Anne C. Taylor, Public Defender
Kootenai County Public Defender
PO Box 9000
Coeur d'Alene, Idaho 83816
Phone: (208) 446-1700; Fax: (208) 446-1701
Bar Number: 5836
iCourt Email: pdfax@kcgov.us

Elisa G. Massoth, PLLC
Attorney at Law
P.O. Box 1003
Payette, Idaho 83661
Phone: 208-642-3797; Fax: 208-642-3799

Assigned Attorney:
Anne C. Taylor, Public Defender, Bar Number: 5836
Jay W. Logsdon, Chief Deputy Public Defender, Bar Number: 8759
Elisa G. Massoth, Attorney at Law, Bar Number: 5647

IN THE DISTRICT COURT OF THE SECOND JUDICIAL DISTRICT OF THE STATE OF IDAHO, IN AND FOR THE COUNTY OF LATAH

STATE OF IDAHO	CASE NUMBER CR29-22-2805
Plaintiff,	
V.	**NOTICE OF FILING DECLARATION OF BICKA BARLOW IN SUPPORT OF DEFENDANT'S THIRD MOTION TO COMPEL**
BRYAN C. KOHBERGER,	
Defendant.	

COMES NOW, Bryan C. Kohberger, by and through his attorney of record, Anne C. Taylor, Public Defender, and hereby files the attached Declaration of Bicka Barlow in support of the Defendant's Third Motion to Compel.

DATED this ___22___ day of June, 2023.

ANNE C. TAYLOR, PUBLIC DEFENDER
KOOTENAI COUNTY PUBLIC DEFENDER

BY: _____

ANNE TAYLOR
PUBLIC DEFENDER
ASSIGNED ATTORNEY

CERTIFICATE OF DELIVERY

I hereby certify that a true and correct copy of the foregoing was personally served as indicated below on the ___22___ day of June, 2023 addressed to:

Latah County Prosecuting Attorney –via Email: paservice@latahcountyid.gov
Elisa Massoth – via Email: legalassistant@kmrs.net
Ingrid Batey – via Email: ingrid.batey@ag.idaho.gov
Jeff Nye – via Email: jeff.nye@ag.idaho.gov

_____

DECLARATION OF BICKA BARLOW

I Bicka Barlow do state and declare:

I am an attorney licensed to practice law in the State of California and I have been retained to assist counsel for Bryan Kohberger in the above matter in the capacity of DNA consultant.

I have a B.S. from the University of California, Berkeley, in Genetics, and a M.S. from Cornell University in Developmental Genetics, with minors in Cellular Biology and Plant Molecular Biology. I received my J.D. from the University of San Francisco, School of Law. I began consulting on DNA cases as a law student in 1994. In 2004, I was hired as the DNA attorney for the San Francisco Public Defender's Office. During my nine years at the office, my practice was focused almost completely on DNA cases. In 2013, I left the office and returned to consulting on DNA cases throughout the United States in both state and federal courts and at all stages of litigation (See Attachment 1).

I have acted as DNA counsel or as a consultant on cases involving DNA evidence in over 200 cases, including capital cases, in both trial and post-conviction proceedings. Additionally, I have been retained in three cases as a "Strickland" expert to assess the performance of trial counsel as it related to DNA evidence. As DNA counsel I have reviewed case files for DNA evidence from many labs throughout California and the United States, including the California Department of Justice, SERI, Forensic Analytical, Bode, Orchid Cellmark, the FBI, the St. Charles, MO, Police Department Crime Lab, Contra Costa County Police Crime Lab, San Francisco Police Department Crime Lab,

Sacramento County District Attorney Crime Lab, Oakland Police Crime Lab, Santa Clara County District Attorney Crime Lab, the Arizona Department of Public Safety lab, and the Phoenix Police Department Lab.

I regularly communicate with attorneys and experts who specialize in DNA evidence. I also regularly train attorneys in the area of DNA evidence.

I was requested by counsel in this case to assist in determining the scope of DNA discovery necessary for defending this case, including any genetic genealogy that was used. Part of this assignment was to identify potential expert witness and legal issues that might arise from the use of genetic genealogy in particular. Because I have been retained on other cases involving genetic genealogy I am aware of the methods used and the type of material that is and has been made available in criminal cases. In order to assist counsel in this and other cases, I have reviewed a multitude of materials, including peer-reviewed articles, magazine and newspaper articles and motions submitted in other cases. I have also interviewed numerous expert witness in order to identify potential areas of interest to the defense.

I have reviewed the material provided to the Defense in this case regarding the DNA testing done as well as numerous news articles regarding this case and the government's Motion for Protective Order, filed on June 16, 2023.

I am informed and believe via news reports, that in this case Mr. Kohberger was initially identified as a possible suspect via a new methodology called investigative or forensic genetic genealogy (hereinafter IGG).

Genealogy Testing and Database Searches

I am familiar with the methods used to conduct IGG searches in criminal cases. I have been retained in other jurisdictions on cases involving this type of search and have received and reviewed reports generated by the private company that conducted both the testing itself and the construction of the family tree that lead to the identification of the defendant in those matters.

Based on my review of that material, I am informed and believe that the use of these databases does not necessarily lead to a single individual as a potential suspect. The testing conducted by the private lab is different in nature from the testing done by forensic labs. The private labs do not generate a "profile" in the same way that forensic labs do and there cannot be a direct comparison between the data obtained in this case by the Idaho State Police Crime (ISP) lab and the private lab. The ISP lab had conducted what is typically called STR based testing which results in a profile that can be uploaded to the state and federal CODIS databases. IGG labs use a different technology where individual SNPs (single nucleotide polymorphisms) are sequenced. The SNP and STR data cannot be directly compared.

Once the IGG lab completes its testing, the SNP data is uploaded to a genetic genealogy database that contains similar data from other individuals. One such database is GEDMatch (GEDMatch has recently been acquired by Verogen Labs, https://verogen.com/a-message-to-verogen-customers-about-the-gedmatch-partnership/).

I am aware of only one other website that allows law enforcement searches. Most websites such as 23andMe, do not allow use of their data by law enforcement and are used by most people for ancestry research.

Data is compared within that database and possible relatives are identified. The determination of who is or is not a relative is subjective and based on the length of DNA shared between two individuals. The comparisons in such a database do not yield an identification of someone identical the uploaded SNP data; rather it would identify possible relatives who might be in the database.

Once a putative relative has been identified, a family tree is created, working backwards to grandparents and possible great-great-grandparents. The family tree is then build down. The construction of these family trees is highly subjective and is based on the use of public records such as marriage and birth certificates. Difficulties with tracing a family tree may arise when there are events that sever a relationship, such as an out of wedlock birth, name change, or adoption.

In some instances, contacting individuals for further family information such as noted above. This process leads to a pool of individuals rather than one specific individual.

I am also aware via news reports in the Golden State Killer case, that multiple individuals are often identified and must be eliminated in order to find the potential suspect. As described in the Washington Post, the technique leads to a pool of relatives not to a single individual. (https://www.washingtonpost.com/local/public-safety/to-find-alleged-golden-state-killer-investigators-first-found-his-great-great-great-

grandparents/2018/04/30/3c865fe7-dfcc-4a0e-b6b2-0bec548d501f_story.html, last visited 6/20/2023). This same issue has come up in other cases where more than one individual was identified and subject to investigation (https://abc.com/shows/the-genetic-detective, last visited 6/20/2023).

In the episode from The Genetic Detective show, "Who Killed Angie Dodge?" (https://abc.com/shows/the-genetic-detective/episode-guide/season-01/03-who-killed-angie-dodge, last visited 6/20/2023), at least one relative of the defendant in that case, was investigated and ruled out. Further complicating the Dodge case was the fact that the defendant in that case had been born after his parents divorce and adopted by his step-father.

