LIVERPOOL JOHN MOORES UNIVERSITY

LJMU Research Online

Weedall, GD, Mugenzi, LMJ, Menze, BD, Tchouakui, M, Ibrahim, SS, Amvongo-Adjia, N, Irving, H, Wondji, MJ, Tchoupo, M, Djouaka, R, Riveron, JM and Wondji, CS

A cytochrome P450 allele confers pyrethroid resistance on a major African malaria vector, reducing insecticide-treated bednet efficacy

http://researchonline.ljmu.ac.uk/id/eprint/12552/

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Weedall, GD, Mugenzi, LMJ, Menze, BD, Tchouakui, M, Ibrahim, SS, Amvongo-Adjia, N, Irving, H, Wondji, MJ, Tchoupo, M, Djouaka, R, Riveron, JM and Wondji, CS (2019) A cytochrome P450 allele confers pyrethroid resistance on a major African malaria vector, reducing insecticide-treated

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

http://researchonline.ljmu.ac.uk/

http://researchonline.ljmu.ac.uk/

Title: A cytochrome P450 allele confers pyrethroid resistance in a major African malaria vector reducing insecticide-treated nets' efficacy

Authors: Gareth D. Weedall[1], Leon M.J. Mugenzi[2,3], Benjamin D. Menze[1,2,3], Magellan Tchouakui[2,3], Sulaiman S. Ibrahim[1,4], Nathalie Amvongo-Adjia [3,5], Helen Irving[1], Murielle J. Wondji[1,2,3], Micareme Tchoupo[2,3], Rousseau Djouaka[6], Jacob M. Riveron[1,2,3], Charles S. Wondji[1,2,3*]

.

Affiliations:

[1]Vector Biology Department, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, United Kingdom.

[2]LSTM Research Unit at the Organisation de Coordination pour la lutte contre les Endémies en Afrique Centrale (OCEAC), PO Box 288, Yaoundé, Cameroon

[3]Centre for Research in Infectious Diseases (CRID), P.O. Box 13501, Yaoundé, Cameroon

[4]Department of Biochemistry, Bayero University, PMB 3011, Kano, Nigeria

[5]Centre for Medical Research, Institute of Medical Research and Medicinal Plants Studies (IMPM), P.O. Box 13033, Yaoundé, Cameroon

[6]International Institute of Tropical Agriculture (IITA), Cotonou, 08 BP 0932, Benin

*Correspondence to: Email: **charles.wondji@lstmed.ac.uk**

Short title: A P450 allele is reducing bednet efficacy

Abstract

Metabolic resistance to insecticides in mosquitoes threatens malaria control. Unless it is managed, recent gains in reducing malaria transmission could be lost. To improve monitoring and assess the impact of this resistance on control interventions, we elucidated the molecular basis of pyrethroid resistance in the major African vector, *Anopheles funestus* in southern Africa, showing that a single cytochrome P450 allele (*CYP6P9*a) is reducing bednet efficacy. Key resistance genes are detected Africa-wide, but vary geographically. Signatures of selection and adaptive evolutionary traits including structural polymorphisms and *cis*-regulatory transcription factor binding sites were detected with evidence of selection by insecticide-treated nets scale-up. A *cis*-regulatory polymorphism driving the over-expression of the major resistance gene *CYP6P9a* allowed us to design the first DNA-based diagnostic assay for P450-mediated pyrethroid resistance. Using this tool to detect and track the spread of resistance revealed that it is almost fixed in southern Africa but absent elsewhere. Furthermore, experimental hut studies demonstrated that *CYP6P9a-R*esistant mosquitoes survived and succeeded in blood-feeding significantly more than susceptible individuals, highlighting the urgent need to introduce new generations of insecticide-treated nets that are not reliant on pyrethroids.

One Sentence Summary: First DNA marker for P450-mediated pyrethroid resistance in malaria vectors enables to show that metabolic resistance is reducing bednet's efficacy.

Introduction

Malaria prevention relies heavily on the use of insecticide-based vector control interventions, notably pyrethroid-based Long Lasting Insecticidal Nets (LLINs). These tools have been credited with more than 70% of the decrease in malaria mortality in the past 15 years, having helped avert more than 663 million clinical cases of malaria (*1*). However, resistance to insecticides, particularly pyrethroids, in malaria vectors threatens their continued effectiveness. Unless it is managed, the recent gains in reducing malaria transmission could be lost (*2*). Elucidating the genetic basis and evolution of resistance is crucial to design resistance management strategies and prevent malaria resurgence (*2*).

Without genetic information on insecticide resistance genes and associated molecular markers, it is difficult to track and anticipate the course of resistance or assess its impact on malaria transmission and on the effectiveness of control tools such as LLINs. The current inability to track metabolic resistance in this way in all major African malaria vectors including *Anopheles gambiae* and *Anopheles funestus* is a major obstacle to the design of rational, evidence-based resistance management strategies. Of the four classes of insecticides used in public health, pyrethroids are by far the most widely used and the main class recommended for use in insecticide-treated nets. Therefore, understanding the mechanisms conferring pyrethroid resistance in mosquitoes is of critical importance.

Two major causes of insecticide resistance are metabolic resistance and target-site insensitivity (*3*). Unlike target-site resistance (e.g. knockdown resistance: *kdr*), metabolic resistance remains less characterized despite posing a greater risk to control interventions (*4*). Although candidate resistance genes have been detected (*5-8*), it has proved difficult to dissect the molecular bases of metabolic resistance and detect associated molecular markers, because of the size of detoxification gene families, redundancy among their members and the multiple mechanisms through which metabolic resistance can arise (*9*). Cytochrome P450

monooxygenases have consistently been associated with pyrethroid resistance but, unlike *kdr*, no DNA-based marker has yet been detected to track P450-mediated resistance and assess its impact on malaria control tools.

Here, we elucidated the complex molecular basis and genomic evolution of metabolic resistance to pyrethroids in the major African malaria vector *An. funestus*. We detected key DNA-based markers to design a field-applicable diagnostic assay to track this resistance and used it to demonstrat that this metabolic resistance reduces the efficacy of insecticide-treated nets.

Results

RNAseq transcriptional profiling of mosquitoes from across Africa identifies candidate pyrethroid-resistance associated genes

To identify genes associated with pyrethroid resistance in *Anopheles funestus* Africa-wide, we performed RNAseq-based transcriptional profiling of mosquitoes from four different African regions: southern (Malawi), East (Uganda), West (Ghana) and Central (Cameroon), in comparison to a laboratory colony (FANG) that is fully susceptible to all insecticides. Pronounced differences in the expression of key candidate genes were observed between the four regions (Fig. 1A-C, Fig. S1A-E, Table S1).

Cytochrome P450s were frequently significantly over-expressed (adjusted p value <0.05)(Text S1), with *CYP6P9a* (AFUN015792; 60.5x) and *CYP6P9b* (AFUN015889; 23.9x) showing extreme over-expression in Malawi compared to other regions (<7x) (Fig 1B-C). Other P450s were more over-expressed in one region than in others (Figure 1B-C) including *CYP9K1* (AFUN007549) which was highly over-expressed in Uganda (5.2x) and moderately so in Ghana (2.9x), but not in Malawi or Cameroon. *CYP6P5* (AFUN015888) was significantly over-expressed in Ghana (6.3x), Cameroon (5.8x) and Uganda (4.1x) but not in Malawi. The

duplicated *CYP6P4a* and *CYP6P4b* are highly over-expressed in Ghana (44.8x and 23.9x respectively), moderately so in Malawi and Uganda (<6x) and not in Cameroon. *CYP325A* is highly over-expressed in Cameroon (26.9x) but less so in other regions (<6x). Other P450s are moderatey over-expressed, including two paralogous *CYP9J11* genes and *CYP6N1*, up-regulated in southern and West Africa, whereas *CYP315A1* is over-expressed in all sites but Malawi. Other detoxification-associated gene families are also over-expressed (Text S1) including a cluster of glutathione S-transferase epsilon genes (*GSTe1, GSTe3, GSTe5* and *GSTe2*, a known DDT resistance gene (*10*)) up-regulated in all regions except East Africa (Text S1). Analysis of Gene Ontology enrichment (Fig. S2A-C) further confirmed these regional differences.

Quantitative RT-PCR with fifteen genes confirmed these regional differences, with a high and significant correlation observed between qRT-PCR and RNAseq results for the 4 countries when compared to FANG (R^2=0.695; P<0.001) (Fig. 1E; Fig S3A to C; Text S1).

Overall, gene expression analyses underlined the importance of P450 monooxygenases in pyrethroid resistance and identified heterogeneities in gene expression among populations from different geographical regions. Notably, the most extreme differences in expression profiles among samples were observed in members of a cluster of CYP6 genes in the *rp1* pyrethroid resistance Quantitative Trait Locus (QTL): the southern African mosquitoes showed massive over-expression of *CYP6P9a* and *CYP6P9b*, while West African mosquitoes from Ghana showed over-expression of *CYP6P5* and *CYP6P4a/b.* These differences suggest that the molecular bases of pyrethroid resistance may vary across Africa, possibly due to a combination of variation in selective pressures and restricted gene flow among regions (*11, 12*).

Whole genome polymorphism analysis of field collected mosquitoes identifies selective sweeps associated with insecticide resistance loci

Because metabolic resistance could be conferred by point mutations in coding and *cis/trans* regulatory regions, we scanned the whole genome for pyrethroid resistance-related signatures of selective sweeps in the highly pyrethroid resistant population from Malawi, to detect resistance loci. We performed pooled-template whole genome sequencing on field-collected population samples from southern Africa where high levels degree of expression of *CYP6P9a* and *CYP6P9b* have been observed. We also comparatively assessed the genomes and compared these to samples of the laboratory resistant FUMOZ-R and the susceptible FANG strains. We detected contrasting patterns of polymorphism between these strains and also between pre- (MWI-2002) and post-bednet intervention (MWI-2014) samples from Malawi (Fig. 2A; Table S3). The major selective sweep for both Malawi and FUMOZ-R was on scaffold KB119169 spanning the *rp1* pyrethroid resistance region on chromosome arm 2R (Fig. 2B). Plotting minor allele frequencies across the region revealed a valley of reduced genetic diversity around the cluster of P450s on the *rp1* pyrethroid resistance QTL correlating a selected *rp1* haplotype with *CYP6P9a* over-expression (Fig. 1D). This selective sweep appears to be at or near fixation in a contemporary Malawian population (as well as in FUMOZ-R) with little diversity observed around the P450 cluster in this highly pyrethroid resistant population (Fig. 2B). No reduced diversity was observed in the susceptible FANG strain, suggesting an association between this selective sweep and pyrethroid resistance in line with the very low *CYP6P9a* expression in FANG (Fig. 2A-B). These results are consistent with previous reports of selection on highly over-expressed resistance genes (*8*).

Complex evolution of the gene cluster of the *rp1* CYP6 genes associated with pyrethroid resistance

As *rp1* was consistently associated with pyrethroid resistance, a fine-scale analysis of this locus (120kb) was performed, revealing evidence of complex molecular evolution likely under insecticide-driven selection.

Inspection of pooled-template whole genome alignments showed two anomalous features in the 8.2kb sequence between *CYP6P9a* and *CYP6P9b*. In some samples, the coverage depth was greater than for the surrounding sequence and some samples showed read pairs in the correct relative orientation but with greater than expected insert sizes. This is indicative of a large indel: a "deletion" in the sequenced genome(s) or an "insertion" in the reference genome (Fig. S4A). This insertion corresponds to 6545 bp and appears fixed in the FUMOZ colony sample with evidences that the inserted sequence is homologous to another region nearby (on the same assembly scaffold) in the genome (Fig. S4B). In contrast, the FANG susceptible strain shows evidence of the "deletion" form of the indel (Fig. S4A). This insertion, nearly fixed across southern Africa (26/27), is absent elsewhere in the continent where only a 1.7kb intergenic region is observed after PCR (Table S4). RNAseq data showed that the inserted region is a transcribed region, showing evidence of splicing and containing three micro-RNAs, but no P450s.

We assessed the composition of this 6.5kb insert to elucidate its role by sequencing the full 8.2kb *CYP6P9a/b* intergenic region and analysing it using GPminer (*13*). This insert contains abundant binding sites for transcription factors including a CpG island (1.3kb), several GATA, TATA (35), CCAAT (12) and GC (11) boxes and over-represented oligonucleotides. It also contains several binding sites for key transcription factors associated with xenobiotic detoxification, including Cap-n-Collar-C (CnCC) (51 sites) and Muscle aponeurosis fibromatosis (Maf), suggesting that this insertion may drive *CYP6P9a/*b over-expression. The insert also contains a microsatellite (FUNR) between 6082bp and 6482bp, only 80bp from the 5' untranslated region (UTR) of *CYP6P9a*. Previous genotyping of this marker Africa-wide revealed marked differences associated with pyrethroid resistance profile (*11*). FUNR is not present within the 1.7kb intergenic region between *CYP6P9a* and *CYP6P9b* in the susceptible FANG strain. It has been shown that microsatellite loci are involved in upregulation of P450 to

confer insecticide resistance in other insects such as aphids (*14*). It has also been shown in Yeast that polymorphic tandem repeats in the promoter regions can activate gene expression by impacting local chromatin structure to act as "evolutionary tuning knobs" to drive rapid evolution of gene expression such in a case of insecticide resistance selection.

Polymorphism of *CYP6P9a* regulatory region associated with pyrethroid resistance

Polymorphism of *CYP6P9a* 5' UTR and upstream region Africa-wide: To detect *cis*-regulatory mutations controlling *CYP6P9a/b*-based pyrethroid resistance, we compared 800bp immediately upstream of *CYP6P9a* Africa-wide. Several locations exhibited a very low or no polymorphism (Fig. S5A). The sample with the highest diversity was the fully susceptible FANG with the highest diversity indexes supporting the selection acting on this gene in field resistant populations. Despite low diversity observed in different regions, southern Africa populations consistently exhibited a different polymorphism pattern to other regions including the presence of an AA insertion 8bp upstream of a putative CCAAT box present only in southern African samples through an A/C substitution. The AA insertion located 359 bp from start codon is tightly associated with other polymorphisms in a haplotype (STH10) which is nearly fixed in southern Africa (63/68) reflecting the marked selective sweep observed around this gene in this region. Analysis of the phylogenetic tree revealed four clusters of haplotypes from different regions notably southern Africa (Malawi, Mozambique and Zambia plus FUMOZ-R), East-Central (Kenya, Uganda and Cameroon), West (Ghana) and West-Central (Benin and Congo) (Fig. 3A). The FANG susceptible strain forms its own cluster divergent from the others.

Closer analysis of the haplotypes using a haplotype network confirmed the presence of 4 major haplotypes corresponding to these geographical clusters, STH10 in southern, EST/CNT24 in East-Central, BEN/DRC21 in West-Central and GHA11 in Ghana only (West)

(Fig. 3B). Surprisingly, the other 3 regions also exhibit predominant haplotype in the populations to near fixation contrary to previous data where they were more polymorphic (*11*). This result suggests that resistance to pyrethroid beyond southern Africa could have also been selected through *CYP6P9a* or other genes in the vicinity on the *rp1* QTL regions. The hypothesis of a possible hitchhiking effect here rather than the direct involvement of *CYP6P9a* is supported by the important differences observed between the four major haplotypes with more than 20 mutational steps of difference between them. Therefore, it is very likely that resistance conferred by the *rp1* locus occurred through independent selective events with likely different genes. This is supported by RNAseq data showing that although *rp1* genes are over-expressed Africa-wide the main genes are different with *CYP6P4a* predominant in Ghana but not in others, *CYP6P5* in Cameroon and Uganda whereas *CYP6P9a* is the major gene in southern Africa. The neighbor-joining phylogenetic tree of the *Nst* genetic distances between different countries (Fig. 3C) correlated with the polymorphism patterns showing countries clustering according to the haplotypes diversity pattern from ML tree or the TCS network.

Scaled-up use of insecticide-treated nets has selected for changes in the regulatory region of *CYP6P9a*: To assess whether the differences observed in the 5'UTR and upstream region of *CYP6P9a* between southern Africa and other regions was a result of selective pressure from insecticides, we compared samples from southern Africa before the scale-up of insecticide-based interventions such as LLINs and samples after scale-up. 39 clones of the 800bp fragment upstream of *CYP6P9a* were obtained and sequenced from 'pre-intervention' mosquitoes from Mozambique and Malawi and 52 from 'post-intervention' samples of both countries. Strikingly this portion was highly polymorphic pre-intervention with many segregating sites (61 and 25 in Malawi and Mozambique, respectively) and haplotypes (19 and 12, a total of 30 for both) and high nucleotide diversity (π=0.027 and 0.012) (Table S5; Fig. S5B). By contrast, the post-intervention samples exhibited very low diversity as revealed by all

indexes with S of 4 and 3, haplotypes number of 2 and 4 with extremely low nucleotide diversity (π=0.00066 and 0.0008), respectively in Malawi and Mozambique. The difference is reflected on the Maximum Likelihood (ML) phylogenetic tree showing that mosquito samples of pre-intervention not only cluster together but are more diverse with several haplotypes (Fig. 4A). In contrast the post-intervention samples cluster noticeably away from the pre-intervention with drastically reduced haplotype number. A haplotype network confirmed that the major haplotype associated with resistance and now nearly fixed in all southern African populations was present also in the pre-intervention sample but at a much lower frequency of only 4/39 (19.23%) contrary to 44/52 (84.6%) in post intervention with other post haplotypes only 1 or 2 mutational steps away from the predominant one (Fig. 4B). The pre-intervention haplotypes were polymorphic and separated with high mutational steps whereas the post-intervention samples showed a marked reduced diversity (Fig. 4B). A detailed analysis of the polymorphisms between pre- and post-intervention samples, revealed that the AA insertion now common in southern populations as well as the CCAAT box was also present in pre-intervention samples but only at a very low frequency (7/24 in Malawi and 0/15 in Mozambique) whereas both AA and CCAAT box are now fixed in all the post-intervention samples (52/52). Furthermore, a second Nrf2:MafK binding site is found only in post-intervention samples and linked with both the AA insertion and the CCAAT box. These major modifications show that scale up of insecticide-treated nets is very likely the major factor that has driven this evolution in *An. funestus* populations in southern Africa.

