Malaria protection due to sickle haemoglobin depends on parasite genotype

Gavin Band

ZJU Global Dialogue Series

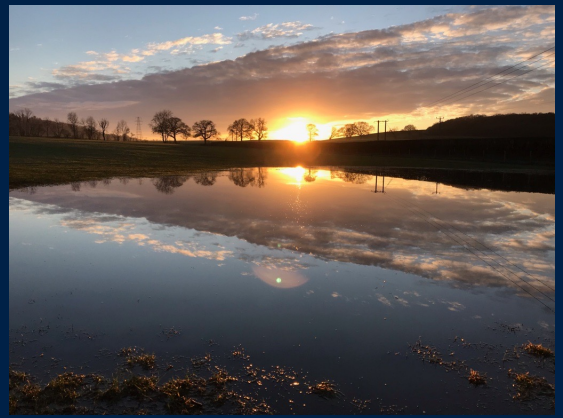
14th July 2022

A bit about Oxford

wellcome centre human genetics

COI
Chinese Academy of Medical Sciences Oxford Institute

Around Oxford



This talk

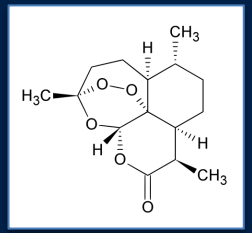
- Background on malaria
- Why study malaria genetics?
- A search for human-parasite genetic interactions
- Parasite population genetics
- Toward biological function

Malaria remains a major killer

WHO >600,000 deaths due to *P.falciparum* malaria in 2020
The vast majority are among young children in Africa.

Current frontline treatments are based on the compound **artemisinin,** famously rediscovered in the 1970s by screening traditional Chinese medicines.

Tu Youyou
Nobel Prize 2015

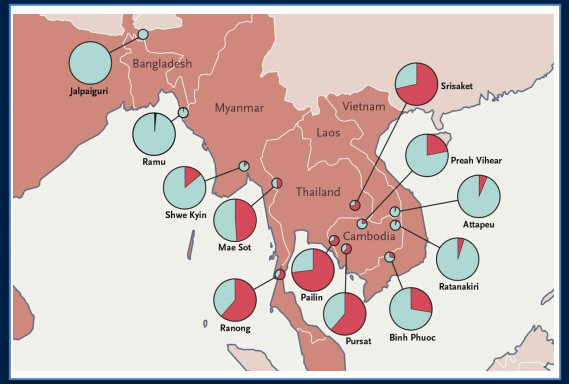


青蒿素
(Qinghaosu)
Artemisinin

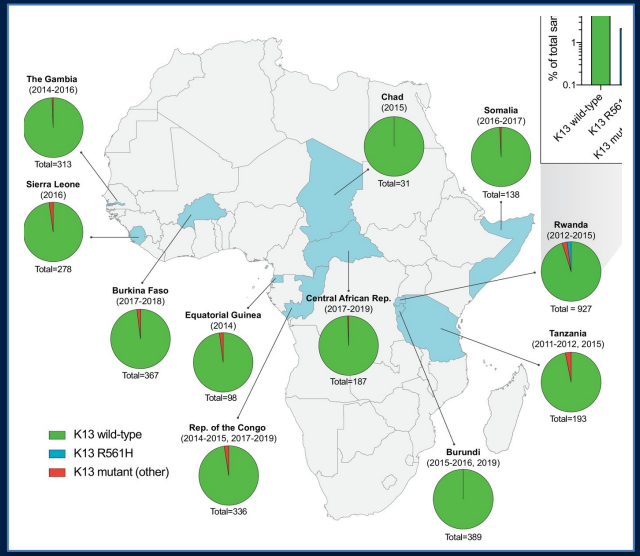
Artemisia annua

Artemisinin works but…

Resistance to artemisinin and other antimalarials is rising



Frequency of
artemisinin resistance



"Recent evidence of the independent emergence of artemisinin partial resistance in [Africa] is of great global concern" - WHO

Figures: Ashley et al, N. Engl. J. Med. 2014, Stokes et al, eLife 2021

Why study malaria genetics?

1. Uncover new aspects of **infection biology**
2. **Study natural selection** and **evolution**

Protection against infection…

Sickle haemoglobin (HbS)

A mutation of the *HBB* gene

PROTECTION AFFORDED BY SICKLE-CELL TRAIT AGAINST SUBTERTIAN MALARIAL INFECTION

BY

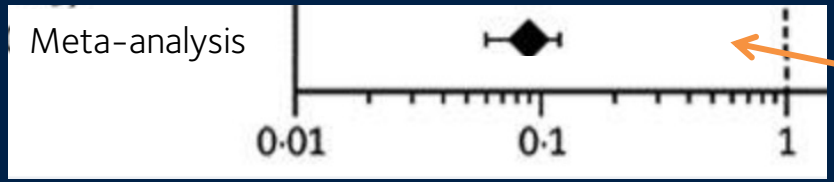
A. C. ALLISON, D.Phil., B.M.*

(From the Clinical Pathology Laboratory, the Radcliffe Infirmary, Oxford)

(1954)

TABLE I

	With Parasitaemia	Without Parasitaemia	Total
Sicklers	12 (27·9%)	31 (72·1%)	43
Non-sicklers ..	113 (45·7%)	134 (53·3%)	247

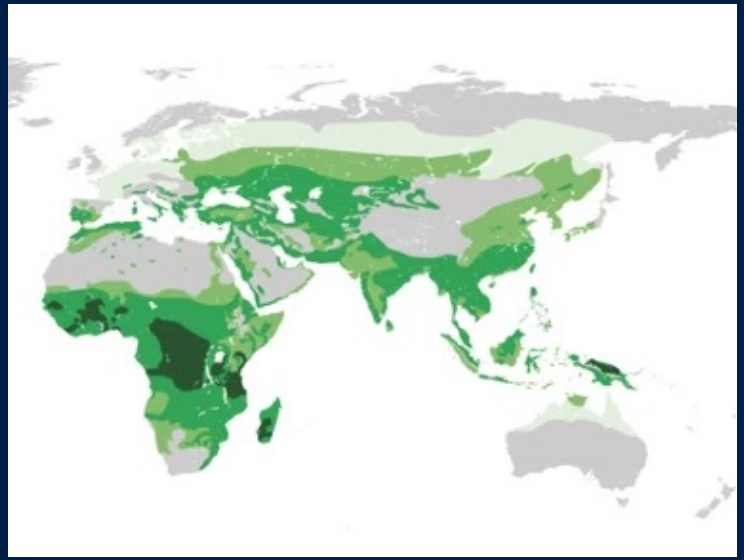


Taylor et al Lancet Infect. Dis 2012

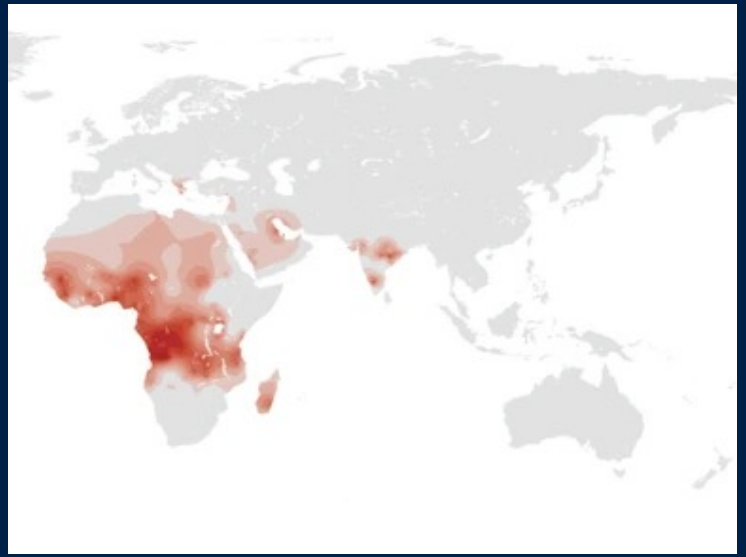
Carrying one copy of HbS is thought to have about a **ten-fold protective effect**

Protection against infection… and natural selection

Malaria endemicity



Frequency of HbS



Piel et al Nature Communications (2010)

O blood type
frequency ~ 50%
RR ~ 0.75
(recessive)

Dantu blood type
frequency ~ 0-10%
RR ~ 0.6
(additive)

ATP2B4 calcium pump variation
frequency ~ 50%
RR ~ 0.66
(recessive)

Sickle haemoglobin (HbS)
Frequency ~ 2-20%
RR ~ 0.1-0.2
(heterozygote)

Allison Br Med. J. (1954)

Estimates from our multi-population GWAS of severe malaria susceptibility - MalariaGEN Nature Communications 2019

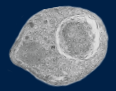
O blood type

Dantu blood type

ATP2B4 calcium pump variation

Sickle haemoglobin (HbS)

?



