

RESEARCH

Open Access



# The *Anopheles coluzzii* range extends into Kenya: detection, insecticide resistance profiles and population genetic structure in relation to conspecific populations in West and Central Africa

Luna Kamau<sup>1\*†</sup>, Kelly L. Bennett<sup>2†</sup>, Eric Ochomo<sup>3,9</sup>, Jeremy Herren<sup>4</sup>, Silas Agumba<sup>3</sup>, Samson Otieno<sup>3</sup>, Diana Omoke<sup>3</sup>, Damaris Matoke-Muhia<sup>1,8</sup>, David Mburu<sup>5</sup>, Joseph Mwangangi<sup>6</sup>, Edith Ramaita<sup>7</sup>, Elijah O. Juma<sup>8</sup>, Charles Mbogo<sup>8</sup>, Sonia Barasa<sup>2,8</sup> and Alistair Miles<sup>2</sup>

## Abstract

**Background** *Anopheles coluzzii* is a primary vector of malaria found in West and Central Africa, but its presence has hitherto never been documented in Kenya. A thorough understanding of vector bionomics is important as it enables the implementation of targeted and effective vector control interventions. Malaria vector surveillance efforts in the country have tended to focus on historically known primary vectors. The current study sought to determine the taxonomic status of samples collected from five different malaria epidemiological zones in Kenya as well as describe the population genetic structure and insecticide resistance profiles in relation to other *An. coluzzii* populations.

**Methods** Mosquitoes were sampled as larvae from Busia, Kwale, Turkana, Kirinyaga and Kiambu counties, representing the range of malaria endemicities in Kenya, in 2019 and 2021 and emergent adults analysed using Whole Genome Sequencing (WGS) data processed in accordance with the *Anopheles gambiae* 1000 Genomes Project phase 3. Where available, historical samples from the same sites were included for WGS. Comparisons were made with *An. coluzzii* cohorts from West and Central Africa.

**Results** This study reports the detection of *An. coluzzii* for the first time in Kenya. The species was detected in Turkana County across all three time points from which samples were analyzed and its presence confirmed through taxonomic analysis. Additionally, there was a lack of strong population genetic differentiation between *An. coluzzii* from Kenya and those from the more northerly regions of West and Central Africa, suggesting they represent a connected extension to the known species range. Mutations associated with target-site resistance to DDT and pyrethroids and metabolic resistance to DDT were found at high frequencies up to 64%. The profile and frequencies

<sup>†</sup>Luna Kamau and Kelly L. Bennett contributed equally to this article and are designated as co-first authors.

\*Correspondence:

Luna Kamau

LKamau@kemri.go.ke

Full list of author information is available at the end of the article



of the variants observed were similar to *An. coluzzii* from West and Central Africa but the *ace-1* mutation linked to organophosphate and carbamate resistance present in *An. coluzzii* from coastal West Africa was absent in Kenya.

**Conclusions** These findings emphasize the need for the incorporation of genomics in comprehensive and routine vector surveillance to inform on the range of malaria vector species, and their insecticide resistance status to inform the choice of effective vector control approaches.

**Keywords** *Anopheles coluzzii*, Malaria vectors, Kenya, Population structure, Insecticide resistance

## Background

Malaria is transmitted through the infectious bite of the female *Anopheles* mosquito and is a major cause of morbidity and mortality in Kenya and sub-Saharan Africa in general. In 2022, there were an estimated 249 million cases of malaria worldwide with 233 million of these occurring in the WHO African region and accounting for about 94% of all cases [1]. In Kenya, approximately 70% of the population is at risk of malaria with the disease accounting for an estimated 13–15% of outpatient consultations [2]. Despite concerted global efforts to control malaria, elimination remains a challenge in both low and high burden settings. Historically, the *Anopheles gambiae* species complex and *Anopheles funestus* group have been known to transmit malaria in Kenya. Within the *An. gambiae* complex, *An. gambiae sensu stricto* (s.s.) and *Anopheles arabiensis* were considered the major vectors, with *Anopheles merus* contributing to transmission in coastal Kenya. Invasive *Anopheles stephensi* has also recently been detected in the country, but its contribution to malaria control is yet to be evaluated [3].

*Anopheles coluzzii* is another member of the *An. gambiae* species complex, which is morphologically indistinguishable from at least ten sibling species of malaria in sub-Saharan Africa (SSA) [4, 5]. It is responsible for a significant proportion of the malaria transmission across SSA along with *An. gambiae* and *An. arabiensis*, although *An. funestus* is of increasing concern in East and Southern Africa [6]. *Anopheles coluzzii* was formally named in 2013 following accumulating evidence of subdivision between the previously described M (Mopti) and S (Savana) molecular forms of *An. gambiae*, evidence which included the presence of pre-mating barriers and genome-wide divergence and independent evolutionary trajectories; consequently, the M form was assigned the name *An. coluzzii* and the S form *An. gambiae* [7]. *Anopheles coluzzii* is widely distributed in West and Central Africa and found in sympatry with other members of the species complex and has also been documented in Somalia [8]. To date however, no record of its presence in Kenya is available. Similar to other vector species within the *An. gambiae* species complex, the distribution and role of *An. coluzzii* in malaria transmission as well as the

development of insecticide resistance varies greatly in different settings [9–11].

The heterogeneities with respect to ecological characteristics and trophic habits of members of *An. gambiae sensu lato* (s.l.) have allowed the expansion of its range and contributed to its success in malaria transmission [12, 13]. Compared to *An. gambiae*, which prefers to breed in unpolluted environments typical of rural areas, *An. coluzzii* possess a greater capacity to survive in ecologically complex environments characterized by the presence of a variety of stressors. *Anopheles coluzzii* is likely to have a greater resistance to desiccation because it predominates in arid regions [14, 15]. Studies have shown *An. coluzzii* to have greater tolerance to salinity as well as xenobiotics and ammonia pollutants in larval habitat compared to *An. gambiae*, enabling extension of its range beyond traditional rural settings to densely urbanized settings [16–18]. With the trends in urban migration seen in many African countries expected to continue and unplanned human settlements in urban settings presenting a potential risk factor of increased malaria transmission [19], comprehensive vector surveillance is crucial. Such surveillance will improve the understanding of vector bionomics thereby enabling the implementation of targeted and effective vector control interventions [20–22]. These objectives are in line with the WHO Global Vector Control Response 2017–2030 strategy's recommendation of strengthening national surveillance systems and integration with health information systems to guide vector control and effectiveness [23].

Most entomological studies and vector surveillance efforts, including those in Kenya tend to focus on the historically known primary vectors of malaria to the exclusion of unanticipated or novel vector species. Recently however, the potential role of secondary vectors in malaria transmission has been highlighted aided by the use of molecular identification tools [24, 25]. The current study used whole genome sequence (WGS) data to investigate the taxonomic relationship of samples collected across five counties in Kenya with different epidemiological parameters. This investigation describes the first report on *An. coluzzii* in Kenya, while seeking to characterize their population structure and insecticide resistance profiles.

