Scientific Working Group on DNA Analysis Methods

Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories

TABLE OF CONTENTS

SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories

The Scientific Working Group on DNA Analysis Methods, better known by its acronym of SWGDAM, is a group of scientists representing Federal, State, and Local forensic DNA laboratories in the United States and Canada. During meetings, which are held twice a year, committees discuss topics of interest to the forensic DNA community and often develop documents to provide direction and guidance for the community. Revisions to these guidelines, drafted by the SWGDAM Lineage Marker Committee, were presented to the SWGDAM membership and approved on March 2, 2022.

The guidelines described herein supersede the *SWGDAM Y-Chromosome STR Interpretation Guidelines* issued in 2014. This document contains several revisions to the 2014 guidelines and

is intended to reflect the current state of forensic Y-STR typing. SWGDAM intends for these guidelines to be applied prospectively. Provided that work (validation, training, analysis, interpretation) performed prior to the issuance of this revision was appropriate and scientifically valid, these revised guidelines are not intended to invalidate or call into question the previous work or to be applied retroactively. This document contains guidelines and not minimum standards. In the event of a conflict between the FBI Quality Assurance Standards for Forensic DNA Testing or DNA Databasing Laboratories (QAS) and these guidelines, the QAS and the QAS related audit guide and documents have precedence over these guidelines. Absent any other directive, the use of the terms *shall* or *must* is not intended to transform these guidelines into standards.

These guidelines are not intended to address the interpretation of analytical results from Y-STR testing using enhanced low template DNA techniques. This document therefore does not offer an opinion as to the validity of any enhanced detection methods (see Caddy et al. 2008 and *SWGDAM Guidelines for STR Enhanced Detection Methods* for more information).

Unlike previous versions, these guidelines are stated without further explanation included here. Refer to the *Supplemental Information for the SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories* document for further explanation of these guidelines and extensive background information.

Introduction

The interpretation of DNA typing results, including the results of Y-STR testing, requires professional judgment and expertise. Additionally, laboratories that analyze DNA samples for forensic casework purposes are required by the Quality Assurance Standards for Forensic DNA Testing Laboratories to establish and follow documented procedures for the interpretation and reporting of DNA typing results. Due to the multiplicity of forensic sample types and the potential complexity of DNA typing results, it is impractical and infeasible to cover every aspect of DNA interpretation by a preset rule. However, the laboratory should utilize written procedures for interpretation of analytical results with the understanding that specificity in the standard operating procedures will enable greater consistency and accuracy among analysts

within a laboratory. It is recommended that standard operating procedures for the interpretation of Y-STR typing results be sufficiently detailed that other forensic DNA analysts can review, understand in full, and assess the laboratory's policies and practices. The laboratory's interpretation guidelines should be based upon validation studies and scientific literature.

Background

Refer to the *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* for general background information regarding forensic DNA analysis and interpretation.

For the purposes of forensic nuclear DNA testing, the typing of autosomal STR loci is generally preferred due to the higher power of discrimination and utility for searching against the National DNA Index System. Y-STR typing is an additional tool that can be used, typically in concert with autosomal typing, in mixed samples with a small proportion of male DNA as compared to total human DNA. Y-STR typing may be used in lieu of autosomal typing for the detection of male DNA in mixtures that contain an overabundance of female DNA. Considering that under certain conditions a minor male contributor in a mixture of male and female DNA may only be detectable by Y-STR typing, laboratories should pursue Y-STR analysis as the most appropriate means of detecting a male contributor(s) in some forensic samples. Due to the transmission of the Y-chromosome within a paternal lineage, Y-STR typing can also aid in the identification of missing persons and familial searches.

Y-STR loci exhibit the same general characteristics as autosomal STR loci, namely:

- Peak height variability, which is inversely proportional to peak heights and manifests as either inconsistencies in mixture proportion, deviations from expected stutter ratios or variation in peak heights between loci.
- Stutter, to include back stutter and forward stutter. The height of a stutter peak is generally dependent on the parent peak's height, locus and allele length.
- Allelic peak heights that are the sum of the contributions from each donor to a multiple contributor profile. Allelic peak heights can also be additive with overlapping stutter.

- Depending on the amplification kit, separating technology, and sample quality/quantity, a general downward trend in peak heights with increasing molecular weight.
- Allelic drop-out, as expected at low DNA quantity/quality (low either due to initial template amount, PCR inhibition and/or DNA degradation).
- Allelic drop-in, occurring at a rate which is dependent on the amplification kit and factors that impact sensitivity of detection (e.g., amplification cycle number, detection instrumentation and detection conditions such as the injection voltage and time for capillary electrophoresis).
- Locus-specific amplification efficiency. The magnitude of differences in amplification efficiency among loci can depend on kit formulation, DNA sample quality and the presence of PCR inhibitors.

