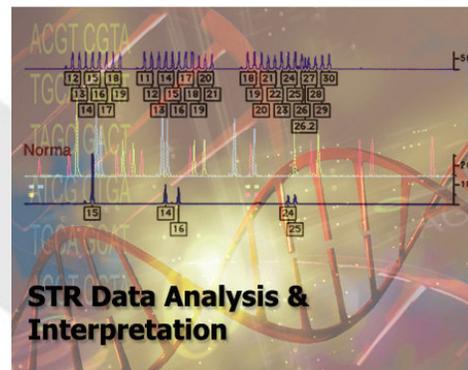


STR Data Analysis & Interpretation for Forensic Analysts

This course is provided free of charge and is part of a series designed to teach about DNA and forensic DNA use and analysis.

Find this course live, online at:
<http://dna.gov/training/strdata>

Updated: October 8, 2008



PRESIDENT'S

DNA

INITIATIVE



www.DNA.gov

About this Course

This PDF file has been created from the free, self-paced online course “Crime Scene and DNA Basics for Forensic Analysts.” To learn more and take this and other courses online, go to <http://www.dna.gov/training/online-training/>. Most courses are free but you must first register at <http://register.dna.gov>.

If you already are registered for any course on DNA.gov, you may login directly at the course URL, e.g., <http://letraining.dna.gov> or you can reach the courses by using the URL <http://www.dna.gov/training> and selecting the “Login and view your courses” link.

Questions? If you have any questions about this file or any of the courses or content on DNA.gov, visit us online at <http://www.dna.gov/more/contactus/>.

Links in this File

Most courses from DNA.Gov contain animations, videos, downloadable documents and/or links to other useful Web sites. If you are using a printed, paper version of this course, you will not have access to those features. If you are viewing the course as a PDF file online, you may be able to use some of these features if you are connected to the Internet.

Animations, Audio and Video. Throughout this course, there may be links to animation, audio or video files. To listen to or view these files, you need to be connected to the Internet and have the requisite plug-in applications installed on your computer.

Links to other Web Sites. To listen to or view any animation, audio or video files, you need to be connected to the Internet and have the requisite plug-in applications installed on your computer.

Legal Policies and Disclaimers

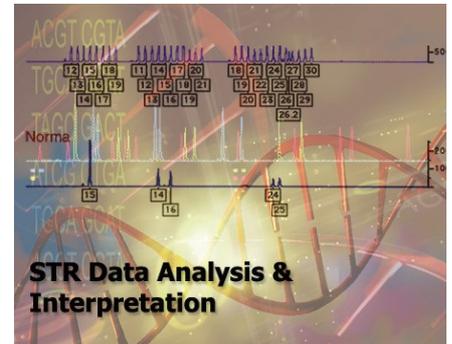
See [Legal Policies and Disclaimers](#) for information on Links to Other Web Sites, Copyright Status and Citation and Disclaimer of Liability and Endorsement.

STR Data Analysis and Interpretation for Forensic Analysts

This course provides information in three lessons.

STR (Short Tandem Repeat) Data Analysis and Interpretation Software.

Learn the basics of data analysis software, become familiar with the purpose of GeneScan® and Genotyper® software, learn the difference between GeneScan® and Genotyper® software and GeneMapper ID® software, and become aware of the unique features of GeneMapper ID® software, and understand FMBIO® Analysis software and STaRCaLL™ software as related to GeneScan® and Genotyper® software.

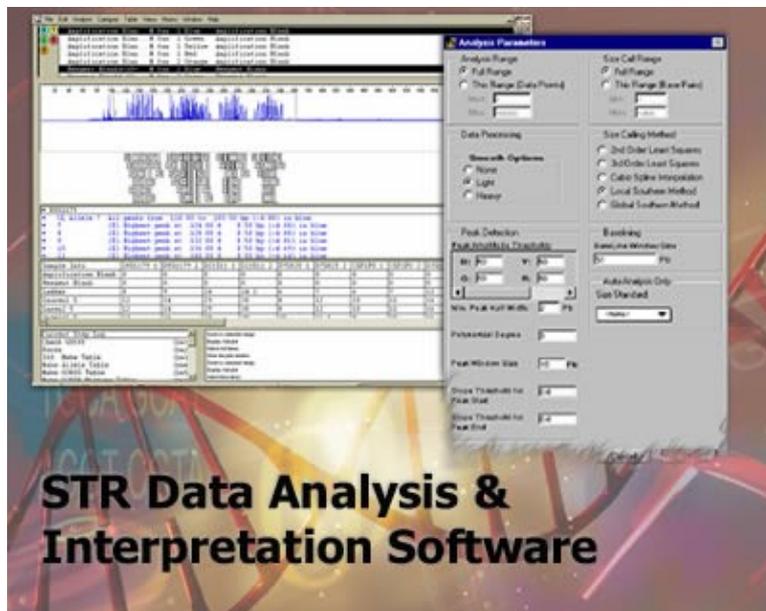


Data Troubleshooting. Learn about commonly used STR data interpretation parameters and thresholds, spurious peaks, pull-up, stutter, microvariants, degradation, stochastic effects, allele dropout, mutations, and controls used in troubleshooting STR data.

Data Interpretation and Allele Calls. Learn the factors that can influence the output from instruments used in DNA analysis, learn about the controls and other techniques used to validate instrument output, compare methods to evaluate data and recognize analysis artifacts, and learn how to analyze data from single source and mixed source samples.

STR (Short Tandem Repeat) Data Analysis & Interpretation Software

Introduction



Following separation of amplified DNA products, the information from the DNA separation must be converted into a common language that is standard from laboratory to laboratory. Software programs provide the means to perform the necessary data analysis and standardize the output.

STR Data Analysis and Interpretation for Forensic Analysts

Objectives

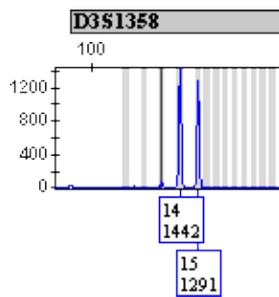
Upon successful completion of this unit of instruction, the student shall be able to:

- Identify data analysis software
- Be familiar with the purpose of GeneScan® and Genotyper® software
- Differentiate between GeneScan® and Genotyper® software and GeneMapper *ID*® software
- Explain the unique features of GeneMapper *ID*® software
- Understand FMBIO® Analysis software and STaRCaLL™ software as related to GeneScan® and Genotyper® software

Overview

Data produced in the separation and characterization of amplified DNA is displayed as:

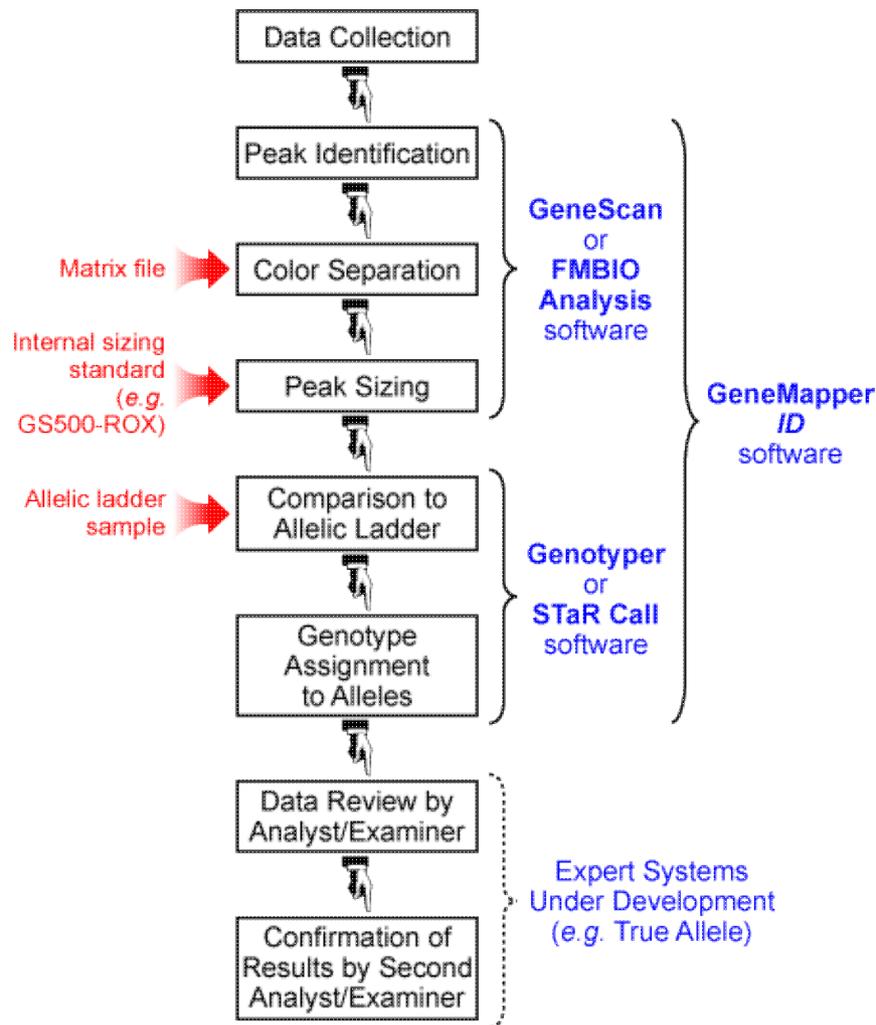
- Peaks (capillary electrophoresis)
- Bands (slab gel electrophoresis)



The DNA fragments are sized, which includes an indirect assessment of quantity present (peak area/height or band density), and genotypes are assigned. The conversion of sized DNA fragments to genotypes is the standardization between all forensic DNA laboratories for comparing data and is essential for laboratories utilizing CODIS to compare profiles. [Read more about CODIS.](#)

The steps for converting fluorescent data/peaks into allele calls are shown below with the corresponding software noted to the right of the specific steps.01

STR Data Analysis and Interpretation for Forensic Analysts



[View an animation about the basic components of an electropherogram.](#)

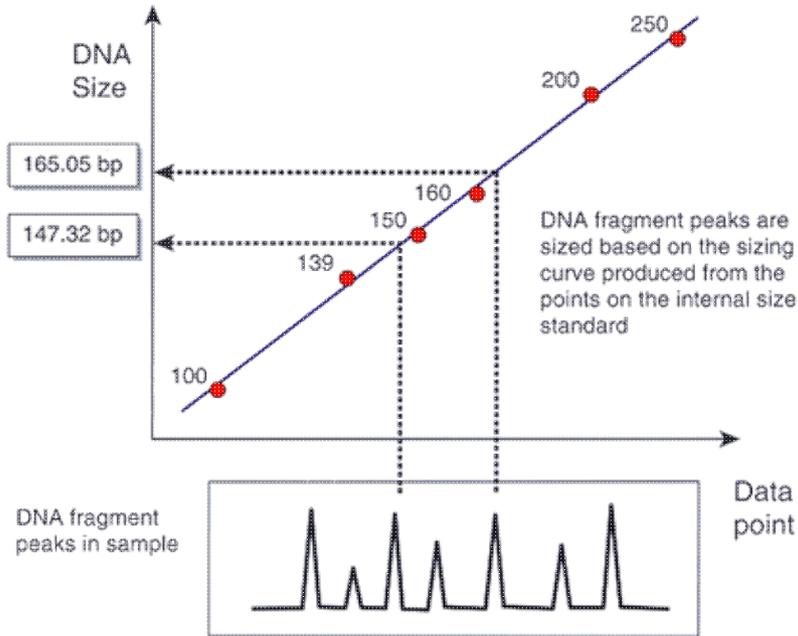
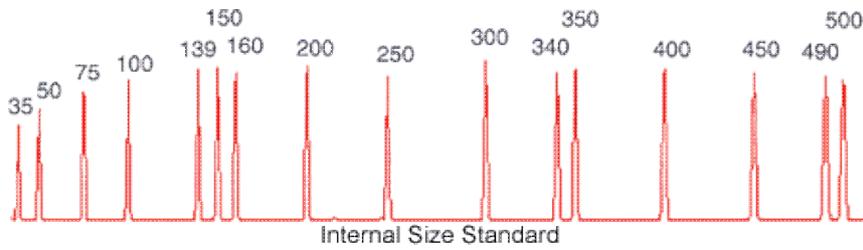
GeneScan®



GeneScan® is a sophisticated software program that converts raw data to analyzed data through the application of a size standard, a matrix file, and specific parameter settings.

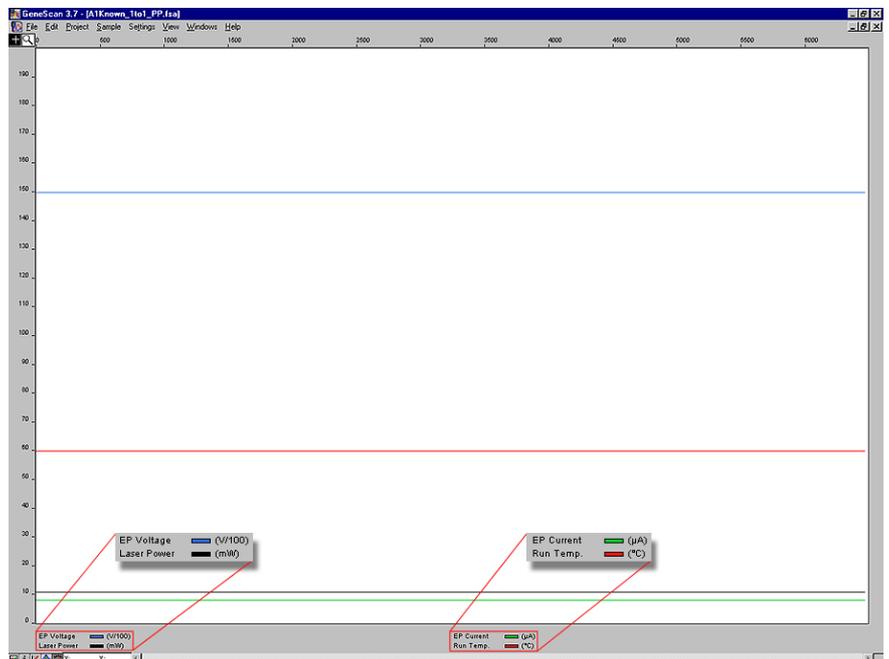
GeneScan® software sizes peaks using the internal size standard added to the sample prior to separation, as shown below.

STR Data Analysis and Interpretation for Forensic Analysts



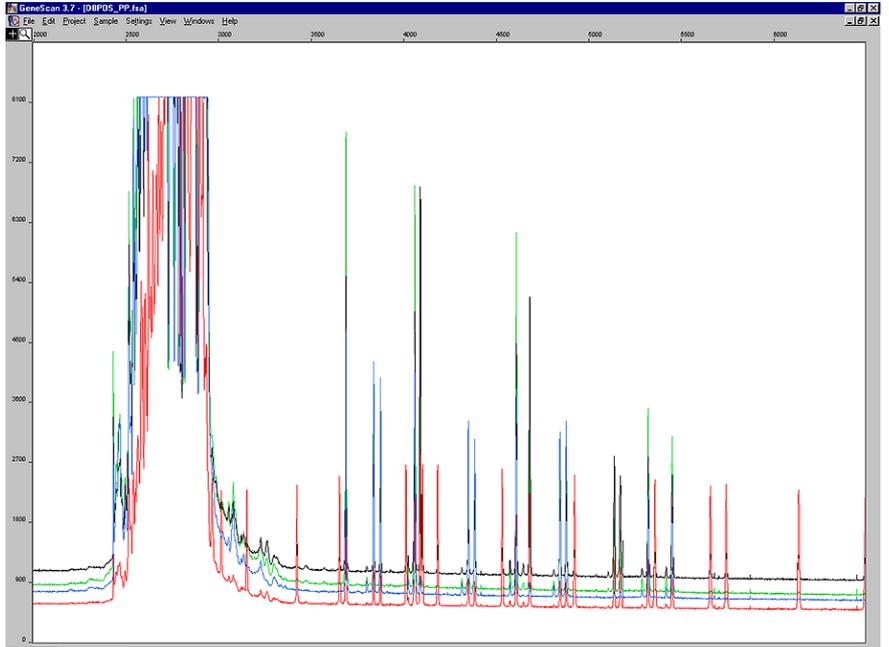
While samples are being processed by the genetic analyzer, sample files (.fsa) are generated during the data collection process. These files are then analyzed by the GeneScan® analysis software. There are five vital pieces of information in the sample file that are used and displayed during GeneScan® data analysis:

- Ept file

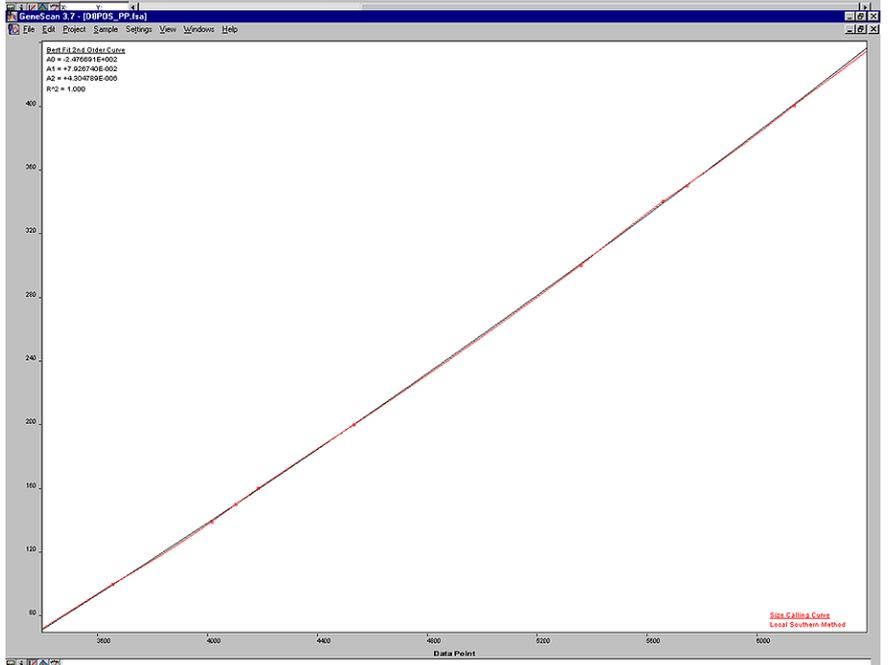


STR Data Analysis and Interpretation for Forensic Analysts

- Raw Data



- Size calling curve



STR Data Analysis and Interpretation for Forensic Analysts

- Sample file information

GeneScan 3.7 - [D8POS_PP.fsa]

File Edit Project Sample Settings View Windows Help

Sample File Information

▼ Run Information

User Name:	TD0000655	Run Date:	Thu, Mar 09, 2006
Instrument:	ABI PRISM 310	Start Time:	10:48:14 PM
Data Coll. Version:	ABI PRISM 310 Collection 3.0.0	Run Duration:	30 Mns 34 Secs
Tube:	D8	Total Points:	6496

▼ Data Collection Settings

Module File:	GS STR POP4 (1 mL) F.md4	Run Voltage:	15000 Volts
Matrix File:	CBS_Matrix_030806_310A.mtx	Inj. Voltage:	15000 Volts
Parameters:	GS600Analysis.gsp	Inj. Duration:	5 Seconds
Size Standard:	GS 500 All.szs	Temperature:	60 °C
		Laser Power:	9 mWatts

▼ Gel Information

Gel Type:		Lot Number:	
Length-To-Detector:	36.0 cm	Expiration Date:	

▼ Sample Information

Sample Name:	POS_PP	
Dye	Sample Info	Comment
B:	POS_PP	
G:	POS_PP	
Y:	POS_PP	
R:	POS_PP	

▼ Analysis Records

▼ B: Analyzed 7:51:53 AM Fri, Mar 10, 2006

Parameters:	Standard: DGS 500 All.szs	Peak Totals
Analysis Range: 3500 - 6495 pts	Dye Std: R	Found In:
Baselined: Yes Multi-Componented: Yes	Size Method: Local Southern Method	- Sample: 25
Data Smoothing: Light	Size Range: 75 - 400 bps	- Dye Std: 11
Peak Detection Threshold: 50	Std Peak Det. Threshold: 50	Std Defined: 10
Peak Detection Min. Half-width: 2	Baseline Window Size: 51	Std Matched: 9
Polynomial Degree: 3 Peak Win: 15	Slope Threshold for Pk Start: 0.0 End: 0.0	

