Genomics and Human Identity

Grades 9-12

Lesson 2

GENOME
UNLOCKING LIFE'S CODE

Inspired by the museum exhibit
Genome: Unlocking Life's Code
unlockinglifescode.org

Table of Contents

Genomics and Human Identity Lesson 2 – The Tiger Shark and the Missing Fishermen

About the Genomics and Human Identity Lessons

This is one of a series of inquiry-based lessons from the National Human Genome Research Institute inspired by the Smithsonian National Museum of Natural History exhibit Genome: Unlocking Life's Code and the website http://unlockinglifescode.org. NHGRI prepared this educational resource to bring the genome research featured in the exhibit into high school classrooms and other venues. These lessons are developed collaboratively with educators, scientists, teachers, and students nationwide.

Introduction to Lesson 2 – The Tiger Shark and the Missing Fishermen

The Tiger Shark and the Missing Fishermen is a CSI-style forensics activity where students analyze individuals' DNA data (16 different loci with multiple variations each) from a real, published police incident. The DNA data extracted from human remains found within a shark were used to determine the identity of a person. Students will determine the identity of the person by comparing the DNA profiles (variations in sequence) of the remains with the DNA of close relatives of two missing fishermen.

The lesson addresses the genetic concepts of inheritance and variation of traits by engaging students in the process and application of genomic research on human genetic variation, identity, and ancestry. Students participate in hands-on inquiry-based learning that emphasizes collecting data, analyzing and comparing data, and drawing conclusions. The lessons are guided by the Next Generation Science Standards. Genomics and Human Identity Lesson 2 is also designed to emphasize the importance of multidisciplinary collaboration in driving discoveries in science and address common misconceptions in genetics and genomics.

Lesson Components

- Teacher's Guide – detailed description of lesson objectives, procedures, and materials needed

- Student Handouts – these worksheets are included at the end of each lesson description in the teacher's guide

- PowerPoint Slides – Each lesson has a PowerPoint presentation to guide you and your students through the activities

- Web – http://unlockinglifescode.org Teacher background information and supplementary activities for students are available online. However, none of the student lessons in this Teachers Guide requires computers or internet access.

Time

This lesson is expected to take the equivalent of two 45 minute class periods.

Key Concepts

This lesson takes a deeper look at the genetic variation we inherit from our parents and pass on to our children. DNA patterns of inheritance among family members are explored.

Learning Objectives

After completing this lesson, students will:

- Understand from the analysis of crime-scene data the patterns of inheritance of genetic variation from parents to children
- Gain knowledge of specific genetic variations, short tandem repeats, in our genome
- Understand how and why forensic scientists study variation of short tandem repeats to identify and compare human DNA samples

Next Generation Science Standards

This lesson addresses the NGSS life science standards **HS-LS3 "Heredity: Inheritance and Variation of Traits"** and **HS-LS4 "Biological Evolution: Unity and Diversity."** Specifically, the lesson addresses these performance expectations and science practices:

Core Idea: HS-LS3 Heredity: Inheritance and Variation of Traits
HS-LS3-1: Ask questions to clarify the relationships about the role of DNA and chromosomes in coding the instructions for characteristic traits passed from parents to offspring
HS-LS3-2: Make and defend a claim based on evidence that inheritable genetic variation may result from: (1) new genetic combinations through meiosis, (2) viable errors occurring during replication, and/or (3) mutations caused by environmental factors
HS-LS3-3: Apply concepts of statistics and probability to explain the variation and distribution of expressed traits in a population.

Core Idea: HS-LS4 Biological Evolution: Unity and Diversity
HS-LS4-2: Construct an explanation based on evidence that the process of evolution primarily results from four factors: (1) potential for a species to increase in number, (2) the heritable genetic variation of individuals in a species due to mutation and sexual reproduction (3) competition for limited resources, and (4) the proliferation of those organisms able to survive and reproduce in the environment.

Science Practices:
Asking Questions and Defining Problems
Analyzing and Interpreting Data
Engaging in Argument from Evidence
Using Mathematics and Computational Thinking
Constructing Explanations and Designing Solutions
Obtaining, Evaluating and Communicating Information

Prerequisite Knowledge

Students should know that human genomes are 99.5% the same.   The 0.5% genetic variation between individuals, i.e., the 10 to 15 million single nucleotide polymorphisms, is the result of sequence differences in their DNA. These variants are most likely in non-coding regions of the genome. Students should know that we inherit our genetic information from the DNA in the nucleus of the egg and sperm of our parents. They should also be familiar with the structure of DNA and recognize that our chromosomes are made up of long stretches of four base pairs in different combinations. Finally, knowledge of how the polymerase chain reaction (PCR) amplifies small pieces of DNA for analysis will be helpful in understanding the forensics data.

