SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF BIO AND CHEMICAL ENGINEERING

DEPARTMENT OF BIOINFORMATICS

UNIT – 1- SBIA1302 – Evolutionary Biology

UNIT I

Evolution is change in the heritable characteristics of biological populations over successive generations. These characteristics are the expressions of genes that are passed on from parent to offspring during reproduction. Different characteristics tend to exist within any given population as a result of mutation, genetic recombination and other sources of genetic variation.[3] Evolution occurs when evolutionary processes such as natural selection (including sexual selection) and genetic drift act on this variation, resulting in certain characteristics becoming more common or rare within a population.[4] It is this process of evolution that has given rise to biodiversity at every level of biological organisation, including the levels of species, individual organisms and molecules.

The scientific theory of evolution by natural selection was conceived independently by Charles Darwin and Alfred Russel Wallace in the mid-19th century and was set out in detail in Darwin's book *On the Origin of Species*. Evolution by natural selection was first demonstrated by the observation that more offspring are often produced than can possibly survive. This is followed by three observable facts about living organisms: (1) traits vary among individuals with respect to their morphology, physiology and behaviour (phenotypic variation), (2) different traits confer different rates of survival and reproduction (differential fitness) and (3) traits can be passed from generation to generation (heritability of fitness). Thus, in successive generations members of a population are more likely to be replaced by the progenies of parents with favourable characteristics that have enabled them to survive and reproduce in their respective environments. In the early 20th century, other competing ideas of evolution such as mutationism and orthogenesis were refuted as the modern synthesis reconciled Darwinian evolution with classical genetics, which established adaptive evolution as being caused by natural selection acting on Mendelian genetic variation.[9]

All life on Earth shares a last universal common ancestor (LUCA)[10][11][12] that lived approximately 3.5–3.8 billion years ago.[13] The fossil record includes a progression from early biogenic graphite,[14] to microbial mat fossils,[15][16][17] to fossilised multicellular organisms. Existing patterns of biodiversity have been shaped by repeated formations of new species (speciation), changes within species (anagenesis) and loss of species (extinction) throughout the evolutionary history of life on Earth.[18] Morphological and biochemical traits are more

similar among species that share a more recent common ancestor, and can be used to reconstruct phylogenetic trees.[19][20]

Evolutionary biologists have continued to study various aspects of evolution by forming and testing hypotheses as well as constructing theories based on evidence from the field or laboratory and on data generated by the methods of mathematical and theoretical biology. Their discoveries have influenced not just the development of biology but numerous other scientific and industrial fields, including agriculture, medicine and computer science

Different types of evolution

Convergent evolution

- When the same adaptations evolve independently, under similar selection pressures.
- For example, flying insects, birds and bats have all evolved the ability to fly, but independently of each other.

Co-evolution

- When two species or groups of species have evolved alongside each other where one adapts to changes in the other.
- For example, flowering plants and pollinating insects such as bees.

Adaptive radiation

- When a species splits into a number of new forms when a change in the environment makes new resources available or creates new environmental challenges.
- For example, finches on the Galapagos Islands have developed different shaped beaks to take advantage of the different kinds of food available on different islands.

History of evolutionary thought

Evolutionary thought, the recognition that species change over time and the perceived understanding of how such processes work, has roots in antiquity—in the ideas of the ancient Greeks, Romans, and Chinese as well as in medieval Islamic science. With the beginnings of modern biological taxonomy in the late 17th century, two opposed ideas influenced Western biological thinking: essentialism, the belief that every species has essential characteristics that are unalterable, a concept which had developed from medieval Aristotelian metaphysics, and that fit well with natural theology; and the development of the new anti-Aristotelian approach to modern science: as the Enlightenment progressed,

evolutionary cosmology and the mechanical philosophy spread from the physical sciences to natural history. Naturalists began to focus on the variability of species; the emergence of paleontology with the concept of extinction further undermined static views of nature. In the early 19th century Jean-Baptiste Lamarck (1744–1829) proposed his theory of the transmutation of species, the first fully formed theory of evolution.

In 1858 Charles Darwin and Alfred Russel Wallace published a new evolutionary theory, explained in detail in Darwin's *On the Origin of Species* (1859). Unlike Lamarck, Darwin proposed common descent and a branching tree of life, meaning that two very different species could share a common ancestor. Darwin based his theory on the idea of natural selection: it synthesized a broad range of evidence from animal husbandry, biogeography, geology, morphology, and embryology. Debate over Darwin's work led to the rapid acceptance of the general concept of evolution, but the specific mechanism he proposed, natural selection, was not widely accepted until it was revived by developments in biology that occurred during the 1920s through the 1940s. Before that time most biologists regarded other factors as responsible for evolution. Alternatives to natural selection suggested during "the eclipse of Darwinism" (c. 1880 to 1920) included inheritance of acquired characteristics (neo-Lamarckism), an innate drive for change (orthogenesis), and sudden large mutations (saltationism). Mendelian genetics, a series of 19th-century experiments with pea plant variations rediscovered in 1900, was integrated with natural selection by Ronald Fisher, J. B. S. Haldane, and Sewall Wright during the 1910s to 1930s, and resulted in the founding of the new discipline of population genetics. During the 1930s and 1940s population genetics became integrated with other biological fields, resulting in a widely applicable theory of evolution that encompassed much of biology—the modern synthesis.

Following the establishment of evolutionary biology, studies of mutation and genetic diversity in natural populations, combined with biogeography and systematics, led to sophisticated mathematical and causal models of evolution. Paleontology and comparative anatomy allowed more detailed reconstructions of the evolutionary history of life. After the rise of molecular genetics in the 1950s, the field of molecular evolution developed, based on protein sequences and immunological tests, and later incorporating RNA and DNA studies. The gene-centered view of evolution rose to prominence in the 1960s, followed by the neutral theory of molecular evolution, sparking debates over adaptationism, the unit of selection, and the relative

importance of genetic drift versus natural selection as causes of evolution.[2] In the late 20th-century, DNA sequencing led to molecular phylogenetics and the reorganization of the tree of life into the three-domain system by Carl Woese. In addition, the newly recognized factors of symbiogenesis and horizontal gene transfer introduced yet more complexity into evolutionary theory. Discoveries in evolutionary biology have made a significant impact not just within the traditional branches of biology, but also in other academic disciplines (for example: anthropology and psychology) and on society at large

Evidence for Evolution

In his book, Darwin offered several pieces of evidence that supported evolution. He attempted to convince the scientific community of the validity of his theory.

Paleontology

One piece of evidence offered by Darwin is found in the science of paleontology. **Paleontology** deals with locating, cataloging, and interpreting the life forms that existed in past millennia. It is the study of fossils—the bones, shells, teeth, and other remains of organisms, or evidence of ancient organisms, that have survived over eons of time.

Paleontology supports the theory of evolution because it shows a descent of modern organisms from common ancestors. Paleontology indicates that fewer kinds of organisms existed in past eras, and the organisms were probably less complex. As paleontologists descend deeper and deeper into layers of rock, the variety and complexity of fossils decrease. The fossils from the uppermost rock layers are most like current forms. Fossils from the deeper layers are the ancestors of modern forms.

Comparative anatomy

More evidence for evolution is offered by **comparative anatomy** (see Figure 12-1). As Darwin pointed out, the forelimbs of such animals as humans, porpoises, bats, and other creatures are strikingly similar, even though the forelimbs are used for different purposes (that is, lifting, swimming, and flying, respectively). Darwin proposed that similar forelimbs have similar origins, and he used this evidence to point to a common ancestor for modern forms. He suggested that various modifications are nothing more than adaptations to the special needs of modern organisms.

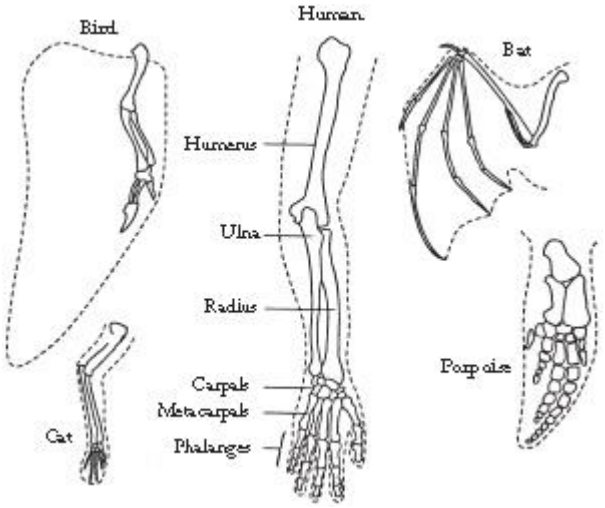


Figure 12-1 The forelimbs of a human and four animals showing the similarity in construction. This similarity was offered by Darwin as evidence that evolution has occurred.

Darwin also observed that animals have structures they do not use. Often these structures degenerate and become undersized compared with similar organs in other organisms. The useless organs or body parts are called *vestigial organs.* In humans, they include the appendix, the fused tail vertebrae, the wisdom teeth, and muscles that move the ears and nose. Darwin maintained that vestigial organs may represent structures that have not quite disappeared. Perhaps an environmental change made the organ unnecessary for survival, and the organ gradually became nonfunctional and reduced in size. For example, the appendix in human ancestors may have been an organ for digesting certain foods, and the coccyx at the tip of the vertebral column may be the remnants of a tail possessed by an ancient ancestor.

Embryology

Darwin noted the striking similarity among embryos of complex animals such as humans, chickens, frogs, reptiles, and fish. He wrote that the uniformity is evidence for evolution. He pointed out that human embryos pass through a number of embryonic stages inherited from their ancestors because they have inherited the developmental mechanisms from a common ancestor. These mechanisms are modified in a way that is unique to an organism's way of life.

The similarities in comparative embryology are also evident in the early stages of development. For example, fish, bird, rabbit, and human embryos are similar in appearance in the early stages. They all have gill slits, a two-chambered heart, and a tail with muscles to move it. Later on, as the embryos grow and develop, they become less and less similar. The branch of biology that focuses on embryos and their development is called **embryology.**

Comparative biochemistry

Although the biochemistry of organisms was not well known in Darwin's time, modern biochemistry indicates there is a biochemical similarity in all living things. This comparison of biochemical processes with ancient species is called **comparative biochemistry.** For example, the same mechanisms for trapping and transforming energy and for building proteins from amino acids are nearly identical in almost all living systems. DNA and RNA are the mechanisms for inheritance and gene activity in all living organisms. The structure of the genetic code is almost identical in all living things. This uniformity in biochemical organization underlies the diversity of living things and points to evolutionary relationships.

Domestic breeding

From observing the domestic breeding experiments of animal and plant scientists, Darwin developed an idea about how evolution takes place. **Domestic breeding** brings about new forms that differ from ancestral stock. For example, pigeon fanciers have developed many varieties of pigeons through domestic breeding experiments. In effect, evolution has taken place under the guidance of human hands. The development of new agricultural crops by farmers and botanists provides more evidence for directed evolution.

Geographic distribution

Darwin was particularly interested in the life forms of the Galapagos Islands. He noticed how many of the birds and other animals on the islands were found only there. The finches were particularly puzzling; Darwin found 13 species of finches not found anywhere else in the world, as far as he knew. He concluded that the finches had evolved from a common ancestral group that probably reached the island many generations earlier. In the isolation of the Galapagos Islands, the original finches had probably evolved into the 13 species.

The **geographic distribution** of species in geographic areas can help to explain evolution. For instance, alligators are located only in certain regions of the world, presumably because they have evolved in those regions. The islands of Australia and New Zealand have populations of animals found nowhere else in the world because of their isolated environments.

From DNA to phenotypes

Genetic information is passed between generations in most organisms by DNA. Variation among individuals of the sequence of their DNA is the raw material of evolution. DNA codes for phenotype by sequences specifying proteins and RNAs and by regulatory elements that control

when and by how much of each is made. Variation in DNA sequence can translate into variation in phenotype, which may cause fitness differences among individuals on which natural selection can act.

The Central Dogma

The so-called central dogma of molecular biology is that, for most organisms, DNA is the genetic material that transmits information between generations, and mRNA is copied from genes on the DNA to carry that information to parts of the cell to produce proteins, which do most of the work of the cell. This "dogma" is largely true, but all its steps can have exceptions. Some organisms use RNA as the genetic material that is passed between generations to convey the information encoded therein. Many genes in DNA code for RNAs that themselves have physical function (and do not code for proteins at all). Some information is transmitted between generations by modifications of the DNA, such as methylation, such that genetic transmission of information to subsequent generations is not entirely composed of data in the sequence of the DNA. Parents also modify the environment of their offspring in many cases, affecting the nature of subsequent generations in nongenetic ways. Such enriching exceptions aside, however, the central dogma explains much of the nature of how information is transmitted and used by living organisms.

Variation in DNA sequence

Evolution by natural selection requires heritable differences betweenindividuals, which occur when different individuals carry alleles that produce different phenotypes. A population that contains more than one allele can have genetic variation, meaning that all individuals are not genetically identical. This variation at the DNA level can mean that there is variation at the protein level, which may translate into variation in phenotype, which can translate into variation among individuals in fitness. Only when fitness varies among individuals, and when that fitness variation has a genetic basis, can a population evolve by natural selection. Evolutionary biologists use a variety of techniques to study this genetic variation, ranging from examination at the DNA sequence level to investigation of the genetic basis of differences in phenotypes between individuals. DNA sequencing now allows relatively inexpensive reading of the genome itself; sequencing all or part of the genome in multiple individuals within a species allows the genetic variability of that species to be measured. DNA sequences at a locus can differ in a variety of ways.

DESCRIPTIONS OF GENETIC VARIATION The most basic unit in the description of the genetic variation is the allele frequency, which is the fraction of all alleles at a locus in the population of interest that have a particular sequence. (Population geneticists often, but not always, use the letters p or q to denote the allele frequency of a particular gene under study. Every individual of a diploid species carries two copies of each autosomal locus, one from each of its parents. The genotype of the individual at that locus therefore must be described by keeping track of both alleles that it carries. If there are only two allelesin the population, there are three possible genotypes: individuals that carry two copies of one allele, two of the other allele, or one of each. If there are more alleles at a locus, then the number of possible genotypes is higher. If both copies at a locus are the same allele, then we say that the individual is homozygous at that locus; if the two copies differ, it is heterozygous. Genotype frequencies can be predicted from the allele frequencies, provided a long list of evolutionary assumptions are true. If we assume that no selection, mutation, or migration affects the frequencies of alleles at a locus, that the population being studied is extremely large, and all individuals in the population are equally likely to mate with all other individuals, then the genotype frequencies can be predicted from the HardyWeinberg equilibrium (HWE)

Phylogenetic trees

Ancestral State Reconstruction.

A procedure that uses a phylogeny as well as character data from extant species to infer likely ancestral character states.

Character State.

Alternative states for a given biological character (e.g., brown, blue, hazel are character states of the character eye color; A, C, G, or T are character states of the character corresponding to base position 12 in the actin gene).

Chronogram.

A phylogenetic tree with branch lengths scaled to represent time (related to phylogram, a tree in which branch lengths are scaled to the amount of character evolution).

Cladogram.

An unscaled phylogenetic tree that shows relationships among organisms, but in which branch lengths are meaningless. The key information retained in a cladogram istopology, which refers to the composition of clades/monophyletic groups and how they are related to one another.

Derived.

Generally the opposite of ancestral; a derived character state is one that has evolved recently relative to an ancestral character state (e.g., scaly skin is an ancestral character state for reptiles, whereas feathers in avian reptiles represent a derived character state).

Monophyletic Group (Clade).

A group of species including a common ancestor and all its descendants; a "natural group" with all members more closely related to each other than to any other species (compare with paraphyletic).

Paraphyletic Group.

A nonmonophyletic group that includes a common ancestor but leaves out some descendants (e.g., "reptiles" leaving out birds). Sometimes contrasted with polyphyletic groups, which include two or more different ancestors and their descendants, for example, bats plus birds without their common ancestor.

Phylogenetic Tree.

A branching diagram showing relationships among organisms (e.g., frequently among different species, or among different individuals).

Phylogeny.

The evolutionary history of a group of species (or any set of taxa, genes, or tips).

Root Node.

The ancestral node at the base of a tree, representing the most recent common ancestor of all species included in that tree.

Shared Derived Character State (Synapomorphy).

A character state that defines a monophyletic group (e.g., the presence of mammary glands for mammals).

Species Tree.

Phylogenetic tree showing relationships among species, generally based onmultipleindependent genes (distinguished from gene tree, a tree showing relationships based on one gene from multiple individuals, or showing relationships among multiple paralogous genes).

MOLECULAR CLOCK DATING

The molecular clock hypothesis provides a simple yet powerful way of dating evolutionary events. Under the clock assumption, the expected distance between sequences increases linearly with time of divergence. When external information about the geological ages of one or more divergence events on a phylogeny is available, based on the fossil record or certain geological events, the distances between sequences or the branch lengths on the tree can be converted into absolute geological times. This is known as molecular clock dating

TESTING THE MOLECULAR CLOCK Several statistical tests have been developed to examine whether the rate of molecular evolution is constant over time. The simplest, known as the relative rate test, examines whether two species a and b evolve at the same rate by using a third out-group species o (figure 1). As species a and b share the same common ancestor y, the distance from y to a should equal the distance from y to b if the hypothesis of the molecular clock is true: dya = dyb (figure 1A). Equivalently, one can formulate the clock hypothesis relative to the out-group as dao = dbo and test whether the difference between the two calculated distances d = dao – dbo is significantly different from 0. The sequence distances and their variances can be calculated under any model of nucleotide or amino acid substitution, and the calculated d and its standard error can be used to construct a test based on the normal distribution.

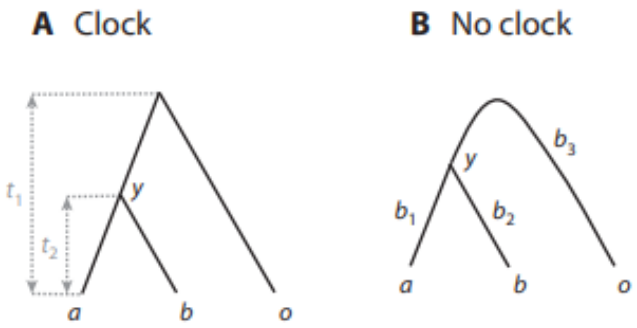


Figure 1. The relative rate test compares the rates of evolution in two species (*a* and *b*) using a third species *o* as the out-group.

