Tracking the spread of insecticide resistance in *Anopheles gambiae* populations

Alistair Miles (@alimanfoo)

MalariaGEN Resource Centre

University of Oxford / Wellcome Sanger Institute

6 June 2019 - WHO/Global Malaria Programme

These slides: http://alimanfoo.github.io/slides/20190606-who-geneva.html

Use cases for genomic surveillance of malaria vector populations

Pyrethroid resistance is widespread in primary African vector species.

How can surveillance improve insecticide resistance management (IRM)?

Use case (1): Procurement and deployment of next-generation LLINs

- "Next-generation" LLINs available, e.g.:
  - Olyset Plus: permethrin + PBO (P450 synergist)
  - Olyset Duo: permethrin + pyriproxyfen (second insecticide)
- More expensive than standard LLINs
- How many to buy?
- Where to deploy them?

Use case (2): IRS deployment strategy

- "Next-generation" IRS formulations available, e.g.:
  - Actellic 300CS: pyrimiphos methyl (organophosphate)
  - SumiShield 50WG: clothianidin (neonicotinoid)
  - Fludora Fusion: deltamethrin + clothianidin
- Preemptive rotation?
- Geographical mosaic?
- Is it working?

Use case (3): Cross-border coordination

- Can countries take decisions in isolation about how to manage insecticide resistance?
- When and where do decisions need to be coordinated across borders?

The *Anopheles gambiae* 1000 Genomes Project (Ag1000G)

- A consortial project using whole-genome sequencing to investigate genetic variation and evolution in natural mosquito populations
- Create an open access data resource to accelerate research and surveillance
- www.malariagen.net/ag1000g

Ag1000G Consortium

Wellcome Sanger Institute / University of Oxford / Liverpool School of Tropical Medicine / Sapienza University of Rome / University of California, Riverside / Liverpool John Moores University / Broad Institute / Institut de Recherche pour le Développement / Virginia Tech / KEMRI Wellcome Trust Research Programme / New Mexico State University / Universidade Nova de Lisboa / University of Minnesota / Université d'Abomey–Calavi, Benin / Indiana University / University of Notre Dame / Washington State University / Imperial College / University of Oregon / University of North Carolina at Chapel Hill / University of Montana / Institut Pasteur / Instituto Nacional de Saúde Pública, Guiné-Bissau / Centre International de Recherches Médicales de Franceville, Gabon / Programa Nacional de Controle da Malária, Angola / Institut de Recherche en Sciences de la Santé, Burkina Faso / University of Bamako, Mali / Infectious Diseases Research Collaboration, Uganda / Organisation de Coordination pour la lutte contre les Endémies en Afrique Centrale, Cameroon

Ag1000G sequencing methods

- Sequence individual mosquitoes collected from the field
- Use whole-genome Illumina (Hi-Seq) sequencing
- Deep coverage (~30X)
- Sequencing performed at and funded by Wellcome Sanger Institute

Ag1000G population sampling

- Aim for broad geographical coverage
  - 18 countries, ~1 site per country
- *An. gambiae*, *An. coluzzii*, *An. arabiensis*
- Sequence >30 individuals per site per species
  - Why 30? Statistical power to make inferences about populations (e.g., gene flow).

Ag1000G data production

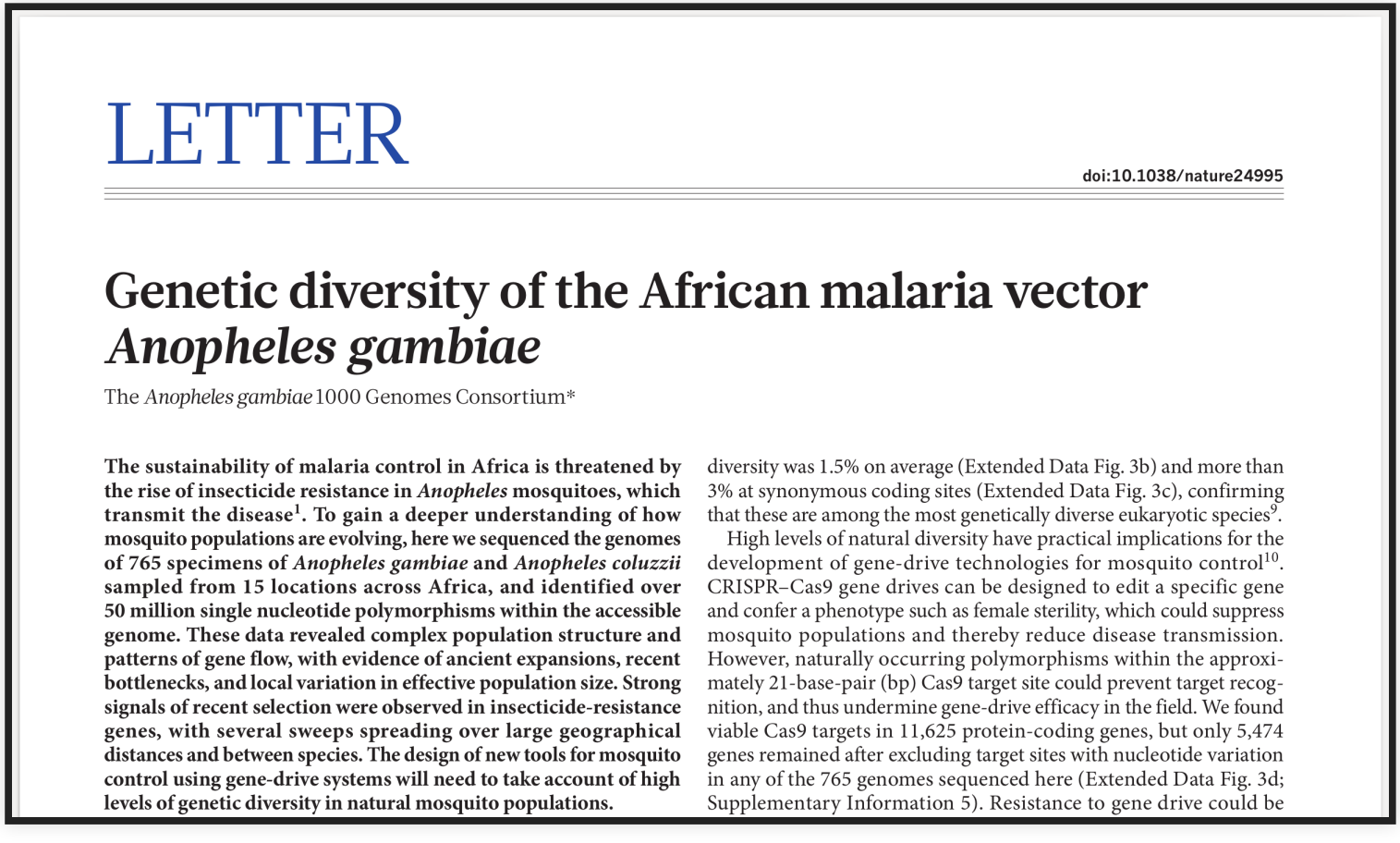
- Raw sequence reads →
- Alignment to reference genome →
- Variant calling →
- Variant filtering and annotation →
- Haplotype phasing →
- **Curated "analysis-ready" variant calls and haplotypes**