In addition, there has been a history of misuse of these IGG databases by law enforcement.[1] Abuses by law enforcement of the GEDMatch resulted in users opting out of the use of their DNA data by law enforcement. The abuses and protests by users led GEDMatch to change their database so that users had to opt in for law enforcement use. Another database that allows law enforcement searches, FamilyTreeDNA, automatically opts users into law enforcement searches. An example of other abuses of this technology, a Florida police officer obtained a warrant for the entire GEDMatch database despite users

---

[1] Video history regarding law enforcement use of GEDMatch.
https://www.youtube.com/watch?v=FiiKfruIvcE (last vistied 6/21/23)

DECLARATION OF BICKA BARLOW 5

choice to opt out of law enforcement searches.[2] Recently the state of Maryland passed a statute regulating the use of IGG searches, requiring court approval at many stages of the process.[3]

> But [I]GG has generated concern, not just acclaim. In early 2019, FamilyTreeDNA faced criticism after the public (and site users) learned that the company had secretly been working with the US Federal Bureau of Investigation (FBI) for nearly a year to conduct [I]GG searches. GEDmatch similarly faced reproach after it secretly permitted law enforcement to search its database to investigate a crime outside the scope of the site's stated list of offenses. In response, both sites unilaterally altered their terms of service and privacy policies to explicitly embrace their challenged conduct (3). But even after those efforts sites have come under attack for violating their own policies, quietly changing their settings, and even failing to delete material as promised.[4]

In some cases, both law enforcement and the genealogists that they have employed, have engage in "deceptive" practices, tricking family members, surreptitiously gathering genetic material from nonsupsect family members, or simply uploading an evidentiary profile to a database that does not allow law enforcement searches[5]

It is imperative to the defense in this case to know how Mr. Kohberger was identified and who else in his family tree might have been identified as a subject of investigation.

---

[2] https://www.technologyreview.com/2019/11/06/132047/a-detective-has-been-given-access-to-private-consumer-dna-data-for-the-first-time/ (last visited 6/21/23).
[3] Ram, et al., Regulating forensic genetic genealogy, (2021) Science Vol. 373, 1444-1446.
[4] Id.
[5] Id.

While in the Dodge case, another relative was ruled out, it is unknown to the defense in this case, whether every lead or possible suspect was further investigated and ruled out by genetic testing. Because the family trees generated by these genealogy searches often identify distant relatives through great or great-great grandparents, a person may be completely unaware of the relatives in the genetic genealogy searches and unable to conduct an independent investigation of possible third party suspects. It is not possible for any defendant to investigate family relatives who are unknown to him and may in fact have been in the area of the crime. This material clearly is *Brady* material in that it would provide investigative leads that are otherwise unavailable to Mr. Kohberger.

IGG cases and similarities to CODIS (DNA database searches)

In my opinion as an attorney who specializes in DNA cases and who has seen both cold hit cases and IGG cases, the IGG database search results raise the same type of questions and areas of investigation that a cold hit case[6] would.

- First, as noted above, who else was included in the pool of putative relatives in the first instance in the IGG database.

- Second, once the IGG search was completed, who in the pool of identified putative suspects, were ruled out by further investigation. For the defense team, how many

_____

[6] A cold hit case is a case in which no suspect had been identified and a DNA profile from evidence was created using standard STR based testing. This DNA profile is then uploaded to a state CODIS (convicted offender) database and the data base is searched for a matching profile from an offender.

DECLARATION OF BICKA BARLOW 7

others were investigated and what caused law enforcement to focus on Mr. Kohberger is key to possible investigation and third party suspects.

- Third, and in many ways, most importantly, how does the search of a large database impact the statistical analysis of the comparison by the state lab using standard STR methods? As with cold hit cases, it is now well recognized that the statical analysis of a comparison of standard STR profiles is impacted by a phenomenom called ascertainment bias.

In the case of cold hits, initially, courts did not allow the fact of a cold hit as well as its associated statistical, the Database Match Probability (DMP), as evidence for reasons similar to those presented by the government in this case (Erin Murphy, INSIDE THE CELL, 2015, at 106-119, *see* Attachment 2). However, it is now is now accepted that in a cold hit case, the DMP is generally accepted and can be calculated by crime labs. The best illustration of this can be found in *U.S. v. Jenkins*, 887 A.2d 1013, 1023 (D.C. Court of App., 2005):

> Database match probability accurately expresses the probability of obtaining a cold hit from a search of a particular database. Balding–Donnelly accurately expresses the probability that the person identified through the cold hit is the actual source of the DNA in light of the fact that a known quantity of potential suspects was eliminated through the database search

The government cites to a California trial court order in *In re Michael Green*, (Case # PDL202000007, El Dorado County Superior Court) case, which relies heavily on *People v. Johnson*, (2006) 139 Cal. App.4th 1135, as support for the proposition that the IGG

material need not be disclosed. In its argument that the IGG search is a mere "tip" and not subject to discovery by the defense. The prosecution misreads *Johnson*. The issue in *Johnson* was when the *government seeks to admit* the fact of the cold hit, should the court require the government to also present the DMP, described above, rather than the standard RMP[7] statistic. (*Id.* at 1144). In fact, in *Johnson,* the fact of the cold hit was admitted *over the defendant's objection.*

In *People v. Turner* (2020) 10 Cal.5th 786, the California Supreme Court confirmed that the DMP statistic in a case in which a defendant is identified via a cold hit is relevant and admissible and that

> there is no controversy in the relevant scientific community as to the accuracy of the various formulas. In other words, the math that underlies the calculations is not being questioned. Each approach to expressing significance of a cold hit DNA match accurately answers the question it seeks to address. The rarity statistic accurately expresses how rare a genetic profile is in a given society. Database match probability accurately expresses the probability of obtaining a cold hit from a search of a particular database. Bayesian analysis accurately expresses the probability that the person identified through the cold hit is the actual source of the DNA in light of the fact that a known quantity of potential suspects was eliminated through the database search.

---

[7] In its Motion, the government misstates the statistical rarity of the comparison to the DNA from the sheath (at pg 6) "the STR profile is at least 5.37 octillion times more likely to be seen if Defendant is the source than if an unrelated individual randomly selected from the general population is the source." This reported statistic for this comparison is a Likelihood Ratio, similar to a RMP, which compares two competing hypotheses. The government's statement is extremely misleading and is essentially the "Prosecutor's Fallacy." "The fallacy is to say that [the probability] is also the probability that the DNA at the crime scene came from someone other than the defendant. . . . It does not say that the odds that the suspect contributed the evidence are 1,000:1." National Research Council, THE EVALUATION OF FORENSIC DNA EVIDENCE, 1996, at pg 133.

DECLARATION OF BICKA BARLOW 9

(*Id.* at 804-805).

I have spoken with a number of experts in the field who have expressed the opinion that an IGG search can impact the statistical rarity of a profile in manner similar to a cold hit search, meaning that the statistic that is generated by an analysis of a IGG search could yield a *relevant* and admissible statistic.

In the case of a cold hit case, even a high RMP can yield a low DMP statistic. (INSIDE THE CELL, at 116-15). For instance an RMP of 1 in 1 million can yield a DMP of 1 in 3, which means that while the profile is rare, the likelihood of getting a coincidental match during the search of the database and *selecting a person who matches but is not the perpetrator*, is high (*Id.* at 112). When this evidence is excluded, jurors wonder why and how the defendant was identified. In one case, where the evidence of the search and the DMP was excluded, jurors inquired of the trial court how he had been identified. (*Id.* at 117). Failing to allow the defense access to the ability to investigate and potentially present evidence about the "tip" that the government intends to exclude could lead to a jury being actively mislead by the strength of the government's evidence as it relates to the DNA on the sheath.

The government's arguments are similar to those that were eventually discarded by the court's when they recognized that the search of a database can be considered as relevant evidence.

