

ARTICLES

GENERAL GUIDELINES FOR CATEGORIZATION AND INTERPRETATION OF MIXED STR DNA PROFILES

RAY WICKENHEISER¹

ABSTRACT

The proposed general DNA interpretation guidelines provide a comprehensive approach for categorization and interpretation of mixed forensic STR DNA profiles. Mixed DNA profiles occur when there is more than one individual contributing biological material to a sample. The guidelines represent a compilation of methods currently in use by forensic laboratories, along with a discussion of the LSD (Least Squares Deconvolution) objective mathematical approach to interpreting two-person mixtures, as well as the Combined Probability of Inclusion/Probability of Exclusion method. DNA sample features and contexts found in commonly recurring case scenarios permits sorting of profiles into one of six categories. An example profile is interpreted under various scenarios to highlight differences in interpretation dependent on case circumstances. Sorting DNA profiles into categories permits a simplified, scientifically and mathematically sound, conservative interpretation without discarding significant forensic data.

RÉSUMÉ

Les consignes générales proposées ci-dessous décrivent une démarche compréhensive pour la classification et l'interprétation de profils génétiques basés sur les RCT (répétitions courtes en tandem), d'origine mixte. Les profils mixtes sont dérivés d'échantillons composés de matière biologique provenant de plusieurs individus. Les conseils ci-dessous rassemblent les méthodes actuellement utilisées dans les laboratoires médico-légaux, et discutent de l'approche mathématique objective de déconvolution des moindres carrés et de la méthode combinée des probabilités d'inclusion et d'exclusion. Les caractéristiques des échantillons d'ADN et les circonstances typiques des cas criminels permettent la classification des profils génétiques en six catégories. L'interprétation d'un profil mixte sous divers contextes démontre les différentes interprétations possibles selon les circonstances particulières d'un cas. La catégorisation proposée simplifie l'interprétation des profils génétiques en suivant une démarche scientifique et mathématique raisonnée et conservatrice, sans exclure des données génétiques significatives.

INTRODUCTION

Forensic DNA is a rapidly evolving discipline within the field of forensic science. From its start with Sir Alec Jeffreys in 1985 (1) to today, the discriminating power of forensic DNA to both include and exclude individuals suspected of a crime has been a tremendous asset to law enforcement, civil liberty, and public safety. With this improved discrimina-

1. Forensic Consultant, 105 Midcrest Drive, Irving, Texas, 75063. www.raywickenheiser.com

tion, a single source DNA profile, which matches that of a known person, may be attributed to that individual (2, 3). This source attribution may become clouded when the DNA from more than one individual results in the development of a mixed DNA profile.

Greater sensitivity of multiplexed Short Tandem Repeat (STR) Polymerase Chain Reaction (PCR) has resulted in the recovery of DNA profiles from a greater number of evidentiary items and biological sources found at crime scenes (2). This has also meant a greater potential to detect mixed DNA profiles. Besides biological material transfer between victims and suspects, sloughed skin cells transferred to garments by their wearer frequently produce endogenous DNA profiles (4). Handled objects such as guns and knives are frequently touched by more than one individual, especially in cases of gang violence, thereby producing mixed DNA profiles. Mixed DNA profiles are often more difficult to interpret than profiles originating from a single source, and therefore warrant special treatment. STR DNA profiles provide two sources of data that can be analyzed and compared, retention time and signal strength. Fragment lengths are separated into alleles. Alleles from crime scene samples are compared to samples from known individuals for inclusion or exclusion. Signal strength corresponding to amounts of DNA provides peak heights and areas, which can be used to determine relative quantitative contributions of various contributors.

Guidelines for interpreting mixed DNA profiles have been proposed to assist the scientist in determining the original contributors to mixed samples (5–11). All laboratories participating in the Combined DNA Index System (CODIS) must be accredited, and must therefore adhere to the FBI Quality Assurance Standards for Forensic DNA Laboratories. These guidelines dictate procedures for DNA interpretation. As a result, the Scientific Working Group on DNA Analysis Methods (SWGDM) has published a set of general mixture interpretation guidelines (12). Methods have focused on identification of mixed stains and their anomalies, as well as mathematical means of expressing the significance of the inclusion of a particular contributor (5, 6). Additional articles have elaborated on statistical interpretation of findings which involve mathematical (7, 8) or expert systems approaches (7, 9). The Least Squares Deconvolution (LSD) method in particular will be examined in greater detail in this article, as it provides a tool for mathematical determination and evaluation of best fit of mixed sample data to theoretical mixture models (8). Veteran forensic scientists have detailed their methods developed over time with case and mixed sample experience (10, 11). Very detailed studies have involved multiple laboratories conducting analyses on repetitive samples, and included detailed examination and comparison of results (13–16).

The forensic significance of DNA findings hinges on case circumstances. DNA findings may be innocently explained, or take on great meaning regarding the perpetrator's identity and circumstances surrounding the crime, depending on the case and sample details (17). With personal freedom often at stake, all DNA findings must be interpreted in a conservative fashion. An unbiased scientist is independent of the prosecution or defense, basing opinions on objective evidence fully supported by data. While it is critical to be conservative, all forensically valid assumptions should be made in order to make full use of all information and data available, including sample context and features within the case circumstances. Relevant information should not be sacrificed by excessively conservative approaches, as the scientist must maximize the value of the evidence.

Statistical interpretations of DNA findings must be straightforward for explanation to, and the understanding of, laypersons and non-scientists who may comprise the jury, members of the bar, and the judiciary. The approach must be sound and accepted scientifically in order to be defended in the court system. Likelihood Ratios are commonly used in

cases of forensic paternity (18). Mixture interpretation protocols using Likelihood Ratios are used successfully in many jurisdictions, supported by thorough study and rationale (19, 20). The Likelihood Ratio is a ratio of two competing hypotheses, the prosecution hypothesis (the DNA is that of the suspect) versus the defense hypothesis (the DNA is that of another individual). Depending on the case circumstances, the number of contributors, and the variety of hypotheses put to the scientist, the use of Likelihood Ratios can become quite complex. Excellent textbooks dealing with the entire field of STR DNA typing include treatment of mixture interpretation (21) and extensive statistical evaluation using Likelihood Ratios (22).

METHOD

Prior to detailed description of the DNA mixture interpretation guidelines, the following 10 steps provide a brief outline of the overall approach (key activities in bold for emphasis):

1. Know your **sample** and its context
2. Develop the DNA **profile**
3. **Survey** the profile (overview of the entire profile)
4. **Sift** – arrive at the callable peaks and profile
5. Do sample **context** or sample features add value?
6. Do **peak heights** add value?
7. **Sort** - if profile is a mixture, determine **mixture type**
8. Can a **sole source** be identified?
9. What confidence can be applied to **minor contributors**?
10. **Compare to known** profiles – if included, with how much confidence?

In the “Sort” step a more detailed method will be developed, which involves determining the type of mixture and detailed interpretation with sample conclusions. This step-by-step guideline for interpretation is performed according to the nature of the DNA profile, based on commonly occurring crime scenarios. Finally, an abbreviated flowchart is provided in the appendix section.

Within these guidelines, peak heights are used as opposed to peak areas. Algorithms used by the Perkin Elmer Applied Biosystems Genescan software to generate peaks create both a peak height and peak area value, however, peak height has been favored with respect to estimating the true amount of pre-PCR template DNA originally contained in a sample². Further, the Genescan algorithm uses peak height in order to generate peak area data³. Therefore, peak height will be utilized in this application. Note that use of peak area is in commonly accepted practice and is a valid approach (23).

The term mass ratio is referred to herein. Mass ratio is defined as the ratio of the peak height of the major contributor normalized to the peak height of the minor contributor. Hence, the amount of DNA (mass) originating from one contributor can be related to the

2. Sheppard M., Personal communication, Advanced Capillary Electrophoresis Course, ABI course in Baton Rouge, LA, March 17-20, 2003.
3. Scherzinger, C., Personal communication, Advanced Capillary Electrophoresis Course, ABI course in Foster City, CA, August 28, 2003.

contribution of the second contributor. For example, a mass ratio of 2:1 indicates that the major contributor has peak heights twice as high as those peaks attributed to the minor contributor. The same calculation could be conducted for peak areas rather than peak heights.

1. Know your **sample** and its context

The forensic scientist should obtain sufficient case details to determine which examinations are probative. Information regarding the origin of the sample can determine which individual is expected to be included in a profile, such as a victim, and which profiles are of crime-solving value, such as a foreign male.

2. Develop the DNA **profile**

As this method focuses on the interpretation of the DNA profile once developed, it will be assumed that a qualified scientist has now obtained a valid DNA profile, and is prepared to move to interpretation of findings at this point.

3. **Survey** the profile (overview of the entire profile)

In this step, the scientist takes an overview of the entire profile, rather than a locus-by-locus approach. By viewing the DNA profile in its entirety, the scientist can gain a sense as to the quality of the original DNA, the amount of degradation, the presence of multiple contributors, and the amount of confidence that can be attributed to subsequent interpretations. The overall quality of the data obtained is assessed in this step.

Discrete peak height ratios, and lower and upper rfu (Relative Fluorescence Units) thresholds and mass ratios have been suggested, but are to act as guidelines only. An individual laboratory's validation studies should be used to establish these levels. Background noise should not be greater than one-third the peak height in order for a peak to be considered valid (24), and hence included in a calculation to impart statistical significance to associative findings. The function of the lower threshold is to ensure that all peaks included in the calculation are definite peaks, which are attributable to DNA. The function of the upper threshold is to ensure that peak heights are indicative of DNA mass. Excess DNA quantity may saturate the detector, thereby compromising the relationship of peak height to DNA quantity.

With this overview of pertinent profile characteristics, such as peak heights, numbers of alleles present, and amelogenin X to Y mass ratio, a number of general conclusions can be drawn. Appropriate conclusions regarding sample attributes include the following:

- a) Indication of single or multiple sources based on peak number and peak height balance. Greater than two peaks at a locus indicates a possible mixture. Significant peak imbalance is frequently produced by two or more individuals contributing DNA to the forensic sample. Peak imbalance is defined as the difference in peak heights of alleles within a locus, typically greater than 30% relative to each other. This level is a suggested point for closer scrutiny.
- b) If it is a mixed profile, female only, male only, or possible male and female mixture based on Y to X amelogenin mass ratio.
- c) Level of degradation based on the profile survey.
A decrease in peak height moving from the low molecular weight loci to the high molecular weight loci indicates degradation. Excessive template DNA can also create this "ski slope" effect. Alleles at the higher molecular weight loci decreasing to the point

that alleles are no longer above the lower threshold for peak acceptance are another indicator of degradation. Likewise, larger alleles within each locus having less height, particularly where peaks appear to originate from a single heterozygous donor demonstrate a degrading sample. The steepness of the slope of the peak height decrease between loci will help determine the level of conservativeness required in interpretation. The more degraded the DNA, the higher the slope, and hence the higher the need for a conservative interpretation.

- d) Level of confidence to be used in comparing to known profiles. If heavily degraded, partial or incomplete profile, or a profile from many contributors, then less weight may be placed on an inclusion. A locus in which there are peaks just below the lower threshold should not be included in a statistical interpretation, but may be considered for comparison purposes.

4. **Sift** – arrive at the callable peaks and profile

a) Account for artifacts

In this step, the scientist determines the value of the peak height and allele data within the profile. The true alleles represented by the data are determined for each locus across the entire profile. Artifacts such as stutter and pull-up are accounted for (5), or eliminated through a re-injection or re-amplification of the sample.

b) Confirmation of a mixed profile

A mixed DNA profile possesses features that include:

- Greater than two peaks in a locus, not attributable to a mutation. This should be seen in several loci across the entire mixed profile, in conjunction with other attributes.
- Variation in autosomal peak height ratios.
- Variation in amelogenin peak height ratios.

Some single source profiles have greater than two peaks at a locus. If only one locus has three peaks, all other loci possess two peaks or less, and each set of two heterozygous peaks are relatively equal in peak height, a single source can be considered. Likewise, single source profiles may have a variation in autosomal and amelogenin peak height ratios due to a mutation in primer binding sites. These hypotheses can be supported or refuted through comparison to a known reference sample.

If there is a distinct major contributor, and evidence of a very minor contributor at a limited number of loci, then whether this is a mixture in the fullest sense, or a single source profile with a trace of a second individual can be debated. Regardless, a strong conclusion may be rendered with respect to the distinct major contributor, while extreme caution must be exercised with respect to a trace minor contributor.

5. Do sample **context** or sample features add value?

Additional interpretative value can be provided from information regarding sample location, or the nature of the sample itself, such as blood under the fingernails of a deceased, or a vaginal swab.

Under appropriate circumstances greater than one crime scene DNA profile produced from the same evidentiary item sample can be used to deduce the necessary individual donor profiles required to produce those composite profiles. Circumstances where this is appropriate include differentially extracted samples (e.g. profiles from differential extractions from personal samples, such as vaginal and oral swabs, penile swabs) and multiple samplings of the same stain (e.g. two cuttings from a mixed bloodstain, each revealing a

different mass ratio of the same two contributors). Multiple samplings of a questioned stain adds confidence to the interpretation with each consistent profile, and could also present differing mass ratios that assist in assigning alleles to particular contributors. As each profile obtains a new picture of the same sample, it is reasonable that data from all profiles be used in concert to provide an improved interpretation.

