Roadmap

- Course mechanics
- Correcting a slide about D (disequilibrium coefficient)
- Finishing up background selection
- Population subdivision
- Gene flow
- Gene flow versus selection

Course mechanics

1. Exam is still on Friday; this lecture **will** be included
2. Sample equation sheet is on the web site; this will be attached to the exam
3. Please bring a calculator
4. Updated syllabus will be posted this week

One-minute responses

- Are there simulators for the coalescent with recombination?
- If N_e is very large and recombination rate is low, do adjacent trees along the chromosome become uncorrelated?

Simulator for the coalescent with recombination

- ms program of Hudson
  http://home.uchicago.edu/rhudson1/source.html

- Very powerful:
  - Recombination
  - Population subdivision
  - Population growth and shrinkage

- Not user friendly

- Many derived programs add things like recombination hotspots, gene conversion

- A few even attempt to handle selection (VERY hard for a coalescent simulator)

Correlation between adjacent trees

- Two neighboring trees along the chromosome normally differ by just one recombination

- They will therefore be correlated, no matter what N_e is

- The higher N_e is:
  - The older, on average, haplotypes in the population are
  - The more recombinations they will contain
  - The shorter the segments with the same tree will be

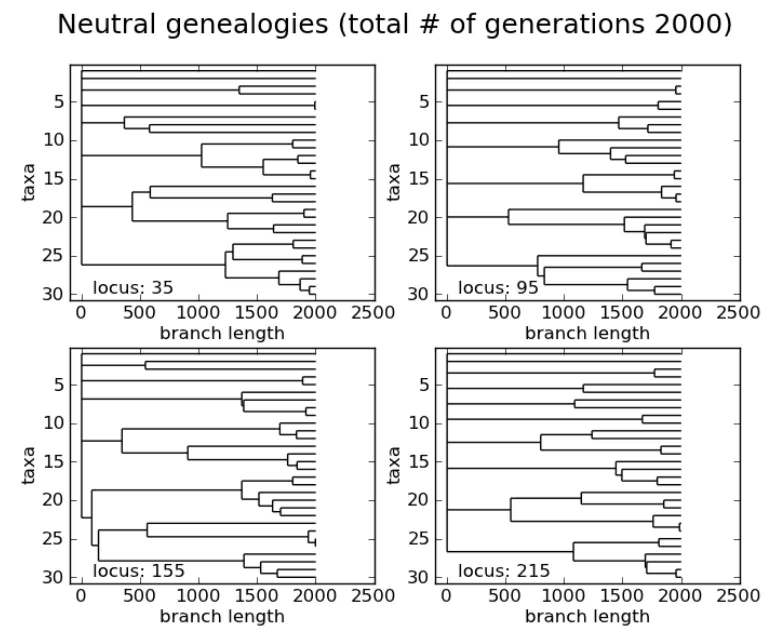
- If N_e and recombination both very high:
  - Segments very short; hard to detect
  - Might have multiple recombinations at the exact same place, de-correlating the trees

Correction to Lecture 10 (linkage)

