DNA Fingerprinting

The New Sherlock Holmes

By Amrita Goyal

On a clear night, anyone with a good telescope can see that the sky is strewn with thousands of satellites—some are manmade satellites, launched into orbit for communications and defense, and some are natural ones, such as our beloved Moon. However, there is another type of satellite that is much closer to home: the human genome is also littered with "satellites."

Genetic satellites are entirely different from the ones overhead. A DNA satellite is a region that consists of a short sequence of DNA repeated over and over again—like an extended genetic stutter, hundreds to thousands of nucleotides long (1). Satellite regions in DNA are apparently functionless and have long been dismissed as "junk" DNA.

Although nature has no readily apparent purpose for satellite DNA, humans have been happy to put these regions to good use: variation between individuals in the lengths of their DNA satellites is the basis for one of the most revolutionary techniques in molecular biology: DNA fingerprinting.

Astoundingly, despite its ubiquity in today's courts and on shows like Law and Order and CSI, DNA fingerprinting was only developed two decades ago (2). In the short time since its invention, this technology has not only transformed criminal forensics and judicial systems around the world, but has also revolutionized fields of biology ranging from evolutionary studies to cancer detection and treatment.

DNA fingerprinting has provided critical evidence in thousands of criminal cases and exonerations (3). In a more familiar example, it was also used to identify another killer, the lethal tainted spinach in the recent E. coli outbreaks, and helped trace the outbreak back to its source in California (4).

Amazingly enough, DNA fingerprinting technology will soon enable doctors to detect as few as 10 cancerous cells in an entire lymph node (5), and will allow them to determine whether a woman who had breast cancer is at risk for relapse (6). In evolutionary studies, DNA fingerprinting has even traced the origin of all of mankind to Africa! (7).

"DNA fingerprinting has not only transformed criminal forensics and judicial systems around the world, but has also revolutionized fields of biology ranging from evolutionary studies to cancer detection and treatment."

Micro and minisatellites galore!

DNA satellites were first identified by biologists in the 1960s, and come in two types—microsatellites and minisatellites (1).

Microsatellites, also known as simple sequence repeats (SSRs), are made of repeating units of nucleotides, each repeat ranging from two to six base pairs long. On the other hand, in a minisatellite, the repeating units range from ten to 100 base pairs. These are also known as variable number tandem repeats, or VNTRs (8).

The characteristic of micro and minisatellites that makes them useful for identification purposes is that they are highly polymorphic – they can vary significantly in length from person to person. The length of each satellite in your DNA is inherited. Just as with any other gene on an autosomal chromosome, you have two alleles: one inherited from your mother and one from your father (2). (see fig. 1)

Variations in satellite length are the result of fortuitous mistakes in the process of crossing-over during meiosis. Because of the repetitive nature of micro- and minisatellite regions, occasionally when the chromosomes undergo crossing over, they do not align properly. When the strands of DNA pair, instead of lining up so the crossover occurs at the same place in the satellite on both chromosomes, they can slip and cross over at the wrong tandem repeat. This leads to a deletion of a repeat on one chromosome and an insertion of a repeat on the other. This generates two new alleles of the satellite, one longer than either of the original alleles, and one shorter (1).

Thus, the lengths of satellite regions are highly variable between people. This wealth of variation across the population makes satellites invaluable for identifica-

tion purposes. The probability that two random people have the same number of repeats at each and every one of multiple satellites is infinitesimally small, nearing one in 5 x 1019 (2). This probability is so low that in forensic analysis, a DNA fingerprint allows near certain positive identification of the source of a sample. The heritability of satellite alleles has also allowed the development of DNA fingerprinting-based paternity and maternity testing (2).

Restriction enzyme-based fingerprinting

Jeffreys' original procedure, restriction fragment length based DNA fingerprinting (RFLP analysis), is relatively straightforward: a restriction enzyme is used to cut the DNA at specific points, yielding fragments of varying lengths. These cleaved pieces of DNA are then subjected to gel electrophoresis to separate them based on size, and finally visualized by Southern blotting (2).

When genomic DNA is cleaved with particular restriction enzymes, the resulting fragments can contain a minisatellite (VNTR) region. Since the number of repeats in the satellite determines its length, and gel electrophoresis separates the digested DNA into bands based on the length of the fragments, the pattern of bands reflects the number of repeats in a minisatellite. When enough loci are examined, each person's DNA will yield a unique banding pattern—his or her unique DNA fingerprint (2).

PCR-based fingerprinting

In 1983, Kary Mullis developed a technique that would revolutionize molecular biology: the polymerase chain reaction, better known as PCR. This gave scientists an in vitro way to exponentially amplify DNA (9).

