Scientific Working Group on DNA Analysis Methods

Supplemental Information for the *SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories*

This document provides supplemental information for the *SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories* in the form of frequently asked questions (FAQs). Where applicable, FAQ responses include cross-references to the specific guideline in the parent document and references to published documents but may also use examples based on laboratory experiences.

FAQ-1: What type of cases can Y-STR testing be used for?

Cross-reference Guideline 1.1.1.

Y-STR testing is useful for criminal cases, particularly sexual assault cases, in which mixtures of male and female DNA are expected, and the amount of female DNA exceeds the amount of male DNA such that the autosomal profile of the male may not be observed.

In a mixture of male and female DNA, the proportion of the male DNA relative to the total DNA present in a sample can generally be predictive of the ability to detect a male DNA contributor using autosomal typing. This proportion can be determined through mixture studies conducted by a laboratory as part of the autosomal STR amplification system validation. Samples for which detection of a male contributor is not expected with autosomal typing based on quantification data should be conserved for Y-STR typing. Y-STR testing can be attempted for any sample where the total male DNA value from the male to female ratio is too high to obtain a usable male profile with autosomal STR testing. This determination should be based on internal

validation studies. Also, Y-STR can be attempted on samples with a limited overall quantity of male DNA.

If a ratio of total to male DNA (or female to male DNA) is not used to determine the suitability of autosomal and Y-STR typing, the laboratory should establish alternative strategies to maximize the potential for detecting male DNA. An example would be performing Y-STR testing directly on sexual assault evidence for which seminal fluid is detected, but sperm cells are not identified.

Y-STR testing is also useful for identification and paternity cases to associate two or more male paternal relatives. The decision to utilize Y-STR typing should be based on the context of the case, sample types, and any results of autosomal STR amplification systems if utilized.

FAQ-2: Do Y-STRs have duplications and/or deletions?

Cross-reference Guidelines 3.1, 3.2, and 7.1.1.

Duplicated sequences are present in the Y-chromosome and generate more than one allele when amplified with a single primer pair. These duplicated sequences are thus considered part of a multi-copy locus (Butler et al. 2005). A single duplication event may result in duplicated sequences that are far apart from one another (e.g., DYS385 a/b) or in close proximity (e.g., DYS437, DYS439 and/or DYS389I/II). The majority of duplications have alleles that differ in size by 1 repeat unit, while about 20% of duplications involve 2-, 3-, 4- and partial-repeat unit differences (Butler et al. 2005).

Triplications of Y-STR loci have been reported in YHRD, Release 64 (337,450 haplotypes), at DYS385a/b (17 different triplications), DYS19 (2 different triplications), DYF387S1 (22 different triplications), and DYS390, DYS439, DYS448, and DYS481 (1 triplication each).

Occasionally, null alleles can occur due to deletion of a portion of the Y-chromosome or a primer-binding site sequence variant can result in the failure to detect one or more Y-STR loci. The majority of null alleles occur at a single locus per haplotype. However, multi-locus null

alleles can occur due to deletion of loci that are within close proximity to one another (e.g., DYS437, DYS439 and DYS389I/II; DYS391 and DYS635).

FAQ-3: How does allelic drop-out impact Y-STRs?

Cross-reference Guideline 3.2.

Allelic drop-out is when an allele present in the sample does not produce a peak above the analytical threshold (AT). It is different from a null allele which is the inability to detect an individual's allele during DNA testing. Both situations produce no detected signal but have different causes.

The ability to assign a null allele at a locus of a single source profile with no detectable data depends on the ability to determine that drop-out has not happened at this locus. This may be done by inspecting the heights of peaks at this locus if duplicated (with respect to the within locus stochastic threshold - ST_W described in FAQ-8) or between loci (with respect to the stochastic threshold among loci - ST_B described in FAQ-9).

Low template amplifications could have drop-out of a duplicated allele at a Y-STR locus that is typically single-copy. However, a reasonable profile probability estimate will generally be obtained by searching the database using the observed allele. Accounting for an undetected second allele should not result in a practical difference in probability.

FAQ-4: What kit-specific artifacts may be encountered in Y-STR kits?

Cross-reference Guideline 4.1.

Stutter peaks due to the PCR process as well as other kit-specific artifacts may be observed in Y-STR kit results. Stutter peaks are typically reproducible and include back stutter, forward stutter, double-back stutter, and half-back stutter. Published works (e.g., Andersen et al. 2011 and Bright et al. 2014) offer guidance on the types and characteristics of these artifacts for the loci frequently used for Y-STR analysis. FAQ-6 describes an approach to account for stutter peaks in the interpretation of results.

Forensic STR kit manufacturers tend to design kits to maximize non-template addition of a 3´ terminal nucleotide by the DNA polymerase on the DNA fragment detected. Failure to attain complete terminal nucleotide addition may result in "split peaks," visualized as two peaks that are one base apart; these peaks are also often referred to as "-A/+A" peaks because adenine is the nucleotide frequently preferred for this phenomenon by Taq polymerase. One allele split into two peaks compromises the sensitivity of detection of that allele and can additionally complicate data interpretation. Laboratories should empirically determine quantitative and/or qualitative interpretation criteria for such peaks. Similar to stutter peaks, non-template nucleotide addition peaks may be characterized based on size and amplitude relative to an allelic peak. Due to the influence of primer design, locus-specific patterns within a Y-STR kit may also be useful for interpretation purposes.

Drop-in peaks have been characterized in the literature as the rare occurrence of spurious, non-reproducible allelic peaks, generally of up to two alleles in a profile (Gill et al. 2012), depending in part on the number of loci tested and the analytical threshold applied (Taylor et al. 2016b, Hansson and Gill 2017). In general, the rate of drop-in tends to increase as sensitivity is increased and is dependent on the amplification kit and detection system used by the laboratory. Drop-in is thought to result from fragments of cells that are introduced into the sample or extract from the laboratory environment or consumables used. Such alleles have been described as arising from DNA of different individuals, rather than a single contaminant which generally manifests as several alleles from one individual; however, these may not always be distinguishable (Moore et al. 2020). Drop-in is most easily detected in reagent blanks and negative amplification controls since no peaks are expected but can also occur in samples containing amplified product.

