

Article

Reduced Genetic Diversity of Key Fertility and Vector Competency Related Genes in *Anopheles gambiae* s.l. Across Sub-Saharan Africa

Fatoumata Seck ^{1,2}, Mouhamadou Fadel Diop ¹, Karim Mané ¹, Amadou Diallo ² , Idrissa Dieng ² ,
Moussa Namountougou ³, Abdoulaye Diabate ³, Alfred Amambua-Ngwa ¹, Ibrahima Dia ² ,
and Benoit Sessinou Assogba ^{1,*}

¹ Medical Research Council Unit, The Gambia at the London School of Hygiene and Tropical Medicine, Banjul P.O. Box 273, The Gambia; seckafatoumata@gmail.com (F.S.); mouhamadou-fadel.diop@lshtm.ac.uk (M.F.D.); karim.mane@lshtm.ac.uk (K.M.); alfred.ngwa@lshtm.ac.uk (A.A.-N.)

² Institut Pasteur de Dakar, Dakar 220, Senegal; idrissa.dieng@pasteur.sn (I.D.); ibrahima.dia@pasteur.sn (I.D.)

³ Institut de Recherche en Sciences de la Santé (IRSS), Bobo-Dioulasso 01 BP 545, Burkina Faso; namountougou_d@yahoo.fr (M.N.); npiediab@gmail.com (A.D.)

* Correspondence: sessinou-benoit.assogba@lshtm.ac.uk

Abstract: Background: Insecticide resistance challenges the vector control efforts towards malaria elimination and proving the development of complementary tools. Targeting the genes that are involved in mosquito fertility and susceptibility to *Plasmodium* with small molecule inhibitors has been a promising alternative to curb the vector population and drive the transmission down. However, such an approach would require a comprehensive knowledge of the genetic diversity of the targeted genes to ensure the broad efficacy of new tools across the natural vector populations. Methods: Four fertility and parasite susceptibility genes were identified from a systematic review of the literature. The Single Nucleotide Polymorphisms (SNPs) found within the regions spanned by these four genes, genotyped across 2784 wild-caught *Anopheles gambiae* s.l. from 19 sub-Saharan African (SSA) countries, were extracted from the whole genome SNP data of the Ag1000G project (Ag3.0). The population genetic analysis on gene-specific data included the determination of the population structure, estimation of the differentiation level between the populations, evaluation of the linkage between the non-synonymous SNPs (nsSNPs), and a few statistical tests. Results: As potential targets for small molecule inhibitors to reduce malaria transmission, our set of four genes associated with *Anopheles* fertility and their susceptibility to *Plasmodium* comprises the mating-induced stimulator of oogenesis protein (*MISO*, AGAP002620), *Vitellogenin* (*Vg*, AGAP004203), *Lipophorin* (*Lp*, AGAP001826), and *Haem-peroxidase 15* (*HPX15*, AGAP013327). The analyses performed on these potential targets of small inhibitor molecules revealed that the genes are conserved within SSA populations of *An. gambiae* s.l. The overall low *Fst* values and low clustering of principal component analysis between species indicated low genetic differentiation at all the genes (*MISO*, *Vg*, *Lp* and *HPX15*). The low nucleotide diversity (>0.10), negative Tajima's *D* values, and heterozygosity analysis provided ecological insights into the purifying selection that acts to remove deleterious mutations, maintaining genetic diversity at low levels within the populations. None of *MISO* nsSNPs were identified in linkage disequilibrium, whereas a few weakly linked nsSNPs with ambiguous haplotyping were detected at other genes. Conclusions: This integrated finding on the genetic features of major malaria vectors' biological factors across natural populations offer new insights for developing sustainable malaria control tools. These loci were reasonably conserved, allowing for the design of



Academic Editor: Domenico Lio

Received: 20 February 2025

Revised: 8 April 2025

Accepted: 9 April 2025

Published: 30 April 2025

Citation: Seck, F.; Diop, M.F.; Mané, K.; Diallo, A.; Dieng, I.; Namountougou, M.; Diabate, A.; Amambua-Ngwa, A.; Dia, I.; Assogba, B.S. Reduced Genetic Diversity of Key Fertility and Vector Competency Related Genes in *Anopheles gambiae* s.l. Across Sub-Saharan Africa. *Genes* **2025**, *16*, 543. <https://doi.org/10.3390/genes16050543>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

effective targeting with small molecule inhibitors towards controlling vector populations and lowering global malaria transmission.

Keywords: genes; genetic diversity; fertility; vector competency; *Anopheles gambiae* s.l.; sub-Saharan Africa

1. Introduction

Caused by *Plasmodium* parasites and transmitted by female *Anopheles* mosquitoes, malaria remains a significant public health concern. According to the latest estimations from the World Health Organization (WHO), malaria caused 597,000 deaths worldwide in 2023, with 97% of those cases occurring in Africa, primarily in sub-Saharan Africa [1].

Between 2000 and 2015, malaria incidence significantly decreased, with an estimated 663 million clinical cases averted. Of these cases, 68% were attributed to the use of insecticide-treated bed nets (ITNs), and 10% to indoor residual spraying (IRS) [2]. This demonstrates the substantial impact of vector control in the programs established to reduce the malaria burden. However, since 2016, malaria cases have been on the rise [1]. The extensive use of massive insecticide-based control interventions has led to the widespread development of insecticide resistance [3–7] and adaptive behavioral changes in vector populations [8–10]. This situation poses challenges to malaria elimination and highlights the need for the development of novel vector control tools.

The Sterile Insect Technique (SIT) holds promise for managing various insect pests, including *Anopheles* mosquitoes. However, this technique has limitations, particularly the requirement for efficient sex-separation systems to prevent the release of female mosquitoes and reduce the risk of disease transmission [11]. Additionally, the dispersal and fitness (survival and competitiveness) of released sterile insects are lower compared to wild populations [12]. Alternatively, direct sterilization of female *Anopheles* by disrupting the reproductive synergy with specific gene-targeted small molecule inhibitors could be an innovative approach to control malaria transmission. Molecular pathways underlying *Anopheles* reproductive biology have been extensively studied over the last decade, and the role of certain proteins involved has been established [13–24].

For this study, we conducted a brief systematic review on known genes involved either in *Anopheles* fertility or infectivity to *Plasmodium* to identify those associated with both biological factors. Studies have shown that investigating mosquito traits of fertility and immunity is necessary for the detection of potential targets for novel vector control strategies [25]. Four thoroughly characterized genes were selected as prospective targets for small molecule inhibitors: the mating-induced stimulator of oogenesis protein (*MISO*), the nutrients transporters, *Vitellogenin* (*Vg*) and *Lipophorin* (*Lp*), and the *Heme-peroxidase 15* (*HPX15*) (see File S1).

Nevertheless, such an approach requires prior knowledge of the genetic structure of the target genes in natural populations for evidence-based control. The genetic variation at the nucleotide level can lead to changes in phenotype as non-synonymous mutations can influence the protein's structure, function, or binding properties [26] and are often subject to natural selection.

Advancements in genome sequencing and the accessibility of next-generation sequencing (NGS) technologies are key to malaria surveillance and intervention programs. They provide updated insights into malaria genomic epidemiology, including demographic and evolutionary trends in both mosquitoes [27] and parasite [28] populations, some imposed by interventions. These technologies also offer opportunities to investigate and predict genetic variations underlying important traits, such as insecticide resistance [29–31] and drug

resistance [32], but have not been adequately explored to assess variation in susceptibility to malaria parasite infections and fertility, two important biological factors that may impact variance in malaria across populations.

Here, we analyzed the genetic variation in selected target genes involved in the fertility and vector competency of the species in the *An. gambiae* complex populations across sub-Saharan African populations. Genetic variants were obtained from whole genomic Single Nucleotide Polymorphisms (SNPs) data of the phase 3 (Ag3.0) of the *An. gambiae* 1000 Genomes project (Ag1000G). The goal was to identify the most conserved genes as suitable targets for vector sterilization and a potential reduction of infectivity to *Plasmodium*.

2. Results

2.1. Potential Targetable Genes Involved in *Anopheles gambiae* s.l. Fertility and *Plasmodium* Infectivity

Many genes express differently for controlling post-mating pathways (e.g., egg production, sperm storage, refracton to subsequent mating) and post-blood-feeding ones (e.g., egg maturation and immune response to pathogens as *Plasmodium*) of mated *Anopheles gambiae* females. The brief systematic review identified 27 genes that are involved in *An. gambiae* s.l. mosquito fertility, and 25 genes in their infectivity to *Plasmodium* (Figure 1). Amongst these, seven genes are associated with both *Anopheles* fertility and infectivity to *Plasmodium* (genes in the overlapping region between the two circles). Amongst these, the *mating-induced stimulator of oogenesis protein* (MISO, AGAP002620), the *Heme-peroxidase 15* (HPX15, AGAP013327), the *major yolk protein precursor Vitellogenin* (Vg, AGAP004203), and the *major vitellogenic lipid transporter Lipophorin* (Lp, AGAP001826) were selected as key targetable genes for vector control (see Supplementary Materials Introduction).

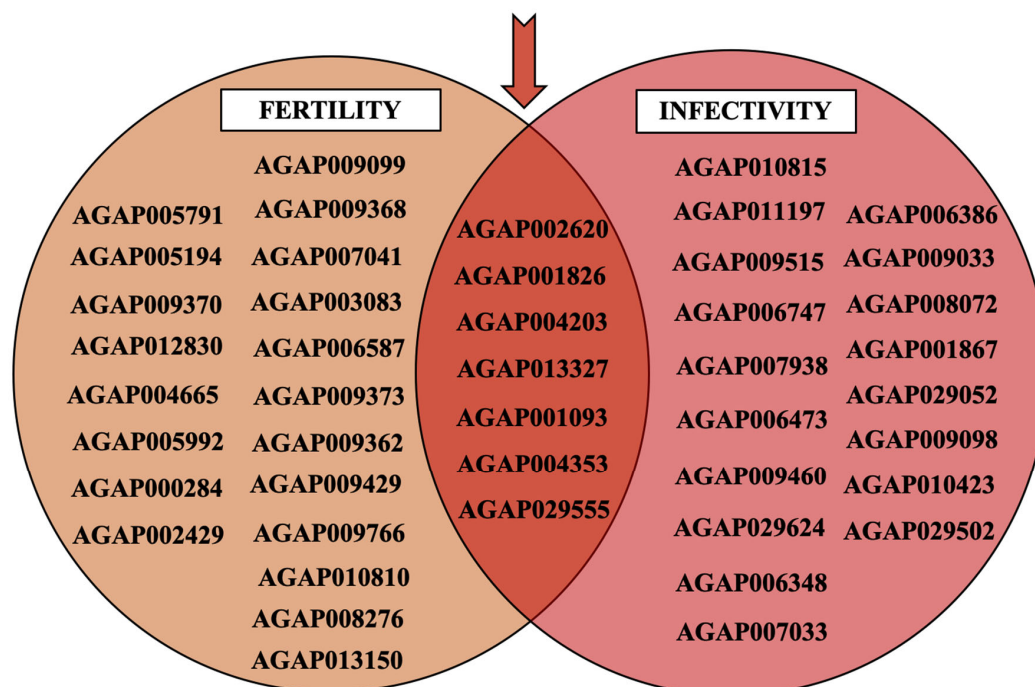


Figure 1. Identified genes involved in the fertility and/or infectivity of *Anopheles gambiae* s.l. This Venn diagram shows the list of genes identified from our literature review. Among the 27 genes identified as involved in *An. gambiae* s.l. mosquito fertility and the 25 genes identified as involved in their infectivity to *Plasmodium*, seven genes have been identified as having a possible implication in both fertility and infectivity (genes in the overlapping region between the two circles). The gene names are referred to the *An. gambiae* s.s. *PEST* gene annotation from Vectorbase.

