Autosomal STR Interpretation

1. *Purpose*

To describe the procedure for STR data analysis and interpretation.

2. *Summary*

Determine the Number of Contributors, Assess Locus Suitability, Perform Comparisons, and Report Conclusions. ArmedXpert™ software is used to assist in statistical calculations and mixture interpretation for both CE and MPS data.

3. *Procedure*

Number of Contributors:

1. Determination of the number of contributors is made by assessing the total number of alleles at each locus and their relationship to each other.
    a. An analyst may adjust the number of contributors reported based on additional factors at their discretion.

1 (Single Source):	No more than two alleles present at a locus.
1+	A single source profile with indications of a possible additional contributor below analytical threshold. A minimum of one contributor may also be used to describe uncertainty if calling only one contributor is not desired.
2 (Two only):	No more than four alleles present at a locus.
2+ (Minimum of two):	No more than four alleles present at a locus with indications of a possible additional contributor. A minimum of two contributors may also be used to describe uncertainty if calling only two contributors is not desired.
3 (Three only):	No more than six alleles present at a locus.
3+ (Minimum of three):	No more than six alleles present at a locus with indication of a possible additional contributor. A minimum of three contributors may also be used to describe uncertainty if calling only three contributors is not desired.
4+ (Minimum of four):	Seven or more alleles present at a locus.

Note: Tri-Alleles may alter the number of observed alleles for a contributor.

Multiple allele height imbalances may indicate the presence of a mixture or an additional contributor. 5+, 6+, etc. can be called if sufficient alleles are present to support the call.

2. The number of assumed contributors will be documented on the DOC-323 Sample Interpretation Worksheet.

Gender Determination

3. Any called allele in a male identifying marker denotes the presence of male DNA.

4. Gender determination may be accomplished with a review of both STR and YSTR data, if co-amplification has been performed (i.e., NGS).

5. Gender determination for single source, major and deduced profiles is at analyst discretion.

Data Interpretation:

6. ArmedXpert™ software is used for assistance in interpretation and performing statistical calculations. The software allows Unrestricted, Restricted, Modified and Conditional RMP statistics. Based on the analyst input, the software will apply the appropriate statistic type. It can be used for interpretation assistance of both CE and MPS data.
    a. ArmedXpert™ provides visualization (Allele Bar Chart/electropherogram, Aggregate/Locus specific contributor ratios) to assist the analyst in interpretation decisions.
    b. ArmedXpert™ allows the analyst to toggle interpretation decisions for alleles/isoalleles obtained in the DNA profile (major, minor, obligate allele, allele/any) and automatically applies the appropriate statistical model for the chosen decision.
    c. Interpretation/technical review should be focused on these interpretation decisions to confirm appropriate alleles are included in the relevant statistics.
    d. ArmedXpert™ documentation of the locus/allele/isoallele interpretation will be exported from the software and included in the case file to document the interpretation decisions for the profile.
    e. Statistical calculations are reported using 3 figures (i.e. 3.54 Duodecillion, 580 Tredecillion, etc.)
    f. Statistics are reported for the Caucasian, African American, Hispanic, and Asian population groups.

Single Source Profile:

7. In a single source profile, all called alleles are determined to be from the same contributor.

8. A fully genotyped locus is either homozygous that is above the stochastic threshold or heterozygous.

9. An obligate allele can be suitable for comparison when there is a single called allele below

the stochastic threshold or at analyst discretion using modified RMP.

10. 1+ profiles
    a. Heterozygous locus
    i. Both alleles above the stochastic threshold
        1. Both are attributable to the profile.
        ii. One allele below the stochastic threshold
            1. Allele above the threshold is marked as an obligate.
            2. Allele below the threshold is marked with “Any.”
        iii. Both alleles below the stochastic threshold
            1. Both alleles are marked with “Any”
    b. Homozygous locus
        i. Above the stochastic threshold plus the analytical threshold
            1. Attributable to the profile
        ii. Below the stochastic threshold plus analytical threshold
            1. Marked as “Any”

11. Refer to the **Minimum Suitability Requirements** section for profile suitability minimums.

Mixtures:

General Guidelines:

12. ArmedXpert™ provides calculations for possible allele pairing based on analyst input parameters for mixtures with up to three contributors. These parameters are used as a starting point for the interpretation.
    a. Refer to IAC-200 ArmedXpert MixtureAce

13. Analytical and Stochastic thresholds are set for each testing kit and are not to be changed for interpretation.

14. Allele (Peak) height ratio is set at 60% but is not intended to be a threshold in assessing possible allele pairings.

15. Assess the Electropherogram or Allele Bar Chart (ABC) to get a full picture of the data and determine the number of contributors before assigning profile alleles.
    a. Assess for elevated stutter and artifacts.
    b. Known SNP locations are documented in Gettings et al 2019 Supplementary Excel document and should be used in assessing called alleles that are not in the NIST database.

16. Begin mixture interpretation with the loci that have the highest number of alleles calling. This assists in establishing contributor ratios.

