

Variable Number Tandem Repeats

Contents

- [1 Polymorphisms](#)
- [2 Tandem Repeats](#)
- [3 CoDIS](#)
- [4 Crime Scene Investigation](#)
- [5 External Resources](#)

Polymorphisms

The difference in nucleotide sequences between humans lies between 0.1-0.4%. That means that people are greater than 99% similar. But when you look around the room at your classmates, you can see that that small difference amounts to quite a bit of variation within our species. The bulk of these differences aren't even within the coding sequences of genes, but lie outside in regulatory regions that change the expression of those genes. Imagine if there were mutations to the coding sequences, this could be very deleterious to the well-being of the organism. We say that the coding sequences of genes that ultimately lead to proteins has a **selective pressure** to remain the same. The areas outside of the coding sequences have a reduced and sometimes non-existent selection pressure. These areas are allowed to mutate in sequence and even expand or contract. Areas of changes or differences are called **polymorphic** (many forms). If you were to read a repetitive set of sequences and count the repetition, you'd make mistakes and lose count. Likewise, DNA polymerase will make errors or stutter in areas of repetitiveness and produce polymorphic regions.

Tandem Repeats

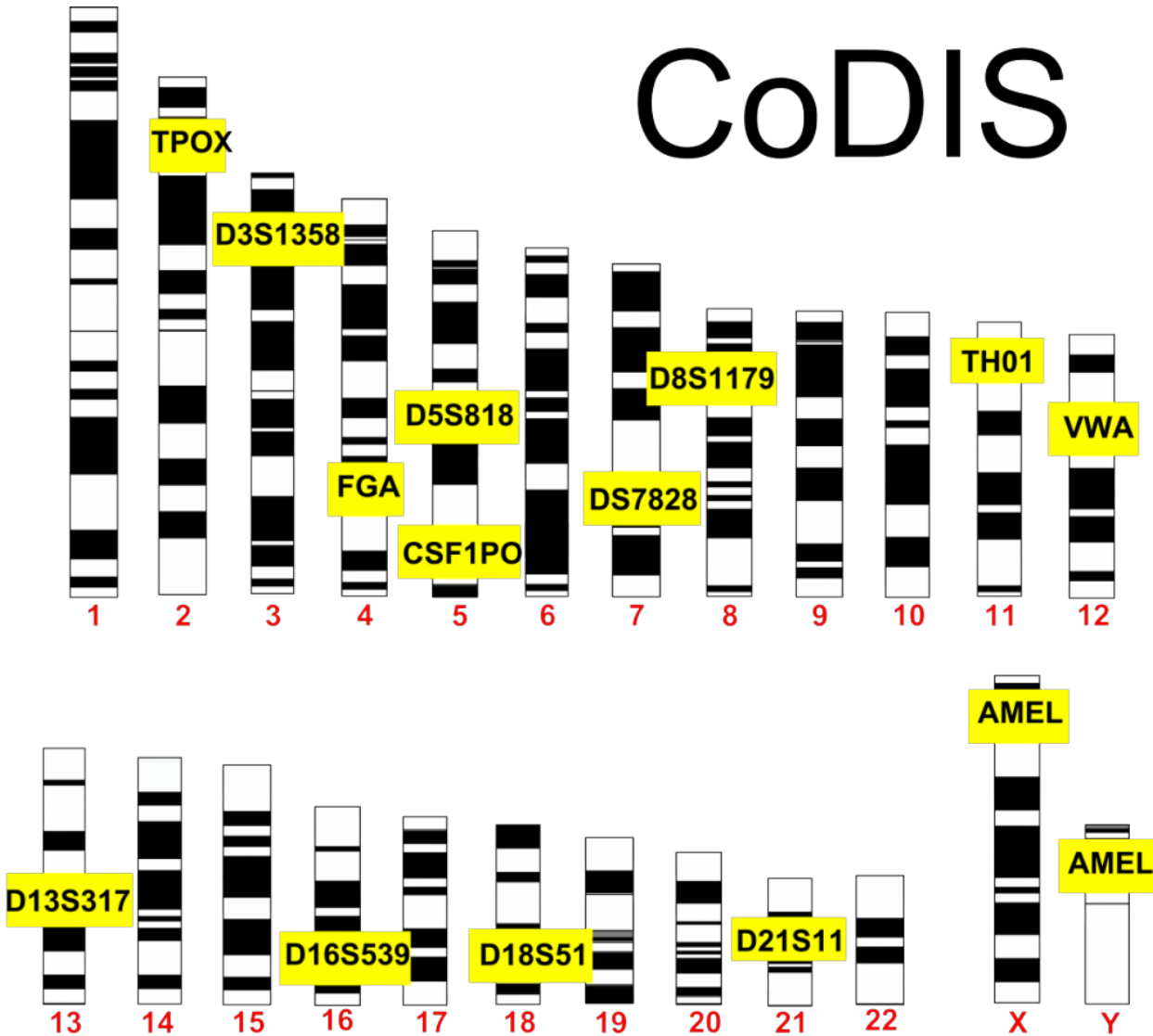
A type of polymorphism occurs due to these repeats expanding and contracting in non-coding regions. These regions are called variable number tandem repeats (**VNTRs**) or sometimes short tandem repeats (**STRs**). Any region or location on a chromosome is referred to as **locus** (loci for plural). Scientists use polymorphic loci that are known to contain VNTRs/STRs in order to differentiate people based on their DNA. This is often used in forensic science or in maternity/paternity cases. Any variation of a locus is referred to as an **allele**. In standard genetics, we often think of an allele as a variation of gene that would result in a difference in a physical manifestation of that gene. In the case of STRs, these alleles are simply a difference in number of repeats. That means the length of DNA within this locus is either longer or shorter and gives rise to many different alleles. VNTRs are referred to as minisatellites while STRs are called **microsatellites**.

