BIOB480/BIOE548 notes 8/27/2024

Introduction

- Request: Please say your name before speaking in class to help me learn!
- Email reminder: not D2L, please

Course topics

Conservation genetics is a relatively young discipline that merges population and evolutionary genetics with conservation biology. It aims to delimit the basic units of biological conservation and understand how their genetic composition is affected by habitat loss, exploitation, and environmental change.

We will cover many of the following topics:

- why inbreeding is harmful on reproduction and survival
- how genetic diversity is lost and why it affects evolutionary potential
- how fragmenting populations and habitat impacts gene flow
- how small populations end up evolving more by random processes (genetic drift) than by deterministic processes (natural selection)
- how harmful mutations are accumulated and purged
- how animals in captivity adapt to their surroundings, and why this can be a problem for reintroduction
- how we determine taxonomy and management units using genetic data
- how outcrossing (mating with distant relatives) can impact the fitness of populations
- unique aspects of the genetics and evolution of invasive species species
- aspects of species' biology that is relevant for conservation

Much of the start of the textbook lays the broader justification for conservation biology. Note that much of this field is shaped by values, ethics, and policy—not just science. The field is mission-driven, but can't tell us what we must do, just provide evidence of most effective way to pursue a particular societally desired goal.

Foundational genetics vocabulary and concepts:

The following terms and concepts from genetics and evolutionary biology are foundational to the field, and will occur throughout the subsequent material:

Mendel's Rules of Inheritance:

- Law of dominance and uniformity: Some alleles are dominant while others are recessive; an organism with at least one dominant allele will display the effect of the dominant allele.
- Law of segregation: During gamete formation, the alleles for each gene segregate from each other so that each gamete carries only one allele for each gene.
- Law of independent assortment: Genes of different traits can segregate independently during the formation of gametes.

Darwin's Theory of Natural Selection: Variation of traits, both genotypic and phenotypic, exists within all populations of organisms. However, some traits are more likely to facilitate survival and reproductive success. Thus, these traits are passed onto the next generation. These traits can also become more common within a population if the environment that favors these traits remain fixed.

gene: The basic unit of heredity OR a sequence of nucleotides that is transcribed.

locus: A specific, fixed position on a chromosome.

allele: A variant in the sequence of nucleotides at a particular locus.

genotype: The complete set of alleles an organism carries.

genome: All the genetic information of an organism.

phenotype: The set of observable charectistics or traits of an organism.

ploidy: The number of complete sets of chromosomes in a cell.

diploid: The state of having two homologous copies of each chromosome.

homozygote: The state of having identitical alleles at the same locus on homologous chromosomes.

heterozygote: The state of having different alleles at the same locus on homologous chromosomes.

dominant: The state of an allele that masks or overriding the effect of a different allele at the same locus.

recessive: The state of an allele that has its effect masked by a dominant allele.

chromosome: A package of DNA with some or all the genetic material in an organism. In most chromosomes, thin DNA fibers are coated with packaging proteins to form nucleosomes (histones in Eukaryotes), which bind and condense DNA to protect it, generating a complex 3D shape.

DNA: A polymer composed of two polynucleotide chains that coil around each other to form a double helix. The polymer carries genetic instructions for the development, functioning, growth and reproduction of all known organisms and many viruses. DNA and ribonucleic acid (RNA) are nucleic acids.

Central Dogma: "DNA makes RNA, and RNA makes protein" (DNA is transcribed to RNA by RNA polymerase, RNA is translated to a protein in a ribosome)

nucleotide: Organic molecules composed of a nitrogenous base, a pentose sugar and a phosphate. They serve as monomeric units of the nucleic acid polymers – deoxyribonucleic acid (DNA) and ribonucleic acid (RNA).

A, C, T, G: Adenine, cytosine, thymine, and guanine. Adenine and guanine are purines, with two rings in the base; cytosine and thymine are pyrimidines, with one ring in the base. A & T or C & G form base pairs.

protein: a biomolecule comrpised of one or more amino acids.

amino acid: organic compounds that contain both amino and carboxylic acid functional groups; incorporated into proteins.

enzyme: A protein that acts as a catalyst.

active site: The region of an enzyme where substrate molecules bind to undergo a reaction.

allele frequency: The proportion of chromosomes with a particular allele at a particular locus in a particular population. **genotype frequency**: The proportion of individuals with a particular set of allele at a particular locus in a particular population.

monomorphic: Only one form; invariant

polymorphic: Many forms; variable

diagnostic locus: A locus where all individuals of a particular population or species of interest carry the same allele; can be used to identify them.

