Ecological Genetics

On-line Study Guide

INDEX

The Index contains a list of all **Study Guide** topics available for the ***Ecological Genetics*** program. For information on how to use the **Study Guide**, press F1.

Overview

The ***Ecological Genetics*** programs is based on the breeding group models of gene dynamics. These models take a fundamentally different view of how populations are structured. The classical view of population structure involves an assemblage of organisms that may have subpopulations that are somewhat isolated from each other. This isolation leads to differentiation among the subpopulations, but within each subpopulation individuals are assumed to mate at random. Thus, in the classical view, genetic drift acts to differentiate the subpopulations and the chance mating of related individuals within the subpopulations leads to an accumulation of inbreeding.

There have been a few theoretical models to remove the constraint of random mating, which obviously doesn't occur in all species. Among these are the models of Crow and Denniston (1988) and Caballero and Hill (1992). Each of these take the tact that with non-random mating, individuals are more or less likely to mate with related individuals than chance predicts. Unfortunately, this approach often ignores the biology of the organisms understudy. Behavioral ecologists have often argued that population geneticist overestimate the levels of inbreeding because they don't recognize that organisms often avoid mating with close relatives. Thus, many of these theories don't fit well with behavioral observations, and they have lead to some misinterpretations of empirical data in population genetics studies.

The reason that the above approaches have lead to some problems interpreting genetic data is two fold. First, as behavioral ecologists have shown, many species neither mate at random, nor do they mate with close relatives. They often from breeding groups with specific patterns of mating and dispersal that minimize the probability that close relatives will mate. These breeding groups are the second problem with previous approaches to studying gene dynamics because the represent a real level of structure in a population that is often ignored when biologist sample for genetic studies. Because each level of structure, from individual to species, is a potential avenue for the loss of genetic information, studies that ignore a level in the middle of a hierarchy will misrepresent the true gene dynamics.

This is where the breeding group models differ. They were developed to take social groups, which is where matings occur, into account. When one applies these models to real data, one often finds is that inbreeding is actually lower than expected from random mating at the level of breeding groups. Of course one could take the social structure into account when sampling and use the previously mentioned models to estimate the loss of genetic variation, but the breeding group models do much more that.

The idea behind the breeding group models was to incorporate how social

affiliations would affect the relatedness of individuals in breeding groups and the entire population, much like one would do with a pedigree. These models use information about the demography and ecology of the organism to determine the correlation of genes within individuals (F), between individuals within a breeding group ( ), and between individuals from different breeding groups ( ). These gene correlations simultaneously give a researcher information about how genetic variation is partitioned within and among the various levels of structure, and about the rate at which genetic variation is lost.

The ***Ecological Genetics*** program takes behavioral information that a researcher collects and from that calculates the gene correlations, fixation indices, and effective population sizes for the population. The model is based on the papers by Chesser (1991a), Chesser (1991b), Chesser et al. (1993a), and Sugg and Chesser (1994). You can find a review of the concepts and application of these models in Sugg et al. (1996).

Distribution Regulations

This copy of ***Ecological Genetics*** is copyrighted shareware. You, as a end user, are entitled to distribute copies of this program under two conditions: 1) you do so free of charge, and 2) you include all the files that came on the original disk, or in the original compressed file if you downloaded the program from a bulletin board. Since this software is targeted for scientists, resource managers and students, every attempt has been made to make the registration price reasonable. So, if you find this software useful, please register it. If you would like more information on registration, see the order form

Scope: Socially Structured Populations

Any behavioral ecologist can tell you that most organisms don't live in randomly mating populations. In fact, most have behavioral, morphological or physiological mechanisms to avoid mating with themselves or close relatives. If this is the case, then why is it most empirical population geneticists report moderate to high levels of inbreeding in populations?

There is one simple answer, we usually ignore social structure when we sample a population. So, why does this make a difference? Isn't inbreeding the same regardless of the social nature of the beast? No, because inbreeding is always relative to some mathematical expectation. Usually the expectation is that of gene correlations among randomly breeding individuals in a population of the same size. This, after all, is exactly what the inbreeding coefficients F_{IS} and
F_{IT} are telling you.
F_{IS} is the level of inbreeding of individuals in a subpopulation compared to what it would be if they had mated a random.
F_{IT} is the same for the total population. When either of these is negative, then it indicates individuals are less inbred than expected at that level of population structure. They should be positive when individuals are more inbred and around zero when there is no inbreeding.

One of the interesting things that seems to have been missed or explained away by many population geneticists is that we should expect is a negative F_{IS} under most situations. You probably know from your own experience that population geneticists often report positive, or at best zero, values for this index. Doesn't this fly in the face of what behavioral ecologists have been telling us about the avoidance of matings with close relatives? The answer to this seeming paradox is really quite simple. If the subscripts I, S, and T, refer to individuals, the lowest level of organization, and the total population, then what is usually reported as
F_{IS} is really more similar to an average
F_{IT} for lots of breeding groups in different populations.

Positive F_{IS} values are often explained by a Wahlund effect (1928). This concept says that
F_{IS} will be positive when two randomly mating populations are mixed. This mixing can result from natural causes, such as the loss of a former barrier that isolated the populations, or it can be caused by sampling two populations without recognizing them as separate units. Still, lots of researchers have argued that they are sure that they have not mixed populations in their samples, so the Wahlund effect can't explain their positive
F_{IS} values. What they don't realize is that this concept applies to all levels of structure, not just populations or subpopulations. If breeding groups

represent randomly mating subunits, the mixing of these will result in an upward bias of the F_{IS}. And if breeding groups represent individuals that are actually avoiding mating with close relatives, the bias will still occur. And breeding groups can be as small as a monogamous pair!

