Spatially-explicit genomics of *An. gambiae s.l* reveals fine-scale population structure and mechanisms of insecticide resistance

Sanjay C. Nagi, John Essandoh, Alex Egyir-Yawson, Eric R. Lucas, Nicholas Ato-Egyir, Chris Clarkson
Kwame Desewu, Alistair Miles, Luigi Sedda, David Weetman, Martin J. Donnelly

Abstract

Progress in malaria control in sub-Saharan Africa is stalling, partly due to the spread of insecticide resistance in the Anopheline vector. Monitoring the evolution of insecticide resistance alleles and their spatial heterogeneity is important for malaria control programmes, and genomic surveillance has emerged as a pivotal tool for this purpose. Earlier genomics research has typically employed convenience-based sampling, and research has yet to be performed to optimise sampling regimens for malaria vector genomics. In this study, we applied a spatially explicit sampling framework that considers the underlying ecology to enable sampling mosquitoes with reduced bias.

Methods

- Mosquitoes were collected using CDC light traps in Obuasi, Ghana (Oct-Dec 2018) following an ecological sampling framework (Sedda et al., 2019), and underwent 30X Illumina whole-genome sequencing at the Wellcome Sanger institute
- Initial bioinformatics analyses performed by MalariaGEN Vector Observatory (Variant calling, QC, Phasing, CNV calling)
- Data was accessed through the malariagen_data python package

Results

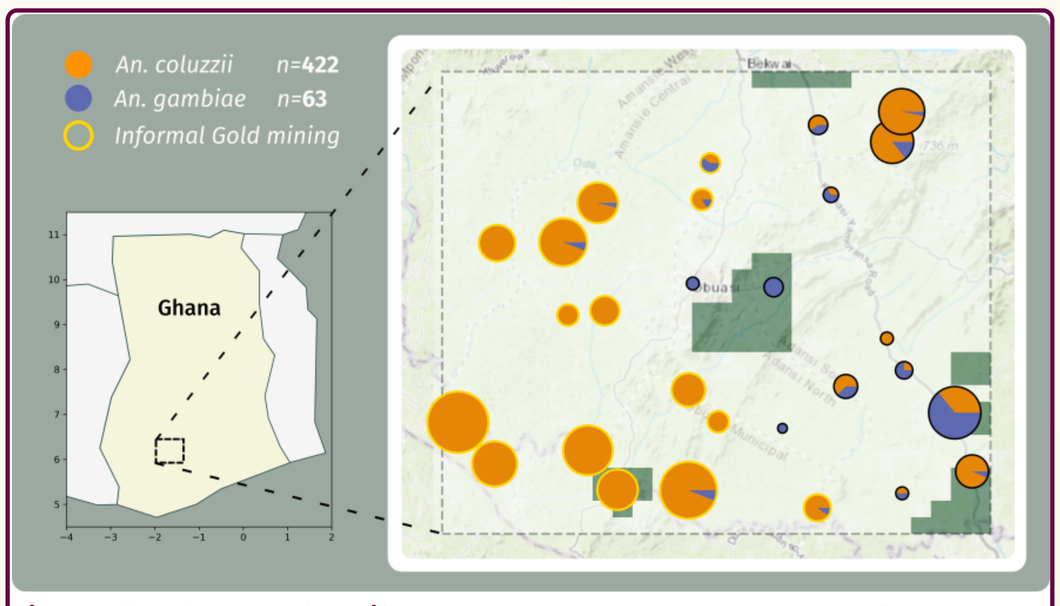

Figure 1. Sample map and manifest. A map of sampling locations from the Obuasi region, overlaid onto the ecological classifications of Sedda et al., 2019.

- Spatially-explicit sampling framework implemented to ensure collections of mosquitoes from different ecotypes
- Majority of sampled mosquitoes were *An. coluzzii*