The general procedure and concept behind PCR is a simple one. First, the genomic DNA is denatured, its strands separated by heating. Then a primer--a synthetic single of stranded piece DNA--binds to the complementary portion of the genomic DNA. The portion of the DNA after primer is then copied by a molecule of DNA polymerase in a third step called extension. Multiple cycles of these three steps allow for exponential amplification of the region copied by the primers (9).

Primers can be designed to amplify regions of DNA containing satellites. If the PCR products are then separated by gel electrophoresis and a Southern Blot performed to image the DNA bands,

the resulting banding pattern reflects the lengths of the satellite regions. The fingerprint created is just like the fingerprint generated by restriction enzyme analysis (10).

While the first DNA fingerprints were made using the restriction enzyme method, the PCR-based method became dominant soon after its development in 1988. The important advantage to PCR-based DNA fingerprinting is that by virtue of the amplification process, it is possible to begin with only a tiny amount of DNA. While RFLP analysis requires a sample of blood at least the size of a quarter, it is possible to fingerprint the DNA of a single cell using PCR-based DNA fingerprinting (10).

It is also possible to analyze old or degraded samples using PCR-based fingerprinting. Because the microsatellites examined in PCR-based fingerprinting are significantly shorter than the minisatellites used in restriction enzyme based analysis, they are much less likely to be degraded. Most of the time, the DNA in blood or semen from a crime scene has been significantly degraded by time and exposure to the elements and would be nearly impossible to fingerprint with restriction analysis (12).

Breast cancer detection

Every year, the lives of hundreds of women are destroyed by rapists and murderers. DNA fingerprinting has helped identify and imprison untold numbers of such violent offenders. But, there is another serial killer of women, far more insidious than any of these criminals. It is one who has taken hundreds of thousands of lives, and who nearly every woman lives in fear of—breast cancer. But soon, DNA fingerprinting will help put a stop to this killer too. DNA fingerprinting will be used to identify renegade cancerous cells (5), and has been shown to be useful in assessing a cancer survivor's likelihood of relapse (6).

All of the hallmarks of malignancy in a cancer cell, such as uncontrolled cell division, secretion of angiogenic factors, and the ability to infiltrate other tissues and metastasize are the result of genetic mutations. Such genetic mutations also result in changes in the DNA fingerprint of that cell—a cancerous cell has a different DNA fingerprint than the rest of the healthy cells in the body.

Because of the amplification of DNA during PCR-based DNA fingerprinting, it is possible to detect the presence of less than 10 cancerous cells in an entire lymph node. This incredible sensitivity will allow doctors to determine whether tissue is cancerous long before a tumor would be visible by simply fingerprinting a sample of it. Such early detection will undoubtedly save countless lives.

Doctors will also be able to see if a cancer has metastasized, or check whether or not they have removed all of a tumor by fingerprinting biopsy

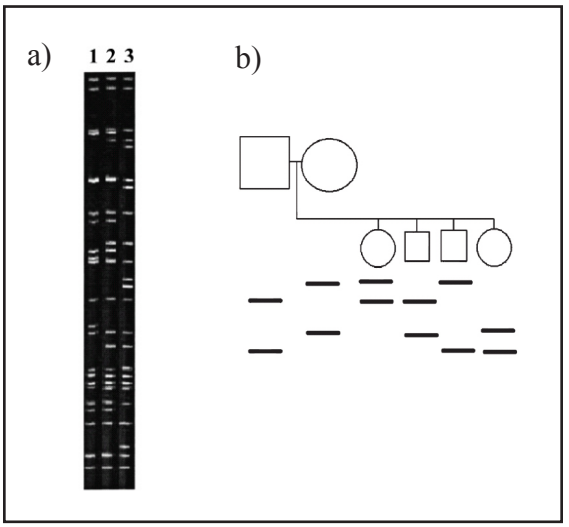

Figure 1. a) A DNA fingerprint from a paternity test, using several satellites. Lane 1 is the mother, 2 is the child, and 3 is the father (Varsha). b) DNA fingerprint for each member of a family, using a single satellite. Each person has two copies of the satellite. Larger DNA fragments with more repeats are closer to the top of the gel. In both a) and b), each of the children's bands is inherited from their parents.

samples. Although the majority of research in cancer cell detection has been done with breast cancer, it is possible to adapt this technology for the detection of other types of cancer (5).

In tumors, cancerous cells divide at an astounding rate. However, in many cases they also die rapidly. When they die they often release their contents, including their DNA, into the surrounding tissue. Studies have shown that with many breast cancers, there is a large amount of free DNA floating in the blood plasma that probably comes from the cancerous cells. Scientists have shown that the higher the concentration of tumor DNA in the blood, the more likely the patient is to relapse. Such a measuring stick for the probability of relapse will undoubtedly help oncologists save lives, by alerting them that a patient is likely to become sick again (6).

