STR Analysis on 3500xL Genetic Analyzers

1 Guiding Principles and Scope

1.1 These guidelines for analysis are applicable for samples amplified using PowerPlex® Fusion 5C (Fusion) and run on 3500xL Genetic Analyzers.

1.2 The purpose of these guidelines is to provide a framework which can be applied to the analysis and interpretation of STR results in casework. The guidelines are based on validation studies, literature references, standard rules, and experience.

1.3 This manual may not cover all situations that arise, and not every situation can be covered by a pre-set rule. Equipped with these guidelines, analysts should rely on professional judgment and expertise as well as their supervisor for further guidance.

2 Allele Calling Criteria

2.1 Electropherograms are analyzed through observation of peaks at the loci simultaneously amplified using the PowerPlex® Fusion 5C system. During amplification and capillary electrophoresis, an allele is characterized by dye-colored locus specific primers and the length of the amplified fragment compared to the size standard. Identification of a peak as an allele is then determined by comparison to the allelic ladder. To eliminate possible background peaks, only peaks that display intensity above the minimum analytical threshold (AT) are labeled as potential alleles.

2.2 Computer program processing steps for raw data:

2.2.1 Recalculation of fluorescent peaks using the instrument-specific spectral file to correct for the overlapping spectra of the fluorescent dyes.

2.2.2 Calculation of the fragment length for the detected peaks using the known internal-lane standard fragments.

2.2.3 Compare and adjust the allele categories to the sizing of the co-electrophoresed allelic ladder by calculating the offsets (the difference between the first allele in a category and the first allele in the allelic ladder at each locus).

2.2.4 Labeling of all sized fragments that are above the AT, exhibit appropriate peak morphology, and fall within or between the locus specific size ranges (see the GeneMarker v3.0 Operation Manual).

2.2.5 Removal of the labels from non-allelic peaks (background noise and/or stutter) according to the AT and stutter filter functions.

2.2.5.1 Analytical thresholds can be found in the corresponding interpretation manuals (Interpretation of PowerPlex® Fusion data run on 3500xL)

2.2.5.2 Stutter filtering is described in Section 3.3.

2.3 Allele Nomenclature:

2.3.1 After the assigning of allele names to the remaining labeled peaks, the software creates a result table where all peaks that meet the above listed criteria are listed as alleles. The allele nomenclature follows the recommendations of the International Society for Forensic Haemogenetics (ISFH), (DNA recommendations, 1994) and reflects the number of core repeat units for the different alleles.

2.3.2 Subtypes displaying incomplete repeat units are labeled with the number of complete repeats and a period followed by the number of additional bases.

2.4 Electropherograms:

2.4.1 Capillary electrophoresis plot data containing case specific samples are a part of each case record. **The electrophoresis plots or electropherograms are the basis for the interpretation of results.**

2.4.2 The electropherogram will display all labeled and unlabeled peaks at each locus, peak height information, and base pair size.

2.4.2.1 **Fusion:** For a single-source sample, in general, a locus can be either homozygous and show one allele or heterozygous and show two alleles.

2.4.3 Reporting analysts will interpret the electropherograms.

2.5 Discrepancies for overlapping loci in different multiplex systems:

2.5.1 The primer-binding site of an allele may contain a mutation.

2.5.1.1 This mutation may make the annealing phase of amplification less efficient.

2.5.1.2 Alternatively, if the mutation is near the 3' end, this may completely block extension (Clayton et al. 1998).

2.5.2 In **Fusion** samples, mutations may result in a pseudo-homozygote type.

2.5.2.1 For a specific set of primers, this is reproducible.

2.5.2.2 However, these mutations are extremely rare, estimated between 0.01 and 0.001 per locus (Clayton et al. 1998).

2.5.2.3 If a pseudo-homozygote type for a locus was generated, evidence and exemplar samples amplified with the same primer sequence can be used for comparison.

2.5.2.4 Fusion does not have the same primer sequences as kits from different manufacturers, for example Identifiler and Minifiler. Therefore, the results at a single locus in Fusion may not be identical when compared with those of Identifiler.

2.5.2.5 If a sample is amplified using two different multiplex systems, it is possible for a locus to have a heterozygote type in one multiplex and a pseudo-homozygote in the other. This may be due to differences in primer sequences. The heterozygote type is the correct type and should be reported.