Biogeography

Biogeography Definition

Biogeography refers to the distribution of various species and ecosystems geographically and throughout geological time and space. Biogeography is often studied in the context of ecological and historical factors which have shaped the geographical distribution of organisms over time. Specifically, species vary geographically based on latitude, habitat, segregation (e.g., islands), and elevation. The subdisciplines of biogeography include zoogeography and phytogeography, which involve the distribution of animals and plants, respectively.

Types of Biogeography

There are three main fields of biogeography: 1) historical, 2) ecological, and 3) conservation biogeography. Each addresses the distribution of species from a different perspective. Historical biogeography primarily involves animal distributions from an evolutionary perspective. Studies of historical biogeography involve the investigation of phylogenic distributions over time. Ecological biogeography refers to the study of the contributing factors for the global distribution of plant and animal species. Some examples of ecological factors that are commonly studied include climate, habitat, and primary productivity (the rate at which the plants in a particular ecosystem produce the net chemical energy). Moreover, ecological biogeography differs from historical biogeography in that it involves the short-term distribution of various organisms, rather than the long-term changes over evolutionary periods. Conservation biogeography seeks to effectively manage the current level of biodiversity throughout the world by providing policymakers with data and potential concerns regarding conservation biology.

How Does Biogeography Support Evolution?

Biogeography provides evidence of evolution through the comparison of similar species with minor differences that originated due to adaptations to their respective environments. Over time, the Earth's continents have separated, drifted apart, and collided, resulting in the creation of novel climates and habitats. As species adapted to these conditions, members of the same species that had been separated geographically diverge, resulting in the eventual formation of distinct species. This knowledge is important, as by understanding how adaptations occurred in response to changing environments in the past, we can apply this knowledge to the future.

Example: The Galapagos Islands

One of the most famous examples of biodiversity in support of evolution is Charles Darwin's study of finches on the Galapagos Islands, which resulted in his book *On the Origin of Species*. Darwin noted that the finches on the mainland of South America were similar to those located on the Galapagos Islands; however, the shape of the bills differed depending on the type of food available on each island. The islands had once been a part of the South American mainland, but the two land masses were subsequently separated and drifted apart. The result was the creation of novel habitats and food sources available for the species residing in each of these regions. Therefore, each finch species had adapted to the local environment through the selection of alleles which promoted survival, eventually resulting in speciation. Islands are excellent for the study of biogeography because they consist of small ecosystems that can easily be compared to those of the mainland and other nearby regions. Moreover, since they are an isolated region, invasive species and the associated consequences for other organisms within the ecosystem can be readily studied. By studying such changes over time, the evolution of distinct species and ecosystems becomes apparent.

Phylogeography

Phylogeography is a field of study concerned with the principles and processes governing the geographic distributions of genealogical lineages, especially those within and among closely related species.[1] Phylogeography is a subdiscipline of biogeography. It is "inherently interdisciplinary, with input from, for example, molecular genetics, population genetics, phylogenetics, demography, climatology, ecology, and historical geography — plus archaeology, anthropology, and linguistics, in the case of humans, and ethology and palaeontology in the case of other species".

In the field of population genetics phylogeography refers to the joint phylogenetic (genealogical) relationships and geographic distributions of genetic lineages. It is a method of mapping the DNA sequence variants to their geography to show the present-day distribution of genetic lineages. Phylogeography is most often applied to mitochondrial DNA and Y chromosome lineages. Phylogeography is a scientifically legitimate term. One can talk of – for example – high or low phylogeographic structure; with the former, closely related lineages are also found to be geographically clustered, with the latter there is little correlation between the relatedness of lineages and their geographic distribution. There is also no doubt that the

phylogeographic structure of a genetic locus (a piece of DNA that is inherited as a chunk, and so has a single phylogenetic tree / genealogy associated with it) contains information on the ancestry of a population, or even an individual (although the amount of information contained can be surprisingly small). The main issue is how do we extract that information to make reliable inferences about the ancestry of a population or even an individual?

Systematic and interpretative phylogeography

While phylogeography in itself is perfectly valid as a description of the spatial distribution of observed lineages, issues arise when it comes to making inferences about the geographic location of unobserved lineages, such as those in the past, and the population history that shaped the phylogeography. Approaches to this can be broadly lumped into two categories: **systematic** and **interpretative**.

A **systematic approach** is one that can be formalized in an algorithm and automated (usually using a computer programme). This means that given the same phylogeographic data, it should give the same answer every time. The attraction of making an inference approach systematic is that it can be tested explicitly to see if it works (by simulating data under a known history and seeing if the approach recovers that history). The first major systematic phylogeographic inference approach was designed by Alan Templeton and was called Nested Clade Phylogeographic Analysis (NCPA).[3] NCPA has been tested and explicitly shown not to work (see Nielsen and Beaumont 2009, and references therein). Other systematic approaches have also been developed and hold out some promise.

The more widely-used approach to inferring the geographic location of unobserved lineages, and the population history that shaped the phylogeography, is known as **interpretative phylogeography**. It is the interpretative aspect that is questioned by the majority of population geneticists. Interpretative approaches are, by definition, not formalised in an algorithm, and so not testable. The analysis is usually performed by interpreting phylogeographic patterns as indicating a particular population history, rather than by systematically exploring a range of population histories to test which best explain a particular phylogeographic pattern (using, for example, computer simulations). What is more, such approaches are easily steered by subjective biases, which may help to explain their popularity among some ancestry ‘testing’ companies.

Interpretative phylogeography has been criticised by population geneticists because of its lack of scientific and statistical rigour, and has been described by many scientists as storytelling.

The scientific approach is to construct explicit models of population histories – incorporating what we know about the inheritance process (population genetics) – to make predictions about those relationships and geographic distributions in the past, and then see which of those models best explain the genetic data.[12] This follows the standard hypothesis testing / hypothesis comparison paradigm that is the cornerstone of modern science. However, a number of people short cut this hypothesis testing / hypothesis comparison step and merely interpret phylogeography *post hoc*. The problems with this approach to inferring ancestry are numerous. Many of the problems are somewhat obvious, such as it is easy for interpretations to be steered by subjective biases. For example, some genetic ancestry testing companies will try and claim an association with a haplogroup and a historical personality. More worryingly, some researchers will attempt to mould a population's history to some nationalist or political agenda. Other problems are somewhat more subtle, such as the failure to take into account the intrinsically noisy nature of genetic inheritance; any phylogeographic pattern could be explained by a very wide range of often very different population histories / ancestries.

Interpretative phylogeography is best viewed "as a means of generating hypotheses (storytelling), whereas explicit models permit those hypotheses (or stories) to be tested".[9]
Despite the criticisms, many researchers have published papers in peer-reviewed scientific journals in the last 25 years using the technique of interpretative phylogeography. Their presence in the scientific literature has been considerable, but is now on the decline, though the debate continues.

Macroevolution

Macroevolution is an evolution that occurs at or above the level of the species. It is the result of microevolution taking place over many generations. Macroevolution may involve evolutionary changes in two interacting species , as in coevolution , or it may involve the emergence of one or more brand new species

Origin of Species

One of the main topics in macroevolution is how new species arise. The process by which a new species evolves is called **speciation**. How does speciation occur? How does one species evolve into two or more new species? To understand how a new species forms, it is important to review what a **species** is. A species

is a group of organisms that can breed and produce fertile offspring together in nature. For a new species to arise, some members of a species must become reproductively isolated from the rest of the species . This means they can no longer interbreed with other members of the species . How does this happen? Usually, they become geographically isolated first.

Allopatric Speciation

Assume that some members of a species become geographically separated from the rest of the species. If they remain separated long enough, they may evolve genetic differences. If the differences prevent them from interbreeding with members of the original species , they have evolved into a new species . Speciation that occurs in this way is called allopatric speciation

Sympatric Speciation

Less often, a new species arises without geographic separation. This is called **sympatric speciation**

Coevolution

Evolution generally occurs in response to changes in the environment. Environmental change often involves other species of organisms. In fact, many species evolve along with other species with which they interact. This is called **coevolution**. As one species changes, the other species must also change in order to adapt. The coevolution of rough-skinned newts and garter snakes is described above. Many other cases of coevolution occur in flowering plants and the species that pollinate them.

Character macroevolution

Evolutionary homology was applied first to organismal anatomy and form. The forelimbs of a human and an orangutan are homologous as vertebrate forelimbs because they descend, with much modification, from the forelimbs of a common ancestral form. Wagner (1989) elaborated in his concept of biological homology the properties that we expect of homologous organismal structures. First, homologies are historically unique; they arise in a particular population lineage at a particular place and time and occur only in the descendants of that lineage. Second, they have evolutionary continuity; two characters are homologous to each other only if there is an unbroken chain of lineal descent connecting them to each other and to their common ancestral origin. Third, homologies are individuated; they exist as semiautonomous components within the context of the organism as a whole. A vertebrate forelimb, for example, has an individual evolutionary history and semiautonomous developmental dynamic within an organism. We construct characters and test hypotheses of their homology also at the cellular level. Among the many cellular-level characteristics used in evolutionary studies are the detailed structures of chromosomes as they appear during cell division. Chromosomes are homologous to each other if they descend with some modification from a common ancestral chromosome. Chromosomes of an orangutan each have homologous chromosomes in human cells despite some minor rearrangements of chromosomal contents. Perhaps most effective for testing precise hypotheses of common descent of species is homology at the level of DNA sequences.

Having defined the concept of homology, we must establish principles for testing hypotheses of character homology, evaluating whether the general principle of homology explains our comparisons for a particular set of characters. There are three diagnostic tests of homology that can be applied at the organismal, cellular, and molecular levels (Patterson 1988). Whether a set of structures passes or fails these tests separates homology from contrasting explanations for character variation and resemblance. The first diagnostic test is that of charactersimilarity. This test is the definitive one at the molecular level. If one compares the DNA base sequences of two pieces of DNA and finds greater than 70 percent sequence similarity, then those sequences undoubtedly trace to a common ancestral DNA sequence. The hereditary pathway that connects them to their common ancestral sequence is nonetheless sometimes a contorted one, as revealed by the second and third diagnostic tests

The second diagnostic test for homology is the conjunction test; if two organismal structures are hypothesized to be homologous to each other, a single organism should not express both

structures at the same stage of development. The hypothesis that human arms and bird wings are homologous as vertebrate forelimbs would be rejected were we to find angels as often depicted in Italian Renaissance artwork; no living or fossil forms, however, have the characteristic arms of humans and wings of birds present in the same organism.

The third diagnostic criterion is the congruence test (figure 1). A homology has by definition only a single origin on the tree of life, and we hypothesize that a homology is transmitted from its lineage of origin to all descendant lineages (unless secondarily lost). If we assume that a new homology spreads throughout its population shortly after arising and is not subsequently lost, the homology should characterize all and only the descendants of that ancestral population. Hence, a homology should characterize a particular clade of the phylogenetic tree.

CHARACTER EVOLVABILITY

Evolutionary biologists sometimes try to evaluate the potential of a species for undergoing further evolutionary change and diversification, including production of new species and new characters. Darwinian theory and its population-genetic models make clear that large geographic distributions and large amounts of genetic variation enhance the opportunities for a species lineage to give rise to new species and new characters. Because genetic variation in a population often stabilizes organismal development rather than expressing itself as greater variation in organismal morphology, a species that is relatively uniform in organismal morphology nonetheless can have the potential to produce a great range of organismal morphologies. If genetic variation is reordered, as might occur in the founding of a new population by a small number of individuals drawn from the ancestral one (Carson and Templeton 1984), organismal development can be destabilized to reveal alternative morphologies whose developmental pathways were latent in the ancestral population.

Fossils

Fossils provide solid evidence that organisms from the past are not the same as those found today; fossils show a progression of evolution. Scientists determine the age of fossils and categorize them all over the world to determine when the organisms lived relative to each other. The resulting fossil record tells the story of the past, and shows the evolution of form over millions of years ([Figure 1]). For example, highly detailed fossil records have been recovered for sequences of species in the evolution of whales and modern horses. The fossil record of

horses in North America is especially rich and many contain transition fossils: those showing intermediate anatomy between earlier and later forms. The fossil record extends back to a dog-like ancestor some 55 million years ago that gave rise to the first horse-like species 55 to 42 million years ago in the genus *Eohippus*. The series of fossils tracks the change in anatomy resulting from a gradual drying trend that changed the landscape from a forested one to a prairie. Successive fossils show the evolution of teeth shapes and foot and leg anatomy to a grazing habit, with adaptations for escaping predators, for example in species of *Mesohippus* found from 40 to 30 million years ago. Later species showed gains in size, such as those of *Hipparion*, which existed from about 23 to 2 million years ago. The fossil record shows several adaptive radiations in the horse lineage, which is now much reduced to only one genus, *Equus*, with several species.

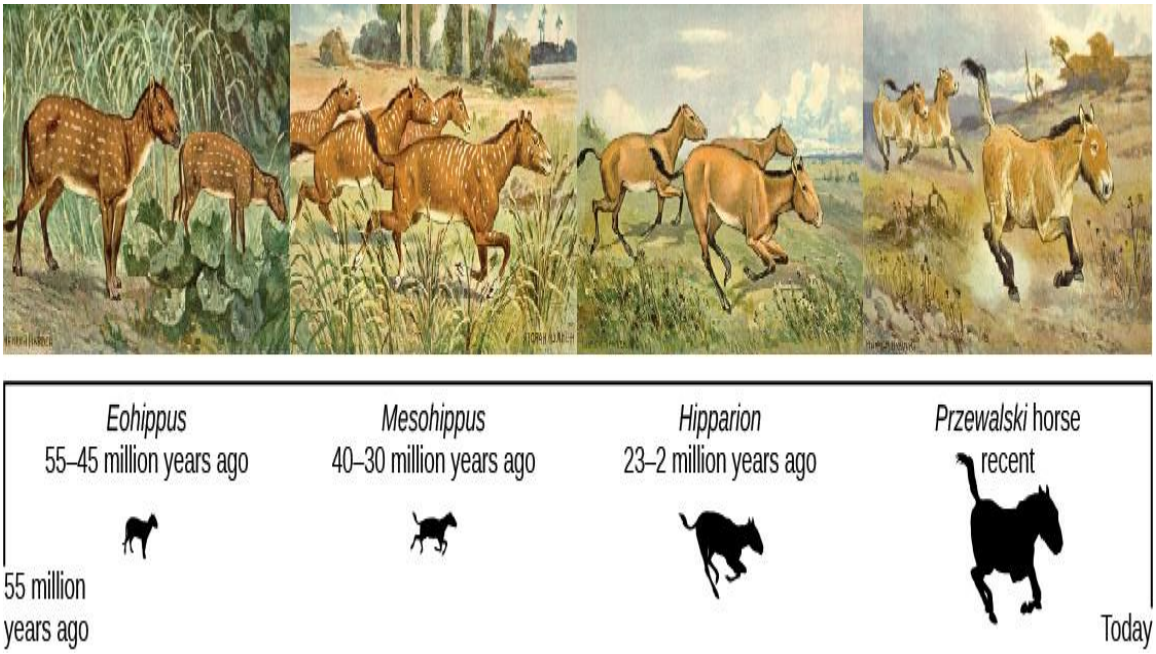


Figure 1: This illustration shows an artist's renderings of these species derived from fossils of the evolutionary history of the horse and its ancestors. The species depicted are only four from a very diverse lineage that contains many branches, dead ends, and adaptive radiations. One of the trends, depicted here is the evolutionary tracking of a drying climate and increase in prairie versus forest habitat reflected in forms that are more adapted to grazing and predator escape through running. Przewalski's horse is one of a few living species of horse.

Anatomy and Embryology

Another type of evidence for evolution is the presence of structures in organisms that share the same basic form. For example, the bones in the appendages of a human, dog, bird, and whale all share the same overall construction ([Figure 2]). That similarity results from their origin in the appendages of a common ancestor. Over time, evolution led to changes in the shapes and sizes of these bones in different species, but they have maintained the same overall layout, evidence of descent from a common ancestor. Scientists call these synonymous parts homologous structures. Some structures exist in organisms that have no apparent function at all, and appear to be residual parts from a past ancestor. For example, some snakes have pelvic bones despite having no legs because they descended from reptiles that did have legs. These unused structures without function are called vestigial structures. Other examples of vestigial structures are wings on flightless birds (which may have other functions), leaves on some cacti, traces of pelvic bones in whales, and the sightless eyes of cave animals.

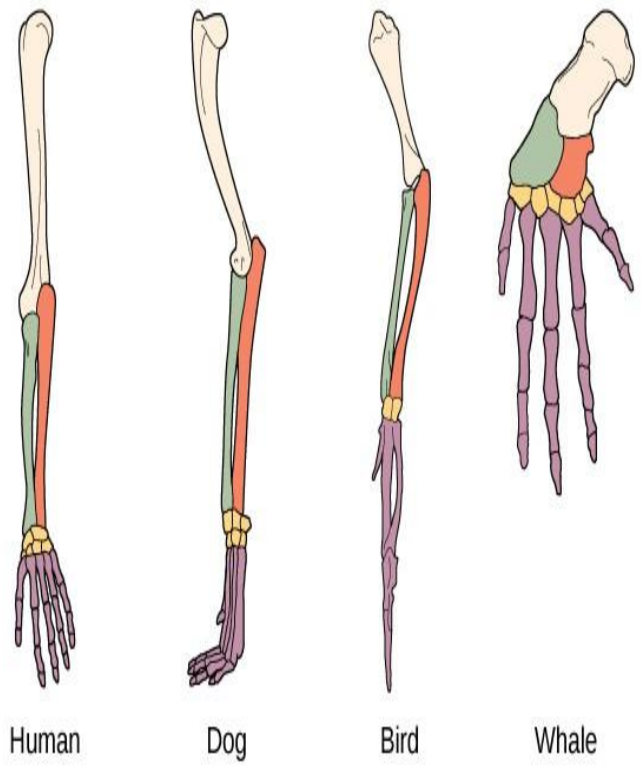


Figure 2: The similar construction of these appendages indicates that these organisms share a common ancestor.

Click through the activities at this interactive site to guess which bone structures are homologous and which are analogous, and to see examples of all kinds of evolutionary adaptations that illustrate these concepts.