Under appropriate circumstances the DNA profile of a known reference sample is used in the deduction of a suspect DNA profile. Endogenous profiles, such as that of the wearer of a clothing item, or the individual from whom a swab of the skin or personal area has been taken, or from fingernails or swabs of bite marks, are commonly obtained from these types of samples. In cases where samples are taken directly from the person, or articles worn by them, the profile of the endogenous donor expected to contribute to the mixed DNA profile can be used in the deduction of a remaining suspect profile.

6. Do **peak heights** add value?

There are two sources of data obtained in the generation of a DNA profile. These sources of data are retention time and signal strength. When compared to allelic ladders, the retention time permits assignment of alleles to individual peaks. The second source of data is signal strength, which generates the peak heights of those alleles. In the appropriate circumstances, peak height has a positive linear relationship to DNA quantity, such that this additional information can assist in interpretation of a mixed DNA profile. Careful scrutiny of the profile will determine whether the positive relationship of peak height to DNA quantity is sufficiently linear to permit reliable use of this data. The adage “if in doubt, throw it out” could be appropriately applied to peak heights. If the data is not reliable, it should not be utilized. The quality of the relationship between peak height and DNA quantity can be demonstrated through observation of the consistency of the mass ratio between loci. If the mass ratio is consistent across the profile, greater confidence can be used to assign the taller corresponding peaks to one contributor, and the shorter peak heights to a second contributor. Conversely, with greater variation seen through degradation and other factors, such as a mixture of equal contributors, there is a point where the scientist determines that alleles cannot be unambiguously assigned to a particular donor using peak height. At this point, peak height data has been determined to be insufficiently reliable, and is not used in the interpretation.

Determining the number of potential contributors to a mixed profile assists in interpretation, as well as providing meaningful information as to the nature of the DNA sample, and hence circumstances surrounding the crime whence it came. A preliminary determination of the likely number of contributors to the mixed profile can be accomplished by examining the loci with the largest number of peaks. If no locus has more than four peaks, a two-person mixture is considered. If no locus has more than six peaks, a three-person mixture is considered, and so on. Next, the peak heights relative to each other within each locus are examined. Are the peak heights of the alleles within each locus consistent with the number of contributors chosen? The level of degradation is recalled relative to the peak height variation seen in each locus. Once again, the entire profile is considered, in addition to a locus-by-locus approach. The greater the level of degradation, the greater is the need to be conservative in terms of interpretation.

If the number of contributors can be established, the next step is assessing the relative contribution of each individual to the profile. Relative contributions of major and minor contributors can be best established by observing loci without shared peaks. An average of the higher height corresponding peaks can be divided by the average of the corresponding peaks with the lower peak heights to produce an estimate. In order to quickly

estimate the major:minor ratio for a two-person mixture; locate a locus bearing four alleles. Divide the sum of the two tallest allele peaks by the sum of the two shortest allele peaks. The result indicates a ratio of major:minor normalized to the minor contributor, which is the mass ratio.

For two-contributor mixtures, the Least Squares Deconvolution (LSD)(8) approach provides an objective mathematical mass ratio independently for each locus, including those with shared alleles. Comparison of mass ratios across loci of the profile is a good diagnostic to determine the degradation level and evenness of amplification of loci relative to each other.

An elevated level of degradation, reflected by an increased level of variation in the mass ratio across the loci of the profile, reduces the confidence in the information conveyed by peak heights, and may warrant consideration of the more conservative Probability of Inclusion/Exclusion approach, which uses only alleles and not peak height information to statistically evaluate the mixed DNA profile.

7. Sort – if the profile is a mixture, determine the **mixture type**.

If the profile is a mixture, by categorizing the mixture according to type, detailed interpretation is facilitated. The following steps 8, 9 and 10 that were mentioned previously can now be incorporated. The Sort step will be discussed in detail in the DNA Mixture Type Categorization section.

8. Can a sole source be identified?

Depending on the mixture type and category selected, a sole source can be unambiguously determined, and thereby treated as a single donor profile for the purposes of comparison to known samples and conducting database searches.

9. What confidence can be applied to minor contributors?

Depending on the mixture type and category selected, the confidence level to be applied to the interpretation of minor contributors is facilitated.

10. Compare to known profiles – if included, with how much confidence?

Crime scene samples are developed such that comparison to known samples can reveal the identity of the perpetrator of a crime. Frequently a known sample is not available. Unambiguous deduction of a contributing profile infers that the profile of a suspect has been deduced. This profile in turn is entered in the appropriate database for comparison to other crime scene profiles, as well as large number of known offender/suspect samples.

The ideal circumstance of fully interpreting mixed DNA profiles prior to consulting suspect known profiles may not be appropriate for minor contributors. In mixed profiles where some DNA peaks are small, shared, or DNA is degraded, this ideal circumstance may not be practical. In cases such as these, minor peaks potentially originating from a minor contributor could be discounted if the known profile is not consulted in the interpretation, thereby causing an incorrect elimination.

As an illustration of the scenario where correct final interpretation may not be possible without consultation of a known profile for comparison, consider the following profile at a single locus:

| Allele | Peak Height |
|--------|-------------|
| 9 | 200 |
| 10 | 1200 |
| 12 | 1000 |

If the laboratory has established a stutter peak threshold cutoff of 17% through their validation studies at this locus, the scientist could establish that the profile is 10, 12. Now, consider a suspect with 9,10 at this locus. This suspect could be a minor contributor to this potentially mixed profile with a mass ratio of 5:1, which would have been established in the overview of the profile previously. Would it be appropriate to then eliminate the suspect by adhering to an inflexible policy of interpreting all profiles without some consultation of the known samples? In this case, the scientist would take into account the low mass ratio of the minor contributor, and the location of the 9 allele, and now declare this locus inconclusive after consideration of the known profile. Therefore, interpretation in this case is best performed with a review of the known sample. While this scenario presents a “less than ideal circumstance”, consulting the known reference profile prior to final interpretation of the mixed DNA profile is correct in this context.

These interpretation guidelines may be used to attach appropriate statistical significance to findings. For simplicity, Random Match Probability and Probability of Inclusion/Exclusion Statistics will be used, as opposed to more detailed statistical approaches. Simply stated, the Random Match Probability is how rare a particular profile is estimated to be within a given population. The Random Match Probability as applied in these guidelines assumes that individuals are unrelated. Where this assumption may not hold true, kinship calculations may be performed to take various relationships into account in statistical significance calculations (28).

Where an individual contributor can be unambiguously deduced from a mixed profile, it is recommended that consideration be given towards use of an identity statement. This topic will be dealt with in the discussion section.

DNA Mixture Type Categorization

A DNA profile that originates from more than one contributor (mixed profile) will usually fit into three broad categories:

- a) Differential extraction (Mixture Type 1)
- b) Forensically valid assumption (Mixture Type 2)
- c) Single mixed non-differential profile with no assumptions (Mixture Types 3, 4, 5, and 6)

Mixture Type 1: Differential Extraction

Mixture Type 1 is a set of two mixed profiles containing a male and a female contributor, specifically including a sperm (male fraction) and an epithelial (female or non-sperm) fraction, such as would appear in a sexual assault case. The term non-sperm has been added, as not all cases have females as the non-sperm donor. As the scientist is considering not one, but two distinct profiles obtained from a single item, additional information can often be derived in this situation.

Commonly occurring evidence includes vaginal swabs, penile swabs, vaginal pools and washes, stains from inside suspects' undershorts, oral swabs, anal swabs, rectal swabs, and stains from panties and sheets. As a differential extraction procedure is used, there will usually be an enhancement or enrichment of one component of the profile relative to another, in the non-sperm fraction profile versus the male fraction profile. Through this enrichment, the non-sperm contributor can be separated from the male contributor, as seen by a shift in peak heights from the non-sperm contributor in relation to peak heights from the male contributor. If this separation is unambiguous, then the non-sperm contributor within the non-sperm fraction profile is predominant, while substantially reduced within

the male fraction profile. Likewise, the male contributor within the non-sperm fraction profile is absent or in the minority in the non-sperm fraction, while substantially increased within the male fraction profile. If this holds true, then a Random Match Probability can be stated for each contributor, provided that the following conditions are met:

1. The profile appears to include only two individuals. Profiles with greater than two contributors should consider interpretation under Mixture Type **5** or **6**.
2. All peaks within the profile are accounted for via comparison to known reference samples. That is, if the non-sperm profile is subtracted, then an unambiguous male profile is determined. Subsequent comparison to a known male profile accounts for all peaks in this profile.
3. There is one sperm and one non-sperm contributor, and an enrichment of each profile relative to the other profile in its corresponding fraction is seen.
4. The lower limit for use of a Random Match Probability for the male contributor is a **1:4** mass ratio of Male: Female across the profile, as seen in the male fraction. This level has been chosen as it is distinctly above common stutter and allele imbalance levels, such that peaks can be distinguished. The scientist must be confident that the deduced male profile has been unambiguously determined after accounting for the contribution of the non-sperm profile. Should this ratio not be met, interpretation under Mixture Type **6** (Probability of Inclusion/Exclusion) should be considered.
5. Appropriately shaped peaks fall between the established lower threshold and upper thresholds (up to scientist discretion), in order for loci to be included in the Random Match Probability.

Sample Conclusion

A mixed DNA profile was obtained from the vaginal swab, exhibit 1 (Jones, Mary), originating from one male and one female contributor. The non-sperm component of the profile matches that of the known blood sample, exhibit 2 (Jones, Mary). The male (semen) component of the profile matches that of the known blood sample, exhibit 3 (Smith, John). *The Random Match Probability (or other means of attaching a statistical significance) would be stated here.* An identity statement may be considered for the male contributor (See identity statement).

Mixture Type 2: Forensically Valid Assumption

Mixture Type **2** is a mixed profile obtained from a single sampling of a specimen, to which a forensically valid assumption can be applied. Unlike the differential extraction outlined in Mixture Type **1**, above, there is a single profile to consider (usually non-semen related), without the benefit of enrichment from one fraction to another. The DNA from the individual from which the swab was taken would likely be present in a mixed profile found on a “personal” or “intimate” sample. Therefore, it is acceptable to account for the contribution of the endogenous profile within this mixture, provided this assumption is stated. A Random Match Probability may be stated for the resulting DNA profile association. Examples would include:

- A swab of blood from the hand of a suspect or complainant/deceased.
- A swab from a bite mark.
- Swabs or clippings from beneath fingernails.

- Blood from clothing on the body of an individual known to be bleeding. This individual's DNA profile may be subtracted. Use of this assumption is based on case specifics.

Using a forensically valid assumption, a Random Match Probability can be stated for each contributor provided that the following conditions are met:

1. The sample was either taken directly from the body of an individual, or from a garment or item seized directly from the individual whose profile is to be subtracted. Use of the assumption in this instance is very case specific, so care must be taken that it is valid.
2. The profile appears to include only two individuals. Profiles with greater than two contributors should consider interpretation under Mixture Type **5** or **6**.
3. All peaks within the profile are accounted for via comparison to known reference samples. If a locus is ambiguous, that is, there are peaks that are overlapping, or cannot be unambiguously determined from the mixed profile, that locus may be omitted from use in the statistical interpretation. The informative loci that can be unambiguously deduced would be included in the statistical interpretation. As an example, this more conservative approach may result in 10 of 13 loci being included in a statistical interpretation, while three loci are not applied towards the statistic.
4. The lower limit for use of a Random Match Probability is a **1:4** mass ratio of Remaining Profile: Subtracted Profile. The remaining profile should be **1:4**, or greater, relative to the profile being accounted for. The scientist must be confident that the deduced profile has been unambiguously determined after accounting for the contribution of the endogenous profile. The assumption of common origin of the non-probative profile and the known donor should be stated in the conclusion (see sample conclusion below). Should this ratio not be met, interpretation under Mixture Type **6** (Probability of Inclusion/Exclusion) should be considered.
5. Appropriately shaped peaks fall between the established lower threshold and upper thresholds (up to scientist discretion), in order for loci to be included in the Random Match Probability.

Sample Conclusion

A mixed DNA profile was obtained from the swab of the bite mark, exhibit 1 (arm of Jones, Mary). If one can assume that the donor of the known blood sample, exhibit 2 (Jones, Mary), is a contributor to the mixed profile found on the swab of the bite mark, exhibit 1 (arm of Jones, Mary), then the remaining profile matches that of the known blood sample, Exhibit 3 (Smith, John). *The Random Match Probability (or other means of attaching a statistical significance) would be stated here.* An identity statement may be considered for the deduced contributor (See identity statement).

Single mixed non-differential profile with no assumptions (Mixture Types 3, 4, 5, and 6)

These categories include those mixed profiles that are not included in Mixture Type **1** or **2**, such as blood on an item of clothing found in a garbage bin. No valid assumptions can be stated regarding the originator, exclusive of forensic results, and a differential extraction would yield no additional probative results. Mixed profiles with greater than two contributors are considered for interpretation using Mixture Types **5** and **6**.

Mixture Type 3: Distinct major contributor and trace minor contributor

Some mixtures have a minor trace contributor at a small number of loci. The major contributor mass ratio is **10:1** or greater, with the minor seen in only a few loci, with an incomplete profile. This level has been chosen as the major profile is very distinct, while the minor profile is not above common stutter and allele imbalance levels, such that peaks can be distinguished. Very limited information can be drawn regarding the trace minor contributor, due to the limited loci and alleles, and the fact that much of the minor profile could be “buried” in the major profile, and stutter peaks. With the small nature of the minor contributor’s contribution to adding to these major peaks, even its additive effect in shared peaks is often lost in the normal peak to peak variation (10%)(17). Trace minor contributors should be interpreted with caution under Mixture Type **6**, or in the case of extremely very low and infrequent peaks, not interpreted statistically. Re-sampling should be considered.

