CENTRE *for* HUMAN GENETICS

NUFFIELD DEPARTMENT *of* MEDICINE

UNIVERSITY OF OXFORD

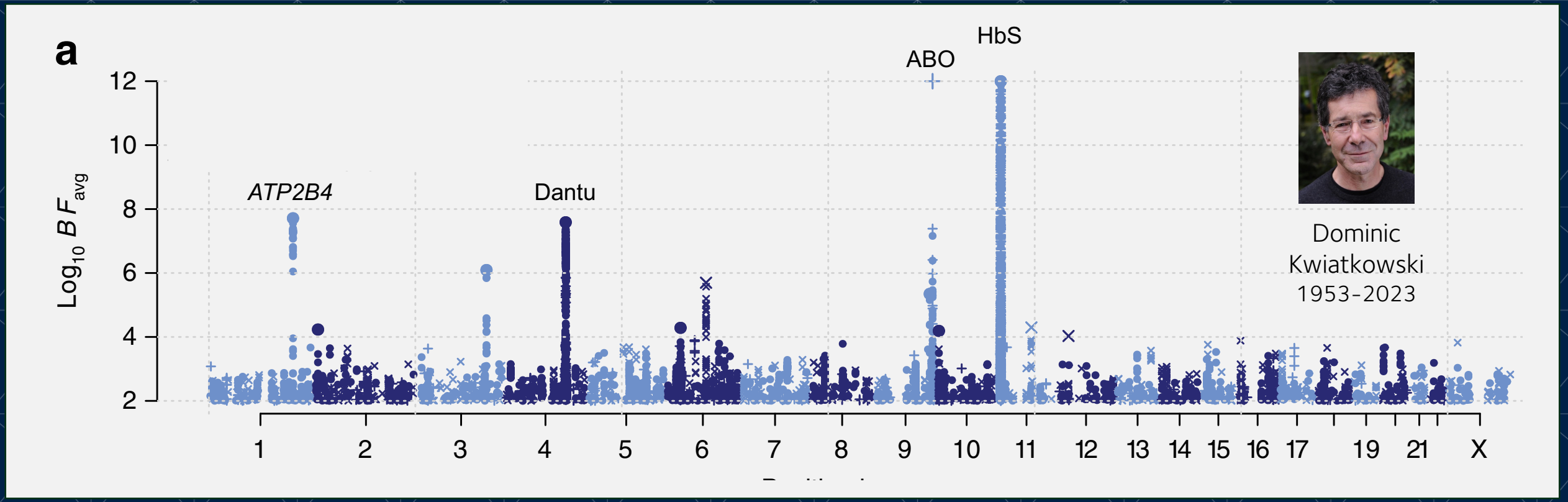
Malaria protection due to sickle haemoglobin depends on parasite genotype – but why?

Gavin Band

Laboratory of malaria & vector research
NIAID
14th May 2024

Oxford

CENTRE *for* HUMAN GENETICS



GWAS of severe malaria susceptibility - MalariaGEN Nature Communications 2019

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)

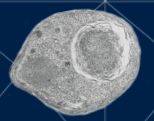
O blood type

Dantu blood type

ATP2B4 calcium pump variation

Sickle haemoglobin (HbS)

?



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

Parasite image: Miller et al J. Exp. Med. 1979 (with apologies!)

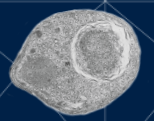
O blood type

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?)

Talk outline

- Malaria protection due to sickle haemoglobin depends on parasite genotype
- *Pfsa* population genetics
- *Pfsa* biology

Plan of our analysis:

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

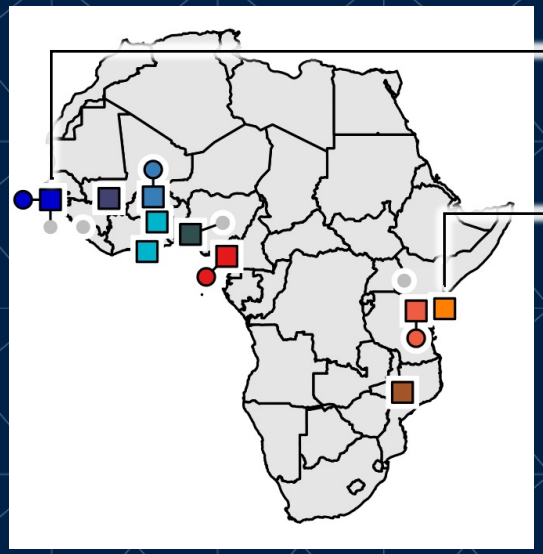
2. Test for association between human and parasite genetic variants

CENTRE *for* HUMAN GENETICS

*MalariaGEN human GWAS, Nature Communications 2019

Investigating human-parasite genetic interaction in severe malaria cases

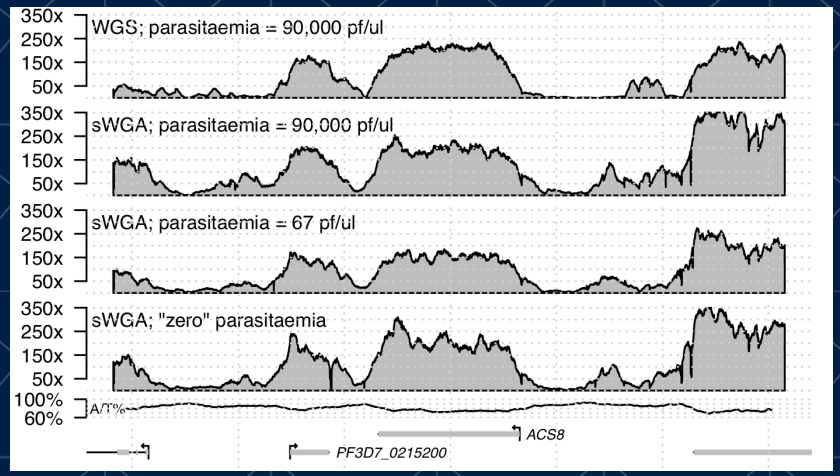
1. Sequence the *P.falciparum* genome in severe malaria cases from the published human GWAS



Banjul, The Gambia
N = 2,721

Kilifi, Kenya
N = 2,375

Collected in 1995-2009



Variant calling and quality control

Overlap with human data
N = 3,346 samples

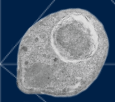
Previously generated human genome-wide genotypes and imputation

CENTRE *for* HUMAN GENETICS

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

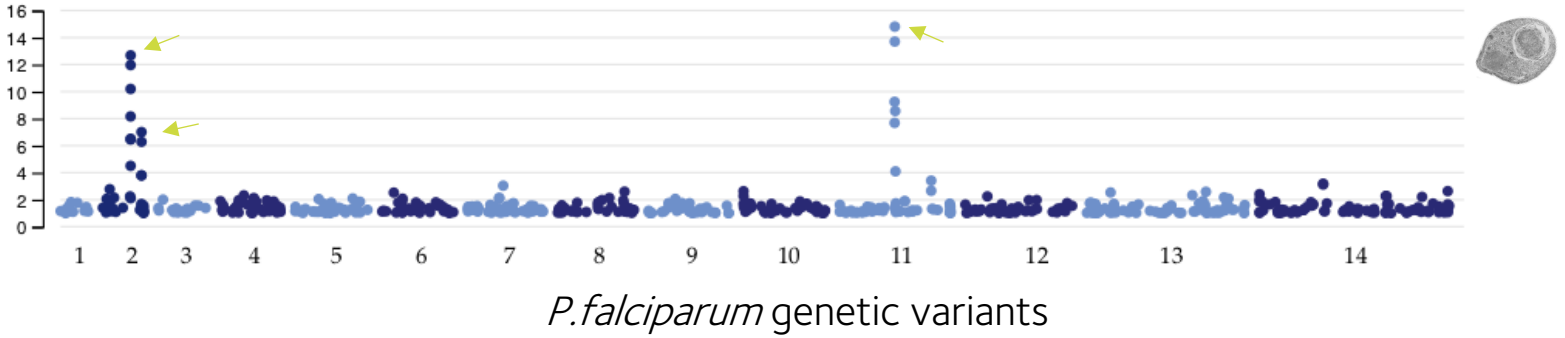
Focus on 'easy' parts:

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

(...excludes multiallelics and complex regions)

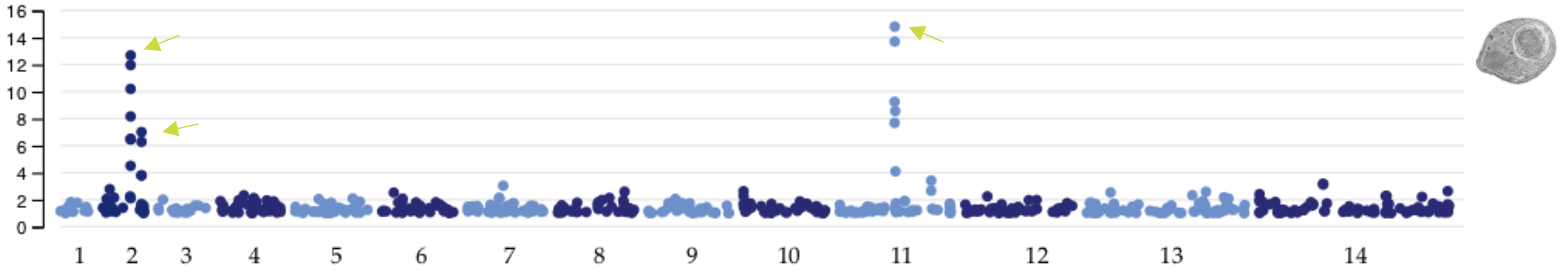
Three regions of the Pf genome are associated…

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

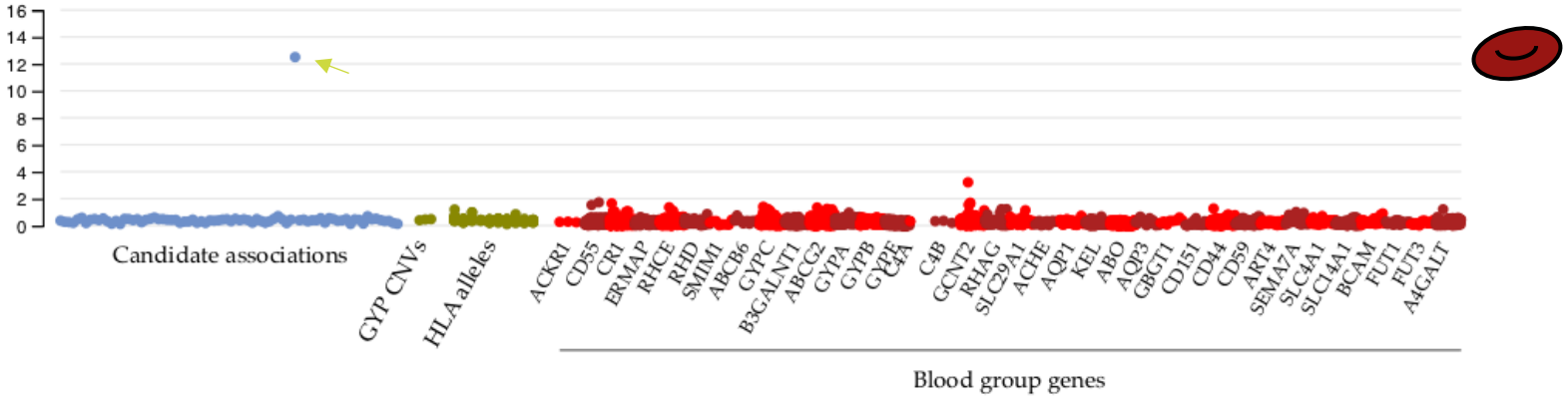


Three regions of the Pf genome are associated with…

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

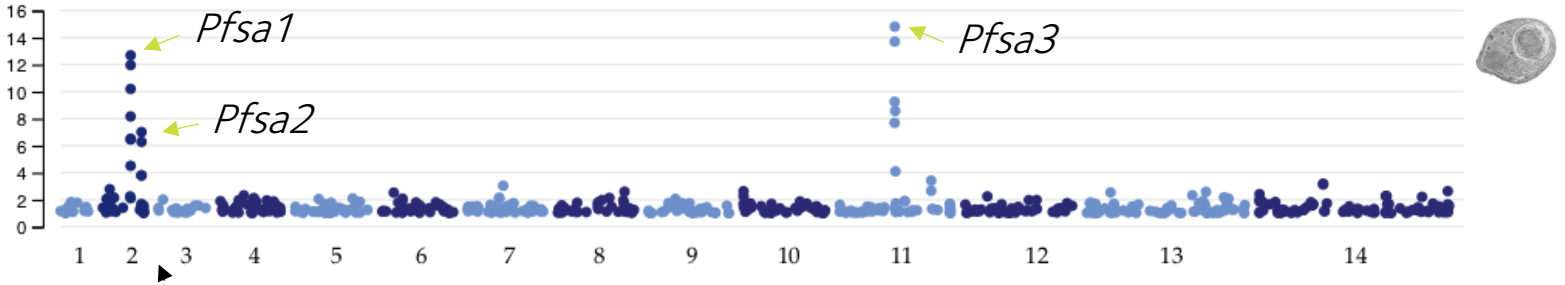


Evidence for association
for human variants
(averaged over *Pf* variants)

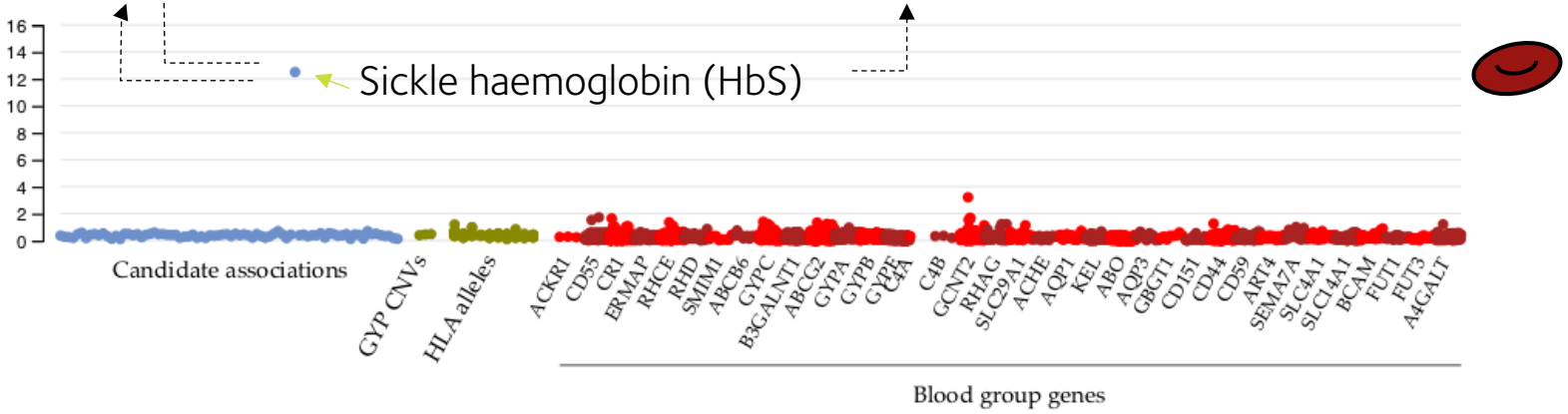


Three regions of the Pf genome are associated with HbS

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



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

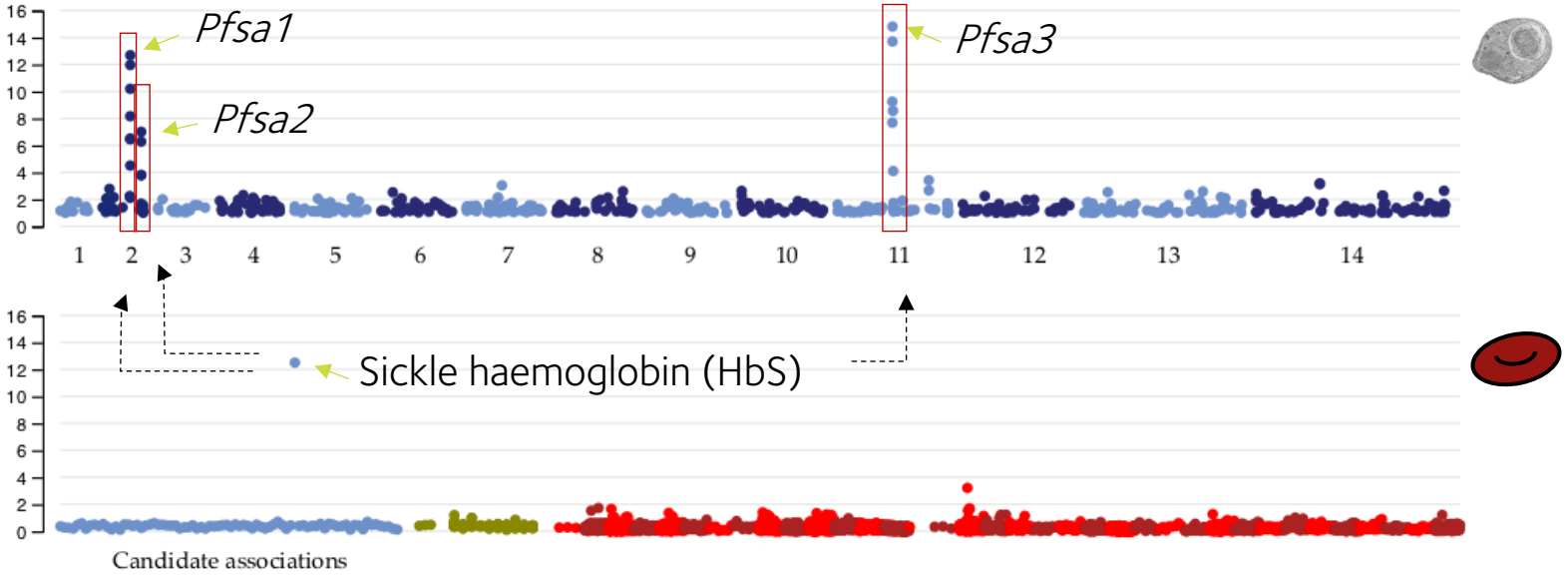


"Plasmodium falciparum sickle-associated"

Three regions of the Pf genome are associated with HbS

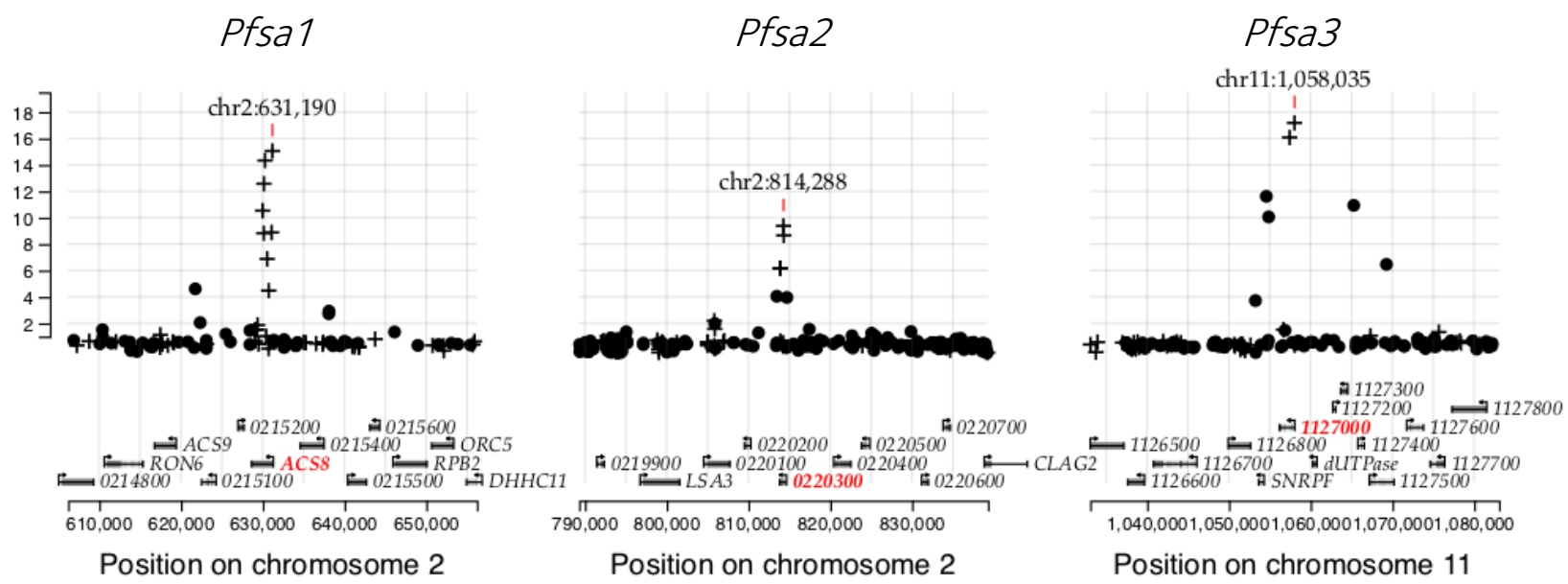
Evidence for association
for *P.falciparum* variants
(averaged over human variants)

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



Zoom in to
Pf genome:

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

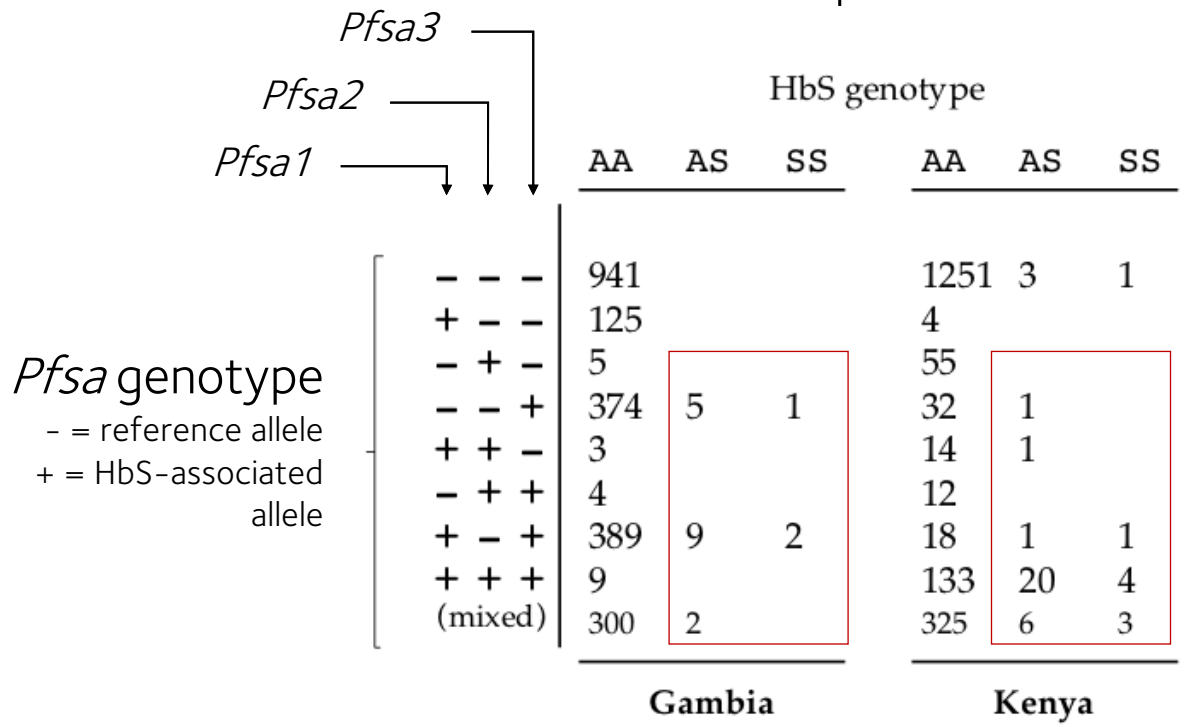


CENTRE

The protective effect of HbS varies with *Pfsa* genotype

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

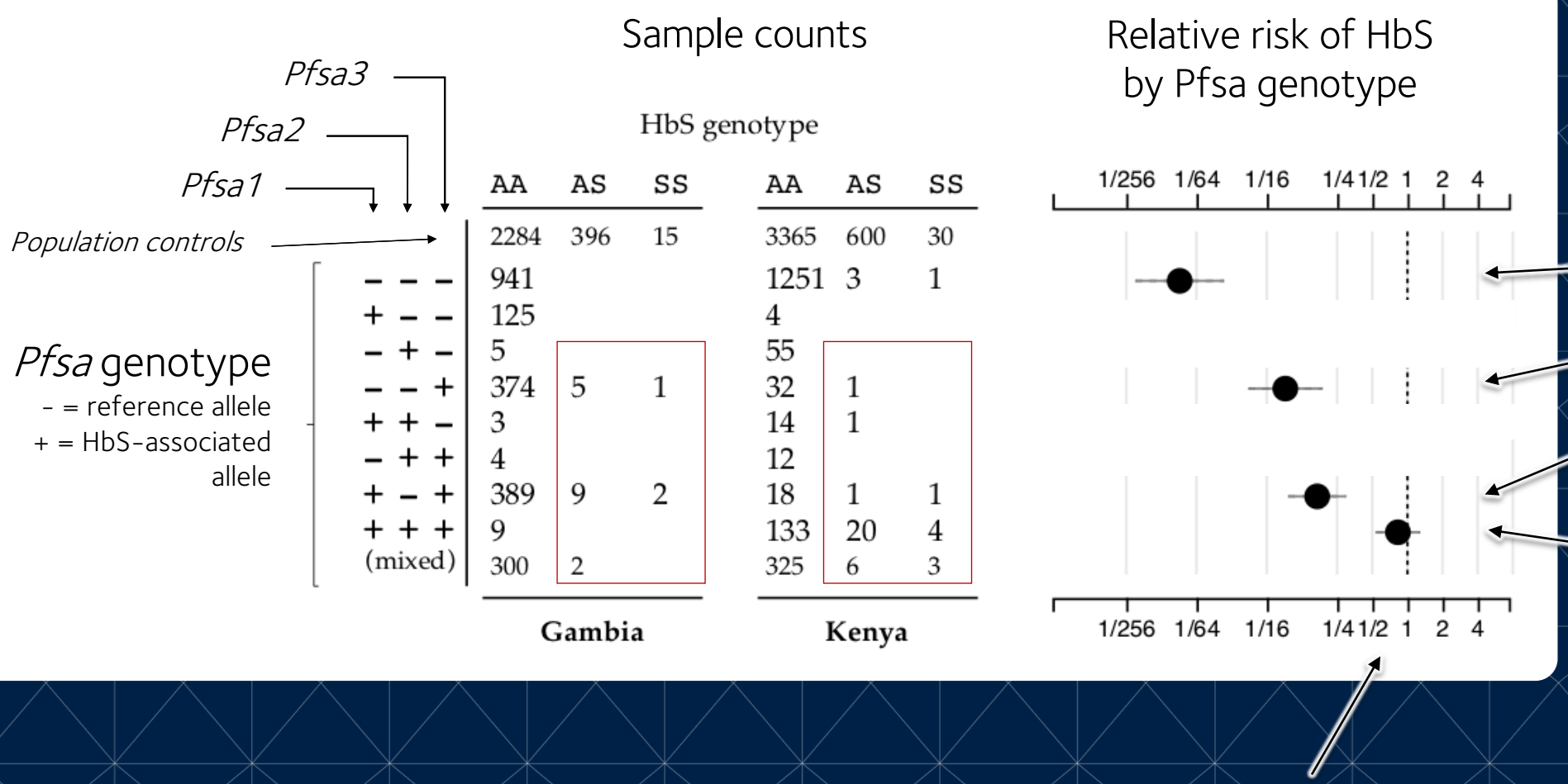
Sample counts



Pfsa genotype	Gambia AA	Gambia AS	Gambia SS	Kenya AA	Kenya AS	Kenya SS
− − −	941			1251	3	1
+ − −	125			4		
− + −	5			55		
− − +	374	5	1	32	1	
+ + −	3			14	1	
− + +	4			12		
+ − +	389	9	2	18	1	1
+ + +	9			133	20	4
(mixed)	300	2		325	6	3

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

The protective effect of HbS varies with *Pfsa* genotype

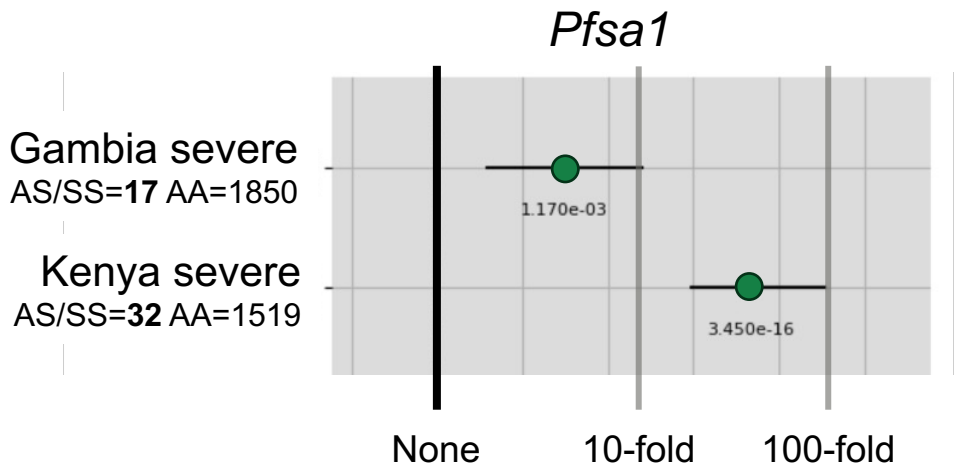


Pfsa+ parasites may have overcome HbS protection

https://doi.org/10.1038/s41586-021-04288-3

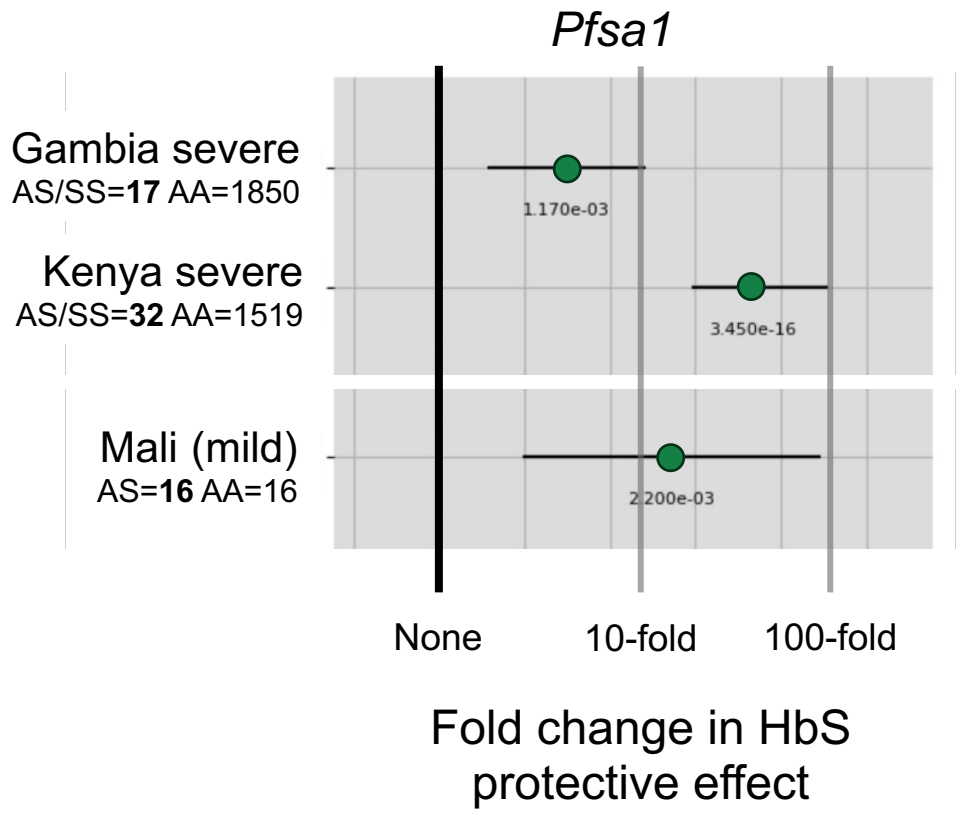
+ caveats!

Replication across contexts



Original study

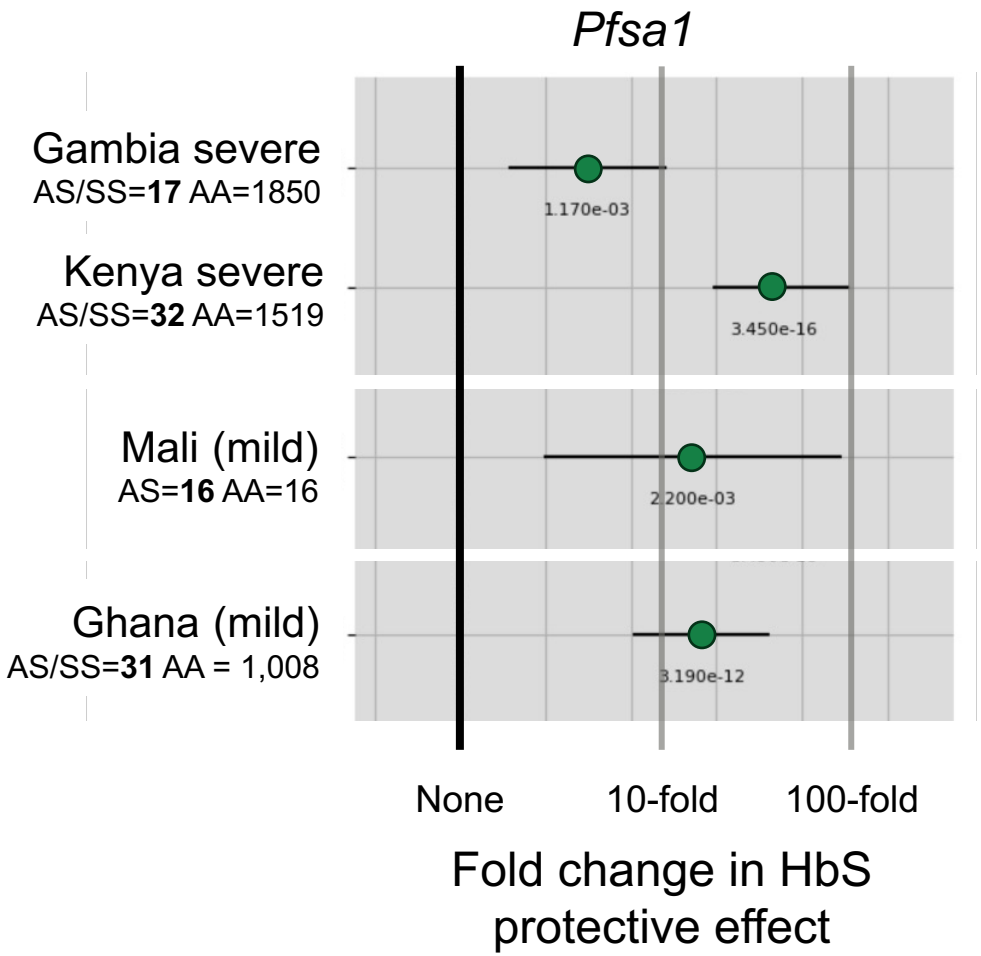
Replication across contexts



Original study

Joseph Saelens, Mahamadou Diakité, Steve M. Taylor mSphere 2021

Replication across contexts

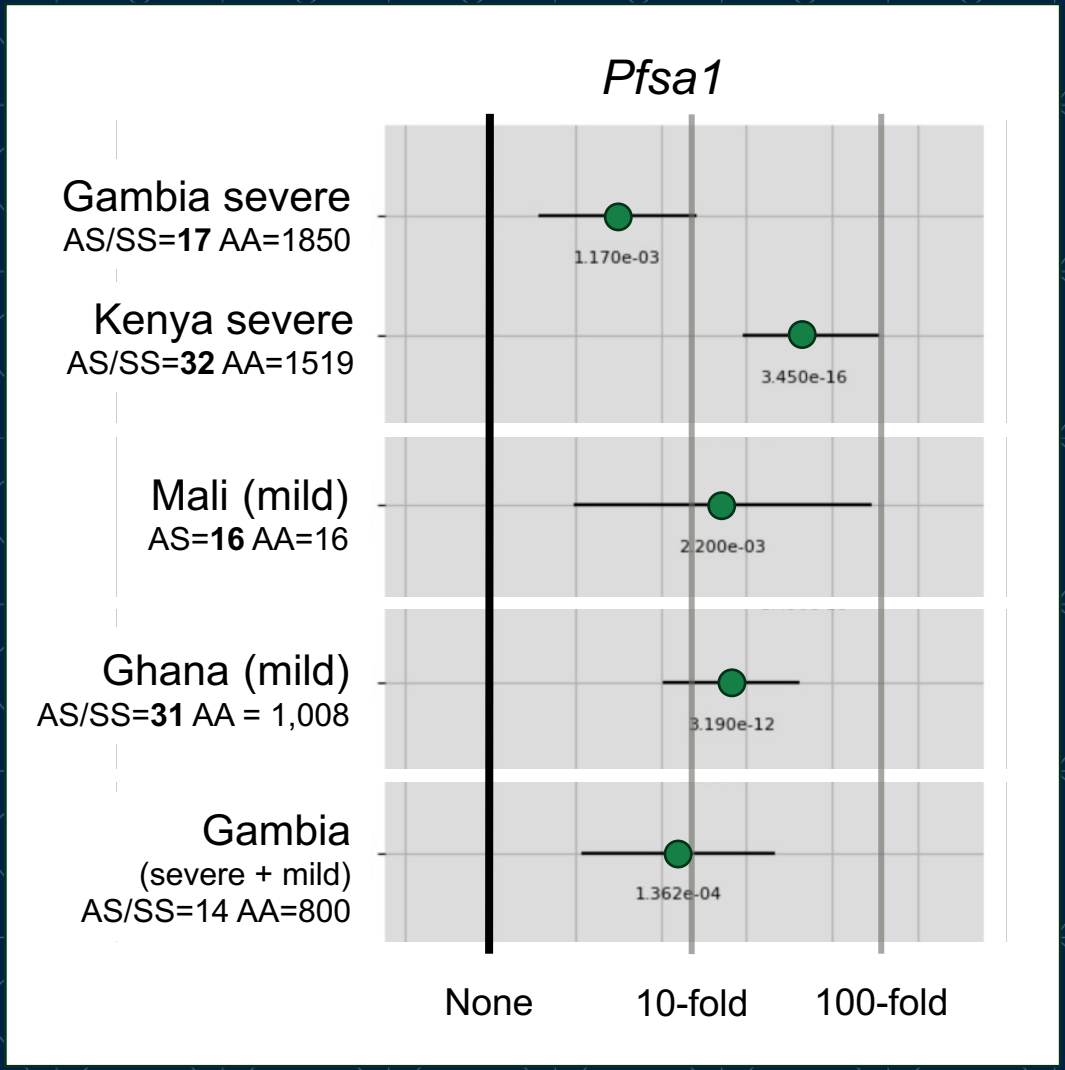


Original study

Joseph Saelens, Mahamadou Diakité, Steve M. Taylor mSphere 2021

Lucas Amenga-Etego, Will Hamilton bioRxiv 2023

Replication across contexts



Original study

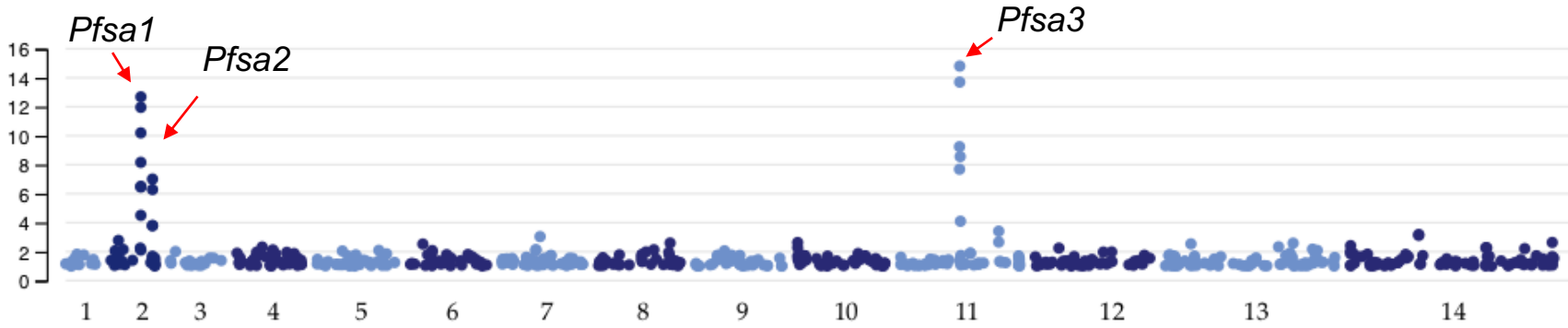
Joseph Saelens, Mahamadou Diakité, Steve M. Taylor mSphere 2021