Another evidence of evolution is the convergence of form in organisms that share similar environments. For example, species of unrelated animals, such as the arctic fox and ptarmigan (a bird), living in the arctic region have temporary white coverings during winter to blend with the snow and ice ([Figure 3]). The similarity occurs not because of common ancestry, indeed one covering is of fur and the other of feathers, but because of similar selection pressures—the benefits of not being seen by predators.

Figure 3: The white winter coat of (a) the arctic fox and (b) the ptarmigan's plumage are adaptations to their environments. (credit a: modification of work by Keith Morehouse)

Embryology, the study of the development of the anatomy of an organism to its adult form also provides evidence of relatedness between now widely divergent groups of organisms. Structures that are absent in some groups often appear in their embryonic forms and disappear by the time the adult or juvenile form is reached. For example, all vertebrate embryos, including humans, exhibit gill slits at some point in their early development. These disappear in the adults of terrestrial groups, but are maintained in adult forms of aquatic groups such as fish and some amphibians. Great ape embryos, including humans, have a tail structure during their development that is lost by the time of birth. The reason embryos of unrelated species are often similar is that mutational changes that affect the organism during embryonic development can cause amplified differences in the adult, even while the embryonic similarities are preserved.

Human Evolution

The study of evolution is one that scientists and researchers have been carrying out for the last few centuries, with the biggest pioneer in the field of evolution being the English scientist Charles Darwin, whose concepts such as natural selection helped form the basis for the theory of evolution as we know it today. In simple terms, evolution refers to "the process by which new species or populations of living things develop from pre-existing forms through successive generations", as defined by Merriam Webster. The changes characterizing evolution do not occur in any particular individual alone- these changes that are beneficiary for one's survival are passed onto the offspring and so on to a whole population. This article seeks to look at human evolution in particular, and give a general overview of the same in layman terms.

What is Human Evolution?

Human evolution refers to "the process by which human beings developed on earth from now-extinct primates" (Tuttle, 2020). The subject of human evolution in of itself is so vast that there is a separate section of scientific study dedicated to it known as Palaeoanthropology. It is a subsection of anthropology wherein scientists seek to investigate the origins of our species and human behaviour by studying human fossils.

Something that many may not know is that *Homo sapiens*, i.e human beings are one of many species that come under the genus *Homo,* with other species before us including *Australopithecus afarensis, Homo erectus,* and *Homo habilis* amongst others. There are an estimated 15 to 20 such species under the category of early humans. The feature that

distinguished hominids (our early ancestors) from other apes is that of the ability to walk on two feet, known as bipedalism, which developed over 4 million years ago. Human beings, chimps and gorillas share a common ancestor with chimps and bonobos being our closest relatives however the chimp-human lineages split around 7 and a half million years ago. Fossil and genetic evidence both suggest that humans originated in Africa, with all fossils of early humans (lived between 2-6 million years) having been found there.

There are different ways in which researchers have mapped out the stages of human evolution. Stages of evolution can be grouped in different sections therefore there can be four or five or son depending on how they have been arranged. Here we are looking at ***four main stages of evolution,*** arranged from the earliest species to the most recent:

1. **Stage one** – between 4-7 million years ago: *Sahelanthropus*, *Orrorin* and *Ardipithecus* all appeared between this time period, and they are known as proto-hominins since it is debated whether they are our ancestors or not. They may have been bi-pedal.
2. **Stage two** -around 4 million years ago: The genus *Australopithecus*, our ancestors who are also bipedal appeared then, found in both East and West Africa.
3. **Stage three** – Around 2.7 million years ago: This is marked by the appearance of species classified under *Paranthropus*, however, it is unsure whether this classifies as a separate genus since amongst paleoanthropologists many classify it as a subsection of the Australopithecus genus.
4. **Stage four** – Between 2.5-1.8 million years ago: This stage consists of the appearance of species that all fall under the genus *Homo*, which as mentioned before consists of us and our immediate ancestor species. *Homo habilis*, the oldest is characterized by the usage of stone tools and a chimpanzee sized brain. *Homo erectus* and *Homo egaster*, the following members had double the brain size and were able to control fire and use relatively advanced Then came along *Homo heidelbergensis* about 800,000 years ago and finally *Homo sapiens* appeared approximately 200,000 years ago. Some of the main traits or features that helped build modern human society – such as culture and symbolic language are ones we acquired 50,000 years ago.

Another angle of approaching it is looking at the four important steps, or rather four important traits that were involved in the making of modern humans. They are:

1. **Terrestriality:** This refers to the adaptation of living on land (terrestrial).

2. **Bipedalism:** As has been mentioned above already, this refers to the ability to walk on two feet. This helped tremendously as now hands could be used for other labour, long-distance travelling became easier, lesser sun exposure, etc. It also resulted in changes to body parts and changes in bodily processes such as gestation.
3. **Encephalization:** An evolutionary adaptation involving an increase in the size of the brain and complexity. This increased the possibility for social learning and learning of languages.
4. **Civilization:** A society that is characterized by advanced cultural and social development, and the making of tools is one such example that is regarded as indicative of evolution.

Other evolutionary adaptations of human beings include reduced differences between male and female sexes (sexual dimorphism), lesser body hair, a hidden estrus (no physically perceptible sign of fertility), etc.

Hypotheses surrounding early human evolution:

There are a couple of hypotheses that have been formed regarding the evolution of early humans – let us look into some of them. The **savannah hypothesis** suggested that hominins became bipedal as they had to adapt to the savannah after being forced out of trees and thus began walking erect using only their feet. The **aridity hypothesis** developed this further to state that evolution also occurred due to an increase in aridity and expansion of the savannah. The **turnover pulse hypothesis** looked into the role that climate and environmental change played and suggested that it led to a higher rate of evolution amongst specialist species and the spreading out of generalist species (as they can grow even with environmental changes).

The study of human evolution is one that is ongoing, not just because evolution is an ongoing process but also because there is still a lot that is unknown and that is being analysed and explored regarding both our origins as a species, our ancestors and our recent evolution, with studies coming out often containing new findings.

SCHOOL OF BIO AND CHEMICAL ENGINEERING

DEPARTMENT OF BIOINFORMATICS

UNIT – 2- SBIA1302 – Evolutionary Biology

UNIT II

What is natural selection?

- Natural selection is a mechanism that will sustain and replicate species that are best suited to an environment.
- This suggests that this variant organism's beneficial alleles are passed on to offspring.
- The mechanism of natural selection over several centuries contributes to evolution.

Examples of natural selection:

1. Peppered moths:

- Most peppered moths were of the pale variety till the industrial revolution in Britain in the early 1800s.
- This indicated that they were camouflaged against the pale birch trees where they rested.
- Moths with a mutant black coloring were quickly detected by birds and eaten.
- This gave an advantage to the white variety, and they were more likely to continue to reproduce.
- Airborne emissions in manufacturing areas blackened the birch tree bark with soot during the second half of the 1800s.
- This indicated that they were now camouflaged by the mutant black moths, while the white variety became more susceptible to predators.
- This offered a benefit for the black type, and they were more likely to survive and replicate.
- The dark moths passed on the black wing color alleles that led to offspring with the phenotype of black wing color.
- Over time, in urban areas, black peppered moths have become much more common than the pale type.
- It should be noted that the shift in phenotype was not due to the moths being darker by pollution.
- The dark variety still existed, but when the climate changed, it was the better suited version.
- It took several years until the moth population was predominantly black in colour.

2. Galapagos finches:

- It is popular example of natural selection from Darwin's voyage.

- Every Galapagos island visited by Darwin had its own kind of finch (14 in all), found nowhere else in the world.
- Some had beaks adapted for feeding large seeds, some for small seeds, some for feeding on buds and fruits had parrot-like beaks, and some for feeding on small insects had slender beaks.
- As some woodpeckers do, one used a thorn to test for insect larvae in wood.
- Eight of them were tree finches and six of them were ground dwellers.
- It seemed that each was slightly altered from an original colonist, most likely the finch on South America's mainland, some 600 miles to the east.
- Adaptive radiation is likely to contribute to the creation of so many species because other birds were few or absent, leaving empty niches to fill; and because there were enough opportunities for geographical isolation in the various Galapagos islands.

The process of natural selection:

- There are four components in the Darwin's process of natural selection:
- **Variation:**
  - In appearance and behavior, species (within populations) exhibit individual variation.
  - Body size, hair color, facial markings, speech properties, or number of offspring can be included in these variants.
  - On the other hand, some characteristics show little or no difference among individuals, such as the number of vertebrate eyes.
- **Inheritance:**
  - Some characteristics are routinely passed on from parent to offspring.
  - These characteristics are heritable, while other traits are highly affected by environmental factors and demonstrate poor heritability.
- **Increasing population growth rate:**
  - Every year, most populations have more offspring than local resources can sustain, leading to a battle for resources.
  - Considerable mortality is faced by each generation.
- **Variability in survival and reproduction:**

- Individuals with characteristics well suited to the fight for local resources will bring more offspring to the next generation.

- **Evolution as a genetic function:**

The concept of natural selection:

- With the presence of genetic variation, the core argument of Darwin's theory of evolution begins.
- It was demonstrated to Darwin that the experiences with the plant and animal breeding led to variations that could be significant for man.
- Thus, he concluded that variations that are favorable or important for organism to survive should exist.
- These favorable variations enhanced the chances for the living and procreation.
- Those beneficial variations were conserved and passed on to generations.
- This process is actually known as natural selection.
- An organism that is well adapted to its environment is the outcome of the process, and evolution also occurs as a consequence.
- Natural selection can then be characterized as the differential reproduction of alternative hereditary variants, defined by the fact that certain variants increase the probability of survival and reproduction of the organisms more effectively than organisms that carry alternative variants.
- Selection may occur by differences in survival, fertility, rate of development, success in mating, or any other aspect of the life cycle.
- All of these differences can be integrated under the term differential reproduction since all result in natural selection to the point that they affect the number of progenies an organism leave.
- It is possible to see evolution as a two-step process.
- First, hereditary variation takes place; second, certain genetic variations are chosen that will be passed on to the subsequent generations most effectively.
- Hereditary variation also contains two mechanisms—the spontaneous mutation of one variant into another and the sexual method that recombines those varieties (see recombination) to form a wide range of variations.

- The variants that occur from mutation or recombination are not equally passed from one generation to another.
- Others will occur more often because they are beneficial to the organism; events of chance, called genetic drift, may decide the frequency of others.

Types of natural selection:

1. Stabilizing selection:

- Through studying its effects on changing gene frequencies, natural selection can be studied, but it can also be studied by examining its effects on the observable characteristics or phenotypes of individuals in a population.
- Distribution scales of phenotypic traits such as height, weight, progeny number, or longevity usually display higher numbers of people with intermediate values and the so-called natural distribution is less and less towards the extremes.
- The selection is said to be stabilizing when individuals with intermediate phenotypes are preferred and extreme phenotypes are chosen against.
- Then the range and dispersion of phenotypes remains roughly the same from one generation to another.
- Stabilizing selection is quite usual.
- Those who have intermediate phenotypic values are the ones that live and reproduce more effectively.
- For example, mortality among newborn infants is greatest when they are either very small or very large; intermediate-size infants are more likely to survive.
- Stabilizing selection is usually noticeable after artificial selection.
- As a consequence of stabilizing selection, populations frequently maintain a steady genetic constitution in relation to many traits.
- This characteristic of populations is called genetic homeostasis.

2. Directional selection:

- In a population, the distribution of phenotypes often systematically shifts in a specific direction.
- The physical and biological aspects of the environment are evolving constantly, and the changes can be important over long periods of time.
- The environment and even the structure of the land or waters differ continuously.

- In biotic environments, that is, in the other species present, whether predators, prey, parasites, or rivals, changes often take place.
- As a result, genetic changes occur as the genotypic fitnesses can shift so that different sets of alleles are preferred.
- When species colonize new habitats where the conditions are different from those of their original habitat, the potential for directional selection often occurs.
- Furthermore, as the new genetic constitution replaces the preexisting one the emergence of a new desirable allele or a new genetic combination may prompt directional changes.
- The procedure of directional selection occurs in spurts.
- The substitution of one genetic constitution with another alters the genotypic fitnesses at other sites, which then modify their allelic frequencies, stimulating further modifications, and so on in a cascade of consequences.
- Directional selection is only possible if genetic variation occurs with regard to the phenotypic characteristics under selection.
- There are vast stores of genetic variation in natural populations, and these are constantly replenished by additional new variations that emerge by mutation.
- Artificial selection's almost universal success and natural populations' rapid response to new environmental challenges show that the current variety provides the materials needed for directional selection.
- Directional selection contributes to significant changes in morphology and ways of life over geologic time.
- Evolutionary changes that remain in a more or less consistent fashion over extended periods of time are defined as evolutionary trends.
- From the tiny brain of Australopithecus, human ancestors three million years ago, which was less than 500 cc in volume, to a brain about three times as large in modern humans, lateral evolutionary improvements expanded the cranial ability of the human lineage.
- Another well-studied example of directional selection is the evolution of the horse from more than 50 million years ago to modern times.

3. Diversifying selection:

- Diversifying selection, much like directional selection, drives the population towards the extremes of the trait.

- This type of selection is also termed disruptive selection.
- Diversifying selection moves the trait both directions, in contrast to directional selection.
- This can happen in a number of ways, but since species can become so distinct, it also leads to speciation.
- However, if only diversified for short periods, the selection will lead to a variety of characteristics that can be shared by one species.
- By diversifying selection, two or more divergent phenotypes in an environment may be preferred simultaneously.
- No natural environment is homogeneous; instead that the environment of any plant or animal population is a mosaic comprised of more or less distinct sub-environments.
- Also, the heterogeneity could be temporal, with change occurring over time, along with spatial.
- Species cope in different ways with environmental heterogeneity.
- One of the strategies is the selection of a generalistic genotype i.e called genetic monomorphism, that is well suited to all of the species' sub-environments.
- Genetic polymorphism, the selection of a diversified gene pool that yields various genotypes, each suited to a particular sub-environment, is another strategy.
- In conditions in which populations living a short distance apart have been genetically distinct, the efficiency of diversifying natural selection is very evident.
- In one example, on heaps of mining waste heavily polluted with metals such as lead and copper, populations of bent grass can be found growing.
- The soil has been so polluted that it is poisonous to most plants, but it has been shown that the thick stands of bent grass growing over these refuse heaps have genes that make them resistant to high lead and copper concentrations.
- But bent grass plants that are not resistant to these metals can be found only a few metres from the polluted soil.
- Bent grasses reproduce mainly by cross-pollination, so that wind-borne pollen from the neighboring non-resistant plants is collected by the resistant grass.
- Since non-resistant seedlings are unable to grow in the polluted soil and the non-resistant seedlings outgrow the resistant ones in the surrounding uncontaminated soil, they retain their genetic differentiation.

- The evolution of these resistant strains has occurred in the lesser than 400 years since the mines were first opened.

4. Sexual selection:

- A significant factor in reproduction is mutual attraction between the sexes.
- Except for the reproductive organs and secondary sexual features, such as the breasts of female mammals, the males and females of many animal species are identical in size and form.
- However, there are species in which striking dimorphism is displayed by the sexes.
- Males are often larger and heavier, more brightly colored, or endowed with conspicuous adornments, especially in birds and mammals.
- However, bright colors make animals more conspicuous to predators-in the best of situations, the long plumage of male peacocks and paradise birds and the large antlers of aged male deer are bulky tons.
- Darwin knew that natural selection could not be predicted to favour the evolution of undesirable traits, and he was capable of offering a solution to this problem.
- He indicated that such characteristics occur by "sexual selection," which does not depend on a fight for existence in relation to other organic beings or external conditions."
- Other things being equal, species with greater fitness are more proficient in securing partners.
- There are two general conditions that lead to sexual selection.
- One is the choice displayed one sex (often females) for individuals of the other sex that display certain traits.
- The other is enhanced strength (usually among males) that produces greater success in attracting mates. The existence of a specific attribute among members of one sex can make them more attractive to the opposite sex in some way.
- In all kinds of species, from vinegar flies to pigeons, rats, dogs, and rhesus monkeys, this form of "sex appeal" has been experimentally illustrated.
- For example, when Drosophila flies are put together, some with yellow bodies as a result of random mutation and others with regular yellowish gray pigmentation, normal males are preferred over yellow males by females with either body color.

5. Kin selection and reciprocal altruism:

- Like other examples of sexual selection, the apparent altruistic behavior of many species is a characteristic that initially appears incompatible with the theory of natural selection.
- Altruism is a type of behavior that favors other people at the cost of the one who performs the action; the altruist's fitness is decreased by his behavior, whereas people who behave selfishly benefit from it at no cost to themselves.
- Accordingly, natural selection may be expected to encourage the production of selfish behaviour and eradicate altruism.
- This outcome is not so convincing when it is realised that the beneficiaries of altruistic behaviour are usually relatives.
- Many of them bear the same genes, including those that foster altruistic behaviour.
- Altruism can grow through the selection of kin, which is simply a form of natural selection in which relatives are taken into account when determining the fitness of an individual.
- Natural selection favors genes that increase their carriers' reproductive success, but it is not mandatory for reproductive success to be greater for all individuals that share a given genotype.
- On average, it is necessary for carriers of the genotype to replicate more effectively than those with alternative genotypes.
- A parent shares half of its genes with each progeny, so if the cost of the behavior to the parent is less than half of its average benefits to the progeny, a gene that promotes parental altruism is preferred by selection.
- Over the generations, such a gene is more likely to increase in frequency than an alternative gene that does not support altruistic behavior.
- Therefore, parental care is a type of altruism readily explained by the selection of kin.
- As it promotes the reproductive success of the parent's genes, the parent spends some energy caring for the progeny.
- Kin selection goes beyond the association between parents and their offspring.
- It promotes the development of altruistic behavior when an individual's energy invested, or the risk incurred, is compensated in excess by the benefits that follow through relatives.
- The finer the relationship between the beneficiaries and the altruist, and the higher the number of beneficiaries, the greater is the altruist's risks and efforts.

- Individuals who live together in a herd or troop are generally related and often act in this way towards each other.
- For instance, adult zebras, instead of fleeing to protect themselves will turn towards an attacking predator to protect the young in the herd.
- Altruism often happens when the action is reciprocal among unrelated people and the cost of the altruist is smaller than the gain to the recipient.
- This reciprocal altruism is noticed in the mutual grooming of chimpanzees and other primates as they scrub each other of lice and other pests.

