

Parallel geographic adaptation: One or many waves of advance?

Graham Coop and Peter Ralph

Evolution and Ecology

And Center for Population Biology

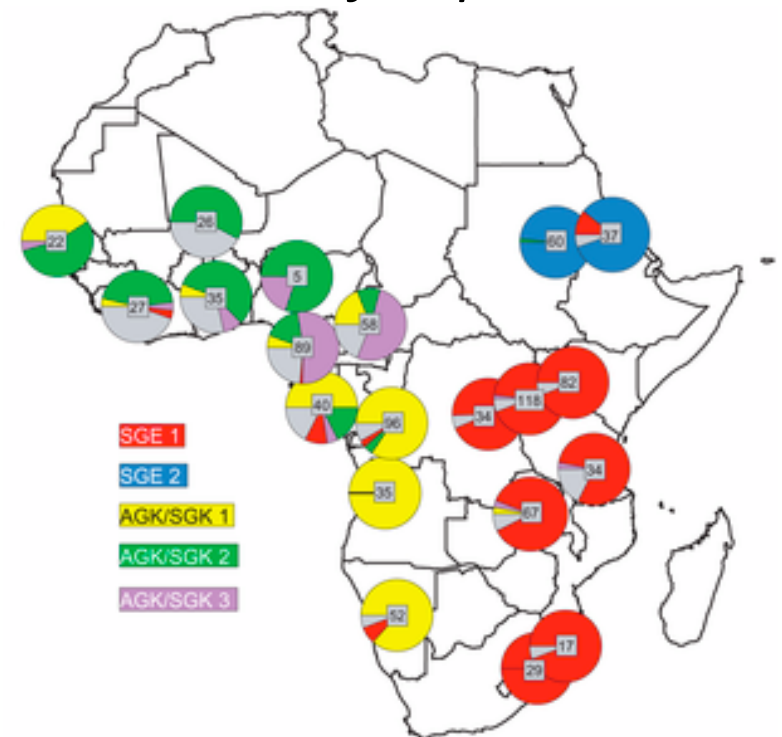
UC Davis



Parallel (convergent) origins of selective alleles

- hypothesized in e.g.
- In insecticide resistance
 - (aphids, *Drosophila*, whiteflies, *Tribolium*)
- In pigmentation
 - (cavefish, pocket mice, *Drosophila*)
- Often, changes at a single base

Drug resistance alleles
at the
dhps locus in
P. falciparum



Pearce et al 2009

Examples of Parallel adaptation in humans

E.g. mutations at Lactase (Tishkoff et al.), have occurred >3 times

Light skin pigmentation: convergent genetic basis between E. Asian and western Eurasia (Norton et al 2007, Edwards et al 2010). SLC24A5 vs OCA2

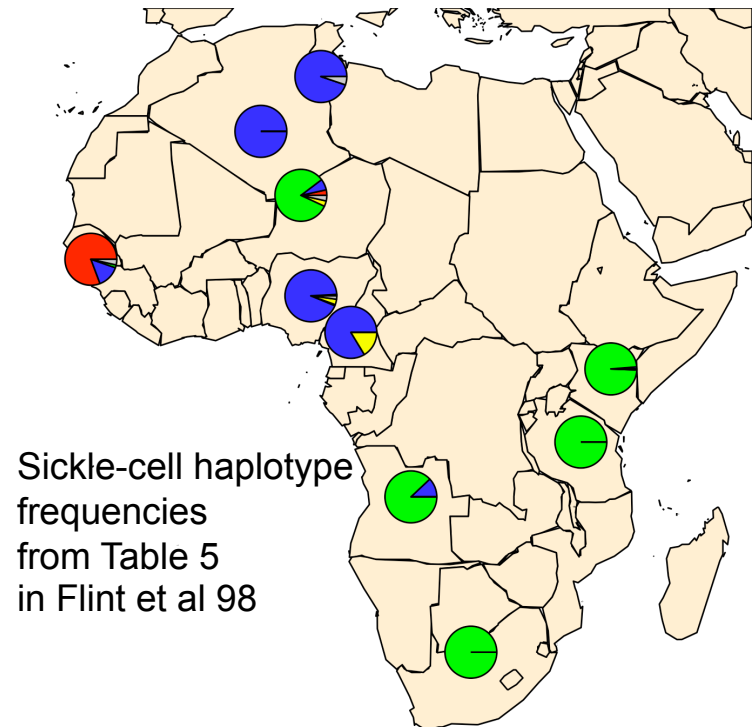
various malaria resistance alleles: G6PD alleles, sickle-cell

