40.1 Purpose:

The primary objective in this SOP is to guide analysts on the CT DSS policies for interpretation of Yfiler Plus results.

40.2 Responsibility:

DNA Section personnel.

40.3 Y-STR Analysis:

40.3.1 This procedure is a general guideline for the analysis and interpretation of Y-STR profiles when using the Yfiler Plus (YFP) PCR Amplification Kit. However, it is not an exhaustive list of all possible casework scenarios. Analyst training, experience, and judgment must be considered when reporting Y-STR profiles. Y-STR data are interpreted by evaluating the results at all loci. These procedures generally follow those outlined in the Applied Biosystems/Life Technologies GeneMapper ID-X Software version 1.5 Reference Guide.

40.3.2 When the YFP amplification kit is employed, 27 Y-STR loci; DYS576, DYS389I, DYS635, DYS389II, DYS627, DYS460, DYS458, DYS19, Y-GATA H4, DYS448 (6 base repeat), DYS391, DYS456, DYS390, DYS438 (5 base repeat), DYS392 (3 base repeat), DYS518, DYS570, DYS437, DYS385 a/b (2 loci amped with one set of primers), DYS449, DYS393, DYS439, DYS481, DYF387S1 (2 loci amped with one set of primers), and DYS533 are typed in a single PCR amplification reaction according to manufacturer's and laboratory protocols.

40.3.3 YFP amplification products are separated and detected using an ABI 3500xl Genetic Analyzer and the 3500 Series Data Collection Software 4 and analyzed with GeneMapper ID-X version 1.6 (ID-X) software.

40.3.4 For the ID-X software for YFP, the Analysis Method (YFP_CTDSS) is as outlined on pages 53-56 of Appendix A, Experimental Procedures for the CT-DSS YFP Internal Validation performed by ThermoFisher except for the following: Under the 'Peak Detector Tab', the 'Analysis Range' is 'Partial Range'. The 'Start' and 'Stop Pts' will be set by the analyst for each set of data but shall include the full range of required size standard peaks (60-460). The 'Sizing Range' is 'Partial Sizes' with the 'Start Size' set to 60 and the 'Stop Size' set to 460. The 'Peak Detection Amplitude Thresholds' will be set to B: 35, G: 55, Y: 35, R: 55, P: 55, and O: 150. For the 'Panel', use

'Yfiler_Plus_Panel_v4.1X-Allele-Specific'. For the 'Size Standard', use 'GS600_LIZ_(60-460').

40.4 Analysis of Y-STR DNA Profiles using GMID:

40.4.1 The results are evaluated for each locus. However, considering the data from all loci tested will assist in the overall interpretation of the results for each sample. See DNA SOP-6 for standard report conclusions.

40.4.2 Off-Scale Data: If too much DNA is added to the amplification reaction, the fluorescent intensity for the PCR products may exceed the linear dynamic range for detection by the instrument. Samples with off-scale peaks may exhibit raised baselines, split peaks, increased minus A peaks, excessive pull-up and/or higher than expected stutter levels, and/or amplification artifacts in one or more of the colors. Examine the peak heights in ID-X software. Any peaks greater than 10,000 RFU may be off-scale. These samples may need to be reinjected for a shorter time or diluted and re-amplified.

40.4.3 Peak identification: The laboratory's analytical threshold (AT) is dye channel specific: B: 35 RFU, G: 55 RFU, Y: 35 RFU, R: 55 RFU, P: 55 RFU, and O: 150 RFU. The identity of peaks called by ID-X are generally assigned to one of the following categories:

40.4.4 Allele Peak: A called allele has a peak height ≥ AT RFU, a fragment size that falls within the base pair range and has the appropriate dye color for the locus. However, not all peaks ≥ AT RFU are typed as alleles. Stutter peaks, pull-up/pull-down peaks, background noise, and amplification or injection artifacts could fall above the RFU threshold, especially if the allele peaks are high. Peaks ≥ AT RFU that are not called by ID-X may be called by the analysts on a case-by-case basis. Such peaks are interpreted with caution and the Technical Reviewer must agree with the insertion.

40.4.5 Detection of variant alleles: Variant alleles have been identified from many STR loci. These alleles may contain full or partial repeats. Variant alleles are labeled as Off Ladder (OL) alleles in ID-X. If a particular variant has been observed before by the DESPP CT Division of Scientific Services (CT-DSS), ThermoFisher (Life Technologies), NIST, or in published literature, the sample does not need to be repeated. In ID-X, change the OL to the proper allele call. However, if the repeat has not been observed, the sample may be re-amplified and/or reinjected to confirm the OL allele as warranted.

40.4.6 Stutter peaks (N-2, N-3, N-4, N-5, N-6, N-8, N-9, N-10, N-12, N+2, N+3, N+4, N+5, N+6—depending on the locus): A stutter peak has a fragment length typically four base pairs smaller than the true allele (see exceptions above). If the peak height in the stutter position is less than the user-defined stutter threshold, the peak is considered a stutter peak and is filtered out by GeneMapper ID-X or DNA QR-345. Allele-specific stutter thresholds (see DNA SOP 41.1.2) were determined by CT-DSS during the YFP Internal Validation. Peaks in stutter positions that are below the Laboratory's stutter thresholds are not listed on the electropherograms and are not used for statistical purposes. If the height of the peak in the stutter position is greater than the CT-DSS stutter thresholds, this indicates a possible mixture, an overloaded sample, or a higher than normal stutter ratio. Additional information provided by other loci may be of assistance in these cases.

40.4.7 Minus A Peaks: The YFP amplification kit is optimized to add an extra nucleotide to the 3' end of the PCR product. The nucleotide is predominantly adenosine (A). Lack of a full A addition may be observed, especially when the amount of input DNA is greater than approximately 1 ng. However, sample degradation can increase the effective concentration of a small amplicon relative to the larger amplicons, leading to minus A product at the smaller loci. Data from other loci may be of assistance in these cases. When appropriate, a smaller quantity of sample may be re-amplified.

