**Restriction to gene flow is linked to changes in the molecular basis of pyrethroid resistance in the malaria vector *Anopheles funestus***

Kayla G. Barnesa,c,d, Helen Irvinga, Martin Chiumiab, Themba Mzilahowab,Michael Colemana, Janet Hemingwaya and Charles S. Wondjia,e

a Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, L3 5QA, United Kingdom

b Malaria Alert Centre, College of Medicine, Blantyre, Malawi.

c FAS Center for Systems Biology, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, United States of America.

d Broad Institute, Cambridge, Massachusetts 02142, United States of America.

e LSTM/OCEAC Research Unit. Organisation de Coordination pour la lutte contre les Endémies en Afrique Centrale, Yaoundé, Cameroon.

To whom reprint requests should be addressed. Email: charles.wondji@lstmed.ac.uk

**Keywords**: Malaria, pyrethroid resistance, *Anopheles funestus*, cytochrome P450, gene flow

***Corresponding authors:***

Janet Hemingway

Vector Biology Department, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L5QA, United Kingdom. Phone: + 44 151 705 3100, Fax: +44 151 705 3370

Email: Janet.Hemingway@lstmed.ac.uk

Charles S. Wondji

Vector Biology Department, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L5QA, United Kingdom. Phone: + 44 151 705 3228, Fax: +44 151 705 3369

Email: charles.wondji@lstmed.ac.uk

**Abstract (249/250)**

Resistance to pyrethroids, the only insecticide class recommended for bednets, threatens the control of major malaria vectors including *Anopheles funestus*. To manage resistance effectively, it is crucial to understand the dynamics and mechanisms driving resistance. Here, using genome-wide transcription and genetic diversity analyses, we show that a shift in the molecular basis of pyrethroid resistance in southern African populations of this species is associated with a restricted gene flow. Across the most highly endemic and densely populated regions in Malawi, *An. funestus* was resistant to pyrethroids, carbamates and organochlorides. Genome-wide Microarray-based transcription analysis revealed that over-expression of cytochrome P450 genes was the main mechanism driving this resistance. The most up-regulated genes included *CYP6P9a*, *CYP6P9b* and *CYP6M7.* However, a significant shift in the over-expression profile of these genes was detected across a south/north transect with CYP6P9a and CYP6P9b more highly over-expressed in the southern resistance front whereas CYP6M7 was predominant in the northern front. A genome-wide genetic structure analysis of southern African populations of *An. funestus* from Zambia, Malawi and Mozambique revealed restriction of gene flow between populations in line with the geographical variation observed in the transcriptomic analysis. Genetic polymorphism analysis of the three key resistance genes CYP6P9a, CYP6P9b and CYP6M7 support barriers to gene flow that are shaping the underlying molecular basis of pyrethroid resistance across southern African. This barrier to gene flow is likely to impact the design and implementation of resistance management strategies in the region.

**Significant Statement (120/120 words)**

Vector control is the cornerstone for malaria prevention. Wide-scale distribution of treated bednets and indoor residual spraying throughout Africa has led to a major decrease in mortality but insecticide resistance threatens this ongoing success. Here we show that across a highly endemic country,

one of the main malaria vectors, *Anopheles funestus*, is resistant to several insecticide classes, including the insecticides used for bednets. This study shows multiple *CYP450s,* genes linked to xenobiotic metabolism, are responsible for insecticide resistance but their implication significantly varies geographically. Genetic diversity analysis revealed that this genomic variation is associated with barriers to gene flow between southern African populations of *An. funestus*. This work will help improve resistance management and vector control strategies throughout southern Africa.

\body

**Introduction**

Vector control including indoor residual spraying (IRS) and long-lasting insecticide treated bed nets (LLINs) are the primary control interventions to prevent malaria transmission. Unfortunately, insecticide resistance is threatening the continued success of these crucial control tools ([1](#_ENREF_1)). In southern African countries such as Malawi, this is especially concerning because *Anopheles funestus*, the primary disease vector is increasingly resistant to insecticides although it remains unclear whether the resistance front is unique or if gene flow is uniform across the region. Malawi has recently experienced a scale up of vector control interventions throughout the country including large-scale distribution of LLINs starting in 2007 and a six-district IRS campaign starting in 2009 along the densely populated Lake Malawi and in the southern agricultural district of Chikwawa. Pyrethroid lambda cyhalothrin (ICON, Syngenta) was first used in Nkhotakota in 2009, followed by pirimiphos methyl (Actellic EC) in Nkhotakota and Salima in 2010 and Nkhotakota in 2011. Simultaneously, insecticide resistance was documented in the southern agricultural area of Chikwawa where the pyrethroid lambda cyhalothrin was being used for IRS as well as the isolated Likoma Island in Lake Malawi ([2](#_ENREF_2), [3](#_ENREF_3)) but the insecticide resistance status of *An. funestus* in the rest of the country was largely unknown. Therefore, an extensive susceptibility profiling of local vector populations spanning the most endemic districts in Malawi is needed to design appropriate control strategies and if the strategies being deployed are sustainable nation-wide and throughout the southern Africa where *An. funestus* dominates.

