National Institute of Justice
Contract Number #2006-DN-BX-K011

Project Title: "Analysis of DNA Forensic Markers Using High Throughput Mass Spectrometry"

PI: Steven A. Hofstadler, Ph.D.
Ibis Biosciences Inc.
A subsidiary of Isis Pharmaceuticals, Inc.
1896 Rutherford Road
Carlsbad, CA 92008
Telephone: (760) 603-2599
Fax: (760) 603-4653
e-mail: shofstad@ibisbio.com

co-PI: Thomas A. Hall, Ph.D

Administrative POC:
LeeAnne Patton, Contracts Administrator
Address: Same as above
Telephone: (760) 603-3521
Fax: (760) 603-3893
e-mail: lpatton@ibisbio.com

Financial POC: Manolita Villavert
Address: Same as above
Telephone: (760) 603-3801
Fax: (760) 603-4653
e-mail: mvillavert@isisph.com

Abstract

This project focused on design, implementation, and validation of a next generation DNA forensics platform based on high throughput electrospray ionization mass spectrometry (ESI-MS). The approach is based on using ESI-MS to "weigh" DNA forensic markers with enough accuracy to yield an unambiguous base composition (i.e. the number of A's, G's, C's and T's in an amplicon) which in turn can be used to derive a DNA profile for an individual. Importantly, these base composition profiles can be referenced to existing forensics databases derived from mtDNA sequence or STR profiles.

Because the mass spectrometer is measuring base composition, not just length, mass spectrometric analysis shows that the standard STR markers often exhibit SNPs within the repeat region which can distinguish same length alleles. In one study using DNA from three distinct population groups, we noted that polymorphisms were found in 10 of the 13 core CODIS loci and a total of 364 polymorphic alleles were detected in 1330 genotype assignments (95 samples typed at 14 loci). These studies showed that alleles in four of the core 13 loci contain SNPs more than 25% of the time.

Throughout this project hundreds of STR profiles and thousands of mtDNA profiles were generated to support assay development and validation studies including sensitivity, specificity, stability, and mixture analysis.

This grant facilitated close collaboration with forensics scientists from several laboratories including the FBI DNA Unit II, AFDIL, and NIST. These collaborators provided numerous DNA test samples and blinded panels, critical feedback on assay design and performance, and expert opinion. The project was divided into four specific aims focused on 1) assay development and validation, 2) software development, 3) automation and testing, and 4) assay kit production and QC.

Table of Contents **Page**

Executive Summary

In this project we focused on design, implementation, and validation of a next generation DNA forensics platform based on high throughput electrospray ionization mass spectrometry (ESI-MS). Although there are a number of methods by which amplified nucleic acids can be interrogated, including sequencing, probe-based hybridization, hybridization to DNA microarrays, melting profiles, and various incantations of slab or capillary-based electrophoresis, it was long overlooked that a significant amount of information can be derived simply from the exact mass of a fragment of DNA.

The approach is based on using ESI-MS to "weigh" DNA forensic markers with enough accuracy to yield an unambiguous base composition (i.e. the number of A's, G's, C's and T's in an amplicon) which in turn can be used to derive a DNA profile for an individual. Importantly, these base composition profiles can be referenced to existing forensics databases derived from mtDNA sequence or STR profiles. Mass spectrometry is remarkably sensitive and can measure the mass and determine the base composition from small quantities of nucleic acids in a complex mixture with a throughput exceeding one electrospray event per minute. The ability to detect and determine the base composition of a large number of amplicons in a mixed sample enables analysis and identification of products in a short amount of time.

We have developed a 24-primer multiplexed assay that amplifies human mitochondrial DNA (mtDNA) hypervariable region 1 (HV1) coordinates 15924 to 16428 (primers span 15893 to 16451) and HV2 coordinates 31 to 493 (primers span 5 to 603). The assay consists of eight tri-plexed PCR reactions that occupy one column of a 96-well plate. The assay produces a profile of base compositions anchored on revised Cambridge Reference Sequence (rCRS) coordinates. By anchoring on standardized coordinates, any base composition profile can be directly compared to any other profile produced in this assay, or to any mtDNA sequence spanning the full range of amplified coordinates for each base composition to be compared (the base composition that would be

produced by any sequence spanning one of our primer pairs' amplified coordinates can be predicted accurately).

This approach facilitates the analysis of samples containing length heteroplasmy in the HV1/HV2 regions without degradation of information; in fact the method captures the extent and type of heteroplasmy in situations that are not amenable to sequencing. Additionally, because the analytical methods employed are highly automated, the technique is very high throughput. During the course of this project many thousands of mtDNA profiles were generated as part of the validation, sensitivity, and specificity studies.

With respect to STR markers, because base compositions are used to derive specific alleles, the MS-based method picks up SNPs within STR regions that go undetected by conventional electrophoretic analyses. For example, all "allele type 13" for the D5S818 marker are not equivalent; some contain a G to T SNP which distinguish them from individuals containing the "normal" allele type 13. Similarly, individuals which are typed as homozygous for this allele may in fact be heterozygotes containing alleles 13 and $13_{G/T\_SNP}$.

Unlike measurement of product mobility in a gel, therefore, the measurement of PCR product masses does not require an allelic ladder to correctly assign a product to the allele it represents. Using accurate mass measurements, we can determine when an allele has a polymorphism within the amplified region relative to the reference allele because the polymorphism changes the base composition of the PCR product.

While the concept of exploiting mass spectrometry-derived base compositions as forensics markers pre-dates this NIJ funded project, this grant provided critical resources for the redesign of the STR assay and a mechanism to thoroughly test and validate both the mtDNA and STR assays. At the beginning of this project the mtDNA assay was considerably further along the development pathway then the STR assay so the majority of resources allocated to the mtDNA assay were used for validation and

software enhancement, while the majority of resources allocated to the STR assay were used for assay development and testing. This grant facilitated our close collaboration with forensics scientists from several laboratories including the FBI DNA Unit II, AFDIL, and NIST. Our collaborators provided us with numerous DNA test samples and blinded panels, critical feedback on assay design and performance, and expert opinion as to how (if) this new platform would fit into the workflow of an accredited (and busy) forensics laboratory.

The project was divided into four specific aims, each designed to develop, validate, or automate a different aspect of the assay. These specific aims touched on various facets of development and/or performance validation of the automated mass spectrometry platform, the design and production of the assay kits, and the software used to interpret and report the results. Below we outline the specific aims of this project and offer results highlights from a few sections. Significantly more detail is found in the body of this report.

Specific Aim 1. Development and validation.

The development portion of this task focused on re-designing the STR assay to include all the CODIS 13 loci (plus amelogenin) into a multiplexed assay. While the initial approach had as many as six STR loci/reaction, it soon became apparent that the additional mass spectral complexity associated with a 6-plex reaction put significantly challenging demands on the data processing software owing to the frequent overlap of peaks. The final assay layout derived during this project, includes all CODIS 13 markers plus amelogenin, and consists of five tri-plexed reactions utilizing 11 loci with each tri-plex containing at least one primer pair that is redundant with another triplex. D21S11, FGA and D18S51 each occupy a single reaction.

The performance of both the STR and mtDNA assays were tested as they relate to the presence of non-human DNA. Neither assay was adversely affected by the presence of 10 ng/reaction (STRs) or 1-2 ng (mtDNA) of feline or canine DNA. Sensitivity of both assays was evaluated by dilution to extinction studies with quantified DNA templates,

the STR assay generated full profiles down to 111 pg/reaction while the mtDNA assay generated full profiles down to 6.25 pg/reaction. Reproducibility, precision, and accuracy were evaluated over hundreds of replicates of the STR assay and thousands of replicates for the mtDNA assay. For example, for the mtDNA reproducibility studies, a positive control was analyzed repeatedly over a 1.5 year period (August 2006 to December 2007) with 2,207 replicates. In this set, 2,130 (97%) of the trials yielded a full, correct profile and 73 (3%) were missing one or more primer pair (generally due to a plugged injector or marginal cleanup) and produced a correct, albeit partial, profile. In all, there were 159 missed calls out of 52,968 expected (99.70% call rate with no erroneous calls, 0.30% no call rate).

Through a NIST collaboration we were able to analyze DNA from diverse population groups and evaluate the prevalence of SNPs in STR markers. In one set of 95 samples from NIST, polymorphisms were found in 10 of the 13 core CODIS loci. A total of 364 polymorphic alleles were detected in 1,330 genotype assignments (95 samples typed at 14 loci). Thirteen distinct variant alleles were identified in D21S11, 13 in vWA, 11 in D3S1358, eight in D5S818, seven in D8S1179, six in D13S317, four in D7S820, three in D18S51, two in FGA and one in D16S539. Surprisingly, sixty percent of alleles observed in the 95 samples for D3S1358 were polymorphic relative to the reference allele.

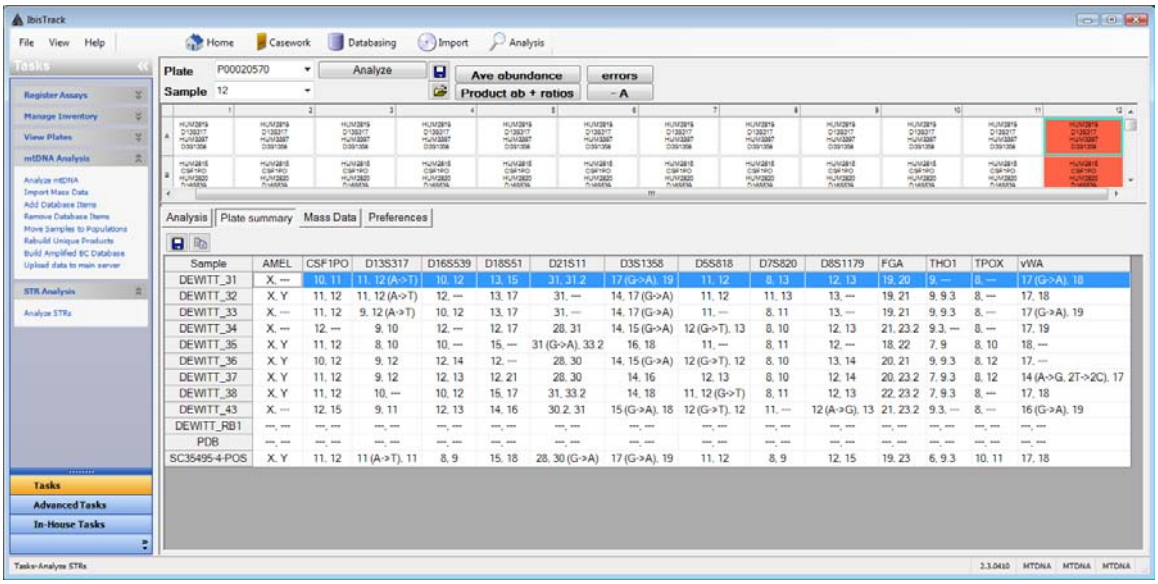
The table below summarizes the frequency of polymorphic (i.e. SNP-containing) alleles from a combined NIST/AFDIL/FBI sample panel. Although this sample size may be too small for meaningful statistics, these data provide tantalizing evidence that there may be a bias in certain populations for specific polymorphisms. For example, 75% of African American individuals from the NIST samples contained one or two polymorphisms, with 39.1% showing 2 G→A SNPs relative to the reference allele compared with 4.8% of Caucasians and 9.4% of Hispanics. See report text for details.

Locus	Number of Alleles with SNPs	Samples tested	% of alleles with one or more SNPs	Number of Same-Length Heterozygous Loci	% of samples heterozygous with same-length alleles
D3S1358	115	95	60.5	10	10.5
vWA	81	142	28.5	11	7.7
D13S317	77	142	27.1	7	4.9
D21S11	50	95	26.3	6	6.3
D5S818	59	142	20.8	10	7.0
D8S1179	35	142	12.3	7	4.9
D7S820	30	142	10.6	2	1.4
D18S51	7	95	3.7	0	0.0
FGA	2	95	1.1	0	0.0
D16S539	1	142	0.4	0	0.0

Specific aim 2. Development of a transferrable software interface

We have developed an interface for analysis of the Ibis mtDNA tiling assay and the Ibis STR assay that operates directly as a module within the IbisTrack sample / assay tracking package. All basic elements of assay setup, tracking, data processing and analysis are controlled through the IbisTrack interface. Assay "kit" templates are configured and stored in a relational database (Oracle) such that sample names simply need to be added to a pre-defined kit layout to register a 96-well assay plate to be run through the Ibis T5000 system. The running and processing of a mitochondrial tiling kit is fully end-to-end automated. Mass spectrometry and data processing steps for the STR assay are also automated, but there is currently a manual step to trigger data analysis and import processed data into our database.

The screen shot on the following page shows the GUI of the STR software and the data derived from nine samples, one extraction blank, one negative control and one positive control run an STR assay plate. Note that proper assignment of polymorphic alleles is part of the automated assignment process.

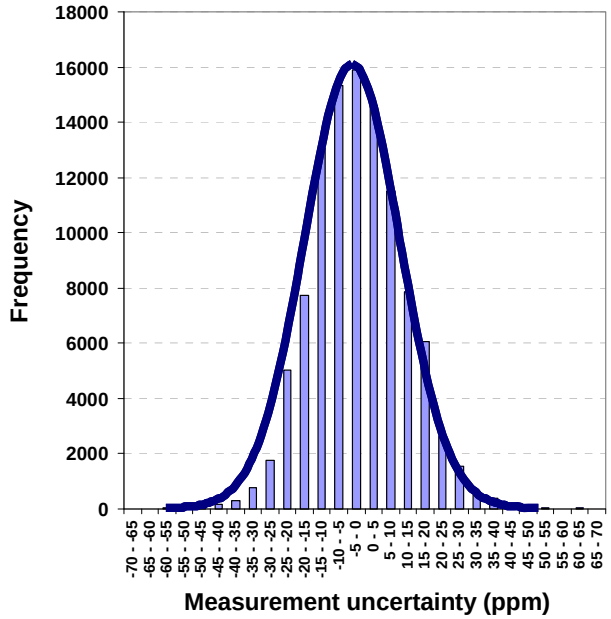


Specific Aim 3. Implementation of ESI-MS forensics methods on automated benchtop platform.

The hardware platform used throughout these studies is an automated mass spectrometry-based platform referred to as the Ibis T5000. The Ibis T5000 takes PCR generated amplicons generated off-line, performs a magnetic bead-based desalting/purification step, then injects purified aliquots of PCR amplicon into an automated electrospray ionization time-of-flight mass spectrometer (ESI-TOF). Mass spectrometry is used to determine the base composition (the number of adenosines (A), cytidines (C), guanosines (G) and thymidines (T)) of the PCR amplicon. Based on the maturity of the T5000 platform and the infrastructure Ibis has built to manufacture and support the instrument, we opted to run all studies on the standard T5000 platform. A number of software modifications, specific to the forensics assays, were effected during the course of this grant under NIJ funding. The DNA II Unit of the FBI laboratory in Quantico (through funding independent of this grant) acquired the T5000 platform during the period of performance of this grant. This afforded the opportunity for us to collaborate with FBI researchers at a significantly more detailed level than originally anticipated and allowed us to transfer SOPs, reagents, QC standards, and software to an outside laboratory for validation. These frequent scientific exchanges with the FBI allowed the Ibis forensics team to customize software functionality, report formats, and

standard QC protocols to be compatible with the established workflow of an accredited forensics laboratory.

To test system performance and establish automation robustness, 37 QC plates were run between Jan 2007 and Jan 2008 which comprised 3,552 PCR wells that were desalted, injected, and analyzed by ESI-MS on the T5000 instrument. Of these 3,552 analyses, 3 wells failed due to problems with the magnetic bead desalting module and 3 wells failed due to problems with the autosampler/injector. These data suggests a 99.83% automation success**.**



Mass measurement error and precision were determined through 105,580 independent mass measurements derived from 2,207 trials of a standard sample run through the mtDNA tiling assay. As shown above, the average (absolute value) of the mass measurement deviation for 105,580 measurements was 10.64 $\pm$ 8.2 ppm. These data support the hypothesis that the internal calibration scheme employed is robust over long periods of time and negates potential adverse affects of changes in environmental parameters such as room temperature, line voltage, or relative humidity.

Specific Aim 4. Kitting of amplification reagents and cleanup reagents

During the course of this project we supplied the FBI with approximately 70 assay plates and magnetic bead purification kits, enough to perform 700 assays at ten assays/plate (with one positive and one negative control/plate). Some of these plates were used for training and/or "practice", while the majority of the kits were used to support internal validation efforts and collaborative validation studies with Ibis. Furthermore, during the course of the project we performed stability studies on mtDNA kits to determine whether there were any shelf-life issues. The plates were tested at 4-6 weeks intervals for the first 6 months and every 12 weeks thereafter. As shown below, each QC plate has passed (i.e. yielded a full mtDNA profile at 25 pg/reaction or less). In addition, one plate was stored at room temperature for 72 hours before being used for PCR. The system validation sample layout was used. The plate passed, giving a full correct profile down to 25 pg input DNA. Although a more detailed study is necessary, this indicates that plates accidently left out on the bench top will still provide the intended performance.

Storage @ -20C	Date tested	Pass/Fail
4 Days	10/10/2006	Pass
4 Days	10/10/2006	Pass
12 Days	10/18/2006	Pass
5 weeks	11/10/2006	Pass
9 weeks	12/4/2006	Pass
13 weeks	1/3/2007	Pass
13 weeks	1/3/2007	Pass
17 weeks	2/6/2007	Pass
22 weeks	3/9/2007	Pass
28 weeks	4/20/2007	Pass
40 weeks	7/13/2007	Pass
52 weeks	10/11/2007	Pass

Impact of NIJ Funding

Through NIJ funding, the ESI-MS based forensics platform was significantly advanced with respect to assay design, extent of validation, and status of software. NIJ funding enabled rigorous assay validation and stability studies and provided a mechanism by which Ibis scientists could collaborate closely with forensic scientists from AFDIL, NIST, and the FBI. The mass spectrometry-based approach demonstrated the ability to rapidly and accurately detect SNPs in the standard STR loci which may have significant implications for information content from partial profiles. This work demonstrated the ability to accurately generate both STR and mtDNA profiles in a high throughput, automated modality. We believe these studies, and the supporting data presented in this report will have significant implications on how DNA forensics will be practiced in the future.

Introduction

From Mass Spectrum to DNA Base Composition

Unlike nucleic acid probes or arrays, mass spectrometry does not require anticipation of products analyzed, but simply measures the masses of the nucleic acids present in the sample. Although there are a number of methods by which amplified nucleic acids can be interrogated, including sequencing, probe-based hybridization, hybridization to DNA microarrays, melting profiles, and various incantations of slab or capillary-based electrophoresis, it was long overlooked that a significant amount of information can be derived simply from the exact mass of a fragment of DNA. We know that DNA amplicons are comprised exclusively of the four DNA building blocks, adenine (A), guanine (G), cytosine (C), and thymine (T). We also know the exact atomic composition of each of these building blocks and thus the exact mass (i.e., molecular weight) of each building block. After accurately "weighing" a piece of DNA it is rather straightforward to calculate its base composition (the number of As, Gs, Cs, and Ts). This approach is commonly used in the field of mass spectrometry to calculate a molecular formula of a compound from an exact mass measurement and knowledge of the exact masses of the elements.

Perhaps a better way to illustrate the concept of base composition determination by mass spectrometry is by analogy to a scale and a collection of coins (Figure 1). If one was given a scale and a table which listed the exact masses of coins (penny, nickel, dime, and quarter) and was then

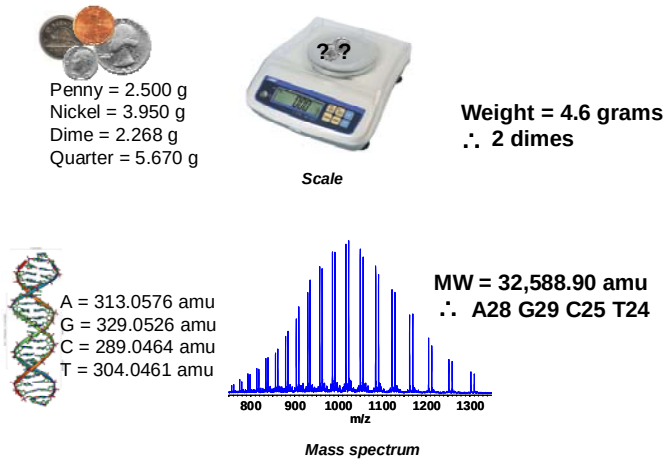


Figure 1. Base composition determination by mass spectrometry. Much like a scale can be used to discern the number and type of coins in a collection of loose change (top), Ibis technology uses mass spectrometry to accurately "weigh" amplified DNA to determine the base composition of amplified regions of chromosomes (bottom). These signatures are used to create base composition signatures which can be used for forensic differentiation of humans.

given an unknown coin and asked to determine what type of coin it was using only the scale, it would be rather straightforward. For example, a measured weight of 2.3 ± 0.1 grams would be consistent with a dime, which weighs 2.268 grams, but not a penny, nickel, or quarter. If the challenge was a few unknown coins it would still be a rather straightforward measurement and would require only simple mathematics (e.g., a measured mass of 4.6 ± 0.2 grams is consistent with two dimes but no other combination of coins). It would be significantly more challenging if one were given an unspecified number of coins which could number as many as 150. First, one would need a very accurate scale that would provide mass measurements to many decimal places with very high precision. Second, one would need considerably more computational resources than a scratch pad or a hand calculator.

This example is analogous to the way base compositions are calculated based on mass spectrometry. The mass spectrometer is our "molecular scale;" it provides extremely accurate measurements of the masses of the amplified DNA products (in the scale analogy above the mass measurement uncertainty was ± 4%, in the mass spectrometers used in the Ibis platform the mass measurement uncertainty is ± 0.001%). Because we know the exact masses of each DNA base, by employing sophisticated processing algorithms that take into account the masses of the building blocks, the base composition of the primers used, and the rules for Watson-Crick base pairing, it is straightforward to calculate an unambiguous base composition for amplicons up to about 150 bases.

Importantly, mass spectrometry is remarkably sensitive and can measure the mass and determine the base composition from small quantities of nucleic acids in a complex mixture with a throughput exceeding one electrospray event per minute. The ability to detect and determine the base composition of a large number of amplicons in a mixed sample enables analysis and identification of products in a short amount of time.

Previous Approaches to MS-based Forensic DNA Analysis.

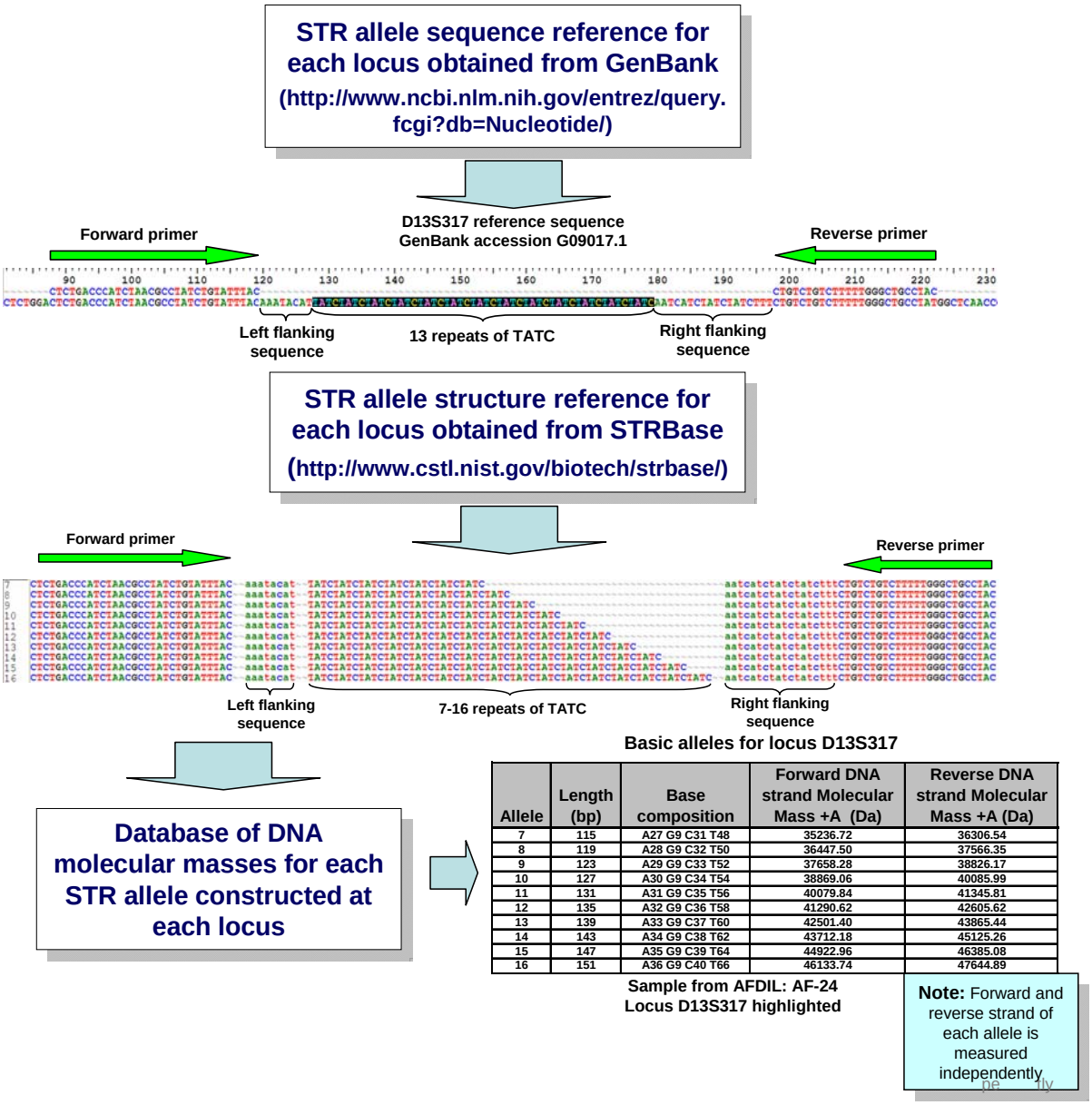
Matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI TOF MS) has been previously employed by others to analyze STR, SNP, and Y-chromosome markers. To obtain routinely the necessary mass accuracy and resolution using MALDI TOF MS, the amplicon size must be less than 100 bp, which often requires strategies such as enzymatic digestion and nested linear amplification. In the MALDI approach, PCR amplicons must be thoroughly desalted and co-crystallized with a suitable matrix prior to mass spectrometric analysis. The size reduction schemes and clean-up schemes employed for STR and SNP analyses in the cited reports resulted in the mass spectrometric analysis of only one strand of the PCR amplicon. By measuring the mass of only one strand of the amplicon, an unambiguous base composition cannot be determined and only the length of the allele is obtained. Even with the size reduction schemes, mass measurement errors of 12 to 60 Daltons (Da) are observed for products in the size range 15,000 to 25,000 Da. This corresponds to mass measurement errors of the 800 to 2,400 ppm. Because of poor mass accuracy and mass resolution typical of MALDI, multiplexing of STRs is difficult and not routine, although in one published report three STR loci were successfully multiplexed. The issue of allelic balance has not been addressed for MALDI-TOF-MS based assays.

The mass accuracy and resolution obtained with electrospray ionization (ESI) TOF MS is significantly improved relative to MALDI TOF MS. With amplicons in the 120-150 bp range, we have been able to obtain mass measurement errors of less than 20 ppm routinely with ESI. This allows multiplexing of STR loci and as shown below, we have successfully multiplexed 6 STR loci. We have also developed a highly automated PCR cleanup scheme that is compatible with ESI TOF. Unlike the MALDI TOF examples above, both strands of the amplicon are observed when using ESI TOF. Observation of both strands of the amplicon allows unambiguous base composition determination and confirmation of allele calls as well as the ability to determine SNPs present in alleles (see below). With automation, ESI TOF systems can analyze a well every 56 seconds. High throughput capacities of 1,536 wells/day can be obtained. Thus, ESI TOF shows great promise for the analysis of PCR amplicons.

STR assay

The principle elements of our STR assay are the measurement of PCR product masses via Electrospray-ionization time-of-flight mass spectrometry (ESI-TOF-MS), determination of product base compositions from their masses[1], and the association of the product base compositions to a database of alleles for each locus. The mass of a PCR product is an inherent property of the product that does not change according to assay conditions. Unlike measurement of product mobility in a gel, therefore, the measurement of PCR product masses does not require an allelic ladder to correctly assign a product to the allele it represents. The mass of a given allele generated with a specific primer pair is static and precise. We populate a database of all known alleles based upon a reference sequence and the published allele structures for each of the loci (obtained from STRBase[2]). The basic outline of generation and use of the database in this assay is outlined in Figure 2. Using accurate mass measurements, we can determine when an allele has a polymorphism within the amplified region relative to the reference allele because the polymorphism changes the base composition of the PCR product.

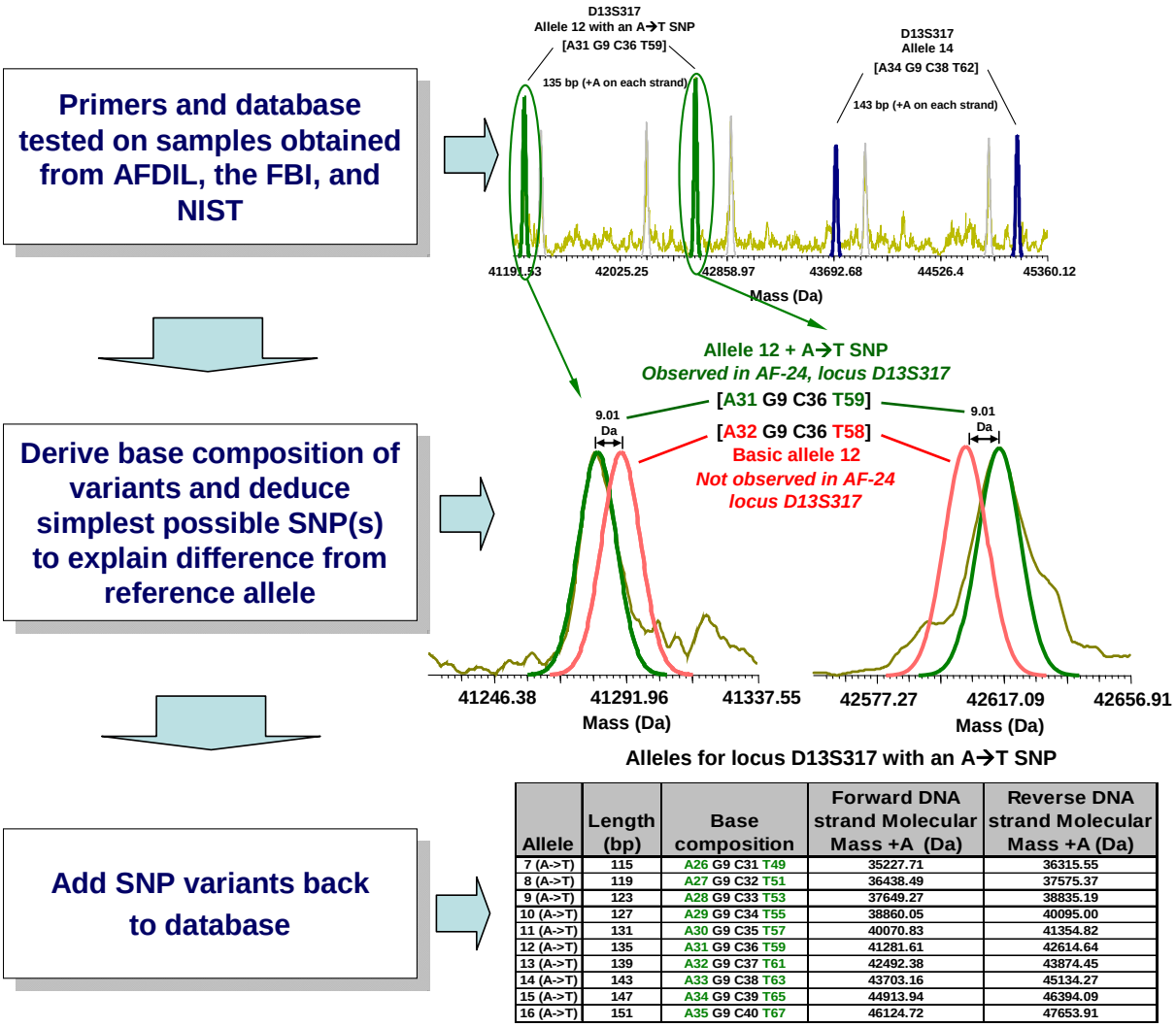
A.



Allele	Length (bp)	Base composition	Forward DNA strand Molecular Mass +A (Da)	Reverse DNA strand Molecular Mass +A (Da)
7	115	A27 G9 C31 T48	35236.72	36306.54
8	119	A28 G9 C32 T50	36447.50	37566.35
9	123	A29 G9 C33 T52	37658.28	38826.17
10	127	A30 G9 C34 T54	38869.06	40085.99
11	131	A31 G9 C35 T56	40079.84	41345.81
12	135	A32 G9 C36 T58	41290.62	42605.62
13	139	A33 G9 C37 T60	42501.40	43865.44
14	143	A34 G9 C38 T62	43712.18	45125.26
15	147	A35 G9 C39 T64	44922.96	46385.08
16	151	A36 G9 C40 T66	46133.74	47644.89

Figure 2. Panel A. The process of generating reference allele entries for an STR allele database is outlined above using D13S317 as an example.

B.



Allele	Length (bp)	Base composition	Forward DNA strand Molecular Mass +A (Da)	Reverse DNA strand Molecular Mass +A (Da)
7 (A->T)	115	A26 G9 C31 T49	35227.71	36315.55
8 (A->T)	119	A27 G9 C32 T51	36438.49	37575.37
9 (A->T)	123	A28 G9 C33 T53	37649.27	38835.19
10 (A->T)	127	A29 G9 C34 T55	38860.05	40095.00
11 (A->T)	131	A30 G9 C35 T57	40070.83	41354.82
12 (A->T)	135	A31 G9 C36 T59	41281.61	42614.64
13 (A->T)	139	A32 G9 C37 T61	42492.38	43874.45
14 (A->T)	143	A33 G9 C38 T63	43703.16	45134.27
15 (A->T)	147	A34 G9 C39 T65	44913.94	46394.09
16 (A->T)	151	A35 G9 C40 T67	46124.72	47653.91

Figure 2. Panel B. The use of an allele database in the absence of an allelic ladder. Correct allele assignments can be made by the direct measurement of product masses and the subsequent calculation of product base compositions. A sequence polymorphism in the allele relative to the reference allele results in shifted masses of both the forward and reverse strands. Calculation of the product base composition reveals the polymorphism(s). Polymorphic alleles can then be added back to the database. The location of the polymorphism remains unknown unless the allele is sequenced. Also, if two cancelling polymorphisms are present (e.g. and A→G SNP and a G→A SNP within the same amplicon), the ESI-TOF-MS assay will not register a polymorphism. This is expected to be quite rare in STR alleles, however.

Previously, we had been developing a proof-of-principle mass spectrometry-based STR assay consisting of two six-plex assays covering a total of 9 core CODIS loci plus the amelogenin marker with a 2-locus overlap between the two primer panels. The identity

and grouping of primer pairs incorporated in the original assay are shown below in Table 1.

Table 1. Primer pairs involved in original proof-of-principle STR multiplexed assays.

Six-plex group	pp num	Locus	Primer sequences
1	2822	TPOX	GGCACAGAACAGGCACTTAGGGA GGTGTCCTTGTCAGCGTTTATTTGCC
	2856	THO1	GGAAATCAAAGGGTATCTGGGCTCTGG CGCTGGTCACAGGGAACACAGAC
	2818	D8S1179	GGGGTTTTGTATTTCATGTGTACATTCGTATC GGGTACCTATCCTGTAGATTATTTTCACTGTGG
	2816	D5S818	GGGTGATTTTCCTCTTTGGTATCCTTATGTAAT GCCAATCATAGCCACAGTTTACAACATTTGTA
1&2	2823	vWA	GGGGAGAATAATCAGTATGTGACTTGGATTG GGGTGATAAATACATAGGATGGATGGATAGATGG
	2824	AMEL	GCCCTGGGCTCTGTAAAGAATAGTG GCATCAGAGCTTAAACTGGGAAGCTG
2	2815	CSF1PO	GGCATGAAGATATTAACAGTAACTGCCTTCATA GTGTCAGACCCTGTTCTAAGTACTTCCT
	2817	D7S820	GGGAACACTTGTCATAGTTTAGAACGAACTA CCCGGAATGTTTACTATAGACTATTTAGTGAGAT
	2819	D13S317	CTCTGACCCATCTAACGCCTATCTGTATTTAC GTAGGCAGCCCAAAAAGACAGACAG
	2820	D16S539	GCTCTTCCTCTTCCCTAGATCAATACAGACA GCTACCATCCATCTCTGTTTTGTCTTTCAATG

Using the two primer pair panels shown in Table 1, we STR-typed 25 blood spot samples provided by AFDIL and 22 buccal swab samples provided by the FBI. All samples were blinded. The 25 samples from AFDIL all had full STR truth data (generated at AFDIL with the Promega PowerPlex 16 system) that were shared with us after our results were communicated to AFDIL. Nine of the FBI samples had full STR typing results, two had 5 of the 10 loci analyzed previously typed, and ten of the samples had not been STR typed at the FBI. The ABI Cofiler and Profiler Plus systems were used by the FBI in their analyses of the samples.

Blood spot samples were received from AFDIL as three punches per sample on filter paper. Samples were purified with a Qiagen QIAamp mini kit spin protocol for dried blood spots and quantified using the ABI Quantifiler assay. Each sample was typed in duplicate using an estimated 1 ng of input DNA per reaction. Buccal swab samples were received from the FBI as single wooden-stick cotton or nylon swabs in separate plastic zip-lock bags. DNA was extracted from each swab using a Qiagen QIAamp mini

kit spin protocol for buccal swabs and quantified using the ABI Quantifiler assay. Each sample was typed in duplicate using an estimated 1 ng of input DNA per reaction.

PCR was performed in 96-well BioRad assay plates in 40 μl reactions containing 1X Qiagen multiplex buffer (Qiagen mat. No. 1022824), 200 nM each primer (grouped as shown in Table 1), and 1 ng of DNA per reaction. Thermocycling was performed in an MJ Research PTC-225 thermocycler and consisted of: 95 ºC, 10 min, [95 ºC, 30 sec., 61 ºC, 2 min., 72 ºC, 25 sec.] (35 cycles), 72 ºC, 4 min., 60 ºC, 30 min. (to promote full adenylation of PCR products). Desalting of PCR products was performed on an automated cleanup-up station using an in-house magnetic bead / anion-exchange protocol modified from Jiang, 2003[3]. ESI-TOF-MS was performed as described in Ecker et al., 2006[4]. The MassCollapse algorithm developed in collaboration with Science Applications International Corp. (SAIC) was used to deconvolve raw data into forward and reverse strand masses for each PCR product and these were associated to STR alleles using in-house-developed software. An example of the deconvolved mass spectrometry results for blinded sample AF-1 are shown in Figure 3. Forward and reverse strand masses for each PCR product are measured separately and both strands contribute to the assignment of a double-stranded allele.

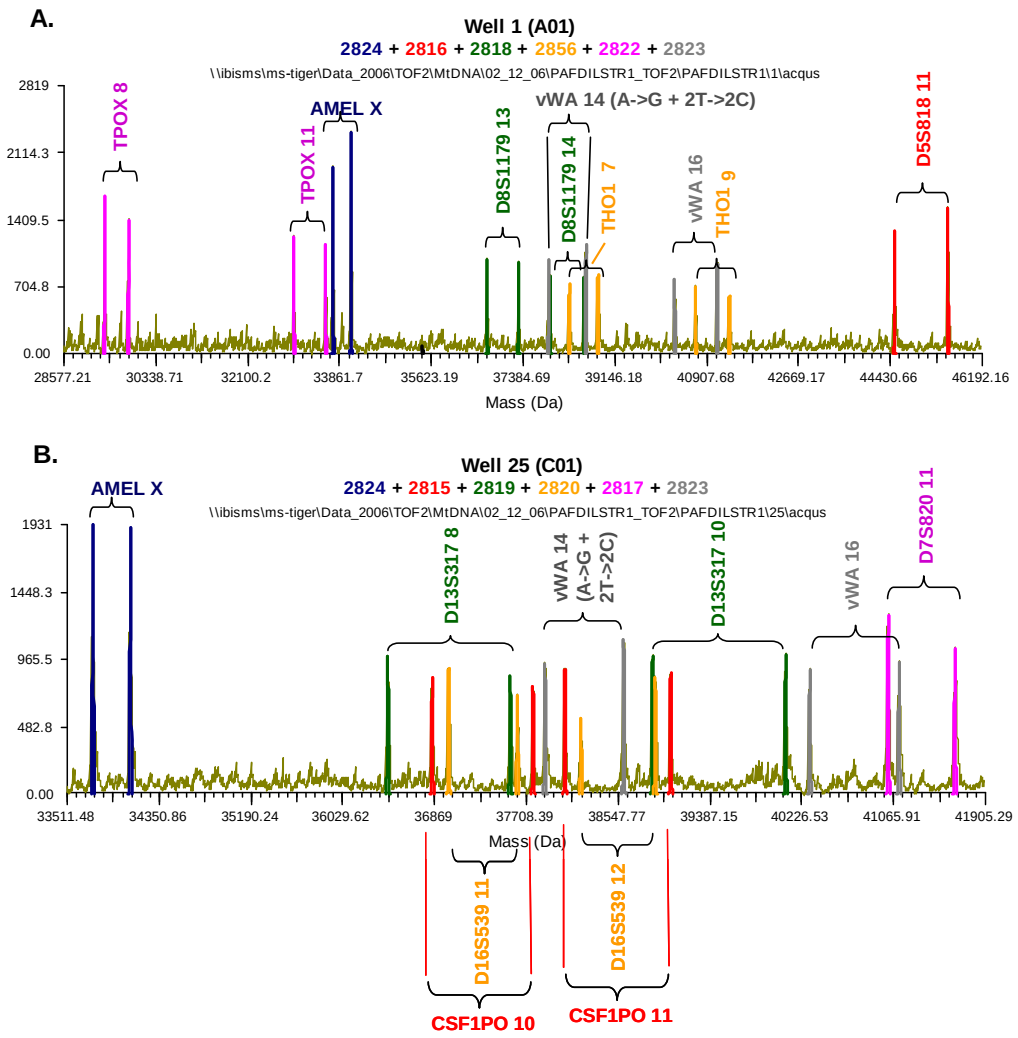


Figure 3. Deconvolved mass spectra representing 1 ng each of blinded sample AF-1 analyzed with the two 6-plex primer pair panels shown in Table 1. Mass peaks are color coded by Ibis primer pair number as shown in the line of 6 numbers near the top of each spectrum. The two product strands for each allele are bracketed and labeled by locus and allele assignment. Each spectrum was deconvolved over a mass range of 20 kDa to 55 kDa. The above figures are zoomed to the relevant region of each spectrum. A.) Multiplex panel 1. B.) Multiplex panel 2 (see Table 1).

Our STR typing results were consistent with conventional typing results at all loci for which we received truth data (Table 2). In addition, sequence-based polymorphisms relative to the reference allele were noted in five of the ten loci tested (Table 3). A total of five variant alleles were observed in D13S317, three variants were seen in D7S820, five variants were seen in D5S818, two variants were seen in D8S1179 and eight variants were seen in vWA. The observed frequency of occurrence of each variant relative to products matching reference alleles are shown in Table 4.

Table 2. Results of typing 25 blood spot samples from AFDIL ("AF-##") and 22 buccal swab samples from the FBI ("FBI-##"). The presence of polymorphisms relative to a reference allele is indicated by cell coloring and letter superscripts explained in the legend below the figure. The precise identity of each SNP was verified in separate, single-plex reactions utilizing ^{13}C-enriched dGTP in place of dGTP to alter the mass of each amplicon strand by ~10 Da per G residue (see Figure 4).