40.4.8 Pull-up Peaks: Pull-up peaks are generally seen when excess amplification product is injected on the 3500xl or when the spectral calibration is sub-optimal. The greater fluorescent intensity causes uncompensated spectral overlap. Pull-up peaks are often displayed in several colors. Pull-up peaks should have approximately the same fragment size in each of the colors as the true allele peak. Additional information provided by other loci may be of assistance in these cases. When evaluating allelic peaks that may contain a pull-up (PU) contribution, an estimate of the PU contribution is subtracted from the height of the potential allelic peak to determine if the remaining peak height is sufficient for allele designation. The estimated pull-up contribution is determined by evaluating the height of unambiguous pull-up peaks (e.g. pull-up with no stutter or allelic contribution) at each locus where pull-up is detected for a particular color.

Note: the 3500 Series Data Collection Software 4 has a pull-up correction function applied to the raw data. This function is not necessarily consistent across the entire base-pair range for the detection of peaks from amplified DNA, therefore care must be taken when identifying the percentage of signal for unambiguous pull-up peaks. Using unambiguous pull-up peak(s) in the same

locus will most often be optimal. In the case where there are no unambiguous pull-up peaks for the dye channel, consider that the median percentage of pull-up peaks observed during the validation was 0.7% with 95.1% of the pull-up peaks being < 3% and 99.4% being < 5%.

Example: The percentage of pull-up peaks as compared to the height of the peaks causing the pull-up is calculated. (PU percentage = PU peak height/height of the peak causing the PU). This percentage is converted to an RFU value (multiply percent by height of peak causing the pull-up) which is then subtracted from the potential allelic peak in question. The resulting "subtracted" peak is called where DNA Section allele calling criteria are met. The same methodology would be applied to potential allelic peaks with a pull-down component (see below).

40.4.9 Pull-down Peaks: Pull-down peaks are generally seen when excess amplification product is run on the CE or when the spectral calibration is sub-optimal for the run. The greater fluorescent intensity causes uncompensated spectral overlap. Pull-down peaks can be displayed in several colors. Pull-down peaks can have approximately the same fragment size, but often, the pull-down peak will be a slightly larger or smaller size as compared to the true allele. Additional information provided by other loci may be of assistance in these cases.

40.4.10 Spikes: fluorescent spikes can be seen within GMID data. Spikes are typically sharp peaks caused by air bubbles or polymer crystals in a capillary, which can be displayed in different colors, and have approximately the same "fragment size". Additional inspection of the raw data or re-injection may assist in cases where spikes are believed to be present. Upon subsequent injection, spikes should not reoccur at the same position.

40.4.11 Dye Artifacts/Dye Blobs: Broad peaks may be the result of dissociated dyes or fluorescent particles that are trapped within the polymer matrix and fluorescently excited by the laser. Additional inspection of other samples, especially negative controls, within the module may assist in interpretation in cases where dye blobs are believed to be present.

40.4.12 Amplification Artifact Peaks: Nonspecific amplification product peaks, generally with good peak morphology, may result when excessive template DNA is amplified. When appropriate, a smaller quantity of sample may be re-amplified.

40.4.13 High Background: Background fluorescence may reach an RFU level above the calling threshold. Generally, such high background peaks exhibit atypical peak morphology. When appropriate, re-injection may assist with interpretation.

40.4.14 Raised Baseline: Elevated baseline of one color over a broader (multiple bin) scan range may occur. Note: Pull-down from two adjacent peaks may cause raised baseline between the pull-down peaks.

40.4.15 Analysis of unknown and known samples amplified with Yfiler Plus

Artifacts, except for any stutter peak, are not labeled in GeneMapper. They are removed from the profile, and a comment is unnecessary. This protocol (section 40.4) has details definitions of each artifact type and its appearance. Please also refer to DNA WI-39 Genemapper Work Instructions for more information. For any stutter peak in evidentiary (forensic unknown) and known profiles, a comment or "label" is added, designating that peak as stutter. All other artifacts require no comment.

40.5 Highly Degraded/Low Template Samples:

Highly degraded/low template samples are interpreted with caution. Degraded samples may "appear" to exhibit high stutter and greater –A at the smaller amplicons with off-scale data. In addition, degraded/low template samples may display unbalanced heterozygous alleles and locus/allele dropout. Partial profiles may be interpreted as determined by the data present for each locus. Degradation may also result in a particular contributor to a profile being low template at the large amplicons but standard template at the small amplicons (see 40.10). Given the possibility of stochastic allele sampling, especially with low quantity and/or degraded samples, single peaks (at loci DYS385 and DYF387S1) less than the stochastic threshold are interpreted with caution (pertains to putative haplotype assignment at that locus). To maximize the number of loci with peaks above the threshold, a sample may be run at Max on the 3500xL. However, any off-scale data at the smaller amplicons are interpreted with caution.

40.6 Inhibited Samples:

40.6.1 Inhibited samples may exhibit allele, locus, or complete profile dropout, unbalanced alleles at loci DYS385 and DYF387S1, and interlocus peak imbalance. To possibly overcome an inhibitor, smaller quantities of sample may be re-amplified. Quantitation data may aid in determining whether re-amplification is warranted as well as the amount of sample to be re-amplified. Before the genotyping amplification is performed, an evaluation of the Quantifiler Trio IPC for each sample is useful. The Quantifiler Trio IPC result for a sample

being "undetected" or having an elevated C_T value may indicate that a sample has an inhibitor present. If this is the case, multiple quantities or dilutions of a sample may be amplified for the original amplification. If necessary, a Microcon purification can be performed to reduce the amount of inhibitors in a sample.