Sickle cell haplotypes in Africa

The human sickle-cell allele found on multiple distinct haplotypes in different geographic ranges

- Consistent with the allele having arisen multiple times
- Other explanations include gene conversion (Flint et al '98)

A 2nd malaria resistance allele, HbC, is found at the same codon in Africa.



Geography and Parallel adaptation

Chance of parallel mutation should be higher in species with large habitats, with high mutation rates, and low dispersal distances

But: Only in species, and between populations, with strong neutral population structure?

Expect more parallel mutation when selection is weak or strong?

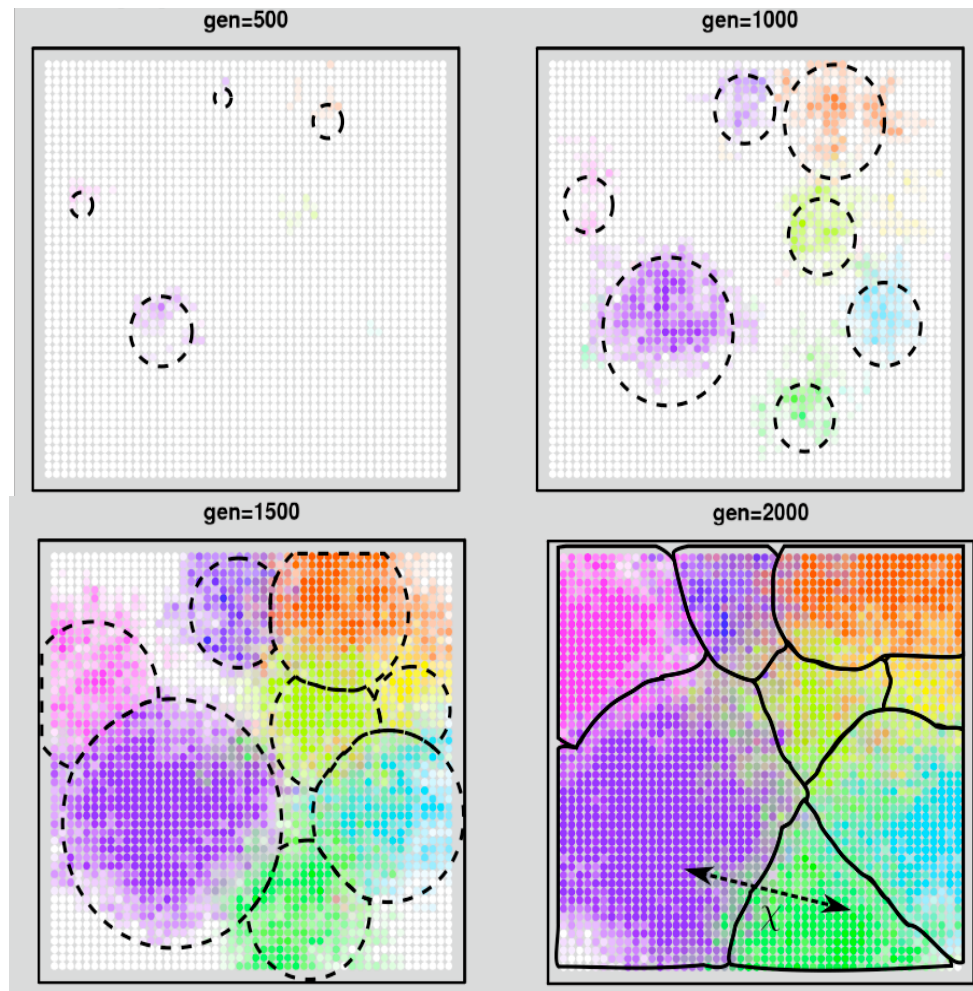
Are parameters in humans, and other species, suitable?