CoDIS



The FBI and local law enforcement agencies have developed a database called the Combined DNA Index System (**CoDIS**) that gathers data on a number of STRs. By establishing the number of repeats of a given locus, law enforcement officials can differentiate individuals based on the repeat length of these alleles. CoDIS uses a set of 13 loci that are tested together. As you would imagine, people are bound to have the same alleles of certain loci, especially if they were related. The use of 13 different loci makes it statistically improbable that 2 different people could be confused for each other. Think about this in terms of physical traits. As you increase the number of physical traits used to describe someone, you are less likely to confuse that person with someone else based on those combinations of traits. Using the CoDIS loci increases the stringency since there are many alleles for each locus. The thirteenth locus in CoDIS (called AMEL) discriminates between male and female.

CoDIS



CoDIS STRs: The FBI utilizes 13 different loci to discriminate between people. AMEL discriminates by gender and is located on the X & Y.

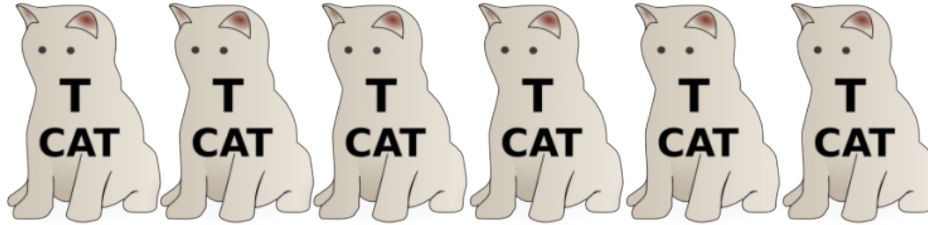
Crime Scene Investigation

This lab uses a CoDIS locus called TH01. TH01 is a locus on chromosome 11 that has a repeating sequence of TCAT. There are reported to be between 3-14 repeats in this locus. With the exception of X and Y in a male, all chromosomes have a homologous partner. Therefore, each individual will have 2 alleles for each CoDIS locus.

The TH01 locus contains repeats of TCAT.

