5.1 Purpose:

The primary objective reflected in this section is for the analyst to render the best interpretation of the data.

5.2 Responsibility:

DNA Section personnel.

5.3 Interpretational Paradigms:

5.3.1 The general goal of DNA testing is to eliminate the suspect from the evidentiary DNA profile where scientifically appropriate.

5.3.2 The absence of detection—barring a sufficiently compelling scientific reason to the contrary—is elimination.

5.4 STR Analysis and Interpretation:

This procedure is a general guideline for the analysis and interpretation of STR profiles when using the Identifiler (ID), Identifiler Plus (IDPlus), Minifiler (MF), Yfiler (YF), and PowerPlex-Y (PPY) STR kits. However, it is not an exhaustive list of all possible casework scenarios. Where appropriate, other interpretational approaches/methods may be used with approval from the DNA Technical Leader (TL). Analyst training, experience, and judgment must be considered when reporting STR profiles. STR data are interpreted by evaluating the results at all loci.

5.5 Autosomal STR:

5.5.1 This procedure is a general guideline for the analysis and interpretation of STR profiles when using the Identifiler, Identifiler Plus and Minifiler Kits. These procedures generally follow those outlined in ABI Prism GeneMapper ID Software v3.2 User Guide.

5.5.2 When the Identifiler or Identifiler Plus Kit is employed, the thirteen core CODIS STR loci, plus two additional loci (D2S1338 and D19S433), and Amelogenin are typed in a single PCR amplification reaction according to the manufacturer's and laboratory protocols. When the Minifiler Kit is employed, the loci D13S317, D7S820, D2S1338, D21S11, D16S539, D18S51, CSF1PO, FGA, and Amelogenin are typed in a single PCR amplification reaction according to the manufacturer's and laboratory protocols.

5.5.3 Identifiler, Identifiler Plus, and Minifiler amplification products are separated and detected using an ABI 3130xl Genetic Analyzer and the supplied Run 3130xl Data Collection v4.0 and GeneMapper ID v3.2 (GMID) softwares.

For the GMID software, the Analysis Method is as outlined in the user's manual except for the following: Under the Allele Tab, confirm that the Bin set is for the appropriate kit and that Marker Specific Stutter Ratio is used. Under the Peak Detector Tab, the Peak Detection Algorithm is Advanced, the Analysis Range is Partial Range but shall include the full range of required size standard peaks, the Sizing Range is Partial Sizes with the Start Size set to 75 (GS500-LIZ) or 60 (GS600-LIZ) for Identifiler and Identifiler Plus and 70-74 (for both GS500-LIZ and GS600-LIZ) for Minifiler and the Stop Size set to at least 450 (GS500-LIZ) or 400 (GS600-LIZ) for Identifiler and Identifiler Plus and 400 (GS500-LIZ) or 320 (GS600-LIZ) for Minifiler. Smoothing is set to Light for all three kits and the Size Calling Method is set to Local Southern Method for Identifiler and Identifiler Plus and 3rd Order Least Squares for Minifiler. For the Panel, use Identifiler_CODIS_v1.1 for Identifiler, Identifiler_Plus_Panels_v1 for Identifiler Plus, and Minifiler_GS500_v1 for Minifiler. For the Size Standard, use CE_G5_HID_GS500 for Identifiler and Minifiler and CE_G5_Identifiler_Plus_GS500 for Identifiler Plus. In panel manager, confirm that the marker specific stutter thresholds match those listed in SOP-6 for the appropriate amplification kit.

The analysis and interpretation of STR results are listed below.

5.5.4 <u>Identifiler, Identifiler Plus, and Minifiler Analysis:</u>

5.5.4.1 A GMID project is created by importing all of the pertinent data files.

5.5.4.2 The standard analysis parameters for samples and controls are B,G,Y,R = 75 RFU, O = 150 RFU (ID-75, IDPlus-75, or Minifiler-75). Evidentiary profiles (forensic unknowns) with peaks >5000 RFU are interpreted with caution with approval from the TL. With known samples and controls, any peaks > 6500 are interpreted with caution. For known samples, including positive controls, the thresholds for B, G, Y, & R may be raised where appropriate, and the RFU threshold for O may be lowered to 50 RFU. See below:

1. The standard analysis parameters for forensic unknowns, knowns, ladders, and controls are B, G, Y, R = 75 RFU, O = 150 RFU. Ladders and any samples with ILS peaks lower than these values are interpreted with caution, with the restriction noted in #2.
2. The minimum analysis parameters for ladder peaks and the internal lane standards are:

a) Ladder: B, G, Y, R, and O > 49 rfu (including 250 peak (and 340 peak for IDPlus)--R>49 for PPY.

b) With all samples and controls, the ILS peaks must be >49 rfu (including 250 peak (and 340 peak for IDPlus).