Research in the area of the impact of an IGG search on the statistical weight of a DNA comparison, using standard STR DNA profiles, is ongoing. The ascertainment bias

DECLARATION OF BICKA BARLOW 10

is accounted for using CODIS and standard STR testing by use of the DMP statistic. The DMP helps a jury understand the likely of a coincidental match when a database is searched which is different that the simple rarity of the profile expressed with an RMP. Without access to the results of the search, including how the first putative relative was identified, how many profiles were in the databases searched, which databases were searched and how the family tree was constructed along with family tree itself, Mr. Kohberger has no means to address possible significant issues with the government's evidence.

Item 4.2: **Match Detail Reports and long-form Candidate Match Reports for the hit(s) including partial hits and hits that are dispositioned to be nonmatching (even if the laboratory has dispositioned a profile as a hit).**

From the discovery in this case, it appears that at least three, and possibly four profiles were uploaded to the CODIS database prior to Mr. Kohberger being charged.

Mr. Kohberger has requested the discovery of these uploads and any resulting hits or candidate matches to these unknown male profile, even if the lab has dispositioned them as non-candidate matches

When a profile is uploaded to the CODIS database, it is searched continuously against the database of convicted offenders and arrestees maintained by the state, generally on a weekly basis. This ensures that as new individuals are added to the database, they will "hit" on evidence sample profiles is they are deemed a "match." These "candidates" are then forwarded directly to the lab that uploaded the evidence profile for further evaluation. The lab will then determine if the individual candidate is a "candidate match"

or non-match. The lab then dispositions the individual in the database as a "candidate match" or non-match. All that the lab sees of the candidates is an anonymized DNA profile. The disposition of a "candidate match" triggers the state to then provide the identity of that individual to the lab for further investigation.

In the last 10 years of my practice, I have encountered a number of cases in which a database search has resulted in more than one candidate and in which, upon court order, the anonymous profiles of the other "non-matching" profiles have been provided to the defense under court order. The number of "non-matching" profiles varies from case to case and include one case in Santa Clara County, *People v George Shirakawa*, case no: 213265 and one case in San Francisco County, *People v. Hernandez*, case no: 12015380. The *Hernandez* case is illustrative of the subjective nature of the interpretation. In *Hernandez*, the lab uploaded multiple interpretations of a mixed sample to the state database. One interpretation returned 5 candidates, and the second returned 28 candidates (including the defendant). In *Hernandez*, the lab opined that two of the candidates could not be excluded as a source while a defense expert reached a different opinion that 15 of the candidates could not be excluded. The lab in the *Hernandez* case provided to the defendant the profiles of the 5 and 28 individuals from the CODIS search. This information was presented at trial in order to impeach the interpretation of the work done by the lab and to undermine the idea that a "match" by a lab is an actual identification of a single individual (see *People v Hernandez*, A144628, https://www.courts.ca.gov/opinions/nonpub/A144628M.PDF, last visited 6/20/23).

DECLARATION OF BICKA BARLOW 12

The fact of this match made by the CODIS system is considered *Brady* material because it provides possible avenues of investigation of other suspects in the case. Given the subjective nature of mixture interpretation, it is reasonable to believe that other experts might disagree with whom actually is the candidate. And in this case, in which the profile at issue is ambiguous and partial, other suspects are an important area of investigation.

Item 11: Unexpected Results and corrective actions

Laboratories are required under accreditation standards to maintain logs of what are called corrective actions or unexpected results along with root cause analysis documents for specific types of incidents encountered in labs. These logs are an important measure of how the test methods are performing in the lab and also can be an indication of ongoing systemic issues as well as analyst specific issues, both of which may be relevant to the testing in this case. The San Francisco Crime Lab routinely produces their entire log to the Office of the District Attorney as *Brady* material and this document is distributed to the defense bar. The last iteration of this document was over 100 pages.

I declare under penalty of perjury that the foregoing is true and correct, and that those matters stated upon information and belief are true to the best of my knowledge.

Executed at San Francisco, California on June 22, 2020

Bicka Barlow
Bicka Barlow

DECLARATION OF BICKA BARLOW 13

BICKA BARLOW
ATTORNEY AND FORENSIC CONSULTANT
2358 MARKET STREET, SAN FRANCISCO CA 94114
(415) 553-4110

EDUCATION

University of San Francisco, School of Law, Juris Doctor, Magna Cum Laude, May 1995

Cornell University
   M.S., Genetics (Minors: Plant Molecular Biology, Cell Biology), January 1990
   McKnight Fellow, 1986 -1988
   National Science Foundation Predoctoral Fellowships, Honorable Mention, 1985, 1986

University of California, Berkeley
   B.S., Genetics, December 1984
   President's Research Fellowship, 1984

MEMBERSHIPS

American Academy of Forensic Science Standards Board, DNA Consensus Body, 2016-2022

Subcommittee on Accreditation and Proficiency Testing, National Commission on Forensic Science, September 2014-2017

American Academy of Forensic Science, 2007-2014

National Association of Criminal Defense Lawyers

State Bar of California, Member #178723
   Member, Executive Committee, Criminal Law Section, 2013-2017

LEGAL EXPERIENCE

Attorney and Forensic Consultant, San Francisco   March 2013-present
Private practice with an emphasis on serious felony and homicide cases involving forensic evidence with an emphasis on DNA evidence.   Services include consultation with lead counsel cases involving DNA evidence, conducting pretrial hearings on DNA evidence, and cross examination of trial experts as well as preparation and presentation of defense experts on DNA evidence

Deputy Public Defender, San Francisco   August 2004 -March 2013
DNA attorney for Public Defender's Office.   Duties include consultation with lead counsel on serious felony cases involving DNA evidence, conducting pretrial hearings on DNA evidence, and cross examination of trial experts as well as preparation and presentation of defense experts on DNA evidence.   In-house training on DNA evidence of attorney, investigators, paralegals and interns.

Private Practice, Criminal Defense, San Francisco   August 1996-August 2004
Represent clients in criminal prosecutions, with a focus on the challenge of DNA evidence at both *Kelly* hearings as well as trial.

BICKA BARLOW
ATTORNEY AND FORENSIC CONSULTANT
2358 MARKET STREET, SAN FRANCISCO CA 94114
(415) 553-4110

Superior Court of San Francisco, Criminal Division
Research Attorney   August 1995 - August 1996
Reviewed pretrial motions in the Criminal Division, including Penal Code § 995 and 1538.5 motions; wrote memoranda to judges analyzing relevant law and facts.   Assisted judges in trial motions, particularly those involving evidentiary issues.

San Francisco Public Defender, San Francisco, CA   January 1995 - May 1995
Intern, Supervising Attorney Michael Burt
Researched and wrote motion for discovery of DNA evidence for a capital case.   Researched state law and scientific issues for a motion to suppress DNA evidence in a capital case.

Federal District Court for the Northern District of California, San Francisco, CA
Legal Extern, The Honorable Judge Eugene Lynch   September - December 1994
Researched and wrote orders on 42 U.S.C. § 1983 civil rights claims and habeas corpus writs.   Reviewed and orally briefed judge on civil matters.

Law Offices of Carmen Gutierrez, Anchorage AK   July - August 1994
Researched and wrote motion to suppress DNA evidence in a criminal trial; interviewed DNA experts; and prepared trial attorney for direct and cross-examination of expert witnesses.

Alaska Office of Public Advocacy, Anchorage, AK
Intern, Supervising Attorney Leslie Hiebert   June - August 1994
Represented misdemeanor clients in hearings for bail, change of plea, sentencing, and revocation of probation; wrote motions and appeals; and interviewed clients.   Advisor on DNA evidence in homicide trial; coordinated DNA testing and data interpretation; and prepared supervising attorney for direct and cross-examination of expert witnesses.

University of San Francisco, School of Law, San Francisco, CA   June - August 1993
Research Assistant, Steve Shatz.
Researched burglary statutes and case law from all fifty states in order to compare the elements of burglary and the proportionality of sentencing to California law.

OTHER EXPERIENCE

ICF Kaiser Engineers, San Francisco, CA
Environmental Associate   January 1990 - August 1992, June - August 1993
Managed group of engineers and geologist responsible for conducting over 100 environmental assessments of hazardous waste sites for the EPA.   Duties included preparation of over 20 site assessment reports requiring government agency record searches, field sampling, and data evaluation.