**Methods**

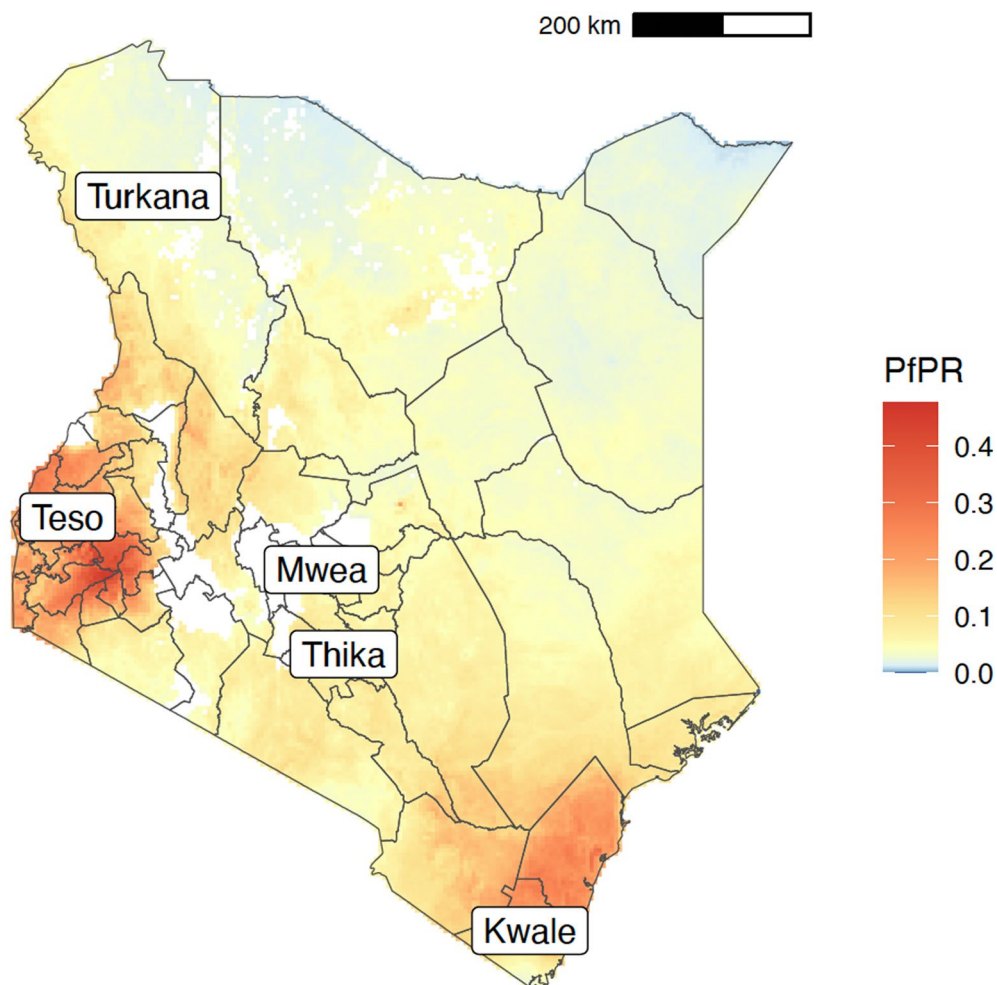
**Mosquito sampling, identification and rearing**

The study utilized archived mosquitoes collected from previous studies in addition to samples collected between December 2019 and February 2021 from five study sites (Fig. 1). These were: (i) Teso in Busia County, (ii) Kwale in Kwale County, (iii) Kakuma in Turkana County, (iv) Mwea in Kirinyaga County and (v) Thika in Kiambu County (Table 1). These five locations represent different ecological and malaria epidemiological zones within Kenya [26]. Teso on the western border of Kenya with Uganda is within the lake endemic zone. Kwale in the southeast is within the coastal endemic zone. Turkana in the northwest is relatively arid and within the seasonal transmission zone. Mwea and Thika are within the central highlands with Mwea being in the seasonal transmission zone and Thika in the low-risk zone.

For both sets of mosquitoes, *An. gambiae* s.l. larvae identified based on morphology [28] were collected, using standard larval 350 ml dippers, from multiple breeding sites to minimize chances of sampling siblings. The larvae were transported to the laboratory for rearing to adults (except for Turkana where rearing was carried out in the field due to the long distance to the laboratory). Larvae were reared in water collected from the larval sampling site or dechlorinated tap water at temperatures between 28 °C–31 °C and humidity between 80%–85% and fed on finely ground *Sera Vipar staple diet*<sup>TM</sup> (Sera, Germany) fish food. The resultant adult mosquitoes aged 2–7 days old were analysed using WGS.

**Sequencing and SNP calling**

Sequencing and single nucleotide polymorphisms (SNP) calling was performed following the Ag1000G phase 3



**Fig. 1** Map of Kenya showing the study sites. The figure shows the sampling locations in Kenya in relation to *Plasmodium falciparum* prevalence rate in 2015 standardized to the age group 2 to 10 years using data obtained from the malariaAtlas R package [27]

**Table 1** Sampling locations in Kenya and the number of samples successfully sequenced, grouped by taxon

Location	Latitude	Longitude	Year	Month	Taxon			
					<i>An. arabiensis</i>	<i>An. coluzzii</i>	<i>An. gambiae</i>	<i>An. quadriannulatus</i>
Kwale	-4.572	39.257	2019	7	7	0	0	2
Mwea	-0.717	37.382	2007	6	18	0	0	0
			2014	9	15	0	0	0
			2020	2	46	0	0	0
			2020	12	26	0	0	0
			2021	1	29	0	0	1
			2021	2	16	0	0	0
Teso	0.626	34.236	2013	8	24	0	4	0
			2019	7	17	0	14	0
			2019	12	43	0	17	0
Thika	-1.061	37.181	2019	7	45	0	1	0
			2019	8	19	0	0	0
Turkana	3.717	34.857	2006	2	19	5	1	0
			2019	1	51	13	0	0
			2019	9	123	8	0	0

project protocol. Briefly, paired-end multiplex libraries were prepared using Illumina's DNA preparation protocol with fragmentation using Covaris Adaptive Focused Acoustics. Multiplexes of 12 tagged individual mosquitoes were sequenced in three replicates using Illumina HiSeq 2000 and the Illumina HiSeq X technologies. Reads were aligned to the AgamP4 reference genome using BWA version 0.7.15 and indel realignment and SNP calling performed using GATK version 3.7.0. Quality control filters applied included the exclusion of individuals with median coverage < 10X, with no coverage across > 50% of the reference genome, or samples identified as cross-contaminated by a percentage of  $\geq 4.5\%$  using the protocols set out by the AG1000G project. Only technical replicates with the best sequencing coverage were retained. Additionally, site filters defined by the Ag1000G project were applied to exclude sites where SNP calling and genotyping was less reliable because the observed genotypes were not consistent with Mendelian inheritance in laboratory crosses.

#### Taxonomic assignment

To investigate taxonomic status, individual mosquitoes were assessed against two sets of ancestry-informative markers (AIMs) used to distinguish *An. gambiae* from its sister taxa *An. coluzzii* and *An. gambiae/An. coluzzii* from *An. arabiensis* using publicly available data from the *Anopheles* 16 genomes project [29–31]. The dataset, described by the *Anopheles gambiae* 1000 Genomes project [32], includes a set of AIMs SNPs informative in distinguishing taxa because they are exclusive to each

taxonomic group discounting multiallelic sites and those with missing data. In total, 2612 and 700 AIMs were used to differentiate *An. gambiae/An. coluzzii* from *An. arabiensis* and *An. gambiae* from *An. coluzzii*, respectively. Individuals assessed against AIMs for distinguishing *An. gambiae* from *An. coluzzii* were called as *An. gambiae* when the fraction of coluzzii-like calls was < 0.12 and *An. coluzzii* where this fraction was > 0.9. Individuals assessed against AIMs distinguishing *An. gambiae/An. coluzzii* from *An. arabiensis* were called as the latter when the fraction of arabiensis-like alleles was > 0.6. Individuals in-between these fractions represent other taxa.