5.5.4.3 Ensure that all required internal lane size standard peaks are called in all samples (75 – 450 nt, GS500-LIZ; 60 – 400 nt for ID & IDPlus or 80-320, Minifiler GS600-LIZ). 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. Both the primary and the second analyst must confirm that the size quality failure was not due to the loss of internal lane size standard peaks. The sample may then be analyzed as normal.

5.5.4.4 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 must meet the requirements.

5.5.5 <u>GMID Analysis-- Control Requirements:</u>

5.5.5.1 Import the appropriate GMID sample data files including controls (RB, NEG, RKO, POS) and at least one ladder. At least one set of controls for each extraction/amplification set must give the expected results. If not, the amplification set must be re-injected, re-amplified, or re-extracted as warranted. Note that the RB may be omitted for re-amplifications using the same volume (or less) of template. The RKO may be omitted for re-amplifications if expected results were previously generated for the extraction set. All re-amplifications must include an amplification Negative and Positive control. The RB and NEG may be omitted for re-injections of equal or lesser time/sensitivity.

Expected results of controls will be verified and documented on DNA QR-4 which is maintained in the case file. Positive controls and extraction controls are additionally verified by using the POS/RKO Concordance Check Excel Spreadsheet for the specific test kit (DNA-QR-45A,C,E,F or DNA-QR-45D) and are maintained in the case file. Instructions to use this worksheet are within the excel workbook.

5.5.5.2 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. Analysts should use their experience and judgment during this process. 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. RB and NEG samples should not have any callable peaks. If callable peaks are detected, the set of samples may be re-run, re-amplified, or re-extracted as warranted. If callable peaks are detected and the evidentiary samples are not affected, the set of extracted samples may be interpreted with caution with TL approval.

If minor peaks are detected (any peaks ≥50 & <75 RFU), the set of samples may be re-run, re-amplified, or re-extracted as warranted. If minor peaks are detected and the evidentiary samples are not affected, the set of extracted samples may be interpreted with caution.

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 in the report or deemed unsuitable for comparison following TL review and approval as required by CT and Federal laws (see Alternate Report Template Statements located within the DNA Report Template).

Procedures listed here in 5.5.5.2 will be followed for all methodologies, on a case by case basis.

5.5.5.3 Select appropriate analysis parameters and proceed with analysis in GMID. Record any editing (including if no edits) of GMID results on the appropriate 3130 analysis worksheet (DNA-QR-16b, QR-16c, QR-16d, QR-16e, QR-16f, or QR-23). Sufficient detail will be noted on the 3130 analysis worksheet comments section so that it is clear to a reviewer. For example: "re-inject at max due to low peaks" or "re-prep sample- Liz failed".

5.5.5.4 Control Profiles: Ensure that all control samples performed as expected. The positive controls (POS, RKO) must give the expected results. For known sample processing, only 1 positive control (either POS or RKO) is required. POS/RKO may be omitted for re-injections.

The Identifiler and Identifiler Plus profiles for the Kit Positive Control DNA 9947A (POS) and the laboratory Extraction Positive Control (RKO) are listed in the table below:

6-FAM (blue)

Control	D8S1179	D21S11	D7S820	CSF1PO
RKO	12,13	28,29	8,11	9,11
POS	13	30	10,11	10,12

VIC (green)

Control	D3S1358	TH01	D13S317	D16S539	D2S1338
RKO	15,18	6,9.3	9	11,12	17
POS	14,15	8,9.3	11	11,12	19,23

NED (yellow)

Control	D19S433	vWA	TPOX	D18S51
RKO	13,15.2	14,18	8,11	14,15
POS	14,15	17,18	8	15,19

PET (red)

Control	Amelogenin	D5S818	FGA
RKO	X,Y	10,12	22,24
POS	X	11	23,24

The Minifiler profiles for the Kit Positive Control DNA 007 (mPOS) and the laboratory Extraction Positive Controls (RKO) are listed in the table below:

6-FAM (blue)

Control	D13S317	D7S820
RKO	9	8,11
mPOS	11	7,12

VIC (green)

Control	Amelogenin	D2S1338	D21S11
RKO	X,Y	17	28,29
mPOS	X,Y	20,23	28,31

NED (yellow)

Control	D16S539	D18S51
RKO	11,12	14,15
mPOS	9,10	12,15

PET (red)

Control	CSF1PO	FGA
RKO	9,11	22,24
mPOS	11,12	24,26

5.5.6 Interpretation of DNA Profiles in GMID:

5.5.6.1 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 "Appendices" for standard report conclusions.

5.5.6.2 Off-Scale Data: If too much DNA is added to the amplification reaction or is loaded onto the 3130, 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 GMID software. Any peaks greater than 5000 RFU may be off-scale. These samples may need to be diluted and re-amplified.