3 Removal of Labels for Non-Allelic Peaks

3.1 Additional **non-allelic peaks** may occur under the following instances (Clark 1988, Walsh et al. 1996, Clayton et al. 1998, Buckleton et al. 2011), which may have labels manually removed. The edit sheet documents all labels that are removed during this process.

3.2 Running a replicate amplification may assist with artifact recognition.

3.3 **Stutter** is an artifact caused during the amplification process due to slippage of one of the DNA strands resulting in an insertion or deletion of a repeating unit(s) or partial unit. It is smaller in height than the main allele and 2 bp or 1-2 full repeats smaller or larger than the main allele.

3.3.1 The stutter types considered in analysis of electropherogram data are as follows:

- **Back stutter** – 1 full repeat smaller than the parent allele
- **Forward stutter** – 1 full repeat larger than the parent allele
- **Double-back stutter** – 2 full repeats smaller than the parent allele
- **Half-back stutter** – 2bp smaller than the parent allele, for some tetranucleotide loci

3.3.2 The analysis software settings for each system includes one or more stutter filters for each locus and/or allele (see the Appendix for PowerPlex® Fusion Stutter).

3.3.2.1 In GeneMarker®, peaks can only be considered stutter by the software if they are within the appropriate stutter position. If a peak is outside of the stutter position of the parent peak, the software will not automatically filter this out as stutter, even if it falls below the stutter filter percentage for that location. See Section 5 for more information.

3.3.2.2 Fusion:

3.3.2.2.1 Allele-specific stutter filters are applied additively. For example, if a peak is in both back and forward stutter position of two different parent peaks, the peak in stutter position will have both back and forward stutter filters applied accordingly.

3.3.2.2.2 Allele-specific stutter filters are not applied within the GeneMarker® software to peaks in stutter position when the parent peak is out-of-bin (OB) or off-ladder (OL) or the stutter peak is off-ladder (OL). See Section 5.

3.3.2.2.3 Allele-specific stutter filters are applied within the GeneMarker® software to OB peaks in stutter position when the parent peak has an allele call. See Section 5.

3.3.3 Labels on stutter peaks may be removed from exemplar and positive control samples.

3.3.3.1 Consider the height of the stutter peaks when evaluating the exemplar for possible mixture or contamination. See Appendix for PowerPlex® Fusion Stutter.

3.3.4 Labels on stutter peaks should NOT be removed from evidence samples.

3.3.4.1 Labels should not be removed on other types of artifacts if they could also be considered a stutter artifact.

3.3.5 For further information on Fusion stutter modeling, see Appendix for PowerPlex® Fusion Stutter.

3.4 Pull-up, pull-down, and over-subtraction spectral artifacts

- Pull-up or pull-down peaks in one dye color can occur when there are high rfu peaks in another dye color. These peaks are spectral artifacts caused by the inability of the software to compensate for the spectral overlap between the different dye colors if the peak height in one color is too high.

3.4.1 Pull-up

- Pull-up peaks in one dye color will have a base pair size very close to the true allelic peak in another color. The pull-up artifact peak will always be shorter than the true allelic peak. For example, a high rfu peak in the blue dye channel could potentially create a pull-up artifact in the green, yellow, red, or orange dye channel.

3.4.2 Pull-down

- Pull-down peaks can appear as a valley in one dye color above or below a true allelic peak in another dye color. For example, an allele in the blue dye channel could cause a pull-down artifact in the green dye channel. This is due to the correction for oversaturation. The artifact peak label may be present on either or both sides of the valley.

3.4.3 Spectral over-subtraction

- Spectral artifacts can manifest as a raised baseline between two high rfu peaks.

- Spectral over-subtraction can also show up as split peaks, see below.

3.4.4 Labels on pull-up, pull-down, and raised baseline spectral over-subtraction artifacts should be removed in all samples.

3.5 Split peaks

- Split peaks are characterized by the main allelic peak appearing as if it is split into two peaks and are treated differently depending on the cause of the split peak.

3.5.1 Split peaks from spectral over-subtraction

- Split peaks may occur from spectral over-subtraction in overblown or high rfu alleles. For example, an overblown peak in the green dye channel may dip at the top where there is another high rfu peak in yellow; these will likely be accompanied by a pull-up peak in blue and red.

3.5.1.1 When the split peak is due to spectral over-subtraction, the sample is usually overblown and will be treated differently depending on the sample. See Section 4 for more information on saturation limits and how to handle saturated samples.

