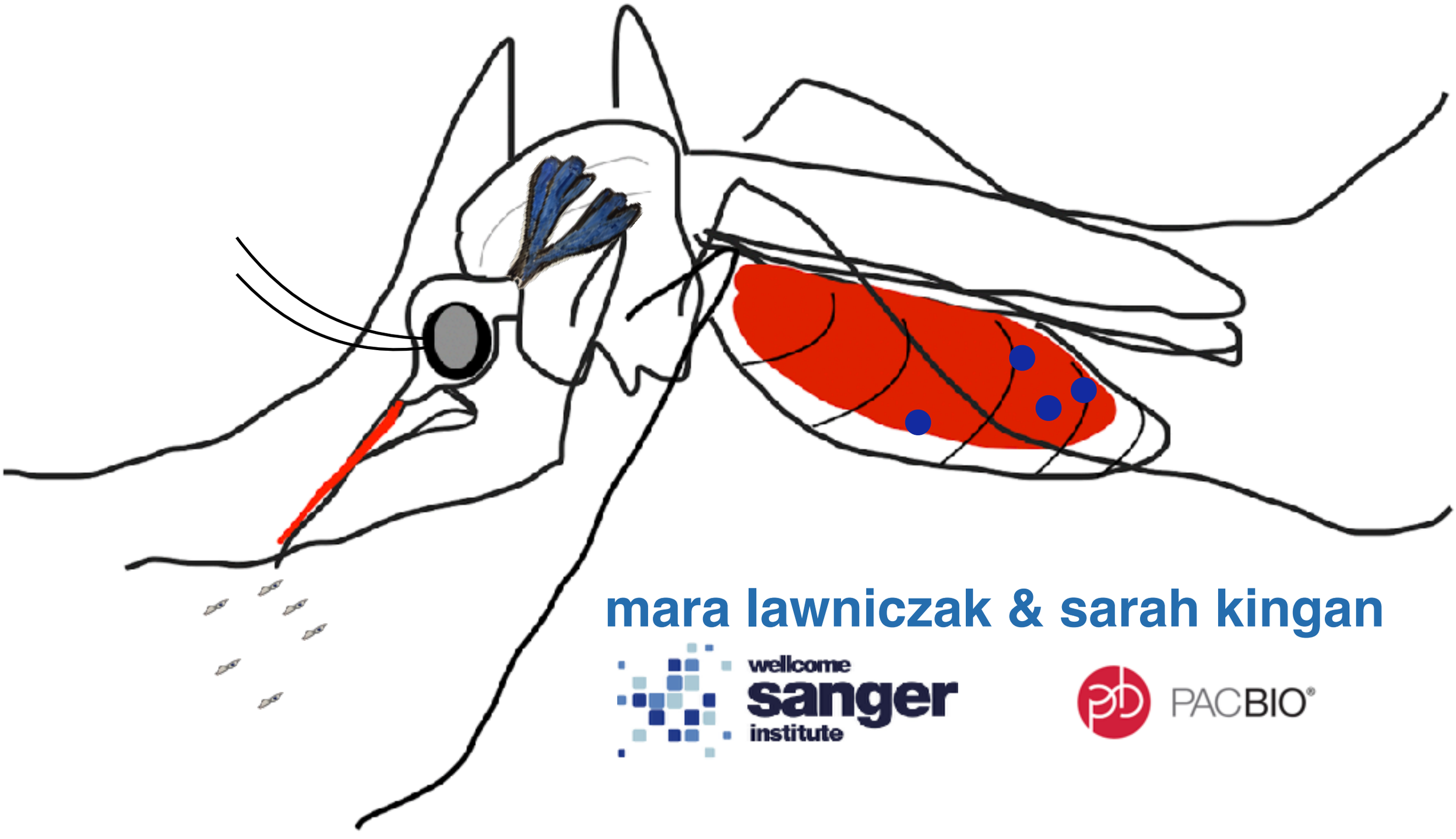
High quality PacBio genomes from single insects: implications for vector research.



mara lawniczak & sarah kingan

wellcome sanger institute

PACBIO®

Sarah Kingan
@drsarahdoom
Staff Scientist
PacBio

Mara Lawniczak
@marakat
Group Leader
Wellcome Sanger Institute

Outline

mara

- Overview of Anopheles species diversity, vector control, population variation, and looking forward to next projects.
- Summary of experiences with HMW DNA extractions on single insects.

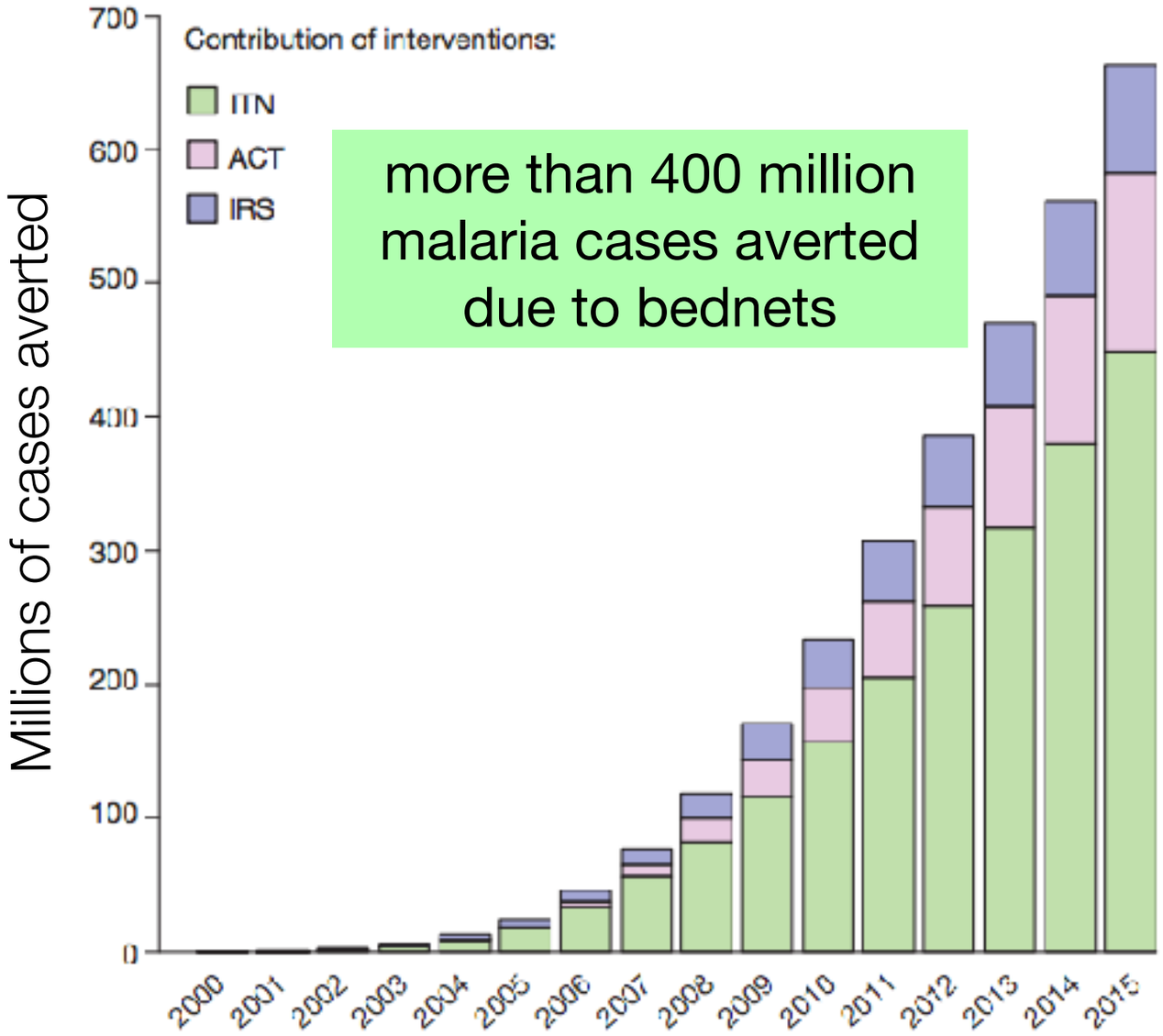
sarah

- Haplotype assembly approaches.
- Low-input protocol: single mosquito assembly, from library prep through assembly, importance of the DNA profile, recommendations for coverage and genome size, application to small organisms.
- What’s new with PacBio, including Sequel II 8M data, HiFi CCS reads, and Iso-Seq for genome annotation.

