4.2: Origins of Mutations

Types of mutations

Mutations may involve the loss (deletion), gain (insertion) of one or more base pairs, or else the substitution of one or more base pairs with another DNA sequence of equal length. These changes in DNA sequence can arise in many ways, some of which are spontaneous and due to natural processes, while others are induced by humans intentionally (or unintentionally) using mutagens. There are many ways to classify mutagens, which are the agents or processes that cause mutation or increase the frequency of mutations. We will classify mutagens here as being (1) biological, (2) chemical, or (3) physical.

Mutations of biological origin

A major source of spontaneous mutation is errors that arise during DNA replication. DNA polymerases are usually very accurate in adding a base to the growing strand that is the exact complement of the base on the template strand. However, occasionally, an incorrect base is inserted. Usually, the machinery of DNA replication will recognize and repair mispaired bases, but nevertheless, some errors become permanently incorporated in a daughter strand, and so become mutations that will be inherited by the cell’s descendents (Figure \(\PageIndex{1}\)).
Figure \(\PageIndex{1}\): Mispairing of bases (e.g. G with T) can occur due to tautomerism, alkylating agents, or other effects. As a result, in this example the AT base pair in the original DNA strand will become permanently substituted by a GC based pair in some progeny. The mispaired GT basepair will likely be repaired or eliminated before further rounds of replication. (Original-Deyholos-CC:AN)

Another type of error introduced during replication is caused by a rare, temporary misalignment of a few bases between the template strand and daughter strand (Figure \(\PageIndex{2}\)). This strand-slippage causes one or more bases on either strand to be temporarily displaced in a loop that is not paired with the opposite strand. If this loop forms on the template strand, the bases in the loop may not be replicated, and a deletion will be introduced in the growing daughter strand. Conversely, if a region of the daughter strand that has just been replicated becomes displaced in a loop, this region may be replicated again, leading to an insertion of additional sequence in the daughter strand, as compared to the template strand.

Figure \(\PageIndex{2}\): Strand-slippage can occur occasionally during replication, especially in regions with short, repeated sequences. This can lead to either deletion (left) or insertion (right) of sequences compared to the products of normal replication (center), depending on whether the template strand or daughter strand “loops-out” during replication. (Original-Deyholos-CC:AN)

**Consequences:** Regions of DNA that have several repeats of the same few nucleotides in a row are especially prone to this type of error during replication. Thus regions with short-sequence repeats (SSRs) are tend to be highly polymorphic, and are therefore particularly useful in genetics. They are called microsatellites.

Mutations can also be caused by the insertion of viruses, transposable elements (transposons), see below, and other types of DNA that are naturally added at more or less random positions in chromosomes. The insertion may disrupt the coding or regulatory sequence of a gene, including the fusion of part of one gene with another. These insertions can occur spontaneously, or they may also be intentionally stimulated in the laboratory as a method of mutagenesis called
Transposon-tagging. For example, a type of transposable element called a *P element* is widely used in Drosophila as a biological mutagen. *T-DNA*, which is an insertional element modified from a bacterial pathogen, is used as a mutagen in some plant species.

### Mutations due to Transposable Elements

Transposable elements (TEs) are also known as mobile genetic elements, or more informally as *jumping genes*. They are present throughout the chromosomes of almost all organisms. These DNA sequences have a unique ability to be cut or copied from their original location and inserted into new locations in the genome. This is called transposition. These insert locations are not entirely random, but TEs can, in principle, be inserted into almost any region of the genome. TEs can therefore insert into genes, disrupting its function and causing a mutation. Researchers have developed methods of artificially increasing the rate of transposition, which makes some TEs a useful type of mutagen. However, the biological importance of TEs extends far beyond their use in mutant screening. TEs are also important causes of disease and phenotypic instability, and they are a major mutational force in evolution.

There are two major classes of TEs in eukaryotes (Figure 1).

- **Class I** elements include retrotransposons; these transpose by means of an RNA intermediate. The TE transcript is reverse transcribed into DNA before being inserted elsewhere in the genome through the action of enzymes such as integrase.
- **Class II** elements are known also as transposons. They do not use reverse transcriptase or an RNA intermediate for transposition. Instead, they use an enzyme called transposase to cut DNA from the original location and then this excised dsDNA fragment is inserted into a new location. Note that the name transposon is sometimes used incorrectly to refer to any type of TEs, but in this book we use transposon to refer specifically to Class II elements.

![Diagrams of the two main types of transposable elements](https://bio.libretexts.org/Bookshelves/Genetics/Book%3A_Online_Open_Genetics_(Nickle_and_Barrette-Ng)/04%3A_Mutations_and_Mutation_Rates/4.07%3A_Retrotransposons_and_Transposons_and_Transposition)

Figure 1: Diagrams of the two main types of transposable elements. (TEs) Class I elements transpose via an ssRNA intermediate, which is reverse transcribed to dsDNA prior to insertion of this copy in a new site in the genome. Class II elements do not involve an RNA intermediate; most Class II elements are cut from their original location as dsDNA, prior to being inserted into a new site in the genome. Although the diagram shows TEs being inserted on the same chromosome as they originated from, TEs can also move to other chromosomes within the same cell.

TEs are relatively short DNA sequences (100-10,000 bp), and encode no more than a few proteins (if any). Normally, the protein-coding genes within a TE are all related to the TE’s own transposition functions. These proteins may include reverse transcriptase, transposase, and integrase. However, some TEs (of either Class I or II) do not encode any proteins at all. These non-autonomous TEs can only transpose if they are supplied with enzymes produced by other, autonomous TEs located elsewhere in the genome. In all cases, enzymes for transposition recognize conserved nucleotide sequences within the TE, which dictate where the enzymes begin cutting or copying.
The human genome consists of nearly 45% TEs, the vast majority of which are families of Class I elements called **LINEs** and **SINEs**. The short, **Alu** type of SINE occurs in more than one million copies in the human genome (compare this to the approximately 21,000, non-TE, protein-coding genes in humans). Indeed, TEs make up a significant portion of the genomes of almost all eukaryotes. Class I elements, which usually transpose via an RNA **copy-and-paste** mechanism, tend to be more abundant than Class II elements, which mostly use a **cut-and-paste** mechanism. But even the cut-paste mechanism can lead to an increase in TE copy number. For example, if the site vacated by an excised transposon is repaired with a DNA template from a homologous chromosome that itself contains a copy of a transposon, then the total number of transposons in the genome will increase.

Besides greatly expanding the overall DNA content of genomes, TEs contribute to genome evolution in many other ways. As already mentioned, they may disrupt gene function by insertion into a gene’s coding region or regulatory region. More interestingly adjacent regions of chromosomal DNA are sometimes mistakenly transposed along with the TE; this can lead to gene duplication. The duplicated genes are then free to evolve independently, leading in some cases to the development of new functions. The breakage of strands by TE excision and integration can disrupt genes, and can lead to chromosome rearrangement or deletion if errors are made during strand rejoining. Furthermore, having so many similar TE sequences distributed throughout a chromosome sometimes allows mispairing of regions of homologous chromosomes at meiosis, which can cause unequal crossing-over, resulting in deletion or duplication of large segments of chromosomes. Thus, TEs are a potentially important evolutionary force, and may not be included as merely "junk DNA", as they once were.