Lucas Amenga-Etego, Will Hamilton bioRxiv 2023

Annie Forster, Alfred Amambua-Ngwa, unpublished

Three regions of the *Pf* genome are associated with HbS

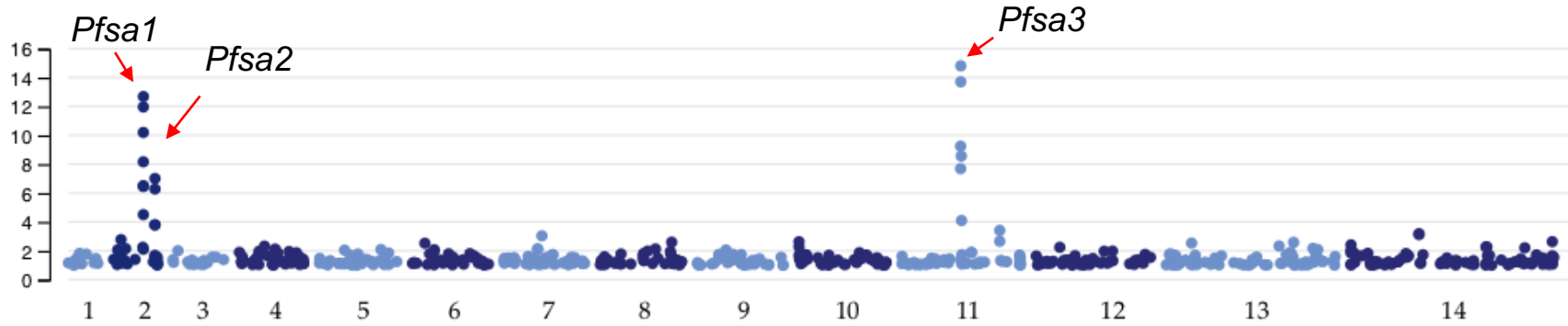
Original study



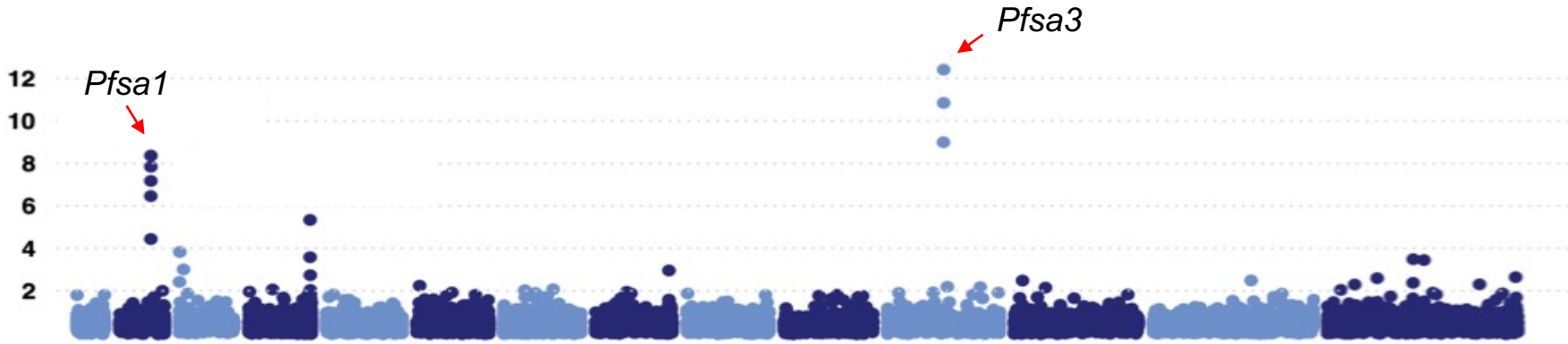
Will Hamilton, … Lucas Amenga-Etego bioRxiv. doi.org/10.1101/2023.09.14.557461

~~Three~~ **Four** regions of the *Pf* genome are associated with HbS

Original study



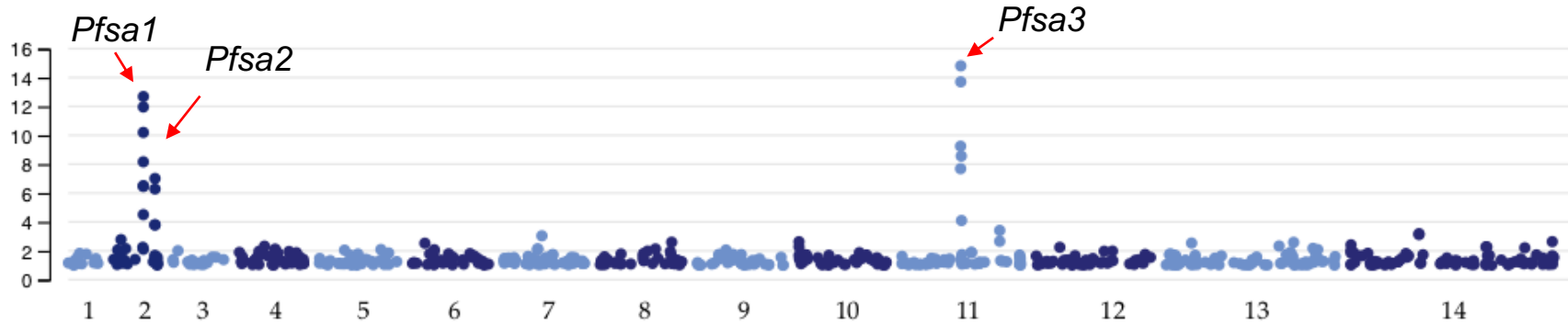
New study in Ghana



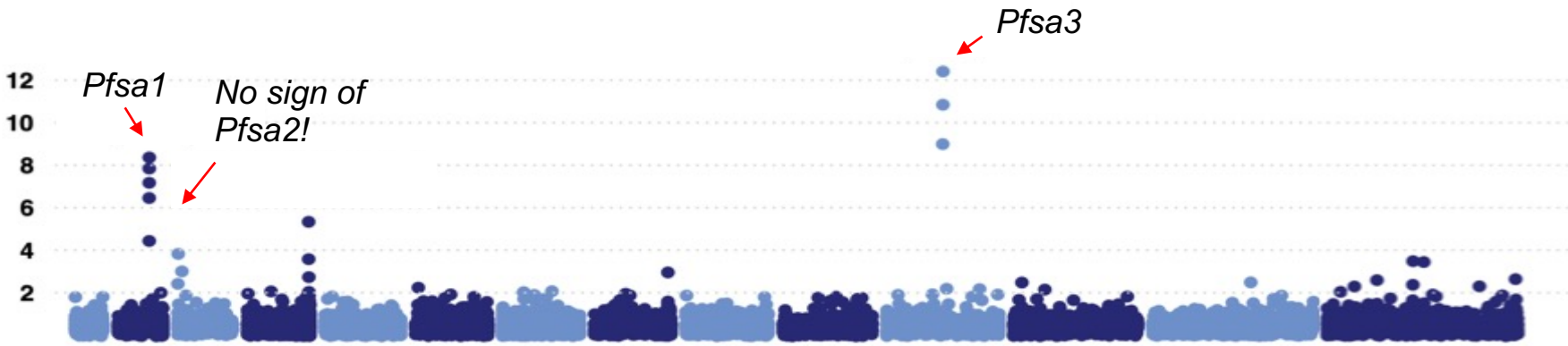
Will Hamilton, … Lucas Amenga-Etego bioRxiv. doi.org/10.1101/2023.09.14.557461

~~Three~~ Four regions of the *Pf* genome are associated with HbS

Original study



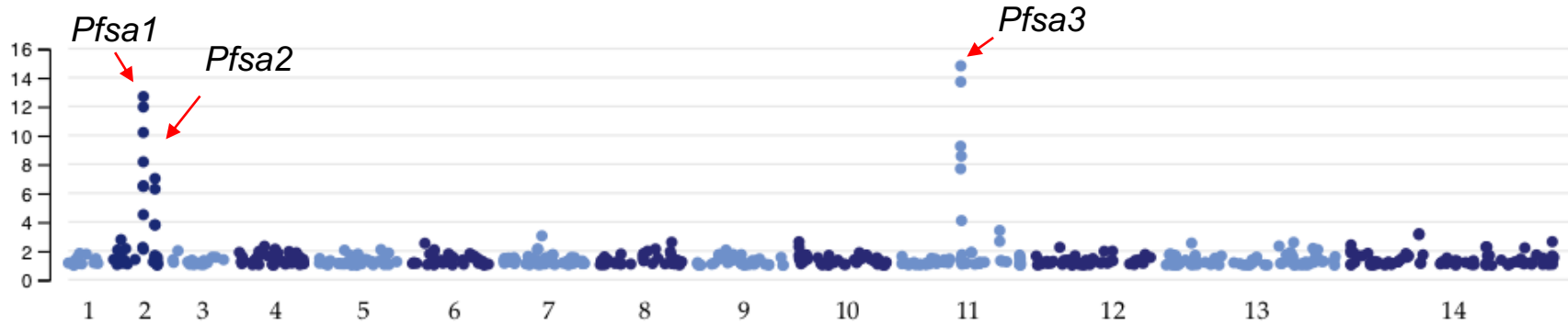
New study in Ghana



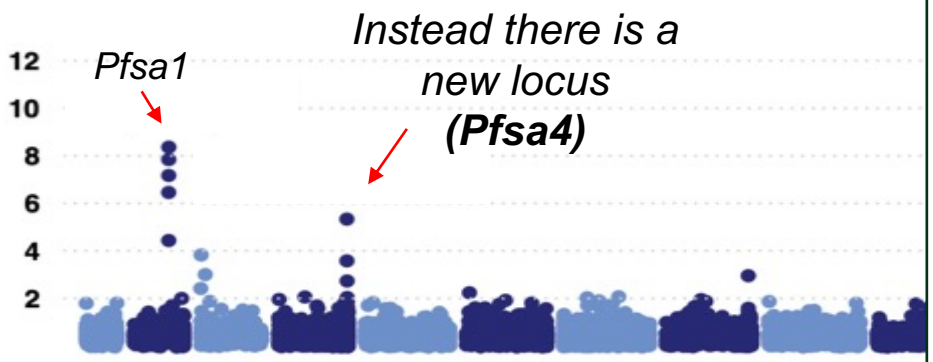
Will Hamilton, … Lucas Amenga-Etego bioRxiv. doi.org/10.1101/2023.09.14.557461

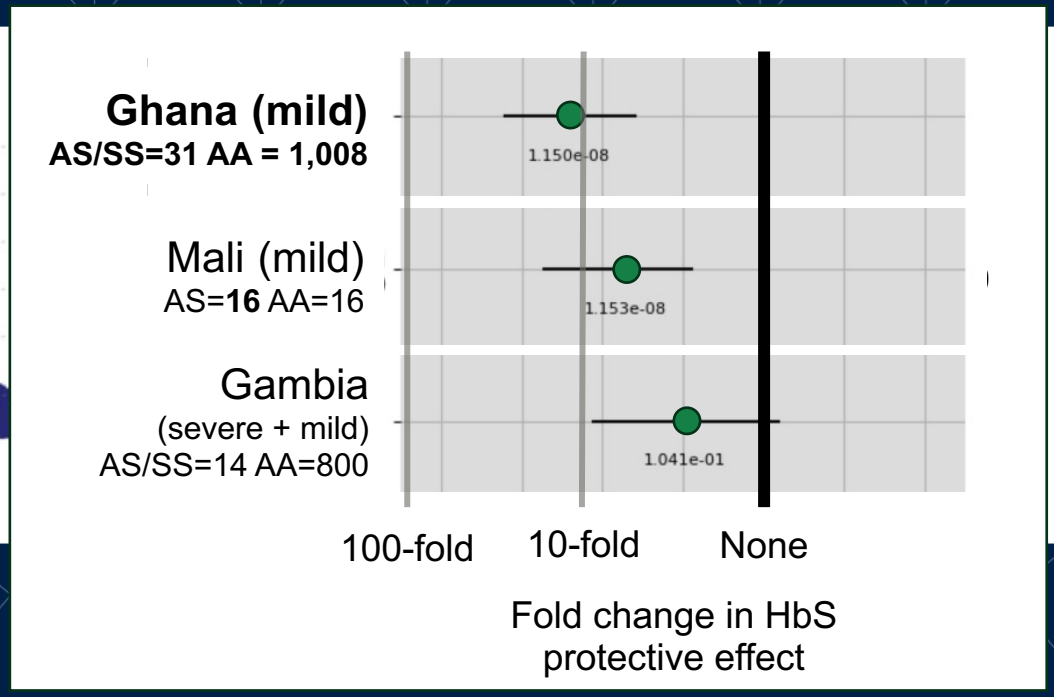
~~Three~~ Four regions of the *Pf* genome are associated with HbS

Original study

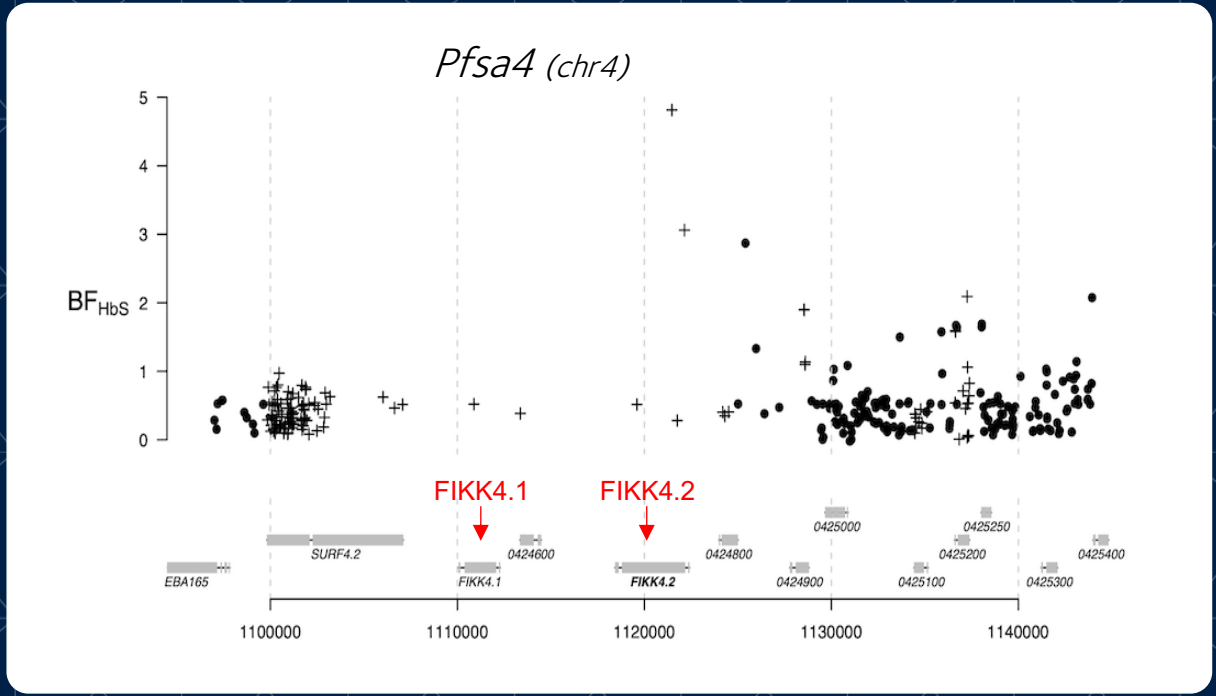


New study in Ghana





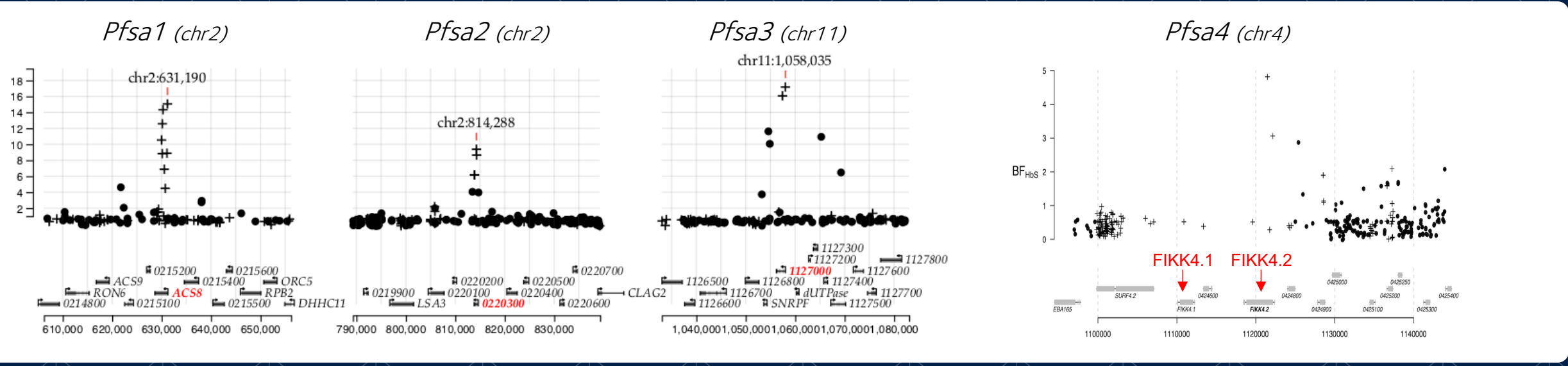
Will Hamilton, … Lucas Amenga-Etego bioRxiv. doi.org/10.1101/2023.09.14.557461



An exported kinase family mediates species-specific erythrocyte remodelling and virulence in human malaria

Heledd Davies[1,8], Hugo Belda[1,8], Malgorzata Broncel[1], Xingda Ye[1,2], Claudine Bisson[3], Viola Introini[4], Dominique Dorin-Semblat[5,6,7], Jean-Philippe Semblat[5,6,7], Marta Tibúrcio[1], Benoit Gamain[5,6,7], Myrsini Kaforou[2] and Moritz Treeck[1]✉

There is lots of work on FIKKs e.g. Davies et al 2020

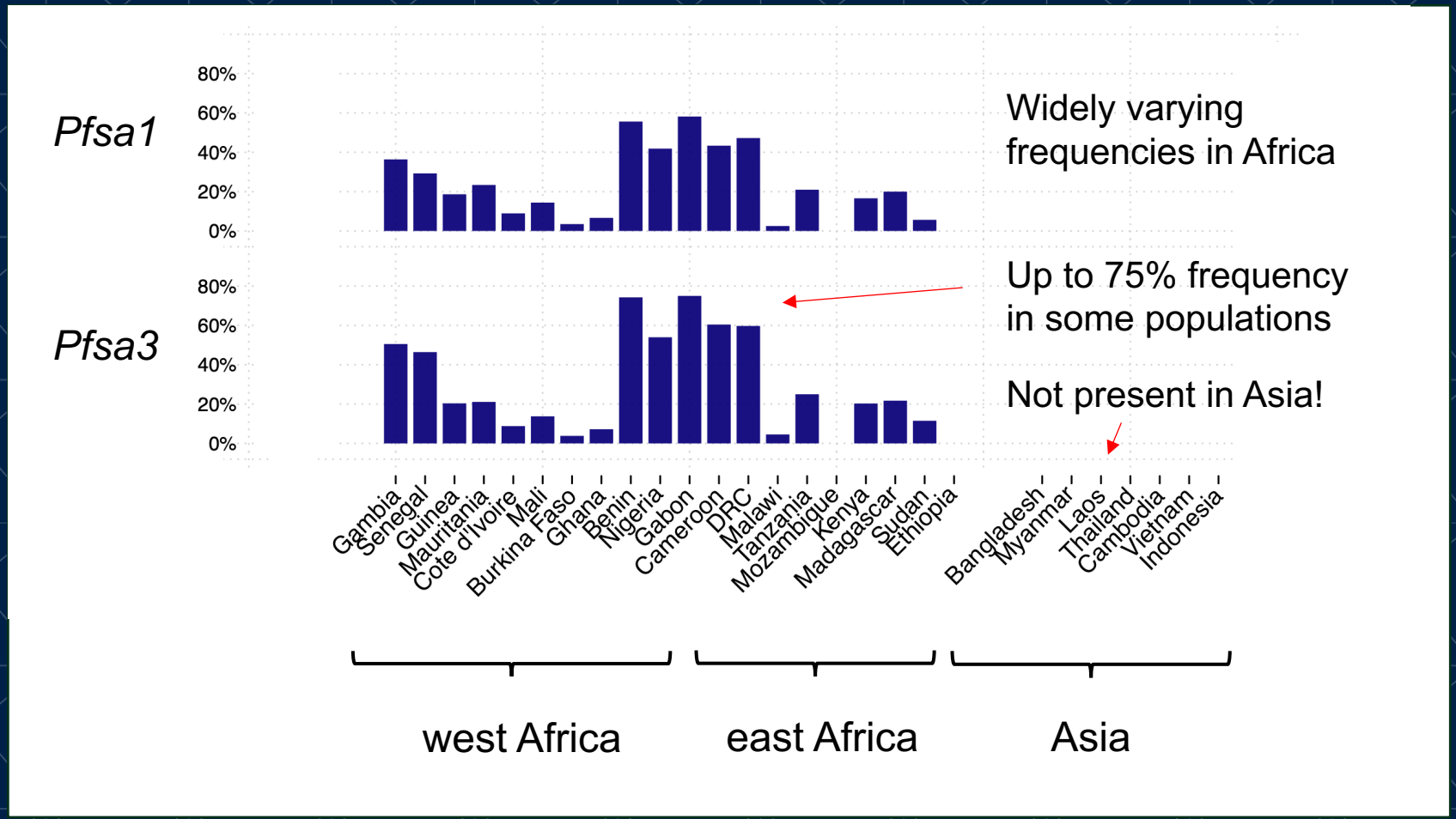


Pfsa

- Four regions of the *Pf* genome are associated with sickle
- Four genes: *PfACS8*, *0020300*, *FIKK4.2*, *1127000*, are implicated
- Almost all infections of HbAS/HbSS individuals carry these mutations
- But there's also complicated variation across populations.

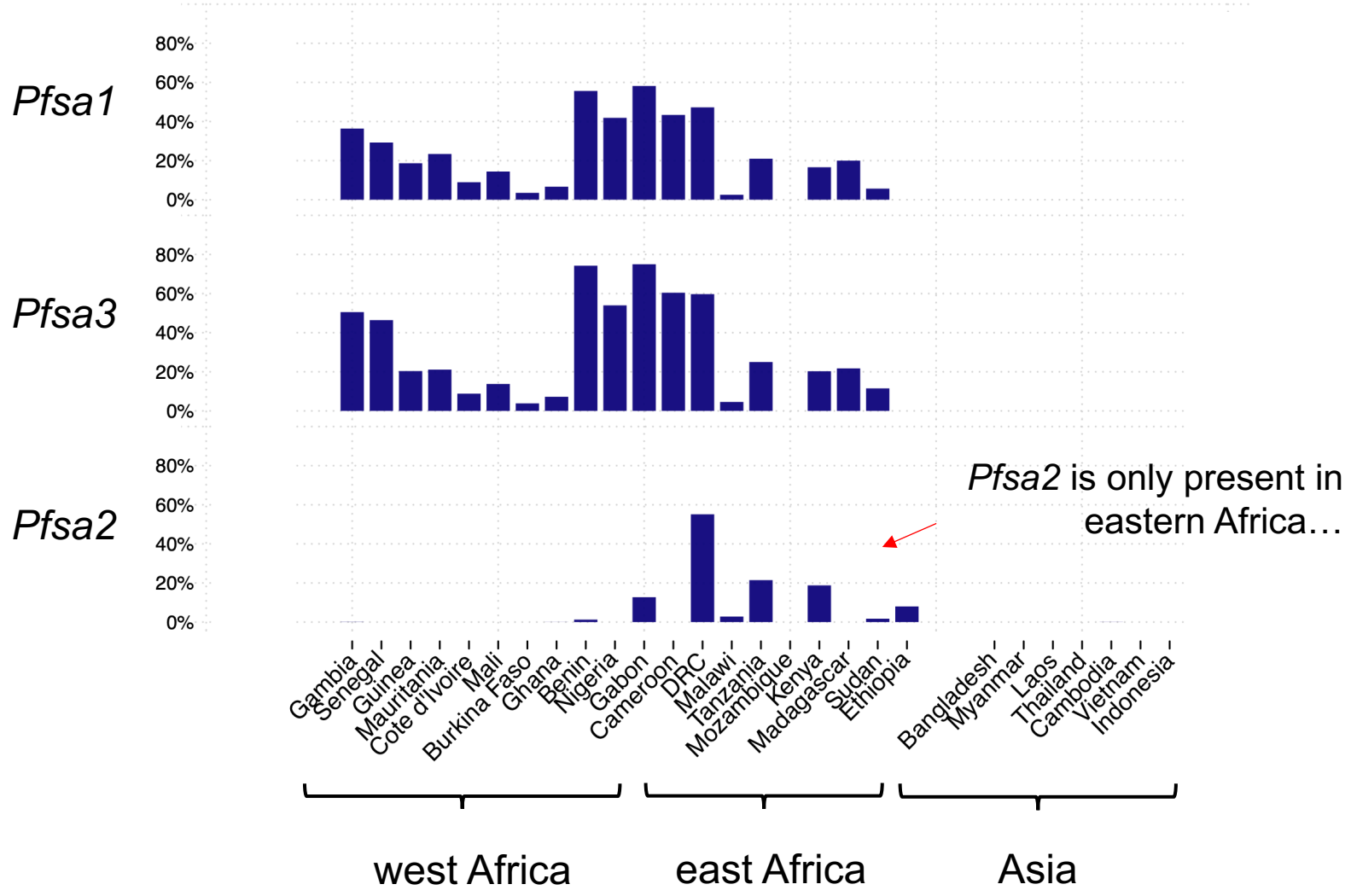
Talk outline

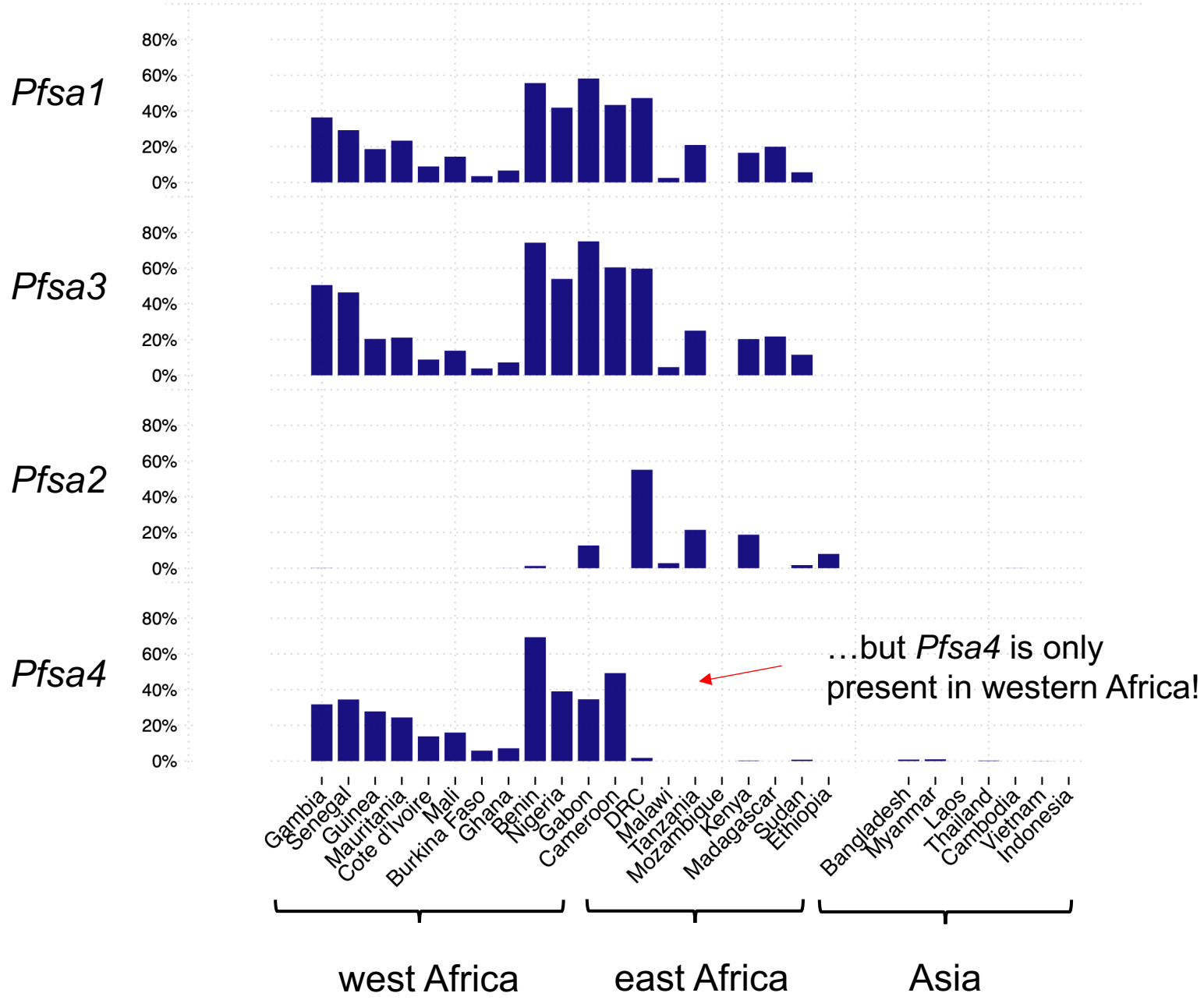
- Malaria protection due to sickle haemoglobin depends on parasite genotype
- *Pfsa* population genetics
- *Pfsa* biology

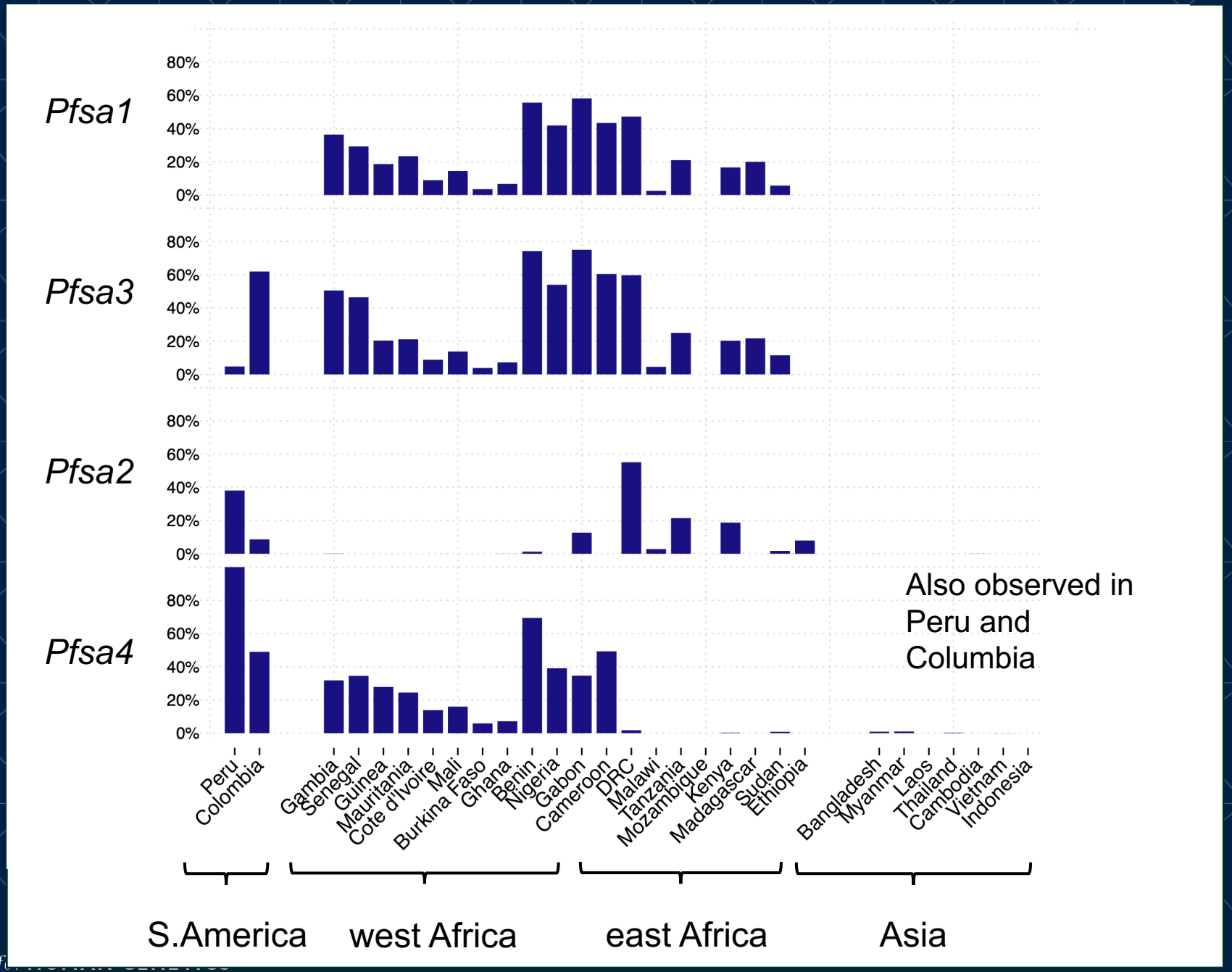


Pfsa global frequencies

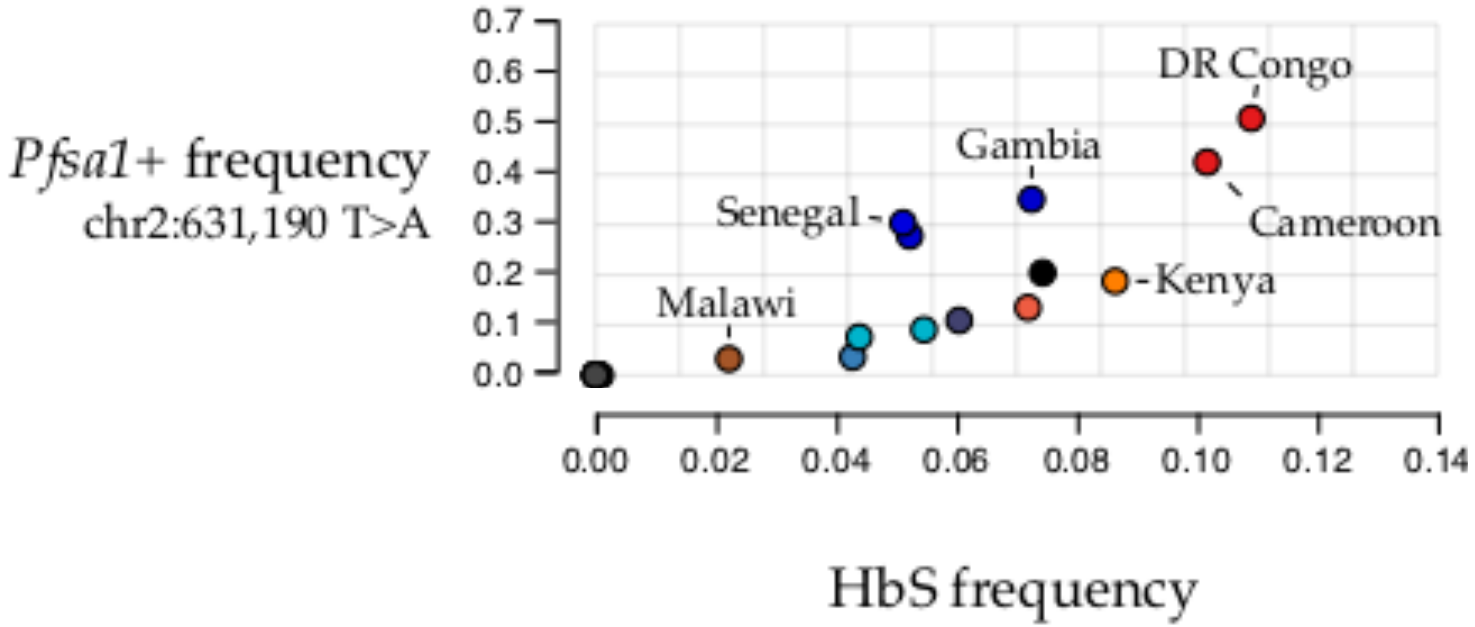
MalariaGEN Pf7
(MalariaGEN Wellcome Open 2023)