5.5.6.3 Peak identification: Only allele peaks with a height of ≥75 RFU in GMID are reported. The laboratory's reporting/allele calling threshold is ≥75 RFU. The laboratory's analytical threshold is ≥ 50 RFU. The identity of peaks called by GMID, is generally assigned to one of the following categories:

5.5.6.4 Allele Peak: A called allele has a peak height ≥75 RFU, a fragment size that falls within the base pair range, and has the appropriate dye color for the locus. However, not all peaks ≥75 RFU are typed as alleles. Stutter peaks, shoulders, pull-up/pull-down peaks, background noise, and amplification artifacts could fall above the RFU threshold, especially if the allele peaks are high. Peaks that do not meet the Laboratory's calling criteria/thresholds are not used for statistical purposes (see 5.10-5.13). Peaks ≥ 75 RFU that are not called by GMID may be called by the analysts on a case-by-case basis. Such peaks are interpreted with caution.

5.5.6.5 * Peaks (GMID Peaks): Low-level peaks are those less than 75 RFU but ≥ 50 RFU that have good peak morphology and are clearly distinguishable from background noise and other artifacts. Any * peaks detected are considered only to evaluate if the known source can be eliminated from the profile. These peaks cannot be called on the case report (or used for statistical purposes) but their presence is noted with an * as appropriate. This

information is part of the overall sample interpretation. Any peaks in the evidentiary samples or controls less than 50 RFU are not evaluated or noted on the worksheets.

5.5.6.6 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 GMID. If a particular variant has been observed before by the DESPP CT Division of Scientific Services (CTDSS), Applied Biosystems, NIST, or in published literature, the sample does not need to be repeated. In GMID, 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.

5.5.6.7 Stutter peaks (generally N-4 bases, but Identifiler Plus also has N+4 bases for all STR loci, Yfiler also has N-2, N-3, N-5, N-6, and N+3 and PowerPlex-Y also has N-2, N-3, N-5, N-8, N-9, and N+3 bases for select loci): 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 the GMID program. CT DSS stutter thresholds are determined by evaluating the ABI and CTDSS empirical stutter data on an allele-by-allele basis (see DNA SOP-6; Section 3). If the height of the stutter peak is greater than the CT stutter threshold, this indicates a possible mixture, an overloaded sample, or a higher than normal stutter percentage. Additional information provided by other loci may be of assistance in these cases. Peaks in the stutter position that are below the Laboratory's stutter thresholds are not listed on the report and are not used for statistical purposes. However, if a peak (below the stutter threshold) in a stutter position is relevant to the comparison of a particular known profile, a "^" may be noted in the allele table. After completion of 2nd analysis of an evidentiary profile, when comparing a specific known profile to that evidentiary profile, a "^" may be used to clarify the interpretation of that comparison. For that specific comparison, the locus/loci containing the "^" will not be used for statistical purposes. In most cases, the stutter percentage of the allele in question should be greater than the average stutter for that locus (see stutter validation study). The "^" signifies that the presence of the allele in the stutter position can neither be confirmed nor eliminated.

5.5.6.8 Minus A Peaks: AmpF*l*STR kits are 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 2.5 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. The level of minus A product may be reduced by adding more Taq or incubating the sample at

60 ^{o}C for an additional 30-90' with approval from the TL. When appropriate, a smaller quantity of sample may be re-amplified.

5.5.6.9 Pull-up Peaks: Pull-up peaks are generally seen when excess amplification product is injected on the CE 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 or asterisk designation. The estimated pull-up contribution is determined by evaluating the height of unambiguous pull-up peaks (pull-up with e.g., no stutter or allelic contribution) at each locus where pull-up is detected for a particular color.
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 (or *) where DNA Section allele calling/ * peak criteria are met. The same methodology would be applied to potential allelic peaks with a pull-down component (see below).

5.5.6.10 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 slightly a larger or smaller size as compared to the true allele. Additional information provided by other loci may be of assistance in these cases.

5.5.6.11 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.

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

5.5.6.13 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.

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

5.5.6.15 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; the "pull-down" designation should be used for the editing notation when raised baseline is caused by pull-down.

5.5.7 Interpretation of Mixed Samples in Identifiler, Identifiler Plus, and Minifiler:

5.5.7.1 Generally, the detection of more than two alleles per locus indicates a mixed sample. Variation in peak height between alleles in a single locus may assist in the interpretation of such results. Peaks below the 75 RFU threshold for GMID may assist in the interpretation of mixed DNA samples. The following guidelines pertain to the interpretation of mixtures.

5.5.7.2 A DNA mixture is conclusively "demonstrated" by the following; (a) the detection of three or more alleles at two or more loci or (b) 4 or more alleles detected at one or more loci.

A DNA mixture can be defined as "consistent with" (but not proof of); (c) 3 alleles at 1 locus, or as discussed below in 5.5.7.3 & 5.5.7.4.

If the criteria for a DNA mixture described in (a), (b), or (c) above are met, the unrestricted CPI is calculated for that item except where mixture deconvolution is performed (see 5.10.3). Three peak patterns (at a single locus) have been observed from single sources, but they are rare.

