Roadmap

- Midterm
- Gene flow versus selection
- Inbreeding
  - Effect on heterozygosity
  - Effect on N_e
  - Interaction with selection
- F Statistics and migration

Midterm

- Distribution of grades: 91, 86, 85, 84, 69, 68, 67
- Exam too long for given time
- All exams met acceptable graduate student standards and grades will be interpreted accordingly

Midterm

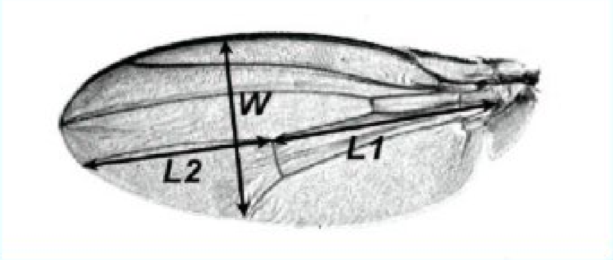
Issues to watch for on future homeworks and final:

- Don't give one-generation solution if problem states "for a long time"
- Check if the explanation you are proposing goes in the right direction:
  - Selection against something should not make it more common
  - Inheritance of very old haplotypes should not reduce diversity
  - ...etc.
- Key is available on web site

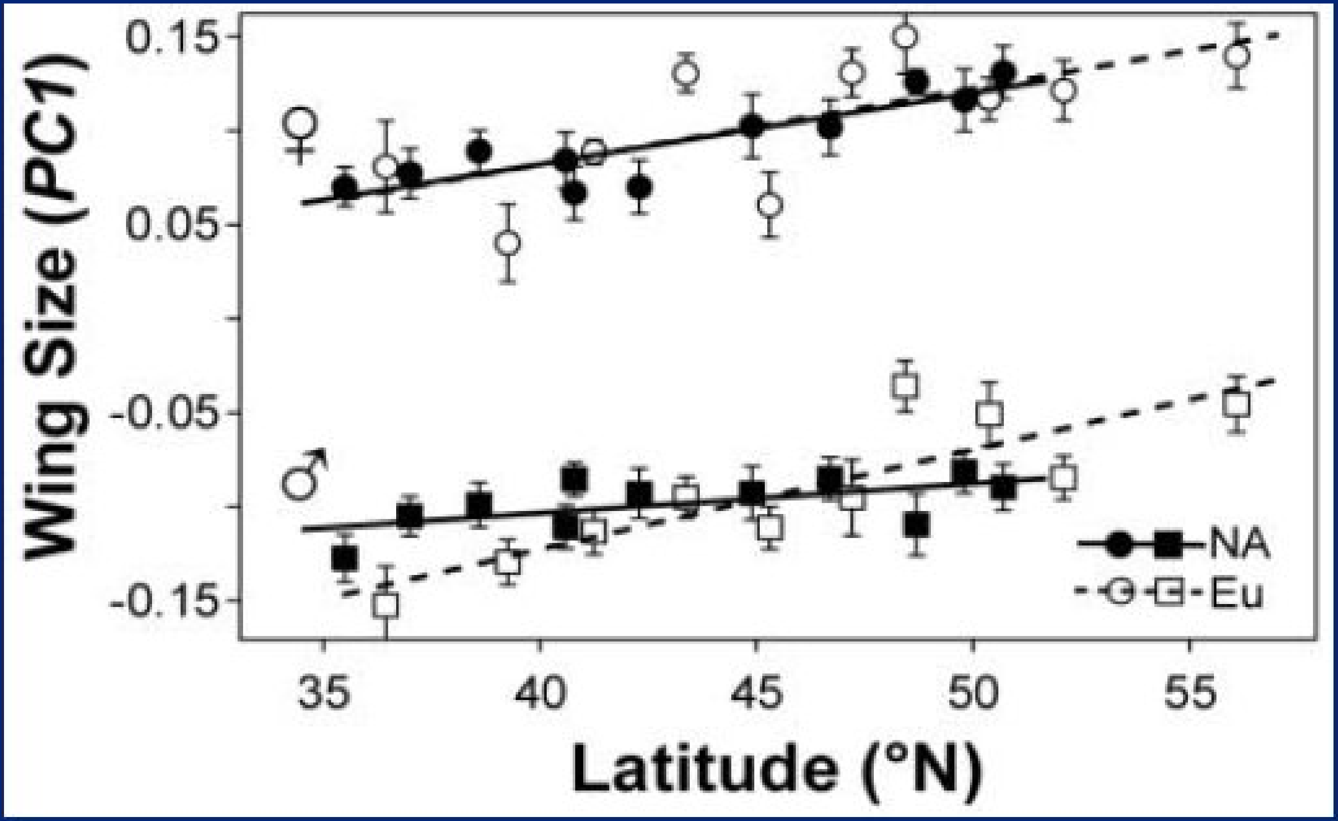
Clines

- 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

Drosophila subobscura



Drosophila subobscura cline



Inbreeding

- Ways to look at inbreeding:
  - Increased chance of mating with a related individual
  - Increased chance of two alleles with a recent common ancestor ending up in the same individual
- Population subdivision can be looked at as a form of inbreeding (more on this later)

