# Scientific Working Group on DNA Analysis Methods

# Guidelines for Missing Persons Casework

## Table Of Contents



## SWGDAM Guidelines for Missing Persons Casework

The Scientific Working Group on DNA Analysis Methods, better known by its acronym of SWGDAM, is a group of approximately 50 scientists representing Federal, State, and Local forensic DNA laboratories in the United States and Canada. During meetings, which are held twice a year, Committees discuss topics of interest to the forensic DNA community and often develop documents to provide direction and guidance for the community. This document was presented to SWGDAM and received approval on January 9, 2014.

This document provides guidelines for the processing and analysis of missing persons casework. Laboratories are encouraged to review their standard operating procedures in light of these guidelines.

## Introduction

The intent of this document is to provide forensic laboratories with foundation information for conducting missing persons casework. These guidelines were developed primarily with regard to the context of missing persons in the United States that may involve the use of the National DNA Index System (NDIS). It is the responsibility of each organization to develop protocols, policies and procedures consistent with those of its criminal casework component as well as the local medico-legal authority.

The comparison of DNA profiles from unidentified human remains to samples attributed to the missing person or biological relatives of a missing person can provide a powerful tool for the medico-legal authority to aid in the identification process. The identification of individuals and remains can be accomplished by several methods. Recognition by the next of kin and the comparison of ante-mortem and post-mortem biometric data such as dental X-rays and fingerprints are common methods of identification. DNA analysis is available when these methods are not successful or additional information is needed.

DNA profiles developed from human remains may be compared directly to references from a family at the request of the investigating agency when the facts of the case lead to a preliminary hypothesis of identity. Alternatively, the DNA profiles from remains without a presumed identity can be compared to a DNA database of missing persons and family reference samples, such as the Combined DNA Index System (CODIS).

## 1. DNA Recovery from Human Remains

Unidentified human remains submitted to laboratories can vary from a blood stain to skeletal samples. DNA analysis success will depend on the type of remains recovered and the degree of decomposition. Blood is the best choice, since it contains a relatively high amount of DNA as compared to skeletal remains. Muscle tissue and nails are useful sources of DNA from decomposed remains. Skeletal remains are processed when other samples are not available or do not yield sufficient DNA. Skeletal material such as long bones and unrestored teeth are the most commonly submitted samples for identification purposes.

At the onset of sample processing, skeletal remains should be cleaned to remove dirt and surface contamination. Typically, this is accomplished by sonication in a detergent solution and/or sanding the surface of the bone. After cleaning, a portion of the bone is excised. Generally, this entails the use of a saw blade attached to a rotary tool. A section or cutting of the skeletal sample to be used in the extraction process should be reduced to a powder form by a freezer mill, blender or drill. The sanding and cutting of bones presents both a safety concern to the analyst as well as a contamination risk for the laboratory. Therefore, these procedures should be performed in a hood, preferably under negative air pressure. Proper personal protective equipment should be used to minimize exposure. Care should be exercised to avoid cutting completely through a bone or taking a section from an anatomically and/or forensically significant location. Such actions could negatively impact anthropological examinations.

In addition to cutting a sample for extraction, it is a good practice to cut and retain a portion of the remains for future DNA analysis. Retained samples are maintained in the event of advances in techniques that increase the possibility of obtaining a complete genetic profile (such as improvements in DNA isolation and amplification) and the adoption of additional DNA markers. The contributor should be notified that a portion of the sample is being retained.

The extraction of DNA from skeletal remains is similar to the DNA extraction from other types of forensic samples. The goal is to isolate the DNA from the sample and purify it in order to obtain a suitable DNA template for the typing of polymorphic autosomal (e.g., STR, SNP) and lineage (e.g., Y-STR, mtDNA) markers. Various methods of DNA extraction can be used in order to provide a suitable DNA extract from the bone sample including both organic and inorganic methods. If necessary, each laboratory should evaluate the suitability of its current methodologies and standard operating procedures for use with human remains. Laboratories interested in adopting procedures for processing skeletal samples for DNA analysis can obtain more information from the referenced papers and request a copy of the DNA extraction procedures for skeletal samples from laboratories currently conducting analysis of unidentified human remains.

Unidentified human remains are often recovered and analyzed months or years after death and are routinely exposed to adverse environmental conditions. These conditions can negatively impact the success of DNA typing through the degradation of DNA and the introduction of PCR inhibitors. The extent of degradation and inhibition depends on temperature, humidity, microbial growth and the length of time in the environment. For samples with highly degraded DNA, it may be necessary to type the sample with testing systems amenable to small template fragment sizes, such as mini-STR kits and mini-primer sets for amplifying mtDNA. Laboratories may attempt to overcome PCR inhibition using alternate purification techniques, dilution of DNA extracts or modified PCR reaction mixes (Alaeddini, 2012).

## 2. Reference Samples

Reference samples are either collected from the relatives of a missing person or are samples attributed directly to the missing person. These reference samples are subsequently compared to unidentified human remains to facilitate identifications. Family reference samples usually consist of buccal samples or dried blood stains. The collection should be conducted and/or witnessed by a law enforcement officer. The donor's identity and relationship to the missing person should be verified and documented. Family references should also be submitted with a consent form signed by the donor (or the donor's legal guardian). The consent form should indicate that the sample was donated voluntarily for the purpose of identification of a missing person. It should also indicate the extent to which the profile may be searched in a DNA database. Reference samples that are not submitted by law enforcement agencies with the appropriate documentation may not be suitable for entry into the National DNA Index System (NDIS).

Reference samples for missing and unidentified persons should be obtained from at least two biological relatives of the missing individual whenever possible (Figure 1). The most informative samples are from individuals closely related to the missing person such as parents, children and siblings, also known as first degree relatives. Parents and children share exactly 50% of their alleles (barring mutation) and have the additional advantage of permitting exclusions to be made, as each locus must have at least one shared allele. A

sibling on average also has 50% allele sharing but can share alleles to a lesser or greater extent, without the ability to make definitive exclusions. In the absence of parents or children, it is very useful to have multiple siblings as references.

