



Ankita Dua and Aeshna Nigam

22.1 Concept of Evolution

Nothing in biology makes sense except in light of evolution—Theodosius Dobzhansky

Evolution may be defined as the changes in gene pool which will lead to progressive adaptation of the population to the environment. The concept of natural selection proposed by Charles Darwin and Alfred Russel Wallace combined with Mendelian inheritance gives an insight into the mechanism of evolution. Evolution is basically a two-step process which can only occur when there are heritable variations in the gene pool.

In words of Darwin, evolution is ‘descent with modification’ that species change over time giving rise to new species sharing a common ancestor. Each species has its own set of heritable differences from the common ancestor, which accumulate over periods of time gradually. In the ‘tree of life,’ repeated branching events produce a multilevel tree that links all organisms. Evolution, with respect to extent of change and geological time scales, can be classified as:

- Microevolution is defined as a systematic change in allele frequencies and chromosomal segments in local populations.
- Macroevolution involves changes of greater magnitude. Changes of macro extent occur in the development of characteristics that distinguish groups such as genera, families, orders, classes and phyla. These take place over a long span of geological time scale.

A. Dua · A. Nigam (✉)
Shivaji College, University of Delhi, New Delhi, India
e-mail: aeshnanigam@shivaji.du.ac.in

Various evidences have supported the theory of evolution. Homologous structures are evidence of presence of a common ancestor, whereas analogous structures evolve independently without common descent and are similar due to inhabitation of similar environment as a result of convergent evolution. Species relatedness can be evaluated by similarities and differences at the molecular level, assessing the DNA and protein structure. Study of biogeography shows distribution of organisms all over the globe and gives clues about relationships between them. One of the most valuable evidences comes from fossil records that unfold the story of timelines of existence of species over various periods (Futuyma 1998; Hall and Hallgrímsson 2008; Ridley 2004).

22.1.1 Theories of Evolution: Lamarckism

Jean Baptiste de Lamarck is a French naturalist of the nineteenth century, known for his speculations on evolution, published in his book *The Philosophie Zoologique* in the year of 1809. 'Inheritance of Acquired characters' is his significant proposition albeit unaccepted widely.

Lamarck's strong doctrine is, species undergo modifications concerning the environment, contradicting its fixity. He claims domestication of plants and animals modifies their structure unrecognizable to the wild variety. For instance, domestic ducks and geese lost their ability to fly compared to the wild birds of their race due to prolonged captivity. If the captivity is extended, even more, there might not be a change only in their ability but also in their morphology, claims Lamarck. He further endorsed, *Ranunculus hederaceus* of terrestrial habitat grown in a damp soil has been found to have a smaller stem and devoid of small segmental leaves which are dissimilar to the same species, *Ranunculus aquatilis* of aquatic habitat. According to Lamarck, the impact of the environment on living species causes an imperceptible alteration in structure and organization. Unconcealed modifications in animals can be observed with substantial changes in the environment leading to novel requirements. The emergence of new habits due to these long-lasting changes in the habitat will lead to the development of a new pertinent organ which further develops stronger and larger with perpetual use. These modifications in animals curtail inefficient organs to disuse. The disappearance of inefficient organs is coupled with prolonged disuse. Lamarck also observed these permanent modifications become inherited giving rise to distinct species. In a nutshell, external stimulus causes a heritable beneficial genomic mutation in species for adaptation (Fig. 22.1). Based on his extensive studies and observation of species, he discerned two laws of nature:

First Law: 'In every animal which has not passed the limit of its development, more frequent and continuous use of any organ gradually strengthens, develops and enlarges that organ, and gives it a power of the proportional length of time it has been so used; while the permanent disuse of any organ imperceptibly weakens and deteriorates it, and progressively diminishes its functional capacity until it finally disappears.'

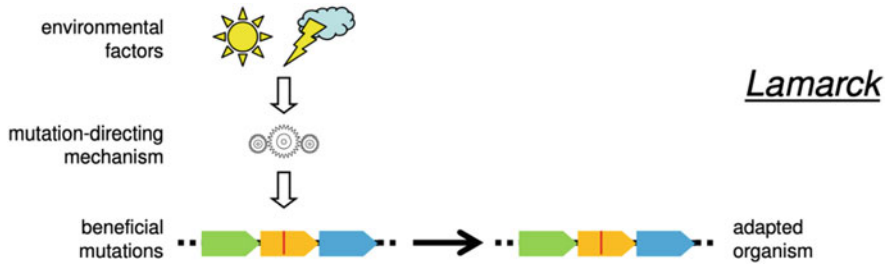


Fig. 22.1 Lamarck's theory of evolution. (Adapted from Koonin and Wolf 2009)

Lamarck has demonstrated the authenticity of the laws with various examples and believed these laws are certainly true and permanent. He gave definitive examples of use and disuse of organs rising due to contemporary habits. Certain changes in the habitat of animals have induced swallowing of the feed without primitive mastication, eventually leading to the absence of teeth in vertebrates (e.g. whale, anteater). Further, Lamarck articulates, disuse of eyes had constricted the organ in moles. Living beneath the soil, where sunlight is arduous to percolate, mole has tiny eyes based on its utility. In addition to that, *Spalax* which lives in the similar habitat of mole rats is blind with the vestiges of the organ, as a result of lack of utility. Snakes are the exceptional class of reptiles which do not have four limbs like crocodiles, frogs and turtles. Lamarck reasoned this unique feature of snakes with two facts: (1) Their peculiar adoption to crawl with the elongated body helped them to hide in the grass and move in confined places with ease (2). Their long legs have put to perpetual disuse, which eventually disappeared. It is unlikely for a snake to have short legs which makes them incompetent, and they cannot have more than four legs under the reptile criteria.

Lamarck illustrated his opinion on the development of a new organ or stronger and prominent development due to recurrent use of an existing organ supporting altered environment with exemplifications. Perpetual use of skin between the three-digit feet to capture aquatic life forms of prey has given rise to the palmate or webbed foot essential for swimming in ducks and geese. Lamarck explained, in a few cases, birds developed long stretchy legs with feathers above the thighs when they are reluctant to swim and depend on prey in the shore. In forests, prey-predator stress is inevitable and especially ruminants like deer need to protect themselves from predators as well as hunters. Lamarck claims, the 'inner feeling' of the animals to safeguard them from dangers has allowed the secretion of a blend of horny-bony substance which gave rise to antlers and horns. He claims giraffe has developed long necks compared to its ancestors. The gradual transformation of African forest grasslands to arid areas coerced the animals to depend on trees for nourishment. This obligation has resulted in elongated necks and limbs in giraffe. Lamarck's other doctrine is, the adaptations in species that help in survival are inherited and are called use-inheritance. This gives rise to the second law of Lamarckism.

Second Law: 'All the acquisitions or losses wrought by nature on individuals, through the influence of the environment in which their race has long been placed,

and hence through the influence of predominant use or permanent disuse of any organ; all these are preserved by reproduction to the new individuals which arise, provided that the acquired modifications are common to both sexes or at least the individuals which produce the young.'

This theory of use-inheritance is considered improbable due to lack of evidence by Lamarck. This theory was condemned by many people, and experiments were conducted to prove or disprove it. August Weismann, a German Evolutionary biologist, was the first person to propose the germplasm theory in animals (Weismann 1893). He proposed any metamorphosis in the somatoplasm does not affect the germplasm. Weismann argued Lamarck's proposition of 'Inheritance of acquired characters,' claiming germ cells give rise to somatic cells, therefore for a variation to occur, preliminary change must occur in the germplasm to be inherited. During one of his lectures delivered in 1888 on 'A supposed transmission of mutilations,' he presented results of his experimental investigation on mutilation inheritance in mice (Weismann 1891). In the first generation, he amputated tails of 12 mice comprising 7 females and 5 males. The offspring from the first generation were found with perfectly grown tail, and even subtle presence of the acquired inheritance was not found. Surprisingly, the fifth generation with 901 offspring developed from the mutilated parents did not show the trivial presence of rudimentary tail defects or tail-less condition. Weismann set forth a plausible assumption that the expression of mutilations in the progeny might take place after many generations. Unfortunately, use-inheritance can be widely accepted if there was at least one proof to support this theory.

Furthermore, McDougall in 1938 (McDougall 1938) conducted experiments on learning as an acquired inheritance. He designed a T-shaped tank, with two exits: one exit with the electric shock was illuminated, whereas the free exit was kept dim. Rats which chose the lighted pathway received an electric shock for 3 s, and animals which chose the dim exit were rewarded. He trained the rats six times daily to accustom to the experiments and halted the training only when the rats learnt to discriminate the exits and chose the dim exit successively. He bred these rats for the second generation. McDougall found mistakes reduced gradually from generation to generation and claimed learning is an acquired trait. Drew (1939) criticized McDougall experiments for biased learning in animals, and inheritance of avoidance behaviour interlinked with various factors is impossible. When repeated, contrast results were obtained by Crew and Agar (Agar et al. 1954; Crew 1936). Further, technical errors found in McDougall's experiment led to severe criticism.

22.1.2 Theories of Evolution: Darwinism and Neo-Darwinism

Charles Darwin and Alfred Russel Wallace independently developed theories of evolution based on natural selection which were communicated to a meeting of Linnaean Society of London on July 1, 1858. Both went on voyages of discovery before stating their ideas: Darwin went around the world on H.M.S. Beagle (Fig. 22.2), and Wallace travelled to Brazil, Malaysia and Indonesia.

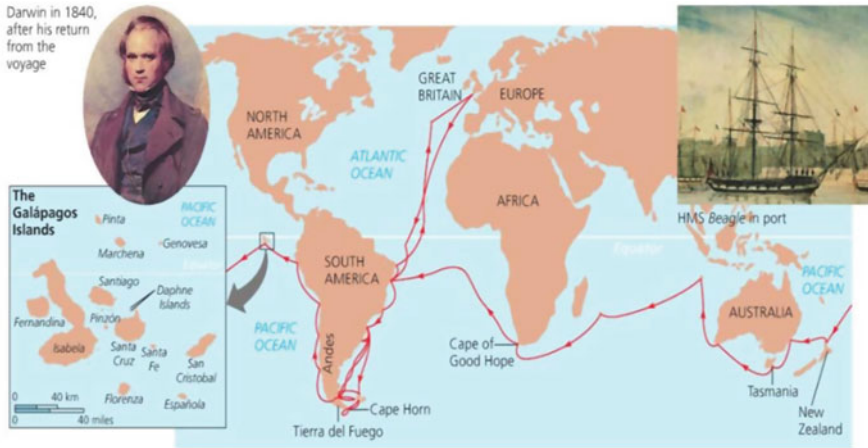


Fig. 22.2 The voyage of HMS Beagle. The path traced by HMS Beagle in 1831 in its 5-year journey that led to Darwin's postulates of natural selection and origin of species. (Adapted from Campbell et al. 2008)

Wallace is considered to have begun the study of biogeography, and both of them were posthumously awarded the 'Gold Medal' by Linnaean Society of London for the 50th anniversary of their publication. Darwin's voyage spanned a period of 5 years from 1826 to 1830 that enabled him to observe the wide range of species and geological forms around the globe. His breakthrough discovery was the exotic collection of flora and fauna evolving on the Galapagos Islands off the coast of Ecuador. These are located in the Pacific Ocean, approximately 960 km west of the South American coast, straddling the equator at the 90th meridian west. The archipelago was made of 13 major islands, 6 smaller islands, over 40 islets and many smaller unnamed islets and rocks, for a total of approximately 8000 km² of land spread over 45 000 km² of water.

He noted that different islands with similar habitats were not always occupied by identical species. He proposed that:

- Excess reproduction and limited resources lead to competition, and as a result of natural selection, only the organisms best adapted to the habitat could survive and pass their characters to the next generation.
- Changing environments and hereditary variations and natural selection together result in modification of existing characters or origin of new characters that become established throughout a species.

Darwin's work was documented in his book *The Origin of Species* in 1859 which is said to have revolutionized the foundation of evolutionary biology.

The most curious fact is the perfect gradation in the size of the beaks in the different species of Geospiza, from one as large as that of a hawfinch to that of a chaffinch, and . . . even to that of a warbler. . . . Seeing this gradation and diversity of structure in one small, intimately related group of birds, one might really fancy that from an original paucity of birds in this

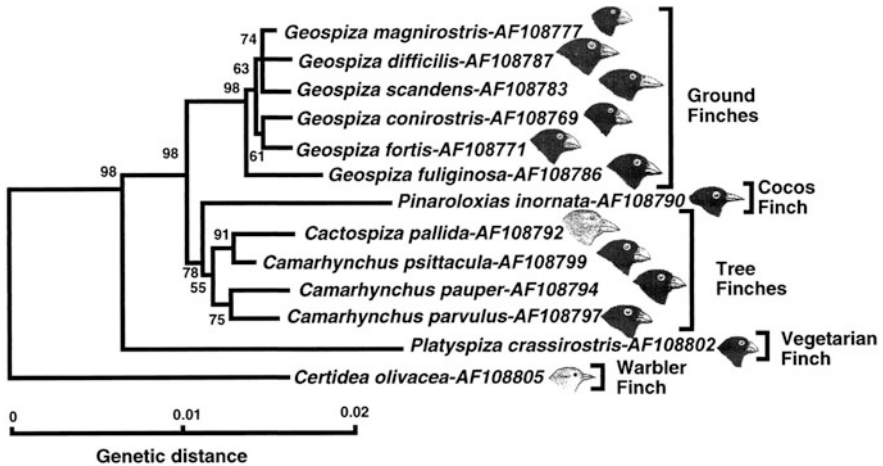


Fig. 22.3 Phylogenetic analysis of Darwin's finches. Combined analysis of the cytb and cr sequences of Darwin's finches done by neighbour-joining tree construction method. Shape of the beak is illustrated by the drawings made on the right side (Sato et al. 1999)

archipelago, one species had been taken and modified for different ends.—Darwin (1839) (Abzhanov 2010)

He studied the concept of adaptive radiation (diversification of a founder population into a collection of species differentially adapted to diverse habitats) under natural selection of 14 closely related finch species (belonging to the Avian order Passeriformes). These were diverse in their beak shapes and sizes. Later on, he asked John Gould at the museum of Zoological Society, London, to catalogue these species. Gould realized that these diverse species actually reflected differences in diets that were closely related to each other and distantly to a South American mainland species. Developmental basis of this variation was analysed by a comparative analysis of expression patterns of the growth factor *Bmp4* of the species of genus *Geospiza* (Abzhanov et al. 2004; Parent et al. 2008). The expression of *Bmp4* in upper beak mesenchyme correlated with the deep and broad beak morphology. When this protein was misexpressed in chick embryo, it caused morphological transformations similar to beak morphology of the large ground finch (Abzhanov et al. 2004). In another analysis, two mitochondrial DNA segment sequences (cytb, cytochrome b, and cr, control region) have been used for evolutionary history of the group. The results reveal that Darwin's finches are a monophyletic group with the closest species to the founder being Warbler's finch, followed by vegetarian finch and two sister groups of ground and tree finches. The Cocos finch found on the Cocos Island of the Pacific Ocean is related to the tree finches (Fig. 22.3). The mtDNA and microsatellite data found was consistent with the theory that the finches originated from a single common ancestor from a possible founder population that

reached the islands from South/Central America (Bluestone 2009; Bowman 1961; Parent et al. 2008).

Post Darwin, a number of workers such as J. Huxley, T. Dobzhansky, J.B.S. Haldane, S. Wright, T.H. Morgan, G.G. Simpson, G.L. Stebbins, E. Mayr and C. Darlington came about the theory of ‘Modern Synthesis’ or Neo-Darwinism.

Neo-Darwinism is an attempt to reconcile Mendelian genetics, which says that organisms do not change with time, with Darwinism, which claims they do.—Lynn Margulis

The term ‘Neo-Darwinism’ was coined by physiologist George Romanes in 1883. It is the modern concept of evolution where a gene pool of a population is a unit of selection versus the individual. There is a synthesis between Darwin’s theory of natural selection and gene mutations playing a central role in production of variation in a population. By this time, Mendel’s work had also been rediscovered, and it was significant in adding to the theory. A connection was established between genes as a unit of evolution and natural selection as the mechanism. Evolution is a gradual process which is a result of variations accumulated at the genetic level in populations over a period of time (via mutations). The result of these variations is the phenotypic changes in the population and in allele frequencies as a result of natural selection (Denis 2011).