17. Assign profile allele(s) locus by locus using analyst discretion and the assistance of the ArmedXpert™ calculations for all contributors.
    a. Include all possible allele combinations determined to be possible for each contributor where possible.

18. For non-discernable mixtures, each locus is assessed for dropout.
    a. 2 person
        i. 4 alleles
            1. Suitable with all alleles
        ii. 3 alleles

1. Dropout is not suspected – Suitable with all alleles
2. Dropout is possible - Suitable for Forced RMP

iii. 2 alleles or 1 allele
1. Dropout is not suspected – Suitable with all alleles
2. Dropout is possible - Not suitable

b. 3 person
i. 6 alleles
1. Suitable with all alleles
ii. 5 alleles
1. Dropout is not suspected – Suitable with all alleles
2. Dropout is possible - Suitable for Forced RMP
iii. 4, 3, 2 alleles or 1 allele
1. Dropout is not suspected – Suitable with all alleles
2. Dropout is possible - Not suitable

19. For 4 person mixtures, only a major or co-major profile can be called:
a. The Major/Minor disparity is calculated by dividing the allele height of the highest minor allele by the peak height of the lowest major allele.
b. Suitability is determined on a locus-by-locus basis, but the entire mixture should have the general visible appearance of having a major profile.

# of minor peaks	Heterozygous	Homozygous
No Minor	---	---
1 Minor Peak	---	---
2 Minor Peaks	10%	5%
3 Minor Peaks	20%	10%
4 Minor Peaks	30%	15%
5 Minor Peaks	40%	30%
6 Minor Peaks	50%	40%

--- Analyst discretion based on general contributor disparities of the mixture as a whole Disparity percentages must be lower than indicated above.

Statistical Calculations:

20. Statistical calculations are included in the case file for each suitable profile on the ArmedXpert interpretation/statistics page.
a. ArmedXpert™ statistical calculations for four population groups and “All”
i. All
ii. “AfAm” (African American)
iii. “Cauc” (Caucasian)
iv. “Hisp” (Hispanic)
v. Asian
b. Only the individual population groups are added to the report document.
c. Databases used:
i. MPS
1. Sequence based US population data for 27 autosomal STR loci_NIST1036_auSTR_Seq_SuppTables
ii. CE

1. NIST 1036-Revised-Allele-Freqs-PopStats-July-19-2017

Known Contributors:

21. When it is reasonable to assume that a sample has a known contributor, entering the reference may provide additional interpretational assistance.
    a. ArmedXpert™ allows assignment of alleles/isoalleles of known profiles to assist in this interpretation.

Minimum Suitability Requirements:

22. A minimum rarity for statistical calculation of 1 in 100 is required for profile comparison suitability. Any profile probability that is more common than 1 in 100 at any one of the four reported statistical calculations (see above) does not meet minimum suitability requirement.
    a. ArmedXpert™ is used to build profile allele possibilities. Any test result that does not meet the **Minimum Suitability Requirements** is deemed not suitable for comparison and no profile possibilities are reported.

23. Mixtures that contain at least five contributors are considered uninterpretable due to complexity of the mixture.

24. Mixture profiles that contain more than three contributors are interpretable only if discernible profiles can be obtained from the major portion of the profile.
    a. A single major or co-major are possible in this scenario.
    b. If individual profiles are not interpretable, four person mixtures are deemed uninterpretable due to complexity of the mixture.

25. Any item that generates a profile (or multiple profiles) that do not meet or exceed these minimum suitability requirements will be deemed not suitable for comparison (inconclusive) and the reason (complexity, rarity) will be included in the report conclusion (see report wording guidelines).

26. Any item that generates a profile (or multiple profiles) that meet or exceed these minimum suitability requirements, the item will be deemed suitable for comparison.
    a. Any item which obtains an individual profile deemed suitable for comparison will be given an “Unknown Profile #” (Unknown Profile 01, Unknown Profile 02, Unknown Profile 03 etc.)
    b. Profiles that are attributable (cannot be excluded from each other) will be grouped together
        i. Example: if three items obtained a profile that cannot be excluded from each other, these profiles would be reported as attributable to same Unknown Profile #.
    c. Non-discernable mixtures will include a comparison to each obtained Unknown Profile.
    d. Appropriate statistics will also be reported for each Unknown Profile #

27. The assessment for suitability and deriving individual profiles will be completed prior to interpretation (where possible) and comparison to known sample(s)
    a. Loci that are unsuitable for comparison and/or statistics will be identified in the case documentation. The reason for the decision will also be included in the case

documentation.

b. Reasons for overall profile unsuitability will be documented on the DOC-323 Sample Interpretation Worksheet.

Probabilistic Genotyping:

28. Probabilistic genotyping software does not use this rule set for data interpretation

Comparisons:

29. All samples with results that are determined to be suitable for comparison are compared to all known reference samples and/or profiles that are provided for that case.