Sample Conclusion

A DNA profile was obtained from the blood identified on the running shoe, exhibit 1 (scene), comprised of a distinct major contributor, and a trace minor contributor. The profile of the major contributor matches that of the known blood sample, exhibit 2 (Smith, John). *The Random Match Probability (or other means of attaching a statistical significance) would be stated here.* A statement regarding the minor component may be considered under Mixture Type **6** below.

As peaks which may have originated from the trace contributor may fall within the stutter peaks or overlap peaks of the major contributor, the known sample may be consulted prior to final interpretation. Loci with potential trace contributor peaks that do not meet minimum stutter threshold requirements may be considered inconclusive rather than exclusionary when interpreted along with the known reference sample. An identity statement may be considered for the major contributor (See identity statement).

Mixture Type 4: A mixed profile with one major contributor and one minor contributor

A Random Match Probability may be stated for the major contributor, provided that the following conditions are met:

1. The profile appears to include only two individuals. Profiles with greater than two contributors should consider interpretation under Mixture Type **5** or **6**.
2. All major peaks within the profile are accounted for via comparison to the known reference sample for the major contributor.
3. The major contributor is present with a mass ratio of greater than **2.3:1** (in a locus with four peaks, the ratio of the lowest peak of the major contributor compared to the highest peak of the minor contributor) for the majority of loci across the DNA profile. This level has been chosen as it is distinctly above a 2:1 ratio, where ambiguity may occur in specific two- and three-peak profiles. This situation is highlighted in the discussion section. The scientist must be confident that the major profile has been unambiguously determined after accounting for the contribution of the minor profile. If a locus is ambiguous, that is, there are peaks that are buried in, or cannot be dissected from, the mixed profile, that locus should be considered for omission from the statistical interpretation. Should this ratio not be met, interpretation under Mixture Type **6** (Probability of Inclusion/Exclusion) should be considered.

4. Appropriately shaped peaks fall between the established lower threshold and upper thresholds (up to scientist discretion), in order for loci to be included in the Random Match Probability.
5. A separate conclusion regarding the minor profile could also be drawn if the minor profile can be unambiguously determined. Application of Mixture Type **6** (Probability of Inclusion/Exclusion) guidelines below may be applied to include minor profiles in a second conclusion, in addition to a conclusion stated separately for the major contributor.

Sample Conclusion

A mixed DNA profile was obtained from the blood identified on the running shoe, exhibit 1 (scene). The profile of the major contributor matches that of the known blood sample, exhibit 2 (Smith, John). *The Random Match Probability (or other means of attaching a statistical significance) would be stated here.* A statement regarding the minor component may be considered under Mixture Type **6** (Probability of Inclusion/Exclusion) below. An identity statement may be considered for the major contributor (See identity statement).

Mixture Type 5: A mixed profile with one major contributor, and more than one minor contributor

A Random Match Probability may be stated for the major contributor, provided that the following conditions are met:

1. The profile appears to include more than two individuals.
2. All major peaks within the profile are accounted for via comparison to the known reference sample for the major contributor.
3. The major contributor is present in a ratio of greater than **3:1** (in a locus with four peaks or more, the ratio of the lowest peak of the major contributor compared to the highest peak of the minor contributors) for the majority of loci across the DNA profile. This level has been chosen as it is distinctly above common stutter and allele imbalance levels, such that peaks can be distinguished, plus an added margin for error given the multiple minor contributors. A common allele could conceivably be shared by greater than one minor contributor, so an extra margin is added. The scientist must be confident that the deduced profile has been unambiguously determined after accounting for the contribution of the minor profiles. Should this mass ratio not be met, interpretation under Mixture Type **6** (Probability of Inclusion/Exclusion) should be considered.
4. Appropriately shaped peaks fall between the established lower threshold and upper thresholds (up to scientist discretion), in order for loci to be included in the Random Match Probability.
5. A separate conclusion regarding the minor profiles could also be drawn if the minor profiles can be unambiguously determined. Application of Mixture Type **6** (Probability of Inclusion/Exclusion) guidelines below may be applied to include minor profiles in a second conclusion, in addition to a conclusion stated separately for the major contributor.

Sample Conclusion

A mixed DNA profile was obtained from the blood identified on the running shoe, exhibit 1 (scene). The profile of the major contributor matches that of the known blood

sample, exhibit 2 (Smith, John). *The Random Match Probability (or other means of attaching a statistical significance) would be stated here.* A statement regarding the minor components may be considered under Mixture Type **6** (Probability of Inclusion/Exclusion) below. An identity statement may be considered for the major contributor (See identity statement).

Mixture Type 6: A mixed profile with no distinct major and/or minor contributors

The Probability of Inclusion/Exclusion statistical treatment of mixtures applies to Mixture Type **6**, which is a mixed profile with no distinct major and/or minor contributors. A more detailed description of the Probability of Inclusion/Exclusion is found in Appendix I. In Mixture Type **6**, the profile is not a differential extraction, and no forensically valid assumptions can be made. No major contributor can be unambiguously determined. As a result, the data carried by peak height is not informative for separating mixture contributors, and is therefore not used. Only the data carried by the allele designations themselves is used in this conservative statistical treatment. Generally, mixtures to which the Probability of Inclusion/Exclusion applies occur under one of the following conditions:

- a) The mixed profile has greater than two contributors, which does not fit the Mixture Type **5** mass ratio **3:1** guideline above, as in there is no distinct major contributor.
- b) In a two-person mixture, the mass ratio of the contributors is between 1:1 and 2.3:1 (major:minor). The 2.3:1 ratio represents just below a 70:30 major:minor mass ratio. Insufficient differences exist between in the peak heights between alleles such that no major contributor can be unambiguously determined as described in previous mixture types.

An example of this case would be a mixture of blood from two individual donors with approximately the same relative quantities of DNA originating from each of the two contributors. No distinct major or minor contributor can be determined.

- c) The mixed profile has a major and a minor contributor. The profile of the major contributor has been unambiguously determined, and a conclusion drawn regarding identity. The Probability of Inclusion/Exclusion statistic is often used as a means of relating the statistical significance to a minor contributor(s) in Mixture Types **3**, **4** or **5** above. In this manner, a more conclusive interpretation may be stated regarding the major contributor, while an appropriately more conservative conclusion may be made with respect to minor contributor(s).

In cases of this type, Pop Stats (CODIS, FBI) or the SDRmix program of Dr. George Carmody (Carleton University, Ottawa, Ontario, Canada)(28) will be considered to facilitate calculation of the Probability of Inclusion/Exclusion statistics. In order that a Probability of Inclusion/Exclusion statistic may be stated, the following conditions must be met:

1. Appropriately shaped peaks fall between the established lower threshold and upper thresholds (up to scientist discretion), in order for loci to be included in the Combined Probability of Inclusion/Exclusion Statistic. Background noise should not be greater than one-third the peak height of any peak used in the calculation (24). The function of the lower limit is to ensure that all peaks included in the calculation are definite peaks.
2. All allele peaks found in the probative known reference profile are also found in the mixed profile, in the absence of degradation, very low peak heights, or another reasonable explanation. Otherwise, an allele found in the known reference profile that is

not found in the mixed profile is grounds for exclusion. If an allele shaped peak corresponding to the known profile is seen below the lower threshold, but above the 3:1 signal-to-noise baseline such that the peak cannot be discounted as noise, the individual is still included, but that locus is not used in the statistical interpretation. If no peak is seen above the 3:1 signal-to-noise baseline, but that allele is present in the known reference sample, in the absence of degradation, very low peak heights, or another reasonable explanation, then the donor of the known reference sample is eliminated as a possible contributor to the mixed profile.

3. No peaks should be seen above the 3:1 signal to noise baseline, to ensure smaller minor components are not discounted. If peaks occur between the 3:1 signal to noise baseline and the lower threshold, one cannot unambiguously determine the number of alleles to include in the Probability of Inclusion/Exclusion Statistic calculation. Therefore, consideration should be given to exclude that locus from the calculation. A combined Probability of Inclusion/Exclusion Statistic can be calculated for any number of loci, as any number of loci can be omitted from the calculation.

Where only one or two loci are being considered for inclusion in the statistical interpretation, the probative value of the DNA findings may be very marginal. Care should be taken not to overstate findings by simply including a contributor as a potential donor to the mixed profile. Reference should be made as to the weight of the inclusion, so those considering the DNA report make no false assumptions.

The resulting calculation will determine how many individuals in a given population are potential contributors to the mixed profile seen. The calculation includes all possible genotypic combinations for each locus, and provides an individual statistic for that locus. The suitable loci are then multiplied together to obtain a combined statistic using the product rule. This combined statistic includes the theta correction factor for the appropriate racial population selected from the available populations' databases (29).

Sample Conclusions

a) Probability of Inclusion:

A mixed DNA profile was obtained from the blood identified on the running shoe, exhibit 1 (scene). The donor of the known blood sample, exhibit 2 (Smith, John) is a potential contributor to the mixed profile. The number of individuals selected randomly from the American Caucasian population that are a potential contributor to the mixed profile is one in one million, or 0.0001 percent.

b) Probability of Exclusion:

A mixed DNA profile was obtained from the blood identified on the running shoe, exhibit 1 (scene). The donor of the known blood sample, exhibit 2 (Smith, John) is a potential contributor to the mixed profile. The percentage of randomly selected individuals from the American Caucasian population that can be excluded as a potential contributor to the mixed profile is 99.99990 percent.

Identity Statement

An identity statement, or single source attribution statement, may be applied to DNA profiles under the following conditions:

1. The DNA profile from the questioned (scene) sample is from a single contributor, and matches at all loci to a known (reference) profile **OR**

2. The DNA profile from the questioned (scene) sample is of mixed origin, but a donor can be unambiguously determined, and matches at all loci to a known (reference) profile **AND**
3. The Random Match Probability of the DNA profile is 1 in 280 billion or less (3) **AND**
4. No identical twin to the donor of the known suspect reference sample is known to exist. If an identical twin is known, this should be reflected in the conclusion.

Sample Conclusions

The DNA profile obtained from the blood identified on the running shoe, exhibit 1 (scene), matches that of the known blood sample, exhibit 2 (Smith, John), and therefore originated from a common source, or an identical twin of the known sample donor, to a reasonable degree of scientific certainty.

If no identical twin is known to exist:

John Smith (exhibit 2) is the source of the DNA obtained from the blood identified on the running shoe, exhibit 1 (scene), to a reasonable degree of scientific certainty.

Note: The term “matches” may be replaced with the term “is indistinguishable from”.

DISCUSSION

The comparison of questioned samples, many of which originate from crime scenes, to samples of known origin, takes on many forms. In many comparisons, features used for comparison are divided into two main categories. The first category is the class characteristic. Many items of a similar nature, for example running shoes with the same size and tread design, have the same class category. While the manufacturer and shoe size may be determined by a shoe print at a crime scene, class characteristics alone are insufficient to identify a shoe print back to a particular shoe, to the exclusion of all other shoes with the same class characteristics. As shoes are mass produced, there may be many shoes of the same type that possess the same class characteristics.

The second feature category is the random individual characteristic. These characteristics are introduced through a series of randomly occurring events. In the shoe example, stepping on various items, such as sharp metal, glass, or stones, randomly imparts variation in the shoe sole, making it individual. These characteristics, while perhaps not capable of identifying an origin uniquely when taken one at a time, are sufficiently rare and accidental in nature, that when taken together in combination are sufficient for identification. Dependent on the rarity and detail seen, one, two, or three random individual characteristics may not be of sufficient discriminating power to eliminate all other potential shoes besides the shoe in question (30). However, if the number and rarity of these random individual characteristics surpasses a certain threshold, an identity statement can be issued. Often, this threshold is determined by the training and experience of the forensic examiner.

In many ways, comparisons of DNA profiles between questioned and known samples are very similar to a shoeprint comparison. A major asset of forensic DNA technology is its ability to both include and exclude suspects with tremendous discriminating power. Each locus analyzed by PCR STR DNA typing is in essence a class characteristic, common to human beings. Similarly, the alleles within each locus are random individual characteristics. No single allele in the areas currently examined is individualizing, however when taken in combination with all of the other alleles from all of the other loci, sufficient power of discrimination is present that an identity statement can be issued.

Within the various traditional fields of forensic science, the first objective of a comparison is to search for significant differences. An unbiased scientist will seek first to exclude. If a very reliable questioned sample profile does not contain an allele borne by a suspect, then that suspect can be eliminated with certainty. What constitutes a very reliable questioned sample profile? It is a profile with peaks above the predetermined threshold, with no smaller peaks visible, and no other reasonable explanation for a peak to be missing, such as severe degradation. With profiles obtained from degraded samples, often no alleles are found in the higher molecular weight loci, or the peaks present are considerably smaller than those of higher molecular weight loci. If those peaks are at or below the lower threshold, those loci may be considered uninformative. A missing allele in this case may not be grounds for elimination. If on the other hand an allele found in the known profile is not found in the very reliable loci in the questioned sample, the donor of the known profile can be eliminated as a potential contributor with full confidence.

Once a suspect profile cannot be excluded, and all random individual features (alleles) match, then conversely, the suspect is included as a potential donor to the questioned profile. The next question becomes one of weight. How much significance can now be placed upon the association established by the match between the questioned and known samples? Assessing the significance of a match inherently includes information surrounding the case circumstances (17). Figures included for discussion will provide points of illustration for treatment of mixture types, and attributed associative weight. If a single donor accounts for all alleles found in the questioned sample, with all loci matching, opinions have been expressed that a source can be identified within a scientific certainty (2, 3). If the deduction of the suspect profile from a mixed DNA profile is unambiguous, this means that no other single source DNA profile can produce those corresponding alleles. Therefore, an identity statement can also be made in these cases.

