig. 4).

Maximum-parsimony analysis was also performed on the data set. The branch-and-bound (b&b) algorithm (Hendy & Penny, 1982), guaranteeing to find all minimum length (maximum parsimony) trees was performed on the thirteen sequences. A single minimum-length tree of 674 steps was found (Fig. 5). Its consistency index (c.i.x), excluding 668 uninformative characters (retaining 197 characters), was 0.57, its rescaled consistency index (r.c.i.x) excluding uninformative characters (equal to the product of c.i.x and the retention index r.i.) was 0.335, which are both high enough to indicate that there is not too much homoplasy in the data set for a meaningful phylogenetic analysis (homoplasy index excluding uninformative characters: h.i.x = 1 - c.i.x = 0.43). The MP

tree nodes were tested for reliability by performing 500 bootstrap replicates on the characters. Moreover, the decay value of each branch was calculated (for a given branch, its decay value =  $d$  if it collapses for the strict consensus of all trees smaller than or equal to the minimal length plus  $d$ ). All polychotomies were treated as 'hard' (i.e. a character having changed in the same way in  $x$  different members of a polychotomy accounted for  $x$  steps).

Like in the preceding trees, the Culicinae are strongly supported as a clade. The *Culex* + *Aedes* node is supported in 98% of bootstrap replicates, and the connecting branch decays at  $d = 15$ ; as do the Culicidae (100%,  $d = 15$ ) and the Ceratopogonidae (100%,  $d = 12$ ). This is also the case for the (*Simulium*, *Androprosopa*) clade, in 100% bootstrap replicates and decay index > 16. The MP tree places the Ceratopogonidae as a sister group to the Culicidae + Chaoboridae clade, but the connecting branch appears in only 75% of the bootstrap replications and decays at  $d = 4$ . In a different evolutionary scenario to the NJ tree, *Paradixa* then branches with the [Ceratopogonidae (Culicidae, Chaoboridae)], but the connecting branch appears at only 59% of bootstraps and decays at  $d = 2$ . Finally the chironomid species is placed again



**Fig. 6.** Distribution of tree lengths after a PAUP exhaustive search on the sequences of eleven species listed in Table 1. A total of 34 459 425 trees were thus evaluated, the shortest being 554 steps long, the longest 800 steps long, with a mean length of 735.2 and standard deviation of 29.5. This left-skewed distribution (negative  $g_1 = -0.782$ ) has a median tree length of 739.5 and a mode of 754 steps.

as basal to all other Culicomorpha, but it must be noted that the branch leading to the latter decays at  $d = 8$  and that the node connecting together the non-chironomids appears only in 87% of bootstrap replicates.

So as to be sure that our sequences effectively contained phylogenetic information, an exhaustive search was performed to be able to analyse the distribution of lengths of all possible trees (Fig. 6). Only the eleven sequences obtained by us were thus analysed, so as to concentrate only on the Culicomorpha as a group, to avoid the bias of having with both *Aedes* two rather similar sequences, and last but not least because we could not allocate three weeks of a Quadra Macintosh CPU solely to this task. The  $g_1$  skewness statistic (Sokal & Rohlf, 1981) of this distribution (i.e. the third central moment divided by the cube of the standard deviation), calculated at  $-0.782$ , was sufficiently negative to indicate that there is much less than 1% probability that such a distribution of characters across the eleven sequences could have arisen randomly. Clearly, our data contains phylogenetic signal.

## Discussion

The phylogeny of Culicomorpha inferred from the 28S DNA sequences confirmed monophyly of Culicidae and Ceratopogonidae. However, with that small available data, paraphyly of the Culicomorpha cannot be excluded. The relationships between the families of the Culicomorpha are partly in disagreement with those previously proposed. The distinction of two superfamilies, the Culicoidea and Chironomidea, accepted in all recent classifications of Culicomorpha (Steyskal, 1974; Saether, 1977; Wood & Borkent, 1989), was not supported by our results. Out of the eight families recognized

in this infraorder, five were grouped here into two clades: (Culicidae, Chaoboridae, Corethrellidae) and (Simuliidae, Thaumaleidae) and three: the Dixidae, Ceratopogonidae and Chironomidae, branched separately.

The phylogenetic relationships between the families Culicidae and Chaoboridae are in agreement with the traditional classification. Both families share in common some morphological characters of the larval stage, unique in Culicomorpha, such as an enlarged thorax and long antennae. The long antennae, present also in Dixidae, are considered synapomorphy for Culicoidea (Hennig, 1973). The phylogeny of the Culicidae agreed with the distinction of two subfamilies, Culicinae and Anophelinae. However, the phylogenetic relationships between the species of Culicinae could be resolved only when the sequence data from the expansion segments were added. In all our analyses, the Corethrellidae are placed as a sister group of the Chaoboridae. This is in disagreement with the placement of the Corethrellidae as a sister group to Chaoboridae + Culicidae proposed by Wood & Borkent (1989) on the basis of larval morphology. In view of our data, the taxonomic status of the family Corethrellidae may need to be revised.

In contrast to all previous classifications, our study demonstrated that Thaumaleidae form a sister group of the Simuliidae. Close relationship of both families was never suggested or expected on the basis of morphological studies. The recent suggestion that Thaumaleidae should be removed from Culicomorpha (Hackman & Väisänen, 1982; Kovalev, 1989, 1990; Krzeminska *et al.*, 1993) was based on cytological evidence that, contrary to the Simuliidae, Culicidae and Chironomidae, the Thaumaleidae have cytologically distinguishable X and Y chromosomes but lack the chiasmata observed in the autosomal bivalents during male meiosis (White, 1949). There are, however, some similarities between Thaumaleidae and Simuliidae adult morphology and larval ecology, both are associated exclusively with running water. The amphipneustic larvae of the Thaumaleidae live in cold streams on wet rocks covered with a film of water. The apneustic larvae of the Simuliidae attach to various submerged objects in many types of running waters. The adults of both families are stout, with short antennae composed of 7–10 flagellomeres in both sexes. Antennae lack plume in males and have a small pedicel bearing the Johnston's organ. The shape of their antennae is unique within the Culicomorpha and may be recognized as a valid synapomorphy.

The present study cannot define precisely the position of the Dixidae. NJ and ML analyses would suggest that the Dixidae form a sister group of Thaumaleidae + Simuliidae but with weak statistical support. Traditionally the Dixidae were associated with the Chaoboridae and Culicidae and their position has never been questioned (Hennig, 1973; Hackman & Väisänen, 1982; Wood & Borkent, 1989; Borkent & McKeever, 1990). The Dixidae, Simuliidae and Thaumaleidae have few common morphological characters. Their males have antennae without plume and with 14 flagellomeres like the Chironomidae, while the males of the Chaoboridae, Culicidae and Ceratopogonidae have antennae with 13 flagellomeres, with well developed plumes. Moreover, in adult Dixidae, Thaumaleidae and Simuliidae the vertex, as in the Ceratopogonidae, lacks a

sagittal suture present in Chironomidae, Culicidae and Chaoboridae.

