History and Science

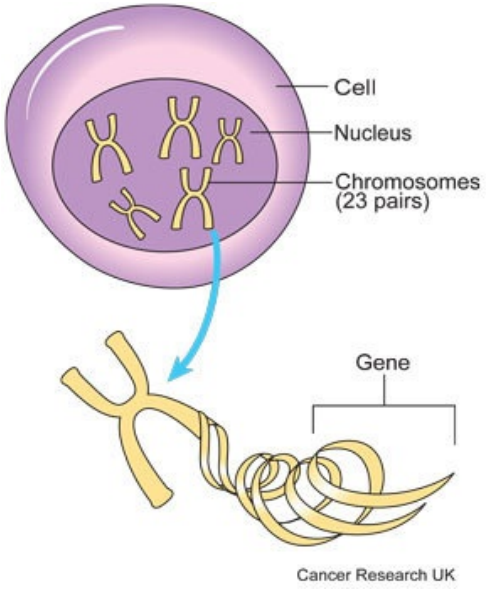
of Forensic DNA Testing

By: Paul Couenhoven

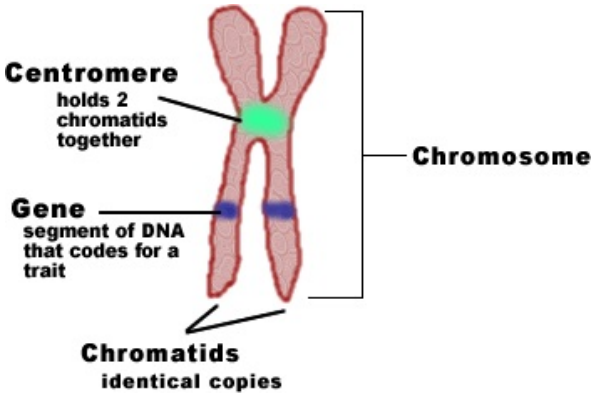
SCIENTIFIC BASICS OF DNA

What is DNA?

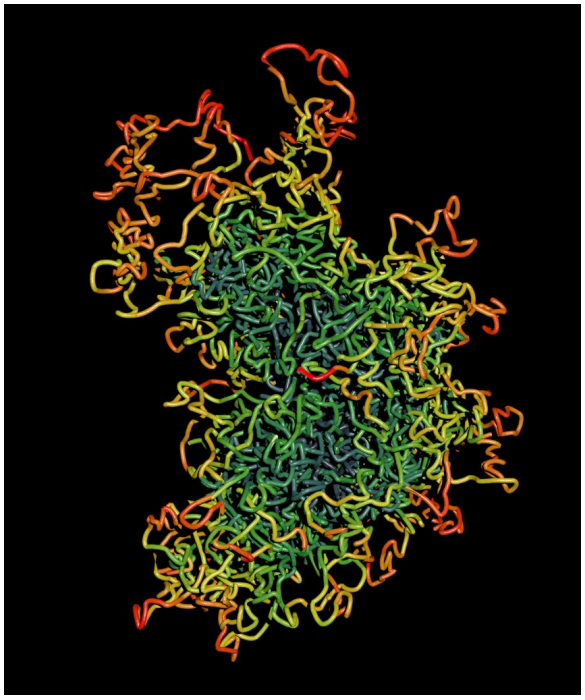
DNA stands for DeoxyriboNucleic Acid. It is the genetic material of a cell. The chromosomes inside the nucleus of the cell are made of DNA. It is very fine and tightly coiled but there may be as much as a meter in a single cell. DNA is really a code. It is divided up into sections. These sections are genes, which carry all the instructions for making up our body, or the body of any living organism. So there is a gene that tells the body to have brown hair, or blue eyes or a certain type of belly button (iny or outy) and so on. So the genes dictate how we are made and what our bodies look like.