Pfsa1
Pfsa3
Pfsa2
80%
60%
40%
20%
0%
Pfsa2 is only present in eastern Africa…
Gambia
Senegal
Guinea
Mauritania
Cote d'Ivoire
Mali
Burkina Faso
Ghana
Benin
Nigeria
Gabon
Cameroon
DRC
Malawi
Tanzania
Mozambique
Kenya
Madagascar
Sudan
Ethiopia
Bangladesh
Myanmar
Laos
Thailand
Cambodia
Vietnam
Indonesia
west Africa
east Africa
Asia

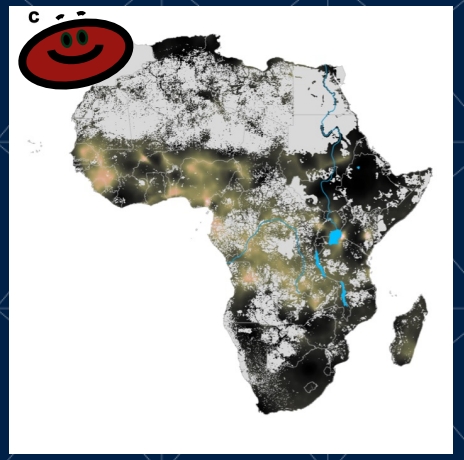

Pfsa1
Pfsa3
Pfsa2
Pfsa4
80%
60%
40%
20%
0%
…but Pfsa4 is only present in western Africa!
Gambia
Senegal
Guinea
Mauritania
Cote d'Ivoire
Mali
Burkina Faso
Ghana
Benin
Nigeria
Gabon
Cameroon
DRC
Malawi
Tanzania
Mozambique
Kenya
Madagascar
Sudan
Ethiopia
Bangladesh
Myanmar
Laos
Thailand
Cambodia
Vietnam
Indonesia
west Africa
east Africa
Asia


Pfsa1
Pfsa3
Pfsa2
Pfsa4
80%
60%
40%
20%
0%
Peru
Colombia
Gambia
Senegal
Guinea
Mauritania
Cote d'Ivoire
Mali
Burkina Faso
Ghana
Benin
Nigeria
Gabon
Cameroon
DRC
Malawi
Tanzania
Mozambique
Kenya
Madagascar
Sudan
Ethiopia
Bangladesh
Myanmar
Laos
Thailand
Cambodia
Vietnam
Indonesia
Also observed in Peru and Columbia
S.America
west Africa
east Africa
Asia

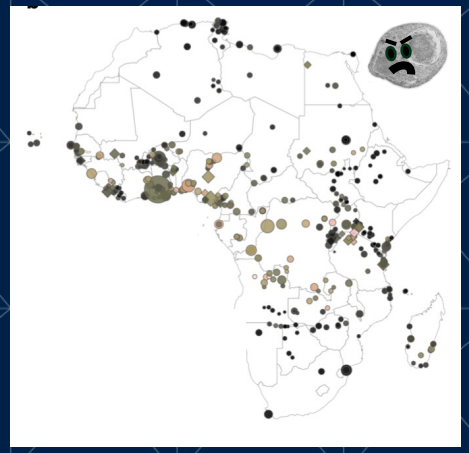
Pfsa+ frequencies are explained by HbS



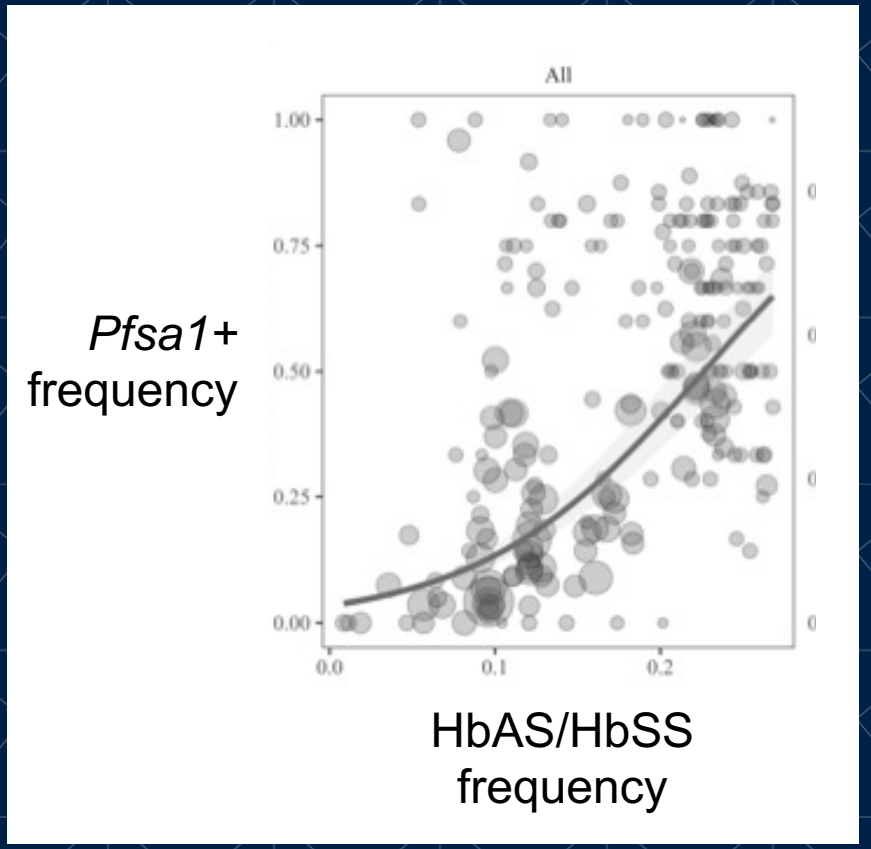
(HbS data from the Malaria Atlas Project – Piel et al Lancet 2013)



+



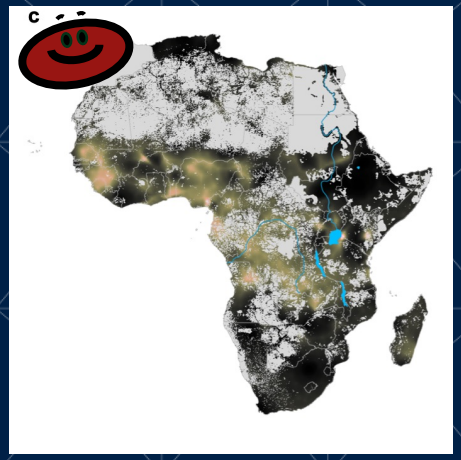
=



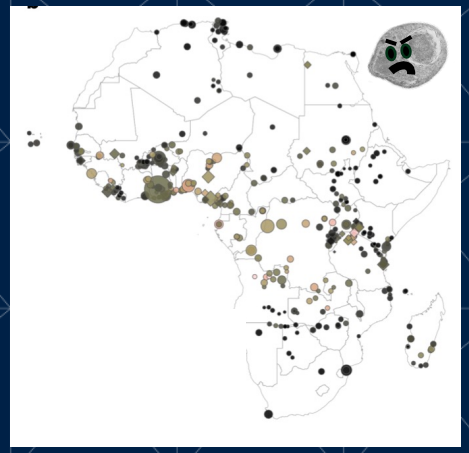
All the HbS allele frequency data
That exists across Africa

All the *Pf* allele frequency data
we could gather across Africa

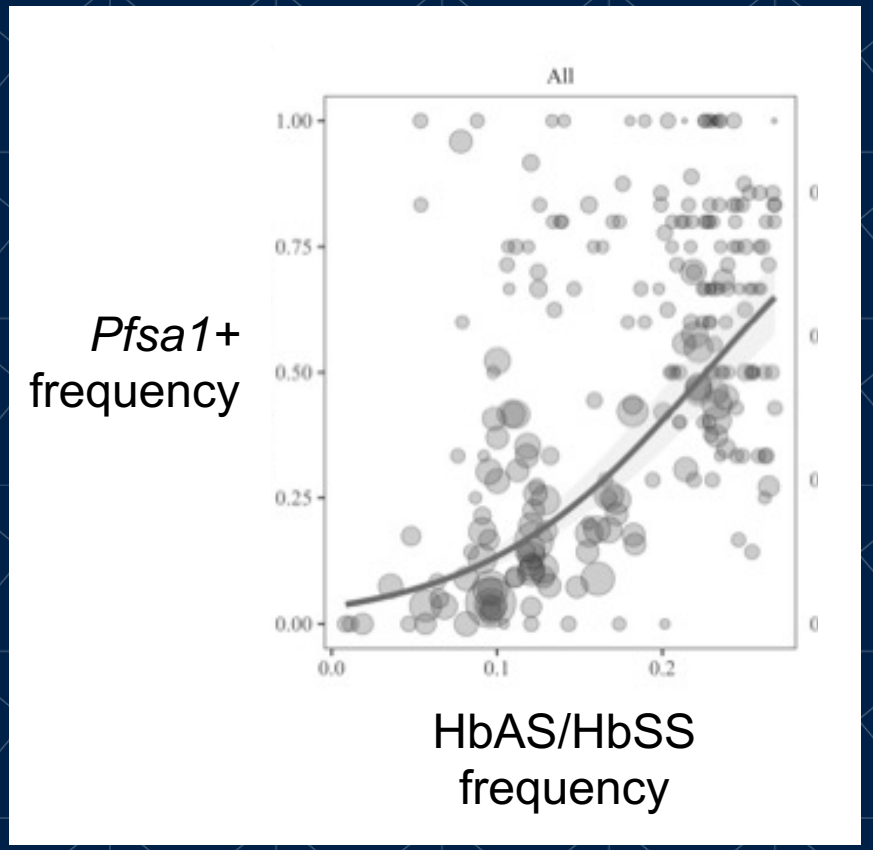
Estimate the relationship at local scales



+



=



All the HbS allele frequency data
That exists across Africa

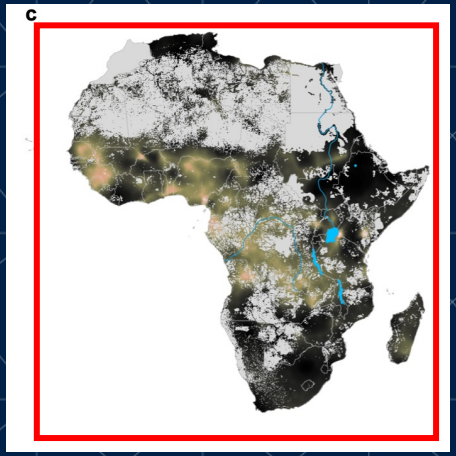
All the *Pf* allele frequency data
we could gather across Africa

Estimate the relationship at local scales

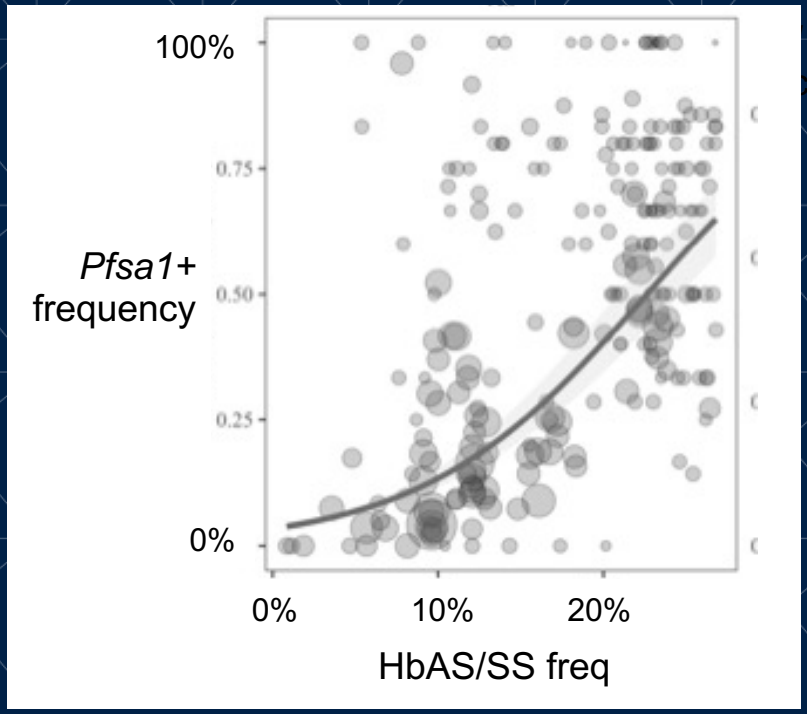
Analysis by Andre Python (Zheijiang University)

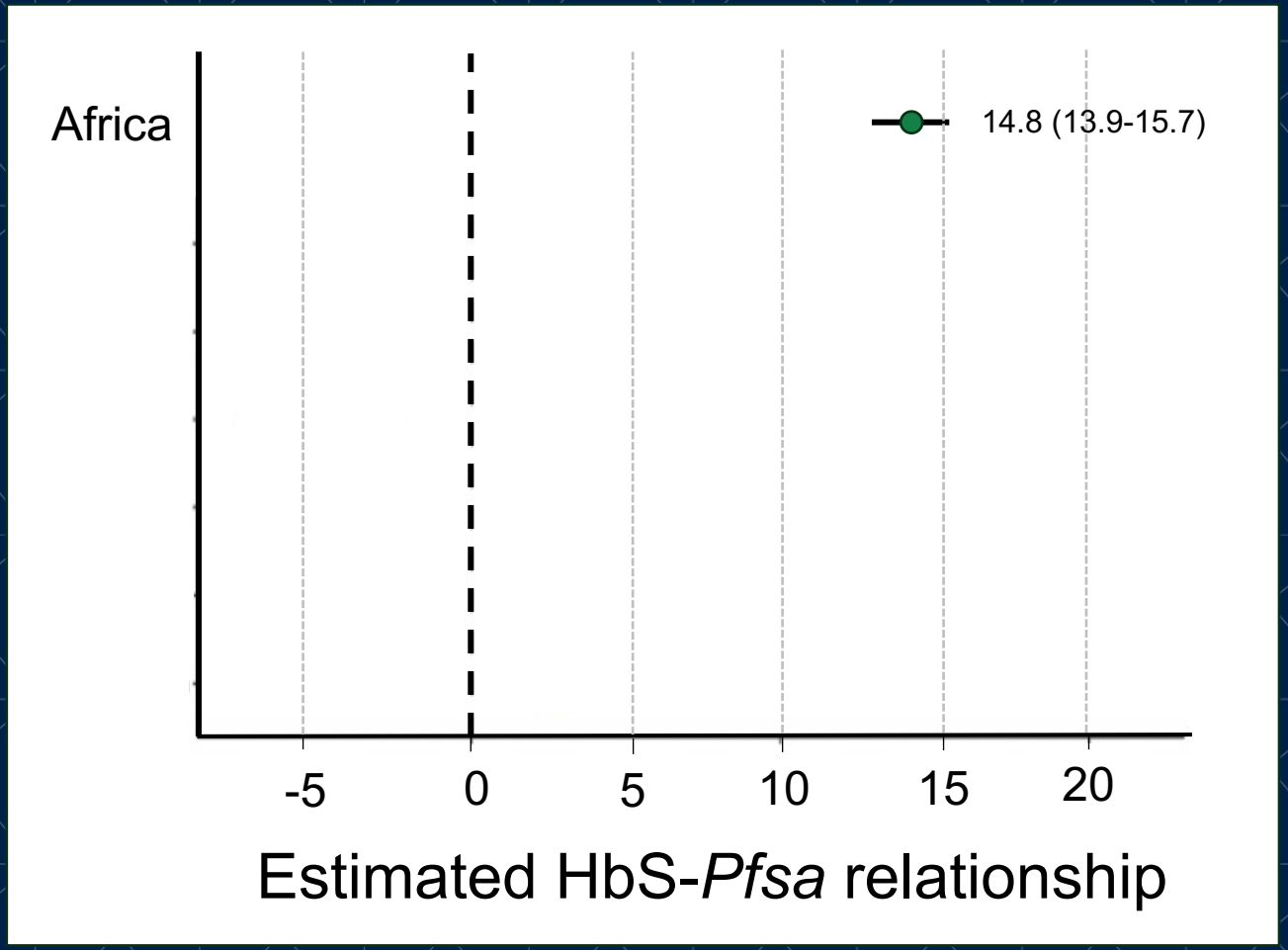
Thanks to: MalariaGEN, Jeffrey Bailey, Robert Verity, Deus Ishengoma for data

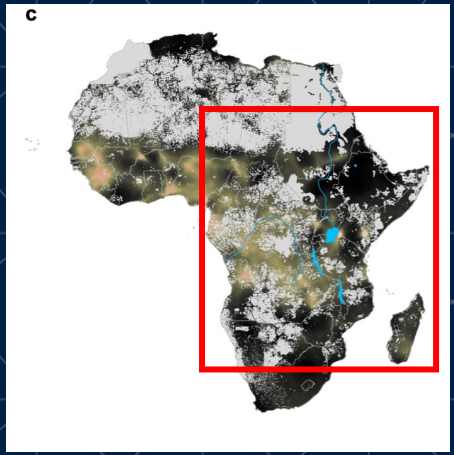
Verity et al 2021 (DRC data) Moser et al 2021 (Tanzania data)


c

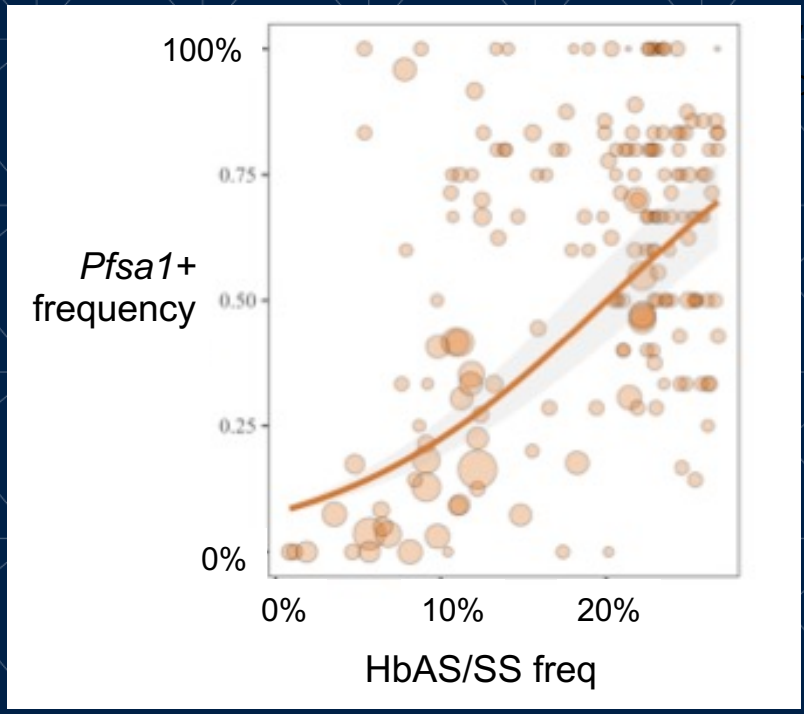
all Africa

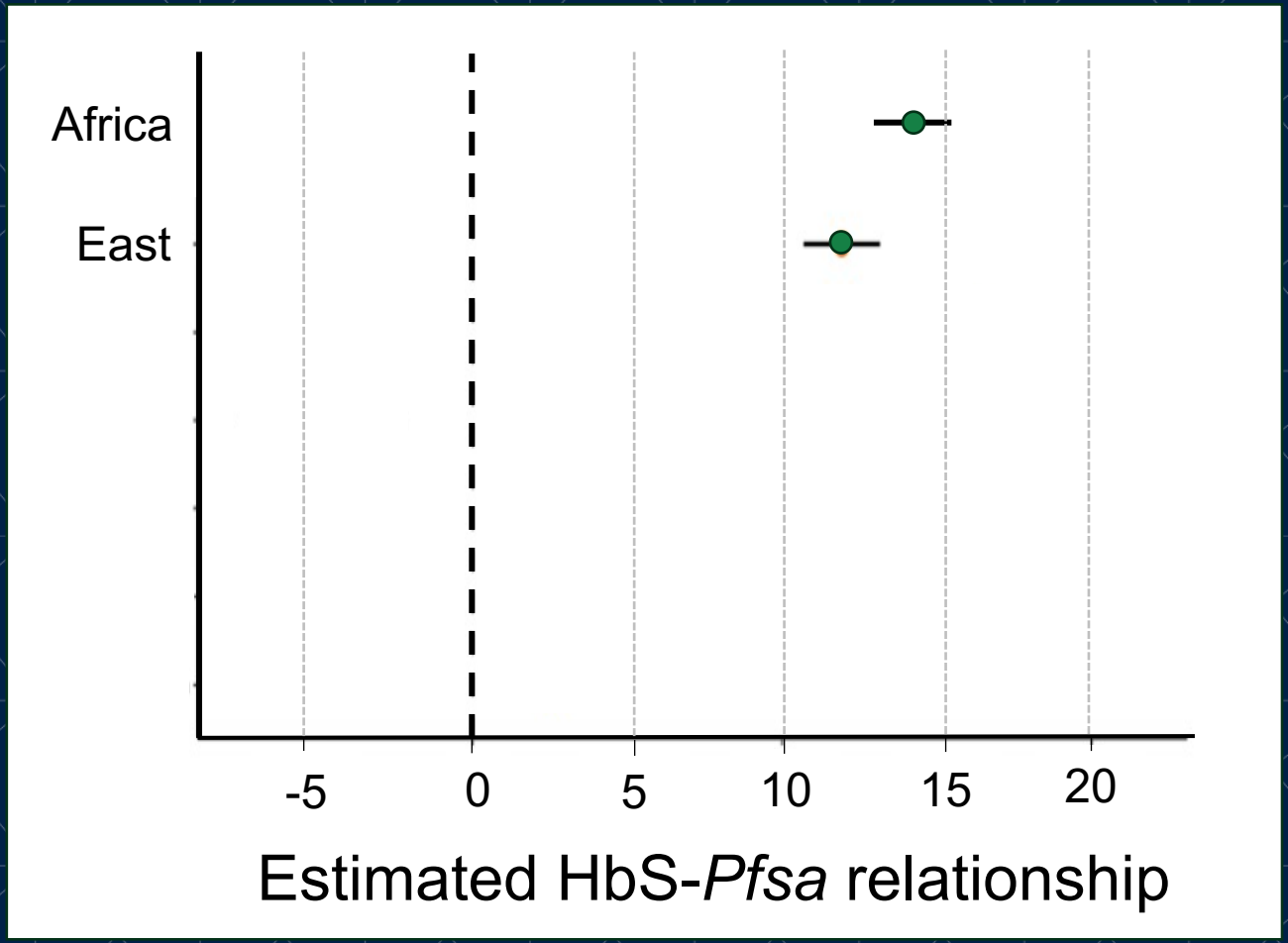

100%
0.75
0.50
0.25
0%
Pfsa1+
frequency
0%
10%
20%
HbAS/SS freq

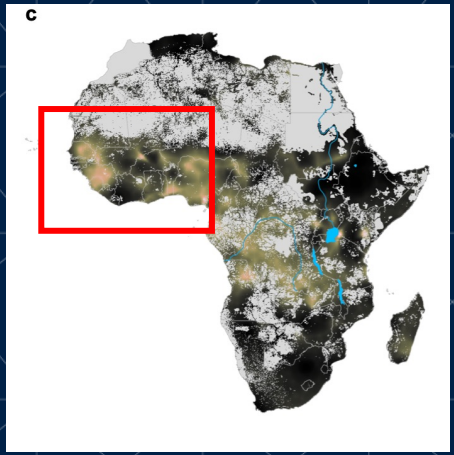

Africa
14.8 (13.9-15.7)
-5
0
5
10
15
20
Estimated HbS-Pfsa relationship



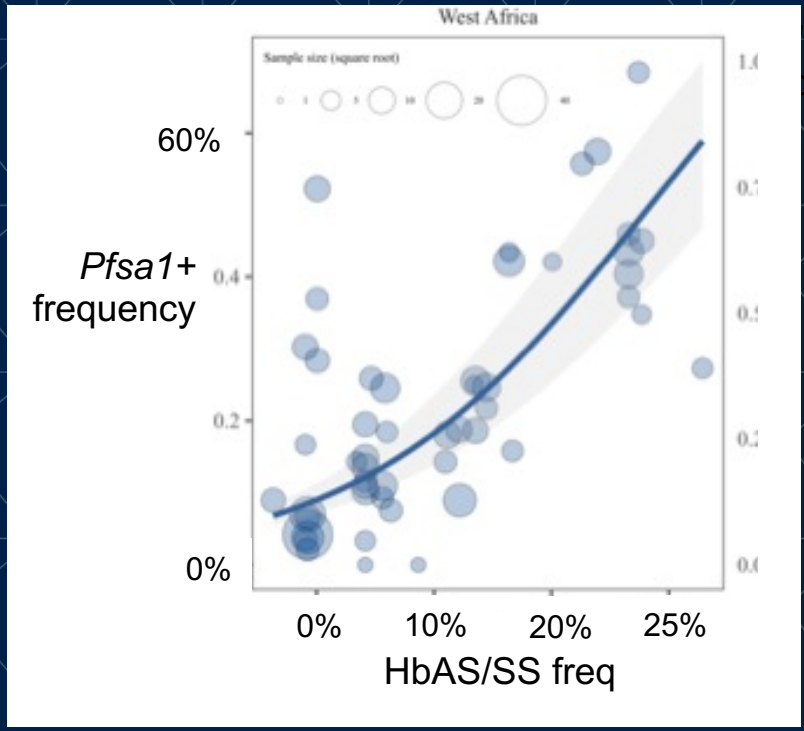
East Africa only
(including DRC)

