Molecular Ecology & Evolution: An Introduction

Molecular Ecology & Evolution: An Introduction

designed for BIO 317 at Wheaton College Massachusetts

DREW DAVINACK

NORTON

Contents

Preface ix

Part I. Introduction

A Brief History of Molecular Ecology 3

Key Concepts and Applications in Molecular Ecology 6

Recommended Reading 11

Part II. Review of Molecular Biology

Nucleic Acids 15

Review of DNA Replication and the Central Dogma 20

Mutation and Genetic Variation 27

Genome Organization 36

Part III. Molecular Markers

Introduction 45

Neutral and Adaptive Markers 47

Early Molecular Markers 50

Development of PCR & Sanger Sequencing 56

Recommended Reading 65

Part IV. Species and Populations

Species Concepts 71
Operational Taxonomic Units 74
Speciation 76
Reproductive Isolating Mechanisms 79
Evolutionary Significant Units 83
Cryptic Species 87
Hybridization and Introgression 91
Recommended Reading 96

Part V. Population Genetics

The Metapopulation Concept 103
Revisiting the Hardy-Weinberg Equilibrium 105
Defining Genetic Diversity 109
Genetic Diversity and Bottlenecks 114
Genetic Drift and Diversity 118
A Brief Overview of Analysis of Molecular Variance 124
Gene Flow and Migration 128
Recommended Reading 130

Part VI. Introduction to Phylogeography

A Brief History 135
Case study: Grey Wolf phylogeography 137
The Pleistocene Epoch 140
Coalescent Theory 143
Phylogeographic Barriers 145
Haplotype Networks 151
Marker Selection in Phylogeography 153

Recommended Reading 155

Part VII. Conservation Genetics

The Biodiversity Crisis 159
Applying Population Genetics to Conservation Biology 165
The Extinction Vortex 170
Inbreeding Effects 176
Genetic Rescue 182
Recommended Reading 187

Part VIII. Concepts in Behavior

Sexual Selection 191
Mating Systems 194
Recommended Reading 203

Part IX. Genomics: The New Frontier of Molecular Ecology

Significance of the Genomic Era 207
Key Genomic Technologies 209
Application of Genomics in Molecular Ecology 213
Recommended Reading 216

Bibliography 217
Appendix I: Required Software 222
Appendix II: Syllabus 224

Preface

Welcome to “Molecular Ecology & Evolution,” an open educational resource (OER) textbook designed for advanced undergraduate students in biology and related fields. This textbook is the result of funding by the Wheaton’s Library, Technology and Learning Committee (LTLC) to provide a resource for understanding the complex interplay between molecular biology and ecology. The material presented in this textbook aligns with the curriculum of the Molecular Ecology & Evolution course (BIO 317 and BIO 317L), taught at Wheaton College, and aims to bridge the gap between theoretical concepts and practical applications in this fascinating field of study. Molecular ecology is an interdisciplinary field that combines molecular biology, genetics, and ecology to unravel the mysteries of biodiversity and the evolutionary processes that shape it. This textbook serves as an introduction to the core concepts and methodologies of molecular ecology, offering insights into topics such as molecular markers, population genetics, phylogeography, and conservation genetics. The book is designed to cater to Biology majors who have only completed an introductory level Biology course but are eager eager to explore the molecular mechanisms underlying ecological interactions and evolutionary patterns. The content was deliberately designed in such a way that each chapter is short and tailored to the specific lecture and lab material explored each week. It was not meant to be exhaustive, with the more in-depth details to be presented in lecture and lab. For those adopting this resource, I suggest assigning the readings to students **prior to coming to class**, which can then be used as material for pre-lecture quizzes. Students are encouraged to engage with the material through discussions, hands-on laboratory exercises, and projects. The textbook is complemented by laboratory exercises that provide practical experience with key techniques, including DNA extraction, PCR, and bioinformatics analysis. The course also culminates in two major projects: an NSF-style research proposal and a manuscript for publication, allowing students to apply their knowledge creatively and collaboratively. Pre-requisites: BIO 114 (Introductory Biology) and any 200-level course.

Learning Objectives

At the end of this course, students should be able to:

- **Analyze and Interpret Molecular Data:** Understand and evaluate molecular data at the population level.
- **Dissect Molecular Ecology Research:** Critically analyze research articles in molecular ecology.
- **Formulate Research Questions:** Develop novel research questions and hypotheses in molecular ecology.
- **Make Conservation Decisions:** Apply molecular ecological knowledge to conservation and environmental decisions.

Molecular Ecology & Evolution is designed to be a dynamic and evolving resource, reflective of the ever-changing field it represents. We hope that this textbook serves as a valuable guide for students embarking on their journey into the world of molecular ecology, fostering a deeper understanding and appreciation of the natural world.

PART I
INTRODUCTION

Molecular ecology is a rapidly evolving field of biology that integrates molecular biology and ecology to understand how evolutionary and ecological processes shape biodiversity. By examining genetic variation within and among populations, molecular ecology provides insights into a wide range of biological phenomena, including species adaptation, population structure, and ecological interactions. The ***primary goal of molecular ecology is to understand how genetic factors influence ecological and evolutionary dynamics.*** The field encompasses various subdisciplines, such as population genetics, phylogeography, and ecological genomics. These areas help scientists address questions about species' responses to environmental changes, patterns of gene flow, and the origins of genetic diversity. In this course, we will be drawing on principal concepts that you may have learned from other courses such as genetics, evolution, ecology, and bioinformatics.

Figure 1. Migration of animals is an important feature of ecosystems and also has important consequences at the molecular level.

Media Attributions

- 7371874564_2bd81aeba9_k © USFWS Mountain Prairie is licensed under a CC BY (Attribution) license

A Brief History of Molecular Ecology

The roots of molecular ecology can be traced back to the field of population genetics, which emerged in the early 20th century. Scientists such as R.A. Fisher, J.B.S. Haldane, and Sewall Wright developed the theoretical framework for understanding the distribution of genetic variation within and between populations. However, it wasn't until the mid-20th century that molecular tools became available to test these theories empirically.

The discovery of **DNA** as the genetic material by James Watson, Francis Crick and Rosalind Franklin in 1953 laid the foundation for modern molecular biology. The development of **protein electrophoresis** in the 1960s allowed scientists to directly examine genetic variation in populations. This technique, which separates proteins based on their electric charge, revealed extensive genetic variation in natural populations, challenging the prevailing view of genetic uniformity.

The Advent of Molecular Markers

In the 1970s, molecular ecology began to take shape as scientists developed more sophisticated molecular markers to study genetic variation. The **restriction fragment length polymorphism (RFLP)** technique, introduced in the late 1970s, enabled researchers to analyze variations in DNA sequences. This was followed by the development of the **polymerase chain reaction (PCR)** in the 1980s, which revolutionized molecular biology by allowing specific DNA sequences to be amplified quickly and efficiently.

The 1990s saw an explosion of new molecular markers, such as

random amplified polymorphic DNA (RAPD), **amplified fragment length polymorphism (AFLP)**, and **microsatellites**. These markers provided the tools to address a wide range of ecological and evolutionary questions, from population structure and gene flow to mating systems and phylogeography.

The Genomics Era

The turn of the 21st century marked the beginning of the genomics era in molecular ecology. The completion of the **Human Genome Project** in 2003 highlighted the potential of whole-genome sequencing for understanding genetic variation and evolutionary history. Advances in **next-generation sequencing (NGS)** technologies made genome-wide studies feasible for non-model organisms, further expanding the scope of molecular ecology.

During this period, the focus of molecular ecology shifted from analyzing a few genetic markers to examining entire genomes or large portions of genomes. This shift enabled researchers to study complex traits, identify adaptive genes, and understand the genomic basis of ecological and evolutionary processes.

Integration with Ecology and Evolution

Throughout its history, molecular ecology has evolved from a primarily genetic field to an interdisciplinary field that integrates molecular biology with ecology and evolutionary biology. The field has expanded to include topics such as **conservation genetics**, which uses molecular tools to understand and protect endangered species, and **environmental DNA (eDNA)**, which uses DNA from environmental samples to monitor biodiversity.

Molecular ecology has also benefited from advances in bioinformatics and computational biology, which have provided the tools to analyze large-scale genetic data and uncover patterns of genetic variation and evolutionary relationships. The integration of molecular data with ecological and evolutionary theories has led to a deeper understanding of the mechanisms underlying biodiversity and adaptation.

Current Trends and Future Directions

Today, molecular ecology is a vibrant and rapidly evolving field that continues to push the boundaries of our understanding of biological diversity. Current trends include the use of **metagenomics** to study microbial communities, the application of **CRISPR-Cas9** for functional genomics, and the exploration of **epigenetics** in ecological and evolutionary contexts.

The future of molecular ecology holds exciting possibilities, such as understanding the genetic basis of adaptation to climate change, exploring the role of microbiomes in host ecology, and developing innovative conservation strategies based on genomic data. As the field continues to evolve, it will play an increasingly important role in addressing pressing ecological and evolutionary questions and informing conservation and environmental management.

Key Concepts and Applications in Molecular Ecology

KEY CONCEPTS IN MOLECULAR ECOLOGY

Genetic Diversity and Population Structure: Genetic diversity refers to the total number of genetic characteristics in the genetic makeup of a species, which is manifested in the variety of inherited alleles, genes, and traits within and among individuals of a population, species, or group of species. It is a critical component of biodiversity and underpins the ability of any living population to adapt to changing environments, ensuring survival and evolutionary development. Population structure, on the other hand, describes the distribution of genetic variation within and among populations. Key metrics used to assess population structure include **allele frequencies** (the proportion of different alleles of a particular gene in a population), **heterozygosity** (the proportion of individuals in a population that are heterozygous for a particular gene) and **F-statistics (**a measure of genetic differentiation among populations). While there are several other metrics we will be exploring throughout this course, grasping these three will be crucial for gaining a comprehensive understanding of the field.

Molecular Markers and Techniques: Molecular ecology relies on various molecular markers to study genetic variation, including **microsatellites (**short, repeating sequences of DNA used for assessing genetic diversity and relatedness), **single Nucleotide polymorphisms (SNPs)**, single-base differences in DNA sequences, useful for fine-scale population studies and **mitochondrial DNA (mtDNA)** (inherited maternally and useful for studying population history and phylogeography). In addition, with decreased costs of sequencing and increasing computational power, the use of **whole**

genomes in exploring species adaptations is on the rise and is likely to become standard for many organismal groups in the near future. These can be used as special 'molecular markers' to answer questions in molecular ecology and evolution. **PCR-Based Markers** utilize an important process known as **Polymerase Chain Reaction** which we will be exploring in more depth in lab. PCR markers include Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP) and Microsatellites. **Non-PCR Based Markers include** techniques like Restriction Fragment Length Polymorphism (RFLP) which were among the earliest molecular markers used. **Reduced Genome Analysis include** techniques like Restriction site Associated DNA sequencing (RAD-seq) which provide focused genome insights without require the entire genome.

IMPORTANCE OF MOLECULAR ECOLOGY

Molecular ecology addresses critical topics and the tools we will be exploring in this course (both in lecture and lab) are important for:

1. **Estimating Extinction Risk:** By analyzing genetic diversity, molecular ecology helps estimate the extinction risk of endangered animals. It plays a crucial role by identifying distinct populations, assessing interbreeding and guiding breeding programs. For example, it can help identify "evolutionary significant units" (ESUs) for conservation efforts.
2. **Biodiversity Research:** The field aids in identifying species (e.g., through DNA barcoding) and understanding hybridization events. Environmental DNA (eDNA) is an exciting new method of identifying species without actually requiring the species to be present! Environmental DNA is collected from environmental samples like soil or water. It allows for non-invasive monitoring of biodiversity and detecting rare or

invasive species, contributing to ecological studies and conservation efforts.

3. **Behavioral Ecology:** Molecular techniques can reveal insights into animal behavior, such as mating systems, kinship, and social structure. For example, genetic analyses can determine parentage or identify relatedness among individuals in a social group.

Case Study: Burmese Python Invasion in Florida

The Burmese python (*Python bivittatus*) is a large constrictor native to Southeast Asia that has become an invasive species in Florida, particularly in the Everglades. These pythons, which can grow over 20 feet long, pose a significant threat to native wildlife. The application of molecular ecology in managing this invasive species provides crucial insights into their population dynamics, origins, and impacts on the ecosystem. One of the key applications of molecular ecology is understanding the genetic diversity and population structure of invasive species. By analyzing the genetic markers of Burmese pythons, scientists have been able to trace their origins and understand their population dynamics in Florida. Studies have revealed that the invasive python population in Florida exhibits a genetic bottleneck, indicating that a small number of individuals initially founded the population. This information helps in understanding how the population grew and spread. Additionally, genetic analyses have shown that the pythons in Florida likely originated from multiple sources, including the pet trade and possibly intentional releases. By studying the genetic diversity of these pythons, molecular ecologists assess their potential for adaptation to new environments. The relatively low genetic diversity observed in Florida's python population suggests limited adaptability, yet their successful invasion indicates that genetic diversity alone does not determine invasive potential. Molecular ecology also aids in understanding the ecological impacts of invasive species through diet analysis. By examining stomach contents and feces, and using DNA barcoding, researchers can identify the prey species consumed by Burmese pythons. Using DNA barcoding, scientists have identified numerous native species in the diet of Burmese pythons, including endangered species like the Key Largo wood rat. This information highlights the impact of the invasive species on native biodiversity and helps prioritize conservation efforts. Future applications of molecular ecology may include genetic control methods, such as gene drives, to reduce the invasive python population. While these approaches are still in development, they offer promising avenues for managing invasive species through molecular interventions.

The US Department of Agriculture's Animal Plant Health and Inspection Service researchers provide training to other biologists on how to trap and handle invasive Burmese pythons.

Media Attributions

- python © USDAgov is licensed under a CC BY (Attribution) license

Recommended Reading

The following are recommended readings – we will be discussing at least one of these articles in next week's group discussion.

Nielsen ES, Hanson JO, Carvalho SB, Beger M, Henriques R, Kershaw F, von der Heyden S (2022) Molecular ecology meets systematic conservation planning. *Trends in Ecology & Evolution* 38: P143 – P155. https://doi.org/10.1016/j.tree.2022.09.006

Andrew RL, Bernatchez L, Bonin A, Buerkle CA, Carstens BC, Emerson BC, *et al* (2013). A roadmap for molecular ecology. *Molecular ecology* 22: 2605 – 2626. https://doi.org/10.1111/mec.12319

Hohenlohe PA, Funk WC, Rajora OP (2021) Population genomics for wildlife conservation and management. *Molecular ecology* 30: 62 – 82. https://doi.org/10.1111/mec.15720

Johnson JB, Peat SM, Adams BJ (2009) Where's the ecology in molecular ecology? *Oikos* 118: 1601 – 1609. doi: 10.1111/j.1600-0706.2009.17557.x

PART II
REVIEW OF MOLECULAR BIOLOGY

LEARNING OBJECTIVES

At the end of this chapter, you should be able to:

1. Be able to explain the primary and secondary structure of deoxyribonucleic
2. acid (DNA).
3. Be able to explain the importance of mutation in generating genetic
4. differences among living organisms.
5. Be able to explain how the information encoded in DNA is accessed and processed.
6. Explain the role of the Central Dogma in understanding genetic diversity in organisms.
7. List and explain the various ways the genome is organized in most living organisms.

Understanding molecular biology is essential before delving into molecular ecology because it provides the foundational knowledge required to appreciate how molecular mechanisms influence ecological and evolutionary processes. Molecular biology explores the structure and function of the molecules that make up living cells, particularly DNA, RNA, and proteins, which are the building blocks of life. This field explains how these molecules interact to carry out the processes essential for the survival and reproduction of organisms. In molecular ecology, researchers apply this fundamental molecular knowledge to address questions about the ecological and evolutionary dynamics of populations. For instance, by understanding DNA replication and gene expression, molecular ecologists can investigate genetic variations within and between populations, helping to uncover the molecular underpinnings of adaptation to different environmental conditions. This approach is critical for studying how species evolve in response to ecological pressures and for identifying the genetic basis of traits that contribute to survival and reproductive success in various environments. Moreover, molecular biology techniques, such as DNA sequencing and PCR (polymerase chain reaction), are indispensable tools in molecular ecology. These techniques allow scientists to examine the genetic makeup of organisms from different ecosystems, providing insights into population structure, genetic diversity, and evolutionary history. For example, the ability to amplify specific regions of DNA through PCR enables the detailed study of genetic markers across individuals and populations, facilitating investigations into gene flow, genetic drift, and selection pressures. Additionally, a solid grasp of molecular biology is necessary to interpret results from molecular ecology studies accurately. Understanding the sources of genetic variation, such as mutations, gene duplication, and horizontal gene transfer, allows ecologists to make informed conclusions about the genetic structure of populations and their adaptive responses to environmental changes.

Nucleic Acids

Nucleic acids are biological molecules that play essential roles in the storage, transmission, and expression of genetic information. The two primary types of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). These molecules are composed of long chains of nucleotides, which serve as the fundamental building blocks.

Each nucleotide comprises three components:

1. **A Phosphate Group**
2. **A Five-Carbon Sugar** – Deoxyribose in DNA and ribose in RNA
3. **A Nitrogenous Base** – Adenine (A), Thymine (T), Cytosine (C), Guanine (G) in DNA, and Uracil (U) replacing Thymine in RNA

The Structure of DNA

DNA has a double-helix structure, consisting of two intertwined strands that are held together by hydrogen bonds between complementary nitrogenous bases. The bases pair specifically:

- Adenine **(A)** pairs with Thymine **(T)**
- Cytosine **(C)** pairs with Guanine **(G)**

The structure of DNA allows for accurate replication, as each strand serves as a template for the formation of a complementary strand.

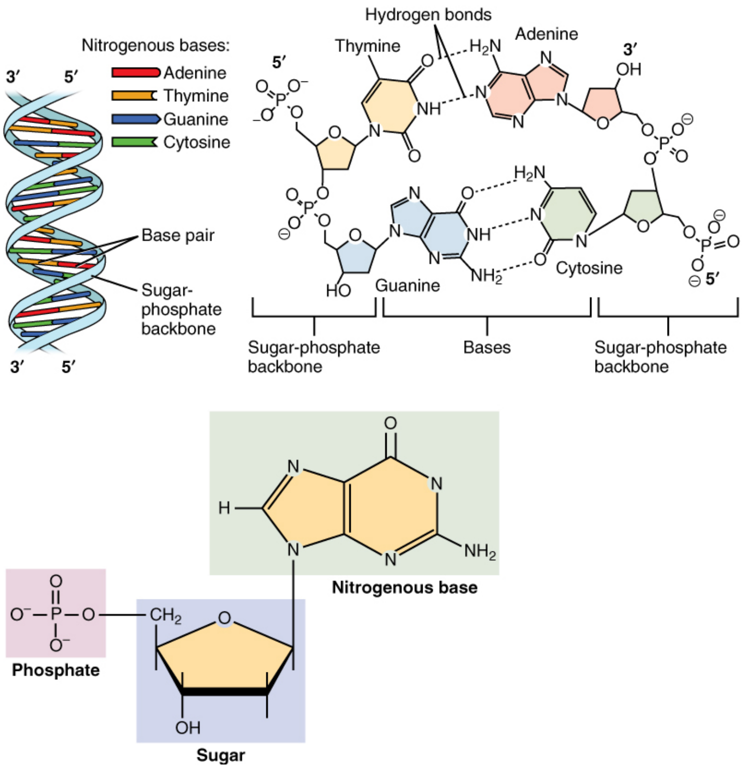


Figure 3. Secondary structure of DNA along with chemical structure of each of the four deoxyribonucleotides

The Structure of RNA

RNA is typically single-stranded but can form complex three-dimensional structures through base pairing within the same strand. The nitrogenous bases in RNA are similar to those in DNA, except that RNA contains Uracil (U) instead of Thymine (T). RNA serves multiple functions, including:

- **Messenger RNA (mRNA):** Carries genetic information from DNA to the ribosome for protein synthesis.

- **Transfer RNA (tRNA):** Helps decode mRNA into proteins.
- **Ribosomal RNA (rRNA):** Forms the core structural and functional components of the ribosome.

The universal presence of nucleic acids in all known life forms suggests a common origin of life. This is further supported by the following: (1) **Genetic Code:** the genetic code, which specifies how nucleotides are translated into amino acids, is nearly universal across all organisms. This points to a shared evolutionary heritage, (2) **Homology:** The structural and functional similarity between the nucleic acids of different organisms indicates a common ancestry. Homologous genes, which are genes inherited from a common ancestor, often perform similar functions across species and (3) **Phylogenetic Evidence:** Molecular phylogenetics, which uses genetic data to reconstruct evolutionary relationships, consistently reveals a tree of life that supports a common origin.

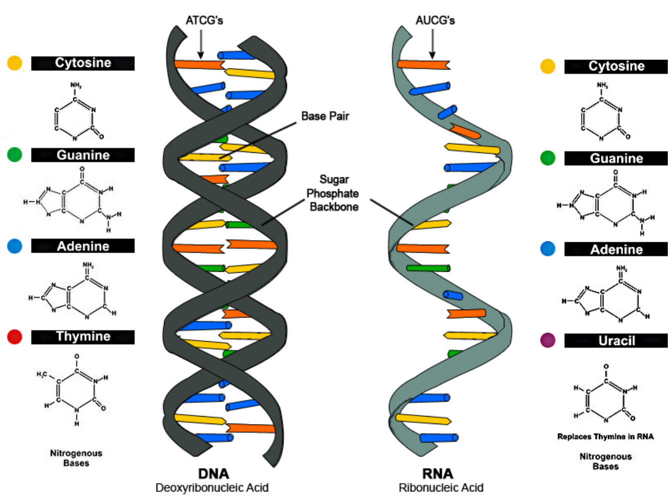


Figure 4. Difference in RNA vs DNA Secondary Structure

The primary structure of DNA refers to the specific sequence of nucleotides. These nucleotides are linked together through covalent bonds, forming a long chain known as a DNA strand. The order of these nucleotides is crucial because it determines the genetic information carried by the DNA. **DNA Sequencing** is the process of determining the precise order of nucleotides in a DNA molecule. It provides the primary structure of DNA, which is fundamental for various applications in molecular ecology. From a molecular ecology perspective, this is crucial for understanding the genetic basis of ecological and evolutionary processes. DNA sequencing technologies enable researchers to decode this information, providing insights into species identification, genetic diversity, evolutionary relationships, and adaptation, all of which are key areas of interest in molecular ecology.

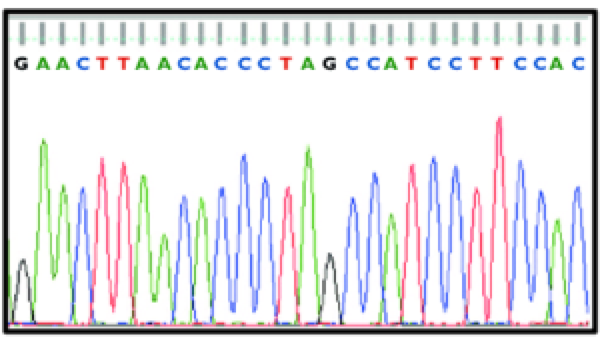


Figure 5. Partial DNA sequence obtained from a chromatogram generated by Sanger sequencing using an ABI Sequencer.

Media Attributions

- 0322_DNA_Nucleotides © OpenStax is licensed under a CC BY (Attribution) license
- dnavsrna © Zappys Technology Solutions is licensed under a CC BY (Attribution) license
- sequencing

Review of DNA Replication and the Central Dogma

DNA replication is a fundamental biological process that ensures the faithful transmission of genetic information from one generation to the next. The process is semi-conservative, meaning each new DNA molecule consists of one original (parental) strand and one newly synthesized strand. This process occurs during the S phase of the cell cycle. The replication of DNA involves several key steps:

1. **Initiation**:
   - **Origin of Replication**: Replication begins at specific sites called origins of replication.
   - **Helicase**: An enzyme called helicase unwinds the DNA helix, creating a replication fork.
   - **Primase**: The enzyme primase synthesizes a short RNA primer complementary to the DNA template to provide a starting point for DNA synthesis.

2. **Elongation**:
   - **DNA Polymerase**: DNA polymerase adds new nucleotides to the 3′ end of the RNA primer, extending the new DNA strand.
   - **Leading and Lagging Strands**:
     - The **leading strand** is synthesized continuously in the direction of the replication fork.
     - The **lagging strand** is synthesized discontinuously, creating short fragments called Okazaki fragments.
   - **DNA Ligase**: DNA ligase joins the Okazaki fragments to

form a continuous strand.

3. **Termination**:
    - Replication terminates when replication forks meet or when specific termination sequences are encountered.
    - The RNA primers are removed and replaced with DNA, and any gaps are sealed.

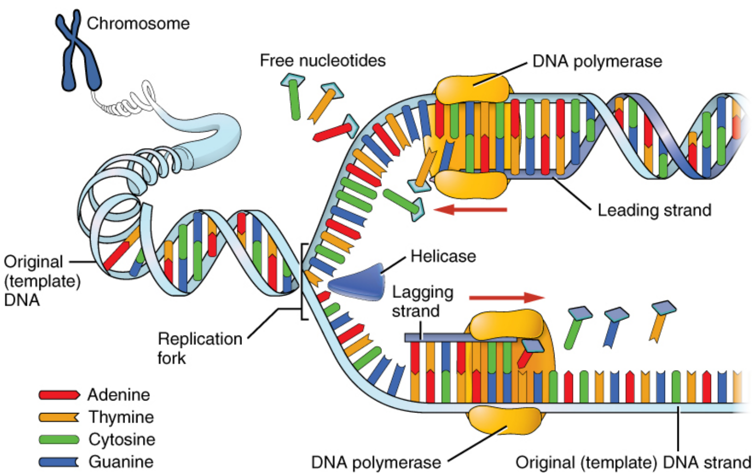


Figure 6. Summary of DNA replication.

DNA replication is remarkably accurate, but errors can occur. The fidelity of replication is maintained through several mechanisms. DNA polymerases have **proofreading abilities**, allowing them to remove incorrectly paired nucleotides immediately after they are added. DNA polymerases have a natural ability to detect when an incorrect nucleotide has been incorporated. This detection is largely based on the instability of a mismatch within the DNA double helix. The correct pairing of DNA bases (adenine with thymine, and cytosine with guanine) forms stable hydrogen bonds, while incorrect pairings do not fit properly and disrupt the helical structure. Many DNA polymerases also have an intrinsic 3' to 5' exonuclease activity. This means that when a mismatch is detected, the polymerase can reverse its direction, remove the incorrect nucleotide (exonucleolytic cleavage), and then resume DNA synthesis. This proofreading activity is crucial for correcting errors. Once the incorrect nucleotide has been excised, the DNA polymerase repositions itself back at the 3' end of the newly synthesized strand and continues to add the correct nucleotides, ensuring the fidelity of the DNA replication process. Despite these mechanisms, replication errors do occasionally occur, leading to **mutations**.

THINK/PAIR/SHARE

1. Considering the role of DNA polymerase's exonuclease activity in maintaining replication fidelity, what might be the consequences for a cell if this proofreading mechanism was genetically disabled or inhibited by a chemical agent? How could this impact genetic stability and the overall health of an organism?
2. DNA replication is tightly regulated and occurs during the S phase of the cell cycle. Why is the timing of replication critical for cellular function and genetic integrity? What potential problems could arise if DNA replication occurred at multiple points throughout the cell cycle?

The Central Dogma of Molecular Biology describes the flow of genetic information within a biological system. It is a framework that describes the flow of genetic information within a biological system. Articulated by Francis Crick in 1958, it outlines the process by which genetic information is transferred from DNA to RNA and then to protein. This fundamental concept consists of three main processes: replication, transcription, and translation, each interconnected and essential for the expression of genes.

1. **Replication:** *The process by which DNA makes a copy of itself* (covered earlier).
2. **Transcription:** *The process of converting DNA into RNA.* Transcription is the first step in the gene expression process, where the sequence of a gene is copied from DNA to **messenger RNA (mRNA)**. RNA polymerase, an enzyme, binds to a specific sequence on the DNA called the promoter, located at the beginning of a gene. This binding unwinds the DNA segment, and RNA polymerase reads the DNA template strand to synthesize a single-stranded RNA molecule with a nucleotide sequence complementary to the DNA template. Unlike in DNA, the RNA nucleotide uracil (U) is used instead of thymine (T), pairing with adenine. Once the mRNA molecule is synthesized, it undergoes processing where introns (non-coding regions) are removed, and a cap and tail are added to stabilize the mRNA before it exits the nucleus.
3. **Translation:** *The process of synthesizing proteins based on the sequence of an mRNA molecule.* Translation is the process by which the genetic code carried by mRNA is decoded to produce a specific sequence of amino acids, ultimately resulting in a polypeptide chain that folds into a functional protein. This occurs in the ribosome, a complex machine in the cell that facilitates the docking of mRNA and the **transfer RNA (tRNA)** molecules that carry amino acids. Each three-nucleotide sequence (codon) on the mRNA corresponds to one amino acid, as specified by the genetic code. For example, the

codon AUG codes for the amino acid methionine and also serves as the start signal for translation. As the ribosome moves along the mRNA, tRNA molecules match their complementary anticodon sequences to the mRNA codons, bringing the appropriate amino acids into place. The ribosome catalyzes the formation of peptide bonds between the amino acids, elongating the polypeptide chain until it reaches a stop codon, which does not code for an amino acid but signals the end of translation.

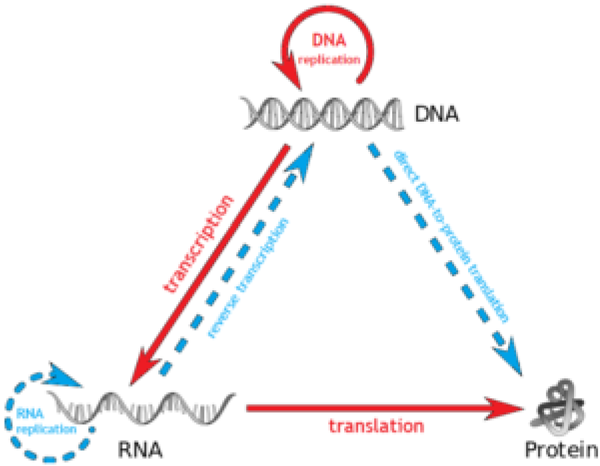


Figure 7. The Central Dogma of Molecular Biology.

Finally, the **standard genetic code** specifies the basic rules by which the nucleotide sequence of mRNA is translated into the amino acid sequence of proteins. It is nearly universal across all organisms, underlining the evolutionary continuity and common origin of life (there are exceptions). The genetic code is redundant, meaning that most amino acids are encoded by more than one codon, providing some tolerance against mutations. This code ensures that genetic information is translated with remarkable precision into the diverse array of proteins that perform various functions within the cell.

Second Base									
First Base	U		C		A		G		**Third Base**
U	UUU	phe	UCU	ser	UAU	tyr	UGU	cys	U
	UUC		UCC		UAC		UGC		C
	UUA	leu	UCA		UAA	STOP	UGA	STOP	A
	UUG		UCG		UAG		UGG	trp	G
C	CUU		CCU	pro	CAU	his	CGU	arg	U
	CUC	leu	CCC		CAC		CGC		C
	CUA		CCA		CAA	gln	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	ile	ACU	thr	AAU	asn	AGU	ser	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	lys	AGA	arg	A
	AUG	met START	ACG		AAG		AGG		G
G	GUU	val	GCU	ala	GAU	asp	GGU	gly	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	glu	GGA		A
	GUG		GCG		GAG		GGG		G

Figure 8. The standard genetic code showing which proteins specific mRNA codes translate to.

Media Attributions

- 0323_DNA_Replication © OpenStax
- Central_dogma_of_molecular_biology_colorized+special_transfer(1) © Philippe Hupé is licensed under a CC BY-SA (Attribution ShareAlike) license

Mutation and Genetic Variation

Mutations are *changes in the DNA sequence of an organism*. They can have various causes, including errors during DNA replication or damage from environmental factors. Mutations are categorized based on their effect on the genetic material, and understanding them is crucial in molecular ecology for interpreting genetic variation and its consequences. Below are the types of mutations classified based on their nature and effects.

