



A review: On emerging trends and technologies in familial DNA searching

Shreya Sharma * and Susmita Das

Department of Forensic Science, Vivekananda Global University, Jaipur, India.

GSC Biological and Pharmaceutical Sciences, 2025, 30(02), 210-216

Publication history: Received on 27 December 2024; revised on 13 February 2025; accepted on 16 February 2025

Article DOI: <https://doi.org/10.30574/gscbps.2025.30.2.0052>

Abstract

Forensic science has made tremendous strides in recent years, making it challenging to draw comparisons with older investigation methods. One of the groundbreaking advancements is the use of DNA technology in forensics. A cutting-edge technique called "familial DNA searching" involves deliberately scouring a DNA database to find matches with a limited set of genetic markers. This approach aims to identify potential relatives of an unknown individual, providing new leads for investigations. The FBI oversees the CODIS database, which consolidates DNA data from various U.S. crime labs. Familial DNA searches rely on identifying common genetic traits, or alleles, and assessing the rarity of these shared alleles in the broader population. The database contains profiles from criminals, arrestees, convicted felons, and others from whom DNA samples have been lawfully collected. When identifying potential familial ties between an individual and a suspect's DNA, the focus is solely on shared genetic markers that don't reveal any specific biological traits apart from gender. The FORENSIC SCIENCE Service. Ltd (FSS) has developed two algorithms to assist in familial DNA searching within the NDNAD (DNA Database). The first algorithm aims to pinpoint DNA profiles that could be linked as parent-child matches to the target DNA profile. The second algorithm prioritizes DNA profiles likely to be siblings of the actual perpetrator, ranking them based on the number of shared alleles with the target DNA profile. Before utilizing a DNA database for familial searches, it's crucial to ensure compliance with existing laws, regulations, court orders, or other governing guidelines. Notably, familial DNA searching is in line with the objectives of the Federal DNA Identification Act.

Keywords: DNA Searching; Criminal DNA Identification; Inherited Alleles; Genetic Markers; Legal Proceedings

1. Introduction

The domain of forensic science has made such enormous progress in recent years that the techniques employed for inquiry hardly resemble those from years ago. One of the newest advancements in forensic technology is DNA, which may also be considered a tool for forensic scientists. "Familial searching" is the deliberate process of looking for a match on a limited subset of the typed loci that are available in a DNA database [1]. The goal is to find relatives that have never been found before, hopefully leading to new lines of inquiry [1]. The FBI funds a computer database called "CODIS," which houses DNA profiles that local states and federal criminal laboratories in the US submit. Depending on the state, DNA profiles against both convicted and arrested criminals. With the use of these DNA profiles, it is also possible to discover anyone in the database who may be a close relative of an alleged criminal. "Familial DNA Searching" is the term for using a DNA database. Law enforcement officials who look into information from searches can use the comprehensive written guidelines and standards that the National DNA Database of the United Kingdom has produced [2].

* Corresponding author: Shreya Sharma.

1.1. How familial DNA searching works?

The number of shared genetic traits, or alleles, and the rarity of those shared alleles among human groups serve as the foundation for Familial DNA searching. This search allows matched subsets of alleles to be used as a standard for comparison, which is different from a straight match search. It is reasonable to assume that close relatives of a targeted criminal will have more common alleles, particularly rare alleles, than unrelated individuals because human alleles are inherited one to one from parents [3].

