Forensic Investigation Approaches of DNA Analysis and Criminal Investigation

Abhishek Sharma Padmanabhan , S. Sapna*

Department of Law, CHRIST (Deemed to be University), S.G. Palya, Bengaluru, Karnataka 560029, India

Corresponding Author Email: sapnas59969@gmail.com

ABSTRACT

Deoxyribonucleic acid (DNA) has been a significant factor in the criminal justice system since it was first used in forensic investigations. The reference sample's DNA profile is typically compared to the DNA profile from the evidence sample from the crime scene criminal cases. Familial DNA analysis can identify a person and provide significant investigation leads even without a reference sample for comparison in a criminal investigation process. The potential source of a forensic biological sample is determined using several indirect database searching techniques. These DNA-based techniques include Mitochondrial DNA (mtDNA) analysis, investigative genetic genealogy (IGG), familial searching, and Y-STR database searching. This study examined these methods and compares them in terms of searching efficiency, database structures, searching methods, genotyping technologies, data security, data quality, and costs. It also raises several possible legal and privacy problems for scientists to consider further. The significance of familial DNA analysis, the procedures used for finding and identifying relatives using familial DNA, and its benefits in forensics are all covered in this paper. Additionally, future options for the appropriate application of this technology and social, legal, and ethical concerns related to familial DNA analysis have been considered.

1. INTRODUCTION

Utilizing the DNA of biological family members to provide leads for identifying an unidentified person is a technique known as familial DNA analysis. Due to the intimate genetic relationships between individuals and others who may be considered close relatives, this information is crucial to the resolution of numerous criminal cases. Family DNA analysis has been an emerging area in criminal investigations recently. The familial relationship is determined by using familial DNA analysis rather than using traditional DNA match searches [1]. Similar matches are identified rather than exact matches at the same loci, the exact matches are not considered. This is used to determine which persons in the comparison are related to one another in the family. When there is no way to collect a reference sample to match the crime scene sample, familial DNA analysis may be the only option available in an investigation.

The underlying theory of familial DNA analysis states that individuals who are genetically related, including siblings, parents, or other close relatives, will share more alleles on analyzed loci than those who are unrelated. The significant investigation leads are provided by 'familial searching' for familial DNA analysis purposes in searching DNA databases where the lack of a reference sample prevented a comparison with the suspect's DNA [2].

When a suspect is unavailable for comparison, it is fairly usual in the course of the inquiry to find the perpetrator's DNA at the scene, consequently, finding any use for the DNA that was recovered from the crime scene as evidence becomes difficult. For comparison purposes, it is also challenging to identify the true offender because the DNA database does not contain the offender's DNA. Searching a database of DNA samples is the primary objective of family DNA testing that is now available to identify someone related to the offender in some way. It is commonly known that family members who are more closely connected are more likely than those who are not related to one another to share alleles at a specific locus [3].

Autosomal STR loci are often typed using forensic biological samples. From crime scenes, Biological forensic samples are collected. The Federal Bureau of Investigation (FBI) Combined DNA Index System (CODIS) has specified the 20 autosomal STR (A-STR) loci that make up the current core loci in the United States. These loci have a high power to discriminate. For "direct" match searching, upload the generated A-STR profile to the National DNA Index System (NDIS), State DNA Index Systems (SDIS), and or Local DNA Index Systems (LDIS), whether a mixture, a mixture's deconvoluted single source, it is known as a "hit" [4]. In the NDIS, although there are 18 million profiles for arrestees and offenders, the hit rates are only between 30 and 50% [5]. Therefore, there are still a lot of cases that haven't been solved. The NDIS contains approximately one million forensic profiles [6, 7].

The critical investigation leads to the resolution of heinous crimes made possible by DNA databases. Moreover, an insignificant portion of DNA database searches still involves family DNA analysis due to interpretation challenges with the database-sourced incomplete DNA profiles, ethical issues, and family DNA analysis's unclear legal implications. Additional measures are required in familial DNA searching rather than a direct comparison to identify the familial. If a public or commercial DNA database system contains the profiles of

relatives of the crime scene sample donor, one indirect method to find a forensic biological sample's sources is to search such profiles [8, 9]. Some techniques include family searching, mitochondrial DNA testing, investigative genetic genealogy (IGG), and Y-STR database searching. Any of these methods search for: 1) private companies or private citizens or private companies manage a genetic genealogy database (e.g., GEDMatch [10]), or a Government-run forensics database (such as NDIS [10]). Based on the information and technology available, further research incorporating genetic and/or non-genetic data is necessary for all three approaches. From a mixed sample, a search may also use a mixture profile [11]. A database search uses a single-source profile through a crime scene sample for descriptive reasons [12, 13]. To resolve cold or active cases, all three methods have been effective. In this study, various methodologies are described and compared regarding search methods, genotyping technologies, searching effectiveness, database architectures, data security, data quality, costs, and a few legal and privacy concerns are brought up for further debate.

2. DATABASES AND SOFTWARE USED IN FAMILIAL DNA ANALYSIS

Familial searching is the indirect database search method for providing leads for an investigation. It was founded using the existing government forensic laboratory systems and DNA database infrastructure. Because many criminals have relatives who have also been found guilty of crimes, familial searches are justified. Thus, DNA databases' lead value is increased and elevated above the direct matching success rates, especially local ones (such as SDIS and LDIS).