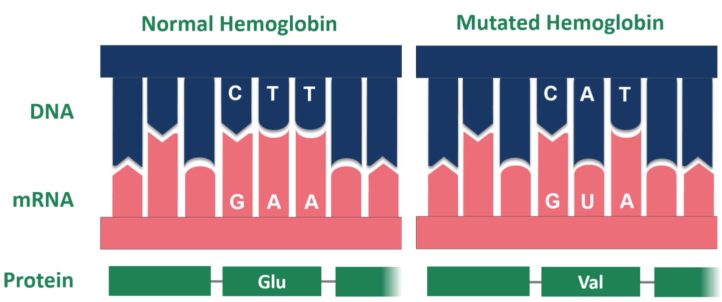


Figure 9. Diagram showing the impact of a single nucleotide mutation on the gene encoding the beta-globin subunit of hemoglobin, which leads to sickle cell disease. The top part of the diagram shows the sequences of a segment of DNA and its corresponding mRNA for both normal and mutated forms. In the normal hemoglobin, the DNA sequence "CTT" leads to the mRNA codon "GAA," which codes for the amino acid glutamic acid (Glu). In the mutated hemoglobin, a point mutation changes the DNA sequence from "CTT" to "CAT," resulting in the mRNA codon "GUA" and the substitution of valine (Val) for glutamic acid in the hemoglobin protein. This amino acid change causes the hemoglobin molecules to form rigid structures that distort red blood cells into a sickle shape, leading to various health complications associated with sickle cell disease.

Point Mutations

Point mutations are changes involving a single nucleotide base pair. They are categorized into:

1. **Substitutions**: A single base is replaced by a different base.
    - **Silent Mutation**:
        - The altered codon still codes for the same amino acid due to the redundancy of the genetic code.
        - These mutations do not affect the protein and are often considered neutral.
    - **Missense Mutation**:
        - The altered codon codes for a different amino acid.
        - This can result in a protein with altered function or no function, depending on the specific amino acid change.
    - **Nonsense Mutation**:
        - The altered codon becomes a stop codon, resulting in premature termination of translation.
        - This typically produces a truncated protein that is often non-functional.

2. **Insertions and Deletions –** Insertions and deletions involve the addition or removal of one or more nucleotide base pairs, which can lead to multiple consequences:
    1. *Frameshift mutations:*
        1. Insertion or deletion of a number of nucleotides that is not a multiple of three, shifting the reading frame of the gene.
        2. Frameshift mutations usually result in non-functional proteins because the entire downstream sequence is altered.

2. *In-frame mutations*:
    1. Insertion or deletion of nucleotides in multiples of three, preserving the reading frame.
    2. These mutations may affect the function of the protein depending on the number and position of the inserted or deleted amino acids.
3. **Structural Mutations – i**nvolve changes to larger sections of DNA, such as:
    1. **Duplications – a** segment of DNA is copied and inserted into the genome. Duplications can provide raw material for evolution, as the extra copy of a gene can accumulate mutations and potentially gain a new function.
    2. **Deletions** – segment of DNA is removed from the genome. Deletions can range from a few base pairs to large chromosomal regions, often resulting in loss of function.
    3. **Inversions** – a segment of DNA is reversed within the genome. Inversions typically do not affect gene function unless they disrupt regulatory regions or involve large segments.
    4. **Translocations** – a segment of DNA is moved to a different position within the genome. Translocations can result in fusion genes or disrupt regulatory elements, leading to altered gene expression.

Mutation by Impact on Function

Mutations can also be classified based on their impact on protein function. **Loss-of-Function mutations** result in a reduced or absent protein function and are often recessive because the presence of one functional copy of the gene can compensate for the loss. **Gain-**

of-Function mutations result in a new or enhanced protein function. These are often dominant because the altered function can manifest even in the presence of a normal gene copy. **Dominant-Negative Mutations** are those where the mutated protein interferes with the function of the normal protein, often through abnormal interactions. These are typically dominant because they affect the function of both alleles.

Mutation by Origin

Lastly, mutations can be classified based on their origins. **Germline mutations** occur in the reproductive cells and are passed on to offspring. These mutations are important in evolution and hereditary disease. **Somatic mutations** occur in non-reproductive cells and are not passed on to offspring. These can lead to cancer and other somatic cell abnormalities.

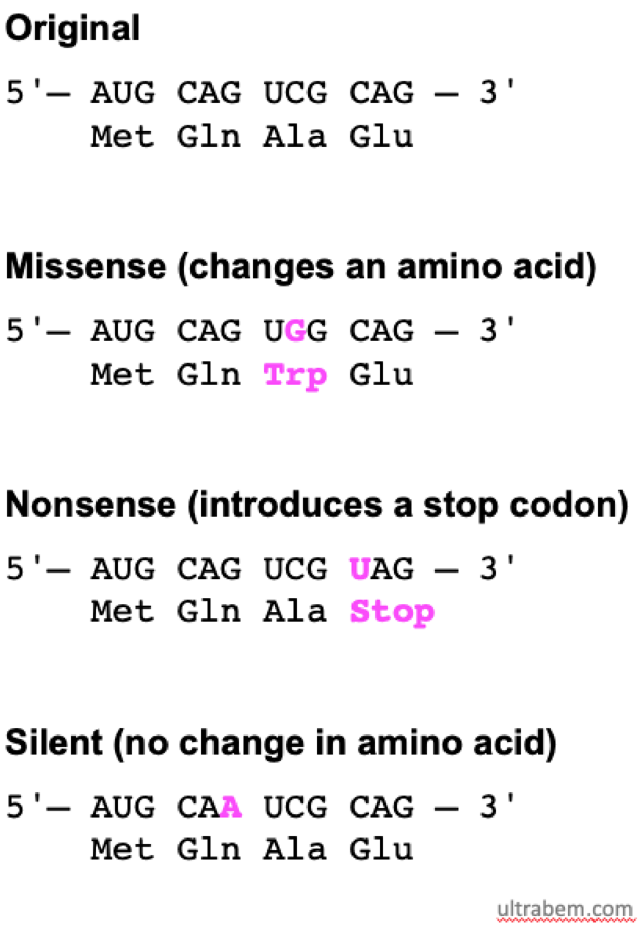


Figure 10. Different types of DNA mutations.

THINK/PAIR/SHARE

1. Consider the different outcomes of point mutations (silent, missense, and nonsense). How might each type of mutation influence the evolutionary fitness of an organism in a rapidly changing environment? Discuss scenarios where a missense or nonsense mutation could provide an adaptive advantage or disadvantage
2. Given the descriptions of various mutation types, such as insertions, deletions, and structural mutations, discuss how these genetic alterations could lead to different genetic disorders or diseases. Can you think of specific examples where a frameshift mutation or a structural mutation has been identified as the cause of a particular disease? How do these mutations disrupt normal cellular function?

Mutational Rates

Mutational rates refer to the frequency at which mutations occur in a genome over a certain period of time or across a generation. Understanding mutation rates is crucial for elucidating the dynamics of genetic variation, which in turn impacts evolutionary processes, adaptation, and biodiversity. Mutation rates vary widely across species and even within genomes. Mutation rates tend to be lower in organisms with larger genomes, possibly due to the increased cost of correcting errors. In viruses, mutation rates are typically higher, which can facilitate rapid adaptation but also lead to higher deleterious mutation loads. Mutation rates also vary across different regions of the genome, often being higher in non-coding regions or areas with repetitive sequences The Y chromosome and mitochondrial DNA typically have higher mutation rates compared to autosomes, due to less effective repair mechanisms.

Table 1: Factors Affecting Mutational Rates

Factor	Example	Explanation
Genetic Factors	1. DNA Polymerase Fidelity 2. Mismatch Repair	1. The accuracy of DNA polymerase during replication affects mutation rates. 2. Post-replication repair mechanisms correct mismatched based pairs, reducing mutation rates.
Environmental Factors	1. Radiation 2. Chemicals 3. Temperature	1. Exposure to UV or ionizing radiation increases mutation rates by damaging DNA. 2. Mutagenic chemicals can induce mutations. 3. High temperatures can increase mutation rates by destabilizing DNA.
Biological Factors	1. Generation Time 2. Metabolic Rate	1. Species with shorter generation times tend have higher mutation rates per year, as they undergo more rounds of replication. 2. Higher metabolic rates can increase mutation rates due to increased production of reactive oxygen species (ROS).

Media Attributions

- Point-Mutation-Sickle-Cell-Normal_and_Mutated-Hemoglobin © Thomas Samuel - ACC Bioinnovation Lab is licensed under a CC BY-SA (Attribution ShareAlike) license
- DNA_mutation(1) is licensed under a CC0 (Creative Commons Zero) license

Genome Organization

The genomes of living organisms, whether prokaryotic or eukaryotic, are the fundamental blueprints that dictate their biological functions and evolutionary potential. The organization of these genomes reflects the different structural and functional requirements of the two groups, influencing their biological complexity, adaptability, and ecological interactions.

Prokaryotic Genomes

Prokaryotic genomes are generally simpler in structure compared to their eukaryotic counterparts. Prokaryotes, which include bacteria and archaea, typically have a single, circular chromosome. This chromosome is located in the nucleoid, a region within the cell that is not membrane-bound. The circular nature of the chromosome is advantageous, as it allows for continuous replication. Additionally, many prokaryotes possess plasmids, which are small, circular DNA molecules that replicate independently of the chromosomal DNA. These plasmids often carry genes that confer advantageous traits, such as antibiotic resistance, facilitating rapid adaptation to changing environments. The genome size of prokaryotes tends to be much smaller than that of eukaryotes, reflecting their generally less complex biological processes. Prokaryotic genomes are densely packed with genes, and intergenic regions are minimal, resulting in a high gene density. This efficient organization reflects the streamlined nature of prokaryotic genomes, which often lack introns, making their genes continuous coding sequences. This arrangement allows for efficient transcription and translation, facilitating rapid cellular responses to environmental stimuli.

Eukaryotic Genomes

Eukaryotic genomes are more complex and varied in their structure. Eukaryotes, which include animals, plants, fungi, and protists, typically have multiple linear chromosomes contained within a membrane-bound nucleus. This structural compartmentalization allows for greater regulation of gene expression and more complex interactions between genetic elements. Eukaryotic chromosomes are associated with histone proteins, which help package the DNA into a compact, organized structure known as chromatin. The chromatin structure plays a crucial role in regulating gene expression, as regions of tightly packed chromatin (heterochromatin) are generally transcriptionally inactive, while loosely packed chromatin (euchromatin) is transcriptionally active.

Eukaryotic genomes are significantly larger than prokaryotic genomes, often containing a vast amount of non-coding DNA. This **non-coding DNA includes introns**, which are non-coding regions within genes that are spliced out during RNA processing, as well as various types of repetitive elements. The presence of introns allows for alternative splicing, a process that *enables a single gene to produce multiple protein variants*, increasing the functional diversity of the proteome. The repetitive elements, which include transposable elements and tandem repeats, contribute to genomic diversity and evolution through mechanisms such as gene duplication and genome rearrangement.

One of the key distinctions between prokaryotic and eukaryotic genomes lies in the regulation of gene expression. In prokaryotes, gene expression is typically regulated at the transcriptional level through mechanisms such as operons, where a single promoter controls the expression of multiple genes. This allows for coordinated expression of genes involved in related functions. In eukaryotes, gene regulation is more complex, involving multiple levels of control, including chromatin remodeling, transcriptional regulation, RNA processing, and post-transcriptional regulation. This complexity allows for precise spatial and temporal control of gene expression, enabling the development and differentiation of multicellular organisms. Another important aspect of genome organization is the presence of organellar genomes in eukaryotes. Eukaryotic cells contain mitochondria and, in the case of plants and algae, chloroplasts, both of which have their own genomes. These organellar genomes are typically circular and resemble prokaryotic genomes, reflecting their evolutionary origin from endosymbiotic bacteria. The organellar genomes encode essential components for energy production, and their presence adds an additional layer of genetic complexity to eukaryotic cells.

An Introduction to Mitochondrial DNA (mtDNA)

Mitochondrial DNA (mtDNA) is a type of extranuclear DNA located in the mitochondria, the energy-producing organelles within eukaryotic cells. Unlike nuclear DNA, which is inherited from both parents, mtDNA is typically maternally inherited, making it a valuable tool for tracing lineage and evolutionary patterns. However, using mtDNA to infer evolutionary relationships has both advantages and disadvantages. Most of the studies we will be exploring in this course will focus heavily on mtDNA and for lab, you will be amplifying a mitochondrial DNA gene-fragment. It is important that you understand the shortcomings of using this type of DNA for inferring evolutionary patterns.

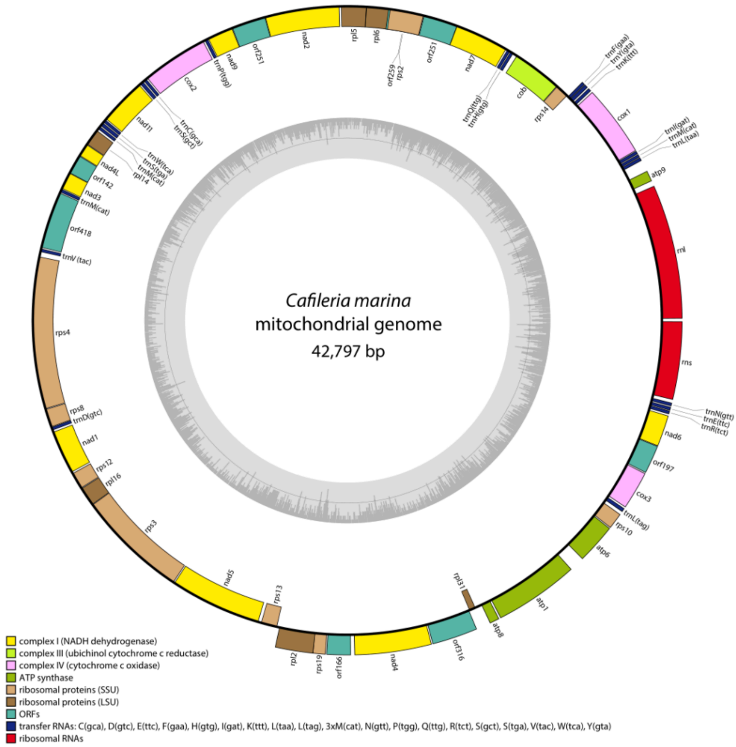


Figure 11. Mitochondrial genome of the protist, Cafileria marina

Advantages of using mtDNA in molecular ecology research

1. **Maternal Inheritance**:
    - mtDNA is passed down from the mother without recombination, allowing for clear lineage tracing.
    - This feature is particularly useful for studying maternal lineages and resolving evolutionary relationships within species and closely related species.

2. **High Mutation Rate**:
    - mtDNA generally has a higher mutation rate compared to nuclear DNA.
    - This characteristic provides greater genetic diversity over shorter evolutionary time scales, making it ideal for studying recent evolutionary events.
3. **Conserved Structure**:
    - The structure of mtDNA is relatively conserved across species, allowing for easy comparison of homologous regions.
    - The consistent gene arrangement simplifies comparative studies and phylogenetic analyses.
4. **Abundance**:
    - Mitochondria contain multiple copies of mtDNA, which makes it easier to extract and amplify compared to nuclear DNA.
    - This abundance is especially advantageous when dealing with degraded or ancient DNA samples, as often found in paleogenomics and forensic studies.
5. **Lack of Recombination**:
    - The absence of recombination in mtDNA provides a straightforward inheritance pattern, which simplifies the interpretation of evolutionary relationships.

Disadvantages of using mtDNA for molecular ecology research

1. **Maternal Inheritance**:

- While maternal inheritance is useful for studying maternal lineages, it provides a biased view of the evolutionary history of an organism.
- It ignores the paternal lineage and provides no information about recombination-based evolution or nuclear DNA contributions.

2. **High Mutation Rate**:

   - The high mutation rate can lead to homoplasy, where unrelated lineages independently acquire similar mutations.
   - Homoplasy can obscure true evolutionary relationships and complicate phylogenetic analyses.

3. **Limited Genomic Information**:

   - The mitochondrial genome is small and contains limited genetic information compared to the nuclear genome.
   - This limitation restricts the amount of evolutionary data available for analysis and may not reflect the broader evolutionary patterns present in the nuclear genome.

4. **Selective Sweeps and Bottlenecks**:

   - mtDNA is subject to selective sweeps and genetic bottlenecks, which can reduce genetic diversity and obscure evolutionary history.
   - These events can result in misleading interpretations of population history and evolutionary relationships.

5. **Nuclear-Mitochondrial DNA (Numts)**:

   - Nuclear copies of mitochondrial DNA, known as NUMTs, can complicate analyses by producing misleading sequences that resemble true mtDNA.

    - The presence of Numts requires careful interpretation of sequencing data to avoid incorrect phylogenetic inferences.

Media Attributions

PART III

MOLECULAR MARKERS

LEARNING OBJECTIVES

At the end of this chapter, you should be able to:

1. Define what a molecular marker is and how they are used.
2. Explain how molecular markers are generated in the laboratory.
3. Explain the difference between neutral versus adaptive markers.
4. Distinguish between PCR and non-PCR based markers.

Molecular markers are indispensable tools in the field of molecular ecology and evolution, providing insights into the genetic makeup of organisms that are otherwise invisible to the naked eye. These markers are specific sequences of DNA that can be used to identify individuals, ascertain genetic diversity, and understand the genetic structure of populations. They play a pivotal role in a myriad of applications including, but not limited to, phylogenetics, population genetics, conservation biology, and the study of evolutionary processes. This chapter will explore the various types of molecular markers such as microsatellites, single nucleotide polymorphisms (SNPs), and sequence-tagged sites, among others. Each marker type offers unique advantages and limitations depending on the research objectives, from fine-scale genetic mapping to broad-scale population studies.

Introduction

What is a molecular marker?

A **molecular marker** is a *specific* DNA *sequence or variation in* DNA *sequence* that can be used to identify individuals, populations, or species. These markers are often variations in the DNA sequence that do not necessarily affect the phenotype (the physical characteristics) of an organism but are inherited and can be passed down through generations.

There are several types of molecular markers commonly used in molecular ecology and evolution studies:

1. **Microsatellites (Simple Sequence Repeats, SSRs)**: Microsatellites are short tandem repeats of DNA sequences, consisting of 1-6 base pairs repeated multiple times. They are highly polymorphic and widely distributed throughout the genome, making them useful for studying genetic diversity, population structure, and relatedness among individuals.
2. **Single Nucleotide Polymorphisms (SNPs)**: SNPs are single base pair differences in DNA sequence that occur commonly throughout the genome. They are the most abundant type of genetic variation and are often used to study population genetics, genome-wide association studies (GWAS), and phylogenetics.
3. **Insertion/Deletion Polymorphisms (Indels)**: Indels are variations in DNA sequence where a segment of DNA is either inserted or deleted. They can be used as molecular markers to study population structure and evolutionary relationships.
4. **Restriction Fragment Length Polymorphisms (RFLPs)**: RFLPs are variations in DNA sequence that result in differences in the lengths of DNA fragments when cut with restriction enzymes.

They were one of the first types of molecular markers used in genetics and have been widely used in population genetics and phylogenetics.

5. **Random Amplified Polymorphic DNA (RAPD)**: RAPD markers are short DNA sequences amplified using PCR (Polymerase Chain Reaction) with random primers. They are useful for studying genetic diversity and population structure due to their ability to generate many markers across the genome.
6. **Sequence-Characterized Amplified Region (SCAR)**: SCAR markers are DNA sequences amplified using PCR with primers designed from known DNA sequences. They are useful for identifying specific genetic traits or genes of interest within populations.

Molecular markers are used in a variety of molecular ecology and evolutionary studies, including population genetics, phylogeography, conservation biology, and evolutionary biology. They provide valuable information about genetic diversity, population structure, gene flow, adaptation, and evolutionary relationships among organisms.

Neutral and Adaptive Markers

Neutral markers are genetic variations that do not confer any selective advantage or disadvantage to an organism. Instead, they evolve primarily under the influence of genetic drift and mutation. These markers are often used to study demographic processes, population history, and gene flow. For example, microsatellites, also known as simple sequence repeats (SSRs), are highly polymorphic neutral markers commonly used in molecular ecology studies. They consist of short tandem repeats of DNA sequences, such as CA or AT, repeated multiple times. Microsatellites evolve rapidly due to high mutation rates in the repeat regions, and mutations in these regions are generally neutral in terms of fitness.

Researchers studying the population genetics of a species of bird might use microsatellites to assess genetic diversity and population structure among different breeding populations. By genotyping individuals from multiple populations at several microsatellite loci, they can estimate gene flow, genetic differentiation, and demographic history without being biased by selection.

DEFINITION BOX 1

In genetics, a "locus" (plural: loci) refers to a specific, fixed position on a chromosome where a particular gene or genetic marker is located. Each locus can have different forms, known as alleles, which can vary among individuals in a population. Loci play a crucial role in genetics, and are often discussed in studies of inheritance, population genetics, and in the mapping of genetic disease. They serve as reference points for identifying the position of genes on chromosomes.

Figure 12. Great blue heron pair flirting in breeding plumage.

Adaptive markers, onthe other hand, are genetic variations that are directly or indirectly associated with fitness-related traits and are subject to natural selection. These markers often show signatures of positive selection or balancing selection, reflecting their role in adaptation to different environmental conditions. For example, **single nucleotide polymorphisms (SNPs)** are the most common type of genetic variation in the genome and can be both neutral and adaptive. In the case of high-altitude adaptation in humans, certain SNPs have been identified as adaptive markers associated with the physiological response to hypoxia (low oxygen levels) at high altitudes. For instance, researchers have found that populations native to high-altitude regions, such as the Tibetan Plateau, have higher frequencies of specific SNPs in genes related to oxygen transport and regulation, such as EPAS1 and EGLN1. These SNPs are thought to confer adaptive advantages, such as increased hemoglobin levels or improved efficiency of oxygen utilization, allowing individuals to thrive in low-oxygen environments. By comparing the frequency of these adaptive SNPs between high-altitude and low-altitude populations, researchers can infer the role of natural selection in shaping genetic variation associated with altitude adaptation.

Media Attributions

Early Molecular Markers

Allozymes – the first true marker

Allozymes are variants of enzymes encoded by different alleles of the same gene, and they can be detected through **electrophoresis**, a technique that separates charged molecules based on their size and charge. Allozyme analysis emerged as one of the earliest methods for studying genetic variation in natural populations. Before the advent of DNA sequencing technologies, scientists relied on protein-based markers, such as allozymes, to study genetic diversity and population structure. The principle behind allozyme analysis is based on the fact that different alleles of a gene may produce enzymes with slightly different amino acid sequences, resulting in variations in enzyme mobility during electrophoresis. These differences can be visualized as distinct bands on a gel, allowing researchers to infer genetic variation within and between populations.

Methodology:

1. **Sample Collection and Preparation:** Researchers collect tissue samples (e.g., blood, muscle, or plant tissue) from individuals within a population.
2. **Protein Extraction:** Proteins are extracted from the tissue samples using biochemical methods.
3. **Electrophoresis:** The extracted proteins are separated based on their charge and size using electrophoresis. This involves applying an electric current to a gel matrix containing the protein samples. Proteins migrate through the gel at different rates based on their charge and size, resulting in distinct bands corresponding to different allozyme variants.
4. **Staining and Visualization:** After electrophoresis, the gel is

stained to visualize the allozyme bands. Each band represents a different allozyme variant encoded by different alleles of the same gene.

5. **Analysis:** Researchers analyze the allozyme banding patterns to estimate genetic diversity, population structure, gene flow, and evolutionary relationships among populations.

Allozyme markers have been widely used in molecular ecology to address various research questions, including (1) *assessing genetic diversity within populations*, (2) *investigating population structure and genetic differentiation*, (3) *estimating gene flow and migration patterns*, (4) *studying evolutionary relationships and phylogenetic reconstruction* and (5) *monitoring genetic responses to environmental changes and anthropogenic impacts*. Despite being an older technology, allozyme analysis laid the groundwork for modern molecular ecology and provided valuable insights into the genetic diversity and evolutionary dynamics of natural populations. While it has been largely replaced by DNA-based markers such as microsatellites and SNPs due to their higher resolution and ease of analysis, allozyme analysis remains relevant in certain contexts and continues to contribute to our understanding of biodiversity and conservation biology.

Limitations of Allozymes

Allozyme analysis, while a valuable tool in molecular ecology and evolutionary biology, has several limitations that researchers should consider. Firstly, allozyme analysis provides relatively **low resolution** compared to newer DNA-based markers like microsatellites and SNPs. This is because allozymes are the products of gene expression rather than the genes themselves, and they may not fully capture the underlying genetic variation present in a population. Secondly, allozyme analysis typically examines a limited number of enzyme loci, often ranging from a few to several dozen loci. This limited number of markers **may not provide sufficient coverage of the genome** to accurately capture genetic diversity and population structure, especially in species with large or complex genomes. Moreover, allozymes are subject to **genetic homoplasy**, where different alleles at different loci may produce the same allozyme phenotype. This can lead to misleading interpretations of genetic relationships and population structure, particularly when using a small number of allozyme loci. Furthermore, allozyme analysis **requires specialized laboratory techniques** for protein extraction, electrophoresis, and staining, which can be time-consuming, labor-intensive, and technically challenging. Additionally, obtaining high-quality allozyme data may require optimization of experimental conditions and protocols. Additionally, allozyme expression can be influenced by environmental factors such as temperature, pH, and substrate availability. **Variation in environmental conditions among sampling locations may confound allozyme patterns** and complicate interpretations of genetic diversity and population structure. Lastly, allozyme analysis focuses on variations in protein-coding genes and **may not capture genetic variation in non-coding regions of the genome**, which can play important roles in evolution and adaptation.

Restriction Fragment Length Polymorphism (RFLP)

RFLPs are variations in DNA sequences that result in differences in the lengths of DNA fragments when cut with restriction enzymes. These variations occur due to differences in the recognition sites of restriction enzymes among individuals or populations. In RFLP analysis, DNA samples are first digested with restriction enzymes, which cut the DNA at specific recognition sites. The resulting DNA fragments are then separated by size using gel electrophoresis. Because individuals may have different alleles with different restriction sites, the resulting banding patterns on the gel can reveal genetic variation within and between populations. RFLPs were one of the earliest types of molecular markers used in genetics and have been widely applied in population genetics, phylogenetics, and forensic science. They have been used to study genetic diversity, population structure, evolutionary relationships, and gene mapping in various organisms.

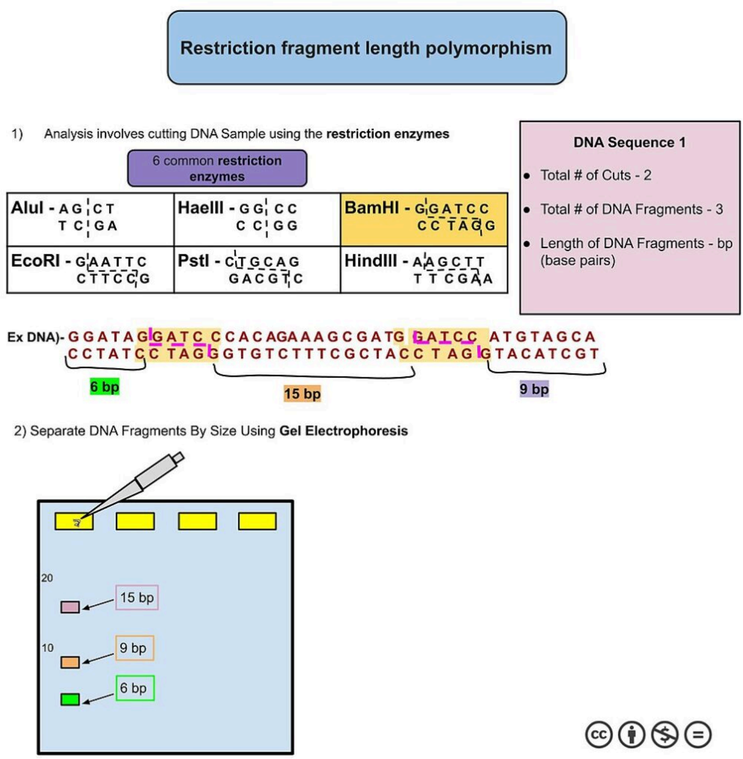


Figure 13. Diagram showing overview of Restriction Fragment Length Polymorphism.

However, RFLP analysis has several limitations:

1. **Low Throughput**: RFLP analysis typically examines a limited number of restriction sites, and the process can be time-consuming and labor-intensive, especially when analyzing multiple loci or large sample sizes.
2. **Technically Challenging**: RFLP analysis requires specialized laboratory techniques for DNA extraction, restriction enzyme digestion, gel electrophoresis, and staining, which can be technically challenging and prone to error.
3. **Limited Resolution**: RFLPs may have lower resolution compared to newer molecular markers such as microsatellites and SNPs. The size differences between DNA fragments may be relatively small, leading to overlapping bands on the gel and difficulty in accurately determining fragment sizes.
4. **Inability to Detect Point Mutations**: RFLPs are not suitable for detecting single nucleotide polymorphisms (SNPs) or point mutations, as they primarily detect variations in DNA fragment length resulting from differences in restriction enzyme recognition sites.
5. **Limited Information Content**: RFLPs provide limited information about the genetic variation within a population compared to DNA sequence data. They may not capture variation in non-coding regions of the genome or provide insights into gene function or gene regulation.

Despite these limitations, RFLP analysis has been instrumental in early genetic studies and has contributed valuable insights into genetic diversity and population structure in various organisms. However, it has largely been supplanted by more advanced molecular markers with higher throughput and resolution, such as microsatellites, SNPs, and next-generation sequencing technologies.

THINK/PAIR/SHARE

1. Given that both allozyme and RFLP analyses have lower resolution compared to modern DNA-based markers like microsatellites and SNPs, discuss the potential scenarios or research contexts where these older techniques might still be advantageous or necessary. What are the trade-offs involved in choosing these methods over newer technologies?

Media Attributions

- 1024px-Restriction_Fragment_Length_Polymorphism © Lolyas39 is licensed under a CC BY-SA (Attribution ShareAlike) license

Development of PCR & Sanger Sequencing

Polymerase Chain Reaction

As you have already learned in Introductory Biology, **Polymerase Chain Reaction (PCR)**, is a powerful molecular biology technique used to amplify specific DNA sequences. Developed in the 1980s by Kary Mullis, PCR revolutionized molecular biology by enabling rapid and efficient amplification of DNA, even from minute quantities of starting material. Before the development of PCR, DNA amplification was a laborious and time-consuming process that often involved cloning DNA fragments into bacterial vectors and culturing them in bacterial cells. Mullis' invention of PCR provided a simple and efficient alternative, allowing researchers to amplify DNA in a test tube through a series of repeated temperature cycles.

Principle of PCR

PCR amplifies specific DNA sequences through a series of three basic steps:

1. **Denaturation:** The DNA template is heated to a high temperature (typically 94-98°C), causing the double-stranded DNA to denature into single strands.
2. **Annealing:** The reaction temperature is lowered to allow primers (short DNA sequences complementary to the target DNA) to bind to their complementary sequences on the template DNA.
3. **Extension:** The reaction temperature is raised, and a DNA polymerase enzyme synthesizes new DNA strands by extending from the primers along the template DNA. This process generates complementary copies of the target DNA

sequence.

By repeating these temperature cycles multiple times (usually 20-40 cycles), PCR exponentially amplifies the target DNA sequence, resulting in millions to billions of copies of the desired DNA fragment.