40.6.2 Input of excessive female DNA may result in inhibition of Y-STR amplification. When appropriate, smaller quantities of sample may be re-amplified. Quantitation data may aid in determining whether re-amplification is warranted as well as the amount of sample to be re-amplified.

40.7 X-homologous peaks:

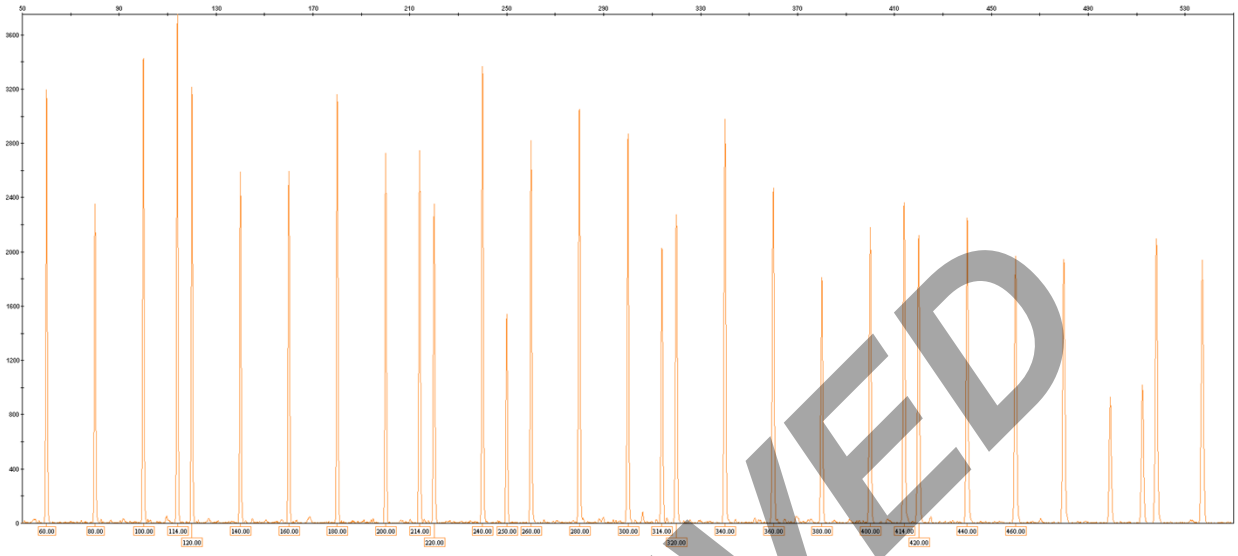
With large quantities of female DNA present in the amplified DNA extract, the Y primers may generate X-homologous or female cross-reactivity peaks. If the known female sample is amplified, any similar? X-homologous peaks may be removed from the evidentiary DNA profile. Note: This may require amplifying the same quantity of female DNA as with the evidentiary sample. The female cross-reactivity peaks are most common in the blue and red dye channels (see the YFP Developmental and CT-DSS Internal Validation results and the YFP User Guide for examples).

40.8 Technical Review:

The technical reviewer of a Y-STR project shall import the primary analyst's project utilizing the GeneMapper Manager. Any discrepancies in the analysis should be noted and retained, but the project shall not be altered by the technical reviewer. If applicable, the analyst will make any appropriate changes to their project, to then be checked by the technical reviewer.

40.9 General Analysis of Yfiler Plus Results:

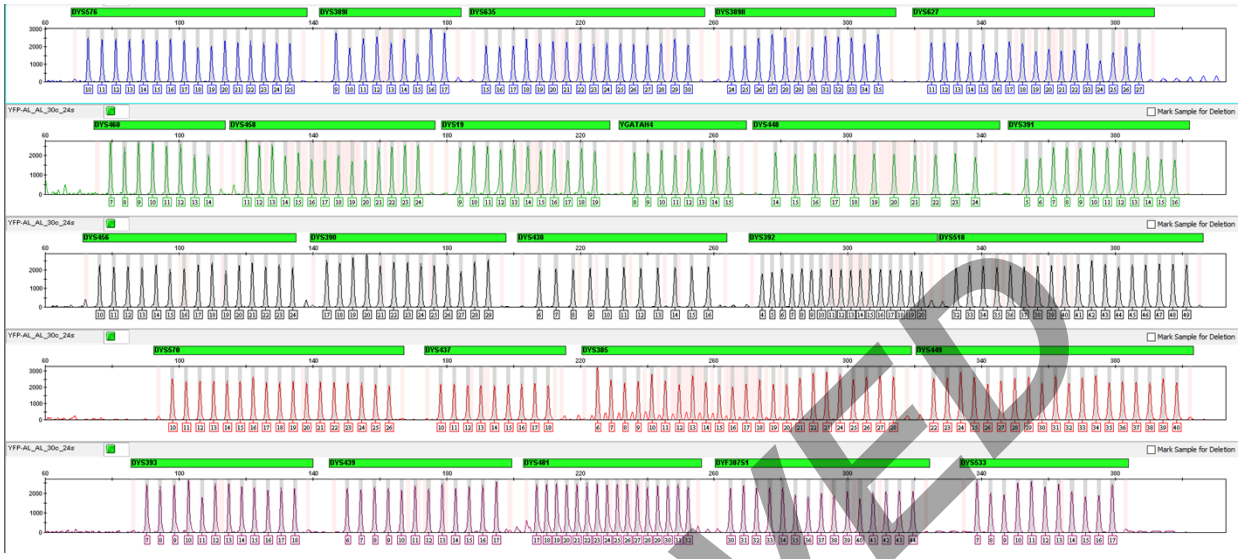
40.9.1 Ensure that all required internal lane size standard (60-460) peaks are called in all samples.