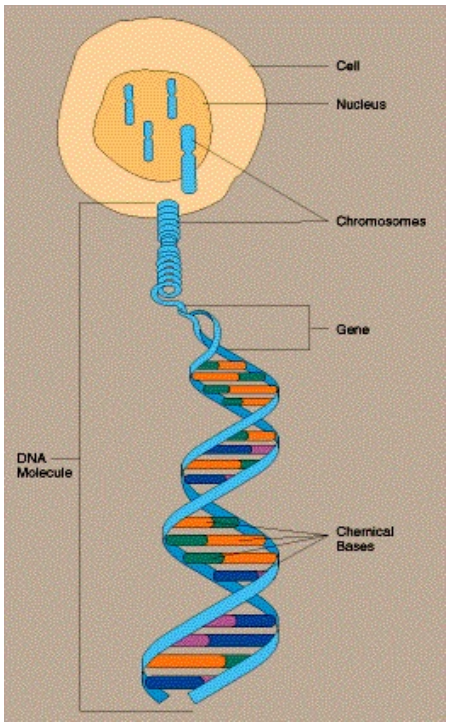
Chromosomes are DNA wrapped around proteins to form an X-shaped structure.



A chromosome actually looks something like this:



Here is another depiction of a chromosome, which shows the familiar DNA double helix design.



Alleles are corresponding pairs of genes located at specific positions in the chromosomes. In humans, one allele is inherited from the mother and the second is from the father. Together, they determine the genotype of their host organism. The genotype of an organism is the inherited map it carries within its genetic code. For example, the alleles for eye color are found on chromosomes 15 and 19, and depending on which ones someone has, he or she may have blue, brown, green, gray, or hazel eyes, and sometimes a mixture of these traits is present. Alleles that determine some aspect of the phenotype, that is, the physical appearance of an organism, are said to be "coding alleles," while "non-coding alleles" or "junk DNA" are those which do not appear to have an impact on phenotype.

THE GENESIS OF DNA TESTING

In 1984, Sir Alec Jeffreys, a British geneticist, discovered the technique of DNA testing to determine a genetic "fingerprint" in a laboratory in the Department of Genetics at the University of Leicester, England. Jeffreys says he had a "eureka moment" in his lab after looking at the X-ray film image of a DNA experiment which unexpectedly showed both similarities and differences between the DNA of different members of his technician's family. Within about half an hour he realized the possible scope of DNA fingerprinting, which uses variations in the genetic code to identify individuals.

Jeffreys' DNA method was first put to use in 1985 when he was asked to help in a disputed immigration case to confirm the family identity of a British boy whose family was originally from Ghana. A family from Ghana immigrated to the UK and became citizens. However, one of the sons went back to Ghana and was stopped from returning to the UK because he had a forged passport. The family's lawyer contacted Jeffreys and asked whether he could confirm that the boy was in fact the mother's son and not her nephew. Samples of DNA were taken from the mother, from the son whose identity was disputed, and from the mother's three undisputed children. The DNA patterns confirmed the relationship between the mother and the son in question. Moreover, the testing confirmed that all four children had the same father Sir Jeffreys said he saw the relief in the mother's face when she heard the results.

DNA fingerprinting was first used in a police forensic test in 1986. Two teenagers had been raped and murdered in Narborough, Leicestershire, in 1983 and 1986 respectively. Although the attacks had occurred 3 years apart, similarities led the police to believe that one person was responsible for

both. A suspect in custody, Richard Buckland, confessed to the most recent murder but not the earlier one. Jeffreys was asked to do DNA profiling on a blood specimen that was collected from the suspect and on tissue specimens and semen collected from the two victims.

The DNA profiling revealed that the semen from both victims was identical, proving that one person had committed both murders. However, the results also proved that Richard Buckland was not the murderer. His confession had, evidently, been false. He was released and became the first suspect to be cleared of a crime by DNA evidence.

A large-scale manhunt was then launched to find the person whose DNA profile matched that of the killer's semen. All adult men who lived in the area were asked to give blood or saliva specimens for testing. More than 5000 specimens were collected and DNA profiling carried out on the 10% of men who had the same blood type as the killer, but no match was found. The police and the public were disappointed that this new and sophisticated test was unable to identify the killer.

Six months after the initial investigation, a woman reported overhearing a man who claimed to have given blood on behalf of a colleague, Colin Pitchfork. Pitchfork was apprehended and his blood tested; the long-sought DNA match was made, and Pitchfork was convicted of both murders.