Units and Levels of Selection

Selection is the primary cause of adaptive evolutionary change. To understand how selection takes place, it is necessary to determine what biological units or entities, at what levels of biological organization, are subject to its effects, and how its prevalence and efficacy varies across levels. Selection itself can usefully be defined as the differential reproduction of biological units due to differences in form or character between these units. 'Units of selection' are defined here as the units whose frequencies are adjusted by natural selection across generations, and 'levels of selection' are defined here as the levels of biological organization where natural selection occurs, within generations. Biological units are arranged in a hierarchy (Table 1), with higher-level units subsuming lower-level ones, and units at the different levels exhibiting different properties with regard to how they reproduce and whether and how they interact with units at different levels and with aspects of the environment. These properties are critical to understanding how and why natural selection, and across-generation responses to selection, take place

Table 1 The primary levels of biological organization, the units at each level, and the properties of the units

Level	Amount of variation	Turnover rate	Inheritance fidelity	Expression of traits	Nature of unit
Genes	High	High	Very high	No	Replicator
Chromosomes	High	High	Medium	No	Replicator
Genotypes	High	High	Low	No	Replicator
Gene products	High	High	N/A	Yes	Vehicle
Cells	Variable	High	N/A	Yes	Vehicle
Individuals	High	High	Low	Yes	Vehicle
Groups	Variable	Variable	Variable	Yes	Vehicle
Species	Variable	Very low	Variable	Yes	Vehicle
Communities	Variable	Low	Variable	Yes	Vehicle

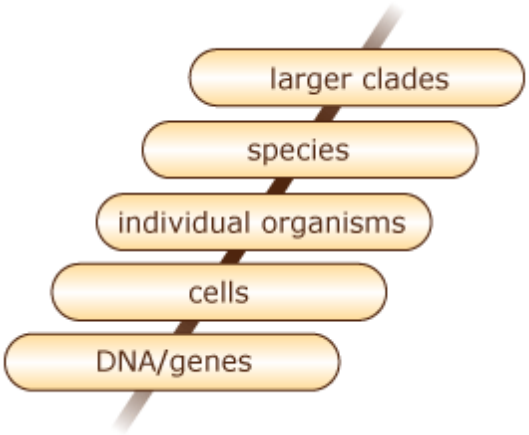
N/A, not applicable.

The degree to which the units at different levels in Table 1 are units of selection depends on three conditions (Lewontin, 1970; Alexander and Borgia, 1978). First, the units must exhibit variation among themselves in their effects, and the potential strength of selection at that level increases with the magnitude of the expressed variation. Second, the units must have some rate of differential reproduction or turnover, which determines the frequency of selective episodes, and this differential reproduction must be causally linked to variation among units. Third, the units must exhibit fidelity of inheritance, such that they persist as unique, replicating variant units for a sufficient number of selective episodes to have their frequencies adjusted by natural selection.

The gene is the primary unit of selection because it is the only unit exhibiting high variation, high turnover rate, and ability to replicate or reproduce with extremely high fidelity. Chromosomes may be units of selection only if rates of recombination are very low or zero, and genotypes may be units of selection only in asexual forms where the absence of recombination and meiosis results in the inheritance of entire genomes intact. To be a unit of selection, being a replicator is key (Dawkins, 1982), and only the lowest-level units can meet this stringent criterion.

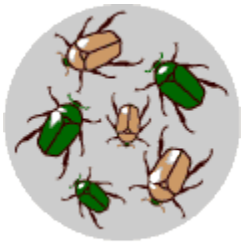
The hierarchy of selection

We most typically think of natural selection working at the level of the individual, favoring those better at leaving behind more individual descendants. However, with a little imagination, we can see how natural selection might work at other levels of biological organization as well. Moving down the hierarchy, natural selection could act on the cells within an individual, favoring those cell lineages better at leaving behind descendent cells. Moving up the hierarchy, natural selection could act on species, favoring those species better at diversifying into descendent species.

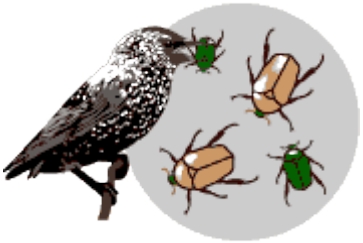


Selection at the individual level

To understand how selection works at different levels, first consider the requirements of natural selection in the most familiar situation, at the level of the individual organism, using a beetle population as an example:



1. Variation in traits.
Some beetles are green and some are brown.



2. Differential birth and death.
Since the environment can't support unlimited population growth, not all individuals get to reproduce to their full potential. In this example, green beetles tend to get eaten by birds and survive to reproduce less often than brown beetles do.

3. Heredity.
Because this trait has a genetic basis, the surviving brown beetles have brown offspring.

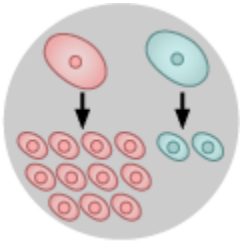
In the end, the advantageous trait, brown coloration, becomes more common in the population.

Selection at the cellular level

This process requires just three basic features to operate: variation, differential birth and death, and heredity. However, these traits are not unique to populations of individuals. As an example, imagine a group of cells within one person's liver:

1. Variation in traits.

Most of the cells are basically the same, but one has experienced a chance mutation, inactivating a gene that controls the cell's growth.



2. Differential birth and death.

Because it has lost the ability to regulate growth, the cell with the chance mutation divides more rapidly than the others.

3. Heredity.

The normal cells pass on normal DNA when they divide, and the mutant cells pass on mutant DNA when they divide.

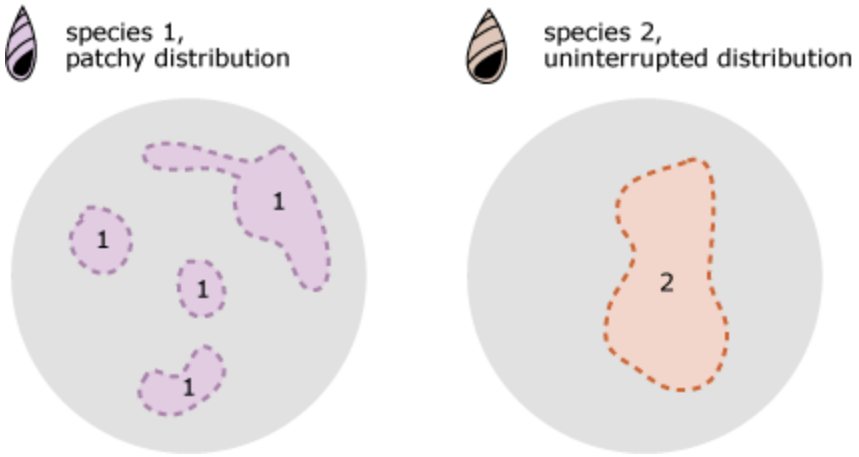
In the end, the cells carrying mutant DNA become much more common in the liver and may even take over, causing the liver to malfunction. In this case, the result of selection is a process better known as cancer. Selection at the cellular level is constantly operating within all multicellular organisms (including humans), but we rarely notice it except when it leads to such detrimental effects. Interestingly, cellular selection can work *against* selection at the level of the individual: what's advantageous for a cell lineage (e.g., replicating out of control) can be disadvantageous for the whole organism (e.g., causing an early death from cancer).

Selection at the species level

Moving up the hierarchy, we can also see how this same process would work on entire lineages of organisms. As an example, imagine a group of closely related species:

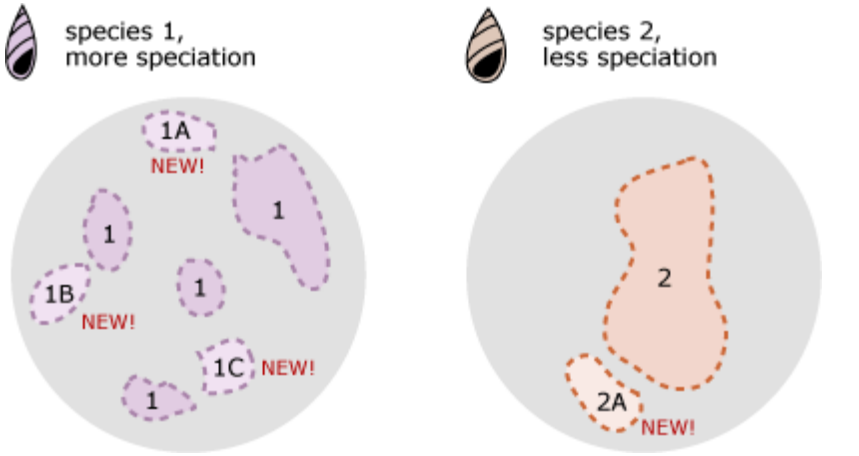
1. Variation in traits.

Some species have patchy distributions and others have uninterrupted distributions. Patchy distributions often result in isolated populations evolving into new species.



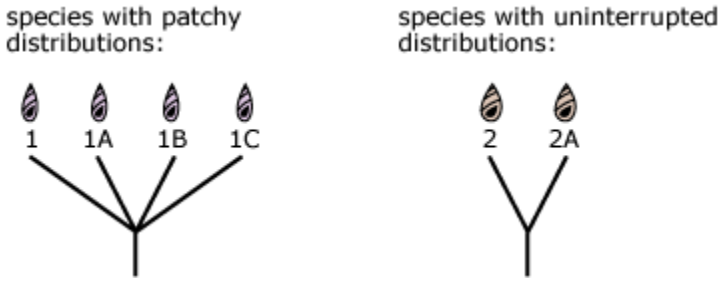
2. Differential birth and death.

Species with patchy distributions are more likely than others to undergo speciation — the analog of "birth" for a species.



3. Heredity.

Species with patchy distributions tend to speciate into lineages with similarly patchy distributions, and species with uninterrupted distributions tend to speciate into lineages with uninterrupted distributions.



In the end, this process favors species with patchy distributions, and those species will eventually become more common than those with uninterrupted distributions. If the patchy lineage happens to be purple and the continuous lineage happens to be brown, the number of purple species will increase faster than the number of brown species.

Selection at the species level is much harder to observe than selection at the cellular, or even the individual, level because species selection takes place over millions of years. Nevertheless, we know that this process theoretically *could* happen. We even have some evidence that it *does* happen. Paleontologist David Jablonksi and his colleagues have found evidence suggesting that geographic range in marine mollusks may evolve this way. They collected data on the occurrence of more than 1000 different mollusk species fossilized in late Cretaceous rocks of southeastern North America and found a few key patterns:

1. Variation in traits.

Geographic range varies across mollusk species. Some have broad ranges and some have small ranges.

2. Differential birth and death.

Those mollusk species with broad ranges are extinction resistant (i.e., have a low "death" rate).

3. Heredity.

When mollusks with broad geographic ranges speciate, they tend to give rise to other species with broad geographic ranges — in other words, geographic range is a heritable trait for these species.

Marine mollusks seem to have all the basic characteristics that would result in species selection — but how important has this process been in their evolutionary history and the history of life in general? Scientists are still working out the answer to that question. While we know that selection can operate at levels above that of the individual, it is still unclear how important this process has been in shaping the history of life.

Evolutionary Limits and Constraints

A population or species might have the potential to evolve, but other factors like the movement of genes among populations and/or trait interactions make evolutionary change difficult despite ongoing selection. Both fundamental limits and other forms of constraints can prevent populations and species from adapting to new environments. Constraints restrict species to living

under a particular set of environmental conditions. In this way, limits and other constraints drive biodiversity; without them, there might be a few common species adapted to a wide range of conditions, rather than a diversity of species, the majority restricted to a narrow range of ecological conditions. Explanations for evolutionary constraints can be divided into two categories: those that reflect the nature of genetic variation required for evolution and adaptation, and those that have their origin in the ecological processes to which populations and species are exposed.

1. LACK OF GENETIC VARIATION AS A LIMIT AND CONSTRAINT

   Natural selection will not change a trait if the trait lacks genetic variation. In the absence of genetic variation, any increase or decrease in the mean value of a trait after selection will not be passed on to the next generation. A way to think about this issue is to imagine a population of individuals derived from a single clone. Barring mutation, all individuals in the population will then be genetically identical, and there will be no genetic variation for any trait in this population. Because of differences in environmental conditions experienced by individuals, they will still differ somewhat in appearance and performance. However, even if these differences affect the fitness of individuals, the differences will not be passed on to subsequent generations because all individuals are genetically identical.

   Because traits are typically affected by a number of genes (and the regulatory mechanisms acting on these genes), genetic variation in a trait may be lost only when there is a cumulative effect of molecular changes at multiple loci, or when a key regulatory gene in a developmental pathway is inactivated. The absence of genetic variation in a trait can be detected through the inability of selection to change the distribution of the trait when artificial or natural selection is imposed in a particular direction

2.TRADE-OFFS

While evolutionary constraints due to DNA decay and loss of genetic variation arise because genes are absent and nonfunctional or lack genetic variation, limits can also arise because of the pleiotropic effects of genes, which occur when the same genes affect multiple traits. Genes have an enormous potential for pleiotropic effects because proteins encoded by genes are embedded in networks of interacting biochemical processes, and these networks in turn are likely to influence the expression of multiple traits. Moreover, genes that regulate the expression of other genes can

have pleiotropic effects by influencing multiple networks. Because of the complex and indirect ways in which genes influence phenotypes (see chapter V.13), selection for a decrease or increase in a trait will favor a set of underlying allelic changes that simultaneously impact other traits.

Evolutionary trade-offs can be studied either by considering allelic variants of genes individually or by examining patterns of genetically based correlations among traits.

3.MULTIVARIATE SELECTION

Although trade-offs due to trait interactions are often regarded as essential for evolutionary constraints, other types of genetic interactions among traits can also prevent selection responses. Evolution is constrained when no genetic variation is available for a population to respond in the direction in which selection acts, and this does not necessarily require trade-offs or negative genetic correlations among traits. It is the way in which multiple traits under selection interact at the genetic level that drives eventual selection limits. Understanding this type of selection limit requires an understanding of selection on multiple traits at once

4. GENE FLOW IN MARGINAL POPULATIONS LIMITING RANGE EXPANSION Gene flow occurs when individuals or propagules move from one population to another and then contribute to the genetic constitution of the other population (see chapter IV.3). This process can both enhance and retard evolutionary adaptation. The former occurs when gene flow increases genetic variation by introducing new genetic variants into a population that can then be selected to increase fitness. On the other hand, when too much gene flow occurs, the effects of selection can be overwhelmed by an influx of nonadapted genotypes. Gene flow can then act as an evolutionary constraint.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF BIO AND CHEMICAL ENGINEERING

DEPARTMENT OF BIOINFORMATICS

UNIT – 3- SBIA1302 – Evolutionary Biology

Evolutionary Processes

Biologists organize their thinking about biological processes using evolution as the framework. There are five key mechanisms that cause a population, a group of interacting organisms of a single species, to exhibit a change in allele frequency from one generation to the next. These are evolution by: mutation, genetic drift, gene flow, non-random mating, and natural selection. Each mechanism of evolution can be characterized by how it affects fitness, adaptation, the average phenotype of a trait in a population, and the genetic diversity of the population.

Genetic drift

The simple definition of genetic drift (also referred to as Sewall Wright effect or allelic drift) is a method of changing the population's frequency of an allele by chance where some individuals with specific allele reproduce more than the others, this process may result in the loss of beneficial alleles or allele fixation of harmful alleles since the gene frequency changes by chance and not because it is a beneficial allele such as the natural selection theory.

Genetic drift contributes to the natural evolution of species, it may lead to the fixation of new alleles that have been recessively fixed in the previous individuals and the development of a new feature in the population. Genetic drifting occurs in all species; however, it is much more significant in small populations in which the number of individuals has been reduced due to a natural disaster (***bottleneck effect***) or when a number of individuals separate from their population to form a new colony (***founder effect***).

Is genetic drift random? In fact, genetic drift is random as it occurs as a result of pure chance, however, it affects small populations significantly unlike large populations that are not susceptible to change due to chance. For example, if a population of 5 members lost 1 member due to chance without leaving any offspring, the population would lose 1/5 of their genetic pool which represents 20 percent of the population's genes. Alternatively, if one member is lost from a 50-member population, the lost genes would represent 1/50 of their genetic pool which represents only 2 percent of the overall population's genes. Therefore, large populations are not greatly affected by the impact of genetic drift.

Unlike natural selection, genetic drift outcomes are unpredictable, it is like flipping a coin without knowing which face you will get, gene drifting outcomes are always due to pure chance.

What causes genetic drift?

Even though genetic drift is the result of chance, however, if you need to define genetic drift outcomes you should know how it is affected by many factors such as the number of individuals within the population, where the effect of genetic drift is more predominant in small populations, the second factor is the number of individuals contributing in genetic drifting since some individuals do not produce offspring. Another factor is the occurrence of natural disasters that affects the population size; accordingly, these disasters will increase the significance of random genetic drift among the population as well as the change in natural gene flow patterns. Habitat fragmentation is one of the main factors that influence gene drifting where humans live in nonurban habitats so nonhuman populations are dispersed or even eliminated. As a result, gene flow among these populations is decreased while gene drifting becomes more significant.

Humans can change the effect of gene drifting where they can move individuals to new environments voluntarily or accidentally due to urbanization, urbanization affects the less-mobile species to an extent more than mobile species by increasing the rate of gene flow among these species. The gene flow facilitated by humans is usually known as “Human-facilitated gene flow” it may also introduce new genes into the population allowing for new alleles and mutations. Dispersion of populations due to urbanization may decrease the influence of genetic drift since the bottleneck effect will decrease.

Genetic Drift Consequences

What is the role of Genetic drift in evolution? Genetic results in loss of diversity in the genetic pool where fixation of certain genes may be similar to its initial frequency in the population, another consequence of genetic drifting is the increase in differentiation among populations where different the population of the same species may have different alleles due to genetic drifting, this occurs if the frequency of one allele is fixed in a population, while its frequency decreases in another population since the frequencies of alleles in the genetic pool proceed in different directions.