40.9.1.1 If a sample's internal size standard fails due to a low sizing quality (e.g., pull-up peaks into the size standard), the option to override the sample's size quality is available and shall be noted on the analysis worksheet. The analyst and the technical reviewer must confirm that the size quality failure was not due to any loss of internal lane size standard peaks. The sample may then be analyzed.

40.9.2 Ensure that each ladder in each injection has all allele peaks detected above the minimum threshold. If a ladder does not meet these requirements, change the sample type from 'Allelic Ladder' to 'Sample'. One ladder in each injection module must meet the requirements.

40.9.2.1 The Yfiler Plus Allelic Ladder shall appear as follows:



40.9.3 Y-STR Control Requirements:

40.9.3.1 Import the appropriate sample files including controls (RB, NEG, EP1, POS control DNA 007) and at least one ladder. At least one set of controls (for each amplification set) must give the expected results. If not, the amplification set must be re-injected, re-amplified, or re-extracted as warranted. When necessary, if one of the positive amplification controls does not give the expected result, the TL may sign off on an extraction positive control that did give expected results as the amplification positive. The RB, if amplified concurrently and on the same Thermal Cycler as the evidence samples, may be used as the negative amplification control. Reagent Blank stringency must be followed per QAS.

40.9.3.2 Potential contamination is assessed by evaluating the evidentiary samples, RB, NEG, and positive controls. All samples within a batch will be cross-checked against each other to detect potential contamination within the batch. Analysts should use their experience and judgment during this process. An excel macro is available to assist with batch cross-comparisons.

**RB and NEG samples should not have any peaks above the analytical threshold. Drop-in may occr, butif more than two callable peaks are detected, the negative samples, rework is warranted and they should be re-prepped and re-injected, re-amplified, or re-extracted as appropriate. Rework should be utilized

to determine the root of the contamination. These contamination events will be documented in a Qualtrax workflow and the TL will be notified.

Note: The RB may be omitted for re-amplifications using the same volume (or less) of template. The RB and NEG may be omitted for re-injections of equal or lesser time/sensitivity.

The appropriate response for a contamination event (including how the sample(s) are reported) is determined on a case-by-case basis with approval of the TL. Instances in which a single peak ≥ analytical threshold is observed in a negative control will not be tracked as a workflow instance and does not require TL signoff. More than one called peak in a negative control will be evaluated with the input of the Technical Leader.

Depending on factors such as the type and extent of contamination, and whether the source of the contamination can be identified, the contaminated sample(s) and/or amplification sets may be interpreted with caution or deemed unsuitable for comparison following TL review and approval (see Contamination statements located within the DNA Report Template).

40.9.3.3 Positive Control Profiles: Ensure that all alleles were included and that the positive controls give the expected profile. The EP1 may be omitted for re-amplifications if expected results were previously generated for the extraction set. The positive controls may be omitted for re-injections if expected results were previously generated.

40.9.4 Evidentiary profiles that contain elevated stutter peaks at a majority of the loci and/or have off-scale peaks and other artifacts visible through the electropherogram due to injecting too much and/or amplifying too much sample may be reinjected using a reduced injection time and/or using less amplified DNA. Samples may also be re-amplified using a reduced amount of template and re-run on the CE instrument. Please refer to DNA SOP-37 YFP Amplification.

40.9.5 If a peak is labeled as off-ladder (OL) or is outside the ladder region and therefore not labeled by the GeneMapper ID-X software, review the data to determine that it is a true microvariant or an off-ladder allele. Peaks of this nature may be confirmed through re-injection or re-amplification, if necessary.

40.9.6 The stochastic threshold (ST) is 500-700 RFU for loci DYS385a/b and DYF387S1 (see table below; ST derived from internal validation).

40.9.6.1 Per the "SWGDAM Guidelines on STR Interpretation": A stochastic threshold (ST) is defined as the "value above which it is reasonable to assume that allelic dropout has not occurred within a single-source sample". For YFP, the stochastic threshold is the peak height above which one can expect that both peaks at loci DYS385a/b and DYF387S1 would be observed. This threshold is applied to reduce the incidences of calling a false "homozygote."

Injections Time Ranges

Stochastic Threshold	3500-1, 2, 3, 5, & 6		3500-4
500 RFU	1.2 kV, 16 sec	LOW	1.2 kV, 12 sec
500 RFU	1.2 kV, 24 sec	STANDARD	1.2 kV, 18 sec
700 RFU	1.6 kV, 32 sec	HIGH	1.2 kV, 24 sec

40.9.6.2 Specific worksheets to be used for YFP amplification and injection on CE instrument are as follows:

DNA QR-324 YFP Amplification Worksheet

DNA QR-328 YFP Injection/Analysis Worksheet

40.10 Determination of Single Source or Mixture Profile & Number of Contributors

Single Source Profile Attributes: A single source profile will generally have no more than 1 allele at any locus except DYS385a/b and DYF387S1 (typically 2 alleles). Expected peak balance at DYS385a/b and DYF387S1 decreases as a function of peak height similar to that seen with autosomal loci. In rare instances, an individual may exhibit a bi/tri-allelic pattern at a locus.

When an extra-allelic pattern is identified, the results must be documented in the case file. The locus can be searched in the YHRD database.