The Chironomidae and Ceratopogonidae are traditionally considered as sister groups (Hennig, 1973). However, according to our data the chironomids could be a sister group to the rest of Culicomorpha, while the position of Ceratopogonidae is uncertain. They branch together either with Culicidae + Chaoboridae (MP, ML) or rather with Dixidae (NJ, Anataxis). The position of Chironomidae seems to be in agreement with palaeontological evidences showing that this family, together with Chaoboridae, is the oldest group of Culicomorpha known from the Lower Jurassic (Kalugina, 1977, 1991). Among the other Culicomorpha, *Corethrella* was described from the Lower Cretaceous (R. Szadziewski, unpubl. data), while the oldest Culicidae are known from the Upper Eocene (Kalugina, 1991).

Our molecular study has brought new questions and information on the origin and evolution of Culicomorpha. This infraorder raises interesting questions on the phenomenon of haematophagy (feeding of the females on blood and haemolymph) in Nematocera. Our results suggest independent origins of haematophagy in two families of exclusive blood-feeders, the Culicidae and Simuliidae. The Chaoboridae, which include a few haematophagous species, share a common ancestor with Culicidae, while the haematophagy in Ceratopogonidae seems to have evolved independently. Among the families of exclusive nectar-feeders and those in which adults do not feed at all, the Chironomidae could be placed as a sister group to the rest of the Culicomorpha, while the Thaumaleidae are placed in the same clade as the haematophagous Simuliidae, and the Dixidae branch separately. It can be concluded that blood feeding is plesiotypic in mosquitoes (Culicidae) and phantom midges (Chaoboridae). Whether it is a plesiotypic feature for all Culicomorpha, as suggested by Borkent (in press), and by the recent discovery of functional piercing mouthparts in an extant chironomid of the genus *Archaeochnus* (Cranston *et al.*, 1987) cannot be confirmed by the present study. It seems more probable, that nectar-feeding was a plesiotypic feeding habit for Culicomorpha, while a supplementary feeding on blood or haemolymph evolved secondarily several times.

In conclusion, the phylogeny of the Culicomorpha inferred from the 18S DNA sequences differs significantly from that based on the morphological studies. The molecular data confirm a close relationship between the Culicidae, Chaoboridae and Corethrellidae, which are traditionally included in the superfamily Culicoidea. The Dixidae, however, which are considered as closely related to the Culicoidea, were placed separately from them in the molecular trees. In view of our data, the taxonomic status of the superfamily Chironomidea, is also questionable. The Thaumaleidae and Simuliidae seem to be much more closely related than suggested by their morphology. On the other hand, the morphological relationship between the Ceratopogonidae and Chironomidae was not confirmed by our data. Indeed, the Chironomidae appear to be a sister group of all other Culicomorpha, however, the sequence data of an outgroup from within the Nematocera are needed to confirm this observation. The additional sequence data are also necessary to determine precisely the phylogenetic position

of the Ceratopogonidae and Dixidae as well as to revise the taxonomic status of the Corethrellidae.

## Acknowledgments

The authors thank Dr Art Borkent, for sending specimens of *Corethrella* and Dr Jean Wuest for comments on the manuscript. This research was supported by a grant from the Swiss National Foundation for Scientific Research (7PLPJ038617).

## References

- Bittar, G. & Carter, L. (1994) New probabilistic and numerical phenetics methods for inferring natural groupings and phylogenesis. *Proceedings of the International Meeting 'Ecology and statistical methods'*, Niort, France, pp. 145–150.
- Bittar, G. (1995) *Phylogénie: histoire, concepts, principes et présentation de nouvelles méthodes d'analyse numérique de l'évolution moléculaire*. PhD thesis, University of Geneva (Switzerland), 420 pp.
- Borkent, A. (1995) *Biting Midges in the Cretaceous Amber of North America (Diptera: Ceratopogonidae)*. Backhuys Publishers, Leiden, The Netherlands, 237 pp.
- Borkent, A. & McKeever, S. (1990) First instar larvae of *Corethrella appendiculata* Grabham have a prothoracic proleg (Diptera: Corethrellidae). *Entomologica Scandinavica*, **21**, 219–223.
- Campbell, B.C., Steffen-Campbell, J.D., Sorensen, J.T. & Gill, R.J. (1995) Paraphyly of Homoptera and Auchenorrhyncha inferred from 18S DNA nucleotide sequences. *Systematic Entomology*, **20**, 175–194.
- Carmean, D., Kimsey, L.S. & Berbee, M.L. (1992) 18S DNA sequences and the holometabolous insects. *Molecular Phylogenetics and Evolution*, **1**, 270–278.
- Cranston, P.S., Edward, D.H.D. & Colless, D.H. (1987) *Archaeochnus* Brundin: a midge out of time (Diptera: Chironomidae). *Systematic Entomology*, **12**, 313–334.
- Donoghue, M.J., Olmstead, R.G., Smith, J.F. & Palmer, J.D. (1992) Phylogenetic relationships of *Dipsacales* based on *rbcL* sequences. *Annals of Missouri Botanical Garden*, **79**, 333–345.
- Felsenstein, J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, **39**, 783–791.
- Felsenstein, J. (1993) PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.
- Grogan, W.L. Jr. & Szadziewski, R. (1988) A new biting midge from Upper Cretaceous (Cenomanian) amber of New Jersey (Diptera: Ceratopogonidae). *Journal of Paleontology*, **62**, 808–812.
- Hackman, W. & Väisänen, R. (1982) Different classification systems in the Diptera. *Annales Zoologici Fennici*, **19**, 209–219.
- Hancock, J.M., Tautz, D. & Dover, G.A. (1988) Evolution of the secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster*. *Molecular Biology and Evolution*, **5**, 393–414.
- Hendy, M.D. & Penny, D. (1982) Branch and bound algorithms to determine minimal evolutionary trees. *Mathematical Biosciences*, **59**, 277–290.
- Hennig, W. (1973) 'Ordnung Diptera (Zweiflüger)'. *Handbuch der Zoologie*, Berlin, **4** (2), 337 pp.
- Higgins, D.G., Bleasby, A.J. & Fuchs, R. (1992) CLUSTAL V:



