

Hidden DNA Evidence: Exonerating the Innocent

Mark W. Perlin, Ph.D., M.D.

In 1989, northwest Indiana was plagued by bump-and-rob road crimes of escalating violence. In the darkness of a cold December night, a woman's car was rear ended on highway I-65. Upon exiting her car, she was dragged into another vehicle, then stripped and raped by five strangers. The men left her in her car, draped by green coveralls. The same night, coworkers Darryl Pinkins, Roosevelt Glenn and William Durden had engine trouble along that highway. They parked their car on the roadside, and went to get help and motor oil. On their return, they found shattered side windows and their work coveralls gone.

Traced to their employer by the crime scene coveralls, Pinkins, Glenn and Durden, along with two other coworkers, were arrested for the I-65 bump-and-rape. RFLP testing of semen DNA left on the victim's jacket and sweater excluded the defendants. But nonspecific serology testing, along with faulty hair evidence and tainted eye witness identification, led to Pinkins' and Glenn's wrongful convictions. Pinkins was found guilty of all charges in May 1991, and sentenced to 65 years in prison. Glenn's 1992 jury deadlocked, but on retrial he was convicted of rape in 1993, and sentenced to 36 years. Despite their incarceration, the bump-and-rob and rape crimes continued unabated. The men's exoneration by science would not happen soon.

DNA science

In the early 1990s, the Human Genome Project was in full gear. A key genetic marker was the short tandem repeat (STR). At an STR genetic location (locus) on a chromosome, naturally occurring variation in the number of repeated words (allele) could track inherited DNA or distinguish different people.






	V	P	G	
 J	-15.39	-18.00	-15.39	<i>Evidence (rows)</i>
 S	-15.39	-15.17	-15.39	Jacket
 H	-15.39	-18.00	-15.39	Sweater
 JJ	-2.78	-8.50	-8.43	Hair
 JS	-10.47	-5.79	-12.60	Jacket-Jacket
				Jacket-Sweater
				<i>Reference (columns)</i>
				Victim
				Pinkins
				Glenn

Figure 1: Match statistics between evidence genotypes (rows) and reference individuals (columns). The entries give the base ten logarithm of the match statistic. For example, +6 log units would represent an inclusionary million (six zeros after the 1) statistic, whereas -6 log units would represent an exclusionary one in a million.

This allelic repeat variation was easily measured on a DNA sequencer—more repeats gave longer sequences, and more allele DNA molecules fluoresced more brightly. STR data have a “stutter” artifact, where an allele produces shorter shadow sequences. The stutter problem was solved mathematically by removing these DNA shadows. This genotyping automation was immediately applicable to genetic diagnosis, gene discovery and human identification. Computers could interpret STR data faster and more accurately than human review. In fact, automated DNA review software eliminated Great Britain's national DNA database backlog of 350,000 offenders. Each cheek swab contained abundant DNA from one person. But



biological evidence is usually a mixture of two or more people. On this more complex data, human review often fails to identify or exclude suspects from DNA mixtures.

A person's genotype at a locus is the pair of alleles they inherit from their parents. A mixture adds together genotype allele pairs from different people. STR mixture data shows allele peaks of different heights that mirror this genotype addition. Computers can mathematically separate the data signal (up to probability) into each contributor's genotype.

Comparing a separated evidence genotype with another genotype, relative to a population, produces a match statistic. Big match numbers (e.g., a million) indicate a DNA association. Small numbers (e.g., one in a million) reflect no association.

DNA analysis

Frances Watson of the Indiana University McKinney Law School, and director of the school's Wrongful Conviction Clinic, took the Pinkins and Glenn case.

In 2001, the Cellmark laboratory developed STR data from the jacket and sweater semen stains. Differential DNA extraction separated the sperm fractions, and major 80% and 90% male contributors were derived from the items. Those two different genotypes did not match Pinkins or Glenn, establishing the presence of two unknown assailants at the rape scene.

This new DNA evidence did not convince the Indiana courts. "2+3=5" went their legal argument: two unknown genotypes (jacket and sweater) plus three accused (Pinkins, Glenn and Durden) equals five attackers on highway I-65. Not exculpatory. Pinkins and Glenn remained in prison.

Seven years passed. The appeals court ruled that nonspecific serology and jailhouse testimony trumped DNA. After 16 years of exemplary behavior, protected by prison gangs impressed with character and innocence, Glenn was released from prison—still tarred a convicted rapist. Pinkins remained behind bars, serving out his 65 year sentence.