Sample	D5S818	D8S1179	THO1	TPOX	AMEL	vWA	CSF1PO	D13S317	D16S539	D7S820
AF-1	11, ---	13, 14	7, 9	8, 11	X, ---	14 [a], 16	10, 11	8, 10	11, 12	11, ---
AF-2	11, ---	12, 13	7, ---	8, 9	X, ---	18 [d], 19	9, 12	8, 11 [c]	12, 13	11, ---
AF-3	11, 12	12, 13 [d]	7, 9	8, 11	X, Y	16, ---	10, 11	9 [c], 11 [c]	12, ---	11, 12
AF-4	9 [b], 11 [b]	13, 14	7, ---	8, ---	X, ---	16, 17	10, 11	8, 11 [c]	10, 11	11, 9
AF-5	11, 13	13, 14	6, 9	8, ---	X, ---	17, 18	10, 13	11, 12	13, 9	8, ---
AF-6	10, 11	11, 13 [d]	7, 9.3	8, 9	X, ---	17, 18	10, 13	12, 13	11, ---	10, 12
AF-7	12 [b], 13	8, 10	6, 7	8, 11	X, ---	17, 19	9, 10	11, 11 [c]	11, 13	9, 11 [e]
AF-8	12 [b], ---	11, 13	6, ---	8, ---	X, ---	16, 18	10, ---	9, 11	11, 13	10, 12 [e]
AF-9	11, 12	10, 12	7, 9.3	8, 11	X, Y	14 [a], 17	11, 12	8, 9	11, 13	10, ---
AF-10	11, 12 [b]	10, 12	6, 9	9, 11	X, ---	17, 20	9, 12	12 [c], 14	8, 12	10, 12
AF-11	12, ---	11, 13	9.3, ---	8, ---	X, ---	14 [a], 18	11, ---	8, 13	11, 13	10, 12 [e]
AF-12	12, ---	14, 15	9, ---	8, 11	X, Y	15 [d], 16 [d]	11, ---	8, 12	12, ---	8, 9
AF-13	12, ---	10, 13	6, ---	8, ---	X, ---	15 [d], 17	12, ---	11, 12	12, ---	7, 12
AF-14	9 [b], 11 [b]	14, 14 [d]	9.3, ---	8, 11	X, ---	15 [d], 17	12, ---	11, 12	12, 13	8, 11
AF-15	11, 13	10, 14	6, 9	8, 11	X, ---	17, 19	10, 11	11, 13	11, 13	9, 11
AF-16	11 [b], 12 [b]	14, 15	7, 9	8, 10	X, ---	14 [a], 17	11, ---	11, 12 [c]	9, 11	10, 12
AF-17	13, ---	13 [d], 14	9, 9.3	8, ---	X, ---	18, 19	10, ---	12, 12 [c]	11, 12	8, 12
AF-18	11, 15	8, 13 [d]	6, 9	8, 11	X, Y	16, 18	9, 10	12 [c], 13	12, ---	10, 12
AF-19	10 [b], 11	8, 10	6, 7	8, 11	X, ---	17, 18	11, ---	12, ---	9, 12	8, 9
AF-20	12, 13	14, ---	6, ---	8, ---	X, ---	16, 18	10, 12	11 [c], 12	9, 11	11, 12
AF-21	9 [b], 12	13, 14	7, 9.3	8, 11	X, Y	17, 18	12, 13	11 [c], 12 [c]	8, 11	8, 10
AF-22	10 [b], 12	13, 14	6, ---	8, 11	X, ---	14 [a], 17	10, 11	9, 13 [c]	13, ---	10, ---
AF-23	12, 12 [b]	13, ---	6, 9.3	8, ---	X, ---	17, ---	12, ---	12, 13	9, 12	8, 11 [e]
AF-24	11, 13	13, ---	6, ---	9, 11	X, ---	16, 19	10, 11	12 [c], 14	12, 14	10, 12
AF-25	11, 12 [b]	11, 12	7, 8	8, ---	X, ---	17, ---	11, 13	11 [c], ---	11, ---	11, 11 [e]
FBI-3	11, 14	12, 14	6, 9.3	8, 11	X, ---	15 [d], 18	10, ---	8, 11 [c]	11, 13	9, ---
FBI-9	13, ---	13, 14 [d]	6, 7	8, 9	X, ---	17, ---	10, 11	11, 14	9, 12	11, ---
FBI-22	11, ---	10, 13	7, 9.3	8, 9	X, ---	18 [f], 19	10, 12	11 [c], 12 [c]	10, 12	10 [e], 12
FBI-28	9 [b], 11	12, 13	7, 9	8, 11	X, Y	19, ---	10, 11	12, 13	10, 11	10 [e], 12
FBI-32	12, ---	9, 10	8, 9.3	8, 10	X, ---	15 [d], 18	12, ---	9, 11 [c]	10, 12	10, 11
FBI-33	11, 12	14, ---	6, 7	8, ---	X, ---	18, ---	9, 11	14 [c], ---	12, ---	8, 12 [e]
FBI-37	11, 13	12, 14	5, 9	8, 10	X, ---	14 [a], 17	11, 12	11, 12	11, 13	8, 10 [e]
FBI-47	12 [b], 14	12, 13	6, 9.3	8, ---	X, ---	17, 19	11, 12	11, 12	10, 12	12 [e], ---
FBI-48	9 [b], 12	13, 16	9, 10	8, 11	X, ---	17, 19	10, 11	8, 12 [c]	9, 11	11, 13
FBI-49	12, ---	12, 15	6, ---	8, 9	X, ---	16 [d], 17	12, 13	12 [c], 14 [c]	9, 13	9, 12
FBI-51	11, 13 [1]	11, 15	6, ---	8, 11	X, Y	17, 19	11, ---	11, 14 [c]	11, 13	10, 12
FBI-57	11, 12	9, 14	6, 9.3	8, 9	X, ---	15 [d], 19	11, ---	8, 13	11, 13	11, 11 [e]
FBI-58	12, ---	11, 14	7, ---	8, 9	X, Y	17, ---	11, 12	10, 11 [c]	11, 12	10, 12
FBI-61	9 [b], 11	10, 12	6, 9.3	8, 11	X, ---	17 [d], 17	10, 12	9, 11 [c]	12, 13	8, ---
FBI-65	11 [b], 12	14, ---	8, 9	8, 9	X, ---	15, 18	12, ---	8, 12	11, 12	10, 12
FBI-66	11, 11 [b]	13, 14	8, 9.3	8, 11	X, ---	16, 19 [f]	10, ---	13, 14	12, ---	8, 12 [e]
FBI-67	11, 12	12, 13	9, ---	8, ---	X, ---	16, 16 [d]	11, 12	11 [c], 14	11, 12	10, 11
FBI-69	11, ---	12, 14	7, 9.3	8, 9	X, ---	15 [d], 16	12, 13	12, 13 [c]	12, 14	8, 9
FBI-70	11, ---	13, 14	6, ---	8, ---	X, Y	18, ---	10, 11	12, ---	9, 11	8, 10
FBI-72	11, 12	10, ---	9, 9.3	8, 11	X, ---	18 [f], 19	11, ---	11 [c], 12	10, 12	8, 11
FBI-75	12, 13 [b]	11, 12	9.3, ---	8, 11	X, ---	18, 18 [f]	11, 12	9, 14	9, 13	9, 11
FBI-82	12, 12 [b]	13, 14	7, 9.3	8, 11	X, ---	14 [g], 17	10, 11	12 [c], 13	11, 12	11, ---

[a] A→G + 2T→2C SNPs
[b] G→T SNP
[c] A→T SNP
[d] G→A SNP
[e] T→A SNP
[f] A→G SNP
[g] G→A + T→C SNP
[1] Allele 13 for D5S818, FBI-51 was missed in first analysis and found after reanalyzing data

Base allele calls confirmed with conventional STR typing

Base allele calls confirmed with conventional STR typing, one or both alleles contain SNP relative to reference

Conventional STR typing not run

Conventional STR typing not run, one or both alleles contain SNP relative to reference

As indicated in the description for Table 2, we employed single-plex reactions for each of the samples/loci for which a polymorphism was observed. Although the presence of a SNP variant can be identified without ambiguity in our assay, there are alternative variants that result in forward and reverse strand masses that are similar (for example, given an allele with a base composition of [A31 G15 C27 T50], a G→A SNP would give a composition of [A32 G14 C27 T50] and a T→C SNP would give a composition of [A31 G15 C28 T49]. The forward and reverse strands of a PCR product of [A32 G14 C27 T50] differ in mass from the forward and reverse strand masses for [A31 G15 C28 T49] by approximately 1 Da each. With sufficient mass accuracy and precision, which is normally achievable with our ESI-TOF-MS analysis, we can distinguish between these possibilities. However, because biochemical reactions contain noise and signals are not always perfect in distribution, ambiguity at this level is possible. The use of ^{13}C-enriched dGTP offsets the masses of the two strands by ~11 Da, rather than 1 Da, making SNP assignment unambiguous. Figure 4 demonstrates this concept graphically. All polymorphisms listed in Tables 2, 3 and 4 were verified using ^{13}C-enriched dGTP as a mass tag.

Table 3. Frequency of observing polymorphisms within STR alleles from 47 unrelated donor samples analyzed at 10 STR loci (see Table 2). Polymorphisms were found in five of the ten loci and at least one representative heterozygous individual was found at each of the five loci for which the locus is called homozygous by conventional typing (the two alleles have the same length but differ in base sequence structure as indicated by PCR product base composition).

Locus	Number of Alleles with SNPs	% of Total Allele Calls Containing a SNP*	Number of Same-Length Heterozygous Loci
D5S818	22	23.4	3
D8S1179	6	6.4	1
vWA	23	24.5	3
D13S317	30	31.9	2
D7S820	12	12.8	2

A.

D13S317		
Allele	**Count**	**Percentage**
8	10	10.6
9	6	6.4
9 (A->T)	1	1.1
10	2	2.1
11	11	11.7
11 (A->T)	15	16.0
12	18	19.1
12 (A->T)	10	10.6
13	9	9.6
13 (A->T)	2	2.1
14	6	6.4
14 (A->T)	4	4.3

B.

D7S820		
Allele	**Count**	**Percentage**
7	1	1.1
8	16	17.0
9	10	10.6
10	18	19.1
10 (T->A)	3	3.2
11	20	21.3
11 (T->A)	4	4.3
12	15	16.0
12 (T->A)	6	6.4
13	1	1.1

C.

D5S818		
Allele	**Count**	**Percentage**
9 (G->T)	6	6.4
10	1	1.1
10 (G->T)	2	2.1
11	30	31.9
11 (G->T)	5	5.3
12	26	27.7
12 (G->T)	9	9.6
13	11	11.7
13 (G->T)	1	1.1
14	2	2.1
15	1	1.1

D.

D8S1179		
Allele	**Count**	**Percentage**
8	3	3.2
9	2	2.1
10	11	11.7
11	7	7.4
12	14	14.9
13	22	23.4
13 (G->A)	4	4.3
14	24	25.5
14 (G->A)	2	2.1
15	4	4.3
16	1	1.1

E.

vWA		
Allele	**Count**	**Percentage**
14 (A->G + 2T->2C)	6	6.4
14 (G->A + T->C)	1	1.1
15	1	1.1
15 (G->A)	7	7.4
16	11	11.7
16 (G->A)	3	3.2
17	28	29.8
17 (G->A)	1	1.1
18	17	18.1
18 (G->A)	1	1.1
18 (A->G)	3	3.2
19	13	13.8
19 (A->G)	1	1.1
20	1	1.1

Table 4. Observed allele frequencies for each of five loci displaying polymorphisms relative to the reference allele. Each percentage is based upon a sample number of 94 (22 FBI samples and 25 AFDIL samples with 2 alleles at each locus).

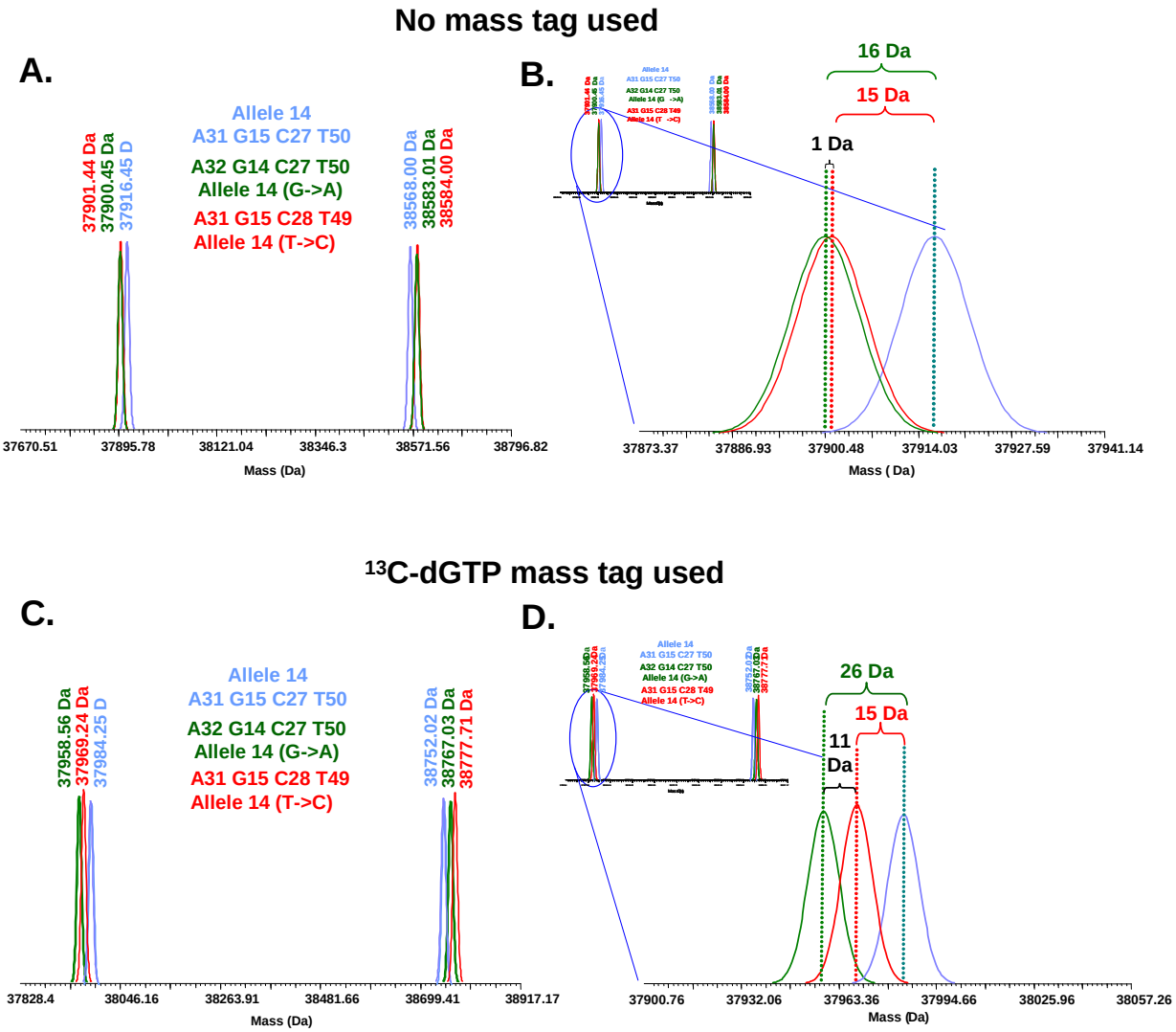


Figure 4. Use of a mass tag to make an unambiguous SNP assignment in a PCR amplicon. The example above shows locus D8S1179, allele 14, amplified with Ibis primer pair 2818. A.) When amplified with natural dNTPs, a G→A and a T→C variant from allele 14 produce amplicons that are very close in mass (about 1 Da difference in each of the product strands). B.) Zoomed-in view of the forward strand masses for each of the three PCR products. There is an unambiguous detection of a SNP from the base allele 14 product, but only a 1 Da difference between masses for the forward strands of the G→A and T→C products, making the precise identity of the SNP potentially ambiguous between two possibilities. C.) When amplified with ^{13}C-enriched dGTP in place of dGTP, a G→A and a T→C variant from allele 14 produce amplicons separated by nearly 11 Da, which allows unambiguous assignment of each SNP variant. D.) Zoomed-in view of the forward strand masses for each of the three PCR products amplified with ^{13}C-enriched dGTP. There is an unambiguous detection of a SNP from the base allele 14 product, and an unambiguous assignment of the base switch involved in the SNP. The basic allele 14 product is separated from the G→A SNP by ~26 Da and from the T→C SNP by ~15 Da. The two SNP variants are separated by ~11 Da.

Mitochondrial DNA control region tiling assay

We have developed a 24-primer multiplexed assay that amplifies human mitochondrial DNA (mtDNA) hypervariable region 1 (HV1) coordinates 15924 to 16428 (primers span 15893 to 16451) and HV2 coordinates 31 to 576 (primers span 5 to 603). The assay consists of eight tri-plexed PCR reactions that occupy one column of a 96-well plate (Figure 5). The basic concepts of the mechanics of this assay are outlined in panel B of Figure 5. Primer pairs within each tri-plex are spaced as far apart as possible and a very short PCR extension cycle (5 sec.) is used to minimize PCR products spanning any two primer pairs (Figure 5, panel A). The assay produces a profile of base compositions anchored on revised Cambridge Reference Sequence (rCRS) coordinates (Figure 6). By anchoring on standardized coordinates, any base composition profile can be directly compared to any other profile produced in this assay, or to any mtDNA sequence spanning the full range of amplified coordinates for each base composition to be compared (the base composition that would be produced by any sequence spanning one of our primer pairs' amplified coordinates can be predicted accurately). The primer pairs used in the Ibis mtDNA tiling assay and their working concentrations are listed in Table 5.

Based upon an analysis of 1,266 human mtDNA sequences from GenBank containing the full range of sequence covered by our tiling assay, the tiling assay retains the ability to differentiate 94.2% of the entries by one or more base composition differences (see Appendix A for a list of GenBank accession numbers for the sequences used in this comparison). The partial loss of information comes from reciprocal SNPs within a single amplified region for one of our primer pairs. If two sequences contain a C→T SNP and a T→C SNP relative to each other, and these SNPs are close enough to be contained within a single amplified region for one of our primer pairs, they will produce identical base compositions for that primer pair if they are the only differences between the sequences. The condition that all differences between two sequences are compensated by reciprocal SNPs is rare, and in some cases where it occurs, they are still split by overlapping primer pairs, so over the span of 24 primer pairs, most profile-to-profile differentiation power is retained. Although the assay retains 94.2% of

individually-discriminating information relative to sequencing, the assay also covers a larger coordinate space than the minimal HV1 / HV2 sequencing routinely performed (rCRS coordinates HV1 16024-16365 and HV2 73-340). For the same 1,266 sequences that are distinct in sequence over our tiling coordinates, only 90.2% are discriminated by sequencing of the minimal HV1 and HV2 coordinates.

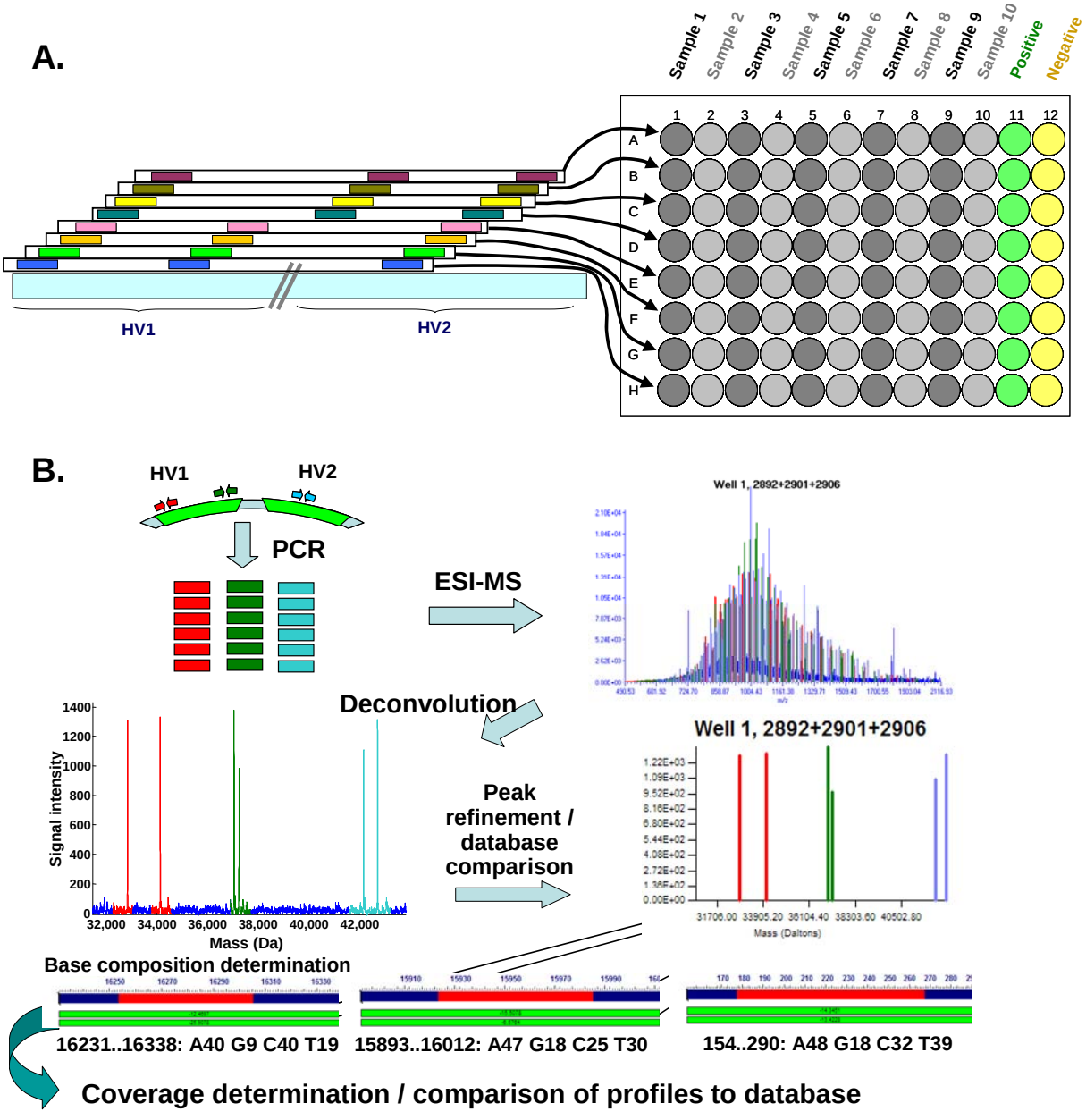


Figure 5. Overview of the mtDNA triplex tiling assay. A. Primer pair groupings by approximate mtDNA spatial locations of grouped primers (left side) and layout of a 12-sample-per-plate assay kit plate. B. Schematic overview outlining the analysis process for one reaction. Results of 8 reactions are automatically tied together to generate a base composition profile.

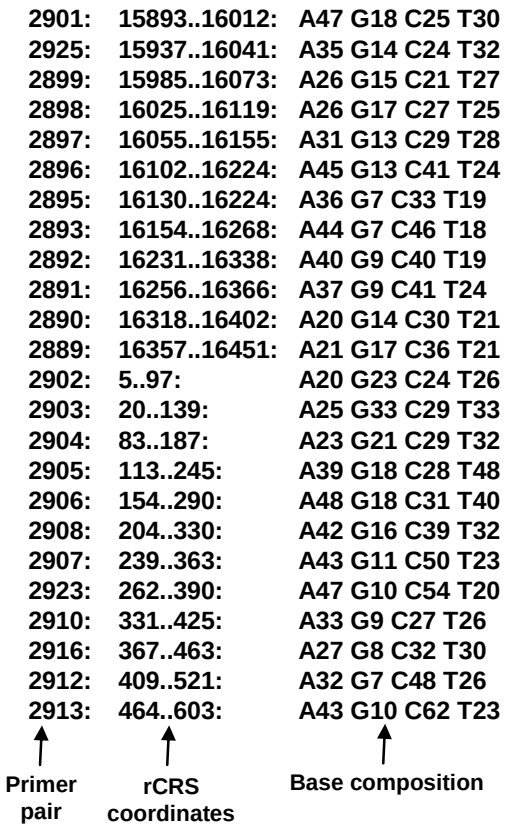


Figure 6. Format of a base composition profile produced with the Ibis mtDNA tiling assay. The base composition is anchored on rCRS coordinates to allow direct comparison of profiles to each other or to a database developed from sequence data, direct base composition measurement, or both.

A.

Triplex	Ibis primer pair number	rCRS start	rCRS end	Primer sequences
1	2901	15893	16012	TGGGGTATAAACTAATACACCAGTCTTGTAA TTAAATTAGAATCTTAGCTTTGGGTGC
	2906	154	290	TCCTTTATCGCACCTACGTTCAAT TGGTTGTTATGATGTCTGTGTGG
	2892	16231	16338	TCACACATCAACTGCAACTCCAA TGCTATGTACGGTAAATGGCTTTATGTACTATG
2	2925	15937	16041	TCCTTTTTCCAAGGACAAATCAGAGA TGCTTCCCCATGAAAGAACAGAGA
	2891	16256	16366	TCACCCCTCACCCACTAGGATACCAAC TGGGACGAGAAGGGATTTGACT
	2908	204	330	TGTGTTAATTAATTAATGCTTGTAGGACAT TCTGTGGCCAGAAGCGG
3	2899	15985	16073	TGCACCCAAAGCTAAGATTCTAATTTAAAC TGGTGAGTCAATACTTGGGTGG
	2890	16318	16402	TGCCATTTACCGTACATAGCACAT TGGTCAAGGGACCCCTATCTG
	2907	239	363	TAACAATTGAATGTCTGCACAGCC TGTTTTTGGGGTTTGGCAGAGAT
4	2898	16025	16119	TCTTTCATGGGGAAGCAGATTTG TCATGGTGGCTGGCAGTAATG
	2889	16357	16451	TCTCGTCCCCATGGATGACC TCGAGGAGAGTAGCACTCTTGTG
	2923	262	390	TGCTTTCCACACAGACATCATAACAAA TCTGGTTAGGCTGGTGTTAGGGT
5	2893	16154	16268	TAGTACATAAAAACCCAATCCACATCAA TGGTGAGGGGTGGCTTTG
	2910	331	425	TCTTAAACACATCTCTGCCAAACC TAAAAGTGCATACCGCCAAAAGAT
	2902	5	97	TCAGGTCTATCACCCTATTAACCACT TGTCTCGCAATGCTATCGCGT
6	2897	16055	16155	TCCAAGTATTGACTCACCCATCA TACAGGTGGTCAAGTATTTATGGTAC
	2903	20	139	TATTAACCACTCACGGGAGCT TTTCAAAGACAGATACTGCGACATA
	2916	367	463	TACCCTAACACCAGCCTAACCA TGGAGGGGAAAATAATGTGTTAGTTG
7	2896	16102	16224	TACTGCCAGCCACCATGAATAT TGGGTTGATTGCTGTACTTGCTT
	2913	464	603	TCTCCCATACTACTAATCTCATCAATACA TGCTTTGAGGAGGTAAGCTACATAAAC
	2904	83	187	TAGCATTGCGAGACGCTGGA TGCCTGTAATATTGAACGTAGGTGC
8	2905	113	245	TCTATGTCGCAGTATCTGTCTTTGA TGGGTTATTATTATGTCCTACAAGCATT
	2895	16130	16224	TTTCCATAAATACTTGACCACCTGTAG TGGGTTGATTGCTGTACTTGCTT
	2912	409	521	TGCGGTATGCACTTTTAACAGT TGTGTGTGCTGGGTAGGATG

B.

Primer pairs	Conc (nM)
2892, 2901, 2906	150, 280, 298
2891, 2908, 2925	150, 700, 350
2890, 2899, 2907	150, 220, 600
2889, 2898, 2923	150, 153, 700
2902, 2893, 2910	150, 450, 453
2897, 2916, 2903	125, 260, 600
2904, 2896, 2913	150, 167, 334
2905, 2895, 2912	570, 150, 150

Table 5. Primer pairs and working concentrations for the Ibis mtDNA tiling assay. A. Primer pairs grouped by triplex Triplexes 1-8 are distributed in rows A-H (in that order) in the assay plate. B. Working concentrations of triplexed primer pairs in the PCR reactions.

Strengths and limitations of the Ibis forensics platform

The Ibis platform provides several significant advantages over conventional assays in both mitochondrial and STR analyses. The mitochondrial assay is capable of resolving mixtures of two templates. When template inputs are asymmetric, the approach is capable of quantifying the relative contribution of each template. Moreover, the Ibis mtDNA assay is unhindered by length heteroplasmy, which causes significant analytical trouble during sequence analysis. For example, figure 7 shows the effect of HV1 poly-C length heteroplasmy on the readout of forward and reverse sequence traces through the poly-C region of HV1 (figure 7, A.). The sequence trace is destroyed by multiple, mixed

templates after reading through the poly-C region. In contrast, Ibis mtDNA analysis of the same sample reveals four clearly-resolved length variants for this region (figure 7, B). Because each DNA species is measured independently as a separate entity, length variants do not hinder performance of the assay.

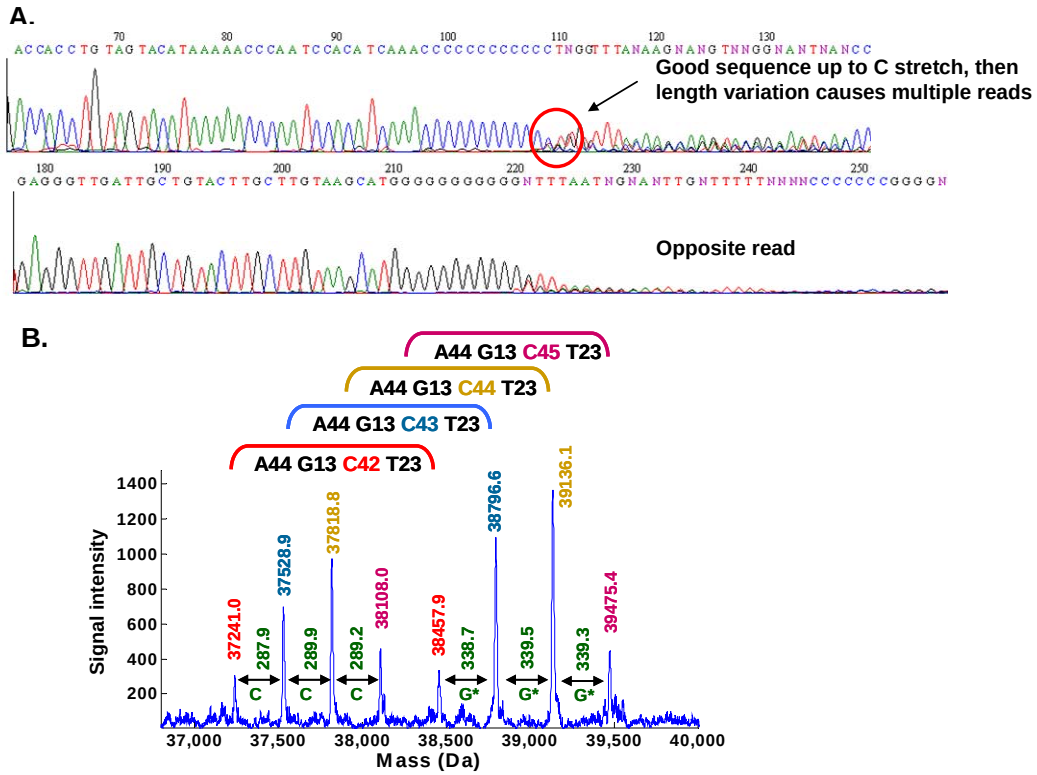


Figure 7. Effect of length heteroplasmy upon mitochondrial DNA sequencing and the Ibis mitochondrial tiling assay. Sequence analysis is severely hindered by mixed templates that differ in length. Panel A shows the forward and reverse strand sequence traces of sequencing reactions developed by direct PCR-product sequencing from a buccal-swab sample provided by the FBI (USA.FBI.000009). The sequence trace becomes unreadable after the poly-C region in either reading direction. Panel B shows the data obtained for the sample using primer pair 2896 (rCRS amplified coordinates 16124..16201). There are four clearly-resolved products that differ by single C-length increments.
* ^{13}C-enriched dGTP was used in these reactions, adding just under 10 Da to the mass of each G residue.

Despite significant advantages, there is a trade-off. Compensating SNPs within an amplified region (e.g. a C→T and T→C within the same amplicon) will be silent in the Ibis system. For example, the common variations at rCRS 150 and 152 may therefore be masked when both occur in unison for two samples being directly compared. It is not possible with the current mtDNA tiling system to interrogate specific phylogenetically-significant mtDNA positions. It is possible to develop assays on our

platform to do this, but that is beyond the purpose of our mitochondrial profiling assay. Also, there is the possibility of measuring DNA strand masses that overlap within a given reaction due to mixed templates or heteroplasmy. These can generally be resolved clearly by manual interrogation of the data with integrated data viewing tools.

The Ibis STR assay brings the advantage of resolving polymorphisms within the common STR alleles that are invisible to conventional typing methods such as capillary electrophoresis (CE). These polymorphisms are quite common in the majority of the core CODIS loci, as demonstrated in sections S.1.4 and S.1.5 of this report. This information can be of benefit when analyzing familial relationships, as the presence of SNPs within repeat structures can lend substance to tracing potential identity-by-decent scenarios in shared alleles between generations. Also, the resolution of polymorphisms expands the allele base for several loci, which may aid in resolution of samples when full profiles cannot be obtained. Moreover, the STR assay required no allelic ladder. DNA mass is an intrinsic property of the PCR product and requires no relative comparison to a standard run in parallel for each assay. Novel variant alleles can be resolved and characterized directly, the first time they are encountered.

The ESI-MS platform is currently limited by the upper size of an amplicon that can be analyzed with high precision. For the unambiguous determination of product base compositions, PCR products should be kept under 150 bp in length. This limit is imposed by three factors: 1.) The mass measurement deviations observed in mass spectrometry analyses are proportional to the mass of the analyte being measured, 2.) spectral congestion increases with product size, as the number of discrete charge states present is proportional to the mass of the analyte, and 3.) The number of possible nucleotide combinations that can add up close to a given mass increases as the mass increases. When potential base compositions are highly constrained (as with STR products), the upper size can be extended to 200 bp or greater because an unambiguous, *de novo* calculation of base composition is not required (rather, the masses can be compared directly to an allele database).

The ESI-MS platform is currently limited in capacity for multiplexing compared to CE-based methods. This is due primarily to product size and spectral congestion limitations. Although we have successfully multiplexed 6 STR loci, we are currently more comfortable with three loci within a single reaction, and the three larger loci (D21S11, FGA and D18S51) are currently single-plexed to limit spectral congestion. Because of the need to divide a sample across eight PCR reactions for our mitochondrial and STR assays, the assays may require more total template than equivalent commercial systems. This is not a concern for most analyses, but can be significant in analysis of trace quantities of DNA.

A current limitation for the Ibis STR assay is the noise baseline of the data and the overall dynamic range afforded by the assay. Forensic analysts may be accustomed to a very flat baseline in a CE genotyping trace, where two products may be resolved at a 100:1 ratio. The ESI-MS system has a ubiquitous, underlying baseline of chemical noise fluctuation and products must be clearly resolvable above that noise baseline before they are confidently assigned. Although we have demonstrated that the analytical platform is capable of dynamic range in excess of $100:1^{5}$, the dynamic range limits within our multiplexed STR and mitochondrial assays is closer to 10:1. This is due primarily to increased spectral congestion and complexity from an increase in product sizes and numbers. We are developing data processing and analysis software with continued increases in robustness and we are currently developing signal-to-noise and statistical weighting calculations to aid in the proper thresholding of mass assignments.

Despite certain limitations, we feel that the advantages provided by the ESI-MS platform for the resolution of mtDNA mixtures, the resolution of polymorphic STR alleles, STR analysis without allelic ladders, and the convenience and throughput afforded by full end-to-end automation greatly outweigh current weaknesses. This technology is at a relatively early stage and will only improve in the near future. The ESI-MS platform has the potential to unify and automate the most common DNA forensic analyses, autosomal STR analysis, mitochondrial DNA analysis and Y-STR analysis, on one common platform.

Specific Aim 1. Development and validation: STR assay

The following section of this report addresses development of the Ibis STR assay and subsections are prepended with an S. The mitochondrial profiling assay is discussed in a following section and subsections therein are prepended with an M.

S.1.0. Development of an updated STR assay and new assay layout

With six primer pairs per reaction, the propensity for signal confusion due to spectral congestion was high enough that we wished to simplify the assay to provide easier, more automatable data processing and analysis. In addition, the original assay was developed with a commercial buffer system (Qiagen multiplex assay buffer) to allow for easy assay development while testing the CODIS STR loci for the presence of sequence polymorphisms. Another goal going forward was to develop our assay into a buffer system that we can manufacture in-house unencumbered by additional licensing/cost issues. In addition to simplifying assay analysis (by decreasing mass spectral congestion), we sought to work in the additional four CODIS loci (D21S11, FGA, D18S51 and D3S1358).

Even after placing primer pairs extremely close to the repeat sequence for loci D21S11, FGA and D18S51, these loci produce amplicons with an upper size for the larger alleles that pushes the limits of ESI-TOF-MS (Table 6). The size limit for detection and resolution is due to several factors, but one key factor is spectral congestion caused by a large number charge states for the analyzed product (ESI produces multiply charged molecules and measures a distribution of charge states in a single spectrum). As amplicons increase in size, the number of distinct charge states detected in the mass spectrum increases and the spacing between them decreases. Each of the primer pairs for loci D21S11, FGA and D18S51 produce amplicons with an upper size limit near or greater than 200 bp. If an amplicon of this size is multiplexed with other primer pairs, closely spaced charge state signals will be likely to interfere with accurate measurement

of other products in the same spectrum. For this reason, these three primer pairs will be used in separate single-plex reactions in our new assay layout.

Table 6. Size minima and maxima for products of primer pairs targeting STR loci D21S11, FGA and D18S51. We have two alternative primer pairs for each locus. The primer pairs in current use are 3390, 3393 and 3394.

Locus	Primer pair	Shortest allele in DB	Min length	Longest allele in DB	Max length
D21S11	3390	24	150	38	206
D21S11	3391		150		206
FGA	3392	12.2	116	50.2	268
FGA	3393		110		262
D18S51	3394	7	117	27	197
D18S51	3395		124		204

We have defined an eight-well STR assay encompassing the 13 core CODIS loci and the sex marker amelogenin. The current assay layout consists of five tri-plexed reactions utilizing 11 loci with each tri-plex containing at least one primer pair that is redundant with another triplex. D21S11, FGA and D18S51 each occupy a single reaction. The 14 primer pairs in current use are shown in Table 7.

Table 7. Current primer pairs used for 14-locus STR assay. Each primer is initiated with a G or a C to promote efficient *Taq* polymerase-mediated adenylation of final PCR products[6].

PP Number	Locus	Primer sequences
2815	CSF1PO	GGCATGAAGATATTAACAGTAACTGCCTTCATA GTGTCAGACCCTGTTCTAAGTACTTCCT
2816	D5S818	GGGTGATTTTCCTCTTTGGTATCCTTATGTAAT GCCAATCATAGCCACAGTTTACAACATTTGTA
2817	D7S820	GGGAACACTTGTCATAGTTTAGAACGAACTA CCCGGAATGTTTACTATAGACTATTTAGTGAGAT
2818	D8S1179	GGGGTTTTGTATTTCATGTGTACATTCGTATC GGGTACCTATCCTGTAGATTATTTTCACTGTGG
2819	D13S317	CTCTGACCCATCTAACGCCTATCTGTATTTAC GTAGGCAGCCCAAAAAGACAGACAG
2820	D16S539	GCTCTTCCTCTTCCCTAGATCAATACAGACA GCTACCATCCATCTCTGTTTTGTCTTTCAATG
2822	TPOX	GGCACAGAACAGGCACTTAGGGA GGTGTCCTTGTCAGCGTTTATTTGCC
2823	vWA	GGGGAGAATAATCAGTATGTGACTTGGATTG GGGTGATAAATACATAGGATGGATGGATAGATGG
2824	AMEL	GCCCTGGGCTCTGTAAAGAATAGTG GCATCAGAGCTTAAACTGGGAAGCTG
2856	THO1	GGAAATCAAAGGGTATCTGGGCTCTGG CGCTGGTCACAGGGAACACAGAC
3390	D21S11	CCCCAAGTGAATTGCCTTCTA GGTAGATAGACTGGATAGATAGACGATAGA
3393	FGA	CCCAATTAGGCATATTTACAAGCTAGTT GTCTGTAATTGCCAGCAAAAAAGAAA
3394	D18S51	GATGTCTTACAATAACAGTTGCTACTATTTCT CTGAGTGACAAATTGAGACCTTGTC
3397	D3S1358	GAAATCAACAGAGGCTTGCATGTAT GACAGAGCAAGACCCTGTCTCAT

The new assay is compatible with a kit layout where each sample is distributed across one column of a 96-well assay plate, so that one plate can contain 12 samples typed at 14 loci each (Figure 8). The new assay was developed and tested with blood-derived samples N31774 and SC35495 originally obtained from Seracare Life Sciences (Oceanside, CA). The current assay has been developed and tested with target inputs of 1 ng template per reaction. The initial reaction formulation was 20 mM Tris-HCl, 100 mM KCl, 3 mM $MgCl_2$, 800 mM betaine, 100 μg/ml BSA (NEB), 200 μM each dNTP, 5 U / reaction Amplitaq Gold (ABI). Primer pairs are used at variable relative concentrations within tri-plex reactions to achieve acceptable inter-locus product balance, as shown in Figure 8, panel A. Prototype pre-fabricated assay kit plates were constructed manually by mixing appropriate primer pairs into buffer to produce 35 μl/well of 1.14X reaction mix to which 5 μl of DNA sample is added prior to thermocycling. Initial thermocycling involved 95 °C for 15 minutes, [95 °C for 30 seconds, 61°C for 2 minutes, 72 °C for 30 seconds] x 40 cycles, 72 °C for 10 minutes, 60 °C for 30 minutes (to promote adenylation), 4 °C hold. Ninety-five reference

samples from NIST were correctly typed using the initial reaction and cycling formulation. These conditions were used for studies outlined in sections S.1.4 and S.1.5. Since that time, the buffer formulation and cycling has been modified to 10 mM Tris-Cl, 75 mM KCl, 3 mM $MgCl_2$, 1 M betaine, 20 mM sorbitol, 200 μM each dNTP, 5 U / reaction Amplitaq Gold to achieve slightly better reproducibility and product balance. Current thermocycling involves 95 °C for 15 minutes, [95 °C for 30 seconds, 61°C for 2 minutes, 72 °C for 45 seconds] x 40 cycles, 72 °C for 4 minutes, 60 °C for 30 minutes, 4 °C hold. Studies outlined in sections S.1.1, S.1.2 and S.1.3 used the latter PCR conditions, which represent the current state of the assay.

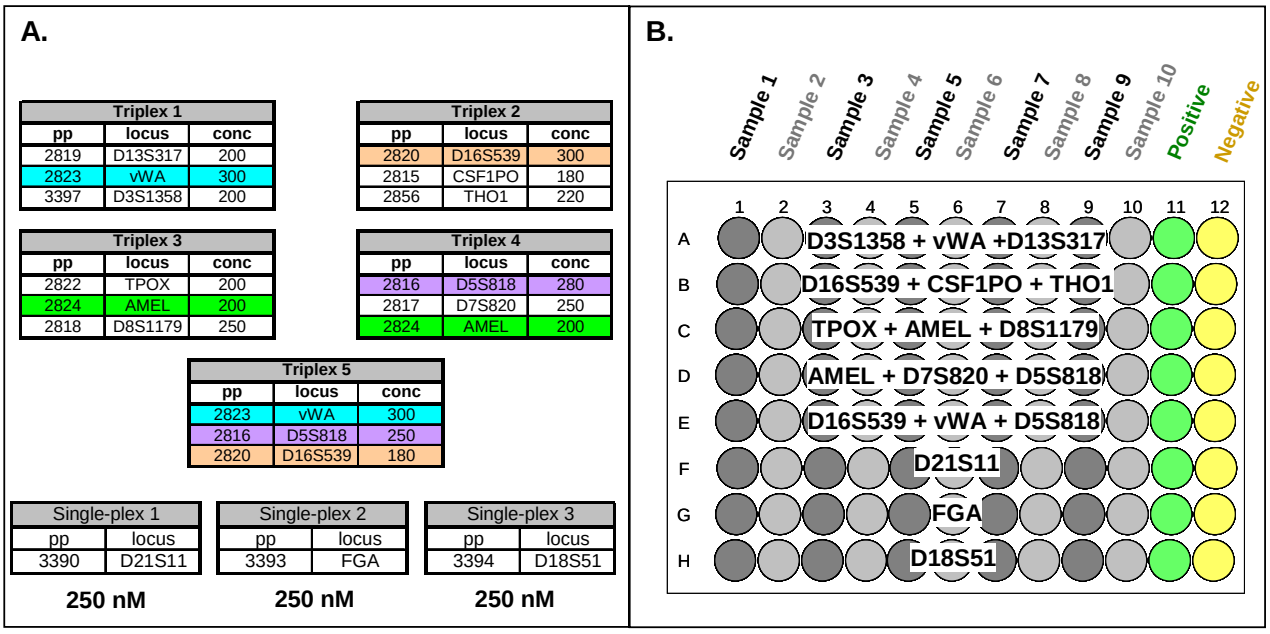


Figure 8. Assay layout for a 14-locus ESI-TOF-MS STR assay. Each sample is distributed across 8 wells in one column each of a 96-well assay plate. A standard layout could consist of 10 samples, one positive control and one negative control on an assay plate. A.) Primer pair groupings for each reaction set. There are eight reaction sets (five tri-plexes and three single-plexes). Primer pair concentrations in tri-plexed reactions are adjusted to achieve acceptable inter-locus product balance. Working concentrations of primers in each primer pair are shown in the right column of each table. All single-plex reactions are performed using 250 nM of each primer. B.) Proposed layout of a standardized assay plate for a 14-locus STR assay.

