RPPR COVER PAGE

Federal Agency and Organization Element to Which Report is Submitted: National Institute of Justice

Federal Grant or Other Identifying Number Assigned by Agency: 2015-DN-BX-K053

Project Title:
Validation of a Probe Capture Next-Generation Sequencing Assay for Whole Mitochondrial Genome Analysis

PD/PI Name, Title and Contact Information (e-mail address and phone number):
Cassandra D. Calloway, PhD
Assistant Scientist

Name of Submitting Official, Title, and Contact Information (e-mail address and phone number), if other than PD/PI:
Robert S. Fries
Vice President, Children's Services Finance, Chief Financial Officer
747 52nd Street
Oakland, CA 94609-1809

Submission Date: October 10th, 2018

DUNS and EIN Numbers:

Recipient Organization (Name and Address):
Children's Hospital & Research Center at Oakland
747 52nd Street
Oakland CA 94609-1809

Recipient Identifying Number or Account Number, if any:
Project/Grant Period (Start Date, End Date): 01/01/2016 – 09/30/2018

Report Term or Frequency (annual, semi-annual, quarterly, other): Draft Final Summary Overview

Signature of Submitting Official

DocuSigned by:
Robert S. Fries
X [redacted]
Robert S. Fries
Vice President, Children's Services Finance, Chief Financial Officer

10/10/2018
Date

Validation of a Probe Capture Next-Generation Sequencing Assay for Whole Mitochondrial Genome Analysis

Draft Final Summary Overview

Submitted to the National Institute of Justice

Principal Investigator: Cassandra Calloway, PhD
Authors: Shelly Shih, M.S, Rachel Gordon, M.S., Jessica Lim, Henry Erlich, PhD, and Cassandra Calloway, PhD
Grant Number: 2015-DN-BX-K053

Children's Hospital Oakland Research Institute
5700 Martin Luther King Jr. Way
Oakland, CA 94609

A. Project Purpose and Goals

The primary goal of this NIJ supported research project was to validate a previously developed and optimized probe capture next-generation sequencing (NGS) assay for sequencing of the entire mitochondrial (mt) genome for forensically challenging samples. The mtgenome probe capture NGS assay can be used on reference and compromised samples, including highly degraded, mixed and limited DNA samples often encountered in forensic cases or mass disasters.

The specific aims of this project were to 1) complete the developmental validation of the mtgenome probe capture NGS assay, 2) test the performance on highly degraded, limited, and mixed forensic type samples, 3) complete internal validation studies (external collaborators), 4) conduct comparative NGS technology studies (Illumina MiSeq and ThermoFisher Ion Torrent), and 5) customize and validate mtDNA analysis tools in collaboration with SoftGenetics (GeneMarker®HTS software).

This probe capture NGS assay has the potential to greatly improve the current method of mtDNA analysis of challenging forensic-type samples by broadening the types of samples that can be successfully analyzed for the whole mtgenome. We demonstrated that our assay can be applied to complex mixtures as well as challenging samples such as hairs, touch DNA, and bones. Challenging complex mixtures with trace minor contributor DNA or those with more than two contributors were successfully analyzed using the phylogenetic based software, Mixemt.

1) Developmental Validation of the mtGenome Probe Capture NGS assay

A. Optimization

Optimized conditions for DNA library preparation using the Kapa Hyper Prep Kit, probe capture using the NimbleGen SeqCap EZ Custom Probe Capture Assay, and NGS on an Illumina MiSeq were previously determined for DNA input amounts ranging from 0.1- 1 ng[1]. Since lower DNA amounts may be encountered for some case samples such as shed hairs, a series of experiments was conducted to determine the optimal PCR cycle number testing 17-28 cycles for DNA sample input amounts ranging from 1ng–1pg and undetectable nDNA amounts from shed hairs. Based on these studies, for samples with very low or undetectable nuclear DNA, we recommend 24–28 PCR cycles to generate sufficient DNA library amounts for downstream analysis.

