ROLE OF CONVERGENT EVOLUTION AND STANDING VARIATION IN LOCAL ADAPTATION

Graham Coop

Evolution & Ecology,
Center for Population Biology,
UC Davis
@graham_coop
www.gcbias.org

Evolution
June, 2014

A PATCHY APPROACH TO LOCAL ADAPTATION

Species often repeatedly adapt to environments that are patchily distributed across their range.

Geography: At what spatial-scale should we expect adaptations to be shared across landscape?

How common is convergent evolution within a species?

What is the role of standing variation in adaptation to discrete patches?

THEORY ON THE CONVERGENT BASIS OF ADAPTATION TO PATCHY ENVIRONMENTS.

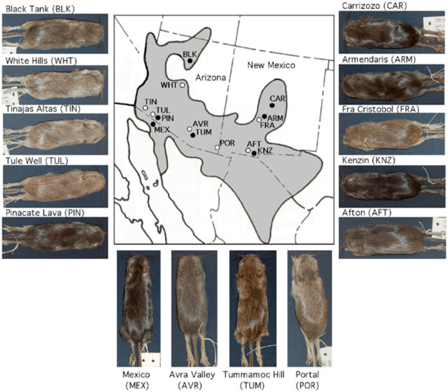
note motivation from: Pennings & Hermisson, messer & petrov (Soft Sweeps), Ralph and Coop 2010 (geographic parallel sweeps).

MOTIVATION: A TAIL OF TWO PATCHES.

image: Hoekstra

- Dark-pigmented mammals and reptiles on volcanic outcrops in the Southwest. (Dice '40, Benson '33)
- "Dark" pigmentation beneficial on outcrops, deleterious elsewhere.

Motivation: A tail of two patches.



- MC1R know to control dark pigmentation in Rock Pocket mice on one outcrop
- but this genetic basis is not shared with mice on other outcrops

(Nachman et al 2003)

PATCHY ENVIRONMENT

Focal allele is beneficial in patches; deleterious between:

A = patch size

R = patch separation

s_b = on-patch selective advantage

$-s_d$ = between-patch selective cost

ρ = pop density

μ = mutation rate

σ = SD dispersal distance

What is the time scale of:

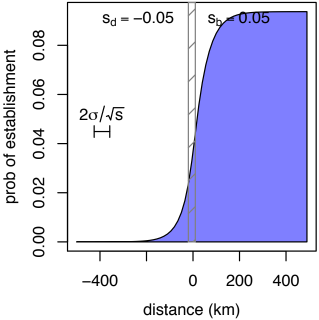
- appearance of new mutations in unadapted patches?
- transit of mutations between patches by migration?

What does the latter look like?

ADAPTATION BY NEW MUTATION

Mutations arise at rate $\rho\mu$. Prob. of establishment depends on where they arise. Inside patch $p_{\text{estab}} = 2s_b$
Approximation: if our patch is of size A the rate of influx is

$$\lambda_{\text{mut}} = A\rho\mu 2s_b \quad \text{per generation.}$$

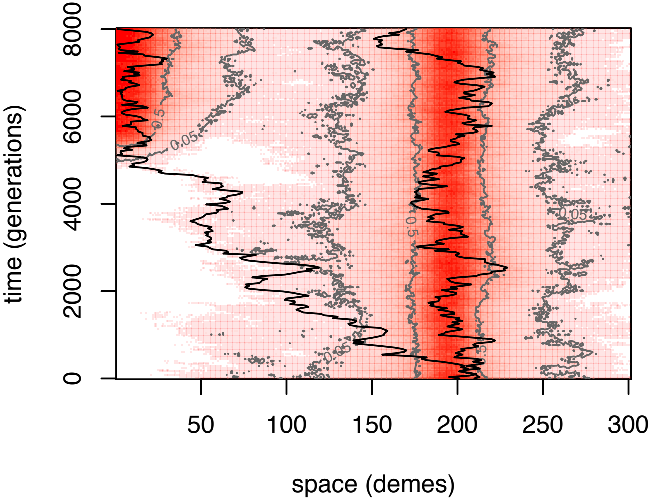


Outside patch, at distance r is $p_{\text{estab}} \simeq \exp(-r\sqrt{s_d}/\sigma)$ (Barton 1987)

ADAPTATION BY MIGRATING ALLELES

Suppose a mutation has fixed locally in one patch.