There are two main sources of information that are of value for interpreting mixed PCR STR DNA profiles. The first is the fragment length size of the individual STR fragments found for each of the multiple loci analyzed. In analysis, this fragment length is determined by retention time, which in turn can be compared to known allelic ladders. Alleles are then designated for each of these fragments. The second source of information is the magnitude of signal or signal strength generated by these alleles as they are detected. As alleles are visualized as peaks, the signal strength is represented in both peak height and peak area. As peak height and area are directly related to each other, for the purposes of this discussion, peak height will be used throughout. The two sources of information, alleles and peak height, are independent from each other. Under proper conditions, these two features can both be used to deduce the individual contributors to the mixed profile.

Peak height bears information useful in interpreting contributors to mixed profiles as peak height is related to amounts of DNA found in the original sample. In order to use peak heights to deduce the DNA profiles of individual contributors to mixed DNA profiles, one must assume that peak heights are positively related to DNA concentration. Original amounts of DNA in the mixture of biological materials are present in a certain ratio, which will be termed the mass ratio. If the starting mass ratio is 3:1, based on the conservation of this mass ratio in the Polymerase Chain Reaction, we expect the peak heights of the major contributor to be 3 times larger than those of the minor contributor. Therefore, when the original mass of DNA template differs sufficiently between major and minor contributor, we can use the peak heights to determine which alleles can be attributed to which contributor. That is, the higher peaks found with 3 times the contribution relative to the lower peaks, can be assigned to the major contributor. Likewise, the lower peaks can be assigned to the minor contributor.

Within the DNA profiles generated in forensic STR PCR DNA profiling systems, there is an optimal peak height window where the peak height is most linearly related to the original DNA mass contribution. Peak heights that fall above or below this window are compromised for use in extrapolating contributors, as the assumption of the positive linear relationship of peak height to original mass is compromised. Peaks above the upper threshold saturate the instrumental detection system, and therefore reach a maximum peak height with high DNA concentrations. As DNA concentrations further increase, higher peaks are held at that saturation point, while lower peaks continue to increase in height relative to their original DNA contribution. As a result, linearity of the peak height to DNA mass relationship is lost between the major and minor components of a mixture.

At the lower end of the optimal peak height window, minute quantities of DNA or degraded DNA may result in peaks with very low peak height. At these lower concentrations, variations in amplification become more pronounced, and peak heights may vary more widely relative to DNA concentration. Degradation also will have a greater effect on peak height with less original template DNA. The lower peak height threshold ensures deduction of profiles with data in the optimal range of instrument sensitivity. The establishment of minimum lower and maximum upper peak height thresholds required for profiles prior to their interpretation is therefore suggested. These lower and upper peak height thresholds may be established for each laboratory, based on validation studies, mixture study publications (13–16), the profile in question, and examiner training and experience. Loci containing peaks above or below these thresholds may be used in the interpretation for inclusion or elimination purposes, but should not be considered in terms of designating a major or minor contributor, nor used in assigning a statistical significance to findings. In cases such as these, a rerun of the sample is suggested after appropriate adjustment of sample concentration.

As previously stated, degraded or otherwise compromised DNA samples are of concern, as mass ratios may not be conserved throughout the profile. Mixtures of biological materials of differing ages, cell types, and varying environmental conditions may degrade at different rates, possibly calling the assumption of the preservation of mass ratios into question. Therefore, it is at the analyzing scientist's discretion and best judgment to observe the sample prior to analysis, and the resultant DNA profile, to determine if the assumption will be held. Compromised and environmentally challenged DNA samples routinely produce interpretable mixed DNA profiles. This is part of the great value of the PCR STR DNA technique. Such profiles must be interpreted with caution, however, to ensure environmental extremes do not adversely influence a correct interpretation. If the peak height information is seen to be of sufficient variation to be compromised in reliably deducing major and minor contributors, allele information may carry useful information. Exclusive of peak height, the presence or absence of alleles in mixtures is very significant, particularly in loci where less degradation is seen. This significant information borne by the alleles may be captured in the more conservative Probability of Inclusion/Exclusion approach.

Use of peak heights in the interpretation of mixed DNA profiles is based upon the following assumptions:

Mass ratios are conserved:

- a. Between individuals within the mixture
- b. From locus to locus
- c. Within the locus
 - i. Within the individual

ii. Between the individuals

iii. Addition of shared peak height/area is conserved

It is suggested that the mixed DNA profile be first approached as a whole, to determine its suitability for interpretation. True alleles must be designated, with artifacts and anomalies accounted for and removed from the profile. With an objective examination the deduction of contributing individual profiles to the mixture is conducted prior to comparison to the known reference profiles. By observing the mixed profile first, the individual profiles of contributing profiles can be deduced by matching together alleles with corresponding peak heights. The notable exception is the case of a major contributor with a trace minor contributor. As discussed previously, consultation of the known reference sample of the trace minor contributor may prevent an incorrect elimination.

Peaks that are shared between contributors are magnified by a contribution (or dose) from more than one individual. Conversely, multiple doses of one allele can also result because of a homozygote, as in an individual with two copies of the same allele, rather than one copy of two different alleles (heterozygote). Where the peak is shared between more than one individual, the height of this multiple dose peak can be estimated by combining the peak height of the lower non-shared peaks. We assume that the height of the larger shared peak is the sum of the smaller non-shared peaks. That is, the addition of shared peak heights is conserved. Where the peak is a result of a homozygote, the peak height is divided by 2 to account for the double dose of DNA received from a single contributor.

Shared peaks can be particularly problematic, as they combine two potential errors. First, there is some variation in the size of each smaller peak, and how close it is to the true mass ratio of its corresponding original biological material. Second, when combined with the allele of the other contributor, this shared peak will also have its own variation from the original mass ratio. This variation, coupled with the potential error made in the assumption that the smaller peaks accurately reflect the shared peak heights, may cause potential errors in interpretation. Peak area may also be consulted to scrutinize whether multiple doses are present. Area may be less affected by these variations as height, so could add a valuable cross check in difficult cases. Therefore, loci with shared alleles must be approached with greater caution than those where alleles are not shared.

While ideally peak heights bear a positive linear relationship to the mass ratio of the original biological material, there is an amount of variation inherent in the PCR system. This variation is dependent on a number of variables, but can be quantified for individual laboratories through repetitive studies conducted in validation. Applied Biosystems Inc. (Foster City, CA) has suggested in their Profiler Plus/Cofiler user's manual that ratios of less than 70% for allele peak height imbalance of heterozygotes are rare (31). Reamplification is suggested in these cases. Generally, peak height balances are in the range of 10 to 15% for a normal range of non-severely degraded samples. Using ABI data, locus D21S11 shows the largest peak imbalance among the 13 core CODIS (Combined DNA Index System – FBI) loci, with a standard deviation of $\pm 7\%$. At a confidence level of 99%, with three standard deviations variation in heterozygote peak height ratio, approximately $\pm 21\%$ indicates the amount of variation most commonly seen in pure single source DNA samples (32). Therefore, with 99% confidence one can conclude that heterozygote peak imbalance should be less than 21% with an ample quantity of non-degraded DNA. Hence, the separation of major and minor contributors in a mixed DNA profile should be outside this range of heterozygote peak height variation to be reliable.

In addition to the “normal” range of heterozygote peak imbalance or variation seen, in order for a peak to be unambiguously assigned to a particular donor, there should also be an additional range of safety added for conservativeness. Severely challenged forensic samples may be beyond the range of samples tested in the system validation. Larger peak height imbalance demonstrates a corresponding erosion of the assumption of the positive linear relationship of peak height to the original DNA mass ratio. Therefore, should evidence be seen of severe degradation, consideration may be given to disregard peak height information as potentially unreliable. Allele designations may still be informative, and may still be used in cases of this type, however.

Likewise, very minute quantities of DNA may affect the assumption of a positive linear peak height to mass ratio. Krenke et al. (33) found that quantities of original template DNA below 0.5 ng frequently result in larger heterozygote peak height variations. This variation causes a greater potential imbalance in peak height ratios, and therefore decreases the confidence within which contributors can be deduced in affected loci. As higher molecular weight loci tend to be the most susceptible to degradation, often the peak heights of these loci will be the first to be compromised. Caution should also be exercised in comparing allele heights within these larger STR loci, as they may show a loss of peak height in larger alleles versus the smaller alleles. This loss of peak height in larger loci and alleles may be accentuated with very minute DNA quantities. Therefore, the entire profile should be examined for this decrease in peak heights, and those loci and alleles showing any marked size reduction should be examined with care. A recent workshop presentation and discussions provide additional detail in the area of mixed profiles and assessing minor components (34).

The LSD program can be used to confirm the deduction of the individuals comprising the mixed DNA profile. The LSD program compares the peak heights of the mixed profile to the gambit of two-person mixture possibilities on a locus-by-locus basis, and provides a measure of best fit for each possible contributor. The term “Least Squares” refers to the method used to compare the data to ideal models. The data from each peak in the mixed profile is compared to the ideal peak height for a particular theoretical mixture. For example, a 2:1 mass ratio could be expected to generate a theoretical four peak profile with peak heights of 1000, 1000, 500, and 500 rfu under ideal conditions. The difference between the real profile and the ideal is calculated and then squared, and summed for the peaks at each locus, creating a best fit value. This calculation is repeated for each possible two-person mixture combination. The ability to compare the fit of the mixture to the various possible combinations is a helpful tool in determining which possibilities are viable, particularly when taken in conjunction with the major:minor ratios also calculated by the LSD program. Mathematically, if the LSD program can unambiguously pick individual(s) from the mixture, this will add confidence to the scientist’s determination through an independent and objective analysis. Depending on the case scenario, if a forensically valid assumption can be made, the LSD program approach may include use of a known reference profile in deducing a second contributor.

Use of the LSD program may reveal trends with respect to limitations in the DNA analytical system, which in turn should be reflected in more conservative interpretations. Minute quantities of initial DNA template may increase the incidence of stochastic effects, and thereby increase the level of error in the assumed positive linear relationship of starting DNA mass ratio to peak height. Interpretation of DNA profiles developed from cellular material and biological material with very low DNA quantity have been dealt with in specific articles (25-27). Caution must be exercised in interpreting these profiles.

Use of the Least Squares Deconvolution (LSD) program can objectively demonstrate the amount of separation in peak height magnitude required in order to distinguish major

and minor contributors. Repeated use of the program has established a baseline where correct donors can reliably be designated. At a 1:1 mass ratio, there is no significant peak height difference between contributors; therefore peak height data is uninformative. As peak heights diverge from the 1:1 mass ratio, peak heights of the major contributor grow relative to the loss of height of the minor contributor. Nearer 1:1, the variation in the PCR STR system itself is sufficiently large as to compromise reliable separation of contributors. Moving away from 1:1, at some point the difference in peak heights allows an unambiguous separation of major and minor contributors for loci with non-shared peaks. The peaks with larger height are paired, as are the peaks with smaller height. The LSD program has demonstrated reliable interpretation of loci with unshared alleles at a 1.5:1 mass ratio and above (8), with the exception of the 2:1 mass ratios found in special circumstances as discussed next.

Loci with multiple dose alleles must be handled differently. The term “multiple dose” means that the peak height represents contributions from more than a single allele. This may be in the form of a homozygote (two doses from one individual), or in a peak shared between more than one individual, or both. As mass ratios approach 2:1, loci with shared peaks experience an interesting limitation, aside from the increased variation in shared peak height discussed earlier. Mathematically, several distinctly different combinations of events can be represented by the same combination of three peaks with equal heights. As demonstrated in Figure 1, any combination of a 2:1 mass ratio of homozygote and non-shared heterozygote produces the same profile. As a result, the peak height data is not fully informative in this case without comparison to another profile.

Likewise, two-peak profiles must be treated specially for mixtures with a 2:1 mass ratio. Two-peak profiles wherein the larger peak is two times the height of the smaller peak could be produced by two distinctly different mixture contributor scenarios (See Figure 2). First, both the major and the minor contributor could be homozygous and unshared. Secondly, the major contributor could be heterozygous, and share one allele with a homozygous minor contributor. Both possible scenarios yield the same mixed DNA profile. Therefore, the profiles of the contributing individuals cannot be unambiguously deduced. In this two-peak profile instance, without another profile to be used in deducing contributors, the peak height data alone is not conclusive. A more conservative approach may therefore be warranted for profiles in these special circumstances, such as dropping the loci from the calculation, or adding the probability of both variations to the statistical calculation.

As a result of these two possible scenarios, it is recommended that mass ratios between 2:1 and 1:1 not be considered for use with LSD for three- and two-peak profiles. Coupled with the variation within the technique as described above, an additional buffer zone to encompass this area increases the mass ratio from 2:1 to 2.3:1. Four-peak loci may be very informative below this point, and may be treated separately. For simplicity, and ease of explanation, a single approach may be desired over a combined approach. Cases with mass ratios between 2.3:1 and 1:1 may be best interpreted using the Probability of Inclusion/Exclusion approach, as described in greater detail as Mixture Type 6 and in Appendix I.