Foundational statistics vocabulary and concepts:

Like most of biology, empirical population and conservation genetics is reliant on applied statistics. A grasp of the basic vocabulary and conceptual basis of that discipline is therefore important:

parameter: any quantity of a statistical population that summarizes or describes an aspect of the population, such as a mean or a standard deviation—describes a *population*

statistic: ny quantity computed from values in a sample which is considered for a statistical purpose—describes a *sample*

confidence interval: an interval which is expected to typically contain the parameter being estimated.

null hypothesis: the hypothesis in which no relationship exists between two sets of data or variables being analyzed

p-value: the probability of obtaining test results at least as extreme as the result actually observed, under the assumption that the null hypothesis is correct—A very small p-value means that such an extreme observed outcome would be very unlikely under the null hypothesis. Does not mean the hypothesis of interest is correct.

normal distribution: Normal distribution is a continuous probability distribution wherein values lie in a symmetrical fashion mostly situated around the mean

R^2: the proportion of the variation in the dependent variable that is predictable from the independent variable(s).

degrees of freedom: the number of degrees of freedom is the number of values in the final calculation of a statistic that are free to vary.

Preliminary Assessment

1) White fur color is recessive to brown in oldfield mice, i.e. a white mouse will have the genotype *bb*, while a brown mouse can be either homozygous dominant (*BB*) or heterozygous (*Bb*). In the absence of selective pressures from scarcity or predation, random mating among mice of different colors will continue to produce all three genotypes and thus individuals of both colors will remain in approximately equal percentages. Sometimes two brown mice with heterozygous genotypes will produce a white offspring (*Bb* x *Bb* = *bb* ~1/4th of the time); a white mouse crossing with a homozygous brown mouse will have brown heterozygous offspring. In the absence of other information or an *evolutionary mechanism* that changes the number of *alleles* (e.g., the number of *Bs*) we expect no directional change in pelt color.

2) There are many possible definitions of evolution, but "Evolution is the change in the heritable characteristics of biological populations over successive generations" is a good one. For this class, we will consider evolution a change in the allele frequencies of a population. (If we know a particular phenotype is heritable, we can track that too.) Importantly, it does *not* need to involve selection: it can be driven by mutation, selection, gene flow, or genetic drift.

3) We read phylogenies—evolutionary trees—by tracing the tips ("leaves"; often but not always species) to the root, or the point at which all branches come together. This movement aligns with the temporal axis of the figure, no matter which direction it is pointing in: the tips of a tree are usually meant to represent a group of organisms alive in the present, while the root represents the common ancestor of all the tips, which is necessarily a group of organisms that existed in the past. To determine whether species A and B or B and C, find the two points where their branches come together (called "nodes"). If A and B share a node that is closer to the tips than B and C, they shared a common ancestor more recently and are thus more closely related (or vice versa). In this figure, the node linking frogs and mice is closer to the tips than the node linking fish and frogs. UC Berkeley hosts a useful website for understanding the basics of phylogenies and other foundational evolutionary topics: https://evolution.berkeley.edu/evolution-101/the-history-of-life-looking-at-the-patterns/understanding-phylogenies/

4) A polymer composed of two polynucleotide chains that coil around each other to form a double helix. The polymer carries genetic instructions for the development, functioning, growth and reproduction of all known organisms and many viruses. DNA and ribonucleic acid (RNA) are nucleic acids.

5) Since PKU is recessive, we know that Ms. Doe must have the genotype *pp* to be a carrier. Since it is not sex linked and Mr. Doe does not have it but his mother did, his genotype must be *Pp*—while his father may have been either *PP* or *Pp*, he clearly passed on a *P*, so this missing information is irrelevant to the problem; his mother necessarily passed on a *p* as she was a homozygote. Putting this together, their child will be the product of a *Pp* x *pp* cross, with a 1/2 x 1/2 = 0.25 probability of inheriting a *P* from dad and mom's first *p*, a 1/2 x 1/2 = 0.25 probability of inheriting a *P* from dad and mom's second *p*, a a 1/2 x 1/2 = 0.25 probability of inheriting a *p* from dad and mom's first *p*, and a 1/2 x 1/2 = 0.25 probability of inheriting a *p* from dad and mom's second *p*. The latter two scenarios involve independent events that lead to a PKU phenotype, so we sum them to come up with a 50% chance their child has the disorder.

6) We can solve this by using probability rules (discussed Thursday). The Bobcats have a 80% chance of winning given snowy conditions, which we express as $P(\text{Cats win} \mid \text{snow}) = 0.80$. They have a 40% chance of winning if it isn't snowing (sad!), which we express as $P(\text{Cats win} \mid \text{no snow}) = 0.40$. We also know that the chance of snow on gameday is 25% ($p(\text{snow}) = 0.25$), meaning there is a 75% chance it doesn't snow ($p(\text{no snow}) = 0.75$). To come up with the overall probability, we add the products of 1) the conditional probability of winning with snow and the probability of snow and 2) the conditional probability of winning given no snow and the probability of no snow, e.g. $(0.8 * 0.25) + (0.4 * 0.75) = 0.50 = 50$. This is because the two possible routes to winning (i.e., the two different forecasts) are *independent*—they represent alternate, nonoverlapping (or "mutually exclusive") chances for victory, and thus should be summed.