Previous investigation of pyrethroid resistance in southern Africa has revealed an absence of the knockdown resistance (*kdr*) target site mutation ([3-5](#_ENREF_3)). Instead, pyrethroid resistance in *An. funestus* is mainly driven by metabolic mechanisms, primarily through the over-expression of the tandemly duplicated cytochrome P450s *CYP6P9a* and *CYP6P9b*, both located on the pyrethroid resistance QTL (*rp1)* ([5](#_ENREF_5), [6](#_ENREF_6)). Recently the contribution of *CYP6M7* to pyrethroid resistance in southern Africa was confirmed, with a more predominant role for this gene in Zambia than Mozambique ([4](#_ENREF_4)), raising the possibility that the molecular basis of pyrethroid resistance might vary across southern African populations. If such a shift exists, the underlying factors behind such difference remains uncharacterized. Malawi bordering both Mozambique and Zambia, is ideally located to test the hypothesis that the mechanisms of pyrethroid resistance vary across this region and this difference is linked to the genetic structure of *An. funestus* throughout southern Africa. In addition, it remains possible that the screening of new populations could also detect new genes driving such resistance. The identification of the full set of genes involved in resistance will help to thoroughly decipher the molecular basis of resistance and potentially detect suitable resistance markers to explain the total genetic variance of pyrethroid resistance and help design DNA-based molecular diagnostic tools to easily detect and track resistance in the field.

Here, using genome-wide transcription and genetic diversity analyses, we show that a shift in the molecular basis of pyrethroid resistance in southern African populations of *An. funestus* is associated with significant reduced gene flow levels between southern and northern populations of this region. Previously, it was thought that P450s located in the *rp1* were mainly responsible for pyrethroid resistance, and while they play a key role in insecticide metabolism, genes located on other chromosomes, mainly *CYP6M7* on the 2L chromosome, are also important. Therefore, insecticide resistance diagnostic assays need to incorporate a multigene panel. Additionally this work evidences that population structure and variation in gene expression are correlative and can be used to inform future interventions especially as new insecticides come to market.

**Results**

**The insecticide resistance profile of *An. funestus* across Malawi**: The mortality rates of all three districts chosen to span the highly endemic region of Malawi (Fig.1*A*), showed high resistance to the pyrethroids, permethrin (32-43% mortality rate) and deltamethrin (16-40%), and to the carbamate, bendiocarb (32-58%) with moderate resistance to DDT (76-85%) (Fig.1*B*). In contrast, there was full susceptibility to the organophosphate, malathion (100%) as previously reported for this species in other African locations ([7](#_ENREF_7)). Pre-exposure of mosquitoes to piperonyl butoxide (PBO), a synergist assay that inhibits P450 activity, followed by permethrin and bendiocarb exposure restored full susceptibility to both insecticides (100% mortality) in all three populations suggesting that cytochrome P450s are driving resistance to carbamates and pyrethroids.

 **Genome-wide transcription profiling of pyrethroid resistance**: Mosquitoes resistant to permethrin, unexposed to insecticide, and a susceptible laboratory strain (Fang) were used to analyze the transcriptome profile. The set of genes differentially expressed for each comparison of resistant versus susceptible (R-S) (Fig.1*C*) and unexposed control versus susceptible (C-S) (*SI Appendix*, Fig.S1*D*) were determined for each location and between each location (>2-fold change, P<0.05). The three districts in Malawi share 192 over-transcribed probes (FC≥2, P<0.01) (*SI Appendix*, Fig.S1*A*). There were additional probes over-transcribed in only two collection sites (CKW-NKK = 302 probes, NKK-SLM = 100 probes, and SLM-CKW = 229 probes) while other probes were specific to a single location.

The list of genes commonly over-expressed across Malawi comprises several gene families among which the most preeminent were the cytochrome P450 genes (*SI Appendix*, Table S1). Similar to previous reports in southern Africa the most highly over-expressed of these P450s in all three locations in Malawi belong to the CYP6 family including *CYP6P9a, CYP6P9b* and *CYP6M7 (*[*4*](#_ENREF_4)*,* [*8*](#_ENREF_8)*,* [*9*](#_ENREF_9)*).* Unlike previous reports the over-expression of these candidate genes significantly varied geographically between southern and northern districts. In southern Malawi (Chikwawa) *CYP6P9a (FC=39.4)* and *CYP6P9b (FC=23.96)* have the highest over-expression levels, whereas a lower fold change is observed for *CYP6M7 (FC=12.5)* (Fig.1*C*). However, the opposite pattern is observed for the northern districts where *CYP6M7* has the highest expression in Salima District [*CYP6P9a (*FC *=8.37), CYP6P9b (FC=8.6)* and *CYP6M7 (FC=59.06)]* and Nkhotakota District *[CYP6P9a (FC=13.75), CYP6P9b (FC= 11.33)* and *CYP6M7 (FC= 59.63)]*. This shift in expression pattern was also observed for other cytochrome P450s commonly over-expressed in all 3 locations. This includes *CYP6AA1* located on the pyrethroid resistance QTL (rp1) with a higher over-expression in northern locations (SLM FC = 12.66, NKK FC = 8.25, CKW FC = 5.20) and *CYP6M4* exhibiting higher expression in the south (CKW FC=12.45, SLM FC=3.82, and NKK FC=3.51). Another P450 from the CYP4 family, *CYP4C27,* was highly over-expressed in the two northern locations of Salima (FC=19.50) and Nkhotakota (FC=26.78) but minimally over-expressed in the southern location of Chikwawa (FC=2.11). However, other P450 genes such as *CYP6Z1* had a similar expression in all three locations(SLM FC = 3.01, NKK FC = 2.95, CKW FC = 3.04) (*SI Appendix,* Table S1). Beside cytochrome P450s, other genes belonging to multiple genes families included several proteases and alcohol and glucose dehydrogenases were up-regulated. Although these genes may be playing a role in insecticide resistance they are likely secondary factors as the PBO-synergist assay established the importance of P450s in relationship to resistance. There were additional detoxifying genes that were up-regulated in only one or two locations including some P450s, with a larger set in southern Malawi, although these all maintained a lower fold change (FC range: 6.78-2.11) than the main genes of interest.