Ag1000G data releases

- **Phase 1**: 765 mosquitoes; 8 countries; *An. gambiae*, *An. coluzzii*
  - Data released 2016
- **Phase 2**: 1,142 mosquitoes; 13 countries; *An. gambiae*, *An. coluzzii*
  - Data released 2017
- **Phase 3**: ~4,000 mosquitoes; 18 countries; *An. gambiae*, *An. coluzzii*, *An. arabiensis*
  - Data in production

Ag1000G further information

- www.malariagen.net/ag1000g
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6026



LETTER

doi:10.1038/nature24995

Genetic diversity of the African malaria vector *Anopheles gambiae*

The *Anopheles gambiae* 1000 Genomes Consortium*

The sustainability of malaria control in Africa is threatened by the rise of insecticide resistance in *Anopheles* mosquitoes, which transmit the disease[1]. To gain a deeper understanding of how mosquito populations are evolving, here we sequenced the genomes of 765 specimens of *Anopheles gambiae* and *Anopheles coluzzii* sampled from 15 locations across Africa, and identified over 50 million single nucleotide polymorphisms within the accessible genome. These data revealed complex population structure and patterns of gene flow, with evidence of ancient expansions, recent bottlenecks, and local variation in effective population size. Strong signals of recent selection were observed in insecticide-resistance genes, with several sweeps spreading over large geographical distances and between species. The design of new tools for mosquito control using gene-drive systems will need to take account of high levels of genetic diversity in natural mosquito populations.

diversity was 1.5% on average (Extended Data Fig. 3b) and more than 3% at synonymous coding sites (Extended Data Fig. 3c), confirming that these are among the most genetically diverse eukaryotic species[9].

High levels of natural diversity have practical implications for the development of gene-drive technologies for mosquito control[10]. CRISPR–Cas9 gene drives can be designed to edit a specific gene and confer a phenotype such as female sterility, which could suppress mosquito populations and thereby reduce disease transmission. However, naturally occurring polymorphisms within the approximately 21-base-pair (bp) Cas9 target site could prevent target recognition, and thus undermine gene-drive efficacy in the field. We found viable Cas9 targets in 11,625 protein-coding genes, but only 5,474 genes remained after excluding target sites with nucleotide variation in any of the 765 genomes sequenced here (Extended Data Fig. 3d; Supplementary Information 5). Resistance to gene drive could be

Pyrethroid target-site resistance

Spread of "knock-down resistance" (*kdr*) mutations in the voltage-gated sodium channel gene (*Vgsc*).