What the breeding group models have shown is that if you take the social structure into account, F_{IS} is usually strongly negative, F_{IT} is usually around zero, and F_{ST} is usually strongly positive. Thus, what the behavioral ecologists tell us is going on can be seen in the genes. The only reason this works is that we let the behavior tell us what the structure is, and then we sample accordingly. More information on these expectations and the interplay of behavior and genetics can be obtained from Chesser (1991a), Chesser (1991b), Chesser et al. (1993a), Chesser et al. (1993b), and Sugg and Chesser (1994).

The breeding group models were designed to use ecological and demographic data. This is because the mathematical proofs requires that one show how these behaviors affect the gene diversity. Such an approach can be an advantage for ecologists who know nothing about electrophoresis or DNA techniques. With much of the information they usually collect they can calculate the expected gene correlations, fixation indices, and effective population size. The model can be applied to genetic data as well, because you can estimate the gene correlations directly. This version of ***Ecological Genetics*** does not use genetic data, but a future version may if enough people express a need for it.

Scope: Ecological Data

The ***Ecological Genetics*** program is geared for the person who wishes to have knowledge about gene diversity in a population. This population can be natural, managed, or captive. What the user needs is to have some information, or at least some educated guesses, about the ecology and demography of the organisms at hand.

There are five broad categories of parameters that are needed to successfully calculate the Ecological Genetics. These are described in detail in the Study Manual section devoted to each:

Census Data	Includes the number of groups and individuals
Female Reproduction	The number of offspring a female produces
Male Reproduction	Male contributions to matings
Multiple Paternity	Number of sires siblings have
Dispersal	Movement of males and females

Once you have reviewed these parameters, you will find that most behavioral ecologists have good estimates of these. Additionally, managers of captive breeding programs will also be able to glean the information from their stud books. But these aren't the only cases in which one my want to use the program. We showed that one could get excellent estimates of the fixation indices and effective sizes using only values already published in the literature (Sugg et al., 1996). You could also play what-if games with the program to see what strategies minimize the loss of genetic variation, perhaps as part of a classroom exercise. I hope that the use of this program will not only provide better estimates of gene dynamics, but it will also provide a better understanding of how genetic variation is apportioned and maintained in socially structured populations.

Scope: Genetic Data

At this time, the ***Ecological Genetics*** program does not support the analysis of genetic data. Although it is possible to incorporate this type of analysis into the program, once one recognizes social structure and samples accordingly, one can obtain estimates of the fixation indices from other analysis programs that many of you are probably already using. Once you have estimates of these parameters, you can use equation 25 in Sugg and Chesser (1994) to estimate the effective population size. Although this is only an approximation, it is usually a good one and it is the best this program could do without the user also having some ecological data.

If you would like to see the analysis of genetic data incorporated into this program, please let us know.

Concept: Ideal Populations & Effective Sizes

An ideal population is a theoretical concept with the following attributes:

1) Separate Sexes
2) Diploid Genome
3) No Selection
4) No Mutation
5) No Migration
6) Random Mating
7) Constant Population Size

Often, effective population sizes are viewed as the size of an ideal population that would loss genetic variation at the same rate as the population understudy. This view of effective sizes can be misleading, because it is possible for an effective size to be a number that no population could really attain. For example, if a population is not losing genetic variation, its effective population size is *infinity*! If the population is gaining genetic variation, then its effective size is *negative*! Obviously, no real population can be infinitely large or negative in size. Thus, one should view the effective size as a rate of change in variation. High positive values indicate a slow rate of loss of variation, but strongly negative values indicate a slow increase in genetic variation.

This does not mean that the concept of ideal populations is useless. The characteristics of an ideal population are all mechanistic attributes that tell us something about what is important for determining the rate of loss in genetic variation. For example, when individuals don't mate at random, then we expect the effective size to be larger (when they avoid inbreeding) or smaller (when they do inbreed) than the real population size.

It is these attributes of an ideal population which form the basis of how models of gene dynamics are developed. Theoreticians try to determine how the biology of organisms affect things like random mating, and in turn how nonrandom mating changes the gene dynamics.

Concept: Inbreeding

Everyone has a gut feeling about what inbreeding is, yet there is some confusion about things like inbreeding coefficients. Is homozygosity an indication of inbreeding? If so, then the correlation of genes within individuals (which is roughly 1 minus the heterozygosity) should be also. The answer to this question is a resounding maybe!

If one could tell every gene in a population by the ancestor it came from, then an individual that had two of the same genes at the same locus would show some level of inbreeding. However, our ability to do this is severely limited with allozyme data, and it is not an easy task even for many of the DNA techniques.

Inbreeding is the result of two related individuals mating. If they are related, then they are said to have some level of coancestry. The correlation of the offspring's genes is directly related to the coancestry of their parents. But using *F* as an indication of inbreeding is meaningless because it really tells us nothing about the opportunity of other possible matings. In a finite population, *F* will always increase (given the population size and dispersal are constant) and inbreeding of this type can't be avoided. When we *talk* about inbreeding in the context of loss of genetic variation, we are really comparing it to the expected level of gene correlations within individuals if they were all mating at random.