**Resistance genes are constitutively over-expressed**: The expression profile obtained by comparing control mosquitoes unexposed to insecticides (C) collected from the same three districts to the FANG susceptible strain (C-S) revealed that the main genes detected in R-S hybridization are also over-expressed in C-S hybridization. This confirms that these genes are constitutively over-expressed in resistant mosquitoes and are not induced by insecticide exposure. The expression profile of C-S is presented in Table S1 (*SI Appendix*). The geographical shift in gene expression of key resistance genes identified using the R-S comparison was also confirmed using the C-S comparison with a higher overexpression of *CYP6M7* (Fold change: SLM=41.11, NKK=24.08, CKW=3.73) found in the northern sites and a higher overexpression of *CYP6P9a* (Fold change: SLM=7.00, NKK=7.17, CKW=69.19) and *CYP6P9b* (Fold change: SLM=8.85, NKK=6.56, CKW=26.56) in the southern site.

**Validation of over transcription with qRT-PCR:** Quantitative real-time PCR (qRT-PCR) broadly confirmed the microarray results for the seven genes tested with a significant correlation observed between qRT-PCR and microarray fold change (R2= 0.519) although variation between the techniques was observed (*SI Appendix*, Fig. S1*B*) (supplementary text). .

**Analysis of genetic diversity across southern Africa**: The significant differences in the gene expression profiles observed between populations of *An. funestus* in Malawi compliment observed changes in expression profiles documented between Zambia and Mozambique ([4](#_ENREF_4)). These variations could suggest the presence of barriers to gene flow that are impacting the spread of resistance genes. If barriers to gene flow exist across Malawi and throughout southern Africa this may significantly impact both the speed and spread of insecticide resistance mediating genes. Unfortunately, knowledge on the population structure of *An. funestus* in southern Africa is limited. The population structure of Southern Africa was analyzed using microsatellites for multiple sites in this region, Zambia (N=3), Malawi (N=3), and Mozambique (N=1). Field caught female mosquitoes (N=36-48) from the seven collections were typed using 17 microsatellite markers spanning the genome.

The 17 microsatellite markers spanning the entire genome were genotyped consistently in all populations showing robustness of the multiplex design. Each population analysed contain ≥36 individuals (≥72 alleles) which is significant for accurately estimating allele frequencies ([10](#_ENREF_10)). The loci FUNR, FUNQ, and AFND12 showed an excess of one allele. FUNR is located in the resistance to permethrin QTL (*rp1)* and may be under selection. The loci FUNQ and AFUB12 are both located on the X chromosome where a loss of diversity was also identified in *An. funestus* in the African-wide analysis ([11](#_ENREF_11)). The number of total alleles typed for each microsatellite marker ranged from 7 to 4.

Individual microsatellite markers were analyzed for Hardy-Weinberg Equilibrium (HWE) with Bonferroni correction for multiple tests. A total of 21 out of 152 (13.8%) markers deviated from HWE due to deficit of heterozygosity (*SI Appendix*, Table S3). To ensure deviation form HWE equilibrium did not affect predicted population structure all analysis was performed on the full panel of 17 microsatellites, 14 microsatellites (minus homogenous markers) and the 11 microsatellites (no HWE deviation in any populations) (*SI Appendix*, Table S4). There was no linkage disequilibrium (LD) observed. For all populations there was a total 1,224 pairwise comparisons, of which, 31 comparisons were in significant LD (p < 0.05) but this constitutes less than 5% of alleles.

Patterns of genetic differentiation based on pairwise *Fst* estimates between all locations revealed the presence of a north-south geographical subdivision. Three Zambia collections plus Salima and Nkhotakota from Malawi show very low and non-significant *Fst* pairwise estimates suggesting a high level of gene flow between these populations (*Fst* range 0.0001-0.04). However, significant *Fst* estimates were observed between the southern Malawi populations of Chikwawa versus populations from northern Malawi and Zambia (Fst range 0.0597-0.0704) and Mozambique (*Fst* range 0.0676-0.0812) suggesting the existence of barriers to gene flow (*SI Appendix*, Table S5). This genetic differentiation pattern was maintained after removing the 6 markers deviating from HWE. The segregation in the Mozambique population (1500km away from Malawi) may be explained by distance but segregation by distance breaks down when comparing western Zambia to central Malawi (*SI Appendix,* Fig. S1*C*). Western Zambia is almost 1200km from central Malawi yet presents as an interbreeding population while southern Malawi is only 350km from central Malawi and shows significant genetic deviation further suggesting the presence of barriers to gene flow across Malawi.

The population structure in southern Africa was also supported by a neighbour-joining tree with genetic distances based on *Fst* estimates (Fig. 2A) showing a cluster of Zambian and central and northern Malawian populations while Chikwawa and southern Mozambique were divergent. To confirm the estimated genetic difference inferred based on *Fst*, Bayesian prediction of population structure (without pre-defined populations based on geographically collected samples) was determined using the full panel of 17 microsatellites. The ‘true’ population size was K=3, predicted using the Evanno method, and the graphical structure was produced for K=3-7 to visualize the multimodality within and between populations. Complementary to the *Fst* predictions, Zambia samples and Salima and Nkhotakota samples from Malawi share a similar (Fig. 2*B*).