▼ G: Analyzed 7:51:53 AM Fri, Mar 10, 2006

Parameters:	Standard: DGS 500 All.szs	Peak Totals
Analysis Range: 3500 - 6495 pts	Dye Std: R	Found In:
Baselined: Yes Multi-Componented: Yes	Size Method: Local Southern Method	- Sample: 83
Data Smoothing: Light	Size Range: 75 - 400 bps	- Dye Std: 11
Peak Detection Threshold: 50	Std Peak Det. Threshold: 50	Std Defined: 10
Peak Detection Min. Half-width: 2	Baseline Window Size: 51	Std Matched: 9
Polynomial Degree: 3 Peak Win: 15	Slope Threshold for Pk Start: 0.0 End: 0.0	

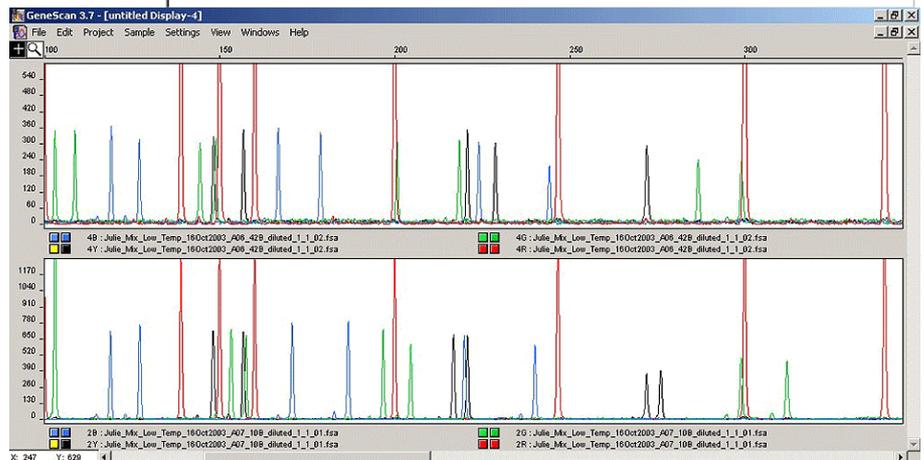
▶ Y: Analyzed 7:51:54 AM Fri, Mar 10, 2006

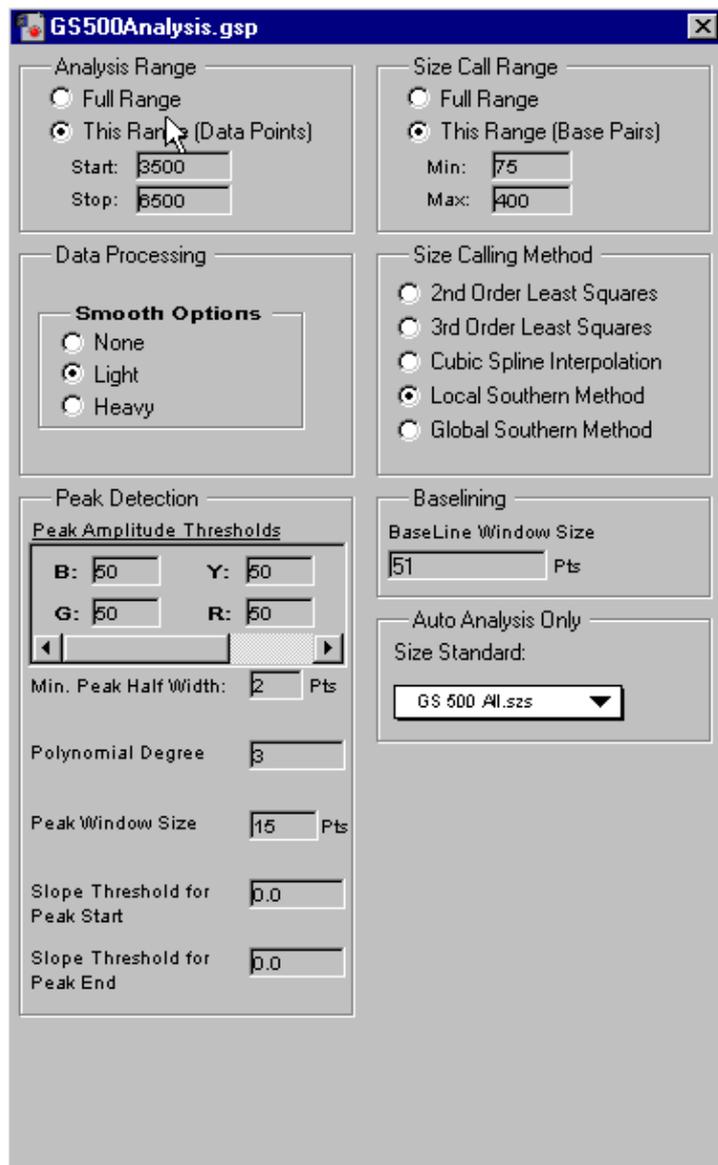
▶ R: Analyzed 7:51:54 AM Fri, Mar 10, 2006

- Analyzed data

Analysis Parameters

Analysis parameters for the GeneScan® 3.7 for Windows NT are configured prior to analyzing data. Two fields that are commonly altered are Analysis Range and Peak Detection.





Size Calling Methods

The analyst selects one of four possible size calling methods (in the analysis parameters) utilized by GeneScan®:

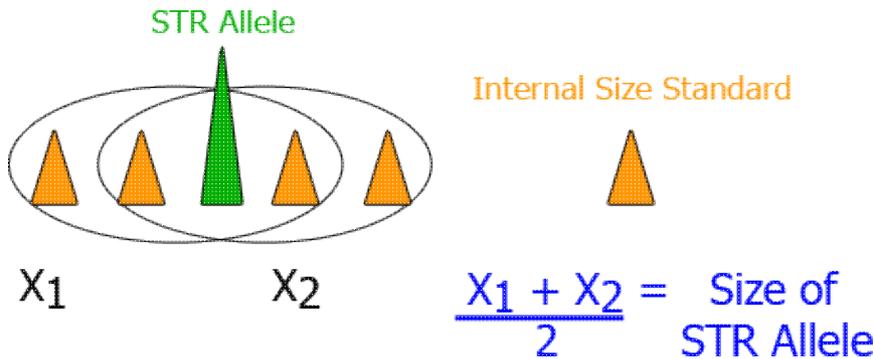
1. Local Southern
2. 2nd (or 3rd) Order Least Squares
3. Cubic Spline
4. Global Southern method

The Local Southern method for size calling is the most commonly used algorithm in forensic DNA analysis. It determines the sizes of fragments by using the reciprocal relationship between fragment length and mobility. The unknown fragment is surrounded by two known sized fragments above and one below then two below and one above. The results are averaged and the size of the allele is determined.

For a review of size calling methods 2-4 listed above, reference the GeneMapper® *ID* software version 3.1

User's Guide.02

Below is a visual representation of the sizing process of the Local Southern method.



In variable-temperature environments, some studies have found that the Global Southern method (rather than the Local Southern method), may provide better sizing precision.[03](#), [04](#)

One important issue concerning the sizing of AmpFSTR® GS500 size standards (ROX or LIZ) is the de-labeling of the 250 base pair fragment. At a minimum, the 250 fragment should not be labeled.[04](#)

Read more about delabeling elsewhere in this PDF file.

The size calling method is one component of the overall parameter settings that are configured prior to analyzing data.

Smoothing

The purpose of smoothing is to reduce the number of false peaks detected by the software. Prior to GeneMapper® ID 3.2, the Macintosh GS software smoothed before the data were analyzed and the GS software for Windows NT smoothed after analysis. This particular difference in the smoothing algorithm between Macintosh and Windows NT versions produced slightly different peak heights for the same sample files. When working with an established threshold, this could cause differences in allele designations.

The algorithm used for GS 3.7.1 for Windows NT (the updater) also tends to increase baseline noise, but operating with GS 3.7 for Windows NT gave similar results for peak height as that of the Macintosh version. A thorough discussion of each of the parameter settings shown above can be found in the User Bulletin of the Windows NT software for GeneScan®.[05](#)

Genotyper®



Genotyper® converts GeneScan® sized peaks into genotype calls using predefined macros, providing defined results. Genotyper® uses the tabular data from GeneScan® to make allele calls using the first ladder

STR Data Analysis and Interpretation for Forensic Analysts

recognized by the program. The two main manufacturer macros for the proprietary kits are Kazam for analysis of AmpFSTR® amplification kits and PowerTyper™ for Promega products.

These macros do three things:

- Calculate the bin offsets (alleles) based on the tabular data for sizing
- Filter stutter as defined in the macro
- Assign the number(s) that represent the genotypes for the profile (based on the ladder used for the sample set)

The macro is simply a step list of actions that are performed sequentially as defined when a particular template is launched. Macros are used to check the size standard by attaching the labeled sizes for confirmation; others can be written to label peaks and to set up Genotyper® tables. After assignment of allele labels, the analyst proceeds with the interpretation of the overall profile.



[Click here to view Genotyper's macros window.](#)

GeneMapper® ID v3.2



GeneMapper® ID (GMID) is an automated genotyping program that combines the functions of the GeneScan® analysis software and Genotyper® software into one package. The software program designates peaks in electropherograms by sizing and makes allele calls through size comparisons to an allelic ladder.

GMID provides the flexibility to analyze data in either a Macintosh (classic analysis mode) or a PC (advanced mode). One difference for the advanced mode using the Windows NT format is in the smoothing algorithm; both classic and advance modes smooth before the data are analyzed.

The unique features in the GMID (v3.1 and v3.2) not previously present in the GS/GT software combination are presented below. 02, 06

Unique Features in the GMID (v3.1 and v3.2)

Unique Feature	Description
CODIS Export	The software can export results in a CODIS recognized format (cmf v1.0 and 3.0).
Process Component-Based Quality Values (PQV)	The PQV system automatically assigns values to the quality of the data in respect to sizing and allele calling. Poor quality samples are those below the user defined thresholds.

STR Data Analysis and Interpretation for Forensic Analysts

GeneMapper® ID Database

GeneMapper® ID database stores the following data:

Note: The database does not store individual sample files.

- Predefined and custom-designed size standard definitions
- Panel, marker (loci), and allele bin definitions
- Analysis methods
- Table profiles (for generating tabular reports)
- Saved projects with sizing and genotyping results
- Matrix files (310 and 377 instruments only)
- Plot settings

Automated concordance checks

- GeneMapper® ID software compares genotype concordance between overlapping loci among different AmpFL STR kits for the same sample(s), or concordance of genotype calls from duplicate amplifications or duplicate injections of the same sample.
- Positive and negative controls give the expected allele calls.

Export Combined Table format

- When exporting from the Samples view, you can now export samples that do not pass sizing along with samples that pass sizing. This feature combines columns from the sample table and the genotype table and exports them as a single table.
- There are two display options when exporting samples: one line per marker and one line per sample (This is similar to the Make Allele table in the Genotyper® Software.)

Changes to the electropherogram displays

- The software now provides the option to display labeled peak assignments for all size standards. The user can quickly identify peaks visually and perform a size precision test. The labeled peak assignments are printable.
- When switching from the "align by base pair" to the "align by data point" views for the x-axis, the labels associated with the peaks are now retained in both views.

Allelic Ladder

Genemapper ID uses the average sizes of the alleles between multiple-run allelic ladders to determine the allelic bin offsets (whereas Genotyper uses a single run ladder sample).

An important feature of this software, which is unique to the first release and carried through to subsequent versions, is the use of PQVs (process quality values). The PQV system is the first step in the direction of expert systems analysis. User-defined process quality values generate notifications to help provide confidence in allele calls and to aid in troubleshooting. It is important for analysts to read about and understand their

purpose and function in the user's manual. Analysts should experiment with settings to ensure that notifications correspond with laboratory procedures. Prior to implementation, laboratories must conduct proper validation. [02](#), [06](#)

Features and Procedures: Yfiler Kit

Features have been added in GeneMapper® *ID* Software v3.2 to facilitate analysis of the AmpF®STR® Yfiler™ PCR Amplification Kit.

Features and Procedures (v3.2)

Feature/Procedure	Description
New feature: Allele calling parameters for new marker repeat types	In GeneMapper <i>ID</i> Software v3.1, allele-calling parameters were only available for markers with tetranucleotide repeat motifs. In GeneMapper® <i>ID</i> Software v3.2, however, allele calling parameters are available for four marker repeat types: tri-, tetra-, penta-, and hexanucleotide. The four allele calling parameters appear in the Allele tab of the Analysis Method Editor. All related analysis values are entered directly into the Tri, Tetra, Penta, or Hexa column fields, which allows for viewing of all values.
New feature: Plus stutter filtering (for the DYS392 locus)	To aid in interpreting genotype profiles, two new fields have been added to the Allele tab of the Analysis Method Editor specifically to filter out the DYS392 plus stutter: <ul style="list-style-type: none"> • Plus Stutter Ratio • Plus Stutter Distance
Procedure: Workaround for the DYS19 locus (-2bp filtering)	Laboratories may choose to implement a work-around for the DYS19 locus by using the Minus A Ratio and Minus A Distance fields in the Allele tab of the Analysis Method Editor. These fields can be used to filter out the -2-bp stutter that is observed at the DYS19 locus.
Procedure: Creating HID analysis methods for the Yfiler kit	Use the Analysis Method Editor in GeneMapper® <i>ID</i> Software v3.2 to set analysis parameter values for analyzing the Yfiler kit data.
Procedure: Creating a table setting and uploading exported haplotype(s) for searching profiles with the Yfiler™ Haplotype Database	Using the Table Setting Editor, a table setting is created in GeneMapper® <i>ID</i> Software v3.2 to export haplotypes specifically for searching the Yfiler Haplotype Database for profile match estimation.

STR Data Analysis and Interpretation for Forensic Analysts

Genotyper® had a feature that allowed review of the sizing standard for precision assessment. This feature was not available in the first version of GMID (v3.1); it is included in version 3.2. However, it requires a work-around to achieve the precision statistics on the 250 base pair fragment. The size standards can be overlaid from all samples in the run. The base pair sizes for the fragments can be displayed, the table exported to Microsoft Excel, and the Excel functions used to calculate the 250 base pair precision from a run.

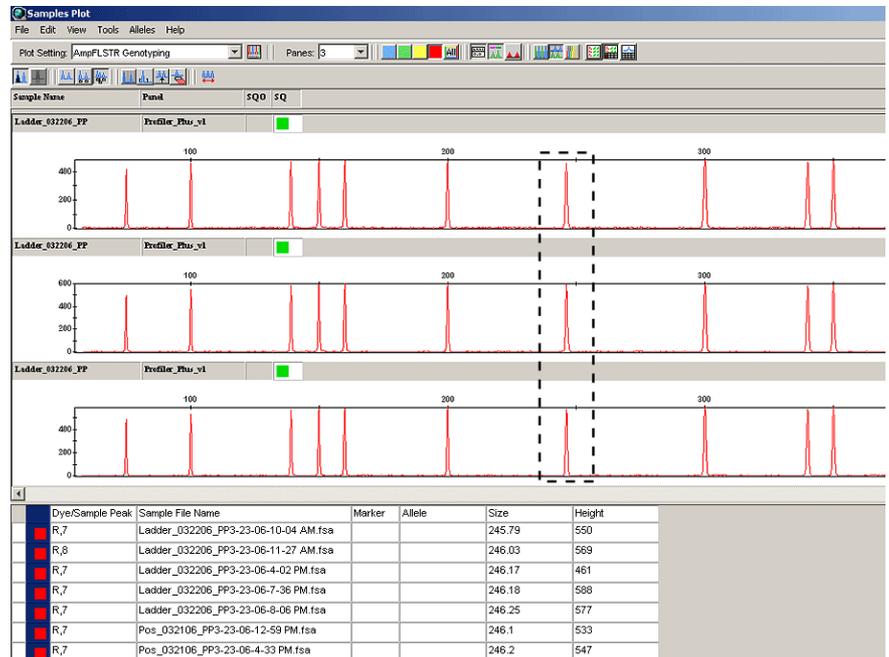
At the right is a screen capture of the size standard overlay with the sizing table beneath:

Analysis Modes

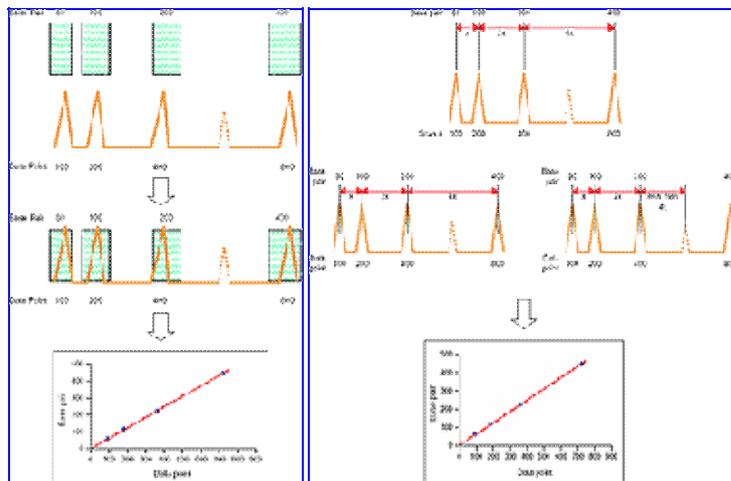
As previously mentioned, there are two analysis modes (classic Macintosh and advanced Windows NT modes). The differences between these modes are found in the sizing method and the flexibility of peak sensitivity settings.

Specifically the mode differences are:

- In the classic mode, size calling is performed by matching the actual size standard fragments of the sample with a defined size standard that must be accurately labeled; it utilizes scan number to assign sizes.
- In the advanced mode, size calling is performed using a function known as "ratio matching." Ratio matching uses an algorithm to determine the distance between the size fragments based on the entry of a set of size fragment values, where it uses the relative distance between the neighboring peaks to size.



The following excerpts from the Applied Biosystems User Bulletin on Size Parameters explain these differences.⁰⁵



STR Data Analysis and Interpretation for Forensic Analysts

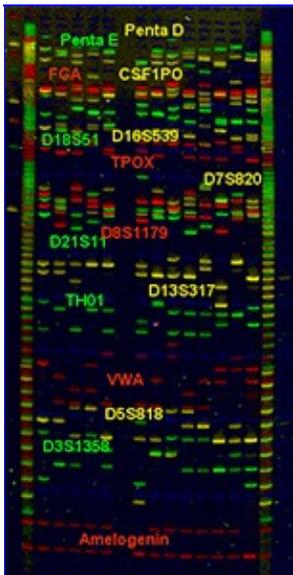
There are similarities between GMID and GS/GT software. However, GMID offers additional features, added flexibility, and efficiency through the combining of the programs. The expert system potential, inherent in the software, will continue to develop as new versions are released.

Note:

To become familiar with the use of GMID 3.2, it is recommended that analysts read through the User's Manual for v3.1 and User Bulletin for v3.2; focusing particular attention on the verification process and the software features and functions.07

FMBIO® Analysis Software

FMBIO® Analysis Software is a program that is associated with slab gel electrophoresis. This program uses the Macintosh operating system.