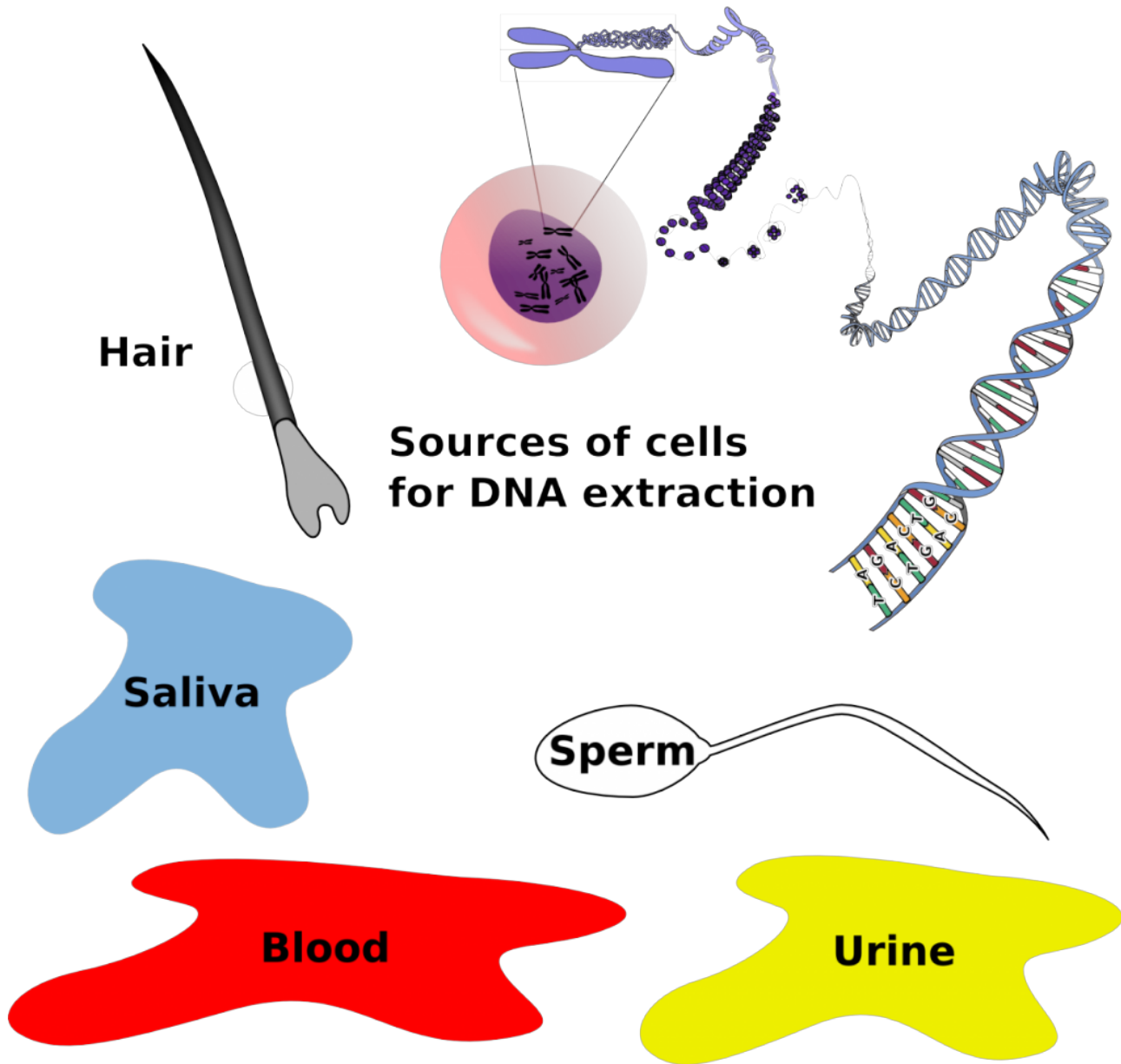
CCC **TCAT** **TCAT** **TCAT** **TCAT** **TCAT** **TCAT** AAA

This example has 6 TCAT repeats.