Detection of a molecular marker associated with pyrethroid resistance gene over-expression

Having confirmed that genomic changes upstream of *CYP6P9a* are associated with pyrethroid resistance, we next searched for the mutations responsible for the over-expression of *CYP6P9a* in resistant mosquitoes. We used luciferase assay to assess the role of

polymorphisms found in the 800bp upstream of the translation start site (including the 5' UTR). The 800bp upstream of the *CYP6P9a* translation start site in both FUMOZ and FANG was successfully cloned and sequenced. To narrow down the region containing the regulatory motifs, four different sized fragments of the 800, 500, 300 and 150 bp immediately upstream of the translation start codon were cloned upstream of a reporter gene in a pGL3 vector. These constructs were used in luciferase reporter gene assays that demonstrated that, while 800 bp inserts from both strains drove reporter gene expression, expression driven by the FUMOZ insert was 3 times higher than that from FANG (Fig. S6A), supporting that this region plays a role in the differential expression of *CYP6P9a* between resistant and susceptible mosquitoes. Progressive deletion of the 800-bp pGL3-FZ-CYP6P9a was performed to identify the major regulatory elements. The first deletion from 800 to 500bp did not impact the activity of the fragment. However, when the fragment was cut from 500 to 300bp removing the AA insert and the CCAAT box, it induced a 33% reduction of activity in the FUMOZ clone ($P<0.001$) (Fig. 5A). Subsequent deletion from 300 to 150bp, removing the AA insert, the CCAAT box and the resistant specific CnCC/MafK binding site, led to a massive 89% reduction in activity ($P<0.001$). This shows that both the CCAAT box and the CnCC/MafK binding sites are key regulatory enhancer elements driving the over-expression of *CYP6P9a*.

Design of a DNA-based diagnostic assay to detect *CYP6P9a*-mediated pyrethroid resistance

To design a DNA-based diagnostic assay to detect *CYP6P9a*-mediated resistance, we screened the most active portion (500bp) for the presence of restriction site polymorphisms that could be used to design a simple PCR-RFLP. We found a restriction site for the TaqI enzyme (cut site 5'-TCGA-3') spanning an A/G mutation located 18bp of the AA insertion (Fig. S6B) and completely tight with the CCAAT box and other regulatory elements on the resistance haplotype. The TaqI enzyme cuts the 450bp fragment from the 'resistant' haplotype into two

fragments of 350bp and 100bp whereas 'susceptible' haplotypes remain uncut (Fig. 5B) allowing us to genotype the resistance allele in single mosquitoes. To validate the robustness of this PCR-RFLP to detect pyrethroid resistance, we used F_8 progeny from a cross between highly resistant (FUMOZ) and highly susceptible (FANG) strains. The genotyping of 46 mosquitoes highly resistant (alive after 180 minutes' exposure) to permethrin (*15*) revealed 9 RR, 35 RS and only 2 SS genotypes. By contrast, 42 highly susceptible mosquitoes (dead after 30 minutes' permethrin exposure) had 0 RR, 1 RS and 41 SS genotypes. Therefore, the odds ratio of the likelihood of surviving exposure to permethrin when homozygote for the resistant allele (RR) of the *CYP6P9a* gene (with the CCAAT box and CnCC/MafK binding sites) is highly elevated at 922 ($P<0.0001$) compared to the homozygous susceptible (SS) (Fig. 5C) demonstrating the reliability of this DNA-based metabolic resistance diagnostic assay.

Geographical distribution of the resistant *CYP6P9a* allele across Africa: Genotyping the CYP6P9a_R allele across Africa revealed that it is nearly fixed in southern Africa and present at an intermediate frequency in Tanzania (East Africa) (55.7%). However, CYP6P9a_R is absent from Central/West Africa (Fig. 5D; Fig. S6C). In DR Congo, a geographical contrast was observed, with the CYP6P9a_R mutation present in the East of the country but absent from the West (Kinshasa). This pattern suggests a 'new' allele/haplotype that arose in southern Africa and is spreading northward. Regional differences in CYP6P9a_R distribution are like those reported for previous markers (*6*, *10*).

CYP6P9a_R reduces the effectiveness of insecticide-treated nets

We next assessed the impact of *CYP6P9a*-based metabolic resistance on the effectiveness of LLINs using experimental hut trials. Mosquitoes from FANG/FUMOZ crosses (F_4) were used after confirming their pyrethroid resistance status (Fig. S7A) and established that resistance was driven additively by *CYP6P9a* [OR = 693 and 131 respectively for RR vs SS and RR vs RS ($P<0.001$)] (Fig. 6A; Fig. S7B-D).

Impact of *CYP6P9a*-mediated metabolic resistance on bed net efficacy using experimental huts: Females from the F_4 generation of the FANG/FUMOZ strain were used in a release-recapture experiment in huts with PermaNet 2.0 (treated with deltamethrin), PermaNet 3.0 (deltamethrin plus PBO) and control nets (untreated). After 4 consecutive nights of release-recapture, analysis of the mosquitoes collected showed no significant induced exophily for both PermaNet 2.0 and PermaNet 3.0 compared to the control ($P>0.05$) (Table 1). The percentage of blood-fed females was significantly lower in both treated nets than in the untreated control, with PermaNet 3.0 showing a significantly lower number of blood-fed mosquitoes than PermaNet 2.0 ($P<0.001$) (Fig. 6B). This is reflected in the percentage of blood feeding inhibition which is significantly higher ($P<0.0001$) for PermaNet 3.0 (76.8%) than for PermaNet 2.0 (49.8%). Both treated nets provided greater personal protection than the untreated nets with, again, higher protection from PermaNet 3.0 (79.04%) than PermaNet 2.0 (61.9%) (Table 1). Analysis of mortality rates revealed very high mortality of the hybrid FANG/FUMOZ strain against PermaNet 3.0 (98.7%) compared to only 33.3% for PermaNet 2.0 ($P<0.001$) (Fig. 6B). Very low mortality was observed in the control untreated net. Mosquitoes that had taken a blood meal through the nets also present a high mortality rate for PermaNet 3.0 (95.5%) but only a low rate for PermaNet 2.0 (40%).

Assessing the role of *CYP6P9a* in the loss of efficacy of PermaNet 2.0: Genotyping of the *CYP6P9a* maker allowed us to assess the impact of P450-based metabolic resistance on the loss of efficacy of PermaNet 2.0. Because most of the mosquitoes released in the PermaNet 3.0 huts died, the impact of the *CYP6P9a* on the ability to survive exposure to LLIN was done only for PermaNet 2.0. To avoid confounding effects from blood feeding or net entry or exophily status, the distribution of the *CYP6P9a* genotypes was assessed first only among unfed mosquitoes collected in the room revealing a highly significant difference in the frequency of the three genotypes between the dead and alive mosquitoes ($Chi^2=375$; $P<0.0001$) (Fig. 6B).

Analysis of the correlation between each genotype and mortality revealed that *CYP6P9a* homozygous resistant mosquitoes (RR) were significantly more able to survive exposure to the PermaNet 2.0 than homozygous susceptible mosquitoes (odds ratio=34.9; CI=15.8-77.1; P<0.0001) (Table S6). Similarly, possessing the heterozygous (RS) genotype of *CYP6P9a* also confers a significant survival advantage than for the homozygous SS genotype (OR=19.9; CI=9.7-40.9; P<0.0001). Although a higher frequency of RR is observed in the alive mosquitoes than for the RS genotype, this difference was not significant (OR=1.75; CI=0.82-3.7; P=0.26). Overall, a single *CYP6P9a* resistant allele (R) significantly confers the ability to survive than the susceptible allele (OR=6.25; CI=3.3-11.7; P<0.0001) (Fig. 6C). The same trend was observed when comparing the mortality in all samples although with at a lower degree (Table S6). The impact of the *CYP6P9a* resistance gene on the ability to blood feed was also assessed, revealing that the distribution of the three genotypes was significantly different between blood fed and unfed mosquitoes for both PermaNet 2.0 (Chi^2=16.9; P<0.0001) and PermaNet 3.0 (Chi2=30.5; P<0.0001). Homozygous RR mosquitoes were significantly more likely to blood feed when exposed to PermaNet 3.0 than both susceptible SS (OR= 4.54; P<0.0001) and heterozygote RS individuals (OR= 2.6; P=0.0012) (Fig. 6C). A similar trend was observed for PermaNet 2.0, although not significant (Fig. S7E-F). This shows that despite the higher efficacy of PermaNet 3.0 compared to PermaNet 2.0, CYP6P9a_R could still increase the risk of malaria transmission.

Discussion

This study presents a comprehensive elucidation of the molecular genetic basis of metabolic resistance to pyrethroids in a major malaria vector detecting the first DNA-based resistance marker for P450-mediated metabolic resistance. A major outcome was the design of a field applicable diagnostic assay to detect and track the spread of this resistance across Africa

and enabled us to establish the direct impact of metabolic resistance to pyrethroids on the efficacy of insecticide-treated nets in experimental huts.

Overall, gene expression analyses underlined the importance of cytochrome P450 monooxygenases in pyrethroid resistance as previously reported in other *An. funestus* populations (*8*, *15*) and in other mosquito species (*16-18*). Important heterogeneities in gene expression were observed among populations from different geographical regions. Most notably, a cluster of CYP6 genes in the *rp1* pyrethroid resistance QTL showed the most extreme differences in expression profiles among locations. Such differences show that the molecular basis of pyrethroid resistance varies across the continent as previously suggested for this species (*19*) and for other malaria vectors such as *An. gambiae* (*7*, *20*).

The association of the *rp1* QTL locus with pyrethroid resistance is further supported by the detection of several adaptive evolutionary features across this locus, including i) signatures of selective sweep; ii) large structural variation with 6.5kb insert; and iii) *cis*-regulatory polymorphisms. The selective sweep detected around this region in southern Africa coincides with the high expression of the duplicated P450s *CYP6P9a* and *CYP6P9b* in Malawi supporting the key role in pyrethroid resistance of these genes previously shown to be efficient pyrethroid metabolizers (*8*). Selective sweeps associated with insecticide resistance have been reported recently in *An. gambiae* across Africa (*21*).

The *cis*-regulatory changes detected here include binding sites for transcription factors such as CnCC/MAfK which has recently been shown to be involved in metabolic resistance in *An. gambiae* (*22*) and in other insects (*23*), supporting that *cis*-regulatory modifications are important drivers of metabolic resistance to insecticides. However, because *cis*-regulatory elements/enhancers are able to drive expression of genes from distant acting locations and can be upstream or downstream of the gene on which they function, future studies will help establish the role of the 6.5kb insert in the over-expression of *CYP6P9a* and *CYP6P9b*. The

strong association of these regulatory variations with pyrethroid resistance provided an excellent opportunity to design a DNA-based diagnostic tool for metabolic resistance. This is the first diagnostic assay for P450-mediated resistance in malaria vectors which constitutes a major achievement in the field of resistance monitoring and surveillance. Indeed, while the first DNA-based diagnostic for target-site resistance (*kdr*) was established two decades ago (*24*), no such tool has been designed for metabolic resistance despite its greater risk to control interventions. Although sets of SNPs associated with pyrethroid resistance have been recently detected in the dengue vector *Aedes aegypti*, no causative markers were detected (*25*). Previous design of a DNA-based diagnostic tool were for a glutathione S-transferase gene (*GSTe2*) in *An. funestus* conferring pyrethroid/DDT resistance in West/Central Africa but using an amino acid change in the coding region of the gene, not the putative causative variant for over-expression as done here for *CYP6P9a* (*10*). However, because this DNA-based assay mainly detects resistance in southern Africa and only applies to *An. funestus*, further efforts are needed to detect similar markers in other regions and other species to comprehensibly track P450-mediated metabolic resistance Africa-wide.

Strikingly, the CYP6P9a_R allele is present mainly in southern Africa, where it is nearly fixed, but completely absent from other regions. Such contrast between African populations of *An. funestus* has previously been observed, notably for the distribution of other resistance markers such as the *L119F-GSTe2* (*10*) and the *A296S-RDL* conferring dieldrin resistance (*26*), which are present in West/Central and East Africa but completely absent from southern Africa. On the other hand, the *N485I-ace-1* carbamate resistance allele is present only in southern Africa (*6*). Furthermore, patterns of F_{ST}-based genetic differentiation indicate a restriction of gene flow and high genetic divergence between southern Africa and other regions (*11, 12*). However, there seems to be a gradual increase of CYP6P9a_R frequency from south to north in southern Africa as seen in Malawi with 98% in the south (Chikwawa), 90% in the Centre

(Salima) and 78 % in the north (Fulirwa). This correlates with previous observations that *CYP6P9a* over-expression was lower in the north and that the spread of this resistance likely originated from the far south and is spreading northwards across the species' range (*27*). It will be important to monitor the spread of such alleles across the continent as there is the risk that super-resistant mosquitoes could be generated if, for instance, *CYP6P9a*-mediated pyrethroid resistance combines with the *GSTe2*-based DDT resistance seen in West/Central Africa. DR Congo will be particularly important to monitor as both resistance mechanisms are present in this country (*28*).

Using the novel CYP6P9a_R assay, we successfully assessed for the first time the direct impact of metabolic resistance on the efficacy of insecticide-treated nets showing that P450-mediated resistance was directly reducing the efficacy of insecticide-treated nets. This evidence further helps clarify the debate about whether pyrethroid resistance is directly impacting the efficacy of insecticide-treated nets (*29*) as it clearly highlights that pyrethroid resistance could jeopardize insecticide-based interventions. Interestingly, nets containing the insecticide synergist PBO, that inhibits the activity of cytochrome P450 enzymes, provided better efficacy than those with pyrethroid alone. However, despite the high mortality observed with the PBO-based net, CYP6P9_R still increases the risk of malaria transmission with this net as resistant mosquitoes are still more likely than susceptible mosquitoes to bite and potentially transmit malaria, suggesting that malaria elimination efforts will be impeded unless the over-reliance on pyrethroids is addressed.

Methods

Study design

The objective of this study was to detect key genetic variants conferring metabolic-mediated pyrethroid resistance in *An. funestus* and design a simple DNA-based assay to monitor such resistance in field populations and assess its impact on the effectiveness of insecticide-based

control tools. Transcriptome profiling of *An. funestus* populations from four African regions [southern (Malawi), East (Uganda), West (Ghana), Central (Cameroon)] were analysed to detect key candidate resistance genes. Because metabolic resistance could also be conferred by point mutations in coding and *cis/trans* regulatory regions, we performed a comparative whole genome sequencing of field permethrin resistant and susceptible mosquitoes to screen for genomic resistance regions and polymorphisms. To comprehensively detect resistance loci, we also scanned the whole genome for pyrethroid resistance-related signatures of selective sweeps in southern Africa. As *rp1* QTL was consistently associated with pyrethroid resistance, a fine-scale analysis of this locus was performed to detect potential structural variants associated with resistance such as Indels and copy number variations. To detect *cis*-regulatory mutations controlling *CYP6P9a/b*-based pyrethroid resistance, we compared 800bp immediately upstream of *CYP6P9a* in resistant and susceptible mosquitoes. The role of insecticide-based interventions in the selection of this *CYP6P9a* cis-regulatory changes was assessed by sequencing mosquitoes collected before and after widespread insecticide-treated nets usage. To establish the specific mutations controlling *CYP6P9a* over-expression, a comparative luciferase assay between resistant and susceptible *CYP6P9a* promoter sequences was performed. To design a DNA-based diagnostic assay to detect *CYP6P9a*-mediated resistance, we screened the promoter for restriction sites to design a PCR-RFLP assay and use it to assess its distribution continent-wide. We next assessed the impact of *CYP6P9a*-based metabolic resistance on the effectiveness of LLINs using experimental hut trials.