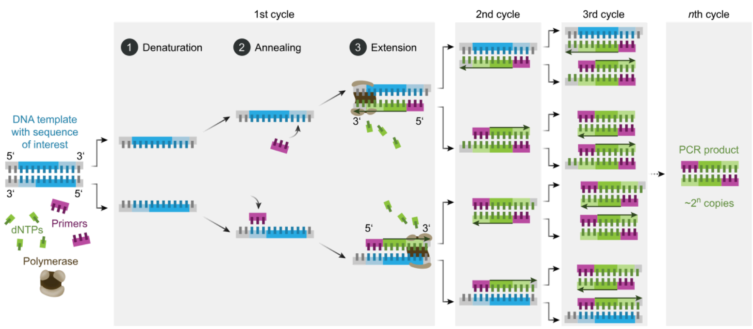


Figure 14. Overview of the Polymerase Chain Reaction and the various ingredients and stages required for amplification of a specific gene region from genomic DNA (gDNA)

Applications of PCR:

PCR has numerous applications in molecular biology, genetics, forensics, and medical diagnostics, including:

- Gene cloning and DNA sequencing
- Gene expression analysis (RT-PCR)
- DNA fingerprinting and forensic analysis
- Pathogen detection and microbial identification
- Environmental DNA (eDNA) analysis

Di-deoxyribonucleotide ('dideoxy') chain-termination sequencing

In the late 1970s, Frederick Sanger and his colleagues at the Medical Research Council Laboratory of Molecular Biology in Cambridge, UK, developed a groundbreaking method for sequencing DNA known as **dideoxy sequencing**, or Sanger sequencing. This method represented a significant advancement in molecular biology, offering a reliable and efficient means of determining the sequence of nucleotide bases in a DNA molecule. Before the advent of dideoxy sequencing, DNA sequencing was a cumbersome and time-consuming process. Early methods, such as Maxam-Gilbert sequencing, relied on chemical treatments to break DNA at specific nucleotide positions and infer the sequence indirectly. These methods were labor-intensive and often yielded limited amounts of sequence data. Dideoxy sequencing revolutionized DNA sequencing by introducing a new approach based on the incorporation of chain-terminating dideoxynucleotides (ddNTPs) during DNA synthesis. Unlike normal deoxynucleotides (dNTPs), which contain a 3′ hydroxyl group required for chain elongation, ddNTPs lack this group, leading to termination of DNA synthesis when they are incorporated into the growing DNA strand. In dideoxy sequencing, the DNA template is first denatured and annealed with a short DNA primer. DNA synthesis is then initiated using DNA polymerase and a mixture of normal dNTPs and a small amount of ddNTPs labeled with fluorescent dyes. As DNA synthesis proceeds, ddNTPs are randomly incorporated into the growing DNA strand, leading to chain termination at specific nucleotide positions.

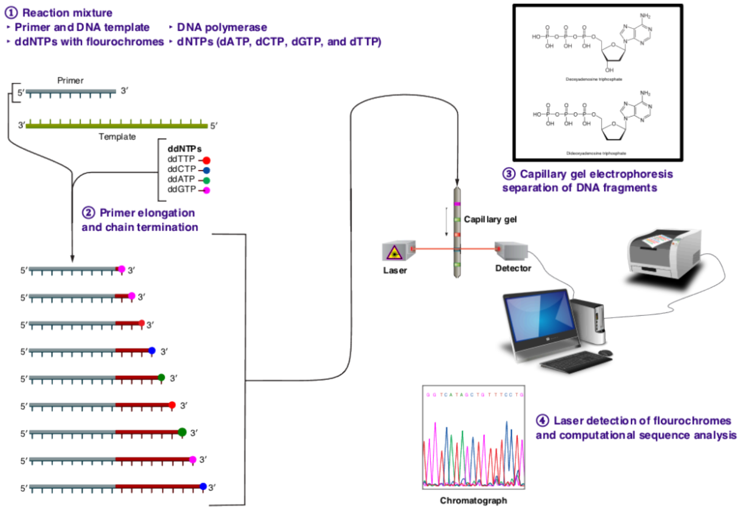


Figure 15. Overview of di-deoxyribonucleotide chain-termination sequencing, also known as Sanger sequencing.

Dideoxy sequencing had a profound impact on molecular biology and genetics, enabling high-throughput DNA sequencing with unprecedented accuracy and reliability. Its development paved the way for large-scale genome sequencing projects, such as the Human Genome Project, which aimed to sequence the entire human genome. Moreover, dideoxy sequencing facilitated the development of various molecular markers, including microsatellites, single nucleotide polymorphisms (SNPs), and restriction fragment length polymorphisms (RFLPs). These markers have become essential tools in molecular ecology and evolutionary biology, allowing researchers to study genetic diversity, population structure, and evolutionary relationships at the molecular level.

PCR & Sanger Sequencing: The Dynamic Duo

The combination of PCR with Sanger sequencing revolutionized molecular ecology and evolutionary biology by enabling rapid and accurate DNA sequencing of specific target regions. Researchers could now amplify target DNA sequences of interest using PCR and then sequence them using Sanger sequencing, providing detailed information about genetic variation, population structure, and evolutionary relationships. This approach facilitated the development of various molecular markers, such as microsatellites, SNPs, and mitochondrial DNA sequences, which have become indispensable tools in molecular ecology and evolutionary biology. These markers allow researchers to study genetic diversity, population dynamics, gene flow, and evolutionary processes in natural populations with unprecedented resolution and precision. Overall, the integration of PCR and Sanger sequencing has had a transformative impact on molecular ecology and evolutionary biology, enabling groundbreaking research and advancing our understanding of the genetic basis of biodiversity, adaptation, and evolution in natural systems.

PCR-based markers rely on the Polymerase Chain Reaction (PCR) to amplify specific DNA sequences of interest. This technique allows for rapid and efficient amplification of target DNA sequences, enabling high-throughput analysis of genetic variation. PCR-based markers are highly sensitive and specific, with the ability to amplify DNA from minute quantities of starting material and detect single nucleotide differences. Examples of PCR-based markers include microsatellites, Single Nucleotide Polymorphisms (SNPs), Amplified Fragment Length Polymorphisms (AFLPs), Restriction Fragment Length Polymorphisms (RFLPs), and mitochondrial DNA (mtDNA) sequences. PCR-based markers are widely used in genetics, genomics, forensics, medical diagnostics, and molecular ecology due to their versatility and scalability.

Non-PCR-based markers on the other hand, do not involve PCR amplification and rely on alternative methods for detecting genetic variation. These markers may include variations in DNA sequence, protein products, or chromosomal structure. While non-PCR-based markers may offer certain advantages such as simplicity and lower cost, they may also have limitations in terms of sensitivity, specificity, and resolution. Examples of non-PCR-based markers include allozymes, restriction site polymorphisms, karyotypes, and isozymes. Non-PCR-based markers have been widely used in molecular ecology, population genetics, and evolutionary biology, providing valuable insights into genetic diversity and evolutionary processes.

Table 2. Summary of differences between PCR and Non-PCR based molecular markers

Aspect	PCR-Based Markers	Non-PCR-Based Markers
Methodology	Amplify specific DNA sequences using PCR	Do not involve PCR amplfication
Amplification	Exponential amplification of target DNA sequences	No amplification step involved
Speed	Rapid amplification within hours	Time-consuming, may take days or weeks
Sensitivity	Highly sensitive, can amplify DNA from minute quantities of starting material	Less sensitive, may require larger amounts of starting material
Specificity	Highly specific, amplifies only target DNA sequences with complementary primers	Less specific, may capture non-target DNA sequences
Accuracy	High accuracy, low error rate	Accuracy may vary depending on the method
Throughput	High throughput, allows for simultaneous amplification of multiple targets	Lower throughput, may analyze fewer loci simultaneously
Applications	Widely used in genetics, genomics, forensics, medical diagnostics, and molecular ecology	Commonly used in molecular ecology, population genetics, and evolutionary biology
Examples	Microsatellites, SNPs, AFLPs, RFLPs, mtDNA sequences	Allozymes, restriction site polymorphisms, karyotypes, isozymes

Cost	Moderate to high, depending on the number of markers and samples	Lower cost for some methods, but may require specialized reagents or equipment
Data Analysis	Requires sequencing or fragment analysis after PCR amplification	Direct analysis of electrophoretic patterns or sequencing data
Resolution	High resolution, can detect single nucleotide differences and subtle genetic variations	Resolution may vary depending on the method, may not detect single nucleotide differences
Flexibility	Highly flexible, allows customization of primer sequences and amplification conditions	Less flexible, dependent on available markers and methods
Ease of Use	Requires specialized equipment and expertise in PCR techniques	May be simpler and require less specialized equipment

Media Attributions

-
-

Recommended Reading

The following are recommended readings – we will be discussing at least one of these articles in next week's group discussion:

1. Karl SA, Toonen RJ, Grant WS, Bowen BW (2012) Common misconceptions in molecular ecology: echoes of the modern synthesis. *Molecular Ecology* 21: 4171 – 4189.

2. Galtier N, Nabholz B, Glemin S, Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology* 18: 4541 – 4550.

3. Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* 17: 81 – 92.

4. Hadrys H, Balick M, Schierwater B (1992) Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Molecular Ecology* 1: 55 – 63.

5. Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology letters* 9: 615 – 629.

6. Mittell EA, Nakagawa S, Hadfield JD (2015) Are molecular markers useful predictors of adaptive potential? *Ecology letters* 18: 772 – 778.

PART IV

SPECIES AND POPULATIONS

LEARNING OBJECTIVES

At the end of the chapter you should be able to:

1. Explain what the 'species problem' is.
2. Explain the practical implications of the species problem.
3. Compare and contrast the three main species concepts.
4. Define and explain what an Operational Taxonomic Unit (OTU) is.
5. Define speciation and explain how species barriers emerge.
6. Distinguish between cryptic, sibling and sister species.
7. Define and explain what evolutionary significant units (ESUs) are.
8. Explain how hybrids can interfere with the speciation process.
9. Explain how molecular methods can be used to distinguish between species.

The **“species problem”** refers to the ongoing debate among biologists and philosophers of science over how to define and identify species. The problem stems from the fact that different concepts and criteria for defining species often lead to inconsistent and conflicting classifications. It highlights the complexities and nuances of defining and identifying species. In molecular ecology and evolution, this issue is particularly pertinent, as researchers rely on clear species boundaries to study genetic diversity, adaptation, and speciation. While different species concepts offer various advantages and limitations, the continued integration of molecular data provides new insights but also presents new challenges in resolving the species problem.

Figure 16. A female moth belonging to the family Anobinae is believed to be part of a larger cryptic species complex (Anoba_sp._trigonoides)

Media Attributions

- Anoba_sp._trigonoides_species_complex,_female_(Itapua,_Paraguay) © Lafontaine JD, Walsh JB is licensed under a CC BY (Attribution) license

Species Concepts

A **species concept** is a theoretical framework used to define and identify species, guiding biologists in distinguishing one species from another. There are several species concepts, each offering a different perspective on what constitutes a species. The most common ones include:

1. **Biological Species Concept (BSC)**
    - **Definition:** A species is a group of interbreeding natural populations that are reproductively isolated from other such groups.
    - **Strengths:** Emphasizes reproductive isolation, which is crucial for maintaining distinct gene pools.
    - **Weaknesses:** Not applicable to asexual organisms, fossils, or instances of hybridization.
2. **Morphological Species Concept (MSC)**
    - **Definition:** A species is defined by its distinct morphological features.
    - **Strengths:** Useful for classifying fossils and organisms with limited genetic data.
    - **Weaknesses:** Subjective and can be misleading due to phenotypic plasticity or cryptic species.
3. **Phylogenetic Species Concept (PSC)**
    - **Definition:** A species is the smallest group of individuals that share a common ancestor and exhibit unique traits.
    - **Strengths:** Applies to asexual and sexual organisms and focuses on evolutionary relationships.
    - **Weaknesses:** Can lead to excessive splitting (over-

splitting) and might ignore ecological aspects.

4. **Ecological Species Concept (ESC)**
    - **Definition:** A species is a group of organisms exploiting a single ecological niche.
    - **Strengths:** Highlights the role of natural selection and ecological factors.
    - **Weaknesses:** Difficult to apply and distinguish overlapping niches.

In the realm of molecular ecology and evolution, the conundrum known as the species problem critically impacts the scientific community's understanding of genetic diversity, population structure, adaptation, and speciation. This narrative explores the multifaceted challenges posed by this problem through four key issues: hybridization and gene flow, cryptic species, speciation and evolutionary history, and conservation and biodiversity.

Hybridization and Gene Flow

One of the primary challenges arises with hybridization and gene flow between closely related species. This blending can obscure the boundaries that define species, making it difficult to identify distinct groups. The impact of such hybridization is profound, as molecular data—like gene sequences—often unveil events of genetic intermingling. This revelation poses significant challenges to traditional species concepts such as the Biological Species Concept (BSC), which relies heavily on reproductive isolation as a marker of species delineation.

Cryptic Species

Another layer of complexity is introduced by the presence of cryptic species. These are groups of organisms that, while morphologically indistinguishable, are genetically distinct. The revelation of these hidden species often comes through advanced molecular techniques. The Morphological Species Concept (MSC) struggles to identify these cryptic species due to its reliance on physical traits alone. In contrast, molecular-based species concepts, like the Genotypic Cluster Species Concept (GCSC) or the Phylogenetic Species Concept (PSC), prove more effective by focusing on genetic distinctions.

Speciation and Evolutionary History

Understanding the processes of speciation and the evolutionary history of organisms also requires clear species delineation. Molecular data plays a critical role here, as it helps construct detailed phylogenies that trace the lineage relationships between species. However, the varied and sometimes conflicting species concepts can lead to inconsistent classifications, complicating the understanding of evolutionary pathways and relationships.

Conservation and Biodiversity

Finally, the species problem has significant implications for conservation and the understanding of biodiversity. Accurate species identification is crucial for setting conservation priorities and managing ecosystems effectively. Misclassifications can lead to either an overestimation or underestimation of biodiversity, which can skew conservation efforts and policy decisions.

In conclusion, each of these issues underscores the ongoing challenges faced by researchers in molecular ecology and evolution due to the species problem. As molecular techniques advance and provide deeper insights into genetic relationships, the scientific community continues to grapple with these complex and evolving questions about what exactly makes a species a distinct entity in the natural world.

Operational Taxonomic Units

Operational Taxonomic Units (OTUs) are a pragmatic solution to a fundamental problem in taxonomy: how to classify organisms, particularly when dealing with molecular data. OTUs are often used in microbial ecology, evolutionary biology, and molecular ecology to categorize groups of closely related individuals. An OTU is a cluster of organisms grouped together based on a set level of similarity, often using DNA sequences. OTUs serve as proxies for species or higher-level taxonomic groups when direct classification is impractical or when detailed taxonomic information is unavailable.

How are OTUs Defined?

OTUs are commonly defined using DNA sequence data. The process generally involves:

1. **Sequencing:** DNA is extracted and sequenced from a sample.
2. **Clustering:** Sequences are grouped into OTUs based on a similarity threshold, commonly set at 97% similarity for bacterial 16S rRNA genes. This threshold is often chosen because it's roughly equivalent to species-level divergence for many bacteria, though this varies.
3. **Identification:** OTUs are then used for ecological or evolutionary analyses, either identified to known species or left as "unclassified" when no match exists

OTUs and the Species Problem

The use of OTUs highlights the practical challenges associated with the species problem, which revolves around how to define and identify species. OTUs provide a functional approach when precise species definitions are elusive, especially in microbial ecology where many organisms lack clear morphological distinctions or well-studied taxonomies.

1. **Flexibility:** OTUs allow researchers to group organisms based

on genetic similarity without requiring formal taxonomic classification, which can be especially useful in understudied or complex groups.

2. **Scalability:** OTUs offer a scalable method for analyzing large datasets from next-generation sequencing, where defining species might be impractical or impossible due to the sheer number of unique sequences.

While OTUs offer a practical solution, they also underscore the limitations and variability in species concepts:

1. **Arbitrariness:** The similarity threshold used to define OTUs is somewhat arbitrary, reflecting a consensus or a specific research focus rather than a fundamental biological reality.
2. **Lack of Reproductive or Ecological Insight:** OTUs based purely on genetic data might not reflect reproductive isolation (as emphasized in the Biological Species Concept) or ecological niche differentiation (as emphasized in the Ecological Species Concept).

THINK/PAIR/SHARE

Considering the advantages and limitations of using Operational Taxonomic Units (OTUs) in molecular ecology, discuss how the arbitrary nature of similarity thresholds might affect ecological and evolutionary studies. What are the potential consequences of relying on OTUs for biodiversity assessments and species identification, and how might this impact conservation efforts and ecological research?

Speciation

Speciation is the evolutionary process by which new biological species arise. This process is central to biodiversity and evolutionary theory. Speciation typically involves genetic, ecological, and behavioral changes that lead to reproductive isolation and divergence between populations. The processes of anagenesis and cladogenesis represent two different patterns of evolutionary change, each with unique implications for biodiversity. Through various modes of speciation, such as allopatric, sympatric, parapatric, and peripatric, organisms diversify and adapt to their environments, highlighting the dynamic nature of evolutionary change.

Anagenesis, also known as phyletic evolution, occurs when a single lineage evolves over time into a new species without branching. In this process, the original species is replaced by a new one. Anagenesis involves gradual changes in the population over time, driven by factors like natural selection, genetic drift, or mutations. The species evolves as a whole, and the original species no longer exists in its ancestral form. For example, in the fossil record, anagenesis might appear as a continuous series of slightly altered forms leading from an ancestor to a descendant.

Cladogenesis, or branching evolution, occurs when a lineage splits into two or more separate species. This results in increased biodiversity because the original species and new species coexist. Cladogenesis typically begins with some form of reproductive isolation between populations, which can be geographic, ecological, behavioral, or genetic. The isolated populations diverge due to natural selection, genetic drift, or mutations, eventually leading to distinct species. A classic example of cladogenesis is adaptive radiation, where one species diversifies into several distinct species to fill different ecological niches.

Both anagenesis and cladogenesis can occur through different *modes of speciation based on the geographic and ecological context of the populations:*

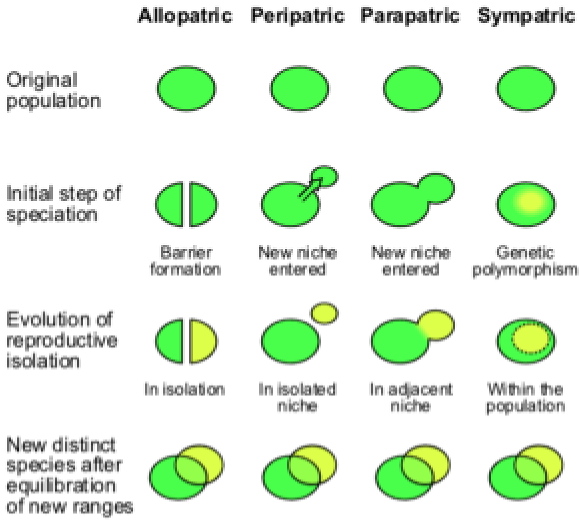


Figure 17. Modes of speciation.

Allopatric speciation occurs when populations are geographically isolated, leading to reproductive isolation and divergence. Geographic barriers, such as mountains or rivers, separate populations, which then evolve independently. Over time, genetic changes accumulate, resulting in speciation.

Sympatric speciation occurs when populations diverge into new species within the same geographic area. Reproductive isolation arises due to factors like polyploidy (common in plants), disruptive selection, or behavioral isolation. Over time, distinct species emerge without physical barriers.

Parapatric speciation occurs when populations are adjacent but not fully overlapping, allowing for limited gene flow. Divergence occurs due to different environmental pressures or mating preferences along a gradient, eventually leading to reproductive isolation and speciation.

Peripatric speciation is a form of allopatric speciation where a small, isolated population diverges from a larger ancestral population. The isolated population, often at the periphery of the ancestral range, experiences rapid evolutionary change due to genetic drift or founder effects, leading to speciation.

Anagenesis and cladogenesis differ in how they contribute to biodiversity. Cladogenesis increases biodiversity by creating new branches in the evolutionary tree, while anagenesis does not increase species diversity but represents evolutionary change within a single lineage. Anagenesis involves linear evolution, while cladogenesis involves branching evolution. In anagenesis, the original species transforms into a new species, whereas in cladogenesis, the original species continues to exist alongside the newly formed species

Media Attributions

Reproductive Isolating Mechanisms

Speciation, the process through which new species arise, is often facilitated by reproductive isolating mechanisms. These mechanisms prevent gene flow between populations, thereby maintaining the distinctiveness of species. Reproductive isolating mechanisms are broadly categorized into two types: **prezygotic** and **postzygotic.** Each type plays a unique role in preventing interbreeding and ensuring reproductive isolation.

Prezygotic Isolating Mechanisms

Prezygotic isolating mechanisms prevent fertilization from occurring. They act before the zygote (fertilized egg) is formed and are particularly effective at preventing different species from mating or ensuring that if mating occurs, fertilization does not happen. These mechanisms include:

1. **Temporal Isolation**
    - **Definition:** Temporal isolation occurs when species breed at different times. This can involve differences in the timing of day, season, or year when mating occurs.
    - **Example:** Two closely related species of frogs might inhabit the same area but breed in different seasons, preventing interbreeding.
2. **Habitat Isolation**
    - **Definition:** Habitat isolation happens when species live in

different habitats or ecological niches, reducing the likelihood of encounters.
    - **Example:** Two species of garter snakes might live in the same geographic area, but one lives primarily in water while the other lives on land, limiting opportunities for mating.

3. **Behavioral Isolation**

    - **Definition:** Behavioral isolation arises from differences in mating behaviors or courtship rituals that prevent different species from recognizing each other as suitable mates.
    - **Example:** Different species of birds might have unique songs or mating dances, preventing cross-species attraction and mating.

4. **Mechanical Isolation**

    - **Definition:** Mechanical isolation occurs when anatomical differences prevent successful mating between species.
    - **Example:** Two species of insects might have differently shaped reproductive organs, making copulation impossible.

5. **Gametic Isolation**

    - **Definition:** Gametic isolation happens when the sperm and egg of different species are incompatible, preventing fertilization.
    - **Example:** In many marine animals, such as corals, different species release sperm and eggs into the water simultaneously, but only sperm and eggs of the same species can successfully fuse.

Postzygotic Isolating Mechanisms

Postzygotic isolating mechanisms come into play after fertilization and typically reduce the viability or reproductive success of hybrid offspring. These mechanisms include:

1. **Hybrid Inviability**
    - **Definition:** Hybrid inviability occurs when hybrid offspring fail to develop properly or die before reaching reproductive maturity.
    - **Example:** Crosses between different species of frogs might result in embryos that fail to develop or larvae that do not survive to adulthood.
2. **Hybrid Sterility**
    - **Definition:** Hybrid sterility happens when hybrid offspring are sterile and cannot produce viable offspring.
    - **Example:** The mule, a hybrid of a horse and a donkey, is sterile and unable to produce offspring, effectively isolating the parent species.
3. **Hybrid Breakdown**
    - **Definition:** Hybrid breakdown occurs when the first-generation hybrids are viable and fertile, but subsequent generations suffer from reduced fitness or sterility.
    - **Example:** Certain hybrid plants might produce viable and fertile offspring initially, but in subsequent generations, these hybrids exhibit abnormalities or sterility.

Reproductive isolating mechanisms are critical in maintaining species boundaries and facilitating speciation. By preventing interbreeding or reducing hybrid fitness, these mechanisms ensure that distinct species retain their unique characteristics. Prezygotic mechanisms, such as temporal, habitat, behavioral, mechanical, and gametic isolation, prevent fertilization, while postzygotic mechanisms, such as hybrid inviability, sterility, and breakdown, act after fertilization to reduce hybrid viability or fertility. Together, these mechanisms illustrate the intricate ways in which nature maintains biodiversity and drives evolutionary divergence.

THINK/PAIR/SHARE

1. Compare and contrast the effectiveness of prezygotic and postzygotic isolating mechanisms in maintaining species boundaries. Which type of mechanism do you think plays a more significant role in the early stages of speciation, and why? Provide examples to support your argument.
2. Discuss the potential evolutionary advantages and disadvantages of hybridization in natural populations. How might postzygotic isolating mechanisms, such as hybrid sterility and hybrid breakdown, influence the evolutionary trajectories of species that occasionally interbreed?

Evolutionary Significant Units

An **Evolutionary Significant Unit (ESU)** is a fundamental concept in conservation biology, referring to a population of organisms that is considered distinct for conservation purposes. This concept helps prioritize populations for protection based on their unique evolutionary traits or ecological roles. The designation of ESUs is particularly relevant for managing and safeguarding endangered or threatened species and subspecies. An ESU is defined as a population or group of populations that is genetically distinct from other populations of the same species. The criteria for designating an ESU typically involve two key aspects:

(1) The population must be substantially reproductively isolated from other populations. This isolation can be evidenced through genetic data showing significant differences in allele frequencies, distinct genetic markers, or unique haplotypes **(Genetic differentiation)**.

(2) The population must represent an important component of the evolutionary legacy of the species. This can be indicated by unique adaptations to local environmental conditions, distinctive ecological roles, or specialized behaviors that are not found in other populations. **(Adaptive significance)**

Importance of ESUs

The concept of ESUs is crucial in conservation biology for several reasons. ESUs help preserve the genetic diversity within a species, which is vital for the species' long-term survival and adaptability. Genetic diversity enables populations to adapt to changing environmental conditions, resist diseases, and maintain ecological functions. Moreover, by focusing on ESUs, conservation efforts maintain the evolutionary potential of a species, allowing it to continue evolving and adapting.

ESUs also guide conservation priorities by highlighting populations that are distinct and therefore irreplaceable. Conservation resources are often limited, and prioritizing ESUs ensures that unique evolutionary lineages or adaptations are preserved.

Applications of ESUs

ESUs have been widely used in the conservation of various species. For instance, Pacific salmon populations in North America are managed as ESUs to preserve their unique genetic and ecological characteristics. Different populations have evolved adaptations to specific river systems, and conserving these ESUs helps maintain the overall diversity and resilience of salmon species. Similarly, different populations of orcas are considered separate ESUs due to their distinct genetic, behavioral, and ecological characteristics. These populations exhibit unique hunting strategies and social structures, highlighting the importance of conserving each ESU to preserve the species' overall diversity. The giant panda has also been divided into several ESUs based on genetic and ecological differentiation, reflecting different adaptations to local environments and underscoring the importance of conserving these units to maintain the panda's evolutionary potential and ecological diversity.

Criticism and Challenges of ESUs

While the concept of ESUs is useful, it faces several criticisms and challenges. The criteria for defining ESUs can be somewhat arbitrary or subjective, leading to inconsistencies. Different studies may use different genetic or ecological markers, leading to varying definitions of what constitutes an ESU. Additionally, some critics argue that ESUs place too much emphasis on genetic differentiation at the expense of ecological or behavioral factors. Populations with significant ecological or behavioral differences may be overlooked if they lack clear genetic distinctions.

Furthermore, the designation of ESUs can lead to conflicts in conservation priorities, especially when different ESUs within the same species are in competition for limited resources or habitats. Balancing the needs of different ESUs within a species can be challenging, particularly when they occupy overlapping or adjacent habitats.

THINK/PAIR/SHARE

1. Considering the criteria for designating Evolutionary Significant Units (ESUs), discuss the potential challenges and limitations of relying primarily on genetic data to identify ESUs. How might the inclusion of ecological and behavioral factors improve the robustness of ESU designations, and what are the potential trade-offs of this more holistic approach?
2. Reflect on the practical implications of managing species based on ESUs. How might the designation of

multiple ESUs within a single species affect conservation strategies, resource allocation, and policy decisions? Provide examples of potential conflicts or synergies that could arise when different ESUs require distinct conservation actions

Cryptic Species

Cryptic species are groups of organisms that are morphologically indistinguishable from one another but are genetically distinct enough to be considered separate species. These species pose significant challenges for taxonomists and ecologists because traditional methods of species identification, which often rely on visible physical traits, fail to distinguish them. Instead, cryptic species can only be reliably identified through genetic, biochemical, or sometimes behavioral analyses. The concept of cryptic species emphasizes the complexity of biodiversity and the need for molecular tools in modern taxonomy. An example of cryptic species is found in the Bocourt's terrapin (*Mauremys rivulata*), a freshwater turtle found in the Mediterranean. Initially, these turtles were thought to be a single species, but genetic studies have revealed that they are a complex of several cryptic species, each with distinct genetic profiles but nearly identical appearances. Another well-known case involves the African elephant, where what was once thought to be a single species has been identified as two distinct species, the African bush elephant (*Loxodonta africana*) and the African forest elephant (*Loxodonta cyclotis*), through DNA analysis.

Figure 18. LEFT – African forest elephant (Loxodonta cyclotis), RIGHT – African bush elephant (Loxodonta africana)

The terms “cryptic species,” “sibling species,” and “sister species” are often used in discussions of species differentiation and speciation but have distinct meanings:

1. **Cryptic Species**: As described, these are species that appear identical or nearly identical in morphology but are genetically distinct and reproductively isolated. Cryptic species are often discovered through molecular techniques rather than traditional morphological examination.
2. **Sibling Species**: Sibling species are a type of cryptic species; they are two or more species that are morphologically similar but genetically distinct. The term “sibling” implies that these species are closely related and may have recently diverged in evolutionary terms. Sibling species are often sympatric (living in the same geographic area) and may occupy slightly different ecological niches or exhibit subtle differences in behavior or physiology.
3. **Sister Species**: Sister species are the closest relatives to each other on the evolutionary tree, meaning they share a most recent common ancestor not shared with any other species. While sister species can be morphologically distinct or similar, they are defined primarily by their phylogenetic relationship rather than their appearance. Sister species can result from recent speciation events or represent lineages that diverged longer ago but have maintained a close evolutionary relationship.

Using Molecular Tools to Understand Cryptic Species

DNA barcoding has emerged as a pivotal technique for species identification. This method relies on sequencing a standardized region of the genome, typically the mitochondrial cytochrome c oxidase I (COI) gene in animals, to produce a unique genetic identifier or “barcode” for each species. These barcodes can then be compared to a comprehensive reference library to confirm species identity. The utility of DNA barcoding extends to delineating cryptic species by uncovering genetic disparities between organisms that appear identical. For example, in the butterfly genus *Astraptes*, what was once thought to be a single species was revealed to be at least ten cryptic species using DNA barcoding. Additionally, this technique aids in discovering new species by identifying barcodes that do not match any known sequences in the database, indicating potentially unrecorded biodiversity.

Metabarcoding extends the principles of DNA barcoding to analyze entire biological communities from environmental samples such as soil, water, or feces. This method extracts DNA from these bulk samples and identifies multiple species simultaneously by comparing the sequences to a reference database. Metabarcoding is particularly effective for delineating cryptic species within complex ecosystems where traditional survey methods may miss less conspicuous or elusive organisms. Moreover, this approach can signal the presence of new species by revealing sequences that are absent in current reference databases, particularly in ecosystems that are difficult to sample or are less studied, like deep-sea or dense rainforest habitats.

Environmental DNA (eDNA) barcoding represents another leap in non-invasive species detection. By extracting DNA directly from environmental matrices such as water or sediment, eDNA barcoding captures the genetic signatures of organisms present in the ecosystem without the need for direct observation or specimen collection. This technique is invaluable for monitoring elusive or rare cryptic species and has proven especially powerful in aquatic environments. For instance, eDNA sampling has successfully identified different salmonid species in rivers, uncovering distribution patterns that are not easily observed through conventional methods. Furthermore, eDNA barcoding can lead to the discovery of new species by capturing genetic material from a range of organisms, including previously unknown microbial, fish, and amphibian species.

Media Attributions

- Picture1 © Matt Muir (African Forest Elephant) and Ray in Manila (African Bush Elephant) is licensed under a CC BY (Attribution) license

Hybridization and Introgression

Hybridization refers to the interbreeding of individuals from two different species or genetically distinct populations, resulting in offspring called hybrids. This process is a naturally occurring phenomenon that has significant implications for the evolutionary trajectory of species. Hybrids can display characteristics that are intermediate between their parent species or, in some cases, traits that surpass those of either parent, a phenomenon known as hybrid vigor or heterosis. An example of natural hybridization occurs between polar bears (*Ursus maritimus*) and grizzly bears (*Ursus arctos*). Their hybrids, known as pizzly or grolar bears, exhibit a combination of physical and behavioral traits from both species. The increasing frequency of these hybrids is attributed to climate change, which causes overlapping ranges as polar bears are forced southward.

