Whole-genome sequencing of major malaria vectors reveals the evolution of new insecticide resistance variants in a longitudinal study in Burkina Faso

Mahamadi Kientega[1,3*], Chris S. Clarkson[2], Nouhoun Traoré[1,3], Tin-Yu J. Hui[4], Samantha O'Loughlin[4], Abdoul-Azize Millogo[1,5], Patric Stephane Epopa[1], Franck A. Yao[1], Adrien M. G. Belem[3], Jon Brenas[2], Alistair Miles[2], Austin Burt[4] and Abdoulaye Diabaté[1*]

Abstract

Background Intensive deployment of insecticide based malaria vector control tools resulted in the rapid evolution of phenotypes resistant to these chemicals. Understanding this process at the genomic level is important for the deployment of successful vector control interventions. Therefore, longitudinal sampling followed by whole genome sequencing (WGS) is necessary to understand how these evolutionary processes evolve over time. This study investigated the change in genetic structure and the evolution of the insecticide resistance variants in natural populations of *Anopheles gambiae* over time and space from 2012 to 2017 in Burkina Faso.

Methods New genomic data have been generated from *An. gambiae* mosquitoes collected from three villages in the western part of Burkina Faso between 2012 and 2017. The samples were whole-genome sequenced and the data used in the *An. gambiae* 1000 genomes (Ag1000G) project as part of the Vector Observatory. Genomic data were analysed using the analysis pipeline previously designed by the Ag1000G project.

Results The results showed similar and consistent nucleotide diversity and negative Tajima's D between *An. gambiae* sensu stricto (s.s.) and *Anopheles coluzzii*. Principal component analysis (PCA) and the fixation index (F_{ST}) showed a clear genetic structure in the *An. gambiae* sensu lato (s.l.) species. Genome-wide F_{ST} and H12 scans identified genomic regions under divergent selection that may have implications in the adaptation to ecological changes. Novel voltage-gated sodium channel pyrethroid resistance target-site alleles (*V402L, I1527T*) were identified at increasing frequencies alongside the established alleles (*Vgsc-L995F, Vgsc-L995S* and *N1570Y*) within the *An. gambiae* s.l. populations. Organophosphate metabolic resistance markers were also identified, at increasing frequencies, within the *An. gambiae* s.s. populations from 2012 to 2017, including the SNP *Ace1-G280S* and its associated duplication. Variants simultaneously identified in the same vector populations raise concerns about the long-term efficacy of new generation bed nets and the recently organophosphate pirimiphos-methyl indoor residual spraying in Burkina Faso.

*Correspondence:
Mahamadi Kientega
mkient54@gmail.com
Abdoulaye Diabaté
npiediab@gmail.com
Full list of author information is available at the end of the article

Kientega *et al. Malaria Journal*   (2024) 23:280

Page 2 of 16

Conclusion These findings highlighted the benefit of genomic surveillance of malaria vectors for the detection of new insecticide resistance variants, the monitoring of the existing resistance variants, and also to get insights into the evolutionary processes driving insecticide resistance.

Keywords Insecticide resistance, Genomic surveillance, *An. gambiae*, Malaria

Background

Malaria remains a major public health problem in Africa [1]. *Anopheles gambiae* sensu lato (s.l.), the main malaria vector in Africa, exhibits high phenotypic and genotypic diversity, allowing rapid adaptation to environmental changes, such as the introduction of new insecticidal control methods [2]. In Burkina Faso, the major malaria vectors *An. gambiae* sensu stricto (s.s.), *Anopheles coluzzii* and *Anopheles arabiensis*, members of the *An. gambiae* complex, can be found living in sympatry, with divergent resting and feeding behaviours [3, 4]. These species, through their ability to colonize different ecological settings, their preference for human blood feeding, and their high susceptibility to a malaria parasite (*Plasmodium falciparum*) infection, are responsible for around 90% of the malaria burden in Burkina Faso and the whole sub-Saharan Africa [5, 6].

The National Malaria Control Programme (NMCP) of Burkina Faso and their partners have invested a lot to increase the coverage of long-lasting insecticidal bed nets (LLINs) and other control tools such as indoor residual spraying (IRS) of insecticides to drop the malaria transmission line in the country. The coverage of LLINs in households increased significantly in the country over the sampling periods, from 5.6% in 2003 to about 80% in 2014, reaching 83% in 2021 [7, 8]. The recent 2019 and 2022 bed net campaigns involved the free distribution of three types of LLINs, especially the piperonyl butoxide (PBO)-synergist, the dual-AI Interceptor G2 (containing the pyrrole insecticide, chlorfenapyr) and the standard LLINs [9]. Despite these achievements, malaria remains a countrywide health problem responsible in 2021 for about 12.2 million cases and 4355 deaths [10].

In Burkina Faso and most sub-Saharan countries, the primary vector surveillance strategies relied on the routine tracking of the vector species density and distribution and the monitoring of phenotypic insecticide resistance (IR) status. Unfortunately, these methods proved unable to adequately explain the genetic and evolutionary processes driving the increasing spread of IR variants and changes in vector behaviour [11]. Recent advances in genomic sequencing technologies, both price and throughput, and new cloud native analysis tools [12] have made genomic surveillance of malaria vectors possible. Genomic surveillance has advantages to provide insights on the genetic and evolutionary phenomena of vectors that could have impact on vector control strategies. Towards this aim, the *Anopheles gambiae* 1000 genomes (Ag1000G) project driven by the MalariaGEN Vector Observatory sequenced more than 15,000 *Anopheles* mosquitoes collected in 25 African countries [13, 14]. A major goal of this project was to support malaria elimination efforts by providing a high quality open access data resource on natural genetic variation within *An. gambiae* s.s., *An. coluzzii* and *An. arabiensis* populations, both for the public health and research communities [15].

Previous studies investigating IR using these genomic data have highlighted associated genetic variants distributed throughout the mosquito genome [16–19]. The phenotypic effect of many of these variants has been demonstrated through *in-vivo* and *in-vitro* experiments and their effects are evident in the low mortality rate observed during the IR bioassays, and in the reduction of the efficacy of the insecticide-based control tools [20–22]. The strength of these resistance phenotypes and the wide geographical distribution of these variants raise concerns about the durability of the new insecticide-based vector control tools being used to overcome resistance, including indoor residual spraying (IRS) with organophosphate insecticides, such as Actellic (pirimiphos-methyl) and pyrethroid bed-nets treated with metabolic resistance blocking piperonyl butoxide (PBO) synergist.

There is an increasingly urgent need for sustainable strategies and effective tools to support malaria elimination efforts in the face of ever evolving vectors. Research of alternative and innovative tools such as genetic modified mosquitoes, new insecticidal compounds as well as natural symbionts has increased in recent years, giving a glimmer of hope to people living in endemic areas [23, 24]. However, given the rapid evolution of the of the malaria vector genome, as demonstrated by many studies [13, 25, 26], any implementation of vector control tools requires a better understanding of the genomic variation of vector species involved, in particular an understanding of the evolutionary processes, emergence and spread of IR-associated variants in natural populations over time. Thus, longitudinal sampling of vectors, followed by whole-genome population genomics becomes a very helpful tool upstream and downstream of the implementation of the control tools to provide insights on the genetic and evolutionary phenomena of vectors that could have any impact on vector control strategies.

Kientega *et al. Malaria Journal*   (2024) 23:280

Page 3 of 16

This study used the genomic data from 1409 *An. gambiae* s.l. samples, collected in Burkina Faso between 2012 and 2017 and sequenced as part of the *Anopheles gambiae* 1000 Genomes Project (2017, 2020) to characterize the spread of known IR variants over space and time, while analysing these natural populations for evidence of the evolution of novel IR mechanisms that could threaten current or future vector control methods.

Methods

Mosquitoes sampling

Mosquito samples were collected by the Target Malaria project in three villages, Bana (11.233, − 4.472), Souroukoudinga (11.235, − 4.535) and Pala (11.150, − 4.235), located in the western part of Burkina Faso (Fig. 1) during the rainy season from 2012 to 2017. These villages have very similar ecological and climatic conditions, i.e. a humid savannah zone surrounding the Bobo-Dioulasso city and are characterized by two seasons: a rainy period from June to October and a dry season from November to April. The average annual rainfall is above 800 mm, with an average annual temperature and humidity of about ∼ 27 °C and ∼ 60%, respectively [27]. The main agricultural practices are cotton and cereal (incl. maize, rice, sorghum) cultivation utilizing large amount of pesticides [28].

Malaria remains endemic in the area surrounding the Bobo-Dioulasso city, causing 338 deaths in the three health districts of the city [10]. The *Anopheles gambiae* complex species remain the major malaria vector in the sampling area [27]. Malaria control relies on the anti-malarial drugs and prevention of mosquito bites using long-lasting insecticidal nets (LLINs), the most widespread vector control tool, and to a lesser extent indoor residual spraying (IRS), sprays and mosquito-repellent coils.

Mosquitoes were collected during the rainy from 2012 to 2017 using three collection methods: Pyrethroid spray catches (PSC), human landing catches (HLC) and swarm collection [29]. PSC and HLC were used for female mosquito collection, whereas males were collected in swarms. After collection, mosquito samples were morphologically identified using the morphological keys [30] and *An. gambiae* s.l. specimens were stored in 80% ethanol and then shipped for the whole genome sequencing at the Wellcome Sanger Institute.

