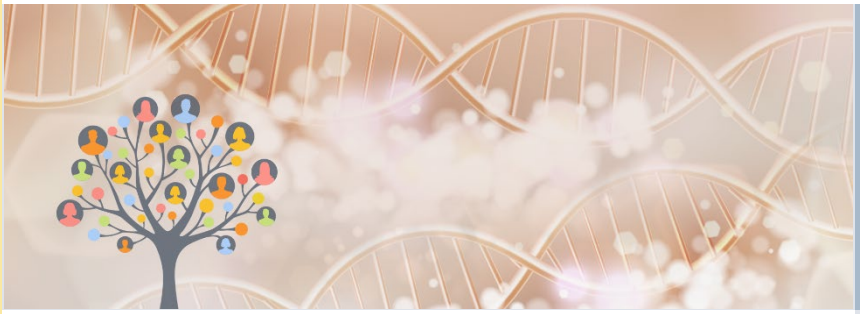
Forensic Technology
CENTER OF EXCELLENCE
A program of the National Institute of Justice

IN-BRIEF

An Introduction to Forensic Genetic Genealogy Technology for Forensic Science Service Providers



"The power of this new partnership between genetic genealogy and law enforcement has unlocked one of the biggest, if not the biggest, crime-fighting breakthroughs in decades."[1]

—CeCe Moore; Parabon Genealogist

Introduction

DNA evidence can play a key role in investigations of cold case violent crimes and in cases of missing and unidentified individuals. Searching DNA profiles against the Federal Bureau of Investigation (FBI)'s Combined DNA Index System (CODIS) is a critical first step toward finding investigative leads—but these searches may not always yield probative matches. When a search does not result in a CODIS match, forensic science service providers (FSSPs) may identify leads using forensic genetic genealogy (FGG), a technique that combines traditional genealogy research with DNA analysis. FGG use is growing across the United States, but many laboratories lack the capacity to generate the single nucleotide polymorphism (SNP) profile required for FGG.

As such, the Department of Justice has issued an **interim policy** that identifies third-party vendors as an option to generate the SNP profile and provides genealogical analysis support. FSSPs collaborate with FGG vendors to identify an appropriate testing strategy for each case, and should advocate for themselves through the FGG process. Thus, understanding how the FGG process works is critical to identifying and prioritizing potential cases to use with this technique.

This brief provides FSSP an overview of the FGG process and factors that they should consider when evaluating potential cases for FGG testing. This brief summarizes information from peer-reviewed literature and conversations with various FGG vendors. The companies interviewed for this brief are not exhaustive of all available vendors; inclusion does not represent NIJ's or FTCoE's recommendation, endorsement, or validation.

Objectives

- Provide an overview of FGG technology to introduce the application of FGG to casework.
- Encourage FSSP participation in multidisciplinary discussions for case submission for FGG testing.
- Provide an overview of major topics that an FSSP should evaluate with the MDT before submitting a case for FGG testing.

FGG can uncover investigative leads in some circumstances when CODIS does not yield probative matches.

To identify unknown victims and suspects, FSSPs may generate a DNA profile and search against CODIS. CODIS includes both the FBI's support of criminal justice DNA databases and the software to run the database.[2] The CODIS database consists of short-tandem repeat (STR) profiles that make up the offender and forensic indexes. CODIS has aided over 545,000 investigations since 1998 and is consistently integrated into DNA analysis workflows[3] but may be limited by the content of its databases. DNA profiles from convicted offenders (and in some states, arrestees) can be legally obtained for testing and upload into CODIS, but they are not collected in some circumstances (in what is considered "lawfully owed" DNA).[4] Inconsistent population of these data can impact database completeness and its ability to yield probative matches (e.g., in cases with unidentified human remains, the individual may have never been reported missing, and would not have family reference samples to search against in CODIS).[5]

When a CODIS search does not yield results, FSSPs may turn to alternative strategies, including FGG. FGG combines genetic and genealogical methods to generate leads for law enforcement entities investigating crimes and identifying human remains.[6, 7] FGG uses profiles of SNPs, which are commonly occuring genetic variations across individuals. SNPs span the entiretey of the human genome and can be passed down through generations. The number of SNP matches links to how closely related the samples are; using high-density SNP profiles (up to hundreds of thousands of SNPs) provides the opportunity to identify more distant relatives beyond immediate family. The concept first started with "direct to consumer" testing services, where consumers submit their DNA for analysis and upload it into genealogy databases to build a genetic family tree and identify relationships to distant relatives. The criminal justice community applied these principles toward identifying suspects or unknown individuals.