Evolutionary Outcomes of Hybridization

The evolutionary outcomes of hybridization are varied and complex, impacting genetic diversity and biodiversity in different ways:

1. **Loss of Genetic Diversity**: When hybridization occurs between a rare species and a more common one, the gene pool of the rare species may be swamped, potentially leading to genetic homogenization and loss of unique genetic diversity. This process, called introgression, can dilute the distinct genetic identity of the rare species and pose a risk of extinction.
2. **Gain of Genetic Diversity**: Conversely, hybridization can introduce new genetic variations into populations. These novel combinations of genes can sometimes confer adaptive advantages, enabling hybrids to exploit new ecological niches

or adapt to changing environmental conditions. This process can lead to the emergence of new hybrid species that are reproductively isolated from their parent species.

3. **Maintenance of Hybrid Zones**: Hybridization can result in the creation of stable hybrid zones, regions where hybrids persist over generations due to ongoing gene flow between parent populations. These zones can act as laboratories of evolution, where the dynamics of genetic mixing and selection provide insights into speciation processes.

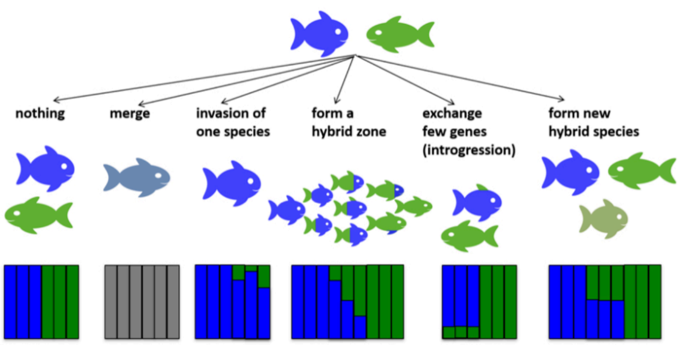


Figure 19. Evolutionary outcomes of hybridization between two different species (represented by blue and green fish). Possible scenarios include: (1) No significant change (nothing), (2) Merging of the two species into one homogeneous population, (3) Invasion of one species into the other's habitat leading to dominance, (4) Formation of a stable hybrid zone where hybrids (blue-green fish) coexist with parent species, (5) Introgression, where a few genes are exchanged between species, and (6) Formation of a new hybrid species distinct from the parent species. Each scenario is depicted with corresponding genetic patterns, highlighting the varying degrees of genetic mixing and integration.

Case Study: Ligers!

Ligers are the hybrid offspring of a male lion (*Panthera leo*) and a female tiger (*Panthera tigris*). These hybrids are known for their extraordinary size, often surpassing both parent species in weight and height. This size is attributed to a phenomenon known as "hybrid vigor" or heterosis, where the hybrid exhibits traits that are superior to either of the parent species, in this case, in terms of growth. Ligers exhibit a blend of physical features from both lions and tigers: they may have a faint striped pattern on a tawny background and sometimes display a mane, albeit less pronounced than that of a lion. Ligers are sterile in most cases, especially the males. This sterility means that these hybrids cannot contribute to the gene pool, leading to a dead-end in terms of genetic lineage. Therefore, while they are a fascinating example of interspecies hybridization, they do not contribute to the genetic diversity of either parent species. Alternatively, although ligers themselves do not contribute to genetic diversity due to their typical sterility, the concept of hybrid vigor observed in ligers shows how hybridization can potentially introduce beneficial traits. In a natural setting, if similar hybrids were fertile, these traits could be passed on, leading to increased genetic variation and potentially new adaptations within a population.

Figure 20. Liger (Panthera leo X *Panthera tigris)*

THINK/PAIR/SHARE

Discuss the potential ecological and evolutionary impacts of increasing hybridization between species, such as polar bears and grizzly bears, due to climate change. What are the implications for conservation strategies, and how should conservationists balance the preservation of distinct species with the potential benefits of hybrid vigor and new genetic variations?

Media Attributions

- outcome of hybridization © Annarunemark is licensed under a CC BY-SA (Attribution ShareAlike) license
- liger © Ali West is licensed under a CC BY (Attribution) license

Recommended Reading

The following are recommended readings – we will be discussing at least one of these articles in next week's group discussion.

1. Herbert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotrophical skipper butterfly A*straptes fulgerator. Proceedings of the National Academy of Sciences* 1010: 14812 – 14817. https://doi.org/10.1073/pnas.0406166101
2. Struck TH, Feder JL, Bendiksby M, Birkeland S, Cerca J, Gusarov VI, Kistenich S, Larsson K-H, Liow LH, Nowak MD, Stedje B, Bachmann L, Dimitrov D (2017) Finding evolutionary processes hidden in cryptic species. *Trends in Ecology and Evolution* 33: 153 – 163. https://doi.org/10.1016/j.tree.2017.11.007
3. Kollar J, Poulickova A, Dvorak P (2021) On the relativity of species, or the probabilistic solution to the species problem. *Molecular Ecology* 31: 411 – 418. https://doi.org/10.1111/mec.16218
4. Collins RA, Cruickshank RH (2013) The seven deadly sins of DNA barcoding. *Molecular ecology resources* 13: 969 – 975. https://doi.org/10.1111/1755-0998.12046
5. Phillips JD, Gillis DJ, Hanner RH (2022) Lack of statistical rigor in DNA barcoding likely invalidates the presence of a true species' barcode gap. *Frontiers in Ecology and Evolution* 10: 859099. https://doi.org/10.3389/fevo.2022.859099
6. Beng KC, Corlett RT (2020) Applications of environmental DNA (eDNA in ecology and conservation: opportunities, challenges and prospects. *Biodiversity and conservation* 29: 2089 – 2121. https://doi.org/10.1007/s10531-020-01980-0
7. Fitzhugh K (2006) DNA barcoding: an instance of technology-driven science? *BioScience* 56: 462 – 463. https://doi.org/10.1641/0006-3568(2006)56[462:DBAIOT]2.0.CO;2

PART V

POPULATION GENETICS

LEARNING OBJECTIVES

At the end of this chapter, you should be able to:

1. Distinguish between populations, subpopulations and metapopulations.
2. Explain the importance of connectivity within a metapopulation.
3. Explain the importance of the Hardy Weinberg Equilibrium (HWE) as it pertains to population dynamics.
4. Explain the relationship between population size (N) and genetic diversity.
5. Distinguish between the three commonly used estimates o f genetic diversity.
6. Calculate F-statistics and Analysis of Molecular Variance (AMOVA).
7. Estimate gene-flow and migration rates from F-statistics.

Population genetics is the field of biology that investigates the genetic composition of populations and how it changes over time. It blends the principles of genetics with the concepts of evolutionary biology to explain *patterns of genetic variation*, *adaptation*, *and speciation*. By analyzing the genetic makeup of populations and how it is shaped by factors like **mutation**, **selection**, **genetic drift**, and **gene flow**, **population genetics** provides insights into the evolutionary processes that mold the diversity of life.

A key concept in population genetics is the **gene pool,** which refers to the complete set of genetic material within a breeding population. Genetic diversity within this pool is crucial for the survival and adaptability of a species. The field studies how genetic variations arise and are maintained within populations and how these variations lead to differences in traits. This, in turn, influences an individual's fitness, or its ability to survive and reproduce in its environment.

Population genetics is grounded in several important principles:

1. **Hardy-Weinberg Equilibrium**: This fundamental principle serves as a null hypothesis that predicts the genetic structure of a population will remain constant across generations in the absence of evolutionary forces. If a population is in Hardy-Weinberg equilibrium, allele and genotype frequencies will remain stable, provided that there is no mutation, selection, migration, genetic drift, or non-random mating. In reality, deviations from this equilibrium indicate the presence of these forces and are thus used to identify evolutionary changes.
2. **Genetic Drift**: This refers to random fluctuations in allele frequencies from one generation to the next due to chance events. Smaller populations are more susceptible to genetic drift, which can lead to a reduction in genetic diversity and even cause certain alleles to become fixed (reach 100% frequency) or lost entirely.
3. **Gene Flow**: The movement of alleles between populations, or gene flow, occurs through migration and interbreeding. Gene flow can counteract the effects of genetic drift by introducing new genetic material and can prevent populations from diverging genetically.
4. **Natural Selection**: Natural selection is a central force driving changes in allele frequencies. It operates when certain alleles

confer a selective advantage or disadvantage in a particular environment. Over time, alleles that increase an organism's fitness become more common in the population.

5. **Mutation**: Mutations are changes in DNA sequences that create new alleles and are the primary source of genetic variation. Although most mutations are neutral or harmful, some may provide adaptive advantages under specific environmental conditions.

The interplay of these forces determines how genetic diversity is structured within populations, influencing phenomena like adaptation to new environments and the formation of new species. For instance, genetic drift in small, isolated populations can lead to the rapid fixation of certain alleles, a phenomenon known as the founder effect. Similarly, strong selective pressures in changing environments may lead to the rapid spread of beneficial alleles, resulting in local adaptations

The Metapopulation Concept

A **metapopulation** is a group of spatially separated populations of the same species that interact at varying levels. This concept, rooted in ecology and evolutionary biology, provides a framework for understanding the dynamics of species that inhabit fragmented landscapes. Each individual population, or subpopulation, within a metapopulation occupies a distinct habitat patch, which can vary in size, quality, and connectivity. Although individual subpopulations may go extinct locally due to environmental changes or demographic factors, the overall metapopulation can persist over time through migration and recolonization of vacant habitat patches.

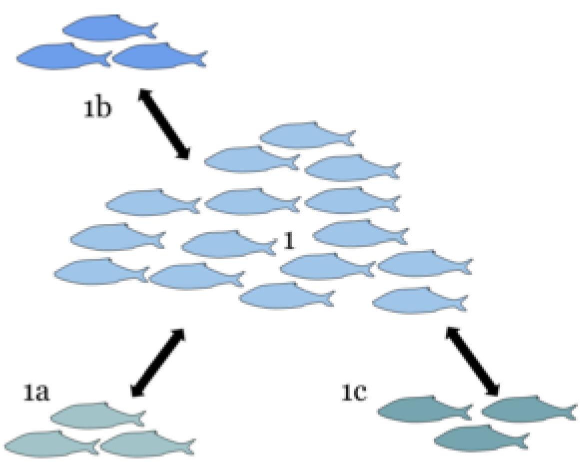


Figure 21. Illustration of a metapopulation where the local population (1.) serves as a source for interbreeding with surrounding subpopulations populations (1.a, 1.b, and 1.c) which may or may not be sinks. The populations are normally spatial separated and independent but spatial overlap during breeding times allow for genetic exchange (gene-flow) between the different populations.

Panmixia refers to a situation where all individuals within a population are equally likely to mate with each other, resulting in random mating. In a truly panmictic population, no barriers to gene flow exist, meaning genetic material is shared freely among all individuals. This results in homogenous allele frequencies throughout the population and no observable sub-structuring. In reality, panmixia is rarely achieved, as most natural populations experience some degree of non-random mating due to geographic barriers, behavioral differences, or social structures. However, in species that can disperse widely across continuous habitats (like certain migratory fish or birds), panmictic tendencies are more pronounced

Media Attributions

Revisiting the Hardy-Weinberg Equilibrium

The Hardy-Weinberg principle posits that if a population is not affected by mutation, selection, migration, or genetic drift, and if random mating occurs, the frequencies of alleles and genotypes will remain in equilibrium across generations. In this idealized state, the frequencies of alleles 'A' (dominant) and 'a' (recessive) are represented by 'p' and 'q', respectively. The principle provides a mathematical formula to predict genotype frequencies:

$$\mathbf{p^2 + 2pq + q^2 = 1}$$

where:

- **p^2** represents the proportion of individuals with the homozygous dominant genotype (AA),
- **2pq** represents the proportion of individuals with the heterozygous genotype (Aa),
- **q^2** represents the proportion of individuals with the homozygous recessive genotype (aa).

This model assumes that the sum of all allele frequencies equals 1 (i.e., p+q=1).

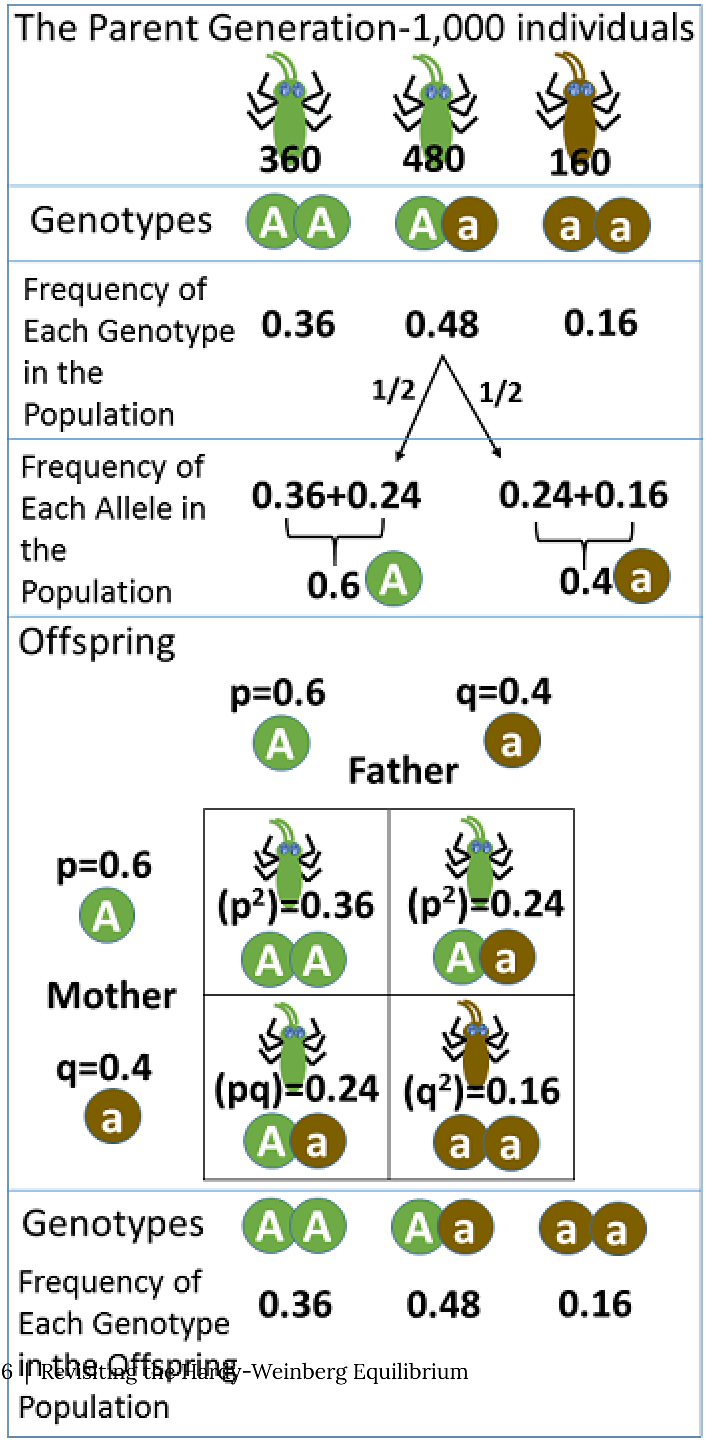

The Parent Generation-1,000 individuals
360
480
160
Genotypes
A A
A a
a a
Frequency of Each Genotype in the Population
0.36
0.48
0.16
1/2
1/2
Frequency of Each Allele in the Population
0.36+0.24
0.24+0.16
0.6 A
0.4 a
Offspring
p=0.6
A
q=0.4
a
Father
p=0.6
A
Mother
q=0.4
a
(p²)=0.36
A A
(p²)=0.24
A a
(pq)=0.24
A a
(q²)=0.16
a a
Genotypes
A A
A a
a a
Frequency of Each Genotype in the Offspring Population
0.36
0.48
0.16

Figure 22. Hardy-Weinberg equilibrium case example.

In molecular ecology and evolutionary biology, the Hardy-Weinberg principle provides a theoretical baseline against which real-world genetic data can be compared. By analyzing deviations from equilibrium, researchers can infer the presence of evolutionary forces or population-level changes:

1. **Natural Selection**: If certain genotypes are found at higher or lower frequencies than expected, this can indicate that natural selection is favoring or disfavoring specific alleles. For instance, in a changing environment, certain traits may confer a selective advantage, leading to changes in allele frequencies.
2. **Genetic Drift**: Small, isolated populations are particularly prone to random fluctuations in allele frequencies, known as genetic drift. Comparing genetic data to the Hardy-Weinberg equilibrium helps identify the impact of drift, which can lead to reduced genetic diversity or fixation of alleles.
3. **Gene Flow**: Migration between populations can introduce new alleles and homogenize genetic differences. Significant deviations from expected allele frequencies can reveal historical or ongoing migration events that impact local genetic structure.
4. **Non-Random Mating**: Mate choice preferences, inbreeding, or assortative mating can alter genotype frequencies, often increasing the proportion of homozygous individuals. These patterns can be detected through a Hardy-Weinberg Equilibrium analysis.
5. **Mutation**: Although mutations are relatively rare, they provide the raw genetic material for evolution. If new mutations confer an adaptive advantage, they may increase in frequency over time, creating a noticeable deviation from Hardy-Weinberg

expectations.

6. **Population Bottlenecks and Founder Effects**: Events like population bottlenecks or founder effects can drastically alter genetic variation. Comparing the genetic composition of populations before and after such events reveals the long-term impacts on allele frequencies.

Media Attributions

- Hartsock_Hardy_Weinberg_Example © Angela Hartsock is licensed under a CC0 (Creative Commons Zero) license

Defining Genetic Diversity

Genetic diversity represents the variation in genetic makeup within a population, species, or entire ecosystem. It is a cornerstone of evolutionary biology and ecology, reflecting the adaptability and resilience of populations to changing environments. High genetic diversity often indicates a robust ability to withstand diseases, environmental changes, and other threats. In contrast, low genetic diversity can make populations more susceptible to these challenges, potentially leading to inbreeding depression or even extinction.

- **Nucleotide Diversity (π)**: This measure assesses genetic variation at the nucleotide level, quantifying the average differences in nucleotides between pairs of sequences in a population. It is expressed as the proportion of nucleotide sites where two randomly chosen DNA sequences differ. Higher nucleotide diversity suggests a more diverse population at the molecular level, reflecting historical mutation rates, selection pressures, or demographic changes. For instance, populations that have experienced recent bottlenecks may have lower nucleotide diversity due to a reduction in genetic variation.
- **Haplotype Diversity (*h*)**: Haplotype diversity refers to the measure of variability of different haplotypes (combinations of alleles at multiple loci that are inherited together) within a population. It indicates the proportion of unique haplotypes and their distribution. A population with high haplotype diversity has many distinct combinations of alleles, signifying a broad genetic base. This measure is particularly useful when studying non-recombining DNA regions, such as mitochondrial DNA in animals or chloroplast DNA in plants. For example, populations of certain fish species may show low haplotype diversity due to historical isolation, while others display high diversity because of frequent gene flow.

- **Expected Heterozygosity (H_e)**: Expected heterozygosity, also known as gene diversity, estimates the likelihood that a randomly chosen individual from a population will be heterozygous at a specific locus. It is derived from allele frequencies, providing an expectation of the proportion of heterozygous genotypes if the population is in Hardy-Weinberg equilibrium. High expected heterozygosity suggests a population with a rich allelic diversity, meaning many different alleles are available at that locus. In contrast, low expected heterozygosity indicates fewer alleles, which could result from inbreeding or population bottlenecks.

Together, these measures of genetic diversity help researchers understand the evolutionary processes shaping populations and guide conservation efforts. By distinguishing between nucleotide diversity, haplotype diversity, and expected heterozygosity, scientists can gain nuanced insights into the genetic structure and health of populations. For example, high nucleotide diversity combined with low haplotype diversity might reflect recent population expansion from a few ancestral lineages. Monitoring genetic diversity is essential for predicting a population's potential adaptability and designing effective management strategies to preserve biodiversity.

Relationship Between Genetic Diversity and Population Size (N)

The census population size (often symbolized as N_c) and effective population size (N_e) are both measures used to describe the size of a population, but they differ in their definitions and implications, particularly in relation to genetic diversity.

- **Census Population Size (N_c):** This is the actual count of individuals within a population, essentially a headcount of all members. It is the total number of organisms that make up a population at a given time. While this is a straightforward metric, it does not account for the proportion of individuals contributing to reproduction or the reproductive potential of individuals. For instance, in populations where only a subset of individuals breed, the census size will overestimate the genetic contribution to the next generation.

- **Effective Population Size (N_e):** The effective population size represents the number of individuals in a hypothetical, idealized population that would exhibit the same amount of genetic drift or inbreeding as the observed population. An idealized population assumes random mating and equal reproductive success among all individuals. N_e is often smaller than the census size due to several factors affecting genetic contribution to future generations, such as skewed sex ratios, variance in reproductive success, overlapping generations, and population fluctuations.

Genetic diversity is influenced heavily by the effective population size because N_e determines the rate at which genetic variation is lost or maintained through processes like ge**netic drift** and **inbreeding**. In small populations with a low effective population size (N_e), **genetic drift** plays a significant role in reducing genetic diversity. This reduction occurs because alleles can become fixed or lost simply by chance. Conversely, a larger N_e helps retain allelic diversity over time, preserving the genetic variation within the population. **Inbreeding** is another consequence of a small effective population size. Increased inbreeding reduces heterozygosity and increases the likelihood of homozygous recessive traits, potentially leading to inbreeding depression, where the fitness of the population decreases due to the expression of deleterious recessive alleles**. Natural selection** operates differently depending on the effective population size. In smaller populations, beneficial mutations may become fixed more slowly, and deleterious alleles might persist due to the influence of genetic drift. This dynamic contrasts with larger populations, where natural selection can more effectively weed out harmful alleles and promote beneficial ones. Effective population size also affects a population's ability to incorporate new genetic material through **gene flow**. Small populations may quickly lose or fix new alleles introduced via migration, limiting their capacity to adapt to changing environments or recover from **genetic bottlenecks.** Events like population bottlenecks or founder effects can sharply reduce N_e, leading to a dramatic loss of genetic diversity even if the census population size (N_c) recovers afterward. These events can have long-lasting impacts on the genetic health and evolutionary potential of the population, emphasizing the importance of maintaining a large and stable Ne to support genetic diversity and resilience.

DEFINITION BOX

- **Genetic Drift**: This is a random change in the frequency of alleles (different versions of a gene) in a population over time. Unlike natural selection, genetic drift doesn't necessarily favor any particular allele; instead, these changes can happen just by chance. For example, if a few individuals happen to reproduce more than others, their genes become more common in the next generation by chance, not because these genes are "better".
- **Inbreeding**: Inbreeding occurs when closely related individuals breed with each other. This leads to an increase in the frequency of offspring inheriting the

same alleles from both parents, which can increase the chance of offspring being affected by recessive or deleterious traits. Essentially, it reduces genetic diversity, which can make populations more vulnerable to diseases and other problems.

- **Natural Selection**: This is a process where individuals with certain traits tend to survive and reproduce more than others in their environment. Over time, these advantageous traits become more common in the population. For example, if a bird species is better camouflaged in its environment, it might avoid predators more effectively and live longer to reproduce, passing on its camouflaging traits to its offspring.
- **Gene Flow**: Gene flow is the transfer of genetic material (alleles) between populations of the same species, often due to individuals moving from one place to another and breeding. This process can introduce new genetic material into a population, helping to maintain or increase genetic diversity, which can help populations adapt to changing environments.
- **Genetic Bottleneck**: A genetic bottleneck occurs when a population's size is significantly reduced for at least one generation. This can be due to events like natural disasters, diseases, or habitat loss. A smaller population size means less genetic diversity, and the genes present in the surviving population may not represent the genetic makeup of the original population. This can affect the population's ability to adapt to new environmental challenges and can increase the likelihood of inbreeding.

Genetic Diversity and Bottlenecks

A **genetic bottleneck** is a sharp reduction in the size of a population, leading to a significant loss of genetic diversity. This event usually results from environmental catastrophes, such as natural disasters, habitat destruction, or disease outbreaks, which suddenly reduce the number of individuals in a population. The concept is crucial in understanding how populations recover and adapt following such drastic reductions.

Mechanism of Genetic Bottlenecks:
During a bottleneck event, many individuals perish, leaving behind a small, non-representative sample of the original population. This drastic reduction means that the surviving individuals' genetic diversity is greatly reduced compared to the original gene pool. Consequently, certain alleles may become completely absent, while others may increase in frequency by chance alone.

Impacts on Genetic Diversity:

- **Loss of Allelic Diversity:** The smaller population that survives a bottleneck will likely have fewer unique alleles than the original, more diverse population. This loss can result in reduced adaptability to future environmental changes.
- **Increased Genetic Drift:** With fewer individuals, the impact of random genetic drift becomes more pronounced. Allele frequencies can fluctuate widely, leading to the fixation or loss of alleles purely by chance rather than through natural selection.
- **Inbreeding:** The reduced population size often forces related individuals to breed, resulting in increased homozygosity and potentially revealing deleterious alleles. Inbreeding depression can further threaten the population's viability.

Long Term Evolutionary Effects

Even if the population size rebounds after the bottleneck event, the genetic diversity lost might not be fully recovered. The surviving population now starts from a smaller gene pool, with fewer alleles to facilitate adaptation to new challenges. This reduction in diversity can make the population more vulnerable to diseases or environmental changes. For example, the cheetah (A*cinonyx jubatus*) is believed to have undergone a genetic bottleneck, leading to low genetic variability and increased susceptibility to disease. Understanding genetic bottlenecks is critical for fields such as conservation biology. Species that have undergone recent population declines may harbor reduced genetic diversity, affecting their ability to survive future environmental changes. Conservation efforts, such as maintaining genetic diversity through captive breeding or managing fragmented populations to enhance gene flow, can help mitigate the long-term effects of bottlenecks.

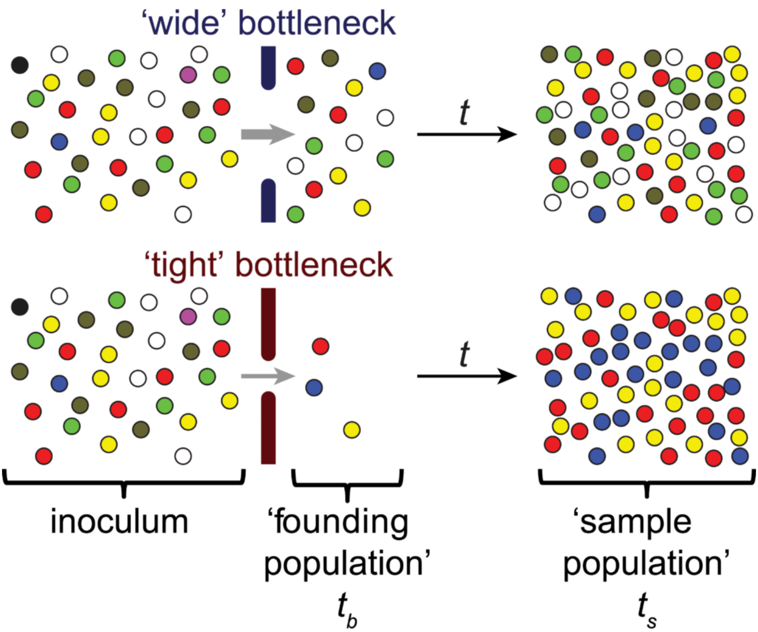


Figure 23. Schematic representation of the effects of bottlenecks on genetic diversity. Wide bottlenecks lead to limited loss of alleles (e.g., the magenta and black spheres) and limited changes in the allelic frequencies (e.g., over-representation of the blue and under-representation of the olive allele). In contrast, tight bottlenecks lead to stochastic loss of alleles and substantial changes in allelic frequencies.

THINK/PAIR/SHARE

SCENARIO: The "Azure Island foxes," endemic to Azure Island, have historically thrived due to diverse microhabitats and minimal human interference. However, a recent volcanic eruption devastated a significant portion of their habitat, resulting in the loss of 75% of their population. The surviving foxes now face limited prey availability and a reduced habitat area, compounding the challenges posed by their significantly diminished genetic diversity.

GROUP DISCUSSION QUESTION: Given the recent genetic bottleneck experienced by the Azure Island foxes due to a volcanic eruption, analyze the potential evolutionary trajectories this population might follow over the next few generations. Consider the impact of genetic drift, inbreeding, and new mutations. Additionally, discuss how environmental pressures such as limited food resources and reduced habitat might influence these genetic factors.

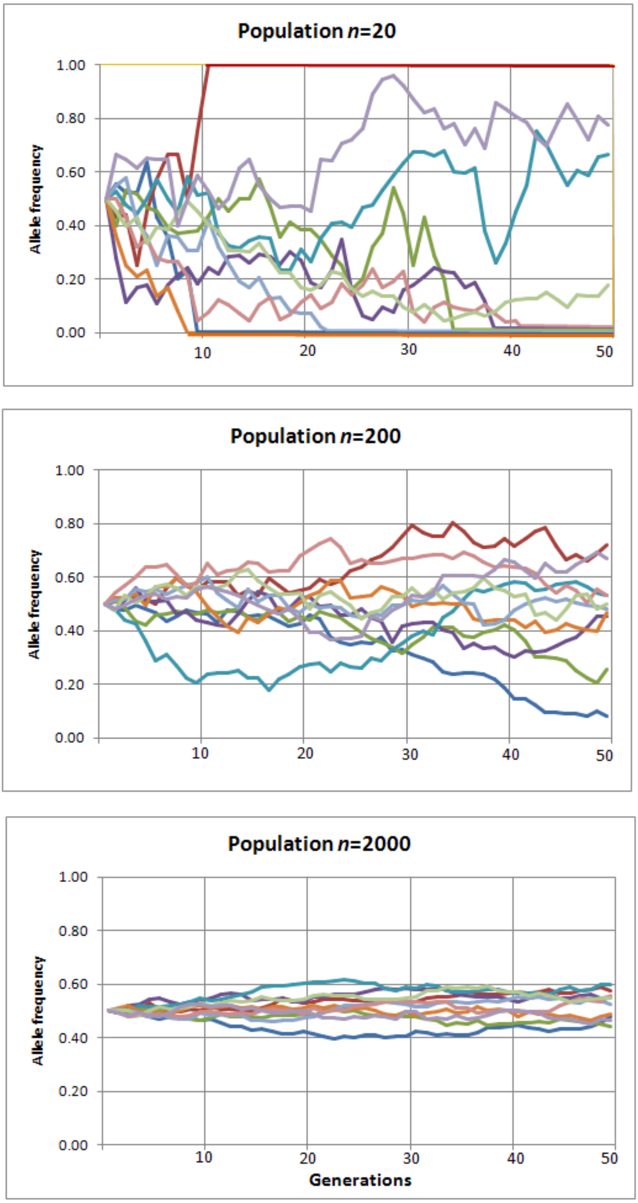
Media Attributions

Genetic Drift and Diversity

Genetic drift is an evolutionary process that results in random changes in allele frequencies within a population from one generation to the next. Unlike natural selection, which consistently favors traits that increase an organism's fitness, genetic drift operates without regard to the adaptive value of alleles. Its effects are most pronounced in small populations, where chance events can significantly impact genetic diversity. Genetic drift occurs because each generation represents only a subset of the genetic variation found in the parental generation. Even if all alleles have an equal chance of being inherited, random fluctuations in which alleles are passed down can lead to significant changes over time. This phenomenon can result in some alleles becoming more common while others decrease in frequency or are lost entirely. Genetic drift can affect genetic diversity in several ways:

1. **Allele Fixation or Loss:** In small populations, alleles can rapidly drift to fixation (meaning they reach a frequency of 100%) or be lost entirely. The rate of fixation is inversely related to population size: the smaller the population, the quicker the fixation or loss.
2. **Reduction in Genetic Variation:** As alleles are lost, the genetic variation within the population decreases. This reduction in variability limits the population's potential to adapt to environmental changes.
3. **Increased Homozygosity:** Genetic drift increases the proportion of homozygous individuals (those with two identical alleles for a particular gene). This can reveal deleterious alleles that may have been masked by heterozygosity.