Identity by descent

- We agree to an artificial distinction:
  - Two alleles with a common ancestor within the time depth we're considering are identical by descent (IBD)
  - Alleles that do not have a common ancestor within this depth are not IBD, even if they are identical (identity by state, IBS)

- The time depth could be:
  - Within the known pedigree
  - An arbitrary cutoff time
  - Back to a non-inbred source population
  - Something nebulous and not discussed

- Note that two alleles IBD might not be IBS if there is a new mutation

Inbreeding coefficient

- Single generation inbreeding of any intensity can be abstracted as:
  - Fraction f of offspring contain two gene copies IBD from the previous generation
  - Fraction $1 - f$ are not IBD (though they may be IBS)
- This is tightly related to the coefficient of kinship
  - Coefficient of kinship between two individuals is the f of a hypothetical offspring

Pedigree examples

- Self-fertilization
- Half siblings
- Full siblings

H-W for a population with average inbreeding f

- In one generation of inbreeding, starting with no IBD
- (Using the convention that pA = p and pa = q)
  $pAA = p^2(1 - f) + pf$
  $pAa = 2pq(1 - f)$
  $paa = q^2(1 - f) + qf$
- A fraction f of heterozygotes is being changed into homozygotes due to inbreeding
- Note that allele frequencies are not changing
- Another way to write this:
  $pAA = p^2 + fpq$
  $pAa = 2pq - 2fpq$
  $paa = q^2 + fpq$

Repeated inbreeding–discuss!

- If allele frequencies are not changing, why does heterozygosity continue to go down over time?
- Previous generation already has alleles IBD
- Repeated self-fertilization:
  - $f = 0.5$
  - In one generation, lose half of heterozygosity
  - Next generation, lose half of what's left; individuals can be IBD with respect to their parent, or failing that their grandparent
- Does this predict that *any* non-zero f eventually makes everyone homozygous?

Selection in the presence of inbreeding

- Mean population fitness:
  $(p^2(1-f)+pf)wAA + 2pq(1-f)wAa + (q^2(1-f)+qf)waa$
- Rearrange terms to get a difference from outbreeding
- Mean fitness is lower in inbred population if wAa higher than mean of wAA and wAa
  - Any degree of dominance of the better allele
  - Overdominance
- Mean fitness is higher in inbred if:
  - Any dominance of the worse allele
  - Underdominance

Why is inbreeding even tolerable?

- In most cases, population mean fitness is lower with inbreeding for a given set of allele frequencies
- However, this causes selection to change the allele frequencies faster:
  - Bad recessives eliminated sooner
  - Good recessives fixed sooner
- Populations like *Arabidopsis* which habitually inbreed do not see a fitness drop due to inbreeding, because bad recessives already purged
- Overdominance is still an issue

Inbreeding and overdominance

- If $f = 1$ no overdominant equilibrium exists:

Genotype	Fitness	Frequency before selection
AA	0.8	p
Aa	1.0	0
aa	0.7	q

- For any non-zero starting frequency, A will fix (barring drift)
- If $f < 0$ the equilibrium exists, but with high f is very vulnerable to drift as heterozygotes are so rare

Self-fertilization vs. cloning

- Self-fertilization:
  - Decrease in heterozygotes every generation
  - Overdominance can't be maintained
  - Recombination can happen, but becomes less useful as everything is homozygous

- Cloning:
  - Random drift of genotype frequencies
  - Overdominant heterozygote can *fix*
  - Recombination probably not happening
  - In the very long run, diploid state likely to be lost (rotifers)

Population subdivision as a form of inbreeding

- An isolated subpopulation is inbred relative to the whole population, even if it has random mating internally

- Wright conceptualized this as the increase in IBD for two alleles from the same subpopulation relative to the total population

- Three terms:
    - F_{IS} – excess IBD in an Individual relative to its Subpopulation (measures inbreeding such as non-random mating)
    - F_{ST} – excess IBD in two alleles drawn from a Subpopulation relative to the Total population
    - F_{IT} – necessarily determined by the other two

- F_{ST} frequently used as a measure of population differentiation

F_{ST} in terms of variances

- $F_{ST} = \frac{\sigma_S^2}{\sigma_T^2} = \frac{\sigma_S^2}{\hat{p}(1-\hat{p})}$
- Here σ_S^2 is the variance in allele frequency across populations, and $\hat{p}$ is a measure of variance within the total population

F_{ST} in terms of inbreeding coefficient

- $F_{ST} = \frac{f_0 - \hat{f}}{1 - \hat{f}}$
- Here f is the probability of IBD of two draws from a subpopulation, and $\hat{f}$ is the probability for two draws from the total population

F_{ST} in terms of the coalescent

- $F_{ST} \approx 1 - \frac{t_S}{t_T}$
- (Note correction on this slide!)
- t_S is the expected time back to the subpopulation common ancestor
- t_T is the expected time back to the total population common ancestor

F_{ST} as it might actually be measured!