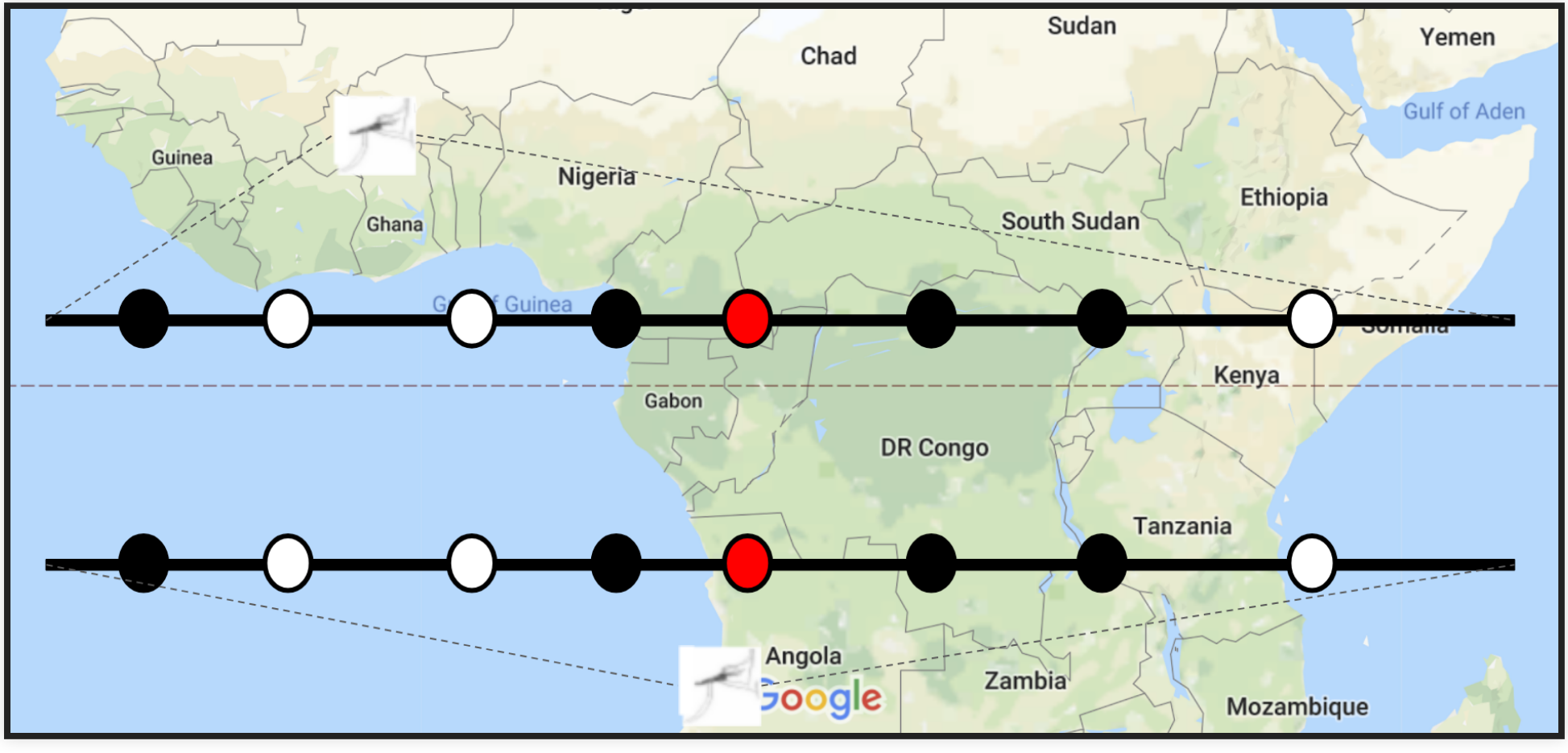
kdr mutations

- Two known *kdr* mutations in *Vgsc* codon 1014
- L1014F found throughout West and Central Africa
- L1014S found throughout East and Central Africa
- Are these mutations spreading?
- Where is gene flow occurring?

Inferring *kdr* gene flow

- Analyse the genetic backgrounds on which *kdr* mutations occur ("*kdr* haplotypes")
- Use all mutations within the *Vgsc* gene
  - 1,710 biallelic SNPs (mostly intronic)
- Same *kdr* haplotype in two different locations:
  - ⇒ gene flow