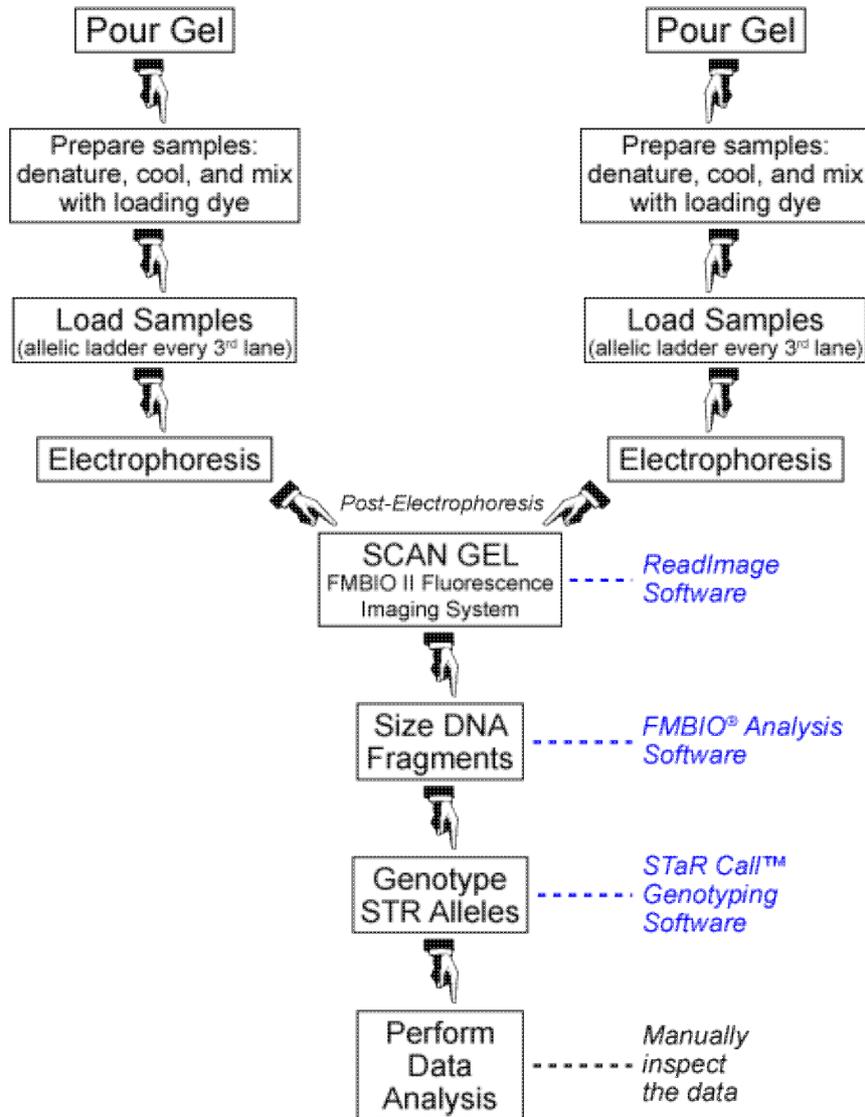


Before using the FMBIO® Analysis Software, the analyst must use the read image program to define the scan area and resolution. It is similar to operating a desktop scanner. Each image captured by the read image program is converted to a Tag Image File Format (TIFF) file.01

Following scanning, the FMBIO® Analysis Software is utilized to size the bands and quantitate peak height and area. This functionality is equivalent to the GeneScan® process discussed previously. The FMBIO® Analysis Software also uses algorithms to size and quantitate. The software includes a band-finding program to aid in the identification of possible DNA fragments, but user review is required.

The following figure depicts the FMBIO® procedure, including the applicable software processes:01

STR Data Analysis and Interpretation for Forensic Analysts



FMBIO® STaRCall™

STaRCall™ is the software program that genotypes gel electrophoresis data for the Hitachi FMBIO® system. Sizes from the FMBIO® Analysis Software are imported into STaRCall, similar to the process used on GS/GT software on the capillary electrophoresis platform. This program is similar in process to Genotyper®.

Allelic ladders are used to size the fragments from adjacent lanes; the closer the ladders are to the samples being run, the more precise the sizing. Where Genotyper® has a bin set associated with its macros, STaRCall™ has STR look-up tables maintained in Microsoft Excel. Optical density units are used to quantitate the fragments as the data are reviewed.

STR Data Analysis and Interpretation for Forensic Analysts

An example of the spreadsheet used to determine genotypes:
Data Troubleshooting

Introduction

PowerPlex 2.1 (TMR)						STR OD/IOD Cutoff Table			
Ranges		Genotype	-	+	Paste Values	Expected	Locus	Cutoff (%)	
From	To							From	To
-1.00	1.00	FGA	46.2	1.00	1.00	444	FGA	0	20
-1.00	1.00	FGA	45.2	1.00	1.00	440	TPOX	0	20
-1.00	1.00	FGA	44.2	1.00	1.00	436	D8S1178	0	20
-1.00	1.00	FGA	43.2	1.00	1.00	432	vWA	0	20
-1.00	1.00	FGA	31.2	1.00	1.00	384			
-1.00	1.00	FGA	30	1.00	1.00	378			
-1.00	1.00	FGA	29	1.00	1.00	374			
-1.00	1.00	FGA	28	1.00	1.00	370			
-1.00	1.00	FGA	27	1.00	1.00	366			
-1.00	1.00	FGA	26	1.00	1.00	362			
-1.00	1.00	FGA	25	1.00	1.00	358			
-1.00	1.00	FGA	24	1.00	1.00	354			
-1.00	1.00	FGA	23	1.00	1.00	350			
-1.00	1.00	FGA	22	1.00	1.00	346			
-1.00	1.00	FGA	21	1.00	1.00	342			
-1.00	1.00	FGA	20	1.00	1.00	338			
-1.00	1.00	FGA	19	1.00	1.00	334			
-1.00	1.00	FGA	18	1.00	1.00	330			
-1.00	1.00	FGA	17	1.00	1.00	326			
-1.00	1.00	TPOX	13	1.00	1.00	290			
-1.00	1.00	TPOX	12	1.00	1.00	286			
-1.00	1.00	TPOX	11	1.00	1.00	282			
-1.00	1.00	TPOX	10	1.00	1.00	278			
-1.00	1.00	TPOX	9	1.00	1.00	274			
-1.00	1.00	TPOX	8	1.00	1.00	270			
-1.00	1.00	TPOX	7	1.00	1.00	266			



Short tandem repeat (STR) data analysis and

interpretation in forensic DNA casework relies on an analyst's professional judgment and expertise. Procedures for analysis and interpretation are based on validation studies, published literature, population studies, and casework experience. Laboratory interpretation procedures are not meant to cover every situation but rather to establish a minimum standard for the interpretation of analytical results. Documented procedures contribute to objectivity and consistency and ensure that the conclusions drawn in casework are scientifically supported by analytical data.

A number of factors can introduce ambiguity into the analysis and interpretation process. Analysts must be able to identify and troubleshoot these ambiguities and understand their implications.

Objectives

Upon successful completion of this unit of instruction, the student shall be able to:

- List commonly used STR data interpretation parameters.

STR Data Analysis and Interpretation for Forensic Analysts

- Explain thresholds, spurious peaks, pull-up, stutter, microvariants, degradation, stochastic effects, allele dropout, mutations, controls used in troubleshooting STR data.

Data Interpretation and Troubleshooting

Interpretation of genotypes is based on a pattern of peaks or bands, which are visual representations of DNA fragments. Laboratories use data compiled from validation studies to establish procedures for data analysis and interpretation. These validation studies provide data so that the laboratory can establish interpretation guidelines. The following are interpretation parameters that validation studies establish:

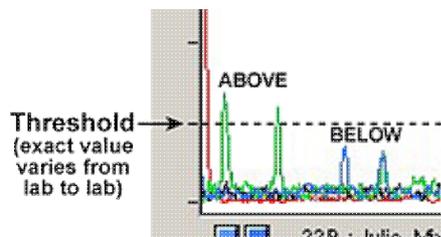
- Sensitivity
- Reproducibility
- Precision
- Heterozygosity
- Mixture assessment

Before analyzing and troubleshooting STR data, the analyst must understand the methods and the issues inherent to the analysis process. The ability to differentiate spurious peaks and artifacts from alleles is imperative.

Thresholds

When the quantity of DNA being analyzed is very low, it may be difficult to distinguish true low-level peaks from technical artifacts, including noise. Consequently, most forensic laboratories have established peak-height thresholds for "scoring" alleles. Only if the peak height, expressed in relative fluorescence unit (RFU), exceeds a standard value will it be accepted.

There are no firm rules for establishing threshold values. Each laboratory must set its own as part of its validation procedure. The threshold may be determined experimentally on the basis of observed signal-to-noise ratios, or may be arbitrarily set to a level established by manufacturers or published data. Applied Biosystems, Inc. (ABI), which sells the most widely used systems for STR typing, has recommended a peak-height threshold of 150 RFU, saying that peaks below this level must be interpreted with caution. However, many crime laboratories that use the ABI system have set lower thresholds based on their own studies, typically 50 to 100 RFU.



The lower threshold is a measure of the sensitivity of the procedure. Most laboratories establish both lower and upper thresholds for data interpretation, thereby establishing a window to interpret data.

A laboratory's threshold can be influenced by a variety of factors. For example, there are sensitivity differences between the types of instrumentation (e.g., capillary electrophoresis (CE) instruments and slab-gel instruments) and within any one type of instrument (e.g., between different ABI 310 instruments).

STR Data Analysis and Interpretation for Forensic Analysts

Many laboratories have noted the varying sensitivities of instruments, which tend to be more sensitive and have better resolution than gel-based systems. Some laboratories have established thresholds within their laboratory that vary depending on the sensitivity of the specific instrument.

The upper threshold is crucial when reviewing data from high quantity DNA samples. Samples with high quantities of amplified DNA will have high RFU values that can oversaturate the instrument's ability to detect the sample. This can lead to difficulty in interpretation because an accurate measurement with respect to the peak heights and/or areas may not be obtained. This can be especially problematic when working with mixed samples.

Note:

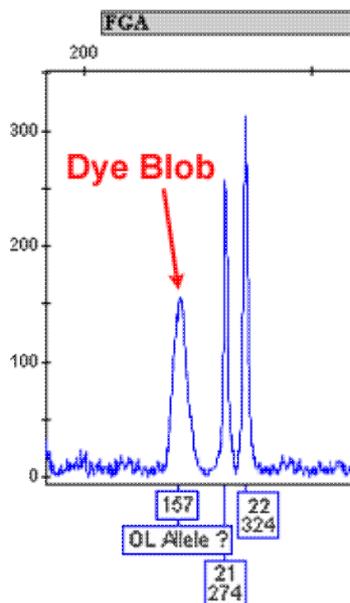
The instrument's software analyzes and makes sizing determinations based on the lower threshold established by the laboratory. Some laboratories have adopted a procedure to evaluate data below their RFU threshold in an attempt to interpret data for exclusionary purposes. For instance, a laboratory can have a lower threshold at 100 RFU. If additional data is detected below that threshold (e.g. 50 to 99 RFU), it could be evaluated for exclusionary purposes.

Spurious Peaks

Chemistry and instrument-related issues can lead to spurious peaks. Spurious peaks, also referred to as artifacts, include dye blobs, spikes, and noise; they may be difficult to differentiate from alleles.

In general, spurious peaks are not reproducible. Samples can be reinjected to determine whether suspected artifacts are reproducible or analysts can rely on their professional experience to differentiate alleles from spurious peaks.

Dye Blobs

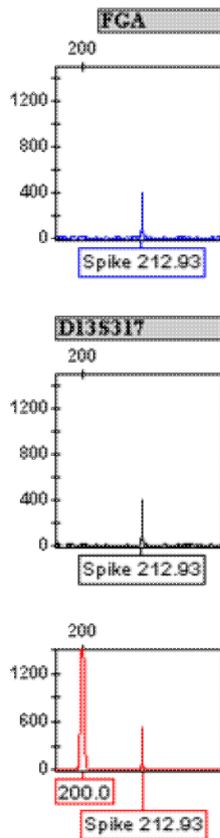


Disassociated primer dyes, more commonly referred to as dye blobs, are fairly common in STR analysis. While it isn't entirely understood why dye blobs occur, evidence suggests that the fluorescent dye tags attached to the primers begin to break down over time.⁰¹ Disassociated primer dyes can show up in the sample analysis range and can mask true data. Dye blobs are usually wider than real peaks and are typically only seen in one color. When the manufacturer's specifications for storage of amplification kits are followed,

problems with disassociated primer dyes can be avoided.

If problems persist, the sample can be reamplified or a filter unit clean-up step (e.g., [link to Microcon® 100](#)) can be performed. Amplified fragments of DNA will attach to the membrane while the disassociated primer dyes pass through the membrane and are filtered out of the sample.

Spikes



Spikes are narrow peaks usually attributed to fluctuation in voltage or the presence of minute air bubbles in the capillary. Spikes can also be caused by crystals in the polymer and/or fluorescent material in the polymer or formamide. Spikes, unlike other artifacts, are generally seen in the same position in all colors. However, it is possible to detect spikes in a single color.

Analysts should view both the raw and analyzed data for each sample. The raw data produce a non-filtered view of the sample run, while analyzed data can obscure detection of a spike. Frequently, spikes are more readily detected in raw data versus analyzed data. The analyzed data provide an exact data point and base pair size for each peak. Spikes that are not obscured in the analyzed data can be assigned an exact data point for each color displayed.

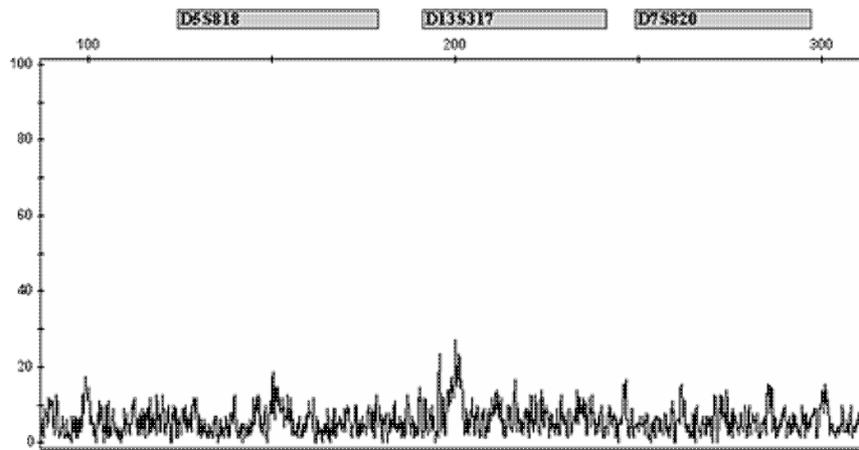
The occurrence of spikes can be minimized by following the instrument manufacturer's procedures for reagent and sample handling. If spikes persist, analysts may need to contact the manufacturer. Frequent electronic spiking can occur due to poorly functioning instruments.

STR Data Analysis and Interpretation for Forensic Analysts

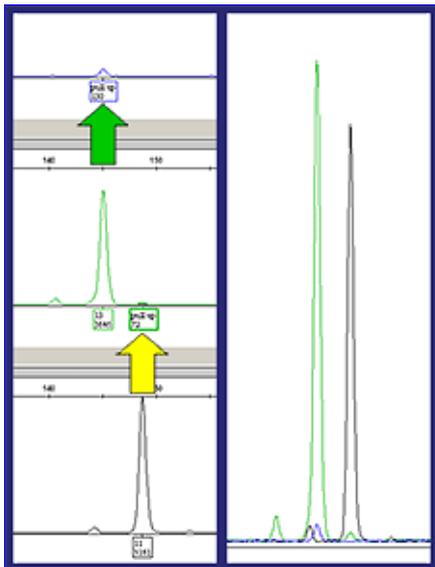
Noise

Noise describes a series of non-reproducible background peaks that occur along the baseline in all samples. A wide variety of factors, including amplified current fluctuations within the electronic circuitry, air bubbles, urea crystals, and sample contamination, can create noise. If large enough (close to the laboratory threshold), they may be confused with an allele or mask alleles.

Analysts may confuse actual alleles with noise and vice versa; they should be familiar with the signal-to-noise ratio of the instrument and/or the specific data. Data interpretation should include viewing both raw and analyzed data to assess the signal-to-noise ratio and distinguish real data from noise. Noise is not reproducible; one way to differentiate alleles from noise is to rerun the sample.



Pull-up



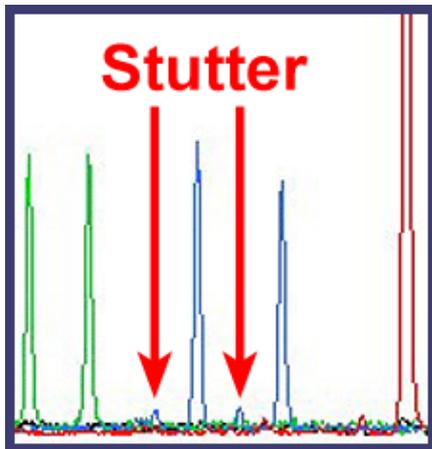
Pull-up, sometimes referred to as bleed-through, represents a failure of the analysis software to discriminate between the different dye colors used during the generation of the data. Oversaturated data can also cause the dyes to "bleed" over or pullup into another color.

If pull-up occurs, the analyst can inject less of the sample or re-amplify the sample with less input DNA.

Reoccurring pull-up (due to too much DNA) may indicate that the quantitation method or the amount of DNA used for amplification should be reevaluated. If this problem is not due to too much DNA, it may be necessary to run a new matrix and apply it to the sample.

A signal from a locus labeled with blue dye, for example, might mistakenly be interpreted as a yellow or green signal, thereby creating false peaks at the yellow or green loci. Pull-up can usually be identified through careful analysis of the position of peaks across the color spectrum, but there is a danger that pull-up will go unrecognized, particularly when the result it produces is consistent with what the analyst expected or wanted to find.

Stutter



Stutter is a by-product of the amplification of STR loci whereby a minor product one repeat smaller than the primary allele is generated. Sequence analysis of stutter products of STR loci has shown that the product is missing one core repeat unit relative to the main allele.^{Q2} Although the mechanism is not entirely understood, stutter occurs in a reproducible and predictable fashion. The proportion of the stutter product relative to the main allele (percent stutter) is measured by dividing the height (or area) of the stutter peak by the height (or area) of the main allele peak.^{Q1}

Typically, stutter is affected by:

- The repeat unit length (2 base pair repeats have higher stutter than 3 basepair, etc).
- The degree of homogeneity of repeats (the more homogenous, the higher the stutter).
- The length of the allele within a locus (the larger the alleles have higher stutter).

In known single source samples, stutter is identifiable by its size and position. However, with mixed samples, stutter and alleles can overlap, complicating interpretation.

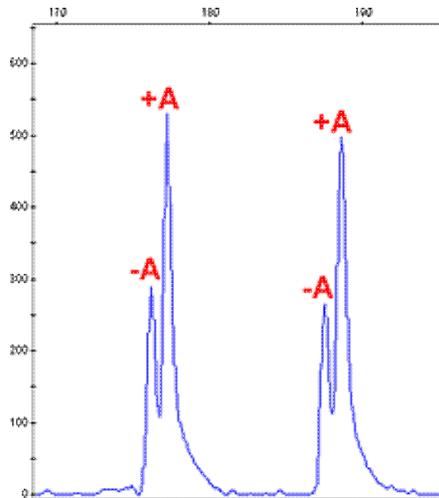
Read more about interpretation elsewhere in this PDF file.

The scientific community, as well as individual forensic laboratories, has conducted validation studies to determine the expected range of stutter percentages. In general, stutter percentages do not vary significantly.

There are two cases in which variability in the stutter percentages can be seen:

- Low-level samples (low RFUs)
- Samples exceeding the detection level of the instrument (excess DNA)

Non-Template Addition



If too much input DNA is added to an amplification reaction, the polymerase may be unable to complete the extension for all amplicons.⁰³ Non-template addition results in a PCR (Polymerase Chain Reaction) product that is one base pair longer than the actual target sequence. When the polymerase is unable to complete the adenine addition on all products, this results in what is commonly referred to as split peaks (+A/-A peaks).⁰¹ To minimize split peaks, the extension phase of the PCR process is designed to drive the addition of adenine, ensuring that all amplicons are the same length. Kit manufacturers have also developed primer sequences that encourage adenine addition.

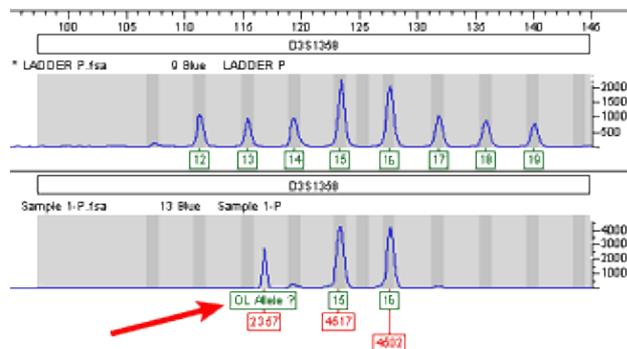
Read more about PCD in the Crime Scene and DNA Basics for Forensic Analysts PDF file.

