



Improved Analysis of

DNA Short Tandem Repeats

With Time-of-Flight
Mass Spectrometry

U.S. Department of Justice
Office of Justice Programs
810 Seventh Street N.W.
Washington, DC 20531

John Ashcroft
Attorney General

Office of Justice Programs
World Wide Web Site
<http://www.ojp.usdoj.gov>

National Institute of Justice
World Wide Web Site
<http://www.ojp.usdoj.gov/nij>

IMPROVED ANALYSIS OF DNA SHORT TANDEM REPEATS WITH TIME-OF-FLIGHT MASS SPECTROMETRY

John M. Butler and Christopher H. Becker

Science and Technology Research Report

October 2001
NCJ 188292



Sarah V. Hart
Director, National Institute of Justice

Lois Tully
Project Monitor

John M. Butler, Ph.D., is currently a research chemist at the National Institute of Standards and Technology and principle investigator on an NIJ-funded project to further develop multiplex PCR and time-of-flight mass spectrometry for future forensic DNA typing assays. He was the first to demonstrate that short tandem repeat typing could be performed with capillary electrophoresis.

Christopher H. Becker, Ph.D., is currently senior director of proteomics technology at Thermo Finnigan in San Jose, California. During the span of this project, he was president and chief operations officer of GeneTrace Systems, Inc.

This project was supported under grant number 97-LB-VX-0003 from the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice.

This document is not intended to create, does not create, and may not be relied upon to create any rights, substantive or procedural, enforceable at law by any party in any matter, civil or criminal.

For further information, contact John M. Butler, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899; phone 301-975-4049; e-mail john.butler@nist.gov.

The National Institute of Justice is a component of the Office of Justice Programs, which also includes the Bureau of Justice Assistance, the Bureau of Justice Statistics, the Office of Juvenile Justice and Delinquency Prevention, and the Office for Victims of Crime.

ACKNOWLEDGMENTS

The project described in this report could not have happened without the hard work and support of a number of people at GeneTrace Systems, Inc. First and foremost, Jia Li did some of the early primer design and STR work to demonstrate that STRs could be effectively analyzed by mass spectrometry. Jia taught us a lot about PCR and was always encouraging of our work. Likewise, Tom Shaler was important in the early phases of this research with his expert advice in mass spectrometry and data processing. The first GeneTrace STR mass spectra were carefully collected by Tom, and thus he and Jia deserve credit for helping obtain the funding for this study. Dan Pollart synthesized numerous cleavable primers for this project, especially in the first year of our work. David Joo and Wendy Lam also prepared PCR and SNP primers for the later part of this work. A number of people assisted in robotic sample preparation and sample cleanup, including Mike Abbott, Jon Marlowe, David Wexler, and Rebecca Turincio. Joanna Hunter, Vera Delgado, and Can Nhan

ran many of the STR samples on the automated mass spectrometers. Their hard work made it possible to focus on experimental design and data analysis rather than routine sample handling.

It was a great blessing to have talented and supportive coworkers throughout the course of this project. Kathy Stephens, Jia Li, Tom Shaler, Yuping Tan, Christine Loehrlein, Joanna Hunter, Hua Lin, Gordy Haupt, and Nathan Hunt provided useful discussions on a number of issues and helped develop assay parameters and tackle automation issues, among other things. Nathan Hunt was especially important to the success of this project because he developed the STR genotyping algorithm and CallSSR software as well as the multiplex SNP primer design software. Kevin Coopman developed the SNP genotyping algorithm and calling software and was always eager to analyze our multiplex SNP samples. Joe Monforte and Roger Walker served as our supervisors for the first year and second year of this project, respectively, which allowed us the opportunity to devote sufficient time to doing the work

described in this project. Last but not least, Debbie Krantz served as an able administrator of these two NIH grants and took care of the financial aspects.

We also were supported with samples and sequence information from a number of scientific collaborators. Steve Lee and John Tonkyn from the California Department of Justice DNA Laboratory provided genomic DNA samples and STR allelic ladders. Debang Liu from Northwestern University provided the D3S1358 DNA sequence used for improved primer design purposes. Peter Oefner and Peter Underhill from the Department of Genetics at Stanford University provided male population samples and Y-chromosome SNP sequences. The encouragement and support of Lisa Forman and Richard Rau from the Office of Justice Programs at the National Institute of Justice propelled this work from an idea to a working product. In addition, Dennis Reeder from the National Institute of Standards and Technology was always a constant source of encouragement at scientific meetings.

CONTENTS

Acknowledgments	iii
Executive Summary	1
Introduction	1
Purpose of the Report	1
Short Tandem Repeats.....	2
Single Nucleotide Polymorphisms.....	5
Conclusions and Implications	6
Project Description	9
STR Grant	9
SNP Grant.....	10
Scope and Methodology	21
Assay Development and Primer Testing.....	22
Sample Cleanup and Mass Spectrometry.....	29
Sample Genotyping.....	31
Comparison Tests With ABI 310 Genetic Analyzer	33
Results and Discussion of STR Analysis by Mass Spectrometry	35
Marker Selection and Feasibility Studies With STR Loci	35
Multiplex STR Work	40
Comparison Tests Between ABI 310 and Mass Spectrometry Results.....	40
PCR Issues.....	41
Analytical Capabilities of This Mass Spectrometry Method	51
Results and Discussion of Multiplex SNPs	63
Mitochondrial DNA Work.....	64
Y-Chromosome Work	66
References.....	73
Published Papers and Presentations	75

List of exhibits included with this report:

Exhibit 1. Schematic of GeneTrace automated time-of-flight mass spectrometer	2
Exhibit 2. PCR product sizes with newly designed primers for commonly used STR loci compared with commercially available primers used in multiplex sets for fluorescence-based assays	3
Exhibit 3. Fluorescent multiplex STR result with AmpF1STR® COfiler™ kit and ABI 310 Genetic Analyzer	4
Exhibit 4. Schematic of expected allele masses for a CSF1PO-TPOX-THO1 (CTT) multiplex involving overlapping allele size ranges.....	4
Exhibit 5. Mass spectrum of an STR triplex involving TPOX, THO1, and CSF1PO	5
Exhibit 6. Mass spectrum of SNP 10-plex assay for screening polymorphic sites in the mtDNA control region	6
Exhibit 7. Overlay of CE electropherograms demonstrating male-specific amplification with the 17-plex set of PCR primers.....	7
Exhibit 8. Mass spectrum of an STR multiplex sample with nonoverlapping alleles	9
Exhibit 9. Mass spectrum of STR multiplex mixture of THO1 and TPOX allelic ladders	10
Exhibit 10. Mass spectrum of a multiplexed THO1 and TPOX sample showing interleaved amplicons that are only 85 Da apart	10
Exhibit 11. STR data collection times for CDOJ samples	11
Exhibit 12. CDOJ CSF1PO results compared with ABI 310 and mass spectrometry methods	12
Exhibit 13. CDOJ TPOX results compared with ABI 310 and mass spectrometry methods	13
Exhibit 14. CDOJ THO1 STR results compared with ABI 310 and mass spectrometry methods	14
Exhibit 15. CDOJ amelogenin results compared with ABI 310 and mass spectrometry methods	15
Exhibit 16. CDOJ D3S1358 STR results compared with ABI 310 and mass spectrometry methods	16
Exhibit 17. CDOJ D16S539 STR results compared with ABI 310 and mass spectrometry methods	17
Exhibit 18. CDOJ D7S820 STR results compared with ABI 310 and mass spectrometry methods	18
Exhibit 19. CDOJ FGA STR results compared with ABI 310 and mass spectrometry methods	19

Exhibit 20. Schematic of the STR assay using the GeneTrace cleavable primer technology	21
Exhibit 21. Schematic of the SNP assay using the GeneTrace cleavable primer technology	22
Exhibit 22. D3S1358 sequence with PCR primer locations	23
Exhibit 23. Expected mass-to-charge ratios of various ions in the mtDNA 10-plex assay	24
Exhibit 24. Primer sequences designed for STR markers tested by mass spectrometry	24
Exhibit 25. CEPH family pedigrees for samples examined	25
Exhibit 26. Mitochondrial DNA primers used for 10-plex SNP reaction	26
Exhibit 27. Multiplex PCR primers used for Y SNP markers	27
Exhibit 28. CE electropherograms of “dropout” experiments conducted on a 9-plex PCR primer set used in developing Y-chromosome SNP markers	28
Exhibit 29. Photo of the automated GeneTrace mass spectrometry ion optics over a sample plate containing 384 different DNA samples	30
Exhibit 30. Mass ladder file for STR loci analyzed in this study	31
Exhibit 31. Mass spectra of CDOJ samples amplified with CSF1PO primers	32
Exhibit 32. Sample sets run on ABI 310 Genetic Analyzer with AmpF1STR® ProfilerPlus™ or AmpF1STR® COfiler™ fluorescent STR kits	33
Exhibit 33. Tetranucleotide markers identified through literature and public database searches as possible candidates for early STR marker development	35
Exhibit 34. STR markers examined at GeneTrace during the course of this project as sorted by their chromosomal position.....	36
Exhibit 35. Mass spectra for CSF1PO, TPOX, TH01, and amelogenin using K562 DNA	37
Exhibit 36. Mass spectra of STR allelic ladders from CSF1PO, TPOX, TH01, and VWA.....	38
Exhibit 37. Mass spectra of a D8S1179 sample illustrating the benefit of a dideoxynucleotide termination approach	39
Exhibit 38. Multiplexing strategies for STR genotyping using FTA™ paper	41
Exhibit 39. CDOJ D8S1179 STR results with the mass spectrometry method	42
Exhibit 40. CDOJ DYS391 STR results with the mass spectrometry method	43
Exhibit 41. Plot of measured masses versus sample number from four different STR loci	44

Exhibit 42. Effects of sequence variation on PCR amplification in or around STR repeat regions.....	45
Exhibit 43. Mass spectra of CDOJ samples amplified with D7S820 primers	46
Exhibit 44. Electropherograms of ABI 310 results for new STR microvariants seen in the Stanford male population samples	47
Exhibit 45. ProfilerPlus™ results from Stanford male population samples.....	48
Exhibit 46. Plot of allele mass offsets (allele 1 versus allele 2) for heterozygous samples from four different loci	49
Exhibit 47. Plot of X allele mass offset versus Y allele mass offset for 88 amelogenin samples	49
Exhibit 48. Mass spectra comparing an STR sample amplified with TaqGold polymerase and Tsp polymerase.....	50
Exhibit 49. Mass spectrum demonstrating detection of stutter products from a particularly stutter-prone dinucleotide repeat locus	50
Exhibit 50. Primer sequence determination with exonuclease digestion and mass difference measurements	51
Exhibit 51. Mass spectrum of AmpF1STR® Green I primer mix	51
Exhibit 52. TH01 STR primer positions for commercially available primers highlighted on the GenBank sequence	52
Exhibit 53. TPOX STR primer positions for commercially available primers highlighted on the GenBank sequence	53
Exhibit 54. CSF1PO STR primer positions for commercially available primers highlighted on the GenBank sequence	54
Exhibit 55. Mass spectrum of a TH01 allelic ladder reamplified from AmpF1STR® Green I allelic ladders	55
Exhibit 56. Mass spectra of D5S818 allelic ladders from two manufacturers	55
Exhibit 57. Mass spectra of TPOX PCR products from various amounts of K562 DNA template material.....	56
Exhibit 58. Mass spectra of CDOJ samples amplified with TPOX primers	57
Exhibit 59. Mass spectra of CDOJ samples amplified with TH01 primers	58
Exhibit 60. Comparison of allele masses collected 6 months apart.....	59
Exhibit 61. Fifteen replicate analyses of a TPOX allelic ladder to measure mass precision and accuracy.....	60
Exhibit 62. Upper strand (TCAT repeat) and lower strand (AATG repeat) mass differences for the TH01 allelic ladder.....	60
Exhibit 63. Comparison of ABI 310 and mass spectrometry allele calls for 90 CEPH/diversity samples	61

Exhibit 64. Histogram of mass difference measurements for 200 samples (50 for each ddN)	63
Exhibit 65. Schematic representation of the mtDNA control region 10-plex SNP assay	64
Exhibit 66. SNP ions impacting multiplex design	64
Exhibit 67. Multiplex PCR information for 17-plex PCR reaction	66
Exhibit 68. CE electropherograms showing a multiplex PCR sample compared with individual PCR reactions	67
Exhibit 69. Primers for testing Y-chromosome SNP markers in singleplexes	68
Exhibit 70. Y SNP multiplex primer information	68
Exhibit 71. Human autosomal SNP markers designed for testing NIH diversity panel	69
Exhibit 72. Characteristics of STR and SNP markers	70
Exhibit 73. STR genotypes for standard DNA templates K562, AM209, and UP006 obtained using AmpF1STR® ProfilerPlus™ and AmpF1STR® COfiler™ fluorescent STR kits	70
Exhibit 74. An ABI 310 CE electropherogram showing a heterozygous sample with a 15.2 microvariant at STR locus D3S1358	71

EXECUTIVE SUMMARY

Introduction

The advent of DNA typing and its use for human identity testing has revolutionized law enforcement investigations in recent years by allowing forensic laboratories to match suspects with minuscule amounts of biological evidence from a crime scene. Equally important is the use of DNA to exclude suspects who were not involved in a crime or to identify human remains in an accident.

