Systematic Entomology

Systematic Entomology (2018), **43**, 355–371 DOI: 10.1111/syen.12279

Phylogenetic relationships and molecular delimitation of *Culicoides* Latreille (Diptera: Ceratopogonidae) species in the Afrotropical region: interest for the subgenus *Avaritia*

M. T. BAKHOUM[1,2], K. LABUSCHAGNE[3,4], K. HUBER[5], M. FALL[1], B. MATHIEU[6], G. VENTER[3,7], L. GARDÈS[2], T. BALDET[2], J. BOUYER[2], A. G. FALL[1], G. GIMONNEAU[8] and C. GARROS[2,9]

[1]Institut Sénégalais de Recherches Agricoles, Laboratoire National de l'Elevage et de Recherches Vétérinaires, BP: 2057 Dakar-Hann Route du Front de Terre, Sénégal, [2]Centre de coopération internationale en recherche agronomique pour le développement, UMR117 ASTRE, Montpellier, France, [3]Agricultural Research Council-Onderstepoort Veterinary Research, Epidemiology, Parasites and Vectors, Onderstepoort, South Africa, [4]Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa, [5]Centre de coopération internationale en recherche agronomique pour le développement, UMR ASTRE, INRA, Montpellier, France, [6]Faculté de Médecine, Institut de Parasitologie et de Pathologie tropicale de Strasbourg (IPPTS), EA7292, Strasbourg, France, [7]Department of Veterinary Tropical Diseases, University of Pretoria, Pretoria, South Africa, [8]Centre de coopération internationale en recherche agronomique pour le développement, UMR INTERTRYP, Montpellier, France and [9]Centre de coopération internationale en recherche agronomique pour le développement, UMR117 ASTRE, Sainte Clotilde, Réunion

Abstract. Phylogenetic relationships of *Culicoides* species of the Afrotropical region are problematic as different authors disagree on the placement of species into specific subgenera or groups. In this study we sequenced two mitochondrial (*COI* and *16S* rDNA) and two nuclear (*CAD* and *28S* rDNA) gene fragments to reconstruct phylogenetic relationships within the *Avaritia*, *Remmia* and *Synhelea* subgenera and the Milnei, Neavei and Similis groups of *Culicoides* using both Bayesian inference and maximum-likelihood approaches. Based on phylogenetic trees, we used the bGMYC (Bayesian General Mixed Yule Coalescent model) and the PTP (Bayesian Poisson Tree Processes) to investigate species boundaries. All species relationships within the studied subgenera and groups were well-supported by using morphological characters and molecular analyses. The subgenus *Avaritia* includes (i) all of the species of the Imicola group, as well as the putative new species, *C. sp*. #22, and we confirmed the monophyly of this group; (ii) the Dasyops group includes *C. kanagai* and *C. sp*. #54 Meiswinkel (new species), shown to be monophyletic; (iii) the *C. sp*. #20 belongs to the Orientalis group; (iv) *C. grahamii*, *C. gulbenkiani* and *C. kibatiensis*. Our results also show that subgenus *Remmia* is monophyletic. Relationships of species of the Milnei group were well-supported and demonstrate the monophyly of this group. Borkent's classification for Similis group is confirmed. In addition, *C. neavei* and *C. ovalis* (Neavei group) are placed in the subgenus *Synhelea*.

Introduction

The biting midges in the genus *Culicoides* Latreille (Diptera: Ceratopogonidae), transmit a number of viruses to domestic and wild ruminants, and equids (Mellor *et al*., 2000; Mullen, 2009). This genus is distributed worldwide and includes about 1358 described species (Borkent, 2016). Currently this genus is classified into 32 subgenera containing a number of groups; 38 groups are unaffiliated with a subgenus (Borkent, 2016). Moreover there is a long list of miscellaneous species, not placed in any group, representing 13% of the world fauna.

The internal classification of the genus is based on the morphological similarity between species that includes wing pattern or the shape of male genital structures, which in no

Correspondence: Mame Thierno Bakhoum, Institut Sénégalais de Recherches Agricoles, Laboratoire National de l'Elevage et de Recherches Vétérinaires, BP: 2057 Dakar–Hann Route du Front de Terre, Sénégal. E-mail: thierno.bakhoum@gmail.com

way reflects real phylogenetic relationships (Borkent, 2016). There is no consensus on the definition of groups or species complexes for the genus *Culicoides*, to the point that the literature is full of identical names for different sets and different names for identical sets (Harrup *et al*., 2015). Testing of the monophyly of subgenera and species groups has been limited and subgenera require a systematic revision at the global level (Mathieu, 2011). The subgeneric classification is almost phenetic and based mostly on regional assessments, with limited or not updated descriptions of subgeneric definition. Indeed, some subgenera show evidence of being polyphyletic – for example, the subgenus *Oecacta* Poey (Szadziewski *et al*., 2016) is commonly called the 'dumping subgenus'.

With the advances made in molecular DNA sequencing, several molecular markers have been assessed positively for inferring phylogenetic relationships or molecular delineation of *Culicoides* species (Harrup *et al*., 2015). These markers are now used widely for species identification, especially for those that are difficult to separate morphologically (Sebastiani *et al*., 2001; Pagès *et al*., 2005, 2009; Nolan *et al*., 2007; Monaco *et al*., 2010), but they have been used rarely to assess phylogenetic relationships. Recently, Bellis *et al*. (2013) published a revision of the *Culicoides* Imicola group using a combination of morphological and molecular analyses. However, the validity of all species groups and subgenera within the systematic classification need to be revised to assess their phylogenetic validity. As an attempt to clarify and help subgeneric affiliation, Harrup *et al*. (2015) published a wing atlas for the described subgenera. Recently, the classification of the subgenus *Avaritia* Fox was revised using molecular phylogeny (Mathieu, 2011; Bellis *et al*., 2013; Gopurenko *et al*., 2015). The critical issue in *Culicoides* systematics is that it requires phylogenetic validity (Harrup *et al*., 2015; Labuschagne, 2016).

Culicoides-borne pathogens (Diptera: Ceratopogonidae) in the Afrotropical region are of interest because of major recent outbreaks affecting livestock (Mellor *et al*., 2000; Mullen, 2009; Purse *et al*., 2015; Carpenter *et al*., 2017) and human populations (Agbolade *et al*., 2006;Simonsen *et al*., 2011 ; Bassene *et al*., 2015 ; Debrah *et al*., 2017). For example, in the last 20 years, African horse sickness outbreaks have been recorded in South Africa (Venter *et al*., 2006), Senegal (Diouf *et al*., 2012) and Namibia (Scacchia *et al*., 2009). Recently in West and Central Africa, high prevalence rates of *Mansonella perstans* were recorded in *Culicoides* specimens and human populations (Simonsen *et al*., 2011; Bassene *et al*., 2015; Debrah *et al*., 2017). However, there have been few studies on the *Culicoides* fauna of the Afrotropical region. Although the first *Culicoides* in this region was described over a century ago (Enderlein, 1908), despite the high and undoubtedly underestimated diversity of *Culicoides*. Today, *Culicoides* species diversity in the Afrotropical region reaches 190 described species (Cornet & Chateau, 1970; Cornet *et al*., 1974; Itoua *et al*., 1987; Meiswinkel & Dyce, 1989; Glick, 1990; Bakhoum *et al*., 2013; Labuschagne, 2016) with about 105 *Culicoides* species recorded in South Africa (Labuschagne, 2016) and 53 species in Senegal (Fall *et al*., 2015). These *Culicoides* species are placed in nine subgenera (*Avaritia*, *Beltranmyia* Vargas, *Culicoides* Latreille, *Meijerehelea* Wirth and Hubert, *Monoculicoides* Khalaf, *Pontoculicoides* Remm, *Remmia* Glukhova, *Synhelea* Kieffer, and *Trithecoides* Wirth and Hubert); nine species groups, unplaced to subgenus (Accraensis, Albovenosus, Bedfordi, Dekeyseri, Inornatipennis, Milnei, Neavei, Nigripennis, and Similis); and miscellaneous species, not placed in any group, representing 28% of the Afrotropical fauna. Borkent placed species of the Accraensis, Bedfordi and Similis groups within the subgenus *Synhelea* (Borkent, 2016); this was in contradiction with Meiswinkel and Dyce who published a study of this subgenus containing only the Tropicalis group (Meiswinkel & Dyce, 1989). The Milnei group contains species of medical and veterinary interest because *C. zuluensis* de Meillon transmit Lesetele virus, bluetongue and Akabane viruses have been isolated from *C. milnei* Austen, and *Onchocerca gutturosa* has been isolated from *C. krameri* Clastrier (Meiswinkel *et al*., 2004). This species group was unplaced in any subgenus by Borkent (2016), whereas Meiswinkel places this group within the subgenus *Hoffmania* Fox (R. Meiswinkel, unpublished data). The classification of *Culicoides* species is problematic because different authors disagree on the placement of species into specific subgenera or groups (Khamala & Kettle, 1971; Boorman & Dipeolu, 1979; Itoua *et al*., 1987; Glick, 1990).

In addition, monographs and catalogues used to identify *Culicoides* species for this region are old, at times inaccurate, with low quality illustrations. Before 1970, only two taxonomic keys were available for adults, limited mainly to the species present in South Africa and East Africa (Kenya, Tanzania and Uganda) (Colçao, 1946; Fiedler, 1951). The revision undertaken by Glick in 1990, in collaboration with Cornet, a West African fauna specialist, includes morphological keys for adult females and males of 55 species identified in Kenya; this remains a vital reference work. In West Africa, taxonomic studies are rare and often limited to subregions. Cornet and their associates worked on groups of interest such as the Schultzei, Milnei and Similis groups (Cornet & Chateau, 1970; Cornet *et al*., 1974; Itoua *et al*., 1987; Cornet & Brunhes, 1994). Some of their works include taxonomic identification keys for adults. Meiswinkel described or redescribed and compiled an identification key for adults and immature stages for the species of the *Culicoides imicola* group (Meiswinkel, 1995; Nevill *et al*., 2007). Therefore, until all species have been compared molecularly as well as morphologically, it remains up to individual authors to either use the current published subgeneric classification by Borkent or the species groups (Labuschagne, 2016).

Besides the problematic systematics, species delimitation is also complicated by large morphological variations observed within certain species. Species identification is made more difficult when the name-bearing specimen type is lost, and the description is old with limited drawings or pictures. To overcome this problem, numerous studies have investigated phylogenetic relationships using molecular data together with species delimitation methods in other insect genera (Toussaint *et al*., 2015). In our opinion, these methods can be used to revise the limit and classification of *Culicoides* species in the Afrotropical region.