Whole genome sequencing and genomic data management

Anopheles gambiae s.l. mosquitoes were sent for individual whole genome sequencing at high coverage (30 ×) using Illumina technology at the Wellcome Sanger Institute as part of the vector observatory. Genomic data were analysed at the MalariaGEN Resource Centre following the previous data analyses and quality clean-up pipelines designed by the Ag1000G project. Briefly, the analyses consisted of the raw data cleaning, the quality control and the mapping to the AgamP4 reference genome using

Fig. 1 Map of the sampling sites around Bobo-Dioulasso city

Kientega *et al. Malaria Journal*   (2024) 23:280

Page 4 of 16

BWA version 0.7.15. Indel realignment was performed using GATK version 3.7-0 RealignerTargetCreator and IndelRealigner. Single nucleotide polymorphisms were called using GATK version 3.7-0 UnifiedGenotyper. Genotypes were called for each sample independently, in genotyping mode, given all possible alleles at all genomic sites where the reference base was not "N". Complete specifications of the alignment and genotyping pipelines are available from the malariagen/pipelines GitHub repository [31]. After variant calling, the raw sequences in FASTQ format and the aligned sequences in BAM format were stored in the European Nucleotide Archive (ENA, Study Accession n° PRJEB42254). The genetic variants in VCF and zarr formats, including the samples' metadata, were stored on Google Cloud and are accessible via the malariagen_data package [32] or are directly downloadable. Full details about the sequencing technology, the raw data clean-up and quality control, the SNPs calling, haplotypes phasing and copy number variants identification including the storage of the genomic data are available in the MalariaGEN homepage [31].

Population genomic and insecticide resistance variant analyses

Genetic diversity and population structure were analyzed using Jupyter Notebook in the Google Colab platform. The genetic variants (SNPs, haplotypes and CNVs) and the reference genome of *An. gambiae* s.l. were accessed using the malariagen_data package for direct analyses of the genomic data in the cloud. This package provides a wide range of functions and properties for cloud data access and for exploratory analysis of large scale genetic variation data [32]. Additional packages such as scikit-allel [33] and other python standard data-management packages were used for the data analyses and the plotting. The number of segregating sites (number of SNPs) was calculated in the whole genome (X, 2RL, 3RL) of *An. gambiae* s.l. The diversity statistics [nucleotide diversity (θ_π), Tajima's D (D), Watterson theta (θ_w), allele frequency spectrum (sfs)] were calculated using the SNPs called in the 3L chromosome in all the mosquito populations. Tukey multiple comparison test was applied to compare the diversity statistic between vector populations over the sampling periods.

Genome wide selection scan was performed using the Garud H12 [34] statistic to detect the genomic region under positive selection. H12 is a haplotype statistic that combines the first and second most frequent haplotype frequencies in a sample into a single combined haplotype frequency and then recalculates haplotype homozygosity using this revised haplotype frequency distribution [34]. F_{ST} and PCA analyses were performed in the 3L chromosome to investigate the genetic structure and the differentiation of *An. gambiae* s.l. populations. The 3L chromosomal arm was used for the genetic diversity and population structure analyses because this region has no polymorphic inversions [35]. The extent and the spread of insecticide resistance variants were also investigated in the *An. gambiae* s.l. populations using the malariagen_data python package. Time series (from 2012 to 2017) SNPs frequencies were analysed in known target-site resistance associated genes, *VGSC* and *ACE1*. The copy number variation (CNV) frequencies were also investigated in genomic regions associated with metabolic resistance, cytochrome P450s, glutathione-s-transferases, acetylcholinesterase, diacylglycerol kinase and carboxylesterases.

Results

Population sampling

Anopheles gambiae mosquitoes collected in Burkina Faso from 2012 to 2017 were sequenced as part of the follow-up project, the MalariaGEN Vector Observatory. In total, 1409 *An. gambiae* s.l. mosquitoes (978 females and 431 males) were collected from three villages (Bana, Souroukoudinga, Pala) surrounding the Bobo-Dioulasso city in the west of Burkina Faso and shipped for deep whole-genome sequenced mosquitoes. The longitudinal mosquito sampling survey was carried out in the three villages during the rainy seasons of the years 2012, 2014, 2015, 2016 and 2017 (Table 1). The species composition of the sample set varied between the sampling sites and the time-points (Table 1). *Anopheles coluzzii* was the most predominant species found, accounting for 52.73% [743/1409] of samples, followed by *An. gambiae* s.s. with 38.96% (549/1409), while *An. arabiensis* only comprised 8.16% (115/1409) of samples collected. Two *An. gambiae* s.s. and *An. coluzzii* hybrids were also identified. As shown by a previous work, *An. coluzzii* remains the most predominant species in Bana (83.87% [489/583]) and Souroukoudinga (60.90% [243/399]) while *An. gambiae* s.s. is prevalent in Pala (70.72% [302/427]) [27].

Genetic diversity

Understanding genetic diversity allows us to compare demography of species over time and space. Analyses of these data showed around 27.94 million of segregating sites [~91% of biallelic SNPs (Single Nucleotide Polymorphism)] in the *An. gambiae* s.l. genome. Approximately 25.68 million of SNPs were identified in the autosomes and 2.262 million of SNPs in the X chromosome. Analyses to assess the genetic diversity (see Methods section) showed almost similar nucleotide diversity (θ_π), Tajima's D and Watterson's theta (θ_w) over all the sampling periods in *An. gambiae* and *An. coluzzii*, as in most natural populations [36] across these species range (Fig. 2), possibly

Kientega *et al. Malaria Journal*   (2024) 23:280

Page 5 of 16

Table 1 Species composition of the *An. gambiae* complex of the samples collected in the three villages surrounding the city of Bobo-Dioulasso during the sampling period (2012–2017)

	An. arabiensis		*An. coluzzii*		*An. gambiae s.s*		gambiae_coluzzii	
	F	M	F	M	F	M	F	M
Bana	1 [0.17%]	0	379 [65.01%]	110 [18.87%]	29 [4.97%]	63 [10.81%]	0	1 [0.17%]
2012	0	0	38 [58.46%]	4 [6.15%]	5 [7.69%]	17 [26.15%]	0	1 [1.54%]
2014	1 [0.78%]	0	63 [48.84%]	33 [25.58%]	13 [10.08%]	19 [14.73%]	0	0
2015	0	0	95 [78.51%]	1 [0.83%]	8 [6.61%]	17 [14.05%]	0	0
2016	0	0	96 [72.18%]	24 [18.05%]	3 [2.26%]	10 [7.52%]	0	0
2017	0	0	87 [64.44%]	48 [35.56%]	0	0	0	0
Pala	78 [18.27%]	35 [8.2%]	11 [2.58%]	0	236 [55.27%]	66 [15.46%]	1 [0.23%]	0
2012	0	0	11 [18.64%]	0	40 [67.8%]	8 [13.56%]	0	0
2014	6 [6.0%]	5 [5.0%]	0	0	68 [68.0%]	21 [21.0%]	0	0
2015	58 [35.37%]	30 [18.29%]	0	0	61 [37.2%]	15 [9.15%]	0	0
2016	14 [20.90%]	0	0	0	31 [46.27%]	22 [32.84%]	0	0
2017	0	0	0	0	36 [97.3%]	0	1 [2.7%]	0
Souroukoudinga	1 [0.25%]	0	141 [35.34%]	102 [25.56%]	101 [25.31%]	54 [13.53%]	0	0
2012	0	0	29 [50.88%]	0	28 [49.12%]	0	0	0
2014	0	0	44 [44.44%]	6 [6.06%]	34 [34.34%]	15 [15.15%]	0	0
2015	0	0	35 [27.56%]	52 [40.94%]	24 [18.9%]	16 [12.6%]	0	0
2016	0	0	30 [34.88%]	18 [20.93%]	15 [17.44%]	23 [26.74%]	0	0
2017	1 [3.33%]	0	3 [10.0%]	26 [86.67%]	0	0	0	0
Total	80 [5.68%]	35 [2.48%]	531 [37.69%]	212 [15.05%]	366 [25.98%]	183 [12.99%]	1 [0.07%]	1 [0.07%]

gambiae_coluzzii: hybrid of *An. gambiae* s.s. and *An. coluzzii*; F: female; M: male

indicating similar demographic histories between these two species. In contrast, *An. arabiensis* showed lower diversity, though with apparent increase over the sampling periods. The diversity findings were reflected in a genome-wide negative Tajima's D, caused by an excess of rare variants due to either positive selection and/or rapid demographic changes through population expansion. In addition, the allelic frequency spectrum (Fig. S1) also showed the same trend highlighting the excess of low frequency variants in the *An. gambiae* s.l. genome.

Population structure

Understanding population structure between, and within, vector species is important not only because of its implications for gene flow of medically relevant alleles or the role of gene drive control methods, but also because hidden structure will confound downstream genomic analyses of the population. Consequently, the genetic structure and the patterns of gene flow were investigated using SNPs on the 3L chromosome (see Methods section). Pairwise F_{ST} was estimated between and within species collected in the different villages and was found to be low (0 to 0.05) between *An. gambiae* and *An. coluzzii* (Fig. 3A). These results support other evidence for ongoing gene flow

between *An. coluzzii* and *An. gambiae* s.s. [37]. However, for *An. arabiensis*, results showed relatively higher F_{ST} between both *An. coluzzii* and *An. gambiae*, ranging from 0.26 to 0.30 (Fig. 3A). No differentiation ($F_{ST} \sim 0$) was observed between populations of the same species collected in different villages, suggesting no population structure on this scale.

The principal component analysis (PCA) showed a clear divergence between the *An. gambiae s.l.* species. Examining the first two principal components, samples fell into three clusters, where individuals of the same species showed higher genetic similarities among themselves compared to individuals of other species. No sub-clusters were observed within each species (Fig. 3B). No change was observed in the population structure of each species throughout the sampling periods, and the sampling sites (Fig. S2); however, some other genomic regions were shown to be under divergent selection (see Organophosphate and carbamate resistance: increasing threats to control plus a potential novel IR mechanism, Fig. S3). The lack of genetic structure (low $F_{ST} \sim 0.05$) between *An. coluzzii* and *An. gambiae* s.s. in the sampling area suggests showed that these species may share advantageous allelic variants, such as those involved in insecticide resistance.