5.5.7.3 Peak imbalance involving the stutter position, e.g., where the N-4 peak is 15-25% of the N peak, may indicate a mixture. With the possibility of allele sharing among individuals, 3 alleles may not be detected especially with degraded samples or partial profiles.

5.5.7.4 General peak imbalance: Heterozygous peak imbalance greater than 40% may indicate a mixture, i.e., the smaller peak is less than 60% of the larger peak. However, microvariants in the primer-binding domain may cause either PCR enhancement or suppression.

5.5.7.5 Mixture deconvolution: Mixtures are deconvoluted for CODIS entry purposes as warranted by the results. Mixtures are deconvoluted for Identifiler and Identifiler Plus DNA profiles on the report and in the allele table (See DNA Report Template) and for statistical purposes only when the following conditions are met:

1. The mixture must not contain more than two person's DNA at any loci. The height of the major peaks is at least 9 times the height of the minor peaks (90:10 major/minor) at all loci in the mixture. Genotype assignment is made only when the peaks of the contributor are above 200 RFU for Identifiler DNA profiles OR the stochastic threshold for Identifiler Plus DNA profiles: 150 RFU for 2-15 second injections, 200 RFU for 16-24 second injections, and 250 RFU for 25-30 second injections at 3kV on the 3130-2 instrument. See table below for stochastic thresholds versus injection times for all 3130 instruments (also see DNA SOP 4; Section 4.9.2.3). If the major contributor genotype is above 200 RFU for Identifiler (ID) DNA profiles OR the stochastic threshold for Identifiler Plus (ID+) DNA profiles but the minor profile is less than 200 RFU (ID) OR the stochastic threshold (ID+), only the genotype of the major contributor may be deduced. For heterozygote contributors, the heterozygous peak balance must be at least 80%.

	Injections Time Ranges				
Stochastic Threshold	**3130-1**	**3130-2**	**3130-3**	**3130-4**	**3130-6**
150 RFU	2 – 18 sec	2 – 15 sec	2 – 10 sec	2 – 10 sec	2 – 10 sec
200 RFU	19 – 28 sec	16 – 24 sec	11 – 16 sec	11 – 16 sec	11 – 16 sec
250 RFU	29 – 35 sec	25 – 30 sec	17 – 20 sec	17 – 20 sec	17 – 20 sec

2. Mixtures are deconvoluted as described above upon agreement by the prosecution and the defense or upon order of the court.

3. All mixture deconvolutions must be reviewed and approved by the TL.

4. Mixture deconvolution shall not be performed for any Minifiler, Y-STR or mtDNA profiles.

5.6 Highly Degraded/Low Template Samples:

5.6.1 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. Given the possibility of stochastic allele sampling, especially with low quantity and/or degraded samples, single peaks less than approximately 200 RFU are interpreted with caution (pertains to putative genotype assignment at that locus). To maximize the number of loci with peaks above the threshold, a sample may be run at increased injection times on the 3130. However, any off-scale data at the smaller amplicons are interpreted with caution.

5.7 Inhibited Samples:

5.7.1 Inhibited samples may exhibit allele, locus, or complete profile dropout, unbalanced heterozygous alleles, 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 Duo IPC for each sample is useful. The Quantifiler Duo IPC result for a sample being undetected or having an elevated CT 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.

5.8 General Interpretation of STR Results:

5.8.1 The following is a general description of the standard DNA report conclusions. Other conclusions may be reported as warranted by the results with approval of the technical leader. (See report templates, DNA SOP-6; Section 2 and DNA Report Template)

5.8.1.1 Elimination: used when J. Smith is not the source (excluded) of the DNA profile detected from item xx. [Note: While chance allele sampling and stochastic fluctuation (allele drop-out) can occur at any locus, extensive validation and case experience has shown that smaller amplicons amplify more efficiently and are less affected by degradation than are larger amplicons. Therefore, the results at the smaller loci are particularly diagnostic]. Either the DNA of J. Smith is not on the evidence or it is not present to a detectable level. Report statement: J. Smith is eliminated as the source of (or contributor to) the DNA profile detected from item xx.

5.8.1.2 Inconclusive: Leading to no conclusion or definitive result. "Based on limited data, the results are inconclusive as to whether J. Smith could be the source of, or a contributor to, the DNA profile from item xx." This statement is made when the data does not provide sufficient support for any conclusion regarding a comparison to a particular known.

5.8.1.3 Insufficient for comparison (IFC): "Insufficient amplification products were detected from item xx for comparison to the known DNA profile of J. Smith." The statement is made for low-level, partial DNA profiles when there is insufficient data from the evidentiary DNA profile for comparison to a particular known.

