Genomic insights into adaptive divergence and speciation among malaria vectors of the *Anopheles nili* group

Caroline Fouet[1]*, Colince Kamdem[1], Stephanie Gamez[1], Bradley J. White[1,2]*

[1]Department of Entomology, University of California, Riverside, CA 92521

[2]Center for Disease Vector Research, Institute for Integrative Genome Biology, University of California, Riverside, CA 92521

*Corresponding authors: caroline.fouet@ucr.edu; bwhite@ucr.edu

Abstract

Ongoing speciation in most African malaria vectors gives rise to cryptic populations, which differ remarkably in their behaviour, ecology and capacity to vector malaria parasites. Understanding the population structure and the drivers of genetic differentiation among mosquitoes is critical for effective disease control because heterogeneity within species contribute to variability in malaria cases and allow fractions of vector populations to escape control efforts. To examine the population structure and the potential impacts of recent large-scale control interventions, we have investigated the genomic patterns of differentiation in mosquitoes belonging to the *Anopheles nili* group, a large taxonomic group that diverged ~3-Myr ago. Using 4343 single nucleotide polymorphisms (SNPs), we detected strong population structure characterized by high F_{ST} values between multiple divergent populations adapted to different habitats within the Central African rainforest. Delineating the cryptic species within the *Anopheles nili* group is challenging due to incongruence between morphology, ribosomal DNA and SNP markers consistent with incomplete lineage sorting and/or interspecific gene flow. A very high proportion of loci are fixed (F_{ST} = 1) within the genome of putative species, which suggests that ecological and/or reproductive barriers are maintained by strong selection on a substantial number of genes.

Key words: *Anopheles nili*, divergent selection, high-F_{ST} regions, speciation.

31 **1. Introduction**

32 One of the principal goals of population genetics is to summarize the genetic
33 similarities and differences between populations (Wright, 1984). This task can be relatively
34 straightforward for some taxa, but the genetic relationship among populations can also be
35 difficult to summarize, especially for species whose evolutionary history is complex and
36 reticulate. The best known mosquito species of the genus *Anopheles* — which includes all
37 vectors of *Plasmodium*, the parasite of human malaria — exhibit very complex range-wide
38 population structure due to the combined effects of cryptic speciation, adaptive flexibility
39 and ongoing gene flow (Harbach, 2013; Krzywinski and Besansky, 2003). For example,
40 almost all the major malaria vectors of the Afrotropical region belong to large taxonomic
41 groups encompassing multiple incipient species relatively isolated reproductively and
42 geographically from one another (reviewed by Sinka et al. 2010; Antonio-Nkondjio and
43 Simard 2013; Coetzee and Koekemoer 2013; Dia et al. 2013; Lanzaro and Lee 2013). These
44 characteristics make them promising model systems to study speciation and the processes
45 which contribute to reproductive barriers (e.g. Turner et al. 2005; Lawniczak et al. 2010;
46 Neafsey et al. 2010; Fontaine et al. 2015; Weng et al. 2016), but can also have far-reaching
47 practical consequences. Both spatial and temporal variability in malaria cases and the
48 effectiveness of vector control measures are greatly impacted by heterogeneity within
49 vector species (Molineaux and Gramiccia 1980; Van Bortel et al. 2001). For these reasons,
50 research on the genetic structure among the major African malaria vector mosquitoes has
51 intensified over the last few decades (Antonio-Nkondjio and Simard, 2013; Coetzee and
52 Koekemoer, 2013; Dia et al., 2013; Lanzaro and Lee, 2013).

53 The recent scaling up of insecticide-treated nets usage and indoor insecticide
54 spraying to a lesser extent have led to a dramatic reduction in malaria morbidity across the
55 continent (WHO, 2016). However, the other implications of these large-scale interventions
56 include increased insecticide resistance (reviewed by Hemingway et al. 2016; Ranson and
57 Lissenden 2016), range shift (e.g. Bøgh et al. 1998; Derua et al. 2012; Mwangangi et al.
58 2013) and profound evolutionary changes among vector populations. In contrast to
59 insecticide resistance and range shift, which have been extensively studied, the recent
60 adaptive changes among mosquito populations have yet to be addressed significantly.
61 These changes — which involve local adaptation and genetic differentiation, introgressive
62 hybridization and selective sweeps across loci conferring resistance to xenobiotics — are
63 particularly evident in the most anthropophilic species (Barnes et al., 2017; Clarkson et al.,
64 2014; Kamdem et al. , 2017; Norris et al., 2015).

65 The ecology, taxonomic complexity, geographic distribution, role in transmission
66 and evolutionary potential of each major African malaria vector are unique. Consequently,
67 further research is needed to specifically resolve the fine-scale population structure and the
68 genomic targets of natural selection in all of the important taxa including currently
69 understudied species. The present work focused on a group of malaria vector species
70 representing a large taxonomic unit named *Anopheles nili* group. Despite the significant role
71 of some of its species in sustaining high malaria transmission, this group has received little
72 attention. To date, four species that occur in forested areas of Central and West Africa and
73 are distinguishable by slight morphological variations are known within the *An. nili* group:
74 *An. nili sensu stricto* (hereafter *An. nili*), *An. ovengensis*, *An. carnevalei*, and *An. somalicus*
75 (Awono-Ambene et al., 2004; Gillies and Coetzee, 1987; Gillies and De Meillon, 1968). These

species are characterized by reticulate evolution and complex phylogenies that have been challenging to resolve so far (Awono-Ambene et al., 2004; Awono-Ambene et al., 2006; Kengne et al., 2003; Ndo et al., 2010, 2013; Peery et al., 2011; Sharakhova et al., 2013). Populations of *An. nili* and *An. ovengensis* are very anthropophilic and efficient vectors of *Plasmodium* in rural areas where malaria prevalence is particularly high (Antonio-Nkondjio et al., 2006).

To delineate genomic patterns of differentiation, we sampled mosquito populations throughout the range of species of the *An. nili* group in Cameroon and used reduced representation sequencing to develop genome-wide SNP markers that we genotyped in 145 individuals. We discovered previously unknown subpopulations characterized by high pairwise differentiation within *An. ovengensis* and *An. nili*. We further explored the genetic differentiation across the genome and revealed the presence of a very high number of outlier loci that are targets of selection among locally adapted subpopulations. These findings provide significant baseline data on the genetic underpinnings of adaptive divergence and pave the way for further genomic studies in this important group of mosquitoes. Notably, a complete reference genome will allow in-depth studies to decipher the functional and phenotypic characteristics of the numerous differentiated loci as well as the contribution of recent selective events in ongoing adaptation.

2. Materials and methods

(a) Mosquito species

We surveyed 28 locations within the geographic ranges of species of the *An. nili* group previously described in Cameroon (Figure 1) (Antonio-Nkondjio et al., 2009; Awono-

99 Ambene et al., 2004; Awono-Ambene et al., 2006; Ndo et al., 2010, 2013). The genetic
100 structure of *Anopheles* species is most often based on macrogeographic or regional
101 subdivisions of gene pools, but can also involve more subtle divergence between larvae and
102 adults, or between adult populations found in or around human dwellings (e.g. Riehle et al.,
103 2011). To effectively estimate the genetic diversity and identify potential cryptic
104 populations within species, we collected larvae and adult mosquitoes within and around
105 human dwellings using several sampling techniques (Service, 1993) in September-October
106 2013 (Table S1). To identify the four currently known members of the *An. nili.* group, we
107 used morphological keys and a diagnostic PCR, which discriminates species on the basis of
108 point mutations of the ribosomal DNA (Awono-Ambene et al., 2004; Gillies and Coetzee,
109 1987; Gillies and De Meillon, 1968; Kengne et al., 2003).