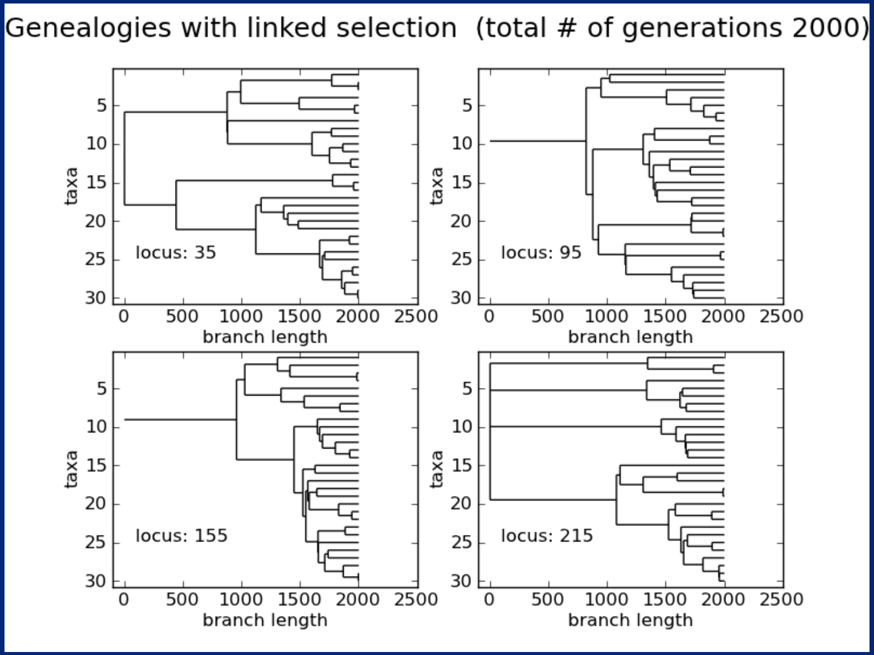
- I said: D ranges from $+1/2$ to -$1/2$
- Hongjie correctly pointed out: $+1/4$ to -$1/4$