**Genetic diversity patterns of *CYP6P9a, CYP6P9b* and *CYP6M7* across Malawi:** Polymorphism patterns of the three main resistance genes, *CYP6P9a, CYP6P9b* and *CYP6M7* were investigated to assess whether there was an association with the shift in gene expression profiles across the South/North transect in Malawi. The assumption being that if there is indeed a potential decreased role of *CYP6P9a* and *CYP6P9b* (both located on the *rp1* QTL) in the north, this would be reflected by a weaker signature of selection compared to southern Malawi. Similarly, analysis on *CYP6M7*, located on the second pyrethroid QTL *rp2* on the 2Lchromosome, could further inform the selection pattern associated with pyrethroids in northern Malawi. The three genes were sequenced from resistant and susceptible mosquitoes from each district and changes in their polymorphism pattern associated with pyrethroid resistance analyzed.

 **Polymorphism analysis of *CYP6P9a* and *CYP6P9b***: A genomic fragment of 2,056bp and 1,812bp of *CYP6P9a* and *CYP6P9b* respectively were sequenced for all three locations in Malawi and analyzed with previously published sequences from Zambia and Mozambique to confirm if the polymorphic patterns observed is complimentary to the trend throughout southern Africa ([4](#_ENREF_4)). The sequences include the full coding region two exons and one intron (*CYP6P9a* = 1584bp; *CYP6P9b* = 1524bp) and the 5’UTR (*CYP6P9a* = 472bp; *CYP6P9b* = 288bp). In the southern district of Malawi (CKW), both *CYP6P9a* and *CYP6P9b* had a significant reduction of genetic diversity in resistant mosquitoes (π: *CYP6P9a* = 0.00036, *CYP6P9b* = 0.0001) versus susceptible mosquitoes (π: 0.0018, *CYP6P9b* = 0.001) (Fig. 3), low substitution numbers in resistant (*CYP6P9a* = 2, *CYP6P9b* = 1) vs. susceptible (*CYP6P9a* = 14, *CYP6P9b* = 12) (*SI Appendix*, Table S6), and a loss in haplotype diversity (*SI Appendix,* Fig. S2*A*). In the northern Malawi district of Salima there was also a reduction in genetic diversity (π resistant: *CYP6P9a* = 0.0001, *CYP6P9b* = 0.0003 vs susceptible: *CYP6P9a* = 0.0033, *CYP6P9b* = 0.0093) and low substitutions in resistant (*CYP6P9a* = 2, *CYP6P9b* = 2) versus susceptible (*CYP6P9a* = 15, *CYP6P9b* = 52). However, in Nkhotakota, the most northern district, both susceptible and resistant samples presented a reduced genetic diversity for *CYP6P9a (*π resistant = 0.00019 vs susceptible = 0.00027) with only 2 substitutions in each phenotype and a low haplotype of 3 while *CYP6P9b* showed variation between resistant (π = 0.0003, S = 3) and suceptible (π = 0.0041, S = 26) mosquitoes (*SI Appendix*, Table S6).

The loss of haplotype diversity, mainly in resistant mosquitoes, is consistent throughout southern Africa. Resistant samples from Malawi, Zambia, and Mozambique reveal a significant reduction in haplotype diversity for *CYP6P9a* with only 10 mutational positions observed in resistant mosquitoes across the region accounting for 10 haplotypes, whereas in susceptible mosquitoes, up to 103 mutational positions are observed generating 25 haplotypes (*SI Appendix,* Fig. S3, S4). *CYP6P9b* also shows a similar loss in diversity in resistant mosquitoes with 9 mutations leading to only 10 haplotypes in resistant mosquitoes but up to 100 mutational positions generating 20 haplotypes in susceptible mosquitoes.

The significant differences in the polymorphism patterns of resistant and susceptible mosquitoes was also observed at the protein level with key amino acid changes detected and likely to impact the metabolic efficacy of both enzymes towards pyrethroids as recently demonstrated ([12](#_ENREF_12)) (SI Text).

Although the pattern of reduced diversity in resistant mosquitoes was consistent throughout southern Africa there was evidence of geographical variation in these genes. The predominant haplotype (H1) in *CYP6P9a* and *CYP6P9b* was consistently higher in southern Malawi (CKW: 90%, 95%) and Mozambique (90%, 90%) compared to cluster 1; SLM (70%, 55% respectively for each gene) NKK (75%, 60%), and west Zambia (70%, 40%) (Fig. 4). The TCS network reveals that *CYP6P9a and CYP6P9b* haplotypes from Chikwawa are only separated by a few mutational steps (<5) while haplotypes from SLM and NKK have more mutational steps (>12-20). For *CYP6P9b* there is an obvious geographic shift where the singleton haplotypes are all from the northern locations (13/40 haplotypes). Most of the CYP6P9b susceptible haplotypes with a few resistant alleles divergent from the H1 or H2R haplotypes. In the north, there are several other resistant haplotypes separated by 1 mutational step from the predominant H1 haplotypes.

The higher diversity of *CYP6P9a* in northern Malawian is confirmed by a Maximum Likelihood (ML) tree where several haplotypes from SLM and NKK form an independent cluster with higher genetic distance from the predominant haplotype (*SI Appendix,* Fig. S4*A*). There is one major clade for both *CYP6P9a* and *CYP6P9b* made mainly of resistant sequences and small divergent clades representing susceptible sequences (*SI Appendix,* Fig. S4*A, B*). The ML tree shows that sequences from southern Mozambique cluster closer to those from southern Malawi whereas sequences from Zambia cluster with northern Malawi. This pattern of diversity further supports the presence of a shift in gene expression of these two important pyrethroid resistance genes across southern Africa. A sliding window analysis of the nucleotide diversity across the full-length of *CYP6P9a* and *CYP6P9b* showed that the loss of diversity was not restricted to the 5’UTR or introns as there is a significant loss of diversity throughout the coding region of resistant mosquitoes compared to susceptible mosquitoes in all three locations (*SI Appendix,* Fig. S4*D,E*).