Additional kit-specific artifacts are routinely observed that do not appear to be the result of the polymerase stuttering, but rather non-specific primer binding, secondary/tertiary structure formation, and non-DNA dye related by-products ("dye blobs"). These artifacts may arise due to various reasons, such as excess quantities of female DNA, manufacturer specific attributes, and storage conditions. Published developmental validation and Y-STR kit manufacturer user guides and bulletins typically provide information on the characteristics of these artifacts. Such artifacts

should also be considered when evaluating the laboratory's internal validation and formulating interpretation criteria.

FAQ-5: What instrument-specific artifacts may be encountered in Y-STR kits?

Cross-reference Guideline 4.1.

The separation and detection technology utilized can present instrument-specific artifacts, in particular "pull-up" and "spikes."

Because the fluorescent dyes used for detection of amplified STR fragments overlap to varying degrees in their emission curves, multicomponent spectral deconvolution analysis is applied to the detected peaks. Pull-up is residual signal of one dye in another dye filter and manifests as a peak in a dye color other than the detected allelic peak. Pull-up peaks are generally small in amplitude, at the same or nearly the same data point, and will be therefore sized at a similar size as the allele. The shape of the pull-up peak may appear similar to a true DNA peak and may or may not be reproducible upon reinjection.

Capillary electrophoretic (CE) data may reveal sharp, narrow peaks often appearing in more than one dye channel, frequently referred to as spikes. This CE artifact is non-reproducible between injections of the same sample and generally occurs intermittently. Causes of spikes can include foreign particles (e.g., dust), air bubbles within the polymer, and transient current due to urea decomposition. Spikes are usually readily distinguishable from a true DNA peak due to morphology.

FAQ-6: How are stutter peak thresholds established with Y-STRs?

Cross-reference Guideline 4.1.

Stutter artifacts should be evaluated during the laboratory's internal validation in which thresholds, or other guidance as relevant, should be established for interpretation.

Back stutter thresholds should be set using at least a per locus basis. Because alleles with longer uninterrupted sequences are known to back stutter more (Bright et al. 2014), setting thresholds

on a per allele basis is preferred. However, this has not been the practice due to limited software capability. Historically, locus-based stutter thresholds have generally been set by calculating a locus stutter ratio mean and adding some number of standard deviations to the mean to create a locus threshold. Peaks detected below the threshold are attributed to stutter in single-source samples, and in the case of mixtures, such peaks may be stutter and/or minor contributor alleles. While the study of stutter ratios is hampered by the fact that these artifacts are small in height, which may result in only the larger values being detected, this can be alleviated by analyzing validation samples using a very low RFU threshold (regardless of the final analytical threshold used for casework interpretation).

As an example, consider a dataset of at least 100 single source profiles that encompasses a wide range of alleles for each locus. Following analysis using a low RFU analytical threshold(s), calculate the back and forward stutter percentages for every allele in the dataset where there is no interference by other stutter or parent peaks. Such interference would typically occur at duplicated or multi-copy loci with alleles that are one repeat unit apart (e.g., given alleles 10 and 11, the height of the 10 allele is accentuated by the stutter peak of the 11 allele). It is recommended that overloaded/overblown samples not be used.

The stutter threshold z is determined by $z = \bar{x} + 3sd$ where $\bar{x}$ is the average and sd is the standard deviation of the stutter ratio in the sample set. Note that in using this threshold, a statistically predictable portion of data is expected to exceed the threshold, especially by stutters from the larger alleles at a given locus.

Back stutter (SR): A peak is assigned as *not* back stutter if $SR = \frac{O_{a-1}}{O_a} > z$

Where: $O_{a\text{-}1}$ is the height of the peak in a position one repeat shorter than allele a (back stutter position). O_a is the height of the peak at position a. z is the back stutter ratio threshold.

A peak in a back stutter position above this threshold may be due in part to a minor contributor allele and should be considered as such when relevant.

Forward stutter (FS): A peak is assigned as *not* forward stutter if $FS = \frac{O_{a+1}}{O_a} > y$

Where: O_a+1 is the height of the peak in a position one repeat longer than *a* (forward stutter position). O_a is the height of the peak at position *a*. *y* is the forward stutter ratio threshold.

A peak in a forward stutter position above this threshold may be due in part to a minor contributor allele and should be considered as such when relevant.

FAQ-7: Is a stochastic threshold applicable to Y-STR typing?

Cross-reference Guideline 5.1.

The stochastic threshold (ST) is the RFU value above which it is reasonable to assume that, at a given locus, allelic drop-out of a sister allele has not occurred. For multi-copy loci such as DYS385 and DYF387S1, a stochastic threshold is useful as it serves to alert the DNA analyst whether all of the DNA typing information has likely been detected at these loci for a given sample. This is referred to as the within-locus ST_W (see FAQ-8).

Furthermore, in order to interpret some mixtures, it is necessary to consider a stochastic threshold which informs the DNA analyst that drop-out is possible at a single-copy locus based on adjacent loci or the whole profile. This is referred to as the between-locus ST_B (see FAQ-9).

FAQ-8: How is the within locus stochastic threshold (ST_W) determined?

Cross-reference Guideline 5.1.

ST_w can be established by assessing peak height ratios across any multi-copy locus, as well as any single-copy loci with duplicate alleles, in a dilution series of DNA amplified in replicate. Methods to determine the probability of drop-out at multi-copy loci are described by Tvedebrink et al. (2009, 2012a) and Buckleton et al. (2014). The stochastic threshold may be set using a probability of drop-out or other methods determined by the laboratory to minimize the risk of misinterpreting the profile.