Sampling error can cause more changes in small populations allele frequencies than the large ones. In small populations, alleles are fixed more rapidly than large populations since the gene pool of large populations tends to be more stable. However, Large populations are affected by genetic drift since they become small as a result of a natural disaster that decreases the population size greatly until the conditions improve to allow the regrowth of the population (bottleneck effect) or when a small population leaves the group to establish a new colony when

the population is threatened (founder effect). In order to protect a small population from extinction, genetic diversity should be maintained, consequently, genetic drift should be minimized and mutation should be maintained to facilitate adaptation.

Types of Genetic Drift

Genetic drift effect does not depend on the benefits of an allele, since a harmful allele may be fixed and a beneficial allele may be lost by chance. Regardless of their effect, rare recessive genes can become more common by the effect of genetic drift when the population is exposed to a natural disaster (bottleneck effect) or when a group of individuals separates from a population (founder effect) where the effect of genetic drift appears greatly in small populations. In more detail, we are discussing the bottleneck effect vs founder effect:

Population bottleneck

What is the Bottleneck Effect? The bottleneck effect definition is the decrease in the number of individuals in a population due to a natural disaster, genetic drift bottleneck effect usually influences the genetic distribution among the population, therefore, the effect of genetic drift becomes more significant. Consequently, the genetic variation among this population will decrease as the number of mating individuals will decrease.

When genetic drift rate increases in a population this leads to the loss or the fixation of some alleles, this phenomenon is described in terms of a decrease in the genetic effective size. Even though populations experiencing a bottleneck may reproduce and become larger in size again, however, the genetic variation among these populations decline at a rate that represents the size of the disaster until new individuals are introduced into the population through migration or when new mutations occur. The bottleneck biology strength is affected by its size and duration, these factors are calculated mathematically to determine the influence of the bottleneck on the population's genetic variation.

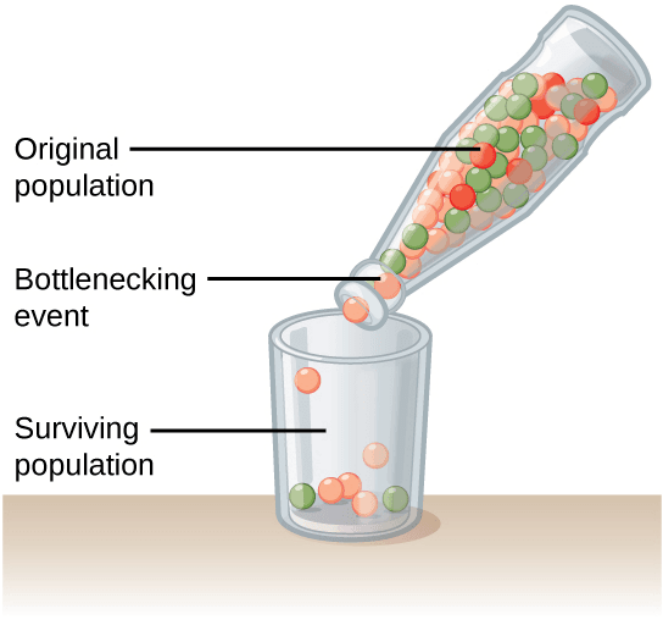


Figure 2: Bottleneck effect example: This population has been affected by a natural disaster where only some individuals survived, the effect of gene drifting will influence the surviving population for many generations (Clark, openstax). Credit: Clark, Openstax.org

Founder effect

The founder effect definition in biology is when a small population may be formed due to the founder effect when a small number of individuals leave their population to start a new colony, these individuals do not necessarily cover the whole genetic set of the population; therefore, gene drifting effect is significant within this small population. An example of founder effect is most commonly found among species of fungi where the spores disperse and colonize in different environments forming new colonies that may not have the same alleles as the population from which they were originated.

Genetic Drift vs Gene Flow

Gene flow is the flow of alleles from one generation to another by means of migration or dispersion, some populations do not usually experience migration or dispersion while others are more flexible, for example, plants and fungi send their pollens or spores away from their population to colonize in different environments. Even though some populations may seem stable, however, they are not as stable as they seem, such as lions that leave their mothers after development to search for a female that is not related to their population. This flow of genes

among populations contribute to the change of the gene pool of every population as well as the introduction of new genes to continue the evolutionary process.

Gene drift is counteracted by gene flow since a population does not usually stay small for a long time to be influenced by genetic drift. However, gene flow can counteract the effect of genetic drift only if the population's gene flow is sufficient to increase the frequency of alleles lost by gene drifting. Gene flow may occur as a result of passive seed dispersal or active migration, studies have shown that only one migrant per generation can increase the genetic differentiation among the population it can also prevent the genetic drifting effect in decreasing the genetic variation among populations. Note that this rule applies only for ideal populations, while the nonideal populations may require more than one migrant to counteract the genetic drift effect.

Genetic Drift in Evolution

Genetic drift contributes to the increase or decrease of a certain allele in each population; therefore, the effect of genetic drift is canceled over the long run in the normal populations, however, the effect of genetic drift cannot be canceled if an allele frequency reached zero unless a mutation produced this allele again. Genetic drifting is important in evolution since it determines the fate of a mutation, it determines whether it will disappear or becomes fixed in the population after a few generations. For nonideal populations Small in size), genetic drift is important even for the common genes.

Normally, if an allele is fixed in one generation it is most likely to increase in the following generations. However, in terms of genetic drifting, what occurs in one generation does not necessarily occur in the following generations, so if one allele increases in one generation, it may increase or decrease in the next generations.

A subdivided population adaptation is a process consisting of two phases, the first phase is genetic drifting where the loss or fixation of some alleles randomly occurs by chance which in turn helps the population to explore new genes, the second phase is characterized by the natural selection of the most beneficial genes that were introduced in phase one, these genes are exported to other populations by migration. The genetic drift theory has a significant role in the evolutionary process of individuals where the balance between mutations and gene drifting creates a state in genetic variation. Since mutations introduce new alleles while gene drifting may eliminate or fix the new alleles.

Gene flow or Gene Migrations

In population genetics, gene flow (also known as gene migration or geneflow and allele flow) is the transfer of genetic material from one population to another. If the rate of gene flow is high enough, then two populations will have equivalent allele frequencies and therefore can be considered a single effective population. It has been shown that it takes only "one migrant per generation" to prevent populations from diverging due to drift.[1] Populations can diverge due to selection even when they are exchanging alleles, if the selection pressure is strong enough.[2][3] Gene flow is an important mechanism for transferring genetic diversity among populations. Migrants change the distribution of genetic diversity among populations, by modifying allele frequencies (the proportion of members carrying a particular variant of a gene). High rates of gene flow can reduce the genetic differentiation between the two groups, increasing homogeneity.[4] For this reason, gene flow has been thought to constrain speciation and prevent range expansion by combining the gene pools of the groups, thus preventing the development of differences in genetic variation that would have led to differentiation and adaption.[5] In some cases dispersal resulting in gene flow may also result in the addition of novel genetic variants under positive selection to the gene pool of a species or population (adaptive introgression.[6])

There are a number of factors that affect the rate of gene flow between different populations. Gene flow is expected to be lower in species that have low dispersal or mobility, that occur in fragmented habitats, where there is long distances between populations, and when there are small population sizes.[7][8] Mobility plays an important role in dispersal rate, as highly mobile individuals tend to have greater movement prospects. Although animals are thought to be more mobile than plants, pollen and seeds may be carried great distances by animals, water or wind. When gene flow is impeded, there can be an increase in inbreeding, measured by the inbreeding coefficient (F) within a population. For example, many island populations have low rates of gene flow due to geographic isolation and small population sizes. The Black Footed Rock Wallaby has several inbred populations that live on various islands off the coast of Australia. The population is so strongly isolated that lack of gene flow has led to high rates of inbreeding

Measuring gene flow

The level of gene flow among populations can be estimated by observing the dispersal of individuals and recording their reproductive success.[4][10] This direct method is only suitable for some types of organisms, more often indirect methods are used that infer gene flow by comparing allele frequencies among population samples.[1][4] The more genetically differentiated two populations are, the lower the estimate of gene flow, because gene flow has a homogenizing effect. Isolation of populations leads to divergence due to drift, while migration reduces divergence. Gene flow can be measured by using the effective population size (N_e) and the net migration rate per generation (m). Using the approximation based on the Island model, the effect of migration can be calculated for a population in terms of the degree of genetic differentiation(F_{ST}).[11] This formula accounts for the proportion of total molecular marker variation among populations, averaged over loci.[12] When there is one migrant per generation, the inbreeding coefficient (F_{ST}) equals 0.2. However, when there is less than 1 migrant per generation (no migration), the inbreeding coefficient rises rapidly resulting in fixation and complete divergence (F_{ST} = 1). The most common F_{ST} is < 0.25. This means there is some migration happening. Measures of population structure range from 0 to 1. When gene flow occurs via migration the deleterious effects of inbreeding can be ameliorated.[1]

$$F_{ST} = 1/(4N_e m + 1)$$

The formula can be modified to solve for the migration rate when F_{ST} is known: $Nm = ((1/F_{ST}) - 1)/4 = \frac{1-F_{ST}}{4*F_{ST}}$, Nm = number of migrants.[1]

Barriers to gene flow

Allopatric speciation

When gene flow is blocked by physical barriers, this results in Allopatric speciation or a geographical isolation that does not allow populations of the same species to exchange genetic material. Physical barriers to gene flow are usually, but not always, natural. They may include impassable mountain ranges, oceans, or vast deserts. In some cases, they can be artificial, man-made barriers, such as the Great Wall of China, which has hindered the gene flow of native plant populations.[13] One of these native plants, Ulmus pumila, demonstrated a lower prevalence of genetic differentiation than the plants Vitex negundo, Ziziphus jujuba, Heteropappus hispidus, and Prunus armeniaca whose habitat is located on the opposite side of the Great Wall of China where Ulmus pumila grows.[13] This is because Ulmus pumila has wind-pollination as its primary means of propagation and the latter-plants carry out pollination through insects.[13] Samples of the same species which grow on either side have been shown to have developed genetic differences, because there is little to no gene flow to provide recombination of the gene pools.

Sympatric speciation

Barriers to gene flow need not always be physical. Sympatric speciation happens when new species from the same ancestral species arise along the same range. This is often a result of a reproductive barrier. For example, two palm species of Howea found on Lord Howe Island were found to have substantially different flowering times correlated with soil preference, resulting in a reproductive barrier inhibiting gene flow.[14] Species can live in the same environment, yet show very limited gene flow due to reproductive barriers, fragmentation, specialist pollinators, or limited hybridization or hybridization yielding unfit hybrids. A cryptic species is a species that humans cannot tell is different without the use of genetics. Moreover, gene flow between hybrid and wild populations can result in loss of genetic diversity via genetic pollution, assortative mating and outbreeding. In human populations, genetic differentiation can also result from endogamy, due to differences in caste, ethnicity, customs and religion

Gene flow between populations

Horizontal gene flow

Horizontal gene transfer (HGT) refers to the transfer of genes between organisms in a manner other than traditional reproduction, either through transformation (direct uptake of genetic material by a cell from its surroundings), conjugation (transfer of genetic material between two bacterial cells in direct contact), transduction (injection of foreign DNA by a bacteriophage virus into the host cell) or GTA-mediated transduction (transfer by a virus-like element produced by a bacterium) .

Viruses can transfer genes between species. Bacteria can incorporate genes from dead bacteria, exchange genes with living bacteria, and can exchange plasmids across species boundaries.[30] "Sequence comparisons suggest recent horizontal transfer of many genes among diverse species including across the boundaries of phylogenetic 'domains'. Thus determining the phylogenetic history of a species can not be done conclusively by determining evolutionary trees for single genes."

Biologist Gogarten suggests "the original metaphor of a tree no longer fits the data from recent genome research". Biologists [should] instead use the metaphor of a mosaic to describe the different histories combined in individual genomes and use the metaphor of an intertwined net to visualize the rich exchange and cooperative effects of horizontal gene transfer.[32]

"Using single genes as phylogenetic markers, it is difficult to trace organismal phylogeny in the presence of HGT. Combining the simple coalescence model of cladogenesis with rare HGT events suggest there was no single last common ancestor that contained all of the genes ancestral to those shared among the three domains of life. Each contemporary molecule has its own history and traces back to an individual molecule cenancestor. However, these molecular ancestors were likely to be present in different organisms at different times."[33]

Hybridization

In some instances, when a species has a sister species and breeding capabilities are possible due to the removal of previous barriers or through introduction due to human intervention, species can hybridize and exchange genes and corresponding traits.[34] This exchange is not always clear-cut, for sometimes the hybrids may look identical to the original species phenotypically but upon testing the mtDNA it is apparent that hybridization has occurred. Differential hybridization also occurs because some traits and DNA are more readily exchanged than others, and this is a result of selective pressure or the absence thereof that allows for easier transaction. In instances in which the introduced species begins to replace the native species, the native species becomes threatened and the biodiversity is reduced, thus making this phenomenon negative rather than a positive case of gene flow that augments genetic diversity.[35] Introgression is the replacement of one species' alleles with that of the invader species. It is important to note that hybrids are sometime less "fit" than their parental generation,[36] and as a result is a closely monitored genetic issue as the ultimate goal in conservation genetics is to maintain the genetic integrity of a species and preserve biodiversity.

Examples

While gene flow can greatly enhance the fitness of a population, it can also have negative consequences depending on the population and the environment in which they reside. The effects of gene flow are context-dependent.

- **Fragmented Population:** fragmented landscapes such as the Galapagos Islands are an ideal place for adaptive radiation to occur as a result of differing geography. Darwin's finches likely experienced allopatric speciation in some part due to differing geography, but that doesn't explain why we see so many different kinds of finches on the same island. This is due to adaptive radiation, or the evolution of varying traits in light of competition for

resources. Gene flow moves in the direction of what resources are abundant at a given time.[37]

- **Island Population:** The marine iguana is an endemic species of the Galapagos Islands, but it evolved from a mainland ancestor of land iguana. Due to geographic isolation gene flow between the two species was limited and differing environments caused the marine iguana to evolve in order to adapt to the island environment. For instance, they are the only iguana that has evolved the ability to swim.
- **Human Populations:** In Europe *Homo sapiens* interbred with Neanderthals resulting in gene flow between these populations.[38] This gene flow has resulted in Neanderthal alleles in modern European population.[39] Two theories exist for the human evolution throughout the world. The first is known as the multiregional model in which modern human variation is seen as a product of radiation of *Homo erectus* out of Africa after which local differentiation led to the establishment of regional population as we see them now.[40][41] Gene flow plays an important role in maintaining a grade of similarities and preventing speciation. In contrast the single origin theory assumes that there was a common ancestral population originating in Africa of *Homo sapiens* which already displayed the anatomical characteristics we see today. This theory minimizes the amount of parallel evolution that is needed.[41]
- **Butterflies:** Comparisons between sympatric and allopatric populations of *Heliconius melpomene*, *H. cydno*, and *H. timareta* revealed a genome-wide trend of increased shared variation in sympatry, indicative of pervasive interspecific gene flow.[42]
- **Human-mediated gene flow:** The captive genetic management of threatened species is the only way in which humans attempt to induce gene flow in ex situ situation. One example is the giant panda which is part of an international breeding program in which genetic materials are shared between zoological organizations in order to increase genetic diversity in the small populations. As a result of low reproductive success, artificial insemination with fresh/frozen-thawed sperm was developed which increased cub survival rate. A 2014 study found that high levels of genetic diversity and low levels of inbreeding were estimated in the breeding centers.[43]
- **Plants:** Two populations of monkeyflowers were found to use different pollinators (bees and hummingbirds) that limited gene flow, resulting in genetic isolation, eventually producing two different species, *Mimulus lewisii* and *Mimulus cardinalis* .[44]

- **Sika deer:** Sika deer were introduced into Western Europe, and they reproduce easily with the native red deer. This translocation of Sika deer has led to introgression and there are no longer "pure" red deer in the region, and all can be classified as hybrids.[45]
- **Bobwhite quail:** Bobwhite quail were translocated from the southern part of the United States to Ontario in order to increase population numbers and game for hunting. The hybrids that resulted from this translocation was less fit than the native population and were not adapted to survive the Northern Winters

Genetic Load

Genetic load is the difference between the fitness of an average genotype in a population and the fitness of some reference genotype, which may be either the best present in a population, or may be the theoretically optimal genotype. The average individual taken from a population with a low genetic load will generally, when grown in the same conditions, have more surviving offspring than the average individual from a population with a high genetic load.[1][2] Genetic load can also be seen as reduced fitness at the population level compared to what the population would have if all individuals had the reference high-fitness genotype.[3] High genetic load may put a population in danger of extinction

Fundamentals

Consider n genotypes $\mathbf{A}_1, \ldots, \mathbf{A}_n$, which have the fitnesses $w_1, \ldots, w_n$ and frequencies $p_1, \ldots, p_n$, respectively. Ignoring frequency-dependent selection, the genetic load L may be calculated as:

$$L = \frac{w_{\max} - \bar{w}}{w_{\max}}$$

where $w_{\max}$ is either some theoretical optimum, or the maximum fitness observed in the population. In calculating the genetic load, $w_1 \ldots w_n$ must be actually found in at least a single copy in the population, and $\bar{w}$ is the average fitness calculated as the mean of all the fitnesses weighted by their corresponding frequencies:

$$\bar{w} = \sum_{i=1}^{n} p_i w_i$$

where the i^{th} genotype is $\mathbf{A}_i$ and has the fitness and frequency w_i and p_i respectively.

One problem with calculating genetic load is that it is difficult to evaluate either the theoretically optimal genotype, or the maximally fit genotype actually present in the population.[4] This is not a problem within mathematical models of genetic load, or for empirical studies that compare the relative value of genetic load in one setting to genetic load in another.

Causes

Deleterious mutation

Deleterious mutation load is the main contributing factor to genetic load overall.[5] Most mutations are deleterious, and occur at a high rate. The Haldane-Muller theorem of mutation–selection balance says that the load depends only on the deleterious mutation rate and not on the selection coefficient.[6] Specifically, relative to an ideal genotype of fitness 1, the mean population fitness is $\exp(-U)$ where U is the total deleterious mutation rate summed over many independent sites. The intuition for the lack of dependence on the selection coefficient is that while a mutation with stronger effects does more harm per generation, its harm is felt for fewer generations.