2.2. Low Population Differentiation in the Four Targetable Genes Involved in *Anopheles gambiae* s.l. Fertility and *Plasmodium* Infectivity

Using SNP data from 2784 wild-caught *Anopheles gambiae* s.l. species collected across 19 sub-Saharan African countries (Figures 2 and S2), the principal component analysis (PCA) identified subpopulations between and within species for *MISO*, *Vg*, and *HPX15*, identifying hybrid cluster between species (Figure 3). The PCA analysis indicated less separation amongst species based on the components analyzed in this dataset at each of the genes (Figure 3A–D). The *LP* gene showed the most significant separation between species cluster (Figure 3C). In contrast, the *MISO* gene showed weaker genetic differentiation with more homogeneous distribution across species (Figure 3A). However, there was an overall low pairwise fixation index (F_{st} , Fixation Index of genetic differentiation between populations) between species low (0.005–0.051, 0.001–0.146, 0.002–0.022, and 0.006–0.048, for *MISO*, *Vg*, *Lp* and *HPX15*, respectively), with the highest divergence being against *An. coluzzii* populations for *MISO*, *Vg*, *Lp* genes and against *An. arabiensis* for *HPX15* gene (Figure 4). For all genes, the global F_{st} were very low (less than 0.03) between *An. gambiae* s.l. species. Interestingly, the F_{st} analysis between *An. gambiae* s.l. species (Figure 4) between geographic populations within species reinforced the observed low differentiation (Figure 5). The F_{st} values were mostly ranged within 0.10–0.15 between populations for *An. arabiensis*, *An. coluzzii* and *An. gambiae* at all four genes (Figure 5), except for *HPX15* in *An. arabiensis* from Cameroon (Figure 5D). Meanwhile, a moderate genetic differentiation between populations was observed within the hybrid *Anopheles coluzzii-gambiae* in all four genes, especially against populations from Kenya and Tanzania against those from Guinea Bissau, Guinea and Gambia in West Africa, reaching F_{st} values of up to 0.6 for *MISO* (Figure 5A).

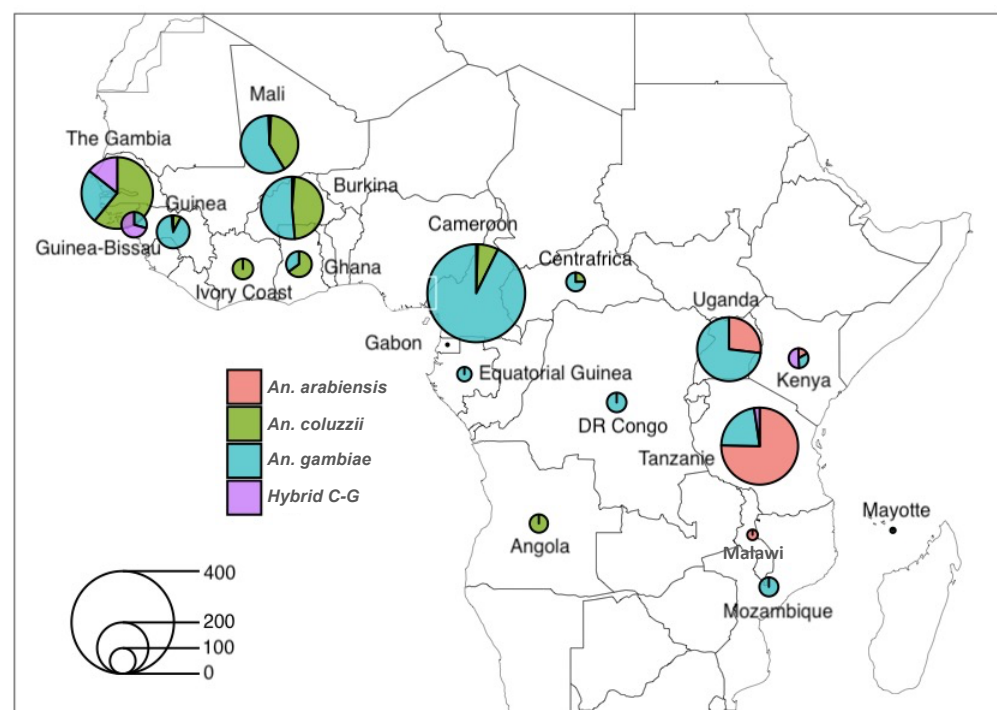


Figure 2. Map of distribution and composition of *Anopheles gambiae* s.l. database used from 19 sub-Saharan African countries. A total of 2784 whole genomes vcf SNP datasets of wild-caught *Anopheles gambiae* s.l. mosquitoes, collected in 19 sub-Saharan African countries, were utilised in this study. These datasets were obtained from the phase 3 Ag1000G project (https://www.malariagen.net/data_package/ag1000g-phase3-snp/, accessed on 6 August 2021). The numbers within the bar plots indicate the sample count for each population or species from a specific country.

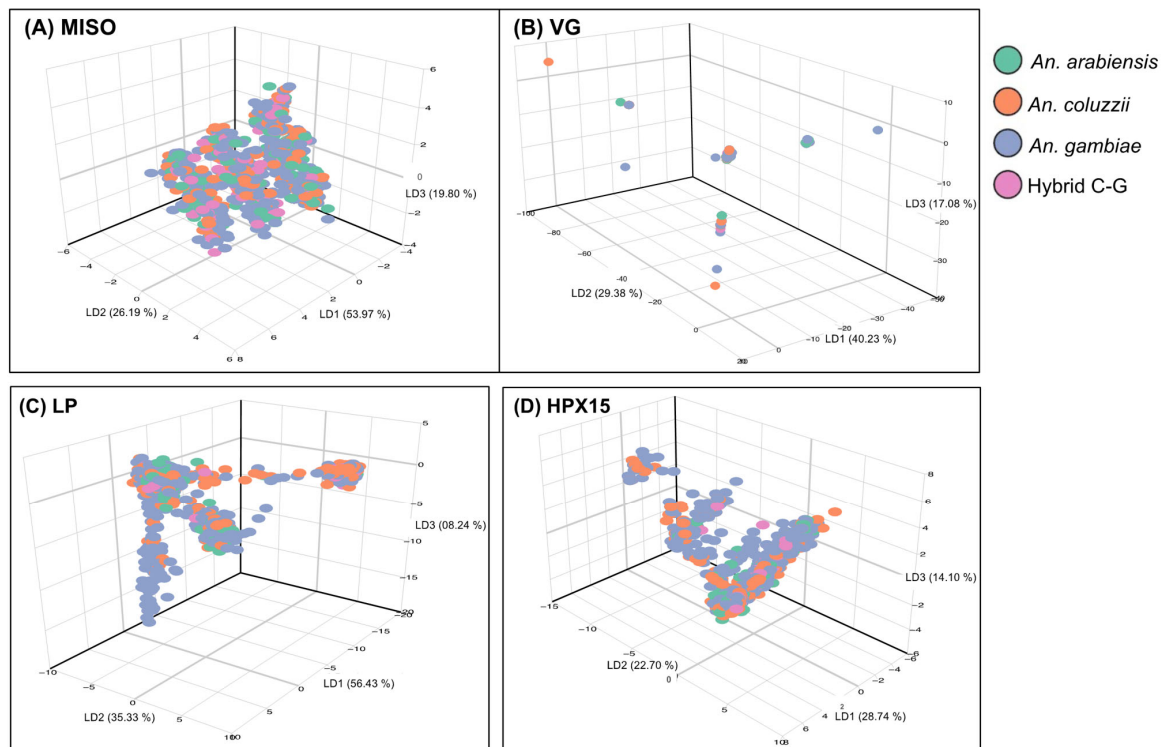


Figure 3. Principal component analysis (PCA) performed on four targetable genes involved in *Anopheles gambiae* s.l. fertility and *Plasmodium* infectivity using sequencing data from Malariagen phase 3 Ag1000G project (see Figure 2). (A) Mating-induced stimulator of oogenesis protein (MISO, AGAP002620), (B) Vitellogenin (VG, AGAP004203), (C) Lipophorin (LP, AGAP001826), and (D) Haem-peroxidase 15 (HPX15, AGAP013327). Within the *Anopheles gambiae* s.l. species, the PCA showed a lower separation among species based on the components analyzed in this dataset at each gene.