The pioneer workers who strengthened the ‘Concept of Modern Synthesis’ were as follows:

- Theodosius Dobzhansky worked on evolution of fruit fly populations (*Drosophila melanogaster*). His major publication was *Genetics and the Origin of Species* in 1937.
- E.B. Ford worked on ‘Ecological Genetics,’ and his book titled the same was published in 1964.
- H.B.D. Kettlewell was the pioneer worker on industrial melanism in peppered moth *Biston betularia*.
- Julian Huxley wrote the book ‘Evolution: The Modern Synthesis’ in 1942 that brought to light the concepts introduced by Fisher, Haldane and Wright.

22.2 Genetic Variation in Population

Variation is an interplay of heredity and environment and is the most impressive characteristic of any sexually breeding population. The capacity to undergo variation is due to a complex set of heritable traits. Variations are the reason present-day organisms have evolved from a few primordial forms of life and diversified into complex life systems; they are seen at phenotypic, chromosomal and molecular levels.

Phenotypic variation is seen in different varieties of species in animals as well as in different flowers in the flower kingdom. These visible morphological or phenotypic differences are called polymorphism, meaning ‘many forms.’ For example, land snails have different coloured bands on their shells, mammals have different coloured coats, and insects have patterned wings. The human species have also been understood to be polymorphic; the most interesting example being their blood

groups where the antigen coded is present on the surface of the blood cell. The antigens are polymorphic. In the Duffy blood typing system, there are two antigens present on the surface of cells. Alleles coding these antigens called the 'Duffy alleles' encoded by a gene on chromosome 1 are often polymorphic. Various human ethnic groups have varied status of Duffy polymorphism (Anstee 2010). Variation in chromosomes is often an indication of polymorphism at the phenotypic level. Researchers found abundant comparative data on comparing polytene chromosomes from various species of *Drosophila* (Zykova et al. 2018). These chromosomes develop from diploid nuclei chromosomes by successive duplication of each chromatid without the segregation. The formed elements associate length-wise and form a cable-like structure. In *Drosophila melanogaster*, they are >100 times longer than regular metaphase chromosomes. Here, level of variation in chromosomes can be studied at an unparalleled level. In every polytene chromosome, banding patterns are significant as there is alternation between compacted and decompact regions of chromosomes known as bands and interbands. Dobzhansky and his team members identified various patterns of banding in *Drosophila* species. Variations are also seen at the nucleotide and at the protein level. Genetic variation in natural population was studied by R.C. Lewontin, J.L. Hubby and H. Harris by application of gel electrophoresis to study amino acid differences in proteins of various species. Amino acids are building blocks of proteins, and their differences in shape, molecular weight and charge can be studied while migrating in gels. This technique was applied to various other creatures as well, and different forms of proteins could be studied as the mobility of a protein was specific through the gel. The ultimate data on genetic variation is obtained on DNA sequencing. All sequences—exons or introns—can be sequenced and analysed. At present high end sequencing technologies have been successful in decoding even the billion base pair human genome.

Variations have been classified by various workers into a number of categories:

22.2.1 Continuous Vs. Discontinuous Variations (Table 22.1)

Evolutionary biologists also classify variations as follows.

Table 22.1 Comparing continuous and discontinuous mutations

Continuous variations	Discontinuous variations
<ul style="list-style-type: none"> • Numerous, small-scale variations occurring between individuals of a population with respect to characters such as height, weight, skin colour, texture, etc. • These are fluctuations on either side of the mean value of a character. 	<ul style="list-style-type: none"> • Sudden, drastic changes in traits of individuals of a population • Usually referred to as mutations which are responsible for variations that lead to formation of a new species

22.2.2 Environmental Variations

These are acquired changes and may not be inherited at the gene level. Environmental influences act on nutrition, competition, disease and biotic and abiotic factors. Phenotypic plasticity is defined as the ability of a genotype to produce more than one phenotype when exposed to different environments. For example, in the semiaquatic plant *Ranunculus*, the leaves that are submerged in water have a dissected leaf lamina, whereas those above water have a single lamina (Cook and Johnson 1968).

22.2.3 Mutational Variations

According to T.H. Morgan, the ultimate source of heritable variation is mutation, and he stressed on the fact that the latter are changes in single genes with effects ranging from minute to severely pronounced. The term ‘gene’ was coined by Danish botanist Johannsen in 1909, and Wilhelm Waagen gave the term ‘mutation.’ However, it was Hugo de Vries and William Bateson who described mutations as major changes in genes at the molecular level leading to deviations from the parental types. Mutations occur in genes bringing about variation, and natural selection favours the advantageous variants and weeds out the disadvantageous ones (Marshall 2002). Mutations can be point mutations or chromosomal aberrations:

- Gene/point mutations (discovered by T.H. Morgan during his work on white-eyed trait in a laboratory population of *Drosophila*) are the elemental source of variations as they bring about changes in a pure line gene of a parent and lead to origin of a new gene that is inheritable. Genes are generally stable and mutate at a slow rate. These changes could be in composition—transition and transversion—or change in number by addition, deletion and frame shift. Mutation rate is defined as ‘probability that an allele copy changes to another allelic form in one generation.’ As mutation is a spontaneous process, the rate of change in gene frequency from mutation is very low. Mutation of a single gene usually has minimal effect on the population; it is amplified by interaction with other genes, for example, cases of multiple allele interactions, epistasis, pleiotropy and polygenic effects. Generally, mutations that are less severe occur at a higher frequency, and a mutant gene may not display its effects under normal environmental conditions but may be deleterious in a different environment.

Table 22.2 Change in number of chromosomes along with amount of DNA

Aneuploidy	Loss or gain of individual chromosome of a normal diploid set of an organism. It can be nullisomy ($2n - 2$), monosomic ($2n - 1$), trisomic ($2n + 1$), tetrasomic ($2n + 2$) and polysomic
Haploidy	Loss of an entire set of chromosome
Polyploidy	Gain of one or more sets of chromosomes. It can be triploidy ($3n$), tetraploidy ($4n$) and polyplody

n refers to the haploid set of chromosomes

- Aberrations on the other hand refer to loss or gain of genes and change in placement or position within the chromosome or between different chromosomes. Structural changes involve deficiency (loss), duplication (repetition of a DNA segment) and polyteny (multiple copies of entire DNA strands) which bring about changes in the amount of total DNA (Table 22.2). Changes in location of genes (no change in DNA amount) are done by inversion (reversal of gene order within same chromosome)—paracentric/pericentric (depending on presence/absence of centromere in the inverted segment) and translocation. Loss or gain of genes with change in amount of DNA is done by change in chromosome number.

22.2.4 Recombination

Variations introduced as a result of permutations and combinations of existing genes are termed as recombination, and these are the raw materials for natural selection and evolutionary changes. Such events may add up with point mutations, and their combined effect makes a great impact. Recombination maybe a result of:

- Crossing over of two homozygous gametes to give rise to a heterozygous offspring.
- Random assortment of genetic material.

22.3 The Neutral Theory of Molecular Evolution

22.3.1 Neutral Theory

The neutral theory of evolution was given by Motoo Kimura in 1968 and it states:

This neutral theory claims that the overwhelming majority of evolutionary changes at the molecular level are not caused by selection acting on advantageous mutants, but by random fixation of selectively neutral or very nearly neutral mutants through the cumulative effect of sampling drift (due to finite population number) under continued input of new mutations (Kimura 1991).

As sequencing technologies developed, more data was available to test this hypothesis. Kimura predicted that in protein molecules the substitutions of amino acids that occur more often are conservative. These replacements do not affect the function of the protein as the changed amino acid shares similar biochemical properties to the original one. Introns and pseudogenes also evolve at a high rate similar to these base substitutions.

There were several similarities of the neutral theory with Darwin's postulates:

- Natural selection is the driving evolutionary force and results in adaptation of organisms to their environment.
- Mutations in the functionally important regions are deleterious, and selection rapidly removes them from the population.

Kimura's theory challenged Darwin's work as:

- It was proposed that intraspecific nucleotide differences are a product of neutral mutations rather than genetic drift.
- Most intraspecific polymorphisms are also neutral.

Alterations or substitutions in protein-coding DNA sequences can be classified as synonymous (do not affect encoded amino acids) and non-synonymous (affect encoded amino acids). Synonymous substitutions are mostly neutral as they do not alter the protein sequence.

The neutral theory focuses on three main ingredients: mutation, genetic drift, and purifying selection:

- Mutations are the driving forces of evolution in proteins as well as DNA. Every generation, approximately 10^8 – 10^9 events of mutation occur. As discussed, mutations can be beneficial or detrimental to the fitness of an organism or may be selectively neutral.
- If the mutations are advantageous, they end up getting fixed in the population. The negative mutations are eliminated from the population by the action of purifying selection.
- Selectively neutral mutations have no effect on fitness, and their fate is dependent on random genetic drift. Most are lost from the population shortly after they appear.

Evolution rate according to the neutral theory depends on the neutral mutation rate which is constant in different lineages over time. Highest rates of evolution are found in molecules in which any mutational change may have least effect on the function. On the contrary, the lowest rate is found in the molecules where selection pressure is the highest (Duret 2008).

22.3.2 Rate of Neutral Substitution and Molecular Clock

Mutations in DNA can be of three types: deleterious (which may affect the fitness of the individual negatively), may increase efficiency of organism and can be neutral. When a mutation has no effect on survival and reproduction of an individual, it is termed as a neutral mutation.

Investigations about the molecular clock started with study of the proteins—haemoglobin, cytochrome c and fibrinopeptides in the early 1960s. E. Zuckerkandl, L.B. Pauling, E. Margoliash, R.F. Doolittle and B. Blomback concluded that

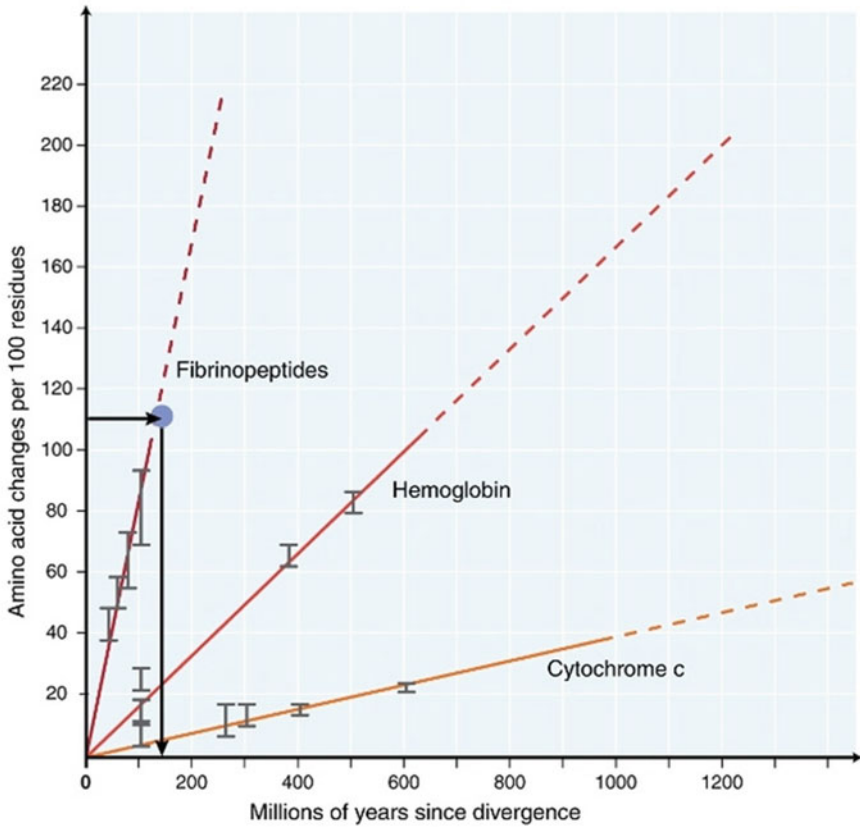


Fig. 22.4 Amino acid changes in cytochrome c, haemoglobin and fibrinopeptides. All three proteins display different rates of changes per unit time, but the rate is constant for each. (Adapted from Yi, S.: Neutrality and Molecular Clocks. Nature Education Knowledge. 4(2), 3 (2013))

differences in the protein sequences of different species were in accordance to their timeline of divergence (Fig. 22.4). Regular ticks of clocks were considered synonymous to accumulation of amino acid/nucleotide substitutions over time in a lineage and hence the name ‘molecular evolutionary clock’ coined by Zuckerkandl and Pauling in 1965. Evolutionary rate or rate of molecular clock is the number of substitutions per a defined unit of time.

The concept of molecular clock measures the absolute time of evolutionary change based on the observation that various genomic regions evolve at steady rates. It talks about orthologous and paralogous genes. Both are homologous, that is, arising from a common ancestor, where orthologous genes diverged post a speciation event and code for proteins with similar function in different species and paralogous genes diverge within the same species and encode for proteins with similar not identical function. Nucleotide substitutions in orthologous genes are

proportional to the speciation event where the species diverged from the common ancestor and in paralogous genes to the time of duplication.

Initially it was hypothesized that the evolutionary force driving these substitutions was natural selection. However, Kimura said that the changes at the molecular level were neutral, that is, no consequence over fitness, and these occurred completely by random chances. Hence, it could not be predicted whether a specific neutral mutation will be or not fixed in a population.

Rate at which neutral substitutions occur in a population depends on the mutation rate and can be predicted as:

Total number of haploid individuals in a population = N .

Rate of neutral mutations = $u/\text{individual/generation}$.

Total number of mutations in one generation = Nu .

As all the mutations are neutral, their success rate is dependent on simply chance probability. Hence all mutations have an equal chance of getting fixed (equivalent to substitution).

Probability that each mutation is fixed = $1/N$.

Rate of substitution = Number of new mutations in each generation (Nu)

× probability that each mutation gets fixed ($1/N$) = u

Therefore, for neutral mutations, rate of substitution = rate of mutation.

It can be predicted that according to the neutral theory, if mutation rates are constant over time, substitutions will occur constantly as well and clock-like regular rates of substitutions will occur constantly over time.

22.4 Natural Selection

Natural selection is defined as a directional, non-random and guiding force that leads to evolution of organisms to a better state of adaptiveness. Natural selection is the differential reproduction of genotypes; it is measured by the relative reproductive successes (fitness) of genotypes.

Salient features of natural selection:

- Key force in driving evolution.
- Shifting of allele frequencies in a large population.
- Fitness is an attribute of a genotype due to its genetic contribution to the future generation. High fitness is a measure of high rate of reproductive success and hence increased genetic contribution to the gene pool. Natural selection clearly favours individuals with higher fitness (Orr 2009).
- Intensity of natural selection acting on genotypes in a population is defined as selection coefficient.

- If a mutation is selected for or against in a population, it is termed as ‘positive or negative’ selection.

Modes of selection in nature vary at different levels of biological organization leading to varied evolutionary outcomes:

- A. At the population level:
 - Directional selection.
 - Stabilizing selection.
 - Disruptive selection.
- B. At the local population level or within species:
 - Kin selection.
 - Group selection.
- C. At the level of individuals:
 - Sexual selection.