The next most informative relatives are second degree relatives (e.g., aunts, uncles, half-siblings, and grandparents). Third-degree relatives include an individual's great-grandparents, great grandchildren, great uncles/aunts, and first cousins. If the only relatives available are distant relatives (i.e., second or third degree), STR markers may not be as useful; however, these relatives may be useful for lineage markers such as mtDNA (i.e., maternal lineage) or Y-STRs (i.e., paternal lineage for a missing male) (Figures 2A, 2B, and 3). For a missing parent, the spouse or parent-in-common submitted with a biological child is also useful.

The resulting profiles can be used to create a pedigree tree. If law enforcement is approached by multiple individuals wishing to donate samples, they should be encouraged to collect a reference from each person. The laboratory can then determine which samples will be best suited for analysis and pedigree tree construction. The importance of verifying the donor's biological relationship to the missing person cannot be overstated. Improperly stated or documented relationships will lead to missed identification opportunities.

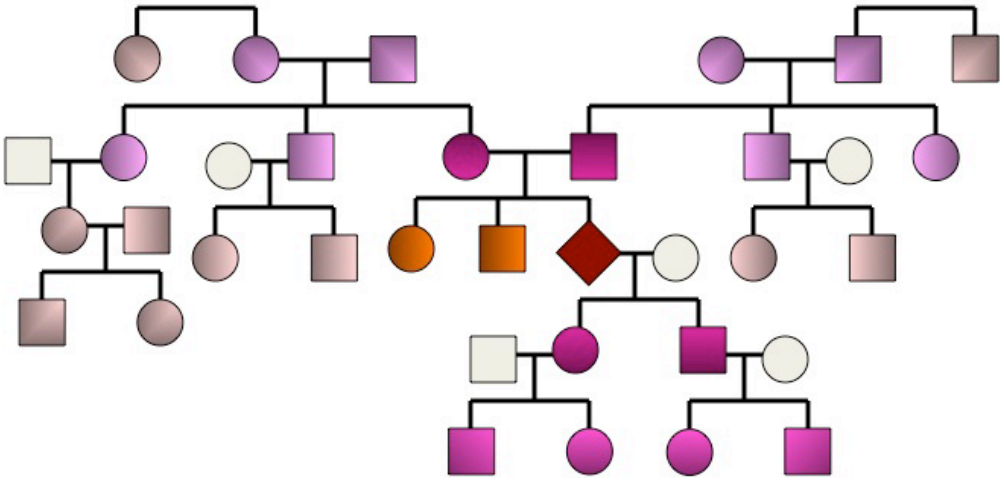

**Figure 1.** Pedigree displaying the dilution of allele sharing from the missing person (diamond). Diffusion of allele sharing is reflected in the shading. Unshaded individuals are genetically unrelated to the missing person (including spouse). (Key: Missing Person-Diamond, Male-Square, Female-Circle)

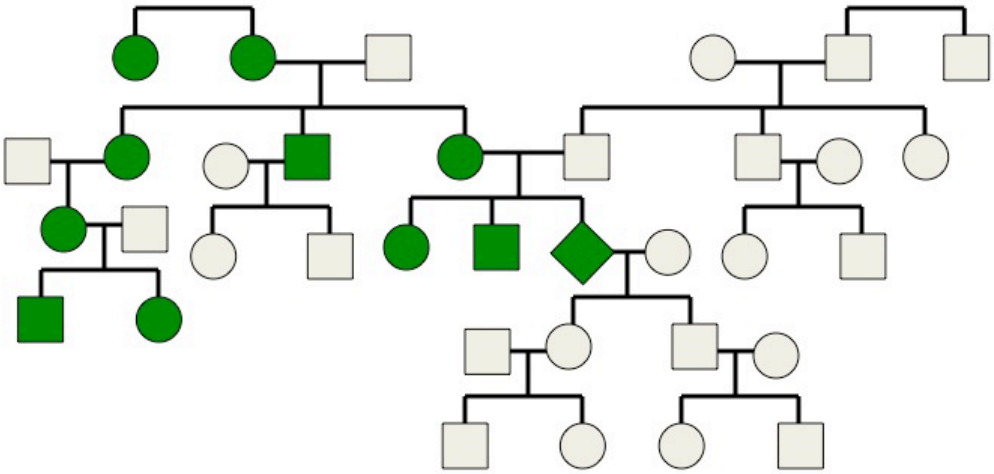

**Figure 2A.** Pedigree depicting pool of reference individuals sharing the same mitochondrial DNA haplotype for a missing male individual (barring mutation) in green. Unshaded individuals are maternally unrelated to the missing person. (Key: Missing Person-Diamond, Male-Square, Female-Circle)

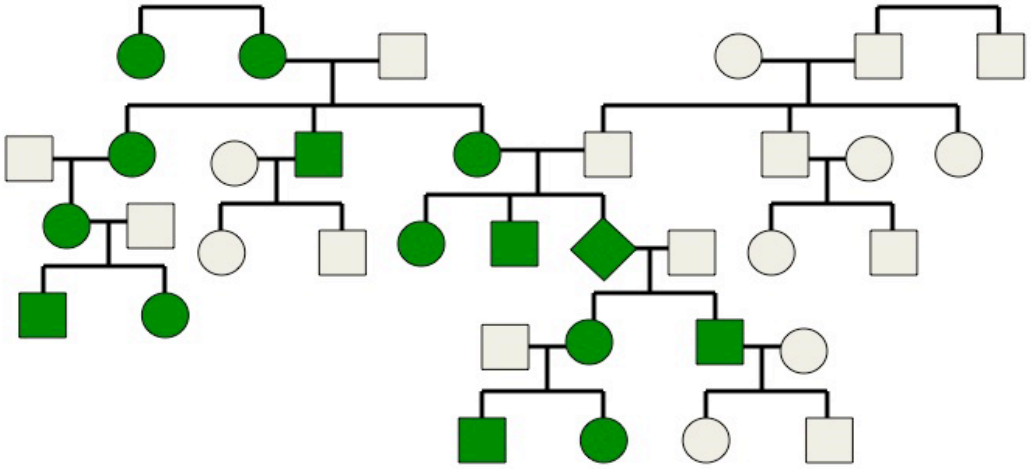

**Figure 2B.** Pedigree depicting pool of reference individuals sharing the same mitochondrial DNA haplotype for a missing female individual (barring mutation) in green. Unshaded individuals are maternally unrelated to the missing person.
(Key: Missing Person-Diamond, Male-Square, Female-Circle)