Collection and rearing of mosquitoes

Two *An. funestus* laboratory colonies were utilised in the study. The FANG colony is a fully insecticide susceptible colony derived from Angola (*30*). The FUMOZ colony is a multi-insecticide resistant colony derived from southern Mozambique. Mosquitoes were collected from 4 primary locations across the continental range of *An. funestus*. Mosquitoes were

447 collected in March 2014 from Obuasi (5°56′ N, 1°37′ W) in Ghana (*31*); in February 2015 from
448 Mibellon (6°46′ N, 11°70′ E) in Cameroon; in March 2014 from Tororo (0°45’ N, 34°5’ E) in
449 Uganda (*32*) and in January 2014 from Chikwawa (16°1’ S, 34°47’ E) in southern Malawi (*33*).
450 Collected mosquitoes were kept until fully gravid and forced to lay eggs using the
451 forced-egg laying method (*34*). All F_0 females that laid eggs were morphologically identified
452 as belonging to the *An. funestus* group according to a morphological key (*35*). Parents (F_0) and
453 egg batches were transported to the Liverpool School of Tropical Medicine under a DEFRA
454 license (PATH/125/2012). Eggs were allowed to hatch in cups and mosquitoes reared to
455 adulthood in the insectaries under conditions described previously (*34*). Insecticide resistance
456 bioassays on these samples have been previously described (*31-33*).
457 **Transcription profiling of pyrethroid resistance using RNAseq**
458 Total RNA was extracted from pools of 10 female mosquitoes (alive after 1h permethrin
459 exposure) using the Arcturus PicoPure RNA isolation kit (Life Technologies), according to the
460 manufacturer’s instructions (Text S1). Pools of libraries were sequenced, 8 per lane of the
461 HiSeq 2500 (Illumina, San Diego, CA, USA) at 2x125 bp paired-end sequencing. Sequence
462 library preparation and sequencing were done at the Centre for Genomic Research (CGR),
463 University of Liverpool. RNAseq data were analysed as described previously (*36*) (Text S1).
464 This involved an initial processing, quality assessment of sequences and alignment to the
465 reference sequence using the AfunF1.4 annotation. Differential gene expression analysis was
466 performed using edgeR and StrandNGS program (Strand Life Sciences, version 3.0) (Text S1).
467 **Whole genome sequencing**
468 Genomic DNA was extracted from whole mosquitoes from F_0 Malawi samples (2014),
469 the pyrethroid resistant FUMOZ-R laboratory strain and the fully susceptible FANG strain
470 using the DNAeasy kit (Qiagen, Hilden, Germany). For each sample, genomic DNA was
471 extracted from individuals and pooled in equal amounts to form pools of DNA. These were

sequenced on an Illumina HiSeq2500 (2x150bp paired-end). Initial processing and quality assessment of the sequence data was performed as for RNAseq data. Alignment of POOLseq R1/R2 read pairs and R0 reads to the reference sequence (the same as used for RNAseq alignment) as well as variant calling were performed as described previously (*11*).

Polymorphism analysis of the promoter region of *CYP6P9a*

To detect potential causative mutations conferring pyrethroid resistance in *An. funestus*, the polymorphism of the cis-regulatory region of the pyrethroid resistance gene *CYP6P9a* was analyzed.

i-Detection of the causative mutations driving up-regulation of *CYP6P9a*: An 800 bp region upstream of the start codon of *CYP6P9a* was amplified and directly sequenced in 15 field collected mosquitoes each from ten countries across different regions of Africa including southern (Mozambique, Malawi and Zambia), East (Uganda, Kenya and Tanzania), Central (DR Congo and Cameroon) and West (Benin and Ghana). Primers are listed in Table S7. Amplification and purification of PCR products was performed as previously described (*11*). Sequences were aligned using ClustalW (*37*) while haplotype reconstruction and polymorphism analysis were done using DnaSPv5.10 (*38*), MEGA (*39*) and TCS (*40*).

ii) Investigation of the content of the *CYP6P9a* and *CYP6P9b* intergenic region: The entire 8.2kb intergenic region between both genes was amplified for the FUMOZ and FANG strains in 2 to 3 fragments using primers listed in Supplementary Table 22. PCR was performed using the Phusion polymerase following the manufacturer's instructions. PCR products were purified and cloned into pJET1.2 plasmid.

iii) Assessing the role of the scale up of insecticide-based intervention in the changes observed in the polymorphism of the promoter region of *CYP6P9a*: The same 800bp region upstream of the *CYP6P9a* was amplified in mosquitoes collected before (pre-intervention) the scale up of insecticide-treated nets and also after the scale up (post-

intervention) in Malawi (2002 and 2014) and Mozambique (2000 and 2016). The PCR products were cloned and sequenced and sequencing data analyzed as described above.

Genotyping of the *CYP6P9a* resistance allele using PCR-RFLP

A restriction site (5'-TCGA-3') for the TaqI enzyme at the A/G mutation located 18bp of the AA insertion and completely tight with the CCAAT box on the resistance haplotype was used to design a PCR-RFLP assay to genotype the CYP6P9a_R allele. The RFLP6P9aF forward primer, 5'- TCCCGAAATACAGCCTTTCAG-3 and RFLP6P9aR 5'-ATTGGTGCCATCGCTAGAAG-3' reverse primers were used to amplify a partial *CYP6P9a* upstream region containing the restriction site. 10µl of the digestion mix made of 1µl of CutSmart buffer, 0.2µl of 2 units of TaqI restriction enzyme enzyme (New England Biolabs, Ipswich, MA, USA), 5µl of PCR product and 3.8µl of dH_20 was incubated at 65°C for 2 hours. Restriction digest was separated on 2.0% agarose gel (Fig. 3D).

Validation of the diagnostic test: To validate the robustness of the PCR-RFLP to detect the pyrethroid resistance in field population, F_8 progeny from a cross between highly resistant (FUMOZ) and highly susceptible (FANG) strains previously used for QTL mapping (*15*) were genotyped and correlation with resistance phenotype established using Odd Ratio.

Geographical distribution of resistant *CYP6P9a* allele across Africa: The geographical distribution of the resistant *CYP6P9a* allele across Africa was established by genotyping the CYP6P9a_R in 30-50 field-collected females of *An. funestus* from several countries in Africa using the PCR-RFLP.

Luciferase reporter assay of *CYP6P9a* core promoter region

The region immediately 5' of *CYP6P9a* from both FUMOZ and FANG strains was amplified using primers 6P9a1F and 6P9R a/b. These primers gave an 817bp product for both FUMOZ and FANG which were cloned into pJET1.2 (Thermo Fischer Scientific) and sequenced. Primers were designed (Supplementary Table 22) to obtain constructs of

progressive serial 5' deletions of the *CYP6P9a* promoter of 800bp, 500bp, 300bp and 150bp for the different primers. The primers incorporated either the SacI or MluI (for the FUMOZ) and KpnI or HindIII (for the FANG) to facilitate cloning in the pGL3 basic vector. Products were amplified with Phusion polymerase (Thermo Fischer Scientific) and cloned into pJET 1.2. The product was then excised from pJET1.2, ligated to pGL 3 basic (Promega) and sequenced. Plasmids were then extracted using Midiprep kit (Qiagen) to obtain high concentrations for the transfection. Dual luciferase assay was undertaken using *An. gambiae* cell line 4a-2 cell line (MRA-917 MR4, ATCC® Manassas Virginia). Approximately 5 $x10^5$ cells (600µl) was sub-cultured from a T75 culture and seeded in each well of 24-well plate, 1 day prior to transfection and allowed to reach 60-70% confluency.

Transfection of the construct was carried out using the Qiagen effectene transfection reagent and the promoter activity measured using the Dual Luciferase Reporter Assay (promega, Madison, WI, USA). 600ng reporter constructs (*CYP6P9a* upstream sequences in pGL3-Basic), pGL3 without insert and LRIM promoter in pGL3 basic were co-transfected with 1ng actin-renilla internal control in 60ml DNA condensation buffer, 4.8 ml enhancer and 6 ml effectene in triplicate. After 48h incubation at 25°C, the cells were washed with PBS and harvested in 100ml passive lysis buffer (Promega) and luciferase activity was measured on a luminometer (EG & GBert-hold, Wildbad, Germany). Renilla luciferase activity was used to normalize the construct luciferase activity. The values obtained after measuring the firefly (LAR II) luciferase activity, which represent the activity of the promoter, was divided by the corresponding Renilla luciferase activity values and the ratio used to compare different promoters.

Evaluation of impact of *CYP6P9a*-based metabolic resistance on efficacy of insecticide-treated nets using experimental huts

Study area and hut description: The study was performed in Mibellon (6°4'60" N and 11°30'0" E), a village in the Adamawa region of Cameroon where we recently built 12 experimental huts of concrete bricks, following the specific design for experimental huts from the West Africa region (*41*).

Mosquito strains: The study was carried out with a hybrid strain generated from reciprocal crossing between the highly pyrethroid resistant strain FUMOZ-R (CYP6P9a_R) and the fully susceptible FANG strain (CYP6P9a_S) (*42*). After the initial F_1 generation obtained from the reciprocal crosses of 50 males and 50 females of both strains, the hybrid strain was reared to F_5 and F_6 generations, which were used for the release in the huts.

Susceptibility profile of the hybrid FANG/FUMOZ strain: WHO bioassays were carried out to assess the susceptibility profile of the two reciprocal hybrid strains for the pyrethroids (0.75% permethrin and 0.05% deltamethrin), DDT (4%) and the carbamate, bendiocarb (0.1%). The bioassays were performed according to WHO protocol (*41*).

Study design: The following three treatments were compared in the experimental huts: (i) Untreated polyethylene net; (ii) PermaNet 2.0® (Deltamethrin incorporated into polyethylene net); (iii) PermaNet 3.0® (PBO + Deltamethrin incorporated into polyethylene net). To simulate a worn net, six holes of 4cm x 4cm were made on each net, according to WHO guidelines. The hybrid FANG/FUMOZ strain was released in each hut for 6 nights (80 mosquitoes per hut).

Ethical clearance: Three adult volunteers were recruited from the Mibellon village to sleep under the nets and collect mosquitoes in the morning. They were provided with a written consent form and given chemoprophylaxis during the trial. Ethical approval was obtained for the National Ethic Committee of the Ministry of the Health in Cameroon.

Mosquito collection in huts: Early in the morning, mosquitoes were collected using glass tubes from: i) the room (the floor, walls and roof of the hut); ii) inside the net; iii) the exit

traps (veranda). Each compartment had its own bag to avoid mixture between samples. Surviving mosquitoes were provided with sugar solution and held for 24 h in paper cups after which delayed mortality was also assessed. Samples were recorded in observation sheets as dead/blood fed, alive/blood fed, dead/unfed, and alive/unfed.

The effect of each treatment was expressed relative to the control (untreated net) by assessing: i) induced exophily (the proportion of mosquitoes that exit early through the exit traps, treatment-induced exiting); ii) the mortality rate, an indicator of the potential mass killing effect of the LNs; iii) the blood feeding rate, an indicator of personal protection.

Genotyping of the *CYP6P9a* metabolic resistance marker: To establish the impact of the *CYP6P9a*-mediated metabolic resistance to pyrethroids on the effectiveness of the insecticide-treated nets, the PCR-RFLP diagnostic assay was used to genotype a subset of each treatment including the dead, alive, blood feed, unfed, mosquitoes on the veranda, in the net and in the room.

Statistical analysis:

Genes differentially expressed in each country and between different countries (generated from Venn diagrams) were detected using DESeq normalization with fold change >2 and performing a multiple test correction using the method of Benjamini and Hochberg, at a false discovery rate of 5% (adjusted p value <0.05). Statistical significance of the Luciferase assays was assessed after an unpaired Student's t-test with respective replicates between FUMOZ and FANG. Correlation between the CYP6P9a_R allele and pyrethroid resistance phenotype was established using Odd Ratio and Fisher's exact test. The effect of both treated nets was established through a direct comparison to the untreated control net. The statistical significance of the difference was estimated by a logistic regression model using Wald statistic that follows a chi-squared distribution (with df=1). Odd Ratio and Fisher's exact tests were

595 used to assess the impact of CYP6P9a_R on the ability to survive and blood-feed after exposure
596 to insecticide-treated nets.
597
598

List of supplementary materials

Text S1: Additional information on results and materials and methods

Fig. S1. Differential gene expression between four permethrin-exposed samples and FANG.

Fig. S2: Gene ontology enrichment of upregulated genes using BLAST2GO

Fig. S3: qRT-PCR validation of the expression profile of the main detoxification genes differentially expressed between resistant and susceptible pyrethroid samples with RNAseq.

Fig. S4: Insertion of a 6.5kb intergenic fragment between *CYP6P9a* and *CYP6P9b* in southern African mosquitoes.

Fig. S5: Genetic diversity patterns of an 800bp cis-regulatory genomic fragment of *CYP6P9a* Africa-wide and before and after LLIN scale up.

Fig. S6: Design of a DNA-based diagnostic assay to detect and track pyrethroid resistance across Africa.

Fig. S7: Impact of the *CYP6P9a*-based metabolic resistance on the efficacy of bed nets using semi-field experimental hut trials

Table S1. Descriptive statistics of RNAseq sequence read data and alignments for different samples

Table S2: Detoxification-associated genes differentially expressed between the four pyrethroid resistant populations and the FANG susceptible strain

Table S3. Descriptive statistics of Whole genome POOLseq sequence read data

Table S4: Counts of reads aligned at the left and right breakpoints of the 6.5 kb insertion supporting different haplotypes

Table S5: Population genetic parameters of the 800bp fragment upstream of *CYP6P9a*

Table S6: Correlation between genotypes of *CYP6P9a* and mortality (PermaNet 2.0) and blood feeding after the experimental hut trial with the FANG/FUMOZ strain

Table S7: Primers used for characterization of the promoter of *CYP6P9a*

References and Notes

1. S. Bhatt, D. J. Weiss, E. Cameron, D. Bisanzio, B. Mappin, U. Dalrymple, K. E. Battle, C. L. Moyes, A. Henry, P. A. Eckhoff, E. A. Wenger, O. Briet, M. A. Penny, T. A. Smith, A. Bennett, J. Yukich, T. P. Eisele, J. T. Griffin, C. A. Fergus, M. Lynch, F. Lindgren, J. M. Cohen, C. L. Murray, D. L. Smith, S. I. Hay, R. E. Cibulskis, P. W. Gething, The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature* **526**, 207-211 (2015).
2. J. Hemingway, The way forward for vector control. *Science* **358**, 998-999 (2017).
3. J. Hemingway, H. Ranson, Insecticide resistance in insect vectors of human disease. *Annual revue of entomology* **45**, 369-389 (2000).
4. WHO, *Global Plan for Insecticide Resistance Management (GPIRM)*. W. G. M. Programme, Ed., (World Health Organization, Geneva, Switzerland, 2012).
5. C. V. Edi, L. Djogbenou, A. M. Jenkins, K. Regna, M. A. Muskavitch, R. Poupardin, C. M. Jones, J. Essandoh, G. K. Ketoh, M. J. Paine, B. G. Koudou, M. J. Donnelly, H. Ranson, D. Weetman, CYP6 P450 enzymes and ACE-1 duplication produce extreme and multiple insecticide resistance in the malaria mosquito Anopheles gambiae. *PLoS Genet* **10**, e1004236 (2014).
6. S. S. Ibrahim, M. Ndula, J. M. Riveron, H. Irving, C. S. Wondji, The P450 CYP6Z1 confers carbamate/pyrethroid cross-resistance in a major African malaria vector beside a novel carbamate-insensitive N485I acetylcholinesterase-1 mutation. *Mol Ecol* **25**, 3436-3452 (2016).
7. S. N. Mitchell, B. J. Stevenson, P. Muller, C. S. Wilding, A. Egyir-Yawson, S. G. Field, J. Hemingway, M. J. Paine, H. Ranson, M. J. Donnelly, Identification and validation of a gene causing cross-resistance between insecticide classes in Anopheles gambiae from Ghana. *Proc Natl Acad Sci U S A* **109**, 6147-6152 (2012).
8. J. M. Riveron, H. Irving, M. Ndula, K. G. Barnes, S. S. Ibrahim, M. J. Paine, C. S. Wondji, 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**, 252-257 (2013).
9. X. Li, M. A. Schuler, M. R. Berenbaum, Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. *Annu Rev Entomol* **52**, 231-253 (2007).
10. J. M. Riveron, C. Yunta, S. S. Ibrahim, R. Djouaka, H. Irving, B. D. Menze, H. M. Ismail, J. Hemingway, H. Ranson, A. Albert, C. S. Wondji, A single mutation in the GSTe2 gene allows tracking of metabolically-based insecticide resistance in a major malaria vector. *Genome Biol* **15**, R27 (2014).
11. K. G. Barnes, G. D. Weedall, M. Ndula, H. Irving, T. Mzihalowa, J. Hemingway, C. S. Wondji, Genomic Footprints of Selective Sweeps from Metabolic Resistance to Pyrethroids in African Malaria Vectors Are Driven by Scale up of Insecticide-Based Vector Control. *PLoS Genet* **13**, e1006539 (2017).
12. A. P. Michel, M. J. Ingrasci, B. J. Schemerhorn, M. Kern, G. Le Goff, M. Coetzee, N. Elissa, D. Fontenille, J. Vulule, T. Lehmann, N. Sagnon, C. Costantini, N. J. Besansky, Rangewide population genetic structure of the African malaria vector Anopheles funestus. *Mol Ecol* **14**, 4235-4248 (2005).
13. T. Y. Lee, W. C. Chang, J. B. Hsu, T. H. Chang, D. M. Shien, GPMiner: an integrated system for mining combinatorial cis-regulatory elements in mammalian gene group. *BMC Genomics* **13 Suppl 1**, S3 (2012).