Note:

Samples displaying -A can be diluted with buffer and reextended in the thermal cycler.

Microvariants and OL Alleles

Allelic ladders represent the most common alleles at each locus and were established through the evaluation of data from several hundred individuals. Alleles within the STR loci are known to vary greatly between individuals, and the kit ladders do not represent all possible types. Alleles that size outside allele categories represented in the ladder are often referred to as off ladder (OL) alleles.



In general, proprietary sizing software is designed to designate allele types if the allele size is within one of the allele categories defined by the ladder. The software designates an allele that falls outside of these allele categories as off ladder unless the manufacturer has established virtual allele categories. Virtual alleles are

STR Data Analysis and Interpretation for Forensic Analysts

alleles that have been previously characterized, but are not present in the allelic ladder. For all defined virtual alleles, the software designates the allele type, rather than designating it as an off ladder allele.

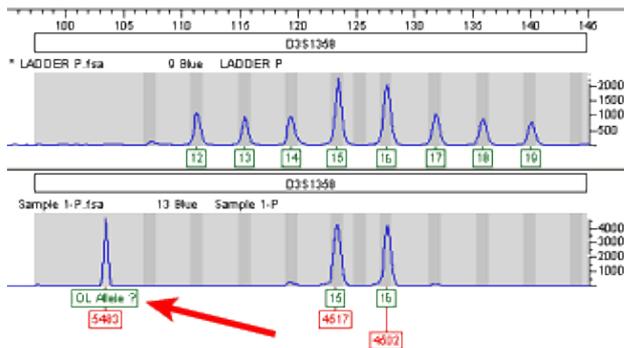
It is not unusual in forensic autosomal STR testing to see microvariants.⁰⁴ A microvariant represents an incomplete repeat for a given allele. For instance, at D18S51 a 15 allele designation means that there are 15 AGAA's along the fragment. However, a 14.2 allele designation means that there are 14 AGAA repeats along the fragment and an additional 2 bases: AG. Microvariants are reported as the number of complete repeat units and are designated as an integer (e.g., 14). Any partial repeat is designated as a decimal, followed by the number of bases in the partial repeat (e.g., ".2").

Many microvariants are represented in allelic ladders or have virtual allele categories within the software program; those that do not have established categories are designated as off ladder alleles.

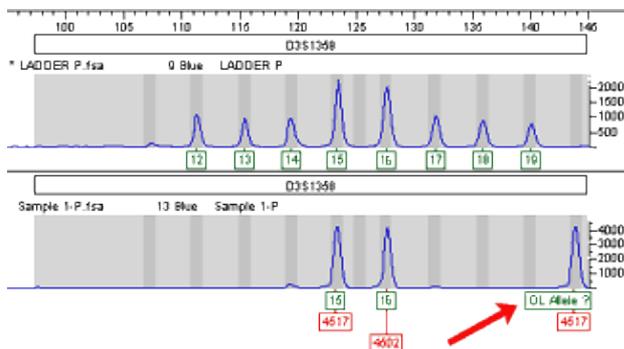
While off ladder alleles have been well documented with forensic STR testing, some may not have been previously characterized. The [National Institute of Standards and Technology \(NIST\) website](#) has a listing of off ladder alleles and can be used as a reference in these instances. If it is determined that an allele has not been characterized, it may be advisable to rerun the sample to confirm the type.

The most common approach for reporting alleles that size higher or lower than the allelic ladder range is as follows:

- Alleles that are less than the lowest allele on the specific ladder are reported as "less than X."



- Alleles that are greater than the largest allele on the specific ladder are reported as "greater than X."



Degradation

DNA degradation is a process by which DNA breaks down into smaller fragments. Environmental factors such as sunlight, heat, and humidity can increase the rate of degradation. DNA samples that are subjected to

STR Data Analysis and Interpretation for Forensic Analysts

environmental factors that promote degradation can pose challenges for data interpretation.

As DNA molecules randomly break down into smaller fragments, the STR regions of the DNA molecule can be fractured. If the STRs do not stay intact, amplification of these regions will not be successful. Degradation is more likely to occur at a large STR locus before occurring in a smaller STR locus.

Generally, degradation can be easily identified because the peak heights exhibit a downward slope across the electropherogram. The process of degradation can reduce the height of some alleles, making them too low to be distinguished from background noise in the data. In severely degraded DNA samples, no results will be obtained. Two or more biological samples that make up a mixture may show different levels of degradation, which can complicate the interpretation of these samples.

[View an animation that further explains partial profiles.](#)

Stochastic Effects

Stochastically induced heterozygote imbalance may be observed in STR analysis due to the effective low copy number of DNA templates in degraded DNA.

The amplification process can produce many copies from a relatively low quantity of DNA. If too small a quantity of DNA is introduced at the beginning of amplification, it is possible that heterozygous alleles may amplify differentially. The first few cycles of the amplification process are extremely important; if imbalanced amplification occurs, it will result in stochastic effects.

An analyst can assess the data by calculating the heterozygosity of alleles. If the heterozygosity is less than 70%, this could indicate a mixture and/or stochastic amplification.

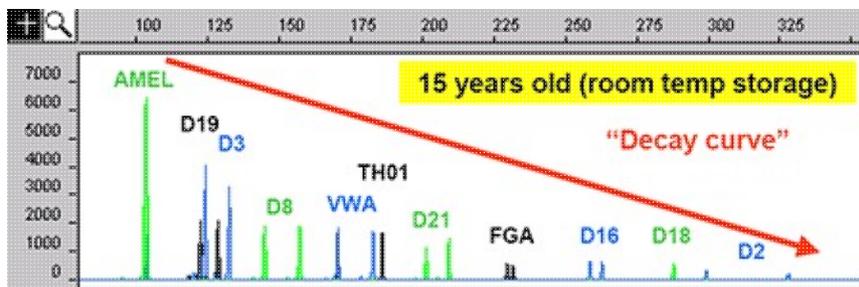
[View an animation that shows an example of the stochastic effect.](#)

Allele Dropout

Allele dropout occurs when a sample is typed and one or more alleles are not present.

This can be due to a variety of factors: $i_c^{1/2}$

- The initial input quantity of DNA is too low, resulting in the failure to amplify one or more alleles in the sample.
- A mutation in the primer binding site is present, which causes a failure in the amplification of the allele.
- An allele sizes outside of the normal calling range for a particular locus and goes undetected.



Mutations

A mutation occurs when DNA is damaged or changed anywhere along the DNA strand. Mutations that occur in primer-binding site regions or along the STR region can complicate interpretation.

The following are types of mutations that can occur:

Types of Mutations

Mutation	Description
Insertion	A base is inserted or added to the original DNA strand
Deletion	A base is deleted or removed from the original DNA strand
Transition	A transition between purine or pyrimidine bases occur in the original DNA strand <ul style="list-style-type: none">• purine: (adenine guanine)• pyrimidine: (cytosine thymine)
	Transitions are more common than transversions
Transversion	A transversion between purine and pyrimidine bases occur in the original DNA strand <ul style="list-style-type: none">• guanine cytosine• guanine thymine• adenine cytosine• adenine thymine

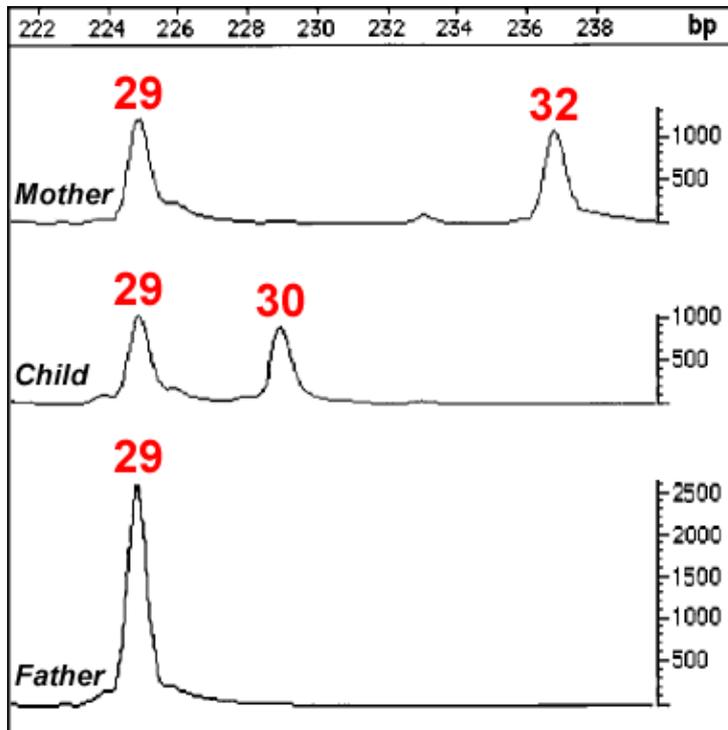
Mutation & Paternity

Mutations occurring during meiosis can affect the interpretation of paternity tests. A mutation during meiosis may result in discrepant results at a locus. For instance, a mother has a 14, 15 type at D3S1358. The assumed biological father has a 17, 18 type at D3S1358. The offspring has a 15, 19 type. When data at this one locus are compared, it appears that the child's assumed biological father can be excluded because the male could not have passed the 19 allele on to the child. During meiosis it is possible for an additional tetranucleotide repeat to be added to the DNA strand. Therefore, the male could have passed along a 19 allele, even though he doesn't have that allele type, if this mutation takes place.

Since mutations occur at known rates,⁰⁵ it is important to develop a procedure to effectively deal with them in paternity cases. Many researchers in the field suggest using a two locus exclusion for STR paternity testing.⁰⁶ Since mutations are relatively rare for the core STR loci, it is unlikely that two mutations would occur during meiosis at two different loci.

Somatic mutations have been known to occur at STR loci used in forensic testing. A somatic mutation occurs within somatic cells and are not inherited.

For example, a somatic mutation can occur in early embryonic development within an STR locus. Although the embryo inherits an 8 allele at D7S820 from its mother and a 9 allele from its father, the mutation results in an additional 10 allele. Therefore, the child will exhibit a tri-allelic pattern at D7S820. A mutation of this type can look like a mixed DNA sample or a contamination event at that locus. After a review of data from other loci, these concerns can be dismissed. It might be advisable to retest the sample or possibly other tissue types from the individual to confirm the DNA profile.



An example of a mutation in the STR locus D21S11 of the father causing allele mismatch in the child. Note the presence of allele 30 in the child, and the peak height of all of the alleles of the mother, the suspicious father, and the child (shown in the vertical scale).

A somatic mutation can occur after tissue differentiation begins in an embryo's development. In this case (for example), it is possible for the DNA profile obtained from a buccal swab to be different from that obtained from a hair or blood sample. If this mutation occurs early in the embryo's development, then it is more likely to be the common type throughout all of the tissues.

Null Alleles

A null allele is an allele that is present in a sample, yet is not amplified. A primer-binding site mutation can inhibit amplification for that allele and result in a null allele. If an individual is heterozygous and has a primer-binding site mutation for one of the alleles, the individual would type as a homozygote.

Note:

The manufacturers of various STR typing kits use different primer sets. If a DNA sample has a mutation in a primer-binding region specific to kit A, but no mutation in the primer-binding region specific to kit B, a rare discordance in allele calls can occur when comparing typing results produced by these two manufacturers' kits.⁰⁷

When comparing DNA typing results from different kits, null alleles (due to primer-binding site mutations) can result in discrepant DNA types at a particular locus. It is important to understand that although null alleles are rare, they must be considered when interpreting potential matches.

Controls

Incorporating controls in STR DNA analysis is an essential and required part of the testing process.^{08, 09} Introducing controls at each step of the DNA process allows the analyst to identify and troubleshoot possible

STR Data Analysis and Interpretation for Forensic Analysts

issues and ensure that the methods used produce accurate and reliable results.

Extraction

An extraction blank is included in the extraction process to assist the analyst in determining if the reagents and/or techniques used may have introduced contamination. The extraction blank should be treated in exactly the same manner as the other samples.

During the amplification process two controls are used:

- A positive control is a known DNA sample that has been previously typed and is added to the sample set. The positive control verifies that the analysis processes are functioning properly. (Manufacturers provide positive control samples in STR kits.)
- A negative control is included in the amplification process to assist the analyst in determining if the reagents and/or techniques used may have introduced contamination. The negative control should be treated in exactly the same manner as the other samples.

Capillary Electrophoresis

Samples are typed using capillary electrophoresis. Internal sizing standards (ISS) must be added to each sample.

Read more about capillary electrophoresis in the [Amplified DNA Product Separation PDF file](#).

Internal size standards (ISS) serve two purposes:

- Sizes of all fragments are established for a sample (relative fragment units). This is accomplished by using the ISS to establish a correlation coefficient, which is used to determine the size of sample fragments.
- The ISS serves as an effective control and provides information about instrument run conditions. For instance, if all of the peaks for the sizing standard are not present, it suggests a temperature, run time, or injection problem.

Laboratory temperatures can vary, and fluctuations occurring during a run can affect electrophoresis. Including allelic ladders with each run corrects for this issue. Allelic ladders must be included in all instrument runs for the instrument's software to properly type the sized fragments for each sample.

Note:

If a laboratory experiences relatively severe temperature variation within a run, it may be advisable to run several allelic ladders so that the run can be broken into smaller sizing projects.

Data Interpretation & Allele Calls



The interpretation of results in casework can be one of the most difficult aspects of forensic DNA analysis and is a matter of professional judgment and expertise.⁰¹ It is not possible or practical to develop interpretation criteria that cover every circumstance. Laboratories should develop interpretation guidelines that are based on validation studies, literature, methodology, and experience. It is expected that laboratory guidelines will evolve as the collective experience of the laboratory and the forensic science community grows. Interpretation guidelines provide a framework for the objective and consistent interpretations of results.

Objectives

Upon successful completion of this unit of instruction, the student shall be able to:

- Describe the factors that can influence the output from instruments used in DNA analysis
- Describe controls and other techniques used to validate instrument output
- Compare methods used to evaluate data and recognize analysis artifacts
- Explain methods used to analyze data from single source and mixed source samples

Overview

Interpretation of genotypes (allele calls) is based on a pattern of peaks or bands on an electropherogram or gel. A peak or band is a visual representation of a DNA fragment.

- In capillary electrophoresis (CE) a peak on the electropherogram rises sharply from the baseline, has smooth sides, and is symmetrical in shape.
- For gel-based systems, a band on the gel or gel image is distinguishable from the background and defined within a gel lane.

Read more about capillary electrophoresis in the [Amplified DNA Product Separation PDF file](#).

Gel-based systems, unlike CE, allow analysts to visually compare alleles without a software program. However, most laboratories use software programs for both CE and gel-based sizing.

STR Data Analysis and Interpretation for Forensic Analysts

The steps in data interpretation are:

1. Assess the internal size standards (ISS), allelic ladders, and controls.
2. Assess each sample for the presence of extraneous peaks and determine if they may interfere with the interpretation process.
3. Assess the data from each sample.

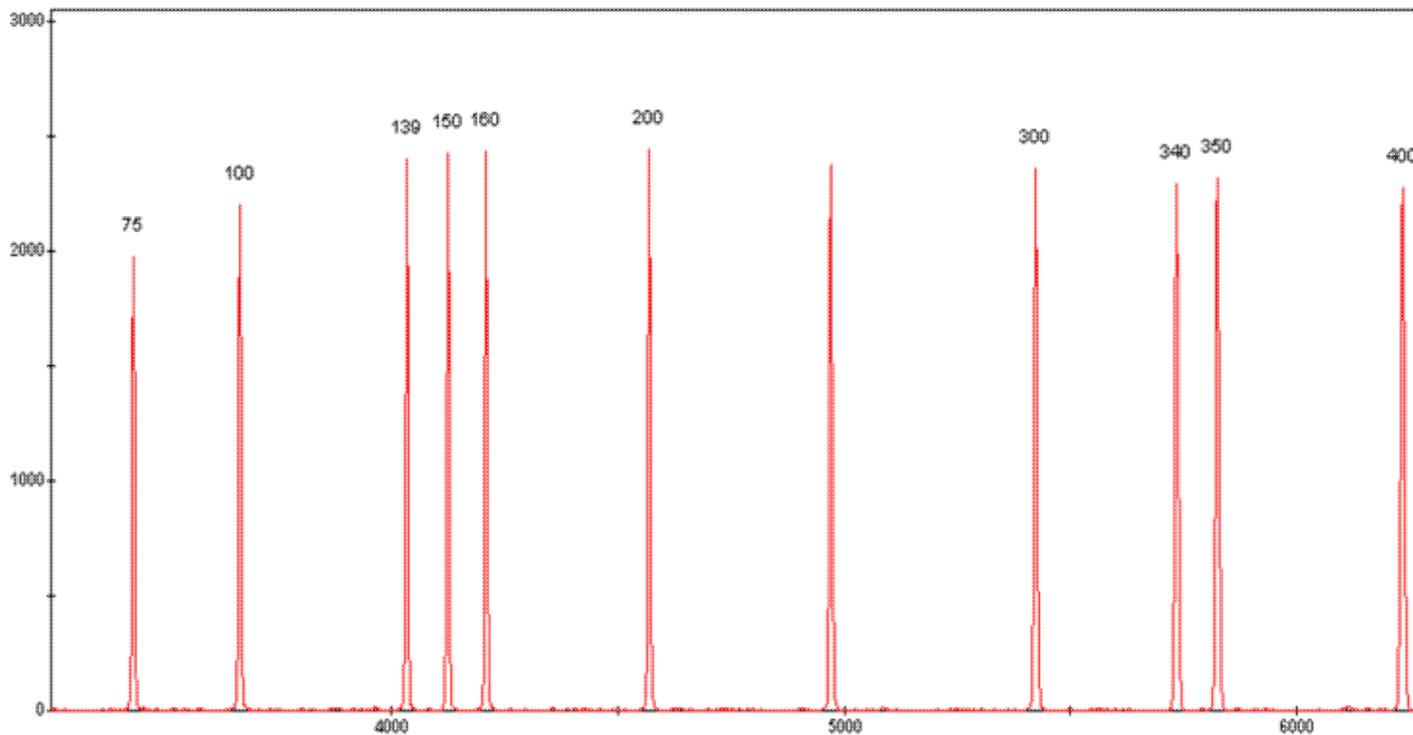
Step 1: Internal Size Standards and Allelic Ladders

Commercial kits for short tandem repeats (STR) typing include allelic ladders and internal size standards (ISS) that are added to each sample.

Read more about capillary electrophoresis in the [Amplified DNA Product Separation PDF file](#).