Materials and Handouts

Lesson 2 PowerPoint

Handout 2-1: A Bad Week for Fishing	1 copy per student
Handout 2-2: DNA Identity Analysis and Allelic Ladder (2-pages)	1 copy per student
Handout 2-3: Forensic DNA Analysis with the E-gram (2-pages)	1 copy per student
Handout 2-4: Forensic DNA Analysis for Shark Scenerio	1 copy per group
Handout 2-5: DNA Forensic Report (2-pages)	1 copy per student

Preparation

Prepare handouts for students.

Procedure

1. Ask students if they have ever seen a TV show or movie where DNA evidence is used to identify a person involved in a crime.

If your class did Lesson 1 of Genomics and Human Identity, remind them that they grouped classmates using four common traits. Four traits were clearly not enough to disinguish one person from another. (If you did not do Lesson 1, go straight to **Slide 2-1**)

Slide 2-1: CSI-style DNA Analysis:

Lesson 2 - CSI-style DNA Analysis

How many "traits" do the police use to make a positive match between a person and evidence from a crime scene?

Genomics and Human Identity   Slide 2-1   Lesson 2

Ask the class for how many "traits" or DNA variants they think the police use in matching DNA evidence from a crime scene.

Allow students to guess and expect a range of answers.

The FBI uses 13 DNA loci from across most of the chromosomes. There is debate now on increasing the number of loci to 20.

Students will now consider an actual DNA forensics case from 2010. At that time most crime labs used 16 loci in identification.

2. Introduce the DNA analysis activity by giving each student **Handout 2-1: A Bad Week for Fishing** (this may also be given out to read as homework the day before). Show **Slide 2-2: Shark Scenario:**

Shark Scenario

On Aug 29, 2010, two fishermen in the Bahamas were reported missing after they tried to swim to a nearby shore following engine trouble with their boat.

Six days later, a local banker caught a 12-foot tiger shark off the coast of New Providence, Bahamas. Once inside the boat, the shark regurgitated what appeared to be a human leg. Three other human remains were found inside the shark.

Did the remains found in the shark belong to either of the fisherman?

Give students time to read **Handout 2-1**.

Ask students how they would solve the case.

Note that this is an actual case that was posted on the Promega Corporation website **http://www. promega.com/resources/teaching-and-training/5-solving-a-crime-using-dna-analysis**. This is a link to the Promega lesson on "Solving a Crime Using DNA Analysis". Promega produces one of the DNA analysis kits used in forensics laboratories. To solve the identity of the remains in the shark, forensic scientists at Forest Green Laboratories looked at variants from 16 **Short Tandem Repeat (STR)** loci using a PCR-based technique that results in an electropherogram.

The tiger shark article is online at: **http://www.promega.com/resources/profiles-in-dna/2013/ dna-identification-of-human-remains-obtained-from-a-tiger-shark/**

You may also wish to add that the medical examiner could not determine the cause of death and that tiger sharks are known scavengers.

Tell students that they will now solve the case by examining DNA evidence.

3. Give each student a copy of **Handout 2-2: DNA Identity Analysis – the E-gram. Show Slide 2-3: Short Tandem Repeats (STRs)**

Short Tandem Repeats (STRs)

TCTTATACTCATGAAATCAACAGAGGCTTGC   Row A
ATGTATCTATCTG TCTA TCTA TCTA TCTA   Row B
TCTA TCTA TCTA TCTA TCTA TCTA TCTA   Row C
TCTA TCTA TCTA TGAGACAGGGTCTTGCTC   Row D
TTCACCCAGATTGGACTGCAGTGGGGGAAT   Row E

TCTA is repeated 14 times

Genomics and Human Identity   Slide 2-3   Lesson 2

As students are looking at Handout 2-2, explain that the DNA variation forensic scientists examine represents short, repeat sequences. Slide 2-3 shows an example of a Short Tandem Repeat, or STR, where TCTA is repeated 14 times in one section of DNA. These STRs are found through out chromosomes in our genome. Different people may have a different number of repeats at a given locus, but we usually inherit these repeats directly from our parents.