14. C. Bass, C. T. Zimmer, J. M. Riveron, C. S. Wilding, C. S. Wondji, M. Kaussmann, L. M. Field, M. S. Williamson, R. Nauen, Gene amplification and microsatellite polymorphism underlie a recent insect host shift. *Proc Natl Acad Sci U S A* **110**, 19460-19465 (2013).
15. C. S. Wondji, H. Irving, J. Morgan, N. F. Lobo, F. H. Collins, R. H. Hunt, M. Coetzee, J. Hemingway, H. Ranson, Two duplicated P450 genes are associated with pyrethroid resistance in Anopheles funestus, a major malaria vector. *Genome Res* **19**, 452-459 (2009).
16. S. S. Ibrahim, J. M. Riveron, R. Stott, H. Irving, C. S. Wondji, The cytochrome P450 CYP6P4 is responsible for the high pyrethroid resistance in knockdown resistance-free Anopheles arabiensis. *Insect Biochem Mol Biol* **68**, 23-32 (2016).
17. I. H. Ishak, J. M. Riveron, S. S. Ibrahim, R. Stott, J. Longbottom, H. Irving, C. S. Wondji, The Cytochrome P450 gene CYP6P12 confers pyrethroid resistance in kdr-free Malaysian populations of the dengue vector Aedes albopictus. *Sci Rep* **6**, 24707 (2016).
18. K. Itokawa, O. Komagata, S. Kasai, M. Masada, T. Tomita, Cis-acting mutation and duplication: History of molecular evolution in a P450 haplotype responsible for insecticide resistance in Culex quinquefasciatus. *Insect Biochem Mol Biol* **41**, 503-512 (2011).
19. J. M. Riveron, S. S. Ibrahim, C. Mulamba, R. Djouaka, H. Irving, M. J. Wondji, I. H. Ishak, C. S. Wondji, Genome-Wide Transcription and Functional Analyses Reveal Heterogeneous Molecular Mechanisms Driving Pyrethroids Resistance in the Major Malaria Vector Anopheles funestus Across Africa. *G3 (Bethesda)* **7**, 1819-1832 (2017).
20. R. M. Kwiatkowska, N. Platt, R. Poupardin, H. Irving, R. K. Dabire, S. Mitchell, C. M. Jones, A. Diabate, H. Ranson, C. S. Wondji, 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**, 98-106 (2013).
21. C. The Anopheles gambiae Genomes, Genetic diversity of the African malaria vector Anopheles gambiae. *Nature* **552**, 96 (2017).
22. V. A. Ingham, P. Pignatelli, J. D. Moore, S. Wagstaff, H. Ranson, The transcription factor Maf-S regulates metabolic resistance to insecticides in the malaria vector Anopheles gambiae. *BMC Genomics* **18**, 669 (2017).
23. M. Kalsi, S. R. Palli, Cap n collar transcription factor regulates multiple genes coding for proteins involved in insecticide detoxification in the red flour beetle, Tribolium castaneum. *Insect Biochem Mol Biol* **90**, 43-52 (2017).
24. D. Martinez-Torres, F. Chandre, M. S. Williamson, F. Darriet, J. B. Berge, A. L. Devonshire, P. Guillet, N. Pasteur, D. Pauron, Molecular characterization of pyrethroid knockdown resistance (kdr) in the major malaria vector Anopheles gambiae s.s. *Insect Mol Biol* **7**, 179-184 (1998).
25. F. Faucon, I. Dusfour, T. Gaude, V. Navratil, F. Boyer, F. Chandre, P. Sirisopa, K. Thanispong, W. Juntarajumnong, R. Poupardin, T. Chareonviriyaphap, R. Girod, V. Corbel, S. Reynaud, J. P. David, Identifying genomic changes associated with insecticide resistance in the dengue mosquito Aedes aegypti by deep targeted sequencing. *Genome Res* **25**, 1347-1359 (2015).
26. C. S. Wondji, R. K. Dabire, Z. Tukur, H. Irving, R. Djouaka, J. C. Morgan, Identification and distribution of a GABA receptor mutation conferring dieldrin resistance in the malaria vector Anopheles funestus in Africa. *Insect Biochem Mol Biol* **41**, 484-491 (2011).

27. K. G. Barnes, H. Irving, M. Chiumia, T. Mzilahowa, M. Coleman, J. Hemingway, C. S. Wondji, Restriction to gene flow is associated with changes in the molecular basis of pyrethroid resistance in the malaria vector Anopheles funestus. *Proc Natl Acad Sci U S A* **114**, 286-291 (2017).
28. J. M. Riveron, F. Watsenga, H. Irving, S. R. Irish, C. S. Wondji, High Plasmodium Infection Rate and Reduced Bed Net Efficacy in Multiple Insecticide-Resistant Malaria Vectors in Kinshasa, Democratic Republic of Congo. *The Journal of infectious diseases* **217**, 320-328 (2018).
29. C. Strode, S. Donegan, P. Garner, A. A. Enayati, J. Hemingway, 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**, e1001619 (2014).
30. R. H. Hunt, B. D. Brooke, C. Pillay, L. L. Koekemoer, M. Coetzee, Laboratory selection for and characteristics of pyrethroid resistance in the malaria vector Anopheles funestus. *Med Vet Entomol* **19**, 271-275 (2005).
31. J. M. Riveron, M. Osae, A. Egyir-Yawson, H. Irving, S. S. Ibrahim, C. S. Wondji, Multiple insecticide resistance in the major malaria vector Anopheles funestus in southern Ghana: implications for malaria control. *Parasit Vectors* **9**, 504 (2016).
32. C. Mulamba, J. M. Riveron, S. S. Ibrahim, H. Irving, K. G. Barnes, L. G. Mukwaya, J. Birungi, C. S. Wondji, Widespread pyrethroid and DDT resistance in the major malaria vector Anopheles funestus in East Africa is driven by metabolic resistance mechanisms. *PLoS One* **9**, e110058 (2014).
33. J. M. Riveron, M. Chiumia, B. D. Menze, K. G. Barnes, H. Irving, S. S. Ibrahim, G. D. Weedall, T. Mzilahowa, C. S. Wondji, Rise of multiple insecticide resistance in Anopheles funestus in Malawi: a major concern for malaria vector control. *Malar J* **14**, 344 (2015).
34. J. C. Morgan, H. Irving, L. M. Okedi, A. Steven, C. S. Wondji, Pyrethroid resistance in an Anopheles funestus population from Uganda. *PLoS One* **5**, e11872 (2010).
35. M. T. Gillies, M. Coetzee, *A supplement to the Anophelinae of Africa south of the Sahara (Afrotropical region)*. (South African Institute for medical research, Johannesburg, 1987), vol. 55, pp. 143.
36. G. D. Weedall, H. Irving, M. A. Hughes, C. S. Wondji, Molecular tools for studying the major malaria vector Anopheles funestus: improving the utility of the genome using a comparative poly(A) and Ribo-Zero RNAseq analysis. *BMC Genomics* **16**, 931 (2015).
37. J. D. Thompson, D. G. Higgins, T. J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680 (1994).
38. P. Librado, J. Rozas, DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452 (2009).
39. S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* **33**, 1870-1874 (2016).
40. M. Clement, D. Posada, K. A. Crandall, TCS: a computer program to estimate gene genealogies. *Mol Ecol* **9**, 1657-1659 (2000).
41. WHO, Test procedures for insecticide resistance monitoring in malaria vector mosquitoes. *World Health Organization*, (2013).

42. S. S. Ibrahim, J. M. Riveron, J. Bibby, H. Irving, C. Yunta, M. J. Paine, C. S. Wondji, Allelic Variation of Cytochrome P450s Drives Resistance to Bednet Insecticides in a Major Malaria Vector. *PLoS Genet* **11**, e1005618 (2015).

Acknowledgments: The authors are grateful to Prof Janet Hemingway and Prof Hilary Ranson for helpful comments on this manuscript. This work was supported by a Wellcome Trust Senior Research Fellowship in Biomedical Sciences to Charles S. Wondji (101893/Z/13/Z). Genomic data generation was carried out by the Centre for Genomic Research, University of Liverpool. The samples of PermaNet 2.0 and 3.0 insecticide-treated nets tested in exprimental hut trials were freely offered by Vestergaard (Helen Jamet).

Author contributions: CSW conceived and designed the study; JMR, RD, SSI and CSW performed the field collection and resistance bioassays; HI, GDW, JMR and CSW performed sample preparation for all next-generation sequencing; GDW and CSW analyzed Next-generation sequencing data. MJW, JMR and MaT performed qRT-PCR; LMJM, GDW and CSW characterized *CYP6P9a* promoter; NAA and CSW analyzed the genetic diversity of *CYP6P9a*; LMJM and CSW designed the PCR-RFLP diagnostic assay; LMJM and MiT generated the lab crosses and performed the validation of the PCR-RFLP; BDM performed the experimental hut experiments with CSW and genotyped *CYP6P9a* with MJW and MaT; GDW and CSW wrote the paper with assistance from LMJM and JMR; All authors read and approved the final draft of the manuscript.

Competing interests: The authors declare no competing financial interests.

Data and materials availability: RNAseq: PRJEB24351, PRJEB10294; **Field PoolSeq**: PRJEB24384 and PRJEB13485; ***CYP6P9a* sequences**: GenBank MG782573-MG782841.

Table 1: Experimental hut results with the FANG/FUMOZ strain

	Untreated Nets	PermaNet 2.0	PermaNet 3.0
Total Mosquitoes	356	270	322
% Exophily (CI)	11.8 (8.45-15.15)	16.7 (12.2-21.1) [ns]	15.8 (11.85-19.8) [ns]
% Blood fed (CI)	29.5 (24.7-34.2)	14.8 (10.6-19.05) [§]	6.8 (4.1-9.6) [§ǂ]
% Blood feeding inhibition	-	49.8[§]	76.84[§ǂ]
% Personal protection (Total blood fed)	-(105)	61.9 (40) [§]	79.04 (22) [§ǂ]
% Blood feeding Mortality (no dead blood fed))	1(1)	40.0 (16) §	95.5 (21) [§ǂ]
% Mortality corrected (CI)		33.3 (27.7-38.9) [§]	98.7 (97.5-99.9) [§ǂ]

For each comparison, estimates not sharing the same symbols (§ or ǂ) are statistically different at $P<0.05$

Figure legends

Figure 1: Transcriptional profiling: A) Venn-diagram showing number of differentially up-regulated genes between different countries relative to FANG at FDR<0.01 and Fold-change >2. B) Volcano plots of differential gene expression between permethrin-exposed samples and the susceptible FANG highlighting differences in the expression of key resistance genes between countries. C) Heatmap showing the contrast in expression of major candidate detoxification genes between African regions. D) Fragments Per Kilobase of transcript per Million mapped reads (FPKM) of genes from *rp1* pyrethroid resistance QTL highlighting differences between regions. E) Correlation between RNAseq and qRT-PCR. The data shown are mean + SEM (n = 3).

Figure 2: Footprints of selective sweep associated with pyrethroid resistance. A) Contrasting polymorphism patterns between lab resistant (FUMOZ) and susceptible (FANG) strains and also between pre- (MWI-2002) and post-bednet intervention (MWI-2014) samples from Malawi (12). Data were aligned to 120kb *rp1* BAC sequence (IGV screenshot). Each track shows the alignment depth (on a log scale for display purposes) respectively (coverage depth is capped at >100x). Grey columns represent bases identical to the reference sequence while coloured columns indicate variant sites with a minor allele frequency >10%. The genes of the P450 cluster are highlighted at the bottom. An increase read coverage is observed between CYP6P9a and CYP6P9b in pyrethroid resistant samples indicating the 6.5kb insert. B) Major signature of selective sweep detected around *rp1* QTL pyrethroid resistance region across the 2R chromosome in southern Africa but not in FANG susceptible strain after plotting Minor allele frequency (MAF).

Figure 3: Genetic diversity of CYP6P9a 5'UTR region across Africa. A) Molecular Phylogenetic analysis of CYP6P9a 5'UTR region by Maximum Likelihood method; The evolutionary history of *CYP6P9a* promoter haplotypes across Africa was inferred by using the Maximum Likelihood method based on the Tamura 3-parameter model. B) Africa-wide TCS network for the *CYP6P9a* haplotypes showing four predominant regional haplotypes in southern Africa (STH10), West (Ghana, GHA11), Central (BEN/DRC21) and East/Central (EST/CNT 24). GHA, Ghana; CMR, Cameroon; MWI, Malawi; MOZ, Mozambique; ZMB, Zambia; TNZ, Tanzania; DRC, Democratic Republic of Congo; FNG, FANG. *, ancestral haplotype. Lines connecting haplotypes and each node represent a single mutation event (respective polymorphic positions are given on each branch). C) Neighbor-Joining phylogenetic tree of CYP6P9a-based genetic distance between ten African populations (N_{ST} estimates).

Figure 4: Impact of scale up of bednet interventions on the genetic diversity of promoter region of *CYP6P9a*. A) Maximum Likelihood phylogenetic tree of *CYP6P9a* showing a cluster of highly diverse haplotypes pre-intervention but a nearly fixed haplotype post-intervention. Pink represents haplotypes before insecticide-treated nets whereas light blue are those post-insecticide-treated nets. B) TCS haplotype network in Malawi and Mozambique Pre- and Post-intervention revealing a major resistant haplotype post-intervention but a very diverse set of haplotypes before pre-insecticide-treated nets.

Figure 5: *CYP6P9a* promoter analysis and design of a DNA-based diagnostic assay for P450-based metabolic pyrethroid resistance. A) Luciferase promoter assay (mean ± SD; n=6) of *CYP6P9a* 5' flanking region with progressive serial deletions to detect the causative variants. B) Agarose gel of TaqI PCR-RFLP of *CYP6P9a*, clearly distinguishing the three genotypes RR, RS and SS. C) Significant correlation (P<0.0001) between *CYP6P9a* resistance allele and permethrin resistance. D) Africa-wide distribution of the *CYP6P9a* resistant allele showing near fixation in southern Africa.

Figure 6: Impact of the *CYP6P9a*-based metabolic resistance on bednet efficacy: A) *CYP6P9a* genotypes correlate with pyrethroid resistance in the hybrid strain of resistant (FUMOZ) and susceptible (FANG) strains supporting the use of this FANG/FUMOZ strain to assess the impact of *CYP6P9a*-mediated resistance on bednet efficacy. B) Blood feeding and mortality rates (mean ± SD; n=4) of FANG/FUMOZ strain after release-recapture in experimental huts with untreated (blue), PermaNet 2.0 (green) and PermaNet 3.0 (blue) nets. C) *CYP6P9a* genotype proportions in dead or alive mosquitoes after PermaNet 2.0 exposure showing that CYP6P9a_R significantly allows mosquitoes to survive bed net exposure ($P<0.0001$). D) Genotype proportions in blood-fed and unfed mosquitoes after PermaNet 3.0 exposure showing that CYP6P9a_R allele increases the ability of resistant mosquitoes of taking a blood meal.

Supplementary Text

Title: A cytochrome P450 allele confers pyrethroid resistance in a major African malaria vector reducing insecticide-treated nets' efficacy

Authors: Gareth D. Weedall[1]*, Leon M.J. Mugenzi[2,3], Benjamin D. Menze[1,2,3], Magellan Tchouakui[2,3], Sulaiman S. Ibrahim[1,4], Nathalie Amvongo-Adjia [3,5], Helen Irving[1], Murielle J. Wondji[1,2,3], Micareme Tchoupo[2,3], Rousseau Djouaka[6], Jacob M. Riveron[1,2,3], Charles S. Wondji[1,2,3*]

Material and Methods

RNA extraction, sequence library preparation and sequencing

Total RNA was extracted from pools of 10 female mosquitoes (alive after 1h permethrin exposure) using the Arcturus PicoPure RNA isolation kit (Life Technologies), according to the manufacturer's instructions and including a DNase treatment step. For each pool, total RNA was rRNA-depleted with Ribo-Zero low input kit for Human/Mouse/Rat (Epicentre, Madison, WI, USA), using 100 ng of starting material. RNAseq libraries were prepared from Ribo-Zero mRNA-enriched material with the ScriptSeq v2 RNAseq library preparation kit (Epicentre), using 15 cycles of PCR amplification. Libraries were purified using Agencourt AMPure XP beads (Beckman and Coulter, Beverly, MA, USA). Each library was quantified using a Qubit fluorometer (Life Technologies, Carlsbad, CA, USA) and the size distribution assessed using the 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA).

Pools of libraries were sequenced, 8 per lane of the HiSeq 2500 (Illumina, San Diego, CA, USA) at 2x125 bp paired-end sequencing with v4 chemistry. Sequence library preparation and sequencing were done at the Centre for Genomic Research (CGR), University of Liverpool.

Analysis of RNAseq data

Initial processing and quality assessment of the sequence data was performed as follows. Basecalling and de-multiplexing of indexed reads was performed by CASAVA version 1.8.2 (Illumina) to produce samples from the pooled sequence data, in fastq format. The raw fastq files were trimmed to remove Illumina adapter sequences using Cutadapt version 1.2.1 (*1*). Reads were further trimmed to remove low quality bases (with a window quality score <20), using Sickle version 1.200 (*2*).

RNAseq R1/R2 read pairs were aligned to the reference sequence using the Subread aligner version 1.4.6 (*3*). The reference annotation available (AfunF1.4) was further improved using BLAST2GO version 4.0.7 (*4*).

Fragments mapped in the sense orientation to annotated *An. funestus* genes (automated predictions from gene set AfunF1.4, 2016-11-20, downloaded from VectorBase and annotated genes from the mitochondrial genome) were counted using featureCounts version 1.4.6 (*5*). Differential gene expression analysis was carried out using edgeR (*6*). Pairwise comparisons were carried out. Normalisation factors were calculated to correct for differences in total tag counts among samples, which may otherwise cause bias in differential gene expression analysis, using the "TMM" (Trimmed Mean M-values) method in edgeR (*6*) with default parameters. P-values associated with logFC were adjusted for multiple testing using the False Discovery Rate (FDR) approach (*7*). Significantly, differentially expressed genes were defined as those with an FDR-adjusted P-value <1% and >2-

fold absolute difference in expression level. Fragments per kilobase of gene sequence per million mapped reads (FPKM) was calculated using tag counts and the total CDS length. Gene ontology enrichment analysis was carried out on differentially expressed gene sets using BLAST2GO (*4*).

The Strand NGS software (Strand Life Sciences, version 3.0) was also used to analyze the data following RNA alignment and RNA-seq analysis pipeline with standard parameters.