P.falciparum
23Mb genome
>5,000 genes
millions of variants

Have parasite populations adapted?

(And is this detectable in current populations?)

Searching for host-parasite genetic interactions

Band et al Nature 2021 doi.org/10.1038/s41586-021-04288-3

Plan of attack:

1. Take DNA samples from a large set of severe malaria cases.
   (These were previously collected and contain both human and parasite DNA)

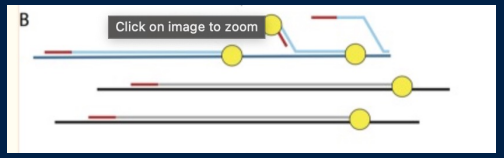
2. Amplify and sequence the parasite DNA.

3. Test for association between human and parasite genetic variants

*MalariaGEN human GWAS, Nature Communications 2019

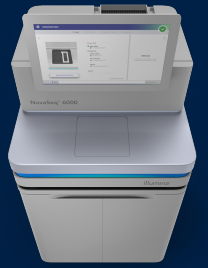
Quick genome sequencing cheat sheet

1. Take DNA sample



2. Amplify parasite DNA
Using A/T-rich primers

3. Fragment DNA into
small (400-500bp)
pieces!



4. Sequence

5. Computationally align reads back to reference genome

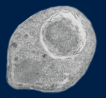
Complex region (difficult to align)

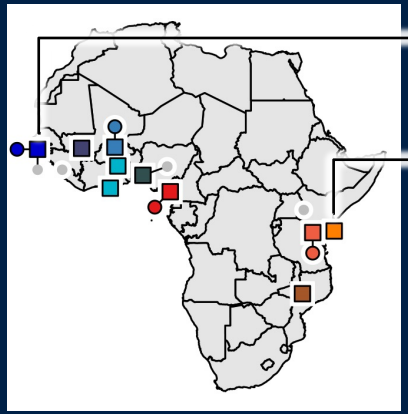
Sequencing error

A mutation relative to the reference

Investigating human-parasite genetic interaction in severe malaria cases

1. Sequence the *P.falciparum* genome
   in severe malaria cases from our previously published human GWAS

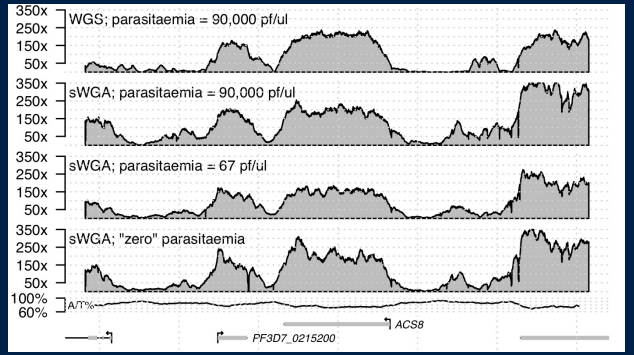




Banjul, The Gambia
N = 2,721

Kilifi, Kenya
N = 2,375

Collected in 1995-2009



Variant calling and quality control

Previously generated human genome-wide genotypes and imputation

Overlap with human data
N = 3,346 samples

2. Test for association pairwise between human and *Pf* variants using a simple logistic regression framework:

$$g_{Pf} \sim g_{\text{human}} + \text{country}$$

Software at: www.well.ox.ac.uk/~gav/hptest

Focus on candidates:

- Known protective mutations
- Further putative associations
- Blood group gene variants
- HLA alleles

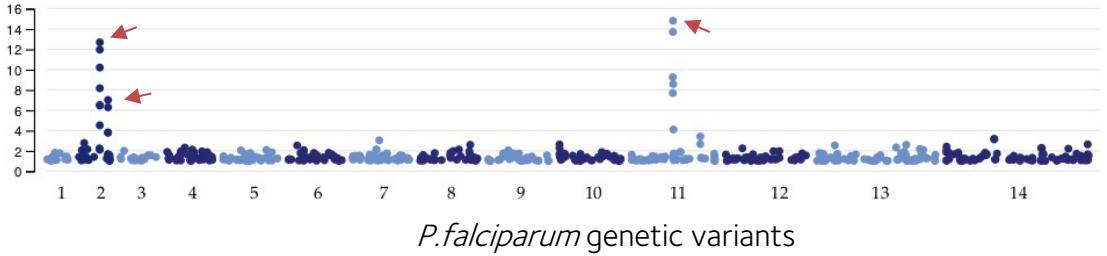
Focus on ‘easy’ parts:

- Biallelic variants in core genome
- Seen in at least 25 infections across the sample.
- 51,552 variants in total

(…excludes multiallelics and complex regions)

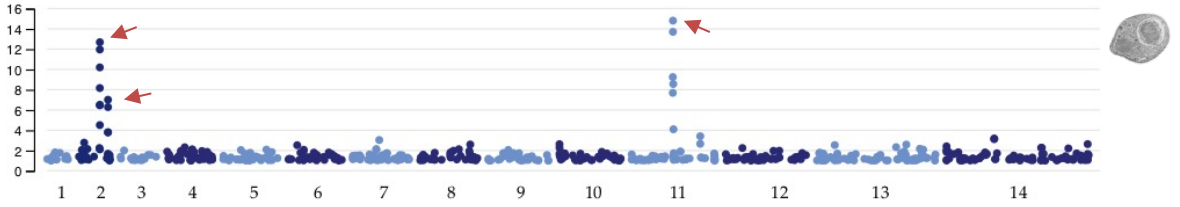
Three regions of the Pf genome are associated...

Evidence for association
for *P.falciparum* variants
(averaged over human variants)

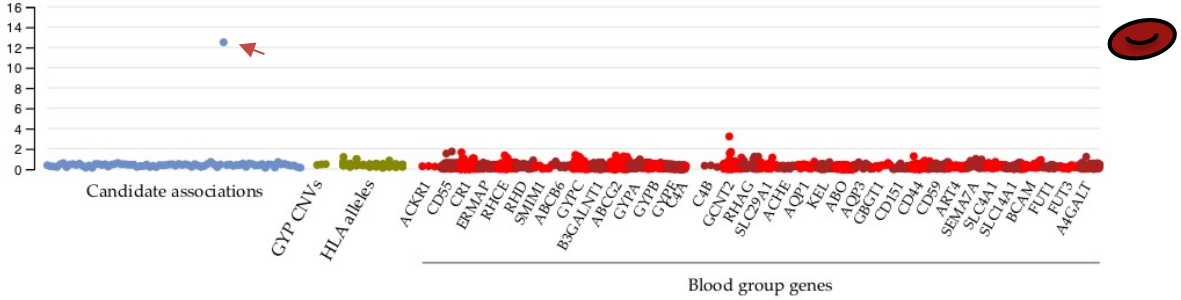


Three regions of the Pf genome are associated with...

Evidence for association
for *P.falciparum* variants
(averaged over human variants)



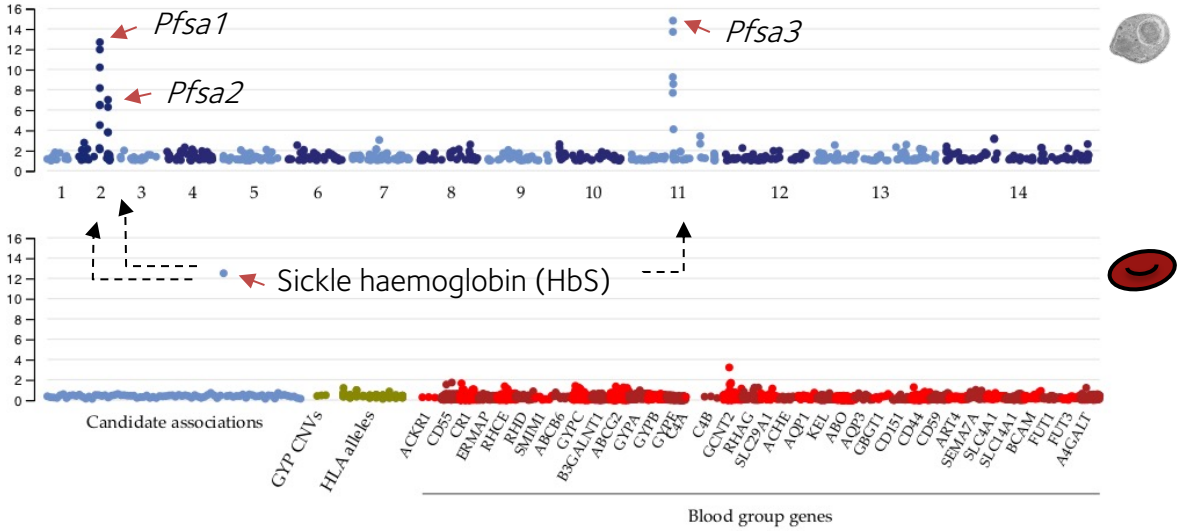
Evidence for association
for human variants
(averaged over *Pf* variants)