 **Polymorphism analysis of *CYP6M7***: Sequencing of a 1.2kb genomic fragment spanning the 5’UTR (610bp) and first exon (590bp) between five resistant (alive after 1 hour) and five susceptible (dead after 1 hour) individual mosquitoes per site revealed a high genetic diversity in *CYP6M7* with no correlation to either the resistance phenotype or geographical location (*SI Appendix*, Table S7). There was high haplotype diversity (0.9-1) (*SI Appendix,* Fig. S2*A*), nucleotide diversity (π = 0.01-0.02), and no dominant haplotype in Malawi or throughout southern Africa, in contrast to *CYP6P9a* and *CYP6P9b*. The TCS network reveals there is no phenotypic grouping in Malawi and 90% of samples have a unique haplotype (Fig. 4). Analysis of individual samples using a ML tree reveals resistant and susceptible mosquitos are equally diverse and divergent and no phenotypic clustering occurs (*SI Appendix,* Fig. S4). Analysis of Tajima D (combined = -0.64, resistant = -0.07, susceptible = -0.41) and Fu and Li’s D\* (combined = 0.78, resistant = 0.43, susceptible = 1.27) estimates did not detect any evidence that *CYP6M7* is under selection (*SI Appendix*, Table S7). Upon closer analysis of the 5’ UTR and coding region of *CYP6M7* there was little difference in Hd estimates either geographically or phenotypically, between resistant and susceptible mosquitoes.

**Discussion**

Insecticide resistance threatens malaria control interventions in Africa, notably a number of highly endemic countries in southern Africa. Coherent vector management plans are required to access bilateral and international support for malaria control. The design and implementation of such plans requires knowledge of the resistance profile of major malaria vectors to the available insecticides, as well as, the elucidation of underlying resistance mechanisms to develop better insecticide resistance diagnostics and future insecticides. To help achieve such goals, this study has established the resistance profile of *An. funestus* on a South/Central transect in Malawi, which encompasses the most densely populated areas of Malawi and the highest incidence of malaria, and detected a geographical shift in the resistance mechanisms corresponding to patterns of population structure throughout southern Africa.

**Widespread distribution of insecticide resistance is a concern for malaria control:** *Anopheles funestus* populations across Malawi are resistant to pyrethroids, which is a concern as pyrethroids are currently the only insecticide class recommended by WHO for use on bed nets. This concern is increased by the presence of resistance to the carbamate bendiocarb, the main alternative to pyrethroids for IRS. The resistance profile described in this study confirms resistance in isolated districts in Malawi ([3](#_ENREF_3), [13](#_ENREF_13)) and is similar to that in neighboring countries, including Mozambique ([14](#_ENREF_14), [15](#_ENREF_15)) and Zambia ([4](#_ENREF_4), [9](#_ENREF_9)), suggesting that pyrethroid and carbamate resistance is widespread in southern Africa but the genetic mechanisms responsible for this resistance are not uniform. The apparent shift in gene expression in *An. funestus* populations observed between southern and northern locations, especially in Malawi, where geographical distance is not a factor, suggests that widespread resistance is not the result of a single mutational event and subsequent gene flow, but likely multiple independent events.

**Multiple cytochrome P450s are driving pyrethroid resistance in Malawi but with geographical variation:** Despite similar resistance levels throughout Malawi, a clear South/North geographical difference was observed in the transcription profile of the main pyrethroid resistance genes, *CYP6P9a* and *CYP6P9b* both under directional selection but more significantly in southern Malawi. Such geographical difference in the role of key metabolic resistance genes have also been observed in the *An. gambiae* genes *CYP6M2* and *CYP6P3* ([16](#_ENREF_16), [17](#_ENREF_17)) and several CYP9 genes in *Aedes aegypti* ([18](#_ENREF_18)).

In contrast to *CYP6P9a* and *CYP6P9b*, no evidence of selection was detected for *CYP6M7*, and it is likely that the molecular basis through which *CYP6M7* confers pyrethroid resistance is different to that of *CYP6P9a* and *CYP6P9b* with little contribution from polymorphisms from coding regions. The over transcription of *CYP6M7* may be caused by changes in the gene’s promoter region or mutations in the transcription factors. Further work is needed to fully elucidate the mechanism driving over-expression in *CYP6M7*. The high polymorphic nature of *CYP6M7* shows that no selective pressure is acting on this gene although it is highly over-expressed and linked to pyrethroid resistance throughout the region ([4](#_ENREF_4), [9](#_ENREF_9)). The higher expression of *CYP6M7* in northern Malawi and Zambia suggests that there are likely two resistance fronts occurring in southern Africa.