- $F_{ST} \approx \frac{\pi_B - \pi_W}{\pi_B}$
- π_B is mean pairwise difference between individuals from different subpopulations
- π_W is mean pairwise difference between individuals from the same subpopulation
- This will react to:
  - Divergence time (longer separation is higher F_{ST})
  - Gene flow (more gene flow is lower F_{ST})

The dark side of F_{ST}

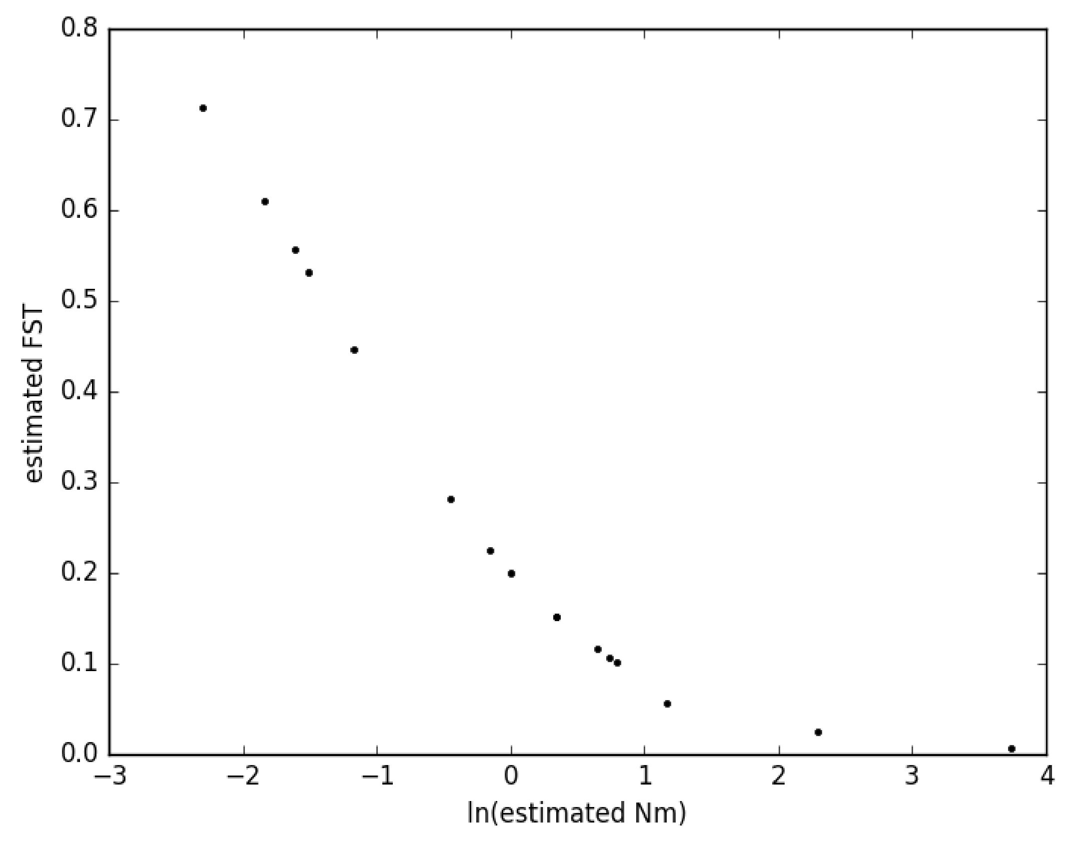
- $F_{ST} = \frac{\sigma_S^2}{\sigma_T^2} = \frac{\sigma_S^2}{\hat{p}(1-\hat{p})}$
- If $\hat{p}$ is very large or small (one allele is very frequent), F_{ST} can't get big
- None the less, people attempt to give it an absolute interpretation
- Need to average across multiple loci to get meaningful results
- Still vulnerable to unexpected allele frequency spectrum (e.g. in growing/shrinking populations)

The dark side of F_{ST}

- $F_{ST} \approx \frac{\pi_B - \pi_W}{\pi_B}$
- Good luck if your sampling yields $\pi_B < \pi_W$, and it can!
- You can do better with ANOVA or AMOVA, or with the coalescent methods to be described later

F_{ST} estimates from Slatkin 1985

Type of organism	Estimated Nm	Estimated F_{ST}
Annual plant	1.4	0.152
Mollusc	42.0	0.006
Insect	9.9	0.025
Insect	1.0	0.200
Fish	3.2	0.056
Frog	1.4	0.152
Salamander	2.1	0.106
Salamander	0.22	0.532
Salamander	0.10	0.714
Salamander	0.64	0.281
Salamander	0.20	0.556
Salamander	0.16	0.610
Lizard	1.9	0.116
Mouse	2.2	0.102
Mouse	0.31	0.446
Gopher	0.86	0.225



Data from Slatkin 1985

Estimate Nm from F_{ST}

- As this graph shows, you can!
- Unfortunately:
  - Assumes same migration in all directions
  - Assumes similar population sizes
  - Quite biased if these assumptions not met
- There are better, but computationally complex, ways to do this

Wednesday

- Note that Monday 2/18 is a holiday! Homework will be assigned Wednesday, due the following Monday
- Quantitative traits
- Heritability
  - What it is
  - (Important) What it isn't

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.