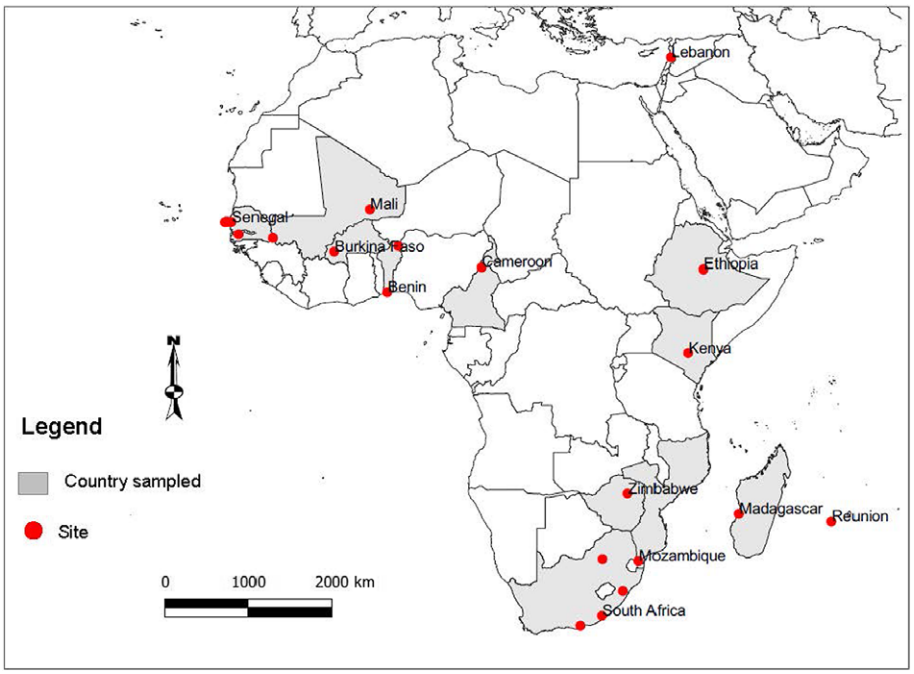


Fig. 1. Geographical map of the *Culicoides* sampling sites. [Colour figure can be viewed at wileyonlinelibrary.com].

In the present study, we aimed at (i) inferring phylogenetic relationships of *Culicoides* species collected mainly in the Afrotropical region using two mitochondrial genes (*COI* and *16S* rDNA) and two nuclear genes (*CAD* and *28S* rDNA) to investigate the monophyly of *Avaritia*, *Remmia* and *Synhelea* subgenera, and Milnei, Neavei and Similis species groups; and (ii) delineating species boundaries using bPTP (Bayesian Poisson Tree Processes) and bGMYC (Bayesian General Mixed Yule Coalescent) methods. This work will revise the current internal systematic classification of *Culicoides* and help future work on the identification of *Culicoides* species in the Afrotropical region.

Materials and methods

Culicoides collection and morphospecies identification

Culicoides specimens were collected in 23 sites located in 12 countries of the Afrotropical region (Fig. 1) and two sites in Lebanon (Palearctic region). *Culicoides* were collected through different field missions between 2009 and 2016 with OVI and CDC light traps set at farms or near equids. Specimens were preserved in 70% ethanol, identified and sexed under a binocular microscope using the available identification keys for the region (Boorman, 1989; Glick, 1990; Cornet & Brunhes, 1994; Labuschagne, 2016). Based on the world systematic catalogue of *Culicoides* species (Borkent, 2016) and Labuschagne's classification (Labuschagne, 2016), specimens belonging to the *Avaritia*, *Remmia* and *Synhelea* subgenera, and Milnei, Neavei and Similis groups were considered in this study. All specimens morphologically identified (or closely related) as species belonging to the above-mentioned subgenera or groups were kept. For each specimen, the wings and genitalia were dissected prior to DNA extraction processing and slide-mounted to record morphological features. All samples are kept as a reference at Cirad, UMR117 ASTRE, Montpellier, France, and are available upon request to the corresponding author.

DNA extraction, amplification and sequencing

Thoraxes of the *Culicoides* were individually homogenized in 50 µL of phosphate buffered saline 1×. After crushing using a piston pellet, genomic DNA was extracted using the *NucleoSpin® Tissue* DNA Kit (Macherey-Nagel, Bethlehem, PA, USA) according to the manufacturer's instructions and maintained at −20°C until further use.

Four fragments from cytochrome oxidase subunit I (*COI*), 16S ribosomal DNA (*16S* rDNA), 28S ribosomal DNA (*28S* rDNA) and *CAD* (carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase) genes were amplified through PCR (Table 1). PCR amplification reactions were performed in a 25-µL total reaction volume containing 1× Qiagen buffer, 1 mM $MgCl^2$, 0.25 mM of each dNTP, 0.2 µM of each primer, 1.25 U Qiagen Polymerase Taq and 0.4 or 0.7 ng/µL genomic DNA depending the gene (*COI*, *16S* rDNA and *28S* rDNA genes or *CAD* gene).

Step-up PCR programs for *COI*, *16S* rDNA and *28S* rDNA included one step of 5 cycles before final step with 35 cycles. The PCR cycling conditions were as follows: an initial denaturation step at 94°C for 5 min followed by: five cycles of 94°C for 30 s; 45°C for *COI*, 42°C for *16S* rDNA or 55°C for *28S* rDNA for 40 s; 72°C for 1 min; 35 cycles of 94°C for 30 s; 51°C for

Table 1. Primers used for PCRs and sequencing in this study.

Gene	Primer name	Sequence (5′–3′)	Length of amplified fragment (bp)	References
COI	LCO1490	GGTCAACAAATCATAAAGATATTGG	710	Folmer *et al*. (1994)
	HCO2198	TAAACTTCAGGGTGACCAAAAAATCA		
16S rDNA	16SF1	CACGTAAGAACTAAATAGTCGAAC	450	Ekrem *et al*. (2010)
	16SR1	GACCGTGCAAAGGTAGCATAATC		
28S rDNA	28S_S3660	GAG AGT TMA ASA GTA CGT GAA AC	657	Dowton & Austin (1998)
	28S_A335	TCG GAA GGA ACC AGC TAC TA		Whiting *et al*. (1997)
CAD	TB1.F TB1.R	TTGGCCGTAAGTTCGAGGAAG AGTTCACGCAAACATCCAACG	674	–
	787F	GGD GTN ACN ACN GCN TGY TTY GAR CC	905	Moulton & Wiegmann (2004)
	1098R	TTN GGN AGY TGN CCN CCC AT		

COI, 55°C for *16S* rDNA or 50°C for *28S* rDNA for 30 s; 72°C for 1 min; and a final extension step at 72°C for 10 min. The touch-down amplification PCR conditions for CAD involved two steps: four cycles of 94°C for 30 s, 51°C for 40 s and 72°C for 1 min, and six cycles of 94°C for 30 s, 47°C for 40 s and 72°C for 1 min. These first two steps were followed by 30 cycles of 94°C for 30 s, 42°C for 40 s, 72°C for 1 min, and a final extension step at 72°C for 10 min. The PCR products were visualized on 1.5% agarose gels with a Gel Red staining after migration of 90 min at 100 volts by electrophoresis for quality control, and the remaining 20 μL were sequenced using the same primers as used in PCR amplifications (https://www.genewiz.com).

Sequence analysis

Amplified sequences were used as query in a BLAST search in the NCBI database to confirm that the amplified sequences were the target genes. The DNA sequences were edited in Geneious R6 (Biomatters, http://www.geneious.com/). Sequences of each gene were independently aligned using MACSE (Multiple Alignment of Coding SEquences accounting for frame shifts and stop codons) (Ranwez *et al*., 2011) for *COI* and *CAD* genes. For *16S* and *28S* rDNA sequences, alignments were generated using Muscle (Edgar, 2004). For each alignment, segments that had too many variable positions or gaps were removed using Gblocks 0.91b (Castresana, 2000) to make alignments more appropriate for phylogenetic reconstruction. The reading frames and sequence statistics were checked under MEGA v6.0 (Tamura *et al*., 2013). A test of substitution saturation (Xia *et al*., 2003) was performed in DAMBE (Xia & Xie, 2001). Rapid detection of selective pressure on individual sites of codon alignments for the *CAD* gene was performed using Datamonkey (Pond & Frost, 2005). Sites under positive or negative selection in this gene were inferred using the single-likelihood ancestor counting (SLAC), fixed-effects likelihood (FEL), mixed effects model of evolution (MEME) and Fast Unconstrained Bayesian AppRoximation (FUBAR) methods as implemented in the Datamonkey server (http://www.datamonkey.org) (Pond & Frost, 2005; Murrell *et al*., 2012). Positive selection for a site was considered to be statistically significant when $P < 0.1$ for the SLAC, FEL and MEME methods, or the posterior probability was <0.9 for the FUBAR method. Selected sites with $P < 0.05$ were reported.

The following sequences were used in this study: *COI* sequences from a *C. bolitinos* Meiswinkel population from Reunion Island (A. Desvars, J.C. Delecolle, F. Biteau, G. Gerbier, F. Roger and T. Baldet, unpublished data) [accession numbers: HQ447061.1 and HQ447062.1], a *C. fulvus* Sen and Das Gupta population from Australasia (Gopurenko *et al*., 2015) [accession numbers: KT352267.1, KT352340.1, KT352547.1, KT352629.1 and KT352696.1], a *C. miombo* Meiswinkel population from Benin (B. Mathieu, C. Garros, T. Balenghien, E. Candolfi, J.C. Delecolle, and C. Cetre-Sossah, unpublished data) [accession numbers: KF417704.1 and KF417705.1] and a *C. similis* Carter population from India (Harrup *et al*., 2016) [accession numbers: KT307841.1-KT307842.1]; and *CAD* sequences from *C. kwagga*, *C. loxodontis* Meiswinkel, *C. gulbenkiani* Caeiro, *C. tuttifrutti* Meiswinkel and *C. bolitinos* populations from South Africa [accession numbers: KJ163032.1, KJ163034.1, KJ163025.1, KJ163044.1, KJ163009.1], and *C. imicola* Kieffer from France [accession number: KJ163028.1] (Bellis *et al*., 2013).

Phylogenetic inferences

Phylogenetic trees were reconstructed for the four markers using Bayesian inference (BI) and maximum-likelihood (ML) under a substitution model found using jModelTest (Darriba *et al*., 2012). The Bayesian information criterion (BIC) implemented within jModelTest was used to determine the most suitable evolutionary model(s). We used BI and ML to reconstruct phylogenetic relationships of all sequenced specimens using *COI*, *16S* rDNA, *28S* rDNA and *CAD* genes separately. For the concatenated alignment including *COI*, *16S* rDNA and *28S* rDNA, the best-fit partitioning scheme and partition-specific substitution model were tested in PartitionFinder v1.1 (Lanfear *et al*., 2012) using the *greedy* algorithm, and the *mrbayes* or *raxml* set of models. The BI analyses were performed using MrBayes 3.2.3 (Ronquist *et al*., 2012). Two simultaneous and independent runs consisting of 16 Metropolis-coupled Markov chain Monte Carlo (MCMC) running 50 million generations were used, with a tree sampling every 1000 generations to

calculate posterior probabilities (PP). In order to investigate the convergence of the runs, we investigated the split frequencies and Effective Sample Size (ESS) of all the parameters, and plotted the log-likelihood of the samples against the number of generations in Tracer 1.5 (http://BEAST.bio.ed.ac.uk/Tracer). A value of ESS > 2710 was found as a good indicator of convergence. The ML analyses were conducted with the best model selected using PhyML 3.0 (Guindon *et al*., 2010) for each dataset and RAxML for the concatenated dataset (Stamatakis, 2014). We performed 1000 bootstrap replicates to investigate the level of support at each node.