The past decade has seen numerous advances in the DNA testing procedures, most notably among them the development of PCR (polymerase chain reaction)-based DNA typing methods. Technologies for measuring DNA variations, both length and sequence polymorphisms, have also advanced rapidly in the past decade. The time needed to determine a sample's DNA profile has dropped from 6–8 weeks to 1–2 days, and with more recent advancements, the time needed to process samples may decrease to as little as a few hours, maybe even a few minutes.

Simultaneous with the evolution of DNA markers and technologies embraced by the forensic community has been the acceptance and use of DNA typing information. The courtroom battles over statistical issues that were common in the late 1980s and early 1990s have subsided as DNA evidence has become more widely accepted.

In the past 5 years, DNA databases have emerged as powerful tools for criminal investigations, much like the fingerprint databases that have been used routinely for decades.

The United Kingdom launched a nationwide DNA database in 1995 that now contains more than 1 million DNA profiles of convicted felons—profiles that have been used to aid more than 75,000 criminal investigations. National DNA databases are

springing up in countries all over the world as their value to law enforcement is being recognized.

In the United States, the FBI has developed the Combined DNA Index System (CODIS) with the anticipation that several million DNA profiles will be entered into this database in the next decade. All 50 States now have laws requiring DNA typing of convicted offenders, typically for violent crimes such as rape or homicide.

While the law enforcement community is gearing up to gather millions of DNA samples from convicted felons, the DNA typing technology needs improvement. Large backlogs of samples exist today due to the high cost of performing the DNA testing and limited capabilities in forensic laboratories. As of the summer of 1999, several States, including California, Virginia, and Florida, had backlogs of more than 50,000 samples. A need exists for more rapid and cost-effective methods for high-throughput DNA analysis to process samples currently being gathered for large criminal DNA databases around the world.

At the start of this project in June 1997, commercially available slab gel or capillary electrophoresis instruments could handle only a few dozen samples per day. While larger numbers of samples can be processed by increasing the number of laboratory personnel and instruments, the development of high-throughput DNA processing technologies promises to be more cost effective in the long run, especially for the generation of large DNA databases. GeneTrace Systems, Inc., a small biotechnology company located in Alameda, California, has developed high-throughput DNA analysis capabilities using time-of-flight mass spectrometry coupled with parallel sample preparation on a robotic workstation. The GeneTrace technology allows several thousand samples to be processed daily. DNA samples can be analyzed in seconds, rather than minutes or

hours, and with improved accuracy compared with conventional electrophoresis methods.

Overall, the mass spectrometry method described in this study is two orders of magnitude faster in sample processing time than conventional techniques.

Purpose of the Report

This NIJ project was initiated to adapt the GeneTrace technology to human identity DNA markers commonly used by forensic DNA laboratories, specifically short tandem repeat (STR) markers. An extension of the original grant was submitted in December 1997 to fund the development of single nucleotide polymorphism (SNP) markers from mitochondrial DNA and the Y chromosome.

Based on the results obtained in this study, the authors believe mass spectrometry can be a useful and effective means for high-throughput DNA analysis, and that it has the capabilities to meet the needs of the forensic DNA community for offender DNA databases.

However, due to limited resources and a perceived difficulty to enter the forensic DNA market, GeneTrace made a business decision to not pursue this market. While the STR milestones on the original grant were met, only the initial milestones were achieved on the SNP portion of the NIJ grant because of the premature termination on the part of GeneTrace.

GeneTrace Systems, Inc., developed an integrated high-throughput DNA analysis system involving the use of proprietary chemistry, robotic sample manipulation, and time-of-flight mass spectrometry. The purpose of this NIJ project was to apply the GeneTrace technology to improve the analysis of STR markers commonly used in forensic DNA laboratories.

Mass spectrometry is a versatile analytical technique that involves the detection of ions and the measurement of their mass-to-charge ratio. Because these ions are separated in a vacuum environment, analysis times can be extremely rapid, often within microseconds. Many advances have been made in the past decade for the analy-

sis of biomolecules such as DNA, proteins, and carbohydrates since the introduction of a new ionization technique known as matrix-assisted laser desorption-ionization (MALDI) and the discovery of new matrixes that effectively ionize DNA without extensive fragmentation. When coupled with time-of-flight mass spectrometry,

this method for measuring biomolecules is commonly referred to as MALDI-TOF-MS. A schematic of MALDI-TOF-MS is presented in exhibit 1.

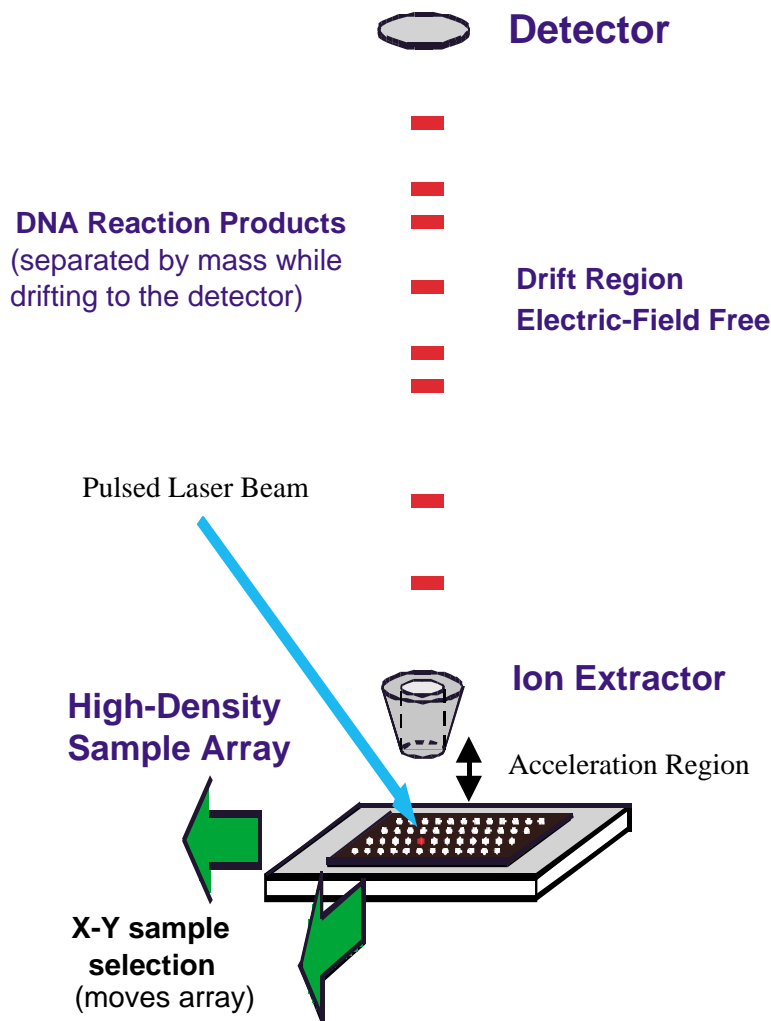
Short Tandem Repeats

Short tandem repeat (STR) DNA markers, also referred to as microsatellites or simple sequence repeats (SSRs), consist of tandemly repeated DNA sequences with a core repeat of 2–6 base pairs (bp). STR markers are readily amplified during PCR by using primers that bind in conserved regions of the genome flanking the repeat region. Forensic laboratories prefer tetranucleotide loci (i.e., 4 bp in the repeat) due to the lower amount of “stutter” produced during PCR. (Stutter products are additional peaks that can complicate the interpretation of DNA mixtures by appearing in front of regular allele peaks.) The number of repeats can vary from 3 or 4 repeats to more than 50 repeats with extremely polymorphic markers. The number of repeats, and hence the size of the PCR product, may vary among samples in a population making STR markers useful in identity testing or genetic mapping studies.

Shortly after this project was initiated, the FBI designated 13 core STR loci for the nationwide CODIS database. These STR loci are TH01, TPOX, CSF1PO, VWA, FGA, D3S1358, D5S818, D7S820, D13S317, D16S539, D8S1179, D18S51, and D21S11. The sex-typing marker, amelogenin, is also included in STR multiplexes that cover the 13 core STR loci. Each sample must have these 14 markers tested to be entered into CODIS.

To illustrate the kinds of numbers involved to analyze the current national sample backlog of ~500,000 samples, more than 7 million genotypes must be generated. Using currently available technologies, an estimated \$25 million (~\$50/sample) and more than 5 years for well-trained and well-funded

Exhibit 1. **Schematic of GeneTrace automated time-of-flight mass spectrometer.** DNA molecules are liberated from a solid-phase matrix environment with a laser pulse. The DNA reaction products are separated by size (mass) in a matter of microseconds, as opposed to hours using conventional methods. For each run, hundreds of samples are prepared in parallel using a robotic workstation and spotted on a sample plate that is introduced to the vacuum environment of the mass spectrometer. The sample plate moves under the fixed laser beam to allow sequential sample analysis.



laboratories would be required to determine those 7 million genotypes. With the high cost and effort required, most of these backlogged samples are being stored in anticipation of future

analysis and inclusion in CODIS, pending the development of new, faster technology or the implementation of more instruments using the current electrophoresis technologies.

Time-of-flight mass spectrometry has the potential to bring DNA sample processing to a new level in terms of high-throughput analysis. However, there are several challenges to using MALDI-TOF-MS for the analysis of PCR products, such as STR markers. Mass spectrometry resolution and sensitivity are diminished when either the DNA size or the salt content of the sample is too large. By redesigning the PCR primers to bind close to the repeat region, the STR allele sizes are reduced to benefit the resolution and sensitivity of the PCR products. Therefore, much of this project involved designing and testing new PCR primers that produced smaller amplicon sizes for STR markers of forensic interest. This research focused on STR loci that have been developed by commercial manufacturers and studied extensively by forensic scientists. These include all of the GenePrint™ tetranucleotide STR systems from Promega Corporation (Madison, WI) as well as the 13 CODIS STR loci that are covered by the Profiler Plus™ and Cofiler™ kits from Applied Biosystems (ABI) (Foster City, CA) (exhibit 2). Where possible, primers were designed to produce amplicons less than 100 bp in size, although it has been possible to resolve neighboring STR alleles as large as 140 bp. For example, TPOX alleles 6–14 ranged from 69–101 bp in size with GeneTrace-designed primers; while with Promega's GenePrint™ primers, the same TPOX alleles ranged in size from 224–256 bp. Unfortunately, due to the long and complex repeat structures of several STR markers, this study was unable to obtain the necessary single-base resolution with the following STR loci: D21S11, D18S51, and FGA (see Results and Discussion of STR Analysis by Mass Spectrometry).