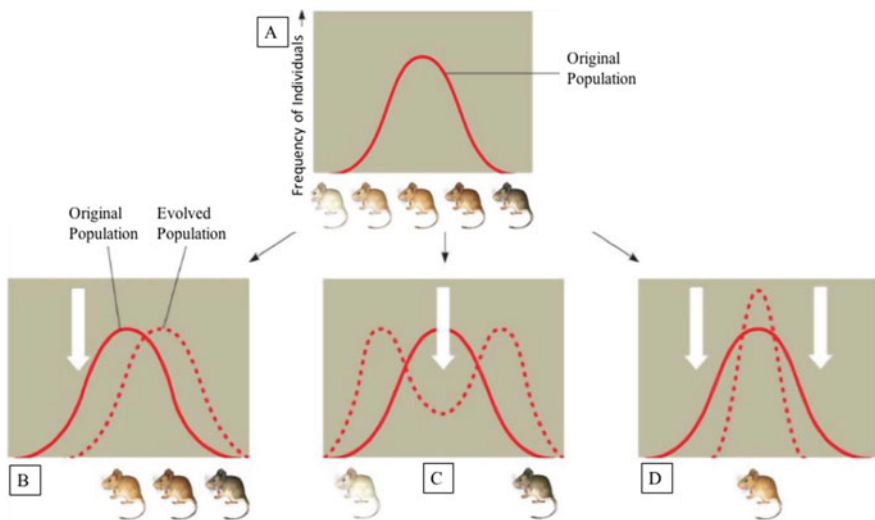


Fig. 22.5 Modes of selection. A hypothetical deer mouse population (with heritable variation in fur colour from light to dark) as an example of three types of natural selection. White arrows indicate patterns of evolution, selective pressures against certain phenotypes: *A*: Original population graph where frequency of individuals is plotted against the fur colour phenotypes. *B*: Directional selection favours the extreme phenotypes, in this case the dark fur individuals. The darker mice are saved from their predators as they take refuge under dark rocks in the environment. *C*: Disruptive selection favours variants at both ends of the spectrum. The mice inhabit patches of light and dark coloured rocks, and the mice of an intermediate colour are at a disadvantage. *D*: Stabilizing selection favours the intermediate/average phenotypes. If the environment consists of intermediate colour rocks, the light and dark mice will be eliminated. (Adapted from Campbell et al. 2008)

22.4.1 Directional, Stabilizing and Disruptive Selection

Effect of the genotype at various loci combined with environmental effects define the phenotype of an organism (Fig. 22.5).

22.4.1.1 Directional/Progressive Selection (Under Changing Environmental Conditions)

Under pressure of changing environmental conditions (biotic and abiotic factors), a species population is faced with the challenge of adjusting and keeping pace with change. Hence the genetic variables that were ideal in the older environment cannot continue to do so in an altered one. With respect to the change in the environment, a new adaptive peak is reached, an extreme being favoured and others eliminated. Also known as progressive selection, a unidirectional change in the genetic composition of the gene pool occurs where the peripheral variants with weak selective advantages originally may find themselves to be better adapted. Phenotypes at one end of the spectrum are selected causing the allele frequency to continuously shift in one direction.

The classical example of directional selection is the phenomenon of industrial melanism in England in the nineteenth century. The role of natural selection in evolution of melanic forms in peppered moths (*Biston betularia*) was studied in detail by H.B.D. Kettlewell and E.B. Ford. Peppered moths exist in two forms—melanic (black in colour; dominant allele C) and non-melanic (mottled grey; homozygous for recessive allele c). Prior to the industrial revolution, the melanic forms were a rare sight, and the non-melanic light coloured moths were abundant in the natural habitat. Predators could easily spot and prey upon the dark melanic forms that rested on the light coloured lichen-covered trees. The non-melanic moths easily blended with the environment and were not easily visible. With the start of the industrial revolution, these numbers underwent a dramatic change as did the environment. The sooty smoke ejected from the factories (heavy industrial areas around Manchester and Birmingham) resulted in attributing a black cover to the trees and prevented growth of lichens. This resulted in protection of the melanic forms and more visibility of the light coloured moths to the predators. Industrial melanism owes its name to the increase in proportion of the melanic forms as selection favoured the moths for the specific character—protective coloration. Kettlewell and Ford who studied this phenomenon in detail reasoned that a single dominant gene for the melanic colour was present in the population and its frequency increased from a mere 1% to 90% with the onset of the revolution in less than 50 generations of the insect. Therefore, changing environment led to change in the genetic constitution of the population as in the changed environment the alleles favoured by selection were different from the ones found in the earlier environment (Cook and Saccheri 2013).

The action of natural selection on a character such as body size which is continuously distributed is interesting to study. Based on the assumption that small-sized individuals have higher fitness (greater reproductive success) compared to larger individuals, natural selection is then directional and favours these, and if the

character is inherited, a decrease in body size is seen over the generations. The same assumption can be vice versa, i.e., if large-sized individuals had higher fitness.

For example, the pink salmon (*Oncorhynchus gorbuscha*) in the Pacific Northwest known for performing extensive migrations has been decreasing in body size. In 1945, fisherman chose the salmon based on not their number but by their pound weight. And for such a screening, they increased the use of gill netting that selects for larger fish. This further led to the increased survival of the smaller fish, and as a result of such a selection, the average weight of salmon decreased by about one-third in the next 25 years.

22.4.1.2 Stabilizing/Normalizing/Centripetal Selection (Under Constant Environmental Conditions)

This selection takes place in a uniform environment, with the sole aim of increasing the adaptive peak by eliminating peripheral variants on both sides of the mean population. The mean population is well adapted to the given environment that remains stable, and the variants arise by mutation, gene flow, recombination and chromosomal segregation. The ill-adapted variants are weeded out, and the well-adapted are preserved and increase in number over time. Hence, this selection aims to keep the population genetically constant for hundreds or thousands of generations. In such conditions, individuals with extreme measurements of any trait are at a disadvantage in comparison to the ones with average measurements. It is unspectacular, ubiquitous and most common mode of selection in nature.

At the genetic level, variability is preserved by stabilizing selection. The intermediate phenotype having the average characteristics is usually heterozygous compared to the extremes that are homozygous and usually disadvantageous.

The most spectacular study of stabilizing selection was published by American scientist H.C. Bumpus at Rhode Island. Post a severe snow storm, he studied 136 immobilized sparrows in his lab of which 72 were revived but 64 died. He identified the sex and measured several morphological traits of the sparrows. His conclusion was that 'the birds which perished had either very long/ short wings as compared to those with average structural measurements.' This indicated that individuals on the extremities of the populations do not survive abnormal or catastrophic events.

Another example of stabilizing selection is the data regarding birth and survival rate of human babies. Karn and Penrose (1952) showed that mortality for infants is greater for those who are either underweight or overweight. Newborns born at average weight (7.5–8.5 lb) show minimum mortality. The underweight babies are born premature (not ready for independent existence), and the overweight babies often suffer physical damage during delivery.

22.4.1.3 Disruptive/Diversifying/Centrifugal Selection (Under Heterogeneous Environmental Conditions)

In disruptive selection, both the extreme phenotypes in a population are favoured, and the intermediates are selected against. Here, genetic uniformity becomes a drawback when a population encounters a range of ecological discontinuities. The

population as a result splits into several groups with different sets of genotypes, each capable of successfully exploiting a different environment. This leads to adaptive polymorphism with respect to ecological opportunities. Diversifying selection hence facilitates a polymorphic population to adapt to different niches of a heterogeneous environment (consisting of different microhabitats). Individuals of the intermediate category have lower fitness compared to the extremes (homozygous for various alleles) and fail to survive. Hence, disruptive selection promotes genetic diversity as a previous homogenous population gets split into different adaptive forms as a result of being subjected to divergent selection pressures.

22.4.2 Kin and Group Selection

Kin selection and group (multilevel) selection are two evolutionary phenomena which form the framework explaining the social behaviour in animals.

22.4.2.1 Kin Selection

Charles Darwin while explaining his theory of natural selection proposed that individuals have a tendency towards increasing the chances of their own survival and reproduction. However, cases have been seen in nature where the close genetic relatives aid each other selectively. This social phenomenon known as altruistic behaviour was a Darwinian puzzle. Why should the individuals save their close relatives at the risk of their own fitness? It was several years later that this question was answered by J.B.S. Haldane, one of the founders of population genetics. This phenomenon was termed as 'indirect selection' by Jerry Brown. This is in contrast to direct selection where the individual increases its own fitness by self-reproduction. Another term 'kin selection' was introduced by John Maynard Smith which includes the parental aid given to the offspring (which are descendent kin) as well as altruistic behaviour for close relative (non-descendent kin). Kin selection has been mainly used to explain why the phenomenon of social and altruistic behaviour has evolved in animals.

The total contribution of an individual including direct and indirect fitness is termed as inclusive fitness. Direct and indirect fitness can be expressed in the same genetic unit. For example, an individual A has one offspring and adopts three nephews. The total genetic contribution of A to the next generation would be:

$$\begin{aligned} \text{Inclusive fitness} &= \text{Direct fitness} + \text{Indirect fitness} \\ &= 1 \times 0.5 + 3 \times 0.25 \\ &= 0.5 + 0.75 = 1.25 \text{ genetic units} \end{aligned}$$

The mathematical concept of inclusive fitness by W.D. Hamilton introduced in 1954 can be used to calculate the adaptive value of altruistic behaviour. The altruistic behaviour would evolve only when the indirect fitness gained through the altruistic

allele is greater than the direct fitness gained by self-reproduction. This is stated as Hamilton's rule and can be represented as:

$$r_b B > r_c C$$

where

' r ' is the coefficient of relatedness.

' B ' is the number of genetic relatives that have survived due to the altruistic act of the individual.

' C ' is the number of offspring produced by the individual.

Using the equation for the previous example:

$$r_b B = 0.25 \times 3 = 0.75$$

$$r_c C = 0.5 \times 1 = 0.5$$

Through calculations we conclude $r_b B > r_c C$; thus, any allele associated with this altruistic act would increase in frequency.

Studies have been carried out to show the universality of kin selection. Łukasiewicz et al. (2017) worked on bulb mites (*Rhizoglyphus robini*) to show how the effect of relatedness promotes female productivity and cooperation in sex. They carried out experiment in two evolutionary groups: one of relatives and the other of non-relatives in the laboratory during the reproductive phase of the cycle. The result was in sync with the kin selection theory, where the evolution in the group of relatives resulted in increased reproductive output by the females (Kin selection, <http://nectunt.bifi.es/to-learn-more-overview/kin-selection/>; Kramer and Meunier 2016).

22.4.2.2 Group Selection

Group selection or multilevel selection, as against the Darwinian concept of selection, refers to the idea that selection does not take place at the individual level but can act on multiple levels of biological organizational level including a whole group of organisms. Group selection looks into the direction and strength selection acting on multiple hierarchical levels. There are several arguments which emerge and discard the theory of group selection.

22.4.3 Sexual Selection

Sexual selection is a subcategory of natural selection which provides advantage to certain individuals over the others to be able to successfully mate. It accounts for the evolutionary phenomenon of the existence of expensive secondary traits which make them much more vulnerable to the predators. Sexual selection is also the evolutionary phenomenon which has also resulted in sexual dimorphism. A common example

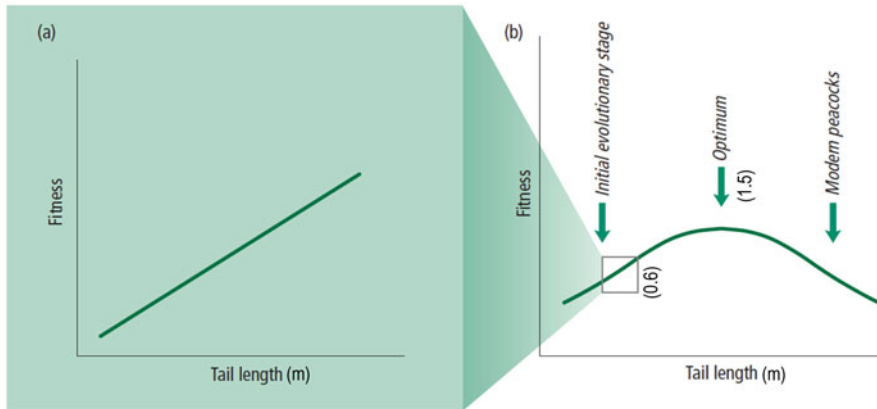


Fig. 22.6 Sexual selection in peacocks. (a) The relation between the tail length of the peacock and fitness; (b) the correlation in the exaggeration of character and fitness results in bell-shaped curve. The peak is the optimal value. The modern peacocks lie on the right of the optimal value (Adapted from Ridley 2004)

of this is the peacock and peahen. Peacock has elaborate tail feathers and is brightly coloured which is costly for its survival. Firstly, there is immense physiological investment in development of these tail feathers. The elaborate courtship display through these colourful plumages consumes time and energy. Lastly these feathers make it easy for the predators to spot them. In contrast the peahen is drab coloured with a short tail exhibiting sexual dimorphism between male and female. Why have these secondary sexual traits evolved in male? Darwin argued that even though these traits are expensive and may shorten the life, but it ensures that the individual is able to contribute its genes to the next generation, thus increasing their fitness. The peacock with the bigger, colourful and elaborate tail feathers would be able to perform an elaborate courtship display, thus ensuring successful reproduction. A peahen would choose such a peacock as the mate because the elaborate and colourful plumage and courtship display would indicate healthy genes, ensuring good genes are passed onto her offspring. The male ornamentation is basically targets of female choice.

There are two types of sexual selection: intersexual (female's choice) and intrasexual (male-male competition). Intersexual selection is when the female has a choice of the mate, and she chooses the mate on the basis of the sexual traits or the exhibition of male dominance (Byers and Waits 2006). The gain obtained could either be direct such as resources and safety, or in other case the gain could be indirect, i.e. creation of offspring with superior quality of genes. The above-stated example also exhibits mate choice of peahen of peacock with highest number of ocelli in the tail feathers, thus exhibiting superior genes. Fisher attempted to explain this evolution of costly characters with his 'runaway theory.' Earlier before the evolution of female choice, the peacock might not have been prevalent with long tails. Randomly a mutant female chose peacock with long tail which also had higher fitness associated with it. The peahen will now produce peacock with averagely long

tail and higher fitness. Slowly, the population will be replaced by peacocks with long tail and peahen who choose the peacock with this attribute. The evolution of long tail of peacock and peahen (with mother's genes of choice and father's genes for long tail) with such preferential choice would reinforce each other resulting in the evolution of long tails (Fig. 22.6).

A study was carried out on the pronghorns of the National Bison Range in northwestern Montana. The sample population was ear tagged and genotyped to ten microsatellite loci. In the study which spanned for 4 years, it was seen that 59% of the fawns were fathered by a small group of males who were physically attractive. Male attractiveness was associated with the offspring survival to weaning with the help of general linear latent and mixed models program, a statistical model. Fawn deaths are basically caused by vulnerability to coyote predation. 50 days post their birth, the fawns are able to gain speed in their sprint and thus escape predation. Thus there would be a differential growth rate of the fawns depending on the rate of weaning. The study showed that the fawns of preferred attractive males had faster growth (hind foot length was measured) and thus greater chances of survival. Thus female investment in mate sampling results in higher fitness by producing offspring with superior genes.

Intrasexual selection mainly includes competition among male to be able to mate with the female. The male-male competition is usually observed in polygynous mating system. This competition often results in intense contests to prove the superiority of the individual. Sometimes the competition is extremely intense resulting in fights. This has resulted in evolution of large body size or modes of fighting (such as horns and antlers). The winner of these intense competitions usually gets access to the females as they display superior and good quality genes.

22.4.4 Concept of Fitness, Selection Coefficient, Genetic Load and Genetic Death

22.4.4.1 Fitness

Fitness in the simplest term may be defined as the ability of an organism, rarely a population or species, to survive and reproduce in its adapted environment. If the organism reproduces successfully, it consequently contributes its genes to the next generation.

Thus, in order to estimate fitness, it is thus important to understand the various components of fitness:

- Viability which defines the survival of the newly formed zygote up to the reproductive age.
- Fecundity which defines the number of offspring produced in the next generation.

Together these components decide the extent of contribution of a genotype or allele to the next generation. Each allele does not enjoy the same fitness through time but changes with the changing environment. The terms dominant and recessive allele, only define the expression of the phenotype at a particular locus. The

changing frequency of allele however is influenced by the caring for young one. The fancy display of feather by male might seem to endanger the survival of the adult, as they become vulnerable to the predator. They are important in attracting the opposite sex and ensuring the survival of the young ones (Alcock 2005). This ensures the increased fitness of the individual and contribution of its genes in the next generation.