- improved software for multiple sequence alignment. *CABIOS*, **8**, 189–191.
- Hillis, D.M. (1991) Discriminating between phylogenetic signal and random noise in DNA sequences. *Phylogenetic Analysis of DNA Sequences* (ed. by M. M. Miyamoto and J. Cracraft), pp. 278–294. Oxford University of Press, Oxford.
- Kalugina, N.S. (1977) The palaeontological data and some problems of the evolution of Culicoidea and Chironomoidea (Diptera). *Systematics and Evolution of Diptera (Insecta)* (ed. by A.V. Gorodkov), pp. 25–30, Zoologicheskii Institut ANSSSR, Leningrad [in Russian].
- Kalugina, N.S. (1991) Novye mesozoiskye Simuliidae i Leptoconopidae i proiskhozhdenye krovososanya y nizhikh dvukrylykh nasekomykh. *Paleontologicheskii Zhurnal*, 69–80 [in Russian].
- Kovalev, V.G. (1989) Geological history and the systematic position of the family Thaumaleidae (Diptera). *Entomologicheskoe Obozrenie*, **69**, 798–808 [in Russian].
- Kovalev, V.G. (1990) Dvukrylye. Muscida. [Dipterans. Muscida]. *Trudy Paleontologicheskogo Instituta*, **239**, 123–177.
- Krzeminska, E., Blagoderov V. & Krzeminski, W. (1993) Elliidae, a new fossil family of the infraorder Axymyiomorpha (Diptera). *Acta zoologica cracoviensis*, **35**, 581–591.
- Lane, R.P. & Crosskey, R.W. (1993) *Medical Insects and Arachnids*. Chapman & Hall, London, 723 pp.
- Larson, N., Olsen, G.J., Maidak, B.L., McCaughey, M.J., Overbeek, R., Macke, T.J., Marsh, T.L. & Woese, C.R. (1993) The ribosomal database project. *Nucleic Acids Research*, **21**, 3021–3023.
- Martin, A.P. & Pashley, S.R. (1992) Molecular systematic analysis of butterfly family and some subfamily relationships (Lepidoptera: Papilionidae). *Annals of the Entomological Society of America*, **85**, 127–135.
- Nadot, S., Bittar, G., Carter, L., Lacroix, L. & Lejeune, B. (1995) A phylogenetic analysis of monocotyledons based on the chloroplast gene *rps4*, using parsimony and a new numerical phenetics method. *Molecular Phylogeny and Evolution*, **4**, 257–282.
- Paskewitz, S.M., Ng, K., Coetzee, M. & Hunt, R.H. (1993) Evaluation of the polymerase chain reaction method for identifying members of the *Anopheles gambiae* (Diptera: Culicidae) complex in Southern Africa. *Journal of Medical Entomology*, **30**, 953–957.
- Pelandakis, M. & Sognac, M. (1993) Molecular phylogeny of *drosophila* based on ribosomal rna sequences. *Journal of Molecular Evolution*, **37**, 525–543.
- Raich, T.J., Archer, J.L., Robertson, M.A., Tabachnick, W.J. & Beaty, B.J. (1993) Polymerase chain reaction approaches to *Culicoides* (Diptera: Ceratopogonidae) identification. *Journal of Medical Entomology*, **30**, 228–232.
- Saether, O.A. (1977) Female genitalia in Chironomidae and other Nematocera: morphology, phylogenies, keys. *Bulletin of the Fisheries Research Board of Canada*, **197**, 1–210.
- Saitou, N. & Nei, M. (1987) The neighbor-joining method: a new method for constructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–425.
- Sokal, R.R. & Rohlf, F.J. (1981) *Biometry*. 2nd edn, W.H. Freeman, San Francisco.
- Steyskal, G.C. (1974) Recent advances in the primary classification of the Diptera. *Annals of the Entomological Society of America*, **67**, 513–517.
- Swofford, D.L. (1993) PAUP: Phylogenetic Analysis Using Parsimony, version 3.1.1. Distributed by the Illinois Natural History Survey, Champaign, IL.
- Tautz, D., Hancock, J.M., Webb, D.A., Tautz, C. & Dover, G. (1988) Complete sequences of the rRNA genes of *Drosophila melanogaster*. *Molecular Biology and Evolution*, **5**, 366–376.
- Walsh, P.S., Metzger, D.A. & Higuchi, R. (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*, **10**, 506–513.
- Weller, S.J., Friedlander, T.P., Martin, J.A. & Pashley, D.P. (1992) Phylogenetic studies of ribosomal RNA variation in higher moths and butterflies (Lepidoptera: Ditrysia). *Molecular Phylogeny and Evolution*, **1**, 312–337.
- Wheeler, W.C. (1989) The systematics of insect ribosomal DNA. *The Hierarchy of Life* (ed. by B. Fernholm, K. Bremer and H. Jörnvall), pp. 307–321. Elsevier, Amsterdam.
- Wheeler, W.C., Schuh, R.T. & Bang, R. (1993) Cladistic relationships among higher groups of Heteroptera: congruence between morphological and molecular data sets. *Entomologica Scandinavica*, **24**, 121–137.
- White, M.J.D. (1949) Cytological evidence on the phylogeny and classification of the Diptera. *Evolution*, **3**, 252–261.
- Wood, D.M. & Borkent, A. (1989) *Phylogeny and Classification of the Nematocera* (ed. by J. F. McAlpine and D. M. Wood), Vol. 3, pp. 1333–1370. Research Branch Agriculture Canada Monograph 32.

Accepted 11 April 1996

**Annex 1.** Alignment of partial LSU DNA sequences of 12 species of Culicomorpha and *Drosophila melanogaster*. Regions retained for phylogenetic analysis are underlined. The divergent domains D4, D5, D6, D7a, D7b and the limits of the processed fragment within the D7a domain in *Drosophila melanogaster* (Hancock *et al.*, 1988) are indicated.

