MIGRATION

ROLES OF MIGRATION IN EVOLUTION

READING: Hedrick pp. 403-436.

• Introduces novel genetic variation into populations.

• Tends to homogenize gene frequencies in different populations.

• Sets the spatial scale for evolution.

• Opposes local adaptation.

• Migration with an evolutionary impact: Gene Flow
  
  – Migration introduces individuals and genotypes (“dispersal”).
  
  – Migrants have no effect on evolution unless their genes are incorporated into a population.

• A One-Island Model
  
  – The simplest model of migration.
  
  – Two alleles A and a. Let \( p \) = frequency of A on island.
  
  – A fraction \( m \) of the island gene pool emigrates from the continent where the frequency of A is \( p_c \).

  \[ \Rightarrow \text{A fraction } (1 - m) \text{ of alleles on the island originated on the island.} \]

  – The continent is too vast to be influenced by migration from the island \( \Rightarrow p_c \) is constant.

  – Then the frequency of A on the island changes according to \( p' = (1 - m)p + mp_c \).
– At equilibrium, set \( p' = p \).

• Solving for \( p \) gives \( \hat{p} = p_c \)

– Rate of approach to equilibrium:

• Rewrite evolutionary equation as

\[
p' - \hat{p} = p' - p_c = (1 - m)p + mp_c - p_c = (1 - m)(p - p_c) = (1 - m)(p - \hat{p})
\]

– Conclusions

(1) At equilibrium, both populations have the same allele frequencies.

(2) Rate of approach to equilibrium (\( \hat{p} = p_c \)) is determined by the migration rate \( m \).

• General Models of Migration

– Same conclusions as one-island model hold.

– Exceptions, however, do exist

• For example, consider two populations with different allele frequencies that switch locations each generation.

• The populations will obviously never homogenize (because there's no real exchange of genes).

– Remark: Have implicitly assumed gene frequencies differ in different locations.

– How could this be?
  – “History.”
  – Genetic drift.
  – Selection favors different alleles in different locations.
MIGRATION AND DRIFT

- Migration introduces novel genetic variation into local populations.

- Drift removes local genetic variation.

Which for dominates?

One answer…

- Wright’s “Island Model”

  Consider a large number of “islands” each with a population of size $N$ ($2N$ alleles per locus)

  Each generation, every island exchanges a fraction $m$ of its gametes with an $\infty$-sized “migrant pool” to which all islands contribute gametes.

  Assume infinite-isoalleles model.

  Let $f_i = \text{Pr}(\text{pair or randomly drawn gametes on a typical island are IBD in generation } t) = \text{average within-island homozygosity}$

  By the same logic used when studying mutation-drift balance:

  $$ f_{i+1} = (1 - m)^2 \left[ \frac{1}{2N} + \left( 1 - \frac{1}{2N} \right) f_i \right] $$

  At equilibrium, $f_{i+1} = f_i = \hat{f} = \frac{1}{1 + 4Nm}$

  – expression resembles that describing diversity maintained by mutation & drift, with $\theta = 4Nu$ replaced by $4Nm$.

  If $4Nm < 1$: Local homozygosity is substantial
  – drift dominates migration

  If $4Nm > 1$: Local diversity (heterozygosity) is substantial
  – migration dominates drift

Note 1

- $4Nm > 1$ same as $2Nm > 1/2$
⇒ Migration dominates drift if at least one migrant gamete is exchanged every other generation!

– Conclusion is independent of \( m \), the rate of gene flow. (Why?)

