Heredity Open
Heredity (2012) 109, 383–392

www.nature.com/hdy

ORIGINAL ARTICLE

Positional cloning of rp2 QTL associates the P450 genes *CYP6Z1*, *CYP6Z3* and *CYP6M7* with pyrethroid resistance in the malaria vector *Anopheles funestus*

H Irving[1], JM Riveron[1], SS Ibrahim[1], NF Lobo[2] and CS Wondji[1]

Pyrethroid resistance in *Anopheles funestus* is threatening malaria control in Africa. Elucidation of underlying resistance mechanisms is crucial to improve the success of future control programs. A positional cloning approach was used to identify genes conferring resistance in the uncharacterised *rp2* quantitative trait locus (QTL) previously detected in this vector using F6 advanced intercross lines (AIL). A 113 kb BAC clone spanning *rp2* was identified and sequenced revealing a cluster of 15 P450 genes and one salivary protein gene (*SG7-2*). Contrary to *A. gambiae*, *AfCYP6M1* is triplicated in *A. funestus*, while *AgCYP6Z2* orthologue is absent. Five hundred and sixty-five new single nucleotide polymorphisms (SNPs) were identified for genetic mapping from *rp2* P450s and other genes revealing high genetic polymorphisms with one SNP every 36 bp. A significant genotype/phenotype association was detected for *rp2* P450s but not for a cluster of cuticular protein genes previously associated with resistance in *A. gambiae*. QTL mapping using F6 AIL confirms the *rp2* QTL with an increase logarithm of odds score of 5. Multiplex gene expression profiling of 15 *P450*s and other genes around *rp2* followed by individual validation using qRT–PCR indicated a significant overexpression in the resistant FUMOZ-R strain of the P450s *AfCYP6Z1*, *AfCYP6Z3*, *AfCYP6M7* and the glutathione-s-transferase *GSTe2* with respective fold change of 11.2, 6.3, 5.5 and 2.8. Polymorphisms analysis of *AfCYP6Z1* and *AfCYP6Z3* identified amino acid changes potentially associated with resistance further indicating that these genes are controlling the pyrethroid resistance explained by the *rp2* QTL. The characterisation of this *rp2* QTL significantly improves our understanding of resistance mechanisms in *A. funestus*.
Heredity (2012) **109,** 383–392; doi:10.1038/hdy.2012.53; published online 5 September 2012

Keywords: *Anopheles funestus*; malaria; insecticide resistance; QTL mapping; P450s

BACKGROUND

Malaria control in Africa relies heavily on vector control through the use of insecticide treated nets, long-lasting insecticide nets and indoor residual spraying. However, resistance to the main insecticides such as pyrethroids is threatening the success of these control methods.

Anopheles funestus, a major vector of malaria throughout much of sub-Saharan Africa (Gillies and De Meillon 1968), is increasingly developing resistance to different classes of insecticides used in public health, such as pyrethroids, carbamates and DDT with the fear that this could disrupt control programs against this vector. Indeed, resistance to pyrethroids, DDT and carbamates has been detected in different regions of Africa, such as Southern Africa (Mozambique (Hargreaves *et al.*, 2000; Casimiro *et al.*, 2006; Cuamba *et al.*, 2010) and Malawi (Hunt *et al.*, 2010)), East Africa (Uganda (Morgan *et al.*, 2010)), West Africa (Ghana (Okoye *et al.*, 2008) and Benin (Djouaka *et al.*, 2011)) and Central Africa (Cameroon (Wondji *et al.*, 2011)). To improve the management of these resistances, it is imperative to thoroughly characterise the underlying mechanisms in order to design appropriate control strategies.

Efforts to characterise resistance mechanisms in *A. funestus* have benefited from recent progress made in the study of this species notably the colonisation of two strains, one resistant to pyrethroids named FUMOZ-R originally from Mozambique and the other FANG fully susceptible to all insecticides and originally from Angola (Hunt *et al.*, 2005). Other progress that have facilitated genetic studies in *A. funestus* include the construction of a map (Sharakhov *et al.*, 2004), an integrated genetic and physical map (Wondji *et al.*, 2005), identification of a set of genome-widely distributed single nucleotide polymorphisms (SNPs) (Wondji *et al.*, 2007a) and recent sequencing of *A. funestus* transcriptome (Crawford *et al.*, 2010; Gregory *et al.*, 2011).

Using the resistant (FUMOZ-R) and the susceptible (FANG) laboratory strains, previous studies have identified a major quantitative trait locus (QTL) associated with pyrethroid, named *rp1* (for *resistance to permethrin*) located on chromosome 2R. This *rp1* QTL detected with F_2 mapping (Wondji *et al.*, 2007b) and also with F_6 and F_8 advanced intercross lines (AIL) (Wondji *et al.*, 2009) explains 87% of the genetic variance in pyrethroid susceptibility in two families from reciprocal crosses between susceptible and resistant strains. Two additional QTLs named *rp2* and *rp3* were also detected in chromosomes 2L and 3L, respectively. A positional cloning approach was used to identify the genes conferring pyrethroid resistance in *rp1* using AIL at F_6 and F_8 generations. This involved the sequencing of a 120-kb BAC clone spanning the *rp1* QTL, which identified fourteen

[1]Liverpool School of Tropical Medicine, Liverpool, UK and [2]The Eck Family Institute for Global Health & Infectious Diseases and Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, USA
Correspondence: Dr CS Wondji, Vector Group, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L5QA, UK.
E-mail: c.s.wondji@liverpool.ac.uk
Received 5 April 2012; revised 13 July 2012; accepted 17 July 2012; published online 5 September 2012

protein coding genes and one putative pseudogene (Wondji *et al.*, 2009). Ten of the fourteen genes encoded cytochrome P450s and expression analysis indicated that three of these P450s (*CYP6P9, AfCYP6P4* and *CYP6AA4*) were upregulated in the resistant strain. Furthermore, *CYP6P9* and *AfCYP6P4*, respectively, 25 and 51 times overexpressed in resistant females, were tandemly duplicated compared with *A. gambiae*.

If *rp1* QTL has been well characterised and the genes involved in pyrethroid resistance detected, this is not the case for the other two QTLs notably *rp2* the second most important QTL. Given the importance to fully characterise the mechanisms of pyrethroid resistance in this species, it is fundamental to also identify genes associated with pyrethroid resistance in *rp2* QTL. Recent observations that *CYP6P9* duplicated genes in *rp1* are also associated at various degrees in pyrethroid resistance in field populations of *A. funestus* in Africa (Morgan *et al.*, 2010; Djouaka *et al.*, 2011) indicates that further analysis of *rp2* and *rp3* QTLs in FUMOZ-R can help to better characterise the resistance in field populations. In *A. gambiae*, the other major malaria vector, genes associated with pyrethroid resistance such as *AgCYP6M2(*Djouaka *et al.*, 2008) and *AgCYP6Z1*(David *et al.*, 2005) are found in the chromosome 3R, which is the equivalent to chromosome 2L in *A. funestus* were *rp2 is* located. It remains to be established whether the orthologues of these genes could be associated with pyrethroid resistance in *A. funestus*.

Here, we report the positional cloning of genes in the *rp2* QTL associated with pyrethroid resistance in *A. funestus*, using the AIL approach with F_6 generation progeny from reciprocal crosses between the FUMOZ-R resistant and the FANG susceptible strains.

MATERIALS AND METHODS

Isolation and sequencing of the BAC clone containing the *rp2* QTL

An *A. funestus* BAC library from The Institute for Genomic Research, Notre Dame University, was screened by PCR using primers from nine P450 genes located within the boundaries of the *rp2* QTL using the synteny projection with the *A. gambiae* chromosomal map: these genes are *CYP6Z3, CYP6Z1, CYP6Y1, CYP6M8, CYP6S1, CYP6M1, CYP4H18, CYP6N2* and *CYP6M7*. The primers used are listed in Supplementary Table S1. DNA of whole 384-well plates was pooled and a PCR carried out for each plate. The positive plates were then subdivided into six column pools and 4 row pools and the PCR screen repeated. Finally, individual colonies from the set of 16 identified from the pooled column and row screen were used as template to identify the individual clone containing the markers of interest. The BAC clone was grown at 37°C overnight and harvested in a glycerol solution and stored at −80°C. The size of the BAC clone was estimated after a restriction digestion using the *Bam*H1 restriction enzyme to separate the insert from the vector. The BAC clone was then fragmented by sonication into shorter, random sequences of around 2–5 kb. These small fragments were cloned into the plasmid vector (pC31) and sequenced. The reads were trimmed using Seqman (http://www.dnastar.com/web/r13.php) and assembled using Phrap (P. Green unpublished; http://www.genome.washington.edu/UWGC/analysistools/Phrap.cfm).