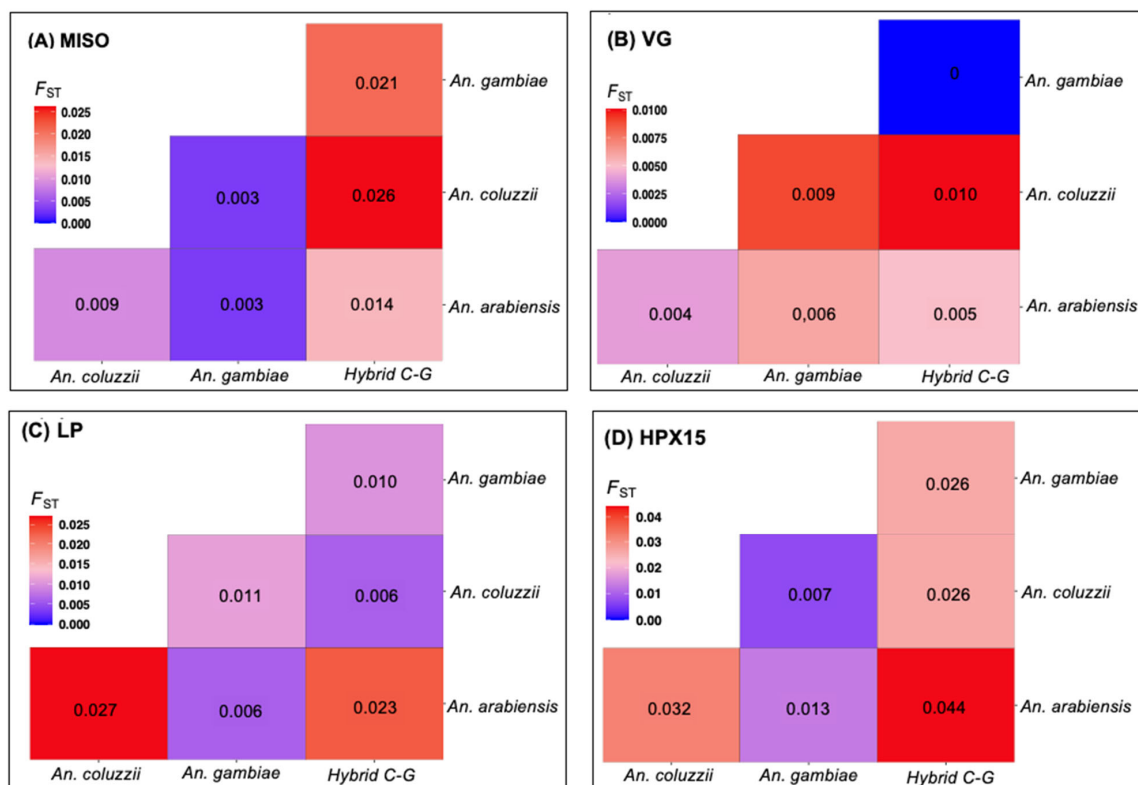


Figure 4. Global pairwise fixation index (F_{ST}) performed on four targetable genes involved in *Anopheles gambiae* s.l. fertility and *Plasmodium* infectivity using sequencing data from Malariagen

phase 3 Ag1000G project (see Figure 2). (A) Mating-induced stimulator of oogenesis protein (*MISO*, AGAP002620), (B) *Vitellogenin* (*VG*, AGAP004203), (C) *Lipophorin* (*LP*, AGAP001826), and (D) *Haem-peroxidase 15* (*HPX15*, AGAP013327). For each gene, the overall F_{ST} between species was generally low, ranging from dark-blue to dark-red.

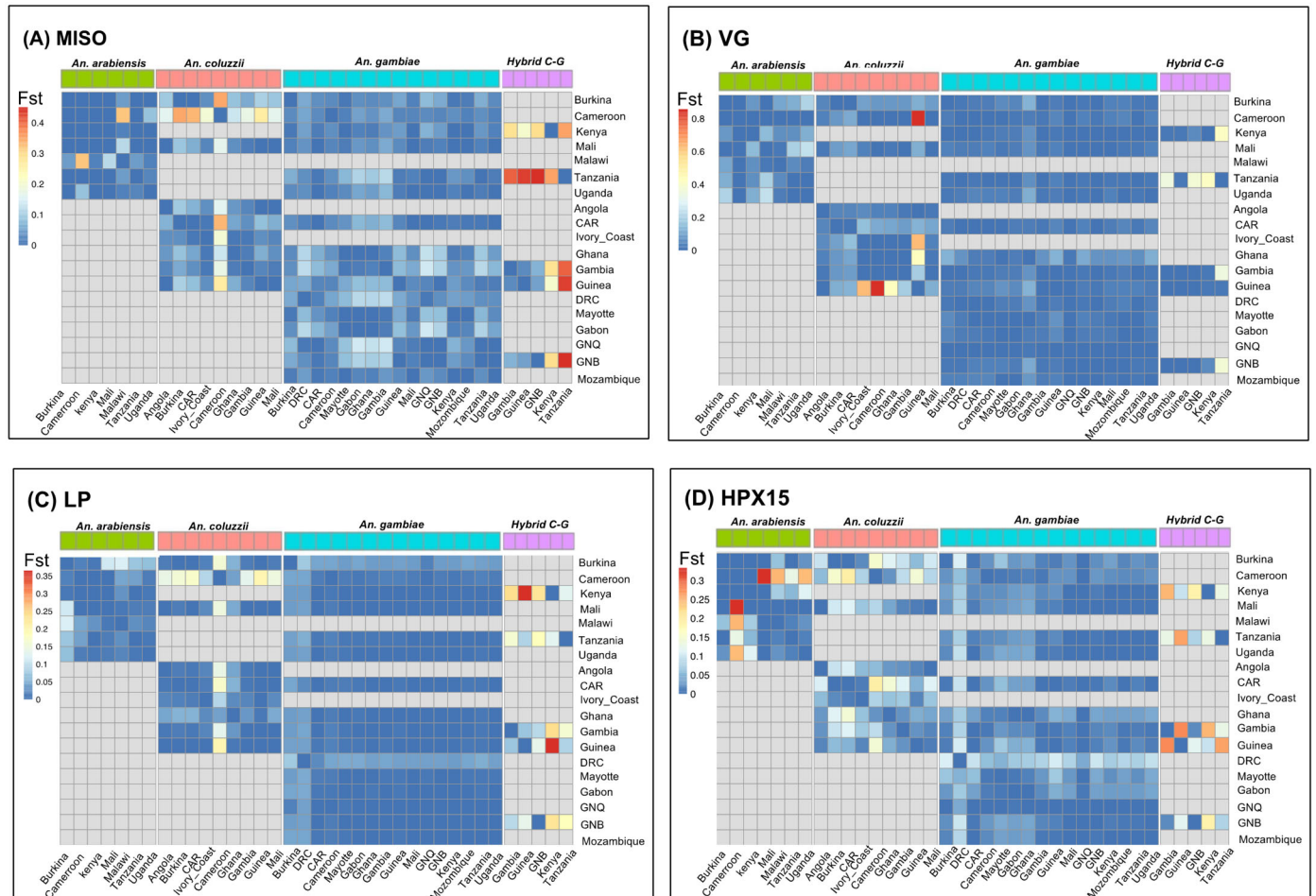


Figure 5. Heatmap of pairwise fixation index (F_{ST}) between populations for each of four species on four targetable genes involved in *Anopheles gambiae* s.l. fertility and *Plasmodium* infectivity using sequencing data from Malariagen phase 3 Ag1000G project. (A) Mating-induced stimulator of oogenesis protein (*MISO*, AGAP002620), (B) *Vitellogenin* (*VG*, AGAP004203), (C) *Lipophorin* (*LP*, AGAP001826), and (D) *Haem-peroxidase 15* (*HPX15*, AGAP013327). Abbreviations: GNB: Guinea Bissau, DRC: Democratic Republic of Congo, CAR: Central African Republic, GNQ: Equatorial Guinea, Hybrid C-G: Hybrid *Anopheles coluzzii-gambiae*. This F_{ST} heatmap showed a low genetic differentiation between populations for *Anopheles arabiensis*, *Anopheles coluzzii*, and *Anopheles gambiae* at all the four genes. However, the heatmap also showed a moderate genetic differentiation between populations of hybrid *Anopheles coluzzii-gambiae* at all four genes.

2.3. Low Genetic Diversity of the Selected Targetable Genes

Genetic diversity for the 2784 wild-caught *Anopheles gambiae* s.l. was estimated by nucleotide diversity (π), Tajima's D selection index, heterozygosity and the linkage disequilibrium (LD) of the non-synonymous SNPs (nsSNPs). Overall, π was consistently low in all populations (Figure 6). Most of the median values were less than 0.05 for *MISO*, *VG* and *HPX15* genes (Figure 6A,B,D), and within 0.5 to 1 for the *LP* gene (Figure 6C). This low nucleotide diversity suggests potential genetic drift, recent population bottlenecks, or selective sweeps, which are a consequence of positive selection, resulting in the rapid

increase in the frequency of one or a limited number of advantageous alleles, reducing the genetic diversity in the populations.