In some circumstances, it is recommended that statistical interpretation be conducted with extra attention. A very complex mixture where there is not a distinct major contributor is one example. Very commonly handled objects often yield complex mixtures. Mixtures with greater than three contributors often hold less significant statistical value attributed to an inclusion. As it is often impossible to place a timeline on the deposition of biological material, as in it may have been deposited prior to the offense, and therefore

| | |
|--------------------------|-------|
| Three-peak mixed profile | ^ ^ ^ |
| Scenario A: | |
| Major contributor | ^ ^ |
| Minor contributor | ^ |
| Scenario B: | |
| Major contributor | ^ ^ |
| Minor contributor | ^ |
| Scenario C: | |
| Major contributor | ^ ^ |
| Minor contributor | ^ |

Figure 1. Three-peak mixed DNA profile with a 2:1 mass ratio.

may bear no connection with the true perpetrator, caution must be exercised. This is not to say that items such as doorknobs, light switches, and the like should not be sampled, but these instances are very case-dependent. If it is obvious that the attacker must have been the last person to handle an item, then the forensic scientist may consider it a worthwhile endeavor to attempt the analysis. This is provided that those individuals with rightful access to that item are not wrongfully implicated. Finding the DNA profile of a person with common access to a crime scene is of limited forensic value. Finding the DNA profile of a person who does not belong at a crime scene may be of great value, and therefore this determination depends heavily upon the case specifics. With the popularity of crime shows in the popular media, the public may now be assigning very large weight to DNA associations. The danger of not including a statistical evaluation of an association is that the value of a DNA inclusion may be over-weighted in the absence of the statistic. Therefore, it is recommended that statistics be provided with associations of limited value, lest they be misinterpreted.

Case for Demonstration

Consider the mixed DNA profile in Figures 3 and 4 under different crime scene scenarios. First, consider this mixed profile as being found on a nondescript white cotton T-shirt found in a dumpster in an alley a block away from a residence where a murder has been committed. A bloodstain was identified, and this mixed DNA profile was produced from that stain using Profiler Plus and Cofiler STR multiplex kits (ABI – Foster City, CA). In this case, no assumptions regarding origin of contributors to the bloodstain can be reasonably made. Observing the loci with unshared peaks, the mass ratio can be calculated as follows:

| | |
|------------------------|-----|
| Two-peak mixed profile | Λ Λ |
| Scenario A: | |
| Major contributor | Λ |
| Minor contributor | Λ |
| Scenario B: | |
| Major contributor | Λ Λ |
| Minor contributor | Λ |

Figure 2. Two-peak mixed DNA profile with a 2:1 mass ratio.

$$\text{VWA: } (1318 + 1200) / (621 + 793) = 2518 / 1414 = 1.78$$

$$\text{D8: } (1431 + 986) / (603 + 560) = 2417 / 1163 = 2.08$$

$$\text{D18: } (1181 + 1029) / (471 + 386) = 2210 / 857 = 2.58$$

$$\text{D16: } (1262 + 1253) / (515 + 514) = 2515 / 1029 = 2.44$$

$$\text{Average Mass Ratio: } (1.78 + 2.08 + 2.58 + 2.44) / 4 = 2.22$$

Using the LSD program (Figure 5), the mass ratios vary from 1.8 (vWA) to 3.0 (THO1). As the ratio is approximately 2.3:1, a deduction of individual contributors could be attempted under Mixture Type 4. Note that with use of the LSD program, mathematically many loci in this mixed profile can be separated into two unambiguous contributors. Note also the loci D21, THO1 and CSF. With the mass ratio near 2:1, as the peaks near the same height, confidence in an unambiguous deduction of source profiles is reduced. The second choices in the LSD program have mass ratios of 1.6 to 1.7. As these mass ratios are close to our expectation, confidence in an unambiguous deduction is reduced.

Likewise, with the two-peak profiles seen in the locus D13, a choice between one of the two competing scenarios is not distinct. This point is reflected in the LSD program (Figure 5), as both choices reflect an acceptable mass ratio. Scenario A of two sole source homozygote contributors (Figure 2) has a mass ratio of 1.9:1. Scenario B of a heterozygote major contributor sharing one peak with a homozygote minor contributor (Figure 2), has a mass ratio of 2.2:1. Therefore, either scenario is a reasonable possibility, and therefore this locus can be deemed ambiguous. This profile was chosen as an example to illustrate that where mass ratios approach 2:1, deduction of source profiles should be done with caution where no known profiles can be used in the interpretation.

As a result of the ambiguous loci found in the deduction process above, a decision is taken to interpret this mixed DNA profile under Mixture Type 6. The Probability of Inclusion/Exclusion approach is used for this mixed crime scene DNA profile under this set of circumstances. Alleles are input into the SDRMix program (28) as indicated in Figure 6. The resulting combined Probability of Inclusion is 1 in 4.83×10^5 or 1 in 483,000, and the Probability of Exclusion is 0.9999979.

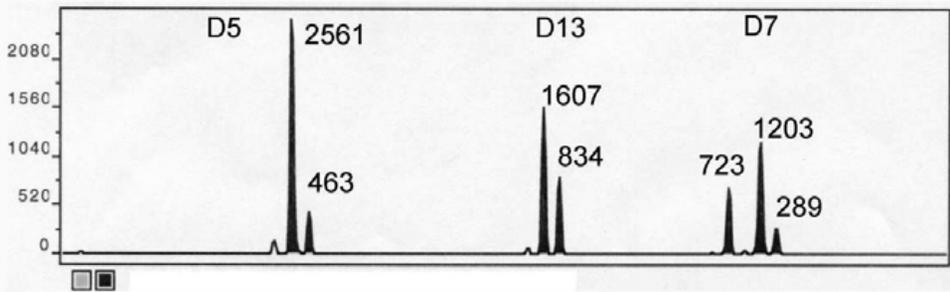
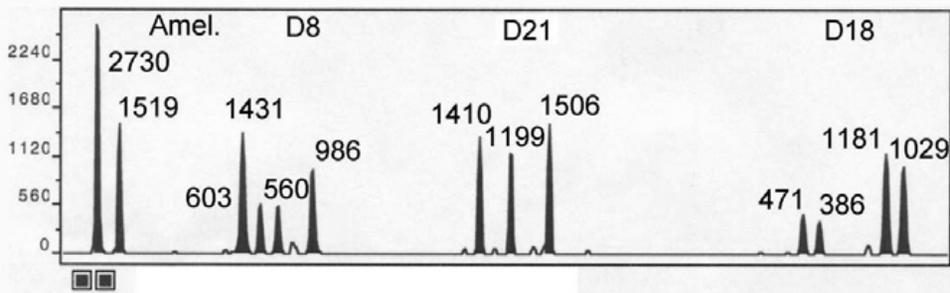
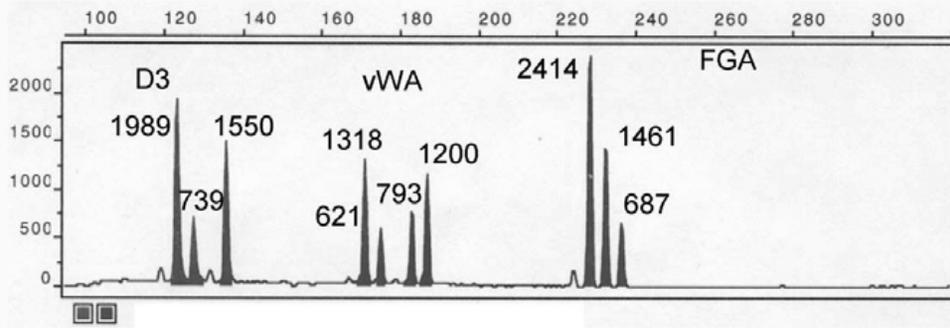


Figure 3. Mixture in Profiler Plus.

The conclusion could be stated as follows:

a) Probability of Inclusion:

A mixed DNA profile was obtained from the blood identified on the white T-shirt, exhibit 1 (dumpster). The donor of the known blood sample, exhibit 2 (Smith, John) is a potential contributor to the mixed profile. The number of individuals selected randomly from the American Caucasian population that are a potential contributor to the mixed profile is 1 in approximately 483,000, or 0.00021 percent.

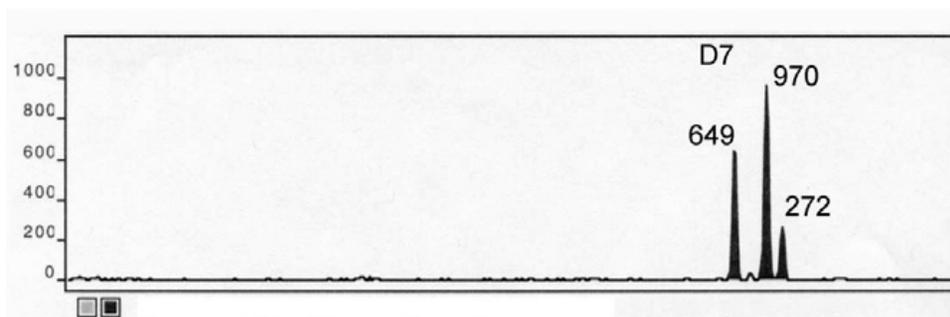
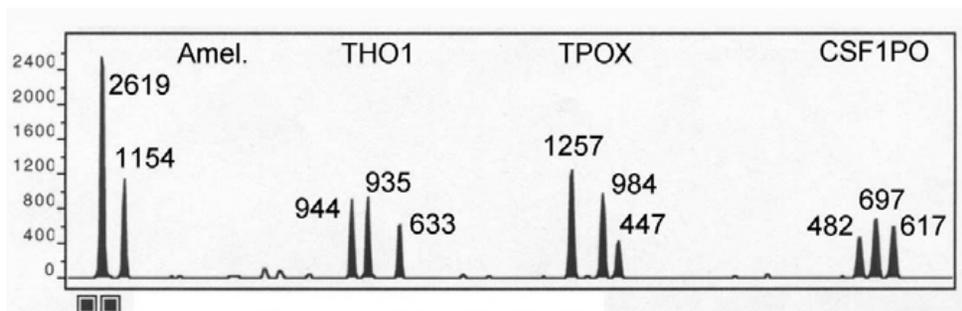
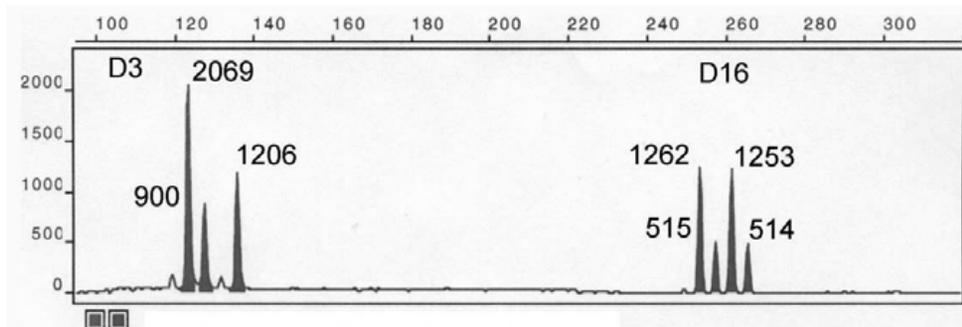


Figure 4. Mixture in Cofiler.

b) Probability of Inclusion/Exclusion:

A mixed DNA profile was obtained from the blood identified on the white T-shirt, exhibit 1 (dumpster). The donor of the known blood sample, exhibit 2 (Smith, John) is a potential contributor to the mixed profile. The percentage of randomly selected individuals from the American Caucasian population that can be excluded as a potential contributor to the mixed profile is 99.99979 percent.

Next, consider the same mixture, but under different case circumstances. This time, the DNA profile is obtained from blood identified under the fingernails of a female com-

| D3S1358- Three Alleles | | | | | | |
|----------------------------------|------|-----------|----------|---------------|-------------|------------|
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 15,16 | 15,18 | 5.5E-02 | 1.0 | 1.0:2.3 |
| <input type="radio"/> | 2 | 16,16 | 15,18 | 0.18 | 3.2 | 1.0:4.8 |
| <input type="radio"/> | 3 | 15,15 | 16,18 | 0.60 | 11 | 1.0:1.2 |
| <input type="radio"/> | 4 | 16,18 | 15,18 | 0.85 | 15 | 1.0:4.6 |
| <input type="radio"/> | 5 | 18,18 | 15,16 | 1.43 | 26 | 1.0:1.8 |
| <input type="radio"/> | 6 | 16,18 | 15,16 | 4.79 | 87 | 1.0:1.7 |
| TH01- Three Alleles | | | | | | |
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 8,8 | 5,6 | 1.0E-04 | 1.0 | 1.0:3.0 |
| <input type="radio"/> | 2 | 5,5 | 6,8 | 0.11 | 1126 | 1.0:1.7 |
| <input type="radio"/> | 3 | 6,6 | 5,8 | 0.12 | 1194 | 1.0:1.7 |
| <input type="radio"/> | 4 | 5,8 | 5,6 | 0.32 | 3205 | 1.0:1.7 |
| <input type="radio"/> | 5 | 6,8 | 5,6 | 0.34 | 3392 | 1.0:1.7 |
| <input type="radio"/> | 6 | 6,8 | 5,8 | 1.29 | 12778 | 1.0:1.0 |
| D21S11- Three Alleles | | | | | | |
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 30,30 | 28,32.2 | 3.2E-03 | 1.0 | 1.0:2.4 |
| <input type="radio"/> | 2 | 32.2,32.2 | 28,30 | 1.5E-02 | 4.8 | 1.0:1.7 |
| <input type="radio"/> | 3 | 28,28 | 30,32.2 | 3.3E-02 | 10 | 1.0:1.9 |
| <input type="radio"/> | 4 | 30,32.2 | 28,32.2 | 0.28 | 88 | 1.0:1.3 |
| <input type="radio"/> | 5 | 28,30 | 28,32.2 | 0.39 | 121 | 1.0:1.4 |
| <input type="radio"/> | 6 | 28,30 | 30,32.2 | 0.68 | 213 | 1.0:1.1 |
| D18S51- Four Alleles | | | | | | |
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 12,13 | 17,18 | 0.10 | 1.0 | 1.0:2.6 |
| <input type="radio"/> | 2 | 13,18 | 12,17 | 3.08 | 30 | 1.0:1.2 |
| <input type="radio"/> | 3 | 12,18 | 13,17 | 3.17 | 31 | 1.0:1.0 |
| D5S818- Two Alleles | | | | | | |