Hitchhiking and related effects reduce N_e

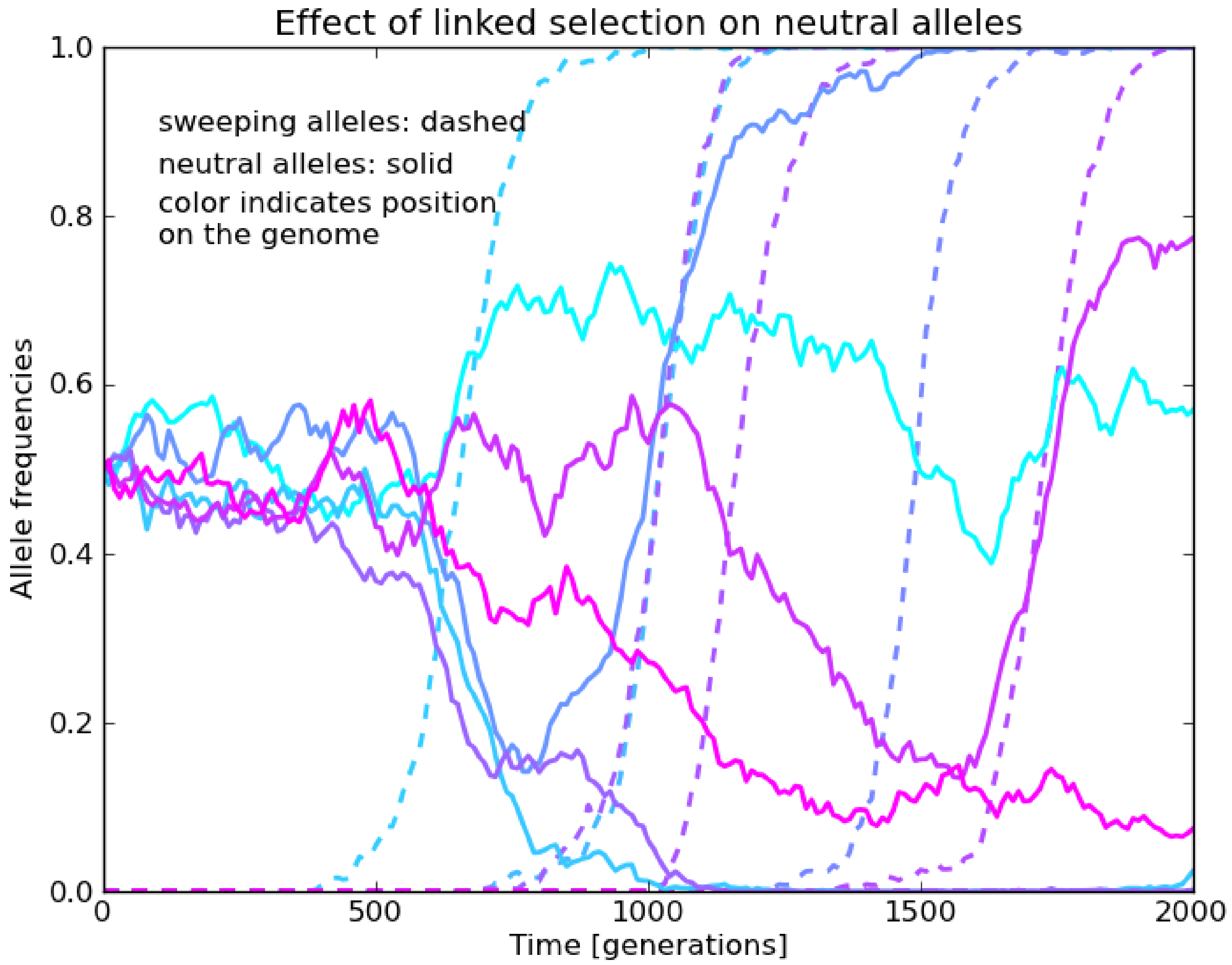
Without background selection



With background selection



Simulated data from
http://webdav.tuebingen.mpg.de/interference/draft.html


Effect of linked selection on neutral alleles
sweeping alleles: dashed
neutral alleles: solid
color indicates position
on the genome
Allele frequencies
0.0
0.2
0.4
0.6
0.8
1.0
0
500
1000
1500
2000
Time [generations]

“Genetic draft”

- Both positive and negative selection reduce N_e for nearby loci
  - Positive $=$ hitchhiking
  - Negative $=$ genetic draft or background selection (same thing)
- Not often considered in applications of the coalescent
- Could the anomalous low N_e for red drum be due to this?
  - Low recombination?
  - Frequent sweeps?

Population subdivision

- Isolated populations drift independently
- If they start out identical they will diverge over time:
  - Different alleles arise by mutation
  - Different alleles increase/decrease by drift
- Selection could make divergence faster or slower:
  - Selection for the same optimum in both populations keeps them more similar (leading to sequence conservation for key genes)
  - Different optima will lead to rapid divergence

How does this depend on N_e?

- The number of *fixed* differences depends on μ
- How fast each one fixes depends on N_e of the two populations
- Split a population into large mainland and small island:
  - Population sizes N_m and N_i
  - After about $4N_i$ generations with no migration, what will this system look like?

Wahlund Effect

- Population with pA=0.6: pAA under H-W is 0.48 (2*pA*pa)

- Two populations:

  1. pA=0.5
  2. pA=0.7

- If populations same size, overall pA still 0.6

- If we mix (not interbreed!) these populations, what is pAA?

Wahlund Effect

- Mixture of populations with different allele frequencies will not be in H-W
- Direction is ALWAYS too many homozygotes (why?)
- Useful for detecting:
  - Cryptic species
  - Hidden barriers to gene flow
  - Admixture zones
- As usual, goes away in 1 generation of truly random mating
- Can persist if populations not fully mixed

Isolation breaking

- Area right along the mixing edge of two populations with different allele frequencies can see transient excess of heterozygotes

- Example:
  - One population fixed for A, one fixed for a
  - Mix the adults–Wahlund effect
  - Interbreed in limited area where both are present–temporary excess of heterozygotes
  - (The local allele frequencies are not the same as the overall ones)

- Alternative hypothesis for increased heterozygotes: non-random mating (self-incompatibility in plants)

Wahlund-like effect on disequilibrium

- If the separate populations have different allele frequencies, the mixed population will have LD
- If they have the same allele frequencies but have LD, the mixed population will usually have LD
- LD decays in the usual way if the populations become fully mixed
- If there is a standing hybrid zone between two populations it may have permanent LD

Gene flow (migration)

