

Forensic Genealogy: A Powerful Investigation Tool for Resolving the Cold Cases

Muhammad Shahid Cholistani^{1,*}, Muhammad Hashim Raza¹ and Muhammad Farhan¹, Bisma Ashraf², Hifz Ur Rehman¹, Muhammad Arjamand¹, Muhammad Luqman¹, Abdul Basit¹, Sabahat Shakoor¹, Areeba Saleem¹ and Muhammad Aleem Ashraf³

¹Institute of Biological Sciences, Khawaja Fareed University of Engineering and Information Technology Rahim yar khan, Pakistan

²Department of Chemistry Government College University, Lahore, Pakistan

³Department of Biosciences and Technology, Emerson University, Multan, Pakistan

***Corresponding Author:** Muhammad Shahid Cholistani, Institute of Biological Sciences, Khawaja Fareed University of Engineering and Information Technology, Rahim yar khan, Pakistan, Tel.: +923006747514, E-mail: shahidcholistani@gmail.com

Citation: Shahid Muhammad Cholistani, Muhammad Hashim Raza, Muhammad Farhan (2024) Forensic Genealogy: A Powerful Investigation Tool for Resolving the Cold Cases, J Forensic Crime Stu 12(3): 301

Received Date: September 08, 2024 **Accepted Date:** October 08, 2024 **Published Date:** October 14, 2024

Abstract

Law enforcement has lately used publicly available genealogical databases to conduct investigations into unresolved cases. In addition, informal discussions with genealogical specifics have not impeded the exchange of information with the eventual objective of disclosure. This review article will examine several real-life illustrations and their corresponding circumstances, in addition to the latest breakthroughs in forensic genealogy and their potential trajectories. Furthermore, it will provide a thorough understanding of how genealogy is currently used in forensic science across many countries. This study aims to enhance our comprehension of different approaches, occurrences, regulations, and other facets of crime scenes that have occurred in the past or may occur in the future. Therefore, it is crucial to build a proficient legal system and strong public safety measures in order to prevent crime. These measures will allow criminals to be captured and convicted, ensuring justice for both wrongdoers and their families.

Keywords: Forensic Genealogy; crime scene investigation; death investigation

Introduction

Forensic Genetic Genealogy (FGG) and Investigative Genetic Genealogy (IGG) have developed effective forensic tools for identifying unidentified human remains and identifying offenders who have evaded detection. Many genetic profiles in IGG include community pedigree DNA from databases, of which there are some tens of thousands. Single-nucleotide polymorphisms (SNPs) are present. These were traced by matching shared DNA fragments that represent biological links to an unknown donor from the 19th century. The key to success is that only a subset of the recreational population wants to be included in the record, theoretically enabling genetic genealogy techniques to identify either sex if he is in the area. According to Ehrlich et al., there is a better than 90% probability that any offspring in the population will have at least a third cousin if 1% of the population contributes to a family tree DNA database [1].

In recent years, DNA datasets have become an important tool for family history research. With an estimated [2, 3] 30 million people in genealogy databases, the likelihood that individuals will find a third or closest sibling through DNA-based searches increases. These advances have made these data sets especially alluring to policing, particularly in the US, and a few data sets currently explicitly permit policing by allowing the transfer of DNA profiles to help with missing individuals or serious wrongdoing cases. Here, policing commonly employs a hereditary genealogist who works with any DNA matches recovered after the transfer of DNA profiles, creating genealogical records, and attempting to identify the individual in question or suspect of interest through their network of cousin matches [4]. More than 200 US law enforcement cases have utilized this method, which is also known as forensic genetic genealogy and investigative genetic genealogy (IGG), with some notable investigative successes. Sweden has utilized IGG to settle a serious homicide case Erlich. The majority of IGG searches currently use the specific protocols for law enforcement matching established by the US-based GED match and Family Tree DNA databases. DNA Tackles, a third US data set, allows clients to transfer their information exclusively for policing [5].