Kientega *et al. Malaria Journal*   (2024) 23:280

Page 6 of 16
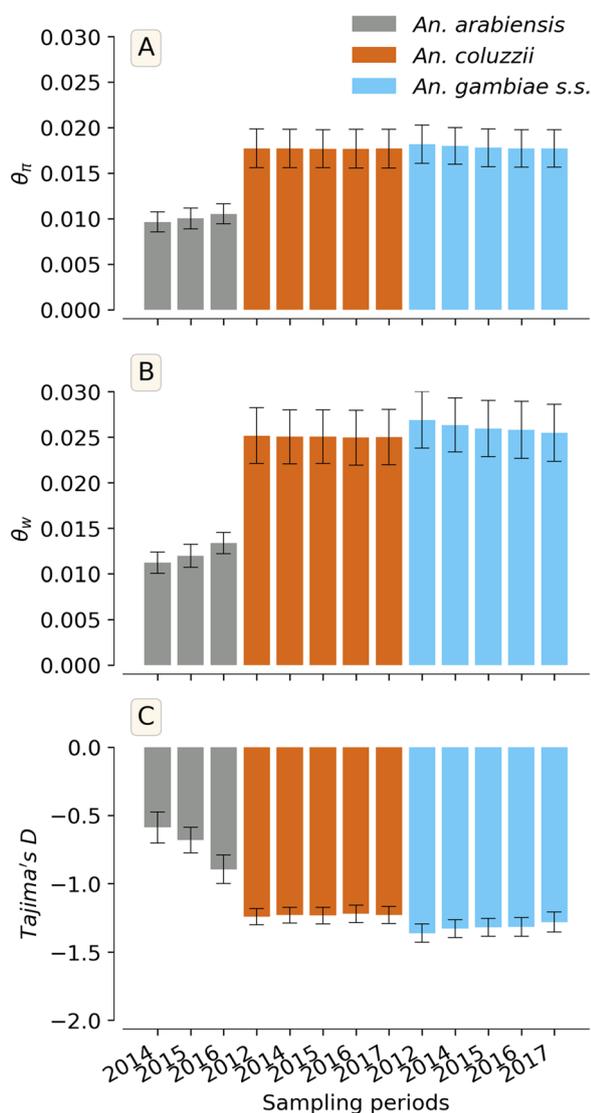
Fig. 2 Average of genetic diversity statistics within the *An. gambiae* s.l. populations in the study sites from 2012 to 2017; The Y-axis shows the diversity statistic, the X-axis shows the sampling periods and the colours show the populations. **A** Nucleotide diversity (θ_π), **B** Watterson theta (θ_w), **C** Tajima's D, error bars correspond to the 95% confidence interval for each year

Insecticide resistance

Pyrethroid target-site resistance: concerning increase of new mutations

Pyrethroid target-site resistance in *An. gambiae* s.l. has, historically, been primarily driven by two main *"kdr (knock down resistance)"* SNPs, causing the amino acid changes *vgsc-L995F* and *vgsc-L995S* [18]. These mutations are widespread and are present at high frequencies among populations in many regions in Africa [13]. Improvements in sequencing technologies have allowed

the identification of many additional SNPs in the voltage gated sodium channel gene *(VGSC [2L: 2358158–2431617])* gene that could enhance the level of pyrethroid resistance over time [39, 40]. This study identified 634 non-synonymous SNPs within the *VGSC* gene in all the *An. gambiae* s.l. populations (Table S1). The *kdr* allele *vgsc-L995F* was identified at high frequencies (> 30%) in all the populations, including *An. arabiensis*, and was found at fixation in *An. gambiae* s.s. collections since 2012 (Fig. 4). The N1570Y allele, previously shown to increase the level of pyrethroid resistance when in concert with *vgsc-L995F* [40], was also present in both *An. gambiae* s.s. and *An. coluzzii* populations at around 30% frequency, but it was not found in *An. arabiensis*.

A *VGSC* haplotype carrying the apparently linked amino acid substitutions *vgsc-V402L+I1527T* was also identified in Burkina Faso *An. coluzzii* (Figs. 4 and 5). Two alleles of the *vgsc-V402L* were detected, the *2L: 2 391 228 G > C (V402L)* with a relatively constant frequency ranging from 6.7% in 2012 to 5.2% in 2017 and the *2L: 2 391 228 G > T (V402L)* with an evolving frequency from 5.5% in 2012 to 28% in 2017. However, the *vsgc- I1527T* is monoallelic and identified with an evolving frequency 12% in 2012 to 33% in 2017 (Fig. 4). A recent in-vivo investigation has shown that the *vgsc-V402L* substitution results in a pyrethroid resistance phenotype with lower fitness costs than the classic *vgsc-L995F kdr* mutation [39]. Upon discovery of *vgsc-V402L+I1527T,* it was suggested that this haplotype may be a "relic" resistance haplotype, pre-dating and being overtaken by a more effective *vgsc-L995F* carrying haplotype [18]. However, the time series data shows that the inverse appears to be the case. The *vgsc-V402L+I1527T* increased from 12.19% in 2012 to 33.12% in 2017, and the sum of the frequencies of the two *vgsc-V402L* alleles (*2L: 2 391 228 G > {C, T}*) is approximately equal to the frequency of *vgsc-I1527T* during the sampling periods (Figs. 4 and 5). The increase of these alleles coincides with the drop in the frequency of the *vgsc-L995F* mutation in *An. coluzzii* populations from 86.58% in 2012 to 66.46% in 2017 (Figs. 4 and 5). These results and previous works [18, 39] suggest these two alleles are increasing in frequency and seem to replace the *vgsc-L995F* alleles in *An. coluzzii* populations, but had not spread to *An. gambiae* s.s. The increased emergence of new *kdr* alleles alongside the existing *kdr* mutations (*vgsc-L995F* and *vgsc-L995S*) at high frequencies may increase the resistance level of vector populations to pyrethroid compounds.

Pyrethroid metabolic resistance: increasing CNV frequencies.

Gene duplications and deletions (copy number variants—CNVs) are known to influence the expression level of genes by deleting or providing additional
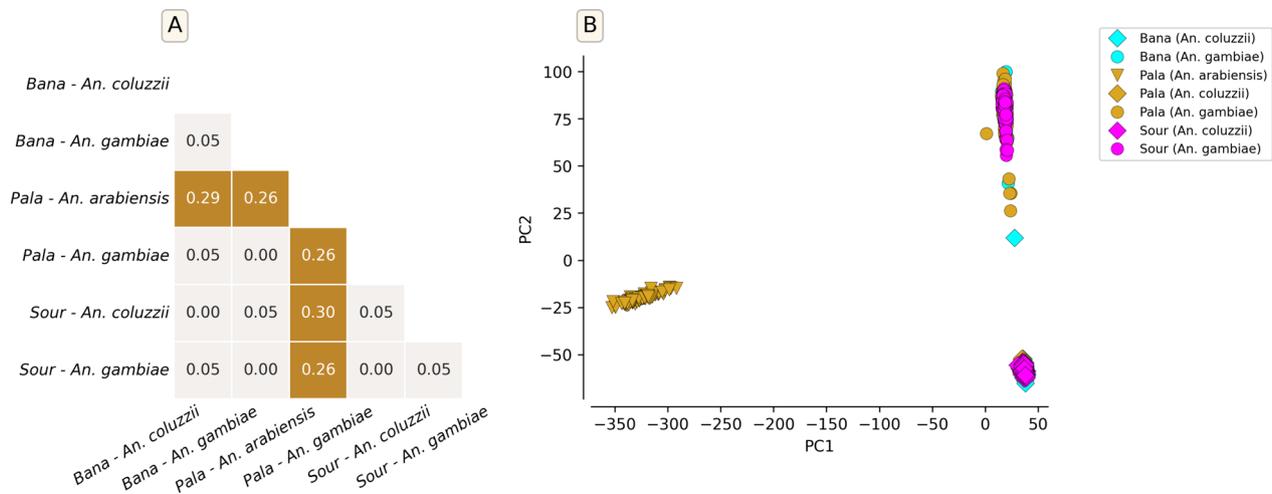
Fig. 3 Genetic differentiation and population structure of *An. gambiae* s.l. using the SNPs identified in the 3L chromosome. Sour: Souroukoudinga; **A** Pairwise F_{ST} between *An. gambiae* s.l. populations collected in different villages; no differentiation between populations of the same species collected in two different villages; **B** PCA showing the genetic structure of *An. gambiae* s.l. populations
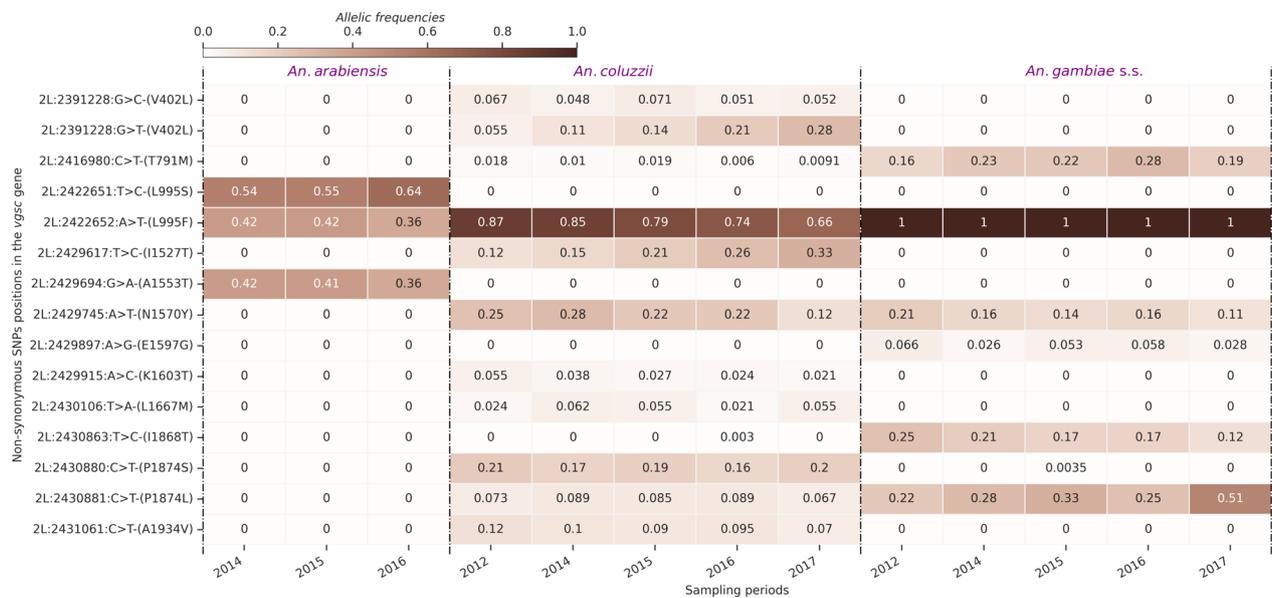
Fig. 4 Heat map showing the frequencies of the non-synonymous SNPs frequencies (max freq > 0.05 in at least one population) in the *VGSC* gene in the *An. gambiae* s.l. populations over time. The X-axis shows the *An. gambiae* s.l. populations and the sampling periods. The Y-axis shows the non-synonymous SNPs positions in the chromosome 2L and the corresponding amino acid change. The gradient colour bar shows the distribution of the allelic frequencies

copies of the original one. Consequently, CNVs of genes involved in xenobiotic detoxification can increase insecticide metabolism or reduce the expression of genes with high fitness cost, producing a resistance phenotype [16]. To help overcome pyrethroid resistance, now commonplace in many natural populations, PBO synergist bed nets have recently been introduced.