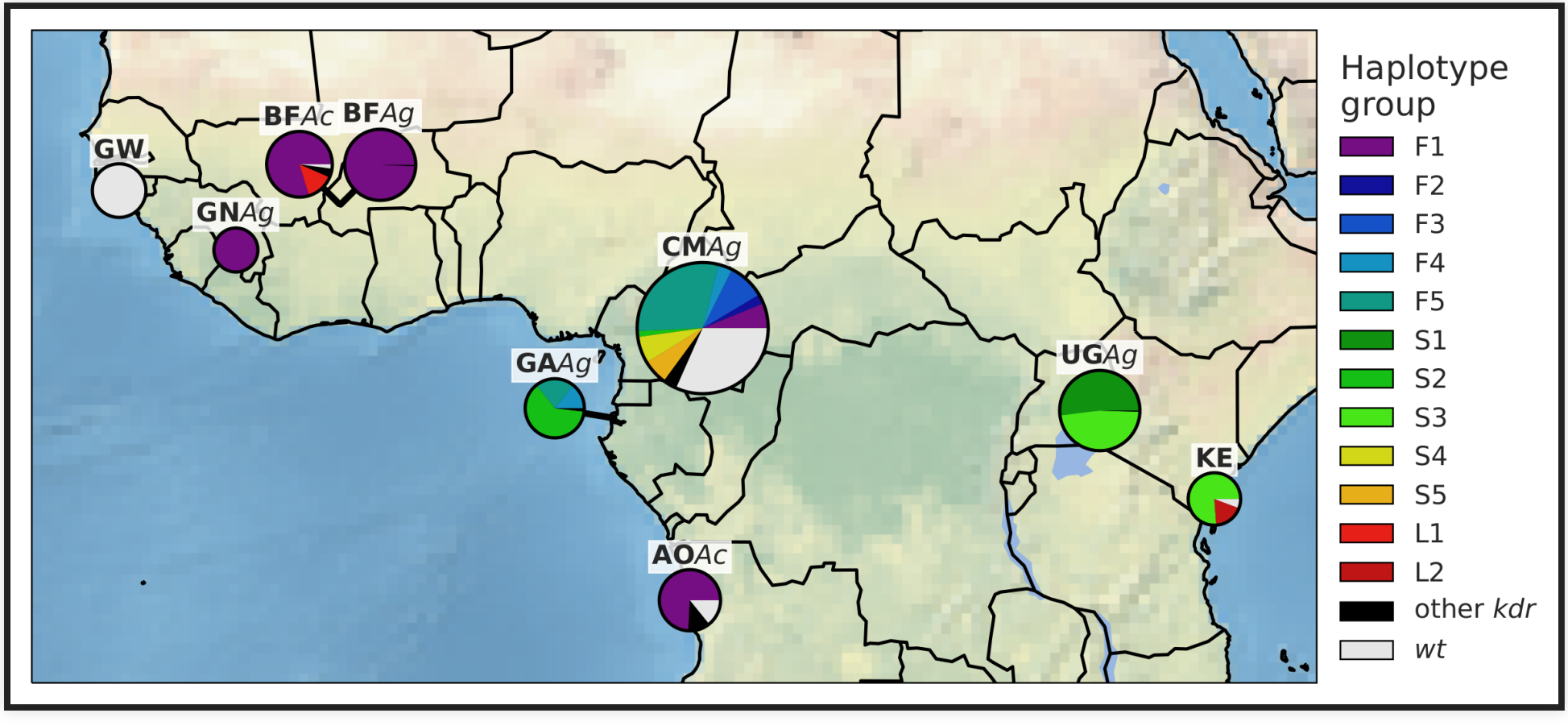
Inferring *kdr* gene flow



kdr haplotypes

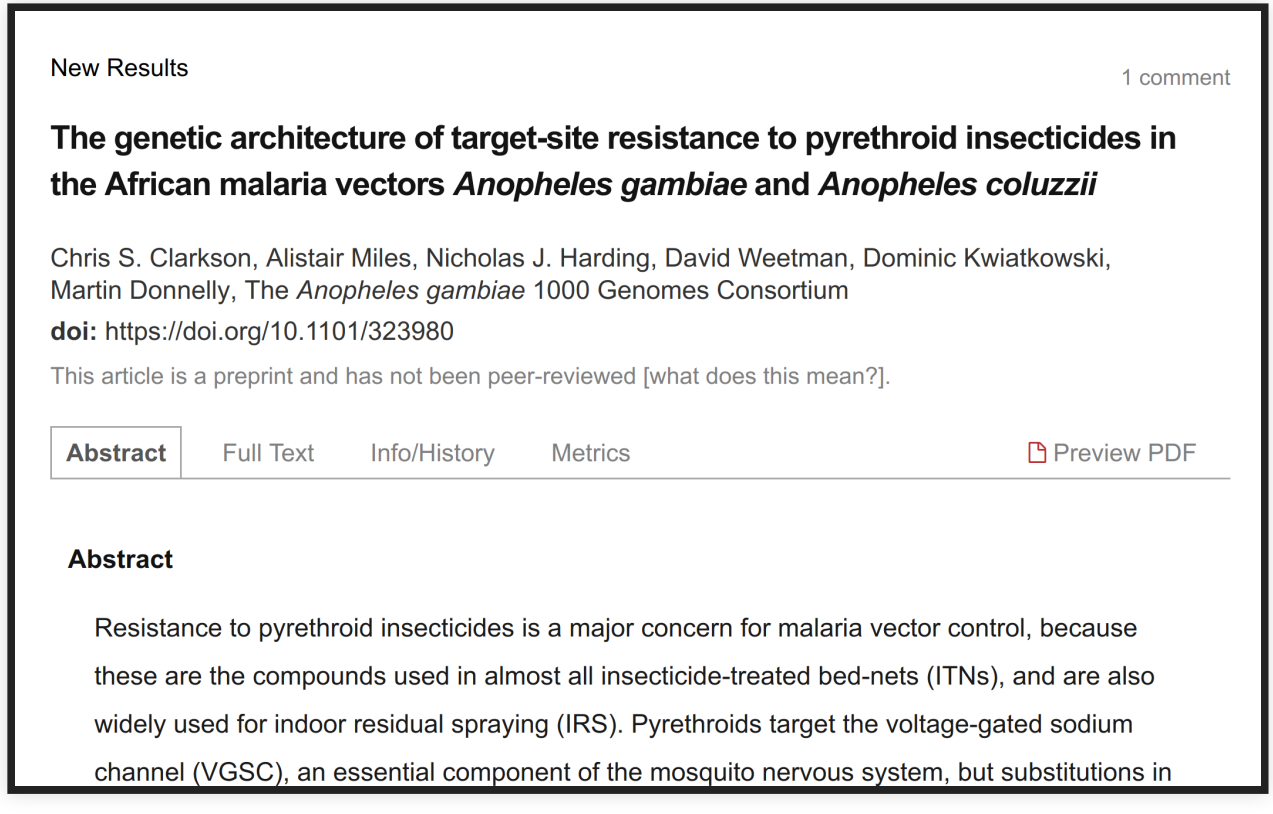
- Analysed data from Ag1000G phase 1 (765 mosquitoes, 8 countries)
- L1014F - 5 major haplotypes (F1-F5)
- L1014S - 5 major haplotypes (S1-S5)

kdr haplotypes



kdr gene flow - further information

https://doi.org/10.1101/323980



New Results

1 comment

The genetic architecture of target-site resistance to pyrethroid insecticides in the African malaria vectors *Anopheles gambiae* and *Anopheles coluzzii*

Chris S. Clarkson, Alistair Miles, Nicholas J. Harding, David Weetman, Dominic Kwiatkowski, Martin Donnelly, The *Anopheles gambiae* 1000 Genomes Consortium

doi: https://doi.org/10.1101/323980

Abstract | Full Text | Info/History | Metrics | Preview PDF

Abstract

Resistance to pyrethroid insecticides is a major concern for malaria vector control, because these are the compounds used in almost all insecticide-treated bed-nets (ITNs), and are also widely used for indoor residual spraying (IRS). Pyrethroids target the voltage-gated sodium channel (VGSC), an essential component of the mosquito nervous system, but substitutions in

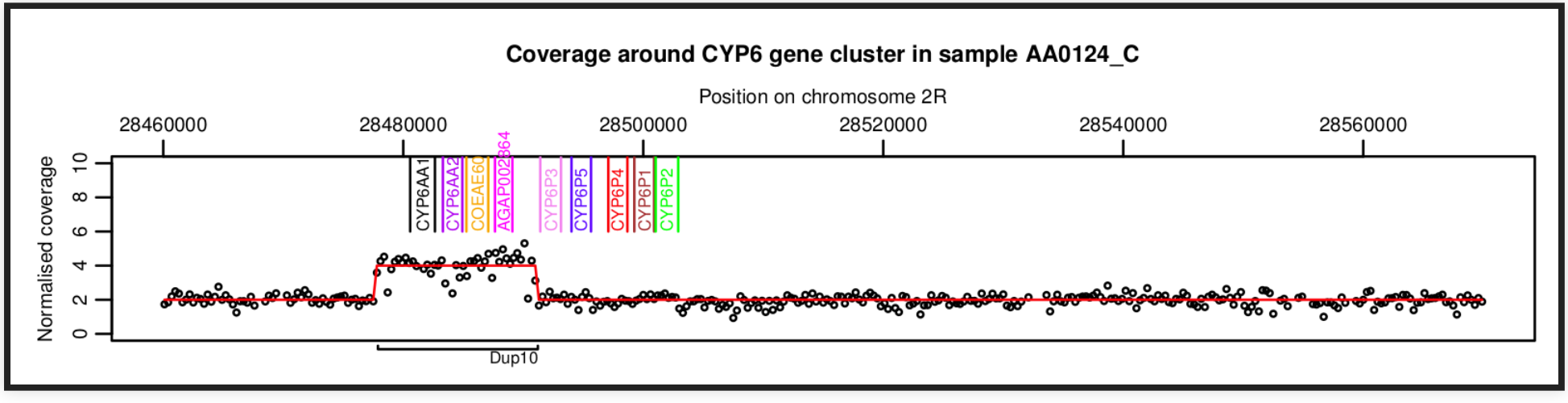
Pyrethroid metabolic resistance

Spread of copy number variations in cytochrome P450 genes.