Vector control is the most effective malaria control strategy

ITN: insecticide treated net

IRS: indoor residual spray



more than 400 million malaria cases averted due to bednets

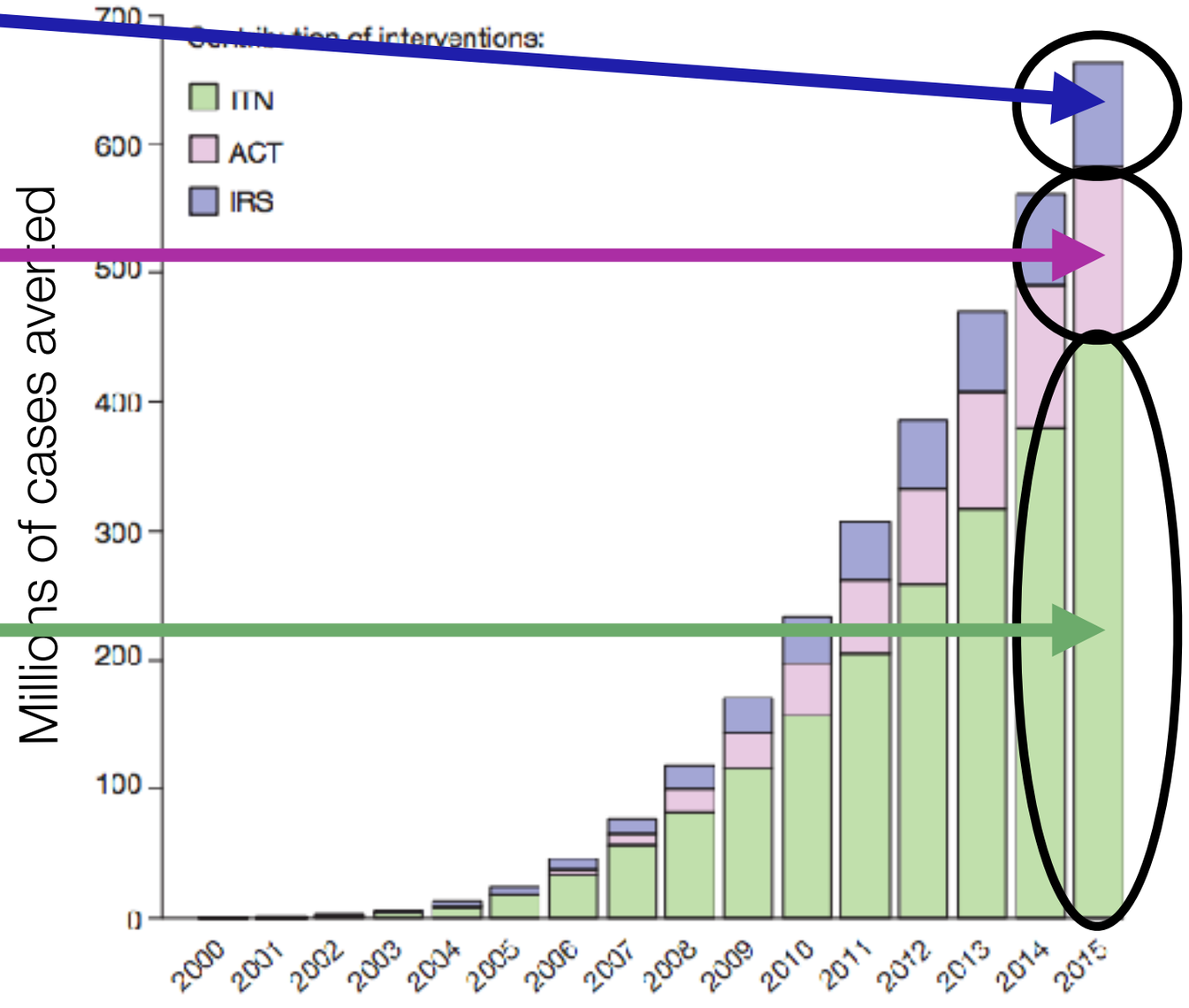
Bhatt *et al*. Nature. 2015

These gains are at risk

resistance to all 4 classes of insecticides on the rise

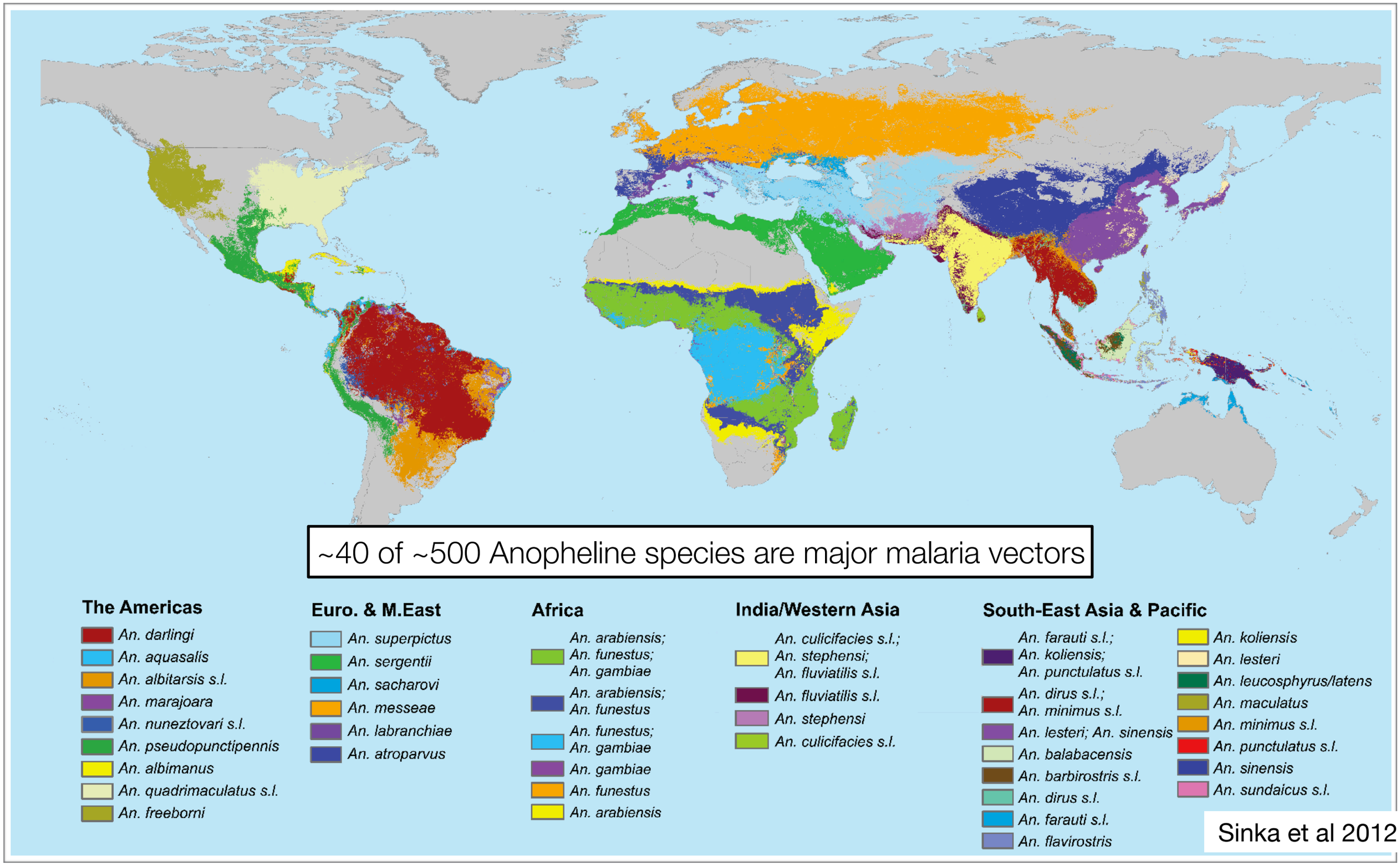
drug resistance on the rise in SEAsia parasites

resistance to pyrethroids is very prevalent

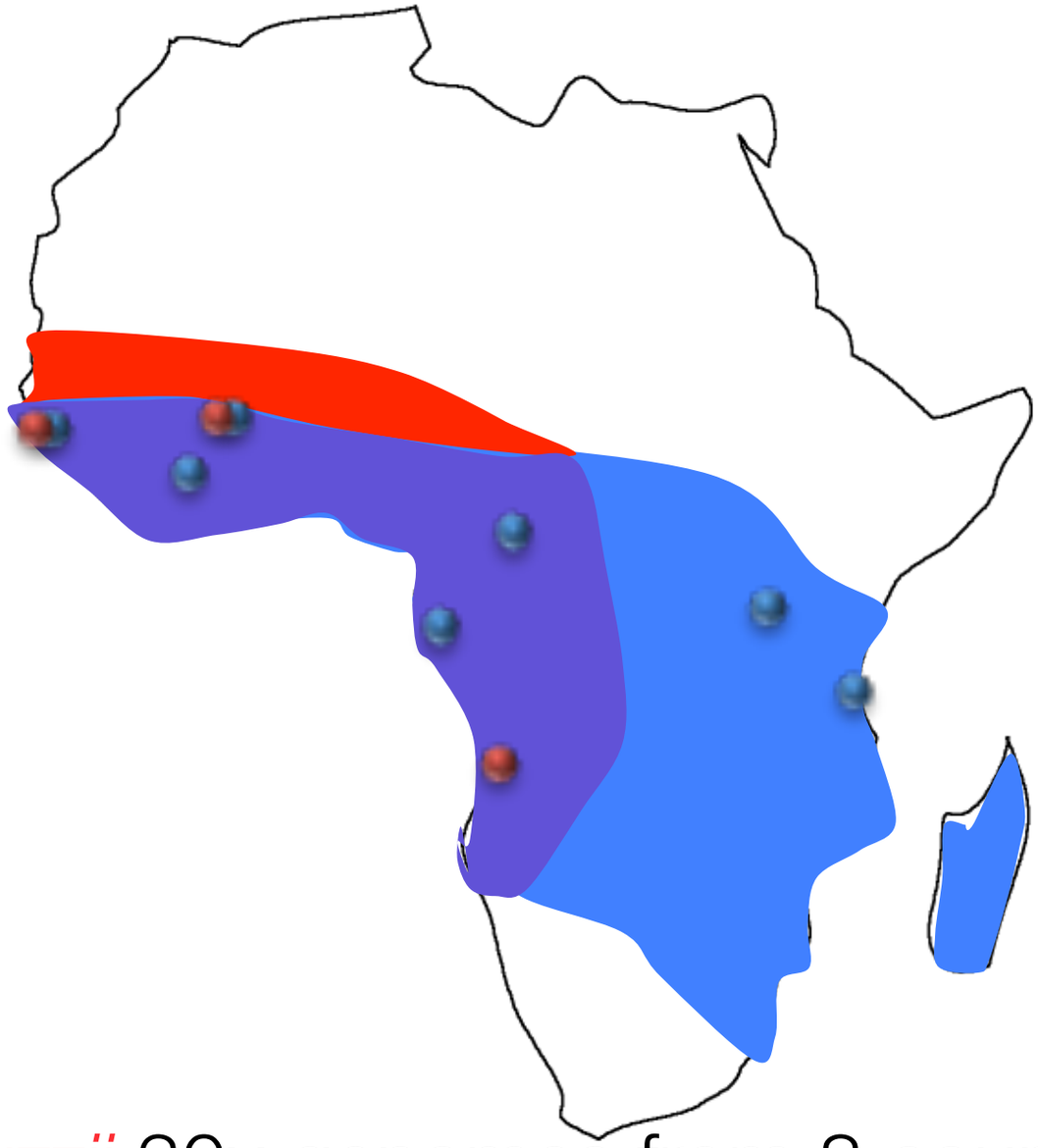


Bhatt *et al*. Nature. 2015

Current methods have bred resistance and also won't work on outdoor biting vector species



The *Anopheles gambiae* 1000 Genomes Project

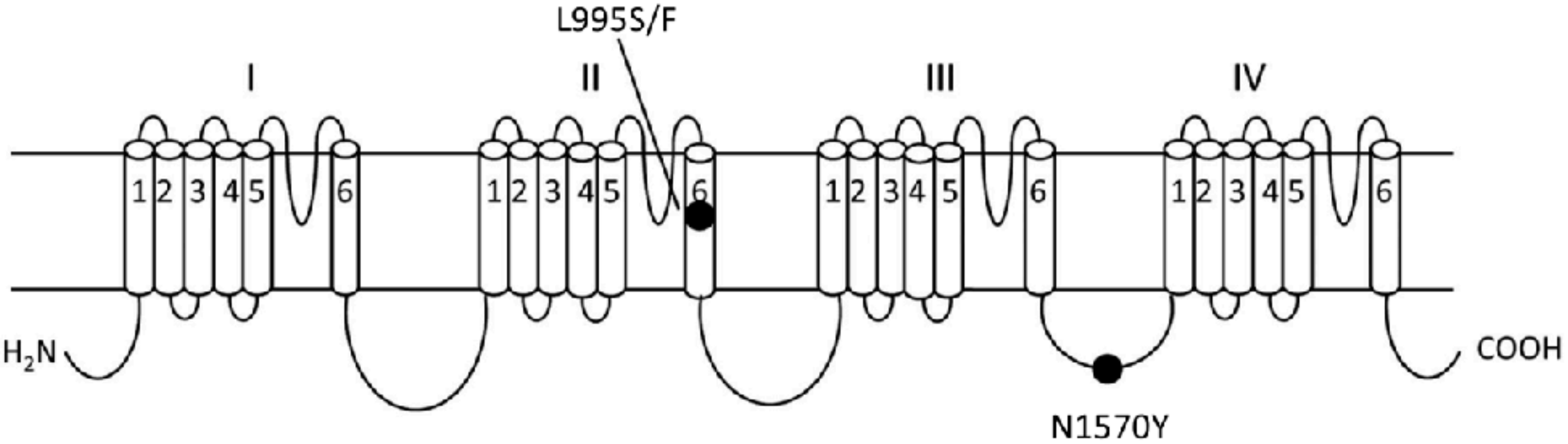


765 wild *An. gambiae* and *An. coluzzii* 30x genomes from 8 countries

Creating a public database of mosquito genomes to advance basic biology and malaria control