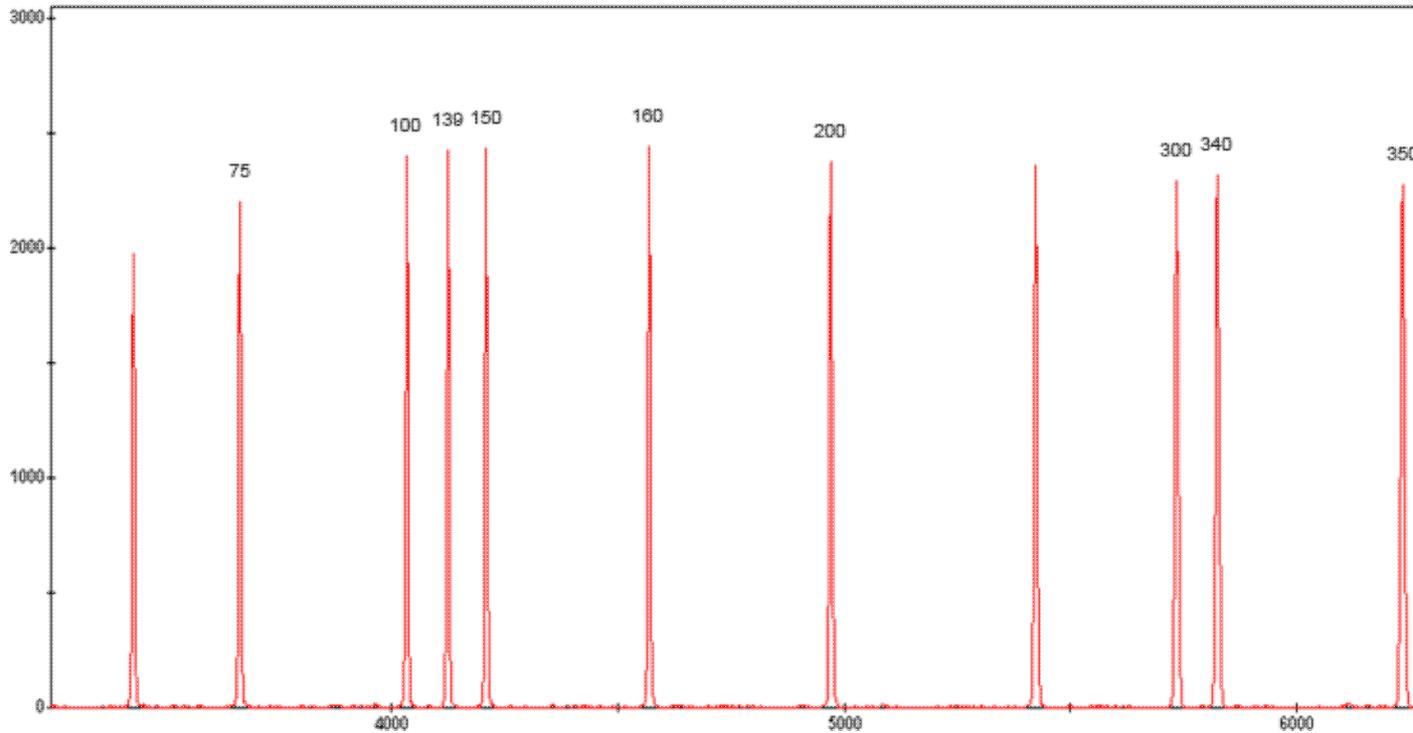
- Each sample from a run or gel should be assessed to determine if the ISS peaks have been called correctly. In general, the peaks or bands from an ISS are uniform in size or intensity. Lack of uniformity or miscalled peaks can indicate problems with the sample, injection, and/or run.

Below are examples of good and bad internal size standards:

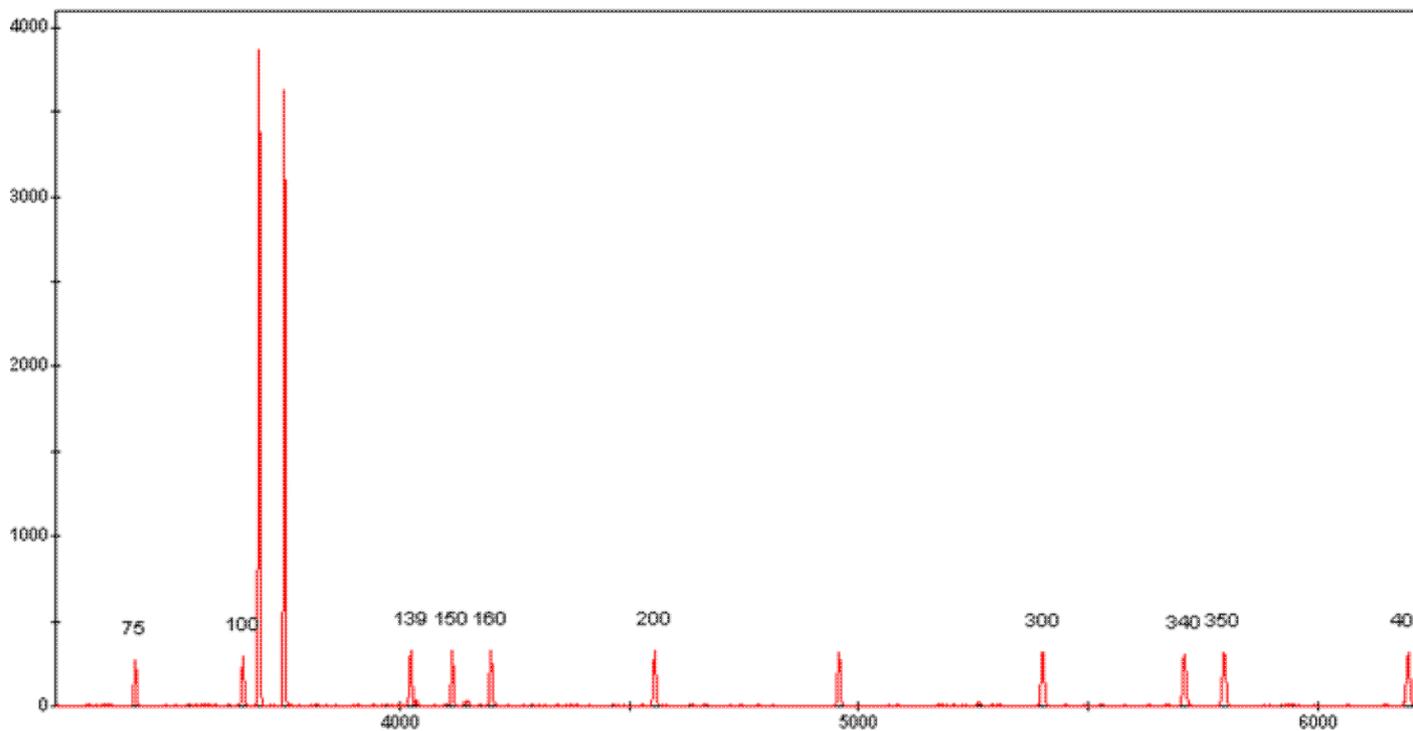


Above can be seen an example of a good internal size standard. Compare this to a miscalled ISS, below. Note that the first peak has not been called, and normal peak calls have begun with the second peak from the left.

STR Data Analysis and Interpretation for Forensic Analysts

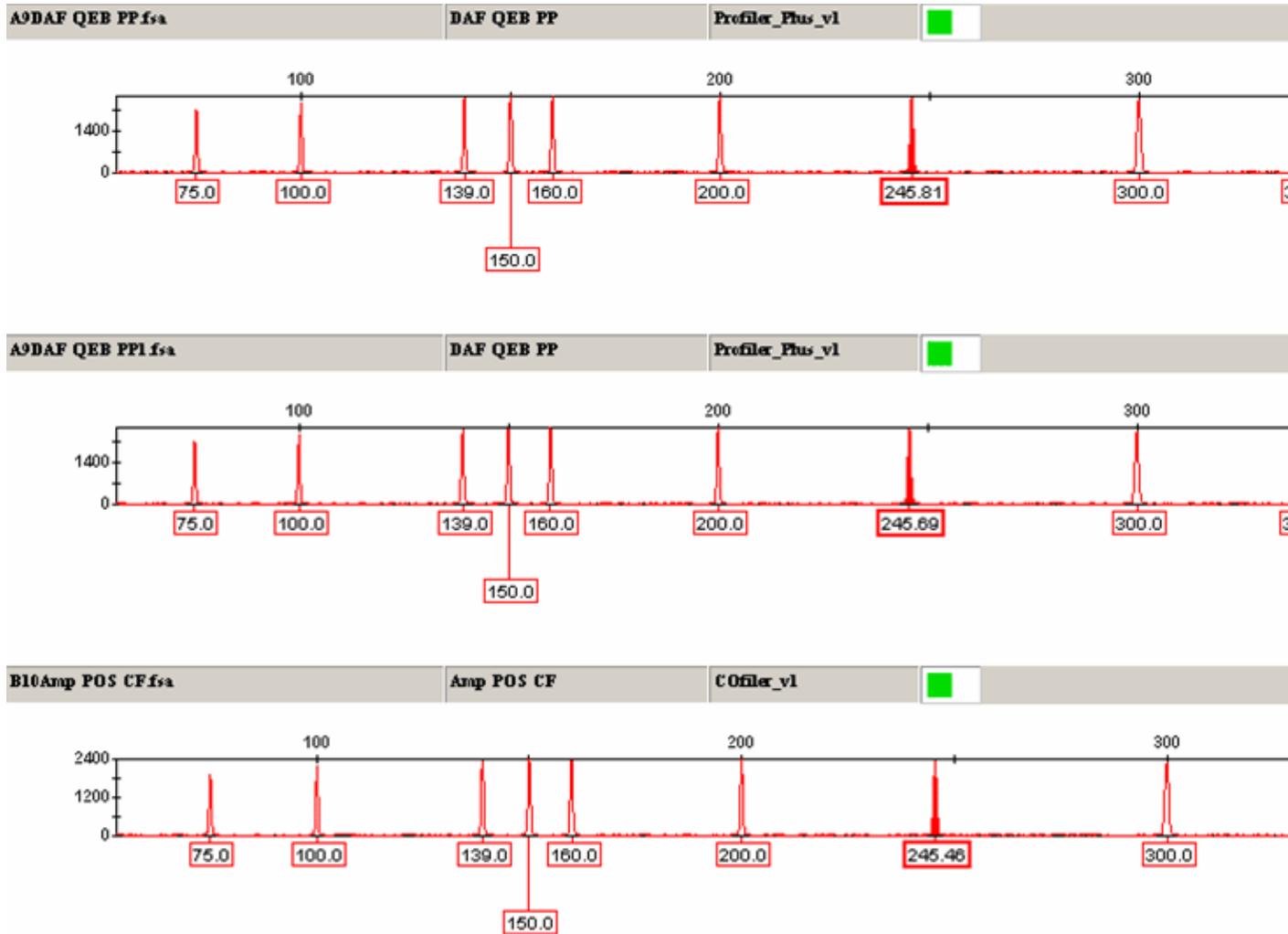


Below is an example of a bad injection, as evidenced by the imbalanced peaks on the ISS.



The ISS is particularly useful in determining the precision of a capillary electrophoresis run. For example, using Applied Biosystems software, the 250 peak is not assigned in the ISS and the size of this peak (in each sample) can be used to determine the in-run precision. Temperature fluctuations can cause the in-run precision to exceed 1 base pair and evaluation of the ISS can assist analysts in identifying this issue.

STR Data Analysis and Interpretation for Forensic Analysts

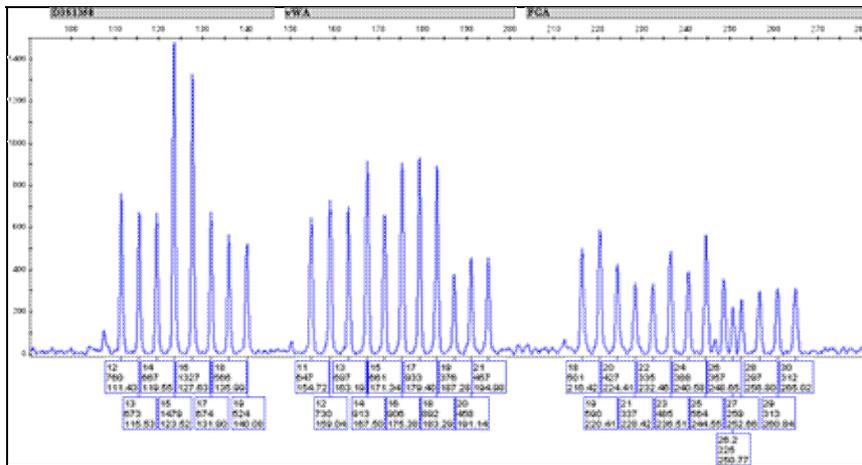


GeneScan™ 500 Size Standards are routinely used with Applied Biosystems methods. The 250 base pair peak has been shown to have abnormal migration of double strands resulting from incomplete separation under denaturing conditions. The 250 base pair peak is not defined for sizing purposes, but can be evaluated to assess in-run precision. In order to accurately size fragments, in-run precision is expected to fall into a size window of one base pair. Temperature fluctuations in the laboratory may cause precision to exceed a one base pair window.

- Allelic ladders should be assessed to ensure that all peaks have been called correctly. Ladder peaks that have not been called or have been miscalled can indicate a problem with the ladder sample, injection, and/or run.

Profiler Plus® allelic ladder (loci D3S1358, vWA, FGA)

STR Data Analysis and Interpretation for Forensic Analysts



Controls

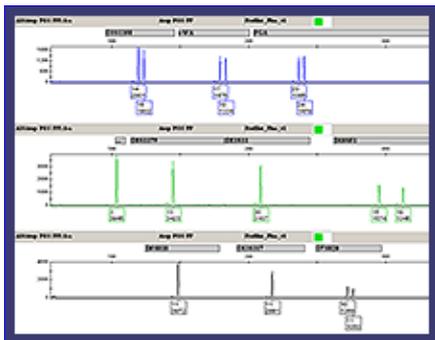
Controls are used to assess the analytical methods. In general, analyses include positive and negative amplification controls, reagent blanks, and in some instances known extraction controls and substrate controls. Laboratories must develop criteria to evaluate controls.

In addition, procedures must be in place for interpretation and documentation of control results that do not perform as expected.

Positive Control(s)

Positive controls are included in most commercial DNA analysis kits used for quantitation or STR typing. Applied Biosystems and Promega STR kits both include 9947A control DNA for use as a positive amplification control. The control is usually amplified with each batch of samples and aids in determining the overall performance of the amplification and typing procedures. 9947A is a female cell line from a 31 year old Caucasian female. The cell type is B-lymphocyte from a blood sample. This cell line is used as a positive control in forensic DNA analysis.

Positive Control 9947A Profiler Plus®



[Read more about 9974A at Coriell Institute.](#)

The positive control should show the expected alleles. If no peaks are seen, this may indicate a problem with the amplification or injection. In the event that the positive control does not yield the correct results, analysts

STR Data Analysis and Interpretation for Forensic Analysts

should troubleshoot the problem and follow the procedures established by the laboratory.

Note:

It is expected that injection issues would have been identified in the internal size standard evaluation step.
Known Extraction Control(s)

Many laboratories include known extraction controls in the analysis process. These controls aid in determining the overall performance of the entire process from extraction through typing.

The known extraction control should show the expected alleles. If no peaks are seen, this may indicate a problem with the amplification or injection. In the event that the known extraction control does not yield the correct results, analysts should troubleshoot the problem and follow the procedures established by the laboratory.

Note:

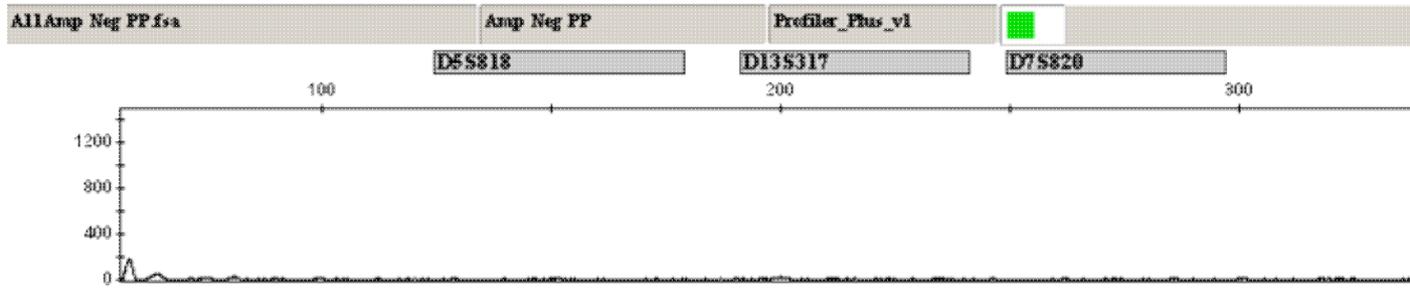
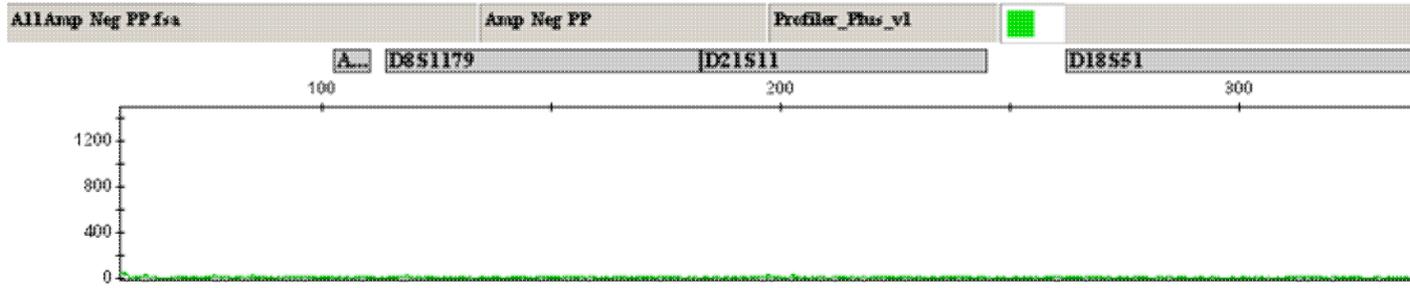
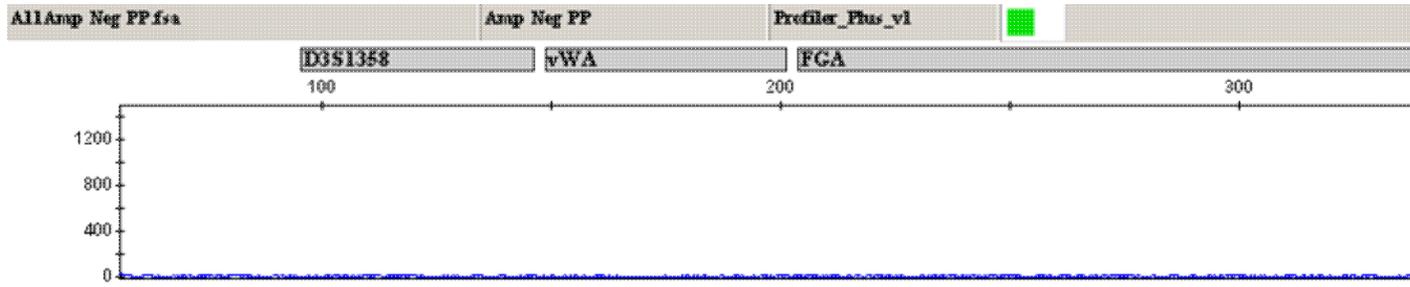
It is expected that injection issues would have been identified in the internal size standard evaluation step.
Negative Control(s) & Reagent Blanks

Reagent blanks are routinely processed with samples. Laboratories include reagent blanks in the extraction, quantitation, and amplification processes to aid in monitoring potential contamination of reagents and/or supplies. These blank samples should show no DNA pattern other than that of the internal size standard. Negative controls or reagent blanks showing unexplained extraneous DNA could indicate contamination.

Analysts should closely evaluate blank samples, including assessment of peaks under the established threshold. If there is a discernable pattern of allelic activity that cannot be attributed to a spike, pull-up, or other artifact, the analyst should troubleshoot the problem and follow the procedures established by the laboratory.