To verify the STR results obtained from the mass spectrometry method, the authors collaborated with the California Department of Justice (CDOJ) DNA Laboratory in Berkeley to generate a large data set. CDOJ provided 88 samples that had been previously genotyped

Exhibit 2. **PCR product sizes with newly designed primers for commonly used STR loci compared with commercially available primers used in multiplex sets for fluorescence-based assays.** The red numbers indicate PCR product size ranges that exceed the recommended 140 bp mass spectrometry detection range.

STR Locus	Known Alleles	GeneTrace Sizes (newly designed primers in this study)	Commerically Available Sizes*	
			Applied Biosystems	Promega
Amelogenin	X, Y	106, 112 bp	106, 112 bp	212, 218 bp
CD4	4–15	81–136 bp	N.A.	
CSF1PO	6–15	87–123 bp	280–316 bp	291–327 bp
F13A1	3–17	112–168 bp	N.A.	279–335 bp
F13B	6–12	110–134 bp	N.A.	169–193 bp
FES/FPS	7–15	76–108 bp	N.A.	222–254 bp
FGA	15–30	118–180 bp	206–266 bp	N.A.
D3S1358	9–20	76–120 bp	101–145 bp	N.A.
D5S818	7–15	89–121 bp	134–166 bp	119–151 bp
D7S820	6–14	66–98 bp	257–289 bp	215–247 bp
D8S1179	8–18	92–130 bp	127–167 bp	N.A.
D13S317	7–15	98–130 bp	201–233 bp	165–197 bp
D16S539	5,8–15	81–121 bp	233–273 bp	264–304 bp
D18S51	9–27	120–192 bp	272–344 bp	N.A.
D21S11	24–38	150–190 bp	186–242 bp	N.A.
DYS19	8–16	76–108 bp	N.A.	
DYS391	9–12	99–111 bp	N.A.	
HPRTB	6–17	84–128 bp	N.A.	259–303 bp
LPL	7–14	105–133 bp	N.A.	105–133 bp
TH01	3–13.3	55–98 bp	160–203 bp	171–214 bp
TPOX	6–14	69–101 bp	217–249 bp	224–256 bp
VWA	11–22	126–170 bp	156–200 bp	127–171 bp
Other STRs				
GATA132B04	10–14	99–115 bp	N.A.	
D22S445	10–16	110–130 bp	N.A.	
D16S2622	4–8	71–87 bp	N.A.	

* Sizes are listed without adenylation (add 1 base for +A form).
N.A. = Not Available.

using validated fluorescent multiplex STR kits from ABI. GeneTrace generated STR results using their primer sets for 9 STR loci (TH01, TPOX, CSF1PO, D3S1358, D16S539, D8S1179, FGA, DYS391, and D7S820) along with the sex-typing marker amelogenin. These experiments compared more than 700 genotypes (88 samples X 8 loci; data from D8S1179 and DYS391 were not available from CDOJ). Although results were not obtained for all possible samples using mass spectrometry, researchers observed almost 100% correlation with the genotypes obtained between the validated fluorescent STR method and GeneTrace's newly developed mass spectrometry technique, demonstrating that the GeneTrace method was reliable (see Results and Discussion of STR Analysis by Mass Spectrometry).

Multiplex STR analysis

To reduce analysis cost and sample consumption and to meet the demands of higher sample throughputs, PCR amplification and detection of multiple markers (multiplex STR analysis) has become a standard technique in most forensic DNA laboratories. STR multiplexing is most commonly performed using spectrally distinguishable fluorescent tags and/or nonoverlapping PCR product sizes. An example of an STR multiplex produced from a commercially available kit is shown in exhibit 3. Multiplex STR amplification in one or two PCR reactions with fluorescently labeled primers and measurement with gel or capillary electrophoresis separation and laser-induced fluorescence detection is becoming a standard method among forensic laboratories for analysis of the 13 CODIS STR loci. The STR alleles from these multiplexed PCR products typically range in size from 100–350 bp with commercially available kits.

Due to the limited DNA size constraints of mass spectrometry, GeneTrace adopted a different approach to multiplex analysis of multiple STR loci. Primers

Exhibit 3. Fluorescent multiplex STR result with AmpFISTR® COfiler™ kit and ABI 310 Genetic Analyzer. The red peaks are an internal sizing standard (GS350-ROX). The DNA size window shown here is 90–350 bp. The allele calls for the seven loci in this multiplex are amelogenin: X,Y; D3S1358: 15, 17; TH01: 6, 7; TPOX: 8, 12; D16S539: 8, 12; D7S820: 9, 11; and CSF1PO: 12, 12.

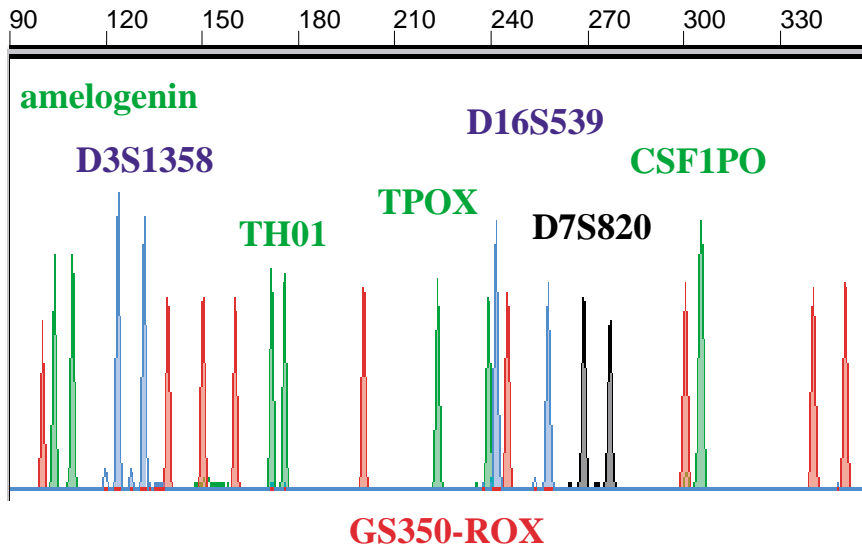


Exhibit 4. Schematic of expected allele masses for a CSF1PO-TPOX-TH01 (CTT) multiplex involving overlapping allele size ranges. All known alleles are fully distinguishable by mass with this interleaving approach.

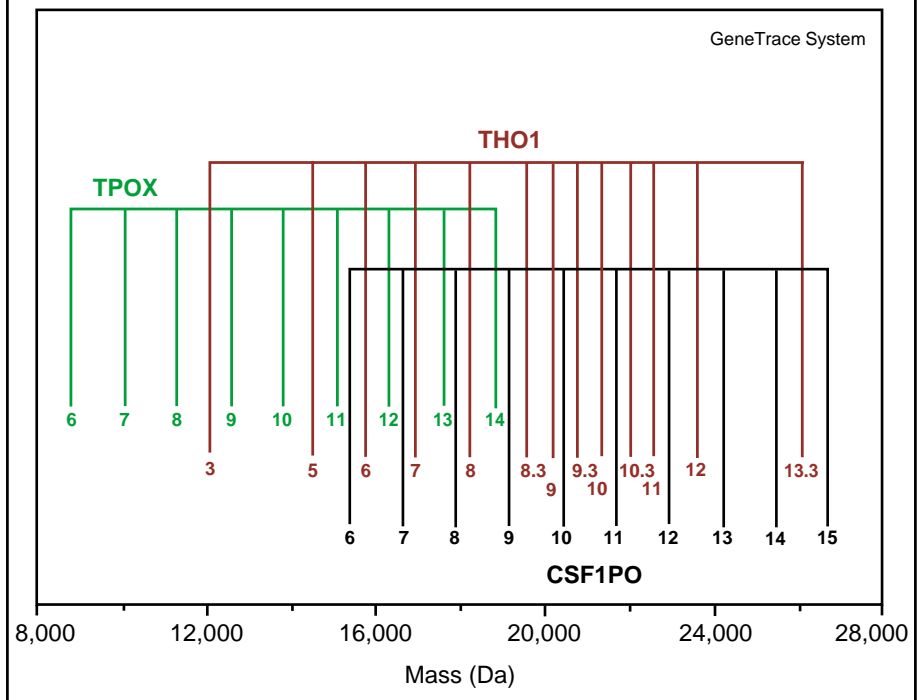
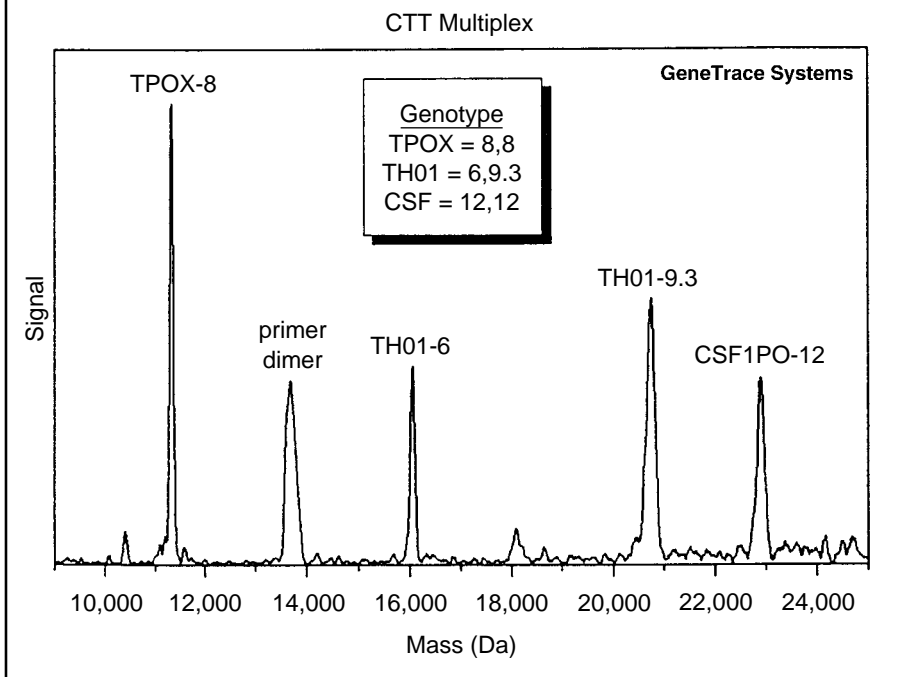


Exhibit 5. Mass spectrum of an STR triplex involving TPOX, TH01, and CSF1PO. This sample is a mass spectrometry result using the interleaving allele approach schematically illustrated in exhibit 4. Multiplex PCR and multiplex primer extension with ddC termination were used to obtain this result.



are designed such that the PCR product size ranges overlap between multiple loci but have alleles that interleave and are resolvable in the mass spectrometer (exhibit 4). As described above, PCR primers are closer to the STR repeat regions than those commonly used with electrophoresis systems. The high accuracy, precision, and resolution of this mass spectrometry approach permits multiplexing STR loci for a limited number of markers. During the study, GeneTrace also developed a TH01-TPOX-CSF1PO STR triplex (exhibit 5).