	1	11	21	31	41	51	61	71	80	
Culicoid	CGAAGTTTCC	CTCAGGATAG	CTGGAGCATA	-GATATTTCT	-----GTAA	-TAATTCTTA	TCTGGTAAAG	CGAATGATTA		72
Dasyhele	CGAAGTTTCC	CTCAGGATAG	CTGGAGCATA	-GATATTTTT	-----GTAA	-TAATTCTTA	TCTGGTAAAG	CGAATGATTA		72
Aedes_sp	CGAAGTTTCC	CTCAGGATAG	CTGGAGCACG	CAACGTTTC-	-----GAAC	TCTATTCTTA	TCTGGTAAAG	CGAATGATTA		73
Aedes_al	CGAAGTTTCC	CTCAGGATAG	CTGGAGCACG	CAACGTTTC-	-----GAAT	CCTATTCTTA	TCTGGTAAAG	CGAATGATTA		73
Culex	CGAAGTTTCC	CTCAGGATAG	CTG-AGGACG	CAACATTTTC-	-----GGAT	TCTATTCTTA	TCTGGTAAAG	CGAATGATTA		72
Anophele	CGAAGTTTCC	CTCAGGATAG	CTGGAGCACG	TAACATTTTC-	-----GAAC	CTTATTCTTA	TCTGGTAAAG	CGAATGATTA		73
Chaoboru	CGAAGTTTCC	CTCAGGATAG	CTGGAGCATA	TAAATTTTT-	-----GAA-	TTTATTCTTG	TCTGGTAAAG	CGAATGATTA		72
Corethre	CGAAGTTTCC	CTCAGGATAG	CTGGAGCATA	AAAACATTT-	-----GAAC	ATTATTCTTA	TCTGGTAAAG	CGAATGATTA		73
Simuloid	CGAAGTTTCC	CTCAGGATAG	CTGGTGCATA	TACCGTTTCA	TAAA--CTTA	ATAAATCTTA	TCTGGTAAAG	CGAATGATTA		78
Andropro	CGAAGTTTCC	CTCAGGATAG	CTGGCACTAT	AGCCG-TTGA	ATT--CTTA	ATAAATCTTA	TCTGGTAAAG	CGAATGATTA		76
Paradixa	CGAAGTTTCC	CTCAGGATAG	CTGGAGCACT	TAAATATTGA	ACAATAATTA	ACAATCTTA	TCTGGTAAAG	CGAATGATTA		80
Camptoch	CGAAGTTTCC	CTCAGGATAG	CTGGAGCA-T	AAAATGGT-	-----TGCTTA	ACAACCTATA	CCTGGTAAAG	CGAATGATTA		73
Drosophi	CGAAGTTTCC	CTCAGGATAG	CTGGTGCAAT	TTAATATT--	---ATATAAA	ATAA-TCCTA	TCTGGTAAAG	CGAATGATTA		74
----- D4 -----										
	81	91	101	111	121	131	141	151	160	
Culicoid	GAGGCATTAG	GATCGAAACG	ATCTTAACT	ATTCTCAAAC	TATAAATGGG	TACG---TTG	AATGACATTC	TTGA-TTGAT		148
Dasyhele	GAGGCCTTAG	GATCGAAACG	ATCTTAACT	ATTCTCAAAC	TATAAATGGG	TACG---TTG	AATAACATTC	TTGA-TTGAT		148
Aedes_sp	GAGGCCTTAG	GTTTCGAAATG	ATCTTAACT	ATTCTCAAAC	TATAAATGGG	TACG---TAC	CATAACATTC	TTGC-ATGAT		149
Aedes_al	GAGGCCTTAG	GTTTCGAAATG	ATCTTAACT	ATTCTCAAAC	TATAAATGGG	TACG---TAC	CATAGCATTC	TTGC-ATGAT		149
Culex	GAGGCCTTAG	GTTTCGAAATG	ATCTTAACT	ATTCTCAAAC	TATAAATGGG	TACG---TAC	CATGACATTC	TTGC-ATGAT		148
Anophele	GAGGCCTTAG	GTTTCGAAATG	ATCTTAACT	ATTCTCAAAC	TATAAATGGG	TACGGGATTG	GGTAGCATGC	TTTG-ATGAT		152
Chaoboru	GAGGCCTTAG	GTTTCGAAATG	ATCTTAACT	ATTCTCAAAC	TATAAATGGG	TACG---TTA	CATAACATGC	TTTA-ATGAT		148
Corethre	GAGGCATTAG	GTTTCGAAATG	ATCTTAACT	ATTCTCAAAC	TATAAATGGG	TACG---TAA	CATAACATTC	TTGA-TTGAT		149
Simuloid	GAGGCATTAG	GTTTCGAAACG	TTCTTAACT	ATTCTCAAAC	TATAAATGGG	TACG---ATA	TATGAGTTTC	TTGA-TTGAA		154
Andropro	GAGGCATTAG	GTTTCGAAACG	TTCTTAACT	ATTCTCAAAC	TATAAATGGG	TACG---AAA	AATGATTCG	TTGA-TTGAA		152
Paradixa	GAGGCATTAG	GTTTCGAAACG	TTCTTAACT	ATTCTCAAAC	TATAAATGGG	TACG---ATA	AATTACATTC	TTGA-TTGAT		156
Camptoch	GAGGCCTTAG	GAACGAAATG	TTCTTAACT	ATTCTCAAAC	TTTAAATGGG	TATG---TAA	GTTAGCATAC	TTGA-ATGAT		149
Drosophi	GAGGCCTTAG	GGTCGAAACG	ATCTTAACT	ATTCTCAAAC	TTTAAATGGG	TAAG---AAC	CTTAACTTTC	TTGATATGAA		151
----- D5 -----										
	161	171	181	191	201	211	221	231	240	
Culicoid	GTCATTGTGA	CTTGATA---	-----	-----	-----CA	CTTGTGTATC	GGTTGTGATA	TTA-ATATGT		196
Dasyhele	GTTATTGTGA	CTTGTTAATG	-----	-----	-----CA	CTTGTGCATT	GGTTGTGATA	TTA-ATATGT		199
Aedes_sp	GTTATTGCAA	-CGTTG----	ACGCCGGGCC	GCCCCCTGTC	AGGGCGGTGC	CACCCAAGTG	GGGTAGAAGA	TAT-CTGTGT		223
Aedes_al	GCTGTTGCAA	ACGTTGGTAA	ACGCCGGGCC	GCCCCCTGTC	AGGGCG-TGC	CAC-----G	GGGTAGAAGA	TAT-CTGTGT		221
Culex	GTCATTGCAA	CCGTTTCG--	ACGTCGGGCC	GCCCCCTGCC	TGGGCCG-TNC	CGG---TCG	ACGTACAAGA	TAT-CTGTGT		220
Anophele	GCTACCTCA	AATCGATGAG	TCGAACGGTG	CGCCCCGTCG	CCCCCGGGT	GGTGGGATGC	ACCGGCTAGA	TAT-CGGTGT		231
Chaoboru	GTTATTGTGA	TAAATGAAAT	TTATTTTCAT	TTAG-----	-----	-----	---TAAAGA	TAT-ATATGT		198
Corethre	GTTATT--TA	CATATATATT	TTTAAATAAT	ATA-----	-----	-----	---TAAAGA	TATTAATGT		197
Simuloid	CTC-TTATTT	CACGCTTTTA	GCGTACCATA	TCAAACATTT	ATTACATGTT	TGTAATTT--	--GGTTTAGA	TTT--ATTGT		227
Andropro	CTCATGTTT	GACTAGT--A	CTTCATCATG	T-----	---CATGTG	ATGTT-----	--GTGAAAGA	TTT-GCTTGT		209
Paradixa	GTAATTTATT	TGCTGGCACT	TTATTTTATA	ATTGGTGC--	-----TG	ATG-----	--GTATTAGA	TTT-TTGTGT		216
Camptoch	GTT----GGC	CATTGGCCGT	TACGCGGCAC	-----	-----	-----	--GT--AAGA	TTG--TATGC		189
Drosophi	GATCAA----	-----	-----	-----	-----	-----	--GGTTATGA	TAT--AAATGT		174
----- D5 -----										
	241	251	261	271	281	291	301	311	320	
Culicoid	GCTCAGTGGG	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGT	GGGATGAACC	AAACGTAATG	TTAAGGCGCC	-AAAATAAAC		275
Dasyhele	GCTCAGTGG-	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGT	GGGATGAACC	AAACGTTATG	TTAAGGCGCC	-AAAATAAAC		277
Aedes_sp	GCTTAGTGGG	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGT	GG-ATGAACC	AAACGTAATG	TTACGGCGCC	T-AAATAAAC		301
Aedes_al	GCTTAGTGGG	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGT	GGGATGAACC	AAACGTAATG	TTACGGCGCC	T-AAATAAAC		300
Culex	TCTTAGTGGG	CCAAGTTTTG	GTAAGCAGAA	CTTGTGCTGT	GG-ATGAACC	AAACGTGATG	TTACGGCGCT	--AAATAAAC		297
Anophele	GCTTAGTGGG	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGT	GG-ATGAACC	AAACGCTATG	TTAAGGCGCC	--AAATAAAC		308
Chaoboru	GCTTAGTGGG	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGT	GGGATGAACC	AAACGCTATG	TTAAGGCGCC	C-AAGTAAAT		277
Corethre	GCTTAGTGGG	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGT	GGGATGAACC	AAACATAATG	TTAAGGCGCC	T-AAATAAAC		276
Simuloid	GCCCAGTGGG	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGT	GGGATGAACC	AATCGTGATG	ATAAGGCGCC	TAAAATAAAC		307
Andropro	GCTTAGTGGG	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGT	GGGATGAACC	AATCGTAATG	TTAAGGCGCC	C-AAATAA-T		287
Paradixa	GCCTAGTGGG	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGT	GGGATGAACC	AATCGCGATG	CTAAGGCCT	--AAATAAAC		294
Camptoch	GCCCAGTGG-	CCAAGTTTTG	GTAAGCAGAA	CTGGTGCTGC	GGGATGAGCC	GATCGCATTT	TTAAGGCGCC	C-AAATAAAC		267
Drosophi	-CCCAGTGGG	CCA-CTTTTG	GTAAGCAGAA	CTGGCGCTGT	GGGATGAACC	AAACGTAATG	TTACG-TGCC	C-AAATTAAC		250

