

Epic FBI DNA database may reveal private health information

Just about any DNA-based investigation in law enforcement involves the U.S. Federal Bureau of Investigation's [Combined DNA Index System](#) (CODIS), a large and growing database of DNA that purportedly can identify criminals. However, new advances in understanding the genome (especially the non-coding parts) suggests that CODIS may be able to [identify health conditions](#), in us. That could pose privacy issues.

[Or maybe not](#). Forensics scientists, legal experts and even [U.S. Supreme Court justices](#) are debating whether or not the specific genetic markers used by CODIS to match crime scene DNA to an individual may overlap with certain epigenetic DNA that may play some regulatory role in the basis of a disease. So far, court cases have contained proclamation from one expert witness or another that these profiles could point to somebody's susceptibility to asthma, schizophrenia and more. The FBI, however, insists that no other personal information besides a match to a criminal's DNA can be gleaned from CODIS

CODIS was created in 1997, and consists of [13 markers](#) called [short tandem repeats](#) (STRs). These are very short sequences of data, usually no more than five base pairs, that are repeated many times in a genome. Variability comes from the number of times these STRs actually repeat. Multiply this variation by 13 different markers, and the chances of two 13-STR profiles being identical are almost nil. This is why the FBI, as well as law enforcement agencies worldwide, adopted this profiling system. The same markers are also used for paternity testing, as well as by other DNA-based identification organizations.

One person's junk, another's treasure

When the 13 CODIS markers were identified for law enforcement use, the human genome had not yet been sequenced, and, most important, the concept of "junk DNA" prevailed in scientific and judicial circles. Anything that didn't code for a protein was considered functionless, or "junk." Since then, [the ENCODE project](#) and other scientific advances have shown us that these strips of non-coding DNA are anything but useless. Parts of non-coding DNA have been shown to play regulatory roles for coding DNA, and have been implicated in cancer, inherited diseases, and even neurological disorders. But epigenetics, as the science is called, is still largely exploring a giant unknown. So, while some legal and forensics experts have claimed that the CODIS DNA data might be used to determine a person's risk for say, Alzheimer's disease, others say that these tandem repeats are too small and not an integral part of any known disease process.

While the CODIS database is expanding rapidly (it contains more than 10 million profiles of individual DNA), the FBI wants [to expand the number](#) of markers from 13 to 20. The current 13 markers are not sufficient to determine kinship or be workable in ever growing genetic databases, the FBI claims. However, some researchers observed that the same sensitivity could be achieved by using fewer markers on the Y chromosome, which reduces the risk of inadvertent genetic information, and eliminates the debate over which markers may be variable enough to be useful, and which may not.

Another area of expansion which has concerned attorneys and forensics experts is the types of crimes

reported to the database. When it was set up, the FBI's main concern was gathering the DNA of people convicted of serious felonies like murder or rape. Now, several states have passed laws that require police to send DNA from people who have been [convicted of misdemeanors](#). New York, in fact, requires that anybody convicted of anything—no matter how serious—has his or her DNA sent into CODIS. This means, that partying too loudly will get you more than a summons and a sentence—it'll get your DNA into the same database as a murderer.

These issues are not arcane scientific arguments. Forensics labs (including the FBI's) [have taken heat](#) over the last few years for shoddy laboratory procedures, grossly inaccurate testimony by law enforcement, and, in a few cases, outright [false documentation](#) of results. While genetics might be able to identify a felonious human, forensics scientists and lawyers agree that the information gathered can't be able to gather more than that. As the Supreme Court wrote in its [Maryland v King](#) decision to allow DNA collection, this issue is ["open to dispute."](#)

[Andrew Porterfield](#) is a writer, editor and communications consultant for academic institutions, companies and non-profits in the life sciences. He is based in Camarillo, California. Follow [@AMPorfield](#) on Twitter.