Single Nucleotide Polymorphisms

Single nucleotide polymorphisms (SNPs) represent another form of DNA variation that is useful for human identity testing. SNPs are the most frequent form of DNA sequence

variation in the human genome and are becoming increasingly popular genetic markers for genome mapping studies and medical diagnostics. SNPs are typically biallelic with two possible nucleotides (nt) or alleles at a particular site in the genome. Because SNPs are less polymorphic (i.e., have fewer alleles) than the currently used STR markers, more SNP markers are required to obtain the same level of discrimination between samples. Current estimates are that 30–50 unlinked SNPs will be required to obtain the matching probabilities of 1 in ~100 billion as seen with the 13 CODIS STRs.

The perceived value of SNPs for DNA typing in a forensic setting include the following:

- ◆ More rapid analysis.
- ◆ Cheaper costs.

- ◆ Simpler interpretation of results because there are no stutter products.
- ◆ Improved ability to handle degraded DNA because of the possibility of smaller PCR product sizes.

While it is doubtful that autosomal SNPs will replace the current battery of STRs used in forensic laboratories in the near future, abundant mitochondrial and Y-chromosome SNP markers exist and have already proven useful as screening tools. These maternal (mitochondrial) and paternal (Y chromosome) lineage markers are effective in identifying missing persons and war casualties and helping answer historical questions such as whether or not Thomas Jefferson fathered a slave child.

The forensic DNA community already has experience with applying SNP markers as a screening process, which can prove very helpful for excluding suspects from crime scenes. Many crime laboratories still use reverse dot blot technology for analyzing the SNPs from HLA-DQA1 and PolyMarker loci with kits from ABI. In addition, mitochondrial DNA (mtDNA) sequencing is currently performed in some forensic laboratories.

In the work performed on multiplex SNP markers at GeneTrace, the authors examined 10 polymorphic sites within the mtDNA control region and 20 Y-chromosome SNPs provided by Dr. Peter Oefner and Dr. Peter Underhill from Stanford University. A multiplex SNP assay was developed for 10 mtDNA SNP sites (exhibit 6). Only limited work was performed on the Y-chromosome SNPs due to the premature termination of the work. However, results demonstrated a male-specific 17-plex PCR of 17 different Y SNP markers (exhibit 7).

Conclusions and Implications

Time-of-flight mass spectrometry offers a rapid, cost-effective alternative for genotyping large numbers of samples. Each DNA sample can be accurately measured in a few seconds. Due to the increased accuracy of mass spectrometry, STR alleles can be reliably typed without comparison with allelic ladders. Mass spectrometry holds significant promise as a technology for high-throughput DNA processing that will be valuable for large-scale DNA database work.

In summary, the positive features of mass spectrometry for STR analysis include:

- ◆ Rapid results—STR typing at a rate of seconds per sample.
- ◆ Accuracy—no allelic ladders.
- ◆ Direct DNA measurement—no fluorescent or radioactive labels.
- ◆ Automated sample preparation and data collection.
- ◆ High-throughput capabilities of thousands of samples daily per system.
- ◆ Flexibility—single nucleotide polymorphism (SNP) assays can be run on the same instrument platform.

This project demonstrates that both STR and SNP analysis are reliably performed with GeneTrace's mass spectrometry technology. Tests were done on a large number of human DNA markers of forensic interest. New primer sets were developed for the 13 CODIS STR loci that may prove useful in the future for situations in which degraded DNA is present and requires smaller amplicons to obtain

successful results. The possibility of developing multiplexed SNP markers also was explored, and a mtDNA 10-plex assay and Y-chromosome, 17-plex, male-specific PCR were demonstrated. Both STR and SNP areas appear promising for future research. In another project, GeneTrace

recently demonstrated a sample throughput of approximately 4,000 STR samples in a single day with a single automated mass spectrometer. Clearly, this is an improvement in the analysis of DNA short tandem repeat markers using time-of-flight mass spectrometry.

Exhibit 6. Mass spectrum of SNP 10-plex assay for screening polymorphic sites in the mtDNA control region. The bottom panel shows the 10 SNP primers prior to the primer extension reaction. The top panel contains the multiplexed reaction products, each labeled with the observed extension product. The results for this K562 PCR product are (in order across HV1 and HV2): H16069 (G), H16129 (C), H16189 (A), H16224 (G), H16311 (A), H00073 (C), L00146 (T), H00152 (A), L00195 (T), and H00247 (C). SNP nucleotide results have been confirmed by sequencing. The primer sequences are listed in exhibit 24. The primer concentrations were 25 pmol MT5, 15 pmol MT8', 15 pmol MT10, 10 pmol MT3', 20 pmol MT9, 25 pmol MT6, 20 pmol MT2g, 27 pmol MT1, 35 pmol MT7, and 20 pmol MT4e.

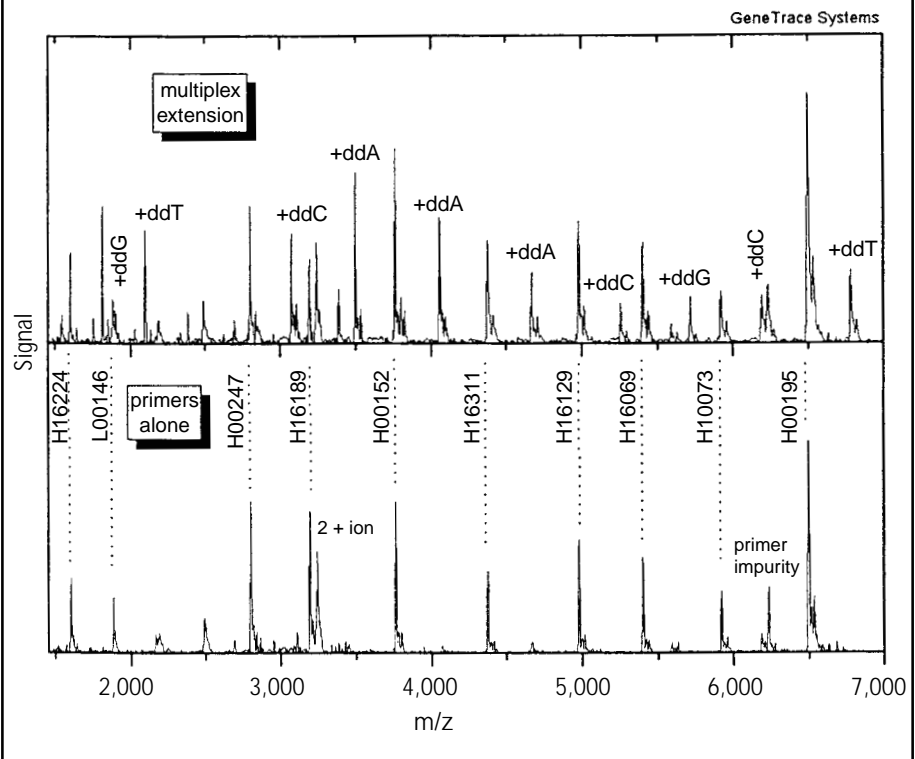
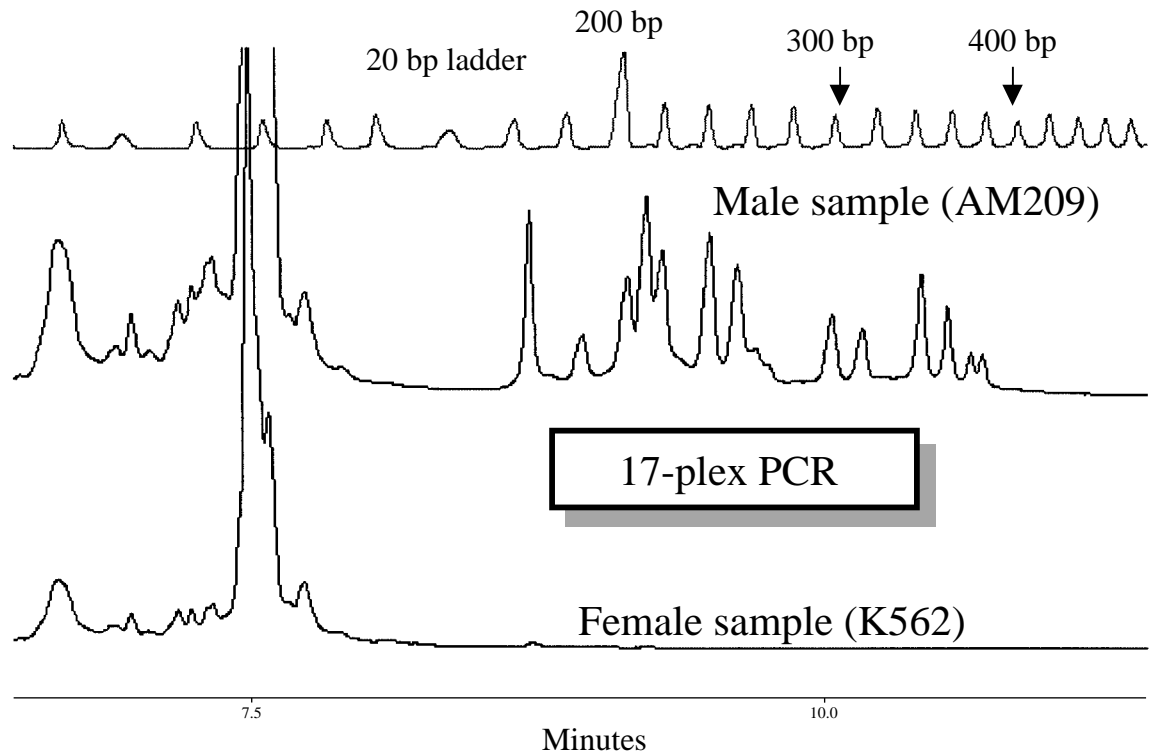


Exhibit 7. **Overlay of CE electropherograms demonstrating male-specific amplification with the 17-plex set of PCR primers.** Note: The female sample (K562) failed to yield any peaks, illustrating that the PCR reaction is Y chromosome-specific. The PCR primers used are described in exhibit 27 with each locus-specific primer set at 0.4 pmole each. For the PCR reaction conditions, see exhibit 67 and details in the text. The 17 amplicons may be seen more clearly in exhibit 68.



PROJECT DESCRIPTION

STR Grant

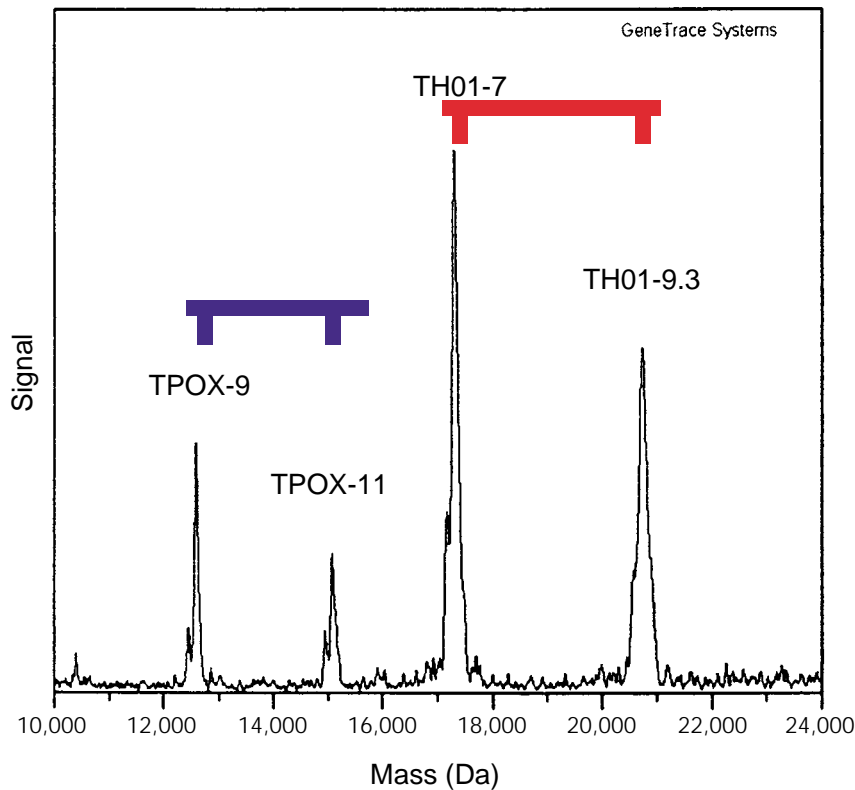
This project focused on the development of a powerful new technology for rapid and accurate analysis of DNA STR markers using time-of-flight mass spectrometry. GeneTrace Systems, Inc., collaborated with the CDOJ DNA Laboratory in Berkeley, California, primarily through Dr. Steve Lee. This collaboration provided the study with the samples used to verify the new GeneTrace technology, which was done by comparing the mass spectrometry results with genotypes obtained using established and validated methods run at CDOJ.