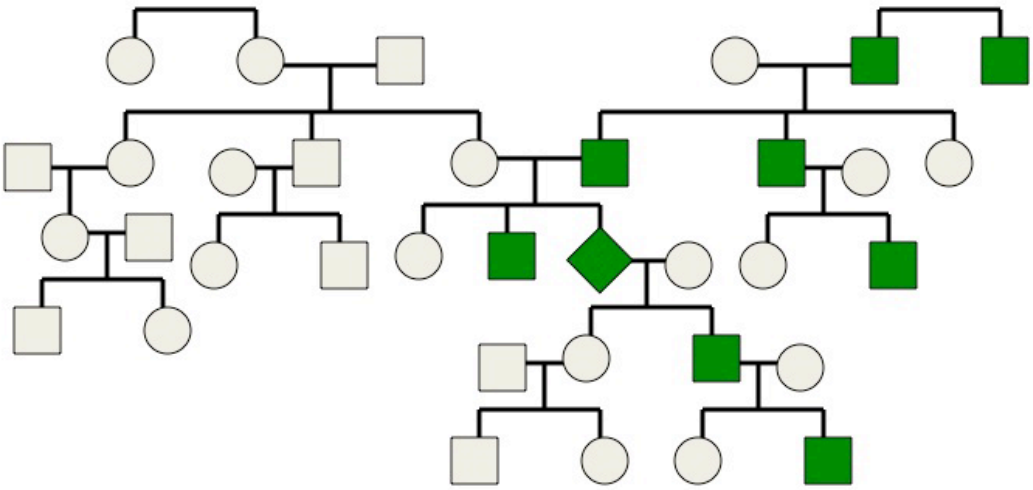

**Figure 3.** Pedigree depicting pool of reference individuals sharing the same Y-STR haplotype (barring mutation) in green. Unshaded individuals are paternally unrelated to the missing person.
(Key: Missing Person-Diamond, Male-Square, Female-Circle)

Direct reference samples are categorized either as a deduced missing person sample or as a missing person sample. A deduced missing person sample is an item that is believed to be attributed to the individual such as a toothbrush, razor, or article of clothing. The deduced missing person DNA typing results should be compared to known family reference samples, when available, to ensure that the results obtained are attributable to the missing person. In contrast, a missing person sample can be independently verified through other documentation as coming directly from the missing person (*e.g.*, a clinical biopsy sample or newborn screening specimen). For a thorough discussion of various sources of DNA from personal effects, see Prinz *et al*. (2007).

## 3. Metadata

Metadata is non-DNA information that is used in conjunction with DNA analysis to help resolve possible associations between missing and unidentified persons (see Table 1). Metadata for the remains is obtained from reports produced by other forensic disciplines such as anthropology, odontology and pathology as well as the law enforcement agency's report. For a missing person, metadata is gathered from police reports, investigators and family members. Information on the date of last contact, the last geographic location, the date of birth, and physical anomalies are just a few examples. After evaluating the genetic data of a potential association, metadata should be compared in an effort to validate or refute the likelihood of relatedness provided by the DNA results. It is important to note that metadata can be incorrect, particularly where date ranges, age estimates and ethnicities are involved. All laboratories and investigating agencies involved in associations should review the sources and limitations of the metadata before any final conclusions are reached.

Prior to conducting DNA analysis, the laboratory should consider involving experts from other disciplines in the identification process. The assistance of a forensic anthropologist can be useful in the recovery of evidence and skeletal remains. For many cases, the forensic anthropological examination can determine if a specimen is a bone and whether a bone is human or non-human. Additionally, an anthropologist can detect and document tool marks as well as ante-mortem and post-mortem skeletal damage. This examination may also determine if the recovered remains represent the commingling of more than one decedent.

Depending on the condition of the remains recovered, an anthropological examination can also provide an estimation of sex, age at death, potential ancestry, stature, and time since death. However, it is important to note that discrepancies can occur between an anthropological determination of sex and the genetic determination of sex (i.e., amelogenin genotype). Prior to processing teeth, a forensic odontologist should perform a dental examination to chart the types of teeth recovered, the position and the condition of the teeth, and any dental restorations. Post-mortem dental X-rays should also be taken. Forensic anthropology and odontology examinations should be performed prior to DNA analysis. If it is not feasible to have these examinations performed prior to DNA analysis, consultation with an anthropologist or odontologist should be considered to ensure that the samples taken for DNA analysis will not affect their examinations later

**Table 1**: Metadata categories used to evaluate potential associations

| **Unidentified Remains** | **Missing Person** |
|---|---|
| Sex | Sex |
| Ethnic Group | Ethnic Group |
| Age Range | Date of Birth (Age = DOB to present) |
| Height (cm) | Height (cm) |
| Dental Records Available | Dental Records Available |
| Specimen Origin (whole/partial remains) | |
| Metadata Comments | Metadata Comments |
| Physical Anomalies | Physical Anomalies |
| Scars/Marks/Tattoos | Scars/Marks/Tattoos |
| Date of Recovery | Date of Last Contact |
| Geographic Location of Recovery | Geographic Location of Last Contact |
| City | City |
| State | State |
| Location | Location |

## 4. Targeted Comparisons

Laboratories often receive requests to conduct specific comparisons when an unidentified person has been linked to a specific missing person through investigations conducted by medico-legal authorities. Typically, the laboratory would receive a sample from an unidentified person and samples collected from the relatives of a missing person with a specific request to compare the DNA results. This could result in strong evidence of identity or exclusion. Alternatively, depending on the amount of genetic data obtained from the remains and/or the relatedness of available references, support or discouragement for the hypothesis of identity could vary over a continuous range. It is important to note that DNA analysis is a tool utilized to assist the medical examiner/coroner in determining the identification of an individual. The laboratory does not make identifications. In situations where the DNA examinations do not conclusively support the proposed relationship, the laboratory can enter the eligible samples into the CODIS database. If an identification is made to a set of partial remains, the decedent's DNA profile should be placed into the CODIS database in the event that additional remains are recovered at a later time.

