

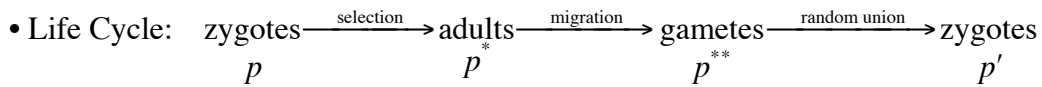
## MIGRATION & SELECTION

• **One-island model with selection**

- *A* favored on island.
- *a* fixed on continent:  $p_c = 0$ .
- *A* is dominant.

• Fitnesses on island:

<b>Genotype</b>	<i>AA</i>	<i>Aa</i>	<i>aa</i>
<b>Fitness</b>	1	1	1 - <i>s</i>



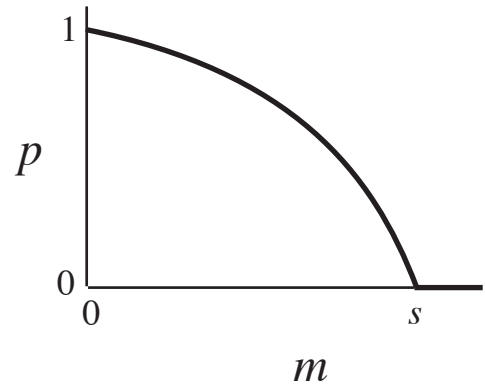
• After selection (before migration):  $p^* = p \frac{1}{1 - q^2 s}$

• After migration & reproduction:  $p' = (1 - m)p^* + m(0) = \frac{p(1 - m)}{1 - q^2 s}$

• To find any equilibria, set  $p' = p$ .

– Solving for  $p$  gives  $\hat{p} = 1 - \sqrt{m/s}$ .

- Require  $0 \leq \hat{p} \leq 1$ .
  - This occurs only when  $m < s$ .
  - Otherwise  $\hat{p} = 0$ .



– Now assume *A* is recessive.

• Fitnesses on island:

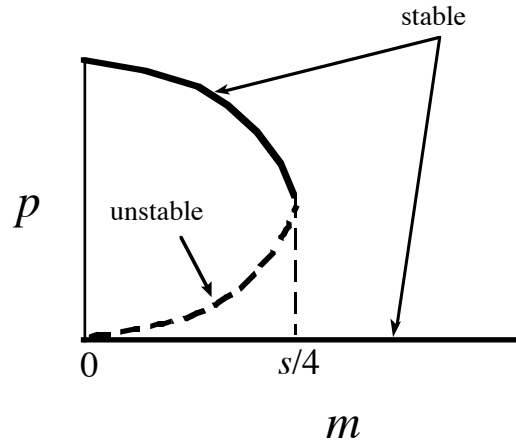
<b>Genotype</b>	<i>AA</i>	<i>Aa</i>	<i>aa</i>
<b>Fitness</b>	1	1 - <i>s</i>	1 - <i>s</i>

• After selection (before migration):  $p^* = \frac{p(1 - qs)}{1 - sq(1 + p)}$

• After migration & reproduction:  $p' = (1 - m)p^* + m(0) = \frac{p(1 - m)(1 - qs)}{1 - sq(1 + p)}$

• To find equilibria, set  $p' = p$  and solve for  $p$ .

- Get cubic equation for  $\hat{p}$ 's (up to 3 possible solutions).
- $\hat{p} = 0$  is always an equilibrium (since  $p_c = 0$ ).
- There are two *polymorphic* equilibria when  $s > 4m$  (assuming  $m$  is small).
  - one equilibrium is stable, the other is unstable.



- Graphically  $\longrightarrow$

• Implications

- If recessive selection is strong enough to maintain  $A$  in the face of migration,  $A$  will spread only if it's initially sufficiently frequent enough. Otherwise, it will be lost.
- In general, unless locally advantageous allele is completely dominant, it must reach a threshold frequency to persist.
- If an allele persists, it won't be found at a low frequency.
- Historical "accidents" play a role.
  - Identical patches will evolve differently if they differ in initial allele frequency.

• **The Levene Model**

**Q:** What happens when a population is made up of a group of distinct subpopulation patches, with different selection pressures occurring in each and migration between locations?

