Lab Exercise #7: Genetic Drift *versus* Selection in *Cepaea nemoralis*

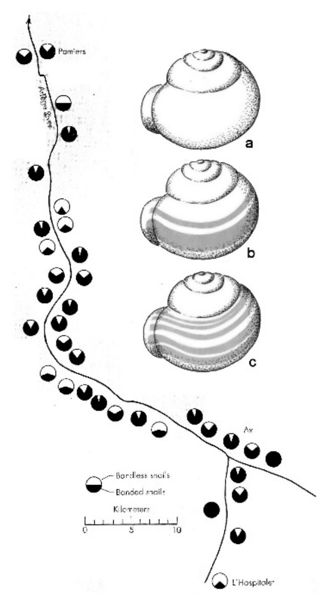
Cepaea nemoralis, the European Land Snail, is the classic example of the interplay of ***deterministic*** (natural selection) and ***stochastic*** (genetic drift) factors in evolution. EB Ford's studies on this species are the part of the foundation of Ecological Genetics.

C. nemoralis is widespread in western Europe and southern Great Britain. Related species occur in North America, including Newfoundland. *C. nemoralis* is phenotypically polymorphic: individuals have dark or light shells, and from zero to five bands [below, left]. The polymorphism is controlled genetically: *Dark* alleles are dominant to *Light*, and *Un-banded* alleles are dominant to *Banded*.

Cepaea, like *Biston betularia* moths, is subject to visual predation by birds, based on **crypsis**. Song thrushes spot the snails, pick them up, and drop them on flat rocks ("*thrush anvils*"). The shells break open, and the birds eat the contents. Local predation patterns can be measured by examining the phenotypes of *broken* shells at the anvils. The incidence of the various shell types varies *inversely* according to habitat type and degree of crypsis. Broken *light* shells are more common at anvils in *dark* "*brown*" backgrounds [*e.g.*, bare soil], and less common in *light* "*green*" backgrounds [*e.g.*, grasslands]. Broken *Un-banded* shells are more common at anvils in "*high contrast*" habitats [*e.g.*, short grass fields], and are less common on "*low contrast*" habitats [*e.g.*, deep-shadow woodlands]. Much of the variation among colonies seems to be correlated with these two environmental variables. Snails that stand out visually are preferentially eaten by thrushes and show up more frequently as broken shells at the anvils: the more cryptic do not show up as frequently.

Genetic drift also plays a role. The snails occur in colonies of tens to thousands of individuals. Colonies are well-separated micro-geographically, and gene flow between colonies is severely limited by the creatures' sedentary habits. Different shell phenotypes typically co-exist in most colonies, and their frequencies vary greatly over short distances, both along and on either side of rivers [below, right], not always in accordance with expectations from crypsis. There is greater genetic variation (measured as variance in shell patterns) among *smaller* colonies than *larger*, consistent with the expectations of genetic drift. Some small colonies may even become '*fixed*' for '*lost*' for the '*wrong*', non-cryptic shell pattern.

This exercise models the behavior of alleles in larger and smaller populations, with greater or lesser degrees intensity of selection, and thus the interplay of **stochastic** Genetic Drift and **deterministic** Natural Selection in determining the degree of genetic polymorphism in natural populations.



Biol4250 – Evolutionary Genetics **Nov 2025**

MatLab simulation of **Natural Selection** / **Genetic Drift** in **finite populations**

Each snail colony is a *finite* population of **N individuals**, with a total of **2N alleles**. The ***Bnd*** locus has two alleles **A** & **B** (**B**anded). **B** is genetically **recessive**. The colonies are in a *uniform, low-contrast environment*, in which *Banded* snails are at a selective *disadvantage*, **s = (1-W2)**. [We count *live snails*, not *broken* shells]. **WriFish** traces **q** = f(**B**) over time. Recall that

f(**AA**) = $(1 - q)^2$ f(**AB**) = $2(q)(1 - q)$ f(**BB**) = q^2 The variance *between generations* is $\sigma^2_q = (q)(1-q)/2N$.

(1) The *deterministic expectation* of **B** subject to *negative* selection is that f(**B**) → **0.0**. In any *finite* population size **N**, the allele frequency is subject to *stochastic genetic drift* over time. We ask: Under what combinations of **s** and **N** can a *disadvantageous* **B** allele drift stochastically to f(**B**) → **1.0**.