Ag1000g consortium, Nature 2017

Target site resistance: the voltage gated sodium channel

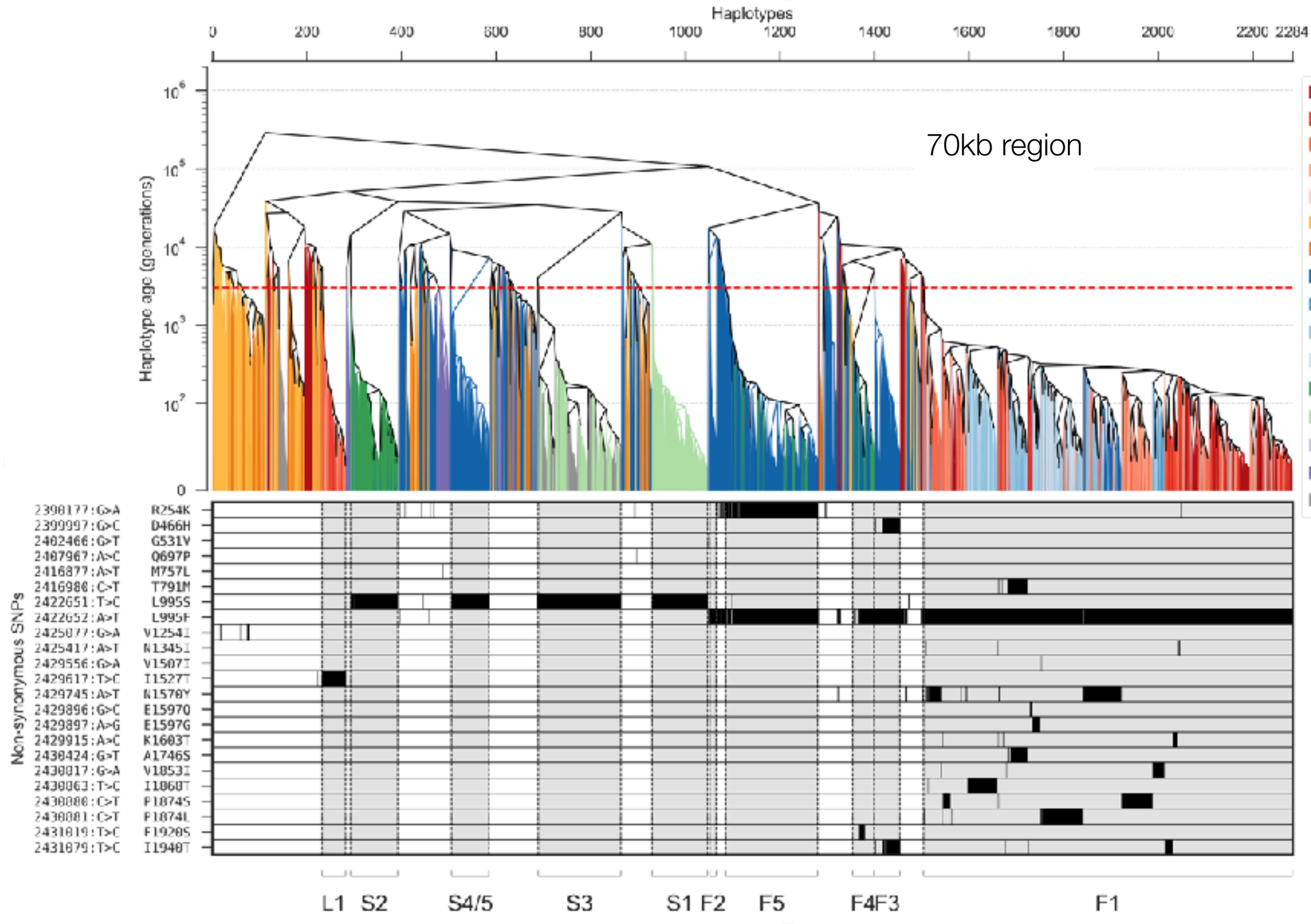


VGSC is an essential component of the insect nervous system.

L995S/F mutation reduces the efficiency of binding to both DDT and pyrethroids, causing knock down resistance (kdr mutation)

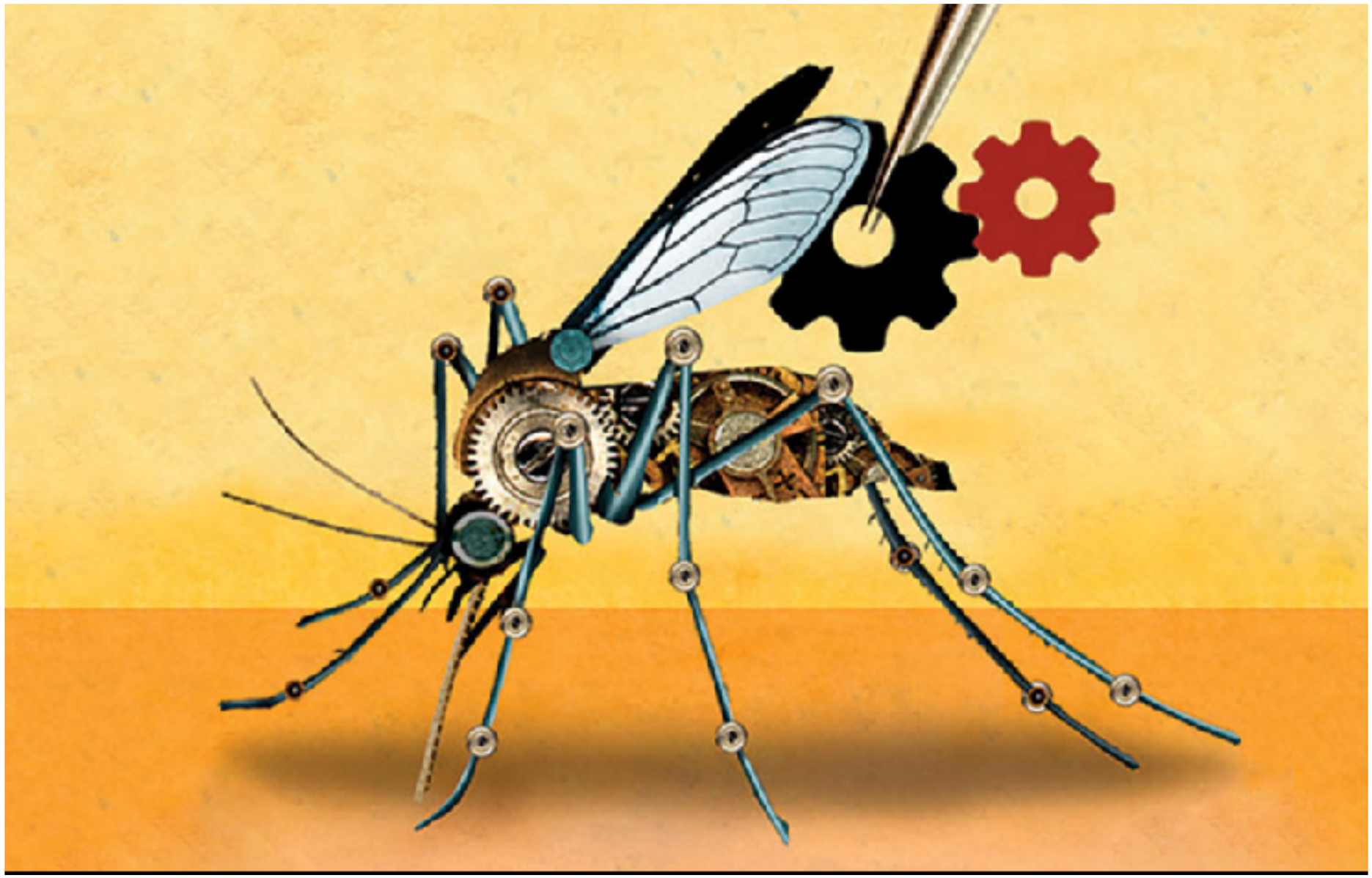
Pyrethroids are the only approved insecticide for bednets

At least ten separate “outbreaks” of resistance at the VGSC gene



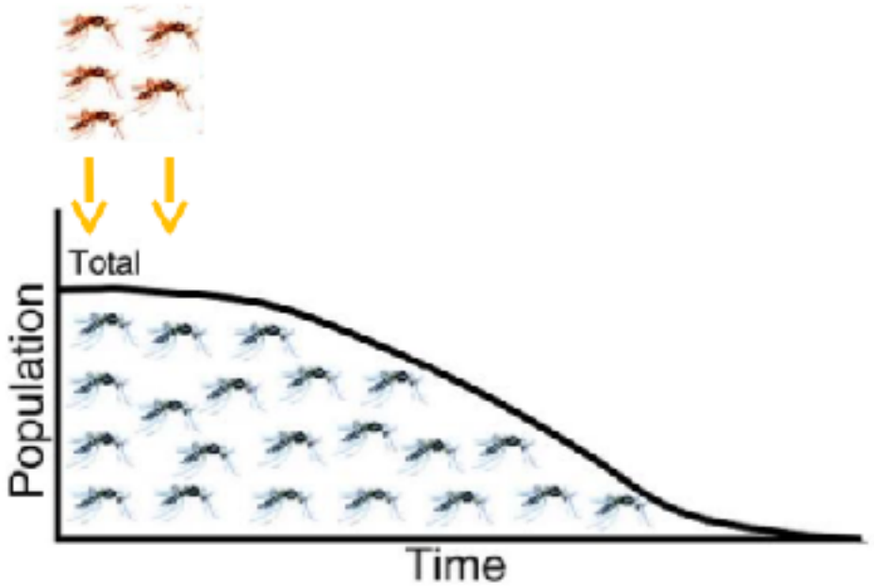
Chris Clarkson
Alistair Miles
Nick Harding