DNA TESTING

The DNA of all human beings is nearly identical. Approximately 99.9% of the sequence of DNA is in the exact same order. This determines common human features such as two eyes, ears on both sides of the head, and long bones in forearms and calves. Although looking at these parts of the DNA molecule might help us determine it is human DNA — rather than, say, banana DNA — it isn’t helpful in distinguishing one human from another.

There are, however, places on the human DNA molecule that are different. Of the approximately 3.2 billion base pairs in the human genome, some 3 million base pairs of DNA (about 0.10 percent of your entire genome) vary from person to person. These variations are at the core of DNA testing.

The DNA sections looked at in forensic science is not currently known to have any function (such as coding for eye color or the potential predisposition toward a genetically inherited disease) —

except for amelogenin , which is used in forensic analysis for gender differentiation. The areas at which forensic analysts look are always found in the same spots on the same chromosomes. Each specific location is called a locus (pronounced “LOW-cuss”). The forensic science community typically uses a minimum of 13 genetic loci (plural for locus , pronounced “LOW-sigh”), referred to as the 13 core CODIS (Combined DNA Index System) loci. This enables laboratories to search profiles against other profiles already in the CODIS databank (although some laboratories test more than the 13 core CODIS loci. Many labs will examine 15 loci if they are comparing DNA from an evidentiary sample to DNA collected from a known suspect.

Several basic steps are performed during DNA testing. The general procedure includes: 1) The collection of samples which might contain DNA; 2) The isolation of the DNA from an evidence sample and the isolation of DNA from a sample from a known individual; 3) The processing of the DNA so that test results may be obtained; 4) The determination of the DNA test results and 5) the comparison and interpretation of the test results from the unknown and known samples to determine whether the known individual is not the source of the DNA or is included as a possible source of the DNA.

Collecting DNA Evidence

The first step in the process is collecting evidence which could contain DNA. All it takes is a few cells to obtain enough DNA information to identify a suspect with near certainty. For this reason, law enforcement officials now try to take great care at crime scenes. Police officers and detectives often work closely with laboratory personnel or evidence collection technicians to make sure evidence isn't contaminated. This involves wearing disposable gloves and using disposable instruments, which can be discarded after collecting each sample. While collecting evidence, officers are instructed to avoid touching areas where DNA evidence could exist. They also are instructed to avoid talking, sneezing and coughing over evidence or touching their face, nose or mouth.

The following are common sources of DNA evidence:

A weapon, such as a baseball bat, fireplace poker or knife, which could contain skin cells, blood or other tissue

A hat or mask, which could contain skin cells, hair or dandruff

A facial tissue or cotton swab, which could contain mucus, skin cells, blood or earwax

A toothpick, cigarette butt, bottle or postage stamp, all of which could contain saliva

A used condom, which could contain semen or vaginal or rectal cells

Bed linens, which could contain skin cells, hair, blood or semen

A fingernail or partial fingernail, which could contain scraped-off skin cells

A swab of genitals and nipples in a case involving suspected sexual assault, which could contain saliva or semen.

A rather famous case in Cook County, Ill., was solved by obtaining DNA from chicken bones — not chicken DNA, but the DNA of the person who ate the chicken off the bones.

When investigators find a piece of evidence, they place it in a paper bag or envelope, not in a plastic bag. This is important because plastic bags retain moisture, which can damage DNA. Direct sunlight and warmer conditions may also damage DNA, so officers try to keep biological materials at room temperature.

When an investigator believes a stain at a crime scene might be blood, semen or saliva, they will often use a presumptive test to determine whether the stain contains human biological material rather than, for example, wine or ketchup. A positive presumptive test suggests the presence of a particular body fluid. The investigator will then take a cut from material which contains the stain, or swab the stain with a Q-tip to collect a sample.

For example, in the recent movie Gone Girl, police used Luminol to detect the presence of blood on the kitchen floor even though it had been mopped up.