S.1.1. Species specificity, STR assay

To test the ability of the Ibis STR assay to detect a correct profile in the presence of an excess of non-human DNA, human blood-derived DNA was amplified in the presence of a 10-fold excess of DNA isolated from buccal swab preparations taken from a male dog (American Eskimo) and a male cat (common house cat mix). Five buccal swabs were taken from each animal specimen and DNA was purified using a Qiagen buccal swab spin protocol (Qiagen QIAmp DNA Mini Kit and QIAmp DNA Blood Mini Kit Handbook, 02/2003, pages 37-38). The swabs were prepared according to the Qiagen protocol with the exception that each set of five extractions were combined and processed over two QIAmp spin columns. The DNA from each column was eluted into 110 μL Qiagen EB buffer. DNA in each of the four preparations was quantified by UV at 260 nm using 1.5 μL of undiluted sample in a NanoDrop UV spectrophotometer. Human blood sample N31774 (from Seracare, Inc.) was previously quantified using PicoGreen (Molecular Probes).

To verify the presence of amplifiable canine and feline DNA in the buccal swab preparations used for specificity testing, primer pairs were created targeting the Y-chromosomal Sry gene Genbank accession U15160.1 (dog, *Canis familiaris*) and AB099654.1 (cat, *Felis catus*)[7-10]. Canine Sry primers used were forward: SRY-DOG_U15160_15_37_F: 5´-GAACGCATTCTTGGTGTGGTCTC-3´ and reverse: SRY-DOG_U15160_124_146_R: 5´-GGCCATTTTTCGGCTTCTGTAAG-3´. Feline Sry primers used were forward: SRY_AB099654_75_96_F: 5´-CTTGCGAACTTTGCACGGAGAG-3´ and reverse: SRY_AB099654_157_177_R: 5´-GCGTTCATGGGTCGTTTGACG-3´.

Figure 13, panel A shows the amplification of specific, predicted Sry PCR products using the canine and feline control primers, respectively, demonstrating the presence of amplifiable DNA in the feline and canine DNA preparations. Figure 13, panel B shows the control reactions with these primers thermocycled in the absence of added template. Figures 9 (canine) and 11 (feline) show the results of running the STR assay in the

presence of 10 ng/reaction of canine or feline DNA, respectively. A single clear product was produced by the amelogenin primer pair (Ibis primer pair 2824) that is included in two of the triplexes (subpanels 3 and 4 in each figure 9 and 11). The product was not consistent with human DNA and did not apparently interfere with correct typing results even when canine or feline DNA was present at a 10-fold excess over human DNA (Figures 10 and 12, respectively). Interestingly, the same product was produced with feline and canine DNA and was identical in eight replicates (two replicates for each cat and dog, and present in two independent reactions in each assay). This product was specific to the animal DNA and was not present in the water control or the human DNA alone control (run in duplicate, not shown). Full, correct profiles were obtained for human DNA for all replicates in the presence of a 10-fold excess of canine or feline DNA (Table 8).

10 ng canine DNA

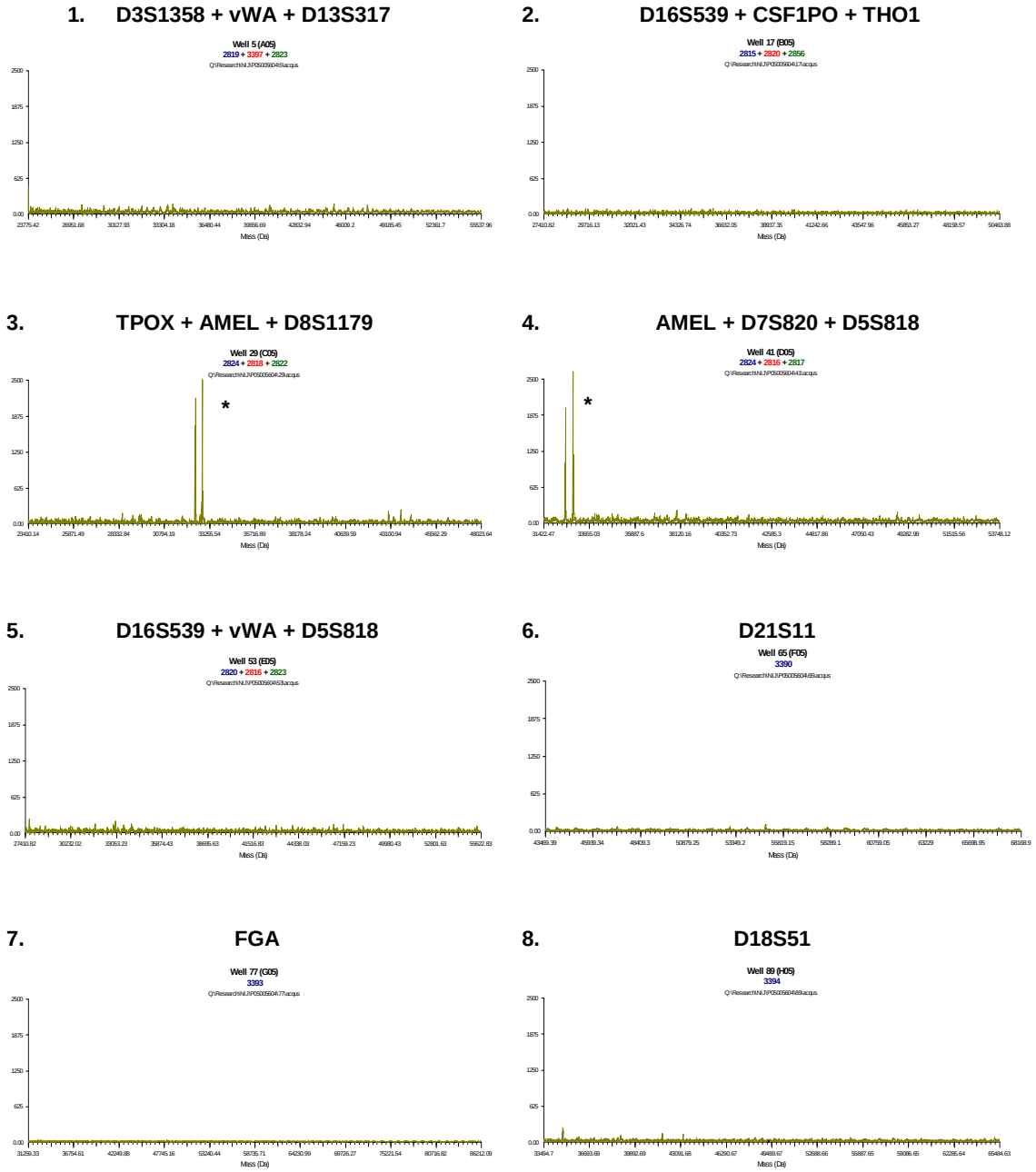


Figure 9. Results of thermocycling 10 ng per reaction of canine DNA in an STR assay plate. No products consistent with human STR PCR products are apparent. In another replicate, weak products corresponding to D18S51 alleles 11 and 14 were detected. This appears to have been introduced (not amplified from the dog DNA), as there is no indication of these products in this replicate. These products are not the same as the control DNA (N31774) used in this experiment.

* Amelogenin amplifies a 105-bp product with a probable base count of [A20 G20 C25 T40]. This is inconsistent with a human product and is shorter than both the female human allele product (109 bp: [A22 G22 C27 T38]) and the male human allele product (115 bp: [A27 G24 C28 T36]). Interestingly, this product is also amplified in feline DNA and appears to be identical in cat and dog.

1 ng N31774 + 10 ng canine DNA

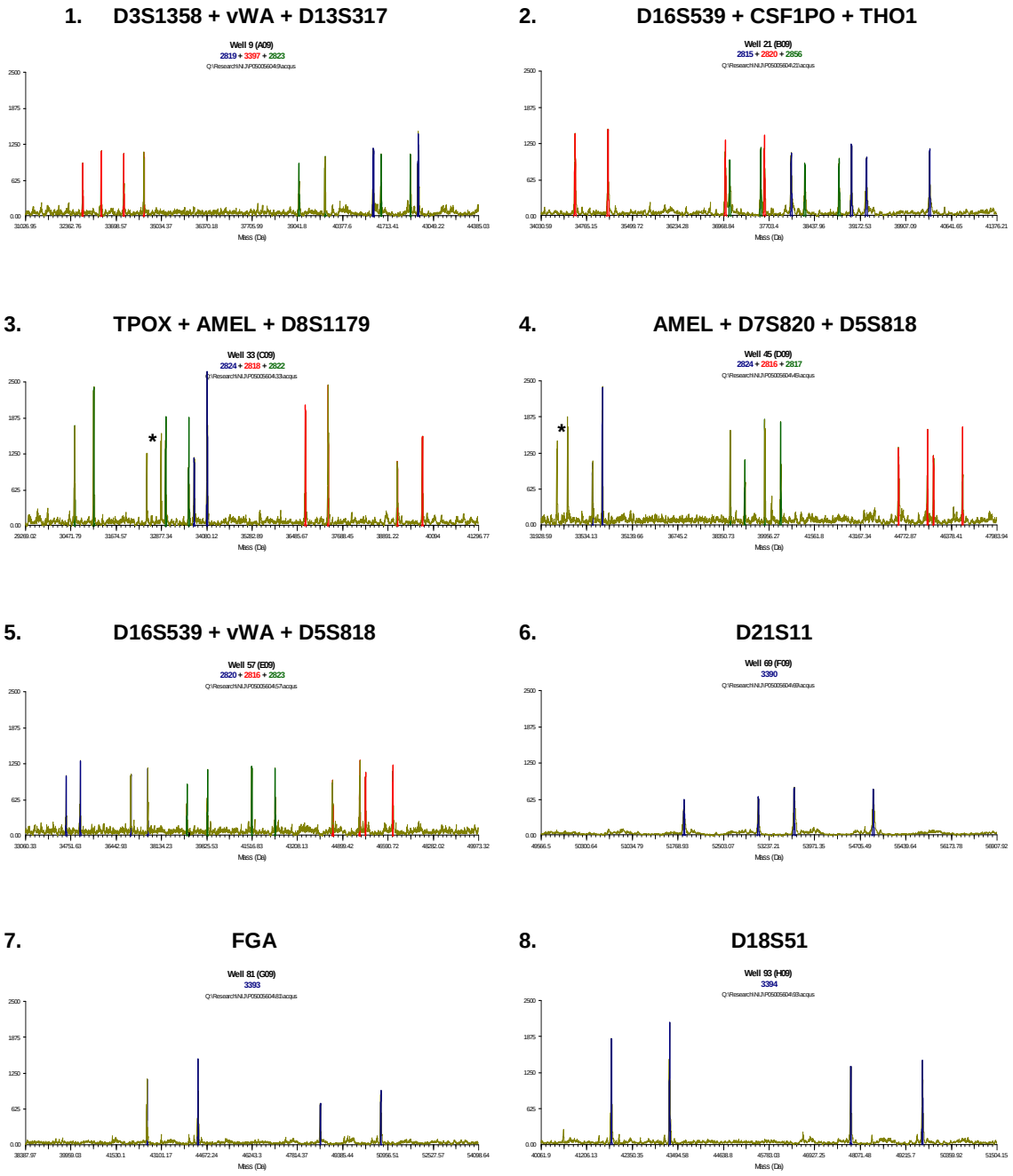


Figure 10. Results of thermocycling 1 ng human blood-derived DNA (N31774) in the presence of 10 ng per reaction of canine DNA in an STR assay plate. A full correct profile is obtained for the human DNA and no additional products are assigned. The full profile obtained is shown in panel G. One of two duplicates is displayed. Results were consistent between duplicates.

* Amelogenin amplifies a 105-bp product with a probable base count of [A20 G20 C25 T40]. This is inconsistent with a human product and is shorter than both the female human allele product (109 bp: [A22 G22 C27 T38]) and the male human allele product (115 bp: [A27 G24 C28 T36]). Interestingly, this product is also amplified in feline DNA and appears to be identical in cat and dog.

10 ng feline DNA

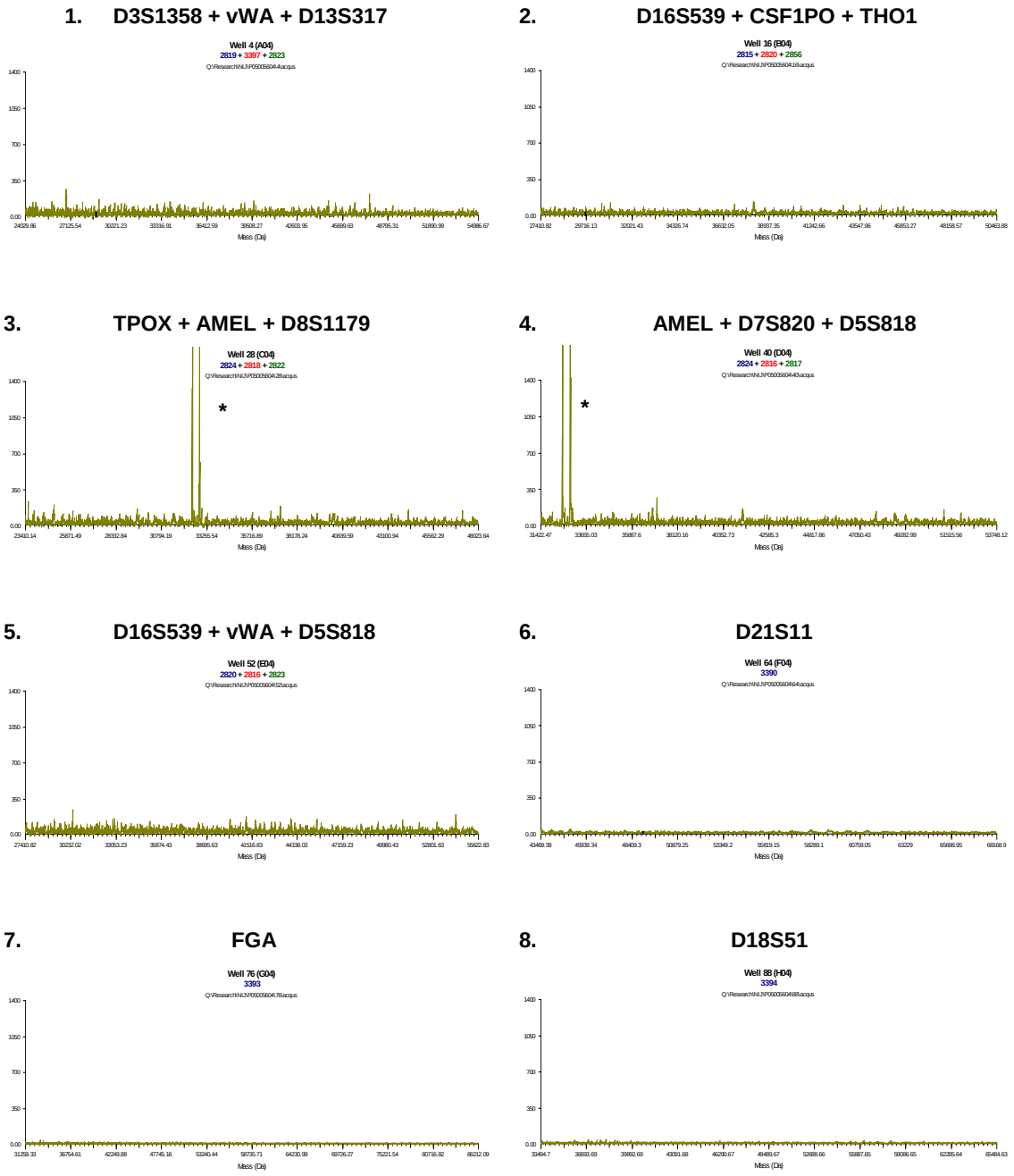


Figure 11. Results of thermocycling 10 ng per reaction of feline DNA in an STR assay plate. No products consistent with human STR PCR products are apparent. In another replicate, a weak product corresponding to D5S818 allele 10 (G->T) was detected. This appears to have been introduced (not amplified from the dog DNA), as there is no indication of this product in this replicate. This product is not the same as the control DNA (N31774) used in this experiment.

* Amelogenin amplifies a 105-bp product with a probable base count of [A20 G20 C25 T40]. This is inconsistent with a human product and is shorter than both the female human allele product (109 bp: [A22 G22 C27 T38]) and the male human allele product (115 bp: [A27 G24 C28 T36]). Interestingly, this product is also amplified in canine DNA and appears to be identical in cat and dog.

1 ng N31774 + 10 ng feline DNA

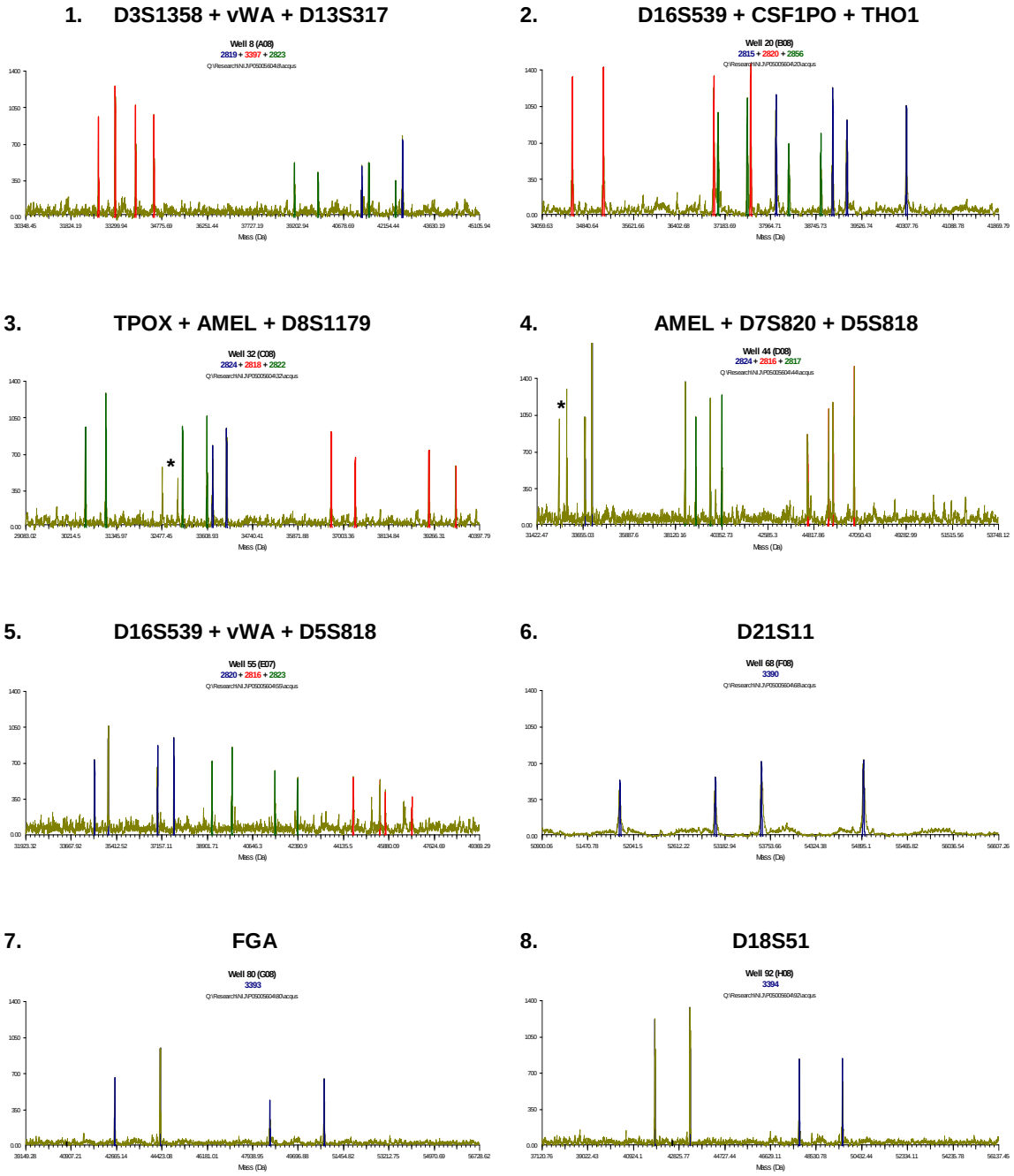


Figure 12. Results of thermocycling 1 ng human blood-derived DNA (N31774) in the presence of 10 ng per reaction of feline DNA in an STR assay plate. A full correct profile is obtained for the human DNA and no additional products are assigned. The full profile obtained is shown in panel G. One of two duplicates is displayed. Results were consistent between duplicates.

* Amelogenin amplifies a 105-bp product with a probable base count of [A20 G20 C25 T40]. This is inconsistent with a human product and is shorter than both the female human allele product (109 bp: [A22 G22 C27 T38]) and the male human allele product (115 bp: [A27 G24 C28 T36]). Interestingly, this product is also amplified in feline DNA and appears to be identical in cat and dog.

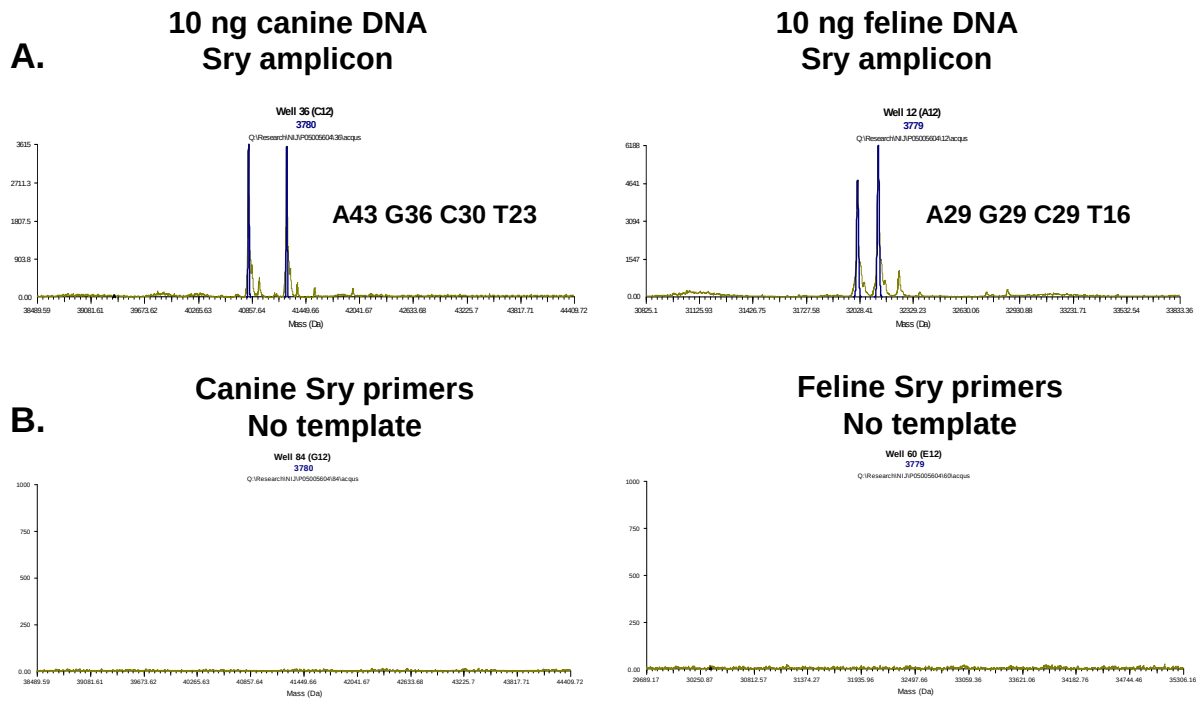


Figure 13. Amplification of canine and feline DNA with species-specific control primers. A.) 10 ng canine DNA (left) or feline DNA (right) in the presence of primer pairs specific to the canine Y-chromosomal Sry gene (left) or the feline Y-chromosomal Sry gene (right). A single strong product with the expected base composition [A43 G36 C30 T23] for canine or [A29 G29 C29 T16] for feline is produced. B.) Canine (left) or feline (right) primer pairs thermocycled in the presence of water rather than template DNA. No products are generated.

Table 8. The genotypes observed in each of duplicate assays for human DNA alone (1 ng, top two rows), 10 ng of canine or feline DNA (next four rows), 1 ng of human DNA in the presence of 10 ng canine or feline DNA (next four rows), or water (last row). Consistent results were observed between duplicates.

* There was a weak detection of alleles 11 and 14 in dog DNA prep #2 for D18S51 and a weak detection of 10 (G→T) in cat DNA prep #1for D5S818. Neither of these repeated in the replicate run at the same input concentration and neither appeared in the presence of 10-fold less specific human DNA run in the presence of the exogenous animal DNA.

Sample	AMEL	CSF1PO	D13S317	D16S539	D18S51	D21S11	D3S1358	D5S818	D7S820	D8S1179	FGA	THO1	TPOX	vWA
N31774, 1 ng	X, ---	11, 12	12 (A->T), 12	9, 11	12, 17	29 (A->G), 30 (G->A)	15 (2G->2A), 16 (2G->2A)	11, 12 (G->T)	9, 10	13, 15 (A->G)	20, 25	6, 7	9, 11	15, 17
N31774, 1 ng	X, ---	11, 12	12 (A->T), 12	9, 11	12, 17	29 (A->G), 30 (G->A)	15 (2G->2A), 16 (2G->2A)	11, 12 (G->T)	9, 10	13, 15 (A->G)	20, 25	6, 7	9, 11	15, 17
Dog DNA prep 1	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
Dog DNA prep 2	---, ---	---, ---	---, ---	---, ---	11, 14*	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
Cat DNA prep 1	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	10 (G->T), --- *	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
Cat DNA prep 2	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
1 ng N31774 + 10 ng dog DNA prep 1	X, ---	11, 12	12 (A->T), 12	9, 11	12, 17	29 (A->G), 30 (G->A)	15 (2G->2A), 16 (2G->2A)	11, 12 (G->T)	9, 10	13, 15 (A->G)	20, 25	6, 7	9, 11	15, 17
1 ng N31774 + 10 ng dog DNA prep 2	X, ---	11, 12	12 (A->T), 12	9, 11	12, 17	29 (A->G), 30 (G->A)	15 (2G->2A), 16 (2G->2A)	11, 12 (G->T)	9, 10	13, 15 (A->G)	20, 25	6, 7	9, 11	15, 17
1 ng N31774 + 10 ng cat DNA prep 1	X, ---	11, 12	12 (A->T), 12	9, 11	12, 17	29 (A->G), 30 (G->A)	15 (2G->2A), 16 (2G->2A)	11, 12 (G->T)	9, 10	13, 15 (A->G)	20, 25	6, 7	9, 11	15, 17
1 ng N31774 + 10 ng cat DNA prep 2	X, ---	11, 12	12 (A->T), 12	9, 11	12, 17	29 (A->G), 30 (G->A)	15 (2G->2A), 16 (2G->2A)	11, 12 (G->T)	9, 10	13, 15 (A->G)	20, 25	6, 7	9, 11	15, 17
Water	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---

S.1.2. Sensitivity, STR assay

To assess the limits of detection of the STR assay, dilution-to-extinction studies were performed using five different blood-derived DNA templates. Four templates were selected from a set of reference samples obtained from NIST (see Figure 18). The fifth was an internal positive control template (SC35495) that we have used in our mitochondrial profiling assay >2,200 times. The four samples from NIST were quantified at NIST and were present in our internal stocks at 5 ng/μL in 10 mM Tris-Cl, 0.1 mM EDTA. The Ibis positive control template (SC35495) was taken from a 10 ng/υL stock in 10 mM Tris-Cl, 0.1 mM EDTA that was quantified using with the Quantifiler assay (ABI). Each dilution series was performed in duplicate and consisted of 3-fold serial dilutions starting at 1 ng template input/reaction. Template inputs to the STR assay were (per reaction) 1 ng, 333 pg, 111 pg, 37 pg, 12 pg and 0 (water control). PCR was run in to 10 mM Tris-Cl, 75 mM KCl, 3 mM $MgCl_2$, 1 M betaine, 20 mM sorbitol, 200 μM each dNTP, 5 U / rxn Amplitaq Gold. Thermocycling parameters were 95 °C for 15 minutes, [95 °C for 30 seconds, 61°C for 2 minutes, 72 °C for 45 seconds] x 40 cycles, 72 °C for 4 minutes, 60 °C for 30 minutes, 4 °C hold.

An example of mass spectrometry outputs for the dilution series of two triplex reactions is shown in Figure 14. As shown in figure 15, each of the five templates produced a full correct profile down to 111 pg per reaction in both duplicates. Four of 10 replicates produced a full profile down to 37 pg input per reaction. At 37 pg, reaction outputs became somewhat random and all alleles are not reliably detected. At 111 pg / reaction, there was an occasional drop-out of vWA alleles from triplex 1 (D3S1358, D13S317 and vWA). However, the vWA primer pair (2824) also exists in triplex 5 (D16S539, D5S818 and vWA). Based upon the observation that ~100 pg input per reaction is sufficient to obtain a full profile, and that 333 pg input/reaction generates solid profiles in all cases, it is currently recommended that a minimum of >200 pg/reaction be used with this assay. This corresponds to 1.6 ng of sample per assay. The limit of full profile detection (111 pg here) corresponds to < 1 ng per assay.

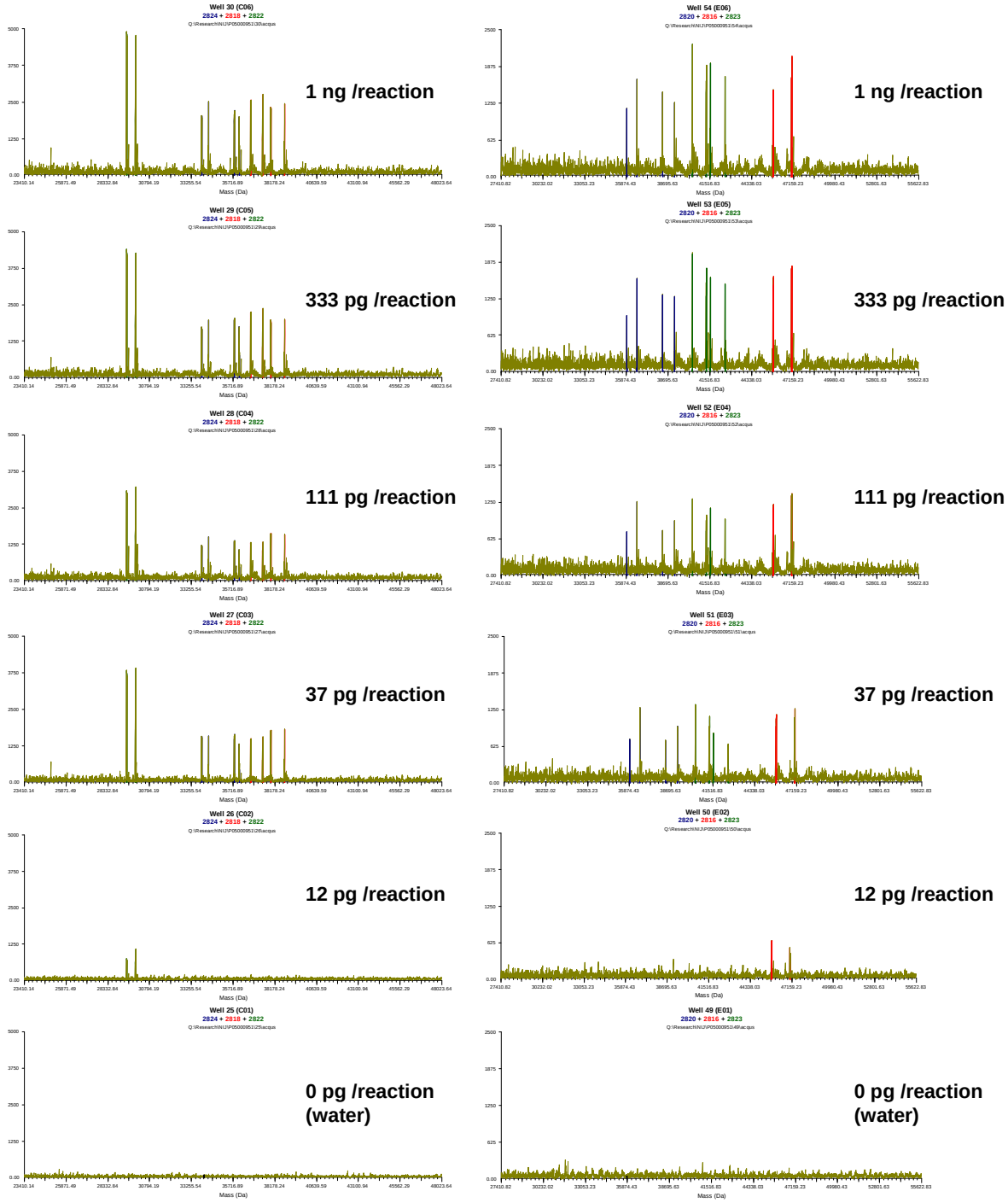


Figure 14. Dilution-to-extinction (DTE) with two STR triplex reactions. 3-fold serial dilutions were performed on 5 blood-derived DNA templates. Deconvolved spectra are shown for one of two duplicates for two STR triplexes at 1 ng, 333 pg, 111 pg, 37 pg, 12 pg, and no template input per reaction.

A. NIST sample UT57312

Amount (pg)	Ave signal intensity	AMEL	CSF1PO	D13S317	D16S539	D18S51	D21S11	D3S1358	D5S818	D7S820	D8S1179	FGA	THO1	TPOX	vWA
0	0	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
12	965.18	X, Y	11, 12	10 (A->T), 11	10, 12	16, 18	---, ---	17 (A->C), ---	12 (G->T), 12	11, 12 (T->A)	---, ---	23, 24	9.3, ---	8, ---	---, ---
37	1311.698	X, Y	11, 12	10 (A->T), 11	10, 12	16, 18	28, 33.2	17 (G->A), 17 (2G->2A)	12 (G->T), 12	11, 12 (T->A)	13, 14 (G->A)	23, 24	9.3, ---	8, ---	16, 17
111	1451.202	X, Y	11, 12	10 (A->T), 11	10, 12	16, 18	28, 33.2	17 (G->A), 17 (2G->2A)	12 (G->T), 12	11, 12 (T->A)	13, 14 (G->A)	23, 24	9.3, ---	8, ---	16, 17
333	1729.905	X, Y	11, 12	10 (A->T), 11	10, 12	16, 18	28, 33.2	17 (G->A), 17 (2G->2A)	12 (G->T), 12	11, 12 (T->A)	13, 14 (G->A)	23, 24	9.3, ---	8, ---	16, 17
1000	1770.391	X, Y	11, 12	10 (A->T), 11	10, 12	16, 18	28, 33.2	17 (G->A), 17 (2G->2A)	12 (G->T), 12	11, 12 (T->A)	13, 14 (G->A)	23, 24	9.3, ---	8, ---	16, 17
0	0	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
12	958.979	X, Y	11, 12	10 (A->T), 11	10, 12	16, 18	28, ---	17 (G->A), 17 (2G->2A)	12 (G->T), 12	11, 12 (T->A)	13, ---	23, 24	9.3, ---	8, ---	16, 17
37	1309.648	X, Y	11, 12	10 (A->T), 11	10, 12	16, 18	28, ---	17 (G->A), 17 (2G->2A)	12 (G->T), 12	11, 12 (T->A)	13, 14 (G->A)	23, 24	9.3, ---	8, ---	16, 17
111	1522.466	X, Y	11, 12	10 (A->T), 11	10, 12	16, 18	28, 33.2	17 (G->A), 17 (2G->2A)	12 (G->T), 12	11, 12 (T->A)	13, 14 (G->A)	23, 24	9.3, ---	8, ---	16, 17
333	1525.692	X, Y	11, 12	10 (A->T), 11	10, 12	16, 18	28, 33.2	17 (G->A), 17 (2G->2A)	12 (G->T), 12	11, 12 (T->A)	13, 14 (G->A)	23, 24	9.3, ---	8, ---	16, 17
1000	1681.712	X, Y	11, 12	10 (A->T), 11	10, 12	16, 18	28, 33.2	17 (G->A), 17 (2G->2A)	12 (G->T), 12	11, 12 (T->A)	13, 14 (G->A)	23, 24	9.3, ---	8, ---	16, 17

B. NIST sample WA29584

Amount (pg)	Ave signal intensity	AMEL	CSF1PO	D13S317	D16S539	D18S51	D21S11	D3S1358	D5S818	D7S820	D8S1179	FGA	THO1	TPOX	vWA
0	NaN	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
12	654.132	X, Y	11, 12	11, ---	11, 12	16, ---	---, ---	---, ---	9 (G->T), 11 (G->T)	11 (T->A), ---	11, ---	22, 25	7, ---	8, 11	---, ---
37	875.442	X, Y	11, 12	11, 14	11, 12	16, ---	---, ---	17, ---	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, ---	8, 11	18, ---
111	1222.261	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
333	1394.78	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
1000	1394.22	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
0	NaN	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
12	683.392	X, Y	11, 12	---, ---	11, ---	16, ---	31.2, ---	---, ---	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, ---	7, ---	8, 11	18, ---
37	863.005	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
111	1258.158	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
333	1272.394	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
1000	1470.912	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---

C. NIST sample PT84222

Amount (pg)	Ave signal intensity	AMEL	CSF1PO	D13S317	D16S539	D18S51	D21S11	D3S1358	D5S818	D7S820	D8S1179	FGA	THO1	TPOX	vWA
0	NaN	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
12	625.704	X, Y	10, 12	12, 13	9, 10	15, 17	---, ---	---, ---	12 (G->T), ---	9, ---	12 (A->G), ---	23, 23.2	8, ---	9, 10	---, ---
37	870.935	X, Y	10, 12	12, 13	9, 10	15, 17	---, ---	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
111	1142.98	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
333	1261.82	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
1000	1599.667	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
0	NaN	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
12	715.223	X, ---	10, 12	12, ---	9, 10	17, ---	---, ---	---, ---	14, ---	9, 11	12 (A->G), 16 (A->G)	23.2, ---	7, ---	9, 10	18, ---
37	730.527	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
111	1039.302	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
333	1275.045	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
1000	1429.58	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18

D. NIST sample ZT80863

Amount (pg)	Ave signal intensity	AMEL	CSF1PO	D13S317	D16S539	D18S51	D21S11	D3S1358	D5S818	D7S820	D8S1179	FGA	THO1	TPOX	vWA
0	NaN	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
12	792.727	X, Y	11, 12	8, 12	12, ---	14, ---	---, ---	---, ---	11, 12	10, 11	---, ---	23, ---	9.3, ---	8, ---	17, 18
37	912.743	X, Y	11, 12	8, 12	11, 12	14, ---	29, ---	15 (2G->2A), 18	11, 12	10, 11	10, 15	23, 26	9.3, ---	8, ---	17, 18
111	1238.926	X, Y	11, 12	8, 12	11, 12	14, ---	29, ---	15 (2G->2A), 18	11, 12	10, 11	10, 15	23, 26	9.3, ---	8, ---	17, 18
333	1448.398	X, Y	11, 12	8, 12	11, 12	14, ---	29, ---	15 (2G->2A), 18	11, 12	10, 11	10, 15	23, 26	9.3, ---	8, ---	17, 18
1000	1611.652	X, Y	11, 12	8, 12	11, 12	14, ---	29, ---	15 (2G->2A), 18	11, 12	10, 11	10, 15	23, 26	9.3, ---	8, ---	17, 18
0	NaN	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
12	648.224	X, Y	11, 12	8, 12	11, 12	14, ---	29, ---	18, ---	11, 12	10, ---	---, ---	23, 26	9.3, ---	8, ---	---, ---
37	1077.403	X, Y	11, 12	8, 12	11, 12	14, ---	29, ---	18, ---	11, 12	10, 11	10, 15	23, 26	9.3, ---	8, ---	---, ---
111	1397.905	X, Y	11, 12	8, 12	11, 12	14, ---	29, ---	15 (2G->2A), 18	11, 12	10, 11	10, 15	23, 26	9.3, ---	8, ---	17, 18
333	1432.167	X, Y	11, 12	8, 12	11, 12	14, ---	---, ---*	15 (2G->2A), 18	11, 12	10, 11	10, 15	23, 26	9.3, ---	8, ---	17, 18
1000	1564.716	X, Y	11, 12	8, 12	11, 12	14, ---	29, ---	15 (2G->2A), 18	11, 12	10, 11	10, 15	23, 26	9.3, ---	8, ---	17, 18

E. Ibis blood sample SC35495

Amount (pg)	Ave signal intensity	AMEL	CSF1PO	D13S317	D16S539	D18S51	D21S11	D3S1358	D5S818	D7S820	D8S1179	FGA	THO1	TPOX	vWA
0	NaN	---, ---	---, ---	---, ---	---, ---	16, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
12	541.65	Y, ---	12, ---	---, ---	---, ---	18, ---	---, ---	---, ---	11, 12	8, 9	---, ---	23, ---	6, ---	10, 11	---, ---
37	710.486	X, Y	11, 12	11 (A->T), 11	---, ---	15, 18	28, 30 (G->A)	19, ---	11, 12	8, 9	12, 15	19, 23	6, 9.3	10, 11	---, ---
111	1039.486	X, Y	11, 12	11 (A->T), 11	8, 9	15, 18	28, 30 (G->A)	17 (G->A), 19	11, 12	8, 9	12, 15	19, 23	6, 9.3	10, 11	17, 18
333	1280.642	X, Y	11, 12	11 (A->T), 11	8, 9	15, 18	28, 30 (G->A)	17 (G->A), 19	11, 12	8, 9	12, 15	19, 23	6, 9.3	10, 11	17, 18
1000	1332.636	X, Y	11, 12	11 (A->T), 11	8, 9	15, 18	28, 30 (G->A)	17 (G->A), 19	11, 12	8, 9	12, 15	19, 23	6, 9.3	10, 11	17, 18
0	NaN	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---
12	577.112	X, Y	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	---, ---	12, ---	8, ---	19, 23	---, ---	10, ---	---, ---
37	797.686	X, Y	11, 12	---, ---	8, 9	15, 18	---, ---	---, ---	11, 12	8, ---	12, 15	---, ---	6, 9.3	10, 11	---, ---
111	852.075	X, Y	11, 12	11 (A->T), 11	8, 9	15, 18	28, 30 (G->A)	17 (G->A), 19	11, 12	8, 9	12, 15	19, 23	6, 9.3	10, 11	17, 18
333	1145.917	X, Y	11, 12	11 (A->T), 11	8, 9	15, 18	28, 30 (G->A)	17 (G->A), 19	11, 12	8, 9	12, 15	19, 23	6, 9.3	10, 11	17, 18
1000	1232.917	X, Y	11, 12	11 (A->T), 11	8, 9	15, 18	28, 30 (G->A)	17 (G->A), 19	11, 12	8, 9	12, 15	19, 23	6, 9.3	10, 11	17, 18

Figure 15. Limit of sensitivity for STR assay. 3-fold serial dilutions were performed on 5 blood-derived DNA templates. In duplicate dilution series, 111 pg template input/reaction produced a full, correct profile.

* One duplicate of the D21S11 reaction in sample ZT80863 failed cleanup or injection on the T5000 instrument, most likely due to a transiently-plugged needle. There was therefore no data generated for one reaction at the 333 pg input for this sample.

S.1.3. Reproducibility, precision and accuracy, STR assay

To assess the reproducibility of the STR typing assay, three templates were selected from the NIST reference sample set and run a combined total of 67 times at 1 ng template input per reaction. Correct typing results were obtained for all runs of each template, with the exception of two wells out of 536 that failed due to no template addition or a plugged injection into the mass spectrometer (Table 9). In each of the two wells, no result was obtained. This represents 0.21% of the individual loci analyzed. A third failed well had three primer pairs (D16S539, vWA and D5S818) that were also present in other triplexes in the same assay, and so caused no loss of genotyping results. In all, 0.56% of the wells in this run failed. Note that there were a total of 12 sequence-polymorphic alleles in the samples used for STR reproducibility (Table 9, D21S11 - 30 (G→A), D3S1358 - 15 (G→A) and vWA - 15 (G→A) for sample ZT80925, D3S1358 – 17 (G→A), D5S818 – 9 (G→T) + 11 (G→T), D7S820 – 11 (T→A) and D8S1179 – 14 (G→A) for sample WA29584, and D21S11 - 30 (G→A), D3S1358 - 15 (2G→2A), D5S818 - 12 (G→T) and D8S1179 - 12 (A→G) + 16 (A→G) for sample PT84222). All variant alleles were reproducibly assigned and it is clear through replication that these variants are genuine. We have also sequenced three of these alleles (D3S1358 - 15 (G→A), D3S1358 - 15 (2G→2A), and D8S1179 - 12 (A→G)) for confirmation (see Population studies, Table 14). Note also that two equal-length heterozygous allele combinations were present in these samples ([15, 15 (G→A)] for vWA in sample ZT80925 and [17, 17 (G→A)] for D3S1358 in sample WA29584). These would be mistaken for homozygous genotypes by conventional STR typing.