Cornell University, Ithaca, NY
Graduate Research, McKnight Fellow   August 1985 - September 1989
Independent research using DNA and protein chemistry methods.   Presented departmental seminars on original research.

PRESENTATIONS

DNA Mixtures and Cognitive Bias, 2023, DNA Bootcamp, Los Angeles County Public Defender, Federal Defender Central District of Cal., Los Angeles Innocence Project, Cal State LA Forensic Sci. Instit.

Admissibility Challenges, with Jennifer Friedman, 2023, DNA Bootcamp, Los Angeles County Public Defender, Federal Defender Central District of Cal., Los Angeles Innocence Project, Cal State LA Forensic Sci. Instit.

Theories of Defense: Touch and Transfer, 2022, Wisconsin State Public Defender, Forensic University

Mixture Issues in DNA Cases, 2022, CACJ/CPDA Capital Case Defense Seminar

Transfer, 2021, San Mateo Private Bar Program

Introduction to DNA, 2021, San Mateo Private Bar Program

Complex Mixtures, 2021, National Forensic College, NACDL-Cardozo National Forensic College

DNA Foundations, 2021, Invited Speaker, Forensic Science & The Law, Santa Clara University School of Law

DNA Mixtures: Part I, 2021, CACJ/CPDA Capital Case Defense Seminar

Genetic Genealogy: the new frontier, Arizona Death-Penalty Seminar 2020: Still Fighting for Life

DNA Transfer: Why do we care?, Arizona Death-Penalty Seminar 2020: Still Fighting for Life

Genetic Genealogy, with Erin Murphy and Barry Scheck, 2020, NACDL-Cardozo National Forensic College

DNA Transfer Evidence, 2020, NACDL Defending Sexual Assault and Child Victims Cases Seminar

Mixture Deconvolution using STRmix, 2020, Wisconsin Bar Association

DNA Foundations, 2020, CACJ/CPDA Death Penalty Seminar, Monterey, CA

DNA Transfer, with Kelly Kulick, 2020, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Advanced DNA, 2020, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Cognitive Bias, 2019, Contra Costa County Public Defender and Northern District of California, Federal Defender, DNA Boot Camp, Oakland CA

Familial DNA Searching as a Law Enforcement Tool, 2019, with Michael Chamberlain and Kevin Kellog, Appellate Judicial Attorneys Institute, Redwood City, CA

BICKA BARLOW
ATTORNEY AND FORENSIC CONSULTANT
2358 MARKET STREET, SAN FRANCISCO CA 94114
(415) 553-4110

What every defense lawyer need to know about SFPD's latest software program: STRmix, 2019, San Francisco Bar Assn, San Francisco Superior Court, San Francisco Public Defenders Office

Introduction to DNA Evidence, 2019, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Complex DNA Mixtures, 2019, with Kate Philpott, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Cold Hits and Database Issues, 2018, Contra Costa County Public Defender and Northern District of California, Federal Defender, DNA Boot Camp, Oakland CA

Evolving Issues in DNA, 2018, NACDL, Annual Forensic Science Seminar, Las Vegas, NV

Cross Examining a DNA Expert, with Jessica Goldthwaite, 2018, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Introduction to DNA Evidence, 2018, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Evolution of Statistical Analysis in DNA Testing of Mixed Samples, 2017, Fall Conference, Habeas Corpus Resource Center, San Francisco, CA

DNA Primer, 2017, Advanced Topics in Criminal Defense, NACDL/UACDL, Provo, Utah

Keeping Pseudo-Science Out of the Courtroom, 2017, Advanced Topics in Criminal Defense, NACDL/UACDL, Provo, Utah

Database Cases 2016, Contra Costa County Public Defender and Northern District of California, Federal Defender, DNA Boot Camp, Oakland CA

Emerging Issues in DNA Cases- Mixtures and Low Copy DNA, with Jennifer Friedman, 2016, Hennepin County Public Defenders, DNA Bootcamp, St. Paul/Minneapolis, MN

Introduction to DNA Evidence, 2016, Hennepin County Public Defenders, DNA Bootcamp, St. Paul/Minneapolis, MN

Just Because It's a Match Doesn't Mean It's a Match: Mixtures, LCN, and CODIS Searches, 2016, DNA and More: Developments in Forensic Science, Cook County Public Defender Office, Forensic Division, Chicago, Il

Litigating DNA Evidence, 2016, CACJ/CPDA Death Penalty Seminar, San Diego, CA

Foundations of a Successful Challenge to DNA Evidence, with Kelly Kulick, 2016, CACJ/CPDA Death Penalty Seminar, San Diego, CA

DNA and Wrongful Convictions, 2005, 2006, 2008, 2009, 2013, 2014, 2015, Wrongful Convictions Course, Invited Speaker, Golden Gate University, School of Law

DNA Basics, 2015, Contra Costa County Public Defender and Northern District of California, Federal Defender, DNA Boot Camp, Oakland CA

Foundations of STR Testing including Mixtures, 2015, NACDL and Cardozo Law's National Forensic College, New York City, NY

Database Cases and Low Copy Number/Enhanced Detection Methods, with Jennifer Friedman and Jessica Goldthwaite, 2015, NACDL and Cardozo Law's National Forensic College, New York City, NY

DNA and Penal Code sec 1473, 2015, Spring Conference, Habeas Corpus Resource Center, San Francisco, CA

PLENARY: New Developments in DNA, with Andrea Roth, 2015, CACJ/CPDA Death Penalty Seminar, Monterey, CA

DNA: Low Copy Numbers and Mixtures, 2015, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Low Copy Number DNA and Mixtures, 2014, California Public Defenders Association, Berkeley CA

Understanding-and Challenging-DNA Evidence, 2014, California Appellate Defense Counsel, Annual Conference, Sacramento, CA

Pattern Impression Evidence, Including Toolmark Comparisons, 2014, CACJ/CPDA Death Penalty Seminar, Monterey, CA

DNA Database Issues, 2014, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Taking DNA to Trial, 2013, Texas Criminal Defense Lawyers Association, 11th Annual Forensic Seminar, Dallas, TX

Recent Developments on the Evidentiary Limits of Expert Testimony and Forensic Evidence 2013, with Michael Chamberlain, Appellate Judicial Attorneys Institute, Redwood City, CA

Basic DNA, 2013 CACJ/CPDA Death Penalty Seminar, Monterey, CA

Why Not All Cold Hits Are Created Equal, 2013, CACJ/CPDA Death Penalty Seminar, Monterey, CA

DNA Training Series, September – November, 2012, San Francisco Public Defender's Office, San Francisco CA

Investigating the Lab and Examiner, 2012, Mississippi Spring Public Defender Conference, Biloxi, MI

DNA Update, with Denise Gragg, 2012, CPDA's 43rd Annual Convention Training Program, Oxnard, CA

DNA: From Basics to Advanced - with Jennifer Friedman 2012, Making Sense of Science, 5th Annual Forensic Science Seminar, NACDL and CACJ, Las Vegas, NV.

Low Copy Number DNA Testing and Mixtures, with Rosette Isip, 2012, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Low Copy Number DNA Testing, Mixtures and Cold Hits, with Rosette Isip, 2011, Los Angeles Public Defender's Office Forensic Science Training, Los Angeles CA

Principles of DNA Evidence, 2011, Making Sense of Science, 4th Annual Forensic Science Seminar, NACDL and CACJ, Las Vegas, NV.

DNA: Not All Cold Hits Are Equal, 2011, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Undermining & Utilizing DNA, 2010, CACJ Annual Criminal Defense Seminar, San Francisco CA

The Fallibility of the Infallibility of DNA Evidence, 2010, Annual Western All-Star Conference and Confabulation, Federal Defenders of Eastern Washington and Idaho, Boise ID.
Lab 101 for Lawyers, 2010, THRILLS AND SKILLS XVI, SNITCHES, SCIENCE & SINSEMILLA, Federal Defenders of Eastern Washington and Idaho, Spokane WA.

