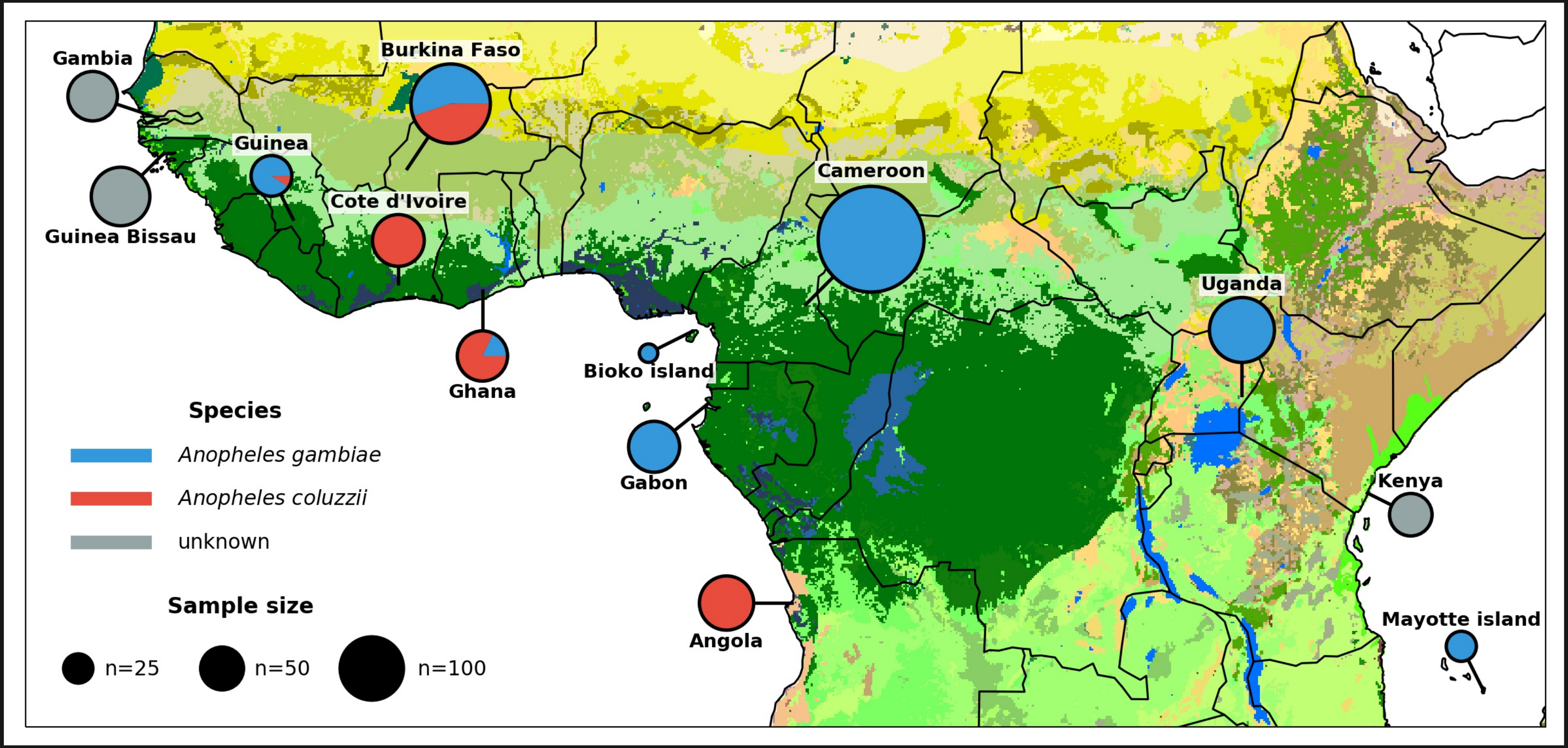
Insights into the evolution and spread of insecticide resistance from whole-genome sequencing of 1,142 *Anopheles gambiae* mosquitoes from 13 countries

Alistair Miles (@alimanfoo) - PAMCA 2019

The *Anopheles gambiae* 1000 Genomes Consortium

These slides: alimanfoo.github.io/slides/pamca2019/

Anopheles gambiae 1000 Genomes Project - phase 2

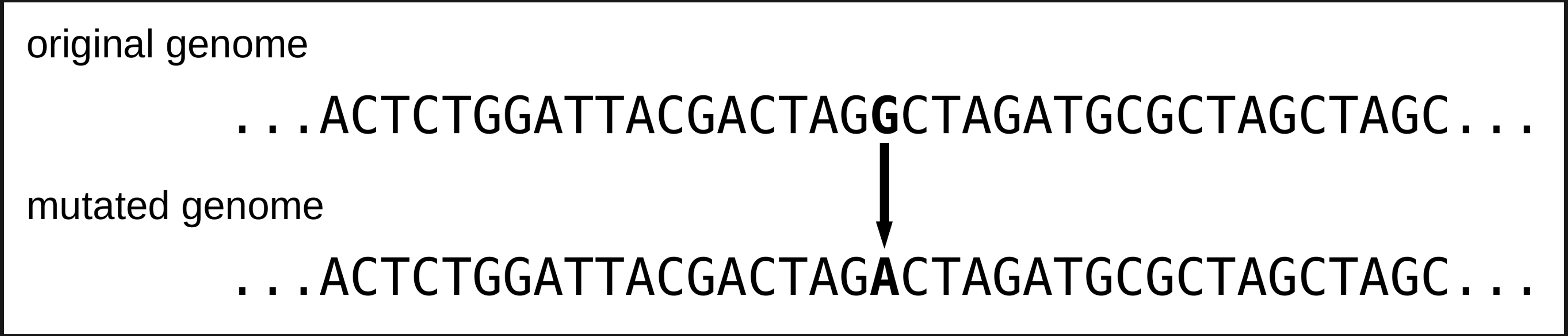


N.B., one location per country; one time point per country; different countries sampled in different years; nothing more recent than 2012.