Tell students that each peak in the e-gram represents one variation of a STR at a specific location. The STR example in the slide has TCTA repeated 14 times. One person may have the same number of TCTA repeats, another may have only six, and so forth.

Remind students of the DNA they examined in Lesson 1 and the repeated sequences. Each of the variants in the 16 loci in this analysis is a **Short Tandem Repeat**, not a single point mutation.

4. Show **Slide 2-4: DNA ID - 16 loci**
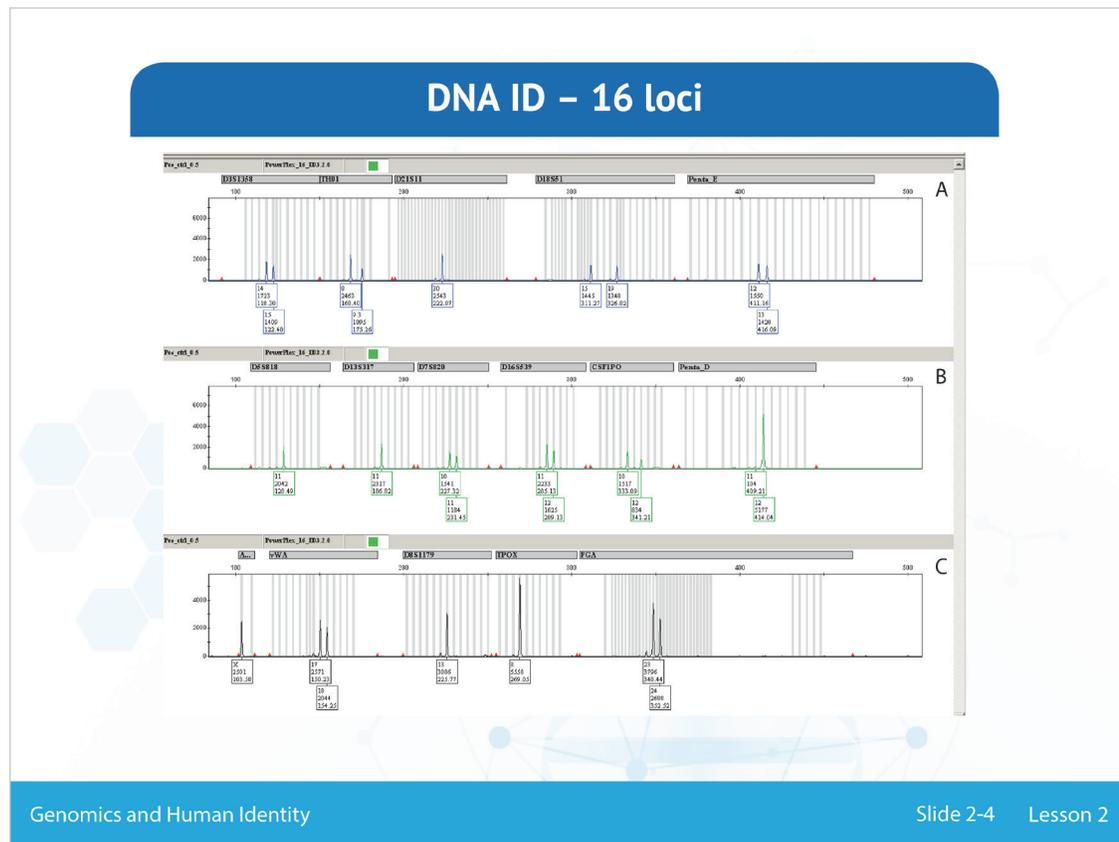
Tell students that this slide shows the type of data that forensic analysts use to do DNA matching. These are the DNA variation results for one person. There are 16 STR loci shown; five are on row A, six on Row B, and five on Row C. Each peak represents a DNA variation.

Note: Students will examine this E-gram data in more detail with Handout 2-3.

[For a more detailed account of this type of DNA analysis, **http://www.promega.com/resources/ profiles-in-dna/2013/dna-identification-of-human-remains-obtained-from-a-tiger-shark/**]

Tell students that this E-gram is not from a single person. It shows all of the possible variants for each of the 16 STR loci. Many people think of a locus as having one or two possible variations – one "normal" and one "mutant." The sickle cell mutation is one such example. With STRs there are usually a number of common variations. This is why forensic scientists use these loci for identification purposes.