1.2. Protocol

DNA databases should be searched using a familial search strategy to see if any current stability, regulatory, legal, or other controlling authority addresses this kind of database use. The interpretation of non-specific statutory language may be necessary to determine if familial searches are permitted in a certain jurisdiction. Any primary rights, interests, or concerns held by those affected must be carefully weighed against the interests of public safety and law enforcement. This group includes the person's family, any suspects discovered as a result of a familial search, and the eligible persons who are ultimately discovered to be potential relatives of the suspected offender [2]. Such searches are only permitted in the UK for very serious crimes. The usage and publication of DNA databases that are subject to statutory authorities are often restricted. State and federal laws forbid the use of DNA data for purposes other than those officially acknowledged by law enforcement, such as criminal identification and missing person identification. Database profiles may contain information on convicted criminals, arrested individuals, and/or other individuals from whose samples have been legally taken. Shared genetic markers only code for gender and no other known biological attribute, hence any first determination of potential kinship between a target perpetrator's DNA profile and an individual is predicated entirely on the existence of these shared genetic markers. In order to confirm that a qualifying individual is, in fact, related to the accused perpetrator, SLT agencies may decide to withhold additional demographic data, such as race, age, and so forth, when a qualifying individual is identified based on a sufficient degree of genetic similarity to the accused perpetrator. When a particular qualifying person is identified because of sufficient genetic similarity to the alleged perpetrator, SLT agencies may choose to relieve other demographic information, including race, age, etc. to further test the hypothesis that the identified person is, indeed related to the alleged perpetrator [4]. Family searching was introduced by Forensic Science Service Ltd. (FSS) in 2002 to aid in the advancement of criminal investigations in cases where a complete DNA profile was available based on biological trace evidence that was allegedly left at the crime scene by the real offenders, but where no matches were found between the profiles of any individual retained on the NDNAD. The availability of both complete target DNA profile and incomplete DNA profiles generated from tiny or damaged DNA samples is necessary for the family DNA search technique to be effective [5].

The FSS designated **two types** of algorithms for Familial Searches

- To locate any DNA profiles in the NDNAD that might be connected to the target DNA profile as a kid or parent.
- To rank the DNA profiles with the NDNAD according to the number of alleles shared with the target DNA profile in order to determine which people are more likely to be connected to as a sibling to the actual offender.

Simulation trials show that this strategy can potentially produce adventitious pairings.

1.3. Effectiveness of familial DNA searching

Amongst several natures of forensic investigations, including automobile breaks and incidents, robberies, sexual assaults, and homicides, familial searches have been effective in identifying the perpetrators. Quality assurance protocols have been developed over the years in DNA testing to boost confidence in the obtained results. Due to the knowledge gathered from organizations like the European DNA Profiling Group (EDNAP), the European Network of Forensic Science Institutes (ENFSI) in Europe, and the scientific working group on DNA Analysis Methods (SWGDM) and DNA Advisory Board (DAB) of the Federal Bureau of Investigation (USA), the DNA testing quality infrastructure is likely more advanced than that of many other forensic disciplines. Globally, DNA testing is also used to combat fraud and abuse issues that come up during family reunions and international migration [6]. Furthermore, relationship testing is used in a lot of criminal and missing person identification investigations. The Likelihood Ratio (LR), which compares the probabilities of witnessing the genetic data given two alternative hypotheses, is typically used to assess a kinship link. A person is linked to another person in a certain relationship as opposed to the two people not being related. The more unrelated hypotheses there are, the higher the LR (usually <1). Factors like testing kits, co-ancestry measurements, mutations, and the frequency of alleles in a specific group play important roles in relationship testing. The selection of a testing kit, including the type and number of markers it contains, is crucial because it determines how accurate the test will be [4]. Back in the early 2000s, most relationship tests used either custom-made kits or commercially available ones like AmpFISTER Profiler and Cofiler. These kits typically had 8 to 11 autosomal Short Tandem Repeat (STR)

markers each. To ensure that they covered the required 13 DNA Index System (CODIS) core locations and to improve the test's accuracy, multiple kits were often used together in a single test. In the early 2000s, relationship testing mainly relied on custom-made or ready-to-use kits like AmpFISTER Profiler and COfiler [7]. These kits typically contained 8-11 autosomal Short Tandem Repeat (STR) markers each. To ensure accuracy and boost discrimination capabilities, multiple kits were often combined to cover the required 13 DNA Index System (CODIS) core loci. As technology advanced, newer kits such as Identifier and Powerplex16 were introduced. These kits were designed to encompass all 13 FBI CODIS core loci, plus a few extra markers streamlining the process [7].

2. Method and techniques of familial dna searching

Interpreting forensic evidence involves comparing the data collected from a crime scene (Q) with known reference samples (K). The accuracy of this comparison largely depends on the quality of the crime scene data and the availability of a suitable reference sample. When we have DNA samples from a suspect or multiple suspects, it's straightforward to compare them at specific genetic markers. However, if we don't have any suspects, we can search through DNA databases created over the last 20 years. These databases help us identify potential matches with profiles from past criminals, essentially allowing us to broaden our search to find the person responsible for the crime [8].