- Important part for population genetics is *immigration*
  - Immigration from a population with different allele frequencies changes the recipient's frequencies
  - Emigration changes nothing (except possibly N_e)
    - Exception: emigration rate depends on genotype (flightless individuals fail to emigrate)
- When this lecture talks about migration, it always means immigration

Quantifying migration

- Migration of adults:
  - Population not in H-W during a generation with migration
- Migration of gametes:
  - Population always in H-W (simpler, but implausible for many organisms)

Two ways to quantify migration

- Proportion of gene pool made up by new immigrant haplotypes: m
- Number of immigrant haplotypes per generation: $2Nm$
- If a pair of lineages does not have a migration in their expected $2N$ generations time to their common ancestor, they will still be in the same population
- This suggests that $2N \times 2m = 1$ is the rough breakpoint
- Above this, migration holds the allele frequencies together; below it, they diverge

This rule varies in the literature

- Lots of sources say $2Nm >> 1$; lots say $4Nm >> 1$
- Felsenstein book says the rule is so rough it doesn't matter
- In reality there is a wide zone where behavior is neither that of isolated populations nor of a single panmictic population

[PopG demonstration goes here]

Migration is more effective than intuition would suggest

- In a neutral system:
  - Migration overcomes drift if number of migrants per generation much more than 1 individual per generation
  - Humans clearly move around more than that!
- Why do humans have any population distinctions at all?

Migration versus selection

- High altitude waterfowl benefit from specialized hemoglobin alleles which seem to be unhelpful at lower altitudes
- Five population pairs of high/low altitude Andean ducks
- We assume ducks don't migrate based on their phenotype (not a totally safe assumption)

Gene flow impeded at selected loci

McCracken et al. (2009) Am Nat.

- Estimated gene flow lower at Hb than five putatively neutral unlinked introns
- Pairs ranged from practically no divergence at non-Hb loci to substantial divergence (the last pair is currently considered two separate species)
- Complications of this research: what is phenotype and fitness of heterozygote? What about epistasis between Hb alpha and beta chains?
- (I thought the less diverged populations might be younger, but they all seem to date to the Ice Ages)

Rule of thumb

- Divergent selection can overcome drift when $s/m >> 1$
- (That s should be from multiplicative fitnesses; dominance complicates matters.)
- Gene flow in the presence of divergent selection reduces fitness
- Population may respond by reducing gene flow

Clines

- Region between two populations where allele frequencies shift
- Could be due to drift:
  - Two populations are isolated and diverge by drift
  - When they re-contact, a cline forms
- Could also be due to divergent selection

