DNA analysis

Coiled inside the nucleus of most of our cells is a 6-ft (2-m) long spiral. Parts of it encode who we are, and what we look like. The remainder contains patterns that repeat a unique number of times for each individual. The ability to count these "stutters" has revolutionized crime investigation.

◀ **DNA**
The double helix, or spiraling spiral, shape of the DNA strand makes it very compact. The entire 6-ft (2-m) length fits inside the minute nucleus of a cell.

This crucial spiral is deoxyribonucleic acid, or DNA. Portions of this long molecule form our genes, or the "genetic code," from which all our characteristics are derived. The structure of DNA is like a curled-up, spiraling ladder with three billion rungs. The rungs are pairs of simple organic chemicals called "bases." There are four kinds of bases (see box below), and the exact order in which they occur along a strand of DNA is different in every individual, except identical twins. This unique character of our DNA is the main reason why it is such an important tool for identifying individuals. It is also readily available—it exists in every living cell that has a nucleus, including hair roots and bone marrow—and only tiny amounts are needed for analysis.

Crucial "junk" DNA

By examining carefully selected sections of DNA taken from a crime scene, we can compare them to a sample taken from a suspect. A match strongly suggests that the suspect was the source of the DNA found at the crime scene. DNA can also be used to help identify a victim. We inherit half our DNA from each parent, so a part match from a missing person's parent can demonstrate a blood relationship to an otherwise unidentified body.

Only a small portion of our DNA actually provides the blueprint for who we are and what we look like. The remainder, so-called "junk" DNA, seems to perform no useful function that we know of. But it is this "junk" DNA that is crucial for identification purposes. This is because much of it consists of short sequences of base pairs that repeat end to end, called "short tandem repeats" (STRs). Though these sequences occur in everybody's DNA, the number of times they repeat varies dramatically between individuals. Counting the repetitions of a number of different sequences (up to thirteen) allows us to identify people uniquely. Locating the repetitions and separating them for analysis is possible because on either side of them there are always characteristic sequences of base pairs that are the same in everybody's DNA.

In this graphic representation of DNA, (I) marks the start of the repeated section. The sequence between (I) and (II) is repeated many times as the strand spirals toward us.

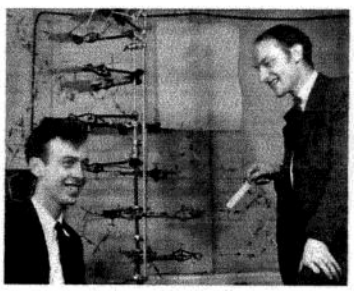
KEY:

Four kinds of bases make up the DNA "ladder." Two of them, adenine (A) and guanine (G), are purines; the other two, cytosine (C) and thymine (T), are pyrimidines. Bases always occur in pairs of purine and pyrimidine bonded together, A with T, and G with C.

G GUANINE
T THYMINE
C CYTOSINE
A ADENINE

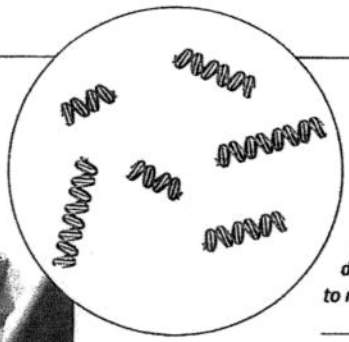
WATSON AND CRICK

The double helix structure of DNA was discovered by James Watson and Francis Crick in 1953, using X-ray diffraction images of the molecule. Some of these images were produced by Maurice Wilkins and Rosalind Franklin. Watson, Crick, and Wilkins shared the Nobel Prize in 1962.



James Watson b. 1928 and Francis Crick b. 1916

ADDING PRIMER ▲
PCR starts with the addition of primers to the samples of DNA in plastic reaction tubes. These are then heated and cooled in a thermal cycler.



◀ COLORED SPIRALS
Attached to each primer is a different colored fluorescent dye. Because the primer has bonded to the DNA, it makes the different sequences easy to recognize at the next step.

Extraction

Analysis begins with extraction of the DNA from the sample. This is usually done using a chloroform and phenol mixture, or a salt solution that separates DNA from the other material in the cell nucleus. This extraction process does not usually produce enough DNA for analysis, so the next step is to artificially increase the quantity of DNA in the sample using a technique called polymerase chain reaction (PCR).

Amplification

Polymerase is an enzyme in our cells that can copy or repair DNA. The "ladder" of DNA splits apart at the middle of each "rung" to create two strands. Since each base can only combine with a complementary base, each strand of DNA forms a template for its other, now missing, part. The polymerase enzyme works down each strand and adds the complementary bases as it goes, and thus reconstructs each strand into a complete, double-stranded DNA molecule.