To accomplish the task of developing a new mass spectrometry technology for STR typing, five milestones were proposed in the original grant application, which included the following:

- ◆ Redesign PCR primers for a number of commonly used STR markers to produce smaller PCR products that could be tested in the mass spectrometer (exhibit 2).
- ◆ Demonstrate multiplexing capabilities to a level of 2 or 3 for detection with the TOF-MS method (exhibits 4–5 and 8–10).
- ◆ Transfer the sample preparation protocols from manual to a highly parallel and automated pipetting robot.
- ◆ Develop a large data set to confirm the accuracy and reliability of this method (exhibits 11–19).
- ◆ Automate and incorporate DNA extraction techniques onto the GeneTrace robots.

As described in the results section, all milestones were met on time except the final one regarding DNA extraction. Two other companies, Rosys and Qiagen, produced robotic systems for DNA extraction after this project began. Meanwhile, GeneTrace remained focused on developing

Exhibit 8. Mass spectrum of an STR multiplex sample with nonoverlapping alleles. The two loci, TH01 and TPOX, were coamplified using the ddC termination approach. Note: The mass accuracy is improved for peaks closer to the calibration standard of 10,998 Da.



Expected Mass	Observed Mass	Δ Mass
TPOX-9 = 12,581 Da	12,584 Da	+3 Da
TPOX-11 = 15,101 Da	15,080 Da	-21 Da
TH01-7 = 17,336 Da	17,305 Da	-31 Da
TH01-9.3 = 20,803 Da	20,751 Da	-52 Da

other steps in DNA sample processing since commercially available solutions had already been developed, thereby eliminating the need to include the DNA extraction portion in this study.

Since this project began in June 1997, a number of advances that impact the ability to perform high-throughput DNA typing have occurred in the biotech field. In early 1998, ABI released a dual 384-well PE9700 (“Viper”) thermal cycler, which makes it possible to prepare 768 PCR samples simultaneously. Beckman Instruments

also came to market with a 96-tip Multitek pipetting robot. At the beginning of this project, GeneTrace used funds from this NIJ grant to purchase an MJ Research 384-well thermal cycler and a custom-built 96-tip robotic pipettor on a CyberLab x-y-z gantry. Both of these pieces of equipment were the state of the art at the time but are now obsolete at GeneTrace for routine operations and have been replaced by the newer and more reliable products from Applied Biosystems and Beckman Instruments.

Exhibit 9. Mass spectrum of STR multiplex mixture of TH01 and TPOX allelic ladders. The PCR products from the two loci differ by 120 Da. The ladders were reamplified from AmpF1STR® Green I kit materials. The TH01 ladder ranges from 5 to 10 repeats with 9.3 included; the TPOX ladder ranges from 6 to 13 repeats with 13 not shown here.

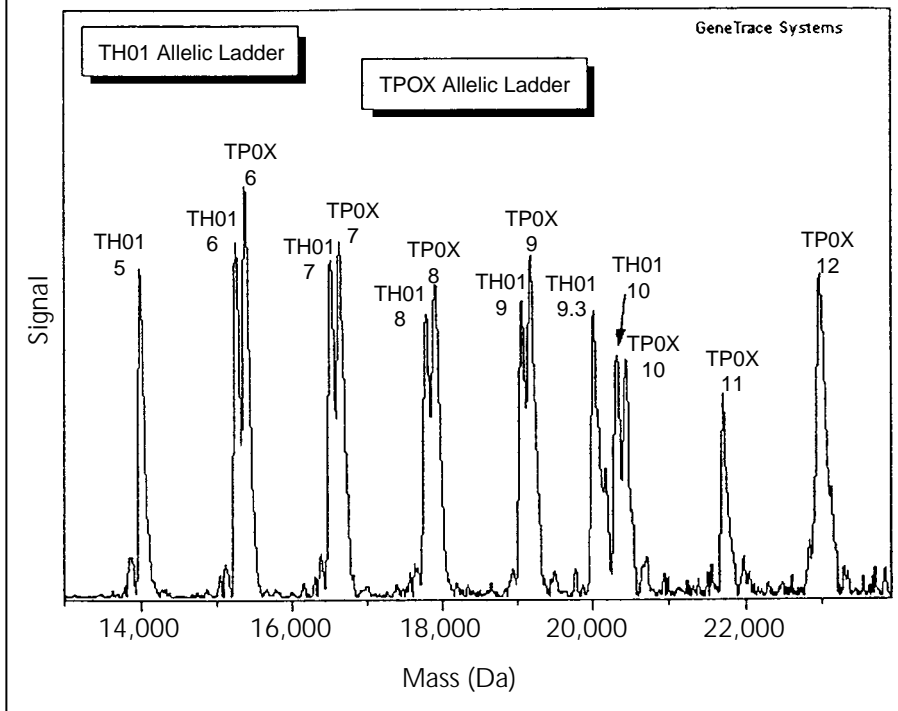
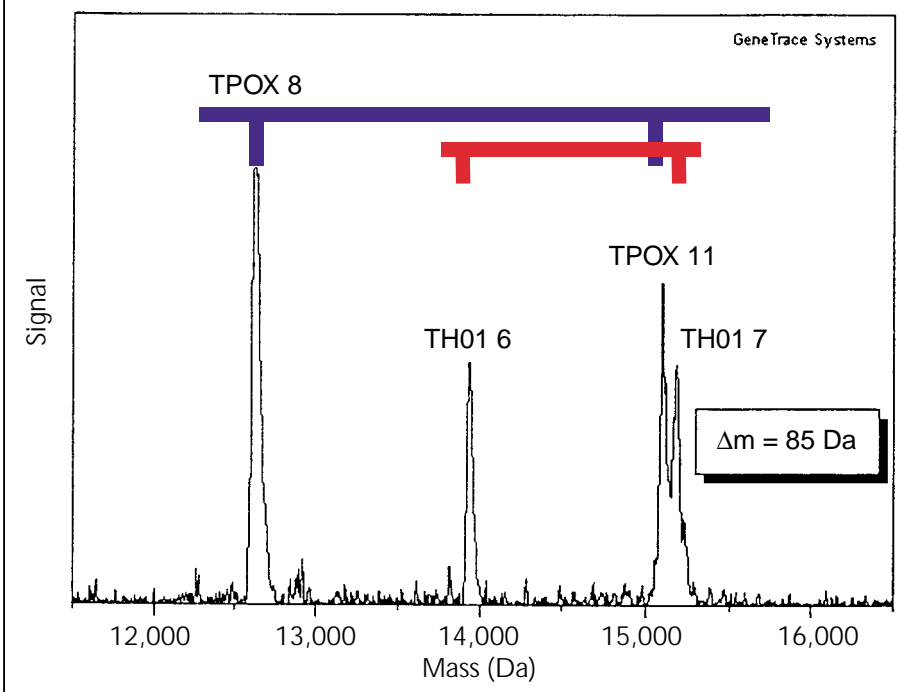


Exhibit 10. Mass spectrum of a multiplexed TH01 and TPOX sample showing interleaved amplicons that are only 85 Da apart



SNP Grant

The grant extension, which began in August 1998, focused on the development of multiplexed SNP markers from mtDNA and the Y chromosome. Although the grant extension was terminated prematurely by GeneTrace management in April 1999, portions of the first four milestones were accomplished. The five milestones described in the original grant extension included the following:

- ◆ Produce and test a set of 10 or more SNP probes for mtDNA control region “hot spots” (exhibit 5).
- ◆ Develop software for multiplex SNP analysis and data interpretation.
- ◆ Examine individual Y-chromosome SNP markers.
- ◆ Develop multiplex PCR and multiplex SNP probes for Y-chromosome SNP loci (exhibit 7).
- ◆ Determine the discriminatory power for a set of Y-chromosome markers by running ~300 samples across 50 Y-chromosome SNP markers.

The goal of the grant extension project was to develop highly multiplexed SNP assays that worked in a robust manner with mass spectrometry and could be genotyped in an automated fashion. GeneTrace planned to select markers with a high degree of discrimination to aid in rapid screening of mitochondrial DNA and Y-chromosome polymorphisms with the capability to handle analysis of large databases of offender DNA. At the time this proposal was written (December 1997), GeneTrace still intended to provide reagents and instruments to large DNA service laboratories or to provide a DNA typing service to the forensic DNA community. The grant extension was prematurely terminated due to a change in business focus and a need to consolidate the research efforts at GeneTrace.

Exhibit 11. **STR data collection times for CDOJ samples.** Comparing ABI 310 run times with multiple mass spectrometry runs indicates that data collection speed increased almost 10-fold during 1998 due to a number of improvements in the GeneTrace process. Data collection speed is a combination of laser rate, total number of shots taken, sample cleanliness, and other factors. Overall, the mass spectrometry data collection speed is approximately two orders of magnitude faster than the ABI 310. Exhibits 12–19, 39, and 40 list the observed masses for these samples measured by GeneTrace’s mass spectrometry method.