110 **(b) Library preparation, sequencing and SNP discovery**

111 We created double-digest Restriction site Associated DNA (ddRAD) libraries as
112 described in Kamdem et al. 2017 using a modified version of the protocol designed by
113 Peterson et al*.* 2012. Briefly, genomic DNA of mosquitoes was extracted using the DNeasy
114 Blood and Tissue kit (Qiagen) and the Zymo Research MinPrep kit for larvae and adult
115 samples respectively. Approximately 50ng (10µl) of DNA of each mosquito was digested
116 simultaneously with *MluC1* and *NlaIII* restriction enzymes. Digested products were ligated
117 to adapter and barcode sequences enabling identification of individuals. Samples were
118 pooled, purified, and 400-bp fragments selected. The resulting libraries were amplified via
119 PCR and purified, and fragment size distribution was checked using the BioAnalyzer. PCR
120 products were quantified, diluted and single-end sequenced to 100 base reads on Illumina
121 HiSeq2000.

122

123 **(b) SNP discovery and genotyping**

124 The *process_radtags* program of the Stacks v 1.35 pipeline (Catchen et al., 2013;
125 Catchen et al., 2011) was used to demultiplex and clean Illumina sequences. Reads that
126 passed quality filters were aligned to the *An. nili* Dinderesso draft genome assembly
127 (Giraldo-Calderon et al., 2015) made up of 51048 short contigs (~200-30512bp long) using
128 Gsnap (Wu and Nacu, 2010). To identity and call SNPs within consensus RAD loci, we
129 utilized the *ref_map.pl* program of Stacks. We set the minimum number of reads required to
130 form a stack to three and allowed two mismatches during catalogue creation. We generated
131 SNP files in different formats for further downstream analyses using the *populations*
132 program of Stacks and Plink v1.09 (Purcell et al., 2007).

133 **(d) Population genomics analyses**

134 We analyzed the genetic structure of *An. nili* sensu lato (s.l.) populations using a
135 Principal Component Analysis (PCA) and an unrooted Neighbor-Joining tree (NJ). We also
136 examined ancestry proportions and admixtures between populations in Admixture v1.23
137 (Alexander et al., 2009) and Structure v2.3.4 (Pritchard et al., 2000). We used the package
138 *adegenet* (Jombart, 2008) to implement the PCA in R (R Development Core Team 2016).
139 The NJ tree was generated from SNP allele frequencies via a matrix of Euclidian distance
140 using the R package *ape* (Paradis et al., 2004). We ran Admixture with 10-fold cross-
141 validation for values of k from 1 through 8. Similarly, we analyzed patterns of ancestry from
142 k ancestral populations in Structure, testing five replicates of k = 1-8. We used 200000
143 iterations and discarded the first 50000 iterations as burn-in for each Structure run.
144 Clumpp v1.1.2 (Jakobsson and Rosenberg, 2007) was used to summarize assignment

145 results across independent runs. To identify the optimal number of genetic clusters in our
146 sample, we applied simultaneously the lowest cross-validation error in Admixture, the ad-
147 hoc statistic deltaK (Earl and VonHoldt, 2012; Evanno et al., 2005) and the Discriminant
148 Analysis of Principal Component (DAPC) method implemented in *adegenet.* To examine the
149 level of genomic divergence among populations, we assessed genetic differentiation (F_{ST})
150 across SNPs using the *populations* program of the Stacks pipeline. Mean F_{ST} values were also
151 used to quantify pairwise divergence between populations. To infer the demographic
152 history of different populations, we used the diffusion approximation method implemented
153 in the package ∂a∂i v 1.6.3 (Gutenkunst et al., 2009). Single-population models were fitted
154 to allele frequency spectra, and the best model was selected using the lowest likelihood and
155 Akaike Information Criterion as well as visual inspections of residuals.
156

157 **3. Results**

158 **(a) SNP genotyping**

159 We collected mosquitoes from four locations out of 28 sampling sites (Figure 1,
160 Table S1) and sequenced 145 individuals belonging, according to morphological criteria
161 and diagnostic PCRs, to two species (*An. nili* (n = 24) and *An. ovengensis* (n = 121)). We
162 assembled 197724 RAD loci that mapped to unique positions throughout the reference
163 genome. After applying stringent filtering rules, we retained 408 loci present in all
164 populations and in at least 50% of individuals in each population. Within these loci, we
165 identified 4343 high-quality biallelic markers that were used to analyze population
166 structure and genetic differentiation.

167 **(b) Morphologically defined species do not correspond to genetic clusters**

168 PCA and a NJ tree show that the genetic variation across 4343 SNPs is best
169 explained by more than two clusters, implying subdivisions within *An. nili* and *An.*
170 *ovengensis* (Figure 2). Three subgroups are apparent within *An. nili* while two distinct
171 clusters segregate in *An. ovengensis*. These five subpopulations are associated with the
172 different sampling sites suggesting local adaptation of divergent populations. Importantly,
173 Structure and Admixture analyses reveal that, at k = 2, one population identified by
174 morphology and the diagnostic PCR as *An. nili* has almost the same ancestry pattern as the
175 largest *An. ovengensis* cluster (Figure 3). Such discrepancies between morphology-based
176 and molecular taxonomies can be due to a variety of processes including phenotypic
177 plasticity, introgressive hybrydization or incomplete lineage sorting (i.e., when independent
178 loci have different genealogies by chance) (Arnold, 1997; Combosch and Vollmer, 2015;
179 Fontaine et al., 2015; Weng et al., 2016). At k = 2 and k = 3, some populations also exhibit
180 half ancestry from each morphological species suggestive of gene flow. We found a
181 conflicting number of genetic clusters in our samples likely reflecting the complex history of
182 subdivisions and admixtures among populations (Figure 4). The Evanno et al., (2005)
183 method, which highlights the early states of divergence between *An. nili* and *An. ovengensis*
184 indicates two probable ancestors. DAPC and the Admixture cross-validation error, which
185 are more sensitive to recent hierarchical population subdivisions, show five or more
186 distinct clusters as revealed by PCA and the NJ tree (Figure 4).

187 As suggested by the long internal branches, which connect subpopulations on the NJ
188 tree, there is strong differentiation between and within morphological species
189 characterized by globally high F_{ST} values (Table 1). Relatively lower F_{ST} values observed
190 between certain clusters may be due to greater inter-population migration and intermixing

191 or more recent divergence. The F_{ST} values do not reflect the morphological delimitation of
192 species, and the level of genetic differentiation is even higher between some subpopulations
193 within the same morphological species. Overall, patterns of genetic structure and
194 differentiation reveal a group of populations whose phylogenies and species status are
195 likely confounded by hybridization and/or incomplete lineage sorting. We argue that the
196 current taxonomy based on morphology and ribosomal DNA does not capture the optimal
197 reproductive units among populations of this group of mosquitoes.