Methods

- Collect mosquitoes
- Extract DNA
- Whole-genome sequencing (Illumina Hi-Seq)
- Align sequence reads
- Identify genetic differences

Single nucleotide polymorphisms (SNPs)



original genome

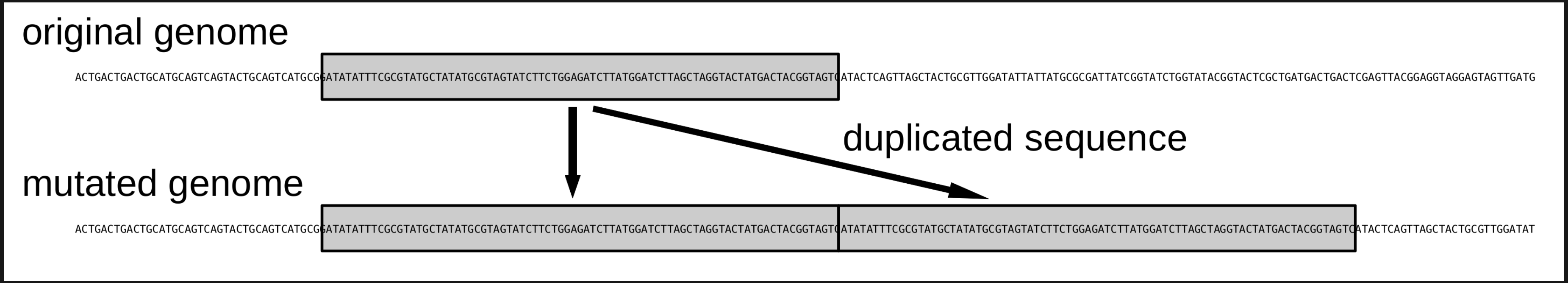
...ACTCTGGATTACGACTAG**G**CTAGATGCGCTAGCTAGC...

mutated genome

...ACTCTGGATTACGACTAG**A**CTAGATGCGCTAGCTAGC...

If a SNP occurs in a gene coding sequence, it may change the encoded amino acid sequence, which may change the behaviour of the protein.

Copy number variants (CNVs)



If a CNV spans a gene, then the gene will occur twice (two copies) in the mutated genome. More gene copies may mean more protein is produced.

Ag1000G phase 2 data resource

Across the whole genome, we find:

- 57,837,885 SNPs
- 31,335 CNVs

Outline

- Pyrethroid resistance
    - Metabolic resistance
    - Target-site resistance
    - Combined view
- Organophosphate resistance
- Emerging/unknown resistance
- Translation to surveillance

Pyrethroid metabolic resistance

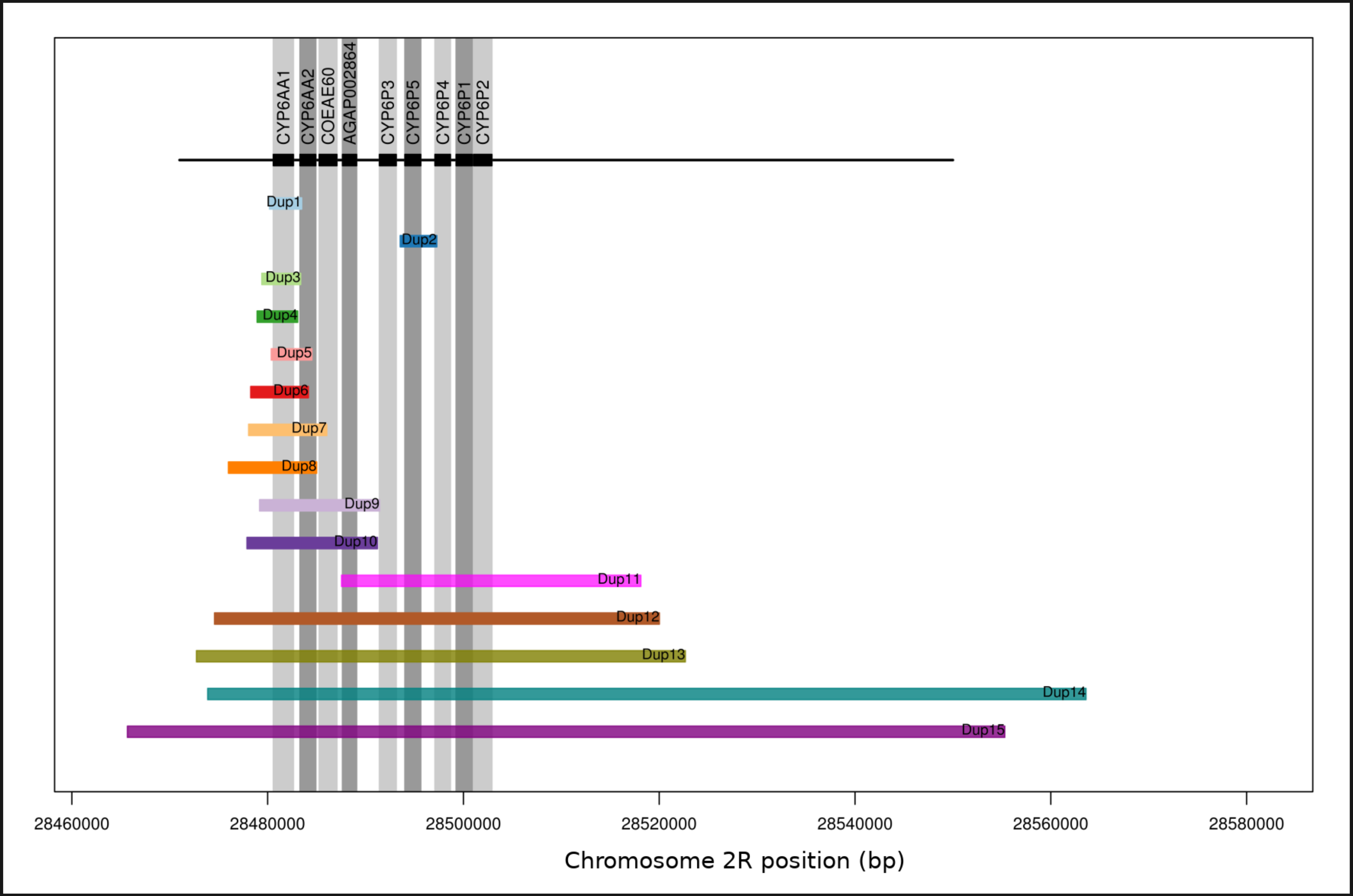
- Pyrethroids are metabolised by cytochrome P450 (CYP) enzymes (a.k.a. mixed-function oxidases; MFOs)
- Increased expression of certain CYP genes causes resistance
- Genetic variants causing metabolic resistance **not previously known** in *An. gambiae* complex
    - Although long suspected that CNVs play a role
    - More gene copies → more protein → faster metabolism → resistance
- N.B., PBO LLINs work by inhibiting CYPs