40.10.1 Determination of the number of Y lineages detected in the DNA profile: Evidence profiles should be assessed for the number of possible male lineages considering data above and below the stochastic threshold as discussed below and documented in the case file:

The following steps are followed to assign the likely number of male lineages in the profile:

1. Review the profile as a whole, assessing the level of degradation, presence of low-level peaks, noisy or clean baseline, and general quality (template) of the profile.
2. Identify likely stutter peaks by referencing the YFP allele-specific stutter ratio (SR) expectations for each allele at each locus using DNA QR-345 (also see plots of SR as needed in validation).
3. DNA QR-345 is used as a tool to calculate allele-specific and low-template stutter values not included in the GeneMapper ID-X panel YFP-CTDSS. The panel used in Genemapper is specifically named "Yfiler _Plus_v4.1X-Allele-Specific. Low-template stutter thresholds are applied to each locus as appropriate based on the estimated quantity of DNA for each contributor to the profile [i.e., (a) when the total quantity amplified is ≤ 100 pg—apply to all contributors or (b) when the total quantity amplified is > 100 pg but the amount of DNA from some contributor(s) is ≤ 100 pg—apply low-template thresholds only to the peaks from the low-template contributors.] Peaks in stutter positions that are below the Laboratory's low-template stutter thresholds (see (a) & (b) above), which are not removed by ID-X are filtered using DNA QR-345 and are not listed on the electropherograms.
4. Find the locus with the highest number of unambiguous allelic peaks, "A", keeping in mind that DYS385 and DYF387S1 routinely have two results. "A" gives the initial assessment of the likely number of contributors to the profile.
5. Review peak height balance at loci DYS385 and DYF387S1. Refer to 'Appendix 1' for an indication of expected heterozygote peak balance for varying template amounts at these loci. Note that peak variance (imbalance) increases as the amount of template in the PCR reaction decreases following a continuous model. With low-template amplifications (≤100 pg), an imbalance greater than the mean minus 1 standard deviation can support an additional contributor to the profile. However, significant peak imbalance at a locus (e.g., imbalance greater than mean minus 1 SD) is not always a strong predictor itself of an additional contributor. Imbalance should be evaluated in the context of the overall profile results to determine the most likely number of contributors to the DNA profile. When peak imbalance is used to infer an additional contributor, the potential impact of additive stutter from adjacent allele(s) must be evaluated.

The detection of more than one allele per locus (other than the DYS385 and DYF387S1 loci) is generally indicative of a mixed sample. Variation in peak

heights between alleles between loci and at DYS385 and DYF387S1 may assist in the interpretation of such results.

40.10.2 Typically, if a maximum of 2 alleles are observed at a locus (except for DYS385 and DYF387S1), the DNA results are consistent with a mixture of two male lineages. Two peak patterns (at a single locus other than DYS385 and DYF387S1) have been observed from single source samples, but they are rare. The analyst should consider the possibility of high stutter peaks, which can complicate interpretation. While counting allele peaks is critical in determining the number of contributors, the analyst must also consider that allele sharing between different haplotype lineages may result in an underestimation of the actual number of contributors. The number of male lineages identified reflects the analyst's assessment of the most likely number of different haplotypes required to reasonably explain the observed profile.

40.10.3 A Y-STR mixture simulation tool has been validated by the Division of Scientific Services to assist in the determination of the number of haplotypes that are present in Yfiler Plus profiles. This tool may be used by analysts to demonstrate the expected amount of allele sharing at each locus per contributor number in the absence of dropout.

Note 1: Amplification results from low-template quantities (particularly less than 100 pg) may exhibit increased stutter and peak variance. Based on the YFP/3500xl validation data, it was noted that profiles from low-template contributors may exhibit multiple high stutter peaks that can be considerably above the standard stutter thresholds (see low-template stutter thresholds). A similar impact may occur with respect to peak imbalance at DYS385 and DYF387S1. Therefore, when these low-template peaks pertaining to stutter or peak imbalance are used to infer an extra lineage, look for data elsewhere in the profile to support this assessment, if possible.

Note 2: The low-template stutter thresholds in DNA QR-345 are included to aid in determining the number of contributors and to highlight the increased expected stutter variance associated with low-template single male lineage profiles and/or profiles with low-template contributors.

40.11 **Qualitative Assessment of YFP Profiles—Comparisons to Knowns:**

40.11.1 Single source haplotypes: If 1 peak was called ≥ AT (except at DYS385 and DYF387S1) with nothing else detected ≥ AT at any other locus, and "Known A"

has that allele, the results would be consistent with "Known A" (or another member of the same male lineage) consistent with being the source of the DNA profile detected. Considering the possibility of "heterozygous" allelic peak imbalance at DYS385 and DYF387S1, an inconclusive conclusion based on limited data detected may be warranted. Significant peak imbalance at DYS385 and/or DYF387S1, may be indicative of a mixture. Note: All YFP results are utilized for comparison purposes. However, less loci are often used for statistics (when the comparison is "inclusion/CBE") as described in 40.13.

40.11.1.1 For single source partial profiles (or mixture partial profiles) with peaks ≥ AT, report "Known A (or another member of the same male lineage) cannot be eliminated (CBE) as the source/or contributor from the DNA profile from item #X as long as the observed alleles are consistent with the known source.

For fully detected single source haplotype DNA profiles, report "Known A (or another member of the same male lineage) is consistent with being the source of the DNA profile from item #X. See DNA SOP-6 for report statements.

Note: in situations where no loci have more than 1 peak ≥ AT (except 2 peaks at DYS385 or DYF387S1) [e.g., such as if there is a significant peak imbalance], it is important to consider if the profile could be a mixture (partial composite of 2 or more individuals).

40.11.2 **Interpretation of DNA mixtures**

40.11.2.1 The results at all YFP loci are utilized for comparison purposes. These general approaches do not account for all possible mixture scenarios. With TL approval, other approaches may be necessary on a case-by-case basis.

The analyst considers the quantity of DNA amplified, the peak heights of the alleles, the extent of degradation/inhibition, the impact of stutter, potential carryover between fractions in a differential extraction, and if a known profile compared to the mixture can be objectively determined to be a contributor (e.g., intimate samples).