Three regions of the Pf genome are associated with HbS

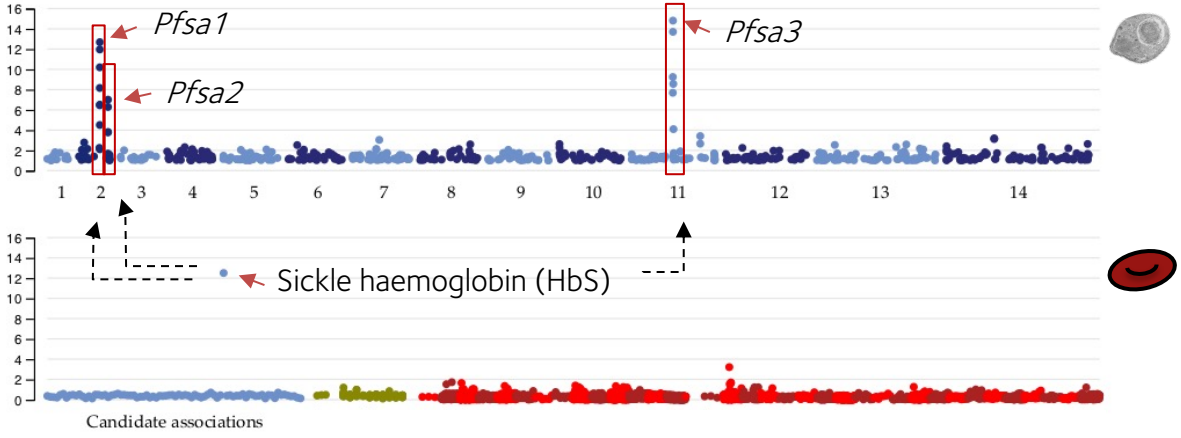
Evidence for association
for *P.falciparum* variants
(averaged over human variants)

Evidence for association
for human variants
(averaged over *Pf* variants)



Three regions of the Pf genome are associated with HbS

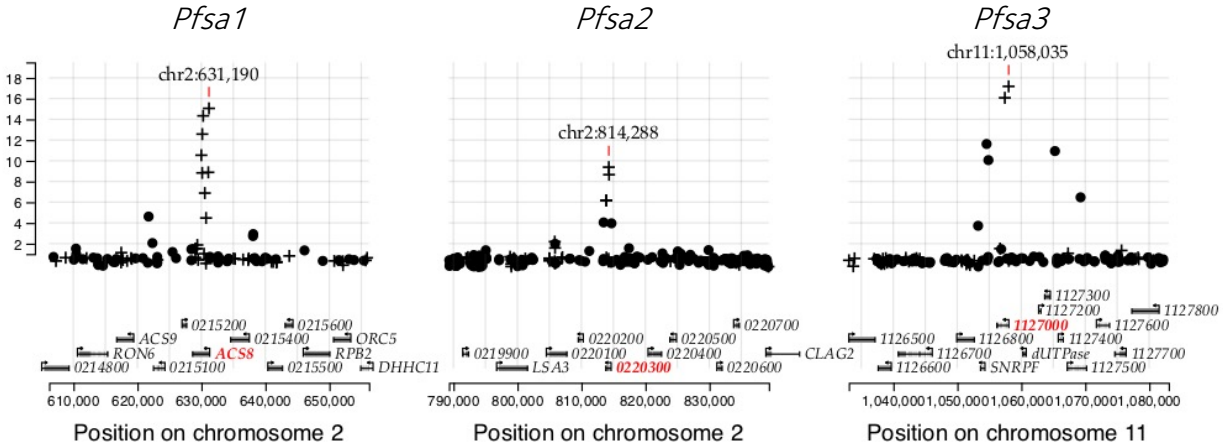
Evidence for association
for *P.falciparum* variants
(averaged over human variants)



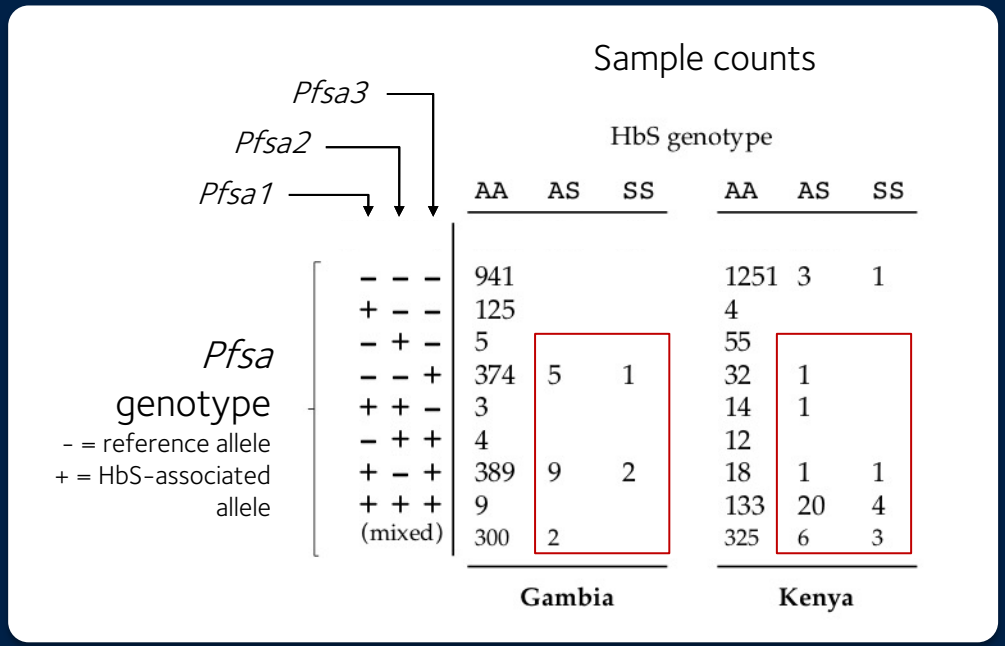
Evidence for association
for human variants
(averaged over *Pf* variants)

Zoom in to
Pf genome:

Evidence for association
for *P.falciparum* variants
with HbS



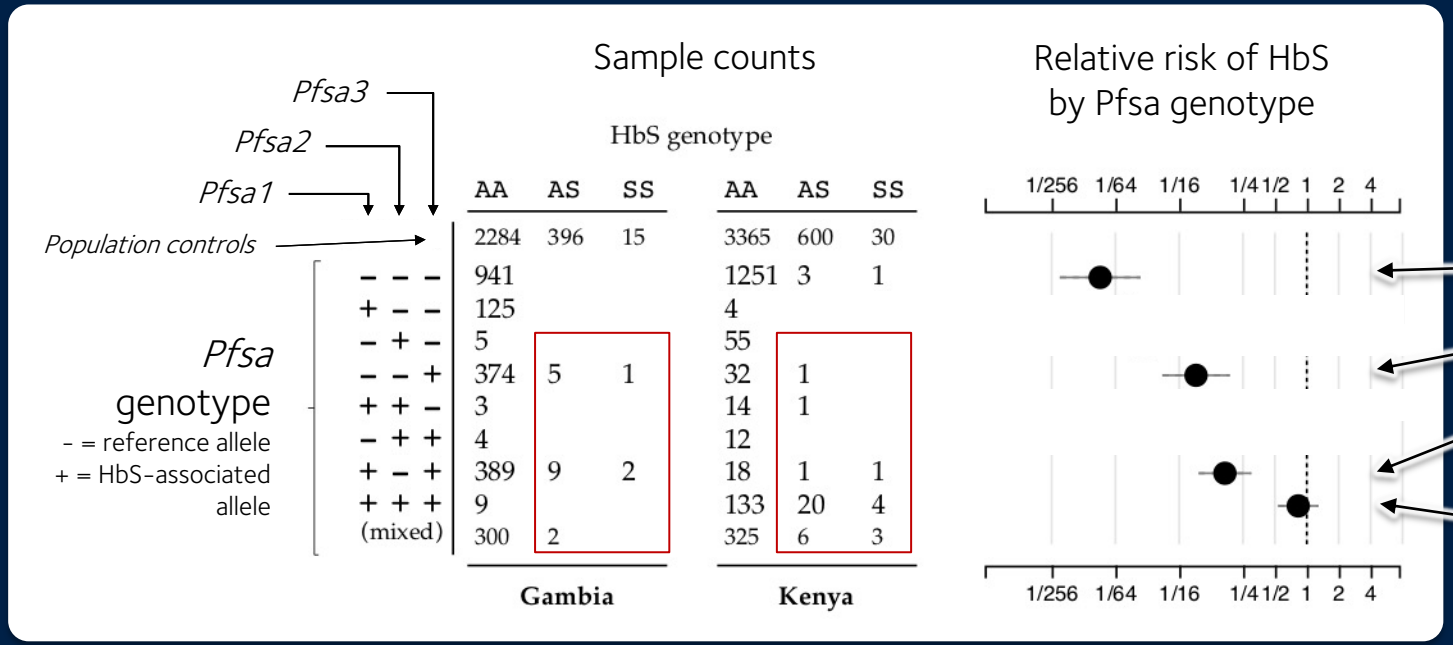
The protective effect of HbS varies with *Pfsa* genotype