Molecular species delimitation

We used the Bayesian Poisson Tree Processes (bPTP) on a molecular phylogenetic tree constructed from *COI*, *16S* rDNA and *28S* rDNA genes concatenated and the Bayesian General Mixed Yule Coalescent (bGMYC) on 100 ultrametric trees from each gene in order to delimit *Culicoides* species. The PTP method (Zhang *et al*., 2013) inferred molecular clades based on our inferred molecular phylogeny. The analyses were conducted on the web server for PTP (available at http://species.h-its.org/ptp/) using the Bayesian topology as advocated for this method (Zhang *et al*., 2013). The bGMYC (Reid & Carstens, 2012) is a Bayesian implementation of the GMYC method. This method searches in an ultrametric gene tree the threshold at which branching patterns represent coalescent events or speciation events (Pons *et al*., 2006). We conducted the bGMYC model using ultrametric gene trees inferred in the BEAST 1.8.0 (Drummond *et al*., 2012) without outgroups under a strict clock model and a Speciation: Yule Process Tree Model. The runs consisted of 10 million generations sampled every 1000 cycles. Convergence was assessed by ESS values. A conservative burn-in of 10% was performed after checking the log-likelihood curves in Tracer 1.5. As recommended by Reid & Carstens (2012), 100 trees sampled at intervals from the posterior distribution of trees using LogCombiner 1.8.0 (Drummond *et al*., 2012) were used to perform the bGMYC analyses. Species delimitation analyses were conducted in R using the package 'bGMYC'. For each of the 100 trees selected, the analyses consisted of 250 000 generations with a burnin of 25 000 and a thinning parameter of 100, as performed in Toussaint *et al.* (2015).

Results

Morphological identification

Based on morphological characteristics 47 101 specimens were identified belonging to 58 *Culicoides* species. Of these, 153 specimens of 33 morphological units (morphospecies) belonging to the *Avaritia*, *Remmia* and *Synhelea* subgenera, Milnei, Neavei and Similis groups were selected. These 33 morphological units were distributed as follow: 16 belonging to *Avaritia*, six to *Remmia* (Schultzei group: *C. enderleini* Cornet and Brunhes, *C. kingi* Austen, *C. nevilli* Cornet and Brunhes, *C. oxystoma* Kieffer, *C. subschultzei* Cornet and Brunhes, and *C. schultzei* Enderlein), one to *C. tropicalis* Kieffer, type species of subgenus *Synhelea*, five to the Milnei group (*C. austeni* Carter, *C. isioloensis* Cornet, Nevill and Walker, *C. milnei*, *C. moreli* Clastrier and *C. zuluensis*), two to the Neavei group (*C. neavei* Austen and *C. ovalis* Khamala and Kettle) and three to the Similis group (*C. exspectator* Clastrier, *C. ravus* de Meillon and *C. similis*). The 16 species of *Avaritia* were placed in the Dasyops group (*C. kanagai* Khamala and Kettle, *C. sp*. #54 dark and pale forms), Grahamii group (*C. grahamii* Goetghebuer), Gulbenkiani group (*C. gulbenkiani*), Imicola group (*C. bolitinos*, *C. imicola*, *C. kwagga*, *C. loxodontis*, *C. miombo*, *C. pseudopallidipennis* Clastrier, *C. sp*. #22 and *C. tuttifrutti*, Cornet and Dyce) and Orientalis group (*C. trifasciellus* Goetghebuer and *C. sp*. #20). *Culicoides kibatiensis* of the subgenus *Avaritia* was not grouped. The female wing pattern of these *Culicoides* species was described (Fig. 2).

DNA sequences

Of 153 samples considered in this study, we obtained 139 sequences for the *COI* gene, 147 for *16S* rDNA and 146 for *28S* rDNA. Many samples failed sequencing to the *CAD* gene (66 of 153 samples amplified). DNA sequences of *COI*, *16S* rDNA and *28S* rDNA genes generated in this study are deposited in GenBank (see Table S1 for additional details): *COI* (MF399674–MF399811); *16S* rDNA (MF422796–MF422942); *28S* rDNA (MF422943–MF423087).

The final concatenated alignment of *COI*, *16S* rDNA and *28S* rDNA yielded 132 sequences of 1493 bp for 31 morphological units corresponding to all studied *Culicoides* species excluding *C. trifasciellus* and *C. sp*. #54 pale form. Information relative to sequence statistics and best-fit partitioning scheme and partition-specific substitution model are provided in Tables 2 and 3. Rapid detection of selective pressure on individual sites of 176 codons for the *CAD* gene found 11 positively selected sites. These 11 sites were removed in order to make a better phylogenetic tree using the *CAD* gene. One hundred and sixty five negatively selected sites were also observed for this gene. Saturation tests as a function of the genetic distance estimated under substitution model JC69 showed low saturation of DNA sequence alignments.

Phylogenetic relationships

Phylogenetic analyses conducted with *COI*, *16S* rDNA and *28S* rDNA genes concatenated, and the *CAD* gene are represented in Figs 3 and 4, respectively.

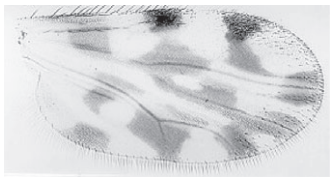
A concatenated *COI*, *16S* rDNA and *28S* rDNA alignment of 132 specimens which included all studied *Culicoides* species excluding *C. trifasciellus* and *C. sp*. #54 pale form, contained six clades: C1–C6 (Fig. 3). *CAD* alignment of 72 sequences, including six sequences from GenBank: KJ163009.1,

Subgenus *Avaritia*

Diagnostics: Eyes contiguous. Sensilla coeloconica on each of flagellomeres: 1, 9-13 or 1, 10-13. The 3rd palpal segment is usually slender with a single sensory pit. Spermathecae: two ovoid well-developed with short necks, third rudimentary spermathecae and presence of a sclerotized ring. Short parameres and usually separate.

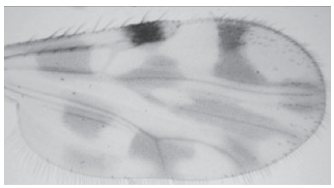
Imicola group

C. imicola (♀):



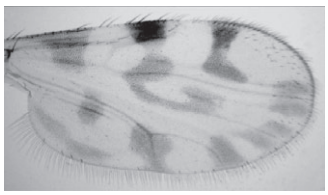
Proximal margin of pale spot in r3 cell is diamond shaped;
Vein M2 is dark at wing margin;
Pale spot above M2 vein is long and narrow.

C. bolitinos (♀):



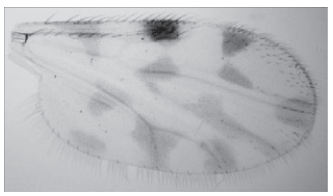
Proximal margin of 3rd post stigmatic pale spot in r3 cell is curved;
Spot over M2 vein at margin is narrowly dark.

C. kwagga (♀):



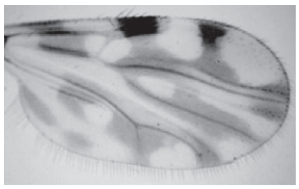
Proximal margin of pale spot in r3 cell is diamond shaped;
Dark spot at angle of anal cell is long.

C. loxodontis (♀):



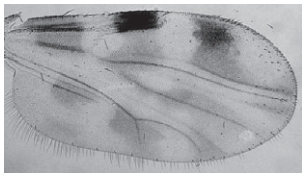
Wing pale with dark spots;
Proximal margin of pale spot in r3 cell is diamond shaped;
Area across M2 vein to wing margin is pale.

C. miombo (♀):



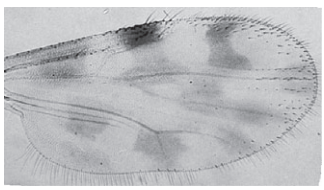
Wing dark with pale smudge;
Dark spot at angle of anal cell is long.

C. sp. # 22 (undescribed) (♀):



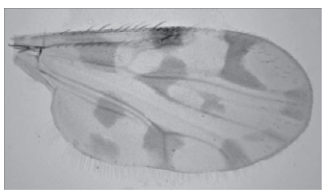
Similar to *C. bolitinos*;
Dark spot between 2nd costal spot and pale spot in r3 cell is trapezoid-shaped.

C. pseudopallidipennis (♀):



Very similar to *C. imicola*;
Spot over M2 vein at margin is narrowly dark.

C. tuttifrutti (♀)



Wing pattern very similar to that of *C. pseudopallidipennis* but spot over M2 vein is narrowly pale.

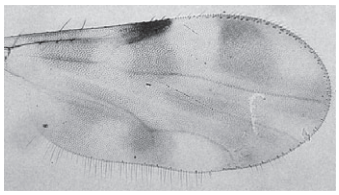
Fig. 2. Female wing pattern of *Culicoides* species of *Avaritia* subgenus Fox, 1955; *Remmia* subgenus Glukhova, 1972; *Synhelea* subgenus *Kieffer*, 1925; and Milnei, Neavei and Similis groups included in our study. The wings were photographed using a ×4 lens. Bars = 200 μm.

KJ163025.1, KJ163028.1, KJ163032.1, KJ163034.1 and KJ163044.1 (Bellis *et al*., 2013), corresponding to 22 species [*C. bolitinos*, *C. gulbenkiani*, *C. imicola*, *C. kanagai*, *C. kibatiensis* Goetghebuer, *C. kwagga*, *C. loxodontis*, *C. miombo*, *C. pseudopallidipennis* and *C. tuttifrutti* (*Avaritia*), *C. austeni*, *C. milnei*, *C. moreli* and *C. zuluensis* (Milnei group), *C. enderleini*, *C. kingi*, *C. nevilli*, *C. oxystoma*, *C. schultzei* and *C. subchultzei* (*Remmia*), *C. tropicalis* (*Synhelea*), and *C. similis*, Ingram and Macfie (Similis group)], contained four clades: C′1 to C′4 (Fig. 4). The clades C1 (Fig. 3) and C′1 (Fig. 4) contained all of the specimens affiliated to the subgenus *Avaritia*. Regarding Fig. 3, *Culicoides grahamii* appeared at

Subgenus *Avaritia*

Grahamii group

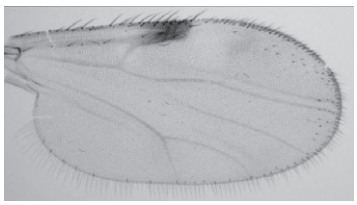
C. grahamii (♀):



Dark wing with of pale spots;
1st costal spot across the r-m crossvein is round;
2nd costal spot is wider than the first costal spot;
Small pale spot in r3.

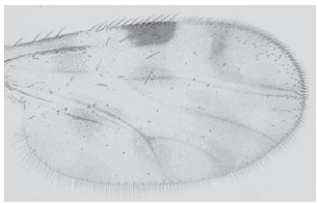
Dasyops group

C. kanagai (♀):



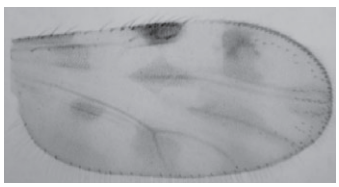
Very pale wing;
Dark area is on costal above and between r-m crossvein and 2nd radial cell.

C. sp. #54 pale form (undescribed) (♀):



Pale wing with faint pattern of dark spots;
Dark spot between pale spot in r3 cell and 2nd costal spot is thin.