**A:** Depends on geography (population structure).

- Natural populations fall somewhere between the following two extremes:

- Unrestricted migration.
- Restricted migration.

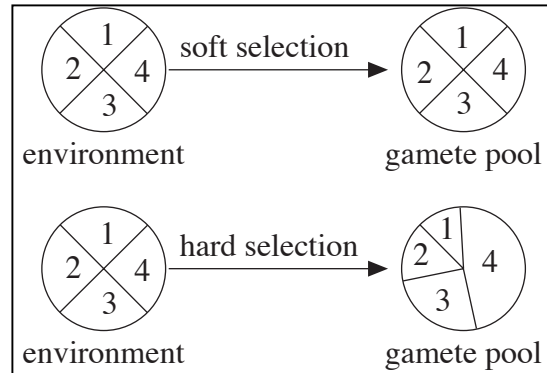
- A simple model of unrestricted migration was presented in 1953 by H. Levene.
- Assumptions of Levene's 1953 model:
  - $n$  patches in which different patterns of selection occur.
  - Frequency of  $A$  among gametes is  $p$ .
  - After fertilization, (diploid) zygotes colonize the different patches (at random).
    - Important: this implies that the zygotes *within* patches are in H-W proportions.
  - $i$ th patch makes up a fraction  $c_i$  of the environment.
  - Fitnesses in the  $i$ th patch:

<b>Genotype</b>	$AA$	$Aa$	$aa$
<b>Fitness</b>	$w_{AA}(i)$	$w_{Aa}(i)$	$w_{aa}(i)$

- Random mating between patches.
  - Individuals from different localities form a single mating (gamete) pool.
- Why study the Levene model?
  - Captures essential features of spatially subdivided population and is mathematically tractable.
  - Is a reasonable representation of certain natural systems as well.
- Back to model...How many gametes does each patch contribute to the gamete pool?
  - Two extremes:
    - (1) Hard selection (due to Dempster, 1955)
      - Patch contributes gametes in proportion to the fraction of survivors.
        - i.e., patches with higher fitness contribute disproportionately more.
      - Implies population size is not regulated within patches.
    - (2) Soft selection
      - Each patch contributes fixed number of gametes to the mating pool *regardless* of local fitnesses.
      - Number of reproducing adults from each patch is the same from one generation to the next.

- Implies population size is regulated within each patch.

- A schematic comparison between soft and hard selection, assuming  $c_1 = c_2 = c_3 = c_4$  and  $\bar{w}_4 > \bar{w}_3 > \bar{w}_2 > \bar{w}_1$



– Levene model with hard selection (“constant number of zygotes”):

- Assumes contribution of genotype from patch  $i$  to the gamete pool is proportional to its fitness in that patch  $[w_{\text{genotype}}(i)] \times$  frequency of  $i$  patch in environment ( $c_i$ ):
  - i.e., total number of survivors of that genotype in patch  $i \propto c_i w_{\text{genotype}}(i)$

- Overall fitness of genotype in population is its average fitness over patches:

– For example, mean fitness of  $AA$ :  $\bar{w}_{AA} = \sum_{i=1}^n c_i w_{AA}(i)$

– Likewise for  $Aa$  and  $aa$ .

- Consider changes in the frequency  $p$  of  $A$  in the gamete pool.

•  $p' = p \frac{p\bar{w}_{AA} + q\bar{w}_{Aa}}{\bar{w}} = p \frac{\bar{w}_A}{\bar{w}}$  where  $\bar{w}_A = p\bar{w}_{AA} + q\bar{w}_{Aa}$  and  $\bar{w} = p^2\bar{w}_{AA} + 2pq\bar{w}_{Aa} + q^2\bar{w}_{aa}$ .

- Looks just like selection with constant fitnesses:  $\bar{w}_{AA}, \bar{w}_{Aa}, \bar{w}_{aa}$

• Consequences

- An allele will spread if it has the highest arithmetic mean fitness across patches.
- Selection will maintain a stable polymorphism if heterozygotes have the greatest arithmetic mean fitness across patches.

- For example, consider two equally sized patches,  $c_1 = c_2 = 0.5$ .