FAQ-9: How is the between-locus stochastic threshold (ST_B) determined?

Cross-reference Guideline 5.1.

The between locus ST_B may be determined using a number of methods which largely start with the analysis of a DNA dilution series focusing on the point at which drop-out begins to occur. An example method is outlined below.

Logistic regression: A sample set of approximately 100 single source amplifications are diluted so that they span the stochastic range. Logistic regression using one of the models described by Tvedebrink et al. (2009, 2012a) and Buckleton et al. (2014) should be undertaken. All of these methods need the input of a parameter α which is the probability that a specific allele has dropped out. The probability α was modelled by Tvedebrink et al. (2009) as being dependent on the stochastic threshold (ST), with the model being expressed as a logistic regression:

$$\ln\left[\frac{1-\alpha}{\alpha}\right] = \beta_0 + \beta_1 \ln(ST)$$

Here β_0 is the intercept and β_1 is the slope of this regression, and can be estimated from data of observed dropout for known ST. Buckleton et al. (2014) reversed the relationship to give:

$$ST = e^{\frac{\left(\ln\frac{1-\alpha}{\alpha}\right)-\beta_0}{\beta_1}}$$

and used this to predict ST values for a specified value of α. Using a value of $\alpha = 0.005$ means setting the ST_B at a probability of drop-out of 0.5%. Alternate values of α could be explored by the laboratory as appropriate to the data in the study and the risk tolerance of the laboratory.

FAQ-10: Are peak height ratios applicable to Y-STR loci?

Cross-reference Guideline 6.1.

The peak height ratio (PHR) concept frequently used for autosomal heterozygous STR loci may also be applied to multi-copy Y-STR loci (e.g., DYS385a/b and DYF387S1), particularly for use in deconvoluting mixtures (cross-reference FAQ-11). A peak height ratio is the relative proportion of two alleles at a given locus, as determined by dividing the peak height of an allele with a lower relative fluorescence unit (RFU) value by the peak height of an allele with a higher RFU value, and then multiplying this value by 100 to express the PHR as a percentage. As the amount of DNA template in a PCR reaction is reduced, PHRs exhibit greater variation due to

stochastic effects. Note that PHRs are only applicable to allelic peaks that meet or exceed the stochastic threshold. PHR interpretation criteria should be established based on the laboratory's internal validation data from single-source samples spanning the range of input DNA template.

FAQ-11: How are major and minor contributors assigned?

Cross-reference Guidelines 7.1, 7.1.1, and 7.2.

Major and minor contributors may be assigned using criteria determined through validation that allow for the resolution of major and minor contributor haplotypes at some or all loci in a Y-STR mixture. These criteria may include but are not limited to peak height ratios, peak heights, within and between locus stochastic thresholds, and mixture proportions in consideration of duplications or other genetic anomalies. In setting these criteria, validations should include a variety of two-person mixture amplifications from known contributors of sufficient number to address the variability in amplification and electrophoresis results across the dynamic range of the detection platform and should span the range of input template for which the rules will be applied. There should be dense sampling around the mixture ratio likely to be key to the designation of major and minor contributor alleles. For example, the sample set could consist of two-person mixtures of varying ratios (e.g., 1:20, 1:10, 1:5, 1:4, 1:3, 1:2, and 1:1) amplified in a dilution series consisting of different input amounts (e.g., 1.00ng, 0.75ng, 0.5ng, 0.25ng, 0.13ng, 0.06ng, 0.03ng).

The validation data should be analyzed to determine when major and minor contributors can be readily distinguished. Laboratories may set mixture ratio thresholds (e.g., a mixture ratio of 1:5 must be met before proceeding with profile deconvolution), and/or peak height and peak height ratio thresholds (e.g., both peaks must be above 300RFU and exhibit a peak height ratio of at least 60% before assigning major and minor alleles). These thresholds should be set to avoid the incorrect assignment of major and minor contributor haplotype alleles in circumstances where the minor contributor yields an allele of greater RFU than the major (inverted major and minor peak heights), when the minor contributor allele is detected above the analytical threshold where the major contributor allele is not (inverted drop-out), or when a minor contributor allele is

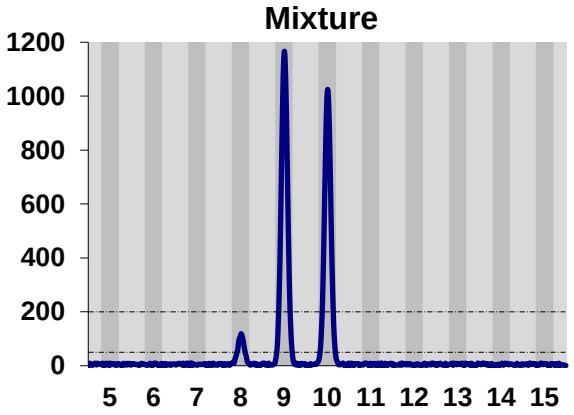
shared with the major contributor as opposed to having dropped out. A laboratory may choose to take a probabilistic approach to setting these thresholds as described in Taylor et al. (2016a).

FAQ-12: What are indistinguishable Y-STR mixtures?

Cross-reference Guidelines 7.2, 8.3, and 8.4.

By definition, an indistinguishable Y-STR mixture is a DNA mixture in which mixture ratios, presented as relative peak height ratios, are insufficient to attribute alleles to individual male contributor(s). When multiple contributor genotypes cannot be distinguished because of similar contribution levels, the sample is an indistinguishable mixture. Mixture profiles deemed indistinguishable may still be interpretable. Individual males may still be included or excluded as possible contributors to an indistinguishable mixture.

As an example, the following single-locus electropherogram was obtained from the vaginal swab of a victim of a sexual assault. The victim also reported recent consensual intercourse with a known individual.