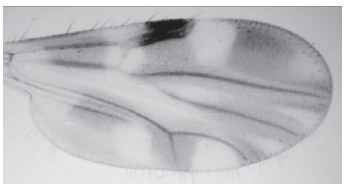
C. sp. #54 dark form (undescribed) (♀):



Similar to that of C. sp. # 54 pale form. Dark spot between spot r3 cell and 2nd costal spot is larger than in *C. sp*. #54 pale form.

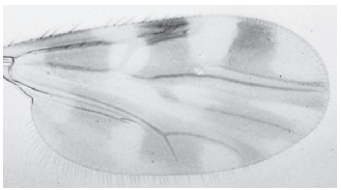
Orientalis group

C. trifasciellus (♀):



1st costal spot is square and 2nd costal spot is round;
Dark spot between 2nd costal spot and small pale spot in r3 is large.

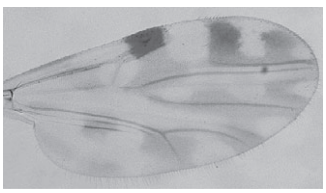
C. sp. # 20 (♀):



Very similar to *C.* trifasciellus;
Pale spot in r3 cell is wider than that of *C. trifasciellus*.

Gulbenkiani group

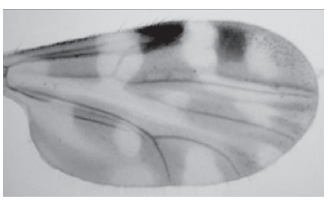
C. gulbenkiani (♀):



Hourglass-shaped pale spot between 2nd costal spot and 3rd post stigmatic pale spot in r3 cell;
Tip of costal vein pale and intrudes into the 2nd costal spot.

Other *Avaritia* (No group)

C. kibatiensis (♀):



Tip of 2nd radial cell pale and intrudes into the 2nd costal spot.
3rd post stigmatic pale spot in r3 cell does not touch the wing margin.

Fig. 2. Continued.

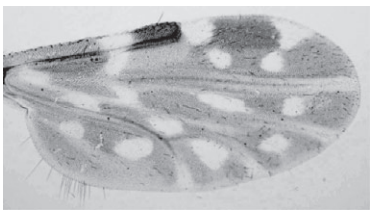
the basal position within clade C1. The internal subdivision of clade C1 (*Avaritia*) is recovered as monophyletic with a strong support for the C2 and C3 clades (PP = 1/ BS = 75 for C2, PP = 1/BS = 83 for C3). The clade C2 contained all the studied species of the Imicola group (erected by Khamala and Kettle in 1971, and completed by Meiswinkel in 1995) as well as putative new species, *C. sp*. #22. Species relationships within the Imicola Group (*C. bolitinos*, *C. imicola*, *C. kwagga*, *C. loxodontis*, *C. miombo*, *C. pseudopallidipennis*, *C. sp*. #22, and *C. tuttifrutti*) were well-supported and resolved at the concatenated phylogenies constructed, but less so at the *CAD* gene. *Culicoides gulbenkiani* sequence from GenBank and *C. kibatiensis* were nested among the Imicola group in the *CAD* phylogeny (Fig. 4). The clade C3 cluster includes two groups

Subgenus *Remmia*

Diagnostics: Eyes narrowly to moderately separate. Sensilla coeloconica are normally on each of flagellomeres 1, 6-8 or 1, 3, 5-8 or 1, 5-8 and occasionally on 3 and 4. The 3rd palpal segment is moderately inflated with a single sensory pit. Distinct wing pattern with 3 to 4 pale spots in r3 cell, first two spots often connected to form an hourglass-shaped spot. Radial cells greatly reduced.

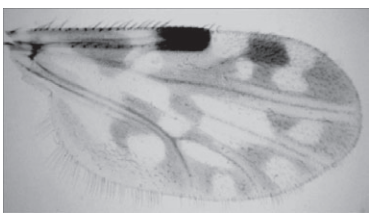
Schultzei group is the only one of this subgenus.

C. enderleini (♀):



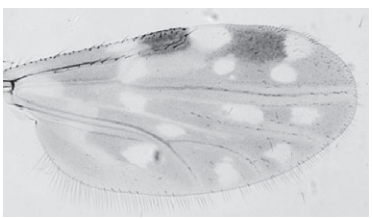
Pale spot in cua1 cell away from wing margin towards cubital fork;

C. kingi (♀):



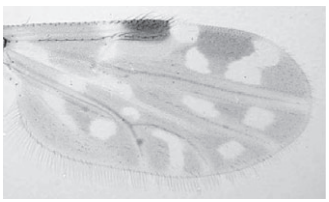
Two pale spots in cell cua1; m2 cell with a basal spot which runs over M2 vein and joins the basal spot in cell m1;
Tips of CuA1 and CuA2 veins are dark-bordered at wing margin.

C. nevilli (♀):



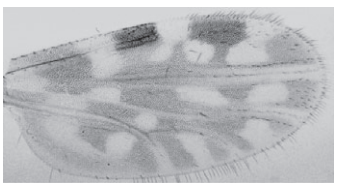
Pale spot in cua1 cell on wing margin;
Tips of CuA1 and CuA2 veins are dark-bordered at wing margin.

C. subschultzei (♀):



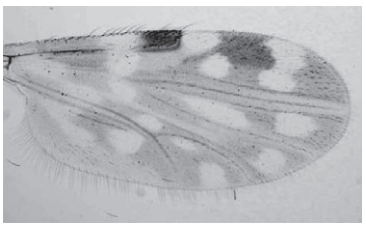
Pale spot in cua1 cell long and narrow touch wing margin but not CuA1 vein.

C. oxystoma (♀):



Wing pattern very similar to that of *C. subschultzei*;
Small pale spot under 2nd radial cell;
Tips of CuA1 and CuA2 veins are pale-bordered at wing margin.

C. schultzei (♀):



Two pale spots in cua1 cell: one is round and touches the wing margin and the other is long next to CuA1 vein.

Fig. 2. Continued.

with strong support: *C. sp*. #20, affiliated to the Orientalis group, and specimens of the Dasyops group, *C. sp*. #54 dark form and *C. kanagai*. In the phylogenetic tree constructed from *COI*, relationship between *C. fulvus* population from Australasia and *C. sp*. #20 was strongly supported in Orientalis group (see Figure S1 for additional details). In addition, *C. sp*. #20 was close to *C. trifasciellus* (Orientalis group) in the phylogenetic tree constructed from the *16S* rDNA gene (see Figure S2 for additional details). Phylogenetic relationships between *C. kanagai*, *C. sp*. #54 dark form and *C. sp*. #54 pale form were confirmed in phylogenetic trees using *16S* rDNA and *28S* rDNA genes (see Figures S2 and S3 for additional details).

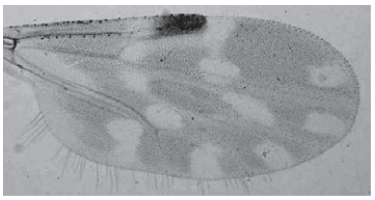
All of the studied species of the Milnei group (*C. austeni*, *C. isioloensis*, *C. milnei*, *C. moreli* and *C. zuluensis*) constituted a monophyletic clade with strong support at the concatenated and *CAD* phylogenies constructed (PP = 1/ BS = 87 for clade C4 in Fig. 3, PP = 1/BS = 100 for clade C′2 in Fig. 4).

All phylogenies (concatenated and *CAD* phylogenies constructed) strongly supported the close relationship between *C. tropicalis*, type species of subgenus *Synhelea*, and species

Tropicalis group (**subgenus *Synhelea***):

Diagnostics: Eyes separated. The 3rd palpal segment is usually slender with a single deep sensory pit. Sensilla coeloconica on each of flagellomeres varies within the group. Spermathecae: two ovoid well developed spermathecae with short necks, 3rd rudimentary spermathecae and presence of sclerotized ring.

C. tropicalis (♀):

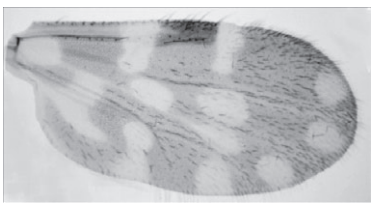


Pale spot below 2nd costal spot crosses vein M1.
1st costal spot does not cross r-m crossvein
.

Neavei group

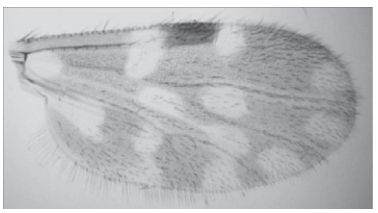
Diagnostics: Eyes separated moderately. No pale spot between 1st and 2nd costal spots. Spermathecae: 2 ovoid spermathecae with 3rd rudimentary spermathecae and sclerotized ring present. Sensilla coeloconica on each of flagellomeres: 1, 8-12 or 1, 9-12.

C. neavei (♀):



Pale spots in r3, m1, m2, cua1 and anal cells touch wing margin.

C. ovalis (♀):

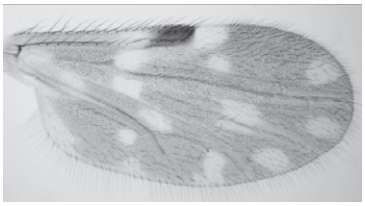


Similar to *C. neavei* but the 2nd costal spot is smaller and round;
Pale spots in basal half of m1 and m2 cells are more square-shaped.

Similis group

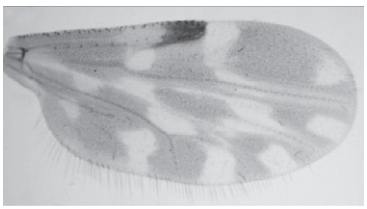
Diagnostics: Eyes very narrowly separated. The 3rd segment of the maxillary palpus inflated with single large, deep sensory pit. Sensilla coeloconica on each of flagellomeres: 1, 3, 5-8 or 1, 5-8 or 1, 8-12 or 1-8. Spermathecae: 2 ovoid spermathecae with 3rd rudimentary spermathecae and sclerotized ring present.

C. similis (♀):



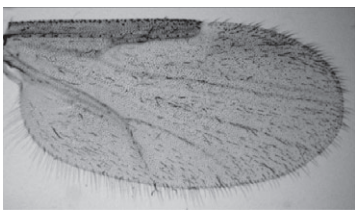
Small pale spots in r2, m1 and m2 cells;
Pale spot slightly below 2nd costal spot does not cross vein M1.

C. exspectator (♀):



Wing with pale streaks situated just above and below middle of vein M2;
Pale spot below the 2nd costal spot touches and merges with vein M1 to form a pale streak.

C. ravus (♀):



No pale markings on the wing;

Fig. 2. Continued.

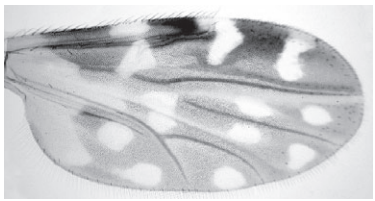
of the Similis group. Clade C5 in Fig. 3 shows strong support (PP = 1/BS = 94) and contains *C.* (*Synhelea*) *tropicalis*, three species of the Similis group (*C. exspectator*, *C. ravus* and *C. similis*) and two species of the Neavei Group (*C. neavei* and *C. ovalis*). Regarding the constructed *CAD* phylogeny (Fig. 4), clade C′3 (PP = 1/BS = 68) contains *C. tropicalis* and *C. similis*.