Thus, if you want to know if inbreeding is going on it a population, you really are interested in the inbreeding coefficients or two of the three fixation indices for this type of population. These coefficients are F_{IS} and F_{IT}. When either is negative, the individuals are less inbreed that expected with random mating at that level of population structure. The opposite is true if these coefficients are positive.

To avoid some of the confusion about inbreeding we use *F* to indicate the correlation of genes within individuals, F_{IS} to indicate the level of inbreeding relative to members of the same breeding group, and F_{IT} to indicate the level of inbreeding relative to individuals of the entire population.

Concept: Genetic Drift

How does genetic drift differ from inbreeding? In some ways it doesn't. One can view the fact that individuals can't avoid mating with relatives in a finite population as a sampling problem. And sampling is the key to drift; not your sampling of the population, but the individuals sampling the genes to pass on to the next generation.

Each time an individual mates, it uses only a sample of the gametes that it can produce. In turn, these gametes don't represent all the possible combinations of genes that are in an individual because of segregational variance.

How we most often view drift is the chance fixation of a gene due to random events. What it means for real populations is, even if they can successfully avoid inbreeding (which they can't), they would become fixed for some genes by chance alone. If there are more than one population, subpopulation, or breeding group, they may be fixed for different genes. This process is termed differentiation.

Just like inbreeding, drift is a possible avenue for the loss of genetic variation. How does that fit into the concept of an ideal population? One way to counteract the differentiation force of drift is to mix the populations. In natural populations this is accomplished by migration (or more appropriately, gene flow). When individuals move from one population unit to another, they take their genes with them. If they come from a unit that is fixed (or nearly so) for one gene to a unit fixed for another gene they decrease the differentiation by making the average individual in both units more similar.

This brings up a very important point about genetic variation. There is variation at each level of structure in a population (i.e., individuals, breeding groups, and population). While drift leads to a loss of variation *within* a breeding group, it also leads to an increase in the variation *among* breeding groups. Thus, each level in the hierarchy acts as a reservoir for genetic variation. The only avenue for unrecoverable loss is at the highest level of the structure, and that results simply from a finite population size. If individuals can successfully minimize inbreeding within breeding groups while simultaneously differentiating, then they can maintain the highest possible effective population size. Therefore, such a population will be losing genetic variation at the slowest possible rate given their census size. Can a real population meet these criteria? Well Sugg et al. (1996) show that prairie dogs in fact do just that!

Concept: Coancestry

The concept of coancestry is not new to population genetics, it is just rarely used. Most often, one sees the use of coancestry (or kinship) in the behavioral literature, but it is an important part of pedigree analyses. In fact, everything that the ***Ecological Genetics*** program does can be accomplished with pedigrees, provided you have a pedigree for the *entire population*! Pedigrees, after all, will give you the *real* correlations of genes within individuals, not just estimates of their expected values. Fortunately for those who don't have good pedigrees, or lack the intestinal fortitude to slog through one, the breeding group model gives a very good approximation to results from a pedigree.

So what is coancestry? Simply put, it is the correlation of genes between two individuals. For the breeding group model, this type of correlation takes two forms. There is the correlation of genes between individuals within the same group (α) and between individuals from different groups (β). You will see from the papers that have been cited here, and for most of the discussions, we call the correlation within groups the coancestry and that between groups the inter-group correlation. This is simply to avoid the obvious confusion that would arise if we used the same term for both.

Why is coancestry important? Well, like the correlation of genes within individuals, this correlation is a measure of genetic variation. As the correlations get large there is less genetic variation at that particular level of population structure. Thus, if one knows how the gene correlations have changed from one generation to the next, one has an estimate of the loss of variation from that level of structure.

Concept: Apportionment of Variation

Everyone is some what comfortable with the concept of inbreeding and how it relates to loss of genetic variation, but they don't often realize that the same processes that lead to gene correlations (and the loss of genetic variation) also lead to the apportionment of genetic variation within a population. Yet, we all deal with the apportionment of variation every time we calculate the fixation indices.

As was discussed earlier, the inbreeding coeffiecients relate the correlation of genes within individuals to that expected if individuals had mated at random. In fact, that expected level of correlation is exactly one minus total variation at the same level. Each of the fixation indices can be described in the same way.

Before we talk about the apportionment of variation, we first need to see why the gene correlations are just an opposite way of looking at genetic variation. When the correlation of genes increases the variation decreases. You are familiar with this concept if you use correlation analyses. When we normally look at the fixation indices we calculate three heterozygosity values to represent the genetic variation in a population. The first of these is the average individual heterozygosity (H_I), and F is equal to $1 - H_I$. Similarly, we use $\overline{H}_s$ to represent the expected heterozygosity in the average subpopulation (or breeding group) and H_T to represent the expected heterozygosity for the total population. Like F, the other correlations have values that are related to these heterozygosities: $q = 1 - \overline{H}_s$ and $a = 1 - H_T$.

The fixation indices can be calculated from formulae that use these estimates of heterozygosity. One can also solve for the fixation indices in terms of the gene correlations, which are:

$$F_{IS} = \frac{F - q}{1 - q}$$

$$F_{IT} = \frac{F - a}{1 - a}$$

$$F_{ST} = \frac{q - a}{1 - a}$$

Concept: Loss of Variation & Effective Size

The rate of loss of genetic variation has become one of the most important aspects of population genetics for many researchers. Conservation biologists study the impacts of human impacts of the survivability of populations, and genetic diversity plays a role in their analyses. Managers of wildlife species and captive populations are often interested in knowing how the plans they have developed will affect genetic variation as well.