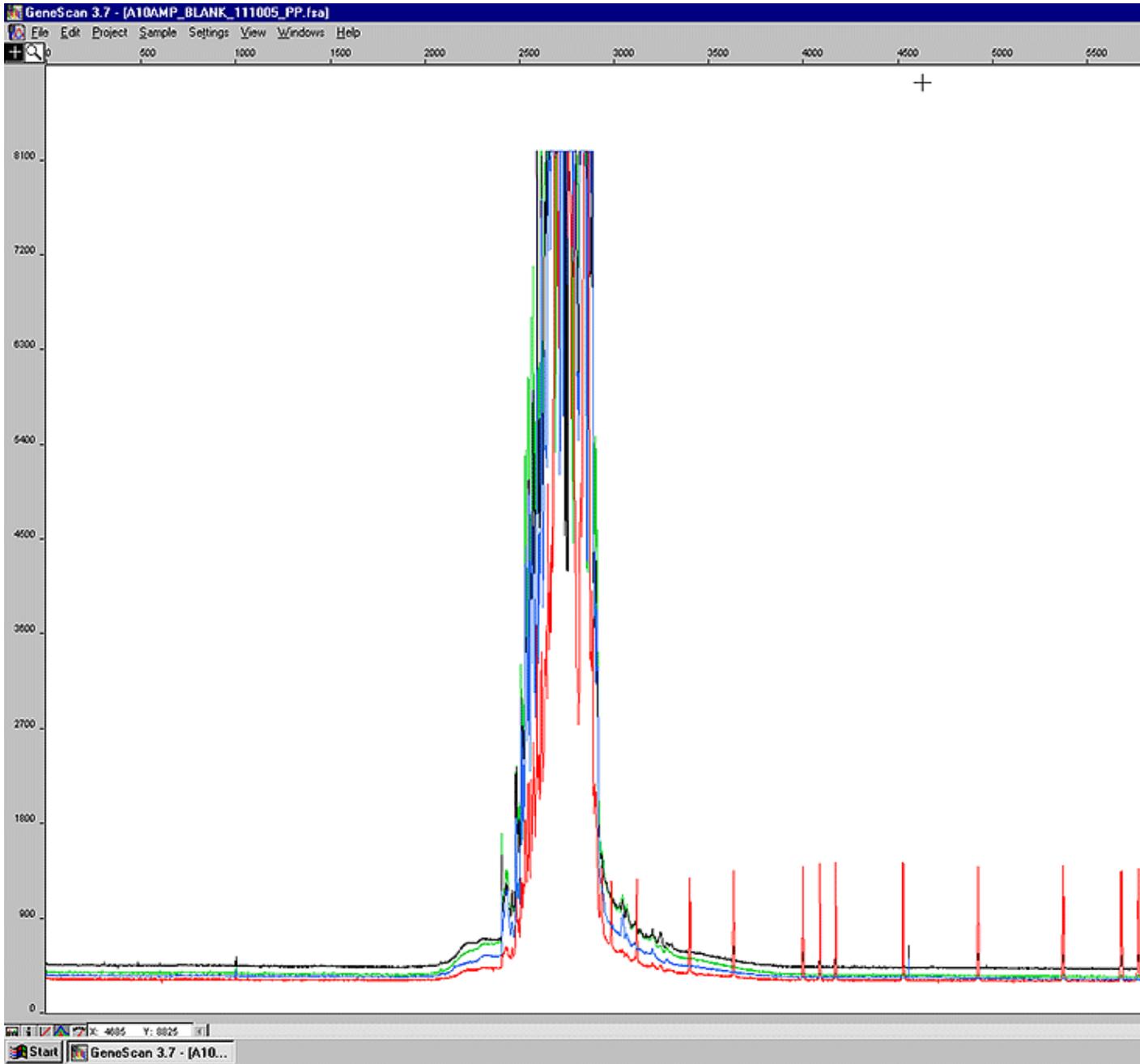
Read more about detecting and preventing contamination in the DNA Amplification PDF file.

STR Data Analysis and Interpretation for Forensic Analysts

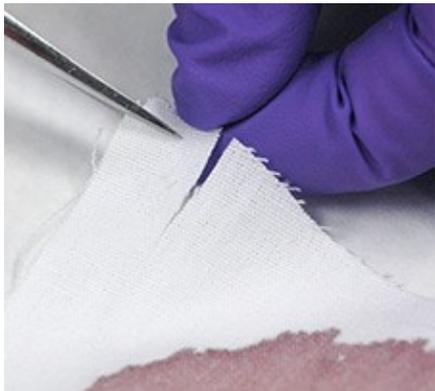


Above can be seen the blue, green, and yellow electropherograms of a negative control. The raw data, below, should also be evaluated; the primer peaks at the beginning of the run verify that an actual negative control was run, and not just an empty injection.

STR Data Analysis and Interpretation for Forensic Analysts



Substrate Controls



STR Data Analysis and Interpretation for Forensic Analysts

Some laboratories may include substrate controls in the process. These controls aid in troubleshooting inhibition arising from the substrate. In addition, substrate controls may show background DNA associated with the source of the biological stain being tested, and therefore assist in mixture interpretation. Evaluation and interpretation of data obtained from substrate controls is assessed on a case-by-case basis.

Read more about Substrates in the DNA Extraction & Quantitation PDF file.

Step 2: Extraneous Peaks

The second step is to assess each sample to determine if there are any extraneous peaks, and if they interfere with the interpretation process. Extra peaks within an allele range should be assessed following laboratory procedures.

Read more about Extraneous Peaks elsewhere in this PDF file.

Artifacts

Various artifacts can complicate and/or interfere with the interpretation process. Prior to interpreting allele designations, the analyst should evaluate each sample to determine if artifacts are present. All samples, to include controls and ladders, should be assessed for the following:

- Stutter
- 3'-A nucleotide addition
- Spurious peaks (spikes, blobs, noise)
- Pull-up

If any of these artifacts are present, the analyst should follow the procedures established by the laboratory.

Step 3: Data Evaluation

The third step in the interpretation process is to assess the data from each sample. Laboratories take both qualitative and quantitative approaches to this assessment. Laboratory procedures should be based upon validation studies and provide a scientific foundation for data interpretation and reporting.

After determining if extraneous peaks interfere with interpretation, many laboratories assess the peak height percentages and determine the contribution from each donor in mixtures.

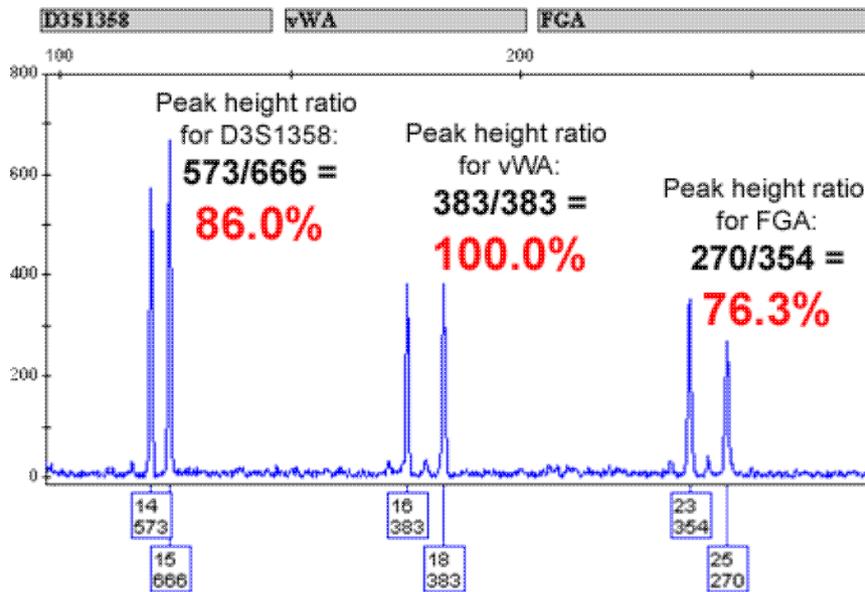
In order to use either of these estimations, the following should be established:

- The laboratory should conduct validation studies to establish a relative fluorescence unit (RFU) range within which consistent peak height percentages are obtained with the method(s) used. For example, low-level and/or degraded DNA may produce data at or around the threshold established by the laboratory. Many validation studies have demonstrated that consistent peak height percentages cannot be obtained at low RFUs.
- The laboratory should conduct validation studies to establish minimum donor contribution. For example, many laboratories have determined that, for a two person mixture, the minor donor must contribute at least 10% in order to be reliably assessed.

Peak Height Percentage

Under optimum conditions (sufficient quantity and quality of DNA), heterozygous peaks within a locus should be similar in height or intensity to each other. In general, heterozygous alleles have peak heights that are within 70% of each other, but laboratories should conduct validation studies to determine an appropriate peak height percentage for the method(s) used.

The peak height percentage for two heterozygous peaks is determined by dividing the peak height of the smaller peak by the peak height of the larger peak. The result is then expressed as a percentage, and is referred to as the peak height percentage.



Many things can affect heterozygous peak height percentages. Data from samples with low-level and/or degraded DNA may have peak height percentages much lower than 70%. In addition, the composition of the multiplex, to include the size of the multiplex, can affect the percentage. Laboratories should use their validation studies to determine an acceptable peak height percentage range for both single source and mixed DNA sample data interpretation.

Note:

The peak area can be used in place of peak height.

Single Source Profiles

Profiles developed successfully from a single contributor will display one or two alleles at each locus (excluding rare mutations). The profiles will have peak height percentages consistent with the laboratory's interpretation procedures, which were established through validation studies. All loci should be evaluated in interpreting profiles.

[View an animation about how profiles are derived from electropherograms.](#)

In general, a homozygous locus will show a single peak that is approximately twice the height of alleles seen at a heterozygous locus within the same dye color. This is due to the doubling of the signal from two alleles of the same size. If the peak height is not approximately twice that of alleles seen at a heterozygous locus, this may indicate a null allele or a primer-binding site mutation at this locus.

STR Data Analysis and Interpretation for Forensic Analysts

The peak height percentages may vary from locus to locus and should be assessed based on the laboratory's interpretation procedures. Low peak height percentages may indicate a mixed DNA sample or mutation at a specific locus.

Mixtures

The interpretation of mixtures can be a complicated process. Laboratories should conduct sufficient validation studies and provide thorough training to ensure that conclusions are supported by data.

The following may indicate a mixture:

- The presence of more than two alleles at any one locus
- A peak in a stutter position that has a higher RFU than what would be expected, based upon the established stutter percentage for that locus
- Imbalanced and/or unexpected peak height percentages

Once the determination of a mixture has been made, it may be helpful to use the case information and/or sample type to aid in the interpretation of the sources of the mixture.

For example:

- It is common to find a mixture from sexual assault evidence
- It is possible to find a mixture in a known blood sample from a deceased individual, if that person was transfused prior to death.
- It is possible to find a mixture when intimate samples are analyzed. Intimate samples are generally swabs collected from a person's body (skin, buccal, vaginal swabs)

[View an example of how mixtures might exist on a piece of evidence.](#)

Mixtures can vary in their complexity, and alternative approaches should be established for each of these categories:

- Mixtures with major and minor contributors
- Mixtures with known contributor(s)
- Mixtures with indistinguishable contributor(s)

Many laboratories quantitatively assess the peak height or area, in an effort to determine the proportion contribution of donors to a mixture.⁰² It is worth noting that mixtures can be *fully* interpreted only when all the alleles from the minor contributor are above the background noise.⁰³

Using percent contribution of donors in a mixture is premised on the following:

- The amplification parameters of the multiplex must be established. This can be accomplished through validation studies. Data from these studies can be used to determine peak area/peak height percentages of heterozygotes and stutter ranges for each locus. Establishing these parameters and understanding how a multiplex performs is crucial when using this mixture interpretation approach.
- The sample must be of suitable quality and quantity to ensure a balanced amplification. Degraded and low-level samples are prone to unbalanced amplification, which could prevent an analyst from reliably using this approach. Laboratories should conduct validation studies on low-level degraded samples, to include mixtures.
- The staff understands and recognizes how the multiplex behaves with these sample types.

Mixtures with Major/Minor Contributors

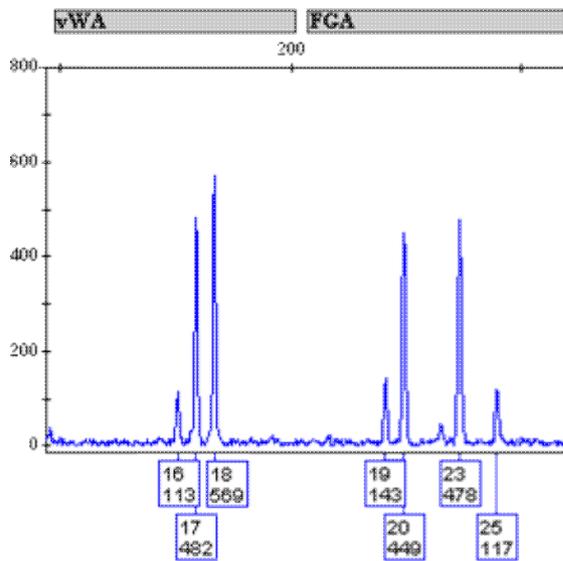
The presence of major and minor contributor(s) is distinguishable when samples display alleles that have distinct contrast in signal intensities. All loci should be used for this evaluation, and laboratories should follow established procedures that are based upon validation studies.

[View an animation about major and minor contributors to a sample.](#)

The calculations of peak height/area percentages and percent contribution of donors can be used to support the declaration of a major and/or minor contributor.

In these instances, the DNA profile of the major source may be easily determined at unambiguous loci.

In general, determining the DNA profile of the major contributor is easier than determining that of the minor source.



Above is an electropherogram demonstrating a mixture.

The set of peaks on the right (FGA locus) represent a clear major and minor contributor: there are two tall peaks and two short peaks.

The set of peaks on the left (vWA locus) represent an ambiguous minor contributor. The major contributor appears to be a 17, 18. It is unclear, however, whether the major contributor is or is not masking the minor contributor, who could be a 16, 16; a 16, 17; or a 16, 18.

Mixtures with Known Contributor(s)

It is common to obtain samples where one of the contributors (e.g., the victim) is known. In these cases, it may be possible to infer an unknown profile by subtracting the contribution of the known donor from the mixed profile.01

In general, this approach is taken on intimate samples such as vaginal swabs, fingernails, breast swabs, neck swabs, and clothing items removed from a person's body in situations where it is likely that there are no more than two contributors. It may be possible to infer unknown profiles from mixtures with more than two contributors, but that would be considerably more complex.

An analyst can use percent mixture contribution and peak height percentages to aid in inferring profiles from mixtures with known contributors.

[View an animation about how a known contributor sample can be used to infer the profile of an unknown contributor to a mixture.](#)

Mixtures with Indistinguishable Contributor(s)

Items falling into this category include non-intimate samples such as swabs from a windshield, sidewalk, wall, floor, weapon, etc. This category applies to mixtures that are not included in the two categories described above.

These mixtures are typically the most difficult category to interpret. In general, the interpretation of two-person mixtures (where each person donates half of the DNA) will not be aided by estimating percent contribution or percent peak height ratios.

Applying statistical interpretation is more simplistic with one contributor (e.g., single source samples or mixtures where a major component can be defined or inferred). Statistical evaluation of complex mixtures involves the use of the likelihood ratio or probability of inclusion/exclusion.

Read more about the likelihood ratio in the Population Genetics & Statistics PDF file.

Read more about the probability of inclusion/exclusion in the Population Genetics & Statistics PDF file.

Example of Percent Contribution

Consider the evaluation of the following:

FGA Allele	Peak Height
19*	250
20	674
23	717
25*	225

* *Visual minor alleles*

There appears to be a major and a minor contributor. Assuming only two donors, the mixture proportions are estimated by summing the peak heights of the two major or minor alleles and dividing by the peak heights of all four alleles. This is then multiplied by 100 to provide a percentage contribution.

- $250 + 225 / 250 + 674 + 717 + 225 = .25 \times 100 = 25\%$ contribution from the minor donor
- $674 + 717 / 250 + 674 + 717 + 225 = .75 \times 100 = 75\%$ contribution from the major donor

This can also be represented as a ratio by summing the peak heights of the two minor alleles and dividing them by the sum of the peak heights from the two major alleles.

- $250 + 225 / 674 + 717 = .34$ or a ratio of approximately 1 to 3

A second check of this hypothesis is evaluation of the peak height percentages. In this case the peak height percentage for genotype 1 (19, 25) is 90% and genotype 2 (20, 23) is 94%. These values are within the acceptable range of peak height percentage for balanced heterozygotes.

STR Data Analysis and Interpretation for Forensic Analysts

This evaluation can be done at any locus with no shared alleles to estimate the proportion from each contributor. Once the estimate is established, loci with two or three alleles can be evaluated.

Note:

Genotype #1	Genotype #2
A,B	C,D
A,C	B,D
A,D	B,C
C,D	A,B
B,D	A,C
B,C	A,D

In some cases major/minor contributors cannot be established, the calculations noted above can be performed for each combination to determine the percent contribution from each donor.

Two Person Mixture with Three Alleles at a Locus

Consider the following information:

D3S1358

Allele	Peak Height
14	1080
16*	320
18	690

* *Visual minor alleles*

Assuming these results came from the same case as the FGA data in the first example above, the FGA data can be used to estimate the mixture percentages. This information can aid in determining the most likely genotype(s) of the major and minor contributors for a locus with three alleles. There are numerous scenarios that may be possible. However, in many instances the analyst can easily dismiss some based on the available data.

Genotypes per Scenario

Scenario	Genotype #1(Major)	Genotype #2(Minor)
1	14,18	14,16
2	14,18	16,16
3	14,18	16,18
4	14,14	16, 18
5	16, 16	14, 18
6	18, 18	14,16
7	14, 16	14, 18
8	16, 18	14, 18
9	14,16	16,18
10	14,16	18,18
11	16,18	14,16
12	16,18	14,14

STR Data Analysis and Interpretation for Forensic Analysts

If it is assumed that the major contributor is approximately 75%, the data only support scenarios 1-3. In this example, the assignment of a genotype for the major contributor is straightforward. When alleles are shared or possibly masked, assigning a possible genotype of the minor contributor is more complex. Use of the above example to assess these three scenarios will assume that there is a 1:1 RFU ratio for heterozygous alleles.

Scenario #1

Major 14, 18; Minor 14, 16

In this scenario, 14 is a shared allele and it is assumed that the contribution is equal to that of the 16 allele, which is 320 RFUs (Genotype #2). This would leave Genotype #1 contributing 760 RFUs for the shared 14. $1080 - 320 = 760$

The estimated peak height percentage for Genotype #1 would be 91%

$690/760 = 91\%$ (note that no estimated peak height percentage is calculated for Genotype #2 since the assumption was a 1:1 contribution from each allele--320 RFU)

The percent contributions would be:

- Major = $690 + 760 / 1080 + 320 + 690 = .69 \times 100 = 69\%$
- Minor = $320 + 320 / 1080 + 320 + 690 = .31 \times 100 = 31\%$

Scenario #2

Major 14, 18; Minor 16, 16

If the minor contributor is a homozygous 16, the estimated peak height percentage for Genotype #1 would be 64%

$690/1080 = 64\%$

The percent contributions would be:

- Major = $690 + 1080 / 1080 + 320 + 690 = .85 \times 100 = 85\%$
- Minor = $320 / 1080 + 320 + 690 = .69 \times 100 = 15\%$

Scenario #3

Major 14, 18; Minor 16, 18

In this scenario, 18 is a shared allele, and it is assumed that the contribution is equal to that of the 16 allele which is 320 RFUs (Genotype #2). This would leave Genotype #1 contributing 370 RFUs to the 18.

$690 - 320 = 370$

The estimated peak height percentage for Genotype #1 would be 34%

$370/1080 = 34\%$

The percent contributions would be:

STR Data Analysis and Interpretation for Forensic Analysts

- Major = $1080 + 370 / 1080 + 320 + 690 = .69 \times 100 = 69\%$
- Minor = $320 + 320 / 1080 + 320 + 690 = .31 \times 100 = 31$

Based on these estimations, Scenario #1 is the best fit, but Scenario #2 should be considered. Scenario #3 has an estimated peak height percentage for Genotype #1 of 34%, making this combination unlikely.

Two Person Mixture with Two Alleles at a Locus

Consider the following information:

D5S818

Allele	Peak Height
9	794
11	1201

The same approach can be used to estimate the mixture percentages and interpret data in a two person mixture case with two alleles at a locus. This information can aid in determining the most likely genotypes of the major and minor contributors.

The following scenarios are possible:

Possible Scenarios

Scenario	Genotype #1 (Major)	Genotype #2 (Minor)
1	9,9	9,11
2	9,11	9,11
3	9,9	11,11
4	9,11	11,11
5	9,11	9,9
6	11,11	9,9
7	11,11	9,11

- **Scenario #1**

This scenario is unlikely because a shared 9 allele would leave a very low percent contribution for Genotype #1.

- **Scenario #2**

If it is assumed that the minor donor contributed 25% and the major donor contributed 75% (based on FGA results).

The estimated peak height percentage would be:

- ◆ Genotype #1 would be $198/300 \times 100 = 66\%$
- ◆ Genotype #2 would be $596/901 \times 100 = 66\%$

- **Scenario #3**

The percent contribution would be:

- ◆ Genotype #1 $(794 / 794 + 1201) \times 100 = 40\%$
- ◆ Genotype #2 $(1201 / 794 + 1201) \times 100 = 60\%$

- **Scenario #4**

In this scenario, 11 is a shared allele, and it is assumed that the contribution is equal to that of the 9

STR Data Analysis and Interpretation for Forensic Analysts

allele which is 794 RFUs (Genotype #1). This would leave Genotype #2 contributing 407 RFUs.