5.8.1.4 Cannot be eliminated (CBE): The data do not support excluding J. Smith A as the source of, or contributor to the DNA profile/results detected for item xx. This conclusion is used when there is a strong genetic linkage/association between a known and the DNA profile/results detected for item xx. However, all the alleles from the known were not detected (above threshold) at all loci for which results were generated. This conclusion is analogous to a partial inclusion/match and would be supported by results such as the presence of * peaks, degraded samples, low template amplification, mixtures, and the probability of allele drop-out. (See also note above in 5.8.1.1). Note that CBE is not equivalent to CBI (cannot be included).

5.8.1.5 Inclusion: All of the alleles of J. Smith were detected (≥75 RFU) in the evidentiary profile at all loci that generate results. "The results are consistent with J. Smith being the source of (or contributor to) the DNA profile from item xx."

5.8.1.6 If only * peaks are present in a profile, the result can be used for elimination but not for inclusion. Report that "J. Smith is eliminated as the source of the minor peaks from item xx."

5.8.2 Interpretation of complex/comprehensive mixtures

For complex/comprehensive 4+ person mixtures where multiple persons are detected throughout all loci, the following general interpretational approach is employed when comparing the evidentiary profile results to a known profile: if peaks consistent with the known source are not detected at all loci (where results are generated) above the analytical threshold (i.e, called and * peaks), the known source is eliminated. Other conclusions may be reported as warranted by the results on a case-by-case basis with approval of the Technical Leader.

5.9 Y-STR:

5.9.1 This procedure is a general guideline for the analysis and interpretation of STR profiles when using the Yfiler or PowerPlex-Y STR kits. These procedures generally follow those outlined in ABI Prism GeneMapper ID Software v3.2 User Guide.

5.9.2 When the PowerPlex-Y kit is employed, twelve Y-STR loci, DYS391, DYS389I, DYS439, DYS389II, DYS438 (5 base repeat), DYS437, DYS19, DYS392 (3 base repeat), DYS393, DYS390, and DYS385 (2 loci amped with one set of primers), are typed in a single PCR amplification reaction according to manufacturer's and laboratory protocols. When the Yfiler kit is employed, seventeen Y-STR loci, DYS456, DYS458, DYS635, Y-GATA H4, DYS448 (6 base repeat), and the twelve loci listed above for the PowerPlex-Y kit.

5.9.3 Yfiler and PowerPlex-Y amplification products are separated and detected using an ABI 3130xl Genetic Analyzer and the supplied Run 3130xl Data Collection v4.0 and GeneMapper ID v3.2 (GMID) softwares.

For the GMID software for PPY, the Analysis Method, Panel, and Size Standard are outlined in the PanelandBinReadme3.2 PDF file that is titled Using the Promega PowerPlex Systems with the GeneMapper ID, Version 3.2 (located in the Panels_Bins folder on the U:\ Drive). Note: The PowerPlex-Y Kit contains one locus that is a 3 base repeat and one locus that is a 5 base repeat, so there are some extra parameter values needed for the Analysis Method under the Allele Tab, Marker Repeat Type. Additionally, under the Peak Detector Tab, the size standard ILS-600 requires a Sizing, Partial Sizing, Start Size of 60 and the Stop Size needs to be at least 400. For the panel, use PowerPlex_Y_ID3.2.0. The size standard is created by the user per the document listed above and is called ILS-600. In panel manager, confirm that the marker specific stutter thresholds match those listed in SOP-6 for the PowerPlex-Y amplification kit. For the GMID software for YF, the Analysis Method is as outlined in the user's manual except for the following: Under the Allele Tab, confirm that the Bin set is for the appropriate kit and that Marker Specific Stutter Ratio is used. Note: The Yfiler Kit contains one locus that is a 3 base repeat, one locus that is a 5 base repeat, and one locus that is a 6 base repeat, so there are some extra parameter values needed for the Analysis Method under the Allele Tab, Marker Repeat Type. Under the Peak Detector Tab, the Peak Detection Algorithm is Advanced, the Analysis Range is Partial Range but shall include the full range of required size standard peaks, the Sizing Range is Partial Sizes with the Start Size set to 75 (GS500-LIZ) or 60 (GS600-LIZ) and the Stop Size set to at least 400 (GS500-LIZ) or 360 (GS600-LIZ). Smoothing is set to Light and the Size Calling Method is set to Local Southern Method. For the Panel, use Yfiler_v2. For the

Size Standard, use CE_G5_HID_GS500 (GS500-LIZ) or GS600LIZ (GS600-LIZ). In panel manager, confirm that the marker specific stutter thresholds match those listed in SOP-6 for the Yfiler amplification kit.

The analysis and interpretation of Yfiler and PowerPlex-Y STR results are the same as for AmpF*l*STR Identifiler or Identifiler Plus except as noted.

5.9.4 PowerPlex-Y Analysis Setup:

5.9.4.1 The standard threshold analysis parameters for samples and controls are B,G,Y = 75 RFU, R= 150 (PPY-75). Profiles (forensic unknowns) with peaks >3000 RFU are interpreted with caution with approval of the TL. Known profiles > 5000 RFU are interpreted with caution with approval of the TL.