40.11.2.2 Factors taken into consideration when interpreting Y-STR mixtures:

40.11.2.2.1 If one person can be objectively determined to be included in the mixture (e.g., consensual partner), the mixture is evaluated by a consideration of what loci have DNA profile results (≥ AT) that could not be from the known male source.

40.11.2.2.2 The total number of dropouts (overall and at each locus) that would be required to still be consistent with a CBE conclusion. While allele dropout (including peaks below the stochastic threshold at DYS385 and DYF387S1) can occur at any locus, extensive validation and case experience have shown that smaller amplicons amplify more efficiently and are less affected by degradation than larger amplicons. Therefore, the results at the smaller amplicons can be particularly diagnostic.

40.11.2.2.3 To what degree would a putative low-level contributor coincidently be included by a high-level contributor due to overlap.

40.11.2.2.4 If the analyst and reviewers feel that the above rules do not adequately apply to a specific mixture comparison, consult the TL.

40.11.3 YFP profiles will be searched against the Staff Index and documented on DNA QR-20 Staff Search Worksheet.

40.12 General Report Conclusions for Y-STR Results:

40.12.1 YFP report templates are described in DNA SOP-6. Additionally, see below examples of general statements for reports.

40.12.2 Elimination: used when the known sample is not the source of, or a contributor to, the DNA profile detected from item xx. The known sample has alleles that are not detected in the evidentiary sample and the lack of detection is not reasonably explainable by degradation/inhibition/allele dropout (stochastic effects) within the evidentiary sample.

40.12.3 Inconclusive: No conclusion can be drawn from the comparison between the known sample and the evidentiary sample due to uncertainty. There are two reasons for inconclusive results: limited data detected and the mixture being too complex. See 40.12.3.1 and 40.12.3.2 below:

40.12.3.1 Based on limited data: This statement is made when the data does not provide sufficient support for any conclusion regarding comparing a mixture to a particular known.

40.12.3.2 Mixture too complex: This statement is made for mixtures of 4 or more male lineages where the known is not eliminated.

40.12.4 Cannot be eliminated (CBE) as the source of, or a contributor to a DNA

profile: This conclusion is used when there is a strong correlation between the known sample and the evidentiary sample, but some of the alleles present in the known sample are not detected in the evidentiary sample above the analytical threshold (for single source profiles, this only applies to the DYS385 and DYF387S1). In this instance, dropout is a reasonable explanation taking into consideration peak height data, overall quantity of template DNA, information regarding the extent of degradation, inhibition, stochastic effects, and masking by stutter or contributor ratio.

40.12.5 Inclusion: All alleles consistent with the known sample are detected (≥ AT) in the evidentiary profile at all loci where results are generated.

40.13 YFP Statistics: (see 40.14 YFP Workflow and Qualification of Loci for Statistics)

40.13.1 For single source YFP profiles, statistics are calculated, for most circumstances, at all loci where DNA results are obtained ≥ the analytical threshold. For DYS385, if both alleles are detected, the locus is also used for statistical purposes. If only one allele is detected at DYS385, that locus can be used for statistical purposes, given that the allele is above the stochastic threshold. The YFP random match probability is calculated using the counting method. A haplotype frequency is generated by counting the number of times a haplotype has been observed in the population database divided by the total number of haplotypes in the Database. Haplotype frequencies are given an upper bound 95% confidence interval (See DNA SOP-6). The Y-Chromosome STR Haplotype Reference Database (YHRD) Y-STR population database (available at https://yhrd.org) is used.

40.13.1.1 On the YHRD website homepage, click "Estimate Frequency" at the top of the site. From there, click on "Manually enter haplotype(s)". Select the Yfiler Plus kit. Enter the YFP profile at all loci, with a comma separating the two alleles (even if a homozygote) at DYS385 and DYF387S1. (Note only enter DYS385 and DYF387S1 if you will be using them for statistics.) After entering the YFP profile, select the "Minimal" Dataset and click on the "Search" button to search the profile. When search is complete, click "Add feature to this report" and select "National Database (with Subpopulations, current SWGDAM-compliant version). Click the "X" to close the "worldwide" results before printing page 1 for the case jacket. Click the back arrow and repeat the process for Y12, Y17, and Y27 Datasets.
Report the statistic for the population with the most common haplotype frequency using the most discrimination dataset as described in 40.13.1.2. Generally, there will be a total of 4 dataset search printouts.

1. From the most common United States data sets, (African American, Caucasian and Hispanic), pick the most conservative haplotype frequency that is represented on the "Minimal", "Y12", "Y17", and "Y27" datasets.
2. Then choose the most discriminating frequency (out of the four picked in step 1) to report.

Ensure the case number, item number, initials, and date are on all four documents.

Note: A dataset would not be searched if the profile does not have results for any loci associated with that dataset. A dataset would also not be searched when the entire profile results being entered in YHRD include only loci present in a smaller (fewer loci) dataset (i.e., no results for the loci only found in the larger dataset).

40.13.1.2 An exact upper bound 95% confidence interval (Clopper-Pearson method) is calculated from the results of a database search to estimate the population frequency of a haplotype using the population (United States data set: African American, Caucasian, or Hispanic) with the most common haplotype frequency. The upper bound estimate is dependent on the size of the database, and these estimates may change as the database size changes.

40.13.1.3 The formula used to calculate the 95% upper bound frequency estimate is:

$$\sum_{k=0}^{x}\binom{n}{k}p_0^k(1-p_0)^{n-k}=0.05$$

https://strbase.nist.gov/pub_pres/SWGDAM-Jan2012-YSTRs.pdf

Clopper, C. and Pearson, E. *The use of confidence or fiducial limits illustrated in the case of binomial,* Biometrika (1934) 26:404-413.