CNV hot spots

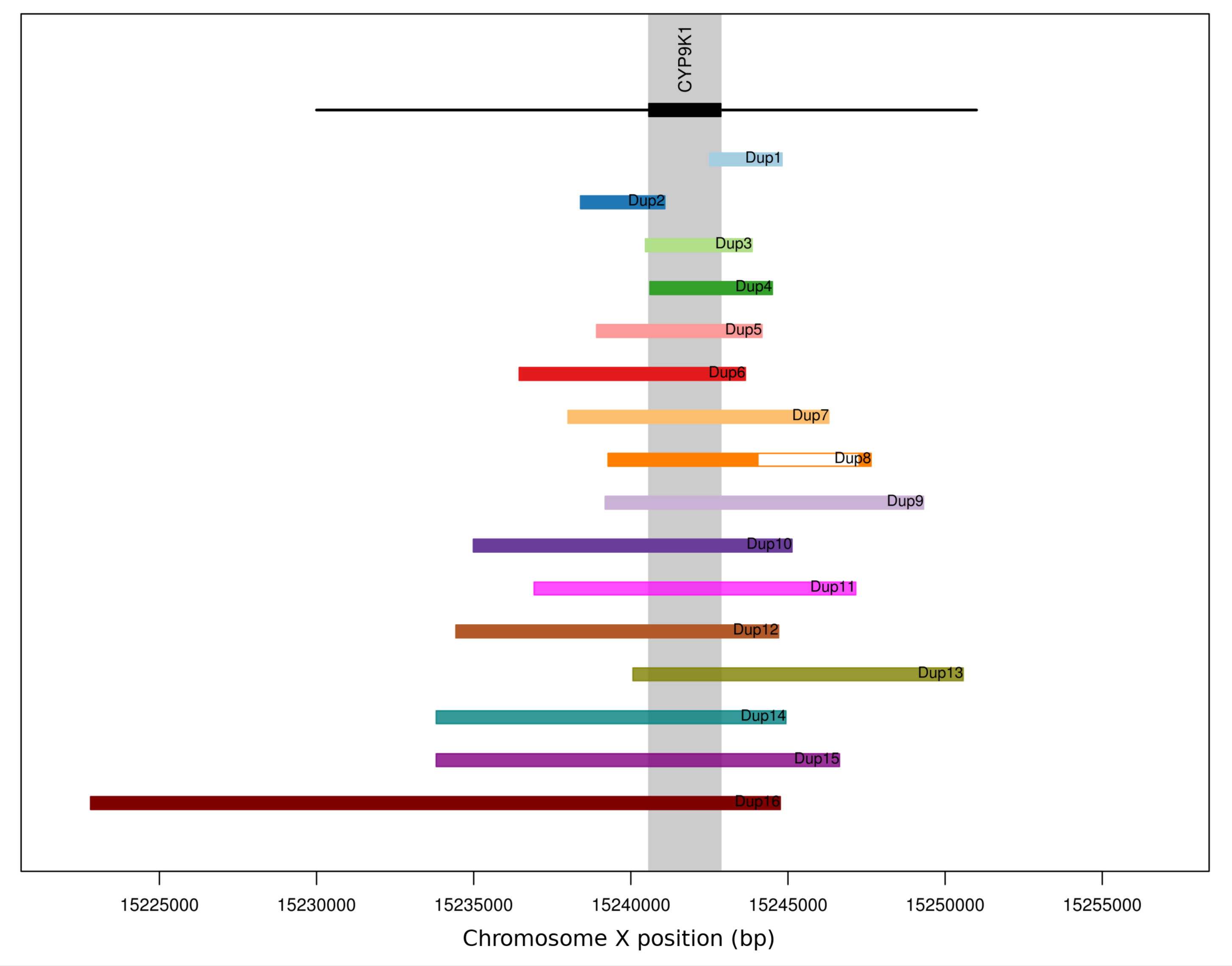
Genome locations where we find a lot of CNVs at high population frequency:

- CYP6P/AA (cytochrome P450 gene cluster)
- CYP9K1 (cytochrome P450 gene)
- GSTE (glutathione S-transferase gene)