How long until it reaches another at distance *r*? (Slatkin 1973)

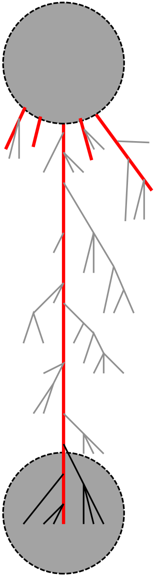


MODELING TRANSITING ALLELES

General idea:

- Model allelic lineages that have escaped original patch as subcritical branching processes branching rate (nearly) homogeneous $1 - s_d$.
- A surviving lineage looks like a single “trunk” with transient family of “branches” (Geiger 1999), all performing random walks.
- we want $\lambda_{\text{mig}} =$ number of lineages leaving patch
  $\times$ probability that the trunk hits the patch
  $\times$ the probability that one of its family establishes (i.i.d $2s_b$).

$$\approx \frac{\rho s_b}{s_d} \left(r\sqrt{2s_d}/\sigma \right)^{-1/2} \exp(-r\sqrt{2s_d}/\sigma).$$



MIGRATION AND MUTATION COMPARED

Mutations arise, and establish, on a new patch at rate

$$\lambda_{\text{mut}} = 2A\rho\mu s_b \tag{1}$$

Migration leads to establishment of migrant allele on patch at rate

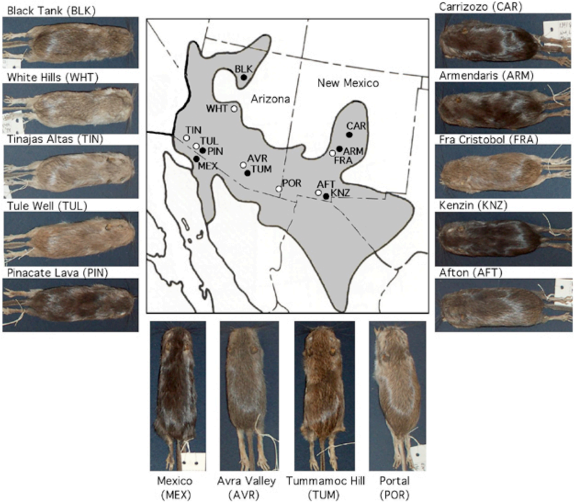
$$\lambda_{\text{mig}} \approx \frac{\rho s_b}{s_d} \left(r\sqrt{2s_d}/\sigma \right)^{-1/2} \exp(-r\sqrt{2s_d}/\sigma). \tag{2}$$

The relevant time scales of each are $T_{\text{mig}} = 1/\lambda_{\text{mig}}$ and $T_{\text{mut}} = 1/\lambda_{\text{mut}}$.

$$\frac{\lambda_{\text{mut}}}{\lambda_{\text{mut}} + \lambda_{\text{mig}}(r)} \approx \frac{A\mu s_d}{2A\mu s_d + \exp\left(-\sqrt{2s_d}\,r/\sigma\right)}, \tag{3}$$

MEANWHILE BACK ON THE LAVA FLOW

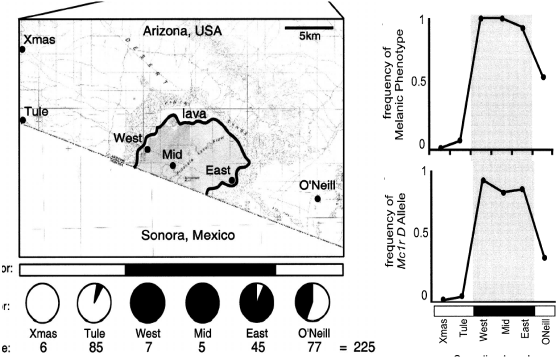
Back to *Chaetodipus intermedius* (Dice, Benson, Nachman, Hoekstra, etc)



APPLICATION TO ROCK POCKET MICE

On one flow (Pinacate) an allele of MC1R is responsible for much of the change to a dark pelage (Nachman et al 2003).

