The phylogeny of Anophelinae revisited: inferences about the origin and classification of *Anopheles* (Diptera: Culicidae)

Ralph E. Harbach & Ian J. Kitching

Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK

Corresponding author: Ralph E. Harbach, Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK. E-mail: r.harbach@nhm.ac.uk

Ralph E. Harbach, Ian J. Kitching, Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK. E-mails: r.harbach@nhm.ac.uk; i.kitching@nhm.ac.uk

Short title: Phylogeny of Anophelinae

Abstract

The evolution of anopheline mosquitoes (Culicidae: Anophelinae) has been the subject of speculation and study for decades, but a comprehensive phylogeny of these insects is far from complete. The results of phylogenetic studies based on morphological and molecular data sets are conspicuously ambiguous. Here we revisit the phylogenetic relationships of anopheline mosquitoes using state-of-the-art software and cladistic methods to analyse the data set of Harbach & Kitching (2005). We present a refined interpretation of relationships based on analyses of a revised data set that includes an additional species. Implied weighting analyses were conducted with TNT with the concavity constant *K* ranging from 1–33. We determined the optimal *K* value by summing the GC supports for each MPC and selected the tree with the highest support, *K* = 30, as the preferred cladogram. We then collapsed the branches with GC support < 1 to obtain the "best" topography of relationships. Genus *Chagasia* is the basalmost taxon of Anophelinae, and genus *Anopheles* is recovered as monophyletic but only if *Anopheles implexus* is excluded and genus *Bironella* is subordinated within it. The Afrotropical *An. implexus* is recovered as the sister to all other anophelines, and *Christya* Theobald, **stat. nov.**, is elevated from synonymy with *Anopheles* Meigen as a subgenus to accommodate it. The other anophelines comprise two large clades. The first includes the reciprocally monophyletic subgenera *Kerteszia* + *Nyssorhynchus*; the second consists of subgenus *Cellia* as the sister to a heterogeneous clade that includes genus *Bironella* and subgenera *Anopheles*, *Baimaia*, *Lophopodomyia* and *Stethomyia* of genus *Anopheles*. The sister relationship of *Cellia* and the heterogeneous clade is lost when the branches with GC < 1 are collapsed. The monophyly and non-monophyly of the informal subordinate taxa of subgenera *Nyssorhynchus*, *Cellia* and *Anopheles*, and also evolutionary scenarios, are discussed in relation to previous studies.

Introduction

Culicidae, mosquitoes, comprise a large and abundant group of ~~3,543~~ 3,546 formally recongized species with distributions in temperate and tropical regions of the world. The species are classified in two subfamilies, Anophelinae (~~482~~ 485 species) and Culicinae (3,061 species). Subfamily Anophelinae comprises three genera: *Anopheles* Meigen (cosmopolitan, ~~469~~ 472 species), *Bironella* Theobald (Australasian, 8 species) and *Chagasia* Cruz (Neotropical, 5 species) (http://mosquito-taxonomic-inventory.info/). Mosquitoes of these genera are known in the vernacular as "anophelines".

Anopheles has been the subject of more taxonomic research than any other genus of mosquitoes because it includes the species that transmit malarial and filarial parasites to humans. The majority of anopheline species belong to this genus, which comprises seven subgenera: *Anopheles s.s*. (cosmopolitan, ~~182~~ 185 species), *Baimaia* (Oriental, 1 species), *Cellia* Theobald (Old World, 224 species), *Kerteszia* Theobald (Neotropical, 12 species), *Lophopodomyia* Antunes (Neotropical, 6 species), *Nyssorhynchus* Blanchard (Neotropical, 39 species) and *Stethomyia* Theobald (Neotropical, 5 species) (http://mosquito-taxonomic-inventory.info/). The subgenera are based principally on the number and positions of specialized setae borne on the gonocoxites of the male genitalia (Christophers 1915; Reid 1968; Harbach & Kitching 2005). The three largest subgenera, *Anopheles*, *Cellia* and *Nyssorhynchus*, are divided into hierarchical systems of informal taxonomic categories (Reid & Knight 1961; Harbach 1994, 2004). Subgenus *Anopheles* is divided into two Sections based on the shape of the pupal trumpet (Reid & Knight 1961) and subgenus *Nyssorhynchus* is divided into three Sections based on characters of the larvae, pupae and adults (Faran 1980; Harbach 1994). Subgenus *Cellia* and the Sections of subgenera *Anopheles* and *Nyssorhynchus* are divided into Series, the larger Series are divided into species Groups, and some Groups are further divided into Subgroups and species Complexes. It is generally

assumed that each level of classification comprises a more or less natural assemblage of species based on morphological similarity. Genus *Bironella* includes three subgenera, but unlike the larger subgenera of *Anopheles*, they are not subdivided into informal group taxa. Genus *Chagasia* is a small homogenous group of species that is also not subdivided (Harbach & Howard 2009).

Sallum *et al*. (2000) conducted the first phylogenetic study of Anophelinae, based on morphological traits. The results indicated that genus *Anopheles* is paraphyletic because it included genus *Bironella*. Subgenera *Kerteszia*, *Nyssorhynchus*, *Cellia*, *Lophopodomyia* and *Stethomyia*, along with genus *Bironella*, were found to be monophyletic taxa dispersed among various Series and species Groups of subgenus *Anopheles*. The Christya Series of subgenus *Anopheles* was placed with *Kerteszia* + *Nyssorhynchus* and this clade was sister to *Cellia* + all other anophelines except *Chagasia*.

Sallum *et al*. (2002) assessed anopheline relationships based on ribosomal (18S, 28S) and mitochondrial (COI, COII) DNA sequences for half of the species included in the analyses of Sallum *et al*. (2000). Although the results of the two studies cannot be compared directly, analyses of the available molecular data corroborated the paraphyly of genus *Anopheles* relative to *Bironella*, the sister-group relationship of *Kerteszia* and *Nyssorhynchus*, and the monophyly of subgenera *Cellia*, *Lophopodomyia* and *Stethomyia* and genus *Bironella*, which was recovered as the sister of *Lophopodomyia* rather than *Stethomyia*.

Harbach & Kitching (2005) revised and expanded the data set of Sallum *et al*. (2000), including reinterpretation of certain homologies (especially the specialized setae of the male gonocoxites that diagnose the subgenera), revised coding of some characters, and addition of new data and two further taxa, to reassess the phylogeny of Anopheline. Parsimony analysis of the data set under implied weighting supported the monophyly of subgenera *Cellia*, *Kerteszia* and *Nyssorhynchus*, and the sister relationship of *Kerteszia* + *Nyssorhynchus*.

Subgenus *Anopheles* was recovered as a polyphyletic lineage basal to a clade consisting of *Cellia* + (*Kerteszia* + *Nyssorhynchus*). *Bironella*, *Lophopodomyia* and *Stethomyia* were nested within subgenus *Anopheles*, and subgenus *Baimaia* was recovered as the sister of *Bironella* + all other *Anopheles*. In contrast to the results of Sallum *et al*. (2000), *Bironella* and *Stethomyia* were recovered as monophyletic groups separate from subgenus *Anopheles*.