$N = 4{,}071$ severe malaria cases

45 of 49 severe infections of individuals with HbS genotypes were with *Pfsa+* parasites

The protective effect of HbS varies with *Pfsa* genotype



45 of 49 severe infections of individuals with HbS genotypes were with *Pfsa+* parasites

+ caveats!

Parasite population genetics

Pfsa frequencies vary widely within and between populations

DOI: 10.1111/mec.15706

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Describing the current status of *Plasmodium falciparum* population structure and drug resistance within mainland Tanzania using molecular inversion probes

Tanzania - Moser et al,Molecular Ecology 2020

Check for updates

ARTICLE

https://doi.org/10.1038/s41467-020-15779-8 OPEN

The impact of antimalarial resistance on the genetic structure of *Plasmodium falciparum* in the DRC

DRC - Verity et al Nat. Comm 2020

Wellcome Open Research

Wellcome Open Research 2021, 6:42 Last updated: 29 MAR 2021

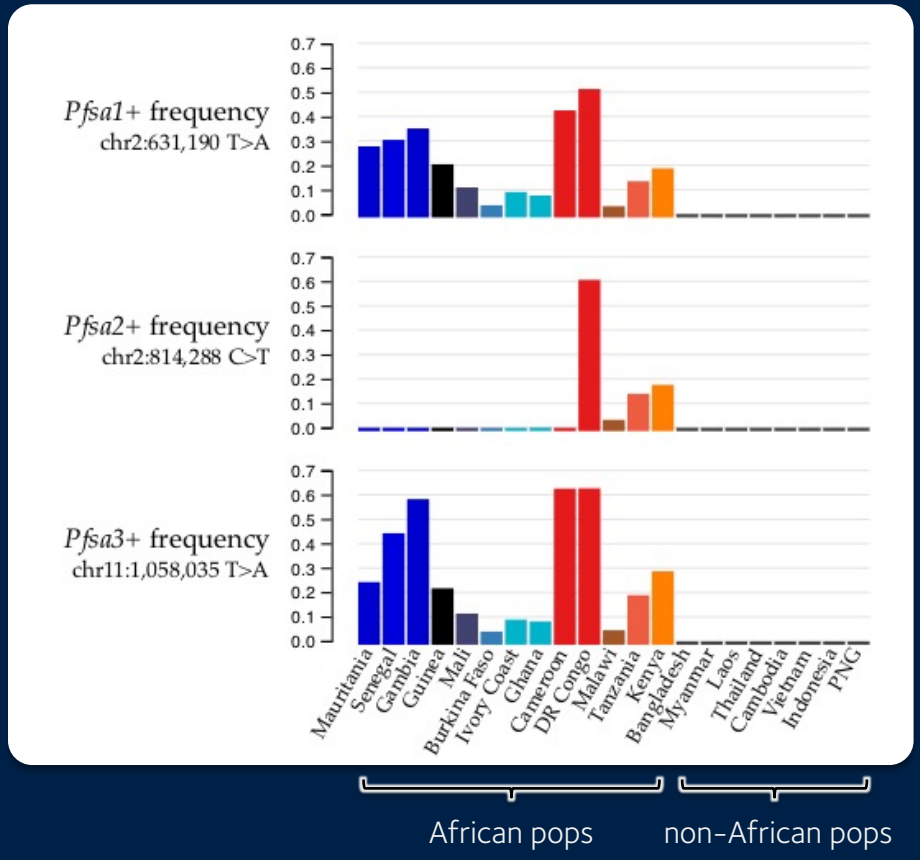
Check for updates

RESEARCH ARTICLE

An open dataset of *Plasmodium falciparum* genome variation in 7,000 worldwide samples

MalariaGEN Pf6

Wellcome Open Research 2021



Pfsa frequencies vary widely within and between populations

DOI: 10.1111/mec.15706

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Describing the current status of *Plasmodium falciparum* population structure and drug resistance within mainland Tanzania using molecular inversion probes

Tanzania - Moser et al,Molecular Ecology 2020

Check for updates

ARTICLE

https://doi.org/10.1038/s41467-020-15779-8 OPEN

The impact of antimalarial resistance on the genetic structure of *Plasmodium falciparum* in the DRC

DRC - Verity et al Nat. Comm 2020

Wellcome Open Research

Wellcome Open Research 2021, 6:42 Last updated: 29 MAR 2021

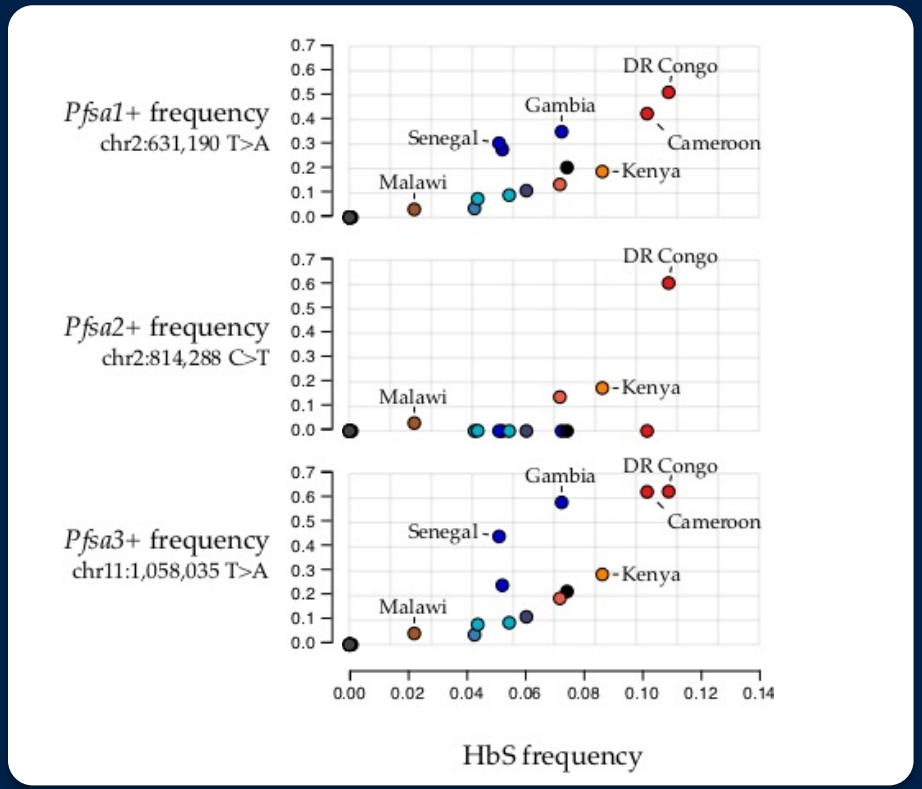
Check for updates

RESEARCH ARTICLE

An open dataset of *Plasmodium falciparum* genome variation in 7,000 worldwide samples

MalariaGEN Pf6

Wellcome Open Research 2021



(Malaria Atlas Project - Piel et al Lancet 2013)

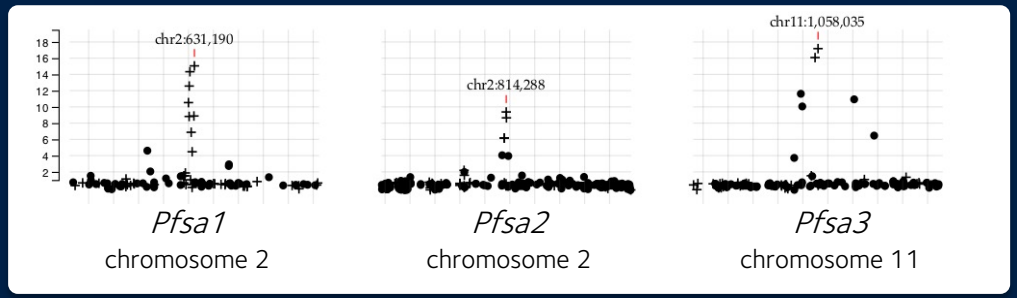
The *Pfsa* alleles are in strong linkage disequilibrium i.e. they co-occur

r=0.80

Kenya:

r=0.75

r=0.66



Gambia:

r=0.43

Correlation between *Pfsa+* alleles in severe malaria cases...