C. Characterization of Genetic Markers

Genetic variation in the hypervariable regions I and II (HVI/II), non-coding D-loop, and coding region of the human mtgenome has been extensively characterized. In this study, we sequenced the mtgenome of 56 different blood derived DNA samples from four population groups- 14 African American, 12 Caucasian, 11 Hispanic, and 19 Japanese – in replicates for a total of 162 DNA libraries sequenced in 11 NGS runs. All major variant reports were concordant across replicates for all samples at a 10% mutation threshold. Relatively uniform coverage was observed across the mtgenome and mtDNA variants did not appreciably affect hybridization/capture, as measured by sequence read recovery. While the density of mtDNA variants was highest in the HVI/II regions, ~3x more variants were observed outside these regions across all four population groups tested, ranging from ~6 to 16 HVI/II variants and ~16 to 50 variants in the variable and coding regions in Caucasian and African American populations, respectively. These results further demonstrate that sequencing the whole mtgenome identifies valuable polymorphic sites in regions outside of HVI/II, thus increasing the discriminatory power of mtDNA analysis and further differentiating the common HVI/II haplotypes.

D. Species Specificity

Species specificity was evaluated with nine non-human species in replicates- three primates (gorilla, chimpanzee, and orangutan) and six animals (cat, dog, pig, cow, rat, and chicken). Low level contamination of human DNA was detected and confirmed in all animal species tested, likely at the DNA extract level as these were >20 year old samples. To bioinformatically remove the human DNA contaminant, the NGS data from each animal was aligned to their respective mtgenome (ranging from 55% - 100% coverage) and then the aligned animal DNA reads were aligned to the human mtgenome. The percent coverage (>100×) of the animal DNA to the human mtgenome ranged from 0.45% (pig) to 4.61% (cat) in the six animals and 22.71% (Orangutan) to 50.43% (Chimpanzee) in higher order primates. Although regions of the animal species tested were homologous to the human mtgenome, the percent coverage was low for non-primate animals, which are most likely to be encountered at a crime scene. In addition, the number of mtDNA variants observed for the animal to human alignment was higher and distinct from true human mtDNA variants.

E. Sensitivity

Sensitivity of the system was evaluated using 35 HL60 control DNA samples varied in DNA input amount, ranging from 200 – 200,000 mtDNA copies (1pg – 1ng genomic DNA) in replicates. Full coverage (100%) of the mtgenome (>100× with and >20× without PCR duplicates) and 100% concordance (15% variant threshold) with the known sequence was observed for ≥10,000 mtDNA copies (50pg). Full coverage (100%) of the mtgenome was observed for 2,000 mtDNA copies (10pg) with PCR duplicates; however, the mtgenome coverage was reduced to 99-100% (>5×) and ~40-75% (>20×) when PCR duplicates were removed. As a result, some regions represent fewer than 5-10 mtDNA copies, and thus the base calling accuracy was decreased when a 10-15% variant threshold was applied due to the low number of mtDNA copies as one artifact would represent 10-20% reads without PCR duplicates. Base calling accuracy was further reduced for samples < 1,000 mtDNA copies (5pg) due to low template effects as evidenced by low percent mtgenome coverage (4 – 75%) with PCR duplicates removed (>20×). Overall, full coverage of the mtgenome can be expected for samples with as few as ~2,000 mtDNA copies (10pg of DNA), however regions covered by fewer than 10-20 unique sequence reads may have lower base calling accuracy. Full mtgenome coverage and 100% base calling accuracy of the major variants can be expected for samples with >10,000 mtDNA copies (≥ 50pg). Base calling accuracy can be improved by removing PCR duplicates and by setting a >10-20× minimum coverage, removing artifacts; and applying a phylogenetic based approach for analysis (e.g. Mixemt).

F. Stability

A mock degradation study was conducted by fragmenting K562 control DNA at 100pg and 1ng to 150bp to mimic degraded samples and processed with our standard conditions including shearing to 250bp. Results were similar to un-fragmented controls with 100% coverage of the mtgenome and uniform coverage for both 1ng and 100pg samples, demonstrating that the assay can capture and sequence degraded samples independent of the fragment size[1].

G. Reproducibility and Repeatability

Reproducibility was evaluated by comparing the concordance of the major variant reports of the same population sequence data generated by different analysts, instruments, and laboratories. In this study, concordance of major variant reports of 23 DNA samples (21 population and 2 positive controls) were

carried out following the same library-preparation-probe-capture-enrichment protocol by two different analysts between the Richmond DNA Laboratory of the Department of Justice (DOJ) and CHORI. On-target rates were 78% for the CHORI run and the 88.5% for DOJ run. Analysis of concordance of the major variant reports across these two runs is ongoing. Additionally, concordance of the major variant reports of the same DNA libraries was evaluated with data generated by different analysts, instruments, and laboratories. In this study, enrichment and sequencing of 14 DNA libraries across three NGS runs between Netherlands Forensic Institute (NFI) and CHORI were carried out following the same probe capture enrichment and sequencing protocol. On-target rates were also comparable for the three NGS runs between NFI and CHORI (73% - 82%). All major mutation reports were concordant with a 10% mutation cutoff across three NGS runs. Concordance of the mutation reports and on-target rates both demonstrate that sequence data generated by this system is highly reproducible starting from both sample extracts and DNA libraries independent of the analyst, instrument and laboratory.