## Annex 1. Continued

	321	331	341	351	361	371	381	391	400	
Culicoid	GACATATCAT	GGATACCATG	AAAGGTGTTG	ATTGCTGAAG	ACAGCAGGAC	GGTGGACATG	GAAGTTGTCA	TCCGCTAAGG	355	
Dasyhele	GACATATCAT	GGATACCATG	AAAGGTGTTG	ATTGCTGAAG	ACAGCAGGAC	GGTGGACATG	GAAGTTGTCA	TCCGCTAAGG	357	
Aedes_sp	GACGCATCAT	AGATACCATG	AAAGGTGTTG	ATTGCTACAG	ACAGCAGGAC	GGTGGACATG	GAAGTCGTCA	TCCGCTAAGG	381	
Aedes_al	GACGCATCAT	AGATACCATG	AAAGGTGTTG	ATTGCTACAG	ACAGCAGGAC	GGTGGACATG	GAAGTTGTCA	TCCGCTAAGG	380	
CuLex	G-CG-ATCAT	AGATACCATG	AAAGGTGTTG	ATTGCTACAG	ACAGCAGGAC	GGTGGACATG	GAAGTTGTCA	TCCGCTAAGG	375	
Anophele	GACGCATCAT	AGATACCATG	AAAGGTGTTG	ATTGCTAAAG	ACAGCAGGAC	GGTGGACATG	GAAGTCGTCA	TCCGCTAAGG	388	
Chaoboru	GACGTAACAC	AGATACCATG	AAAGATGTTG	ATTGCTAAAG	ACAGCAGGAC	GGTGGACATG	GAAGTTGTCA	TCCGCTAAGG	357	
Corethre	GACGCATCAC	AGATACCATG	AAAGGTGTTG	ATTGCTAAAG	ACAGCAGGAC	GGTGGACATG	GAAGTTGTCA	TCCGCTAAGG	356	
Simuloid	GACACACCTT	AGA--CCATA	AAAGATGTTG	ATTGCTGAAG	ACAGCAGGAC	GGTGGACATG	GAAGTTGTCA	TCCGCTAAGG	385	
Andropo	GACTCACCTT	AGACACCATG	AAAGATGTTG	ATTGCTGAAG	ACAGCAGGAC	GGTGGACATG	GAAGTTGTCA	TCCGCTAAGG	367	
Paradixa	GACGTATCAT	GGATACCATG	AAAGGTGTTG	ATTGCTGAAG	ACAGCAGGAC	GGTGGACATG	GAAGTTGTCA	TCCG-TAAGG	373	
Camptoch	GACGCATCAT	CGATACCATG	AAAGGTGTTG	ATTGCTGAAG	ACAGCAGGAC	GGTGGACATG	GAAGTTGTCA	TCCGCTAAGG	347	
Drosophi	AACTCATG-C	AGATACCATG	AAAGGCGTTG	GTTGCTTAA	ACAGCAGGAC	GGTGATCATG	GAAGTCGAAA	TCCGCTAAGG	329	
	401	411	421	431	441	451	461	471	480	
Culicoid	AGTGTGTAAC	AACTCACCTG	CCGAAGCAAT	TAGCCCTTAA	AATGGATGGC	GCTTAAGTCG	TTTGCCCTATA	CATTACCGCT	435	
Dasyhele	AGTGTGTAAC	AACTCACCTG	CCGAAGCAAT	TAGCCCTTAA	AATGGATGGC	GCTTAAGTCG	TTTGCCCTATA	CATAACCGCT	437	
Aedes_sp	AGTGTGTAAC	AACTCACCTA	CCGAAGCAAT	TAGCCCTTAA	AATGGATGGC	GCTTAAGTCG	TTTGCCCTATA	CATTACCGCT	461	
Aedes_al	AGTGTGTAAC	AACTCACCTG	CCGAAGCAAT	TAGCCCTTAA	AATGGATGGC	GCTTAAGTCG	TTTGCCCTATA	CATTACCG-T	459	
CuLex	AGTGTGTAAC	AACTCACCTG	CCGAAGCAAT	TAGCCCTTAA	AATGGATGGC	GCTTAAGTCG	TTTGCCCTATA	CATTACCGCT	455	
Anophele	AGTGTGTAAT	AACTCACCTG	CCGAAGCAAT	TAGCCCTGAA	AACGGATGGC	GCTTAAGTCG	TTTGCCCTATA	CATTGCCGCT	468	
Chaoboru	AGTGTGTAAC	AACTCACCTG	CCGAAGCAAT	TAGCTCTTAA	AATGGATGGC	GCTTAAGTCG	TTTGCCCTATA	CATGCCGCTT	437	
Corethre	AGTGTGTAAC	AACTCACCTG	CCGAAGCAAT	TAGCCCTTAA	AATGGATGGC	GCTTAAGTCG	TTTGCCCTATA	CATTATTGTC	436	
Simuloid	AGTGTGTAAC	AACTCACCTG	CCGAAGCAAT	TAGCTCTGAA	AATTGATGGC	GCTCTAGTCG	TTTGCCGATT	CATCACTGCT	465	
Andropo	AGTGTGTAAC	AACTTACCTG	CCGAAGCAAT	TAACTCTTAA	AATGGATGGC	GCTCAAGTCG	TTTGCCGATA	CATTACTGCT	447	
Paradixa	AGTGTGTAAC	AACTCACCTG	CCGAAGCAAT	TAGCCCTTAA	AATGGATGGC	GCTCAAGTCG	TTTGCCGATT	CATTGCCGCT	453	
Camptoch	AGTGTGTAAC	AACTCACCTG	CCGAAGCAAT	TAGCCCTTAA	AATGGATGGC	GCTCAAGTCG	TTTGCCGATA	CATGCCGCTC	427	
Drosophi	AGTGTGTAAC	AACTCACCTG	CCGAAGCAAC	TAGCCCTTAA	AATGGATGGC	GCTTAAGTTG	TATACCTATA	CATTACCGCT	409	