Analysis of BAC clone sequence and annotation of the P450 genes

The programme Genemark.hmm version 2.2 was used to locate genes in the assembled sequence of the BAC clone. Putative genes were annotated by using BlastX (http://www.ncbi.nlm.nih.gov/blast/) and the predicted transcripts were also compared with *A. gambiae* transcripts in Vectorbase (www.vectorbase.org). Further detailed annotation of the P450 genes was aided by the P450 site (http://p450.sophia.inra.fr/). Sequence alignments of *A. funestus* and *A. gambiae* genes were carried out using ClustalW (Thompson *et al.*, 1994). Exact boundaries of the exon/intron were manually inspected and compared with that of *A. gambiae* using the DNAstar sequence analysis package. MEGA 4.0 (Tamura *et al.*, 2007) was used to construct a Neighbour-Joining tree of the *A. funestus* genes in comparison with *A. gambiae*.

Mosquito samples

Two AIL families at F_6 generations were used for this study. These families (Family 1 and Family 10) were generated from reciprocal crosses between the susceptible and the resistant strains as previously described (Wondji *et al.*, 2009). Briefly, a two exposure times approach of 30 min and 2 h was used to select the most susceptible and the most resistant mosquitoes, respectively, as described previously in order to minimise the level of phenotype misclassification and increases the power of QTL detection (Lander and Botstein, 1989).

SNP identification and genotyping in the reciprocal isofemale lines

In order to carry out a fine-scale mapping of the *rp2* QTL, SNPs were identified in all the genes detected in the BAC clone and in other genes spanning the *rp2* QTL boundaries in the 2L chromosome such as the glutathione-s-transferase *GSTe2*, the cuticular protein genes *CPLC5, CPLC9* and *CPLC8*, the P450s *CYP9M1, CYP303A1* and *CYP4H18*, and AGAP007980 and AGAP009073. The cuticular protein genes were amplified using primers listed in Supplementary Table S2. Genomic DNA was extracted using the LIVAK method (Livak, 1984) and amplified using the primers listed in Supplementary Table S3. For each family, the F_0 parental female was amplified with $3F_1$ in order to detect the informative SNPs to be used for genetic mapping. The PCR were carried out using 10 pmol of each primers and 30 ng of genomic DNA as template in 25 ml reactions. PCR products were purified using the QIaquick PCR purification kit (Qiagen, Hilden, Germany) and directly sequenced on both strands. Sequences for each gene were analysed to detect the polymorphic sites manually using BioEdit and as sequence differences in multiple alignments using ClustalW (Thompson *et al.*, 1994).

SNPs were identified as transitions or transversions in coding and non-coding regions. Genetic diversity analysis was performed using DnaSP 5.1 (Rozas *et al.*, 2003). The nucleotide diversity π, was calculated for each gene as well as the haplotype diversity. The average number of synonymous substitutions per synonymous site (Ks) and non-synonymous substitutions per non-synonymous site (Ka) was computed.

SNP genotyping. Equal numbers of surviving and dead F_6 progeny were genotyped for each of the two reciprocal families (75 dead and 75 alive for family 1, 48 dead and 48 alive for family 10) with the informative SNP using pyrosequencing method according to the manufacturer's instructions using the PSQ 96 SNP Reagent Kit (Biotage AB, Uppsala, Sweden) and as previously described (Wondji *et al.*, 2007b). Primers details are given in Supplementary Table S4.

QTL mapping. The JoinMap 2.0 package (Stam and Van Ooijen, 1995) was used to build a genetic linkage map for each of the two families using the same parameters as described previously (Wondji *et al.*, 2007b). Windows QTL Cartographer 2.5 (Wang *et al.*, 2005) software was used to plot the genetic map. Associations between genotypes at each locus and the resistance phenotype were assessed using a contingency χ^2-analysis. The null hypothesis was that susceptibility to permethrin is equal in each genotype class. For loci with a significant χ^2, we analysed the inheritance of the alleles at these loci. The *a priori* hypothesis was that a higher mortality rate would occur among F_6 individuals with one or both alleles inherited from the susceptible parent. The JoinMap linkage map and the genotype/phenotype data were entered into Windows QTL Cartographer 2.5 (Wang *et al.*, 2005). Interval mapping, composite-interval mapping and multiple-interval mapping procedures were implemented for each family as previously described (Wondji *et al.*, 2007b).

Multiplex GeXP expression profiling of *rp2* and 2L chromosome genes

The expression pattern of 28 genes located in the *rp2* BAC clone or in 2l chromosome was compared between the resistant strain FUMOZ-R and the susceptible strain FANG using the GeXP multiplex gene expression profiling method from Beckman Coulter as previously described (Wondji *et al.*, 2009). Total RNA was extracted from three batches of ten (1–3-day old) female and male adult mosquitoes using a PicoPure RNA isolation kit (Arcturus, Mountain View, CA, USA) according to manufacturer's instructions. Total RNA quantity and quality were assessed using Nanodrop spectrophotometer (Nanodrop Technologies, Oxfordshire, UK) and Bioanalyzer (Agilent, Santa

Clara, CA, USA), respectively. The GenomeLab eXpress Profiler programme was used for automated primer design, calculation of relative gene expression values, data checking and data analysis. The quantitative PCR reaction was carried out using the GenomeLab GeXP Start Kit (Beckman and Coulter, Brea, CA, USA) according to the protocol provided and using the primers listed in Supplementary Table S5. The amplified qPCR products were diluted 10 times and added in a 96-well microplate with the DNA size standard-400 and ran on the GenomeLab GeXP genetic analysis system. The expression level of the *RSP7* ribosomal gene was used to normalise for variation in total cDNA concentration. A two-sample t-test was used to compare the results between the two strains.

Expression analysis of potential candidate genes

Genes with a significant overexpression in the FUMOZ-R strain from the GeXP multiplex expression profiling was further analysis using individual quantitative reverse transcription PCR (qRT–PCR). To increase the specificity of the amplification, reverse primers were designed in the 3′UTR region for all the genes. Primers are listed in Supplementary Table S6. One microgram of total RNA from each of the three biological replicates for FUMOZ-R and FANG was used as template for cDNA synthesis using the Superscript III (Invitrogen, Carlsbad, CA, USA) with oligo-dT20 and RNase H, according to the manufacturer's instructions. A serial dilution of cDNA was used to establish standard curves for each gene in order to assess PCR efficiency and quantitative differences between samples. The qRT–PCR amplification was carried out in a MX 3005 real-time PCR system (Agilent) using Brilliant III Ultra-Fast SYBR Green QPCR Master Mix (Agilent). Ten nanogram of cDNA from each sample was used as template in a three-steps programme involving a denaturation at 95°C for 3 min followed by 40 cycles of 10 s at 95°C and 10 s at 60°C and a last step of 1 min at 95°C, 30 s at 55°C and 95°C at 30 s. The relative expression and fold change of each target gene in FUMOZ-R relative to FANG was calculated according to the $2^{-\Delta\Delta CT}$ method incorporating PCR efficiency after normalisation with the housekeeping *RSP7* ribosomal protein S7 (AGAP010592) and the actin 5C (AGAP000651) genes. A two-sample *t*-test was used to compare the results between the two strains.

Analysis of polymorphism patterns of candidate genes

In order to detect any potential polymorphism associated with resistance, the full length of the candidate genes detected after qRT–PCR was amplified using the cDNA synthesised from total RNA as template between the resistant and the susceptible strains. Amplification was performed with the Phusion polymerase with the following conditions: 1 cycle at 95 °C for 5 min; 35 cycles of 94 °C for 20 s, 57 °C for 30 s and elongation at 72 °C for 60 s; followed by 1 cycle at 72 °C for 5 min. The PCR products were purified using the QIaquick PCR purification kit (Qiagen) and cloned into the pJET1.2/blunt cloning vector using the CloneJET PCR cloning kit (Fermentas, Burlington, ON, Canada). Positive clones were purified and sequenced on both strands. Sequences were analysed as described above.

RESULTS

BAC clone isolation and sequencing

Nine positive 384-well plates containing one or more of the loci spanning *rp2* were identified from the *A. funestus* BAC library. A single individual BAC clone presented a positive PCR result for eight of the nine genes tested (*CYP4H18* PCR was negative). The estimation of the size of this clone indicated that it was above 100 kb and likely to span the *rp2* QTL.

BAC clone sequence analysis. A total of 1152 reads were generated from the sequencing of the BAC clone from which 91 were excluded either for poor base quality or contamination with vector sequence. The remaining 1061 reads (844 ± 141 bp) were assembled in six contigs of a final size of 113.59 kb with an average coverage of 6.57. The summary of these sequencing statistics is presented in Supplementary Table S7 and the full sequence is submitted to Genbank (accession number: JQ711183). An analysis of the BAC clone sequence for gene prediction indicated the presence of sixteen open reading frames. Annotation using BlastX revealed that fifteen of these open reading frames correspond to cytochrome P450 genes organised in a cluster as seen in *A. gambiae* on Chromosome 3R (equivalent to 2L chromosome in *A. funestus*), while one salivary protein named *gSG7-2* in *A. gambiae* was also identified. All the 15 P450 genes belong to the CYP6 family of cytochrome P450 genes. A direct comparison between this P450 cluster in *A. funestus* and that observed in *A. gambiae* reveals first that there is a triplication of the *CYP6M1* gene with three different copies in *A. funestus* compared with just one in *A. gambiae*. Second, the *AgCYP6Z2* gene present in *A. gambiae* is missing in *A. funestus*. The *AgCYP6Z2* gene is located between *AgCYP6Z1* and *AgCYP6Z3* in *A. gambiae* but no open reading frame was detected between these two genes in this *rp2* BAC clone. The P450 gene called *AfCYP6Z3* in this *rp2* clone exhibited a same similarity to both *AgCYP6Z3* and *AgCYP6Z2* in *A. gambiae* when compared by ClustalW alignment with 79% similarity for DNA sequence and 85% for the protein sequences. However, when using Blastx (NCBI), it indicates an 86% similarity of amino acid sequence to *AgCYP6Z3* in *A. gambiae* and 85% to *AgCYP6Z2*. Therefore, it is called *AfCYP6Z3* in this study, although it is very close to both genes. Third, only one salivary protein gene was detected in the *rp2* clone, while three of these *SG7* genes are found in this cluster in *A. gambiae*.