The percent contribution would be:

$$\blacklozenge \text{Genotype \#1} = (794 + 794 / 794 + 1201) \times 100 = 80\%$$

$$\blacklozenge \text{Genotype \#2} = (407 / 794 + 1201) \times 100 = 20\%$$

- **Scenario #5**

This scenario unlikely since a shared 9 allele would leave a very low percent contribution for Genotype #2.

- **Scenario #6**

See scenario #3, as this is the reverse.

The percent contribution would be:

$$\blacklozenge \text{Genotype \#1} (1201 / 794 + 1201) \times 100 = 60\%$$

$$\blacklozenge \text{Genotype \#2} (794 / 794 + 1201) \times 100 = 40\%$$

- **Scenario #7**

See scenario #4, as this is the reverse

In this scenario, 11 is a shared allele and it is assumed that the contribution is equal to that of the 9 allele which is 794 RFUs (Genotype #2). This would leave Genotype #1 contributing 407 RFUs.

The percent contribution would be:

$$\blacklozenge \text{Genotype \#1} = (407 / 794 + 1201) \times 100 = 20\%$$

$$\blacklozenge \text{Genotype \#2} = (794 + 794 / 794 + 1201) \times 100 = 80\%$$

Based on these seven scenarios:

- Scenarios 2, 4, and 7 are possible
- Scenarios 3 and 6 are less likely
- Scenarios 1 and 5 are not likely

Reporting Guidelines

A report contains the conclusions that an analyst has made based on the scientific data and established interpretation procedures.

An inclusion or match is reported when all the loci from which a result is obtained match between a questioned sample and a known sample. With the exception of paternity testing, if the result for even one locus does not match and is discordant, an exclusion is declared. A minimum of a two-locus exclusion is needed in order to declare an exclusion for paternity.

Read more about the rules for parentage and relatedness in the Population Genetics & Statistics PDF file.

[View an explanation of a match.](#)

Exclusion/Non-match

An exclusion or non-match is reported when loci from which results are obtained are discordant between a questioned sample and a known.

[View an explanation of a non-match.](#)

Inconclusive/ Uninterpretable

There are occasions when an analyst cannot interpret data or the data is inconclusive. Some mixtures may be too complex to render conclusions and will be reported as uninterpretable. Low-level and/or degraded DNA samples may result in inconclusive results at some or all loci.

[View an explanation of an inconclusive electropherogram.](#)

No Results

No results are reported when there is no discernable allelic activity observed at a locus.

[View an explanation of an electropherogram that shows no results.](#)

Read more about Reporting in the Communicating Results PDF file.

Author: Steve O'Dell

Steve O'Dell, MSFS, CCSA, is the DNA Technical Manager and Crime Scene Response Section Supervisor for the Phoenix Police Department.

Author: Russell Vossbrink

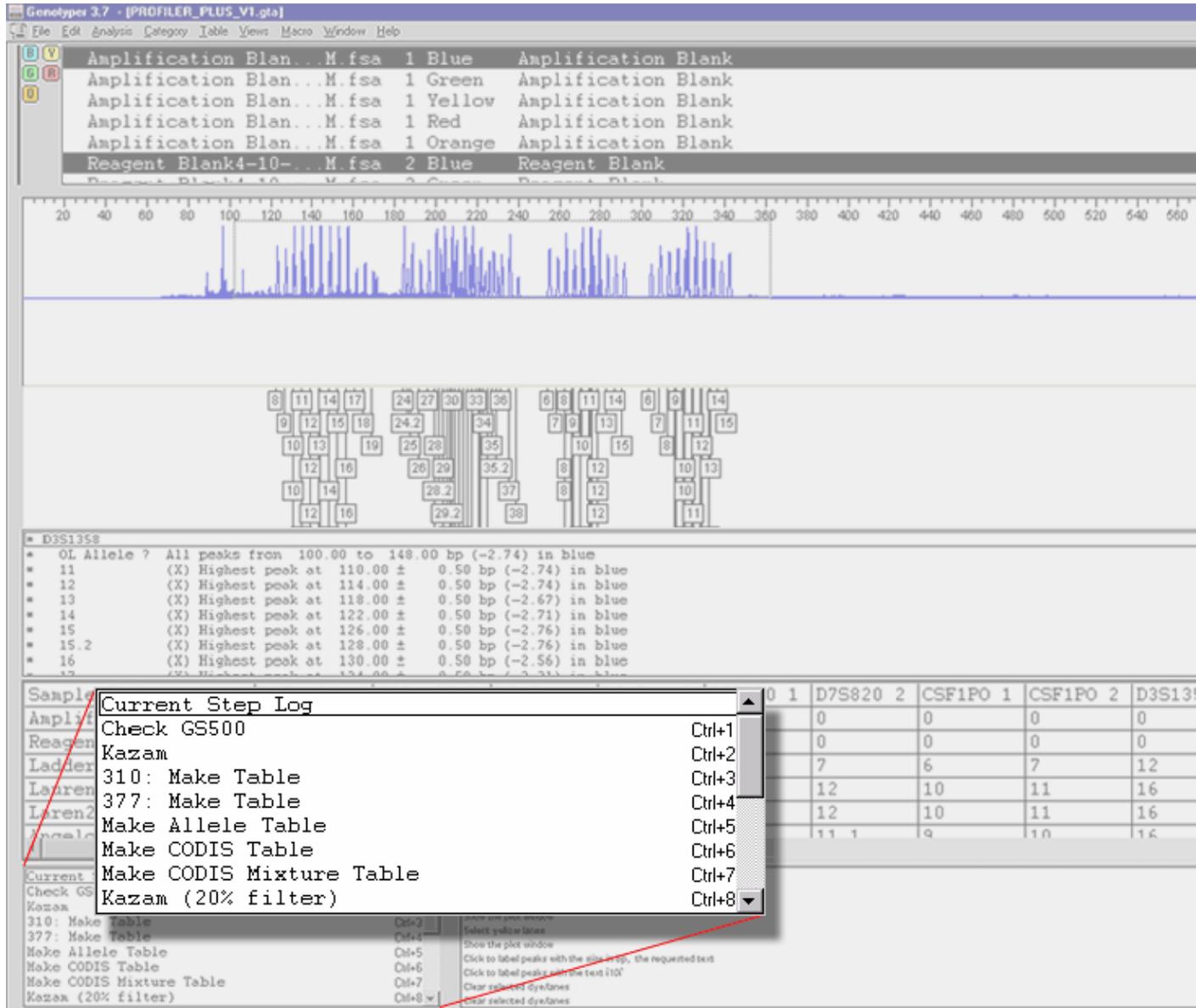
Russell Vossbrink is the technical supervisor of the Regional FBI mtDNA lab at the Arizona Department of Public Safety. Russell is a member of the American Academy of Forensic Sciences (AAFS) and he has been a regular member and invited guest of the Scientific Working Group for DNA Analysis Methods (SWGDM). Russell has presented at Promega, Combined DNA Index System (CODIS), AAFS, California Association of Criminalists (CAC), Southwestern Association of Forensic Scientists (SWAFS), California Association of Crime Laboratory Directors (CACLD) and the Bode Technology Workshop.

Author: Debbie Figarelli

Debbie Figarelli serves as DNA Technical Leader at the National Forensic Science Technology Center. Debbie assists with the development of DNA training programs and participates in compliance audits of DNA laboratories.

Below is a screen shot of Genotypere^{1/2}, focusing on the macro window. Click on the macro labeled "Kazam" to view the steps of the macro as they appear in text format.

STR Data Analysis and Interpretation for Forensic Analysts



Macro: Kazam

Find ÒladderÓ in sample info, in blue dye/lanes

Run macro ÒCalculate D3S1358 offsetsÓ

Run macro ÒCalculate vWA OffsetsÓ

Run macro ÒCalculate FGA OffsetsÓ

Find ÒladderÓ in sample info, in green dye/lanes

Run macro ÒCalculate AMEL OffsetsÓ

Run macro ÒCalculate D8S1179 offsetsÓ

STR Data Analysis and Interpretation for Forensic Analysts

Run macro "Calculate D21S11 Offsets"

Run macro "Calculate D18S51 offsets"

Find "ladder" in sample info, in yellow dye/lanes

Run macro "Calculate D5S818 Offsets"

Run macro "Calculate D13S317 Offsets"

Run macro "Calculate D7S820 Offsets"

Clear Labels

Select all categories

Unmark selected categories

Select category: D3S1358

Mark selected categories

Select blue lanes

Label category peaks with the category's name

Following Step is Sample Allele Stutter Filter

Remove labels from peaks followed by a 80% higher, labeled peak within 3.25 to 4.75 bp

Find all "ladder" in sample info, in blue dye/lanes

Label category peaks with the category's name

Remove labels from peaks whose height is less than 20% of the highest peak in a category's range

Select all categories

Unmark selected categories

Select category: vWA

Mark selected categories

Select blue lanes

Label category peaks with the category's name

Following Step is Sample Allele Stutter Filter

Remove labels from peaks followed by a 80% higher, labeled peak within 3.25 to 4.75 bp

STR Data Analysis and Interpretation for Forensic Analysts

Find all "ladder" in sample info, in blue dye/lanes

Label category peaks with the category's name

Remove labels from peaks whose height is less than 25% of the highest peak in a category's range

Select all categories

Unmark selected categories

Select category: FGA

Mark selected categories

Select blue lanes

Label category peaks with the category's name

Following Step is Sample Allele Stutter Filter

Remove labels from peaks followed by a 80% higher, labeled peak within 3.25 to 4.75 bp

Find all "ladder" in sample info, in blue dye/lanes

Label category peaks with the category's name

Remove labels from peaks whose height is less than 30% of the highest peak in a category's range

Select all categories

Unmark selected categories

Select category: AMEL

Mark selected categories

Select green lanes

Label category peaks with the category's name

Following Step is Sample Allele Stutter Filter

Remove labels from peaks whose height is less than 3% of the highest peak in a category's range

Find all "ladder" in sample info, in green dye/lanes

Label category peaks with the category's name

Remove labels from peaks whose height is less than 40% of the highest peak in a category's range

Select all categories

STR Data Analysis and Interpretation for Forensic Analysts

Unmark selected categories

Select category: D18S51

Mark selected categories

Select green lanes

Label category peaks with the category's name

Following Step is Sample Allele Stutter Filter

Remove labels from peaks followed by a 525% higher, labeled peak within 3.25 to 4.75 bp

Find all "ladder" in sample info, in green dye/lanes

Label category peaks with the category's name

Remove labels from peaks whose height is less than 10% of the highest peak in a category's range; then remove labels from peaks followed by a 222% higher, labeled peak within 0.00 to 5.00 bp

Select all categories

Unmark selected categories

Select category: D21S11

Mark selected categories

Select green lanes

Label category peaks with the category's name

Following Step is Sample Allele Stutter Filter

Remove labels from peaks followed by a 669% higher, labeled peak within 3.25 to 4.75 bp

Find all "ladder" in sample info, in green dye/lanes

Label category peaks with the category's name

Remove labels from peaks whose height is less than 20% of the highest peak in a category's range

Select all categories

Unmark selected categories

Select category: D8S1179

Mark selected categories

STR Data Analysis and Interpretation for Forensic Analysts

Select green lanes

Label category peaks with the category's name

Following Step is Sample Allele Stutter Filter

Remove labels from peaks followed by a 733% higher, labeled peak within 3.25 to 4.75 bp

Find all "ladder" in sample info, in green dye/lanes

Label category peaks with the category's name

Remove labels from peaks whose height is less than 30% of the highest peak in a category's range

Select all categories

Unmark selected categories

Select category: D5S818

Mark selected categories

Select yellow lanes

Label category peaks with the category's name

Following Step is Sample Allele Stutter Filter

Remove labels from peaks followed by a 900% higher, labeled peak within 3.25 to 4.75 bp

Find all "ladder" in sample info, in yellow dye/lanes

Label category peaks with the category's name

Remove labels from peaks whose height is less than 25% of the highest peak in a category's range

Select all categories

Unmark selected categories

Select category: D13S317

Mark selected categories

Select yellow lanes

Label category peaks with the category's name

Following Step is Sample Allele Stutter Filter

Remove labels from peaks followed by a 900% higher, labeled peak within 3.25 to 4.75 bp

STR Data Analysis and Interpretation for Forensic Analysts

Find all ladders in sample info, in yellow dye/lanes

Label category peaks with the category's name

Remove labels from peaks whose height is less than 25% of the highest peak in a category's range

Select all categories

Unmark selected categories

Select category: D7S820

Mark selected categories

Select yellow lanes

Label category peaks with the category's name

Following Step is Sample Allele Stutter Filter

Remove labels from peaks followed by a 101% higher, labeled peak within 3.25 to 4.75 bp

Find all ladders in sample info, in yellow dye/lanes

Label category peaks with the category's name

Remove labels from peaks whose height is less than 25% of the highest peak in a category's range

Select all categories

Unmark selected categories

Select categories: D3S1358, vWA, FGA, AMEL, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820