2.1 Combined DNA Index System (CODIS) and Familial DNA Searching

Additional stringency is required in familial DNA analysis than regular DNA testing processes for profile matching. Allele numbers from the DNA profile created by the Genetic Analyzer are added to CODIS compared to other currently included DNA profiles. Three levels of CODIS software can be employed depending on how stringently DNA profiles are matched: low, moderate, and high [14]. The partial matches are only produced by the low and middle stringency. For high stringency, the DNA profile's exact match is required. High stringency (exact match) DNA profiles are widely regarded as credible in criminal investigation processes. But having low stringency (partial match) capabilities can also be beneficial. When considering DNA's genetic inheritance (related people exhibit more resemblance than unrelated peoples). Additionally, CODIS is not intended for familial searches, thus over time, many software programs that enable efficient familial DNA analysis have been developed [15].

Numerous FDS software has been utilized for family searches to overcome these constraints. The familial searches are evaluated by these software using state, Likelihood ratio mixture of these two statistical methods. Comparatively to other methods, the FDS is advantageous for family DNA analysis due to additional features. Many nations, notably New Zealand, the United Kingdom (UK), and the United States (US), have employed FDS software to resolve various criminal cases. But because of several moral and legal concerns voiced by academics and legal organizations, the usage of FDS is constrained. This software is used to obtain the partial searches and must also be evaluated using statistical analysis with higher stringency policies [16].

2.2 GED match

Another database used for familial matching is GEDMatch, where autosomal DNA profiles are matched to data files from any further DNA analysis laboratory. The database must be updated with the already-created SNPs-based profile; the development of DNA profiles is not offered by GEDMatch. It outputs DNA segment chunks that show if the query profile matches or doesn't match the profiles already present in the database. With the capture of the culprit DeAgnelo in 2018, the database that was first created in 2010 gained notoriety. People use this database to locate family members or learn more about their parents. With the use of this software, various factors can be determined, including genetic-distance calculator, ethnicity, ethnicity calculator, and many more [17, 18]. Since Verogen, Inc. purchased this software, the emphasis has switched from preventing crimes to solving them.

2.3 Ancestry DNA

Some other program utilized for ancestry research is Ancestry DNA. When using this software, the person whose ancestry is being investigated provides samples. The software compares the DNA sequence, represented by SNPs, with the database's other ancestry DNA samples and outputs the results as an "ethnicity estimate" [19].

2.4 23 and Me

The first business to provide its consumers with genetic ancestry information was 23 and Me. A customer's saliva is collected by a specialized personal genomics company, which analyses the DNA and provides the consumer with health-related and ancestry information. Since 2018, this company has genotyped about three million individuals [20, 21].

3. METHODS USED IN FAMILIAL DNA ANALYSIS

3.1 Familial searching

Familial searching was the first indirect database search method for generating investigation leads. It was created using the government's existing DNA database infrastructure and the forensic laboratory system. Because a large percentage of criminals have relatives who have also been found responsible for their crimes, familial searches are justified. Thus, DNA databases' lead value can be greater than direct matching success rates, especially local ones like LDIS and SDIS [22].

From a crime scene sample, an arrestee or offender database is first searched against an A-STR profile created to find direct matches in familial searching (i.e., hits). If unsuccessful, potential first-degree kinship relationships prioritize similar but distinct profiles by applying metrics including shared allele probability and likelihood ratio. The donor of the forensic profile's close relative could be included in a list of potential matches with similar but distinguishable profiles [23].

Y-STR typing is performed on the list's top male candidates (for example, the top 100). Using additional non-genetic information, we only evaluate candidates with a matched Y-

STR profile for further analysis. Candidates are disqualified from being first-degree paternal relatives if their Y-STR profiles do not like those of the crime scene sample. As a result, the risk of incorporating unrelated people is reduced due to the two-step process, which reduces the possibility of privacy intrusion. The two-step process is, A-STR typing is done first, then Y-STR typing. To our knowledge, no male relative has ever been mistakenly identified using familial searches and reported as such. In the United States, strict guidelines, such as a list of specific and limited criminal types, an approval procedure, and an affirmative statement from particular stakeholders are often included that use familial searching [24].

3.2 Y-STR database searching

In government databases, familial searching uses huge collections of A-STR profiles and other manual search techniques. When no results were found from the direct A-STR search, another option would be to search for a Y-STR profile created if a Y-STR forensic database was available. A greater chance that matched Y-STR profiles belong to males from the same lineage depends on how many Y-STR loci make up the haplotype. In addition, there may be a strong likelihood that closely matched Y-STR profiles of more distant relatives. As a result, their A-STR profiles are further used for researching these matching Y-STR profiles' relatives in familial searching, if available, and other nongenetic information. The ability to extend among first relatives in the search when employing a Y-STR database search [25, 26].

The YHRD and forensic Y-STR databases are different. The profiles in YHRD are anonymous because it is used for research and assessing YHRD frequencies. Creating investigative leads is another use for a forensic Y-STR database. Furthermore, the Y-STR data found in open-access genealogy databases have helped with several successful Y-STR searching cases. Also, for determining the source of a sample, the effectiveness of Y-STRs is discussed in various studies [18].

3.3 Investigative Genetic Genealogy (IGG)

IGG is recently used by specific governmental organizations for identifying biological evidence, or human remains, mainly from cold cases. Even though CODIS's 20 core A-STR loci are highly effective at individuating a sample, they often cannot relate to relatives beyond the first degree because they lack the necessary authority and are occasionally even unhelpful for first-degree relatives [27]. The more distant relatives are located by requiring more genetic markers and typically, 50% of their alleles are shared by siblings who do not naturally share them. The familial DNA searching (FDS) process is given in Figure 1.