Analysis of the P450 gene cluster in *rp2* BAC clone

A complete genomic organisation of this P450 cluster is presented in Figure 1. The genomic arrangement of the fifteen P450 genes is exactly the same as in *A. gambiae*. The size of the intergenic spaces is also similar between the two species with the lowest intergenic space in this cluster being 78 bp between *AfCYP6S1* and *AfCYP6R1* (This is also the lowest in *A. gambiae* with 71bp), while the longest intergenic space is between *AfCYP6M4* and the salivary protein *gSG7-2* (24 739 bp), which is similar in *A. gambiae*.

All the 15 P450 genes possess two exons separated by short introns from 58 bp for *AfCYP6S1* to 87 bp for *AfCYP6M7*. These intron sizes are also similar to that of *A. gambiae* (Supplementary Table S8). The amino acid sequences of the fifteen P450 genes varied from 492 amino acids for *AfCYP6Z1* and *AfCYP6Z3* to 514 for *AfCYP6Y2*, similar to *A. gambiae* with the only difference that *AfCYP6Y2* is 520 amino acids long in this species. The highly conserved motifs C-helix WxxxR, ExLR and PERF are all present as indicated in the multiple alignment. Similarly, the haem-binding loop (the P450 signature sequence) PFxxGxRxCxG/A (Feyereisen, 2011) and the membrane targeting/anchor hydrophobic N-terminal region are present in all fifteen P450s (Supplementary Figure S1). The GxE/DTT/S is also present in all genes apart from *AfCYP6R1* as observed in *A. gambiae*. The orthology of the 15 P450 genes to the respective *gambiae* genes is confirmed in the Neighbour-joining tree in Supplementary Figure S2 showing that the *A. funestus* P450 genes are primarily closer to their orthologues in *A. gambiae*. This is also the case for the three *AfCYP6M1* genes all closer to *AgCYP6M1* in *A. gambiae* with similarity rates of 79%, 82% and 83%, respectively, for *AfCYP6M1b, AfCYP6M1a* and *AfCYP6M1c*.

The triplication of *AfCYP6M1* was confirmed by PCR amplification and sequencing of the 5′flanking region of each copy of the three genes using genomic DNA from laboratory and field samples. The deduced amino acid sequences of the three genes are all of the same length with 498 amino acids each. There are 59 amino acid substitutions between *AfCYP6M1a* and *AfCYP6M1b* among which 13 replacements between amino acid of opposite physicochemical properties resulting in an 88% similarity (Supplementary Figure S3).

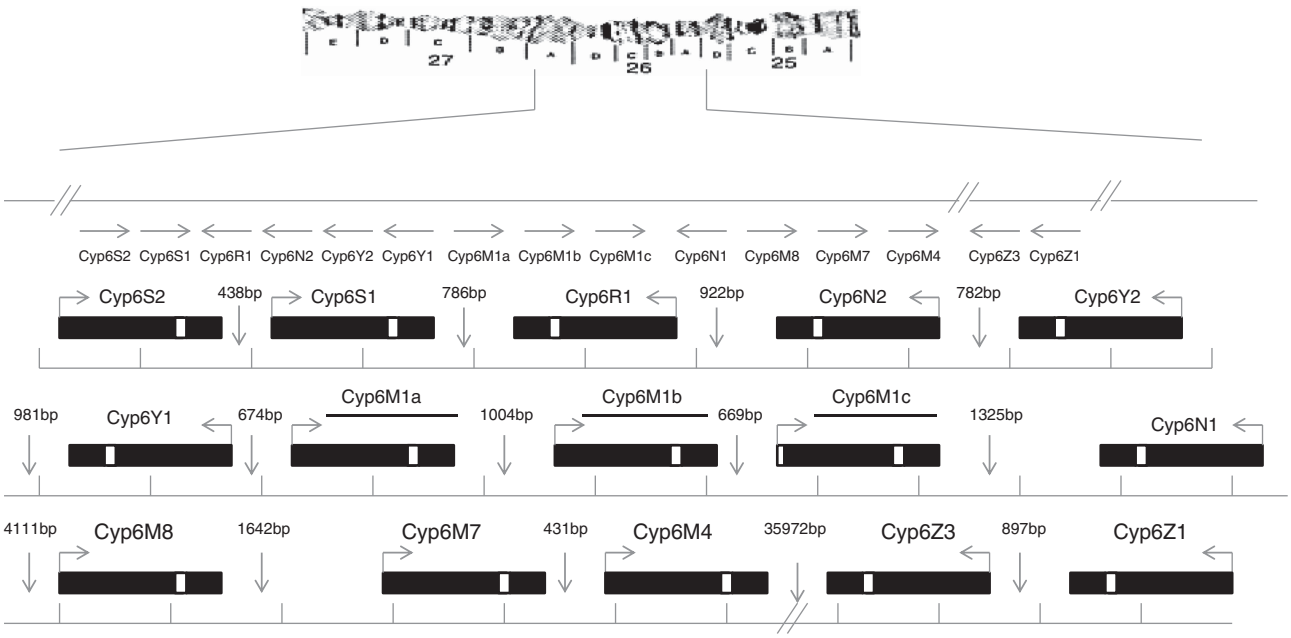

Figure 1 Schematic representation of the gene organisation in the BAC clone spanning the *rp2* QTL in *A. funestus*. Vertical arrows indicate intergenic regions while horizontal arrows indicate the 5′—3′orientation of each gene. The triplicated copies of *AfCYP6M1* gene are underlined. A full color version of this figure is available at the *Heredity* journal online.

AfCYP6M1a and *AfCYP6M1c* have 45 variant amino acids (90% similarity) among which 8 replacements between amino acid of opposite physicochemical properties (Supplementary Figure S3). *AfCYP6M1b* and *AfCYP6M1c* have 57 variant amino acids (88% similarity) among which 12 replacements between amino acid of opposite physicochemical properties (Supplementary Figure S3).

Sequencing of cuticular protein genes

Four CPLCG genes were successfully amplified and sequenced in *A. funestus* samples. As some of these genes have been previously associated with insecticide resistance in *A. gambiae* (Vontas *et al.*, 2005; Djouaka *et al.*, 2008), the aim was to assess the correlation between these genes and *rp2* QTL. Detailed information about these genes is presented in Supplementary Table S9. Overall, high similarity is observed between the four CPLCG genes and their orthologues in *A. gambiae* (Supplementary Table S9) with percentage of protein sequence similarity between 67 and 86%. The *CPLCG3* protein in *A. funestus*, previously associated with possible reduced penetration resistance mechanism in *A. gambiae* (Vontas *et al.*, 2005; Djouaka *et al.*, 2008), had seven amino acid differences with the *A. gambiae* orthologue.

SNP identification in *rp2* genes

In total, 565 SNPs from 28 genes screened were identified after analysing 20 431 bp of sequencing traces indicating that one SNP is present every 36 bp. The patterns of polymorphism observed in this study is similar to the patterns previously described in this species from 50 genes (Wondji *et al.*, 2007a). This is the case for the ratio of transition/tranversion substitutions with transition SNPs significantly predominant in the total (60.9%) over transversion substitutions (39.1%) compared with a ratio of 62% vs 38% in the previous study (Wondji *et al.*, 2007a). This bias towards transition substitutions was also more pronounced in the coding region with 62.7% (271 SNPs out of 432) of transition and 37.3% (161 SNPs) of transversion slightly lower (but not significantly) than the 66.3% vs 33.7% seen previously (Wondji *et al.*, 2007a). More transversion SNPs were significantly observed in non-coding regions (47.5%) than in coding region (37.3%) as also observed in other species such as *A. gambiae* (Morlais *et al.*, 2004) and *Aedes aegypti* (Morlais and Severson, 2003), confirming that SNPs occur more frequently as transition in coding regions than in non-coding regions. This difference between coding and non-coding regions is also reflected in the frequency of SNPs, which is one SNP every 39 bp in coding regions vs one SNP every 24 bp in non-coding regions (5′UTR, 3′UTR and intron combined).

A higher frequency of SNPs was observed at the third codon position (71.9%) than at the first or second position (Table 1). More than 2/3 (308) of the 432 coding SNPs were synonymous, whereas $<1/3$ (124) were non-synonymous or replacement SNPs. Indels were also detected (1–6 bp long) or inferred owing to overlapping peaks in the sequencing traces. Twenty-five triallelic and six tetrallelic SNPs were detected confirming that multiallelic SNPs are present in this mosquito species although at low frequency.