Ag1000g consortium, Nature 2017

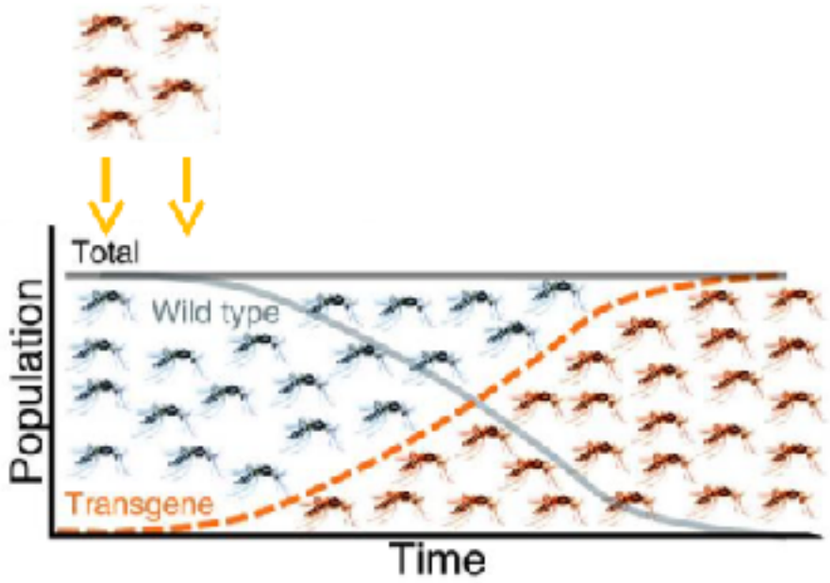


Synthetic gene drive systems for mosquito control

Gene drive: super-Mendelian inheritance

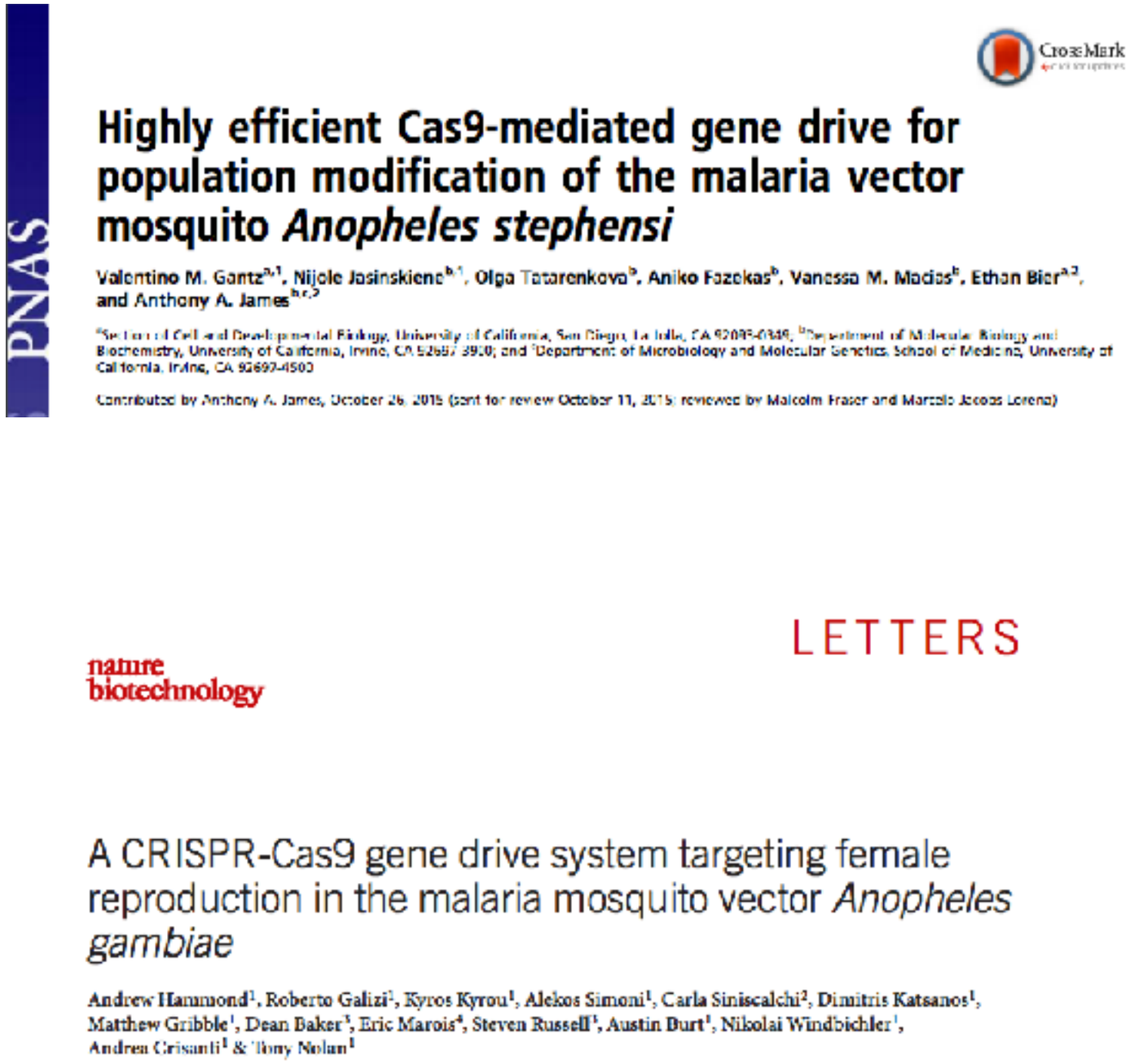


Population suppression
[elimination/sex ratio distortion]



Population replacement
[no transmission]

Both population suppression and replacement have now been engineered in the lab using Cas9



PNAS

Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*

Valentino M. Gantz[a,1], Nijole Jasinskiene[b,1], Olga Tatarenkova[b], Aniko Fazekas[b], Vanessa M. Macias[b], Ethan Bier[a,2], and Anthony A. James[b,c,2]

[a]Section of Cell and Developmental Biology, University of California, San Diego, La Jolla, CA 92093-0349; [b]Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697-3900; and [c]Department of Microbiology and Molecular Genetics, School of Medicine, University of California, Irvine, CA 92697-4500

Contributed by Anthony A. James, October 26, 2015 (sent for review October 11, 2015; reviewed by Malcolm Fraser and Marcelo Jacobs-Lorena)

population replacement with cassette expressing single chain antibodies that target transmission-critical parasite and mosquito proteins.

nature biotechnology

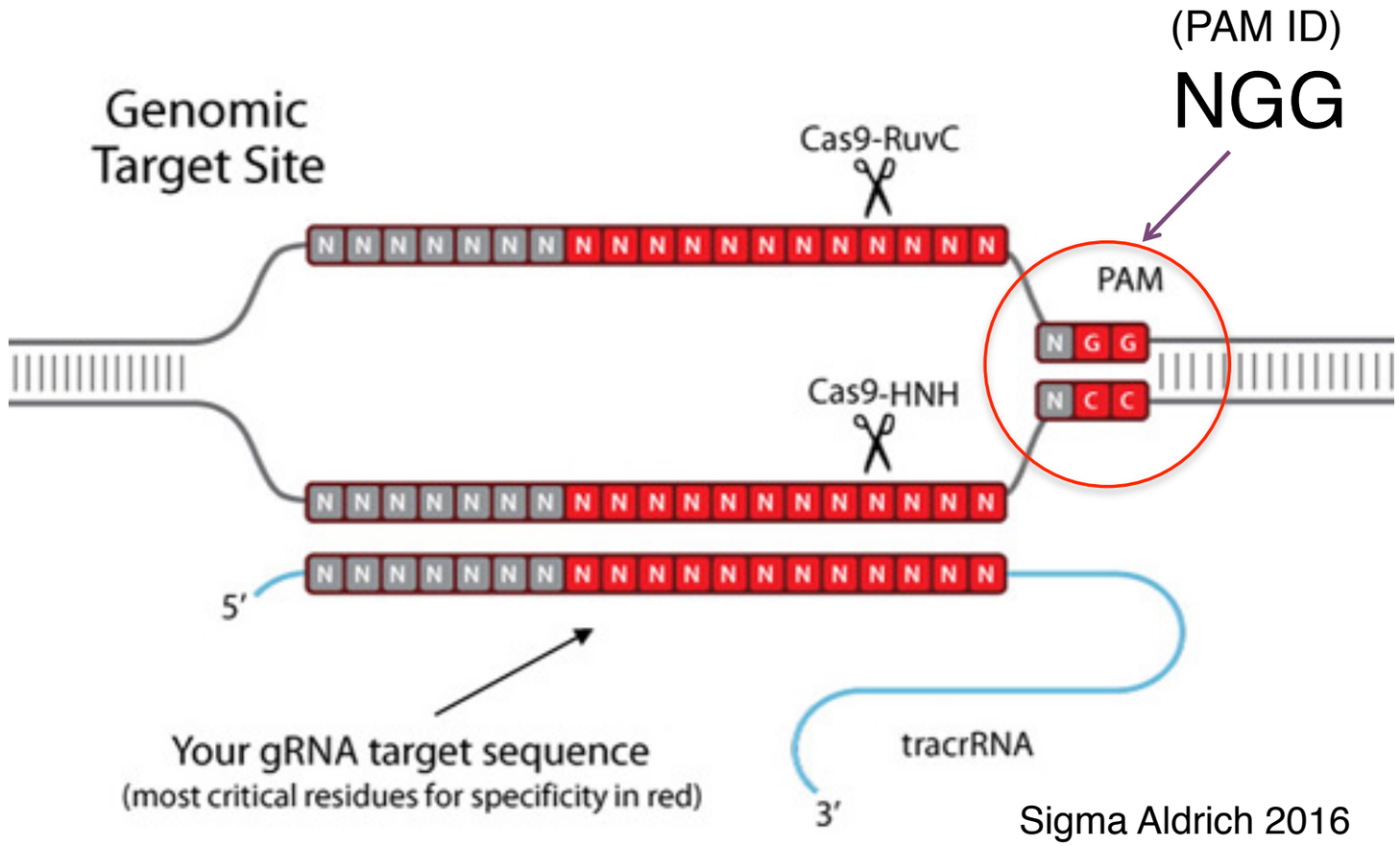
LETTERS

A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*

Andrew Hammond[1], Roberto Galizi[1], Kyros Kyrou[1], Alekos Simoni[1], Carla Siniscalchi[2], Dimitris Katsanos[1], Matthew Gribble[1], Dean Baker[3], Eric Marois[4], Steven Russell[1], Austin Burt[1], Nikolai Windbichler[1], Andrea Crisanti[1] & Tony Nolan[1]

population suppression using a gRNA that targets a fertility gene, whose interruption causes female sterility only when homozygous.