In summary, the phylogenetic studies conducted to date disclose the following principal conclusions about the phylogeny and classification of anopheline mosquitoes: (1) subfamily Anophelinae is a monophyletic lineage basal to all other Culicidae; (2) genus *Chagasia* is basal to the other anopheline taxa; (3) genus *Anopheles* is not demonstrably monophyletic with regard to genus *Bironella* and subgenera *Lophopodomyia* and *Stethomyia*; (4) subgenera *Kerteszia*, *Nyssorhynchus* and *Cellia* are each monophyletic (subgenus *Baimaia* is monobasic); (5) *Kerteszia* and *Nyssorhynchus* are sister taxa; and (6) the monophyly of the informal Sections and most Series of subgenera *Anopheles*, *Cellia* and *Nyssorhynchus* is doubtful (Harbach & Kitching 1998; Sallum *et al*. 2000, 2002; Krzywinski *et al*. 2001a, b; Harbach & Kitching 2005; Harbach 2007, 2013).

~~Harbach (2013) maintained that the preferred cladogram of Harbach & Kitching (2005: figs 2 and 3) is currently the best hypothesis of anopheline phylogeny because it is based on analyses of a greater number of taxa and morphological characters than all other published hypotheses. But is it really the best estimate of relationships? We thought so until~~ The monophyly of the Cyclopeppteron Series of subgenus *Anopheles* has been in doubt since it was redefined by Reid & Knight (1961) to include only *An. annulipalpis* Lynch Arribálzaga (known only from Argentina and Uruguay) and *An. grabhamii* Theobald (endemic to the West Indies). In 2014, we received a request from Gustavo Rossi (Consejo Nacional de Investigaciones Científicas y Técnicas, La Plata, Argentina) ~~questioned the inclusion of *An. annulipalpis* Lynch Arribálzaga (known only from Argentina and Uruguay),~~

~~along with *An. grabhamii* Theobald (endemic to the West Indies), in the Cyclolepteron Series (only two species) of subgenus *Anopheles* and asked us~~ to re-run the ~~2005~~ data set of Harbach & Kitching (2005) with the addition of character data for *An. annulipalpis* (*An. grabhamii* was already included). The morphology-based phylogenetic study of subgenus *Anopheles* conducted by Collucci & Sallum (2007) had indicated that the two species are not closely related, and Gustavo wanted to confirm this based on cladistic analysis of a larger data set. ~~We immediately realized that the data set should not be analysed using the same software and methods used in 2005.~~ In our previous study, we analysed the data using implied weights, implemented by PIWE version 3.0 (for Windows) (Goloboff 1997), with the default value of the concavity constant, $K = 3$. More recently, commencing with our collaborative study on the phylogeny of Aedini (Reinert *et al*. 2009), we conducted cladistic analyses using TNT version 1.1 (Willi Hennig Society Edition) (Goloboff *et al*. 2008), in which any value of *K* can be applied, and exploring the results of analyses using a much broader range of *K* values. Furthermore, ~~we now agree~~ subsequent to our 2005 paper, ~~with~~ Sereno (2007) ~~that~~ provided strong and logical arguments for separating the neomorphic and transformational elements of multistate characters ~~should be coded separately~~ (using "contingent coding"; Forey & Kitching 2000), arguments that we find compelling, and ~~.~~ ~~C~~consequently, recoding of the multistate characters employed by Harbach & Kitching (2005) became necessary. We also deemed it appropriate to apply more recently developed and stringent methods of assessing clade support. With this as background, we present here a refined interpretation of anopheline relationships based on analyses of a revised data set, with the inclusion of *An. annulipalpis*.

Materials and methods

Morphology

As indicated in the Introduction, we added *An. annulipalpis* to the data set of 66 anopheline species analysed by Harbach & Kitching (2005: Appendix). The full data set of 69 taxa thus comprised an ingroup of 67 species of Anophelinae and two outgroup species from subfamily Culicinae: *Aedeomyia* (*Aedeomyia*) *squamipennis* (Lynch Arribálzaga) and *Uranotaenia* (*Uranotaenia*) *lowii* Theobald. The data set of Harbach & Kitching (2005) included 46 compound multistate characters that required subdivision into one or more characters, resulting in an increase from 167 to 224 morphological characters (Appendix S1) coded from adults (55), male genitalia (35), fourth-instar larvae (95) and pupae (39). The coded data are provided in Appendix S2.

Phylogenetic analyses

Parsimony analyses were implemented with TNT version 1.1 (Willi Hennig Society Edition, August 2011) (Goloboff *et al*. 2008) using both equal weighting (EW) and implied weighting (IW) with values of the concavity constant, *K*, ranging from 1–33. The value of *K* indicates inversely the weighting "strength" applied, with low values weighting more strongly against homoplastic characters (measured as the number of extra steps required to fit the cladogram topology in question) and higher values weighting less strongly (Goloboff 1993). The individual character weights are summed to produce the overall "fit" and the most parsimonious cladogram (MPC) is that with the greatest fit. Heuristic searches were conducted using the new technology search options: sectorial searches, ratchet, tree drifting and tree fusing. For the ratchet, the up/downweighting probabilities were set to 5% and the number of replicates to 200. The number of cycles of tree drifting was set to 50. All other search parameters remained at their default settings. Analyses were terminated once the MPC had been found 100 times. The maximum number of trees held was set to 10,000. Multistate

characters were treated as unordered and cladograms were rooted between *Ad. squamipennis* and the remaining taxa.

The relative support for each node was assessed using symmetric resampling, as implemented in TNT, recording the frequency differences, i.e. "Groups present / Contradicted" or GC values (Goloboff *et al*. 2003). This metric does not suffer from the frequency distortions seen in other methods that resample to assess group support, such as the bootstrap and jackknife, and which are particularly pernicious when applied to weighted data (Goloboff *et al*. 2003). The GC values assess the difference between the absolute frequency with which a clade is found in the resampled matrices and that in the most frequent alternative topology in which the clade is not recovered. The GC values range from 100, where the clade is recovered in all resampled matrices, to -100, where an alternative arrangement is found in all resampled matrices (Goloboff *et al*. 2003). A zero value indicates that levels of support and contradiction are equal. Due to time constraints, we calculated GC values using the traditional search options, with 100,000 replicates and the default change probability, and searches constrained to use only those groups found in the MPC. We then summed the GC supports across all groups on each MPC and used this as the optimality criterion to select the best topology and its associated value of *K* (Goloboff *et al*. 2003; González-Santillán & Prendini 2015). Cladograms were prepared and morphological character mappings investigated using WinClada ver. 1.00.08 (Nixon 1999–2002).

Results

Each analysis produced only a single MPC, except for $K = 3$, 4 and 8, where eight, three and two MPCs, respectively, were found. In contrast to the study of González-Santillán & Prendini (2015: fig. 7), which found unimodal distributions of summed GC ($\sum$GC) values, each with a single clear maximum value, our results (Fig. 1) show three peaks of $\sum$GC at *K*

values of 3, 8 and 30–33, of which the highest is that for *K*= 30/32/33 (∑GC = 1837). However, as the ∑GC for *K* = 31 was slightly lower at 1835, we checked whether this decrease was simply a stochastic artefact. We replicated the GC calculations a further nine times for *K* = 30 and *K* = 31. The *K* = 30 tree yielded ∑GC values ranging from 1835–1837 (average = 1836.5) and the *K* = 31 tree yielded ∑GC values ranging from 1834–1838 (average = 1836.2). There is thus clearly a stochastic element to the ∑GC calculations due to the pseudorandom nature of the resampling procedure, and so we also repeated the procedure for the two MPCs found with *K* = 8, which had yielded the next highest ∑GC (1830). Tree 0 yielded ∑GC values ranging from 1819–1825 (average = 1822.1) and tree 1 yielded ∑GC values ranging from 1825–1830 (average = 1826.9). The highest value for *K* = 8 (1830) was well separated from the lowest ∑GC for *K* = 30/31 (1835), confirming our choice of the *K* = 30–33 topology as our preferred pattern of relationships among anophelines.