Bin	RFU
8	109
9	1157
10	1021

The above example uses an analytical threshold of 50 RFU and a stutter threshold of 10%. Based upon this locus, the mixture may be assumed to be from two individuals contributing DNA of roughly the same proportion (the 9 and the 10 alleles). The 8 peak is below the stutter threshold and is therefore assumed to be stutter in this example. Since the 9 and 10 alleles are of similar signals, it cannot be determined which allele was contributed by each of the two different contributors. Thus 9 and 10 are indistinguishable. If the consensual partner's genotype is a 12,

the consensual partner is excluded as a contributor to the mixture obtained. If the consensual partner's genotype is a 9, the indistinguishable mixture may be further refined into a deduced haplotype foreign to the consensual partner by assigning the 10 to the deduced contributor.

FAQ-13: How does a laboratory report indistinguishable mixtures?

Cross-reference Guideline 8.4.1.

The statistical subject matter experts have not yet reached consensus on a statistical approach for estimating the occurrence of a combination of haplotypes in a population.

A laboratory choosing to report inclusionary Y-STR typing results from indistinguishable mixtures that are determined to be relevant in the context of a case must perform statistical analysis in support of any inclusion. The statistical method employed must be supported by empirical data and internal validation.

FAQ-14: Why has the U.S. forensic community switched from using the U.S. Y-STR Database to YHRD?

Cross-reference Guideline 9.1.

The U.S. Y-STR Database was managed by the National Center for Forensic Science at the University of Central Florida since 2007 through funding from the National Institute of Justice (Ballantyne et al. 2006). To mitigate encumbrances in the administration of resources and to ensure long-term operational stability, the U.S. Y-STR Database haplotypes were permanently transferred to the Y- Chromosome Haplotype Reference Database (YHRD, Willuweit and Roewer 2007).

An announcement titled *Notice to U.S. Forensic Laboratories on the status of the U.S. Y-STR Database* was released in November 2018 and made available through various websites (swgdam.org, yhrd.org, and http://usystrdatabase.org). For more information, refer to the following presentation: Transition from U.S. Y-STR Database to YHRD (https://docs.wixstatic.com/ugd/4344b0_7c1fe1eaa2e04d48be9f4ea3101c6e3e.pdf).

FAQ-15: Can searches in YHRD accommodate haplotypes with allelic drop-out?

Cross-reference Guideline 9.1.

For single copy Y-STR loci exhibiting allelic drop-out, the locus may either be dropped from a search in YHRD or, if a null allele is suspected, it may be searched as a "0" (see FAQ-3 for more information on drop-out *v.* null alleles).

At the time of publication of this document, YHRD is not able to accommodate a multi-copy Y-STR locus with allelic drop-out. Any locus with suspected allelic drop-out should be excluded from the search in the database.

FAQ-16: If a profile or match probability was calculated using the U.S. Y-STR Database, should a laboratory recalculate a profile or match probability using the YHRD?

Cross-reference Guideline 9.2.4.

Profile and match probability estimates between the U.S. Y-STR Database and YHRD [National Database (with subpopulations) United States] are expected to be similar and within the same order of magnitude since many of the same sets of Y-STR data from the U.S. have historically been contributed to both databases by researchers, commercial entities, and practitioners. The impact on profile and match probabilities as a result of switching to YHRD from the U.S. Y-STR Database should be minimal. Previous statistics generated using the U.S. Y-STR Database are valid and should not require a recalculation with YHRD. However, as the database size of YHRD increases, the differences between profile and match probability calculations may become significant. Each laboratory should decide if and/or when a previously calculated profile or match probability using the U.S. Y-STR Database needs to be recalculated using YHRD.

FAQ-17: How can a laboratory address new releases of the YHRD?

Cross-reference Guideline 9.2.

With each new release of YHRD, a laboratory should review the context of the database changes to sizes and/or relevant populations used for statistics, determine if the updates affect their

procedures for searching and/or calculating profile and match probabilities, and as appropriate, perform any necessary validations or performance checks for statistical software per the FBI Quality Assurance Standards. For example, a laboratory can identify an appropriate set of Y-STR haplotypes to evaluate subsequent new releases of YHRD and then compare the search results and calculations to the results from the previous release.

FAQ-18: What is a reduced locus search?

Cross-reference Guidelines 9.2.1, 9.2.1.1, and 9.2.1.2.

Reduced locus searches are performed to identify the most informative result from the search of an evidentiary haplotype against a population database when using counting method-based approaches (recommendation 9.2.2 and 9.2.3). This search method addresses the counting method paradox in which the search of a newer multiplex with a more discriminating set of loci seemingly yields a higher estimated profile probability when searching a database. The logical discrepancy is due to the database containing fewer reference haplotypes with the newer multiplex than with an older, less discriminating multiplex that had fewer loci. By reducing the search to a smaller locus set (e.g., searching just the Yfiler loci when the evidentiary profile was based upon the PowerPlex Y23 loci), the larger database size of that set can be incorporated into the sample match frequency.

When no matching haplotypes are observed regardless of locus set, the smaller locus set will always provide a lower frequency due to the larger database size. Even though the search was performed with a less discriminating set of loci, a profile with no observations in the database using the smaller locus set could not "match" if you added additional loci. The resulting sample frequency better represents the discrimination potential of the full Y-STR profile.

When the reduced locus search yields "matches" at the reduced locus-count, additional searches in YHRD are performed in a manner that excludes any "matches" that would have been non-matches had more of the evidence profile been searched. For example, a "match" between the evidence and a population database sample at the 8 loci in YHRD's minimal haplotype would not be included as a match for statistical purposes if the profiles differed at any additional loci

for which they both had information. The most informative sample frequency is best represented by whichever is lower, the final reduced locus frequency or the original full profile frequency.

FAQ-19: How is a reduced locus search performed using YHRD?