198 **(c) Genomic signatures of divergent selection and demographic history**

199 We analyzed patterns of genetic differentiation across SNP loci throughout the
200 genome. Pairwise comparisons are based on filtered variants that satisfy all criteria to be
201 present in both populations, which explains the discrepancy in the number of SNPs
202 observed between specific paired comparisons (Figure 5). The distribution of locus-specific
203 F_{ST} values between the five subpopulations revealed a U-shape characterized by two peaks
204 around 0 and 1. The great majority of SNPs have low to moderate divergence, but a
205 substantial number of variants are extremely differentiated between populations. The
206 maximum F_{ST} among SNPs is 1 and the proportion of loci with F_{ST} = 1 varies from 6.52%
207 between the populations we termed *An. nili* group 1 and *An. ovengensis* group 1 to 44.74%
208 between the subgroups called *An. nili* group 2 and *An. nili* group 3 (Figure 5). This pattern of
209 genome-wide divergence suggests that a very high number of sites with abrupt
210 differentiation — which likely contain genes that contribute to divergent selection and/or
211 reproductive isolation — coexist with regions of weak divergence that can be freely
212 exchanged between species. As is the case with the overall genetic differentiation,
213 morphology is not a reliable predictor of locus-specific divergence. Precisely, the lowest

214 percentage of fixed SNPs is found between *An. ovengensis* from Nyabessan and *An. nili*
215 collected from Mbébé and Ebebda (Figure 1, Figure 5). In contrast, the greatest proportion
216 of fixed loci is observed between locally adapted subgroups within the same morphological
217 species, *An. nili*. The draft reference genome made up of short contigs did not allow us to
218 test hypotheses about the genomic distribution of differentiated loci. For example, it
219 remains unknown whether the numerous SNPs that are fixed among populations are
220 spread throughout the entire genome or clustered in genomic regions of low recombination
221 including chromosomal inversions and chromosome centers (Nosil and Feder, 2012; Roesti
222 et al., 2012).

223 Models of population demography indicate that all subgroups have experienced an
224 increase in effective size in a more or less recent past (Table 2). Nevertheless, confidence
225 intervals of population parameters are high in some populations, and our results should be
226 interpreted with the necessary precautions. The population growth is less significant in *An.*
227 *nili* group 1.

228

229 **4. Discussion**

230 **(a) Genetic differentiation**

231 Advances in sequencing and analytical approaches have opened new avenues for the
232 study of genomes of disease vectors. We have focused on malaria mosquitoes of the *An. nili*
233 group, whose taxonomy and population structure have been challenging to resolve with
234 low-resolution markers. We analysed the genetic structure using genome-wide SNPs and
235 found strong differentiation and local adaption among populations belonging to the two
236 morphologically defined species *An. nili* and *An. ovengensis*. The exact number of

237 subpopulations remains contentious, with the suggested number of divergent clusters
238 varying from two to five within both species. Significant population structure at eight
239 microsatellite loci has been described among *An. nili* populations from Cameroon, with F_{ST}
240 values as high as 0.48 between samples from the rainforest area (Ndo et al., 2013). By
241 contrast, *An. ovengensis* was discovered recently and the genetic structure of this vector
242 remains understudied. This species was initially considered as a sibling of *An. nili* (Awono-
243 Ambene et al., 2004; Awono-Ambene et al., 2006; Kengne et al., 2003), but more recent
244 studies have started to challenge the assumed relatedness between the two species due to
245 the high divergence of their polytene chromosomes (Sharakhova et al., 2013). Our findings
246 call for a careful review of the current taxonomy within this group of species, which is a
247 necessary first step for accurately delineating the role played by the different
248 subpopulations in malaria transmission.

249 Our samples were collected from locations characterized by a more or less degraded
250 forest in the rainforest area of Cameroon. In these habitats, larvae of *An. nili* s.l. exploit
251 relatively similar breeding sites consisting of slow-moving rivers (Antonio-Nkondjio et al.,
252 2009). The ecological drivers of genetic differentiation remain unknown, and will be
253 difficult to infer from our data given the apparent similarity of habitats among populations.
254 Further study is needed to clearly address the environmental variables that may be
255 correlated with ongoing adaptive divergence at adult and larval stages. One of the most
256 expected outcomes of current large-scale malaria control measures that are underway in
257 Sub-Saharan African countries concerns the effects of increased insecticide exposure on the
258 genetic diversity and population demography of vectors. A substantial population decline
259 that may considerably affect the adaptive potential of vector species has been occasionally

260 reported following the distribution of insecticide-treated bed nets and/or indoor residual
261 house spraying (e.g. Athrey et al., 2012). The inferred demographic history of the different
262 subpopulations within the *An. nili* group does not reveal signatures of bottlenecks that can
263 be potentially correlated with increased usage of insecticides and insecticide-treated nets.
264 This result is consistent with the demography of several other important malaria vectors of
265 the Afrotropical region, including *An. gambiae*, *An. coluzzii*, *An. funestus* and *An. moucheti*,
266 which reveals a substantial population increase suggesting that recent intense insecticide
267 exposure has yet to leave deep or detectable impacts on patterns of genetic variation among
268 mosquito populations (Fouet et al., 2017; Kamdem et al., 2017; O'Loughlin et al., 2014).

269 **(b) Genomic architecture of geographic and reproductive isolation**

270 Understanding the genomic architecture of reproductive isolation may reveal crucial
271 information on the sequence of events that occur between the initial stages of divergence
272 among populations and the onset of strong reproductive barriers (e.g. Turner et al., 2005;
273 Harr 2006; Nadeau et al., 2012; Ellegren et al., 2012; Carneiro et al., 2014; Burri et al.,
274 2015). One influential concept of speciation coined the "genic view of species" proposes
275 that boundaries between species are properties of individual genes or genome regions and
276 not of whole organisms or lineages (Barton and Hewitt, 1985; Harrison and Larson, 2014;
277 Harrison, 1990; Key, 1968; Nosil and Feder, 2012; Rieseberg et al., 1999; Wu, 2001). We
278 have discovered a dramatically high number of SNPs that are strongly differentiated
279 between populations and often fixed within subgroups of *An. nili* s.l. Interpreting this
280 intriguing pattern of genomic differentiation is not straightforward due to the complex
281 interactions between numerous forces — including positive or negative selection,
282 recombination, introgressive hybridization and incomplete lineage sorting — that can affect

283 the level of divergence among SNPs (Begun and Aquadro, 1992; Cutter and Payseur, 2013;
284 Harrison and Larson, 2016; Nachman and Payseur, 2012; Roesti et al., 2012). Some of these
285 variants exhibiting high divergence among populations certainly contain markers of
286 ecological and/or reproductive isolation. However, as far as reproductive barriers are
287 concerned, recent studies have indicated a complex relationship between genetic
288 differentiation and gene flow at the genome level (e.g. Gompert et al., 2012; Hamilton et al.,
289 2013a, b; Larson et al., 2013, 2014; Parchman et al., 2013; Taylor et al., 2014). Highly
290 divergent genomic regions do not necessarily coincide with regions of reduced gene flow.
291 Several alternative interpretations exist for the numerous high-F_{ST} regions we detected in
292 all pairwise comparisons (Cruickshank and Hahn, 2014; Delmore et al., 2015; Nachman and
293 Payseur, 2012; Noor and Bennett, 2009). Nevertheless, careful examination of these outliers
294 of differentiation may reveal significant insights into the wide range of genes and traits that
295 contribute to ecological divergence and/or reproductive isolation between subgroups of *An.*
296 *nili* s.l. A complete genome assembly will be necessary to better delineate specific regions of
297 the genome under natural selection, and thereby clarify the genomic basis of phenotypic
298 fitness differences between divergent populations. This will also help understand the extent
299 to which recent selection associated with human interventions contribute to local
300 adaptation and genetic differentiation as observed in *An. gambiae* and *An. coluzzii* (Kamdem
301 et al., 2017).