Methods used in Familial DNA Analysis

2.1. CODIS – Combined DNA Index System

Familial DNA profiling Analysis necessitates more precision. The analyser for the DNA profile is added to CODIS in the form of an allele number, and it is assessed against all possibilities in the program. Depending on the level of stringency, CODIS software can be utilized at low, moderate, or high levels for matching DNA profiles. High stringency demands a precise match in the DNA profile.

- Partial matching results are produced when we discuss moderate or low stringency. Regarding criminal investigative protocols, DNA profiles with great stringency (exact match) are seen to be more practical.
- There are two main parts to Familial DNA Searching (FDS)
 - Software exists that can compare an unidentified DNA sample to known DNA profiles in a database.
 - Additionally, lineage testing is used to further confirm or refute the potential family ties of an individual by analyzing genetic similarities commonly found in close relatives. This method reduces the chances of false matches.
 - One key technique within lineage testing is Y-STR testing, which focuses on unique genetic markers found on the Y chromosome to determine if there's a paternal relationship, especially between males.
 - Another method is Mitochondrial DNA testing, which helps trace maternal lineage. This kind of DNA analysis has been instrumental in criminal investigations in countries like the US, UK, and New Zealand.

3. GED match

Curtis Rogers and John Olsen are two dedicated genealogy enthusiasts behind the website GED Match. Established in 2010, their platform initially offered a tool to compare family names, aiding researchers in understanding y-DNA matches. Users could upload their DNA data to sites like GED Match and Family Tree DNA, and then trace back their family trees to find the most closely related genetic match. Over time, genetic genealogy, especially using y-chromosomal DNA (Y-DNA), has become intertwined with various fields, including forensic science. Additionally, mitochondrial DNA (Mt DNA) has also been a valuable tool in both genealogy and forensic investigations.

4. Y-STR databases

In forensic or genealogical databases, paternal relations can be located using Y-STR chromosomes. The advantage of Y-STR Haplotypes over those with autosomal genetic markers is the ability to identify more distant relationships. Nonetheless, the STRs can highlight variations amongst paternal relatives that are closely linked. Using a reference sample of buccal swabs, a CODIS Core Loci STR Profile is created using an entirely automated approach. Without the need for human interaction, the "swab in – profile out" procedure includes automated extraction process, amplification of gene, segregation, detection, and allele calling. In forensic or genealogical databases, paternal relations can be located using Y-STR chromosomes. The advantage of Y-STR Haplotypes over those with autosomal genetic markers is the ability to identify more distant relationships. Nonetheless, the STRs can highlight variations amongst paternal relatives that are closely linked. Using a reference sample of buccal swabs, a CODIS Core Loci STR Profile is created using an

entirely automated approach. Without the need for human interaction, the "swab in – profile out" procedure includes automatic extraction, amplification process, separation, detection, and allele calling.

4.1. Single nucleotide polymorphisms

The ABI 3500XL genetic analyser, produced by Thermo Fisher Scientific, is a popular device used for separating and determining the size of STR alleles in DNA samples. Thermo Fisher Scientific has also created advanced software designed to help with the genotyping process of these DNA samples. The main functions of this software include checking the conditions of the electrophoretic run, identifying the specific wavelengths needed to capture the fluorochromes using virtual filters within the CCD, and ensuring accurate data collection. The software helps in creating a checklist of samples needed for setting up an electrophoresis run. This includes details like the sequence and conditions for injecting samples, the settings for the electrophoretic run, and the virtual filter to be applied. After running the electrophoresis, each sample generates a raw data file along with a graph. This graph plots the number of data points on one axis against fluorescence units on the other. To interpret this data and analyse the STRs, the raw data is processed using specialized software like Genescan, Genotyper, or Genemapper, which transform it into a readable genetic profile.

- Identifies the peaks by the threshold value and also determine the height and area parameters of the various peaks;
- Distinguishes the emission spectra of fluorochromes based on the matrix;
- Assigns the size to the fragments based on comparison with the internal standards peaks.

After the peaks are identified and sized using the Genotyper program by comparing them to ladder peaks, the results are displayed in an electropherogram. This graph shows multiple lines, each representing a different color, and displays the various loci with their corresponding alleles sorted from shortest to longest. Thermo Fisher Scientific introduced the GeneMapper ID software in November 2003, enhancing the capabilities of GeneScan and Genotyper. One notable feature is the automated Quality value system, which assigns quality scores during allele calling and size determination, helping to pinpoint any potential issues during preparation for analysis. This system allows for easy integration of results into databases, printing, or exporting for further analysis.