How are the gene correlations and fixation indices related to the loss of genetic variation? As we saw earlier, the gene correlations are a measure of the amount of genetic variation. As time goes on, the gene correlations change, indicating that there has been a change in the amount of variation. The simplest way to calculate effective population size is with the equation $N_e = 1/2\Delta r$, where r is the correlation at some level. One calculates the change in the correlations from the following equations:

$$\Delta F = \frac{F_{t+1} - F_t}{1 - F_t}$$

$$\Delta q = \frac{q_{t+1} - q_t}{1 - q_t}$$

$$\Delta a = \frac{a_{t+1} - a_t}{1 - a_t}$$

The effective size based on *F* is often called the inbreeding effective size (N_{eI}), that based on a is called the coancestral effective size (N_{eq}), and the one for a is called the intergroup effective size (N_{ea}). The intergroup effective size is same as the variance effective size when the population size is stable.

Sugg and Chesser (1994) give the different effective sizes in terms of the parameters used in this models. However, they also recognized that how variation is apportioned among the levels of structure also indicates the rate at which it is lost. Thus, one can get an estimate of the *asymptotic effective size* from the fixation indices and the number of groups.

So what is an asymptotic effective size? It is the value of all the effective sizes when they have converged on a common value. That may seem odd, but that is what is expected when all of ecological variables are constants (i.e., breeding tactics are constant), and the population size is stable. The breeding group model, like pedigree analysis, assumes that all individuals are unrelated (no gene correlations between individuals) and there is no gene correlation within individuals at some point in the past. Each

generation, these gene correlations change until all are increasing at the same rate. When the rate of change in the gene correlations is the same, then the fixation indices attain constant values. Also, because the gene correlations are changing at the same rate, the rate of loss of genetic variation is the same for each level of population structure; hence, all the effective sizes are identical. The asymptotic values are determined by the breeding structure of the population, but the time it takes for asymptotic conditions to be obtained is largely due to the dispersal rates. The lower the dispersal rate, the longer it takes. When there is no dispersal, then each of the effective sizes remains distinct.

There are two things to note about this outcome: 1) it doesn't matter how correlated the genes are at the start of the process, the asymptotic estimate will still be the same, and 2) an asymptote is only achieved if all the parameters are constants and there is some dispersal. When the demography and ecology of the organism changes over time, the effective sizes remain distinctly different.

The Model: Why it was Developed

The breeding group model was developed to better understand the interplay between the ecology of organisms and the dynamics of genetic variation in natural populations. Simply put, the behavior and ecology of organisms result in population structure, not the other way around. With this in mind, one should not use gene diversity to determine the breeding patterns and levels of inbreeding, one should use the breeding patterns to understand the genetics.

It is hoped that this model will not only aid or understanding of how genetic variation is apportioned and lost in natural populations, but that it will also provide biologists with a tool to study natural populations in greater detail.

The Model: Why is it Different

The major difference between the breeding group model and all others is that it recognizes that populations (or subpopulations) are subdivided into breeding groups. If genetic data is obtained with this in mind, the best estimates of inbreeding and rate of loss of genetic variation can be obtained. Furthermore, ecological data can be used to estimate the gene dynamics.

Genetic data will give you a snapshot in time. If one has several generations of data, then these can be used to obtain average values.

Ecological data will furnish all the information for asymptotic values, but they cannot tell a researcher the current values of the gene correlations. Nevertheless, the fixation indices and effective sizes will still be the same because of their asymptotic nature. As with genetic data, one is better off with several generations worth of data.

The Model: When to use it

You should use the breeding group model whenever you think the population you are studying is socially structured. However, if the population turns out not to be subdivided, the model simplifies to the more classical approaches. In other word, you won't go wrong if you use the breeding group model with correctly estimated parameters, but you can go wrong with models that ignore social structure.

The Model: Equations

This is a listing of the most important equations from Sugg and Chesser (1994). To see the equation, simply click on the definition. The parameters used in these equations are described in the sections labeled **Parameters**.

Breeding Parameters

- Probability progeny within a group share the same father
- Probability progeny within a group share the same mother
- Probability progeny of a mother share the same father

Gene Correlations

- Change in *F*
- Change in *q*
- Change in *a*

Fixation Indices

- Inbreeding within Groups
- Inbreeding within the Population
- Differentiation among Groups

Effective Sizes

- Inbreeding
- Coancestral
- Intergroup

Parameters: Census Data

There are three parameters that can be obtained from census data. These are:

s = the number of breeding groups
n = average number of breeding females in each group
m = average number of breeding males in each group

Limits -- values must be greater than or equal to one.

Note: *m* and *n* may also represent the actual number of adult males and females, respectively, if one calculates the means and variances in reproductive success for males and females accordingly.

Parameters: Female Reproduction

These parameters are only estimable from behavioral data or breeding records. They are:

k = the average number of progeny produced that survive to reproduce during the lifetime of a female.
var(k) = the variance in the above number.

Limits -- *k* must be greater than or equal to one and less than or equal to 10. The variance must be greater than or equal to zero, with its upper limit being dynamically determined based on the other parameters.