Fitness mathematically involves two terms: absolute and relative fitness.

Absolute fitness is the total fitness of a genotype which includes viability, successful reproduction, no. of viable offspring produced, etc.; it is represented as W and can be greater than or equal to 1.

The geneticists, however, more often use the term ‘fitness’ for relative fitness of an organism. It is represented by w and may be defined as the survival/reproductive rate of a genotype in comparison to the maximum survival/reproductive rate of other genotypes in the environment. It is also known as survival or adaptive value.

For example, $w_{AA} = 1$ represents the relative fitness of the genotype AA, $w_{Aa} = 0.8$ represents the relative fitness of genotype Aa and $w_{aa} = 0.7$ represents the relative fitness of genotype aa. This means all individuals of genotype AA, 80% of genotype Aa and 70% of the genotype aa would survive in the given environment. Of the three genotypes, AA is considered to be most fit.

22.4.4.2 Selection Coefficient

The measure of fitness and effect of selection is measured in terms of selection coefficient. It is the measure of the extent of natural selection against the contribution of a genotype to the next generation. It is symbolized as ‘ s ’ and lies between 0 and 1. Let us look at a hypothetical case where the population comprises the genotype AA, Aa and aa to understand this concept. These genotypes have different chances of survival as seen in Table 22.3.

The selection acts against the recessive allele ‘ a ’, and thus there will be a smaller number of surviving adults of genotype ‘aa’. If $s = 0.1$ then the chances of survival of genotype ‘aa’ are 90% as against 100% survival of genotypes AA and Aa. This is a mathematical model; however, in a real environment, the chances of survival, from birth to reproductive age, of an organism of even the best genotype are only 50%.

22.4.4.3 Genetic Load and Genetic Death

The concept of genetic load was introduced by Haldane where he observed the loss of population fitness due to selection-mutation balance. Therefore, in order to understand genetic load, we need to once again refer to our concept of fitness. As we have already discussed, a population comprises genotypes of varying fitness. The genotype with highest fitness is assigned a relative fitness value of 1. The rest of the genotype have relative fitness values less than 1. In spite of the fact that the selection

Table 22.3 Chances of survival of different genotypes

Genotype	Chances of survival
AA, Aa	1
aa	1-s

Table 22.4 Relative fitness of different genotypes

Genotype	Survival to reproduce	Relative fitness
AA	40%	0.8
Aa	50%	1
aa	30%	0.6

is continually acting, these alleles with suboptimal fitness remain in the gene pool either because the allelic variant resulting in reduction in fitness is being replenished in the gene pool due to mutation (known as mutational load) or because they remain in combination with advantageous alleles (known as segregational load). We can also measure the average fitness of a population. It is defined as mean fitness (w) and equals to the frequency multiplied by fitness of genotype. The genetic load basically measures the relative chance that an average individual will die before the reproductive age and thus have no contribution to the next generation because of the deleterious allele in it. It can also be defined as the sum of deleterious genes in the genome. It is symbolized as 'L' and lies between 0 and 1.

$$L = 1 - w$$

If all the individuals of a population have fitness 1, then there is no load on the population.

Let us look at an example to understand genetic load: The frequency of the two alleles A and a is 0.5.

Calculating from Table 22.4:

$$\text{Mean fitness} = 0.25(0.8) + 0.5(1) + 0.25(0.6) = 0.85$$

$$\text{The load of the population } L = 1 - 0.85 = 0.15$$

That means 15% of the offspring will die before the reproductive age, i.e. undergo genetic death.

The failure of individuals to produce offspring and contribute to the next generation is called **genetic death**. High genetic load can put the population in the danger of extinction. The marine life has been observed to have maximum genetic load in comparison to freshwater or terrestrial species. The bivalves among the marine species show highest genetic load due to small population size and high mutation rate.

22.5 Population Genetics

A population is a group of organisms of the same species living in a defined geographical area and can reproductively interbreed. Population genetics is an important aspect of evolution and speaks about gene and allele frequencies. It is

the study of genetic make-up of the population, i.e. gene pool and change in the gene pool over time.

Evolution and variation in gene pool are population phenomena so they are best understood as changes in allele frequencies. There are four characteristics which account for most of the changes in allele frequencies:

1. Mutation: produces genetic variation in gene pool and contributes to the first step of evolution.
2. Natural selection: adaptive, directional changes in allele frequencies.
3. Genetic drift: random, non-adaptive and non-directional changes in allele frequencies.
4. Migration: presence of gene flow.

The rates of mutation are very low in nature, so the primary contribution of mutation is only in production of genetic variation. It is migration, genetic drift or natural selection which acts on these genetic variations to produce change in the allele frequencies.

It might seem that a population which is well adapted to a geographic area would actually harbour high levels of homozygosity. But to the surprise of many evolutionary biologists, there is considerable genetic variation in the gene pool. In fact, for a population to succeed, it should have genetic variability.

22.5.1 Hardy-Weinberg Model

While studying single gene locus in a population, we observe the changes in allele and genotype frequencies. The determination of the change in allele and genotype frequencies of the population, from one generation to the other, forms the major study of the population geneticists. The relation between changing allele frequencies leading to change in genotype frequencies has been explained by Hardy-Weinberg law. G.H. Hardy and Wilhelm Weinberg in 1908 independently formulated a simple equation which can be used to trace the allele and genotype frequencies of the population in an ideal scenario. The Hardy-Weinberg law states that in an ideal, infinitely large population with random mating, on which no evolutionary force is acting, the allele frequency does not change and the genotype frequencies stabilize after one generation.

Assumptions made in Hardy-Weinberg law:

1. Infinitely large population.
2. Random mating.
3. Absence of evolutionary forces like:
 - (a) Natural selection—thus all individuals have equal rates of survival and reproduction.
 - (b) Migration—absence of gene flow.
 - (c) Mutation—as a result no new alleles are created.

Fig. 22.7 Punnett square showing permutation and combinations of genotype produced by random matings. Each allele has equal probability of pairing with other alleles producing genotypes p^2 , $2pq$, q^2

sperm egg	f(A)=p	f(a)=q
f(A)=p	$p \times p = p^2$	$p \times q = pq$
f(a)=q	$q \times p = qp$	$q \times q = q^2$

A population which meets the above criteria are said to be in Hardy-Weinberg equilibrium. The Hardy-Weinberg law is a mathematical model and uses the Mendelian laws of segregation.

If two alleles A and a of a single locus are considered, then the genotype frequencies AA , Aa and aa if in Hardy-Weinberg equilibrium would be p^2 , q^2 and $2pq$, where the frequency of $A = p$ and frequency of $a = q$.

The above genotype can only be obtained when there is random mating in the population resulting in the random combination of gametes. This can be understood by the study of Punnett square often used by the twentieth-century geneticists to predict probability of genotypes.

From the Punnett square (Fig. 22.7), it is clear that each allele has equal probability with pairing with the other producing the genotypes p^2 , $2pq$ and q^2 if there is random mating.

The Hardy-Weinberg equilibrium equation is:

$$p^2 + 2pq + q^2 = 1$$

Since we are taking into consideration a single locus with two alleles (A and a), these alleles should account for the 100% frequency of the gene in the gene pool.

Or in other words, $p + q = 1$

To demonstrate how Hardy-Weinberg law of equilibrium can be used, let us consider a population with $T = 0.7$ and $t = 0.3$.

$$f(T) + f(t) = p + q = 0.7 + 0.3 = 1$$

If the population undergoes random mating, then we attain three genotypes TT , Tt and tt in the proportion as discussed above: p^2 , $2pq$ and q^2 :

$$f(TT) = p^2 = 0.7 \times 0.7 = 0.49$$

$$f(tt) = q^2 = 0.3 \times 0.3 = 0.09$$

$$f(Tt) = 2pq = 2 \times 0.3 \times 0.7 = 0.42$$

$0.49 + 0.42 + 0.09 = 1$ showing that we have accounted for all zygotes formed.

Assuming they follow Hardy-Weinberg law, these offspring have equal chances of survival, and they become adults. They also have equal probability to reproduce and mate randomly. Forty-nine percent of the gametes would be contributed by the genotype TT , 42% by Tt and only 9% by tt .

The gamete T would be contributed by both the genotypes TT and Tt . Thus

$$f(T) = 0.49 + \frac{1}{2} 0.42 = 0.7$$

The gamete t would be contributed by both the genotypes Tt and tt . Thus

$$f(t) = 0.09 + \frac{1}{2} 0.42 = 0.3$$

The initial gene pool that we started with had the frequency of alleles T and t in the proportion 0.7 and 0.3. After a generation of random mating, the alleles remain in the same proportion. Thus, we can say in absence of evolutionary force, this population has remained in Hardy-Weinberg equilibrium and has not exhibited any change in allele frequencies.

Thus, we can draw two inferences from this model:

1. Allele frequencies remain the same from one generation to another.
2. Genotype frequencies stabilize after one generation of random selection.

However, the real population is rarely in Hardy-Weinberg equilibrium. Then why is this model of great importance to the population geneticists? The Hardy-Weinberg model helps the population geneticists to understand what the evolutionary forces are acting on the real-world population and changing the allele frequencies (Jankowska et al. 2011; Klugs et al. 2012).

22.5.1.1 Application of Hardy-Weinberg Law on Human Population

Hardy-Weinberg law can be applied on human population to test if the population is undergoing evolution or not. It is used to calculate the percentage of population carrying autosomal recessive alleles causing inherited diseases and for biomedical research. One such study was done on population of the USA to study the occurrence of phenylketonuria (PKU), a metabolic disease due to homozygosity of recessive allele. It results in mental retardation, stunted growth and other symptoms. The frequency of occurrence of this disease is 1 in every 10,000 babies born in the USA. Considering that no new mutations for PKU are being added into the gene pool, we can use this data to calculate the percentage of carriers in the population.

Using the Hardy-Weinberg equation: $p^2 + 2pq + q^2 = 1$
 $q^2 = 1/10000 = 0.0001$

$$q = \sqrt{0.0001} = 0.01$$

$$p + q = 1$$

$$p = 1 - q = 1 - 0.01 = 0.099$$

$$\text{The number of carrier (heterozygotes)} = 2pq = 2 \times 0.099 \times 0.01 = 0.198.$$

Approximately 2% of the population are the carriers of the recessive allele. This estimate is not exact, but we can get an idea of an approximate idea of the carriers in the population.

Another such genetic disease is haemophilia, a genetic disease which impairs the body's ability to make blood clot. The occurrence of haemophilia A and B in the population is 1/12000. Using Hardy-Weinberg's equation, we can estimate the carriers in population.

$$q^2 = 1/12000 = 0.00008333$$

$$q = \sqrt{0.00008333} = 0.0091$$

$$p + q = 1$$

$$p = 1 - q = 1 - 0.0091 = 0.9909$$

$$\text{The number of carrier (heterozygotes)} = 2pq = 2 \times 0.9909 \times 0.0091 = 0.01803.$$

Thus the frequency of occurrence of carriers in the population is 0.01803 which is very low (Klugs et al. 2012; Pierce 2012).

22.5.2 Genetic Drift

When the size of the population is small, then chance alone can result in the change in allele frequency. The smaller the size, the greater can be the degree of fluctuation of allele frequencies. Any random, non-adaptive, non-directional change in allele frequency occurring in a small population is called genetic drift. Once it begins, the phenomenon of genetic drift will continue in subsequent generations till an allele is either completely lost or fixed in a population. The concept of genetic drift was introduced by one of the founding fathers of population genetics, Sewall Wright in 1931, and is also known as Sewall Wright effect (Wade 2008).

This concept can be understood by random sampling in genetics. Let us consider a single locus gene with two alleles A and a, in a population of ten individuals. The allele frequency of A and a is 0.5 each in a gene pool of genotype AA, Aa and aa. Each of these genotypes has equal probability of survival and reproduction, i.e. in absence of natural selection, the fitness of all the genotypes is 1. In accordance with the Hardy-Weinberg principle, the allele frequency should remain the same in the next generation. However random sampling may change the scenario, and by chance the allele 'A' might get a better environment for reproduction resulting in increase in the frequency of allele 'A'. The allele A has no adaptive value for the environment, and this increase in frequency is just a matter of chance. Such random change in allele frequency is significant only in a small population and results in subsequent changes in genotype frequencies.

The genetic drift can act through two phenomena:

- (a) Founder's effect.
- (b) Bottleneck effect.

22.5.2.1 Founder's Effect

The establishment of new population formed by a few founding members from the parental population which are not representative of the original gene pool is known as founder's effect. This concept was introduced by Ernst Mayr in 1942.

Several evidences of founder's effect have been seen in humans such as prevalence of some recessive diseases in selective populations. These usually occur in small or isolated population as observed in Finns. There are 30 recessive diseases which are common in Finland than rest of the world. Diseases like cystic fibrosis and phenylketonuria are common in Caucasian population. Studies have shown that these diseases have arisen from one remote mutation in the immigrants of the founder's population, thus exhibiting the founder's effect (Peltonen 2001).

22.5.2.2 Bottleneck Effect

Bottleneck effect occurs usually because of a disaster or natural calamity which drastically reduces the population. The surviving population does not represent the genetic make-up of the parental population.

An example illustrating bottleneck effect is the loss of genetic variation in Northern elephant seals. Due to unchecked hunting, their number reduced to as low as 20 at the end of the nineteenth century. Conservation acts resulted in rebound of their population but with low genetic variation showing the mark of bottleneck effect.

Similar loss of genetic diversity has also been observed in cheetah (*Acinonyx jubatus*). Cheetahs are inbred with extremely low genetic diversity. Investigations have been carried out by Stephen J. O'Brien and his colleagues where they examined the genomes of seven cheetahs, four from Namibia and three from Tanzania. Genome sequencing confirmed the fact that there is very low genetic diversity in the gene pool of this animal. Low genetic diversity poses a threat to the extinction of the species.

The initial loss of diversity happened around 10,000 years ago, when they just managed to survive the ice age. Studies have been carried out on hypervariable minisatellite loci and mitochondrial DNA to date this bottleneck event to the late Pleistocene epoch. Modern times have increased the problems of habitat encroachment, hunting and poaching making them vulnerable to extinction. Conservation biologists are working hard to repopulate them, but their gene pool exhibits low genetic diversity due to these events (Raymond and O'Brien 1993; Emerling 2016).

22.5.3 Role of Mutation and Migration in Changing Allele Frequency

22.5.3.1 Mutation

Mutation is any change in DNA sequence of the gene within the chromosome which occurs due to error in DNA replication. Mutation in fact is the only phenomenon which can produce new alleles. All other evolutionary forces only reshuffle the gene pool to produce variable genotypes. Mutations are very important as they produce variation in the gene pool on which evolution acts. In other words, mutations are the

raw material for evolution or are the engine of evolution. Mutations are random events without any adaptation value. They can either be selected and become abundant or completely lost from the gene pool.

Mutation, in itself, is a weak evolutionary force to change allele frequency but a strong force to create genetic variation. However, it is difficult to measure mutation in a diploid organism as most mutations are recessive in nature.

22.5.3.2 Migration

Migration is the movement of a subpopulation from one place to another. The migrating population carries its own ancestral genes and interbreeds with the native subpopulation resulting in sudden influx of alleles. This transfer of genes results in gene flow. Thus if the two populations have different set of genes and in absence of any selection phenomenon, migration alone can result in change in genotype frequency. In order to understand it, let us look at an example where a hypothetical species has two alleles A and a. The species also has two populations, one residing in the mainland and the other in the island. The frequency of A on mainland is represented by p_m and frequency of A on island by p_i . If migration takes place from the mainland to island, then the frequency of A in the next generation on island p_{i1} is represented by:

$$p_{i1} = (1 - m) p_i + m p_m$$

where m represents the migrants from mainland to island. Putting in the value of $p_m = 0.7$ and $p_i = 0.3$ and that at a time 10% of the migrants have moved from mainland to island,

$$\begin{aligned} p_{i1} &= (1 - 0.1) 0.3 + 0.1 \times 0.7 \\ &= 0.9 \times 0.3 + 0.07 \\ &= 0.27 + 0.07 \\ &= 0.34 \end{aligned}$$

These calculations show change in allele frequency due to migration from 0.3 to 0.34 in one generation itself.