Interpreting and reporting DNA typing results for human identification requires a high level of expertise and professional judgment. This complex process relies on empirical data as well as the scientist's knowledge and experience. According to the Quality Assurance Standards for Forensic DNA Testing Laboratories by SWGDAM (2017), forensic labs must establish and follow documented protocols for reporting and interpreting DNA results. The correct interpretation of STR analysis data should adhere to the SWGDAM guidelines for autosomal STR typing, which were approved in 2017.

5. Kinship relationship techniques

The U.S. loci are utilized to assess relatedness between individuals in addition to identifying each individual for forensic purposes. Rather of pinpointing an exact match between two DNA profiles, Kinship Analysis evaluates any genetic information shared by potential relatives under different scenarios or by approximating an unidentified relationship. The child will share alleles that are **identical by descent (IBD)** at every tested marker if the purported father is, in fact, the biological father. This is because the allele came from a common ancestor in father-child and mother-child couples. The likelihood of an IBD allele being shared by a pair of relatives is known [6].

The strength with which the genetic data support the proposed link will be assessed using the population allele frequencies, the genetic data, the proposed relationship, the associated IBD probability, and the LR calculation. Applications for kinship analysis include those involving civil and criminal paternity issues, immigration applications, cases involving missing persons, identifications of disaster victims, and searches of offender databases by relatives. A putative father-child link in paternity cases can be verified using data from genome sequencing of DNA and statistical analysis [8].

DNA testing is commonly used to determine relationships especially in paternity cases, DNA samples from mother, child and potential father are mainly analyzed using 15-20 specific genetic markers known as forensic STR markers (short tandem repeat). These markers are often tested using commercial kits designed to identify 15 Specific markers nowadays it is being used more than 20 STR markers. These markers are often tested using commercial kits specially designed to identify these particular genetic markers in DNA. These kits usually target 13 standard CODIS locations along with D2S1338/D19S433 or Penta D/ Penta E, depending on the kit's manufacturer [9].

For most cases involving a mother, child and alleged father, analyzing these forensic loci is usually enough to confirm or refute paternity. However the field of forensic DNA analysis involving kinship and familial searching. This means looking for indirect matches and dealing with increased uncertainties [9].

In situations like identifying disaster victims or conducting familial searches, searching a database of 13-15 STR profiles can yield many coincidental matches, especially for close biological relatives like parents and children or full siblings, while most true parent-child matches show strong support for their relatedness, there have been instances where unrelated individuals had test results that mistakenly suggested a parent child relationship [7].

5.1. Defining threshold frequencies

An LR threshold was chosen for this investigation based on the concept of an LR, designating a positive or negative connection as being more than or less than one, respectively. To attain specificity and sensitivity thresholds, however, various LR thresholds might actually be necessary (Gaytmenn et al. 2002). If greater LRs are used, there will be a trade-off between more erroneous exclusions and fewer false inclusions. Finding a balance is important and can vary based on the application (such as immigration tests, familial searches, paternity testing, etc.). Unlike a discrete LR threshold, a grey zone technique may be helpful in setting a likelihood ratio range that does not entirely eliminate misunderstanding regarding the relationship status.

5.2. Evidence of familial DNA searching

In 2003, Gladys Godfrey was killed in Nottinghamshire, UK, and the murderer was initially identified through the use of familial DNA research. Although few DNA samples from the scene of crime were collected by the investigators, the murderer had not yet been apprehended, thus they were unable to locate an accurate match on the UK's National DNA Database. Instead, the investigators chose to search for close matches in their DNA profiles, who might be the murderer's relatives. Through further refinement of near matches utilizing additional factors, the UK's Forensic Intelligence Bureau was able to produce a "top 20" list of persons of interest. Gladys's killer was identified, apprehended, and charged as a result of this strategy. But because this strategy made use of STRs, there were very few [10].