Note: these parameters are calculated using all the females in the population, regardless of which breeding group they come from.

Parameters: Male Reproduction

These parameters are also only estimable from behavioral or breeding data. The parameters are:

b = the average number of females mated by a male during his lifetime that produce progeny that survive to reproduce.
var(b) = the variance in the above number.
l = the average number of males mated by each female.

Limits -- *b* must be less than or equal to the number of females (*n*) but greater than or equal to one. The limits for the variance are determined dynamically, based on the values of the other parameters, but will always be greater than or equal to zero.

Note: these parameters are calculated using all the males in the population, regardless of which breeding group they come from. Also note that you need not calculate l separately because it is equal to *k*/*p*.

Parameters: Multiple Paternity

Multiple paternity in the breeding group model is considered to include multiple sires of a single brood/litter as well as serial monogamy. The parameters are:

p = the average number of progeny produced by a female that are sired by a single male.
var(p) = the variance in the above number.
l = the average number of males mated by each female.

Limits -- *p* must be greater than or equal to one and less than or equal to *k*. The variance in *p* will always be greater than or equal to zero with the upper limit dynamically determined based on the values of the other parameters.

Note: these parameters are calculated using all the males and females in the population, regardless of which breeding group they come from. Also note that you need not calculate l separately because it is equal to *k*/*p*.

Parameters: Dispersal

The dispersal rates are best obtained from behavioral data or breeding rates. They are:

dm = the proportion of males reproduce in more than one breeding group.

df = the proportion of females that reproduce in more than one breeding group.

Limits -- values must be between 0 and 1, inclusive.

Cyclic Population Dynamics

Many natural populations go through fairly regular changes in population size, often termed population cycles. The consequences of cycling can be profound for gene dynamics, especially for how genetic variation is partitioned. Because these types of populations are nonequilibrial, it is best to use the harmonic mean of the values to estimate the long-term effective sizes. The ***Ecological Genetics*** program does this any time the user chooses cyclic or stochastic variation. These values are reported in the status panels below the results pane.

Populations can be made to change in size by checking the Cyclic box. This action will provide the user with another entry panel which contains edit boxes for the period and magnitude of the cycles, and checkboxes for the parameters that should vary.

The period of the cycle can be between 1 and 100 generations, and the scale is dynamically determined based on the values of other parameters already entered. Using a negative value for the scale will cause the population to initially decline and then increase; using a positive value will initially increase in size. The default values of the period and scale are 1 and 0, respectively. The default values result in no cycling. Additionally, if none of the checkboxes corresponding to the parameters are checked, then the population will not cycle.

Cyclicity is incorporated by changing the variables in the following manners. For the number of females (n), males (m) the formulae are:

$$n_t = n_0\left[1 + Scale \cdot Sin\left(\frac{tp}{0.5\,Period}\right)\right]$$

$$m_t = m_0\left[1 + Scale \cdot Sin\left(\frac{tp}{0.5\,Period}\right)\right]$$

Of course, any time the number of adults changes, then there has been a change in the net reproductive rate. Therefore, the number of surviving offspring (k) must also change. The mean and variance in this parameter are given by:

$$k_t = 2\frac{n_t + m_t}{n_{t-1} + m_{t-1}}$$

$$\mathrm{var}(k)_t = k_t\left[f_{f_0}(k_t n_t - 1) - k_t + 1\right]$$

When the number of offspring per male (p) and the number of mates per male (b) are made cyclic, their means and variances are determined by:

$$p_t = p_0 \frac{k_t}{k_0}$$

$$\mathrm{var}(p)_t = p_t \left[f_{w_0} (k_t - 1) - (p_t - 1) \right]$$

$$b_t = b_0 \frac{n_t + m_t}{n_0 + m_0}$$

$$\mathrm{var}(b)_t = \frac{k_t n_t f_{m_0} (n_t - 1)}{m_t p_t} - b_t (b_t - 1)$$

Cycling *p* and *b* results in minimal (or no) variation in the breeding parameters. Thus, one should choose to cycle these parameters if it is believed that the organism in question has fairly fixed tactics for mating (i.e., constant degree of polygyny), and not cycle them if mating tactics are density dependent. One interesting comparison to make is to make a run with *n*, *m*, *p*, and *b* all cyclic and another with only *n* and *m* cyclic. You will see that the gene correlations and fixation indices remain fairly constant with the former but not the latter. However, the effective population sizes will vary under either comparison because they are so sensitive to changes in the gene correlations.

For dispersal rates, the cycles are given by:

$$d_{m_t} = 2 d_{m_0} Min\left(d_{m_0}, 1 - d_{m_0}\right)^{\frac{m_t - m_0}{m_0}}$$

$$d_{f_t} = 2 d_{f_0} Min\left(d_{f_0}, 1 - d_{f_0}\right)^{\frac{n_t - n_0}{n_0}}$$

When the user chooses to make the population cycle, an edit box appears to allow the user to enter the number of generations to carry out the analysis. This is necessary because cyclic populations do not reach an asymptote. The default value is 250 generations (the maximum allowed by the program). The user will probably find it most convenient to change this to about three time the period of the cycle.

