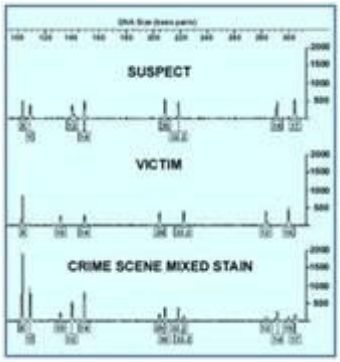
Posted 4/16/10

DNA: PROCEED WITH CAUTION

Subjectivity can affect the interpretation of mixed samples



> *"It's an irony that the technique that's been so useful in convicting the guilty*
> *and freeing the innocent may wind up leading to wrongful convictions in mixture cases, especially those with very low amounts of starting DNA."*

By Julius (Jay) Wachtel. Some might consider these words unduly alarmist. After all, no less an authority than the National Academy of Sciences has declared DNA to be *the* gold standard: "With the exception of nuclear DNA analysis...no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source."

Yet for years there have been troubling signs that interpreting mixed DNA – meaning DNA that's a blend from different contributors – isn't as straightforward as some forensic "experts" claim. Consider the case of John Puckett, who was mentioned in "Beat the Odds, Go to Jail," a post about random match probability, the likelihood that a particular DNA match could have happened by chance alone.

In 2003 a partial DNA profile from an unsolved, decades-old rape/murder was compared against the California DNA database. Although the biological specimen was badly degraded and had fewer than the minimum number of markers the state usually requires to call a "match," what was there was consistent with the DNA profile of Puckett, a convicted sex offender. Although nothing else connected him to the victim or the crime scene, Puckett was tried and convicted. Jurors said they were swayed by a prosecution expert who testified that the probability that the evidence DNA *wasn't* Puckett's was one in a million. It's since been suggested that the

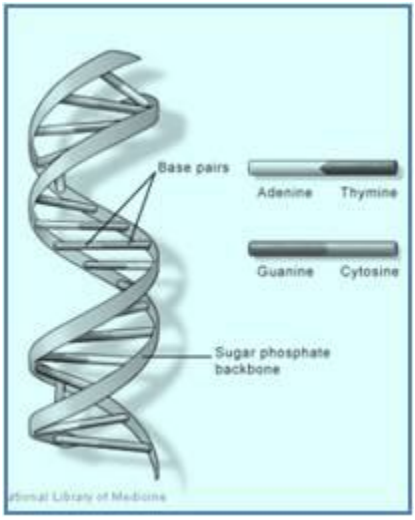
government's logic was faulty and that the true chance of a mismatch was actually *one in three*.

Since then the trustworthiness of the DNA processing has also come under attack. After sitting on a shelf for twenty-one years the biological sample was badly degraded, leaving only a tiny bit of DNA, and that being a mixture from both the victim and perpetrator. A growing chorus of scientists (and even police labs) warn that such factors can introduce dangerous uncertainties into DNA typing, making matching far more subjective that what one would expect.

But let's turn this over to a *real* expert. Greg Hampikian, Ph.D (the source of the introductory quote) is professor of biology at Boise State University and director of the Idaho Innocence Project. One of the nation's foremost authorities on forensic DNA, Dr. Hampikian jets around the globe giving advice and testimony and helping set up innocence projects. He graciously took the time from his busy schedule to give us a primer on DNA and the issues that attend to mixed samples.

An interview with Greg Hampikian, Ph.D.

What is DNA?

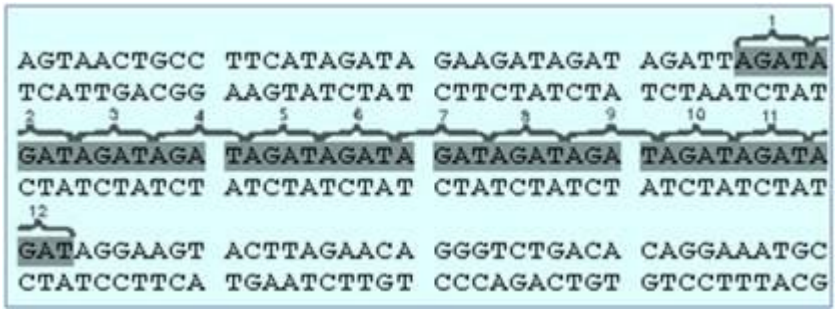


DNA is the repository of all hereditary information. It provides the recipes for all the proteins that can be made by an organism. A stringy acid, it's comprised of a chain of subunits or "bases," the chemicals Adenine, Guanine, Cytosine and Thymine. These are linked in pairs, with A only binding to T, and G to C.

Nuclear DNA, the kind most often used for identification, is found in twenty-three pairs of chromosomes – one inherited from each parent – that occupy the nucleus of every cell (except mature red blood cells). The complete set of nuclear chromosomes (all 23 pairs) is known as the "genome." Sperm and egg cells contain only one of each of the 23 chromosomes, and thus have half the DNA of other body cells.