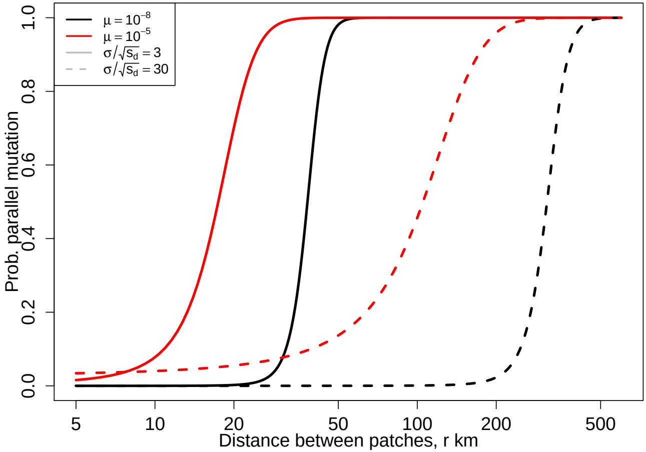
- Hoekstra et al (2004), cline width suggests $3km < \sigma/\sqrt{s_d} < 30km$
- Other studies $\sigma \approx 0.28km$ . . . so maybe $10^{-4} < s_d < 0.01$



APPLICATION TO ROCK POCKET MICE

Probability of parallel adaptation, i.e. adaptation by mutation before migration, is

$$\frac{\lambda_{\text{mut}}}{\lambda_{\text{mut}} + \lambda_{\text{mig}}(r)} \approx \frac{A\mu s_d}{2A\mu s_d + \exp\left(-\sqrt{2s_d}\, r/\sigma\right)}, \tag{4}$$



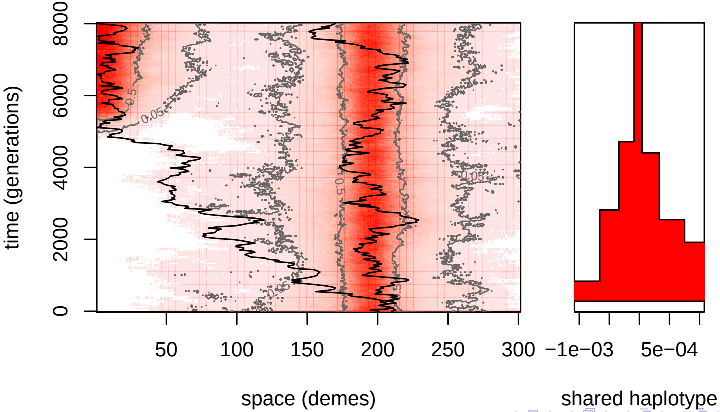
CAN WE DETECT ALLELES SHARED BY MIGRATION?

Alleles shared by migration will share a haplotype; how long?

- Haplotype whittled down by recombination between patches:
- if transit between patches takes t generations,
- shared haplotype will be $\approx$ Exponential with mean $1/t$.

The shared haplotype is initially of length:

$$\mathbb{E}[L] = \sigma\sqrt{2s_d}/r + \sigma^2/r^2 \tag{5}$$



CONCLUSIONS

- Convergent adaptation among patches should be common if alleles are deleterious outside patches.
- When adaptation by migration occurs, size of shared genomic region may be quite small.

ADAPTATION TO EXTREME ENVIRONMENTAL PATCH

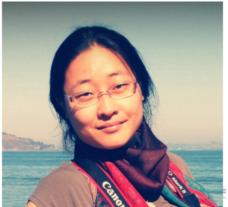
ADAPTATION TO EXTREME ENVIRONMENTAL PATCH

- Plants growing on toxic mine-tailings are a classic e.g. of rapid local adaptation. Role of standing variation?
- *Mimulus guttatus* has repeatedly adapted to Copper mines in California foothills. (Macnair & colleagues from 1980s onwards.)
- Mining began in 1860 at Copperopolis, so strong adaptation in $\sim$ 150 years.

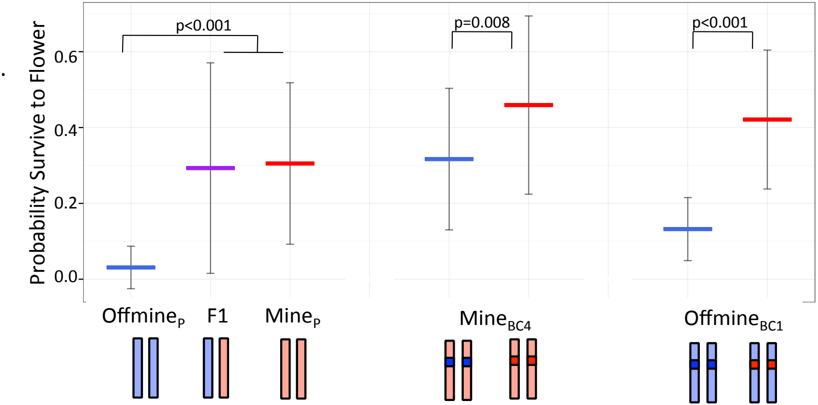
COPPEROPOLIS OR BUST!