30. A statistical value greater than 1 in 100 is required for all unknown to known sample comparisons.
    a. If this requirement cannot be met due to a partial reference profile, a no meaningful comparison conclusion can be made.

31. All matches and inclusions not considered to be intimate require statistical calculations to be performed and reported.
    a. Intimacy to an item is reported as an assumed contributor.

32. When all comparable alleles of the unknown profile are consistent with a known profile, a **Match** is declared.

33. When a profile can be included in a comparable non-discernable mixture, an **Inclusion** (as a possible contributor) is declared.

34. Any indication that profiles do not match, or a profile could not be a possible contributor of a mixture, results in an **Exclusion**.

Legacy Data Interpretation or comparisons:

35. Data interpretations must be performed using the thresholds of the laboratory where that data is generated.

36. Comparisons between profiles generated using different typing methods (Capillary Electrophoresis vs. MPS) use the allele frequencies for the CE data to generate statistical calculations.

Thresholds:

37. Analytical

KIT	Threshold
Globalfiler	Blue – 42 RFU Green – 65 RFU Yellow – 44 RFU Red – 61 RFU Purple – 67 RFU
Investigator 24Plex	Blue – 36 RFU Green – 59 RFU Yellow – 80 RFU Red – 87 RFU Purple – 62 RFU

ForenSeq	60 Reads

38. Stochastic
(For STR kits (Globalfiler/Investigator 24Plex), this is an RFU value. For NGS (ForenSeq), this is a read count).

KIT	Locus	Threshold
Globalfiler		690
Investigator 24Plex		430
ForenSeq	Amelogenin	253
	D1S1656	256
	TPOX	739
	D2S441	963
	D2S1338	307
	D3S1358	791
	D4S2408	115
	FGA	556
	D5S818	615
	CSF1PO	470
	D6S1043	1084
	D7S820	302
	D8S1179	791
	D9S1122	805
	D10S1248	522
	TH01	2073
	vWA	297
	D12S391	640
	D13S317	823
	PentaE	504
	D16S539	447
	D17S1301	852
	D18S51	524
	D19S433	572
	D20S482	1101
	D21S11	343
	PentaD	449
	D22S1045	514

4. *References*

ArmedXpert™ User Manual

Gettings et al 2019

Supplementary Excel document

5. *Definitions*

Allele Bar Chart (ABC): An allele chart generated by ArmedXpert™ software using MPS data from the Verogen Forenseq Sequencing kit.

Allele Height Ratio: The relative ratio of two alleles at a given locus, as determined by dividing the height of an allele with a lower relative fluorescence unit (RFU) value or read count by the height of an allele with a higher RFU value or read count, and then multiplying this value by 100 to express the allele height ratio as a percentage. This ratio is used as an indication of which alleles may be heterozygous pairs and also to determine the possible presence of a DNA mixture.

Analytical Threshold: The minimum height requirement at and above which detected alleles can be reliably distinguished from background noise; data above this threshold are generally not considered noise and are either artifacts or true alleles.

Electropherogram: The graphic representation of the separation of DNA fragments by electrophoresis in which data appear as peaks along a line.

Exclusion: When the genotype comparison can only be explained by the two samples originating from different contributors.

Heterozygote: Presence of different alleles at a locus.
Homozygote: Presence of identical alleles at a locus.

Inclusion: When the compared STR profiles can be included as a possible contributor in a mixture.

Inconclusive: “Inconclusive” may be used to describe a specific allele, locus, or entire DNA result.

Inconclusive Allele: Unable to determine whether allele is real DNA or artifact and will only be used in determining number of contributors.

Inconclusive Locus: Locus will not be used for comparison, typically due to incomplete results in a positive control.

Inconclusive Interpretation Result: DNA results do not meet the minimum requirements for interpretation.

Indiscernible Mixture: A DNA mixture in which relative allele height ratios are insufficient to attribute alleles to individual contributor(s).

Iso-allele: An allele that is sequenced by NGS and may have sequence variation that is indistinguishable using CE fragment analysis STR.

Match: When two profiles have no unexplainable differences when compared.

Mixture: The presence of DNA from multiple contributors.

Non-discernable Mixture: A mixture where individual profiles are not able to be determined.
Obligate Allele: An allele which belongs to a specific contributor which may or may not be paired with another allele. **Probabilistic Genotyping**: The use of biological modeling, statistical theory,

computer algorithms, and probability distributions to calculate Likelihood ratios and/or infer genotypes for DNA typing results.

Random Match Probability (RMP):

Unrestricted – considers all possible allele combinations, no dropout Restricted – analyst eliminates improbable alleles

Modified – calculation considers potential dropout. Conditional – calculation considers an assumed contributor

RFU: (relative fluorescent units): measurement of signal intensity.

Stochastic Threshold: The allele height value above which it is reasonable to assume that, at a given locus, allelic drop-out of a sister allele has not occurred.

Tri-allelic Pattern: Result from extra chromosomal fragments being present in a sample or DNA sequence where the primer's allele is duplicated on one chromosome thus resulting in an extra allele at a single locus.