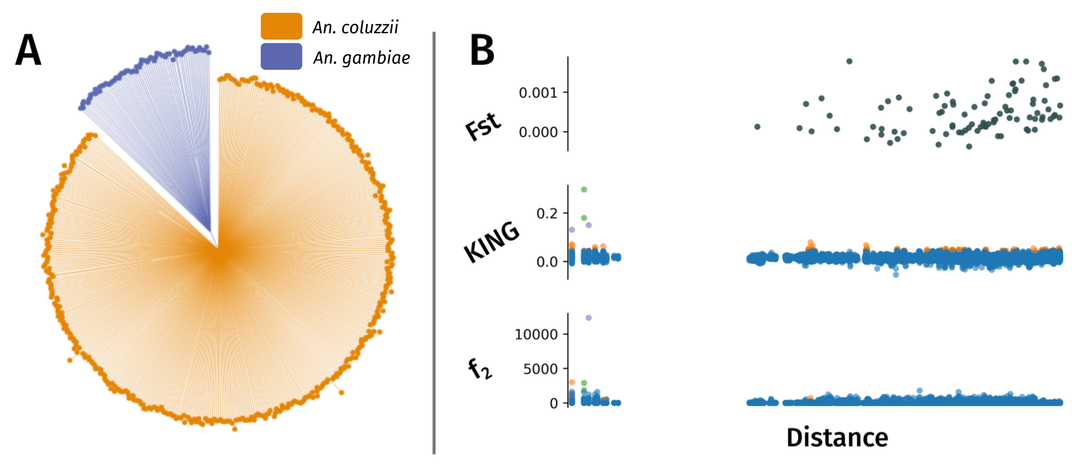

Figure 2. Population structure and isolation-by-distance. A) An unrooted neighbour-joining tree computed from pairwise genetic distances using biallelic SNP variants from the 3L chromosomal arm (3L:15,000,000-44,000,000). B) Measures of genetic distance vs log geographic distance in kilometres. Fst is calculated between sample collections sites and only for *An. coluzzii*. KING is a measure of genetic relatedness based on linkage disquilibrium. f2 = proxy for rare variants

- Little population structure, however, we were able to detect fine-scale isolation-by-distance (IBD). We found that distance, rather than ecology, drives IBD at these spatial scales.
- We found close-kin in villages 7km apart. Estimation of kinship is hindered by the presence of chromosomal inversions. Kin in mosquitoes have high variance in genetic relatedness.

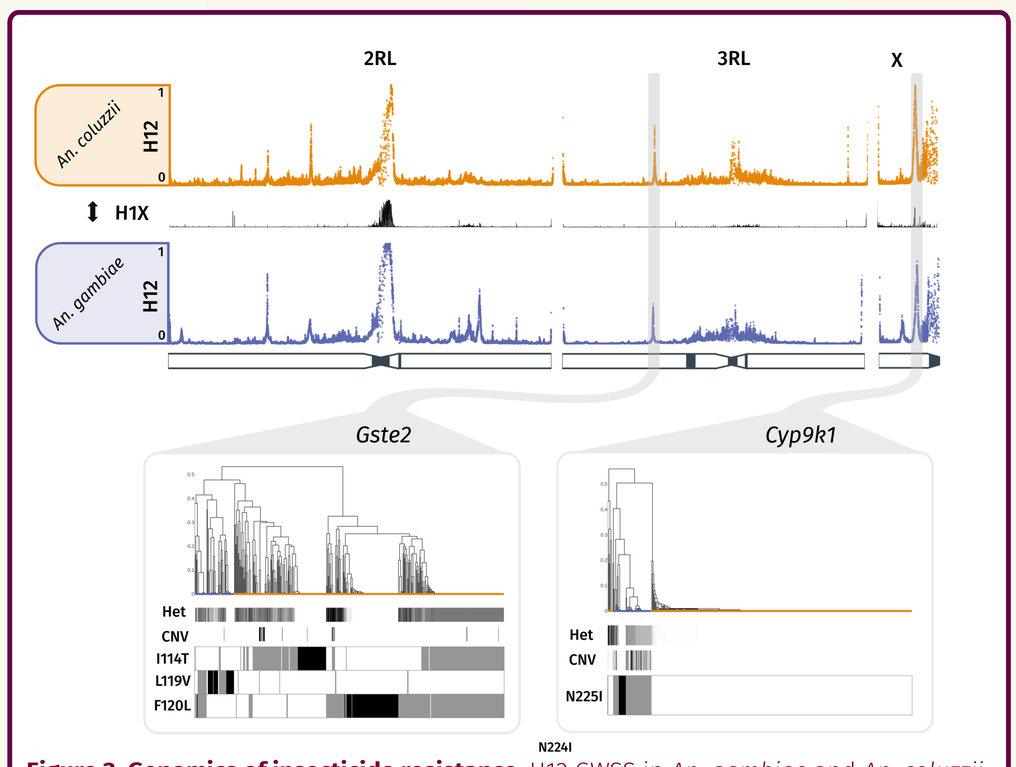

Figure 3. Genomics of insecticide resistance. H12 GWSS in *An. gambiae* and *An. coluzzii* from Obuasi, are separated by a H1X introgression scan. A schematic of the An. gambiae PEST chromosomes is shown beneath the scans, with heterochromatic highlighted in black. Diplotype clustering at two loci, *Gste2* and *Cyp9k1* is shown.

- Using diplotype clustering, we find novel resistance variants driving selective sweeps:
  - *Cyp9k1*-N224I, which modelling shows sits at the P450 substrate channel entrance.
  - *Cyp6aa1*-D155N
  - *Gste2*-F120L
- H1X shows a signal of introgression at the *Vgsc* (L995F). The alternative Kdr haplotype V402L and I1527T is at 14% frequency in *An. coluzzii*.
- CNVs at Cyp6p locus and Cyp9k1 at high frequency
- Novel selective sweeps

Summary

- Fine-scale population structure driven by geographic distance rather than ecology at this scale
- Segregating chromosomal inversions hinder kinship estimation
- Diplotype clustering reveals novel resistance mutations for future surveillance
- Strategic sampling enhances genomic surveillance