The MPC for *K* = 30–33 is shown in Figure 2. As demonstrated in previous studies (see Introduction), genus *Chagasia* is the basalmost taxon of Anophelinae. Genus *Anopheles* is recovered as monophyletic but only if genus *Bironella* is subordinated within it (see below). *Anopheles* (*Ano*.) *implexus* is placed as the sister to all other anopheline taxa, which fall into two large clades. The first comprises the reciprocally monophyletic subgenera *Kerteszia* and *Nyssorhynchus*. Within the second, subgenus *Cellia* is recovered as monophyletic and sister to the remaining taxa. The first to branch off within this latter clade is the representative of subgenus *Lophopodomyia*, followed by a grade comprising members of subgenus *Anopheles*. However, within the largest subclade of this latter group, subgenera *Stethomyia* and *Baimaia* and genus *Bironella* are arrayed in a terminal clade.

However, the GC supports also suggest that some groups are very poorly supported or unsupported. Collapsing those groups with GC < 1 on any one of the 10 replicates of the GC calculations gave the topology shown in Figure 3 (maximum GC values > 0 are shown, so the

∑GC of the figure is greater than 1837). The GC values of six of the remaining branches varied among the replicates but only by (a decrease of) 1 and never all at once in any one replicate. There is some loss of resolution within subgenera *Nyssorhynchus* and *Cellia*, but they each remain monophyletic. The sister-group relationship between *Cellia* and the clade comprised of the remaining taxa is also lost with the formation of a three-way polytomy. Within the third clade, the sister-group relationship between subgenus *Lophopodomyia* and the remaining taxa is lost within a basal eleven-way polytomy, as is also much of that within subgenus *Anopheles*. The subclade comprising *An*. (*Ano*.) *punctipennis* to *An*. (*Ano*.) *sinensis* has some loss of resolution but remains monophyletic. Most importantly, however, all the relationships within the clade comprising *An*. *aitkenii* + (*An*. *corethroides* + (*Stethomyia* + (*Baimaia* + *Bironella*))) survive.

Discussion

The monobasic Christya Series includes only the African species *An. implexus*. ~~The~~ Whereas this species was recovered as the sister to subgenera *Cellia* + (*Kerteszia* + *Nyssorhynchus*) in our previous study (Fig. S2), the results of the present analyses shown in Fig. 3 suggest that ~~*An. implexus*~~ it is sister to all Anophelinae except genus *Chagasia*, although the support for the latter group (GC = 2) is weak. Reid & Knight (1961) noted that *Anopheles implexus* ~~shows~~ shares a number of ~~affinities~~ features with *Chagasia*, including speckled legs, tarsi with unusually broad basal pale bands and a simple pupal trumpet. The basal position of the tarsal pale bands in particular is unusual for an anopheline and more like ~~certain~~ some species of subfamily Culicinae. ~~These characters indicate that *An. implexus* is a primitive member of genus *Anopheles*.~~ Reid & Knight ~~(1961)~~ also observed that *An. implexus* shows affinities with species of the Anopheles Series of subgenus *Anopheles* as well as groups of the Laticorn Section, but also suggested a relationship with subgenus *Nyssorhynchus*~~suggesting a~~

~~somewhat intermediate position within the subgenus~~. In contrast, *An. implexus* and two species of the Arribalzagia Series comprised a terminal clade within a polyphyletic Myzorhynchus Series in the phylogeny of subgenus *Anopheles* recovered in the morphology-based study of Collucci & Sallum (2007). Thus, the phylogenetic relationships of *An. implexus* are rather ambiguous. The results of the present study do not support a close affinity with subgenus *Anopheles*, and therefore not with species of the Anopheles and Myzorhynchus Series (Fig. 3). Features of ~~the parabasal setae of the male gonocoxite~~the male genitalia seem to indicate a closer relationship with other taxa. The gonocoxite of *An. implexus* bears a single parabasal seta that arises from an elongate prominence, whereas most species of subgenus *Anopheles* have multiple parabasal setae borne on a small protuberance. ~~Exceptions include~~There are exceptions, however. ~~*An.*~~*Anopheles algeriensis* and *An. corethroides*, ~~the latter~~ which is recovered as sister to a terminal clade comprised of *Stethomyia* + (*Baimaia* + *Bironella*) (Fig. 3), ~~—~~both have a single parabasal seta and that of the former species is borne on an elongate prominence. Another distinctive feature of *An. implexus* is the presence of posterolateral tufts of piliform scales on abdominal tergum IV, a feature that is shared with *An. oiketorakras* and *An. nimbus* of the Neotropical subgenera *Lophopodomyia* and *Stethomyia*, respectively, both of which appear to be more closely related to subgenus *Anopheles* than to *An. implexus* (Figs 2, 3). The scale-tufts are usually absent in species of subgenus *Anopheles*, and are spatulate rather than piliform when present. Larval seta 1-M of *An. implexus* is unique in having a long and broad rachis. The rachis is short or long and narrow in the other anopheline species included in the analyses. Somewhat similar to the results reported here, Sallum *et al*. (2000) found that a clade comprised of *An. implexus* + (*Kerteszia* + *Nyssorhynchus*) was the earliest lineage of genus *Anopheles*. The relationship, however, was very weakly supported (bootstrap < 50%; Bremer support 2). ~~Reid & Knight (1961) suggested a relationship between *An. implexus* and subgenus *Nyssorhynchus*~~

based on the mutual occurrence of extensive pale scaling on the hindtarsi of adults and the unserrated leaflets of the abdominal palmate setae of larvae. In addition to the characters mentioned above, Wilkerson & Peyton (1990) found that the wing of *An. implexus* is distinctive in having both apical and preapical pale spots, and noted that the presence of speckled legs, prominent abdominal scale tufts and pale hindtarsomeres suggested that *An. implexus* is "a possible ancestor of the New World species of the subgenera *Nyssorhynchus*, *Kerteszia*, *Lophopodomyia*, and the Arribalzagia Series".

Subgenera Kerteszia *and* Nyssorhynchus

Phylogenetic studies of anopheline mosquitoes based on both morphological and molecular data support the reciprocal monophyly and sister relationship of subgenera *Kerteszia* and *Nyssorhynchus* (Sallum *et al*. 2000, 2002; Collucci & Sallum 2003; Harbach & Kitching 2005). The sister relationship recovered in the collapsed tree (Fig. 3) is supported by six characters (Fig. S1), including three that are unique and not contradicted (63:1, accessory setae of gonocoxite present; 68:1, parabasal setae inserted on margin of gonocoxite; 77:1, specialized apical seta of ventral claspette present, spiniform and foliform). Relationships of the informal group taxa represented by the species that comprise the *Nyssorhynchus* clade are delineated in Figure 4. The monophyly of the subgenus is strongly supported by 12 characters (Fig. S1), two of which are unique and not contradicted (57:1, ventromesal connection of gonocoxite developed as a truncate process; 204:0, pupal seta 10-VI absent). However, the Myzorhynchella and Argyritarsis Sections are not monophyletic; the Myzorhynchella Section is paraphyletic relative to species of the Argyritarsis and Albimanus Sections. The Argyritarsis Series is polyphyletic and species of the Albimanus and Oswaldoi Series of the Albimanus Section fall in an unresolved clade together with a monophyletic Albitarsis Series of the Argyritarsis Section. The monophyly of the Albitarsis Series, however, is only weakly

supported (GC = 8 or 9). Despite the more stringent assessment of clade support, it is worth noting that the same pattern of relationships was recovered as in our previous study (Fig. S3).