The average nucleotide diversity per gene ($\pi = 0.0105$) was higher than the estimate of 0.0072 in the previous study (Wondji *et al.*, 2007a). The average nucleotide diversity in non-coding DNA (0.013) was lower than in synonymous sites of the coding regions (0.028). This is an indication that non-coding regions are under greater purifying selection than synonymous sites within coding regions. The Ka/Ks ratio was 0.11, which is lower than the estimate of 0.18 seen in the previous study.

Fine-scale linkage and QTL mapping at F_6

Linkage mapping. The mapping of *rp2* BAC clone SNP markers and the other additional SNPs selected from chromosome 2L indicated that they all genetically map to a single-linkage group corresponding to chromosome 2L. This was the case for each family and when they were both combined in a single map (Supplementary Figure S4). The genetic distance between the *rp2* BAC clone loci are very low around 0.2–0.8 cM, indicating a very limited recombination rate between these P450 SNPs as expected from markers located in such a small

Table 1 Nucleotide polymorphism in *A. funestus* genes

	Coding region													*Non-coding region*							
	Polymorphic sites													*Nucleotide diversity*				*Polymorphic sites*			
Gene	*L*	*S*	π	*h*	L_c	*Ts*	*Tv*	*First*	*Second*	*Third*	*Syn*	*Rep*	Σ	π_c	*Ks*	*Ka*	L_{nc}	*Ts*	*Tv*	Σ	π_{nc}
CYP4C35	915	10	0.63	4	600	4	4	4	3	1	1	7	8	0.78	0.36	0.91	315	2	0	2	0.37
CYP4H18	503	11	0.89	7	333	1	1	0	0	2	1	1	2	0.25	0.63	0.12	170	7	2	9	2.1
CYP6AF1	438	12	0.74	6	363	5	0	0	0	5	5	0	5	0.38	1.5	0.0	73	4	3	7	2.5
CYP6AH1	901	25	1.3	4	699	13	4	0	1	16	15	2	17	1.15	4.3	0.2	203	4	4	8	2.04
CYP6M1a	700	13	0.68	10	444	2	1	0	0	3	3	0	3	0.24	1.01	0.0	254	4	6	10	1.4
CYP6M1b	911	27	1.3	6	648	11	12	4	4	15	16	7	23	1.5	4.6	0.59	263	3	1	4	0.73
CYP6M1c	804	39	1.9	15	657	25	14	5	5	29	29	10	39	2.01	7.3	0.6	145	3	1	4	1.2
CYP6M3	522	17	1.24	11	438	8	4	1	1	10	10	2	12	1.03	3.7	0.24	81	2	4	6	2.5
CYP6M4	814	37	2.01	7	750	27	9	10	0	26	28	8	36	2.02	7.03	0.67	62	0	1	1	0.64
CYP6M8	799	25	1.05	11	726	17	4	4	2	15	17	4	21	1.04	3.7	0.22	73	2	2	4	1.13
CYP6N1	866	40	1.9	14	804	22	9	2	1	28	23	8	31	1.6	5.5	0.44	59	5	4	9	5.4
CYP6R1	809	26	1.56	8	744	13	8	1	7	12	12	9	21	1.4	2.8	0.89	59	2	3	5	4.1
CYP6S1	829	33	1.5	12	829	19	14	6	2	25	29	4	33	1.5	5.4	0.27	/	/	/	/	/
CYP6S2	905	29	1.3	16	834	16	8	0	2	22	20	4	24	1.13	4.2	0.19	70	3	2	5	2.9
CYP6Y1	924	15	0.66	8	846	8	6	1	2	11	10	4	14	0.69	2.2	0.23	76	0	1	1	0.26
CYP6Y2	868	9	0.46	5	868	6	3	1	1	7	7	2	9	0.46	1.5	0.14	/	/	/	/	/
CYP6Z1	924	28	1.33	14	510	13	6	4	2	13	12	7	19	1.6	4.1	0.76	413	2	7	9	1.02
CYP9B2	561	15	0.94	9	459	5	6	4	4	3	3	8	11	0.87	1.5	0.7	102	2	2	4	1.3
CYP9M1	891	25	0.94	9	891	15	10	7	4	14	15	10	25	0.94	2.6	0.41	/	/	/	/	/
CYP303A1	765	9	0.48	8	765	6	3	2	1	6	6	3	9	0.46	1.6	0.17	/	/	/	/	/
7980	649	17	0.9	8	336	0	2	0	1	1	0	2	2	0.19	0.0	0.26	311	6	9	15	1.7
9073	734	33	1.57	13	510	7	7	1	1	12	11	3	14	0.89	3.3	0.23	224	15	4	19	3.1
9624	788	24	1.04	10	786	13	11	6	2	16	17	7	24	1.04	3.5	0.28	/	/	/	/	/
CPLC5	390	3	0.17	4	390	2	1	0	0	3	3	0	3	0.17	0.6	0.0	/	/	/	/	/
CPLC8	253	8	1.14	9	253	3	6	2	2	5	3	6	9	1.14	3.2	0.34					
CPLC9	317	17	2.05	7	117	3	2	0	0	5	5	0	5	1.4	5.5	0.0	200	6	6	12	2.4
CPR*	895	5	0.18	4	792	3	2	1	1	3	4	1	5	0.21	0.75	0.03	101	0	0	0	0
GSTe2	756	13	0.54	13	612	4	4	2	3	3	3	5	8	0.36	0.7	0.25	143	1	4	5	1.13
Total	20431	565			17004	271	161	68	52	311	308	124	432				3397	73	66	139	
Average			1.05	8.7										0.91	2.86	0.31					1.31

L, length of the nucleotide sequence; S, number of mutations; π, average number of nucleotide substitution per site; h, number of haplotypes; L_c, Length of coding region; Ts, transitions; Tv, transversions Syn, synonymous substitutions; Rep, replacement substitutions; Σ, total; π_c, nucleotide diversity in coding region; Ks, per synonymous site; Ka, per non-synonymous site; π_{nc}, nucleotide diversity in non-coding region. * CPR, cytochrome P450 reductase

genomic region. The markers order on the linkage map is different to that seen in *A. gambiae* but is in accordance with the synteny pattern observed between *A. funestus* and *A. gambiae* for chromosome 2L (equivalent of 3R in *A. gambiae*) as indicated previously (Sharakhov *et al.*, 2004).

Genotype/phenotype correlation for rp2 loci. χ^2 goodness-of-fit tests carried out identified loci significantly associated ($P<0.05$) with pyrethroid resistance in *rp2* in both reciprocal families (Figure 2). In family 1, all the seven informative SNPs genotyped from the *rp2* clone showed a significant correlation with the phenotype with a higher mortality rate observed in individuals homozygous for the allele of the susceptible parent (Figure 2a). In average, mortality rates ranged from 21–33% when the individuals have no allele of the susceptible parent, 42–47% for heterozygotes and 58–77% for individuals with two alleles of the susceptible parent indicating an additive effect in the resistance associated with the *rp2* QTL as the mortality increases with the number of alleles from the susceptible parent.

When analysing other loci of the 2L chromosome away from the BAC clone, a significant correlation was also detected for three other markers, the P450 *CYP9M1*, the cytochrome c BU10 (orthologue of AGAP009537-RA in *A. gambiae*) and BU29, a putative sensory appendage protein orthologue of *SAP-3* in *A. gambiae* (AGAP008054-RA). However, the additive effect seen for the *rp2* BAC clone genes was not observed for these three loci as individuals with no allele of the susceptible parent and the heterozygotes had a similar mortality rate. No correlation was observed between the two SNPs in the cuticular protein genes and the resistance phenotype.

For Family 10, although a similar trend was observed (Figure 2b), no significant difference was recorded. Contrary to Family 1, the range of mortality rates between the three genotypes was closer in this family. In average, mortality rates ranged from 42–47% when the individuals have no allele of the susceptible parent, from 50–55% for heterozygotes and 58–80% for those homozygote susceptible mosquitoes indicating an additive effect of the *rp2* QTL as observed in Family 1. The parental female of this family repeatedly exhibited a heterozygous haplotypic pattern to most of the genes sequenced in *rp2* similar to the F_1 progeny of family 1. This heterozygote status could explain this lack of significant correlation in Family 10 despite the existence of a trend between the mortality rate and the resistance phenotype.