Fitness in patch:	$AA$	$Aa$	$aa$
# 1	0	0.75	1
# 2	1	0.75	0
Average:	0.5	0.75	0.5

- Selection maximizes arithmetic mean fitness across environments

– Levene model with soft selection (“constant number of adults”):

– Within each patch, selection operates as usual.

– Fitness in patch  $i$ :  $\frac{AA}{w_i} \quad \frac{Aa}{1} \quad \frac{aa}{v_i}$

– After selection, frequency of  $A$  in patch  $i$  is

$$p^*(i) = p \frac{pw_i + q(1)}{p^2w_i + 2pq(1) + q^2v_i} = p \frac{\bar{w}_A(i)}{\bar{w}(i)}$$

– Density regulation occurs independently in each patch.

– Survivors contribute to gamete pool in proportion to the size (= relative proportion of adults) of the patch,  $c_i$ :

$$p' = \sum_{i=1}^n c_i p^*(i) = \sum_{i=1}^n c_i p \frac{pw_i + q}{p^2w_i + 2pq + q^2v_i}$$

– Equilibrium: set  $p' = p$  and solve for  $p$ .

- Results in polynomial of degree  $2n + 1$  in  $p$ .  
 $\Rightarrow$  as many as  $2n + 1$  equilibria,  $\hat{p}$ , are possible!
- Mathematically too difficult to find all these.

– Alternative: protected polymorphism analysis:

- Near  $p = 0$ ,  $p' \approx \sum_{i=1}^n \left( \frac{c_i}{v_i} \right) p = p \frac{1}{\tilde{v}}$

where  $\tilde{v} = 1 / \left[ \sum_{i=1}^n c_i \left( \frac{1}{v_i} \right) \right]$  is the harmonic mean fitness of  $aa$  homozygotes.

- Note that  $p' > p$  (i.e.,  $\Delta p > 0$ ) whenever  $1/\tilde{v} > 1 \Leftrightarrow \tilde{v} < 1$   
 – i.e., whenever the “harmonic mean fitness of  $aa$  homozygotes” < “mean fitness of heterozygotes”
- Likewise, near  $p = 1$ , ( $q = 0$ ),  $q' > q$  whenever  $\tilde{w} < 1$ .
- Conclude: protected polymorphism occurs with soft selection whenever there is harmonic mean overdominance in fitness across patches:  $\tilde{w} < 1 > \tilde{v}$ .

– Bottom line(s) for soft selection

- Harmonic mean fitness across patches is the relevant fitness measure if  $p \approx 0$  or 1.
- Turns out, however, that selection maximizes *geometric* mean fitness.

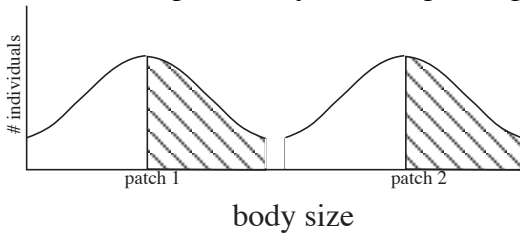
– Hard versus Soft Selection

- Conditions exist in which an allele will increase under soft selection but not hard selection.
  - I.e., polymorphisms can be maintained under a broader range of conditions with soft selection versus hard selection.
  - Intuitively follows because under soft selection, individuals compete selectively only against "patch-mates".
    - With hard selection, all compete.
  - Mathematically follows because harmonic mean is never larger than the

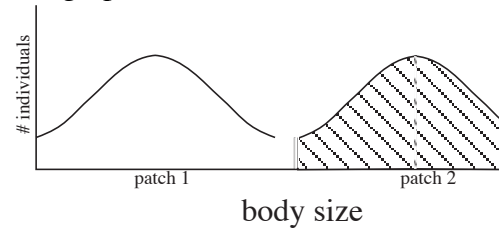
arithmetic mean:  $\tilde{v} \leq \bar{v} = \sum_{i=1}^n c_i v_i$ .

Q: Why does soft selection seem “hard” (density regulation; intense local competition) while hard selection seems “soft” (little competition; no density regulation)?

A: It all depends on your viewpoint (genetic vs. demographic).



– Soft selection: top 50% in each patch selected.



– Hard selection: top 50% selected (regardless of patch).