Inter-locus product balance was assessed by comparing the average ratios between total signal outputs for each combination of loci for all runs performed for each sample (Figure 16, panels A-E.). The sample (or combination of observed alleles for each sample) appeared to have a strong effect on the observed balance between products in the triplexes in some cases. For example, in Figure 16, panel B. (triplex 2), sample ZT80925 produced much stronger CSF1PO signals relative to D16S539 and THO1 signals compared to the ratios observed for the same loci on sample PT84222, whereas ratios between D163539 and THO1 were very even in all samples in the same triplex.

Likewise, in triplex 5 (Figure 16, panel E), sample WA29584 produced stronger signals for D5S8181 and D16S539 relative to vWA compared to the other two samples. Never-the-less, the sample-dependent variations in inter-locus product balance did not interfere with the ability to assign the correct genotype for each sample.

Measurement precision was assessed by comparing the predicted mass to the observed mass (automatically assigned by our data processing software) for each allele assignment for the 67 assay runs used in the reproducibility study. As forward and reverse DNA strands from each PCR product are separated and measured separately in the mass spectrometer, two independent measurements are used to accurately assign each product. Both measurements were included in this analysis for all assignments and all DNA strand mass measurements were treated as independent. Figure 17 shows the distribution of mass measurement deviations observed over 4,406 independent product assignments for the STR assay. Mass measurements were normally distributed around a deviation of 0, suggesting that there is no overall systematic bias in the assay system in terms of accuracy of mass measurements. Approximately 92% of the measurements were within -35 ppm and 35 ppm of the predicted mass for the assigned product (Figure 17). A deviation of 35 ppm corresponds to approximately 1 Da (1 proton mass) for a DNA strand of about 100 bases, or about 30 kDa, or about 2 Da for the largest single-plexed products we normally expect to observe, which approach about 60 kDa. The average measurement deviation was -0.054 $\pm$ 20.3 ppm, with an average deviation magnitude (absolute value) of 15.2 $\pm$ 13.5 ppm. This precision is beyond what is required to make unambiguous allele assignments and allows us to accurately assign base composition polymorphisms for equal-length alleles. The mass spectrometry resolution we see in this assay is also sufficient to simultaneously assign different same-length alleles within the same individual when they differ by sequence polymorphisms.

Table 9. Results of STR typing of three samples a total of 67 times. Three templates were selected from the set of reference samples supplied by NIST and run 19 times (ZT80925, first 19 runs), or 24 times (WA29584 and PT84222, next 48 runs). 1 ng per reaction was used in all reactions. There were three failed wells out of the 536 wells run (0.56%), causing the loss of results for 2 loci (0.21%). A failed well in run 55 (D18S51) appears to have received no template. Another failed locus (D21S11 in run 9) appears to have been an injection failure into the mass spectrometer. The three primer pairs in the third failed well (well E09 in run 55) were redundant with other triplex reactions in the same assay run and caused no loss of genotyping results.

Run	Sample	AMEL	CSF1PO	D13S317	D16S539	D18S51	D21S11	D3S1358	D5S818	D7S820	D8S1179	FGA	THO1	TPOX	vWA
1	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
2	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
3	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
4	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
5	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
6	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
7	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
8	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
9	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	---, ---	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
10	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
11	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
12	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
13	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
14	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
15	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
16	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
17	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
18	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
19	NIST-ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 (G->A)	15 (G->A), 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 (G->A)
20	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
21	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
22	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
23	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
24	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
25	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
26	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
27	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
28	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
29	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
30	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
31	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
32	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
33	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
34	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
35	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
36	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
37	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
38	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
39	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
40	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
41	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
42	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
43	NIST-WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 (G->A)	9 (G->T), 11 (G->T)	8, 11 (T->A)	11, 14 (G->A)	22, 25	7, 9.3	8, 11	18, ---
44	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
45	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
46	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
47	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
48	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
49	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
50	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
51	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
52	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
53	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
54	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
55	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	---, ---	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
56	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
57	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
58	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
59	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
60	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
61	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
62	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
63	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
64	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
65	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
66	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18
67	NIST-PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 (G->A)	15 (2G->2A), ---	12 (G->T), 14	9, 11	12 (A->G), 16 (A->G)	23, 23.2	7, 8	9, 10	17, 18

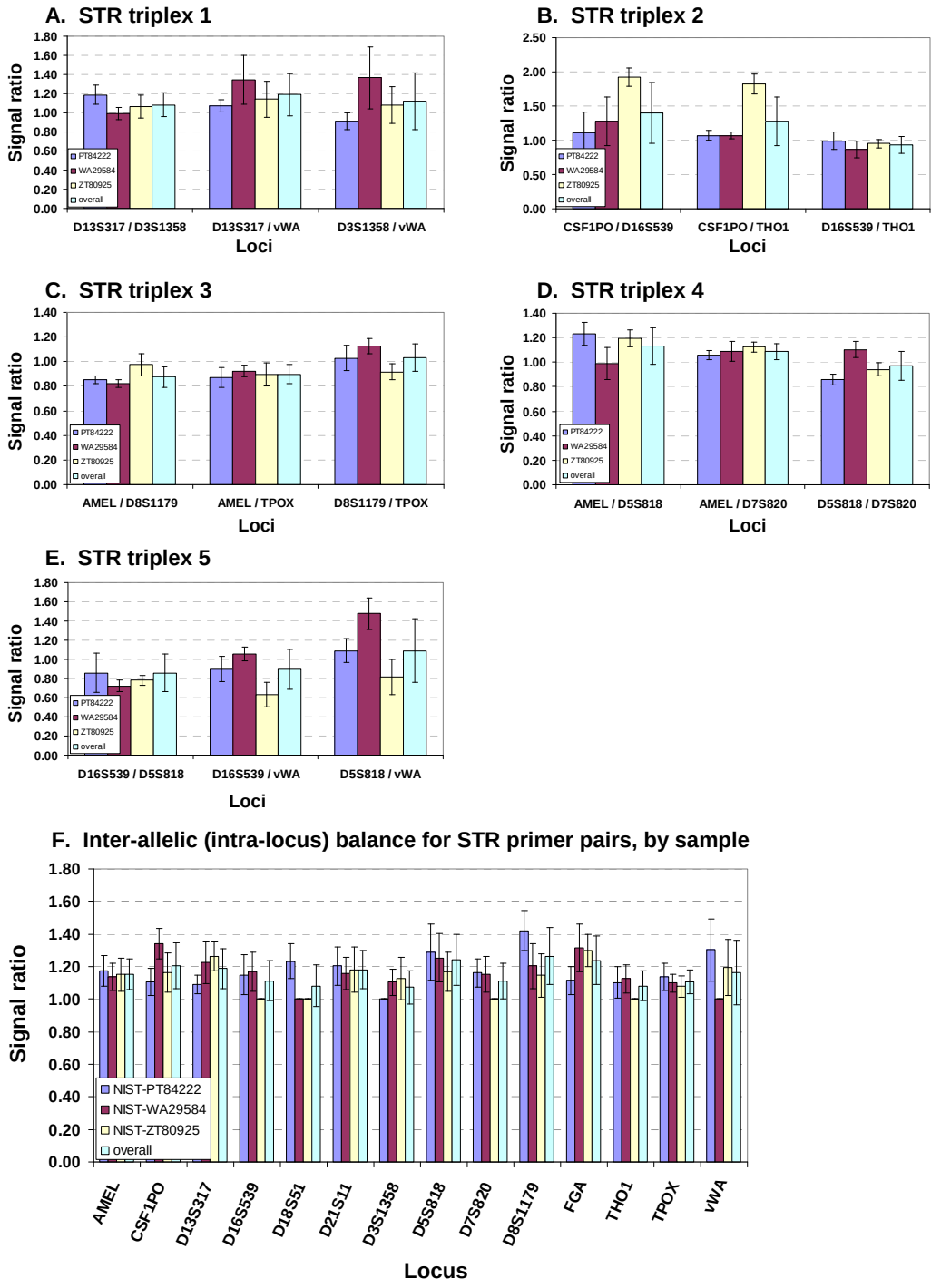


Figure 16. Inter-locus and intra-locus product balance for STR primers. Total signal output ratios for each pair of loci in each triplex are plotted by sample for each STR triplex over the three samples (67 independent assay runs) used in the reproducibility study in panels A-E. The right-most bar in each graph is the average locus:locus ratio for all samples combined. All ratios are expressed as the total signal ratio of the locus on the left side of each label compared to the locus on the right side of each label. F.) The average allele:allele ratio for each primer pair in the STR assay plotted by sample and overall for all samples combined (right-most bar for each locus). Ratios are all expressed as the higher signal allele divided by the lower signal allele.

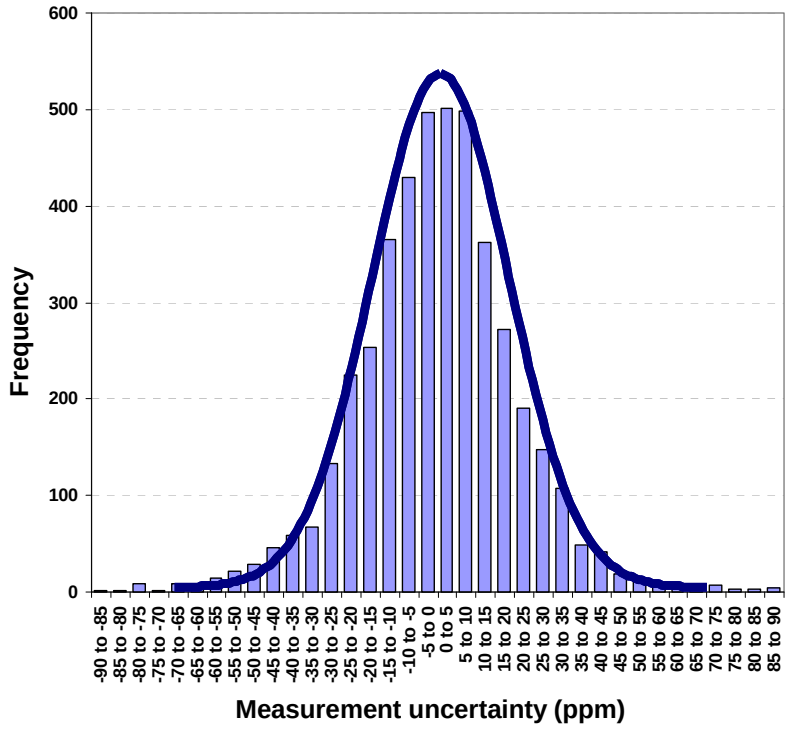


Figure 17. Mass measurement precision for STR assay. 4,406 mass-product assignments were assessed over the 67 repeated samples in the STR assay. Observed mass measurement deviations in parts per million were binned in 5 ppm increments and plotted by frequency of occurrence. Observed measurement deviations are normally distributed around zero and there does not appear to be a strong systematic bias toward measurement errors lower or higher than the true product masses. 91.8% of the measurements were within -35 to 35 ppm of the predicted product mass. 35 ppm corresponds to ~1 Da (one proton mass) for 30 kDa DNA molecule (about 100 nucleotides), or about 2 Da for our largest single-plexed products that can approach 60 kDa.

S.1.4 Reference samples: STR analysis of NIST samples

We received 95 DNA samples from John Butler at NIST that were tested in our updated STR assay. Samples were received as dried DNA (original extracts were derived from blood samples) that was resuspended to a final stock concentration of 5 ng/μl. Each sample was tested in the 14-locus assay (shown in Figure 8) using 1 ng of input DNA per reaction and the buffer and cycling conditions described in section S.1.0. The samples received from NIST were tested in a blinded manner, but were in fact published reference samples consisting of DNA derived from 31 Caucasian, 32 African American, and 32 Hispanic individuals (Figure 18). Reference to these samples can be found in Butler et. al., 2003[11] at STRbase

(http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm).

	Caucasian				African American				Hispanic			
	1	2	3	4	5	6	7	8	9	10	11	12
A	Blank	UT57318	WT51362	WA29594	JT51471	OT05897	PT84223	PT84232	GT37778	GT37900	TT51422	ZT80786
B	UT57300	WT51342	WT51373	WA29612	JT51499	OT05898	PT84224	PT84234	GT37812	GT37913	TT51435	ZT80815
C	UT57301	WT51343	WT51378	ZT81387	OT05888	OT05899	PT84225	PT84236	GT37828	JT52076	TT51483	ZT80826
D	UT57302	WT51345	WT51381	MT94859	OT05890	OT05901	PT84226	PT84239	ZT80932	OT07280	TT51511	ZT80863
E	UT57303	WT51354	WT51386	MT94866	OT05892	PT84214	PT84227	PT84240	GT37862	PT85612	TT51530	ZT80865
F	UT57310	WT51355	BC11352	MT94868	OT05893	PT84215	PT84228	PT84241	GT37864	PT85658	ZT80731	ZT80869
G	UT57312	WT51358	MT97172	MT94869	OT05894	PT84216	PT84230	PT84242	GT37869	TT51399	ZT80737	ZT80870
H	UT57317	WT51359	WA29584	MT94875	OT05896	PT84222	PT84231	PT84243	GT37888	TT51407	ZT80782	ZT80925

Figure 18. Reference samples obtained from NIST.

Thirty of these samples were repeated for verification. In addition, all samples were analyzed separately at all loci in single-plex reactions using the 14 primer pairs in Table 7 for cross-comparison of results. All samples were additionally analyzed in single-plex reactions containing ^{13}C-dGTP in the five loci that displayed polymorphisms in the 47 AFDIL and FBI samples previously tested (D13S317, D5S818, D7S820, D8S1179 and vWA, Tables 2, 3 and 4). An example of mass spectrometry results for a representative NIST sample (African American sample PT84222) is shown in Figure 19. A detailed view of one spectrum containing a tri-plexed reaction for the same sample is shown in Figure 20. Typing results for the 95 samples are shown in Table 10. Our typing results were 100% concordant with the typing results obtained at NIST using the Promega PowerPlex 16 and ABI Minifiler systems on the same samples for the base allele calls (John Butler, personal communication). The Minifiler, PowerPlex, and Ibis systems all demonstrated two minor discrepancies with the Identifiler system originally reported for the reference samples: 1.) In D18S51 from sample ZT80731, Identifiler gives a genotype of [15, 15] whereas in our assay, PowerPlex, and Minifiler it gives a genotype of [13, 15]. 2.) In CSF1PO from sample TT51483, Identifiler gives a genotype of [10.3,

11], whereas in our assay, PowerPlex Minifiler and PowerPlex it gives a genotype of [11, 11] (John Butler, personal communication).

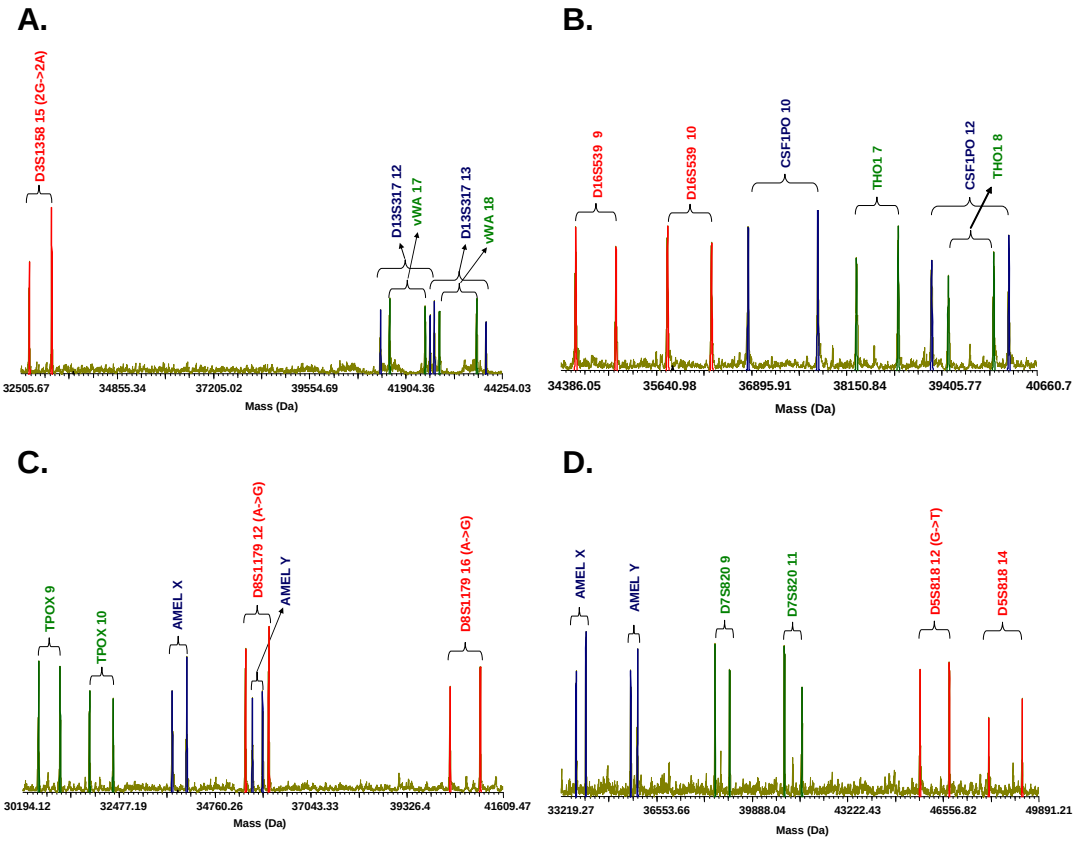


Figure 19. Part 1. Example spectra for STR typing of representative sample PT84222 obtained from NIST. In each spectrum, forward and reverse strands of PCR products are color coded according to primer pair / locus. Allele assignments are indicated above each forward / reverse strand product pair. A.) D3S1358, D13S317 and vWA. This individual was homozygous at D3S1358 for an allele with the same length as allele 15, but containing two G→A differences from allele 15. B.) D16S539, CSF1PO and THO1. C.) AMEL, TPOX and D8S1179. The individual presented variant allele 12 and 16 at locus D8S1179. Both alleles contained an A→G SNP relative to the reference. D.) AMEL, D7S820 and D5S818. The individual presented a variant allele 12 with a G→T SNP at D5S818.

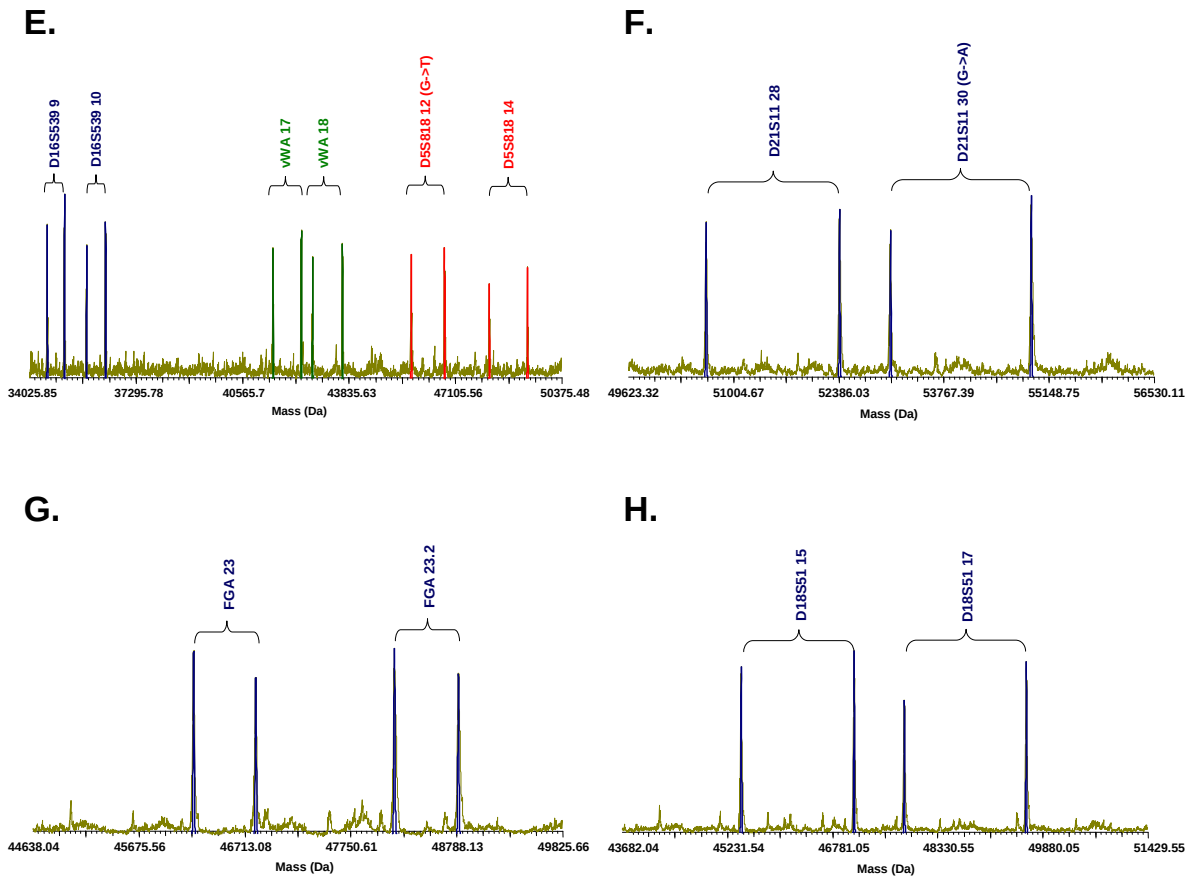


Figure 19, Part 2. Example spectra for STR typing of representative sample PT84222 obtained from NIST. In each spectrum, forward and reverse strands of PCR products are color coded according to primer pair / locus. Allele assignments are indicated above each forward / reverse strand product pair. E.) D16S539, D5S818 and vWA. As in panel D, there was a variant allele 12 with a G→T SNP at D5S818. F.) D21S11. A variant allele 30 (with a G→A SNP) was observed for D21S11. G.) FGA. H.) D18S51.

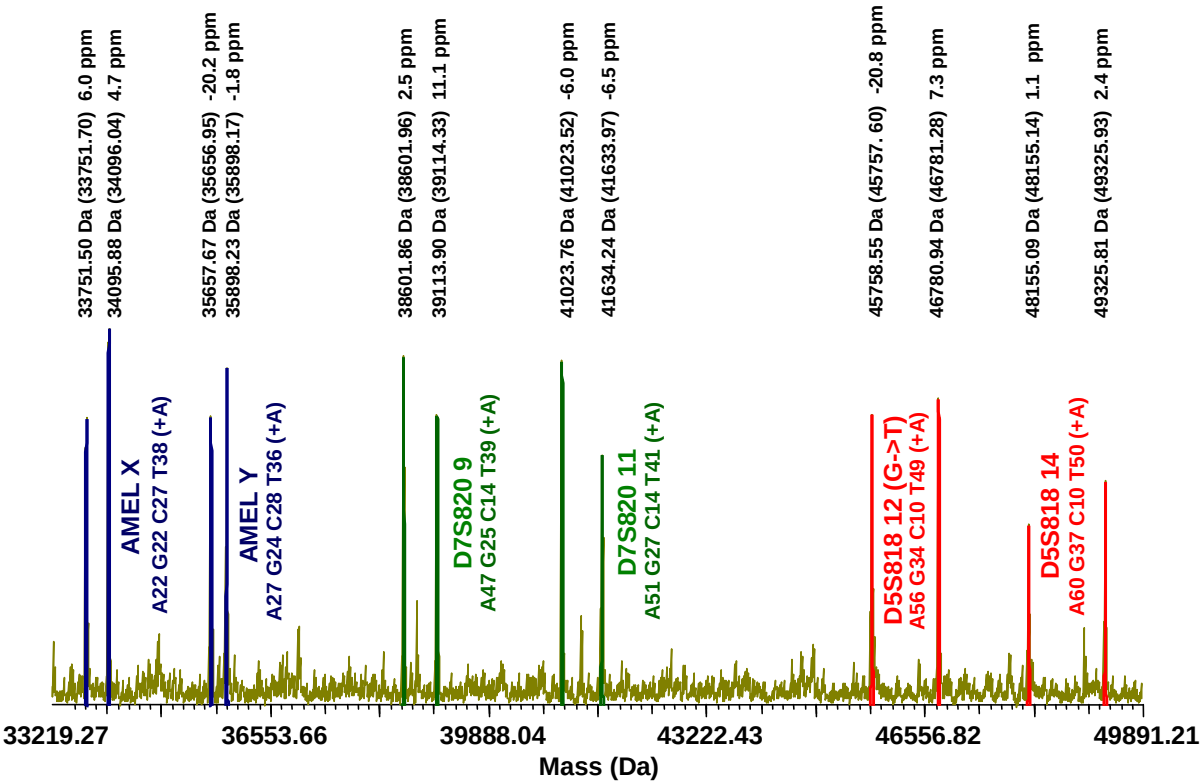


Figure 20. Detailed view of reaction set 4 for representative sample PT84222 obtained from NIST. The assignment of each locus and allele is shown next to color-coded mass peak pairs representing the forward and reverse strands of each PCR product. The observed mass of each peak is shown above each assigned mass peak. In parentheses are the expected masses for each assigned product. The measurement error in parts per million (ppm) is shown on the right of each label. The ppm error is determined here as ((expected mass-observed mass)/expected mass) * 1,000,000.

Table 10. Part 1. Results of STR typing of NIST samples. Caucasian and African American samples are shown. Polymorphic alleles are indicated by coloring of cells as indicated in the bottom legend. The nature of each SNP is indicated by a letter designation next to the base allele call.

	Sample	AMEL	CSF1PO	D13S317	D16S539*	D18S51*	D21S11*	D3S1358*	D5S818	D7S820	D8S1179	FGA*	THO1	TPOX	vWA
Caucasian	UT57300	X, Y	10, 12	8, 12	9 f, 12	10, 13	28, ---	15, 16 d	12, ---	8, 13	10, 15	20, 21	9.3, ---	8, 9	15 d, 19
	UT57301	X, Y	11, 12	8, ---	11, 12	15, 17	28, ---	15 d, 16 d	11 b, 13	8, 11	13, 14	23, 25	6, ---	8, ---	17, 18
	UT57302	X, Y	10, 12	11 c, 12 c	11, ---	14, 17	28, 29	15 d, 17	11, ---	8, 13	10, 13	21, 24	8, 9	8, 9	14 g, 17
	UT57303	X, Y	10, 12	11 c, 13	9, 12	13, 14	29 d, 30	15 d, 17	11, 13	10, 12 e	12, 13	21, 22	8, 9	8, 11	16, 17
	UT57310	X, Y	12, ---	8, 13	9, 13	12, 13	30, 30.2	17 d, 18	12, ---	9, 10	12, ---	20, 21	9.3, ---	9, 11	14 a, 15 d
	UT57312	X, Y	11, 12	10 c, 11	10, 12	16, 18	28, 33.2	17 d, 17 j	12 b, 12	11, 12 e	13, 14 d	23, 24	9.3, ---	8, ---	16, 17
	UT57317	X, Y	11, 12	12 c, 13	9, 12	12, 17	29 d, 30.2	15 d, 15 j	11, ---	8, 11	13, 14 d	22, 23	6, 9.3	8, ---	14 a, 17
	UT57318	X, Y	11, 13	11 c, 12 c	11, ---	16, 18	25.2, 30	15 d, 16 d	10, 12 b	10, 11	12, ---	21, 22	9.3, ---	8, 11	15 d, 17
	WT51342	X, Y	10, 12	11, 14	12, 13	13, 14	29, 31 d	18, ---	12, 13	9, 12	12, 14	24, 25	9, 9.3	9, 11	18, ---
	WT51343	X, Y	11, 12	11, 13	11, ---	14, 16	28, 31.2	16 d, 17 d	11, 12	10, 12 e	12, 13 d	22, ---	6, 7	8, ---	17, 18
	WT51345	X, Y	11, 12	11 c, ---	10, 13	13, ---	29, ---	15 d, 16 d	11, ---	8, 10	13, 14 d	20, 22	7, 9.3	8, ---	17, 19
	WT51354	X, Y	10, 12	11, 14	11, 12	16, ---	30 d, 32.2	15 d, 16	8, 11	7, 9	11, 13	18, 26	6, 9.3	8, 9	13 l, 16 d
	WT51355	X, Y	10, 13	11, 12	12, ---	15, 17	30, 31.2	16 d, 18	11, ---	8, 10	13, 13 d	20, 24	7, 9.3	10, 11	14 a, 16
	WT51358	X, Y	11, 12	11, 12 c	9, 13	20, 22	30, 31 d	14, 18	10, ---	11, 13	11, 13	19, ---	7, 9.3	9, 11	15 d, 16
	WT51359	X, Y	12, ---	13, ---	11, 13	13, 20	27 f, 32.2	14, 15 d	13, ---	10, 12	14, 16	21, 24	9, 9.3	11, ---	15 d, 18 d
	WT51362	X, Y	11, 12	8, 13	9, 11	16, 19	30, 31	16, 17 d	11, 12	10 e, 12 e	10, ---	23, 24	9.3, ---	10, 11	17, 18
	WT51373	X, Y	10, 12	11 c, 12 c	11, 13	14, 17	28, ---	15 d, 18	10, 11	8, 11	11, 14	21, 22	8, 9	8, 11	15 d, 17 d
	WT51378	X, Y	10, ---	8, 12	9, 11	12, 15 i	30, 31 d	15, 18	11, ---	9, ---	13, ---	19, 23	6, 9	8, ---	16, 16 d
	WT51381	X, Y	12, ---	8, 11	9, 11	12, 18	30 d, ---	15 d, 17	12, 14	10, ---	10, 16	22, 24	6, ---	8, 9	15 d, 18
	WT51386	X, Y	12, 13	11, ---	11, 12	17, ---	28, 29	17, 18	11 b, 12	9, 10	11, 14 d	23, ---	9, 9.3	9, 11	19, ---
	BC11352	X, Y	11, 13	13, ---	8, 12	13, 15	28, 30 d	14, 17	10, 11	10, 12 e	13, 14	24, ---	6, 9.3	8, ---	17, ---
	MT97172	X, Y	11, 12	11, 12	9, 12	16, 18	29, 31	15 d, 15 j	11, 12 b	9, 12	13, 15	23, ---	7, ---	8, ---	16, 18
	WA29584	X, Y	11, 12	11, 14	11, 12	16, ---	31.2, 32.2	17, 17 d	9 b, 11 b	8, 11 e	11, 14 d	22, 25	7, 9.3	8, 11	18, ---
	WA29594	X, Y	11, 12	11, 12	11, ---	12, 15	30, 30.2 d	17, 18	11, 12	7, 11	12, 13	22, 25	6, 9	8, ---	14 a, 14 g
	WA29612	X, Y	11, 12	11 c, 13	11, 12	12, 14	28, 30 d	14, 17 d	12, ---	11, 12	13, ---	22, 23.2	6, 9.3	8, 11	14 g, 19
	ZT81387	X, Y	11, ---	9, 13	12, 13	16, 18	28, 32.2	18, ---	11 b, 11	10 e, 11	13, 15	21, 22	7, ---	6, 9	16, 19
	MT94859	X, Y	10, 13	9, 11	9, 12	13, 17	29, 30	16, 17 d	11, 13 b	9, 11 e	11, ---	20, 26	6, 9	8, 11	18, ---
	MT94866	X, Y	11, 12	8, 12 c	12, 13	12, 16	29.2 f, 30	15 d, ---	11, 12	10, 11 e	12, 13 d	22, 22.2	9.3, ---	8, 9	17, 20 e
	MT94868	X, Y	12, ---	9, 11 c	9, 12	12, 17	27 f, 32.2	16 d, 17 d	11, ---	8, 11	12, 14	20, 23	9.3, ---	8, 11	14 a, 16
	MT94869	X, Y	12, ---	11 c, 12	12, 13	15, 19	28, 30	16, ---	11, 12 b	10, 12 e	13, 14 d	20, 23	6, 9	8, ---	14 a, 17 d
	MT94875	X, Y	9, 10	11 c, 12 c	12, ---	14, 16	29 d, 32.2	16 d, 18	11, 12	8, 9	8, 13	18, 25	6, ---	8, 12	16, 19
African American	JT51471	X, Y	10, 13	12, ---	9, 13	16, ---	28, 33.2	15 d, 16 d	11, ---	8, 10	12 f, 13	28 h, 31.2	6, 7	8, ---	16, ---
	JT51499	X, Y	11, 12	11, ---	11, 12	13, 16	30 d, ---	15 d, 18	9 b, 12	10, 11	14, 14 d	22, 26	7, 9.3	8, ---	16 d, 18
	OT05888	X, Y	10, 11	11 c, 11	10, 13	16, 20 h	31.2, 36 o	15 d, 17	12 b, 13 b	9, 11	13, 14	19, 20	7, 8	8, 11	16, 18
	OT05890	X, Y	7, 8	10 c, 12 c	11, 12	12, 18	30, 32.2	16 d, 17	11, 12	9, 10	16, ---	19, 24	7, 8	6, 11	16, 18
	OT05892	X, Y	8, 11	12, ---	9, 12	13, 17	29, 34.1 e	15 j, 18 d	11 b, 11	10, ---	12 f, 14	23, 24	7, 8	9, 10	17, ---
	OT05893	X, Y	11, ---	11 c, ---	10, 13	15, 17	28, 30 d	15.2, 16 j	10, 11	8, 12	13, 14	24, 26	7, 9	8, 9	16, 17
	OT05894	X, Y	7, 10	12 c, 13	11, ---	13.2 h, 16	29, 37 k	15 d, 15 j	12, 13 b	8, 12	14, 15	23, 25	7, ---	8, 11	18, ---
	OT05896	X, Y	11, 12	12, ---	11, ---	18, ---	30 d, 30	15 j, 17 d	11, 12	10, ---	14, 15	19, 25	8, 9	8, 10	15, 18
	OT05897	X, Y	12, 13	11, 12	11, 13	15, 17	27, 30 d	16, 18	12, ---	8, ---	13, ---	22, 23	7, 8	11, ---	16, 18
	OT05898	X, Y	8, 11	12, 13	11, ---	17, ---	29, 32.2	14, 15 j	8 b, 12	8, 13	13, 15	19, 22	8, 9	6, 9	15, 17
	OT05899	X, Y	7, 11	12, ---	9, 11	15, 16	28, 32.2	13 d, 16 j	8 b, 11	11, 12	13 d, 14	22, ---	7, ---	7, 10	16 d, 17
	OT05901	X, Y	10, 11	8, 11	9, 10	18, ---	27 f, 28	15 j, 18 j	11, 12	9, ---	13, 16	22, 23	7, ---	6, 11	18 o, 18
	PT84214	X, Y	11, 12	12 c, 12	12, 13	12, 13.2 h	28, 32.2	16 d, 17 d	12, ---	10, ---	14, 15	21, 24	6, 8	10, 11	17, 17 d
	PT84215	X, Y	7, 11	12, ---	9, 11	15, 16	31.2, 32 f	15 d, 16 j	12 b, 12	8, 9	12 f, 14	22, ---	8, 9.3	8, 11	14 h, 16
	PT84216	X, Y	10, ---	11, 12	12, 13	12, 16	30 d, ---	15 j, 16 j	12, 13 b	10, ---	13, 14	24, ---	7, 8	6, 8	13 m, 18
	PT84222	X, Y	10, 12	12, 13	9, 10	15, 17	28, 30 d	15 j, ---	12 b, 14	9, 11	12 f, 16 f	23, 23.2	7, 8	9, 10	17, 18
	PT84223	X, Y	12, ---	11 c, ---	11, 13	15, 19	29 f, 29	17, 17 j	11, 12	8, 10	14, 15	21, 22	7, ---	8, 11	16, 19 o
	PT84224	X, Y	10, 12	11 c, 12	8, 9	16, 17	30 d, 31	15, 16 j	12, ---	10, ---	14, ---	22, 26	7, ---	11, ---	15, 17
	PT84225	X, Y	10, 12	9, 11 c	9, 12	15, 17	31, ---	17 d, 18 d	12, ---	10, 11	13, ---	24, 25	7, ---	10, 12	14 a, 16 d
	PT84226	X, Y	11, 12	8, 12 c	11, 12	17, ---	28, 32.2	16 d, 16 j	11, 12 b	10, ---	14, ---	23, 26	8, 9.3	8, 11	15 d, 18
	PT84227	X, Y	8, 10	11 c, 12	9, 11	12, 22	31.2, 32.2	15 d, 17 d	9 b, 12 b	10, 11	12, 15	20, 22	6, 8	8, 9	15, 16 d
	PT84228	X, Y	11, 12	11 c, 12	11, ---	12, 20 h	29, 30.2	15 d, 17	10, 12	8, 10	13 d, 14	24, 26 d	7, 9	8, 10	16, 19
	PT84230	X, Y	11, ---	11, 12	9, 13	13, 17	28, 29 f	16 j, 17 f	11, ---	9, 11	13, 14	19, 24	7, ---	8, ---	18, 20 f
	PT84231	X, Y	11, 12	12, ---	12, 13	14, 17	27 f, 29	15 j, 17 d	8 b, 13	8, 9	14, 15 f	22, 24	7, 9	8, 11	16 d, 18
	PT84232	X, Y	9, 10	12 c, 13	11, 12	14, 20 h	27, 28	15 j, 16 j	9 b, 10	8, 10	13, 14	20, 25	7, ---	8, 9	15, 19
	PT84234	X, Y	10, 12	13, ---	9, 11	15, 16	29, 31	15 j, 16	12, 13	10, 11	13 d, 15	19.2, 25	6, 7	6, ---	16, 16 d
	PT84236	X, Y	7, 12	12, 13	11, ---	12, 23	28, 29	16 d, 16 j	12 b, 13	8, 9	12 f, 14	25, 26	7, ---	9, ---	14 h, 17
	PT84239	X, Y	12, 13	12, 13	10, 11	15, 20 h	30, 31 d	15 j, 18	10, 12	11, ---	13, 14	23, 24	7, 8	9, ---	17, 18
	PT84240	X, Y	10, 11	11 c, 12	12, 13	14, 19	28, 29	16 d, 17 j	12 b, 12	9, 10	11, 13	23, 27'	6, 9	9, 11	16, 17
	PT84241	X, Y	12, 13	12, 13	11, 13	16, 17	29 f, 29	14, 18	11 b, 12	10, ---	10, 14	23, 25	6, 7	8, ---	18, ---
	PT84242	X, Y	9, 10	12 c, 12	8, 11	12, 17	30 d, ---	16, 17 d	11, 13	8, 11	13, ---	23, 25	7, 9.3	8, 9	14 a, 17
	PT84243	X, Y	10, 12	11, 12 c	10, 11	13, 16	28, 30 d	15 d, 16 j	13, 14 b	10, 11	13, 14	23, 25	8, 9	6, ---	16 d, 17

a A->G + 2T->2C
b G->T
c A->T
d G->A
e T->A
f A->G
g G->A + T->C
h T->C
i T->G
j 2G->2A
k 3A->3G
l G->A + C->T
m C->T
n C->G
o 2A->2G
p G->C

One or both allele had a polymorphism

Both alleles were the same length, but differentiable by polymorphism(s)

* Exact nature of every observed polymorphism at loci D16S539, D18S51, D21S11, D3S1358 and FGA has not yet been confirmed independently with a mass tag or sequencing (some have been sequenced). Until final verification, there is a small chance that an observed G->A could be a T->C, for example. All polymorphisms at loci D13S317, D5S818, D7S820, D8S1179 and vWA have been confirmed using ^{13}C-enriched dGTP.

Table 10. Part 2. Results of STR typing of NIST samples. Hispanic samples are shown. Polymorphic alleles are indicated by coloring of cells as indicated in the bottom legend. The nature of each SNP is indicated by a letter designation next to the base allele call.

	Sample	AMEL	CSF1PO	D13S317	D16S539	D18S51*	D21S11*	D3S1358*	D5S818	D7S820	D8S1179	FGA*	THO1	TPOX	vWA
Hispanic	GT37778	X, Y	10, 12	9, 13	10, 12	14, 20	28, 30 d	17 d, 18	11, ---	9, 12	14, ---	23, 24	7, ---	9, 11	15, 17 d
	GT37812	X, Y	12, 13	12, 13	11, 13	15, ---	29, 32.2	14 d, 15 j	14 b, 14	8, 12	13, 13 d	22, ---	6, 9.3	11, ---	15, 16
	GT37828	X, Y	10, ---	9, 14	9, 11	14, 17	31.2, 31.2 d	15 d, 16 j	11, 12 b	11, 12	10, 15	22, 23	7, ---	8, ---	16, 18
	ZT80932	X, Y	12, 13	10, 11	11, 13	17, 18	29, 32.2	14, 17 d	11, 12	11, 13	13, ---	23, 24	8, 9	8, ---	15 d, 16
	GT37862	X, Y	10, 11	8, 12	10, 11	11, 15	29, 29 d	16, 17	9 b, 11	9, 11	15, 16	19, 22	6, 7	8, 12	15, 17
	GT37864	X, Y	12, ---	9, 11 c	10, 12	13, 15	29, ---	15 d, 16 d	11, 12	9, 12 e	12, 13	19, ---	7, ---	8, ---	16, ---
	GT37869	X, Y	10, 11	11, 12	9, 12	15, 16	30, 31.2	16 d, 17 d	7, 12	10, 11	13, 13 d	24, 25	6, 7	8, 11	15, 15 d
	GT37888	X, Y	10, 12	11, 12	9, 13	17, 19	30 d, 30.2	13, 15 d	12, ---	10 e, 11	14, 15	23, 25	7, 9	8, ---	15 d, 17
	GT37900	X, Y	11, 12	11, 12	10, 12	12, 13	29, 31.2	14, 17 f	11, ---	9, 11	13, 14 d	23, 26	6, 7	11, ---	17, 19
	GT37913	X, Y	11, 12	11, ---	11, ---	13, 15	29, 31 d	16, 17 d	7, 11	10, 13	12, 13	19, 22	6, 9.3	8, 11	16 d, 17
	JT52076	X, Y	12, ---	8, ---	11, 12	16, 21	30, 32.2	15 d, 16	11, ---	11, 12	12, 14	21, 22	8, 9.3	8, 11	15 d, 18
	OT07280	X, Y	10, 12	9, 13	12, 13	13, 14	30 d, 30	15, 16 d	7, 12	8, 11	12, 13	22, 24	7, 8	8, ---	17, 20
	PT85612	X, Y	10, 12	11 c, 11	12, ---	14, 17	29, ---	15 d, 16 d	7, 10	12, 13 e	12 f, 12	21, 24	7, 8	12, ---	14 a, 19
	PT85658	X, Y	12, 14	9 c, 14	12, 13	14, 15	30, ---	16, 17	10 b, 13	10, 11	10, 12	19, 23	6, 7	8, 9	18, 19
	TT51399	X, Y	10, 12	10 c, ---	10, 11	12, 13	29, 32.2 d	15 d, 15 j	7, 11	11, 12 e	13, 13 n	25, 27	6, 7	8, 12	16, 18
	TT51407	X, Y	10, ---	9, 12 c	11, 13	12, 19	28, 30 d	18, ---	11, ---	10 e, 11	10, 11	22, 23	7, ---	8, 11	14 a, 16
	TT51422	X, Y	10, 12	9, 13 c	13, ---	15, 18	29, 31 d	16, 17	11, 13	10, 12	11, 14	20, 24	6, 8	8, 11	16, 17 d
	TT51435	X, Y	9, 12	9, 10 c	11, 14	12, 17	28, 30.2	15 d, 16 d	10, 13 p	10, 11	10, 12	23, 26	7, 9.3	8, 9	16, 17
	TT51483	X, Y	11, ---	11 c, 11	11, ---	18, 19	30 d, 31	14 d, 18	12 b, 12	8, 10	13, 15	21, 26	9, 9.3	9, 10	14 a, 18
	TT51511	X, Y	11, 12	9, ---	11, 12	13, 22	30 d, 32.2	15 d, 17 d	12, ---	8, 10 e	12, 14	20, 25	7, ---	11, ---	15, 16
	TT51530	X, Y	10, 11	9, 10 c	10, 13	13, 15	31, 31.2	15 d, 18	11, 12	10, 11	12 f, 14	23, 26	7, 9.3	11, 12	16, 18
	ZT80731	X, Y	10, 11	11, 12	12, 13	13, 15	30 d, 30.2	14, 16 j	11, 13	10, ---	11, 13	18, 24	9.3, ---	8, 11	15, 17
	ZT80737	X, Y	11, 13	12, 13	9, 12	12, 16	29, 30.2	15 d, 15 j	12, ---	10, ---	14, 15	21, ---	9.3, ---	9, 12	16, 17
	ZT80782	X, Y	12, ---	8, 10 c	13, ---	14, 15	30 d, 32.2 d	15 d, ---	11, ---	11, ---	10, 14 d	21, 23	8, 9	8, 12	17, 18
	ZT80786	X, Y	10, 11	9, 12 c	9, ---	15, 16	29 d, 30	15 d, 18	11, 12	9, 11	13, 15	20, 23	7, 8	11, 12	15 d, 20
	ZT80815	X, Y	10, 13	13, 15 c	10, 12	14, 16	28, 29 d	14, 15 d	10, 11	11, 12	13, ---	24, ---	7, 8	10, 11	16 d, 17
	ZT80826	X, Y	10, 12	10, 12	9, 11	12, 18	28, 31.2	17, 18 d	9 b, 12	10, 12 e	11, 17 d	21, ---	9.3, ---	8, ---	17, 17 d
	ZT80863	X, Y	11, 12	8, 12	11, 12	14, ---	29, ---	15 j, 18	11, 12	10, 11	10, 15	23, 26	9.3, ---	8, ---	17, 18
	ZT80865	X, Y	10, 12	8, 11	13, ---	16, 17	29, 31.2	15 d, 17	11, 14	9, 11	13 d, 14	21, 23	7, 9.3	8, ---	17, 19
	ZT80869	X, Y	11, 12	9, 11 c	12, 13	15, 16	30, 32.2	14, 15 d	11, 15	9, 10	14, ---	20, 22	6, 9	11, 12	16 d, 17
	ZT80870	X, Y	11, 12	9, 13 c	11, 12	13, 20	29, 34	17, 18	12 b, 13	8, ---	11, 14	21, 23	6, 9	8, 12	15 d, 16
	ZT80925	X, Y	10, 11	11, 14	12, ---	17, ---	27, 30 d	15 d, 17	11, 13	8, ---	16, 18	18, 21	7, ---	8, 11	15, 15 d

a A->G + 2T->2C
b G->T
c A->T
d G->A
e T->A
f A->G
g G->A + T->C
h T->C
i T->G
j 2G->2A
k 3A->3G
l G->A + C->T
m C->T
n C->G
o 2A->2G
p G->C

One or both allele had a polymorphism

Both alleles were the same length, but differentiable by polymorphism(s)

* Exact nature of every observed polymorphism at loci D18S51, D21S11, D3S1358 and FGA has not yet been confirmed independently with a mass tag or sequencing (some have been sequenced). Until final verification, there is a small chance that an observed G->A could be a T->C, for example. All polymorphisms at loci D13S317, D5S818, D7S820, D8S1179 and vWA have been confirmed using ^{13}C-enriched dGTP.