302 Signals consistent with gene flow between *An. nili* and *An. ovengensis* are apparent
303 in our data despite significant time since divergence (~3M-yr) (Ndo et al., 2013). Some
304 individuals display almost half ancestry from each morphological species. The
305 disagreement between morphology/PCR and molecular taxonomies observed in Structure

306 and Admixture analyses also suggests that incongruent genealogies may be widespread
307 along chromosomes due to hybridization. However, hybridization can be difficult to detect
308 because other factors such as incomplete lineage sorting or technical artefacts can leave
309 signatures that are similar to those of interspecific gene flow (Liu et al., 2014; Patterson et
310 al., 2012). A complete reference genome is also needed to analyze the detailed distribution
311 of genealogies across small genomic windows and to disentangle the relative contribution
312 of processes that generate the putative admixtures and species confusion observed among
313 divergent populations (Fontaine et al., 2015; Martin et al., 2013; Weng et al., 2016).
314

315 **5. Conclusions and implications**

316 Delineating the fine-scale population structure of mosquito populations is crucial for
317 understanding their epidemiological significance and their potential response to vector
318 control measures. Moreover, recent malaria control efforts are affecting interspecific gene
319 flow, genetic differentiation, population demography and natural selection in mosquitoes
320 (Athrey et al., 2012; Barnes et al., 2017; Clarkson et al., 2014; Kamdem et al., 2017; Norris et
321 al., 2015). Deciphering the signatures of these processes across mosquito genomes is
322 important to minimize their negative impacts on vector control. Our findings shed some
323 light on the complex evolutionary history and provide a framework for future
324 investigations into the genetic basis of ecological and reproductive barriers among species
325 of the *An. nili* group.
326
327

Acknowledgements

Funding for this project was provided by the University of California Riverside and NIH grants 1R01AI113248 and 1R21AI115271 to BJW. We thank inhabitants and administrative authorities of the sampling sites included in this study for their collaboration.

Author contributions

Conceived and designed the experiments: CF CK BJW. Performed the experiments: CF CK SG BJW. Analysed the data: CF CK. Wrote the paper: CF CK BJW.

Data Archiving Statement

Data for this study are available at: to be completed after manuscript is accepted for publication

References

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664.

Antonio-Nkondjio, C., Kerah, C. H., Simard, F., Awono-Ambene, P., Chouaibou, M., Tchuinkam, T., & Fontenille, D. (2006). Complexity of the malaria vectorial system in Cameroon: contribution of secondary vectors to malaria transmission. *Journal of Medical Entomology*, *43*, 1215–1221. https://doi.org/10.1603/0022-2585(2006)43[1215:COTMVS]2.0.CO;2

Antonio-Nkondjio, C., Ndo, C., Costantini, C., Awono-Ambene, P., Fontenille, D., & Simard, F. (2009). Distribution and larval habitat characterization of Anopheles moucheti, Anopheles nili, and other malaria vectors in river networks of southern Cameroon. *Acta Tropica*, *112*(3), 270–276. https://doi.org/10.1016/j.actatropica.2009.08.009

Antonio-Nkondjio, C., & Simard, F. (2013). Highlights on Anopheles nili and Anopheles moucheti, Malaria Vectors in Africa. In S. Manguin (Ed.), *Anopheles mosquitoes - New insights into malaria vectors* (INTECH).

Arnold, M. L. (1997). *Natural Hybridization and Evolution*. *Oxford University Press, Oxford*.

Athrey, G., Hodges, T. K., Reddy, M. R., Overgaard, H. J., Matias, A., Ridl, F. C., … Slotman, M. a. (2012). The effective population size of malaria mosquitoes: large impact of vector control. *PLoS Genetics*, *8*(12), e1003097. https://doi.org/10.1371/journal.pgen.1003097

Awono-Ambene, H. P., Kengne, P., Simard, F., Antonio-Nkondjio, C., & Fontenille, D. (2004). Description and bionomics of Anopheles (Cellia) ovengensis (Diptera: Culicidae), a new malaria vector species of the Anopheles nili group from south Cameroon. *Journal of Medical Entomology*, *41*, 561–568. https://doi.org/10.1603/0022-2585-41.4.561

Awono-Ambene, H. P., Simard, F., Antonio-Nkondjio, C., Cohuet, A., Kengne, P., & Fontenille, D. (2006). Multilocus enzyme electrophoresis supports speciation within the Anopheles nili group of malaria vectors in Cameroon. *The American Journal of Tropical Medicine and Hygiene*, *75*, 656–658. https://doi.org/75/4/656 [pii]

Barnes, K. G., Weedall, G. D., Ndula, M., Irving, H., Mzihalowa, T., Hemingway, J., & Wondji, C.

S. (2017). Genomic Footprints of Selective Sweeps from Metabolic Resistance to Pyrethroids in African Malaria Vectors Are Driven by Scale up of Insecticide-Based Vector Control. *PLoS Genetics*, *13(2): e10*, 1–22. https://doi.org/10.1371/journal.pgen.1006539

Barton, N., & Hewitt, G. . (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, *16*, 113–148.

Begun, D., & Aquadro, C. (1992). Levels of naturally occurring DNA polymorphism correlate with recombination rates in D. melanogaster. *Nature*, *356*, 519–520.

Bøgh, C., Pedersen, E. M., Mukoko, D. a, & Ouma, J. H. (1998). Permethrin-impregnated bednet effects on resting and feeding behaviour of lymphatic filariasis vector mosquitoes in Kenya. *Medical and Veterinary Entomology*, *12*(1), 52–9.

Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., … Ellegren, H. (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Research*, *25*(11), 1656–1665. https://doi.org/10.1101/gr.196485.115

Carneiro, M., Albert, F. W., Afonso, S., Pereira, R. J., Burbano, H., Campos, R., … Ferrand, N. (2014). The genomic architecture of population divergence between subspecies of the European Rabbit. *PLoS Genetics*, *10*(8). https://doi.org/10.1371/journal.pgen.1003519

Catchen, J., Hohenlohe, P. a., Bassham, S., Amores, A., & Cresko, W. . (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140. https://doi.org/10.1111/mec.12354

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: building and genotyping Loci de novo from short-read sequences. *G3 (Bethesda, Md.)*, *1*(3), 171–82. https://doi.org/10.1534/g3.111.000240

Clarkson, C. S., Weetman, D., Essandoh, J., Yawson, A. E., Maslen, G., Manske, M., … Donnelly, M. J. (2014). Adaptive introgression between Anopheles sibling species eliminates a major genomic island but not reproductive isolation. *Nature Communications*, *5*(May), 4248. https://doi.org/10.1038/ncomms5248

Coetzee, M., & Koekemoer, L. (2013). Molecular Systematics and Insecticide Resistance in the Major African Malaria Vector Anopheles funestus. *Annual Review of Entomology*. https://doi.org/10.1146/annurev-ento-120811-153628

Combosch, D. J., & Vollmer, S. V. (2015). Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific Pocillopora corals. *Molecular Phylogenetics and Evolution*, *88*, 154–162. https://doi.org/10.1016/j.ympev.2015.03.022

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*(13), 3133–57. https://doi.org/10.1111/mec.12796

Cutter, A. D., & Payseur, B. A. (2013). Genomic signatures of selection at linked sites: unifying the disparity among species. *Nature Reviews. Genetics*, *14*(4), 262–74. https://doi.org/10.1038/nrg3425

Delmore, K. E., Hübner, S., Kane, N. C., Schuster, R., Andrew, R. L., Câmara, F., … Irwin, D. E. (2015). Genomic analysis of a migratory divide reveals candidate genes for migration and implicates selective sweeps in generating islands of differentiation. *Molecular Ecology*, (24), 1873–1888. https://doi.org/10.1111/mec.13150

Derua, Y. a, Alifrangis, M., Hosea, K. M., Meyrowitsch, D. W., Magesa, S. M., Pedersen, E. M., &

Simonsen, P. E. (2012). Change in composition of the Anopheles gambiae complex and its possible implications for the transmission of malaria and lymphatic filariasis in north-eastern Tanzania. *Malaria Journal*, *11*(1), 188. https://doi.org/10.1186/1475-2875-11-188

Dia, I., Guelbeogo, M., & Ayala, D. (2013). Advances and Perspectives in the Study of the Malaria Mosquito Anopheles funestus. In S. Manguin (Ed.), *Anopheles mosquitoes - New insights into malaria vectors* (INTECH).