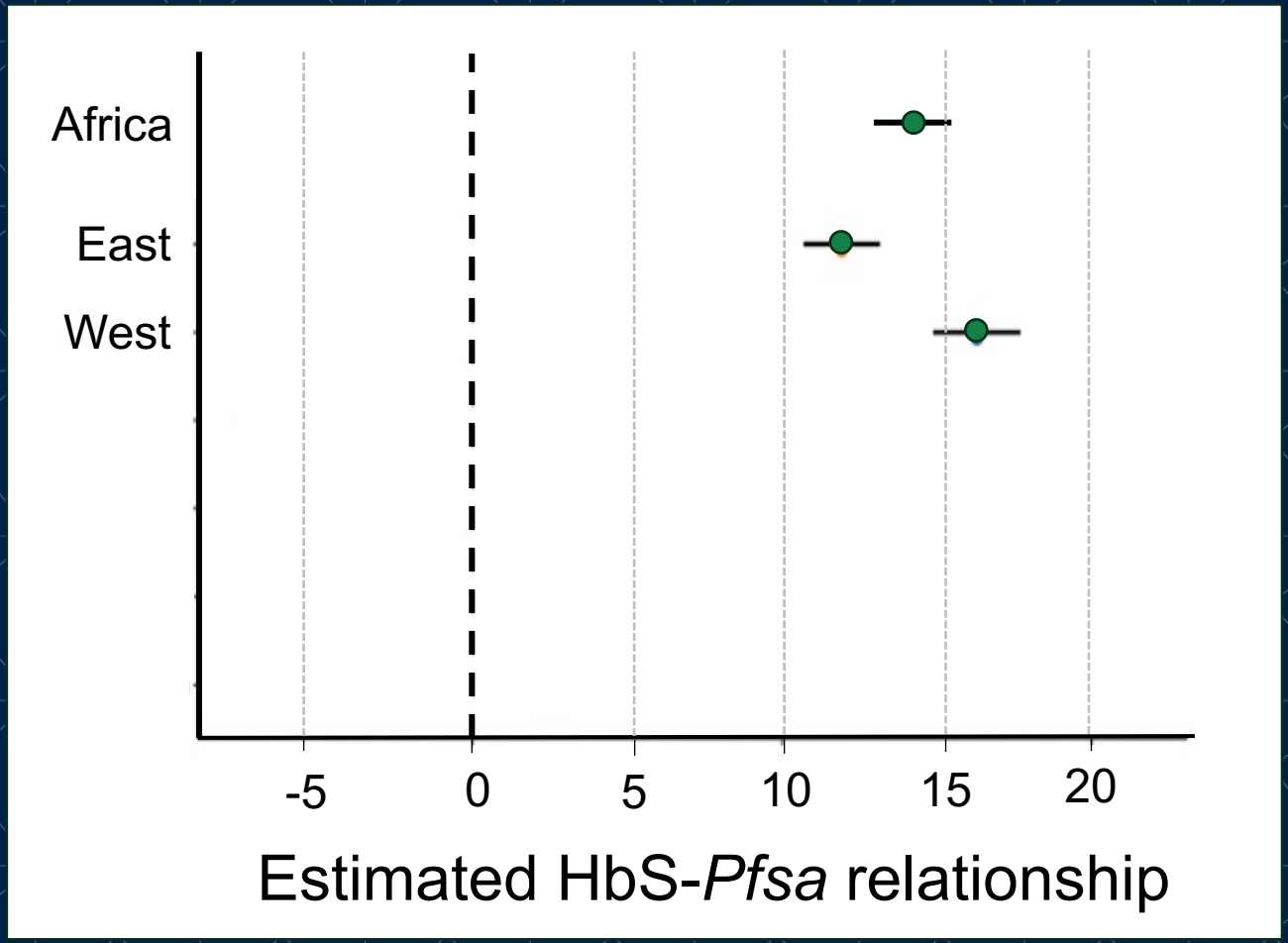


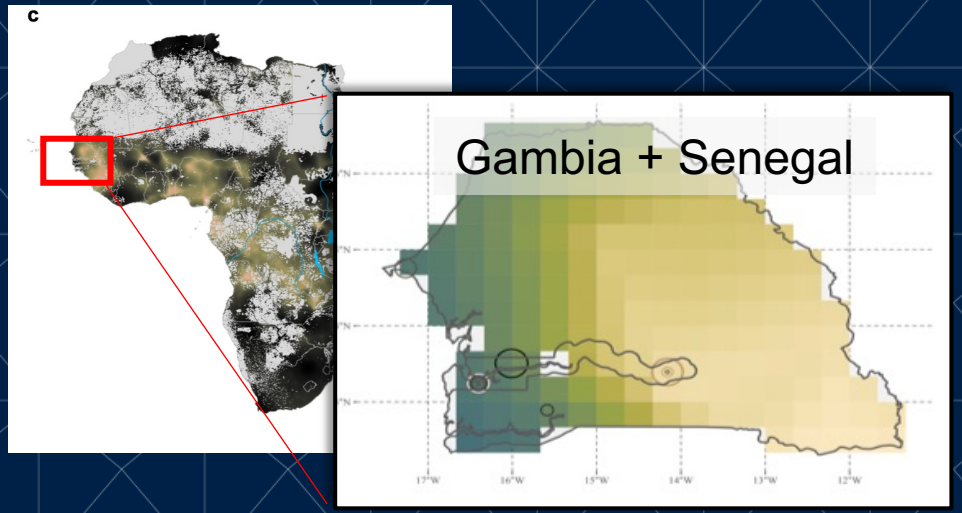


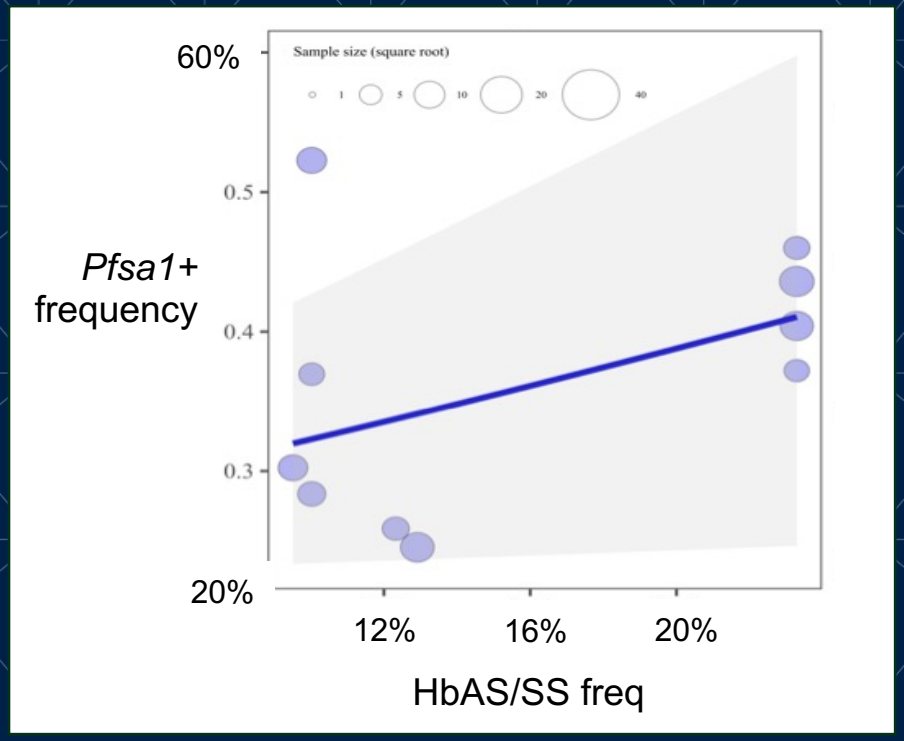


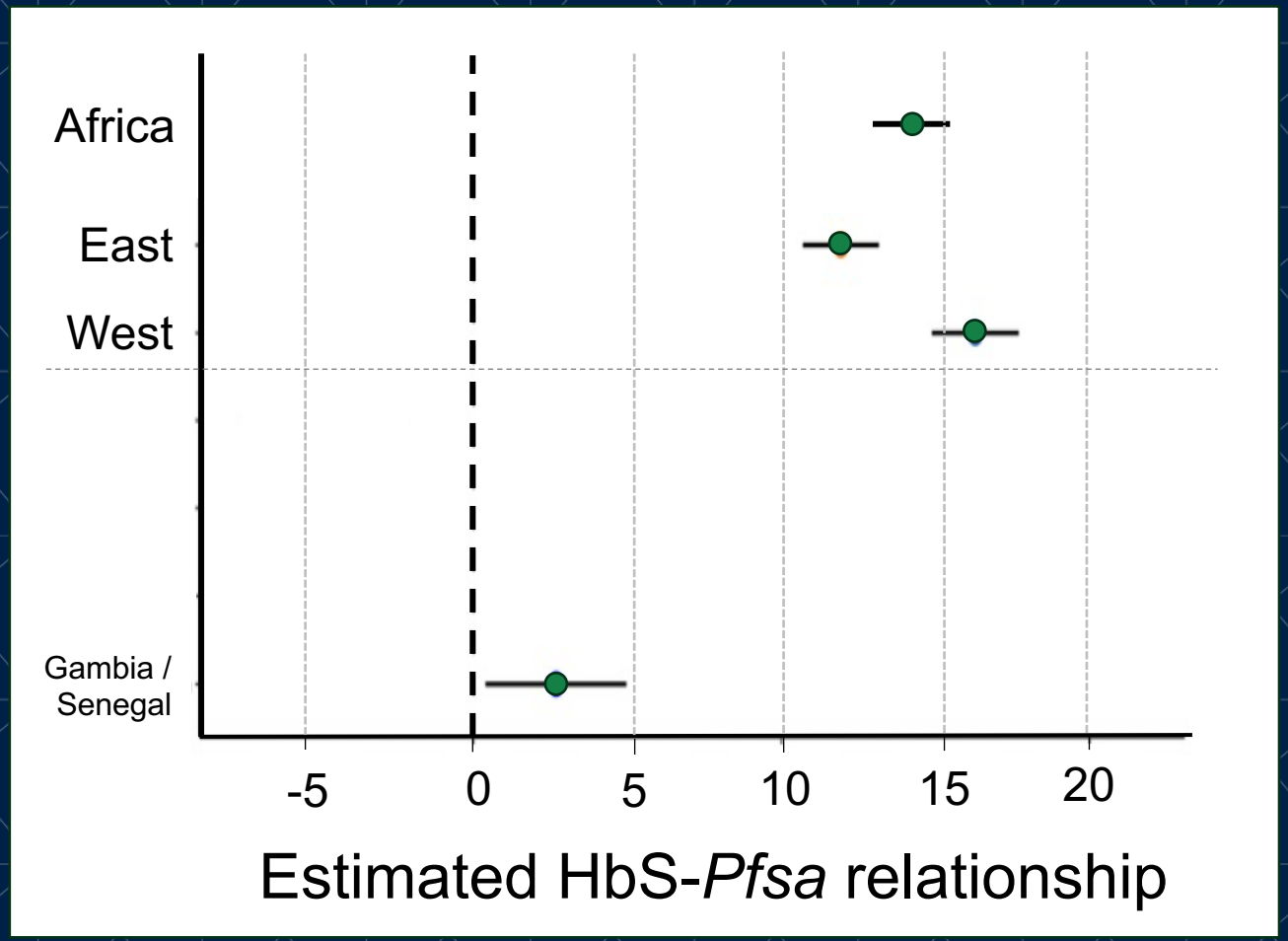
West Africa only

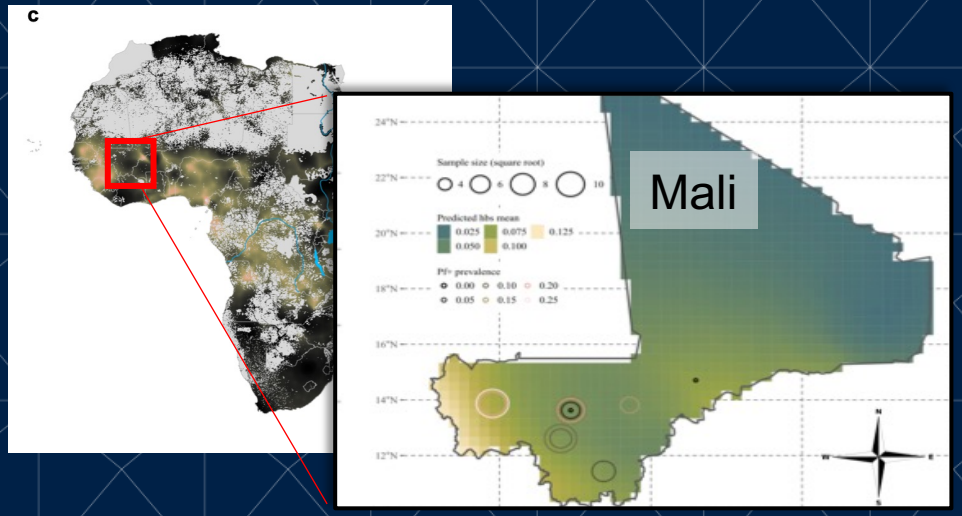


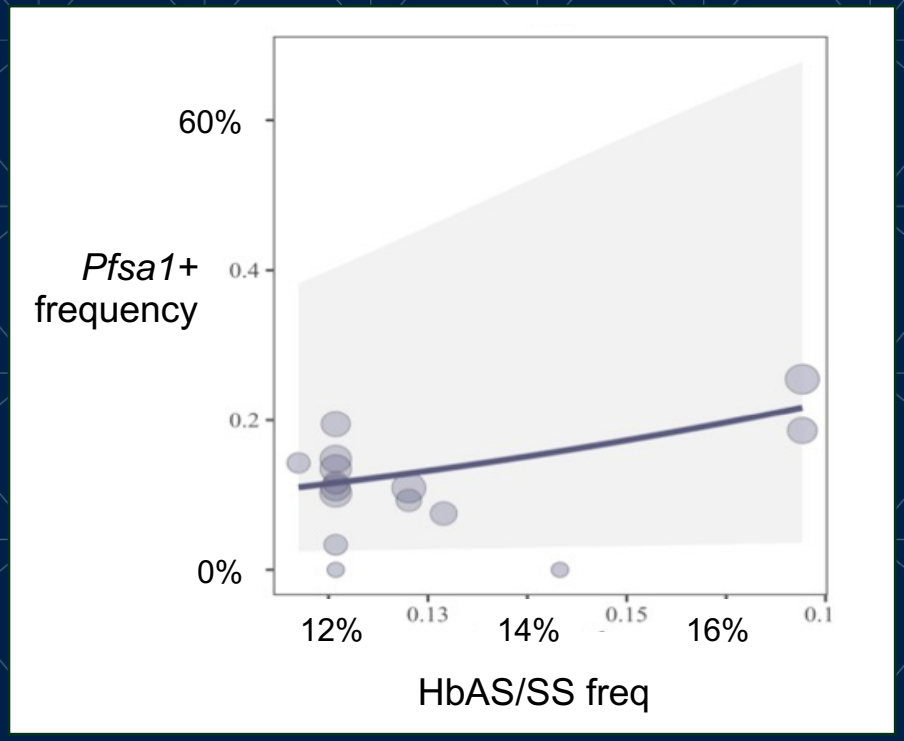


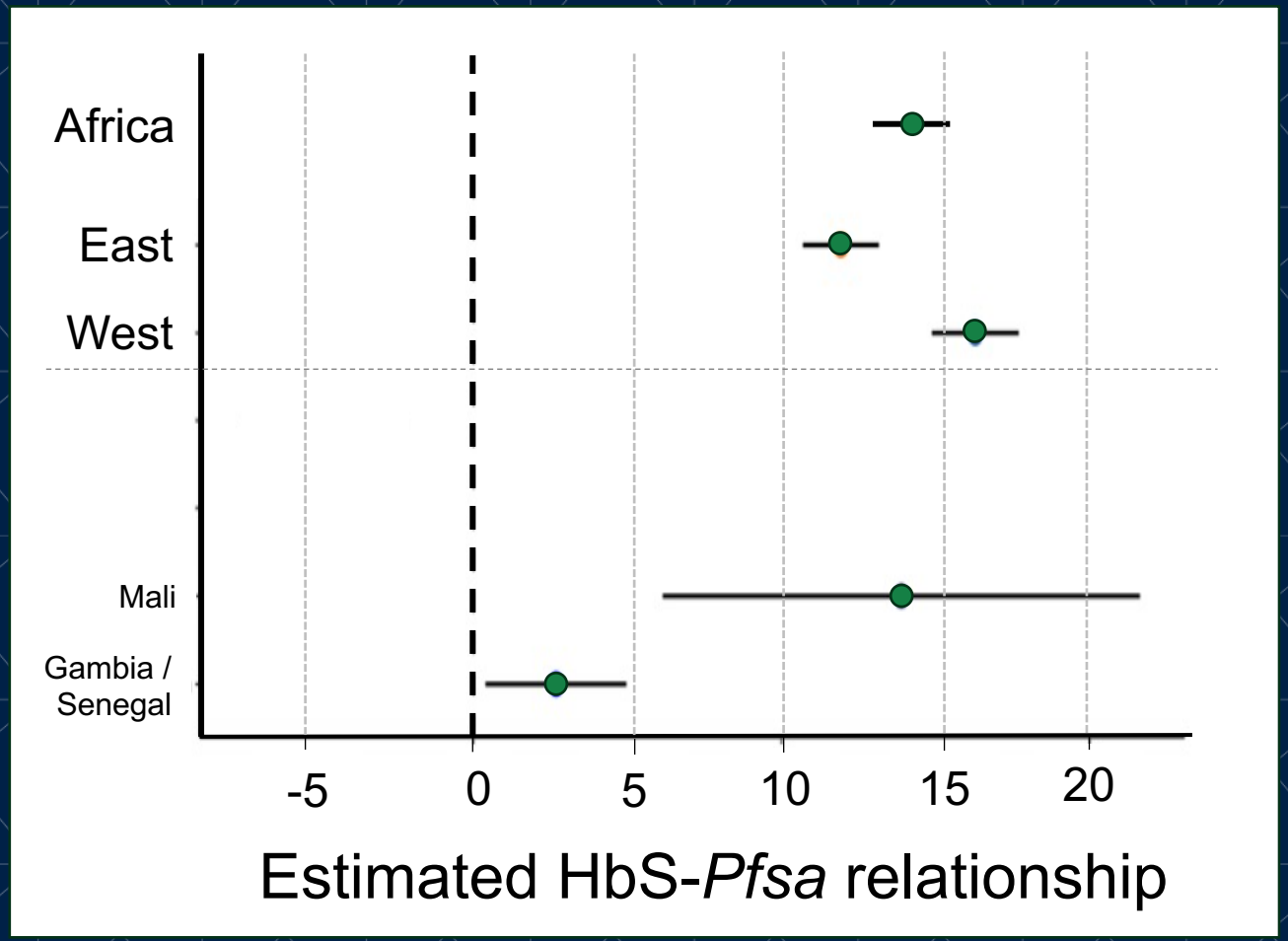

c
Gambia + Senegal

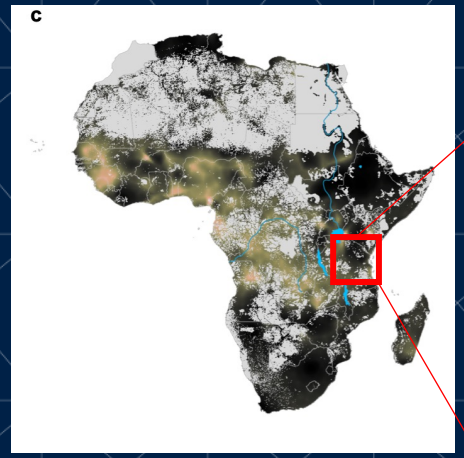

Sample size (square root)
1
5
10
20
40
60%
0.5
0.4
0.3
20%
Pfsa1+
frequency
12%
16%
20%
HbAS/SS freq

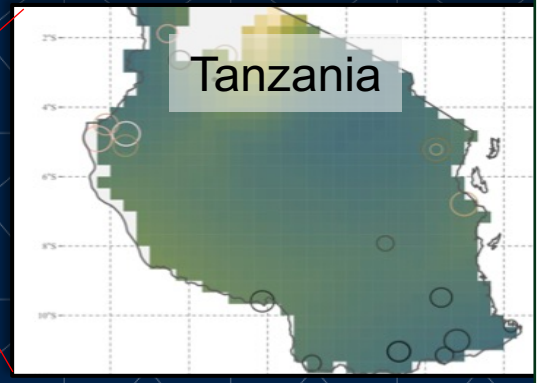

Africa
East
West
Gambia /
Senegal
-5
0
5
10
15
20
Estimated HbS-Pfsa relationship

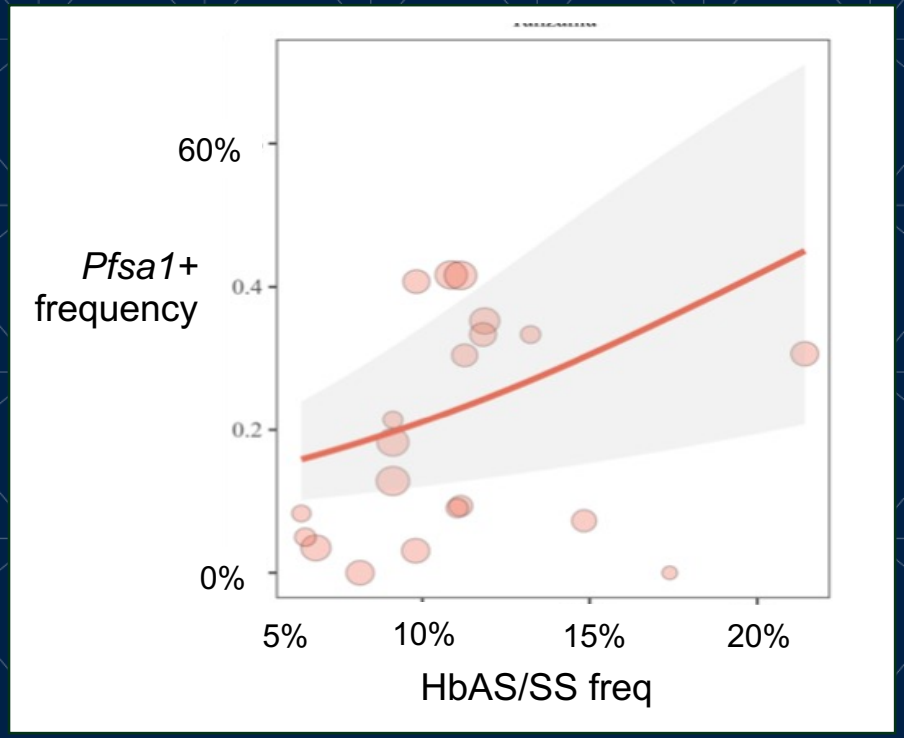

c
Mali

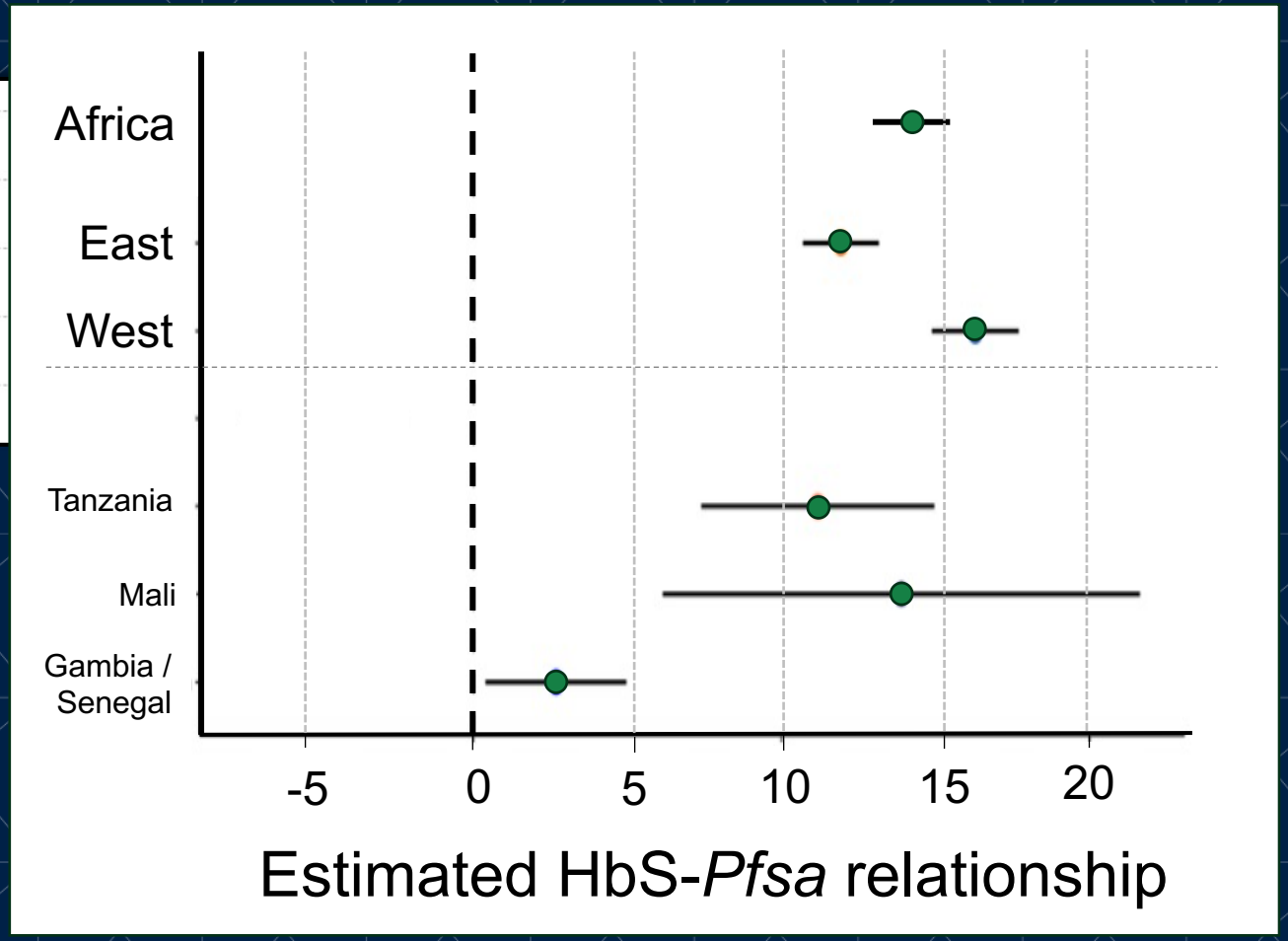

60%
Pfsa1+
frequency
0.4
0.2
0%
12%
0.13
14%
0.15
16%
0.1
HbAS/SS freq

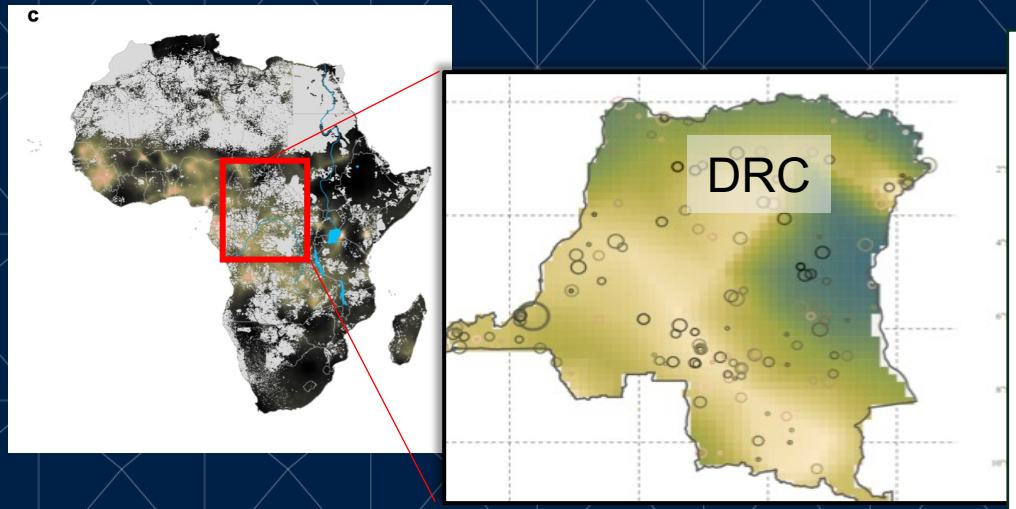

Africa
East
West
Mali
Gambia /
Senegal
-5
0
5
10
15
20
Estimated HbS-Pfsa relationship

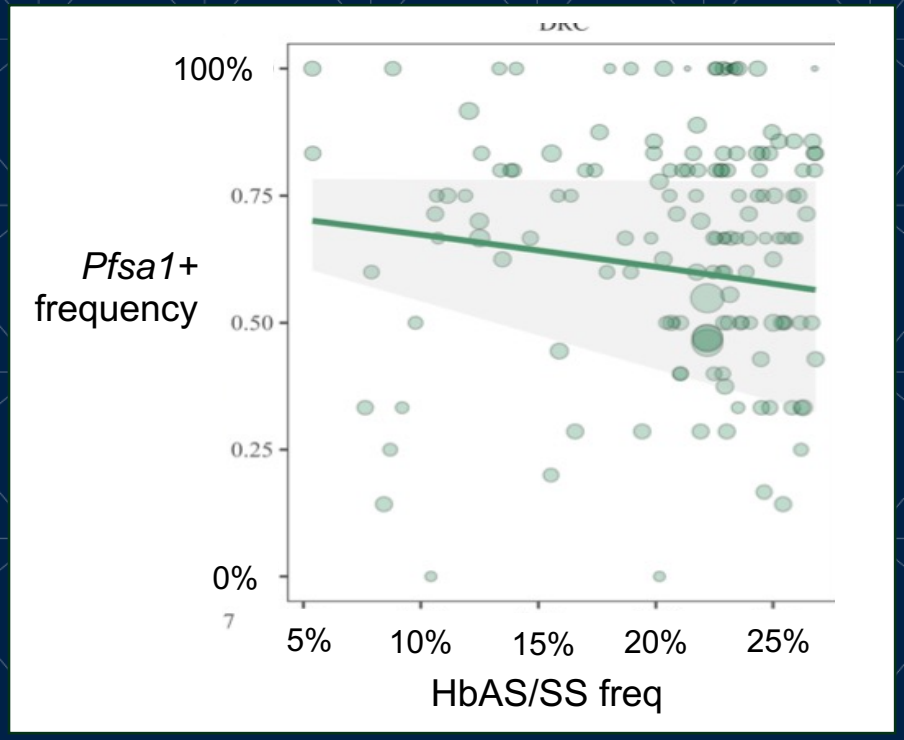

c

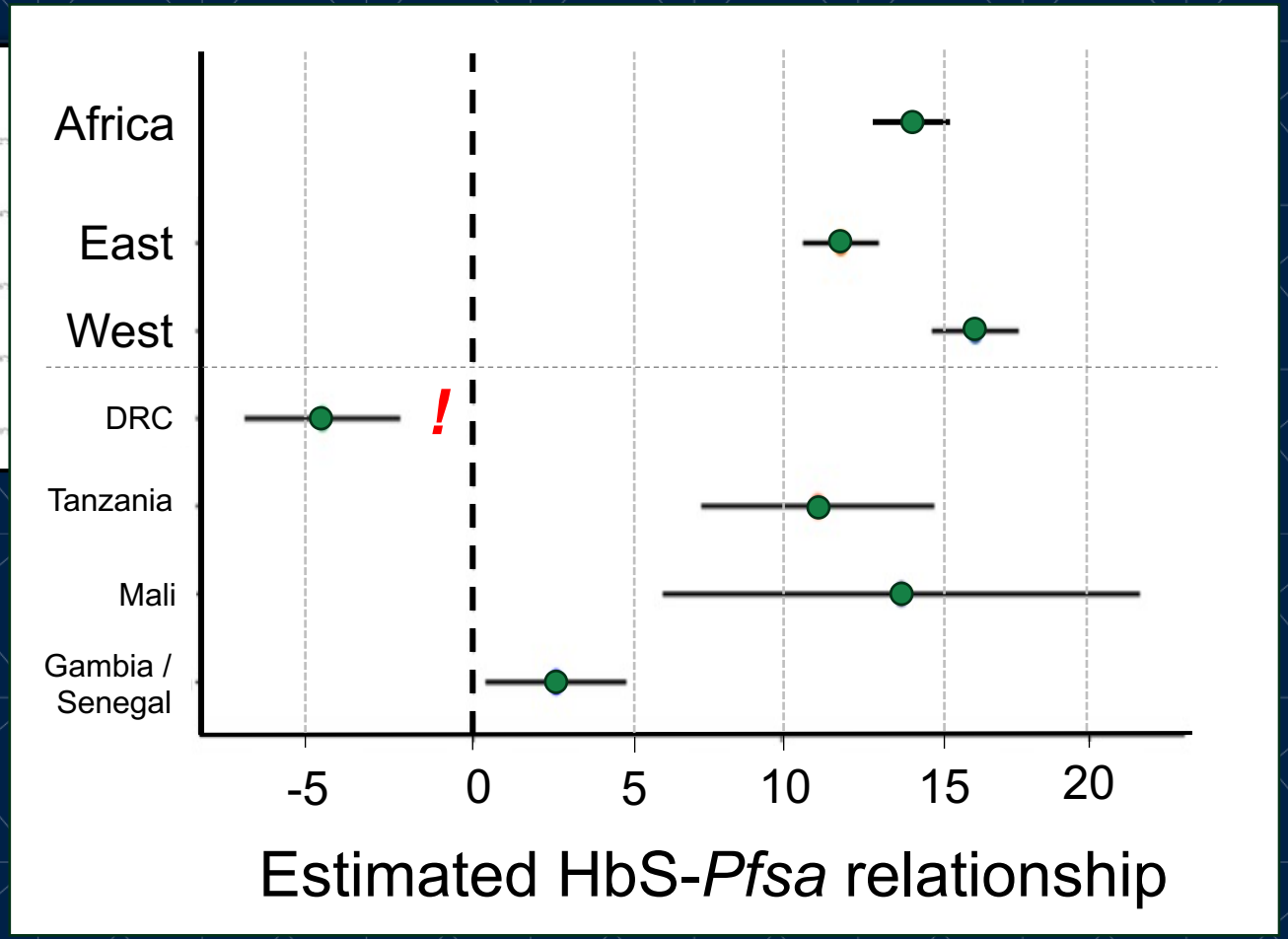

Tanzania

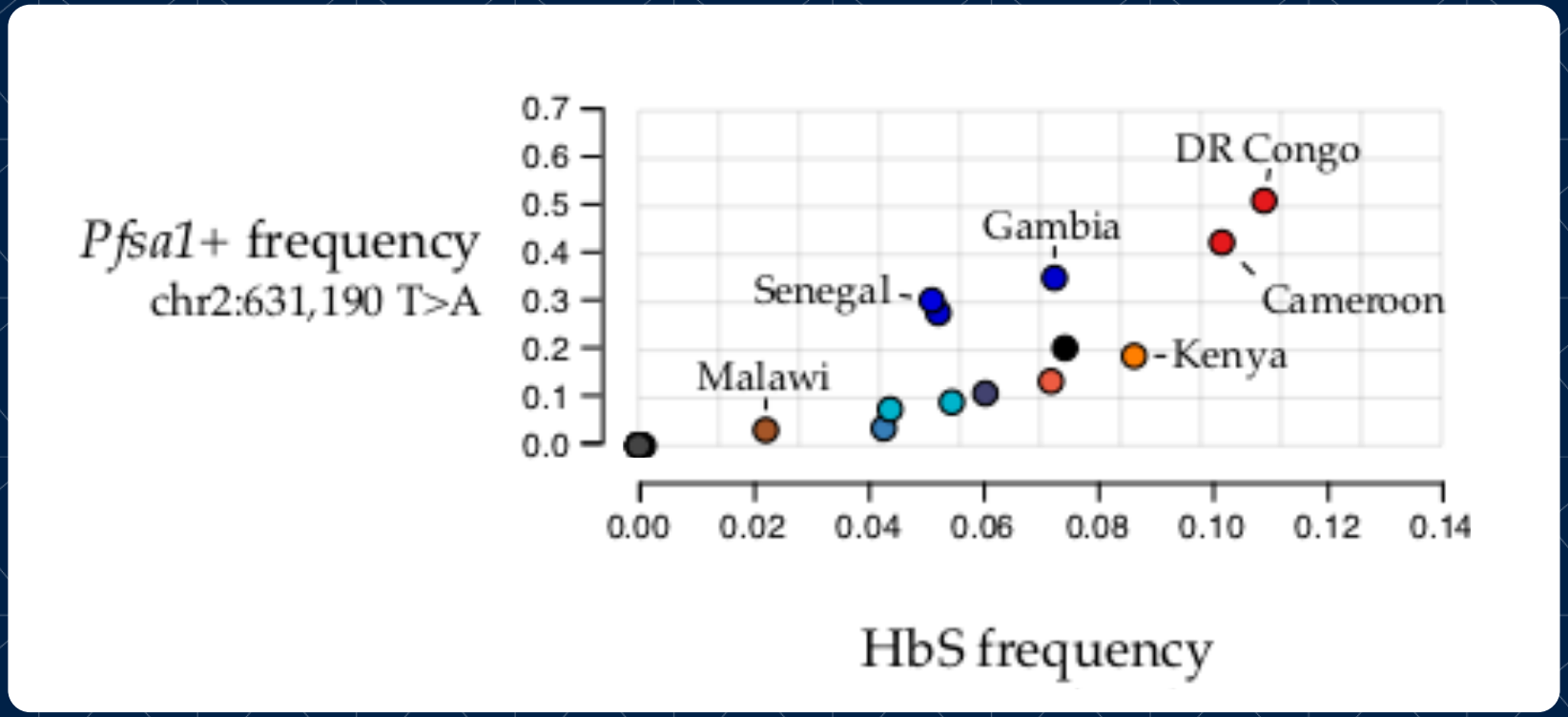

60%
Pfsa1+
frequency
0.4
0.2
0%
5%
10%
15%
20%
HbAS/SS freq


Africa
East
West
Tanzania
Mali
Gambia /
Senegal
-5
0
5
10
15
20
Estimated HbS-Pfsa relationship


c
DRC

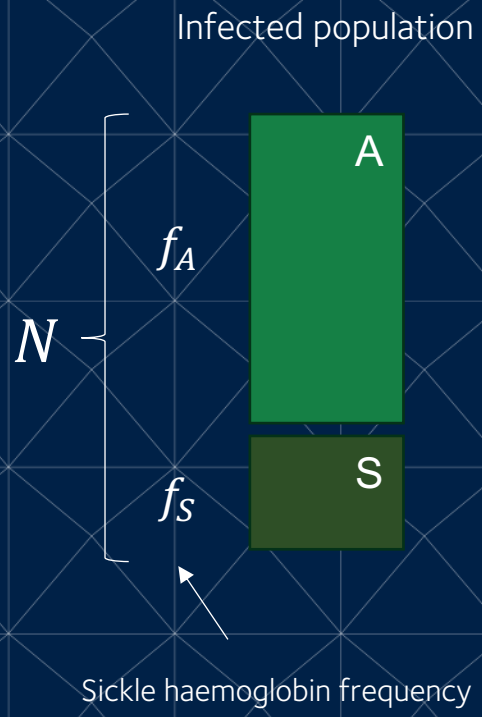

100%
0.75
0.50
0.25
0%
Pfsa1+
frequency
5%
10%
15%
20%
25%
HbAS/SS freq


Africa
East
West
DRC
Tanzania
Mali
Gambia /
Senegal
!
-5
0
5
10
15
20
Estimated HbS-Pfsa relationship



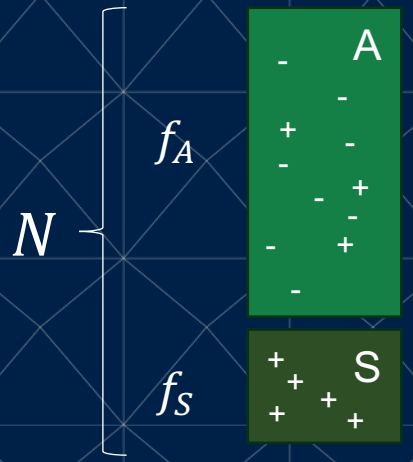
Hypothesis: *Pfsa+ mutations are +vely selected for by HbS but prevented from spreading by a balancing fitness cost*

Simple model of parasite evolution



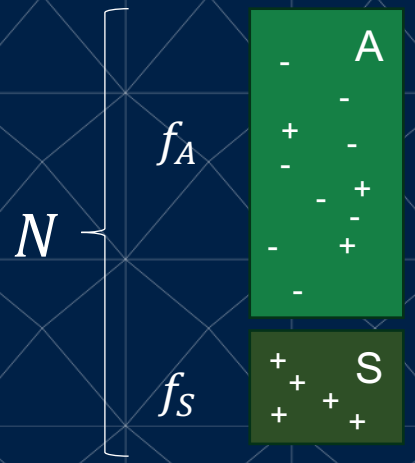
N = "population size"

I assume this is constant



f_+

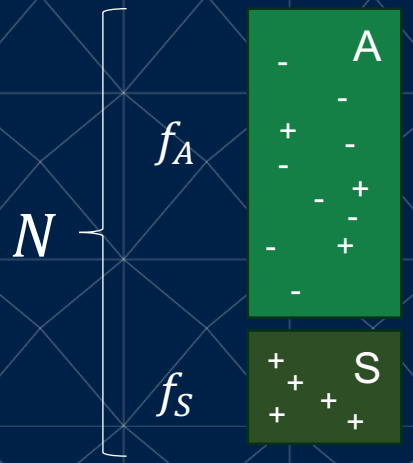
(= frequency of *Pf*+ allele)



f_+
(= frequency of *Pf*+ allele)

Parasite fitnesses
when infecting A and S individuals

$$\begin{pmatrix} & A & S \\ - & 100\% & 0 \\ + & & 100\% \end{pmatrix}$$

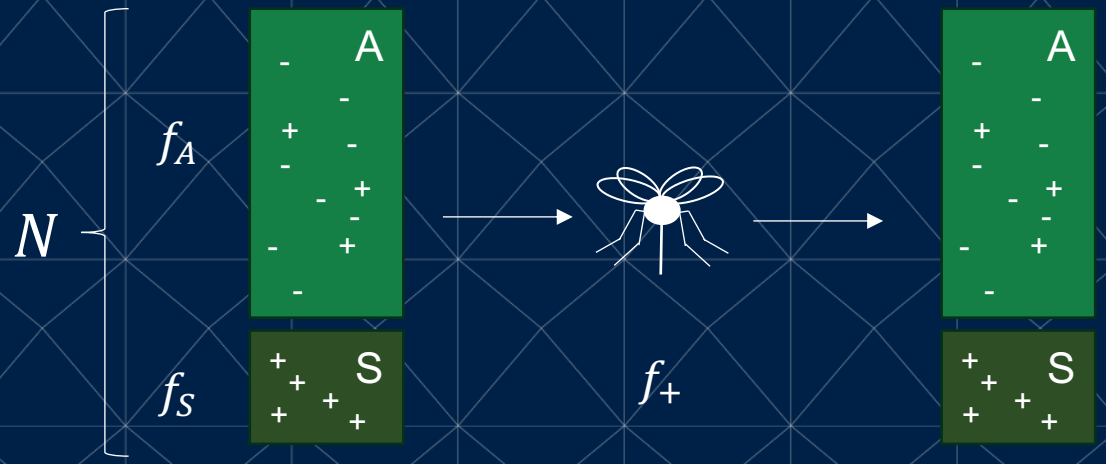


f_+
(= frequency of *Pf*+ allele)

Parasite fitnesses
when infecting A and S individuals

$$\begin{pmatrix} & A & S \\ - & 100\% & 0 \\ + & 90\% & 100\% \end{pmatrix}$$

Competing fitness costs



Parasite fitnesses
when infecting A and S individuals

$$\begin{pmatrix} & A & S \\ - & 100\% & 0 \\ + & 90\% & 100\% \end{pmatrix}$$

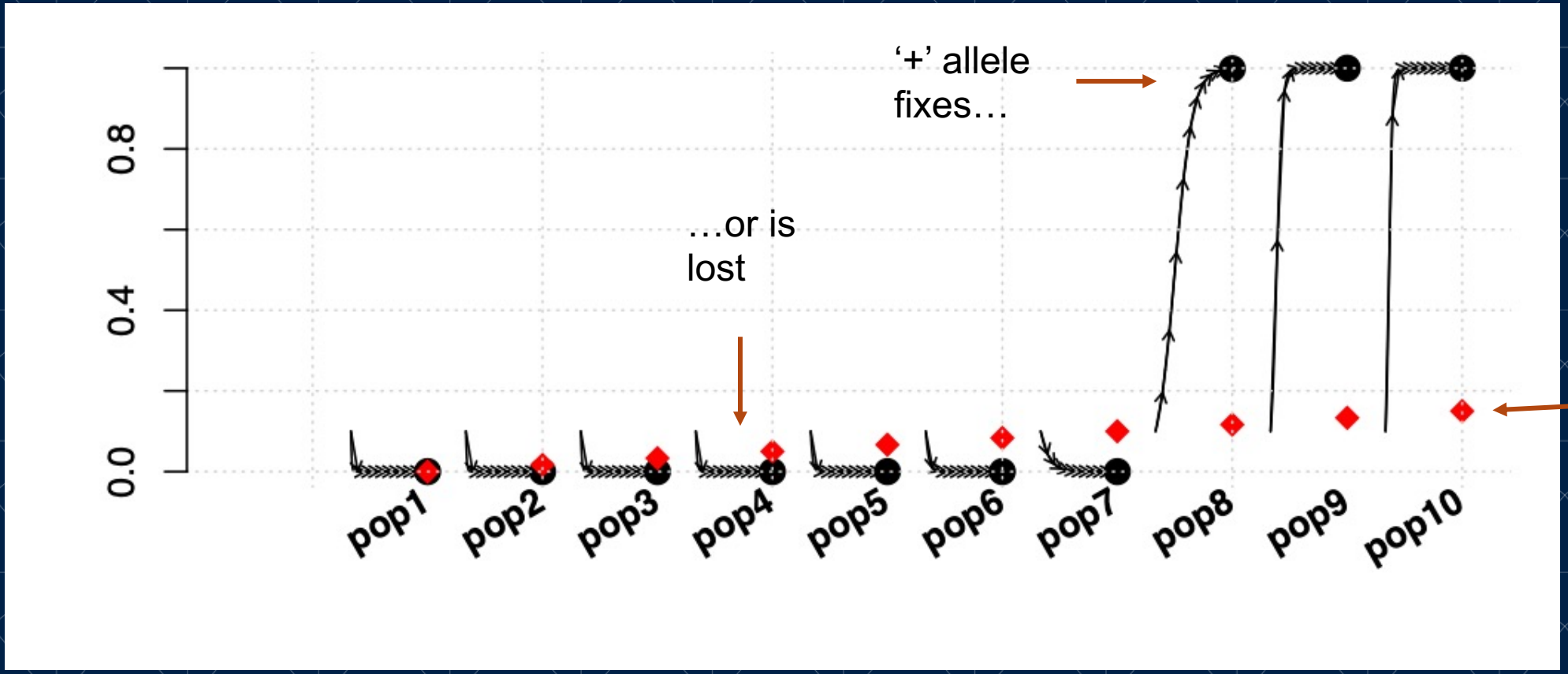
Competing fitness costs

f_+
(= frequency of *Pf*+ allele)

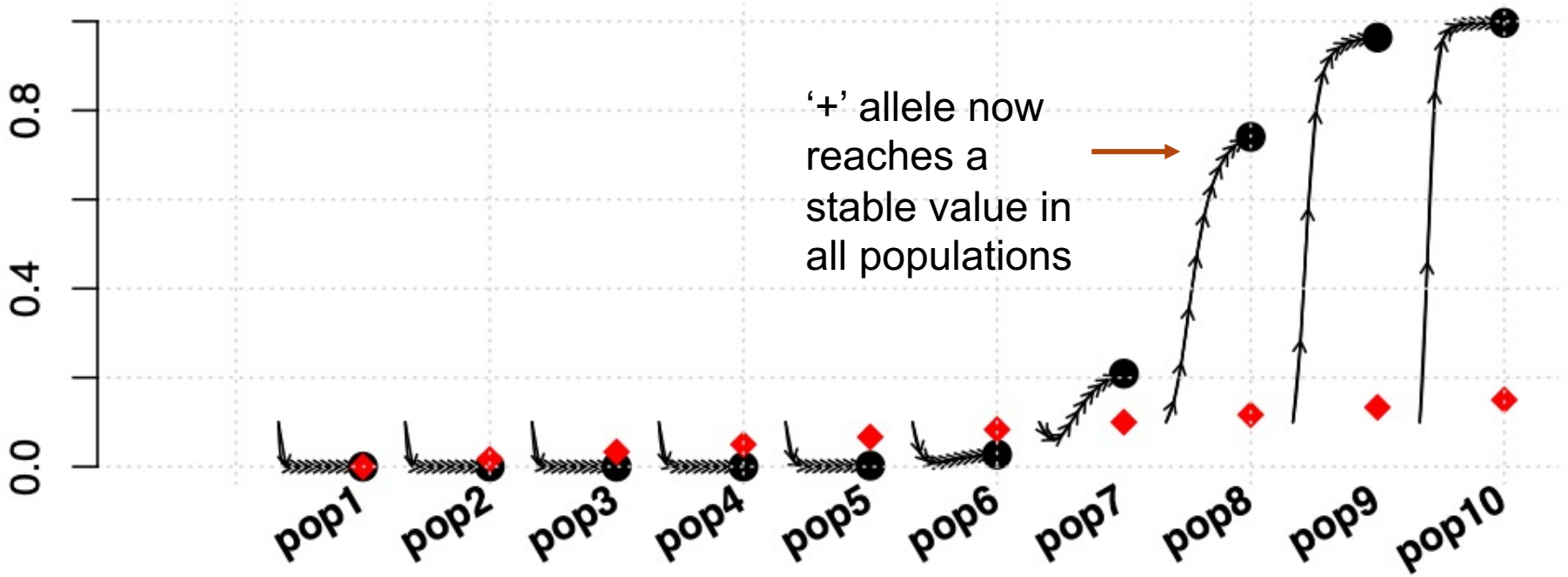
$$\text{new } f_+ = \frac{N f_A f_+ \Gamma_{+A} + N f_S f_+ \Gamma_{+s}}{(\text{denominator})}$$

Simulation: start at 10% frequency, see what happens.