IGG, generated by whole-genome sequencing (WGS) and microarray, takes advantage of high-density SNP profiles. A high-density SNP profile database is searched against an SNP profile using measures like a pair of profiles combined with lengths of standard DNA segments to search for possible relatives in the database [28]. IGG approaches are capable of determining relationships of the third degree or higher with reasonable accuracy, far superior to the existing A-STR-based methods. IGG-specific policies are given in Table 1.
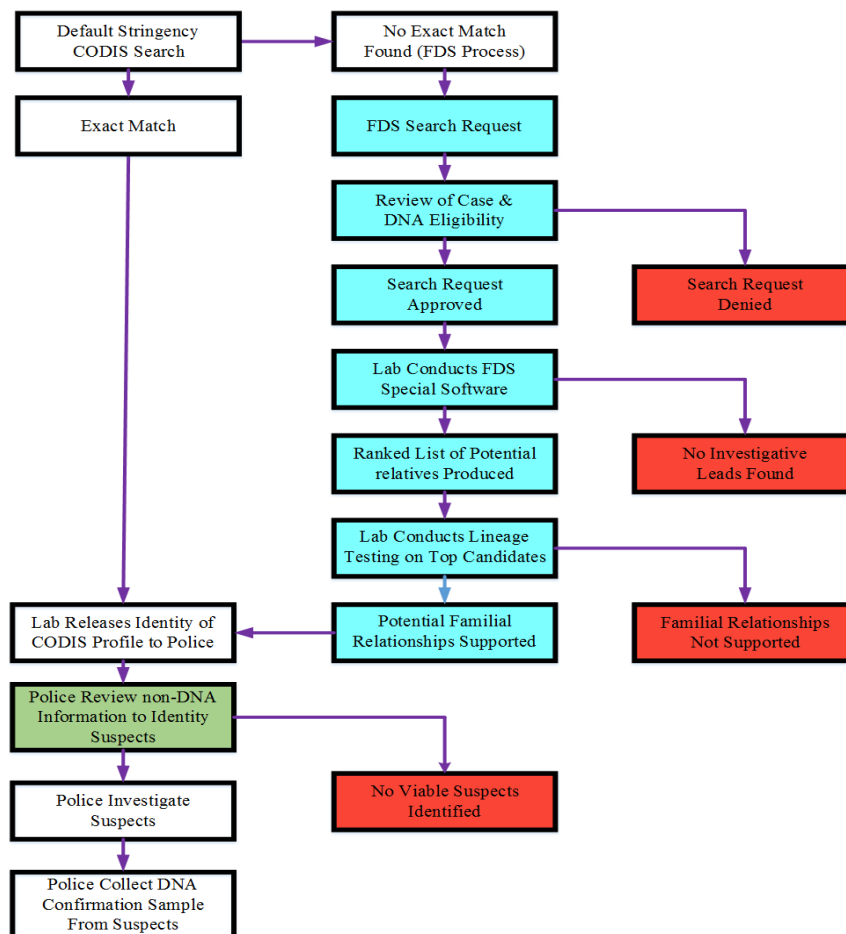

Figure 1. Familial DNA Searching (FDS) processes

Table 1. IGG-specific policies

Policy	Date	Applicability
Family-based CODIS search [7]	Various	While specific legal systems forbid it, others make searching CODIS for family members legal. Family search is permitted in some states, depending on the state protocols or policies. Familial searching is prohibited in certain states, but follow-up on accidental CODIS partial matches is permitted. Many states do not address the issue.
Genealogy Standards of the Board for Certification of Genealogists [29]	2019	Quality and integration of DNA data for inferring family histories, the evaluation of the DNA test results, and planning for DNA tests are all covered in Genealogy Standards' second edition.
A court order requires GEDmatch to allow searches of its public database [30]	November-2019	After the terms of usage were changed to require GEDmatch data producers to consent to law enforcement searches, in a case that was being investigated before the change, the complete data set might be searched, according to a Florida court's ruling.
Interim IGG policy from the US Department of Justice [31]	September-2019	By the interim policy, only unidentifiable remains and violent cases may use IGG as of November 2019, IGG-led DNA data cannot be uploaded to CODIS without having DNA data from the crime scene. After verification of a questionable match, demands IGG data removal from records
Surreptitious DNA sampling [21]	Various	Genetic data is viewed as property in some states, and its owners have certain rights (e.g., Georgia, Florida, Colorado, and Alaska). Other than those states that don't have applicable statutes, they all allow law enforcement to conduct covert DNA sampling. Law enforcement's compliance with state property restrictions hasn't been put to the test in court systems.
Best practices for consumer genomics from the Future Privacy Forum [32, 33]	July-2018	Guidelines for consumer genome firms forbid the distribution of DNA information to law enforcement without a valid court order or warrant and demand that businesses notify customers immediately after disclosing personal information. Companies must also report on demands for protected data made by law enforcement, according to these regulations.

Golden State Killer is one of the most famous IGG instances [27]. In this instance, the public database GEDMatch was uploaded with an extensive SNP profile created from a crime scene sample [28]. The GEDMatch search revealed a high quantity of prospective relatives, and several of those profiles revealed that third cousins or further removed relatives might be included in the database. The genealogy portion of the inquiry started after potential relatives were connected using DNA data. Investigators could focus their search on a specific person by thoroughly searching genealogical records based on distant relatives and other details. Once analyzed, the A-STR profiles obtained from the suspect and the evidence were determined to match. Figure 2 provides the basic steps of a criminal investigation that involves IGG.