Pennings and Hermisson (2006) investigated parallel mutation in a panmictic population, termed these soft sweeps

Geographic parallel mutation assumptions:

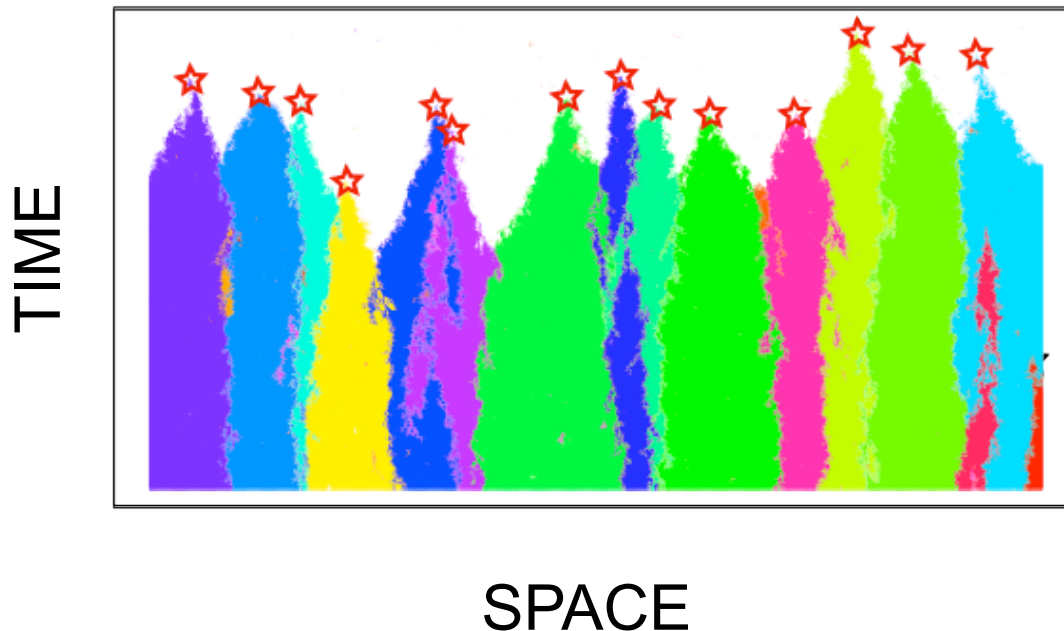
- New mutations
- Selection pressure is geographically uniform
- Parallel selected mutations are selectively equivalent (e.g. same base pair, gene, or on same pathway) and so spatially exclude each other