guide RNA cut sites must have sequence conservation



Sigma Aldrich 2016

AGAP007280 is a promising target

BBC Sign in News Sport Weather iPlayer T

NEWS

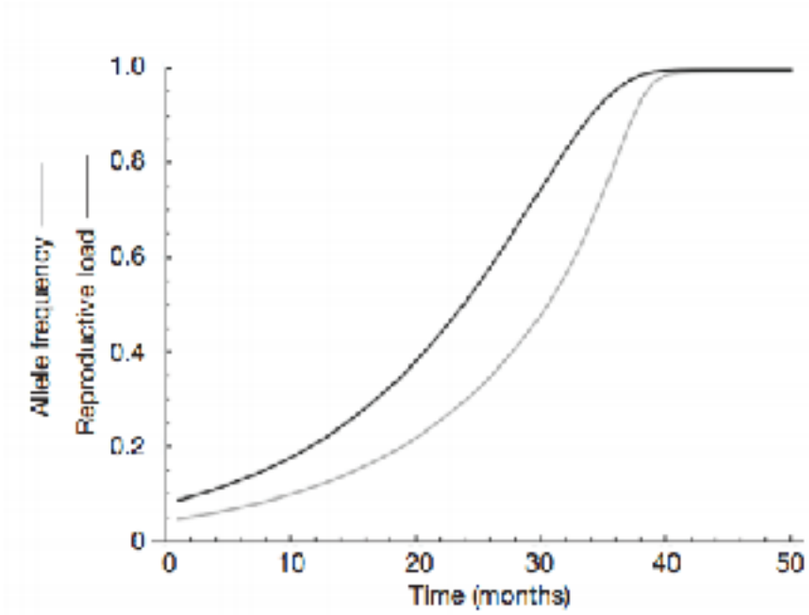
Home UK World Business Politics Tech Science Health Education

Health

Scientists create infertile mosquitoes

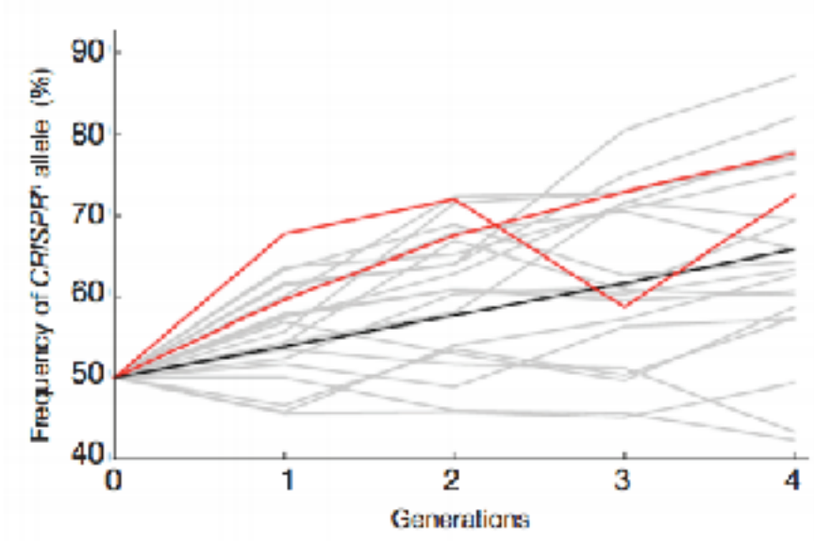
By Michelle Roberts
Health editor, BBC News online

7 December 2015 | Health



Modeling shows that the construct could spread through the population, imposing a reproductive load as it does so.

Allele increased in frequency over 4 generations from 50% to 75%



Homing depends on a 21bp guideRNA matching the sequence at the cut site.

Hammond et al 2015 Nature Biotech

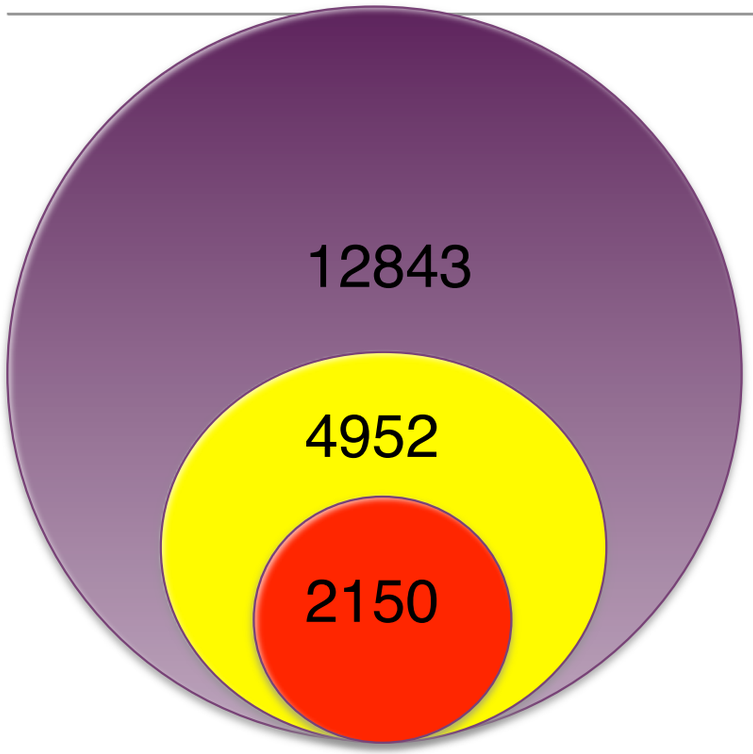
21bp gene drive target site has 8 SNPs segregating just among the Ag1000g samples

General problem as other targets investigated suffered same issue.

AGAP005958: **5 SNPs**

AGAP011377: **7 SNPs**

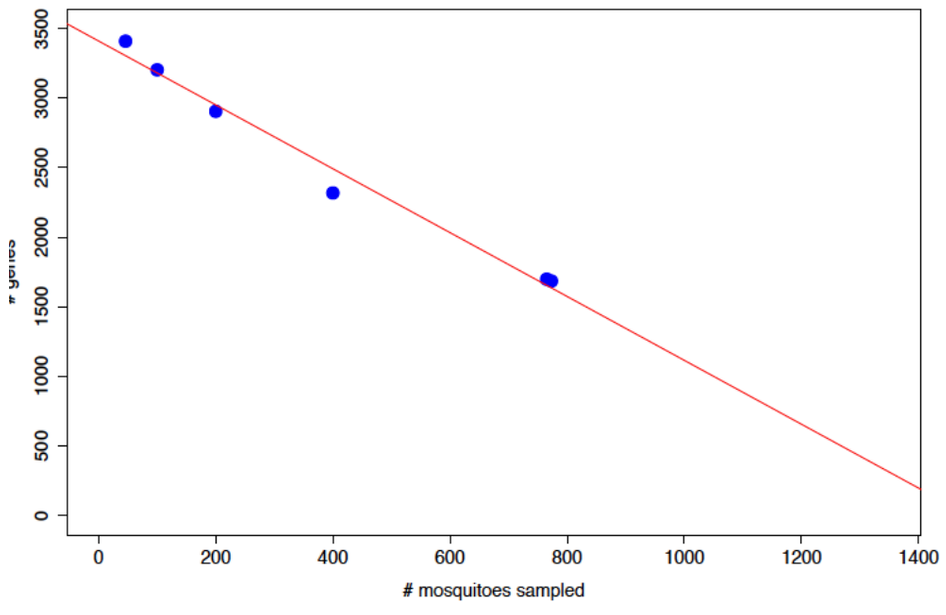
Extended this analysis to the whole genome: How many genes have invariant gRNA targets among these 765 wild sequenced individuals?



All protein coding genes

Genes with at least 1 invariant gRNA site

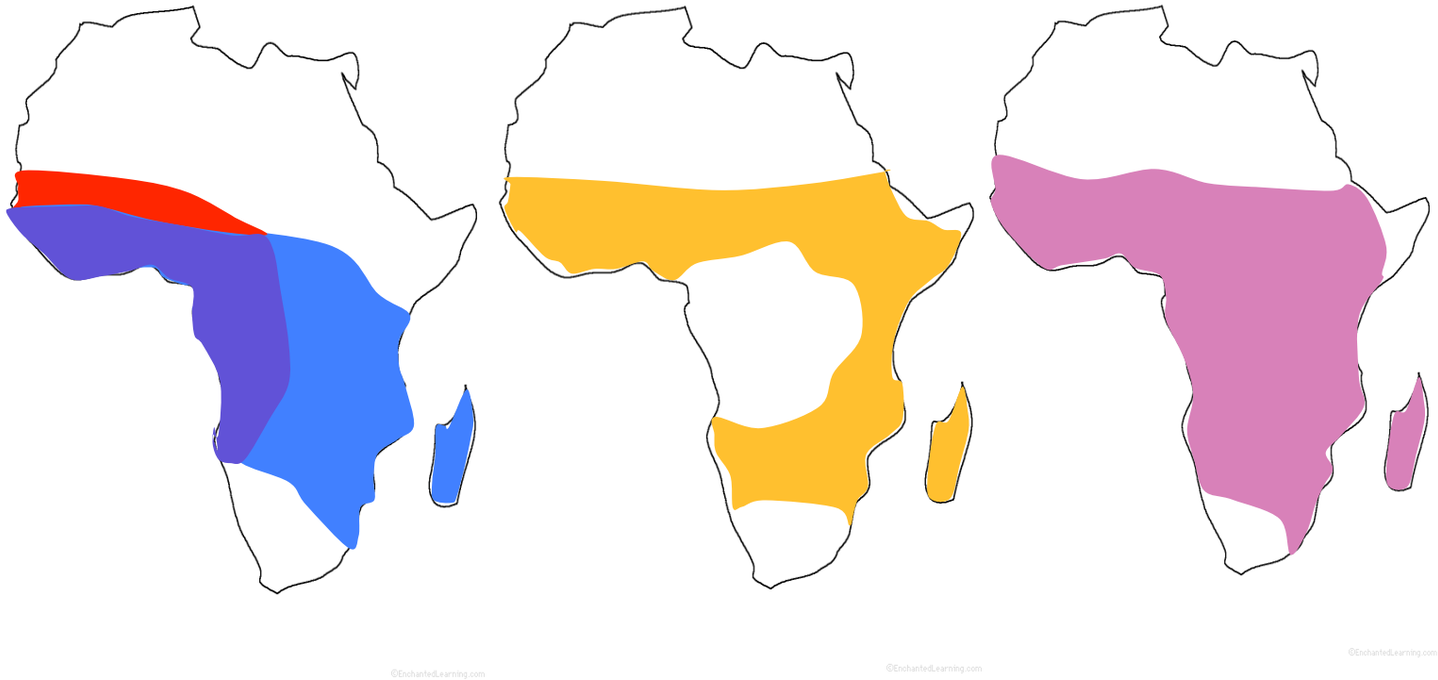
Genes with > 3 invariant gRNA sites



Sub-sampling the Ag1000g data shows that the number of targetable genes declines rapidly with additional data.