#### Population structure

To compare the genomic composition of *An. coluzzii* in Kenya with other *An. coluzzii* cohorts, a Principal Component Analysis (PCA) dimensionality reduction on the allele counts of 100,000 biallelic SNPs equally distributed across chromosome three so as to exclude known introgressions was performed. Chosen SNPs had a minor allele frequency greater than 0.2% and no missing data. Using the same criteria for SNP selection, comparison of the evolutionary relationships of African *An. coluzzii* was carried out by constructing an unrooted Neighbour-Joining tree with a city block distance metric. To determine whether *An. coluzzii* across Africa are connected to *An. coluzzii* in Kenya, genomic differentiation between populations using Hudson's pairwise  $F_{ST}$  was computed [33]. To further investigate whether *An. coluzzii* in Kenya have a similar demography to other *An. coluzzii*, informative summary statistics including Nucleotide diversity ( $\theta\pi$ ),

Watterson's theta ( $\theta_w$ ) and Tajima's D were calculated. All analyses were performed using the open source and freely available malariagen\_data python package.

### Insecticide resistance

To investigate whether *An. coluzzii* in Kenya have target site mutations associated with insecticide resistance similar to other African *An. coluzzii*, the malariagen\_python package was used to calculate amino acid substitution frequencies based on the occurrence of non-synonymous SNPs at genomic sites of interest. These included the gene targeted by pyrethroid insecticides, the voltage-gated sodium channel (*Vgsc*; AGAP004707), the glutathione S-transferase gene conferring resistance to DDT (*Gste2*; AGAP009194), the Resistance to dieldrin gene (*Rdl*; AGAP006028) and the organophosphate target gene, acetylcholinesterase (*Ace1*; AGAP001356). To account for sequencing error and remove substitutions unlikely to be under selection, only amino acid substitutions present at a frequency greater than 5% in at least one population were retained.

## Results

### Population sampling and sequencing

A total of 1130 individual mosquitoes were collected during this study from five locations in Kenya (Fig. 1; Table 1). All locations included relatively recent sampling (2019–2021) and samples from earlier collections (2006–2014) were also available for Mwea, Teso and Turkana. A total of 744 mosquitoes with good quality extracted DNA were submitted for WGS, of which 564 passed all data quality control filters, with an average median coverage of 36 X and minimum of 10 X. After alignment to the AGAMP4 reference genome, we discovered a total of 83,052,633 SNPs segregating within the samples from this study, of which 43,701,680 passed all site quality filters previously established by the *Anopheles gambiae*1000 Genomes Project phase 3 [32].

### Taxon assignment

Taxon assignment within the *An. gambiae* complex is challenging because taxa are morphologically indistinguishable, and conventional genetic markers are based on a single locus that does not always reflect the ancestry of the rest of the genome [28, 29]. All samples in this study were morphologically identified as *An. gambiae* s.l., then the genomic data used to investigate the species. Using a set of ancestry-informative markers (AIMs) previously ascertained from samples with known species status [30, 31], 498 samples were identified as *An. arabiensis* (*arabiensis* AIM fraction > 0.85; Additional file 1: Figure S1). Of the remaining samples, 37 were identified as *An. gambiae* (*coluzzii* AIM fraction < 0.1) and 26 were identified

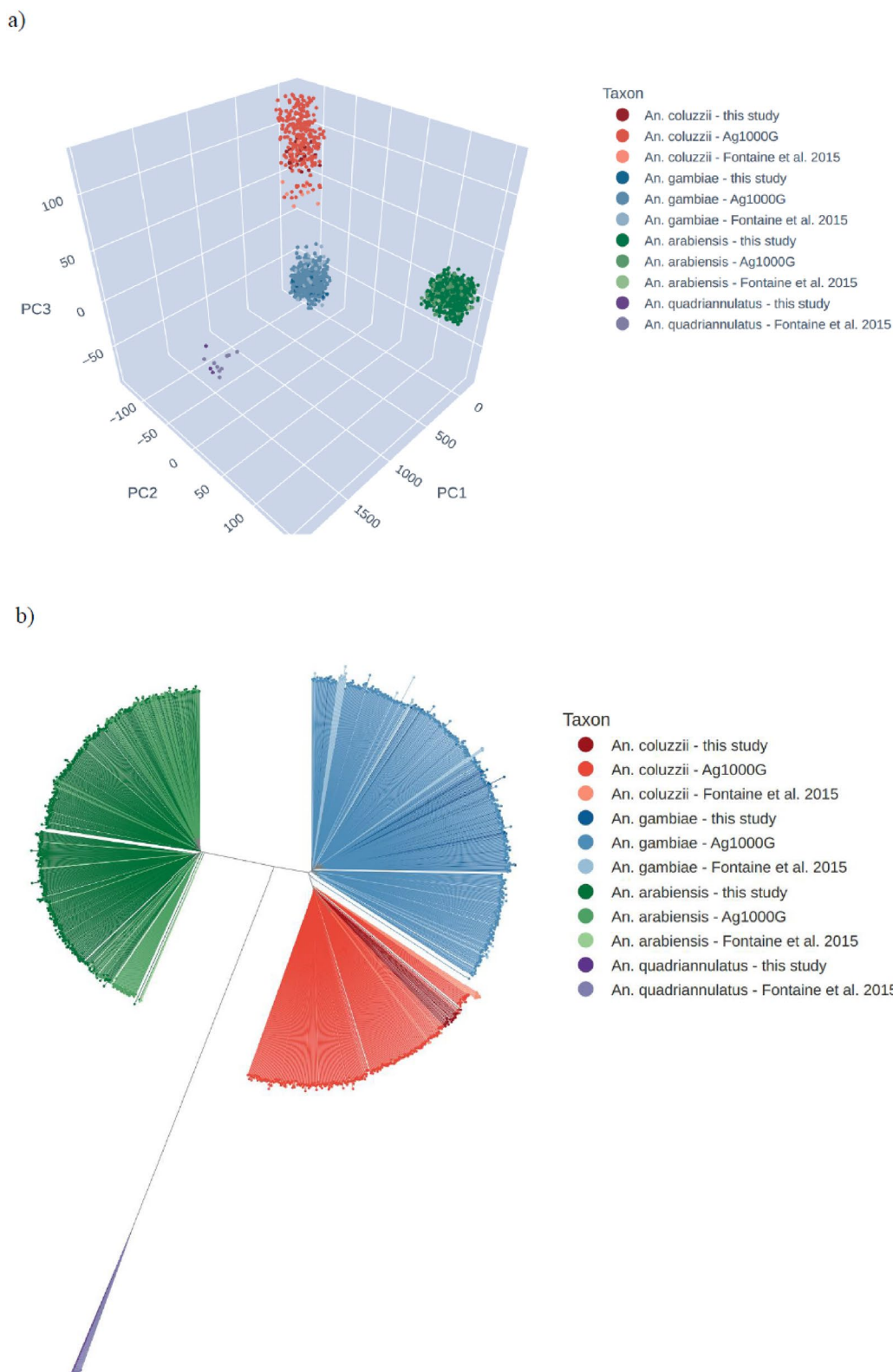
as *An. coluzzii* (*coluzzii* AIM fraction > 0.9). To provide additional confirmation of taxonomic status, a principal components analysis (PCA) was performed, and a neighbour-joining tree (NJT) constructed using genomic data from this study together with samples from other African countries from the *Anopheles gambiae* 1000 Genomes Project and the study of Fontaine et al. [31] (Fig. 2). These analyses showed a clear grouping by species, with the position of Kenyan *An. gambiae*, *An. arabiensis* and *An. coluzzii* samples entirely consistent with the AIM results. On comparison of data originating from multiple taxa [31], a further 3 individuals that could not be identified via AIMs were confirmed to be *Anopheles quadriannulatus* based on the PCA and NJT.

