STR Introduction

STR Typing

- Forensic DNA interpretation has been centered on the analysis of STRs (*short tandem repeats*), i.e. short DNA sequences that are repeated several times.

- These repeat patterns are located in areas called *loci* and vary among individuals.

- Variants for a given locus are called *alleles* and it is this variation (called polymorphism) that allows us to associate a particular DNA sample with an individual person.

Mutations

Mutations are the cause of the variation encountered in DNA and one of the reasons that STR loci render highly informative markers in forensic genetics. Most of the mutations are caused by an error during DNA replication (although other mechanisms and external influences can also lead to a change in DNA sequence).

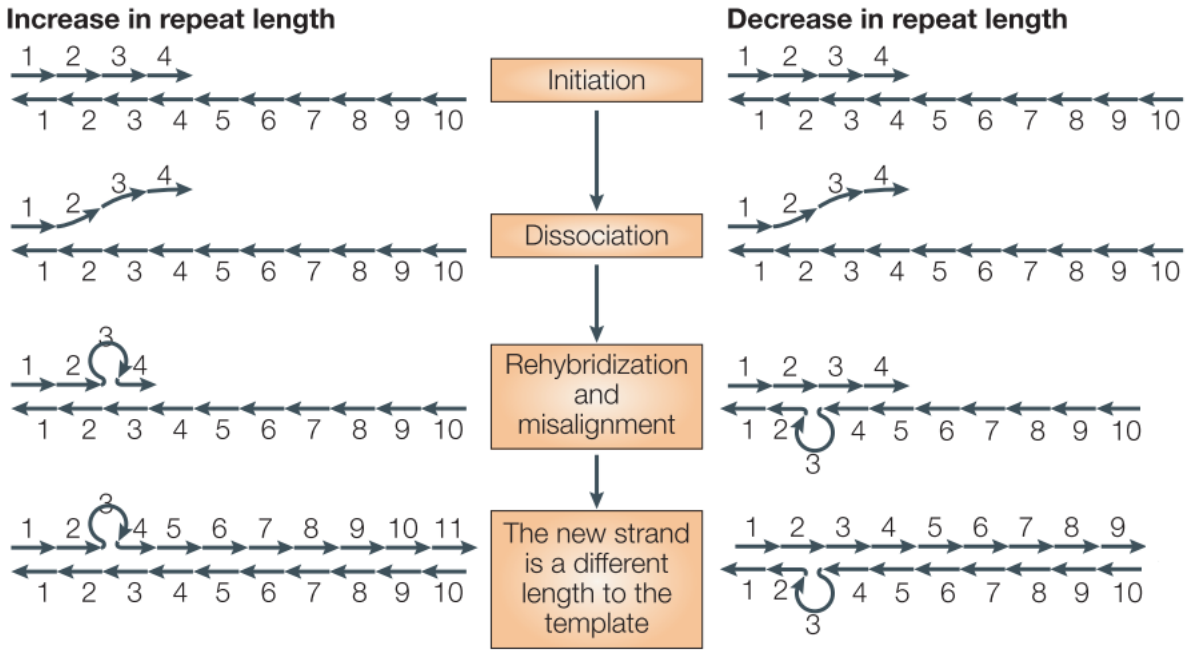
Examples of mutations:

- **Substitutions:** A point mutation where one base is substituted for another, such as a SNP.

- **Indels:** Small insertions/deletions due to the addition of one or more extra nucleotides into the DNA or the loss of a section of DNA.

Slipped Strand Mispairing

STR polymorphisms derive mainly from variability in length. A proposed mechanism for these genetic variations is the *slipped strand mispairing* (SSM) mechanism: the dissociation of replicating DNA strands followed by misaligned re-association.



Source: Microsatellites: simple sequences with complex evolution (Ellengren, 2004).

STR Classes

STR loci may be categorized in three different classes, based on how well alleles conform to the core repeat pattern:

- **Simple STRs**: only show variation in the number of repeats without additional sequence variation.

- **Compound STRs**: consist of several adjacent repeats of the same repeat unit length.

- **Complex STRs**: contain repeats of variable length as well as sequences.

Sequence Variation

STR loci may be categorized in three different classes, based on how well alleles conform to the core repeat pattern:

Class	Locus	Allele sequence
Simple	CSF1PO	$[TCTA]_8$
Simple	Penta D	$[AAAGA]_{12}$
Compound	vWA	$[TCTA][TCTG]_4[TCTA]_{13}$
Compound	D22S1045	$[ATT]_7ACT[ATT]_2$
Complex	FGA	$[TTTC]_3TTTTTT[CTTT]_{11}CTCC[TTCC]_2$
Complex	D1S1656	$[TAGA]_4TGA[TAGA]_{13}TAGG[TG]_5$

Same-length variants (i.e. *isoalleles*) have the same allele number, but differ at sequence level. This means that traditional DNA typing methods have less discriminatory capability than is potentially available via sequencing techniques.

Locus	Allele number	Allele sequence
D3S1358	15	$[TCTA][TCTG]_3[TCTA]_{11}$
D3S1358	15	$[TCTA][TCTG]_2[TCTA]_{12}$
D18S51	20	$[AGAA]_{20}$
D18S51	20	$[AGAA]_{16}GGAA[AGAA]_3$