Bourke *et al*. (2010) performed a phylogenetic study of relationships among 21 species of subgenus *Nyssorhynchus* based on sequences for the mitochondrial ND6 and nuclear *white* genes. When the two genes were combined in a single analysis, the Myzorhynchella Section was recovered as a monophyletic group in a basal relationship to the Albimanus and Argyritarsis Sections, neither of which, in agreement with the results of the present study, were recovered as a monophyletic group. It is interesting to note that species of the Myzorhynchella Section were recovered as two separate clades in a polytomy with a third clade comprised of the other sections when the ND6 sequence data were analysed alone. The preponderance of evidence, gleaned from analyses of morphological and molecular data (Sallum *et al*. 2000, 2002; Harbach & Kitching 2005; Bourke *et al*. 2010; present study), convincingly indicates that the current internal classification of subgenus *Nyssorhynchus* (Harbach 2013; http://mosquito-taxonomic-inventory.info/subgenus-ltemgtnyssorhynchusltemgt) is not based on natural relationships.

Subgenus Cellia

Phylogenetic analyses of morphological and molecular data ~~have shown~~indicate that subgenus *Cellia* is a monophyletic group (Krzywinski *et al*. 2001a, b; Sallum *et al*. 2000, 2002; Harbach & Kitching 2005; Mohanty *et al*. 2009; Wang *et al*. 2014; Norris & Norris 2015). The *Cellia* clade shown in the MPC (Fig. 2) is supported by a combination of four characters, none of which are unique (Fig. S1). However, the clade is fairly well supported as indicated by a GC value of 47 (Fig. 3). It is interesting to note that in the analyses of Wang *et al*. (2014), *Cellia* was recovered as a monophyletic group based on D2 rDNA sequences of 28 species of *Cellia* and eight species of subgenus *Anopheles*, but was polyphyletic in a D3

rDNA phylogeny that included sequences for 49 species of *Cellia* and the same eight species of subgenus *Anopheles* due to the exclusion of *An*. (*Cellia*) *maculatus* from the clade comprising subgenera *Anopheles* plus the remaining *Cellia*.

~~Mohanty *et al*. (2009) stated that they conducted neighbour-joining (NJ) and maximum parsimony analyses of sequences for the COI and COII loci of mtDNA and the D3 and ITS regions of rDNA obtained from different numbers of *Cellia* species (18, 21, 26 and 26 genetic species respectively), but only illustrated neighbour-joining trees, noting for each locus that the "neighbor-joining and maximum-parsimony methods produced equivalent topologies". It is doubtful that "equivalent" indicates that the trees had identical topologies. As neighbour-joining is a phenetic method and the grouping of species, unlike the clades constructed from shared-derived characters in phylogenetic methods, is based on overall similarity; hence, the relationships portrayed in the NJ trees of Mohanty *et al*. are not "evolutionary relationships". Consequently, their results are not relevant to the present discussion.~~

The internal classification of *Cellia* stems from the framework of Edwards (1932) and the revised and updated schemes of Grjebine (1966), Reid (1968), Gillies & de Meillon (1968) and Harbach (1994). The subgenus is divided into six principal informal groups, the Cellia, Myzomyia, Neocellia, Neomyzomyia, Paramyzomyia and Pyretophorus Series, each of which has some or all of the included species classified in one or more species Groups. The subgenus includes 224 formally named species~~. The Cellia Series includes eight species, two of which form the subordinate Squamosus Group; the Myzomyia Series comprises 65 species, 17 unplaced and 48 divided between four species Groups; the Neocellia Series has 32 species, 15 unplaced and 17 classified in three species Groups; the Neomyzomyia Series includes 92 species, 18 unplaced and 74 divided between 10 species Groups; the Paramyzomyia Series has only six species split between two species Groups; and the~~

~~Pyretophorus Series encompasses 21 species, 10 unplaced and 11 divided between three species Complexes~~ (~~Harbach, 2013;~~ http://mosquito-taxonomic-inventory.info/node/11370). The present analysis only included 21 species (9%) of the subgenus,~~:~~ ~~two from the Cellia Series (one unplaced and one representing the Squamosus Group); six from the four species Groups of the Myzomyia Series; two from~~representing all the six series and all but two of the species groups of the Neocellia Series. ~~(one unplaced and one representing one of the three species Groups); seven representing six of the 10 species Groups of the Neomyzomyia Series; two from the Paramyzomyia Series, representatives of the two species Groups; and two from the Pyretophorus Series, an unplaced species and a member of the Gambiae Complex.~~ Despite the severely limited taxon representation of this and previous morphological (Sallum *et al*. 2000; Harbach & Kitching 2005) and molecular studies (Sallum *et al*. 2002; Mohanty *et al*. 2009; Norris & Norris 2015), it is obvious that a significant part of the current internal classification of *Cellia* does not reflect evolutionary relationships. In comparison with our previous study (Fig. S3), the relationships within *Cellia* recovered in the present study (Fig. 5) are more poorly resolved, but this, in part, is due to our now more stringent application of branch support. Whereas the Cellia, Myzomyia, Paramyzomyia and Pyretophorus Series were recovered previously as monophyletic groups (Fig. S3), ~~The relationships of the informal group taxa delineated in Fig. 5 reveal that the Myzomyia, Neocellia, Neomyzomyia and Paramyzomyia Series are not monophyletic. The~~ only the Cellia Series now appears to be monophyletic (strongly supported, GC = 95)~~,~~ ~~but~~ However, the monophyly of the Pyretophorus Series is questionable due to the placement of *An. vagus* and *An. gambiae* as separate terminals within the eleven-way polytomy. These results mirror the results of Sallum *et al*. (2000), whose study included 61 fewer morphological characters for 64 of the 67 species included herein. As shown in ~~many~~ other studies (e.g. Sallum *et al*. 2007; Wang *et al*. 2014), the Funestus Group (Myzomyia Series) and the Leucosphyrus

Group (Neomyzomyia Series) are monophyletic assemblages. It is noteworthy, however, that *An. theileri* Edwards of the Wellcomei Group of the Myzomyia Series was recovered within the Funestus Group in the maximum parsimony analyses of COI mtDNA and ITS2 rDNA sequence data conducted by Norris & Norris (2015), indicating that the group may not be monophyletic.