While genetic drift and natural selection are distinct processes, they can interact. In small populations where drift predominates, even beneficial alleles can be lost by chance. Conversely, in larger populations where selection has more influence, genetic drift plays a lesser role.



Population n=20
Allele frequency
1.00
0.80
0.60
0.40
0.20
0.00
10
20
30
40
50
Population n=200
Allele frequency
1.00
0.80
0.60
0.40
0.20
0.00
10
20
30
40
50
Population n=2000
Allele frequency
1.00
0.80
0.60
0.40
0.20
0.00
10
20
30
40
50
Generations

Figure 24. Effect of population size on genetic drift: Ten simulations each of random change in the frequency distribution of a single hypothetical allele over 50 generations for different sized populations; first population size n=20, second population n=200.

Founder Effect

The founder effect is a phenomenon that occurs when a small group of individuals separates from a larger population to establish a new, isolated population. This subset typically carries only a fraction of the genetic diversity found in the original group. As a result, the new population may exhibit allele frequencies that are significantly different from those of the source population, leading to a rapid and sometimes substantial shift in genetic composition. When a few individuals colonize a new habitat or region, their genetic makeup defines the genetic starting point of the new population. Since only a limited number of alleles are represented in this founding group, certain alleles may become more common simply due to chance, while others may be entirely absent. Over time, genetic drift in this small population can cause these changes to be further exaggerated.

Genetic consequences of the founder effect:

1. **Reduced Genetic Variation:** The new population starts with reduced genetic diversity compared to the original population. The small initial pool of alleles limits the genetic variation that can be passed to subsequent generations.
2. **Altered Allele Frequencies:** The alleles present in the founders may become disproportionately represented in the new population, resulting in frequencies that differ markedly from the source population. This alteration could lead to a prevalence of traits not common in the original group.
3. **Inbreeding:** The limited genetic diversity can increase the likelihood of mating between closely related individuals,

leading to higher levels of homozygosity and a potential increase in deleterious alleles.

Island populations often show strong founder effects. For example, the genetic makeup of Darwin's finches on the Galapagos Islands differs significantly from that of mainland finches due to the isolation of each island population. In addition, certain human populations, such as the Amish or the inhabitants of Tristan da Cunha, exhibit genetic traits not commonly found elsewhere due to their descent from a small number of founders. The founder effect can lead to rapid speciation and adaptive radiation in isolated environments, as seen in island ecosystems. However, the reduced genetic diversity may also increase susceptibility to diseases and limit adaptability to environmental changes.

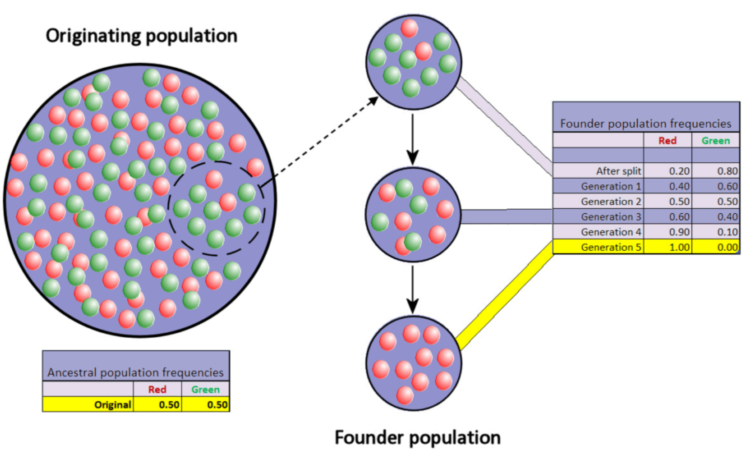


Figure 25. Representation of the founder effect: the colored balls represent the two alleles for a specific locus which are present in a hypothetical population; once a random subgroup of a population becomes separated from its ancestral population, the allele frequencies in the two groups' subsequent generation can diverge widely within a relatively short period of time as a consequence of random selection of alleles for reproduction.

Media Attributions

- Random_genetic_drift_chart © Professor marginalia is licensed under a CC BY-SA (Attribution ShareAlike) license
- Founder_effect_with_drift © Professor marginalia is licensed under a CC BY-SA (Attribution ShareAlike) license

A Brief Overview of Analysis of Molecular Variance

Analysis of Molecular Variance (AMOVA) is a statistical method used extensively in molecular ecology to quantify genetic variation within and among populations. Developed by Laurent Excoffier in the early 1990s, AMOVA allows researchers to partition genetic diversity at multiple hierarchical levels, such as within individuals, among individuals within populations, and among populations themselves. *It is important to have a rudimentary understanding of this analysis as we will be using it in lab.* AMOVA is based on the premise that genetic variation can be dissected into components corresponding to different levels of biological organization. By comparing genetic sequences or allele frequencies across these levels, AMOVA calculates the proportion of total genetic variation attributable to each level. This partitioning is crucial for understanding the distribution of genetic diversity and the processes driving population structure.

Methodology

1. **Data Input:** AMOVA typically uses molecular data, such as DNA sequences, microsatellite genotypes, or single nucleotide polymorphisms (SNPs). The input data are arranged according to the levels of hierarchical structure being studied.
2. **Distance Calculations:** AMOVA computes genetic distances or dissimilarities between all pairs of individuals or haplotypes, often using measures like F-statistics, which are based on Wright's F-statistics concept.
3. **Variance Partitioning:** The total genetic variance is partitioned

into components. For a simple two-level hierarchy (within and among populations), AMOVA calculates FST, analogous to Wright's FST, to measure genetic differentiation among populations. For more complex structures, AMOVA can compute FSC (among groups within populations) and FCT (among groups).

4. **Statistical Significance:** To assess the significance of the variance components, AMOVA typically uses permutation tests, randomly reallocating individuals or haplotypes to different groups and recalculating the variance components to create a distribution against which observed values can be tested.

AMOVA is instrumental in various research scenarios in molecular ecology:

- **Population Structure Analysis:** By quantifying the proportion of genetic variance at different hierarchical levels, AMOVA helps in detecting and describing the genetic structure of populations and species. It can reveal whether genetic differentiation is primarily within populations, among populations, or among groups of populations.
- **Conservation Genetics:** AMOVA informs conservation strategies by identifying the main sources of genetic diversity. If most genetic variation is within populations, strategies might focus on protecting many small populations; if among populations, preserving genetic exchange and connectivity becomes crucial.
- **Phylogeography and Historical Demography:** AMOVA can elucidate historical processes, such as migration, isolation, or expansion, by showing how genetic diversity is partitioned across landscapes or among lineages over time.

Wright's Fixation Index (F_{ST})

Wright's Fixation Index, F_{ST} , compares the genetic diversity observed within individual subpopulations to the total genetic diversity observed across all subpopulations. Essentially, it measures the proportion of genetic variance that can be attributed to differences among subpopulations. The value of F_{ST} ranges from 0 to 1, where **0** indicates no differentiation, meaning the subpopulations are genetically indistinguishable and share the same allele frequencies and **1** indicates complete differentiation, where the subpopulations are completely distinct, with no shared alleles. To calculate F_{ST}, researchers typically use allele frequency data from multiple loci across the subpopulations. The formula involves computing the average heterozygosity within subpopulations (often referred to as H_S) and the total heterozygosity across the entire population (H_T). The equation is:

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

EXAMPLE: Calculate genetic differentiation (F_{ST}) between the following two population:

Population 1 (Pop1):

Frequency of allele A (p_1): 0.7
Frequency of allele a (q_1): 0.3

Population 2 (Pop2):

Frequency of allele A (p_2): 0.4
Frequency of allele a (q_2): 0.6

Step-by-Step Calculation of F_{ST}

1. **Calculate the Average Allele Frequencies across Populations:** For allele A:

$$p_{avg} = \frac{p_1 + p_2}{2} = \frac{0.7 + 0.4}{2} = 0.55$$

For allele a:

$$q_{avg} = \frac{q_1 + q_2}{2} = \frac{0.3 + 0.6}{2} = 0.45$$

2. **Calculate the Expected Heterozygosity within Each Population:** This is $2 \times p_i \times q_i$ for each population. For **Pop1**:

$$H_{S1} = 2 \times p_1 \times q_1 = 2 \times 0.7 \times 0.3 = 0.42$$

For **Pop2**:

$$H_{S2} = 2 \times p_2 \times q_2 = 2 \times 0.4 \times 0.6 = 0.48$$

3. **Calculate the Average Expected Heterozygosity within Populations (H_S):**

$$H_S = \frac{H_{S1} + H_{S2}}{2} = \frac{0.42 + 0.48}{2} = 0.45$$

4. **Calculate the Total Expected Heterozygosity if the Entire Population were Mixed (H_T):**

$$H_T = 2 \times p_{avg} \times q_{avg} = 2 \times 0.55 \times 0.45 = 0.495$$

5. **Calculate F_{ST}:**

$$F_{ST} = \frac{H_T - H_S}{H_T} = \frac{0.495 - 0.45}{0.495} \approx 0.09$$

An F_{ST} value of 0.09 suggests that there is a moderate level of genetic differentiation between these two populations.

Media Attributions

- FST
- solution

Gene Flow and Migration

Gene flow refers to the movement of genetic material within or between populations. It's the process by which alleles (different versions of a gene) are shared, increasing genetic diversity. The result of gene flow is the alteration of allele frequencies in the recipient population. It can occur due to the movement of individuals carrying new alleles or through the dispersal of reproductive material like seeds or pollen. Gene flow can prevent populations from diverging too much genetically, helping maintain a species' cohesion. **Migration** is the movement of individuals from one place to another. In a biological context, it often describes the seasonal movement of animals to and from breeding or feeding areas. When individuals migrate to a new population and breed, their genes become incorporated into the new gene pool, leading to gene flow. However, not all migration results in gene flow if, for instance, the newcomers don't breed or don't survive long enough to reproduce. Migration rate (N_m) can be measured using the following equation:

$$N_m = \frac{1}{4}\left(\frac{1}{F_{st}} - 1\right)$$

GROUP REFLECTION QUESTIONS (*Choose one!*)

1. Choose a small, isolated population (e.g., an island species, or a conservation-reliant species in a zoo) and think about how genetic drift might affect its genetic diversity over multiple generations. Discuss potential scenarios where genetic drift could either lead to the fixation of deleterious alleles or possibly aid in rapid adaptation to new environmental pressures.
2. Evaluate the role of gene flow in counteracting the effects of genetic drift and selection in a fragmented habitat (e.g., a species living in an urbanized area e.g. rats in New York City or a species with habitat divided by human activity like roads or agriculture for e.g. deer in rural Massachusetts). What are the potential barriers to gene flow in this scenario, and how might they affect the population's genetic structure? Propose practical interventions or habitat management strategies that could enhance gene flow and thus genetic diversity.

Media Attributions

- Migration rate

Recommended Reading

The following are recommended readings – we will be discussing at least one of these articles in next week's group discussion.

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology* 19: 3038 – 3051. https://doi.org/10.1111/j.1365-294X.2010.04688.x

Pante E, Puillandre N, Viricel A, Arnaud-Haond S, Aurelle D, Castelin M, Chenuil A, Destombe C, Forcioli D, Valero M, Viard F, Samadi S (2014) Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Molecular Ecology* 24: 525 – 544. https://doi.org/10.1111/mec.13048.

Ellegren H, Galtier N (2016) Determinants of genetic diversity. *Nature Reviews Genetics* 17: 422 – 433. https://doi.org/10.1038/nrg.2016.58

Meirmans PG (2007) Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* 60: 2399 – 2402. https://doi.org/10.1111/j.0014-3820.2006.tb01874.x

Morjan CL, Rieseberg LH (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology* 13: 1341 – 1356. https://doi.org/10.1111/j.1365-294X.2004.02164.x

Hellberg ME (2009) Gene flow and isolation among populations of marine animals. *Annual Review of Ecology, Evolution and Systematics* 40: 291 – 310. https://doi.org/10.1146/annurev.ecolsys.110308.120223

Wilkes MA, Webb JA, Pompeu PS, Silva LGM, Vowles AS, Baker CF, Franklin P, Link O, Habit E, Kemp PS (2019) Not just a migration problem: Metapopulations, habitat shifts, and gene flow are also important for fishway science and management. *River Research and Applications* 35: 1688 – 1696. https://doi.org/10.1002/rra.3320

Miles LS, Rivkin LR, Johnson MTJ, Munshi-South J, Verrelli BC (2019)

Gene flow and genetic drift in urban environments. *Molecular Ecology* 28: 4138 – 4151. https://doi.org/10.1111/mec.15221

PART VI

INTRODUCTION TO PHYLOGEOGRAPHY

Learning Objectives

At the end of this chapter, you should be able to:

1. Define phylogeography
2. Explain the use mitochondrial DNA in inferring phylogeographic patterns
3. Explain coalescent theory and its applications
4. Interpret haplotype networks
5. Explain what a phylogeographic break is
6. Distinguish between vicariance and dispersal

Phylogeography is a scientific discipline that merges principles of phylogenetics and population genetics to understand *the geographic distribution of genetic lineages across spatial and temporal scales*. It explores how historical events like glaciations, river formations, and mountain uplifts, combined with ecological factors, have shaped the genetic structure and distribution of species. By tracing the geographical and historical patterns of genetic variation, researchers can infer past demographic events such as population expansions, contractions, and migrations. These inferences rely on analyzing genetic markers, such as mitochondrial DNA or microsatellites, to reveal the evolutionary history of populations. For instance, in studying a species spread across different geographical regions,***phylogeography helps identify distinct genetic lineages that may correspond to historical barriers or periods of isolation***. This information enhances our understanding of evolutionary processes by linking genetic data with past climatic and geological changes, ultimately providing insight into biodiversity patterns and guiding conservation efforts.

A Brief History

Origins and Foundational Work (1980s-1990s):
The term "phylogeography" was formally introduced in 1987 by John Avise, in his landmark paper that laid the foundation for this emerging field. Avise's work highlighted the need to understand how historical and geographical factors shape genetic patterns. Using mitochondrial DNA (mtDNA) data from diverse vertebrates, he showed that certain geographic features, such as rivers or mountain ranges, can significantly influence the distribution of genetic lineages within and among species. Early studies focused on distinguishing phylogeographic patterns caused by historical factors, like glaciations, from those driven by contemporary ecological interactions.

Technological Advances and Expansion

(1990s-2000s):
The 1990s and early 2000s saw rapid advancements in molecular techniques, leading to a proliferation of genetic markers like nuclear microsatellites and DNA sequencing. This enabled scientists to uncover deeper insights into population history. Concurrently, coalescent theory emerged, providing a robust statistical framework for modeling genetic variation over time and space. This theoretical advancement allowed researchers to test explicit demographic hypotheses, making phylogeography increasingly predictive.

Global Expansion and Interdisciplinary Growth (2000s-2010s):
As genetic data from diverse taxa across the globe became available, phylogeography expanded to encompass both regional and global biogeographic patterns. Collaboration with paleoecology, geology, and climate science enriched phylogeographic analysis, leading to multidisciplinary studies that could infer how historical climatic

events influenced current biodiversity. The rise of statistical phylogeography made it possible to build more detailed models linking genetic variation with ancient demographic events like population bottlenecks, expansions, and migrations.

Genomics Era (2010s-present):
The advent of high-throughput sequencing technologies in the late 2000s revolutionized phylogeography. Whole-genome data could now be collected from hundreds of individuals, providing unprecedented resolution for detecting fine-scale genetic patterns. This has enabled researchers to resolve complex demographic histories, identify cryptic species, and assess the effects of recent environmental changes on genetic diversity. Additionally, the development of GIS (geographical information systems) has helped to overlay genetic data with environmental and geographical maps for comprehensive spatial analysis.

Current and Future Directions:
Today, phylogeography is increasingly focused on understanding the impacts of recent and ongoing changes, such as habitat fragmentation, invasive species, and climate change. Furthermore, the growing field of landscape genetics blends genetic data with landscape ecology to understand how physical features and human activities affect gene flow. Researchers are now also incorporating ecological niche modeling to predict how species distributions and genetic patterns may shift under future climate scenarios.

Case study: Grey Wolf phylogeography

Case Study: Phylogeography of Grey Wolves (*Canis lupus*)

The gray wolf (*Canis lupus*) provides a fascinating example of how phylogeography can illuminate the complex evolutionary history and demographic changes of a species. As a widely distributed carnivore with populations across North America, Europe, and Asia, the gray wolf has undergone significant fluctuations in range and abundance due to both natural and anthropogenic factors.

Figure 26. Pack of grey wolves (Canis lupus)

Historical Context

The gray wolf's global range has contracted significantly over the past few centuries due to habitat loss, hunting, and persecution. However, during the last glacial maximum (LGM), they were widespread across the Northern Hemisphere. The glaciers acted as a physical barrier, driving isolation and genetic differentiation between Eurasian and North American populations.

Genetic Analysis and Findings

Recent phylogeographic studies have used mitochondrial DNA (mtDNA) and nuclear microsatellites to reveal patterns of genetic diversity and divergence across different populations. Researchers found that:

- European and Asian wolf populations form a distinct lineage from North American wolves, indicating a long period of isolation.
- Within North America, wolves exhibit regional genetic structure due to historical ice barriers and local adaptation.
- Wolves in Alaska and western Canada have retained ancient genetic lineages from the Pleistocene epoch.
- Wolves in the western United States, recolonizing after extirpation in the 20th century, have lower genetic diversity and reflect a relatively recent genetic bottleneck.

Implications for Conservation

These findings have significant implications for conservation. Regional populations possess unique genetic characteristics that should be preserved. For instance, the European populations represent a distinct lineage with a limited distribution. In North America, efforts should focus on maintaining genetic diversity among remaining wolf populations while facilitating connectivity across fragmented habitats.

Media Attributions

- wolves © Ronnie Macdonald is licensed under a CC BY (Attribution) license

The Pleistocene Epoch

The Pleistocene Epoch, which lasted from about 2.6 million to 11,700 years ago, is crucial in phylogeography due to its profound impact on shaping the current distribution and genetic structure of species. Known as the "Ice Age," this epoch was marked by repeated glacial cycles that drove significant ecological changes, influencing the dispersal, isolation, and evolution of numerous species across the globe.

Figure 27. During the Pleistocene, mammoths were widespread across North America, Europe, Asia, and Africa. The most well-known species, the woolly mammoth (Mammuthus primigenius), lived during the later part of the Pleistocene and was well adapted to the cold environments of the time. Mammoths, along with many other megafaunal species, went extinct towards the end of the Pleistocene, around the time of the last glacial maximum and the subsequent warming period.

Glacial Cycles and Habitat Shifts

The Pleistocene was characterized by a series of glacial and interglacial periods where large ice sheets periodically expanded and retreated. These glaciations covered much of the Northern Hemisphere, particularly North America and Eurasia. Glacial advances forced species to migrate to refugia—areas that remained ice-free and habitable. These refugia acted as sanctuaries where species survived in isolated pockets. When glaciers receded during interglacial periods, species dispersed outward, recolonizing previously uninhabitable regions. These cycles created a pattern of repeated isolation and secondary contact that significantly influenced genetic differentiation and speciation.

Refugia and Genetic Differentiation

Refugia played a critical role in shaping phylogeographic patterns. Populations isolated in different refugia evolved independently, leading to genetic differentiation. When species recolonized after glacial retreats, some populations remained genetically distinct due to barriers like mountains or rivers. In other cases, secondary contact resulted in hybrid zones where different lineages interbred. The location of historical refugia and patterns of recolonization are often reflected in modern genetic variation, helping researchers reconstruct past distributions and migration routes.

Evolutionary Impacts and Adaptive Responses

The Pleistocene's dynamic climate and habitat changes also influenced natural selection and adaptation. Species adapted to survive in different climates, and some evolved specific traits or physiological mechanisms to endure colder temperatures. Genetic markers associated with these adaptations are sometimes identifiable in modern populations, providing clues about their evolutionary history. Additionally, populations that persisted through multiple glacial cycles often show greater genetic diversity due to their long-term persistence.

Extinctions and the Rise of Modern Biotas

The Pleistocene also witnessed significant extinctions, particularly during the Late Pleistocene, when many large mammals (megafauna) vanished, possibly due to climate change and human activities. The loss of these species dramatically altered ecosystems, influencing the distribution and genetic structure of surviving species. Consequently, the modern biotas reflect the legacy of these extinctions and the reorganization of ecological communities.

Applications in Conservation

Understanding the Pleistocene's impact on phylogeography has crucial implications for conservation. Recognizing regions with high genetic diversity or unique lineages can help identify areas with historical significance, which might harbor greater evolutionary potential. Additionally, studying Pleistocene refugia and migration corridors provides a framework for predicting how species might respond to future climate changes.

Media Attributions

- mammoth © Amy

Coalescent Theory

Coalescent theory is a fundamental concept in population genetics that traces the ancestral lineage of gene copies within a population back to a common ancestor. Developed in the 1980s by John Kingman, this mathematical framework models how genetic variation emerges and is shaped over generations. In phylogeography, it provides powerful insights into the historical and demographic processes that influence the genetic structure of populations. The basic premise of coalescent theory is that any two gene copies in a population can be traced back in time until they converge, or "coalesce," at a common ancestor. The further back we go, the more gene copies merge until reaching the most recent common ancestor for the entire population. The theory simplifies this complex process by providing probabilistic models that predict the distribution of coalescence times under different demographic scenarios, such as population growth, bottlenecks, or migrations.

In phylogeography, coalescent theory helps decipher how historical events have shaped the spatial distribution of genetic diversity. For instance, it allows scientists to estimate the time since populations diverged from one another or how recent demographic changes, like expansions or contractions, have affected genetic diversity. By analyzing genealogical relationships, the theory can infer past events like population bottlenecks, which occur when populations shrink significantly, leaving a genetic signature due to reduced diversity. Coalescent theory is crucial in phylogeography because it provides a framework to model gene flow, divergence, and historical population sizes based on genetic data. For example, researchers can assess whether populations remained isolated during historical climatic events like the Pleistocene glaciations or whether secondary contact led to admixture between previously distinct populations. The theory also accounts for the effects of natural selection, genetic drift, and geographic barriers on the genetic makeup of populations.

A significant strength of coalescent theory is its flexibility. It can handle various genetic markers, such as mitochondrial DNA, autosomal genes, or whole-genome data. Additionally, it works across different scales, from single populations to entire species ranges. Its probabilistic models incorporate the effects of mutations, recombination, and demographic changes, providing more accurate reconstructions of population histories than traditional genetic analyses. The importance of coalescent theory in phylogeography extends to practical applications, such as conservation genetics. It helps identify evolutionary significant units by revealing how genetically distinct populations are and how much gene flow occurs between them. This knowledge can guide conservation strategies by prioritizing populations with unique genetic legacies or identifying connectivity corridors that maintain genetic diversity.

Phylogeographic Barriers

Phylogeographic barriers (also known as phylogeographic "breaks") are *physical or environmental features that limit or prevent the movement of organisms across a landscape, seascape or riverscape, leading to genetic differentiation among populations*. These barriers can profoundly influence the distribution of genetic diversity, often resulting in distinctive lineages or haplotypes that reflect the historical separation of populations. Their study helps researchers understand how geography and history have shaped the genetic structure of species. One of the most significant types of phylogeographic barriers is a geographical feature that directly obstructs movement. Mountain ranges, large rivers, oceans, and deserts are classic examples. For instance, the Himalayan mountains present a formidable barrier between northern and southern Asia, isolating the flora and fauna on each side. Similarly, the Sahara Desert restricts gene flow between sub-Saharan Africa and the Mediterranean region.

Climatic barriers also play a crucial role in shaping phylogeographic patterns. During the Pleistocene glaciations, ice sheets expanded and contracted, forcing many species to retreat into ice-free refugia. This created isolated populations that evolved separately. For example, North America's glacial ice sheets separated eastern and western populations, and similar patterns were observed in Europe and Asia. The legacy of these isolated refugia is visible today, where populations exhibit distinct genetic lineages that trace back to these historical separations. In addition to physical and climatic barriers, ecological differences can create invisible yet significant boundaries. Some species are highly specialized to particular habitats, such as freshwater lakes, specific forest types, or certain altitudes. The transition between these ecological zones can limit dispersal and cause genetic differentiation. For instance, high-elevation species often show unique genetic structures due to the sharp ecological boundary between montane and lowland environments.

Human activities can also establish artificial barriers that fragment populations. Urbanization, road construction, and agricultural expansion increasingly divide natural habitats, disrupting connectivity. This can cause genetic drift or inbreeding within isolated populations. Conservation efforts often focus on maintaining or restoring habitat corridors to ensure gene flow and genetic health. The effects of phylogeographic barriers are evident in species that display regional patterns of genetic variation. Populations separated by barriers often have reduced gene flow between them, resulting in higher genetic differentiation. Sometimes, this differentiation is subtle but detectable through genetic markers. Other times, the barrier is so strong that it leads to the formation of new species, a process known as allopatric speciation.

Understanding phylogeographic barriers helps clarify the historical processes that have shaped modern biodiversity. It allows scientists to trace how ancient geographic and climatic shifts influenced the genetic structure of current populations. Moreover, this knowledge aids in conservation by identifying regions where unique genetic lineages persist and highlighting habitat corridors that maintain connectivity. As environmental changes continue, the role of phylogeographic barriers remains crucial in predicting how species will respond and adapt to future challenges.

Example: Strait of Gibraltar

The Strait of Gibraltar, the narrow body of water separating southern Europe from northern Africa, is an excellent example of a barrier where "gene flow leakage" occurs. Despite being a significant natural boundary between the Atlantic Ocean and the Mediterranean Sea, it allows limited but detectable genetic exchange between populations of various species. The strait is only about 14 km (9 miles) wide at its narrowest point, separating the southern tip of Spain and northern Morocco. This narrow distance seems small, yet it represents a significant geographic boundary between two distinct biogeographical regions—Europe and Africa. For terrestrial species, crossing this water body presents a challenge, while the strong currents make migration difficult for marine species. For terrestrial animals and plants, the strait acts as a strong barrier, limiting movement between Europe and Africa. However, some species exhibit gene flow leakage across this divide. Birds are one of the most notable examples, as migratory species regularly travel between the two continents, carrying genetic material with them. For instance, certain bird populations like the white stork (*Ciconia ciconia*) migrate seasonally between Europe and Africa, leading to gene flow between populations.

Figure 27. The Strait of Gibraltar is well known phylogeographic break that separates the eastern Atlantic from the Mediterranean Sea

The Strait of Gibraltar, the narrow body of water separating southern Europe from northern Africa, is an excellent example of a barrier where “gene flow leakage” occurs. Despite being a significant natural boundary between the Atlantic Ocean and the Mediterranean Sea, it allows limited but detectable genetic exchange between populations of various species. The strait is only about 14 km (9 miles) wide at its narrowest point, separating the southern tip of Spain and northern Morocco. This narrow distance seems small, yet it represents a significant geographic boundary between two distinct biogeographical regions—Europe and Africa.

For terrestrial species, crossing this water body presents a challenge, while the strong currents make migration difficult for marine species. For terrestrial animals and plants, the strait acts as a strong barrier, limiting movement between Europe and Africa. However, some species exhibit gene flow leakage across this divide. Birds are one of the most notable examples, as migratory species regularly travel between the two continents, carrying genetic material with them. For instance, certain bird populations like the white stork (*Ciconia ciconia*) migrate seasonally between Europe and Africa, leading to gene flow between populations.

In the marine environment, the strait represents a significant bottleneck (do not confuse this with a *genetic* bottleneck!) between the Atlantic Ocean and the Mediterranean Sea. The strong, complex currents make it difficult for many marine organisms to traverse, often resulting in genetic differentiation between populations on either side. However, gene flow leakage has been observed in several marine species. The European anchovy (*Engraulis encrasicolus*), for example, shows genetic differences between Atlantic and Mediterranean populations, but some mixing still occurs across the strait. Other examples include certain bivalves and small fish species that occasionally disperse through the strait's currents, maintaining genetic connectivity. Gene flow leakage across the Strait of Gibraltar helps scientists understand the complexity of migration barriers. Even a seemingly strong boundary can allow limited exchange, influencing genetic diversity and adaptation. This phenomenon is crucial for conservation, as identifying connectivity across barriers can inform strategies to maintain gene flow and preserve genetic diversity.

THINK/PAIR/SHARE

1. Reflect on the different types of phylogeographic barriers mentioned in the reading—geographical, climatic, ecological, and anthropogenic. Choose one type of barrier and discuss how it could impact the genetic structure of a species you are familiar with. How might this barrier influence the population's genetic diversity and potential for adaptation?
2. The reading provides the Strait of Gibraltar as an example of a phylogeographic barrier with "gene flow leakage." Choose another natural or artificial barrier and explore the concept of gene flow leakage in that context. How does limited gene flow across this barrier affect the genetic diversity and evolutionary potential of the populations on either side?

Media Attributions

-

Haplotype Networks

A **haplotype network** is a graphical representation that illustrates the relationships between different haplotypes—a group of closely linked genetic markers inherited together—within a population or species. It provides insights into the evolutionary history and genetic diversity of these groups. In phylogeography, haplotype networks are valuable because they help researchers trace patterns of genetic variation across geographic regions, offering clues about the historical processes that have shaped the distribution of species.

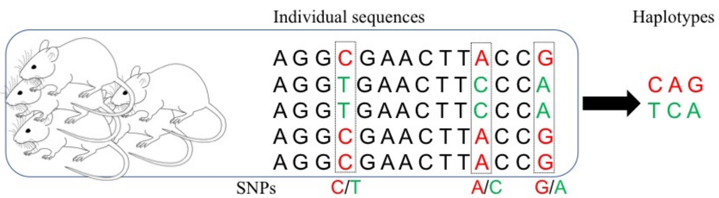


Figure 28. This figure shows the sequences of multiple individuals in a population, highlighting how single nucleotide polymorphisms (SNPs) define different haplotypes. Each row represents an individual sequence, with variations at specific positions indicated. The SNPs shown are C/T, A/C, and G/A, which distinguish the haplotypes. For example, the first and fourth sequences share the same haplotype, defined by the SNPs C at position 1, A at position 2, and G at position 3, while the second and third sequences represent a different haplotype with T, C, and A at the same positions. Mouse vector illustration by , via Wikimedia Commons" data-url="http://Gwilz, CC BY-SA 4.0 , via Wikimedia Commons">Gwilz (CC BY-SA 4.0)

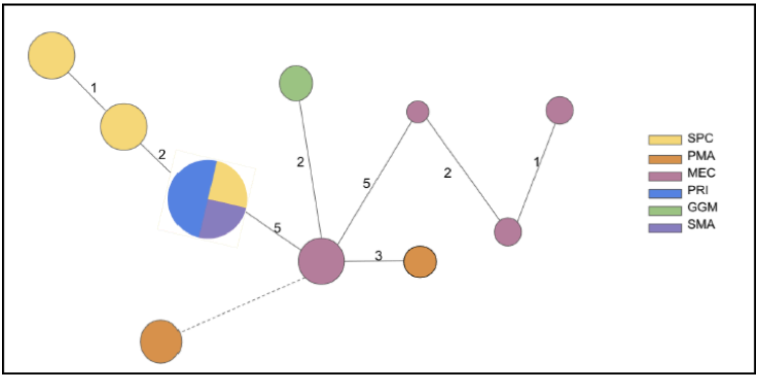


Figure 29. . Haplotype network of the blue mussel, Mytilus edulis, showing the distribution of genetic variation across the New England Coast. Each color represents a different population while each circle represents a specific haplotype. Lines represent a single mutational step, numbers above lines represent additional mutational steps and dashed lines represent mutation steps that exceed 60. SPC (Stonington Point, Connecticut), PMA (Plymouth, Massachusetts), MEC (Maine), PRI (Point Judith, Rhode Island), GGM (Maine – Site 2), SMA (Sandwich, Massachusetts). (Network constructed by Wheaton College Biology majors Megan Kelleher & Madison Faulkingham: Genetic connectivity of Mytilus edulis from the New England coast – Spring 2024 BIO 317L, Wheaton College.)