There is considerable influence of human migration in distribution of ABO blood group (Mourant et al. 1976). Karl Landsteiner, an Austrian physician, has been credited with the discovery of ABO blood group. The ABO blood group is controlled by a single gene with three alleles, present on chromosome 9. The native Americans can be traced to a founder population of 10–20 individuals who had migrated to the American mainland. This has been illustrated through study of mtDNA and Y chromosome. The Americans thus have a high percentage of O blood group. The D_i blood group polymorphism tracks the migration of humans from East Asia to America. Another such example is the prevalence of D_i antigen in South East Poland which provides the measure of extent of invasion of Mongolians in recent times in Europe.

22.5.4 Impact of Positive Selection

Selection can be of two types: positive and negative or purifying selection. Purifying selection prohibits the spread of deleterious mutations in the gene pool.

Positive selection is also known as Darwinian selection. It is the phenomenon of natural selection by which advantageous mutation becomes fixed in a population, or in other words it promotes the spread of advantageous mutation in a gene pool. Positive selection thus promotes the emergence of new phenotype.

Charles Darwin in his explanation of selection had stated that those organisms which have the best attributes are the ones that survive in an environment. He was mainly concerned with phenotypic evolution. As we are looking at evolution in terms of genetics, let us redefine this outlook. The organisms that harbour mutations which increase their fitness in the environment, are the ones which survive and reproduce (Forsdyke 2007).

Whether the mutations that occur in the nucleic acid sequence result in positive or negative selection depends on which part of the gene product or protein they are affecting. If a mutation occurs in the active site of the enzyme such that it lowers the catalysing rate of the enzyme, it might result in lowering the fitness of the organism. In other case, mutation in the antigen might enhance the ability of the pathogen to invade the host, thus increasing its fitness and resulting in positive selection.

Extensive studies have been done on positive selection. One such study reviews large number of genes of the human population which have undergone positive selection (Wu and Zhang 2008). Darwinian selection has intensively acted in the modern human population resulting in high genetic diversity which has resulted in differences in appearances, metabolism of drugs and resistance to diseases. One such set of genes are those that are involved in the development of the brain. The brain size has increased in primates specially in the *Homo sapiens* and the species closely related to them. Some genes involved in development of brain have undergone rapid positive selection. Microcephalin is the key regulating gene of the brain of human and is still evolving. *FOXP2* is another gene present both in human and birds. It regulates the singing ability of the birds and speech expression in man. The copy of *FOXP2* in human has high evolutionary rate under the influence of positive selection.

Another such example was a study carried out by Zhang et al. (2002), on the evolution of a duplicated pancreatic ribonuclease gene (*RNASE1*) of leaf-eating colobines. Like ruminants these old world monkeys extract nutrients by breaking down the symbiotic bacteria with a set of enzymes including *RNASE1*. Phylogenetic analyses of the *RNASE1* gene of the non-colobine monkeys with the Asian colobine douc langur (*Pygathrix nemaeus*) show the substantial difference in the sequence. A closer examination shows that one copy of the gene *RNASE1* has remained conserved but the other copy of the gene *RNASE1B* has accumulated many non-synonymous substitutions post recent duplication. These rapid substitutions have accumulated due to positive selection pressure for adaptation of enhanced ribonuclease activity at low pH of the colobine intestine (Zhang 2010).

22.6 Speciation

22.6.1 Definition of Species

Several workers have tried to define the concept of 'species,' and the most important fact to be noted is that these individuals belonging to a species co-exist in a particular span of time.

22.6.1.1 Typological Species Concept

This concept suggests that 'variety in nature can be reduced to a few distinct types.' Varied individuals often belong to a single type of species that is defined by a certain set of norms. The concept was proposed centuries ago by Plato and fails to recognize a species as a unit of evolution. Hence it was not accepted by modern evolutionists.

22.6.1.2 Biological Species Concept

The 'Biological Species Concept' proposed by Harvard evolutionist Ernst Mayr (1963) states 'Species are groups of interbreeding natural populations that are reproductively isolated from other such groups.' Species is an ecological unit sharing resources of the environment and interacting as a unit with other such groups. This concept focuses on species, in genetic terms as a gene pool, a unit within which the gene frequencies can change. The movement of genes through a species via mechanisms such as immigration, emigration and interbreeding is termed as gene flow. Natural selection favours inheritance of genes that are beneficial to the organisms as well as that help them in better co-existence with the environment. A gene that does not favour these aspects is selected against. An important limitation of the biological species concept is that it does not apply to the organisms that reproduce asexually.

22.6.1.3 Recognition Species Concept

The concept states 'species as a collection of individuals with shared specific mate recognition system (SMRS).' It was given by H. Paterson in 1993. The idea behind placing individuals together as a species is interbreeding and production of a viable, fertile offspring. And this concept highlights the method by which potential conspecific mates are recognized for breeding. The sensory systems may be visual, acoustic, chemical, etc. An example citing the same is a population of crickets within a single habitat in the USA. Their courtship begins with the males singing for attracting the females. The songs are species specific and hence confine interbreeding to a single species.

22.6.1.4 Morphological Species Concept

Species are characterized by body shape and other structural features and can be applied to sexual as well as asexual organisms.

22.6.1.5 Ecological Species Concept

The ecological niche, that is, the microenvironment of a species, defines a species population. The concept highlights the interaction of the members with the living and non-living parts of their environment.

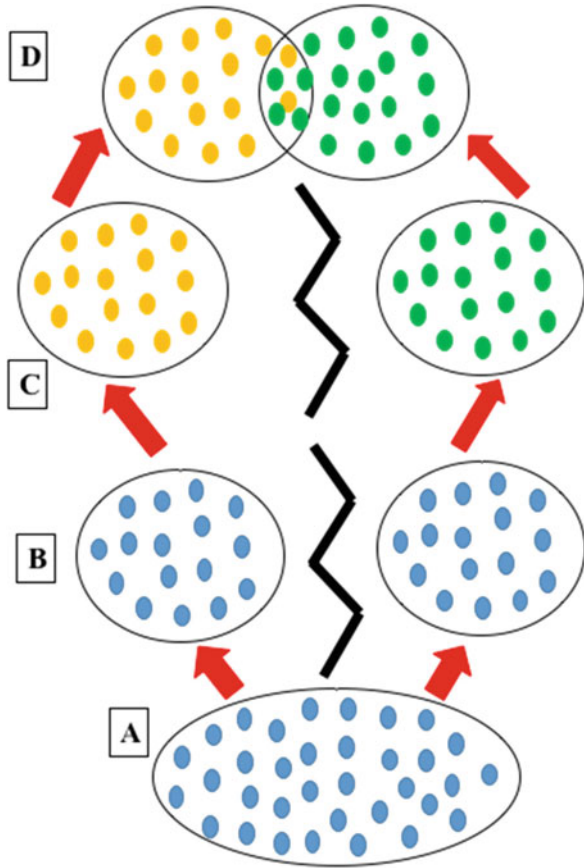
In a population that is sexually reproducing, speciation leads to the division of a parental gene pool into two or more distinct gene pools. The concept of genetic changes accumulating to form a new species is a macroevolutionary event. The underlying cause of speciation is the adaptation of populations to discontinuous ecological niches, leading to reproductive isolation that allows them to evolve independently.

22.6.2 Modes of Speciation

Speciation is classified as allopatric, sympatric, peripatric and parapatric (Fitzpatrick et al. 2009).

- If a geographical barrier is responsible for divergence of a population from its ancestor, it is termed as **allopatric speciation** (Fig. 22.8). The physical barrier prevents mixing of the population, interbreeding and hence gene flow. The divided populations hence formed start developing independently of each other, accumulating genetic differences between them as a result of mutations, genetic drift and natural selection. The presence of related flora and fauna in Galapagos Islands off the west coast of South America is a key evidence of allopatric speciation. These have been discussed in detail by Charles Darwin in his book *Origin of Species*. Over 14 species of finches are distributed all over these islands that have accumulated differences from their ancestors with respect to their food preferences and beaks. The antelope squirrel *Ammospermophilus harrisi* inhabits the southern rim of the Grand Canyon and the northern rim by the white-tailed *Ammospermophilus leucurus*. Small rodents are unable to cross such barriers unlike birds and hence have diverged into two different species.
- **Sympatric speciation** occurs in geographically overlapping areas (Fig. 22.9). Divergence occurs due to reduced gene flow by polyploidy, sexual selection and habitat differentiation. An example is presence of cichlid species in East African Great Lakes that are known to be monophyletic (Schliewen et al. 1994). Analysis of their mitochondrial DNA revealed that these flocks have undergone sympatric diversification. Sympatric speciation is also said to be underway in pest of apples, North American maggot fly *Rhagoletis pomonella*. Its original habitat was the hawthorn tree, but now it colonizes apple trees. Apples mature sooner compared to hawthorn fruits; natural selection favoured the apple-feeding flies. The apple-feeding flies show temporal isolation from the hawthorn feeders providing a prezygotic isolating barrier to gene flow as well.
- Ernst Mayr described the founder effect speciation which was later termed as **peripatric speciation** by him. His idea elaborated that in isolated populations with restricted distributions, the peripherally located species to the parent species

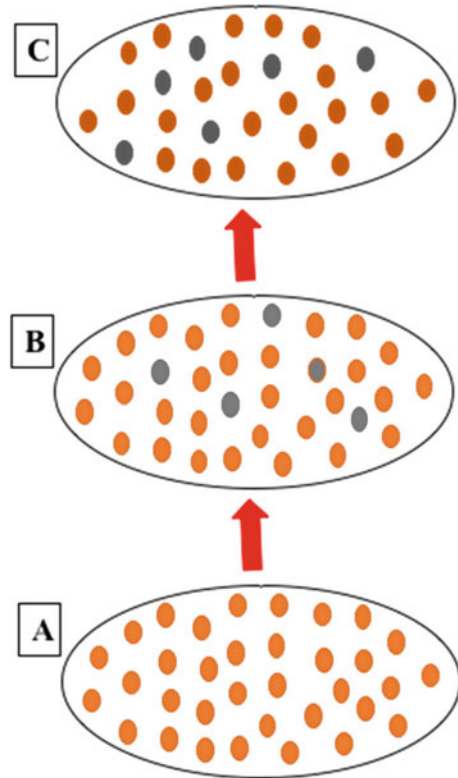
Fig. 22.8 Sequence of events in allopatric speciation. *A*: The original parent population. *B*: The population is split by a geographical barrier. *C*: Owing to the distance and isolating mechanisms that act, the populations become genetically different by means of drift and selection. *D*: As a result of the events, the populations have become independent species, have adapted to their unique niche and cannot interbreed anymore. Speciation is complete when no gene flow occurs even if they occupy the same geographical area



tend to diverge genetically and form new genera. He reasoned that the allele frequencies at various loci differed from the parent population due to genetic drift. A common example is of the paradise kingfishers *Tanysiptera* in New Guinea. The *T. galatea* is present throughout lowlands of New Guinea, whereas several distinct forms (*T. riedelii*, *T. carolinae*) are distributed on the islands along its coast.

- If gene flow is weak between populations residing in adjacent regions with varied selection pressures, it leads to **parapatric speciation**. The hybrids so formed may be weak, with lower fitness, inviable and sterile. Steady genetic divergence would lead to complete reproductive isolation. An example is of the three-spined sticklebacks (*Gasterosteus aculeatus*) in lakes (each with outlet streams) of Vancouver Island in western Canada. Though there was an absence of any physical barriers between streams and lake populations, the subpopulations have evolved with various different morphological features. Genomic analysis has shown that genetic differences between the populations were pronounced in the central chromosomal regions (Roesti et al. 2012).

Fig. 22.9 Sequence of events in sympatric speciation. *A*: The original parent population. *B*: The subpopulations start getting formed in the absence of any geographical barrier and begin their genetic differentiation. *C*: The subspecies hence formed become completely reproductively isolated and cannot interbreed, thus forming independent species



22.6.3 Isolating Mechanisms

The barriers that lead to these events may be physical, geographical, physiological, temporal or ethological and are collectively termed as reproductive isolating mechanisms. These gene flow barriers evolve as a result of divergence between populations that accumulate genetic differences over time. Accumulation of these isolating mechanisms results in the process of speciation. The strength of isolation is assessed if the species merge back into a single lineage on coming back into contact with each other. Speciation is complete if no intermixing takes place and no hybrids are formed. However, if isolating mechanisms are weak and overcome by gene flow, they end up merging the species into a single lineage.

These are classified into two main categories with respect to their occurrence pre- or post-mating (Table 22.5):

- (a) Prezygotic isolating mechanisms: These act before fertilization and hence no zygote formation takes place. As a result of these mechanisms, no mating can take place.
- (b) Post-zygotic isolating mechanisms: These act post fertilization, the population members are willing to mate and the hybrid hence formed isn't fit to be either

Table 22.5 Isolating mechanisms operating in nature

S. no.	Mechanism	Example
Pre-mating/prezygotic mechanisms: no mating, no zygote formation (interspecific pairs do not form)		
1.	Geographical: Differences in preferred habitat/biotope hinder mating in animals and pollen exchange in plants	Two species of fern <i>primula</i> , <i>P. vulgaris</i> and <i>P. elatior</i> , have different soil, temperature and moisture preferences (Keller et al. 2016)
2.	Seasonal/temporal: Members of different populations differ in breeding time/time of attaining sexual maturity or flowering	Species of toads: <i>Bufo americanus</i> and <i>Bufo fowleri</i> have different breeding times (early and late in spring, respectively)
3.	Ethological/behavioural/psychological: Isolation occurs due to incompatible behaviour of the males and females of different species	Behaviour, courtship signals are species specific and enable mate recognition. For example, in birds and arthropods, songs, courtship dances, display and tactics are responsible for conspecific mating. Marine blue footed booby birds are recognized by their blue feet, a sexually selected trait. They mate post a courtship display that calls for the male to high step that enables the female to focus on his bright feet
4.	Mechanical: Reproductive organs (genitalia in animals, stamen and pistil in plants) of the species are physically incompatible	The 'Lock & key' theory has been proposed for genitalia of animals to correspond and mate with each other. For entomophilous flowers, floral parts become highly specialized to attract a single kind of pollinator and prevent interspecific pollination. Shells of two species of snail <i>Bradybaena</i> spiral in different directions (clockwise and counter-clockwise). As a result, their genital openings are not aligned, and mating is unsuccessful
Post-mating prezygotic mechanisms: Mating takes place, but the gametes do not get fertilized		
1.	Gamete viability: The sperms do not survive in the female reproductive tract; pollen fails to germinate	In some crosses of <i>drosophila</i> , an insemination reaction can result in swelling in the vagina of the female and hence unsuccessful fertilization (Turissini et al. 2018)
2.	Gamete recognition: Sperms and eggs of the two species are incompatible with each other; pollen and egg nuclei are also incompatible/fail to attract each other	Sea urchins of the genus <i>Echinometra</i> have external fertilization, where the sperms and eggs are laid into surrounding water and fuse to form zygotes. Hetero-specific gametes do not fuse with each other as the proteins on their surfaces cannot bind to each other
Post-zygotic mechanisms: mating takes place, and hybrids hence formed are sterile/non-viable		
1.	Hybrid inviability: Interspecific hybrids are weak, are non-viable with lower	Salamander subspecies <i>Ensatina</i> occasionally hybridize, and the resultant hybrids either do not complete their

(continued)

Table 22.5 (continued)

S. no.	Mechanism	Example
	fitness and hence get eliminated at any stage before attaining sexual maturity	development or, if they do, are frail in stature
2.	Hybrid sterility: Hybrids are sterile due to abnormal developments of their gonads or failure of gametes to develop	Genetic causes of such an event are at three levels: (a) Genic: Incompatible gene combinations create a disharmony and failure to form gametes (common in animal hybrids) (b) Chromosomal: Different gene arrangements and chromosomal aberrations (inversions, translocations, etc.) lead to sterility (common in plant hybrids) (c) Cytoplasmic: Unfavourable interactions between nuclear and cytoplasmic factors
3.	F₂ breakdown: Inviability or sterility of F ₂ hybrid even if F ₁ is vigorous	

viable or reproduce further. Hybrids are sterile, or even if they manage to reproduce, the progeny does not have high levels of fitness. Post-zygotic mechanisms are uneconomical as the hybrids hence produced have no genetic contributions to the gene pool and hence a waste of sexual energy. Mostly the parent populations have diverged so much that hybrids are inviable or sterile.