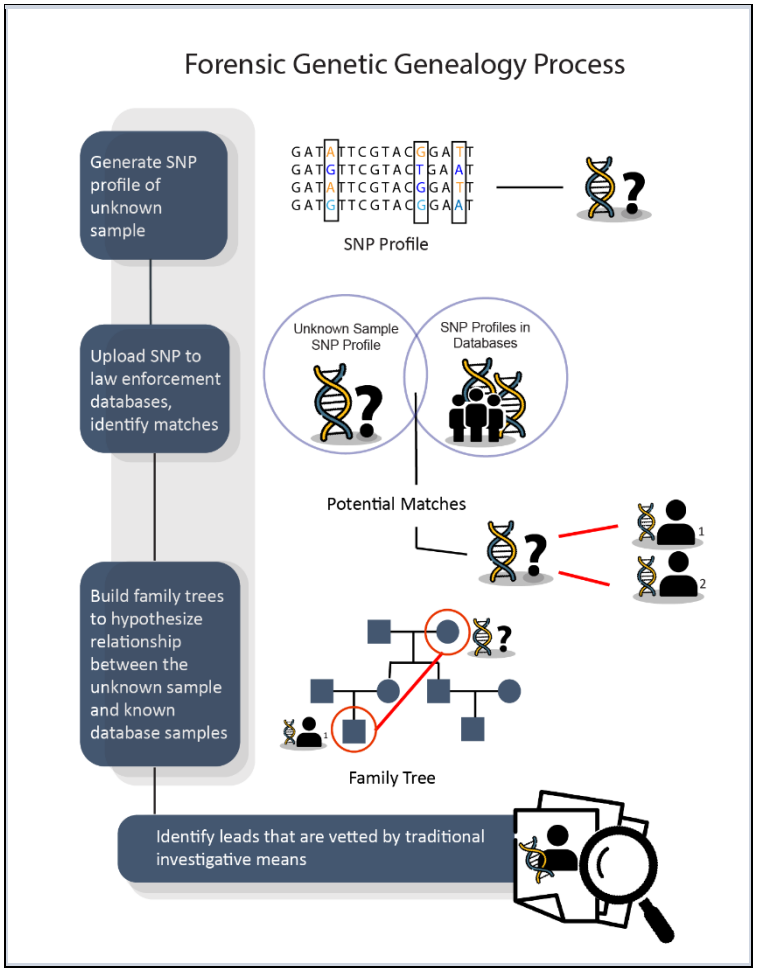


Exhibit 1. Overview of steps taken during the FGG process.

Exhibit 1 outlines the typical FGG process and outputs. First, the case sample is sequenced to obtain a SNP profile. This profile is then uploaded to law enforcement-specific genetic databases to compare the case sample against available profiles that have been uploaded by other consenting individuals. The similarities between the case sample profile and consenting individual profiles can help law enforcement identify individuals who are related to the sample of interest, creating investigative leads that are confirmed with additional STR analysis.

Six things FSSPs should know about FGG

FGG use is becoming more commonplace in the forensics community. FSSPs are a critical mediator between the vendor laboratory and the larger multidisciplinary team (MDT), an allied team of criminal justice professionals who collaborate to determine which cases may be appropriate for FGG. In addition to FSSPs, an MDT may include law enforcement investigators, prosecutors, victim or family advocates, and sexual assault nurse examiners. FSSPs considering FGG should know the following:

1. **The DOJ interim policy identifies third-party vendors, which can provide the SNP profile generation and analysis support, as an option for implementing FGG.**

Many FSSPs lack the capacity to develop a SNP profile in-house or build out the genealogical tree. As such, the DOJ notes that FSSPs can build this capacity using a third-party vendor.[8] FGG vendors offer SNP sequencing capabilities and technical support to interpret and analyze lineages detected from the genealogy databases. These vendors can help create an appropriate testing strategy and provide support with bioinformatics analysis, freeing the FSSP to perform other tasks. Examples of FGG vendors include Astrea Forensics, Bode Technology, DNA Labs International, DNA Solutions, Gene by Gene, Intermountain Forensics, Othram, and Parabon.

FGG uses an SNP profile, which generates different genetic data than STR testing. These high-density profiles, which contain 6K-300K SNPs, enable analysts to identify shared segments of DNA between an unknown suspect or victim and potential relatives.[8,9] Currently, the two most employed methods are SNP microarray and whole genome sequencing (WGS), though some vendors offer targeted SNP assays.[6] **Exhibit 2** provides an overview of the sequencing techniques often employed for FGG, their benefits, and limitations.