Repeatability of the mtgenome probe capture assay was evaluated by comparing the concordance of major variant reports generated by the same analyst. In this study, library construction, probe capture enrichment, and sequencing of 156 population (56 unique) samples were carried out by the same analyst on the same NGS instrument for a total of eight sequencing runs. All samples exhibited full coverage of the mtgenome and all major variant reports are concordant between replicates with 10% mutation threshold, demonstrating that our system generates highly repeatable data for the same samples.

H. Mixtures

Performance of the mtgenome probe capture NGS assay was assessed for analysis of mixtures by testing four contrived mixtures in quadruplicate with the minor contributor DNA ranging from 15% - 45%. The results demonstrated reliable and uniform capture and sequence read recovery of the minor contributor haplotype in samples with minor contributor proportion as low as 15% with no minor variant dropouts. The narrow range of the standard deviation values also demonstrate the precision of our system for quantification of minor contributor DNA in mixtures.

I. Precision and Accuracy

Base calling precision and accuracy were calculated by determining the number of correctly called mtgenome bases (16,569 bases) across 11 HL60 1ng control sample replicates compared to the published HL60 mtgenome sequence (accuracy) or the majority base called across the 11 replicates (precision). At a 15% mutation threshold setting, all 11 HL60 replicates were 100% concordant with the published HL60 sequence[2] with 100% base calling accuracy and precision with detection of the reported heteroplasmy at position 12071 and a single base called at all other positions. At a 10% mutation cutoff, the major mutation reports were 100% concordant for all 11 samples and base calling accuracy and precision was 100% for both major and minor variants (≥10%) reported when PCR duplicates were removed; however, the base calling accuracy and precision were slightly reduced (99.998%) due to the detection of three minor variants in one replicate which were slightly above the 10% threshold analyzed with PCR duplicates. As the mutation filter was lowered to 5% and 3%, major mutation reports were still concordant, but base calling accuracy of the minor variants was slightly reduced to 99.991% and 99.944% as more pseudogenes and artifacts were detected. Overall, both 100% accuracy and precision can be achieved by setting a ≥10% mutation threshold or by removing the BLAST-confirmed pseudogenes and artifacts when lower mutation threshold settings are used with GeneMarker®HTS. Alternatively, a phylogenetic based software (e.g. Mixemt) can be used when mixtures are suspected with trace or low minor frequency to improve base calling accuracy of minor variants.

2) Analysis of highly degraded, limited, and mixed forensic type samples

The mtgenome probe capture NGS system was applied to a range of challenging forensic-type samples including the following: 54 shed hairs, 22 cut hairs, 14 ancient bone specimens, 16 touch DNA from spent brass cartridges, 24 contrived 2-person and 3-person mixtures, six mock case hair mixtures, and nine heteroplasmic samples.

A. Shed Hair

A total of 54 2cm telogen hair roots with mtDNA copies ranging from 1,300 - >2,000,000 mtDNA copies were tested, and 51 exhibited 100% coverage of the mtgenome (>100×) with the PCR duplicates and ~99% coverage of the mtgenome (>5×) without the PCR duplicates. Partial mtgenome coverage was observed for

three telogen hair root portions with ~90% coverage with the PCR duplicates (>100×) and ~60% coverage after removing the PCR duplicates (>5×).

B. Hair Shafts

Results from HVI/II PCR amplification were directly compared to the mtgenome probe capture NGS system for 22 cut telogen hair shafts (4 cm divided for the two analyses) ranging from ~128 – 43,000 mtDNA copies. Using the HVI/II PCR, nine cut hairs yielded sufficient HVI/II PCR product for Sanger sequencing, four were too low, and nine failed to amplify. Of the 13 hair shafts that showed weak or negative HVI/II amplification results, full or partial mtgenome coverage (>96%) was observed for all 13 cut hairs with eight hairs showing 100% coverage of the mtgenome (with the PCR duplicates) using the probe capture NGS system. After removing the PCR duplicates, five hairs still showed 100% coverage of the mtgenome and eight hairs showed >90% coverage of the mtgenome. Full coverage of the HVI/II regions was observed for all cut hair shafts except for two hair shafts with partial coverage in the HVI region (16024 – 16365). These results demonstrate that our system can recover a significant portion of the mtgenome in the hair shafts that initially failed HVI/II amplification.