Reproductive isolation is not a product of a single gene mutation between two isolated populations; rather, it arises by gradual accumulation of differences in gene combinations, frequencies, arrangements and interactions between them.

22.6.4 Genetics of Speciation and Reproductive Isolation

Differences in various chromosomal segments and genes lead to strengthening of isolating mechanisms and finally result in speciation (Noor and Feder 2006). The hybrids formed due to a hetero-specific mating are inviable, weak and sterile as both paternal and maternal sets of chromosomes form disharmonious association. This may lead to a loss in function of several indispensable housekeeping genes due to epistasis or occurrence of homologous alleles on different chromosomes. Prezygotic mechanisms hence are more economical and do not allow investment of biological energy which may lead to formation of a wasteful end product (Wolf et al. 2010).

In prezygotic isolation, several differences in gene combinations and chromosomal segments have been established firmly by natural selection over a long span of time. When two such populations come in proximity, then do not seek potential mates within each other. Post-zygotic mechanisms reveal that the two populations under scrutiny haven't yet accumulated enough differences yet. Strongest genetic

dissimilarities exist in the mechanism resulting in inviable hybrids compared to others. Hence, natural selection is yet to intensify its divergent action on them. Hence, it can be concluded that species do not arise at a single step, but through an accumulation of various genic and chromosomal differences. At the genetic level, four problems are known to cause hybrid inviability/sterility/breakdown: chromosomal rearrangements, incompatibility at the gene level, ploidy and interaction between endosymbionts and the nuclear genome. When a gene at a specific locus from one parent does not interact well with a gene at another locus from the other parent, such genic incompatibilities affect the hetero-specific hybrid viability.

Also proposed in the 1930s was the Dobzhansky-Muller theory for post-zygotic isolation. According to the theory, an ancestral population splits into different populations in which gene flow is absent. Each population is well adapted to its local environment as a result of genetic changes that have accumulated over time. If the populations encounter each other later and mate, the genetic changes at the various loci will now allow the hybrids to be successful, and hence they will be sterile. The main aspect to be impressed upon is that post-zygotic isolation is due to interaction of multiple loci. An interesting pattern in post-zygotic isolation was inferred by J.B.S. Haldane, called Haldane's rule, that 'the heterogametic hybrid of a population will have lower fitness compared to the homogametic one.'

22.7 Evolution of Man

Apart from fossil evidences, the origin and evolution of man have been widely studied with the help of DNA sequences to trace the history of modern man. The relationship among man and its nearest relatives 'The great apes' has long been studied. There are major morphological differences including bipedalism, presence of an opposable thumb and body proportions distinguishing chimpanzees, gorillas and man.

All early hominid fossils have been procured from Africa, but the first fossils found outside were of *Homo erectus* (China and Indonesia) (Fig. 22.10). It is postulated that *H. erectus* gave rise to archaic European, Asian and African populations. The best known of such hominids are the Neanderthals (*Homo neanderthalensis*) that lived in Europe and Western Asia about 3,00,000 years ago. These fossils were obtained from Feldhofer cave in Germany and Mezmaiskaya cave in Caucasus Mountains east of Black Sea. Mitochondrial DNA analysis from Neanderthal fossils with 2000 present-day human samples suggests that they did not contribute to the mitochondrial DNA of *Homo sapiens*. They may have competed with ancestors of modern humans and lost out in the competition and became extinct. Another approach was the whole genome sequencing of the Neanderthal species. Their genome is made up of 3.2 billion base pairs and is 99.7% identical to the modern human genome. Comparisons of the Neanderthal genome, five present-day humans and chimpanzee genome revealed that there are amino acid coding differences in 78 genes. Interestingly it was revealed that 1–4% of Neanderthal sequences were found in humans from Europe and Asia but not Africa. Phylogenetic

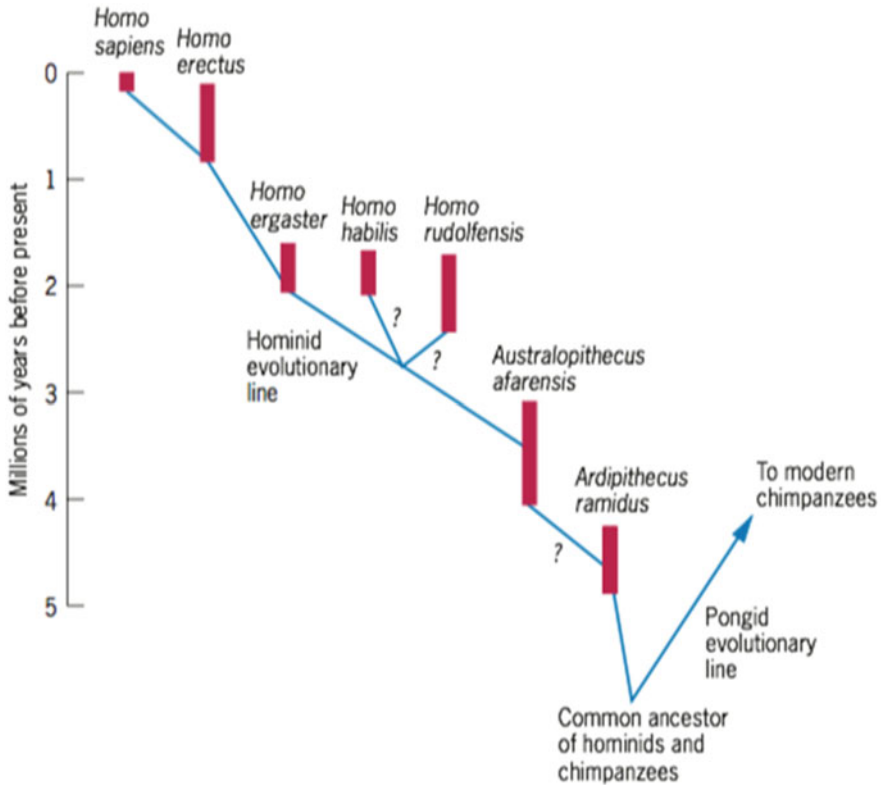


Fig. 22.10 Evolutionary history of the modern man. *Homo sapiens* evolved from a common ancestor in parallel with chimpanzees as traced by fossil evidences procured from over the globe. Uncertainties in the line are indicated by question marks. (Adapted from Snustad & Simmons; 6th Edition)

analysis done to compare and analyse the divergence of the species revealed that humans and Neanderthals last shared an ancestor 706,000 years ago (Fig. 22.11). The mitochondrial genome of 12 Neanderthal specimens has been completely sequenced, and these are quite different from the known human mtDNA. It is unlikely that their mtDNA made any significant contributions to modern human mtDNA. Hence, modern man and Neanderthal man are considered as clear biological species (Hartl and Jones 2009).

The initial publication of the chimpanzee (*Pan troglodytes*) genome and comparison with the human genome (The Chimpanzee Sequencing and Analysis Consortium, 2005) has shed light on the formation of the human species and the complex speciation between the two. In addition to this was the sequencing of the genome of rhesus macaque (*Macaca mulatta*) by the Rhesus Macaque Genome Sequencing and Analysis Consortium in 2007. It was possible to compare the three primate genomes and construct an ancestral primate genome. Comparative genomics could also

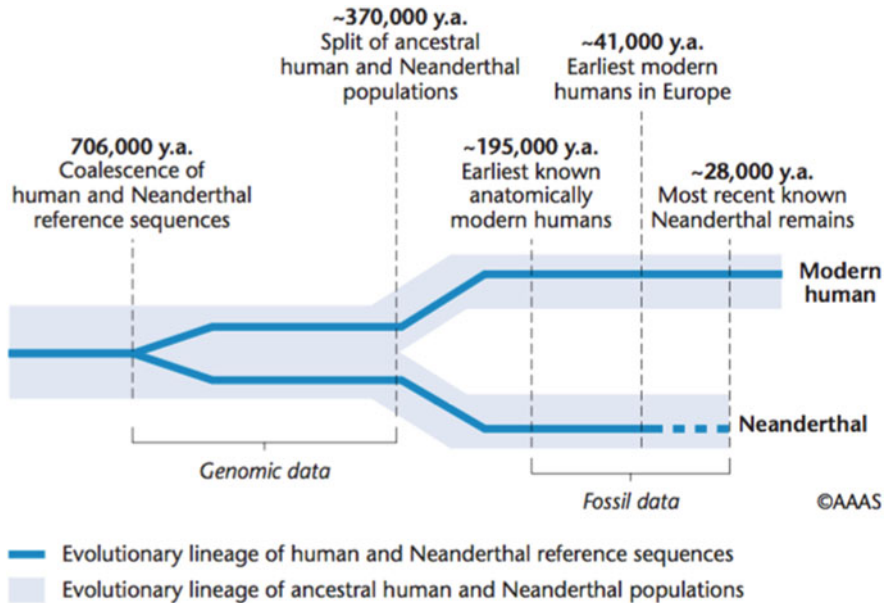


Fig. 22.11 Divergence between the human and the Neanderthal species. The separation events led to major evolutionary events in both populations. Data obtained on sequencing comparisons between the genomes of modern humans and the Neanderthal DNA. (Adapted from Concepts of Genetics by Klug – 10th Edition)

determine the regions of the ancestor that could have contributed to human evolution. Chimpanzees and humans have 98% nucleotide level similarity (Portin 2007). The main difference between the haploid chromosomal sets of the two in a karyotype is a big metacentric chromosome 2 in man vs acrocentric in chimpanzee. The sequencing of the Y chromosome of both the species revealed that there is an accelerated rate of evolution of the chromosome vs the entire genome. It is a huge challenge till date to explain the reasons of emergence of man post their separation less than 6.3 million years ago.

As of today, genetic data such as variations at the level of blood groups, restriction fragment length polymorphisms, lengths of repeat DNA sequences and DNA composition have been used to investigate relatedness amongst various populations, races and ethnic groups. Most analysis of human evolution has been done using mitochondrial DNA as it evolves faster than the nuclear DNA and is passed on only through the maternal parent. Hence, in evolutionary terms researchers can detect changes at a genic level over short period of time and trace it back to a common female ancestor.

22.8 Concept of Molecular Phylogeny

The concept of classification and systematic organization of biological hierarchy is an age-old concept introduced by the naturalist Linnaeus in the eighteenth century. He classified mainly to arrange the organisms in a hierarchical pattern, but unknowingly he also paved path for understanding phylogeny of organisms. Phylogeny is the study of evolutionary relationships. Earlier the direct phylogenetic studies were mainly based on the fossil evidence. Other studies included comparison of morphological traits. The main problem with these methods was that the fossil data was often incomplete and comparison of morphological traits was often biased. Moreover, study on microorganisms could not use any of these mentioned methods. The introduction of molecular phylogeny broadened the horizons where phylogenetic relationships could also be derived among evolutionary distant organisms. The molecular phylogeny uses the sequences of the biomolecules to look into the evolution of organisms. It is basically selection acting on the mutation in the biomolecules. The accumulated mutations over millions of years can act as molecular fossils. Though sequencing was first introduced only in the 1970s, but Nuttall had used immunological assays in the early 1900s to understand the evolutionary position of man in relation to other primates. Even though Nuttall had displayed the successful use of biomolecules in understanding phylogeny as early as 1904 yet it gained momentum only in the 1950s because of several technical challenges. There has now been gradual shift towards molecular phylogeny, which involves the use of DNA, RNA and protein as molecular markers to study modern phylogenetics. The major reason of this shift is to be able to obtain large datasets for the studies. The evolution of sequencing has also ensured easy availability of molecular data. The molecular data can be easily converted into mathematical and statistical data which makes it easier to analyse. Unlike the morphological traits, the molecular data (A, T/U, G, C for DNA/RNA and amino acids for proteins) are unambiguous (Brown 2002).

22.8.1 Phylogenetic Tree

The molecular data has been largely used to construct phylogenetic trees. A phylogenetic tree (Fig. 22.12) is a visual display of the evolutionary relationships among organisms. Even though the phylogenetic tree is being constructed now to display phylogenetic events, but tree-like illustrations (Fig. 22.13) were also observed in Darwin's book *Origin of Species*, where he used it to show that accumulation of slow modification can lead to speciation event.

22.8.1.1 Objectives of Molecular Phylogeny

1. To infer a tree displaying true phylogenetic relationships among organisms.
2. To study and recover the order of evolutionary events and represent it as phylogenetic tree, which is a graphical representation of phylogenetic relationship.

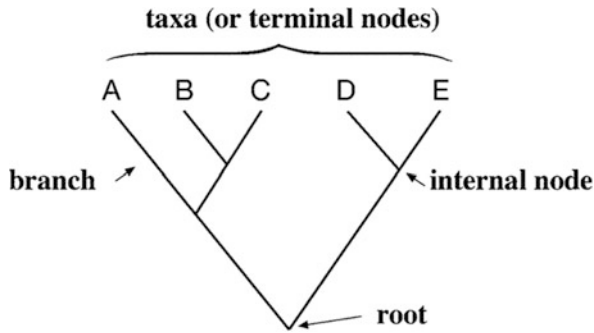


Fig. 22.12 A typical rooted phylogenetic tree. Diagram showing various parts of the tree—terminal nodes or taxa (operational taxonomic unit, OTU) are the extant species, internal nodes are the recent common ancestors, branching shows the event of divergence and root is the common ancestor. Often the details of common ancestor are not available, so to root a tree an outgroup species is used. Outgroup is the species which is distantly related to the group of organisms. The pattern of branching of the tree is known as tree topology. Ninety-nine percent of the species on earth have become extinct. A tree like this gives a visual display inferring what would have been the phylogenetic relationship of the extant species with the extinct species

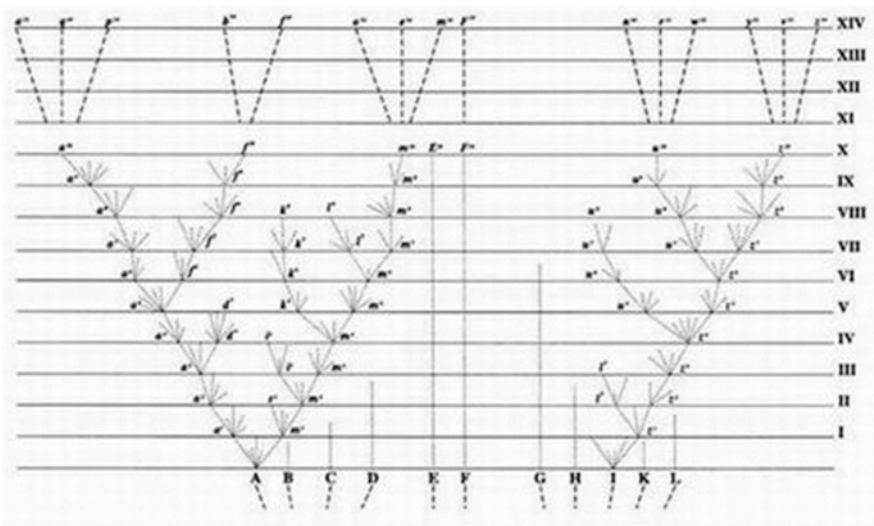


Fig. 22.13 Darwin’s illustration. Darwin too in his book *Origin of species* had made a tree-like illustration which expressed evolution. (Adapted from Karen Dowell 2008)

3. To be able to estimate the evolutionary date of divergence of organisms from their common ancestor.

Since it is thought that all organisms have arisen from the Last Universal Common Ancestor (LUCA), objectively there should be a single tree of life. However, it is