The Golden State Killer, a well-known American murderer, was identified more recently thanks to a more sophisticated SNP-based familial search technique [4]. The Golden State Killer was thought to have carried out a spate of rapes, killings, and robberies in California (USA) during the 1970s and 1980s. Investigators had samples of his DNA, but until 2018, they had been unable to identify him since there was no profile that matched his in the FBI's nationwide database, CODIS (Combined DNA Index System) [10].

Subsequently, the scientists took the very odd step of building an SNP-based profile, which they then uploaded to the GED match genealogy website. Authorities were able to concentrate their search on retired police officer Joseph James DE Angelo after discovering that several of the Golden State Killer's relatives had profiles on GED match. Confirmation came after officers took two distinct DNA samples from his trash and processed them [11].

5.3. Future prospects of familial DNA searching

Advances in technology have revealed that using familial DNA analysis in criminal investigations can be beneficial. When it comes to specifications, this analysis is more strict than other forensic DNA tests. The searching capacity of familial DNA analysis can be expanded by the use of probability calculation software with additional correction factors based on the ethnicity in question. DNA comparison can be combined with Y-STR, mitochondrial, and SNP analysis to yield partial matches. In the future, employing multiple comparison DNA techniques can serve two purposes: it could mitigate numerous ethical, legal, and societal difficulties related to familial DNA analysis and reduce the rate of discrepancies.

5.4. Downsides of familial DNA searching

Family DNA analysis can be problematic due to potential false matches. This method can sometimes mistakenly link individuals to crimes based on partial DNA similarities. Many countries lack a unified policy on how to use family DNA analysis [12]. There are concerns about the security of DNA profiles if proper protocols aren't followed during searches, risking the unauthorized sharing of personal DNA data. Moreover, comparing a specific DNA profile with various databases can compromise an individual's privacy [11].

Another challenge with familial DNA matching involves statistical analysis. Partial matches in DNA comparisons require careful interpretation. Adjustments in statistical methods, like the likelihood ratio and random match probability, are needed to ensure accurate results. Failing to update these statistical techniques could lead to misinterpretations and unjust use of family DNA evidence [10].

When it comes to using family DNA analysis, statistical analysis is a key factor to consider. Deciding if a match is valid requires additional scrutiny because of the partial similarities found in family DNA searches. It's important to adjust certain statistical methods, like the likelihood ratio and random match probability, to get accurate results [8]. If these statistical approaches aren't updated correctly, it could lead to misunderstandings and the improper use of family DNA findings [7].

6. Conclusion

Familial DNA searching is an effective investigative tool for solving cases by discovering relatives of unknown individuals through genetic links, as technology progresses, its involvement in forensic investigations is certain to grow, providing law enforcement with new means to solve cold cases and identify perpetrators while posing ethical and privacy problems that must be properly addressed. Balancing its benefits with responsible use will be critical to its long-term utilization.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