*Anopheles coluzzii* populations in West Africa are commonly found to have experienced adaptive introgression of genetic material from *An. gambiae* within a genomic region towards the centromere of chromosome arm 2L, driven by selection for pyrethroid target-site resistance alleles [35–37]. AIM profiles revealed the majority of *An. coluzzii* from Kenya are either homozygous (38%) or heterozygous (46%) for introgression from *An. gambiae* towards the centromere of 2L (Additional file 1: Figure S1).

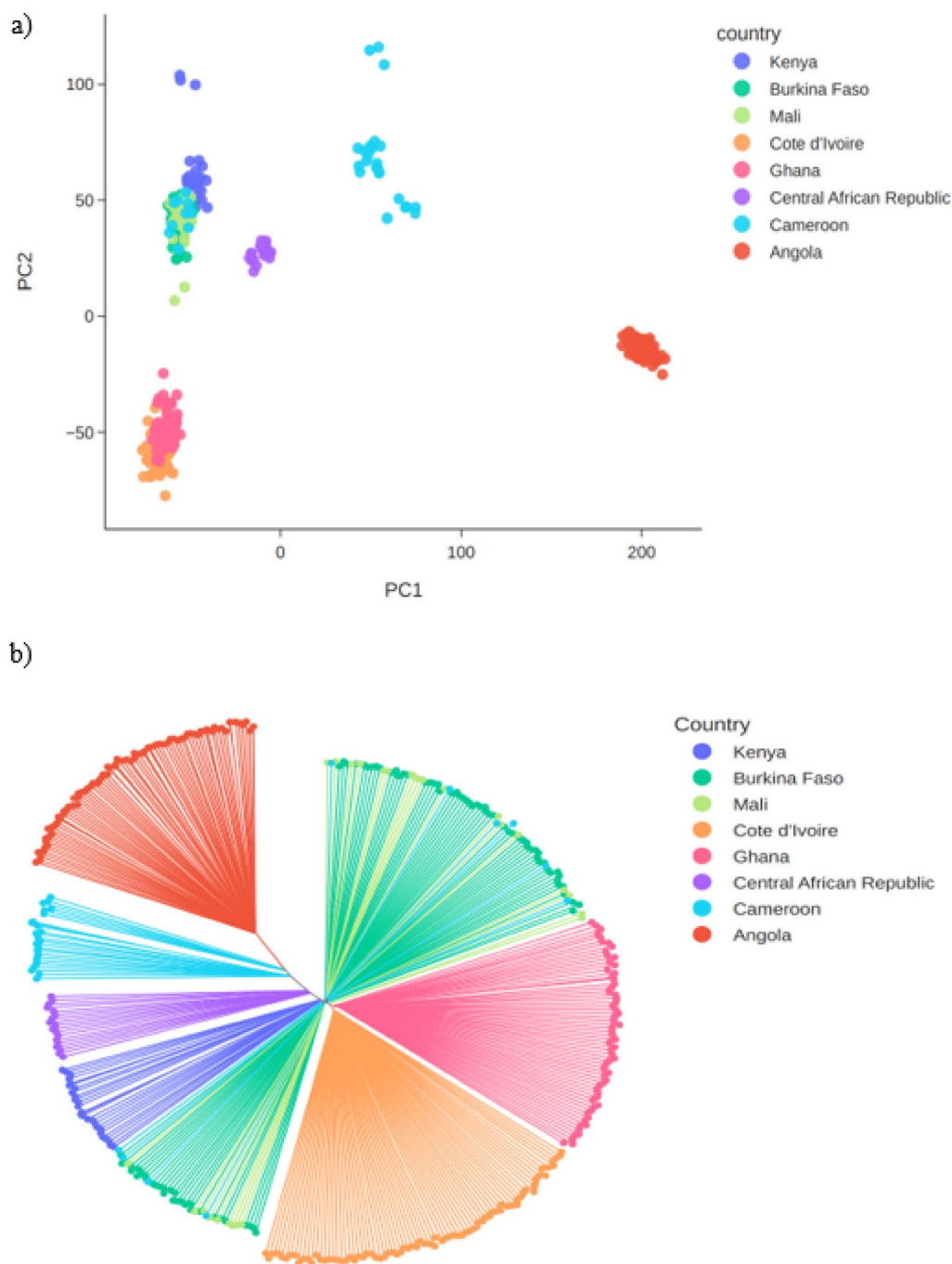
*Anopheles coluzzii* has not previously been reported in Kenya, but is a highly competent malaria vector in West and Central Africa. In the current study, the species were detected only in Turkana and across three different sampling time points at frequencies of 20.0% (5 out of 25) in month 2 of 2006, 20.3% (13 out of 64) in month 1 of 2019 and 6.1% (8 out of 131) in month 9 of 2019. Given the importance of this finding for malaria vector surveillance and control, the remainder of this report focuses on a full characterization of the Kenyan *An. coluzzii*. Analysis of genomic data from the other *Anopheles* taxa sequenced in this study will be reported separately.

### Geographical population structure and genetic diversity

To explore the genetic relationship between the Kenyan *An. coluzzii* and conspecific populations from other countries, data from this study was combined with previous sequence data of *An. coluzzii* populations from inland West Africa (Burkina Faso, Mali), coastal West Africa (Côte d'Ivoire, Ghana), and Central Africa (Cameroon, the Central African Republic, Angola) [34]. Single Nucleotide Polymorphisms (SNPs) from Chromosome 3, which is free from polymorphic inversions, were used to perform a principal components analysis (PCA), compute a neighbour-joining tree (NJT) and quantify the degree of allele frequency differentiation ( $F_{ST}$ ) between cohorts from different locations. The PCA and NJT analyses grouped the Kenyan *An. coluzzii* most closely with *An. coluzzii* from inland West Africa (Mali, Burkina



**Fig. 2** Taxon confirmation of Kenya *An. coluzzii*. Confirmation of taxon assignment using principal components analysis (PCA) and neighbour-joining tree (NJT) analysis, comparing samples in this study with reference samples from the *Anopheles gambiae* 1000 genomes project from inland West Africa (Burkina Faso, Mali), coastal West Africa (Cote d'Ivoire, Ghana), Central Africa (Cameroon, the Central African Republic, Angola) and the study of Fontaine et al. [31]. **a** PCA. **b** NJT