Figure 5. LSD Results (Courtesy of Tse Wei Wang)(8).

| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
|----------------------------------|------|----------|----------|---------------|-------------|------------|
| <input type="radio"/> | 1 | 13,13 | 12,12 | 0.0E+00 | | 1.0:5.5 |
| <input checked="" type="radio"/> | 2 | 12,13 | 12,12 | 0.0E+00 | | 1.0:2.3 |
| <input type="radio"/> | 3 | 13,13 | 12,13 | 0.0E+00 | | 1.0:-2.4 |
| <input type="radio"/> | 4 | 12,13 | 12,13 | 10.27 | | 1.0:1.0 |
| D13S317- Two Alleles | | | | | | |
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input type="radio"/> | 1 | 12,12 | 11,11 | 0.0E+00 | | 1.0:1.9 |
| <input checked="" type="radio"/> | 2 | 11,11 | 11,12 | 0.0E+00 | | 1.0:2.2 |
| <input type="radio"/> | 3 | 12,12 | 11,12 | 0.0E+00 | | 1.0:-4.2 |
| <input type="radio"/> | 4 | 11,12 | 11,12 | 0.43 | | 1.0:1.0 |
| D7S820- Three Alleles | | | | | | |
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 10,11 | 8,10 | 0.15 | 1.0 | 1.0:2.2 |
| <input type="radio"/> | 2 | 8,11 | 10,10 | 1.13 | 7.7 | 1.0:1.2 |
| <input type="radio"/> | 3 | 11,11 | 8,10 | 1.38 | 9.5 | 1.0:6.7 |
| <input type="radio"/> | 4 | 8,11 | 8,10 | 2.36 | 16 | 1.0:29.0 |
| <input type="radio"/> | 5 | 8,8 | 10,11 | 5.00 | 34 | 1.0:2.1 |
| <input type="radio"/> | 6 | 8,11 | 10,11 | 10.70 | 73 | 1.0:3.7 |
| D16S539- Four Alleles | | | | | | |
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 12,14 | 11,13 | 1.6E-04 | 1.0 | 1.0:2.4 |
| <input type="radio"/> | 2 | 13,14 | 11,12 | 2.09 | 13465 | 1.0:1.0 |
| <input type="radio"/> | 3 | 12,13 | 11,14 | 2.09 | 13465 | 1.0:1.0 |
| CSF1PO- Three Alleles | | | | | | |
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 10,10 | 11,12 | 1.4E-02 | 1.0 | 1.0:2.7 |
| <input type="radio"/> | 2 | 11,11 | 10,12 | 3.9E-02 | 2.8 | 1.0:1.6 |
| <input type="radio"/> | 3 | 12,12 | 10,11 | 9.9E-02 | 7.2 | 1.0:1.9 |
| <input type="radio"/> | 4 | 10,11 | 11,12 | 0.23 | 17 | 1.0:1.4 |
| <input type="radio"/> | 5 | 10,12 | 11,12 | 0.45 | 33 | 1.0:1.7 |
| <input type="radio"/> | 6 | 10,12 | 10,11 | 0.99 | 72 | 1.0:1.2 |

Figure 5. Continued.

| VWA- Four Alleles | | | | | | |
|----------------------------------|------------|----------|----------|---------------|-------------|------------|
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 16,18 | 15,19 | 5.6E-02 | 1.0 | 1.0:1.8 |
| <input type="radio"/> | 2 | 16,19 | 15,18 | 0.79 | 14 | 1.0:1.2 |
| <input type="radio"/> | 3 | 15,16 | 18,19 | 0.84 | 15 | 1.0:1.0 |
| D8S1179- Four Alleles | | | | | | |
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 13,14 | 12,16 | 0.32 | 1.0 | 1.0:2.1 |
| <input type="radio"/> | 2 | 14,16 | 12,13 | 1.38 | 4.3 | 1.0:1.3 |
| <input type="radio"/> | 3 | 13,16 | 12,14 | 1.44 | 4.5 | 1.0:1.3 |
| TPOX- Three Alleles | | | | | | |
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 8,11 | 8,10 | 5.1E-02 | 1.0 | 1.0:2.4 |
| <input type="radio"/> | 2 | 11,11 | 8,10 | 0.19 | 3.7 | 1.0:5.0 |
| <input type="radio"/> | 3 | 8,8 | 10,11 | 0.72 | 14 | 1.0:1.1 |
| <input type="radio"/> | 4 | 10,11 | 8,10 | 0.86 | 17 | 1.0:4.9 |
| <input type="radio"/> | 5 | 10,10 | 8,11 | 1.64 | 33 | 1.0:1.7 |
| <input type="radio"/> | 6 | 10,11 | 8,11 | 5.37 | 106 | 1.0:1.7 |
| FGA- Three Alleles | | | | | | |
| Select | Case | Person 1 | Person 2 | Fitting Error | Error Ratio | Mass Ratio |
| <input checked="" type="radio"/> | 1 | 21,23 | 21,22 | 5.0E-02 | 1.0 | 1.0:2.0 |
| <input type="radio"/> | 2 | 22,23 | 21,21 | 0.63 | 13 | 1.0:1.1 |
| <input type="radio"/> | 3 | 23,23 | 21,22 | 0.96 | 19 | 1.0:5.6 |
| <input type="radio"/> | 4 | 22,23 | 21,22 | 1.90 | 38 | 1.0:13.3 |
| <input type="radio"/> | <TD ALIGN= | CENTER | | | | |

Figure 5. *Concluded.*

plainant of a brutal assault. Under these circumstances, a forensically valid assumption can be made with respect to the female contributor to the mixed DNA profile. Mixture Type 2 guidelines would therefore apply. Again, the mixed DNA profile is viewed independently of known samples, and alleles designated. Next, the contribution from the DNA profile from the assumed source is accounted for within the mixed profile, enabling the deduction of a single source DNA profile (See Figures 7 and 8, and Table 1). This profile can now

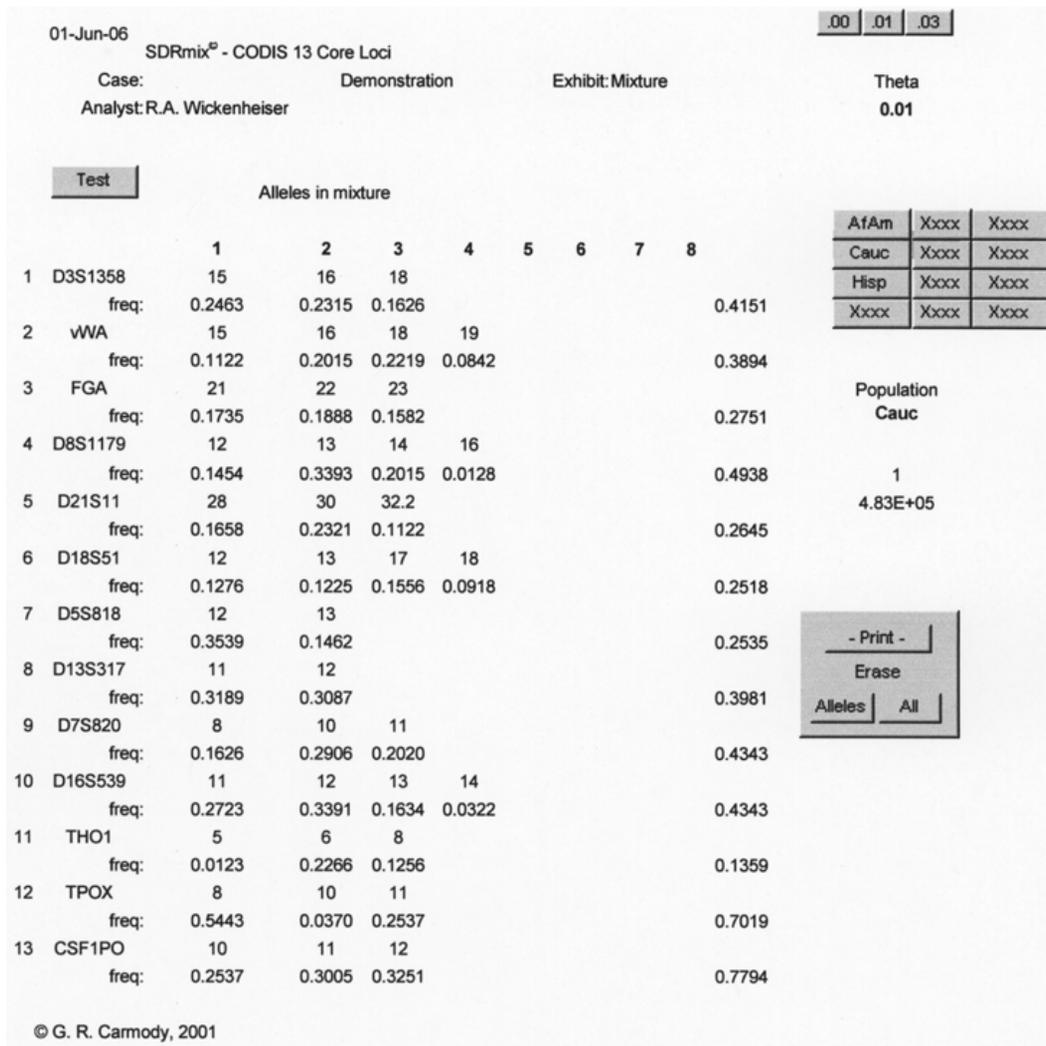


Figure 6. A screen shot of the SDRmix program (Courtesy of George Carmody)(28).

be compared to suspect profiles. Care should be taken that resulting allele heights and the resultant profile fit the expected outcome for a single source profile. Note that with the forensically valid assumption (Mixture Type 2), the male contributor can be unambiguously deduced. If that resulting deduced profile is subsequently matched to a known reference sample, an identity statement can be issued regarding the source (See identity statement).

Finally, consider the same mixed DNA profile, but now as the male fraction of a differential profile from a vaginal swab from a complainant of a sexual assault. In this case, the “known” profile in Figures 7 and 8 is the female fraction of the vaginal swab. This scenario would dictate use of Mixture Type 1 guidelines. Once again, the female profile from Figures 7 and 8 is used to deduce the male component from the mixed DNA profile. If there is premature lyses of sperm or epithelial cells from the male component seen in the non-sperm fraction, this additional information is used to assist in assigning alleles to the female or the male donor. Again, if an unambiguous profile can be deduced from the mix-

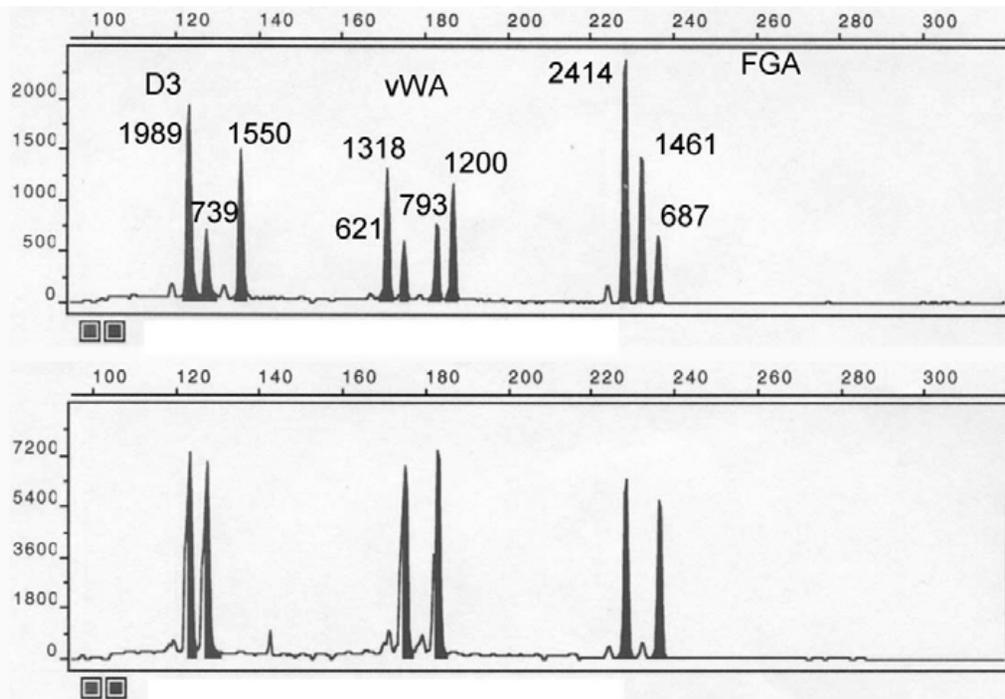


Figure 7. Mixture with single source sample in Profiler Plus.

ture, and this profile is subsequently matched to a known reference sample, an identity statement can be issued regarding the source (See identity statement).

CONCLUSION

Deduction of individual contributors in mixed STR DNA profiles can be a very complicated, laborious process. As described in these guidelines, separation of mixed DNA profiles into six specific types simplifies this process, and provides a stepwise comprehensive framework for interpretation.