Simulation:
2d island
model grid
4 time points



The process in more detail

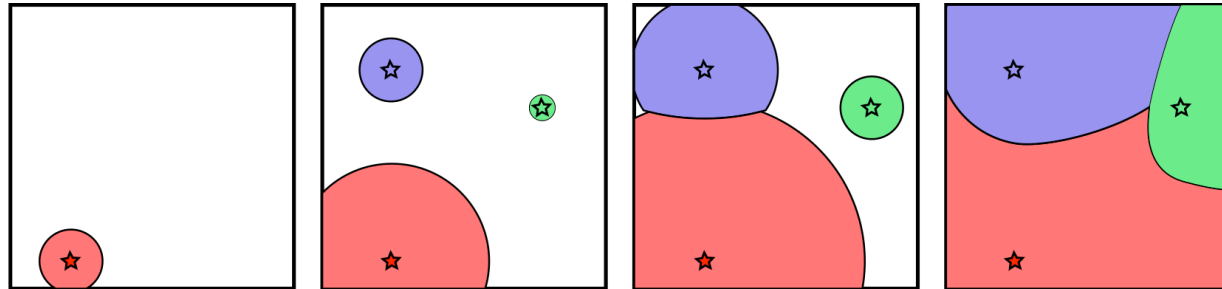
- ▶ Mutations rare (μ); those escaping drift rarer ($\mu 2s/\xi^2$)
- ▶ *Rare* \Rightarrow a **Poisson process** (of rate $2\rho \times \mu 2s/\xi^2$)
- ▶ *Successful* mutations spread outwards radially at speed $v = \sigma\sqrt{s}$ (F; KPP 1937)
- ▶ Mutations **exclude** each other (in initial patchwork)



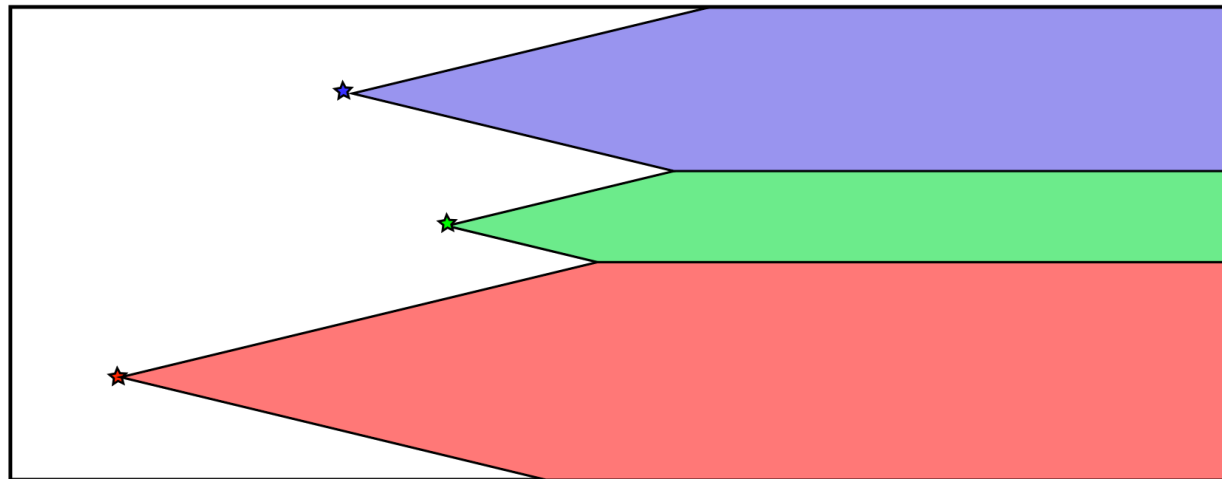
s = selection coefficient
 ρ = pop density
 μ = mutation rate
 σ = SD dispersal distance
 ξ = SD # of offspring
 d = spatial dimensions (1 or 2)
 $\omega(d)$ = a constant

This model has been studied before: (Kolmogorov-)Johnson-Mehl model of crystallization. Crystals appear at new nucleation sites, and spread, at constant rate, until they run into each other

2d population
4 time points



1d population
Space-time
diagram



time →

Møller (1992) studied various properties of crystals.

We have extended Moller's results to:

A population genetic setting

Non-constant wave speeds (fat tailed long-distance dispersal)

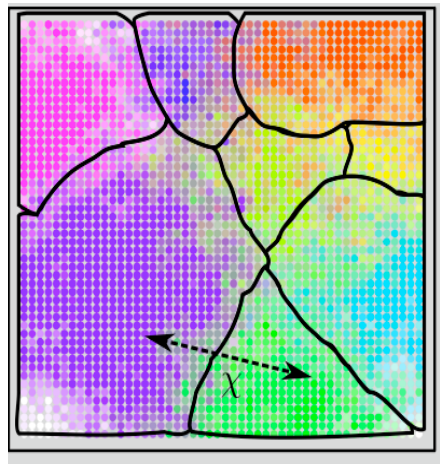
Stochastic spread of waves.