Limits

Time -- 1 to 250
Period -- 1 to 100
Scale -- $\frac{1 - m}{m}$ to $\frac{1 - m}{-m}$

Stochastic Variation

Generation-to-generation variation in populations is the rule rather than the exception in natural populations. These vagaries lead to nonequilibrim conditions, so natural populations rarely attain the asymptotic conditions of the general model. However, the fluctuations have only transitory effects on the gene dynamics, and the results tend to track those for asymptotic conditions. For this reason, the harmonic means of the effective sizes are the best estimate of the long-term rate of loss of genetic variation. These estimates are provided in the status panels below the results pane.

The user can incorporate stochastic variation by checking the Stochastic Checkbox. Checking this box will open an additional part of the parameter entry panel and a generations edit box. The generation box is used to enter the time to carry out the analysis. The default time is 250 (the maximum allowed by the program), but the user will probably find it better to use fewer generations.

In the new entry panel the user needs to provide the level of variation to use in the variance edit box. This value scales the variation of the parameters to their maximum possible values (e.g, a value of 100 means the variation is equal to 100% of the maximum variation). Additionally, one can choose the correlated check box to incorporate autoregression. Autoregression causes the values of one generation to be correlated with the value of the previous generation. This is used to prevent values of being very large one generation and very small in the next. It is up to the user to decide which approach (correlated or uncorrelated) is best. If the user chooses correlated, then an edit box is displayed for entry of the degree of correlation, a value for -0.5 to 0.5 (0 is the default).

To incorporate stochastic variation a random deviate $\left(\tilde{N}_{(0,z)}\right)$ is calculated. This deviate has a normal distribution with mean of zero and variance of $z = V\sqrt{\frac{x_t - 1}{3}}$, where *V* is the variance factor entered in the entry panel and *xt* is the present value of either *n* or *m*. This normal deviate is used to calculate then error term $e_t = r e_{t-1} + \tilde{N}_{(0,z)}$, where *r* is the correlation value from the entry panel. And the new values for them mean number of females and males are given by:

$$n_{\acute{F}} = n_t + e_t$$
$$m_{\acute{F}} = m_t + e_t$$

As was discussed for cyclic variation, the mean and variance in progeny number must also change if the population size is changing:

$$k_t = 2\frac{n_t + m_t}{n_{t-1} + m_{t-1}}$$

$$\mathrm{var}(k)_t = k_t\left[f_{f_0}(k_t n_t - 1) - k_t + 1\right]$$

The mean and variance in the *p* and *b*, as well as the mean dispersal rates, are also calculated in the same way as for cyclic variation:

$$p_t = p_0\frac{k_t}{k_0}$$

$$\mathrm{var}(p)_t = p_t\left[f_{w_0}(k_t - 1) - (p_t - 1)\right]$$

$$b_t = b_0\frac{n_t + m_t}{n_0 + m_0}$$

$$\mathrm{var}(b)_t = \frac{k_t n_t f_{m_0}(n_t - 1)}{m_t p_t} - b_t(b_t - 1)$$

$$d_{m_t} = 2d_{m_0}Min\left(d_{m_0}, 1 - d_{m_0}\right)^{\frac{m_t - m_0}{m_0}}$$

$$d_{f_t} = 2d_{f_0}Min\left(d_{f_0}, 1 - d_{f_0}\right)^{\frac{n_t - n_0}{n_0}}$$

As with cyclic variation, the user must check the parameters for which stochastic variation is to be applied. Making *p* and *b* cyclic is the same as making breeding tactics constant; therefore, only slight variations in the gene correlations and fixations indices will be seen. The default values for the variation and correlation result in no stochasticity, as is also the case when no parameters are checked.

Limits

Time -- 0 to 250
Variance -- 0 to 100
Correlation -- -0.5 to 0.5

Generations

Generations in the ***Ecological Genetics*** program refer to the number of generations you wish the program to calculate the gene correlations for. This *does not* refer to the generation time of the organisms understudy. The model assumes that there are nonoverlapping generations. Although this may seem to be a problem for most organisms, Hill (1979) has shown that having overlapping generations doesn't alter the ultimate values of the estimates in equilibrium populations, it simply increases the time necessary to attain them.

The ***Gene Dynamics*** program searches for the asymptote and stops when it is found. When you choose either stochastic or cyclic dynamics, the program will simply runs until for time you specify. When choosing a generation time, keep in mind that the larger this value is, the longer the program will take to run, the longer it will take to display charts, and the more memory and resources the program will consume.

Limits -- values must be in the range of 1 to 250.

Breeding Parameter

The breeding parameters make up a crucial part of the breeding group models, and they are therefore important in the ***Ecological Genetics*** program. Each of the breeding parameters defines a probability of an event that effects the relationship of organisms in the population. The probability that randomly chosen progeny in a breeding group share the same father (f_m) is the degree of genetic polygyny. Values near zero indicate little polygyny, and values near 1 indicate that a single male is the sire of all the offspring. The probability that randomly chosen progeny in a group share the same mother (f_f) is termed the degree of single maternity. Values near zero indicate all offspring in the group have different mothers, and values near one indicate they share the same mother. The probability that offspring of a single mother are the product of a single father(f_w) is the degree of single paternity. Values near 1 indicate that a female is successfully mating with one male, while values near zero indicate females are successfully mating with all the males in the group. The effects of the mating systems on gene dynamics have been covered in Chesser (1991a), Chesser (1991b), Chesser et al. (1993a), and Sugg and Chesser (1994).