C. Bones

Solid tissues dating to ~100 years (N = 6, femurs and humerus), ~2,000 years (N = 6, teeth), and ~4,000 years (N = 2, teeth) were tested using our system. The ~100-year-old bone samples had previously failed conventional STR and HVI/II mtDNA analysis. Although the mtDNA was highly degraded as confirmed using a mtDNA qPCR degradation assay, we were able to capture and sequence DNA fragments as short as 30bp and recover 52.75% – 100% of the mtgenome with an on-target rate of 0.268% – 12.57%. Additionally, six ~2,000 –year-old teeth showed ~26 – 90% coverage of the mtgenome with an on-target rate of 0.8 - ~52%. The system was also successfully applied to the ~4,000-year-old teeth from two individuals with initial coverage of the mtgenome ranging from 93.3% – 99.8% (>5×) and the on-target rate ranging from ~10% – ~94%. The overall coverage of the mtgenome was increased to 94.7% - 100% when multiple sequence runs of the same DNA libraries were combined (>5×), showing that by combining the different sequence data for the same samples, we not only increased the mtgenome coverage, but also increased the read depth and percent aligned reads (%) for the less successfully sequenced samples.

D. Touch DNA

A total of 16 touch DNA samples collected with flocked swabs from spent brass cartridges were captured and sequenced. When analyzed with the PCR duplicates included, 11 of the 16 DNA samples collected with flocked swabs exhibited full coverage of the mtgenome (>100×) while five samples exhibited partial coverage of the mtgenome ranging from 47 - >99%. After removing the PCR duplicates, six samples still exhibited full coverage of the mtgenome, six samples had >99% coverage, and four samples showed 18% - 96% coverage of the mtgenome (5×). All 16 samples were mixtures, as expected. Using the phylogenetic-based software Mixemt, we were able to detect the contributor's haplotype in 11/16 samples with the lowest minor contributor proportion detected at 7.04%.

E. Contrived Mixtures

Six two-person mixtures and 18 three-person mixtures at varying minor contributor ratios were created to evaluate the assay and to compare analysis with frequency-based (GeneMarker®HTS; HTS) and phylogenetic-based (Mixemt) software. Using the frequency-based HTS, we were able to de-convolute two-person mixtures with minor contributor proportions between 15% - 40%. As expected, using HTS alone, de-convolution of 50:50 mixtures was difficult since variants were incorrectly assigned to minor or the major report, leading to incorrect haplogroup determination. De-convolution of trace minor amounts (95:5) using HTS alone is difficult as noise such as pseudogenes can occur at ≤ 10%, thus confounding the results. Using the phylogenetic approach, we were able to de-convolute 50:50 and 95:5 mixtures as well as three-person mixtures, which could not be interpreted using HTS analysis alone. While Mixemt is capable of separating complex mixtures, the software does not consider Indels and is not able to assign some un-linked private mutations. Using both frequency based HTS and phylogenetic-based Mixemt software, we can more accurately de-convolute complex mixtures.

F. Mock-case Mixtures

Six mock-case mixtures were created using a single hair mixed with trace amounts of blood, semen, or saliva and extracted without undergoing the typical hair washing procedure prior to extraction to evaluate our assay and software for forensic applications. One of the hair/semen mixtures was determined to be single source hair using both softwares. Overall, the frequency-based HTS was able to de-convolute 3/5

mixtures while Mixemt was able to separate all five mixtures with a minor contributor frequency as low as 0.3%. Using Mixemt, we were able to detect the minor contributor of the hair/trace blood mixture at 0.3% frequency (**Figure 1**), which would be filtered out as background noise using HTS analysis.