All Y-STR loci are physically linked on the Y-chromosome. Due to the lack of genetic recombination, the entire Y-chromosome haplotype must be treated as a single locus. Subsequent to a match between two samples using Y-STR testing, a single-source, major, or deduced Y-STR haplotype may be searched against a database of Y-STR haplotypes to obtain the sample frequency of the profile, and, as needed, to calculate profile and/or match probabilities. It is noted that two specimens that exhibit the same Y-STR haplotype may have originated from either a common individual source, from males within the same paternal lineage, or from unrelated individuals. A paternal lineage consists of those male relatives to whom the same Y-chromosome has been transmitted from a common ancestor. Barring mutation, all male relatives within the same paternal lineage have the same Y-STR profile. Attribution of the Y-STR typing results to a single individual, to the exclusion of relatives in the paternal lineage, is generally not possible based on Y-chromosome loci. However, loci with higher mutation rates may enhance the ability to distinguish relatives in the same paternal lineage (Ballantyne et al. 2012).

1. Application of Y-STR Typing

1.1 The laboratory should establish guidelines that define the parameters under which samples are subjected to Y-STR typing.

1.1.1 Based on STR amplification system and quantification system validations, the laboratory should establish guidelines for when detection of a male contributor to a mixture is not expected with autosomal typing (e.g., based on male to total DNA quantities).

2. Preliminary Evaluation of Data

Refer to the *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* for general guidance regarding the following topics: analytical threshold, internal standards, allelic ladders, controls and concordance of redundant loci.

3. Allele Designation

Refer to the *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* for general guidance regarding locus and allele designation.

3.1 Alleles should be designated in accordance with recommendations of the DNA Commission of the International Society of Forensic Genetics (Gill et al. 2001 and Gusmão et al. 2006).

3.2 The laboratory should establish guidelines based on experimental studies for the identification of null alleles. The guidelines should ensure that a null allele can be distinguished from an undetected allele resulting from low template amounts, DNA degradation or inhibition (i.e., allelic drop-out).

4. Identification of Non-Allelic Peaks

Y-STR typing results generated with the current Y-STR typing kits exhibit the same non-allelic peaks observed in autosomal STR typing results. Refer to the *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* for general guidance regarding non-allelic peaks and off-scale data.

4.1 The laboratory should establish a method based on validation to document the designation of a peak as an artifact or an allele.

5. Application of the Stochastic Thresholds to Allelic Peaks

Refer to the *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* for general guidance regarding the establishment and usage of the stochastic threshold.

5.1 The laboratory should establish a stochastic threshold for known multi-copy Y-STR loci (e.g., DYS385 and DYF387S1) based on empirical data derived within the laboratory and specific to the quantification and amplification systems and the detection instrumentation used.

6. Application of Peak Height Ratios to Multi-copy Y-STR Loci

Refer to the *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* for general guidance regarding the establishment and usage of peak height ratios.

6.1 The laboratory should establish an expected peak height ratio for known multi-copy Y-STR loci (e.g., DYS385 and DYF387S1) based on empirical data derived within the laboratory and specific to the amplification systems and the detection instrumentation used.

7. Number of Contributors to a Y-STR Profile

Refer to the *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* for general guidance regarding the recognition of mixtures, the minimum number of contributors to a mixture, and the generation of composite profiles.

7.1 A laboratory should assess the number of contributors to a Y-STR profile. Y-chromosome profiles typically show one allele per locus except for multi-copy loci (e.g., DYS385 and DYF387S1). A specimen is generally considered to have originated from more than one male individual if two or more alleles are present at two or more single-copy loci.

7.1.1 For a given locus, a laboratory can assess repeat-unit differences among the detected alleles to aid in distinguishing a mixed sample from a single-source sample that exhibits duplication.

7.2 The laboratory should establish guidelines based on an assessment of peak heights and peak height ratios for evaluating potential sharing of allelic peaks between major and minor contributors and for determining whether the alleles of the contributors to a mixed DNA typing result are distinguishable.

8. Comparison of DNA Typing Results

Refer to the *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* for general guidance regarding the following topics: interpretation of evidentiary profiles relative to that of known profiles, partial profiles, possible conclusions, full accounting of mixed results, documentation of assumptions, and results for which no comparisons will be made.

8.1 The laboratory should establish guidelines, based on internal validations, for determining whether a Y-STR typing result is suitable for comparisons.