## 5. Database Comparison Using CODIS Software

CODIS describes the FBI's program of support for criminal justice DNA databases as well as the software used to run these databases. NDIS is considered one part of CODIS, the national level, containing the DNA profiles contributed by federal, state, and local participating forensic laboratories.

The DNA Identification Act of 1994 (42 U.S.C. §14132) authorized the establishment of this National DNA Index. The DNA Act specifies the categories of data that may be maintained in NDIS (convicted offenders, arrestees, legal, detainees, forensic (casework), unidentified human remains, missing persons and relatives of missing persons) as well as requirements for participating laboratories relating to quality assurance, privacy and expungement.

For missing persons, relatives of missing persons and unidentified human remains samples, additional DNA methods other than autosomal STR typing (such as mtDNA or Y-STR

typing) should always be considered, if relevant. For example, if the missing person is a female, then Y-STR typing would not be relevant. The lack of an additional typing system will not render a sample ineligible for entry into CODIS, but analysis using an additional appropriate system is recommended to ensure the most robust searching possible.

If a single allele is detected in an unidentified human remains specimen that is below the laboratory's stochastic threshold, that allele should not be entered into CODIS. Instead, this locus should be left blank, since it may result in a false exclusion.

A pedigree tree utilizing the relatives of the missing person should be created in the CODIS software. A pedigree tree is a graphical representation that describes the relationships among the missing person and his/her relatives. The more informative pedigree trees have at least one first degree relative of the missing person and both STR and lineage marker data.

### 5.1 Pairwise Profile Searches

Unidentified human remains profiles can be compared in a pairwise manner to profiles in various CODIS indexes in an attempt to discover a match to the source of the unidentified profile, or an association to profiles of closely related family members of the missing person. A pairwise search can provide very strong evidence for identification if the unidentified remains come from an individual whose profile had previously been entered into a searchable index. Using CODIS, close relatives can be found with a low stringency search to find profiles sharing at least one allele per STR locus and consistent results with at least one lineage marker. Under certain circumstances, such as when there may be inconsistencies in the stated relationships within the reference pedigree, pairwise comparisons can indicate a relationship association that is erroneously excluded by pedigree searches.

### 5.2 Pedigree Searches

A pedigree search compares an unidentified person to all known family members at once. This allows for a more robust search than the pairwise comparisons mentioned above. After the search is performed, a rank may be returned involving a pedigree and

an unidentified human remains specimen. For each pedigree tree, a ranked list of unidentified human remains will be generated using a likelihood ratio (see section 5.3 below). The same unidentified person may be returned as a rank to multiple pedigree trees among different laboratories. However, the true biological relationship may not be represented by the highest likelihood ratio, since that statistic is dependent on the number and types of relatives present in a particular pedigree tree. In other words, a weak pedigree tree, consisting of more distant relatives, may not return as high of a likelihood ratio as a stronger pedigree tree, with first degree relatives. The weak pedigree tree may contain the true biological relatives of the unidentified person while the stronger pedigree tree may have been returned as a fortuitous rank.

5.3 Statistical Relevance

CODIS uses several likelihood ratios to evaluate and rank candidate associations during a pedigree search. The statistical value can then be used as a threshold to limit the number of false associations returned during a search. The use of a threshold strikes a balance between limiting the number of false associations returned against the potential of not returning the true association if the statistic falls below the threshold.

The statistic used to evaluate kinship associations is based on a likelihood ratio (LR). Generally speaking, this is the ratio of the following two probabilities:

Ho: The probability of observing the genetic profiles from the unidentified remains sample and the family reference samples given that the remains represent the missing individual in the established pedigree of the family references;

Ha: The probability of observing the genetic profiles from the unidentified remains sample and the family reference samples given that the remains represent an individual unrelated to the established pedigree of the family references.

CODIS calculates likelihood ratios for each of the genetic typing systems present in a rank. For example, the Joint Pedigree Likelihood Ratio (JPLR) evaluates autosomal STRs. A Y-STR LR and an mtDNA LR are computed for each of those systems. All

three LR values, if present, are multiplied together to generate the Combined Likelihood Ratio (CLR). The threshold for returning a rank is based on the CLR.

It is important to stress that the LR values calculated by CODIS are only for the ranking of pedigree tree associations, not for statistical reporting. This is because CODIS combines several population groups to generate allele frequencies, whereas reported LR values are typically calculated separately by population group. PopStats or another statistical tool should be used for the calculation of LR values for reporting purposes. Prior to reporting combined statistics for autosomal STR, mtDNA and Y-STR results, the laboratory issuing the report should determine that each population used demonstrates independence among the genetic systems

## 6. Resolving Associations

Following a targeted comparison or as a result of a database search, additional genetic and non-genetic data may be useful in evaluating the validity of the proposed relationship between unidentified remains and reference samples. Depending on the particular family members present in a pedigree tree, the amount of genetic data and the metadata, it may not be possible to exclude a biological relationship between a pedigree tree and an unidentified remains sample, whether there is a true biological relationship or not. The laboratory should adopt report wording that accurately conveys the strength of the association to the submitting agency.