Cross-reference Guideline 9.2.1.2.

Below is a diagram of the reduced locus search process when searching with a Yfiler Plus profile and reducing the loci all the way to the minimal haplotype.

NOTES:

- The search process always begins at the locus set best representing the full evidentiary profile. For example, when casework or missing persons analyses were performed using the Yfiler kit, the process would start there. No searches would be performed using the Yfiler Plus setting in yhrd.org.

- For profiles generated using the PowerPlex Y23 kit, searches would instead start with that locus set.
- Only when the profile has been typed with both PowerPlex Y23 and Yfiler would it be beneficial to start with the maximal locus set.

The following is a demonstration of the calculations performed when a reduced locus search "matches" a database profile. This example will focus on the Yfiler Plus and Yfiler locus sets, but the concepts are the same for the remainder of the locus set pairs. It is noted that the x and n values are just examples and are not from any actual YHRD search.

- Search a Yfiler Plus profile against database samples with full Yfiler Plus profiles:

$$x_{YFP} = 1$$

$$n_{YFP} = 500$$

$$x_{YFP}/n_{YFP} = 1/500$$

- Reduce the search profile to just the loci found in Yfiler, and search it against database samples with full Yfiler profiles (of which the Yfiler Plus samples are a subset):

$$x_{YF} = 5$$

$$n_{YF} = 7000$$

$$x_{YF}/n_{YF} = 1/1400$$

- To examine how many of those 5 Yfiler profiles were already tested in the subset of Yfiler Plus database profiles, either as inclusions or exclusions, search Yfiler Plus database profiles with the reduced Yfiler profile.

$$x_{YFP\ using\ YF} = 4$$

  - Since $x_{YFP} = 1$, this means that 3 of the 4 $x_{YFP\ using\ YF}$ profiles must have been exclusions at one or more of the loci found in Yfiler Plus but not Yfiler.
  - Since $x_{YFP\ using\ YF} = 4$, this means that 1 of the 5 x_{YF} profiles must not have been typed for the additional Yfiler Plus loci.

$$x_{YF\ only} = x_{YF} - x_{YFP\ using\ YF} = 1$$

- Combined, there were 2 Yfiler profiles that matched the search profile at all of the available loci.

$$x_{YFP} = 1$$

$$x_{YF\ only} = 1$$

$$x_{YF\ total} = x_{YFP} + x_{YF\ only} = 2$$

$$n_{YF} = 7000$$

$$x_{YF\ total}/n_{YF} = 1/3500$$

- In this fictitious example, the reduced locus search process would capture 7X more evidential weight than when just searching the Yfiler Plus database samples, and 2.5X more evidential weight than just searching the Yfiler database samples.

For labs that report 95% Upper Confidence Interval (UCI), the locus set to use for reporting would be the set with the lowest UCI. While the YHRD provides the 95% UCI for each search, the 95% UCI will need to be manually calculated for the reduced locus search process. The Excel formula below applies the beta distribution to provide a highly accurate approximation. It is functionally equivalent to the version described at https://sigmazone.com/binomial-confidence-intervals/.

BETAINV (0.95, x + 1, n – x).

FAQ-20: How can a laboratory conduct a performance check for reduced locus searches using YHRD?

Cross-reference Guideline 9.2.1.2.

The profiles of matching and non-matching haplotypes are not available to view when searching YHRD. Therefore, just as is the case with YHRD searches when using all of the available loci, reduced locus search results cannot be independently verified. There are certain logical assessments that could be tested. For example, the same or fewer matches should be obtained when searching higher locus counts.

FAQ-21: How can a laboratory report the use of a reduced locus search?

Cross-reference Guideline 9.2.1.2.

An example of report wording is provided below if a laboratory elects to include the use of a reduced locus search within a report.

The evidence haplotype was compared to [n_{RLS}] haplotypes in the YHRD database. There were [x_{RLS}] haplotypes that could not be differentiated, for a sample frequency of [x_{RLS}/n_{RLS}].

Where

x_{RLS} The number of haplotypes at the reduced locus count that were not shown to be different at a higher locus count

n_{RLS} The number of haplotypes at the reduced locus count

Both x_{RLS} and n_{RLS} can be augmented with one extra observation, per recommendation 9.2.2.2.

FAQ-22: What are the differences between the Counting Method and the Augmented Counting Method?

Cross-reference Guidelines 9.2.2, 9.2.2.1, and 9.2.2.2.

The counting method is the simplest method to report the prevalence of haplotypes in a sampling of a population.

Estimate of the population proportion:

$$p = x / n$$

where x is equal to the number of times the haplotype is observed in a database containing n number of haplotypes. For example, if a haplotype has been observed twice in a database of n = 2000, the frequency of that haplotype will be: 2/2000 = 0.001.

In most Y-STR databases, the majority of haplotypes are only observed once in the database. Therefore, when a haplotype is searched in a database, the number of haplotypes that match is often zero. Using the counting method, the estimate of the population proportion would be zero (unique). However, further sampling of the population may identify other instances of the haplotype.

To compensate for zero matches in the database, the use of the augmented counting method by adding the observed haplotype to both x (in the numerator) and n (in the denominator) is

recommended by the DNA Commission of the International Society for Forensic Genetics (ISFG) (Roewer et al. 2020).

$$p = (x + 1) \,/\, (n + 1)$$

FAQ-23: What is the Clopper and Pearson 95% confidence interval?

Cross-reference Guideline 9.2.3.

A confidence interval proposes a range of values (an interval) having a confidence level (e.g., 95%) that the true parameter (the estimate of the population proportion of a particular haplotype) is within this proposed range. Previous methods such as the normal approximation of the binomial distribution have been replaced with the Clopper and Pearson exact method (Clopper and Pearson, 1934). The interval is based on the cumulative probabilities of the binomial distribution rather than an approximation of the interval. The Clopper and Pearson 95% confidence interval can be determined from either the counting method or the augmented counting method.