The separation of mixed DNA profiles into distinct types retains significant case specific information by considering relevant known DNA profiles. In differential extractions and sub-sampled items, consideration of a known profile and/or more than one profile in deducing another contributor is appropriate. If a forensically valid assumption can be made regarding the presence of a contributor expected to be present, then consideration of the contribution of a known sample can be used in deducing another contributor. If no assumptions can be made based on case specifics, a more conservative interpretative approach is warranted. Where appropriate, indication of a sole source may be possible. Should peak height data prove uninformative, a more conservative Probability of Inclusion/Exclusion approach to statistical interpretation is suggested.

These mixture interpretation guidelines can be adapted to reflect inter-lab variations through incorporating upper and lower thresholds and mass ratios established in validation studies. As a result, the guidelines may serve as a template, which can be modified to fit different forensic DNA laboratories.

This approach to DNA mixture interpretation provides a conservative, informative, straightforward, and systematic approach to the interpretation and resolution of a DNA profile containing more than one contributor.

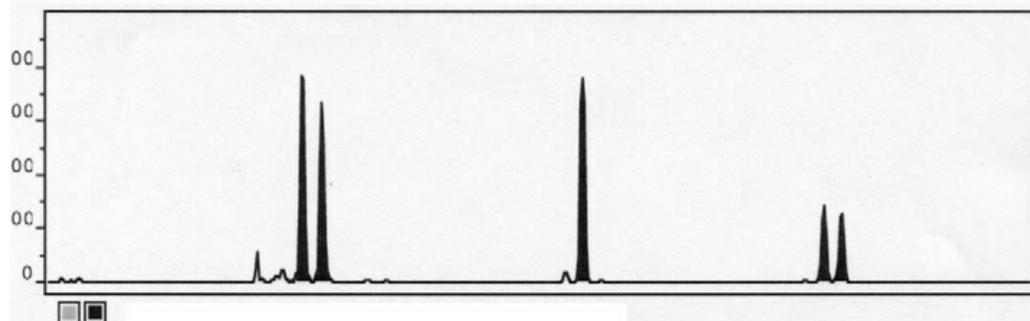
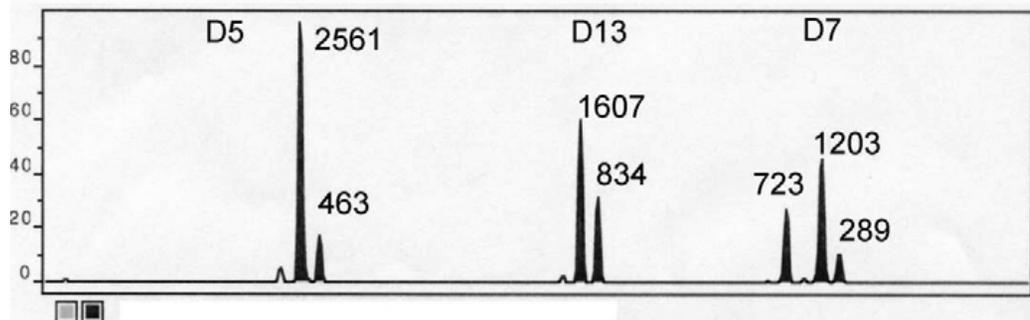
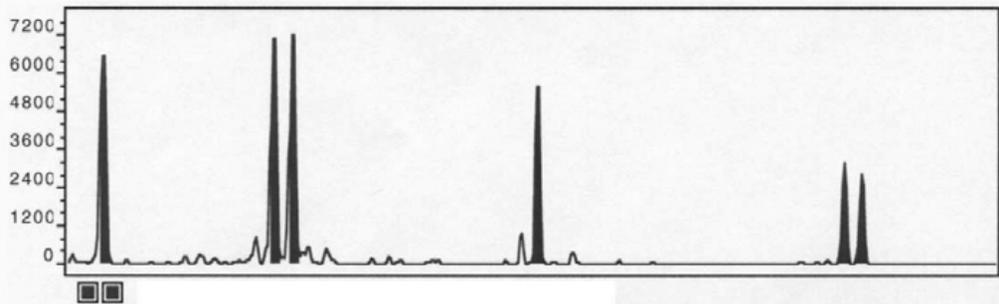
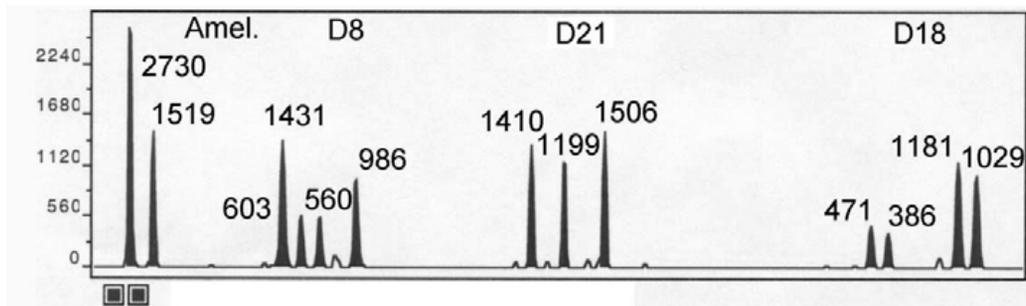


Figure 7. Continued.

TABLE 1
Mixture With Known Samples
[X] – Indicates Peak Height

| Exhibits Locus \ | Complainant | Crime Scene Mixed DNA Profile | Suspect |
|------------------|-------------|--|------------|
| D3S1358 | 15 16 | 15 [1989] 16 [739] 18 [1550] | 15 18 |
| VWA | 16 18 | 15 [1318] 16 [621] 18 [793] 19 [1200] | 15 19 |
| FGA | 21 23 | 21 [2414] 22 [1461] 23 [687] | 21 22 |
| Amelogenin | X | X [2730] Y [1519] | X Y |
| D8S1179 | 13 14 | 12 [1431] 13 [603] 14 [560] 16 [986] | 12 16 |
| D21S11 | 30 | 28 [1410] 30 [1199] 32.2 [1506] | 28 32.2 |
| D18S51 | 12 13 | 12 [471] 13 [386] 17 [1181] 18 [1029] | 17 18 |
| D5S818 | 12 13 | 12 [2561] 13 [463] | 12 |
| D13S317 | 11 | 11 [1607] 12 [834] | 11 12 |
| D7S820 | 10 11 | 8 [723] 10 [1203] 11 [289] | 8 10 |
| D16S539 | 12 14 | 11 [1262] 12 [515] 13 [1253] 14 [514] | 11 13 |
| TH01 | 8 | 5 [944] 6 [935] 8 [633] | 5 6 |
| TPOX | 8 11 | 8 [1257] 10 [984] 11 [447] | 8 10 |
| CSF1P0 | 10 | 10 [482] 11 [697] 12 [617] | 11 12 |

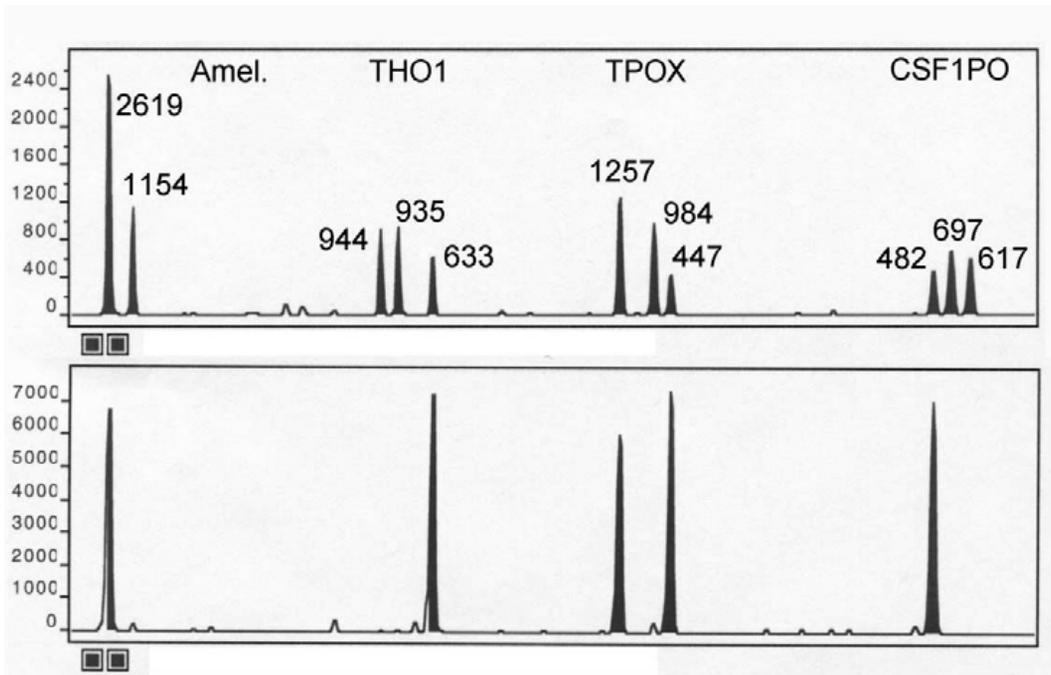
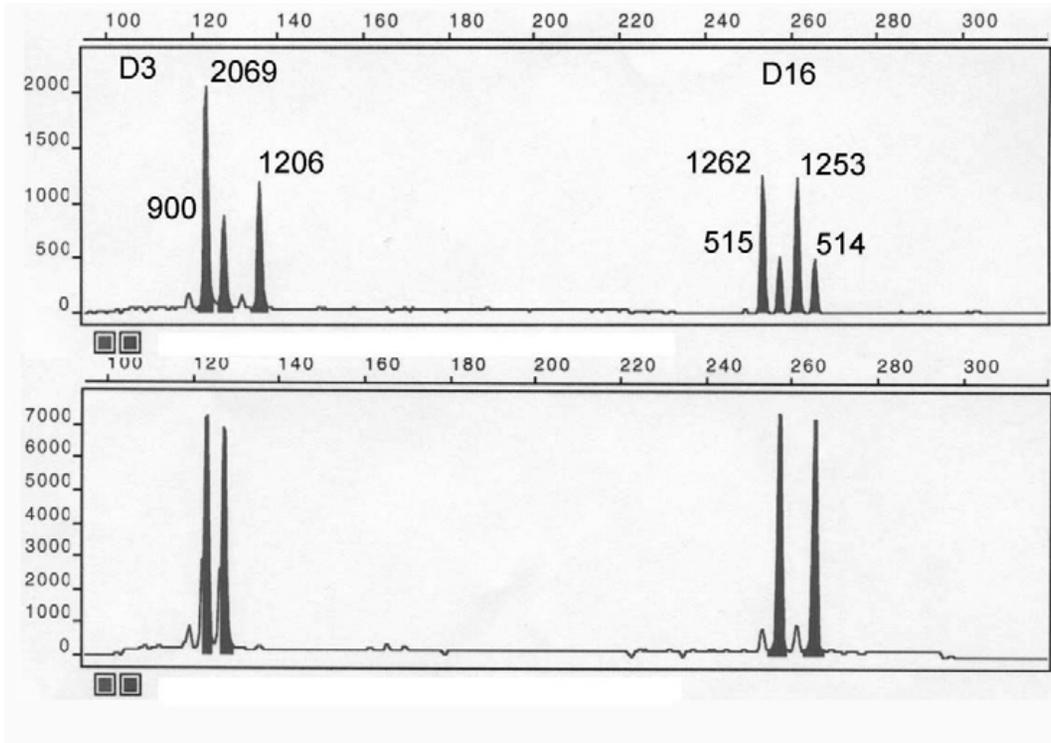


Figure 8. Mixture with single source sample in Cofiler.

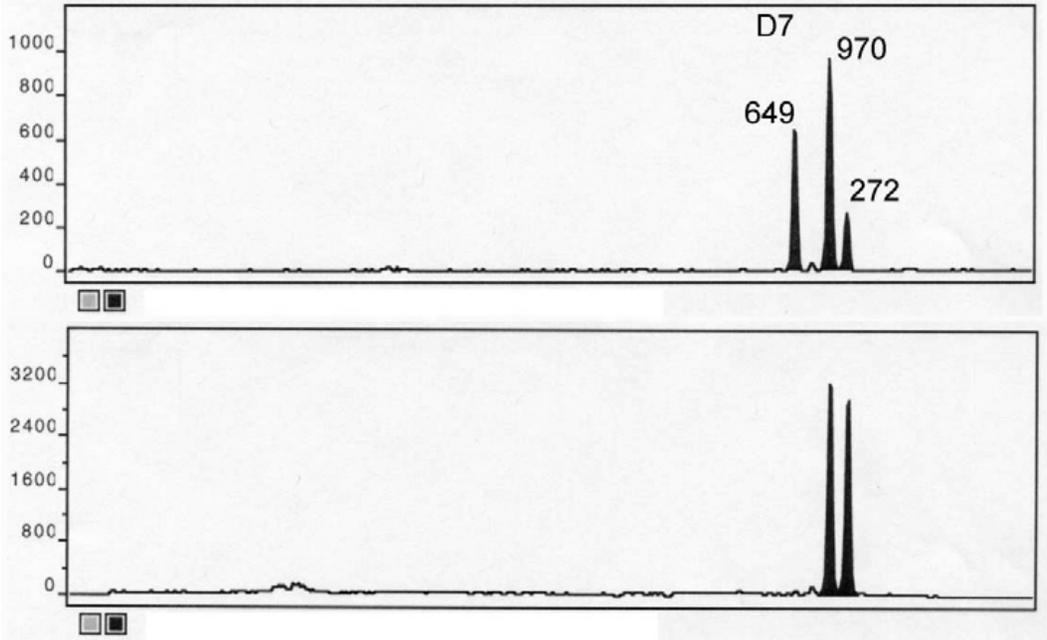


Figure 8. *Continued.*

TABLE 2
Genotypic Combinations of Potential Contributors to Mixed DNA Profiles

| Number of Alleles Present | Alleles Represented in the Mixed Profile | Number of Genotypic Combinations of Potential Contributors | Genotypic Combinations of Potential Contributors |
|---------------------------|--|--|---|
| 1 | A | 1 | AA |
| 2 | A,B | 3 | AA,AB,BB |
| 3 | A,B,C | 6 | AA,BB,CC,AB,AC,BC |
| 4 | A,B,C,D | 10 | AA,BB,CC,DD,AB,AC,AD,BC,BD,CD |
| 5 | A,B,C,D,E | 15 | AA,BB,CC,DD,EE,AB,AC,AD,AE,BC,BD,BE,CD,CE, and DE |
| 6 | A,B,C,D,E,F | 21 | AA,BB,CC,DD,EE,FF,AB,AC,AD,AE,AF,BC,BD,BE,BF,CD,CE,CF,DE,DF, and EF |

ACKNOWLEDGEMENTS

The author would like to gratefully acknowledge the assistance of Dr. Norma Szakacs of RCMP Forensic Laboratory Regina and Mr. George Schiro of Acadiana Criminalistics Laboratory in the review of the guidelines, and the work of Dr. Tse Wei Wang and Dr. George Carmody in furthering the area of forensic DNA interpretation.