DNA Analysis: Traditional Technique

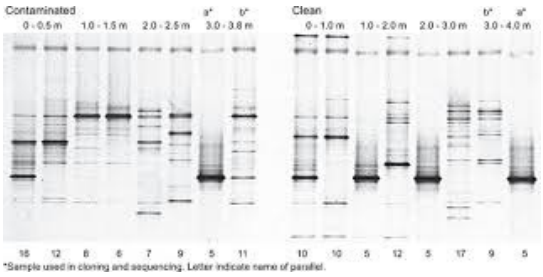
Evidence is examined and tested for DNA at a forensic laboratory. The DNA is isolated from the cells collected in the sample. The current preferred method is called short tandem repeat (STR) analysis. STR testing examines regions of the DNA molecule that tend to repeat themselves in short, adjacent, or tandem segments. This method can be used with a much smaller sample of DNA than prior methods. Scientists amplify the sample through a process known as polymerase chain reaction,

or PCR. PCR makes copies of the DNA much like DNA copies itself in a cell, producing almost any desired amount of the genetic material. Some forensic scientists liken this to molecular Xeroxing.

The process of making multiple copies of the DNA is achieved through heating and cooling. A tube or well — containing the forensic sample (DNA), the primers for the primer binding sites with fluorescent dyes attached, additional bases, and an enzyme to replicate the DNA — is heated. This causes the DNA molecule (the double helix or "ladder") to split into two strands. At a high temperature, the DNA strands will stay apart. At a slightly higher temperature, copies of the DNA are made by the enzyme that replicates the DNA, adding complementary bases to each of the single DNA strands. The temperature is then lowered, and the primers will bind to their corresponding complementary bases on the original DNA strands. As the PCR process continues, we heat up the sample and reagents again and the strands split again; cool it down and the primers bind; heat it up a little more and additional bases in the tube bind in that same predictable pattern. The yield after two PCR cycles is, therefore, four copies of the areas of the original DNA strand in which we are interested. This process of heating and cooling takes place in a thermal cycler and is done approximately 30 times . At the end of the process, there are literally billions of copies of areas of interest.

Once the laboratory has a large number of copies, the generated DNA pieces or fragments are separated by size. The most commonly used method for separation is via the use of a capillary electrophoresis (CE) instrument. A capillary (shaped like a very thin straw) is inserted into the tube or well which contains the DNA and draws out a small amount of the amplified product mixture. The DNA travels up and through the straw in a predictable manner — smaller DNA fragments moving faster than larger DNA fragments. Once a piece of DNA reaches the end of the capillary, it passes over a laser light. This excites the fluorescent dye incorporated during the PCR process and causes the bound dye to fluoresce (light up). A camera captures and measures the emitted light, which is reproduced in the corresponding dye color in an electropherogram. When the alleles come across the laser light, the color as well as the length of the fragment is recorded. The color of the alleles, along with their length, indicates which alleles go with which other alleles as well as what locus they come from.

Example of an electropherogram:



DNA Analysis: Specialized Techniques

Although most labs today use STR techniques for their DNA analysis, there are situations that require a different approach. One such situation is when there are multiple male contributors of genetic material, which sometimes happens in sexual assault cases. The best way to resolve the complex mixture and sort out exactly which men were involved is Y-marker analysis. As its name suggests, this technique examines several genetic markers found only on the Y chromosome. Because the Y chromosome is transmitted from father to son but not to daughters, DNA on the Y chromosome can be used to identify DNA from different males.

Another situation involves identifying old remains or biological evidence lacking nucleated cells, such as hair shafts, bones and teeth. STR testing can't be used on these materials because they require DNA found in the nucleus of a cell. In these cases, investigators often use mitochondrial DNA, which uses DNA from a cell's mitochondria. Mitochondria are cellular organelles (like a cell's organs) that convert chemical energy from food into a form that cells can use. Investigators have found mtDNA testing to be very useful in solving cold cases, which are murders, missing-person cases or suspicious deaths that are not being actively investigated. Cold cases often have biological evidence in the form of blood, semen and hair that has been stored for a long time or improperly stored. Submitting those degraded samples for mtDNA testing can sometimes break the case open and help detectives find the perpetrator.