40.13.2 Y-STR haplotypes are inherited independently from autosomal loci. Therefore, Y haplotype (and cumulative haplotype) frequencies may be combined with autosomal STR RMPs or CPIs using the product rule to generate an overall profile frequency estimate (Applies to outsourced cases only with no STRmix analysis).

40.13.3 Y-STR Mixtures: (see 40.14 YFP Workflow and Qualification of Loci for Statistics)

40.13.3.1 The expected frequency of haplotypes that could contribute to a YFP mixture is calculated using the Combined Counting Method (CCM). A cumulative haplotype frequency is generated by counting how many times each of the possible haplotypes are observed in the population database, summing them, and dividing by the total number of haplotypes in the database (i.e. cumulative haplotype frequency = x/N, where x is the number of potentially contributing haplotypes observed in a population and N is the total number of haplotypes searched in that population). An exact upper bound 95% confidence interval is then applied using the population (African American, Caucasian, or Hispanic) with the most common CCM frequency. Note: The frequency and confidence interval for a mixture profile are calculated using the software tool "Y-Mix", which contains the U.S. portion of the YHRD database.

The cumulative haplotype frequency (with an upper bound 95% confidence interval) is calculated for samples with an additional allele at any locus, i.e., 2 alleles at DYS393, DYS19, DYS389II, DYS390, DYS391, or 3 alleles at DYS385.

40.13.3.2 The file for the "Y-Mix" software tool is located on the S: Drive. The loci in the excel file are listed in alphanumeric order. Enter the mixture profile under the appropriate YF/YFP loci, with one row per allele. (Note: for each locus with data entered, add a "0" along with the alleles detected if a null has been seen in the database at that locus). Defaults (Limit database to samples with all the loci entered above? Yes, Treat this profile as a single source sample?: No, Desired UCI: 95%, Use (x+1)/(N+1)?: No) are to be unchanged. Click "compare the profile to the database" when profile entry is complete. Print page when results are obtained.

40.14 **YFP Workflow and Qualification of Loci for Statistics:**

Note: If it so happens that Y-STR loci will be used for statistics, the same guidelines for Y-STR loci will follow (i.e. DYF387S1 will follow the guidelines set forth for DYS385).

40.14.1 Evaluate data at all YFP loci (≥ AT) to determine if profile is a single haplotype or a mixture (see 40.10).

40.14.2 Qualify results for statistics at DYS385 using a stochastic threshold (ST) of ≥ 500 or 700 RFU (based on injection time) and at other loci as discussed below. Statistics are calculated as described in 40.13 for 1-3 male lineage profiles.

40.14.3 Compare known(s) to evidentiary profiles as in 40.11 and 40.12.

40.14.4 Single Haplotype Profile:

40.14.4.1 2 alleles detected at DYS385: use data (above & below ST) at DYS385 and other Yfiler loci where included.

40.14.4.2 1 allele detected at DYS385: (A) Qualify locus if allele is above ST and use data for statistics at other loci where included. (B) Disqualify locus for statistics if allele is below ST and use data at remaining YFP loci where included. Note: the known would be eliminated as the source if the known does not have the allele found in the evidence profile.

40.14.5 Mixture of 2 Haplotypes (CCM): Evaluate results at all loci as described in 40.14.1-3.

40.14.5.1 Mixture evident at DYS385 with < 4 alleles and all alleles > ST at DYS385: qualify locus for stats and use data for statistics at other YFP loci where the mixture is detected.

40.14.5.2 Mixture evident at DYS385 with < 4 alleles and some/all allele(s) below ST at DYS385: Disqualify locus for stats and use data at other loci where the mixture is detected. Note: the known (2 alleles at DYS385) would be eliminated as a contributor if the known does not have at least 1 allele found in the evidence profile. If the known has only l allele at DYS385, the known may not be eliminated depending on the results at other loci. Note that DYF387S1 is utilized in the same manner as DYS385.

40.14.5.3 Mixture with 4 alleles at DYS385 and allele(s) above or below ST at DYS385: qualify locus for statistics and use data for statistics at other YFP loci where the mixture is detected.

40.14.5.4 Mixture not evident at DYS385, e.g., 2 alleles detected at DYS385 and no significant peak imbalance at DYS385 (where the minor allele is > 20% of the major allele): disqualify locus for statistics. Use data for statistics at other YFP loci where the mixture is detected.

40.14.5.5 Statistics are calculated for loci that detect the mixture (e.g., 2 alleles at one locus).

40.14.6 Mixture of 3 Haplotypes (CCM): Evaluate results at all YFP loci as described in 40.14.1-3.

40.14.6.1 Mixture with 2-4 alleles at DYS385 and all alleles > ST at DYS385: disqualify locus for stats and use data for statistics at other YFP loci where all 3 contributors are detected.

40.14.6.2 Mixture with 2-4 alleles at DYS385 and some/all allele(s) below ST at DYS385: disqualify locus for stats and use data at other YFP loci where all 3 contributors are detected.

40.14.6.3 Mixture with 5 alleles at DYS385 and all alleles > ST at DYS385: qualify locus for stats and use data for statistics at other YFP loci where all 3 contributors are detected.

40.14.6.4 Mixture with 5 alleles at DYS385 and some/all alleles below ST at DYS385: disqualify locus for stats and use data for statistics at other YFP loci where all 3 contributors are detected.

40.14.6.5 Mixture with 6 alleles at DYS385 and allele(s) above or below ST at DYS385: qualify locus for statistics and use data for statistics at other YFP loci where all 3 contributors are detected.

40.14.6.6 Mixture not evident at DYS385: disqualify locus for statistics. Use data for statistics at other YFP loci where all 3 contributors are detected.