**Note 2**

– Recall from discussion of F statistics: \( \bar{H}_S = \operatorname{Avg}(H_{S,i}) \approx 1 - \hat{f} \), since \( \hat{f} \) is the average local homozygosity and there is no additional inbreeding

– Also, \( H_T = 1 - \Pr(\text{pair of randomly chosen gametes from entire population are IBD}) = 1 - 0 = 1 \)

\[ \Rightarrow F_{ST} = \frac{H_T - \bar{H}_S}{H_T} = 1 - \left(1 - \hat{f}\right) \frac{1}{1} = \hat{f} = \frac{1}{1 + M} \text{, where } M = 4Nm. \]

– Suggests way to estimate rate of migration from \( F_{ST} \):

\[ \hat{M} = \frac{1 - F_{ST}}{F_{ST}}. \]

– Careful: estimate requires lots of assumptions (island model, equilibrium, etc.) to be valid.

**Migration and Selection**

- One-island model with selection
  - \( A \) favored on island.
  - \( a \) fixed on continent: \( p_c = 0 \).
  - \( A \) is dominant.

  - Fitnesses on island:

    | Genotype | AA | Aa | aa |
    |----------|----|----|----|
    | Fitness  | 1  | 1  | 1 - s |

  - Life Cycle: \( p \) zygotes \( \xrightarrow{\text{selection}} \) adults \( p \) \( \xrightarrow{\text{migration}} \) gametes \( p^* \) \( \xrightarrow{\text{random union}} \) zygotes \( p' \)

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

- To find any equilibria, set \( p' = p \).
  - Solving for \( p \) gives \( \hat{p} = 1 - \sqrt{ms} \).
  - 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:
    | Genotype | AA | Aa | aa |
    |----------|----|----|----|
    | Fitness  | 1  | 1-s| 1-s|

  - 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

- 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:

<table>
<thead>
<tr>
<th>Genotype</th>
<th>( AA )</th>
<th>( Aa )</th>
<th>( aa )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitness</td>
<td>( w_{AA}(i) )</td>
<td>( w_{Aa}(i) )</td>
<td>( w_{aa}(i) )</td>
</tr>
</tbody>
</table>
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 $w^*_4 > w^*_3 > w^*_2 > w^*_1$.

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

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

\[
p' = p \frac{p \bar{w}_{AA} + q \bar{w}_{Aa}}{\bar{w}} = p \frac{\bar{w}_A}{\bar{w}} \quad \text{where} \quad \bar{w}_A = p \bar{w}_{AA} + q \bar{w}_{Aa} \quad \text{and} \quad \bar{w} = p^2 \bar{w}_{AA} + 2pq \bar{w}_{Aa} + q^2 \bar{w}_{aa}.
\]

• 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, \( c_1 = c_2 = 0.5 \).

<table>
<thead>
<tr>
<th>Fitness in patch:</th>
<th>AA</th>
<th>Aa</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>0</td>
<td>0.75</td>
<td>1</td>
</tr>
<tr>
<td>#2</td>
<td>1</td>
<td>0.75</td>
<td>0</td>
</tr>
<tr>
<td>Average</td>
<td>0.5</td>
<td>0.75</td>
<td>0.5</td>
</tr>
</tbody>
</table>

• 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 \( i \): \( \bar{w}_{i} \)
  – 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, \( c_i \):
\[ p' = \sum_{i=1}^{n} c_i p^*(i) = \sum_{i=1}^{n} c_i p \frac{p w_i + q}{p^2 w_i + 2 pq + q^2 v_i} \]

- Equilibrium: set \( p' = p \) and solve for \( p \).

  - Results in polynomial of degree \( 2n + 1 \) in \( p \).
    \[ \implies \text{as many as } 2n + 1 \text{ equilibria, } \hat{p}, \text{ are possible!} \]
  
  - Mathematically too difficult to find all these.

- Alternative: protected polymorphism analysis:

  - Near \( p = 0 \), \( p' = \sum_{i=1}^{n} \left( \frac{c_i}{v_i} \right) p = p \frac{1}{\tilde{v}} \)
    
    where \( \tilde{v} = \frac{1}{\sum_{i=1}^{n} \left( \frac{1}{v_i} \right)} \) 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 \( p = 0 \) 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: \( \bar{v} \leq \bar{\nu} = \sum_{i=1}^{n} c_i \nu_i \).

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).