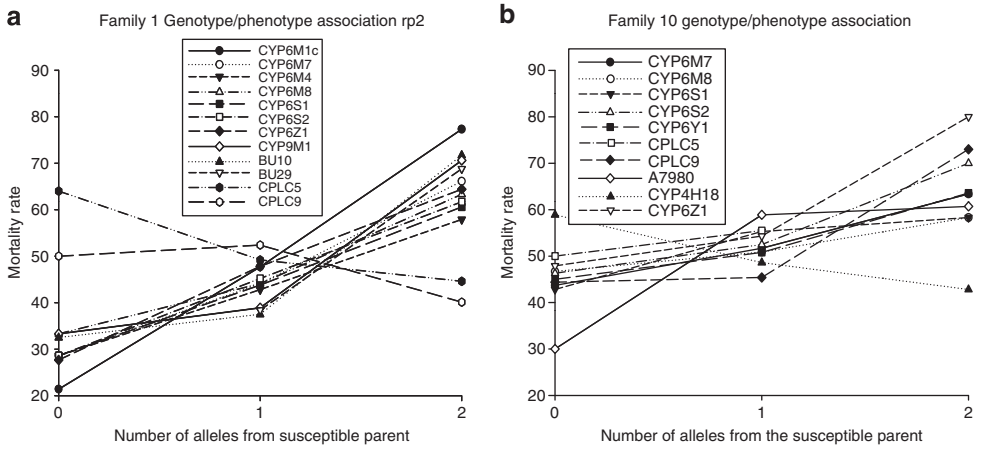

Figure 2 Plot of mortality as a function of alleles inherited from susceptible parents at loci associated with resistance in rp2 QTL using F6 generation for Family 1 (**a**) (FUMOZ-R ♂ X FANG ♀) and (**b**) for Family 10 (FUMOZ-R ♀ X FANG ♂).

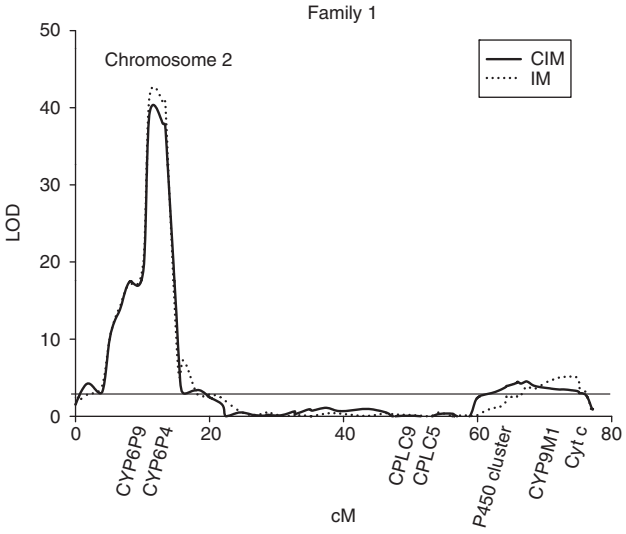

Figure 3 Plot of logarithm of odd scores for the pyrethroid resistance *rp2* QTL in Family 1 using F_6 progeny. The y axis indicates logarithm of odd ratio scores, and the x axis indicates chromosome positions. Solid lines represent logarithm of odd estimated by composite-interval mapping and dashed lines represent logarithm of odd estimated by interval mapping. The straight line along the top of the graph represents the threshold value for logarithm of odd as determined by permutations. Names of markers are listed around QTL locations.

QTL mapping. The fine-scale mapping of the markers on the 2L chromosome confirms the presence of the *rp2* QTL in Family 1 with a small increase of the logarithm of the odd from 4 in the previous study (Wondji *et al*., 2009) (Figure 3) to 5. This logarithm of odd, estimated with both interval mapping and composite-interval mapping is still significantly lower than that of the *rp1* QTL on chromosome 2R (Figure 3). The boundaries of this *rp2* QTL encompass the cluster of the 15 P450 genes on the 113 kb BAC clone but extend also to the *AfCYP9M1* P450 gene and the cytochrome c gene (BU10) confirming the pattern observed for the genotype/phenotype association. The loci from two cuticular protein genes do not map within this *rp2* QTL indicating as already shown by the genotype/phenotype association that these genes found to be associated with the reduced penetration resistance in *A. gambiae*

Table 2 Multiple-interval mapping estimates of QTL position and associated genetic, environmental, phenotypic, additive and dominance effects associated with pyrethroid resistance in Family 1 *A. funestus* using F_6 progeny

σ_g^2(% σ_p^2)	σ_e^2(% σ_p^2)	σ_p^2	*Nearest marker*	*QTL*	*Genetic distance (cM)*	*Effect*	% σ_g^2
22.16	2.84	25	CYP6P9	*rp1*	13.7	A, −4.77	65.8
(88.7%)	(11.3%)				(7.3–14.3)	D, 4.02	19.6
			CYP6M4	*rp2*	65.9	A, −0.76	4.1
					(60.5–69.5)	D, 0.45	−0.8
			BU92	*rp3*	92	A, −0.10	0.2
					(85.9–100.9)	D, 0.31	−0.3

σ_g^2, σ_p^2, σ_e^2, respectively, for genetic, phenotypic (in parentheses) and environmental variance; A for additive; D for dominance; confidence intervals for QTL position are in parentheses near the position estimate.

(Djouaka *et al*., 2008) probably have no role in the resistance associated with *rp2* QTL. Multiple-interval mapping indicated that *rp2* QTL explains about 3.3% of the genetic variance with *rp1* still the major QTL explaining 85.4% of the genetic variance (Table 2). The genetic variance (σ_g^2) explained by the three QTLs now accounts for 88.7% of the phenotypic variance (σ_p^2), whereas 11.3% are explained by environment variance (σ_e^2) (Table 2).

The *rp2* QTL was not detected in family 10 in line with the lack of significant genotype/phenotype correlation on chromosome 2L in this family. This absence of *rp2* in Family 10 is probably due to the heterozygote profile of the parental female of this family, which originated from the resistant strain FUMOZ-R. The combined genetic map of both families did not change much for *rp2* as a similar logarithm of odd was observed.

Multiplex gene expression profiling of genes around *rp2* QTL

In order to detect the candidate genes associated with the *rp2* QTL, the expression profile of the 15 P450 genes from the *rp2* BAC clone was assessed in comparison with the reference ribosomal protein S7 gene (*RSP7*) between the resistant strain FUMOZ-R and the susceptible FANG strain. This multiplex qPCR experiment detected no expression signal for *AfCYP6M1b*, *AfCYP6S1* and *AfCYP6S2* either in resistant or in susceptible samples. The gene with the highest expression level was *AfCYP6M8* orthologue of *AgCYP6M2*

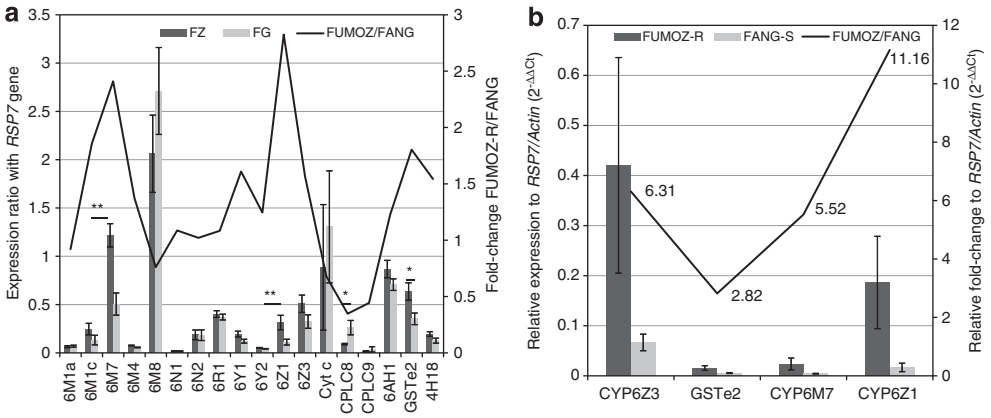

Figure 4 Gene expression profiling of candidate genes: **a** is the expression profile of some of the 28 tested P450 genes in females of the resistant strain FUMOZ-R and the susceptible strain FANG using the Beckman GeXP multiplex method. The normalised expression ratio of each gene against the *RSP7* gene is represented on the primary vertical axis, while the secondary vertical axis (curve) represents the fold change of each gene between FUMOZ-R and FANG. Significant differential expressions are indicated by asterisks: ** for $P<0.01$ and * if $P<0.05$. **b** is the qRT–PCR expression patterns of the four candidate genes in females of the resistant strain FUMOZ-R and the susceptible strain FANG. The normalised ($2^{-\Delta\Delta Ct}$) relative expression ratio of each gene against the *RSP7* and *Actin* genes is represented on the primary vertical axis while the secondary vertical axis (curve) represents the fold change of the $2^{-\Delta\Delta Ct}$ of each gene between FUMOZ-R and FANG.

gene regularly associated with pyrethroid resistance in *A. gambiae* (Djouaka *et al.*, 2008; Muller *et al.*, 2008; Stevenson *et al.*, 2011). However, this high expression of *AfCYP6M8* is observed in both resistant and susceptible samples indicating that contrary to its *A. gambiae AgCYP6M2* orthologue, *AfCYP6M8* is not associated with pyrethroid resistance in this FUMOZ-R strain. A significant differential expression ($P<0.01$; *t*-test) was observed at two genes, *AfCYP6M7* and *AfCYP6Z1*, between the resistant FUMOZ-R and the susceptible FANG, with, respectively, a 2.4- and 2.82-fold-change overexpression in the resistant strain (Figure 4a). No significant difference was observed for the remaining genes, although a 1.7-fold change was also observed for *AfCYP6Z3* but at $P=0.065$ (*t*-test).