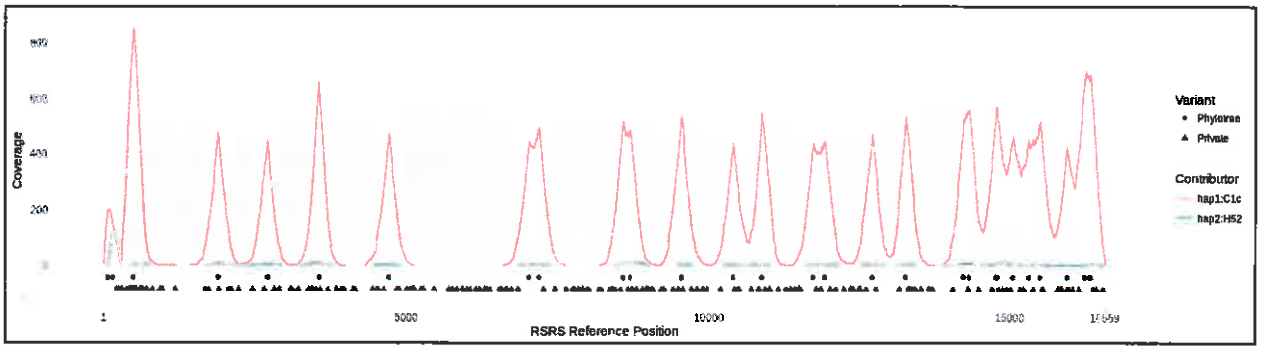

Figure 1. Mtgenome Coverage curve generated by Mixemt of hair + trace blood mixture.

G. Heteroplasmy Samples

Nine different population samples with heteroplasmic sites were captured and sequenced at least two times. All heteroplasmic sites were detected with a minor variant frequency from ~8% - 43%. The minor variant frequency at the heteroplasmic positions deviated $< \pm 2\%$ with an exception of one sample with a deviation of $\pm 3.6\%$. The results show that our system is precise in estimating the minor variant frequency and can be useful for detection of minor contributor components in a mixture.

3) Internal validation studies with external collaborator (CA DOJ)

Internal validation studies were conducted to validate the mtgenome probe capture NGS assay between CA DOJ and CHORI labs with HL60 sensitivity samples, seven mixtures, 22 case-type and non-probative samples. A hair shaft that failed duplex HVI/II amplification ((444 bp and 415 bp), part of current Sanger sequencing protocol) was successfully sequenced with the mtgenome Probe Capture NGS assay. **Sensitivity** was evaluated with HL60 DNA at 1pg to 5ng with 17 – 28 cycles. Reliable results were observed for samples with >2,000 mtDNA copies (10pg). When input amounts of DNA were lowered to ≤ 10pg, artifacts increased due to low level template effects. Seven **mixtures** (five two-person and two three-person) were made for mixture evaluation. Two-person mixtures at 5%, 10%, and 33% minor contributor frequencies were successfully de-convoluted using HTS alone. The three-person mixtures could not be separated using HTS software, the only available software at the time. Twenty-two **non-probative** and **mock-case** samples were tested, including three artificially degraded, four hair shafts, 11 non-probative,

and four artificially inhibited samples. Two of the three artificially degraded samples exhibited full coverage of the mtgenome with <50% on-target rates. Four hair shafts exhibited partial coverage (2cm hairs) and full coverage (4cm) of the mtgenome. Six of the eight bone samples yielded full mtDNA types. Humic acid (90 and 180 ng/µl) and hematin (150 and 300 µM) were added to control DNA to create artificially inhibited samples. Samples with the lower inhibitor concentration exhibited partial mtgenome coverage (73 and 85%, respectively) and the samples with higher inhibitor concentration showed very little coverage of the mtgenome. Contamination of the assay was assessed with 12 positive and 12 negative controls. Minor contamination was detected in all negative controls with the average coverage of 9.5 ± 6.4. Low-level contamination is expected in mtDNA assays and can be used to assess background noise level and thresholds necessary to produce interpretable data.

4) Comparative NGS Technology Studies: (MiSeq vs Ion Torrent S5 XL)

Comparative NGS studies were conducted to demonstrate that our mtgenome probe capture NGS system can be applied to both MiSeq (Illumina) and Ion Torrent (ThermoFisher) by using the respective Kappa library preparation methods. A subset of previously sequenced (on MiSeq) population samples (N=8), heteroplasmic samples (N =2), HL60 limiting dilution samples at 1 ng and 100 pg (N =5), NA24149 control DNA at 1ng (N =3), and two-person mixtures (N =2) were captured and sequenced on the Ion Torrent S5 XL to assess initial capture and sequencing performance. Library preparation for limiting dilutions of HL60 were performed for 1 pg to 1 ng in duplicates, with the exception of 1 ng libraries that were made in triplicate, to assess sensitivity. Lower sensitivity was observed due to limitations with Ion Torrent shotgun library preparation methods which did not consistently produce libraries that were quantitatively or qualitatively suitable for capture with DNA amounts less than 100 pg, resulting in only 1 ng and 100 pg libraries suitable for enrichment and sequencing. At 1 ng and 100 pg, 100% and >99% coverage of the whole mtgenome, respectively was obtained with PCR duplicates (> 100×). Mtgenome coverage was less uniform in Ion Torrent data with areas of low coverage (<500×) observed in homopolymer regions. Uniformity of coverage was improved when PCR duplicates were removed; this improvement could be attributed to the use of emulsion PCR for template preparation prior to sequencing in Ion Torrent technology