Because these parameters are so important, the user should be aware of their values. Each time the user makes a change to the parameter entry panel, the values of these parameters are recalculated and displayed on the tool bar. If any of these parameters are outside of the allowable range (0 to 1), the run button is disabled. This prevents the program from crashing or producing unrealistic results. The parameters also help the user make guesses when they are uncertain about some of the parameter values.

Parameters: Making Guesses

In many cases it will be almost impossible to obtain estimates for all of the necessary parameters. This is not a reason to panic, there are a couple of rules of thumb.

1) If you don't know a parameter, make an educated guess. If the results seem outlandish, try again. You can play what-if games that will help you understand the roles that reproduction and dispersal play in gene diversity.

2) If you don't know a parameter, use the classical assumptions. For example, for the Crow and Denniston's (1988) model you would use $s = 1$, $dm = df = 1$, $m = n$, $b = 1$ and the variance in b is zero. For random union of gametes you can also assume $p = 1$ and the variance in that number is zero.

3) If you know the mean of a number, but not the variance, assume a Poisson distribution where the variance and the mean are equal. The variance in most of the parameters is usually less than or equal to this value in real populations, so you are likely to bias your results in a conservative manner using this tactic.

4) Use the breeding parameters to help you. If you know a species tends to be highly polygynous, then f_m should be relatively large. If you know there is little or no multiple paternity, then f_w should be around 1. If there is a dominant female that does most of the reproduction, then f_f should be around 1.

The most important thing you can do is use as many years of data to estimate the parameters as you can. One year will only give you a picture of what would happen if things stayed the way they are. But if you use values averaged over some year the asymptotic values will be close approximations of the long-term values even when the parameters show generation-to-generation variation. Incorporation of stochastic variation will allow you to see what changes in these parameters will do.

Crow, J.F., and C. Dennistion. 1988. Inbreeding and variance effective population numbers. *Evolution* **42:**482-495.

Caballero, A., and W.G. Hill. 1992. Effective size in nonrandom mating populations. *Genetics* **130:**909-916.

Chesser, R.K. 1991a. Gene diversity and female philopatry. *Genetics* **127:**437-447.

Chesser, R.K. 1991b. Influence of gene flow and breeding tactics on gene diversity within populations. *Genetics* **129:**573-583.

Chesser, R.K., O.E. Rhodes, Jr., D.W. Sugg, and A.F. Schnabel. 1993a. Effective sizes for subdivided populations. *Genetics* **135:**1221-1232.

Sugg, D.W., and R.K. Chesser. 1994. Effective population sizes with multiple paternity. *Genetics* **137:**1147-1155.

Sugg, D.W., F.S. Dobson, R.K. Chesser, and J.L. Hoogland. 1996. Population genetics meets behavioral ecology. *Trends in Ecology and Evolution* **(in press)**.

Chesser, R.K., D.W. Sugg, O.E. Rhodes, Jr., J.M. Novak, and M.H. Smith. 1993b. The evolution of mammalian social structures. *Acta Theriologica* **38(Suppl.2):**163-174

$$N_e @ \frac{4s - 3F_{IT} - 1}{6(F_{ST} - F_{IT})}$$

$$F_{IS} = \frac{\overline{H}_s - H_I}{\overline{H}_S}$$

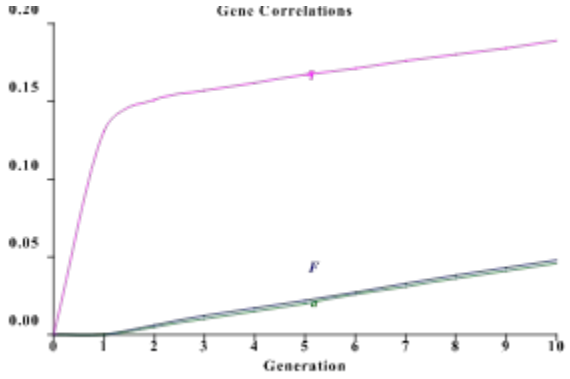
$$F_{IT} = \frac{H_T - H_I}{H_T}$$

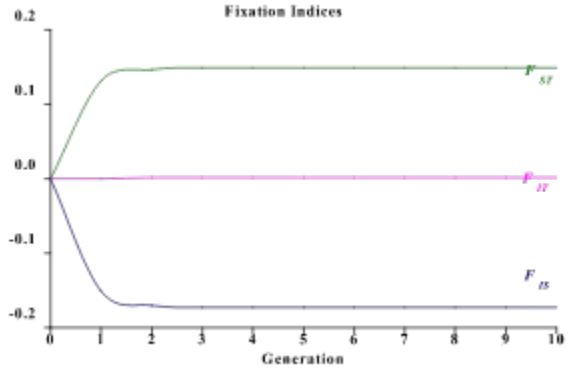
$$F_{ST} = \frac{H_T - \overline{H}_s}{H_T}$$