5.9.4.2 Ensure that all required internal lane size standard peaks are called in all samples (60 – 400 nt, PPY), (75 – 400 YF, GS500-LIZ), (60 – 360 YF, GS600-LIZ). See Section 5.5.4.2.

5.9.5 Y-STR Control Requirements:

5.9.5.1 Import the appropriate sample files including controls (RB, NEG, RKO*, {POS-9948 or POS 2800M for PPY}, {POS control DNA 007 for YF) 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-run, re-amplified, or re-extracted as warranted. *With extraction sets where a male laboratory control (RKO) is not available, importing the PPY or YF kit positive control is sufficient.

5.9.5.2 **RB and NEG samples should not have any callable peaks. If callable peaks are detected, the set of samples may be re-run, re-amplified, or re-extracted as appropriate. If callable peaks are detected and the evidentiary samples are not affected, the set of extracted samples may be interpreted with caution with TL approval. If minor peaks (any peaks ≥ 50 rfu & <75 rfu) are detected, the set of samples may be re-run, re-amplified, or re-extracted as appropriate. If minor peaks are detected and the evidentiary samples are not affected, the set of extracted samples may be interpreted with caution (see 5.5.5.2). Note that 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

5.9.5.3 **Positive Control Profiles: Ensure that all alleles were included and that the positive controls give the expected profile. The RKO 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. The PPY profiles for the kit positive control DNA 9948 (POS) or 2800M (POS) and the laboratory extraction positive control (RKO) are listed in the table below:

	FL (blue)				JOE (green)	
Control	**DYS391**	**DYS389**	**DYS439**	**DYS389II**	**DYS438**	**DYS437**
9948 POS	10	13	12	31	11	15
2800M POS	10	14	12	31	9	14
RKO	11	13	11	29	12	15

	TMR (yellow)				
Control	**DYS19**	**DYS392**	**DYS393**	**DYS390**	**DYS385**
POS	14	13	13	24	11,14
2800M POS	14	13	13	24	13,16
RKO	14	13	13	24	11,15

The YF profiles for the kit positive control DNA 007 (POS) and the laboratory extraction positive control (RKO) are listed in the table below:

6-FAM (blue)

Control	DYS456	DYS389I	DYS390	DYS389II
RKO	15	13	24	29
POS 007	15	13	24	29

VIC (green)

Control	DYS458	DYS19	DYS385
RKO	17	14	11,15
POS 007	17	15	11,14

NED (yellow)

Control	DYS393	DYS391	DYS439	DYS635	DYS392
RKO	13	11	11	24	13
POS 007	13	11	12	24	13

PET (red)

Control	Y GATA H4	DYS437	DYS438	DYS448
RKO	12	15	12	19
POS 007	13	15	12	19

5.9.6 Interpretation of Y-STR DNA Profiles in GMID:

5.9.6.1 Stutter peaks: 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 GMID. CT DSS stutter thresholds are determined by evaluating the Promega, Life Technologies, and CT DSS empirical stutter data on an allele-by-allele basis (see Appendix 6.3). Any peaks < the CT stutter threshold are manually edited as required. If the height of the peak in the stutter position is greater than the CT stutter threshold, this indicates a possible mixture, an overloaded sample, or a higher than normal stutter percentage. Additional information provided by other loci may be of assistance in these cases.

5.9.6.2 X-homologous peaks: Rarely, (especially with large quantities of female DNA) the Y primers may generate X-homologous or non-specific peaks. If the known female sample is amplified, any X-homologous peaks may be removed from the evidentiary DNA profile.

5.9.7 Interpretation of Mixed Samples for Y-STR:

5.9.7.1 Generally, the detection of more than one allele per locus (other than the DYS385 locus) indicates a mixed sample. Variation in peak height between alleles in a single may assist in the interpretation of such results. Peaks below 75 RFU threshold may assist in the interpretation of mixed DNA samples. The following guidelines pertain to the interpretation of Y-STR mixtures:

5.9.7.2 The detection of two or more alleles at two or more loci typically identifies a sample mixture (note: 3 or more peaks at DYS385 and 2 + peaks at another locus satisfies this criterion). The presence of 4 alleles at DYS385 or 3 + peaks at any of the other Y STR loci also indicates a mixture. Other results are consistent with (but not proof of) a mixture as discussed below in section 5.10.3. Two peak patterns (at a single locus other than DYS385) have been observed from single source samples, but they are rare.

5.9.8 Inhibited Samples:

5.9.8.1 Input of excessive female DNA may result in inhibition of Y-STR amplification. YF was less sensitive than PPY to this type of inhibition. 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.

5.9.9 Report Conclusions for Y-STR Results:

5.9.9.1 Y-STR report templates are described in DNA SOP-6; Section 2 and DNA Report Template. The discussion in 5.8.1 applies to Y-STR conclusions except as listed below:

5.9.9.2 For a CBE conclusion: "J. Smith (or another member of the same paternal lineage) cannot be eliminated as the source of (or contributor to) the Y-STR DNA profile from item xx."