These bed nets enhance the pyrethroid efficacy by inhibiting the cytochrome P450 enzymes that can be responsible for metabolic pyrethroid resistance [41]. Often, mechanisms driving the resistance in a given area remain unclear [20], however, whole genome sequencing allows CNV calling, and the detection of resistance associated alleles [17].
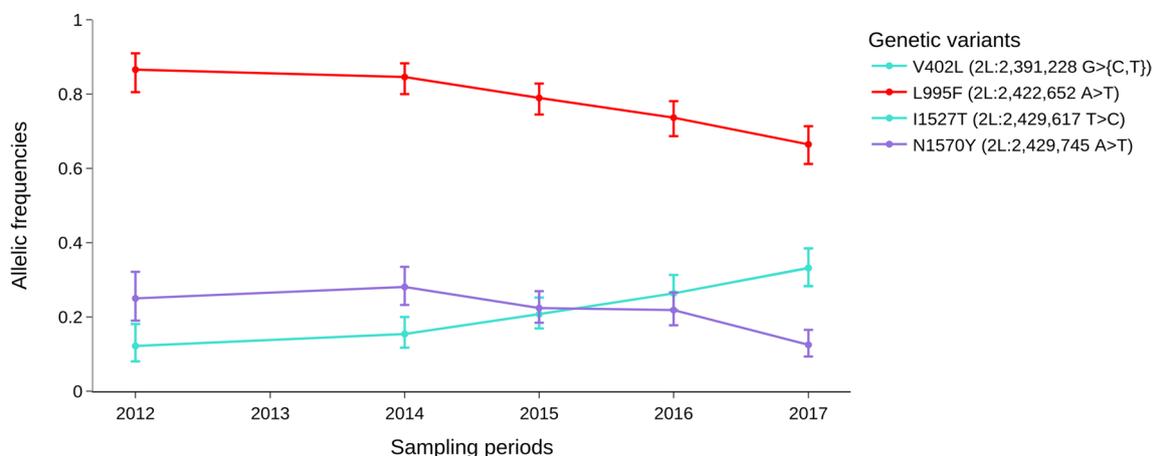
Fig. 5 Dynamics of *vgsc-L995F*, *vgsc-V402L+I1527T* and N1570Y allele frequencies in the *An. coluzzii* populations; *vgsc-V402L* and *vgsc-I1527T* are shown to be linked and exhibited the same frequencies from 2012 to 2017, the X-axis shows the sampling periods and the Y-axis shows the allelic frequencies

CNVs were identified in 136 Cytochrome P450 genes distributed across the whole genome and most of these (~63.97%) were gene amplifications (Table S2). The majority of these CNVs were species-specific, and some were found at increasing frequencies over time in the three species. Some genes showing low CNVs frequencies in 2012 displayed increasing CNVs frequencies in the following years (Fig. 6). This situation was observed in the

CYP6Z cluster of genes (*CYP6Z1*, *CYP6Z2 and CYP6Z3*), known to be associated with pyrethroid resistance [16], whose CNVs frequencies increased from 17% in 2012 to up to 60% in 2017 in *An. gambiae* s.s. populations. Additionally, the *CYP9K1* gene, also previously shown to be involved in pyrethroid resistance, showed CNVs at high frequencies up to 94% from 2012 to 2017. In the *An. coluzzii* populations, CNVs were identified at frequencies
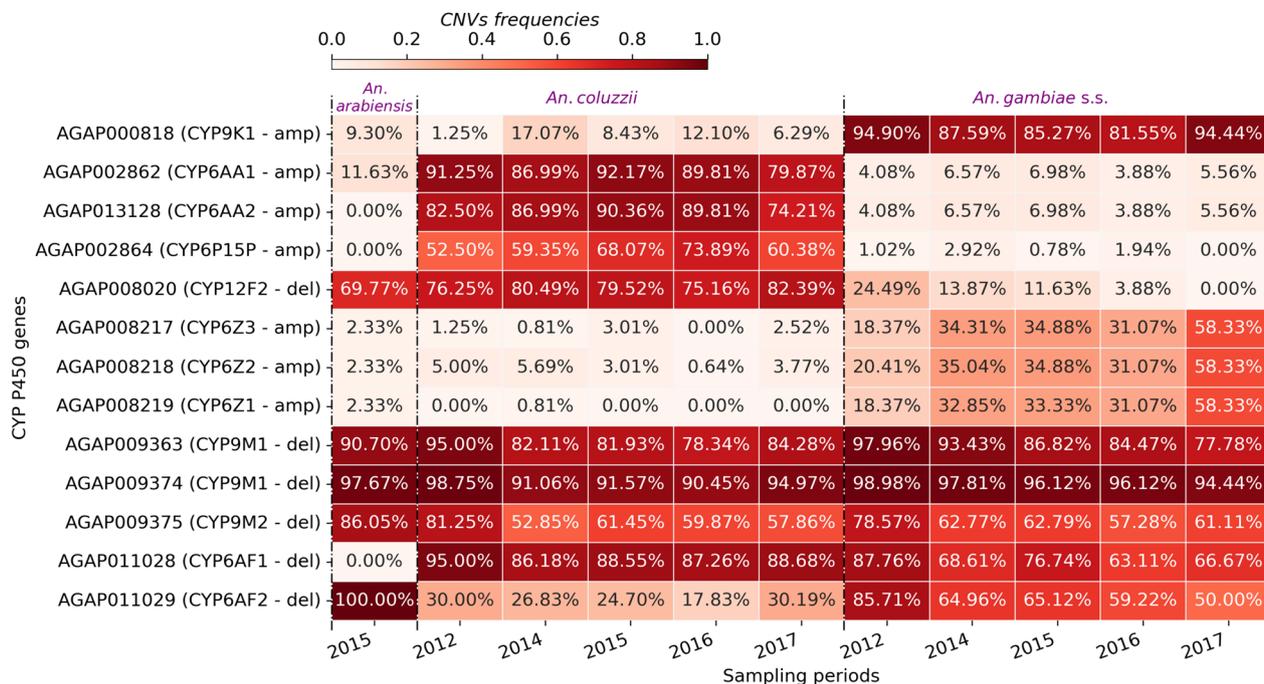
Fig. 6 Heat map of the cytochrome P450 genes showing high CNVs frequencies in the *An. gambiae* s.l. populations. The X-axis shows the *An. gambiae* s.l. populations and the sampling periods. The Y-axis shows the positions of the P450 genes ID and the CNV type (del: deletion or amp: amplification). The gradient colour bar shows the distribution of the CNV frequencies

Kientega *et al. Malaria Journal*   (2024) 23:280

Page 9 of 16

ranging from 50% in 2012 to up to 90% in 2017 within the *CYP6AA1, CYP6AA2, CYP6P15P* and *CYP12F2* genes. CNVs related to gene deletions rate was shown to be high in the vector populations and seems to be associated with metabolic resistance. The results showed approximately 36.03% of CYP P450 genes having CNVs related to gene deletion in the vector populations. Interestingly, some of these genes such as the *CYP6AF* (*CYP6AF1* and *CYP6AF2*) and *CYP9M* cluster (*CYP9M1 and CYP9M2*) were shown to be deleted in 50–100% of the vectors populations except *An. arabiensis* for the gene *CYP6AF1* (Fig. 6). The presence of both CNVs types in the same mosquito populations can increase and/or decrease the expression level of some genes to enhance mosquito fitness response to face the environmental forces like insecticide pressure. Further studies are strongly needed to understand the impact of these new variants in the mosquito fitness and vector control tools.

Organophosphate and carbamate resistance: increasing threats to control plus a potential novel IR mechanism

Indoor residual spraying (IRS) is one of the major vector control strategies in Africa [42]. However, the coverage of this strategy dropped considerably over the sampling periods because of the emergence of resistance to the pyrethroid compounds and also the low residual efficacy of carbamates compounds [43]. Recently, IRS with Actellic 300 CS containing pirimiphos-methyl, an organophosphate as the active ingredient, showed interesting results including a prolonged residual efficacy up to seven months and a reduction of vector density in many sprayed areas in Africa [44–46]. This new formulation demonstrated better ability to target pyrethroid resistant and indoor resting mosquitoes and also showed

additional benefit when used in combination with LLINs [47, 48].

Longitudinal sampling was used to assess the extent of organophosphate and carbamate resistance in Burkina Faso that could threaten the long-term efficacy of pirimiphos-methyl-based IRS. A sliding window genome-wide F_{ST} scan showed a year to year (from 2012 to 2017) genetic difference in some regions of the genome especially on chromosomes 2 and X in *An. gambiae* s.s. and *An. coluzzii* samples. Most of these changes were observed in genomic regions previously shown to be associated with insecticide-resistance suggesting a signal of positive selection in these genes (Figs. 8 and 9).