CNVs at CYP6P/AA



CNVs at CYP9K1



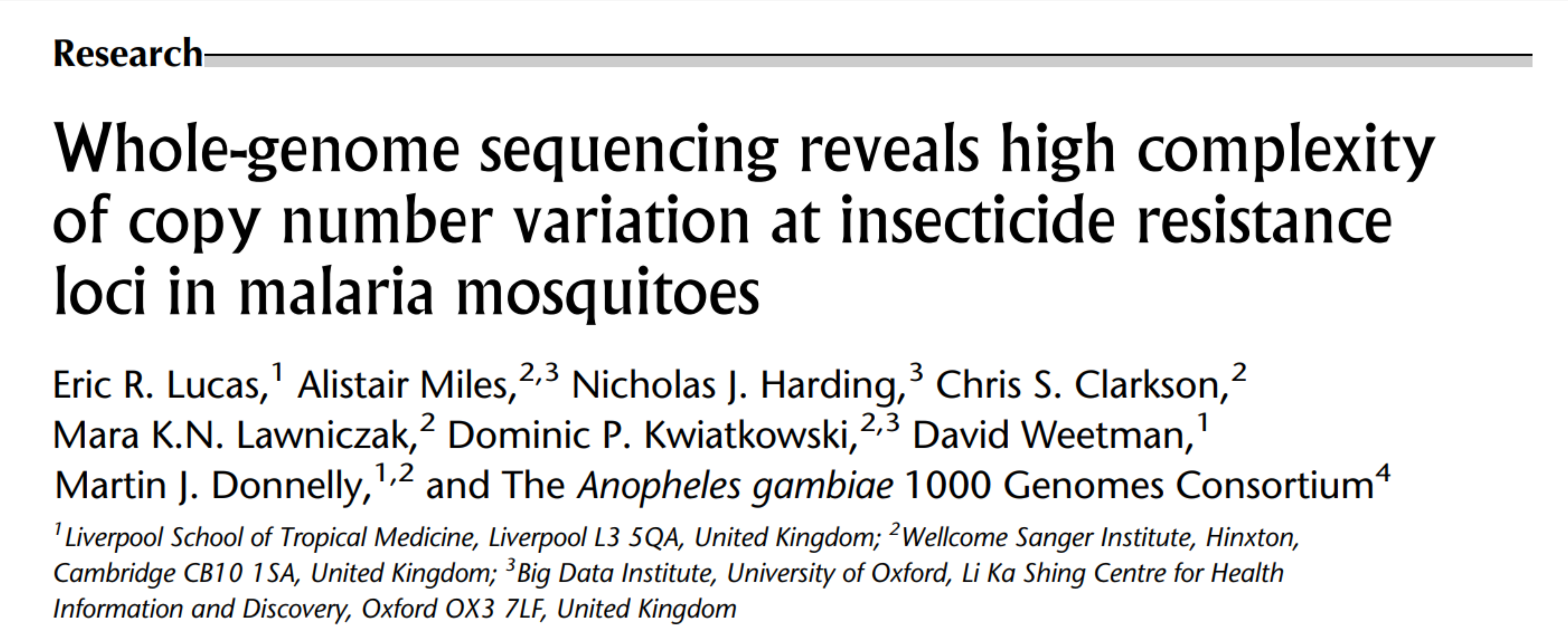
Prevalence and spread of CYP6P/AA CNVs

- Dup1 - *An. gambiae*: Uganda (58%)
- Dup7 - *An. coluzzii*: Burkina Faso (44%), Cote d'Ivoire (32%), Ghana (5%), Guinea (75%)
- Dup10 - *An. coluzzii*: Burkina Faso (49%), Ghana (5%)
- Dup11 - *An. coluzzii*: Burkina Faso (41%), Ghana (5%)
- Dup14 - *An. coluzzii*: Burkina Faso (3%), Cote d'Ivoire (46%)
- Dup15 - *An. coluzzii*: Burkina Faso (1%), Cote d'Ivoire (39%)

Some CNVs are very common and spreading, especially in West African *An. coluzzii*.

Surveillance of metabolic resistance

- Next-generation LLINs: how many, what type, where to put them?
- Which populations have CYP-mediated metabolic resistance?
- CNVs are likely to be a good marker of CYP-mediated metabolic resistance
- Although, N.B., there could be other genetic variants causing metabolic resistance



Research

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

Eric R. Lucas,[1] Alistair Miles,[2,3] Nicholas J. Harding,[3] Chris S. Clarkson,[2] Mara K.N. Lawniczak,[2] Dominic P. Kwiatkowski,[2,3] David Weetman,[1] Martin J. Donnelly,[1,2] and The *Anopheles gambiae* 1000 Genomes Consortium[4]

[1]Liverpool School of Tropical Medicine, Liverpool L3 5QA, United Kingdom; [2]Wellcome Sanger Institute, Hinxton, Cambridge CB10 1SA, United Kingdom; [3]Big Data Institute, University of Oxford, Li Ka Shing Centre for Health Information and Discovery, Oxford OX3 7LF, United Kingdom

See also "Insecticide resistance gene copy number variation in malaria mosquitoes" on

Pyrethroid target-site resistance

- Pyrethroids bind to the voltage-gated sodium channel (VGSC) protein
- SNPs in the VGSC gene can change the protein and cause resistance
- Known resistance SNPs
  - L995F ("*kdr* west")
  - L995S ("*kdr* east")
  - L995F + N1570Y

Prevalence of *kdr* L995F

- *An. gambiae*: Cameroon (53%), Ghana (100%), Burkina Faso (100%), Guinea (100%), Gabon (33%)
- *An. coluzzii*: Angola (84%), Ghana (82%), Burkina Faso (85%), Cote d'Ivoire (91%), Guinea (88%)

Prevalence of *kdr* L995S

- *An. gambiae*: Cameroon (16%), Gabon (67%), Uganda (100%), Kenya (76%)