FGG vs. Familial DNA Searching

FGG is different from familial DNA searching. Familial DNA searching is the process of running additional CODIS searches in pursuit of biological relatives; this uses STR profiles containing 13–20 markers.[6] FGG uses SNP profiles, which contain thousands of individual markers that recombine as "blocks" in predictive ways between generations. These SNP blocks can be referenced against private genealogy databases to identify **both** close and distantly related individuals. According to DOJ guidelines, familial DNA searching—in states where this technique is allowed—must have been attempted prior to the use of FGG.[8]

SNP Sequencing Technique	Description	Technique is Helpful for….	Limitations
SNP Microarray	A specific type of DNA microarray used to detect polymorphisms within a population.	Single-source and high quality sample types	Often unsuccessful with degraded samples or samples with a large amount of non-human DNA
WGS	A technique used to amplify DNA and identify all sequences in a sample genome.	Degraded samples	Difficult bioinformatics analysis and increased cost
Targeted SNP Assay	Targeted arrays genotype specific sets of specific SNPs to focus research efforts on only the most biologically meaningful variants.	Samples with a small amount of DNA, or degraded sample	Not compatible with all law enforcement databases

Exhibit 2. Overview of SNP sequencing approaches that an FGG vendor may employ.[9]

2. FSSPs play a key role in selecting cases that meet the DOJ interim policy criteria for FGG.

To identify appropriate cases for FGG, the FSSP may work with an MDT, who can provide information and perspectives that can help determine which cases meet DOJ interim policy criteria.[10] Before engaging with a vendor, FSSPs should collaborate with an MDT to review case information and identify which evidence has the highest likelihood of returning viable results (e.g., reviewing sample degradation, prior testing results, contamination concerns, evidence storage and handling). The DOJ interim policy, which must be followed in certain criminal investigations (e.g., when DOJ has jurisdiction or the DOJ provides funding to the investigating agency for FGG purposes), provides guidance on when and how to use FGG. Case criteria include:[8]

The *United States Department of Justice Interim Policy, Forensic Genetic Genealogical DNA Analysis and Searching* can be found here.

- The case must be an unsolved violent crime.
- The forensic sample must be from a putative suspect or unidentified homicide victim.
- DNA profiles derived from the forensic samples must already be in CODIS without a probative match.
- All reasonable leads have been exhausted and the case sample must have previously been uploaded to the Violent Criminal Apprehension Program (ViCAP) and the National Missing and Unidentified Persons System (NamUs), when applicable.

3. The FSSP, in collaboration with the FGG vendor and MDT, should consider submission guidelines when prioritizing cases.

Understanding the FGG process can help FSSPs prioritize cases where FGG could have the most impact. Discussing the following parameters can facilitate an informed, productive discussion with vendors about a testing strategy.

- **Sample Quantity**. The amount of DNA that vendor laboratories require for FGG is constantly changing as the preparation kits to generate SNP profiles become more efficient. Generally, vendors require approximately 1 nanogram of DNA diluted in elution buffer or water. Samples with large DNA quantities are preferred because an aliquot can be made to avoid exhausting all of the original sample. In rare situations, samples with as low as 120 picograms of DNA have been used with FGG[11]; the amount of DNA required varies greatly by sample origin. For example, vendor laboratories may require more DNA from bone-derived sources than they will for blood samples.
- **Sample Quality**. DNA quality dictates whether the sample is sequenced using a microarray or WGS. DNA degradation into short fragments and microbial DNA contamination interferes with this process and affects SNP call rates (i.e., the proportion of an individual's SNP information that is not present in the sample).[12] Microarrays perform poorly with low sample quality[13] and are not recommended in samples with greater than 50% of non-human DNA. WGS sequencing may be advised in highly contaminated samples comprising small DNA base fragments.
- **Sample Mixtures**. Samples of mixed origin complicate the FGG process, and FSSPs should first look for samples of single human origin. When multiple contributors are present in the case sample, some vendors can accept samples in which the majority contributor comprises greater than 50% of the sample.[i] Each vendor has their own rules and guidelines for processing mixed samples, which may depend on the sequencing method.

[i] These values are a general recommendation based on conversations with FGG vendors.