5.9.9.3 For an inclusion: "The results are consistent with J. Smith (or another member of the same paternal lineage) being the source of (or contributor to) the Y-STR DNA profile from item xx."

5.10 Statistics—Random Match Probabilities (RMP)

Note: For standard reporting purposes statistics will be calculated for the African American, Caucasian, and Hispanic populations

5.10.1 Autosomal STR Statistics:

Match probabilities are calculated according to the Hardy-Weinberg equilibrium and the product rule with the exceptions listed below. A statistical evaluation of any scientifically relevant result is included on the case report where comparisons to known profiles are possible except for where the results are inconclusive, IFC, exclusionary or if there is a comparison of a victim to their own non-probative intimate sample, such as a vaginal swab, epithelial rich sample. An intimate sample is defined as any sample collected directly from the body, e.g., body cavity, skin swabs, nail clippings/swabbings/scrapings, etc.

5.10.2 Autosomal RMP Ceiling:

5.10.2.1 For court presentation, the CT Division of Scientific Services will report a RMP ceiling of 1.43×10^{-10}. Any match probability (or CPI) less than 1.43×10^{-10} (1 in 7 billion) will be reported as follows: the expected frequency of individuals who could be the source of

the observed DNA profile is less than 1 in 7 billion. This does not imply that the product rule calculation (without a ceiling) is scientifically unwarranted. In the event that the ceiling is not reached, match probabilities are reported according to FBI Quality Assurance Standards and as described in NRC II:

Heterozygotes	$2pq$
Homozygotes	$p^2 + p(1-p)\theta$; where $\theta = 0.01$

5.10.3 Autosomal Mixtures:

5.10.3.1 The unrestricted combined probability of inclusion (CPI) is calculated for mixtures (demonstrated and consistent with), (see 5.5.7). Mixtures are not deconvoluted on the report or for statistical purposes in court except as specified in 5.5.7.5.

$PI = P^2$, where P is the sum of the observed allelic frequencies.
PI is the probability of inclusion at one locus.
$CPI = (PI_i)(PI_j)(PI_k)\ldots(PI_n)$ = probability of inclusion at multiple loci.

When mixtures are deconvoluted, use the statistical approach listed in 5.10.2.1 regarding the deduced genotype following the review and approval of the TL.

5.10.3.2 A locus is used in the CPI calculation only when the allele(s) of the known sources have been detected in the evidentiary profiles (peak height = at least 75 rfu). Loci with * peaks detected in addition to called alleles, are evaluated with caution to consider the impact of possible allelic dropout, but may be used for the CPI calculation. The * peak(s) in those loci are not used in the CPI calculation.

5.10.3.3 With low-level mixtures where there is clear evidence of the minor contributor(s) only at the smaller amplicons; the larger amplicons (including where the known is not eliminated) may be omitted from the CPI calculation.

5.11 Y-STR Statistics:

5.11.1 For single source Y-STR samples a Y-STR random match probability (Y-RMP) is calculated. A haplotype frequency is generated by counting the number of times a haplotype has been observed in the Database divided by the total number of haplotypes in the Database (the counting method). Haplotype frequencies are given an upper bound 95% confidence interval (See DNA SOP-6; Section 2, DNA Report Template, and QR-43). The National Y-STR population database (available at www.usystrdatabase.org) or the CT Y-STR population database can be used as appropriate.

5.11.1.1 An upper bound 95% confidence interval is calculated from the results of a database search in order to estimate the population frequency of a haplotype. The upper bound estimate is dependent on the size of the database, and these estimates may change as the database size changes.

5.11.1.2 The following formulas are used to calculate the upper bound estimate:

5.11.1.2.1 In cases where the haplotype has been observed in a population database (i.e. African American, Caucasian, or Hispanic):

$$\leq p + \left(1.96 \sqrt{\frac{p(1-p)}{N}} \right)$$

p = x/N, where x is the number of times a haplotype has been observed in a population and N is the number of haplotypes in that population.

5.11.1.2.2 In cases where the haploype has <u>not</u> been observed in a population database:

$$\leq 1 - \left(0.05^{(1/N)} \right)$$

0.05 is the confidence coefficient for an upper bound 95% confidence interval and N is the number of haplotypes in the population.

5.11.2 <u>Y-STR Mixtures:</u>

5.11.2.1

The expected frequency of haplotypes that could contribute to a Y-STR mixture is calculated using a 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.CCM = 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 upper bound 95% confidence interval is then applied (See 5.11.1.2.1 and 5.11.1.2.2 (above), DNA SOP-6; Section 2 and QR-43). The

CCM (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. Y-STR mixtures are not deconvoluted on the report or for statistical purposes in court.