All species of the subgenus *Remmia* studied were well-supported and congruently resolved in the concatenated and *CAD* phylogenies (PP = 1/BS = 100 for clade C6 in Fig. 3, PP = 1/BS = 100 for clade C′4 in Fig. 4). The subgenus *Remmia* was recovered as monophyletic with strong support (1/100) and contained the six studied species (*C. enderleini*, *C. kingi*, *C. nevilli*, *C. oxystoma*, *C. subschultzei*

Milnei group:

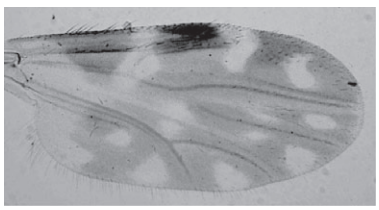
Diagnostics: Wing with prominent pattern of distinct pale spots; distal portion of 2nd radial cell is pale. Spermathecae: two ovoid well-developed with 3rd rudimentary spermathecae and sclerotized ring present at junction of ducts. Sensilla coeloconica on each of flagellomeres: 1, 9-13 or 3, 11-15. The 3rd segment of maxillary palpus is with more than 1 sensory pit.

C. milnei (♀):



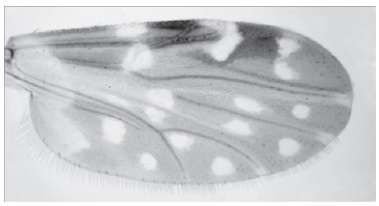
1st costal spot triangular and merges with pale spot above costal vein;
Pale spots in r3, m1 and m2 cells do not touch wing margin;
Single pale spot in cua1 cell on the wing margin.

C. austeni (♀):



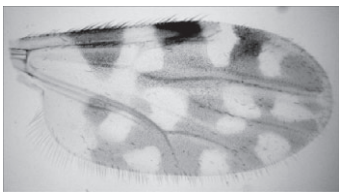
Very similar to that of *C. milnei*;
Pale spots on either side of the middle of the M2 vein are reduced.

C. moreli (♀):



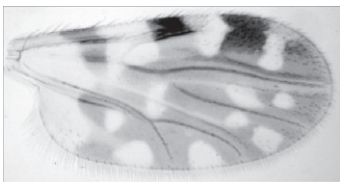
Two pale spots in cua1 cell: one near wing margin and the other at junction of veins CuA2 and CuA1 (cubital fork);
Tips of veins M1, M2, CuA1 are pale.

C. isioloensis (♀):



Double spot straddling vein M2, midway between base and tip of vein and these are connected with streaks to spot in cell m;
Apex of wing pale between M1 and M2 veins;
Two equal pale spots in anal cell, one at wing margin other near vein;
Single pale spot in cua1 cell on the wing margin.

C. zuluensis (♀):



Tips of CuA1, CuA2, M1 and M2 veins are dark;
1st costal spot large and square;
Single pale spot in cua1 cell on the wing margin.

Fig. 2. Continued.

and *C. schultzei*). The phylogenetic relationships between the different clades are not well-supported and do not allow strong conclusions.

Molecular species delimitation

Using *COI*, *16S* rDNA and *28S* rDNA, we found that the number of putative species varied depending on the method and molecular markers used (Fig. 3). Based on the bGMYC method with *16S* rDNA and the *COI* gene, results were similar unlike *28S* rDNA that had a lower resolution. Analysis based on the bPTP method using a concatenated phylogenetic tree yielded very similar results to the clusters formed by morphological identification, unlike analysis based bGMYC method using the genes separately. Putative MOTUs (Molecular operation taxonomic units) were observed in *C. bolitinos*, *C. pseudopallidipennis*, and *C. oxystoma* (Fig. 3). The Cluster of *C. bolitinos* from Madagascar, Mozambique and Reunion (with posterior probability = 0.85) was

Table 2. Sequence statistics of the four gene fragments.

	COI	*16S* rDNA	*28S* rDNA	*CAD*
Length (pb)	567	259	667	495
Percentage C + G (%)	30.53	21.3	39.67	50
Number of variable sites	313	107	606	233
Number of parsimony informative sites	270	90	477	192

Table 3. Partition models and implemented parameters of the BI analysis.

Partitioned dataset	Nucleotide model under BIC	Implemented model and related parameters
①*COI 1st*	GTR + I + G	nst = 6, rates = invgamma
②*COI 2nd*	GTR + I + G	nst = 6, rates = invgamma
③*COI 3rd*	GTR + I + G	nst = 6, rates = invgamma
④*16SrDNA*	GTR + I + G	nst = 6, rates = invgamma
⑤*28SrDNA*	HKY + I + G	nst = 2, rates = invgamma

separated from that *C. bolitinos* from South Africa and Kenya (PP = 0.65) (Fig. 3). We also observed a high level of divergence within *C. pseudopallidipennis* and *C. oxystoma* (Fig. 3). The bPTP analysis conducted on the concatenated phylogenetic tree showed two distinct clusters for *C. pseudopallidipennis* and also *C. oxystoma*. Some specimens of *C. pseudopallidipennis* from Senegal (PP = 0.98) were separated from *C. pseudopallidipennis* from Benin and some other specimens from Senegal (PP = 0.80). Molecular delineation of *C. oxystoma* based on the bPTP analysis using the concatenated phylogenetic tree was very strong (Fig. 3). The cluster of *C. oxystoma* from Lebanon was separated from those of Senegal and Mali with PPs of 1.0 and 0.98, respectively. However, all of these putative MOTUs were not well-supported by bGMYC conducted on individual genes.

Discussion

The aim of this study was to demonstrate the monophyly of subgenera and groups of the genus *Culicoides*, previously described in the literature, with special focus on the Afrotropical region. We used an integrative taxonomic approach incorporating four genes (*COI*, *16S* rDNA, *28S* rDNA and *CAD*) and morphological data to examine species boundaries as well as phylogenetic relationships of the *Avaritia*, *Remmia* and *Synhelea* subgenera, and Milnei, Neavei and Similis groups.

Because of missing sequences for the *CAD* gene, we performed separate phylogenetic analysis for the *CAD* gene and concatenated *COI*, *16S* rDNA and *28S* rDNA genes. Missing data were observed for the *CAD* gene as a consequence of amplification and sequencing difficulties despite several attempts with previous published primer sets (Moulton & Wiegmann, 2004; Bellis *et al*., 2013; Gopurenko *et al*., 2015).

Relationships between *Culicoides* species from the first phylogenetic tree using the *CAD* gene did not fit that generated by other genes or delimited by morphology. Moreover, rapid detection of selective pressure on individual sites of 176 codons for the *CAD* gene found 11 positively selected and 165 negatively selected sites. Selective pressures may have the ability to generate phylogenetic signal that is different from ancestry (Massey *et al*., 2008). Because of the amplification issue and the positive signal for selective pressure, unlike the other molecular markers studied, we do not consider the *CAD* gene to be a good maker for phylogeny studies of Afrotropical *Culicoides* species. Other primers sets from other arthropod groups could be used in future works (Sikes & Venables, 2013). The phylogenetic tree using the *CAD* gene without positively selected sites showed relationships between *Culicoides* subgenera or groups in line with the morphology. Indeed, an increase in the positive selection rate has two possible modes of action on phylogenies: (i) it causes long branch attraction, and (ii) it generates convergence or parallel evolution (homoplasy) through similar selective pressures (Philippe *et al*., 2000; Massey *et al*., 2008). Negative selection also has potential modes of action on phylogenies (Townsend, 2007; Massey *et al*., 2008).

All species relationships within the studied subgenera and groups were well-supported and congruently resolved in the concatenated phylogenetic tree (*COI*, *16S* rDNA and *28S* rDNA), but less so for each gene individual. In concatenated and *CAD* phylogenetic trees, relationships between the different clades were not well-supported and do not allow strong conclusions. In our study, phylogenetic relationships and molecular delimitation of species using bPTP on the concatenated phylogenetic tree were in accord with that delimited by morphology, but less so with bGMYC based on *COI*, *16S* rDNA and *28S* rDNA genes, and showed potential cryptic species within the clusters of *C. bolitinos*, *C. pseudopallidipennis* and *C. oxystoma*. In fact, Cornet and Brunhes suggested that *C. oxystoma* is a species complex (Cornet & Brunhes, 1994); this is corroborated by the highest level of intraspecific divergence being observed in *C. oxystoma* based on *COI* sequences in previous studies (Bakhoum *et al*., 2013; Harrup *et al*., 2016). In our opinion, further investigation of *C. oxystoma* specimens from the distribution area of this species (West Africa, Saharo-Arabian, Oriental and Australian regions) is necessary in order to delineate species within the Oxystoma group.

Molecular analyses using the *CAD* gene and concatenated *COI*, *16S* rDNA and *28S* rDNA genes provided strong evidence that subgenus *Avaritia* includes species of Imicola group, *C. kibatiensis* (not grouped), *C. sp*. #20 (Orientalis group), species of Dasyops group (*C. kanagai* and *C. sp*. #54 dark form), *C. gulbenkiani* (Gulbenkiani group) and *C. grahamii* (Grahamii group). The Imicola group including *C. bolitinos*, *C. imicola*, *C. kwagga*, *C. loxodontis*, *C. miombo*, *C. pseudopallidipennis*, *C. tuttifrutti* and *C. sp*. #22 (a putative new species) is monophyletic. This monophyletic group is regarded as a natural species complex within the subgenus *Avaritia* (Meiswinkel, 1995; Bellis *et al*., 2013). Based on adult characters, Meiswinkel (2004) separated the Imicola and Orientalis groups. Using morphological characters, *C. sp*. #20 from Senegal was closely