The analysis of six other genes of interest on the 2L chromosome away from the *rp2* BAC clone indicated a significant downexpression of the cuticular protein *CPLCG3* (CPLC8) gene in the resistant strain contrary to previous observation made in *A. gambiae* associating this gene with pyrethroid resistance in Djouaka *et al.*, 2008. A significant overexpression of the *GSTe2* gene ($P=0.017$) was also observed in the resistant FUMOZ-R strain. No informative SNP was identified in this gene and therefore was not included in the QTL mapping to assess its correlation with resistance phenotype. No differential expression was observed for the cytochrome c oxidase subunit I gene (orthologue of the *A. gambiae* AGAP009537 gene), which exhibited a significant genotype/phenotype association.

The expression profile of *AfCYP6Z1*, *AfCYP6M7* and *GSTe2* upregulated in FUMOZ-R through the multiplex GeXP expression profiling was further assessed individually using the qRT–PCR. This was also done for *AfCYP6Z3*, which showed a 1.7-fold change although not significant. All genes tested had PCR efficiency between 90 and 102%. A significant overexpression was observed for all the four genes when their ($2^{-\Delta\Delta Ct}$) relative expression was compared with that of the two housekeeping genes *RSP7* and Actin either individually or when combined (Figure 4b). The highest fold change is observed for *CYP6Z1* with an 11.16-fold change upregulation in FUMOZ-R compared with the susceptible FANG. *AfCYP6Z3* is 6.3-fold overexpressed in FUMOZ-R, while *AfCYP6M7* is also upregulated in this resistant strain at 5.5-fold. *GSTe2* also showed a significant upregulation in FUMOZ-R but at a lower fold change of 2.82.

Table 3 Summary statistics for polymorphism of *AfCYP6Z1* and *AfCYP6Z3* between the FANG susceptible and FUMOZ-R resistant strains

Samples	*N*	*S*	*Syn*	*Nonsyn*	*π (k)*	*D (Tajima)*	*D* (Fu and Li)*
AfCYP6Z1							
FANG-S	4	32	19	13	0.0114 (16.8)	-0.37^{ns}	-0.27^{ns}
FUMOZ-R	4	28	18	10	0.0094 (14.0)	-0.86^{ns}	-0.86^{ns}
Total	8	49	30	19	0.0115 (17.1)	-0.52^{ns}	-0.56^{ns}
AfCYP6Z3							
FANG-S	5	39	33	6	0.0116 (17.2)	-0.61^{ns}	-0.61^{ns}
FUMOZ-R	6	34	29	5	0.0117 (17.3)	1.04^{ns}	1.34^{ns}
Total	11	59	50	9	0.015 (22.7)	0.61^{ns}	0.38^{ns}

N, number of sequences (n); Nonsyn, Non-synonymous mutations; ns, not significant; S, number of polymorphic sites; Syn, Synonymous mutations; Tajima's D and Fu and Li's D statistics; π, nucleotide diversity (k = mean number of nucleotide differences).

Polymorphism analysis of *AfCYP6Z1* and *AfCYP6Z3*

The polymorphism of the *AfCYP6Z1* and *AfCYP6Z3* was analysed in order to identify potential mutations associated with pyrethroid resistance. The full length of the coding region of *AfCYP6Z1* was sequenced for eight clones (four FANG and four FUMOZ-R), whereas 11 clones (five FANG and six FUMOZ-R) were sequenced for *AfCYP6Z3*. The two genes exhibited a high polymorphism with a total of 49 and 59 polymorphic sites observed, respectively, for *AfCYP6Z1* and *AfCYP6Z3*, while 19 amino acid changes were observed for *AfCYP6Z1* (Supplementary Figure S5) and 9 for *AfCYP6Z3* (Supplementary Figure S6). This level of polymorphism is similar between the FANG and the FUMOZ-R strain with, nevertheless, a slightly higher number of mutations in the susceptible strain for the two genes. Other parameters of this polymorphism are summarised in Table 3. A Neighbour-joining tree of the haplotypes of both genes indicated a cluster of haplotypes specific to each strain for both genes (Figure 5) but there are also haplotypes similar between the two strains.

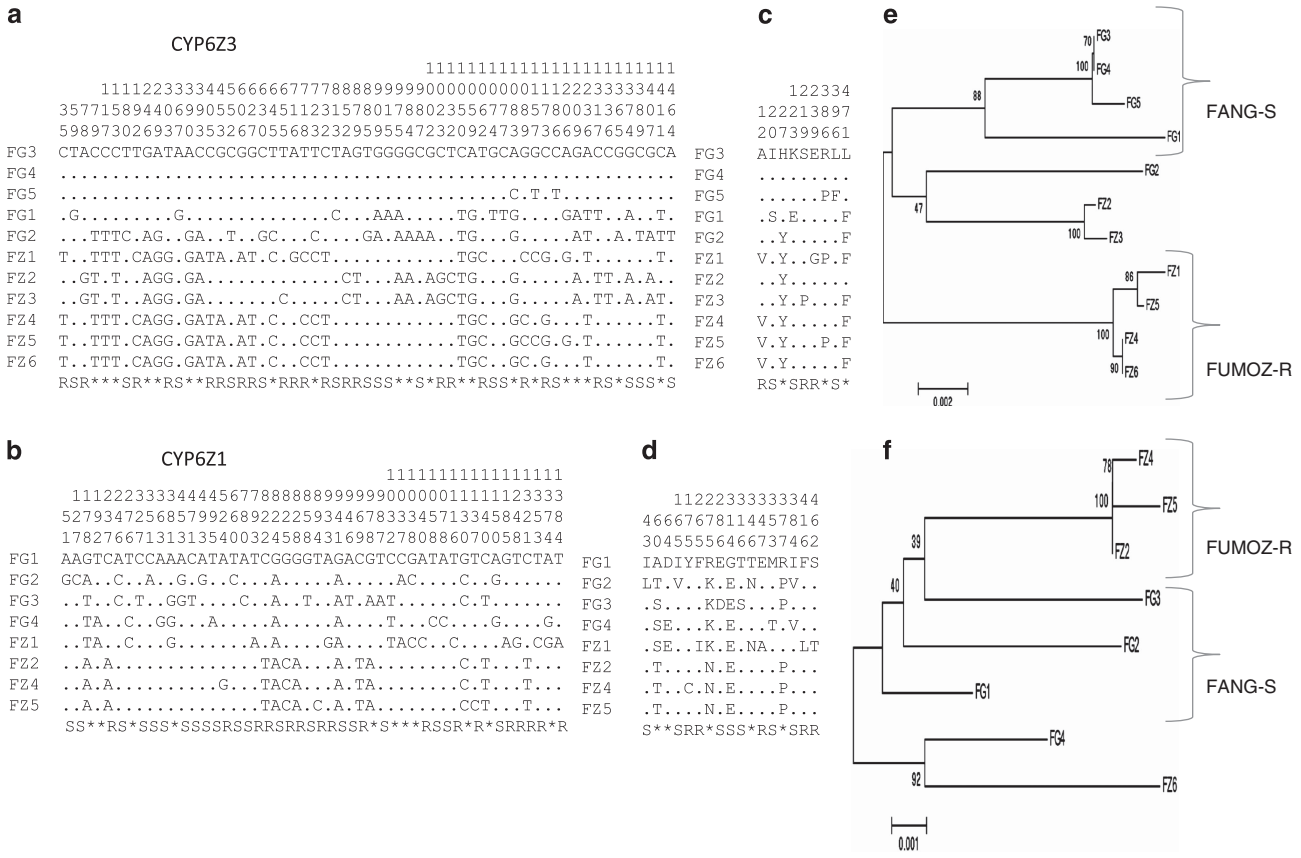

Figure 5 Schematic representation of haplotypes of *AfCYP6Z3* and *AfCYP6Z1* genes between the susceptible FANG and the resistant FUMOZ-R strains. **a**, **b** show the nucleotide polymorphic sites for both *AfCYP6Z3* and *AfCYP6Z1*, respectively, while (**c**, **d**) are for the amino acid polymorphism. Below the list of sequences, R or S indicates the positions that are polymorphic in the resistant or susceptible mosquitoes, respectively, while an asterisk (*) marks a position polymorphic in both phenotypes. (**e**) Neighbour-joining tree of *AfCYP6Z3* showing the clades specific to each phenotype while (**f**) is for *AfCYP6Z1*.

DISCUSSION

The characterisation of mechanisms of insecticide resistance is a key prerequisite for the successful management of vector control programms. In order to fully elucidate the underlying mechanisms conferring pyrethroid resistance in a resistant strain of the major malaria vector *A. funestus*, we successfully carried out a fine-scale mapping of the *rp2* QTL associated with this resistance.

The successful isolation and sequencing of the 113 kb BAC clone was a significant step in allowing the positional mapping of *rp2* QTL. The main observation from this *rp2* clone sequencing was the identification of another copy number increase of a P450 gene in *A. funestus* compared with *A. gambiae* with three genes related to *AgCYP6M1* detected in *A. funestus*. Two other duplication events were already previously reported for *AfCYP6P9* and *AfCYP6P4* genes in a 120 kb BAC clone spanning *rp1* QTL (Wondji *et al.*, 2009). However, the triplication of *AfCYP6M1* is not associated with the resistance mechanism as this event is ancient as indicated by the high number of amino acid changes and the even lower similarity between the three copies (88 to 90%) than that seen between copies of the two duplicated P450 genes *AfCYP6P9a* and *AfCYP6P9b* in *rp1* (95% similarity) and *AfCYP6P4a* and *AfCYP6P4b* (94% similarity).