### 6.1 Statistical Relevance

When a rank is evaluated, the CODIS core loci and the available lineage markers that are part of the rank are evaluated first. Many times, additional STR data are obtained during the analysis process and are available for comparison purposes. Those additional loci should be reviewed in order to determine if they can be used to exclude the unidentified remains from the pedigree association. If appropriate population databases are available, those loci can also be used in subsequent LR calculations.

The necessity for additional DNA information should also be evaluated at this time. Based on the information provided with the association, it may be useful to obtain more genetic information. For example, if an unidentified remains sample was only typed using autosomal STRs and was potentially ranked to a pedigree containing two siblings, both with autosomal STRs and one sibling that also had mtDNA, then it may also be informative to develop a mtDNA profile for the unidentified remains to use in the comparison.

6.2 Review Metadata

The metadata should be evaluated secondarily to the DNA typing results and used in order to potentially include or exclude the rank. Metadata information for both the missing person and unidentified remains should be evaluated to assess consistency. For example, a date of last contact for the missing person after the date the body was found would exclude the association. Another example would be a different sex for the unidentified remains and the missing person. Vastly differing ages between the missing person and the unidentified remains may also be a reason to exclude an association. However, it is important to note that metadata can be incorrect, particularly where date ranges, age estimates and ethnicities are involved.

6.3 Facilitate Communication

If an exclusion cannot be made solely on the DNA results and the metadata provided, each laboratory is responsible for communicating with the other laboratory. Each laboratory is also responsible for contacting the respective submitting agencies in order to obtain additional information that could help in evaluating an association.

Additional family member reference samples may be necessary in order to build a stronger pedigree. Another sample from the unidentified remains may also be required in order to obtain additional DNA data, such as mini-STRs, Y-STRs or mtDNA.

## 7. Reporting Associations

Typically, the laboratory that entered the profile for the unidentified remains will take the lead in writing the report resulting from a CODIS association. It is important to understand that CODIS will allow for mutational events, and the reporting laboratory should be capable of making such calculations.

### 7.1 Statistical Evaluation

Missing persons casework encompasses a myriad of possible scenarios requiring appropriate statistical interpretation for an association or competing associations. Simple cases may require only a standard parentage or reverse parentage calculation, in which the missing person (represented by the unidentified remains profile) is a parent or child, respectively, of a trio. These types of calculations can be performed using the existing tools in PopStats. However, most cases present a more complex pedigree of known samples provided by family members of the missing person. Comparison of the genetic data from an unidentified remains sample to a pedigree of family reference samples is the appropriate and most powerful statistical approach for evaluating a database association. Statistical evaluation of pedigree data is most commonly done utilizing modifications and amendments to the algorithm described by Elston and Stewart (1971). Examples of such calculations are discussed in Appendix A.

It is imperative that the laboratory examines the pedigree for consistency prior to its use in comparison to the unidentified remains, especially if additional reference samples were added following an association. Different hypotheses may be considered during the statistical analysis to determine the best fit of the unidentified remains data to a reference pedigree or alternate pedigrees. Instances involving mutational events occur with STRs. The mutation rate of a single STR locus is low (approximately 10-3 per generation). However, the probability of observing a mutation increases when multiple family members, especially siblings, are included in the pedigree or a large number of loci are typed. These inconsistencies must be evaluated in terms of alternate

hypotheses surrounding the observation of the particular set of genetic data (Brenner 2004).

Another possibility is that the defined relationships in the pedigree are incorrect as reported (e.g., a presumed full sibling may actually be a half sibling). Alternate relationships in the pedigree should be investigated statistically, and if possible, through clarification by the submitting agency of the reference samples. Unaccounted for inconsistencies will result in a reduction of the statistical power inherent in the genetic data. Laboratories should have policies in place with regard to the reporting of relationships that are different than those stated by the family (e.g., non-paternity; Parker et al., 2012).

Statistical statements for reporting must clearly state the hypotheses examined that yielded the likelihood ratio. Laboratories may choose to adopt the convention of the relationship testing community (AABB, 2011) and report the LR based on the relevant population group (e.g., that of the pedigree members, hence the missing person); or report statistics in the manner used for criminal casework as determined by their laboratory protocol (see Appendix A).

## 7.2 Developing Reporting Policies

It is important for every laboratory to develop policies for the release of personally identifying information regarding an association, such as the names and the biological relationships of the tested individuals. It may be the laboratory policy to only report the findings to the contributor of the remains sample, where other laboratories may also send a copy of the report to the laboratory(ies) that submitted the family reference samples. If two laboratories are involved in an association, only one laboratory should generate a statistical report for the association.

## 8. Resources

Laboratories that have developed missing persons programs are available to assist laboratories in establishing similar programs. Technical procedures, reporting criteria, consent forms and submission guidelines can be obtained from each of the following laboratories:

California Department of Justice, Bureau of Forensic Services, Jan Bashinski DNA Laboratory (http://www.oag.ca.gov/bfs)

Federal Bureau of Investigation (www.fbi.gov/about-us/lab)

Minnesota Bureau of Criminal Apprehension (https://dps.mn.gov/divisions/bca/bca-divisions/forensic-science

University of North Texas Center for Human Identification (www.untfsu.com)

All states have missing persons clearinghouses that can be of assistance in obtaining information such as metadata or obtaining additional family reference samples to support an association (www.missingkids.com/clearinghouses). In addition, several organizations are involved in the efforts to match missing persons and unidentified remains using non-DNA information and are listed below (in alphabetical order). Many of these websites will indicate whether DNA is available and the laboratory contact information, but only NDIS conducts national DNA searches between missing and unidentified persons.

**Doe Network** (www.doenetwork.org)

The Doe Network is a web-based repository of information regarding missing and unidentified persons. This public website was started and is maintained by private citizens. A similar website is the Charley Project (www.charleyproject.org).