Sallum *et al*. (2002) conducted a molecular analysis of anopheline relationships based on ribosomal (18S, 28S) and mitochondrial (COI, COII) DNA sequences. Contrary to the findings reported here and by Sallum *et al*. (2000), the Myzomyia, Neocellia, Neomyzomyia and Pyretophorus Series were recovered as monophyletic groups based on analyses of the ribosomal and combined ribosomal and mitochondrial sequence data. ~~Those results, however, cannot be construed to confirm the monophyly of the four series because significantly fewer taxa were included in the analyses.~~ Whereas the present study and that of Sallum *et al*. (2000) included 21 species of subgenus *Cellia*, the molecular study of Sallum *et al*. (2002) only included sequence data for nine species. In addition to five of the species included here, *An. dirus*, *An. farauti* (Neomyzomyia Series), *An gambiae* (Pyretophorus Series), *An. funestus* and *An. minimus* (Myzomyia Series), the molecular study included *An. stephensi* Liston (Neocellia Series, unplaced; substitution for *An. superpictus*), and three species, in addition to *An. gambiae*, of the Pyretophorus Series, *An. arabiensis* Patton (Gambiae Complex), *An. sundaicus* (Rodenwaldt) (Sundaicus Complex) and *An. subpictus* Grassi (Subpictus Complex). Despite the inclusion of three additional members of the Pyretophorus Series, in the absence of molecular data for *An. vagus* (Fig. 5), and probably also *An. multicolor* (sister to *An. vagus* in Fig. 2), the monophyly of the series must remain in doubt. Likewise, the Myzomyia and Neomyzomyia Series cannot be inferred to be monophyletic in the absence of data for those species included in the present study that cause these groups to be polyphyletic. Furthermore, the Neocellia Series cannot be regarded as being monophyletic based on the

inclusion of only one species of the group (i.e. *An. stephensi*). It is interesting to note, however, that Bayesian analysis of D2 rDNA sequence of 28 species of *Cellia* (*An. vagus* absent) support the finding of Sallum *et al*. (2002), whereas their parsimony analysis of D3 rDNA sequence for 49 species of the subgenus (*An. vagus* included) only recovered the Neomyzomyia Series (four species) as monophyletic.

~~One is tempted to assume that t~~The molecular phylogeny of Neafsey *et al*. (2015) based on protein sequences of 1,085 single-copy orthologs for seven species of the Pyretophorus Series – five species of the Gambiae Complex, *An. christyi* (unplaced) and *An. epiroticus* (Sundaicus Complex) – and the phylogeny of Anthony *et al*. (1999) based on morphological data for 10 species of the group, convincingly support~~s~~ the monophyly of the series~~,~~. ~~but in fact it is illogical to infer the monophyly of the entire series based on limited taxon sampling.~~ The phylogeny of Neafsey *et al*. also included lineages that correspond to the Myzomyia Series (three species), Neocellia Series (two species) and Neomyzomyia Series (two species), but the probability that these clades accurately reflect a common ancestry for all species of each group is doubtful in view of the relationships recovered for significantly more species included in the morphological and other molecular studies conducted to date.

~~It is interesting to compare the relationships of the seven species of the Pyretophorus Series included in the study of Neafsey *et al*. (2015) with the relationships recovered among 10 species of the group in the morphological phylogenetic study of Anthony *et al*. (1999). The Oriental *An. epiroticus* was recovered as the basal taxon in the former study whereas the two Afrotropical species included in the latter study (*An. christyi* and *An. gambiae*) were basal and paraphyletic relative to the Oriental species, with *An. christyi* in the most basal position. *Anopheles christyi* was also recovered basal to the Afrotropical species (Gambiae Complex) in the phylogeny of Neafsey *et al*. (2015).~~

Subgenus Anopheles

The clade comprising subgenus *Anopheles* in the MPC (Fig. 2) is supported by a combination of six characters (Fig. S1), three of which are unique and not contradicted (66:1, gonocoxite with multiple parabasal setae; 67:1, parabasal seta(e) of gonocoxite differentiated; and 69:1, parabasal setae borne on a small protuberance or swelling). However, despite this, the clade receives zero GC support and so does not appear in the collapsed tree (Fig. 3). The inclusion of *Bironella* within genus *Anopheles* contradicts the results of other analyses based on molecular and morphological data (Besansky & Fahey 1997; Foley *et al*. 1998; Harbach & Kitching 1998; Krzywinski *et al*. 2001b), but is consistent with the analyses of Sallum *et al*. (2002) based on ribosomal (18S, 28S) and mitochondrial (COI, COII) DNA sequences that support the paraphyly of genus *Anopheles* relative to *Bironella*.

Relationships of the informal group taxa of the clade are delineated in Figure 6. As is the case with subgenus *Cellia*, the relationships within subgenus *Anopheles* are much less resolved than in our previous study (Fig. S2). The Angusticorn and Laticorn Sections, the Anopheles and Cycloleppteron Series of the former and the Myzorhynchus Series of the latter are not monophyletic. The Anopheles Series is not monophyletic because it excludes *An. punctipennis* (as previously), and two species Groups (Aitkenii and Stigmaticus) fall within a paraphyletic relationship to a clade comprised of *Stethomyia* + (*Baimaia* + *Bironella*), the three of which were basal and paraphyletic to all other *Anopheles* in the previous study (Fig. S2). The two species of the Cycloleppteron Series are unrelated, with *An. annulipalpis* sister to a clade comprised of the Coustani + Hyrcanus Groups. Only the Lophoscelomyia Series of the Angusticorn Section and the Arribalzagia Series of the Laticorn Section are recovered as monophyletic groups.

When Reid & Knight (1961) proposed the classification of subgenus *Anopheles*, they listed the informal group taxa in order from the Laticorn Section, with the Christya,

Arribalzagia and Myzorhynchus Series, and the Angusticorn Section, with the Cycloleppteron, Lophoscelomyia and Anopheles Series. This order was thought to reflect relationships "reasonably well", with the Christya Series being the "more primitive" and Anopheles Series the "more advanced". As noted above, the results of the present study support the "more primitive" position of the Christya Series, but polyphyletic arrangement of the Anopheles Series (Fig. 6) fairly convincingly shows that it does not include the "more advanced" species of the subgenus. The Anopheles Series was also found to be polyphyletic in previous studies, with its members interspersed in a complexity of inter-group relationships (Sallum *et al*. 2000) and with *Bironella*, *Stethomyia* and *Lophopodomyia* interspersed within it (Harbach & Kitching 2005). The present results differ in that part of the Anopheles Series is recovered in a sister relationship to a terminal clade comprised of *Stethomyia* + (*Baimaia* + *Bironella*) (Fig. 6). It is interesting to note that collapsing those groups in Fig. 2 with a GC value of less than 1 does not affect the relationships within the clade comprising *An. aitkenii* + (*An. corethroides* + (*Stethomyia* + (*Baimaia* + *Bironella*))) (Fig. 3). Sallum *et al*. (2000) recovered *Bironella* and *Stethomyia* as monophyletic sister groups in the absence of *Baimaia*, and concluded that they should be classified as informal groups within genus *Anopheles* rather than generic-level taxa. In their subsequent molecular study (Sallum *et al*. 2002), *Bironella* was placed as sister to *Lophopodomyia* in a clade that also included *Nyssorhynchus* and *Kerteszia*, and suggested that *Bironella* might be treated as a subgenus of *Anopheles*. Although originally introduced as a genus (Theobald 1905), *Bironella* was treated as a subgenus of *Anopheles* by Christophers (1924) and some other authors between 1924 and 1938 (Marks *et al*. 1963). Belkin (1962) considered *Bironella* to be an ancient group that shares a number of features with subgenus *Anopheles*, notably the absence of cibarial armature in adult females and the position and development of seta 1-A on the antennae of larvae. The placement of *Bi. gracilis* basal to *An. atroparvus* (subgenus

Anopheles) + *An. gambiae* (subgenus *Cellia*) in the Bayesian likelihood tree of Reidenbach *et al*. (2009) derived from a data set consisting of six nuclear genes seems to support Belkin's contention; however, the monophyly of genus *Anopheles* was not supported in all of their analyses – the relationship of *Bi. gracilis* + *An. atroparvus* was also recovered. Considering the discordant relationships observed in the phylogenetic studies conducted thus far, it is not possible to determine objectively whether *Bironella* should retain generic status or be classified as a subgenus or species group within *Anopheles*. If the relationships indicated in Fig. 3 prove to be correct and *Bironella* is deemed to warrant generic status, then firmly established monophyletic lineages within the current concept of subgenus *Anopheles* may also require generic status. Two such groups are the Arribalzagia and Lophoscelomyia Series, both of which were originally established as genera (*Arribalzagia* Theobald, 1903a and *Lophoscelomyia* Theobald, 1904). Species of the Arribalzagia Series are the only anophelines that have the subcostal vein of the wing ending in an isolated dark spot distal to the sector dark spot (Wilkerson & Peyton 1990), which in addition to the six homoplastic characters shown in Fig. S1, confirms the monophyly of the series (Fig. 3: GC = 53). The monophyly of the Lophoscelomyia Series is strongly supported (Fig. 3: GC = 94) by a combination of five homoplastic characters (Fig. S1). Sallum *et al*. (2000), ~~and~~ Harbach & Kitching (2005) and Collucci & Sallum (2007) also found strong support for the monophyly of this group. Adults of the Lophoscelomyia Series are distinctive in having white scales on the coxae, dark-scaled legs, and dorsal and ventral scales on the last abdominal segment. Males lack leaflets of the aedeagus.

