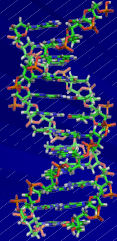


DISMANTLING THE MARYLAND DNA CONVICTED OFFENDER DATABASE



Sharon B. Benzil
Office of the Attorney General of the State of Maryland
Deputy Counsel, Maryland State Police

The Science

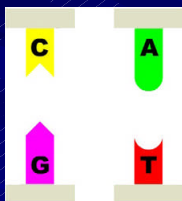


- DNA is the genetic material present in the nucleus of cells which is inherited half from each biological parent. An individual's DNA is unique except for identical twins (who have different fingerprints)
- DNA is sometimes called a genetic blueprint because it contains all of the instructions that determine an individual's genetic characteristics
- A person's DNA is the same in all cells in the body (except red blood cells which have no DNA)

Everyone inherits one copy of DNA from their mother & one from their father



DNA Structure



- DNA is comprised of 4 building blocks called bases
- The building blocks are: Cytosine, Guanine, Thymine, & Adenine, commonly referred to as C G T A
- It is the order (sequence) of these blocks that determines a person's genetic characteristics
- The 4 letter DNA alphabet always follow certain rules: C always bond with G; T only bonds with A
- These are called base pairings

Base pairing

- All of the DNA in a cell is known as the genome
- The human genome has about 3 billion base pairs
- Because there can be trillions of base pair combinations, scientists can reliably report on DNA analyses
- It is the specific sequence of these base pairings that determines each person's genetic individuality

Polymorphism

- 99.9% of all DNA is the same among human beings
- Scientists use a small amount of the remaining .1% for testing because of the high amount of variability in that remaining amount (called polymorphism)

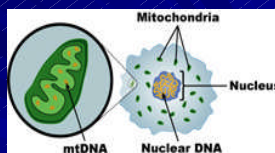
Interesting Nuclear DNA Facts

- My nuclear DNA is 99 % identical to your nuclear DNA...yet we are so different
- Human nuclear DNA is 98 % identical to chimpanzee DNA



Types of DNA

- Nuclear DNA is located in the nucleus of the cell
- Mitochondrial DNA (mtDNA) is found in cell cytoplasm
- Both are used in forensic DNA identity testing



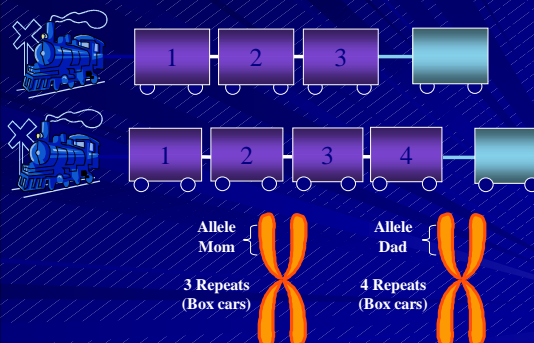
"Junk" DNA

- About 95% of DNA is noncoding, which means that does not code for any physical characteristics
- The "junk" DNA is valuable for identity testing but does not reveal any physical or genetic attributes of the person whose DNA is being tested

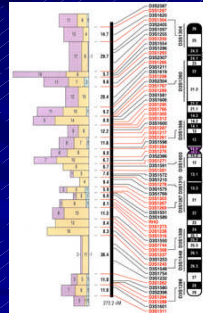
Basic Biology of Forensic DNA Testing

- "DNA marker" refers to a specific chromosomal location that is analyzed in the forensic DNA laboratory.
- The most widely used DNA markers are defined by their 'Short Tandem Repeat' (STRs) characteristics on the chromosome.
- Multiple types of STRs can be analyzed in one test, or multiplexed, thus making the analysis process faster than previous technologies.
- Multiplexed STRs are very valuable because they can produce results that are highly reliable for identification, even with old or minute biological samples.

STR's

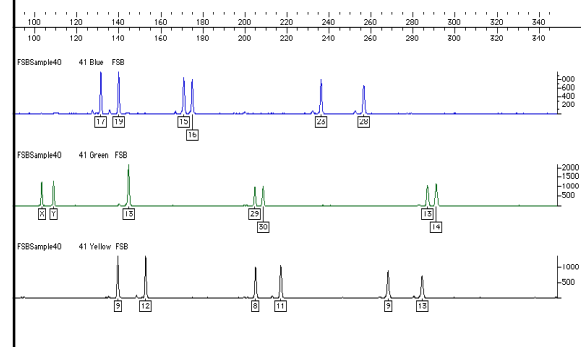


- Research was conducted to select DNA markers from the polymorphic (variable) noncoding regions ("junk DNA") of the human genome for the purposes of forensic identity testing.