(2) For a new **B**rown allele arising in a green snail colony on *brown, bare soil*, **B** will be at a selective *advantage*. We ask the inverse question: does an *advantageous* **B** allele necessarily increase to f(**B**) → **1.0** ? Under what combinations of **s** and **N**?

(3) For (2) above, consider selection with **semi-dominance** such that **W0 = 1**, **W1 = 1 – 0.5 s** , and **W2 = 1 - s**

WriFish Model

The **MatLab** program **WriFish** generates **two** alleles **A** or **B**, according to **q = f(B)**. The two alleles define an individual with genotype **AA**, **AB**, or **BB**. The appropriate fitness **W0**, **W1**, or **W2** is applied as a *probability of survival* of the newly-generated individual genotype. If the individual does not '*survive*', two new alleles are sampled, and the fitness test applied until one survives: this individual is added to the population. This "***Draw & Replace***" is repeated until the sample reaches **N** individuals, in each of N_{pop} populations. **N** remains constant between pre- & post selection populations. **WriFish** corresponds to a "***soft selection***" model of Natural Selection.

Instructions

1) Click on the program **WriFish.m**. This will open the program in **MatLab**.
Click the green **Run** button [center, top].
A new screen will appear, with a box "**Model Parameters**" in the upper left corner.
The screen will ask for the following. Run the scenarios in **Part 3** :

Initial value of **q** = f(**B**) in the range **0.0 < q < 1.0**
Population size **N** for any value **≤ 100,000**
Number of populations **Npop**, **> 1 ~ 1,000**
Number of generations **Ngen** (see **Suggestion** below)
Fitness values **W0**, **W1**, & **W2** for **AA**, **AB**, & **BB**:
B is **recessive** to **A**; set **W0** = **W1** = **1** for **AA** & **AB**, set **W2= 1 - 2** for **BB** according to the scenario

2) Run the program by clicking on yellow "**Run Evolution**" in the lower left corner

The program will plot f(**B**) = **q**, for each population, for the model parameters.
For each of the following scenarios, run **10** [or more] **simulations simultaneously**
The trajectories of f(**B**) are plotted for all populations in the simulation.
Counts of populations that go to **q = 0** or **1** (loss & fixation of **B**) are reported in the left column.

Scenarios

3) **Run the scenarios** on the **last page** of this handout. Start the simulations at **q = 0.5**, then repeat for **0.1 & 0.9**

Record the count of populations that go to **fixation** (**q = 1**) or **loss** (**q = 0**)
Note the approximate range of generations required for loss / fixation
Copy & Save **JPG** graphs to a drawing program or **WORD** for your discussion:
Label graphs with starting parameters: suggested format **q.1 N10000 s.1.jpg**
Suggestion: Run **Npop = 100** or **1,000** to get a more accurate ratio of fixation & loss;
repeat with **Npop = 10 or 100**, each repeated **10** times to see the variance.

1. Over time, in which colony size are allele frequencies more variable ***among populations***? Why?

2. Is a species that comprises ***multiple small populations*** more variable than one composed of a ***few large populations***? Explain.

3. Can **deleterious** alleles go to **fixation** in a population? Under what circumstances? Explain.

4. What is the ***ultimate fate*** of alleles, even in '*large*' populations? Explain.

5. For what values of **s** and **N** do the effects of selection and drift achieve an approximate equilibrium of **Δq** ?

q = 0.1									
N= s=1-W2	0.1	0.01	0.001	0.0001	0	-0.0001	-0.001	-0.01	-0.1
10									
100									
1,000									

q = 0.5									
N= s=1-W2	0.1	0.01	0.001	0.0001	0	-0.0001	-0.001	-0.01	-0.1
10									
100									
1,000									

q = 0.9									
N= s=1-W2	0.1	0.01	0.001	0.0001	0	-0.0001	-0.001	-0.01	-0.1
10									
100									
1,000									

N= s=1-W2	0.1	0.01	0.001	0.0001	0	-0.0001	-0.001	-0.01	-0.1
10									
100									
1,000									