4. Vendors generate and upload the SNP profile to genealogy databases and help the FSSP interpret results.

Vendors are responsible for generating the SNP profile from the forensic sample supplied by the FSSP. Vendors will create a genotype file from the sequenced sample and evaluate the sample for call rate and homozygosity. Vendors can upload the sample and search for SNP similarities in GEDMatch Pro and FamilyTreeDNA because the terms of service of these databases allow searching by law enforcement for forensic purposes. Many vendors have staff trained specifically in building family trees who can compile a list of familial matches based on genetic relatedness of known individuals to the unknown DNA sample in a final report. The final report will vary by vendor but consists of several elements:

FSSPs should consider incorporating prosecutors in their MDT to ensure that vendor contracts align with their privacy and security requirements. Prosecutors may also help ensure that the DNA profile has been uploaded in accordance with the database's terms and conditions, and that the conditions of the DOJ interim policy are met.

- **Context.** This section contains information related to the case type and background information useful for the MDT to reference in the future. The section may include information specific to the sample run, including sequencing methods used and quality parameters.
- **List of Familial Matches.** Information specific to the top familial matches, with charts indicating possible familial relationships and information about the database that was used.
- **Conclusion.** A hypothesis regarding the relationship of the individual to the top matches and recommendations for next steps.

As the vendor laboratory sequences and interprets the SNP profile, the FSSP should remain involved in the ongoing operations and processing of the case sample. Vendors should schedule meetings throughout the process to provide updates on the status of the work and to gather any additional information from the FSSP that might help further the vendor's recommendations for next steps and any follow-up testing required by the FSSP. As a service, the vendor may walk the FSSP through the results of the database search and help them decide the next best steps. FSSPs should advocate for and take advantage of this service whenever possible.

5. Additional testing will be required to validate the familial tree.

An FGG output is a list of potential individuals who may be related to the case sample. With these data, there are two types of testing that will need to be performed after the SNP profile is developed:

- **Reference Testing.** Sometimes the resulting familial tree is incomplete and requires additional information. Genealogists search the developed family tree and identify additional individuals through public records. Further sampling of these individuals with FamilyTreeDNA kits builds out the family tree and helps genealogists narrow down which branch of the family tree to focus on further. This step is not the FSSP's responsibility.
- **Confirmation Testing.** Once the genealogist has provided a lead, law enforcement will facilitate the collection of a new DNA sample from the newly identified lead individual and the FSSP will process the sample for further STR analysis to confirm a DNA match. This STR profile is compared to the forensic evidence STR profile; if there is a match, law enforcement may be authorized to arrest the individual.

6. FSSPs should consider different FGG vendor offerings.

There are a growing number of vendors offering varied FGG services. Before choosing an FGG vendor, FSSPs should note differences in the following:

- **Front-end processing capabilities** (if needed), including DNA extraction and purification. The FSSP may be responsible for these steps if the vendor is not capable of performing them.
- **Ability to generate the SNP profile in-house**. Some vendors may outsource this step, which may have chain of custody or turnaround time implications.
- **Types of samples accepted**. Although most vendors accept a wide variety of biological samples (e.g., bone or hair) some may accept sample types that other vendors do not accept (e.g., cigarette butts or touch DNA samples).
- **Pricing models**. Some vendors offer a flat fee or an hourly rate for services. The choice of sample sequencing and other case factors may impact the costs (e.g., microarrays for SNP generation are typically less expensive than performing WGS). Vendors may pass on additional costs, such as the cost of uploading the SNP profile to either GEDmatch or FamilyTreeDNA.

Conclusion

FGG is a technique in the FSSP's toolkit that can provide significant value when CODIS does not yield probative results. Consistent and regimented use of the technique is imperative to realize its potential impact in resolving cases. FSSPs, in collaboration with an MDT, play a key role in prioritizing cases and items of evidence while maintaining DOJ interim guideline compliance to optimize success in FGG cases. Although vendors provide comprehensive support, it is the FSSP's responsibility to ensure the privacy and security of casework biological samples. Understanding the expectations of and how to engage with FGG vendors is a necessary first step to successfully implementing FGG in practice.