Results

RNAseq transcriptional profiling of mosquitoes from across Africa identifies candidate pyrethroid-resistance associated genes

The annotation gene set AfunF1.4 includes 13,506 protein coding gene annotations. However, only 5,095 of these were functionally annotated. To improve the functional annotation of the genome, we used BLAST2GO (*4*) to assign descriptions and gene ontologies using sequence similarity. This analysis assigned putative descriptions to 12,196 protein coding genes and gene ontology descriptions to 10,072.

All read libraries were aligned to the reference genome. The quality metrics and the alignment parameters are presented in Table S1. Analysis with EdgeR and StrandNGS generated similar results. The number of differentially expressed genes between the four populations and the FANG susceptible strain or between populations is represented in the Venn diagram (Fig. 1A; Fig. S1A) whereas the expression profile is shown in the volcano plot for each country (Fig. 1B-E).

Gene Ontology (GO) enrichment: After quality control and analyses (Table S1), GO enrichment was performed to assess the generic metabolic terms associated with resistance. Permethrin resistant mosquitoes in Malawi showed significant enrichment of gene ontologies associated with cytochrome P450 genes in genes over-expressed relative to the fully susceptible FANG. These GO terms include heme binding, tetrapyrrole binding, oxidoreductase activity and iron ion binding (Fig. S2A). These GO terms among others are also enriched in the Ghana over-expressed gene set (Fig. S2B). In Uganda, more GO terms are enriched among the over-expressed genes but not directly associated with detoxification activities (Fig. S2C).

Candidate resistance genes commonly up-regulated Africa-wide:

To elucidate the continent-wide drivers of pyrethroid resistance in *An. funestus*, we first detected the sets of genes commonly over-expressed in all four regions. Cytochrome P450 genes were the most predominant detoxification genes on this list (Table S2). The two duplicated P450 genes *CYP6P9a* and *CYP6P9b* were the most

abundantly expressed notably in Malawi with fold change (FC) of 60.5x and 23.9x respectively. Although these two genes are also up-regulated in other regions of East, West and central Africa, it is with a much-reduced fold change with FC of 2.1, 6.3 and 2.7 only respectively in Cameroon, Ghana and Uganda for *CYP6P9a*. This suggests that this gene is more specific to the southern region (Fig. 1B-D). Besides these two genes, other P450s commonly up-regulated had lower FKPM <1000. Among these, the C*YP325A* gene presented a marked increase in Central Africa with FC of 26.9 in Cameroon in contrast to other regions where FC<6 suggesting that the over-expression of this gene could be more specific of this region. Two glutathione S-transferases *GSTe6* and *GSTD3* were also commonly up-regulated in all four regions at similar FC between 2.6 and 4.9.

Genes commonly up-regulated in three regions: Among the genes commonly expressed in three out of four regions, are two P450s. The duplicated *CYP6P4a* and *CYP6P4b* are highly over-expressed in Ghana with FC of 44.8 and 23.9 respectively but lower FC (<5.9) in Malawi and Uganda but not up-regulated in Cameroon. The other P450, *CYP6P5*, also located on the *rp1* pyrethroid resistance QTL region is over-expressed in all regions but not in southern Africa (Fig. 1D). Other P450s differentially expressed have lower FC including *CYP4C36* (not over-expressed in Uganda), *CYP306A1* (not in Ghana), *CYP315A1* (not in Malawi). Strikingly a set of GSTs from the epsilon class where up-regulated in all regions except East Africa including *GSTe2* previously shown to confer DDT and pyrethroid resistance, which nevertheless exhibits higher FC in Cameroon and Ghana than Malawi (Table S2). These epsilon GSTs also include *GSTe1*, *GSTe3* and *GSTe5*. The *GSTD1* presented a high FPKM>10k in Malawi, Ghana and Uganda suggesting a role for this gene in these regions but not in Cameroon where it is not significantly over-expressed.

Genes commonly up-regulated only in two regions: Analysis of the sets of genes up-regulated only in two regions revealed the presence of key P450s. Among these are *CYP6M7* and *CYP9J11* both with abundant reads counts, up-regulated only in southern and West Africa with FC around 2 and 3. These genes have previously been shown to metabolise pyrethroids (*8*, *9*). Another P450, *CYP6N1* located on the *rp2* QTL chromosome, is also up-regulated in southern and West Africa only. The *CYP9K1* P450 gene is up-regulated in East and West Africa although with higher FC in East (FC5.2). Among the genes only up-regulated in West and Central Africa, are a carboxylesterase AFUN002514 with FC of 5.5 and 3.6 respectively in Cameroon and Ghana but also the glutathione S-transferase *GSTe4*. Among the list of genes only present in Central and Southern Africa are also a carboxylesterase (AFUN000422) and a GST, AFUN007291 (*GSTt2*). Other detoxification genes are up-regulated only in one regions and usually have a relatively low FC (Table S2).

970 **Quantitative RT-PCR**: To validate the RNAseq transcription profile, the expression levels of fifteen
971 detoxification genes was also assessed by qRT-PCR for both permethrin resistant (Fig. S3A) and unexposed (Fig.
972 S3B) mosquitoes. These genes included 12 P450s differentially expressed across the different regions as well as
973 one GST, one aldehyde oxidase and one carboxylesterase. Primers were previously published (*8, 10*). Firstly,
974 control mosquitoes not exposed to insecticides also showed a strong correlation (R^2=0.695; P=0.002 in Malawi)
975 with permethrin resistant samples used for RNAseq (Fig. 1E) supporting a constitutive expression of these
976 candidate resistance genes to confer resistance. Overall, a high and significant correlation was observed between
977 qRT-PCR and RNAseq results for the 4 countries when compared to FANG (R^2=0.85; P<0.001) (Fig. S3C).
978 However, significant differences were observed for the expression of some genes such as the *CYP6Z1* P450 which
979 for qRT-PCR exhibited a high up-regulation in Malawi (FC 66) and Ghana (FC13.4) but not in RNAseq. This
980 gene has previously also been shown to be over-expressed in southern Africa using microarray and qRT-PCR (*11*)
981 suggesting that RNAseq could have missed or that the primers was not efficient with the susceptible FANG strain.
982 Similarly, the aldehyde oxidase gene (AFUN000093) was significantly over-expressed with qRT-PCR in all
983 regions except Cameroon but not with RNAseq. This could be due to the poor annotation of some of these genes
984 in the current *An. funestus* genome with certain genes been mistakenly combined.
985
986 **Complex evolution of the gene cluster of the rp1 CYP6 genes associated with pyrethroid resistance**
987 Due to the strong evidence of selection at *rp1* across southern Africa from RNAseq profiles, a detailed
988 analysis of this major resistance locus was performed revealing that a large scale structural polymorphism may
989 affect the evolution of the *rp1* locus:
990 **Identification of a 6.5 kb insertion between *CYP6P9a* and *CYP6P9b***: Close inspection of the pooled-
991 template whole genome alignment to the 120 kb BAC sequence containing the CYP6 cluster showed two
992 anomalous features in the 8.2 kb sequence between the paralogous genes *CYP6P9a* and *CYP6P9b*. In some
993 samples, the coverage depth was greater than for the surrounding sequence and some samples showed read pairs
994 in the correct relative orientation (><) but with greater than expected insert sizes, indicative of a large indel: a
995 "deletion" in the sequenced genome(s) or an "insertion" in the reference genome (Supplementary Figure 4A; Table
996 4). The FUMOZ alignment contains reads that are left-clipped (the leftmost part of the read, as aligned to the
997 reference, is clipped irrespective of the read's orientation: so the 5' end of a read aligned to the positive strand and
998 the 3' end of a read aligned to the negative strand) between BAC sequence positions 37409 and 37410 (37410
999 being the leftmost base included in the insertion). It also contains reads that are right-clipped between positions

43954 and 43955. This defines a region of 6545 bp. The presence of only left-clipped reads on the left of the region and right-clipped reads on the right of the region indicates two things in FUMOZ: (i) that the "insertion" form of the indel is fixed in the FUMOZ sample (there is no evidence for presence of the "deletion" form), and (ii) that the inserted sequence is homologous to part of a larger sequence found elsewhere in the genome (indicated by the "clipped" parts of the reads).

In FANG the situation is more complicated. At the left end of the insertion there are reads left-clipped between positions 37409 and 37410 (as for FUMOZ) but also some reads right-clipped slightly further left, between positions 37404 and 37405. At the right end of the insertion there are reads right-clipped between positions 43954 and 43955 (as for FUMOZ) but also some reads left-clipped between the same positions. In addition, further to the right there are some reads right-clipped between positions 44053 and 44054 and some left-clipped between positions 44070 and 44071. Detailed inspection of the clipped reads showed that the reads right-clipped at 37404/37405 and left-clipped at 43954/43955 indicate the "deletion" form of the indel, as the clipped parts of the reads from the left and right end of the insertion overlap each other (but also contain a short length of DNA that does not match FUMOZ). The clipping at 44053/44054 and 44070/44071 is due to a region of 35 bp in FANG (TAA TAC CGG GAG ATA CAT GGA GCT CGT GTA AAA GA) that does not align with the FUMOZ reference (ATA TGT CGG AGG TTT AT) at the same location. Overall, FANG shows evidence of the "deletion" form of the indel in addition to the presence of the large homologous sequence elsewhere in the genome. This makes simple inspection of the alignment misleading, as rather than a loss of coverage across the 6.5 kb indel, coverage is seen due to reads originating from sequence elsewhere in the genome.

Identifying the presence or absence of the insertion in samples from different geographical locations in Africa: To determine the geographical extent of the FUMOZ-like "insertion" haplotype, a PCR amplification of the 8.2kb was performed with two possible fragments obtained; a 1.7kb in case of absence of the 6.5kb insert or the 8.2kb in presence of the insert. The results indicate that the 6.5 kb insertion between *CYP6P9a* and *CYP6P9b* was present only in southern Africa population of Malawi, where it was nearly fixed (only a single read in Malawi supported the deletion haplotype) (Supplementary Table 5). However, populations from other parts of Africa showed no evidence of the insertion haplotype. Evidence that the insertion existed (albeit at low frequency) in the early 2000s comes from its presence in the FUMOZ colony, which was colonized form the field in Mozambique in 2000, and subsequently selected for insecticide resistance, which appears to have fixed the insertion haplotype in colony.

Investigating the genomic origin of the 6.5kb insertion: To identify the genomic origin of the inserted 6.5 kb sequence, it was extracted from the BAC sequence and used to search the *An. funestus* FUMOZ AfunF1 reference genome assembly using BLASTn implemented in the VectorBase web resource. The results indicated that the sequence occurred at two different locations in the genome, both of which were on scaffold KB669169. KB669169 is 1,771,395 bp long and contains the CYP6 cluster between positions 1,340,840 (the 3' end of *CYP6AA1*) and 1,431,230 (the 5' end of *CYP6AD1*). The first location was from 1,378,027 to either 1,403,987 or 1,410,005. This is the "inserted" location, between *CYP6P9a* and *CYP6P9b*. The two different right-hand positions may be due to the poor quality of the assembly across the CYP6 cluster, with maybe gaps in the assembly in this region, and the possible inclusion of the same sequence twice when the contigs should have been merged. The second location was between 1,109,272 and 1,118,666, approximately 260 kb away from the CYP6 cluster on the same scaffold (therefore, on the same chromosome). In addition, short (100 bp) sequences from the left and right ends of the insertion were used to conduct BLASTn searches of AfunF1 and confirmed the results obtained with the full-length insertion sequence. Finally, clipped sequences from immediately to the left and right of the insertion were used to conduct BLASTn searches of AfunF1. The results (matches only adjacent to the KB669169:1109272-1118666 region) confirmed that the "parent" sequence of the insertion between *CYP6P9a* and *CYP6P9b* came from KB669169:1109272-1118666. This putative genomic "parent" sequence of the insert contains no annotated protein coding genes but there is a large assembly gap in the region. The orthologous region in the *Anopheles gambiae* genome is on chromosome arm 2R. The protein coding genes flanking the insertion sequence, AFUN008344 and AFUN008346, are orthologous to AGAP002842 and AGAP002845, respectively. *An. gambiae* has no annotated protein coding genes between AGAP002842 and AGAP002845, suggesting that *An. funestus* may not have also. Three micro-RNAs annotated in both species are outside of the insertion sequence. One of these (mir-317; AFUN015669) has its 5' end approximately 130 bp away from the leftmost breakpoint of the insertion. Despite the lack of annotated genes, the region is transcribed and shows a large transcribed region, with some evidence of splicing, covering the three annotated micro-RNAs. Whether this transcript is processed to form mature micro-RNAs is not known.

Detection of a molecular marker associated with pyrethroid resistance

i-Comparative analysis of the *cis*-regulatory region of *CYP6P9a* between susceptible and resistant mosquitoes: To have a full view of the potential regulatory elements driving over-expression of *CYP6P9a*, we amplified and sequenced the full 8.2kb intergenic region between *CYP6P9a* and *CYP6P9b* in individual resistant (FUMOZ_R) and susceptible (FANG) mosquitoes. Amplifications revealed that while resistant mosquitoes have

the full 8.2 kb region, the susceptible ones only present a 1.7kb size for this intergenic region confirming the insertion of a 6.5 kb fragment in resistant mosquitoes. This insertion is present in the lab resistant FUMOZ strain (reference genome) as well as all southern Africa to near fixation (Malawi, Mozambique and Zambia). However, mosquitoes from other parts of the continent were similar to the lab susceptible strain FANG. To understand why such insertion occurred in resistant mosquitoes in southern Africa, we analyzed the composition of the 6.5kb fragment which occur at 830bp from stop codon of *CYP6P9b* and at 905 from start codon of *CYP6P9a*. Using GPminer (*12*), we detected, that it is full of transcription factor sites including an cpg island of 1.3kb and several GATA sites. It also contains several TATA (35), CCAAT (12) and GC (11) sequences. Furthermore, this 6.5kb is rich in over-represented (OR) oligonucleotides. Using the Alggen Promo program, to search for key transcription factors associated with regulation of genes involved in xenobiotics detoxification, revealed an abundant number of binding sites (51) for the Cap n Collar C (CnCC) and the Muscle aponeurosis fibromatosis (Maf) transcription factors sites which are known xenobiotic sensors in insects. Other of these sensors were also detected in the 6.5 kb including 14 binding sites for Ahr:Arnt, 13 for SXR:PXR and also for HNF4. The richness of this 6.5kb in regulatory factors suggests that this insertion contributes to increase regulation of *CYP6P9a* gene. Noticeably this 6.5kb contains a microsatellite (FUNR) (*13*), 388bp in size and located between 6082bp and 6482bp, only 80bp from the 5'UTR of *CYP6P9a*. Previous genotyping of this marker Africa-wide revealed significant differences associated with pyrethroid resistance profile. This FUNR is not present within the 1.7kb intergenic region between *CYP6P9a* and *CYP6P9b* for the FANG susceptible strain or before bednets distribution. It has been shown that microsatellite loci are involved in upregulation of P450 to confer insecticide resistance in other insects such as *Aphid* (*14*). It has also been shown in Yeast that polymorphic tandem repeats in the promoter regions can activate gene expression by impacting local chromatin structure to act as "evolutionary tuning knobs" to drive rapid evolution of gene expression such in a case of insecticide resistance selection. However, the FUNR has been detected in mosquitoes lacking this 6.5kb across Africa suggesting that for those populations' portions of this 6.5 kb are located in other regions of the genome close to 2R chromosome. There is also the possibility of the 6.5kb in resistant mosquitoes is causing a chromatin remodeling thereby promoting the access of the transcription factors to the promoter (*15*).

Geographical distribution of the resistant *CYP6P9a* allele across Africa: To establish the geographical distribution of the resistant *CYP6P9a* allele, field populations of *An. funestus* collected from several countries in Africa were genotyped using PCR-RFLP. This revealed that the CYP6P9a-R allele is mainly present in southern Africa where it is close to fixation in Mozambique, Malawi and Zambia. The allele is also present in

Tanzania in East Africa, although at lower frequency than in southern Africa (55.7%). However, in correlation with sequencing data, the CYP6P9a-R is completely absent in Central and West Africa (Fig. 5D; Fig. S6C). This total absence of CYP6P9a-R in West-Central Africa supports the restriction to gene flow we are observing for this species maybe because of the Rift Valley. In DR Congo, a contrast was observed between East and West as the CYP6P9a-R mutation is present in East but absent in the West (Kinshasa). It is important to monitor the spread of this marker across DRC to assess the speed and direction of spread of this resistance allele.

CYP6P9a_R reduces the effectiveness of insecticide-treated bed nets

To assess the impact of the CYP6P9a-R haplotype on the effectiveness of LLINs, we opted to use lab strains as this mutation is nearly fixed in the field in southern Africa. We crossed the highly resistant laboratory strain FUMOZ-R (where CYP6P9a_R is fixed) with the fully susceptible laboratory strain FANG (where CYP6P9a_R is completely absent). Using reciprocal crosses between the two strains we generated a hybrid strain at the F_4 generation that we used for semi-field studies in experimental huts.

Susceptibility profiles of the FANG/FUMOZ and FANG/FUMOZ strains: The bioassays performed with the reciprocal FANG/FUMOZ strains revealed that both hybrid strains were resistant to pyrethroids and carbamates and moderately resistant to DDT (93% mortality) (Fig. S7A). As expected, the level of resistance was lower than in the fully resistant strain FUMOZ_R, with a mortality rate of 76.1-80.7% when exposed to permethrin. However, a significant difference was observed for deltamethrin with a higher mortality rate recorded for the strain generated from crossing females FUMOZ_R to males FANG (48.5%) than in the strain from females FANG and males FUMOZ_R (77.3%). This difference could indicate the role of some candidate genes in the X chromosome for deltamethrin resistance (*CYP9K1*, for instance). Resistance pattern was similar for the carbamate bendiocarb in both reciprocal strains.