Short-range cline around a zinc mine

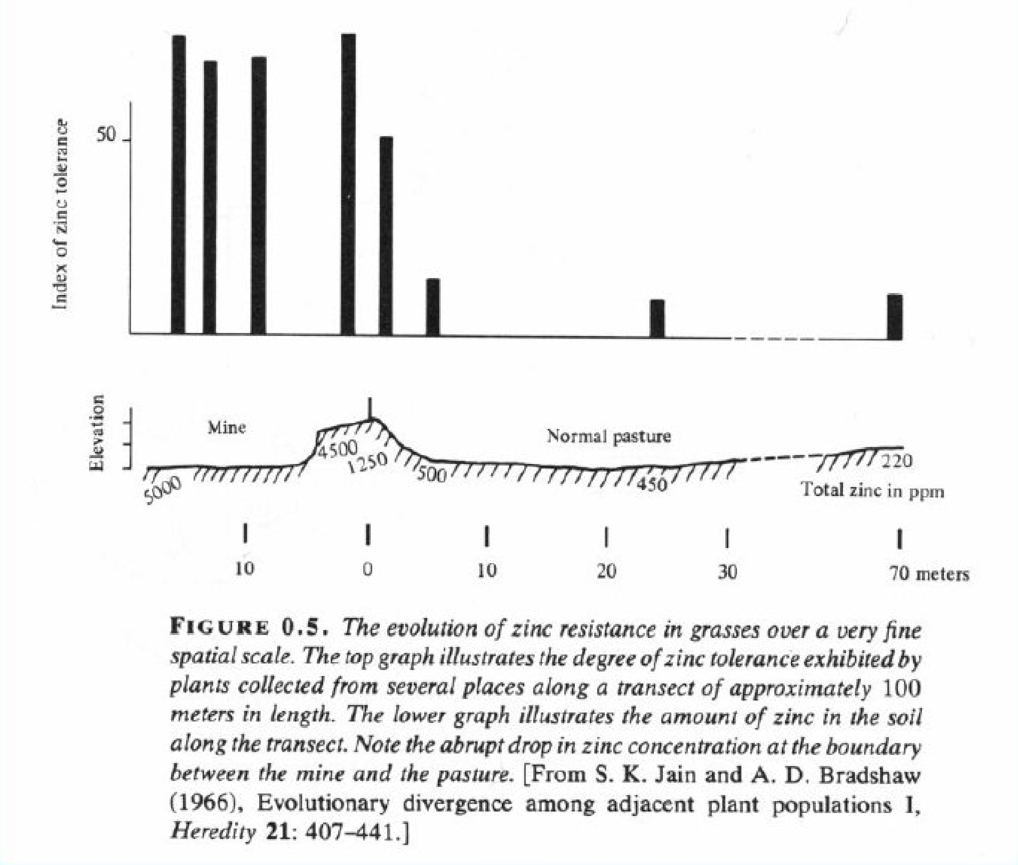
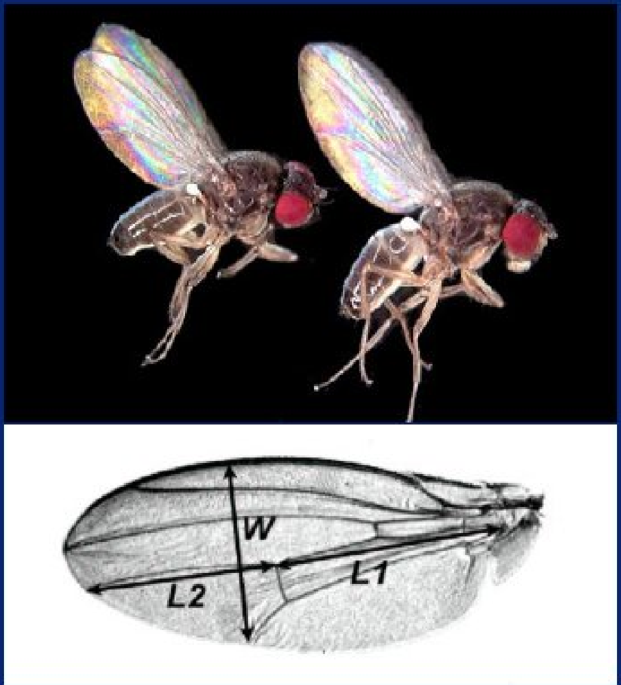


FIGURE 0.5. *The evolution of zinc resistance in grasses over a very fine spatial scale. The top graph illustrates the degree of zinc tolerance exhibited by plants collected from several places along a transect of approximately* 100 *meters in length. The lower graph illustrates the amount of zinc in the soil along the transect. Note the abrupt drop in zinc concentration at the boundary between the mine and the pasture.* [From S. K. Jain and A. D. Bradshaw (1966), Evolutionary divergence among adjacent plant populations I, *Heredity* **21**: 407–441.]