3.5.1.1.1 If the sample can be edited per Section 4 and both sides of the split peak are labeled, the non-allelic label should be removed.

3.5.2 Split peaks from incomplete adenylation

- These can occur due to the Taq polymerase activity that causes the addition of a single "A" to the terminus of the amplified product ("N+1" band). Since allele calling is based on N+1 bands, a complete extra "A" addition is desired. Split peaks due to incomplete non-nucleotide template "A" addition should not occur for samples with low amounts of DNA.

3.5.2.1 When the split peak is coming from incomplete nucleotide template "A" addition, the N-1 label should be removed in all samples.

3.6 Shoulder peaks

- Shoulder peaks are approximately 1-4 bp smaller or larger than main allelic peaks and can be recognized by their shape. They appear as a continuation of the main allelic peak and have low peak heights.

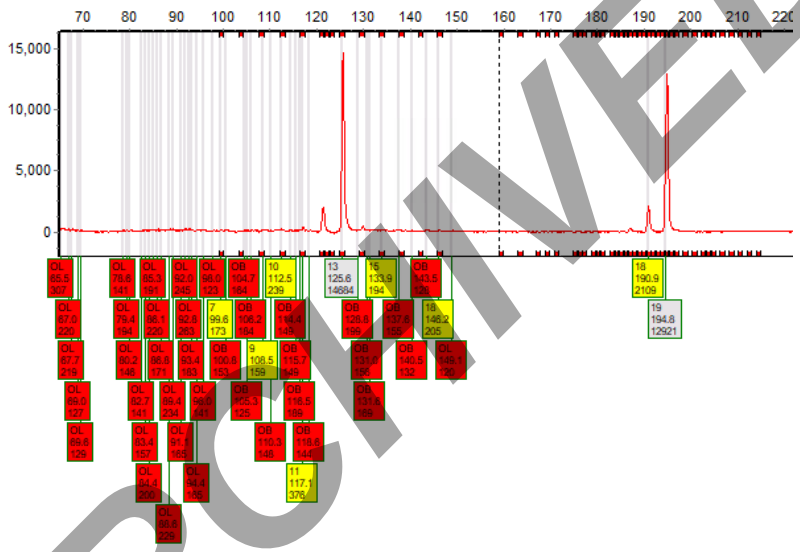
3.6.1 Labels on shoulder peak should be removed in all samples.

3.7 Elevated baseline

- Elevated or noisy baseline may have labels but do not resemble allelic peak morphology. This may sometimes occur adjacent to a shoulder peak or high rfu peaks.

3.7.1 Labels on elevated baseline should be removed in all samples.

3.7.2 In Fusion on 3500xL, elevated baselines have been shown to occur even in samples without saturated peaks. This appears as multiple low RFU allele-designated peaks and OB/OL peaks to be labeled at the lower molecular weight loci. This may appear within multiple dye channels, in multiple samples within an injection, and across multiple injections of the same sample.



3.7.2.1 Label removal when this occurs is treated differently depending on the sample.

3.7.2.1.1 Exemplars and controls may be interpreted if artifacts are attributable to the elevated baseline and can be edited. Otherwise, see Section 6 for handling of failed controls and retesting strategies.

3.7.2.1.2 Caution should be taken when evaluating evidence samples with elevated baseline as true allelic peaks could be masked by artifacts. If identifiable, artifacts may be edited manually.

3.7.2.1.3 If artifacts cannot be differentiated from true alleles, all peak labels must be removed, and the electropherogram with deleted labels will be included in the file. Enter the appropriate re-run code.

3.8 Dye artifacts

- Dye artifacts are caused by fluorescent dye that is not attached to the primers or is from unincorporated dye-labeled primers. These "color blips" can occur in any color and typically appear as low level, wide, and hill-like peaks above the baseline.
- These artifacts may or may not appear in all samples but are particularly apparent in samples with little or no DNA such as the negative controls.

3.8.1 Labels on dye artifacts should be removed in all samples.

3.9 Primer Front

- Primer front is a type of dye artifact referring to the common artifacts seen at low molecular weights which result from the detection of additional dye-labeled primer molecules not attached to the DNA in the sample. These artifacts usually appear as off-ladder and/or out-of-bin artifacts at the very beginning of a dye color.

3.9.1 Labels on primer front should be removed in all samples.

3.10 Spikes

- Generally, a spike is an electrophoresis artifact that looks like a vertical line or a peak and is usually present in all colors at the same location. They can be easily distinguished from true DNA peaks by looking at all other dye channels including size standard and/or other samples on the same electrophoresis run.
- Spikes may be caused by power surges, crystals, or air bubbles traveling past the laser detector window during electrophoresis.