Validating the role of CYP6P9a-R in pyrethroid resistance in the hybrid FANG/FUMOZ strains: Before any field studies with the hybrid FANG/FUMOZ strains, the role of the CYP6P9a_R allele in the observed pyrethroid resistance was confirmed. WHO bioassays showed a mortality of 39.0% and 42.3% after 30 minutes' exposure and mortality rates of 81.3% and 86.3% after 90 minutes' exposure, respectively to permethrin and deltamethrin (Fig. S7B). The odds ratio of surviving exposure to permethrin when homozygous for the resistant CYP6P9a_R allele (RR) was high at 693 (CI 88-5421; $P<0.0001$) compared to the homozygous susceptible (SS) (Fig. S7C-D). The OR was 131 (CI 27-978; $P<0.0001$) when comparing RR to RS indicating that the resistance conferred by *CYP6P9a* is additive.

References

1. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10-12 (2011).
2. N. A. Joshi, J. N. Fass, Joshi NA, Fass JN. (2011). Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files *Available at https://github.com/najoshi/sickle.* **(Version 1.33) [Software].** , (2011).
3. Y. Liao, G. K. Smyth, W. Shi, The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res* **41**, e108 (2013).
4. A. Conesa, S. Gotz, J. M. Garcia-Gomez, J. Terol, M. Talon, M. Robles, Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674-3676 (2005).
5. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
6. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
7. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing. . *Journal of the Royal Statistical Society* **Series B** 289-300. (1985).
8. J. M. Riveron, S. S. Ibrahim, E. Chanda, T. Mzilahowa, N. Cuamba, H. Irving, K. G. Barnes, M. Ndula, C. S. Wondji, 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**, 817 (2014).
9. J. M. Riveron, S. S. Ibrahim, C. Mulamba, R. Djouaka, H. Irving, M. J. Wondji, I. H. Ishak, C. S. Wondji, Genome-Wide Transcription and Functional Analyses Reveal Heterogeneous Molecular Mechanisms Driving Pyrethroids Resistance in the Major Malaria Vector Anopheles funestus Across Africa. *G3 (Bethesda)* **7**, 1819-1832 (2017).
10. J. M. Riveron, H. Irving, M. Ndula, K. G. Barnes, S. S. Ibrahim, M. J. Paine, C. S. Wondji, 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**, 252-257 (2013).
11. S. S. Ibrahim, M. Ndula, J. M. Riveron, H. Irving, C. S. Wondji, The P450 CYP6Z1 confers carbamate/pyrethroid cross-resistance in a major African malaria vector beside a novel carbamate-insensitive N485I acetylcholinesterase-1 mutation. *Mol Ecol* **25**, 3436-3452 (2016).
12. T. Y. Lee, W. C. Chang, J. B. Hsu, T. H. Chang, D. M. Shien, GPMiner: an integrated system for mining combinatorial cis-regulatory elements in mammalian gene group. *BMC Genomics* **13 Suppl 1**, S3 (2012).
13. A. Cohuet, F. Simard, A. Berthomieu, M. Raymond, D. Fontenille, M. Weill, Isolation and characterization of microsatellite DNA markers in the malaria vector Anopheles funestus. *Molecular Ecology Notes* **2**, 498-500 (2002).
14. C. Bass, C. T. Zimmer, J. M. Riveron, C. S. Wilding, C. S. Wondji, M. Kaussmann, L. M. Field, M. S. Williamson, R. Nauen, Gene amplification and microsatellite polymorphism underlie a recent insect host shift. *Proc Natl Acad Sci U S A* **110**, 19460-19465 (2013).

15. M. D. Vinces, M. Legendre, M. Caldara, M. Hagihara, K. J. Verstrepen, Unstable tandem repeats in promoters confer transcriptional evolvability. *Science* **324**, 1213-1216 (2009).

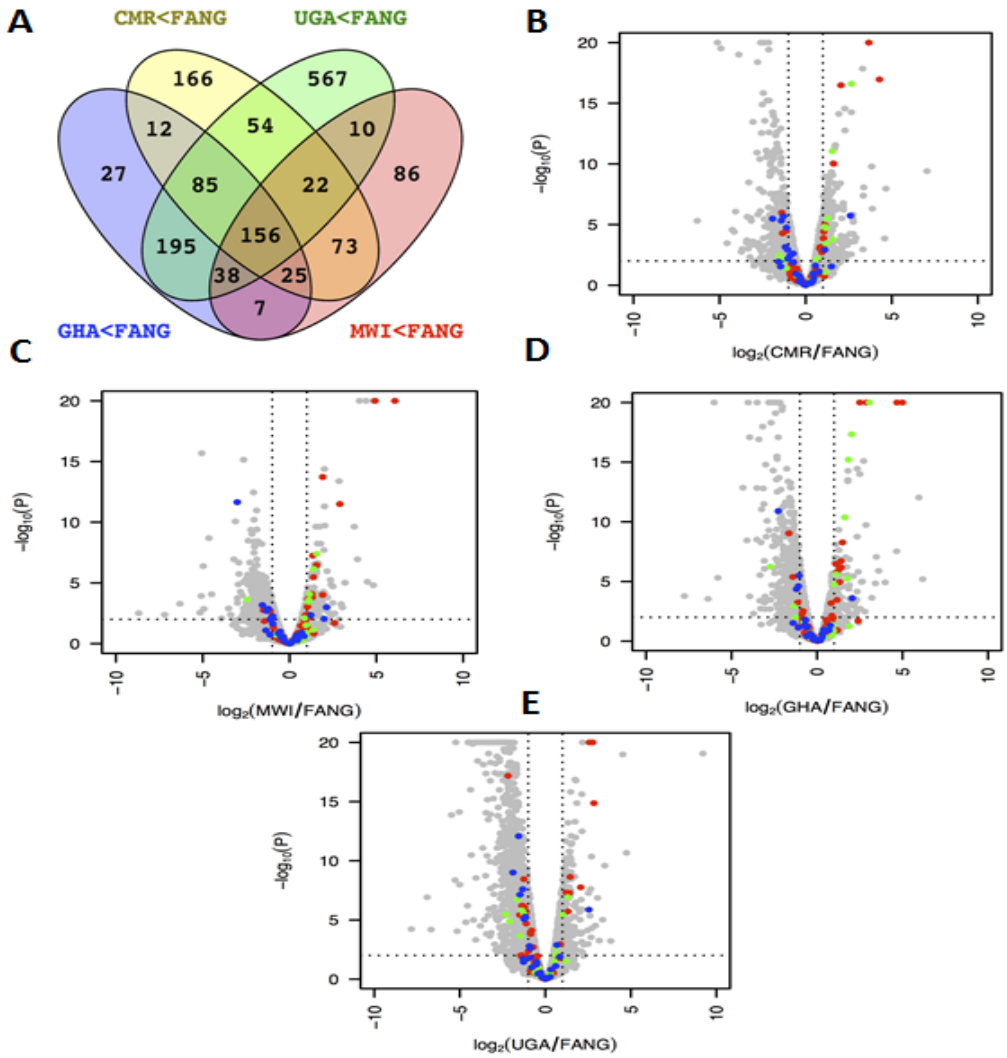

Fig. S1. Differential gene expression between four permethrin-exposed samples and FANG. A) Venn-diagram showing number of differentially downregulated genes between different countries relative to FANG at FDR<0.05 and Fold-change > 2. B) Volcano plot of the expression of between CMR and FANG susceptible strain. The X-axis shows $\log_2$ fold-change (positive values are up-regulated relative to FANG). The Y-axis shows $-\log_{10}$ transformed P values (adjusted for multiple testing; values greater than 40 were displayed as 40). The horizontal dashed line marks P=1% and the vertical dashed lines indicate two-fold expression difference among conditions. Red points indicate genes annotated as cytochrome P450s and green points indicate GST and blue points represent carboxylesterases. C) is for Malawi, D) for Ghana and E) is for Uganda.

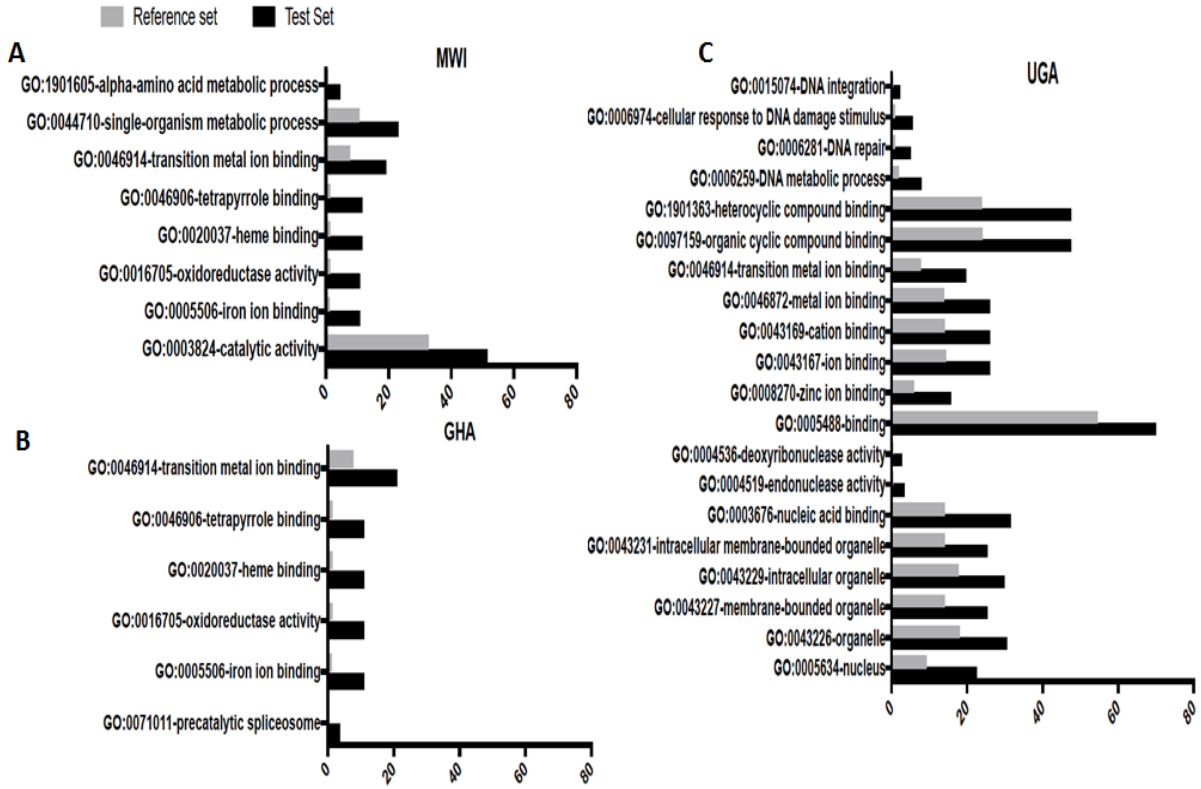

Fig. S2: Gene ontology enrichment of upregulated genes using BLAST2GO: A) Malawi, B) Uganda, C) Ghana. The test set represents the transcripts up-regulated while the reference set is made of the entire *An. funestus* transcript set (Afun1.4). Significance of the enrichment was assessed using a Benjamini and Hochberg multiple testing correction ($P < 0.05$). No significant terms detected for Cameroon at FDR of 0.

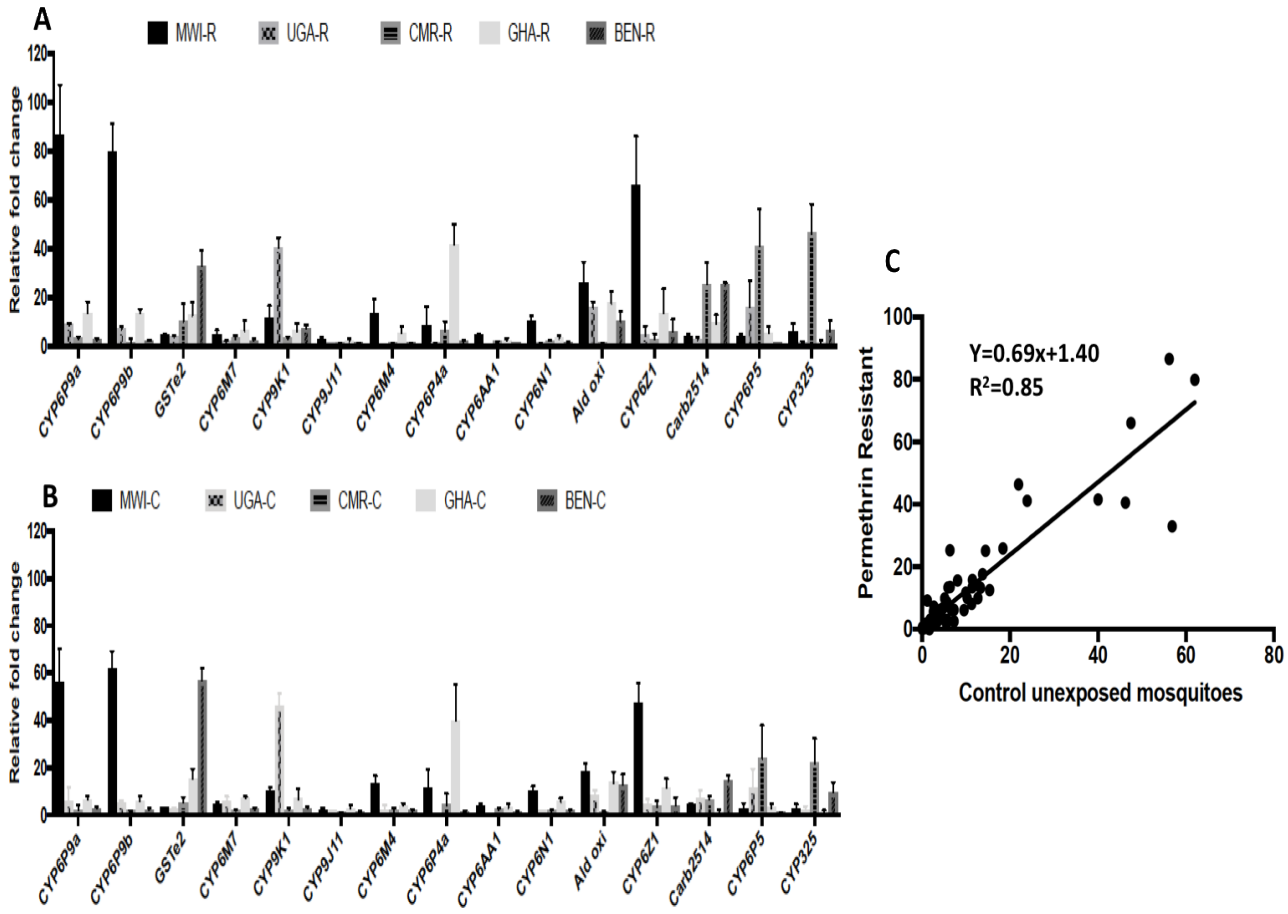

Fig. S3: qRT-PCR validation of the expression profile of the main detoxification genes differentially expressed between resistant and susceptible pyrethroid samples with RNAseq. A) qRT-PCR expression profile between R-S; B) qRT-PCR expression profile between C-S; C)Correlation between qRT-PCR of R_S and C-S. The data shown are mean + SD (n = 3).

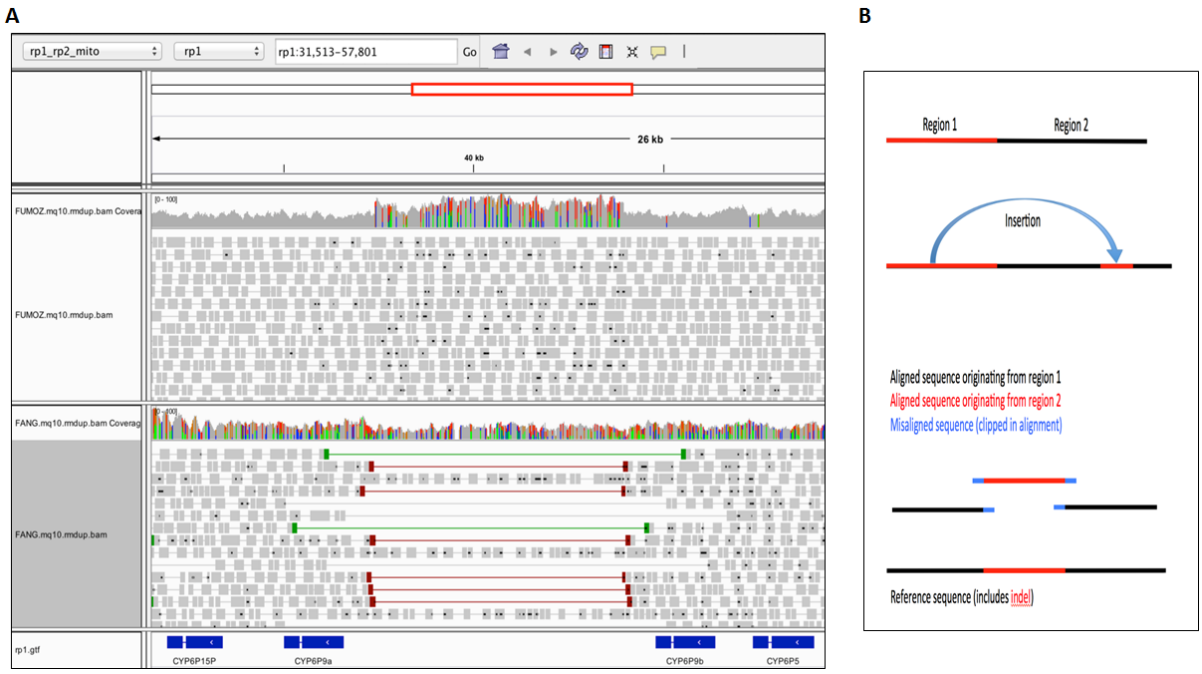

Fig. S4: Insertion of a 6.5kb intergenic fragment between *CYP6P9a* and *CYP6P9b* in southern African mosquitoes. A) Screenshot from the integrative genomics viewer (IGV), showing coverage depth and aligned reads for FUMOZ (upper) and FANG (lower) pooled template whole genome sequence alignments. The coverage depth plots show deeper coverage in this region in FUMOZ but not in FANG. The FANG alignment contains read pairs with unusually long insert sizes, indicated in red in the lower panel (thick lines represent reads, read pairs are linked by thin lines). B) Schematic representation of the process of insertion from a one region of the genome to another, generating 2 homologous sequences in different parts of the genome.