ABI 310 Runs

January 5, 1999–January 8, 1999:

- Multiplex STR detection and analysis (COfiler™ kit amplifies 6 STRs + amelogenin simultaneously: TH01, TPOX, CSF1PO, D3S1358, D16S539, D7S820)
- ~54 hours of continuous operation (107 injections that included 88 samples, 9 allelic ladders, and reinjections of 10 samples at ~30 min per injection)
- ~2209 sec/sample or 316 sec/genotype (~5 min/genotype)

Mass Spectrometry Runs

Date	STR Loci Tested	Number of samples	Time Required for Data Collection	Average Time per Sample
Feb 11, 1998	TH01	88 + 8 controls	76 min for 96 samples	~50 sec/sample
May 6, 1998	TH01	6 samples	2.5 min for 6 samples	~25 sec/sample
October 1, 1998	TPOX, CSF1PO, D3S1358, D7S820	72 + 8 controls for each locus	60 min for 320 samples	11.3 sec/sample
Jan 12, 1999	TH01, TPOX, CSF1PO, D7S820	88 + 8 controls for each locus	30 min for 384 samples	4.7 sec/sample
Jan 14, 1999	TH01, D16S539, D3S1358, Amelogenin	88 + 8 controls for each locus	37 min for 384 samples	5.8 sec/sample
Feb 12, 1999	D3S1358, FGA, D8S1179, DYS391	88 + 8 controls for each locus	41 min for 384 samples	6.4 sec/sample
March 26, 1999	TH01, TPOX, CSF1PO, D16S539	88 + 8 controls for each locus	51 min for 384 samples	8.0 sec/sample

Exhibit 12. CDOJ CSF1PO results compared with ABI 310 and mass spectrometry methods									
Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)	Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)
A1	11,12	11,12	27,823	29,060	A7	11,12	11,12	27,867	29,000
B1	10,12	10,12	26,520	29,033	B7	7,13	7,13	22,809	30,271
C1	7,10	7,10	22,917	26,684	C7	10,13	10,13	26,529	30,253
D1	10,13	10,13	26,586	30,344	D7	10,11	10,11	26,621	27,794
E1	8,12	8,12	24,077	29,075	E7	10,11	10,11	26,609	27,834
F1	8,10	8,10	24,107	26,639	F7	7,12	7,12	22,923	29,202
G1	12,13	12,13	29,144	30,373	G7	11,12	11,12	27,958	29,244
H1	11,12	11,12	27,812	29,018	H7	10,12	10,12	26,497	29,005
A2	12,12	12,12	29,033		A8	12,12	12,12	29,129	
B2	10,10	10,10	26,594		B8	8,11	8,11	24,117	27,862
C2	10,10	10,10	26,550		C8	7,12	7,12	22,913	29,071
D2	10,11	10,11	26,609	27,823	D8	10,10	10,10	26,575	
E2	10,12	10,12	26,620	29,129	E8	10,12	10,12	26,552	29,007
F2	10,10	10,10	26,584		F8	8,10	8,10	24,061	26,572
G2	11,12	11,12	27,862	29,080	G8	12,12	12,12	29,064	
H2	10,10	10,10	26,605		H8	11,12	11,12	27,796	29,026
A3	12,12	12,12	28,987		A9	8,12	8,12	24,085	29,091
B3	10,13	10,13	26,614	30,380	B9	11,12	11,12	27,819	29,022
C3	10,10	10,10	26,618		C9	11,12	11,12	27,810	29,000
D3	11,11	11,11	27,823		D9	8,13	8,13	24,093	30,348
E3	11,12	11,12	27,817	28,998	E9	11,12	11,12	27,878	29,098
F3	11,12	11,12	27,871	29,093	F9	7,12	7,12	22,836	29,104
G3	10,11	10,11	26,616	27,812	G9	9,13	9,13	25,387	30,389
H3	11,12	11,12	27,849	29,024	H9	7,11	7,11	22,866	27,836
A4	8,10	8,10	24,048	26,499	A10	9,10	9,10	25,391	26,607
B4	11,13	11,13	27,941	30,412	B10	11,12	11,12	27,844	29,044
C4	11,11	11,11	27,873		C10	10,10	10,10	26,586	
D4	10,11	10,11	26,638	27,794	D10	8,11	8,11	24,091	27,838
E4	7,10	7,10	22,835	26,553	E10	11,12	11,12	27,843	29,040
F4	10,14	10,14	26,592	31,565	F10	10,12	10,12	26,607	29,111
G4	12,14	12,14	29,113	31,625	G10	8,11	8,11	24,029	27,807
H4	9,10	9,10	25,360	26,582	H10	8,10	8,10	24,099	26,609
A5	10,10	10,10	26,590		A11	12,12	12,12	29,011	
B5	7,11	7,11	22,818	27,838	B11	7,12	7,12	22,836	29,095
C5	10,10	10,10	26,628		C11	7,11	7,11	22,826	27,834
D5	10,11	10,11	26,647	27,925	D11	10,11	10,11	26,567	27,784
E5	11,12	11,12	27,847	29,031	E11	10,12	10,12	26,577	29,078
F5	10,10	10,10	26,601		F11	10,11	10,11	26,611	27,825
G5	10,12	10,12	26,603	29,113	G11	10,10	10,10	26,605	
H5	11,13	11,13	27,856	30,389	H11	10,12	10,12	26,584	29,078
A6	11,11	11,11	27,832		SUMMARY: Complete agreement observed among all genotypes.				
B6	No data	10,12	26,654	29,180					
C6	11,11	11,11	27,862						
D6	11,12	11,12	27,818	29,079					
E6	11,12	11,12	27,838	29,049					
F6	8,10	8,10	24,206	26,739					
G6	11,12	11,12	27,843	29,062					
H6	11,11	11,11	27,773						

Exhibit 13. CDOJ TPOX results compared with ABI 310 and mass spectrometry methods

Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)	Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)
A1	8,10	8,10	17,914	20,431	A7	9,9	5,5	Gas-phase dimer	
B1	6,8	6,8	15,376	17,897	B7	8,8	8,8	17,820	
C1	9,11	9,11	19,137	21,663	C7	10,12	10,12	20,361	22,830
D1	11,12	11,12	21,562	22,801	D7	11,12	6,6	Gas-phase dimer	
E1	9,10	9,10	19,131	20,374	E7	6,8	6,8	15,366	17,850
F1	9,12	9,12	19,106	22,889	F7	10,10	10,10	20,372	
G1	8,8	8,8	17,838		G7	8,9	8,9	17,820	19,070
H1	8,8	8,8	17,825		H7	9,9	9,9	19,098	
A2	8,8	8,8	17,837		A8	9,10	5,5	Gas-phase dimer	
B2	8,8	8,8	17,841		B8	9,9	5,5	Gas-phase dimer	
C2	8,9	8,9	17,860	19,101	C8	8,10	8,10	17,848	20,366
D2	8,9	8,9	17,785	19,049	D8	8,11	8,11	17,843	21,596
E2	8,11	8,11	17,829	21,577	E8	9,9	9,9	19,117	
F2	8,12	8,12	17,834	22,816	F8	9,12	9,12	19,108	22,793
G2	7,10	7,10	16,525	20,257	G8	8,8	8,8	17,905	
H2	8,9	8,9	17,870	19,171	H8	8,10	8,10	17,850	20,366
A3	8,10	8,10	17,844	20,366	A9	8,11	8,11	17,820	21,563
B3	9,10	9,10	19,063	20,320	B9	11,11	11,11	21,575	
C3	9,11	9,11	19,178	21,676	C9	8,8	8,8	17,841	
D3	7,10	7,10	16,626	20,370	D9	10,11	10,11	20,379	21,600
E3	9,11	9,11	19,146	21,665	E9	8,11	8,11	17,829	21,569
F3	8,10	8,10	17,839	20,357	F9	7,8	7,8	16,607	17,837
G3	8,8	8,8	17,853		G9	8,10	8,10	17,904	20,381
H3	6,9	6,9	15,313	19,070	H9	8,11	5,5	Gas-phase dimer	
A4	6,8	6,8	15,359	17,895	A10	9,10	9,10	19,076	20,342
B4	11,11	11,11	21,579		B10	8,11	8,11	17,827	21,569
C4	8,9	8,9	17,865	19,126	C10	8,12	8,12	17,855	22,850
D4	8,9	8,9	17,839	19,088	D10	8,9	8,9	17,846	19,093
E4	6,9	6,9	15,289	19,070	E10	9,10	9,10	19,086	20,346
F4	9,11	5,5	Gas-phase dimer		F10	9,12	9,12	19,065	22,801
G4	8,8	8,8	17,827		G10	11,12	11,12	21,554	22,795
H4	8,8	8,8	17,798		H10	8,11	8,11	17,841	21,584
A5	9,11	9,11	19,126	21,663	A11	6,9	6,9	15,299	19,076
B5	8,11	8,11	17,846	21,586	B11	8,9	8,9	17,851	19,095
C5	10,11	10,11	20,264	21,550	C11	9,11	9,11	19,120	21,592
D5	9,12	9,12	19,137	22,905	D11	8,8	8,8	17,829	
E5	9,10	9,10	19,076	20,337	E11	11,11	11,11	21,569	
F5	9,11	9,11	19,081	21,571	F11	11,11	11,11	21,548	
G5	8,8	8,8	17,831		G11	11,11	11,11	21,579	
H5	6,11	5,5	Gas-phase dimer		H11	8,9	8,9	17,870	19,095
A6	8,11	5,5	Gas-phase dimer		SUMMARY:	<u>Allele</u>	<u>Number</u>	<u>Avg. mass</u>	<u>Std. Dev.</u>
B6	No data	11,11	21,573			6	5	15,343	34.3
C6	8,8	8,8	17,844			7	3	16,586	53.7
D6	8,12	5,10	Gas-phase dimer			8	43	17,846	26.5
E6	8,12	8,12	17,831	22,816		9	31	19,102	33.1
F6	8,11	8,11	17,846	21,600		10	18	20,353	40.7
G6	9,11	9,11	19,070	21,565		11	24	21,590	37.9
H6	9,10	9,10	19,151	20,368		12	10	22,830	39.7

Exhibit 14. CDOJ TH01 STR results compared with ABI 310 and mass spectrometry methods. The shaded samples were run with a different primer set and have different masses.

Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)	Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)
A1	7,7	7,7	18,854		A7	7,8	7,8	18,877	20,104
B1	7,9.3	7,9.3	17,404	20,854	B7	8,9.3	8,9.3	20,128	22,337
C1	7,8	7,8	18,894	20,108	C7	8,9	8,9	20,193	21,415
D1	6,9	6,9	17,574	21,294	D7	6,9.3	6,9.3	16,167	20,876
E1	7,9.3	7,9.3	18,938	22,387	E7	9.3,9.3	9.3,9.3	22,349	
F1	9,9	9,9	21,415		F7	9,9.3	9,9.3	21,417	22,345
G1	7,9	7,9	18,919	21,408	G7	7,8	7,8	18,911	20,147
H1	5,7	5,7	16,403	18,926	H7	9,9.3	9,9.3	21,400	22,277
A2	6,10	6,10	16,145	21,170	A8	8,8	8,8	20,091	
B2	6,6	6,6	17,662		B8	7,9.3	7,9.3	18,938	22,391
C2	7,8	7,8	18,895	20,115	C8	7,9	7,9	17,436	19,957
D2	7,9	7,9	18,838	21,309	D8	7,7	7,7	18,944	
E2	7,9.3	7,9.3	18,904	22,329	E8	7,7	7,7	17,376	
F2	7,9	7,9	18,831	21,302	F8	7,9	7,9	17,387	19,891
G2	7,7	7,7	17,391		G8	9.3,9.3	9.3,9.3	22,263	
H2	9,9.3	9,9.3	19,871	20,685	H8	7,9.3	7,9.3	18,829	22,267
A3	7,7	7,7	18,856		A9	9,9	9,9	19,907	
B3	7,8	7,8	18,917	20,114	B9	7,7	7,7	18,938	
C3	7,9.3	7,9.3	17,380	20,828	C9	9.3,9.3	9.3,9.3	22,351	
D3	7,9	7,9	17,427	19,944	D9	7,8	7,8	18,845	20,097
E3	7,9.3	7,9.3	18,922	22,364	E9	8,9.3	8,9.3	18,684	20,790
F3	9,9.3	9,9.3	21,377	22,321	F9	6,9.3	6,9.3	17,622	22,287
G3	8,9.3	8,9.3	20,193	22,384	G9	7,9	7,9	17,380	19,885
H3	7,9.3	7,9.3	18,894	22,292	H9	6,7	6,7	17,677	18,926
A4	8,8	8,8	20,126		A10	7,8	7,8	17,389	18,652
B4	6,9.3	6,9.3	17,606	22,294	B10	9,9	9,9	21,377	
C4	6,7	6,7	17,586	18,838	C10	7,9	7,9	17,294	19,812
D4	7,9.3	7,9.3	18,897	22,312	D10	7,8	7,8	17,277	18,547
E4	7,9	7,9	17,401	19,917	E10	6,7	6,7	17,570	18,819
F4	7,9	7,9	18,836	21,334	F10	8,8	8,8	18,640	
G4	8,9.3	8,9.3	20,077	22,258	G10	6,7	6,7	17,644	18,906
H4	7,9.3	7,9.3	18,899	22,304	H10	8,9	8,9	18,643	19,902
A5	7,7	7,7	17,205		A11	6,7	6,7	17,540	18,782
B5	6,9	6,9	16,121	19,874	B11	6,8	6,8	17,682	20,191
C5	8,9.3	8,9.3	20,126	22,360	C11	7,7	7,7	18,927	
D5	6,8	6,8	17,679	20,184	D11	7,7	7,7	17,477	
E5	8,10	8,10	20,155	22,658	E11	6,6	6,6	17,596	
F5	7,9.3	7,9.3	18,870	22,285	F11	7,8	7,8	17,400	18,656
G5	6,7	6,7	17,624	18,906	G11	8,8	8,8	18,710	
H5	7,9	7,9	17,404	19,921	H11	6,7	6,7	17,600	18,895
A6	6,8	6,8	17,648	20,139					
B6	7,7	7,7	18,910						
C6	8,9.3	8,9.3	18,626	20,823					
D6	8,8	8,8	20,147						
E6	8,9.3	8,9.3	19,961	22,143					
F6	8,8	8,8	20,195						
G6	7,9	7,9	18,858	21,383					
H6	8,9	8,9	18,704	19,969					

SUMMARY:

All 88 samples were in agreement.