Earl, D. A., & VonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(359– 361).

Ellegren, H., Smeds, L., Burri, R., Olason, P. I., Backström, N., Kawakami, T., … Wolf, J. B. W. (2012). The genomic landscape of species divergence in Ficedula flycatchers. *Nature*, *491*(7426), 756–60. https://doi.org/10.1038/nature11584

Evanno, G., Goudet, J., & Regnaut, S. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, *14*, 2611–2620.

Fontaine, M. C., Pease, J. B., Steele, a., Waterhouse, R. M., Neafsey, D. E., Sharakhov, I. V., … Besansky, N. J. (2015). Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science*, *347*(6217), 1258524–1258524. https://doi.org/10.1126/science.1258524

Fouet, C., Kamdem, C., Gamez, S., & White, B. J. (2017). Extensive genetic diversity among populations of the malaria mosquito Anopheles moucheti revealed by population genomics. *Infection, Genetics and Evolution*, *48*, 27–33. https://doi.org/10.1016/j.meegid.2016.12.006

Gillies, M. T., & Coetzee, M. (1987). *A supplement to the Anophelinae of Africa south of the Sahara*. Johannesburg: The South African Institute for Medical Research.

Gillies, M. T., & De Meillon, B. (1968). *The Anophelinae of Africa South of the Sahara* (Second Edi). Johannesburg: Publications of the South African Institute for Medical Research.

Giraldo-Calderon, G. I., Emrich, S. J., MacCallum, R. M., Maslen, G., Dialynas, E., Topalis, P., … Lawson, D. (2015). VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Research*, *43*(D1), D707–D713. https://doi.org/10.1093/nar/gku1117

Gompert, Z., Lucas, L. K., Nice, C. C., Fordyce, J. A., Forister, M. L., & Buerkle, C. A. (2012). Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution*, *66*(7), 2167–2181. https://doi.org/10.5061/dryad.f0b2f083

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, *5*(10), e1000695. https://doi.org/10.1371/journal.pgen.1000695

Hamilton, J. a., Lexer, C., & Aitken, S. N. (2013). Differential introgression reveals candidate genes for selection across a spruce (Picea sitchensis × P. glauca) hybrid zone. *New Phytologist*, *197*(3), 927–938. https://doi.org/10.1111/nph.12055

Hamilton, J. A., Lexer, C., & Aitken, S. N. (2013). Genomic and phenotypic architecture of a spruce hybrid zone (Picea sitchensis × P. glauca). *Molecular Ecology*, *22*(3), 827–841. https://doi.org/10.1111/mec.12007

Harbach, R. E. (2013). The Phylogeny and Classification of Anopheles. In S. Manguin (Ed.),

Anopheles mosquitoes - New insights into malaria vectors (INTECH).
Harr, B. (2006). Genomic islands of differentiation between house mouse subspecies, 730–737. https://doi.org/10.1101/gr.5045006.entiation
Harrison, R. G. (1990). Hybrid zones: windows on evolutionary process. In D. Futuyma & J. Antonovics (Eds.) (Oxford Uni, pp. 69–128). New York.
Harrison, R. G., & Larson, E. L. (2014). Hybridization, introgression, and the nature of species boundaries. *Journal of Heredity*, *105*(S1), 795–809. https://doi.org/10.1093/jhered/esu033
Harrison, R. G., & Larson, E. L. (2016). Heterogeneous genome divergence, differential introgression, and the origin and structure of hybrid zones. *Molecular Ecology*, *25*, 2454–2466. https://doi.org/10.1111/mec.13582
Hemingway, J., Ranson, H., Magill, A., Kolaczinski, J., Fornadel, C., Gimnig, J., … Hamon, N. (2016). Averting a malaria disaster: Will insecticide resistance derail malaria control? *The Lancet*, *387*(10029), 1785–1788. https://doi.org/10.1016/S0140-6736(15)00417-1
Jakobsson, M., & Rosenberg, N. (2007). CLUMPP: a cluster matching and permutation program for dealing with multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801– 1806.
Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405.
Kamdem, C., Fouet, C., Gamez, S., & White, B. J. (2017). Pollutants and insecticides drive local adaptation in African malaria mosquitoes. *Molecular Biology and Evolution*. https://doi.org/DOI: 10.1093/molbev/msx087 [Epub ahead of print]
Kengne, P., Awono-Ambene, H. P. Antonio-Nkondjio, C., Simard, F., & Fontenille, D. (2003). Molecular identification of the Anopheles nili group African malaria vectors. *Med Vet Entomol*, *17*, 67–74.
Key, K. H. L. (1968). The concept of stasipatric speciation. *Syst. Zool.*, *17*, 14–22.
Krzywinski, J., & Besansky, N. J. (2003). Molecular systematics of Anopheles : from subgenera to subpopulations. *Annual Review of Entomology*, *48*(1), 111–139. https://doi.org/10.1146/annurev.ento.48.091801.112647
Lanzaro, G. C., & Lee, Y. (2013). Speciation in Anopheles gambiae — The Distribution of Genetic Polymorphism and Patterns of Reproductive Isolation Among Natural Populations. In S. Manguin (Ed.), *Anopheles mosquitoes - New insights into malaria vectors* (INTECH).
Larson, E. L., Andrés, J. A., Bogdanowicz, S. M., & Harrison, R. G. (2013). Differential introgression in a mosaic hybrid zone reveals candidate barrier genes. *Evolution*, *67*(12), 3653–3661. https://doi.org/10.1111/evo.12205
Larson, E. L., White, T. A., Ross, C. L., & Harrison, R. G. (2014). Gene flow and the maintenance of species boundaries. *Molecular Ecology*, *23*(7), 1668–1678. https://doi.org/10.1111/mec.12601
Lawniczak, M. K. N., Emrich, S. J., Holloway, a K., Regier, a P., Olson, M., White, B., … Besansky, N. J. (2010). Widespread divergence between incipient Anopheles gambiae species revealed by whole genome sequences. *Science (New York, N.Y.)*, *330*(6003), 512–4. https://doi.org/10.1126/science.1195755
Liu, K. J., Dai, J., Truong, K., Song, Y., Kohn, M. H., & Nakhleh, L. (2014). An HMM-Based Comparative Genomic Framework for Detecting Introgression in Eukaryotes. *PLoS*

Computational Biology, *10*(6). https://doi.org/10.1371/journal.pcbi.1003649

Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., … Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in Heliconius butterflies. *Genome Research*, *23*(11), 1817–28. https://doi.org/10.1101/gr.159426.113

Molineaux, L., & Gramiccia, G. (1980). *The Garki Project. Research on the epidemiology and control of malaria in the sudan savanna of West Africa*. World Health Organization.