Does not produce polymorphism in independent pops…

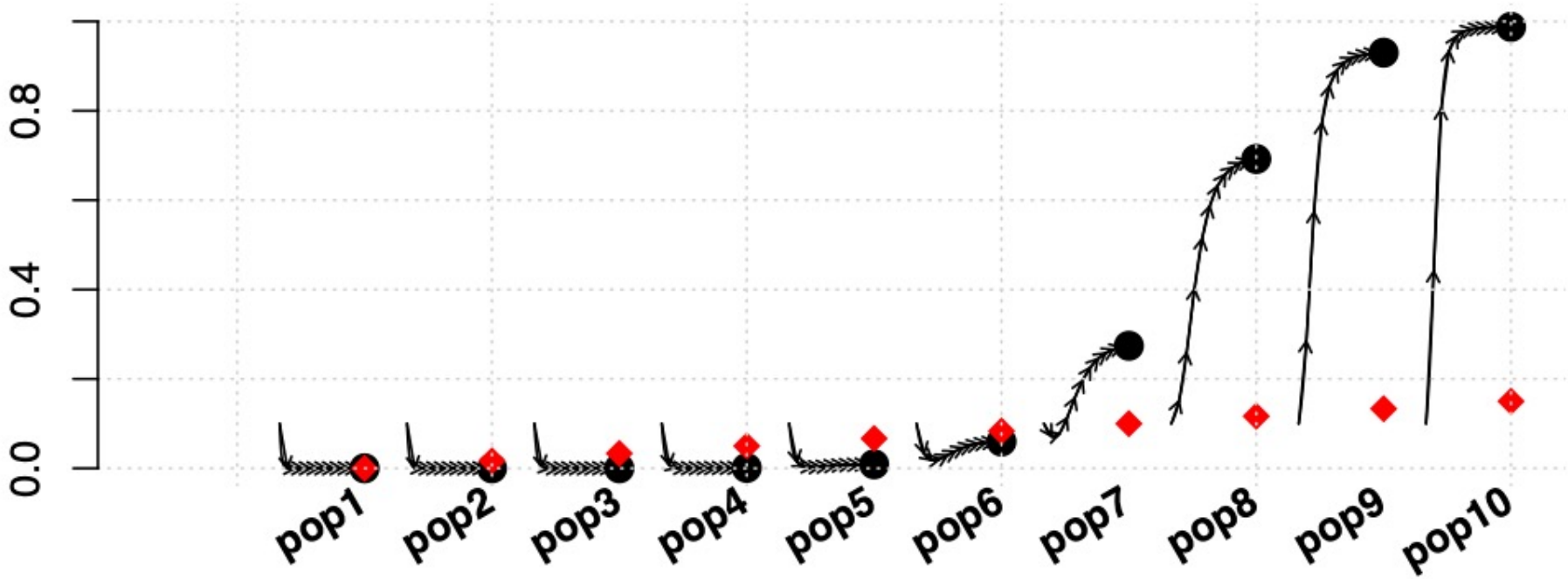


…but does in spatially related pops.



With 1% migration between neighbouring populations

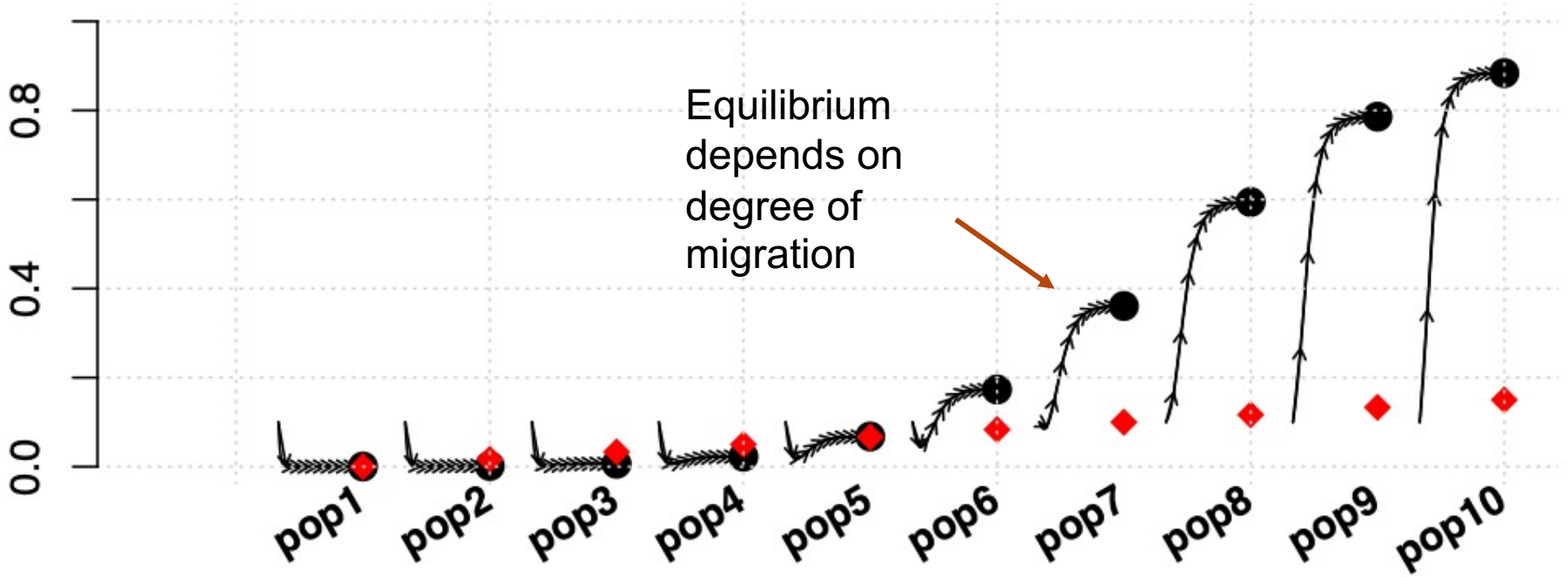
…but stable polymorphism in spatially related pops



With 5% migration between neighbouring populations

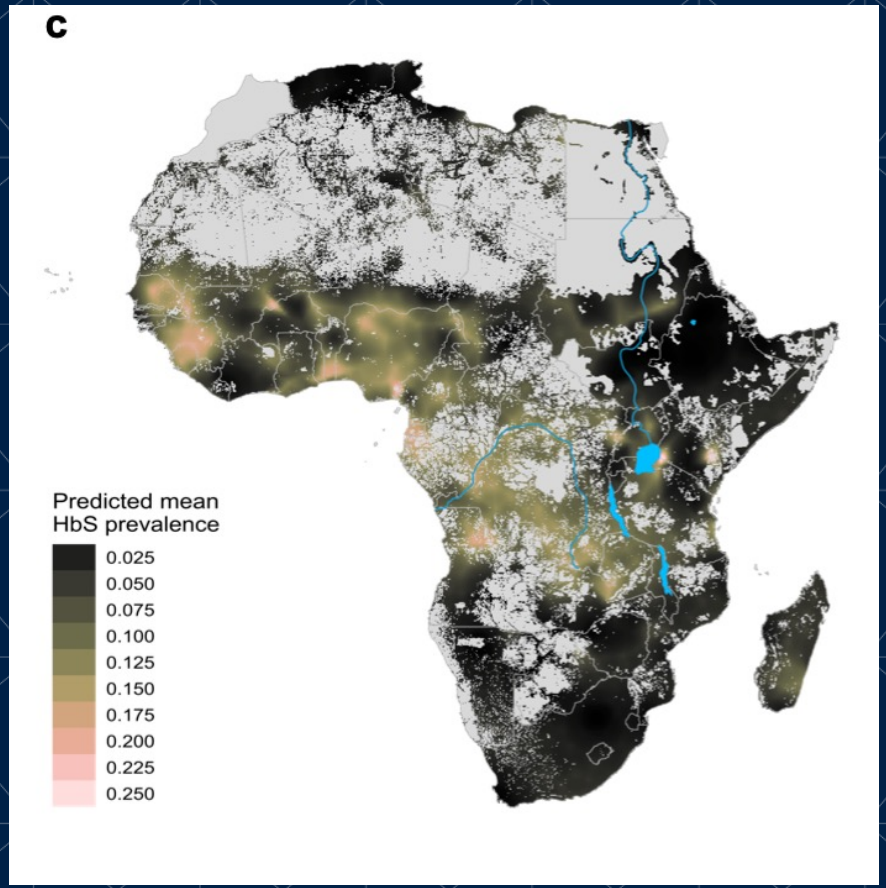
5% migration between neighbouring populations

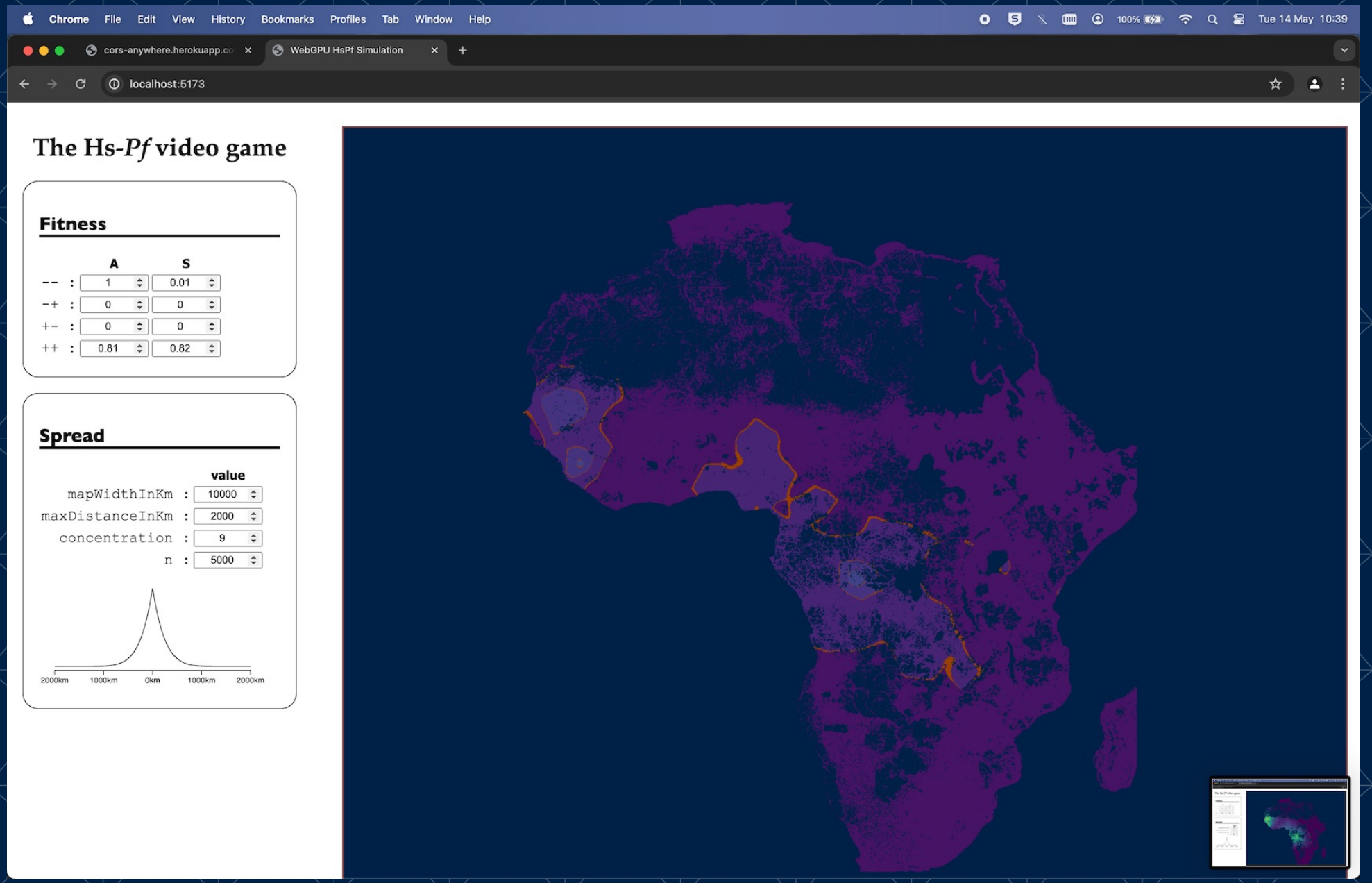
…but stable polymorphism in spatially related pops

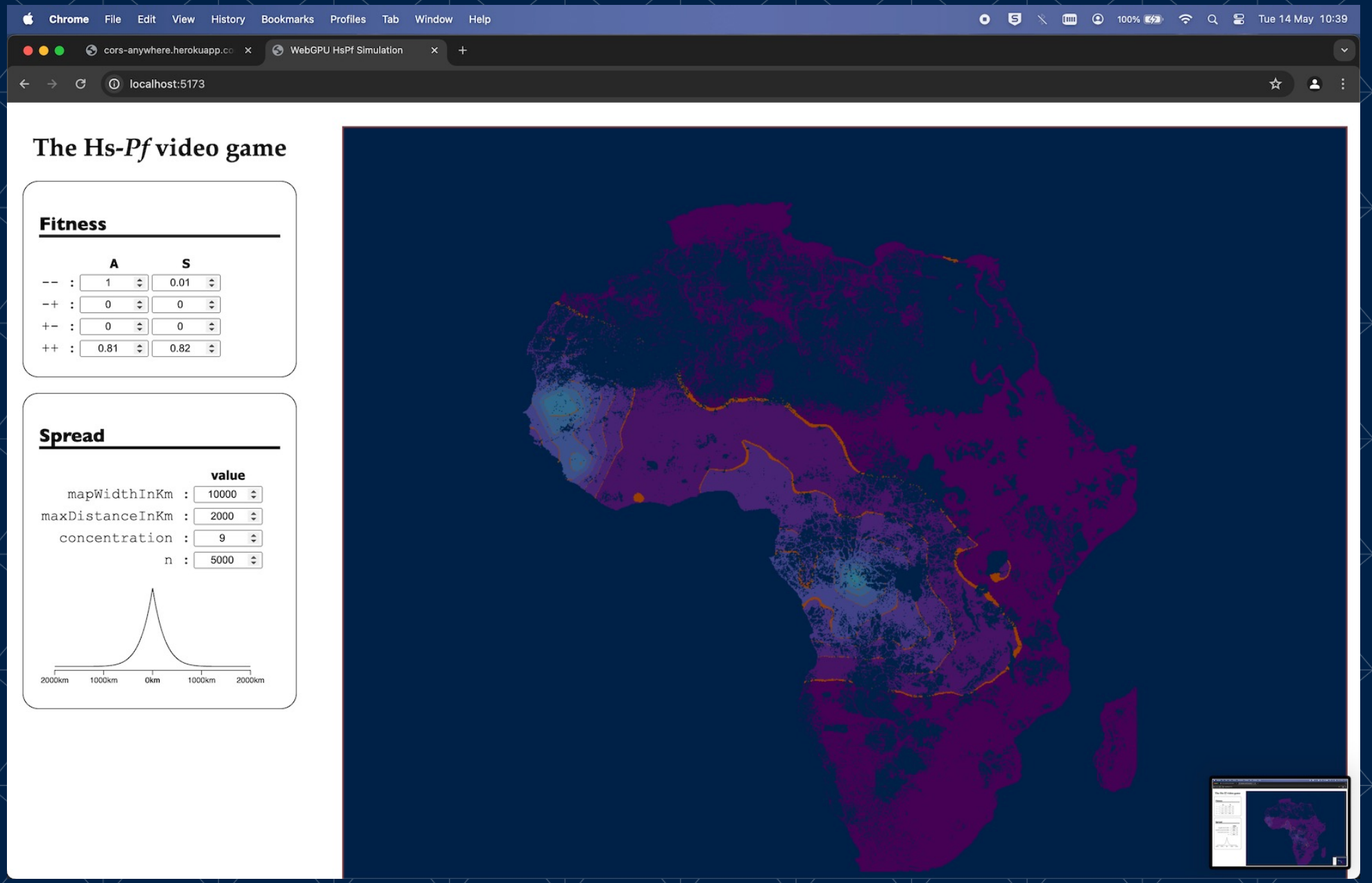


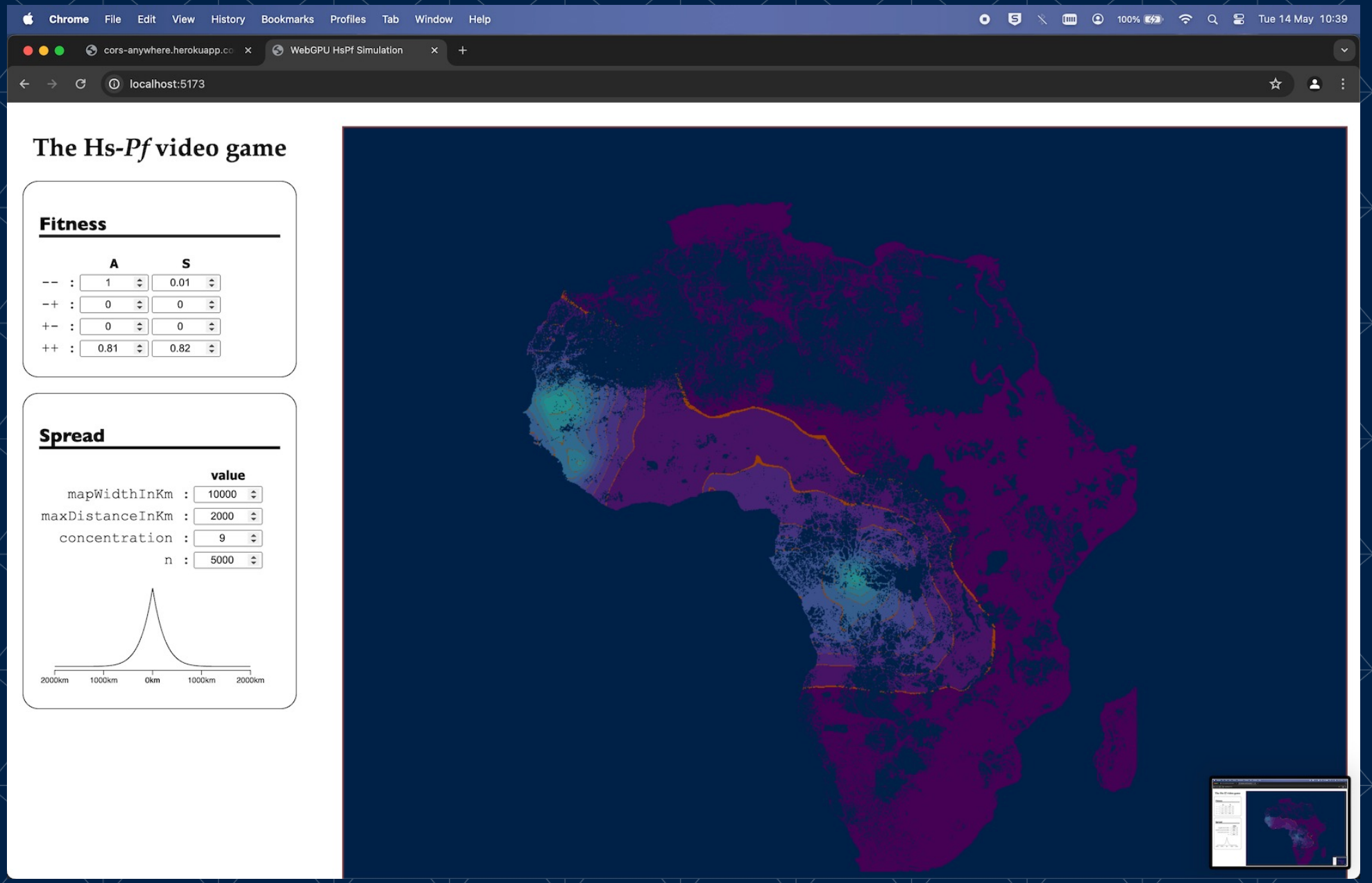
With 10% migration between neighbouring populations

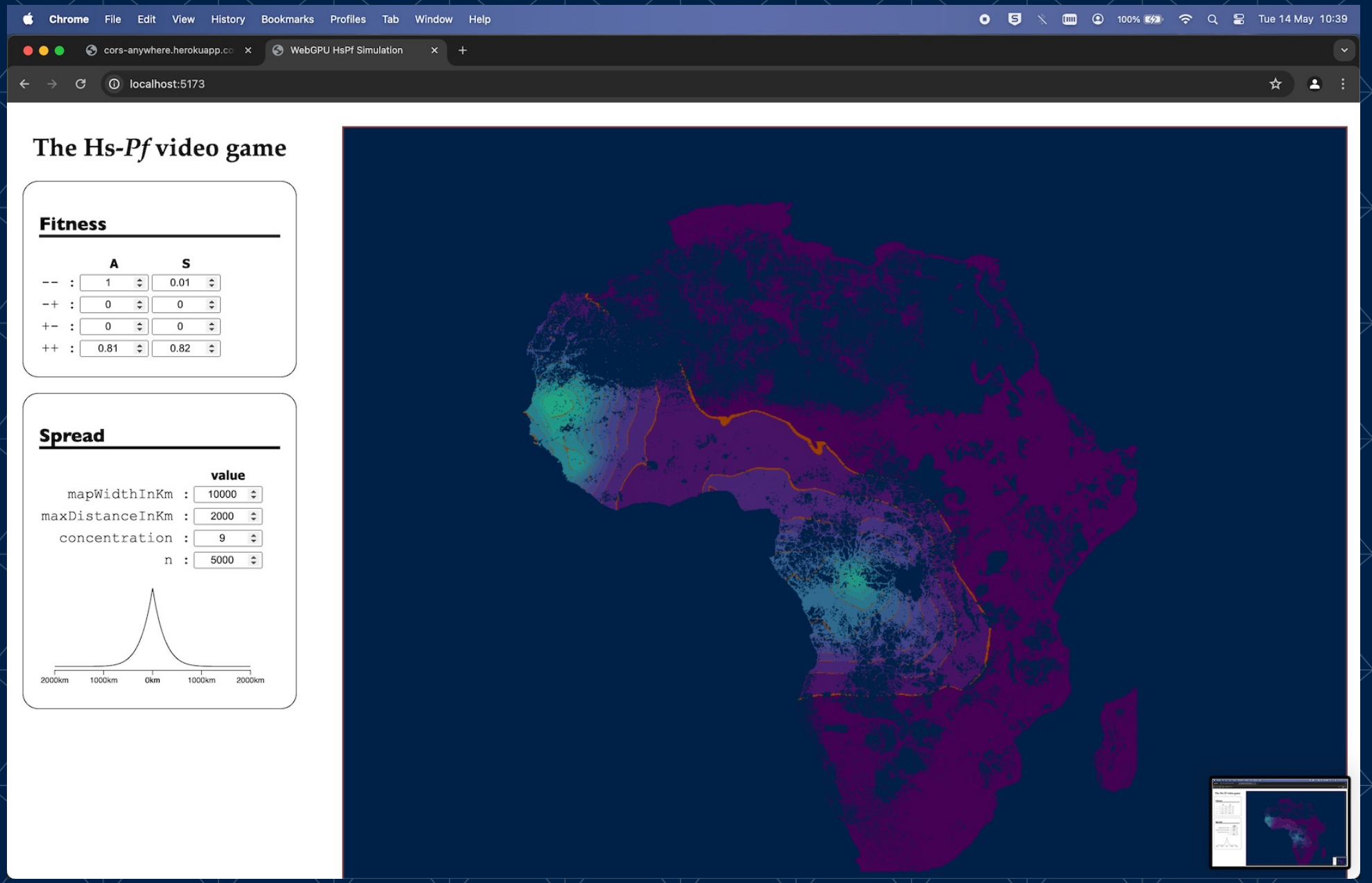
What happens if we overlay this model on the real map of HbS frequency?