Media Attributions

Marker Selection in Phylogeography

Currently, there is a vast array of molecular markers which can be employed for phylogeographic research. When choosing a molecular marker, it is essential to consider the marker's mutation rate, inheritance pattern, and the kind of genetic information it provides. Some examples of commonly used (and older) markers are provided below.

Mitochondrial DNA (mtDNA)

Mitochondrial DNA (mtDNA) is a circular, maternally inherited DNA often utilized for studying maternal lineages. One of its primary advantages is the high mutation rate, which provides detailed resolution of genetic differences. The single-parent inheritance simplifies lineage tracing, and its abundance in cells makes it easy to extract and amplify. However, mtDNA only provides information on maternal lineage and its lack of recombination can obscure past hybridization events. Additionally, selective sweeps can reduce genetic variability within mtDNA.

Microsatellites

Microsatellites are short, repetitive DNA sequences dispersed throughout the genome, making them useful for detecting genetic diversity and structure. They exhibit high polymorphism, offering detailed genetic differentiation, and co-dominant markers, allowing the detection of heterozygotes. Microsatellites are effective for fine-scale population studies. Nevertheless, developing species-specific primers for microsatellites is relatively expensive and time-consuming. Their high mutation rates can also lead to homoplasy, where independent mutations result in identical patterns.

Single Nucleotide Polymorphisms (SNPs)

Single nucleotide polymorphisms (SNPs) are single base pair variations within the genome, occurring in both coding and non-coding regions. SNPs can be analyzed in high throughput using genotyping arrays or sequencing, providing genome-wide information. They are suitable for detecting recent genetic differentiation. However, SNPs require prior knowledge of the genome for effective identification and may offer less information per marker due to limited allelic variation.

Restriction Site-Associated DNA Sequencing (RAD-seq)

Restriction Site-Associated DNA Sequencing (RAD-seq) generates genome-wide SNP markers by sequencing regions near restriction enzyme cut sites. RAD-seq offers thousands of markers across the genome and is suitable for non-model organisms. It is effective for understanding fine-scale genetic differentiation. On the downside, RAD-seq requires significant bioinformatics processing, and the sequence coverage can vary between samples, complicating data analysis.

Allozymes

Allozymes are enzyme variants detected by electrophoresis, representing one of the earliest molecular markers used in population genetics. They are co-dominant markers that can detect heterozygosity and are relatively inexpensive and quick to analyze, making them useful for broad-level population studies. However, allozymes have limited resolution compared to modern markers, detect genetic variation only in protein-coding regions, and are less sensitive due to low mutation rates.

Whole-Genome Sequencing

Whole-genome sequencing involves the comprehensive sequencing of an organism's entire genome, providing the most detailed and exhaustive genetic data available. It can detect rare and previously unknown genetic variants, enabling genome-wide studies of adaptation and population history. Nonetheless, whole-genome sequencing requires advanced sequencing infrastructure, significant computational resources, and extensive data analysis and bioinformatics expertise. The high cost associated with whole-genome sequencing can also limit the sample size.

Recommended Reading

The following are recommended readings – we will be discussing at least one of these articles in next week's group discussion.

1. Avise JC (2009) Phylogeography: retrospect and prospect. *Journal of biogeography* 36: 3 – 15. https://doi.org/10.1111/j.1365-2699.2008.02032.x
2. Gauther MR, Toonen RJ, Bowen BW (2012) Coming out of the starting blocks: extended lag time rearranges genetic diversity in introduced marine fishes of Hawai'i. *Proceedings of the Royal Society B: Biological Sciences* 279: 3948 – 3957. https://doi.org/10.1098/rspb.2012.1481
3. Bowen BW, Gaither M, DiBattista JD, Briggs JC (2016) Comparative phylogeography of the ocean planet. *Proceedings of the National Academic of Sciences* 113: 7962 – 7969. https://doi.org/10.1073/pnas.1602404113
4. Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM (2013) Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular Ecology* 22: 2953 – 2970. https://doi.org/10.1111/mec.12228
5. Monteiro FA, Peretolchina T, *et al.* (2013) Phylogegoraphic pattern and extensive mitochondrial DNA divergence disclose a species complex within the Chagas disease vector, *Triatoma dimidiata.* PLoS ONE 8: e70974. https://doi.org/10.1371/journal.pone.0070974
6. Larmuseau MHD, Raeymaekers JAM, Hellemans B, Van Houdt JKJ, Volckaert (2010) Mito-nuclear discordance in the degree of population differentiation in a marine goby. *Heredity* 105: 532 – 542. https://doi.org/10.1038/hdy.2010.9
7. Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege

in avian phylogeography. *Molecular Ecology* 17: 2107 – 2121. https://doi.org/10.1111/j.1365-294X.2008.03737.x
8. Godinho R, Crespo EG, Ferrand N (2008) The limits of mtDNA phylogeography: complex patterns of population history in a highly structured Iberian lizard are only revealed by the use of nuclear markers. *Molecular Ecology* 17: 4670 – 4683. https://doi.org/10.1111/j.1365-294X.2008.03929.x
10. Nieberding CM, Durette-Desset M-C, Vanderpoorten A, Casanova JC, Ribas A, Deffontaine V, Feliu C, Morand S, Libois R, Michaux JR (2008) Geography and host biogeography matter for understanding the phylogeography of a parasite. *Molecular Phylogenetics and Evolution* 47: 538 – 554. https://doi.org/10.1016/j.ympev.2008.01.028
11. Godefroid M, Rasplus J-Y, Rossi J-P (2016) Is phylogeography helpful for invasive species risk assessment? The case study of the bark beetle genus *Dendroctonus. Ecography* 39: 1197-1209. https://doi.org/10.1111/ecog.01474

PART VII

CONSERVATION GENETICS

LEARNING OBJECTIVES

At the end of this chapter, you should be able to:

1. Explain how population genetics and phylogeography can be used to study and remediate the biodiversity crisis.
2. Explain the basis of inbreeding and be able to calculate the inbreeding co-efficient (*F*)
3. Explain the evolutionary adaptations some organisms have developed to avoid inbreeding in nature.
4. Explain the concept of “genetic rescue” for endangered species.
5. Discuss the various forms genetic rescues including their pros and cons.

The Biodiversity Crisis

The biodiversity crisis is a rapidly unfolding global phenomenon marked by unprecedented rates of species extinction. While extinction is a natural part of life's history, with species continuously emerging and disappearing over geological time scales, current extinction rates are alarmingly high. This pattern reflects profound changes in the way ecosystems function and directly threatens the ecological services that sustain human societies.

Extinction is a natural process driven by several factors. Throughout Earth's history, species have gone extinct due to environmental changes, competition, predation, disease, and stochastic events. Paleontological records show that over 99% of all species that ever existed are now extinct. This natural turnover has led to continuous evolutionary adaptations and the diversification of life. In geological history, mass extinctions—periods characterized by the loss of a significant proportion of species over relatively short periods—have punctuated natural extinction rates. Five such mass extinction events are recognized, including the end-Permian event (~252 million years ago), which wiped out over 90% of marine species, and the more recent Cretaceous-Paleogene event (~66 million years ago) that led to the extinction of the dinosaurs.

In stark contrast to the natural baseline, the current extinction rate is estimated to be 100 to 1,000 times higher than historical averages. This accelerated rate is largely attributed to human activities. Unlike past mass extinctions, which resulted from catastrophic natural events like volcanic eruptions or asteroid impacts, the present crisis is primarily driven by anthropogenic causes:

1. **Habitat Destruction and Fragmentation:** Urbanization, agriculture, logging, and infrastructure development continue to

destroy and fragment habitats, reducing biodiversity by displacing species and disrupting ecological connectivity.

Figure 30. Fragmented forest in Uganda due to anthropogenic disturbances.

2. **Overexploitation:** Unsustainable hunting, fishing, and logging deplete species populations, leading to localized extinctions.

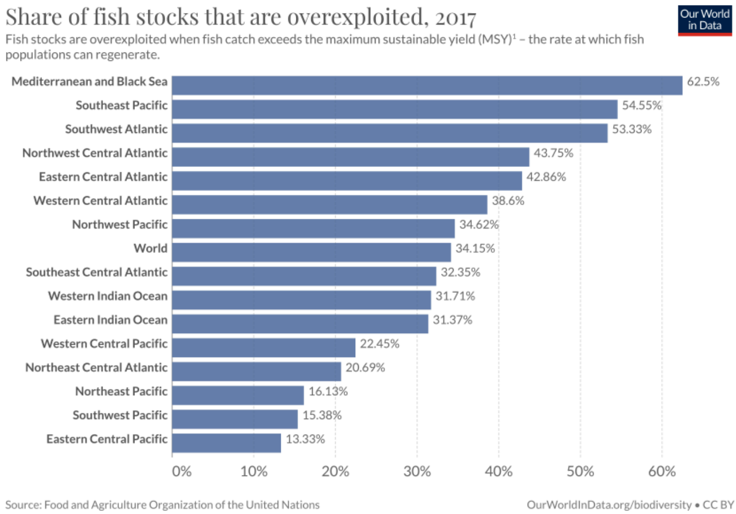


Figure 31. Overexploited fish stocks.

3. **Invasive Species:** Non-native species introduced by human activity often outcompete, predate, or spread disease to native species, causing their decline.

Figure 32. The invasive Rosemary beetle, Chrysolina americana.

4. **Pollution:** Industrial, agricultural, and urban pollution harms both terrestrial and marine life, while climate change exacerbates the effects of pollution.

Figure 33. Plastic pollution in Ghana. The International Union for the Conservation of Nature lists plastic pollution as "a major driver of biodiversity loss and ecosystem degradation" due to its impact on all of earth's biomes.

5. **Climate Change:** Rapid changes in temperature, precipitation patterns, and extreme weather events due to climate change challenge species' ability to adapt, causing population declines.

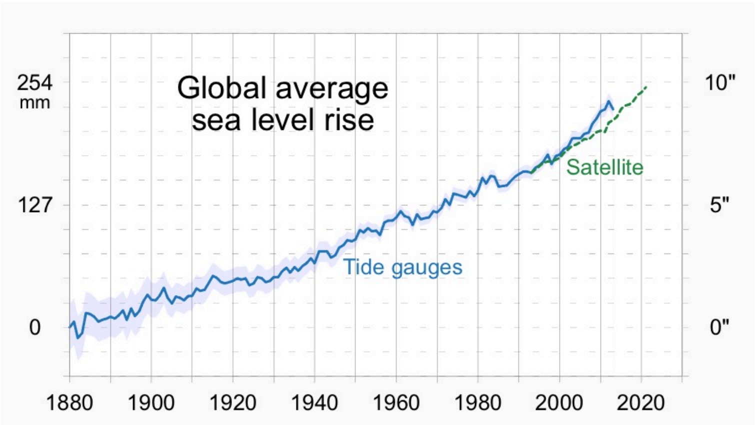


Figure 34. Chart showing cumulative changes in sea level for the world's oceans since 1880, based on a combination of long-term tide gauge measurements and recent satellite measurements.

Media Attributions

- After the Rainforest © Rod Waddington is licensed under a CC BY-SA (Attribution ShareAlike) license
- Share-of-fish-stocks-overexploited_(OWID_0995)(1) © Our World In Data is licensed under a CC BY (Attribution) license
- rosemary beetle © gbohne is licensed under a CC BY-SA (Attribution ShareAlike) license
- 1024px-Plastic_Pollution_in_Ghana © Muntaka Chasant is licensed under CC BY-SA (Attribution ShareAlike) license

Applying Population Genetics to Conservation Biology

Conservation biology increasingly relies on the understanding of genetic principles to inform strategies aimed at preserving biodiversity. Three key concepts—**genetic diversity**, **effective population size**, and **gene flow**—are pivotal in maintaining the resilience and adaptability of populations. Each of these concepts contributes to how we approach conservation challenges in different ecosystems.

Genetic diversity forms the foundation for populations to adapt to changing environmental conditions, disease pressures, and other stressors. Populations with high genetic diversity harbor a broad range of traits, increasing the likelihood that some individuals can survive and reproduce under new conditions. Such adaptability is crucial in the face of climate change, habitat fragmentation, and emerging diseases. Low genetic diversity, on the other hand, can lead to inbreeding depression, where the accumulation of deleterious recessive alleles reduces fitness and increases susceptibility to disease. This reduction in fitness can further shrink already endangered populations, setting them on a path toward extinction. For example, the genetic bottlenecks experienced by cheetahs and the inbreeding issues in Florida panthers illustrate the detrimental effects of reduced genetic diversity. Effective population size is a critical measure in conservation because it determines the rate at which genetic drift—the random loss of alleles—affects a population. Small effective population sizes increase the rate of genetic drift, leading to the rapid loss of genetic diversity and fixation of deleterious alleles. Populations with low Ne are at higher risk of inbreeding depression, reducing their adaptive potential. Maintaining an effective population size above critical thresholds is essential to ensure long-term genetic health and viability. Genetic management plans for endangered species, such as California condors and the Tasmanian devil, emphasize increasing Ne through captive breeding, managed introductions, and genetic rescue. Gene flow, or the transfer of genes between populations, plays a vital role in conservation biology by maintaining genetic diversity and reducing the risks associated with inbreeding. Connectivity between populations ensures that genes can move across the landscape, allowing populations to adapt and remain resilient. This connectivity counters genetic drift and introduces beneficial alleles into smaller populations. Barriers to gene flow, such as habitat fragmentation due to agriculture or urbanization, isolate populations, increasing their susceptibility to genetic issues. For instance, highways and other infrastructure often disrupt migration corridors for large mammals. In aquatic environments, dams can block fish migrations, leading to isolated populations and reducing genetic diversity. Restoring gene flow through wildlife corridors, habitat restoration, and assisted migration can help re-establish connectivity. Corridors enable the natural movement of individuals between fragmented habitats, while habitat restoration efforts create stepping stones that facilitate gene flow. Assisted migration moves individuals or entire populations to areas where they can thrive, particularly for species with limited dispersal capabilities.

Genetic Management of Endangered Species Using Captive Breeding Programs

Captive breeding programs for endangered species are critical tools for preventing extinction and maintaining biodiversity. They offer a controlled environment to support population growth, genetic management, and eventual reintroduction into the wild. To achieve these goals effectively, population genetic principles such as genetic diversity, effective population size, and gene flow are carefully incorporated into the design and management of these programs. The primary goal in captive breeding is to maximize and maintain genetic diversity, which ensures that the population retains a broad range of traits to adapt to environmental change.

Figure 35. A captive Scottish wildcat (Felis silvestris) at the Wildwood Trust Wildcat Captive Breeding Program. It is a critically endangered species with a census population size of below 500.

Captive breeding programs begin with a genetically diverse founding population, sourced from a wide range of wild individuals. This initial diversity forms the basis for long-term management, ensuring that as many alleles as possible are represented. Detailed records of breeding pairs, family histories, and individual genetic markers are then used to minimize inbreeding. By tracking **pedigrees,** managers avoid mating closely related individuals, thereby reducing inbreeding depression. The reproductive contribution of each founder's lineage is balanced across generations. Breeding pairs are chosen to ensure that no particular genetic lineage dominates, preventing the loss of alleles due to genetic drift. Effective population size is managed to limit genetic drift and maintain genetic diversity over time. This involves pairing individuals strategically over time ensures that the genetic contributions are as equal as possible, preventing dominance by a few individuals and reducing skewed reproductive success (**breeding rotation**), monitoring generation intervals and life history traits allows for optimizing breeding schedules. By staggering the ages of breeding individuals, the effective population size is increased, slowing the loss of genetic variation. Finally, sufficient numbers of breeding individuals are maintained to keep the effective population size high. This often requires maintaining larger captive populations than the minimum needed for short-term survival.

Captive breeding programs face challenges, such as maintaining natural behaviors and avoiding domestication effects. Genetic management requires adaptive strategies that respond to new data, environmental changes, and emerging genetic issues. Managers must continuously assess genetic markers, monitor population health, and adjust breeding plans accordingly.

GROUP REFLECTION QUESTION

Considering the concepts outlined in the assigned reading above, evaluate the ethical implications and potential long-term ecological impacts of using captive breeding programs as a primary strategy for conserving endangered species. How do we balance the immediate need to prevent extinction with the potential risks associated with captivity, such as loss of natural behaviors and genetic adaptation to captivity rather than the wild? Provide specific examples and suggest alternative or complementary conservation strategies that could mitigate these risks.

Media Attributions

- scottish wildcat © Smudge9000 is licensed under a CC BY-SA (Attribution ShareAlike) license

The Extinction Vortex

What is an extinction vortex?

The extinction vortex is a self-reinforcing process in which small populations are increasingly likely to decline toward extinction due to a combination of genetic, demographic, and environmental factors. As a population diminishes, it becomes increasingly vulnerable to genetic and ecological challenges, creating a feedback loop that accelerates its path toward extinction. This concept was first formally described by conservation biologist Michael Soulé in the 1980s. He categorized the factors contributing to the extinction vortex into two major categories: **deterministic factors** (*such as habitat destruction and overhunting*) and **stochastic factors** (*random events like natural disasters, disease outbreaks, or genetic issues*). Together, these factors amplify each other's effects.

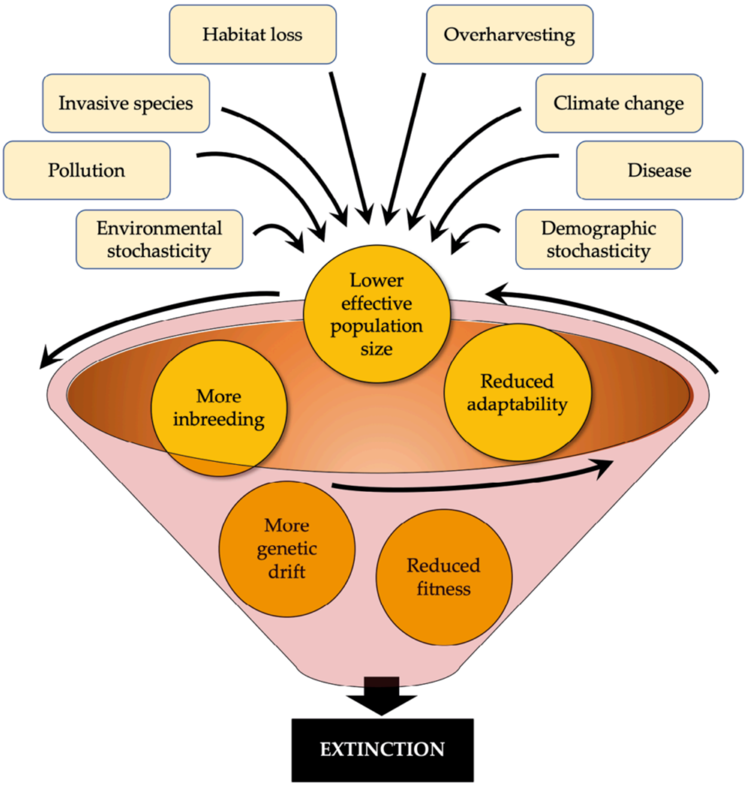


Figure 36. Diagram illustrating the extinction vortex, which describes a process whereby the factors that affect small populations can drive its size progressively downward towards extinction.

Genetic factors play a significant role in driving the extinction vortex. Two critical concepts in this regard are genetic load and inbreeding depression:

1. **Genetic Load**
   Genetic load refers to the accumulation of deleterious alleles in a population's gene pool, reducing the overall fitness of individuals. In large populations, selection can effectively eliminate harmful alleles. However, in small populations,

genetic drift—a random change in allele frequencies—may cause these alleles to become more frequent, increasing the genetic load. This results in a general decline in population fitness, making individuals more susceptible to disease and reducing reproductive success.

2. **Inbreeding Depression**
   As a population shrinks, the likelihood of close relatives breeding increases, leading to inbreeding. Inbreeding depression describes the reduced biological fitness observed in offspring due to the increased expression of recessive deleterious alleles. This results in individuals with lower fertility, higher mortality rates, and greater susceptibility to disease, further reducing the population's ability to recover.

Demographic and Environmental Impacts

Genetic problems are compounded by demographic and environmental challenges that further weaken populations such as **demographic stochasticity,** whereby small populations are disproportionately affected by random fluctuations in birth and death rates. This demographic stochasticity can lead to skewed sex ratios and poor recruitment, accelerating the decline. In **environmental stochasticity,** small populations are highly vulnerable to environmental changes, such as droughts, floods, and temperature fluctuations and a single catastrophic event can cause local extinction in small, isolated groups. Finally, reduced population sizes also lead to **Allee effects**, where individuals struggle to find mates or cooperate effectively. This reduces reproduction and survival rates.

The Feedback Loop

The extinction vortex arises because these genetic, demographic, and environmental factors form a feedback loop that worsens over time. As genetic diversity declines, inbreeding depression increases, which, coupled with demographic and environmental stochasticity, reduces reproductive success and survival rates. This further shrinks the population size, exacerbating inbreeding and the genetic load, and making the population increasingly vulnerable to stochastic events.

Breaking the Vortex

Breaking the extinction vortex requires intensive conservation efforts to restore populations above critical sizes. Strategies include:

- **Genetic Management:** Introducing new genetic material through translocation or managed breeding can reduce genetic load and inbreeding depression.
- **Habitat Restoration:** Improving habitats and creating wildlife corridors can reduce isolation and promote natural gene flow.
- **Captive Breeding and Reintroduction:** Establishing ex-situ populations can help maintain genetic diversity and serve as a source for reintroducing individuals into the wild.

Genetic purging is an intriguing concept in conservation genetics that suggests populations can, under certain conditions, rid themselves of deleterious alleles and recover from the brink of an extinction vortex. This process operates within small, isolated populations where the effects of inbreeding and genetic drift are pronounced. In the context of genetic purging, the increased inbreeding that typically contributes to inbreeding depression can paradoxically become a mechanism for recovery. As inbreeding intensifies in a small population, deleterious recessive alleles are more frequently expressed due to the pairing of similar genes. This leads to individuals with harmful traits manifesting these traits more clearly, which often results in reduced survival or reproductive success. When these individuals with detrimental traits have lower fitness and fail to reproduce as effectively, the harmful alleles they carry are less likely to be passed on to the next generation. Over time, this can lead to a reduction in the population's genetic load — the collective burden of these harmful alleles. The critical factor here is the balance between the speed at which deleterious alleles are purged through natural selection and the rate at which genetic diversity is lost due to genetic drift and inbreeding.

For genetic purging to be beneficial, a population must have enough genetic diversity to survive the initial increase in inbreeding depression. If a population is too small or its genetic diversity is too low, the immediate negative impacts of inbreeding depression might overwhelm any potential long-term benefits from purging. Thus, while genetic purging can help a population recover, it is not a panacea and works within specific limits. Natural examples of genetic purging are rare but insightful. Some conservationists observe this process in captive breeding programs where the initial generations suffer high levels of inbreeding depression, but later generations show signs of recovery as harmful alleles are eliminated. This suggests that careful management, even of very small populations, can harness the purging process to improve population viability over time.

However, relying solely on genetic purging is risky. Conservation efforts often complement potential purging with strategies like introducing new genetic material to boost diversity, establishing protected areas to support larger, more stable populations, and enhancing habitat connectivity to facilitate gene flow. These measures ensure that while purging may aid recovery, it is part of a broader strategy to restore and sustain biodiversity. In essence, genetic purging is a nuanced process where the detrimental effects of inbreeding can lead to a long-term reduction in genetic load, potentially aiding population recovery. Yet, this process requires careful consideration and often needs to be supported by proactive conservation measures to ensure the overall health and resilience of endangered populations.

Media Attributions

- The_extinction_vortex_describes_a_process_whereby_the_factors_that_affect_small_populations_can_drive_its_size_progressively_downward_towards_extinction © John W. Wilson, Richard B. Primack is licensed under a CC BY (Attribution) license

Inbreeding Effects

Inbreeding is the *mating of individuals that are closely related genetically*, resulting in an increase in homozygosity and a corresponding decrease in genetic diversity within a population. This phenomenon can have significant consequences for the fitness and survival of populations, particularly those that are small or isolated.

Inbreeding Coefficient

The **inbreeding coefficient (F)** is a measure that quantifies the probability that an individual has inherited identical alleles from both parents due to their ancestors having been related. This measure is crucial in conservation genetics as it helps in understanding the extent of inbreeding within a population and its potential impacts on genetic health and variability.

The following formula provides the calculation of *F*:

$$F = 1 - \frac{H_0}{H_T}$$

Here:

- F is the inbreeding coefficient.
- H_0 (often denoted as H_O) is the observed heterozygosity in the population, which is the actual proportion of heterozygous individuals.
- H_T is the expected heterozygosity under Hardy-Weinberg equilibrium, assuming no inbreeding (also known as theoretical or expected heterozygosity).

In this formula, H_T represents the heterozygosity that would be expected if the population were in Hardy-Weinberg equilibrium and there was no inbreeding. The observed heterozygosity, H_O, is what is actually measured in the population. The difference between these values, normalized by H_T, gives the inbreeding coefficient F. This approach allows researchers to quantify the impact of inbreeding on genetic diversity by comparing the actual genetic diversity observed to what would be expected if all individuals were mating randomly and there was no inbreeding. A decrease in H_O relative to H_T indicates inbreeding and its associated effects on the genetic structure of the population.

Evolutionary Adaptations for Avoiding Inbreeding

Organisms have evolved a variety of strategies to minimize the risks of inbreeding depression, which arises from the increased expression of deleterious alleles in genetically similar individuals. These adaptations can be behavioral, physiological, or ecological, and vary widely across different taxa. Here are some notable adaptations:

Dispersal Mechanisms: Many species have evolved mechanisms that promote the dispersal of offspring to new territories. This dispersal reduces the likelihood of mating with close relatives. In *natal dispersal*, offspring leave their birth location to establish territories elsewhere, a common behavior in birds and mammals. In many mammals, males often disperse farther than females, reducing the risk of inbreeding. In plants, seeds are dispersed through wind, water, or animals to increase the likelihood of germinating far from the parent plant, reducing the chances of inbreeding.

Kin Recognition: Some mammals and birds can distinguish their siblings or other relatives through chemical cues, vocalizations, or other means, leading to mate avoidance. One notable example is found in Belding's ground squirrels (*Urocitellus beldingi*), which can recognize their kin through scent. These ground squirrels produce unique chemical cues from specialized glands, allowing them to identify close relatives and avoid mating with them, thereby reducing the risk of inbreeding.

Self-Incompatibility Systems: In flowering plants, self-incompatibility systems prevent pollen from fertilizing ovules if the genetic relationship is too close. This promotes cross-pollination and genetic diversity.

Social Structures: The organization of social structures within animal societies also plays a role. For example, African elephants (*Loxodonta africana*) live in matriarchal family groups led by a dominant female, the matriarch, and are composed of related females and their offspring. As male elephants reach sexual maturity around 14-15 years old, they leave their natal groups to join loose associations with other males, known as bachelor groups. This dispersal is a primary mechanism for preventing inbreeding, ensuring that males do not mate with closely related females from their birth groups. Additionally, elephants possess a well-developed sense of kin recognition, which helps them avoid mating with close relatives even in complex social environments. Elephants' ability to travel long distances further facilitates the mixing of genetic material between different populations, reducing the chances of inbreeding.

Dominance Hierarchies: In some groups, dominant individuals control mating opportunities, often excluding siblings or other close relatives from breeding. For example, in wolf packs (*Canis lupus*), dominance hierarchies play a crucial role in preventing inbreeding and maintaining genetic diversity. These packs are typically structured around an alpha pair, consisting of the dominant male and female, who are the primary breeders within the group. This alpha pair exerts control over mating opportunities, effectively excluding subordinate members, often their offspring, from reproducing. This hierarchical structure ensures that only the most fit and typically unrelated individuals within the pack breed, thus reducing the likelihood of inbreeding.

Cooperative Breeding: Some species form groups where only a few individuals breed, while others act as helpers. These helpers are often closely related and avoid inbreeding by not reproducing within the group. One notable example is found in meerkats (*Suricata suricatta*). In meerkat societies, only the dominant pair typically breeds, while subordinate members of the group act as helpers. These helpers, often closely related to the dominant pair, assist in raising the offspring by providing food, protection, and other forms of care. This social structure significantly reduces the likelihood of inbreeding because the helpers do not reproduce within the group. By refraining from breeding, these related helpers avoid mating with close relatives. Additionally, this cooperative breeding system enhances the survival rate of the dominant pair's offspring, increasing the overall fitness of the group.

Mating Systems: Certain mating systems are designed to maximize genetic diversity. For example, in lek mating systems, males congregate in specific display areas known as leks to perform competitive displays and attract females. The females visit these leks, observe the males, and choose their mates based on the quality of the displays. This system promotes genetic diversity in several ways. First, females have the opportunity to select mates from a large pool of males **(female choice),** often choosing those with the most impressive displays. This selection process increases the likelihood of mating with genetically diverse and high-quality males, which enhances the genetic fitness of the offspring. Second, in many lekking species, females may mate with multiple males during a single breeding season, further increasing genetic diversity among their offspring **(multiple mates).** Third, the lek mating system encourages genetic mixing by bringing together males from various locations, ensuring that genes are widely dispersed across the population. This reduces the risk of inbreeding and maintains a healthy gene pool **(genetic admixture).**

Multiple Mating: In polyandrous systems (one female mates with multiple males), females can avoid inbreeding depression by increasing the genetic diversity of their offspring. Similarly, polygynous systems can prevent inbreeding by expanding the pool of potential mates.

Outbreeding Preference: Some species exhibit a preference for unrelated mates, selecting for individuals that are genetically different based on physical traits or pheromones. For example, house mice (*Mus musculus*) have evolved a highly sensitive olfactory system that allows them to detect pheromones—chemical signals used to communicate reproductive status and genetic compatibility. Mice can distinguish between the pheromones of related and unrelated individuals, and they tend to prefer the scent of potential mates with different genetic backgrounds, particularly those with dissimilar major histocompatibility complex (MHC) genes. This preference enhances immune system diversity in offspring.

Temporal and Spatial Separation: Plants, fungi, and some marine organisms have evolved mechanisms that separate the timing and location of gamete release or maturation.

Spatial Separation: Some plant species have physical barriers or mechanisms that separate male and female gametes spatially within the plant, reducing self-pollination.

These adaptations highlight the various strategies that species use to avoid inbreeding depression, ensuring the continued genetic health of their populations and increasing the likelihood of survival in the long term.

GROUP DISCUSSION QUESTION

Reflect on the various evolutionary adaptations organisms have developed to avoid inbreeding, as described in the reading. Choose one adaptation (e.g., dispersal mechanisms, kin recognition, self-incompatibility systems, social structures, dominance hierarchies, cooperative breeding, mating systems, multiple mating, or outbreeding preference) and discuss its effectiveness in maintaining genetic diversity and reducing inbreeding depression. Consider both the advantages and potential limitations of this adaptation. How might this knowledge inform conservation strategies for endangered species? Provide specific examples to support your discussion.

Media Attributions

- inbreedingcoefficient

Genetic Rescue

Genetic rescue is a strategy aimed at *improving the genetic diversity and fitness of a small, inbred population by introducing individuals from another population.* This process helps to reduce inbreeding depression and increase the population's adaptability and resilience to environmental changes. By enhancing genetic variability, genetic rescue can prevent the extinction of endangered species and support their long-term survival.

Outbreeding

populations or genetically distinct groups to enhance genetic diversity. In the context of endangered species conservation, this strategy is used as a genetic rescue technique to mitigate the negative effects of inbreeding depression and to increase the genetic health of small populations.