Figure 2. Basic steps of a criminal investigation that involves IGG

3.3.1 Direct-to-consumer testing

The four primary DTC firms are described in the most recent discussions on genetic genealogy such as FTDNA, My Heritage, 23andMe, and AncestryDNA. Each firm offers insights into a person's health risks and/or ancestry using SNP microarrays, additionally the chance to discover connections to distant relatives who match for a predetermined minimum percentage of chromosomal regions. Putative IBD segments are detected by using slightly different approaches in each company, frequently without sharing all the specifics of how their algorithm is implemented. For a match to be declared, they each use different thresholds. 20% of matches are false positives because of the limitations of microarray technology [34]. Customers must opt-in for the majority of DTC's comparative analysis. Customers must opt-in for the majority of DTC's comparative analysis. Raw SNP data from AncestryDNA and 23andMe can be uploaded by users of My Heritage and FTDNA to increase the number of possible family connections [35]. The performance comparison of familial searching, Y-STR database searching, IGG, and Mitochondrial DNA Analysis are shown in Table 2.
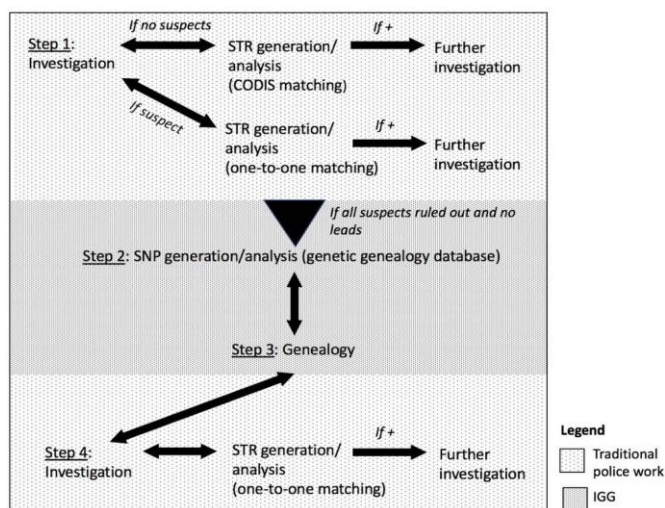
3.4 Mitochondrial DNA analysis

Numerous forensic uses of mtDNA analysis include missing person investigations, human remains identification, and criminal investigations. Only this kind of DNA may be obtained from specific sample types, including hair shafts. When comparing the same person's samples, different mtDNA sequences can result from mtDNA heteroplasmy within and between different tissue types. Based on sequences of the control region, forensic samples including hair, teeth, old bones, and other biological samples with low DNA contents are examined using whole genomic sequence analysis or mtDNA typing. Examining biological issues carefully is necessary for assessing and reporting the results [36].

An improved technique for obtaining putative GVP profiles and entire genome mtDNA from a single small hair sample. The procedure entails extracting hair proteins using urea, followed by protease digestion and buffer exchange. A 30 kDa membrane is used to elute peptides, which are then subjected to standard proteome analysis. The filter is then used to extract DNA, which is examined utilizing whole mt-genome analysis. A diverse cohort of 22 people's hair samples, including arm,

pubic, and head hair was used to validate the approach. 22 people had presumptive GVP profiles and mtDNA haplotypes that matched those in buccal swab samples from the same donor. Additionally, the method's usefulness was confirmed in two other labs. The technique can be used for mt-genome and proteomic-based GVP analysis in forensic research applications [37].

Most mtDNA-related research publications employ the control region's hypervariable regions I and II (HVR I and HVR II) databases, which are highly polymorphic and have the greatest amounts of variation and are therefore appropriate for identifying purposes. This makes the area suitable for utilizing analytical tools to infer genetic differences [38].

Table 2. Comparison of familial searching, Y-STR database searching, IGG, and Mitochondrial DNA analysis

	IGG	Y-STR database searching	Familial searching	mtDNA analysis
Expertise	DNA analysts, genealogists, and Service lab	DNA lab analysts, Recognized laboratories	DNA lab analysts, Recognized laboratories	DNA lab analysts, Recognized laboratories
Genotyping technology	Forensic laboratories have not verified SNP panels with a high density such as whole-genome sequencing or microarray	Validation of Y-STR and A-STR by forensic laboratories	Validation of Y-STR and A-STR by forensic laboratories	A-STR validated by forensic laboratories
Database to search	DNASolves, Family tree, GEDmatch, etc.	Forensic Y-STR database	Forensic A-STR database (e.g., NDIS)	Forensic A-STR database (e.g., HV1 and HV2)
DNA quantity	Parabon NanoLabs suggests >10 ng; Bode suggests >20 ng; At least 1 ng; potential data leaks; more susceptible to assaults; less secure	Lower input may be effective; 500-1000 pg is advised.	Lower input may be effective; 500-1000 pg is advised.	Lower input may be effective; 500-1000 pg is advised.
Data security		Secured	Secured	Secured
Methods to measure the relationship	IBD and IBS segment's overall length	Number of mismatched steps	Number of shared alleles; Likelihood ratio	Number of mismatched steps; Likelihood ratio
Data quality	Unidentified profiles	Identified profiles	Identified profiles	Unidentified profiles
Relationships	According to the number of SNPs, is capable of detecting third-degree or more distant relationships	Those males who share the same ancestry	Relations of the first degree (full sibling, parent-child)	Relations of the first degree (full sibling, parent-child)
Time to results	Primarily as a result of intensive non-DNA genealogical searching, weeks to months	Days to weeks	Days to weeks	Family Finder – 2 to 4 weeks. mtFull Sequence – 6 to 8 weeks
Advantages	Finding female relatives and distant relatives is possible.	Associate to surnames; high-performance male lineage search; safe and secured database; Tested and validated for proficiency; Existing technologies	A protected and secure database; Tested and validated for proficiency; existing legislation; Existing database and technologies	It enables forensic scientists to gather information from small evidence with minimal biological content and outdated evidence related to cold cases.
Disadvantage	Unlike previous methods, this one generates much more intrusive private information, possibly insecure data; is cost-prohibitive for routine analyses; significantly uses public SNP databases and genealogists; forensically unreliable technology	May require the help of genealogists for distant matches; building a Y-STR database at a high initial cost if all reference profiles are retyped; identifying only male ancestry relatives	Uses Y-STR filtering as a basis; large candidate list after searching	The individuals of the same maternal lineage are indistinguishable by mtDNA analysis. lower discrimination power and lower discrimination power
Cost of DNA tests	$3,000 to $10,000 for each IGG service case: additionally, forensic laboratories' A-STR typing	Based on the number of A-STR profiles, $100 per case.	Based on the various Y-STR profiles typed during filtering, $1,000 per case.	Based on the number of A-STR profiles typed during filtering, each case will cost $399.