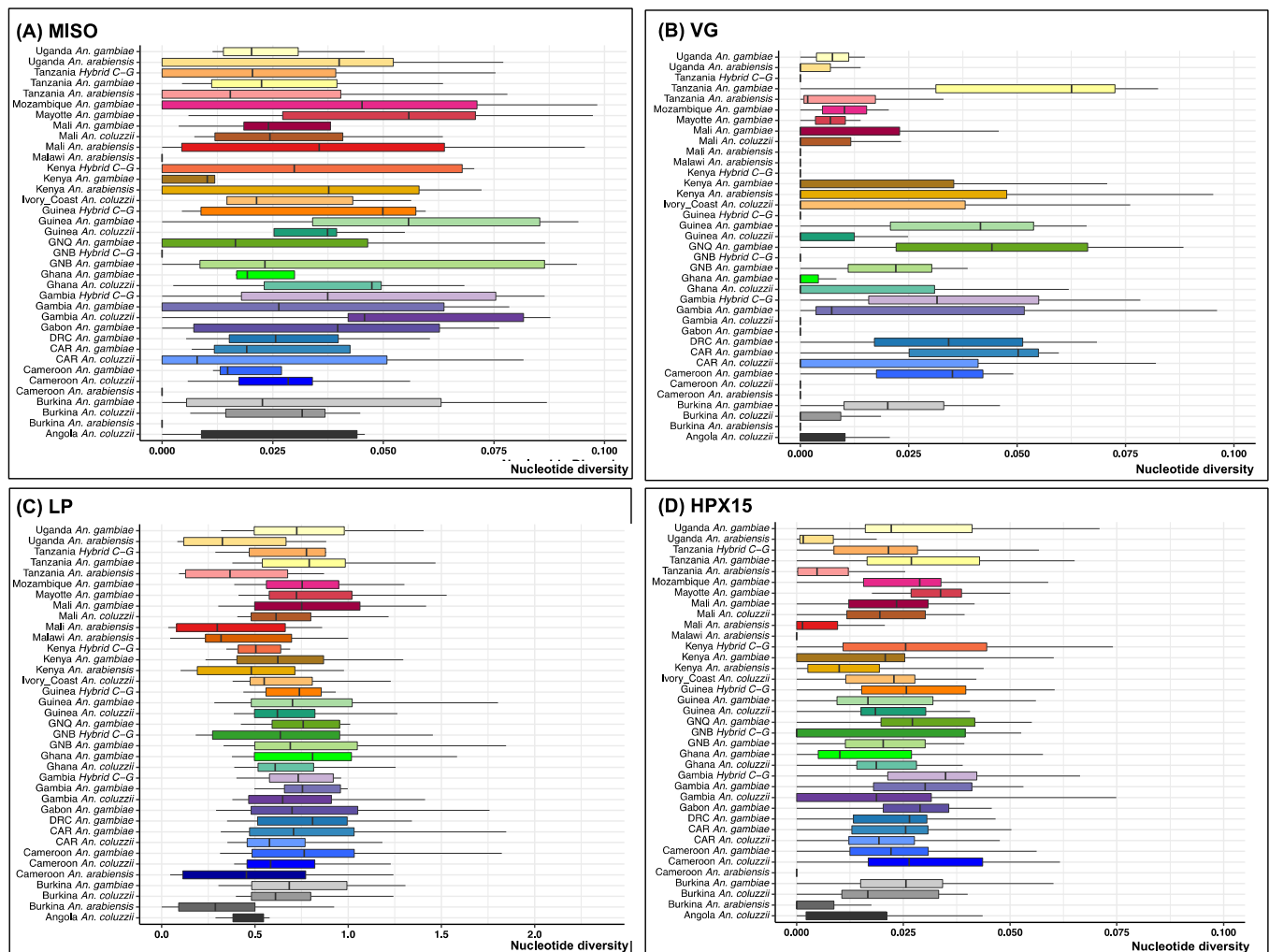


Figure 6. Boxplot of nucleotide diversity (π) for each of *Anopheles* species populations on four targetable genes involved in *Anopheles gambiae* s.l. fertility and *Plasmodium* infectivity using sequencing data from Malariagen phase 3 Ag1000G project. (A) Mating-induced stimulator of oogenesis protein (MISO, AGAP002620), (B) Vitellogenin (VG, AGAP004203), (C) Lipophorin (LP, AGAP001826), and (D) Haem-peroxidase 15 (HPX15, AGAP013327). Abbreviations: GNB: Guinea Bissau, DRC: Democratic Republic of Congo, CAR: Central African Republic, GNQ: Equatorial Guinea, Hybrid C-G: Hybrid *Anopheles coluzzii-gambiae*. The nucleotide diversity was overall low (>0.10) in all the population at MISO, VG, LP, and HPX15 genes (see panels (A–D)).

Furthermore, the Tajima's D values for each target locus within each *Anopheles* species population were mainly negative (Figure 7), suggesting an excess of rare alleles due to population growth of positive selection that reduces genetic diversity. However, there were exceptions of positive Tajima's D values for some species and populations. For MISO, Tajima D was positive for *An. gambiae* s.s. from Mozambique hybrids from Kenya, and *An. coluzzii* from Kenya (Figure 7A). Four *An. gambiae* populations (from Mozambique, Kenya, Ghana and Gambia), *An. coluzzii* from Angola, and hybrid C-G (from Kenya and Tanzania) had positive values at the LP gene (Figure 7C). At the HPX15 gene, positive Tajima's D was observed only for the hybrid population from Kenya and Tanzania (Figure 7D).

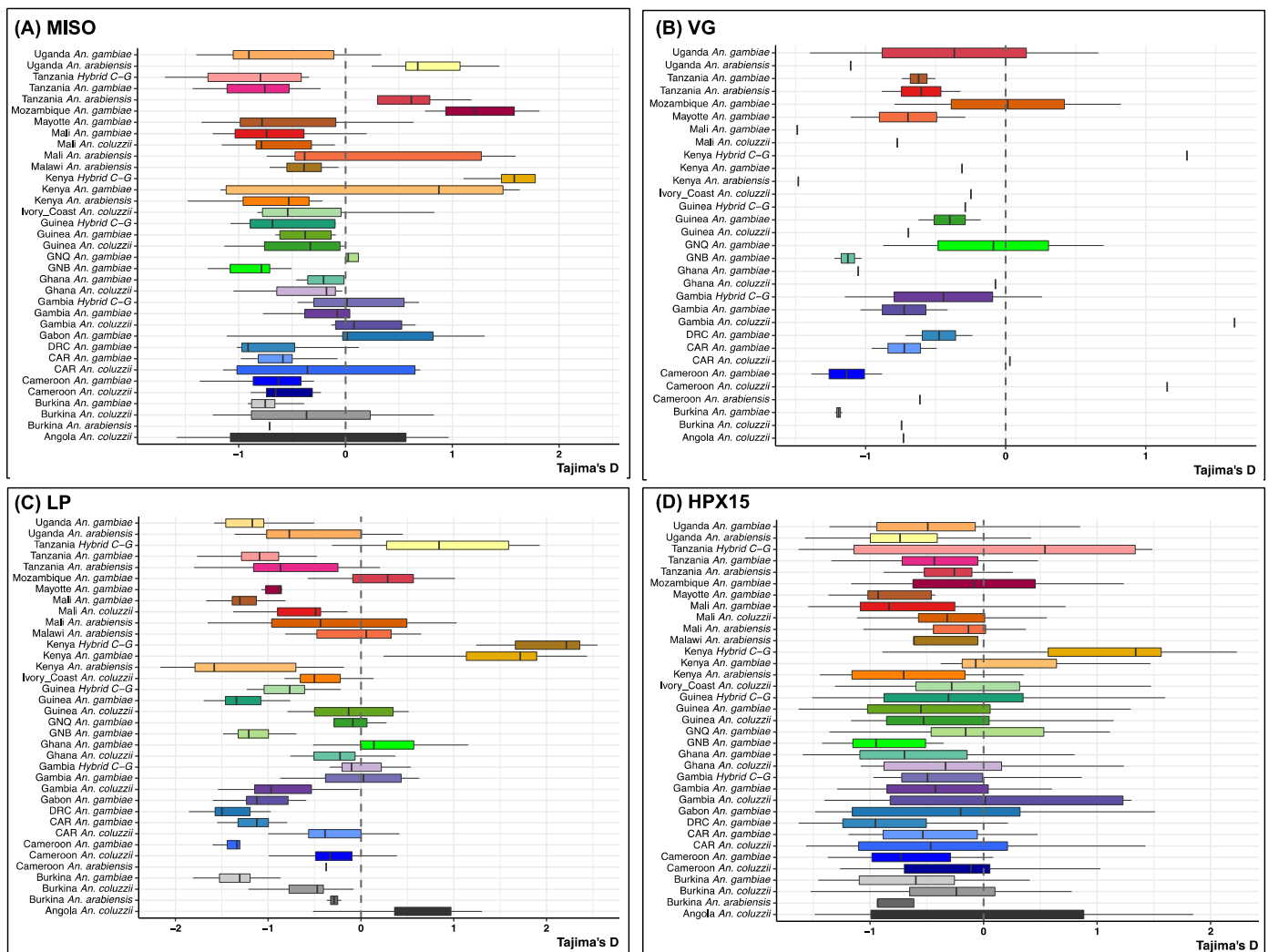


Figure 7. Boxplot of population selection index (Tajima's D) within each of *Anopheles* species populations on four targetable genes involved in *Anopheles gambiae* s.l. fertility and *Plasmodium* infectivity using sequencing data from Malariagen phase 3 Ag1000G project. (A) Mating-induced stimulator of oogenesis protein (*MISO*, AGAP002620), (B) *Vitellogenin* (*VG*, AGAP004203), (C) *Lipophorin* (*LP*, AGAP001826), and (D) *Haem-peroxidase 15* (*HPX15*, AGAP013327). Abbreviations: GNB: Guinea Bissau, DRC: Democratic Republic of Congo, CAR: Central African Republic, GNQ: Equatorial Guinea, Hybrid C-G: Hybrid *Anopheles coluzzii-gambiae*. The Tajima's D within each of the *Anopheles* species populations on four targetable genes was overall negative, suggesting a purifying selection which eliminates deleterious mutations and leads to a reduction in genetic diversity.

Additionally, the observed heterozygosity and the expected heterozygosity (potential genetic diversity within the population under the assumption of Hardy–Weinberg equilibrium) were estimated for each species population across the four target genes (Supplementary Materials Figure S3). In each case, the heterozygosity values were low, and the observed heterozygosity did not significantly deviate from the expected heterozygosity. These results suggest populations with reduced genetic diversity as observed with π and Tajima's D. Likewise, the pairwise linkage disequilibrium (LD) measurement between non-synonymous SNPs (nsSNPs) across the four genes was weak (Supplementary Materials Figure S4), with no associated SNPs found at the *MISO* gene. Some linked nsSNPs were found in the other genes, especially in the *HPX15*. However, no dominant haplotype (pairs of consecutive SNPs on the same chromosome with high LD, thus inherited together) was detected, with the exception of three SNPs in the *HPX15* gene (between 3L_10786531 and 3L_10786542).

3. Discussion

The burden of malaria in Africa remains staggering, with approximately half a million fatalities and two hundred million clinical cases reported annually [1]. Despite the widespread distribution and use of insecticide-treated bed nets, the efficacy of malaria control efforts has increasingly been jeopardized by the emergence of insecticide-resistant *Anopheles* mosquito populations. In light of the pressing need for novel and complementary vector control strategies, a promising approach to reducing malaria transmission involves the direct sterilization of female *Anopheles* mosquitoes. This innovative strategy could be achieved by applying small molecule inhibitors targeting critical genes of reproductive processes, thereby disrupting the reproductive synergy necessary for population maintenance.

Our study focused on several key genes: *MISO*, *Vg*, *Lp*, and *HPX15*, identified as significant contributors to fertility and vector competency in *Anopheles* mosquitoes. To elucidate the genetic patterns associated with these genes, we conducted a comprehensive genetic analysis of these loci in the natural populations of *Anopheles gambiae* s.l. across 19 countries in sub-Saharan Africa, utilizing SNP data from phase 3 of the Ag1000G project (Ag3.0). These data provided key insights into the genetic homogeneity of the four fertility and vector competency-related genes, required for refining the design of small molecule inhibitors against *Anopheles* as an innovative tool to control malaria transmission.