	481	491	501	511	521	531	541	551	560	
Culicoid	AAAA-TAAAA	TTTAATCTTT	ATCA-CTTGT	GGTAATGA--	-----	-TTTTGATAT	TTTAGTGAGT	AGGAGGGTAT	500	
Dasyhele	AAAA-TAAAA	TT-AATTTAT	AATATCTTGT	ATATTATAAA	-----	-TTTTGATAT	TTTAGTGAGT	AGGAGGGTAT	504	
Aedes_sp	GGCG-TACAA	GTNGTGACAG	GCGTTGCGAC	G-ACTGCA--	-----	-CTTTGAGAC	GCCAGCGAGT	AGGAGGGTCT	526	
Aedes_al	GGCG-CACAA	GTGGGGCA-C	GCCT-CGAGC	G--CGTCTGC	-----	-CCTTGAGGC	GCCA-CGAGT	AGGAGGGTCT	522	
CuLex	GGCG-TACAA	GTNGTGCGGC	GGGCCCTAGC	CTCCGCTCCG	A-----	-CTTTGAGAC	GCCAGCGAGT	AGGAGGGTCT	524	
Anophele	AGCGGTACAA	CGAGTAGCAA	GCCGGCGTGC	T---GTGC-T	A-----	-CTTCGAGAC	CCTAGTGAGT	AGGAGGGTAC	534	
Chaoboru	-GTTATAGAA	ATGATAATAT	AATTTATTTT	ATA-TTA---	-----	-TTTTGAAAT	AACAACGAGT	AGGAGGGTCT	501	
Corethre	-CATATAAAA	TT-AT-ATAT	ATAAAATTAA	AAATTTATAT	AT-----	-ATTTGAGAT	GTGGATAAGT	AGGAGGGTAT	504	
Simuloid	AAAG-CATAA	TT-AATACTC	ACTAAGTACT	GTGATGTA--	-----	-TTTTGAAAT	TTTAGTAAGT	AGGAGGGTCT	530	
Andropo	AATA-CATAA	TTC AATATGT	GCCAAATGTA	CACACAGCCA	AGTGCACAAT	ATTTTGAAGT	ATTAGTAAGT	AGGAGGGTCT	526	
Paradixa	ATTA-TAAAA	GT-AATATAT	-ACATTTTAT	GTATTATA--	-----	-TTTTAATAT	AATAGTGAGT	GGGAGGGTCT	517	
Camptoch	AAGA-TAAAA	T-CAATGTCG	TGCGCTTCAA	GTGCGTTCA-	-----	-TTTTGAGAT	CTTGACGAGT	AGGAAGGTAT	493	
Drosophi	AAAG-TAGAT	G-ATTTATAT	TACTTGTGAT	ATAAAA----	-----	-TTTTGAAAC	TTTAGTGAGT	AGGAAGGTAC	471	
	D6									
	561	571	581	591	601	611	621	631	640	
Culicoid	GATGATGTGC	GTTGAAGTAC	CTAGCGTGAG	TTGGTATGAA	GCCGTCATTA	GCACAGATCT	TG-TGG-TAG	TAGCAAA-TA	577	
Dasyhele	GATGATGTGC	GTTGAAGTAT	CTAGCGTAAN	-TGATATGAA	GCCGTCATTA	GCACAGATCT	TGGTGG-TAG	TAGCAAA-TA	581	
Aedes_sp	GGTGGTGTGC	GTTGAAGTGC	CTGGCGTAA-	-CGACATGGA	GCCGCCACTA	GCACAGATCT	TGGTGG-TAG	TAGCAAA-TA	602	
Aedes_al	GGTGGTGTGC	GTTGAAGTGC	CTGGCGTAA	CCGACATGGA	GCCGCCACTA	GCACAGATCT	TGGTGG-TAG	TAGCAAA-TA	600	
CuLex	GGTGGTGTGC	GTTGAAGTGC	CTGGCGTAA-	-CGACATGGA	GCCGCCACTA	GCACAGATCT	TGGTGG-TAG	TAGCAAA-TA	600	
Anophele	GGTGGTGTGC	GTTGAAGTGC	TTGGCGCAAG	-CAACATGGA	GCCGCCACCG	GCACAGATCT	TGGTGG-TAG	TAGCAAA-TA	610	
Chaoboru	GATGATGTGC	ATTGAAGTGT	TTAGCGTAA	CTGGCATGGA	GCCATCATTA	GCACAGATCT	TGGTGG-TAG	TAGCAAA-TA	579	
Corethre	TATGATA-GC	ATTGAAGTAT	TTAACGTAAG	TTGATATGAA	GCTATGATAA	GCACAGATCT	TGGTGG-TAG	TAGCAAA-TA	581	
Simuloid	GATGGTGTGC	GTTGAAGTGC	CAACCGTAA	GTGGTATGGA	GCCGCCATTA	GCACAGATCT	TGGTGG-TAG	TAGCAAAATA	609	
Andropo	AGTGGTGTGC	TTTGAAGTGT	TAACCGTAA	GTAGCATGGA	GCCGCCATTG	GCACAGATCT	TGGTGG-TAG	TAGCAAA-TA	604	
Paradixa	AATGGTGTGC	CTTGAAGTAT	CTAGCGTAA	-TGATATGGA	GCCACCATTG	GCACAGATCT	TGGTGG-TAG	TAGCAAA-TA	594	
Camptoch	TATGGTATGC	GTTGAAGTGC	CTAACGTGAG	TTGGTATGGA	GCTGCCATAA	GCACAGATCT	TGGTGG-TAG	TAGCAAA-TA	571	
Drosophi	AATGGTATGC	GTAGAAGTGT	TTGGCGTAA	CCTGCATGGA	GCTGCCATTG	GTACAGATCT	TGGTGGATAG	TAGCAAA-TA	550	