3.10.1 Labels on spikes should be removed in all samples.

3.11 Non-specific artifacts

- Non-specific artifacts are labeled peaks caused by a not-otherwise-categorized technical problem or by non-specific priming in a megaplex reaction. These artifacts are usually easily recognized due to their low peak height and their position outside of the allele marker range.

3.11.1 Labels on non-specific artifacts should be removed in all samples.

3.12 Other common artifacts observed in Fusion data as reported in the PowerPlex® Fusion System for Use on the Applied Biosystems® Genetic Analyzers Technical Manual #TMD039 are included below:

3.12.1 Artifacts have been observed in samples in the following dye channels within the listed base pair ranges or at the following locations, where n represents 'number of bases' in the allelic peak. (For example, n-1 represents a peak 1 bp smaller than the allelic peak.)

Fluorescein	~58-59 bp[a], ~61-63 bp[a], 83-86 bp[a] ~63-68 bp	Amelogenin	n-1
		D3S1358	~88-112 bp[b]
		D1S1656	n+2
		D2S441	n-1
		D13S317	n+2
JOE	~62-67 bp[a] ~73-85 bp, ~214 bp, ~247 bp	D18S51	n+2
TMR-ET	~58-62 bp[a] ~66-72 bp, 172-176 bp	D21S11 D7S820 D5S818	n+2 n+2 n+2, n-8 to n-9[c]
CXR-ET	~175-183 bp	D12S391 D19S433	n+2, n-3 n+2

[a] These were also observed in samples without human genomic DNA present.
[b] Artifacts may fall in allelic bins.
[c] Low intensity peaks (~50-200rfu) in front of allele may represent DNA secondary structure.

3.12.2 Labels on these known artifacts **should be removed in all samples**.

4 Over-Saturation

4.1 Over-saturation may affect the **size standard**. All samples with a failing size standard caused by oversaturation should be diluted and rerun where possible. Fusion Direct exemplars should be recut for extraction and quantitation.

4.2 If an **evidence sample** has few peak(s) above or approaching 30,000 RFU but minimal artifacts, editing may be performed.

4.3 If an **evidence sample** has many saturated peaks >30,000 RFU, multiple split peaks, or requires excessive editing, this sample should be rerun at a dilution.

4.3.1 If there are many instances of pull-up, pull-down, and/or elevated baseline, the sample may be rerun at a dilution. All peak labels must be removed, and the electropherogram with deleted labels will be included in the file. Enter the appropriate re-run code and dilution factor.

4.4 **Exemplars or positive controls** with saturated peaks may be interpreted if artifacts are easily recognizable and can be edited.

4.4.1 Exemplars may be rerun at a dilution if needed, except for Fusion Direct samples.

4.4.1.1 Fusion Direct samples that cannot be edited should be recut for extraction and quantitation instead of being rerun.

5 Off-Ladder (OL) Alleles and Out-of-Bin (OB) Alleles

5.1 In the GeneMarker® HID software, OL refers to Off-Ladder alleles which are peaks outside of the marker range. OB refers to Out-of-Bin alleles which are peaks within the marker range but outside of a bin (microvariant alleles are frequently labeled as "OB").

5.2 A peak labeled as an OL or OB allele may be a true allele not represented in the allelic ladder or may be from migration of a true allele.

5.3 Examine the OL or OB allele closely in comparison to the ladder and other alleles present at that locus. If it is not at least one full base pair from an allele bin, it is more likely to be from migration than a true OL or OB.

5.4 If an OL or OB allele does not appear to be a true OL or OB allele (ex. if it is 0.55 bp away from the closest allelic ladder allele call) and appears to be a **migration of a true allele**, this allele may be assigned the appropriate allele call based on its measurement in comparison to the allelic ladder and other alleles present at that locus.

5.4.1 Compare the results of the run with other samples in the case and other amplifications of the sample for verification of the identity of the allele.

5.4.2 It may be helpful to examine the allelic peak in relation to the bins within GeneMarker® and the peak's position within the bin, especially in relation to other allelic peaks (and their positions in relation to their bins) within the same locus.

5.4.3 Allele frequency tables and/or online resources (ex. STRBase) may be referred to in order to determine if a particular microvariant has been seen before.