or to the PCR duplicate removal algorithm. Variants, including confirmed heteroplasmic sites, and haplogroup calling for HL60 and population samples was concordant between both sequencing platforms using a 10% variant threshold. Analysis to determine concordance between the sequencing platforms is ongoing for population samples and contrived mixtures. Further optimization studies are needed to improve low-template DNA library construction for sequencing on Ion Torrent platforms.

5) Validation and refinement of reporting features and tools for mtDNA analysis using customized GeneMarker®HTS software (in collaboration with SoftGenetics)

GeneMarker®HTS (HTS) software is a software developed for analysis of high-throughput sequencing data. New features and tools were added to HTS as suggested by CHORI. HTS now allows the analyst to copy individual sequence reads to search against BLAST, which is useful for determining if artifacts arise from pseudogenes or other sources of contaminants. Individual reads can be manually removed in order to reflect the true variant frequencies. The major variant strings can be directly searched against EMPOP. Another useful feature is the in-project PCR duplicate removal option which allows us to determine the starting number of DNA molecules in order to establish the thresholds for stochastic variation. We are currently working with SoftGenetics team to implement a pseudogene filter.

B. Impact to criminal justice system, forensic science and other fields of science

The ability to analyze whole mtgenome sequences in degraded forensic samples as well as mixture samples will have a major impact on the criminal justice system. The Power of Discrimination increases significantly when analyzing the whole mtgenome sequence rather than just HVI/II regions as seen in the population samples with higher number of variants observed outside of the HVI/II regions. The customized system of library preparation, probe capture, and NGS on Ilumina MiSeq is a robust system as demonstrated in the data generated by CHORI, CA DOJ, and NFI. Using both frequency-based (GeneMarker®HTS) and phylogenetic-based (Mixemt) software, we can expand the types of complex mixtures that can be analyzed. The successful completion of the developmental validation studies for the mtgenome probe capture NGS system described here, will provide an alternative method for analysis of mtDNA for applications to highly degraded, limited, and mixed DNA samples often encountered in missing person identification, mass disaster cases, and forensic casework.

References

1. Shih, S.Y., Bose, N., Goncalves, A.B.R., Erlich, H.A. & Calloway, C.D. Applications of Probe Capture Enrichment Next Generation Sequencing for Whole Mitochondrial Genome and 426 Nuclear SNPs for Forensically Challenging Samples. *Genes (Basel)* **9**(2018).
2. Riman, S., Kiesler, K.M., Borsuk, L.A. & Vallone, P.M. Characterization of NIST human mitochondrial DNA SRM-2392 and SRM-2392-I standard reference materials by next generation sequencing. *Forensic Sci Int Genet* **29**, 181-192 (2017).

Appendix: List of Deliverables and availability information:

I. Protocols:

Name: Library Preparation and Capture Protocol for Limited and High Quantity DNA Samples. (v. 2.1 07/06/2018)
Description: Combined, optimized single protocol based on the KAPA Hyper Prep Protocol (KR096-v1.14) and the NimbleGen SeqCap EZ Library User SR (v5.1) for DNA library construction and probe capture enrichment for pristine and casework samples (degraded/limited/mixed) with starting amounts ranging from 200 – 200,000 mtDNA copies (1pg – 1ng genomic DNA).
Accessibility: Protocol can be obtained from PI, Dr. Calloway (scalloway@chori.org) upon request

II. Probe Capture Panel:

Name: Custom whole mitochondrial genome SeqCap EZ Choice DNA probe solution (Calloway Lab Design)
Description: DNA probes with high on target capture redundancy for efficient capture of up to 24 samples at equal amounts.
Accessibility: Available from Roche NimbleGen with prior permission from Dr. Cassandra Calloway (scalloway@chori.org).