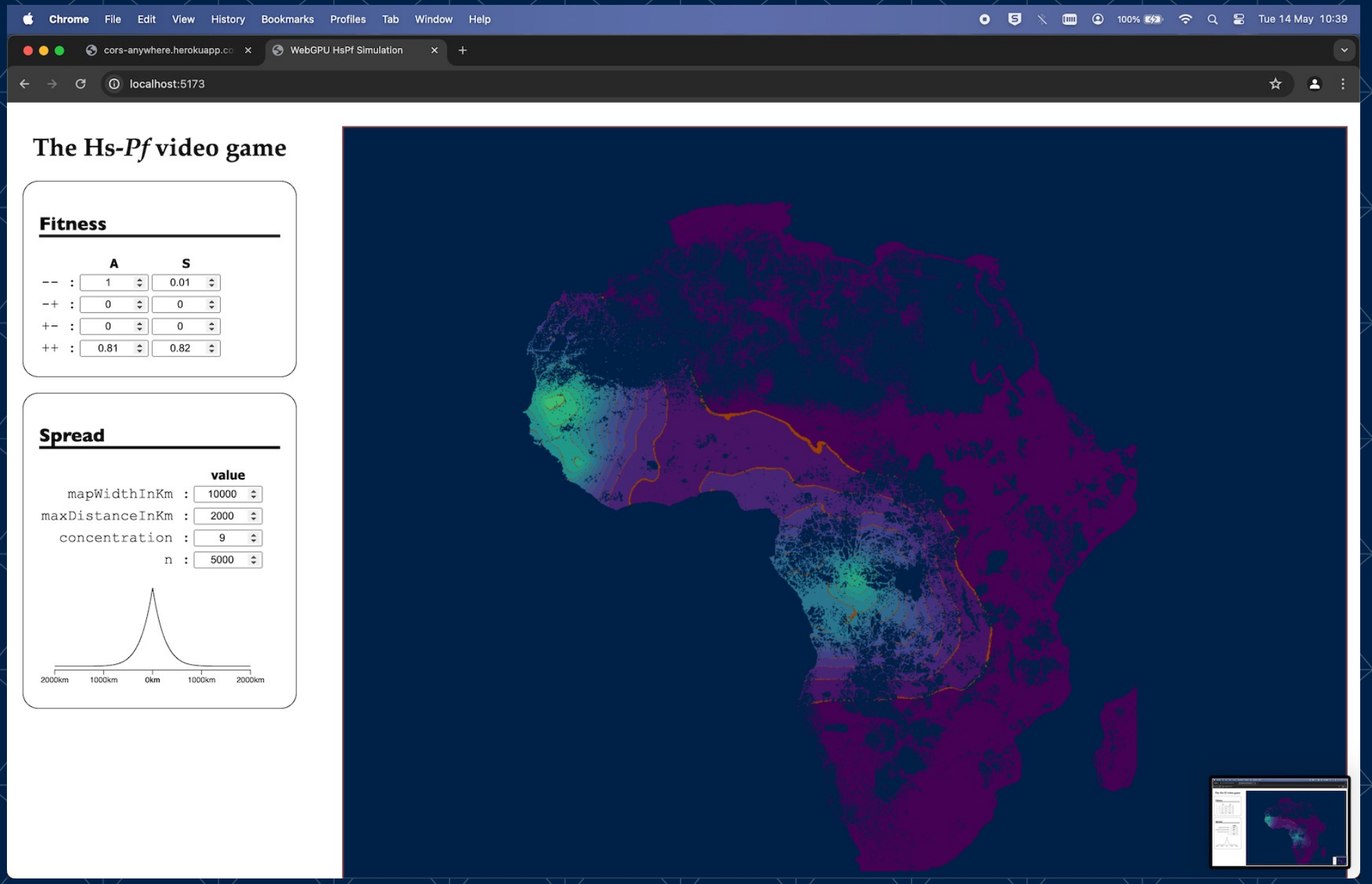


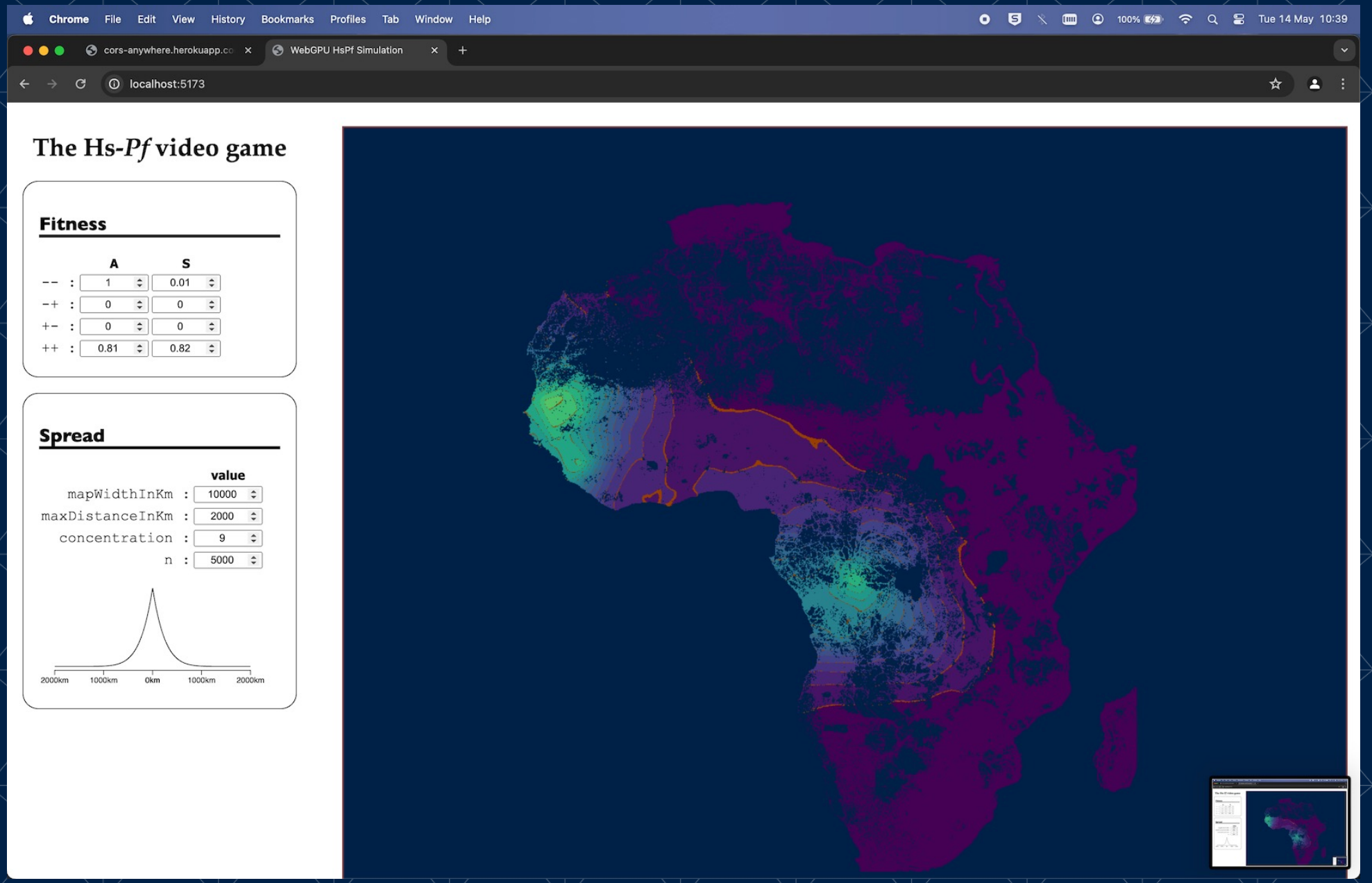

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km

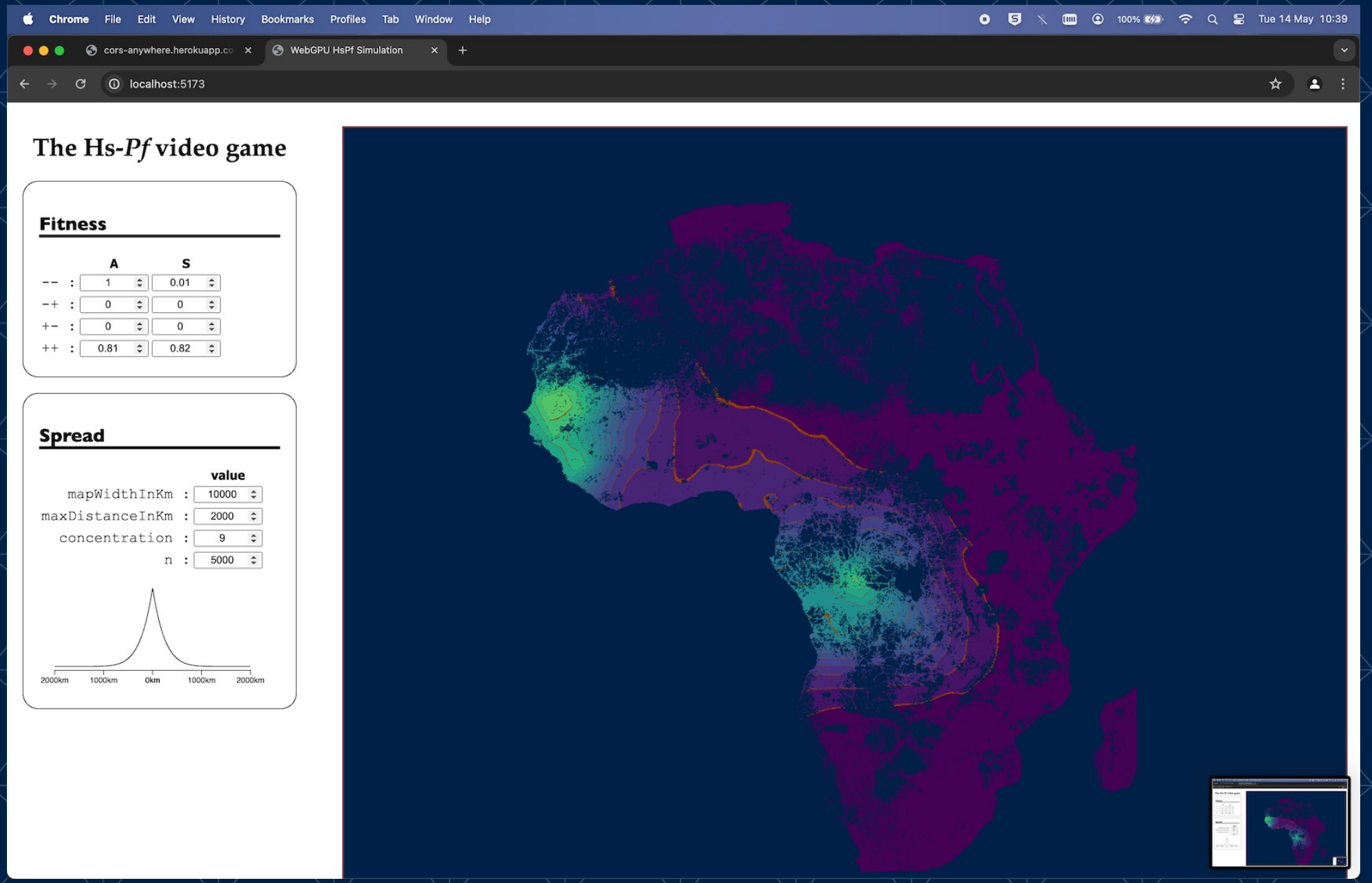

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km

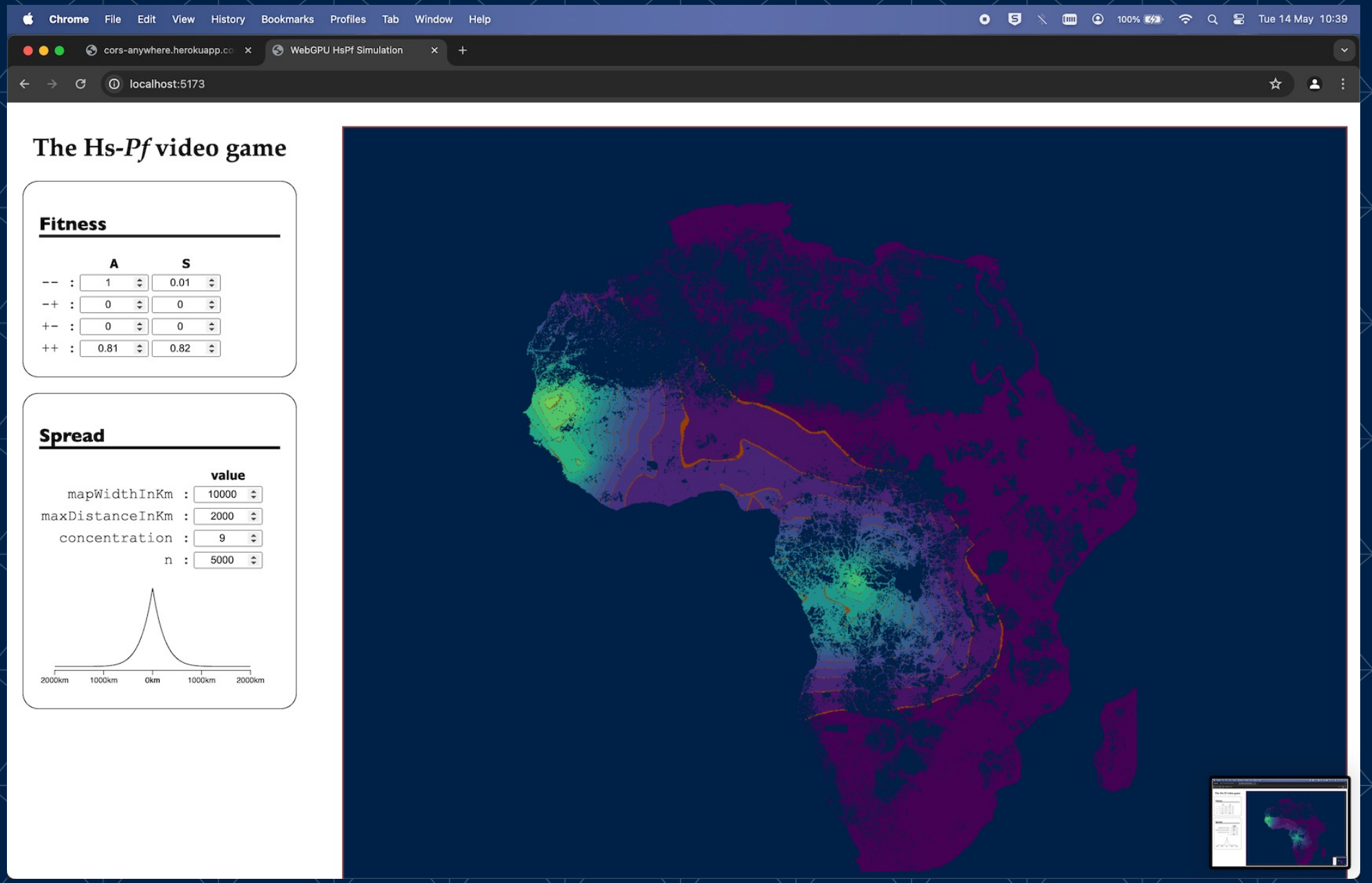

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km

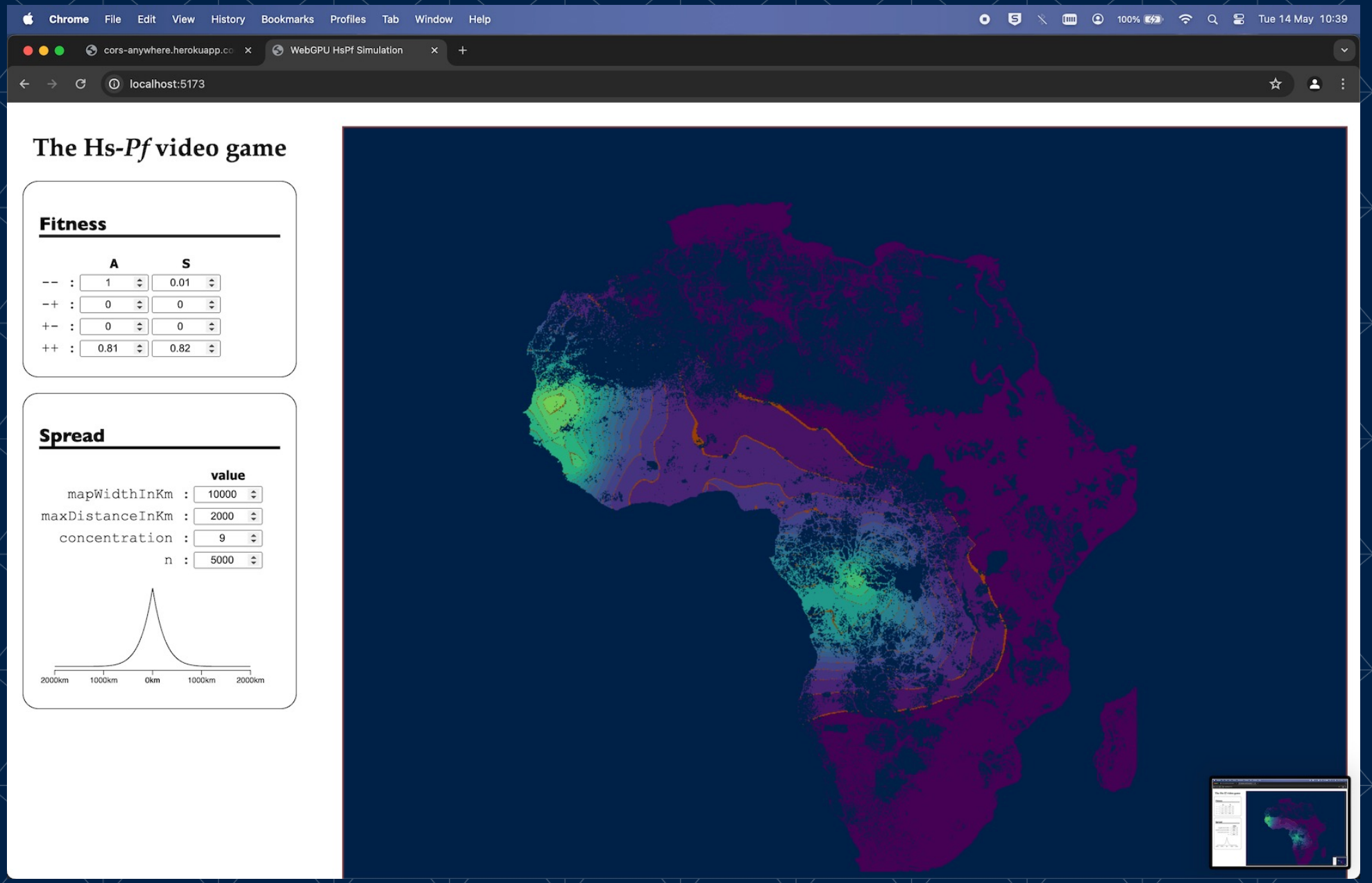

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km

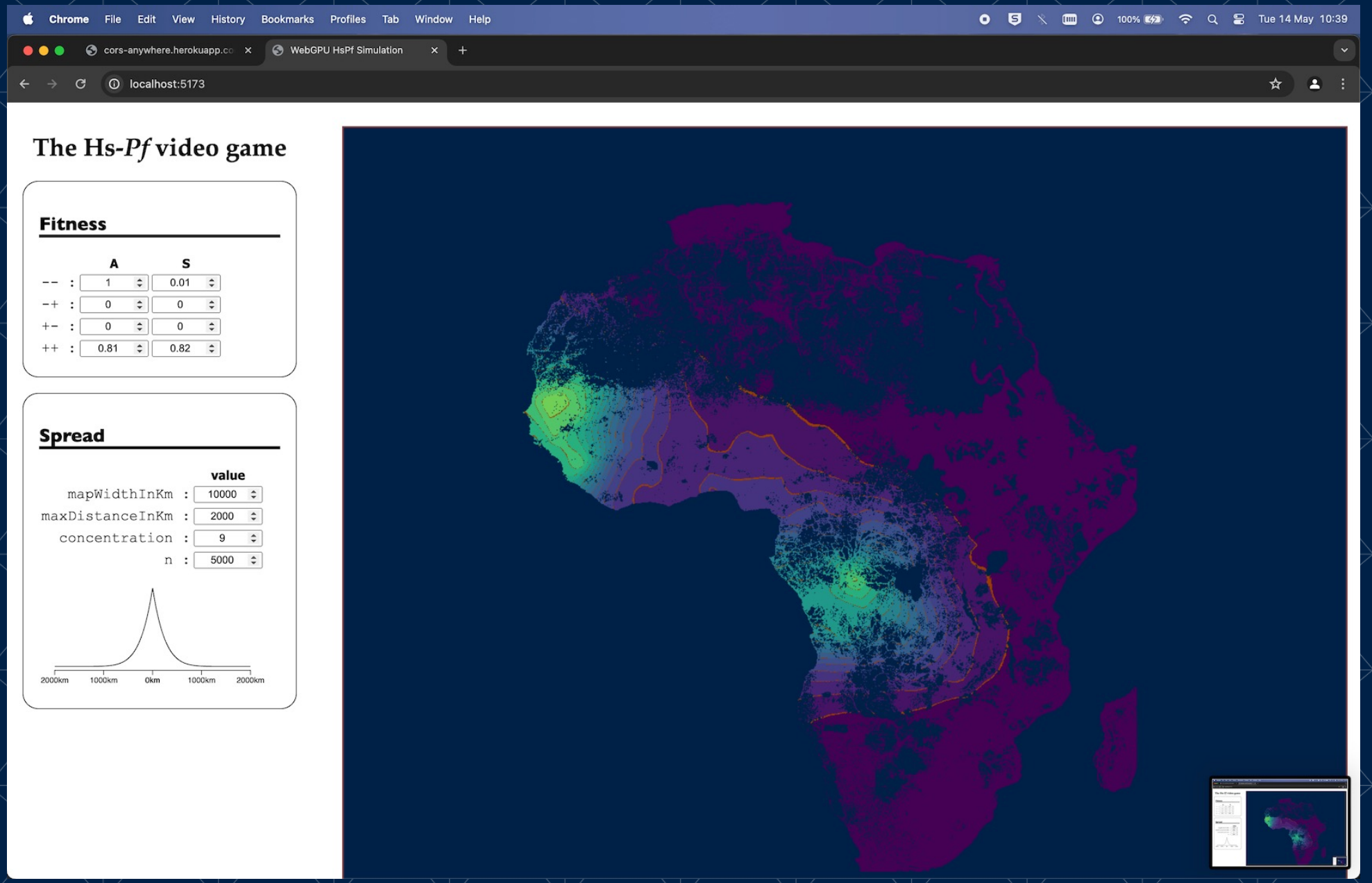

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km 1000km 0km 1000km 2000km

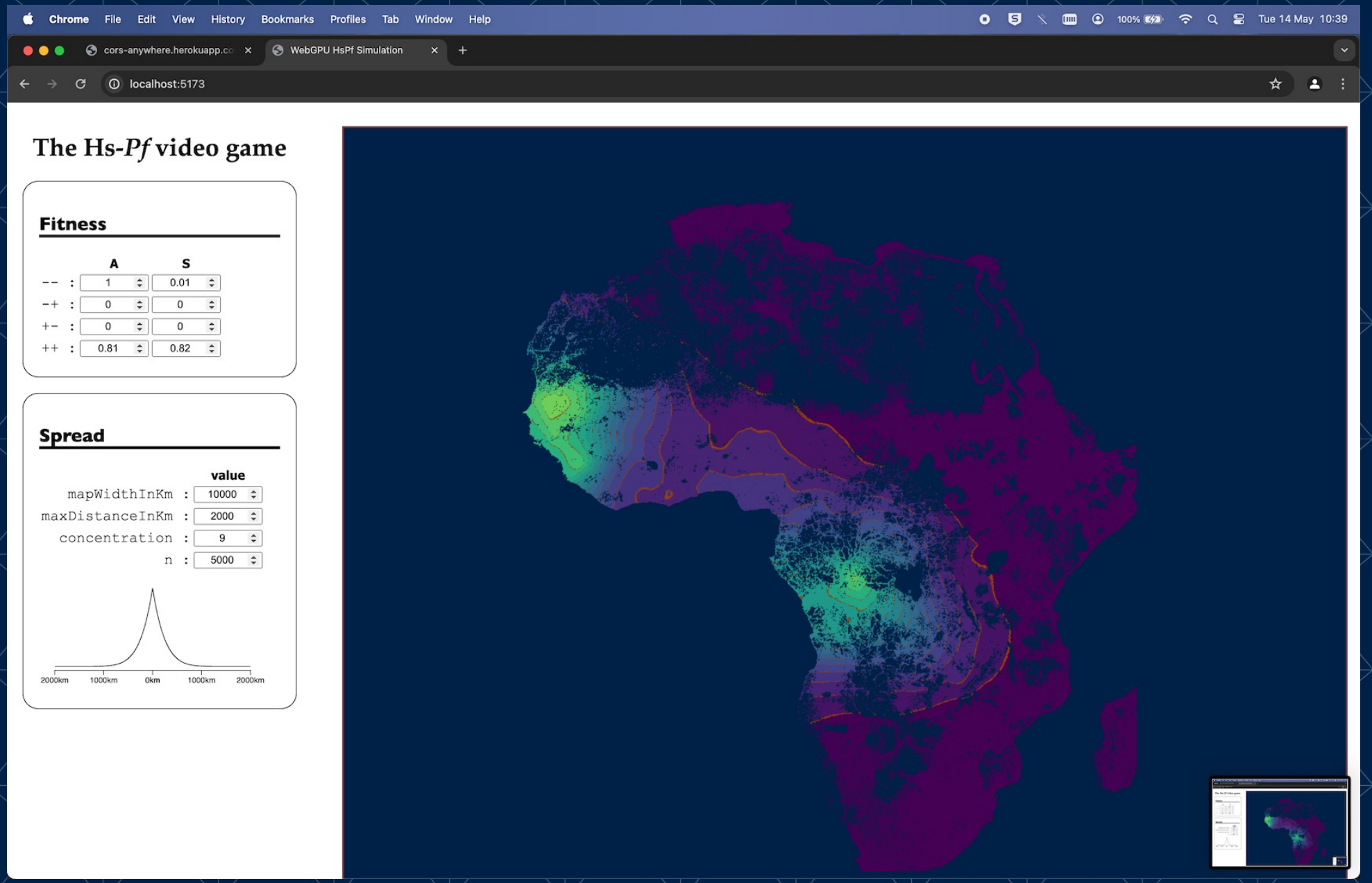

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km

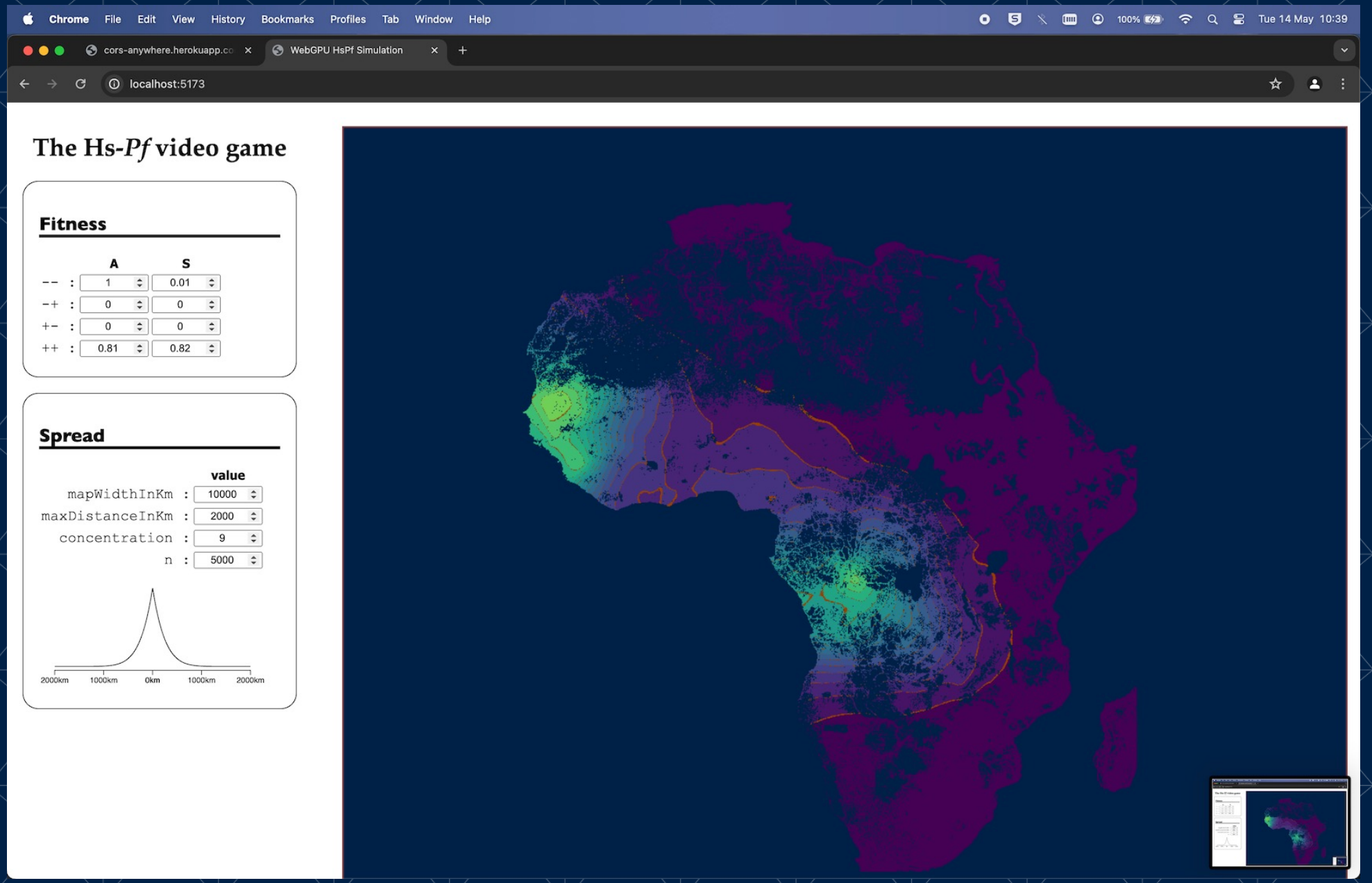

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km

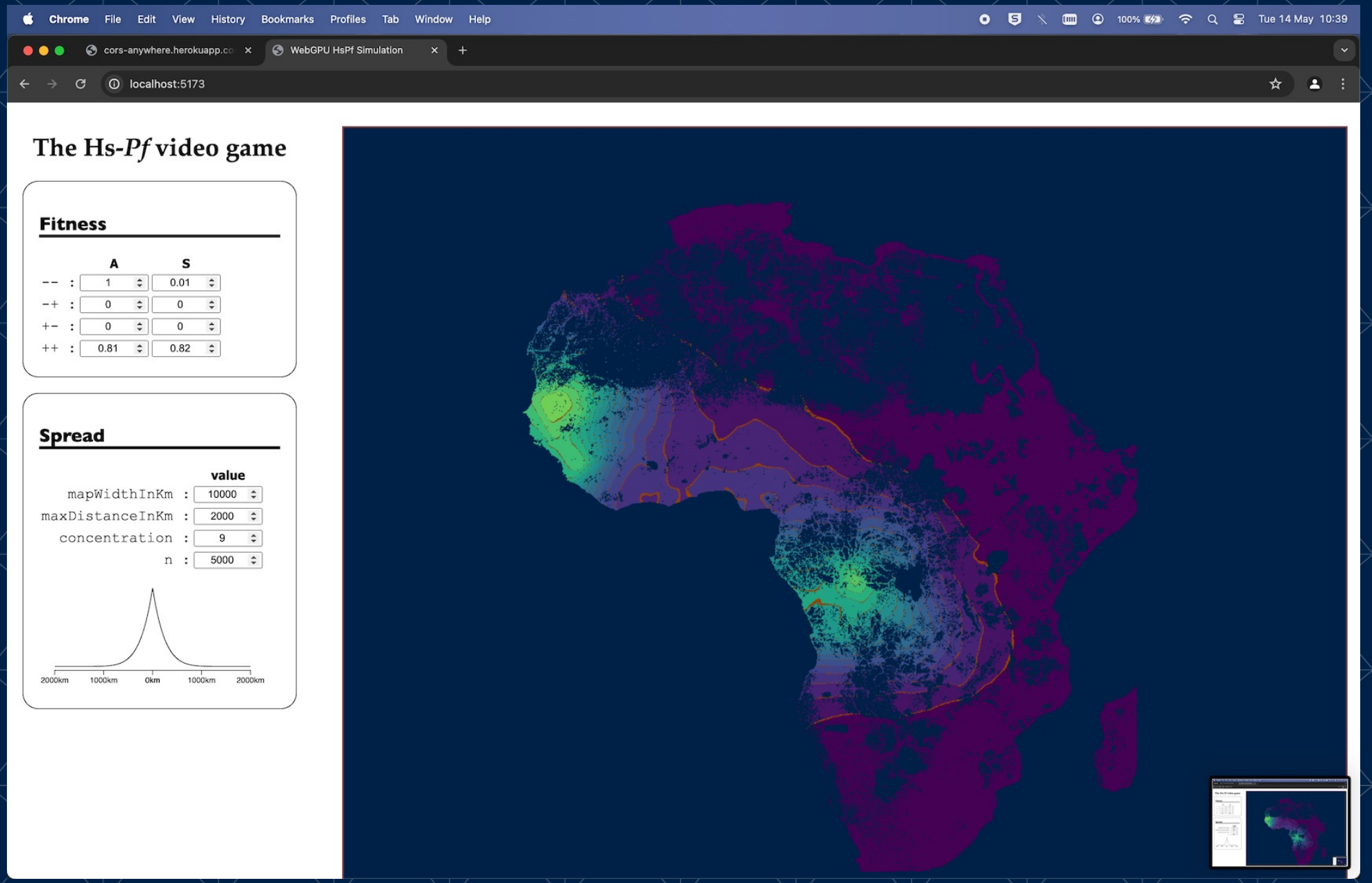

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km

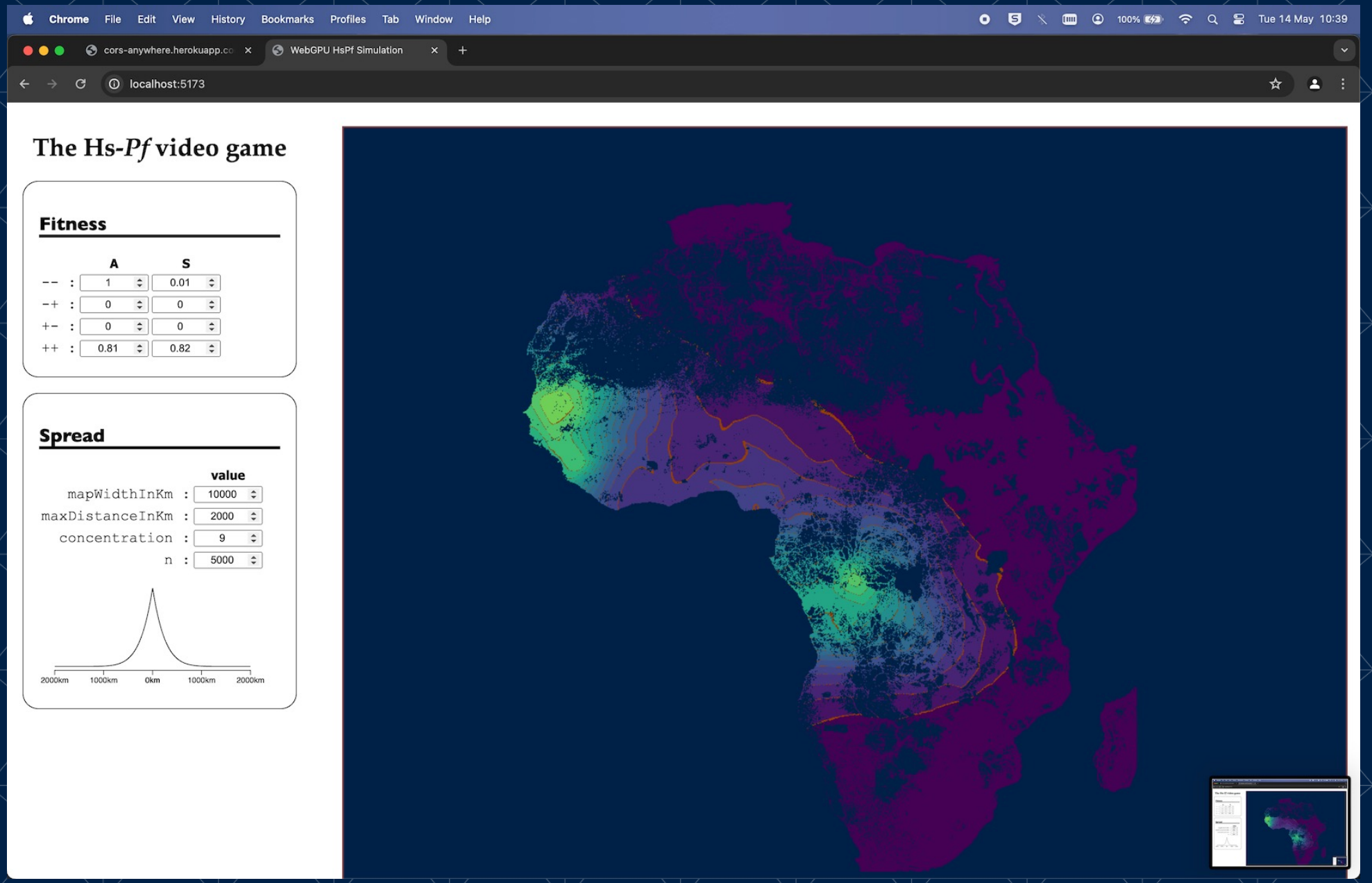

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km

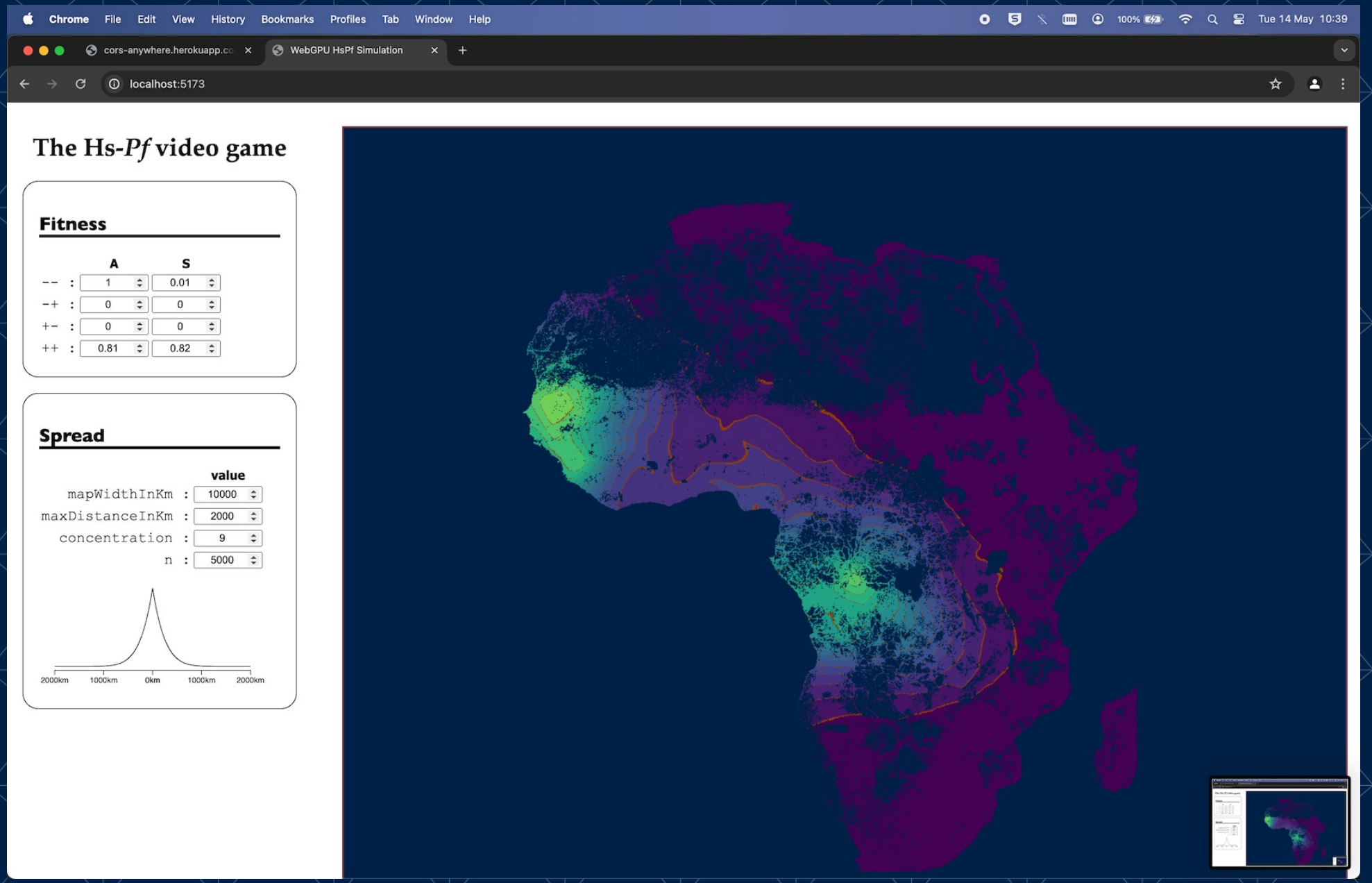

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km

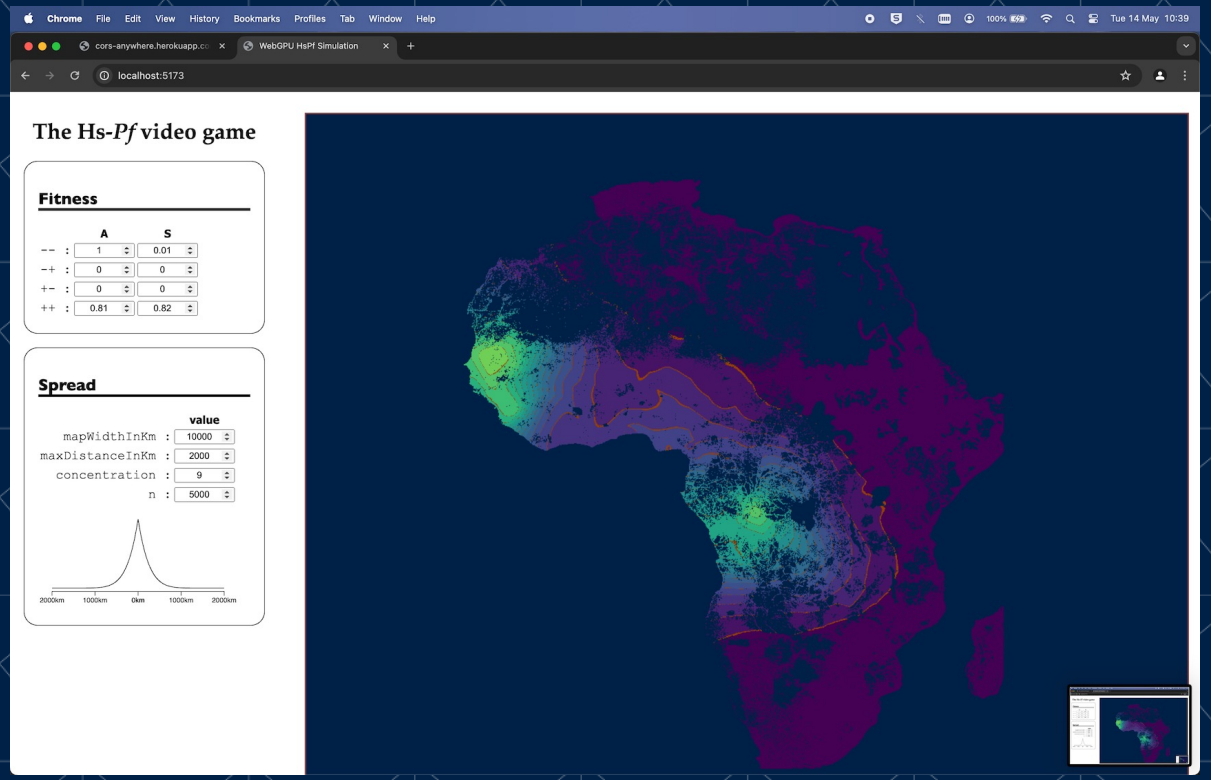

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km

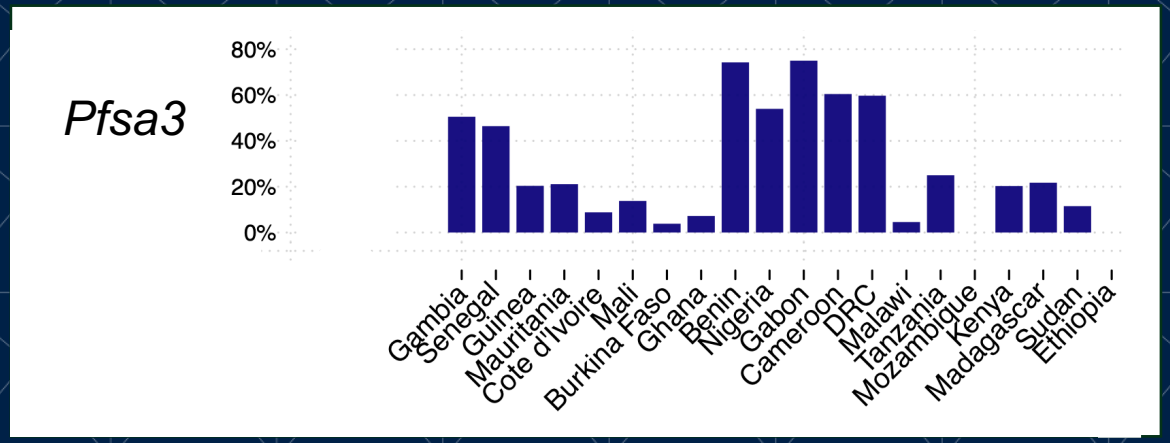

Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km


Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km


Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km


Chrome File Edit View History Bookmarks Profiles Tab Window Help
Tue 14 May 10:39
cors-anywhere.herokuapp.co
WebGPU HsPf Simulation
localhost:5173
The Hs-Pf video game
Fitness
A
S
-- : 1 0.01
-+ : 0 0
+- : 0 0
++ : 0.81 0.82
Spread
value
mapWidthInKm : 10000
maxDistanceInKm : 2000
concentration : 9
n : 5000
2000km
1000km
0km
1000km
2000km





Caution: this is a *very* simple model of parasites evolving on the real HbS map.

However, it converges to a steady state (with frequencies between ~5% and ~80% everywhere, for the right choice of parameters)

Acknowledgement: Peter Todd

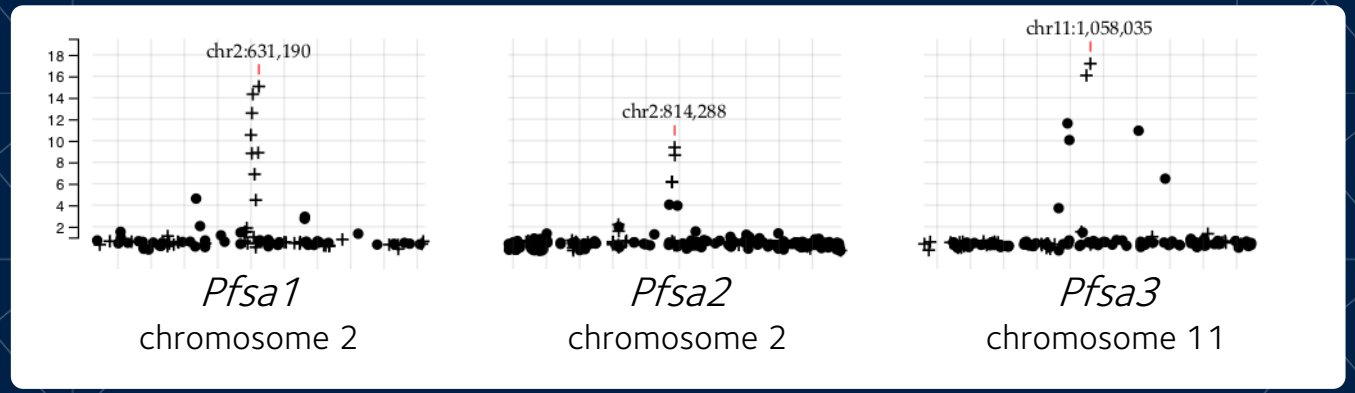
The *Pfsa* alleles are in strong linkage disequilibrium i.e. they co-occur

r=0.80

Kenya:

r=0.75

r=0.66



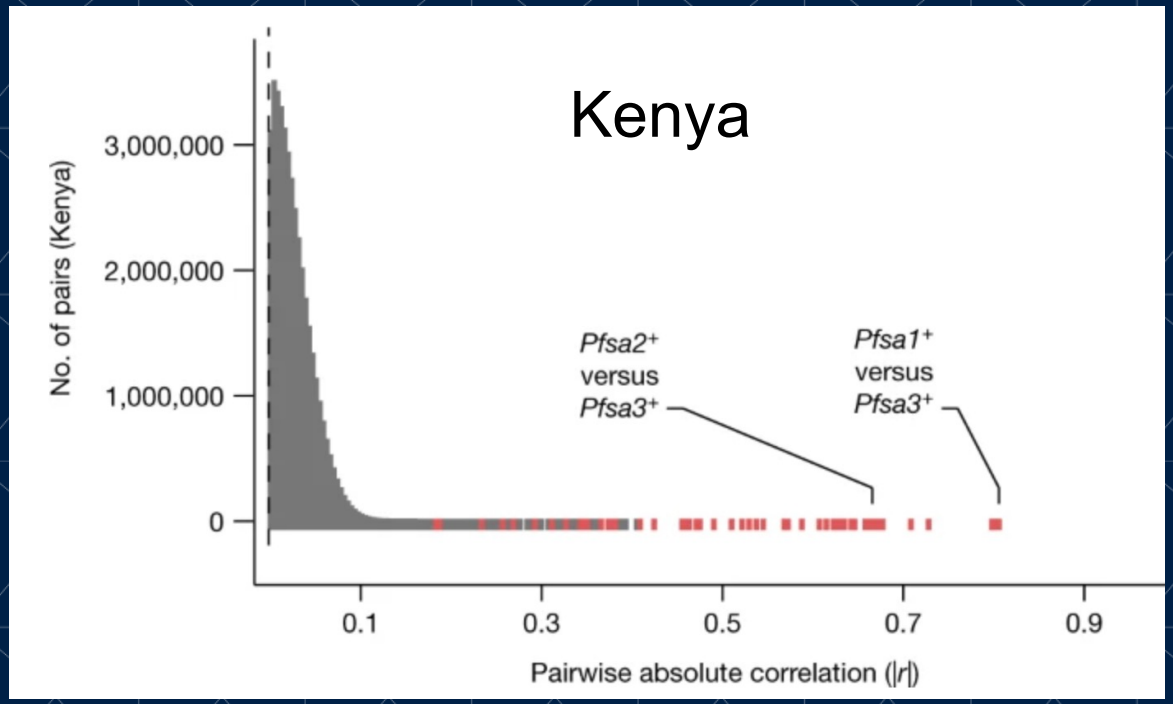
Gambia:

r=0.43

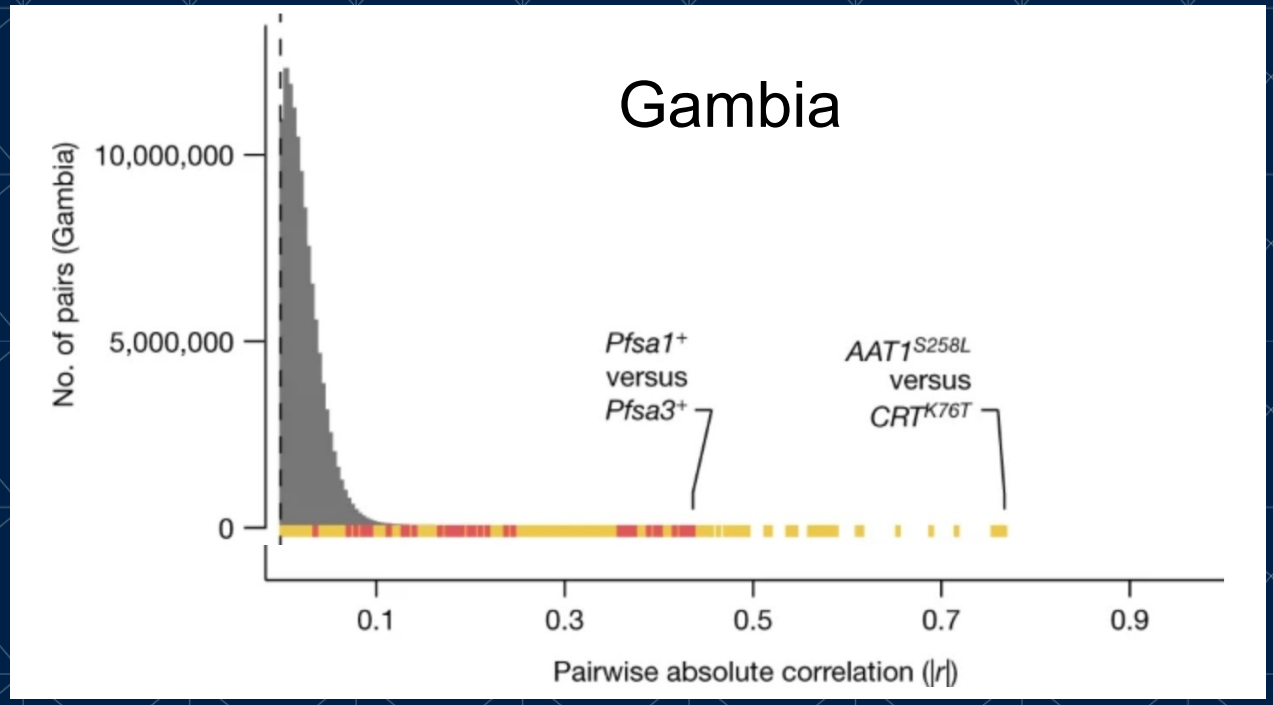
Pfsa1+ vs *Pfsa3+*

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



In Kenya, *Pfsa* variants account for all of the strongest between-chromosome LD
Other genes like CLAGs and REX1 turn up with elevated LD



In Gambia, the strongest between-chromosome LD is for drug resistance mutations - CRT vs AAT1
Next down the list is Pfsa1 vs Pfsa3.