Ag1000g consortium, Nature 2017

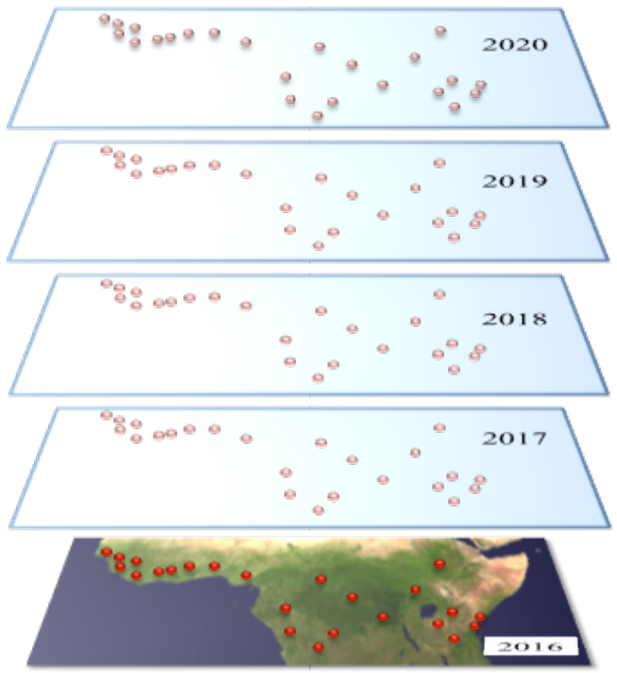
Vector and parasite observatory in Africa and Asia



An. gambiae
An. coluzzii

An. arabiensis

An. funestus



P. falciparum

We aim to understand vector and parasite evolution in order to inform and achieve sustainable disease control.

Sequence 30,000 vectors and 30,000 parasites in Asia and Africa 2016-2021

All analyses just presented depended on a reference genome… but very few species have high quality reference genomes.

An. gambiae PEST reference genome was the second insect ever assembled (2002).

Many years of effort, millions of dollars.

All mosquito genomes until now have required pooling mosquitoes to overcome DNA input requirements.

Heterozygosity headache….can be overcome by inbreeding, but this is never complete and is also laborious.

Better to reduce the number of individuals required for the assembly.

Collaboration between Sanger and PacBio groups to drop the input requirements…

HMW Extractions

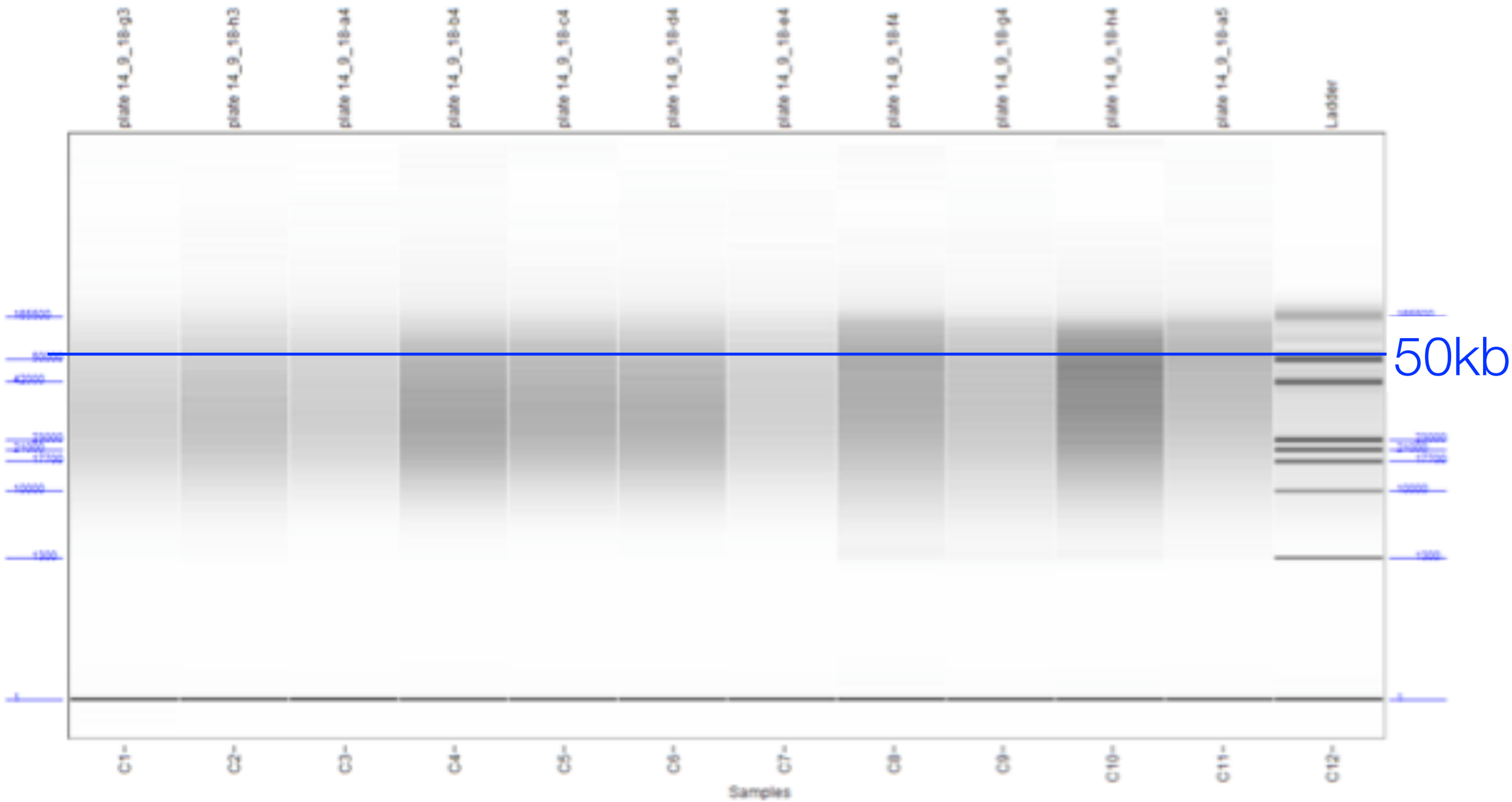
Qiagen MagAttract HMW DNA Kit: requires a magnetic rack ($500-1000); $333/£196 for 48 rxns.

- "**standard**" (as per instructions)
- "**10X modified**" (pg 6-8 of "Chromium™ Genome Reagent Kits v2 User Guide")

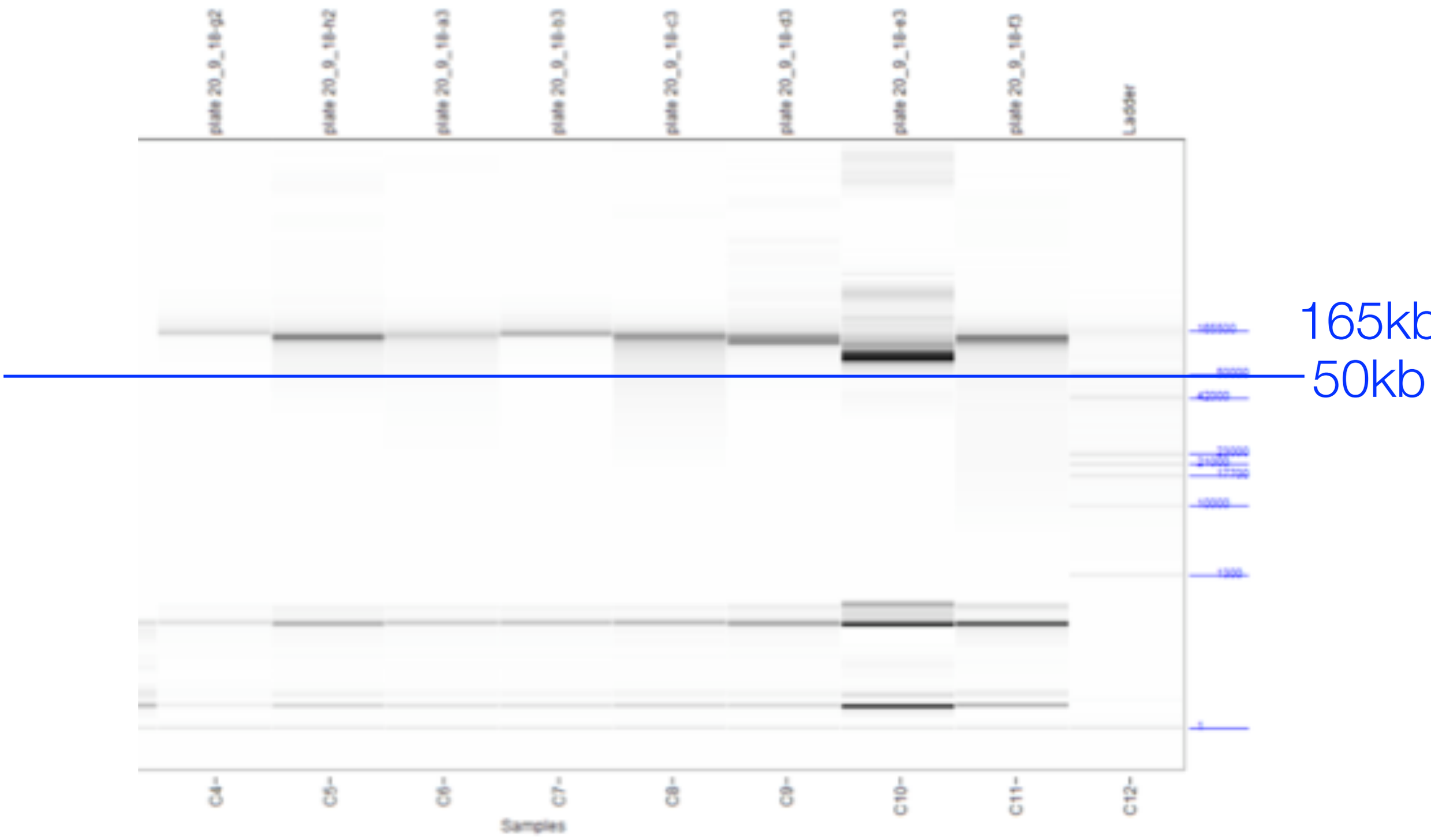
Note: kit designed for much higher quantity of DNA extracted than from single small insect, so we reduce the reagents we use for each extraction by 50%, apart from the beads which are reduced even further to 10ul per extraction.

How much tissue do you need? Single fresh *Anopheles gambiae* female weighs about 2 milligrams and gives about 200 nanograms of HMW DNA.
We guesstimate that about 0.01% of the weight of a fly is DNA.