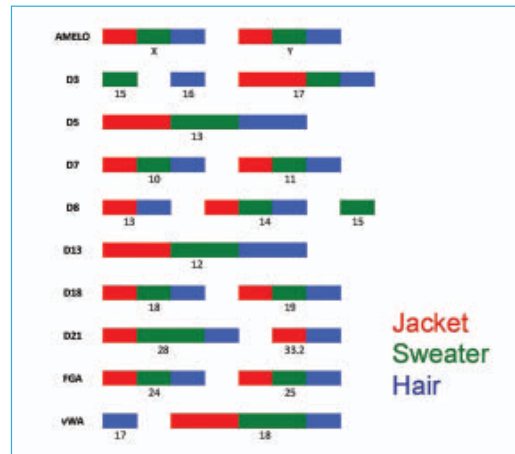


Figure 2: The three genotypes from the jacket (red), sweater (green) and hair (blue) DNA evidence items. Each row is for a different genetic locus. The numbers designate alleles.

Another seven years passed. Greg Hampikian of Boise State University, and director of the Idaho Innocence Project, had testified during the appeals about the limitations of serology. He predicted that better DNA mixture analysis would arrive one day to free these men. He reached out to the CBS 48 Hours news team, gaining their interest.¹

Automated DNA software (TrueAllele) separated the DNA mixture data into genotypes. Two easy results concurred with Cellmark—there were distinct major male genotypes from the jacket and sweater. The computer went on to find exclusionary DNA match statistics (Figure 1), proving that these two unknown genotypes were not from Pinkins, Glenn or the victim. A third nonmatching evidence genotype was found from a *hair* disclosed at Glenn's trial.

The computer's positive match statistics suggested similarity between the three unknown genotypes. Indeed, these genotypes show considerable allele overlap, and are identical at half the loci (Figure 2).

For kinship analysis, the computer inferred sibling genotypes, one for each evidence item. Comparing these siblings with the three unknown evidence genotypes gave match statistics in the thousand to million range. This meant that the



sweater, jacket and hair major contributors were most likely brothers. Yet defendants Pinkins, Glenn and Durden were not related.

Statistical computing can simultaneously assess evidence from different experiments. Examining duplicate STR amplifications from the jacket, joint computer analysis separated out a 10% minor contributor. Comparing this minor evidence genotype with Pinkins, Glenn and victim reference genotypes gave exclusionary match statistics (Figure 1). The jacket-jacket minor furnished a fourth unknown genotype, different from the accused, jacket, sweater and hair.

Jointly examining jacket and sweater mixture data, the computer found a fifth unknown genotype. Representing about 5% of the DNA samples, this minor jacket-sweater genotype was statistically different from the known references, as well as the four unknown evidence genotypes.

The computer derived far more information from Cellmark's data.² The lab's DNA data had laid dormant for 15 years. The machine's capability surpassed human review by:

1. Comparing evidence with evidence
2. Calculating exclusionary match statistics
3. Revealing 5% and 10% minor contributors
4. Jointly analyzing DNA mixture data
5. Showing three perpetrators were brothers

The more complete scientific argument reached the needed "5+0=5". Producing five unknown genotypes—jacket, sweater, hair, jacket-jacket and jacket-sweater—fully accounted for the five highway I-65 attackers. No defendant was linked to the crime. Would this better science have an impact on criminal justice?

DNA justice

In June 2015, armed with newly discovered evidence of five unknown genotypes, Watson filed a successor petition for Pinkins. Unusually, it was approved. But days before the scheduled April 2016 hearing, the county prosecutor conceded to the better science and released Pinkins. After 24 wrongful years in prison, Pinkins was exonerated and free. Based on the same scientific evidence, Glenn was exonerated in January 2017.

The three definite genotypes from the jacket, sweater and hair DNA have not yet been compared with the national CODIS database to possibly identify the real bump-and-rape criminals.

Incomplete mixture interpretation silences DNA evidence. For 15 years, existing DNA data awaited more informative analysis. Pinkins and Glenn were fortunate. Modern scientific computing, conducted for the public benefit outside a crime lab, awoke their exculpatory evidence.

Hundreds of thousands are not as lucky. Some defendants are denied effective assistance of DNA. Many languish in prison, some on death row, for crimes they did not commit.

DNA is power. Properly interpreted, it can find truth in evidence, and assist courts in delivering just verdicts. But misinterpreted DNA abuses that power. Pinkins' and Glenn's 40 wasted prison years teach us how true justice can benefit from scientific truth. ●

References

1. "Guilty Until Proven Innocent," *CBS News 48 Hours*, Feb 5, 2017. <https://www.cbsnews.com/videos/guilty-until-proven-innocent>
2. M.W. Perlin, F. Watson and G. Hampikian, "When DNA alone is not enough: exoneration by computer interpretation," *American Academy of Forensic Sciences 69th Annual Meeting*, New Orleans, LA, Feb 16, 2017.

Dr. Mark Perlin is chief scientist and executive at *Cybergenetics*, a Pittsburgh-based bioinformatics company that provides patented *TrueAllele* automation for accurate and objective interpretation of complex DNA evidence.