After excluding HbS individuals

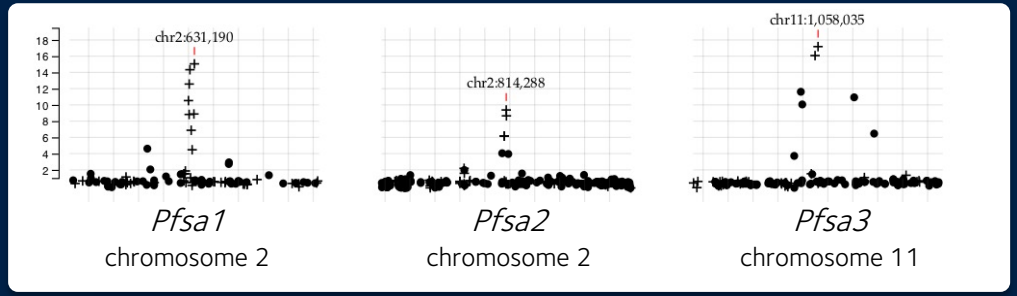
The *Pfsa* alleles are in strong linkage disequilibrium i.e. they co-occur

r=0.80

Kenya:

r=0.75

r=0.66



Gambia:

r=0.43

Correlation between *Pfsa+* alleles in severe malaria cases...

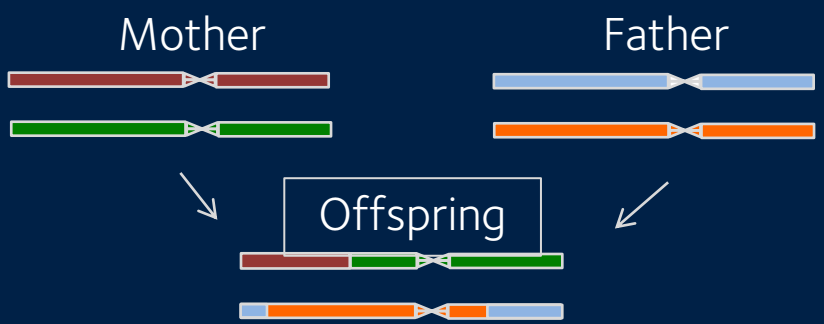
After excluding HbS individuals

...or in milder infections:

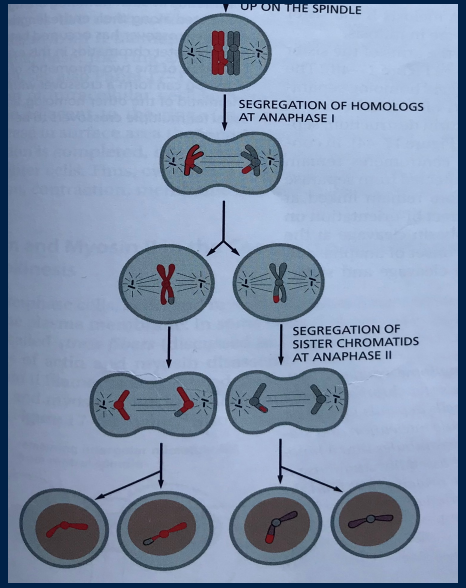
Pfsa1+ vs *Pfsa3+*

Country	N	r
Gambia	169	0.20
Guinea	133	0.79
Mali	379	0.84
Ghana	807	0.86
Cameroon	174	0.52
Congo	241	0.64
Malawi	239	0.79
Tanzania	282	0.59
Kenya	89	0.71

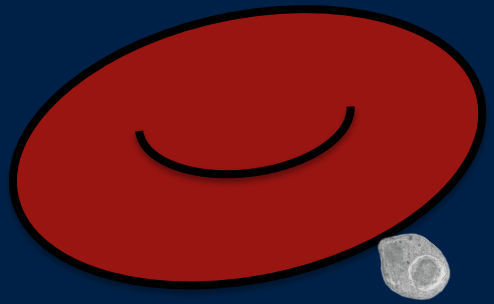
MalariaGEN Pf6



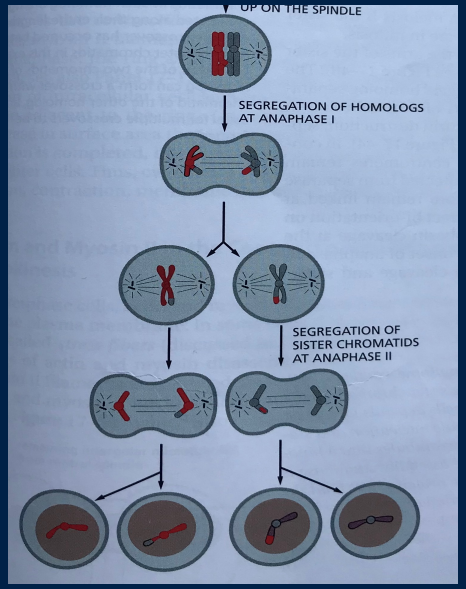
Meiosis (germline cell division)



Random assortment of chromosomes and chromosome segments



Parasites are haploid (one genome copy) and replicate clonally

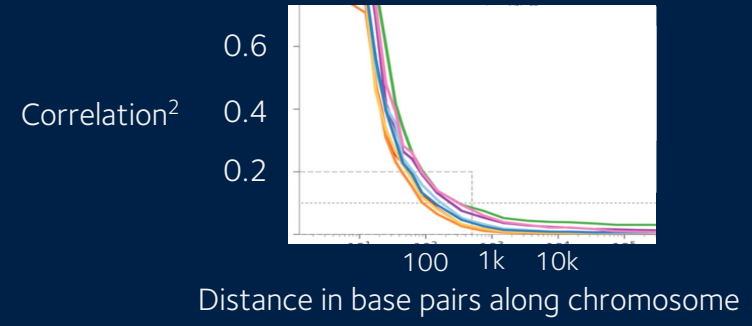


Parasites fuse and undergo meiosis

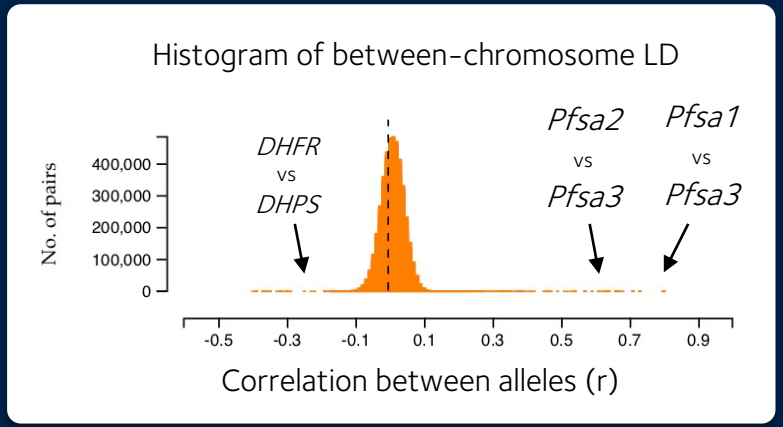
The *Pfsa* alleles are in strong linkage disequilibrium i.e. they co-occur

Parasites undergo sexual reproduction (meiosis) in mosquitos

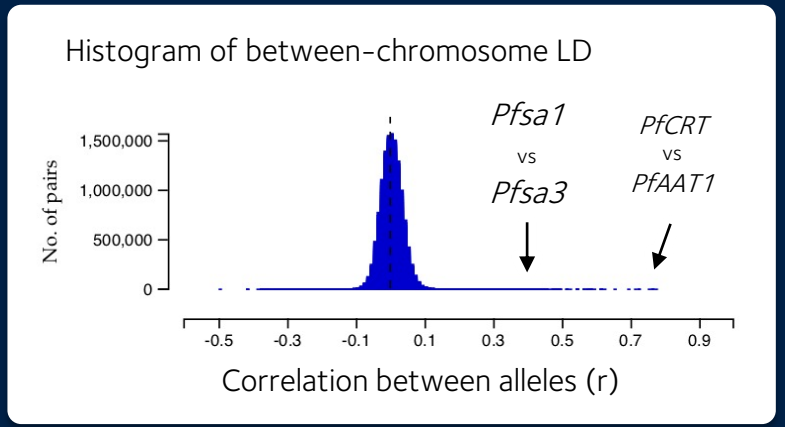
This breaks down correlations over short genetic distances in the genome.