The main finding of this study is the observed genetic conservation in the four selected fertility and vector competency-related genes amongst *An. gambiae* s.l. populations across sub-Saharan Africa. The absence of genetic clustering and a low pairwise fixation index across the four genes (*MISO*, *Vg*, *Lp*, and *HPX15*) indicate a low differentiation between populations, with a slight divergence observed only in the hybrid population for *MISO*, *Lp*, and *HPX15* genes. Additionally, low heterozygosity (*He*) and nucleotide diversity (*Pi*), and overall Tajima's *D* suggest purifying selection that eliminates deleterious mutations or inbreeding, which reduces genetic diversity. Moreover, the limited ns-SNPs with low LD and no clear haplotype block supports the overall negative Tajima's *D* observed. However, there were exceptions, like positive Tajima's *D* observed for the hybrid population from Kenya, which could result from positive selection or demographic factors.

Genetic variation serves as the driving force of evolution, enabling organisms to survive and adapt to their environmental conditions [33–35]. Species within the complex *An. gambiae*, including the main vectors *An. gambiae*, *An. coluzzii*, and *An. arabiensis*, are crucial in malaria transmission in sub-Saharan Africa. However, *An. arabiensis* populations exhibit slightly distinct plasticity, behavior, and preferences compared to *An. gambiae* and *An. coluzzii* [36–38]. They also display unique patterns of seasonal and ecological distributions, such as a greater tolerance to dry environments [38–45]. These geographical contrasts, reproductive behaviors, or environmental adaptive selections may explain the slight genetic distance between the species. Recent genome-wide studies have shown the *An. gambiae* complex as the most polymorphic vector, contributing to its adaptability to insecticides [27,29]. However, we observed genetic diversity conservation across all four genes and *Anopheles* populations in sub-Saharan Africa, emphasizing evolutionary conservation of mechanisms driving fertility and vector competency. These findings could pave the way for the development of innovative tools targeting those genes and all vectors to sustain malaria control efforts in sub-Saharan Africa. These could include targeted interventions aimed at disrupting mosquito reproduction and overcoming the insecticide resistance challenge. Indeed, the reduced genetic diversity of the identified genes suggests that small molecule inhibitors designed to target these specific genes could have broad applicability across diverse *An. gambiae* s.l. populations as is well documented for drug or vaccine [46,47]. This could lead to cost-effective non-insecticide-based vector

control tools such as small molecule inhibitors that can block these enzymes and inhibit malaria transmission by a wide range of mosquitoes. Targeting these conserved genes has the advantage of developing the same tools against multiple species across different geographical regions. Moreover, the potential small molecule inhibitors targeting these genes can be integrated into existing vector control measures (LLINs or IRS approaches to optimize the reproductive disruption and transmission blocking strategies [48–50]), and to create synergies and enhance the overall efficacy of malaria control programs.

4. Conclusions

Our results present rare integrated findings on the biology factors and genetic features of the major malaria vector in natural populations. By revealing this significant degree of conservation of the genes (*MISO*, *Vg*, *Lp* and *HPX15*) within the members of the complex *An. gambiae* throughout sub-Saharan Africa, our data offer new insights for the development of sustainable malaria control tools. These genes could serve as effective targets for managing vector populations and reducing global malaria transmission. However, the genetics of these conserved genes should not be interpreted as a static condition. Continuous monitoring of mosquito populations is essential to detect any evolving genetic changes or the potential development of resistance against small molecule inhibitors.

5. Methods

5.1. Selection of Potential Targetable Genes Involved in *Anopheles gambiae* s.s PEST Fertility and *Plasmodium* Infectivity

An initial exploratory literature review was conducted to identify genes that are associated with either the *An. gambiae* fertility [13–22,24,25,51–62] or its infectivity to *Plasmodium* [15,21–23,58–78] (see Supplementary Materials Introduction and Table S1 for more details). Amongst these genes, we selected the set of genes that are involved in both the fertility and infectivity to *Plasmodium* for a further review. At the end of the review process, four genes were retained based on the following criteria that have been well characterized with publication evidence (Supplementary Materials Introduction). Those four genes were annotated based on the VectorBase [62] and named according to the *An. gambiae* s.s PEST genes IDs.

5.2. SNP Data Used

We used Single Nucleotide Polymorphisms (SNPs) data from the phase 3 *Anopheles gambiae* s.l. 1000 genomes (Ag1000G) project (Ag3.0), accessible online at the MalariaGEN website (<https://malariagen.github.io/vector-data/ag3/download.html#snp-calls-vcf-format>, accessed on 6 August 2021) [79]. These SNPs were genotyped from 2784 wild-caught *Anopheles* mosquitoes (including *An. arabiensis*, *An. gambiae* s.s, *An. coluzzii* and hybrids *An. gambiae-coluzzii*) that were collected from 19 sub-Saharan Africa countries (Supplementary Materials Introduction). They were used to investigate the genetic diversity of the selected genes. These data was released in February 2021 on the mosquito data-sharing home page of the Malaria Genomic Epidemiology Network (MalariaGEN) [80].

An. gambiae s.s was the predominant and widely dispersed species across sub-Saharan Africa, representing 57% of the dataset. *An. coluzzii* was collected mainly in West Africa (24%), where the highest rate of hybrid *An. gambiae-coluzzii* was observed (6%). *An. arabiensis* was primarily collected in East Africa, accounting for 13% of the dataset (Supplementary Materials Figure S2).

A comprehensive description of the population sampling, sequencing, alignment, genotyping, quality control, variant calling and filtering methods are available on the Ag3.0 MalariaGEN vector data user guide (<https://malariagen.github.io/vector-data/ag3>

[/methods.html](#), accessed on 6 August 2021). Briefly, mosquito samples were sequenced at high coverage using Illumina technology at the Wellcome Sanger Institute, and the MalariaGEN Resource Centre Team conducted a quality control check and preliminary analysis generating the VCF SNPs data as previously described [79].

The whole genome VCF files, metadata samples and other relevant data were downloaded and stored in the high-performance computing server (HPC) at the Medical Research Council Unit of The Gambia's (MRCG) at LSHTM for the subsequent analysis. Using the *Anopheles gambiae* s.s reference genome annotations, AgamP4 PEST [81] from VectorBase variant data for the four target genes of interest were extracted from the whole genome data using bcftools v1.9 [82]. The data were filtered with vcftools v0.1.17 [83] to exclude low-quality SNPs, retaining only sites with less than 10% genotype missing across individuals, a Phred score above 30, and a read depth beyond 10 but not exceeding 50 using PLINK v1.9 [84]. Finally, SNPs with a minor allele frequency greater than 1% were retained (Supplementary Materials Figure S5).

5.3. Data Analysis

Population structure analysis: We measured the genetic variance due to genetic structure in populations based on the Wright (1951) fixation index (F_{st}) [85]. F_{st} values were estimated between species and populations following the WC84 method (Weir & Cockerham, 1984) [86] implemented in the R package adegenet v2.1.7 [87]. The same R package was also used to perform a principal component analysis (PCA) to identify the main axes of variation in the multivariate dataset as described by Jolliffe [88]. To reduce bias in the F_{st} and PCA analyses, we pruned out SNPs with pairwise linkage disequilibrium (LD) values $r^2 > 0.25$ within a 50 bp window in each gene. This pruning was done with a step size of 5. After this pruning step, the remaining SNPs located on the shared chromosomes between the species were used to estimate the observed and expected heterozygosity using the R package hierfstat [89].

Genetic diversity analysis: To assess the level of polymorphism between and within populations, we calculated the nucleotide diversity (π) at each locus using the R package PopGenome [90]. The π was defined as the average number of nucleotide differences per site between two sequences [91]. Furthermore, we determined Tajima's D index per gene, a neutrality derived from comparing the mean number of segregating-sites and the average pairwise differences [92]. Tajima's D allows for detecting evolutionary departure from neutrality as proposed by Kimura [93] that may be due to selective pressure within populations. Also, as the linkage disequilibrium (LD) could provide highlights on the history of natural selection, gene conversion, mutation, and other forces that cause gene-frequency evolution [94]. We used the R packages LDheatmap v.1.0-6 [95] and Gaston v.1.5.7 to compute pairwise linkage disequilibrium (LD) between different non-synonymous SNP (nsSNPs) sites across all species using the genotype correlation coefficient (r^2) index between alleles at physically separate loci. Assessing LD between nsSNPs is valuable for identifying genomic regions that may harbor functionally significant genetic variations.

The observed pairwise LD (r^2) was averaged for each inter-SNP distance, and r^2 measures for each gene were calculated between all pairs of SNPs within each population.

Data and codes used to analyze the data are available in the GitHub 3.4.9 repository: https://github.com/mfdiop/Genetic_Diversity_Anopheles_Gambiae, accessed on 6 August 2021.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes16050543/s1>, SI1 Introduction: Genes involved in fertility and vector competency of *Anopheles gambiae*. Figure S2. Composition of *Anopheles gambiae* s.l. database used. A total of 2784 whole genomes vcf SNP datasets of wild-caught *Anopheles gambiae* s.l. mosquitoes, collected in

19 sub-Saharan African countries, were utilized in this study. These datasets were obtained from the phase 3 Ag1000G project (https://www.malariagen.net/data_package/ag1000g-phase3-snp/, accessed on 6 August 2021). The numbers within the bar plots indicate the sample count for each population or species from a specific country. Figure S3. Observed Heterozygosity (Ho in Blue) and Expected Heterozygosity (He in Gray) for each species on the four targetable genes involved in *Anopheles gambiae* s.l. fertility and *Plasmodium* infectivity using sequencing data from the MalariaGEN phase 3 Ag1000G project. (A) Mating-induced stimulator of oogenesis protein (MISO, AGAP002620), (B) Vitellogenin (VG, AGAP004203), (C) Lipophorin (LP, AGAP001826), and (D) Haem-peroxidase 15 (HPX15, AGAP013327). At each of the genes, the observed heterozygosity was not significantly deviated from the expected heterozygosity providing insights into the low genetic diversity within the populations. Figure S4. Pairwise linkage disequilibrium (LD) measurement between non-synonymous SNPs across the four targetable genes involved in *Anopheles gambiae* s.l. fertility and *Plasmodium* infectivity using sequencing data from the MalariaGEN phase 3 Ag1000G project. (A) Mating-induced stimulator of oogenesis protein (MISO, AGAP002620), (B) Vitellogenin (VG, AGAP004203), (C) Lipophorin (LP/AGAP001826), and (D) Haem-peroxidase 15 (HPX15, AGAP013327). No associated non-synonymous SNPs (nsSNPs) were found at the MISO gene, whereas few and low linked nsSNPs with no clear haplotype block were found in the other genes except the potential haplotype on the HPX15 gene between 3L_10786531 and 3L_10786542 that might be taken into consideration for further investigation. Figure S5. Global and intra-species minor allele frequencies (MAF) distribution of SNPs at four targetable genes involved in *Anopheles gambiae* s.l. fertility and *Plasmodium* infectivity using sequencing data from the MalariaGEN phase 3 Ag1000G project. (A) Mating-induced stimulator of oogenesis protein (MISO, AGAP002620), (B) Vitellogenin (VG, AGAP004203), (C) Lipophorin (LP/AGAP001826), and (D) Haem-peroxidase 15 (HPX15, AGAP013327). The MAF has shown a skewed distribution with most SNPs having low MAF (close to 0) indicating that most SNPs are monomorphic (only one allele is common) in the population, indicating low genetic diversity. Table S1. List of genes reviewed.