15 of the STR loci have at least 8 common variations. FGA (the last locus in row C) has the most variation with 28 alleles. The Amelogenin STR locus (the first locus in row C) only has two variants and is used to determine the sex of the individual. The person in Handout 2-2 has only one peak on the X chromosome and is female. A second peak from the Y chromosome appears in males.

5. Show **Slide 2-5: Short Tandem Repeats (STRs)**

Each locus has multiple variants
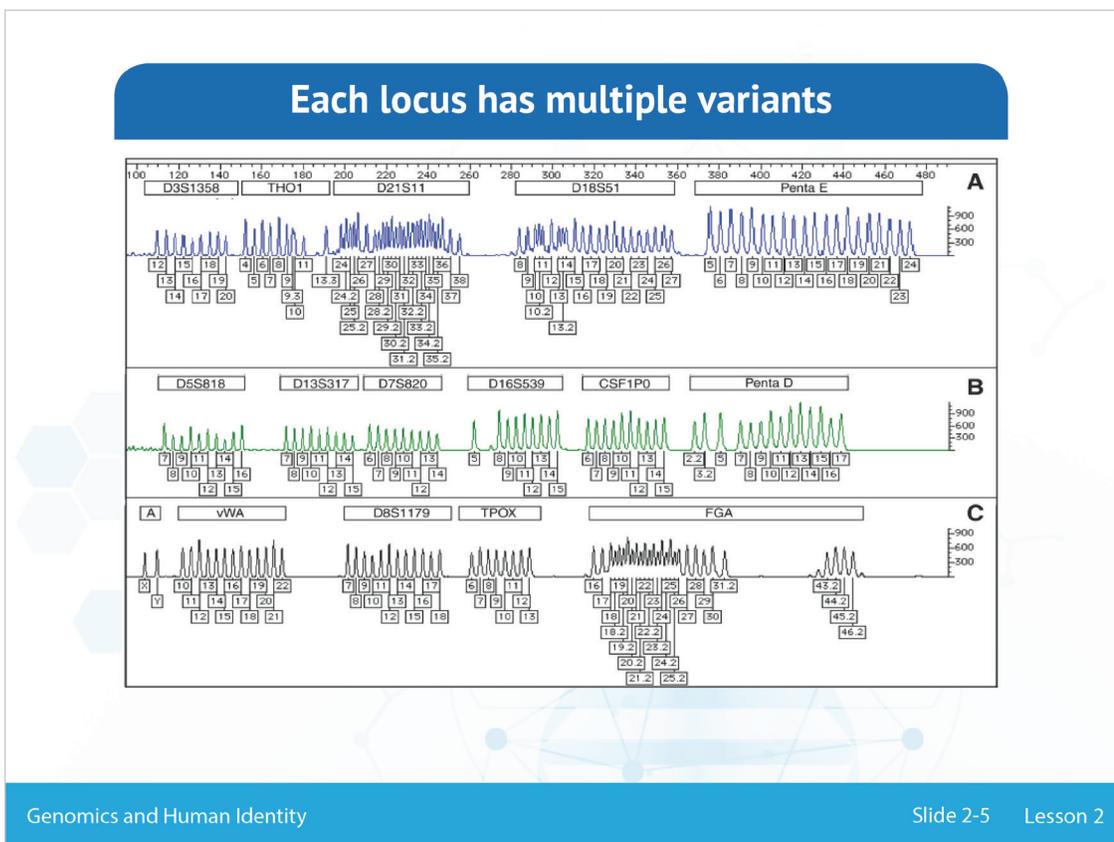
6. Give each student a copy of **Handout 2-3: Forensic DNA Analysis with the E-gram**. This handout walks students through the key points in understanding the E-gram and prepares them to understand and analyze the data table in Handout 2-4. (Students could also do this as homework).

7. If students are having difficulty with the E-gram, you may also wish to refer to **Slide 2-6: STR D3S1358** while they are doing **Handout 2-3**.

Slide 2-6 focuses on a single STR, the first one shown in each E-gram on **Handout 2-2** and **2-3**.

8. After students have completed **Handout 2-3**, review the answers with the class.

1. What are the variation numbers for each of the remaining STR loci? Complete the DNA profile of this person:

STR Locus	Variation number(s)
D3S1358	14, 15
THO1	8, 9.3
D21S11	30
D18S51	15, 19
Penta E	12, 13

2. There are two different variations at STR locus THO1. Where did these variations come from?

One variant is inherited from the person's mother, the other from the father.