Collucci & Sallum (2007) erroneously stated that in contrast "to the hypothesis of Harbach and Kitching (2005), the monophyly of the Cycloleppteron Series" was not supported by the results of their study. However, the Cycloleppteron Series was only represented by *An. grabhamii* in the study of Harbach & Kitching, and monophyly cannot be

demonstrated from only a single included taxon. As in the present study, Collucci & Sallum found that *An. annulipalpis* and *An. grabhamii* are unrelated and the Cycloleppteron Series is a monobasic category that only includes the latter species. Whereas the two species were recovered in a paraphyletic relationship basal to a clade comprised of species of the Arribalzagia, Christya and Myzorhynchus Series in the implied weighting analysis of Collucci & Sallum, the results of the present study suggest that the two species are more distantly related, with *An. annulipalpis* sister to a terminal clade comprised of the Coustani and Hyrcanus Groups (Fig. 6). The association of *An. annulipalpis* with those groups of species is fairly well supported (GC = 38) by three homoplastic characters (Fig. S1). The Coustani-Hyrcanus lineage comprises a homogeneous assemblage of species that are distinguished by the unique presence of lateral patches of scales on the clypeus (Fig. S1, ~~ch.~~character 6:1). Clypeal scales are absence in *An. annulipalpis*. In the absence of further information, it is not possible to assign *An. annulipalpis* to an existing taxonomic group, so for the time being it is retained in the Angusticorn Section as an unplaced species.

The terminal clade consisting of *An. aitkenii* + (*An. corethroides* + (*Stethomyia* + (*Baimaia* + *Bironella*))) sits in stark contrast to the relationships of these taxa at the base of genus *Anopheles* in the phylogeny of Harbach & Kitching (2005), expressed parenthetically as *Baimaia* + (*Bironella* + (*Stethomyia* + (*An. corethroides* + (all other *Anopheles*)))), with *An. algeriensis* basal to three other species (*An. aitkenii*, *An. judithae* and *An. sintonoides*) in a clade that is sister to the remaining *Anopheles* species. The placement of these taxa at opposite ends of the *Anopheles* topologies is perhaps not as important as their apparent phyletic associations. The Oriental *An. aitkenii*, unlike the Palaearctic *An. algeriensis*, which resembles species of the Australian Stigmaticus Group, represented in the analyses by *An. corethroides*, in having a single parabasal seta, has multiple parabasal setae; this is perhaps why its placement as the sister to *An. corethroides* + (*Stethomyia* + (*Baimaia* + *Bironella*)) is

only weakly supported by a GC value of 5 (Fig. 3). The larger sister group is supported by three characters (Fig. S1), one of which is unique and not contradicted (140:1, ventral ramus (sclerotized) reaching the ventral edge of the labiohypopharynx) and a GC support of 23 or 24. The monophyly of subgenus *Stethomyia* and that of genus *Bironella* are very strongly supported (GCs = 100 and 97 respectively) whereas the sister relationship of *Baimaia* and *Bironella* is only weakly supported by two homoplastic characters and a GC support of 25 (Fig. 3). If *Bironella*, *Baimaia* and *Stethomyia* are to retain their generic/subgeneric status, then it would seem appropriate to afford the same rank to the Stigmaticus Group.

The results of the present analyses do not agree with the conclusions of Krzywinski *et al*. (2001a, b), Sallum *et al*. (2002), Collucci & Sallum (2003) and Neafsey *et al*. (2015) that subgenus *Anopheles* is monophyletic. Sallum *et al*. (2000), who concluded likewise, included 64 species of *Anopheles* in their analyses of morphological data, whereas Sallum *et al*. (2002) only included 32 species in their molecular analyses. The molecular phylogenetic studies of Krzywinski *et al*. were based on only five species representing the Anopheles (Maculipennis and Pseudopunctipennis Groups), Arribalzagia and Myzorhynchus (Coustani Group) Series, and those of Neafsey *et al*. included only two species of the subgenus, representing the Maculipennis and Hyrcanus Groups. We are confident that the inclusion of more species of all currently recognized informal groups of the subgenus will confirm the polyphyly of the subgenus. The inclusion of representatives of other taxonomic groups and additional DNA markers in molecular analyses will show, in agreement with morphological data, that subgenus *Anopheles* is a polyphyletic assemblage of species.

Evolution of Anopheles

If *An. implexus* and the species of *Lophopodomyia*, *Stethomyia*, *Baimaia* and *Bironella* are ignored, the three large clades shown in Fig. 2 are the same as those constructed from the

protein sequences of 1,085 single-copy orthologs in the study of Neafsey *et al*. (2015). The topology is the same, with *Nyssorhynchus* sister to a clade comprised of *Anopheles* and *Cellia* as monophyletic sister taxa: *Nyssorhynchus* + (*Anopheles* + *Cellia*). If *An. implexus* is indeed the most primitive species of *Anopheles* (see above), then it must have evolved from an ancestor that predates the splitting of the lineage that gave rise 100 million years ago to the ancestral lineages of *Nyssorhynchus* and *Anopheles* + *Cellia* (Neafsey *et al*. 2015). Although the separation of *An. implexus* from the other *Anopheles* is only weakly supported (Fig. 3, GC = 2), we nevertheless feel justified, based on the morphological distinctions and similarities with genus *Chagasia* noted above, in recognizing it as the monotypic member of a separate subgenus; hence, we hereby resurrect the generic name *Christya* Theobald, 1903, **stat. nov.**, from synonymy with *Anopheles* Meigen, 1818 and recognize it as a valid subgenus for this species. Theobald (1903b) originally introduced *Christya* as a genus with *Christya implexa* as the only included species. In view of the similarities that *An*. (*Christya*) *implexus* shares with species of *Chagasia* (see above), it might be appropriate to recognize *Christya* as a separate genus, but doing so in the absence of molecular data for a greater number of *Anopheles* species would be premature. If, however, *Christya* is found to be distinct enough to be afforded generic status, available morphological and molecular evidence (Sallum *et al*. 2002; Neafsey *et al*. 2015; present study) suggests that *Kerteszia* + *Nyssorhynchus* is the most basal group of *Anopheles*. This would support the suppositions of Belkin (1962), Krzywinski *et al*. (2001b) and Harbach & Kitching (2005). Belkin (1962) hypothesized that anophelines initially differentiated in the American Mediterranean Region; Harbach & Kitching (1998) suggested a possible New World origin of Anophelinae based on the basal placement of *Chagasia* relative to *Anopheles* + *Bironella* in their phylogeny of mosquito genera; and Krzywinski *et al*. (2001b) provided support for the South American origin of Anophelinae based on a phylogeny of 16 anopheline species inferred from sequences of two protein-

coding nuclear genes and the Neotropical distributions of *Chagasia* and subgenera *Anopheles*, *Lophopodomyia*, *Nyssorhynchus* and *Stethomyia* of *Anopheles*. However, an alternate hypothesis was more recently proffered by Harbach (2013), who advanced the theory of Christophers (1933) that the ancestral lineage of *Anopheles* existed before the breakup of Pangaea and subsequently diversified into the extant subgenera after the separation of the continents. The earlier evolution of the lineages that gave rise to *Chagasia* (South America) and *Christya* (Africa) favours the Christophers-Harbach evolutionary scenario, but a great deal more work needs to be done before the origins and genealogical relationships of Anophelinae are known with certainty, and a natural classification of the subfamily can be realized.