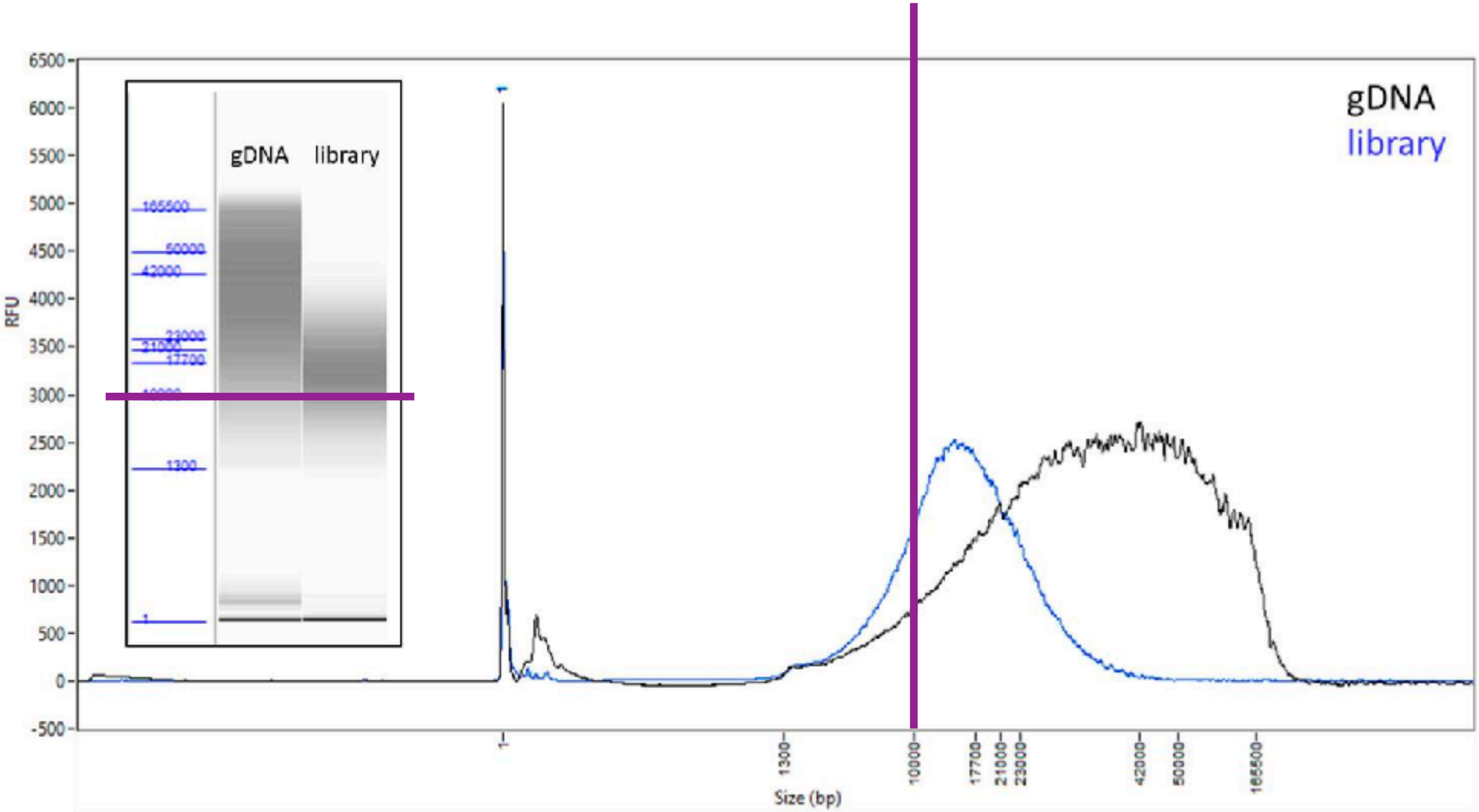
Typical profiles for “standard” MagAttract extraction on a fresh or a freshly killed and stored at -80C single mosquito. Typically get 200ng of HMW DNA.



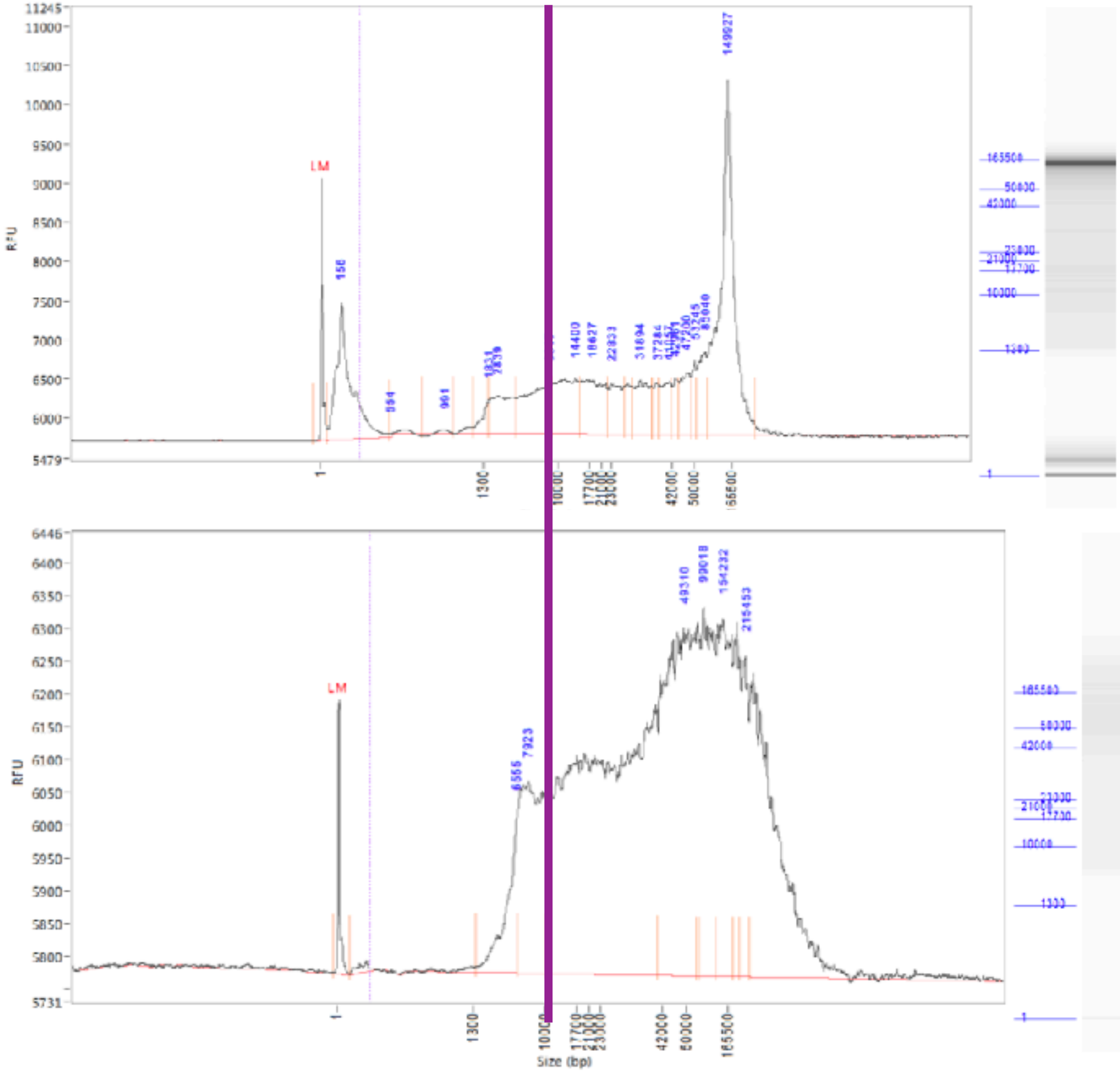
Typical profiles for “10X modified” MagAttract
(note, this is also typical of Phenol/Chloroform extractions). Typically get 200ng of HMW DNA.



Mosquito number 7—*An. coluzzii* Ngousso colony
(actual extraction before shipment looked like single band, but shattering happened during shipment, changing the profile quite dramatically)

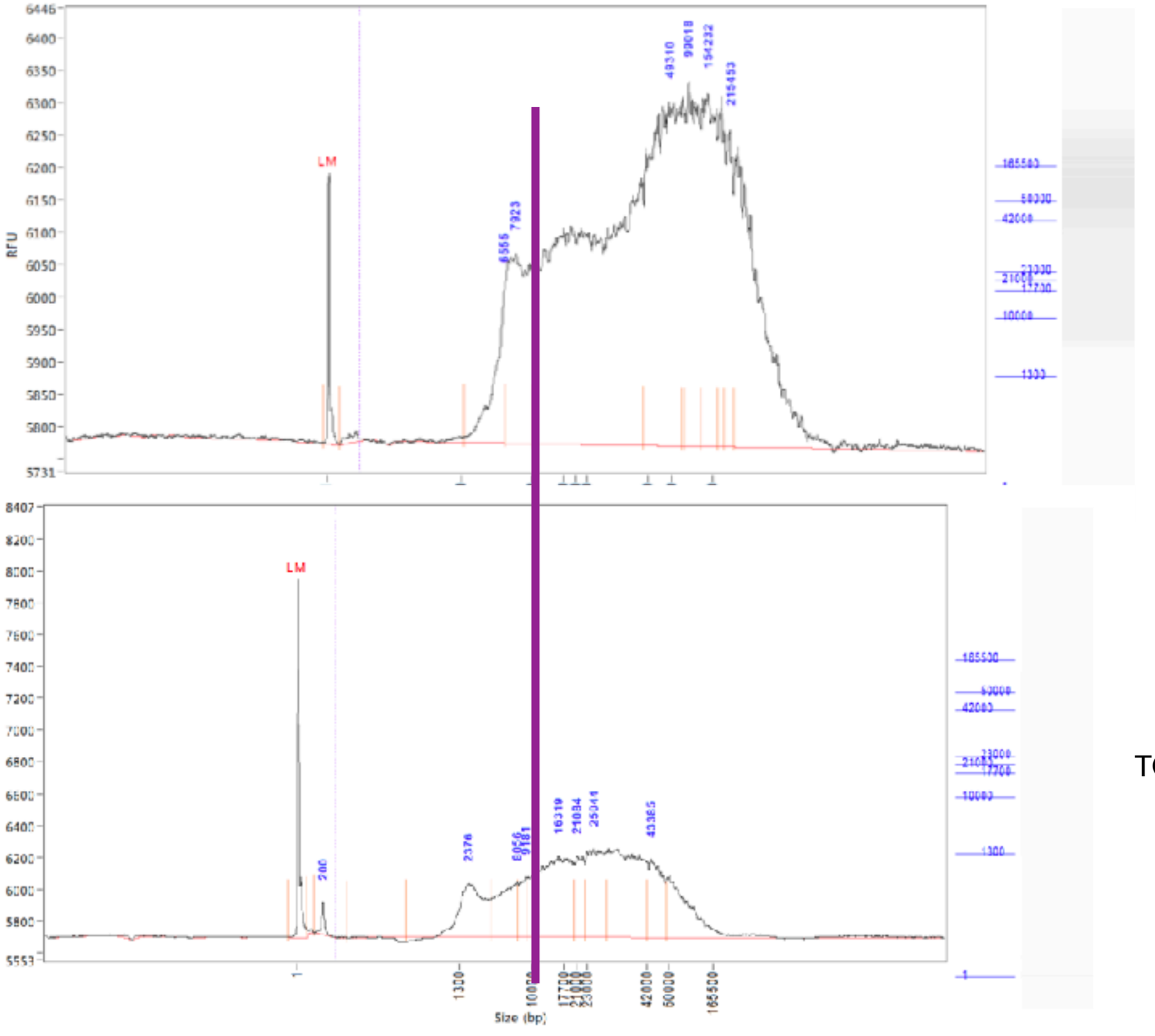


Test two DNA extractions (5.1 and 8.10) at Sanger to replicate mosquito number 7 success. “10X Modified magattract” (8.10) shown here.