The existence of geographical difference in the role of key resistance genes in Malawi suggests the presence of barriers to gene flow or variation in selection forces between populations from south to north. Previous work predicted southern Africa was one interbreeding population based on continent wide analysis ([19](#_ENREF_19)) but genome-wide genetic diversity analysis using *Fst*comparisons and Bayesian structure analysis revealed 3 population clusters in southern Africa. Cluster 1, which includes Zambia and northern Malawi (Salima and Nkhota-kota) corresponds to higher over expression of *CYP6M7*. Southern Malawi forms a unique cluster (cluster 2) and this is likely due to its location at the end of the rift valley which has previously been shown to constitute a barrier to gene flow in the other major malaria vector *An. gambiae* (Lehman et al 1999). The difference in the gene expression profile between Southern and Northern Malawi further supports the presence of barriers to gene flow. Mozambique (cluster 3) is over 1,500km from southern Malawi and the level of genetic differentiation observed with Southern Malawi could be a consequence of an isolation-by-distance as suggested by the analysis (Fig. 2C). A similar expression profile between southern Malawi and Mozambique further support the role of isolation-by-distance (Fig 2C). Future studies with a larger sampling across the southern Africa region will help to further define pattern of gene flow in *An. funestus* or other malaria vectors. Furthermore, the number of genetic markers could also be increased and new next-generation sequencing tools could be used such as ddRADseq or whole genome sequencing across the regions to further elucidate the genetic structure of this malaria vector in southern Africa.

Difference in the molecular basis of resistance within a given country means that national resistance management strategies without characterization of underlying resistance mechanisms from multiple localities may be flawed. The similarity of resistance profile in Malawian *An. funestus* suggests that the same resistance management strategy could be implemented nationally although attention should be paid to the independent evolution of resistance mechanisms across the region. The full susceptibility observed against organophosphate suggests that an IRS campaign using an insecticide from this class, such as pirimiphos methyl could be a viable alternative, although the higher cost of this insecticide and its lower residual efficacy could be a limiting factor. This was already the case in when pirimiphos methyl was sprayed for two years in Nkhotakota and 1 year in Salima but the insecticide showed low efficiency after 2-months. This may explain some of the differences in gene expression profile in central Malawi versus southern Malawi but does not explain the higher overexpression of *CYP6M7* observed in Zambia, a country that has used mainly pyrethroid and carbamate insecticides for IRS. Overall, the predominant role of cytochrome P450s in pyrethroid resistance in Malawian suggests that rolling out the new generation of bed nets that combine pyrethroids and a P450 inhibitors ([20](#_ENREF_20)), could improve LLIN effectiveness across this region. Most importantly, this study highlights pyrethroids should be eliminated from IRS use to prevent resistance becoming fixed.

Broadly, this study highlights that cytochrome P450s play a major role in pyrethroid resistance but the underlying molecular basis of resistance in one area does not necessarily extent to other regions and this should be taken into account in designing resistance management strategies. The role of P450s in resistance is not limited to *An. funestus*. Cytochrome P450s have been linked to insecticide resistance in other major malaria vector including *An. gambiae (*[*21*](#_ENREF_21)*)* and *An. minimus (*[*22*](#_ENREF_22)*)*, as well as, the major Dengue virus and Zika virus vector *Aedes aegypti (*[*18*](#_ENREF_18)*,* [*23*](#_ENREF_23)*)*. Additionally many agricultural pests including aphids ([24](#_ENREF_24)), whiteflies ([25](#_ENREF_25)), and plant hoppers ([26](#_ENREF_26)) utilize the expression of multiple P450s to metabolize pesticides. Understanding the population structure of a species can help inform potential genetic differences in resistance and the breath of sampling needed to determine mechanisms driving this resistance. As existing insecticides and pesticides are scaled up and as new insecticides and pesticides reach the public market, molecular studies of multiple collections are needed to fully elucidate the mechanisms of resistance to ensure the longevity of both vector and pest control.

**Methods**

**Mosquito collection and insecticide bioassays:** Gravid indoor resting female mosquito collections were carried out in three highly endemic districts: Chikwawa, Salima and Nkhotakota (Fig. 1A). Three to five collections for each site were conducted over two rainy seasons, November 2010-February 2011 and November 2011-February 2012. Two seasonal collections were needed to generate adequate numbers to test the full spectrum of insecticides. Eggs were collected and female F0 mosquitoes used for oviposition were species identified using the funestus complex cocktail PCR ([27](#_ENREF_27)). Insecticide resistance was assessed using 2-5 day-old F1 adult mosquitoes. Standardized WHO insecticide impregnated papers were used for all experiments including the 4% PBO. One hour bioassays for the insecticides: 0.1% bendiocarb (carbamate), 4% DDT (organochlorine), 0.05% deltamethrin (pyrethroid), 0.75% permethrin (pyrethroid), and 5% malathion (organophosphate) were carried out with methods detailed in the WHO manual ([28](#_ENREF_28)) using 185-200 adult females and 100-150 adult males plus parallel control tubes. Resistant mosquitoes (alive after 24h) were stored at -80oC and dead on silica gel. Mosquitoes collections from Zambia occurred in 2009 and 2010 ([29](#_ENREF_29)) and from Mozambique in 2009 ([15](#_ENREF_15)) (*SI Appendix,* Fig S5).

**Microarray and qrt-PCR:** 30 Custom 8x60k Agilent microarrays (Agilent, Santa Clara, CA, USA) (A-MEXP-2374) ([4](#_ENREF_4)) were used to analyze the genome-wide transcription profile for the three locations. RNA was extracted from three pools of 10 female mosquitoes per location per phenotype: C=unexposed to insecticide, R=resistant after exposure to 0.75% permethrin and S=susceptible colony Fang ([30](#_ENREF_30)), using the Picopure RNA Isolation Kits (Applied Biosystems, Paisley, PA, USA). Transcription profiles where determined using Genespring GX 12.0 software. cDNA was synthesized from four biological replicates for R, C and S mosquitoes per districts and carried out using methods previously published ([16](#_ENREF_16)). Fold change was calculated using three technical replicates and then by normalizing with two housekeeping genes; Actin and SP7 using the 2-ΔΔCT method ([31](#_ENREF_31)).