3. The STR locus D21S11 only has one variation. Why?

The individual has two copies of variation #30, one on each chromosome.

4. The e-gram on Handout 2-2 and above is from a woman. She has variations 14 and 15 at STR D3S1358. If she has children with a man who has variations 12 and 19 at the same STR, what are the possible combinations of variations that their children would have?

A child will inherit variant 14 or 15 from the mother and 12 or 19 from the father. The four possible combinations are: 12, 14; 12, 15; 14, 19; and 15, 19. Note that variants are listed in numerical order.

Now that students have had a chance to examine the STR data on an E-gram, they will be ready to analyze the real data from the shark case.

9. Students will work in small groups to analyze the Shark Scenario data. Show **Slide 2-2** to reintroduce the scenario.
Give each student group a copy of **Handout 2-4: Forensic DNA Analysis for Shark Scenario**.

Shark Scenario

On Aug 29, 2010, two fishermen in the Bahamas were reported missing after they tried to swim to a nearby shore following engine trouble with their boat.

Six days later, a local banker caught a 12-foot tiger shark off the coast of New Providence, Bahamas. Once inside the boat, the shark regurgitated what appeared to be a human leg. Three other human remains were found inside the shark.

Did the remains found in the shark belong to either of the fisherman?

Genomics and Human Identity   Slide 2-2   Lesson 2

Give each student a copy of **Handout 2-5: Forensic DNA Report**.

Have students analyze the data on **Handout 2-4** to answer the questions on **Handout 2-5**. Each team should determine if the remains in the shark are from one or both fishermen and give evidence for their decsions.

10. **Slides 2.7, 2.8** and **2.9** are available to show your class if they need guidance in analyzing the data.

11. Once students have completed **Handout 2-5**, call on different teams to answer each question as you decide identity of the remains in the shark. Students should reach the following conclusions:

1. There are two missing fishermen. Multiple human remains were found in the shark. Are the remains (Samples 01, 02, and 03) from one person, two, or three? Show your evidence.

Samples 01 and 02 are identical, except where there are missing data points. This suggests that the remains are from one person.

2. Samples 04 and 05 are from the children of Fisherman #1. If samples 01 or 02 are from Fisherman #1, what STR variations would you expect to see in his children inherit?

The children should each have one of the father's variations for every STR locus.

a. Do any of the 16 STR loci suggest that samples 01 and 02 are from the father of the children in columns 04 and 05?

The variations at THO1, D21S11, D8S1179, and TPOX are consistent with Fisherman #1 being the father

b. Which of the 16 STR loci suggest that samples 01 and 02 are NOT from the father of the children in columns 04 and 05?

The variations at the other STR alleles do not match as expected. That is, there are variations in the children that are not present in the father. For example, neither child has variation 8 or 10 at D5S818 that are present in Fisherman #1.

3. Sample 06 is from the mother of Missing Fisherman #2. If samples 01 or 02 are from Fisherman #2, what STR variations would you expect to see in him inherit from his mother?

> Note: Your students may suggest that Fisherman #1 is not the biological father of the two children. Cases of mistaken paternity are sometimes brought to light through DNA analysis even though that is not the goal of the testing. Deciding if/how to present this data to the family raises ethical issues for forensic scientists.

Samples 01 and 02 should have at least one match with sample 06 for all of the STR loci. A son inherits half of his variations from his mother.

a. Do any of the 16 STR loci suggest that sample 06 is from the mother? Do variants match between columns 01, 02, and 06?

At least one variation from sample 06 is present at <u>every</u> STR locus of samples 01 and 02.

b. Do any of the 16 STR loci suggest that sample 06 is NOT from the mother? Do any of the variants in samples 01 or 02 not match a pattern of inheritance from sample 06?

No. All loci have a pattern of inheritence that is consistent with 06 being the mother.

4. Sample 07 is from the half-brother of Missing Fisherman #2. They share a father but have different mothers. If samples 01 or 02 are from Fisherman #2, how many STR variations would you expect the half brothers to share?

Each brother will inherit one of the father's two variations. There is a 50% chance that half-brothers will inherit the same variation at any given locus.

a. Do any of the variations at 16 STR loci suggest that sample 07 and samples 01 or 02 have the same father?

Samples 01, 02 and 07 share a variation in 11 of the 16 loci. This is consistent with the individuals having the same father and a different mother.

b. Do any of the variations at 16 STR loci suggest that sample 07 and samples 01 or 02 have the a different father?