8.10 pre-cleanup (0.45XAMPure)
Estimated to have 380ng by Femto
(560ng by Picogreen)

8.10 post-cleanup

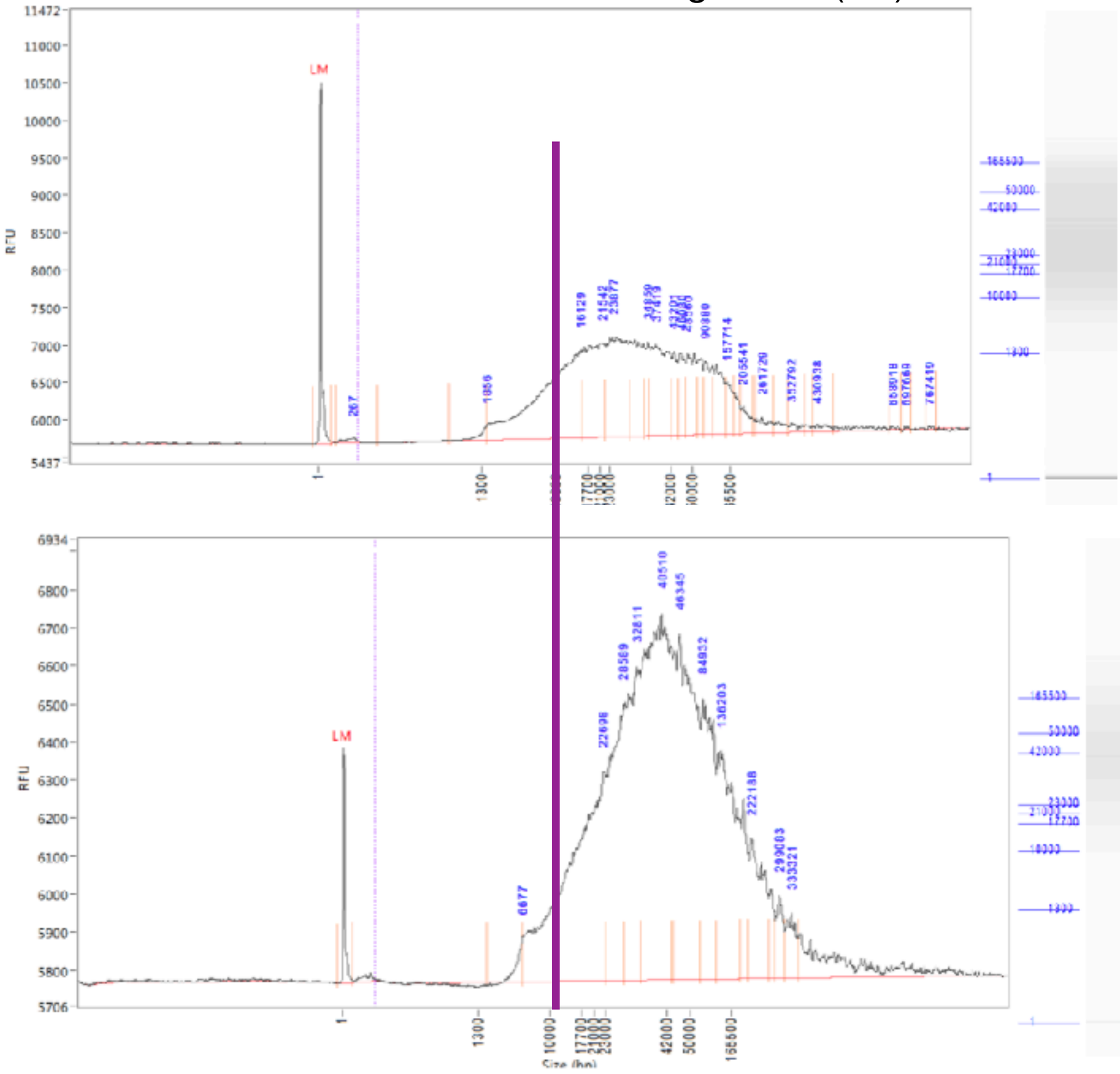


8.10 post-cleanup

8.10 library
TOO many molecules < 10kb.
Did not sequence this.

* new approaches to remove <3kb fragments are under development at PacBio, which might rescue these smaller DNAs

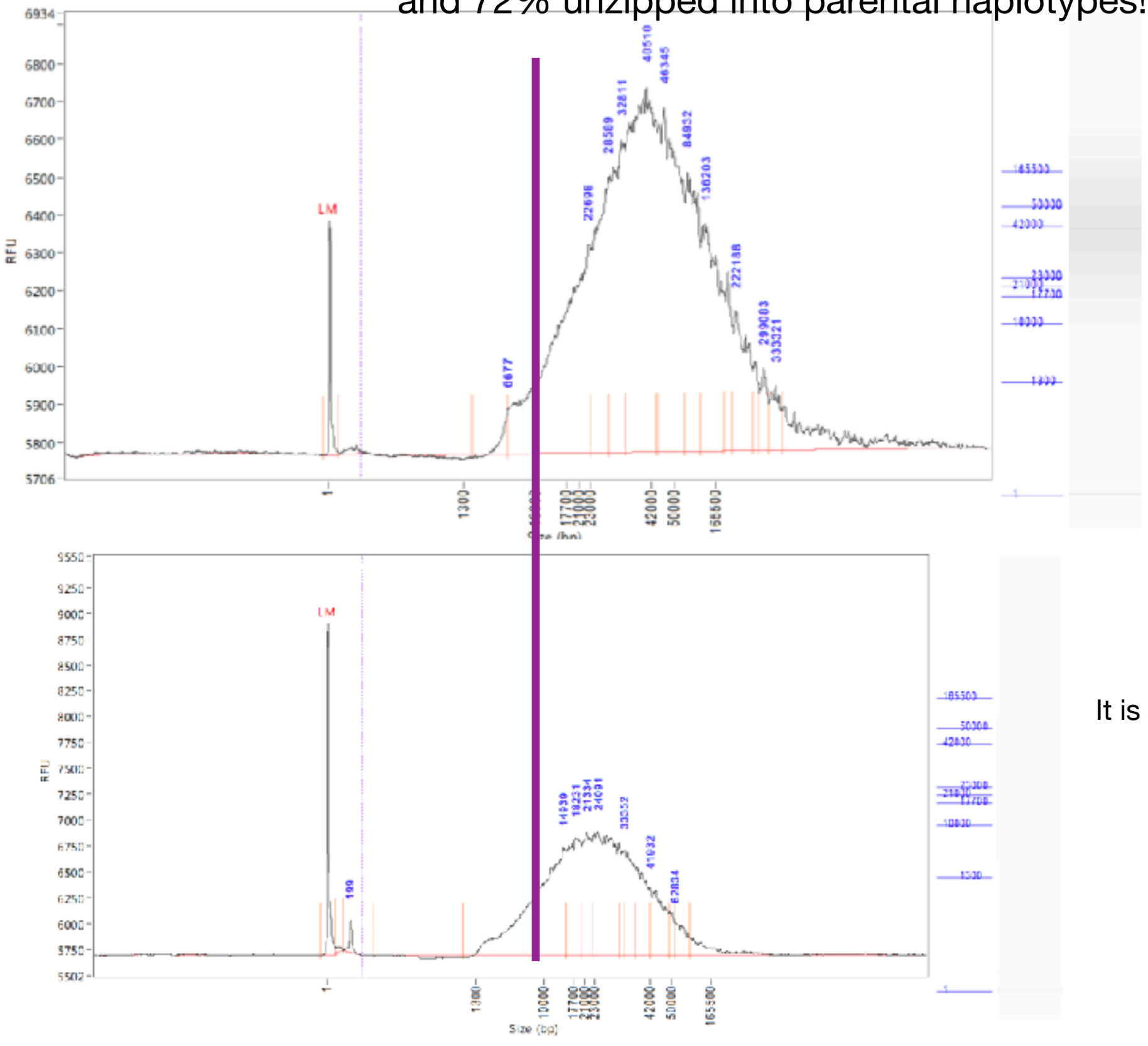
Test two DNA extractions (5.1 and 8.10) at Sanger to replicate mosquito number 7 success. “Standard” MagAttract (5.1) shown here.



5.1 pre-cleanup (0.45XAMPure)
Estimated to have 380ng by Femto
(800ng by Picogreen)

5.1 post-cleanup

“Standard” MagAttract (5.1) gives a much better looking library. Sequenced on 3 SMRT Cells and 72% unzipped into parental haplotypes!



5.1 post-cleanup

5.1 library
Sequenced this.
It is a wild *An. arabiensis* mosquito from Kenya

For now, we use the MagAttract Kit “Standard” protocol for good high molecular weight extractions well-suited for Low Input PacBio library prep.

Take care with extractions…no twisting when grinding with pestle (3-5 up/down squish strokes), use extremely slow pipetting with wide bore tips, don’t drop tubes, keep DNA at 4C if using soon (avoid freeze/thaws).

The specimen must be fresh or else killed and immediately stored in -80C. (We are now in the process of evaluating 7 different storage buffers and 4 different temperature conditions for the possibility of longer term storage when extracting from fresh is not possible….this is coming soon.)

Plans for new Anopheles vector mosquito assemblies

Testing multiplexing on the 8M Smart cell: using single mosquitoes from an F1 hybrid cross between colony mosquitoes (mother=*An. coluzzii*, father = *An. gambiae*).

Collaboration with Dan Neafsey and a variety of partners in South America and Africa to generate a “blueprint” for creating high quality reference genomes for Anopheles mosquitoes. Funded by Bill and Melinda Gates Foundation.

Partners are collecting wild bloodfed females, believed to have mated once, from which they will rear up a full brood and:

One offspring will be kept for low input PacBio sequencing.
One offspring will be kept as morphological voucher.
One offspring will be kept for RNA sequencing for genome annotation.

All other offspring will be pooled for technologies that still require higher amounts of DNA (e.g., bionano, ONT, Hi-C) and are necessary for piecing together chromosomes.

Connecting communities

Matt Loose and Nick Loman have started **@longreadclub longreadclub.org**
will include videos on youtube channel.
post protocols on protocols.io
chat on social media @longreadclub

all great…still a gap in easy chat so I have set up a fully open slack workspace:

“all.things.up.to.assembly” (e.g., sample preservation, sample extractions, sequencing technologies, modifications to reduce input required, etc).

atuta.slack.com

invite link (due to never expire) is:

https://tinyurl.com/y4dvuo74

y4dvuo74

acknowledgements

PacBio
Jonas Korlach
Sarah Kingan
Primo Baybayan
Christine Lambert
Brendan Galvin

Lawniczak lab
Haynes Heaton
Katharina Wyschetski
Juli Cudini

Sanger Long Read Team
Karen Oliver
Michelle Smith
Craig Corton