A slightly deleterious mutation may not stay in mutation–selection balance but may instead become fixed by genetic drift when its selection coefficient is less than one divided by the effective population size.[7] In asexual populations, the stochastic accumulation of mutation load is called Muller's ratchet, and occurs in the absence of beneficial mutations, when after the most-fit genotype has been lost, it cannot be regained by genetic recombination. Deterministic accumulation of mutation load occurs in asexuals when the deleterious mutation rate exceeds one per replication.[8] Sexually reproducing species are expected to have lower genetic loads.[9] This is one hypothesis for the evolutionary advantage of sexual reproduction. Purging of deleterious mutations in sexual populations is facilitated by synergistic epistasis among deleterious mutations.[10]

High load can lead to a small population size, which in turn increases the accumulation of mutation load, culminating in extinction via mutational meltdown.[11][12]

The accumulation of deleterious mutations in humans has been of concern to many geneticists, including Hermann Joseph Muller,[13] James F. Crow,[10] Alexey Kondrashov,[14] W. D. Hamilton,[15] and Michael Lynch.[16]

Beneficial mutation

New beneficial mutations create fitter genotypes than those previously present in the population.[citation needed] When load is calculated as the difference between the fittest genotype present and the average, this creates a substitutional load. The difference between the theoretical maximum (which may not actually be present) and the average is known as the "lag load".[17] Motoo Kimura's original argument for the neutral theory of molecular evolution was that if most differences between species were adaptive, this would exceed the speed limit to adaptation set by the substitutional load.[18] However, Kimura's argument confused the lag load with the substitutional load, using the former when it is the latter that in fact sets the maximal rate of evolution by natural selection.[19]

More recent "travelling wave" models of rapid adaptation derive a term called the "lead" that is equivalent to the substitutional load, and find that it is a critical determinant of the rate of adaptive evolution

Inbreeding

Inbreeding increases homozygosity. In the short run, an increase in inbreeding increases the probability with which offspring get two copies of a recessive deleterious alleles, lowering

fitnesses via inbreeding depression.[22] In a species that habitually inbreeds, e.g. through self-fertilization, recessive deleterious alleles are purged.[23][24]

Recombination/segregation

Combinations of alleles that have evolved to work well together may not work when recombined with a different suite of coevolved alleles, leading to outbreeding depression. Segregation load is the presence of underdominant heterozygotes (i.e. heterozygotes that are less fit than either homozygote). Recombination load arises through unfavorable combinations across multiple loci that appear when favorable linkage disequilibria are broken down. Recombination load can also arise by combining deleterious alleles subject to synergistic epistasis, i.e. whose damage in combination is greater than that predicted from considering them in isolation.

Migration

Migration load is the result of nonnative organisms that aren't adapted to a particular environment coming into that environment. If they breed with individuals who are adapted to that environment, their offspring will not be as fit as they would have been if both of their parents had been adapted to that particular environment.[27][28][29] Migration load can also occur in asexually reproducing species, but in this case, purging of low fitness genotypes is more straightforward.

Inbreeding

Inbreed is the production of offspring from the mating or breeding of individuals or organisms that are closely related genetically.[2] By analogy, the term is used in human reproduction, but more commonly refers to the genetic disorders and other consequences that may arise from expression of deleterious or recessive traits resulting from incestuous sexual relationships and consanguinity.

Inbreeding results in homozygosity, which can increase the chances of offspring being affected by deleterious or recessive traits.[3] This usually leads to at least temporarily decreased biological fitness of a population[4][5] (called inbreeding depression), which is its ability to survive and reproduce. An individual who inherits such deleterious traits is colloquially referred to as inbred. The avoidance of expression of such deleterious recessive alleles caused by inbreeding, via inbreeding avoidance mechanisms, is the main selective reason for outcrossing.[6][7] Crossbreeding between populations also often has positive effects on fitness-

related traits,[8] but also sometimes leads to negative effects known as outbreeding depression. However increased homozygosity increases probability of fixing beneficial alleles and also slightly decreases probability of fixing deleterious alleles in population.[9] Inbreeding can result in purging of deleterious alleles from a population through purifying selection.

Inbreeding is a technique used in selective breeding. For example, in livestock breeding, breeders may use inbreeding when trying to establish a new and desirable trait in the stock and for producing distinct families within a breed, but will need to watch for undesirable characteristics in offspring, which can then be eliminated through further selective breeding or culling. Inbreeding also helps to ascertain the type of gene action affecting a trait. Inbreeding is also used to reveal deleterious recessive alleles, which can then be eliminated through assortative breeding or through culling. In plant breeding, inbred lines are used as stocks for the creation of hybrid lines to make use of the effects of heterosis. Inbreeding in plants also occurs naturally in the form of self-pollination.

Inbreeding can significantly influence gene expression which can prevent inbreeding depression.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF BIO AND CHEMICAL ENGINEERING

DEPARTMENT OF BIOINFORMATICS

UNIT – 4- SBIA1302 – Evolutionary Biology

GENOME EXPANSION AND RESTRUCTURING

How does genome restructuring occur? What processes result in changes in genome size? Several mechanisms are thought to play a role in large-scale changes in genome architecture, including those that shuffle genotypes and, thus, alter the structure of chromosomes, as well as processes that result in addition or relocation of new DNA sequences in the genome. Recombination

Recombination, the repair of double-stranded breaks (DSBs) in DNA, has important influences on organismal evolution, including both generating and reducing genetic variation (see chapter IV.4). DSBs may be incurred exogenously through exposure to environmental agents

at any point during an organism's life cycle, or endogenously during meiosis in eukaryotes. Repair of these breaks frequently involves using a homologous piece of DNA as template—typically a sister chromatid or homologous chromosome. When a reciprocal exchange of DNA between homologous chromosomes occurs, it is referred to as a *crossover event*. Efficient repair of these breaks is critical, because their presence will disrupt replication and transcription. Errors in recombination can be devastating to an organism, but when they occur in the germ line, they also provide heritable restructuring events in genomes that contribute to genomic evolution in eukaryotes.

Repetitive elements and self-replicating mobile elements are present throughout eukaryotic genomes. Because these elements have multiple homologous templates in the genome, recombination can potentially occur between any two, even if located on different chromosomes. When ectopic recombination occurs between these elements as a result of their sequence similarity, large-scale chromosomal rearrangements can occur, including sequence duplications, deletions, or inversions of large sections of chromosomes, and translocation of a chromosomal section from one chromosome to another. These changes can disrupt protein-coding sequences directly, as well as remove or add regulatory sequences that can result in aberrant expression of genes.

During meiosis, DSBs are induced and repaired in a process mediated by a cell's machinery. Because of the inherent risk associated with the formation of these breaks, it is unsurprising that meiotic recombination appears a tightly regulated and evolutionarily constrained process. More unexpected are the constraints limiting the number of these breaks that result in a crossover event. Furthermore, these crossovers do not occur equally across the genome; rather, they are concentrated at *hot spots*, where rates of recombination are higher by several orders of magnitude than their flanking genome regions, or *cold spots*, in which crossover rates are extremely low. These hot spots are rapidly evolving and dynamic. Organisms as closely related as humans and chimpanzees share no overlap in the genomic locations of hot spots, and intraspecific variation has even been observed within humans. Despite this fast rate of evolution of hot spots, there is mounting evidence that their location is sometimes dictated by specific sequence motifs. In the fission yeast *Schizosaccharomyces pombe*, several discrete sequences seven base pairs in length have been identified at active hot spots. In humans, one degenerate thirteen base-pair motif has been characterized at 41 percent of identified hot spots. Furthermore, in humans, the transcription factor Prdm9 is required for activation of these hot spots, and the amino acids that interact with the thirteen base-pair motif are under strong positive selection. Prdm9 may therefore act as a driver in hot spot evolution, or may be evolving rapidly in response to changes in hot-spot sequence motifs.

This observed specificity in DNA sequence at active hot spots is puzzling. If a specific sequence is required for increased recombination, that sequence should also be lost as a result of the very recombination that it induces. A model was recently proposed to explain this apparent paradox. If a specific sequence is required for hot-spot activation, then a single base-pair change will inactivate it. Conversely, there are many sites in the genome that require a single base-pair change in order to become an activated hot spot; therefore, an evolutionary equilibrium may exist in which hot spots are degraded and introduced through these single base-pair changes. This explanation, coupled with the rapid evolution of hot-spot activators like Prdm9, may explain the dynamic nature of genomic hot-spot locations even in very closely related organisms.

Transposable Elements

An astonishingly large fraction of many eukaryotic genomes is composed of mobile DNA elements. These self-replicating pieces of DNA, which frequently contain their own protein-coding and regulatory sequences, make up about 50 percent of the human genome. Generally, there are two classes of mobile elements characterized primarily by their mode of replication. First there are DNA transposons, which replicate through a "cut-and-paste" mechanism, in which an enzyme (transposase)—which may be encoded by the transposon itself or by a separate transposable element—excises the DNA sequence prior to its insertion into a new genomic location. Proliferation of these elements relies on the horizontal transfer of new elements from one organism to another, such as the transmission of small circular chromosomes containing the elements between prokaryotes. The second class of mobile elements is collectively referred to as *retrotransposons*. These elements replicate by "copy-and-paste" mechanisms, in which an RNA intermediate is produced and reverse transcribed (by a retrotransposon-encoded reverse transcriptase) before insertion. Such elements can proliferate horizontally, as described for DNA transposons, as well as vertically, when they proliferate within cells in the germ line and can then be transmitted to the next generation.

It is easy to see how the replication and insertion of mobile elements in a host genome could be slightly or strongly deleterious. For example, transposon insertion into a protein-coding region would most likely result in a frameshift, premature stop codon, or otherwise-aberrant protein sequence. Potentially, for this reason, a host has mechanisms to defend its genome from such elements.

These include transcriptional silencing of the elements by chromatin modifications and transcription of small interfering RNA molecules that target mRNA produced by the element for destruction, thus depriving the transposable element of the machinery it needs for proliferation. Because successful proliferation of a mobile element depends on the success of a host genome, it is also possible that transposable elements have built-in self-regulatory mechanisms preventing them from uncontrolled proliferation that would drive a host to extinction; however, such mechanisms have not been characterized.

Importantly, as with all forms of mutation, mobile element insertion can on rare occasion give rise to evolutionary novelty. Because mobile elements encode their own machinery, multiple consequences can arise following their insertion into a new location. First, the elements contain protein-coding sequences and thus can introduce new coding regions into the genome (see chapter V.6). Second, these protein-coding sequences in mobile elements frequently have their own regulatory elements that can modify gene expression patterns of sequences, especially when adjacent to the insertion site. For example, the promoter region of a gene in the transposable element may recruit transcriptional machinery to a location near a host gene that has tight temporal or spatial regulation, causing it to be transcribed when it is normally silent. Indeed, it is hypothesized that centromeres and telomeres are often derived from mobile elements, and in some cases (e.g., *Drosophila*), mobile elements provide a mechanism for telomere maintenance. Finally, there is also evidence that mobile elements play a role in DNA double-strand break repair by using double-strand breaks as sites of insertion.

Noncoding Elements

Noncoding DNA sequences are those that do not determine a functional product. This chapter will consider the evolution of two types of noncoding elements, untranslated regions (UTRs) and introns. UTRs are parts of genes that are transcribed but not translated into an amino acid sequence, and are found both preceding the translation initiation site (5′ UTRs) and following the termination of translation (3′ UTRs). The addition of 5′ UTRs to eukaryotic genes is a risky prospect when the potential inclusion of an alternative translation initiation site is considered. Mutation of the 5′ UTR to contain such a site could have dramatic effects on the resulting amino acid sequence, resulting in a nonfunctional product. Because of this increased mutation risk, it is not clear what, if any, advantage eukaryotes gain through the addition of 5′ UTRs, but their presence and length are consistent across eukaryotic diversity. Although the addition of 3′ UTRs to eukaryotic genes does not appear to carry the same risks as 5′ UTRs, these elements are important in several aspects of mRNA regulation. The 3′ UTRs are critical for mRNA stability and nuclear export, and they have important regulatory functions in several aspects of translation. It is likely these features arose subsequent to the evolution of the 3′ UTR itself; therefore they cannot provide an explanation for the addition of this element.

The mechanisms for evolution and origins of introns are much better understood than 5′ UTRs. Despite the similarity in the length and number of protein-coding genes across eukaryotic diversity, there is substantial variation in the amount of intronic DNA. In eukaryotes, introns in nuclear genes (spliceosomal introns) are processed by a nucleoprotein complex—the spliceosome—which is present in all eukaryotes and thus likely present in the most recent eukaryotic ancestor. In humans, an average gene contains 7.7 introns, with an average intron length of 4.66 kilobases (kb). Compared to the average length of a human exon sequence (0.15 kb), it is clear that the total length of a human (and in general any eukaryotic) gene is dominated by introns. This density of introns allows for a large number of potential transcripts per locus through alternative splicing, which in humans is responsible for the average 2.6 transcripts produced per gene. Although the current importance of introns is at least partly understood (alternative splicing, regulatory element content, etc.), the origin and evolutionary mechanisms responsible for the proliferation of introns in eukaryotes remains unclear.

Debate over spliceosomal intron origin has been divided into two camps: those that propose the early evolution of introns prior to the divergence of eukaryotes and prokaryotes, and those that posit a later origin exclusively in eukaryotes. The resolution of this debate rests primarily on the hypothesized relationship of eukaryotic spliceosomal introns with the self-splicing group II introns found in some prokaryotes, which some argue are homologous. Whether spliceosomal introns arose early or late, there has been massive divergence in intron content in eukaryotes, making our understanding of the mechanisms underlying intron gain and loss of great importance.

Both intron loss and gain can be mediated by recombination, with intron loss hypothesized to result from replacement of a genomic gene copy with a reverse transcribed mRNA transcript of that gene (see chapter V.6), while hypotheses for mechanisms of intron gain include ectopic insertion of DNA fragments during an alternative DNA repair mechanism known as *nonhomologous end joining* (NHEJ). During NHEJ, fragments of DNA with very little sequence identity (microhomology) may be joined to repair DSBs, and aberrant insertion of a DNA fragment within a coding region may

explain the origin of novel introns. NHEJ may be an intron loss mechanism as well, if microhomology between an intron's splice junctions is used for repair. Consistent with this, species that are intron poor have high conservation of their splice sites, whereas intron-rich species have more degeneracy in their intron splice sites. The hypothesized role of NHEJ in intron gain is supported by the observation that intron-rich species use NHEJ more frequently during DNA repair.

3. DRIVERS OF GENOME EVOLUTION

A challenge that remains for our understanding of genome evolution is explaining how the addition of DNA to the genome and the existence of more complex genomic elements are possible. The presence of these elements is inherently risky, as they provide additional locations at which deleterious mutations can occur. For example, the addition of an intron to a protein-coding region now adds splice junctions, a branch point, and other regulatory elements that are evolutionarily constrained. One could argue that such genomic complexity is necessary for the evolution of organismal complexity; however, the diversity in content of these complex elements suggests otherwise. Adaptive and neutral arguments for the evolution of genomic complexity are described below.

Adaptive Evolution

What evolutionary pressures might be acting on genome size? Some data suggest that the forces may be mutation bias, such that small (<400 kb) deletions occur more frequently than insertions, resulting in reduction in genome size over time. For example, work performed in *Caenorhabditis elegans* demonstrated that at genomic sites not under selective constraints (i.e., pseudogenes), the rate of deletion was 2.8-fold higher than the rate of insertion. These data offer an explanation for the relatively compact size of the *C. elegans* genome, and suggest a more generalizable trend of deletions outnumbering insertions: selective pressures may favor a smaller genome.

Some other, less generally supported hypotheses suggest selective pressures might underlie the evolution of genome size as a result of the phenotypic consequences of these differences, primarily the effect of genome size on cell size. For example, the *nucleoskeletal hypothesis* proposes that increasing genome size requires an increase in the size of the nucleus, which coevolves with cell size. According to this hypothesis, a larger cell has greater requirements for transcription and translation, and thus requires a larger nucleus and genome to meet its needs; however, the nucleoskeletal hypothesis does not account for accommodation of a larger cell's needs through increased rates of transcript production as opposed to increased DNA content.

Because beneficial outcomes are extremely unlikely for the majority of transposable element insertions, adaptive hypotheses for the existence of these elements can be excluded for the most part. These elements are more frequently thought of as parasitic or selfish because of their lack of dependence on host machinery for replication, and their likely detrimental effects on host fitness. The role of mobile elements in genome evolution is therefore referred to as the *selfish DNA hypothesis*, which posits that genome expansion is mediated by proliferation of mobile elements, and that such elements will spread until the point at which their impact on host fitness is so great that natural selection prohibits their further proliferation. This hypothesis does not account for the role of other elements present in eukaryotic genomes, such as introns, and therefore cannot fully explain the increased genome size in eukaryotes.

There are also several hypotheses for adaptive mechanisms underlying intron evolution in eukaryotes. First, large introns within genes increase the likelihood that incorrect splicing will result in the introduction of a premature stop codon that will be recognized early and will result in the degradation of the mRNA—a process known as nonsense-mediated decay. Second, the presence of one or more introns allows for alternative splicing to occur, in which introns can be excised or retained, exons can be skipped, or exon length can vary depending on the usage of specific splice junctions. This diversity in mRNA products from a single locus greatly increases the number of potential protein products resulting from that locus and allows for increased variation and complexity in molecular pathways (see chapter V.3). Further, the modular nature of genes that result from the inclusion of introns may have allowed for exon shuffling, in which mixing of domains from several different genes gives rise to genes with novel functions (see chapter V.6).