Rule of thumb: a characteristic length

Summary of properties: **characteristic length**

$$\chi = \left(\frac{\sigma\xi^2}{\rho\mu\sqrt{2s}\omega(d)} \right)^{1/(d+1)}$$

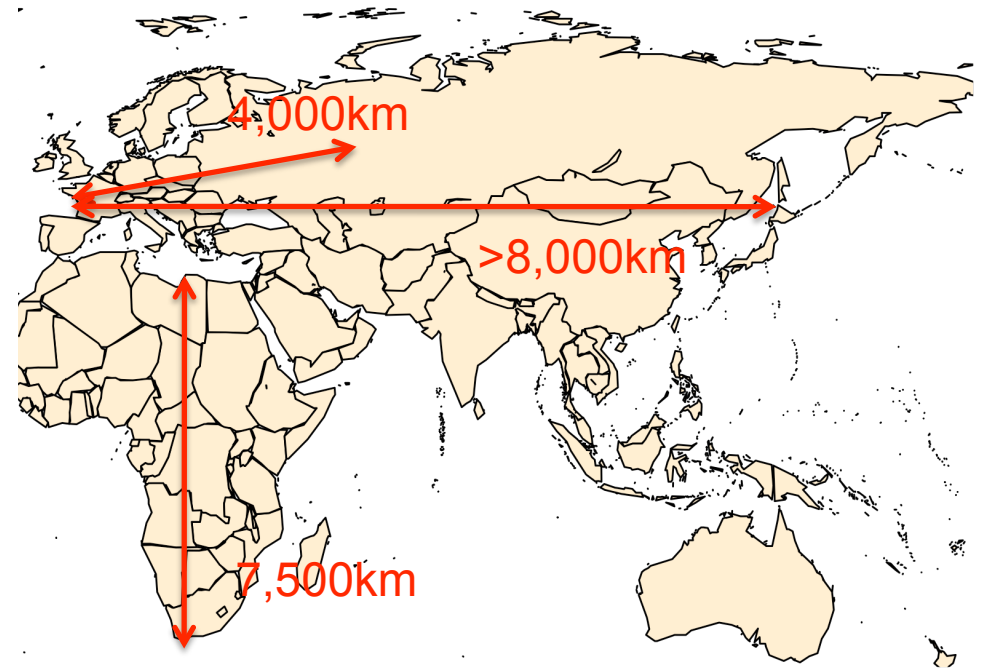
the distance spread by a selected allele before expected to encounter another.



s = selection coefficient
 ρ = pop density
 μ = mutation rate
 σ = SD dispersal distance
 ξ = SD # of offspring
 d = spatial dimensions (1 or 2)
 $\omega(d)$ = a constant

If range is **large relative to** χ ,
parallel mutation is very **likely**.

- population densities: $\rho = .002 = 10000 / (\text{area of Europe})$.
 $\rho = 2\text{km}^2$ human population density $\sim 5,000$ years ago



Pop density	Mut. rate	
ρ	μ	Gaussian
2.000	10^{-8}	2240.9649
0.002	10^{-8}	22409.6485
2.000	10^{-5}	224.0965
0.002	10^{-5}	2240.9649

Dispersal std. dev = $\sigma = 100\text{km}$, selection coeff. = $s = 1\%$

Parallel adaptive mutations likely when width of population's range $>$ characteristic length

Suggesting a mutational target size $\sim 1000\text{bp}$ is sufficient

These conclusions are very rough as obviously human history/migration is complicated

