18.4D: Gene Duplications and Divergence

Learning Objectives

- Explain the mechanisms of gene duplication and divergence

Gene Duplication

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

Ectopic Recombination

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

Replication Slippage

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

Retrotransposition

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

Aneuploidy

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

Gene duplication as an evolutionary event

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

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

Another possible fate for duplicate genes is that both copies are equally free to accumulate degenerative mutations, so long as any defects are complemented by the other copy. This leads to a neutral “subfunctionalization” model, in which the functionality of the original gene is distributed among the two copies. Neither gene can be lost, as both now perform important non-redundant functions, but ultimately neither is able to achieve novel functionality. Subfunctionalization can occur through neutral processes in which mutations accumulate with no detrimental or beneficial effects. However, in some cases subfunctionalization can occur with clear adaptive benefits. If an ancestral gene is pleiotropic and performs two functions, often times neither one of these two functions can be changed without affecting the other function. In this way, partitioning the ancestral functions into two separate genes can allow for adaptive specialization of subfunctions, thereby providing an adaptive benefit.

Divergence

Genetic divergence is the process in which two or more populations of an ancestral species accumulate independent genetic changes through time, often after the populations have become reproductively isolated for some period of time. In some cases, subpopulations living in ecologically distinct peripheral environments can exhibit genetic divergence from the remainder of a population, especially where the range of a population is very large. The genetic differences among divergent populations can involve silent mutations (that have no effect on the phenotype) or give rise to significant morphological and/or physiological changes. Genetic divergence will always accompany reproductive isolation, either due to novel adaptations via selection and/or due to genetic drift, and is the principal mechanism underlying speciation.
Genetic drift or allelic drift is the change in the frequency of a gene variant (allele) in a population due to random sampling. The alleles in the offspring are a sample of those in the parents, and chance has a role in determining whether a given individual survives and reproduces. A population’s allele frequency is the fraction of the copies of one gene that share a particular form. Genetic drift may cause gene variants to disappear completely and thereby reduce genetic variation. When there are few copies of an allele, the effect of genetic drift is larger, and when there are many copies the effect is smaller. These changes in gene frequency can contribute to divergence.

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

Key Points

- Ectopic recombination occurs when there is an unequal crossing-over and the product of this recombination are a duplication at the site of the exchange and a reciprocal deletion.
- Gene duplications do not always result in detrimental mutations; they can contribute to divergent evolution, which causes genetic differences between groups to develop and eventually form new species.
- Replication slippage can occur when there is an error during DNA replication and duplications of short genetic sequences are produced.
- Retrotranspositions occur when a retrovirus copies their genome by reverse transcribing RNA to DNA and aberrantly attach to cellular mRNA and reverse transcribe copies of genes to create retrogenes.
- Aneuploidy can occur when there is a nondisjunction even at a single chromosome thus, the result is an abnormal number of chromosomes.
- Genetic divergence can occur by mechanisms such as genetic drift which contribute to the accumulation of independent genetic changes of two or more populations derived from a common ancestor.

Key Terms

- **paralogous**: having a similar structure indicating divergence from a common ancestral gene
- **nondisjunction**: the failure of chromosome pairs to separate properly during meiosis
- **retrogene**: a DNA gene copied back from RNA by reverse transcription
- **genetic drift**: an overall shift of allele distribution in an isolated population, due to random fluctuations in the frequencies of individual alleles of the genes