**Genetic Population Structure of *An. funestus* in southern Africa**: 17 microsatellites spanning the *genome were* optimized and genotyped into three plexes (*SI Appendix,* Table S2) for Zambia (N=3), Malawi (N=3), and Mozambique (N=1). 48 gDNA samples from each population (N=432) where run using the Type-it Microsatellite PCR kit (Qiagen) and scored. HWE, LD, F-statistics, null-allele and Fst were determinedusing GENEPOP 4.0.10 ([32](#_ENREF_32)). Bayesian analysis of population structure was generated using the program STRUCTUREv2 ([33](#_ENREF_33)). The genetic distance (determined by pairwise Fst scores) was compared to geographical distance using Isolation by Distance ([34](#_ENREF_34)).

**Genetic variability of candidate genes:** For the resistance genes of interest, *CYP6P9a, CYP6P9b* and *CYP6M7,* the full-length coding region as well as the 5’ UTR were amplifiedin all three districts in Malawi. gDNA was extracted on five resistant and five susceptible female F1 mosquitoes using the Livak DNA extraction method or DNAeasy kit (Qiagen). Amplification followed previously published methods (6). Aligned sequences were haplotype phased, the coding regions were defined using Blastx, and genetic parameters were determined using dnaSP 5.1 ([35](#_ENREF_35)).

Full analysis is detailed in *SI Appendix, Material and Methods*. All DNA sequences in this study have been deposited in the GenBank database (KP984806-KP984983) and all microarray data was submitted to Array Express (E-MTAB-3342).

**Acknowledgments**

This work was supported by a Wellcome Trust Research Career Development Fellowship (083515/Z/07/Z) and Senior Research Fellowship to CSW (101893/Z/13/Z) and by the National Institute of Health (NIH) grant IU01 A1058271-01 to JH.

**Author contributions**

CSW conceived the study. KGB, MCh and TM carried out sample collections; KGB performed laboratory work; KGB and CSW performed the sequencing of resistance genes; MC and JH contributed toward the conception and significant insights; KGB and CSW analyzed data and wrote the manuscript. All authors read and approved the manuscript.

**Accession Numbers**

Microarray data from this study were submitted to Array Express: E-MTAB-3342. The DNA sequences reported in this paper have been deposited in the GenBank: KP984806-KP984983.

**Supplementary data**

Supplementary Data are available online.

**Reference**

1. WHO (2012) *Global Plan for Insecticide Resistance Management (GPIRM)* (World Health Organization, Geneva, Switzerland).

2. Hunt R, Edwardes M, & Coetzee M (2010) Pyrethroid resistance in southern African Anopheles funestus extends to Likoma Island in Lake Malawi. *Parasit Vectors* 3:122.

3. Wondji CS*, et al.* (2012) Impact of pyrethroid resistance on operational malaria control in Malawi. *Proc Natl Acad Sci U S A* 109(47):19063-19070.

4. Riveron JM*, et al.* (2014) The highly polymorphic CYP6M7 cytochrome P450 gene partners with the directionally selected CYP6P9a and CYP6P9b genes to expand the pyrethroid resistance front in the malaria vector Anopheles funestus in Africa. *BMC Genomics* 15(1):817.

5. Riveron JM*, et al.* (2013) Directionally selected cytochrome P450 alleles are driving the spread of pyrethroid resistance in the major malaria vector Anopheles funestus. *Proc Natl Acad Sci U S A* 110(1):252-257.

6. Wondji CS*, et al.* (2009) Two duplicated P450 genes are associated with pyrethroid resistance in Anopheles funestus, a major malaria vector. *Genome Res* 19(3):452-459.

7. Coetzee M & Koekemoer LL (2013) Molecular systematics and insecticide resistance in the major African malaria vector Anopheles funestus. *Annu Rev Entomol* 58:393-412.

8. Riveron JM*, et al.* (2013) Directionally selected cytochrome P450 alleles are driving the spread of pyrethroid resistance in the major malaria vector Anopheles funestus. *P Natl Acad Sci USA* 110(1):252-257.

9. Thomsen EK*, et al.* (2014) Underpinning sustainable vector control through informed insecticide resistance management. *PLoS One* 9(6):e99822.

10. Hale ML, Burg TM, & Steeves TE (2012) Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. *Plos One* 7(9):e45170.

11. Temu EA, Hunt RH, & Coetzee M (2004) Microsatellite DNA polymorphism and heterozygosity in the malaria vector mosquito Anopheles funestus (Diptera: Culicidae) in east and southern Africa. *Acta Trop.* 90(1):39-49.

12. Ibrahim SS*, et al.* (2015) Allelic Variation of Cytochrome P450s Drives Resistance to Bednet Insecticides in a Major Malaria Vector. *PLoS Genet* 11(10):e1005618.

13. Vezenegho SB*, et al.* (2013) Characterization of the Anopheles funestus group, including Anopheles funestus-like, from Northern Malawi. *Trans R Soc Trop Med Hyg* 107(12):753-762.

14. Casimiro S, Coleman M, Mohloai P, Hemingway J, & Sharp B (2006) Insecticide resistance in Anopheles funestus (Diptera: Culicidae) from Mozambique. *J Med Entomol* 43(2):267-275.