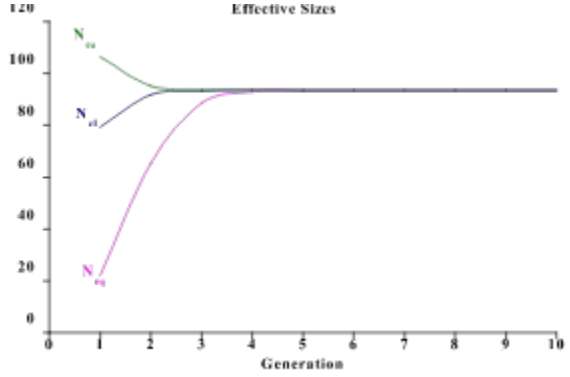
$$F_{IS} = \frac{F - q}{1 - q} = \frac{(1 - H_I) - (1 - \overline{H}_S)}{1 - (1 - \overline{H}_S)} = \frac{\overline{H}_s - H_I}{\overline{H}_S}$$

$$F_{IT} = \frac{F - a}{1 - a} = \frac{(1 - H_I) - (1 - H_T)}{1 - (1 - H_T)} = \frac{H_T - H_I}{H_T}$$

$$F_{ST} = \frac{q - a}{1 - a} = \frac{(1 - \overline{H}_S) - (1 - H_T)}{1 - (1 - H_T)} = \frac{H_T - \overline{H}_S}{H_T}$$


Gene Correlations
0.20
0.15
0.10
0.05
0.00
F
0
1
2
3
4
5
6
7
8
9
10
Generation


Fixation Indices
0.2
0.1
0.0
-0.1
-0.2
0
1
2
3
4
5
6
7
8
9
10
Generation
F ST
F IT
F IS


Effective Sizes
120
100
80
60
40
20
0
0
1
2
3
4
5
6
7
8
9
10
Generation

$$f_m = \frac{m\left[s_b^2 + b(b-1)\right]}{1n(n-1)}$$

$$f_f = \frac{s_k^2 + k(k-1)}{k(kn-1)}$$

$$f_w = \frac{1\left[s_p^2 + p(p-1)\right]}{k(k-1)}$$

$$F_{t+1} = q_t + \left(1 - \frac{1}{s}\right)\left(d_m + d_f - d_m - d_f\right)\left(a_t - q_t\right)$$

$$q_{t+1} = \frac{f_f(1+f_w)+(1-f_f)f_m}{8}(1+F_t) + \frac{\left(2-\left(1-\frac{1}{s}\right)(d_m+d_f-d_m d_f)\right)+f_f(1-f_w)\left(1-\left(1-\frac{kn-1}{kns-1}\right)d_m\right)+(1-f_f)\left(2-f_m-\left(1-\frac{kn-1}{kns-1}\right)(d_m(1-f_m)+d_f)\right)}{4} q_t + \frac{\left(1-\frac{1}{s}\right)(d_m+d_f-d_m d_f)+(1-f_w)\left(1-\frac{kn-1}{kns-1}\right)d_m+(1-f_f)\left(1-\frac{kn-1}{kns-1}\right)(d_m(1-f_m)+d_f)}{4} a_t$$

$$a_{t+1} = \frac{2\dfrac{d_m + d_f - d_m d_f}{s} + \dfrac{kn-1}{kns-1}\left(d_m - d_f\right)}{4} q_t + \left(1 - \frac{1}{2s}\left(d_m + d_f - d_m d_f\right) - \frac{kn-1}{4(kns-1)}\left(d_m - d_f\right)\right) a_t$$

$$F_{IS_i} = \frac{F_i - q_f}{1 - q_f}$$

$$F_{IT_r} = \frac{F_t - a_t}{1 - a_t}$$

$$F_{ST_i} = \frac{q_i - a_i}{1 - a_i}$$

$$N_{cI} = \frac{4}{\left(1 + f_w\right)f_f\left[1 - \left(1 - \frac{1}{s}\right)\left(d_m + d_f - d_m d_f\right)\right]}$$

$$@\frac{1 - F_{IT}}{2\left[1 - \left(1 - \frac{1}{s}\right)\left(d_m + d_f - d_m d_f\right)\right] + F_{ST} - F_{IT}}$$

$$N_{eq} = \frac{4}{\left[f_f\left(1+f_w\right)+\left(1-f_f\right)f_m\right]\left(1+F_t\right)} + \frac{2}{\left\{2\left[1-\frac{1}{s}\left(d_m+d_f-d_m d_f\right)\right]+f_f\left(1-f_w\right)\left[1-\frac{kn-1}{kns-1}d_m\right]+\left(1-f_f\right)\left[2-f_m-\left[1-\frac{kn-1}{kns-1}\left(d_m\left(1-f_m\right)+d_f\right)\right]\right]\right\}t} + \frac{2}{\left\{2\left[1-\frac{1}{s}\left(d_m+d_f-d_m d_f\right)\right]+f_f\left(1-f_w\right)\left[1-\frac{kn-1}{kns-1}d_m\right]+\left(1-f_f\right)\left[1-\frac{kn-1}{kns-1}\left(d_m\left(1-f_m\right)+d_f\right)\right]\right\}t}$$

$$N_{ea} = \frac{4}{(1 - f_w) f_f \left[\frac{d_m + d_f - d_m d_f}{2s} + \frac{(kn - 1)(d_m + d_f)}{4(kns - 1)} \right]}$$

$$@ \frac{1}{2F_{ST} \left[\frac{d_m + d_f - d_m d_f}{2s} + \frac{(kn - 1)(d_m + d_f)}{4(kns - 1)} \right]}$$

Hill, W.G. 1979. A note on the effective population size with overlapping generations. *Genetics* **92:**317-332.

Wahlund, S. 1928. Zusammensetzung von population und korrelationserscheinung vom standpunkt der vererbungslehre aus betrachtet. *Hereditas* **11:**65-105.