Can Cold Hit Cases Be Tried Fairly?: Confrontation Clause and DNA Analysis, 2010, 62nd Annual Scientific Meeting, American Academy of Forensic Science, Seattle WA

Cold Hit Statistics and Database Access, 2010, 62nd Annual Scientific Meeting, American Academy of Forensic Science, Seattle WA

Cutting Edge Issues in DNA Evidence, with Andrea Roth, 2010, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Basic DNA Evidence, 2010, CACJ/CPDA Death Penalty Seminar, Monterey, CA

West Coast Forensic Trial College: DNA, 2009, Organizer and Trainer, San Francisco Public Defenders Office

DNA Database Issues, 2009, NACDL and CACJ, 2nd Annual Forensic Science and Technology Seminar, Las Vegas, NV

Access to DNA Databases and Familial Searching, 2009, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Understanding and Debunking DNA Statistical Misinformation, 2009, CPDA, DNA Evidence Seminar, Monterey, CA

Cold Hit Statistics, 2008, Solano County Public Defenders Office, Fairfield CA

Introduction to DNA, 2008, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Basic DNA, 2008, Forensic Science Course, Invited Speaker, University of California, Hastings School of Law

BICKA BARLOW
ATTORNEY AND FORENSIC CONSULTANT
2358 MARKET STREET, SAN FRANCISCO CA 94114
(415) 553-4110

Using DNA Evidence to Your Advantage, 2007, Habeas Corpus Research Center, Summer Conference, San Francisco, CA

DNA Evidence: The Meaning of a Match, 2007, Habeas Corpus Research Center, Summer Conference, San Francisco, CA

West Coast Forensic Trial College: DNA, 2007, 2009, Organizer and Trainer, San Francisco Public Defenders Office

California Commission on the Fair Administration of Justice, 2007, Testimony

DNA Analysis, 2006, Solano County Public Defenders Office, Fairfield CA

Understanding DNA, 2006, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Brainstorming DNA, 2006, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Litigating DNA Discovery, 2005, 2006, Pre Trial Discovery Course, Invited Speaker, University of California, Boalt School of Law

CSI and the Urban Myth of DNA, 2006, Women Defenders, Fall Seminar, San Francisco, CA

DNA Cross Examination College, 2005, Trainer, Public Defender Service, Washington DC

DNA and Serology, 2005, Habeas Corpus Research Center, Fall Conference, San Francisco, CA

DNA, 2005, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Bring Your DNA Case Brainstorming with the Experts, 2005, CACJ/CPDA Death Penalty Seminar, Monterey, CA

Neuroscience and the Law, 2003, Invited Participant, American Association for the Advancement of Science and The Dana Foundation, Washington D.C.

Basic DNA, 2003, CACJ/CPDA Death Penalty Seminar, Monterey, CA

The Basics of Defending a Case in Which DNA Evidence Implicates your Client, 2003, San Francisco Bar Association

DNA is Different, 2002, Panelist with Dr. Henry Lee and Ronald Weich, American Bar Association Annual Convention, Washington D.C.

DNA Evidence: Motions and Challenges, 2001, Arizona Capital Defense Seminar

Attacking DNA Evidence, 2001, California Public Defender's Association, Scientific Evidence Seminar

BICKA BARLOW
ATTORNEY AND FORENSIC CONSULTANT
2358 MARKET STREET, SAN FRANCISCO CA 94114
(415) 553-4110

Forensic DNA Evidence: Is it as Good as You Think?, 2001, Cornell University, Department of Genetics and Molecular Biology

Trial Strategies for Forensic DNA Evidence, 2001, Office of the Public Defender, State of Colorado

DNA Workshop, February 2001, CACJ/CPDA Death Penalty Seminar, Monterey, CA

DNA and the Crime Lab: A DNA Primer, 1999, California Public Defenders Association, Scientific Evidence Seminar

PUBLICATIONS

DNA Warrants Violate Rights, U.S.A. Today, Editorial, November 1, 1999, pg. 28A (with Kim Kruglick)

Severe Penalties for the Destruction of "Potential Life"--Cruel and Unusual Punishment? 29 **U.S.F. L. Rev**. 463 (Winter 1995)

Regulation of the Expression of Pet 122 in the Yeast Saccharomyces Cerevisiae, Thesis, Cornell University (1990)

A Highly Conserved Brassica Gene with Homology to the S-Locus-Specific Glycoprotein Structural Gene, 1 **The Plant Cell** 249 (Feb. 1989) (with Beth LaLonde, Mikhail E. Nasrallah, Kathleen G. Dwyer, Che-Hong Chen, June B. Nasrallah)

INSIDE
the
CELL
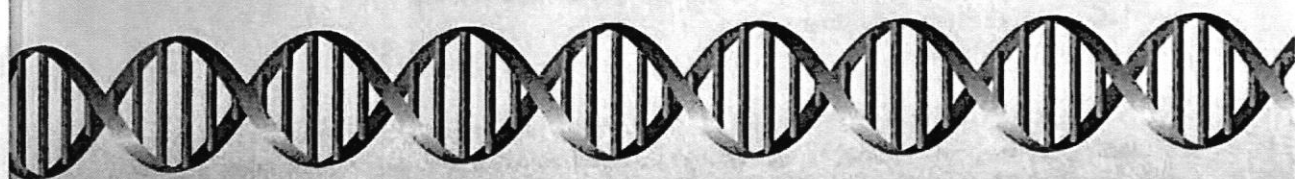
THE DARK SIDE *of* FORENSIC DNA

Erin E. Murphy

Dangers of the Database:
Cold Hits and Coincidental Matches

IN 2001, STATES had just begun to make major inroads in collecting DNA from known offenders and storing the profiles in DNA databases. Kathryn Troyer worked as an analyst in the DNA unit of Arizona's state crime laboratory. At work one day, she received notice of something interesting.[1] Two seemingly unrelated individuals—one white and one black—shared the same two markers at nine of the thirteen places in the standard DNA profile. Indeed, they also shared one marker at the remaining four places—of twenty-six possible genetic differences, they had twenty-two in common. Yet that nine-locus genetic profile should have been exceedingly rare. According to the standard means of computing the random match probability, if you plucked a non-Hispanic white person at random from the population, there would be only a 1 in 754 million chance of finding that profile. For African Americans, the number was 1 in 561 billion, and for Southwest Hispanics, 1 in 113 trillion.[2] And yet here, in a database of just 65,493 people, it was appearing twice—and in people of different races.

Importantly, at that time, one of the most popular testing kits for forensic DNA analysts came in two parts—a nine-locus kit and a six-locus kit.[3] Because the kits were expensive and conducting DNA analysis took up a lot of lab time, many analysts did not bother typing a full thirteen-locus profile in every case. Instead, they routinely did only half the tests, reporting nine-locus matches that quickly turned into

efendants' convictions. Analysts also routinely wrote in their reports, nd testified under oath, that nine- and ten-locus matches were highly nusual and rarely seen.[4] For instance, Harvard professor Frederick Bie- er, an early champion of aggressive programs of forensic DNA typing, estified in one criminal case that he had only seen a nine-locus match etween two people once, and that was in a case of brothers.[5]

Intrigued by the prospect that such matches were in fact not uncom- non, Troyer and two of her colleagues wrote up a quick summary of their indings and submitted the results to a major international conference on orensic DNA typing.[6] Her observations came to the attention of Bicka Barlow, a public defender in San Francisco. Barlow, a passionate attorney who held a master's degree in genetics along with her law degree, took immediate interest. As it happened, she was in the midst of defending a California man, John Puckett, accused of a rape and murder from thirty years earlier.