Cytochrome P450 genes

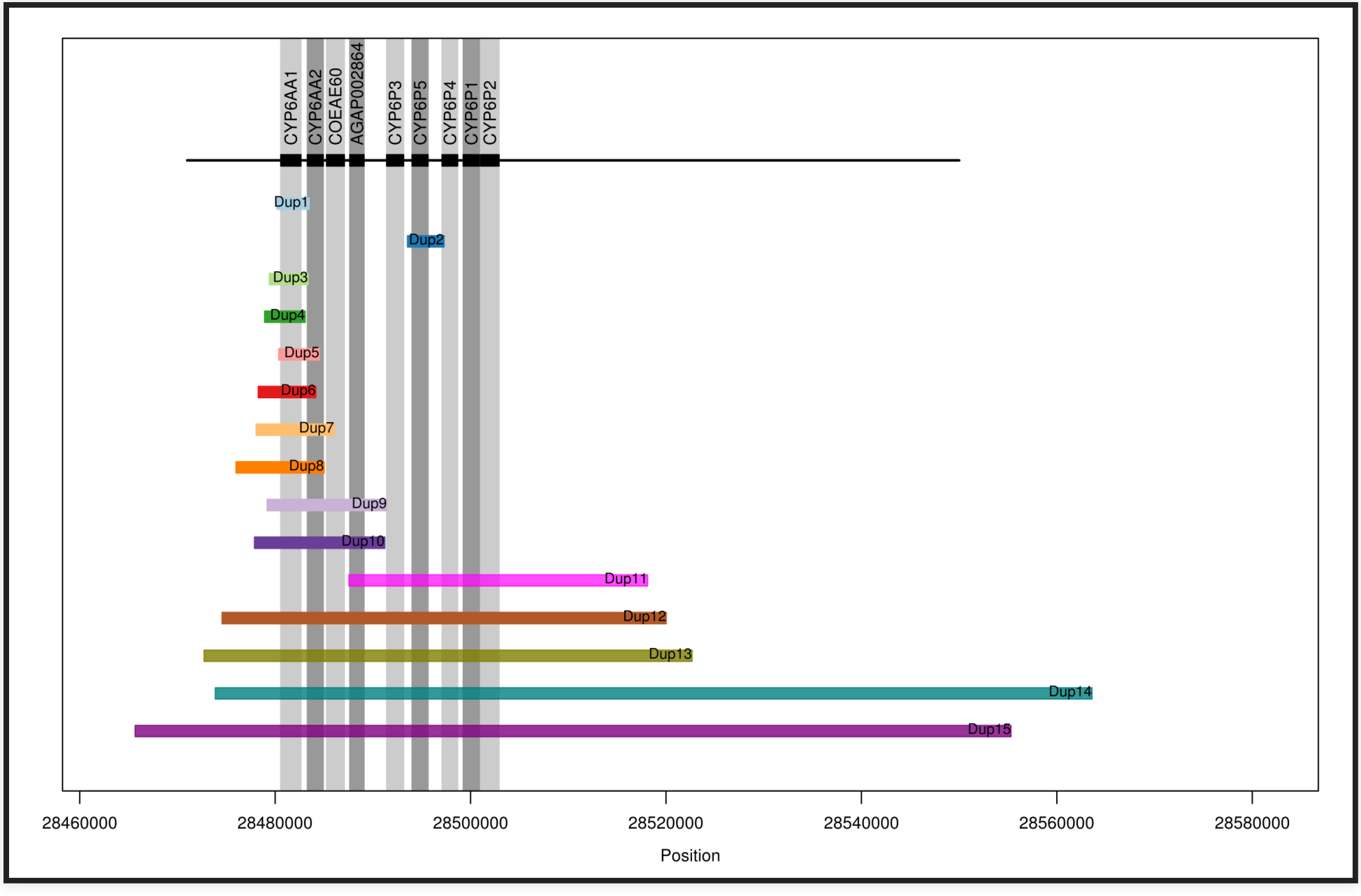
- Known to metabolise pyrethroids
- Neutralised by PBO synergist in next-gen LLINs
- Multiple P450 genes in genome, e.g.:
  - *Cyp6p/aa*
  - *Cyp9k1*
- Increased expression ⇒ pyrethroid resistance
- Increased gene copy number ⇒ increased expression

Detecting copy number variation



P450 copy number variation

- Analysed data from Ag1000G phase 2 (1,142 mosquitoes, 13 countries)
- Gene amplifications are common at two P450 loci:
  - *Cyp6p/aa*
  - *Cyp9k1*


CYP6AA1
CYP6AA2
COEAE60
AGAP002864
CYP6P3
CYP6P5
CYP6P4
CYP6P1
CYP6P2
Dup1
Dup2
Dup3
Dup4
Dup5
Dup6
Dup7
Dup8
Dup9
Dup10
Dup11
Dup12
Dup13
Dup14
Dup15
28460000
28480000
28500000
28520000
28540000
28560000
28580000
Position

Cyp6p/aa CNV gene flow

- Dup1 - BFcol (8%), UGgam (58%)
- Dup7 - BFcol (44%), CIcol (32%), GHcol (5%), GNcol (75%)
- Dup8 - BFgam (3%), GNgam (3%)
- Dup10 - BFcol (49%), GHcol (5%)
- Dup11 - CIcol (41%), GHcol (5%)
- Dup14 - BFcol (3%), CIcol (46%)
- Dup15 - BFcol (1%), CIcol (39%)

CNVs further information

https://doi.org/10.1101/399568

New Results

Comment on this paper

Whole genome sequencing reveals high complexity of copy number variation at insecticide resistance loci in malaria mosquitoes