**NamUs** (www.namus.gov)

NamUs is an on-line repository of physical information for both missing and unidentified persons. It is publicly accessible and searchable, with restricted access to law enforcement and coroners/medical examiners for case information entry. The availability, location and type (STR or mtDNA) of DNA data can be noted for each case.

**NCIC**

The National Crime Information Center is a repository of information restricted to law enforcement agencies. It includes information on both people and property, involving several categories of crime such as wanted persons and stolen vehicles. There are also records of missing persons and unidentified persons. These records may include metadata such as the date of birth, sex, race, and dental records for a missing person. For unidentified persons, the records may include the estimated age, sex, race, estimated weight and height, date found, personal effects, and dental records. There is also information on whether DNA results are available and which laboratory processed the samples.

**NCMEC** (www.missingkids.com)

The National Center for Missing and Exploited Children is a non-profit entity chartered by Congress. NCMEC provides assistance to families and law enforcement in cases of missing children.

**ViCAP**

The FBI's Violent Criminal Apprehension Program was developed to track violent crime data. In addition to homicides and sexual assaults, ViCAP also contains information related to missing and unidentified persons, including photographs and other biographical data.

The information contained in these various databases, such as dental charting in NCIC, allows for searches to be conducted between missing persons and unidentified remains with potential matches to be returned. The investigating agencies that receive these potential matches may then request that specific DNA comparisons be performed. Similarly, requests for comparisons may be initiated by members of the public based on similarities between missing and unidentified persons on a public website.

Although requests for comparison may be initiated by one of the groups listed above or investigators in particular cases, every laboratory should have a policy for how they will handle such requests.

Some laboratories may decline the request for comparison if it is known that all sufficient data is included in the CODIS software and meets the specific searching parameters. As long as all data is housed at NDIS, then the CODIS software is already performing these comparisons. Laboratories should have written policies in place regarding information being released to entities other than the original contributor of the samples.

## 9. References

AABB, *Standards for Relationship Testing Laboratories*, (2011) 10th edition.

Alaeddini, R. *Forensic Implications of PCR Inhibition – A review*, Forensic Science International Genetics, 6 (3), (2012) 297 – 305; doi: 10.1016/j.fsigen.2011.08.006

Brenner, C.H. *Multiple mutations, covert mutations and false exclusions in paternity casework*, International Congress Series 1261 (2004), 112-114. Available at http://www.isfg.org/files/df2bf40f0c01d744f73028c3e723eb953de311ac.03018430_598204918881.pdf.

Elston, R.C., Stewart, J. *A general model for the genetic analysis of pedigree data*, Human Heredity, 21 (1971) 523–542; doi:10.1159/000152448.

Parker, L.S., London, A.J., Aronson, J.D. *Incidental findings in the use of DNA to identify human remains: An ethical assessment*, Forensic Science International Genetics, 7 (2013), 221-229; doi: 10.1016/j.fsigen.2012.10.002.

Prinz, M., Carracedo, A., Mayr, W.R., Morling, N., Parsons, T.J., Sajantila, A., Scheithauer, R., Schmitter, H., Schneider, P.M. *DNA Commission of the International Society for Forensic Genetics (ISFG): Recommendations regarding the role of forensic genetics for disaster victim identification (DVI)*, Forensic Science International Genetics, 1(1), (2007) 3-12; doi: 10.1016/j.fsigen.2006.10.003.

## Appendix A. Pedigree Statistical Evaluation Example

Unlike routine identity comparisons or single reference relationship analysis (*e.g.*, parent-offspring, sibling-sibling, etc.), evaluation of more extensive pedigrees to unidentified remains associations are more complex. Ultimately, the working hypotheses of these evaluations are: $H_0$: Given the observed genetic data (the unidentified remains profile and the pedigree data), the unidentified remains represent the missing person in the pedigree, or $H_a$: Given the observed genetic data (the unidentified remains profile and the pedigree data), the unidentified remains originate from an unknown individual unrelated to the pedigree. It is important to note that in certain circumstances, the unidentified remains may actually represent a different missing member of the same family. These additional alternate hypotheses should be evaluated if the case scenario warrants.

Each of the hypotheses can be represented by a conditional probability that takes into account the following three items: (1) the particular relationships among the members of the pedigree and the unidentified individual; (2) the probability associated with allele sharing given the degree of relationship among the members of the pedigree and the unidentified individual; and (3) the frequencies of the obligate alleles/genotypes that contribute to the structure of the pedigree.

Although the actual process of calculation will not be discussed here, the end result will be individual likelihoods from each system tested comparing the pair of hypotheses referred to as Kinship Indices (KI). The individual KIs are multiplied across all tested loci/systems (STR, Y-STR haplotype, mtDNA haplotype) to provide a Combined KI for the overall evaluation. During reporting and testimony, it is crucial that the value obtained be correctly presented in the context of the specific association. Sample wording for the various comparisons is provided in Appendix B.

As a working example, begin with the following pedigree:

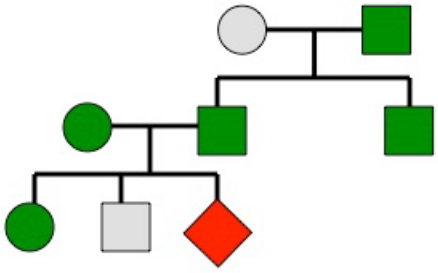

**Figure 1. Example Pedigree**

The red diamond represents the missing person from the pedigree and the green circles and squares (females and males, respectively) represent possible family reference samples obtained in these cases. Rarely will a pedigree contain all of the members shown in this example. In the case scenario below, different combinations of these family members will be used to demonstrate the statistical value various relationships can provide.

The case scenario is as follows:

- A young man went missing some number of years ago.
- The family filed a missing persons report.
- Cold-case detectives began gathering family reference samples in more recent years.
- An association was made in a search of NDIS between an unidentified remains sample submitted from an adjoining state and the pedigree assembled from the family reference samples.