Both *ACE1 (2R: 3484107–3495790)* and Diacylglycerol Kinase (DGK or AGAP000519 [X: 9215504–9266532]) loci showed no selection signal in 2012 in any vector populations sampled, but by 2017 selection signals were detected at both loci in *An. gambiae s.s.* populations. The increasing gene amplification signal in the *ace1* locus, from 32.65% in 2012 to up to 91.67% in 2017, was corroborated by the increase in the insecticide resistance associated with *ace1-G280S* variant previously shown to be involved in organophosphate and carbamate resistance [49] (from 15.81% in 2012 to 65.27% in 2017), but only in *An. gambiae s.s.* populations (Fig. 7). A total of 243 other non-synonymous SNPs in the *ace1* gene (Table S3). CNV amplifications were also found in *An. coluzzii* populations, but at low and constant frequencies (∼3%), but not in *An. arabiensis* populations.

Genetic variants of the *GSTe* genes have also been reported to be involved in resistance to organophosphates [50]. The analyses identified 41 *GSTe* genes with CNVs including 24 genes with amplifications and 17 genes with deletions. These CNVs are distributed across
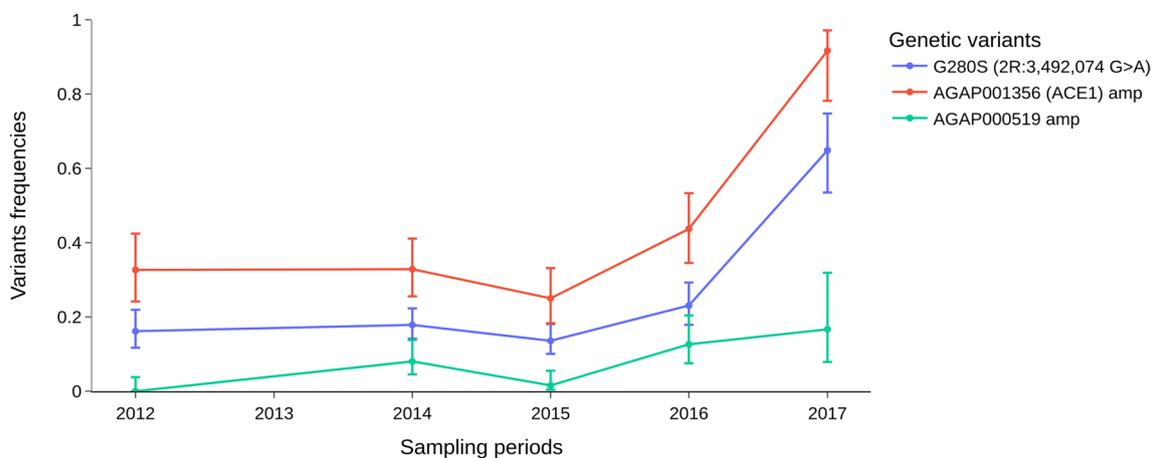
Fig. 7 Evolution of variant frequencies of the *ACE1* (*Ace1-G280S, AGAP001356 (ACE1) amp*) and Diacylglycerol Kinase (*AGAP000519 amp*) genes in the *An. gambiae* s.s. populations. The X axis shows the sampling periods and the Y axis, the allele frequencies of the genetic variants
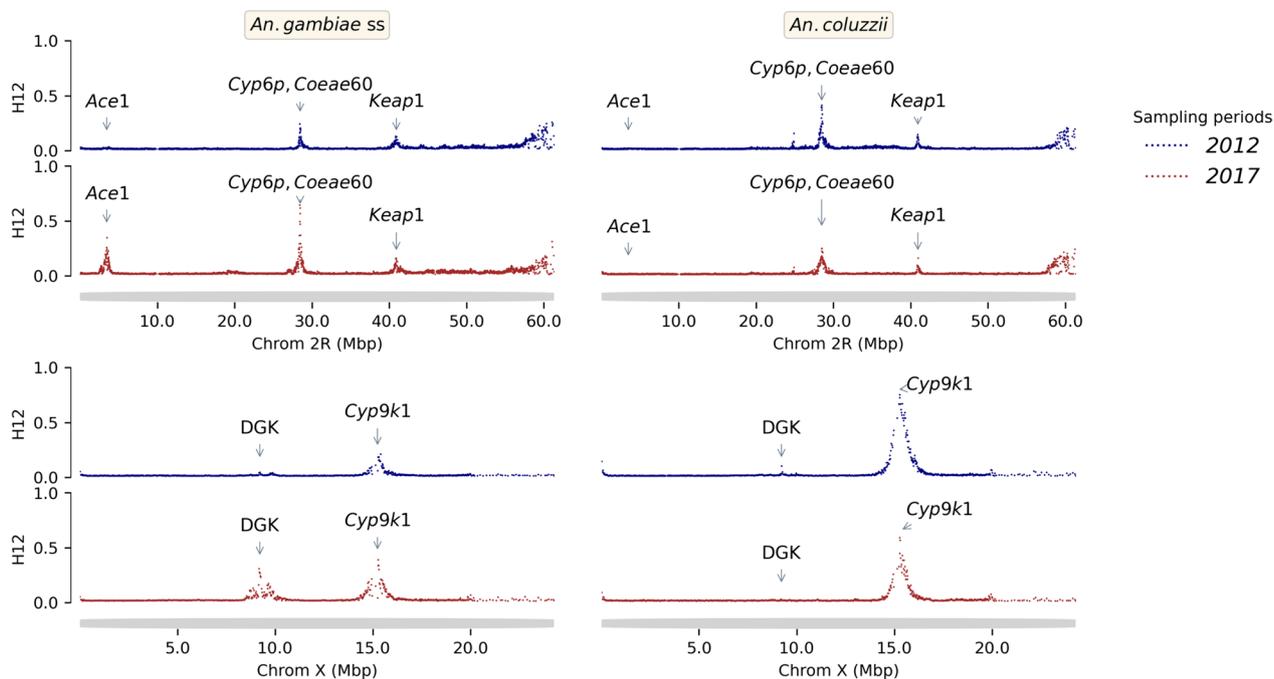
Fig. 8 Genome-wide scan for recent selection using Garud's H12 across chromosome X and arm 2R (5 kb window) of the *An. gambiae* s.s. and *An. coluzzii* populations from 2012 to 2017 showed positive selection signals mainly in the genomic region involved in insecticide resistance. The X-axis shows the genome (chromosome) positions and the Y-axis shows the Garud' H12 values. The top and bottom left panels show the selection signal in *An. gambiae* s.s. populations, and the top and bottom right panels show the signal in *An. colozzii* populations. The colours represent the sampling periods, where the navy colour is 2012 and the brown colour is 2017. High values of H12 indicate a signal of positive selection in the corresponding genomic region within the populations: Ace1: Acetylcholinesterase gene; Cyp6p: Cytochrome P450 gene; Coeae60: Carboxyl-esterase gene; Keap1: Kelch-like ECH-Associated Protein 1; Dgk: Diaglycerol Kinase (AGAP000519) gene; Cyp9k1: Cytochrome P450 gene

the genome (Table S4). Most of these CNVs were found at frequencies below 50% except the deletions of *GSTD5* in 60% to 100% of all the vector populations (Fig. S4, Table S4).

The carboxylesterases (COEs) enzymes have been reported to be involved in the rapid metabolization of both pyrethroid and organophosphate compounds [51] and consequently could have an impact on the benefit of the combined organophosphate-based IRS and pyrethroid-based LLINs approach. In this study, CNVs were identified in 39 genes across the genome, most of these CNVs were gene amplifications, only 25.64% were deletions (Table S5). The frequencies of these CNVs were relatively low, except those of the *COEAE*60 gene which were identified at frequencies above 60% in *An. coluzzii* populations from 2012 to 2017 (Fig. S5). Additionally, H12 analysis revealed a positive selection signal in the *COEAE*60 gene (Fig. 8). The *COEAE*G* cluster genes (*COEAE2G, COEAE3G, COEAE5G, COEAE6G* and *COEAE7G*) and the *Coeae3H* gene showed CNVs at increasing frequencies from 11% in 2012 to up to 25% in 2017 in the *An. gambiae* s.s. populations. The *COE-JHE3E* gene showed CNVs at 32.56% frequencies in *An.*

arabiensis, and at frequencies higher than 30% in *An. gambiae* s.s. populations (except for sampling in 2017 where its frequency was 19.4%) (Fig. S5).

The emergence of new forms of insecticide resistance may impact vector control strategy in unexpected ways. Recent studies reported new forms of resistance, especially behavioural changes (early and/or outdoor biting) in *An. gambiae* mosquitoes, in response to the intense use of LLINs causing residual malaria transmission [52, 53]. Here, a signal of positive selection was detected at the DGK locus and an increased frequency of CNVs of this gene within the *An. gambiae* s.s. populations (Figs. 8 and 9). No selection signal and no CNVs were detected in vector populations in 2012. However, in 2017, a signal of positive selection was detected in the DGK locus, with the presence of CNVs at increasing frequency (from 0.00 in 2012 to 16.67% in 2017) in vector populations (Figs. 7, 8 and 9). The evolutionary pattern of the DGK gene and its reported role in the adaptation of *Caenorhabditis elegans* and *Drosophila* spp. to environmental changes [54, 55], suggest that this gene may play an important role in the modulation of *An. gambiae* biology in response to the presence of vector control tools.