Eric R. Lucas, Alistair Miles, Nicholas J. Harding, Chris S. Clarkson, Mara K. N. Lawniczak, Dominic P. Kwiatkowski, David Weetman, Martin J. Donnelly,
The Anopheles gambiae 1000 Genomes Consortium

doi: https://doi.org/10.1101/399568

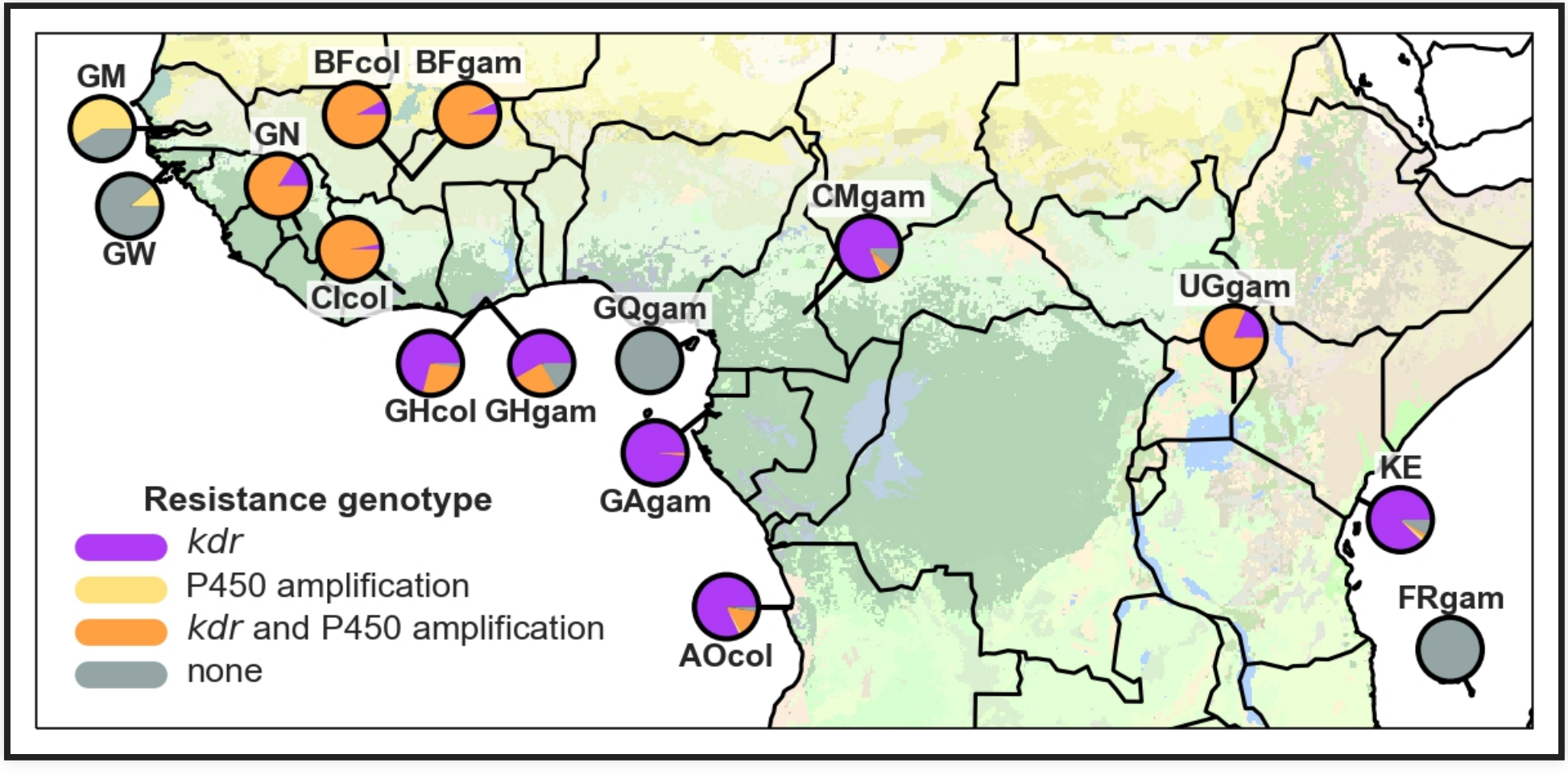
Abstract

Background Polymorphisms in the copy number of a genetic region can influence gene expression, coding sequence and zygosity, making them powerful actors in the evolutionary process. Copy number variants (CNVs) are however understudied, being more difficult to

Summary & discussion

- Target-site (*kdr*) and metabolic (P450 CNV) pyrethroid resistance are spreading via gene flow
- Multiple independent outbreaks of resistance
  - Some spreading, some localised
- Long distance gene flow, e.g.:
  - *kdr*-F1 found in GN, BF, CM and AO
  - *Cyp6p*-Dup1 found in BF and UG
- *kdr*, *Cyp6p/aa* and *Cyp9k1* show different patterns of spread

Where to deploy PBO LLINs?



Open questions

- Geographical origins?
- Direction and routes of gene flow?
- Timing?
- Rate of movement?

Next steps

- Scale up genome sequencing of vector populations
- Increase geographical coverage
- Regular (seasonal) sampling
- Other vector species (e.g., *An. funestus*)

MalariaGEN Vector Observatory

- Aim to sequence 10,000 mosquitoes per year
- Coupled with routine ento surveillance
  - Follow sentinel sites over time
  - Link genomic and epi/ento data
- Partnerships
  - PAMCA/BMGF, GAARDian, Target Malaria, ...
- Open data

The Anopheles gambiae 1000 Genomes Consortium*

▾ Author information ▸ Copyright and License information Disclaimer

*Lists of participants and their affiliations appear at the end of the paper

Corresponding authors. Alistair Miles[1,2], Mara K. N. Lawniczak[1], Martin J. Donnelly[3,1], Dominic P. Kwiatkowski[1,2].

Data analysis group. Alistair Miles[1,2] (project lead), Nicholas J. Harding[2], Giordano Bottà[4,2], Chris S. Clarkson[1,3], Tiago Antão[5,3,2], Krzysztof Kozak[1], Daniel R. Schrider[6], Andrew D. Kern[6], Seth Redmond[7], Igor Sharakhov[8,9], Richard D. Pearson[1,2], Christina Bergey[10], Michael C. Fontaine[11], Martin J. Donnelly[3,1], Mara K. N. Lawniczak[1], Dominic P. Kwiatkowski[1,2] (chair).