In 1972, nurse Diana Sylvester was found naked near her Christmas tree. She had been sexually assaulted and fatally stabbed. Police had col- lected forensic evidence, but DNA typing was still decades away. The case sat open until, over thirty years later, investigators dusted off the badly degraded DNA samples and tested them, running the results through the state database. A partial match linked then seventy-year-old, wheelchair- bound John Puckett to the only testable evidence in the case—sperm found on the body that showed reliable results at just five and a half loci, and possible results at another one and a half places. On the basis of this match, prosecutors charged Puckett with murder.[7]

Barlow promptly contacted Troyer for more information, but Todd Griffith, the head of the lab, intervened and denied her request. Shut out, Barlow sought a subpoena from an Arizona court to compel the lab to disclose its findings. Barlow called Troyer to testify at a hearing to decide whether to grant the request. Troyer testified that she had found not just one matching nine-locus pair, but ninety such paired matches. When the lab offered no explanation for why 1 in 1 trillion events were happening regularly in the Arizona database, the judge granted the subpoena. The order also included a directive to the lab to conduct a full search of the known offender database and report back all matching pairs.

Ultimately, that report showed that there were actually quite a large number of these pairwise matches. The Arizona database had only 65,493

people in it, each identified by the two markers at thirteen places that constituted his or her supposed DNA profile. Yet 122 sets of people shared the same genetic markers at 9 places of the 13, twenty pairs matched at 10 places, one pair matched at 11 places, and one pair matched at 12 places. Only the 11- and 12-locus matches were confirmed siblings. It is akin to assuming that you have a fairly unique identifier—such as twenty-six digits that represent birthday, bank account, and social security numbers all combined together—only to learn that a significant number of people share most of those numbers, and in the same order, as you.

As news of these unexpected pairings swept the nation, lawyers in other cities pressed for similar searches. After all, if there were 122 pairwise nine-locus matches in a roughly 65,000-person sized database, how many such matches might be found in the 11 million–person national database? A back-of-the-envelope calculation suggested the figure could be as high as 20,000. But rather than embrace the inquiry, the FBI effectively instituted a crackdown. Officials called Troyer's results "misleading" and "meaningless," and moved to suppress her findings. The FBI also used its power as the guardian of the national database to bully states into refusing to conduct similar studies in their own databases. FBI leaders reprimanded the Arizona lab, claiming that disclosing the results violated its agreement with the FBI. They further threatened to cut off access to the national database to any lab that independently conducted their own such studies, although there was evidence that the threat was intended to scare judges more than lab officials.[10]

Why were Troyer's findings so explosive? The answer turns half on an understanding of math, and half on an understanding of law. And as is so often the case with forensic evidence, the gap between those two worlds proved critical.

THE COLD HIT CASE

At the time of Troyer's findings, state and national DNA databases had started to blossom. In the early days of DNA testing, most people thought of it as a tool to confirm the identity of a person that police had identified as a suspect in a crime. But it was on the brink of becoming something much more significant. The idea of "big data"—the use of vast networks of computers to churn unprecedented amounts of information—was on the cusp of taking off. For instance, although law enforcement agencies

had amassed an incredible trove of fingerprint data, it was not until 1999—approximately the same time that DNA databases were born—that computerized searching became commonplace.[11]

With the advent of computerized networks of biological information, a new kind of case came to the fore. Specifically, the "cold hit" case was born. The FBI built the architecture for its large national repository of DNA profiles, and then put it to work. At first once a week—now twice a week—CODIS software automatically searches the database for associations between all the profiles contained therein.[12] Matches are called cold hits, because they are associations prompted by genetic identity, rather than on conventional investigative leads or information. Suppose a burglary occurs at a home. Crime scene analysts recover some blood by a broken window, which they expect belongs to the burglar. They type the sample, yielding a thirteen-locus DNA profile. Analysts can enter that profile into the forensic index. There, the profile will be searched against other crime scene evidence, as well as the convicted and arrested offender indices. If there is a match, the submitting laboratories will be notified. They can in turn exchange pertinent identifying information, such as the identity of a known offender whose profile matched that from the blood from the broken window. These kind of hits are known as *offender hits*.

Matches may also be made between crime scene samples, otherwise known as *forensic, scene-to-scene,* or *case-to-case* hits. So, suppose that the burglary occurs, but there is no match in the database to a known offender. The profile thus sits in the forensic database. Later, another burglary occurs across the same state. The crime lab recovers evidence, obtains a DNA profile, and uploads that profile to the forensic database. Although there may still not be a match to a known offender, the database will return a match to the earlier burglary. Now investigators know these two incidents are related, and may be part of a broader pattern of criminality by the same person.

Some cold hit cases become "hot" immediately upon investigation. Suppose that the burglar in our example took a valuable painting from the home of the victim. If police use a DNA database to identify a suspect and go to the home of that person, it certainly helps to confirm the individual's guilt if the missing painting is hanging in the suspect's living room. Similarly, scene-to-scene matches might help investigators narrow down a list of suspects by considering unique connections between

people or places that inculpate a particular individual. In other words, a match in the DNA database can always be placed in context with the rest of the evidence, such as corroborating witness identifications or recovery of missing items. Even a criminal record of behavior similar to that alleged, while not always admissible in court, might assuage some concern that the DNA match is purely coincidental.

But some cold hit cases stay cold. Despite investigative efforts, little to no additional evidence may link a suspect to the crime. Or perhaps the mere fact of a match does not conclusively prove that the suspect committed the crime. A DNA match between the suspect and an intimate swabbing from a rape kit may appear fairly clear cut, as there are few innocent explanations for the presence of the suspect's DNA. But in other cases the evidence may be more ambiguous, such as a DNA match to a sample taken from a half-smoked cigarette at the scene, which might have been left by the perpetrator or by an innocent bystander.

Prosecutors proceed in both kinds of cases, not just those in which DNA is part of a constellation of evidence against the defendant. They have proved willing to press for conviction based on the DNA match alone, even in the absence of confirmatory evidence. Or they have defined "confirmatory" evidence in anemic terms—such the suspect's being the right age or gender. Even a prior record may be misleading; after all, virtually every person in the DNA database will have a prior record of some kind. In short, DNA databases raise the stakes of DNA testing. They may change what would otherwise be *confirmatory* evidence into the *sole* inculpatory evidence in a case. They create what are in effect one witness cases—and one *genetic* witness at that.[13]

In such cases, findings like Troyer's—and the uncertainty they engender and which started a national debate among mathematicians, lawyers, and forensic scientists—become indisputably important. The heart of the explanation for Troyer's matches lies in a mathematical parable known as the birthday problem. The lesson goes something like this: How many people must there be in a group before there is just over a 50 percent chance that two people in that group will have the same birthday? For example, the probability that a person's birthday is March 18 is 1 in 365, as there are 365 days in the year (excluding the complication of a leap-year birthday). Yet that is different from asking whether, in a group of people, any two people share any one of those 365 possible birthdays. Rather than keep one side of the equation "fixed" (March 18), this inquiry opens

up both sides of the equation to chance—any single birthday as held by any two persons. Crunching the numbers, it turns out that it takes only twenty-three people before it is more likely than not that two share the same birthday.

This difference between asking "Does anyone in the database match this evidence?" versus "Does anyone in the database match anyone else?" explains why some forensic experts dismissed the Arizona findings as both expected and inconsequential. A sophisticated understanding of match statistics would have led anyone to expect some number of shared nine-locus pairs in the database. But a sophisticated appreciation of criminal trial practice explains why Troyer's findings took criminal justice actors by surprise. Police, prosecutors, testifying lab analysts, and even some defense attorneys found it hard to believe, even infuriating, that nine-locus matches were likely to be common in a large databases. Cases routinely proceeded on the basis of only a nine-locus database match, even without other clearly condemning evidence, because such matches were treated by lawyers and courts alike as conclusive proof of guilt.