- [1] (Mateen)Author links open overlay panelRana Muhammad Mateen, AbstractSince the discovery of Deoxyribonucleic acid (DNA) capability in forensic investigation, Krenke, B. E., Gershaw, C. J., Bryc, K., Shafique, M., Larmuseau, M. H. D., Kling, D., Starinsky-elbaz, S., Karantzali, E., Wickenheiser, R. A., Maguire, C. N., Giill, P., Butler, J. M., Codis, Melton, T., Fortier, L., Debus-sherrill, S., Jeong, S. J., ... Suter, S. M. (2020a, November 4). Familial DNA analysis and criminal investigation: Usage, downsides and privacy concerns. Forensic Science International. <http://www.sciencedirect.com/science/article/abs/pii/S0379073820304382>
- [2] Elaine Ortyl. (n.d.-a). DNA and the Fourth Amendment: Would a defendant succeed ... - sage journals. DNA and the Fourth Amendment: Would a Defendant Succeed on a Challenge to a Familial DNA Search? <https://journals.sagepub.com/doi/abs/10.1177/0098858819892746>
- [3] C.N. Maguire a, C. N. M., a, b, c, d, AbstractThe National DNA Database (NDNAD) of England and Wales was established on April 10th 1995. The NDNAD is governed by a variety of legislative instruments that mean that DNA samples can be taken if an individual is arrested and detained in a police , Reid, T. M., Curran, J. M., Kloosterman, A., Rothstein, M. A., Bieber, F. R., O'Connor, K. L., Haimes, E., Gabel, J. D., Gal, M., Graham, J., Glaze, L. E., Shelton, N., Bernasco, W., ... Unit, H. O. P. S. (2013, September 6). Familial searching: A specialist forensic DNA profiling service utilising the national DNA database® to identify unknown offenders via their relatives-the UK experience. Forensic Science International: Genetics. <https://www.sciencedirect.com/science/article/abs/pii/S1872497313001543>
- [4] Author links open overlay panelCassandra J. Gershaw, C. J. G., AbstractDNA evidence is widely recognized as an invaluable tool in the process of investigation and identification, Koblinksy, L., Steinberger, E., Bieber, F. R., & Budowle, B. (2010, August 23). Forensic utilization of familial searches in DNA databases. Forensic Science International: Genetics. <https://www.sciencedirect.com/science/article/abs/pii/S187249731000116X>
- [5] Elena Pilli. (n.d.). Forensic DNA analysis. Google Books. https://books.google.co.in/books?hl=en&lr=&id=A9MREAAAQBAJ&oi=fnd&pg=PA77&dq=working%2Bof%2By-strs%2Bin%2Bfamilial%2Bdna%2Bsearching&ots=PaeX8j0eRX&sig=6e9R-7-Kl4wAtIW5CH8e7ig83ns&redir_esc=y#v=onepage&q=working%20of%20y-strs%20in%20familial%20dna%20searching&f=false
- [6] Author links open overlay panelKaye N. Ballantyne a 1 2, K. N. B., a, 1, 2, 3, b, c, AbstractThe panels of 9–17 Y-chromosomal short tandem repeats (Y-STRs) currently used in forensic genetics have adequate resolution of different paternal lineages in many human populations, Prinz, M., Budowle, B., Krenke, B. E., Mohyuddin, A., Hedman, M., Xue, Y., Kayser, M., Zerjal, T., D'Amato, M. E., Rodig, H., Lessig, R., ... Calacal, G. C. (2011, May 25). A new future of forensic Y-Chromosome Analysis: Rapidly mutating Y-strs for differentiating male relatives and paternal lineages. Forensic Science International: Genetics. <https://www.sciencedirect.com/science/article/abs/pii/S1872497311000937>

- [7] Kristen. Lewis “O” Connor. (n.d.). Evaluating the effect of additional forensic loci on likelihood ratio ... Evaluating the effect of additional forensic loci on likelihood ratio values for complex kinship analysis. https://strbase-archive.nist.gov/pub_pres/OConnor_Promega%20Proc%202010_Additional%20Loci%20Kinship.pdf
- [8] Abigail Sawyer. (n.d.). Who owns your DNA? | biotechniques - future-science.com. Who owns your DNA? <https://www.future-science.com/doi/full/10.2144/btn-2021-0059>
- [9] Hillary L. Kody. (n.d.). William & Mary Law Review. Standing t Standing to Challenge F o Challenge Familial Sear amilial Searches of Commer ches of Commercial DN cial DNA Databases. <https://scholarship.law.wm.edu/cgi/viewcontent.cgi?article=3831&context=wmlr>
- [10] Ge, J., & Budowle, B. (n.d.). How many familial relationship testing results could be wrong?. PLOS Genetics. <https://journals.plos.org/plosgenetics/article?id=10.1371%2Fjournal.pgen.1008929>
- [11] John M. Butler. (n.d.). The Future of Forensic DNA analysis - Royal Society. The future of forensic DNA analysis. <https://royalsocietypublishing.org/doi/10.1098/rstb.2014.0252>
- [12] Author links open overlay panelDebbie Kennett, D. K., AbstractIn the last year direct-to-consumer (DTC) genetic genealogy databases have been used to identify suspects and missing persons in over fifty cold cases, Phillips, C., Greytak, E. M., Wagner, J. K., King, T. E., Wang, C.-C., Willuweit, S., Murphy, E., Court, D. S., Augenstein, S., Wamsley, L., Gafni, M., Lee, A., Murphy, H., Molteni, M., NanoLabs, P., Hernandez, S., Guerrini, C. J., ... Gymrek, M. (2019, May
- [13] Using genetic genealogy databases in missing persons cases and to develop suspect leads in violent crimes. Forensic Science International. <https://www.sciencedirect.com/science/article/abs/pii/S0379073819302014>