Prevalence of *kdr* L995F + N1570Y

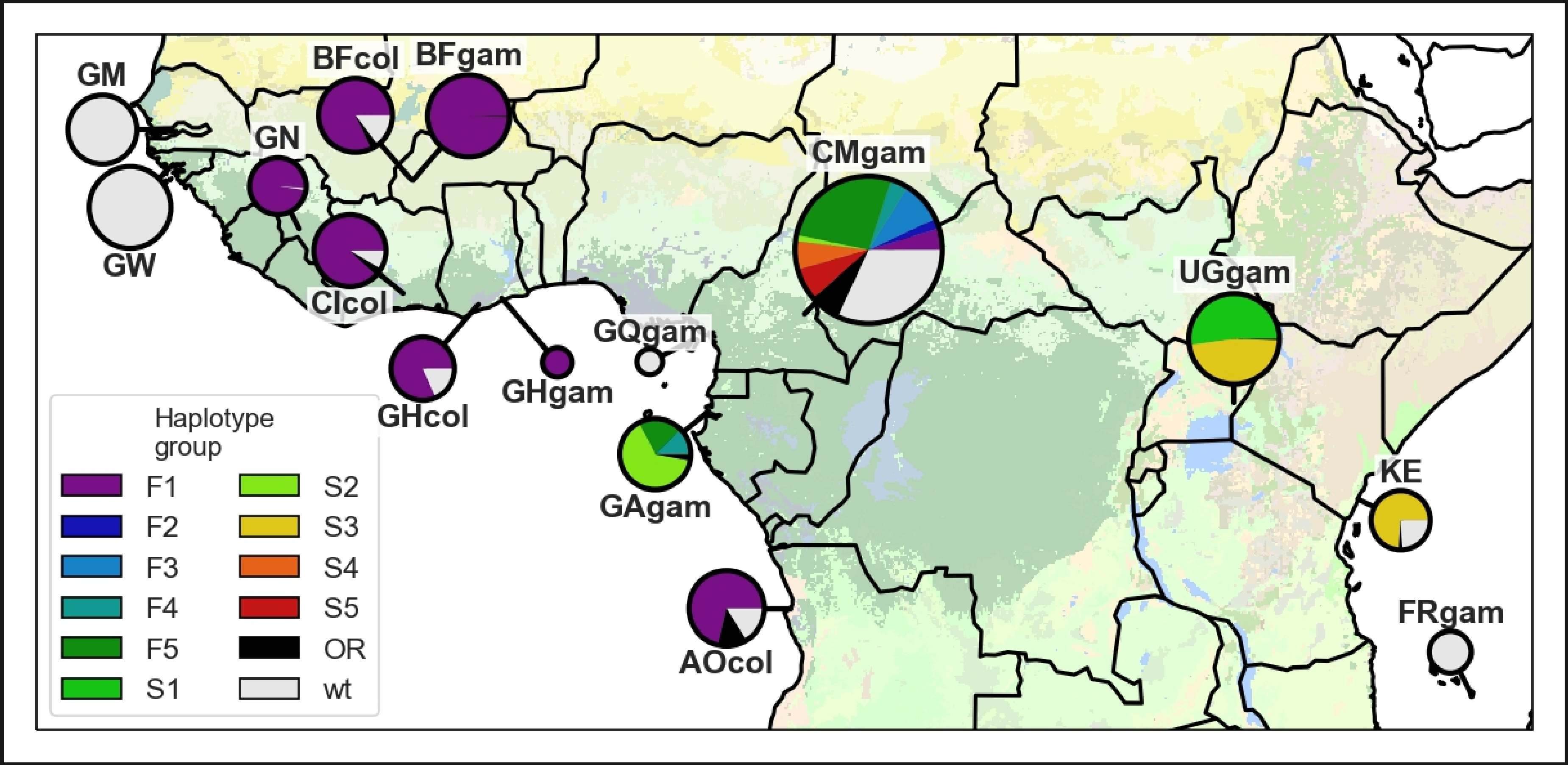
- *An. gambiae*: Cameroon (6%), Ghana (17%), Burkina Faso (21%), Guinea (9%)
- *An. coluzzii*: Burkina Faso (27%)

New resistance SNPs?

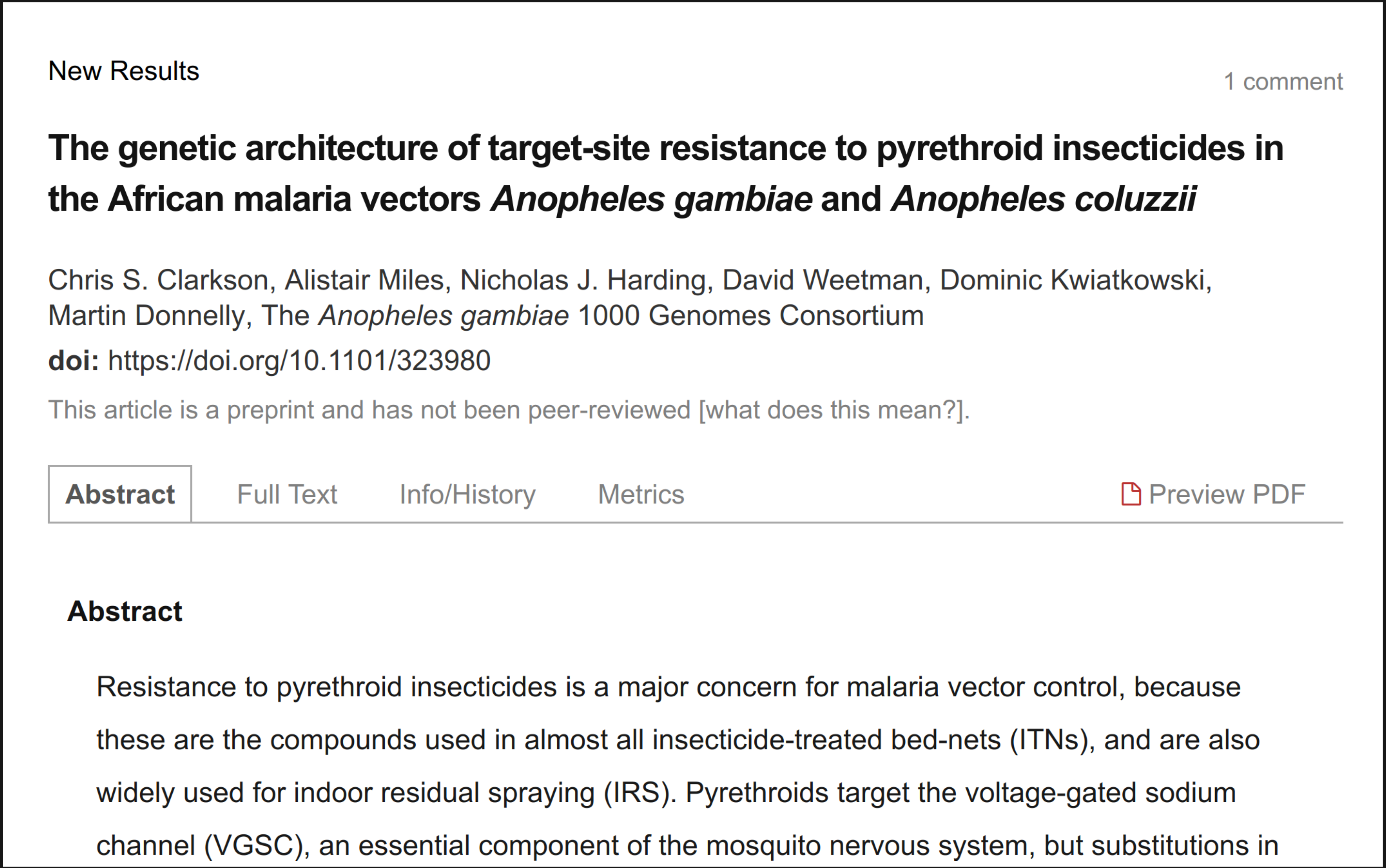
- L995F + R254K
- L995F + D466H + I1940T
- L995F + T791M + A1746S
- L995F + E1597G
- L995F + K1603T
- L995F + V1853I
- L995F + I1868T
- L995F + P1874S
- L995F + P1874L
- L995F + F1920S
- L995F + A1934V
- V402L + I1527T

There is **much** more to pyrethroid target-site resistance than just *kdr*!