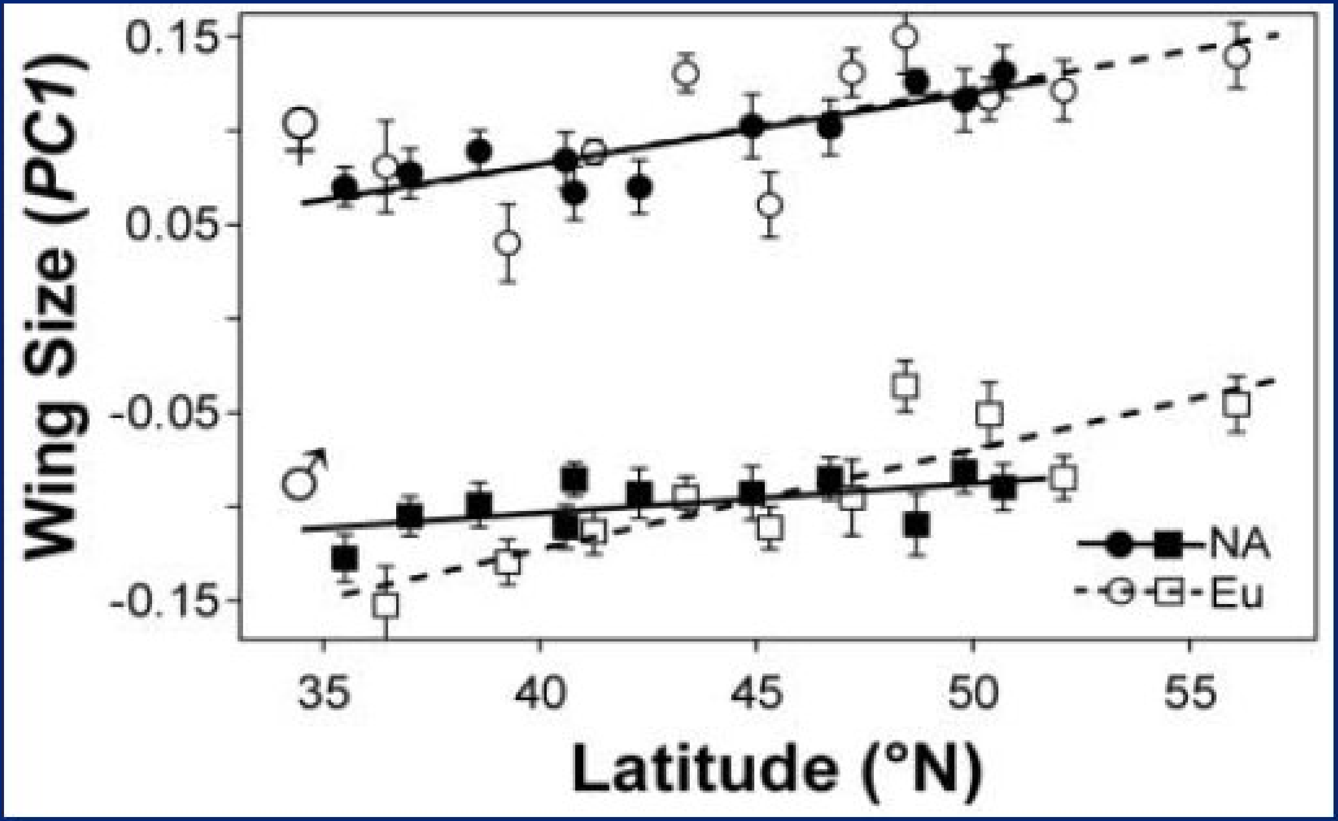
Clines and speciation

- The zinc cline is potentially costly
- Mating with nearby non-resistant plants can be disastrous
- How can the plants avoid this?
  - Self-fertilization
  - Mating with nearby plants (less mobile seeds/pollen?)
  - Asexual reproduction (runners)
  - Mating only with one's own kind
- This could be a first step to speciation

Drosophila subobscura



Drosophila subobscura cline



Estimating gene flow: the F statistics

- Sewall Wright estimated gene flow by comparing variability between and within populations
- I will cover this Monday as it is closely related to the math for inbreeding

Monday

- Inbreeding:
  - Effect on heterozygosity
  - Effect on N_e
  - Interaction with selection
- F statistics and migration

One-minute responses

- Please:
  - Tear off a slip of paper
  - Give me one comment or question on something that worked, didn't work, needs elaboration, etc.