Neutral Evolution

Because eukaryotic genome expansion likely gave rise to sources of vast phenotypic novelty, it is tempting to develop adaptive hypotheses for their origination, such as those described above. However, the main explanation for the existence of these novel features may lie in *neutral evolutionary processes*—those that result from changes that have little or no effect on host fitness, but arise and become fixed in a population through genetic drift. A general framework for understanding the origin and evolution of such genomic novelties was proposed by Arlin Stoltzfus in 1999, which he called "constructive neutral evolution." Expanding on this, Michael Lynch proposed a synthetic hypothesis that posits neutral

processes as being largely responsible for the origin of genomic elements that,in turn, gave rise to the expanded genome size observed in eukaryotes. Since eukaryotic cells are typically much larger than prokaryotic cells, which generally result in much smaller population sizes for eukaryotes, the effects of genetic drift are amplified, making it much more likely that neutral or

even slightly deleterious mutations—including unusual genetic features—will become fixed in a population (see chapter IV.1). As described above,incorporation of the featuresmost responsible for increased genome size would have been a very risky prospect for early eukaryotes. In particular, noncoding elements like introns and UTRs dramatically increase the number of sites at which deleterious mutation may occur. Further, the origin of these elements would have been extremely dangerous, as their addition would interrupt protein-coding regions, potentially causing frameshifts, premature stop codons, or alternative translation start sites. The inclusion of these elements therefore would likely have immediate deleterious effects on an organism, or at best would not confer an immediate benefit to be acted on by natural selection. Instead, neutral processes may account for the initial fixation of these features in early eukaryotic populations. Small eukaryotic populations increased the impact of genetic drift and reduced the efficacy of selection such that these genomic elements could become fixed despite not conferring an advantage on a cell. Any beneficial effects these elements currently have were therefore subsequently acquired and may contribute to their maintenance in a population, but adaptive arguments are unlikely to explain their original fixation in eukaryotes. Our understanding of the evolution of genome structure and content has substantially improved in the past decade. Fast-moving advances in DNA sequencing technologies have provided unfettered access to complete genomes from across the entire tree of life. Decoding the content of these genomes has been only a first step in understanding their biology. A deeper and more satisfying view of genome biology is emerging in which genomes are not only repositories of genes but also evolving entities with emergent and sometimes-unusual properties that are increasingly explicable within a solid theoretical framework.

Evolution of sex chromosomes

Sex chromosome evolution (a) Genetic sex determination and recombination suppression The accepted theory of the evolution of heteromorphic sex chromosomes (figure 1) starts with a pair of homologous autosomes that gain a major sexdetermining function through one or several genes [2,3,8]. This can happen in a system that already has a sex chromosome pair (and in that case it results in a so-called turnover, figure 1c,d) or in a hermaphrodite ancestor [2]. Two mutations are needed in order for separate sexes to evolve from hermaphroditism—one suppressing male fertility and the other suppressing female fertility, usually at different loci—

otherwise a mixed mating system results (e.g. gynodioecy with females and hermaphrodites, which is the most common mixed system in plants) [8–10]. In case of a turnover, the new sex-determining gene needs to cause a fitness increase compared to the old sex-determining gene in order to invade [1]. Next, sex-specific genes become linked to the sexdetermining region, and suppression of recombination evolves in the heterozygous sex since it is advantageous for these genes to be inherited together [2]. Recombination between the proto-X and proto-Y sex chromosomes (protoZ and -W in female heterogametic systems) can be hindered either through gradual reduction with genetic modifiers or large inversions [8]. The recombination suppression region of the proto-sex chromosomes can expand further via the accumulation of sexually antagonistic genes (i.e. genes that are beneficial for one sex but detrimental for the other), near the sex-determining region [8,11].

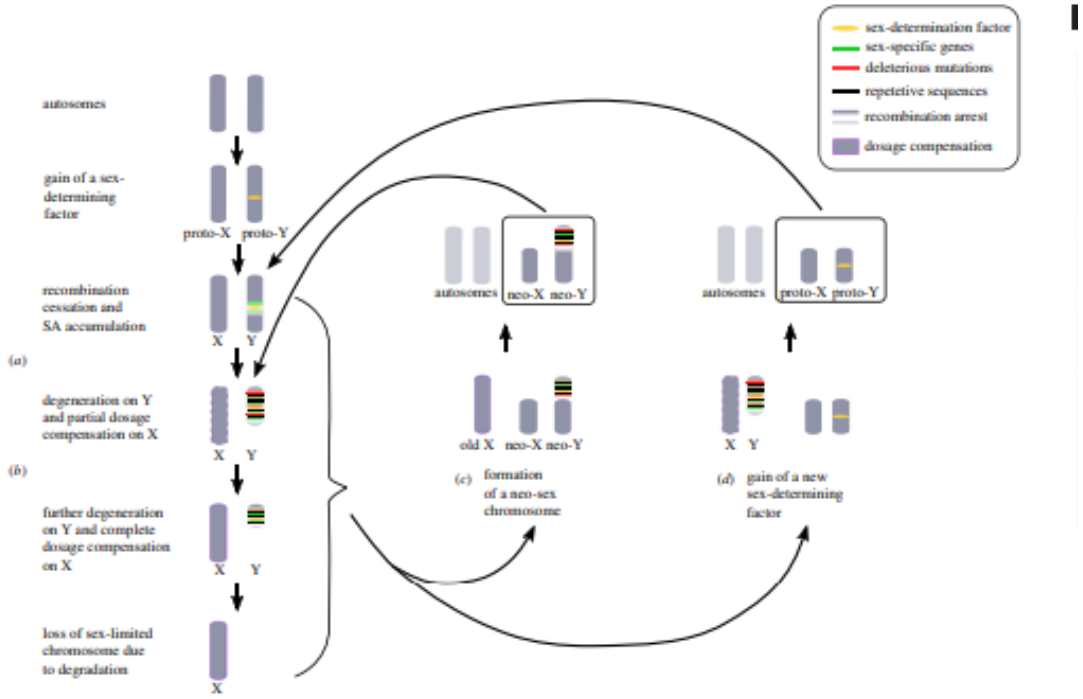


Figure 1. Overview of the dynamic evolution of sex chromosomes, illustrated in a male heterogametic system. Top left corner: an autosome pair in a hermaphrodite gains a sex-determining factor that evolves to become a highly heteromorphic pair of sex chromosomes, via cessation of recombination, degeneration (*a*) and evolution of dosage compensation (*b*). This progression can however be perturbed by a turnover event, such as the formation of a neo-sex chromosome (*c*) or a gain of a new sex-determining factor (*d*). In (*c*), the moderately degenerated Y chromosome fuses with an existing autosome, forming a new sex chromosome pair with an old sex-determining factor. In (*d*), an autosomal pair gains a new sex-determining factor, creating a completely new sex chromosome pair. The old Y is lost. In both (*c*) and (*d*), the old X may eventually gain diploidy through non-disjunction and subsequently lose dosage compensation, becoming an ordinary autosome pair. Figure adapted from [7]. Note that although (*c*) and (*d*) are shown as leading to chromosome turnovers, this progression is not inevitable. SA, sexually antagonistic allele.

(b) Degeneration and dosage compensation

The increase of the non-recombining region results in strongly differentiated sex chromosomes, as genes decay via accumulation of deleterious mutations on the sex-limited Y chromosome [2,11]. Following Y degeneration (figure 1a), the homogametic sex (XX females) will have two copies of X-linked genes compared to the heterogametic sex's (XY males) one, resulting in unequal expression between the sexes. The solution is dosage compensation (figure 1b), which can be achieved in multiple ways (e.g. X chromosome inactivation in female mammals [12], or X hyperexpression in male Drosophila [12,13]). Dosage compensation is a common phenomenon taxonomically, but varies in its extent; it is almost complete in mammals, but is partial in birds and some snakes.

(c) Sex chromosome turnovers Though some organisms have lost the Y chromosome completely (e.g. crickets and dragonflies), not all sex chromosomes end up highly differentiated [1,2]. There are two main hypotheses: occasional recombination between X and Y due to sex-reversals and frequent turnover events. Sex chromosomes in sex-reversed female frogs (i.e. with an XY genotype) recombine as much as in XX-females, introducing new genetic variance on the Y [14]. However, this only works for species with relatively undifferentiated sex chromosomes—strongly differentiated sex chromosomes cannot recombine successfully [14]. Sex chromosome turnovers are very common in fishes and may result from the evolution of a new sex-determining gene on an autosome or transposition of a sex-determining locus to an autosome (figure 1d), or fusions between autosomes old X autosomes neo-X neo-Y neo-X neo-Y X X Y X Y Y (a) (b) gain of a new sex-determining factor autosomes proto-X proto-Y sex-determination factor sex-specific genes deleterious mutations repetetive sequences recombination arrest dosage compensation formation of a neo-sex chromosome proto-X proto-Y X Y autosomes gain of a sexdetermining factor recombination cessation and SA accumulation degeneration on Y and partial dosage compensation on X further degeneration on Y and complete dosage compensation on X loss of sex-limited chromosome due to degradation X (c) (d) Figure 1. Overview of the dynamic evolution of sex chromosomes, illustrated in a male heterogametic system. Top left corner: an autosome pair in a hermaphrodite gains a sex-determining factor that evolves to become a highly heteromorphic pair of sex chromosomes, via cessation of recombination, degeneration (a) and evolution of dosage compensation (b). This progression can however be perturbed by a turnover event, such as the formation of a neo-sex chromosome (c) or a gain of a

new sex-determining factor (d). In (c), the moderately degenerated Y chromosome fuses with an existing autosome, forming a new sex chromosome pair with an old sex-determining factor. In (d), an autosomal pair gains a new sex-determining factor, creating a completely new sex chromosome pair. The old Y is lost. In both (c) and (d), the old X may eventually gain diploidy through non-disjunction and subsequently lose dosage compensation, becoming an ordinary autosome pair. Figure adapted from [7]. Note that although (c) and (d) are shown as leading to chromosome turnovers, this progression is not inevitable. SA, sexually antagonistic allele. rspb.royalsocietypublishing.org Proc. R. Soc. B 284: 20162806 2 Downloaded from https://royalsocietypublishing.org/ on 08 July 2021 and existing sex chromosomes (formation of a neo-sex chromosome; figure 1c) [15]. (d) Our changing views of sex chromosomes Although most research has been carried out on highly heteromorphic sex chromosomes, we do know that sex chromosomes are diverse across living organisms, from the mammal XY and bird ZW to the less-studied haploid UV sex chromosomes (found in e.g. bryophytes [3,16,17]). We also know that there is a large variation in the level of degeneration of heteromorphic sex chromosomes, a variety of mechanisms of dosage compensation and a high frequency of sex chromosome turnovers in some groups but not others [1], making general patterns in sex chromosome evolution far from ‘general’ [1]. This is a relatively recent insight stemming from the explosion of sequencing technologies (see below) and suggests that our theories of sex chromosome evolution have likely been biased towards mammalian-style XY systems and shaped largely by studies of model organisms [18]. We, therefore, argue that a historic overview of key findings and empirical discoveries will put current thinking into context and help us better understand where to go next. To this end, we have compiled a timeline of sex chromosome evolution research (table 1), which illustrates the progress over time of our understanding of various stages in sex chromosome evolution. Although the points we include are inevitably somewhat subjective, we have attempted to cover all major discoveries in the evolution of sex chromosomes.

Gene Duplication

Gene duplication is the process by which a region of DNA coding for a gene is copied. Gene duplication can occur as the result of an error in recombination or through a retrotransposition event. Duplicate genes are often immune to the selective pressure under which genes normally exist. This can result in a large number of mutations accumulating in the duplicate gene code. This may render the gene non-functional or in some cases confer some benefit to the organism. There are multiple mechanisms by which gene duplication can occur.

Ectopic Recombination

Duplications can arise from unequal crossing-over that occurs during meiosis between misaligned homologous chromosomes. The product of this recombination is a duplication at the site of the exchange and a reciprocal deletion. Ectopic recombination is typically mediated by sequence similarity at the duplicate breakpoints, which form direct repeats. Repetitive genetic elements, such as transposable elements, offer one source of repetitive DNA that can facilitate recombination, and they are often found at duplication breakpoints in plants and mammals.

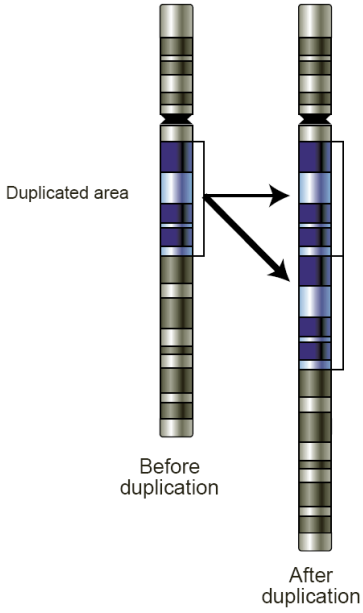


Figure 18.4D.118.4D.1: **Gene Duplication**: This figure indicates a schematic of a region of a chromosome before and after a duplication event. Ectopic recombination is typically mediated by sequence similarity at the duplicate breakpoints, which form direct repeats.

Replication Slippage

Replication slippage is an error in DNA replication, which can produce duplications of short genetic sequences. During replication, DNA polymerase begins to copy the DNA, and at some point during the replication process, the polymerase dissociates from the DNA and replication stalls. When the polymerase reattaches to the DNA strand, it aligns the replicating strand to an incorrect position and incidentally copies the same section more than once. Replication slippage is also often facilitated by repetitive sequence but requires only a few bases of similarity.

Retrotransposition

During cellular invasion by a replicating retroelement or retrovirus, viral proteins copy their genome by reverse transcribing RNA to DNA. If viral proteins attach irregularly to cellular mRNA, they can reverse-transcribe copies of genes to create retrogenes. Retrogenes usually lack intronic sequence and often contain poly A sequences that are also integrated into the genome. Many retrogenes display changes in gene regulation in comparison to their parental gene sequences, which sometimes results in novel functions.

Aneuploidy

Aneuploidy occurs when nondisjunction at a single chromosome results in an abnormal number of chromosomes. Aneuploidy is often harmful and in mammals regularly leads to spontaneous abortions. Some aneuploid individuals are viable. For example, trisomy 21 in humans leads to Down syndrome, but it is not fatal. Aneuploidy often alters gene dosage in ways that are detrimental to the organism and therefore, will not likely spread through populations.

Gene duplication as an evolutionary event

Gene duplications are an essential source of genetic novelty that can lead to evolutionary innovation. Duplication creates genetic redundancy and if one copy of a gene experiences a mutation that affects its original function, the second copy can serve as a ‘spare part’ and continue to function correctly. Thus, duplicate genes accumulate mutations faster than a functional single-copy gene, over generations of organisms, and it is possible for one of the two copies to develop a new and different function. This is an examples of neofunctionalization.

Gene duplication is believed to play a major role in evolution; this stance has been held by members of the scientific community for over 100 years. It has been argued that gene duplication is the most important evolutionary force since the emergence of the universal common ancestor.

Another possible fate for duplicate genes is that both copies are equally free to accumulate degenerative mutations, so long as any defects are complemented by the other copy. This leads to a neutral “subfunctionalization” model, in which the functionality of the original gene is distributed among the two copies. Neither gene can be lost, as both now perform important non-redundant functions, but ultimately neither is able to achieve novel functionality.

Subfunctionalization can occur through neutral processes in which mutations accumulate with no detrimental or beneficial effects. However, in some cases subfunctionalization can occur with clear adaptive benefits. If an ancestral gene is pleiotropic and performs two functions, often times neither one of these two functions can be changed without affecting the other function. In this way, partitioning the ancestral functions into two separate genes can allow for adaptive specialization of subfunctions, thereby providing an adaptive benefit.

Divergence

Genetic divergence is the process in which two or more populations of an ancestral species accumulate independent genetic changes through time, often after the populations have become reproductively isolated for some period of time. In some cases, subpopulations living in ecologically distinct peripheral environments can exhibit genetic divergence from the remainder of a population, especially where the range of a population is very large. The genetic differences among divergent populations can involve silent mutations (that have no effect on the phenotype) or give rise to significant morphological and/or physiological changes. Genetic divergence will always accompany reproductive isolation, either due to novel adaptations via selection and/or due to genetic drift, and is the principal mechanism underlying speciation.

Genetic drift or allelic drift is the change in the frequency of a gene variant (allele) in a population due to random sampling. The alleles in the offspring are a sample of those in the parents, and chance has a role in determining whether a given individual survives and reproduces. A population's allele frequency is the fraction of the copies of one gene that share a particular form. Genetic drift may cause gene variants to disappear completely and thereby reduce genetic variation. When there are few copies of an allele, the effect of genetic drift is larger, and when there are many copies the effect is smaller. These changes in gene frequency can contribute to divergence.

Divergent evolution is usually a result of diffusion of the same species to different and isolated environments, which blocks the gene flow among the distinct populations allowing differentiated fixation of characteristics through genetic drift and natural selection.Divergent evolution can also be applied to molecular biology characteristics. This could apply to a pathway in two or more organisms or cell types. This can apply to genes and proteins, such as nucleotide sequences or protein sequences that are derived from two or more homologous genes. Both orthologous genes (resulting from a speciation event) and paralogous genes (resulting from gene duplication within a population) can be said to display divergent evolution.

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF BIO AND CHEMICAL ENGINEERING

DEPARTMENT OF BIOINFORMATICS

UNIT – 5- SBIA1302 – Evolutionary Biology

Epigenetics

Epigenetics is the study of heritable changes in gene expression (active versus inactive genes) that do not involve changes to the underlying DNA sequence — a change in phenotype without a change in genotype — which in turn affects how cells read the genes. Epigenetic change is a regular and natural occurrence but can also be influenced by several factors including age, the environment/lifestyle, and disease state. Epigenetic modifications can manifest as commonly as the manner in which cells terminally differentiate to end up as skin cells, liver cells, brain cells, etc. Or, epigenetic change can have more damaging effects that can result in diseases like cancer. At least three systems including DNA methylation, histone modification and non-coding RNA (ncRNA)-associated gene silencing are currently considered to initiate and sustain epigenetic change.**1** New and ongoing research is continuously uncovering the role of epigenetics in a variety of human disorders and fatal diseases.

How Does Epigenetics Work?
Epigenetic changes affect gene expression in different ways. Types of epigenetic changes include:
DNA Methylation
DNA methylation works by adding a chemical group to DNA. Typically, this group is added to specific places on the DNA, where it blocks the proteins that attach to DNA to "read" the gene. This chemical group can be removed through a process called demethylation. Typically, methylation turns genes "off" and demethylation turns genes "on."

Histone modification
DNA wraps around proteins called histones. DNA wrapped tightly around histones cannot be accessed by proteins that "read" the gene. Some genes are wrapped around histones and are turned "off" while some genes are not wrapped around histones and are turned "on." Chemical groups can be added or removed from histones and change whether a gene is unwrapped or wrapped ("on" or "off").