The family members indicated they were of African American origin in their sample submission documents and anthropological examination of the remains is consistent with the remains originating from a male of African descent. The unidentified remains sample yielded a partial STR profile (10 loci including Amelogenin), a partial Y-STR haplotype (11 loci), and a complete mtDNA haplotype (16000-16386; 52-399). The autosomal STR allelic composition for the remains and reference samples are provided in Table 1.

**Table 1. Autosomal STR allelic composition of unidentified remains and potential family reference samples.**

| STR Locus | Unidentified Remains | Sister | Mother | Father | Paternal Grandfather | Paternal Uncle |
|---|---|---|---|---|---|---|
| D8S1179 | 11,12 | 11,13 | 11,13 | 11,12 | 11,12 | 11,14 |
| D21S11 | 29,30 | 32.2,36 | 30,32.2 | 29,36 | 29,30 | 30,36 |
| D3S1358 | 16,16 | 16,16 | 16,18 | 16,17 | 15,16 | 15, 18 |
| TH01 | 7,9.3 | 7,7 | 7,9.3 | 7,11 | 9.3,11 | 8,9.3 |
| D13S317 | 12,14 | 9,14 | 12,14 | 9,12 | 12,13 | 9,13 |
| D19S433 | 12,12 | 14,15 | 12,15 | 12,14 | 14,14 | 12,14 |
| vWA | 16,19 | 14,16 | 16,16 | 14,19 | 14,15 | 15,19 |
| TPOX | 8,11 | 8,11 | 8,8 | 11,11 | 11,11 | 9,11 |
| D5S818 | 12,13 | 12,13 | 12,13 | 11,12 | 12,12 | 12,13 |
| Amelogenin | XY | XX | XX | XY | XY | XY |

Working pedigrees can be made from various combinations of the family references available. Additionally, lineage marker testing can provide additional strength to a putative association, or exclude the unidentified remains from being a member of the pedigree. It is critical that the pedigree composition is verified to ensure that accurate relationships are depicted between the pedigree members. Potential contributors for lineage data and results from the assessment of Y-STRs and mtDNA are found in Table 2.

**Table 2. Database statistics for lineage testing performed on the unidentified remains sample.** Check marks indicate individuals that can contribute lineage data with these marker systems.

| | Unidentified Remains | Sister | Mother | Father | Paternal Grandfather | Paternal Uncle |
|---|---|---|---|---|---|---|
| **Y-STR** | ✓* | | | ✓ | ✓ | ✓ |
| * Partial Profile 11 loci | | | | | | |
| 0 observations in 4953 profiles | | | (African American) | | | |
| Upper 95% CI using US YSTR Database: | | | Freq = 0.000604 | | LR= | **1656** |
| **mtDNA** | ✓ | ✓ | ✓ | | | |
| 0 observations in 643 profiles | | | (African American) | | | |
| Upper 95% CI using Popstats mtDNA Database: | | | Freq = 0.0046 | | LR= | **217** |

Utilizing the data presented, several testing scenarios are possible in this sample case. One possibility is that only a single reference sample from this set is available for comparison to the unidentified remains. Below are examples of statistical results that would be obtained from pairwise comparisons of the unidentified remains to the putative relatives (Table 3). Each of these calculations can be performed in the PopStats module of CODIS.

**Table 3. Statistical results from kinship evaluations based on single family reference sample pairwise comparison to the unidentified remains profile.** NA indicates relative that cannot contribute the specified lineage marker results.

| Locus | Mother | Father | Sister | Paternal Uncle | Paternal Grandfather |
|---|---|---|---|---|---|
| | KI | KI | KI | KI | KI |
| D8S1179 | 6.923076923 | 9.23076923 | 3.71153846 | 3.961538462 | 5.115384615 |
| D21S11 | 1.3984375 | 1.31617647 | 0.25 | 1.19921875 | 1.857306985 |
| D3S1358 | 1.627906977 | 1.62790698 | 4.5279881 | 0.5 | 1.313953488 |
| TH01 | 2.953931204 | 0.56756757 | 0.81756757 | 1.693181818 | 1.693181818 |
| D13S317 | 7.401956425 | 0.51734104 | 3.69230769 | 0.5 | 0.75867052 |
| D19S433 | 4.394736842 | 4.39473684 | 0.25 | 2.697368421 | 0.5 |
| VWA | 1.855670103 | 3.46153846 | 0.71391753 | 2.230769231 | 0.5 |
| TPOX | 1.357142857 | 1.79027356 | 2.65387538 | 1.055851064 | 1.611702128 |
| D5S818 | 1.725852273 | 0.703125 | 2.55113636 | 1.362926136 | 1.203125 |
| Combined | 6582.36 | 111.21 | 15.33 | 17.00 | 7.77 |
| with mtDNA | 1428372 | NA | 3327 | NA | NA |
| with Y-STR | NA | 184158 | NA | 28152 | 12873 |
| with mtDNA & Y-STR | NA | NA | NA | NA | NA |

Having two family reference samples available for comparison in a pedigree provides marked improvement in the weight of the statistical evaluation performed (Table 4). Although the Reverse Parentage calculation (using the mother and father) can be performed in the PopStats module of CODIS, other combinations require the use of pedigree-based software tools.[1]

The inclusion of additional family reference samples in a pedigree will generally improve the statistical strength of an evaluation. The selection of additional pedigree members can be strategically made to permit the inclusion of lineage markers, such as Y-STRs, as shown in Table 4. The inclusion of the more distant paternal relatives alone does not substantially improve the power of the KI. There are additional scenarios possible and one often overlooked is the testing

[1] All pedigree calculations performed for this example utilized the KIn CALc 4.0 spreadsheet kindly provided by Steven Myers, California Department of Justice, Richmond, CA.

of multiple siblings[2] of a missing person *in lieu* of having one or both parents available (Table 5). Similarly, the inclusion of a spouse and children of the missing person with a parent and/or sibling can greatly improve the statistical evaluation of the genetic data. Strategically selecting available family references to reconstruct any missing parent(s) and provide lineage markers (mtDNA and Y-STR) should be actively discussed with investigators when there is a need to request additional family reference samples.