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

FAQ-24: What is the difference between a profile probability and a match probability?

Cross-reference Guidelines 9.2.2, 9.2.3, 9.2.4.

For Y-STR haplotypes, the strength of the evidence is presented when there is a match between the evidence and the reference haplotype. The presentation of a probability (profile probability or match probability) depends upon the question being asked. For a profile probability, the relevant question is rather straightforward: How rare (or common) is this haplotype in the population? The answer can be easily determined by searching the haplotype in a relevant database and presenting the number of times the haplotype was observed in that database. Termed, "the counting method" this frequency can be accompanied by a confidence interval such as the Clopper and Pearson exact method.

A match probability addresses a different question from having a match between the evidence haplotype and the reference haplotype. Here, the relevant question is, "the haplotype has been observed already in the evidence – what is the probability of observing an unrelated individual in the relevant population with the same haplotype as the reference?" A match probability is commonly incorporated into a likelihood ratio (LR), a comparison of two conditional probabilities to explain the match. The first conditional probability (in the numerator) is the probability of observing the evidence if the person of interest (POI) is the contributor to the profile versus the second conditional probability (in the denominator) of observing the evidence if a random, unrelated individual in the relevant population is the contributor to the profile. Note that the LR is not a probability, but a ratio of two mutually exclusive conditional probabilities, and is not presented as a "1 in number" frequency like a profile probability (See FAQ 33). The DNA Commission of the ISFG (Roewer et al. 2020) is a resource for more information on profile and match probabilities.

FAQ-25: What methods can be used to calculate a match probability?

Cross-reference Guideline 9.2.4.

The lack of independence among Y-STR loci makes it difficult to quantify the strength of matching Y-STR profiles. It is known, however, that it becomes increasingly unlikely that two different unrelated men share the same Y-STR profile as more loci are included in the profile.

There are several methods to calculate a match probability. They include use of theta, the kappa method, and the discrete Laplace method. See FAQs below for further details. SWGDAM urges the continued development and publication of these and related approaches that can offer guidance to forensic practitioners.

FAQ-26: How is theta used to assign a match probability?

Cross-reference Guideline 9.2.4.

Theta is a correction factor to account for substructure within a population and is most often used when calculating match probabilities of diploid autosomal markers. Using theta for haploid

markers is still a topic of much debate. Haploid markers are not a primary means of identification and are most powerful when used for exclusionary purposes. The use of theta was outlined in the 2014 SWGDAM *Interpretation Guidelines for Y-Chromosome STR Typing* but without guidance for population structure within a single ethnic group.

The subpopulation correction using theta (Buckleton et al. 2011, Weir and Goudet, 2017) gives an estimate $\hat{p}_i$ of the probability for haplotype *i* of

$$\hat{p}_i = \theta + (1 - \theta)\left(\frac{x}{n}\right)$$

This is used when x is the observed number of haplotype *i* in a database of n profiles. If x is zero, then θ serves as a lower bound on the estimate. Laboratories should establish the value of θ they wish to use, using published values. If the database is for a particular ethnicity, then the θ value for that ethnicity should be used. It is possible to estimate θ with data from populations within an ethnicity. The estimate is (Mw-Mb)/(1-Mb) where Mw is the proportion of matching pairs of haplotypes among all pairs within one population, averaged over populations, and Mb is the proportion of matching pairs of haplotypes, one from each of two populations, averaged over pairs of populations.

There is not currently a publication for Y-STR θ values from a world-wide survey as there is for autosomal STRs (Buckleton et al. 2016). Such a publication is forthcoming. It is likely that values of 10^{-4} or less are appropriate for 15 or more Y-STR loci, and 10^{-5} or less are appropriate for 20 or more Y-STR loci.

FAQ-27: How is the application of theta to match probabilities different in YHRD as compared to how it was applied in the U.S. Y-STR Database?

Cross-reference Guideline 9.2.4.

Important differences existed between the theta-corrected match probabilities reported in each of the databases. YHRD limits theta-corrections to profiles with fewer than 23 loci, regardless of which loci are searched and the multiplex selected, while the U.S. Y-STR Database applied theta to all searches of any number of loci as long as the Yfiler Plus kit locus order was not selected

for profile entry. Although the same theta values described in Appendix 1 of the 2014 SWGDAM *Interpretation Guidelines for Y-Chromosome STR Typing* have been applied by both databases, U.S. Y-STR Database separated the theta-corrected match probabilities by major population group (African American, Asian, Caucasian, Hispanic, and Native American), while YHRD combines all populations (without and, where data exists, with the Native American population) to calculate the "Overall" theta-corrected match probabilities. Relevant case information regarding the pool of possible alternate contributors may be used as a guide when selecting between YHRD match probabilities that exclude or include the Native American data. If desired, the population-level match probabilities that are not supplied by YHRD can be calculated outside of that website using the YHRD search results for each population, Eq. 3 from the 2014 SWGDAM *Interpretation Guidelines for Y-Chromosome STR Typing*, and theta values from Appendix 1 of the 2014 SWGDAM *Interpretation Guidelines for Y-Chromosome STR Typing*.

FAQ-28: How is the kappa method used to calculate a match probability?

Cross-reference Guideline 9.2.4.

The term kappa (κ) denotes the fraction of haplotypes that have been observed only once, i.e., singletons, in the database augmented by *x*. As defined here, this gives a match probability (Brenner 2010). If the count of the POI's haplotype x_p in the database of D individuals is assigned as C_{xp}, then

$$\hat{p}_x = \frac{(C+1)(1-\kappa)}{D+1}$$

FAQ-29: How is the discrete Laplace method used to calculate a match probability?

Cross-reference Guideline 9.2.4.