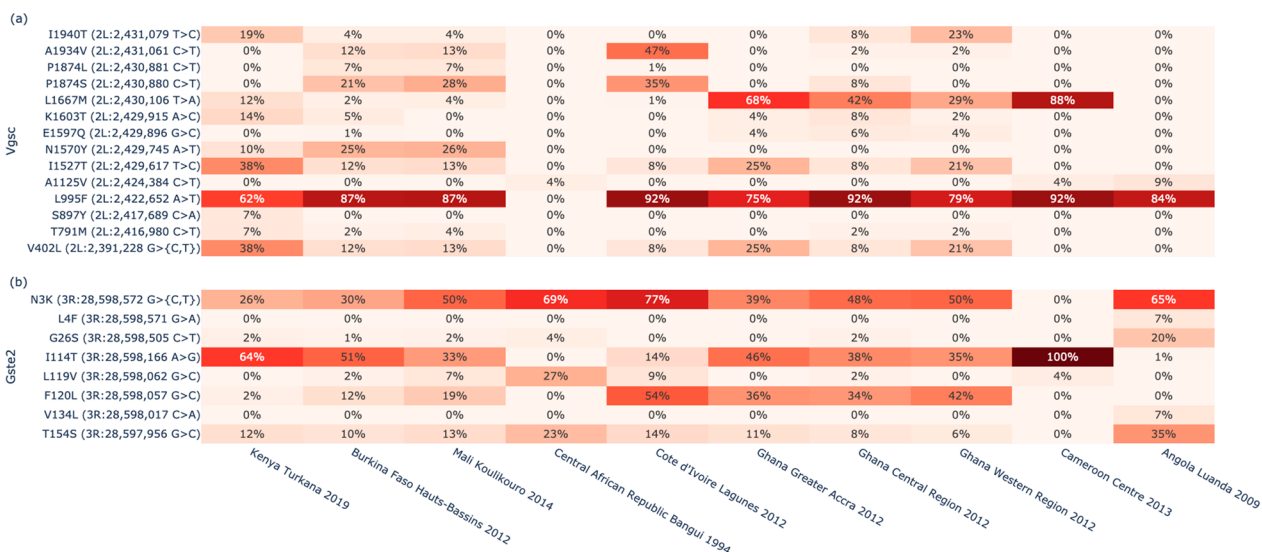


**Fig. 3** Population genetic structure of Kenyan *An. coluzzii*. The figure shows the analysis of geographical population structure within *An. coluzzii*, comparing samples from Turkana, Kenya collected in this study with reference samples from the Ag1000G project and Fontaine et al. [31]. Kenyan *An. coluzzii* are most closely related to *An. coluzzii* from inland West Africa (Mali, Burkina Faso) and Northern Cameroon. **a** PCA. **b** NJT

Faso) and northern Cameroon (Fig. 3). The  $F_{ST}$  results were consistent with these analyses, finding the lowest  $F_{ST}$  between Kenya, Mali and Burkina Faso ( $F_{ST}$  0.006–0.007 for both comparisons; Additional file 1: Table S1). These results show a lack of strong population structure

between *An. coluzzii* from Kenya and more northerly regions of West and Central Africa.

To further explore whether Kenyan *An. coluzzii* share a genomic profile similar to northern West and Central Africa, a PCA targeting the 2La and 2Rb regions of the



**Fig. 4** Insecticide resistance profiles of Kenyan *An. coluzzii*. The figure shows amino acid substitution frequencies in genes associated with target-site resistance to DDT and pyrethroids. **a** Vgsc (AGAP004707). **b** Gste2 (AGAP009194)

genome was performed to investigate whether the populations share the same 2La and 2Rb inversion karyotype. *Anopheles coluzzii* from Kenya have the same 2La/2Rb karyotype found in *An. coluzzii* from Burkina Faso and Mali only, supporting the finding that they are genetically most similar to these populations (Additional file 1: Figure S2). To investigate whether the Kenyan *An. coluzzii* have a similar demographic history to *An. coluzzii* populations from other countries in Africa, genetic diversity summary statistics for mosquito cohorts grouped by geographical region and year of sampling were computed. Nucleotide diversity, the density of segregating sites (Watterson’s theta) and allele frequency spectra (Tajima’s D) in Kenyan *An. coluzzii* were similar to *An. coluzzii* from West Africa (Nucleotide diversity, 0.026; Watterson’s theta, 0.033; Tajima’s D, -0.961; Additional file 1: Figure S3), suggesting lack of genetic isolation.

**Insecticide resistance**

To investigate whether Kenyan *An. coluzzii* have mutations associated with target-site resistance to insecticides, we computed amino acid allele frequencies across four genes encoding insecticide binding targets for cohorts with at least ten data points: Vgsc (AGAP004707), Gste2 (AGAP009194), Rdl (AGAP006028) and Ace1 (AGAP001356). Kenyan *An. coluzzii* were found to have the Vgsc-L995F substitution associated with resistance to DDT and pyrethroids, at 62% frequency (Fig. 4a) [38]. This allele is also found at high frequency in *An. coluzzii* populations throughout West and Central Africa. In addition, Kenyan *An. coluzzii* displayed the double

mutant Vgsc-V402L + I1527T substitution at 38% frequency. This double mutant is also present in other *An. coluzzii* populations. Additionally, Kenyan *An. coluzzii* have the Gste2-I114T mutation at high frequency (64%), which confers metabolic resistance to DDT [39] (Fig. 4b) and also observed at high frequency in *An. coluzzii* populations from West and Central Africa.

Two haplotypes have been previously associated with resistance to dieldrin in *An. coluzzii*, Rdl-A296G/T345M widely distributed across Africa and Rdl-A296S/T345S previously only reported from Burkina Faso [40]. The 296/345 substitution pair was observed in Burkina Faso and also in Mali and Kenya, but the frequency was very low, ~5%, in *An. coluzzii* from Kenya (Additional file 1: Figure S4). A mutation in the Acetylcholinesterase gene, Ace1-G280S, was previously linked to organophosphate and carbamate resistance [41]. Although the mutation is present in *An. coluzzii* from coastal West Africa (i.e., Côte d’Ivoire and Ghana), it was not present in *An. coluzzii* from Kenya (Additional file 1: Figure S5).

**Discussion**

A proper understanding of vector bionomics within the context of the malaria transmission system is important for the choice and successful implementation of vector control interventions. The current study sought to improve this understanding by analyzing the distribution of *An. gambiae* s.l. sibling species from five sites in Kenya, one each in the different malaria epidemiological zones. Based on analysis of whole-genome data we report, for the first time in Kenya, the presence of *An. coluzzii*.



*Anopheles coluzzii* is a highly competent malaria vector in West and Central Africa. All the *An. coluzzii* samples found in this study were collected from Turkana in Northwest Kenya. This included mosquitoes collected across multiple years and months spanning both the dry and rainy season, confirming the presence of *An. coluzzii* in the region since at least 2006 (Table 1). Seasonal shifts in *An. gambiae* species complex composition linked to adaptation to ecology are known to occur [42, 43]. Similar to mosquito collection from Turkana, mosquitoes that were analyzed from the other sites included collections from multiple years and seasons, discounting the possibility that the failure to detect *An. coluzzii* from these sites could be related to seasonality. Past failure to detect *An. coluzzii* in Kenya is likely associated with the fact that malaria vector surveillance and identification has tended to focus on historically known primary vectors using species-specific markers. For example, in Turkana, a past study reported all mosquitoes identified as *An. arabiensis*, but only 84.8% of samples reacted in PCR diagnostic tests based on species-specific nucleotide sequences in the ribosomal DNA intergenic spacers of the *An. gambiae* species complex [44, 45]. Failure to identify specimens in such molecular assays is usually explained away as failure in the assay itself due to factors such as poor-quality DNA or other associated factors. The current analysis exploited WGS allowing the use of a large number of markers distributed across the entire genome to detect cryptic and/or unknown species demonstrating how whole genomic sequencing can be configured to support routine surveillance to accurately identify present and shifting vector distributions, adaptive evolutionary changes and populations structuring, integral to understanding of malaria transmission dynamics [46].