S.1.5. Population studies, STR assay

In the 95 samples from NIST, polymorphisms were found in 10 of the 13 core CODIS loci. A total of 364 polymorphic alleles were detected in 1330 genotype assignments (95 samples typed at 14 loci). Thirteen distinct variant alleles were identified in D21S11, 13 in vWA, 11 in D3S1358, eight in D5S818, seven in D8S1179, six in D13S317, four in D7S820, three in D18S51, two in FGA and one in D16S539. Table 11 shows the frequency of each allele observed for each of the three population groups, with polymorphic alleles annotated with the SNP(s) present relative to the reference allele. Sixty percent of alleles observed in the 95 samples for D3S1358 were polymorphic relative to the reference allele. For this locus, each allele length appears to have three polymorphic variants in the population: one base allele (the choice of reference is somewhat arbitrary), one with a G→A SNP, and one with two G→A SNPs. Two previous studies reporting sequencing of this locus report a complex repeat structure with TCTA and TCTG repeats[12, 13]. In one of these studies, multiple same-length allele pairs were reported with different sequence structure for D3S1358, vWA and FGA, consistent with our results[13]. In addition, our results indicate the presence of more than two same-length variants for multiple allele lengths in D3S1358, vWA and D21S11 (we observed three variants each for several allele lengths at each of these loci). Table 12 shows the overall frequency of observations for polymorphic alleles in the 47 FBI and AFDIL samples plus the 95 NIST samples combined.

Table 11, part 1. Observed frequency of each allele by population in 31 Caucasian, 32 African American and 32 Hispanic samples from NIST.

		Count			Percentage		
Locus	Allele	Caucasian	African American	Hispanic	Caucasian	African American	Hispanic
AMEL	X	31	32	32	50.00	50.00	50.00
	Y	31	32	32	50.00	50.00	50.00
CSF1PO	7	0	5	0	0.00	7.81	0.00
	8	0	4	0	0.00	6.25	0.00
	9	1	2	1	1.61	3.13	1.56
	10	11	15	19	17.74	23.44	29.69
	11	16	17	15	25.81	26.56	23.44
	12	29	17	24	46.77	26.56	37.50
	13	5	4	4	8.06	6.25	6.25
	14	0	0	1	0.00	0.00	1.56
D13S317	8	8	2	6	12.90	3.13	9.38
	9	3	1	13	4.84	1.56	20.31
	9 (A->T)	0	0	1	0.00	0.00	1.56
	10	0	0	2	0.00	0.00	3.13
	10 (A->T)	1	1	5	1.61	1.56	7.81
	11	13	8	11	20.97	12.50	17.19
	11 (A->T)	10	10	4	16.13	15.63	6.25
	12	6	26	9	9.68	40.63	14.06
	12 (A->T)	7	7	2	11.29	10.94	3.13
	13	11	9	5	17.74	14.06	7.81
	13 (A->T)	0	0	2	0.00	0.00	3.13
	14	3	0	3	4.84	0.00	4.69
	15 (A->T)	0	0	1	0.00	0.00	1.56
D16S539	8	1	2	0	1.61	3.13	0.00
	9	10	11	7	16.13	17.19	10.94
	9 (A->G)	1	0	0	1.61	0.00	0.00
	10	2	6	7	3.23	9.38	10.94
	11	18	24	16	29.03	37.50	25.00
	12	21	10	18	33.87	15.63	28.13
	13	9	11	15	14.52	17.19	23.44
	14	0	0	1	0.00	0.00	1.56
D18S51	10	1	0	0	1.61	0.00	0.00
	11	0	0	1	0.00	0.00	1.56
	12	8	7	6	12.90	10.94	9.38
	13	9	4	9	14.52	6.25	14.06
	13.2 (T->C)	0	2	0	0.00	3.13	0.00
	14	7	3	9	11.29	4.69	14.06
	15	5	9	13	8.06	14.06	20.31
	15 (T->G)	1	0	0	1.61	0.00	0.00
	16	12	12	7	19.35	18.75	10.94
	17	9	14	8	14.52	21.88	12.50
	18	5	5	4	8.06	7.81	6.25
	19	2	2	3	3.23	3.13	4.69
	20	2	0	2	3.23	0.00	3.13
	20 (T->C)	0	4	0	0.00	6.25	0.00
	21	0	0	1	0.00	0.00	1.56
	22	1	1	1	1.61	1.56	1.56
	23	0	1	0	0.00	1.56	0.00

		Count			Percentage		
Locus	Allele	Caucasian	African American	Hispanic	Caucasian	African American	Hispanic
D21S11	25.2	1	0	0	1.61	0.00	0.00
	27	0	2	1	0.00	3.13	1.56
	27 (A->G)	2	2	0	3.23	3.13	0.00
	28	14	12	5	22.58	18.75	7.81
	29	7	10	16	11.29	15.63	25.00
	29 (G->A)	3	0	3	4.84	0.00	4.69
	29 (A->G)	0	3	0	0.00	4.69	0.00
	29.2 (A->G)	1	0	0	1.61	0.00	0.00
	30	11	3	7	17.74	4.69	10.94
	30 (G->A)	5	12	9	8.06	18.75	14.06
	30.2	2	1	4	3.23	1.56	6.25
	30.2 (G->A)	1	0	0	1.61	0.00	0.00
	31	2	4	2	3.23	6.25	3.13
	31 (G->A)	3	1	2	4.84	1.56	3.13
	31.2	3	3	6	4.84	4.69	9.38
	31.2 (G->A)	0	0	1	0.00	0.00	1.56
	32 (A->G)	0	1	0	0.00	1.56	0.00
	32.2	6	6	5	9.68	9.38	7.81
	32.2 (G->A)	0	0	2	0.00	0.00	3.13
	33.2	1	1	0	1.61	1.56	0.00
	34	0	0	1	0.00	0.00	1.56
	34.1 (T->A)	0	1	0	0.00	1.56	0.00
	36 (2A->2G)	0	1	0	0.00	1.56	0.00
	37 (3A->3G)	0	1	0	0.00	1.56	0.00
D3S1358	13	0	0	1	0.00	0.00	1.56
	13 (G->A)	0	1	0	0.00	1.56	0.00
	14	4	2	5	6.45	3.13	7.81
	14 (G->A)	0	0	2	0.00	0.00	3.13
	15	2	1	1	3.23	1.56	1.56
	15 (G->A)	13	8	17	20.97	12.50	26.56
	15 (2G->2A)	2	12	4	3.23	18.75	6.25
	15.2	0	1	0	0.00	1.56	0.00
	16	5	3	5	8.06	4.69	7.81
	16 (G->A)	8	6	5	12.90	9.38	7.81
	16 (2G->2A)	0	10	2	0.00	15.63	3.13
	17	7	4	7	11.29	6.25	10.94
	17 (A->G)	0	1	1	0.00	1.56	1.56
	17 (2G->2A)	1	2	0	1.61	3.13	0.00
	17 (G->A)	8	6	5	12.90	9.38	7.81
	18	12	4	8	19.35	6.25	12.50
	18 (G->A)	0	2	1	0.00	3.13	1.56
	18 (2G->2A)	0	1	0	0.00	1.56	0.00

Table 11, part 2. Observed frequency of each allele by population in 31 Caucasian, 32 African American and 32 Hispanic samples from NIST.

		Count			Percentage		
Locus	Allele	Caucasian	African American	Hispanic	Caucasian	African American	Hispanic
D5S818	7	0	0	5	0.00	0.00	7.81
	8	1	0	0	1.61	0.00	0.00
	8 (G->T)	0	3	0	0.00	4.69	0.00
	9 (G->T)	1	3	2	1.61	4.69	3.13
	10	5	4	3	8.06	6.25	4.69
	10 (G->T)	0	0	1	0.00	0.00	1.56
	11	25	13	25	40.32	20.31	39.06
	11 (G->T)	4	2	0	6.45	3.13	0.00
	12	15	22	15	24.19	34.38	23.44
	12 (G->T)	4	7	3	6.45	10.94	4.69
	13	5	5	5	8.06	7.81	7.81
	13 (G->T)	1	3	0	1.61	4.69	0.00
	13 (G->C)	0	0	1	0.00	0.00	1.56
	14	1	1	2	1.61	1.56	3.13
	14 (G->T)	0	1	1	0.00	1.56	1.56
	15	0	0	1	0.00	0.00	1.56
D7S820	7	2	0	0	3.23	0.00	0.00
	8	10	13	8	16.13	20.31	12.50
	9	9	10	7	14.52	15.63	10.94
	10	13	25	14	20.97	39.06	21.88
	10 (T->A)	2	0	3	3.23	0.00	4.69
	11	10	12	19	16.13	18.75	29.69
	11 (T->A)	3	0	0	4.84	0.00	0.00
	12	4	3	7	6.45	4.69	10.94
	12 (T->A)	6	0	3	9.68	0.00	4.69
	13	3	1	2	4.84	1.56	3.13
	13 (T->A)	0	0	1	0.00	0.00	1.56
D8S1179	8	1	0	0	1.61	0.00	0.00
	10	5	1	6	8.06	1.56	9.38
	11	7	1	5	11.29	1.56	7.81
	12	10	1	8	16.13	1.56	12.50
	12 (A->G)	0	5	2	0.00	7.81	3.13
	13	19	17	14	30.65	26.56	21.88
	13 (G->A)	3	3	3	4.84	4.69	4.69
	13 (C->G)	0	0	1	0.00	0.00	1.56
	14	6	23	12	9.68	35.94	18.75
	14 (G->A)	6	1	2	9.68	1.56	3.13
	15	3	7	7	4.84	10.94	10.94
	15 (A->G)	0	1	0	0.00	1.56	0.00
	16	2	3	2	3.23	4.69	3.13
	16 (A->G)	0	1	0	0.00	1.56	0.00
	17 (G->A)	0	0	1	0.00	0.00	1.56
	18	0	0	1	0.00	0.00	1.56

		Count			Percentage		
Locus	Allele	Caucasian	African American	Hispanic	Caucasian	African American	Hispanic
FGA	18	2	0	2	3.23	0.00	3.13
	19	3	5	5	4.84	7.81	7.81
	19.2	0	1	0	0.00	1.56	0.00
	20	7	3	4	11.29	4.69	6.25
	21	8	2	11	12.90	3.13	17.19
	22	13	12	9	20.97	18.75	14.06
	22.2	1	0	0	1.61	0.00	0.00
	23	11	11	14	17.74	17.19	21.88
	23.2	1	1	0	1.61	1.56	0.00
	24	9	11	9	14.52	17.19	14.06
	25	5	9	4	8.06	14.06	6.25
	26	2	5	5	3.23	7.81	7.81
	26 (G->A)	0	1	0	0.00	1.56	0.00
	27	0	1	1	0.00	1.56	1.56
	28 (T->C)	0	1	0	0.00	1.56	0.00
	31.2	0	1	0	0.00	1.56	0.00
THO1	6	15	6	10	24.19	9.38	15.63
	7	9	33	25	14.52	51.56	39.06
	8	3	14	8	4.84	21.88	12.50
	9	10	7	6	16.13	10.94	9.38
	9.3	25	4	15	40.32	6.25	23.44
TPOX	6	1	8	0	1.61	12.50	0.00
	7	0	1	0	0.00	1.56	0.00
	8	33	21	29	53.23	32.81	45.31
	9	10	12	5	16.13	18.75	7.81
	10	2	7	2	3.23	10.94	3.13
	11	15	14	18	24.19	21.88	28.13
	12	1	1	10	1.61	1.56	15.63
vWA	13 (G->A + C->T)	1	0	0	1.61	0.00	0.00
	13 (C->T)	0	1	0	0.00	1.56	0.00
	14 (T->C)	0	2	0	0.00	3.13	0.00
	14 (G->A + T->C)	3	0	0	4.84	0.00	0.00
	14 (A->G + 2T->2C)	6	2	3	9.68	3.13	4.69
	15	0	5	7	0.00	7.81	10.94
	15 (G->A)	7	1	7	11.29	1.56	10.94
	16	9	11	13	14.52	17.19	20.31
	16 (G->A)	2	7	3	3.23	10.94	4.69
	17	12	13	14	19.35	20.31	21.88
	17 (G->A)	2	1	3	3.23	1.56	4.69
	18	11	16	8	17.74	25.00	12.50
	18 (G->A)	1	0	0	1.61	0.00	0.00
	18 (2A->2G)	0	1	0	0.00	1.56	0.00
	19	7	2	4	11.29	3.13	6.25
	19 (2A->2G)	0	1	0	0.00	1.56	0.00
	20	0	0	2	0.00	0.00	3.13
	20 (T->A)	1	0	0	1.61	0.00	0.00
	20 (A->G)	0	1	0	0.00	1.56	0.00

Table 12. Frequency of observation of polymorphic alleles in 95 samples typed at 13 CODIS loci (Table 11) and 47 samples typed at 9 CODIS loci (Table 2).

Locus	Number of Alleles with SNPs	Samples tested	% of alleles with one or more SNPs	Number of Same-Length Heterozygous Loci	% of samples heterozygous with same-length alleles
D3S1358	**115**	**95**	**60.5**	**10**	**10.5**
vWA	**81**	**142**	**28.5**	**11**	**7.7**
D13S317	**77**	**142**	**27.1**	**7**	**4.9**
D21S11	**50**	**95**	**26.3**	**6**	**6.3**
D5S818	**59**	**142**	**20.8**	**10**	**7.0**
D8S1179	**35**	**142**	**12.3**	**7**	**4.9**
D7S820	**30**	**142**	**10.6**	**2**	**1.4**
D18S51	**7**	**95**	**3.7**	**0**	**0.0**
FGA	**2**	**95**	**1.1**	**0**	**0.0**
D16S539	**1**	**142**	**0.4**	**0**	**0.0**

A subset of alleles from the NIST samples was sequenced for verification. Alleles that were sequenced are shown in red in Table 13. The majority of these were done as base allele – variant allele pairs to verify confirm the SNP(s) and to identify the relative location. Allele sequencing confirmed the presence of a sequence polymorphism relative to the reference allele in each case and allowed the identification of its position (Table 14).

Table 13. Selected alleles that were sequenced from NIST samples for verification of polymorphisms observed in the ESI-TOF-MS assay.

Sample	Well	Locus	Alleles	PP
NIST-PT84223	A07	D3S1358	**17, 17 (2G->2A)**	3397
NIST-PT84232	A08	D3S1358	**15 (2G->2A), 16 (2G->2A)**	3397
NIST-JT52076	C10	D3S1358	**15 (G->A), 16**	3397
NIST-OT07280	D10	D3S1358	**15, 16 (G->A)**	3397
NIST-GT37864	F09	D7S820	9, **12 (T->A)**	2817
NIST-MT97172	G03	D7S820	9, **12**	2817
NIST-JT51471	A05	D8S1179	**12 (A->G)**, **13**	2818
NIST-WT51343	C02	D8S1179	**12, 13 (G->A)**	2818
NIST-TT51399	G10	D8S1179	13, **13 (C->G)**	2818
NIST-PT84232	A08	D18S51	14, **20 (T->C)**	3394
NIST-WT51378	C03	D18S51	12, **15 (T->G)**	3394
NIST-OT05899	C06	D18S51	**15**, 16	3394
NIST-PT84223	A07	D21S11	**29 (A->G), 29**	3390
NIST-OT07280	D10	D21S11	**30 (T->C), 30**	3390

Table 14. Repeat sequence structures determined for sequenced STR alleles.

Sample	Locus	Allele	Repeat region sequence	Reference
NIST-OT07280	D3S1358	15	TCTA $(TCTG)_3$ $(TCTA)_{11}$	
NIST-JT52076	D3S1358	15 (G->A)	TCTA $(TCTG)_2$ $(TCTA)_{12}$	
NIST-PT84232	D3S1358	15 (2G->2A)	TCTA TCTG $(TCTA)_{13}$	
NIST-JT52076	D3S1358	16	TCTA $(TCTG)_3$ $(TCTA)_{12}$	13
NIST-OT07280	D3S1358	16 (G->A)	TCTA $(TCTG)_2$ $(TCTA)_{13}$	
NIST-PT84232	D3S1358	16 (2G->2A)	TCTA TCTG $(TCTA)_{14}$	
NIST-PT84223	D3S1358	17	TCTA $(TCTG)_3$ $(TCTA)_{13}$	
NIST-PT84223	D3S1358	17 (2G->2A)	TCTA TCTG $(TCTA)_{15}$	
NIST-MT97172	D7S820	12	$(GATA)_{12}$ GACAGATTGATAGTTTTTTTTT	14
NIST-GT37864	D7S820	12 (T->A)*	$(GATA)_{12}$ GACAGATTGATAGTTTTTTTT**A**	
NIST-WT51343	D8S1179	12	$(TCTA)_{12}$	
NIST-JT51471	D8S1179	12 (A->G)	$(TCTA)_2$ TCTG $(TCTA)_9$	15
NIST-JT51471	D8S1179	13	TCTA TCTG $(TCTA)_{11}$	
NIST-WT51343	D8S1179	13 (G->A)	$(TCTA)_{13}$	
NIST-TT51399	D8S1179	13 (C->G)	TCTA TCTG TGTA $(TCTA)_{10}$	
NIST-OT05899	D18S51¥	15	$(TTCT)_{15}$	
NIST-WT51378	D18S51¥	15 (T->G)	$(TTCT)_{13}$ TGCT TTCT	15
Not sequenced	D18S51¥	20	$(TTCT)_{20}$	
NIST-PT84232	D18S51¥	20 (T->C)	$(TTCT)_3$ TTCC $(TTCT)_{16}$	
NIST-PT84223	D21S11	29	$(TCTA)_4$ $(TCTG)_6$ $(TCTA)_3$ TA $(TCTA)_3$ TCA $(TCTA)_2$ TCCA TA $(TCTA)_{11}$	
NIST-PT84223	D21S11	29 (A->G)	$(TCTA)_4$ $(TCTG)_7$ $(TCTA)_3$ TA $(TCTA)_3$ TCA $(TCTA)_2$ TCCA TA $(TCTA)_{10}$	16
NIST-OT07280	D21S11	30	$(TCTA)_6$ $(TCTG)_5$ $(TCTA)_3$ TA $(TCTA)_3$ TCA $(TCTA)_2$ TCCA TA $(TCTA)_{11}$	
NIST-OT07280	D21S11	30 (T->C)	$(TCTA)_6$ $(TCTG)_5$ CCTA $(TCTA)_2$ TA $(TCTA)_3$ TCA $(TCTA)_2$ TCCA TA $(TCTA)_{11}$	

¥ The T→A SNP observed in allele 12 for locus D7S820 did not occur in the repeat region, but was one base in from the 3′ end of the reverse primer.

¥ The reference strand used for D18S51 was based upon the reported sequence for GenBank accession AP001534.2. Where allele 20 is reported as (AGAA)20 on STRBase[2], we have reported it as $(TTCT)_{20}$. What we are reporting as a T→C SNP (for allele 20) would be an A→G SNP on the complementary strand. Likewise, what we report as a T→G SNP (for allele 15) would be an A→C SNP on the complementary strand.

S.1.6. Mixture studies. Mixture analysis with the Ibis STR assay has not been addressed at this point.

S.1.7 PCR-based procedure-specific validation, STR assay

S.1.7.1 Potential for differential amplification among loci.

Inter-allelic product ratios for 67 runs of the Ibis STR assay with three different samples are shown in Figure 21, panels A-E. The observed ratios for 1 ng, 333 pg, 111 pg, 37 pg and 12 pg template inputs per reaction for five different templates run in duplicate are shown in Figure 21. All ratios are expressed in terms of the higher signal allele to the lower signal allele for each locus. For homozygous loci, the ratio is artificially fixed at 1.0. At 1 ng template input per reaction, all ratios between alleles within a given locus are less than 1.5. At levels of 111 pg and below, the ratios become quite variable and show substantial sample dependence, with some ratios approaching 2:1 (Figure 21, panels C and D). At 12 pg input (Figure 21, panel E.), too many alleles and entire loci drop out to allow useful interpretation of product ratios. STR typing at levels of 37 pg and below with this assay is not reliable (see section S.1.2 – Sensitivity). The inter-allelic balance ratios within the 14 loci used in the STR assay suggest that at levels of 333 pg and above for a major contributor, a minor contributor present at below 2/3 the level of the major contributor should be identifiable as the minor contributor (provided that the minor contributor is above the detectable limits of the assay - ca. 40 to 100 pg). At contributor ratios more even than 1.5, allelic signal output ratios will begin to overlap with normal ratios for a pure template and it will be difficult to tell which alleles belong to each contributor.

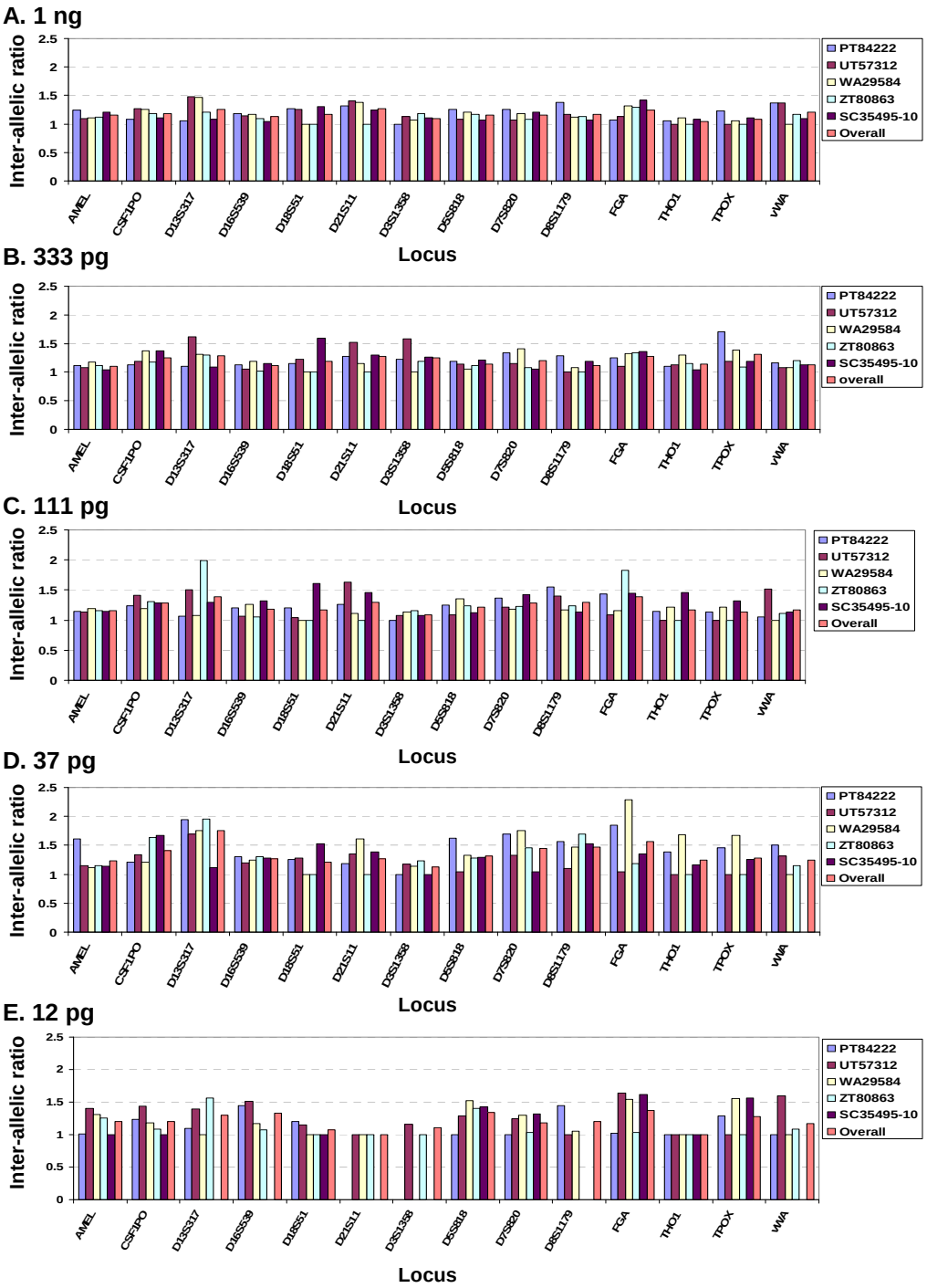


Figure 21. Inter-allelic balance for STR loci over five templates at five template input levels. At 1 ng input per reaction (A.), the observed inter-allelic output ratios (expressed as higher signal-to-lower signal in all cases) are all below 1.5. This is consistent with the observation for 3 templates run 67 times in section 1.3 (see Figure 1.3.1, panel F.). At levels of 111 pg and below, substantial template / allelic combination dependence becomes apparent (panels C and D) and some ratios approach 2.0. At levels of 37 pg and below (panels D. and E.), several alleles and full loci begin to drop out entirely. Interpretation of allelic balance becomes problematic or impossible at 12 pg (panel E.) due to sporadic representation of products. For each template, reactions were run in duplicate and results shown are the average of two replicates. Thus, no standard deviations are shown.

S.1.7.2. Effects of coamplification of loci.

STR '-A' artifacts

Outputs from the 67 reproducibility runs for section S.1.3 were examined for the presence of '-A' artifacts (the presence of product peaks corresponding to incomplete 3´ adenylation of PCR products). Out of 4,408 product assignments, peaks with sufficient signal to pass our current processing thresholds that corresponded to non-adenylated products occurred 133 times (3%) with an average relative intensity of 21% of the combined signal of the adenylated and non-adenylated product (Figure 22). Note that the lowest percentage detected in this analysis was 9.7%, which is consistent with the observed dynamic range we have observed for our processing with mixtures of mitochondrial data (~10%). Non-adenylated artifacts were most prevalent in FGA, and were not observed in 6 of the 14 loci (Figure 22).

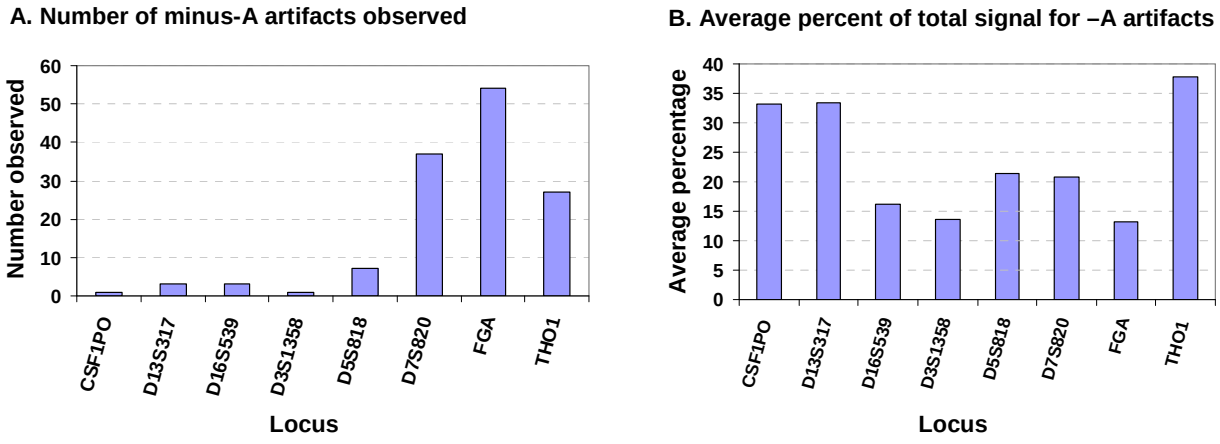


Figure 22. Non-adenylated PCR artifacts from 67 reproducibility assay runs. Reproducibility data from sections S.1.3 was examined for the presence of automatically-assigned mass peaks corresponding to non-adenylated PCR products. A.) The number of –A products observed for each locus. Non-adenylated artifacts were not observed for AMEL, D18S51, D21S11, D8S1179, TPOX or vWA. B. The average percent of total product signal accounted for by non-adenylated products when they were detected by our automated mass-calling software.

Specific Aim 1. Development and validation: Mitochondrial control region tiling assay

The following section of this report addresses development of the Ibis mitochondrial control region tiling assay and subsections are prepended with an M. The STR assay is discussed in a preceding section and subsections therein are prepended with an S.

M.1.1 Species specificity, Mitochondrial control region tiling assay

For mitochondrial casework, the evidence samples are usually hair or bone which are considered to be sole source samples. Thus, in mitochondrial work, one would typically not be examining mixtures. To examine the species specificity of the Mitochondrial tiling assay, buccal swabs from a dog and a cat were collected as well as hair samples from a goat. The samples were isolated using modified Qiagen protocols for buccal swabs and a modified version of the FBI protocol for hair extraction. As a control for the goat hair, a human hair was isolated in parallel to show that the extraction method worked. We do not have any primers to verify that there is goat DNA and assumed that if the human hair gave the full correct profile the extraction method was working correctly. As mentioned in the STR section (S.1.1), the isolation of the dog and cat DNA was verified by assaying the samples with primers designed to amplify dog and cat DNA. Both the goat hair extract and the human hair extract were diluted from a volume of 20 μL to 50 μL. To each well of the mitochondrial tiling assay, 5 μL of the diluted extract was added. The goat DNA produced no amplified products in the mitochondrial assay while the human hair extract produced a complete correct profile. For the dog and cat extracts, 2 μL of extract were diluted to 50 μL since we typically perform a 1:25 dilution of this sample type. To each well of the assay, 5 μL of the diluted extract was added. This is equivalent to using approximately 1 to 2 ng of the animal DNA per reaction. Both the dog and cat extracts produce amplified products in the mitochondrial tiling assay for triplexes A, C, D and E. However, the products were not consistent with human products for those primers and were outside the typical mass range for the primer pair most likely to amplify the DNA (Figure 23). Thus, the assay would not

identify dog or cat DNA as human and would not give a human profile for a dog or a cat sample.

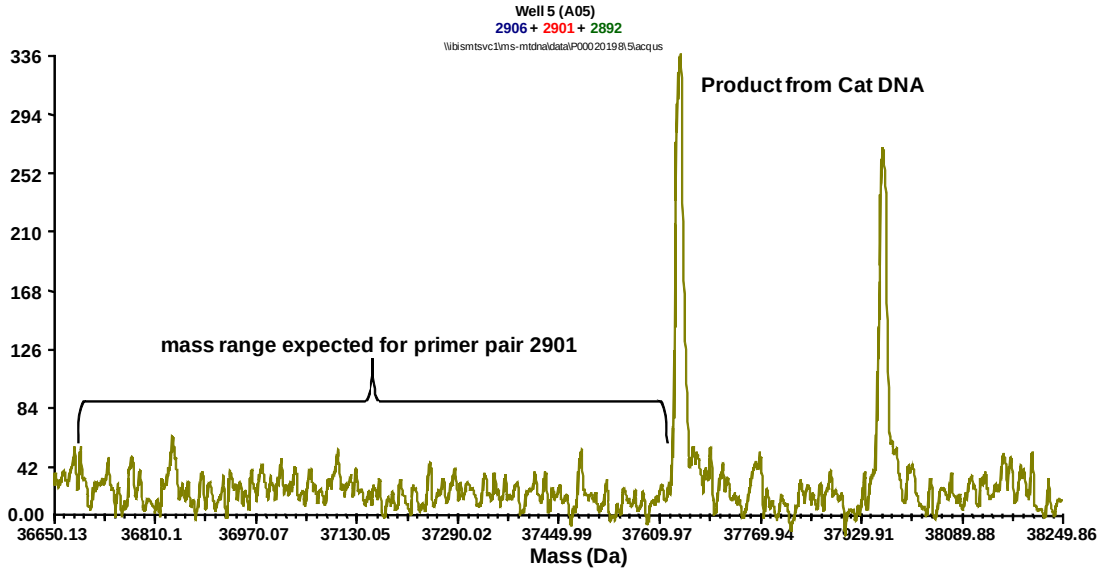


Figure 23. A deconvolved mass spectrum showing the PCR product from cat DNA in triplex A of the mitochondrial assay. The masses for the two strands of the PCR product do not overlap with the expected mass range for primer pair 2901.

M.1.2 Sensitivity, Mitochondrial control region tiling assay

The sensitivity of the Mitochondrial Tiling Assay was tested in two phases. The sensitivity of the primer pairs was tested individually and in the appropriate triplexes to make sure that multiplexing of the primer pairs did not decrease the sensitivity of the assay. Initially, the sensitivity of the individual primer pairs was tested against one template. The starting concentration was 500 pg/well and 5 fold serial dilutions were made down to 0.8 pg/well. The template was DNA derived from Seracare blood sample SC35495 and was quantified using Applied Biosystems Quantifiler Human DNA Quantification Kit. In addition, this template was selected for this study since it does not show C-length or SNP heteroplasmies for any of the primer pairs. The base composition profile for this template was confirmed by sequencing the HV1, HV2 and HV3 mitochondrial regions. The cycling conditions for PCR are shown in Figure 24, with the exception that for single-plex reactions, the annealing time was decreased to 30 seconds and the extension time was increased to 30 seconds. Thirty-six cycles was used to minimize potential amplification of background due to the sensitivity of mitochondrial PCR. The masses of the PCR products from the reactions were measured using ESI-TOF mass spectrometry and the masses were converted to unambiguous base compositions. The base compositions were compared to the base composition expected from the sequence of the region amplified by the primer pairs. For all primer pairs tested in single-plex, the correct product was assigned down to 4 pg of input template, and for 8 primer pairs the sensitivity went down to 0.8 pg of template (Figure 25).

Figure 24. PCR cycling conditions for the mitochondrial control region tiling assay

1. 96°C, 10 min
2. 96°C, 20 secs
3. 50°C, 90 secs
4. 72°C, 5 secs
5. Go to step 2, 35 times
6. 72°C, 4 min
7. 4°C, hold

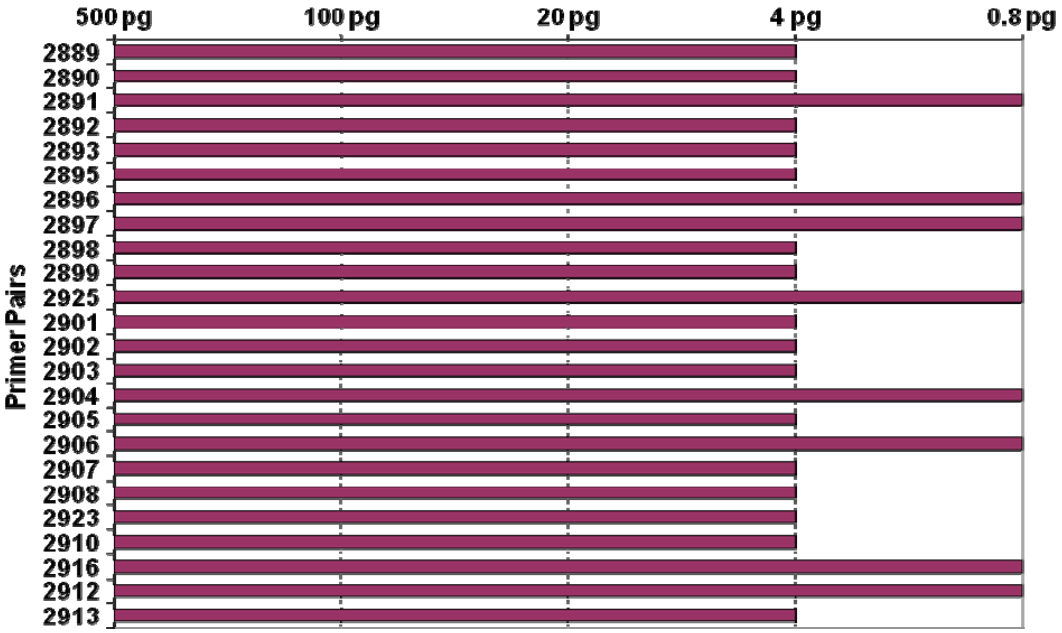


Figure 25. Sensitivity of primer pairs when screened individually is depicted in the above graph. The correct product is automatically detected down to at least 4 pg for all primer pairs.

The sensitivity study was repeated using 5 templates and using the primer pairs in the final assay format triplexes. Three templates were selected from the blinded samples sent by AFDIL (AF-6, AF-18 and AF-25) and two templates were selected from the blinded samples sent by the FBI (FBI-58 and FBI-67). The AFDIL samples were blood punches isolated using an appropriate Qiagen protocol. The FBI samples were buccal swabs isolated using an appropriate Qiagen protocol. Because two different samples types were chosen, the sensitivity study also examined the effect of DNA isolated from different sample types. All of the templates chosen exhibited no C-length or SNP heteroplasmy. All of the templates were quantified using Applied Biosystems Quantifiler Human DNA Quantification Kit. The template concentrations were examined over the range 100 pg to 0.098 pg per well. Two-fold serial dilutions were made of each template. The reactions were thermocycled using the conditions listed in Figure 24. The lowest concentration at which the correct product was automatically detected for each primer pair and the lowest concentration at which a full correct profile was automatically detected for each sample were examined (Figure 26).

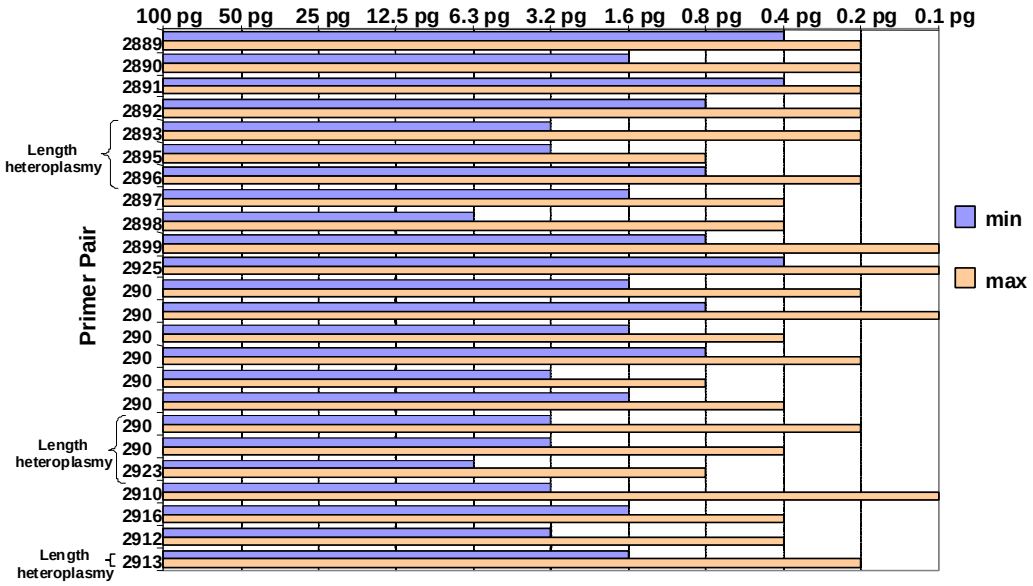


Figure 26. Sensitivity of each mtDNA tiling primer pair in the multiplex assay. For each primer pair, the minimum and maximum template dilution at which the correct product is automatically assigned is shown. The sensitivity of a primer pair is not affected by the presence of other primer pairs in the triplex.

Figure 26 shows the minimum and maximum template dilution for which the correct product was automatically detected for each primer pair. For primer pairs where C-length heteroplasmies are expected, there is a greater difference between the minimum and maximum template dilution where correct products are automatically assigned. In all cases, the primer pair sensitivity in the corresponding triplex is comparable to that observed for the primer pairs screened individually. Table 15 shows the lowest dilution at which the correct product was automatically detected for a primer pair as a function of template and the lowest dilution at which the full correct profile was automatically detected. The individual primer pair

Table 15. Primer pair sensitivity and full profile detection sensitivity as function of template.

Primer Pair	Amplified coordinates	Template				
		AF-6	AF-18	AF-25	FBI-58	FBI-67
2889	16377..16428	0.195	0.391	0.391	0.391	0.195
2890	16342..16381	0.195	0.391	1.563	0.391	0.195
2891	16283..16344	0.391	0.195	0.391	0.391	0.195
2892	16254..16305	0.391	0.391	0.781	0.195	0.781
2893	16182..16250	0.195	3.125	0.781	0.195	0.195
2895	16157..16201	0.781	0.781	3.125	0.781	0.781
2896	16124..16201	0.195	0.195	0.781	0.391	0.391
2897	16078..16129	0.391	1.563	1.563	0.781	0.391
2898	16048..16098	0.391	0.781	0.781	3.125	6.250
2899	16015..16051	0.781	0.391	0.391	0.098	0.098
2925	15963..16017	0.391	0.195	0.391	0.391	0.098
2901	15924..15985	0.391	0.391	0.195	1.563	0.781
2902	31..76	0.391	0.195	0.781	0.195	0.098
2903	41..114	0.391	1.563	1.563	0.391	0.391
2904	103..162	0.195	0.195	0.781	0.391	0.195
2905	138..217	1.563	3.125	3.125	0.781	0.781
2906	178..267	0.781	0.391	0.781	1.563	0.781
2907	263..340	0.781	0.391	0.781	0.195	3.125
2908	234..314	0.391	0.391	3.125	0.391	0.781
2923	289..371	3.125	0.781	0.781	3.125	6.250
2910	355..402	0.195	3.125	0.781	0.391	0.098
2916	389..437	0.391	1.563	1.563	0.781	0.391
2912	431..501	0.391	1.563	3.125	0.781	0.781
2913	493..576	0.195	0.391	1.563	0.391	0.195
complete profile		3.125	3.125	3.125	3.125	6.250

sensitivity varied depending on the template tested. For the 5 templates tested, a full correct profile was observed down to at least 6.25 pg/well of total DNA. The source of the DNA did not affect the ability to detect correct full profiles for the samples. With these data, and the addition of >60 dilution trials of our standard positive control run in a routine system validation format, we have set a conservative limit of sensitivity of 25 pg/reaction for the mtDNA assay. This is 4X the limit observed in this trial and is defined as the sensitivity that must be achieved for a system validation plate to pass our sensitivity criteria.

M.1.3. Reproducibility, precision and accuracy, Mitochondrial control region tiling assay

For an initial reproducibility study, samples AF-25 (no length heteroplasmy) and sample FBI-75 (length heteroplasmy) were selected. The sample FBI-75 displayed length heteroplasmy in three locations: the HV1 and HV2 poly-C tracts, and in the region covered by primer pair 2913 (amplified coordinates 493 to 576). The latter is presumably in the poly-C tract located at 568-573. For FBI-75, three to four products were observed for each primer pair spanning the HV1 poly-C stretch (primer pairs 2893, 2895 and 2896), and three products were observed in primer pair 2913. In addition, two length variants were observed for each primer pair spanning the HV2 poly-C stretch (2907, 2908 and 2923). These length variant observations were very consistent across the 55 replicates. The reproducibility study was performed with 100 pg/well of input DNA. This amount of template is well above the reliable range determined in section M.1.2. Using premade PCR plates, 55 replicates of each sample were analyzed as well as 5 replicate no-template added controls to show that there was no contamination. Five microliters of template was added manually to each well of a plate using a multichannel pipette. One measure of reproducibility is how many primer pairs do not produce a PCR product; this

Table 16. Primer pair reproducibility in Mitochondrial tiling assay.