Questions

Does between-locus LD imply that there is functional epistasis between the genes?

Or is it just an emergent feature of the population genetics and natural selection?

Talk outline

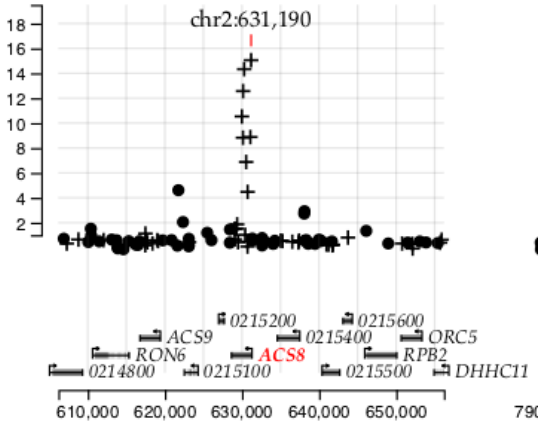
- Malaria protection due to sickle haemoglobin depends on parasite genotype
- Population genetics and host-parasite coevolution
- *Pfsa* biology

How does sickle provide protection?

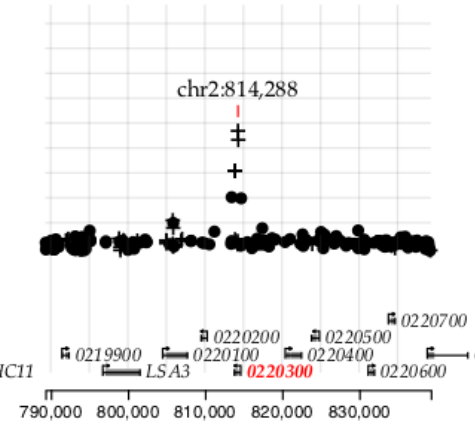
Theory 1: sickle induces alterations to the red cell cytoskeleton, that reduce the ability of parasites to put their molecules (notably PfEMPs) on the surface.

Theory 2: parasites are unable to grow in sickle trait erythrocytes, in the low-oxygen conditions likely to be found in real infections.

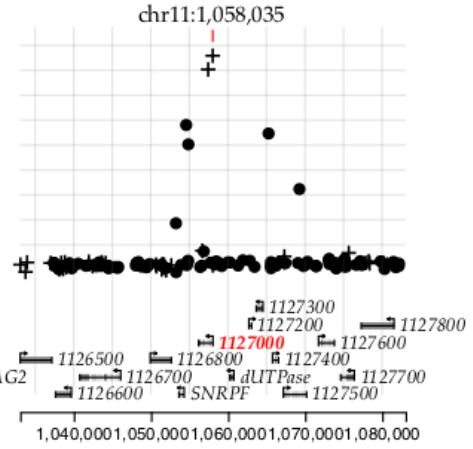
Pfsa1 (chr2)



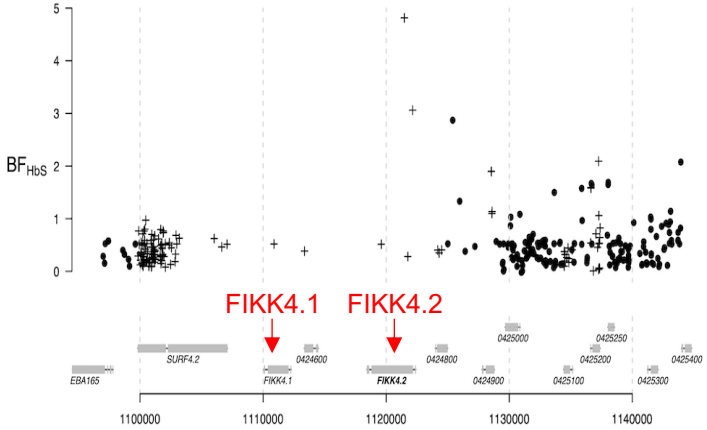
Pfsa2 (chr2)



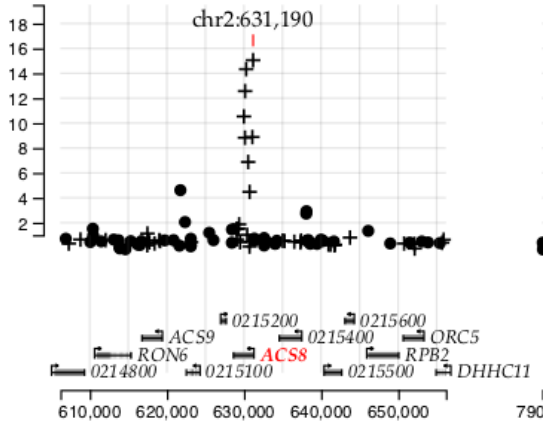
Pfsa3 (chr11)



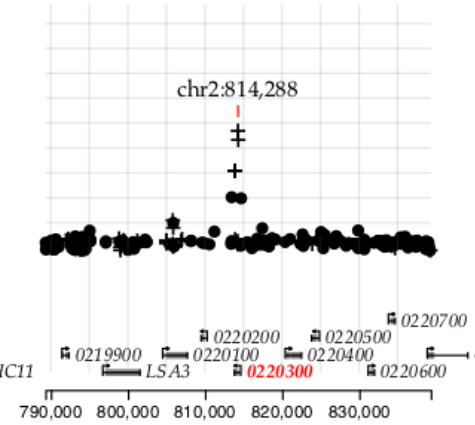
Pfsa4 (chr4)



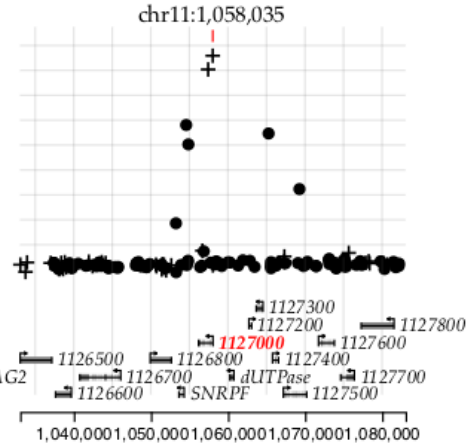
Pfsa1 (chr2)



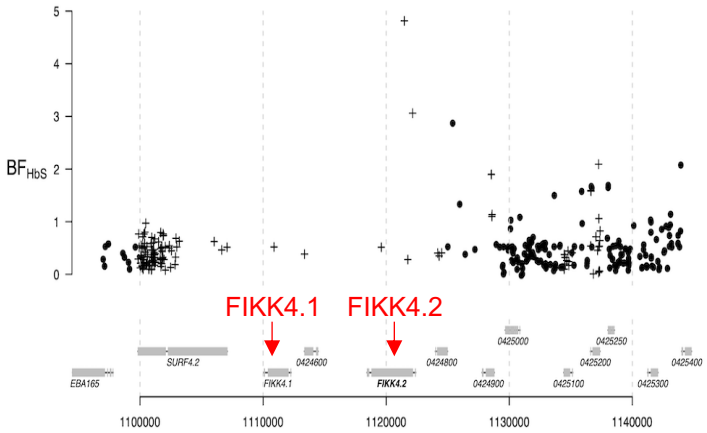
Pfsa2 (chr2)



Pfsa3 (chr11)



Pfsa4 (chr4)



PfACS8 is an Acyl co-A synthetase family member. To my knowledge it has not been observed exported (unlike other family members)

0020300 is an exported protein, unknown function.
Has been observed in cytosol
Tamira Butler PhD Thesis
Daniel Golberg lab

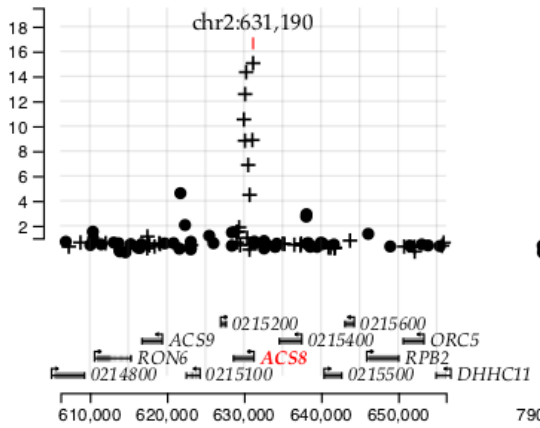
1127000 is a predicted exported, putative Tyrosine phosphatase

FIKK4.1/4.2 are exported Serine/Threonine Kinases

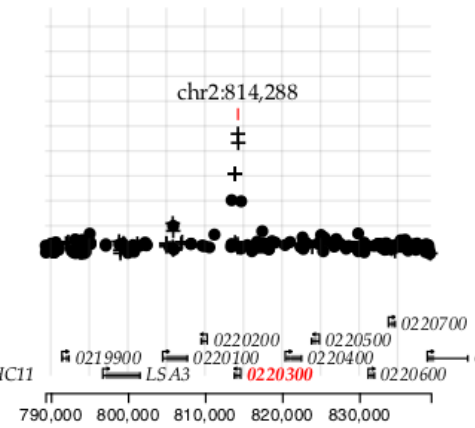
None are thought to be essential (for parasite growth in vitro)

CENTRE *for* HUMAN GENETICS

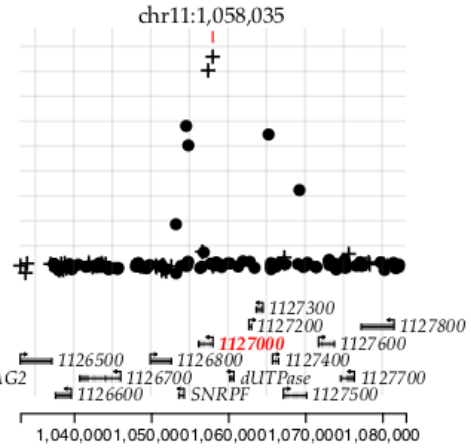
Pfsa1 (chr2)



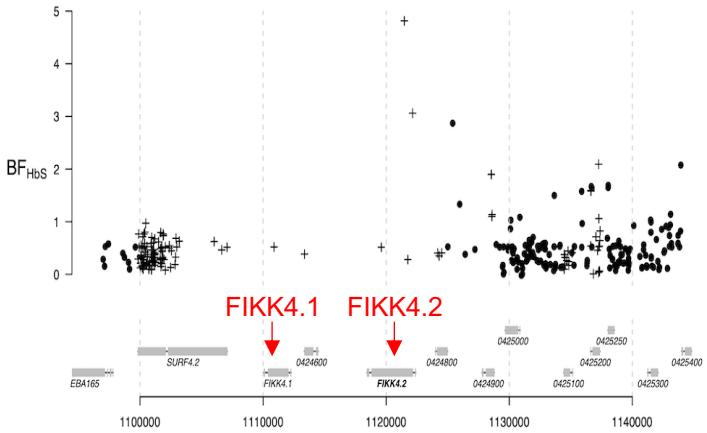
Pfsa2 (chr2)



Pfsa3 (chr11)

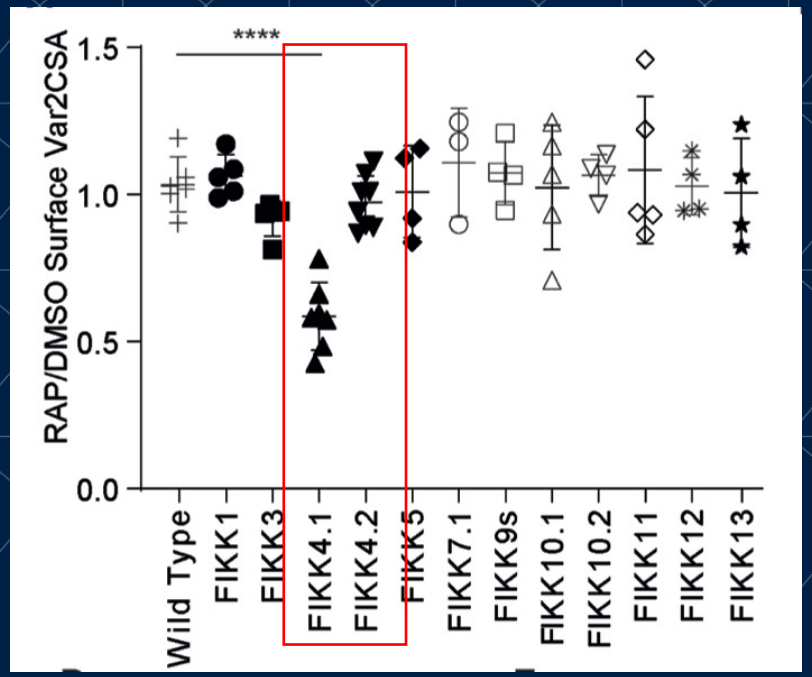


Pfsa4 (chr4)



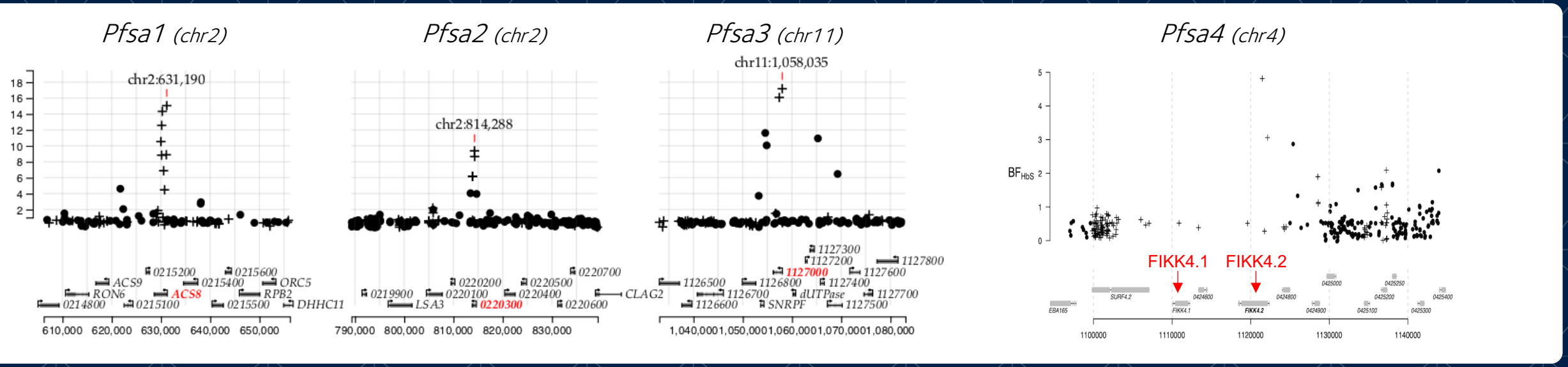
Evidence that FIKKs affect var2csa expression:

Davies et al 2023



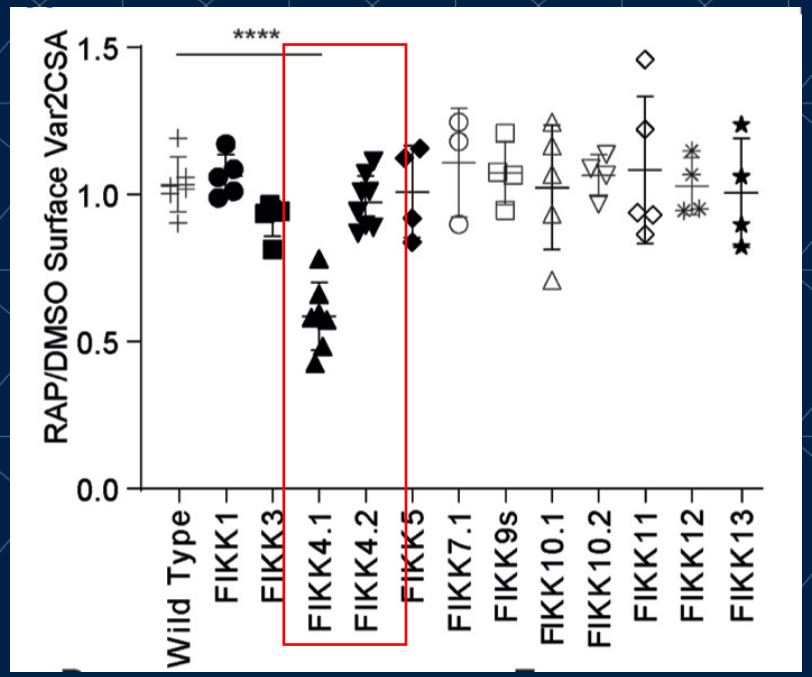
FIKK4.1/4.2 are exported Serine/threonine Kinases

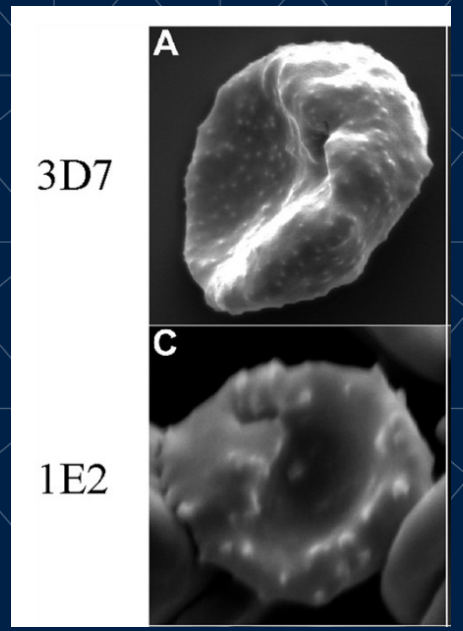
CENTRE *for* HUMAN GENETICS



Evidence that FIKKs affect var2csa expression:

Davies et al 2023



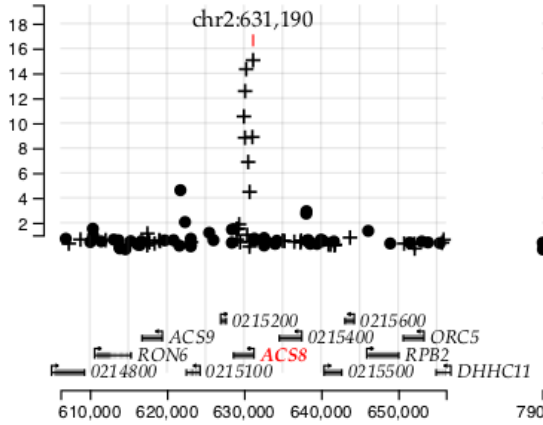


FIKK4.2 knockout

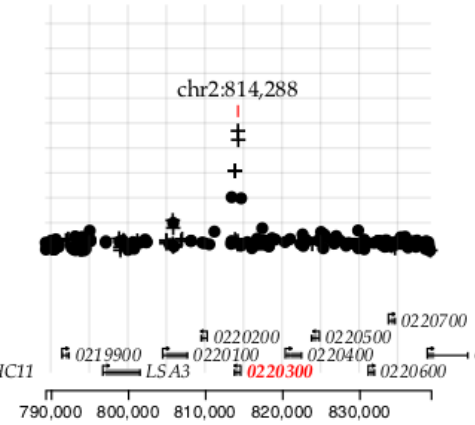
Kats et a 2014

CENTRE for HUMAN GENETICS

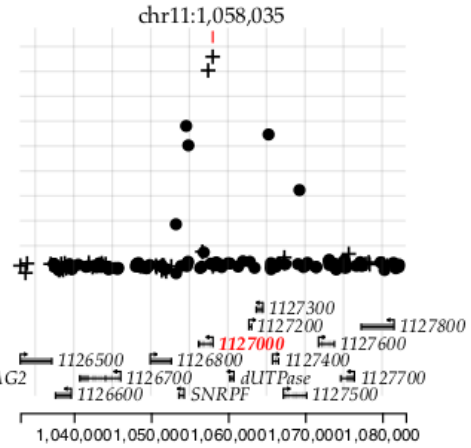
Pfsa1 (chr2)



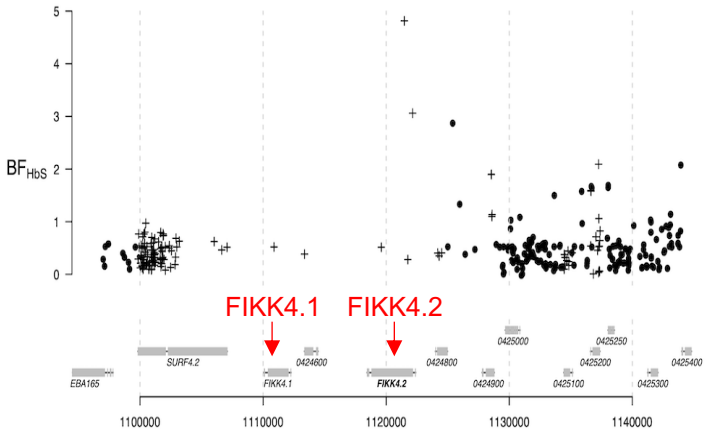
Pfsa2 (chr2)



Pfsa3 (chr11)

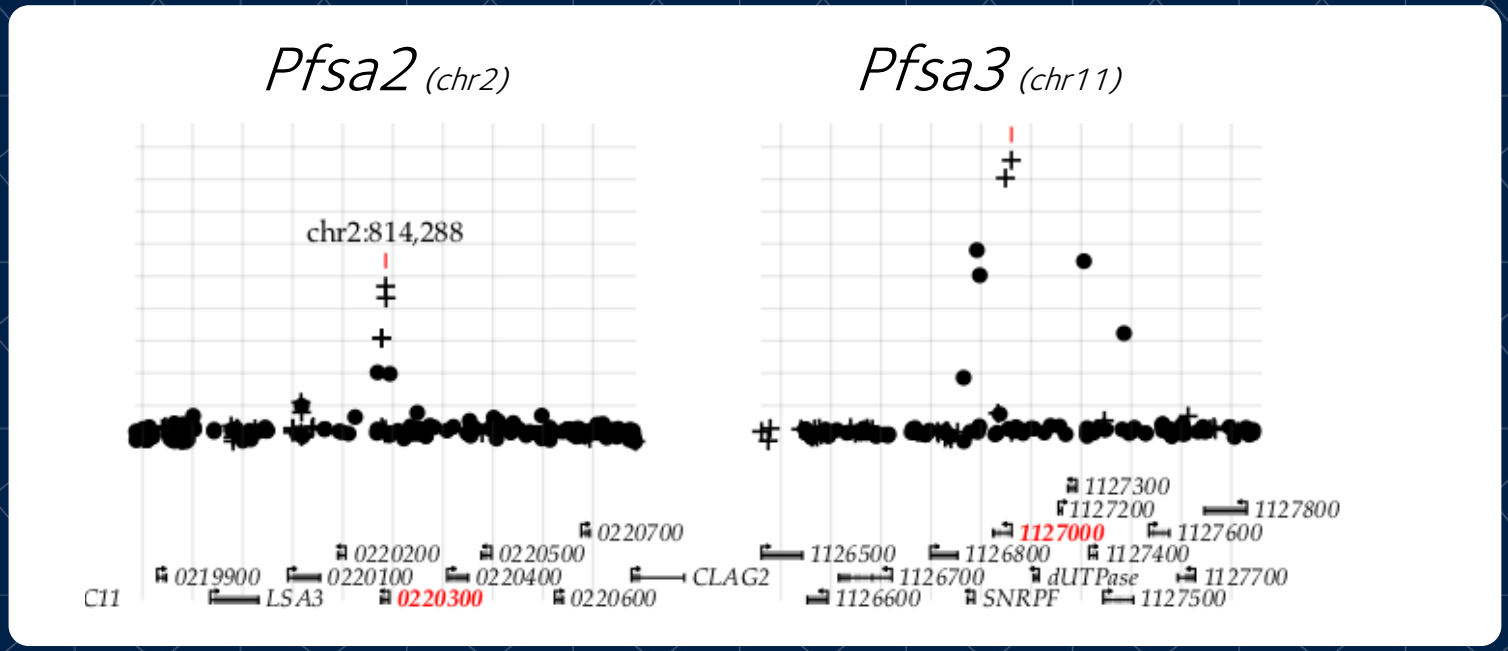


Pfsa4 (chr4)



Both ACSs and FIKKs are part of large gene families that have expanded in laverania
(The sub-clade of ape-infecting malaria that includes *P.falciparum*)

Otto et al, “Genomes of all known members of a Plasmodium subgenus reveal paths to virulent human malaria”, Nat. Microbiol 2018
Davies et al, “An exported kinase family mediates species- specific erythrocyte remodelling and virulence in human malaria”, Nat. Microbiol 2020

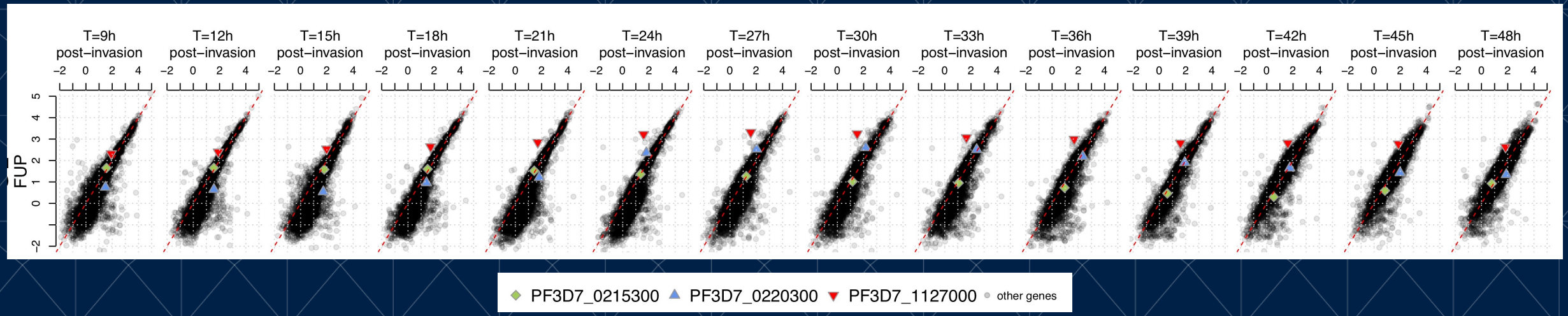


Both 0220300 and 1127000 contain PEXEL motifs

These mediate export via cleavage by Plasmepsin V

Possible hypothesis: do the *Pfsa* mutations affect export?

Locus	Gene identifier, base pair and amino acid location of PEXEL motif in Pf3D7	MOTIF sequence	*Pfsa* mutation	Codon change
Pfsa2	*PF3D7_0220300* Pf3D7_02_v3 814,307 - 814,321 AA: 49-53	RTLTE	814,288 C>T (+7)	GAT -> AAT (Aspartic acid -> Asparagine)
			814,329 A>G (-3)	ATA -> ACA (Isoleucine -> Threonine)
Pfsa3	*PF3D7_1127000* Pf3D7_11_v3 1,057,438 – 1,057,452 AA: 71-75	RCLNY*	1,057,437 T > C (+1)	AAA -> GAA (Lysine -> Glutamic acid)

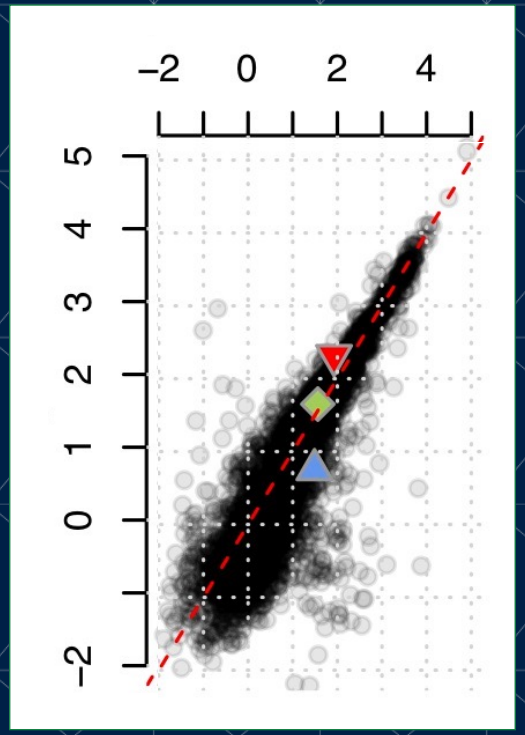


A different hypothesis: the *Pfsa+* mutations increase expression of Pfsa genes

Joseph Saelens, Mahamadou Diakité, Steve M. Taylor, mSphere 2021

Parasite gene expression (RNA levels)