- In 2013 Wright et al (PLOS Bio.) mapped the copper tolerance to one large QTL (TOL1).
- Sequenced pooled data from Copperopolis and McNulty mines and two nearby off-mine sites.
- F_{ST} low between all pairs ($F_{ST} = 0.049 - 0.079$).
- Any sweep in 150 gens will wipe out diversity over $\sim$ 100's of kb

Kevin Wright, Chenling Xu, Uffe Hellsten, Dan Rokshar, & John Willis

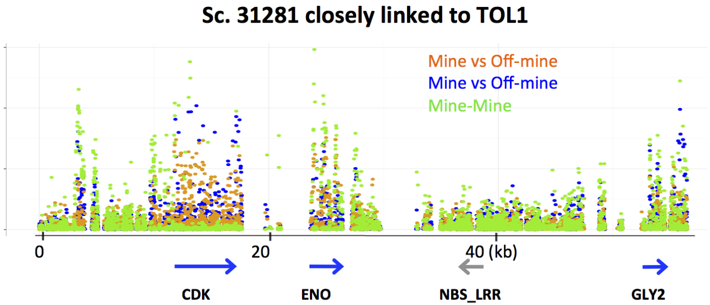


TOL1 IS A MAJOR FITNESS QTL ON COPPER MINE.

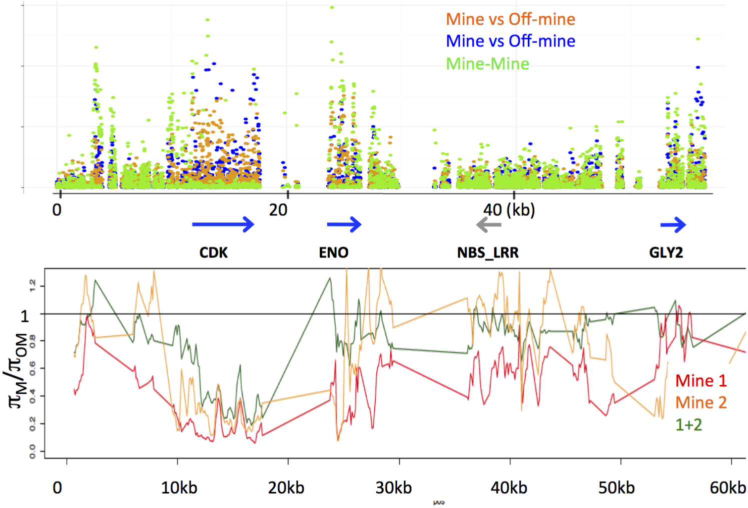


Difference in fitness $s \approx 0.7$ in favour of TOL1 allele on mine.
If due to new mutation, π would only recover to 50 % over 4Mb in TOL1 region .
No obv. fitness cost of TOL1 allele off-mine.

GENETIC DIVERSITY IN THE TOL1 REGION.



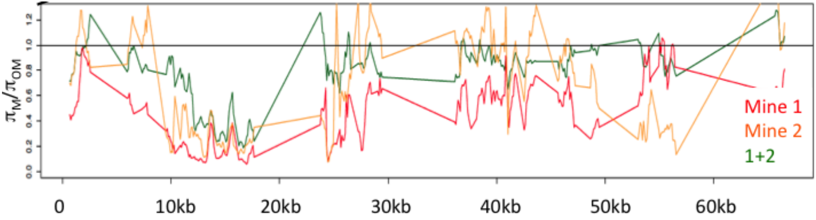
GENETIC DIVERSITY IN THE TOL1 REGION.