- In the United States, the National DNA Database (CODIS) requires the use of 13 STR markers

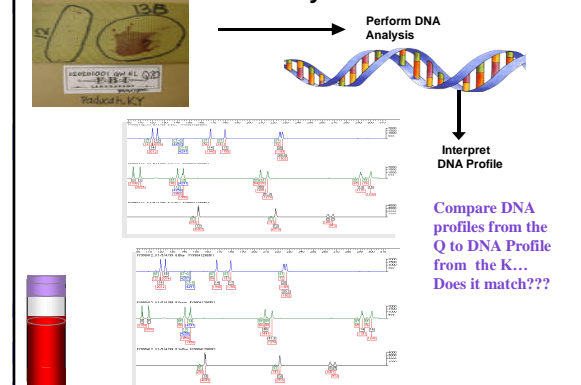
STR Profile



Forensic DNA Profile

- A forensic DNA profile is the combination of individual genotypes for all of the DNA markers or loci that have been analyzed.
- For forensic identity testing, a DNA profile is compared to other DNA profiles from biological samples such as crime scene evidence or samples from known individuals.

DNA Analysis



Alleles

- At a particular location on the DNA strand, a person carries a maximum of 2 alleles, one of which is contributed by the person's mother and the other is contributed by the father.
- When the allele contributed by both mother and father are the same, the profile shows only one numbered allele. When analyzed, the DNA profile report shows, for each of the identified loci, 1 or 2 numbers representing the alleles at that location.

DNA DATABASES

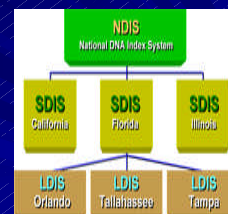
CODIS

- In 1990, the Federal Bureau of Investigation (FBI) Laboratory began a pilot project called the Combined DNA Index System (CODIS) creating software that enables Federal, State, and local laboratories to exchange and compare DNA profiles electronically.
- CODIS uses two main indexes: The Forensic Index and the Offender Index
 - The Forensic Index contains DNA profiles from crime scene evidence.
 - The Offender Index contains DNA profiles of individuals who have been convicted of various offenses defined by State and/or Federal law.

CODIS Systems

- The CODIS system operates on three levels:

- National DNA Index System (NDIS)
- State DNA Index System (SDIS)
- Local DNA Index System (LDIS)



Comparing Genetic Profiles



- Forensic DNA analysts compare the genetic profile obtained from crime scene evidence to the profile from a known individual (e.g., suspect, victim).
- If the DNA profiles from the evidentiary and known samples are the same at each locus, laboratory analysts can provide a determination of the statistical significance of the evidence. In some cases, no conclusive interpretation can be made.

Statistical Interpretation of the Evidence

- In a given population, any particular 13-locus short tandem repeat (STR) profile is rare because there are many distinct alleles at each of these STR loci. Therefore, it is possible to distinguish between individuals.
- Practical issues (e.g., small amounts of DNA) can limit the ability to obtain results; additionally, not all laboratories routinely type the 13 STR loci required for the national DNA database. Nonetheless, it is still possible to distinguish between individuals.
- CODIS matches require all 13 loci to match at high stringency

Statistical Interpretation of the Evidence

- Many courts require statistical interpretation of a DNA match. **Random match probabilities** are most often used to interpret evidence from single source samples.
- Generally, once a DNA match is determined, a statistical computation is performed to estimate how often a random *unrelated* person would be found with that particular DNA profile.
- Once a DNA match is observed, forensic scientists estimate the chance of finding that DNA profile in particular human populations. This calculation is necessary to inform the jury of the rarity of the profile.

Probability and Population Databases

- Once an individual's 13-locus STR profile is identified, it is statistically improbable that anyone else in the world will have the same profile, unless that person has an identical twin.
- Population databases of major racial and ethnic groups are used to determine estimates of the rarity of DNA profiles. These databases sometimes consist of as few as 100 profiles from unrelated persons, yet allow a reliable estimate of the chance of observing a given DNA profile in a larger population.

The History of the Attack: The Arizona Search

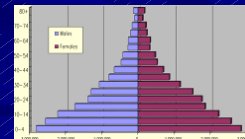


- 2001: Search of the Arizona convicted offender database yielded matches at 9 loci between 2 apparently unrelated individuals, 1 a Caucasian and 1 an African American
- Later, Arizona determined that there were 122 pairs of 9-loci matches, and 20 pairs of 10-loci matches in a 66,000 profile database

- Random Match Probability of this match was 1 in 754 million Caucasians and 1 in 531 million African Americans
- The Arizona database, at the time, contained 22,000 profiles
- Population of the US in 2001 was 261 +/- million