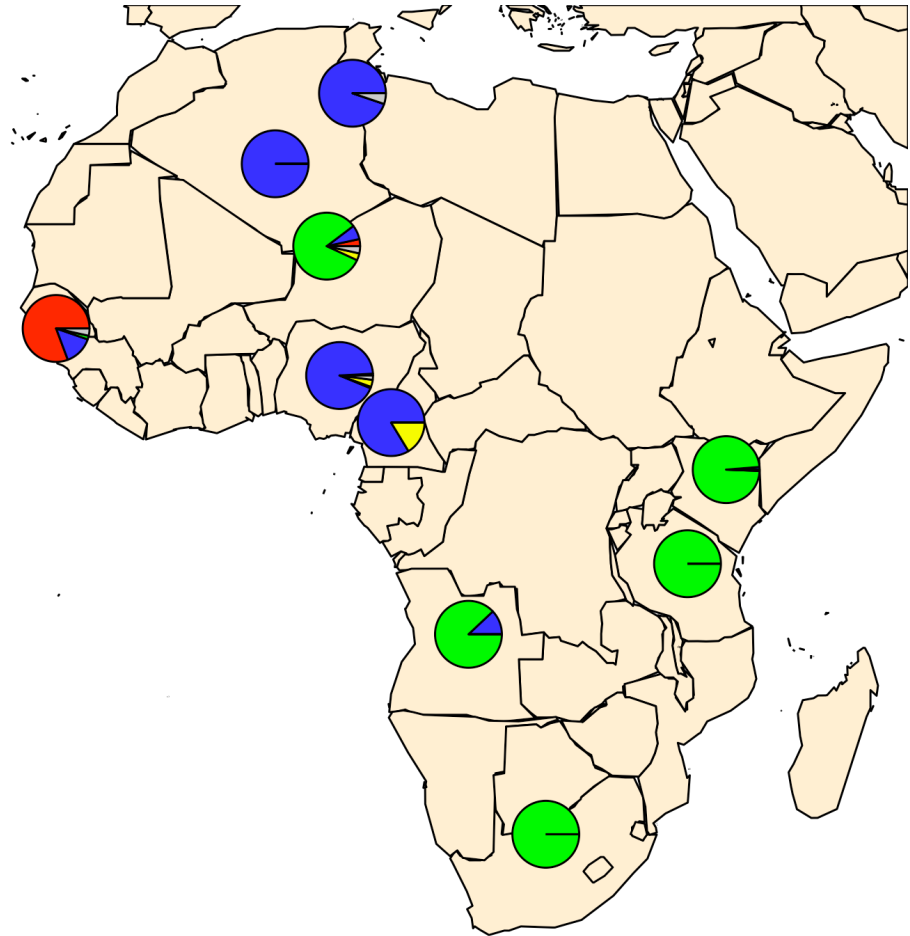
Sickle cell haplotypes in humans

Different (putative) origins of sickle-cell mutations separated by a few 1000km

For an $s=5\%$ (Currat et al 2002) and $\mu=10^{-8}$

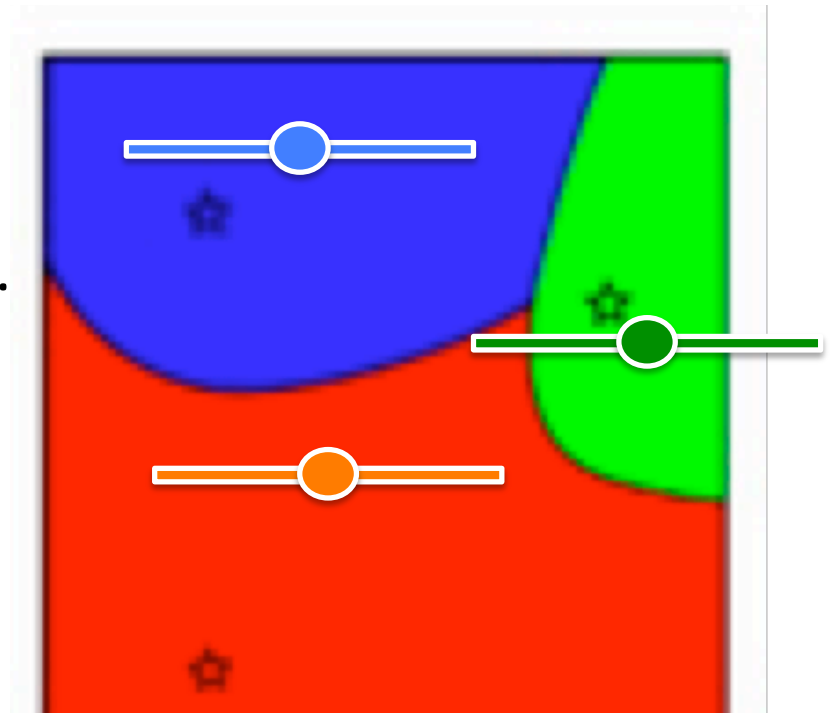
A $\sigma=10\text{km}$ and $\rho=2 \text{ people/km}^2$ give
A characteristic length $\sim 1000\text{km}$

These parameters do not seem too unrealistic for 5-10 k years ago.



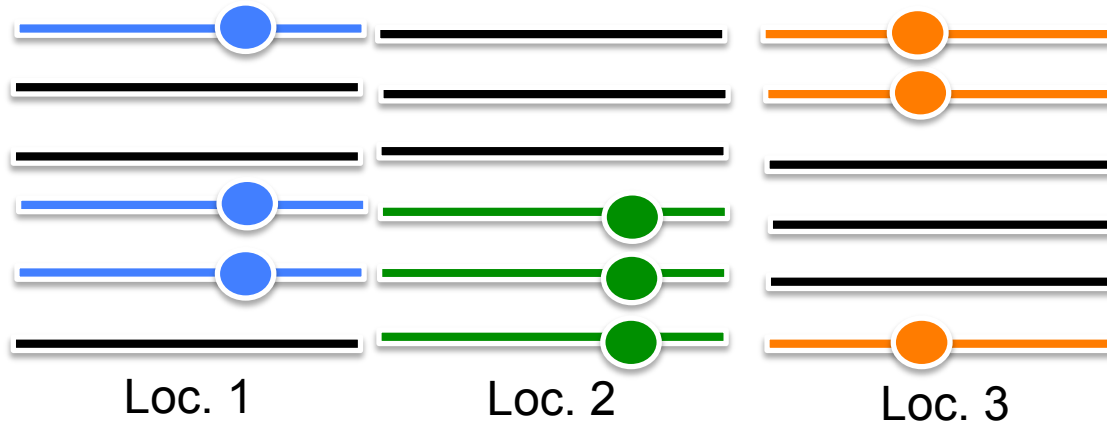
Sickle-cell haplotype info from Flint et al 98

- relatively 'sharp' domains formed by mutations
 - may be mistaken for local adaptation.
 - Especially if the parallel mutations occur at unlinked loci.
- The edges between selected types will mix slowly by migration and drift.



After mixing:

If mutations at unlinked loci
-A set of partial sweeps results



If mutations all at one locus
-A soft sweep results



Is strong structure needed to create parallel adaptation?

- A lack of strong structure is not telling you that gene flow is fast, only that it is fast compared to local drift.
- Thus species in with little neutral structure at neutral markers, parallel mutation may be common simply because there population density in high.

Conclusions

- Dispersal rates are key to our interpretation of population genomics signals of selection.
- Geographic parallel mutation is likely common
 - Predict many more will be found as more traits are dissected.
 - Are full sweeps the rare exception?
- The signal of such parallel sweeps may be quite complex.

Thanks!

- Peter Ralph
- Sebastian Schreiber, Dave Begun, Chuck Langley, John Novembre, Michael Turelli
- paper available at
- <http://web.eve.ucdavis.edu/plralph>
- Funding: Sloan Foundation Fellowship