REFERENCES

1. Jeffreys A.J., Wilson V., and Thein S.L. Individual-specific ‘fingerprints’ of human DNA. *Nature* 1985; 316: 76–79.

2. Wickenheiser R.A. Trace DNA: a review, discussion of theory, and application of the transfer of trace quantities of DNA through skin contact, *J Forensic Sci* 2002; 47(3): 442–450.
3. Budowle B., Chakraborty R., Carmody G., and Monson K.L. Source attribution of a forensic DNA profile, *Forensic Science Communications*, July 2000; 2(3).
4. Blackett R. DNA from dirty clothes, *Proceedings of the 13th Annual Symposium on Human Identification*, Phoenix, Arizona, October 7–10, 2002.
5. Clayton T.M., Whitaker J.P., Sparkes R., and Gill P. Analysis and interpretation of mixed forensic stains using DNA STR profiling, *Forensic Science International* 1998, 91(1): 55–70.
6. Fung W.K., Hu Y.Q. Interpreting DNA mixtures based on the NRC-II recommendation 4.1, *Forensic Science Communications*, October 2000; 2(4).
7. Perlin M.W., Szabady B. Linear mixture analysis: a mathematical approach to resolving mixed DNA samples. *J Forensic Sci* 2001; 46(6): 1372–78.
8. Wang T., Xue N., and Wickenheiser R. Least-square deconvolution (LSD): a new way of resolving STR/DNA mixture samples, *13th International Symposium on Human Identification*, (Online: <http://www.promega.com/geneticidproc/ussymp13proc/contents/wang.pdf>), Phoenix, Arizona, October 7–10, 2002.
9. Bill M., Gill P., Curran J., Clayton T., Pinchin R., Healy M., and Buckleton J. Pendulum – a guideline based approach to the interpretation of STR mixtures, *For Sci Int* 2005; 148(2–3); 181–9.
10. Baer, D.B. An ExcelTM Spreadsheet for Resolving STR Mixtures, *Southern Association of Forensic Science Newsletter*, March 2003; XXXI(1): 20–22.
11. Jankowski L.B., Petersack J.R., and LaRue E.J. A novel method for determining and annotating DNA STR mixtures using a set of mixture interpretation guidelines developed for complex cases, *Proceedings of the 13th Annual Symposium on Human Identification*, Phoenix, Arizona, October 7–10, 2002.
12. Scientific Working Group on DNA Analysis Methods (SWGDM), Short Tandem Repeat (STR) Interpretation Guidelines, *Forensic Science Communications*, 2000;2(3) <http://www.fbi.gov/hq/lab/fsc/backissu/july2000/strig.htm>
13. Duewer D.L., Kline M.C., Redman J.W., Newall P.J., and Reeder D.J. NIST Mixed Stain Studies #1 and #2: interlaboratory comparison of DNA quantification practice and short tandem repeat multiplex performance with multiple-source samples, *J For Sci* 2001; 46: 1199–1210.
14. Tomsey C.S., Kurtz M., Flowers B., Fumea J., Giles B., and Kucherer S. Case work guidelines and interpretation of short tandem repeat complex mixture analysis, *Croat Med J* 2001; 42(3): 276–280.
15. Kline M.C., Duewer D.L., Redman J.W., and Butler J.M. NIST mixed stain study 3: DNA quantitation accuracy and its influence on short tandem repeat multiplex signal intensity, *Anal Chem* 2003; 75: 2463–2469.
16. Duewer D.L., Kline M.C., Redman J.W., and Butler J.M. NIST Mixed Stain Study #3: signal intensity balance in commercial short tandem repeat multiplexes, *Anal Chem* 2004; 76: 6928–6934.
17. Newman J., Almer J., Frappier R., Greenfield A., Hageman C., and Laird J. Addressing the significance of a crime scene profile when the identity of the donor is not an issue, *Proceedings of the 13th Annual Symposium on Human Identification*, Phoenix, Arizona, October 7–10, 2002.
18. Wickenheiser R.A., MacMillan C.E. and Challoner C.M. Case of identification of severely burned human remains via paternity testing with PCR DNA typing, *Can Soc Forens Sci J* 1999; 32(1): 15–24.
19. Weir B.S., Triggs C.M., Starling L., Stowell L.I., Walsh K.A.J., and Buckleton J. Interpreting DNA Mixtures, *J For Sci* 1997; 42(2): 213–222.
20. Curran JM, Triggs CM, and Buckleton J. Weir B.S. Interpreting DNA mixtures in structured populations, *J For Sci* 1999; 44(5): 987–995.
21. Butler J.M. *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*, 2nd Ed, Elsevier Academic Press, USA, 2005; 154–167.
22. Buckleton J.S., Triggs C.M., and Walsh S.J. (eds) *Forensic DNA Evidence Interpretation*, CRC Press 2005: 1–552.

23. Gill P., Sparkes R., Pinchin R., Clayton T., Whitaker J., and Buckleton J. Interpreting simple STR mixtures using allele peak areas. *For Sci Int* 1998; 91(1): 41–53.
24. Rubinson K.A., Rubinson J.F. *Contemporary Instrumental Analysis*, Prentice Hall, Upper Saddle River, NJ, 2000; 150–158.
25. Gill P., Whitaker J., Flaxman C., Brown J., and Buckleton J. An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *For Sci Int* 2000; 112(1): 17–40.
26. Murray C., Lowe A., Richardson P., Wivell R., Gill P., Tully G., and Whitaker J., Use of low copy number (LCN) DNA in Forensic Inference, Proceedings of the 12th International Symposium on Human Identification, 2001.
27. Gill P. Application of Low Copy Number DNA Profiling, *Croat Med J* 2001; 42(3): 229–232.
28. Carmody, G. DNA Mixture Interpretation Workshop (Part I and II), presented at the 8th CODIS User's Conference, Arlington, Virginia, November 6 and 7, 2002.
29. The Evaluation of Forensic DNA Evidence (NRCII), National Research Council, Committee on DNA Forensic Science: An Update Commission on DNA Forensic Science, National Academy Press, Washington, D.C., 1996; 29–30, 102–106, 114–116.
30. Bodziak W.J. *Footwear impression evidence, detection, recovery and examination*, 2nd Edition, CRC Press, Boca Raton, USA, 2000: 372–374.
31. AmpF ℓ STR Profiler Plus PCR Amplification Kit User's Manual, Results and Interpretation 9-(24–34), Perkin Elmer Applied Biosystems, Foster City, California, 1997.
32. Berenson M.L., Levine D.M., and Krehbiel T.C. *Basic Business Statistics: Concepts and Applications*, 8th Ed. Prentice Hall, Upper Saddle River, NJ, USA, 2002: 123–124.
33. Krenke B.E., Tereba A., Anderson S.J., Buel E., Culhane S., Finis C.J., Tomsey C.S., Zachetti J.M., Masibay A., Rabbach D.R., Amriott E.A, and Sprecher C.J. Validation of a 16-locus fluorescent multiplex system, *J. Forensic Sci* 2002; 47(4): 773–785.
34. Butler J.M., McCord B.R. *Advanced Topics in STR Mixture Interpretation: DNA Mixture Interpretation*, Amer Acad of For Sci 2006 Workshop #6, Seattle, WA, February 20, 2006. http://www.cstl.nist.gov/biotech/strbase/ppt/AAFS2006_5_Mixture_Interpretation.pps#1

Appendix I Probability of Inclusion/Exclusion

The Probability of Inclusion Statistic will demonstrate how many individuals in a given population are potential contributors to a mixed DNA profile. Conversely, the Probability of Exclusion Statistic will demonstrate how many individuals in a given population can be eliminated as potential contributors to the mixed profile. The Probability of Exclusion is one minus the Probability of Inclusion, as indicated in the following equation:

$$P_{\text{excl}} = 1 - P_{\text{incl}}$$

The Probability of Inclusion Statistic is calculated by first determining the alleles present in the mixed profile for each loci. Next, every possible genotypic combination that could be a contributor to the mixed profile is determined for each individual locus. This is depicted in table 2.

The formula used for calculating the frequency of homozygotes is p^2 , and heterozygotes is $2pq$, as dictated by the Hardy-Weinberg equilibrium equation:

$$P = p^2 + 2pq + q^2$$

Allelic frequencies are obtained from a DNA database. Using the equations obtained above for every possible genotypic combination contributing to the profile at that locus, frequencies for each genotypic combination are calculated. These resulting frequencies are then added together to form a single statistic per locus. As indicated by NRCII, a Theta correction value (29) can be applied to account for any minor departure from ideal Hardy-Weinberg equilibrium. Using the product rule, all suitable loci are then multiplied together to obtain a combined statistic. The resulting statistic is the Probability of Inclusion. The Probability of Exclusion is calculated as indicated above.

Figure 6 demonstrates the calculation of the Probability of Inclusion through the use of the STRMix program (28). Below each allele is the individual allele frequency. These allele frequencies are combined through the variety of allelic combinations as shown in Table 2 to provide a single statistic per locus, as seen at the end of each row. These probabilities are then multiplied together to provide the combined Probability of Inclusion, as highlighted in the right center of Figure 6.

A short method to determine the Probability of Inclusion rather than using the expanded version from Table 2 is to simply add the individual frequencies for each allele and square that value. The resulting equation is as follows:

$$P_{\text{incl}} = (P^A + P^B + P^C + \dots)^2$$

Again, the combined Probability of Inclusion is then determined by multiplying the P_{incl} for each locus together.

It is strongly suggested that both Probability of Inclusion and Probability of Exclusion statistics are included in mixture conclusions to avoid bias, and to present findings in an understandable manner (see sample conclusions under Mixture Type 6).

Appendix II Flowchart for interpretation of mixed DNA STR profiles

1. Determine that the DNA profile is mixed (more than one contributor). Are there more than 2 peaks at a locus that are not explained by anomalies, stutter, known reference sample (mutation, chimera, mosaicism), and significant peak imbalance?

Yes = mixed profile, continue to 2.

No = explain extra peak(s), peak imbalance, exit and consider repeating sample.

2. As in a differential extraction, can more than one profile be used to deduce a suspect profile?

Yes = Mixture Type 1, go to 8.

No = continue to 3.

3. Can a forensically valid assumption be made to account for the contribution of a known profile presumed to be present in the mixed profile?

Yes = Mixture Type 2, go to 9.

No = continue to 4.

4. Determine the number of contributors. Are there more than 4 peaks in a locus that are not explained by anomalies, as in step 1?

Yes = greater than 2 person mixture, go to 5.

No = 2 person mixture, go to 6.

5. Is there a major contributor with greater than one minor contributor where the major contributor has a mass ratio of greater than 3:1 over the most significant minor contributor?

Yes = Mixture Type 5, go to 12. Report sole source for major. Consider Probability of Inclusion/Exclusion approach for minor contributor(s), see 13.

No = continue to 13.

6. Is there a distinct major contributor with a trace minor contributor with a mass ratio of 10:1 or greater?

Yes = Mixture Type 3, go to 10. Report sole source for major. Consider Probability of Inclusion/Exclusion approach for minor contributor, see 13, or do not interpret trace minor contributor if it is below threshold.

No = continue to 7.

7. Is there a major contributor with a minor contributor where the mass ratio is between 2.3:1 and 10:1?

Yes = Mixture Type 4, go to 11. Report a sole source for major. Consider Probability of Inclusion/Exclusion approach for minor contributor, see 13.

No = Mixture Type 6, go to 13.

8. Mixture Type 1, differential extraction permits that greater than one profile is used to deduce a suspect profile.

9. Mixture Type 2, forensically valid assumption permits accounting for the contribution of an endogenous donor to deduce a profile from a suspect donor.

10. Mixture Type 3, major contributor with trace minor contributor with a mass ratio of 10:1 or higher.
11. Mixture Type 4, major contributor with greater than 2.3:1 mass ratio over a single minor contributor (2 person mixture).
12. Mixture Type 5, major contributor with greater than 3:1 mass ratio over the most significant of a number of minor contributors (3 or more person mixture).
13. Mixture Type 6, the Probability of Inclusion/Exclusion approach.