Forensic Genetic Genealogy (FGG) has rapidly gained popularity as a valuable technique in criminal investigations since its inception in 2018. FGG is an innovative investigation tool that has been used in numerous unsolved cold cases in the United States to produce investigative leads and determine the identity of unidentified individuals. The popularity of consumer DNA testing and the growing interest in personal DNA and genetic genealogy have significantly expanded the accessibility of human genetic data. The area of genetic genealogy has attracted the attention of both amateur and professional genetic genealogists for some years. They utilize consumer DNA data to investigate genetic relationships within family trees [6].

FGG has predominantly been employed to resolve cases related to serial and sexual violence against women and vulnerable individuals, as well as those involving victimization by unknown individuals - cases that have historically presented greater challenges in terms of resolution. Approximately 80% of the victims were specifically targeted for sexual violence, while over 28% belonged to social categories that exhibit heightened vulnerability to criminal and sexual exploitation [7]. Approximately 79% of the individuals accused of a crime and 48% of the individuals who were harmed were of European descent. However, the ancestral background of numerous victims remains unidentified due to their status as victims of sexual assault [8]. FGG investigations in the U.S. were mainly carried out in lower-income counties and had a higher representation in rural regions, while being less utilized in the 50 major metropolitan areas. Out of all the cases handled by the ten major police agencies in the U.S., just 2% were successfully resolved and a total of 5 suspects were identified. The mean duration for FGG to resolve a case was determined to be 12.1 months [9-11].

Genotyping Techniques for Molecular Genes

Since the beginning of 1980, advancements in medicinal applications and DNA inquiry have evolved simultaneously. At first, we enhanced the finite component length polymorphism and fine-tuned it using Southern blotting. Short-paired repeats

(STRs) discovered by PCR were included. The solution emerged when the development of thin electronics equalized the opportunities after the fluorescence localization of PCR fragments. In the past, court systems struggled to keep pace with the swift progress in technology [12]. Nevertheless, this choice remained dormant until the advent of capillary electrophoresis, which made STR identification possible. The incorporation of DNA into the less dependable realm of forensic research was embraced as indispensable, in juxtaposition to the analytical constraints of current forensic techniques. In contrast to clinical or scholarly professions, which rely on SNP fragments for genetic code investigations, forensics has not incorporated SNPs as regular technical tools. Short tandem repeats (STRs) are more effective than single nucleotide polymorphisms (SNPs) in generating a highly accurate genetic profile from a small amount of DNA, which is typically the case with limited samples obtained from crime scene documentation. By employing a compact central configuration of Short Tandem Repeats (STRs), we also considered fundamental data while dividing the wards. Instead of analyzing a subset of one million Single Nucleotide Polymorphisms (SNPs), a single ward might transmit a list of hetero sequence widths for the typical 13 Short Tandem Repeats (STRs) via fax [13].

Additionally, a single extremely diverse tetra nucleotide repeat provides a significantly greater number of repeat variants compared to a single SNP with only two to four types. Achieving a distinct genetic profile necessitates multiplying fewer loci due to STRs' variability. Additionally, selecting fewer loci could disperse the selected markers across the genome, thereby reducing the likelihood of a reciprocal relationship [14]. Moreover, the use of a small number of indicators simplified the process of transferring DNA data between sources, showcasing the potential of newly developed technologies to scan these markers and provide likely implications for each one. Furthermore, the considerable diversity of the indicators suggests that linking an individual repeat-length gene to any attribute or disease is unlikely. Compared to SNPs, a number of these STR-related factors are less diverse and easier to link to specific characteristics and situations. In light of these considerations and the updated framework for debating the ethical implications of SNP use, we must reevaluate our next steps now that this study is applying SNP genetic analysis to challenging scenarios [15, 16].

DNA Databases for Direct and Partial Matching

The National DNA Index System (NDIS), the State DNA Index System (SDIS), and the Local DNA Index System (LDIS) are the three tiers that make up CODIS, permitting every hierarchy to operate within the boundaries of its respective legislation. Accordingly, depending on the nature and gravity of an offense, some locations may permit or prohibit searching for culpable parties' DNA. Subsequently, the framework sorts the records using indicators primarily associated with criminal records, criminals, captives, mysterious remains, people who have vanished, and relatives of missing individuals. The relevant jurisdictional rules and regulations permit the matching of DNA samples from crime scenes to other forensic samples after their inclusion in the forensic database, offender data, and so on, first at the LDIS level and then at the SDIS and NDIS levels [16]. CODIS was intended designed to facilitate the interfacing of violations with each other and to distinguish suspects through immediate, careful matches. However, it also considered less severe pursuits to require DNA information from crime location for example, a brief DNA sample could only genotype 30 of the 40 potential alleles at 20 loci. potential alleles at 20 loci. Furthermore, the use of CODIS to distinguish deceased people implied that calculations should also allow some degree of connection investigation. Therefore, CODIS allows for less stringent searches to identify close biological relatives, preferring a parent, child, or full sibling [17].