The discrete Laplace method is a statistical model (Andersen et al. 2013a and Andersen et al. 2013b) that can be used to estimate population frequencies of Y-STR haplotypes based on a reference database. An estimated population frequency can serve as a match probability when the reference database is a random sample from the suspect population. The discrete Laplace

method assumes a number of latent clusters with shared ancestry exists, each of which is represented by a central haplotype. The haplotypes in the population are then spread around these central haplotypes (caused by neutral stepwise mutations). This is recommended by the DNA Commission of the ISFG (Roewer et al. 2020).

To estimate a haplotype frequency for a given Metapopulation (e.g.,"African American" or "Native American“) using the discrete Laplace method, one can use the "Metapopulation“ feature of the standard YHRD search.

Additional guidance on the discrete Laplace method is provided by Mikkel Meyer Andersen and David Balding and can be accessed at https://mikl.dk/ytalks/.

FAQ-30: Can the likelihood ratio calculated in the pedigree search function of CODIS be reported?

Cross-reference Guideline 10.1.

The Pedigree Search function of the CODIS software generates a combined likelihood ratio for autosomal, mtDNA and Y-STR results for missing person searches to rank potential candidates. This combined value is solely designed to generate a ranked list of potential investigative leads and is not appropriate for reporting purposes.

FAQ-31: Can the match probabilities from Y-STR, mtDNA, and/or autosomal STRs be combined into a match probability?

Cross-reference Guideline 10.1.

If there is reasonable expectation of genetic independence, match probabilities from any combination of mtDNA, Y-STR and/or autosomal STRs may be combined; however, the statistical subject matter experts have not yet reached consensus at this time regarding the suitability of combining the likelihood ratios from lineage markers or combining an autosomal likelihood ratio with one or both lineage markers. Additional research examining the independence of lineage markers and autosomal markers is needed.

FAQ-32: What is an example of wording for reporting a frequency with an upper confidence interval?

Cross-reference Guideline 9.2.3.

Example when all population groups are reported separately:

The (major/minor) Y-chromosomal DNA profile obtained from item # has been observed in the population groups as follows:

U.S. Population Group	Observations	Database Size	Upper Limit Frequency (%)	Which equates to approximately
African American				
Caucasian				
Hispanic				

Therefore, we would not expect to see the Y-STR profile obtained from item # more than once in *X* African Americans, once in *X* Caucasians or once in *X* Hispanics.

Example when all population groups are reported separately with exclusions:

The Y-STR profile obtained from item # was compared to the YHRD (yhrd.org, release #). The frequency of occurrence of this profile in the YHRD of the major U.S. population groups is given below. Furthermore, the frequency of this profile was determined by applying the 95% Upper Confidence Interval. The inclusion and exclusion probabilities in the major U.S. population groups are as follows:

African American: Inclusion 1 in XX; XX% excluded
Caucasian: Inclusion 1 in XX: XX% excluded
Hispanic: Inclusion 1 in XX; XX% excluded

U.S. Population Database	Frequency
African American	XX in XXXX
Caucasian	XX in XXXX

Hispanic	XX in XXXX

Example when all population groups are reported together:

In a search of XXXX U.S. male Y-STR profiles, this profile was observed X times. Applying a statistical confidence interval of 95%, this profile is not expected to occur more frequently than 1 in XXX U.S. males.

FAQ-33: What is an example of wording for reporting a likelihood ratio (LR)?

Cross-reference Guideline 9.2.5.

Note: The verbal scales used in direct comparison Examples 1 and 2 are from the Recommendations of the SWGDAM Ad Hoc Working Group on Genotyping Results Reported as Likelihood Ratios document found at swgdam.org.

Examples of direct comparison:

Example 1:

The Y-STR typing results from item # were interpreted as originating from one individual. The Y-STR typing results from item # are 150 times more likely if NAME is the contributor than if an unknown, unrelated male is the contributor.[1]

Person of Interest (POI)	Likelihood Ratio (LR)[2]	Level of Support[3]
NAME	150	Moderate Support for Inclusion

Example 2:

The Y-STR typing results from item # were interpreted as originating from two individuals. The major contributor profile from item # is 150 times more likely if NAME is the major contributor than if an unknown, unrelated male is the major contributor.[1]

Likelihood Ratio (LR)[2]	Level of Support[3]
150	Moderate Support for Inclusion

Reports can utilize the following statements to further clarify comparisons:

[1] Barring mutation, any male relative within the same paternal lineage has the same Y-STR profile and would also be expected to be included/excluded as a possible contributor.

[2] The likelihood ratio is a statistical approach that compares the probabilities of observing the DNA results under two alternative propositions. Calculations were performed using the African American, Caucasian, and Hispanic populations in the Y Chromosome Haplotype Reference Database (release R63). The lowest calculated likelihood ratio is reported.

[3] These likelihood ratio ranges provide the following support for Y-STR conclusions:

Likelihood Ratios:	Qualitative Equivalent:
1	Uninformative
2-99	Limited Support
100-9,999	Moderate Support
10,000-999,999	Strong Support
≥1,000,000	Very Strong Support

Examples of familial comparison:

Example 1:

Information provided by the contributor identifies NAME1 as the biological brother of NAME2. The Y-STR typing results obtained from item # and NAME1 are the same; therefore, item # could have originated from NAME2. These results are X times more likely if item # is from NAME2 than if item # is from an unknown, unrelated male.

Example 2:

Based on the genetic results, the most conservative estimate indicates that these data are *X* times more likely to be observed under the scenario that the unidentified human remains originate from a paternal relative of NAME as compared to originating from an unrelated male from the general population**.**

Example 3:

The Y-chromosomal DNA profile obtained from the femur (*ITEM X)* matches the Y-chromosomal DNA profile obtained from NAME (*ITEM Y).* Kinship analysis calculations for a paternal relative relationship for the Y-chromosomal profiles obtained from *ITEM X* and Y are as follows:

U.S. Population Group	Likelihood Ratio
African American	
Caucasian	
Hispanic	
Other population groups as appropriate	

Based on the above results, the most conservative estimate indicates that the Y-chromosomal DNA profile obtained from the femur is X times more likely to be observed in a paternal relative of NAME than in someone unrelated to NAME.