DNA Matching

The electropherogram is like a DNA "picture" with columns of dark-colored parallel bands. To identify the owner of a DNA sample, the DNA "fingerprint," or profile, must be matched, either to

DNA from a suspect or to a DNA profile stored in a database.

When there is an identified suspect, investigators take a DNA sample from the suspect (usually by taking a buccal swab), send it to a lab and receive a DNA profile. Then they compare that profile to a profile of DNA taken from the crime scene. There are three possible results:

1. Inclusions -- If the suspect's DNA profile matches the profile of DNA taken from the crime scene, then the results are considered an inclusion. In other words, the suspect is included (cannot be excluded) as a possible source of the DNA found in the sample.

2. Exclusions -- If the suspect's DNA profile doesn't match the profile of DNA taken from the crime scene, then the results are considered an exclusion. Exclusions almost always eliminate the suspect as a source of the DNA found in the sample.

3. Inconclusive results -- Results may be inconclusive for several reasons. For example, contaminated samples often yield inconclusive results. So do very small or degraded samples, which may not have enough DNA to produce a full profile.

Sometimes, investigators have DNA evidence but no suspects. In that case, law enforcement officials can compare crime scene DNA to profiles stored in a database. Databases can be maintained at the local level (the crime lab of a sheriff's office, for example) or at the state level. A state-level database is known as a State DNA index system (SDIS). It contains forensic profiles from local laboratories in that state, plus forensic profiles analyzed by the state laboratory itself. The state database also contains DNA profiles of convicted offenders. Finally, DNA profiles from the states feed into the National DNA Index System (NDIS).

To find matches quickly and easily in the various databases, the FBI developed a technology platform known as the Combined DNA Index System, or CODIS. The CODIS software permits laboratories throughout the country to share and compare DNA data. It also automatically searches for matches. The system conducts a weekly search of the NDIS database, and, if it finds a match, notifies the laboratory that originally submitted the DNA profile. These random matches of DNA from a crime scene and the national database are known as "cold hits," and they are becoming

increasingly important. Some states have logged thousands of cold hits in the last 20 years, making it possible to link otherwise unknown suspects to crimes.

DNA Profile Frequency Calculations

Currently, time and expense limit an examination of an individual's entire genome, which would show unique identity. Due to the fact that DNA typing is only an examination of a DNA sample's sequence and/or length at 13 to 15 discrete locations, a match in DNA typing is always a statistical exercise. In order to determine the probability that a particular genotype might occur at random in a population, population data must be compiled to make an estimate of the frequency of each possible allele and genotype. Based on work done in the field of statistics, it has been determined that a minimum sampling of 100 people can be used to infer the frequency of occurrence of each of these alleles in the entire population. Typically, the sample population is a group of 100 to 200 persons in an identified group (Caucasian, Hispanic, African-American, Asian). Some labs have their own databases, and some labs rely on work conducted by the FBI or other labs for their databases.

Remember the scientist is only comparing 13 or 15 different places on the genome. It's not logically impossible for two entirely unrelated people to match just by chance. We need to know roughly how unlikely that is. This can be calculated from the proportion of the population that has each of the different numbers of repeats at the 13 to 15 locations. These allele frequencies are estimated from databases that are collected from forensic laboratories.

At each core CODIS locus, the possible types one can have are labeled by number. At TH01 (the name of one of the 13 CODIS loci), for example, the types that have been observed are 5, 6, 7, 8, 9, 9.3, 10 and 11. Generally, each person on the planet has two of these: one from mom and one from dad. These types are referred to as alleles (pronounced "uh-LEELS"). If the two alleles in a profile are identical (in other words, the person received a 5 from mom and a 5 from dad), they are homozygous. If the two alleles are different, say, a 5 from mom and an 8 from dad, they are heterozygous at that locus.