Exhibit 15. CDOJ amelogenin results compared with ABI 310 and mass spectrometry methods

Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)	Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)
A1	X,Y	X,Y	25,676	27,566	A7	X,Y	X,Y	25,774	27,687
B1	X,Y	X,Y	25,730	27,596	B7	X,Y	X,Y	25,659	27,555
C1	X,Y	X,Y	25,693	27,667	C7	X,Y	X,Y	25,797	27,695
D1	X,Y	X,Y	25,686	27,607	D7	X,Y	X,Y	25,641	27,533
E1	X,Y	X,Y	25,684	27,765	E7	X,Y	X,Y	25,674	27,583
F1	X,Y	X,Y	25,693	27,592	F7	X,Y	X,Y	25,676	27,585
G1	X,Y	X,Y	25,674	27,574	G7	X,Y	X,Y	25,668	27,568
H1	X,Y	X,Y	25,653	27,551	H7	X,Y	X,Y	25,666	27,564
A2	X,Y	X,Y	25,676	27,574	A8	X,Y	X,Y	25,659	27,555
B2	X,Y	X,Y	25,674	27,579	B8	X,Y	X,Y	25,653	27,544
C2	X,Y	X,Y	25,666	27,559	C8	X,Y	X,Y	25,668	27,566
D2	X,Y	X,Y	25,768	27,626	D8	X,Y	X,Y	25,674	27,566
E2	X,Y	X,Y	25,670	27,568	E8	X,Y	X,Y	25,843	27,763
F2	X,Y	X,Y	25,657	27,544	F8	X,Y	X,Y	25,649	27,581
G2	X,Y	X,Y	25,672	27,646	G8	X,Y	X,Y	25,661	27,561
H2	X,Y	X,Y	25,641	27,523	H8	X,Y	X,Y	25,757	27,585
A3	X,Y	X,Y	25,663	27,559	A9	X,Y	X,Y	25,641	27,531
B3	X,Y	X,Y	25,770	27,581	B9	X,Y	X,Y	25,684	27,579
C3	X,Y	X,Y	25,691	27,635	C9	X,X	X,X	25,645	
D3	X,Y	X,Y	25,722	27,594	D9	X,Y	X,Y	25,653	27,553
E3	X,Y	X,Y	25,749	27,628	E9	X,Y	X,Y	25,674	27,594
F3	X,Y	X,Y	25,682	27,577	F9	X,Y	X,Y	25,720	27,646
G3	X,Y	X,Y	25,645	27,536	G9	X,Y	X,Y	25,766	27,637
H3	X,Y	No data			H9	X,Y	No data		
A4	X,Y	X,Y	25,666	27,553	A10	X,Y	X,Y	25,697	27,618
B4	X,Y	X,Y	25,726	27,637	B10	X,Y	X,Y	25,684	27,590
C4	X,Y	X,Y	25,736	27,583	C10	X,Y	No data		
D4	X,Y	X,Y	25,782	27,670	D10	X,Y	X,Y	25,657	27,564
E4	X,Y	X,Y	25,701	27,789	E10	X,Y	X,Y	25,657	27,549
F4	X,Y	X,Y	25,674	27,585	F10	X,Y	X,Y	25,653	27,551
G4	X,Y	X,Y	25,599	27,572	G10	X,Y	No data		
H4	X,Y	X,Y	25,682	27,579	H10	X,Y	X,Y	25,722	27,654
A5	X,Y	X,Y	25,816	27,719	A11	X,Y	X,Y	25,676	27,572
B5	X,Y	X,Y	25,672	27,570	B11	X,Y	X,Y	25,655	27,549
C5	X,Y	X,Y	25,680	27,583	C11	X,Y	X,Y	25,666	27,577
D5	X,Y	X,Y	25,688	27,661	D11	X,Y	X,Y	25,618	27,464
E5	X,Y	X,Y	25,643	27,523	E11	X,Y	X,Y	25,697	27,676
F5	X,Y	X,Y	25,666	27,572	F11	X,Y	X,Y	25,653	27,557
G5	X,Y	X,Y	25,778	27,672	G11	X,Y	X,Y	25,661	27,553
H5	X,Y	X,Y	25,738	27,568	H11	X,Y	X,Y	25,643	27,540
A6	X,Y	X,Y	25,695	27,587					
B6	X,Y	X,Y	25,649	27,523					
C6	X,Y	X,Y	25,655	27,551					
D6	X,Y	No data							
E6	X,Y	X,Y	25,720	27,605					
F6	X,Y	X,Y	25,674	27,568					
G6	X,Y	X,Y	25,881	27,745					
H6	X,Y	No data							

SUMMARY:

	<u>X allele</u>	<u>Y allele</u>
Count	82	81
Average mass	25,690 Da	27,594 Da
Std. dev.	50.4 Da	59.6 Da

Eighty-two samples were in agreement.
"No data" resulted from mass spectrometry in six samples.

Exhibit 16. CDOJ D3S1358 STR results compared with ABI 310 and mass spectrometry methods

Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)	Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)
A1	17,17	17,17	28,182		A7	16,17	16,17	27,027	28,112
B1	16,16	16,16	26,967		B7	14,19	14,19	24,305	30,503
C1	15,17	15,17	25,668	28,139	C7	14,14	14,14	24,254	
D1	12,17	12,17	21,820	27,991	D7	15,16	15,16	25,524	26,713
E1	15,15	15,15	25,741		E7	16,16	16,16	26,822	
F1	16,16	16,16	26,762		F7	18,18	18,18	29,491	
G1	14,15	14,15	24,342	25,501	G7	15,16	15,16	25,772	26,830
H1	15,16	15,16	25,802	26,941	H7	16,17	16,17	27,018	28,273
A2	15,17	15,17	25,638	28,134	A8	14,16	14,16	24,331	26,856
B2	16,17	16,17	26,795	27,889	B8	16,17	16,17	26,843	28,056
C2	15,16	15,16	25,661	26,869	C8	15,16	15,16	25,638	26,766
D2	13,16	13,16	23,171	27,091	D8	15,15	15,15	25,724	
E2	14,17	14,17	24,352	28,126	E8	14,15	14,15	24,343	25,477
F2	15,17	15,17	25,549	28,091	F8	15,16	15,16	25,662	26,703
G2	15,17	15,17	25,641	28,119	G8	16,16	16,16	26,907	
H2	14,15	14,15	24,311	25,507	H8	15,15	15,15	25,618	
A3	17,17	17,17	28,184		A9	14,16	14,16	24,323	26,832
B3	18,18	18,18	29,384		B9	15,16	15,16	25,537	26,739
C3	15,15	15,15	25,645		C9	15,15	15,15	25,499	
D3	15,15	15,15	25,645		D9	15,15	15,15	25,522	
E3	16,17	16,17	26,967	27,945	E9	14,15	14,15	24,440	25,459
F3	15,16	15,16	25,573	26,790	F9	16,17	16,17	26,745	28,021
G3	17,18	17,18	28,025	29,302	G9	17,18	17,18	28,372	29,315
H3	14,15	14,15	24,360	25,530	H9	15,16	15,16	25,761	26,777
A4	17,17	17,17	28,001		A10	16,16	16,16	26,913	
B4	15,16	15,16	25,518	26,811	B10	15,18	15,18	25,539	29,347
C4	15,16	15,16	25,705	26,722	C10	15,16	15,16	25,477	26,741
D4	15,17	15,17	25,589	28,110	D10	14,15	14,15	24,214	25,341
E4	16,17	16,17	27,035	28,023	E10	14,15	14,15	24,252	25,507
F4	14,16	14,16	24,374	26,886	F10	16,17	16,17	26,871	28,139
G4	14,17	14,17	24,342	28,115	G10	16,17	16,17	27,033	27,934
H4	15,18	15,18	25,655	29,378	H10	15,16	15,16	25,716	26,715
A5	15,16	15,16	25,727	27,110	A11	15,18	15,18	25,545	29,293
B5	16,17	16,17	26,988	28,300	B11	15,15.2	15,15.2	25,734	26,379
C5	15,15	15,15	25,684		C11	16,17	16,17	26,937	27,845
D5	14,16	14,16	24,520	26,941	D11	15,16	15,16	25,544	26,672
E5	15,17	15,17	25,678	28,132	E11	16,16	16,16	26,899	
F5	16,16	16,16	26,813		F11	15,16	15,16	25,706	26,678
G5	9,17	9,17	18,092	28,285	G11	16,16	16,16	26,841	
H5	17,17	17,17	27,893		H11	16,17	16,17	26,881	28,246
A6	16,17	16,17	26,999	28,178					
B6	15,17	15,17	25,645	28,182					
C6	15,16	15,16	25,681	26,710					
D6	15,15	15,15	25,833						
E6	14,16	14,16	24,301	26,816					
F6	14,14	14,14	24,437						
G6	14,15	14,15	24,425	25,524					
H6	15,15	15,15	25,523						

SUMMARY:

All 88 samples were in agreement.

ABI 310 CE electropherogram of boxed data is shown in exhibit 74.