Kenya:



Gambia:

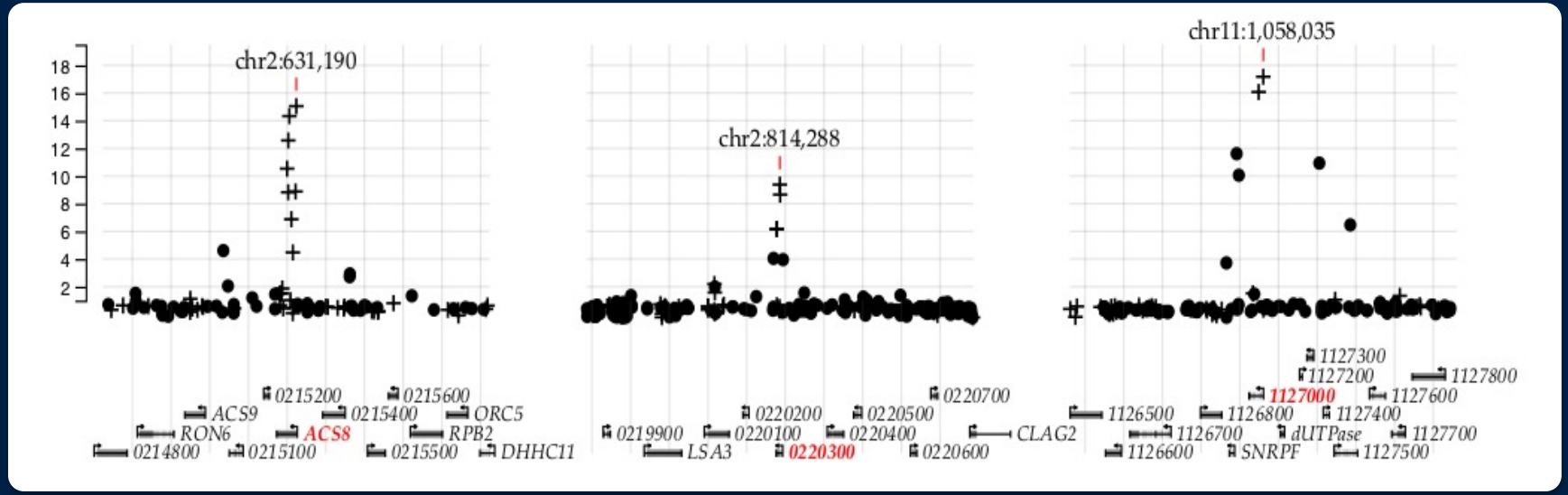


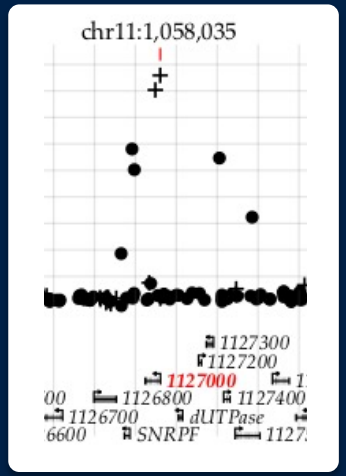
Hypothesis

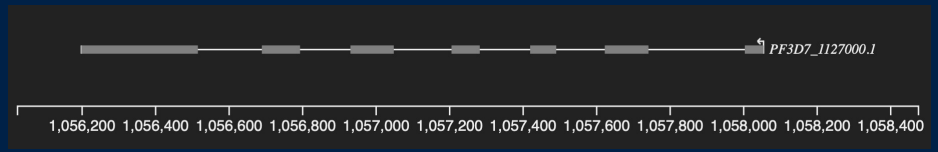
- *Pfsa*-carrying parasites are able to infect and cause disease in HbS-carrying individuals
- They are presumably +vely selected in individuals carrying HbS.
- Yet they have not been driven to 100% in any population
- It is possible that *epistasis* (fitness interaction between the three loci) is one of the factors contributing to this.

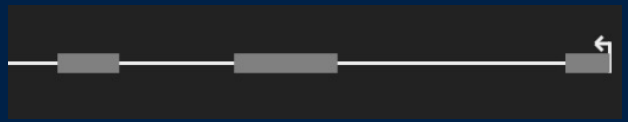
Towards understanding biological function

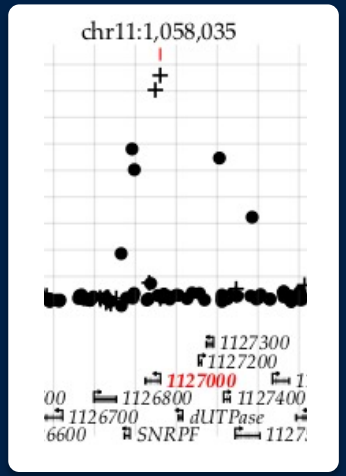
Puzzles and questions

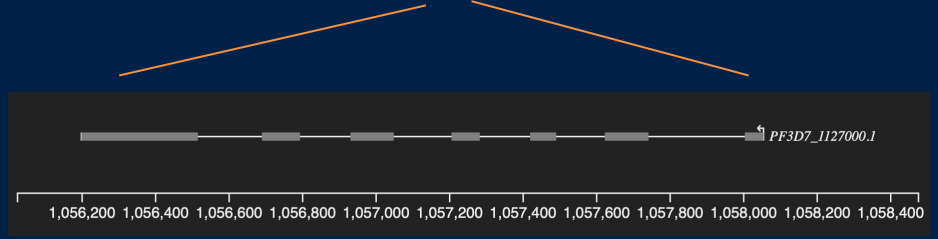


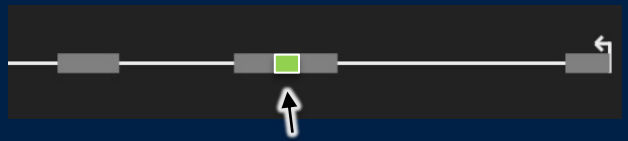

chr11:1,058,035
1127300
1127200
1127000
1126800
1127400
1126700
dUTPase
SNRPF


PF3D7_1127000.1
1,056,200 1,056,400 1,056,600 1,056,800 1,057,000 1,057,200 1,057,400 1,057,600 1,057,800 1,058,000 1,058,200 1,058,400