III. Software Programs:

Name: Mixemt
Description: Mixemt uses estimation maximization algorithm and phylogenetic information on each sequence reads to estimate the probability of each reads originating from all possible haplogroup candidates
Accessibility: Open source on Github https://github.com/svohr/mixemt

IV. Webinar:

1. Cassandra Calloway. (2017, March). Development of a Probe Capture NGS System for Forensics. FTCOE Webinar. https://rticqpub1.connectsolutions.com/content/connect/c1/7/en/events/event/shared/1178106013/event_landing.html?sco-id=1208236424&_charset_=utf-8

   Description: A targeted probe capture based method of enrichment for clonal, massively parallel sequencing of the mitochondrial genome and nuclear SNPs can address the challenge of analyzing complex and low level mixtures, highly degraded and limited DNA samples.
   Accessibility: Available as archived webinar through FTCOE and RTI International.

V. Publications:

A. Peer Reviewed Publications

1. Shih, Shelly, Nikhil Bose, Anna Beatriz R. Gonçalves, Henry A. Erlich, and Cassandra D. Calloway. Applications of Probe Capture Enrichment Next Generation Sequencing for Whole Mitochondrial Genome and 426 Nuclear SNPs for Forensically Challenging Samples. (2018) Genes, 9(1):49. PMID: 29361782.

2. Shih, Y. Shelly, Rachel Gordon, Henry A. Erlich, and Cassandra D. Calloway. Developmental Validation of a Probe Capture Next Generation Sequencing System for Whole Mitochondrial Genome of Forensically Challenging Samples. (manuscript in preparation).

B. Theses:

1. Cassandra Taylor. (December, 2016). *Analysis of Highly Degraded DNA from Bone Samples using Probe Capture and Next Generation Sequencing*. (UC Davis Master Thesis). Retrieved from ProQuest Dissertations and Theses.

2. Guillermina Almada. (December, 2016). *Improving Telogen Hair Analysis by Predicting Nuclear and Mitochondrial DNA Success for Massively Parallel Sequencing Using Microscopic and qPCR Methods*. (UC Davis Master Thesis). Retrieved from ProQuest Dissertations and Theses.

3. Shelly Shih. (March, 2017). *Characterization of Germline Heteroplasmy in Mother-Offspring Pairs using Next Generation Sequencing*. (UC Davis Master Thesis). Retrieved from ProQuest Dissertations and Theses.

4. Symone Watson. (March, 2018). *A Comparison of Touch DNA Collection Methods from Spent Cartridge Casings with Next-Generation Sequencing*. (UC Davis Master Thesis). Retrieved from ProQuest Dissertations and Theses.

5. Mary Wisner. (In preparation, December, 2018). *Resolution of mtDNA Mixtures using a Probe Capture Next Generation Sequencing System and Custom Analysis Software*. (UC Davis Master Thesis). Retrieved from ProQuest Dissertations and Theses.

C. Invited Talks

1. Cassandra Calloway. (2017, March). Targeted Massively Parallel Sequencing of mtDNA and Nuclear SNP Markers Using Probe Capture for Analysis of Mixtures and Degraded DNA. Oral Presentation at the 2nd Annual Genetics in Forensics Congress. London, UK.

2. Cassandra Calloway. (2017, March). Development of a Probe Capture NGS System for Forensics. Oral Presentation at Netherlands Forensic Institute, Hague Netherlands.

3. Cassandra Calloway. (2017, June). Development of a Probe Capture NGS System for Forensics. Oral Presentation at the ISABS Conference on Forensic and Anthropologic Genetics and Mayo Clinic Lectures in Individualized Medicine. Dubrovnik, Croatia.

4. Cassandra Calloway. (2017, August 18th). Analysis of Nuclear SNP Markers and Mitochondrial Genome Using a Novel Probe Capture Massively Parallel Sequencing Assay for Forensic Applications. Cambridge HealthTech Institute's Second Annual Next Generation Summit, NGS for DNA Forensics, Washington, DC

5. Cassandra Calloway. Probe Capture NGS System for Forensics and Clinical Applications. Thermofisher, Oyster pt. CA, February, 2018.

6. Cassandra Calloway. Probe Capture NGS System for Forensics and Clinical Applications. Thermofisher, Oyster pt. CA, April, 2018.