Investigative Genetic Genealogy

Law enforcement noted the emergence of GED match and commercial genome firms because they saw the possibility for exploiting the data to solve unresolved crimes. With DNA validation. Analysts utilized a court order to get the girl's Pap smear sample of a person of interest in a 2008 investigation into the BTK continual murderer. Her DNA results were sufficient for the court to issue a warrant for her dad's arrest and strong guilt once the evidence was forwarded to the university medical center

[18]. This started a pattern wherein suspect identification might be made using DNA information other than CODIS. In 2014, investigators tried to use Ancestry.com's Y-STR data to identify an individual. Culprit in the Angie Dodge homicide investigation. After creating a profile of 35 Y-STRs using semen taken from the scene of the crime, they were notified by Ancestry.com to investigate the STR profile using their database of Y-chromosome information. A contest relating specialists to New Orleans, Louisiana-based movie producer Michael Usry Jr. was held in 34 out of 35 venues. Usry was compelled by a court order to provide a DNA sample, which eliminated him as a suspect [19].

Genealogy Databases

The imprisonment of Joseph James DeAngelo on April 24, 2018 has created a serious privacy crisis in civilization. DeAngelo is considered the Golden State Killer, who committed over thirteen murders and fifty rapes. The case took a turn for the better when detectives matched DNA obtained at homicide and death scenes with other DNA data in open genealogical databases GED match, this is the collection of DNA samples for forensic qualification in the National DNA Index System (NDIS) (in is not equal) since 1989, People who have been arrested or convicted of specific crimes Law enforcement agencies have drawn the public's attention to non-judicial database searches, and many want to know how often those searches occur [20]. Every state and federal government in the United States has passed legislation which stipulates which people who have been imprisoned or convicted must cooperate in providing DNA samples for analysis and inclusion in the NDIS database. The NDIS contains more than 12 million pieces of information, often used to compare DNA from crime scenes to identify potential offenders. However, it is only useful if the suspect—or a family member of the suspect has been incarcerated or charged with a crime and their DNA has been taken and stored [21].

In May 2021, Maryland implemented the inaugural legislation in the United States, and globally, that thoroughly governs the utilization of consumer genetic data by law enforcement agencies for crime investigations [22]. Thus far, the main limitation on law enforcement has originated from consumer genetics platforms themselves, with certain platforms refusing to collaborate, some secretly collaborating, and a few publicly collaborating with criminal investigators. Courts have predominantly adopted a non-interventionist stance, with limited attempts at supervision arising from the US Department of Justice (DOJ), which introduced a provisional policy in November 2019 [23, 24].

Major Challenges

Re-Identification Risks

Police now have access to readily accessible genetic data, so it is important to address the safety and confidentiality of DNA data. A need to strike a harmony between the perils related with auxiliary purposes of freely accessible information and the advantageous impacts of sharing information with the end goal of revelation has for some time been at the core of the secrecy of hereditary information banter. Different hereditary assets had recently been made for established researchers before direct-to--customer individual genome firms opened the entryway for the development of freely accessible DNA datasets. To work with the trading of hereditary information among researchers, the Data set of Genotypes and Aggregates was created by the Public Establishments of Wellbeing [20].

Probability of Recognizing Someone from a Public Data Set

Investigators have looked at and assessed the risks associated with utilizing genetic data to identify a person in both the commercial and therapeutic domains. These investigations have included reidentification risk examination from huge information bases as well as models for reidentification inside total informational indexes. Gierek et al. showed that it is feasible to deduce last names from genomic information by joining factual and genealogical strategies, pushing the limits of security guidelines.