Benefits of Outbreeding for Genetic Rescue:

1. **Increased Genetic Diversity:**
   Introducing new genetic material can replenish the genetic diversity lost through inbreeding and genetic drift. This diversity can enhance the overall adaptability and resilience of the population to environmental changes and disease.
2. **Reduction of Inbreeding Depression:**
   By introducing alleles from an unrelated population, the prevalence of harmful recessive alleles can be diluted, reducing the expression of genetic disorders and increasing fitness.
3. **Improved Reproductive Success and Viability:**
   Increased genetic diversity often leads to improved fertility rates, higher survival rates among offspring, and a more robust immune system. This can enhance the long-term viability of the endangered population.

4. **Enhanced Evolutionary Potential:**
   A genetically diverse population is better equipped to respond to environmental changes and challenges, which is crucial for species recovery and adaptation in a changing climate.

Problems and Risks of Outbreeding:

1. **Outbreeding Depression:**
   While outbreeding can improve genetic diversity, it can also result in outbreeding depression. This occurs when two distinct populations have evolved different local adaptations, and their hybrid offspring lack the specialized traits needed to survive in either parent's environment. Hybrid offspring may exhibit reduced fitness due to the breakdown of coadapted gene complexes.
2. **Genetic Swamping:**
   Introducing new genetic material can overwhelm the local genetic structure, erasing unique local adaptations and homogenizing populations. This may threaten the survival of endemic traits critical for the species' adaptation to specific ecological niches.
3. **Behavioral and Social Disruption:**
   Mixing individuals from different populations can lead to conflicts due to differences in mating systems, social structures, or behaviors, affecting reproduction and social cohesion.
4. **Disease Transmission:**
   Movement of individuals between populations risks introducing new pathogens to a naïve population, potentially causing disease outbreaks that can further threaten the endangered species.

Figure 37. The Florida Panther (Puma concolor coryi). To combat inbreeding depression, a genetic rescue initiative introduced Texas cougars to the dwindling Florida panther population.

De-extinction Initiatives

De-extinction technology, the process of reviving extinct species, has gained increasing attention in recent years as a potential tool for genetic rescue. While this concept has been popularized through speculative fiction, significant scientific advancements are turning de-extinction into a tangible, albeit challenging, possibility. The objective of using de-extinction as a genetic rescue strategy is to enhance biodiversity and restore lost genetic traits to current ecosystems.

There are three primary methods through which de-extinction might be achieved: back-breeding, cloning, and genetic engineering. Each has its potential benefits and limitations.

Back-breeding involves selectively breeding closely related extant species to recreate the phenotypes and traits of the extinct species. For instance, by choosing individuals that possess particular ancestral characteristics, conservationists can emphasize these traits through generations of breeding. One prominent example is the attempt to recreate the Aurochs, an extinct wild cattle species. Scientists have selectively bred modern cattle breeds to produce a genetically similar animal that mimics the appearance and behaviors of the original.

Cloning is the process of creating genetically identical copies of an organism by using preserved cells of the extinct species. In cloning, scientists take a nucleus from a cell of the extinct species and insert it into an egg cell of a closely related living species that has had its nucleus removed. The egg is then stimulated to develop into an embryo and implanted into a surrogate mother. However, cloning is technically challenging due to the availability and preservation quality of ancient DNA. Moreover, cloning often faces ethical concerns and low success rates in developing healthy offspring.

Genetic Engineering represents the most cutting-edge approach, leveraging CRISPR/Cas9 gene-editing technology to modify the DNA of a closely related living species. By comparing the genomes of extinct and living species, scientists can identify genetic differences responsible for specific traits. The genome of the living species is then altered to reflect the extinct species' traits. An example includes the project to bring back the woolly mammoth by editing the genome of Asian elephants. The goal is to produce elephants with traits adapted to colder climates, akin to the extinct mammoths.

While de-extinction is theoretically feasible, its practical application as a means of genetic rescue remains contentious. Proponents argue that de-extinction can restore lost ecological functions and enhance genetic diversity, particularly for closely related endangered species. For instance, reviving the passenger pigeon could re-establish its role in maintaining forest dynamics in North America. However, critics point out the high costs involved and the potential ecological risks of releasing genetically modified or resurrected species into modern ecosystems. Moreover, focusing on de-extinction could divert resources away from protecting existing endangered species that still require urgent attention. The ethical implications are also significant, raising questions about humanity's role in shaping nature. How do we decide which species should be resurrected, and how will their reintroduction affect current ecosystems? Furthermore, even if an extinct species were successfully revived, it would face significant challenges in adapting to current environmental conditions, which are vastly different from those in which it evolved.

Despite these challenges, de-extinction remains a fascinating frontier. It exemplifies the innovative thinking required to address the current biodiversity crisis and offers a new perspective on humanity's relationship with nature. Scientists and policymakers will need to carefully balance the technological potential with ethical and ecological considerations to determine the role that de-extinction should play in future conservation strategies.

Media Attributions

- floridapanther2 © Larry W. Richardson/USFWS is licensed under a CC BY (Attribution) license

Recommended Reading

The following are recommended readings – we will be discussing at least one of these articles in next week's group discussion:

1. Stuart KC, Cardilini APA, Cassey P, Richardson MF, Sherwin WB, Rollins LA, Sherman CDH (2020) Signatures of selection in a recent invasion reveal adaptive divergence in a highly vagile invasive species. *Molecular Ecology* 30: 1419 – 1434. https://doi.org/10.1111/mec.15601
2. Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology and Evolution* 22: 454 – 464. https://doi.org/10.1016/j.tree.2007.07.002
3. Foote AD, Martin MD, *et al.* (2019) Killer whale genomes reveal a complex history of recurrent admixture and vicariance. *Molecular Ecology* 28: 3427 – 3444. https://doi.org/10.1111/mec.15099
4. Hu Y, Qi D, Wang H, Wei F (2010) Genetic evidence of recent population contraction in the southernmost population of giant pandas. *Genetica* 138: 1297 – 1306. https://doi.org/10.1007/s10709-010-9532-2
5. Willoughby JR, Fernandez NB, Lamb MC, Ivy JA, Lacy RC, DeWoody JA (2014) The impacts of inbreeding, drift and selection on genetic diversity in captive breeding populations. *Molecular Ecology* 24: 98 – 110. https://doi.org/10.1111/mec.13020
6. Vandergast AG, Wood DA, Thompson AR, Fisher M, Barrows CW, Grant TJ (2016) Drifting to oblivion? Rapid genetic differentiation in an endangered lizard following habitat fragmentation and drought. *Diversity and Distributions* 22: 344 – 367. https://doi.org/10.1111/ddi.12398
8. Shapiro B (2017) Pathways to de-extinction: how close can we get

to resurrection of an extinct species? *Functional Ecology* 31: 996 – 1002. https://doi.org/10.1111/1365-2435.12705

PART VIII

CONCEPTS IN BEHAVIOR

LEARNING OBJECTIVES

At the end of this chapter, you should be able to:

1. List and explain the five different mating systems found among animals.
2. Explain how molecular data is used to understand animal mating systems.
3. Define and explain the concept of Extra Pair Paternity (EPP).
4. Explain how male and female reproductive behavior can alter genetic diversity patterns within a species.
5. Define and explain the concept of sexual conflict in animals.

Behavioral ecology bridges the gap between the actions of organisms and the genetic mechanisms that underpin these behaviors, offering insights into how evolutionary pressures shape not only the physical traits but also the behavioral strategies of species. From mating rituals and social structures to foraging patterns and habitat selection, behavioral ecology provides a context for interpreting molecular data in a broader ecological and evolutionary framework. By integrating behavioral observations with molecular genetics, we can uncover the genetic basis of behavior and its evolutionary consequences, thereby enriching our understanding of biodiversity and species adaptation.

Media Attributions

- © Claire Gribbin is licensed under a CC BY (Attribution) license

Sexual Selection

Revisiting Sexual Selection

Sexual selection is a specific form of natural selection that acts on an organism's ability to obtain or successfully reproduce with a mate. It was first introduced by Charles Darwin as a mechanism that could explain the evolution of traits that do not necessarily provide a direct survival advantage but instead enhance mating success. These traits can include physical characteristics, behaviors, or physiological features that make an individual more attractive or competitive in the mating arena. Sexual selection is often divided into two main categories: intersexual selection and intrasexual selection.

Intersexual selection refers to the process where individuals of one sex (typically females) choose mates based on certain desirable traits in the opposite sex (typically males). For instance, female peafowls tend to select males with the most elaborate and colorful plumage. This preference can lead to the evolution of extravagant traits, like bright feathers or complex courtship displays, even if these features might otherwise be disadvantageous in terms of survival.

Intrasexual selection, on the other hand, involves competition within the same sex (typically males) for access to mates. This competition can result in the development of traits that enhance fighting ability or dominance, such as larger body size, weapon-like structures (like antlers or tusks), or increased aggressiveness. In elephant seals, for instance, males fight for control of harems, with larger and stronger males monopolizing access to females.

In comparison to sexual selection, **natural selection** encompasses the broader concept of differential survival and reproduction based on heritable traits that increase an individual's fitness in its environment. It explains the adaptation of species over generations, leading to the evolution of traits that improve survival, resource acquisition, and overall reproductive success. Examples include changes in fur color for camouflage, the development of venom in predators, or physiological adaptations to extreme temperatures.

While sexual selection can sometimes work in concert with natural selection (traits that attract mates can also confer survival advantages), it often opposes natural selection. For example, the peacock's elaborate tail, while attractive to females, can make it more vulnerable to predators. This conflict can lead to evolutionary trade-offs where the benefits of increased reproductive success outweigh the costs to survival. Conversely, natural selection can limit the extent of sexual selection if a trait becomes too detrimental for survival.

Figure 38. Male fiddler crab, distinguishable by its conspicuously enlarged claw. This trait has evolved primarily through sexual selection, serving both as a signal to potential mates and as a weapon in male-male competition. While the large claw enhances mating success, it may also pose a trade-off by hindering mobility and increasing vulnerability to predators, illustrating the complex interplay between sexual selection and natural selection in evolutionary biology.

Media Attributions

- Fiddler crab © Gail Hampshire is licensed under a CC BY (Attribution) license

Mating Systems

Mating Systems and Genetic Diversity

Mating systems describe the patterns of mate association and reproductive strategies that different species adopt. They are shaped by evolutionary pressures and ecological factors, influencing genetic diversity and social structures within populations. Four main types of mating systems are generally recognized in animals: monogamy, polygyny, polyandry, and promiscuity. These systems often have profound effects on the genetic diversity of a species.

Monogamy is characterized by a pair bond between one male and one female. The bond may last for a single breeding season or extend over multiple seasons. Monogamous pairs often share parental responsibilities, cooperating in activities such as nest building, guarding, and feeding offspring. In birds like albatrosses and some mammals like wolves, monogamous pairs form stable family units, enhancing offspring survival through biparental care. Despite this outward stability, genetic studies have revealed that extra-pair copulations are common in many monogamous species, resulting in offspring sired by individuals outside the primary pair bond. This behavior can increase genetic diversity by introducing additional alleles into the gene pool and reducing the risk of inbreeding.

Polygyny involves one male mating with multiple females. This system is widespread among mammals and some bird species and can manifest in several forms. In "resource defense polygyny," males control access to resources such as food or nesting sites that females require, attracting mates. In "female defense polygyny," males directly defend a group of females (a harem) and prevent other males from gaining access to them. The reproductive success of males is highly skewed, with dominant individuals achieving the most matings, often resulting in reduced genetic diversity. However, female choice and competition among males can help maintain genetic variation.

Polyandry, the reverse of polygyny, is where one female mates with multiple males. This system is relatively rare in nature but can occur in species like the spotted sandpiper or the Jacana bird. Here, males invest heavily in parental care, often incubating eggs or caring for young, while females mate with several males. This system promotes genetic diversity as multiple fathers contribute to the offspring, increasing the genetic variation within broods. In some cases, polyandry can enhance offspring survival by increasing genetic compatibility or reducing the impact of harmful mutations.

Promiscuity involves both males and females mating with multiple partners without forming exclusive bonds. It is common in many primates and small mammals. This system can enhance genetic diversity by ensuring a wider distribution of alleles and reducing the likelihood of inbreeding depression. It can also increase the chances of successful reproduction through genetic bet-hedging. However, it may also lead to intense sperm competition among males, resulting in the evolution of larger testes or specialized reproductive strategies to increase their success.

The interplay between these mating systems and genetic diversity is significant. In monogamy, fidelity can reduce genetic diversity, while extra-pair copulations can counterbalance this. Polygyny often skews reproductive success toward dominant males, reducing genetic diversity in the next generation. In contrast, polyandry and promiscuity generally maintain higher genetic diversity by spreading reproductive opportunities across more individuals. Thus, understanding these systems provides crucial insights into the evolutionary pressures shaping genetic structure and adaptation within animal populations.

Application of Molecular Tools in Understanding Mating Systems

DNA fingerprinting revolutionized our understanding of mating systems by allowing researchers to genetically analyze parent-offspring relationships, revealing the true nature of reproductive behavior within animal populations. Before this technology, many assumptions about monogamy, polygamy, or other mating systems were based on observed behaviors or social structures, which often did not reflect genetic reality. DNA fingerprinting involves the analysis of highly variable regions in the genome known as "minisatellites," which consist of short, repeating DNA sequences. The variation in the number and sequence of these repeats is unique for each individual (except identical twins), creating a distinctive genetic profile that can be used to identify relationships between individuals. By comparing the genetic profiles of potential parents with those of their offspring, researchers have discovered surprising patterns in many species, revealing a prevalence of extra-pair copulations and multiple paternities in socially monogamous systems. For instance, DNA fingerprinting showed that extra-pair paternity is common in birds traditionally thought to be monogamous. In tree swallows and blue tits, more than half of the offspring in a single brood may be fathered by males other than the social partner of the female. This discovery led to new insights into the complexities of avian mating systems and the evolution of reproductive strategies like sperm competition.

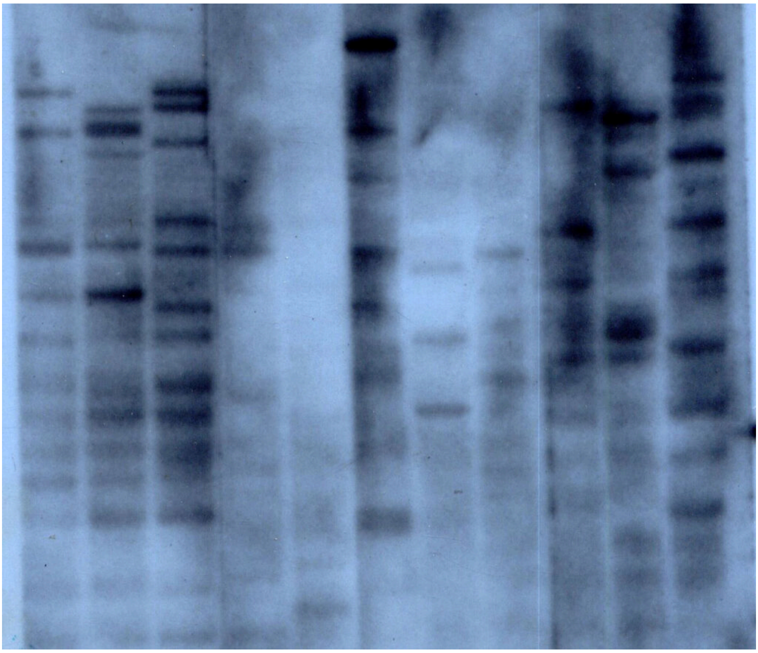


Figure 39. The first DNA fingerprint. The first three lanes contain DNA from a woman, her mother and her father respectively. Lanes 4 – 11 contain DNA from assorted other species including mouse, baboon, lemur, cow, grey seal and tobacco (last lane). The DNA probe used in this experiment detected minisatellites whose length varies between individuals.

Microsatellites, also known as short tandem repeats (STRs), are another type of highly variable genetic marker. While conceptually similar to minisatellites, they consist of shorter repeating DNA sequences, generally 2-6 base pairs in length. Their high mutation rates result in numerous alleles, making them ideal for genetic studies. Unlike DNA fingerprinting, which analyzes patterns across multiple minisatellite regions simultaneously, microsatellite analysis examines a smaller number of specific loci. By focusing on these loci, researchers can determine individual genotypes and assess genetic variation, parentage, and population structure more precisely.In parentage analysis, microsatellites offer a high level of discrimination due to the large number of alleles available per locus. This makes them particularly useful for reconstructing pedigrees in wildlife populations and understanding the genetic consequences of different mating systems. For instance, researchers have used microsatellites to determine paternity in polygynous mammals like seals and lions, showing that reproductive success is often skewed toward dominant males. Similarly, studies on socially monogamous birds using microsatellites have confirmed widespread extra-pair paternity.

Extra Pair Paternity (EPP)

Extra-pair paternity refers to the phenomenon where offspring within a brood or litter are sired by males other than the social partner or mate of the female. This occurrence challenges traditional views of animal mating systems, particularly in species previously thought to exhibit strict monogamy. Extra-pair paternity is widespread across many taxa, from birds to mammals, and plays a significant role in shaping the genetic structure and evolutionary dynamics of populations. The classic case study illustrating extra-pair paternity involves the superb fairy-wren (*Malurus cyaneus*), a small Australian bird that appears monogamous at first glance. Males and females form long-term pair bonds, jointly defending territory and raising young. However, genetic studies using DNA fingerprinting and microsatellites revealed a surprising truth: a high percentage of chicks (often more than 70%) were sired by males outside the primary pair bond.

This discovery in superb fairy-wrens has significant consequences and implications. First, it underscores the complexity of avian mating systems, demonstrating that social and genetic monogamy can be decoupled. While the social pair works together to raise the young, females mate with multiple males, potentially seeking genetic benefits such as increased genetic diversity and quality in their offspring. This behavior can lead to a variety of evolutionary and ecological outcomes.

Figure 40. The superb fairywren (Malurus cyaneus).

1. **Genetic Diversity:** Extra-pair paternity introduces new genetic material into the brood, enhancing the genetic diversity of the offspring. This diversity is crucial for the adaptive potential of the population, allowing it to better respond to changes in the environment and reducing the risk of inbreeding depression.
2. **Sexual Selection:** The phenomenon supports the role of sexual selection in evolution. Females may choose extra-pair mates based on traits that signal genetic fitness, such as vibrant plumage, vigorous displays, or good health. This selection pressure encourages the development of these traits in males, driving evolutionary change.
3. **Sperm Competition:** With the possibility of multiple males fertilizing the same set of eggs, sperm competition becomes a significant factor. Males may evolve strategies such as increased sperm count or faster-swimming sperm to ensure

their paternity over competitors.

4. **Parental Investment:** The revelation of extra-pair paternity can also affect parental care patterns. Males might adjust their investment in offspring based on the certainty of their paternity. In some species, males reduce their contribution to parental care if there is a high risk of cuckoldry, while in others, the risk is mitigated by the benefits of maintaining a pair bond or due to difficulty in assessing paternity.
5. **Social Structure and Conflict:** The dynamics of extra-pair paternity can lead to complex social interactions and conflicts within populations. It may influence the formation of territories, the stability of pair bonds, and the interactions among neighbors and conspecifics.

Extra-pair paternity (EPP) presents a significant evolutionary challenge for males in many species, particularly those that invest heavily in parental care. The risk of investing in offspring that are not genetically their own can lead to significant evolutionary losses for males. As a result, males across various species have evolved a range of strategies to counteract the risk of EPP, enhancing their chances of reproductive success and ensuring that their investments are directed towards their own genetic offspring.

One of the primary strategies is **mate guarding**, where males closely monitor and control the movements and interactions of their mates, especially during fertile periods. This behavior is widespread among birds, mammals, and even insects. For example, in many bird species, males stay close to their mates during the breeding season, following them closely and often interrupting interactions with other males. This reduces the opportunities for females to engage in extra-pair copulations.

Increased vigilance and aggression toward potential rivals is another strategy males use to prevent EPP. Males may display aggressive behaviors toward other males that approach their mate or territory. This not only discourages rivals but also signals strength and fitness to their mate, potentially reinforcing the pair bond and deterring the female from seeking extra-pair mates. In many mammalian species, such as deer and primates, males often fight to establish dominance, and the dominant male gets exclusive mating rights with females in the area, reducing the likelihood of EPP.

Sperm competition is an evolutionary adaptation that has arisen in response to EPP. In species where females mate with multiple males, the sperm of different males compete to fertilize the eggs. This has led to various adaptations, including increased sperm production, changes in sperm morphology to enhance speed and competitiveness, and even the development of copulatory plugs in some species, which block further matings by the female. For instance, in many rodent species, males produce copulatory plugs after mating, which can physically prevent other males from mating with the female immediately afterward.

Cryptic female choice is a phenomenon where females have the ability to influence which male's sperm fertilizes their eggs. In response, males have evolved traits that may influence this choice in their favor. This includes producing more attractive or stimulating seminal fluids that can influence the female's reproductive tract to favor their sperm. In some bird species, males provide more parental care or resources to the females, which can increase the female's reliance on the male and reduce her motivation to seek extra-pair copulations.

Paternity assurance through frequent copulations is another strategy where males attempt to ensure their paternity by increasing the frequency of mating with their partner, especially close to the time of ovulation. This ensures a higher likelihood that their sperm will be the ones to fertilize the eggs. This is seen in many species, including some primates and birds, where males initiate copulations more frequently when the female is most fertile.

Media Attributions

Recommended Reading

The following are recommended readings – we will be discussing at least one of these articles in next week's group discussion:

1. Rafter MA, McCulloch GA, Daglish GJ, Gurdasani K, Walter GH (2017) Polyandry, genetic diversity and fecundity of emigrating beetles: understanding new foci of infestation and selection. *Journal of Pest Science* 91: 287 -298. https://doi.org/10.1007/s10340-017-0902-8
2. McLeod L, Marshall DJ (2009) Do genetic diversity effects drive the benefits associated with multiple mating? A test in a marine invertebrate. *PLoS One* 4: e6347. https://doi.org/10.1371/journal.pone.0006347
3. Madsen T, Ujvari B, Bauwens D, Gruber B, Georges A, Klaassen M (2023) Polyandry and non-random fertilization maintain long-term genetic diversity in an isolated island population of adders (V*ipera berus*). *Heredity* 130: 64 – 72. https://doi.org/10.1038/s41437-022-00578-2
4. Dolotovskaya S, Roos C, Heymann EW (2020) Genetic monogamy and mate choice in a pair-living primate. *Scientific Reports* 10: 20328. https://doi.org/10.1038/s41598-020-77132-9
5. Daly-Engel TS, Grubbs RD, Feldheim KA, Bowen BW, Toonen RJ (2010) Is multiple mating

beneficial or unavoidable? Low multiple paternity and genetic diversity in the shortspine spurdog *Squalus mitsukurii*. *Marine Ecology Progress Series* 403: 255 – 267. https://doi.org/10.3354/meps08417

PART IX

GENOMICS: THE NEW FRONTIER OF MOLECULAR ECOLOGY

LEARNING OBJECTIVES

At the end of this chapter you should be able to:

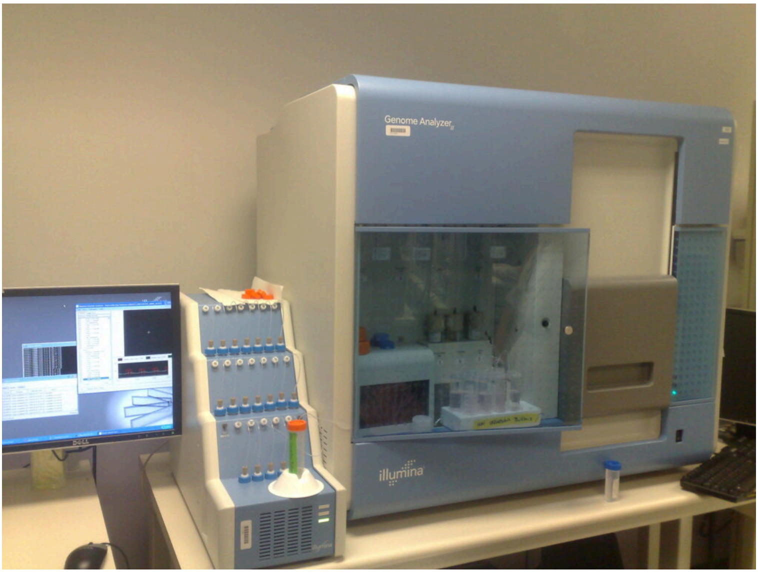
- Identify key genomic technologies and describe their roles in advancing molecular ecology.
- Explain the principles behind Next-Generation Sequencing, Third-Generation Sequencing, and other genomic tools.
- Evaluate the contributions of comparative genomics to understanding evolutionary relationships and mechanisms.
- Evaluate case studies that demonstrate the use of genomics in species conservation, invasive species management, and microbial ecology.
- Describe the challenges associated with employing genomic data into molecular ecology research.

Significance of the Genomic Era

Genomics is *the comprehensive study of the complete genetic material of organisms—their entire genomes.* This discipline encompasses the sequencing, mapping, and analysis of genomes, which include not just the coding regions, which are the genes themselves, but also the non-coding regions that regulate gene expression. In molecular ecology and evolution, genomics provides unparalleled insights into genetic diversity, species relationships, and the genetic underpinnings of traits, adaptations, and behaviors. The significance of genomics in molecular ecology is profound. It allows researchers to study organisms in ways that were not possible with traditional molecular markers. Genomics facilitates the examination of genetic variation at the finest scale across entire populations, enabling detailed studies of genetic drift, selection, gene flow, and speciation. Furthermore, it aids in understanding how genetic variation influences ecological interactions and evolutionary processes, thus providing a clearer picture of biodiversity and ecosystem functioning. The field of genomics truly took shape with the initiation of the Human Genome Project (HGP) in 1990, a large international effort to sequence the entire human genome. Completed in 2003, the HGP utilized Sanger sequencing, a method developed in the 1970s that became the gold standard for DNA sequencing. The project's success demonstrated the feasibility of large-scale sequencing and led to significant reductions in the cost of DNA sequencing. The introduction of NGS technologies in the mid-2000s marked a turning point in genomics, dramatically increasing sequencing speed and reducing costs even further. NGS platforms, such as those developed by Illumina and 454 Life Sciences, could generate massive datasets, enabling whole-genome sequencing on a much broader scale and facilitating comparative genomics among diverse species. The development of genomic technologies has had a transformative impact on biological sciences. In ecology and evolution, genomics has led to the emergence of new subfields such as evolutionary developmental biology (evo-devo), phylogenomics, and conservation genomics. Researchers can now link genetic variation directly to ecological outcomes, study the genetic basis of adaptation, and track the evolutionary histories of species over time scales previously unimaginable. Genomics has also enhanced our ability to conserve biodiversity, manage natural resources, and predict ecological responses to environmental changes, solidifying its role as a cornerstone of modern biological research. These developments reflect a field that is rapidly evolving, continuously enhanced by technological advances that expand the scope and depth of research in molecular ecology and evolution.

Key Genomic Technologies

Next Generation Sequencing (NGS) technologies have revolutionized genomic studies by allowing rapid, cost-effective sequencing of DNA and RNA. For instance, **Illumina's** sequencing platforms have been extensively used to perform whole-genome sequencing, metagenomics, and transcriptomics. A notable application is the sequencing of the Neanderthal genome, which provided insights into human evolution and our genetic overlap with Neanderthals, illustrating the potential of NGS in answering profound evolutionary questions.



Third generation sequencing technologies such as **PacBio's** Single Molecule, Real-Time (SMRT) sequencing and **Oxford Nanopore's** platforms offer longer read lengths, which are beneficial for genome assembly, identification of structural variants, and epigenetic modifications. For example, the use of PacBio sequencing to assemble the complete genome of the *Aedes aegypti* mosquito has significantly advanced our understanding of the genetic factors that confer disease vector capabilities, critical for controlling vector-borne diseases.

High-throughput genotyping technologies like **SNP chips** and genotyping by sequencing (GBS) are crucial for assessing genetic variation across populations at a reduced cost. A practical application can be seen in agricultural ecology, where SNP chip technology has been used to enhance breeding programs for crops like rice and wheat by rapidly identifying and selecting for traits such as drought resistance and improved yield.

Challenges in Analyzing Genomic Data

An integral part of analyzing genomic data is a comprehensive understanding of advanced analytics (Bioinformatics) which is required for processing and analyzing the vast amounts of data generated. This includes tasks such as read alignment, genome assembly, variant calling, and functional annotation. For example, bioinformatics tools like BLAST for sequence alignment and QIIME for analyzing microbial communities have become staples in ecological genomics. The primary challenges include data storage, computational requirements for data processing, and the need for sophisticated algorithms to handle noise and errors in the data. Additionally, interpreting the biological significance of genomic data requires robust statistical models and machine learning techniques. A case in point is the Earth BioGenome Project, aiming to sequence all eukaryotic life forms. The project faces immense bioinformatics challenges in data integration, requiring the development of new tools and collaboration platforms.

Case Example: Earth BioGenome Project

The Earth BioGenome Project (EBP) is an ambitious initiative that aims to sequence and catalog the genomes of all of Earth's eukaryotic biodiversity. Launched in 2018, the project seeks to provide a complete DNA sequence for each of the approximately 1.5 million known animal, plant, protozoan, and fungal species on Earth. This monumental task not only promises to revolutionize our understanding of biology and biodiversity but also presents significant challenges, particularly in the realm of bioinformatics. One of the foremost challenges is the sheer volume of data generated. Sequencing 1.5 million species means potentially petabytes of genomic data, encompassing a wide variety of genome sizes and complexities. For instance, the genome of a simple fungus may be relatively small and straightforward, while the genome of a complex flowering plant may be large and full of repetitive elements, which are difficult to sequence and assemble accurately. Integrating this vast amount of data from diverse organisms and multiple sequencing centers worldwide requires sophisticated data management strategies. The EBP must develop and maintain a standardized, open-access framework to ensure that genomic data, along with associated metadata (e.g., species information, geographical location, sample source), are accessible and usable. This integration is critical for comparative genomic analyses, which can reveal evolutionary relationships, genetic diversity, and functional genomics across the tree of life. Handling and processing this data also require immense computational resources. The bioinformatics pipeline for the EBP involves data storage, sequence assembly, annotation, and analysis, each of which demands significant computational power and efficient algorithms. The project therefore relies on cloud computing solutions and high-performance computing centers to manage the workflow.

Existing genomic analysis tools are often not scalable to the magnitude required by the EBP. There is a need for the development of new computational tools that can more efficiently process large-scale sequence data, accurately assemble genomes from complex and repetitive DNA regions, and perform in-depth comparative analyses across highly divergent species.

Media Attributions

- illumina © Jon Callas is licensed under a CC BY (Attribution) license
- © Konrad Förstner is licensed under a CC BY (Attribution) license

Application of Genomics in Molecular Ecology

Genome Wide Association Studies (GWAS)

Genome-wide association studies (GWAS) are a powerful tool used to identify genetic variants associated with particular traits across the genome. Unlike methods that focus on specific candidate genes, GWAS examines the entire genome, usually by scanning for single nucleotide polymorphisms (SNPs) that occur more frequently in individuals with a particular trait than in those without. This approach has become especially valuable in molecular ecology for linking genetic variation to phenotypic traits that are crucial for survival and reproduction in natural environments. GWAS can help elucidate the genetic basis of adaptive traits, such as drought resistance in plants or disease resistance in animals, by correlating specific genomic regions with ecological responses. This genomic approach allows researchers to uncover the polygenic nature of complex traits that are influenced by multiple genetic factors and environmental interactions. In an important study, Bac-Molenaar et al. (2015) used GWAS to investigate the genetic basis of heat and drought stress responses in A*rabidopsis thaliana*. By analyzing the genomes of hundreds of different A. *thaliana* accessions, they identified specific genetic variants associated with survival under extreme environmental conditions. This research highlighted alleles in genes related to heat shock proteins and water-use efficiency, providing insights into the plant's adaptation mechanisms. Such studies are crucial for understanding how genetic variation contributes to ecological adaptability and can guide breeding programs aimed at enhancing crop resilience to climate change.