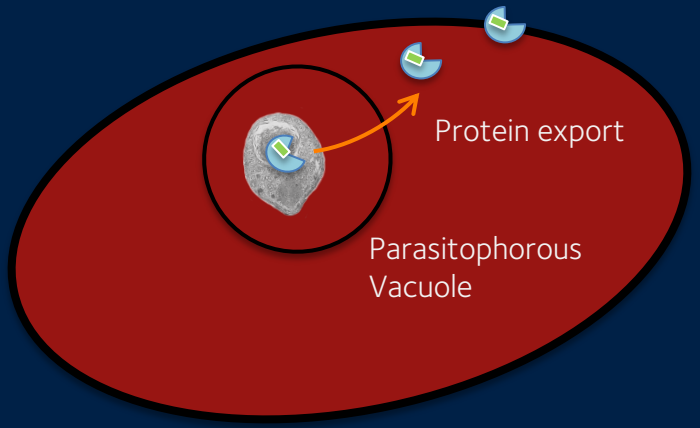






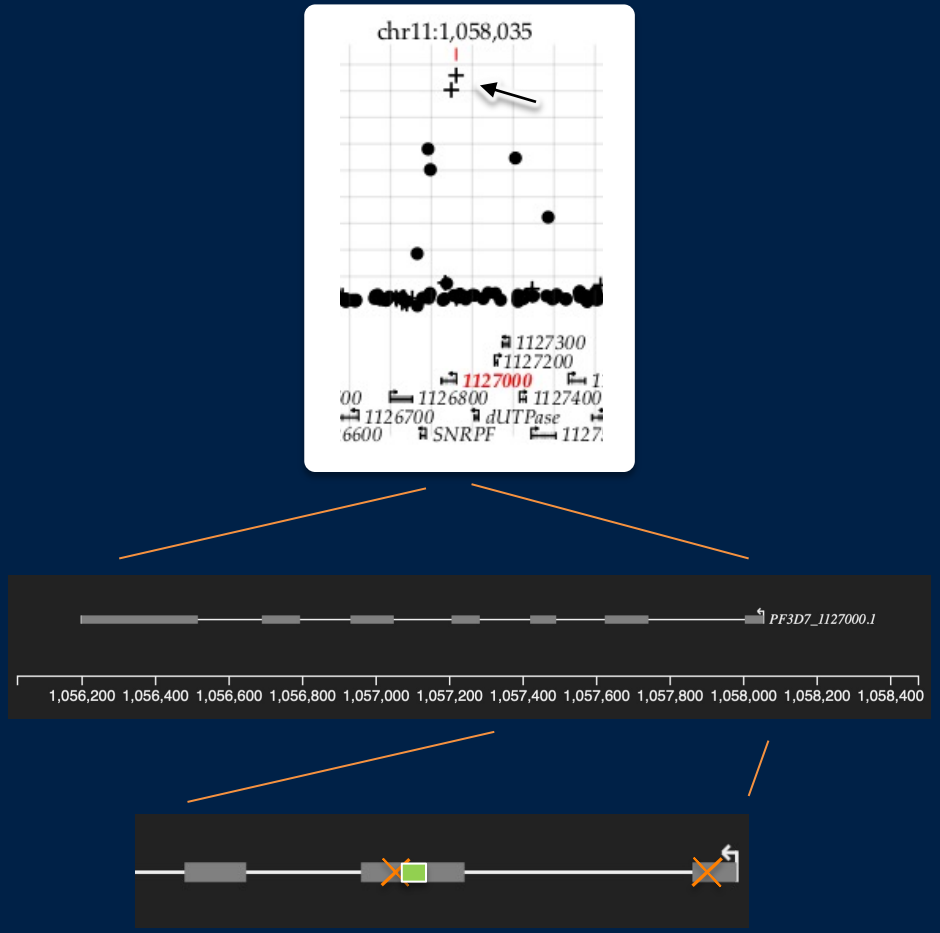


'PEXEL' amino acid motif

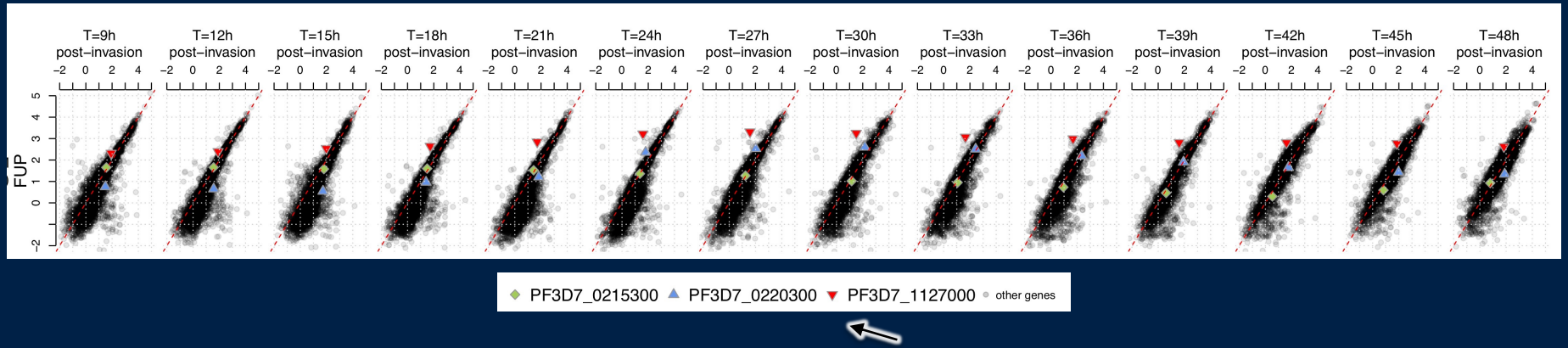


Red cell

The 1127000 protein product contains a Plasmodium export element (PEXEL) motif. It is predicted to be exported to the host cytosol.



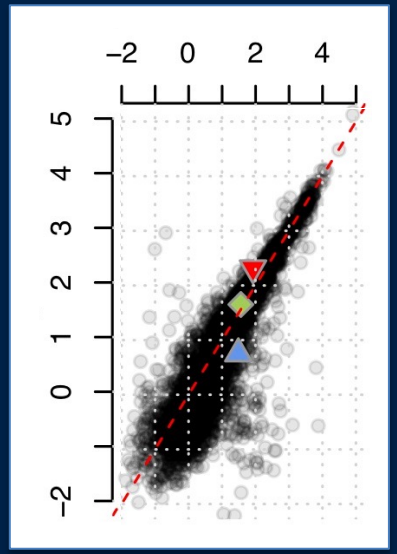
Hypothesis: the Pfsa+ mutations affect export of the protein to the host cell (likely increasing the degree of export).



Hypothesis: the *Pfsa+* mutations increase expression of the1127000 gene (and presumably the corresponding protein)

Parasite gene expression (RNA levels)

sickle-associated (*Pfsa+*) parasite



Pfsa- parasite

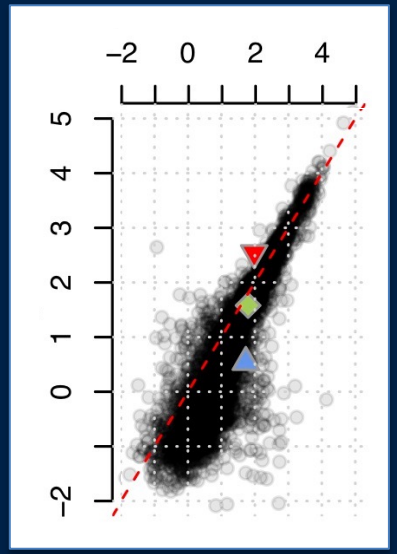
T=9

Hours post-invasion

PF3D7_1127000 other genes

The parasite takes about 48 hours to replicate within red cells (then they burst and the parasites re-invade). What does gene expression look like across this cycle?