Troyer's findings also rekindled the debate about the accuracy of DNA statistics, and led to a public call to allow qualified researchers access to the DNA database to test the validity of the assumptions of independence underlying match statistics. Over forty scientists and academics signed onto a letter published in *Science* magazine, calling for a range of "real-world tests of propositions that previously have been addressed only by simulation."[14] Granting such access was entirely in keeping with both law and practice: the law creating the national database specifically includes a provision for access if "personally identifiable information is removed, for a population statistics database, for identification research and protocol development purposes, or for quality control purposes." And in fact the FBI has granted that access to researchers aligned with its interests.[15] But given possible adverse findings, the FBI shut down all outside inquiries.

Instead, the FBI claimed that the database was not representative of the general national population—it contained too many close relatives at one extreme, and wildly divergent population groups at the other extreme. As leading scientists at the FBI wrote, "[o]bserved departures from expectations will occur using these databases."[16] Accordingly, they asserted that "9-, 10-, 11- and 12-locus (out of 13 loci) matching profiles have been observed, are expected, and do not call into question the reliability

of statistical practices." No effort was made to explain the scope or degree of such expected departures, nor square their existence with the practice of generating match statistics that presume greater randomness.

Intriguingly, one population geneticist tried to reverse-engineer the Arizona findings, using the statistical practices that would be employed in the ordinary criminal case to see whether they squared with the large number of matches found in Arizona. Using models that assume the presence of siblings and different subpopulation groups, he found that "even for the best models, the probability of the Arizona observations is only 9%–12%"—that is, a very narrow set of conditions must hold for what was seen in Arizona to comport with how match statistics are routinely computed.[17] Without knowing if Arizona's matches were typical, researchers cannot be sure whether "modification of the underlying probability models may be required."

Instead of viewing the Arizona matches as an opportunity to refine and improve on the criminal justice system's use of DNA evidence, the government chose instead to try to bury any data at odds with its interest. Indeed, that is what happened next in John Puckett's specific case. The prosecutor proposed to tell the jury the random match probability, which was calculated as 1 in 1.1 million.[18] Partly in light of the Arizona findings, Barlow pressed the court to allow her to present to the jury an alternative match statistic computed in her case to be 1 in 3. That statistic, commonly referred to as the database match probability, or DMP, aims to discount the impact of the match by factoring in the effect of a search in the database. In other words, it tries to account statistically for the difference between a truly random match, and a match made among a finite pool of candidates. Barlow did not invent the DMP; rather, it was put forward as the proper method by a blue-ribbon panel of experts in what is considered the single most authoritative report on DNA evidence in criminal cases.[19]

The government contends that the database match statistic is misleading, because it artificially deflates the match statistic based on the size of the database searched, and does not account for the many ways in which a coincidental cold hit may be undermined—such as a match's failing to square with the suspect in terms of sex, age, or geography. But it is possible to take such information into account. And besides, it is unclear whether narrowing those demographics could unfairly incriminate

swaths of innocent individuals. Most crime is committed by men, within a certain defined age range, and close to home. If sex, age, and geography are the way in which we differentiate a true match from a false one, those characteristics may in many cases provide little safeguard against wrongful accusation.

Amazingly, these two approaches were not the only ways of representing the statistical significance of Puckett's match.[20] Another method, sometimes called the Balding-Donnelly approach, after its major proponents, treated a search in a database as especially informative, because a single match also constituted an exclusion of all the other persons in the database.[21] For instance, if a search in the 11 million–person national database returns a match to only one individual, it signifies that attempts to match everyone else in the database failed. The net result can be a statistic even more inculpatory than would be obtained by using either the database match or random match methods.

Finally, still another approach—and probably the one most helpful to the jury—would have asked, "Of all the men who lived in the metropolitan area at the time of the killing, and who were the right age to have committed the offense, how many would likely match the crime scene evidence?"[22] This approach, nicknamed the n*p statistic by one of its proponents, helps "place[ ] the match probability p in perspective."[23] In Puckett's case, the result of such calculation was that at least two other people living in the area at that time matched the evidence.[24]

Each of these statistics generates very different interpretations of the significance of the match. Yet all are legitimate in one way or another, and there remains a lack of consensus as to which one deserves priority within the criminal justice system. A group of twenty-five renowned statisticians signed a joint letter stating they could all agree that the fact that a match was made through a database carried statistical meaning, even while acknowledging that they could not agree on a single method to express the significance of that match. Defense lawyers have argued that this disagreement among experts, and the vastly different picture painted by varying approaches, require courts to reject database match cases altogether. Alternatively, they have sought additional confirming testing, or at the very least, presentation of conflicting statistics.

Yet courts have rejected their entreaties, for understandable reasons. Throwing out database cases too quickly disregards the value of cases

that are made through a database match. At the same time, ignoring the conflict paints an unfair picture of uniformity, and risks overstating the meaning of the match. Although presenting a bunch of different statistics may leave jurors in the awkward position of having to resolve among themselves a debate that even experienced statisticians cannot decide, it may be the best solution to an otherwise intractable problem.

As databases grow, and cold-hit searches continue, these questions become increasingly important. The determination by the FBI to move from a thirteen-locus CODIS standard to a twenty-plus-locus CODIS standard is animated in part by recognition that adventitious hits may occur already.[25] After all, large US databases continue to expand. A 2014 report by the European Network of Forensic Science Institutes (ENFSI) spelled it out in plain language: "[a]s DNA-databases become larger, the chance of finding adventitious matches also increases, especially with partial and mixed profiles and DNA-profiles of relatives, which have higher random match probabilities."[26] The report gives the example of a DNA profile that has a random match probability of 1 in 1 million. The mean result of searching such a profile in a 3 million–person database is "three matches and none of them may be the actual originator of the crime stain DNA-profile." ENFSI, which since its inception has paid careful attention to the adventitious or coincidental match problem, recommends that "every DNA-database manager ... determine the chance of finding adventitious matches in his/her database."

To facilitate that process, ENFSI even provides a table of the likelihood of a coincidental match in a particular size database, given a particular profile's random match probability, though with the caveat that the calculation is complicated by the fact that far more than one crime scene profile is searched against the database each year. That is, the estimate must take into account not just the result of a single search of a profile against a large database, but a large number of searches against a large number of profiles in the database. For instance, as depicted in Table 7.1, ENFSI provides an estimate of the expected number of coincidental hits when roughly 70,000 crime scene profiles of differing random match probability values are compared to a database with 4 million known persons (roughly a third the size of the national database).[27]

In light of these findings, ENFSI's recommendation is twofold. First, acknowledging the risk of adventitious matches whenever DNA-database matches are found, it counsels:

TABLE 7.1 Expected number of adventitious matches when searching a DNA database of a given size with a DNA profile with a given random match probability.

DNA database size	RMP crime-related stain	Number of searches	Expected number of adventitious matches
	1 : 10.000.000.000	50.000	20
	1 : 1.000.000.000	10.000	40
4.000.000	1 : 100.000.000	5000	200
	1 : 10.000.000	3000	1200
	1 : 1.000.000	2000	8000
Total		70.000	

Source: Reprinted with permission from ENFSI DNA Working Group, *DNA-Database Management: Review and Recommendations* (April 2014), 31, Table 6.

DNA-databases contain large numbers of DNA-profiles of known persons and of biological traces related to unsolved crimes. When the number of DNA-profiles in a DNA-database increases, so does the chance of getting an adventitious match with a person who is not the actual donor of the trace. This is especially true for partial-DNA profiles and mixed-DNA profiles because the chance that they would match with a randomly chosen person is greater than the chance that a full single DNA-profile would match a randomly-chosen person. If there are doubts if the matching person is the donor of the trace, for instance because there is no other tactical or technical evidence which links the person to the crime, the possibility to do additional DNA-testing can be considered. This point of attention particularly applies to matches which are found as a result of . . . large scale . . . DNA-profile comparisons.[28]

Second, ENFSI recommends that managers of DNA databases keep a record of the number of adventitious matches, along with the "conditions under which they were found (size of the database, number of searches, etc.) for future analysis." ENFSI further advises "a warning should be included . . . when reporting a DNA database match." In contrast, here in the United States neither SWGDAM nor the NDIS operational manual—the FBI's DNA guidance documents—discuss the problem, much less possibility, of adventitious hits with any depth. The FBI's response to the issue continues to mirror its response to the revelation of the Arizona matches—to truncate any questioning.