Is the full genome used for identification?

No. A genome is comprised of millions of linked pairs, far too much information to process efficiently. And it's not necessary. Instead, identification relies on comparing repetitive sequences (for example AGAT in figure) which can be found at various chromosomal locations, or "loci." These "short tandem repeats," or STR's, can take various forms.



For example, at locus CSF1PO (in chromosome 5) it's always AGAT. Each locus actually has two STR sequences: one is the "allele" or gene variant contributed by the mother's chromosome, and one is the allele contributed by the father's chromosome. In this example, one of the two CSF1PO's alleles has twelve AGAT repeats. According to population studies alleles at CSF1PO can have between six and sixteen AGAT repeats.

Wouldn't there be many people who have the same number of repeats at this locus?

Locus	Sample	Suspect
CSF1PO	11	11
D13S317	12	12
D16S539	12,13	12,13
D18S51	10,20	10,20
D21S11	29,30	29,30
D3S1358	16,18	16,18
D5S818	12,13	12,13
D7S820	9,11	9,11
D8S1179	11,15	11,15
FGA	23,24	23,24
TH01	6,9.3	6,9.3
TPOX	8	8
vWA	17,19	17,19

Yes. Numerous persons have the same alleles at one or more loci. But when one compares alleles at thirteen loci, the number required under the FBI's CODIS system, the probability that a biological sample will be tied to an innocent person (the so-called "random match probability") is infinitesimally small, far less than one over the population of the Earth.

This example demonstrates a perfect match at each of the thirteen loci used by CODIS. Repeat sizes are reported for both alleles. (If both parents contributed the same number of repeats only a single number appears.)

So this suspect must be the source of the DNA sample.

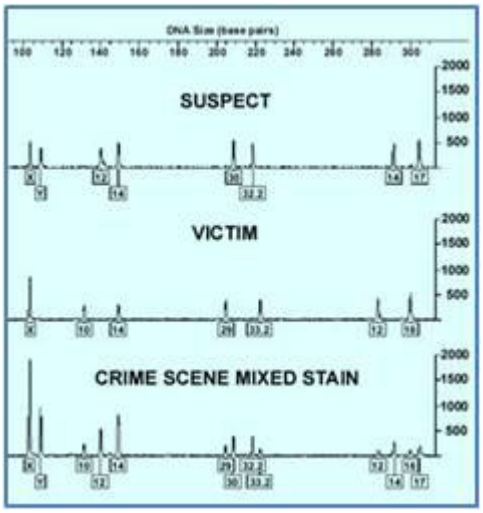
Yes, most likely, unless they have a twin. Analysts will testify that a match at thirteen loci establishes a positive identification. However, the statistics are less impressive when low amounts of DNA or degradation makes it impossible to type a biological sample at all thirteen loci. CODIS does accept DNA profiles from forensic

samples with as few as ten loci, which also yield high match probabilities, but are not unique. Some State systems may allow fewer.

Does subjectivity ever intrude into DNA identification?

It can. When evidence DNA is from a single source there is general agreement on computing random match probabilities. But interpretation is more difficult when samples are mixed; for example, a rape with multiple assailants. Mixed DNA is like mixing names made with scrabble tiles. For each person you add to the mix, the number of possible names you can pull out soars, so excluding anyone becomes problematic.

Mixture electropherograms, the charts used to detect alleles, can become crowded with peaks, making contributors extremely difficult to distinguish. We know from laboratory studies that an allele may sometimes be undetectable because one contributor's DNA is in a low concentration and a few alleles have "dropped out." Other times an allele may be obscured by someone else's peak. When two people touch an object, one profile might dominate while the other may be completely absent. These difficulties and differences in protocols can lead labs to vary a billion-fold when estimating mixed-sample match probabilities from the same data.



And there's another problem that becomes more of an issue with mixtures – the possibility of bias. Most labs train analysts not to look at suspect profiles before performing mixture analysis. However, since it's always easier to traverse a maze backwards, the goal of true blind testing is frequently violated. Analysts who have suspect DNA profiles on hand are susceptible to bias and could be less likely to exclude a suspect in a complicated mixture. Also, while most lab protocols require a second, independent analysis, the second analyst is often a close colleague who may have access to the first analyst's conclusions.

What suggestions do you have for the future?

There needs to be a lot more study and experimentation with mixed-sample DNA. There's no accepted standard for interpreting mixed samples, nor is there general agreement among experts as to when to exclude a suspect. Studies by independent researchers are also needed to help labs avoid bias, and enforcement of true independent analysis should be instituted. Defense lawyers and prosecutors are by and

large uninformed about these issues, and courts tend to leave it to jurors to work out any apparent contradictions. It's an irony that the technique that's been so useful in convicting the guilty and freeing the innocent may wind up leading to wrongful convictions in mixture cases, especially those with very low amounts of starting DNA.