References:

Andersen, M.M., Olofsson, J.K., Mogensen H.S., Eriksen, P.S. and Morling, N. (2011) *Estimating stutter rates for Y-STR alleles*. FSI Genetics Supplemental Ser 3:e192-e193.

Andersen, M.M., Eriksen, P.S. and Morling, N. (2013a) *The discrete Laplace exponential family and estimation of Y-STR haplotype frequencies.* J Theor. Biol 329: 39-51.

Andersen, M.M., Caliebe, A., Jochens, A., Willuweit, S. and Krawczak, M. (2013b) *Estimating trace-suspect match probabilities for singleton Y-STR haplotypes using coalescent theory.* FSI Genetics 7:264-271.

Ballantyne, J., Fatolitis, L. and Rower, L. (2006) *Creating and Managing Effective Y-STR Databases.* Profiles in DNA 9(2) 10-13.

Brenner, C.H. (2010) *Fundamental problem of forensic mathematics – the evidential valude of a rare haplotype.* FSI Genetics 4:281-291.

Brenner, C.H. (2014) *Understanding Y haplotype matching probability.* FSI Genetics 8:233-243.

Bright, J-A., Curran, J.M. and Buckleton, J.S. (2014) *Modelling PowerPlex® Y stutter and artefacts.* FSI Genetics 126-136.

Buckleton J, Kelly, H. Bright, J-A., Taylor, D, Tvedebrink, T. and Curran, J.M. (2014) *Utilising allelic dropout probabilities estimated by logistic regression in casework.* FSI Genetics 9:9-11.

Buckleton, J.S, Bright J-A. and Taylor, D. (2016) *Forensic DNA Evidence Interpretation*, CRC Press, Taylor and Francis Group, 2nd edition.

Buckleton, J., Kelly, H. Bright, J-A., Taylor, D., Tvedebrink, T. and Curran, J.M. (2014) *Utilising allelic dropout probabilities estimated by logistic regression in casework.* FSI Genetics 9:9-11.

Buckleton J.S., Krawczak, M., Weir B.S. (2011) *The interpretation of lineage markers in forensic DNA testing.* FSI Genetics 5:78-83.

Buckleton J.S., Curran J.M., Goudet J., Taylor D., Thiery A., Weir B.S. (2016) *Population-specific Fst values: A worldwide survey.* FSI Genetics 23:91-100.

Butler, J.M., Decker, A.E., Kline, M.C. and Vallone, P.M. (2005) *Chromosomal Duplications Along the Y-Chromosome and Their Potential Impact on Y-STR Interpretation.* J Forensic Sci 50:853-859.

Clopper, C.J. and Pearson, E.S. (1934) *The use of confidence or fiducial intervals illustrated in the case of the binomial.* Biometrika 26:404-13.

Gill, P., Gusmao, L., Haned, H., Mayre, W.R., Morling, N., Parson, W., Prieto, L., Prinz, M., Schneider, P., and Weir, B.S. (2012) *DNA Commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include drop-out and/or drop-in using probabilistic methods.* FSI Genetics 6:679-688.

Hansson, O. and Gill, P. (2017) *Characterisation of artefacts and drop-in events using STR-validator and single-cell analysis.* FSI Genetics 30: 57-65.

Moore, D., Clayton, R. and Thomson, J. (2020) *A comprehensive study of allele drop-in over an extended period of time.* FSI Genetics 48:102332.

Roewer, L., Andersen, M.M., Ballantyne, J., Butler, J.M., Caliebe, A., Corach, D., D'Amato, M.E., Gusmao, L., Hou, Y., de Knijff, P., Parson, W., Prinz, M., Schneider, P.M., Taylor, D., Vennemann, M. and Willuweit, S. (2020) *DNA Commission of the International Society of Forensic Genetics (ISFG): Recommendations on the interpretation of Y-STR results in forensic analysis.* FSI Genetics 48:102308.

Taylor, D., Bright, J-A. and Buckleton, J. (2016a) *Using probabilistic theory to develop interpretation guidelines for Y-STR profiles.* FSI Genetics 21:22-34.

Taylor, D., Bright, J-A., McGoven, C., Hefford, C., Kalafut, T., and Buckleton, J. (2016b) *Validating multiplexes for use in conjunction with modern interpretation strategies.* FSI Genetics 20:6-19.

Taylor, D., Curran, J. and Buckleton, J. (2018) *Likelihood ratio development for mixed Y-STR profiles.* FSI Genetics 35:82-96.

Tvedebrink, T., Erikesen, P.S., Aspund, M., Mogensen, H.S. and Morling, N. (2012a) *Allelic drop-out probabilities estimated by logistic regression – Further considerations and practical implementation.* FSI Genetics 6:263-267.

Tvedebrink, T., Erikesen, P.S., Mogensen, H.S. and Morling, N. (2009) *Estimating the probability of allelic drop-out of STR alleles in forensic genetics.* FSI Genetics 3:222-226.

Tvedebrink, T., Erikesen, P.S., Mogensen, H.S. and Morling, N. (2012b) *Statistical model for degraded DNA samples and adjusted probabilities for allelic drop-out.* FSI Genetics 6:97-101.

Weir B.S., Goudet J. (2017) *A unified characterization for population structure and relatedness*. Genetics 206:2085-2103.

Willuweit S., Roewer L. (2007) *International Forensic Y Chromosome User Group. Y chromosome haplotype reference database (YHRD): update.* FSI Genetics 1(2):83-87.

718 # Revision History

Document Version	Revision History
July 2021	Original. Drafted to provide supplemental information for the 2021 version of the *SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories*.

719