Mwangangi, J. M., Mbogo, C. M., Orindi, B. O., Muturi, E. J., Midega, J. T., Nzovu, J., … Beier, J. C. (2013). Shifts in malaria vector species composition and transmission dynamics along the Kenyan coast over the past 20 years. *Malaria Journal*, *12*, 13. https://doi.org/10.1186/1475-2875-12-13

Nachman, M. W., & Payseur, B. A. (2012). Recombination rate variation and speciation : theoretical predictions and empirical results from rabbits and mice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 409–421. https://doi.org/10.1098/rstb.2011.0249

Nadeau, N. J., Whibley, A., Jones, R. T., Davey, J. W., Dasmahapatra, K. K., Baxter, S. W., … Jiggins, C. D. (2012). Genomic islands of divergence in hybridizing Heliconius butterflies identified by large-scale targeted sequencing, 343–353. https://doi.org/10.1098/rstb.2011.0198

Ndo, C., Antonio-Nkondjio, C., Cohuet, A., Ayala, D., Kengne, P., Morlais, I., … Simard, F. (2010). Population genetic structure of the malaria vector Anopheles nili in sub-Saharan Africa. *Malaria Journal*, *9*, 161. https://doi.org/10.1186/1475-2875-9-161

Ndo, C., Simard, F., Kengne, P., Awono-Ambene, P., Morlais, I., Sharakhov, I., … Antonio-Nkondjio, C. (2013). Cryptic Genetic Diversity within the Anopheles nili group of Malaria Vectors in the Equatorial Forest Area of Cameroon (Central Africa). *PLoS ONE*, *8*(3), 1–12. https://doi.org/10.1371/journal.pone.0058862

Neafsey, D. E., Lawniczak, M. K. N., Park, D. J., Redmond, S. N., Coulibaly, M. B., Traoré, S. F., … Muskavitch, M. a T. (2010). SNP genotyping defines complex gene-flow boundaries among African malaria vector mosquitoes. *Science (New York, N.Y.)*, *330*(6003), 514–7. https://doi.org/10.1126/science.1193036

Noor, M. a F., & Bennett, S. M. (2009). Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity*, *103*(6), 439–444. https://doi.org/10.1038/hdy.2010.13

Norris, L. C., Main, B. J., Lee, Y., Collier, T. C., Fofana, A., Cornel, A. J., & Lanzaro, G. C. (2015). Adaptive introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated bed nets. *Proceedings of the National Academy of Sciences*, 201418892. https://doi.org/10.1073/pnas.1418892112

Nosil, P., & Feder, J. L. (2012). Widespread yet heterogeneous genomic divergence. *Molecular Ecology*, *21*(12), 2829–2832. https://doi.org/10.1111/j.1365-294X.2012.05580.x

O'Loughlin, S. M., Magesa, S., Mbogo, C., Mosha, F., Midega, J., Lomas, S., & Burt, A. (2014). Genomic Analyses of Three Malaria Vectors Reveals Extensive Shared Polymorphism but Contrasting Population Histories. *Molecular Biology and Evolution*, 1–14. https://doi.org/10.1093/molbev/msu040

Paradis, E., Claude, J., & Strimmer, K. (2004). Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, *20*(2), 289–290.

Parchman, T. L., Gompert, Z., Braun, M. J., Brumfield, R. T., McDonald, D. B., Uy, J. a C., … Buerkle, C. a. (2013). The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. *Molecular Ecology*, *22*(12), 3304–17. https://doi.org/10.1111/mec.12201

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., … Reich, D. (2012). Ancient admixture in human history. *Genetics*, *192*(3), 1065–1093. https://doi.org/10.1534/genetics.112.145037

Peery, A., Sharakhova, M. V, Antonio-Nkondjio, C., Ndo, C., Weill, M., Simard, F., & Sharakhov, I. V. (2011). Improving the population genetics toolbox for the study of the African malaria vector Anopheles nili: microsatellite mapping to chromosomes. *Parasites & Vectors*, *4*(1), 202. https://doi.org/10.1186/1756-3305-4-202

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLoS ONE*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M., Bender, D., … Sham, P. (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, *81*.

Ranson, H., & Lissenden, N. (2016). Insecticide Resistance in African Anopheles Mosquitoes: A Worsening Situation that Needs Urgent Action to Maintain Malaria Control. *Trends in Parasitology*, *32*(3), 187–196. https://doi.org/10.1016/j.pt.2015.11.010

R Development Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Riehle, M. M., Guelbeogo, W. M., Gneme, A., Eiglmeier, K., Holm, I., Bischoff, E., … Vernick, K. D. (2011). A cryptic subgroup of Anopheles gambiae is highly susceptible to human malaria parasites. *Science (New York, N.Y.)*, *331*(6017), 596–8. https://doi.org/10.1126/science.1196759

Rieseberg, L. H., Whitton, J., & Gardner, K. (1999). Hybrid Zones and the Genetic Architecture of a Barrier to Gene Flow Between Two Sunflower Species.

Roesti, M., Hendry, A. P., Salzburger, W., & Berner, D. (2012). Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. *Molecular Ecology*, *21*(12), 2852–62. https://doi.org/10.1111/j.1365-294X.2012.05509.x

Service, M. W. (1993). *Mosquito ecology: field sampling methods*. (U. K. Elsevier Applied Science, London, Ed.).

Sharakhova, M. V., Peery, A., Antonio-Nkondjio, C., Xia, A., Ndo, C., Awono-Ambene, P., … Sharakhov, I. V. (2013). Cytogenetic analysis of Anopheles ovengensis revealed high structural divergence of chromosomes in the Anopheles nili group. *Infection, Genetics and Evolution*, *16*, 341–348. https://doi.org/10.1016/j.meegid.2013.03.010

Sinka, M. E., Bangs, M. J., Manguin, S., Coetzee, M., Mbogo, C. M., Hemingway, J., … Hay, S. I. (2010). The dominant Anopheles vectors of human malaria in Africa, Europe and the Middle East: occurrence data, distribution maps and bionomic précis. *Parasites & Vectors*, *3*(1), 117. https://doi.org/10.1186/1756-3305-3-117

Taylor, S. A., Curry, R. L., White, T. A., Ferretti, V., & Lovette, I. (2014). Spatiotemporally

consistent genomic signatures of reproductive isolation in a moving hybrid zone. *Evolution*, *68*(11), 3066–3081. https://doi.org/10.1111/evo.12510

Turner, T. L., Hahn, M. W., & Nuzhdin, S. V. (2005). Genomic islands of speciation in Anopheles gambiae. *PLoS Biology*, *3*(9), 1572–1578. https://doi.org/10.1371/journal.pbio.0030285

Van Bortel, W., Harbach, R. E., Trung, H. D., Roelants, P., Backeljau, T., & Coosemans, M. (2001). Confirmation of Anopheles varuna in vietnam, previously misidentified and mistargeted as the malaria vector Anopheles minimus, *65*(April 1999), 729–732.

Weng, D., Yu, Y., Hahn, M. W., & Nakhleh, L. (2016). Reticulate evolutionary history and extensive introgression in mosquito species revealed by phylogenetic network analysis. *Molecular Ecology*, *Special is*. https://doi.org/10.1111/mec.13544

WHO. (2016). *World malaria report 2016*.

Wright, S. (1984). *Evolution and the Genetics of Populations: Genetics and Biometric Foundations v. 1 (Genetic & Biometric Foundations); New Edition*. University of Chicago press.

Wu, C. (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology*, *14*, 851–865.

Wu, T. D., & Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*, *26*(7), 873–881.

626 **Tables**

627

628 **Table 1**: Pairwise F_{ST} between divergent subpopulations of *An. nili* s.l.