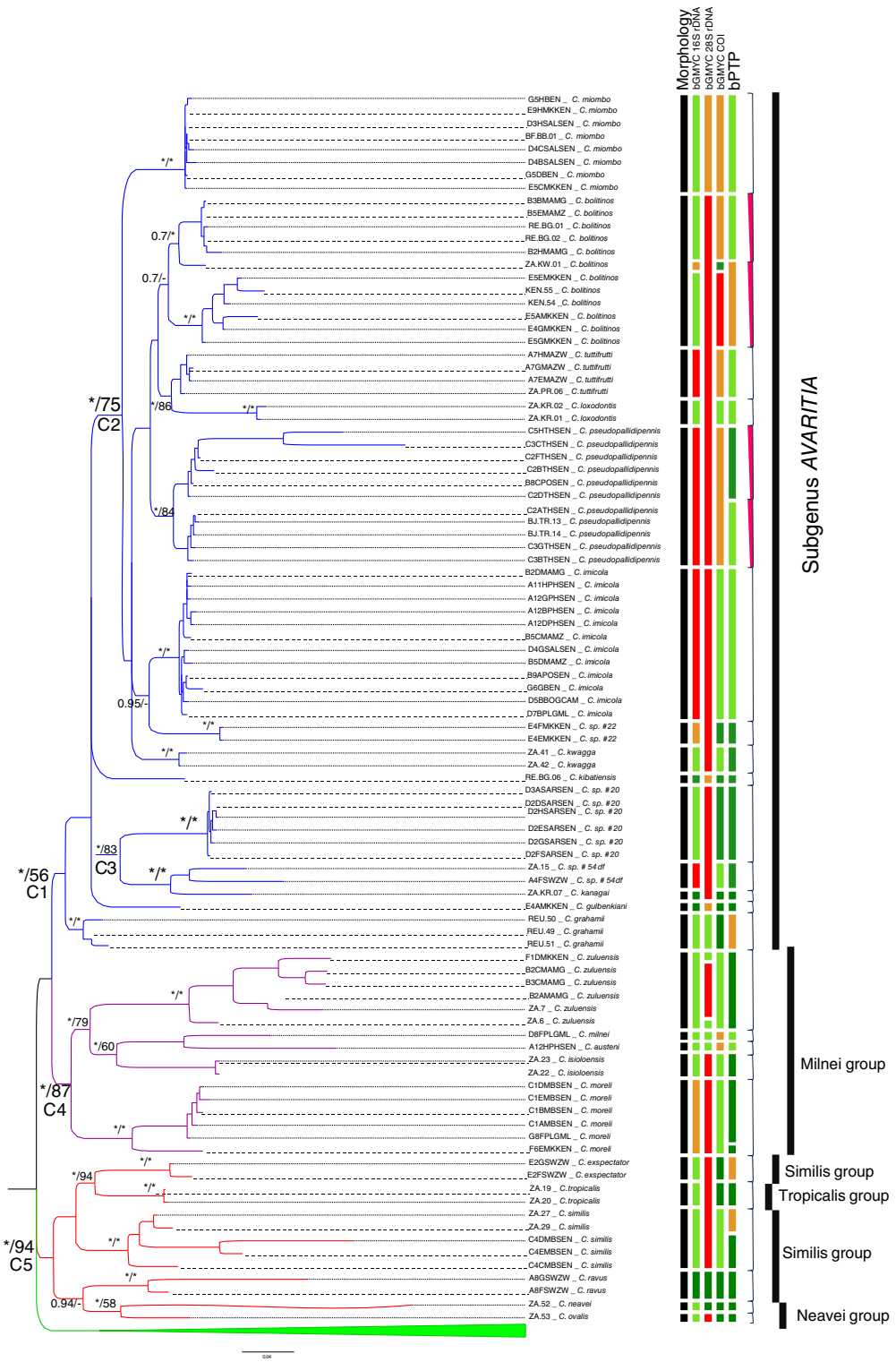


Fig. 3. *Culicoides* molecular phylogenetic relationships and species boundaries using COI, 16S rDNA and 28S rDNA. Posterior probabilities and bootstrap values from the RAxML analysis are presented for the most important nodes (asterisks indicate PPP 0.95 or BSP 95; – indicate that the node was not recovered in the RAxML topology). [Colour figure can be viewed at wileyonlinelibrary.com].

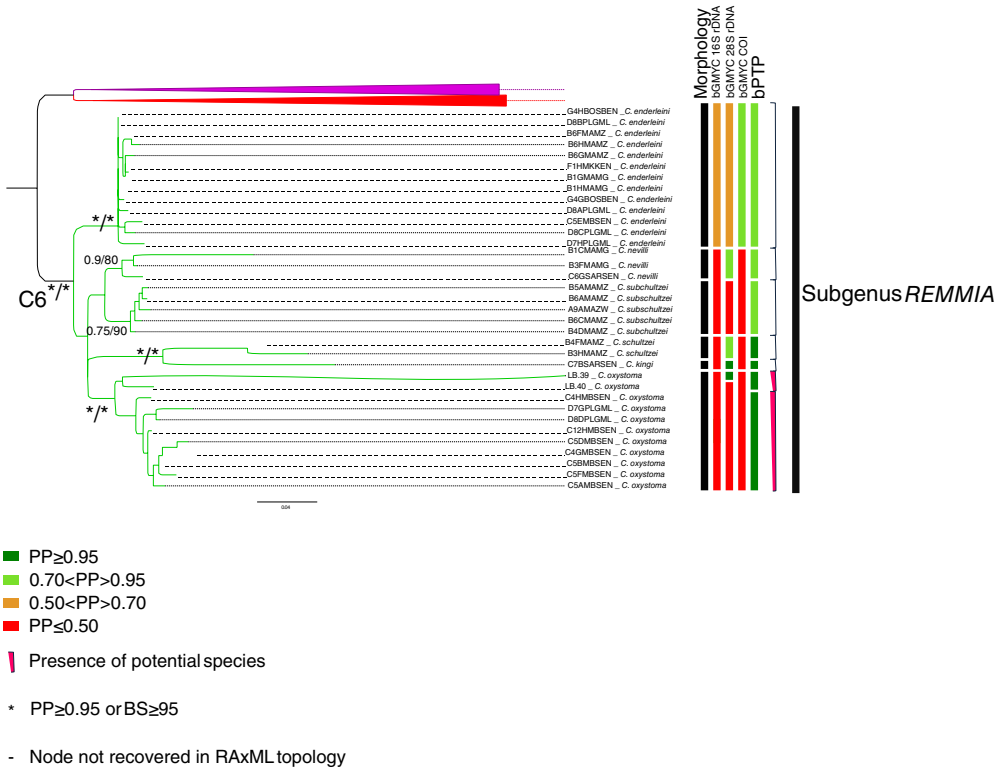


Fig. 3. Continued.

related to species within the Orientalis group. In the Afrotropical region, the Orientalis group includes *C. brosseti*, *C. dubitatus* and *C. trifasciellus* (Kremer *et al*., 1975; Meiswinkel, 2004). Relationships between *C. trifasciellus* and *C. sp*. #20 in phylogenetic trees constructed from *16S* rDNA and *28S* rDNA were strongly supported. The status of *C. sp*. #20 still needs to be clarified in future studies. The Dasyops group, as suggested by Meiswinkel (1987) with a redescription of *C. kanagai* Khamala and Kettle in 1971, is monophyletic and *C. kanagai* and *C. sp*. #54 dark form were well-supported. Species within the Dasyops group are typically Afrotropical. This group includes *C. alticola*; *C. kanagai*; *C. dasyops* and recently *C. sp*. #54 Meiswinkel (new species, not described) (Nevill *et al*., 2009; Labuschagne, 2016). The subgenus *Avaritia* was erected by Fox in 1955. Definition of *Avaritia* was completed later, based on adult morphology (Blanton & Wirth, 1979; Wirth & Hubert, 1989). Several species belonging to *Avaritia* are of considerable veterinary importance as vectors of important arboviruses such as African horse sickness virus (AHSV), Bluetongue virus (BTV) and Epizootic hemorrhagic disease virus (EHDV) (Venter *et al*., 1998, ; Meiswinkel *et al*., 2004).

Relationships between *C. austeni*, *C. isioloensis*, *C. milnei*, *C. moreli* and *C. zuluensis* were well-supported. These species belong to the Milnei group, as defined by Cornet *et al.* (1974). Our molecular analyses using the *CAD* gene and concatenated *COI*, *16S* rDNA and *28S* rDNA genes indicated that the Milnei group is monophyletic as reported previously in another study based on *COI* and *28S* rDNA sequences (Augot *et al*., 2017). Some species of the Milnei group are of medical and veterinary interest (Labuschagne, 2016), such as *C. austeni* which is suspected in the transmission of *Mansonella perstans* to humans, and *C. milnei* of BTV to livestock (Labuschagne, 2016). According to Borkent's classification (Borkent, 2016), species of the Milnei group are not in any subgenus, whereas their morphological characters are similar to the subgenus *Hoffmania*. Meiswinkel placed this group in the subgenus *Hoffmania* (Meiswinkel,). Future investigations will take into account the morphological characteristics and molecular analyses of the Milnei group and *C. insignis* Lutz, type species of the subgenus *Hoffmania*, in order to place this group within *Hoffmania* or to create a new subgenus for this monophyletic group.

Species relationships within the subgenus *Synhelea*, with *C. tropicalis* as type species, were revised by Meiswinkel and Dyce in 1989 based on morphological characters. They limited this subgenus to the Tropicalis group that includes *C. camicasi* Cornet and Chateau, *C. congolensis* Clastrier, *C. dispar* Clastrier, *C. dutoiti* de Meillon, *C. moucheti* Cornet and Kremer, *C. pellucidus* Khamala and Kettle, *C. perettii* Cornet and Chateau, *C. tauffliebi* Clastrier, *C. tropicalis* and *C. vicinus* Clastrier (Meiswinkel & Dyce, 1989). Regarding molecular analyses using the *CAD* gene, *C. tropicalis* is recovered as sister taxon to *C. similis*. Borkent's classification placed species from the Similis group in subgenus *Synhelea* (Borkent, 2016). Concatenated *COI*, *16S* rDNA and *28S* rDNA genes provide strong relationships between *C. tropicalis* and the Similis group

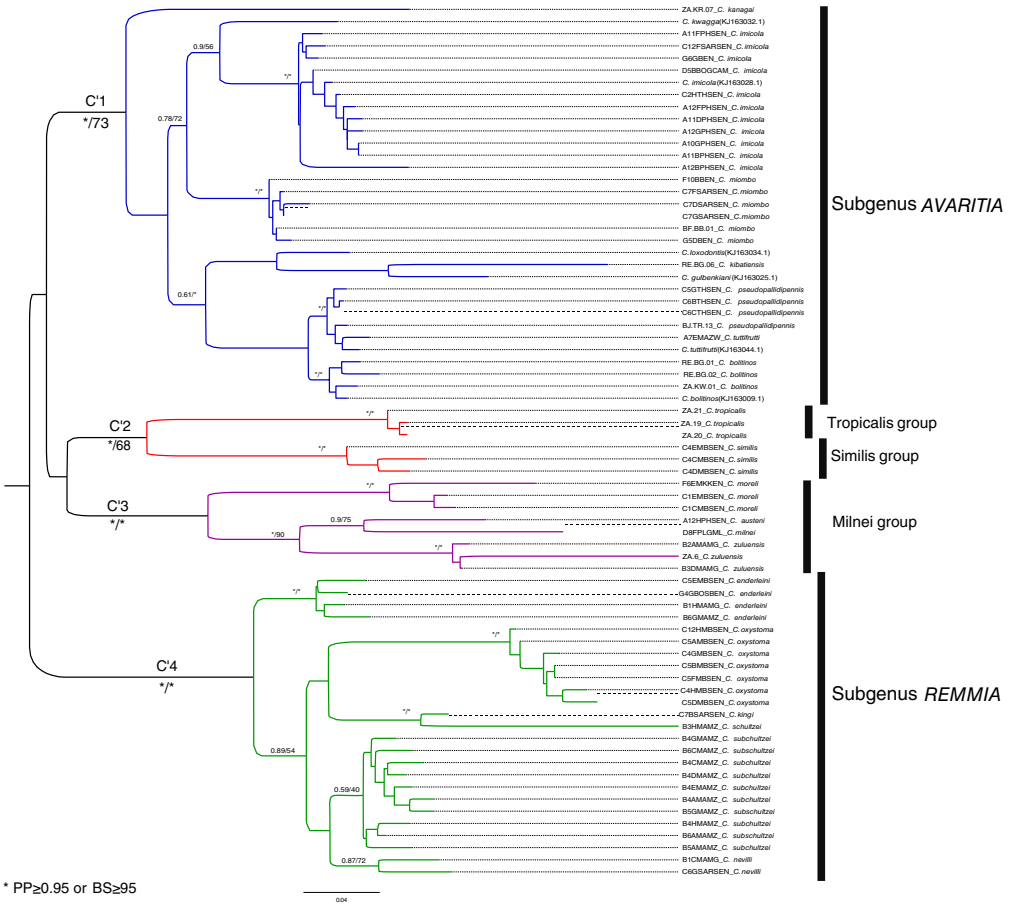


Fig. 4. *Culicoides* molecular phylogenetic relationships using CAD gene. Posterior probabilities and bootstrap values are presented for the most important nodes (asterisks indicate PPP 0.95 or BSP 95). [Colour figure can be viewed at wileyonlinelibrary.com].