15. Cuamba N, Morgan JC, Irving H, Steven A, & Wondji CS (2010) High level of pyrethroid resistance in an Anopheles funestus population of the Chokwe District in Mozambique. *PLoS One* 5(6):e11010.

16. Kwiatkowska RM*, et al.* (2013) Dissecting the mechanisms responsible for the multiple insecticide resistance phenotype in Anopheles gambiae s.s., M form, from Vallee du Kou, Burkina Faso. *Gene* 519(1):98-106.

17. Tene Fossog B*, et al.* (2013) Physiological correlates of ecological divergence along an urbanization gradient: differential tolerance to ammonia among molecular forms of the malaria mosquito Anopheles gambiae. *BMC ecology* 13:1.

18. Strode C*, et al.* (2008) Genomic analysis of detoxification genes in the mosquito Aedes aegypti. *Insect Biochem Mol Biol* 38(1):113-123.

19. Michel AP*, et al.* (2005) Rangewide population genetic structure of the African malaria vector Anopheles funestus. *Mol Ecol* 14(14):4235-4248.

20. Strode C, Donegan S, Garner P, Enayati AA, & Hemingway J (2014) The impact of pyrethroid resistance on the efficacy of insecticide-treated bed nets against African anopheline mosquitoes: systematic review and meta-analysis. *Plos Med* 11(3):e1001619.

21. Ranson H*, et al.* (2002) Molecular analysis of multiple cytochrome P450 genes from the malaria vector, Anopheles gambiae. *Insect Mol Biol* 11(5):409-418.

22. Rodpradit P, Boonsuepsakul S, Chareonviriyaphap T, Bangs MJ, & Rongnoparut P (2005) Cytochrome P450 genes: molecular cloning and overexpression in a pyrethroid-resistant strain of Anopheles minimus mosquito. *J Am Mosq Control Assoc* 21(1):71-79.

23. Ishak IH*, et al.* (2016) The Cytochrome P450 gene CYP6P12 confers pyrethroid resistance in kdr-free Malaysian populations of the dengue vector Aedes albopictus. *Sci Rep* 6:24707.

24. Puinean AM*, et al.* (2010) Amplification of a cytochrome P450 gene is associated with resistance to neonicotinoid insecticides in the aphid Myzus persicae. *PLoS Genet* 6(6):e1000999.

25. Karatolos N*, et al.* (2012) Over-expression of a cytochrome P450 is associated with resistance to pyriproxyfen in the greenhouse whitefly Trialeurodes vaporariorum. *PLoS One* 7(2):e31077.

26. Bass C*, et al.* (2011) Overexpression of a cytochrome P450 monooxygenase, CYP6ER1, is associated with resistance to imidacloprid in the brown planthopper, Nilaparvata lugens. *Insect Mol Biol* 20(6):763-773.

27. Koekemoer LL, Kamau L, Hunt RH, & Coetzee M (2002) A cocktail polymerase chain reaction assay to identify members of the Anopheles funestus (Diptera: Culicidae) group. *Am J Trop Med Hyg* 66(6):804-811.

28. WHO (2013) Test procedures for insecticide resistance monitoring in malaria vector mosquitoes. *WHO Press*

29. Chanda E*, et al.* (2011) Insecticide resistance and the future of malaria control in zambia. *PLoS One* 6(9):e24336.

30. Wondji CS*, et al.* (2005) An integrated genetic and physical map for the malaria vector Anopheles funestus. *Genetics* 171(4):1779-1787.

31. Schmittgen TD & Livak KJ (2008) Analyzing real-time PCR data by the comparative C-T method. *Nature Protocols* 3(6):1101-1108.

32. Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour* 8(1):103-106.

33. Pritchard JK, Stephens M, & Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945-959.

34. Jensen JL, Bohonak AJ, & Kelley ST (2005) Isolation by distance, web service. *BMC Genet* 6:13.

35. Rozas J (2009) DNA sequence polymorphism analysis using DnaSP. *Methods Mol Biol* 537:337-350.

**Figure Legends**

**Figure 1**: **Insecticide resistance and gene expression profiling: A)** Map - Malaria incidence level in Malawi collection sites: Chikwawa (CKW), Salima (SLM), and Nkhotakota (NKK) (Adapted from Gething et al. 2011) **B)** Susceptibility profile of Malawian *An. funestus* population to the main insecticides and a synergist (PBO). Mean of 185-200 mosquitoes (100 for Malathion) with standard error (SEM). **C)** Fold Change using microarray and qRT-PCR. For qRT-PCR error bars represent standard deviation (SD).

**Figure 2: Population structure of southern Africa: A)** Neighbor joining tree based on the *Fst* pairwise scores using 17 microsatellites. **B)** Bayesian-based structure patterns of southern Africa. K=3,5,7 shows separation of CKW, Malawi and Chokwe, Mozambique.

**Figure 3: Haplotype distribution analysis of *CYP6P9a* and *CYP6P9b***: **A)** Nucleotide diversity based on Pi x 10-3 for *CYP6P9a* and *CYP6P9b*. Significance between resistant (red) and susceptible (blue) including SD and significance P<0.05\*, P<0.01\*\*.

**Figure 4: Comparative analysis of haplotype diversity across Malawi:** TCS network was constructed for the coding region for all samples from Malawi (light blue = CKW, green = NKK, blue = SLM) for **A)** *CYP6P9a* ***B)*** *CYP6P9b and* ***C****) CYP6M7*. Haplotypes with multiple hits are denoted by an H and resistance haplotypes are orange squares. Unique haplotypes are shown as circles with sample information including a 3-digit location code and susceptible (S) or resistant (R).