**Table 4. Statistical results from kinship evaluations based on pedigrees containing two family reference samples compared to the unidentified remains profile.** NA indicates relative that cannot contribute specified lineage marker results.

| Locus | Mother & Sister | Father & Sister | Mother & Father | Paternal Grandfather & Sister | Paternal Uncle & Sister | Paternal Uncle & Mother |
|---|---|---|---|---|---|---|
| | KI | KI | KI | KI | KI | KI |
| D8S1179 | 3.461538462 | 4.615384615 | 31.95266272 | 4.761015683 | 3.586538462 | 3.461538462 |
| D21S11 | 0.69921875 | 0.658088235 | 3.681181066 | 0.678653493 | 0.46536045 | 0.69921875 |
| D3S1358 | 3.464034613 | 3.464034613 | 2.650081125 | 4.782207077 | 3.996011357 | 0.813953488 |
| TH01 | 4.185810811 | 0.283783784 | 2.708845209 | 2.234797297 | 1.526182432 | 2.831388206 |
| D13S317 | 3.700978213 | 7.382058693 | 7.123388173 | 3.631336149 | 1.932643989 | 3.700978213 |
| D19S433 | 2.197368421 | 2.197368421 | 19.31371191 | 0.179184549 | 0.981165288 | 11.85422438 |
| VWA | 0.927835052 | 8.154242665 | 12.84694687 | 0.612452523 | 1.708564631 | 7.351308485 |
| TPOX | 3.696048632 | 4.129179331 | 3.017477204 | 2.843939029 | 2.488708054 | 2.18731003 |
| D5S818 | 2.301136364 | 1.789772727 | 1.438210227 | 2.712883707 | 2.558041001 | 2.301136364 |
| Combined | 2252.24 | 2918.69 | 6.48E+06 | 106.17 | 209.94 | 9055.00 |
| with mtDNA | 488736 | 633355 | 1.41E+09 | 23039 | 45558 | 1.96E+06 |
| with Y-STR | NA | 4.83E+06 | 1.07E+10 | 175819 | 347669 | 1.50E+07 |
| with mtDNA & Y-STR | NA | 1.05E+09 | 2.33E+12 | 3.82E+07 | 7.54E+07 | 3.25E+09 |

[2] Additional sibling profiles not shown.

**Table 5. Statistical results from kinship evaluations based on pedigrees containing three or four family reference samples compared to the unidentified remains profile.**

| Locus | Mother, Sister & Paternal Grandfather<br>KI | Mother, Sister & Paternal Uncle<br>KI | Mother, Sister & Brother<br>KI | Mother, 2 Sisters & Brother<br>KI |
|---|---|---|---|---|
| D8S1179 | 7.724819559 | 1.73076923 | 2.75089149 | 31.95266272 |
| D21S11 | 1.840590533 | 0.34960938 | 0.68958764 | 0.671100034 |
| D3S1358 | 3.658519075 | 3.05705787 | 3.84539744 | 2.650081125 |
| TH01 | 4.063267813 | 4.12453931 | 2.70884521 | 2.708845209 |
| D13S317 | 3.561694086 | 1.85048911 | 7.12338817 | 7.123388173 |
| D19S433 | 1.834802632 | 9.38783095 | 1.83480263 | 14.39060888 |
| VWA | 0.818677987 | 3.67565424 | 0.86984536 | 12.84694687 |
| TPOX | 4.125457114 | 3.59079168 | 4.12545711 | 4.63829353 |
| D5S818 | 2.516867898 | 2.58877841 | 2.5168679 | 1.438210227 |
| Combined | 11741.35 | 4528.69 | 2332.59 | 1.4E+06 |
| with mtDNA | 2.55E+06 | 982726 | 506173 | 2.93E+08 |
| with Y-STR | 1.94E+07 | 7.50E+06 | 3.86E+06 | 2.24E+09 |
| with mtDNA & Y-STR | 4.22E+09 | 1.627E+09 | 8.38E+08 | 4.86E+11 |

## Appendix B. Statistical Report Wording Examples

Reporting the results of a statistical evaluation from a missing person association encompasses the hypotheses tested and all of the genetic data utilized in the evaluation. Appropriate wording for statistical conclusions includes:

1. Mother and Father

The genetic data (autosomal STRs, mtDNA, & Y-STRs) are approximately 2.3 trillion times more likely to be observed under the scenario that the unidentified remains originated from a biological child of John Smith and Jane Smith as opposed to the unidentified remains originating from an unrelated individual from the African American population.

2. Mother, Sister and Paternal Uncle

The genetic data (autosomal STRs, mtDNA, & Y-STRs) are approximately 1.63 billion times more likely to be observed under the scenario that the unidentified remains originated from a biological child of Jane Smith, from a biological sibling of Judy Smith and a paternal nephew of Ralph Smith as opposed to the unidentified remains originating from an unrelated individual from the African American population.

3. Mother, Two Sisters and Brother

The genetic data (autosomal STRs, mtDNA, & Y-STRs) are approximately 483 billion times more likely to be observed under the scenario that the unidentified remains originated from a biological child of Jane Smith and from a biological sibling of Judy Smith, Mary Smith and Emmitt Smith as opposed to the unidentified remains originating from an unrelated individual from the African American population.

Along with the conclusions (*e.g.*, the unidentified remains sample cannot be excluded from being a child of…etc.) and the statistical weight of the genetic evidence developed in the case, the following statements may be included in the report:

> *Investigators are strongly encouraged to evaluate all associated case information in addition to the provided genetic results before declaring identity of the remains.*

*In the event that an identification is rendered by the appropriate legal authority, a copy of the death certificate or written communication of identification is requested by the laboratory so that all genetic data obtained from the associated family reference sample(s) can be expunged from the CODIS database.*