Other biological applications

Humans are not the only creatures who can be fingerprinted—anything with a DNA or RNA genome, including animals, plants, bacteria, and viruses can be fingerprinted the same way. DNA fingerprinting thus has an incredible number of biological applications. These applications range from monitoring endangered species to tracking down the source of tainted food.

For example, DNA fingerprinting of samples of tissue left behind by animals has allowed scientists to track the size and genetic diversity within populations of endangered species. Fingerprinting is used to help them determine the number of unique individuals in an area and ascertain how closely related they are (8).

As DNA fingerprinting can be used in humans to match a DNA sample with its source, it can also be used to trace the source of finished meat products back to specific animals. This can be very useful in outbreaks of diseases such as mad cow disease, since other animals that could be infected need to be destroyed, and contaminated meat products need to be recalled to prevent further spread of the disease. Similarly, fingerprinting can also be used to trace plants and plant products back to their sources (8).

DNA fingerprinting has also made significant contributions to the study of human evolution and migration. In 2003, scientists conducted an analysis of 377 satellites in each of over a thousand people, hailing from every corner of the globe. The amount of diversity between and within populations was then analyzed, and this statistical analysis was used to establish definitive support for the "out-of-Africa theory of human evolution and migration (7).

It is also possible to fingerprint bacteria and viruses. Such analyses are extremely important, and have widespread applications. For example, during the spinach recalls of this past fall, DNA fingerprinting of the bacteria on spinach leaves was used to determine which packages of spinach were contaminated with the deadly E. Coli O157:H7 bacteria. This information was eventually used to trace the source of the infection back to a handful of farms in California (4). Fingerprinting has also been used to identify viruses of all types that infect humans, including DNA viruses, RNA viruses, and even retroviruses (14).

As is evidenced by this multitude of applications, although DNA fingerprinting is only two decades old, in the short time since its birth it has revolutionized nearly every facet of biology (not to mention the justice system). Fingerprinting has given us a method to establish paternity, and a method to identify nearly any living thing. DNA fingerprinting has already saved untold numbers of lives by both putting away violent criminals and by helping identify deadly bacteria and viruses. Soon it will save even more lives by helping to cure cancer.

Crime and punishment

In 1990, the FBI created a national DNA databank known as CODIS. This extremely valuable database contains the DNA fingerprints of convicted sex offenders and violent criminals, as well as DNA profiles from evidence taken from crime scenes. CODIS is responsible for having helped bring thousands of criminals to justice. In fact, in over 30,000 cases, matching of crime scene evidence to DNA profiles of known criminals and inmates has led the police to a suspect. To take these fingerprints, the FBI uses a standard panel of 13 microsatellites, which together yield a remarkably miniscule 1 in 575 trillion chance of a random match (12).

The most common use of DNA evidence in courts is to help positively identify a rapist or murderer. This is done by comparing the DNA fingerprint of a sample of blood or semen left at the crime scene to that of the suspect. Fingerprints can also be made using DNA in urine, saliva, and any number of other bodily sources. In paternity testing, the DNA fingerprints of the child, the mother, and the father are compared. Paternity is the determined

by assessing whether it would be possible for the child to have inherited the combination of bands in their DNA fingerprint from the alleged father and mother.

With DNA fingerprinting, investigators have been able to re-examine evidence from old cases, leading to the arrests of violent killers and rapists who would have otherwise gone free. However, re-examination of evidence has also exonerated dozens of unjustly imprisoned people, and even saved innocent people from being executed. Project Innocence, a non-profit legal clinic, has freed 153 innocent men on the basis of post-conviction DNA testing (3). In Illinois in 2003, thirteen men on death-row were exonerated on the basis of DNA evidence. DNA fingerprinting saved thirteen innocent men who would have otherwise been executed for crimes they did not commit (12).

Genetic discrimination

While the 13 markers used by the FBI are useful for matching a DNA sample to its source, these loci are otherwise uninformative, providing no information as to any physical characteristics of the person (11). In 1997, the National DNA Advisory Board deliberately chose not to include markers associated with genes that denote ancestral origin or genetic diseases (13). However, there are many loci that could be used to provide information about a person's hair color, eye color, skin color, and even ethnicity.

Even the possibility of using such loci gives rise to a number of intractable ethical questions primarily concerning genetic discrimination based on ethnicity and hereditary predisposition to disease (11). It is incredibly difficult to strike a balance between using all information available to put violent killers behind bars, and respecting our individual right to privacy. The possibility of using profiling to determine a person's ethnicity brings up a number of questions about racial profiling, an issue that is particularly pertinent in today's post-9/11 world.