Kientega *et al. Malaria Journal*   (2024) 23:280

Page 11 of 16
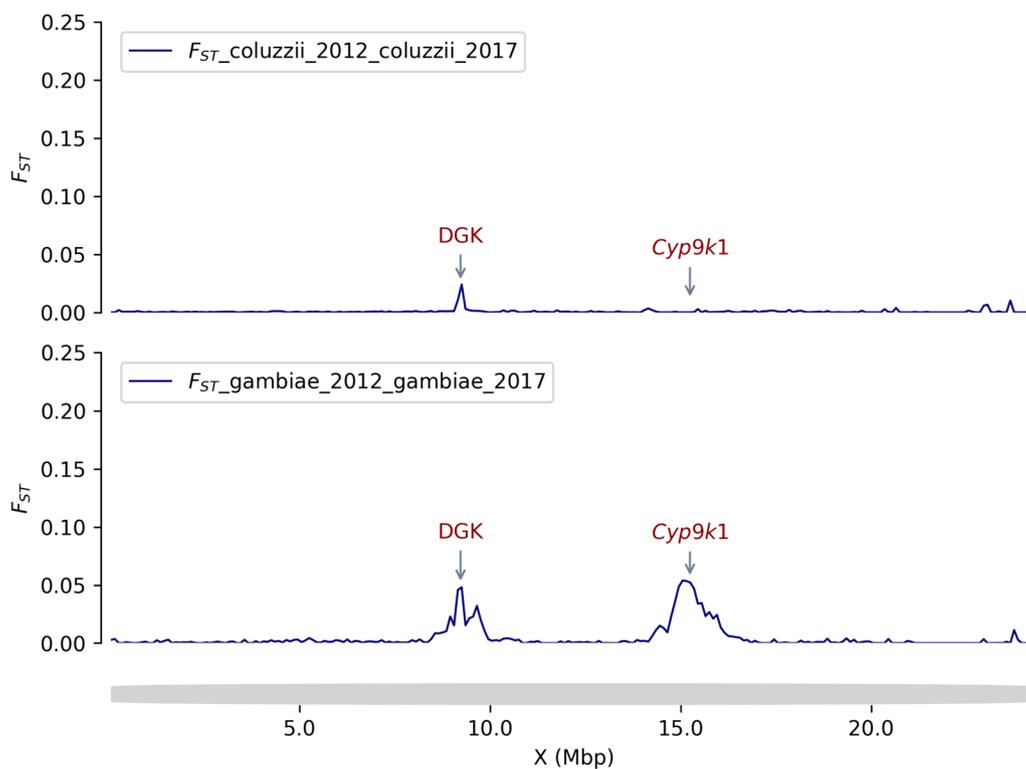
Fig. 9 Genome-wide selection scan using genetic differentiation (F_{ST}) across chromosome X (10 kb window) between populations of *An. gambiae* s.s. and *An. coluzzii* collected in 2012 and in 2017. The X-axis shows the positions of the chromosome X and the Y-axis shows the values of the 10 kb windowed F_{ST}. Signals of positive selection (high genetic differentiation, $F_{ST} > 0$) were observed in the genomic regions corresponding to *Cyp9k1* and *DGK* genes within the *An. gambiae* s.s populations

Discussion

Significant advances have been made in the reduction of malaria burden using LLINs and IRS, but the efficacy of these tools is at risk due to the high adaptability of the vector populations to the anthropogenic ecological changes [42]. A better understanding of the molecular, ecological and evolutionary processes allowing vector populations to adapt to the environmental changes would help in the development of innovative tools to support the current control strategies in malaria elimination. This study generated whole genome sequencing data to investigate the genetic diversity, the population structure and the evolution of insecticide resistance variants in the *An. gambiae* s.l. populations in the western part of Burkina Faso.

Population structure and vector control

Anopheles gambiae s.s. and *An. coluzzi*, originated from the previously described *An. gambiae* M and S forms have similar historical genomic and demographic backgrounds [36]. These results are consistent with this evidence and confirm the similar demographic history between these two species. The results also showed low F_{ST} between *An. coluzzii* and *An. gambiae* and are consistent with ongoing gene flow between *An. gambiae* and *An. coluzzii*. This supports previous studies in the same area showing strong evidence of gene flow between *An. gambiae* and *An. coluzzii* populations in west Africa [13]. In the context of genetic control, this lack of population structure between these species supports the ongoing hypothesis of the spread of gene drive constructs between *An. gambiae* s.s. and *An. coluzzii*, as previously observed with the insecticide resistance variants [38, 56]. The gene flow also contributes to the spread and the maintenance of genetic variants associated with insecticide resistance in vector populations [57].

Bed nets—evidence and impacts of target-site and metabolic pyrethroid resistance

The insecticides pressure on malaria vectors lead to an emergence of new insecticide resistance variants that confer fitness benefits to vector populations [58]. This seems to be consistent and evident for the SNPs *vgsc-V402L+I1527T* which were identified at increasing frequencies in *An. coluzzii* populations. The rapid increase

in their frequencies within *An. coluzzii* populations was expected as a recent study demonstrated the contribution of the *vgsc-V402L* to the pyrethroid resistance phenotype at a reduced fitness cost compared to *vgsc-L995F* [39]. The simultaneous emergence of different types of genetic variants in genes and gene clusters known to be involved in pyrethroid metabolism and target site mutations within the *An. gambiae s.l.* populations raises major concerns about the long-term efficacy of new pyrethroid insecticide-based control strategies such as LLINs.

Insecticide resistance has been widely studied in Africa, and recent works have reported multiple resistance variants spreading in vectors populations, jeopardizing NMCP efforts in control and elimination [18, 19, 59]. The implementation of LLINs incorporating piperonyl butoxide (PBO) aims to strengthen the efficacy of bed nets by inhibiting the cytochromes P450 enzymes responsible for the rapid detoxifying of pyrethroid insecticide [60]. High expression of the P450 enzymes was shown to increase the pyrethroid resistance and pre-exposure to PBO restores the susceptibility of malaria vector to pyrethroid [41]. In most cases, the main genetic variants responsible for the insecticide resistance in a given area remain unclear because of the high heterogeneity of the genetic markers mediating metabolic resistance. The results showed a high deletion rate of some CYP P450 genes in up to 80% of the vector populations. This situation could probably reduce the expression level of certain genes that are mostly targeted by PBO and increase the expression level of the least targeted one. However, the causes and the magnitude of the genes deletion and amplification in the mosquito genome remain unclear, but these variants do have a high importance in the rapid adaptation of vector populations to different ecological changes [16]. Moreover, the emergence of these additional new variants could enhance the level of pyrethroid resistance impacting the efficacy of the new vector control tools.

Indoor residual spraying (IRS)—evidence and impacts of organophosphates resistance

Another major concern related to this multi-modal resistance is its threat on the effectiveness of the IRS. IRS is efficient when it is used in combination with LLINs because of its ability to target indoor resting pyrethroid resistant vectors [61]. Although the SNP *ace1-G280S* has long been identified in *An. gambiae* populations, its frequencies remained low in the populations probably due to its fitness cost in association with the kdr mutation [62]. In fact, the duplication of the *ace1* gene (*ACE1^D*) seems to have completely solved the problem of genetic cost yielded by the SNP *ace1-G280S* [63]. Recent reports showed a wide distribution of this duplication in West African *An. coluzzii* and *An. gambiae* s.s. populations

[19, 63] probably in response to the use of organophosphate and carbamates compounds in vector control. Similarly, the duplication of *Ace1* gene follows the same trend as the SNP *ace1-G280S* raising up in frequencies in *An. gambiae* s.s. populations. This clearly showed the close related link between these two variants in conferring high fitness to the mosquito populations. In addition, the CNVs identified at high frequencies in some of the *GSTe* and COE genes could increase the expression level of these genes. The simultaneous presence of these variants (CNVs and SNPs) in *ACE1*, *GSTe* and esterases genes within the same vector populations strongly contribute to increase the survival capacity of the vector against organophosphates and carbamates. This situation threatens the long-term efficacy of vector control interventions including the recently introduced pirimiphos-methyl-based IRS [19].

Whole genome surveillance enables detection of potential novel insecticide resistance loci

Genome-wide scans for positive selection have proved their usefulness in the identification of genomic regions under divergent selection. In malaria vectors, under high insecticide mediated evolutionary pressure, these regions often have implications in insecticide resistance [64]. These analyses showed a signal of positive selection in a genomic region of the chromosome X corresponding to the DGK gene. Though there is currently no evidence of this gene's association with insecticide resistant phenotypes in malaria vectors, studies in *Drosophila melanogaster* and *Caenorhabditis elegans* showed its potential involvement in the adaptation to environmental changings [54, 55]. In *D. melanogaster,* the *rdgA* gene, an orthologue of the *dgk* gene, is eye-specific, strongly involved in light signaling. Any genetic variant influencing the functioning of this gene would probably affect its sensitivity to light, as reported previously [54, 65]. Consequently, if light signaling is regulated the same way in mosquito species, the *dgk* gene may also be involved in the regulation of *An. gambiae* vision. In *An. gambiae*, a previous study showed a rhythmic expression of the genes involved in light signaling under the control of the circadian clock [66]. The expression pattern of these genes suggests their role in calibration of the mosquito visual system to start the nocturnal activities such as flight and biting. In response to the high coverage of LLINs, major changes were reported in vector biting behaviour particularly early biting and later biting activities in many areas of Africa [53, 67]. Even though no evidence was highlighted, these changes could be related to alterations in the expression of genes involved in the mosquito visual system.

Another possible implication of the *dgk* gene in the adaptation of vector populations is its indirect implication to insecticide resistance via the regulation of the acetylcholine release at synaptic junctions. In *C. elegans*, the *dgk* gene regulates the amount of acetylcholine in synaptic junctions and its inactivation causes a high susceptibility to aldicarb, a carbamate insecticide [55]. This report clearly showed the potential implication of the *dgk* gene in organophosphates and carbamates resistance *C. elegans.* If synaptic junctions are regulated in the same way in mosquitoes, any selective variant in the DGK gene would affect the acetylcholine release in synapses. These two hypotheses of the potential implication of the *dgk* gene in adaptations, either insecticide resistance or biting time change, were previously drawn up to explain the selection signal observed in the *dgk* gene [64]. These hypotheses appear to be different in both organisms, but the molecular mechanisms underlying the DGK gene function remain similar. Given this intriguing role of *dgk*, further studies are needed to deeply investigate in depth the role of the *dgk* gene (AGAP000519) in modulating malaria vector behaviour and its potential involvement in behavioural and genetic resistance. Therefore, future research would focus on understanding the genetic variation associated with dry season adaptation and population maintenance, as well as the molecular mechanisms underlying changes in vector behaviour.