3.5 Genotyping technologies

A-STR typing and STR profiles are required for both family searching and Y-STR database searches, respectively. A few additional common STR loci are often present in an A-STR reference profile. Depending on the commercial kit, the number of Y-STR loci in a Y-STR profile can range from 23 to 40. Thermo Fisher's S5 and Illumina's MiSeq, two recent examples of massive parallel sequencing (MPS) technology, have recently improved the precision and throughput of STR allele analysis and calling. Due to the ability to transform sequence data into length-based allele data, CE-based STR typing outcomes and STR's MPS are backward-compatible in global DNA databases. Although 500–1,000 pg of DNA is the recommended maximum input for STR typing methods, less DNA can yield a complete DNA profile [25, 39].

Microarray or WGS technologies can produce high-density SNP profiles, which IGG needs. Own SNP data is provided by the largest direct-to-consumer genetic testing businesses, such as My Heritage, Ancestry.com, and 23andMe. SNP genotyping is typically contracted out to private service laboratories by the major IGG providers including DNA Doe Project, Bode Technology, and Parabon NanoLabs. Due to the accuracy (99.9%) and affordability of microarray genotyping, these service laboratories frequently use them to create SNPs. Determining sufficient information is provided by the genotyping accuracy of distant relationships with different confidence degrees [40].

Although the forensic laboratories have not verified the microarray or the WGS technologies, the existing service providers have not disclosed any supporting data for those procedures. Currently, each association formed by IG must receive direct A-STR matching confirmation to be accepted by the US Department of Justice. This confirmation method may eliminate part of the IGG validation criteria when appropriate, but considerations for sample consumption, privacy, and any probable cause issues must still be made [41, 42].

3.6 Implications of familial DNA analysis

Recent years have seen a huge increase in the value of familial DNA analysis. Significant investigational clues are provided by the familial DNA analysis in the absence of any standardized reference samples [23]. During inquiry processes, this circumstance arises in a wide range of situations. From the crime scene, numerous crucial investigation leads have been generated due to partial DNA matches between samples and those in the database. This, in those cases, finally led to the culprit's being found.

Joseph James DeAngelo, the 'Golden State Killer,' is a good example of the significance of family DNA research. The DeAngelo had been involved in over fifty cases of sexual assault and several killings over two decades, the 1970s and 1980s. The lack of the suspect's DNA made it challenging to positively identify him, even though his profile was found on numerous pieces of evidence gathered at the crime scene. However, after a forty-year hiatus, a breakthrough finally materialized when Californian law enforcement officers searched the GEDMatch genealogy database for the DNA profile. It was discovered that the profile match belonged to DeAngelo's third cousin. DeAngelo was identified through family tree research [43, 44]. The offender was apprehended as a result of the investigation, and this case resulted in the global expansion of numerous ancestry-based databases.

3.7 Difficulties of familial DNA analysis

Familial DNA analysis is one method that can result in false positives. Comparing two DNA profiles, this method depends on the partial matches discovered, resulting in people unrelated to criminal activity. There isn't a consistent policy regulating familial DNA analysis utilization [45, 46].

The subject of "security" is also crucial. The DNA profile is not secure since there are no suitable guidelines for entering and evaluating DNA profiles in the various databases that are available for comparison. The risk of disclosure of a person's DNA profile information has increased with familial DNA searching. Additionally, a particular DNA profile is checked for matches across several DNA databases, this may compromise a person's DNA profile's security [47].

Another essential consideration for dealing with familial DNA matching is statistical analysis. When determining a "hit," further investigation is necessary after analyzing the partial matches discovered through DNA comparison during family searches. A few adjustment factors should be included when applying some statistical computation techniques. Familial DNA evidence may be misinterpreted and given an incorrect weight if the aforementioned statistical computation procedures are unchanged [26].

3.8 Legal and ethical challenges with familial DNA analysis

There are several ethical concerns with the expanding use of family DNA analysis. It may be possible to investigate more people because family DNA searching depends on probability estimates and partial matches. In addition, racial bias is prevalent in several databases used for familial DNA analysis, holding more information about people of a specific color than about people of other colors. There are more DNA profiles in most criminal investigation databases from criminal offenders than non-criminals. A biased, partial hit may result from this [48].

The growth of databases for familial searches in recent years has been predicated on inequity, meaning that individuals of a particular ethnicity or those convicted of a crime are more likely to have relatives included in the investigation than others. This might result in the suspect being incorrectly identified, which would then lead to the person being wrongfully convicted [17]

Familial DNA analysis raises essential legal and moral questions when considering family structures and how they handle their relationships. The DNA profile of a person that has been added to the database presents significant issues because it may affect the person's family and relations. Furthermore, it can violate the person's right to "not know" what is occurring with his DNA profile in the database and how it would affect his or her family. Additionally, some genomic loci connected to genetic conditions have been discovered. Probabilistic probabilities of such disorders being inherited from parents and passed on to kids significantly impact a person's life. The DNA database's DNA profile thus raises several consequences for a person's fundamental rights to consider [49].