- The defense bar argued that this 9 loci match should not have happened in a database of only 22,000 profiles since it should have more rare than the entire population of the United States!
- Therefore, they argued that the method of calculated the random match probability and perhaps the entire foundation of the databases and "hits" were not sound

"Such a dramatic difference between the theoretical model and the empirical data cannot be reconciled and demands further discovery as to the statistical foundation" of any "hit" (match) comparing an evidentiary sample and a profile from the convicted offender database



The OPD Attack

- Capital murder case involving a local but not Maryland (State) convicted offender database
- 13 loci match in the LDIS of only 1000 convicted offender profiles: offender to various pieces of evidence (hats, steering wheel, and more)

OPD Discovery Motion

- "With the requested discovery, the Defendant may be able to demonstrate that the State's methods of preparing statistics is flawed and does not reflect the actual statistical significance of a "match"
- Requested complete disclosure of the Maryland DNA database

Additional Discovery

- Later, the OPD requested that the Maryland Crime Lab to do a "pair-wise" search of the convicted offender database to report on matches at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, and 13 loci

Maryland's Response



- Collaborated with Illinois and California, using materials previously filed in those jurisdictions
- Responded with affidavits, arguments and testimony

Response in 3 Parts: Legal, Logistic and Scientific

- **Legal**
 - No Maryland database employed
 - Discovery request not authorized by *any* law or rule of court
 - No Due Process or Confrontation Clause authority for this request
 - Violation of State-CODIS MOU and software user agreement

Logistics

- No software parameters in system; Feds would need to help
 - Estimates of time to produce
 - 5 – 10 min per search
 - Manual evaluation of each potential ("candidate") match
 - Research for duplicates, etc. in the database
- Conclusion: Months to Produce – stopping entire database system and all lab staff for the duration

January	February	March
April	May	June
July	August	September
October	November	December

Logistical Nightmare

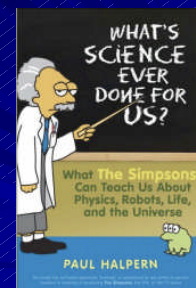


To compare every profile to every other profile to see which pairs match would mean 312,000,000 pair-wise comparisons (every profile would match multiple other profiles at 1 loci, many at 2, somewhat fewer at 3, and so on)

Conclusion: Months to Produce – stopping entire database system and all lab staff for the duration of the search

Scientific Refutation

- Samuel Baechtel
 - One cannot be misled by an inappropriate comparison of the frequency by which one would expect to find a specified profile in a convicted offender database with the frequency by which one would expect to find a specified profile in the general population
 - Apples to oranges comparison



2 DISTINCTLY DIFFERENT DATABASES: Offender versus Population

- Offender Databases *cannot* and *are not* used to estimate how frequently a particular DNA profile is likely to occur
 - A convicted offender database is not a random database
 - It is not cleansed of duplicate samples or relatives
 - In other words, it is not a database established and recognized as representative of the random population

Offender versus Population Databases



- The frequency of similarities between profiles in an offender database can be explained by:
 - Close relatives (population databases do not contain)
 - Duplicate samples (twice collected, aliases)
 - Inbred or insular population groups which might not bear on the general population

NOT ALL "MATCHES" ARE EQUAL



- The frequency of 9 or 10 loci matches are like matching 6 of 8 lottery numbers or 7 of 8 numbers and letters on a license plate
- If it's not a 13 loci match, similarities are interesting but not considered CODIS matches

Statistics and Lies

- The fact that there may be pairs of entries in a database that coincidentally match at a given number of loci does nothing to discredit the rareness of any given profile
- The random match probability statistic indicates the probability that a person randomly selected from a given population will possess a given DNA profile



The Result

- Court initially ordered State database administrator to perform pair-wise comparison of all profiles AND to report which profiles matched at 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 and 13 loci

CODIS (lack of) Assistance

- FBI decided that they would not assist Maryland administrator
- Maryland, left without support and resources, reported to Court
- Court threatened contempt
- Maryland prepared for an interlocutory appeal

The Court's Compromise



- Ultimately, the Court modified its Order to permit a production of the same information produced in the "Arizona" search
- Maryland relented and produced the report of matches at 9, 10, 11, 12 and 13 loci matches

THANKS TO California, Illinois, Georgia



THE FUTURE

- Familial (Partial) Searches
 - Partial results to be used as investigative leads



Additional Resources

- American Prosecutors Research Institute
 - www.apri.org
- President's DNA Initiative
 - www.dna.gov
- Federal Bureau of Investigation
- National Commission on the Future of DNA Evidence
 - Office of Justice Programs
 - National Institute of Justice
 - 810 Seventh Street N.W.
 - Seventh Floor
 - Washington, DC 20531
 - 202-307-0645
 - www.ojp.usdoj.gov/nij/topics/forensics/dna/commission/