When the Kenya *An. coluzzii* were compared with populations from the northerly regions of West and Central Africa, they showed a lack of strong population genetic structure evidenced by clustering on the PCA and NJT and low  $F_{ST}$ , suggesting relatively unrestricted gene flow across a northerly belt spanning continental Africa. In support of this notion, Kenyan *An. coluzzii* had a similar genetic diversity to other West and Central African populations. Exploration of the 2La/2Rb inversion karyotypes, which has previously been linked to aridity and increases in frequency with this cline [47, 48] revealed that Kenyan *An. coluzzii* have the same inversion karyotype found in Burkina Faso and Mali at both loci, supporting the finding of genetic similarity to these populations. The finding that the 2La/2Rb karyotype is the same as that found in arid West and Central Africa is consistent with its

environmental association, since Turkana in Kenya is arid to semi-arid, experiencing high temperatures and highly seasonal rainfall. The presence of *An. coluzzii* has also been documented in Somalia, which shares a similar arid to semi-arid ecosystem [8]. These findings suggest Kenyan *An. coluzzii* are not an isolated population, nor have they experienced any recent bottlenecks or other distinct demographic events.

Investigation of the insecticide resistance profiles of Kenyan *An. coluzzii* revealed the presence of the *Vgsc-L995F* substitution associated with resistance to DDT and pyrethroids [38] at 62% frequency and the *Gste2-I114T* mutation which confers metabolic resistance to DDT [39] at a frequency of 64%. These alleles are also found at high frequency in *An. coluzzii* populations from West and Central Africa [49–51]. In addition, Kenyan *An. coluzzii* displayed the double mutant *Vgsc-V402L+I1527T* substitution at 38% frequency also present in other *An. coluzzii* populations and recently observed to be increasing in frequency in *An. coluzzii* in Burkina Faso [52]. The V402L substitution has been functionally validated to confer insecticide resistance with reduced fitness cost to the mosquito when compared to L995F [53] and the fact that V402L is almost exclusively found in combination with I1527T suggests a strongly synergistic effect that may further increase fitness in the presence of insecticides.

Previous studies have documented resistance to pyrethroids in *An. gambiae* s.l. in Kakuma (Turkana County) with enzymatic resistance being implicated [54]. Both ITNs and IRS have historically been used to prevent malaria in the area. The finding of mutations associated with resistance to insecticides is therefore not surprising and fits with the picture of widespread and increasing insecticide resistance in the country [55]. The presence of the V402L+I1527T substitution in Kenya is concerning, because variants conferring stronger pyrethroid resistance could compromise the efficacy of new pyrethroid+PBO LLINs, currently considered among the most effective defense against resistant populations [56].

Increased urbanization promotes retention of surface water and is associated with water distribution and drainage systems that provide suitable habitat for *Anopheles* mosquitoes. The refugee camp from which the samples were collected is a highly populated region that experiences low level local transmission and has recently suffered recurrent epidemics [46]. Historically, the area has received little attention regarding malaria intervention, because the climate was considered unsuitable for known vector species, except *An. arabiensis*, considered a less competent vector.

## Conclusion

The observation that *An. coluzzii*, an efficient vector of malaria which can tolerate more arid conditions and thrive in both rural and urban settings occurs in Kenya but has not been reported to date means that this vector is potentially contributing to malaria transmission in Turkana County and malaria control interventions currently in place may be ineffective against it. This finding alongside the recent finding of *An. stephensi* in Turkana and other regions of northern Kenya emphasizes the need for re-evaluation of the distribution, bionomics and epidemiological significance of the local vector populations in the country. This is the only way the country will be able to ensure vector control approaches are sufficiently targeted at the myriad of *Anopheles* vectors responsible for transmission in the different settings in Kenya.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12936-024-04950-x>.

**Additional file 1: Table S1.** Allele frequency differentiation (FST) between different *An. coluzzii* cohorts across Africa. **Figure S1.** Taxon assignment using ancestry informative marker (AIM) genotypes. **Figure S2.** Principal components analysis of the 2La (2L:20,528,089–42,165,182) and 2Lb (2R:19,444,433–26,313,071) inversion regions to compare the karyotype of Kenyan *An. coluzzii* to other *An. coluzzii* cohorts across Africa. **Figure S3.** Bar plots of nucleotide diversity, Watterson's theta and Tajima's D comparing Kenyan *An. coluzzii* to other *An. coluzzii* cohorts. **Figure S4.** Amino acid frequencies for the resistance to dieltrin gene Rdl in *An. coluzzii*. **Figure S5.** Amino acid frequencies for the Ace1 gene in *An. coluzzii*.

## Acknowledgements

This study was supported by the MalariaGEN Vector Observatory which is an international collaboration working to build capacity for malaria vector genomic research and surveillance and involves contributions by the following institutions and teams. Wellcome Sanger Institute: Lee Hart, Kelly L. Bennett, Anastasia Hernandez-Koutoucheva, Jon Brenas, Menelaos Ioannidis, Chris Clarkson, Alistair Miles, Julia Jeans, Paballo Chauke, Victoria Simpson, Eleanor Drury, Osama Mayet, Sónia Gonçalves, Katherine Figueroa, Tom Maddison, Kevin Howe, Mara Lawniczak; Liverpool School of Tropical Medicine: Eric Lucas, Sanjay Nagi, Martin Donnelly; Broad Institute of Harvard and MIT: Jessica Way, George Grant; Pan-African Mosquito Control Association: Jane Mwangi, Edward Lukyamuzi, Sonia Barasa, Ibra Lujumba, Elijah Juma. The authors would like to thank the staff of the Wellcome Sanger Genomic Surveillance unit and the Wellcome Sanger Institute Sample Logistics, Sequencing and Informatics facilities for their contributions. The MalariaGEN Vector Observatory is supported by multiple institutes and funders. The Wellcome Sanger Institute's participation was supported by funding from Wellcome (220540/Z/20/A, 'Wellcome Sanger Institute Quinquennial Review 2021-2026') and the Bill & Melinda Gates Foundation (INV-001927). The Liverpool School of Tropical Medicine's participation was supported by the National Institute of Allergy and Infectious Diseases (NIAID) R01-AI116811, with additional support from the Medical Research Council (MR/P02520X/1). The latter grant is a UK-funded award and is part of the EDCTP2 programme supported by the European Union. Martin Donnelly is supported by a Royal Society Wolfson Fellowship (RSWF/FT/180003). The Pan-African Mosquito Control Association's participation was funded by the Bill and Melinda Gates Foundation (INV-031595).

## Author contributions

LK, EO, JH, DMM, DM, JM, EJ, CM conceptualized and designed the study, data interpretation and assisted in drafting the manuscript. LK, EO, SA, SO, JM, designed the study, conducted sample collection and processing, data

collection. LK, KLB, EO, DO, SB, AM conducted data analysis and interpretation and assisted in drafting the manuscript. All authors read, reviewed and approved this manuscript.

## Funding

This study was funded by the Bill and Melinda Gates Foundation (Grant Number OPP1210319) to the Kenya Medical Research Institute (KEMRI), PI: LK and supported by the Pan African Mosquito Control Association (PAMCA).

## Availability of data and materials

The sequences for the specimens identified in this study were submitted to The European Nucleotide Archive (ENA) (accession nos. ERR11811571–ERR11811786, ERR11812004–ERR11812101 and ERR11840239–ERR12031958).

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests. No chatbot or artificial intelligence tool were used in any of this work.

### Author details

<sup>1</sup>Centre for Biotechnology Research and Development (CBRD), Kenya Medical Research Institute, PO Box 54840, Nairobi 00200, Kenya. <sup>2</sup>Malaria Vector Genomic Surveillance, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. <sup>3</sup>Centre for Global Health Research (CGHR), Kenya Medical Research Institute, Nairobi, Kenya. <sup>4</sup>International Center for Insect Physiology and Ecology (Icipe), Nairobi, Kenya. <sup>5</sup>Pwani University Biosciences Research Centre (PUB-REC), Kilifi, Kenya. <sup>6</sup>Centre for Geographic Medicine Research-Coast (CGMR-C), Kenya Medical Research Institute, Nairobi, Kenya. <sup>7</sup>Ministry of Health-National Malaria Control Programme (NMCP), Kenya, Nairobi, Kenya. <sup>8</sup>Pan African Mosquito Control Association (PAMCA), Nairobi, Kenya. <sup>9</sup>Liverpool School of Tropical Medicine, Liverpool, United Kingdom.