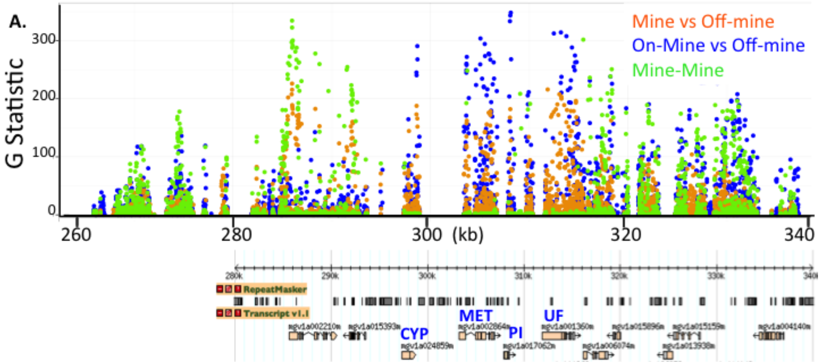
Sc. 31281 closely linked to TOL1

GENETIC DIVERSITY IN THE TOL1 REGION

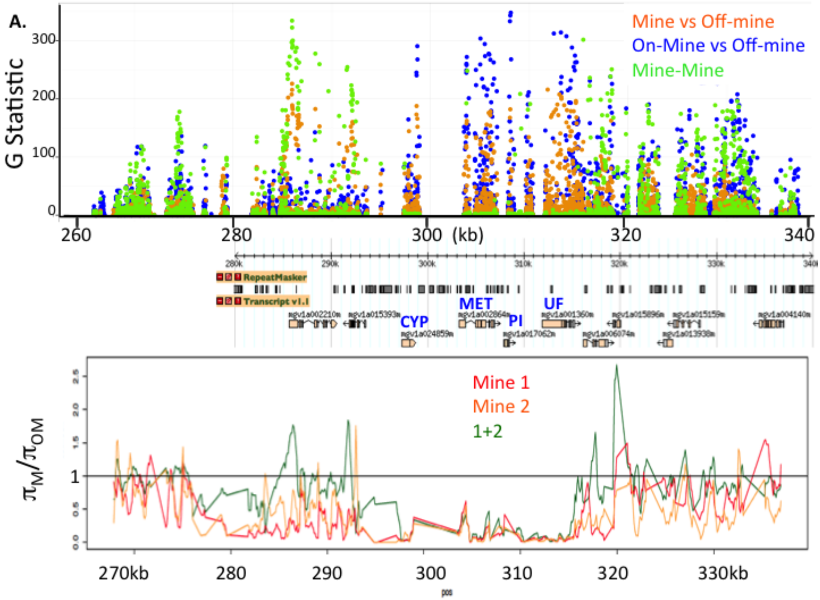
TOL1 region: no signal of single super sweep.
Instead evidence of small swept haplotype(s?), that are shared between populations.
"swept" haplotype old: e.g. 3,200 and 8,720 generations in the Copperopolis and McNulty mines.
Assuming that $s \approx 0.7$, then allele must have been at standing freq. $0.001 < f < 0.01$ in original pop. (calc. Xu, Berg,Coop)



GENETIC DIVERSITY IN TOP "SWEEP REGION".

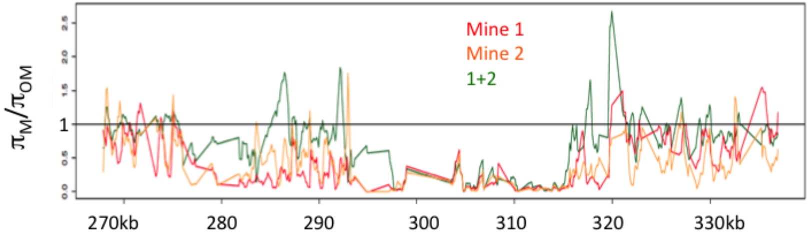


GENETIC DIVERSITY IN TOP "SWEEP REGION".



GENETIC DIVERSITY IN TOP "SWEEP REGION".

Top sweep candidate shared across populations.
We estimate the haplotype to be 659 and 801 generations old in the Copperopolis and McNulty mines
and 2500 generations between mine populations.



CONCLUSIONS

- Similar patterns hold for our top ten putative sweep regions.
- All stand out strongly against genomic background.
- All shared across mines, all older than mines.
- All likely adaptation from standing variation
- Rough calcs (Chenling, Jeremy Berg and Coop) suggest these alleles coming from $< 1\%$

THANKS

Kevin Wright, Uffe Hellsten, Dan Rokshar, John Willis, & Michael Nachman.
NIH and NSF

Chenling Xu

Peter Ralph

Simon Aeschbacher, Kristin Lee, Alisa Sedghifar, Gideon Bradburd, Jeremy Berg, & ???

LENGTH OF HAPLOTYPE SHARED BY MIGRATION

Trunk lineage: $\approx$ Brownian, killed at rate s_d.
The shared haplotype is initially of length:

$$\mathbb{E}[L] = \sigma\sqrt{2s_d}/r + \sigma^2/r^2 \tag{6}$$