8.2 Single-source Y-STR haplotypes, including partial haplotypes, may be used for comparison purposes, inclusionary and exclusionary. The laboratory should establish the minimum number of loci from an evidentiary profile required to perform a comparison to a reference profile.

8.3 Mixtures of DNA from more than one male individual may also be used for comparison purposes when the contributors can be distinguished based on means such as peak height

ratio comparisons. These deduced haplotypes are then used for comparison as in 8.2. Such haplotypes include (a) those of major, and potentially minor, contributor(s) to a distinguishable mixture, and (b) for an indistinguishable mixture, those foreign alleles derived from separation of a conditional known sample type (e.g., from the consensual partner).

8.4 The laboratory should establish guidelines for identifying mixtures for which no major or minor contributor can be discerned. Interpretation and comparison of indistinguishable Y-STR mixtures shall be supported by internal validations.

9. Statistical Analysis of Y-STR Typing Results

Refer to the *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* for general guidance regarding the following topics: when to perform statistical analysis, data appropriate for use in statistical analysis, and reporting of statistical analysis.

9.1 The Y-Chromosome Haplotype Reference Database (YHRD, Willuweit and Roewer 2007) available at https://yhrd.org/ contains a U.S. Y-STR population database which should be used for estimation of profile (i.e., haplotype) probabilities and match probabilities. As this database is regularly updated with population data from new population studies (for example, Carrecedo et al. 2010), it is advised to state the YHRD release for traceability.

9.2 Statistical Analysis of Single-Source and Deduced Single-Source Y-STR Haplotypes

9.2.1 The laboratory should establish guidelines for the number of Y-STR loci used for searches of population databases. Due to the challenge of small database sizes for the larger multiplex systems, it is acceptable to perform additional searches of the population database using reduced locus sets in an attempt to obtain the most informative result for that combination of evidence and population database profiles.

9.2.1.1 Regardless of the number or selection of loci searched, the most informative search is generally the one which gave the lowest proportion of matching haplotypes per number of profiles compared.

9.2.1.2 When performing reduced locus-count searches, any "matches" that would have been non-matches had more of the evidence profile been searched must be excluded. 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.

9.2.2 The laboratory should determine which of the following methods will be used to determine the haplotype sample frequencies.

9.2.2.1 A Y haplotype sample frequency can be determined using the observed counting method. The Y haplotype sample frequency (p) is calculated using $p = x/n$ formula, where x is equal to the number of times the haplotype is observed in the database containing n number of haplotypes in the database.

9.2.2.2 A Y haplotype sample frequency can be determined using the augmented counting method. This Y haplotype sample frequency (p) is calculated using the $p = (x+1)/(n+1)$ formula, where x is equal to the number of times the haplotype is observed in a database and n is equal to the number of haplotypes in the database (Gjertson, et al. 2007).

9.2.3 A Y-STR haplotype upper bound profile probability estimate can be calculated from the observed or augmented haplotype frequencies by including a confidence interval (generally 95% or greater) to capture the effect of database size (Clopper and Pearson 1934). The laboratory should establish if an upper confidence limit for the haplotype probability or a confidence interval for the haplotype probability will be calculated and reported.

9.2.4 The laboratory should determine if match probabilities will be used and the method to calculate them. There is no consensus from statistical subject matter experts as to the best or preferred method for calculating match probabilities. See Brenner (2010), Budowle (2009), and Weir and Goudet (2017) and the *Supplemental Information for the SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories* document.

9.2.5 The laboratory should determine if likelihood ratios will be used to provide quantitative assessments of the value of the matches using relevant populations. An approach is described in Evett and Weir (1998) and is recommended by the DNA Commission of the International Society of Forensic Genetics (Roewer, et al. 2020).

9.3 Statistical Analysis of Indistinguishable Y-STR Mixtures

9.3.1 A laboratory choosing to report inclusionary Y-STR typing results from indistinguishable mixtures must perform statistical analysis in support of any inclusion determined to be relevant in the context of the case.

10. Combining Lineage Markers and Autosomal Results

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. Such an expectation could arise from large scale independence testing or strong population genetic models (Walsh and Hammer 2008, Buckleton and Myers 2014).

References

Ballantyne, K.N., Keerl, V., Wollstein, A., Choi, Y., Zuniga, S.B., Ralf, A., Vermeulen, M., de Knijff, P., and Kayser, M. (2012) *A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages*. FSI Genetics 6:208–218.

Brenner, C. H. (2010) *Fundamental Problem of forensic mathematics-The evidential value of a rare haplotype.* FSI Genetics 4:281-291.

Buckleton, J., Myers, S. (2014) *Combining autosomal and Y chromosome match probabilities using coalescent theory*. FSI Genetics 11:52-55.