F_{ST}	***An. nili* group 1**	***An. nili* group 2**	***An. nili* group 3**	***An. ovengensis* group 1**	***An. ovengensis* group 2**
***An. nili* group 1**	-				
***An. nili* group 2**	0.374	-			
***An. nili* group 3**	0.506	0.552	-		
***An. ovengensis* group 1**	0.135	0.275	0.364	-	
***An. ovengensis* group 2**	0.432	0.458	0.492	0.349	-

629

630 **Table 2:** Demographic models of different subgroups of *An. nili* s.l.

Population	**Best Model**	**Log Likelihood**	**Final Population Size[a] (95% CI)**	**Time[b] (95% CI)**
An. nili group 1	*Growth*	-18.42	6.41 (5.326 - 20.71)	3.70 (1.11 - 13.31)
An. nili group 2	*Two-epoch*	-19.97	17.87 (9.33 - 35.50)	11.27 (4.93 - 19.64)
An. ovengensis group 1	*Growth*	-112.18	13.04 (12.15 - 17.26)	0.70 (0.58 - 1.08)
An. Ovengensis group 2	*Growth*	-22.98	19.95 (14.45 - 45.70)	5.11 (2.33 - 15.13)

[a] Relative to ancestral population size.

[b] Expressed in units 2Ne generations from start of growth to present.

Figures

Figure 1: Map showing the sampling locations and relative frequencies of the morphological species *An. nili* and *An. ovengensis*. Small and large black dots indicate respectively the 28 locations surveyed and the four sampling sites where mosquitoes were collected.

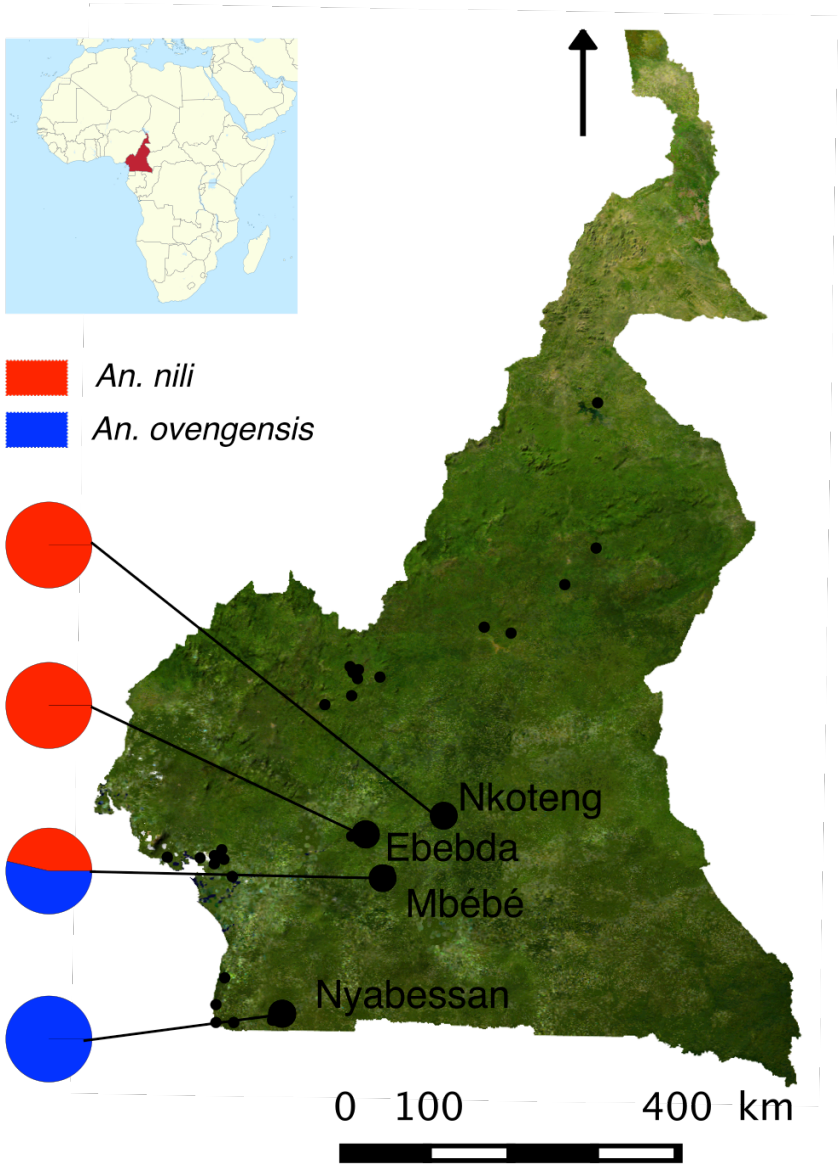


Figure 2: Population genetic structure inferred from 4343 SNPs using PCA (A) and a neighbor-joining tree (B). The percentage of variance explained is indicated on each PCA axis.

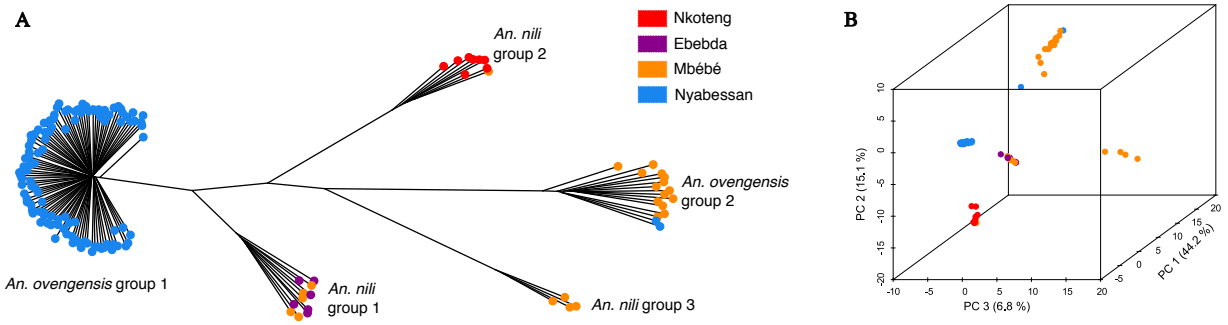


Figure 3: Ancestry proportions inferred in Admixture with k = 2 – 8.