References

1. Gearty, R. "DNA, Genetic Genealogy Made 2018 the Year of the Cold Case: "Biggest Crime Fighting Breakthrough in Decades". *Fox News*, https://www.foxnews.com/us/dna-genetic-genealogy-made-2018-the-year-old-the-cold-case-biggest-crime-fighting-breakthrough-in-decades.
2. U.S. Department of Justice. "Frequently Asked Questions on CODIS and NDIS." Federal Bureau of Investigation, https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet.
3. U.S. Department of Justice, Federal Bureau of Investigation. "The FBI's Combined Data Index System (CODIS) Hits Major Milestone." Federal Bureau of Investigation, https://www.fbi.gov/news/press-releases/press-releases/the-fbis-combined-dna-index-system-codis-hits-major-milestone.
4. U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Assistance. "National Sexual Assault Kit Initiative Training and Technical Assistance Program. Collecting Lawfully Owed DNA." https://www.sakitta.org/resources/docs/Collecting-Lawfully-Owed-DNA.pdf.
5. surviverape.org. "Forensices for Survivors. Healing. Justice." https://www.surviverape.org/forensics/sexual-assault-forensics/codis.
6. Kling, D., Phillips, C., Kennett, D., and Tillmar, A. "Investigative Genetic Genealogy: Current Methods, Knowledge and Practice." *Forensic Sci Int Genet* 52 (May 2021): 102474. https://doi.org/10.1016/j.fsigen.2021.102474.
7. Kennett , D. "Sense About Genealogical DNA Testing." Accessed 6 Nov, 2018. Blog. *Sense About Science*, https://archive.senseaboutscience.org/blog.php/41/sense-about-genealogical-dna-testing.html.
8. U.S. Department of Justice. "United States Department of Justice Interim Policy: Forensic Genetic Genealogical DNA Analysis and Searching." U.S. Department of Justice, https://www.justice.gov/olp/page/file/1204386/download.
9. Oefelein, R. "'DNA Technologies, Laboratory Processes, and Considerations for Choosing Vendors.' Virtual Forensic Genetic Genealogy Training from the National Sexual Assault Kit Initiative Training and Technical Assistance Program." 2022, March 30.
10. McEwen, J., Pino, N., Raphael, A., Renna, K., Boyer, J., and Brody, L. C. "Investigative Genetic Genealogy: Ethical, Legal, and Social Issues and Directions for Future Research." *Forensic Genomics.* (2021): 91-98. https://doi.org/10.1089/forensic.2021.0008.
11. Vogen, M. "Las Vegas Metropolitan Police Department and Othram Partner to Identify 1989 Murderer of Stephanie Isaacson." DNASolves.com, https://dnasolves.com/articles/stephanie_isaacson_murder/.
12. Reed, E., Nunez, S., Kulp, D., Qian, J., Reilly, M. P., and Foulkes, A. S. "A Guide to Genome-Wide Association Analysis and Post-Analytic Interrogation." *Stat Med* 34, no. 28 (Dec 10 2015): 3769-92. https://doi.org/10.1002/sim.6605.
13. Verogen. "High-Quality Outcomes from Low-Quality Samples with ForenSeq Kintelligence." https://verogen.com/wp-content/uploads/2021/07/high-quality-outcomes-low-quality-samples-app-note-vd2021002-b.pdf.

NIJ | National Institute of Justice
STRENGTHEN SCIENCE. ADVANCE JUSTICE.

Forensic Technology
CENTER OF EXCELLENCE

RTI INTERNATIONAL

Published September 2022

FTCoE Contact

Jeri Ropero-Miller, PhD, F-ABFT
Director, FTCoE, RTI International
jerimiller@rti.org

NIJ Contact

Jennifer Love, PhD, D-ABFA
Physical Scientist
Office of Investigative and Forensic Sciences
Jennifer.love@usdoj.gov

Technical Contacts

Ashley Rodriguez, MA
RTI International
ashleyrodriguez@rti.org

Cody Sorrell, MS
RTI International
csorrell@rti.org

Disclaimer

The NIJ FTCoE, led by RTI International, is supported through a Cooperative Agreement from the NIJ (2016-MU-BX-K110), Office of Justice Programs, U.S. Department of Justice. Neither the U.S. Department of Justice nor any of its components are responsible for, or necessarily endorse, this in-brief. NIJ is the research, development, and evaluation agency of the U.S. Department of Justice. NIJ is dedicated to improving knowledge and understanding of crime and justice issues through science. NIJ provides objective and independent knowledge and tools to inform the decision-making of the criminal and juvenile justice communities to reduce crime and advance justice, particularly at the state and local levels. The NIJ Office of Investigative and Forensic Sciences (OIFS) is the federal government's lead agency for forensic science research and development. OIFS's mission is to improve the quality and practice of forensic science through innovative solutions that support research and development, testing and evaluation, technology, information exchange, and the development of training resources for the criminal justice community.

Public Domain Notice

Suggested Citation

Sorrell, Cody, Ashley Rodriguez, Rebecca Shute, Shannon Krauss, and Patricia Melton. "An Introduction to Forensic Genetic Genealogy Technology for Forensic Science Service Providers." Research Triangle Park, NC: RTI International, September 2022.