40.14.6.7 Statistics are calculated for loci that detect the mixture (e.g., 3 alleles at one locus).

40.15 Deconvolution of Major in 2 Male Lineage Mixtures

Note: All loci able to be deconvoluted will be deconvoluted as the major profile. The deconvolution of a major is done independently to determining which loci will be used for statistical purposes.

Note: A macro has been created to assist in the determination of YFP deconvolutions as well as Y-Mix.

40.15.1 A major contributor to mixtures of 2 male lineages may be deconvoluted in lieu of the Combined Counting method (CCM) discussed in 40.14.5.

40.15.2 Deconvolution of mixtures at all YFP loci except DYS385 and DYF387S1: A major contributor may be deconvoluted where the height of the major allele is at least 5X greater than the minor allele (minor is ≤ 20% of major). Example: if the minor allele is 100 RFU, the major allele must be ≥ 500 RFU.

40.15.3 Where only 1 allele is detected at any locus except DYS385 and DYF387S1, the "major contributor allele" must be ≥ 500 or 700 RFU, depending on the injection time (see 40.9.6.1). If not, the locus is not eligible for deconvolution.

40.15.4 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 1 allele is present, the major contributor may be deduced if the peak height is ≥ the stochastic threshold.

40.15.5 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 2 alleles are present, the "heterozygous" major contributor may be deduced if the peaks are ≥ the stochastic threshold and "heterozygous" peak balance is at least 50%.

40.15.6 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 2 alleles are present, ("homozygous" major with significant peak imbalance), the major contributor may be deduced if the major peak is ≥ the stochastic threshold and the mixture ratio at DYS385 or DYF387S1 is at least 5:1.

40.15.7 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 3 alleles are present, ("heterozygous" major): the major contributor may be deduced if both peaks are ≥ the stochastic threshold, the mixture ratio of the lowest major allele to the minor allele is at least 5:1, and "heterozygous" peak balance is at least 50%.

40.15.8 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 3 alleles are present, ("homozygous" major): the major contributor may be deduced if the major allele is ≥ the stochastic threshold and the mixture ratio of the major allele to the highest minor allele is at least 5:1.

40.15.9 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 4 alleles are present; the major contributor may be deduced if the two highest peaks are ≥ the stochastic threshold, the mixture ratio is at least 5:1 (lowest major allele must be at least 5X greater than the highest minor allele), and "heterozygous" peak balance is at least 50%.

40.15.10 Once deconvolution is completed, the known(s) are compared to the profile of the deconvoluted major contributor (see 40.11 & 40.12).

40.15.11 If an individual is eliminated as a potential source of the major contributor but cannot be eliminated from the overall mixture (i.e., a potential minor contributor), the CCM statistic is calculated as discussed in 40.14.5.

40.16 Deconvolution of Major in 3 Male Lineage Mixtures

Note: All loci able to be deconvoluted will be deconvoluted as the major profile. The deconvolution of a major is done independently to determining which loci will be used for statistical purposes.

Note: A macro has been created to assist in the determination of YFP deconvolutions as well as Y-Mix.

40.16.1 A major contributor to mixtures of 3 male lineages may be deconvoluted in lieu of the combined counting method (CCM) discussed in 40.14.5 as follows;

40.16.2 Deconvolution of mixtures at all loci except DYS385 and DYF387S1: A major contributor may be deconvoluted where the height of the major allele is at least 5X greater than the highest minor allele (minor is ≤ 20% of major).

40.16.3 Where only 1 allele is detected at any locus except DYS385 or DYF387S1, the "major contributor allele" must be ≥ 500 or 700 RFU depending on the injection time (see 40.9.6.1). If not, the locus is not eligible for deconvolution.

40.16.4 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 1 allele is present, the major contributor may be deduced if the peak height is ≥ the stochastic threshold and the mixture ratio at all loci where the mixture is detected is at least 5:1.

40.16.5 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 2 alleles are present, the "heterozygous" major contributor may be deduced if the peaks are ≥ the stochastic threshold, the mixture ratio at all loci where the mixture is detected is at least 5:1, and "heterozygous" peak balance is at least 50%.

40.16.6 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 2 alleles are present ("homozygous" major with significant peak imbalance), the major contributor may be deduced if the major peak is ≥ the stochastic threshold and the mixture ratio at DYS385 and DYF387S1 is at least 5:1.

40.16.7 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 3-6 alleles are present ("heterozygous" major): the major contributor may be deduced if both peaks are ≥ the stochastic threshold, the mixture ratio of the lowest major allele to the highest minor allele is at least 5:1, and "heterozygous" peak balance is at least 50%.

40.16.8 To qualify the results at DYS385 and DYF387S1 for deconvolution: where 3-5 alleles are present ("homozygous" major): the major contributor may be deduced if the major allele is ≥ the stochastic threshold and the mixture ratio of the major allele to the highest minor allele is at least 5:1.

40.16.9 Once deconvolution is completed, the known(s) are compared to the profile of the deconvoluted major contributor (see 40.11 & 40.12).

40.16.10 If an individual is eliminated as a potential source of the major contributor but cannot be eliminated from the overall mixture (i.e., a potential minor contributor), the CCM statistic is calculated as discussed in 40.14.5 and 40.14.6.

Appendix 1: Expected Heterozygous Peak Balance at DYS385 & DYF387S1

Data set 1:

16-125 pg –Standard injection:

AVE	0.671	MIN	0.23
SD	0.2232	MAX	0.99
COUNT	30		

Max Injection:

16-63 pg:

AVE	0.6650	MIN	0.2261
SD	0.2217	MAX	0.9889
COUNT	21		

Data set 2:

16-63 pg –Standard injection:

AVG	0.608107
STD	0.203004
MIN	0.3027
MAX	0.9272
COUNT	28

Max Injection:

4-63 pg:

AVG	0.601754
STD	0.217629
MIN	0.2178
MAX	0.9639
COUNT	35