## Annex 1. Continued

	641	651	661	671	681	691	701	711	720
Culicoid	TTCGAATGAG	ATC-TTGGAT	GGCTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	TCAGTCTCTAA	656
Dasyhele	TTCGAATGAG	ATC-TTGGAT	GGCTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	TCAGTCTCTAA	660
Aedes_sp	TTCGAATGAG	ATC-TTGGAT	GACTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	CCAATCCTAA	681
Aedes_al	TTCGAATGAG	ATC-TTGGAT	GACTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	CCAATCCTAA	679
Culex	TTCGAATGAG	ATC-TTGGAT	GACTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	CCAATCCTAA	679
Anophele	TTCGAATGAG	ATC-TTGGAT	GACTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	CCAATCCTAA	689
Chaoboru	TTCGAATGAG	ATCCTTGGAT	GACTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	CCAGTCTCTAA	659
Corethre	TTCGAATGAG	ATC-TTGGAT	GACTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	CCAATCCTAA	660
Simuloid	TTCGAATGAG	ATC-TTGGAT	GACTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	CCAATCCTAA	688
Andropo	TTCGAATGAG	ATC-TTGGAT	GACTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	CCAATCCTAA	683
Paradixa	TTCGAATGAG	ACC-TTGGAT	GACTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGTGTGTTGAA	CACGAGTTAG	TCAGTCTCTAA	673
Camptoch	TTCGAATGAG	ATC-TTGGAT	GGCTGAAGTG	GAGGAGGTTT	TCGTGTCAAC	AGCAGTTGAA	CACGAGTTAG	TCAGTCTCTAA	650
Drosophi	ATCGAATGAG	AGCCTTGGAG	GACTGAAGTG	GAGGAGGTTT	TCGTGTGAA	AGTGGTTGAT	CACGAGTTAG	TCGGTCTCTAA	630
	721	731	741	751	761	771	781	791	800
Culicoid	GGTCTAAGGG	AAACCTGAAC	TAAT--GTGT	ATGATACATA	AATATTTTATT	GTGCAGAAAT	--TAGTATCA	ATTCC--TG	729
Dasyhele	GGTCTAAGGG	AAACCTGAAC	TAAT--GCGT	ATGATAATTA	AATTTTTTTTT	GATTAATAAAT	ACTAAAAATTA	GTATCATTTT	738
Aedes_sp	GCTCTATGGG	AAACCTGATA	TATATTAAGC	ATTTAACCAA	ATACGAACGC	CGTACGTC--	--GATTCGAT	GCAAC--ATC	755
Aedes_al	GCTCTATGGG	AAACCTGATA	TATATTAAGC	ATTTAACCAA	ATAC--AACAA	CGTATACCAG	C-GTGTGAT	GCAACCGGATA	757
Culex	GCTCTATGGG	AAACCTGATA	TATATTAAGC	ATTTAACCAA	ATTC--AACCA	CGCGTGT-G	CTGTGTTGAT	GCAAC--ATC	755
Anophele	GCTATATGGG	AAATCTGATT	CAAA---CGC	-----	----GATCCA	CCGTACGCGG	CTG----GAA	ACA-CACATC	747
Chaoboru	GCTCTATGGG	AAACCTGATT	TATTTTTAGT	GGTTTAATAA	TAATATTTTAC	TACGTAACAT	TCAGGTAATAA	CATAA-----	734
Corethre	GCTCTATGGG	AAATCTGATT	CAATATATAT	GCATATGTAT	GTATAAATAT	ATATATA---	-----	-----	717
Simuloid	GCTCTAAGGG	AAACCTGAAC	ATAT--GAGT	A-GATAC-AG	AAAT-TTAAT	TAT-TGATGT	---AATC--G	AAAATTAGTA	757
Andropo	GCTCTGAGGG	AAACCTGAAC	ATAT--GAGT	A-GATAATAA	TAATAATAAT	AATATGATT-	ATCCATC---	--AATTTGTA	754
Paradixa	GCTCTAAGGG	AAACCTGAAC	ATAT--GAGT	AAGATATTGA	T-TCCCTTAA	TATAAATAAC	--GTATCTAA	AAAATTGAT	748
Camptoch	GCCCTAGGGG	AAACCTGAAA	CGAT--GAGT	A-GATATAAA	GTTTTACCCT	ATCAAAAAATC	TAATCAAGT	TTGACAAATTG	727
Drosophi	GTTCAAGGCG	AAAGC-GAAA	ATTTTCAAGT	AAAACA----	-----	-----	-----	-----	668
	801	811	821	831	841	851	861	871	880
Culicoid	TATATAT---	-----	-----	--AAAGACCG	AAAGGGAATC	CGGTTACAAT	TCCGGAACCT	GTTAAGTATA	784
Dasyhele	CATATAC---	-----	-----	--AAAGACCG	AAAGGGAATC	CGGTTACAAT	TCCGGAACCT	GTTGAGTATA	793
Aedes_sp	CTAGTATATA	TT-----	-----	-GAATGAGCG	AAAGGGAATC	CGGTTACAAT	TCCGGAGCCT	GTTGGGTATA	816
Aedes_al	C-GATATATA	TT-----	-----	-GAATGAGCG	AAAGGGAATC	CGGTTACAAT	TCCGGAGCCT	GTTGGGTATA	817
Culex	GCCGTATATA	TT-----	-----	-AAATGAGCG	AAAGGGAATC	CGGTTACAAT	TCCGGAGCCT	GTTGGGTATA	816
Anophele	GCGCGC---	-----	-----	-GATCAAGCG	AAAGGGAATC	CGGTTACAAT	TCCGGAGCCT	GTTGAGTATA	802
Chaoboru	-----	-----	-----	-TATGGAGCG	AAAGGGAATC	CGGTTACAAT	TCCGGAGCCT	GTTGAGTATA	783
Corethre	-----	-----	-----	-TTATGAGCG	AAAGGGAATC	CGGTTACAAT	TCCGGAGCCT	GTTGAGTATA	766
Simuloid	TCACCCTCAT	T-----	-----	-TATGAGCG	AAAGGGAATC	CGGTTACAAT	TCCGGAACCT	GTTGAGTATA	816
Andropo	TCATCCTCAT	T-----	-----	-CAAAGAGCG	AAAGGGAATC	AGGTTACAAT	TCCTGAACCT	ATTGAGTATA	814
Paradixa	CACCCT-ATA	AA-----	-----	-TAATGAGCG	AAAGGGAATC	TGGTTACAAT	TCCAGAACCT	ATTGAGTATA	808
Camptoch	GCTTATGCGA	ACGGTGCATA	GTTGTCTCTA	GCAAAGGGCG	AAAGGGAATA	CGGTTCCAAT	TCCGGAGCCT	ATTGGGTATA	807
Drosophi	-----	-----	-----	TA ATTTTGAACG	AAAGGGAATA	CGGTTCCAAT	TCCGTAACCT	GTTGAGTATC	717
	881	891	901	911	921	931	941	951	960
Culicoid	CGTTTGCATA	TATCACCTTT	ATGGTCGAT-	-----	-----	-----TAT	GCAATCTCGG	TAACGAGAAA	836
Dasyhele	CGTTTGCATA	TGGTATCTTG	TA-----	-----	-----	-----TAT	G-AATCTCGG	TAACGGGAAA	837
Aedes_sp	CGTTTGCATT	GGCGT----	-CTCATACCG	G--AAACGGT	AGCGCC----	-----TTT	GTAATCATGG	CAACATGAAT	877
Aedes_al	CGTTTGCATT	GGCGT----	-CCCATACCG	G--CAACGGT	AGCGCC----	-----TTT	GTAATCATGG	CAACATGAAT	878
Culex	CGTTTGCATT	GGCGG----	--TAATACCG	G--CAACG-T	AGCGCC----	-----TTT	GTAATCATGG	CAACATGAAT	875
Anophele	CGTTTGTCT-	GGCGTGTTCG	GTTCCCTCCG	GGGGATCG-T	ACCGCC----	-----TT	GCGATCATGG	CAACATGAAT	868
Chaoboru	CGTTTGCATA	GTTGTTAAT	TATTTAATTA	ATTGCAATT-	-----	-----TAT	GTAATCATGG	TAACATGAAT	845
Corethre	CGTTTATAAT	GAAATTTTTT	TTTATGAAAT	TTTTT----	-----	-----ATT	ATAATCATGG	CAACGTGAAT	824
Simuloid	CGTTTGCATT	ATGTGACCG-	CTA----CAT	TT-ATGTAGG	TTAT-----	---CCATTAT	GTA-TCTCGG	CAACGAGAA	880
Andropo	CGTTTGCATT	ATGTGACCGT	CTCTCCTCAT	ATCATGTGTG	TGTGAGACTG	ATATCATAAT	GTA-TCTTGG	CAACAGAAG	893
Paradixa	CGTTTGTCT-	ATGTATTTTT	AAATATATAT	TA-----	-----	-ATATACTGT	TCAATCTCGG	TAACGAGAAA	868
Camptoch	CGTTTGTCTA	TTGTTGATTT	CGGTGCGCG-	-----	-----	-----ATT	AGCATCTCGG	TAACGAGAAG	859
Drosophi	CGTTTGTCTA	TAAATAT---	---GGGCT	C-----	-----	-----GTG	CTCATCTCGG	CAACAGGAAC	764