Exhibit 17. CDOJ D16S539 STR results compared with ABI 310 and mass spectrometry methods

Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)	Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)
A1	11,12	11,12	26,862	28,095	A7	11,11	11,11	26,873	
B1	9,14	9,14	24,421	30,665	B7	9,10	9,10	24,338	25,595
C1	9,11	9,11	24,360	26,888	C7	11,13	11,13	26,894	29,471
D1	10,11	10,11	25,622	26,831	D7	11,11	11,11	26,879	
E1	11,13	11,13	26,903	29,367	E7	9,12	9,12	24,425	28,200
F1	10,11	10,11	25,543	26,830	F7	9,12	9,12	24,368	28,160
G1	11,12	11,12	26,852	28,070	G7	11,12	11,12	26,997	28,145
H1	8,9	8,9	23,157	24,360	H7	12,12	12,12	28,104	
A2	9,11	9,11	24,342	26,805	A8	9,13	9,13	24,360	29,362
B2	12,13	12,13	28,191	29,323	B8	9,9	9,9	24,394	
C2	11,11	11,11	26,907		C8	9,13	9,13	24,352	29,369
D2	13,13	13,13	29,445		D8	11,11	11,11	26,933	
E2	12,12	12,12	28,128		E8	8,12	8,12	23,332	28,486
F2	12,13	12,13	28,202	29,436	F8	11,12	11,12	26,937	28,128
G2	10,11	10,11	25,638	26,869	G8	9,13	9,13	24,336	29,351
H2	10,11	10,11	25,576	26,864	H8	11,14	11,14	26,864	30,573
A3	9,12	9,12	24,340	28,106	A9	11,13	11,13	26,869	29,378
B3	9,12	9,12	24,356	28,117	B9	11,12	11,12	26,888	28,078
C3	12,13	12,13	28,141	29,483	C9	12,12	12,12	28,106	
D3	9,9	9,9	24,417		D9	11,13	11,13	26,849	29,342
E3	11,13	11,13	27,014	29,603	E9	11,13	11,13	26,899	29,396
F3	11,11	11,11	26,899		F9	12,15	12,15	28,115	31,845
G3	9,11	9,11	24,376	26,931	G9	11,13	11,13	26,886	29,389
H3	8,10	8,10	23,161	25,649	H9	11,13	11,13	26,888	29,387
A4	9,11	9,11	24,427	26,935	A10	9,11	9,11	24,423	26,920
B4	9,12	9,12	24,480	28,248	B10	9,12	9,12	24,449	28,126
C4	13,13	13,13	29,318		C10	12,13	12,13	28,302	29,498
D4	11,12	No data			D10	8,9	8,9	23,072	24,317
E4	13,13	13,13	29,389		E10	9,13	9,13	24,376	29,402
F4	9,11	9,11	24,374	26,894	F10	11,13	11,13	26,856	29,436
G4	13,14	No data			G10	9,9	9,9	24,358	
H4	11,13	11,13	26,837	29,336	H10	10,13	10,13	25,653	29,441
A5	12,12	12,12	28,117		A11	9,12	9,12	24,427	28,128
B5	8,9	8,9	23,181	24,417	B11	11,12	11,12	26,862	28,043
C5	10,11	10,11	25,670	27,007	C11	11,12	No data		
D5	12,12	12,12	28,265		D11	11,11	11,11	26,890	
E5	9,11	9,11	24,482	26,984	E11	11,13	11,13	26,928	29,619
F5	9,11	9,11	24,329	26,854	F11	9,10	No data		
G5	10,12	10,12	25,688	28,222	G11	10,12	No data		
H5	11,12	11,12	26,918	28,041	H11	11,11	11,11	26,894	
A6	9,11	9,11	24,372	26,888	<p>SUMMARY: Eighty-one samples were in agreement. "No data" resulted from mass spectrometry in five samples. "No data" resulted from ABI 310 in one sample. Gas-phase trimer caused an error in one sample.</p>				
B6	No data	9,12	24,405	28,208					
C6	9,13	9,13	24,380	29,398					
D6	9,13	9,13	24,465	29,416					
E6	12,12	12,12	28,170						
F6	13,13	13,13	29,329						
G6	11,13	11,13	26,867	29,371					
H6	11,11	9,9	Gas-phase trimer						

Exhibit 18. CDOJ D7S820 STR results compared with ABI 310 and mass spectrometry methods

Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)	Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)
A1	11,13	11,13	20,322	22,807	A7	10,12	10,12	19,097	21,621
B1	10,11	10,11	19,081	20,333	B7	11,12	11,12	20,545	21,592
C1	8,11	8,11	16,641	20,379	C7	8,10	10,10	19,360	
D1	9,10	9,10	17,857	19,131	D7	11,12	11,12	20,383	21,646
E1	8,10	10,10	19,072		E7	9,12	9,12	17,841	21,583
F1	10,12	10,12	19,088	21,594	F7	8,9	8,8	16,631	
G1	8,10	8,10	16,643	19,174	G7	11,12	11,12	20,374	21,575
H1	8,10	8,10	16,616	19,088	H7	11,12	11,12	20,366	21,581
A2	8,11	8,11	16,646	20,448	A8	10,12	10,12	19,189	21,703
B2	10,11	10,11	19,169	20,381	B8	10,11	10,11	19,265	20,400
C2	8,8	8,8	16,534		C8	10,10	10,10	19,095	
D2	10,10	10,10	19,189		D8	9,10	9,10	17,848	19,093
E2	8,8	8,8	16,641		E8	8,10	10,10	19,169	
F2	8,8	8,8	16,616		F8	8,9	No data	<i>Double null alleles?</i>	
G2	8,10	10,10	19,101		G8	9,12	9,12	17,949	21,745
H2	8,9	8,9	16,550	17,834	H8	8,10	10,10	19,187	
A3	11,12	11,12	20,372	21,569	A9	8,11	11,11	20,571	
B3	10,11	10,11	19,104	20,366	B9	8,8	8,8	16,534	
C3	8,12	8,12	16,532	21,548	C9	10,10	10,10	19,169	
D3	12,13	12,13	21,560	22,789	D9	7,8	7,8	15,383	16,638
E3	10,12	10,12	19,187	21,695	E9	9,10	9,10	17,844	19,090
F3	10,11	10,11	19,198	20,381	F9	10,12	10,12	19,395	21,866
G3	8,11	11,11	20,374		G9	8,10	8,10	16,651	19,201
H3	8,10	10,10	19,097		H9	8,11	8,11	16,826	20,363
A4	9,9	9,9	17,848	17,848	A10	12,13	12,13	21,586	22,820
B4	11,11	11,11	20,528		B10	11,12	11,12	20,387	21,586
C4	8,12	8,12	16,639	21,667	C10	8,10	10,10	19,330	
D4	8,9	8,9	16,636	17,860	D10	8,11	8,11	16,554	20,342
E4	9,11	11,11	20,392		E10	10,11	10,11	19,086	20,348
F4	10,11	10,11	19,185	20,396	F10	9,10	9,10	17,851	19,097
G4	10,10	10,10	19,101		G10	9,10	10,10	19,398	
H4	10,12	12,12	21,579		H10	10,11	10,11	19,382	20,377
A5	9,10	9,10	17,836	19,083	A11	10,10	10,10	19,254	
B5	11,12	11,12	20,631	21,678	B11	10,11	10,11	19,093	20,353
C5	11,12	11,12	20,363	21,573	C11	9,10	9,10	17,886	19,187
D5	10,10	10,10	19,088		D11	8,11	8,11	16,643	20,519
E5	9,11	9,11	17,851	20,372	E11	10,10	10,10	19,252	
F5	8,11	8,11	16,634	20,377	F11	9,12	9,12	17,831	21,569
G5	11,12	11,12	20,590	21,787	G11	10,12	10,12	19,142	21,653
H5	10,10	10,10	19,322		H11	10,12	10,12	19,160	21,883
A6	8,10	8,10	16,556	19,074					
B6	8,9	8,8	16,643						
C6	8,11	11,11	20,374						
D6	8,10	8,10	16,614	19,090					
E6	8,9	9,9	17,851						
F6	8,8	8,8	16,831						
G6	8,10	10,10	19,155						
H6	9,9	9,9	18,169						

SUMMARY:
 Seventy samples were in agreement.
 "No data" resulted from mass spectrometry (null alleles?) in one sample.
 Seventeen samples disagreed because of null alleles.

Exhibit 19. CDOJ FGA STR results compared with ABI 310 and mass spectrometry methods

Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)	Position	ABI 310	Mass Spec	Allele 1 (Da)	Allele 2 (Da)
A1	20,21	No data			A7	23,24	24,24	40,518	Not resolved
B1	20,23	20,23	35,987	39,547	B7	22,25	22,25	38,290	41,811
C1	23,25	23,25	39,390	41,917	C7	24,25	24,24	41,079	Not resolved
D1	19,23	19,23	34,858	39,547	D7	22,22	21,21	37,930	
E1	23,24	23,23	40,317	Not resolved	E7	24,27	24,27	40,680	44,130
F1	18.2,26	18.2,26	34,066	43,221	F7	22,22	22,22	38,295	
G1	21,21	21,21	37,011		G7	22,23	22,22	38,582	Not resolved
H1	No data	23,23	39,555		H7	21,26	21,26	37,148	43,116
A2	22,23	23,23	39,446	Not resolved	A8	OL?,24	16,24	31,415	40,688
B2	21,22	21,21	37,570	Not resolved	B8	22,24	22,24	38,226	40,874
C2	21,25	21,25	37,146	41944	C8	19.2,25	19.2,25	35,193	41,718
D2	19,19.2	19,19	34,699	Not resolved	D8	24,25	24,24	41,184	Not resolved
E2	23,27	23,26	39,310	43958	E8	18.2,24	18,23	33,960	40,247
F2	19,25	19,25	34,667	41575	F8	23,25	23,25	39,482	41,914
G2	22,22	22,22	38,554		G8	25,25	24,24	41,668	
H2	22,25	22,25	38,244	41,861	H8	21,23	21,23	37,113	39632
A3	22,24	22,24	38,110	40,604	A9	19,25	19,25	34,631	41,827
B3	30.2, OL?	30,30	48,114		B9	20,24	20,24	35,888	40,688
C3	21,24	21,24	37,158	40,777	C9	20,24	19,23	35,667	40,312
D3	23,23	23,23	39,475		D9	21,24	21,24	36,973	40,573
E3	19,23	19,23	34,769	39,547	E9	21,25	21,24	36,921	41,549
F3	23,24	23,23	39,852	Not resolved	F9	25,26	25,25	42,218	Not resolved
G3	22,23	22,22	38,582	Not resolved	G9	20,21	21,21	36,901	Not resolved
H3	18.2,23	18.2,23	34,138	39,423	H9	18.2,25	18.2,25	34,111	41,774
A4	22,23	22,22	39,056	Not resolved	A10	24,28	24,28	40,706	45,467
B4	19,26	19,26	34,578	42761	B10	21,24	21,24	37,018	40,636
C4	24,24	24,24	40,649		C10	18,23	18,23	33,560	39,490
D4	26,27	No data			D10	22,23	22,22	38,537	Not resolved
E4	22,22	22,22	38,442		E10	18,23	18,23	33,292	39,289
F4	22,23	22,22	39,081	Not resolved	F10	19,22	19,22	34,829	38,402
G4	23,24	23,23	39,741	Not resolved	G10	22,24	22,24	38,216	40,581
H4	23,23	23,23	39,562		H10	20,26	20,26	36,127	43,407
A5	20,22	20,22	35,856	38,364	A11	24,26	24,26	40,659	43,043
B5	23,23	22,22	39,017		B11	23,28	23,28	39,493	45,376
C5	22,23	22,22	38,659	Not resolved	C11	24,25	24,24	40,869	Not resolved
D5	23,26	22,25	39,233	42,879	D11	22,25	22,25	38,361	41,989
E5	22,23	22,22	38,554	Not resolved	E11	25,26	25,25	42,333	Not resolved
F5	19,25	19,25	34,795	41,994	F11	22,24	22,24	38,209	40,560
G5	21,24	21,24	37,366	40,853	G11	23,25	23,25	39,511	41,909
H5	21,22	21,21	37,565	Not resolved	H11	23,26	23,26	39,408	42,965
A6	22,24	22,24	38,155	40,570	SUMMARY:				
B6	24,27	24,27	40,612	44,288	Fifty-three samples were in agreement.				
C6	21,25	21,25	37,136	41,997	"No data" resulted from mass spectrometry in two samples.				
D6	20,23	20,23	35,972	39,477	"No data" resulted from ABI 310 in one sample.				
E6	22,23	22,22	38,644	Not resolved	One sample was too large for detection.				
F6	25,28	25,27	41,737	45,299	Nine samples resulted in wrong calls from poor calibration.				
G6	18.2,24	18.2,24	34,387	40,869	Twenty-two samples resulted in wrong calls from poor resolution.				
H6	21,23	21,23	37,113	39,534					