Partner working group. Martin J. Donnelly[3,1] (chair), Diego Ayala[12,13], Nora J. Besansky[10], Austin Burt[14], Beniamino Caputo[4], Alessandra della Torre[4], Michael C. Fontaine[11], H. Charles J. Godfray[15], Matthew W. Hahn[16], Andrew D. Kern[6], Dominic P. Kwiatkowski[1,2], Mara K. N. Lawniczak[1], Janet Midega[17], Daniel E. Neafsey[7], Samantha O'Loughlin[14], João Pinto[18], Michelle M. Riehle[19], Igor Sharakhov[8,9], Kenneth D. Vernick[20], David Weetman[3], Craig S. Wilding[21,3], Bradley J. White[22].

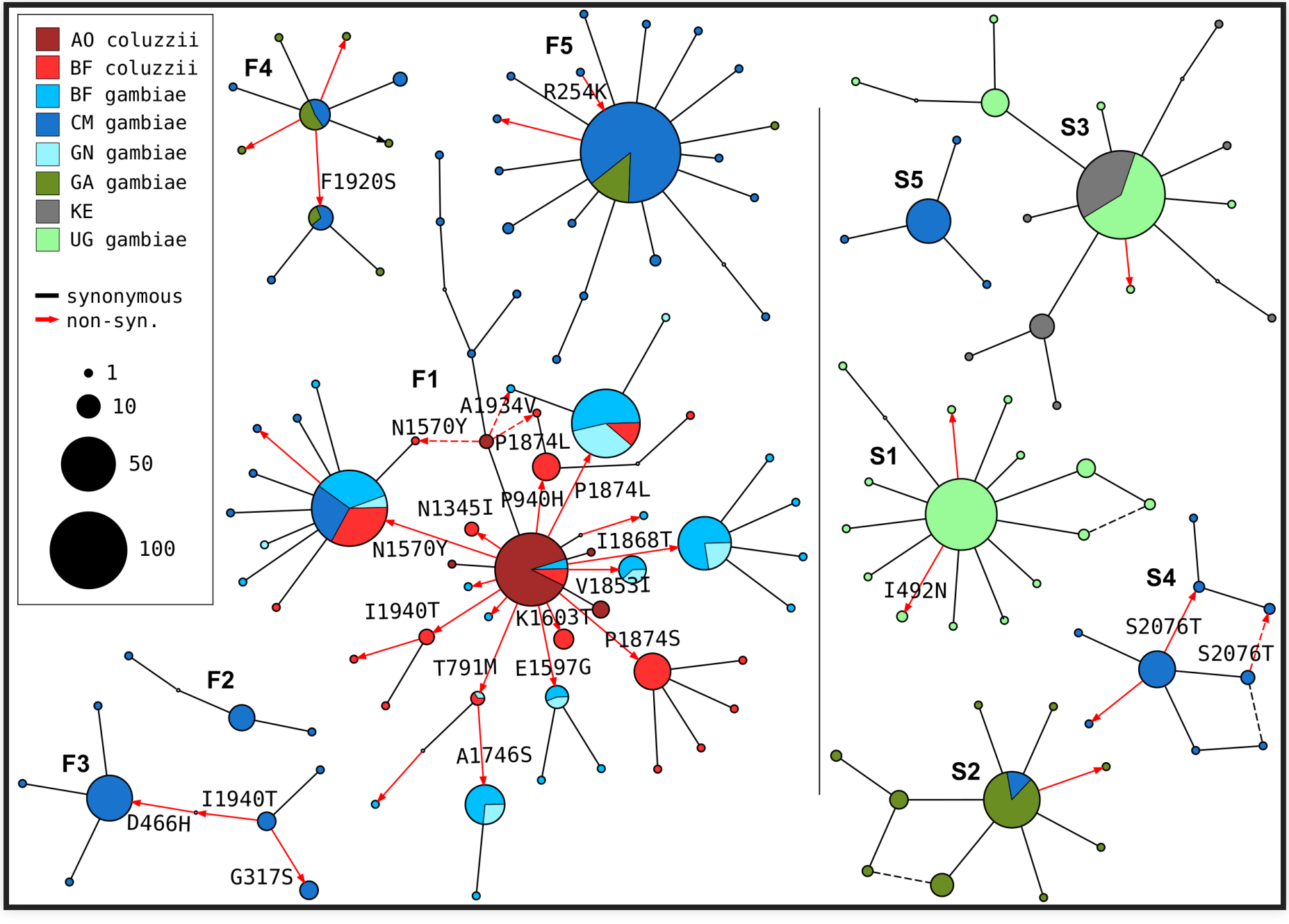
Sample collections. Angola: Arlete D. Troco[23], João Pinto[18]; **Burkina Faso:** Abdoulaye Diabaté[24], Samantha O'Loughlin[14], Austin Burt[14]; **Cameroon:** Carlo Costantini[13,25], Kyanne R. Rohatgi[10], Nora J. Besansky[10]; **Gabon:** Nohal Elissa[12], João Pinto[18]; **Guinea:** Boubacar Coulibaly[26], Michelle M. Riehle[19], Kenneth D. Vernick[20]; **Guinea-Bissau:** João Pinto[18], João Dinis[27]; **Kenya:** Janet Midega[17], Charles Mbogo[17], Philip Bejon[17]; **Uganda:** Craig S. Wilding[21,3], David Weetman[3], Henry D. Mawejje[28], Martin J. Donnelly[3,1]; **Crosses:** David Weetman[3], Craig S. Wilding[21,3], Martin J. Donnelly[3,1].