Plate number	Number of primer pairs missed		Profile match percentage
	AF-25	FBI-75	
1	8	5	100
2	6	0	100
3	0	0	100
4	3	10	100
5	0	0	100
Total	17/1320	15/1320	
% missed	1.3%	1.2%	

indicates how reproducibly the template was added to the plate (Table 16). If only primer masses are observed in the mass spectrum, the most likely explanation is that no or insufficient template was added to the well to observed PCR. In most cases, when no PCR products were observed for primer pairs, all of the primer pairs were from the same triplex. When a full or partial profile was observed, it was correct and there was no indication of contamination. For each template, 1.2 to 1.3% of the primer pairs were missed. The error rate did not depend on the template. Currently, liquid handling robots are used for template addition which has lowered the error rate.

The reproducibility of Seracare sample SC35495, which is currently the positive control for this assay, was examined over a 1.5 year period (August 2006 to December 2007) with 2,207 replicates. In this set, 2,130 (97%) of the trials yielded a full, correct profile and 73 (3%) were missing one or more primer pair (generally due to a plugged injector or marginal cleanup). In all, there were 159 missed calls out of 52,968 expected (0.30%). There were 19 mis-assignments in the automated first pass (0.035%) which are corrected during manual inspection. These were generally associated with wells flagged as failures that contained low signal and salt adducts. There were 92 artifact assignments (0.17%) which are also corrected during manual inspection. To asses the precision and accuracy of profile assignment and mass measurements, the observed mass measurement deviation in ppm was plotted versus the frequency of occurrence. As is shown in Figure 27, the plot is centered around 0 ppm mass measurement deviation and indicates that there is not a bias in the mass measurement uncertainty. The average measurement deviation magnitude for 105,580 measurements was 10.64 $\pm$ 8.2 ppm.

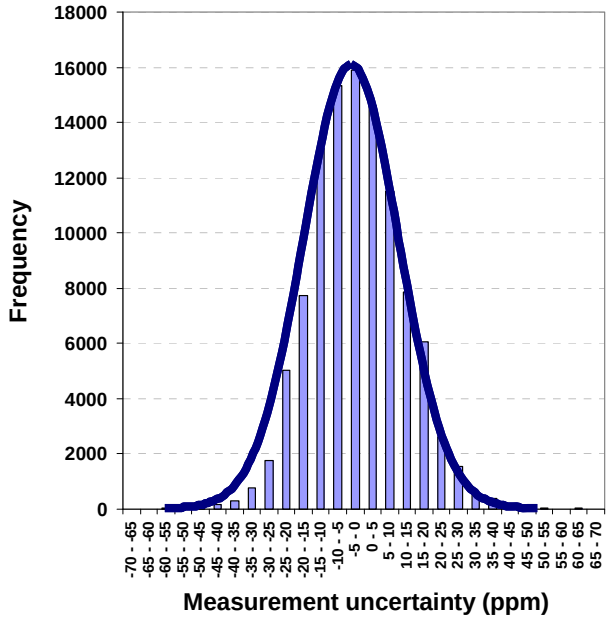


Figure 27. Mass measurement precision for mitochondrial control region tiling assay. 105,580 independent, correct mass-product assignments were assessed over 2,207 trials of the same template run through the mtDNA tiling assay. Observed mass measurement deviations in parts per million were binned in 5 ppm increments and plotted by frequency of occurrence. Observed measurement deviations are normally distributed around zero and there does not appear to be a strong systematic bias toward measurement errors lower or higher than the true product masses. 99% of the measurements were within -35 to 35 ppm of the predicted product mass. 35 ppm corresponds to ~1 Da (one proton mass) for 30 kDa DNA molecule (about 100 nucleotides), or about 1.6 Da for our largest known mtDNA products (150 bp, or about 46.5 kDa). All but three of our primer pairs generate products less than 130 bp in length. The average measurement deviation magnitude for 105,580 measurements was 10.64 $\pm$ 8.2 ppm.

M.1.4. Reference samples: Mitochondrial control region tiling analysis of NIST samples.

The same 95 samples from NIST that were typed with the STR assay were also analyzed with the mtDNA tiling assay. The samples were analyzed blindly and sequence data were received from Peter Vallone for comparison to our base composition typing results after typing was performed. All samples gave complete base composition profiles. Ten samples were repeated internally for QC / verification before communicating results to NIST. Sequence data were received as a series of text file exports from the Sequencher program (Gencodes, Inc.). The sequence profiles were imported into the Ibis database and converted into stored base composition profiles by an in-house automated procedure incorporated in the Ibis mtDNA profiling system.

All NIST samples were run at 500 pg per reaction in the standard assay layout shown in Figure 5, panel A using the primer pairs and triplex groupings from Table 5. Reactions are formulated in prefabricated kit plates with 35 μl per well of 1.14X master mix containing the target concentrations (for a 40 μl final volume) of 10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, 400 mM betaine, 5 U Amplitaq Gold (ABI), and the primer pair concentrations shown in Table 5, panel B for each triplex. To these pre-made reaction mixes, 5 μl of appropriately-diluted DNA sample (100 pg/μl in this case) were added prior to thermocycling. Thermocycling consisted of 96 ºC, 10 min, [96 ºC, 20 sec., 50 ºC, 1.5 min., 72 ºC, 5 sec] (36 cycles), 72 ºC, 4 min., 4 ºC hold. After thermocycling, samples were desalted and subjected to ESI-TOF-MS on an automated Ibis T5000 as described[4]. An automated in-house processing system deconvolved the raw mass spectrometry data and imported mass assignment results into the Ibis mtDNA profiling database system. In-house analysis software was then used to assign and register base composition profiles into the Ibis mtDNA database.

Available sequence data were consistent with the base compositions we obtained in all cases for the NIST reference sample set (data not shown). There were sequence data for 94 of the 95 samples for HV1 coordinates 16024-16569 and HV2 coordinates 1-577. This allowed comparison of base composition profiles to sequences for 21 out of our 24 regions for 94 out of the 95 samples analyzed. Several sequence profiles contained IUPAC codes (ambiguous characters). In all cases these were observed in the base composition profiles as either a base composition consistent with one of the possible identities for the IUPAC code or, more often, as a mixture of the two IUPAC possibilities (heteroplasmy). Our system readily identifies both SNP and length heteroplasmy and is capable of quantification of the relative contribution of major and minor contributors. However, the current data processing algorithms occasionally do not automatically identify both peaks of an asymmetric mixture of two DNA strands differing by an A←→G, C←→T or A←→T SNP. In these cases, manual data interrogation tools built into our interface software are relied upon to correctly assign both components.

We revised our technology transfer plan to emulate case work samples considerably from the original proposal. The original strategy we contemplated to initiate technology transfer to other practitioners in the forensics community was to consist of Ibis sending assay kits and SOPs to collaborators at AFDIL and/or the FBI with forensics scientists at these sites generating STR and/or mtDNA amplicons to be sent back to Ibis for mass spectrometric analysis and interpretation of results. As the FBI, through independent funding, was able to bring the T5000 hardware platform in house (DNA Unit II – see details in Specific Aim 3), we opted for a more "hands-on" technology transfer approach aimed at supporting FBI scientists in running kitted assays all the way through the process. Based on the maturity of the assays and the focus of DNA Unit II, we focused efforts on transfer and implementation of the mtDNA assay to the FBI in support of Specific Aim 1.4.

During the course of this project we supplied the FBI with approximately 70 assay plates, enough to perform 700 assays at ten assays/plate (with one positive and one negative control/plate). Some of these plates were used for training and/or "practice", while the majority of the kits were used to support internal validation efforts and collaborative validation studies with Ibis. Ibis scientists transferred detailed SOPs for assay plate setup, PCR amplification, T5000 operation, and data analysis to FBI researchers. In addition to document and reagent transfer, several highly interactive face-to-face meetings were held both at Ibis headquarters in Carlsbad, CA, and at the FBI labs in Quantico, VA. The mtDNA mixture studies described in section M.1.6 were acquired by FBI researchers on the instrument in the DNA Unit II Laboratory in Quantico.

M.1.5. Population studies, Mitochondrial control region tiling assay

Figure 28 shows intra-population cross comparisons of all NIST samples analyzed with the Ibis mtDNA tiling assay. The number of sequence base differences that must exist between two samples over the tiling assay coordinate range is calculated as the minimum number of changes possible to get from one base composition profile to another base composition profile. This calculation takes into account length differences and also corrects for overlapping amplified regions between primers so that differences are not possibly counted twice. It is possible to underestimate the total number of sequence differences (because there can be reciprocal SNPs in an amplicon, or two overlapping amplicons could have an independent sequence difference each that can only be counted once because it is not known *a priori* from base compositions if there is simply one difference in the region overlapped by the two amplicons). However, it is not possible to overestimate the number of differences by this method, hence it is the most conservative calculation for the number of differences that must exist between two profiles. The Caucasian group is generally more homogenous than the other two groups, with an average of 5.9 $\pm$ 3.3 differences between any two profiles. The African American group was the most diverse, with an average of 12.0 $\pm$ 3.4 differences between any two profiles. The Hispanic group had an average of 10.6 $\pm$ 3.6 minimum differences between profiles. Overall, the average minimum difference between any two profiles from the 95 samples was 10.3 $\pm$ 3.8 differences.

A. Caucasian group

Sample		UT57300	UT57301	UT57302	UT57303	UT57310	UT57312	UT57317	UT57318	WT51342	WT51343	WT51345	WT51354	WT51355	WT51358	WT51359	WT51362	WT51373	WT51378	WT51381	WT51386	BC11352	MT97172	WA29584	WA29594	WA29612	ZT81387	MT94859	MT94866	MT94868	MT94869	MT94875
	Well	B01	C01	D01	E01	F01	G01	H01	A02	B02	C02	D02	E02	F02	G02	H02	A03	B03	C03	D03	E03	F03	G03	H03	A04	B04	C04	D04	E04	F04	G04	H04
UT57300	B01	0	4	6	6	3	3	9	3	5	3	2	3	4	6	4	7	4	3	5	6	4	9	4	4	4	5	6	12	11	2	4
UT57301	C01	4	0	5	5	1	3	6	4	3	3	2	4	1	6	2	5	2	1	3	6	2	8	2	4	1	4	6	12	12	3	4
UT57302	D01	6	5	0	9	4	6	8	8	6	6	5	5	5	10	5	9	3	4	6	10	3	8	3	7	5	7	9	14	14	7	6
UT57303	E01	6	5	9	0	5	6	6	6	7	7	5	7	5	5	4	6	6	5	7	9	6	9	6	6	6	8	8	14	11	5	8
UT57310	F01	3	1	4	5	0	2	6	4	2	2	1	3	1	6	1	5	1	0	2	6	1	8	1	3	1	3	5	11	12	3	4
UT57312	G01	3	3	6	6	2	0	8	4	4	4	1	4	3	5	3	7	3	2	4	7	3	9	3	4	3	5	7	14	13	3	6
UT57317	H01	9	6	8	6	6	8	0	9	7	8	7	7	5	6	6	8	5	6	8	11	7	9	5	9	6	9	11	17	14	8	6
UT57318	A02	3	4	8	6	4	4	9	0	6	4	3	5	4	6	5	7	5	4	6	3	5	9	5	5	5	5	7	13	12	1	6
WT51342	B02	5	3	6	7	2	4	7	6	0	4	3	5	3	8	3	7	3	2	4	8	3	10	3	5	3	5	7	11	14	5	6
WT51343	C02	3	3	6	7	2	4	8	4	4	0	3	3	3	6	3	7	3	2	4	6	3	10	3	3	3	3	5	10	12	3	4
WT51345	D02	2	2	5	5	1	1	7	3	3	3	0	3	2	6	2	6	2	1	3	6	2	8	2	3	2	4	6	12	12	2	5
WT51354	E02	3	4	5	7	3	4	7	5	5	3	3	0	4	7	4	8	2	3	5	8	4	8	2	4	4	5	5	11	13	4	3
WT51355	F02	4	1	5	5	1	3	5	4	3	3	2	4	0	6	2	5	2	1	3	6	2	8	2	4	1	4	6	13	12	3	4
WT51358	G02	6	6	10	5	6	5	6	6	8	6	6	7	6	0	6	7	7	6	8	7	7	8	7	7	7	8	7	13	12	5	8
WT51359	H02	4	2	5	4	1	3	6	5	3	3	2	4	2	6	0	5	2	1	3	6	2	8	2	4	2	4	5	11	12	4	5
WT51362	A03	7	5	9	6	5	7	8	7	7	7	6	8	5	7	5	0	6	5	7	9	5	4	6	8	6	8	9	14	9	6	8
WT51373	B03	4	2	3	6	1	3	5	5	3	3	2	2	2	7	2	6	0	1	3	7	2	7	0	4	2	4	6	12	13	4	3
WT51378	C03	3	1	4	5	0	2	6	4	2	2	1	3	1	6	1	5	1	0	2	6	1	8	1	3	1	3	5	11	12	3	4
WT51381	D03	5	3	6	7	2	4	8	6	4	4	3	5	3	8	3	7	3	2	0	6	3	10	3	3	3	3	7	13	12	5	6
WT51386	E03	6	6	10	9	6	7	11	3	8	6	6	8	6	7	6	9	7	6	6	0	7	10	7	6	7	5	8	14	11	4	7
BC11352	F03	4	2	3	6	1	3	7	5	3	3	2	4	2	7	2	5	2	1	3	7	0	9	2	4	2	4	6	12	13	4	5
MT97172	G03	9	8	8	9	8	9	9	9	10	10	8	8	8	8	8	4	7	8	10	10	9	0	7	10	9	11	10	15	10	8	9
WA29584	H03	4	2	3	6	1	3	5	5	3	3	2	2	2	7	2	6	0	1	3	7	2	7	0	4	2	4	6	12	13	4	3
WA29594	A04	4	4	7	6	3	4	9	5	5	3	3	4	4	7	4	8	4	3	3	6	4	10	4	0	4	3	7	12	12	4	6
WA29612	B04	4	1	5	6	1	3	6	5	3	3	2	4	1	7	2	6	2	1	3	7	2	9	2	4	0	4	6	13	13	4	5
ZT81387	C04	5	4	7	8	3	5	9	5	5	3	4	5	4	8	4	8	4	3	3	5	4	11	4	3	4	0	7	11	12	4	6
MT94859	D04	6	6	9	8	5	7	11	7	7	5	6	5	6	7	5	9	6	5	7	8	6	10	6	7	6	7	0	9	13	6	6
MT94866	E04	12	12	14	14	11	14	17	13	11	10	12	11	13	13	11	14	12	11	13	14	12	15	12	12	13	11	9	0	18	12	12
MT94868	F04	11	12	14	11	12	13	14	12	14	12	12	13	12	12	12	9	13	12	12	11	13	10	13	12	13	12	13	18	0	11	12
MT94869	G04	2	3	7	5	3	3	8	1	5	3	2	4	3	5	4	6	4	3	5	4	4	8	4	4	4	6	12	11	0	5	
MT94875	H04	4	4	6	8	4	6	6	6	6	4	5	3	4	8	5	8	3	4	6	7	5	9	3	6	5	6	6	12	12	5	0

B. African American group

Sample		JT51471	JT51499	OT05888	OT05890	OT05892	OT05893	OT05894	OT05896	OT05897	OT05898	OT05899	OT05901	PT84214	PT84215	PT84216	PT84222	PT84223	PT84224	PT84225	PT84226	PT84227	PT84228	PT84230	PT84231	PT84232	PT84234	PT84236	PT84239	PT84240	PT84241	PT84242	PT84243
	Well	A05	B05	C05	D05	E05	F05	G05	H05	A06	B06	C06	D06	E06	F06	G06	H06	A07	B07	C07	D07	E07	F07	G07	H07	A08	B08	C08	D08	E08	F08	G08	H08
JT51471	A05	0	9	11	10	16	8	9	10	10	9	9	8	10	10	12	4	9	13	10	9	9	8	13	15	11	12	18	9	11	10	6	10
JT51499	B05	9	0	3	11	14	12	8	14	8	1	10	10	14	11	4	12	3	12	13	10	7	10	3	14	13	15	14	10	15	9	10	13
OT05888	C05	11	3	0	13	13	13	11	14	9	2	12	14	16	13	5	13	4	14	12	11	8	12	6	12	14	15	14	11	18	11	13	15
OT05890	D05	10	11	13	0	15	10	16	13	10	11	16	13	14	6	14	14	10	15	15	15	8	12	15	15	14	15	18	14	15	12	14	15
OT05892	E05	16	14	13	15	0	13	16	13	14	14	17	15	16	15	15	19	15	13	15	18	12	17	17	2	13	13	4	15	17	11	18	18
OT05893	F05	8	12	13	10	13	0	13	5	9	12	13	12	13	8	15	12	10	14	13	12	9	12	16	13	7	6	15	9	15	11	10	13
OT05894	G05	9	8	11	16	16	13	0	14	10	9	4	7	14	12	11	11	11	11	12	8	11	9	7	16	14	16	17	7	12	7	6	9
OT05896	H05	10	14	14	13	13	5	14	0	10	14	15	13	12	10	17	14	12	12	14	14	10	13	17	13	4	4	15	11	13	12	12	16
OT05897	A06	10	8	9	10	14	9	10	10	0	8	12	10	15	10	10	13	8	11	14	13	2	12	11	12	10	12	14	9	16	9	11	14
OT05898	B06	9	1	2	11	14	12	9	14	8	0	10	12	14	11	3	11	2	12	12	10	7	10	4	14	13	15	14	9	16	9	11	13
OT05899	C06	9	10	12	16	17	13	4	15	12	10	0	9	14	13	12	12	12	12	14	7	12	8	10	17	15	17	18	7	13	8	8	9
OT05901	D06	8	10	14	13	15	12	7	13	10	12	9	0	10	11	14	12	14	12	14	10	10	7	12	14	13	15	15	8	10	7	9	10
PT84214	E06	10	14	16	14	16	13	14	12	15	14	14	10	0	12	16	14	14	7	15	13	13	10	17	16	15	15	17	12	4	12	12	13
PT84215	F06	10	11	13	6	15	8	12	10	10	11	13	11	12	0	14	14	11	12	16	12	8	10	15	14	12	12	17	9	13	10	10	13
PT84216	G06	12	4	5	14	15	15	11	17	10	3	12	14	16	14	0	12	5	14	14	13	10	11	5	15	16	18	16	12	17	12	14	16
PT84222	H06	4	12	13	14	19	12	11	14	13	11	12	12	14	14	12	0	11	15	10	12	13	12	15	19	15	16	21	12	14	14	10	14
PT84223	A07	9	3	4	10	15	10	11	12	8	2	12	14	14	11	5	11	0	13	13	12	7	12	6	15	11	13	15	11	16	11	11	15
PT84224	B07	13	12	14	15	13	14	11	12	11	12	12	12	7	12	14	15	13	0	14	12	11	11	14	13	15	14	15	11	9	8	12	11
PT84225	C07	10	13	12	15	15	13	12	14	14	12	14	14	15	16	14	10	13	14	0	10	12	11	16	15	14	16	16	10	14	13	12	14
PT84226	D07	9	10	11	15	18	12	8	14	13	10	7	10	13	12	13	12	12	12	10	0	11	6	14	17	14	16	19	8	13	8	8	10
PT84227	E07	9	7	8	8	12	9	11	10	2	7	12	10	13	8	10	13	7	11	12	11	0	10	11	11	10	12	14	9	14	7	10	13
PT84228	F07	8	10	12	12	17	12	9	13	12	10	8	7	10	10	11	12	12	11	11	6	10	0	13	16	13	15	17	7	10	9	7	8
PT84230	G07	13	3	6	15	17	16	7	17	11	4	10	12	17	15	5	15	6	14	16	14	11	13	0	17	17	18	15	13	19	12	14	17
PT84231	H07	15	14	12	15	2	13	16	13	12	14	17	14	16	14	15	19	15	13	15	17	11	16	17	0	13	13	5	15	16	11	17	17
PT84232	A08	11	13	14	14	13	7	14	4	10	13	15	13	15	12	16	15	11	15	14	14	10	13	17	13	0	4	14	11	16	12	12	16
PT84234	B08	12	15	15	15	13	6	16	4	12	15	17	15	15	12	18	16	13	14	16	16	12	15	18	13	4	0	15	13	15	14	12	16
PT84236	C08	18	14	14	18	4	15	17	15	14	14	18	15	17	17	16	21	15	15	16	19	14	17	15	5	14	15	0	17	17	13	19	18
PT84239	D08	9	10	11	14	15	9	7	11	9	9	7	8	12	9	12	12	11	11	10	8	9	7	13	15	11	13	17	0	12	7	8	9
PT84240	E08	11	15	18	15	17	15	12	13	16	16	13	10	4	13	17	14	16	9	14	13	14	10	19	16	16	15	17	12	0	10	10	10
PT84241	F08	10	9	11	12	11	11	7	12	9	9	8	7	12	10	12	14	11	8	13	8	7	9	12	11	12	14	13	7	10	0	8	10
PT84242	G08	6	10	13	14	18	10	6	12	11	11	8	9	12	10	14	10	11	12	12	8	10	7	14	17	12	12	19	8	10	8	0	7
PT84243	H08	10	13	15	15	18	13	9	16	14	13	9	10	13	13	16	14	15	11	14	10	13	8	17	17	16	16	18	9	10	10	7	0

Figure 28, Part 1. Intra-population cross-comparison of base composition profiles obtained with the Ibis mtDNA tiling assay for Caucasian and African American reference samples from NIST. The minimum number of differences that must exist between each pair of samples is indicated in the matrix at the intersection of the column and row for the two samples. The minimum number of differences takes into account the overlapping regions between primer pairs and does not count differences twice that could be explained by a single difference. A. Caucasian group. The average minimum number of differences between two Caucasian samples is 5.9 $\pm$ 3.3. B. African American group. The average minimum number of differences between two African American samples is 12.0 $\pm$ 3.4. The average overall difference for any two samples from all three population groups is 10.3 $\pm$ 3.8.

C.

Sample		GT37778	GT37812	GT37828	ZT80932	GT37862	GT37864	GT37869	GT37888	GT37900	GT37913	JT52076	OT07280	PT85612	PT85658	TT51399	TT51407	TT51422	TT51435	TT51483	TT51511	TT51530	ZT80731	ZT80737	ZT80782	ZT80786	ZT80815	ZT80826	ZT80863	ZT80865	ZT80869	ZT80870	ZT80925
	Well	A09	B09	C09	D09	E09	F09	G09	H09	A10	B10	C10	D10	E10	F10	G10	H10	A11	B11	C11	D11	E11	F11	G11	H11	A12	B12	C12	D12	E12	F12	G12	H12
GT37778	A09	0	13	14	15	13	13	13	11	13	12	12	14	2	5	12	14	14	13	11	9	14	5	12	3	12	12	11	13	12	3	12	10
GT37812	B09	13	0	4	8	11	13	14	9	7	10	14	13	13	12	12	5	16	1	8	7	3	15	9	13	14	15	9	0	14	13	2	10
GT37828	C09	14	4	0	8	10	12	13	8	7	9	14	12	14	14	11	1	15	3	11	6	6	16	8	14	14	14	8	4	13	14	4	9
ZT80932	D09	15	8	8	0	14	12	17	11	7	12	15	15	15	13	15	9	15	7	12	11	8	16	11	15	14	14	11	8	13	15	7	12
GT37862	E09	13	11	10	14	0	5	12	5	13	2	11	5	13	12	11	11	16	11	9	6	12	14	6	13	11	9	6	11	7	12	10	6
GT37864	F09	13	13	12	12	5	0	14	8	16	7	13	8	13	12	13	13	15	13	11	8	15	14	7	12	11	11	9	13	5	12	13	8
GT37869	G09	13	14	13	17	12	14	0	9	15	10	11	11	15	14	13	14	14	14	10	9	15	16	10	15	10	11	8	14	9	15	13	9
GT37888	H09	11	9	8	11	5	8	9	0	10	3	9	7	11	9	8	9	14	9	7	4	10	13	4	11	9	8	2	9	8	11	8	3
GT37900	A10	13	7	7	7	13	16	15	10	0	11	16	14	13	11	13	8	17	6	10	8	8	14	10	13	17	13	9	7	16	12	5	11
GT37913	B10	12	10	9	12	2	7	10	3	11	0	10	4	12	10	8	10	15	10	8	5	11	14	5	12	10	7	4	10	7	12	9	4
JT52076	C10	12	14	14	15	11	13	11	9	16	10	0	9	11	11	13	15	16	15	11	10	15	15	9	13	10	10	8	14	14	13	14	7
OT07280	D10	14	13	12	15	5	8	11	7	14	4	9	0	15	12	10	13	17	13	11	8	14	15	5	15	10	7	7	13	9	14	13	7
PT85612	E10	2	13	14	15	13	13	15	11	13	12	11	15	0	5	12	14	14	13	11	9	14	5	12	3	13	12	10	13	13	3	12	10
PT85658	F10	5	12	14	13	12	12	14	9	11	10	11	12	5	0	12	14	15	13	10	8	15	7	9	5	12	9	9	12	11	5	12	8
TT51399	G10	12	12	11	15	11	13	13	8	13	8	13	10	12	12	0	10	15	12	12	5	12	11	8	11	14	10	9	12	12	12	11	9
TT51407	H10	14	5	1	9	11	13	14	9	8	10	15	13	14	14	10	0	16	4	12	7	7	16	9	14	15	15	9	5	14	14	5	10
TT51422	A11	14	16	15	15	16	15	14	14	17	15	16	17	14	15	15	16	0	14	13	13	18	15	14	14	15	15	13	16	14	14	14	14
TT51435	B11	13	1	3	7	11	13	14	9	6	10	15	13	13	13	12	4	14	0	8	7	3	15	9	13	14	15	9	1	14	13	1	10
TT51483	C11	11	8	11	12	9	11	10	7	10	8	11	11	11	10	12	12	13	8	0	7	9	13	8	11	10	11	6	8	10	11	7	5
TT51511	D11	9	7	6	11	6	8	9	4	8	5	10	8	9	8	5	7	13	7	7	0	8	11	4	9	10	8	4	7	9	9	6	5
TT51530	E11	14	3	6	8	12	15	15	10	8	11	15	14	14	15	12	7	18	3	9	8	0	16	10	12	15	16	10	3	15	14	3	11
ZT80731	F11	5	15	16	16	14	14	16	13	14	14	15	15	5	7	11	16	15	15	13	11	16	0	14	6	13	12	13	15	14	6	14	12
ZT80737	G11	12	9	8	11	6	7	10	4	10	5	9	5	12	9	8	9	14	9	8	4	10	14	0	12	9	7	4	9	8	12	9	4
ZT80782	H11	3	13	14	15	13	12	15	11	13	12	13	15	3	5	11	14	14	13	11	9	12	6	12	0	13	12	10	13	13	2	12	10
ZT80786	A12	12	14	14	14	11	11	10	9	17	10	10	10	13	12	14	15	15	14	10	10	15	13	9	13	0	10	11	14	10	14	13	8
ZT80815	B12	12	15	14	14	9	11	11	8	13	7	10	7	12	9	10	15	15	15	11	8	16	12	7	12	10	0	8	15	10	12	14	7
ZT80826	C12	11	9	8	11	6	9	8	2	9	4	8	7	10	9	9	9	13	9	6	4	10	13	4	10	11	8	0	9	8	10	8	4
ZT80863	D12	13	0	4	8	11	13	14	9	7	10	14	13	13	12	12	5	16	1	8	7	3	15	9	13	14	15	9	0	14	13	2	10
ZT80865	E12	12	14	13	13	7	5	9	8	16	7	14	9	13	11	12	14	14	14	10	9	15	14	8	13	10	10	8	14	0	13	13	8
ZT80869	F12	3	13	14	15	12	12	15	11	12	12	13	14	3	5	12	14	14	13	11	9	14	6	12	2	14	12	10	13	13	0	12	10
ZT80870	G12	12	2	4	7	10	13	13	8	5	9	14	13	12	12	11	5	14	1	7	6	3	14	9	12	13	14	8	2	13	12	0	9
ZT80925	H12	10	10	9	12	6	8	9	3	11	4	7	7	10	8	9	10	14	10	5	5	11	12	4	10	8	7	4	10	8	10	9	0

Figure 28, Part 2. Intra-population cross-comparison of base composition profiles obtained with the Ibis mtDNA tiling assay for Hispanic reference samples from NIST. The minimum number of differences that must exist between each pair of samples is indicated in the matrix at the intersection of the column and row for the two samples. The minimum number of differences takes into account the overlapping regions between primer pairs and does not count differences twice that could be explained by a single difference. C. Hispanic group. The average minimum number of differences between two Hispanic samples is 10.6 $\pm$ 3.6. The average overall difference for any two samples from all three population groups is 10.3 $\pm$ 3.8.

M.1.6. Mixture studies, Mitochondrial control region tiling assay

To test the ability of the mtDNA control region tiling assay to interrogate mixtures of DNA templates, and as part of our effort to transfer technology to collaborators, researchers at the FBI mixed DNA templates in specific ratios and ran the Ibis mtDNA tiling assay on their own instrument at the FBI laboratories. Automated data processing was also performed at the FBI laboratory as part of their integrated system operation. The full FBI database and specific batches of raw data were transferred to Ibis using a replication utility developed to provide a portable analysis mode for the mtDNA system.

DNA was isolated from buccal swabs taken from FBI staff members using the FBI DNA Unit II standard DNA isolation protocol (pages 15 to 19 of the FBI Mitochondrial DNA Analysis Protocol, Revision: 0, 06/09/06). DNA samples used in this study are referred to be numerical codes and are labeled FBI-57, FBI-61, FBI-65, FBI-70 and FBI-72. Mitochondrial copies in each sample were estimated with a real-time PCR quantification method developed in the FBI research division. The real-time PCR assay involves the amplification of a PCR product spanning revised Cambridge Reference Sequence coordinates 16517 to 16567 and employs a fluorescent probe spanning coordinates 16528-16546. A dilution series of a plasmid clone carrying the control region from the mtDNA genome from human cell line HL-60 was used for a standard curve for quantification of mtDNA genomes from DNA samples. DNA mixtures were constructed on the basis of estimated mitochondrial copies.

Mixtures of FBI-65 and FBI-70 were made using a total of 500,000 or 100,000 mtDNA copies. All other mixtures were made with a total of 100,000 mtDNA copies, as shown in Table 17. Mixtures were made at 99:1, 95:5, 90:10, 75:25, 25:75, 10:90, 5:95 and 1:99. Mixtures of 99:1 and 1:99 produced a single, clean profile in all cases. There are two implications of this result: 1.) The Ibis mtDNA tiling assay is currently not capable of detecting and resolving mixtures where the minor contributor is present at 1% or lower of the total DNA in a sample, and 2.) the presence of a contaminant at a level of 1% or below the total DNA input will not interfere with simple, automated analysis of mitochondrial DNA using the Ibis mtDNA tiling assay. The point at which a minor

contributor hovers between resolvable and undetectable appears to be ~5% of the total DNA input (Table 17). Mixtures are not reliably resolved for a minor contributor present at 5% of the total DNA input. With manual interrogation of the data, mixtures containing 10% minor contributor are resolvable (Table 17). Mixtures containing 25% minor contributor are easily identified with the automated data processing and full profiles can be resolved clearly with manual data interrogation. Moreover, relative mass spectrometry signal outputs for mixed products track well with template input ratios (Table 17), suggesting that, at minor contributor levels above 10%, relative minor-to-major contributor levels can be determined directly from relative product signal intensities. Figure 29 demonstrates the detection of peaks specific product peaks from a 25:75 mixture of two templates (FBI-65 and FBI-70). In Figure 29, panel B, length multiple length variants resulting from heteroplasmy in HV1 from FBI-70 are clearly resolved as double-stranded products along with the distinct product contributed by FBI-65. Four independent double-stranded products are observed in the view shown in Figure 29, B. The ratio of the signal amplitude from FBI-65 to the total combined amplitude for all four products is close to 25%. There is never-the-less elevated quantification error in regions with length heteroplasmy, presumably because some length variants fall below detection limits and the total signal for the variants from a heteroplasmic template becomes diluted by splitting the signal across multiple products. Figure 30 shows the quantification of relative signal outputs for products that differ between the mixed templates for ratios of 90:10, 75:25, 25:75 and 10:90 of FBI-65 and FBI-70.

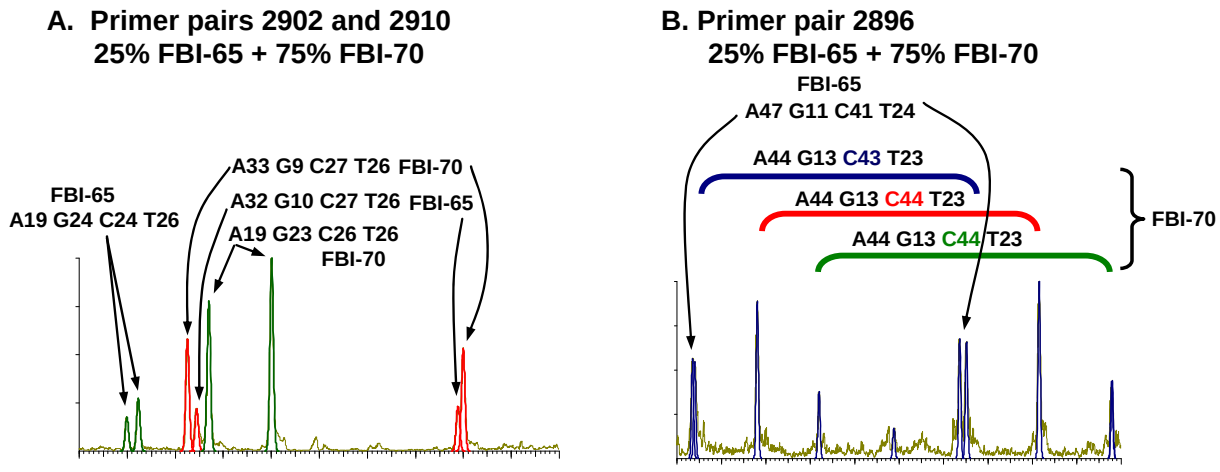


Figure 29. Deconvolved spectra from a 25:75 mixture of FBI-65 and FBI-70. Detailed view of mass spectra zoomed in to specific regions of triplex reactions performed on a 25:75 mixture of FBI-65 and FBI-70 using a total of 100,000 mtDNA genome copies. A.) Products from two primer pairs (2902 and 2910) of triplex 5 are shown demonstrating the resolution of distinct products for the two input templates at signal output ratios corresponding approximately to the intended template input ratios. B.) Products from primer pair 2896 from triplex 7 in a 25:75 mixture of FBI-65 and FBI-70. FBI-70 has multiple C-length heteroplasmies in the HV1 poly-C stretch, which is spanned by primer pair 2896. Three distinct length-variant products are displayed here for template FBI-70 (a fourth is also present but not in zoomed view), A separate distinct product [A47 G11 C41 T24] is present that corresponds to the product generated by template FBI-65. The signal amplitude observed for the FBI-65 product [A47 G11 C41 T24] is ~23% of the combined signal amplitude for the four FBI-70 products [A44 G13 C42/43/44/45 T23].

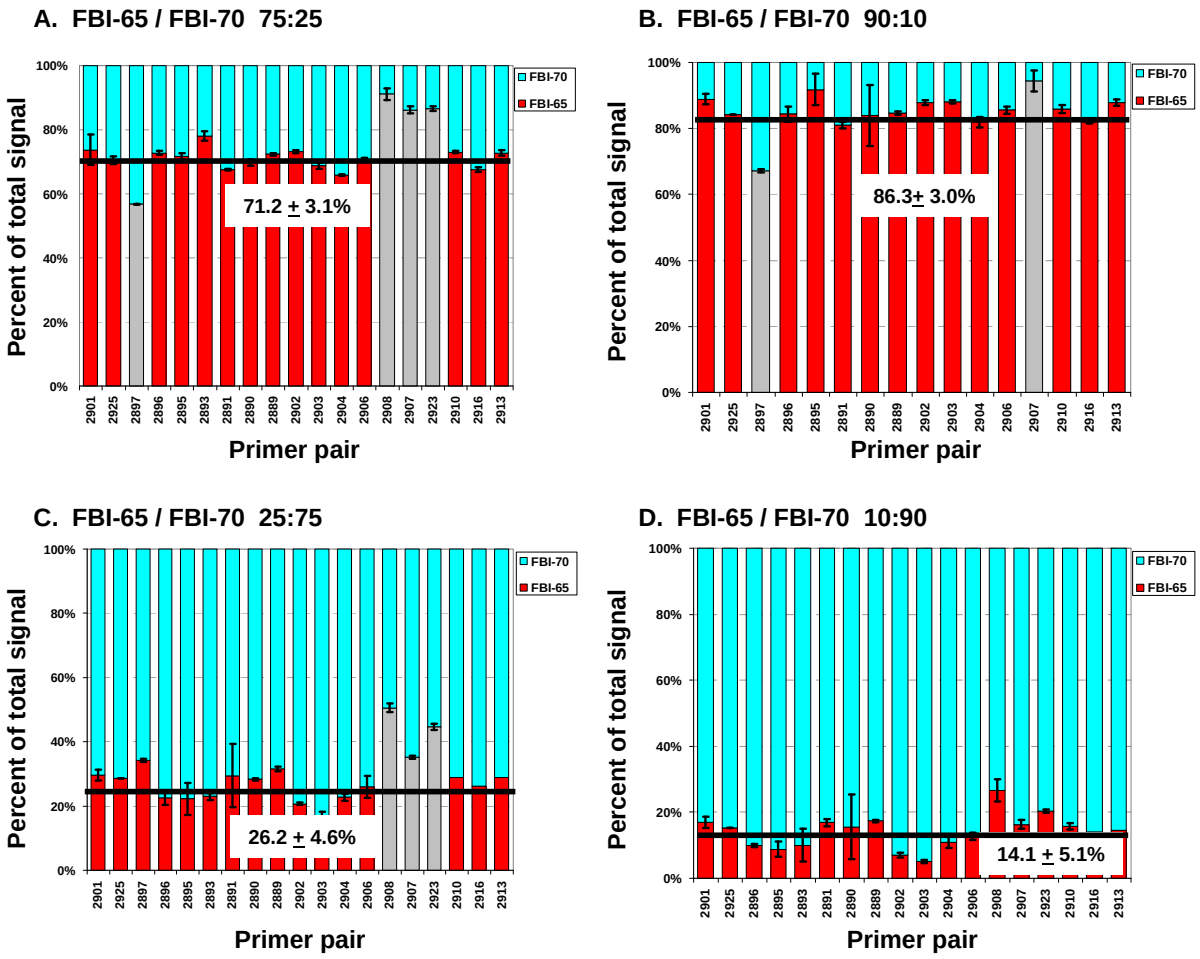


Figure 30. Quantification of relative abundances of mixed templates in the mtDNA tiling assay. Combinations of 75:25, 90:10, 25:75 and 10:90 (copies:copies) were prepared to provide an estimated 500,000 total mtDNA copies to each PCR reaction in the mtDNA tiling assay. The mtDNA tiling assay was run in triplicate for each mixture using mtDNA tiling kits plates supplied to the FBI by Ibis. All reactions were analyzed on an Ibis T5000 instrument at the FBI DNA Unit II laboratory (Quantico VA). Automated data processing was performed at the FBI and processed and raw data were sent to Ibis for analysis. Template FBI-75 had length heteroplasmy in HV1 and HV2. In HV2, length heteroplasmy appears to cause a substantial splitting of signal and loss of combined signal amplitude for FBI-70, resulting in an artificially-inflated estimate of FBI-65-to-FBI-70 product ratios for products that span the HV2 poly-C stretch (2907,2908 and 2923). For this reason, these regions were excluded from the calculation of estimated template ratios (gray bars above). Primer pair 2897 was excluded from this analysis because the forward strands of the two observed products [A31 G13 C29 T28] and [A33 G12 C28 T28] overlap in mass, causing co-addition of amplitudes for one strand of both products.

Table 17. Observed profiles and signal output ratios for mixtures of templates in the Ibis mitochondrial control region tiling assay.

Template 1	Template 2	Total amount (copies)	Template 1 : Template 2 Ratio	Observed profile	Percent estimate for template 1	Relative signal quantification estimate
FBI-65	FBI-70	500,000	99:1	FBI-65	100%	ND
FBI-65	FBI-70	500,000	95:5	FBI-65	100%	ND
FBI-65	FBI-70	500,000	90:10	FBI-65 + FBI-70	86.3 + 3.0%	86:14
FBI-65	FBI-70	500,000	75:25	FBI-65 + FBI-70	71.2 ± 3.1%	71:29
FBI-65	FBI-70	500,000	25:75	FBI-65 + FBI-70	26.2 ± 4.6%	26:74
FBI-65	FBI-70	500,000	10:90	FBI-65 + FBI-70	14.1 ± 5.1%	14:86
FBI-65	FBI-70	500,000	5:95	FBI-70	0%	ND
FBI-65	FBI-70	500,000	1:99	FBI-70	0%	ND
FBI-65	FBI-70	100,000	99:1	FBI-65	100%	ND
FBI-65	FBI-70	100,000	99:1	FBI-65	100%	ND
FBI-65	FBI-70	100,000	95:5	FBI-65	100%	ND
FBI-65	FBI-70	100,000	95:5	FBI-65	100%	ND
FBI-65	FBI-70	100,000	75:25	FBI-65 + FBI-70	79.7 ± 6.5%	80:20
FBI-65	FBI-70	100,000	75:25	FBI-65 + FBI-70	77.9 ± 6.5%	78:22
FBI-65	FBI-70	100,000	25:75	FBI-65 + FBI-70	37.9 ± 4.3%	38:62
FBI-65	FBI-70	100,000	25:75	FBI-65 + FBI-70	33.8 ± 4.5%	34:66
FBI-65	FBI-70	100,000	5:95	FBI-70	0%	ND
FBI-65	FBI-70	100,000	5:95	FBI-65 + FBI-70	12.0 ± 4.4	12:88
FBI-65	FBI-70	100,000	1:99	FBI-70	0%	ND
FBI-65	FBI-70	100,000	1:99	FBI-70	0%	ND
FBI-61	FBI-72	100,000	99:1	FBI-61	100%	ND
FBI-61	FBI-72	100,000	99:1	FBI-61	100%	ND
FBI-61	FBI-72	100,000	95:5	FBI-61	100%	ND
FBI-61	FBI-72	100,000	95:5	FBI-61	100%	ND
FBI-61	FBI-72	100,000	75:25	FBI-61 + FBI-72	66.6 ± 7.6%	67:33
FBI-61	FBI-72	100,000	75:25	FBI-61 + FBI-72	72.5 ± 5.4	73:27
FBI-61	FBI-72	100,000	25:75	FBI-61 + FBI-72	21.9 ± 3.4%	22:78
FBI-61	FBI-72	100,000	25:75	FBI-61 + FBI-72	30.7 ± 4.3	31:69
FBI-61	FBI-72	100,000	5:95	FBI-72	0%	ND
FBI-61	FBI-72	100,000	5:95	FBI-61 + FBI-72	10.0 ± 3.1%	10:90
FBI-61	FBI-72	100,000	1:99	FBI-72	0%	ND
FBI-61	FBI-72	100,000	1:99	FBI-72	0%	ND
FBI-65	FBI-72	100,000	99:1	FBI-65	100%	ND
FBI-65	FBI-72	100,000	95:5	FBI-65	100%	ND
FBI-65	FBI-72	100,000	75:25	FBI-65 + FBI-72	75.2 ± 7.5%	75:25
FBI-65	FBI-72	100,000	25:75	FBI-65 + FBI-72	26.7 ± 6.7%	27:73
FBI-65	FBI-72	100,000	5:95	FBI-72	0%	ND
FBI-65	FBI-72	100,000	1:99	FBI-72	0%	ND
FBI-57	FBI-70	100,000	99:1	FBI-57	100%	ND
FBI-57	FBI-70	100,000	95:5	FBI-57	100%	ND
FBI-57	FBI-70	100,000	75:25	FBI-57 + FBI-70	64.9 ± 2.1	65:35
FBI-57	FBI-70	100,000	25:75	FBI-57 + FBI-70	20.6 ± 4.4	21:79
FBI-57	FBI-70	100,000	5:95	FBI-70	0%	ND
FBI-57	FBI-70	100,000	1:99	FBI-70	0%	ND

M.1.7 PCR-based procedure-specific validation, Mitochondrial control region tiling assay

Potential for differential amplification among loci.