Non-coding RNA
Your DNA is used as instructions for making coding and non-coding RNA. Coding RNA is used to make proteins. Non-coding RNA helps control gene expression by attaching to coding RNA, along with certain proteins, to break down the coding RNA so that it cannot be used to make proteins. Non-coding RNA may also recruit proteins to modify histones to turn genes "on" or "off."
How Can Your Epigenetics Change?
Your epigenetics change as you age, both as part of normal development and aging and in response to your behaviors and environment.

1. **Epigenetics and Development**
   Epigenetic changes begin before you are born. All your cells have the same genes but

look and act differently. As you grow and develop, epigenetics helps determine which function a cell will have, for example, whether it will become a heart cell, nerve cell, or skin cell.

Example: Nerve cell vs. Muscle cell

Your muscle cells and nerve cells have the same DNA but work differently. A nerve cell transports information to other cells in your body. A muscle cell has a structure that aids in your body's ability to move. Epigenetics allows the muscle cell to turn "on" genes to make proteins important for its job and turn "off" genes important for a nerve cell's job.

2. **Epigenetics and Age**
   Your epigenetics change throughout your life. Your epigenetics at birth is not the same as your epigenetics during childhood or adulthood.

Example: Study of newborn vs. 26-year-old vs. 103-year-old

DNA methylation at millions of sites were measured in a newborn, 26-year-old, and 103-year-old. The level of DNA methylation decreases with age. A newborn had the highest DNA methylation, the 103-year-old had the lowest DNA methylation, and the 26-year-old had a DNA methylation level between the newborn and 103-year-old (1).

3. **Epigenetics and Reversibility**
   Not all epigenetic changes are permanent. Some epigenetic changes can be added or removed in response to changes in behavior or environment.

Example: Smokers vs. non-smokers vs. former smokers

Smoking can result in epigenetic changes. For example, at certain parts of the *AHRR* gene, smokers tend to have less DNA methylation than non-smokers. The difference is greater for heavy smokers and long-term smokers. After quitting smoking, former smokers can begin to have increased DNA methylation at this gene. Eventually, they can reach levels similar to those of non-smokers. In some cases, this can happen in under a year, but the length of time depends on how long and how much someone smoked before quitting (2).

Epigenetics and Health

Epigenetic changes can affect your health in different ways:

1. **Infections**
   Germs can change your epigenetics to weaken your immune system. This helps the germ survive.

Example: Mycobacterium tuberculosis

Mycobacterium tuberculosis causes tuberculosis. Infections with these germs can cause changes to histones in some of your immune cells that result in turning "off" the *IL-12B* gene. Turning "off" the *IL-12B* gene weakens your immune system and improves the survival of *Mycobacterium tuberculosis*.

2. **Cancer**
   Certain mutations make you more likely to develop cancer. Likewise, some epigenetic

changes increase your cancer risk. For example, having a mutation in the *BRCA1* gene that prevents it from working properly makes you more likely to get breast and other cancers. Similarly, increased DNA methylation that results in decreased *BRCA1* gene expression raises your risk for breast and other cancers (4).While cancer cells have increased DNA methylation at certain genes, overall DNA methylation levels are lower in cancer cells compared with normal cells. Different types of cancer that look alike can have different DNA methylation patterns. Epigenetics can be used to help determine which type of cancer a person has or can help to find hard to detect cancers earlier. Epigenetics alone cannot diagnose cancer, and cancers would need to be confirmed with further screening tests.

Example: Colorectal Cancer

Colorectal cancers have increased methylation at the *SEPT9* gene. Some commercial epigenetic-based tests for colorectal cancer look at DNA methylation levels at the *SEPT9* gene. When used with other diagnostic screening tests, these epigenetic based tests can help find cancer early.

3. **Nutrition During Pregnancy**
   A pregnant woman's environment and behavior during pregnancy, such as whether she eats healthy food, can change the baby's epigenetics. Some of these changes can remain for decades and might make the child more likely to get certain diseases.

Example: Dutch Hunger Winter Famine (1944-1945)

People whose mothers were pregnant with them during the famine were more likely to develop certain diseases such as heart disease, schizophrenia, and type 2 diabetes (7). Around 60 years after the famine, researchers looked at methylation levels in people whose mothers were pregnant with them during the famine. These people had increased methylation at some genes and decreased methylation at other genes compared with their siblings who were not exposed to famine before their birth. These differences in methylation could help explain why these people had an increased likelihood for certain diseases later in life

What is Speciation?

- In evolution, speciation is the process that results in the formation of new and distinct species that are isolated from one another.
- However, biologists have developed two different pathways for the speciation to occur.
- Allopatric speciation, referring to speciation in other homelands," includes a spatial separation from a parent species and eventual evolution of populations.
- Sympatric speciation, referring to speciation in the 'same homeland', includes speciation taking place within a parent species while staying in a same location.

What are the causes of speciation?

- There are various factors that cause speciation. The two main causes for speciation are listed as:

i. Geographical isolation:

- Speciation results from a splitting event in which a parent species is separated into two separate species, often as a result of geographical isolation or some driving force involving population separation.
- Separation could occur either due to physical barriers such as huge ocean expanses, mountain ranges, glaciers, deep river valleys, large rivers or deserts, or a substantial distance due to wider geographical range.
- The free-flow of alleles is prevented when populations become geographically isolated.
- The two species are able to evolve into different trajectories when the separation continues for a period of time.
- Thus, their allele frequencies progressively become more and more different at various genetic loci as new alleles in each population independently emerge through mutation.
- Usually, environmental conditions for the two groups, such as climate, resources, predators, and competitors, will vary, causing natural selection in each group to favor divergent adaptations.

ii. Reproductive isolation:

- The reproductive isolation which is central to the process of speciation takes place because of reproductive barriers, that are formed as a result of genetic, behavorial or physical differences emerging between the new species.
- These are either pre-zygotic processes i.e. differences in courtship behaviors, non-compatible genitalia, or gametes that are unable to fertilize between species.
- On other way, they are post zygotic for instance zygote mortality or the production of sterile offspring.

Types of speciation:

- There are total of 4 types of speciation i.e. allopatric, parapatric, peripatric and sympatric speciation. However, artificial speciation is also included sometimes.
- **Allopatric speciation** occurs when speciation via geographical separation takes place whereas **sympatric speciation** occurs when speciation occurs without geographic isolation.
- **Peripatric** and **parapatric speciation**, however are similar to allopatric speciation as they occur when populations are isolated.

1. Allopatric speciation (Speciation via Geographical separation)

- The process of speciation that take place when the members of the population are isolated geographically from each other, where they are unable to mate and hence genetic exchange is prevented or interfered is termed as **allopatric speciation.**
- There may be a number of ways of isolating populations leading to allopatric speciation: from a river forming a new branch, erosion forming a new valley, or a group of species migrating to a new location without the opportunity to return, such as seeds floating to an island across the ocean.

- The essence of the geographical separation needed to separate populations is entirely dependent on the organism's biology and its capacity for dispersion.
- If two flying insect populations took up residence in different neighboring valleys, it is possible that individuals would fly back and forth from each population, continuing gene flow.
- If two rodent populations were split by the creation of a new lake, however, continued gene flow would be unlikely; speciation would therefore be more likely.
- Allopatric processes are classified into two groups.
    - If a few individuals of a species migrate to a new geographical area, this is called dispersal.
    - If a natural condition happens to physically separate species, this is called vicariance.

How does allopatric speciation occurs?

- Allopatric speciation occurs due to geographical separation of population.
- The geographical separation of population may be due to geological shifts, such as the formation of a mountain by a volcano, the formation of islands, the division of ecosystems by glaciers and rivers, or the destruction of habitats due to human activity.
- As a result of various selective pressures acting on populations, the separated populations then experience divergence in genotypic or phenotypic traits.
- When mutations occur within species, this allows natural selection to induce genetic drift.
- Over time, because of adaptation to their new environment, the separate populations can evolve morphologically different characteristics.
- The characteristics can become so markedly distinct that there is reproductive isolation, preventing the inbreeding of populations and thereby generating new species.
- If this is the case then it is suggested that allopatric speciation has taken place.
- If populations become sufficiently different to be classified as new species, but not sufficiently distinct for the occurrence of reproductive isolation, the species can return into contact and mate, creating hybrids.

Example of Allopatric speciation; Darwin's Finches:

- In the Galapagos finches that Charles Darwin studied, a significant example of allopatric speciation occurred.
- On the islands of the Galapagos, there are about 15 different species of finches, each of which looks different and has specialized beaks for consuming various kinds of food, such as insects, seeds and flowers.
- All these finches originated from a single species of ancestor that must have emigrated to the several islands.
- When populations on the islands were created, they were isolated from each other and numerous mutations emerged.
- In their respective habitats, the mutations that caused the birds to be more powerful became more and more common, and several different species evolved over time.

- If several new organisms evolve in a relatively rapid geological period from one common ancestor, this is called adaptive radiation.

2. Peripatric speciation:

- It occurs when the individuals lying on the periphery, or border of a huge population split off from the main group and result to a new species in course of time.
- Differentiating it from allopatric speciation can be hard.
- When the population that branch off enters a distinct biological niche, like feeding on different food or surviving in a different environment, peripatric speciation occurs.
- Often these new populations that split away from the existing one are typically small, so this can have an effect on the proportion of some characteristics in the new population compared with the old one.
- Say for instance, that there is a bird population that is mostly blue, but some are red.
- A smaller group of birds splits out of the main group, and red is the majority of this smaller group.
- It is probable that their descendants will also be mainly red, which is different from the main group.
- This type of change in gene frequency is referred to as genetic drift.
- Many changes can take place over time, and these, combined with the effects of genetic drift, can cause new species to evolve.

Example of peripatric speciation; the London underground mosquitoes

- The London Underground mosquito is a type of mosquito found in the Underground area of London.
- Because of to its edacious biting, biologists called the London Underground mosquito *Culex pipiens f. molestus*.
- It eventually adapted to human-made underground structures, from being a local above-ground *Culex pipiens*.
- Recent evidence indicates it is a southern mosquito variety related to *C. pipiens* that has modified to the warm underground spaces of northern cities.
- The proof for this specific mosquito becoming a distinct species from *C. pipiens* comes from studies done by Kate Byrne and Richard Nichols.
- The species have very unique features and are particularly difficult to mate.
- More precisely, the *C. pipiens f. molestus* is cold intolerant and bites rodents, and humans, and can breed all year round, while the above-ground species is also cold tolerant, but hibernates in the winter and targets only birds.
- The eggs were infertile when these two varieties were cross-bred, indicating reproductive isolation.

3. Parapatric speciation:

- Parapatric speciation occurs when subpopulations of the same species are largely isolated from each other however have a small region where their ranges overlap.

- This could be caused by a partial geographical barrier or an uneven distribution of members of subpopulation.
- It has very less chances to occur.
- It can occur between several neighboring subpopulations where all the neighboring populations can interbreed, but each subpopulation is so slightly different that it would not be possible for the members on the extreme ends to interbreed with each other. This is referred as ring species.
- That means within the group, the population does not mate randomly, but rather individuals mate with their nearest geographical neighbors more generally resulting in unequal gene flow.
- Non-random mating could increase the rate of dimorphism within populations, in which differed morphological aspects of the same species are exhibited.
- Parapatric speciation results in one or more distinct sub-populations (termed as 'sister species') that have small continuous overlaps in their biogeographic range and are genotypically dimorphic.

Example of parapatric speciation; *Agrostis tenuis*:

- In populations of the grass *Agrostis tenuis* that span mine tailings and natural soils, the best-known example of ongoing parapatric speciation occurs.
- Heavy metal tolerant individuals, a heritable trait, live well on polluted soil, but poorly on soil that is not contaminated.
- For intolerant populations, the reverse happens.
- Gene flow occurs between sub-populations on and off mine tailings, but small variations in flowering time between the two locations inhibit hybridization.

4. Sympatric speciation (Speciation without geographical separation):

- It is the evolutionary process by which organisms are created from a single ancestral species while occupying the same geographical area.
- The distribution ranges of organisms that evolve by sympatry may be similar, or they may only overlap, as contrasted to allopatric speciation.
- Instead of geographical distance causing a reduction in gene flow between populations, sympatry occurs as members of one population make use of a new niche.
- For example, this could occur if a herbivorous insect starts feeding on a new or noble source of plants with which it is not ancestrally associated, or if a new plant species is introduced into the geographical range of the species.
- As insects normally reproduce or lay eggs within the type of fruit in which they were born, the individuals will specialize in feeding and mating on specific fruits over time.
- As a result, gene flow between populations that specialize in different fruits would be decreased, leading to populations being reproductively isolated.
- As new species emerge from populations living in highly overlapping or even similar environments, sympatric speciation is very distinct from the other forms.

- It may be more prevalent in bacteria than in multicellular organisms because when they split, bacteria may shift genes to each other as well as transfer genes to offspring.

How does sympatric speciation occurs?

- One type of sympatric speciation can start with a chromosomal defect during meiosis or the formation of a hybrid individual with large number of chromosomes.
- A condition in which there is an additional set of chromosomes, or sets, in a cell or organism is termed as polyploidy.
- Polyploidy results from a meiosis defect in which, instead of dividing, all the chromosomes pass into one cell.
- There are two major types of polyploidy that could result in reproductive isolation of an individual in the polyploid state.
- One is autopolyploidy where polyploid individuals will possess two or more complete set of chromosomes from its own species.
- For instance, if a plant species with 2n=6 results in autopolyploid gametes which are also diploid, the gametes now possess twice as many chromosome they should possess.
- These new gametes are incompatible with the usual gametes that this species of plant produces.
- However, they could either self-pollinate or reproduce with several other autopolyploid plants with gametes that have the same diploid number.
- In this way, sympatric speciation will occur rapidly by producing offspring with 4n called a tetraploid.
- Only those of this new kind and not those of the ancestral species will be able to reproduce immediately with these individuals.
- Allopolyploidy is the other form of polyploidy where individuals of two different species reproduce to yield a fertile offspring.
- The examples of allopolyploids are cultivated forms of wheat, cotton, and tobacco plants.
- Sympatric speciation can occur in ways other than polyploidy, as well.
- If we consider a species of fish residing in a lake.
- Competition for food increased as the population grew.
- Under pressure to find food, if we assume that the genetic versatility of a group of these fish was to discover and feed off another resource unused by the other fish.
- If the new food source was discovered at a different depth of the lake then, Those fed on the second food source would communicate more with each other over time than the other fish; they would therefore also breed together.
- The offspring of these fish are likely to act and live in the same area as their parents and feed, keeping them apart from the original population.
- If this group of fish continued to remain separate from the first population, as more genetic differences accumulated between them, sympatric speciation would eventually occur.

Example of sympatric speciation; Cichlid fish

- In Tanzania, cichlid fish that live in a small volcanic crater lake are seen as one such example of sympatric speciation.

- There are two very distinct ectomorphs or forms in the population: a yellow-green one that lives along the shore, and a blue-black one that lives at the bottom of the lake.
- By looking at the DNA of the fishes, researchers could see that the two ectomorphs were genetically very distinct.
- It is assumed that these two forms are in the gradual speciation phase at present.

5. Artificial speciation:

- The type of speciation that can be accomplished through the input of human intervention is termed as artificial speciation.
- Human beings may create new, distinct species by separating populations and thus preventing reproduction, or by purposely breeding individuals with desired morphological or genotypical traits.
- This is also known as 'artificial selection'; artificial selection has been undertaken for most modern domesticated animals and plants.
- While evolution of our modern crops and livestock has taken thousands of years, it is possible to imagine the process of artificial selection in organisms that have short life cycles.

Example of artificial speciation; *Drosophila melanogaster*:

- In fruit fly (*Drosophila melanogaster*) species, artificial selection has been most effectively displayed.
- The changes that occur as flies adapt to each environment are shown by experiments in which flies are placed in environments that contain different resources or habitats.
- The flies are removed from the experimental zone after several generations and allowed to cohabit, although the populations are unable to mate due to the process of reproductive isolation that occurred while in isolation.

ADAPTIVE RADIATION

In evolutionary biology, **adaptive radiation** is a process in which organisms diversify rapidly from an ancestral species into a multitude of new forms, particularly when a change in the environment makes new resources available, alters biotic interactions or opens new environmental niches.[1][2] Starting with a single ancestor, this process results in the speciation and phenotypic adaptation of an array of species exhibiting different morphological and physiological traits. The prototypical example of adaptive radiation is finch speciation on the Galapagos ("Darwin's finches"), but examples are known from around the world

Characteristics

Four features can be used to identify an adaptive radiation:[2]

1. A common ancestry of component species: specifically a *recent* ancestry. Note that this is not the same as a monophyly in which *all* descendants of a common ancestor are included.
2. A phenotype-environment correlation: a *significant* association between environments and the morphological and physiological traits used to exploit those environments.
3. Trait utility: the performance or fitness advantages of trait values in their corresponding environments.
4. Rapid speciation: presence of one or more *bursts* in the emergence of new species around the time that ecological and phenotypic divergence is underway.

Conditions

Adaptive radiations are thought to be triggered by an ecological opportunity[3] or a new adaptive zone.[4] Sources of ecological opportunity can be the loss of antagonists (competitors or predators), the evolution of a key innovation or dispersal to a new environment. Any one of these ecological opportunities has the potential to result in an increase in population size and relaxed stabilizing (constraining) selection. As genetic diversity is positively correlated with population size[5] the expanded population will have more genetic diversity compared to the ancestral population. With reduced stabilizing selection phenotypic diversity can also increase. In addition, intraspecific competition will increase, promoting divergent selection to use a wider range of resources. This ecological release provides the potential for ecological speciation and thus adaptive radiation.[3]

Occupying a new environment might take place under the following conditions:[6]

1. A new habitat has opened up: a volcano, for example, can create new ground in the middle of the ocean. This is the case in places like Hawaii and the Galapagos. For aquatic species, the formation of a large new lake habitat could serve the same purpose; the tectonic movement that formed the East African Rift, ultimately leading to the creation of the Rift Valley Lakes, is an example of this. An extinction event could effectively achieve this same result, opening up niches that were previously occupied by species that no longer exist.
2. This new habitat is relatively isolated. When a volcano erupts on the mainland and destroys an adjacent forest, it is likely that the terrestrial plant and animal species that used to live in the destroyed region will recolonize without evolving greatly. However, if a newly formed habitat is isolated, the species that colonize it will likely be somewhat random and uncommon arrivals.
3. The new habitat has a wide availability of niche space. The rare colonist can only adaptively radiate into as many forms as there are niches