Specialists in biodata have recently settled a connection between hereditary markers and last names got from family and parentage research. The association of Y-STR haplotypes with surnames provided an early indication of the sensitivity of DNA data specific to identity [20].

In the study of DNA profiles, uncertainty derives from various sources including sampling error, individual bands, and correlations within and between loci. One crucial component among them is kinship, where both criminals and innocent suspects may have one or more genetic connections through their lineage from a shared ancestor. Disregarding this lack of knowledge is continually unjust towards innocent individuals who are under suspicion. The impact is typically minor, but might be significant in certain instances. The US National Research Committee's report suggested an intricate, improvised, and excessively cautious approach to addressing some of these issues [22, 26].

Case Studies

Jessica Baggen

In 1996, investigators discovered Jessica Baggen, a teenager from Sitka, strangled and raped in a depression behind a downed tree. The neighborhood mourned the death of a bright teenager known for her independence and free spirit. In the last ninety-six months, charges in two additional cold case homicides in Alaska including Sophie Sergie and Shelley Connolly have resulted from the use of an emerging technique known as genetic genealogy, which the Alaska Department of Public Safety stated it had used to solve Baggen's case. The 66-year-old Steve Branch, who had previously resided in Sitka but had relocated to a small Arkansas town ten years prior, was the subject of the evidence. He committed suicide last week after cold case investigators drove there to confront him and get a sample of his DNA. An autopsy connected his DNA to Baggen's body. The innovative and potent method of genetic genealogy, which compares genetic material from crime scenes with internet DNA genealogical databases, has solved three prominent unresolved murder cases in Alaska, including the Baggen case. Since none of the cases have gone to trial yet, it is unclear how well the methods used to acquire genealogy evidence will hold up in court. Branch's death means we will never resolve the Baggen case. Nonetheless, Public Safety Commissioner Amanda Price stated, "We can finally say that Jessica's case is solved [13].

Lesa Lopez

The Alameda County Sheriff's Office has taken Lesa Lopez, 52, into custody in connection with the 32-year-old cold case murder of Baby John Doe in Castro Valley. The discovery of the baby alive on Madison Avenue in 1988 led to its subsequent slaying. The creation of a DNA profile in 2005, following decades of uncertainty over the baby's identity, is believed to establish a link between the murderer and the child's mother. Multiple ACSO investigators have examined the case over the previous 32 years. Lesa Lopez was taken into custody on July 23, and the Alameda County District Attorney's Office filed a murder complaint against her [13].

Murders of Lloyd Duane Bogle and Patricia Kalitzke

Three Montanan men in a lover's path near the Sun River discovered Lloyd Duane Bogle and Patricia Kalitzke's remains in 1956. Despite the case remaining unsolved for over 60 years, detectives continued their efforts and assigned Detective Sgt. Jon Kadner to it in 2012. In 2019, investigators approached the Kalitzke/Bogle case and requested Bode Technology's assistance in solving it. Forensic genealogists were able to use conserved samples and collaborate with partner labs to construct a DNA profile of the offender. They then combed public databases for possible matches. Most of the time, those profiles can lead to connections with the offender's distant relatives, like a second or third cousin. Forensic genealogists build a family tree based on public record searches that can lead them directly to the culprit even if they have not submitted their DNA to any public

databases.

Thanks to DNA testing, investigators discovered Kenneth Gould, a man who had resided in the Great Falls region with his spouse and children at the time of the killings. Testing the DNA of his surviving relatives was the only method to establish his guilt or innocence. After visiting Missouri and speaking with Gould's children, authorities finally named their father as a suspect. After more than 60 years, Kadner had to get in touch with the victims' surviving relatives and provide the closure that came with the test results indicating Gould was the murderer. In an effort to provide more families with the closure they deserve and, in many cases, have been waiting for years, the sheriff's office is now thinking of creating a cold case task team. Authorities believe they will be able to use this constantly developing science to solve cold cases that date back even further, as demonstrated by the resolution of the Kalitzke/Bogle case, one of the first criminal cases to be solved using forensic genealogy.