sickle-associated (*Pfsa+*) parasite



Pfsa− parasite

T=9

Hours post-invasion

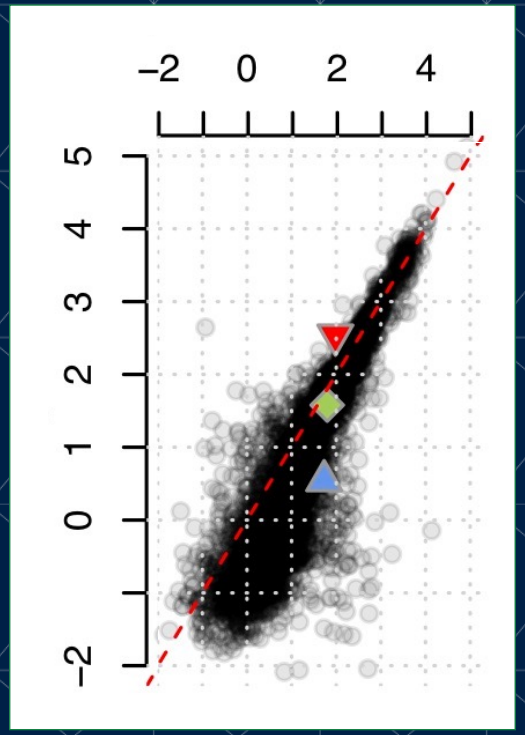
PF3D7_0215300 — PfACS8, *Pfsa1*

PF3D7_0220300 — *Pfsa2*

PF3D7_1127000 — *Pfsa3*

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



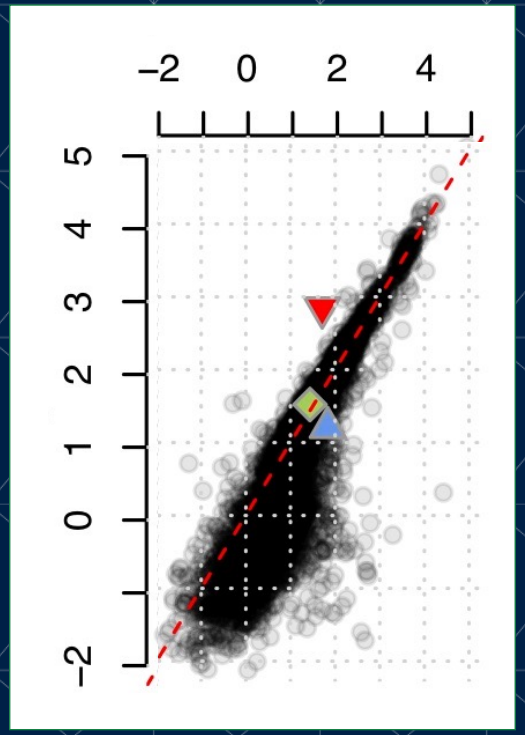
Pfsa− parasite

T=15

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated
(*Pfsa+*) parasite



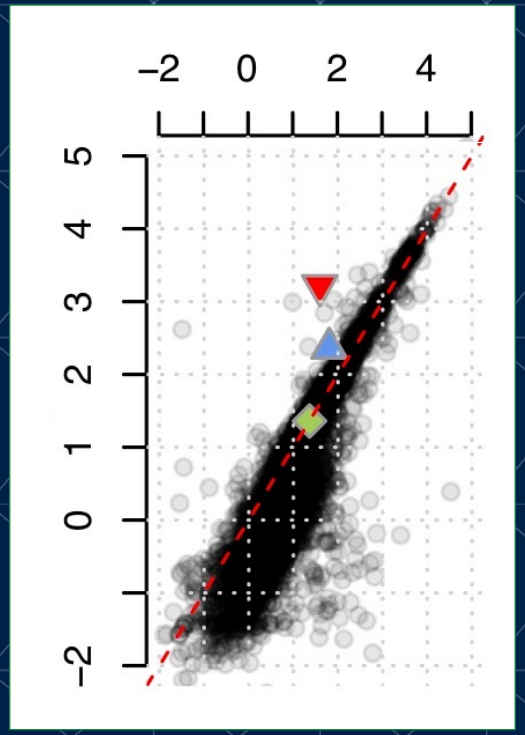
Pfsa− parasite

T=21

Hours post-invasion

▼ PF3D7_1127000 ● other genes

sickle-associated
(*Pfsa+*) parasite



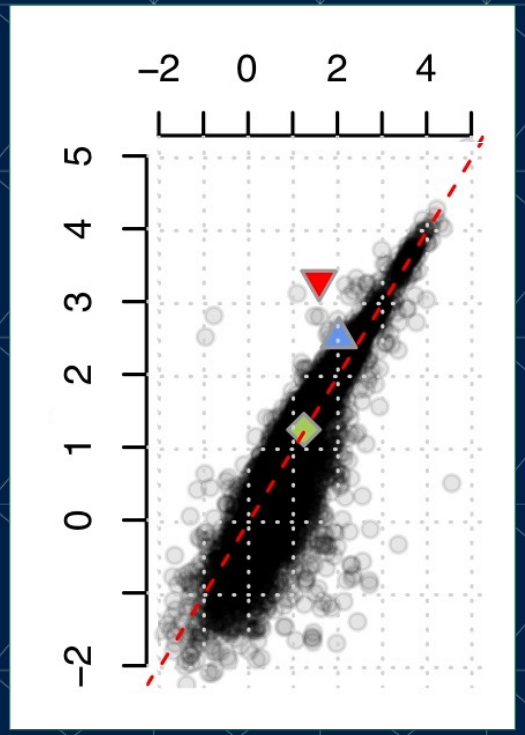
Pfsa− parasite

T=24

Hours post-invasion

▼ PF3D7_1127000 ● other genes

sickle-associated (*Pfsa+*) parasite



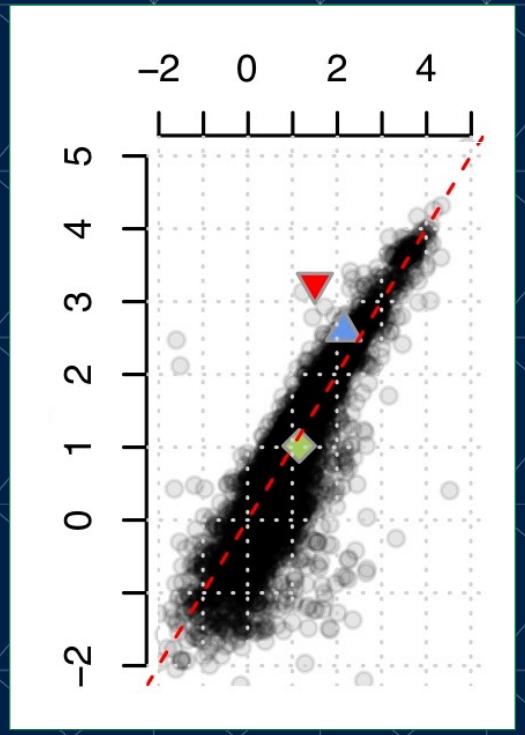
Pfsa− parasite

T=27

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated
(*Pfsa+*) parasite



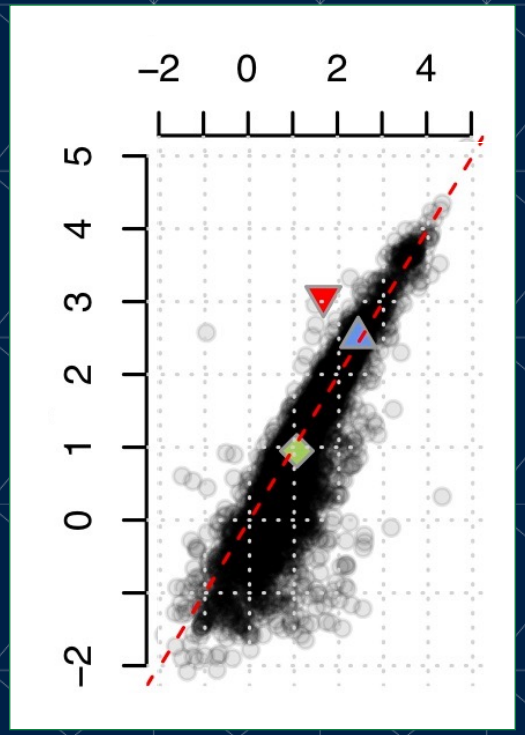
Pfsa− parasite

T=30

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated (*Pfsa+*) parasite



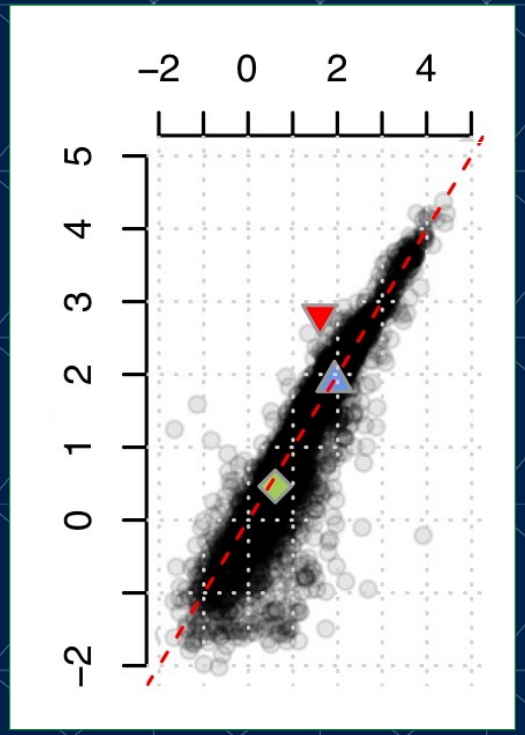
Pfsa− parasite

T=33

Hours post-invasion

PF3D7_1127000 other genes

sickle-associated
(*Pfsa+*) parasite



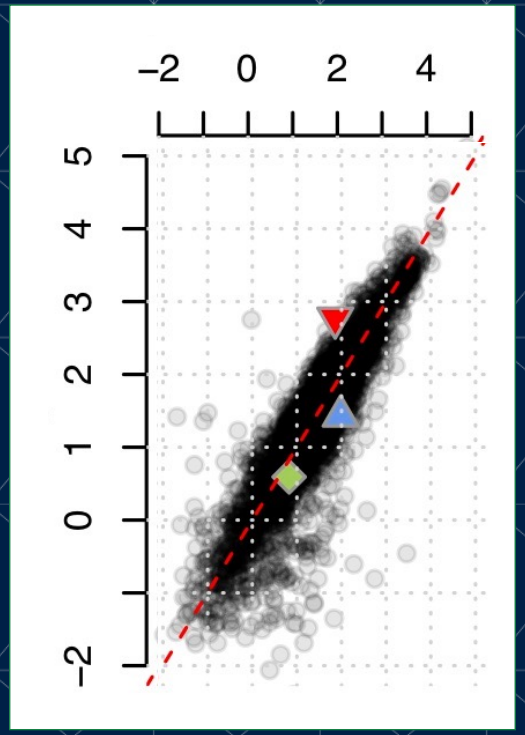
Pfsa− parasite

T=39

Hours post-invasion

▼ PF3D7_1127000 ● other genes

sickle-associated (*Pfsa+*) parasite

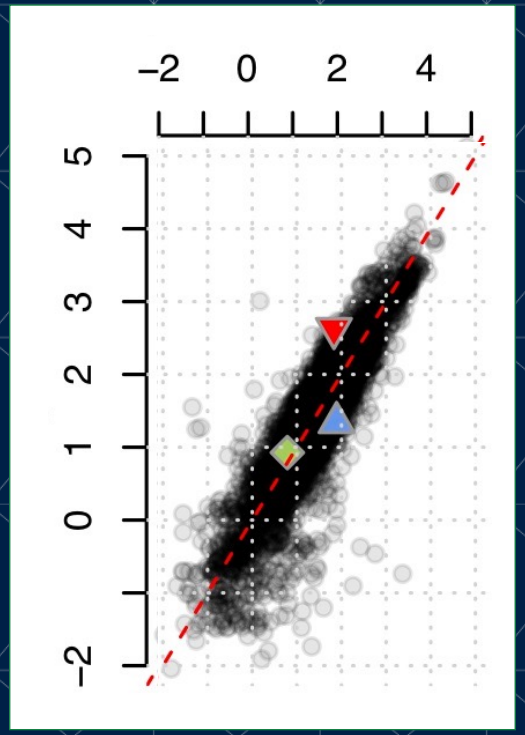


Pfsa− parasite

T=45

Hours post-invasion

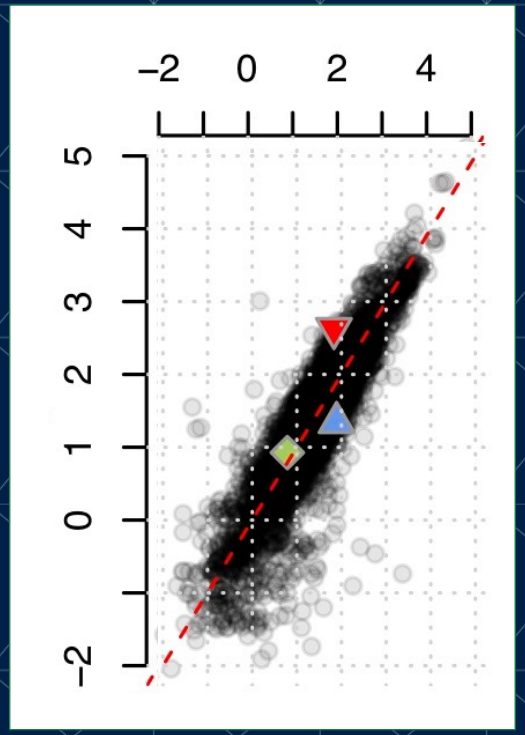
▼ PF3D7_1127000 ● other genes



T=48

Hours post-invasion

▼ PF3D7_1127000 ● other genes



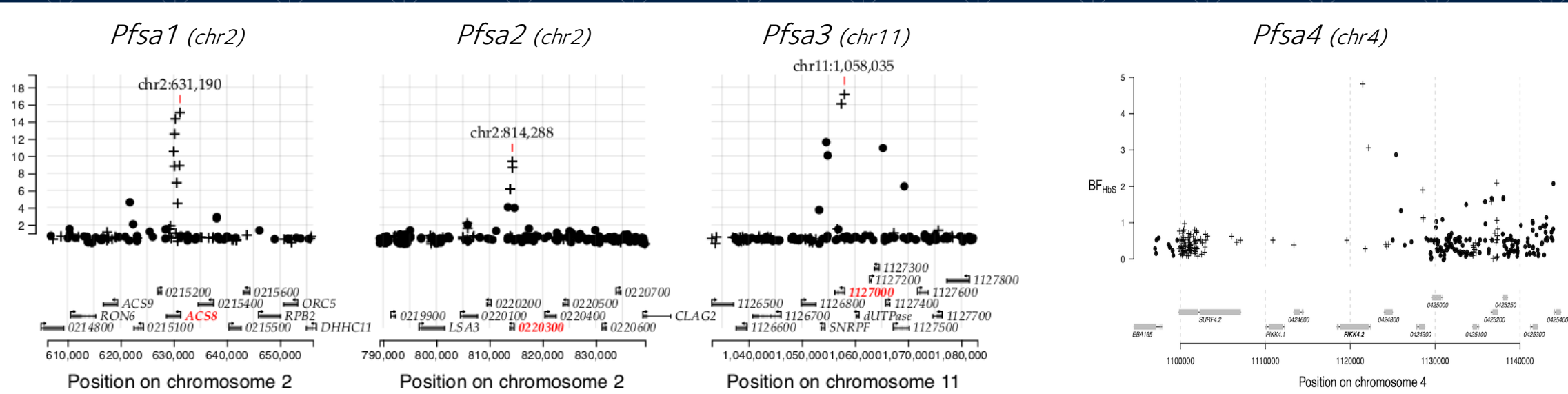
T=48

Hours post-invasion

PF3D7_1127000 other genes

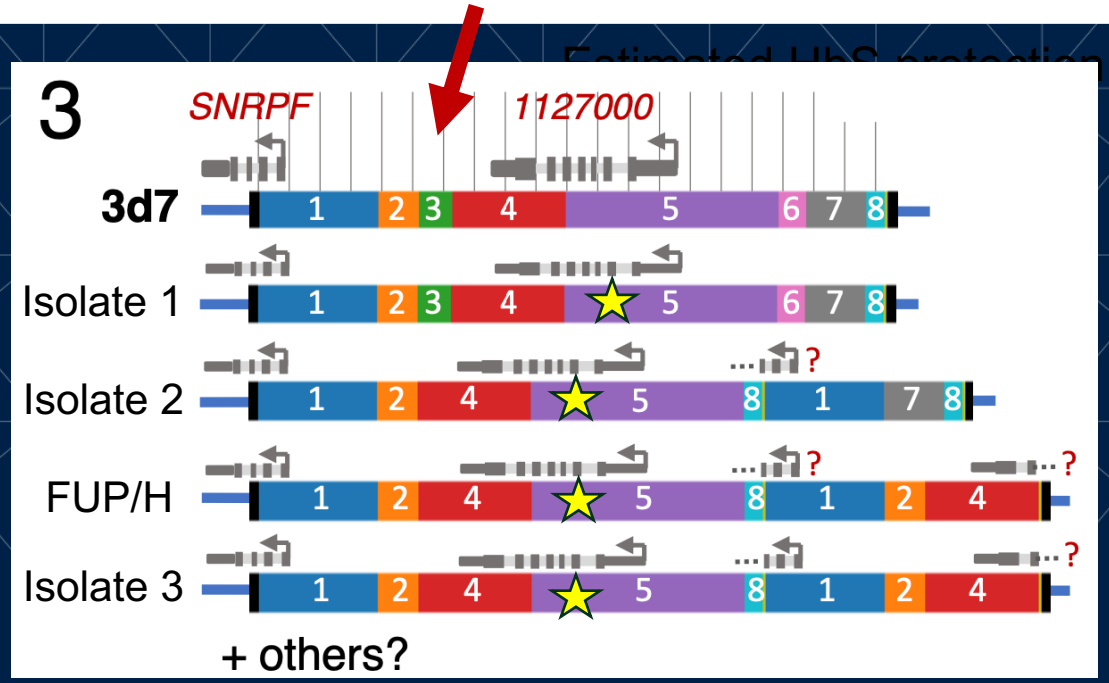
1127000 (Pfsa3) is over-expressed at Troph stage in Pfsa+ parasites.

Structural variation at *Pfsa3*



However…

The *Pfsa3*+ mutation (★) is linked to a set of structural variants.



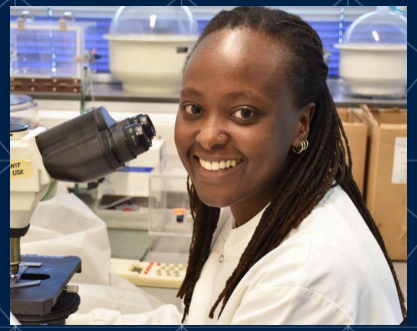
Annie Forster

Jason Hendry
Anna Jeffreys
Katie Healey

CENTRE *for* HUMAN GENETICS

Prelim work – MSIF funded

Tom Williams



Silvia Kariuki

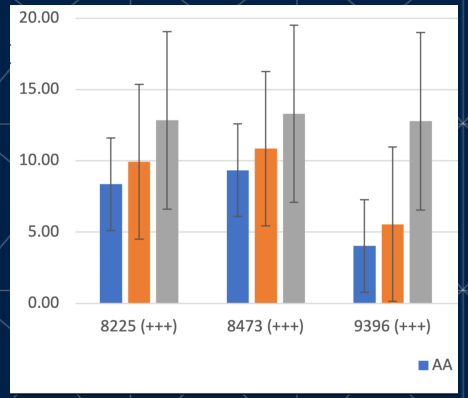
Alex Macharia

Patrick Ombati

Julian Rayner (CIMR)

Revive parasites from 30-year old frozen stocks, and conduct functional assays under controlled conditions.

James Docker, Katie Healey long-read assemblies



Prototype invasion expt.

Genetic association study work

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

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

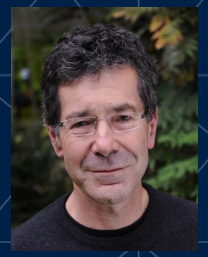
University of Utah:
Ellen Leffler

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

Duke University
Steve M. Taylor
Joseph P. Saeelens

USTTB Mali:
Mahamadou Diakite

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



Dominic Kwiatkowski
1953-2023

Parasite revival / functional work

KEMRI-Wellcome, Kilifi, Kenya
Alexander W. Macharia
Patrick Ombati
Silvia Kariuki
Julian Rayner

Pfsa+ structural variation work:
Annie Forster
Jason Hendry

Spatial mapping work: Andre Python

MIP data:
Deus Ishengoma
Robert Verity
Jeffrey Bailey

"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

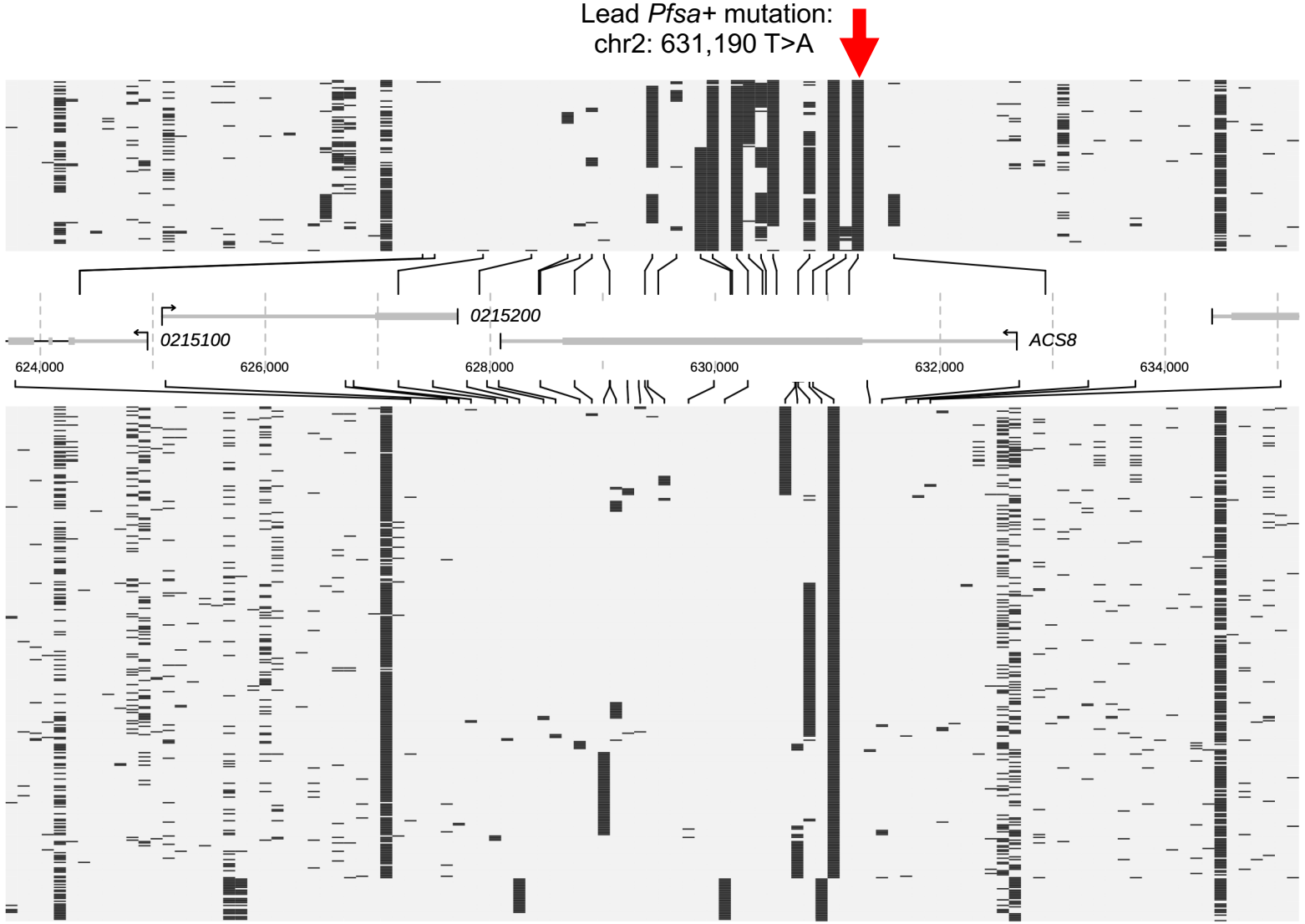
Thank you!

Pfsa1+ haplotypes:

(Observed in infections of *HbAA, HbAS, and HbSS individuals)*

Pfsa1- haplotypes:

(Largely only observed in infections of *HbAA individuals)*

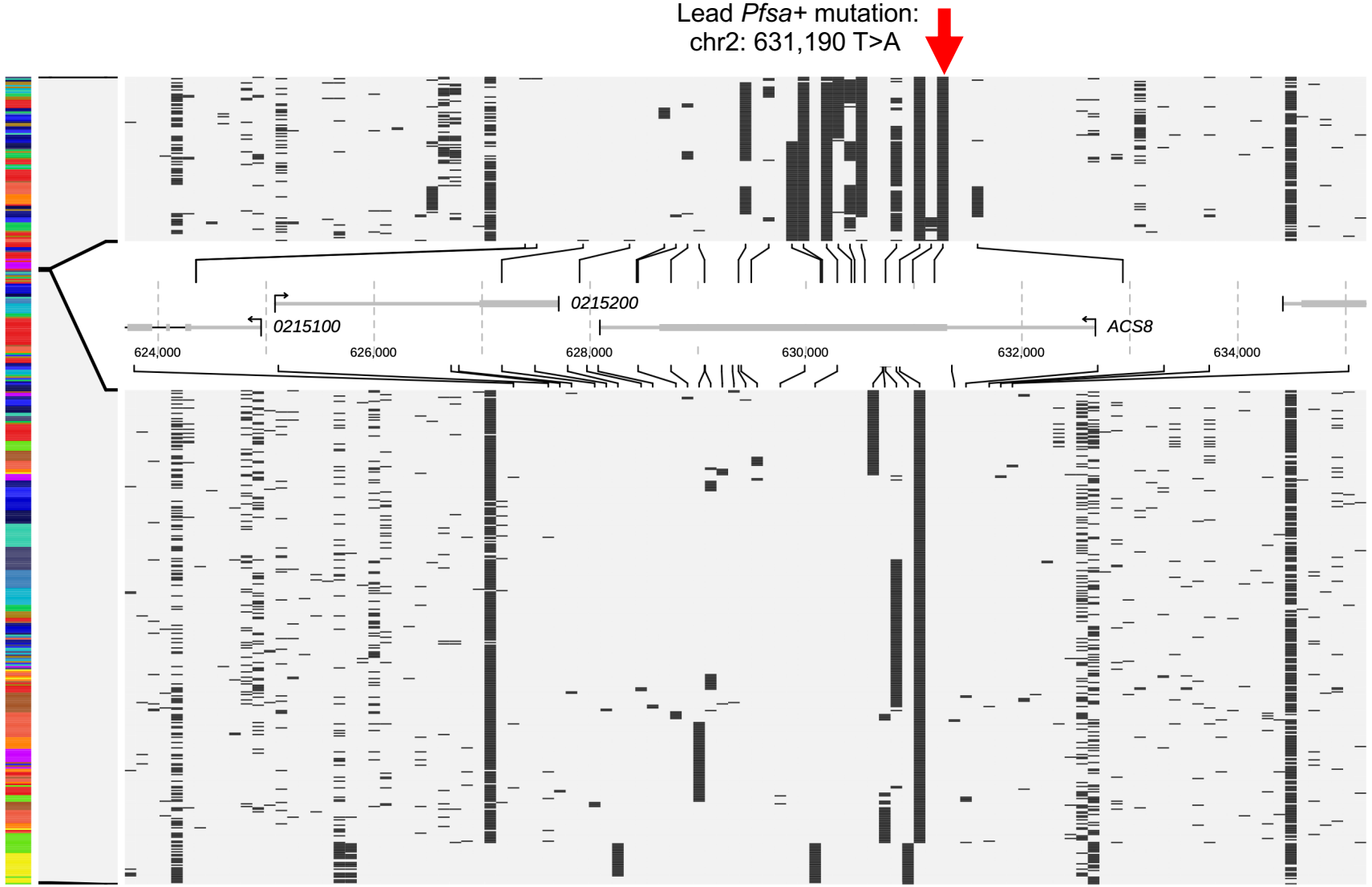


■ Ancestral allele ■ Derived (i.e. mutant) allele

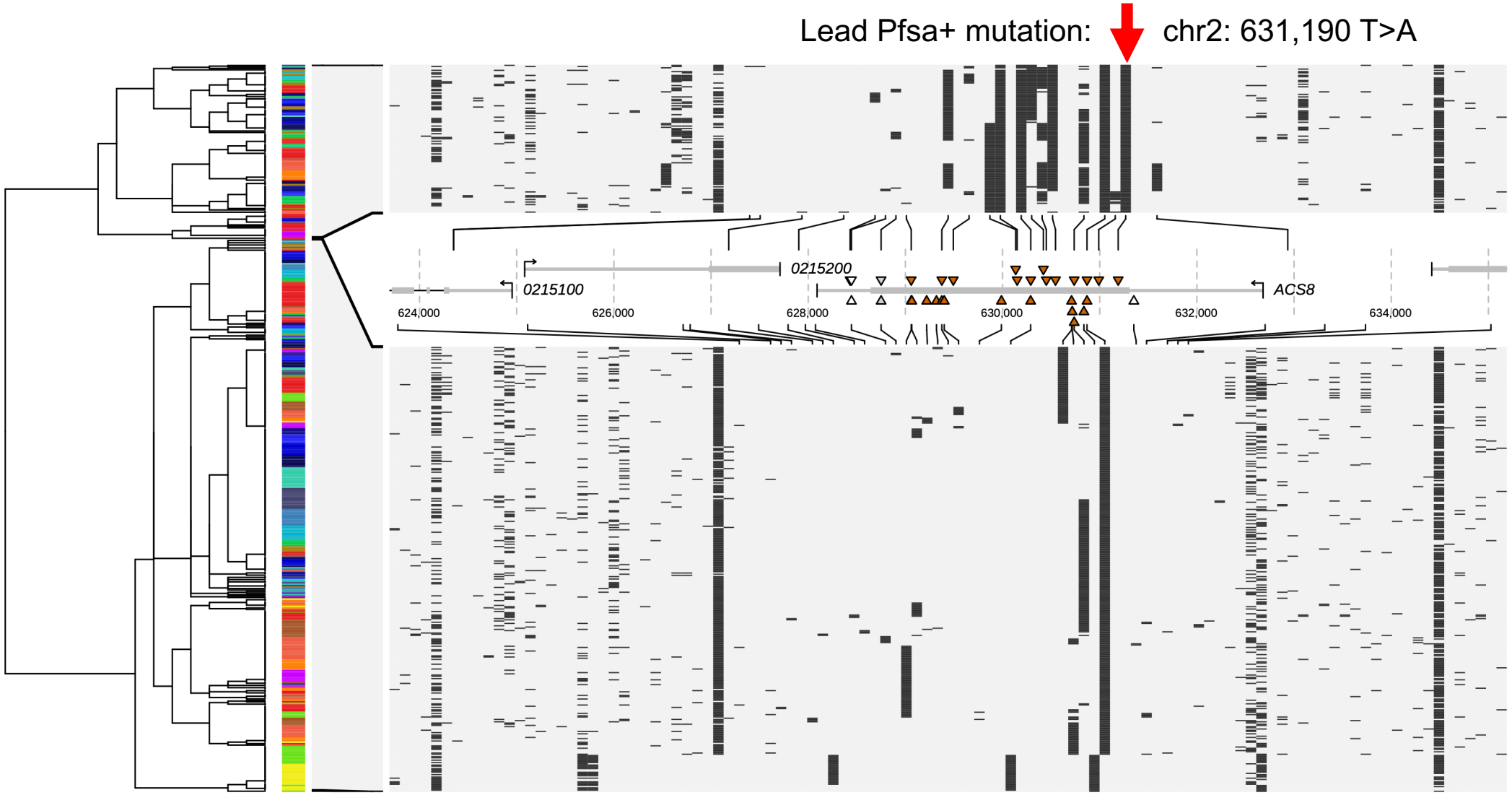
Mauritania
Gambia
Senegal
Guinea
Mauritania
Mali
Burkina Faso
IvoryCoast
Cote dIvoire
Ghana
Benin
Nigeria
Cameroon
Gabon
Congo DR
DRC
Sudan
Uganda
Malawi
Tanzania
Mozambique
Kenya
Ethiopia
Madagascar
Bangladesh
Myanmar
Laos
Thailand
Cambodia
Vietnam
Indonesia
PNG

Pfsa1+ haplotypes:

Pfsa1- haplotypes:

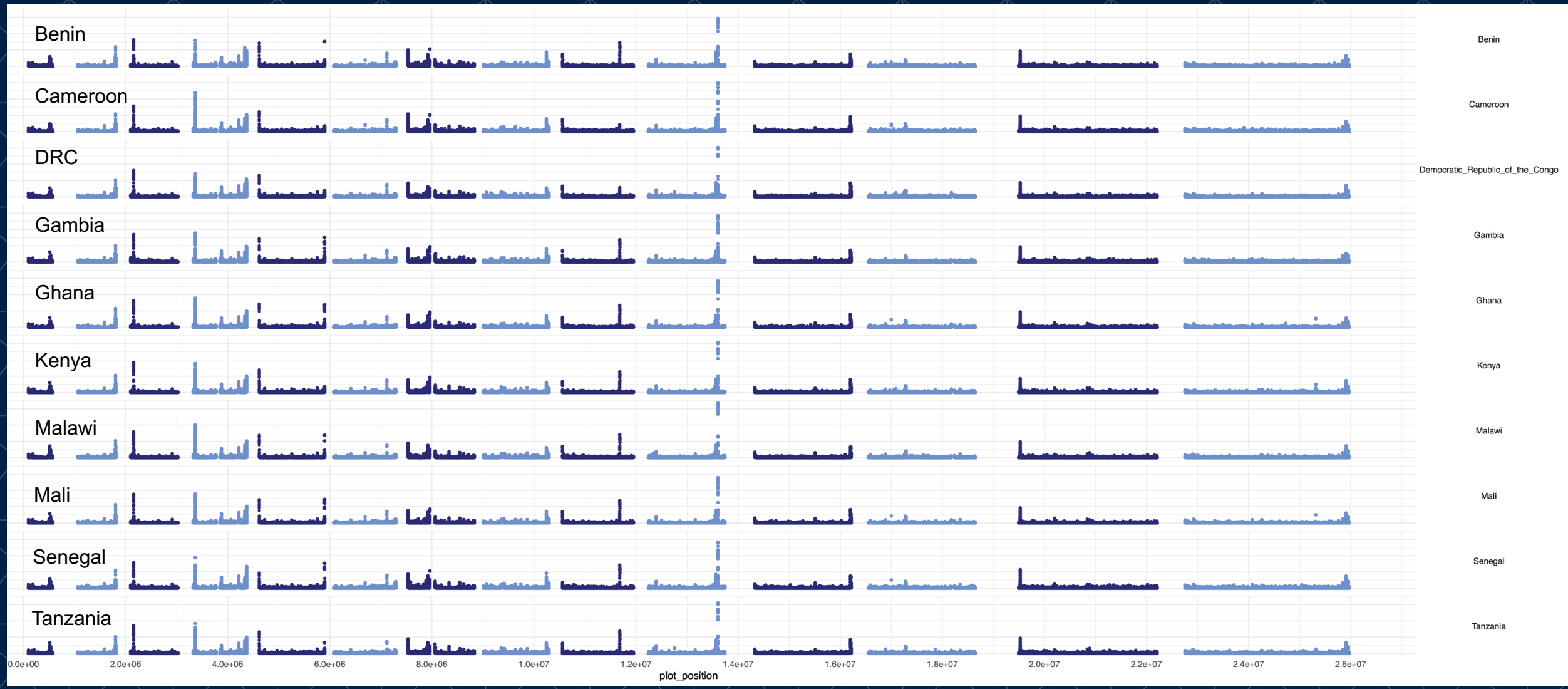


Ancestral allele Derived (i.e. mutant) allele



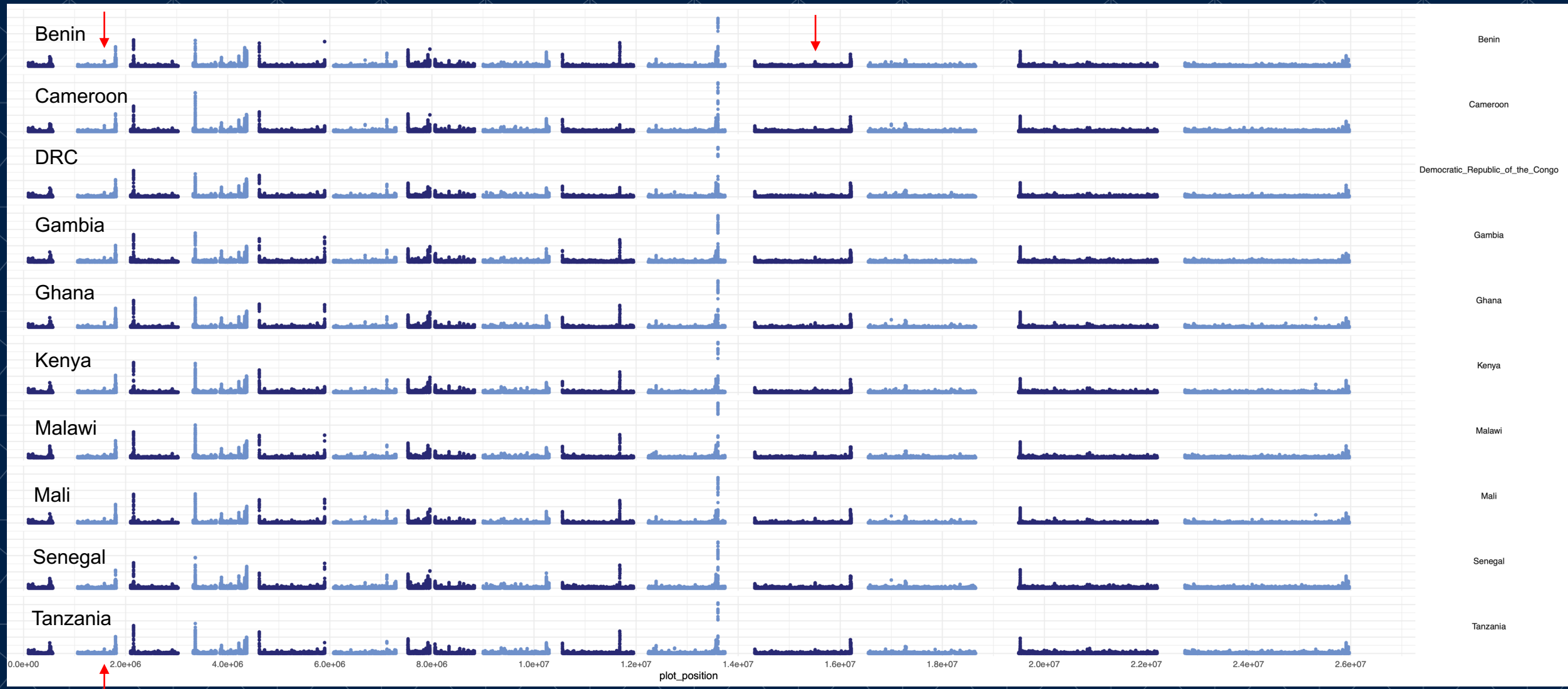
Ancestral allele Derived (i.e. mutant) allele

Genome-wide scan for balancing selection

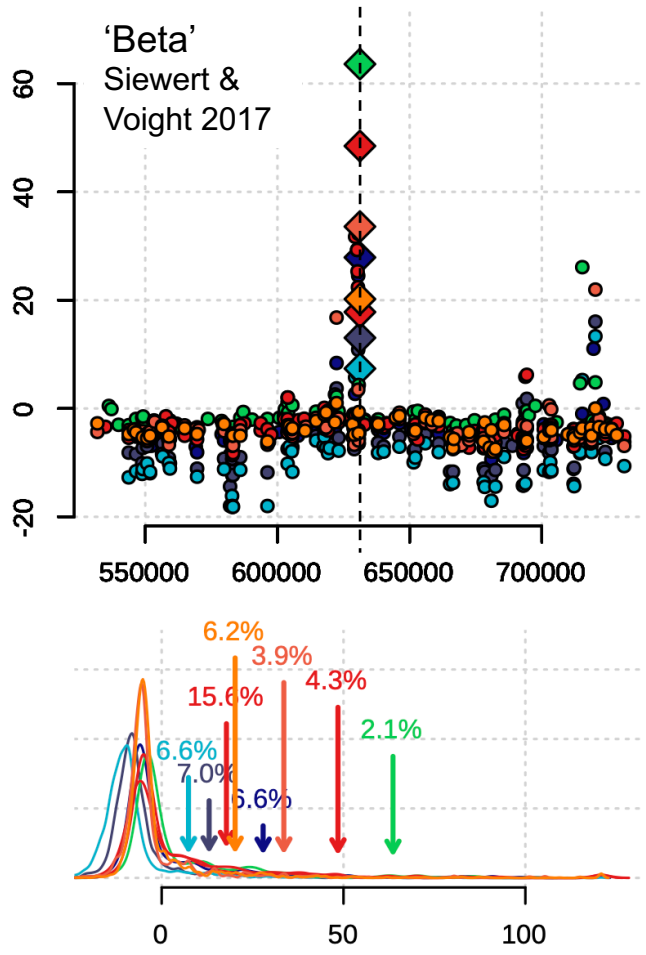


Pfsa1

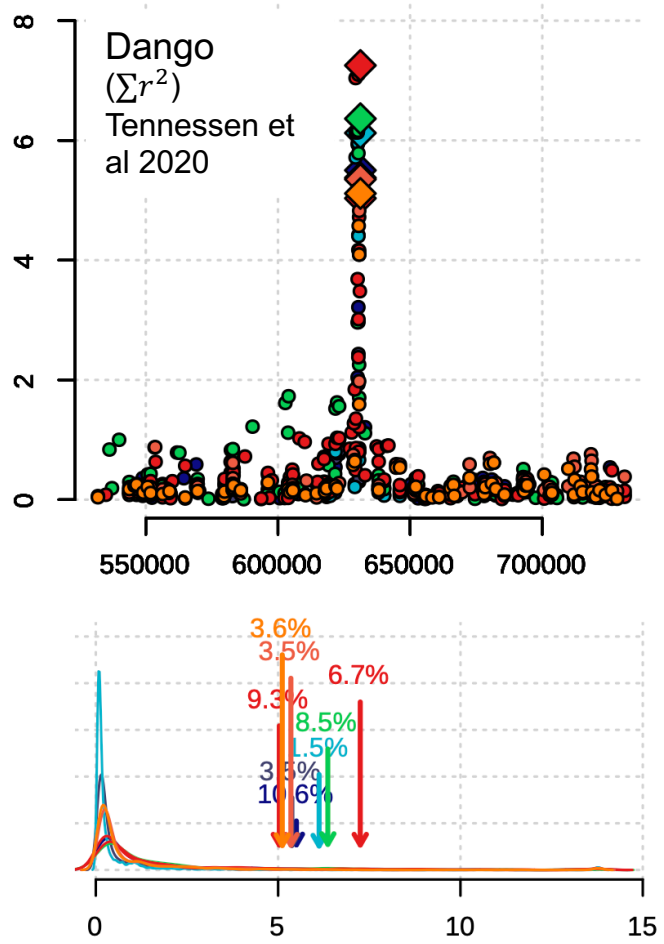
Pfsa3


Benin
Cameroon
DRC
Gambia
Ghana
Kenya
Malawi
Mali
Senegal
Tanzania
0.0e+00
2.0e+06
4.0e+06
6.0e+06
8.0e+06
1.0e+07
1.2e+07
1.4e+07
1.6e+07
1.8e+07
2.0e+07
2.2e+07
2.4e+07
2.6e+07
plot_position
Benin
Cameroon
Democratic_Republic_of_the_Congo
Gambia
Ghana
Kenya
Malawi
Mali
Senegal
Tanzania

Evidence for balancing selection



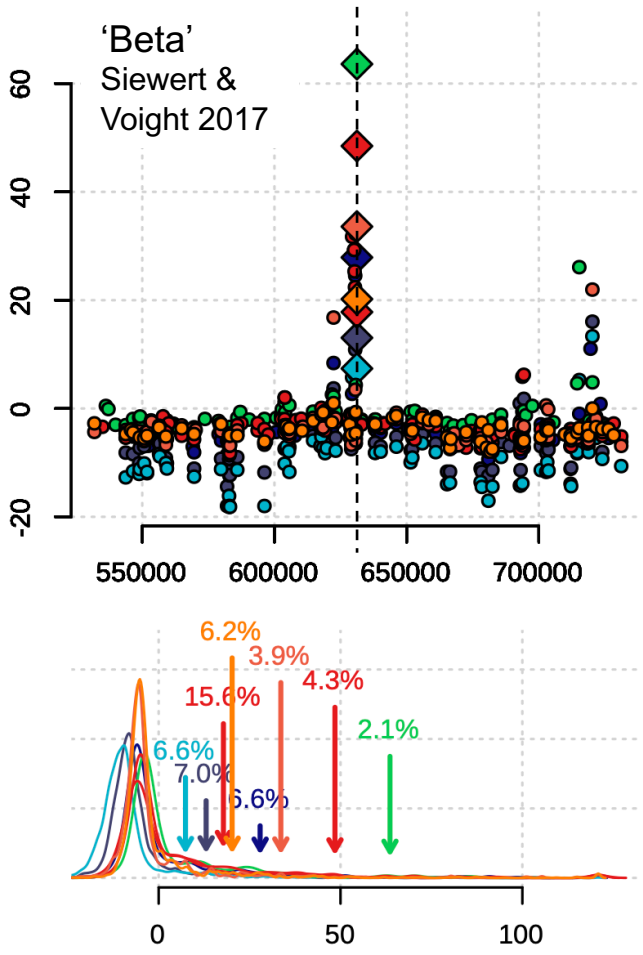
Distribution of genome-wide *Beta* at same frequency
(5kb window)



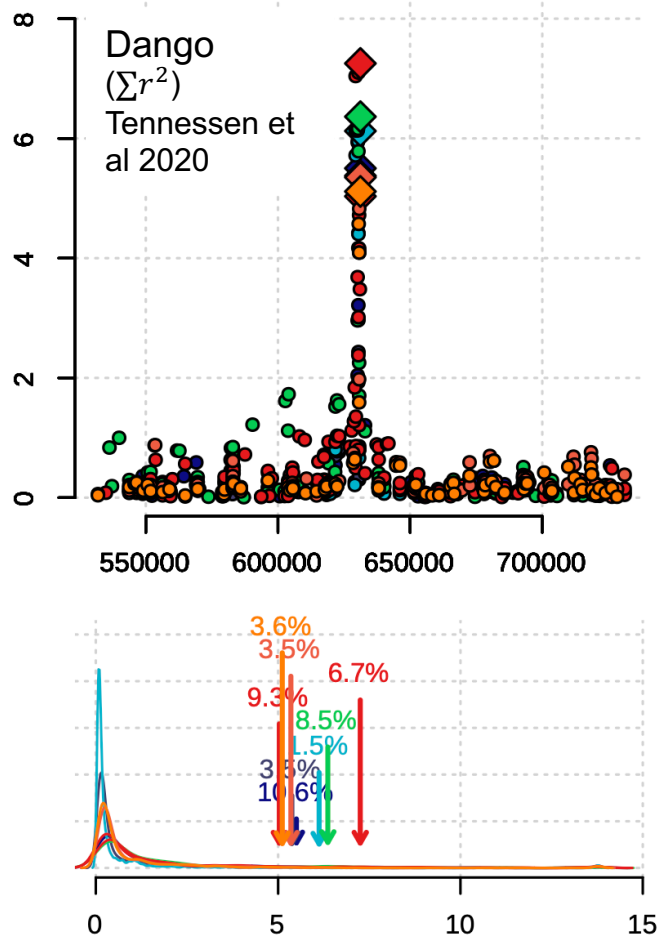
Distribution of genome-wide $\sum r^2$ at same frequency
(5kb window)

There is some evidence for balancing selection.

Evidence for… balancing selection?



Distribution of genome-wide *Beta*
at same frequency
(5kb window)



Distribution of genome-wide $\sum r^2$
at same frequency
(5kb window)