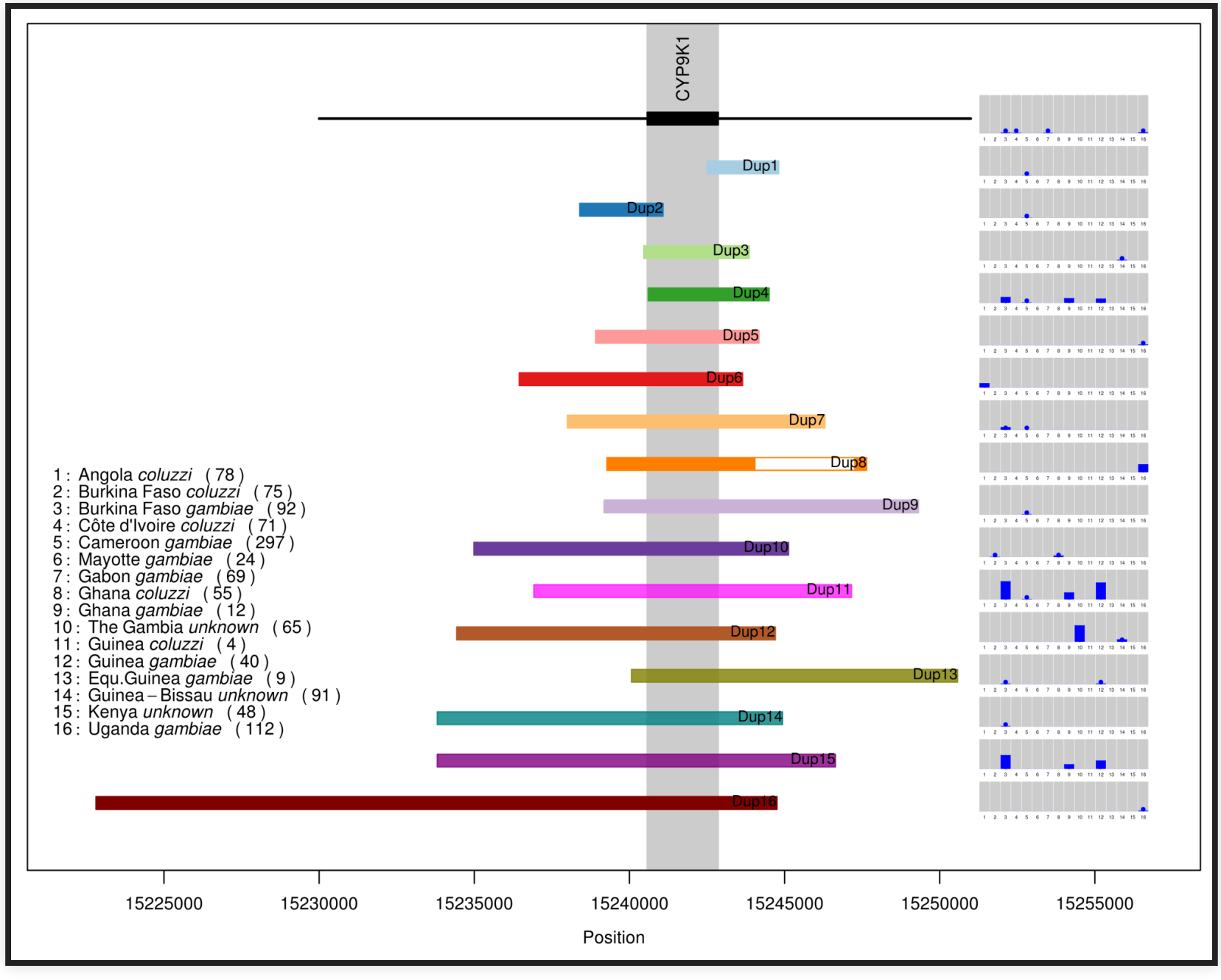
Sequencing and data production. Jim Stalker[1], Kirk Rockett[2], Eleanor Drury[1], Daniel Mead[1], Anna Jeffreys[2], Christina Hubbart[2], Kate Rowlands[2], Alison T. Isaacs[3], Dushyanth Jyothi[1], Cinzia Malangone[1].

Web application development. Paul Vauterin[2], Ben Jeffrey[2], Ian Wright[2], Lee Hart[2], Krzysztof Kluczyński[2].

Project coordination. Victoria Cornelius[2], Bronwyn MacInnis[29], Christa Henrichs[2], Rachel Giacomantonio[1], Dominic P. Kwiatkowski[1,2].

Extra slides


AO coluzzii
BF coluzzii
BF gambiae
CM gambiae
GN gambiae
GA gambiae
KE
UG gambiae
synonymous
non-syn.
1
10
50
100
F4
F1920S
F5
R254K
F1
A1934V
N1570Y
P1874L
N1345I
P940H
P1874L
N1570Y
I1868T
V1853I
I1940T
K1603T
P1874S
T791M
E1597G
A1746S
F2
F3
I1940T
D466H
G317S
S3
S5
S1
I492N
S4
S2076T
S2076T
S2


CYP9K1
Dup1
Dup2
Dup3
Dup4
Dup5
Dup6
Dup7
Dup8
Dup9
Dup10
Dup11
Dup12
Dup13
Dup14
Dup15
Dup16
1 : Angola coluzzi (78)
2 : Burkina Faso coluzzi (75)
3 : Burkina Faso gambiae (92)
4 : Côte d'Ivoire coluzzi (71)
5 : Cameroon gambiae (297)
6 : Mayotte gambiae (24)
7 : Gabon gambiae (69)
8 : Ghana coluzzi (55)
9 : Ghana gambiae (12)
10 : The Gambia unknown (65)
11 : Guinea coluzzi (4)
12 : Guinea gambiae (40)
13 : Equ.Guinea gambiae (9)
14 : Guinea−Bissau unknown (91)
15 : Kenya unknown (48)
16 : Uganda gambiae (112)
15225000
15230000
15235000
15240000
15245000
15250000
15255000
Position

Cyp9k1 CNV gene flow

- Dup4 - BFgam, CMgam, GHgam, GNgam
- Dup7 - BFgam, CMgam
- Dup10 - BFcol, GHcol
- Dup11 - BFgam, CMgam, GHgam, GNgam
- Dup12 - GM, GW
- Dup13 - BFgam, GNgam
- Dup15 - BFgam, GHgam, GNgam