Another noticeable difference between *A. funestus* and *A. gambiae* is the absence in this *rp2* BAC clone of the *A. gambiae AgCYP6Z2* gene previously associated with insecticide resistance (Muller *et al.*, 2007) although not a pyrethroid metabolizer (McLaughlin *et al.*, 2008). In *A. gambiae*, the *AgCYP6Z1*, *AgCYP6Z2* and *AgCYP6Z3* genes are found together in a 7-kb-long genomic region on the 3R chromosome. In this *rp2* clone, only *AfCYP6Z1* and *AfCYP6Z3* were detected and both are separated by an intergenic space of 897 bp and no other gene was detected within the next 20 kb indicating that *AgCYP6Z2* orthologue is probably missing in *A. funestus* either through a deletion event in *A. funestus* or through a duplication event in *A. gambiae*. The 92% protein similarity observed between *AgCYP6Z2* and *AgCYP6Z3* in *A. gambiae* indicates that these two genes are probably duplicates of an ancestral gene orthologue to the *AfCYP6Z3* found in *A. funestus*. This is further supported by the fact that the similarity between *AfCYP6Z3* and *AfCYP6Z1* in *A. funestus* is only 68%. It is not impossible that this *AgCYP6Z2* orthologue may still be found in *A. funestus* on the 2L chromosome but beyond the *rp2* BAC clone. However, this is unlikely because of the strong conservation observed in the gene organisation pattern of the P450 clusters in *rp1* and in *rp2* between *A. gambiae* and *A. funestus*.

SNP identification for fine-scale mapping indicated a ratio of one SNP every 36 bp, which is slightly higher than the estimate of one SNP every 41 bp found previously in a similar study in a diverse set of 50 genes in *A. funestus* (Wondji *et al.*, 2007a). This increase and other differences between the two studies could be due to the higher number of P450 genes sequenced (20 out of the 28) in this study contrary to just 10 out of 50 in the previous one. P450 genes have been shown to be a highly polymorphic gene family in *A. funestus* and *A. gambiae* (Wilding *et al.*, 2009). Overall, the similarity of the polymorphism pattern observed in this study to previous reports indicates

the robustness of the SNP identification process in this study. This is further confirmed by a successful genotyping of a subset of these SNPs in the isofemale lines progeny. These 565 SNPs located in genes of interest such as P450s constitute an additionally genomic tool for this species for which less genomic resources are available compared with *A. gambiae*.

Although the *rp2* QTL spans a broad genomic region, it can be concluded that the P450 cluster from the BAC clone exhibits the strongest association with the resistance phenotype and therefore contains the likely genes involved in the pyrethroid resistance. This is seen through the linear and additive correlation observed between SNPs in the P450 cluster and the resistance phenotype contrary to the partial association seen for the cytochrome c and *AfCYP9M1* genes or no correlation for the cuticular genes. This is further confirmed by the significantly overexpression of *AfCYP6Z1, AfCYP6Z3* and *AfCYP6M7* genes in the resistant FUMOZ-R strain after qRT–PCR, while no such overexpression was seen for either cuticular protein genes or the *AfCYP9M1* (data not presented) or the cytochrome c.

The association of *AfCYP6Z1* with pyrethroid resistance is not surprising as its orthologue in *A. gambiae* has previously been associated with pyrethroid (David *et al.*, 2005) and is able to metabolise DDT (Chiu *et al.*, 2008). *AfCYP6M8* was not associated with resistance in this *A. funestus* strain despite having a 92% similarity with its *A. gambiae* orthologue *AgCYP6M2*, a metaboliser of pyrethroids (Stevenson *et al.*, 2011).

The 6.3-fold change observed for *AfCYP6Z3* in FUMOZ-R indicates that this gene is probably also associated with the pyrethroid resistance explained by the *rp2* QTL. This will not be surprising as *AgCYP6Z3* have also been linked with insecticide resistance in *A. gambiae* (Muller *et al.*, 2007).

The overexpression of *AfCYP6M7* observed in this study confirmed a recent report that this gene was overexpressed in this FUMOZ-R strain after microarray (Christian *et al.*, 2011). The orthologue of *AfCYP6M7* in *A. gambiae*, *AgCYP6M3*, is not commonly associated with insecticide resistance. It was found to be marginally overexpressed in a field permethrin *A. gambiae* population by microarray but not confirmed by qPCR (Muller *et al.*, 2008).

The absence of upregulation of the cuticular protein gene in FUMOZ-R correlates with the lack of significant correlation between genotypes of markers at these genes and phenotype. This is an indication that this cluster of cuticular protein genes on 2L is not associated with the cuticle thickening observed recently in resistant mosquitoes of this FUMOZ-R strain (Wood *et al.*, 2010).

An overexpression of *GSTe2* also was observed in this study, although at a lower fold change of 2.82 compared with the three P450 genes. Unfortunately, no informative SNP was observed in this gene for both family 1 and 10, which prevented its mapping and to assess its association with *rp2* QTL. However, when taking in consideration the synteny between *A. gambiae* and *A. funestus* chromosomal map, *GSTe2* appears to be located within the boundaries of *rp2*. GSTe2 has been validated as DDT metabolizer in various mosquito species (Ding *et al.*, 2005; Lumjuan *et al.*, 2005). Because the FUMOZ-R strain is fully susceptible to DDT, the overexpression of GSTe2 in this study is probably associated with pyrethroid resistance as previously observed in *A. gambiae* (David *et al.*, 2005) and in other insects, such as *Nilaparvata lugens* (Vontas *et al.*, 2001). As GSTs are known to conjugate insecticides or their metabolites (phase II), it is possible that GSTe2 might be involved in the conjugation of permethrin metabolites.

The allelic variation of Af*CYP6Z1* and *AfCYP6Z3* between the susceptible and resistant was assessed to establish a possible correlation with the resistance phenotype. Such correlation has been previously established for the CYP6AB3 P450 gene in the insect *Depressaria pastinacella* (parsnip webworm). Five amino acid changes observed in this gene enhance the metabolism of plant allelochemicals by altering a proximal surface residue and potential interactions with cytochrome P450 reductase (Mao *et al.*, 2007). Although the polymorphisms observed between the two FANG and FUMOZ-R strains for each gene could reflect their different genetic background, it could also be a consequence of a correlation with the resistance phenotype. This is supported by the location and possible impact on catalytic properties of some amino acid changes between the two strains observed in the two genes. For example, in *AfCYP6Z1*, three amino acid changes exclusively present in the resistant FUMOZ-R strain could impact the function of the gene. These are an Y175C which is the first residue in the highly conserved E helix; the F265I located within the highly variable H helix and the F416L located within the conserved PERF region. These changes could after the catalytic activity of *AfCYP6Z1* in FUMOZ-R. For *AfCYP6Z3*, several of the amino acid predominantly found in FUMOZ-R could also enhance the catalytic activity of this gene in association with resistance to pyrethroid. Among these amino acid substitutions are: an A12V and H27Y both located within the hydrophobic region that anchors the protein to the membrane; a S219P mutation within the hydrophobic region located between F and G helices, thought to penetrate lipid bilayer, making contact with environment from which many substrates can enter the active site; a E239G located within the SRS-3 and may enhance catalysis; and a L471F, the first residue in the SRS-6, which may enhance catalysis of *AfCYP6Z3* in FUMOZ-R. Further analysis of such mutations will confirm their importance and role in conferring pyrethroid resistance in the FUMOZ-R strain. Because mutations associated with resistance are also likely to be found in the regulatory regions of these over-transcribed genes such as the upstream region and UTRs, future work will also investigate these regions. Additionally, the sequencing of other P450 genes in this *rp2* QTL should also be envisaged in order to identify non-synonymous mutations, which, although not necessarily associated with an increase in transcription of the gene, could still modify the conformation of the enzyme and its substrate specificity (higher affinity for the chemical in the resistant strain).

The positional cloning of *rp2 QTL* in this study is a significant contribution to the full elucidation of mechanisms of pyrethroid resistance in this important malaria vector. Although *rp2* is not the major QTL in the FUMOZ-R, it is possible that its relative importance may vary in field populations as it may have a bigger role in other field populations of this species in Africa as seen in *A. gambiae* for which the main genes conferring metabolic resistance to pyrethroid varied from one region to another. Therefore, this elucidation is significant and will help to improve the understanding of resistance mechanisms in this species. Further work is needed to confirm the role had by each of the overexpressed P450 genes and *GSTe2* (such as their ability to metabolise or interact with pyrethroids) and to see whether they are also associated with the resistance in field populations of *A. funestus* across Africa.

DATA ARCHIVING

The sequence sets containing all the SNPs have been deposited at Genbank (accession number: JX431304-JX431487). Genotype data for genetic and QTL mapping, qRT–PCR and DNA sequences have been deposited at Dryad: doi:10.5061/dryad.v218g.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by a Wellcome Trust Research Career Development Fellowship to CSW. We thank Professor Maureen Coetzee for earlier work done on crossing the FUMOZ and FANG strains. We also thank Professor Janet Hemingway and Professor Hilary Ranson for initial contribution in this work. JMR was supported by a fellowship from the Fundacion Ramon Aceres.