(*C. exspectator*, *C. similis* and *C. ravus*). *Culicoides ravus* of the Similis group is more closely related to *C. neavei* and *C. ovalis* of the Neavei group. These two species are recovered as sister taxon with *C. ravus*. Thereby *C. tropicalis*, the Similis group (*C. exspectator*, *C. ravus* and *C. similis*) and the Neavei group (*C. neavei* and *C. ovalis*) are recovered as monophyletic with strong support (1.0/94). Based on our molecular analyses, Borkent's classification for the Similis group is maintained. And it is probable that the Neavei group, unplaced to subgenus (Borkent, 2016), belongs to subgenus *Synhelea*.

The subgenus *Remmia* is monophyletic with strong support (1.0/100) based on our molecular analyses using the *CAD* gene and concatenated *COI*, *16S* rDNA and *28S* rDNA genes. This subgenus, with *C. schultzei* as type species, includes the Schultzei group with species of veterinary interest, such as *C. kingi* involved in the transmission of *Onchocerca gutturosa*, a widespread parasite of Sudanese cattle (El Sinnary & Hussein, 1980) or *C. oxystoma*, potential vector of Akabane virus in Japan (Kurogi *et al*., 1987) and AHSV in Senegal (Fall *et al*., 2015; Bakhoum *et al*., 2016). Molecular delineation using bPTP and bGMYC methods for *C. oxystoma* from Lebanon and Senegal showed that *C. oxystoma* is a complex of sibling species, as previously noted by several authors (Cornet & Brunhes, 1994; Bakhoum *et al*., 2013; Harrup *et al*., 2016). In order to examine potential species within *C. oxystoma*, we suggest the use of Bayesian species delimitation implemented in Bayesian Phylogenetics and Phylogeography (BPP) (Rannala & Yang, 2003; Yang & Rannala, 2010) in future investigations. This approach generates the posterior probabilities of species assignments taking account of uncertainties due to unknown gene trees and the ancestral coalescent process (Toussaint *et al*., 2015).

We conclude that all species relationships within studied subgenera and groups were well-supported. However, we recorded a new species *C. sp*. #22 within the Imicola group which was revealed as monophyletic within the subgenus *Avaritia*. The Milnei group was regarded as monophyletic with strong support. Considered as monophyletic (Bakhoum *et al*., 2013; Augot *et al*., 2017), monophyly of the Schultzei group (subgenus *Remmia*) was confirmed with *C. oxystoma* as a potential complex of sibling species. In our study, all studied species of the Similis group were placed in the subgenus *Synhelea* in accordance with Borkent's classification. In addition, *C. neavei* and *C. ovalis* (Neavei group) were placed in the subgenus *Synhelea*.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: 10.1111/syen.12279

Figure S1. *Culicoides* molecular phylogenetic relationships using the *COI* gene. Posterior probabilities and bootstrap values are presented for the most important nodes (asterisks indicate PPP 0.95 or BSP 95).

Figure S2. *Culicoides* molecular phylogenetic relationships using the *16S* rDNA gene. Posterior probabilities and bootstrap values are presented for the most important nodes (asterisks indicate PPP 0.95 or BSP 95).

Figure S3. *Culicoides* molecular phylogenetic relationships using the *28S* rDNA gene. Posterior probabilities and bootstrap values are presented for the most important nodes (asterisks indicate PPP 0.95 or BSP 95).

Table S1. *Culicoides* specimen details, including *COI*, *16S* rDNA and *28S* rDNA sequences (-indicate missing sequence), GenBank accession numbers, and sample location information.

Acknowledgements

This study was funded by UMR 117 ASTRE, Cirad. M.T.B. was granted by the Islamic Development Bank. M.T.B. would like to thank Christine Chevillon, IRD, UR 224 MIVEGEC, Montpellier, France; and Nabil Haddad, Laboratory of Immunology and Vector Borne Diseases, Lebanese University. The authors are grateful to all the people who gave assistance in *Culicoides* collection. This work was supported by the high-performance cluster of Cirad.

M.T.B. and C.G. designed the study; K.L. and M.F. contributed the collection and identification of *Culicoides*; M.T.B., K.H., B.M. and C.G. analysed the data; K.L., B.M., K.H., G.V., A.G.F., T.B., L.G. G.G. and J.B. contributed to the manuscript after its first draft by M.T.B. and C.G.; all authors read and commented on the final manuscript version.

References

Agbolade, O.M., Akinboye, D.O., Olateju, T.M., Ayanbiyi, O.A., Kuloyo, O.O. & Fenuga, O.O. (2006) Biting of anthropophilic *Culicoides fulvithorax* (Diptera: Ceratopogonidae), a vector of *Mansonella perstans* in Nigeria. *The Korean Journal of Parasitology*, **44**, 67–72.

Augot, D., Mathieu, B., Hadj-Henni, L. *et al.* (2017) Molecular phylogeny of 42 species of *Culicoides* (Diptera, Ceratopogonidae) from three continents. *Parasite*, **24**, 23.

Bakhoum, M.T., Fall, M., Fall, A.G. *et al.* (2013) First record of *Culicoides oxystoma* Kieffer and diversity of species within the Schultzei group of *Culicoides* Latreille (Diptera: Ceratopogonidae) biting midges in Senegal. *PLoS ONE*, **8**, e84316.

Bakhoum, M.T., Fall, M., Seck, M.T. *et al.* (2016) Foraging range of arthropods with veterinary interest: new insights for Afrotropical *Culicoides* biting midges (Diptera: Ceratopogonidae) using the ring method. *Acta Tropica*, **157**, 59–67.

Bassene, H., Sambou, M., Fenollar, F. *et al.* (2015) High prevalence of *Mansonella perstans* filariasis in rural Senegal. *The American Journal of Tropical Medicine and Hygiene*, **93**, 601–606.

Bellis, G., Dyce, A., Gopurenko, D., Yanase, T., Garros, C., Labuschagne, K. & Mitchell, A. (2013) Revision of the *Culicoides* (*Avaritia*) Imicola complex Khamala & Kettle (Diptera: Ceratopogonidae) from the Australasian region. *Zootaxa*, **3768**, 401–427.

Blanton, F.S. & Wirth, W.W. (1979) The Sandflies (*Culicoides*) of Florida (Diptera: Ceratopogonidae). *Arthropods of Florida and Neighboring Land Areas* (ed. by H.V. Weems, Jr.), 10, pp. 201–204. Florida Department of Agriculture and Consumer Services, Division of Plant Industry, Tallahassee, Florida.

Boorman, J. (1989) *Culicoides* (Diptera : Ceratopogonidae) of the Arabian peninsula with notes on their medical and veterinary importance. *Fauna of Saudi Arabia*, **10**, 160–224.

Boorman, J. & Dipeolu, O.O. (1979) A taxonomic study of adult Nigerian *Culicoides* Latreille (Diptera: Ceratopogoaidae) species. *Occasional Publications of the Entomoplogical Society of Nigeria*, **22**, 1–121.

Borkent, A. (2016) World Species of Biting Midges (Diptera: Ceratopogonidae) [WWWdocument]. URL http://wwx.inhs.illinois.edu/files/4514/6410/0252/CeratopogonidaeCatalog.pdf [accessed on 16 May 2016].

Carpenter, S., Mellor, P.S., Fall, A.G., Garros, C. & Venter, G.J. (2017) African horse sickness virus: history, transmission, and current status. *Annual Review of Entomology*, **62**, 343–358.

Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, **17**, 540–552.

Colçao, T.F. (1946) Some *Culicoides* of the Transvaal. *Anais do Instituto de Medicina Tropical*, **2**, 235–266.

Cornet, M. & Brunhes, J. (1994) Révision des espèces de *Culicoides* apparentées à *C. shultzei* (Enderleini, 1908) dans la région Afro-tropicale (Diptera: Ceratopogonidae). *Bulletin de la Société entomologique de France*, **92**, 149–164.

Cornet, M. & Chateau, R. (1970) Les *Culicoides* de l'Ouest africain (2° note) Espèces apparentées à *C. similis* Carter, Ingrain et Macfie, 1920 (Diptera, Ceratopogonidae). *Cahiers ORSTOM, sêrie Entomologie mêdicale et Parasitologie*, **13**, 141–173.

Cornet, M., Nevill, E.M. & Walker, A.R. (1974) Note sur les *Culicoides* (Diptera, Ceratopogonidae) du groupe de *C. milnei* Austen, 1909, en Afrlque orientale et australe. *Cahiers ORSTOM, sêrie Entomologie mêdicale et Parasitologie*, **12**, 231–243.

Darriba, D., Taboada, G.L., Doallo, R. & Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

Debrah, L.B., Nausch, N., Opoku, V.S. *et al.* (2017) Epidemiology of *Mansonella perstans* in the middle belt of Ghana. *Parasites & Vectors*, **10**, 15.

Diouf, N.D., Etter, E., Lo, M.M., Lo, M. & Akakpo, A.J. (2012) Outbreaks of African horse sickness in Senegal, and methods of control of the 2007 epidemic. *The Veterinary Record*, **172**, 152.

Dowton, M. & Austin, A.D. (1998) Phylogenetic relationships among the microgastroid wasps (*Hymenoptera*: Braconidae): combined analysis of 16S and 28S rDNA genes and morphological data. *Molecular Phylogenetics and Evolution*, **10**, 354–366.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.

Ekrem, T., Willassen, E. & Stur, E. (2010) Phylogenetic utility of five genes for dipteran phylogeny: a test case in the Chironomidae leads to generic synonymies. *Molecular Phylogenetics and Evolution*, **57**, 561–571.

El Sinnary, K. & Hussein, H.S. (1980) *Culicoides kingi*, Austen: a vector of Onchocerca gutturosa (Neumann, 1910) in the Sudan. *Annals of Tropical Medicine and Parasitology*, **74**, 655–656.

Enderlein, G. (1908) Neue Ceratopogoninen aus Südafrika. *Denkschriften der Medicinish-Naturwissenschaftlichen Gesellschaft zu Jena*, **13**, 459–461.

Fall, M., Diarra, M., Fall, A.G. *et al.* (2015) *Culicoides* (Diptera: Ceratopogonidae) midges, the vectors of African horse sickness virus a host/vector contact study in the Niayes area of Senegal. *Parasites & Vectors*, **8**, 39.

Fiedler, O.G.H. (1951) The South Afican biting midges of the genus *Culicoides* (Cerutopogonid., Dipt.). *The Onderstepoort Journal of Veterinary Research*, **21**, 3–33.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Glick, J.I. (1990) *Culicoides* biting midges (Diptera: Ceratopogonidae) of Kenya. *Journal of Medical Entomology*, **27**, 85–195.