D. Oral Presentations:

1. Cassandra Taylor, George Sensabaugh, Henry Erlich, and Cassandra Calloway. Analysis of Highly Degraded DNA from Bone Samples Using Probe Capture Enrichment of the Entire Mitochondrial Genome and Next Generation Sequencing. Oral Presentation at the AAFS Scientific Meeting. New Orleans, LA. February 17th, 2017.

2. Henry Erlich, Rachel Gordon, Sam Vohr, Richard Green, Cassandra Calloway. De-Convolution of Forensic Mixtures Using NGS Analysis of Mitochondrial Genome Sequences. Oral Workshop Presentation at the Sequencing, Finishing, and Analysis in the Future (SFAF) Meeting. Santa Fe, New Mexico. (May, 2017).

3. Cassandra Calloway et al. "Probe Capture NGS System for Forensics and Clinical Applications" Oral Presentation at Thermofisher. South San Francisco, CA, February 12th, 2018.

4. Shelly Y. Shih, Symone M. Watson, Guillermina M. Almada, Anna B. Gonçalves, Nikhil Bose, Henry H. Erlich, and Cassandra D. Calloway. "Analysis of Challenging Forensic Samples using Probe Capture Next Generation Sequencing." Oral Presentation at the AAFS Scientific Meeting. Seattle, WA, February 22nd, 2018.

5. Symone M. Watson, Shelly Y. Shih, Henry H. Erlich, Cassandra D. Calloway. "A Comparison of Touch DNA Collection Methods from Spent Cartridge Casings with Next Generation Sequencing." Oral Presentation at the CAC DNA Workshop. Concord, CA, May 8, 2018.

E. Poster Presentations

1. Nikhil Bose, Rachel Gordon, Cassandra Taylor, Shelly Shih, Samuel Vohr, Guillermina Almada, Anna Beatriz Gonscalves, George Sensabaugh, Richard Green, Henry Erlich, and Cassandra Calloway. Mitochondrial Genome and Nuclear SNP Probe Capture Next-Generation Sequencing System for Analyzing Degraded and Mixed DNA Samples. Poster Presentation at ISFG. Seoul, South Korea. (Aug. 28th – Sep. 2nd, 2017)

2. Cassandra Calloway, Nikhil Bose, Rachel Gordon, Cassandra Taylor, Shelly Shih, Guillermina Almada, George Sensabaugh, and Henry Erlich. Mitochondrial Genome and Nuclear SNP Probe Capture Next-Generation Sequencing System for Analyzing Degraded and Mixed DNA Samples. Poster Presentation at ISHI. Seattle, WA, United States. (Oct. 2nd – Oct. 5th, 2017)

3. Symone Watson, Shelly Shih, and Cassandra Calloway. Whole Mitochondrial Genome Coverage of Touch DNA Recovered from Spent Cartridge Casings using Double Swab Method with Cotton and Flocked Swabs. Poster Presentation at ISHI. Seattle, WA, United States. (Oct. 4th, 2017)

4. Nikhil Bose, Rachel Gordon, Shelly Shih, Cassandra Taylor, George Sensabaugh, Henry Erlich, and Cassandra Calloway. "Mitochondrial Genome and Nulcear SNP Probe Capture/Next-Generation Sequencing System for Analyzing Degraded and Mixed DNA Samples. Poster Presentation. PITTCON Conference & Expo 2018. Orange County Convention Center, Orlando, Florida, USA. (Mar. 1st, 2018)

5. Jessica Lim, Henry H. Erlich, and Cassandra D. Calloway. "Improving Mixture Analysis utilizing Single Nucleotide Polymorphism (SNP) Probe Capture Enrichment and Massively Parallel Sequencing." Poster Presentation at the California Association of Criminalists Spring 2018 Seminar. Concord, CA, May 9th, 2018.

6. Mary Wisner, Shelly Shih, MS, Henry Erlich, PhD, and Cassandra Calloway, PhD. "Resolution of mtDNA Mixtures using a Probe Capture Next Generation Sequencing System and Custom Analysis Software." Poster presentation at the spring CAC conference. Concord, CA, May 9, 2018.

F. Lectures

1. Cassandra D. Calloway. Mitochondrial DNA Analysis for Forensic Applications. Guest Lecturer FOR280 Forensic DNA Analysis. UC Davis, Forensic Science Graduate Program. Fall quarter (12 hours total; 3hours/class). (March 2016, March 2017, November 2017, and December 2018)

2. Cassandra Calloway. Probe Capture NGS System for Forensics and Clinical Applications. CHORI faculty seminar. Oakland, CA, May 23rd, 2018.