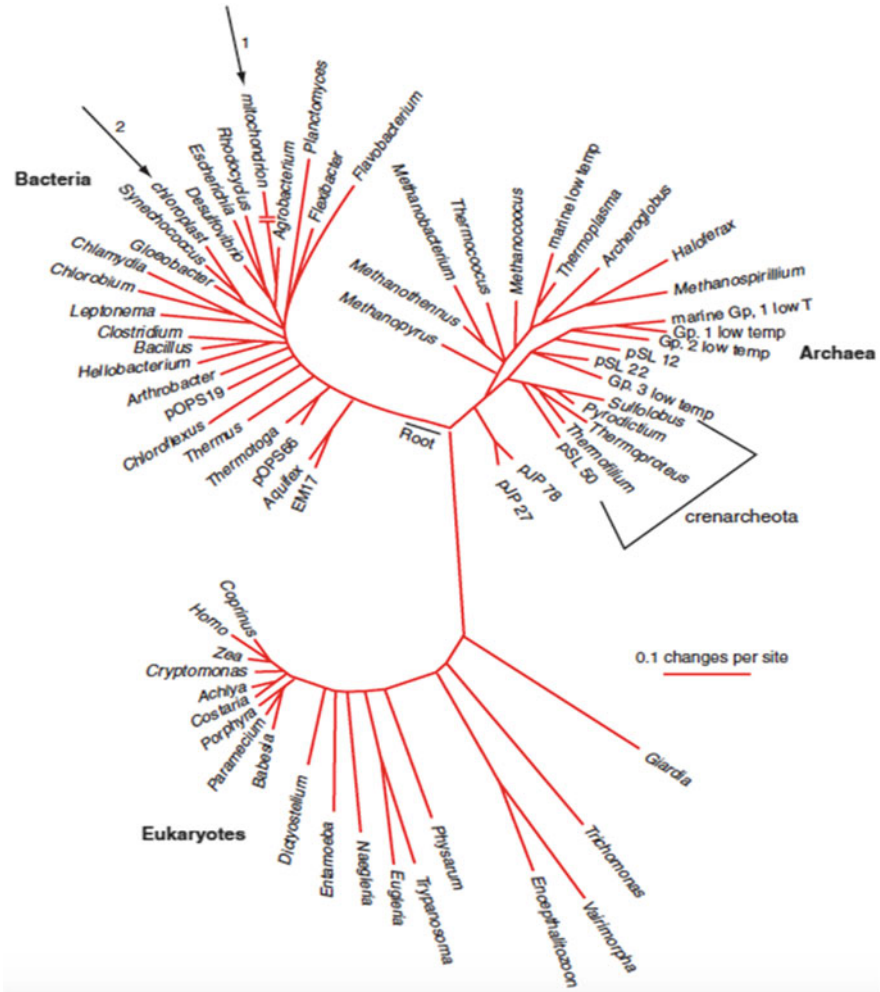


Fig. 22.14 16S rRNA tree of life. This is a rooted tree of life made by analyses of 16 s rRNA gene. It has three major branches—bacteria, archaea and eukaryotes. This is a phylogram (scale 0.1 changes per site) (explained in the next section) showing hypothetically how life originated 3.8 bya from primordial soup and diverged into various life forms. (Adapted from Pevsner 2009)

close to impossible to construct a true tree of life; rather, we construct ‘inferred trees’ which are based on the mutation in biomolecules or available data, showing hypothesized phylogenetic relationships.

The most popular tree of life has been constructed based on the phylogenetic analyses of 16 s rRNA gene (molecular marker). This tree has three branches showing the major divergence—bacteria, archaea and eukarya (Fig. 22.14).

A tree may be rooted showing the common or may be unrooted (Fig. 22.15) without the common ancestor. Often the data for the ancestors is not available that is

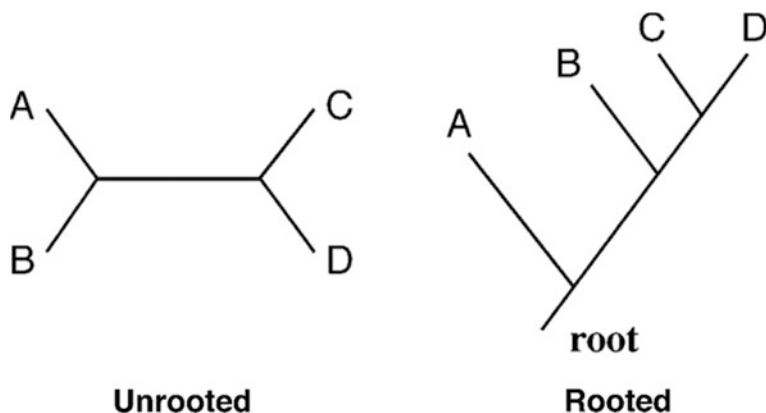


Fig. 22.15 A comparison between unrooted and rooted tree. The unrooted tree does not have any common ancestor. On the other hand, the rooted tree always shows the divergence from the common ancestor

when the unrooted tree is constructed which just shows the phylogenetic relationship among the organisms. A phylogenetic tree is not always constructed to observe the phylogeny of various species; it can also be constructed to chart the evolutionary path of the individual gene. Such a study is known as gene phylogeny. The evolutionary path of the gene might not overlap with the speciation events. A phylogenetic tree of species results from the evolution of genome (total genetic make-up of the organism).

22.8.2 Types of Tree Representation

The topology of the tree can be defined in two ways: cladogram and phylogram. A cladogram (Fig. 22.16) is a basic representative tree. It is a relative tree based on the order of phylogenetic events. The branch is unscaled and of the same length. Phylogram (Fig. 22.16) on the other hand has scaled branches. The branch represents the amount of evolution that has taken place since the time of divergence from the ancestor.

22.8.2.1 Clade

Clade is a group which includes the ancestor and the descendants. The tree also exhibits different type of branching pattern or the types of clade formed. There are three types of clade—monophyletic, paraphyletic and polyphyletic. Monophyletic clade includes the recent common and all its descendants. The paraphyletic clade excludes a few of the descendants, and polyphyletic clade includes distantly related species (OTU) (Fig. 22.17).

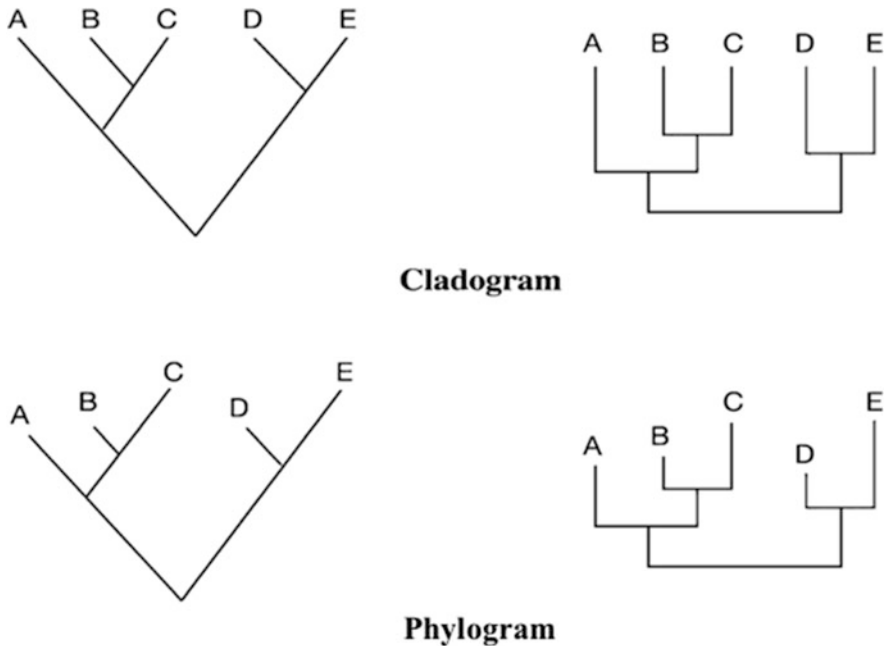


Fig. 22.16 Types of tree representation. Cladogram which has unscaled branches and phylogram which has scaled branches showing amount of evolution that has taken place. (Adapted from Jin Xiong 2006)

22.8.3 Procedure for Tree Construction

The phylogenetic tree construction involves the following steps (Fig. 22.18):

1. Choice of the molecular marker and collecting data.
2. Alignment of the data.
3. Choice of evolutionary model.
4. Constructing the phylogenetic tree.
5. Testing the reliability of the tree.

Let us look into the details of each of these steps:

1. Choice of the molecular marker and assembling data: Molecular marker is the biomolecule whose sequence would be taken into consideration to study the evolution. It may be nucleotide or protein sequence. The correct choice of molecular marker is an important step as it helps in construction of a true tree. If we are working with closely related organisms, then the rapidly evolving nucleotides should be the choice. For studying the evolution of slightly divergent organisms, the relatively conserved rRNA gene should be used. For more divergent organisms, protein sequences are used as they are relatively more conserved

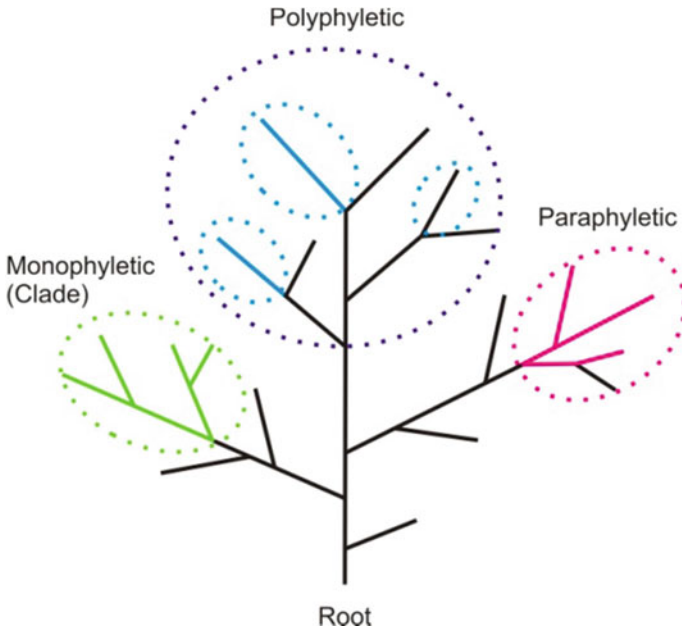


Fig. 22.17 The three types of clades. Monophyletic (green), polyphyletic (blue) and paraphyletic (pink). (Adapted from Karen Dowell 2008)

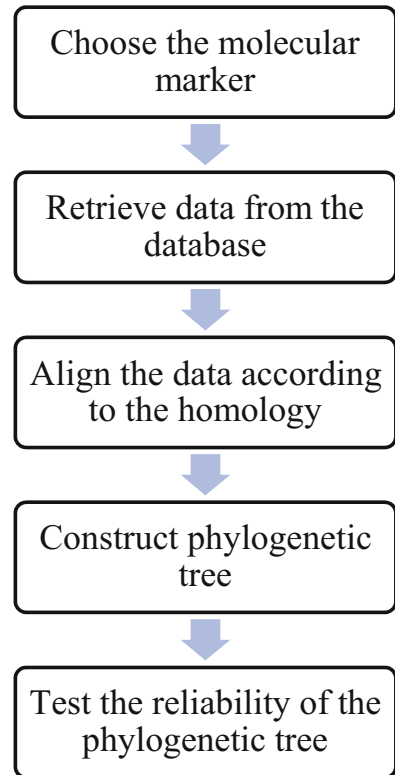
due to degeneracy of genetic code. DNA sequences are also biased than protein due to preferential usage of codon in some organisms. The protein has 20 amino acids as against only 4 bases of nucleotides and thus can be used for sensitive alignment. Globins are popularly used as molecular marker and were one of the first proteins to be sequenced. They are also used as molecular clocks (concept explained in earlier section).

Molecular marker can be used to spot positive and negative selection. For this it is important to distinguish between synonymous (results in no change in amino acid sequences) and non-synonymous substitution (results in change in amino acid). If non-synonymous substitution is higher, then it means a part of protein is undergoing evolution to bring about change in function of protein.

Once the molecular marker is chosen, the next step is to assemble the data of the organisms. For this there are several databases available from which the data can be extracted. For DNA the databases are DNA Data Bank of Japan (DDBJ), GenBank, etc. and for protein are SWISSPROT, etc. There are online tools like BLAST which can carry out the search and extract the data from databases.

2. Alignment of the data: Once the data is collected, the next step is to align these sequences according to homology in these sequences. The homology describes the phylogenetic relationships. There are two types of homologs—orthologs and paralogs. Orthologs are genes which have same ancestor but have diverged due to speciation event. Paralogs are duplicated genes of the same ancestor. Sequence

Fig. 22.18 Brief overview of the steps involved in phylogenetic tree construction. The flowchart enlists the various steps sequentially described in the text to construct a phylogenetic tree



alignment helps in the identification of homologous region and thus defines the evolutionary path. The multiple sequence alignment can be done by several tools—CLUSTAL, MSA, T-Coffee, etc.

3. Choice of evolutionary/substitution model: Substitution models are statistical methods to analyse the amount of evolution taking place. Several models are available for scoring the nucleotide substitution. One of the simplest models is Jukes-Cantor model which assumes that each nucleotide is replaced with equal probability. The other slightly more complex model is Kimura 2 parameter which differentiates between transition (mutation from purine to purine or pyrimidine to pyrimidine) and transversion (mutation from purine to pyrimidine and vice versa). In accordance with this model, transition occurs much more frequently than transversion, which is logical. For amino acid substitution, there are models like PAM. However these models assume that all positions in sequence have equal mutation rates. But this is not the case. For example, the wobble position of the codon mutates at a faster rate than others.
4. Construction of phylogenetic tree: There are two basic methods for tree construction: character based and distance based. The character-based method considers the molecular sequence as character, and after alignment each of these characters shares homology. It is also assumed that each of these characters evolves

independently and thus is considered separate evolutionary units. The character-based methods are maximum parsimony and maximum likelihood. On the other hand, the distance-based method calculates the dissimilarity between the sequences during sequence alignment and converts it into a matrix based on which phylogenetic tree of the organisms can be constructed. The branch lengths are additive, i.e. the evolutionary distance between two organisms can be obtained by adding the length of all the branches connecting them. The commonly used distance-based methods are unweighted pair group method using arithmetic average (UPGMA) and neighbour-joining method.

5. Testing the reliability of the tree: The last step of the phylogenetic tree construction involves statistically analysing the reliability of the inferred tree. This can be done through a statistical analysis model, bootstrapping. It tests for the sampling errors of the phylogenetic tree. This is done by repeated sampling of the datasets by introducing slight changes in them. If there is error in alignment, it will result in construction of biased tree. Random fluctuation in dataset will result in formation of altered tree. However if the alignment is correct, this random fluctuation in datasets will produce the same tree showing statistical confidence in the tree. Bootstrapping is of two types—parametric and non-parametric bootstrapping. When the changes introduced in the datasets are random, it is known as non-parametric bootstrapping. However when the new datasets are generated on the basis of particular sequential changes, it is known as parametric bootstrapping. All phylogenetic trees constructed after bootstrapping are summarized into a single consensus tree with each node of the branching pattern displaying the bootstrap value. This value evaluates the confidence level of the statistical analyses (Fig. 22.19) (Xiong 2006).

Box 22.1: Scientific Concept

Eweleit, L., Reinhold, K., Sauer, J.: Speciation Progress: A Case Study on the Bushcricket *Poecilimon veluchianus*. PLoS One 10(10): e0139494. <https://doi.org/10.1371/journal.pone.0139494> (2015).