Gopurenko, D., Bellis, G.A., Yanase, T. *et al.* (2015) Integrative taxonomy to investigate species boundaries within *Culicoides* (Diptera: Ceratopogonidae): a case study using subgenus *Avaritia* from Australasia and Eastern Asia. *Veterinaria Italiana*, **51**, 345–378.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, **59**, 307–321.

Harrup, L.E., Bellis, G.A., Balenghien, T. & Garros, C. (2015) *Culicoides* Latreille (Diptera: Ceratopogonidae) taxonomy: current challenges and future directions. *Infection, Genetics and Evolution*, **30C**, 249–266.

Harrup, L.E., Laban, S., Purse, B.V. *et al.* (2016) DNA barcoding and surveillance sampling strategies for *Culicoides* biting midges (Diptera: Ceratopogonidae) in southern India. *Parasites & Vectors*, **9**, 461.

Itoua, A., Cornet, M., Vattier-Bemard, G. & Trouillet, J. (1987) The *Culicoides* (Diptera: Ceratopogonidae) of Central Africa. *Cahiers ORSTOM, sêrie Entomologie mêdicale et Parasitologie*, **25**, 127–134.

Khamala, C.P.M. & Kettle, D.S. (1971) The *Culicoides* Latreille (Diptera: Ceratopogonidae) of East Africa. *Transactions of the Royal Entomological Society of London*, **123**, 1–95.

Kremer, M., Rebholtz-Hirtzel, C. & Delecolle, J.C. (1975) Description d'une espèce nouvelle: *C. dubitatus* n. sp. (Diptera: Ceratopogonidae) de la Région Ethiopienne. *Cahiers ORSTOM, sêrie Entomologie mêdicale et Parasitologie*, **13**, 233–236.

Kurogi, H., Akiba, K., Inaba, Y. & Matumoto, M. (1987) Isolation of Akabane virus from the biting midge *Culicoides oxystoma* in Japan. *Veterinary Microbiology*, **15**, 243–248.

Labuschagne, K. (2016) *The* Culicoides *Latreille (Diptera: Ceratopogonidae) species of South Africa*. PhD Thesis, University of Pretoria, South Africa.

Lanfear, R., Calcott, B., Ho, S.Y. & Guindon, S. (2012) Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, **29**, 1695–1701.

Massey, S.E., Churbanov, A., Rastogi, S. & Liberles, D.A. (2008) Characterizing positive and negative selection and their phylogenetic effects. *Gene*, **418**, 22–26.

Mathieu, B. (2011) *Les espèces de* Culicoides *du sous-genre Avaritia (Diptera: Ceratopogonidae) dans le monde: révision systématique et taxonomique des espèces d'intérêt dans la transmission d'*Orbivirus. PhD Thesis, Université de Strasbourg, France.

Meiswinkel, R. (1987) Afrotropical Culicoides: a redescription of *C. (Avaritia) kanagai* Khamala and Kettle, 1971, reared from elephant dung in the Kruger National Park, South Africa. *Onderstepoort Journal of Veterinary Research*, **54**, 585–590

Meiswinkel, R. (1995) *Afrotropical Culicoides*: *biosystematics of the Imicola group, subgenus Avaritia (Diptera: Ceratopogonidae). With special reference to the epidemiology of African horse sickness*. MSc Thesis, University of Pretoria, South Africa.

Meiswinkel, R. (2004) Adult characters defining and separating the Imicola and Orientalis species complexes of the subgenus *Avaritia* Fox, 1955 (*Culicoides*, Diptera: Ceratopogonidae). *Veterinaria Italiana*, **40**, 345–351.

Meiswinkel, R. & Dyce, A. (1989) Afrotropical *Culicoides*: *Synhelea* Kieffer, 1925, resurrected as subgenus to embrace 10 species (Diptera: Ceratopogonidae). *The Onderstepoort Journal of Veterinary Research*, **56**, 147–163.

Meiswinkel, R., Venter, G.J. & Nevill, E.M. (2004) Vectors: *Culicoides* spp. *Infectious Diseases of Livestock* (ed. by J.A.W. Coetzer and R.C. Tustin), pp. 93–136. Oxford University Press, Cape Town.

Mellor, P.S., Boorman, J. & Baylis, M. (2000) *Culicoides* biting midges: their role as Arbovirus vectors. *Annual Review of Entomology*, **45**, 307–340.

Monaco, F., Benedetto, L., Di Marcello, V., Lelli, R. & Goffredo, M. (2010) Development and preliminary evaluation of a real-time polymerase chain reaction for the identification of *Culicoides obsoletus* sensu strictu, *C. scoticus* and *C. montanus* in the Obsoletus complex in Italy. *Veterinaria Italiana*, **46**, 209–214.

Moulton, J.K. & Wiegmann, B.M. (2004) Evolution and phylogenetic utility of CAD (rudimentary) among Mesozoic-aged Eremoneuran Diptera (Insecta). *Molecular Phylogenetics and Evolution*, **31**, 363–378.

Mullen, G.R. (2009) Biting Midges (Ceratopogonidae). *Medical and Veterinary Entomology* (ed. by G. Mullen and L. Durden), Chapter 12, pp. 169–188. Academic Press, San Diego, California.

Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K. & Pond, S.L.K. (2012) Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, **8**, e1002764.

Nevill, H., Venter, G.J., Meiswinkel, R. & Nevill, E.M. (2007) Comparative descriptions of the pupae of five species of the *Culicoides imicola* complex (Diptera, Ceratopogonidae) from South Africa. *The Onderstepoort Journal of Veterinary Research*, **74**, 97–114.

Nevill, H., Nevill, E.M. & Venter, G.J. (2009) Description and comparison of the pupae of a further two *Culicoides* (*Avaritia*) species from the dung of large herbivores in South Africa (Diptera: Ceratopogonidae). *The Onderstepoort Journal of Veterinary Research*, **76**, 277–284.

Nolan, D.V., Carpenter, S., Barber, J., Mellor, P.S., Dallas, J.F., Mordue Luntz, A.J. & Piertney, S.B. (2007) Rapid diagnostic PCR assays for members of the *Culicoides obsoletus* and *Culicoides pulicaris* species complexes, implicated vectors of bluetongue virus in Europe. *Veterinary Microbiology*, **124**, 82–94.

Pagès, N., Sarto, I. & Monteys, V. (2005) Differentiation of *Culicoides obsoletus* and *Culicoides scoticus* (Diptera: Ceratopogonidae) based on mitochondrial cytochrome oxidase subunit I. *Journal of Medical Entomology*, **42**, 1026–1034.

Pagès, N., Muñoz-Muñoz, F., Talavera, S., Sarto, V., Lorca, C. & Núñez, J.I. (2009) Identification of cryptic species of *Culicoides* (Diptera: Ceratopogonidae) in the subgenus *Culicoides* and development of species-specific PCR assays based on barcode regions. *Veterinary Parasitology*, **165**, 298–310.

Philippe, H., Lopez, P., Brinkmann, H. *et al.* (2000) Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. *Proceedings of the Royal Society of London Series B*, **267**, 1213–1221.

Pond, S.L. & Frost, S.D. (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics*, **21**, 2531–2533.

Pons, J., Barraclough, T., Gomez-Zurita, J. *et al.* (2006) Sequence-based species delimitation for the DNA taxonomy of Undescribed insects. *Systematic Biology*, **55**, 595–609.

Purse, B.V., Carpenter, S., Venter, G.J., Bellis, G. & Mullens, B.A. (2015) Bionomics of temperate and tropical *Culicoides* midges: knowledge gaps and consequences for transmission of *Culicoides*-borne viruses. *Annual Review of Entomology*, **60**, 373–392.

Rannala, B. & Yang, Z. (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics and Molecular Biology*, **164**, 1645–1656.

Ranwez, V., Harispe, S., Delsuc, F. & Douzery, E.J. (2011) MACSE: multiple alignment of coding SEquences accounting for frameshifts and stop codons. *PLoS ONE*, **6**, e22594.

Reid, N.M. & Carstens, B.C. (2012) Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evolutionary Biology*, **12**, 196.

Ronquist, F., Teslenko, M., van der Mark, P. *et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Scacchia, M., Lelli, R., Peccio, A. *et al.* (2009) African horse sickness: a description of outbreaks in Namibia. *Veterinaria Italiana*, **45**, 265–274.

Sebastiani, F., Meiswinkel, R., Gomulski, L.M., Guglielmino, C.R., Mellor, P.S., Malacrida, A.R. & Gasperi, G. (2001) Molecular differentiation of the old world *Culicoides imicola* species complex (Diptera, Ceratopogonidae) inferred using random amplified polymorphic DNA markers. *Molecular Ecology*, **10**, 1773–1786.

Sikes, D.S. & Venables, C. (2013) Molecular phylogeny of the burying beetles (Coleoptera: Silphidae: Nicrophorinae). *Molecular Phylogenetics and Evolution*, **69**, 552–565.

Simonsen, P.E., Onapa, A.W. & Asio, S.M. (2011) *Mansonella perstans* filariasis in Africa. *Acta Tropica*, **120**, 109–120.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.

Szadziewski, R., Filatov, S. & Dominiak, P. (2016) A redescription of *Culicoides griseidorsum* Kieffer, 1918, with comments on subgeneric position of some European taxa (Diptera: Ceratopogonidae). *Zootaxa*, **3**, 413–422.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, **30**, 2725–2729.

Toussaint, E.F., Moriniere, J., Muller, C.J., Kunte, K., Turlin, B., Hausmann, A. & Balke, M. (2015) Comparative molecular species delimitation in the charismatic Nawab butterflies (Nymphalidae, Charaxinae, Polyura). *Molecular Phylogenetics and Evolution*, **91**, 194–209.

Townsend, J.P. (2007) Profiling phylogenetic informativeness. *Systematic Biology*, **56**, 222–231.

Venter, G.J., Paweska, J.T., Van Dijk, A.A., Mellor, P.S. & Tabachnick, W.J. (1998) Vector competence of *Culicoides bolitinos* and *C. imicola* (Diptera: Ceratopogonidae) for South African bluetongue virus serotypes 1, 3 and 4. *Medical and Veterinary Entomology*, **12**, 378–385.

Venter, G.J., Koekemoer, J.J.O. & Paweska, J.T. (2006) Investigations on outbreaks of African horse sickness in the surveillance zone in South Africa. *Revue scientifique et Technique (International Office of Epizootics)*, **25**, 1097–1109.

Whiting, M.F., Carpenter, J.C., Wheeler, Q.D. & Wheeler, W.C. (1997) The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. *Systematic Biology*, **46**, 1–68.

Wirth, W.W. & Hubert, A.A. (1989) *The Culicoides of Southeast Asia (Diptera: Ceratopogonidae)*. Memoirs of the American Entomological Institute, Gainesville, Florida.

Xia, X. & Xie, Z. (2001) Data analysis in molecular biology and evolution. *The Journal of Heredity*, **92**, 371–373.

Xia, X., Xie, Z., Salemi, M., Chen, L. & Wang, Y. (2003) An index of substitution saturation and its application. *Molecular Biology and Evolution*, **26**, 1–7.

Yang, Z. & Rannala, B. (2010) Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 9264–9269.

Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. (2013) A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, **29**, 2869–2876.

Accepted 6 October 2017
First published online 16 November 2017