Author Contributions: Conceptualization: F.S., A.D. (Abdoulaye Diabate), A.A.-N., I.D. (Ibrahima Dia) and B.S.A. Data curation: A.A.-N. and B.S.A. Formal analysis: F.S., M.F.D., K.M., A.D. (Amadou Diallo), I.D. (Idrissa Dieng), I.D. (Ibrahima Dia) and B.S.A. Funding acquisition: A.A.-N., I.D. (Ibrahima Dia) and B.S.A. Investigation: F.S., M.F.D., K.M., A.D. (Amadou Diallo), I.D. (Idrissa Dieng), M.N., A.D. (Abdoulaye Diabate), A.A.-N., I.D. (Ibrahima Dia) and B.S.A. Methodology: F.S., M.F.D., K.M., I.D. (Ibrahima Dia) and B.S.A. Project administration: B.S.A. Resources: I.D. (Ibrahima Dia) and B.S.A. Supervision: I.D. (Ibrahima Dia) and B.S.A. Validation: I.D. (Ibrahima Dia), B.S.A. Visualization: I.D. (Ibrahima Dia) and B.S.A. Writing—original draft: F.S., M.F.D., I.D. (Ibrahima Dia) and B.S.A. Writing—review & editing: M.F.D., K.M., M.N., A.D. (Abdoulaye Diabate) and A.A.-N. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported through the sub-awards from MARCAD DELTAS Africa Initiative grant DEL-15-010 of the African Academy of Sciences (AAS)’s Alliance for Accelerating Excellence in Science in Africa (AESA) and the Crick Africa Network (CAN), which receives its funding from the UK’s Global Challenges Research Fund (MR/P028071/1). B.S.A. was supported by MARCAD and CAN fellowships. F.S. was supported by the fellowships from WANIDA (West African Network of Infectious Diseases African Centre of Excellence) and CEA ITECH-MTV (African Centre of Excellence in Biotechnological Innovations for Vector-borne Diseases Elimination).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author. Raw sequence data are available on MalariaGEN (https://www.malariagen.net/data_package/ag1000g-phase3-snp/, accessed on 6 August 2021).

Acknowledgments: We thank Alistair Miles and Mara Lawniczak from MalariaGEN Vector Observatory, which is an international collaboration working to build capacity for malaria vector genomic surveillance. 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 to make the data available.

Conflicts of Interest: The authors state that they have no competing financial or non-financial interest.

References

- World Health Organization. World Malaria Report 2024—Addressing Inequity in the Global Malaria Response. Available online: <https://www.who.int/publications/i/item/9789240104440> (accessed on 3 February 2025).
- Bhatt, S.; Weiss, D.J.; Cameron, E.; Bisanzio, D.; Mappin, B.; Dalrymple, U.; Battle, K.; Moyes, C.L.; Henry, A.; Eckhoff, P.A.; et al. The effect of malaria control on *Plasmodium falciparum* in Africa between 2000 and 2015. *Nature* **2015**, *526*, 207–211. [CrossRef] [PubMed]
- Ranson, H.; N’guessan, R.; Lines, J.; Moiroux, N.; Nkuni, Z.; Corbel, V. Pyrethroid resistance in African anopheline mosquitoes: What are the implications for malaria control? *Trends Parasitol.* **2011**, *27*, 91–98. [CrossRef] [PubMed]
- Ranson, H.; Lissenden, N. Insecticide Resistance in African Anopheles Mosquitoes: A Worsening Situation that Needs Urgent Action to Maintain Malaria Control. *Trends Parasitol.* **2016**, *32*, 187–196. [CrossRef] [PubMed]
- Hemingway, J.; Ranson, H.; Magill, A.; Kolaczinski, J.; Fornadel, C.; Gimnig, J.; Coetzee, M.; Simard, F.; Roch, D.K.; Hinzoumbe, C.K.; et al. Averting a malaria disaster: Will insecticide resistance derail malaria control? *Lancet* **2016**, *387*, 1785–1788. [CrossRef]
- Riveron, J.M.; Tchouakui, M.; Mugenzi, L.; Menze, B.D.; Chiang, M.-C.; Wondji, C.S. Insecticide Resistance in Malaria Vectors: An Update at a Global Scale. In *Towards Malaria Elimination—A Leap Forward*; IntechOpen: London, UK, 2018. [CrossRef]
- Hancock, P.A.; Hendriks, C.J.M.; Tangena, J.-A.; Gibson, H.; Hemingway, J.; Coleman, M.; Gething, P.W.; Cameron, E.; Bhatt, S.; Moyes, C.L. Mapping trends in insecticide resistance phenotypes in African malaria vectors. *PLoS Biol.* **2020**, *18*, e3000633. [CrossRef]
- Sougoufara, S.; Doucouré, S.; Backé Sembéne, P.M.; 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. [CrossRef]
- Sougoufara, S.; Ottih, E.C.; Tripet, F. The need for new vector control approaches targeting outdoor biting Anopheline malaria vector communities. *Parasit. Vectors* **2020**, *13*, 295. [CrossRef]
- Sanou, A.; Nelli, L.; Guelbéogo, W.M.; Cissé, F.; Tapsoba, M.; Ouédraogo, P.; Sagnon, N.; Ranson, H.; Matthiopoulos, J.; Ferguson, H.M. Insecticide resistance and behavioural adaptation as a response to long-lasting insecticidal net deployment in malaria vectors in the Cascades region of Burkina Faso. *Sci. Rep.* **2021**, *11*, 17569. [CrossRef]
- Mashatola, T.; Ndo, C.; Koekemoer, L.L.; Dandalo, L.C.; Wood, O.R.; Malakoane, L.; Poumachu, Y.; Lobb, L.N.; Kaiser, M.; Bourtzis, K.; et al. A review on the progress of sex-separation techniques for sterile insect technique applications against *Anopheles arabiensis*. *Parasit. Vectors* **2018**, *11*, 646. [CrossRef]
- Yao, F.A.; Millogo, A.-A.; Epopa, P.S.; North, A.; Noulain, F.; Dao, K.; Drabo, M.; Guissou, C.; Kekele, S.; Namountougou, M.; et al. Mark-release-recapture experiment in Burkina Faso demonstrates reduced fitness and dispersal of genetically-modified sterile malaria mosquitoes. *Nat. Commun.* **2022**, *13*, 796. [CrossRef]
- Atella, G.C.; Silva-Neto, M.A.C.; Golodne, D.M.; Arefin, S.; Shahabuddin, M. Anopheles gambiae lipophorin: Characterization and role in lipid transport to developing oocyte. *Insect Biochem. Mol. Biol.* **2006**, *36*, 375–386. [CrossRef] [PubMed]
- Rogers, D.W.; Whitten, M.M.A.; Thailayil, J.; Soichot, J.; Levashina, E.A.; Catteruccia, F. Molecular and cellular components of the mating machinery in *Anopheles gambiae* females. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 19390–19395. [CrossRef] [PubMed]
- Rono, M.K.; Whitten, M.M.A.; Oulad-Abdelghani, M.; Levashina, E.A.; Marois, E. The major yolk protein vitellogenin interferes with the anti-plasmodium response in the malaria mosquito *Anopheles gambiae*. *PLoS Biol.* **2010**, *8*, e1000434. [CrossRef] [PubMed]
- Baldini, F.; Gabrieli, P.; Rogers, D.W.; Catteruccia, F. Function and composition of male accessory gland secretions in *Anopheles gambiae*: A comparison with other insect vectors of infectious diseases. *Pathog. Glob. Health* **2012**, *106*, 82–93. [CrossRef]
- Baldini, F.; Gabrieli, P.; South, A.; Valim, C.; Mancini, F.; Catteruccia, F. The interaction between a sexually transferred steroid hormone and a female protein regulates oogenesis in the malaria mosquito *Anopheles gambiae*. *PLoS Biol.* **2013**, *11*, e1001695. [CrossRef]
- Le, B.V.; Nguyen, J.B.; Logarajah, S.; Wang, B.; Marcus, J.; Williams, H.P.; Catteruccia, F.; Baxter, R.H.G. Characterization of *Anopheles gambiae* transglutaminase 3 (AgTG3) and its native substrate PlugIn. *J. Biol. Chem.* **2013**, *288*, 4844–4853. [CrossRef]
- Gabrieli, P.; Kakani, E.G.; Mitchell, S.N.; Mameli, E.; Want, E.J.; Mariezcurrena Anton, A.; Serrao, A.; Baldini, F.; Catteruccia, F. Sexual transfer of the steroid hormone 20E induces the postmating switch in *Anopheles gambiae*. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 16353–16358. [CrossRef]