5.11.3 Y-STR haplotypes (and mtDNA haplotypes) are inherited independently from autosomal loci. Therefore, Y-RMP's or CCM's may be combined according to the product rule with autosomal STR RMPs or CPIs to generate an overall profile frequency estimate.

5.12 Criminal Parentage Testing:

Standard Identifiler/Identifiler Plus (and Minifiler for degraded samples) loci are used for criminal parentage testing. Standard statistical methods for Identifiler, Identifiler Plus and Minifiler results will be used as described below:

5.12.1 The statistic calculated (for inclusions) is the expected frequency (parentage inclusion probability) of individuals who could contribute the paternally (or maternally) transmitted alleles. The formula used is: $p^2 + 2p(1-p) = 2p-p^2$, where p is the frequency of the obligatory allele.

5.12.2 No statistics are required for non-matches. However, due to the occurrence of mutations between generations, an individual must be excluded at more than two loci (when 13-15 STRs are typed) in order to be eliminated as a potential parent. Regarding apparent mutations using STR systems, the repeat # difference of the putative mutation (child's obligatory allele vs. alleged parent's alleles) may be relevant to the final conclusion. In the event of apparent mutations, additional testing may be conducted as appropriate.

5.13 Three-Band Patterns: (Autosomal STRs):

5.13.1 Three band patterns from single sources have been noted infrequently. Three band pattern matches are significant findings. However, for the purposes of match probabilities, the three-band locus is not factored into the calculation. This does not imply that a modified counting approach (i.e., 5/2N) is not valid.

5.14 Interpretation of Identifiler or Identifiler Plus Results (Databasing):

5.14.1 This procedure is a general guideline for the analysis and interpretation of single source STR profiles when using the Identifiler or Identifiler Plus kits. The analysis and interpretation of database Identifiler or Identifiler Plus STR results are as described in sections 5.4-5.7 with the exceptions listed below:

5.14.2 Identifiler and Identifiler Plus Analysis Set-up:

5.14.2.1 The standard analysis parameters for convicted offenders are B,G,Y,R = 75 RFU, O = 150 RFU. The standard analysis range is 75 to 6500 RFU. Peaks >6500 RFU are interpreted with caution.

5.14.3 Identifiler and Identifiler Plus Control Requirements:

5.14.3.1 Import the appropriate GMID sample files including controls (RB, NEG, RKO, POS, QC) and at least one ladder. If * peaks or called peaks arise in the RB or the NEG on database plates, interpret with caution, and see CODIS administrator for approval/documentation of the control. If the CODIS administrator does not approve the control it will be necessary to re-amplify or re-extract affected samples.

5.14.3.2 GMID projects are archived; GMID edits are not recorded on analysis worksheets.

5.14.3.3 Control Profiles: Ensure that all control samples performed as expected (see 5.5.5). The positive controls (RKO, POS, QC) must give the expected results. For convicted offender sample processing, only 1 "positive control" (either POS or RKO) is required. The Identifiler and Identifiler Plus profiles for the kit positive control DNA 9947A (POS) and the laboratory extraction positive control (RKO) are listed in section 5.5.5.4.

5.14.3.4 Technical Review of Identifiler and Identifiler Plus Convicted Offender DNA profiles is limited to a review of the following data parameters using the first analyst's data files:

All required size standard peaks are called.
All expected alleles are called in the ladder(s) used.
All controls performed as expected.
All offender profiles are correctly called.
Specimen categories for CODIS are assigned correctly.

5.15 Minifiler Analysis and Interpretation

5.15.1 Minifiler DNA profiles are analyzed and interpreted as described in section 5.4 through section 5.8.

5.15.2 Minifiler DNA results are listed in a separate allele table in the DNA Report. See DNA SOP-6; Section 2 and DNA Report Template.

5.15.3 Minifiler stutter thresholds are listed in DNA SOP-6; Section 3.

5.16 Composite Interpretation of Identifiler or Identifiler Plus and Minifiler Results

5.16.1 For the purposes of determining a mixture (see section 5.5.7, DNA SOP-6; Section 2, and DNA Report Template), a composite profile of Identifiler or Identifiler Plus and Minifiler is used.

5.16.2 Separate interpretations are performed for Identifiler or Identifiler Plus and Minifiler. When interpretations are identical, a combined conclusion statement will be reported. See DNA SOP-6; Section 2 and DNA Report Template.

5.16.3 Where appropriate, a composite interpretation of Minifiler and Identifiler or Identifiler Plus results may be reported. For the overlapping loci, the combined statistic must include all observed alleles in Identifiler or Identifiler Plus and Minifiler where the known is fully included in at least one system.

5.16.4 Overlapping loci must be used only once in the RMP/CPI calculations. If any loci in either Identifiler, Identifiler Plus or Minifiler "demonstrate" or are "consistent with" a mixture, the CPI statistic is calculated for the entire profile.

5.16.5 When the Identifiler or Identifiler Plus and Minifiler results cannot be combined, a separate conclusion is reported for each test result.