Concluding comments

In conclusion, there seems to be little agreement between the phylogenetic relationships of anopheline mosquitoes gleaned from studies conducted thus far. ~~The results of molecular studies in general do not agree with the results of morphological studies because they are based on fewer and different species. Clearly, the exemplar approach is of limited use for resolving deeper relationships~~ and a natural classification of Anophelinae will not be realized until phylogenetic analyses include both morphological and molecular data for most or all species of the subfamily. For the time being, we must continue to use the current generic, subgeneric and informal group taxa as a framework for analysing species relationships and testing phylogenetic hypotheses.

Acknowledgements

We are grateful to Gustavo Rossi, Centro de Estudios Parasitológicos y de Vectores (CCT La Plata-CONICET–UNLP), La Plata, Argentina, for providing the impetus for the study.

References

Anthony, T. G., Harbach, R. E. & Kitching, I. J. (1999). Phylogeny of the Pyretophorus Series of *Anopheles* subgenus *Cellia* (Diptera: Culicidae). *Systematic Entomology*, *24*, 193–205.

Belkin, J. N. (1962). *The mosquitoes of the South Pacific* (*Diptera, Culicidae*) [*sic*]. *Vols 1 and 2*. Berkeley and Los Angeles, CA: University of California Press.

Besansky, N. J. & Fahey, G. T. (1997). Utility of the *white* gene in estimating phylogenetic relationships among mosquitoes (Diptera: Culicidae). *Molecular Biology and Evolution*, *14*, 442–454.

Bourke, B. P., Foster, P. G., Bergo, E. S., Calado, D. C. & Sallum, M. A. M. (2010). Phylogenetic relationships among species of *Anopheles* (*Nyssorhynchus*) Diptera, Culicidae) based on nuclear and mitochondrial gene sequences. *Acta Tropica*, *114*, 88–96.

Christophers, S. R. (1915). The male genitalia of *Anopheles*. *Indian Journal of Medical Research*, *3*, 371–394 + 6 pls.

Christophers, S. R. (1924). Provisional list and reference catalogue of the Anophelini. *Indian Medical Research Memoirs*, *3*, 1–105.

Christophers, S.R. (1933). *The fauna of British India, including Ceylon and Burma. Diptera.Vol. IV. Family Culicidae. Tribe Anophelini.* London: Taylor and Francis.

Collucci, E. & Sallum, M. A. M. (2007). Cladistic analysis of the subgenus *Anopheles* (*Anopheles*) Meigen (Diptera: Culicidae) based on morphological characters. *Memórias do Instituto Oswaldo Cruz*, *102*, 277–291.

Edwards, F. W. (1932). *Genera Insectorum. Diptera, Fam. Culicidae*. Fascicle 194. Brussels: Desmet-Verteneuil.

Faran, M. E. (1980). Mosquito studies (Diptera, Culicidae) XXXIV. A revision of the Albimanus Section of the subgenus *Nyssorhynchus* of *Anopheles*. *Contributions of the American Entomological Society*, *15(7)*, 1–215.

Foley, D. H., Bryan, J. H., Yeates, D. & Saul, A. (1998). Evolution and systematics of *Anopheles*: insights from a molecular phylogeny of Australasian mosquitoes. *Molecular Phylogenetics and Evolution*, *9*, 262–275.

Forey, P. L. & Kitching, I. J. (2000). Experiments in coding multistate characters, pp. 54–80. *In*: Scotland, R. & Pennington, R. T. (eds), *Homology and systematics: coding characters for phylogenetic analysis*. London: Taylor and Francis.

Gillies, M. T. & de Meillon, B. (1968). The Anophelinae of Africa south of the Sahara (Ethiopian Zoogeographical Region). *Publications of the South African Institute for Medical Research*, *54*, 1–343.

Goloboff, P. A. (1993). Estimating character weights during tree search. *Cladistics*, *9*, 83–91.

Goloboff, P. **(**1997). Pee-Wee: a program for parsimony analysis under implied weights. Program and documentation, available at ftp.unt.edu.ar/pub/parsimony.

Goloboff, P. A., Carpenter, J. M., Arias, J. S. & Esquivel, D. R. M. (2008). Weighting against homoplasy improves phylogenetic analysis of morphological datasets. *Cladistics*, *24*, 1–16.

Goloboff, P. A., Farris, J. S., Källersjö, M., Oxelman, B., Ramírez, M. J. & Szumik, C. A. (2003). Improvements to resampling measures of group support. *Cladistics*, *19*, 324−332.

González-Santillán, E. & Prendini, L. (2015). Phylogeny of North American vaejovid scorpion subfamily Syntropinae Kraepelin, 1905, based on morphology, mitochondrial and nuclear DNA. *Cladistics*, 31, (in press).

Grjebine, A. (1966). *Faune de Madagascar*. *XXII*. *Insecies Diptéres Culicidae Anophelinae*. Paris: Centre National de la Recherche Scientifique, Office de la Recherche Scientifique et Technique Outre-Mer.

Harbach, R. E. (1994). Review of the internal classification of the genus *Anopheles* (Diptera: Culicidae): the foundation for comparative systematics and phylogenetic research. *Bulletin of Entomological Research*, *84*, 331–342.

Harbach, R. E. (2004). The classification of genus *Anopheles* (Diptera: Culicidae): a working hypothesis of phylogenetic relationships. *Bulletin of Entomological Research*, *95*, 537–553.

Harbach, R. E. (2007). The Culicidae (Diptera): a review of taxonomy, classification and phylogeny, pp. 591–688. *In*: Zhang, Z.-Q. & Shear, W. A. (eds), *Linnaeus Tercentenary: Progress in Invertebrate Taxonomy*. *Zootaxa*, *1668*, 1–766.

Harbach, R. E. (2013). The phylogeny and classification of *Anopheles*, pp. 3–55. *In*: Manguin, S. (ed.), *Anopheles* mosquitoes - New insights into malaria vectors. Rijeka, Croatia: InTech.

Harbach, R. E. & Howard, T. M. (2009). Review of the genus *Chagasia* (Diptera: Culicidae: Anophelinae). *Zootaxa*, *2210*, 1–25.

Harbach, R. E. & Kitching, I. J. **(**1998). Phylogeny and classification of the Culicidae (Diptera). *Systematic Entomology*, *23*, 327–370.

Harbach, R. E. & Kitching, I. J. (2005). Reconsideration of anopheline phylogeny (Diptera: Culicidae: Anophelinae) based on morphological data. *Systematics and Biodiversity*, *3*, 345–374.