Fig. S5: Genetic diversity patterns of an 800bp cis-regulatory genomic fragment of *CYP6P9a* Africa-wide and before and after LLIN scale up. (A) Polymorphic sites and haplotypes across Africa. Haplotypes are labeled with prefixes from the country where they are predominant. CMR, Cameroon, GHA is Ghana, BEN is Benin, TNZ is Tanzania, DRC is Democratic Republic of Congo, MOZ is Mozambique, STH is Southern Africa, MWI is Malawi, ZMB is Zambia, EST/CNT is East/Central Africa. B) Polymorphic sites and haplotypes before and after scale up of LLINs in Mozambique and Malawi. Haplotypes are labeled with prefixes from the sample where they are predominant. PreMWI is for Pre-intervention Malawi, PreMoz is Pre-intervention Mozambique; PostMWI is Post-intervention Malawi; PostMoz is for Post-Intervention Mozambique.

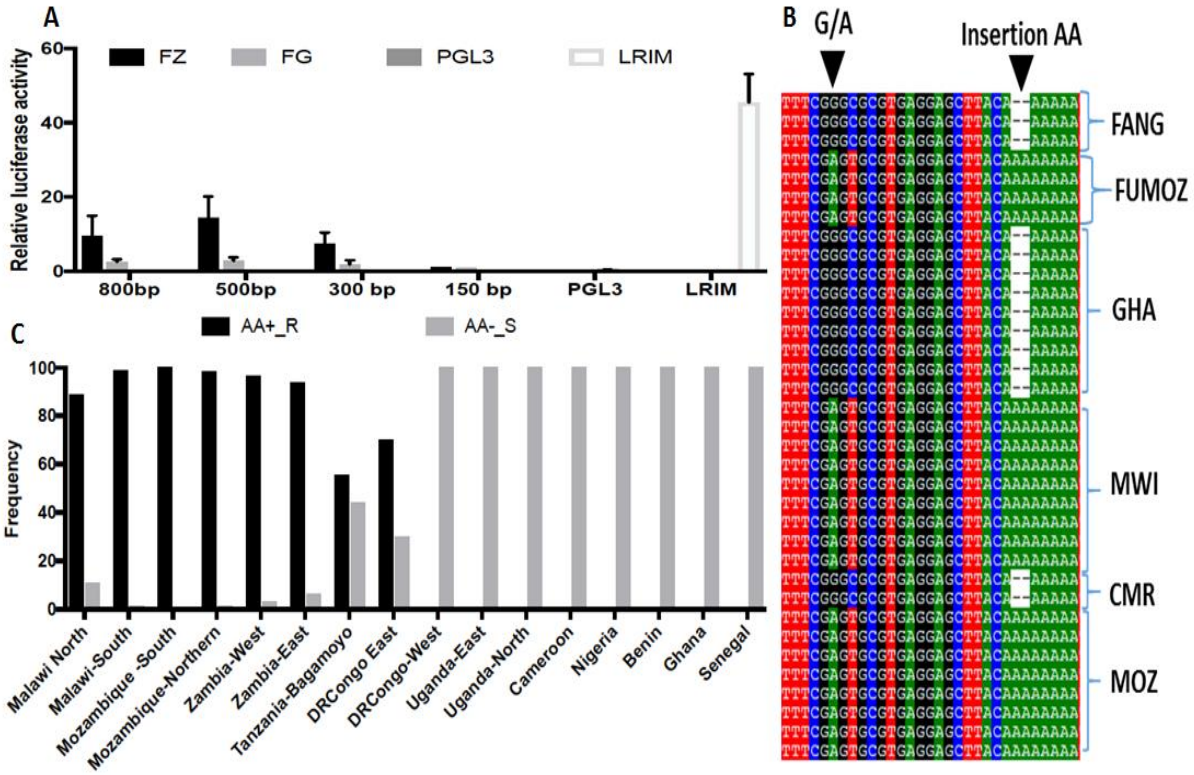

Fig. S6: Design of a DNA-based diagnostic assay to detect and track pyrethroid resistance across Africa. A) Comparative luciferase assay between promoter fragment from the highly resistant FUMOZ and highly susceptible (FANG) lab strains with progressive serial deletions of *CYP6P9a* 5' flanking region to detect the causative variants. Bars represent the mean ± S.D. of four independent transfections of three replicates (n = 6). A significant difference between promoter constructs ($p < 0.001$, Tukeys's t-test) is indicated by an asterisk. B) Schematic alignment of sequences across Africa showing fixed differences associated with pyrethroid resistance including the AA insert only found in resistant mosquitoes and the A/G variant tightly linked to the resistant haplotype and generating a restriction site for the TaqI restriction enzyme. C) Frequency of the CYP6P9a_R allele across Africa showing that it is only present in southern Africa where it is close to fixation but also in East Africa at moderate level. However, it is completely absent in Central and West Africa.

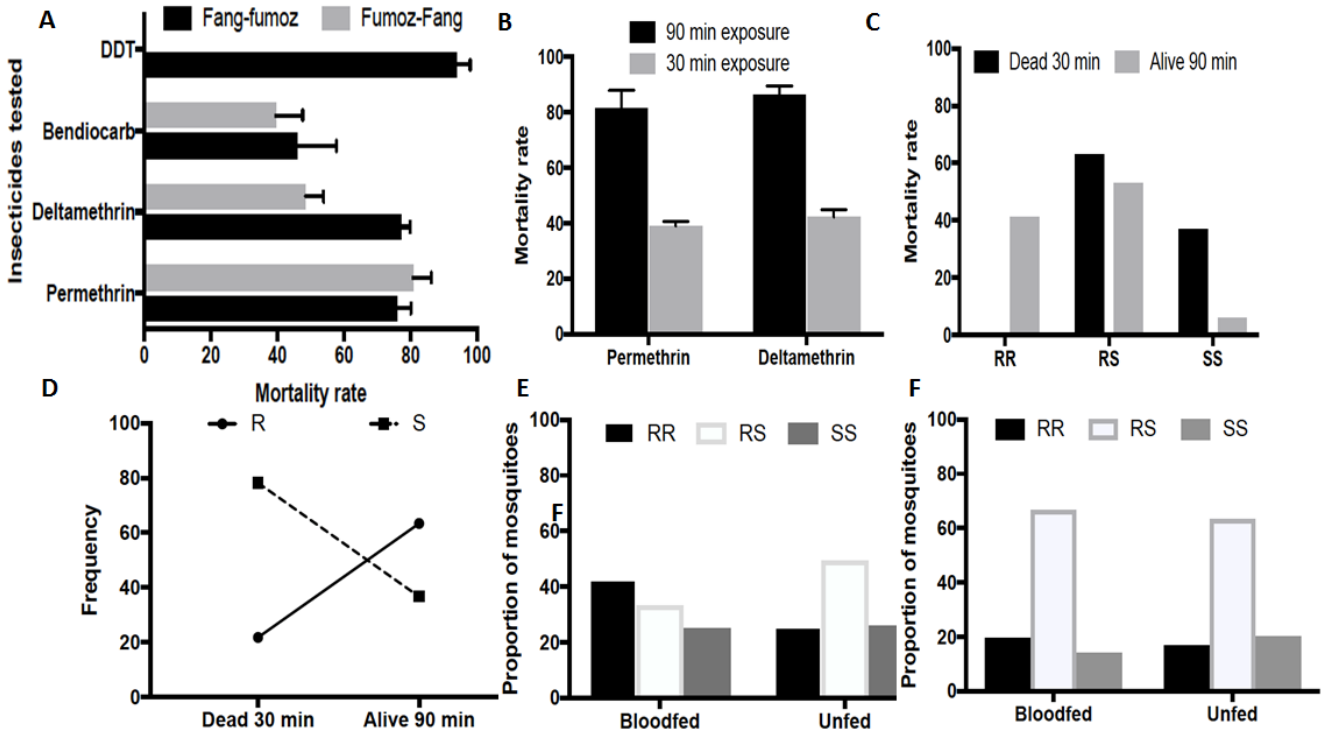

Fig. S7: Impact of the *CYP6P9a*-based metabolic resistance on the efficacy of bed nets using semi-field experimental hut trials: A) Susceptibility profile of the hybrid strain generated from crossing the highly resistant (FUMOZ) and highly susceptible (FANG) strains. Perm is permethrin, Delt is deltamethrin and Ben is bendiocarb. B) Mortality rates of the FANG/FUMOZ hybrid strain at two time-points exposure to validate the correlation between *CYP6P9a* and resistance phenotypes to pyrethroids. C) Distribution of the *CYP6P9a* genotypes according to resistance phenotypes. D) Strong correlation between *CYP6P9a* alleles and survival to PermaNet 2.0 exposure in experimental huts. E) CYP6P9a-R allele increases the ability of resistant mosquitoes of taking a blood meal in contrast to susceptible ones when exposed to PermaNet 2.0 but not for the untreated nets F).

Table S1. Descriptive statistics of RNAseq sequence read data and alignments for different samples

Sample ID	**Untrimmed reads**	**Trimmed reads**	**R1/R2 pairs** [1]	**R0 reads (%)** [2]	Reads to align (R1+R2)	Aligned reads (%) 3	Aligned R1 (%) 4	Aligned R2 (%) 4	Aligned in pair (%) 4	Properly paired (%) 4,5
FNG-UNX-000A-01	61,161,962	60,814,246	30,239,816	334,614 (0.55%)	60,479,632	34,529,287 (57%)	17,175,707 (50%)	17,353,580 (50%)	32,558,814 (94%)	30,926,294 (90%)
FNG-UNX-000A-02	44,384,316	43,991,605	21,866,954	257,697 (0.59%)	43,733,908	38,784,564 (89%)	19,455,342 (50%)	19,329,222 (50%)	37,497,734 (97%)	35,662,300 (92%)
FNG-UNX-000A-03	55,329,478	55,003,414	27,346,953	309,508 (0.56%)	54,693,906	41,539,393 (76%)	20,695,859 (50%)	20,843,534 (50%)	39,451,822 (95%)	37,642,472 (91%)
FNG-UNX-000A-07	61,708,692	60,522,621	30,147,658	227,305 (0.38%)	60,295,316	52,613,292 (87%)	26,304,258 (50%)	26,309,034 (50%)	51,355,756 (98%)	49,077,570 (93%)
GHA-PER-060A-01	56,393,812	56,004,973	27,858,160	288,653 (0.52%)	55,716,320	47,383,161 (85%)	23,661,891 (50%)	23,721,270 (50%)	45,701,704 (96%)	43,711,894 (92%)
GHA-PER-060A-02	56,146,206	55,778,128	27,745,363	287,402 (0.52%)	55,490,726	47,385,531 (85%)	23,689,592 (50%)	23,695,939 (50%)	45,668,698 (96%)	43,627,620 (92%)
GHA-PER-060A-03	46,339,060	45,931,585	22,853,504	224,577 (0.49%)	45,707,008	35,745,010 (78%)	17,841,177 (50%)	17,903,833 (50%)	34,393,520 (96%)	32,935,172 (92%)
CMR-PER-060A-01	45,564,004	45,336,332	22,563,410	209,512 (0.46%)	45,126,820	29,760,118 (66%)	14,809,654 (50%)	14,950,464 (50%)	28,136,026 (95%)	26,881,776 (90%)
CMR-PER-060A-02	49,616,994	49,369,073	24,565,403	238,267 (0.48%)	49,130,806	36,825,328 (75%)	18,551,507 (50%)	18,273,821 (50%)	34,658,598 (94%)	32,960,464 (90%)
CMR-PER-060A-03	56,448,220	55,238,405	27,521,162	196,081 (0.35%)	55,042,324	44,353,985 (81%)	22,394,524 (50%)	21,959,461 (50%)	42,605,426 (96%)	40,679,444 (92%)
UGA-PER-060A-02	63,577,506	62,233,533	30,965,317	302,899 (0.49%)	61,930,634	49,080,456 (79%)	24,430,318 (50%)	24,650,138 (50%)	47,225,124 (96%)	45,294,140 (92%)
UGA-PER-060A-03	74,281,024	72,571,899	36,050,602	470,695 (0.65%)	72,101,204	58,908,769 (82%)	29,401,087 (50%)	29,507,682 (50%)	56,385,738 (96%)	53,742,224 (91%)
UGA-PER-060A-04	57,574,274	55,614,929	27,652,993	308,943 (0.56%)	55,305,986	43,887,550 (79%)	21,868,261 (50%)	22,019,289 (50%)	42,252,224 (96%)	40,393,696 (92%)
MWI-PER-060A-06	72,542,372	72,016,218	35,777,777	460,664 (0.64%)	71,555,554	62,665,469 (88%)	31,430,727 (50%)	31,234,742 (50%)	60,517,686 (97%)	57,445,876 (92%)
MWI-PER-060A-07	65,422,634	64,501,521	32,092,002	317,517 (0.49%)	64,184,004	56,244,764 (88%)	28,121,863 (50%)	28,122,901 (50%)	54,471,126 (97%)	52,005,820 (92%)

[1] Forward (R1) and reverse (R2) read pairs after trimming.
[2] Reads unpaired after trimming (% of total trimmed reads).
[3] % of reads to align.
[4] % of aligned reads.
[5] Properly paired means both read and its mate are mapped to opposing strands of the reference sequence, with 3' ends innermost and 5' ends within the allowed distance from each other (50-600 bp).

Table S2: Detoxification-associated genes differentially expressed between the four pyrethroid resistant populations and the FANG susceptible strain

Gene ID	CMR	GHA	MWI	UGA	Description	FG_S	CMR	GHA	MAL	UG
AFUN010543	6.6	5.4	30.7	48.9	CCChain Structural Basis	0.28	0.98	2.8	7.1	8.9
AFUN011806	4.2	2.1	13.2	6	chymotrypsin-like elastase family member 2A	0.17	0.65	0.4	3.3	1.2
AFUN001444	14	6.2	4.4	3.2	chymotrypsin-like elastase family member 2A	4.5	56.8	31.5	28.9	16.3
AFUN015830	2.4	2.6	3.9	2.2	Cytochrome P450, CYP325C	1.43	3.1	4.2	8.15	3.5
AFUN015966	26.9	6	5.1	2.2	Cytochrome P450, CYP325A	0.74	17.4	4.8	5.2	1.8
AFUN015894	3.1	2.3	7.5	3.7	Cytochrome P450, CYP4H26	0.3	0.9	0.85	3.6	1.4
AFUN015792	2.1	6.3	60.5	2.7	Cytochrome P450, CYP6P9a	14.9	28.7	106.2	1355	45.6
AFUN015889	4.4	6.9	23.9	7	Cytochrome P450, CYP6P9b	11.5	45.3	88.8	401.7	92.2
AFUN015777	2.6	2.3	5.3	3.5	Cytochrome P450, CYP4C26	0.5	1.1	1.2	3.6	1.9
AFUN015839	2.6	4.9	3.3	3	glutathione S-transferase, GSTD3	25.3	60.2	137.8	122.5	87.1
AFUN016008	4.8	3.5	4.1	3.4	glutathione S-transferase, GSTE6	8.3	36.3	32.7	50.2	32.3
AFUN004582	2.8	3.7	2.8	3.1	methyltransferase 2-A	0.7	1.8	2.8	2.8	2.4
AFUN010812	4.9	2.9	9.6	3.1	General odorant-binding 45	0.4	1.2	0.9	4	1
AFUN015887	2.6	2.3	2.2	2.5	Gustatory receptor 68a	7.5	17.4	19	23.9	21.8
AFUN006831	3.6	2.2	3.7	4.4	General odorant-binding	0.4	0.8	0.6	1.3	1.2
AFUN007097	2.8	8.2	4.6	4	cuticular_protein_RR	0.5	0.9	3.1	2.3	1.6
AFUN005949	3.3	4.6	6.2	2.7	General odorant-binding 70	0.3	1	1.7	3	1.1
AFUN004398	5.5	4.4	7.5	7.5	cuticular_protein_TWDL_family_(TWDL12)	0.7	1.1	1	2.3	1.8
AFUN015818	6	3.6	11.8	4.1	Gustatory receptor for sugar taste 64b	0.4	1.1	0.8	3.5	0.9
AFUN001299	3.5	3	2.6	3.7	NTF2-related export 2	6.9	21.5	23.4	26.5	29
AFUN003480	2.4	2.7	3.7	2.7	serine protease 27-like	0.9	2	2.8	5	2.9
AFUN010184	2.2	2.4	2	3.3	transcription factor GATA-3-like isoform X1	1.3	2.6	3.5	3.9	5
AFUN007410	2.9	2.5	2.9	2.1	transmembrane protease serine 9-like	110.5	288.2	304	461.2	269.1
AFUN007052	5.1	3.5	6	4.2	zinc finger 345-like isoform X1	1.1	1.8	1.5	3.3	1.8
AFUN009311	2.3	2.3	2.4	2.3	zinc finger 391-like isoform X1	11.5	23.4	29.1	40.9	30.6
AFUN010612	3.2	3	2.4	3.9	zinc finger 883-like isoform X2	1.7	4.9	5.7	6.1	7.8
AFUN015890		44.8	5.9	2.1	Cytochrome P450, CYP6P4	2.3	3.6	114.2	19.7	5.5
AFUN015891		23.9	3.7	2.7	cytochrome P450 CYP6P4-like	17.1	16.2	456.2	91.8	52.1
AFUN016010		2.2	2	2.2	glutathione S-transferase, GSTD1	117.8	186.4	292	348.5	297.9
AFUN000474	3.2		4.7	2.1	Gustatory receptor for sugar taste 43a	1.5	4.2	2.7	9.8	3.6
AFUN000064	3.1		4.7	3.1	Cuticular Protein as AGAP010105-PA	0.7	2.1	1.6	5	2.7