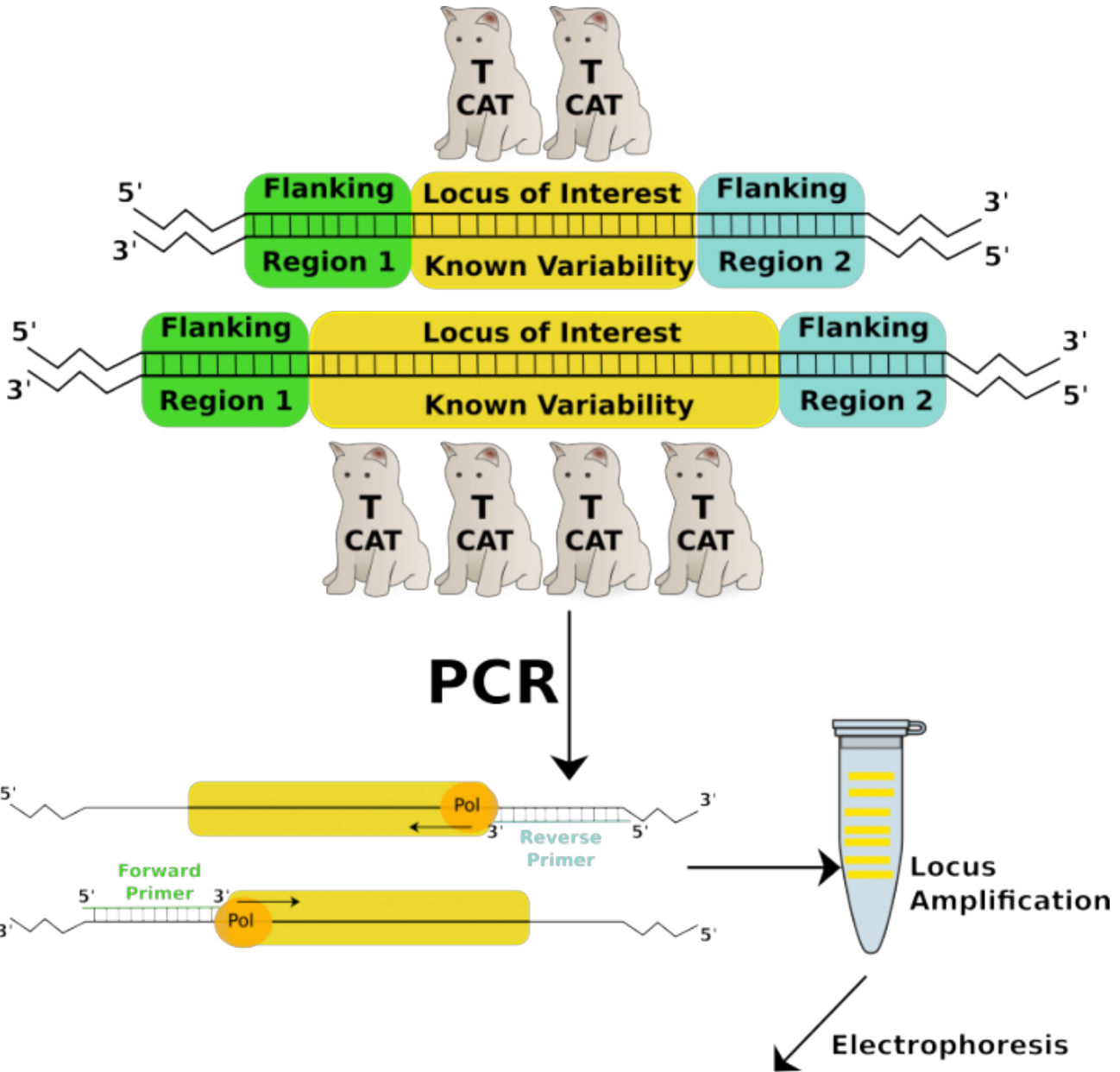
TH01 STR: Outside of the STR, there is flanking areas of known sequence. The primers that amplify TH01 in PCR recognize these flanking sequences to amplify the TCAT repeats.

At a crime scene, criminals don't often leave massive amounts of tissue behind. Scant evidence in the form of a few cells found within bodily fluids or stray hairs can be enough to use as DNA evidence. DNA is extracted from these few cells and amplified by PCR using the specific primers that flank the STRs used in CoDIS.



DNA evidence from a crime scene: DNA can be extracted from cells found from various sources at a crime scene. PCR can amplify this small amount of DNA.

Amplified DNA will be separated by gel electrophoresis and analyzed. Size reference standards and samples from the crime scene and the putative suspects would be analyzed together. In a paternity test, samples from the mother, the child and the suspected father would be analyzed in the same manner. A simple cheek swab will supply enough cells for this test.



Size Standard	PCR Mom	PCR Dad	PCR Sample
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>



TH01 locus used in a Paternity/Maternity test: Individual PCR reactions are run for each sample (mom, dad, child). The TH01 primer pair specifically amplifies the locus. Each amplified sample is run on the same gel to resolve the different alleles of TH01 from each individual. From this test the sample could be the offspring from these 2 parents but use of more STRs would make it more definitive. Count the TCATs.

External Resources

- Flash animation walking through what a STR is <http://www.dnalc.org/view/15981-DNA-variations.html>

Tags: [quantitative reasoning](#), [analysis](#), [visual communication](#), [integration of knowledge](#), [breadth of knowledge](#)