Krzywinski, J., Wilkerson, R. C. & Besansky, N. (2001a). Evolution of mitochondrial and ribosomal gene sequences in Anophelinae (Diptera: Culicidae): implications for phylogeny reconstruction. *Molecular Phylogenetics and Evolution*, *18*, 479–487.

Krzywinski, J., Wilkerson, R. C. & Besansky, N. (2001b). Toward understanding Anophelinae (Diptera, Culicidae) phylogeny: insights from nuclear single copy genes and the weight of evidence. *Systematic Biology, 50*, 540–556.

Marks, E. N., Mackerras, I. M., Lee, D. J. & Iyengar, M. O. T. (1963). *Bironella gracilis* Theobald, 1905 (Insecta, Diptera): proposed validation under the plenary powers. *Bulletin of Zoological Nomenclature, 20*, 206–209.

Mohanty, A., Swain, S., Kar, S. K. & Hazra, R. K. (2009). Analysis of the phylogenetic relationship of *Anopheles* species, subgenus *Cellia* (Diptera: Culicidae) and using it to define the relationship of morphologically similar species. *Infection, Genetics and Evolution*, *9*, 1204–1224.

Neafsey, D. E. *et al*. [120 authors] (2015). Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes. *Science, 347*: DOI 10.1126/science.1258522.

Nixon, K. C. (1999–2002). Winclada version 1.0000, published by the author. Ithaca, New York. Available at http://www.cladistics.org.

Norris, L.C. & Norris, D.E. 2015. Phylogeny of anopheline (Diptera: Culicidae) species in southern Africa, based on nuclear and mitochondrial genes. *Journal of Vector Ecoloogy, 40*, 16–27.

Reid, J. A. 1968. Anopheline mosquitoes of Malaya and Borneo. *Studies from the Institute for Medical Research Malaya, 31*, 1–520.

Reid, J. A. & Knight, K. L. **(**1961). Classification within the subgenus *Anopheles* (Diptera, Culicidae). *Annals of Tropical Medicine and Parasitology, 55*, 474–488.

Reidenbach, K. R., Cook, S., Bertone, M. A., Harbach, R. E., Wiegmann, B. M. & Besansky, N. J. (2009). Phylogenetic analysis and temporal diversification of mosquitoes (Diptera: Culicidae) based on nuclear genes and morphology. *BMC Evolutionary Biology*, *9*, 298.

Reinert, J. F., Harbach, R. E. & Kitching, I. J. (2009). Phylogeny and classification of tribe Aedini (Diptera: Culicidae). *Zoological Journal of the Linnean Society*, *157*, 700–794 + 2 online appendices.

Sallum, M. A. M., Foster, P. G., Li, C., Sithiprasasna, R. & Wilkerson, R. C. (2007). Phylogeny of the Leucosphyrus Group of *Anopheles* (*Cellia*) (Diptera: Culicidae) based on mitochondrial gene sequences. *Annals of the Entomological Society of America*, *100*, 27–35.

Sallum, M. A. M., Schultz, T. R., Foster, P. G., Aronstein, K., Wirtz, R. A. & Wilkerson, R. C. (2002). Phylogeny of Anophelinae (Diptera: Culicidae) based on nuclear ribosomal and mitochondrial DNA sequences. *Systematic Entomology*, *27*, 361–382.

Sallum, M. A. M., Schultz, T. R. & Wilkerson, R. C. (2000). Phylogeny of Anophelinae (Diptera Culicidae) based on morphological characters. *Annals of the Entomological Society of America*, *93*, 745–775.

Sereno, P. C. (2007). Logical basis for morphological characters in phylogenetics. *Cladistics*, *23*, 565–587.

Theobald, F. V. (1903a). *A monograph of the Culicidae or mosquitoes.* Volume 3. London: British Museum (Natural History).

Theobald, F. V. (1903b). Report on a collection of mosquitoes and other flies from Equatorial East Africa and the Nile provinces of Uganda. *Reports of the Sleeping Sickness Commission of the Royal Society*, *7*, 33–42, map.

Theobald, F. V. (1904). New Culicidae from the Federated Malay States. *Entomologist*, *37*, 12–15, 36–39, 77–78, 111–113, 163–165, 211–213, 236–239.

Theobald, F. V. (1905). A catalogue of the Culicidae in the Hungarian National Museum, with descriptions of new genera and species. *Annales Musei Nationalis Hungarici*, *3*, 61–119, figures legends + 4 pls.

Wang, G., Li, C., Guo, X., Xing, D., Dong, Y. & Zhao, T. (2014). Molecular phylogenetic analysis of the subgenera *Anopheles* and *Cellia* (Diptera: Culicidae) based on nuclear ribosomal sequences. *African Entomology*, *22*, 660–669.

Wilkerson, R. C. & Peyton, E. L. (1990). Standardized nomenclature for the costal wing spots of the genus *Anopheles* and other spotted-wing mosquitoes (Diptera: Culicidae). *Journal of Medical Entomology*, *27*, 207–224.

Figure legends

Fig. 1 Graph showing the summed GC supports at *K* values 1–33 and a terminal equal weighted analysis using the groups found in the *K* = 30 MPC (see text for details).

Fig. 2 The most parsimonious cladogram found for *K* = 30.

Fig. 3 Tree obtained from the *K* = 30 MPC when GC values < 1 are collapsed (only values > 0 are shown). An asterisk denotes values that may be 1 less (e.g. 8 instead of 9 for *An. albitarsis* + *An. braziliensis*) in one or more of the 10 replicates (see text for details).

Fig. 4 The *Kerteszia* + *Nyssorhynchus* clade of figure 3 with the informal Sections and Series delineated.

Fig. 5 The *Cellia* clade of figure 3 with the informal Series and species Groups delineated.

Fig. 6 The terminal clade of figure 3 with the informal Sections, Series and species Groups of subgenus *Anopheles* delineated.

Supporting information

Fig. S1. Character support for the clades of the MPC for *K* = 30 shown in Fig. 2. Numbers on the branches correspond to the characters and codes listed in Appendix S1 and the data set shown in Appendix S2. Closed circles indicate "unique" character states that can be placed onto the cladogram in only a single position, although they may be interpreted as undergoing subsequent transformation or secondary reversal. Open circles represent homoplastic character states that are placed on more than one branch of the cladogram.

Fig. S2. Phylogeny of subfamily Anophelinae, modified from Harbach & Kitching (2005), indicating relationships within subgenus *Anopheles*. Filled circles indicate Bremer support values greater than 0.8.

Fig. S3. Phylogeny of subgenera *Cellia*, *Kerteszia* and *Nyssorhynchus*, modified from Harbach & Kitching (2005), indicating relationships within subgenera *Cellia* and *Nyssorhynchus*. Filled circles indicate Bremer support values greater than 0.8.

Appendix S1. Annotated list of morphological characters scored for the two outgroup and 67 ingroup taxa included in the cladistic analyses. The morphological terminology used herein is listed and defined in the Anatomical Glossary of the Mosquito Taxonomic Inventory (http://mosquito-taxonomic-inventory.info/). Numbers in brackets following character numbers are those that Harbach & Kitching (2005) used in whole or in part for the same characters. See the data matrix in Appendix S2.

Appendix S2. Data matrix for the 69 taxa and 224 morphological characters included in the cladistic analyses. Missing data are indicated by "?"; characters that could not be scored due to absence of homologous structures are indicated by "–". Polymorphic characters are explicitly coded as such. See Appendix S1 for character descriptions.