

# The genetic architecture of target-site resistance to DDT and pyrethroids in the malaria vectors *Anopheles gambiae* and *Anopheles coluzzii*

DRAFT

Chris S. Clarkson<sup>1</sup>, Alistair Miles<sup>2,1</sup>, Nicholas J. Harding<sup>2</sup>, @@TODO<sup>1</sup>, Dominic Kwiatkowski<sup>1,2</sup>, Martin Donnelly<sup>3,1</sup>, and The *Anopheles gambiae* 1000 Genomes Consortium<sup>4</sup>

<sup>1</sup>Sanger @@TODO

<sup>2</sup>Oxford @@TODO

<sup>3</sup>Liverpool @@TODO

<sup>4</sup>MalariaGEN @@TODO

WIP

## Abstract

TODO edits from Chris... this is a drill, this is a drill

## Introduction

The malaria vectors *Anopheles gambiae* and *Anopheles coluzzii* are evolving insecticide resistance. Asdlkj dsalkj daslkjd aslkjdas lkadsj lkadsj adslkj adslkja dslkadsj lkadsj lkasd jlkadsj lkads jlkads alksdj asdlk jaslkd adslk jaslkd adslkj adslkj adslkj adslkj adslkj adslkj adslkj jasd.

This is the second paragraph of the introduction. Asdlkjio weipo ewrpoi ewrpoi rwepoi werpoiwe poi rewpoi rwepoi ewrpoi rewpoi rwepoi rwepoi rwepoi rewpo iwrepoi rwepoi wrepoi wrepoi rwepo irwpo irwepoi rwepo ipoewi rpow ierwe.

Third paragraph zcx,m ncxz,mxczn ,mxczn,mxcz n,mcnxz ,mxczn ,mz ncx,mzcxn.

Fourth paragraph asdlkj dsakljdsalkj dsalkj daslkj daslkj dsalkj daslkj sda [1].

TODO xz,m.cxzm.,czx.,m czx.,m zcx.,m czx.,m czx.,m zcx.

## Results

Let's add some results qweoi qewoiewqoip peqwpoi ewqpoi ewqpoieqwipo ewqipo eqwpio eqwipo ewqoip ewqoip ewqioip eqw.

Isn't Figure 1 interesting! Table 2 is pretty interesting too.

Mutation			Population allele frequency (%)									LD ( $D'$ )	
Position <sup>1</sup>	$Ag^2$	$Md^3$	AOAc	BFAc	GNAg	BFAg	CMAg	GAAg	UGAg	KE	GW	L995F	L995S
2,390,177 G>A	R254K	R261	0	0	0	0	32	21	0	0	0	NA	NA
2,391,228 G>C	V402L	V410	0	7	0	0	0	0	0	0	0	NA	NA
2,391,228 G>T	V402L	V410	0	7	0	0	0	0	0	0	0	NA	NA
2,399,997 G>C	D466H	-	0	0	0	0	7	0	0	0	0	NA	NA
2,400,071 G>A	M490I	M508	0	0	0	0	0	0	0	18	0	NA	NA
2,400,071 G>T	M490I	M508	0	0	0	0	0	0	0	0	0	NA	NA
2,416,980 C>T	T791M	T810	0	1	13	14	0	0	0	0	0	NA	NA
2,422,651 T>C	L995S	L1014	0	0	0	0	15	64	100	76	0	NA	NA
2,422,652 A>T	L995F	L1014	86	85	100	100	53	36	0	0	0	NA	NA
2,424,384 C>T	A1125V	K1133	9	0	0	0	0	0	0	0	0	NA	NA
2,425,077 G>A	V1254I	I1262	0	0	0	0	0	0	0	0	5	NA	NA
2,429,617 T>C	I1527T	I1532	0	14	0	0	0	0	0	0	0	NA	NA
2,429,745 A>T*	N1570Y	N1575	0	26	10	22	6	0	0	0	0	NA	NA
2,429,897 A>G	E1597G	E1602	0	0	6	4	0	0	0	0	0	NA	NA
2,429,915 A>C	K1603T	K1608	0	5	0	0	0	0	0	0	0	NA	NA
2,430,424 G>T	A1746S	A1751	0	0	11	13	0	0	0	0	0	NA	NA
2,430,817 G>A	V1853I	V1858	0	0	8	5	0	0	0	0	0	NA	NA
2,430,863 T>C	I1868T	I1873	0	0	18	25	0	0	0	0	0	NA	NA
2,430,880 C>T	P1874S	P1879	0	21	0	0	0	0	0	0	0	NA	NA
2,430,881 C>T	P1874L	P1879	0	7	45	26	0	0	0	0	0	NA	NA
2,431,061 C>T	A1934V	A1939	0	12	0	0	0	0	0	0	0	NA	NA
2,431,079 T>C	I1940T	I1945	0	4	0	0	7	0	0	0	0	NA	NA

**Table 1. Non-synonymous mutations in the voltage-gated sodium channel gene.** All mutations are at 5% frequency or above in one or more of the 9 Ag1000G phase 1 populations, with the exception of 2,400,071 G>T which is only found in the CMAg population at 0.4% frequency but is included because another mutation (2,400,071 G>A) is found at the same position causing the same amino acid substitution (M490I). Substitutions marked with an asterisk (\*) failed conservative variant filters applied genome-wide in the Ag1000G phase 1 AR3 callset, but appeared sound on manual inspection of read alignments.

<sup>1</sup>Position relative to AgamP3 reference sequence, chromosome arm 2L.

<sup>2</sup>Codon numbering according to transcript AGAP004707-RA in geneset AgamP4.4.

<sup>3</sup>Codon numbering according to *Musca Domestica Vgsc* EMBL accession X96668 (Williamson *et al.*, 1996).



**Figure 1.** Demo figure.

TODO

Foo	Bar	Baz
1	a	True
2	b	False

**Table 2.** This is a table.

## Discussion

TODO

## Methods

TODO

## References

- [1] Nandita R. Garud et al. ‘Recent Selective Sweeps in North American *Drosophila melanogaster* Show Signatures of Soft Sweeps’. In: *PLoS Genetics* 11.2 (2015), pp. 1–32. issn: 15537404. arXiv: 1303.0906.