Melissa Lee

Snohomish County Jail has taken Alan Edward Dean into custody. Dean is a suspect in the 1993 kidnapping and murder of 15-year-old Melissa Lee. Dean faced charges of first-degree murder and one count of first-degree kidnapping after his arrest without any incident. Investigators are interviewing witnesses and compiling evidence for their case. Sharon Lee, the victim's mother, expressed her joy at being present for this, as well as her hope that Dean will burn in hell. We used genetic genealogy, which combines traditional genealogical methods with DNA testing to establish a person's relationship to their ancestors, to identify Dean as a potential candidate. DNA technology business Parabon Nano Labs (Parabon) in Virginia handled the genetic genealogy analysis for the case. After Parabon's genealogists identified Dean, investigators obtained a discarded DNA sample from a cigarette butt he had used. It positively matched the DNA profile found in the evidence recovered from the crime site, according to the Washington State Patrol crime lab. This marks the third time that Parabon has assisted in the arrest of a murder suspect in Snohomish County. The autopsy determined that Lee died from manual strangulation, leading to the ruling of his death as a homicide. According to a toxicology analysis, Lee's system included substances such as heptane and ethyl ether, but no narcotics or alcohol.

Legal Implications of IGG and FGG in Investigation

Investigative Genetic Genealogy (IGG), commonly referred to as Forensic Genetic Genealogy (FGG), is a relatively new technique in the investigative toolkit. It has shown promise in identifying the source of biological samples, whether they were taken from an unidentified decedent or were deposited at a crime scene. The public was abruptly made aware of FGG on April 24, 2018, when Joseph James DeAngelo, also referred to as the "Golden State Killer," was taken into custody for a string of rapes and killings in California. In June 2020, he entered a guilty plea [5, 9]. Forensic genetic genealogy (FGG), also known as investigative genetic genealogy (IGG), has become a potent forensic instrument that can produce vital leads for identifying unidentified victims and unidentified human remains [25-27].

Use of GEDmatch and DNA family tree in forensic investigation

GEDmatch is one of the most prominent third-party databases which provide data analysis from raw data submitted by its service user. Curtis Rogers and John Olson founded this privately held website in 2010 to assist in the information gathering process for both novice and expert genealogists. Although it does not offer DTC genetic testing, it does take raw DNA samples from other testing providers. As a result, it enables consumers to compare exams taken at different vendor laboratories across platforms. Using this database allowed for the capture of the Golden State Killer. About a million people were registered in the GEDmatch database as of November 2018, and 1800 new users were added every day [28, 29].

Jim Warren, Bennett Greenspan, and Max Blankfeld's concept served as the foundation for Family Tree DNA. It is a division of Houston, Texas-based Gene by Gene, a for-profit genetic testing business. For those interested in genealogy, it provides autoso-

mal, Y-, and mitochondrial DNA genetic testing services. Two million records are in the company database. It is the fourth most well-known company in the world for autosomal DNA and the most well-known company for YDNA and mitochondrial DNA [29].

Conclusion

Regardless of whether it has considerably aided in achieving the objective of previous crime cases and provided comfort to the groups of people in question, more research and administrative actions are anticipated to address the moral and legitimate implications of genealogy discovery. In the future, ethical and correct use of genealogy tools in scientific investigations will require a coordinated, multidisciplinary effort among quantifiable researchers, geneticists, ethicists, and administrators. Furthermore, because of the advancement of perceptive hereditary parentage as a powerful tool, law enforcement increasingly uses publicly accessible genealogical databases to investigate unsolved cases.