Mark selected categories

Sort categories by: (1) ascending dye color, (2) ascending size or scan

Sort dye/lanes by: (1) ascending lane number, (2) ascending dye color

Select blue lanes

And select green lanes

And select yellow lanes

Unmark selected dye/lanes

Find ladders in sample info, in blue dye/lanes

Mark selected dye/lanes

STR Data Analysis and Interpretation for Forensic Analysts

Find ÒladderÓ in sample info, in green dye/lanes

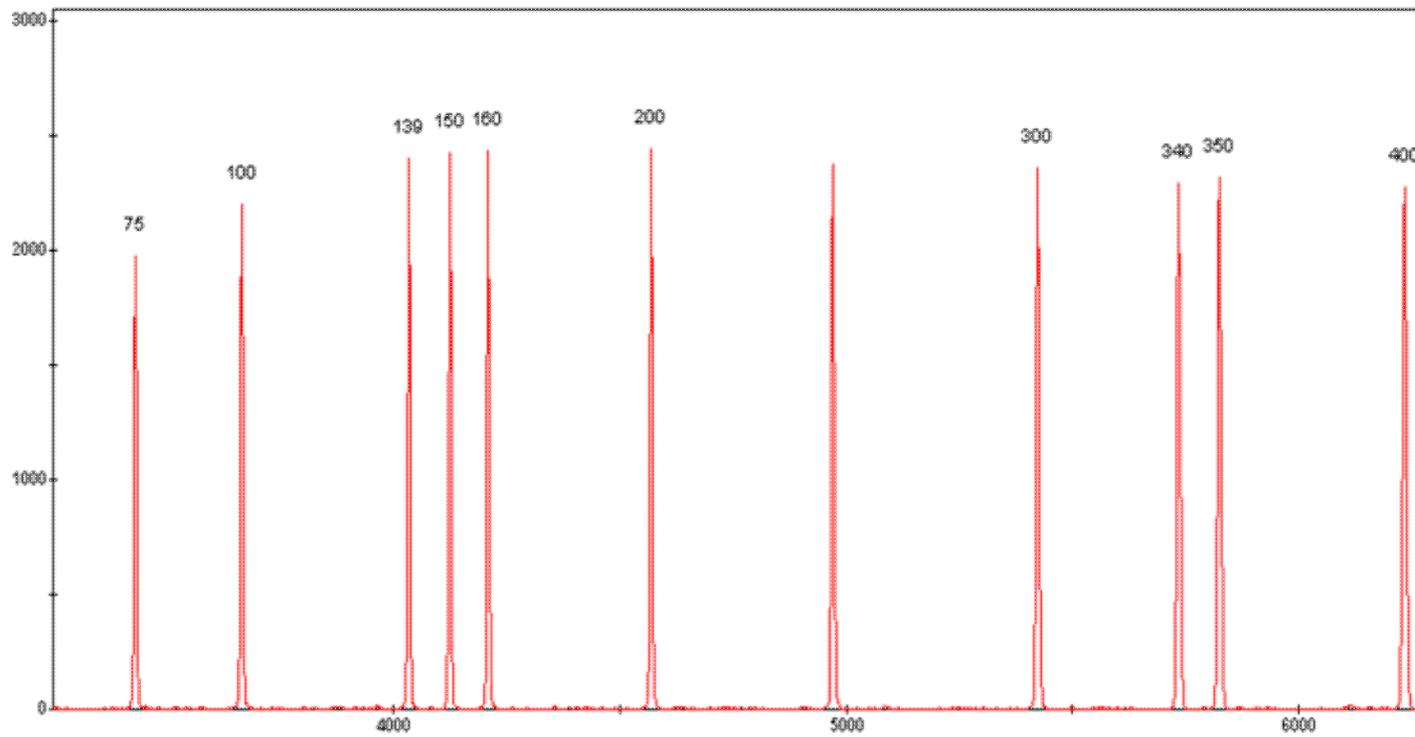
Mark selected dye/lanes

Find ÒladderÓ in sample info, in yellow dye/lanes

Mark selected dye/lanes

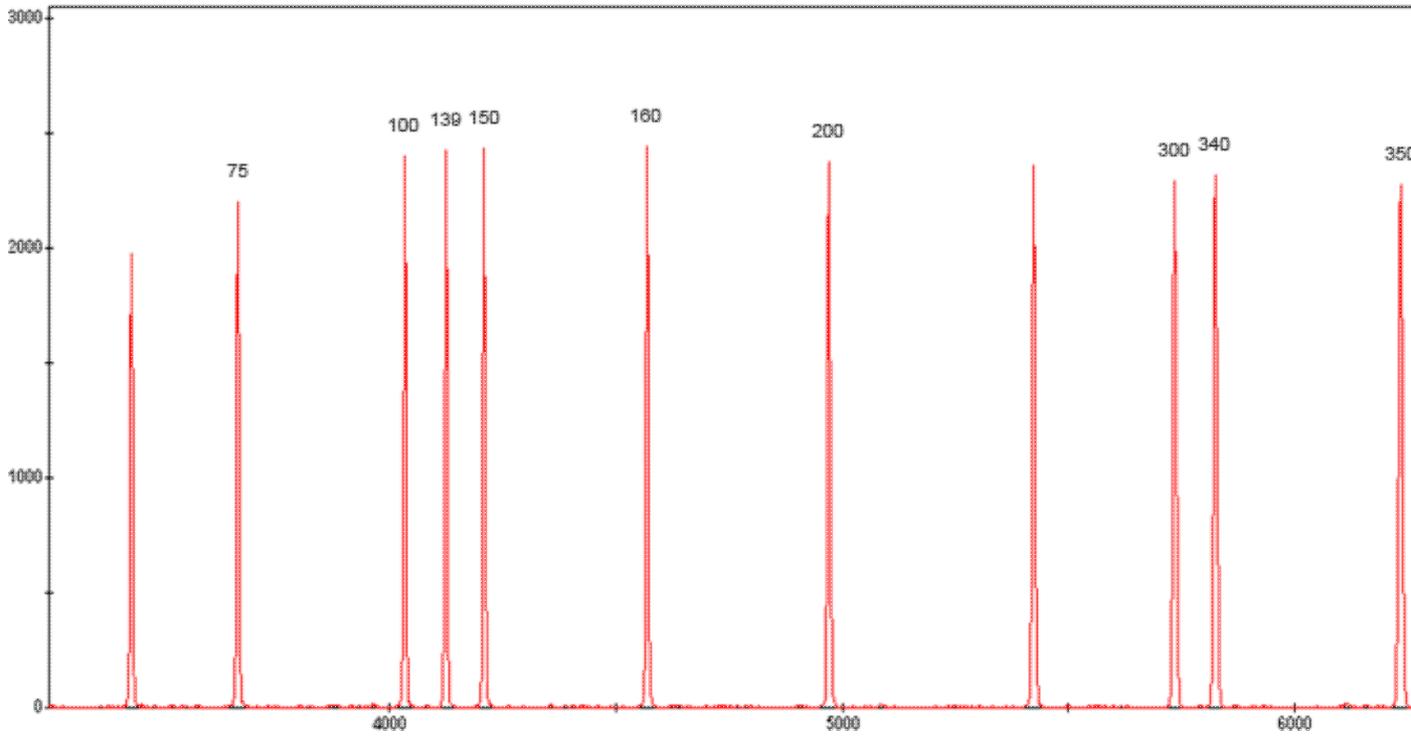
Select blue lanes

Show the plot window

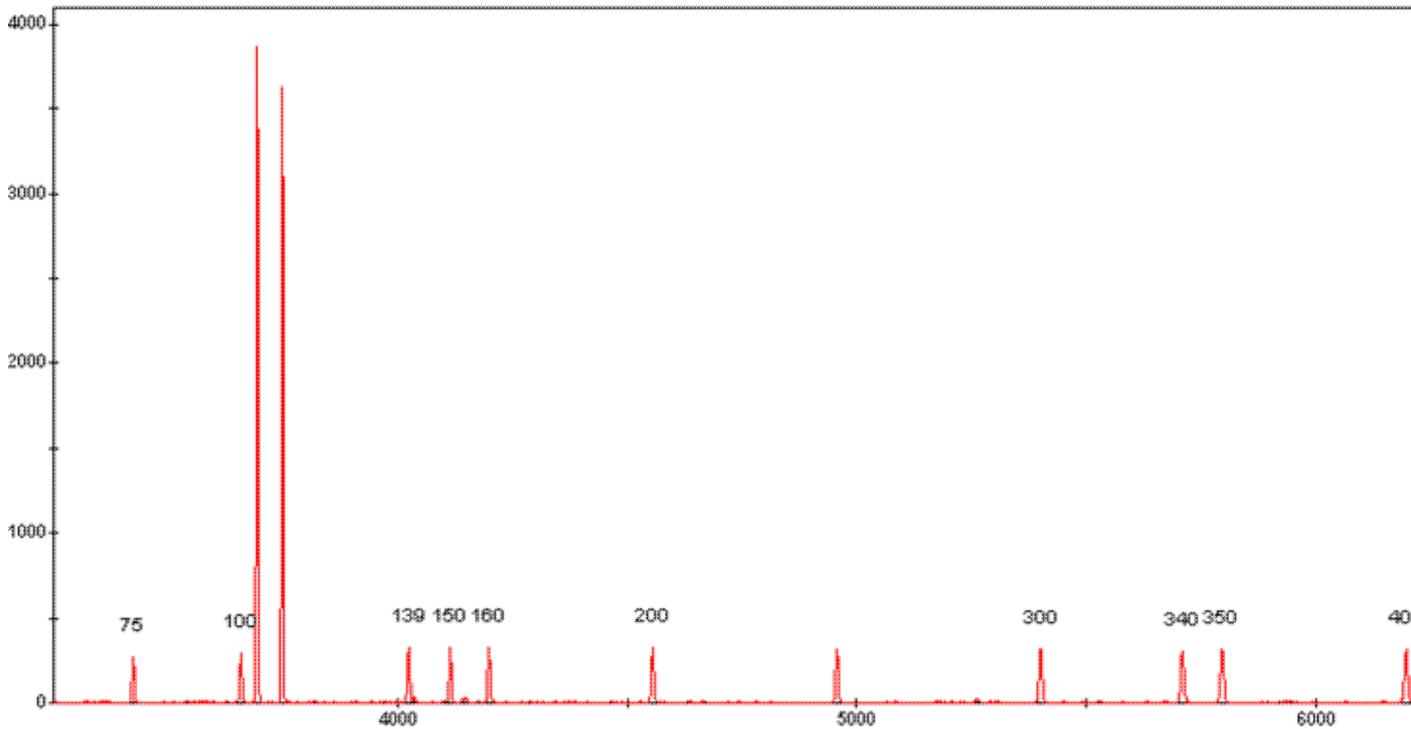


Above can be seen an example of a good internal size standard. Compare this to a miscalled ISS, below. Note that the first peak has not been called, and normal peak calls have begun with the second peak from the left.

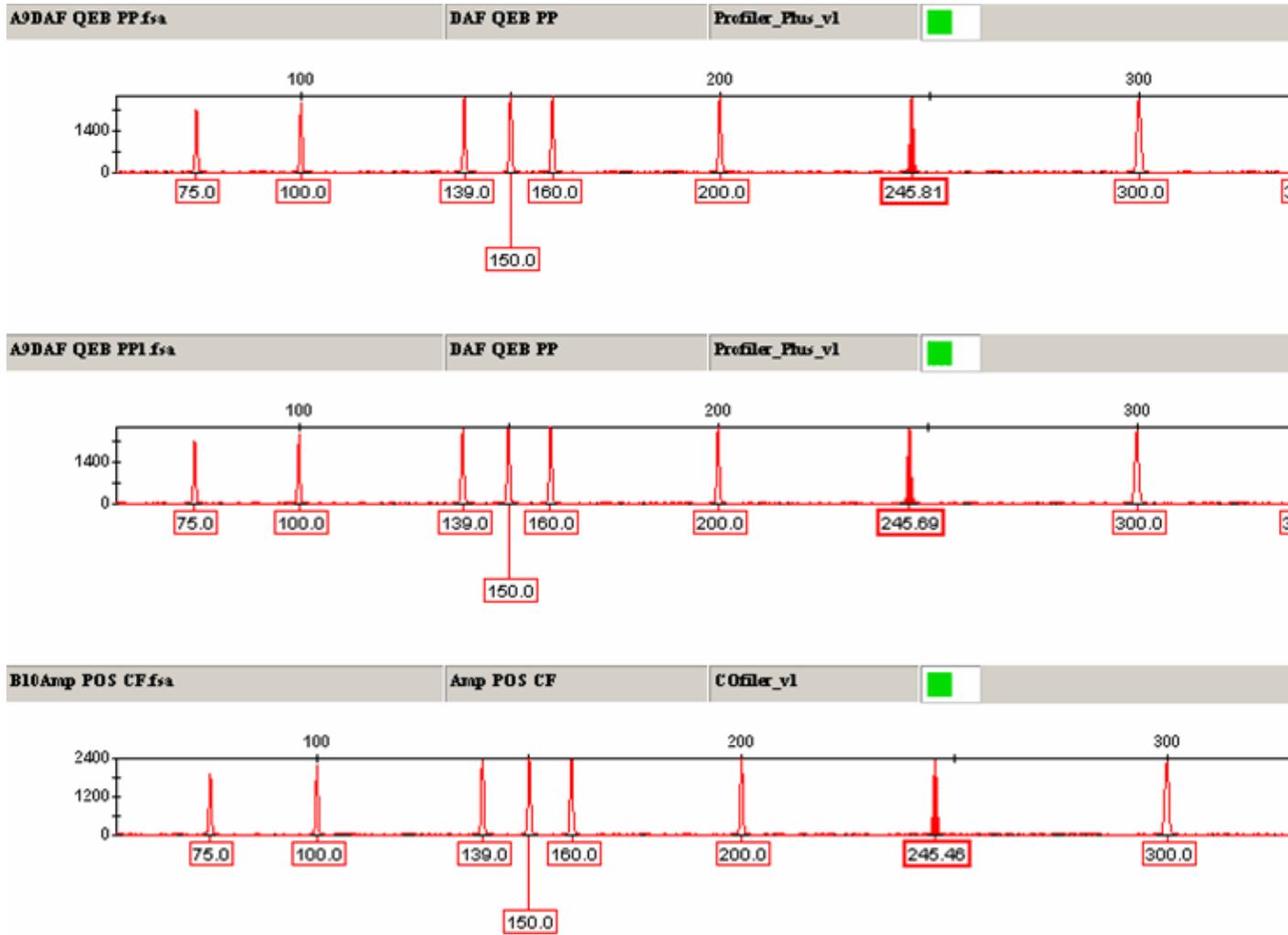
STR Data Analysis and Interpretation for Forensic Analysts



Below is an example of a bad injection, as evidenced by the imbalanced peaks on the ISS.

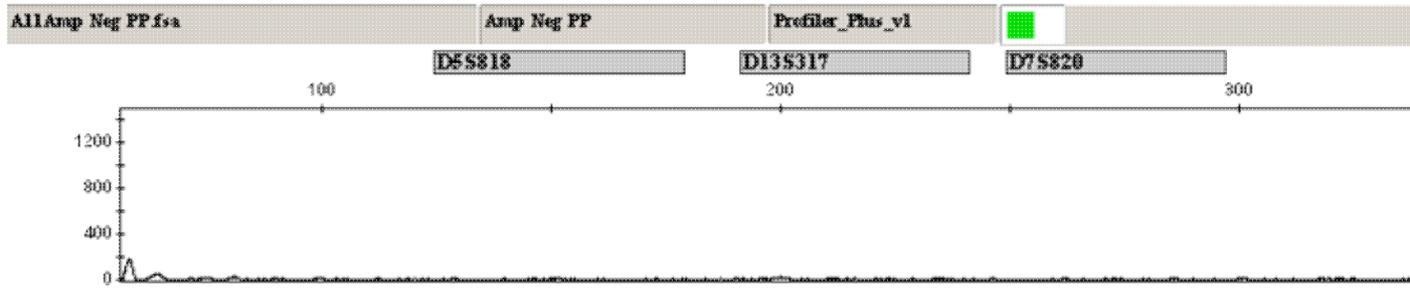
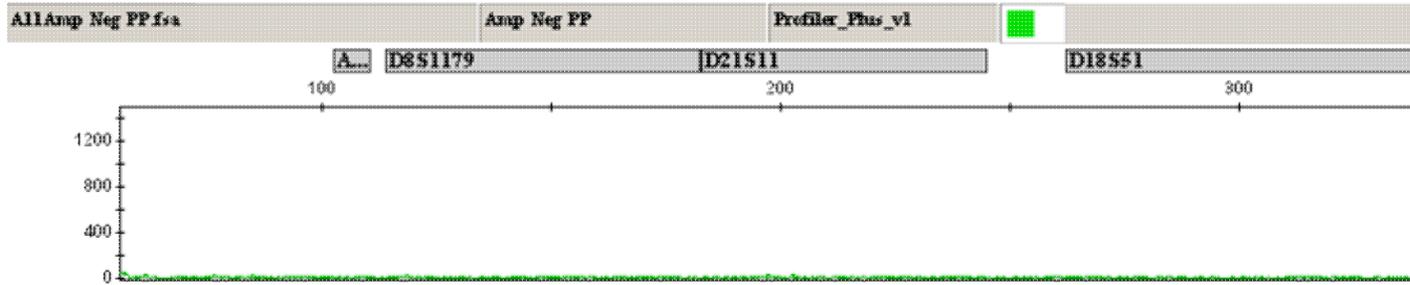
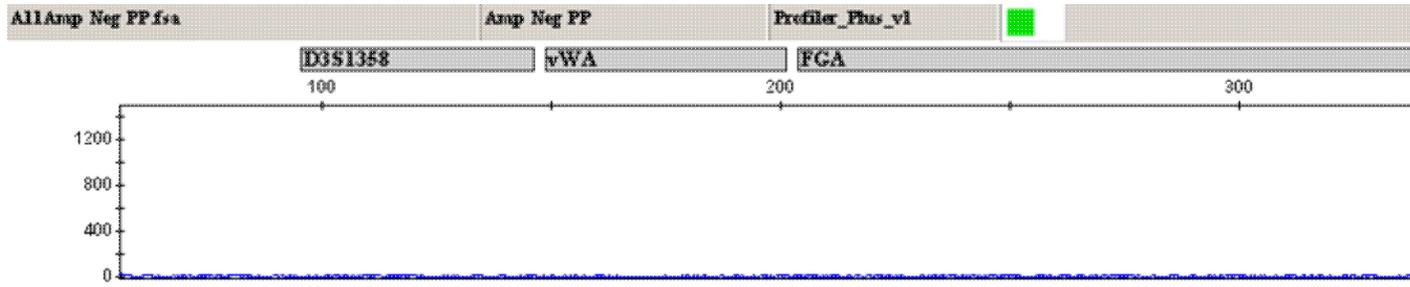


STR Data Analysis and Interpretation for Forensic Analysts



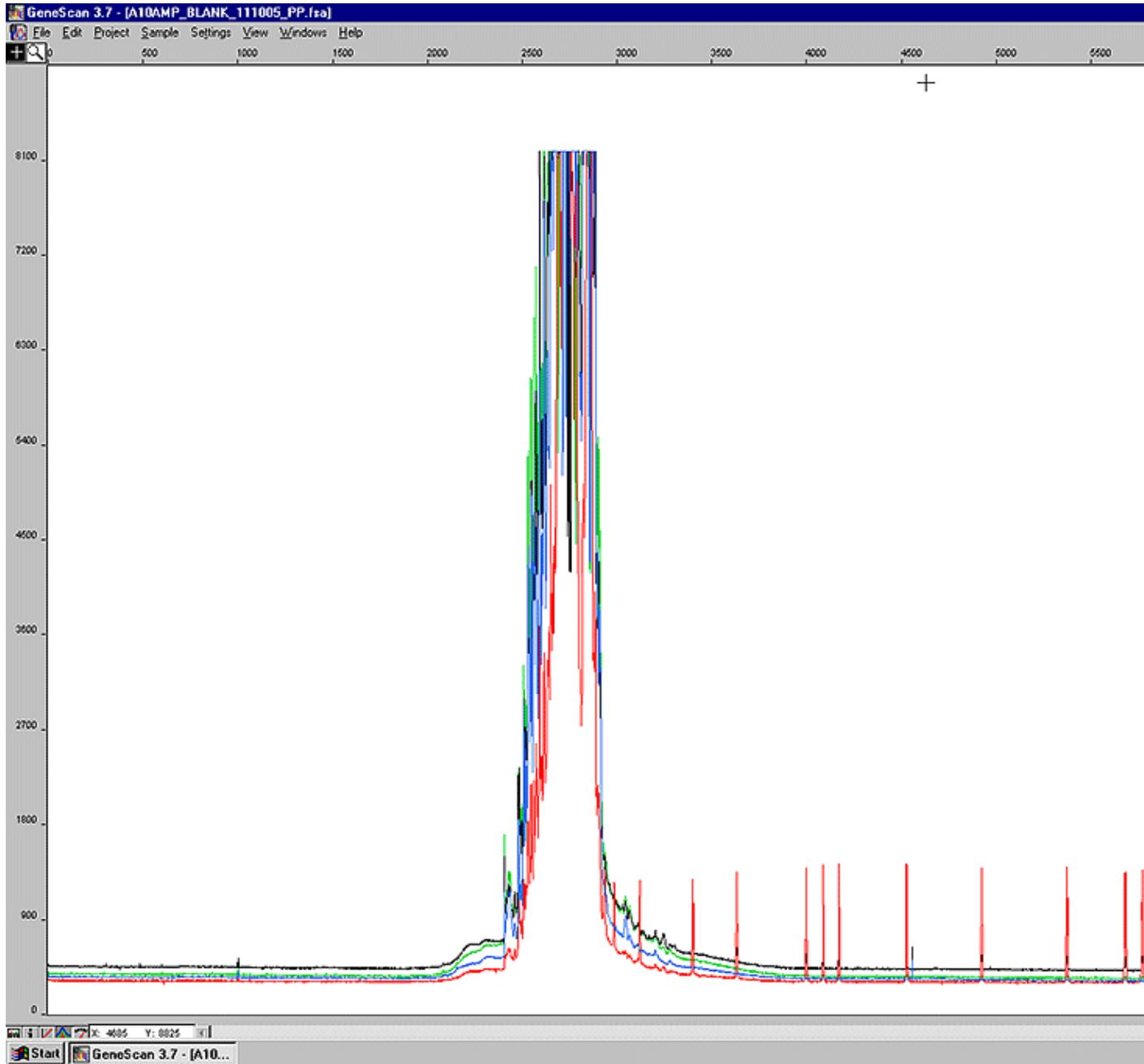
GeneScan™ 500 Size Standards are routinely used with Applied Biosystems methods. The 250 base pair peak has been shown to have abnormal migration of double strands resulting from incomplete separation under denaturing conditions. The 250 base pair peak is not defined for sizing purposes, but can be evaluated to assess in-run precision. In order to accurately size fragments, in-run precision is expected to fall into a size window of one base pair. Temperature fluctuations in the laboratory may cause precision to exceed a one base pair window.

STR Data Analysis and Interpretation for Forensic Analysts



Above can be seen the blue, green, and yellow electropherograms of a negative control. The raw data, below, should also be evaluated; the primer peaks at the beginning of the run verify that an actual negative control was run, and not just an empty injection.

STR Data Analysis and Interpretation for Forensic Analysts



Works Cited & Online Links



1. Butler, John M. 2005. *Forensic DNA typing: Biology, technology, and genetics of STR markers*. 2nd ed. Burlington, MA: Elsevier Academic Press.

STR Data Analysis and Interpretation for Forensic Analysts

2. Applied Biosystems. 2003. GeneMapper™ ID Software Version 3.1: Human Identification Analysis: User Guide, no. 4338775C, rev. C.
3. Hartzell, B., K. Graham, and B. McCord. 2003. Response of short tandem repeat systems to temperature and sizing methods. *Forensic Sci Int*, 133 (3):228–34.
4. Klein, S. B., J. M. Wallin, and M. R. Buoncristiani. 2003. Addressing ambient temperature variation effects on sizing precision of AmpFISTR® Profiler Plus™ alleles detected on the ABI Prism® 310 Genetic Analyzer. *Forensic Science Communications* 5 (1). <http://www.fbi.gov/hq/lab/fsc/backissu/jan2003/klein.htm>.
5. Applied Biosystems. 200_. ABI PRISM® GeneScan® Analysis Software for the Windows NT® Operating System: User Bulletin.
6. Applied Biosystems. 2004. GeneMapper® ID Software: Installation Procedures and New Features for GeneMapper® ID Software v3.2: User Bulletin, part 4352543, rev. A.
7. Applied Biosystems. 2004. GeneMapper® ID Software Versions 3.1 and 3.2: Human Identification Analysis: Tutorial, part 4357520, rev. A.

Online Links

- [GMID Software Tutorial](http://www.appliedbiosystems.com)
<http://www.appliedbiosystems.com>

Works Cited & Online Links



1. Applied Biosystems. 1988. AmpF STR® Profiler Plus™ PCR Amplification Kit user's manual.
2. Walsh, P. S., N. J. Fildes, and R. Reynolds. 1996. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Res* 24 (14): 2807–12.
3. Clark, J. M. 1988. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res* 16 (20): 9677–86.
4. Butler, John M., and Dennis J. Reeder. 2006. NIST standard reference database SRD 130: Variant allele reports. http://www.cstl.nist.gov/div831/strbase/var_tab.htm (accessed August 30, 2006).
5. Butler, John M., and Dennis J. Reeder. 2006. NIST standard reference database SRD 130: Apparent mutations observed at STR loci in the course of paternity testing. <http://www.cstl.nist.gov/div831/strbase/mutation.htm> (accessed August 30, 2006).
6. Chakraborty, R., and D. N. Stivers. 1996. Paternity exclusion by DNA markers: Effects of paternal mutations. *J Forensic Sci* 41 (4): 671–77.
7. Budowle, B. 2000. STR allele concordance between different primer sets: A brief summary. *Profiles in DNA*, Promega publication 3 (3): 10–11. http://www.promega.com/profiles/303/ProfilesinDNA_303_10.pdf.
8. DNA Advisory Board. 1998. *Quality assurance standards for forensic DNA testing laboratories*. *Forensic Science Communications* 2 (3). <http://www.fbi.gov/hq/lab/codis/forensic.htm>.
9. National Research Council, Committee on DNA Technology in Forensic Science. 1996. *The evaluation of forensic DNA evidence: An update*. Washington, D.C.: National Academy Press.

STR Data Analysis and Interpretation for Forensic Analysts

Online Links

- National Institute of Standards and Technology (NIST)
<http://www.cstl.nist.gov/div831/strbase/>

Works Cited & Online Links



1. Scientific Working Group on DNA Analysis Methods (SWGDM). 2000. Short tandem repeat (STR) interpretation guidelines. *Forensic Science Communications* 2 (3). <http://www.fbi.gov/hq/lab/fsc/backissu/july2000/strig.htm>.
2. Gill, P. 2002. Role of short tandem repeat DNA in forensic casework in the UK: Past, present, and future perspectives. *Biotechniques* 32 (2): 366–8, 370, 372 passim.
3. Buckleton, John, Christopher M. Triggs, and Simon J. Walsh, eds. 2005. *Forensic DNA evidence interpretation*. Boca Raton, FL: CRC Press.

Online Links

- 9974A
http://ccr.coriell.org/nigms/nigms_cgi/display.cgi?GM09947