AFUN010577	2.1		2.9	2.2	chymotrypsin-like elastase family member 2A	1.5	2.9	2.7	6.5	3.8
AFUN006858	2.3		2.1	2.1	cytochrome P450 CYP306A1	8.2	17.4	17.8	24.9	19.8
AFUN006563	2.7		3.9	3.3	Cuticular protein	0.4	0.11	0.078	0.26	0.18
AFUN015888	6.3	5.8		4.1	Cytochrome P450, CYP6P5	6.3	36.2	41.1	8.5	29.6
AFUN005715	2.2	2.3		2.3	Cytochrome P450, CYP315A1	10	19.8	25.5	28.3	26.7
AFUN003202	2.2	2.1		2.1	nucleoporin Nup43	6.8	13.7	15.8	18.8	16.2
AFUN006467	2.2	2.7		3.2	zinc finger 391-like isoform X1	7.2	14.5	21.4	17.9	26.9
AFUN002969	2.6	2		2.4	zinc transporter ZIP10-like	10.6	24.8	24	22.9	29.3
AFUN006135	2.2	2.9	2.6		Cytochrome P450, CYP4C36	8.8	17.5	29	33.3	14.5
AFUN015808	2.5	3.3	2.3		glutathione S-transferase GSTE3	63.9	143.7	238.6	210.8	124
AFUN015807	3.5	3.7	2.8		glutathione S-transferase GSTE1	5.2	16.5	21.7	20.9	9.3
AFUN015809	5.9	8.3	2.3		glutathione S-transferase GSTE2	42.7	227	398.4	143.4	78.2
AFUN015811	3	2.3	2.2		glutathione S-transferase GSTE5	28.3	76.9	73	89.4	49.4
AFUN011266	3.1	2.7	4.9		UDP-glucuronosyltransferase	2.3	6.5	6.9	16.3	4.3
AFUN015817	3.8	2.3	7.2		Gustatory receptor for sugar taste 64e	0.4	0.7	0.5	2.1	0.4
AFUN000622	2.5	2.1	2.9		solute carrier family 23 member 2	7.2	16.2	16.7	30.1	12
AFUN006863	2.1	2.1	2		zinc finger 345-like isoform X4	6.7	12.6	15.6	20	13.4
AFUN000799	2.2	2	2.3		zinc finger CCHC-type	12.7	25.7	28.7	43.4	27
AFUN015907			2.2	2.4	gastrula zinc finger -like isoform X1	8.4	14.9	18.6	26.7	23.6
AFUN010481			4.4	3.1	Cuticular protein	0.11	0.3	0.06	0.53	0.3
AFUN010539			2.3	2	cuticular_protein_RR-2_family_(CPR143)	2.1	3.2	4.6	7	4.9
AFUN004166			2.5	3.1	gustatory receptor 28b	2.8	4.8	5.8	10.3	10
AFUN008855			3.9	2.8	nuclear receptor subfamily 2 group C	0.9	0.5	0.8	4.8	2.7
AFUN005273			2	2.6	zinc finger 883-like isoform X2	3.7	5.1	6.4	11	11
AFUN001382		2.7	2.4		Cytochrome P450, CYP9J11	69.5	87.3	206.1	241.7	144
AFUN001383		2.4	3.1		Cytochrome P450, CYP9J11	20.5	23.5	54.2	93	36.4
AFUN015795		2.2	2.2		Cytochrome P450, CYP6M7	118.7	119.1	297.8	372	150.8
AFUN015767		3	2.9		glutathione S-transferase, GSTD11	1.8	2.3	6	7.6	2.2
AFUN012021		2.4	5.9		Cuticular potein	0.16	0.22	0.33	1.1	0.17
AFUN007247		2.1	2.3		Odorant binding 45	1.2	1.2	2.7	3.8	2.5
AFUN001415		2.1	2.2		polypeptide N-acetylgalactosaminyltransferase 11	7.7	12	17.9	24.3	12.9
AFUN010918		2.1	3.2		Cytochrome P450, CYP6N1	18.7	29.6	43.3	86.6	21.6

AFUN007143	2.5		2.2	alkaline phosphatase	18.7	42.2	35.1	59.7	22.9
AFUN009142	2.9		2.4	aminopeptidase Ey-like	20.2	52.4	32.2	70	16.8
AFUN003099	3.8		3.2	arginase	6.6	22.6	13.5	30.8	7.8
AFUN000422	2		2.1	Carboxylesterase	0.8	1.4	0.9	2.4	0.7
AFUN007080	2.2		3.7	caspase-9	5.7	11.1	11	30.8	8.3
AFUN014849	2		2	CCR4-NOT transcription complex subunit 11	5.8	10.5	10.1	17.2	10.6
AFUN008239	2.4		3.9	cytosolic sulfotransferase 3-like isoform X2	15.3	33.6	33.7	86.1	21.7
AFUN007291	2.4		2.1	glutathione S-transferase GSTT2?	13.8	30.2	28.4	42.6	20.6
AFUN015936	2.6		2.7	Gustatory receptor for sugar taste 64a	4.9	11.4	9.5	19	6.4
AFUN006160	2.8		2.5	oocyte zinc finger 6-like	4.1	10.5	8	15.2	9.4
AFUN015810	2.4	3.6		glutathione S-transferase, GSTE4	69.1	151.6	276.7	164	69.6
AFUN002514	5.5	3.6		Carboxylesterase	32.2	160.5	130.6	67.5	43.9
AFUN000001	3.7	2.1		probable chitinase 3	26	85.8	60.5	64.5	15.8
AFUN007549		2.9	5.2	cytochrome_P450 CYP9K1	35.5	36.5	115.9	102.8	212
AFUN002602	2.1			cytochrome b561 domain-containing 2	34.5	65.2	50.3	87.9	41.2
AFUN007526	2.3			cytochrome c oxidase assembly COX19	17.1	35.6	30.4	28.8	29.9
AFUN015723	2			Cytochrome P450, CYP6AH1	22.2	40.5	32	55.9	21.1
AFUN006288	2.6			glycine dehydrogenase (decarboxylating)	57.3	133.8	81.8	157.2	68.4
AFUN008852	2			Glycosyltransferase involved in cell wall bisynthesis	6.6	12	12.9	18.7	14.3
AFUN002311	2.1			monothiol Grx4 family	76.3	141.2	126.6	175.9	142.7
AFUN007415	2.3			7 kDa salivary gland allergen	153.3	313.8	305	276.8	104.5
AFUN004413	2.1			nucleoporin NDC1 isoform X1	12.9	25.1	22.9	32.7	23.4
AFUN008376	2.1			V A-type H+-transporting ATPase subunit B	245.5	473.3	298.5	400	484.2
AFUN008872	2.4			Zinc carboxypeptidase	191	412.8	367.9	356.6	289.7
AFUN009750		2		alkaline phosphatase	7.6	8.9	17.1	20.7	14.5
AFUN002978		2		Cytochrome P450, CYP314A1	5.3	8.3	12.2	15.2	10.8
AFUN008819		2		glutathione S-transferase, GSTMS3	8	4.8	18.3	17.1	6.4
AFUN001429			2.4	alkaline phosphatase	3.7	5.6	3.1	13	2.7
AFUN010814			2.5	Alpha beta hydrolase family	0.8	0.8	0.4	2.8	0.7
AFUN004002			2.4	argininosuccinate lyase	177.4	257	252.8	616.9	170.4
AFUN007079			2.9	caspase-9	18.5	33.3	29.2	78.8	21.8

AFUN009199	2.5		chitin synthase	19.8	27.3	42.3	71.3	34.7
AFUN014173	2.7		chymotrypsin-like elastase family member 2A	13.2	21.8	19.9	51.8	18.1
AFUN004870	2.7		chymotrypsin-like elastase family member 2A	3.5	3.9	7.6	13.5	6.5
AFUN015895	2.4		Cytochrome P450, CYP4H25	3.3	2.5	4.5	11.7	3.6
AFUN015785	2.2		Cytochrome P450, CYP6AA2	9.1	13.8	15.4	29.6	11
AFUN015909	2.3		cytochrome P450	3	4.4	5	9.7	6
AFUN015841	2.2		glutathione S-transferase GSTD4	4.4	2.3	6.5	13.9	3.8
AFUN015768	2.3		glutathione S-transferase, GSTD11	13.5	21.5	27.9	45	14.4
AFUN008560	2.6		glutathione S-transferase	8.4	11.3	17	31.6	18.2
AFUN002910	3.4		lipase member H-like	2.1	2.7	1.9	10.6	1.9
AFUN007911	2.1		Solute carrier family 46 member 3	40.4	60.9	82.1	120.7	58.6
AFUN009946	3.6		Chymotrypsin-elastase inhibitor ixodidin	0.1	0.1	0.1	0.4	0.1
AFUN007416	2.7		37 kDa salivary gland allergen	4.2	7.3	5.8	16.7	4.5
AFUN015951	2.7		Gustatory receptor 8a	1.2	1.2	0.8	4.9	1.5
AFUN010684	3.6		Cuticular protein RR	0.1	0.19	0.03	0.34	0.18
AFUN016007	3.1		Gustatory receptor for sugar taste 43a	1	1.5	0.7	4.4	1.3
AFUN015901	2.2		Odorant receptor Or2	2.6	1.9	1.5	8.2	2.3
AFUN008524	2.2		Odorant receptor 22c	2.1	2.2	1.1	6.8	1.5
AFUN015721	2.6		Odorant receptor 4	1.1	1.2	0.6	4	1
AFUN010920	5.5		cytochrome P450 partial, CYP6M1b	0.18	0.09	0.1	0.99	0.13
AFUN015933	3.2		gustatory receptor 2a	0.5	0.8	0.6	2.1	1
AFUN008427	2.1		Gustatory receptor for sugar taste 64a	3.8	6.4	5.8	11.8	3.1
AFUN000851	3.1		N-alpha-acetyltransferase 40	2.3	3.8	3.3	10.2	2.7
AFUN001353	2.4		odorant-binding OBP56d- partial	0.9	0.5	0.7	3.1	1
AFUN010884	2		oocyte zinc finger 6-like	6.2	10.8	12.3	18.5	14
AFUN000174	3.5		peritrophin-1 [Culex quinquefasciatus]	1.2	0.8	1.4	5.8	2.5
AFUN015801	2.5		cytochrome P450, CYP6P2	27.3	23	36.9	98.9	16.8
AFUN010874	2		zinc finger 2 homolog isoform X1	8.5	13.6	14.2	25.3	15.2
AFUN007113	2		zinc finger 2 homolog isoform X1	6.6	8.3	11	19.5	11
AFUN014520		2.1	CCR4-NOT transcription complex subunit 7 isoform X1	12.2	21.1	23.7	20.3	29.1
AFUN006799		45	lipase member H-like	0	0	0	0	15.8

AFUN011526	2.2	microtubule-associated RP EB family member 3 isoform X2	17.2	23.1	32.4	44.1	42.7
AFUN000877	2	OR94A_DROME ame: Full=Odorant receptor 94a	4.9	4.7	5.8	10.5	11.5
AFUN006418	4	cuticular_protein_RR-3_family_(CPR111)	0.7	1.2	0.8	1.3	3.3
AFUN011733	2.8	Carboxylesterase	13.6	23.9	23.4	37.7	44.4
AFUN005433	2.1	solute carrier family 35 member F5	11.4	11.3	23.5	22.7	27.3

Table S3. Descriptive statistics of Whole genome POOLseq sequence read data

Sample name	Untrimmed reads	Trimmed reads	R1/R2 pairs [1]	R0 reads (%) [2]
FUMOZ	98,333,448	97,312,008	48,166,137	979,734 (1.01%)
FANG	103,497,770	102,476,042	50,736,627	1,002,788 (0.98%)
MWI-PER-060-DEAD-A	122,744,088	121,558,527	60,196,626	1,165,275 (0.96%)

[1] Forward (R1) and reverse (R2) read pairs after trimming

[2] Reads unpaired after trimming (% of total trimmed reads)

Table S4: Counts of reads aligned at the left and right breakpoints of the 6.5 kb insertion supporting different haplotypes

Location/Colony	Year	Genomes	Insertion (left)	Insertion (right)	Deletion (left)	Deletion (right)	"Parental" (left)	"Parental" (right)
FUMOZ	n/a	38	**43**	**44**	0	0	22	16
FANG	n/a	40	0	0	24	30	15	23
Malawi	2014	40	**11**	**15**	0	1	9	12

"Insertion"=insertion present between *CYP6P9a* and b; "Deletion"=insertion absent between CYP6P9a and b;

"Parental"=read originating from elsewhere in the genome, from where the inserted sequence was derived.

Table S5: Population genetic parameters of the 800bp fragment upstream of *CYP6P9a*

Total	**N**	**S**	**h**	**hd**	**π**	**k**	**D**	**D***
Malawi Pre-bednet	24	61	19	0.97	0.027	18.26	0.13ns	0.14ns
Mozambique Pre-bednet	15	25	12	0.96	0.012	8.36	0.36ns	-072ns
Total Pre-bednet	39	74	30	0.98	0.025	17.21	-0.28ns	-1.04ns
Malawi Post-bednet	18	4	2	0.11	0.00066	0.44	-1.85*	-2.5*
Mozambique Post-bednet	34	3	4	0.365	0.0008	0.53	-0.6ns	0.93ns
Total Post-bednet	52	7	5	0.283	0.00077	0.52	-1.74ns	-1.90ns
Total all	91	76	34	0.72	0.019	13.1	-0.62	-1.94

N= number of sequences (2n); S, number of polymorphic sites; h, number of haplotypes; hd, haplotype diversity; π, nucleotide diversity (k= mean number of nucleotide differences); D and D* Tajima's and Fu and Li's statistics; ns, not significant; * significant P<0.05.

Table S6: Correlation between genotypes of *CYP6P9a* and mortality (PermaNet 2.0) and blood feeding after the experimental hut trial with the FANG/FUMOZ strain

		OR	**P**	**CI**
		Mortality		
Unfed	RR vs SS	34.9	<0.0001	15.8-77.1
	RS vs SS	19.9	<0.0001	9.7-40.9
	RR vs RS	1.75	0.26	0.81-3.8
	R vs S	6.25	<0.0001	3.3-11.7
All samples	RR vs SS	10.82	<0.0001	5.6-20.8
	RS vs SS	5.3	<0.0001	2.8-9.8
	RR vs RS	2.04	0.0002	1.1-3.7
	R vs S	3.17	0.02	1.78-5.65
		Blood feeding		
PermaNet 2.0	RR vs SS	1.75	0.19	0.82-3.7
	RR vs RS	2.5	0.052	1.09-5.75
	RS vs SS	0.7	0.67	0.28-1.7
	R vs S	1.43	0.26	0.82-2.5
PermaNet 3.0	RR vs SS	4.54	<0.0001	2.3-8.7
	RR vs RS	2.6	0.0012	1.43-4.7
	RS vs SS	1.74	0.17	0.87-3.47
	R vs S	2.14	0.18	1.17-3.19

Mortality rates were assessed only for PermaNet 2.0 because mortality levels were very high for PermaNet 3.0.

Table S7: Primers used for characterization of the promoter of *CYP6P9a*

Primer name	**sequence**
6P9a1F	TCCCGAAATACAGCCTTTCAG
6P9Ra/b	TACACTGCCGACACTACGAAG
6P9a5F	AGCGGAAGGGGTTTTTGTAG
6P9a5R	CTTCTGTGATGCCCCAAAAT
6P9a3.2F	CAATGCTGCTTTCCTTCACA
SacIFU-6P9a0.8	CGAGCTCGTCCCGAAATACAGCCTTTCAG
SacIFU-6P9a0.5F	CGAGCTCGATCCCTAACTATTAAAAGGCAAT
SacIFU-6P9a.03F	CGAGCTCGTGCAGGGAAAAGGAGGACAT
Sac1FU-6P9a0.15F	CGAGCTCGCACGCACACTGACATGATGT
MluIFU-6P9Ra/b	CGACGCGTCGTACACTGCCGACACTACGAAG
KpnIFA-6P9a0.8F	CGGGGTACCCCGTCCCGAAATACAGCCTTTCAG
KpnIFA -6P9a0.5F	CGGGGTACCCCG ATCCCTAACTATTAAAAGGCAAT
KpnIFA -6P9a0.3F	CGGGGTACCCCGTGCAGGGAAAAGGAGGACAT
KpnIFA -6P9a0.15F	CGGGGTACCCCGCACGCACACTGACATGATGT
HindIII -6P9Ra/b	CCCAAGCTTGGGCCGTACACTGCCGACACTACGAAG

15