Budowle, B., Ge, J., Aranda, X.G., Planz, J.V., Eisenberg A.J., Chakraborty, R. (2009) *Texas population substructure and its impact on estimating the rarity of Y STR haplotypes from DNA evidence*. J. Forensic Sci. 54(5):1016-21.

Caddy, B., Taylor, G.R., Linacre, A.M.T. (2008) *A review of the science of low template DNA analysis*. Available at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/117556/Review_of_Low_Template_DNA_1.pdf.

Carracedo A., Butler J.M., Gusmão L., Parson W., Roewer L., Schneider P.M. *Publication of population data for forensic purposes.* (2010) FSI Genetics. 4(3):145-147.

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

Evett, L.W., Weir, B.S. (1998) *Interpreting DNA evidence: statistical genetics for forensic scientists*. Sinauer Associates. Sunderland.

Gill, P., Brenner, C., Brinkmann, B., Budowle, B., Carracedo, A., Jobling, M.A., De Knijff P., Kayser, M., Krawczak, M., Mayr, W.R., Morling, N., Olaisen B., Pascali, V., Prinz, M., Roewer, L., Schneider, P.M., Sajantila, A., Tyler-Smith, C. (2001) *DNA Commission of the International Society of Forensic Genetics: Recommendations on forensic analysis using Y-chromosome STRs*. Forensic Sci. Int. 124(1):5-10.

Gjertson, D.W., Brenner, C.H., Baur, M.P., Carracedo, A., Guidet, F., Luque, J.A., Lessig, R., Mayr, W.R., Pascali, V.L., Prinz, M., Schneider, P.M., Morling, N. (2007) *ISFG: Recommendations on biostatistics in paternity testing*. FSI Genetics 1:223-231.

Gusmão, L., Butler, J.M., Carracedo, A., Gill, P., Kayser, M., Mayr, W.R., Morling, N., Prinz, M., Roewer, L., Tyler-Smith, C., Schneider, P.M. (2006) *DNA Commission of the International Society of Forensic Genetics: An update of the recommendations on the use of Y-STRs in forensic analysis*. Forensic Sci. Int. 157(2-3):187-197.

Roewer L., Andersen M.M., Ballantyne J., Butler J.M., Caliebe A., Corach D., D'Amato M.E., Gusmão L., Hou Y., de Knijff P., Parson W., Prinz M., Schneider P.M., Taylor D., Vennemann M., 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.

Walsh, B., Redd, A.J., Hammer, M.F. (2008) *Joint match probabilities for Y chromosomal and autosomal markers*. Forensic Sci. Int. 174:234-238.

Weir, B., Goudet, J. (2017) *A Unified Characterization of 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.

Glossary

Refer to the *SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories* for definitions of terms. Additional terms used in this document are defined below.

Back stutter: A PCR artifactual peak typically one repeat unit shorter than the parent allele. Two repeat units or partial repeat units may also be observed.

Deduced Single-Source Y-STR Haplotype: A Y-haplotype from one contributor determined from a mixture, usually by inference of an unknown contributor's DNA profile after taking into consideration the contribution of a known male contributor's alleles, if appropriate, or quantitative peak height information.

Drop-in: The random appearance of non-reproducible allelic peaks in a profile thought to arise from fragments of cells introduced into the extract and not from the principal donors.

Drop-out: The event where an allele in the sample does not produce a peak above the analytical threshold.

Forward stutter: A PCR artifactual peak typically one repeat unit longer than the parent allele. Two repeat units or partial repeat units may also be observed.

Indistinguishable mixture: a DNA mixture in which relative peak height ratios are insufficient to attribute alleles to individual contributor(s).

Null (silent) allele: An allele which cannot be detected due to lack of amplification product, often caused by a mutation in the primer binding site, or deletion of the primer binding site or locus.

Revision History

Document Version	Revision History
January 2009	Original. Published in Forensic Science Communications in January 2009, Vol. 11, No. 1.
January 2014	The document was revised to reflect the current state of forensic Y-STR typing. Additional information was added regarding multi-copy and duplicated loci as well as null alleles. The statistical treatment of Y-STR haplotypes was expanded to include an appendix of theta values, references to the US Y-STR database and the calculation of match probabilities. The use of Clopper-Pearson as an upper confidence interval was added.
January 2022	This document contains several revisions to the 2014 guidelines and is intended to reflect the current state of forensic Y-STR typing. The statistical treatment of Y-STR haplotypes includes the use of YHRD, augmented counting method, and multiple match probability options. Removal of U.S. Y-STR database use and appendix of theta values.