For the mitochondrial assay, the relative signal intensities of the PCR products in each multiplexed reaction were examined over the replicate reactions performed in section M.1.3. First, the amplitude of the forward strand of a PCR product was plotted against the amplitude of the second strand of a PCR product for each PCR product in a triplex. All PCR products for a given triplex reaction are plotted together. This was done for each triplex with the 55 replicates from both sample AF-25 and sample FBI-75. Figures 31-34 show that the two strands of each PCR product are linearly correlated over 55 replicates. Figure 31 shows the plot for triplex reaction A for sample AF-25 and Figure 32 shows the plot for triplex reaction A for sample FBI-75. For triplex reaction A, no difference is observed between the two different templates. Figure 33 shows the plot for triplex reaction E for sample AF-25 and Figure 34 shows the plot for triplex reaction E for sample FBI-75. Because sample FBI-75 exhibits C-length heteroplasmy for primer pair 2893, the plots for the two templates are different. The signal intensities for the two strands of the products from primer pair 2893 are decreased relative to the two other primer pairs in the triplex reaction. The signal intensity for products from primer pair 2893 is spread among multiple products due to the C-length heteroplasmy. In contrast, sample AF-25 does not exhibit C-length heteroplasmy for primer pair 2893 and no decrease in signal intensity is observed for primer pair 2893 relative to the other two primer pairs in the triplex reaction.

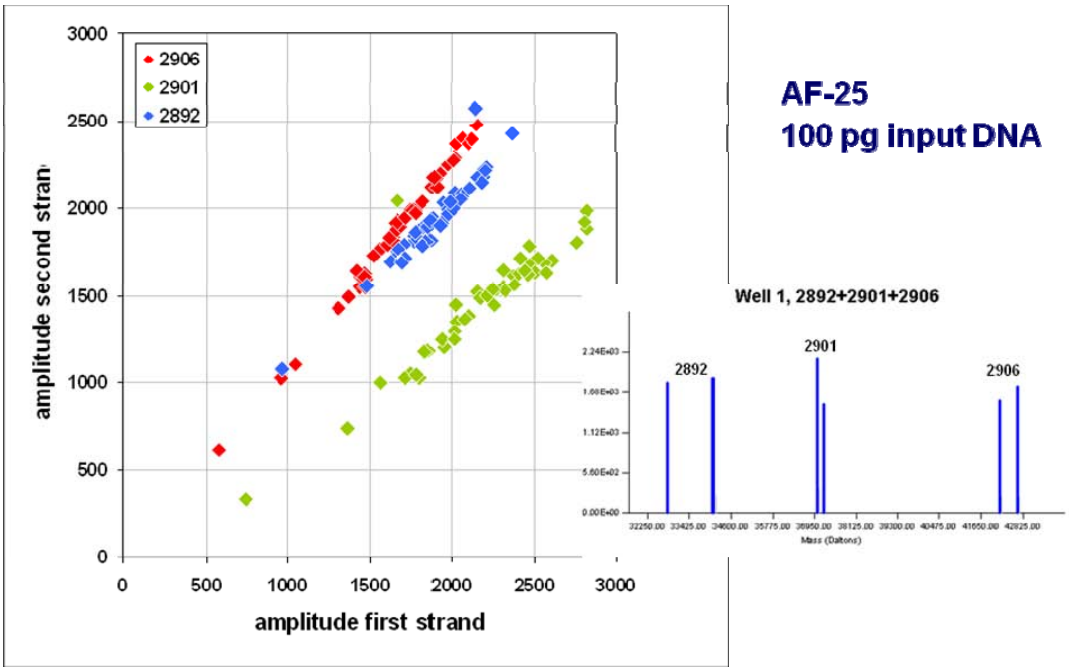


Figure 31. For triplex reaction A of sample AF-25, the amplitude of the first strand is plotted verses the amplitude of the second strand for each PCR product. Triplex A contains primer pairs 2892 (♦), 2901 (♦), and 2906 (♦).

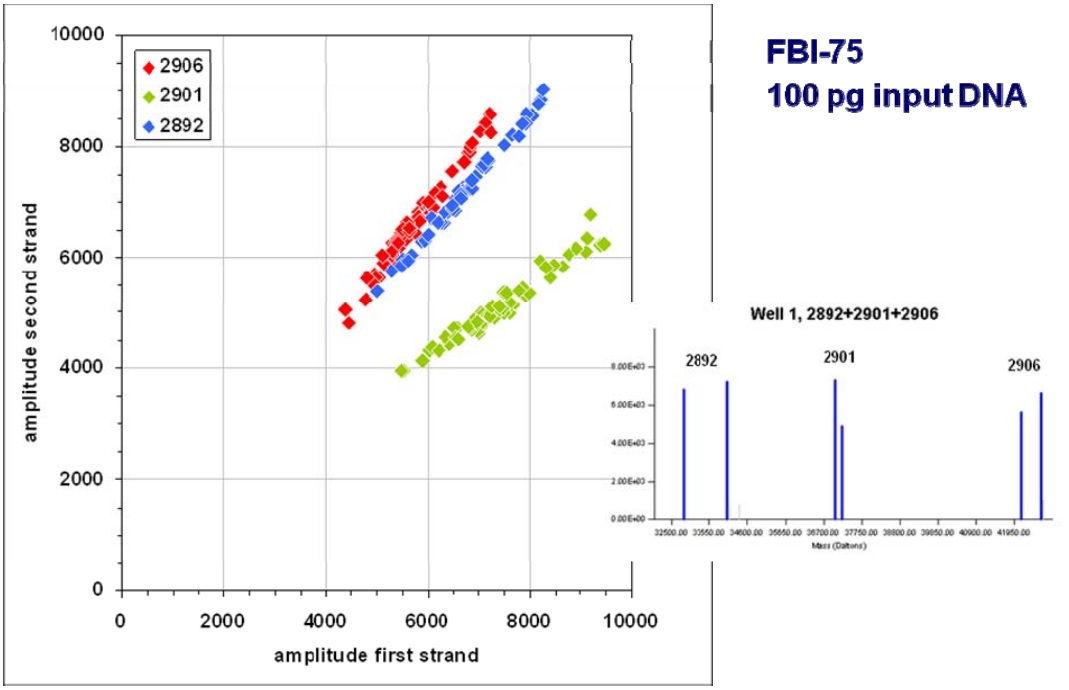


Figure 32. For triplex reaction A of sample FBI-75, the amplitude of the first strand is plotted verses the amplitude of the second strand for each PCR product. Triplex A contains primer pairs 2892 (♦), 2901 (♦), and 2906 (♦).

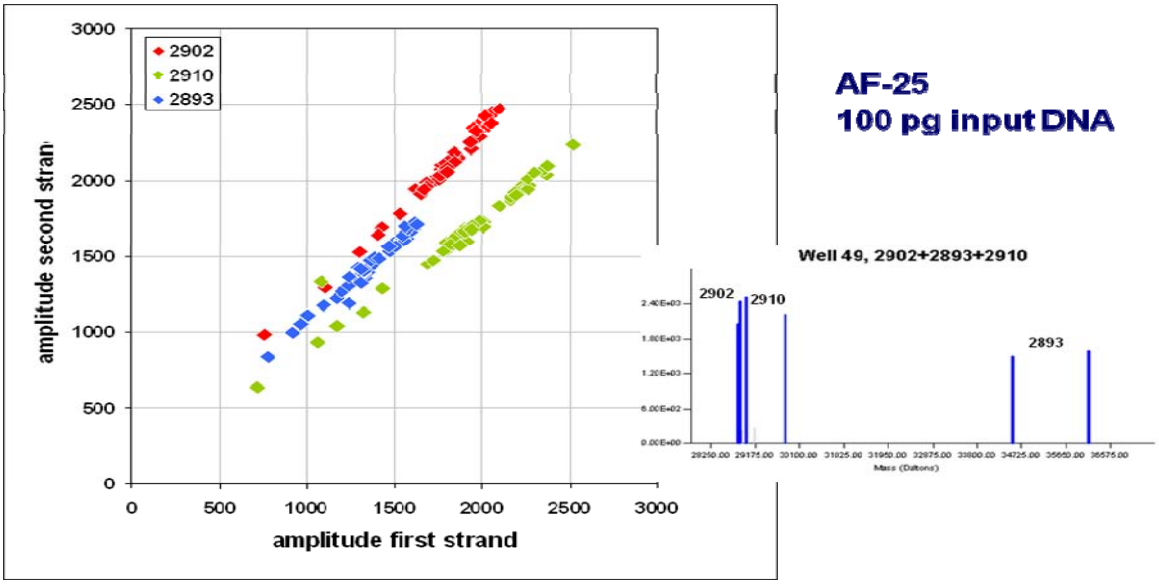


Figure 33. For triplex reaction E of sample AF-25, the amplitude of the first strand is plotted verses the amplitude of the second strand for each PCR product. Triplex E contains primer pairs 2893 (♦), 2910 (♦), and 2902 (♦).

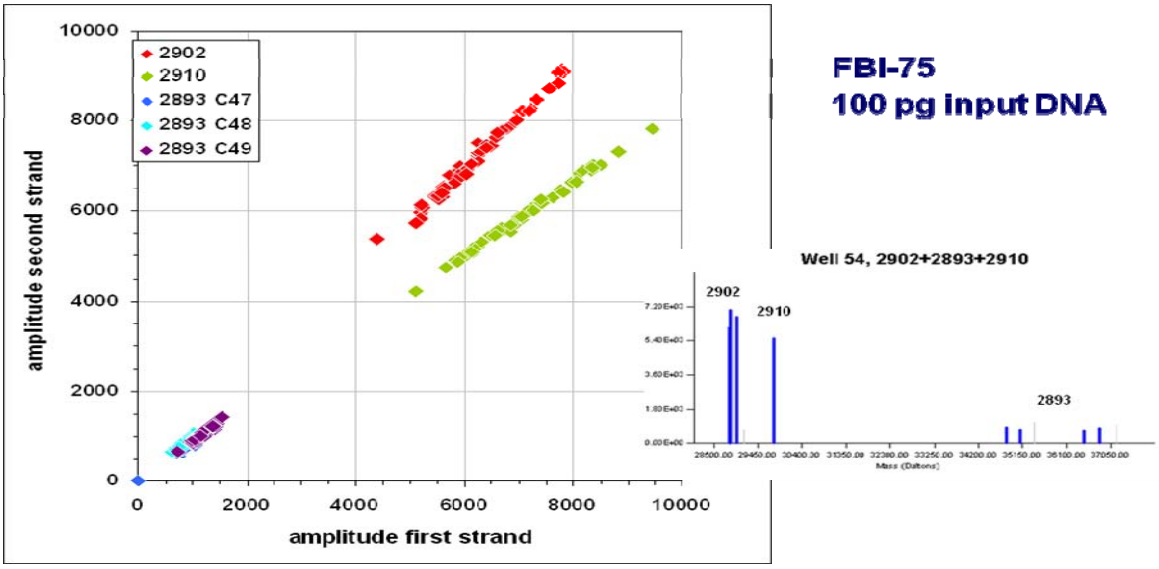


Figure 34. For triplex reaction E of sample FBI-75, the amplitude of the first strand is plotted verses the amplitude of the second strand for each PCR product. Triplex E contains primer pairs 2893 C47 (♦), 2893 C48 (♦), 2893 C49 (♦), 2910 (♦), and 2902 (♦).

In order to determine acceptable product ratios within a triplex, the ratio between the minimum and maximum intensity products within a triplex was calculated for each triplex as a function of the input template (AF-25 or FBI-75). First, the intensity of the two strands for each PCR product was averaged. Then, the ratio of the minimum intensity product to the maximum intensity product was calculated for each of the 55 replicates. For the 55 replicates, the average ratio and standard deviation was determined for each triplex and is shown in Table 18. For triplexes that do not have primer pairs that can exhibit C-length heteroplasmy (Triplex A and F), the ratio does not change substantially between the two templates. In contrast, when the product ratios are compared between a template (FBI-75) that exhibits C-length heteroplasmy and a template (AF-25) that does not exhibit C-length heteroplasmy, the product ratio changes by a factor of 1.5 or greater for triplexes with a primer pair that exhibits C-length heteroplasmy (Triplex B, C, D, E, G and H). When no C-length heteroplasmy is present, this study indicates that reproducible ratios between products in a triplex reaction will be observed across different templates.

Although it is generally the case that product balance is reproducible across templates, in practice, factors other than heteroplasmy can affect the inter-product balance observed for each primer pair triplex in the presence of different templates. During the course of sample analysis outside the scope of this project, we have observed a substantial template-to-template variation in product balance in the three reactions that span the HV2 poly-C stretch located at rCRS 303-315 (data not shown). The observed balance is reproducible on a given template, and appears to be influenced by the amplified region, more than the primer-binding sites. In the small number of templates we have sequenced that gave problems amplifying through the HV2 poly-C region, we did not observe obvious primer/target mismatches to account for lost amplification efficiency. Moreover, all three HV2 poly-C-spanning primer pairs (2907, 2908 and 2923) are generally affected equivalently with problematic templates. Other primer pairs are on rare occasions affected presumably by primer-target mismatches near the 3´ and of one or the other primer and a primer pair may occasionally drop out of a reaction. We did not observe primer pair drop-outs in the samples ran in the scope of

this project, nor in validation studies with the FBI and AFDIL, but we have seen them in other samples processed in our laboratory.

Table 18. The average and standard deviation of the ratio of the minimum intensity product to maximum intensity product for each triplex as a function of input template.

Triplex	Primer Pairs	AF-25 ratio	FBI-75 ratio
A	2892	0.96 ± 0.09	0.89 ± 0.03
	2901		
	2906		
B	2891	0.46 ± 0.03	0.22 ± 0.11
	2908		
	2925		
C	2890	0.75 ± 0.02	0.48 ± 0.02
	2899		
	2907		
D	2889	0.68 ± 0.05	0.38 ± 0.06
	2898		
	2923		
E	**2893**	0.73 ± 0.08	0.17 ± 0.01
	2902		
	2910		
F	2897	0.62 ± 0.03	0.64 ± 0.03
	2903		
	2916		
G	**2896**	0.61 ± 0.04	0.22 ± 0.06
	2904		
	2913		
H	**2895**	0.54 ± 0.08	0.24 ± 0.02
	2905		
	2912		

As an additional way to assess the general range of balance between products generated by different primers within each triplex, the ratio of higher signal-to-lower signal product for all primer pair combinations in each triplex was measured in 55 replicates for two input templates (AF-25 and FBI-75). Figure 35 shows the average inter-product ratios observed in each triplex for the two templates. The largest apparent

sample dependence on the signal ratios observed between primer pairs is in regions that display length heteroplasmy (Figure 35, panels B, C, D, E, G and H).

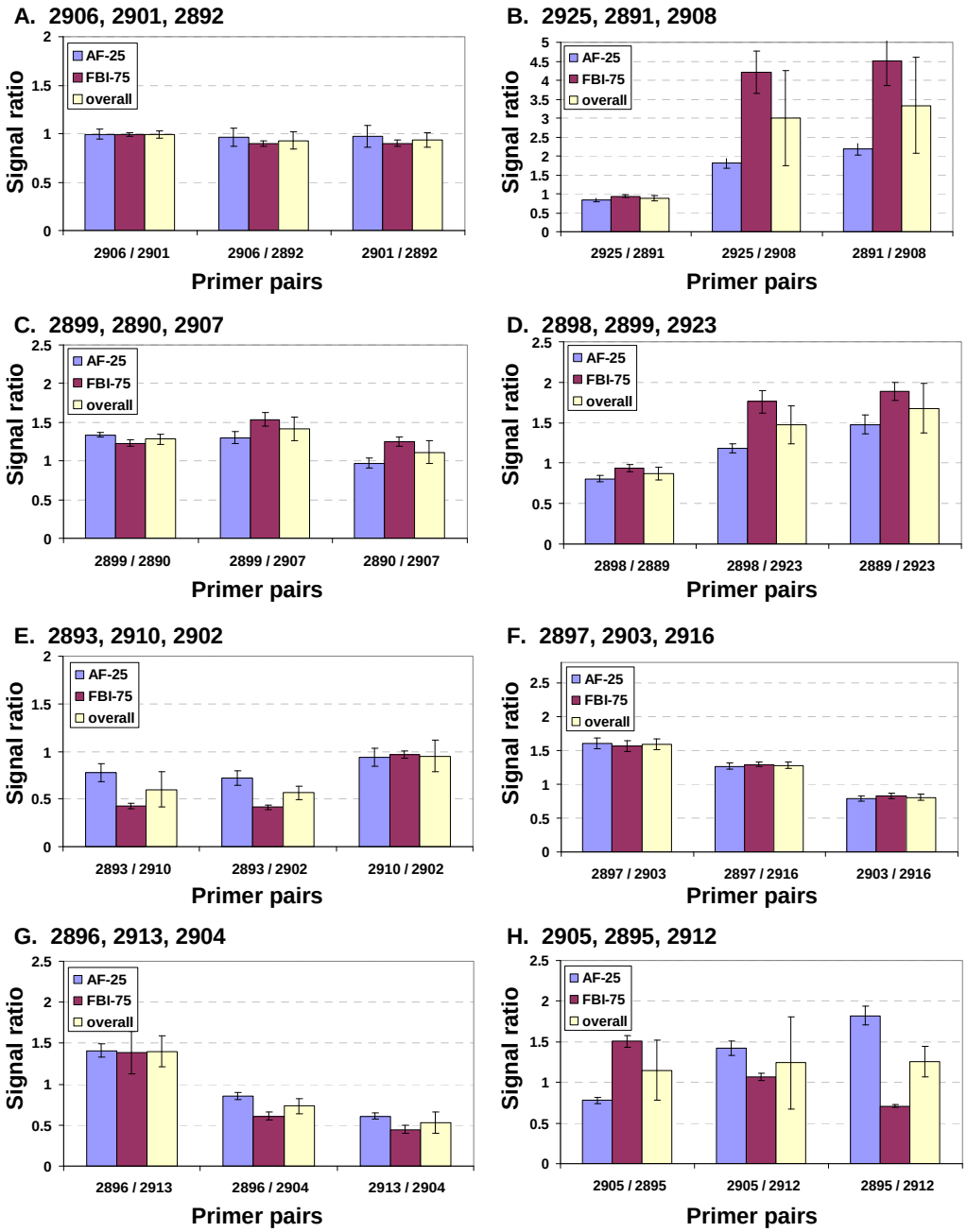


Figure 35. Inter-primer pair product balance for mitochondrial DNA primers. Total signal output ratios for each pair of primer pairs in each triplex are plotted by sample for each mtDNA tiling triplex over two samples run 55 times each. The right-most bar in each graph is the average pair:pair ratio for all samples combined. All ratios are expressed as the total signal ratio of the locus on the left side of each label compared to the locus on the right side of each label.

Positive and negative controls.

We include a fully-characterized DNA template in each assay as a positive control and a water control (no template) with each plate of assay runs. For example, Figure 36 shows the results of 2,207 positive and negative controls run together in automated runs of our mitochondrial control region tiling assay. Over the course of these 2,207 trials, we have had a full correct, profile automatically assigned for the positive control ~98% of the time and a clean negative control ~98% of the time. When either of these controls is flagged, samples must be interrogated carefully and results must be questioned thoroughly.

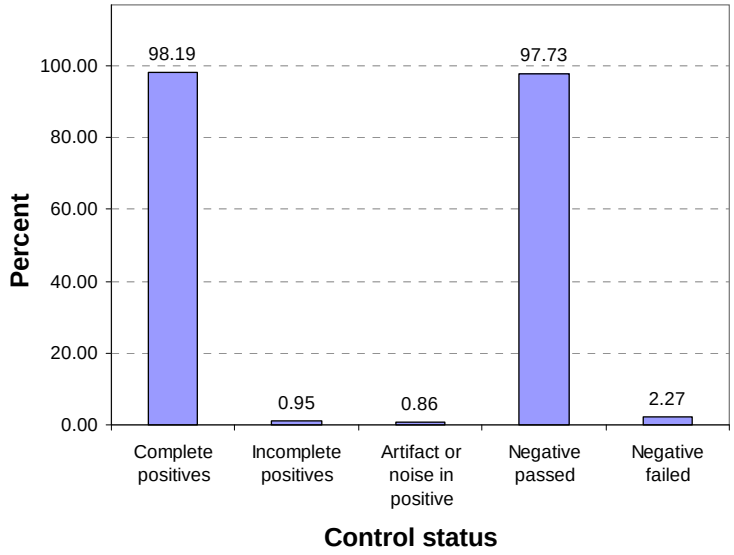


Figure 36. Results of automated software processing on 2,207 positive and negative controls in the mitochondrial control region tiling assay.

Specific aim 2. Development of a transferrable software interface

We have developed an interface for analysis of the Ibis mtDNA tiling assay and the Ibis STR assay that operates directly as a module within the IbisTrack sample / assay tracking package. All basic elements of assay setup, tracking, data processing and analysis are controlled through the IbisTrack interface. Assay "kit" templates are configured and stored in a relational database (Oracle) such that sample names simply need to be added to a pre-defined kit layout to register a 96-well assay plate to be run through the Ibis T5000 system. The running and processing of a mitochondrial tiling kit is fully end-to-end automated. Mass spectrometry and data processing steps for the STR assay are also automated, but there is currently a manual step to trigger data analysis and import processed data into our database.

General software overview

The mtDNA tiling and STR assays are heavily dependent upon software. All steps are performed or controlled by software from the point of DNA addition to a PCR plate through final sample analysis and visual verification. The discussion below is expressed mainly in terms of the mitochondrial control region tiling assay, but the same software core is used for analysis of the STR assay. At the interface layer, however, the analysis software is custom-tailored to the mtDNA or STR assay, depending on which assay is being analyzed. Figure 37 shows a flow diagram of software control through the mtDNA assay process. Briefly, an assay is pre-registered into a

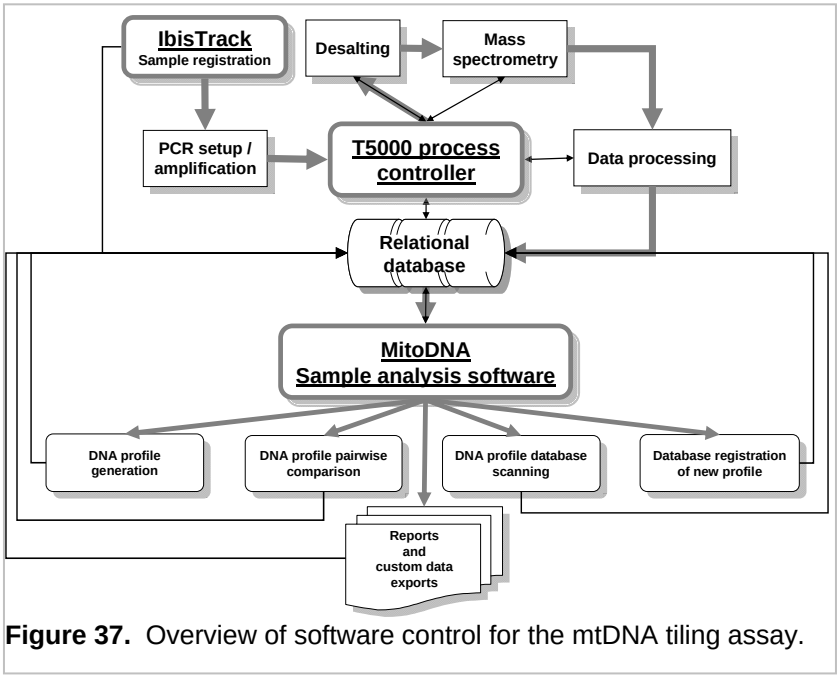


Figure 37. Overview of software control for the mtDNA tiling assay.

relational database through the sample tracking software IbisTrack, and then DNA is added (either manually or robotically) to a pre-made PCR plate. After PCR amplification in a thermocycler, the amplified PCR plate is turned over to the Ibis T5000 process controller. The controller directs the desalting of PCR products and then injection and measurement in the mass spectrometer.

Subsequent to data collection for a plate, the controller converts raw data to a standardized format and transfers all raw spectral data to a defined storage location under a barcode for each assay plate such that spectra are stored in subdirectories by assay plate well. The controller then triggers data processing software that converts raw (mass-to-charge ratio) spectra to deconvolved (mass) spectra and identifies individual masses present within each spectrum with a high degree of accuracy. Upon notification of completed data processing, the controller registers processed masses for each spectrum into the relational database indexed by assay plate barcode and well, along with the storage path to the raw data for each spectrum. The identity and assay layout for each sample is directly tied to raw and processed data through the relational database.

After automated processing steps. Samples are ready for software-assisted analysis and manual verification. The MitoDNA module is directly integrated into and controlled from the IbisTrack sample tracking application. Assay plates that have processed data ready are presented in a list and a plate is chosen by barcode and analyzed with a button-click. The MitoDNA software associates measured product masses with potential mtDNA PCR products for the primer pairs in the mtDNA assay and develops a base composition profile for each sample on the assay plate. The base composition profile is presented to the user with links to all raw and deconvolved mass spectra. The user can then browse the analyzed data and make insertions, deletions or corrections to automated product assignments if necessary. Finalized base composition profiles can then be registered into the relational database for later retrieval, comparison to other profiles, or contribution to a search database.

Deconvolution

The initial result of an mtDNA tiling assay for an individual sample is a collection of eight mass spectra. Processing software is employed that takes each mass spectrum and analyzes it to: 1.) remove general background noise, 2.) fine-tune calibration of the spectrum, based upon the known mass of each internal peptide calibrant, and 3.) "deconvolve" the raw spectrum to produce a simplified spectrum consisting of mass on the X-axis and signal intensity on the Y-axis. Conversion of the raw (mass-to-charge ratio) spectrum to the deconvolved (mass) spectrum employs a software process, termed MassCollapse. MassCollapse coadds all spectral signal in a mass-to-charge ratio spectrum at each mass along a continuous mass scale, resulting in a one-dimensional spectrum of signal intensity over a defined mass range. The resulting mass spectrum is traversed to determine mass peak positions and final masses are reported in a list that is imported directly into a database for storage and retrieval during sample analysis.

Product assignment / Complementarity rules

All masses calculated from a mass spectrum are compared to a database of product base compositions for previously characterized samples analyzed by sequencing or base composition analysis to develop a list of possible base compositions associated with each mass assigned within the spectrum. Masses are compared to calculated masses for characterized product compositions and assigned as consistent with any calculated mass within a measurement uncertainty window. The default uncertainty tolerance is 70 parts per million (ppm) on either side of the calculated mass (this is equivalent to about 2.1 Da for a 30,000 Da piece of DNA, or a 100 base single DNA strand).

For a final mtDNA product to be assigned and exposed to the analysis interface for a given primer pair, there must be a mass consistent with a composition possible for the forward primer of the primer pair, and a complementary base composition consistent with the reverse primer of the pair. The A's, G's, C's and T's of one strand must equal the T's, C's, G's and A's of the complementary strand (in that order). An additional

constraint is imposed that limits the discrepancy in differences between calculated and measured masses for forward and reverse strands of a product. The current default is an 80 ppm maximum difference between the calculated vs. measured mass discrepancies. For example, if a measured mass is 25 ppm less than the calculated mass for a particular DNA product, and another mass is 60 ppm higher than the calculated mass for a product complementary in base composition to the first product, the difference is 85 ppm, and no double stranded product will be assigned. In addition, the signal intensities of the two strands of a double-stranded product must be within a defined ratio of each other (the current default is that the more intense peak cannot be more than 2.5 times the signal of the less intense peak).

In the event that no product is assigned for a primer pair, the detected masses are automatically analyzed for any pair of masses that could produce a double-stranded product within any combination of $\pm$ 1 A, G, C and/or T from a previously characterized mtDNA product. If no product is assigned for two masses of similar intensity, and those masses appear to potentially be a pair associated with a double-stranded DNA product, the interface offers a mechanism to manually interrogate the mass spectrometry data for the presence of double-stranded products.

Automated profile development

Once products have been assigned for each primer pair in each PCR reaction, a composite base composition profile is assembled for the sample, ordered by product coordinates relative to the rCRS. From the addition of DNA sample to a pre-fabricated assay kit, all steps through thermocycling (amplification), desalting, mass spectrometry, data processing, mass assignment, and initial profile development are automated. A schematic overview of the assay setup and analysis process is shown in Figure 3 of the introduction. It is after the automated development of an initial profile that the interface is turned over to the user for a visual verification and manual interaction if necessary.

Manual mass assignment

There are cases where a mass appears to be present in a spectrum, but was not called automatically by the data processing software, or the mass was assigned incorrectly due to a misshaping of the deconvolved peak or a merging of multiple adjacent mass peaks. In these cases, it is possible to interrogate the deconvolved data with an interactive viewer and add a mass assignment according to user judgment. There is a software-assisted peak-fitting incorporated into the deconvolved data viewer that uses a quadratic maximization similar to that used by the automated peak assignment, but directed to a data window by the user using the mouse. There are also built-in graphical features that allow browsing possible products for any given primer pair and to look for artifacts such as salt adducts and adenylated peaks (or non-adenylated peaks in the case of the STR assay).

Product filtering

Although rare, it is possible that a product could be assigned within allowed tolerance levels that can be identified by a trained user as incorrect or questionable. The analysis interface provides a mechanism to manually filter an assignment that is deemed incorrect or is questionable such that it would be preferable to eliminate data for a primer pair rather than chance an incorrect assignment.

Non-templated adenylation

Amplitaq Gold (the polymerase enzyme used in the mtDNA tiling assay) is the DNA polymerase from the bacterium *Thermus aquaticus*. This enzyme has an inherent terminal deoxyribonucleotide transferase activity that has a tendency to add a non-templated nucleotide to the end of a synthesized DNA strand, one base past the last templated addition. The activity strongly favors addition of a single adenosine. It has been empirically observed that this activity is strongly enhanced in elevated magnesium and strongly reduced in low magnesium. Also, experience shows that the activity is greatly reduced when a synthesized DNA strand naturally ends in an adenosine (Taq polymerase does not appear to favor adding an ‘A’ to the end of an existing ‘A’). All primers for the mtDNA tiling assay have a ‘T’ at the 5′ end, which causes all amplified

strands to end in an ‘A’. Also, the magnesium concentration used is 1.5 mM, which helps to keep adenylation levels low. Never-the-less, some level of non-templated adenylation will be observed. Conversely, in the STR assay, products are driven to full adenylation by addition of a ‘G’ or ‘C’ and running the assay in 3 mM $MgCl_2$. When this assay is analyzed, software tools will automatically expose the option to look for non-adenylated artifacts.

Salt Adducts

Crude PCR products must be stripped of salt ions (primarily potassium and magnesium from the PCR reaction) prior to mass spectrometry. If salt is introduced into the mass spectrometer with the DNA, it will adhere to the DNA backbone randomly, causing multiple peaks for each specific PCR product and diluting signal for each peak. This complicates the spectrum and can cause problems in sample analysis.

Despite the fact that there is no sodium in the PCR reaction mix used for the mtDNA tiling assay, sodium adducts are the primary problem in rare cases of heavy salt adduction. Sodium ions can adhere strongly to DNA, forming sodium-adducted peaks that appear at ~22 Da intervals. Again, even though there is no sodium in the PCR reactions buffer or clean-up buffers, and the appearance of large adducts is rare, it does happen occasionally and can be easily recognized. If adduct level are low, they generally will not interfere with confident product assignments. If adduction levels approach the levels of non-adducted products in a well, they can clutter the spectrum, remove amplitude from the correct signals, and look like viable product peaks to the analysis software, forcing a repeat of the affected sample(s). Sodium adducts could presumably occur if successfully desalted DNA is introduced into a recipient assay plate that has had sodium introduced prior to use (e.g. from dust, dander, oil from skin, sweat, possibly at the plate production facility), or if sodium got into a fluid line.

In the case of compromised desalting, or if the desalting system needs cleaning or maintenance, potassium adducts are more likely, as they are present in the PCR buffer. Potassium adducts will present peaks at intervals of ~39 Da above the primary peak,

but otherwise appear much like sodium adducts. Sodium or potassium adducts can be recognized in the spectral thumbnails as fuchsia or orange peaks, respectively, and in the spectral data viewer using the base composition browser.

We have developed a full software user manual for the mitochondrial control region tiling assay. A software manual will be developed as well for the STR assay when development and optimization of the assay and the analysis software are further established.

STR-specific software features.

When an STR assay is analyzed in the IbisTrack interface, products are assigned and labeled with standard STR allele nomenclature, as opposed to the base composition profile list developed for the mtDNA tiling assay. An example of the interface for STR analysis is shown in Figure 38. A dynamic summary of profile assignments for samples on an assay plate is available in the "Plate summary" tab shown in Figure 39. Figure 39 shows a profile summary table for nine samples run on one plate along with one extraction blank, one general negative control, and one positive control (SC35495).

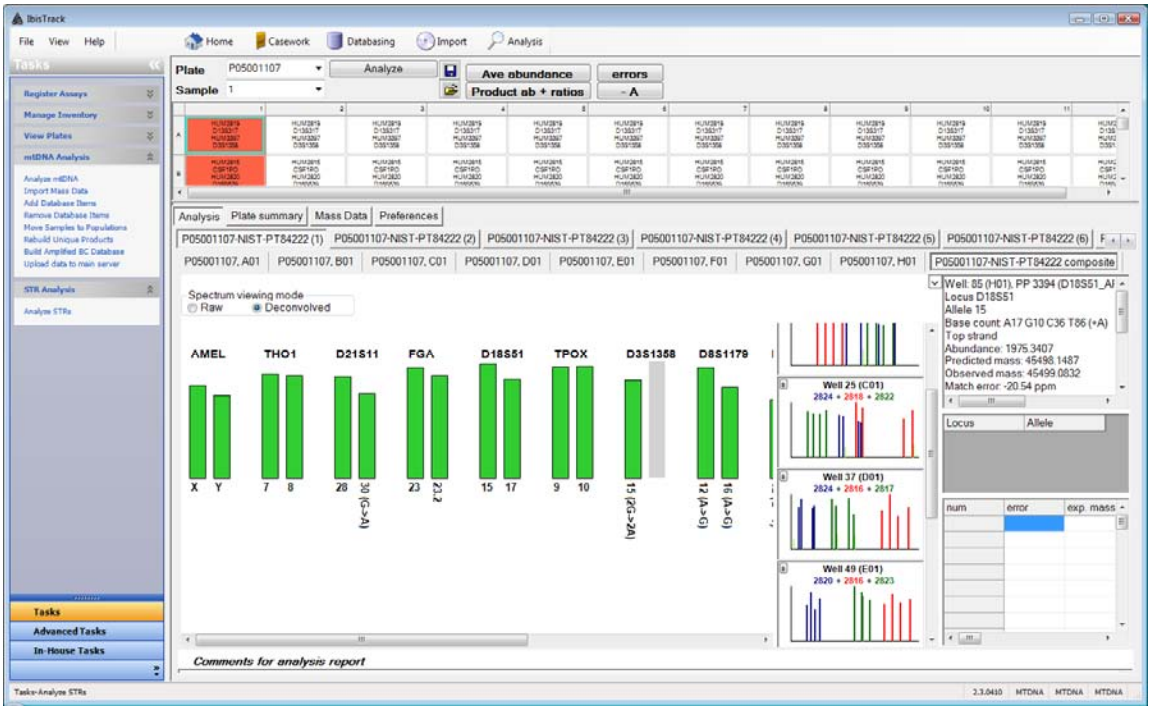


Figure 38. STR analysis interface. The interface displays allele assignments by locus that can be interrogated by clicking on the green assignment bars. Like the mtDNA analysis interface, the assigned products are integrated with the mass data views. Moving the mouse over an allele cartoon in the main window highlights both associated (forward and reverse strand) mass peaks in the associated spectral thumbnail(s) on the right. The same mass data viewer and features used in for the mtDNA tiling analysis are available for STR analysis.

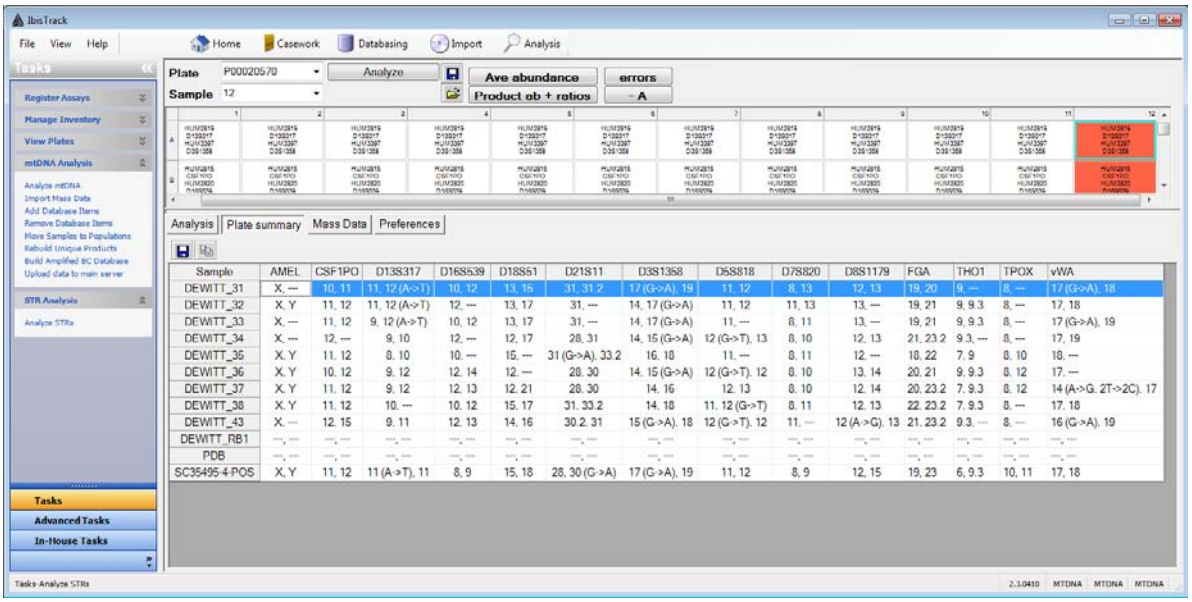


Figure 39. Profile summary for an STR assay plate. The above example shows nine samples, one extraction blank, one negative control and one positive control run an STR assay plate. Note that proper assignment of polymorphic alleles is part of the automated assignment process.

Specific Aim 3. Implementation of ESI-MS forensics methods on automated bench top platform.

The hardware platform used throughout these studies is an automated mass spectrometry-based platform named the Ibis T5000. The Ibis T5000 takes PCR generated amplicons generated off-line, performs a magnetic bead-based desalting/purification step, then injects purified aliquots of PCR amplicon into an automated electrospray ionization time-of-flight mass spectrometer (ESI-TOF). Mass spectrometry is used to determine the base composition (the number of adenosines (A), cytidines (C), guanosines (G) and thymidines (T)) of the PCR amplicon. In T5000 applications of pathogen detection, the PCR primers employed generate amplicons from bacteria or viruses of interest and the base compositions derived from multiple amplicons are used to identify the pathogens, strain type the pathogens, and/or provide information about the drug resistance and virulence properties of the pathogen(s). In forensics applications, the base compositions derived from PCR primers targeting mtDNA (HV1/HV2) segments, or CODIS markers are used generate forensic signatures. The analytical methods employed to analyze PCR products by high performance ESI-MS were originally developed under DARPA funding with a focused application of bioweapons surveillance. More recently, as the applications of the Ibis technology have expanded into clinical applications, food safety, pharmaceutical process control, and human forensics.

The original Ibis system was developed using DARPA funding and was designed for high-throughput analysis of air samples for biological pathogens. The system was comprised of two HEPA-enclosed decks. One deck effected automated nucleic acid extraction using conventional Qiagen methodologies, while the second deck performed PCR amplification, amplicon purification, and mass spectrometric analysis.

While these decks are fully capable of executing the missions they were designed to fulfill, their large footprints, high cost, and inflexible work flow were not conducive to deployment in clinical settings, and are clearly incompatible with spatial constraints and

workflow requirements of a forensics laboratory. Accordingly, we developed a commercial bench top version of the Ibis technology platform called the Ibis T5000 (Figure 40). This configuration is presently deployed to a number of government research laboratories including the CDC (Centers for Disease Control and Prevention, Atlanta, GA), USAMRIID (United States Army Institute for Infectious Disease, Ft. Dietrich, MD), DHS (Department of Homeland Security, Silver Springs, MD), DoD (Naval Health Research Center, San Diego, CA), PFPA (Pentagon Force Protection Agency, Alexandria, VA), and the FBI (Federal Bureau of Investigation, Quantico, VA). This configuration leverages existing laboratory infrastructure (both personnel and equipment) and does not monopolize the nucleic acid extraction components or PCR thermocyclers. The system is capable of analyzing 1536 PCR wells in 24 hours which, at 8 wells/assay, is equivalent to nearly 200 forensic samples/day.

Figure 40. The Ibis T5000 performs fully automated desalting/purification and electrospray ionization mass spectrometry of PCR-generated amplicons. This platform was used for both the mtDNA studies and the STR studies during the course of the project (See text).

Based on the maturity of the T5000 platform and the infrastructure Ibis has built to manufacture and support the instrument, we opted to run all studies on the standard T5000 platform. While the basic run mode of the T5000 (described below) is the same whether one is analyzing for infectious diseases or performing forensic profiling of a human DNA sample, a number of software modifications, specific to the forensics assays, were effected during the course of this grant under NIJ funding (see section 2).

The DNA II Unit of the FBI laboratory in Quantico (through funding independent of this grant) acquired the T5000 platform during the period of performance of this grant. This afforded the opportunity for us to collaborate with FBI researchers at a significantly more detailed level than originally anticipated and allowed us to transfer SOPs, reagents, QC standards, and software to an outside laboratory for validation. These frequent scientific exchanges with the FBI allowed the Ibis forensics team to customize software functionality, report formats, and standard QC protocols to be compatible with the established workflow of an accredited forensics laboratory.

T5000 Instrumentation Description

Plate storage, plate sealing, and plate handling robotics Prior to ESI-MS analysis, each plate of amplified material must be rigorously desalted to avoid inlet contamination and cation-adduction in the mass spectra. For each previously amplified 96-well plate, both a 96-well desalting plate (containing a magnetic bead-based DNA-binding resin), and a clean 96-well “elution” plate are loaded onto the system. The T5000 utilizes a ThermoElectron Catalyst Express robotic arm which has a total plate storage capacity of (45) 96-well plates, thus in this configuration can accommodate up to fifteen (15) 96-well PCR plates for desalting and sampling into the ESI-TOF. All liquid-containing plates are heat sealed prior to loading on the storage robot. The T5000 utilizes a WASP automated heat sealer capable of applying heat seals to plates after they have been pierced. Heat sealing mitigates potential amplicon contamination and minimizes solvent evaporation. The plate handing robot is capable of shuttling plates from the plate storage locations to the automated PCR purification station, the automated heat sealer, and the fluidic module which injects samples into the ESI-TOF.

Automated PCR product purification Following PCR amplification, 96-well plates containing amplicon mixtures are rigorously desalted using a protocol based on the weak anion exchange (WAX) method published elsewhere[3]. A custom 8 channel, fixed needle liquid handling robot based on a mini LEAP autosampler is used to dispense and aspirate rinse buffers and move sample aliquots. Initially, aliquots of PCR solutions are transferred to the desalting plate where the amplicons are bound to a weak anion exchange resin. Unconsumed dNTP's, salts, polymers, and other low MW species that might interfere with subsequent ESI-MS analysis are removed by rinsing the resin with solutions containing volatile salts and organic solvents. Elution of the final purified/desalted PCR products is accomplished by rinsing the resin with an aliquot (typically 25 μL) of a high pH buffer. For optimal calibration of the mass spectrometer, internal mass standards (that bracket the m/z range of the charge state envelope of PCR products) are added to the elution buffer to allow internal calibration of each mass spectrum. An entire 96-well PCR plate is desalted at a rate of 1 plate in < 45 minutes.

Sample injection fluidics module Unlike most of the upstream processing steps, the ESI-TOF is limited to a single channel for sample analysis (i.e. only one PCR well is analyzed at a time). To maximize the number of samples which can be processes through the ESI-TOF it is coupled with a custom fluidics module based on Cavro programmable pumps and switching valves which is tuned to maximize the data acquisition time and minimize the dead time between samples. Samples are processed at a rate of ~59 seconds/sample with 10 seconds of overhead between each sample resulting in an 85% duty cycle. Between each 20 μL sample injection a bolus of rinsing solvent is rapidly injected through the sample transfer line to minimize carryover between samples. After the samples are injected the flow rate is quickly reduced to be compatible with the ESI sprayer (200 μL/hr). After the data acquisition is complete the ESI voltages are quickly turned off and the nebulizer gas pressure is reduced prior to rinsing and sample injection to ensure the high flow of solvents does not comprise the vacuum integrity of the ESI source which can produce elevated background pressure in the ESI-TOF mass spectrometer.

ESI-TOF Mass Spectrometry A Bruker Daltonics (Billerica, MA) MicrOTOF ESI time-of-flight (TOF) mass spectrometer is a key component of the T5000. Ions from the ESI source undergo orthogonal ion extraction and are focused in a reflectron prior to detection. Ions are formed in the standard MicroTOF ESI source which is equipped with an off-axis sprayer and glass capillary. We have tailored solution and interface conditions to facilitate the ionization of the intact PCR amplicons with desolvation conditions which effect strand separation in the gas phase without cleavage of the products. We have published details of these parameters previously[1, 4, 5, 17].