Received: 13 February 2024 Accepted: 15 April 2024

Published online: 26 April 2024

## References

- WHO. World malaria report. Geneva: World Health Organization; 2023.
- Division of National Malaria Programmed (DNMP) [Kenya], ICF. Kenya malaria indicator survey 2020. Nairobi: Ministry Health; 2021.
- Ochomo EO, Milanoi S, Abongo B, Onyango B, Muchoki M, Omoko D, et al. Detection of *Anopheles stephensi* mosquitoes by molecular surveillance, Kenya. *Emerg Infect Dis*. 2023;29:2498–508.
- Loughlin SO. The expanding *Anopheles gambiae* species complex. *Pathog Glob Health*. 2020;114:1.
- Barrón MG, Paupy C, Rahola N, Akone-Ella O, Ngangue MF, Wilson-Bahun TA, et al. A new species in the major malaria vector complex sheds light on reticulated species evolution. *Sci Rep*. 2019;9:14753.
- Msugupakulya BJ, Urrio NH, Jumanne M, Ngowo HS, Selvaraj P, Okumu FO, et al. Changes in contributions of different *Anopheles* vector species to malaria transmission in east and southern Africa from 2000 to 2022. *Parasit Vectors*. 2023;16:408.
- Coetzee M, Hunt RH, Wilkerson R, Della Torre A, Coulibaly MB, Besansky NJ. *Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex. *Zootaxa*. 2013;3619:246–74.
- Kyalo D, Amratia P, Mundia CW, Mbogo CM, Coetzee M, Snow RW. A geocoded inventory of anophelines in the Afrotropical Region south of the Sahara: 1898–2016. *Wellcome Open Res*. 2017;2:57.
- Padonou GG, Zoungbédji DM, Sovi A, Salako AS, Konkon AK, Yovogan B, et al. Trophic preferences of *Anopheles coluzzii* (Diptera: Culicidae): what implications for malaria vector control in Benin? *J Med Entomol*. 2023;60:554–63.

10. Djamouko-Djonkam L, Mouchili-Ndam S, Kala-Chouakeu N, Nana-Ndjangwo SM, Kopya E, Sonhafouo-Chiana N, et al. Spatial distribution of *Anopheles gambiae sensu lato* larvae in the urban environment of Yaoundé, Cameroon. *Infect Dis Poverty*. 2019;8:84.
11. Santolamazza F, Calzetta M, Etang J, Barrese E, Dia I, Caccone A, et al. Distribution of knock-down resistance mutations in *Anopheles gambiae* molecular forms in west and west-central Africa. *Malar J*. 2008;7:74.
12. Coetzee M, Craig M, le Sueur D. Distribution of African malaria mosquitoes belonging to the *Anopheles gambiae* complex. *Parasitol Today*. 2000;16:74–7.
13. Derua YA, Alifrangis M, Hosea KM, Meyrowitsch DW, Magesa SM, Pedersen EM, et al. Change in composition of the *Anopheles gambiae* complex and its possible implications for the transmission of malaria and lymphatic filariasis in north-eastern Tanzania. *Malar J*. 2012;11:188.
14. Slotman MA, Della Torre A, Calzetta M, Powell JR. Differential introgression of chromosomal regions between *Anopheles gambiae* and *An. arabiensis*. *Am J Trop Med Hyg*. 2005;73:326–35.
15. Costantini C, Ayala D, Guelbeogo WM, Pombi M, Some CY, Bassole IHN, et al. Living at the edge: biogeographic patterns of habitat segregation conform to speciation by niche expansion in *Anopheles gambiae*. *BMC Ecol*. 2009;9:16.
16. Kamdem C, Tene Fossog B, Simard F, Etouana J, Ndo C, Kengne P, et al. Anthropogenic habitat disturbance and ecological divergence between incipient species of the malaria mosquito *Anopheles gambiae*. *PLoS ONE*. 2012;7: e39453.
17. Tene Fossog B, Antonio-Nkondjio C, Kengne P, Njikou F, Besansky NJ, Costantini C. Physiological correlates of ecological divergence along an urbanization gradient: Differential tolerance to ammonia among molecular forms of the malaria mosquito *Anopheles gambiae*. *BMC Ecol*. 2013;13:1.
18. Tene Fossog B, Ayala D, Acevedo P, Kengne P, Ngomo Abeso Mebuy I, Makanga B, et al. Habitat segregation and ecological character displacement in cryptic African malaria mosquitoes. *Evol Appl*. 2015;8:326–45.
19. De Silva PM, Marshall JM. Factors contributing to urban malaria transmission in sub-saharan Africa: a systematic review. *J Trop Med*. 2012;2012: 819563.
20. Sinka ME, Rubio-Palis Y, Manguin S, Patil AP, Temperley WH, Gething PW, et al. The dominant *Anopheles* vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic précis. *Parasit Vectors*. 2010;3:117.
21. Killeen GF, Seyoum A, Sikaala C, Zomboko AS, Gimnig JE, Govella NJ, et al. Eliminating malaria vectors. *Parasit Vectors*. 2013;6:172.
22. Kelly-Hope L, Ranson H, Hemingway J. Lessons from the past: managing insecticide resistance in malaria control and eradication programmes. *Lancet Infect Dis*. 2008;8:387–9.
23. WHO. Global vector control response 2017–2030: an integrated approach for the control of vector borne diseases. Geneva: World Health Organization; 2017.
24. Mustapha AM, Musembi S, Nyamache AK, Machani MG, Kosgei J, Wamuyu L, et al. Secondary malaria vectors in western Kenya include novel species with unexpectedly high densities and parasite infection rates. *Parasit Vectors*. 2021;14:252.
25. Lobo NF, St. Laurent B, Sikaala CH, Hamainza B, Chanda J, Chinula D, et al. Unexpected diversity of *Anopheles* species in Eastern Zambia: implications for evaluating vector behavior and interventions using molecular tools. *Sci Rep*. 2015;5:17952.
26. Githinji S, Noor AM, Malinga J, Macharia PM, Kiptui R, Omar A, et al. A national health facility survey of malaria infection among febrile patients in Kenya, 2014. *Malar J*. 2016;15:591.
27. Pfeffer DA, Lucas TCD, May D, Harris J, Rozier J, Twohig KA, et al. MalariaAtlas: an R interface to global malariometric data hosted by the Malaria Atlas Project. *Malar J*. 2018;17:352.
28. Gillies MT, Coetzee M. A supplement to the Anophelinae of Africa South of the Sahara. *Publ South African Inst Med Res*. 1987;55:1–143.
29. Neafsey DE, Waterhouse RM, Abai MR, Aganezov SS, Alekseyev MA, Allen JE, et al. Highly evolvable malaria vectors: the genomes of 16 *Anopheles* mosquitoes. *Science*. 2015;347:1258522.
30. Neafsey DE, Lawnciczak MKN, Park DJ, Redmond SN, Coulibaly MB, Traoré SF, et al. SNP genotyping defines complex gene-flow boundaries among African malaria vector mosquitoes. *Science*. 2010;330:514–7.
31. Fontaine MC, Pease JB, Steele A, Waterhouse RM, Neafsey DE, Sharakhov IV, et al. Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science*. 2015;347:1258524.
32. The *Anopheles gambiae* 1000 Genomes Consortium (2021): Ag1000G phase 3 SNP data release. MalariaGEN. [https://www.malariagen.net/data\\_package/ag1000g-phase3-snp/](https://www.malariagen.net/data_package/ag1000g-phase3-snp/).
33. Hudson RR, Slatkint M, Maddison WP. Estimation of gene flow from DNA sequence data. *Genetics*. 1992;132:583–9.
34. Clarkson CS, Miles A, Harding NJ, Lucas ER, Batten CJ, Amaya-Romero JE, et al. Genome variation and population structure among 1142 mosquitoes of the African malaria vector species *Anopheles gambiae* and *Anopheles coluzzii*. *Genome Res*. 2020;30:1533–46.
35. Clarkson CS, Weetman D, Essandoh J, Yawson AE, Maslen G, Manske M, et al. Adaptive introgression between *Anopheles* sibling species eliminates a major genomic island but not reproductive isolation. *Nat Commun*. 2014;5:4248.
36. Miles A, Harding NJ, Bottà G, Clarkson CS, Antão T, Kozak K, et al. Genetic diversity of the African malaria vector *Anopheles gambiae*. *Nature*. 2017;552:96–100.
37. Norris LC, Main BJ, Lee Y, Collier TC, Fofana A, Cornel AJ, et al. Adaptive introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated bed nets. *Proc Natl Acad Sci USA*. 2015;112:815–20.
38. Clarkson CS, Miles A, Harding NJ, O'Reilly AO, Weetman D, Kwiatkowski D, et al. The genetic architecture of target-site resistance to pyrethroid insecticides in the African malaria vectors *Anopheles gambiae* and *Anopheles coluzzii*. *Mol Ecol*. 2021;30:5303–17.
39. Mitchell SN, Rigden DJ, Dowd AJ, Lu F, Wilding CS, Weetman D, et al. Metabolic and target-site mechanisms combine to confer strong DDT resistance in *Anopheles gambiae*. *PLoS ONE*. 2014;9: e92662.
40. Grau-Bové X, Lucas E, Pipini D, Rippon E, van't Hof AE, Constant E, et al. Resistance to pirimiphos-methyl in West African *Anopheles* is spreading via duplication and introgression of the ace1 locus. *PLoS Genet*. 2021;17: e1009253.
41. Elanga-Ndille E, Nouage L, Ndo C, Binyang A, Assatse T, Nguiffo-Nguete D, et al. The g119s acetylcholinesterase (Ace-1) target site mutation confers carbamate resistance in the major malaria vector *Anopheles gambiae* from Cameroon: a challenge for the coming IRS implementation. *Genes (Basel)*. 2019;10:790.
42. Futami K, Dida GO, Sonye GO, Lutiali PA, Mwanja MS, Wagalla S, et al. Impacts of insecticide treated bed nets on *Anopheles gambiae* s.l. populations in Mbita district and Suba district, Western Kenya. *Parasit Vectors*. 2014;7:63.
43. Ndenga B, Githeko A, Omukunda E, Munyekenye G, Atieli H, Wamai P, et al. Population dynamics of malaria vectors in Western Kenya highlands. *J Med Entomol*. 2006;43:200–6.
44. Bayoh MN, Akhwale W, Omok M, Sang D, Engoki SC, Koros D, et al. Malaria in Kakuma refugee camp, Turkana, Kenya: facilitation of *Anopheles arabiensis* vector populations by installed water distribution and catchment systems. *Malar J*. 2011;10:149.
45. Scott JA, Brogdon WG, Collins FH. Identification of single specimens of the *Anopheles gambiae* complex by the polymerase chain reaction. *Am J Trop Med Hyg*. 1993;49:520–9.
46. Meredith HR, Wesolowski A, Menya D, Esimit D, Lokoel G, Kipkoeh J, et al. Epidemiology of *Plasmodium falciparum* infections in a semi-arid rural African setting: evidence of reactive case detection in Northwestern Kenya. *Am J Trop Med Hyg*. 2021;105:1076–84.
47. Simard F, Ayala D, Kamdem GC, Pombi M, Etouana J, Ose K, et al. Ecological niche partitioning between *Anopheles gambiae* molecular forms in Cameroon: the ecological side of speciation. *BMC Ecol*. 2009;9:17.
48. Ayala D, Zhang S, Chateau M, Fouet C, Morlais I, Costantini C, et al. Association mapping desiccation resistance within chromosomal inversions in the African malaria vector *Anopheles gambiae*. *Mol Ecol*. 2019;28:1333–42.
49. Gueye OK, Tchouakui M, Dia AK, Faye MB, Ahmed AA, Wondji MJ, et al. Insecticide resistance profiling of *Anopheles coluzzii* and *Anopheles gambiae* populations in the southern senegal: role of target sites and metabolic resistance mechanisms. *Genes (Basel)*. 2020;11:1403.
50. Ibrahim SS, Muhammad A, Hearn J, Weedall GD, Nagi SC, Mukhtar MM, et al. Molecular drivers of insecticide resistance in the Sahelo-Sudanian populations of a major malaria vector *Anopheles coluzzii*. *BMC Biol*. 2023;21:125.