Spread of target-site resistance



Use genetic backgrounds (haplotypes) to infer outbreaks of resistance. E.g., "F1" is a major outbreak driven by *kdr* L995F, has spread throughout West and Central Africa, and spread between *An. gambiae* and *An. coluzzii*.



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

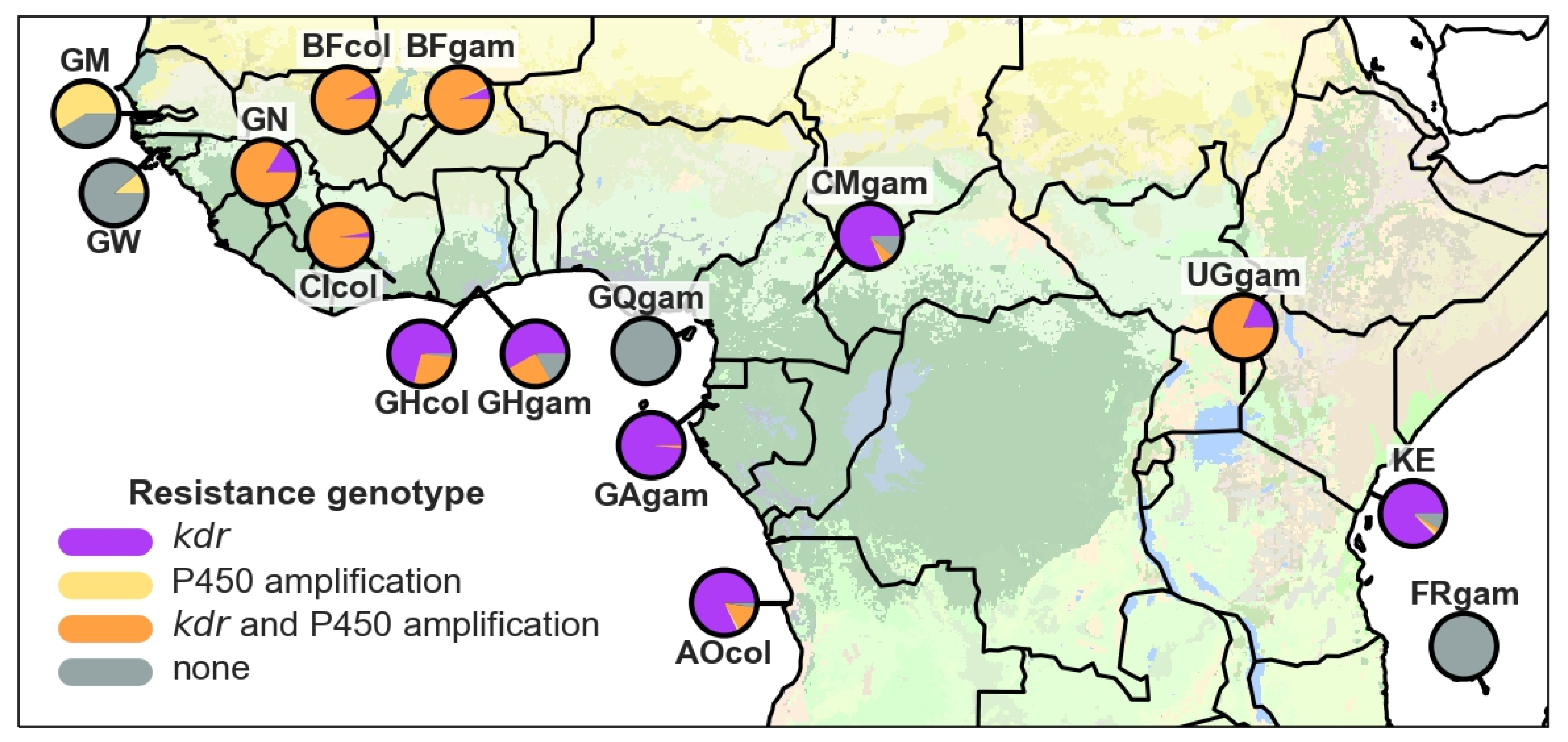
This article is a preprint and has not been peer-reviewed [what does this mean?].

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

(Will be updated with new data from Ag1000G phase 2 soon.)