3.9 Future aspects of familial DNA analysis

Recent years have seen success in using family DNA analysis in criminal investigations. Comparatively, this forensic DNA analysis requires more strict conditions. The

capability of family DNA analysis for searching can be improved by using probability calculating software by the society in consideration. The Y-STR analysis, mitochondrial analysis, and the application of SNPs can all be combined with partial matches discovered by DNA comparison. Future solutions to two issues can be achieved by utilizing numerous DNA comparison approaches, such as decreased error risk, and numerous familial DNA analysis-related moral, legal and societal concerns can be reduced.

4. DISCUSSION

A crucial component of IGG and typically the most time-consuming step in the procedure is genealogy research, even though research time will vary depending on a variety of circumstances, such as how closely the matches match, the availability of genealogical records, family size, and the supporting network of matches. It took volunteer teams hundreds of hours to research certain cases for the DNA Doe Project. Only because of the vast amounts of global genealogy records, IGG is feasible, which have been digitized and indexed during the past two decades. In a wide range of crimes, human identification-related issues are effectively resolved by IGG, but it still faces some difficulties since it interferes with family and personal relationships, poses privacy hazards, and is not always correct.

The Y-STR database search allowed for the highly accurate identification of the relationships among several generations. Comparing family searching to genealogical research, it is also possible to identify more distant male ancestors for generating higher extensive investigative leads. The Y-STR database search has the significant drawback of not providing information on maternal relatives. However, maternal lineage information could be obtained through mitochondrial DNA sequencing if needed. Most criminal cases of interest should benefit from the Y-STR database because men conduct the majority of crimes, especially violent ones. This is especially true for several sexual assault cases with unreliable results from A-STR typing and generating a single-source Y-STR profile is possible.

The number of human identification-related cases such as missing person cases, mass disasters, acts of terrorism, and violent crimes are effectively solved by widely using mtDNA typing throughout the world over the last 25 years. There are still specific issues with the admissibility of mtDNA testing in court, particularly those involving the heteroplasmy problem and, more recently, the potential for biparental inheritance. A thorough understanding of the molecular processes behind mtDNA biparental inheritance, the capacity to predict the circumstances under which this is most likely to take place, and the capacity to precisely identify and describe heteroplasmy are essential issues, these significant challenges must be resolved by using mtDNA analysis, it remains a valuable and viable tool for human forensic identification.

In criminal cases like identifying missing persons and violent crimes, the possible investigative leads are extended by using searching DNA databases, which is a successful and effective approach for crime investigation. For using DNA databases to find relatives, this study discusses the use of IGG, familial searching, and Y-STR database searching as indirect investigation approaches. From different perspectives, the details of these techniques are compiled in Table 1. The value of DNA databasing can be increased in three ways, all of

which are practical and should be considered. IGG, mtDNA, and Y-STR database searches are advantageous since they should be able to help find distant relatives in more situations than familial searching. However, using at least state-level databases and existing laboratory systems makes familial searching easier to carry out. Every nation or state should choose the most appropriate strategies for its own cultural, technical, legal, and economic conditions. Whatever method is used, the result should be the same: create leads for investigations and close open and closed criminal cases.

5. CONCLUSION

The application of family DNA analysis in forensic science has given the criminal investigation system a new dimension. When there is no available reference sample for comparison, it has assisted law enforcement authorities in locating a suspect. It is possible to quickly compare different profiles through various private and government-founded databases and laboratories. But under strictly controlled conditions, the procedure should be carried out. False convictions may occur from incorrect data being interpreted because DNA profile comparison only yields partial matching results. Additionally, when using family DNA analysis for criminal investigation, some moral, social, and legal concerns still need to be addressed, prompting calls for consistent legalization and the policymaking of this procedure.

When identifying distant relatives, mtDNA analysis, IGG, and Y-STR database searching, are more helpful than familial searching since they can help in more situations. However, with current laboratory technologies, family searching is easier. Every nation or state should choose the most appropriate strategies for its own legal, cultural, technical, and economic circumstances. The overall objective should be the same regardless of the strategy, to ensure public safety, develop leads for investigations and resolve open and closed crime cases while adhering to strict security guidelines meant to preserve citizens' privacy.

REFERENCES

[1] Debus-Sherrill, S., Field, M.B. (2019). Familial DNA searching-an emerging forensic investigative tool. Science & Justice, 59(1): 20-28. https://doi.org/10.1016/j.scijus.2018.07.006

[2] Daeid, N.N., Hackman, L., Haddrill, P.R. (2021). Developments in forensic DNA analysis. Emerging Topics in Life Sciences, 5(3): 381-393. https://doi.org/10.1042/ETLS20200304

[3] Fortier, A.L., Kim, J., Rosenberg, N.A. (2020). Human-genetic ancestry inference and false positives in forensic familial searching. G3: Genes, Genomes, Genetics, 10(8): 2893-2902. https://doi.org/10.1534/g3.120.401473

[4] Peng, D., Zhang, Y., Ren, H., Li, H., Li, R., Shen, X., Wang, N., Huang, E., Wu, R., Sun, H. (2020). Identification of sequence polymorphisms at 58 STRs and 94 iiSNPs in a Tibetan population using massively parallel sequencing. Scientific Reports, 10(1): 12225. https://doi.org/10.1038/s41598-020-69137-1

[5] Myers, S. (2021). Searching CODIS with binary conversions of STRmix interpretations. Forensic Science

International: Genetics, 55: 102569. https://doi.org/10.1016/j.fsigen.2021.102569

[6] Edge, M.D., Coop, G. (2020). Attacks on genetic privacy via uploads to genealogical databases. Elife, 9: e51810. https://doi.org/10.7554/eLife.51810

[7] Stanciu, F., Dragomir, E., Bratu, T. (2021). Solving the first murder case in Romania using familial searching. Romanian Journal of Legal Medicine, 29(3): 295-298. https://doi.org/10.4323/rjlm.2021.295

[8] Formici, G. (2021). From "familial searching" to "forensic genetic genealogy": New frontiers–and challenges–of DNA analysis in criminal investigations. BioLaw Journal-Rivista di BioDiritto, (1S): 305-328. https://doi.org/10.15168/2284-4503-788

[9] Kaiser, J. (2018). We will find you: DNA search used to nab Golden State Killer can home in on about 60% of white Americans. Science.