Conclusion

Investigation of genome variation in vector populations is crucial for understanding dynamics of populations in a given area and for guiding strategies of malaria control programs. This study showed a weak differentiation and strong genetic connection between *An. gambiae* and *An. coluzzii* populations in western Burkina Faso. The emergence of new variants together with the already existing pyrethroids resistance variants at high frequencies raise concerns about the long-term efficacy of new-generation nets impregnated with pyrethroids compounds. The long-term efficacy of pirimiphos-methyl-based IRS is also threatened because of the increasing frequency of the SNP *ace1-G280S* and the duplication of the *ace1* gene, which confer high fitness to vector populations carrying them. Genome-wide selection scans detected strong signals of positive selection in the genomic regions having an implication in insecticide resistance. This study highlighted the benefit of the molecular surveillance of malaria vectors for detection of new insecticide resistance variants, for monitoring the existing resistance variants and for getting insights in evolutionary processes underlying these variants.

Supplementary Information

Supplementary Material 1: Fig. S1. Site frequency spectra in the 3L chromosome of the *An. gambiae* s.l. populations from 2012 to 2016 in the three villages). The X axis of each figure shows the minor allele frequency and the Y axis, the density of the variants. The colour boxes are indicating the sampling periods.

Supplementary Material 2: Fig. S2. PCA showing the year-to-yeargenetic structure of *An. gambiae* s.l. populations using the whole SNPs identified in the 3L chromosome. B. Year-to-year genetic structure of *An. coluzzii* populations; C. Year-to-year genetic structure of *An. gambiae* s.s. populations; D. Year-to-year genetic structure of *An. arabiensis* populations; These figures demonstrated the lack of geographic substructure within each species of the *An. gambiae* s.l. populations collected in different villages over the years.

Supplementary Material 3: Fig. S3. Genome-wide selection scan using genetic differentiationacross the genomeof *An. gambiae* s.s. and *An. coluzzii* collected in 2012 and in 2017. Signals of positive selection were observed in the genomic regions shown to be involved in insecticides resistance: *ace1*: acetylcholinesterase gene, *cyp6p*: Cytochrome P450 gene, *keap1*: Kelch-like ECH-associated protein 1, *vgsc*: voltage-gated sodium channel, *rdl*: resistance to dieldrin gene, *Dgk*: Diaglycerol Kinasegene, *cyp9k1*: Cytochrome P450 gene, *Mbp*: megabases.

Supplementary Material 4: Fig. S4. Heat map showing the CNVs frequenciesof the glutathione-s-transferase genes in the *An. gambiae* s.l. populations. The X axis shows the *An. gambiae* s.l. populations and the sampling periods. The Y axis shows the positions of the glutathione-s-transferase genes ID and the CNV type. The gradient colour bar shows the distribution of the allelic frequencies.

Supplementary Material 5: Fig. S5. Heat map showing the CNVs frequenciesof the carboxylesterase genes in the *An. gambiae* s.l. populations. The X axis shows the *An. gambiae* s.l. populations and the sampling periods. The Y axis shows the positions of the carboxylesterase genes ID and the CNV type. The gradient colour bar shows the distribution of the allelic frequencies.

Supplementary Material 6: Table S1. Non-synonymous SNPs frequencies of the *Vgsc* genewithin *An. gambiae* s.l. populations.

Supplementary Material 7: Table S2. Copy number variations of the cytochrome p450 genes and their frequencies within *An. gambiae* s.l. populations.

Supplementary Material 8: Table S3. Non-synonymous SNPs frequencies of the *ace1* genewithin *An. gambiae* s.l. populations.

Supplementary Material 9: Table S4. Copy number variations of the glutathione-s-transferase genes and their frequencies within *An. gambiae* s.l. populations.

Supplementary Material 10: Table S5. Copy number variations of the carboxylesterase genes and their frequencies within *An. gambiae* s.l. populations.

Author contributions

Funding

Availability of data and materials
Jupyter Notebooks and scripts to reproduce all the analyses, tables and figures are available in the GitHub repository: https://github.com/mkient/BF_Ag100 0G_analyses. The SNPs and haplotypes data are available on the homepage of MalariaGEN and can be accessed using the malariagen_data package. The raw sequences in FASTQ format and the aligned sequences in BAM format were stored in the European Nucleotide Archive (ENA, Study Accession n° PRJEB42254).

Declarations

Ethics approval and consent to participate
All methods in this paper have been implemented in accordance with the relevant guidelines/regulations/legislation in Burkina Faso. No ethics approval was required to run all the activities related to this paper.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Institut de Recherche en Sciences de la Santé (IRSS), 01 BP 545 Bobo-Dioulasso 01, Burkina Faso. [2]Vector Surveillance Programme, Genomic Surveillance Unit, Wellcome Sanger Institute, Hinxton, Cambridge, UK. [3]Université Nazi Boni, 01 BP 1091 Bobo-Dioulasso, Burkina Faso. [4]Department of Life Sciences, Imperial College London, Silwood Park, Ascot SL5 7PY, UK. [5]Institut des Sciences des Sociétés, 03 BP 7047 Ouagadougou 03, Burkina Faso.

References
1. WHO. World malaria report 2023. Geneva: World Health Organization; 2023. https://iris.who.int/bitstream/handle/10665/374472/9789240086 173-eng.pdf?sequence=1. Accessed 5 Jan 2024
2. Hemming-Schroeder E, Zhong D, Machani M, Nguyen H, Thong S, Kahindi S, et al. Ecological drivers of genetic connectivity for African malaria vectors *Anopheles gambiae* and *An. arabiensis*. Sci Rep. 2020;10:19946.
3. 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.
4. Fournet F, Cussac M, Ouari A, Meyer PE, Toé HK, Gouagna LC, et al. Diversity in anopheline larval habitats and adult composition during the dry and wet seasons in Ouagadougou (Burkina Faso). Malar J. 2010;9:78.
5. Epopa PS, Millogo AA, Collins CM, North A, Tripet F, Benedict MQ, et al. The use of sequential mark-release-recapture experiments to estimate population size, survival and dispersal of male mosquitoes of the *Anopheles gambiae* complex in Bana, a west African humid savannah village. Parasit Vectors. 2017;10:376.
6. Soma DD, Zogo BM, Somé A, Tchiekoi BN, de Hien DFS, Pooda HS, et al. *Anopheles* bionomics, insecticide resistance and malaria transmission in southwest Burkina Faso: a pre-intervention study. PLoS ONE. 2020;15: e0236920.
7. INSD, ICF. Enquête Démographique et de Santé 2021: Rapport de synthèse. Rockville, Maryland; 2023. https://dhsprogram.com/pubs/pdf/SR279/SR279.pdf. Accessed 4 Nov 2023
8. Ministère de la Santé. Rapport de revue des performances du programme national de lutte 2011–2015—Burkina Faso. Ouagadougou; 2016.
9. U.S. President's Malaria Initiative. Burkina Faso Malaria Operational Plan FY 2023. 2023. https://www.pmi.gov/resources/malaria-operational-plans-mops/. Accessed 8 Nov 2023
10. Ministère de la Santé. Annuaire statistique 2021. Ouagadougou; 2022. http://cns.bf/IMG/pdf/annuaire_mshp_2021_signe.pdf. Accessed 20 Jan 2023.
11. Coulibaly ZI, Gowelo S, Traore I, Mbewe RB, Ngulube W, Olanga EA, et al. Strengthening adult mosquito surveillance in Africa for disease control: learning from the present. Curr Opin Insect Sci. 2023;60:101110.
12. Alonso J, Orue-Echevarria L, Casola V, Torre AI, Huarte M, Osaba E, et al. Understanding the challenges and novel architectural models of multi-cloud native applications—a systematic literature review. J Cloud Comput. 2023;12:6.
13. The *Anopheles gambiae* 1000 Genomes Consortium. Genetic diversity of the African malaria vector *Anopheles gambiae*. Nature. 2017;552:96–100.
14. The *Anopheles gambiae* 1000 Genomes Consortium. 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.
15. MalariaGEN. The *Anopheles gambiae* 1000 genomes project (Ag1000G). MalariaGEN, 2014. https://www.malariagen.net/mosquito/ag1000g. Accessed 8 Apr 2023.
16. Weetman D, Djogbenou LS, Lucas E. Copy number variation (CNV) and insecticide resistance in mosquitoes: evolving knowledge or an evolving problem? Curr Opin Insect Sci. 2018;27:82–8.
17. Lucas ER, Miles A, Harding NJ, Clarkson CS, Lawniczak MKN, Kwiatkowski DP, et al. Whole-genome sequencing reveals high complexity of copy number variation at insecticide resistance loci in malaria mosquitoes. Genome Res. 2019;29:1250–61.

18. 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.

19. Grau-Bové X, Lucas E, Pipini D, Rippon E, van'tHof 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.

20. Moyes CL, Athinya DK, Seethaler T, Battle KE, Sinka M, Hadi MP, et al. Evaluating insecticide resistance across African districts to aid malaria control decisions. Proc Natl Acad Sci USA. 2020;117:22042–50.

21. Amoudji AD, Ahadji-Dabla KM, Hien AS, Apétogbo YG, Yaméogo B, Soma DD, et al. Insecticide resistance profiles of *Anopheles gambiae* s.l. in Togo and genetic mechanisms involved, during 3-year survey: is there any need for resistance management? Malar J. 2019;18:177.

22. Sovi A, Keita C, Sinaba Y, Dicko A, Traore I, Cisse MBM, et al. *Anopheles gambiae* (s.l.) exhibit high intensity pyrethroid resistance throughout Southern and Central Mali (2016–2018): PBO or next generation LLINs may provide greater control. Parasit Vectors. 2020;13:239.

23. Wilson AL, Courtenay O, Kelly-Hope LA, Scott TW, Takken W, Torr SJ, et al. The importance of vector control for the control and elimination of vector-borne diseases. PLoS Negl Trop Dis. 2020;14: e0007831.

24. Sougoufara S, Ottih EC, Tripet F. The need for new vector control approaches targeting outdoor biting anopheline malaria vector communities. Parasit Vectors. 2020;13:295.