Pyrethroid resistance mechanisms

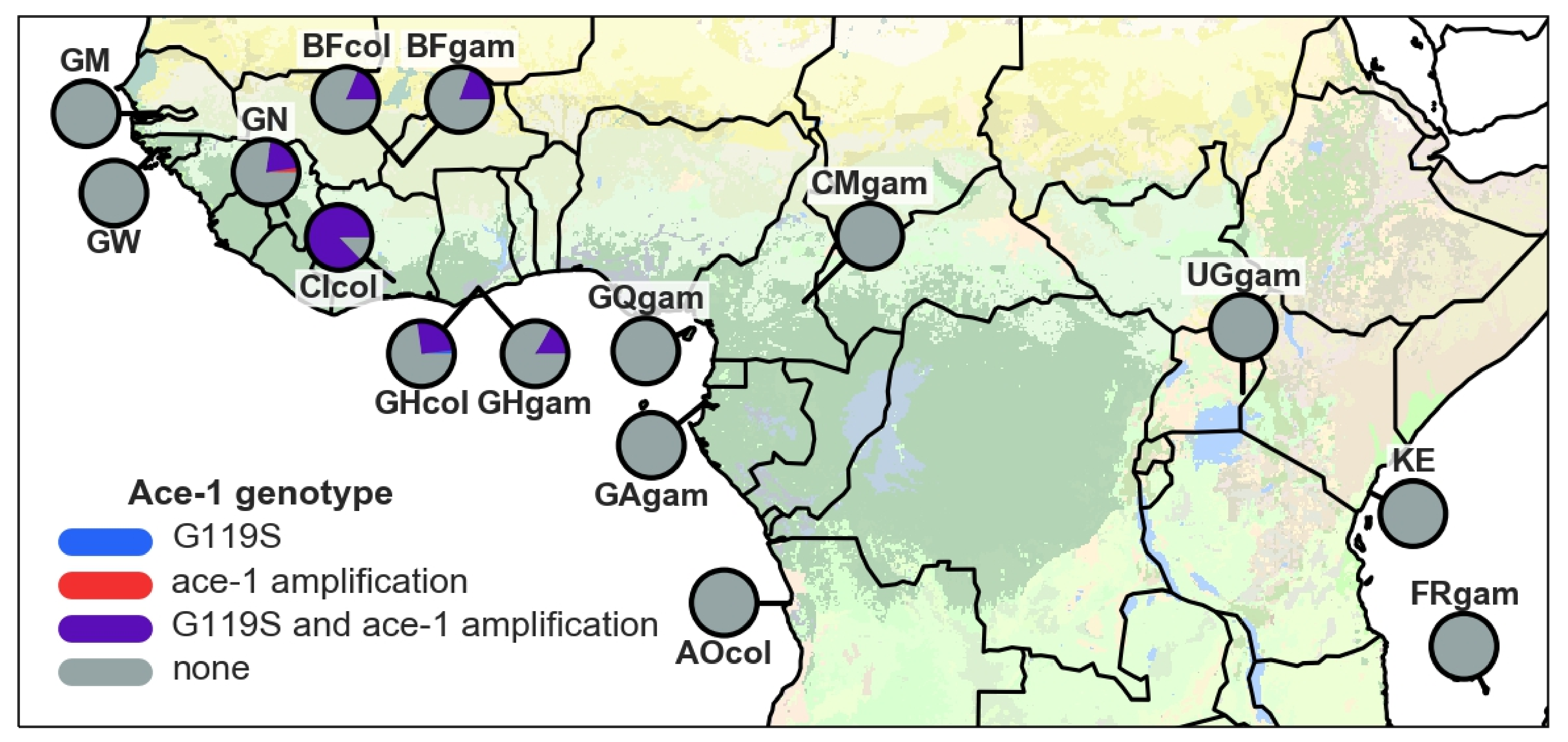


Possible to combine data on target-site and metabolic resistance, to show which molecular mechanisms are present in which populations.

Organophosphate resistance

- Organophosphates bind to acetylcholinesterase (ACE1) enzyme
- Genetic changes in ACE1 gene associated with resistance
    - SNPs (G119S) and CNVs

Prevalence of ACE1 SNPs and CNVs



G119S and CNV always found together. Spreading in West Africa, both *An. gambiae* and *An. coluzzii*.

Unknown/emerging resistance mechanisms

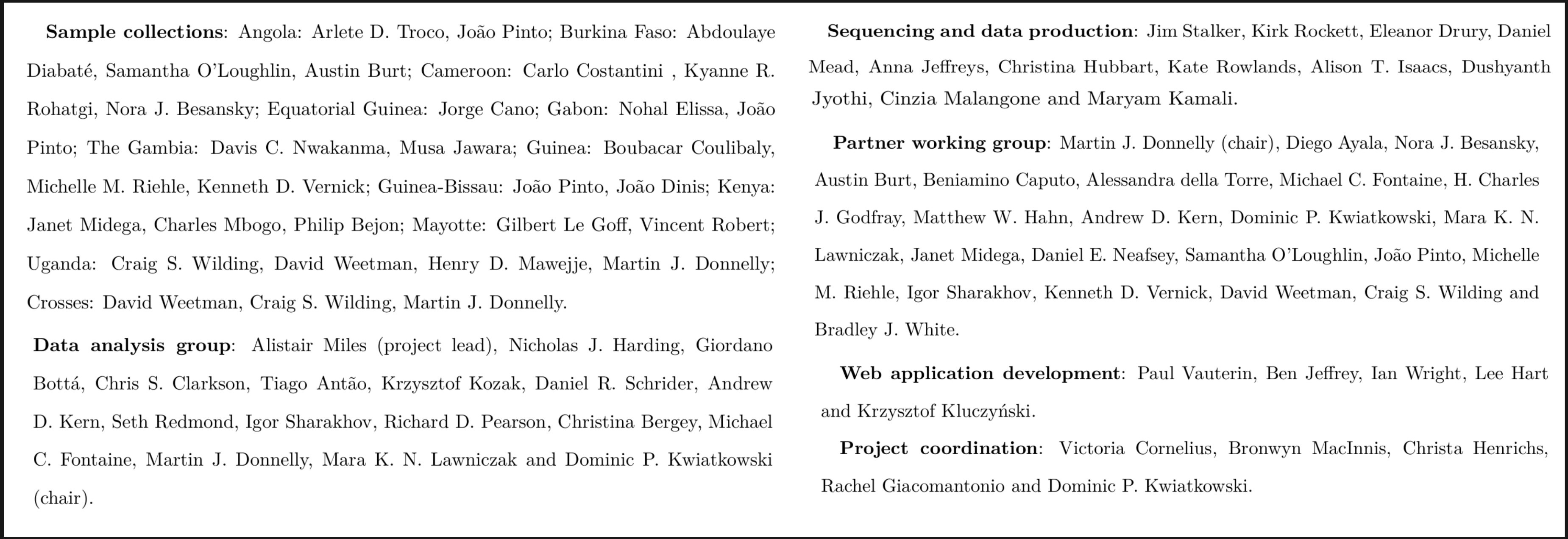
- GSTE CNVs - major role in metabolic resistance of pyrethroids?
- Carboxylesterases - role in organophosphate resistance?
- Emerging signals at other genes in acetylcholine regulation pathway?