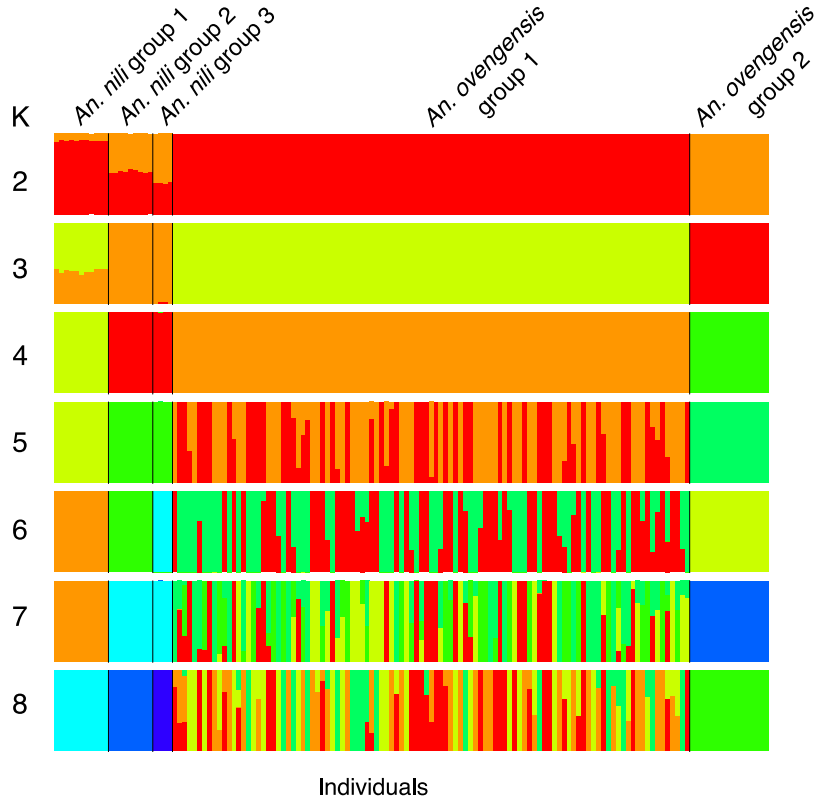
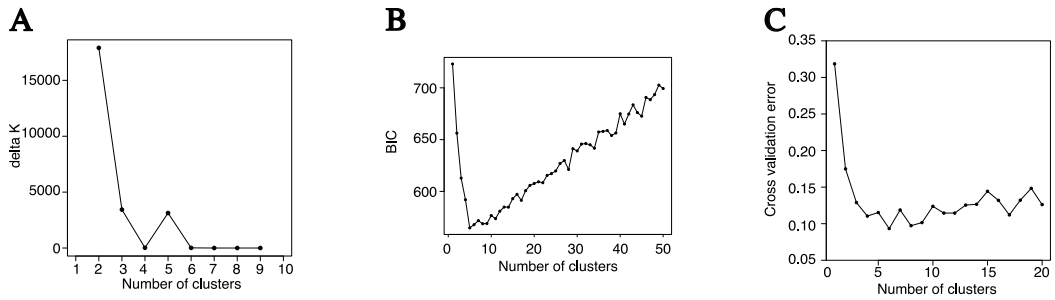
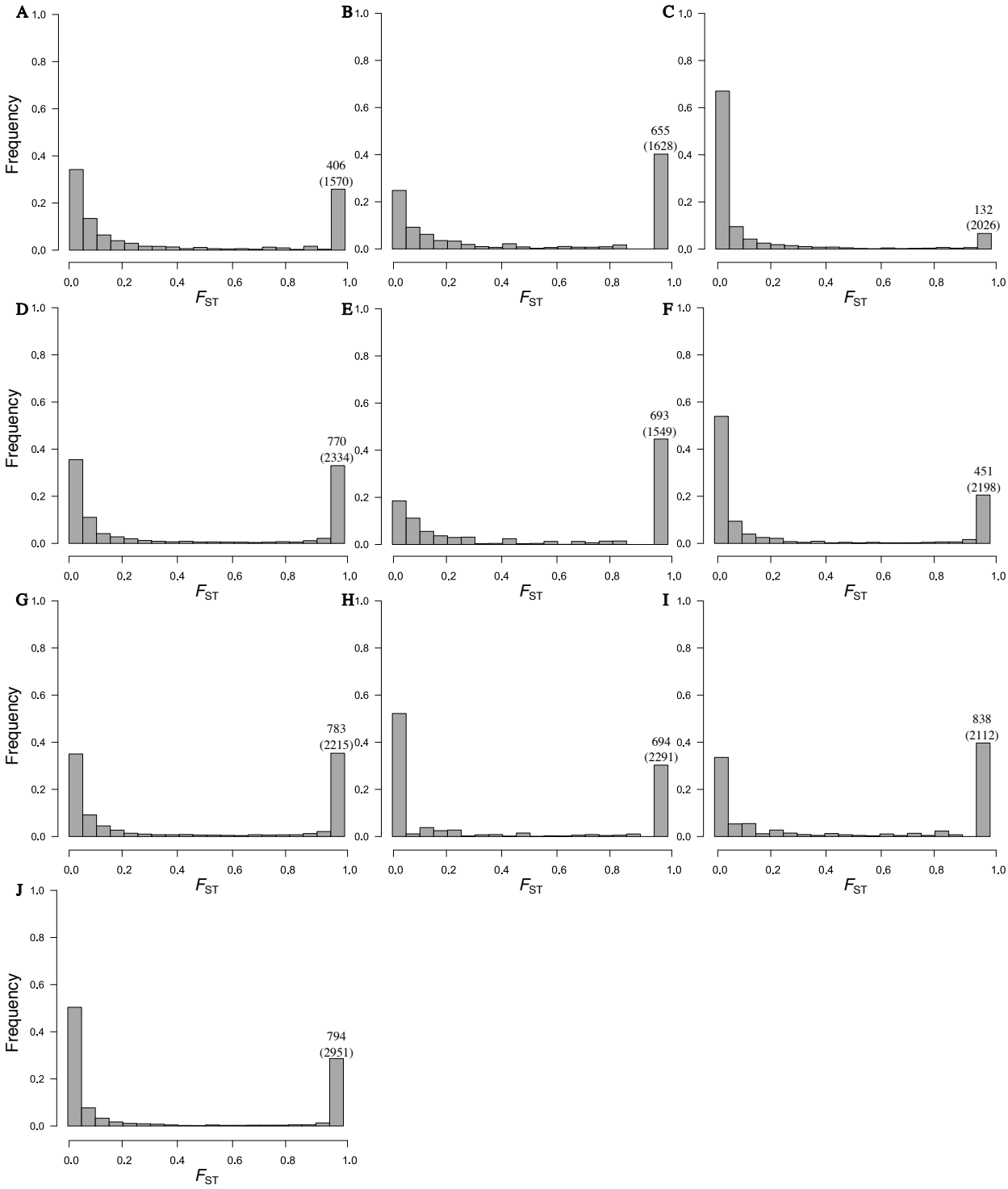


Figure 4: Identification of the optimal number of genetic clusters using the delta k method of Evanno et al., (2005) (A), DAPC (B) and 10-fold cross-validation in Admixture (C). The lowest Bayesian Information Criterion (BIC) and cross-validation error and the highest delta k indicate the most probable number of clusters.



660 **Figure 5:** Distribution of F_{ST} values throughout the genome between *An. nili* group 1 and *An.*
661 *nili* group 2 (A); *An. nili* group 1 and *An. nili* group 3 (B); *An. nili* group 1 and *An. ovengensis*
662 group 1 (C); *An. nili* group 1 and *An. ovengensis* group 2 (D); *An. nili* group 2 and *An. nili*
663 group 3 (E); *An. nili* group 2 and *An. ovengensis* group 1 (F); *An. nili* group 2 and *An.*
664 *ovengensis* group 2 (G); *An. nili* group 3 and *An. ovengensis* group 1 (H); *An. nili* group 3 and
665 *An. ovengensis* group 2 (I); *An. ovengensis* group 1 and *An. ovengensis* group 2 (J). The
666 number of SNPs with F_{ST} = 1 is indicated in each pairwise comparison as well as the total
667 number of SNPs in parenthesis.
668


A
B
C
D
E
F
G
H
I
J
Frequency
F_{ST}
406
(1570)
655
(1628)
132
(2026)
770
(2334)
693
(1549)
451
(2198)
783
(2215)
694
(2291)
838
(2112)
794
(2951)

671 **Supplemental information**

672

673 **Table S1:** Information on mosquito samples included in this study.

Sampling locations	Geographic coordinates	Sampling methods			Total
		HLC-OUT	HLC-IN	LC	
Ebebda	4°20'00"N, 11°17'00"E			6	6
Nkoteng	4°31'00"N, 12°02'00"E			8	8
Nyabessan	2°24'00"N, 10°24'00"E	63	44		107
Mbébé	4°10'00"N, 11°04'00"E	13	3	8	24
Total		76	47	22	145

HLC-OUT, human landing catches performed outdoor; HLC-IN, human landing catches performed indoor; LC, larval collection

674