25. Crawford JE, Riehle MM, Markianos K, Bischoff E, Guelbeogo WM, Gneme A, et al. Evolution of GOUNDRY, a cryptic subgroup of *Anopheles gambiae* s.l., and its impact on susceptibility to Plasmodium infection. Mol Ecol. 2016;25:1494–510.

26. 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.

27. Epopa PS, Collins CM, North A, Millogo AA, Benedict MQ, Tripet F, et al. Seasonal malaria vector and transmission dynamics in western Burkina Faso. Malar J. 2019;18:113.

28. Naré RWA, Savadogo PW, Gnankambary Z, Nacro HB, Sedogo MP. Analyzing risks related to the use of pesticides in vegetable gardens in Burkina Faso. Agricult Forest Fish. 2015;4:165.

29. Epopa PS, Millogo AA, Collins CM, North AR, Benedict MQ, Tripet F, et al. *Anopheles gambiae* (s.l.) is found where few are looking: assessing mosquito diversity and density outside inhabited areas using diverse sampling methods. Parasit Vectors. 2020;13:516.

30. Gillies MT, Coetzee M. A supplement to the Anophelinae of Africa South of the Sahara (Afrotropical region). South African Inst Med Res Johannesburg; 1987.

31. Malariagen. Pipelines for processing malaria parasite and mosquito genome sequence data. MalariaGEN pipelines—releases 9 v0.0.9. 2022. https://github.com/malariagen/pipelines. Accessed 7 Nov 2023.

32. Miles A. GitHub-malariagen/malariagen-data-python: a Python package providing functions for accessing and analysing MalariaGEN data. malariagen/malariagen-data-python. 2022. https://github.com/malariagen/malariagen-data-python. Accessed 6 May 2022.

33. Miles A, Pyup/io Bot, Rodrigues MF, Ralph P, Harding N, Pisupati R, et al. Scikit-allel—Explore and analyse genetic variation. Zenodo. 2021. https://zenodo.org/record/4759368. Accessed 6 May 2022.

34. Garud NR, Messer PW, Buzbas EO, Petrov DA. Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. PLoS Genet. 2015;11: e1005004.

35. Petrarca V, Nugud AD, Ahmed MAE, Haridi AM, Di Deco MA, Coluzzi M. Cytogenetics of the *Anopheles gambiae* complex in Sudan, with special reference to *An. arabiensis*: relationships with East and West African populations. Med Vet Entomol. 2000;14:149–64. https://doi.org/10.1046/j.1365-2915.2000.00231.x.

36. Crawford JE, Lazzaro BP. The demographic histories of the M and S molecular forms of *Anopheles gambiae* s.s. Mol Biol Evol. 2010;27:1739–44. https://doi.org/10.1093/molbev/msq070.

37. 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.

38. Hemingway J, Ranson H. Insecticide resistance in insect vectors of human disease. Annu Rev Entomol. 2000;45:371–91.

39. Williams J, Cowlishaw 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.

40. Jones CM, Liyanapathirana M, Agossa FR, Weetman D, Ranson H, Donnelly MJ, et al. Footprints of positive selection associated with a mutation (N1575Y) in the voltage-gated sodium channel of *Anopheles gambiae*. Proc Natl Acad Sci USA. 2012;109:6614–9.

41. 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.

42. Ranson H. Current and future prospects for preventing malaria transmission via the use of insecticides. Cold Spring Harb Perspect Med. 2017;7: a026823.

43. Sherrard-Smith E, Griffin JT, Winskill P, Corbel V, Pennetier C, Djénontin A, et al. Systematic review of indoor residual spray efficacy and effectiveness against *Plasmodium falciparum* in Africa. Nat Commun. 2018;9:4982.

44. Loha E, Deressa W, Gari T, Balkew M, Kenea O, Solomon T, et al. Long-lasting insecticidal nets and indoor residual spraying may not be sufficient to eliminate malaria in a low malaria incidence area: results from a cluster randomized controlled trial in Ethiopia. Malar J. 2019;18:141.

45. Kenea O, Balkew M, Tekie H, Deressa W, Loha E, Lindtjørn B, et al. Impact of combining indoor residual spraying and long-lasting insecticidal nets on *Anopheles arabiensis* in Ethiopia: results from a cluster randomized controlled trial. Malar J. 2019;18:182.

46. Soma DD, Poda SB, Hien AS, Namountougou M, Sangaré I, Sawadogo JME, et al. Malaria vectors diversity, insecticide resistance and transmission during the rainy season in peri-urban villages of south-western Burkina Faso. Malar J. 2021;20:63.

47. Rowland M, Boko P, Odjo A, Asidi A, Akogbeto M, N'Guessan R. A new long-lasting indoor residual formulation of the organophosphate insecticide pirimiphos methyl for prolonged control of pyrethroid-resistant mosquitoes: an experimental hut trial in Benin. PLoS ONE. 2013;8: e69516. https://doi.org/10.1371/journal.pone.0069516.

48. Hamainza B, Sikaala CH, Moonga HB, Chanda J, Chinula D, Mwenda M, et al. Incremental impact upon malaria transmission of supplementing pyrethroid-impregnated long-lasting insecticidal nets with indoor residual spraying using pyrethroids or the organophosphate, pirimiphos methyl. Malar J. 2016;15:100.

49. Weill M, Malcolm C, Chandre F, Mogensen K, Berthomieu A, Marquine M, et al. The unique mutation in ace-1 giving high insecticide resistance is easily detectable in mosquito vectors. Insect Mol Biol. 2004;13:1–7.

50. Riveron JM, Yunta C, Ibrahim SS, Djouaka R, Irving H, Menze BD, et al. A single mutation in the GSTe2 gene allows tracking of metabolically based insecticide resistance in a major malaria vector. Genome Biol. 2014;15:R27.

51. Cui F, Li M-X, Chang H-J, Mao Y, Zhang H-Y, Lu L-X, et al. Carboxylesterase-mediated insecticide resistance: quantitative increase induces broader metabolic resistance than qualitative change. Pestic Biochem Physiol. 2015;121:88–96.

52. Moiroux N, Damien GB, Egrot M, Djenontin A, Chandre F, Corbel V, et al. Human exposure to early morning *Anopheles funestus* biting behavior and personal protection provided by long-lasting insecticidal nets. PLoS ONE. 2014;9: e104967.

53. Finda MF, Moshi IR, Monroe A, Limwagu AJ, Nyoni AP, Swai JK, et al. Linking human behaviours and malaria vector biting risk in south-eastern Tanzania. PLoS ONE. 2019;14: e0217414.

54. Masai I, Okazaki A, Hosoya T, Hotta Y. Drosophila retinal degeneration A gene encodes an eye-specific diacylglycerol kinase with cysteine-rich zinc-finger motifs and ankyrin repeats. Proc Natl Acad Sci USA. 1993;90:11157–61.

55. Miller KG, Emerson MD, Rand JB. Goα and diacylglycerol kinase negatively regulate the Gqα pathway in *C. elegans*. Neuron. 1999;24:323–33.

56. Djogbénou L, Chandre F, Berthomieu A, Dabiré R, Koffi A, Alout H, et al. Evidence of Introgression of the ace-1R mutation and of the ace-1 duplication in West African *Anopheles gambiae* s.s. PLoS ONE. 2008;3: e2172.

57. Maliti D, Ranson H, Magesa S, Kisinza W, Mcha J, Haji K, et al. Islands and stepping-stones: comparative population structure of *Anopheles gambiae sensu stricto* and *Anopheles arabiensis* in Tanzania and implications for the spread of insecticide resistance. PLoS ONE. 2014;9: e110910.

Kientega *et al. Malaria Journal*   (2024) 23:280

Page 16 of 16

58. Benelli G, Beier JC. Current vector control challenges in the fight against malaria. Acta Trop. 2017;174:91–6.

59. Sougoufara S, Doucouré S, Sembéne PMB, Harry M, Sokhna C. Challenges for malaria vector control in sub-Saharan Africa: resistance and behavioral adaptations in anopheles populations. J Vector Borne Dis. 2017;54:4–15.

60. Feyereisen R. Insect P450 inhibitors and insecticides: challenges and opportunities. Pest Manag Sci. 2015;71:793–800.

61. Okumu FO, Moore SJ. Combining indoor residual spraying and insecticide-treated nets for malaria control in Africa: a review of possible outcomes and an outline of suggestions for the future. Malar J. 2011;10:208.

62. Djogbénou L, Noel V, Agnew P. Costs of insensitive acetylcholinesterase insecticide resistance for the malaria vector *Anopheles gambiae* homozygous for the G119S mutation. Malar J. 2010;9:12.

63. Assogba BS, Djogbénou LS, Milesi P, Berthomieu A, Perez J, Ayala D, et al. An ace-1 gene duplication resorbs the fitness cost associated with resistance in *Anopheles gambiae*, the main malaria mosquito. Sci Rep. 2015;5:14529.

64. Miles A. Genomic epidemiology of malaria vectors in the *Anopheles gambiae* species complex. PhD Thesis, University of Oxford; 2021. https://ora.ox.ac.uk/objects/uuid:d0fc0f47-e24f-4804-9022-26c4e3cf1428/download_file?safe_filename=thesis.pdf&type_of_work=Thesis. Accessed 14 Apr 2023.

65. Delgado R, Muñoz Y, Peña-Cortés H, Giavalisco P, Bacigalupo J. Diacylglycerol activates the light-dependent channel TRP in the photosensitive microvilli of *Drosophila melanogaster* photoreceptors. J Neurosci. 2014;34:6679–86.

66. Rund SSC, Hou TY, Ward SM, Collins FH, Duffield GE. Genome-wide profiling of diel and circadian gene expression in the malaria vector *Anopheles gambiae*. Proc Natl Acad Sci USA. 2011;108:E421–30.

67. Sherrard-Smith E, Skarp JE, Beale AD, Fornadel C, Norris LC, Moore SJ, et al. Mosquito feeding behavior and how it influences residual malaria transmission across Africa. Proc Natl Acad Sci USA. 2019;116:15086–95.

Publisher's Note