3.1 Automation Stress Testing.

The automation methods and component functionality core to the T5000 were routinely tested as part of a normal system QC runs performed over the course of the project. We found frequent analysis of system QC plates (described in detail in section 4.2) containing a dilution to extinction series (500 pg, 250 pg, 100 pg, 25 pg, 10 pg, 4 pg, 1 pg, 0 pg) of DNA from the standard SC35495 extract, an excellent metric with which to assess overall system performance. 37 QC plates were ran between Jan 2007 and Jan 2008 which comprised 3,552 PCR wells that were desalted, injected, and analyzed by ESI-MS on the T5000 instrument. Of these 3,552 analyses, 3 wells failed due to problems with the magnetic bead desalting module (e.g. air bubbles in the fluid lines, debris in the wells, etc) which were manifested as salt adducted or unusable mass spectra, and 3 wells failed due to problems with the injection (e.g. air bubbles in the fluid lines, debris in the wells, etc) which were manifested as “blank” spectra. These failures resulted in the Failure of 2 QC plates. These data suggests a 99.83% automation success. Similarly, as previously described in section M.1.3, of the 2,207 replicates of Seracare sample SC35495, of the 52,968 expected calls (i.e. 2,207 samples x 24 primer pairs/sample), there were 159 missed calls or 0.3% suggesting a 99.70% automation success.

3.2 Automation and Analytical Reproducibility.

As described above in section M.1.7, mass measurement error and precision were determined through 105,580 independent mass measurements derived from 2,207 trials of the Seracare SC35495 sample run through the mtDNA tiling assay. Observed mass measurement deviations were distributed around zero and did not exhibit a systematic bias toward measurement errors lower or higher than the true product masses. These data support the hypothesis that the internal calibration scheme employed is robust over long periods of time and negates potential adverse affects of changes in environmental parameters such as room temperature, line voltage, or relative humidity. The average (absolute value) of the mass measurement deviation for 105,580 measurements was 10.64 $\pm$ 8.2 ppm (see section M.1.3 and associated figure). Importantly, of the 35 mtDNA QC plates that passed QC over the course of the project, the average input DNA level at which a full profile was generated was 7.4 $\pm$ 4 pg/reaction. With the exception of 2 plates that yielded a full Seracare SC35495 profile at 25 pg/reaction, all of these plates yielded a full (and correct) mtDNA profile at 10 pg/reaction or less, with 11 and 6 plates yielding a full and correct profile at 4 pg and 1 pg of input material/reaction, respectively. This high level of reproducibility speaks not only to the robustness of the automation associated with amplicon desalting and mass spectrometric analysis, it suggests that the process used to fabricate the mtDNA assay plates and mtDNA QC plates is efficient and reproducible (see section 4).

Specific Aim 4. Kitting of amplification and cleanup reagents

4.1 Base reagent preparation

All critical reagents for the mitochondrial assay are independently tested for performance and template contamination prior to kitting. One of the most important critical reagents is the primer pairs and subsequent triplexes that are used in the mitochondrial assay. This process is thoroughly document with SOPs for making the primer pairs and triplexes as well as test methods for quality control (QC) of the material at critical steps. The first step in making primer pairs is the resuspension of individual primers to 750 μM based on vendor provide quantification and then quantification by UV spectrometry. Accurate quantification is necessary to ensure that the primer pairs are mixed together at the same concentration. The primers are then diluted to a uniform concentration of 500 μM based on the UV quantification. Thus, when the primer pairs are mixed one –to –one, the resulting solution is 500 μM total (250 μM each primer). According to an appropriate test method, the primer pairs are tested for performance and contamination using our current positive control of Seracare blood sample SC35495. This step is designed to make sure that the intra-strand ratios of the PCR products are within the defined tolerances as well as looking for performance and contamination. In several instances, primer pairs were found that were contaminated or that did perform PCR at all. These primer pairs were destroyed and not used in the next step. Once all twenty-four primer pairs pass QC, the primer pairs are then combined into their appropriate triplexes and undergo performance and contamination testing again. In addition, the product balance in each triplex is also examined using the positive control. To test the performance of the triplexes, PCR reactions are performed using 25 pg, 100 pg, and 500 pg of input DNA. Full correct profiles must be obtained for all three dilutions in order for the triplexes to pass QC. In addition, three replicate PCR reactions with no template added must show no contamination in order for the triplexes to pass QC.

The next critical reagent is the dNTP mix used for PCR. We make our own dNTP mix because we use a carbon 13-enriched dGTP in the mitochondrial assay. Since the

dGTP is enriched in carbon 13, the percent enrichment needs to be determined prior to its use in kitting plates. In addition, the dNTP mix must past performance and contamination criteria. Using previously passed triplexes, three replicate no-template-added PCR reactions are used to test for contamination. To test the performance of the dNTP mix, PCR reactions are performed using 25 pg, 100 pg, and 500 pg of input DNA. These replicate reactions are also used to determine the percent enrichment of the carbon 13-enriched dGTP.

A third critical reagent is the PCR master mix base used in kitting the plates. The PCR master base contains all the reagents necessary for PCR except for the dNTPs, the triplexes and the Taq enzyme. This reagent is tested for contamination and performance in the same manner as the triplexes and the dNTP mix using previously passed lots of triplexes and dNTP mix. In addition, the PCR product balance within a triplex is examined. Various lots of PCR master mix base have been shown to influence the product ratios and thus, the product ratios must be within acceptable tolerances in order for the PCR master mix base to pass QC.

Each lot of Taq enzyme is also tested for contamination and performance. To date, the Taq enzyme has always meet specifications. The 96-well PCR plates that are used for kitting the reagents are UV treated for 10 minutes in a UV cross linker oven to minimize contamination in the plates. The plates are UV treated prior to the kitting and sealed with a removable heat seal that allows the plates to be stored contamination-free.

Once all of these reagents pass QC, kitted plates are made using a closed system to minimize the possibility of outside contamination. Typical kitting runs produce 400 plates. In process QC is performed on the first and every one hundredth plate to ensure that plate volumes are in acceptable range. Prior to dispensing reagents, the weight of the empty plate is recorded. After dispensing the reagents, the weight of the plate is recorded again. The difference between the two weights is divided by 96 and multiplied by 1000 to give an average volume per well per plate. If the average volume per well does not meet specifications, the dispenser is adjusted and tested as above until the

volume is within the acceptable range. These same plates that are used for in-process QC are used for functional and contamination QC. For each first and one hundredth plate, performance and contamination is evaluated by putting primer dilution buffer in columns 1, 2, 5, 6, 9 and 11 as a no-template-added control as well as 500 pg of the positive control in columns 3, 4, 7, 8, 10 and 12. These plates indicate how reproducible the PCR is across the plate as well as test for contamination in the kitting process. Full correct profiles must be obtained for all positive controls and the negatives must be free of contamination in order for the plates to pass. In addition the last plate of the batch is used to test for sensitivity. The sensitivity is tested by using 1 pg, 4 pg, 10 pg, 25 pg, 100 pg and 500 pg inputs of the positive control DNA. A full correct profile must be obtained down to 25 pg of input DNA in order for the plates to pass QC. Typically, full correct profiles are observed between 4 pg and 10 pg of input DNA.

The above processes are all documented and the methods are in the process of being transferred to our manufacturing facilities. Once the formulation for the STR assay has been finalized, the same processes employed to quality control the reagents and kitted plates for the mitochondrial assay will be applied to this assay.

4.2 Preparation of mtDNA kits

Over the course of this grant, ~175 kitted mitochondrial plates were made to support the aims. All of these plates were kitted using the processes described in section 4.1. Seventy of the kitted plates were sent to the FBI for validation and testing on their instruments. Theses plates were also used to conduct the mixture studies described in section 1.6. In addition, 15 system validation plates were sent to the FBI for diagnosing issues with and performance checks of their instruments. The system validation plate sample layout is shown in Figure 41. The 5 samples with 500 pg of input DNA are meant to test the reproducibility of the instrument while the lower amount of input DNA samples test the sensitivity of the instrument. Additionally, the presence or absence of salt adducts on the PCR products checks the cleanup module of the instrument. The system validation plates are run through PCR at Ibis and sent to the customer as

needed for instrument checks. Twenty of the kitted plates were reserved for stability test (see section 4.4).

S-1	S-2	S-3	S-4	S-5	S-6
PDB 1 DF	SC35495-10 1 pg	SC35495-10 4 pg	SC35495-10 10 pg	SC35495-10 25 pg	SC35495-10 100 pg
S-7	S-8	S-9	S-10	S-11	S-12
SC35495-10 250 pg	SC35495-10 500 pg	SC35495-10 500 pg	SC35495-10 500 pg	SC35495-10 500 pg	SC35495-10 500 pg

Figure 41. The samples layout for system validation plates. Each sample consists of 8 wells on a PCR plate.

4.3 Preparation of STR kits

Over the course of this grant, ~36 plates were made to test and validate the STR assay. While the primer pairs and triplex reactions have been finalized, work is continuing on optimizing the PCR master mix base for this assay to achieved optimum allele balance and sensitivity. Once the PCR master mix base has been finalized, stability studies and larger kitting runs will be done.

4.4 Stability testing

For the mitochondrial tiling assay, a batch of plates was made on October 6, 2006 and 20 pates were set aside for stability testing. The plates were stored in a walk-in freezer at -20°C in the cardboard boxes. The temperature of the walk-in freezer is monitored to make sure the temperature stays within an acceptable range. The system validation plate sample layout described above in section 4.2 was used to test the performance of the plates. The plates were tested at 4-6 weeks intervals for the first 6 months and then the plates were tested every 12 weeks. To demonstrate that the plates are still performing as expected, a correct full profile must be obtained down to at least 25 pg input DNA. Table 19 lists the PCR barcode tested, the ESI plate barcode used for mass spectrometric analysis, the date tested and the results. All of the plates tested met specifications and passed.

Table 19. Summary of stability data for kitted mitochondrial tiling assay plates.

PCR Barcode	ESI Barcode	Storage @ -20C	Date tested	Pass/Fail
C00009621	P00006639	4 Days	10/10/2006	Pass
C00009622	P00006640	4 Days	10/10/2006	Pass
C00009641	P00006675	12 Days	10/18/2006	Pass
C00009683	P00006735	5 weeks	11/10/2006	Pass
C00009733	P00006770	9 weeks	12/4/2006	Pass
C00009796	P00006849	13 weeks	1/3/2007	Pass
C00009797	P00006857	13 weeks	1/3/2007	Pass
C00009803	P00008100	17 weeks	2/6/2007	Pass
C00009804	P00008280	22 weeks	3/9/2007	Pass
C00009806	P00008518	28 weeks	4/20/2007	Pass
C00009805	P00010888	40 weeks	7/13/2007	Pass
C00009807	P05000080	52 weeks	10/11/2007	Pass

The intra-strand ratio for primer pair 2891 was also monitored as a function of both amount of input DNA and storage time (Figure 42). The graph shows that the ratio is stable over the input DNA range of 25 pg to 100 pg. The ratio will fluctuate more at the lower in of the amount of input DNA. In addition, one plate was stored at room temperature for 72 hours before being used for PCR. The system validation sample layout was used. The plate passed, giving a full correct profile down to 25 pg input DNA. Although a more detailed study is necessary, this indicates that plates accidently left out on the bench top will still provide optimal performance. Over all, the plates with all the reagents necessary to perform PCR are stable for at least one year. The stability study is ongoing with the goal of having real time stability data for 2 years.

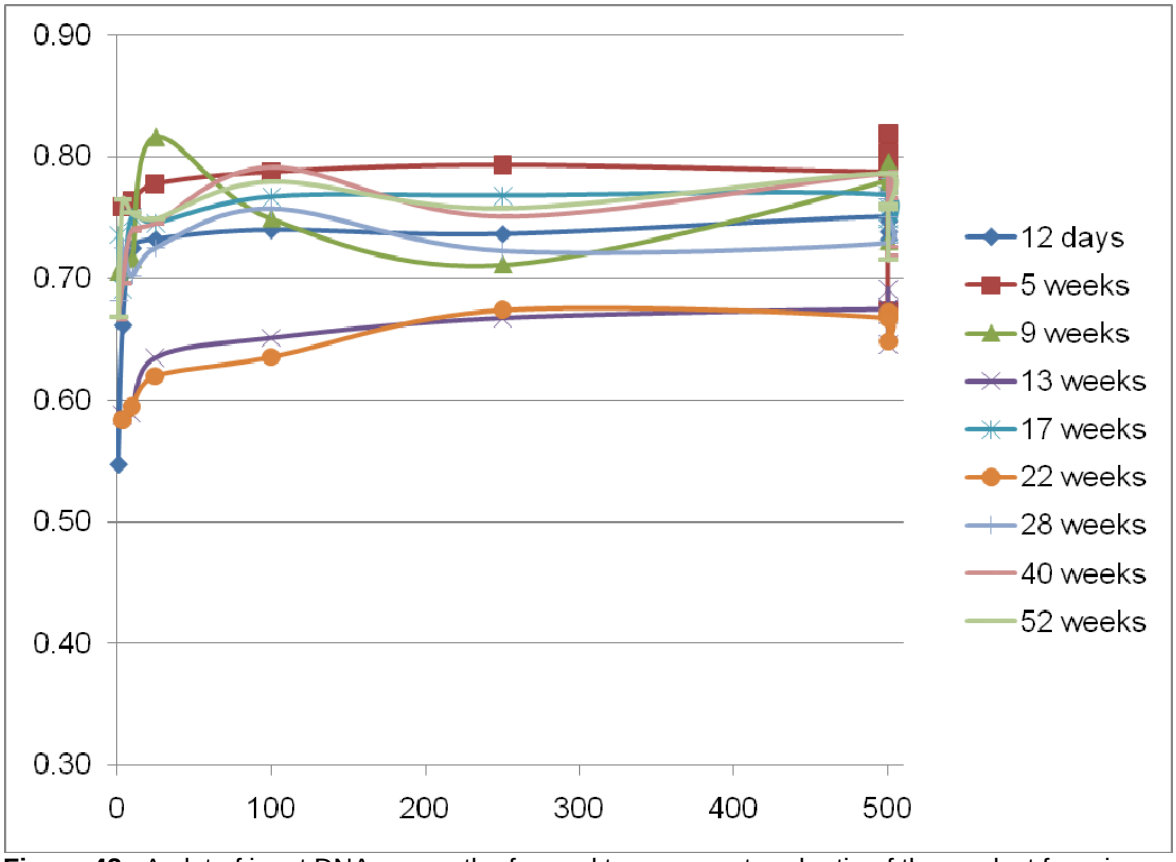


Figure 42. A plot of input DNA versus the forward to reverse strand ratio of the product for primer piar 2891 as function of storage time of the plates. The ratio is stable over the input range of 25 pg to 500 pg.

4.5 Cleanup reagents

The PCR products need to be rigorously desalted and purified before mass spectrometric analysis. Cleanup reagents are provided as part of the mitochondrial tiling assay kit and work with the cleanup module of the Ibis T5000. The reagents are magnetic bead plates that are used to bind the PCR products during cleanup, cleanup reagent 1 which is used to remove salts from the PCR products, cleanup reagent 2 which is used to remove salts as well as polymers and cleanup reagent 3 which is used to elute the PCR product and to spray the cleaned PCR product. All of the reagents are tested prior to release for kitting. Initially, the reagents are test for performance with a Standard QC plate. In addition to verifying that the correct products are obtained without any cleanup failures, the salt adducts are monitored and must be below 10% of

the parent peak in order to pass QC. Following this initial QC, the reagents are then tested using a pre-PCR'd system validation plate to ensure that a full correct profile is obtained down to 25 pg and that salt adducts do not cause any additional products to be called. These two QC steps ensure that only quality reagents are used for the Ibis T5000 cleanup module.

Dissemination

Since the beginning of this effort, we have collaborated closely with members of the DNA Unit II Laboratory division at the FBI, Quantico, VA, in a successful effort to transfer this technology into the hands of forensic scientists outside of our own laboratory. The FBI DNA Unit II (mitochondrial analysis division) has purchased two Ibis T5000 instruments and multiple mtDNA tiling kits for use with the mitochondrial control region tiling assay. We have held four hands-on training sessions with FBI staff members and have transferred detailed SOPs and full software user's manuals to FBI forensics scientists. We have also presented portions of this work at multiple conferences, workshops and seminars and are preparing manuscripts to introduce the forensics community to our mitochondrial profiling assay and polymorphism-revealing STR assay.

Seminars, workshops and conferences where portions this work have been presented:

1. Electrospray-Ionization Mass Spectrometry for Utilization of Sequence Variation in Human Short Tandem Repeats. Thomas A. Hall, Kristin A. Sannes-Lowery, Amy S. Schink, Theodore D. Anderson, and Steven A. Hofstadler. Presented at the American Academy of Forensic Sciences 59th Annuall Meeting in San Antoni, Texas, February 19-24, 2007.
2. High-throughput Analysis of Amplified Nucleic Acids with Mass Spectrometry: Applications in Pathogen Detection, Epidemiology and Human Forensics. Steven A. Hofstadler. Presented at Washington State University, March 26, 2007.
3. High-throughput Analysis of Amplified Nucleic Acids with Mass Spectrometry: Applications in Pathogen Detection, Epidemiology and Human Forensics. Steven A. Hofstadler. Presented at University of California, Davis, March 28, 2007.

4. Electrospray-Ionization Mass Spectrometry for Elucidation of Sequence Variation in Human Short Tandem Repeats. Thomas A. Hall, Kristin A. Sannes-Lowery., Amy S. Schink, Theodore D. Anderson and Steven A. Hofstadler. Presented at the 55th ASMS Conference on Mass Spectrometry, American Society for Mass Spectrometry, Indianapolis, Indiana, June 3-7, 2007.
5. High-throughput Analysis of Amplified Nucleic Acids with Mass Spectrometry: Applications in Human Forensics. Steven A. Hofstadler. Presented at the NIJ Conference, July 23, 2007.
6. Multiplex PCR electrospray-ionization mass spectrometry (ESI-MS): application to forensic mitochondrial DNA examinations. Leslie D. McCurdy, Lora J. Gioeni, Thuy-Trang Pennella, Constance L. Fisher, Alice R. Isenberg, Thomas A. Hall, Kristin A. Sannes-Lowery, Steven A. Hofstadler, and Bruce Budowle. Presented at the 22nd Congress of the International Society for Forensic Genetics. 21–25 August 2007 - Copenhagen – Denmark.
7. High-throughput Analysis of Amplified Nucleic Acids with Mass Spectrometry: Applications in Human Forensics. Steven A. Hofstadler. Presented at the DNA Forensics Working Group Fall 2007 meeting in Las Vegas, NV, September 24-25, 2007.
8. Mitochondrial DNA Mixture Resolution by Base Composition Analysis. Leslie D. McCurdy, Thomas A. Hall, Thuy-Trang Pennella, Lora J. Gioeni, Constance L. Fisher, Kristin A. Sannes-Lowery, Alice R. Isenberg, Steven A. Hofstadler, and Bruce Budowle. Presented at the Promega 18th International Symposium on Human Identification, Hollywood, CA. October 1-4, 2007.
9. High-throughput Analysis of Amplified Nucleic Acids with Mass Spectrometry: Applications in Pathogen Detection, Epidemiology and Human Forensics. Steven A. Hofstadler. Presented at Marshall University, Huntington, WV, December 13, 2007.

References

(1) Hofstadler, S. A.; Sampath, R.; Blyn, L. B.; Eshoo, M. W.; Hall, T. A.; Jiang, Y.; Drader, J. J.; Hannis, J. C.; Sannes-Lowery, K. A.; Cummins, L. L.; Libby, B.; Walcott, D. J.; Schink, A.; Massire, C.; Ranken, R.; White, N.; Samant, V.; McNeil, J. A.; Knize, D.; Robbins, D.; Rudnik, K.; Desai, A.; Moradi, E.; Ecker, D. J. *Inter. J. Mass Spectrom.* **2005**, *242*, 23-41.

(2) Ruitberg, C. M.; Reeder, D. J.; Butler, J. M. *Nucleic Acids Res* **2001**, *29*, 320-322.

(3) Jiang, Y.; Hofstadler, S. A. *Anal Biochem* **2003**, *316*, 50-57.

(4) Ecker, D. J.; Drader, J.; Gutierrez, J.; Gutierrez, A.; Hannis, J.; Schink, A.; Sampath, R.; Ecker, J. A.; Blyn, L. B.; Eshoo, M. W.; Hall, T. A.; Tobarmosquera, M.; Jiang, Y.; Sannes-Lowery, K.; Cummins, L.; Libby, B.; Walcott, D. J.; Massire, C.; Ranken, R.; Manalili, S. M.; Ivy, C.; Melton, R.; Levene, H.; Harpin, V.; Li, F.; White, N.; Pear, M.; Samant, V.; Knize, D.; Robbins, D.; Rudnick, K.; Hajjar, F.; Hofstadler, S. A. *JALA* **2006**, *11*, 341-351.

(5) Hall, T. A.; Budowle, B.; Jiang, Y.; Blyn, L.; Eshoo, M.; Sannes-Lowery, K. A.; Sampath, R.; Drader, J. J.; Hannis, J. C.; Harrell, P.; Samant, V.; White, N.; Ecker, D. J.; Hofstadler, S. A. *Anal Biochem* **2005**, *344*, 53-69.

(6) Brownstein, M. J.; Carpten, J. D.; Smith, J. R. *Biotechniques* **1996**, *20*, 1004-1006, 1008-1010.

(7) Meyers-Wallen, V. N.; Palmer, V. L.; Acland, G. M.; Hershfield, B. *Mol Reprod Dev* **1995**, *41*, 300-305.

(8) King, V.; Goodfellow, P. N.; Pearks Wilkerson, A. J.; Johnson, W. E.; O'Brien, S. J.; Pecon-Slattery, J. *Genetics* **2007**, *175*, 1855-1867.

(9) Eichmann, C.; Berger, B.; Steinlechner, M.; Parson, W. *Forensic Sci Int* **2005**, *151*, 37-44.

(10) Menotti-Raymond, M. A.; David, V. A.; Wachter, L. L.; Butler, J. M.; O'Brien, S. J. *J Forensic Sci* **2005**, *50*, 1061-1070.

(11) Butler, J. M.; Schoske, R.; Vallone, P. M.; Redman, J. W.; Kline, M. C. *J Forensic Sci* **2003**, *48*, 908-911.

(12) Mornhinweg, E.; Luckenbach, C.; Fimmers, R.; Ritter, H. *Forensic Sci Int* **1998**, *95*, 173-178.

(13) Lazaruk, K.; Wallin, J.; Holt, C.; Nguyen, T.; Walsh, P. S. *Forensic Sci Int* **2001**, *119*, 1-10.

(14) Lins, A. M.; Micka, K. A.; Sprecher, C. J.; Taylor, J. A.; Bacher, J. W.; Rabbach, D. R.; Bever, R. A.; Creacy, S. D.; Schumm, J. W. *J Forensic Sci* **1998**, *43*, 1168-1180.

(15) Barber, M. D.; Parkin, B. H. *Int J Legal Med* **1996**, *109*, 62-65.

(16) Moller, A.; Meyer, E.; Brinkmann, B. *Int J Legal Med* **1994**, *106*, 319-323.

(17) Hofstadler, S. A.; Sannes-Lowery, K. A.; Hannis, J. C. *Mass Spectrom Rev* **2005**, *24*, 265-285.

Appendix A. GenBank accession numbers for human mtDNA genomes used to compare discriminating power between sequence and the Ibis mtDNA tiling assay.

DQ404447.3, DQ404446.3, DQ404445.3, DQ404444.3, DQ404443.3, DQ404442.3, DQ404441.3, DQ404440.3, AF346965.1, AF346964.1, AF346963.1, D38112.1, AF346967.1, AF346968.1, AF346977.1, AF346976.1, AF346980.1, AF346985.1, AF346987.1, AF346986.1, AF346992.1, AF346994.1, AF346995.1, AF346997.1, AF346996.1, AF346999.1, AF347000.1, AF347009.1, AF347008.1, NC_001807.4, AF347014.1, AY882407.1, AY195780.1, AY195777.1, AY195789.2, AY195783.2, AY195776.2, AY195788.2, AY195766.1, AY195785.2, AY195784.2, AY195782.2, AY950299.2, AY950297.2, AY950296.2, AY950293.2, AY950291.2, AY950289.2, AY950288.1, AY950287.1, DQ301818.1, DQ301816.1, DQ301815.1, DQ301813.1, DQ301812.1, DQ301810.1, DQ301809.1, DQ301808.1, DQ301807.1, DQ301806.1, DQ301805.1, DQ301803.1, DQ301802.1, DQ301800.1, DQ301799.1, DQ301797.1, DQ301796.1, DQ301795.1, DQ301794.1, DQ301793.1, DQ301792.1, DQ301791.1, DQ301790.1, AF346966.1, AY963572.1, AY963584.2, AY963576.2, AY963577.2, AY275529.1, AY195760.2, AY195770.2, AY195763.2, AY195753.2, AY195771.1, AY195791.2, AY195762.2, AY195755.2, AY195792.2, AY289067.1, AY289066.1, AY289065.1, AY289064.1, AY289063.1, AY289062.1, AY289061.1, AY289060.1, AY289059.1, AY289058.1, AY289057.1, AY289056.1, AY289055.1, AY289054.1, AY289053.1, AY289052.1, AY289051.1, DQ658411.1, AY495123.2, AY495324.1, AY495316.1, AY495312.1, AY495306.1, AY495302.1, AY495301.1, AY495293.1, AY495290.1, AY495287.1, AY495286.1, AY495284.1, AY495273.1, AY495265.1, AY495261.1, AY495260.1, AY495259.1, AY495257.1, AY495256.1, AY495254.1, AY495246.1, AY495245.1, AY495242.1, AY495238.1, AY495234.1, AY495231.1, AY495229.1, AY495223.1, AY495221.1, AY495219.1, AY495217.1, AY495208.1, AY495207.1, AY495204.1, AY495202.1, AY495195.1, AY495188.1, AY495187.1, AY495182.1, AY495175.1, AY495172.1, AY495166.1, AY495164.1, AY495154.1, AY495153.1, AY495152.1, AY495146.1, AY495144.1, AY495139.1, AY495125.1, AY495124.1, AY495109.1, AY495100.1, AY495096.1, DQ513522.1, AF346978.1, AF346973.1, AF346972.1, DQ462234.1, DQ462233.1, DQ462232.1, DQ418488.1, AY963575.2, AY963573.2, AY255139.1, AY255176.1, AY255158.1, AY255149.1, AY255147.1, AY255135.1, AY255178.2, AY255169.1, AY255180.1, AY255177.1, AY255163.1, AY255161.1, AY255160.1, AY255140.1, AY255136.2, AY255171.1, AY255156.1, AY255154.1, AY255152.1, AY255137.1, AY255179.1, AY255166.1, AY255155.1, AY255150.1, AY255173.1, AY255167.1, AY255157.1, AY255153.1, AY255138.1, AY255175.1, AY255172.1, AY255170.1, AY255168.1, AY255165.1, AY255164.1, AY255162.1, AY255151.1, AY255148.1, AY255141.1, AY255133.1, AY255159.1, AY255142.1, DQ519035.1, AY255146.1, AY972053.1, AY289069.1, AY289068.1, DQ372884.1, AF346974.1, DQ341073.1, DQ341072.1, DQ341071.1, DQ341062.1, DQ341059.1, DQ341058.1, AF346975.1, DQ272126.1, DQ272125.1, DQ272124.1, DQ272122.1, DQ272121.1, DQ272120.1, DQ272119.1, DQ272118.1, DQ272117.1, DQ272116.1, DQ272115.1, DQ272114.1, DQ272113.1, DQ272112.1, DQ272111.1, DQ272110.1, DQ272109.1, DQ272108.1, DQ272107.1, AY738973.1, DQ341063.1, DQ341082.1, DQ341081.1, DQ341080.1, DQ341079.1, DQ341077.1, DQ341076.1, DQ341074.1, DQ341069.1, DQ341068.1, DQ341067.1, DQ341066.1, DQ341065.1, DQ341064.1, DQ341061.1, DQ341060.1, AY882416.1, AY882389.1, AY882406.1, AY882403.1, AF347006.1, AY195773.1, AY195775.2, AY195758.2, AY195752.2, AY195746.2, AY195757.1, AY195751.1, AY195769.2, AY195778.2, AY195774.2, AY195754.2, AY195765.2, AY195756.2, AY195767.2, AY195745.2, AY195764.1, AY195781.2, AY195768.2, AY289070.1, AY339410.2, AY339581.2, AY339565.2, AY339564.2, AY339563.2, AY339562.2, AY339558.2, AY339557.2, AY339556.2, AY339554.2, AY339529.2, AY339513.2, AY339511.2, AY339510.2, AY339507.2, AY339506.2, AY339497.2, AY339495.2, AY339490.2, AY339489.2, AY339474.2, AY339472.2, AY339465.2, AY339464.2, AY339463.2, AY339460.2, AY339593.1, AY339592.1, AY339590.1, AY339589.1, AY339588.1, AY339586.1, AY339585.1, AY339582.1, AY339580.1, AY339579.1, AY339578.1, AY339577.1, AY339576.1, AY339575.1, AY339571.1, AY339570.1, AY339569.1, AY339568.1, AY339567.1, AY339566.1, AY339551.1, AY339550.1, AY339549.1, AY339548.1, AY339547.1, AY339546.1, AY339545.1, AY339544.1, AY339538.1, AY339536.1, AY339533.1, AY339531.1, AY339528.1, AY339527.1, AY339524.1, AY339523.1, AY339518.1, AY339502.1, AY339501.1, AY339500.1, AY339499.1, AY339459.1, AY339458.1, AY339457.1, AY339456.1, AY339452.1, AY339451.1, AY339450.1, AY339448.1, AY339445.1, AY339444.1, AY339443.1, AY339441.1, AY339438.1, AY339437.1, AY339433.1, AY339432.1, AY339431.1, AY339430.1, AY339429.1, AY339428.1, AY339427.1, AY339425.1, AY339424.1, AY339423.1, AY339421.1, AY339420.1, AY339419.1, AY339417.1, AY339415.1, AY339412.1, AY339411.1, AY339405.1, AY339403.1, DQ904242.1, AF382013.1, DQ408680.2, DQ408679.2, DQ408678.2, DQ408677.2, DQ408676.2, DQ408675.2, DQ408674.2, DQ408672.2, AY922308.1, AY922307.1, AY922306.1, AY922305.1, AY922304.1, AY922303.1, AY922302.1, AY922301.1, AY922300.1, AY922299.1, AY922298.1, AY922297.1, AY922296.1, AY922295.1, AY922294.1, AY922293.1, AY922292.1, AY922291.1, AY922290.1, AY922289.1, AY922288.1, AY922287.1, AY922286.1, AY922285.1, AY922284.1, AY922283.1, AY922282.1, AY922280.1, AY922279.1, AY922278.1, AY922277.1, AY922276.1, AY922275.1, AY922274.1, AY922273.1, AY922272.1, AY922271.1, AY922270.1, AY922269.1, AY922268.1, AY922267.1, AY922266.1, AY922265.1, AY922264.1, AY922263.1, AY922262.1, AY922261.1, AY922259.1, AY922258.1, AY922257.1, AY922256.1, AY922255.1, AY922254.1, AY922253.1, AY714050.1, AY714049.1, AY714048.1, AY714047.1, AY714046.1, AY714045.1,

AY714044.1, AY714043.1, AY714042.1, AY714041.1, AY714040.1, AY714039.1, AY714038.1, AY714037.1, AY714036.1, AY714034.1, AY714033.1, AY714032.1, AY714031.1, AY714030.1, AY714029.1, AY714028.1, AY714027.1, AY714026.1, AY714025.1, AY714024.1, AY714023.1, AY714021.1, AY714020.1, AY714019.1, AY714018.1, AY714017.1, AY714015.1, AY714014.1, AY714013.1, AY714012.1, AY714011.1, AY714010.1, AY714009.1, AY714008.1, AY714007.1, AY714006.1, AY714005.1, AY714004.1, AY714003.1, AY714002.1, AY714001.1, AY714000.1, AY713999.1, AY713998.1, AY713996.1, AY713995.1, AY713994.1, AY713993.1, AY713992.1, AY713991.1, AY713990.1, AY713989.1, AY713988.1, AY713987.1, AY713986.1, AY713985.1, AY713983.1, AY713982.1, AY713981.1, AY713979.1, AY713978.1, AY713977.1, AY713976.1, DQ246829.1, DQ246824.1, DQ246819.1, DQ246825.1, DQ246820.1, DQ246817.1, DQ246823.1, DQ246827.1, DQ246818.1, DQ246816.1, DQ246815.1, DQ246813.1, DQ246831.1, DQ246821.1, DQ246832.1, DQ246828.1, DQ246814.1, DQ246826.1, DQ246812.1, AF347010.1, AY738977.1, AY738961.1, AY738942.1, AY963586.1, AY882415.1, AY882413.1, AY882411.1, AY882410.1, AY882409.1, AY882402.1, AY882400.1, AY882399.1, AY882397.1, AY882395.1, AY882394.1, AY882393.1, AY882387.1, AY739001.1, AY739000.1, AY738999.1, AY738998.1, AY738997.1, AY738995.1, AY738994.1, AY738993.1, AY738990.1, AY738988.1, AY738986.1, AY738982.1, AY738980.1, AY738978.1, AY738975.1, AY738974.1, AY738972.1, AY738970.1, AY738969.1, AY738968.1, AY738967.1, AY738965.1, AY738964.1, AY738962.1, AY738960.1, AY738959.1, AY738958.1, AY738957.1, AY738956.1, AY738955.1, AY738953.1, AY738952.1, AY738951.1, AY738950.1, AY738949.1, AY738946.1, AY738945.1, AY738944.1, AY738943.1, AY738941.1, DQ341090.1, DQ341089.1, DQ341088.1, DQ341083.1, AP008259.1, AP008258.1, AP008257.1, AP008256.1, AP008255.1, AP008254.1, AP008253.1, AP008252.1, AP008250.1, AP008343.1, AP008341.1, AP008339.1, AP008338.1, AP008337.1, AP008336.1, AP008334.1, AP008329.1, AP008328.1, AP008327.1, AP008322.1, AP008320.1, AP008319.1, AP008317.1, AP008316.1, AP008315.1, AP008313.1, AP008311.1, AP008310.1, AP008308.1, AP008307.1, AP008306.1, AP008305.1, AP008304.1, AP008303.1, AP008299.1, AP008298.1, AP008297.1, AP008295.1, AP008294.1, AP008289.1, AP008287.1, AP008285.1, AP008284.1, AP008283.1, AP008282.1, AP008281.1, AP008279.1, AP008277.1, AP008275.1, AP008274.1, AP008273.1, AP008272.1, AP008271.1, AP008270.1, AP008269.1, AP008268.1, AP008267.1, AP008266.1, AP008261.1, AP008260.1, AP008823.1, AP008818.1, AP008816.1, AP008814.1, AP008813.1, AP008811.1, AP008809.1, AP008808.1, AP008806.1, AP008805.1, AP008804.1, AP008803.1, AP008801.1, AP008800.1, AP008798.1, AP008796.1, AP008795.1, AP008794.1, AP008793.1, AP008790.1, AP008789.1, AP008788.1, AP008787.1, AP008785.1, AP008784.1, AP008783.1, AP008782.1, AP008781.1, AP008780.1, AP008778.1, AP008777.1, AP008776.1, AP008774.1, AP008772.1, AP008771.1, AP008768.1, AP008767.1, AP008764.1, AP008760.1, AP008758.1, AP008756.1, AP008755.1, AP008753.1, AP008752.1, AP008751.1, AP008745.1, AP008744.1, AP008741.1, AP008740.1, AP008737.1, AP008736.1, AP008734.1, AP008733.1, AP008731.1, AP008920.1, AP008917.1, AP008916.1, AP008915.1, AP008914.1, AP008912.1, AP008911.1, AP008909.1, AP008906.1, AP008903.1, AP008902.1, AP008901.1, AP008900.1, AP008898.1, AP008897.1, AP008896.1, AP008895.1, AP008894.1, AP008890.1, AP008889.1, AP008888.1, AP008887.1, AP008886.1, AP008879.1, AP008878.1, AP008875.1, AP008874.1, AP008873.1, AP008870.1, AP008866.1, AP008865.1, AP008863.1, AP008862.1, AP008860.1, AP008857.1, AP008856.1, AP008853.1, AP008852.1, AP008850.1, AP008849.1, AP008847.1, AP008845.1, AP008842.1, AP008841.1, AP008840.1, AP008839.1, AP008838.1, AP008837.1, AP008835.1, AP008833.1, AP008829.1, AP008828.1, AP008827.1, AP008825.1, AP008536.1, AP008535.1, AP008534.1, AP008533.1, AP008531.1, AP008529.1, AP008526.1, AP008525.1, AP008521.1, AP008520.1, AP008519.1, AP008518.1, AP008516.1, AP008515.1, AP008514.1, AP008512.1, AP008509.1, AP008508.1, AP008507.1, AP008505.1, AP008504.1, AP008503.1, AP008501.1, AP008499.1, AP008497.1, AP008496.1, AP008492.1, AP008491.1, AP008489.1, AP008488.1, AP008487.1, AP008486.1, AP008485.1, AP008484.1, AP008483.1, AP008482.1, AP008481.1, AP008478.1, AP008477.1, AP008475.1, AP008474.1, AP008472.1, AP008471.1, AP008470.1, AP008469.1, AP008467.1, AP008466.1, AP008465.1, AP008464.1, AP008463.1, AP008461.1, AP008459.1, AP008457.1, AP008455.1, AP008453.1, AP008452.1, AP008451.1, AP008450.1, AP008448.1, AP008447.1, AP008446.1, AP008445.1, AP008444.1, AP008442.1, AP008441.1, AP008439.1, AP008437.1, AP008436.1, AP008434.1, AP008433.1, AP008432.1, AP008431.1, AP008430.1, AP008428.1, AP008424.1, AP008423.1, AP008422.1, AP008420.1, AP008418.1, AP008417.1, AP008416.1, AP008414.1, AP008413.1, AP008412.1, AP008408.1, AP008404.1, AP008402.1, AP008401.1, AP008400.1, AP008397.1, AP008395.1, AP008394.1, AP008390.1, AP008388.1, AP008387.1, AP008385.1, AP008384.1, AP008382.1, AP008381.1, AP008379.1, AP008378.1, AP008376.1, AP008375.1, AP008374.1, AP008373.1, AP008370.1, AP008369.1, AP008365.1, AP008361.1, AP008360.1, AP008355.1, AP008354.1, AP008353.1, AP008351.1, AP008350.1, AP008347.1, AP008346.1, AP008345.1, AP008727.1, AP008726.1, AP008725.1, AP008723.1, AP008722.1, AP008721.1, AP008720.1, AP008717.1, AP008716.1, AP008715.1, AP008714.1, AP008713.1, AP008712.1, AP008709.1, AP008708.1, AP008706.1, AP008705.1, AP008704.1, AP008701.1, AP008690.1, AP008689.1, AP008688.1, AP008687.1, AP008686.1, AP008685.1, AP008684.1, AP008683.1, AP008682.1, AP008679.1, AP008678.1, AP008677.1, AP008673.1, AP008671.1, AP008669.1, AP008668.1, AP008667.1, AP008666.1, AP008664.1, AP008662.1, AP008661.1, AP008660.1, AP008658.1, AP008657.1, AP008655.1, AP008654.1, AP008653.1, AP008650.1, AP008649.1, AP008648.1, AP008647.1, AP008645.1, AP008644.1, AP008643.1, AP008642.1, AP008639.1, AP008638.1, AP008637.1, AP008635.1, AP008634.1, AP008631.1, AP008630.1, AP008628.1, AP008627.1, AP008624.1, AP008623.1, AP008621.1,

AP008620.1, AP008619.1, AP008618.1, AP008617.1, AP008616.1, AP008614.1, AP008613.1, AP008612.1, AP008611.1, AP008610.1, AP008609.1, AP008608.1, AP008607.1, AP008606.1, AP008605.1, AP008604.1, AP008603.1, AP008600.1, AP008599.1, AP008598.1, AP008596.1, AP008595.1, AP008594.1, AP008591.1, AP008590.1, AP008588.1, AP008585.1, AP008583.1, AP008582.1, AP008580.1, AP008579.1, AP008578.1, AP008577.1, AP008575.1, AP008572.1, AP008571.1, AP008570.1, AP008569.1, AP008568.1, AP008567.1, AP008566.1, AP008565.1, AP008563.1, AP008562.1, AP008561.1, AP008558.1, AP008557.1, AP008556.1, AP008555.1, AP008554.1, AP008553.1, AP008551.1, AP008550.1, AP008548.1, AP008547.1, AP008545.1, AP008543.1, AP008539.1, AP008538.1, AP008537.1, AF346989.1, AF381996.2, AF382000.1, AF381999.1, AF381998.1, AF381997.1, AF381995.1, DQ200805.1, AF346991.1, AY963581.2, AY963580.2, AY963579.3, AY963583.2, AY963578.2, AF381994.2, AF381992.2, AF381991.1, AF381981.1, AY275535.1, AY275531.1, AY963582.2, AY963574.2, AY289075.1, DQ372883.1, DQ372882.1, DQ372881.1, DQ372880.1, DQ372879.1, DQ372878.1, DQ372875.1, DQ372877.1, DQ437577.1, AY882412.1, AF381984.2, AF382008.2, AF381988.1, AF381987.1, AF381986.1, AF381985.1, AF381983.1, AY275534.1, AF381989.1, AY275536.1, AY275527.1, DQ661681.1, AF347001.1, AY195786.2, AY195759.2, AY195749.2, AY195787.1, AY195748.1, DQ372887.1, DQ341078.1, DQ341075.1, DQ341070.1, AY882391.1, AY882390.1, AY882381.1, AY882380.1, AY882379.1, AY738940.1, DQ137410.1, DQ137409.1, DQ137401.1, AY956414.1, DQ137408.1, DQ137399.1, DQ137400.1, DQ137404.1, AY956413.1, DQ137406.1, DQ137405.1, DQ372873.1, DQ372871.1, DQ372870.1, DQ137398.1, AY956412.1, AY289083.1, AY289082.1, AY289081.1, AY289080.1, AY289079.1, AY289078.1, AY289077.1, AY289092.1, AY289091.1, AY289090.1, AY289089.1, AY289088.1, AY289087.1, AY289086.1, AY289085.1, AF347003.1, AF347002.1, AF347005.1, AF347004.1, AF382012.1, AY882398.1, AY882396.1, AY882388.1, AY882384.1, AF346970.1, AF346971.1, AF346979.1, AY738992.1, AY738971.1, AF346982.1, DQ306709.1, DQ902709.1, DQ902708.1, AY519496.2, AY519495.2, AY519494.2, AY519493.2, AY519492.2, AY519491.2, AY519490.2, AY519489.2, AY519484.2, AY570525.2, AY615360.2, AY615359.1, AY570524.1, AY519488.1, AY519486.1, DQ372885.1, AY289094.1, AF347007.1, DQ523679.1, DQ523678.1, DQ523676.1, DQ523675.1, DQ523671.1, DQ523670.1, DQ523668.1, DQ523666.1, DQ523665.1, DQ523663.1, DQ523662.1, DQ523658.1, DQ523656.1, DQ523654.1, DQ523653.1, DQ523651.1, DQ523645.1, DQ523644.1, DQ523643.1, DQ523642.1, DQ523640.1, DQ523639.1, DQ523638.1, DQ523637.1, DQ523629.1, DQ523625.1, DQ523624.1, DQ523623.1, DQ523621.1, DQ523620.1, DQ904241.1, DQ904240.1, DQ904239.1, DQ904238.1, DQ904237.1, DQ904236.1, DQ904235.1, DQ904234.1, AF346984.1, AF347013.1, AY289071.1, AY289073.1, AY289072.1, AY289074.1, AY882417.1, AY882401.1, AY882392.1, AY882386.1, AY882382.1, AY738991.1, AY738989.1, AY738984.1, AY738963.1, AY738948.1, AF382011.1, DQ200804.1, DQ200803.1, DQ200802.1, DQ200801.1, AY275537.1, AY275533.1, AY275528.1, AF382010.1, AF382009.2, AF381982.1, AY275530.1, AF382007.1, AF382005.1, AF382004.1, AF382003.1, AF382002.1, AF382001.1, AY275532.1, X93334.1, DQ902694.1, DQ902707.1, DQ902706.1, DQ372869.1, DQ372868.1, AY289098.1, AY289097.1, AY289096.1, AY289095.1, AJ842749.1, AJ842748.1, AJ842745.1, AJ842744.1, AJ842746.1, AJ842750.1, AJ842751.1, AJ842747.1, AY289101.1, AY289100.1, AY289099.1, AY289102.1, AY963585.2, DQ856317.1, DQ856316.1, DQ830736.2, DQ834261.1, DQ834260.1, DQ834259.1, DQ834258.1, DQ834257.1, DQ834256.1, DQ834255.1, DQ834254.1, DQ834253.1, DQ785296.1, DQ787109.1, DQ781338.1, DQ358977.2, DQ358976.2, DQ358974.2, DQ358973.2, DQ156213.1, DQ156212.1, DQ156211.1, DQ156210.1, DQ156209.1, DQ156208.1, V00662.1, DQ826448.1, AY245555.1, AY882385.1, AY882383.1