5.4.4 In Fusion, if the OB (or microvariant allele such as 10.4, 11.4, etc.) is observed at a **Penta** locus refer to section 5.7.

5.5 If an OL or OB allele or stutter peak does appear to be a **true OB/OL allele not represented in the allelic ladder**, assign the appropriate allele call based on its measurement in comparison to the allelic ladder. It may fall above or below the range of the allelic ladder but can still be assigned based on comparison.

5.5.1 The peak label shows the length in base pairs; this value can be used to determine the proper allele nomenclature by comparing this value to the allelic ladder and other peaks present within the sample at that locus. For example, at locus D12S391 in Fusion, a peak with base pair size of 143.5 could be resolved as a 16.1 allele if the ladder showed allele 16 at 142.5 base pairs and 17 at 146.5 base pairs.

5.6 In the rare event that you are unable to assign an allele call, **re-injection or re-amplification** of the sample may be attempted to confirm the allele call.

5.6.1 Reinjections may help resolve a migration issue, however, results from multiple injections of the same evidence sample cannot be combined for STRmix™ analysis. Only results from separate amplifications of an evidence sample may be combined for interpretation.

5.6.2 In some situations, if the OB/OL cannot be resolved, the locus may be ignored for STRmix™ analysis.

5.7 OB or labeled microvariant alleles (ex. 10.4, 11.4, etc.) observed at the Fusion **Penta** loci:

5.7.1 If the OB microvariant allele is not one full base pair from a non-microvariant allele (0.5-0.6 bp away from the closest allele ladder allele call) this may be an indication of a migration shift.

5.7.2 Conduct a closer inspection of the peak in relation to the allelic bin(s) within GeneMarker®. If the allele appears to be bordering a bin (or two), the OB or microvariant allele call may be due to a migration shift and not the presence of an actual microvariant allele.

5.7.2.1 For OB alleles that appear to be a migration shift where the allele can be assigned to a non-microvariant allele call, the allele call can be assigned.

5.7.2.2 For OB alleles where the allelic call cannot be assigned, the locus may be deemed INC.

5.7.2.2.1 For Fusion, the locus may be dropped for STRmix™ analysis.

5.7.3 When a labeled microvariant alleles (ex. 10.4, 11.4, etc.) is assigned a microvariant allele call in one replicate amplification but assigned a non-microvariant allele call in another amplification (ex. 11, 12, etc.), where the microvariant call appears to be due to a migration shift, the non-microvariant allele call may be assigned.

5.7.4 When a labeled microvariant allele (ex. 10.4, 11.4, etc.) persists through multiple amplifications but appears to be due to a migration shift, the locus may be deemed INC.

5.7.4.1 For Fusion, the locus may be dropped for STRmix™ analysis.

5.8 For Fusion, if an OB appears to be a slight **shift of a stutter peak** and is not filtered by the software refer to the Appendix for PowerPlex® Fusion Stutter to determine if the label on the OB should be removed or assigned an allele call.

5.9 If an allele or stutter is labeled as OB or OL for a sample, a copy of the allelic ladder for that run must be included within the case file.

6 Analysis and Troubleshooting

6.1 Refer to the GeneMarker v3.0 Operation Manual for analysis software instructions and example screenshots of controls.

6.2 Electrophoresis Controls

6.2.1 An electrophoresis run passes when there is **both** a passing ladder and a passing positive control present.

6.2.2 Allelic Ladder

6.2.2.1 The allelic ladder is checked for all expected alleles

6.2.2.2 If there is a passing ladder in the analysis set, check the positive control to determine if the run passes.

6.2.2.3 If OL/OB peaks are observed in one ladder, try re-analyzing with a different ladder in the run, if possible.

6.2.2.4 If there are no passing ladders in the analysis set, the run fails. See Electrophoresis Run Failure section 6.2.4.

6.2.3 Amplification Positive Control

6.2.3.1 The positive control is checked for all expected alleles.

6.2.3.2 There must be at least one amplification positive control present on each STR plate, including all reruns.

6.2.3.3 Positive control samples can be edited as described above in the manual editing of peak labels section. A positive control that does not generate a complete genotype or gives an incorrect genotype fails.

6.2.3.4 Electrophoresis run with failed positive control on a plate containing only one positive control

6.2.3.4.1 The run fails and no control or sample data from that run is usable. See Electrophoresis Run Failure section 6.2.4.