[10] Federal Bureau of Investigation. CODIS - NDIS statistics. https:// www.fbi.gov/services/laboratory/biometric-analysis/codis/ndisstatistics, accessed on 11 Sept. 2020.

[11] GEDMatch. https://www.gedmatch.com/, accessed on 11 Sept. 2020.

[12] de Vries, J.H., Kling, D., Vidaki, A., Arp, P., Kalamara, V., Verbiest, M.M., Piniewska-Róg, D., Parsons, T.J., Uitterlinden, A.G., Kayser, M. (2022). Impact of SNP microarray analysis of compromised DNA on kinship classification success in the context of investigative genetic genealogy. Forensic Science International: Genetics, 56: 102625. https://doi.org/10.1016/j.fsigen.2021.102625

[13] Kim, M.Y., Cho, S., Lee, J.H., Seo, H.J., Lee, S.D. (2018). Detection of innate and artificial mitochondrial DNA heteroplasmy by massively parallel sequencing: Considerations for analysis. Journal of Korean Medical Science, 33(52): e337. https://doi.org/10.3346/jkms.2018.33.e337

[14] Yudianto, A., Setiawan, F. (2020). The effectiveness of mini primer STR CODIS in DNA degradation as the effect of high-temperature exposure. Analytical Cellular Pathology, 2020: 2417693. https://doi.org/10.1155/2020/2417693

[15] Crouse, C.A., Bauer, L., Sessa, T., Looper, A., Sikorsky, J., Yeatman, D. (2019). Combined DNA Index System (CODIS)-based analysis of untested sexual assault evidence in Palm Beach County Florida. Forensic Science International: Synergy, 1: 253-270. https://doi.org/10.1016/j.fsisyn.2019.09.005

[16] Piquado, T., Matthies, C.F., Strang, L., Anderson, J.M. (2019). Forensic familial and moderate stringency dna searches. Research Reports. https://doi.org/10.7249/RR3209

[17] Ney, P., Ceze, L., Kohno, T. (2020). Genotype Extraction and False Relative Attacks: Security Risks to Third-Party Genetic Genealogy Services Beyond Identity Inference. In NDSS.

[18] Erlich, Y., Shor, T., Pe'er, I., Carmi, S. (2018). Identity inference of genomic data using long-range familial searches. Science, 362(6415): 690-694. https://doi.org/10.1126/science.aau4832

[19] Jin, S., Chase, M., Henry, M., Alderson, G., Morrow, J.M., Malik, S., Ballard, D., McGrory, J., Fernandopulle, N., Millman, J., Laird, J. (2018). Implementing a biogeographic ancestry inference service for forensic casework. Electrophoresis, 39(21): 2757-2765.

https://doi.org/10.1002/elps.201800171

[20] Panofsky, A., Donovan, J. (2019). Genetic ancestry testing among white nationalists: From identity repair to citizen science. Social Studies of Science, 49(5): 653-681. https://doi.org/10.1177/0306312719861434

[21] Kennett, D. (2019). Using genetic genealogy databases in missing persons cases and to develop suspect leads in violent crimes. Forensic Science International, 301: 107-117. https://doi.org/10.1016/j.forsciint.2019.05.016

[22] Kitnick, J. (2019). Killer's code: Familial DNA searches through third-party databases under carpenter. Cardozo L. Rev., 41: 855.

[23] Karantzali, E., Rosmaraki, P., Kotsakis, A., Le Roux-Le Pajolec, M.G., Fitsialos, G. (2019). The effect of FBI CODIS Core STR Loci expansion on familial DNA database searching. Forensic Science International: Genetics, 43: 102129. https://doi.org/10.1016/j.fsigen.2019.07.008

[24] Shen, X., Li, R., Li, H., Gao, Y., Chen, H., Qu, N., Peng, D., Wu, R., Sun, H. (2021). Noninvasive prenatal paternity testing with a combination of well-established SNP and STR markers using massively parallel sequencing. Genes, 12(3): 454. https://doi.org/10.3390/genes12030454

[25] Li, M., Zhou, W., Zhang, Y., Huang, L., Wang, X., Wu, J., Meng, M., Wang, H., Li, C., Bian, Y. (2020). Development and validation of a novel 29-plex Y-STR typing system for forensic application. Forensic Science International: Genetics, 44, 102169. https://doi.org/10.1016/j.fsigen.2019.102169

[26] Fan, G., Pan, L., Tang, P., Zhou, Y., Liu, M., Luo, X. (2021). Developmental validation of a novel 41-plex Y-STR system for the direct amplification of reference samples. International Journal of Legal Medicine, 135: 409-419. https://doi.org/10.1007/s00414-020-02326-9

[27] Wickenheiser, R.A. (2019). Forensic genealogy, bioethics and the Golden State Killer case. Forensic Science International: Synergy, 1: 114-125. https://doi.org/10.1016/j.fsisyn.2019.07.003

[28] Greytak, E.M., Moore, C., Armentrout, S.L. (2019). Genetic genealogy for cold case and active investigations. Forensic Science International, 299: 103-113. https://doi.org/10.1016/j.forsciint.2019.03.039

[29] Board Certif. Geneal. (2019). Genealogy Standards. Washington, DC: Ancestry.com. 2nd ed.