Samples 01, 02 and 07 have no common variations in 5 of the 16 loci. These loci are D3S1358, Penta E, D5S818, vWA, and D8S1179. This is consistent with the individuals having the same father and a different mother.

5. Look at your analysis from questions 1 through 3. Whose remains are in the shark? What is your best evidence?

The observation that all loci between sample 06 and samples 01 and 02 share at least one variation provide the strongest evidence that the remains of Missing Fisherman #2 are in the shark. This is exactly the pattern of inheritence one would expect between a mother and her child. The matches between Fisherman #2 and his half brother also support this conclusion, although the evidence with the mother is much stronger.

The lack of a common variant between the samples 01 and 02 and the children of Fisherman #1 at 11 STR loci suggest that his remains are not in the shark – assuming that he is the biological father of the children.

6. Why are some data points missing? Do these missing data points draw any doubts on your conclusion?

These are real results from a forensics lab and sometimes DNA reactions do not happen as planned due to degraded samples (such as 03, perhaps), contaminated reactions, and other lab errors. Students will recognize that experiments sometimes don't work. The DNA from the samples in the shark show the most inconclusive data is likely due to the degredation. These inconsistencies do not draw significant doubt on the hypothesis that samples 01 and 02 are from Fisherman #2. Without data for sample 03, however, it is impossible to conclude that this sample is also from Fisherman #2.

12. Display **Slide 2-10** Shark Scenario Report
After reviewing Handout 2-5, ask class to summarize their findings.

13. Conclude this lesson by telling students that the crime lab analysis concluded that:

"Based on the DNA profiles obtained from the evidence, the mother of one of the missing fishermen cannot be excluded as the biological mother of the individual associated with the bone and tissue.

This is a rare case of human identification from biological remains recovered from a tiger shark. Significant in this process was the relatively short time between the incident and recovery of the remains. This factor and the quick response of the police for the recovery and preservation of evidence made DNA identification possible and gave a quick resolution to this case."

Glossary

Allele: One of two or more versions of a gene. An individual inherits two alleles for each gene, one from each parent. If the two alleles are the same, the individual is homozygous for that gene. If the alleles are different, the individual is heterozygous. Though the term allele was originally used to describe variation among genes, it now also refers to variation among non-coding DNA sequences.

DNA fingerprinting: A laboratory technique used to establish a link between biological evidence and a suspect in a criminal investigation. A DNA sample taken from a crime scene is compared with a DNA sample from a suspect. If the two DNA profiles are a match, then the evidence came from that suspect. Conversely, if the two DNA profiles do not match, then the evidence cannot have come from the suspect. DNA fingerprinting is also used to establish paternity.

DNA sequencing: A laboratory technique used to determine the exact sequence of bases (A, C, G, and T) in a DNA molecule. The DNA base sequence carries the information a cell needs to assemble protein and RNA molecules. DNA sequence information is important to scientists investigating the functions of genes. The technology of DNA sequencing was made faster and less expensive as a part of the Human Genome Project.

Polymerase chain reaction (PCR): A laboratory technique used to amplify DNA sequences. The method involves using short DNA sequences called primers to select the portion of the genome to be amplified. The temperature of the sample is repeatedly raised and lowered to help a DNA replication enzyme copy the target DNA sequence. The technique can produce a billion copies of the target sequence in just a few hours.

Short Tandem Repeats (STRs): STRs are typically two to ten nucleotides that repeat multiple times in a stretch of DNA sequence. For example, the sequence TGATGATGATGACCTA contains an STR. The nucleotides TGA are repeated four times. STRs vary among individuals and have become important for human identity DNA testing. These repeated regions of DNA can be grouped depending on the number and size of repeats in a specific DNA locus or region.

Variation: Genetic variation refers to diversity in gene frequencies. Genetic variation can refer to differences between individuals or to differences between populations. Mutation is the ultimate source of genetic variation, but mechanisms such as sexual reproduction and genetic drift contribute to it as well.

Resources

- **Solving a Crime Using DNA Analysis**

  http://www.promega.com/resources/teaching-and-training/5-solving-a-crime-using-dna-analysis

- **DNA Identification of Human Remains Obtained from a Tiger Shark**

  http://www.promega.com/resources/profiles-in-dna/2013/dna-identification-of-human-remains-obtained-from-a-tiger-shark/

Acknowledgements