PCR reproduces this process of synthesis in a test tube. Technicians mix the DNA sample with buffer (salt water), polymerase, the four bases, and primers. Primers are short, ready-made fragments of DNA that attach to either side of a chosen sequence. Different primers find different repeating sequences of base pairs—so as to analyze up to thirteen STRs.

Heating this mixture to just below the boiling point of water "unzips" the two strands of DNA. When the solution cools, the primers bond to the start and end of the target sequences. Warming the mixture allows the polymerase to reconstruct the second, complementary strand of DNA between the primers. The process is repeated 25–40 times. Each cycle replicates the target portions—those containing the short tandem repeats—while leaving the remainder of the DNA unchanged.

PCR can increase the size of a DNA sample millions of times over. This makes it possible to analyze samples as small as a couple of nanograms of blood—roughly one forty-millionth of a drop.

Analyzing the sample

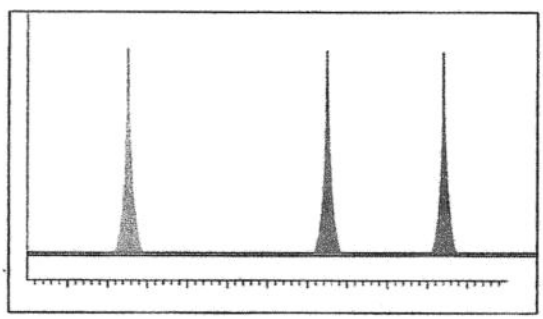
The product of PCR is a mass of DNA fragments. Their lengths vary according to the number of times the chosen sequences of base pairs are repeated. Each sequence will have two lengths, one from each of the parents. To analyze them, the fragments are sorted out using the process of electrophoresis—a kind of electrical race. DNA has a negative electrical charge, and is pulled toward a positive electrode, just as north and south poles on magnets attract each other.

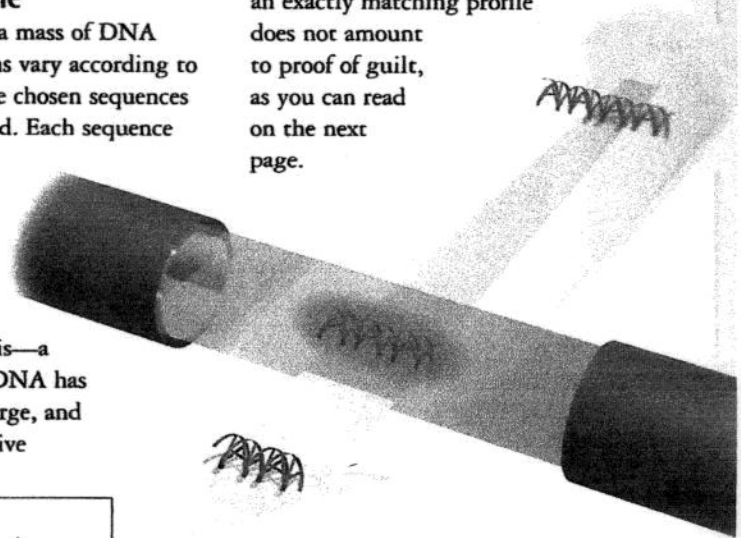
In electrophoresis, the movement of the fragments is artificially slowed by forcing them to travel through a stiff jelly, or down a very narrow tube. The smallest bits speed along, but longer pieces move more slowly, so the process neatly ranks the DNA fragments according to length.

A race is pointless if you cannot tell who has won, so each bit of DNA must have a unique marking. The primers added during the PCR "amplification" phase provide this. Each primer is tagged with a different fluorescent dye.

In the most automated form of DNA analysis, a powerful electric charge drives the DNA fragments along a tube. At its end, they move between a laser and a color-sensitive detector that is connected to a computer. The laser beam makes each of the dyed fragments fluoresce with its primer color. The detector captures the flashes of light, and software plots the fluorescence as peaks on a graph.

To evaluate the test results, the peaks from the suspect's sample are simply compared with those from the crime scene. However, in the absence of other evidence, an exactly matching profile does not amount to proof of guilt, as you can read on the next page.





THE FINISHING LINE
A high voltage drives the fragments of DNA along a capillary tube and past a photodetector, producing characteristic peaks (left) for comparison on a computer screen. Running a parallel control sample containing fragments of known length makes it possible to determine how many repeats are in the test samples.