Translation to surveillance

- Use cases in insecticide resistance management
    - Deployment of next-generation LLINs
    - Rotation of next-generation IRS
- Need recent samples
- Need better geographical coverage
- Need more species (e.g., *An. arabiensis*, *An. funestus*)
- Scale up whole-genome sequencing at genome centres
    - N.B., Wellcome Sanger Institute can now sequence 10,000 mosquitoes/year
- Scale up targeted (amplicon) sequencing at local/region labs

Acknowledgments

This work was carried out by the *Anopheles gambiae* 1000 Genomes Consortium:



Sample collections: Angola: Arlete D. Troco, João Pinto; Burkina Faso: Abdoulaye Diabaté, Samantha O'Loughlin, Austin Burt; Cameroon: Carlo Costantini , Kyanne R. Rohatgi, Nora J. Besansky; Equatorial Guinea: Jorge Cano; Gabon: Nohal Elissa, João Pinto; The Gambia: Davis C. Nwakanma, Musa Jawara; Guinea: Boubacar Coulibaly, Michelle M. Riehle, Kenneth D. Vernick; Guinea-Bissau: João Pinto, João Dinis; Kenya: Janet Midega, Charles Mbogo, Philip Bejon; Mayotte: Gilbert Le Goff, Vincent Robert; Uganda: Craig S. Wilding, David Weetman, Henry D. Mawejje, Martin J. Donnelly; Crosses: David Weetman, Craig S. Wilding, Martin J. Donnelly.

Data analysis group: Alistair Miles (project lead), Nicholas J. Harding, Giordano Bottá, Chris S. Clarkson, Tiago Antão, Krzysztof Kozak, Daniel R. Schrider, Andrew D. Kern, Seth Redmond, Igor Sharakhov, Richard D. Pearson, Christina Bergey, Michael C. Fontaine, Martin J. Donnelly, Mara K. N. Lawniczak and Dominic P. Kwiatkowski (chair).

Sequencing and data production: Jim Stalker, Kirk Rockett, Eleanor Drury, Daniel Mead, Anna Jeffreys, Christina Hubbart, Kate Rowlands, Alison T. Isaacs, Dushyanth Jyothi, Cinzia Malangone and Maryam Kamali.

Partner working group: Martin J. Donnelly (chair), Diego Ayala, Nora J. Besansky, Austin Burt, Beniamino Caputo, Alessandra della Torre, Michael C. Fontaine, H. Charles J. Godfray, Matthew W. Hahn, Andrew D. Kern, Dominic P. Kwiatkowski, Mara K. N. Lawniczak, Janet Midega, Daniel E. Neafsey, Samantha O'Loughlin, João Pinto, Michelle M. Riehle, Igor Sharakhov, Kenneth D. Vernick, David Weetman, Craig S. Wilding and Bradley J. White.

Web application development: Paul Vauterin, Ben Jeffrey, Ian Wright, Lee Hart and Krzysztof Kluczyński.

Project coordination: Victoria Cornelius, Bronwyn MacInnis, Christa Henrichs, Rachel Giacomantonio and Dominic P. Kwiatkowski.

Thank you to the staff of the Wellcome Sanger Institute Sample Logistics, Sequencing and Informatics facilities.

Thank you to Wellcome Trust, MRC UK, DFID UK, NIH, NIAID, Bill and Melinda Gates Foundation.

Extra slides

Collection						***Anopheles* species counts**		
Country	Location	Site	Year	Latitude	Longitude	*gambiae*	*coluzzii*	Unknown
Angola	Luanda		2009	-8.8210	13.2910	0	78	0
Burkina Faso	Bana		2012	11.2330	-4.4720	20	40	0
	Pala		2012	11.1500	-4.2350	46	10	0
	Souroukoudinga		2012	11.2350	-4.5350	26	25	0
Cameroon	Daiguene		2009	4.7770	13.8440	96	0	0
	Gado Badzere		2009	5.7470	14.4420	73	0	0
	Mayos		2009	4.3410	13.5580	105	0	0
	Zembe Borongo		2009	5.7470	14.4420	23	0	0
Cote d'Ivoire	Tiassale		2012	5.8984	-4.8229	0	71	0
Equatorial Guinea	Bioko		2002	3.7000	8.7000	9	0	0
France	Mayotte	Bouyouni	2011	-12.7378	45.1417	1	0	0
		Combani	2011	-12.7787	45.1429	5	0	0
		Karihani Lake	2011	-12.7965	45.1217	3	0	0
		Mont Benara	2011	-12.8570	45.1552	2	0	0
		Mtsamboro Forest Reserve	2011	-12.7027	45.0811	1	0	0
		Mtsanga Charifou	2011	-12.9907	45.1557	8	0	0
		Sada	2011	-12.8521	45.1039	4	0	0
Gabon	Libreville		2000	0.3840	9.4550	69	0	0
Gambia, The	Njabakunda	Kerr Birom Kardo	2011	13.5500	-15.9000	0	0	19
		Kerr Sama Kuma	2011	13.5500	-15.9000	0	0	8
		Maria Samba Nyado	2011	13.5500	-15.9000	0	0	18
		Sare Illo Buya	2011	13.5500	-15.9000	0	0	20
Ghana	Koforidua		2012	6.0945	-0.2609	0	1	0
	Madina		2012	5.6685	-0.2193	12	12	0
	Takoradi		2012	4.9122	-1.7740	0	20	0
	Twifo Praso		2012	5.6086	-1.5493	0	22	0
Guinea	Koraboh		2012	9.2500	-9.9170	22	0	0
	Koundara		2012	8.5000	-9.4170	18	4	0
Guinea-Bissau	Antula		2010	11.8910	-15.5820	0	0	58
	Safim		2010	11.9569	-15.6492	0	0	33
Kenya	Kilifi	Junju	2012	-3.8620	39.7450	0	0	16
		Mbogolo	2012	-3.6350	39.8580	0	0	32
Uganda	Tororo	Nagongera	2012	0.7700	34.0260	112	0	0