Probability calculations are based on knowing allele frequencies for each STR locus for a

representative human population. For example, from studying people's genes, scientists have determined the frequency of alleles at the TH01 location in the general population as follows:

Allele of TH01	Frequency of this specific allele in the general population
4	0.001
5	0.001
6	0.266
7	0.160
8	0.135
9	0.199
9.3	0.200
10	0.038
11	0.001

Each person has two alleles at each locus; with 100 people, there would be 200 potential alleles in the database at each locus; with 150 people, there would be 300 potential alleles at each locus. The probability can be an extremely low number when all 13 CODIS STR markers are included in the DNA profile. Think of it in terms of the California Super Lotto. You have to match 7 numbers to win the Super Lotto. You may guess 1 or 2 or even 3 correctly, but the odds of matching all 7 numbers are 1 in 18 million. This means if one person purchases 50 Lotto tickets each week, they will win the jackpot about once every 5,000 years. In the context of DNA, there are 13 loci to match, and numerous possibilities at each marker, as seen in the example for TH01, above. When all 13 loci are a match, the statistical odds of this occurring randomly are extremely low.

Results are frequently reported as an inclusion or nonexclusion along with a random match/man probability (RMP) or other frequency estimate. RMP is the probability of randomly selecting an unrelated person from the population who could be a potential contributor to the DNA profile obtained from crime scene evidence. It is the theoretical "chance" that, if you sample one person at random from the population, they will have the same DNA profile as the one obtained from the evidence sample.

RMP values are typically associated with single-source DNA profiles, but they can be calculated for mixture samples as well. When applied to mixtures, this calculation is referred to as a modified RMP, which includes an assumption of the number of contributors to the mixture.

Use of an RMP statistic in a DNA report will be similar to the following:

The approximate frequency of the DNA profile obtained from [item of evidence] is:

1 in 3.3 sextillion in the Caucasian-American population.

1 in 75 sextillion in the African-American population.

1 in 38 quintillion in the Hispanic-American population.

1 in 22 septillion in the Asian-American population.

INTERESTING DNA STORIES

The use of DNA is not limited to evidence from humans. For example, detectives in Phoenix, Arizona, were able to link a suspect to a murder victim by testing the DNA of a palo verde tree found at the crime scene. In 1992 Phoenix police found a pager at a murder scene which was traced to Mark Alan Bogan . Bogan said his pager had been stolen, and denied ever having been to the area where the murder occurred. Police found some palo verde seed pods when they searched Bogan's truck. To prove that the pods came from a palo verde tree located at the crime scene and not some other palo verde tree, detectives turned to DNA analysis. The pods found in the truck matched each other -- and matched the pods taken from the tree at the crime scene. It was the first time the DNA fingerprint of a plant was used in a criminal trial.

In the JonBenet Ramsey case, investigators scraped clothing that JonBenet had been wearing. There was enough evidence in two different places to create a DNA profile that matched one already created from blood -- both of which belong to a male not related to JonBenet. This convinced prosecutors that the Ramsey family could not have been responsible for JonBenet's death.

Any DNA profile can give a false result if it is contaminated. Although there have been no documented cases of a laboratory worker intentionally contaminating a DNA sample, DNA samples have been contaminated or even faked by criminals in order to avoid prosecution. In 1992, Dr. John Schneeberger was accused of raping one of his patients while she was sedated. A DNA profile was created using the sample that he left on the victim. A profile from a sample of his blood did not match the crime-scene sample, and the case was closed. The victim persisted, and eventually Dr. Schneeberger was convicted after additional DNA samples showed a match. He was able to avoid

the initial match by implanting a drain in his arm filled with another man's blood and an anticoagulant, and skillfully getting the technician who drew his blood to do so from that spot. Such a ploy would be difficult today when samples are collected by swabbing the inside of a person's cheek rather than by collecting a blood sample. It would be difficult to coat the inside of one's mouth with someone else's skin cells and saliva.

Note that DNA analysis has also resulted in the reversal of many convictions. Post-conviction DNA testing really began with a 1996 National Institute of Justice report that spotlighted 28 people convicted of rape and murder who had been exonerated due to later DNA testing. Between1989 and 2012, 323 convicted criminals have been released where DNA testing was part of the evidence exonerating them. (Report by the National Registry of Exonerations, a joint project of the University of Michigan Law School and the Center on Wrongful Convictions at Northwestern University School of Law).