In fact, the only "bad" cold hits that come to light are those in which law enforcement seriously blunders. There are no public statistics on how many database hits generate an investigation that stalls before the suspect is notified or arrested. Instead, what comes to public attention are cases in which the suspects are fortunate enough to have ironclad alibis. The first example of such a mistake occurred in 2000 in the United Kingdom. Using a six-locus match, police arrested a forty-nine-year-old man by the name of Raymond Easton for a burglary that occurred two hundred miles away. One account placed the rarity of that profile as 1 in 37 million.[29] Trouble was, Easton was severely disabled by late-stage Parkinson's disease, and thus was physically incapable of committing the crime. Additional testing eventually exonerated him.[30]

A similar case occurred in the United States, when a woman in Chicago was arrested and charged with a burglary based on DNA evidence found at the scene. Although the precise details have not been released, it seems that lab technicians communicated the match as a "hit" to police investigators, when in fact it was only a partial match. The error came to light only after the woman offered an indisputable alibi: she had been incarcerated on the day of the offense.[31]

Finally, another burglary prosecution, this one in Ohio, perhaps most directly illustrated the dangers of allowing misunderstood DNA evidence to overshadow what is known about a case. The owner of the home wrestled with an invader—described as short, stout, and balding—and managed to grab a few hairs from his head. A six-locus profile with a random match probability of 1.6 million was entered into the DNA database, where it sat for years until a newly loaded offender profile returned a match to Stephen Myers. Myers, a tall, slender man who would have been fifteen at the time of the burglary and had no connection to the city, was indicted for the offense. Investigators waved away the physical discrepancy by noting that appearances change. Fortunately, the prosecutors assigned to the case began to have second thoughts and ordered further testing. Those results proved that Myers was not the burglar, and he was released from jail—after seven months' awaiting trial.[32] In other cases, false matches have been revealed by the "perpetrator's" failing to match the right demographic characteristics, which hardly provides strong reassurance that there are adequate safeguards against the rush to convict based on a cold hit.[33]

As for John Puckett, Bicka Barlow took his case to trial. Ultimately, it is hard to know whether Puckett was guilty of the offenses or not. On the

one hand, no other evidence directly connected him to the crime. None of the many fingerprints found at the scene matched him. At the time of the incident, the sole eyewitness's description had caused police to focus on a different man, but none of the original police investigators was able to testify, given the passage of time. That man had escaped from a mental institution just before the murder, and was suspected in two different sex offenses within a close radius of the victim's apartment. A drop of blood in the man's van had matched the victim, but when Barlow requested that it be tested, it was missing from the evidence file. That man had died several years after the killing, so he could not be confronted.

On the other hand, Puckett was the right age, gender, and race, and had been in the area at that time—all conditions that diminish the probability of coincidental match. He mostly matched the description given by the sole eyewitness, who unfortunately had never been shown Puckett's picture and had died before trial. The jury also heard from three women whom Puckett had threatened with a weapon and sexually assaulted around the same time—the convictions that had landed him in the database. Was Puckett rightly snared by genetic technologies that did not even exist at the time of his offense? Or had police found a coincidental database match, seen a record that fit, and just assumed it must be him?

Thanks to the judge's order, the jury heard only part of the story recounted here. The court excluded all evidence of the alternative suspect. Jurors were informed of the government's probability statistic—that there was a 1 in 1.1 million chance that a person picked at random would match the crime scene DNA. They even heard another government expert claim that the right way to calculate the statistic would have asked, "How many individuals do I have to examine before finding one whose DNA . . . would produce the profile seen" in the crime scene evidence?[34] By that metric, the jury was again told, the relevant number was "about 1 in 1.7 million."[35]

But although the jury was told the random match probability, it never heard that Puckett had been picked as a result of a nonrandom trawl through a police database. The court excluded this evidence even after the jury sent a note during deliberations, asking just how Puckett came to be identified.[36] Thus the jurors were not equipped to contextualize the random match probability information in the context of a database search. Without that information, the jurors would likely not understand that a profile considered "rare"—had other evidence pointed to

the defendant—becomes less improbable when the defendant is found through a trawl in a large DNA database. In short, if a profile's probability is 1 in 1 million, then the government should expect a match if it looks for that profile in a 1 million–person database. The nature of statistics tells us that the profile ought to appear, whether that person committed the crime or not.

The jurors also never heard about the Arizona matches, or how they brought to life the fact that sharing alleles at nine loci is not uncommon. They did not learn that, even using the government's own probability statistic, around forty other people in California matched that crime scene evidence; or that, according to the database match statistic endorsed by the bible of forensic DNA, the National Research Council report, the probability of a match in the database searched by the government was 1 in 3. They never learned that it was likely that two other people in the area also matched the same evidence. In the end, the jury convicted and Puckett was sentenced to life without parole.

Since Puckett's case, every major court that has considered the question of how to treat match probabilities in a cold hit case has rejected the idea that cases should stall until the scholarly community determines the right approach.[37] However, most courts have also not taken the opposite tack—the one absurdly enforced by Puckett's trial court—that only the random match probability should be introduced. Instead, courts have tended to treat the scientific validity of the statistic as a factual question for juries to resolve—allowing both sides to present whichever statistics they deem most favorable.

But a handful of misguided courts continue to unjustly cabin the presentation of match statistics. One court even went so far as to declare that "the means by which a particular person comes to be suspected of a crime . . . is irrelevant to . . . that person's guilt or innocence."[38] Because the suspect is tested again to make the match, the court reasoned, that second test erases the effect of the database search. But in the words of one scholar, the court's theory is "patently fallacious."[39] The "argument that later searches replicate the match from the trawl is not responsive to the concern that an initial trawl dilutes the probative value of the matching DNA."[40] In other words, if you look in a database to find a redhead and do in fact find one, it means little to confirm that the redhead found does, in fact, have red hair. What matters is that you combed a database to find that characteristic. It's the difference between asking whether

you'll pick a redhead if you open a yearbook and point randomly or can find one in an entire graduating class.

The problem of database match statistics also points to something more fundamentally wrong in the way in which forensic evidence is used by the criminal justice system—that enthusiasm for technological solutions to the enduring question of "whodunit?" often outstrips the actual capacity of the technology. As unsatisfying as it feels, there may be cases in which the absence of other evidence inculpating the defendant counsels against giving undue credence to questionable evidence of a DNA match. But rather than admit that not all questions can be answered, the criminal justice system tends to embrace flawed techniques and gloss over genuine concerns about their responsible implementation.

Unfortunately for John Puckett, there will be no further testing in his case. There will not even be vindication of his attorney's position, which later prevailed in the California courts, because Puckett died while his case was pending appeal. Besides, of all the evidentiary samples that were taken, only one matched to him, and the test that was done consumed the entirety of the sample. Puckett's case exemplifies why simply retesting the evidence, or testing more loci to achieve greater certainty about identity, is often not feasible. Old evidence makes for bad verification material.

The problem with cold hits is not likely to disappear. In fact, there is reason to believe that there will only be more, not fewer, such cases going forward. The federal government has encouraged police departments to reopen old unsolved cases and check for possible DNA evidence, and Congress has allocated funds in support.[41] Advocates for sexual assault victims have pressed investigators to look through their troves of untested rape kits in search of possible evidence. These efforts should be applauded, but they should also proceed with caution. New technologies have vastly improved technicians' ability to wrest typeable results from old and degraded samples, but those results are also often more contestable than the profile gleaned from a high-quality, well-preserved stain.

Nonetheless, matches will be made. In many cases, no other meaningful evidence will develop, and a jury or judge will be asked to decide whether a probabilistic statement of guilt—which may be more or less convincing, depending on how it is presented—alone is enough for a conviction.