sickle-associated (*Pfsa+*) parasite



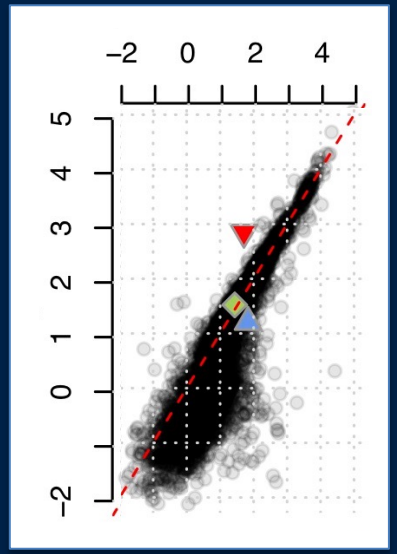
Pfsa− parasite

T=15

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated (*Pfsa+*) parasite



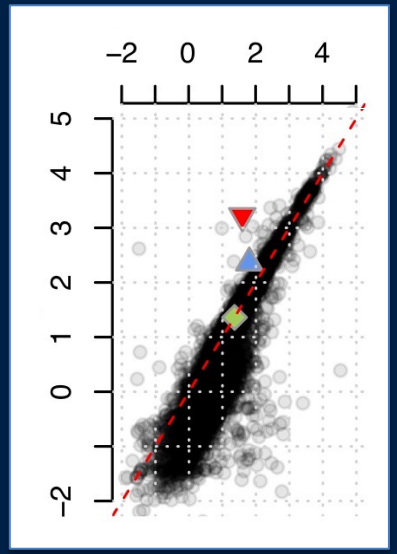
Pfsa− parasite

T=21

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated (*Pfsa+*) parasite



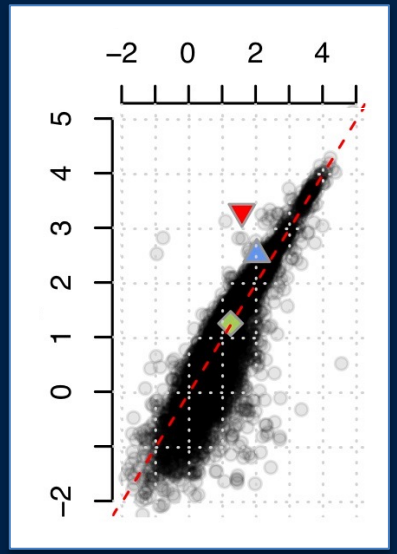
Pfsa- parasite

T=24

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated (*Pfsa+*) parasite



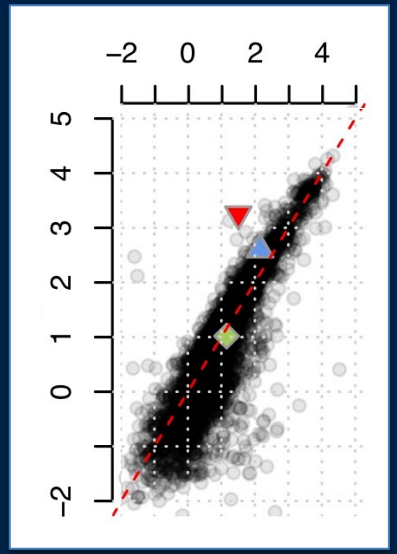
Pfsa- parasite

T=27

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated (*Pfsa+*) parasite



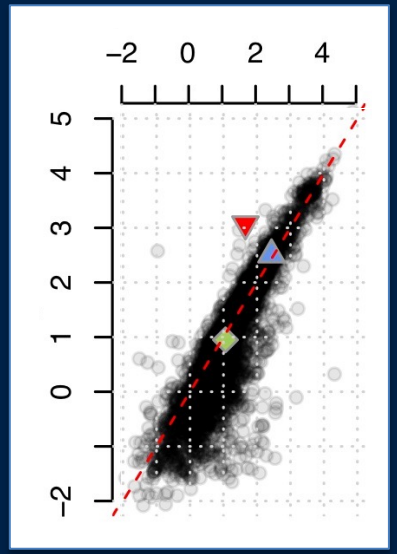
Pfsa− parasite

T=30

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated (*Pfsa+*) parasite



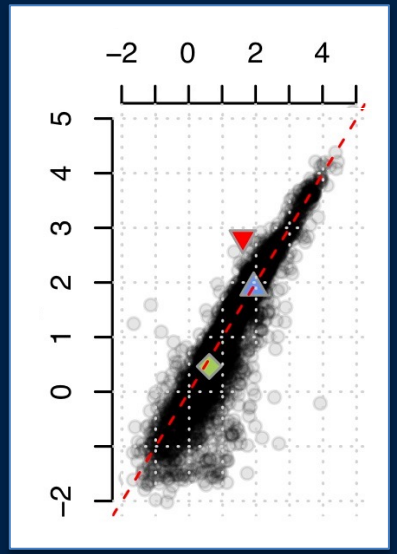
Pfsa− parasite

T=33

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated (*Pfsa+*) parasite



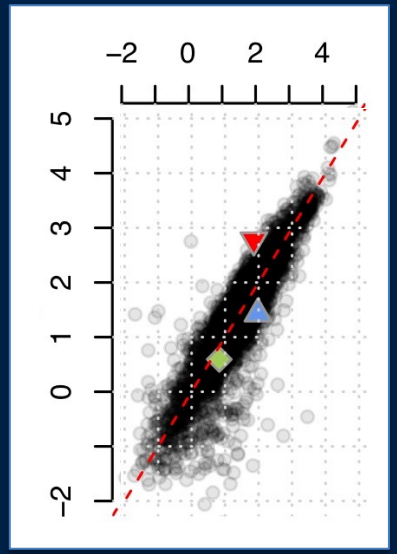
Pfsa− parasite

T=39

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated
(*Pfsa+*) parasite

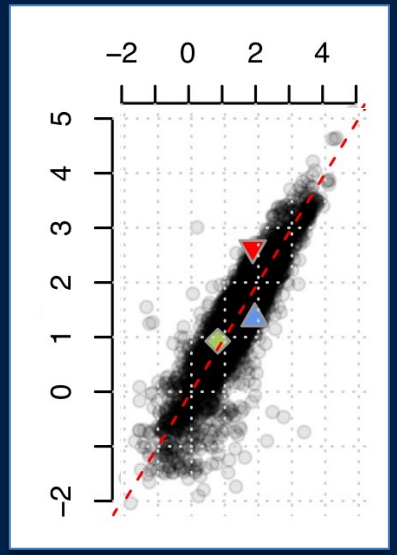


Pfsa- parasite

T=45

Hours post-invasion

PF3D7_1127000 other genes

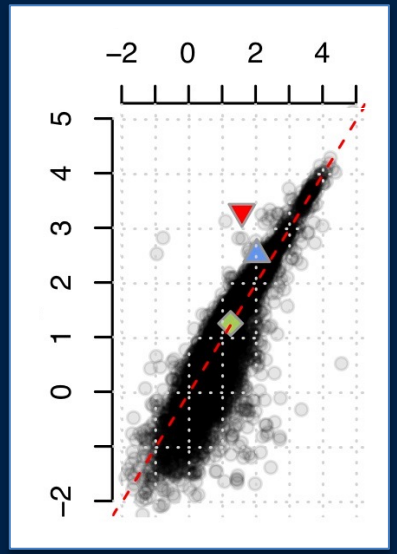


T=48

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated (*Pfsa+*) parasite



Pfsa− parasite

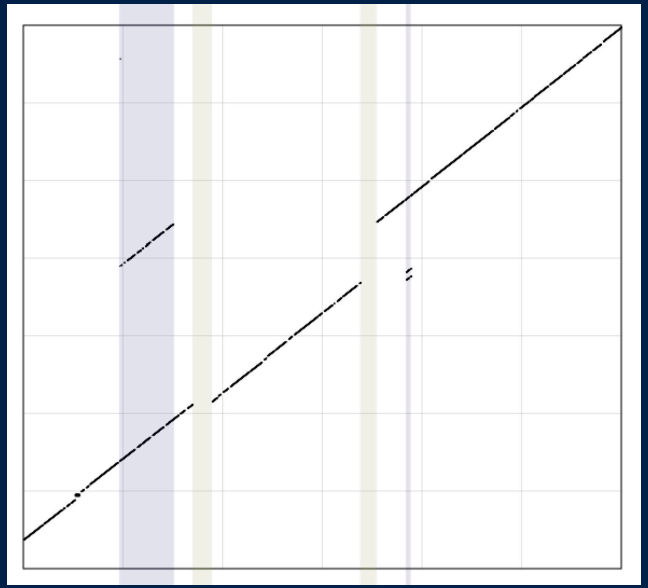
T=27

Hours post-invasion

▼ PF3D7_1127000 ● other genes

Hypothesis: the *Pfsa+* mutations increase expression of the1127000 gene at trophozoite stage (blood feeding / growth stage). And presumably the corresponding protein.

sickle-associated (*Pfsa+*) parasite



Pfsa- parasite

The Pfsa3 locus contains genome structural variation. Does that influence the expression?

Conclusions

- Studying malaria genetics can lead to new insights into infection biology
- There seem to be parasite mutations that can overcome the protective effects of sickle haemoglobin
- We do not know the biology of these variants yet, but have some clues.
- The variants have very unusual population genetic properties that tell a tale of natural selection – and possibly epistasis.

(and genetics is fun!)

Dominic Kwiatkowski

Kirk Rockett

Ellen Leffler

Muminatou Jallow

Tom Williams

MRC The Gambia @ LSHTM
Muminatou Jallow
Fatoumatta Sisay-Joof
Kalifa A. Bojang
David J. Conway
Giorgio Sirugo
Umberto d'Alessandro

KEMRI-Wellcome, Kilifi, Kenya
Carolyne M. Ndila
Alexandar W. Macharia
Norbert Peshu
Kevin Marsh
Thomas N. Williams

University of Oxford
Kirk A. Rockett
Anna E. Jeffreys
Christina Hubbart
Kate Rowlands
Dominic P. Kwiatkowski

Wellcome Sanger Institute
Thuy Nguyen
Sonia M. Goncalves
Cristina V. Ariani
Jim Stalker
Richard D.Pearson
Roberto Amato
Eleanor Drury
Dominic P. Kwiatkowski

Duke University
Steve M. Taylor
Joseph P. Saeelens

USTTB Mali:
Mahamadou Diakite

Thanks also:
Annie Forster
Jia-Yuan Zhang
Andre Python

www.malariagen.net

"We thank the patients and staff at the Paediatric Department of the Royal Victoria Hospital in Banjul, Gambia, and at Kilifi County Hospital and the KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya for their help with this study."

Publication at:
doi.org/10.1038/s41586-021-04288-3