6.2.3.4.2 Indicate in LIMS that the positive control will be rerun. Re-aliquot or reinject the positive control with the entire amplification set and complete the STRControlReview column in LIMS.

6.2.3.4.3 If the rerun positive control fails, the amplification set fails See Electrophoresis Run Failure section 6.2.4.

6.2.3.5 Electrophoresis run with failed positive control on a plate containing multiple positive controls

6.2.3.5.1 If one of the positive controls on the run passes, the run may pass, and no deviation is needed.

6.2.3.5.2 The failed positive control may need to be rerun.

6.2.4 Electrophoresis **Run Failure**

6.2.4.1 When an **electrophoresis run fails, all controls** should still be reviewed and their STRReRun and STRControlReview columns in the data entry screen in LIMS need to be filled out before determining a retesting strategy. For example, if there is no size standard present for the whole plate, it should be re-aliquoted and not reinjected.

6.2.4.2 If the **run fails** for a non-passing ladder or positive control, **complete a deviation** against all output samples in the analysis set in LIMS to generate an electrophoresis failure report.

6.2.4.2.1 If the run fails due to an **Allelic Ladder** failure, it will read "Fail" and all other controls will read "No Data."

6.2.4.2.2 If the allelic ladders passes and the **Positive Control** fails, the Allelic Ladder will read "Pass", the Positive Control will read "Fail", and all others must be evaluated and filled out accordingly.

6.3 Negative Controls

6.3.1 If no peaks attributed to DNA are detected, the control passes.

6.3.1.1 Negative controls can be edited for artifacts that are by-products of the STR or amplification process such as dye artifacts, primer front, excessively elevated/noisy baseline (see section 3.7), and spikes. If these are so abundant that amplified DNA might be masked, the negative control can be rerun.

6.3.2 If peaks attributed to DNA are detected in an extraction negative, microcon negative and/or amplification negative control, **the first step is to rerun the control** by realiquot. The second step is to reamplify, if necessary. See below.

6.3.2.1 If an **extraction negative control fails** after being realiquoted and then fails after being reamplified, the extraction (or extraction set for a differential extraction) fails, and all associated samples should be marked Not Suitable for Comparison/Inconclusive and re-extracted where possible. Contact QA if an extraction negative fails.

6.3.2.2 If a **microcon negative control fails** after being realiquoted and then fails after being reamplified, the microcon batch fails, and all associated samples should be marked Not Suitable for Comparison/Inconclusive. Contact QA if a microcon negative control fails.

6.3.2.3 If an **amplification negative control fails** after being realiquoted, the amplification fails, and all associated samples should be reamplified.

6.4 **Retesting strategies** for samples and controls

6.4.1 In general, a control or sample that fails for size standard or other electrophoresis related issues (including extremely elevated/noisy baseline) should be **reinjected and/or realiquoted before being reamplified**.

6.4.1.1 Before **reinjecting a sample or control**, check that there is not a consistent issue for the whole plate; it may need to be realiquoted.

6.4.1.2 Before **scheduling reruns**, check that the sample or control has not already been rerun. If it has already been rerun, check for reasons the control or sample may have failed and consider additional options (especially if it has been repeatedly reinjected) such as running at a dilution, reamplifying or allowing the reporting analyst to make a case-specific decision (for an evidence sample).

6.4.2 **Fusion Direct** samples are not quantified, and the batches cannot be realiquoted.

6.4.2.1 If the Fusion Direct exemplar has complete data at fewer than 6 locations, evaluate the case scenario and, if needed, schedule the exemplar for a recut in another system with a quantification step (DNA IQ or EZ1).

6.4.2.2 If the exemplar has complete data at 6 or more locations, the reporting analyst may choose to use the data or recut in another system with a quantification step (DNA IQ or EZ1) depending on the case scenario.

6.4.2.3 If a Fusion Direct exemplar is overblown, evaluate the case scenario and, if needed, recut in another system with a quantification step (DNA IQ or EZ1) instead of reinjecting or reamplifying.

7 References

7.1 Office of the Chief Medical Examiner- NYC PowerPlex® Fusion System Amplification Kit on the Applied Biosystems 3130xl Genetic Analyzer using GeneMarker® HID Analysis Software Validation Report, August 29, 2016.

7.2 Internal Validation of STRmix™ V2.4 for Fusion NYC OCME

7.3 PowerPlex® Fusion System for Use on the Applied Biosystems® Genetic Analyzers Technical Manual #TMD039