20. Shaw, W.R.; Teodori, E.; Mitchell, S.N.; Baldini, F.; Gabrieli, P.; Rogers, D.W.; Catteruccia, F. Mating activates the heme peroxidase HPX15 in the sperm storage organ to ensure fertility in *Anopheles gambiae*. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5854–5859. [\[CrossRef\]](#)
21. Kajla, M.; Choudhury, T.P.; Kakani, P.; Gupta, K.; Dhawan, R.; Gupta, L.; Kumar, S. Silencing of *Anopheles stephensi* Heme Peroxidase HPX15 Activates Diverse Immune Pathways to Regulate the Growth of Midgut Bacteria. *Front. Microbiol.* **2016**, *7*, 1351. [\[CrossRef\]](#)
22. Kajla, M.; Kakani, P.; Choudhury, T.P.; Kumar, V.; Gupta, K.; Dhawan, R.; Gupta, L.; Kumar, S. *Anopheles stephensi* Heme Peroxidase HPX15 Suppresses Midgut Immunity to Support Plasmodium Development. *Front. Immunol.* **2017**, *8*, 249. [\[CrossRef\]](#)
23. Cui, Y.; Niu, G.; Li, V.L.; Wang, X.; Li, J. Analysis of blood-induced *Anopheles gambiae* midgut proteins and sexual stage *Plasmodium falciparum* interaction reveals mosquito genes important for malaria transmission. *Sci. Rep.* **2020**, *10*, 14316. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Marcenac, P.; Shaw, W.R.; Kakani, E.G.; Mitchell, S.N.; South, A.; Werling, K.; Marrogi, E.; Abernathy, D.G.; Yerbanga, R.S.; Dabiré, R.K.; et al. A mating-induced reproductive gene promotes *Anopheles* tolerance to *Plasmodium falciparum* infection. *PLoS Pathog.* **2020**, *16*, e1008908. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Le, B.V.; Klöck, C.; Schatz, A.; Nguyen, J.B.; Kakani, E.G.; Catteruccia, F.; Khosla, C.; Baxter, R.H.G. Dihydroisoxazole inhibitors of *Anopheles gambiae* seminal transglutaminase AgTG3. *Malar. J.* **2014**, *13*, 210. [\[CrossRef\]](#)
26. Bhattacharya, R.; Rose, P.; Burley, S.; Prlić, A. Impact of genetic variation on three dimensional structure and function of proteins. *PLoS ONE* **2017**, *12*, e0171355. [\[CrossRef\]](#)
27. *Anopheles Gambiae* 1000 Genomes Consortium; Data Analysis Group; Partner Working Group. Genetic diversity of the African malaria vector *Anopheles gambiae*. *Nature* **2017**, *552*, 96–100. [\[CrossRef\]](#)
28. Amambua-Ngwa, A.; Amenga-Etego, L.; Kamau, E.; Amato, R.; Ghansah, A.; Golassa, L.; Randrianarivelojosia, M.; Ishengoma, D.; Apinjoh, T.; Maïga-Ascofaré, O.; et al. Major subpopulations of *Plasmodium falciparum* in sub-Saharan Africa. *Science* **2019**, *365*, 813–816. [\[CrossRef\]](#)
29. *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*, 1533–1546. [\[CrossRef\]](#)
30. Diallo, M.; Hamid-Adiamoh, M.; Sy, O.; Sarr, P.C.; Manneh, J.; Ndiath, M.O.; Gaye, O.; Faye, O.; Konaté, L.; Sesay, A.K.; et al. Evolution of the Pyrethroids Target-Site Resistance Mechanisms in Senegal: Early Stage of the Vgsc-1014F and Vgsc-1014S Allelic Frequencies Shift. *Genes* **2021**, *12*, 1948. [\[CrossRef\]](#)
31. Diallo, M.; Kolley, E.S.; Dia, A.K.; Oboh, M.A.; Seck, F.; Manneh, J.; Sesay, A.K.; Diédhiou, S.M.; Sarr, P.C.; Sy, O.; et al. Evolution of the Ace-1 and Gste2 Mutations and Their Potential Impact on the Use of Carbamate and Organophosphates in IRS for Controlling *Anopheles gambiae* s.l., the Major Malaria Mosquito in Senegal. *Pathogens* **2022**, *11*, 1021. [\[CrossRef\]](#)
32. Ishengoma, D.S.; Saidi, Q.; Sibley, C.H.; Roper, C.; Alifrangis, M. Deployment and utilization of next-generation sequencing of *Plasmodium falciparum* to guide anti-malarial drug policy decisions in sub-Saharan Africa: Opportunities and challenges. *Malar. J.* **2019**, *18*, 267. [\[CrossRef\]](#)
33. Höglund, J. Genetic diversity in changing environments. In *Evolutionary Conservation Genetics*; Höglund, J., Ed.; Oxford University Press: Oxford, UK, 2009; pp. 60–80. [\[CrossRef\]](#)
34. Cardoso, J.G.R.; Andersen, M.R.; Herrgård, M.J.; Sonnenschein, N. Analysis of genetic variation and potential applications in genome-scale metabolic modeling. *Front. Bioeng. Biotechnol.* **2015**, *3*, 13. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Birader, K. Genetic Diversity and the Adaptation of Species to Changing Environments. *J. Biodivers. Endanger. Species* **2023**, *11*, 3. [\[CrossRef\]](#)
36. Costantini, C.; Sagnon, N.; della Torre, A.; Coluzzi, M. Mosquito behavioural aspects of vector-human interactions in the *Anopheles gambiae* complex. *Parassitologia* **1999**, *41*, 209–217.
37. Main, B.J.; Lee, Y.; Ferguson, H.M.; Kreppel, K.S.; Kihonda, A.; Govella, N.J.; Collier, T.C.; Cornel, A.J.; Eskin, E.; Kang, E.Y.; et al. The Genetic Basis of Host Preference and Resting Behavior in the Major African Malaria Vector, *Anopheles arabiensis*. *PLoS Genet.* **2016**, *12*, e1006303. [\[CrossRef\]](#)
38. Ashine, T.; Getachew, D.; Demisse, M.; Lobo, N.F.; Tadesse, F.G. *Anopheles arabiensis*. *Trends Parasitol.* **2023**, *40*, 91–92. [\[CrossRef\]](#)
39. Coetzee, M.; Craig, M.; le Sueur, D.; Coetzee, M.; Craig, M.; le Sueur, D. Distribution of African Malaria Mosquitoes Belonging to the *Anopheles gambiae* Complex. *Parasitol. Today* **2000**, *16*, 74–77. [\[CrossRef\]](#)
40. Koenraadt, C.J.M.; Githeko, A.K.; Takken, W. The effects of rainfall and evapotranspiration on the temporal dynamics of *Anopheles gambiae* s.s. and *Anopheles arabiensis* in a Kenyan village. *Acta Trop.* **2004**, *90*, 141–153. [\[CrossRef\]](#)
41. Gray, E.M.; Bradley, T.J. Physiology of desiccation resistance in *Anopheles gambiae* and *Anopheles arabiensis*. *Am. J. Trop. Med. Hyg.* **2005**, *73*, 553–559. [\[CrossRef\]](#)
42. Sinka, M.E.; Bangs, M.J.; Manguin, S.; Coetzee, M.; Mbogo, C.M.; Hemingway, J.; Patil, A.P.; Temperley, W.H.; Gething, P.W.; Kabaria, C.W.; et al. The dominant *Anopheles* vectors of human malaria in Africa, Europe and the Middle East: Occurrence data, distribution maps and bionomic précis. *Parasit. Vectors* **2010**, *3*, 117. [\[CrossRef\]](#)

43. Ng'habi, K.R.; Knols, B.G.; Lee, Y.; Ferguson, H.M.; Lanzaro, G.C. Population genetic structure of *Anopheles arabiensis* and *Anopheles gambiae* in a malaria endemic region of southern Tanzania. *Malar. J.* **2011**, *10*, 289. [\[CrossRef\]](#)
44. Derua, Y.A.; Alifrangis, M.; Hosea, K.M.; Meyrowitsch, D.W.; Magesa, S.M.; Pedersen, E.M.; Simonsen, P.E. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Coetzee, M.; Hunt, R.H.; Wilkerson, R.; Della Torre, A.; Coulibaly, M.B.; Besansky, N.J. *Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex. *Zootaxa* **2013**, 3619, 246–274. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Verkhivker, G.M. Leveraging Structural Diversity and Allosteric Regulatory Mechanisms of Protein Kinases in the Discovery of Small Molecule Inhibitors. *Curr. Med. Chem.* **2017**, *24*, 4838–4872. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Oboh, M.A.; Asmorom, N.; Falade, C.; Ojurongbe, O.; Thomas, B.N. High genetic and haplotype diversity in vaccine candidate Pfceltos but not Pfrh5 among malaria-infected children in Ibadan, Nigeria. *PeerJ* **2023**, *11*, e16519. [\[CrossRef\]](#)
48. Koama, B.; Namountougou, M.; Sanou, R.; Ndo, S.; Ouattara, A.; Dabiré, R.K.; Malone, D.; Diabaté, A. The sterilizing effect of pyriproxyfen on the malaria vector *Anopheles gambiae*: Physiological impact on ovaries development. *Malar. J.* **2015**, *14*, 101. [\[CrossRef\]](#)
49. Paton, D.G.; Childs, L.M.; Itoe, M.A.; Holmdahl, I.E.; Buckee, C.O.; Catteruccia, F. Exposing *Anopheles* mosquitoes to antimalarials blocks *Plasmodium* parasite transmission. *Nature* **2019**, 567, 239–243. [\[CrossRef\]](#)
50. Paton, D.G.; Probst, A.S.; Ma, E.; Adams, K.L.; Shaw, W.R.; Singh, N.; Bopp, S.; Volkman, S.K.; Hien, D.F.S.; Paré, P.S.L.; et al. Using an antimalarial in mosquitoes overcomes *Anopheles* and *Plasmodium* resistance to malaria control strategies. *PLoS Pathog.* **2022**, *18*, e1010609. [\[CrossRef\]](#)
51. Pondeville, E.; Maria, A.; Jacques, J.-C.; Bourgouin, C.; Dauphin-Villemant, C. *Anopheles gambiae* males produce and transfer the vitellogenic steroid hormone 20-hydroxyecdysone to females during mating. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 19631–19636. [\[CrossRef\]](#)
52. Rogers, D.W.; Baldini, F.; Battaglia, F.; Panico, M.; Dell, A.; Morris, H.R.; Catteruccia, F. Transglutaminase-mediated semen coagulation controls sperm storage in the malaria mosquito. *PLOS Biol.* **2009**, *7*, e1000272. [\[CrossRef\]](#)
53. Dottorini, T.; Nicolaides, L.; Ranson, H.; Rogers, D.W.; Crisanti, A.; Catteruccia, F. A genome-wide analysis in *Anopheles gambiae* mosquitoes reveals 46 male accessory gland genes, possible modulators of female behavior. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16215–16220. [\[CrossRef\]](#)
54. Liang, J.; Hodge, J.M.; Sharakhov, I.V. Asymmetric phenotypes of sterile hybrid males from reciprocal crosses between species of the *Anopheles gambiae* Complex. *Front. Ecol. Evol.* **2021**, *9*, 660207. [\[CrossRef\]](#)
55. Thailayil, J.; Gabrieli, P.; Caputo, B.; Bascuñán, P.; South, A.; Diabate, A.; Dabire, R.; Della Torre, A.; Catteruccia, F. Analysis of natural female post-mating responses of *Anopheles gambiae* and *Anopheles coluzzii* unravels similarities and differences in their reproductive ecology. *Sci. Rep.* **2018**, *8*, 6594. [\[CrossRef\]](#)
56. Kumar, S.; Molina-Cruz, A.; Gupta, L.; Rodrigues, J.; Barillas-Mury, C. A peroxidase/dual oxidase system modulates midgut epithelial immunity in *Anopheles gambiae*. *Science* **2010**, *327*, 1644–1648. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Bascuñán, P.; Gabrieli, P.; Mameli, E.; Catteruccia, F. Mating-regulated atrial proteases control reinsemination rates in *Anopheles gambiae* Females. *Sci. Rep.* **2020**, *10*, 21974. [\[CrossRef\]](#)
58. Peirce, M.J.; Mitchell, S.N.; Kakani, E.G.; Scarpelli, P.; South, A.; Shaw, W.R.; Werling, K.L.; Gabrieli, P.; Marcenac, P.; Bordonni, M.; et al. JNK signaling regulates oviposition in the malaria vector *Anopheles gambiae*. *Sci. Rep.* **2020**, *10*, 14344. [\[CrossRef\]](#)
59. Mendes, A.M.; Schlegelmilch, T.; Cohuet, A.; Awono-Ambene, P.; De Iorio, M.; Fontenille, D.; Morlais, I.; Christophides, G.K.; Kafatos, F.C.; Vlachou, D. Conserved mosquito/parasite interactions affect development of *Plasmodium falciparum* in Africa. *PLoS Pathog.* **2008**, *4*, e1000069. [\[CrossRef\]](#)
60. Vlachou, D.; Schlegelmilch, T.; Christophides, G.K.; Kafatos, F.C. Functional genomic analysis of midgut epithelial responses in *Anopheles* during *Plasmodium* Invasion. *Curr. Biol.* **2005**, *15*, 1185–1195. [\[CrossRef\]](#)
61. Upton, L.M.; Povelones, M.; Christophides, G.K. *Anopheles gambiae* blood feeding initiates an anticipatory defense response to *Plasmodium berghei*. *J. Innate Immun.* **2015**, *7*, 74–86. [\[CrossRef\]](#)
62. Green, E.; Jaouen, E.; Klug, D.; Olmo, R.P.; Gautier, A.; Blandin, S.; Marois, E. A population modification gene drive targeting both saglin and lipophorin impairs *Plasmodium* transmission in *Anopheles* mosquitoes. *Elife* **2023**, *12*, e93142. [\[CrossRef\]](#)
63. Christophides, G.K.; Zdobnov, E.; Barillas-Mury, C.; Birney, E.; Blandin, S.; Blass, C.; Brey, P.T.; Collins, F.H.; Danielli, A.; Dimopoulos, G.; et al. Immunity-related genes and gene families in *Anopheles gambiae*. *Science* **2002**, *298*, 159–165. [\[CrossRef\]](#)
64. Wormald, S.; Hilton, D.J. Inhibitors of cytokine signal transduction. *J Biol Chem* **2004**, *279*, 821–824. [\[CrossRef\]](#) [\[PubMed\]](#)
65. Frolet, C.; Thoma, M.; Blandin, S.; Hoffmann, J.A.; Levashina, E.A. Boosting NF-kappaB-dependent basal immunity of *Anopheles gambiae* aborts development of *Plasmodium berghei*. *Immunity* **2006**, *25*, 677–685. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Garver, L.S.; Dong, Y.; Dimopoulos, G. Caspar controls resistance to *Plasmodium falciparum* in diverse anopheline species. *PLoS Pathog.* **2009**, *5*, e1000335. [\[CrossRef\]](#)