Landscape & Seascape Genomics

Landscape and seascape genomics are emerging fields that integrate genomic data with geographic and environmental information to study how environmental factors shape genetic variation and population structure. Landscape genomics typically focuses on terrestrial environments, while seascape genomics deals with marine settings. Landscape and seascape *genetic* studies often focus on a relatively small number of traditional markers, which provides a limited view of the genome. While they are excellent for answering specific questions about genetic connectivity or population history. they lack the resolution to detect subtle or complex patterns of genetic adaptation and diversity. Genomics provides a high-resolution view of the genome, examining thousands to millions of genetic markers (like single nucleotide polymorphisms, SNPs) across the genome. This extensive data allows for a more detailed examination of genetic patterns and processes. In addition, unlike traditional methods that focus on a few specific markers, genomic studies can assess the entire genome. This comprehensive analysis can uncover regions of the genome associated with adaptation to environmental gradients or stressors, providing insights into the genetic basis of adaptation. Finally, and probably most importantly, beyond identifying genetic variation, genomics can help link this variation to functional outcomes. For example, it can identify which genes are involved in adaptation to specific environmental challenges (like temperature extremes or salinity changes) and how these genes are regulated. For example Tine et al. (2014) applied genomics based approach to study the European sea bass (*Dicentrarchus labrax*), a species found in diverse marine environments across the European coast. These researchers integrated genomic data with environmental variables such as salinity and temperature to assess how these factors influence genetic variation. The study revealed that population structure and adaptive traits in sea bass are significantly shaped by local environmental conditions, particularly salinity gradients.

These findings are pivotal for fisheries management and conservation strategies, ensuring the sustainability of sea bass across its natural range.

Comparative Genomics

Comparative genomics involves the analysis and comparison of genomes from different species. This field provides insights into the genetic differences and similarities that underpin the diversity of life. Comparative genomic studies can reveal conserved genes that have remained unchanged across evolutionary timescales, as well as adaptations unique to particular lineages. This approach is particularly useful for understanding evolutionary relationships, the function of genes, and the genetic basis of traits and diseases. By comparing genomes across a wide array of species, from microbes to mammals, scientists can trace the evolutionary history of genes and genomes, uncovering the molecular underpinnings of speciation, adaptation, and phenotypic diversity. A notable example is the comparative analysis of the immune genes in humans, mice, and other mammals. This research has uncovered conserved and species-specific immune responses, shedding light on how different mammals have evolved distinct defense mechanisms against pathogens. By comparing the genomes of species that have unique adaptations to their environments (e.g., bats with their viral resistance or elephants with cancer resistance), scientists can pinpoint genetic modifications that confer these unique traits. Such comparative studies not only elucidate evolutionary relationships but also advance our understanding of human diseases and potential treatments.

Recommended Reading

The following are recommended readings – we will be discussing at least one of these articles in next week's group discussion:

Tine M, Kuhl H, Gagnaire P-A *et al.* (2014) European sea bass genome and its variation provide insights into adaptation and to euryhalinity and speciation. *Nature Communications* 5: 5770. https://doi.org/10.1038/ncomms6770

Bac-Molenaar JA, Fradin EF, Becker FFM *et al.* (2015) Genome-wide association mapping of fertility reduction upon heat stress reveals developmental stage-specific QTLs in *Arabidopsis thaliana. The Plant Cell* 27: 1857 – 1874. https://doi.org/10.1105/tpc.15.00248

Bourgeois YXC, Warren BH (2021) An overview of current population genomics methods for the analysis of whole-genome resequencing data in eukaryotes. *Molecular Ecology* 30: 6036 – 6071. https://doi.org/10.1111/mec.15989

Davey JW, Blaxter ML (2011) RADSeq: Next generation population genetics. *Briefings in Functional Genomics* 9: 416 – 423. https://doi.org/10.1093/bfgp/elr007

Bibliography

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), 422-433.

Rowe, G., Sweet, M., & Beebee, T. J. C. (2017). *An introduction to molecular ecology*. Oxford University Press.

Hadrys, H., Balick, M., & Schierwater, B. (1992). Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Molecular ecology*, *1*(1), 55-63.

Vieira, M. L. C., Santini, L., Diniz, A. L., & Munhoz, C. D. F. (2016). Microsatellite markers: what they mean and why they are so useful. *Genetics and molecular biology*, 39, 312-328.

Karl, S. A., Toonen, R. J., Grant, W. S., & Bowen, B. W. (2012). Common misconceptions in molecular ecology: echoes of the modern synthesis. *Molecular Ecology*, *21*(17), 4171-4189.

Ekblom, R., & Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, 107(1), 1-15.

Ross, K. G. (2001). Molecular ecology of social behaviour: analyses of breeding systems and genetic structure. *Molecular ecology*, *10*(2), 265-284.

Kirk, H., & Freeland, J. R. (2011). Applications and implications of neutral versus non-neutral markers in molecular ecology. *International journal of molecular sciences*, *12*(6), 3966-3988.

McKinney, G. J., Larson, W. A., Seeb, L. W., & Seeb, J. E. (2017). RAD seq provides unprecedented insights into molecular ecology and evolutionary genetics: comment on Breaking RAD by Lowry et al.(2016). *Molecular ecology resources*, *17*(3), 356-361.

Andrew, R. L., Bernatchez, L., Bonin, A., Buerkle, C. A., Carstens, B. C., Emerson, B. C., … & Rieseberg, L. H. (2013). A road map for molecular ecology. *Molecular ecology*, *22*(10), 2605-2626.

Ford, M. J. (2002). Applications of selective neutrality tests to molecular ecology. *Molecular Ecology*, *11*(8), 1245-1262.

Bensch, S., & Åkesson, M. (2005). Ten years of AFLP in ecology and evolution: why so few animals?. *Molecular ecology*, *14*(10), 2899-2914.

Johnson, J. B., Peat, S. M., & Adams, B. J. (2009). Where's the ecology in molecular ecology?. *Oikos*, *118*(11), 1601-1609.

Moritz, C. (1994). Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular ecology*, *3*(4), 401-411.

DeSalle, R., & Goldstein, P. (2019). Review and interpretation of trends in DNA barcoding. *Frontiers in Ecology and Evolution*, *7*, 302.

Valentini, A., Pompanon, F., & Taberlet, P. (2009). DNA barcoding for ecologists. *Trends in ecology & evolution*, *24*(2), 110-117.

Hebert, P. D., & Gregory, T. R. (2005). The promise of DNA barcoding for taxonomy. *Systematic biology*, *54*(5), 852-859.

Pires, A. C., & Marinoni, L. (2010). DNA barcoding and traditional taxonomy unified through Integrative Taxonomy: a view that challenges the debate questioning both methodologies. *Biota Neotropica*, *10*, 339-346.

Hey, J. (2006). On the failure of modern species concepts. *Trends in ecology & evolution*, *21*(8), 447-450.

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic biology*, *56*(6), 879-886.

Mayr, E. (2000). The biological species concept. *Species concepts and phylogenetic theory: a debate. Columbia University Press, New York*, 17-29.

Moritz, C. (1994). Defining 'evolutionarily significant units' for conservation. *Trends in ecology & evolution*, *9*(10), 373-375.

Casacci, L. P., Barbero, F., & Balletto, E. (2014). The "Evolutionarily Significant Unit" concept and its applicability in biological conservation. *Italian Journal of Zoology*, *81*(2), 182-193.

Barton, N. H. (2001). The role of hybridization in evolution. *Molecular ecology*, *10*(3), 551-568.

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J., Bierne, N., *et al.* (2013). Hybridization and speciation. *Journal of evolutionary biology*, *26*(2), 229-246.

Chan, W. Y., Hoffmann, A. A., & van Oppen, M. J. (2019).

Hybridization as a conservation management tool. *Conservation Letters*, 12(5), e12652.

Allendorf, F. W., Leary, R. F., Spruell, P., & Wenburg, J. K. (2001). The problems with hybrids: setting conservation guidelines. *Trends in ecology & evolution*, 16(11), 613-622

Bohling, J. H. (2016). Strategies to address the conservation threats posed by hybridization and genetic introgression. *Biological Conservation*, 203, 321-327

Buggs, R. J. A. (2007). Empirical study of hybrid zone movement. *Heredity*, 99(3), 301-312.

Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hübner, S., *et al.* (2016). What we still don't know about invasion genetics. *Invasion genetics: The baker and stebbins legacy*, 346-370

Ficetola, G. F., Bonin, A., & Miaud, C. (2008). Population genetics reveals origin and number of founders in a biological invasion. *Molecular Ecology*, 17(3), 773-782

Lynch, M., Ackerman, M. S., Gout, J. F., Long, H., Sung, W., Thomas, W. K., & Foster, P. L. (2016). Genetic drift, selection and the evolution of the mutation rate. *Nature Reviews Genetics*, 17(11), 704-714.

Petrie, M., Doums, C., & Møller, A. P. (1998). The degree of extra-pair paternity increases with genetic variability. *Proceedings of the National Academy of Sciences*, 95(16), 9390-9395

Brouwer, L., & Griffith, S. C. (2019). Extra-pair paternity in birds. *Molecular ecology*, 28(22), 4864-4882.

Heckel, G., & Von Helversen, O. (2003). Genetic mating system and the significance of harem associations in the bat Saccopteryx bilineata. *Molecular Ecology*, 12(1), 219-227.

Nunney, L. (1993). The influence of mating system and overlapping generations on effective population size. *Evolution*, 47(5), 1329-1341.

Charpentier, M., Hossaert-McKey, M., Wickings, E. J., & Peignot, P. (2005). Consequences of a one-male harem reproductive system and inbreeding in a captive group of Cercopithecus solatus. *International Journal of Primatology*, 26, 697-710.

Avise, J. C. (2009). Phylogeography: retrospect and prospect. *Journal of biogeography*, 36(1), 3-15.

Knowles, L. L. (2009). Statistical phylogeography. *Annual Review of Ecology, Evolution, and Systematics*, 40, 593-612.

Beheregaray, L. B. (2008). Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Molecular ecology*, 17(17), 3754-3774

Arbogast, B. S., & Kenagy, G. J. (2001). Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, 28(7), 819-825

DeSalle, R., & Amato, G. (2009). The expansion of conservation genetics. *Conservation genetics in the age of genomics*, 1-24

Frankham, R., Briscoe, D. A., & Ballou, J. D. (2002). *Introduction to conservation genetics*. Cambridge university press.

Ouborg, N. J., Pertoldi, C., Loeschcke, V., Bijlsma, R. K., & Hedrick, P. W. (2010). Conservation genetics in transition to conservation genomics. *Trends in genetics*, 26(4), 177-187.

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, 30(1), 62-82.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, 17(2), 81-92.

Shapiro, B. (2017). Pathways to de-extinction: how close can we get to resurrection of an extinct species?. *Functional Ecology*, 31(5), 996-1002.

Beaumont, M. A. (2005). Adaptation and speciation: what can Fst tell us?. *Trends in ecology & evolution*, 20(8), 435-440

Williams, S. E., & Hoffman, E. A. (2009). Minimizing genetic adaptation in captive breeding programs: a review. *Biological conservation*, 142(11), 2388-2400.

Attard, C. R. M., Möller, L. M., Sasaki, M., Hammer, M. P., Bice, C. M., Brauer, C. J., ... & Beheregaray, L. B. (2016). A novel holistic framework for genetic-based captive-breeding and reintroduction programs. *Conservation Biology*, 30(5), 1060-1069.

Bowen, B. W., Gaither, M. R., DiBattista, J. D., Iacchei, M., Andrews, K. R., Grant, W. S., *et al.* (2016). Comparative phylogeography of the ocean planet. *Proceedings of the National Academy of Sciences*, 113(29), 7962-7969.

Blomqvist, D., Pauliny, A., Larsson, M., & Flodin, L. Å. (2010). Trapped in the extinction vortex? Strong genetic effects in a declining vertebrate population. *BMC Evolutionary Biology*, 10, 1-9.

Appendix I: Required Software

The following software is required for this course, specifically the laboratory section (BIO 317L). They are all free to download and compatible for both Windows and Macintosh operating systems. If you need a loaner laptop, please see me ASAP.

Jalview

Jalview is a free cross-platform program for multiple sequence alignment editing, visualisation and analysis. Use it to align, view and edit sequence alignments, analyse them with phylogenetic trees and principal components analysis (PCA) plots and explore molecular structures and annotation. We will be using it for sequence alignment and editing in lab.

Download Link: https://www.jalview.org/download/

UGENE

UGENE is a free open-source bioinformatics software for Windows, macOS and Linux. It will be especially important in lab for us to view our .abi files to check for sequence quality.

Download Link: http://ugene.net/download.html

PopART

PopART (Population Analysis with Reticulate Trees) is free population genetics software that was developed as part of the

Allan Wilson Centre Imaging Evolution Initiative. We will be using it specifically to build haplotype networks in lab.

Download Link: https://popart.maths.otago.ac.nz/download/

Python

We will be using the Python programming language to perform the majority of our sequence analyses (tree construction, AMOVAs, neutrality tests, diversity analyses and data visualization). **You do not need to have any experience in coding – all scripts will be provided in a "plug and play" format for you to use.** In addition, we will have a crash-course on the first week of class to get you comfortable with using the command line, and understanding the bare basics of coding (language-specific syntax, code annotation, etc.).

There are several Python environments you can use to carry out your analysis. If you have never written code before, I strongly suggest **THONNY** which can be downloaded here. If you have taken BIO 241, then you should already have the **ANACONDA** platform downloaded on your laptop and you can continue to use that for this course. If you no longer have that platform, you can download it here. A video tutorial on how to install both of these platforms has been posted to this course's Canvas site. If you run into any issues downloading and installing them on your computers, please see me during student hours.

Appendix II: Syllabus

BIO 317: Molecular Ecology & Evolution
Spring 2024
Tentative Lecture Schedule

LECTURE TOPIC	**WEEK**	**ASSIGNED READING**
Introduction	1	OER – Part I
Review of Molecular Biology	2	OER – Part II
Molecular Markers	3 & 4	OER – Part III
Species & Populations	5 & 6	OER – Part IV
Population Genetics	7, 8, 9 & 10	OER- Part V
Phylogeography	11 & 12	OER – Part VI
Behavioral Ecology	13	OER – Part VIII
Conservation Genetics	14 & 15	OER – Part VII
Genomics & Molecular Ecology	16	OER – Part IX

Instructor: Andrew A. Davinack, Ph.D.
office: Mars Center for Science & Technology, room 1134
lab: Mars Center for Science and Technology, room 1118
phone: 508-286-3944
e-mail: davinack_drew@wheatoncollege.edu

MEETING TIMES

Class times:

Lecture – Tuesdays & Thursdays: 9:30 am – 10:50 am, Discovery Center 1313

Lab – Thursdays: 2:00 pm – 4:50 pm, Mars Science Center 2141

Student Hours:

Mondays 9:00 am – noon

Tuesdays: 1:30 pm – 2:30 pm (or by appointment, if none of these times work)

Student hours are designated times set aside for you to come and have a one-on-one discussion with me about anything pertaining to the class, your career goals, research or internship opportunities, coursework, classes or anything else on your mind! Furthermore, if you score a C or below on any of the exams, it is strongly advised that you meet with me within a week of receiving your grade. You can attend student hours by either dropping in to my office (SC 1134) or if you would like to lock in a specific date and time, you can use the following link https://calendly.com/drewdavinack/office-hours (url is also available on Canvas). You can also email me directly to set up a virtual meeting if none of the above days and times work for your schedule.

Course Description: This comprehensive course in Molecular Ecology integrates both theoretical concepts and practical laboratory skills, designed for advanced undergraduates in biology or related fields. It offers an in-depth exploration of molecular tools and techniques used to understand ecological processes and patterns. The course is divided into lecture and laboratory components, each complementing the other to provide a holistic understanding of the subject. The lecture series are interspersed with discussions on scholarly articles and practical applications of molecular ecology in research and conservation. The laboratory component will be a mix of wet lab and dry lab practicals including key tools every molecular biologist needs to have in their repertoire including: DNA extraction and PCR skills, gel electrophoresis, programming skills and data analytical skills. The lecture component will culminate in the preparation and submission of an NSF-style research proposal where you must develop a novel ecological project. The lab component will culminate in the preparation of a manuscript for publication in the journal *Molecular Ecology* based on the data collected and analyzed in lab. Compass curriculum attribute: Natural Science.

Learning Outcomes for Lecture: At the end of the lecture component of this course, students will be able to:

- Explain the basics of molecular genetics as they apply to ecological studies.
- Analyze and interpret molecular data in the context of species and population studies.
- Critically evaluate scientific literature related to molecular ecology topics.
- Explain the principles of phylogeography and its significance in understanding species distribution and evolution.
- Explain the role of molecular techniques in conservation genetics and their application in biodiversity preservation.
- Develop the ability to formulate research questions and hypotheses in molecular ecology.

Recommended text and materials for lecture:

No textbook required for this course. We will be using an Open Educational Resource (OER) textbook specifically designed for BIO 317.

A **laptop** i.e., a fully functional computer with its own hard disk storage unit and full OS (e.g., Windows, Mac or Linux). NO Chromebooks, tablets or Ideapads! None of these machines have good performance capabilities for the type of analyses you will be performing. *If you need a loaner laptop for the semester, please contact me* ASAP. The software you will need for this course will be used in both lecture and lab (though primarily in lab) and is all free.

Lecture Grade Breakdown

Exams	15%
Article Discussions	10%
In-Class Participation	10%
NSF-Research Proposal	25%
LECTURE TOTAL	**60%**

Lecture Components

Exams: All exams will be open book and open note, and must be completed in class except for exam #2 which will be completed and submitted online. If you have accommodations (e.g., time and a half/ double time) please email me with your official accommodations ASAP and we can arrange an alternative time for you to take your exam. Questions will be 50% multiple-choice/Mix & Match/ True or False and 50% short answer. These questions are designed to test your understanding of the basics of molecular ecology and will be based on articles covered in class in addition to material presented during lecture.

Article Discussions: For article discussions, we will be using a technique to encourage deeper reading/discussion (the material below is modified from the technique developed by Parrott & Cherry, 2011). This technique assigns roles to each group member to support a more in-depth examination of the research and shared responsibility for analysis of the work with your group.

Below is a summary of the roles and your role assignment for each paper is on the next page. For each paper, you will fulfill one of these roles and come prepared to support your group's discussion in that capacity (as defined below). Your groups will be pre-arranged:

Discussion Leader and Concept Connector. Your job is to develop *at least three possible discussion questions* that you can discuss in groups to help everyone understand the main points of the assigned reading. Your task is to help people understand the theory in the papers, the purpose/objectives of the study, the methods and the main findings/conclusions. In doing so you should be facilitating the group discussion (NOT presenting the findings to the group). In addition, you should help everyone make connections to other important concepts we have covered in this course, you have learned about in previous courses or through your own experience. *You should have at least three discussion questions* *prepared with your own brief answers. In addition, you should have least two connections prepared to facilitate discussion of concepts.*

Passage Master. Your job is to *locate a few special passages that are important in the reading assignment.* These may give key information, back up the information given, or summarize the author's key points. They might also be passages that inspired your interest, are particularly well written, or might be controversial or contradictory with other passages or other information learned in class. *You will need to note at least two important passages per reading, including a summary of the passage in everyday terminology (in other words, how you would explain the passage to your roommate), and an explanation of why you think the passage is important.*

Devil's Advocate. *Your job is to challenge the ideas/findings of the article by developing a list of critical, thoughtful questions that might be raised in terms of issues resulting from experimental design, lack of controls, conclusions that go beyond the findings, etc. If it is a review paper, your tasks will be developing a list of critiques that challenges the conclusions of the article. You should prepare at least two challenging questions or issues you have with the paper and its findings/evaluations.*

****Reporter.*** *Your job is to prepare a one-page report that summarizes the discussion within your group and amongst the class.* ***The report should be emailed to me within 72 hours following the discussion.*** *Failure of the designated reporter to submit their report will result in a one-point deduction for the ENTIRE group every day that it is late.*

**Any of the three roles (Discussion Leader, Devil's Advocate, Passage Master) may serve as reporter and this will be pre-assigned prior to the discussion.*

Class Participation: As this is an upper-level course, a lot of the work must be completed on your own. This includes assignments, homework and readings. Class participation will involve attendance, discussion of material in class and in-class assignments. Please note that some points will ONLY be awarded in class and therefore it is in your best interest to attend all class sessions. Should a situation arise where you will be unable to attend due to a legitimate reason, you must contact me at least 3-5 days before that particular class. Aside from this, **no in-class activities can be made up (this is non-negotiable).**

Research Grant Proposal: As a capstone project for BIO 317, you will be required to develop an authentic research proposal which follows the United States' National Science Foundation (NSF) template which emphasizes the "Intellectual Merit" and "Broader Impacts" of a proposed research plan. For this course you will be following the NSF's Division of Environmental Biology (DEB) directorate. DEB encourages research that elucidates fundamental principles that identify and explain the unity and diversity of life and its interactions with the environment over space and time. Research may incorporate field, laboratory or collection-based approaches; observational or manipulative studies; synthesis activities; phylogenetic discovery projects; or theoretical approaches involving analytical, statistical, or computational modeling. You may focus your proposed research on one or multiple species from either an aquatic (marine, freshwater, estuarine, etc.) or a terrestrial system. DEB also encourages interdisciplinary proposals that cross conceptual boundaries and integrate over levels of biological organization or across multiple spatial and temporal scale. ***You may work as a group (no more than three students per group) to put together a joint proposal if you wish.*** Specific guidelines regarding the detailed format of your proposal will be provided separately.

Attendance Policy for Lecture: Your attendance at regular lecture meetings is an essential part of our work together. While attendance is not being tracked directly, missed in-class assignments will reflect your absences in this course. As a result, it is your responsibility to (1) let me know you will be absent in advance so you complete the in-class assignment upon your return without penalty and (2) identify/learn the material covered in class (we can meet during office to review the material that you missed). *Per Wheaton's* *policy, I have to report students that miss two classes without prior communication to me.* If you anticipate multiple absences due to any personal circumstances, please reach out to me, Susan Friedman in Advising (friedman_susan@wheatoncollege.edu), or one of the support resources listed later in this document, so we can provide you with the necessary support.

Notification of Canvas usage: All exam and project grades will be posted to the course's canvas page once they have been graded. In addition, all handouts, lecture slides, relevant articles and videos will be posted under the weekly modules.

BIO 317L – Molecular Ecology & Evolution Laboratory
Tentative Laboratory Schedule
Spring 2024
WL: Wet Lab Practical, DL: Dry Lab Practical

Lab Practical	**Week**
Lab Orientation & Pipetting Activity	1
WL: Mussel gDNA extraction & Quality Check	2 & 3
DL: Introduction to Python	4
DL: Navigating GenBank and BoLD database	5
WL: PCR – Round 1	6
WL: Gel Electrophoresis & Purification	7
DL: PCR Troubleshooting	8

WL: Gel Electrophoresis & Purification	9
DL: Sequence Editing & Alignment	10
DL: Phylogenetic Tree Construction & Population Genetic Analyses	11
DL: Haplotype Network Construction	12
Lab Debriefing Session	13
Manuscript write-up (in-class): Methods & Results	14

Manuscript write-up (in-class): Introduction & Discussion	15
NO LAB (Manuscript Submission Due)	16

Learning Outcomes for Lab: At the end of the lab component of this course, students will be able to:

- Work collaboratively in a lab setting to conduct experiments and analyze data.
- Develop a moderate level of competency in gDNA extraction, PCR and gel electrophoresis.
- Develop proficiency in bioinformatics tools, including Python programming for DNA sequence handling and analysis.
- Prepare a scientific manuscript from inception to submission for publication.
- Conduct population genetic analyses using computational tools

Laboratory Grade Breakdown

Lab Notebooks	10%
Research Manuscript	15%
Bioinformatics Practicals	10%
Pre-lab quizzes	5%
LABORATORY TOTAL	**40%**

Lab Components

Lab Notebooks: As part of the laboratory component of this course, you will be required to maintain an organized lab notebook with the genetic data that you collected. The laboratory project for this course will be written up for publication in a peer-reviewed scientific journal which means that your data must be clearly and concisely recorded in a very consistent manner. While you will be working collaboratively on the lab project, each student must maintain their own lab notebook. The guidelines for organizing your notebook are available on Canvas. Your notebook will be checked three times during the semester (specific dates are on the lab schedule).

Research Manuscript: The data collected by each group will be collated towards the end of the semester and each lab group will be required to write a standard scientific manuscript to be submitted to the journal *Molecular Ecology*. As you group, you will have to decide amongst yourself who will be responsible for each component of the manuscript (Introduction, Methods, Results and Discussion) and you will all have to agree on a final version that must be submitted by the due date (see lab schedule).

Bioinformatics Practicals: During "dry lab" days you will be working directly with DNA sequence data on your computers. There will several practicals completed in lab where you will be required to exhibit your proficiency at retrieving batch sequences from the GenBank database, processing them (editing & aligning) using specific software and carrying out various types of phylogenetic and population genetic analyses using Python.

Pre-lab quizzes: There will be a total of four pre-lab quizzes throughout the semester that are front-loaded in the first half of the semester. These quizzes are short (10 questions) and are based on the theory behind the laboratory techniques/data analyses we will be covering that day. You will need to read the material posted in the lab module prior to taking the quiz. You will have unlimited attempts to complete the quiz but it will be due before lab begins on Thursday.

Attendance Policy for Laboratory: Labs are an essential component of this course because you will be applying your knowledge gained in lecture to carry out a real molecular ecology project. This project will be collaborative but you will be responsible for individual lab notebooks. Because of the way the lab has been set up along with the perishable and costly nature of the reagenets being used, **there are no opportunities to make up any of the wet-lab sessions.** It is the course policy that if a student has more than three absences from BIO 317 lab over the course of the semester, that student may not be able to pass the course. If you know in advance that you will have to miss a lab, please reach out to me as soon as possible. If you miss a lab due to illness, you should also contact me to let me know.

Course Policies

Honor Code: You are expected to uphold the Wheaton Honor Code throughout the course. It will be clearly specified in the course which assignments are collaborative and which must be done on an individual basis. If in doubt, ask the instructor.

DEI Statement: At Wheaton I committed to creating a learning environment that supports a diversity of thoughts, perspectives and experiences and honors your identities (including race, gender, class, sexuality, religion, ability, etc.). To accomplish this if you have a name and/or set of pronouns that differ from those that appear in your official Wheaton records, please let me know. If you have any concerns during the course, please do not hesitate to come and talk with me and you can also submit anonymous feedback if you wish.

Accessibility Accommodations: Wheaton is committed to ensuring equitable access to programs and services and to prohibit discrimination in the recruitment, admission, and education of students with disabilities. Individuals with disabilities requiring accommodation or information on accessibility should contact Autumn Grant – Associate Director for Accessibility Services at the Filene Center for Academic Advising and Career Services:

accessibility@wheatoncollege.edu or (508) 286-8215. If you already have accommodations in place, please provide the instructor with documentation during the first two weeks of the semester so that timely and appropriate arrangements can be made. All lecture slides will be published on Canvas after each class period.

Counseling and Mental Health: Your wellness is important to me. If you would benefit from mental health support, please call the Counseling Center at 508-286-3905 at any time of day or night, to schedule an appointment, or follow the voicemail prompts to be connected to a clinician.

Student Recording of Class Sessions: Wheaton College prohibits recording (as defined above) and transmission of activities (e.g. lectures, discussions) that occur as part of a classroom session by a student unless: (1) Accessibility Services has authorized recording as an academic accommodation for a qualified student with a disability and the student has notified the instructor of that authorization by presenting their accommodation letter; or (2) permission from the course instructor has otherwise been granted. In these circumstances, all students participating in the course as well as any guest speakers will be informed that audio/video recording may occur. If a student is given authorization to record any portion of a classroom session, that student understands and agrees that the recording is for the sole use of the individual student and may not be reproduced, sold, posted online, or otherwise distributed. A student does not have permission to reproduce or post the information on any social media (e.g., YouTube, FaceBook, etc.), or other public or private forum that would infringe on the privacy rights of others represented in the recording. Public distribution of such materials may constitute copyright infringement in violation of federal or state law, or College policy. Violation of this policy may subject a student to disciplinary action under the College's Community Standards or other relevant policies.

Wheaton Student Support & Wellness Resources:

- The Counseling Center is the confidential and FREE mental health resource on campus for all students. To learn about services, check out the website, or give the office a call at 508-286-3905. Even when the Counseling Center is closed, or staff are unavailable, After Hours Mental Health Support is available by calling the front desk 508-286-3905 and following voicemail prompts to be connected to a clinician (24/7, available in languages other than English, and accessible from anywhere you are in the world).
- The Filene Center strives to support your learning pathway by fostering successful academic, career, and personal development. The academic advising staff will work collaboratively with you, faculty and campus resources to ensure that you have the access and guidance to become a confident and reflective learner at Wheaton and beyond. Contact us at advising@wheatoncollege.edu.
- Many other offices on campus can also help support the holistic wellness of students. For students who identify as low-income, first-gen, LGBTQ+, or have a faith or spiritual practice they adhere to, the Center for Social Justice and Community Impact and Center for Religious and Spiritual Life (the Base) are good places for support and engagement. The Marshall Center for Intercultural Learning supports BIPOC students and those working towards breaking down barriers across difference, and the Center for Global Education supports international students, and students seeking educational opportunities abroad. We encourage you to reach out to any and all of these offices for support.
- Health Services through Norton Medical Center is available to support students with a variety of physical health needs including specialty support for GYN and STI care. Contact the office at 508-286-4500 to make an appointment for care.

There is no copay for visits and most services are free, with select procedures and labs billed to insurance.

Policy on Plagiarism (adopted from Prof. Robert Morris)

Rule: Cite properly all those whom you draw ideas, words, or work

Plagiarism is the unacknowledged use of another's ideas or words with the intentional or unintentional result of presenting those ideas or words as your own. In the course, all written assignments are designed for you to explore your own ideas and thoughts however in cases where you use explicitly use published material to supplement these ideas, you will be required to cite the source of the study. Try to avoid using direct quotations and instead paraphrase as much as you can in your own words. This course adopts Wheaton College's policy on plagiarism and affirms that "authors who fail to acknowledge their sources are, at the very least, guilty of being ignorant about the ethics governing the wider community of scholars; at the worst, they are guilty of blatant dishonesty. In any case, plagiarism in any form constitutes a serious violation of the most basic principles of scholarship and cannot be tolerated".

In the course, examples of plagiarism include:

1. Passing of written material that is not your own: this refers to the actual *wording*. Copying someone's paragraph or even a few lines of sentences word for word is considered as plagiarism even if you cite it. You must paraphrase all material coming from outside sources.
2. For the final project of the semester, you will be required to develop your own host-parasite model system which will most likely require you to mine images from online sources. Any images used in assignments must have the source either in the figure legend or provide a separate "Reference" slide with the list of sources for the images used.

Consequences of plagiarism in this course: The FIRST instance of plagiarism of any form on any portion of any assignment will result in no credit award for the entire assignment and ALL MEMBERS of the group will be penalized. The SECOND instance of plagiarism of any form on any portion of any assignment will result in no credit award for the entire assignment and ALL MEMBERS of the group will be penalized. The THIRD instance of plagiarism of any form on any portion of any assignment will result in a final grade of F for the course.

The best way to avoid plagiarism altogether is to just ask! Either meet with me or send me a draft of your work and I can give you instant answer. You will not be penalized for drafts, only final versions.

Policy on ChatGPT Usage: You may use ChatGPT to brainstorm ideas for your grant proposal. However, keep in mind that ChatGPT occasionally "hallucinates" resulting in inaccurate information, and worse yet, conjures up false references to support its conclusions. Therefore, you will need to investigate any ideas that ChatGPT comes up with yourself. You will be held liable if you proposed a project and included incorrect information and supplied references that do not exist.

ChatGPT Usage Acknowledgement: If you do choose to use ChatGPT to brainstorm ideas for your proposals, this needs to be stated in a special *ChatGPT Disclaimer Section* that should appear at the end of your references.

.