[30] Hill, K., Murphy, H. (2019). Your DNA profile is private? A Florida judge just said otherwise. New York Times.

[31] US Dep. Justice. (2019). Interim policy: forensic genetic genealogical DNA analysis and searching. Policy Doc., US Dep. Justice, Washington, DC. https://www.justice.gov/olp/page/file/1204386/download.

[32] Martinez, C. (2018). Privacy best practices for consumer genetic testing services. Future of Privacy Forum, July 31. https://fpf.org/2018/07/31/privacy-best-practices-for-consumer-genetic-testingservices.

[33] Carr, N. (2020). Cold cases freeze: law enforcement locked out of DNA database used for investigative genealogy after consumers object to being genetic informants. Washington University Law Review, 98.

[34] Petter, E., Schweiger, R., Shahino, B., Shor, T., Aker, M., Almog, L., Weissglas-Volkov, D., Naveh, Y., Navon, O., Carmi, S., Li, J.H., Berisa, T., Pickrell, J.K., Erlich, Y.

(2020). Relative matching using low coverage sequencing. bioRxiv, 2020-09. https://doi.org/10.1101/2020.09.09.289322

[35] Kling, D. (2019). On the use of dense sets of SNP markers and their potential in relationship inference. Forensic Science International: Genetics, 39: 19-31. https://doi.org/10.1016/j.fsigen.2018.11.022

[36] Mason, K.E., Paul, P.H., Chu, F., Anex, D.S., Hart, B.R. (2019). Development of a protein-based human identification capability from a single hair. Journal of Forensic Sciences, 64(4): 1152-1159. https://doi.org/10.1111/1556-4029.13995

[37] Catlin, L.A., Chou, R.M., Goecker, Z.C., Mullins, L.A., Silva, D.S.B.S.S., Spurbeck, R.R., Parker, G.J., Bartling, C.M. (2019). Demonstration of a mitochondrial DNA-compatible workflow for genetically variant peptide identification from human hair samples. Forensic Science International: Genetics, 43: 102148. https://doi.org/10.1016/j.fsigen.2019.102148

[38] Govender, P., Fashoto, S.G., Maharaj, L., Adeleke, M.A., Mbunge, E., Olamijuwon, J., Akinnuwesi, B., Okpeku, M. (2022). The application of machine learning to predict genetic relatedness using human mtDNA hypervariable region I sequences. Plos One, 17(2): e0263790. https://doi.org/10.1371/journal.pone.0263790

[39] Mo, X.T., Zhang, J., Ma, W.H., Bai, X., Li, W.S., Zhao, X.C., Ye, J. (2019). Developmental validation of the DNATyper™ Y26 PCR amplification kit: an enhanced Y-STR multiplex for familial searching. Forensic Science International: Genetics, 38: 113-120. https://doi.org/10.1016/j.fsigen.2018.10.008

[40] Dagnall, C.L., Morton, L.M., Hicks, B.D., et al. (2018). Successful use of whole genome amplified DNA from multiple source types for high-density Illumina SNP microarrays. BMC Genomics, 19(1): 1-10. https://doi.org/10.1186/s12864-018-4572-6

[41] Tillmar, A., Fagerholm, S.A., Staaf, J., Sjölund, P., Ansell, R. (2021). Getting the conclusive lead with investigative genetic genealogy–A successful case study of a 16 year old double murder in Sweden. Forensic Science International: Genetics, 53: 102525. https://doi.org/10.1016/j.fsigen.2021.102525

[42] Daeid, N.N., Hackman, L., Phillips, C., de la Puente, M. (2021). The analysis of ancestry with small-scale forensic panels of genetic markers. Emerging Topics in Life Sciences, 5(3): 443-453. https://doi.org/10.1042/ETLS20200327

[43] Starinsky-Elbaz, S., Ariel, T., Issan, Y. (2019). DNA kinship analysis of unidentified human remains that led to a murder investigation. Forensic Science International, 300: e20-e23. https://doi.org/10.1016/j.forsciint.2019.03.045

[44] Wickenheiser, R.A. (2019). Forensic genealogy, bioethics and the Golden State Killer case. Forensic science international: Synergy, 1: 114-125. https://doi.org/10.1016/j.fsisyn.2019.07.003

[45] Kruijver, M., Taylor, D., Bright, J.A. (2021). Evaluating DNA evidence possibly involving multiple (mixed) samples, common donors and related contributors. Forensic Science International: Genetics, 54: 102532. https://doi.org/10.1016/j.fsigen.2021.102532

[46] Murphy, E., Tong, J.H. (2020). The racial composition of forensic DNA databases. Calif. L. Rev., 108: 1847.

[47] Snedecor, J., Fennell, T., Stadick, S., Homer, N., Antunes, J., Stephens, K., Holt, C. (2022). Fast and accurate kinship estimation using sparse SNPs in relatively large database searches. Forensic Science International: Genetics, 61: 102769. https://doi.org/10.1016/j.fsigen.2022.102769

[48] Oosthuizen, T., Howes, L.M. (2022). The development of forensic DNA analysis: New debates on the issue of fundamental human rights. Forensic Science International: Genetics, 56: 102606. https://doi.org/10.1016/j.fsigen.2021.102606

[49] McGlynn, K.E. (2019). Remedying wrongful convictions though DNA testing: Expanding post-conviction litigants' access to DNA database searches to prove innocence. BCL Rev., 60: 709.