References

1. Wickenheiser RAJFSIS (2022) Expanding DNA database effectiveness. 4: 100226.
2. Redmonds G, King T, Hey D (2011) Surnames, DNA, and family history. OUP Oxford.
3. Stallard M, De Groot J (2020) Things are coming out that are questionable, we never knew about”: DNA and the new family history. *Journal of family history*, 45: 274-94.
4. Virdiramo V, Shaikh MT, Wolfram EJVoFS (2021) A Glimpse into Genetic Genealogy–Identifying Unknown Persons and Generating Suspect Leads in Violent Crimes, 1: 59-80.
5. Tillmar A, Fagerholm SA, 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.
6. Glynn, CL (2022) Bridging disciplines to form a new one: the emergence of forensic genetic genealogy. *Genes*, 13: 1381.
7. García Ó (2021) Forensic genealogy. Social, ethical, legal and scientific implications. *Spanish Journal of Legal Medicine*, 47: 112-9.
8. Zabel J (2019) The killer inside us: Law, ethics, and the forensic use of family genetics. *Berkeley J. Crim. L*, 24: 47.
9. Dowdeswell TL (2022) Forensic genetic genealogy: A profile of cases solved. *Forensic Science International: Genetics*, 58: 102679.
10. Katsanis SH (2020) Pedigrees and perpetrators: uses of DNA and genealogy in forensic investigations. *Annual review of genomics and human genetics*, 21: 535-64.
11. Greytak EM, Moore C, Armentrout SL (2019) Genetic genealogy for cold case and active investigations. *Forensic science international*, 299: 103-13.
12. Sutton MQ (2022) *Discovering World Prehistory: Interpreting the Past through Archaeology*: Routledge.
13. Smirnova D, Elkins KM (2024) Application of next-generation sequencing

14. Salinas OCJJofs (2024) Eligible offender samples are missing in CODIS: A statewide approach to performing a lawfully owed DNA census.
15. Rogalla-Ładniak UJAMSiK (2022) The overview of forensic genetic genealogy, 72: 211-222.
16. Rogalla-Ładniak, U. (2022). *archiwum medycyny sądowej i kryminologii*.
17. Ram N, Guerrini CJ, McGuire ALJS (2018) Genealogy databases and the future of criminal investigation. 360: 1078-9.
18. Pilli E, Castoldi E, Cattaneo C (2021) Forensic Anthropology Issues: A Synergy between Physical and Molecular Methods: The Contribution of Degraded DNA Analysis to Physical and Forensic Anthropology. In *Forensic DNA Analysis*, 289-336, Apple Academic Press.
19. Milton C (2017) Combined DNA Index System (CODIS) In *Cold case homicides* (pp. 365-370): CRC Press.
20. Kling D, Phillips C, Kennett D, Tillmar A (2021) Investigative genetic genealogy: Current methods, knowledge and practice. *Forensic Science International: Genetics*, 52: 102474.
21. Katsanis SHJArrog. genetics h (2020) Pedigrees and perpetrators: uses of DNA and genealogy in forensic investigations, 21: 535-64.
22. Stern J, Zhang S (2021) The victims left behind by genetic genealogy. *Atlantic*, 27.
23. Ram N, Murphy EE, Suter SM (2021) Regulating forensic genetic genealogy. *Science*, 37: 1444-6.
24. St John P (2020) The untold story of how the Golden State Killer was found: a covert operation and private DNA. *Los Angeles Times*, 8.
25. Katsanis SH (2020) Pedigrees and perpetrators: uses of DNA and genealogy in forensic investigations. *Annual review of genomics and human genetics*, 21: 535-64.
26. Cohen JE (1990) DNA fingerprinting for forensic identification: potential effects on data interpretation of subpopulation heterogeneity and band number variability. *American journal of human genetics*, 46: 358.
27. Erlich Y, Shor T, Pe'er I, Carmi S (2018) Identity inference of genomic data using long-range familial searches. *Science*, 362: 690-4.
28. Kennett D (2019) Using genetic genealogy databases in missing persons cases and to develop suspect leads in violent crimes. *Forensic science international*, 301: 107-17.
29. Mehar13 PA (2021) *Forensic Genetic Genealogy: An Investigative Aid*.
30. Balding DJ, Nichols RA (1994) DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. *Forensic science international*, 64: 125-40.
31. Murphy H (2018) *How an Unlikely Family History Website Transformed Cold Case Investigations*.

Submit your next manuscript to Annex Publishers and benefit from:

- ▶ Easy online submission process
- ▶ Rapid peer review process
- ▶ Online article availability soon after acceptance for Publication
- ▶ Open access: articles available free online
- ▶ More accessibility of the articles to the readers/researchers within the field
- ▶ Better discount on subsequent article submission

Submit your manuscript at

<http://www.annexpublishers.com/paper-submission.php>