D7b

## Annex 1. Continued

	961	971	981	991	1001	1011	1021	1031	1040	
Culicoid										
Dasyhele	C--TTTCATT	GAGAAGCCAT	TAGGAGGTAT	CGGAAGAGTT	TTCTTTTCTG	TTTTACAGTT	GT-----	ACTCACCATG	906	
Aedes_sp	C--TTTCATG	AAGAAGC-AT	CAGGAGGTAT	CGGAAGAGTT	TTCTTTTCTG	TTTTACAGTT	GT-----	ACTCACCATG	906	
Aedes_al	CCTTTTCTTC	GAGAAGCCAA	CAAGAGATAT	CGGAAGAGTT	TTCTTTTCTG	TTTTACAGTC	AC-----	ACTAGCCATG	949	
Culex	CCTTTTCTTC	GAGAAGCCAA	CAAGAGATAT	CGGAAGAGTT	TTCTTTTCTG	TTTTACAGTC	AC-----	ACTAGCCATG	950	
Anophele	CCTTTTCTTC	GAGAAGCCAA	CGGGAGGTAC	TGGAAGAGTT	TTCTTTTCTG	TTTTACAGTC	AC-----	ACTGCCCATG	947	
Chaoboru	CCTTTTCTTT	GAGAAGCCAA	CGGGAGGTAC	TGGAAGAGTT	TTCTTTTCTG	TTTTACAGTC	ACC-----	ACTGACCATG	941	
Corethre	CCTTTTCTTT	GAGAAGTCAA	TAGGATGTAT	CGGAAGAGTT	TTCTTTTCTG	TTTTACAATC	AT-----	ACTAACCATG	917	
Simuloid	CCTTTTCTTT	GAGAAGCCAA	CAGGAGATAT	CGAAAAGAGTT	TTCTTTTCTG	TTTTACAATC	TTATAAATTA	ATTAACCATG	904	
Andropro	CC--AATATA	AAGAAGCTAT	CGGGATGTAA	CAAAAAGAGTT	TTCTTTTCTG	TTTTACAGTT	GT-----	ACGCACCATG	950	
Paradixa	GC--TAATATA	GAGATGCTAT	CGGGAGGTAA	TAAAAGAGTT	TTCTTTTCTG	TTTTACAAT	ATCACCCCC	ACATACCATG	972	
Camptoch	--TTTTTATA	GAGATGCCGA	GGAGATGTAC	CGGGAGAGTT	TTCTTTTCTG	TTTTACAGTT	GT-----	ACATACCATG	938	
Drosophi	C--TTCATA	GAGACGTCAA	TGGGAGATAT	CGAAAAGAGTT	TTCTTTTCTG	TTTTACAGTC	GT-----	ACGCACCATG	929	
	GA---CCATA	AAGAAGCCGT	CGAGAGATAT	CGGAAGAGTT	TTCTTTTCTG	TTTTATAGCC	GT-----	AC-TACCATG	832	
	1041	1051	1061	1071	1081	1091	1101	1111	1120	
Culicoid										
Dasyhele	GAAGTCTTTC	ATAGAGAGAT	ATGGTGTAAAT	AAATTGGTAG	AGCATGGCAT	TAAATTGCTG	TGTCG-ATAC	TCTCTTACTG	985	
Aedes_sp	GAAGTCTTTC	ATAGAGAGAT	ATGGTGTAAAT	AAACTGGTAG	AGCATGGCAT	TAAATTGCTG	TGTCG-ATAT	TCTCTTGCTG	985	
Aedes_al	GAAGTCTTTC	ATAGAGAGAT	ATGGTGGAC	AGGCTGGTAG	AGCATGGTAT	TAAATTGCTG	TGTCG-ATAT	TCTCTTCTTG	1028	
Culex	GAAGTCTTTC	ATAGAGAGAT	ATGGTGGAC	AGGCTGGTAG	AGCATGGTAT	TAAATTGCTG	TGTCG-ATAT	TCTCTTCTTG	1029	
Anophele	GAAGTCTTTC	ATAGAGAGAT	ATGGTGGAC	AAGCTGGTAG	AGCATGGTAT	TAAATTGCTG	TGTCG-ATAT	TCTCTTCTTG	1026	
Chaoboru	GAAGTCTTTC	ATAGAGAGAT	ATGGTGGAC	AGCTGGTAG	AGCATGGTAT	TAAATTGCTG	TGTCG-ATAT	TCTCTTCTTG	1020	
Corethre	GAAGTCTTTC	ATAGAGAGAT	ATGGTGGAC	GCACTGGTAG	AGCATGGTAT	TAAATTGCTG	TGTCG-GTAC	TCTCTTCTTG	1020	
Simuloid	GAAGTCTTTC	ATAGAGAGAT	ATGGTGGAC	AGATTGGTAG	AGCATGGCAT	TAAATTGCTG	TGTCG-ATAC	ACTCTT-TTG	995	
Andropro	GAAGTCTTTC	ATAGAGAGAT	ATGGTGGAC	AGATTGGTAG	AGCATGGCAT	TAAATTGCTG	TGTCG-ATAT	ACTCTTCTTG	983	
Paradixa	GAAGTCTTTC	ATAGAGAGAT	ATGGTGGAC	AGATTGGTAG	AGCATGGCAT	TAAATTGCTG	TGTCG-ATAT	ACTCTTCTTG	983	
Camptoch	GAAGTCTTTC	ATAGAGAGAT	ATGGTGGAC	AGGCTGGTAG	AGCATGGCAT	TCAGTTGCTG	TGTTG-TTAC	AT-CTCCTTA	1028	
Drosophi	GAAGTCTTTC	ATAGAGAGAT	ATGGTGGAC	AGATTGGTAG	AGCATGGCAT	TCAGTTGCTG	TGTTA-TTAC	TCTCTCCTTA	1051	
	GAAGTCTTTC	ATAGAGAGAT	ATGGTATAAT	AGACTGGTAG	AGCATGGCAT	TAAATTGCTA	TGTCGGGTAT	ATTCTTCTCG	1018	
	GAAGTCTTTC	ATAGAGAGAT	ATGGTATAAC	AGACTGGTAG	AGCATGGCAT	TAAATTGTTG	TGTCG-GTAT	TCTCTCATTG	1008	
	GAAGTCTTTC	GCAGAGAGAT	ATGGTA-GAT	GGGCTAGAAG	AGCATGACAT	-ATACTGTTG	TGTCG-ATAT	TTTCTCCTCG	909	
	1121	1131	1141	1151						
Culicoid										
Dasyhele	GACCTTGAAA	ATCAATGAAA	G---CTACGC	AAATT	1017					
Aedes_sp	GACCTTGAAA	-TT-ATGAAA	G-GGC-ACGN	AAATT	1016					
Aedes_al	GACCTTGAAA	ATCGAAGACT	G-GGGCACGC	AAACT	1062					
Culex	GACCTTGAAA	ATCGAAGACT	G-GGGCACGC	AAACT	1063					
Anophele	GACCTTGAAA	ATCGAAGACT	G-GGGCACGC	AAACT	1060					
Chaoboru	GACCTTGAAA	ATCGAAGACT	G-GGGCACGC	AAACT	1054					
Corethre	AACCTTGAAA	ATCAAAGACT	G-GGGCACGC	AAACT	1029					
Simuloid	GACCTTGAAA	ATCAAAGACT	G-GGGCACGC	AAACT	1017					
Andropro	GACCTTGAAA	ATT-ATATT-	G-GGTCACGC	AAACT	1060					
Paradixa	GACCTTGAAA	-TCTATATAA	G-GGT-ACGC	AAACT	1083					
Camptoch	GACCTTGAAA	ATCTATGAAA	GTGGT-ACGC	AAACT	1052					
Drosophi	GACCTTGAAA	-TCTATGAA-	--GGGCACGT	AAATT	1039					
	GACCTTGAAA	ATTTATGGTG	G-GGACACGC	AAACT	943					