A speciation study was carried out on subspecies of the flightless bushcricket *Poecilimon veluchianus*, endemic to Central Greece. These are *P. v. veluchianus* and *P. v. minor* and are differentiated on body size, timing of male signalling and sperm transfer rate. The two subspecies are parapatrically distributed in a V-shaped zone in Central Greece. The Iti mountain could be a geographical barrier to gene flow. Also, as speciation is a long process that could span over centuries, it can be considered that fragmentations occurring previously could be barriers to gene flow. Laboratory experiments done earlier suggest that females do not differentiate between songs of the two subspecies and there is presence of partial post-zygotic isolation that diminishes fertility

(continued)

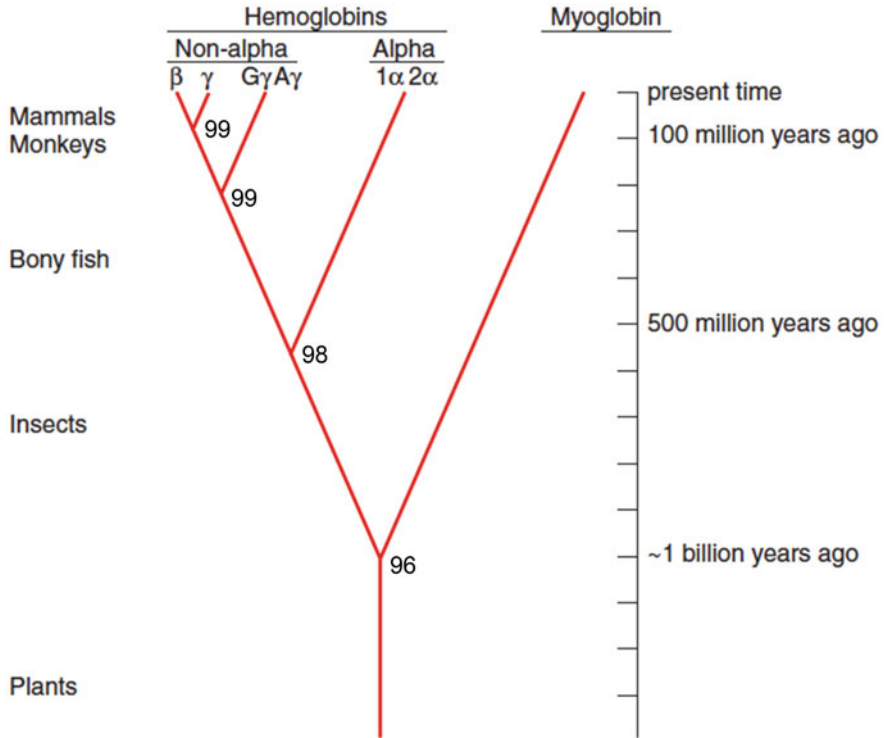


Fig. 22.19 A representative dendrogram. It is a representation of tree where branches have a scale showing evolutionary time showing evolution of globin family of genes (bootstrap values are shown at branching points)

Box 22.1 (continued)

of F1 by reduction in amount of sperm transfer. Also, F1 female hybrids were mostly fertile compared to males with lower sperm count. Hence, there is evidence for presence of 'pre-mating barriers vs post-mating barriers' and it indicates that speciation is an ongoing process.

Genetic differences between the subspecies were evaluated in this study using the sequences of mitochondrial control region (CR marker) and internal transcribed spacers (ITS marker 1 and 2). As mentioned earlier, due to absence of pre-mating isolating mechanisms, hybridization is possible in the area of contact zone that would encourage gene flow in comparison to the distantly located sites. The occurrence of shared haplotypes was investigated in the contact zone vs the distant sites to predict the site of hybridization. Single site substitution differences were found in between various haplotypes. Also to

(continued)

Box 22.1 (continued)

determine if geographic isolation has shaped the population structure, a distance-based redundancy analysis (dbDA) was done.

The results demonstrated that the aligned sequences of the CR dataset contained 794 bp and the ITS dataset consisted of 706 bp and both showed high levels of genetic variation especially due to low number of variable sites. A characteristic of *Poecilimon* species is a high number of exclusive haplotypes occurring at one of the different sampling sites in the mitochondrial CR marker. In contrast there was lesser diversity in the nuclear ITS marker that displayed a lower genetic variation.

The genetic analysis done based on the ITS marker revealed one main barrier to gene flow, hence indicating incomplete reproductive isolation. The contact zone has been proposed to extend from north-east of Central Greece to the south-west. The CR marker in contrast does not clearly support the speciation with formation of two subspecies, restricted gene flow and a clear contact zone.

No influence of sex was observed on the genetic pattern for *P. veluchianus* on investigation with dBRDA. But on testing for isolation by distance (IBD) for both sexes, 19% variability was found in females versus 10% in males. Hence, it can be concluded that IBD has a stronger impact on females. Males produce sounds prior to mating, waiting for the females, and this result indicates that females need to walk around to locate the correct partner.

Individuals of both subspecies are tougher to distinguish in the field as they are phenotypically similar with the exception of body size. This feature depends on the size of the mother, and hence the hybrids' body size depends on the identity of the mothers' subspecies. Laboratory experiments have revealed that hybrids with a *P. v. veluchianus* mother grow bigger than pure *P. v. minor* individuals and hybrids with a *P. v. minor* mother stay smaller than pure *P. v. veluchianus* individuals.

Speciation is in progress for these subspecies, as there is a lack of a strong prezygotic isolation barrier between these two parapatrically distributed subspecies. It could also be predicted that the species experienced a bottleneck and are now in a phase of range expansion. IBD and sexual selection have shown to have a great influence on the population structure, and it has been hypothesized that *P. veluchianus* may be a case of widely distributed ring species. To investigate this further, closely related species of *Poecilimon* are necessary. Both subspecies are also distributed in various altitudes with *P. v. veluchianus* occurring above 380 m and *P. v. minor* occurring also below this altitude, and the lack of a strong prezygotic barrier probably supports a secondary contact zone. It is suggested that a secondary contact after an allopatric phase is likely for the two subspecies. The missing pre-mating barrier suggests a rather weak selection against hybrids and might also indicate speciation in progress. Further study using microsatellite data as well as AFLP could shed light on this ongoing speciation process.

22.9 Summary

- Evolutionary genetics is the modern field of study which integrates genetics with the Darwinian view of evolution. It attempts to account for any change in nature in terms of allele, genes and genotypes and how the variations at population level can bring permanent variations in the species leading from microevolutionary to macroevolutionary changes.
- Mutation, natural selection, genetic drift and migration are the microevolutionary changes. Mutation is the most important variation which acts as the raw material for the evolution of the gene pool. Most of the mutations are neutral, but some are positive which might improve the fitness of the organism in its environment resulting in adaptive evolution. The other variations may occur at the chromosomal level through recombination and aberrations.
- The Hardy-Weinberg law was introduced to understand how these variations affect the allele and genotype frequencies in a population. However, Hardy-Weinberg law functions under ideal conditions in an infinite population in absence of evolutionary forces. But in a real, finite population, the evolutionary forces like natural selection, genetic drift and migration act to affect the variation in gene pool every generation.
- Natural selection is the force which results in the adaptation of the fittest in the environment. It is directional in nature and acts on a large population. Selection can occur at individual level, population level or sexual level.
- On the contrary genetic drift is non-adaptive and results in random fixation of allele in a small population. Migration or gene flow also affects the variation in a gene pool. Magnification of these variations over a long period of time will lead to speciation.
- Earlier evolution was studied through fossil record. However fossil records were often incomplete leaving a number of question marks. With the onset of new technology and development of molecular biology techniques, the field of molecular phylogeny gained momentum. Using biomolecules and genes as markers, phylogenetic tree can be constructed which gives a bird-eye's view of the phylogenetic relationship among organisms.
- Human evolution studies have also been carried out using the mitochondrial DNA which has helped chart out the divergence of **Homo neanderthalensis** and **Homo sapiens**.

References

- Abzhanov A (2010) Darwin's Galapagos finches in modern biology. *Philos Trans R Soc Lond B Biol Sci* 365(1543):1001–1007
- Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ (2004) Bmp4 and morphological variation of beaks in Darwin's finches. *Science* 305(5689):1462–1465
- Agar W, Drummond F, Tiegs O, Gunson M (1954) Fourth (final) report on a test of McDougall's Lamarckian experiment on the training of rats. *J Exp Biol* 31(3):307–321

- Alcock J (2005) *Animal behaviour: an evolutionary approach*, 8th edn. Sinauer Associates, Sunderland, MA
- Anstee DJ (2010) The relationship between blood groups and disease. *Blood* 115:4635–4643
- Bluestone CD (2009) Galapagos: Darwin, evolution and ENT. *Laryngoscope* 119(10):1902–1905
- Bowman RI (1961) Morphological differentiation & adaptation in Galapagos finches. *Univ Calif Publ Zool* 58:1–302
- Brown TA (2002) *Molecular phylogenetics. genomes*, 2nd edn. Wiley-Liss, U.K, Oxford
- Byers JA, Waits L (2006) Good genes sexual selection in nature. *PNAS* 103(44):16343–16345
- Campbell NA, Reece JB, Urry LA, Cain ML, Wasserman SA, Minorsky PV, Jackson RB (2008) *Biology*, 8th edn. Pearson Education Inc., San Francisco, CA
- Cook SA, Johnson MP (1968) Adaptation to heterogeneous environments. I. Variation in heterophylly in *Ranunculus flammula*. *Evolution* 22:496–516
- Cook LM, Saccheri IJ (2013) The peppered moth and industrial melanism: evolution of a natural selection case study. *Heredity* 110(3):207–212
- Crew F (1936) A repetition of McDougall's Lamarckian experiment. *J Genet* 33(1):61–102
- Denis N (2011) Neo-Darwinism, the modern synthesis and selfish genes: are they of use in physiology? *J Physiol* 589(5):1007–1015
- Dowell K (2008) *Molecular Phylogenetics: an introduction to computational methods and tools for analysing evolutionary relationships*. http://www.math.umaine.edu/~khalil/courses/MAT500/papers/MAT500_Paper_Phylogenetics.pdf
- Drew G (1939) McDougall's experiment of the inheritance of acquired habits. *Nature* 143:188–191
- Duret L (2008) Neutral theory: the null hypothesis of molecular evolution. *Nat Educ* 1(1):218
- Emerling CA (2016) Will evolution doom the cheetah? *Understanding Evolution* https://evolution.berkeley.edu/evolibrary/news/160201_cheetahs
- Fitzpatrick BM, Fordyce JA, Gavrilets S (2009) Pattern, process & geographic modes of speciation. *J Evol Biol* 22:2342–2347
- Forsdyke DR (2007) Positive Darwinian selection: does the comparative method rule? *JBS* 15:95–108
- Futuyma DJ (1998) *Evolution*, 3rd edn. Oxford University Press, Boston, MA
- Hall BK, Hallgrímsson B (2008) *Strickberger's evolution*, 4th edn. Jones & Bartlett Publishers, LLC, Burlington, MA
- Hartl DL, Jones EW (2009) *Genetics: analysis of genes & genomes*, 7th edn. Jones & Bartlett Publishers, LLC, Burlington, MA
- Jankowska D, Milewski R, Górska U, Milewska AJ (2011) Application of Hardy-Weinberg law in biomedical research. *Stud InLog Gramm Rhetor* 25(38)
- Karn MN, Penrose LS (1952) Birth rate and gestation time in relation to maternal age, parity and infant survival. *Ann Eugenics* 16:147–164
- Keller B, Vos JM, Schmidt-Lebuhn AM, Thomson JD, Conti E (2016) Both morph- and species-dependent asymmetries affect reproductive barriers between heterostylous species. *Ecol Evol* 6(17):6223–6244
- Kimura M (1991) The neutral theory of molecular evolution: a review of recent evidence. *Jpn J Genet* 66(4):367–386
- Klug WS, Cummings MR, Spencer CA, Palladino MA (2012) *Concepts of genetics*, 10th edn. Pearson Education Inc., San Francisco, CA
- Koonin EV, Wolf YI (2009) Is evolution Darwinian or/and Lamarckian? *Biol Direct* 4(42):1–14
- Kramer J, Meunier J (2016) Kin and multilevel selection in social evolution: a never-ending controversy? *F1000 Res* 5:Faculty Rev-776. <https://doi.org/10.12688/f1000research.8018.1>
- Łukasiewicz A, Szubert-Kruszyńska A, Radwan J (2017) Kin selection promotes female productivity and cooperation between the sexes. *Evol Biol* 3(3):e1602262
- Marshall JH (2002) On the changing means of mutation. *Hum Mutat* 19:76–78
- McDougall W (1938) Fourth report on Lamarckian experiment. *Br J Psychol* 28(3):321–325
- Mourant AE, Kopeć AC, Domaniewska-Sobczak K (1976) *The distribution of the human blood groups and other polymorphisms*, 2nd edn. Oxford University Press, London

- Noor MAF, Feder JLF (2006) Speciation genetics: evolving approaches. *Nat Rev Genet* 7:851–861
- Orr HA (2009) Fitness and its role in evolutionary genetics. *Nat Rev Genet* 10(8):531–539
- Parent CE, Cacccone A, Petren K (2008) Colonization and diversification of Galapagos terrestrial fauna: a phylogenetic & biogeographical synthesis. *Philo Trans R Soc Lond B Biol Sci* 363(1508):3347–3361
- Peltonen L (2001) Founder effect. <https://www.sciencedirect.com/sdfe/pdf/download/eid/3-s2.0-B0122270800004742/first-page-pdf>
- Pevsner J (2009) *Bioinformatics and functional genomics*, 2nd edn. John Wiley & Sons, Hoboken, NJ
- Pierce BA (2012) *Genetics: a conceptual approach*, 4th edn. W.H. Freeman & Company, New York
- Portin P (2007) Evolution of man in the light of molecular genetics: a review. Part I. our evolutionary history and genomics. *Hereditas* 144(3):80–95
- Raymond MM, O'Brien SJ (1993) Dating the genetic bottleneck of the African cheetah. *PNAS* 90(8):3172–3176
- Ridley M (2004) *Evolution*, 3rd edn. Blackwell Publishing, Malden, MA
- Roesti M, Hendry AP, Salzburger W, Berner D (2012) Genome divergence during evolutionary diversification as revealed in replicate lake–stream stickleback population pairs. *Mol Ecol* 21: 2852–2862
- Sato A, O'hUigin C, Figueroa F, Grant PR, Grant R, Tichy H, Klein J (1999) Phylogeny of Darwin's finches as revealed by mtDNA sequences. *PNAS* 96(9):5101–5106
- Schliwen UK, Tautz D, Pääbo S (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* 368:629–632
- Snustad DP, Simmons MJ (2012) *Principles of genetics*, 6th edn. John Wiley & Sons, Inc., Hoboken, NJ
- Turissini DA, McGirr JA, Patel SS, David JR, Matute DR (2018) The rate of evolution of post mating-prezygotic reproductive isolation in drosophila. *Mol Biol Evol* 35(2):312–334
- Wade M Evolutionary genetics. In: Zalta EN (ed) *The Stanford encyclopedia of philosophy* (Fall 2008 edition). <https://plato.stanford.edu/archives/fall2008/entries/evolutionary-genetics/>
- Weismann A (1891) A supposed transmission of mutilations. In: *Essays upon heredity and kindred biological problems*. Oxford University Press, Oxford
- Weismann A (1893) *The germplasm: a theory of heredity*. Charles Scribner's and Sons, New York
- Wolf JBW, Lindell J, Backstrom N (2010) Introduction speciation genetics: current status and evolving approaches. *Phil Trans R Soc B* 365:1717–1733
- Wu DD, Zhang Y (2008) Positive Darwinian selection in human population: a review. *Chin Sci Bull* 53(10):1457–1467
- Xiong J (2006) *Essential bioinformatics*, 1st edn. Cambridge University Press, New York
- Yi S (2013) Neutrality and molecular clocks. *Nat Educ Knowl* 4(2):3
- Zhang J (2010) Positive darwinian selection in gene evolution. <https://pdfs.semanticscholar.org/6c91/7d8d705af18a3d0e5e139076b5335f046f6c.pdf>
- Zhang J, Zhang YP, Rosenberg HF (2002) Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf eating monkey. *Nat Genet* 30:411–415
- Zykova TY, Levitsky VG, Belyaeva ES, Zhimulev IF (2018) Polytene chromosomes—a portrait of functional organization of the *drosophila* genome. *Curr Genomics* 19(3):179–191