Author contributions: HI, JMR and CSW carried out the experiments; SI contributed to the data analysis; NFL contributed to the *rp2* BAC clone identification from the *Anopheles funestus* BAC library at Notre Dame; CSW designed the study, analysed the data and drafted the paper. All authors read and approved the final paper.

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

Chiu TL, Wen Z, Rupasinghe SG, Schuler MA (2008). Comparative molecular modeling of Anopheles gambiae CYP6Z1, a mosquito P450 capable of metabolizing DDT. *Proc Natl Acad Sci USA* **105**: 8855–8860.

Christian RN, Strode C, Ranson H, Coetzer N, Coetzee M, Koekemoer LL (2011). Microarray analysis of a pyrethroid resistant African malaria vector, Anopheles funestus, from southern Africa. *Pest Biochem Physiol* **99**: 140–147.

Crawford JE, Guelbeogo WM, Sanou A, Traoré A, Vernick KD, Sagnon N *et al.* (2010). De novo transcriptome sequencing in Anopheles funestus using Illumina RNA-seq technology. *PLoS One* **5**: e14202.

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

David JP, Strode C, Vontas J, Nikou D, Vaughan A, Pignatelli PM *et al.* (2005). The Anopheles gambiae detoxification chip: a highly specific microarray to study metabolic-based insecticide resistance in malaria vectors. *Proc Natl Acad Sci USA* **102**: 4080–4084.

Ding Y, Hawkes N, Meredith J, Eggleston P, Hemingway J, Ranson H (2005). Characterization of the promoters of Epsilon glutathione transferases in the mosquito Anopheles gambiae and their response to oxidative stress. *Biochem J* **387**: 879–888.

Djouaka R, Irving H, Tukur Z, Wondji CS (2011). Exploring Mechanisms of Multiple Insecticide Resistance in a Population of the Malaria Vector Anopheles funestus in Benin. *PLoS One* **6**: e27760.

Djouaka RF, Bakare AA, Coulibaly ON, Akogbeto MC, Ranson H, Hemingway J *et al.* (2008). Expression of the cytochrome P450s, CYP6P3 and CYP6M2 are significantly elevated in multiple pyrethroid resistant populations of Anopheles gambiae s.s. from Southern Benin and Nigeria. *BMC Genomics* **9**: 538.

Feyereisen R (2011). Arthropod CYPomes illustrate the tempo and mode in P450 evolution. *Biochim Biophys Acta* **1814**: 19–28.

Gillies MT, De Meillon B (1968). *The anophelinae of Africa South of the Sahara*. The South African Institute for Medical Research, Johannesburg.

Gregory R, Darby AC, Irving H, Coulibaly MB, Hughes M, Koekemoer LL *et al.* (2011). A de novo expression profiling of Anopheles funestus, malaria vector in Africa, using 454 pyrosequencing. *PLoS One* **6**: e17418.

Hargreaves K, Koekemoer LL, Brooke BD, Hunt RH, Mthembu J, Coetzee M (2000). Anopheles funestus resistant to pyrethroid insecticides in South Africa. *Med Vet Entomol* **14**: 181–189.

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

Hunt RH, Brooke BD, Pillay C, Koekemoer LL, Coetzee M (2005). Laboratory selection for and characteristics of pyrethroid resistance in the malaria vector Anopheles funestus. *Med Vet Entomol* **19**: 271–275.

Lander ES, Botstein D (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185–199.

Livak KJ (1984). Organization and mapping of a sequence on the Drosophila melanogaster X and Y chromosomes that is transcribed during spermatogenesis. *Genetics* **107**: 611–634.

Lumjuan N, McCarroll L, Prapanthadara LA, Hemingway J, Ranson H (2005). Elevated activity of an Epsilon class glutathione transferase confers DDT resistance in the dengue vector, Aedes aegypti. *Insect Biochem Mol Biol* **35**: 861–871.

Mao W, Rupasinghe SG, Zangerl AR, Berenbaum MR, Schuler MA (2007). Allelic variation in the Depressaria pastinacella CYP6AB3 protein enhances metabolism of plant allelochemicals by altering a proximal surface residue and potential interactions with cytochrome P450 reductase. *J Biol Chem* **282**: 10544–10552.

McLaughlin LA, Niazi U, Bibby J, David JP, Vontas J, Hemingway J *et al.* (2008). Characterization of inhibitors and substrates of Anopheles gambiae CYP6Z2. *Insect Mol Biol* **17**: 125–135.

Morgan JC, Irving H, Okedi LM, Steven A, Wondji CS (2010). Pyrethroid resistance in an Anopheles funestus population from Uganda. *PLoS One* **5**: e11872.

Morlais I, Poncon N, Simard F, Cohuet A, Fontenille D (2004). Intraspecific nucleotide variation in Anopheles gambiae: new insights into the biology of malaria vectors. *Am J Trop Med Hyg* **71**: 795–802.

Morlais I, Severson DW (2003). Intraspecific DNA variation in nuclear genes of the mosquito Aedes aegypti. *Insect Mol Biol* **12**: 631–639.

Muller P, Donnelly MJ, Ranson H (2007). Transcription profiling of a recently colonised pyrethroid resistant Anopheles gambiae strain from Ghana. *BMC Genomics* **8**: 36.

Muller P, Warr E, Stevenson BJ, Pignatelli PM, Morgan JC, Steven A *et al.* (2008). Field-caught permethrin-resistant Anopheles gambiae overexpress CYP6P3, a P450 that metabolises pyrethroids. *PLoS Genet* **4**: e1000286.

Okoye PN, Brooke BD, Koekemoer LL, Hunt RH, Coetzee M (2008). Characterisation of DDT, pyrethroid and carbamate resistance in Anopheles funestus from Obuasi, Ghana. *Trans R Soc Trop Med Hyg* **102**: 591–598.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.

Sharakhov I, Braginets O, Grushko O, Cohuet A, Guelbeogo WM, Boccolini D *et al.* (2004). A microsatellite map of the African human malaria vector Anopheles funestus. *J Hered* **95**: 29–34.

Stam P, Van Ooijen JW (1995). *JoinMap (tm) version 2.0: Software for the calculation of genetic linkage maps*. CPRO-DLO: Wageningen, The Netherlands.

Stevenson BJ, Bibby J, Pignatelli P, Muangnoicharoen S, O'Neill PM, Lian LY *et al.* (2011). Cytochrome P450 6M2 from the malaria vector Anopheles gambiae metabolizes pyrethroids: Sequential metabolism of deltamethrin revealed. *Insect Biochem Mol Biol* **41**: 492–502.

Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**: 1596–1599.

Thompson JD, Higgins DG, Gibson TJ (1994). 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.

Vontas J, Blass C, Koutsos AC, David JP, Kafatos FC, Louis C *et al.* (2005). Gene expression in insecticide resistant and susceptible Anopheles gambiae strains constitutively or after insecticide exposure. *Insect Mol Biol* **14**: 509–521.

Vontas JG, Small GJ, Hemingway J (2001). Glutathione S-transferases as antioxidant defence agents confer pyrethroid resistance in Nilaparvata lugens. *Biochem J* **357**: 65–72.

Wang S, Basten CJ, Zeng Z-B (2005). *Windows QTL Cartographer 2.5*. Department of Statistics, North Carolina State University: Raleigh, NC.

Wilding CS, Weetman D, Steen K, Donnelly MJ (2009). High, clustered, nucleotide diversity in the genome of Anopheles gambiae revealed through pooled-template sequencing: implications for high-throughput genotyping protocols. *BMC Genomics* **10**: 320.

Wondji CS, Dabire RK, Tukur Z, Irving H, Djouaka R, Morgan JC (2011). 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.

Wondji CS, Hemingway J, Ranson H (2007a). Identification and analysis of Single Nucleotide Polymorphisms (SNPs) in the mosquito Anopheles funestus, malaria vector. *BMC Genomics* **8**: 5.

Wondji CS, Hunt RH, Pignatelli P, Steen K, Coetzee M, Besansky N *et al.* (2005). An integrated genetic and physical map for the malaria vector Anopheles funestus. *Genetics* **171**: 1779–1787.

Wondji CS, Irving H, Morgan J, Lobo NF, Collins FH, Hunt RH *et al.* (2009). Two duplicated P450 genes are associated with pyrethroid resistance in Anopheles funestus, a major malaria vector. *Genome Res* **19**: 452–459.

Wondji CS, Morgan JC, Coetzee M, Hunt RH, Steen K, Black WC *et al.* (2007b). Mapping a Quantitative Trait Locus conferring pyrethroid resistance in the African malaria vector Anopheles funestus. *BMC Genomics* **8**: 34.

Wood OR, Hanrahan S, Coetzee M, Koekemoer LL, Brooke BD (2010). Cuticle thickening associated with pyrethroid resistance in the major malaria vector Anopheles funestus. *Parasit Vectors* **3**: 67.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)