Even five years later, efforts are still underway to recover the remains of all 2749 victims of the terrorist attacks of September 11th, 40% of whose remains have not been identified. Sadly, in most cases, these remains come in the form of tiny fragments of shattered bone; it is often impossible to separate the remains of one person from those of another. DNA fingerprinting has been the most important tool in this effort, and has been used to reunite families with the remains of their loved ones (13).

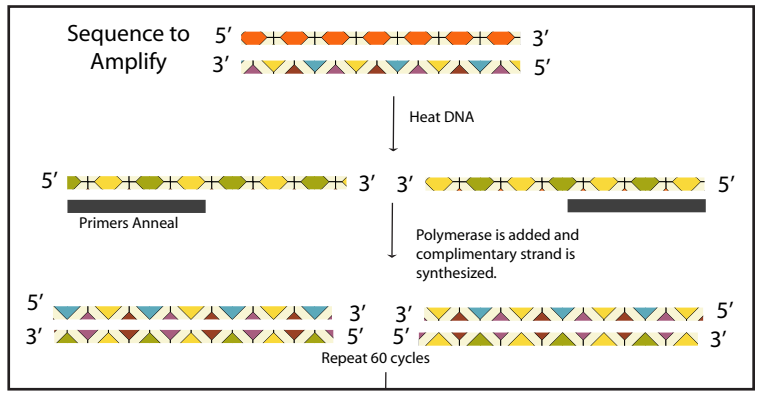

Figure 2. PCR-based DNA fingerprinting uses PCR to amplify satellite regions of DNA. DNA strands are separated by heating, and complementary primers then bind to the DNA. DNA polymerase replicates the satellite. Multiple cycles of this lead to exponential amplification of the satellite. The PCR products are then run out on a gel to yield the DNA fingerprint.

In one case, remains determined to belong to at least two people were recovered from United Airlines Flight 93. Since only one of those people could be identified, the remains were returned to that man's family. However, the man's family refused to bury the remains with the rest of his body parts if any of the intermingled unidentified remains might belong to a terrorist. Subsequent DNA fingerprinting of satellites linked to genes indicating ethnicity was used to determine with 95% certainty that the tissue did not belong to a person of Middle Eastern descent (13).

DNA fingerprinting provides us with the power to make such phenotypic determinations. The question that we are now faced with is should we use that power, or is the risk of abuse too great?

Nonetheless, it is hard to imagine that there was a time when DNA fingerprinting did not exist. Twenty years ago, the only satellites of any importance in our lives were the ones far above us, beaming information back and forth to us on Earth. Next time you look up at the stars, and Earth's satellite, the moon, be glad that we realized the importance of the immense amount of information contained in our DNA satellites. H

—Amrita Goyal '09 is a Chemistry concentrator in Lowell House.

References:

1. Moxon, E. Richard. "DNA Microsatellites: Agents of Evolution?" *Sci. Am.* Jan 1999.
2. Gill, P., Jeffreys, A., and Werrett, D. "Forensic Applications of DNA Fingerprints." *Nature*. 318 (1985): 577-579.
3. Broeders, APA. "Of Earprints, fingerprints, scent dogs, cot death, and cognitive contamination." *Forensic Sci. Int'l*, 159 (2006) 148-157.
4. FDA. "Nationwide E. Coli O157:H7 Outbreak: Questions and Answers." US Food and Drug Admin. <http://www.cfsan.fda.gov/~dms/spinacqa.html#collect >. Sept 25, 2006. Accessed Sept 29, 2006.
5. Ellenhorn, J., Gall, K. "DNA Fingerprinting and Breast Cancer." *Ann. of Surgical Oncology*. 11.5 (2004): 551-552.
6. Silva, J., et al. "Tumor DNA in Plasma at Diagnosis of Breast Cancer Patients." *Clin. Cancer Research*. 8 (2002):3761-3766.
7. Zhivotovsky, L., et al. "Features of Evolution and Expansion of Modern Humans." *Am. J. of Human. Gen.* 72 (2003): 1171-1186.
8. Varsha. DNA Fingerprinting in the Criminal Justice System." *DNA and Cell Biology*. 25 (2006): 181-188.
9. Mullis, K.B. "Target amplification for DNA analysis." *Annales de biol. clin.* 48.8 (1990): 579-82.
10. Jeffreys, A., et al. "Amplification of human minisatellites." *Nucleic Acids Research*. 23 (1988): 10953-10971.
11. Benecke, M. "Coding or non-coding, that is the question." EMBO Reports. 3 (2002): 498-501.
12. Kelly, J. "DNA Takes the Stand." *Invention and Technology*. Fall 2006: 44-56.
13. Snyder-Sachs. "DNA and a New Kind of Racial Profiling." *Popular Science*. Dec 2003: 16-20.
14. Arens, M. "Methods for Subtyping and Molecular Comparison of Human Viral Genomes." *Clin. Microbiol. Rev*. 12(1999), 612-626.