51. Omotayo AI, Ande AT, Oduola AO, Adelaja OJ, Adesalu O, Jimoh TR, et al. Multiple insecticide resistance mechanisms in urban population of *Anopheles coluzzii* (Diptera: culicidae) from Lagos, South-West Nigeria *Acta Trop.* 2022;227: 106291.
52. Kientega M, Clarkson CS, Traoré N, Hui TYJ, O'Loughlin S, Millogo A, et al. Whole-genome sequencing of major malaria vectors reveals the evolution of new insecticide resistance variants in a longitudinal study in Burkina Faso. *bioRxiv.* 2023;2023.11.20.567800.
53. Williams J, Cowlshaw R, Sanou A, Ranson H, Grigoraki L. In vivo functional validation of the V402L voltage gated sodium channel mutation in the malaria vector *An. gambiae*. *Pest Manag Sci.* 2022;78:1155–63.
54. U.S. President's Malaria Initiative. The PMI Vectorlink Project Annual Report. Rockville, USA, 2018. <https://pmivectorlink.org/wp-content/uploads/2023/07/Year-1-PMI-VectorLink-Annual-Report.pdf>
55. Ondeto BM, Nyundo C, Kamau L, Muriu SM, Mwangangi JM, Njagi K, et al. Current status of insecticide resistance among malaria vectors in Kenya. *Parasit Vectors.* 2017;10:429.
56. Protopopoff N, Mosha JF, Lukole E, Charlwood JD, Wright A, Mwalimu CD, et al. Effectiveness of a long-lasting piperonyl butoxide-treated insecticidal net and indoor residual spray interventions, separately and together, against malaria transmitted by pyrethroid-resistant mosquitoes: a cluster, randomised controlled, two-by-two fact. *Lancet.* 2018;391:1577–88.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Dr. Luna Kamau** is a Senior Principal Research Scientist and Deputy Director, Centre for Biotechnology Research and Development at the Centre for Biotechnology Research and Development, Kenya Medical Research Institute in Nairobi Kenya. Her research focuses on molecular and population genetic of disease vectors, vector ecology and behaviour, vector surveillance and insecticide resistance, development of tools and strategies for disease control and emerging infectious diseases.

**Dr. Kelly L. Bennett** is a Senior Data Scientist at the Genomic Surveillance Unit of the Wellcome Sanger Institute. Her research focuses on the ecology and evolution of mosquito and arboviral vectors using genomic technology.