67. Gupta, L.; Molina-Cruz, A.; Kumar, S.; Rodrigues, J.; Dixit, R.; Zamora, R.E.; Barillas-Mury, C. The STAT pathway mediates late-phase immunity against *Plasmodium* in the mosquito *Anopheles gambiae*. *Cell Host Microbe* **2009**, *5*, 498–507. [\[CrossRef\]](#)
68. Garver, L.S.; de Almeida Oliveira, G.; Barillas-Mury, C. The JNK pathway is a key mediator of *Anopheles gambiae* antiplasmodial immunity. *PLOS Pathog.* **2013**, *9*, e1003622. [\[CrossRef\]](#)
69. Dong, Y.; Das, S.; Cirimotich, C.; Souza-Neto, J.A.; McLean, K.J.; Dimopoulos, G. Engineered *Anopheles* immunity to *Plasmodium* infection. *PLOS Pathog.* **2011**, *7*, e1002458. [\[CrossRef\]](#)
70. Fraiture, M.; Baxter, R.H.G.; Steinert, S.; Chelliah, Y.; Frolet, C.; Quispe-Tintaya, W.; Hoffmann, J.A.; Blandin, S.A.; Levashina, E.A. Two mosquito LRR proteins function as complement control factors in the TEP1-mediated killing of *Plasmodium*. *Cell Host Microbe* **2009**, *5*, 273–284. [\[CrossRef\]](#)
71. Povelones, M.; Upton, L.M.; Sala, K.A.; Christophides, G.K. Structure-function analysis of the *Anopheles gambiae* LRIM1/APL1C complex and its interaction with complement C3-Like protein TEP1. *PLOS Pathog.* **2011**, *7*, e1002023. [\[CrossRef\]](#)
72. de Almeida Oliveira, G.; Lieberman, J.; Barillas-Mury, C. Epithelial nitration by a peroxidase/NOX5 system mediates mosquito antiplasmodial immunity. *Science* **2012**, *335*, 856–859. [\[CrossRef\]](#)
73. Horton, A.A.; Wang, B.; Camp, L.; Price, M.S.; Arshi, A.; Nagy, M.; Nadler, S.A.; Faeder, J.R.; Luckhart, S. The mitogen-activated protein kinase from *Anopheles gambiae*: Identification, phylogeny and functional characterization of the ERK, JNK and P38 MAP Kinases. *BMC Genom.* **2011**, *12*, 574. [\[CrossRef\]](#)
74. Blandin, S.; Shiao, S.-H.; Moita, L.F.; Janse, C.J.; Waters, A.P.; Kafatos, F.C.; Levashina, E.A. Complement-like protein TEP1 Is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. *Cell* **2004**, *116*, 661–670. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Dong, Y.; Dimopoulos, G. *Anopheles* fibrinogen-related proteins provide expanded pattern recognition capacity against bacteria and malaria parasites. *J. Biol. Chem.* **2009**, *284*, 9835–9844. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Nsango, S.E.; Pompon, J.; Xie, T.; Rademacher, A.; Fraiture, M.; Thoma, M.; Awono-Ambene, P.H.; Moyou, R.S.; Morlais, I.; Levashina, E.A. AP-1/Fos-TGase2 axis mediates wounding-induced *Plasmodium falciparum* Killing in *Anopheles gambiae*. *J. Biol. Chem.* **2013**, *288*, 16145–16154. [\[CrossRef\]](#)
77. Adedeji, E.O.; Ogunlana, O.O.; Fatumo, S.; Beder, T.; Ajamma, Y.; Koenig, R.; Adebisi, E. *Anopheles* metabolic proteins in malaria transmission, prevention and control: A review. *Parasites Vectors* **2020**, *13*, 465. [\[CrossRef\]](#)
78. Giraldo-Calderón, G.I.; Emrich, S.J.; MacCallum, R.M.; Maslen, G.; Dialynas, E.; Topalis, P.; Ho, N.; Gesing, S.; the VectorBase Consortium; Madey, G.; et al. VectorBase: An updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Res.* **2015**, *43*, D707–D713. [\[CrossRef\]](#)
79. Ag3.0 (Ag1000G Phase 3) SNP Data Release | MalariaGEN. Available online: https://www.malariagen.net/data_package/ag1000g-phase3-snp/ (accessed on 6 August 2021).
80. Home | MalariaGEN. Available online: <https://www.malariagen.net/> (accessed on 6 August 2021).
81. Sharakhova, M.V.; Hammond, M.P.; Lobo, N.F.; Krzywinski, J.; Unger, M.F.; Hillenmeyer, M.E.; Bruggner, R.V.; Birney, E.; Collins, F.H. Update of the *Anopheles gambiae* PEST genome assembly. *Genome Biol.* **2007**, *8*, R5. [\[CrossRef\]](#)
82. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. 1000 Genome Project data processing subgroup the sequence alignment/mapformat and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [\[CrossRef\]](#)
83. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [\[CrossRef\]](#)
84. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [\[CrossRef\]](#)
85. Wright, S. The genetical structure of populations. *Ann. Eugen.* **1951**, *15*, 323–354. [\[CrossRef\]](#)
86. Weir, B.S.; Cockerham, C.C. Estimating F-Statistics for the analysis of population structure. *Evolution* **1984**, *38*, 1358–1370. [\[CrossRef\]](#) [\[PubMed\]](#)
87. Jombart, T. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **2008**, *24*, 1403–1405. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Jolliffe, I. Principal Component Analysis (2ed., Springer, 2002) (518s)—MVsa—PDF | PDF. Available online: <https://fr.scribd.com/document/484920903/Jolliffe-I-Principal-Component-Analysis-2ed-Springer-2002-518s-MVsa-pdf> (accessed on 3 February 2025).
89. Goudet, J. Hierfstat, a package for R to compute and test hierarchical F-Statistics. *Mol. Ecol. Notes* **2005**, *5*, 184–186. [\[CrossRef\]](#)
90. Pfeifer, B.; Wittelsbürger, U.; Ramos-Onsins, S.E.; Lercher, M.J. PopGenome: An efficient swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* **2014**, *31*, 1929–1936. [\[CrossRef\]](#)
91. Nei, M.; Li, W.H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 5269–5273. [\[CrossRef\]](#)
92. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **1989**, *123*, 585–595. [\[CrossRef\]](#)
93. Kimura, M. *The Neutral Theory of Molecular Evolution*; Cambridge University Press: Cambridge, UK, 1983; ISBN 978-0-521-31793-1.

94. Slatkin, M. Linkage Disequilibrium—Understanding the evolutionary past and mapping the medical Future. *Nat. Rev. Genet.* **2008**, *9*, 477–485. [[CrossRef](#)]
95. Shin, J.-H.; Blay, S.; McNeney, B.; Graham, J. LDheatmap: An R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J. Stat. Softw.* **2006**, *16*, 1–9. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.